Discrete-to-Continuum Limits of Transport Problems and Gradient Flows in the Space of Measures

by

Lorenzo Portinale

June, 2021

A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Committee in charge:

Georgios Katsaros, Chair
Jan Maas
Julian Fischer
Nicola Gigli

IST AUSTRIA
Institute of Science and Technology

The thesis of Lorenzo Portinale, titled *Discrete-to-Continuum Limits of Transport Problems and Gradient Flows in the Space of Measures*, is approved by:

Supervisor: Jan Maas, IST Austria, Klosterneuburg, Austria

Signature: ______________________________

Committee Member: Julian Fischer, IST Austria, Klosterneuburg, Austria

Signature: ______________________________

Committee Member: Nicola Gigli, SISSA, Trieste, Italy

Signature: ______________________________

Defense Chair: Georgios Katsaros, IST Austria, Klosterneuburg, Austria

Signature: ______________________________

Signed page is on file

IST Austria Thesis, ISSN: 2663-337X

ISBN:

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: ______________________________

Lorenzo Portinale
June, 2021

Signed page is on file

Abstract

This PhD thesis is primarily focused on the study of discrete transport problems, introduced for the first time in the seminal works of Maas [Maa11] and Mielke [Mie11] on finite state Markov chains and reaction-diffusion equations, respectively. More in detail, my research focuses on the study of transport costs on graphs, in particular the convergence and the stability of such problems in the discrete-to-continuum limit. This thesis also includes some results concerning non-commutative optimal transport.

The first chapter of this thesis consists of a general introduction to the optimal transport problems, both in the discrete, the continuous, and the non-commutative setting.

Chapters 2 and 3 present the content of two works, obtained in collaboration with Peter Gladbach, Eva Kopfer, and Jan Maas, where we have been able to show the convergence of discrete transport costs on periodic graphs to suitable continuous ones, which can be described by means of a homogenisation result. We first focus on the particular case of quadratic costs on the real line and then extending the result to more general costs in arbitrary dimension. Our results are the first complete characterisation of limits of transport costs on periodic graphs in arbitrary dimension which do not rely on any additional symmetry.

In Chapter 4 we turn our attention to one of the intriguing connection between evolution equations and optimal transport, represented by the theory of gradient flows. We show that discrete gradient flow structures associated to a finite volume approximation of a certain class of diffusive equations (Fokker–Planck) is stable in the limit of vanishing meshes, reproving the convergence of the scheme via the method of evolutionary Γ-convergence and exploiting a more variational point of view on the problem. This is based on a collaboration with Dominik Forkert and Jan Maas.

Chapter 5 represents a change of perspective, moving away from the discrete world and reaching the non-commutative one. As in the discrete case, we discuss how classical tools coming from the commutative optimal transport can be translated into the setting of density matrices. In particular, in this final chapter we present a non-commutative version of the *Schrödinger problem* (or *entropic regularised optimal transport* problem) and discuss existence and characterisation of minimisers, a duality result, and present a non-commutative version of the well-known *Sinkhorn algorithm* to compute the above mentioned optimisers. This is based on a joint work with Dario Feliciangeli and Augusto Gerolin.

Finally, Appendix A and B contain some additional material and discussions, with particular attention to Harnack inequalities and the regularity of flows on discrete spaces.

Acknowledgements

A mia zia Enrica,

First and foremost, I am deeply thankful to my supervisor Jan Maas, who has been of constant guide, support, and inspiration throughout my four years at IST.

I would also like to thank the members of my thesis committee, Nicola Gigli and Julian Fischer, for very useful confrontations and suggestions during my PhD.

A big, sincere thank goes to my friend, colleague, and collaborator Dario Feliciangeli. During these four years we shared inspiring mathematical discussions and an invaluable friendship, and for this I'll always be thankful.

I am grateful to all my collaborators, from young students to experienced professors, with whom I found myself sharing mathematical challenges and friendly evenings. I learned a lot from you and our interaction has made me grow a lot as a mathematician.

I want to thank my family, in particular my mum, dad, and sister, and my friends. From Italy to Vienna, you supported me during this journey, and helped me to pursue my goals with happiness and serenity. To the Drama Queens, to the Bicerin, to my Pisa friends, to my family: I have been lucky to have you around these years.

Finally, a special thank goes to Alice. Having your last year of the PhD in the middle of a pandemic can't be easy, but it becomes simpler when you have a person next to you that supports you and believes in you, as she did for me.

The author gratefully acknowledges support by the Austrian Science Fund (FWF), grants No W1245 and No F65.

About the Author

Lorenzo Portinale completed a Bachelor in Mathematics at the University of Torino in 2015 before moving to Scuola Normale Superiore in Pisa, where he obtained in 2017 his Master in Mathematics, under the supervision of Prof. Luigi Ambrosio. He joined IST Austria as PhD student in Fall 2017 where he joined the group of Prof. Jan Maas. He is also member of the Doctoral Program (DK) "Dissipation and Dispersion in Nonlinear PDEs" funded by the Austrian Science Fund FWF (project number: W1245) and member of the project SFB "Taming Complexity in Partial Differential Systems" (project number: F65), in collaboration with the University of Vienna and the Vienna University of Technology . His research focuses on the theory of optimal transport, with particular attention to discrete transport problems and their connection to evolutive equations and gradient flows. He will start a PostDoc position at the Hausdorff Center of Mathematics in Bonn in Fall 2021, under the mentoring of Prof. Karl-Theodor Sturm and Prof. Franca Hoffmann.

List of Collaborators and Publications

During my PhD at IST Austria, I had the possibility to collaborate with different people from Vienna and abroad. The list of my collaborators includes Jan Maas (supervisor), Dario Feliciangeli, Dominik Forkert, Marco di Francesco, Augusto Gerolin, Peter Gladbach, Eva Kopfer, Simone di Marino, Emanuela Radici, and Ulisse Stefanelli.

The following works are included in this thesis:

- Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Homogenisation of one-dimensional discrete optimal transport. *J. Math. Pures Appl. (9)*, 139:204–234, 2020.
- Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Discrete-to-continuum limits of dynamical transport problems on periodic graphs. *In preparation*, 2021.
- Dominik Forkert, Jan Maas, and Lorenzo Portinale. Evolutionary Γ-convergence of entropic gradient flow structures for Fokker–Planck equations in multiple dimensions. *Submitted*, arXiv:2008.10962, 2020.
- Dario Feliciangeli, Augusto Gerolin, and Lorenzo Portinale. A non-commutative entropic optimal transport approach to quantum composite systems at positive temperature. *Submitted, arxiv:2106.11217*, 2021.

The following works are not included in this thesis, but are the result of scientific collaborations during my PhD at IST Austria as well:

- Lorenzo Portinale and Ulisse Stefanelli. Penalization via global functionals of optimal-control problems for dissipative evolution. *Adv. Math. Sci. Appl., 28*, pages 425–447, 2019.
- Marco Di Francesco, Simone Di Marino, Lorenzo Portinale, and Emanuela Radici. Transport problems with non linear mobilities: a particle approximation result. *In preparation*, 2021.

Table of Contents

CHAPTER 1

Introduction

The theory of Optimal Transport has been a trending topic in the community of calculus of variations for the last twenty years. Based on very intuitive model of transportation introduced by Monge [Mon81] in the 18th century and the works of Kantorovich [Kan42] in the first half of the 20th century, the study of the problem had shown deep and somehow surprising developments in several directions, not only limited to the analysis of the model itself but rather towards various applications and links to a broader class of mathematical topics.

In simple words, the optimal transport problem consists in finding the cheapest way to transport a given distribution of mass μ_0 into a target one μ_1, with respect to some cost function $c(x, y)$ that determines the price of moving the mass from a location x to a final location y.

In mathematical terms, this is modeled by two probability measures $\mu_0, \mu_1 \in \mathcal{P}(X)$ over a space X, a cost function $c : X \times X \to [0, +\infty]$, and the goal is to find the minimal transport cost, that means to solve the optimisation problem

$$\inf \left\{ \int c(x, y) \, \mathrm{d}\pi(x, y) \ : \ \pi \in \Pi(\mu_0, \mu_1) \right\} , \tag{1.1}$$

where $\Pi(\mu_0, \mu_1) \subset \mathcal{P}(X \times X)$ denotes the associated set of *couplings*, namely probability measures on $X \times X$ having μ_0 and μ_1 as marginals (i.e. projections onto the first and second coordinate, respectively). A coupling $\pi \in \Pi(\mu_0, \mu_1)$ describes then one possible way to move the mass from the initial distribution to the target one.

In the particular case when X is endowed with a metric structure, that is (X, d) is a metric space, a typical choice for the cost function is to consider a power of the distance, i.e. $c(x, y) := d(x, y)^p$, for some $p \in [1, +\infty]$. This defines the p-Kantorovich-Rubinsthein-Wasserstein distance $\mathbb{W}_p$ as

$$\mathbb{W}_p(\mu_0, \mu_1)^p = \min \left\{ \int d^p(x, y) \, \mathrm{d}\pi(x, y) \ : \ \pi \in \Pi(\mu_0, \mu_1) \right\} . \tag{1.2}$$

The original problem proposed by Monge corresponds to the linear cost $p = 1$, which quite surprisingly turned out to be one of the most challenging cases to understand (see [San15, Chapter 3] for a detailed discussion about the Monge's problem).

Throughout this introduction, we shall focus on the the particular case of $p = 2$ and restrict our discussion to the euclidean setting, that is $X = \mathbb{R}^d$ (or a convex, bounded domain $\Omega \subset \mathbb{R}^d$), although many results hold true for more general cost functions and spaces. The theory for the quadratic cost is particularly rich: first and foremost, the well-known Brenier's theorem [Bre87] provides a characterisation of the optimal couplings when μ_0 is absolutely continuous with respect to the Lebesgue measure $\mathscr{L}^d$ in $\mathbb{R}^d$. In this case, the optimal $\pi^* \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ is induced by a *transport map* $T : \mathbb{R}^d \to \mathbb{R}^d$, in the sense that $\pi^* = (\mathsf{id}, T)_\# \mu_0$, where $T = \nabla\varphi$ is the gradient of a convex potential φ. See also [McC95], [Gig11] for generalisations of this result.

For the sake of the exposition, we shall not discuss here every major aspect of this extremely broad theory (for a detailed and thorough discussion, see the classical book of Villani [Vil08]), but we shall rather focus on some features which play a crucial role inside this thesis.

The definition (1.2) of $\mathbb{W}_2$ is often referred to as the *static formulation* of the optimal transport problem. In the euclidean setting (or more generally on Riemannian manifolds), the Benamou–Brenier Theorem [BB00] provides a surprising equivalent formulation of the distance $\mathbb{W}_2$ in $\mathcal{P}_2(\mathbb{R}^d)$ which is of *dynamical nature*. The celebrated result indeed states that the quadratic optimal transport problem can be equivalently recast as

$$\mathbb{W}_2(\mu_0, \mu_1)^2 = \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|\nu_t|^2}{\mu_t} \, \mathrm{d}x \, \mathrm{d}t \ : \ (\mu_t, \nu_t)_t \in \mathsf{CE}(\mu_0, \mu_1) \right\} , \tag{1.3}$$

where $\mathsf{CE}(\mu_0, \mu_1)$ denotes the set of all the solutions to the continuity equation

$$\partial_t \mu_t + \nabla \cdot \nu_t = 0 \quad \text{on } \mathbb{R}^d, \qquad \text{a.e. } t \in [0,1] \, , \qquad \mu_{t=0} = \mu_0, \ \mu_{t=1} = \mu_1 \, , \tag{1.4}$$

interpreted in the sense of distributions $\mathcal{D}'((0,1) \times \mathbb{T}^d)$.

In this new dynamical interpretation of $\mathbb{W}_2$, the cheapest transport cost is achieved minimising, between all possible evolutions (i.e. solutions to the continuity equation) $t \mapsto \mu_1$ from μ_0 to μ_1, the one that corresponds to the minimal *total kinetic energy*. In this picture, ν_t describes the flux of mass at time $t \in [0,1]$. Indeed, for any solution $(\mu_t, \nu_t)_t \in \mathsf{CE}(\mu_0, \mu_1)$ of the form $\nu_t = v_t \mu_t$, the curve $(v_t)_t$ represents the time-dependent velocity field associated with the evolution of the mass. In other words, μ_t describes a flow of particles X_t in $\mathbb{R}^d$ evolving accordingly to the equation $\dot{X}_t = v_t(X_t)$, for $t \in [0,1]$.

A key consequence of the result in (1.3) is the reinterpretation of the Wasserstein distance as a geodesic distance once we give to the space of probability measure an interpretation (at least formally) of an infinite dimensional Riemannian manifold. This has been firstly proposed by Otto in his seminal work [Ott01] and then more intensively investigated in the setting of general metric spaces by Ambrosio, Gigli, and Savaré in [AGS08].

The Benamou–Brenier formula (1.3) plays a central role as the link between the theory of optimal transport and different fields of mathematics, including partial differential equations [JKO98], functional inequalities [OV00] and the novel notion of Lott–Villani–Sturm's synthetic Ricci curvature bounds for metric measure spaces [LV07], [LV09], [Stu06].

The main object of this thesis is a notion of optimal transport in discrete settings which is based on a similar approach, structured as a dynamical formulation *à la* Benamou–Brenier as in (1.3). This notion of transport is a natural discrete counterpart of (1.3), introduced in independent works by Maas [Maa11] (in the setting of Markov chains) and by Mielke [Mie11] (in the context of reaction-diffusion systems).

Discrete optimal transport

There are different, equivalent ways to introduce the setting of discrete optimal transport problems. One possibility is to consider transportation problems between measures over finite-states Markov chains satisfying a *detailed-balance condition* (or, in other words, reversible chains), as in the original work of Maas [Maa11]. Equivalently, this can be recast in terms of *symmetric, weighted graphs* with a finite number of nodes and edges. In this thesis, and in particular in this introduction, we adopt the second point of view.

We consider a weighted graph $(\mathcal{X}, \mathcal{E}, \omega)$ with finite set of *nodes* $\mathcal{X}$, symmetric set of *edges* $\mathcal{E} \subset \mathcal{X} \times \mathcal{X}$, and *weight function* $\omega : \mathcal{E} \to [0, +\infty)$. We write $y \sim x$ if $(x, y) \in \mathcal{E}$. Moreover, we fix a reference probability measure $\pi \in \mathscr{P}(\mathcal{X})$ and for any $m \in \mathscr{P}(\mathcal{X})$, we denote by $r = m/\pi$ its density.

Following [Maa11], one introduces a discrete transportation metric $\mathcal{W}$ on $\mathscr{P}(\mathcal{X})$ as

$$\mathcal{W}(m_0, m_1)^2 := \inf \left\{ \int_0^1 \frac{1}{2} \sum_{(x,y)\in\mathcal{E}} \frac{1}{\omega(x,y)} \frac{|j_t(x,y)|^2}{\widehat{r}_t(x,y)} \,\mathrm{d}t \;:\; (m_t, j_t)_t \in \mathsf{CE}_{\mathcal{X}}(m_0, m_1) \right\}, \tag{1.5}$$

for $m_0, m_1 \in \mathscr{P}(\mathcal{X})$, where $\mathsf{CE}_{\mathcal{X}}(m_0, m_1)$ denotes the set of solutions to the discrete continuity equation, i.e. curves of measures satisfying for $t \in [0,1]$, $x \in \mathcal{X}$

$$\partial_t m_t(x) + \frac{1}{2} \sum_{y \sim x} \left(j_t(x,y) - j_t(y,x) \right) = 0, \quad m_{t=0} = m_0, \; m_{t=1} = m_1 \,, \tag{1.6}$$

and $\widehat{r}_t(x,y) := \theta(r_t(x), r_t(y))$ for some continuous, 1-homogeneous, and positive function $\theta : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$.

The intuition behind the above definition comes from the natural discretisations of the energy and the continuity equation in the continuous setting appearing in the Benamou–Brenier formulation (1.3). In particular, for any solution $(m_t, j_t)_t \in \mathsf{CE}_{\mathcal{X}}$ and any given edge $(x,y) \in \mathcal{E}$, the value of $j_t(x,y)$ represents the flux of mass flowing from the node $x \in \mathcal{X}$ to the node $y \in \mathcal{X}$, and the difference appearing in the (discrete) continuity equation (1.6) plays the role of the divergence in the continuous setting.

A key difference between the discrete and the continuous framework is the presence of the average function θ. This reflects some freedom of choice in defining the mass (or the density) of a measure $m \in \mathcal{M}(\mathcal{X})$ over the edges.

In particular, the definition of $\mathcal{W}$ depends on the whole structure: the weighted graph $(\mathcal{X}, \mathcal{E}, \omega)$ itself, the reference measure π, and the choice of θ. This dependence turns out to be less trivial than expected even in very simple cases, as we are going to see throughout this thesis.

Finite-volume. A notable example of a discrete setting is the *finite-volume framework*. Given $\Omega \subset \mathbb{R}^d$ any open, convex, and bounded set, we consider a finite partition $\mathcal{T}$ of Ω into convex sets $K \subset \Omega$ and points $x_K \in K$ such that $x_L - x_K \perp \partial K \cap \partial L$, whenever K and L are neighboring cells. We associate to it the graph structure $\mathcal{X} = \mathcal{T}$ and $x_K \sim x_L$ if and only if $\mathscr{H}^{d-1}(\partial K \cap \partial L) > 0$ (see Figure 1.1).

Following [GKM20], a natural choice of the reference measure and of the weight function is

$$\pi(x_K) := \mathfrak{m}(K), \quad \omega(x_K, x_L) := \frac{\mathscr{H}^{d-1}(\partial K \cap \partial L)}{|x_K - x_L|} \,, \tag{1.7}$$

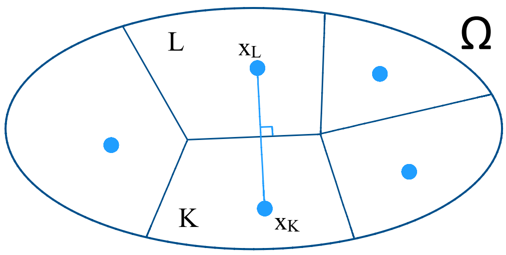

Figure 1.1: An admissible mesh with cells $K, L, \dots$ on a domain $\Omega \subset \mathbb{R}^d$.

where $\mathfrak{m} \in \mathcal{P}(\bar{\Omega})$ is a given reference measure. We also denote by $[\mathcal{T}]$ the maximum diameter of any $K \in \mathcal{T}$.

The first question we would like to pose in this thesis is the following.

$$\textbf{Question}\text{: what can we say about the behaviour of } (\mathscr{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}}) \text{ as } [\mathcal{T}] \to 0? \tag{1.8}$$

The first result in this direction has been obtained in [GM13], where the authors proved the convergence of $(\mathscr{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}})$ to the Wasserstein space $\left(\mathscr{P}_2(\mathbb{T}^d), \mathbb{W}_2\right)$ in the Gromov–Hausdorff sense in the special case of the orthogonal lattice on the torus $\mathbb{T}^d$. A result for more general meshes has been obtained in the subsequent work [GKM20], where the authors proved that the convergence to the 2-Wasserstein space is essentially equivalent to an *asymptotic isotropy condition* on the mesh. In the special case of symmetric average θ, this condition reads as

$$\frac{1}{2} \sum_{y \in \mathcal{T}} \omega(x,y)\,(y-x) \otimes (y-x) \le \pi(x)\,(I_d + \eta_{\mathcal{T}}(x))\,, \quad \forall x \in \mathcal{T}\,. \tag{1.9}$$

where $\sup_{x \in \mathcal{T}} \|\eta_{\mathcal{T}}(x)\| \to 0$ as $[\mathcal{T}] \to 0$. Both [GM13] and [GKM20] work with $\mathfrak{m} = \mathscr{L}^d$.

As evident from the first results of [GM13], [GKM20], an important feature of the problem is the interplay between the *geometry* of the graph and the *energies* that define the transport problem (in the particular case (1.5) of $\mathcal{W}_{\mathcal{T}}$, the choice of θ). In this work, we are going to investigate in detail how the change of geometry effects the limit behavior of $\mathcal{W}_{\mathcal{T}}$ and how the limit transport problem can be computed, starting from the discrete transport energies, with particular attention to the periodic setting.

In the following, we discuss the main contributions of this thesis concerning the study of discrete transport problems. For each different setting, we first introduce the problem, discuss the main results, and analyse several applications and examples of interest.

Discrete-to-continuum convergence of transport costs

One of the main contribution of this thesis is the study of transport costs as in (1.5) on periodic graphs in $\mathbb{R}^d$ [GKMP21], [GKMP20]. We start with one-dimensional, quadratic problems on the circle $\mathcal{S}^1$, whereas in the second part we study more general transport problems in arbitrary dimension and beyond the quadratic cost. Our analysis covers the setting of periodic finite-volume partitions in $\mathbb{R}^d$, both the isotropic and the non-isotropic case, and several different examples of energy functionals.

One-dimensional case, quadratic cost.

The first problem we address is a periodic, one-dimensional setting. We consider the circle $\mathbb{S}^1$ and a partition $\mathcal{T}$ in K cells (intervals) of diameters π_k and points z_k, r_k (see Figure 1.2).

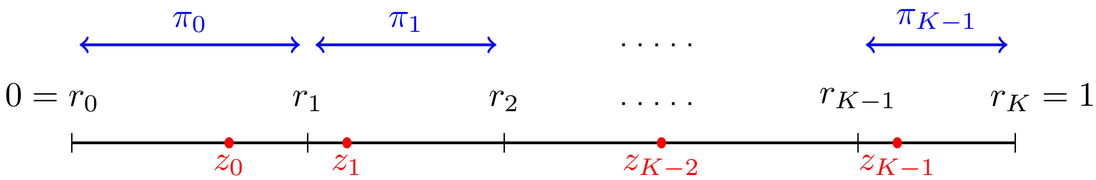

Figure 1.2: Partition of $\mathbb{S}^1$ in K cells.

For every $N \in \mathbb{N}$, we then rescale the problem and consider the sequence of $1/N$-periodic partitions $\mathcal{T}_N = \mathcal{T}/N$.

Our result describes the behaviour of the associated discrete transport costs $\mathcal{W}_{\mathcal{T}_N}$.

Precisely, we are able to show that the metric space $(\mathscr{P}(\mathcal{T}_N), \mathcal{W}_{\mathcal{T}_N})$ converges in the Gromov–Hausdorff sense to $(\mathscr{P}(\mathbb{S}^1), c(\theta, \{z_k\}, \{\pi_k\})\mathbb{W}_2)$ as $N \to \infty$, where

$$c\Big(\theta, \{z_k\}, \{\pi_k\}\Big) := \inf \left\{ \sum_{k=0}^{K-1} \frac{|z_{k+1} - z_k|}{\theta\left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\right)} \ : \ \sum_{k=0}^{K-1} m_k = 1, \ m_K = m_0 \right\}. \tag{1.10}$$

Quite surprisingly, the limit distance is not always the 2-Wasserstein distance, despite the convergence of the corresponding gradient flows (see Remark 2.1.3).

The limit distance is obtained by minimising the distribution of the mass according to the geometry of the mesh and the average θ, which translates into a non-trivial effective mobility $c\Big(\theta, \{z_k\}, \{\pi_k\}\Big)$. In particular, our result quantitatively shows how the discrete transport can take advantage of the graph microstructure in order to reduce the total cost of transportation, with respect to the usual euclidean one.

The key observation needed to understand this phenomena is that the discrete transport problems are "sensitive" to microscopic oscillations of the densities. To present a formal argument to explain this, suppose we consider a smooth solution to the continuity equation $\partial_t \mu + \partial_x j = 0$ and fix a discrete measure $\alpha \in \mathcal{P}(\mathcal{T})$. We define the discrete measure $m \in \mathcal{P}(\mathcal{T}_N)$ by assigning mass $\alpha(k)\mu\Big([\frac{n}{N}, \frac{n+1}{N}]\Big)$ to the corresponding cell $A_{n;k}$, see Figure 1.3.

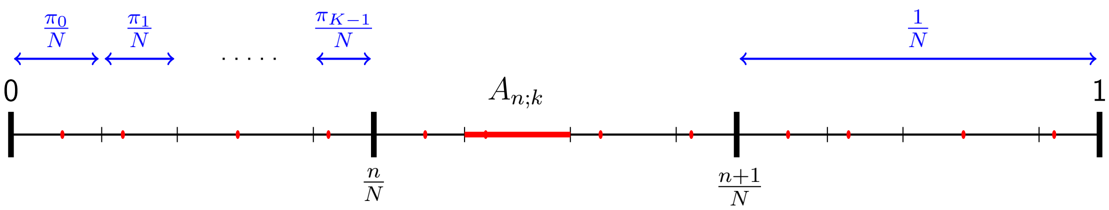

Figure 1.3: The mesh $\mathcal{T}_N$ on $\mathcal{S}^1$ with cells $\{A_{n;k}\}_{n,k}$.

Note that each interval of the form $\left([\frac{n}{N}, \frac{n+1}{N}]\right)$ receives the same mass at the discrete and at the continuous level, equals to $\mu\left([\frac{n}{N}, \frac{n+1}{N}]\right)$. Nonetheless, within each intervals of this form, the measure α introduces a discrete density oscillation.

This operation is "invisible" at the continuous level because the mesh is getting finer. In contrast, the discrete transport cost is quantitatively sensitive to the oscillation. Let J be the associated discrete momentum vector field that solves the continuity equation for m. Assuming such vector field to be regular enough, we may then estimate the discrete energy by

$$\begin{aligned}\sum_{n=0}^{N-1}\sum_{k=0}^{K-1} d_{k,k+1}\frac{J_t^2(n;k,k+1)}{\theta_{k,k+1}\left(\frac{Nm_t(n;k)}{\pi_k}, \frac{Nm_t(n;k+1)}{\pi_{k+1}}\right)} &\approx \frac{1}{N}\sum_{n=0}^{N-1}\frac{J_t^2(n;0,1)}{\mu\left([\frac{n}{N}, \frac{n+1}{N}]\right)}\sum_{k=0}^{K-1}\frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{\alpha(k)}{\pi_k}, \frac{\alpha(k+1)}{\pi_{k+1}}\right)} \\ &= \sum_{k=0}^{K-1}\frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{\alpha(k)}{\pi_k}, \frac{\alpha(k+1)}{\pi_{k+1}}\right)}\int\frac{|j|^2}{\mu} \ .\end{aligned}$$

The cheapest discrete transport cost thus corresponds to choose α^* as the minimiser of the corresponding cell problem

$$\alpha^* \in \operatorname*{argmin}_{\alpha\in\mathcal{P}(\mathcal{T})}\left\{\sum_{k=0}^{K-1}\frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{\alpha(k)}{\pi_k}, \frac{\alpha(k+1)}{\pi_{k+1}}\right)}\right\} ,$$

hence recovering the continuous energy appearing in the Benamou–Brenier formula up to a multiplicative correction, which indeed suggests our main result.

A rigorous argument based on this heuristic requires suitable spatial regularity results for m and J. To this purpose, a central result (Proposition 2.5.3) in our work is to show that discrete curves can be approximated by curves of similar energy, which enjoys good Lipschitz bounds for J as well as good Lipschitz bounds for m up to oscillations within each cell.

As a corollary of the previous result we obtain that for smooth mobilities θ, the cell problem (1.10) is equal to 1 if and only if the isotropy condition holds, which in our periodic one-dimensional setting means that there exist constants $\lambda_{k,k+1}, s \in (0,1)$ such that the following conditions hold for $k = 0,\dots,K-1$:

$$\begin{aligned} r_{k+1} &= \lambda_{k,k+1}z_{k+1} + (1-\lambda_{k,k+1})z_k + s \ , \\ \theta_{k,k+1}(a,b) &\le \lambda_{k,k+1}a \quad\ + (1-\lambda_{k,k+1})b \qquad \text{for any } a,b\ge 0 \ .\end{aligned}$$

Thus, in order to obtain $\mathbb{W}_2$ in the limit, the asymmetry of the means $\theta_{k,k+1}$ should reflect the relative location of the points z_k, r_{k+1}, and z_{k+1}.

For non-isotropic meshes, the optimal mass configuration presents oscillations which exploit the lack of symmetry of the mesh, giving rise to a strictly cheaper cost. On one side, the theorem includes the convergence to $\mathbb{W}_2$ for isotropic meshes. On the other side, it provides explicit counterexamples to this convergence in case of non-isotropic meshes.

Arbitrary dimension, generic cost.

In a subsequent work [GKMP21], we extend our previous one-dimensional result to arbitrary dimensions and to general cost functions. We consider locally finite, $\mathbb{Z}^d$-periodic graphs $(\mathcal{X}, \mathcal{E})$

in $\mathbb{R}^d$ (see Picture 1.4), and study the behavior of energies of the form

$$\inf_j \left\{ \int_0^1 F(m_t, j_t)\, \mathrm{d}t \;:\; (m_t, j_t)_{t\in[0,1]} \text{ solves } \mathrm{CE}_{\mathcal{X}} \right\}, \quad (m_t)_t \subset \mathcal{M}_+(\mathcal{X}) = \mathbb{R}_+^{\mathcal{X}},$$

where $\mathrm{CE}_{\mathcal{X}}$ is a discrete continuity equation and F is a given lower-semicontinuous, convex, and local function which grows more than linearly with respect to the second variable. This is the content of Chapter 3.

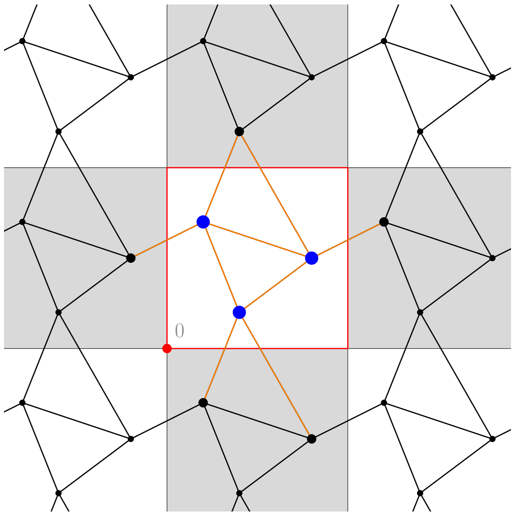

Figure 1.4: A $\mathbb{Z}^2$-periodic graph in $\mathbb{R}^2$. In red, the unitary cube $[0,1]^2 \subset \mathbb{R}^2$.

The precise abstract setup is the following. We consider a set $\mathcal{X} = \mathbb{Z}^d \times \mathbb{V}$, where $\mathbb{V}$ is a finite set. We shall use coordinates $x = (z, v) \in \mathcal{X}$ and denote them by $x_{\mathsf{z}} := z$, $x_{\mathsf{v}} := v$. The set of edges $\mathcal{E} \subset \mathcal{X} \times \mathcal{X}$ is symmetric and $\mathbb{Z}^d$-periodic, and we write $x \sim y$ whenever $(x, y) \in \mathcal{E}$.

One can use the following geometric interpretation of $(\mathcal{X}, \mathcal{E})$, regarding $\mathcal{X}$ as a $\mathbb{Z}^d$-periodic subset of $\mathbb{R}^d$: we choose $\mathbb{V}$ to be any finite subset of $[0,1)^d$ and use the identification $(z, v) \equiv z + v$ (as in Figure 1.4). It turns out that the embedding plays no role in the formulation of the discrete problem (and hence the results), which is why we work with the abstract setting.

We then fix the cost function: we denote by $\mathbb{R}_a^{\mathcal{E}}$ the set of discrete, skew-symmetric momentum vector fields and we pick $F : \mathbb{R}_+^{\mathcal{X}} \times \mathbb{R}_a^{\mathcal{E}} \to \mathbb{R} \cup \{+\infty\}$ a convex, lower-semicontinuous function, which is local and with at least linear growth in the momentum variable (see Assumption 3.2.3 in Chapter 3 for a precise definition).

In order to avoid the problem to be too degenerate, we make an additional assumption of the domain of F. In particular, we assume that there exist $\mathbb{Z}^d$-periodic functions $m^\circ \in \mathbb{R}_+^{\mathcal{X}}$ and $J^\circ \in \mathbb{R}_a^{\mathcal{E}}$ divergence-free, such that $(m^\circ, J^\circ) \in \mathsf{D}(F)^\circ$,

Examples of periodic graphs. A particular class of $\mathbb{Z}^d$-periodic graphs can be associated to $\mathbb{Z}^d$-periodic finite volume partitions (FVPs) of $\mathbb{R}^d$ (or equivalently, finite volume partitions

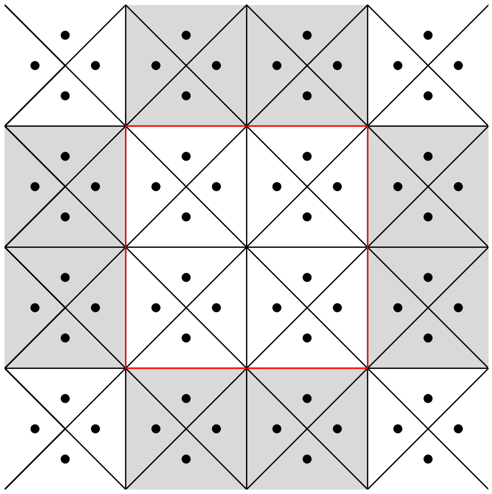

Figure 1.5: A $\mathbb{Z}^2$-periodic finite volume partition of $\mathbb{R}^2$. In red, the unitary cube $[0,1]^2 \subset \mathbb{R}^2$.

of $\mathbb{T}^d$), in the sense of (1.7), see Figure 1.5 for a two-dimensional example. The analysis of quadratic transport problems over this special class of graphs is the content of Section 3.9 in Chapter 3, with particular attention to the role of the isotropic condition in this context.

Example of transport energies. In many interesting cases, the function F takes one of the following forms:

$$F(m,J) = \sum_{x \in \mathcal{X}^Q} F_x\Big(m(x), \big(J(x,y)\big)_{y\sim x}\Big), \quad F(m,J) = \sum_{(x,y)\in\mathcal{E}^Q} F_{xy}\Big(m(x), m(y), J(x,y)\Big).$$

We refer to these classes of examples respectively as *vertex-based* and *edge-based*. A particular example of the second group is the *Wasserstein-like* type of costs (which includes the ones presented in the previous section in the one-dimensional setting), which are of the form

$$F(m,J) = \frac{1}{p} \sum_{(x,y)\in\mathcal{E}^Q} \frac{|J(x,y)|^p}{\Lambda\Big(q_{xy} m(x), q_{yx} m(y)\Big)^{p-1}} \tag{1.11}$$

where $q_{xy}, q_{yx} > 0$ are fixed parameters defined for $(x,y) \in \mathcal{E}^Q$, $p \geq 1$, and Λ is a suitable mean.

It is interesting to note that we are able to include the linear case $p = 1$. Nonetheless, the lack of superlinear growth will necessarily be reflected in weaker results, as we are going to discuss.

The rescaled problem. Let $\mathbb{T}^d_\varepsilon = (\varepsilon\mathbb{Z}/\mathbb{Z})^d = \{[\varepsilon z] \ : \ z \in \mathbb{Z}^d\}$ be the discrete torus of mesh size ε. The rescaled graph $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$ is constructed by rescaling the $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$ and wrapping it around the torus. More precisely, define the map $T^{\bar{z}}_\varepsilon : \mathcal{X} \to \mathcal{X}_\varepsilon$, $T^{\bar{z}}_\varepsilon(z,v) = \big([\varepsilon(\bar{z}+z)], v\big)$. The rescaled graph is defined as

$$\mathcal{X}_\varepsilon := \mathbb{T}^d_\varepsilon \times \mathbb{V} \quad \text{and} \quad \mathcal{E}_\varepsilon := \Big\{\Big(T^0_\varepsilon(x), T^0_\varepsilon(y)\Big) \ : \ (x,y) \in \mathcal{E}\Big\}.$$

In a similar way, we define the rescaled cost function $\mathcal{F}_\varepsilon : \mathbb{R}_+^{\mathcal{X}_\varepsilon} \times \mathbb{R}_a^{\mathcal{E}_\varepsilon} \to \mathbb{R} \cup \{+\infty\}$ as

$$\mathcal{F}_\varepsilon(m, J) = \sum_{z \in \mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{\tau_\varepsilon^z m}{\varepsilon^d}, \frac{\tau_\varepsilon^z J}{\varepsilon^{d-1}}\right),$$

where $\tau_\varepsilon^z m := m \circ T_\varepsilon^z$ and $\tau_\varepsilon^z J := J \circ (T_\varepsilon^z, T_\varepsilon^z)$, see Figure 1.6.

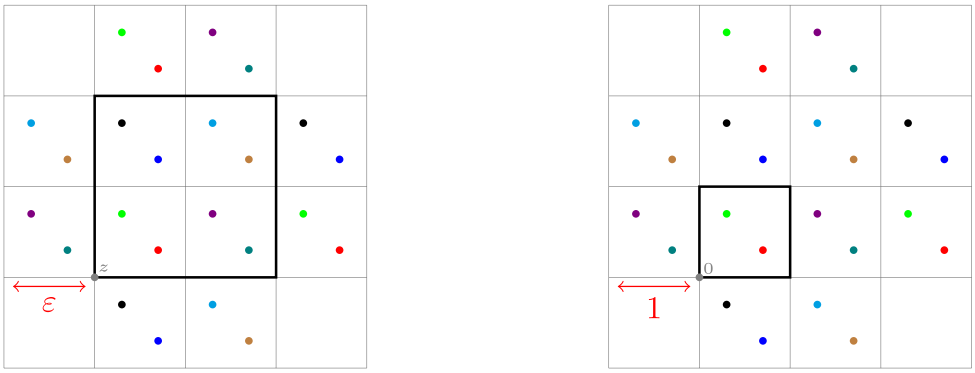

Figure 1.6: On the left, the value of a function $\psi : \mathcal{X}_\varepsilon \to \mathbb{R}$ correspond to different colors over the nodes. On the right, the corresponding values of $\tau_\varepsilon^z \psi : \mathcal{X} \to \mathbb{R}$.

The second step is to describe the evolution of discrete measures. We fix $T > 0$ and we say that a pair $(\mathbf{m}, \mathbf{J})$ solves the discrete continuity equation on the rescaled graph $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$ if $\mathbf{m} : [0, T] \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ is continuous, $\mathbf{J} : [0, T] \to \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ is Borel measurable, and

$$\partial_t m_t(x) + \operatorname{div} J(x) = 0\,, \quad \operatorname{div} J(x) = \sum_{y \sim x} J_t(x, y)\,,$$

for all $x \in \mathcal{X}_\varepsilon$ in the sense of distributions. We use the notation $(\mathbf{m}, \mathbf{J}) \in \mathsf{CE}_\varepsilon$.

The main goal is then to study the rescaled dynamical transport problem

$$\mathcal{A}_\varepsilon(\mathbf{m}) := \inf_{\mathbf{J}} \left\{ \int_0^T \mathcal{F}_\varepsilon(m_t, J_t)\, \mathrm{d}t \ : \ (\mathbf{m}, \mathbf{J}) \in \mathsf{CE}_\varepsilon \right\}$$

and their asymptotic behavior as $\varepsilon \to 0$.

In order to compare the discrete with the continuous setting, we shall introduce suitable embedding maps. In this periodic setting, it is natural to consider for $\varepsilon > 0$, the functions

$$\iota_\varepsilon : \mathbb{R}_+^{\mathcal{X}_\varepsilon} \to \mathcal{M}_+(\mathbb{T}^d)\,, \quad \iota_\varepsilon m := \varepsilon^{-d} \sum_{z \in \mathbb{Z}_\varepsilon^d} \left(\sum_{\substack{x \in \mathcal{X}_\varepsilon \\ x_z = z}} m(x) \right) \mathscr{L}^d|_{Q_\varepsilon^z}\,,$$

for every discrete measure $m \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$.

It turns out that the discrete rescaled energies $\mathcal{A}_\varepsilon$ Γ-converge as $\varepsilon \to 0$ to a homogenised energy $\mathbb{A}_{\mathrm{hom}}$ which is the relaxation of the functional

$$\inf_{\boldsymbol{\nu}} \left\{ \int_{(0,T)\times\mathbb{T}^d} f_{\mathrm{hom}}\left(\frac{\mathrm{d}\mu}{\mathrm{d}t \otimes \mathrm{d}x}, \frac{\mathrm{d}\xi}{\mathrm{d}t \otimes \mathrm{d}x} \right) \mathrm{d}t \otimes \mathrm{d}x \ : \ (\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathsf{CE} \right\},$$

for a suitable homogenised density $f_{\hom}$. Here CE denotes the set of continuous solutions to the continuity equation $\partial_t \boldsymbol{\mu} + \nabla \cdot \boldsymbol{\nu} = 0$ in the sense of distributions $\mathcal{D}'((0,T) \times \mathbb{T}^d)$, where $\boldsymbol{\mu} = \mathrm{d}t \otimes \mu_t \in \mathcal{M}_+((0,T) \times \mathbb{T}^d)$ and $\boldsymbol{\nu} = \mathrm{d}t \otimes \nu_t \in \mathcal{M}^d((0,T) \times \mathbb{T}^d)$.

The limit energy density $f_{\hom}$ can be explicitly computed as follows. Set $\mathcal{X}^Q := \{x \in \mathcal{X} \ : \ x_{\mathsf{z}} = 0\}$ and $\mathcal{E}^Q := \{(x,y) \in \mathcal{E} \ : \ x \in \mathcal{X}^Q\}$.

We say that that $m \in \mathbb{R}_+^{\mathcal{X}}$ *represents* $\rho \in \mathbb{R}_+$ if m is $\mathbb{Z}^d$-periodic and $\sum_{x \in \mathcal{X}^Q} m(x) = \rho$. We also say that a vector field $J \in \mathbb{R}_a^{\mathcal{E}}$ *represents* a vector $j \in \mathbb{R}^d$ if J is $\mathbb{Z}^d$-periodic, divergence free (that is $\operatorname{div} J = 0$), and its *effective flux* equals j, i.e.

$$\mathsf{Eff}(J) := \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} J(x,y)(y_{\mathsf{z}} - x_{\mathsf{z}}) = j\,.$$

We can then express $f_{\hom}$ via the following *cell formula*

$$f_{\hom}(\rho, j) := \inf \Big\{ F(m,J) \ : \ (m,J) \in \mathsf{Rep}(\rho,j) \Big\}\,, \quad \forall \rho \in \mathbb{R}_+\,, j \in \mathbb{R}^d\,.$$

The limit energy $f_{\hom}(\rho,j)$ thus corresponds to the minimal cost to distribute mass ρ and flux j among all the discrete vector fields with null divergence.

In the one dimensional setting, where F is the quadratic cost, we recover the original cell formula presented in the previous section. Our result covers several examples, including quadratic costs in arbitrary dimension, isotropic meshes of $\mathbb{T}^d$, and flow-based models. Once again, we observe surprising behaviors, even in the case of quadratic transport-like costs as in (1.11) with $p = 2$. Despite the Riemannian-structure of the discrete problems, the limit problem is in general not Riemannian, but only Finsler. We can indeed prove that, in this case, $f_{\hom}(1,j) = \|j\|_{\hom}$, where $\|\cdot\|_{\hom}$ is a homogenised norm which is general fails to be induced by a scalar product, if no additional symmetry property on $(\mathcal{X}, \mathcal{E})$ is assumed, such as the isotropy condition in the finite-volume framework.

As in the one-dimensional case, the previous result relies on new regularisation techniques at the discrete level. Moreover, the nature of the proof is substantially different from the one-dimensional setting as it is based on some crucial analysis of discrete divergence equations on periodic graphs. A major additional difficulty arises from the fact that in arbitrary dimension, one must deal with non-trivial divergence-free vector fields J, which reduce to the constant ones in $d = 1$. Furthermore, the lack of homogeneity in the momentum variable for F does not allow us to apply similar techniques as in [GKMP20], which instead strongly rely on the 2-homogeneity of the transport cost.

Topology and compactness. Our convergence result covers the linear growth case. In this generality though, one has to deal with some lack of regularity.

For instance, we have no guarantee that measures $\boldsymbol{\mu} \in \mathcal{M}_+((0,T) \times \mathbb{T}^d)$ with finite energy $\mathbb{A}_{\hom}(\boldsymbol{\mu}) < \infty$ admit a momentum vector field $\boldsymbol{\nu} \in \mathcal{M}^d((0,T) \times \mathbb{T}^d)$, with $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathsf{CE}$, which admits a nice disintegration in time $\boldsymbol{\nu} = \mathrm{d}t \otimes \nu_t$. A simple example of this phenomenon corresponds to the total momentum energy $f_{\hom}(\rho,j) = |j|$ and $\boldsymbol{\mu} := \mathrm{d}t \otimes \delta_{\gamma_t}$, where $\gamma \in \mathsf{BV}((0,1);\mathbb{T}^d)$ is the Heaviside function. It is not hard to see that there is no $\boldsymbol{\nu} = \mathrm{d}t \otimes \nu_t \in \mathcal{M}^d((0,1) \times \mathbb{T}^d)$ which solves the continuity equation for $\boldsymbol{\mu}$. A solution is instead given by

$$\boldsymbol{\nu} := \delta_{1/2} \otimes \mathscr{H}^1|_{[0,1]}\,, \quad \text{and} \quad \mathbb{A}_{\hom}(\boldsymbol{\mu}, \boldsymbol{\nu}) = |\boldsymbol{\nu}|((0,1) \times \mathbb{T}^d) = 1 < \infty\,.$$

Taking these considerations into account, we can only hope to obtain good properties of finite energy curves of measures in a weaker sense. Precisely, suppose that $\mathbf{m}^\varepsilon \in \mathbb{R}_+^{(0,T)\times\mathcal{X}_\varepsilon}$ is a sequence of measures satisfying

$$\sup_{\varepsilon>0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) < \infty \quad \text{and} \quad \sup_{\varepsilon>0} \mathbf{m}^\varepsilon((0,T)\times\mathcal{X}_\varepsilon) < \infty\,. \tag{1.12}$$

Then the sequence $\{\iota_\varepsilon \mathbf{m}^\varepsilon\}_\varepsilon \subset \mathcal{M}_+((0,T)\times\mathbb{T}^d)$ is compact with respect to the narrow topology (i.e. in duality with continuous and bounded functions) and any limit point $\boldsymbol{\mu}$ admits a disintegration $\boldsymbol{\mu} = \mathrm{d}t \otimes \mu_t$ with $t \mapsto \mu_t \in \mathsf{BV}((0,T);\mathcal{M}_+(\mathbb{T}^d))$.

The examples above show that no better regularity can be sought without further assumptions on F. The right condition to impose in order to obtain higher regularity is a *superlinearity* assumption on F. Precisely, assume that there exists a a function $\theta : [0,\infty) \to [0,\infty)$ with $\lim_{t\to\infty} \frac{\theta(t)}{t} = \infty$ and a constant $C \in \mathbb{R}$ such that

$$F(m,J) \geq \left(\theta\left(\frac{J_0}{m_0}\right) - C\right) m_0 - C\,, \quad J_0 = \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)|, \quad m_0 = \sum_{\substack{x\in\mathcal{X} \\ |x|_{\ell^d_\infty}\leq R}} m(x)\,.$$

In this case, we say that F has *superlinear growth* (in the momentum variable). Examples of cost functions with superlinear growth include (1.11) as well as the edge-based costs $F(m,J) = \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)|^p$, both with $1 < p < \infty$.

Under this stronger assumption, we are able to show that curves of measures $\mathbf{m}^\varepsilon$ with uniform bounds on the mass and energy (1.12) convergences (up to a non-relabeled subsequence) $\iota_\varepsilon m_t^\varepsilon \rightharpoonup \mu_t$ uniformly in $(0,T)$. Moreover, $\boldsymbol{\mu} = \mathrm{d}t \otimes \mu_t$ with $t \mapsto \mu_t \in \mathsf{W}^{1,1}((0,T);\mathcal{M}_+(\mathbb{T}^d))$. In particular, $t \mapsto \mu_t$ is continuous with respect to the vague topology of $\mathcal{M}_+(\mathbb{T}^d)$ and there exists a corresponding momentum field $\boldsymbol{\nu} = \mathrm{d}t \otimes \nu_t$ such that $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathsf{CE}$.

Convergence of boundary value problems. Under the superlinear growth assumption for F and the consequent uniform compactness result, we have a nice application of our main result, which is the convergence of boundary value problems.

Consider for $m^0, m^1 \in \mathcal{M}_+(\mathcal{X}_\varepsilon)$ with $m^0(\mathcal{X}_\varepsilon) = m^1(\mathcal{X}_\varepsilon)$ and $\mu^0, \mu^1 \in \mathcal{M}_+(\mathbb{T}^d)$ with $\mu^0(\mathbb{T}^d) = \mu^1(\mathbb{T}^d)$ the minimal actions

$$\mathcal{MA}_\varepsilon(m^0, m^1) := \inf\left\{\mathcal{A}_\varepsilon(m) \,:\, m_0 = m^0, m_1 = m^1\right\}\,,$$
$$\mathbb{MA}(\mu^0,\mu^1) := \inf\left\{\mathbb{A}(\mu) \,:\, \mu_0 = \mu^0, \mu_1 = \mu^1\right\}.$$

Note that the values μ_0 and μ_1 are well-defined under superlinear growth thanks to the fact that finite energy measures admit a disintegration which is continuous in time. Then we can show that the minimal actions $\mathcal{MA}_\varepsilon$ Γ-converge to $\mathbb{MA}$ in the narrow topology of $\mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}_+(\mathbb{T}^d)$.

It is worth noting that even under linear growth, μ_0 and μ_1 can still be defined using the trace theorem in BV, but we cannot prove the convergence result statement in that case. We plan to work on this problem in the future.

Evolutionary Γ-convergence of gradient flow structures

The behavior of the transport costs is only one of the questions arising. In Chapter 4, we focus on the study of gradient flows associated with discrete transport distances and discuss discrete-to-continuum convergence results.

At the continuous level, in their seminal work Jordan, Kinderlehrer and Otto [JKO98] showed that the heat flow in $\mathbb{R}^d$ can be seen as the gradient flow of the relative entropy $\mathsf{Ent}(\rho \, \mathrm{d}x) := \int \rho \log \rho \, \mathrm{d}x$ with respect to the $2-$Wasserstein cost.

There are several interpretations of what a gradient-flow structure means, in this setting. On one side, one can formally introduce an infinite dimensional Riemannian structure on the space of probability measures $\mathcal{P}_2(\mathbb{R}^d)$ whose associated geodesic distance coincides with $\mathbb{W}_2$. If $\mathfrak{g}_2$ is the corresponding metric, then the heat equation formally corresponds to the gradient-flow evolution $\dot{\mu} = -\nabla_{\mathfrak{g}_2} \mathsf{Ent}(\mu)$.

The first mathematically rigorous approach to this problem, proposed in [JKO98], was to consider suitable discretisation in time of the heat equation, via the so-called *minimising movements*, or from the authors' name, the *JKO discretisation scheme*. This represents the infinite dimensional counterpart of the well-known implicit Euler scheme for gradient flows in finite dimension.

An alternative approach is the metric one, which can be applied to general metric spaces, in this case $(\mathcal{P}(\overline{\Omega}), \mathbb{W}_2)$, where $\Omega \subset \mathbb{R}^d$ is a bounded, convex domain of $\mathbb{R}^d$. We refer to the book of Ambrosio, Gigli, and Savaré [AGS08] for an overview on the topic. Let us explain the special case of the $\mathbb{W}_2$ framework. We are going to consider a slightly more general problem, which is the Fokker–Planck equation

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot (\mu_t \nabla V) \quad \text{on } (0,T) \times \Omega \,, \tag{1.13}$$

with Neumann (no-flux) boundary conditions. Here $V \in C^2(\Omega) \cap C(\overline{\Omega})$ is a driving potential, with bounded second derivative. For constant potentials V, we recover the usual heat equation.

The equation (1.13) admits a unique steady state, that is $\mathrm{d}\mathfrak{m}(x) = Z_V \exp(-V(x)) \, \mathrm{d}\mathscr{L}^d(x)$, where Z_V is a renormalisation constant. The equation corresponds to the gradient flow of the relative entropy functional $\mathsf{Ent}_\mathfrak{m}$ given by

$$\mathsf{Ent}_\mathfrak{m} : \mathcal{P}(\overline{\Omega}) \to [0, +\infty] \,, \quad \mathsf{Ent}_\mathfrak{m}(\mu) := \begin{cases} \int_\Omega \rho \log \rho \, \mathrm{d}\mathfrak{m} & \text{if } \mathrm{d}\mu = \rho \, \mathrm{d}\mathfrak{m} \,, \\ +\infty & \text{otherwise} \,. \end{cases} \tag{1.14}$$

The metric formulation of the associated gradient flows, which is equivalent to (1.13), is expressed in terms of the *energy dissipation inequality* (EDI)

$$\mathsf{Ent}_\mathfrak{m}(\mu_T) + \int_0^T \mathbb{A}^*(\mu_t, \dot{\mu}_t) + \mathbb{A}\Big(\mu_t, -\mathsf{DEnt}_\mathfrak{m}(\mu_t)\Big) \, \mathrm{d}t \le \mathsf{Ent}_\mathfrak{m}(\mu_0) \,, \tag{1.15}$$

where $\mathbb{A}(\mu, \varphi) := \frac{1}{2} \int_{\overline{\Omega}} |\nabla \varphi|^2 \, \mathrm{d}\mu$ denotes the (weighted) Sobolev seminorm of $\varphi \in H^1(\Omega)$ and $\mathbb{A}^*(\mu, \cdot)$ its Legendre transform in the second variable, whereas $\mathsf{DEnt}_\mathfrak{m}$ denotes the $L^2(\Omega)$-differential of the relative entropy. The EDI formulation of the Fokker–Planck equation describes the evolution of the relative entropy along the solutions, and quantifies its dissipation via the functionals $\mathbb{A}$, $\mathbb{A}^*$.

One of the strengths of these variational interpretations of the Fokker–Planck equation is that they can be easily translated in the discrete framework. In the very same spirit, the discrete counterpart of the JKO theorem has been proved in [Maa11] and [Mie11], independently, in the setting of discrete heat flow and discrete relative entropy on finite state Markov chains for a suitable choice of the average θ in (1.5), given by the logarithmic mean $\theta_{log}(x,y) = (x-y)\log^{-1}(x/y)$.

In particular, the EDI formulation of the evolution admits a clear discrete counterparts. Consider an admissible (in the sense of [EGH00], see also Section 4.2.2 in Chapter 4) finite volume partition $\mathcal{T}$ of Ω (see Figure 1.1) and the corresponding discrete transport distance $\mathcal{W}_{\mathcal{T}}$, as described in (1.7), with mobility $\theta = \theta_{\log}$.

A natural discretisation of the continuous EDI is as follows. A curve $(m_t)_{t\in[0,1]}$ in the space $(\mathcal{P}(\mathcal{T}), \mathcal{W})$ is said to be a gradient flow of a certain energy $\mathcal{E} : \mathcal{P}(\mathcal{T}) \to [+\infty]$ if it satisfies the discrete *Energy Dissipation Inequality* (EDI$_{\mathcal{T}}$):

$$\mathcal{E}(m_T) + \int_0^T \left[\mathcal{A}^*_{\mathcal{T}}(m_t, \dot{m}_t) + \mathcal{A}_{\mathcal{T}}\Big(m_t, -\mathsf{D}\mathcal{E}(m_t)\Big)\right] \mathrm{d}t \le \mathcal{E}(m_0)\,, \tag{1.16}$$

where $\mathcal{A}_{\mathcal{T}}(m,f) = \frac{1}{4}\sum_{x,y\in\mathcal{T}} \omega_{xy}\widehat{r}(x,y)(f(y)-f(x))^2$ denotes the discrete Sobolev seminorm of a function $f : \mathcal{T} \to \mathbb{R}$ weighted with the measure $m \in \mathcal{P}(\mathcal{T})$ with density r with respect to π, whereas $\mathcal{A}^*_{\mathcal{T}}(m,\cdot)$ denotes its Legendre transform in the second variable.

In Chapter 4, we study the stability of these gradient flows structures in the discrete-to-continuum limit. More in detail, we consider a sequence of meshes $\{\mathcal{T}_N\}_N$ of vanishing diameter $[\mathcal{T}_N] \to 0$ and define the discrete entropy functional $\mathsf{Ent}_N : \mathcal{P}(\mathcal{T}_N) \to \mathbb{R}$ given by

$$\mathsf{Ent}_N(m) := \sum_{x\in\mathcal{T}_N} m(x)\log\left(\frac{m(x)}{\pi(x)}\right), \quad \forall m \in \mathcal{P}(\mathcal{T}_N)\,,$$

for π as defined in (1.7). Define $(m_t^N)_t$ to be the correspondent gradient flow of Ent_N according to (1.16).

We are interested in the limit behavior of the discrete evolutions as $N \to \infty$, and we want to approach this problem passing to the limit directly at the level of the energy dissipation inequality (1.16). The first result in this direction has been obtained by Disser and Liero in [DL15], where the authors showed the convergence of the discrete Fokker–Planck equations on isotropic, one-dimensional, finite-volume meshes passing to the limit directly at the level of the correspondent gradient-flow structure.

Together with my collaborators Dominik Forkert and Jan Maas, in [FMP20] we extended this result to finite volume meshes in arbitrary dimension, even without assuming any additional isotropy on the meshes. In more details, asssume that the initial data are well-posed, namely $Q_N m_0^N \rightharpoonup \mu_0 \in \mathcal{P}(\overline{\Omega})$ and $\mathsf{Ent}_N(m_0^N) \to \mathsf{Ent}_{\mathfrak{m}}(\mu_0)$ as $N \to +\infty$, where we consider the embedding maps

$$\mathsf{Q}_N : \mathcal{P}(\mathcal{T}_N) \to \mathcal{P}(\overline{\Omega}), \quad \mathsf{Q}_N m = \sum_{K\in\mathcal{T}_N} m(K)\mathscr{L}^d \llcorner K \quad \text{for } m \in \mathcal{P}(\mathcal{T}_N).$$

Then the sequence $\mu_t^N := Q_N m_t^N$ is compact in $C([0,T], (\mathcal{P}(\overline{\Omega}), \mathbb{W}_2))$ and any limit point $\mu_t \in \mathcal{P}(\overline{\Omega})$ satisfies

$$\liminf_{N\to\infty} \mathsf{Ent}_N(m_T^N) \ge \mathsf{Ent}_{\mathfrak{m}}(\rho_T), \tag{1.17}$$

$$\liminf_{N\to\infty} \int_0^T \mathcal{A}_{\mathcal{T}_N}\Big(m_t^N, -\mathsf{DEnt}_N(m_t^N)\Big)\,\mathrm{d}t \ge \int_0^T \mathbb{A}\Big(\mu_t, -\mathsf{DEnt}_{\mathfrak{m}}(\mu_t)\Big)\,\mathrm{d}t, \tag{1.18}$$

$$\liminf_{N\to\infty} \int_0^T \mathcal{A}^*_{\mathcal{T}_N}(m_t^N, \dot{m}_t^N)\,\mathrm{d}t \ge \int_0^T \mathbb{A}^*(\mu_t, \dot{\mu}_t)\,\mathrm{d}t. \tag{1.19}$$

In particular μ_t solves the continuous EDI (1.15), or equivalently the Fokker–Planck equation (1.13).

Our proof is in nature very different from the one dimensional case of [DL15]. In arbitrary dimension, one cannot rely on explicit interpolation estimates available on the real line. To prove the result in such generality, we make use of some variational techniques introduced in [BFLM02] and [AC04], and rely on some discrete, uniform Hölder regularity estimates we prove to hold in the finite-volume framework, whose discussion is the content of Appendix A.

Let us briefly describe the strategy of our proof. The first bound (1.17) is a consequence of the weak lower-semicontinuity of relative entropy functionals together with Fatou's lemma. The main difficulty is the proof of the two remaining lower bounds for the dissipation energies. The crucial tool to this purpose is a Mosco convergence result for the discrete Dirichlet energies

$$\mathcal{F}_N : \mathbb{R}^{\mathcal{T}_N} \to \mathbb{R}_+ \,, \quad \mathcal{F}_N(f) := \mathcal{A}_{\mathcal{T}_N}(m_N, f) \,, \quad \forall f \in \mathbb{R}^{\mathcal{T}_N} \,,$$

for a given sequence of discrete measures $m_N \in \mathbb{R}^{\mathcal{T}_N}$. We define the corresponding embedded continuous functionals $\tilde{\mathbb{F}}_{\mathcal{T}} : L^2(\Omega) \to [0, +\infty]$ as

$$\tilde{\mathbb{F}}_N(\varphi) := \begin{cases} \mathcal{F}_N\big(P_{\mathcal{T}}\varphi\big) & \text{if } \varphi \in \mathsf{PC}_N, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{where} \quad \big(P_N\varphi\big)(x_K) := \varphi(x_K) \text{ for } \varphi : \Omega \to \mathbb{R}$$

and PC_N denotes denotes the space of all functions in $L^2(\Omega)$ that are constant a.e. on each cell K in $\mathcal{T}_N$.

Assuming that the $\mu_N := P_{\mathcal{T}_N} m_N \in \mathcal{P}(\Omega)$ have densities (with respect to $\mathscr{L}^d$) which are bounded from above and away from zero, and weakly convergent to some limit measure $\mu \in \mathcal{P}(\Omega)$, we can prove that the energies $\tilde{\mathbb{F}}_N$ Mosco-converge in $L^2(\Omega)$ to a continuous Dirichlet form $\mathbb{F}_\mu$. The proof is based on a compactness and representation procedure, following the ideas of [AC04] and [BFLM02] and extending them to the finite-volume framework.

The final step consist in the identification of the limit $\mathbb{F}_\mu$. Intuitively, one expects to prove that for $\varphi \in L^2(\Omega)$

$$\mathbb{F}_\mu(\varphi) = \mathbb{A}(\mu, \varphi) = \begin{cases} \dfrac{1}{2} \displaystyle\int_\Omega |\nabla\varphi|^2 \,\mathrm{d}\mu & \text{if } \varphi \in H^1(\Omega), \\ +\infty & \text{otherwise}\,. \end{cases} \tag{1.20}$$

We have already seen in the previous sections that, for a general admissible mesh $\mathcal{T}$, we expect discrete energies to be very sensitive to possible oscillations of the densities. Therefore it should not surprise that some additional regularity conditions on μ_N are required in order to obtain the sought convergence result. A possible formulation reads as follows: we say that the *pointwise condition* holds if μ has density $\rho = \frac{\mathrm{d}\mu}{\mathrm{dm}}$ and $\mu_N := \mathsf{Q}_N m_N \rightharpoonup \mu$ with μ_N satisfying for a.e. $x_0 \in \Omega$:

$$\lim_{\varepsilon\to 0} \liminf_{N\to\infty} \sup_{x\in Q_\varepsilon(x_0)} \rho_N(x) \le \rho(x_0) \le \lim_{\varepsilon\to 0} \limsup_{N\to\infty} \inf_{x\in Q_\varepsilon(x_0)} \rho_N(x). \tag{pc}$$

Here, $Q_\varepsilon(x_0)$ denotes the open cube of side-length $\varepsilon > 0$ centered at x_0, and $\rho_N(x) := r_N(x_K)$ for $x \in K$. This clearly prevents the densities μ_N to be strongly oscillating, and with this additional assumption we are indeed able to prove (1.20).

It would be interesting to study the limit behavior of the discrete Dirichlet forms $\tilde{\mathbb{F}}_N$ without the regularity assumption (pc) but assuming more regularity at the level of the meshes, e.g. isotropy or periodicity. This will be the focus of future investigation.

Finally, the flexibility of this approach does not seem limited to the Fokker–Planck setting and opens the door to possible generalisations to non-linear and higher-order equations.

In the last section of this introduction, we make a slight change of perspective, and move our attention away from the setting of discrete transport problems to focus on the non-commutative framework. Interestingly, in a similar fashion as we have seen working with discrete spaces, many classical tools coming from the theory of optimal transport can be translated into the non-commutative world as well.

1.0.1 Non-commutative optimal transport

We have seen in the previous sections how many ideas and techniques classically used in a continuous setting (such as euclidean domains or Riemannian manifolds) can find their counterparts in the setting of discrete Markov chains (or equivalently weighted graphs). Another framework where the classical ideas of optimal transport can find interesting translations and applications is the *non-commutative* one. This is the topic of the last chapter of this thesis.

The easiest examples are *density matrices* over a finite dimensional, complex Hilbert space $\mathfrak{h}$. A density matrix over $\mathfrak{h}$, whose set we denote by $\mathfrak{P}(\mathfrak{h})$, is defined as a hermitian operator Γ with trace one. In the simpler finite dimensional setting, they reduce to the subset of $d \times d$ matrices

$$\mathfrak{P}(\mathbb{C}^d) =: \mathfrak{P}^d = \left\{ \Gamma \in \mathcal{S}^d \ : \ \mathrm{Tr}(\Gamma) = 1 \right\} , \quad \text{for } d \in \mathbb{N} ,$$

where $\mathcal{S}^d \subset \mathcal{M}^d := \mathcal{M}^d(\mathbb{C})$ is the set hermitian $d \times d$ matrices with complex entries.

A density matrix is the non-commutative analogue of a probability measure, where the trace plays the role of the integration. Using the same perspective, one define *couplings* between density matrices. Suppose that $\mathfrak{h} = \mathfrak{h}_1 \otimes \mathfrak{h}_2$, for two given Hilbert spaces $\mathfrak{h}_1$, $\mathfrak{h}_2$. Then one says that $\Gamma \in \mathfrak{P}(\mathfrak{h})$ is a coupling between $\gamma_1 \in \mathfrak{P}(\mathfrak{h}_1)$ and $\gamma_2 \in \mathfrak{P}(\mathfrak{h}_2)$ if

$$\mathrm{Tr}_1(\Gamma) = \gamma_1 \quad \text{and} \quad \mathrm{Tr}_2(\Gamma) = \gamma_2 ,$$

where $\mathrm{Tr}_i(\Gamma)$ denotes the i-th partial trace (5.23) of Γ. We refer to γ_1 and γ_2 as the *marginals* of Γ, as in the commutative setting, and we write $\Gamma \mapsto (\gamma_1, \gamma_2)$. Finally, the cost function is here represented by an hermitian operator H over $\mathfrak{h}$.

The corresponding static, non-commutative optimal transport problem, in strict analogy to the classical Monge–Kantorovich one (1.1), is then given for every $\gamma_1 \in \mathfrak{P}(\mathfrak{h}_1)$, $\gamma_2 \in \mathfrak{P}(\mathfrak{h}_2)$ by

$$\mathfrak{F}(\gamma_1, \gamma_2) := \inf \left\{ \mathrm{Tr}(\mathrm{H}\, \Gamma) \ : \ \Gamma \in \mathfrak{P}(\mathfrak{h}) \, , \ \Gamma \mapsto (\gamma_1, \gamma_2) \right\} .$$

This has been for example considered in [CGP18], where the authors study this problem and a suitable dual formulation, in the same spirit of the Kantorovich dual formulation

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int c(x,y) \, \mathrm{d}\pi(x,y) = \sup_{\phi, \varphi \in \mathrm{Lip}_b(X)} \left\{ \int \varphi \, \mathrm{d}\mu_1 + \int \psi \, \mathrm{d}\mu_0 \ : \ \varphi \oplus \psi \leq c \right\} , \quad (1.21)$$

which holds for bounded, lower semicontinuous cost functions $c : X \times X \to [0, +\infty]$.

The static formulation is certainly not the only attempt to introduce a notion of optimal transportation between density matrices. In their seminal works Carlen and Maas [CM14], [CM17] proposed a non-commutative optimal transport distance between density matrices based on a quantum analogue of the Benamou–Brenier formula (1.3). Nonetheless the connections (if any) between the static and the dynamical formulation remain unclear.

In this thesis, we only consider the static approach to the non-commutative problem. In Chapter 5, we study the corresponding entropic regularised problem, or else known in the commutative setting as *Schrödinger problem* [Sch31], [L14]. For a given regularisation parameter $\varepsilon > 0$, the ε-entropic regularised transport problems on $\mathcal{P}(X)$ is obtained by adding an entropy contribution, in the form

$$\inf\left\{\int c(x,y)\,\mathrm{d}\pi(x,y) + \varepsilon \mathsf{Ent}_{\mathfrak{m}}(\pi) \ : \ \pi \in \Pi(\mu_0,\mu_1)\right\}, \quad \mu_0,\mu_1 \in \mathcal{P}(X),$$

where $\mathfrak{m} \in \mathcal{P}(X \times X)$ is a given reference measure and $\mathsf{Ent}_{\mathfrak{m}}$ denotes the corresponding relative entropy, as in (1.14).

We are interested in the corresponding non-commutative analogue on the space of density matrices. To this purpose, it is natural to consider the (quantum) Shannon entropy of a density matrix $\Gamma \in \mathfrak{P}(\mathfrak{h})$ given by $S(\Gamma) := \mathrm{Tr}(\Gamma \log \Gamma)$, where the logarithm of a hermitian operator is defined using the spectral calculus. In particular, if $\dim \mathfrak{h} = d \in \mathbb{N}$ and $\{\lambda_j\}_{j \le d}$ are the eigenvalues of Γ, then $S(\Gamma) = \sum_{j=1}^d \lambda_j \log \lambda_j$.

Assume now that $\mathfrak{h} = \mathfrak{h}_1 \otimes \cdots \otimes \mathfrak{h}_N$, with $\dim \mathfrak{h}_i = d_i \in \mathbb{N}$, and fix a set of N marginals $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)$, where $\gamma_i \in \mathfrak{h}_i$. For every $\varepsilon > 0$, one then defines the corresponding *multimarginal quantum Schrödinger problem* as

$$\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \inf\left\{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon S(\Gamma) \ : \ \Gamma \in \mathfrak{P}(\mathfrak{h}),\ \Gamma \mapsto \boldsymbol{\gamma}\right\},$$

where $\Gamma \mapsto \boldsymbol{\gamma}$ means that Γ has γ_i as i-th marginal, for every $i = 1, \dots, N$.

From the point of view of quantum physics, this corresponds to the study of the ground state energy of a finite dimensional composite quantum system at positive temperature $\varepsilon > 0$, *conditionally* to the knowledge of the states of all its subsystems (here represented by $\{\gamma_i\}_i$). Every $\Gamma \mapsto \boldsymbol{\gamma}$ represents an admissible state of the system, whereas the ground state energy corresponds to the energy level of the minimiser Γ^ε in the definition of $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$.

A special case of this is what in the physics literature is known as *one-body reduced density matrix functional theory* (in short 1RDMFT). This corresponds to indistinguishable particles $\mathfrak{h}_i = \mathfrak{h}_0$, $\gamma_i = \gamma$ for all $i = 1, \dots, N$, with additional symmetry constraints enforced on the problem, which can be either *bosonic* or *fermionic*.

Our contribution, which is the result of a joint collaboration with D. Feliciangeli and A. Gerolin [FGP21], is twofold: the first result we show is a non-commutative equivalent dual formulation of the Schrödinger problem, given by

$$\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \mathfrak{D}^\varepsilon(\boldsymbol{\gamma}) = \sup\left\{\sum_{i=1}^N \mathrm{Tr}(U_i\gamma_i) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{\bigoplus_{i=1}^N U_i - \mathrm{H}}{\varepsilon}\right]\right) \ : \ U_i \in \mathcal{S}^{d_i}\right\} + \varepsilon,$$

in the case of marginals with trivial kernel (a very similar formula holds in general as well, see Remark 5.3.9). This is the non-commutative counterpart of the dual formula for the classical Schrödinger problem in the commutative setting [DMG20a]. We call the optimal operators

U_i Kantorovich potentials, which classically refers to φ, ψ in the dual formulation of optimal transport (1.21).

It is interesting to note that the well-known Pauli principle (see e.g. [LS10, Theorem 3.2]), which provides necessary and sufficient conditions for γ to be the one-reduced density matrix of an N-body fermionic density matrix, finds a variational interpretation in our discussion. Indeed, in the antisymmetric case we are able to show (see Proposition 5.2.8) that γ satisfies the Pauli principle (resp. satisfies the Pauli principle *strictly*) if and only if the supremum of the dual functional $\mathfrak{D}^\varepsilon(\gamma)$ is finite (resp. is attained), as it is to be expected.

We also prove existence of maximisers U_i^ε for $\mathfrak{D}^\varepsilon(\gamma)$ and take advantage of the duality theorem to show that the optimisers for the primal problem $\mathfrak{F}^\varepsilon(\gamma)$ can be written in the form

$$\Gamma^\varepsilon = \exp\left(\frac{\bigoplus_{i=1}^N U_i^\varepsilon - \mathrm{H}}{\varepsilon}\right). \tag{1.22}$$

One of the strengths of the dual formulation of the Schrödinger problem is the possibility of computing the optimisers in a very efficient way. In the commutative setting this is the well-known *Sinkhorn algorithm* [CP19]. The second contribution of our work is to introduce a non-commutative, multimarginal analogue of this algorithm, and prove the convergence to the corresponding optimal density matrix Γ^ε in the definition of $\mathfrak{F}^\varepsilon(\gamma)$.

This algorithm exploits the shape of the minimizer obtained in (5.2), in order to construct a sequence $\Gamma^{(k)}$ of density matrices converging to Γ^ε of the form

$$\Gamma^{(k)} = \exp\left(\frac{\bigoplus_{i=1}^N U_i^{(k)} - \mathsf{H}}{\varepsilon}\right), \tag{1.23}$$

where the vector $(U_1^{(k)}, \dots, U_N^{(k)})$ is iteratively updated by progressively imposing that $\Gamma^{(k)}$ has at least one correct partial trace. We prove the convergence and the robustness of this algorithm in Section 5.5.

This represents an attempt to extend typical commutative techniques to the quantum case, with the hope of giving some insights to better understand the complex mathematics behind the infinite dimensional picture (for example appearing in quantum mechanics). Our result are for the moment only limited to the finite dimensional framework. Extension to infinite dimension certainly requires extra work and additional technical difficulties, and we postpone the discussion to future collaborations.

1.0.2 Additional works

In this last section, I present a quick overview of some works I have been doing during my PhD at IST Austria, and which will not be inserted in this thesis.

Optimal control

The variational principles which are behind the energy dissipation inequalities and the gradient flows have a huge impact in dealing with various situations, including optimal control problems. In a joint work [PS19] in collaboration with Ulisse Stefanelli, we consider several abstract optimal control problems of the form

$$\min\{F(u,y) \ : \ y \in S(u)\}, \tag{1.24}$$

where $u : [0, T] \to H$ is a time-dependent admissible control, H is a Hilbert space and $y : [0, T] \to H$ belongs to the set $S(u)$ of solutions to a nonlinear evolution equation suitably depending on u. F is a nonnegative *target* functional defined on the trajectories u and y.

Examples include gradient flows, generalised gradient flows, monotone and pseudomonotone flows, and GENERIC flows. The key property of these non linear equations is the possibility of describing them as zeros of a certain functional G, namely $y \in S(u) \Leftrightarrow G(u, y) = 0$. For example, in case of gradient flows $y + \partial\phi(y) \ni u$ for $\phi : H \to (-\infty, +\infty]$ convex energy, one can write the correspondent equation as (u, y) being a zero the *Brezis–Ekeland–Nayroles* functional

$$G_{\mathrm{BEN}}(u, y) = \int_0^T \Big(\phi(y) + \phi^*(u - y') - (u, y)\Big)\,\mathrm{d}t + \frac{1}{2}\|y(T)\|^2 - \frac{1}{2}\|y_0\|^2.$$

In our work [PS19], we propose an approximation result for problems of the form (1.24) with free (usually convex) minimisation ones. Precisely, the differential constraint is penalized by augmenting the target functional by a nonnegative global-in-time functional G which is null-minimized if and only if the evolution equation is satisfied. We present different possible applications of this idea for various variational settings, showing the convergence of the method and some numerical examples.

Generalised gradient flows

In a project in collaboration with Marco di Francesco, Simone di Marino, and Emanuela Radici we focus on conservation laws and gradient flows with respect to generalised Wasserstein distances in $\mathbb{R}^d$.

Introduced in [DNS12], [LM10], they are obtained for any $\rho_0, \rho_1 \in \mathscr{P}(\Omega)$ as

$$\mathbb{W}_{2,m}(\rho_0, \rho_1)^2 := \inf\left\{\int_0^1\int_{\mathbb{R}^d} |v_t|^2 m(\rho_t)\,\mathrm{d}x\,\mathrm{d}t \;:\; \partial_t\rho_t + \nabla\cdot(m(\rho_t)v_t) = 0,\; \rho_{t=i} = \rho_i\right\},$$

where $m : [0, +\infty) \to [0, +\infty)$ is a concave mobility function. The corresponding (formal) gradient flows equation of an energy $\mathcal{E} : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$ with respect to this generalised distance is given by

$$\partial_t\rho_t - \nabla\cdot(m(\rho_t)\nabla(\mathrm{D}\mathcal{E}(\rho_t)) = 0\,.$$

From a modeling point of view, the previous equation can describe an evolution that aims at minimising the free energy $\mathcal{E}$ while including at the same time possible additional effects and/or constraints, such as overcrowding preventions. The simplest example is the case $m(\rho) = \rho(M - \rho)_+$, for $M > 0$, which vanishes whenever $\rho > M$. In particular, no motion happens when the particle density reaches the upper level M, which represents a threshold for the amount of particles allowed in the model. This is a typical assumption for example in crowd motions and traffic flows models.

In [DFDMPR21], we propose, in the scalar case, a *space discretisation* in the framework of non-linear mobilities, adopting a *Lagrangian* point of view. Firstly, we show an approximation result for the generalised transport distances at the continuous level using systems of N-ordered particles. Subsequently, we take advantage of this discrete-in-space approximation and show the stability at the level of the corresponding gradient-flow structures, providing a finite-dimensional approximation of the associated evolution equation.

CHAPTER 2

Homogenisation of one-dimensional discrete optimal transport

In this chapter we present the one-dimensional homogenisation result for Wasserstein-like distances in a periodic setting. This is the content of the article [GKMP21], obtained in collaboration with Peter Gladbach, Eva Kopfer, and Jan Maas.

This work deals with dynamical optimal transport metrics defined by spatial discretisation of the Benamou–Benamou formula for the Kantorovich metric $\mathbb{W}_2$. Such metrics appear naturally in discretisations of $\mathbb{W}_2$-gradient flow formulations for dissipative PDE. However, it has recently been shown that these metrics do not in general converge to $\mathbb{W}_2$, unless strong geometric constraints are imposed on the discrete mesh. In this paper we prove that, in a 1-dimensional periodic setting, discrete transport metrics converge to a limiting transport metric with a non-trivial effective mobility. This mobility depends sensitively on the geometry of the mesh and on the non-local mobility at the discrete level. Our result quantifies to what extent discrete transport can make use of microstructure in the mesh to reduce the cost of transport.

2.1 Introduction

In the past decades there has been intense research activity in the area of optimal transport, cf. the monographs [Vil03, Vil08, San15, PC19] for an overview of the subject. In continuous settings, a key result in the field is the *Benamou–Brenier formula* [BB00], which expresses the equivalence of static and dynamical formulations of the optimal transport problem. In discrete settings, the equivalence between static and dynamical optimal transport breaks down, and it turns out that the dynamical formulation (introduced in [Maa11, Mie11]) is essential in applications to evolution equations, discrete Ricci curvature, and functional inequalities, see, e.g., [CHLZ12, EM12, Mie13, EM14, EMT15, FM16, EHMT17, EF18].

However, the limit passage from discrete dynamical transport to continuous optimal transport turns out to be nontrivial. In fact, it has been shown in [GKM20] that seemingly natural discretisations of the Benamou–Brenier formula do *not* necessarily converge to the Kantorovich distance $\mathbb{W}_2$, even in one-dimensional settings. The main result in [GKM20] asserts that, for a sequence of meshes on a bounded convex domain in $\mathbb{R}^d$, an isotropy condition on the meshes is required to obtain the convergence of the discrete dynamical transport distances to $\mathbb{W}_2$.

It remained an open question to identify the limiting behaviour of the discrete metrics in situations where the isotropy condition fails to hold. The aim of the current paper is to answer this question in the one-dimensional periodic setting.

We start by informally introducing the main objects of study in this paper and present the main result. For more formal definitions we refer to Section 2.2 below.

Continuous optimal transport

Let $\mathcal{P}(\mathcal{S})$ (resp. $\mathscr{M}(\mathcal{S})$) denote the set of Borel probability measures (resp. signed measures) on a Polish space $(\mathcal{S}, \mathsf{d})$. We will work on the one-dimensional torus $\mathcal{S}^1 = \mathbb{R}/\mathbb{Z}$ and use the convention that arithmetic operations are understood *modulo* 1.

The *Kantorovich metric* $\mathbb{W}_2$ (also known as *Wasserstein metric*) on $\mathcal{P}(\mathcal{S})$ is defined by

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf_{\gamma \in \Gamma(\mu_0,\mu_1)} \left\{ \int_{\mathcal{S}\times\mathcal{S}} \mathsf{d}^2(x,y) \,\mathrm{d}\gamma(x,y) \right\} \tag{2.1}$$

for $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S})$. Here, $\Gamma(\mu_0,\mu_1)$ denotes the set of probability measures on $\mathcal{S}\times\mathcal{S}$ with marginals μ_0 and μ_1 respectively. For $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$ the *Benamou–Brenier formula* yields the equivalent dynamical formulation

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf_{\mu, v} \left\{ \int_0^1 \int_{\mathcal{S}^1} \frac{|j|^2}{\mu} \ : \ \partial_t \mu + \partial_x j = 0 \right\} , \tag{2.2}$$

where the infimum runs over all curves $\mu : [0,1] \to \mathcal{P}(\mathcal{S}^1)$ connecting μ_0 and μ_1, and all vector fields $j : [0,1]\times\mathcal{S}^1 \to \mathbb{R}$ satisfying the stated continuity equation. Here, $\int_0^1 \int_{\mathcal{S}^1} \frac{|j|^2}{\mu}$ is to be understood as $\int_0^1 \int_{\mathcal{S}^1} |v_t(x)|^2 \,\mathrm{d}\mu_t(x)\,\mathrm{d}t$ if $j \ll v$ with $\frac{\mathrm{d}j}{\mathrm{d}\mu} = v$, and $+\infty$ otherwise.

Discrete dynamical optimal transport

Let $\mathcal{X}$ be a finite set endowed with a reference probability measure $\pi \in \mathcal{P}(\mathcal{X})$. Let $R : \mathcal{X}\times\mathcal{X} \to \mathbb{R}_+$ denote the transition rates of an irreducible continuous time Markov chain on $\mathcal{X}$. We assume that the *detailed balance condition* holds, i.e., $\pi(x)R(x,y) = \pi(y)R(y,x)$ for all $x, y \in \mathcal{X}$.

Let $\{\theta_{xy}\}_{x,y\in\mathcal{X}}$ be a collection of admissible means, i.e., each $\theta_{xy} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is concave, 1-homogeneous, and satisfies $\theta(1,1) = 1$. We assume that $\theta_{xy}(a,b) = \theta_{yx}(b,a)$ for any $a, b \geq 0$.

The *discrete dynamical transport metric* associated to $(\mathcal{X}, R, \pi)$ is defined by

$$\mathcal{W}^2(m_0, m_1) = \inf_{m,J} \left\{ \frac{1}{2} \int_0^1 \sum_{x,y\in\mathcal{X}} \frac{J_t^2(x,y)}{\theta_{xy}\Big(m_t(x)R(x,y), m_t(y)R(y,x)\Big)} \,\mathrm{d}t \right\} .$$

Here the infimum runs over all curves $m : [0,1] \to \mathcal{P}(\mathcal{X})$ connecting m_0 and m_1, and all discrete vector fields $J : [0,1] \to \mathcal{V}(\mathcal{X})$ satisfying the discrete continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(x) + \sum_{y\in\mathcal{X}} J_t(x,y) = 0 \qquad \text{for all } x \in \mathcal{X} ,$$

where $\mathcal{V}(\mathcal{X})$ denotes the set of all anti-symmetric functions $V : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$. The definition of $\mathcal{W}$ is a direct analogue of (2.2) with one additional feature: between any pair of points x and y an admissible mean θ_{xy} needs to be chosen to describe the mobility.

Discrete optimal transport on 1-dimensional meshes

In this paper we consider discrete transport metrics induced by a finite volume discretisation of $\mathcal{S}^1$.

Fix $0 = r_0 < \ldots < r_1 < \ldots < r_K = 1$ for some $K \geq 1$. We write $\pi_k := r_{k+1} - r_k$ and $A_k := [r_k, r_{k+1})$, so that $\mathcal{T} := \{A_k\}_{k=0}^{K-1}$ is a partition of $\mathcal{S}^1$ into disjoint half-open intervals. We also consider a sequence of points $\{z_k\}_{k=0}^{K-1}$ such that each z_k lies in the interior of A_k. The distance between z_k and $z_{k'}$ in $\mathcal{S}^1$ will be denoted by $d_{kk'}$. Here and below we will often perform calculations *modulo* K.

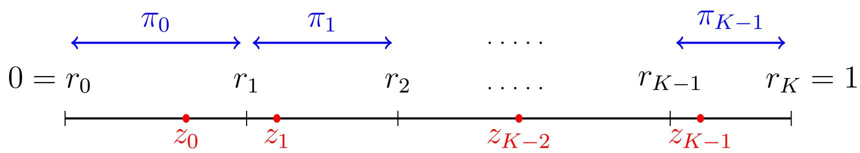

Figure 2.1: The mesh $\mathcal{T}$ on $\mathcal{S}^1$.

We endow the discrete state space $\mathcal{T}$ with the natural reference measure $\pi \in \mathcal{P}(\mathcal{T})$ given by $\pi(A_k) = \pi_k$. The main object of study in this paper is the transport metric $\mathcal{W}_\mathcal{T}$ on $\mathcal{P}(\mathcal{T})$ induced by the Markov transition rates on $\mathcal{T}$ given by

$$R(A_k, A_{k'}) := R_{kk'} := \frac{1}{\pi_k d_{kk'}}$$

if $|k - k'| = 1$, and $R_{kk'} = 0$ otherwise. Then we have the detailed balance condition $\pi_k R_{kk'} = \pi_{k'} R_{k'k}$. The rates are chosen to ensure that solutions to the discrete diffusion equation (i.e., the Kolmogorov forward equation associated to the Markov chain given by R) converge to solutions of the diffusion equation $\partial_t \mu = \partial_x^2 \mu$ in the limit of vanishing mesh size [EGH00]. A gradient flow approach in one dimension can be found in [DL15].

The periodic setting

For any mesh $\mathcal{T}$ as above and $N \geq 1$ one can construct an inhomogeneous periodic mesh $\mathcal{T}_N$ with NK cells $A_{n;k}$ by concatenating N rescaled copies of $\mathcal{T}$.

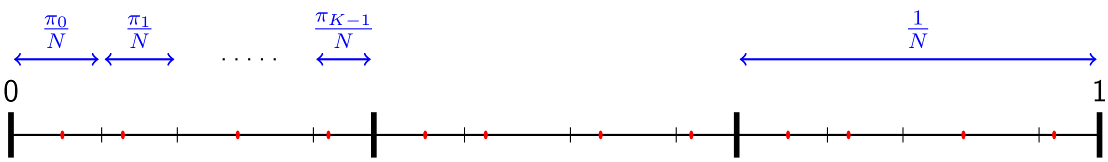

Figure 2.2: The mesh $\mathcal{T}_N$ on $\mathcal{S}^1$.

We then consider the transport metric $\mathcal{W}_N := \mathcal{W}_{\mathcal{T}_N}$ on $\mathcal{P}(\mathcal{T}_N)$ as defined above. Explicitly, we have

$$\mathcal{W}_N^2(m_0, m_1) = \inf_{m,J} \left\{ \frac{1}{N} \int_0^1 \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(n; k, k+1)}{\theta_{k,k+1}\left(\frac{N m_t(n;k)}{\pi_k}, \frac{N m_t(n;k+1)}{\pi_{k+1}} \right)} \, \mathrm{d}t \right\} ,$$

where the infimum runs over all curves $m : [0,1] \to \mathcal{P}(\mathcal{T}_N)$ and $J : [0,1] \to \mathcal{V}(\mathcal{T}_N)$ satisfying the discrete continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(n;k) + J_t(n;k,k+1) - J_t(n;k-1,k) = 0$$

for all $n = 0, \ldots, N-1$ and $k = 0, \ldots, K-1$. Here we use the shorthand notation $m(n;k) = m(A_{n;k})$ and $J(n;k,k+1) = J(A_{n;k}, A_{n;k+1})$. Moreover, we use the convention that $m(n;K) = m(n+1;0)$ and $J(n;K-1,K) = J(n+1;-1,0)$. The main goal of this paper is to analyse the limiting behaviour of $\mathcal{W}_N$ as $N \to \infty$.

The discrete-to-continuous limit

The first convergence result for discrete dynamical transport metrics (in the sense of Gromov–Hausdorff) was obtained in [GM13]. There it is shown that the discrete transport metric associated to the cubic mesh on the d-dimensional torus converges to $\mathbb{W}_2$ in the limit of vanishing mesh size.

The limiting behaviour of discrete dynamical transport metrics on more general meshes turns out to be a delicate issue. In fact, it follows from the multi-dimensional results in [GKM20] that the discrete transport metrics $\mathcal{W}_N$ converge to $\mathbb{W}_2$ if and only if the means $\theta_{kk'}$ are carefully chosen to satisfy an appropriate "balance condition" that reflects the geometry of the mesh $\mathcal{T}$. In our one-dimensional periodic setting, these results imply that $\mathcal{W}_N$ converges to $\mathbb{W}_2$ if and only if there exist constants $\lambda_{k,k+1}, s \in (0,1)$ such that the following conditions hold for $k = 0, \ldots, K-1$:

$$\begin{aligned} r_{k+1} &= \lambda_{k,k+1} z_{k+1} + (1-\lambda_{k,k+1}) z_k + s \ , \\ \theta_{k,k+1}(a,b) &\leq \lambda_{k,k+1} a \qquad + (1-\lambda_{k,k+1}) b \qquad \text{for any } a, b \geq 0 \ . \end{aligned} \tag{2.3}$$

Thus, to fulfill this condition, the asymmetry of the means $\theta_{k,k+1}$ should reflect the relative location of the points z_k, r_{k+1}, and z_{k+1}. We refer to Section 2.4 below for a full discussion.

The main contribution of the current paper is the identification of the limiting behaviour of $\mathcal{W}_N$ in the general one-dimensional periodic setting, without assuming (2.3). To state the result, we introduce the canonical projection operator $P_{\mathcal{T}} : \mathcal{P}(\mathcal{S}^1) \to \mathcal{P}(\mathcal{T})$ defined by

$$(P_{\mathcal{T}}\mu)(\{A\}) = \mu(A) \tag{2.4}$$

for $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $A \in \mathcal{T}$. For brevity we write $P_N := P_{\mathcal{T}_N}$.

The following homogenisation result asserts that $\mathcal{W}_N$ converges to a Kantorovich metric with an effective mobility determined by the geometry of the mesh and by the choice of the means $\theta_{k,k+1}$.

Theorem 2.1.1 (Main result)**.** *Fix a mesh $\mathcal{T}$ on $\mathcal{S}^1$, and consider the induced periodic meshes $\mathcal{T}_N$ for $N \geq 1$. For any $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have*

$$\lim_{N\to\infty} \mathcal{W}_N(P_N\mu_0, P_N\mu_1) = \sqrt{c^\star(\theta, \mathcal{T})}\, \mathbb{W}_2(\mu_0, \mu_1) \ ,$$

where

$$c^\star(\theta, \mathcal{T}) := \inf \left\{ \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\right)} \ : \ m \in \mathcal{P}(\mathcal{T}) \right\} . \tag{2.5}$$

Moreover, as $N \to \infty$ we have Gromov–Hausdorff convergence of metric spaces:

$$(\mathcal{P}(\mathcal{T}_N), \mathcal{W}_N) \to \left(\mathcal{P}(\mathcal{S}^1), \sqrt{c^\star(\theta, \mathcal{T})}\mathbb{W}_2\right) ,$$

in the sense of Definition 2.7.1.

Remark 2.1.2 (Upper bound and isotropic case)*.* We show in Section 2.4 that $c^*(\theta, \mathcal{T}) \le 1$. Moreover, if the compatibility conditions (2.3) are satisfied, it follows that $c^\star(\theta, \mathcal{T}) = 1$, and we recover the result of [GKM20].

Remark 2.1.3 (Convergence of gradient flows)*.* We stress that the limiting behaviour at the level of the transport metrics is in stark contrast with the convergence results of the level of the gradient flow equation. Indeed, consider the discrete transport metric $\mathcal{W}_N$ in the case where each $\theta_{k,k+1}$ is equal to the *logarithmic mean* $\theta_{\log}(a,b) = \int_0^1 a^{1-s}b^s \,\mathrm{d}s$. Then the discrete diffusion equation is the gradient flow equation in $(\mathcal{P}(\mathcal{T}_N), \mathcal{W}_N)$ for the relative entropy with respect to the natural reference measure π_N; cf. [CHLZ12, Maa11, Mie11]. Similarly, the continuous diffusion equation is the gradient flow in $(\mathcal{P}(\mathcal{S}^1), \mathbb{W}_2)$ for the relative entropy with respect to the Lebesgue measure on $\mathcal{S}^1$ [JKO98]. Convergence of solutions of the discrete heat equation to solutions of the continuous heat equation is well known, see, e.g., [EGH00]. Nevertheless, our main result shows that the discrete transport metrics $\mathcal{W}_N$ converge to a limiting metric that is different from $\mathbb{W}_2$, unless the mesh is equidistant. For a systematic study of convergence of gradient flow structures we refer to [Mie16b, Mie16a, DFM19]; see also [ARM17] for a discussion in the context of finite volume discretisations.

Remark 2.1.4 (Convergence on geometric graphs)*.* A convergence result for discrete transport distances on a large class of geometric graphs associated to point clouds on the d-dimensional torus has been obtained in [GT20]. This result applies in particular to iid points sampled from the uniform distribution on the torus. As the results in that paper apply to sequences of graphs with increasing degree, they do not overlap with the results obtained here.

Heuristics

We briefly sketch a non-rigorous argument that makes Theorem 2.1.1 plausible. For this purpose we consider a smooth solution to the continuity equation $\partial_t \mu + \partial_x j = 0$, and fix $\alpha \in \mathcal{P}(\mathcal{T})$. Suppressing the time variable, we define a discrete measure m that assigns mass $m(k) := \alpha(k)\mu\big([\frac{n}{N}, \frac{n+1}{N})\big)$ to each cell $A_{n;k}$ in $\mathcal{T}_N$. This ensures that each interval of the form $[\frac{n}{N}, \frac{n+1}{N})$ receives the same mass at the discrete and the continuous level, but within each such interval, the measure α introduces discrete density oscillations.

Let J be the discrete momentum vector field that solves the continuity equation for m. If this vector field is sufficiently regular, we may estimate the discrete energy by

$$\begin{aligned}
&\sum_{n=0}^{N-1}\sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(n;k,k+1)}{\theta_{k,k+1}\Big(\frac{Nm_t(n;k)}{\pi_k}, \frac{Nm_t(n;k+1)}{\pi_{k+1}}\Big)} \\
&\qquad \approx \frac{1}{N}\sum_{n=0}^{N-1} \frac{J_t^2(n;0,1)}{\mu\big([\frac{n}{N}, \frac{n+1}{N})\big)} \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\Big(\frac{\alpha(k)}{\pi_k}, \frac{\alpha(k+1)}{\pi_{k+1}}\Big)} \approx c^\star(\theta, \mathcal{T}) \int \frac{|j|^2}{\mu} ,
\end{aligned}$$

after minimisation over $\alpha \in \mathcal{P}(\mathcal{T})$. We thus recover the continuous energy appearing in the Benamou–Brenier formula up to a multiplicative correction, which indeed suggests our main result.

A rigorous argument based on this heuristics clearly requires suitable spatial regularity results for m and J. Indeed, we will show in Section 2.5 below that any discrete curve can be approximated by a curve of similar energy, which enjoys good Lipschitz bounds for J as well as good Lipschitz bounds for m up to oscillations within each cell.

Organisation of the paper

In Section 2.2 we collect the basic definitions and preliminary results that are used in this paper. In Section 2.3 we give a simple approach to some of the main convergence results, which only applies in the special case where $\mathcal{T}$ consists of exactly 2 cells. In Section 2.4 we analyse the formula (2.5) for the effective mobility $c^\star(\theta, \mathcal{T})$ and discuss its relation to the geometric conditions from [GKM20].

The bulk of the proof of the main result is contained in Sections 2.5 and 2.6, which deal with the lower and upper bounds for $\mathcal{W}_N$ respectively. The key results in these sections are Theorems 2.5.4 and 2.6.6. In Section 2.7 we finish the proof of the main result by proving the Gromov–Hausdorff convergence.

2.2 Preliminaries

2.2.1 Continuous optimal transport on $\mathcal{S}^1$

For $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, let $\mathbb{CE}(\mu_0, \mu_1)$ denote the set of all distributional solutions to the continuity equation

$$\partial_t \mu + \partial_x j = 0 \tag{2.6}$$

with boundary conditions $\mu_t|_{t=0} = \mu_0$ and $\mu_t|_{t=1} = \mu_1$. More precisely, this means that $(\mu_t)_t$ is a weakly continuous family of measures in $\mathcal{P}(\mathcal{S}^1)$ with the given boundary conditions, $(j_t)_t$ is a Borel family of measures in $\mathscr{M}(\mathcal{S}^1)$ satisfying $\int_0^1 |j_t|(\mathcal{S}^1)\,\mathrm{d}t < \infty$, and (2.6) holds in the sense that

$$\int_0^1 \int_{\mathcal{S}^1} \partial_t \xi(t,x)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t + \int_0^1 \int_{\mathcal{S}^1} \partial_x \xi(t,x)\,\mathrm{d}j_t(x)\,\mathrm{d}t = 0$$

for any test function $\xi \in C_c^1(\mathcal{S}^1 \times (0,1))$. For $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $j \in \mathscr{M}(\mathcal{S}^1)$ we set

$$\mathbb{A}(\mu, j) = \int_{\mathcal{S}^1} \left| \frac{\mathrm{d}j}{\mathrm{d}\mu} \right|^2 \mathrm{d}\mu \ ,$$

if $j \ll \mu$, and $\mathbb{A}(\mu, j) = +\infty$ otherwise. With this notation, the Benamou–Brenier formula [BB00] asserts that

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \mathbb{A}(\mu_t, j_t)\,\mathrm{d}t \ : \ (\mu_t, j_t)_t \in \mathbb{CE}(\mu_0, \mu_1) \right\} , \tag{2.7}$$

see, e.g., [AGS08, Lemma 8.1.3] for more details.

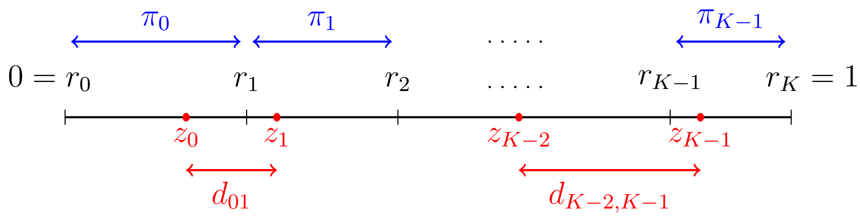

Figure 2.3: The mesh $\mathcal{T}$ on $\mathcal{S}^1$.

2.2.2 Discrete optimal transport on one-dimensional meshes

As in Section 2.1, we fix a mesh $\mathcal{T} = \{A_k\}_{k=0}^{K-1}$ on $\mathcal{S}^1$, and use the notation r_k, π_k, z_k, $d_{kk'}$. The set $\mathcal{V}(\mathcal{T})$ of discrete vector fields is naturally identified with the set of real-valued functions on $\{(k, k+1)\}_{k=0}^{K-1}$.

Definition 2.2.1 (Discrete continuity equation)**.** A pair $(m_t, J_t)_{t\in[0,1]}$ is said to satisfy the *discrete continuity equation* if

(i) $m : [0, 1] \to \mathcal{P}(\mathcal{T})$ is continuous;

(ii) $J : [0, 1] \to \mathcal{V}(\mathcal{T})$ is locally integrable;

(iii) the continuity equation holds in the sense of distributions:

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(k) + J_t(k, k+1) - J_t(k-1, k) = 0 \qquad \text{for all } k = 0, \dots, K-1\,. \tag{2.8}$$

We write $\mathsf{CE}_{\mathcal{T}}(m_0, m_1)$ to denote the collection of pairs $(m_t, J_t)_{t\in[0,1]}$ satisfying $m|_{t=0} = m_0$ and $m|_{t=1} = m_1$.

Definition 2.2.2 (Admissible mean)**.** An *admissible mean* is a function $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ that is concave, 1-homogeneous, and satisfies $\theta(1, 1) = 1$.

Note that we do *not* impose that admissible means are symmetric. Let us briefly recall some properties of admissible means that will be used in the sequel.

Lemma 2.2.3 (Properties of admissible means)**.** *For any admissible mean $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ the following statements hold:*

(i) *For $a, b \geq 0$ we have $\min\{a, b\} \leq \theta(a, b) \leq \max\{a, b\}$.*

(ii) *The map $\mathbb{R}_+ \times \mathbb{R}_+ \ni (a, b) \mapsto \frac{1}{\theta(a,b)} \in (0, +\infty]$ is jointly lower semicontinuous.*

(iii) *θ is locally Lipschitz on $(0, +\infty)^2$.*

Proof. For all $a, b, s, t \geq 0$ we obtain, using 1-homogeneity and concavity,

$$\theta(a + s, b + t) = 2\theta(\tfrac{a+s}{2}, \tfrac{b+t}{2}) \geq \theta(a, b) + \theta(s, t) \geq \theta(a, b),$$

hence θ is non-decreasing with respect to the first and the second variable. Thus, if $a \leq b$, it follows that $a = \theta(a, a) \leq \theta(a, b) \leq \theta(b, b) = b$. Since the same argument applies if $a \geq b$, we obtain (i).

The claims in (ii) and (iii) are easy consequences of the assumptions on θ. □

Definition 2.2.4 (Discrete dynamical transport distance)**.** Let $\mathcal{T}$ be a mesh on $\mathcal{S}^1$, and $\{\theta_{k,k+1}\}_{k=0}^{K-1}$ be a family of admissible means.

1. The energy functional $\mathcal{A}_\mathcal{T} : \mathcal{P}(\mathcal{T}) \times \mathcal{V}(\mathcal{T}) \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\mathcal{A}_\mathcal{T}(m, J) = \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1}\left(\frac{m(k)}{\pi_k}, \frac{m(k+1)}{\pi_{k+1}}, J(k, k+1)\right) ,$$

where $f_{k,k+1}(\rho, \tilde{\rho}, J) = F\Big(\theta_{k,k+1}(\rho, \tilde{\rho}), J\Big)$, and

$$F(\rho, J) = \begin{cases} \frac{J^2}{\rho} & \text{if } \rho > 0 \ , \\ 0 & \text{if } \rho = 0 \text{ and } J = 0 \ , \\ +\infty & \text{otherwise} \ . \end{cases}$$

2. The discrete dynamical transportation distance between $m_0, m_1 \in \mathcal{P}(\mathcal{T})$ is given by

$$\mathcal{W}_\mathcal{T}(m_0, m_1) = \inf\left\{\sqrt{\int_0^1 \mathcal{A}_\mathcal{T}(m_t, J_t)\,\mathrm{d}t} \ : \ (m_t, J_t)_t \in \mathsf{CE}_\mathcal{T}(m_0, m_1)\right\} .$$

The infimum in the the previous definition is attained; cf. [EM12, Theorem 3.2]. In the sequel we apply these definitions to the periodic meshes $\mathcal{T}_N$ defined in Section 2.1. We will then simply write $\mathcal{A}_N$ and $\mathcal{W}_N$ as a shorthand for $\mathcal{A}_{\mathcal{T}_N}$ and $\mathcal{W}_{\mathcal{T}_N}$ respectively.

2.2.3 A priori bounds

In this section we collect some coarse bounds that will be useful in the sequel. To compare discrete and continuous measures, we consider the canonical embedding $\iota_\mathcal{T} : \mathcal{P}(\mathcal{T}) \to \mathcal{P}(\mathcal{S}^1)$ defined by

$$\iota_\mathcal{T} m = \sum_{k=0}^{K-1} m_k \mathcal{U}_{A_k} \qquad \text{for } m \in \mathcal{P}(\mathcal{T}) \ ,$$

where $\mathcal{U}_{A_k}$ denotes the uniform probability measure on A_k. Note that $\iota_\mathcal{T}$ is a right-inverse of the projection map $P_\mathcal{T}$ defined by (2.4). We will often write $\iota_N = \iota_{\mathcal{T}_N}$ for brevity.

The following notion of mesh regularity can be found in a multi-dimensional setting in [EGH00, Section 3.1.2].

Definition 2.2.5 (ζ-regularity)**.** Let $\zeta \in (0, 1]$. We say that a mesh $\mathcal{T}$ is ζ-regular, if $\zeta < \frac{\min_k\{z_k - r_k, r_{k+1} - z_k\}}{\max_k \pi_k}$ for all $k = 0, \dots, K-1$.

A mesh $\mathcal{T}$ is ζ-regular if and only if the ball of radius $\zeta \max_k \pi_k$ around z_k is contained in the interior of the cell A_k for each k. Clearly, any mesh $\mathcal{T}$ on $\mathcal{S}^1$ is ζ-regular for some $\zeta \in (0, 1]$.

Remark 2.2.6. If $\mathcal{T}$ is ζ-regular for some $\zeta \in (0, 1]$, then each $\mathcal{T}_N$ is ζ-regular as well.

Let $[\mathcal{T}]$ denote the size of the mesh, i.e., the maximal diameter of its cells:

$$[\mathcal{T}] := \max\{\pi_k : k = 0, \dots, K-1\} \ .$$

The following result provide a coarse upper bound for $\mathcal{W}_N$ in terms of $\mathbb{W}_2$.

Proposition 2.2.7 (Coarse upper bound for $\mathcal{W}_{\mathcal{T}}$)**.** *Let $\zeta \in (0,1]$. There exists a constant $C < \infty$ depending only on ζ such that for any ζ-regular mesh $\mathcal{T}$ of $\mathcal{S}^1$ and all $m_0, m_1 \in \mathcal{P}(\mathcal{T})$ we have*

$$\mathcal{W}_{\mathcal{T}}(m_0, m_1) \leq C\Big(\mathbb{W}_2(\iota_{\mathcal{T}} m_0, \iota_{\mathcal{T}} m_1) + [\mathcal{T}]\Big) . \tag{2.9}$$

Proof. This result has been proved in [GKM20, Lemma 3.3] for convex domains in $\mathbb{R}^d$; the proof on $\mathcal{S}^1$ proceeds *mutatis mutandi.* □

The following result provides a coarse bound in the opposite direction.

Proposition 2.2.8 (Coarse lower bound for $\mathcal{W}_{\mathcal{T}}$)**.** *Fix $\delta, \zeta \in (0,1)$. There exists a constant $C < \infty$ depending only on δ and ζ, such that for any ζ-regular mesh $\mathcal{T}$ on $\mathcal{S}^1$, and for any solution $(m_t, J_t)_t$ to the discrete continuity equation (2.8) satisfying $\delta \leq \frac{m_t(k)}{\pi_k} \leq \delta^{-1}$ for all $t \in [0,1]$ and $k = 0, \ldots K-1$, we have*

$$\mathbb{W}_2^2(\iota_{\mathcal{T}} m_0, \iota_{\mathcal{T}} m_1) \leq C \int_0^1 \mathcal{A}_{\mathcal{T}}(m_t, J_t) \, \mathrm{d}t . \tag{2.10}$$

Proof. We define

$$\mu_t = \iota_{\mathcal{T}} m_t , \qquad j_t(x) = \frac{r_{k+1} - x}{\pi_k} J_t(k-1,k) + \frac{x - r_k}{\pi_k} J_t(k,k+1) ,$$

for $x \in A_k$. It follows that (μ_t, j_t) solves the continuous continuity equation. Moreover,

$$\begin{aligned} \mathbb{A}(\mu_t, j_t) &= \sum_{k=0}^{K-1} \frac{\pi_k}{m_t(k)} \int_{r_k}^{r_{k+1}} \left(\frac{r_{k+1} - x}{\pi_k} J_t(k-1,k) + \frac{x - r_k}{\pi_k} J_t(k,k+1) \right)^2 \mathrm{d}x \\ &\leq \frac{1}{2} \sum_{k=0}^{K-1} \frac{\pi_k^2}{m_t(k)} \Big(J_t^2(k-1,k) + J_t^2(k,k+1) \Big) \\ &= \frac{1}{2} \sum_{k=0}^{K-1} J_t^2(k,k+1) \left(\frac{\pi_k^2}{m_t(k)} + \frac{\pi_{k+1}^2}{m_t(k+1)} \right) . \end{aligned}$$

Write $\rho_t(k) = \frac{m_t(k)}{\pi_k}$. In view of the bounds on $m_t(k)$, we have

$$\begin{aligned} \theta_{k,k+1}(\rho_t(k), \rho_t(k+1)) &\leq \max\{\rho_t(k), \rho_t(k+1)\} \\ &\leq \delta^{-2} \min\{\rho_t(k), \rho_t(k+1)\} \leq 2\delta^{-2} \left(\frac{1}{\rho_t(k)} + \frac{1}{\rho_t(k+1)} \right)^{-1} . \end{aligned}$$

Since $2\zeta[\mathcal{T}] \leq d_{k,k+1}$, we have

$$\frac{\pi_k^2}{m_t(k)} + \frac{\pi_{k+1}^2}{m_t(k+1)} \leq [\mathcal{T}] \left(\frac{\pi_k}{m_t(k)} + \frac{\pi_{k+1}}{m_t(k+1)} \right) \leq \frac{1}{\zeta\delta^2} \frac{d_{k,k+1}}{\theta_{k,k+1}(\rho_t(k), \rho_t(k+1))} .$$

It follows that

$$\mathbb{A}(\mu_t, j_t) \leq \frac{1}{2\zeta\delta^2} \sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(k,k+1)}{\theta_{k,k+1}(\rho_t(k), \rho_t(k+1))} = \frac{1}{2\zeta\delta^2} \mathcal{A}_{\mathcal{T}}(m_t, J_t) ,$$

which implies the result with $C = \frac{1}{2\zeta\delta^2}$. □

2.3 A simple proof of the lower bound in the 2-periodic case

In this section we focus on the simplest non-trivial periodic setting, which corresponds to taking $K = 2$ in Figure 2.1. In this setting we present a short proof of the lower bound in Theorem 2.1.1 by connecting the problem to known results from [GM13, GKM20]. This approach does not appear to generalise to $K \geq 3$.

We fix a parameter $r \in (0, 1)$, and consider the mesh $\mathcal{T}$ with $K = 2$, and $r_0 = 0$, $r_1 = r$, and $r_2 = 1$. To be able to apply the simple argument in this section, we define the points $z_0 = \frac{r}{2}$ and $z_1 = \frac{r+1}{2}$ to be the midpoints of the cells, so that $d_{01} = d_{12} = \frac{1}{2}$.

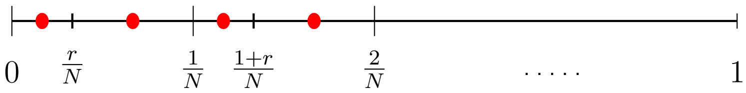

Figure 2.4: A 2-periodic mesh $\mathcal{T}_N$ on $\mathcal{S}^1$.

Throughout this section we make the standing assumption that $\theta_{01} = \theta_{21}$, and we simply write $\theta := \theta_{01}$. This implies that the constant $c^\star(\theta, r) := c^\star(\theta, \mathcal{T})$ is given by

$$c^\star(\theta, r) = \inf_{\alpha \in [0,1]} \frac{1}{\theta\left(\frac{\alpha}{r}, \frac{1-\alpha}{1-r}\right)} . \tag{2.11}$$

The notation $m(k) = m(A_k)$ allows us to canonically identify measures on $\mathcal{T}_N$ with measures on the equidistant mesh corresponding to $r = \frac{1}{2}$. We write $\mathcal{T}_{r,N}$ to emphasise the dependence of $\mathcal{T}_N$ on r, and write $\mathcal{A}^\theta_{r,N}$ and $\mathcal{W}^\theta_{r,N}$ to denote the corresponding energy and metric. The cells in $\mathcal{T}_{r,N}$ will be labeled $0, \ldots, 2N-1$.

2.3.1 Lower bound

The following lemma compares the discrete transport metric on the mesh $\mathcal{T}_{r,N}$ with the corresponding quantity on the equidistant mesh $\mathcal{T}_{\frac{1}{2},N}$.

Lemma 2.3.1. *Let $r \in (0, 1)$ and $N \geq 1$. For any $m_0, m_1 \in \mathcal{P}(\mathcal{T}_{r,N})$ we have*

$$\mathcal{W}^\theta_{r,N}(m_0, m_1) \geq \sqrt{c^\star(\theta, r)}\, \mathcal{W}^{\theta_a}_{\frac{1}{2},N}(m_0, m_1) \tag{2.12}$$

where θ_a denotes the arithmetic mean.

Proof. Note that

$$\mathcal{A}^\theta_{r,N}(m, J) = \frac{1}{2N^2} \sum_{n=0}^{N-1} \left\{ \frac{\big(J(2n, 2n+1)\big)^2}{\theta\left(\frac{m(2n)}{r}, \frac{m(2n+1)}{1-r}\right)} + \frac{\big(J(2n-1, 2n)\big)^2}{\theta\left(\frac{m(2n)}{r}, \frac{m(2n-1)}{1-r}\right)} \right\} .$$

The key observation is that the mean θ of the densities $\frac{m(2n)}{r}$ and $\frac{m(2n\pm1)}{1-r}$ on the mesh $\mathcal{T}_{r,N}$ can be estimated in terms of the *arithmetic mean* θ^a of the corresponding densities $\frac{m(2n)}{1/2}$ and $\frac{m(2n\pm1)}{1/2}$ on the symmetric mesh $\mathcal{T}_{\frac{1}{2},N}$. Indeed, as we can write

$$m(2n) = \alpha_n^\pm \Big(m(2n) + m(2n \pm 1)\Big) \quad \text{and} \quad m(2n \pm 1) = (1 - \alpha_n^\pm)\Big(m(2n) + m(2n \pm 1)\Big)$$

for some $\alpha_n^\pm \in [0,1]$, the 1-homogeneity of θ yields

$$\begin{aligned}\theta\bigg(\frac{m(2n)}{r}, \frac{m(2n\pm 1)}{1-r}\bigg) &= \big(m(2n)+m(2n\pm 1)\big)\theta\bigg(\frac{\alpha_n^\pm}{r}, \frac{1-\alpha_n^\pm}{1-r}\bigg) \\ &\leq \theta^a\bigg(\frac{m(2n)}{1/2}, \frac{m(2n\pm 1)}{1/2}\bigg)\frac{1}{c^\star(\theta,r)} \ .\end{aligned}$$

Consequently,

$$\mathcal{A}^\theta_{r,N}(m,J) \geq c^\star(\theta,r)\mathcal{A}^{\theta^a}_{\frac{1}{2},N}(m,J) \ .$$

As the continuity equation does not depend on r, this implies the result. □

The sought lower bound for $\mathcal{W}^\theta_{r,N}$ can now be easily obtained.

Corollary 2.3.2. *Fix $r \in (0,1)$. For any $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have*

$$\liminf_{N\to\infty} \mathcal{W}_{r,N}(P_N\mu_0, P_N\mu_1) \geq \sqrt{c^\star(\theta,r)}\mathbb{W}_2(\mu_0,\mu_1) \ .$$

Proof. This follows by applying Lemma 2.3.1 to the measures $m_i := P_N\mu_i$ and using the known convergence result for symmetric meshes [GM13, GKM20], which asserts that

$$\lim_{N\to\infty} \mathcal{W}^{\theta_a}_{\frac{1}{2},N}(P_N\mu_0, P_N\mu_1) = \mathbb{W}_2(\mu_0,\mu_1) \ .$$

□

For proving the corresponding upper bound

$$\limsup_{N\to\infty} \mathcal{W}_{r,N}(P_N\mu_0, P_N\mu_1) \leq \sqrt{c^\star(\theta,r)}\mathbb{W}_2(\mu_0,\mu_1) \ ,$$

the 2-periodic setting does not offer conceptual simplifications compared to the general K-periodic setting. Therefore, we will directly treat the K-periodic setting in Section 2.6.

2.3.2 Examples

We finish this section by explicitly computing the value of $c^\star(\theta,r)$ in a number of cases. We write

$$g_{\theta,r}(\alpha) := \theta\bigg(\frac{\alpha}{r}, \frac{1-\alpha}{1-r}\bigg) \ , \qquad \text{so that} \qquad c^\star(\theta,r) = \inf_{\alpha\in(0,1)} \frac{1}{g_{\theta,r}(\alpha)} \ .$$

Example 2.3.3 (θ is r-balanced)**.** Suppose that $\theta(a,b) \leq ra + (1-r)b$ for any $a,b \geq 0$ (i.e., θ is r-balanced in the sense of Definition 2.4.6 below). Applying this inequality to $a = \frac{\alpha}{r}$ and $b = \frac{1-\alpha}{1-r}$ we immediately obtain $g_{\theta,r}(\alpha) \leq 1$ for all $\alpha \in [0,1]$. Since $g_{\theta,r}(r) = \theta(1,1) = 1$ by assumption, it follows that

$$c^\star(\theta,r) = \frac{1}{g_{\theta,r}(r)} = 1 \ .$$

Example 2.3.4 (Geometric mean)**.** Let $\theta(a,b) = \sqrt{ab}$. Then $g_{\theta,r}(\alpha) = \sqrt{\frac{\alpha(1-\alpha)}{r(1-r)}}$ is uniquely maximised at $\alpha = \frac{1}{2}$, and we obtain

$$c^\star(\theta, r) = 2\sqrt{r(1-r)} \ .$$

Note that the fact that $\alpha^\star = \frac{1}{2}$ means that the mass is equally distributed among large and small cells, irrespectively of the value of r. Thus, there will be no oscillations for the optimal discrete measures; however, this means that oscillations at the level of the density do occur.

Example 2.3.5 (Harmonic mean)**.** Let $\theta(a,b) = \frac{2ab}{a+b}$. In this case we have $g(\alpha) := \frac{1}{g_{\theta,r}(\alpha)} = \frac{1}{2}\Big(\frac{1-r}{1-\alpha} + \frac{r}{\alpha}\Big)$, and $g'(\alpha) = \frac{1-r}{2(1-\alpha)^2} - \frac{r}{2\alpha^2}$. It follows that g' vanishes at $\alpha^\star = \frac{\sqrt{r}}{\sqrt{r}+\sqrt{1-r}}$, which is indeed the unique minimiser of g. Consequently,

$$c^\star(\theta, r) = g(\alpha^\star) = \frac{1}{2}\Big(\sqrt{r} + \sqrt{1-r}\Big)^2 \ .$$

Example 2.3.6 (Arithmetic mean)**.** Let $\theta(a,b) = \frac{a+b}{2}$. Then $g_{\theta,r}(\alpha) = \frac{1}{2}\Big(\frac{1-\alpha}{1-r} + \frac{\alpha}{r}\Big)$ is affine in α. If $r < \frac{1}{2}$ (resp. $r > \frac{1}{2}$), the maximum is attained at $\alpha^\star = 1$ (resp. $\alpha^\star = 0$). In both cases, this means that all the mass will be assigned to the small cells. It follows that

$$c^\star(\theta, r) = 2\min\{r, 1-r\} \ .$$

Example 2.3.7 (Minimum)**.** Let $\theta(a,b) = \min\{a,b\}$. In this case, $g_{\theta,r}(\alpha) = \min\{\frac{1-\alpha}{1-r}, \frac{\alpha}{r}\}$ is uniquely maximised at $\alpha^\star = r$. This means that the assigned mass is proportional to the size of the cells, hence there are no oscillations at the level at the density. We find

$$c^\star(\theta, r) = \frac{1}{g_{\theta,r}(\alpha^\star)} = 1 \ .$$

2.4 Analysis of the effective mobility

In this section we investigate some basic properties of the effective mobility $c^\star(\theta, \mathcal{T})$ defined in (2.5), and relate its value to certain geometric properties of the mesh $\mathcal{T}$ that have been considered in [GKM20]. Recall:

$$c^\star(\theta, \mathcal{T}) := \inf \left\{ \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\Big(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\Big)} \ : \ m \in \mathcal{P}(\mathcal{T}) \right\} \ . \tag{2.13}$$

We start with a simple observation.

Proposition 2.4.1. *For any mesh $\mathcal{T}$ on $\mathcal{S}^1$ and any family of means $\theta = \{\theta_{k,k+1}\}_{k=0}^{K-1}$ we have $c^\star(\theta, \mathcal{T}) \leq 1$.*

Proof. This follows by using the competitor $m_k = \pi_k$ in (2.13). □

In view of this result, Theorem 2.1.1 implies the upper bound

$$\limsup_{N\to\infty} \mathcal{W}_N(P_N\mu_0, P_N\mu_1) \leq \mathbb{W}_2(\mu_0, \mu_1) \ ,$$

which had already been proved in [GKM20].

Proposition 2.4.2. *The infimum in* (2.13) *is attained.*

Proof. This readily follows using the lower-semicontinuity result from Lemma 2.2.3. □

In the remainder of this section we shall investigate under which conditions on θ and $\mathcal{T}$ we have $c^\star(\theta, \mathcal{T}) = 1$. For this purpose, we consider two geometric conditions:

Definition 2.4.3 (Geometric conditions on the mesh)**.** Fix $\{\lambda_{k,k+1}\}_{k=0}^{K-1} \in [0,1]^K$, and set $\lambda_{k+1,k} = 1 - \lambda_{k,k+1}$. We say that a mesh $\mathcal{T} = \mathcal{T}_{\pi,z}$ on $\mathcal{S}^1$ satisfies

1. the *center-of-mass condition* with parameters $\{\lambda_{k,k+1}\}_{k=0}^{K-1}$ if, for all k,
$$r_{k+1} = \lambda_{k+1,k} z_k + \lambda_{k,k+1} z_{k+1} \ ; \tag{2.14}$$
2. the *isotropy condition* with parameters $\{\lambda_{k,k+1}\}_{k=0}^{K-1}$ if, for all k,
$$\pi_k = \lambda_{k,k-1} d_{k-1,k} + \lambda_{k,k+1} d_{k,k+1} \ . \tag{2.15}$$

Both of these conditions have been studied for meshes on bounded convex domains in $\mathbb{R}^d$ in [GKM20]. The center-of-mass condition asserts that the center of mass of the cell interfaces lie on the line segment connecting the support points of the respective cells. In dimensions $d \geq 2$, this condition poses a strong geometric condition on the mesh. However, in our one-dimensional context, the condition is always satisfied, for a unique choice of the parameters $\{\lambda_{k,k+1}\}_k$. The isotropy condition is weaker than the center-of-mass condition: it holds with the same parameters, but there is an additional degree of freedom, as the following result shows.

Proposition 2.4.4. *Let* $\mathcal{T} = \mathcal{T}_{\pi,z}$ *be a mesh on* $\mathcal{S}^1$ *and set* $\overline{\lambda}_{k,k+1} = \frac{r_{k+1} - z_k}{z_{k+1} - z_k}$ *for* $k = 0, \ldots, K-1$. *For* $\{\lambda_{k,k+1}\} \subseteq [0,1]$ *the following assertions hold:*

1. *The center-of-mass condition holds if and only if for any* $k = 0, \ldots, K-1$,
$$\lambda_{k,k+1} = \overline{\lambda}_{k,k+1} \ .$$
2. *The isotropy condition holds if and only if there exists* $s \in [-\min_k \overline{\lambda}_{k,k+1} d_{k,k+1}, \min_k \overline{\lambda}_{k+1,k} d_{k,k+1}]$ *such that for any* $k = 0, \ldots, K-1$,
$$\lambda_{k,k+1} = \overline{\lambda}_{k,k+1} + \frac{s}{d_{k,k+1}} \ .$$

Proof. This follows immediately by solving the corresponding linear systems. □

Remark 2.4.5 (Relation to the asymptotic isotropy condition). Recall from [GKM20, Definition 1.3] that a family of meshes $\{\mathcal{T}\}$ (in any dimension) is said to satisfy the *isotropy condition* with parameters $\{\lambda_{KL}\}$ if, for any $K \in \mathcal{T}$,
$$\sum_{L \in \mathcal{T}} \lambda_{KL} \frac{\mathscr{H}^{d-1}(\partial K \cap \partial L)}{|z_K - z_L|} (z_K - z_L) \otimes (z_K - z_L) \leq |K| \Big(I_d + \eta_{\mathcal{T}}(K) \Big) \tag{2.16}$$
where $\sup_{K \in \mathcal{T}} |\eta_{\mathcal{T}}(K)| \to 0$ as $\max\{\operatorname{diam}(A) \,:\, A \in \mathcal{T}\} \to 0$.

Applying this condition to the family of one-dimensional periodic meshes $\mathcal{T}^N$ constructed from $\mathcal{T}$, it reduces to

$$\lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1} \leq \pi_k(1+\eta_N(k))$$

for all $N \geq 1$ and $k = 0, \dots, K-1$, where $\eta_N(k) \to 0$ as $N \to \infty$. As the left-hand side does not depend on N, this condition in turn simplifies to

$$\lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1} \leq \pi_k \tag{2.17}$$

for all $k = 0, \dots, K-1$.

Clearly, (2.15) implies (2.17). To see that both assertions are equivalent, we note that (2.17) can be written as

$$\lambda_{k,k+1}d_{k,k+1} - \lambda_{k-1,k}d_{k-1,k} \leq \pi_k - d_{k-1,k} \ . \tag{2.18}$$

To obtain a contradition, suppose that we have strict inequality in (2.18) for some $k = \bar{k}$. Summation over $k = 0, \dots, K-1$ yields

$$0 = \sum_{k=0}^{K-1} \Big(\lambda_{k,k+1}d_{k,k+1} - \lambda_{k-1,k}d_{k-1,k}\Big) < \sum_{k=0}^{K-1} \Big(\pi_k - d_{k-1,k}\Big) = 0 \ ,$$

which is absurd.

In summary, we conclude that the isotropy condition (2.15) is equivalent to the asymptotic isotropy condition (2.16) for the family of meshes $\{\mathcal{T}^N\}$.

The next definition will be used to connect geometric properties of the mesh to properties of the means in the definition of the transport distance.

Definition 2.4.6 (Adaptedness)**.** Let $\lambda, \lambda_{k,k+1} \in [0,1]$ for $k = 0, \dots, K-1$.

1. A mean θ is said to be λ-*balanced* if $\theta(a,b) \leq \lambda a + (1-\lambda)b$ for any $a, b \geq 0$.
2. A family of means $\{\theta_{k,k+1}\}$ is said to be *adapted* to the parameters $\{\lambda_{k,k+1}\}$ if $\theta_{k,k+1}$ is $\lambda_{k,k+1}$-balanced for each k.

Remark 2.4.7. Each continuously differentiable mean θ is λ-balanced for exactly one value of $\lambda \in [0,1]$, namely

$$\lambda = \partial_1\theta(1,1) \ . \tag{2.19}$$

A non-smooth mean θ can be λ-balanced for multiple values of λ, e.g., the mean $(a,b) \mapsto \min\{a,b\}$ is λ-balanced for any $\lambda \in [0,1]$.

Now we are ready to state the main result of this section. The result is consistent with the main result in [GKM20], which asserts that the asymptotic isotropy condition is necessary (and essentially sufficient) for Gromov–Hausdorff convergence of the discrete transport distance to $\mathbb{W}_2$.

Theorem 2.4.8 (Isotropy is (essentially) equivalent to $c^\star(\theta,\mathcal{T}) = 1$)**.** *Let $\{\theta_{k,k+1}\}$ be a family of means that are adapted to $\{\lambda_{k,k+1}\}$.*

1. *If $\mathcal{T}$ satisfies the isotropy condition with parameters $\{\lambda_{k,k+1}\}$, then $c^\star(\theta,\mathcal{T}) = 1$.*

2. *Assume that each mean $\theta_{k,k+1}$ is continuously differentiable. If $c^\star(\theta,\mathcal{T}) = 1$, then $\mathcal{T}$ satisfies the isotropy condition with parameters $\{\lambda_{k,k+1}\}$.*

Remark 2.4.9 (Minimum mean). In view of Proposition 2.4.4, every mesh $\mathcal{T}$ satisfies the isotropy condition for a suitable choice of $\{\lambda_k\}$. Since the minimum mean $(a,b) \mapsto \min\{a,b\}$ is λ-balanced for any value of $\lambda \in [0,1]$, it thus follows from Theorem 2.4.8 that $c^\star(\theta,\mathcal{T}) = 1$ if $\theta_{k,k+1} = \min$ for each k.

Proof. To prove (1), take any sequence $\{m_k\}_k$ with $\sum_{k=0}^{K-1} m_k = 1$. Using Jensen's inequality, the adaptedness, the periodicity, and the isotropy condition, we obtain

$$\begin{aligned}\sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\right)} &\geq \left(\sum_{k=0}^{K-1} d_{k,k+1}\theta_{k,k+1}\left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\right)\right)^{-1} \\ &\geq \left(\sum_{k=0}^{K-1} d_{k,k+1}\left(\lambda_{k,k+1}\frac{m_k}{\pi_k} + \lambda_{k+1,k}\frac{m_{k+1}}{\pi_{k+1}}\right)\right)^{-1} \\ &= \left(\sum_{k=0}^{K-1} \frac{m_k}{\pi_k}\big(\lambda_{k,k+1}d_{k,k+1} + \lambda_{k,k-1}d_{k-1,k}\big)\right)^{-1} \\ &= \left(\sum_{k=0}^{K-1} m_k\right)^{-1} = 1\ .\end{aligned}$$

Taking the infimum over $\{m_k\}_k$, we obtain $c^\star(\theta,\mathcal{T}) \geq 1$. In view of Proposition 2.4.1 we infer that $c^\star(\theta,\mathcal{T}) = 1$.

To prove (2), we consider the probability measures γ_α^k defined by

$$\gamma_\alpha^k = (\pi_0, \ldots, \pi_{k-1}, \pi_k + \alpha, \pi_{k+1} - \alpha, \pi_{k+2}, \ldots, \pi_{K-1})$$

for $|\alpha|$ sufficiently small. Let us write

$$h_{\theta,\mathcal{T}}(m) = \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}}\right)}\ .$$

As $c^\star(\theta,\mathcal{T}) = 1$, we have $h_{\theta,\mathcal{T}}(m) \geq 1$ for all m. Thus, since $h_{\theta,\mathcal{T}}(\gamma_0^k) = h_{\theta,\mathcal{T}}(\pi) = 1$, it follows that $\frac{\mathrm{d}}{\mathrm{d}\alpha}\Big|_{\alpha=0} h_{\theta,\mathcal{T}}(\gamma_\alpha^k) = 0$. A direct computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\Big|_{\alpha=0} h_{\theta,\mathcal{T}}(\gamma_\alpha^k) = B_{k+1} - B_k \quad \text{where} \quad B_k := \frac{\lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1}}{\pi_k}\ .$$

As this holds for every k, we infer that there exists a constant $\beta > 0$ such that $B_k = \beta$ for every $k = 0, \ldots, K-1$. The latter means that

$$\beta\pi_k = \lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1}$$

for all $k = 0, \ldots, K-1$. Summation over k yields

$$\beta = \beta\sum_{k=0}^{K-1}\pi_k = \sum_{k=0}^{K-1}(1-\lambda_{k-1,k})d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1} = \sum_{k=0}^{K-1} d_{k-1,k} = 1\ ,$$

which proves the isotropy condition with parameters $\{\lambda_{k,k+1}\}_k$. □

2.5 Proof of the lower bound

The goal of this section is to prove Theorem 2.5.4, which yields the lower bound in Theorem 2.1.1. The crucial ingredient is Proposition 2.5.3, which ensures the existence of approximately optimal curves with good regularity properties.

To formulate this result, we fix a non-negative function $\eta \in C_c^\infty(0, \frac{1}{2})$ with $\int_0^1 \eta(x)\,\mathrm{d}x = 1$. We set $\eta_\lambda(x) = \frac{1}{\lambda}\eta(\frac{x}{\lambda})$ for $x \in [0,1)$, and consider its periodic extension to $\mathcal{S}^1$. For $\lambda \in (0,1]$ we define a discrete spatial mollifier by

$$\eta_\lambda^N(n) := \frac{N}{\lambda}\int_{\frac{n}{N}}^{\frac{n+1}{N}} \eta\Big(\frac{x}{\lambda}\Big)\,\mathrm{d}x\ , \qquad n = 0, \dots, N-1\ ,$$

and we extend η_λ^N to $\mathbb{Z}$ periodically modulo N, so that it can be regarded as a function on the discrete torus $\mathbb{T}_N = \mathbb{Z}/N\mathbb{Z}$. It follows that $\frac{1}{N}\sum_{n=0}^{N-1}\eta_\lambda^N(n) = 1$, and the following kernel bounds hold for $n = 0, \dots, N-1$:

$$|\eta_\lambda^N(n)| \leq \frac{\|\eta\|_\infty}{\lambda}\ , \qquad |\eta_\lambda^N(n_1) - \eta_\lambda^N(n_2)| \leq \frac{\|\eta'\|_\infty}{\lambda^2}\frac{|n_1 - n_2|}{N}\ , \tag{2.20}$$

We consider the convolution operators $\mathbb{M}_\lambda : L^1(\mathcal{S}^1) \to L^\infty(\mathcal{S}^1)$ given by

$$\Big(\mathbb{M}_\lambda f\Big)(x) = \int_{\mathcal{S}^1} \eta_\lambda(x-y) f(y)\,\mathrm{d}y\ ,$$

as well as the analogous discrete convolution operators $\mathcal{M}_\lambda^N : L^1(\mathbb{T}_N) \to L^\infty(\mathbb{T}_N)$ defined by

$$\Big(\mathcal{M}_\lambda^N \psi\Big)(n) = \frac{1}{N}\sum_{j=0}^{N-1} \eta_\lambda^N(n-j)\psi(j)\ .$$

The kernel bounds (2.20) imply the following L^1-L^∞ and L^1-Lipschitz bounds:

$$\sup_n |\mathcal{M}_\lambda^N\psi(n)| \leq \frac{\|\eta\|_\infty}{\lambda N}\sum_{n=0}^{N-1}|\psi(n)|\ , \tag{2.21}$$

$$\sup_n |\mathcal{M}_\lambda^N\psi(n_1) - \mathcal{M}_\lambda^N\psi(n_2)| \leq \frac{\|\eta'\|_\infty}{\lambda^2}\frac{|n_1-n_2|}{N^2}\sum_{n=0}^{N-1}|\psi(n)|\ . \tag{2.22}$$

The following result contains some basic properties of convolution operators that will be used in the sequel.

Lemma 2.5.1 (Bounds for convolution operators). *Let $\lambda \in (0,1]$ and $N \geq 2$. For any $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $m \in \mathcal{P}(\mathbb{T}_N)$ we have*

$$\mathbb{W}_2(\mu, \mathbb{M}_\lambda\mu) \leq C\lambda\ , \tag{2.23}$$

$$\mathbb{W}_2(\iota_N\mathcal{M}_\lambda^N m, \mathbb{M}_\lambda\iota_N m) \leq \frac{\lambda}{2} + \frac{2}{N}\ , \tag{2.24}$$

where $C < \infty$ depends only on η.

Proof. The inequality (2.23) follows straightforwardly using the coupling $\gamma(\mathrm{d}x,\mathrm{d}y) = \eta_\lambda(y - x)\,\mathrm{d}\mu(x)\,\mathrm{d}y$.

To prove (2.24), let δ_i be the Dirac mass at i, and note that

$$\mathbb{W}_2^2(\iota_N \mathcal{M}_\lambda^N m, \mathbb{M}_\lambda \iota_N m) \leq \sum_{i=0}^{N-1} m_i \mathbb{W}_2^2(\iota_N \mathcal{M}_\lambda^N \delta_i, \mathbb{M}_\lambda \iota_N \delta_i)$$

by convexity of $\mathbb{W}_2^2$. Thus it suffices to prove the lemma for $m = \delta_i$. Since $\mathrm{d}\Big(\iota_N\delta_i\Big)(x) = N\mathbb{1}_{\left[\frac{i}{N},\frac{i+1}{N}\right]}(x)\,\mathrm{d}x$, we have

$$\mathrm{d}\Big(\mathbb{M}_\lambda \iota_N \delta_i\Big)(x) = N\Big(\eta_\lambda * \mathbb{1}_{[\frac{i}{N},\frac{i+1}{N}]}\Big)(x)\,\mathrm{d}x\ . \tag{2.25}$$

On the other hand, we have

$$\mathrm{d}\Big(\iota_N \mathcal{M}_\lambda^N \delta_i\Big)(x) = \sum_{n=0}^{N-1} \eta_\lambda^N(n-i)\mathbb{1}_{[\frac{n}{N},\frac{n+1}{N}]}(x)\,\mathrm{d}x\ .$$

Since $\operatorname{supp}\eta_\lambda \subset \Big(0,\frac{\lambda}{2}\Big)$, we obtain

$$\operatorname{supp}\mathbb{M}_\lambda\iota_N\delta_i \subseteq \Big[\tfrac{i}{N},\tfrac{i+1}{N}+\tfrac{\lambda}{2}\Big] \quad \text{and} \quad \operatorname{supp}\iota_N\mathcal{M}_\lambda^N\delta_i \subseteq \Big[\tfrac{i}{N},\tfrac{i+1}{N}+\tfrac{\lambda}{2}\Big]\ ,$$

hence

$$\operatorname{diam}\Big(\operatorname{supp}(\iota_N\mathcal{M}_\lambda^N\delta_i)\cup\operatorname{supp}(\mathbb{M}_\lambda\iota_N\delta_i)\Big) \leq \frac{\lambda}{2}+\frac{2}{N}\ .$$

This easily yields the desired result. □

Before stating the crucial regularisation result, we formulate a lemma which asserts that we can decrease the energy at the discrete level by a suitable regularisation. Here it is crucial that the regularisation is performed by averaging the density at spatial locations $nK + k$ and $n'K + k$ that differ by a multiple of the period K. A "naive" regularisation consisting of locally averaging the density, without taking the periodic structure into account, would in general not decrease the energy. We emphasise that the operator $\mathcal{M}_\lambda^N$ is understood to act on the variable n in the result below, namely

$$(\mathcal{M}_\lambda^N m)(n;k) = \frac{1}{N}\sum_{j=0}^{N-1}\eta_\lambda^N(n-j)m(j;k).$$

With this notation we have the following result.

Lemma 2.5.2 (Energy bound under periodic smoothing). *Let $\lambda \in (0,1]$. For any $m \in \mathcal{P}(\mathcal{T}_N)$ and any $J \in \mathcal{V}(\mathcal{T}_N)$ we have*

$$\mathcal{A}_N(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J) \leq \mathcal{A}_N(m,J)\ .$$

Proof. For brevity we write

$$G_{k,k+1}(m,J,n) = \frac{d_{k,k+1}}{N} f_{k,k+1}\bigg(N\frac{m(n;k)}{\pi_k}, N\frac{m(n;k+1)}{\pi_{k+1}}, J(n;k,k+1)\bigg)\ .$$

Applying Jensen's inequality to the jointly convex functions $f_{k,k+1}$ we obtain

$$\begin{aligned}\mathcal{A}_N(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J) &= \sum_{k=0}^{K-1}\sum_{n=0}^{N-1} G_{k,k+1}(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J, n)\\ &\le \sum_{k=0}^{K-1}\sum_{n=0}^{N-1} \frac{1}{N}\sum_{\ell=0}^{N-1} \eta_\lambda^N(n-\ell) G_{k,k+1}(m,J,\ell)\\ &= \sum_{k=0}^{K-1}\sum_{\ell=0}^{N-1}\left(\frac{1}{N}\sum_{n=0}^{N-1}\eta_\lambda^N(n-\ell)\right) G_{k,k+1}(m,J,\ell)\\ &= \sum_{k=0}^{K-1}\sum_{\ell=0}^{N-1} G_{k,k+1}(m,J,\ell) = \mathcal{A}_N(m,J)\ ,\end{aligned}$$

where we used that $\frac{1}{N}\sum_{n=0}^{N-1}\eta_\lambda^N(n) = 1$. □

We are now ready to state the main regularisation result of this section. As we expect that (approximately) optimal densities exhibit oscillations, we cannot expect spatial regularity for such densities. Nevertheless, the lemma above allows us to obtain a restricted form of regularity for such densities, in the sense that good Lipschitz bounds hold if one only compares values of the density at spatial locations $nK + k$ and $n'K + k$ that differ by a multiple of the period K.

Note that the vector field J enjoys better regularity properties: in (2.27e) we even obtain a Lipschitz bound for neighbouring cells.

Proposition 2.5.3 (Space-time regularisation)**.** *Fix $N \ge 1$, and let $(m_t, J_t)_t$ be a solution to the discrete continuity equation (2.8) in $\mathcal{P}(\mathcal{T}_N)$ satisfying*

$$A := \int_0^1 \mathcal{A}_N(m_t, J_t)\,\mathrm{d}t < \infty\ .$$

Then, for any $\varepsilon > 0$ there exists a solution $(\tilde m_t, \tilde J_t)_t$ to (2.8) such that:

1. $\mathbb{W}_2(\iota_N \tilde m_t, \iota_N m_t) \le \varepsilon + \frac{C}{N}$ *for all $t \in [0,1]$, where $C < \infty$ depends only on $\mathcal{T}$;*
2. *the following action bound holds:*
$$\int_0^1 \mathcal{A}_N(\tilde m_t, \tilde J_t)\,\mathrm{d}t \le \int_0^1 \mathcal{A}_N(m_t, J_t)\,\mathrm{d}t\ ; \tag{2.26}$$
3. *the following regularity properties hold, for some constants $c_0, \ldots, c_5 < \infty$ depending on ε and A, but not on N:*
$$c_0^{-1} \le \min_{n,k} N\tilde m_t(n;k) \le \max_{n,k} N\tilde m_t(n;k) \le c_1\ , \tag{2.27a}$$
$$\sup_{t\in[0,1]}\max_{n,k}\left|N\partial_t \tilde m_t^N(n;k)\right| \le c_2\ , \tag{2.27b}$$
$$\sup_{t\in[0,1]}\max_{n,k}\left|N\tilde m_t(n;k) - N\tilde m_t(n+1;k)\right| \le \frac{c_3}{N}\ , \tag{2.27c}$$
$$\sup_{t\in[0,1]}\max_{n,k}\left|\tilde J_t^N(n;k,k+1)\right| \le c_4\ , \tag{2.27d}$$
$$\sup_{t\in[0,1]}\max_{n,k}\left|\tilde J_t(n;k,k+1) - \tilde J_t(n;k-1,k)\right| \le \frac{c_5}{N}\ . \tag{2.27e}$$

Proof. Let $\mathcal{U}_N := P_N \mathscr{L}^1|_{S^1} \in \mathcal{P}(\mathcal{T}_N)$ denote the probability measure that assigns mass $\frac{\pi_k}{N}$ to $A_{n;k}$. Fix a mollifier η as above. For $\lambda, \tau, \delta > 0$ we define a space-time regularisation by

$$\tilde{m}_t(n;k) := \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{M}_\lambda^N \Big[(1-\delta) m_u + \delta \mathcal{U}_N \Big](n;k) \, \mathrm{d}u \ , \tag{2.28a}$$

$$\tilde{J}_t(n;k,k+1) := \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{M}_\lambda^N J_u(n;k,k+1) \, \mathrm{d}u \ . \tag{2.28b}$$

In both expressions, the operator $\mathcal{M}_\lambda^N$ is understood to act on the variable n, i.e., the spatial averaging takes place over cells whose distance is an integer multiple of the period K. Moreover, we use the convention that $m_u = m_0$ and $J_u = 0$ for $u < 0$, and $m_u = m_1$ and $J_u = 0$ for $u > 1$. We claim that this approximation satisfies all the sought properties.

An explicit computation shows that $(\tilde{m}_t, \tilde{V}_t)_t$ solves the discrete continuity equation.

To prove (2.26), we note that by a trifold application of the joint convexity of $\mathcal{A}_N$,

$$\begin{aligned}
\mathcal{A}_N(\tilde{m}_t, \tilde{J}_t) &\leq \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N \Big(\mathcal{M}_\lambda^N \big[(1-\delta) m_u + \delta \mathcal{U}_N \big], (1-\delta) \mathcal{M}_\lambda^N J_u \Big) \, \mathrm{d}u \\
&\leq \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N \Big(\big[(1-\delta) m_u + \delta \mathcal{U}_N \big], (1-\delta) J_u \Big) \, \mathrm{d}u \\
&\leq \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N(m_u, J_u) \, \mathrm{d}u \ .
\end{aligned}$$

Here we used the crucial regularisation bound from Lemma 2.5.2. The desired inequality (2.26) follows.

Moreover, since $\mathcal{M}_\lambda^N$ preserves positivity, we deduce the lower bound in (2.27a) with $c_0^{-1} = \delta \min_k \pi_k$.

To prove the upper bound in (2.27a), we use the fact that m_t is a probability measure and the L^1-L^∞ bound (2.21) to obtain

$$N \tilde{m}_t(n;k) \leq \frac{\|\eta\|_\infty}{\lambda} =: c_1 \ .$$

To prove (2.27b), we observe that

$$\partial_t \tilde{m}_t = \frac{1-\delta}{2\tau} \mathcal{M}_\lambda^N \Big[m_{t+\tau} - m_{t-\tau} \Big] \ .$$

Therefore, by another application of the L^1-L^∞-bound in (2.21), we arrive at

$$N |\partial_t \tilde{m}_t(n;k)| \leq \frac{\|\eta\|_\infty}{\tau \lambda} =: c_2 \ ,$$

which proves (2.27b).

The inequality (2.27e), with $c_5 = c_2$, follows immediately from (2.27b) and the fact that $(\tilde{m}_t^N, \tilde{J}_t)_t$ solves the continuity equation.

Furthermore, since

$$|\tilde{m}_t(n;k) - \tilde{m}_t(n+1;k)| \leq \sup_s |\mathcal{M}_\lambda^N m_s(n;k) - \mathcal{M}_\lambda^N m_s(n+1;k)| \leq \frac{\|\eta'\|_\infty}{\lambda^2 N^2} \ ,$$

we obtain (2.27c) with $c_3 = \frac{\|\eta'\|_\infty}{\lambda^2}$, in view of the Lipschitz bound in (2.22).

Finally, to obtain the L^∞-bound on the vector field (2.27d), we use (2.21) again to infer

$$\begin{aligned}\sup_{n,k} |\tilde{J}_t(n;k,k+1)| &\leq \frac{1-\delta}{2\tau}\int_{t-\tau}^{t+\tau} \sup_{n,k} |\mathcal{M}_\lambda^N J_u(n;k,k+1)| \,\mathrm{d}u \\ &\leq \frac{\|\eta\|_\infty}{\lambda N}\frac{1-\delta}{2\tau}\int_{t-\tau}^{t+\tau} \sup_k \sum_{n=0}^{N-1} |J_u(n;k,k+1)| \,\mathrm{d}u \ .\end{aligned}$$

Writing $\theta_{n;k,k+1} = \theta_{k,k+1}\Big(N\frac{m_u(n;k)}{\pi_k}, N\frac{m_u(n;k+1)}{\pi_{k+1}}\Big)$ for brevity, we infer that

$$\begin{aligned}\frac{1}{N}\bigg(\sum_{n,k} |J_u(n;k,k+1)|\bigg)^2 &\leq \bigg(\sum_{n,k} \frac{d_{k,k+1}}{N}\frac{J_u^2(n;k,k+1)}{\theta_{n;k,k+1}}\bigg)\bigg(\sum_{n,k}\frac{\theta_{n;k,k+1}}{d_{k,k+1}}\bigg) \\ &= \mathcal{A}_N(m_u, J_u)\sum_{n,k}\frac{\theta_{n;k,k+1}}{d_{k,k+1}} \ .\end{aligned}$$

Using the bound $\theta_{k,k+1}(a,b) \leq a+b$ we obtain

$$\sum_{n,k}\frac{\theta_{n;k,k+1}}{d_{k,k+1}} \leq \frac{N}{\min_k d_{k,k+1}}\sum_{n,k}\bigg(\frac{m_u(n;k)}{\pi_k} + \frac{m_u(n;k+1)}{\pi_{k+1}}\bigg) \leq 2BN \ ,$$

where $B = (\max_k \pi_k^{-1})(\max_k d_{k,k+1}^{-1})$. Combining these bounds, we arrive at

$$\begin{aligned}\sup_{n,k} |\tilde{J}_t(n;k,k+1)| &\leq \frac{\|\eta\|_\infty\sqrt{2B}}{2\tau\lambda}\int_{t-\tau}^{t+\tau}\sqrt{\mathcal{A}_N(m_u, J_u)}\,\mathrm{d}u \\ &\leq \frac{\|\eta\|_\infty}{\lambda}\sqrt{\frac{B}{\tau}\int_0^1 \mathcal{A}_N(m_u,J_u)\,\mathrm{d}u} \ ,\end{aligned}$$

which yields (2.27d) with $c_4 := \frac{\|\eta\|_\infty}{\lambda}\sqrt{\frac{AB}{\tau}}$. As we will choose $\delta, \lambda, \tau > 0$ depending on ε, the bounds (2.27a)–(2.27e) follow.

It remains to show that $\mathbb{W}_2(\iota_N\tilde{m}_t, \iota_N m_t) \leq \varepsilon + \frac{C}{N}$ for suitable values of δ, λ and τ. We consider the effect of the three different regularisations separately. First we apply the convexity of $\mathbb{W}_2^2$ to obtain for any $m \in \mathcal{P}(\mathcal{T}_N)$,

$$\mathbb{W}_2^2\Big(\iota_N m, \iota_N[(1-\delta)m + \delta\mathcal{U}_N]\Big) \leq \delta\mathbb{W}_2^2(\iota_N m, \mathscr{L}^1|_{\mathcal{S}^1}) \leq \frac{\delta}{4} \ , \tag{2.29}$$

since the diameter of $(\mathcal{P}(\mathcal{S}^1), \mathbb{W}_2)$ is equal to $\frac{1}{2}$. Moreover, for $m \in \mathcal{P}(\mathcal{T}_N)$, Lemma 2.5.1 yields

$$\begin{aligned}\mathbb{W}_2(\iota_N m, \iota_N\mathcal{M}_\lambda^N m) &\leq \mathbb{W}_2(\iota_N m, \mathbb{M}_\lambda\iota_N m) + \mathbb{W}_2(\mathbb{M}_\lambda\iota_N m, \iota_N\mathcal{M}_\lambda^N m) \\ &\leq C\bigg(\lambda + \frac{1}{N}\bigg) \ ,\end{aligned} \tag{2.30}$$

where $C < \infty$ depends only on η. Furthermore, set $\bar{m}_t = \mathcal{M}_\lambda^N\Big((1-\delta)m_t + \delta\mathcal{U}_N\Big)$ and $\bar{J}_t = (1-\delta)\mathcal{M}_\lambda^N J_t$. It then follows that $c_0 \leq N\bar{m}_t \leq c_1$ and $\int_0^1 \mathcal{A}_N(\bar{m}_t, \bar{J}_t)\,\mathrm{d}t \leq A$. Thus, for $s \le t$, Proposition 2.2.8 yields a constant $\kappa < \infty$ depending on c_0 and c_1 (hence on δ and λ) such that,

$$\begin{aligned}\mathbb{W}_2^2(\iota_N \bar{m}_s, \iota_N \bar{m}_t) &\le \kappa \int_0^1 \mathcal{A}(\bar{m}_{(1-a)s+at}, (t-s)\bar{J}_{(1-a)s+at})\,\mathrm{d}a \\ &\le \kappa(t-s)\int_s^t \mathcal{A}(\bar{m}_u, \bar{J}_u)\,\mathrm{d}u \\ &\le \kappa A(t-s)\ .\end{aligned}$$

By convexity of $\mathbb{W}_2^2$, we obtain

$$\mathbb{W}_2^2\Big(\iota_N \bar{m}_t, \iota_N\Big[\frac{1}{2\tau}\int_{t-\tau}^{t+\tau} \bar{m}_u\,\mathrm{d}u\Big]\Big) \le \frac{1}{2\tau}\int_{t-\tau}^{t+\tau} \mathbb{W}_2^2(\iota_N \bar{m}_t, \iota_N \bar{m}_u)\,\mathrm{d}u \le \frac{\kappa\tau A}{2}\ . \tag{2.31}$$

Applying the estimates (2.29) with $m = m_t$, (2.30) with $m = (1-\delta)m_t + \delta\mathcal{U}_N$ and (2.31), we arrive at

$$\mathbb{W}_2\big(\iota_N \tilde{m}_t, \iota_N m_t\big) \le C\Big(\sqrt{\delta} + \lambda + \frac{1}{N} + \sqrt{\kappa\tau A}\Big)\ ,$$

for some $C < \infty$ depending only on $\mathcal{T}$ and on η. Thus, choosing first λ and δ sufficiently small, and then τ sufficiently small depending on δ, the result follows. □

We are now ready to prove the lower bound in Theorem 2.1.1.

Theorem 2.5.4 (Lower bound for $\mathcal{W}_N$). *For any mesh $\mathcal{T}$ and any family of admisible means $\{\theta_{k,k+1}\}_k$ we have*

$$c^\star(\theta,\mathcal{T})\mathbb{W}_2^2(\mu_0,\mu_1) \le \liminf_{N\to\infty} \mathcal{W}_N^2(P_N\mu_0, P_N\mu_1)\ ,$$

uniformly for all $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$. More precisely, for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \ge \bar{N}$ and $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have

$$c^\star(\theta,\mathcal{T})\mathbb{W}_2^2(\mu_0,\mu_1) \le \mathcal{W}_N^2(P_N\mu_0, P_N\mu_1) + \varepsilon\ . \tag{2.32}$$

Proof. Fix $\varepsilon > 0$. Applying Proposition 2.5.3 to an approximate $\mathcal{W}_N$-geodesic between $P_N\mu_0$ and $P_N\mu_1$, we infer that there exists a curve $(m_t, J_t)_t$ satisfying the bounds

$$\mathbb{W}_2(\iota_N m_i, \iota_N P_N \mu_i) \le \varepsilon + \frac{C}{N} \quad \text{for } i = 0,1\ , \tag{2.33}$$

$$\int_0^1 \mathcal{A}_N(m_t, J_t)\,\mathrm{d}t \le \mathcal{W}_N^2(P_N\mu_0, P_N\mu_1) + \varepsilon\ , \tag{2.34}$$

as well as the regularity properties (2.27a)–(2.27e).

For brevity we write

$$\mathcal{A}_N(m,J) = \frac{1}{N}\sum_{n=0}^{N-1} \mathcal{A}_N^n(m,J)\ ,$$

where

$$\mathcal{A}_N^n(m,J) = \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1}\bigg(N\frac{m(n;k)}{\pi_k}, N\frac{m(n;k+1)}{\pi_k}, J(n;k,k+1)\bigg) .$$

We set

$$\widehat{m}_t(n) := \sum_{k=0}^{K-1} m_t(n;k) , \quad \widehat{J}_t(n) := J_t(n;-1,0) ,$$

and define $\alpha_t : \{0,\ldots,N-1\} \times \{0,\ldots,K\} \to \mathbb{R}$ by

$$\alpha_t(n;k) = \frac{m_t(n;\sigma(k))}{\widehat{m}_t(n)} \quad \text{for } k = 0,\ldots,K ,$$

where $\sigma(k) = k$ for $k = 0,\ldots,K-1$, and $\sigma(K) = 0$. Here it is important to note that $\alpha_t(n;K) \neq \alpha_t(n+1;0)$. Observe that, for $k = 0,\ldots,K-1$,

$$\frac{1}{\theta_{k,k+1}\Big(\frac{\alpha_t(n;k)}{\pi_k}, \frac{\alpha_t(n;k+1)}{\pi_{k+1}}\Big)}\frac{|\widehat{J}_t(n)|^2}{N\widehat{m}_t(n)} = f_{k,k+1}\bigg(N\frac{m_t(n;\sigma(k))}{\pi_k}, N\frac{m_t(n;\sigma(k+1))}{\pi_{k+1}}, \widehat{J}_t(n)\bigg) .$$

Note that, for any n and k,

$$\frac{c_0^{-1}}{\max_\ell \pi_\ell} \le N\frac{m_t(n;k)}{\pi_k} \le \frac{c_1}{\min_\ell \pi_\ell} \quad \text{and} \quad |J_t(n;k,k+1)| \le c_4 .$$

Therefore, since the functions $f_{k,k+1}$ are Lipschitz on the set $[\frac{c_0^{-1}}{\max_\ell \pi_\ell}, \frac{c_1}{\min_\ell \pi_\ell}]^2 \times [-c_4, c_4]$, it follows that, for $k = 0,\ldots,K-1$,

$$\begin{aligned}
&\bigg|f_{k,k+1}\bigg(N\frac{m_t(n;k)}{\pi_k}, N\frac{m_t(n;k+1)}{\pi_{k+1}}, J_t(n;k,k+1)\bigg) - \frac{1}{\theta_{k,k+1}\Big(\frac{\alpha_t(n;k)}{\pi_k}, \frac{\alpha_t(n;k+1)}{\pi_{k+1}}\Big)}\frac{|\widehat{J}_t(n)|^2}{N\widehat{m}_t(n)}\bigg| \\
&\le [f_{k,k+1}]_{\mathrm{Lip}}\bigg(\frac{N}{\pi_{k+1}}\Big|m_t(n;k+1) - m_t(n;\sigma(k+1))\Big| + \Big|J_t(n;k,k+1) - \widehat{J}_t(n)\Big|\bigg) \\
&\le \frac{[f_{k,k+1}]_{\mathrm{Lip}}}{N}\bigg(\frac{c_3}{\pi_{k+1}} + Kc_5\bigg) =: \frac{C}{N} ,
\end{aligned} \tag{2.35}$$

for some $C < \infty$ depending on ε (through c_0, ..., c_5) and on $\mathcal{T}$.

Since $\sum_{k=0}^{K-1} \alpha_t(n;k) = 1$, the sequence $\{\alpha_t(n;k)\}_{k=0}^{K-1}$ is, for any n, a competitor for the cell problem (2.5). Taking into account that $\alpha_t(n;0) = \alpha_t(n;K)$, it follows from (2.35) and the definition (2.5) of $c^\star(\theta,\mathcal{T})$ that

$$\begin{aligned}
\mathcal{A}_N^n(m_t,J_t) &\ge \frac{|\widehat{J}_t(n)|^2}{N\widehat{m}_t(n)} \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\Big(\frac{\alpha_t(n;k)}{\pi_k}, \frac{\alpha_t(n;k+1)}{\pi_{k+1}}\Big)} - \frac{C}{N} \\
&\ge c^\star(\theta,\mathcal{T})\frac{|\widehat{J}_t(n)|^2}{N\widehat{m}_t(n)} - \frac{C}{N} .
\end{aligned} \tag{2.36}$$

At the continuous level, we define a curve of measures $(\mu_t^N)_t$ with piecewise constant densities, and a vector field j_t^N by piecewise affine interpolation of $\widehat{J}_t^N$; more precisely,

$$\begin{aligned}
\mu_t^N &= \sum_{n=0}^{N-1} \widehat{m}_t \mathcal{U}_{\widehat{A}_n} , \\
j_t^N(x) &= \sum_{n=0}^{N-1} \chi_{\widehat{A}_n}(x)\Big[(n+1-Nx)\widehat{J}_t(n) + (Nx-n)\widehat{J}_t(n+1)\Big] .
\end{aligned}$$

As before, $\mathcal{U}_{\widehat{A}_n}$ denotes the normalised Lebesgue measure on $\widehat{A}_n := [\frac{n}{N}, \frac{n+1}{N})$.

We observe that the density ρ_t^N of μ_t^N satisfies

$$\partial_t \rho_t^N(x) = N \sum_{k=0}^{K-1} \partial_t \widehat{m}_t(n;k) = N\Big(\widehat{J}_t(n) - \widehat{J}_t(n+1)\Big) = -\partial_x j_t^N(x)$$

for any $x \in \left(\frac{n}{N}, \frac{n+1}{N}\right)$, which implies that $\left(\mu_t^N, j_t^N\right)_t$ solves the continuity equation.

To estimate the continuous energy, we find

$$\begin{aligned}
\mathbb{A}(\mu_t^N, j_t^N) &= \frac{1}{N}\sum_{n=0}^{N-1} \frac{1}{\widehat{m}_t(n)} \int_{\frac{n}{N}}^{\frac{n+1}{N}} \Big[(n+1-Nx)\widehat{J}_t(n) + (Nx-n)\widehat{J}_t(n+1)\Big]^2 \,\mathrm{d}x \\
&\le \frac{1}{N}\sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2 + \widehat{J}_t(n+1)^2}{2N\widehat{m}_t(n)} \\
&= \frac{1}{N}\sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{\theta_{\mathsf{h}}(N\widehat{m}_t(n), N\widehat{m}_t(n+1))}
\end{aligned}$$

where $\theta_{\mathsf{h}}(a,b) = \frac{2ab}{a+b}$ denotes the harmonic mean. Note that (2.27c) implies the Lipschitz bound

$$|N\widehat{m}_t(n) - N\widehat{m}_t(n-1)| \le \frac{Kc_3}{N} .$$

and (2.27a) yields a lower bound on the density: $N\widehat{m}_t(n) \ge Kc_0^{-1}$. Furthermore, (2.27d) yields the estimate $|\widehat{J}_t(n)| \le c_4$. Thus, in view of the identity $\frac{1}{\theta_{\mathsf{h}}(a,b)} = \frac{1}{a} + \frac{a-b}{2ab}$ we obtain

$$\mathbb{A}(\mu_t^N, j_t^N) \le \left(\frac{1}{N}\sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{N\widehat{m}_t(n)}\right) + \frac{C}{N} , \tag{2.37}$$

with $C < \infty$ depending on ε (through the c_i's) and on $\mathcal{T}$.

Putting things together, it follows from (2.36), (2.37) and (2.34) that

$$\begin{aligned}
c^\star(\theta,\mathcal{T})\mathbb{W}_2^2(\mu_0^N,\mu_1^N) &\le c^\star(\theta,\mathcal{T}) \int_0^1 \mathbb{A}(\mu_t^N, j_t^N)\,\mathrm{d}t \\
&\le c^\star(\theta,\mathcal{T}) \int_0^1 \frac{1}{N}\sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{N\widehat{m}_t(n)}\,\mathrm{d}t + \frac{C}{N} \\
&\le \int_0^1 \frac{1}{N}\sum_{n=0}^{N-1} \mathcal{A}_N^n(m_t, J_t)\,\mathrm{d}t + \frac{C}{N} \\
&= \int_0^1 \mathcal{A}_N(m_t, J_t)\,\mathrm{d}t + \frac{C}{N} \\
&\le \mathcal{W}_N^2(P_N\mu_0, P_N\mu_1) + \varepsilon + \frac{C}{N} .
\end{aligned} \tag{2.38}$$

Finally we note that, for $i = 0, 1$, (2.33) yields

$$\begin{aligned}
\mathbb{W}_2(\mu_i, \mu_i^N) &\le \mathbb{W}_2(\mu_i, \iota_N P_N \mu_i) + \mathbb{W}_2(\iota_N P_N\mu_i, \iota_N m_i) + \mathbb{W}_2(\iota_N m_i, \mu_i^N) \\
&\le \frac{1}{N} + \left(\varepsilon + \frac{C}{N}\right) + \frac{1}{N} \\
&\le \varepsilon + \frac{C}{N} ,
\end{aligned}$$

which implies that

$$\mathbb{W}_2(\mu_0, \mu_1) \leq \mathbb{W}_2(\mu_0^N, \mu_1^N) + 2\Big(\varepsilon + \frac{C}{N}\Big) . \tag{2.39}$$

Combining (2.38) and (2.39) we obtain the desired result. □

2.6 Proof of the upper bound

In this section we present the proof of the upper bound for $\mathcal{W}_N$. The idea of the proof of the upper bound is to start from optimal curves of measures at the continuous level, and to introduce the optimal oscillation in their discretations, as determined by the formula for the effective mobility (2.5).

Let $\alpha^\star = \{\alpha_k^\star\}_{k=0}^{K-1}$ be an optimiser in (2.5), and define $P_N^\star : \mathcal{P}(\mathcal{S}^1) \to \mathcal{P}(\mathcal{T}_N)$ by

$$\big(P_N^\star \mu\big)(n;k) := \alpha_k^\star \mu\big(\widehat{A}_n\big) , \qquad \text{where} \quad \widehat{A}_n := \Big[\frac{n}{N}, \frac{n+1}{N}\Big) , \tag{2.40}$$

as before. Slightly abusing notation, we also define $P_N^\star : C(\mathcal{S}^1;\mathbb{R}) \to \mathcal{V}(\mathcal{T}_N)$ by

$$\big(P_N^\star j\big)(n;k,k+1) := \Bigg(\sum_{\ell=k+1}^{K-1} \alpha_\ell^\star\Bigg) j\big(\tfrac{n}{N}\big) + \Bigg(\sum_{\ell=0}^{k} \alpha_\ell^\star\Bigg) j\big(\tfrac{n+1}{N}\big) .$$

Since $\sum_{k=0}^{K-1} \alpha_k^\star = 1$, the right-hand side is a convex combination of $j\big(\frac{n}{N}\big)$ and $j\big(\frac{n+1}{N}\big)$.

Proposition 2.6.1 (Discretisation of the continuity equation)**.** *Let $(\mu_t)_{t\in[0,1]}$ be a Borel family of probability measures, and let $(j_t)_{t\in[0,1]}$ be a Borel family of continuous functions satisfying the continuity equation $\partial_t \mu + \partial_x j = 0$ on $\mathcal{S}^1$. Then the pair $(m_t, J_t)_{t\in[0,1]}$ defined by*

$$m_t := P_N^\star \mu_t , \quad J_t := P_N^\star j_t ,$$

solves the continuity equation on $\mathcal{T}_N$.

Proof. As $(\mu_t, j_t)_t$ satisfies the continuity equation, we have

$$\int_0^1 \Big(\int_{\mathcal{S}^1} \partial_t \phi_t(x) \,\mathrm{d}\mu_t(x) + \int_{\mathcal{S}^1} \partial_x \phi_t(x) j_t(x) \,\mathrm{d}x \Big) \,\mathrm{d}t = \int_{\mathcal{S}^1} \phi_1(x) \,\mathrm{d}\mu_1(x) - \int_{\mathcal{S}^1} \phi_0(x) \,\mathrm{d}\mu_0(x)$$

for any smooth function $\phi : [0,1] \times \mathcal{S}^1 \to \mathbb{R}$.

Let $\psi : [0,1] \to \mathbb{R}$ be smooth, and define $\eta^\varepsilon : \mathcal{S}^1 \to \mathbb{R}$ by $\eta^\varepsilon = \chi_{\widehat{A}_n} * \xi^\varepsilon$ for a smooth mollifier ξ^ε supported in an ε-neighbourhood of 0. Set $\phi_t^\varepsilon(x) = \psi(t)\eta^\varepsilon(x)$. Applying the weak formulation of the continuity equation to ϕ^ε, and passing to the limit $\varepsilon \downarrow 0$, we obtain

$$\int_0^1 \psi'(t) \mu_t(\widehat{A}_n) \,\mathrm{d}t + \int_0^1 \psi(t) \Big(j_t\big(\tfrac{n+1}{N}\big) - j_t\big(\tfrac{n}{N}\big) \Big) \,\mathrm{d}t = \psi(1)\mu_1(\widehat{A}_n) - \psi(0)\mu_0(\widehat{A}_n) .$$

Multiplying this identity by $\alpha_k^\star$, and using the fact that

$$\alpha_k^\star \Big(j_t\big(\tfrac{n+1}{N}\big) - j_t\big(\tfrac{n}{N}\big) \Big) = J_t(n;k,k+1) - J_t(n;k-1,k) ,$$

we obtain

$$\int_0^1 \psi'(t) m_t(n;k)\,\mathrm{d}t + \int_0^1 \psi(t)\Big(J_t(n;k,k+1) - J_t(n;k-1,k)\Big)\,\mathrm{d}t \\ = \psi(1) m_1(n;k) - \psi(0) m_0(n;k) \ ,$$

which is the distributional form of the discrete continuity equation (2.8). □

Lemma 2.6.2 (Consistency). *For all $\mu \in \mathcal{P}(\mathcal{S}^1)$ we have*

$$\mathbb{W}_2(\mu, \iota_N P_N^\star \mu) \leq \frac{1}{N} \ .$$

Proof. This readily follows from the definitions; see [GKM20, Lemma 3.2] for a similar result. □

The following proposition is the key result of this section. It proves the required upper bound for the discrete energy under suitable regularity conditions.

For $\delta > 0$, it will be useful to write

$$\mathcal{P}_\delta(\mathcal{S}^1) := \left\{ \mu = \rho\,\mathrm{d}x \in \mathcal{P}(\mathcal{S}^1) \ : \ \rho \geq \delta > 0, \quad \mathrm{Lip}(\rho) \leq \frac{1}{\delta} \right\} .$$

Proposition 2.6.3 (Discrete energy upper bound). *Let $\delta > 0$. There exists $C < \infty$ and $\bar{N} \in \mathbb{N}$ (depending on δ), such that for any $N \geq \bar{N}$, all $\mu \in \mathcal{P}_\delta(\mathcal{S}^1)$, and all vector fields $j : \mathcal{S}^1 \to \mathbb{R}$ with $\|j\|_{L^\infty} + \mathrm{Lip}(j) \leq \delta^{-1}$, we have*

$$\mathcal{A}_N(P_N^\star \mu, P_N^\star j) \leq c^\star(\theta, \mathcal{T}) \mathbb{A}(\mu, j) + \frac{C}{N} \ .$$

Proof. Write $m = P_N^\star \mu$ and $J = P_N^\star j$ for brevity, and set $\bar{\rho}(n) := N\mu(\widehat{A}_n)$. Recall that

$$\mathcal{A}_N(m, J) = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{A}_N^n(m, J)$$

where

$$\mathcal{A}_N^n(m,J) = \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1}\left(N\frac{m(n;k)}{\pi_k}, N\frac{m(n;k+1)}{\pi_{k+1}}, J(n;k,k+1) \right) \\ = \frac{1}{\bar{\rho}(n)} \sum_{k=0}^{K-1} d_{k,k+1} \frac{(J(n;k,k+1))^2}{\theta_{k,k+1}\left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \ ,$$

with $\alpha(n;k) := \frac{m(n;k)}{\mu(\widehat{A}_n)}$ for $k = 0, \ldots, K$. Note that $\alpha(n;k) = \alpha_k^\star$ for $k = 0, \ldots, K-1$, but

$$\alpha(n;K) = \frac{m(n+1;0)}{\mu(\widehat{A}_n)} = \alpha_0^\star \frac{\mu(\widehat{A}_{n+1})}{\mu(\widehat{A}_n)} \ ,$$

which is not necessarily equal to $\alpha_K^\star = \alpha_0^\star$. Therefore, $\{\alpha(n;k)\}_{k=0}^K$ is not necessarily an admissible competitor in (2.5). Write $\mathrm{d}\mu(x) = \rho(x)\,\mathrm{d}x$. We claim that the following estimates hold for sufficiently large N, with $C < \infty$ depending only on θ and $\mathcal{T}$:

$$|J(n;k,k+1) - j(x)| \leq \frac{\mathrm{Lip}(j)}{N} \,, \tag{2.41}$$

$$\left| c^\star(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\Big(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\Big)} \right| \leq C \frac{\mathrm{Lip}(\rho)}{\inf \rho} \frac{1}{N} \,, \tag{2.42}$$

the first one being valid for any $x \in \left[\frac{n}{N}, \frac{n+1}{N}\right]$ and $k = 0, \ldots, K-1$. Indeed, writing $\lambda_k = \sum_{\ell=k+1}^{K-1} \alpha_\ell^\star$, we obtain

$$|J(n;k,k+1) - j(x)| \leq \lambda_k |j(\tfrac{n}{N}) - j(x)| + (1-\lambda_k)|j(\tfrac{n+1}{N}) - j(x)| \leq \frac{\mathrm{Lip}(j)}{N} \,,$$

which proves (2.41). Furthermore,

$$\begin{aligned} E &:= \left| c^\star(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\Big(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\Big)} \right| \\ &= d_{K-1,K} \left| \frac{1}{\theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha_K^\star}{\pi_K}\Big)} - \frac{1}{\theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha(n;K)}{\pi_K}\Big)} \right| . \end{aligned}$$

Note that

$$|\alpha_K^\star - \alpha(n;K)| = \alpha_K^\star \left| 1 - \frac{\mu(\widehat{A}_{n+1})}{\mu(\widehat{A}_n)} \right| = \alpha_K^\star \frac{|\mu(\widehat{A}_n) - \mu(\widehat{A}_{n+1})|}{\mu(\widehat{A}_n)} \leq \frac{\alpha_K^\star}{N} \frac{\mathrm{Lip}(\rho)}{\inf \rho} \,.$$

If $\alpha_K^\star = 0$, we infer that $E = 0$, in which case the claim is proved. If $\alpha_K^\star > 0$, we observe that the latter inequality yields

$$\alpha(n;K) \geq \frac{\alpha_K^\star}{2} \tag{2.43}$$

for N sufficiently large (depending on δ). Since $\theta_{K-1,K}$ is concave, we have for any $a \geq 0$ and $0 < b \leq y_1 < y_2$,

$$\frac{\theta_{K-1,K}\big(a, y_2\big) - \theta_{K-1,K}\big(a, y_1\big)}{y_2 - y_1} \leq \frac{\theta_{K-1,K}\big(a, b\big) - \theta_{K-1,K}\big(a, 0\big)}{b} \,,$$

thus $\theta_{K-1,K}(a, \cdot)$ is Lipschitz on $[b, \infty)$. Let $L < \infty$ denote the Lipschitz constant of $\theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \cdot\Big)$ on $\Big[\frac{\alpha_K^\star}{2}, \infty\Big)$. For N sufficiently large we obtain

$$\begin{aligned} E &\leq \frac{1}{\Big(\theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha_K^\star}{2\pi_K}\Big)\Big)^2} \left| \theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha_K^\star}{\pi_K}\Big) - \theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha(n;K)}{\pi_K}\Big) \right| \\ &\leq \frac{L}{\Big(\theta_{K-1,K}\Big(\frac{\alpha_{K-1}^\star}{\pi_{K-1}}, \frac{\alpha_K^\star}{2\pi_K}\Big)\Big)^2} \frac{\alpha_K^\star}{\pi_K} \frac{\mathrm{Lip}(\rho)}{\inf \rho} \frac{1}{N} \end{aligned}$$

which yields our claim (2.42).

Taking into account that $\bar{\rho}(n) \geq \delta$ and $\|j\|_\infty \leq \delta^{-1}$, it follows from (2.41) and a twofold application of (2.42) that

$$\begin{aligned}
\left| \mathcal{A}_N^n(m, J) - c^\star(\theta, \mathcal{T}) \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} \right| &\leq \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} \left| c^\star(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \right| \\
&\quad + \frac{1}{\bar{\rho}(n)} \sum_{k=0}^{K-1} d_{k,k+1} \frac{\left| J^2(n; k, k+1) - j^2\left(\frac{n}{N}\right) \right|}{\theta_{k,k+1}\left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \\
&\leq \frac{C}{N} + \frac{C}{N} \sum_k \frac{d_{k,k+1}}{\theta_{k,k+1}\left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \\
&\leq \frac{C}{N} ,
\end{aligned}$$

where $C < \infty$ depends on $\mathcal{T}$, θ, and δ. Consequently,

$$\mathcal{A}_N(m, J) \leq \frac{c^\star(\theta, \mathcal{T})}{N} \sum_{n=0}^{N-1} \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} + \frac{C}{N} .$$

By the arithmetic-harmonic mean inequality,

$$\frac{\left| j\left(\frac{n}{N}\right) \right|^2}{\bar{\rho}(n)} \leq N \left| j\left(\frac{n}{N}\right) \right|^2 \int_{\widehat{A}_n} \frac{1}{\rho(x)} \, \mathrm{d}x \leq N \int_{\widehat{A}_n} \frac{|j(x)|^2}{\rho(x)} \, \mathrm{d}x + \frac{C}{N} .$$

We infer that

$$\mathcal{A}_N(m, J) \leq c^\star(\theta, \mathcal{T}) \mathbb{A}(\mu, j) + \frac{C}{N} ,$$

which completes the proof. □

The previous result shows that the sought upper bound can be achieved once we assume some regularity of the solution of the continuity equation. Therefore in order to conclude the proof of Theorem 2.1.1 we seek once again for a regularization procedure.

The following result collects some well-known properties of the heat semigroup $(H_s)_{s \geq 0}$ on $\mathcal{P}(\mathcal{S}^1)$.

Lemma 2.6.4 (Regularisation by heat flow)**.** *Let $s > 0$. There exists a constant $\delta > 0$ such that for any $\mu \in \mathcal{P}(\mathcal{S}^1)$ we have $H_s\mu \in \mathcal{P}_\delta(\mathcal{S}^1)$. Moreover, $\mathbb{W}_2(\mu, H_s\mu) \leq \sqrt{2s}$.*

Proof. See, e.g., [GM13, Proposition 2.9] for a proof of these well-known facts. □

We continue with a well-known regularisation result. For the convenience of the reader we include a simple proof.

Lemma 2.6.5 (Smooth approximate action minimisers)**.** *Let $\delta > 0$ and let $\varepsilon > 0$. Then there exists $\tilde\delta \in (0,\delta)$, such that the following assertion holds: for any $\mu_0, \mu_1 \in \mathcal{P}_\delta(\mathcal{S}^1)$ there exists a curve $(\mu_t, j_t) \in \mathbb{CE}(\mu_0,\mu_1)$ with $\mu_t \in \mathcal{P}_{\tilde\delta}(\mathcal{S}^1)$ and $\|j_t\|_{L^\infty} + \mathrm{Lip}(j_t) \le \tilde\delta^{-1}$ for any $t \in (0,1)$, such that*

$$\int_0^1 \mathbb{A}(\mu_t, j_t)\,\mathrm{d}t \le \mathbb{W}_2^2(\mu_0,\mu_1) + \varepsilon\ .$$

Proof. Let $(\mu_t)_{t\in[0,1]}$ be a $\mathbb{W}_2$-geodesic connecting μ_0 and μ_1, and let $(j_t)_{t\in[0,1]}$ be a vector field such that

$$\int_0^1 \mathbb{A}(\mu_t, j_t)\,\mathrm{d}t = \mathbb{W}_2^2(\mu_0,\mu_1)\ .$$

The idea of the proof is to regularise μ_0 and μ_1 by applying the heat flow for a short time $s > 0$, and then to connect the regularised measures $H_s\mu_0$ and $H_s\mu_1$ using the natural candidate $(H_s\mu_t)_{t\in[0,1]}$.

Firstly, for $i = 0, 1$ and $s > 0$, set $\gamma_t^{i,s} = H_{st}\mu_i$ for $t \in [0,1]$, and let $\rho_t^{i,s}$ be the density of $\gamma_t^{i,s}$ with respect to the Haar measure. Then: $\partial_t\gamma_t^{i,s} = s\partial_x^2\gamma_t^{i,s}$, thus the continuity equation $\partial_t\rho_t^{i,s} + \partial_x k_t^{i,s} = 0$ holds with $k_t^{i,s} = -s\partial_x\rho_t^{i,s}$. Using the contractivity of the Fisher information under the heat flow, and the fact that $\mu_i \in \mathcal{P}_\delta(\mathcal{S}^1)$, we obtain

$$\int_0^1 \mathbb{A}(\gamma_t^{i,s}, k_t^{i,s})\,\mathrm{d}t = s^2 \int_0^1 \int_{\mathcal{S}^1} \frac{|\partial_x\rho_t^{i,s}(x)|^2}{\rho_t^s(x)}\,\mathrm{d}x\,\mathrm{d}t \le s^2 \int_{\mathcal{S}^1} \frac{|\partial_x\rho_0^{i,s}(x)|^2}{\rho_0^s(x)}\,\mathrm{d}x \le \frac{s^2}{\delta^3}\ .$$

Secondly, for any $s > 0$, we note that $(H_s\mu_t, H_sj_t)_{t\in[0,1]}$ solves the continuity equation, and, by the joint convexity of $\mathbb{A}$ and the fact that H_s is given by a convolution kernel,

$$\mathbb{A}(H_s\mu_t, H_sj_t) \le \mathbb{A}(\mu_t, j_t)\ .$$

Fix $\tau \in (0,\frac14)$, and consider now the curve $(\tilde\mu_t, \tilde j_t)_{t\in[0,1]} \in \mathbb{CE}(\mu_0,\mu_1)$ defined by

$$\tilde\mu_t := \begin{cases} H_{ts/\tau}\mu_0 \\ H_s\mu_{(t-\tau)/(1-2\tau)} \\ H_{(1-t)s/\tau}\mu_1 \end{cases}, \qquad \tilde j_t := \begin{cases} \frac{1}{\tau}k_{t/\tau}^{0,s} & t \in (0,\tau) \\ \frac{1}{1-2\tau}H_sj_{(t-\tau)/(1-2\tau)} & t \in (\tau, 1-\tau) \\ -\frac{1}{\tau}k_{1-t/\tau}^{1,s} & t \in (1-\tau, 1) \end{cases},$$

It follows from the bounds above, using the fact that $\frac{1}{1-2\tau} \le 1 + 4\tau$ and $\mathbb{W}_2^2 \le \frac14$, that

$$\begin{aligned}\int_0^1 \mathbb{A}(\tilde\mu_t, \tilde j_t)\,\mathrm{d}t &= \int_0^1 \frac{\mathbb{A}(\gamma_t^{0,s}, k_t^{0,s})}{\tau} + \frac{\mathbb{A}(H_s\mu_t, H_sj_t)}{1-2\tau} + \frac{\mathbb{A}(\gamma_t^{1,s}, k_t^{1,s})}{\tau}\,\mathrm{d}t \\ &\le \frac{s^2}{\delta^3\tau} + \frac{\mathbb{W}_2^2(\mu_0,\mu_1)}{1-2\tau} + \frac{s^2}{\delta^3\tau} \\ &\le \frac{s^2}{\delta^3\tau} + (\mathbb{W}_2^2(\mu_0,\mu_1) + \tau) + \frac{s^2}{\delta^3\tau}\ .\end{aligned}$$

Let $\varepsilon > 0$, and choose $\tau = \varepsilon/2$, and $s^2 = \delta^3\tau\varepsilon/4$. Then: $\int_0^1 \mathbb{A}(\tilde\mu_t, \tilde j_t)\,\mathrm{d}t \le \mathbb{W}_2^2(\mu_0,\mu_1) + \varepsilon$.

Moreover, by Lemma 2.6.4, $\tilde\mu_t$ belongs to $\mathcal{P}_{\tilde\delta}(\mathcal{S}^1)$ for some $\tilde\delta > 0$ depending on δ and s. Furthermore,

$$\|H_sj_t\|_{L^\infty(\mathcal{S}^1)} + \|\partial_xH_sj_t\|_{L^\infty(\mathcal{S}^1)} \le C(s)\|j_t\|_{L^1(\mathcal{S}^1)} \le C(s)\sqrt{\mathbb{A}(\mu_t,j_t)} = C(s)\mathbb{W}_2(\mu_0,\mu_1)\ ,$$

where the last inequality follows from the Cauchy-Schwarz inequality. □

We are now ready to prove the upper bound in Theorem 2.1.1.

Theorem 2.6.6 (Upper bound for $\mathcal{W}_N$)**.** *For any mesh $\mathcal{T}$ and any family of admissible means $\{\theta_{k,k+1}\}_k$ we have*

$$\limsup_{N\to\infty} \mathcal{W}_N^2(P_N\mu_0, P_N\mu_1) \le c^\star(\theta,\mathcal{T})\mathbb{W}_2^2(\mu_0,\mu_1)\ ,$$

uniformly for all $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$. More precisely, for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \ge \bar{N}$ and $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have

$$\mathcal{W}_N^2(P_N\mu_0, P_N\mu_1) \le c^\star(\theta,\mathcal{T})\mathbb{W}_2^2(\mu_0,\mu_1) + \varepsilon\ . \tag{2.44}$$

Proof. Let $\mu_0,\mu_1 \in \mathcal{P}(\mathcal{S}^1)$ and $\varepsilon \in (0,1]$. By Lemma 2.6.4 there exist $s \ge 0$ and $\delta > 0$ such that $\tilde{\mu}_i := H_s\mu_i$ belongs to $\mathcal{P}_\delta(\mathcal{S}^1)$, and

$$\mathbb{W}_2(\mu_i,\tilde{\mu}_i) \le \varepsilon \quad \text{for } i = 0,1\ .$$

Using that $\mathbb{W}_2 \le \frac{1}{2}$, it follows that

$$\mathbb{W}_2^2(\tilde{\mu}_0,\tilde{\mu}_1) \le \mathbb{W}_2^2(\mu_0,\mu_1) + 2\varepsilon\ . \tag{2.45}$$

Lemma 2.6.5 yields $\tilde{\delta} \in (0,\delta)$ and a curve $(\tilde{\mu}_t,\tilde{j}_t)_t \in \mathbb{CE}_{\tilde{\delta}}(\tilde{\mu}_0,\tilde{\mu}_1)$ such that $\tilde{\mu}_t \in \mathcal{P}_{\tilde{\delta}}(\mathcal{S}^1)$ and $\|\tilde{j}_t\|_{L^\infty} + \mathrm{Lip}(\tilde{j}_t) \le \tilde{\delta}^{-1}$ for any $t \in (0,1)$, and

$$\int_0^1 \mathbb{A}(\tilde{\mu}_t,\tilde{j}_t)\,\mathrm{d}t \le \mathbb{W}_2^2(\tilde{\mu}_0,\tilde{\mu}_1) + \varepsilon\ . \tag{2.46}$$

Set $\tilde{m}_t^N := P_N^\star\tilde{\mu}_t$ and $\tilde{J}_t^N = P_N^\star\tilde{j}_t$. By Proposition 2.6.3 there exist $\bar{N} \in \mathbb{N}$ and $C_1 < \infty$ depending on ε (through $\tilde{\delta}$) such that for $N \ge \bar{N}$,

$$\mathcal{W}_N^2(\tilde{m}_0^N,\tilde{m}_1^N) \le \int_0^1 \mathcal{A}(\tilde{m}_t^N,\tilde{J}_t^N)\,\mathrm{d}t \le c^\star(\theta,r)\int_0^1 \mathbb{A}(\tilde{\mu}_t,\tilde{j}_t)\,\mathrm{d}t + \frac{C_1}{N}\ . \tag{2.47}$$

Set $m_i^N := P_N^\star\mu_i$ for $i = 0,1$. By Proposition 2.2.7, Lemma 2.6.2, and Lemma 2.6.4, there exists $C_2 < \infty$ depending only on $\mathcal{T}$ (possibly varying from line to line) such that

$$\begin{aligned}
\mathcal{W}_N(m_i^N,\tilde{m}_i^N) &= \mathcal{W}_N(P_N^\star\mu_i, P_N^\star H_s\mu_i)\\
&\le C_2\Big(\mathbb{W}_2(\iota_N P_N^\star\mu_i, \iota_N P_N^\star H_s\mu_i) + \frac{1}{N}\Big)\\
&\le C_2\Big(\mathbb{W}_2(\mu_i, H_s\mu_i) + \frac{1}{N}\Big)\\
&\le C_2\Big(\sqrt{s} + \frac{1}{N}\Big)\ .
\end{aligned}$$

Thus, the triangle inequality yields

$$\mathcal{W}_N(m_0^N,m_1^N) \le \mathcal{W}_N(\tilde{m}_0^N,\tilde{m}_1^N) + C_2\Big(\sqrt{s} + \frac{1}{N}\Big)\ ,$$

and by another application of Proposition 2.2.7,

$$\begin{aligned}
&\mathcal{W}_N^2(m_0^N,m_1^N) - \mathcal{W}_N^2(\tilde{m}_0^N,\tilde{m}_1^N)\\
&\quad\le \Big(\mathcal{W}_N(m_0^N,m_1^N) + \mathcal{W}_N(\tilde{m}_0^N,\tilde{m}_1^N)\Big)C_2\Big(\sqrt{s} + \frac{1}{N}\Big) \le C_2\Big(\sqrt{s} + \frac{1}{N}\Big)\ .
\end{aligned} \tag{2.48}$$

Combining (2.45), (2.46), (2.47), and (2.48), we obtain

$$\mathcal{W}_N^2(m_0^N,m_1^N) \le c^\star(\theta,r)\mathbb{W}_2^2(\mu_0,\mu_1) + 3\varepsilon + \frac{C_1}{N} + C_2\Big(\sqrt{s} + \frac{1}{N}\Big)\ .$$

Choosing s small enough and N large enough depending on ε, we obtain the result. □

2.7 Proof of the Gromov–Hausdorff convergence

We conclude this work with the proof of the Gromov–Hausdorff convergence in Theorem 2.1.1. First we recall one of the equivalent definitions; cf. [BBI01] for more details.

Definition 2.7.1 (Gromov–Hausdorff convergence)**.** A sequence of compact metric spaces $\{\mathcal{X}_N, \mathrm{d}_N\}_N$ is said to converge in the sense of Gromov–Hausdorff to a compact metric space $(\mathcal{X}, \mathrm{d})$, if there exist maps $f_N : \mathcal{X} \to \mathcal{X}_N$ with the following properties:

- ε-isometry: for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and any $x, y \in \mathcal{X}$, we have:

$$|\,\mathrm{d}_N(f_N(x), f_N(y)) - \mathrm{d}(x,y)| \leq \varepsilon \ ;$$

- ε-surjectivity: for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and any $z \in \mathcal{X}_N$ there exists $x \in \mathcal{X}$ satisfying

$$\mathrm{d}_N(f_N(x), z) \leq \varepsilon \ .$$

Proof of Theorem 2.1.1. As the desired lower and upper bounds for the distance have been proved in Theorems 2.5.4 and 2.6.6, it remains to prove the Gromov–Hausdorff convergence. We will show that the conditions above hold with $f_N := P_N$.

Let $\varepsilon > 0$. It follows from Theorems 2.5.4 and 2.6.6 that there exists $\bar{N} \in \mathbb{N}$ such that, for any $N \geq \bar{N}$ and $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$,

$$\left|\mathcal{W}_N(P_N\mu_0, P_N\mu_1) - \sqrt{c^\star(\theta, \mathcal{T})}\,\mathbb{W}_2(\mu_0, \mu_1)\right| \leq \varepsilon \ .$$

This shows that the map P_N is ε-isometric.

The ε-surjectivity of P_N holds trivially, since it is even surjective. □

CHAPTER 3

Discrete-to-continuum limits of dynamical transport problems on periodic graphs

In this chapter we present a generalisation of the homogenisation result presented in Chapter 2 to arbitrary dimension and generic convex costs. This is the content of the work [GKMP21], obtained in collaboration with Peter Gladbach, Eva Kopfer, and Jan Maas.

We consider discrete dynamical transport problems on periodic graphs, obtained as minimisation of functionals defined on curves of measures, where the cost function is some given local, lower-semicontinuous, and convex function, with at least linear growth with respect to the momentum variable. This is a generalisation of the one-dimensional, quadratic problem studied in [GKMP20], corresponding to the spatial discretisation of the Wasserstein distance $\mathbb{W}_2$. We prove that the rescaled discrete energies converge to a homogenised continuous one which can be computed via a suitable *cell-formula*. Moreover, we prove that sequences of discrete measures with bounded mass and energy are compact in BV$(0,T) \times \mathbb{T}^d)$. Under the stronger assumption of superlinear growth for the cost function, we are able to improve the compactness result to W$^{1,1}((0,T) \times \mathbb{T}^d)$ and consequently show the convergence of the corresponding boundary value data problems. Several examples are discussed in detail, including finite-volume discretisation of optimal transport distances on $\mathbb{T}^d$, flow-based models, and limit behaviour of discrete Riemannian structures.

3.1 Introduction

In the past decades there has been intense research activity in the field of optimal transport, both in pure mathematics and in applied areas. In continuous settings, a central result in the field is the *Benamou–Brenier formula* [BB00], which establishes the equivalence of static and dynamical optimal transport. It asserts that the classical Monge–Kantorovich problem, in which a cost functional is minimised over couplings of given probability measures μ_0 and μ_1, is equivalent to a dynamical transport problem, in which an energy functional is minimised over all solutions to the continuity equation connecting μ_0 and μ_1.

In discrete settings, the equivalence between static and dynamical optimal transport breaks down, and it turns out that the dynamical formulation (introduced in [Maa11, Mie11]) is essential in applications to evolution equations, discrete Ricci curvature, and functional inequalities. Therefore, it is an important problem to analyse the discrete-to-continuum limit of dynamical optimal transport in various setting.

This limit passage turns out to be highly nontrivial. In fact, seemingly natural discretisations of the Benamou–Brenier formula do not necessarily converge to the expected limit, even in one-dimensional settings [GKMP20]. The main result in [GKM20] asserts that, for a sequence of meshes on a bounded convex domain in $\mathbb{R}^d$, an isotropy condition on the meshes is required to obtain the convergence of the discrete dynamical transport distances to $\mathbb{W}_2$. This is quite in constrast with the scaling behaviors of the corresponding gradient flows, where no additional symmetry on the meshes is required to ensure the convergence of the discrete evolutions to the continuous one [FMP20] (see also [DL15] for a one-dimensional analysis).

The goal of this paper is to investigate the large-scale behaviour of dynamical transport on graphs with a $\mathbb{Z}^d$-periodic structure. Our main contribution is a homogenisation result that describes the effective behaviour of the discrete problems in terms of a continuous optimal transport problem, in which the effective energy density depends non-trivially on the geometry of the discrete graph and the discrete energy density.

Main results

We give here an informal presentation of the main results of this paper, ignoring certain technicalities for the sake of readability. Precise formulations and a more general setting can be found from Section 3.2 onwards.

Dynamical optimal transport in the continuous setting

For $1 \le p < \infty$, let W_p be the Wasserstein–Kantorovich–Rubinstein distance between probability measures on a metric space (X, d): for $\mu^0, \mu^1 \in \mathcal{P}(X)$,

$$W_p(\mu^0, \mu^1) := \inf_\gamma \left\{ \int_{\mathbb{T}^d \times \mathbb{T}^d} \mathsf{d}(x,y)^p \,\mathrm{d}\gamma(x,y) \right\}^{1/p},$$

where $\Gamma(\mu^0, \mu^1)$ denotes the set of couplings of μ^0 and μ^1, i.e., all measures $\gamma \in \mathcal{P}(X \times X)$ with marginals μ^0 and μ^1. For $p > 1$, the Benamou–Brenier formula [BB00] (for a proof in full generality, see [AGS08]) provides an equivalent dynamical formulation, namely

$$W_p(\mu^0, \mu^1) = \inf_{(\rho, j)} \left\{ \int_0^1 \int_{\mathbb{T}^d} \frac{|j_t(x)|^p}{\rho_t^{p-1}(x)} \,\mathrm{d}x \,\mathrm{d}t \right\}^{1/p}, \tag{3.1}$$

where the infimum runs over all solutions (ρ, j) to the continuity equation $\partial_t \rho + \nabla \cdot j = 0$ with boundary conditions $\rho_0(x)\,\mathrm{d}x = \mu^0(\mathrm{d}x)$ and $\rho_1(x)\,\mathrm{d}x = \mu^1(\mathrm{d}x)$.

In this paper we consider general convex energy densities $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ under suitable (super)-linear growth conditions. (The Benamou–Brenier formula above corresponds to the special case $f(\rho, j) = \frac{|j|^p}{\rho^{p-1}}$). For a sufficiently regular curve $\boldsymbol{\mu} = (\mu_t)_{t \in (0,1)}$, we consider its action

$$\mathbb{A}(\boldsymbol{\mu}) = \inf_{\boldsymbol{\nu}} \left\{ \int_0^1 \int_{\mathbb{T}^d} f\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\mathscr{L}^d}, \frac{\mathrm{d}\nu_t}{\mathrm{d}\mathscr{L}^d} \right) \mathrm{d}x \,\mathrm{d}t \right\}. \tag{3.2}$$

Here, the infimum runs over all time-dependent vector-valued measures $\boldsymbol{\nu} = (\nu_t)_{t\in\mathcal{I}}$ satisfying the continuity equation $\partial_t \boldsymbol{\mu} + \nabla \cdot \boldsymbol{\nu} = 0$ in the sense of distributions.

The goal of this work is to study suitable discrete counterparts of these energies in the setting of $\mathbb{Z}^d$-periodic graphs.

Discrete dynamical optimal transport on $\mathbb{Z}^d$-periodic graphs

For an (undirected) graph $(\mathcal{X}, \mathcal{E})$ with finite set of nodes $\mathcal{X}$ and edges $\mathcal{E} \subset \mathcal{X} \times \mathcal{X}$, we consider the dynamical transport problem associated with a continuous curve $\mathbf{m} = (m_t)_{t\in(0,1)} \subset \mathcal{P}(\mathcal{X})$ given by

$$\mathcal{A}(\mathbf{m}) = \inf_{\mathbf{J}} \left\{ \int_0^1 \sum_{(x,y)\in\mathcal{E}} F_{xy}\Big(m_t(x), m_t(y), J_t(x,y)\Big)\,\mathrm{d}t \ : \ (\mathbf{m}, \mathbf{J}) \in \mathcal{CE} \right\}, \tag{3.3}$$

where $\mathcal{CE}$ denotes the class of all solutions to the discrete continuity equation on the graph, i.e., all curves of probability measures $\mathbf{m} : [0,1] \to \mathcal{P}(\mathcal{X})$ and all time-dependent discrete vector fields (i.e., anti-symmetric functions) $\mathbf{J} : [0,1] \to \mathbb{R}^{\mathcal{E}}$ satisfying in the sense of distributions the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(x) + \operatorname{div} J_t(x) = 0\,, \quad \operatorname{div} J_t(x) := \sum_{y:y\sim x} J_t(x,y) \quad \text{(discrete divergence)}\,,$$

where we write $y \sim x$ iff $(x,y) \in \mathcal{E}$

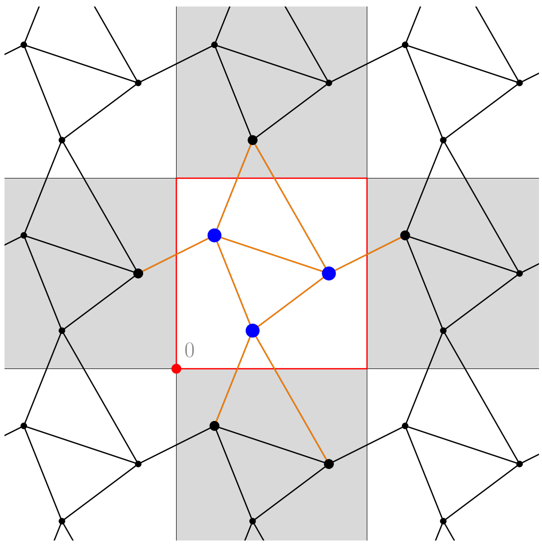

Figure 3.1: A fragment of a $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$. In red, the unitary cube $Q := [0,1]^d \subset \mathbb{R}^d$. In blue and in orange, respectively, $\mathcal{X}^Q$ and $\mathcal{E}^Q$.

The cost functions $F_{xy} : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ are assumed to be lower semicontinuous and convex.

In this work, we fix a $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$ embedded in $\mathbb{R}^d$, as in Figure 3.1. For any $\varepsilon > 0$ with $1/\varepsilon \in \mathbb{N}$, we consider the rescaled graphs $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$ defined by

$$\mathcal{X}_\varepsilon = \varepsilon\mathcal{X}/\mathbb{Z}^d \quad \text{and} \quad \mathcal{E}_\varepsilon = \varepsilon\mathcal{E}/\{(z,z) : z \in \mathbb{Z}^d\},$$

which is naturally embedded as a (finite) graph on the torus $\mathbb{T}^d$. If $\mathcal{CE}_\varepsilon$ denotes the solutions to the discrete continuity equation on the graph $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$, the rescaled transport cost is then given by

$$\mathcal{A}_\varepsilon(\mathbf{m}) = \inf_{\mathbf{J}} \left\{ \int_0^1 \sum_{(x,y)\in\mathcal{E}_\varepsilon} \varepsilon^d F_{xy}\left(\frac{m_t(x)}{\varepsilon^d}, \frac{m_t(y)}{\varepsilon^d}, \frac{J_t(x,y)}{\varepsilon^{d-1}} \right) \mathrm{d}t \ : \ (\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon \right\}.$$

The convergence result

Loosely speaking, our main result asserts that these discrete transport problems converge to a continuous transport problem with a homogenized cost function, as $\varepsilon \to 0$.

The limiting energy is of the form (3.2) with an effective energy density $f = f_{\text{hom}}$ that can computed via a *cell formula* which depends non-trivially on $(\mathcal{X}, \mathcal{E})$ and the discrete costs F_{xy}.

Precisely, if we set $\mathcal{X}^Q := \mathcal{X} \cap [0,1)^d$ and $\mathcal{E}^Q := \left\{(x,y) \in \mathcal{E} \ : \ x \in \mathcal{X}^Q\right\}$ (see Figure 3.1), then $f_{\text{hom}} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ is given by

$$f_{\text{hom}}(\rho, j) := \inf_{m,J} \left\{ \sum_{(x,y)\in\mathcal{E}^Q} F_{xy}\big(m(x), m(y), J(x,y)\big) \ : \ (m,J) \in \mathsf{Rep}(\rho, j) \right\}, \tag{3.4}$$

where $\mathsf{Rep}(\rho, j)$ denotes the set of *representatives* of $\rho \in \mathbb{R}_+$ and $j \in \mathbb{R}^d$, which is given by all $\mathbb{Z}^d$-periodic functions $m : \mathcal{X} \to \mathbb{R}_+$ and all $\mathbb{Z}^d$-periodic discrete, divergence-free vector fields (i.e., all anti-symmetric functions $J : \mathcal{E} \to \mathbb{R}$ with $\operatorname{div} J = 0$) satisfying

$$\sum_{x\in\mathcal{X}^Q} m(x) = \rho \quad \text{and} \quad \mathsf{Eff}(J) := \frac{1}{2} \sum_{(x,y)\in\mathcal{E}^Q} J(x,y)(y-x) = j\,. \tag{3.5}$$

In the special case where the discrete transport cost is associated to a Riemannian gradient-flow structure for a Markov chain (as in [Maa11, Mie11]), our result implies that the limiting metric is a 2-Wasserstein metric associated to a (not necessarily Riemannian) Finsler metric.

The rigorous formulation of our main result is given in terms of Γ-convergence for curves in the space of probability measures. We also establish compactness of bounded-energy sequences in Theorems 3.5.3 and 3.5.8. In the first compactness result, we assume at least linear growth of the discrete energies F_{xy} at infinity to show that limit curves lie in the space $\mathrm{BV}_{\mathrm{KR}}\big((0,1); \mathcal{M}_+(\mathbb{T}^d)\big)$ of curves with bounded variation, with convergence for almost every $t \in (0,1)$. In the second compactness result, if the costs F_{xy} has at least superlinear growth, then limit curves lie in the space $W^{1,1}_{\mathrm{KR}}((0,1); \mathcal{M}_+(\mathbb{T}^d))$ of absolutely continuous curves, with uniform convergence for $t \in [0,1]$. We refer to the Appendix for precise definitions of these spaces.

In the framework of superlinear energies, we are able to combine the convergence of the energies and the compactness result in $W^{1,1}$ to show the Γ-convergence of the associated boundary value problems, namely the variational problems

$$\mathcal{MA}_\varepsilon(m^0, m^1) := \inf_{\mathbf{m}} \left\{ \mathcal{A}_\varepsilon(\mathbf{m}) \ : \ m_0 = m^0, \ m_1 = m^1 \right\},$$

for every m^0,$m^1 \in \mathcal{P}(\mathcal{X})$.

In the final part of the paper, we analyse several examples, including finite-volume discretisations of $\mathbb{T}^d$, discussing the role of the geometry of the partitions in the approximation of the continuous Wasserstein distances (including the non-linear mobility case, as studied in [DNS09], [LM10]), in the same spirit of [GKM20].

We are in fact able to prove the convergence result for slightly more general discrete energies than the one in (3.3), we refer to Section 3.2.1 for the precise setting.

Organisation of the paper

Sections 3.2 and 3.3 contain all necessary definitions as well as the assumptions we use throughout the article, whereas Section 3.4 includes the definition of the homogenised continuous problem. In Section 3.5 we present the rigorous statements of our main results, including the Γ-convergence of the discrete energies to the effective homogenised limit and the compactness theorems for curves of bounded discrete energies. The proof of our main results can be found in Section 3.6 (compactness and convergence of the boundary value problems) and Sections 3.7 and 3.8 (Γ-convergence of $\mathcal{A}_\varepsilon$). Finally, in Section 3.9, we discuss several examples and apply our results to some common finite-volume and finite-difference discretisations.

3.1.1 Sketch of the proof of Theorem 3.5.1

In the last part of this section, we shortly sketch a non-rigorous proof of our main result on the convergence of $\mathcal{A}_\varepsilon$ to the homogenised limit described by $f_{\rm hom}$ (Theorem 3.5.1). Crucial tools to show both the liminf and the Γ-limsup inequality in Theorem 3.5.1 are regularisation procedures for solutions to the continuity equation, both at the discrete and at the continuous level.

In this section, we use the informal notation $\lesssim$ and $\gtrsim$ to mean that the corresponding inequality holds up to a small error in $\varepsilon > 0$, e..g $A_\varepsilon \lesssim B_\varepsilon$ means that $A_\varepsilon \leq B_\varepsilon + o_\varepsilon(1)$ where $o_\varepsilon(1) \to 0$ as $\varepsilon \to 0$.

For $x \in \mathcal{X}_\varepsilon \subset \mathbb{T}^d$, we denote by $x_{\sf z}$ the unique element of $\mathbb{Z}^d_\varepsilon$ satisfying $x \in Q_\varepsilon^{x_{\sf z}} = [0,\varepsilon)^d + \varepsilon x_{\sf z}$. Note that $\{Q_\varepsilon^z \ : \ z \in \mathbb{Z}_\varepsilon^d\}$ defines a partition of $\mathbb{T}^d$.

In order to compare discrete and continuous measures, we make use of the embedding maps for $m \in \mathcal{P}(\mathcal{X}_\varepsilon)$ and anti-symmetric $J : \mathcal{E}_\varepsilon \to \mathbb{R}$

$$\begin{aligned}
\iota_\varepsilon m &:= \varepsilon^{-d} \sum_{x \in \mathcal{X}_\varepsilon} m(x) \mathscr{L}^d|_{Q_\varepsilon^{x_{\sf z}}} \in \mathcal{P}(\mathbb{T}^d)\,, \\
\iota_\varepsilon J &:= \varepsilon^{-d+1} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \frac{J(x,y)}{2} \bigg(\int_0^1 \mathscr{L}^d|_{Q_\varepsilon^{(1-s)x_{\sf z}+sy_{\sf z}}} \,\mathrm{d}s \bigg) (y_{\sf z} - x_{\sf z}) \in \mathcal{M}^d(\mathbb{T}^d)\,,
\end{aligned}$$

as they preserve the continuity equation: if $(\mathbf{m},\mathbf{J}) \in \mathcal{CE}_\varepsilon$, then $(\iota_\varepsilon \mathbf{m}, \iota_\varepsilon \mathbf{J}) \in \mathbb{CE}$.

We also use the notation $\mathcal{F}_\varepsilon(m,J) := \sum_{(x,y)\in\mathcal{E}_\varepsilon} \varepsilon^d F_{xy}\bigg(\frac{m(x)}{\varepsilon^d}, \frac{m(y)}{\varepsilon^d}, \frac{J(x,y)}{\varepsilon^{d-1}}\bigg)$.

Sketch of the Γ-liminf inequality. Consider curves $(m_t^\varepsilon)_{t\in(0,1)} \subseteq \mathcal{M}_+(\mathcal{X}_\varepsilon)$ and let $\mathbf{m}^\varepsilon \in \mathcal{M}_+((0,1) \times \mathcal{X}_\varepsilon)$ be the corresponding measure on space-time defined by $\mathbf{m}^\varepsilon(\mathrm{d}x, \mathrm{d}t) =$

$m_t^\varepsilon(\mathrm{d}x)\,\mathrm{d}t$. Suppose that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ vaguely in $\mathcal{M}_+((0,1)\times\mathbb{T}^d)$ as $\varepsilon \to 0$. The goal is to show the *liminf inequality*

$$\liminf_{\varepsilon\to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) \geq \mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu}). \tag{3.6}$$

We can assume that $\mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) = \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \leq C < \infty$ for every $\varepsilon > 0$, for some sequence of vector fields $\mathbf{J}^\varepsilon$ such that $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon$. As we are going to see in (3.41), the embedded solutions to the continuity equation $(\iota_\varepsilon \mathbf{m}^\varepsilon, \iota_\varepsilon \mathbf{J}^\varepsilon) \in \mathbb{CE}$ defines curves of measures with densities with respect to $\mathscr{L}^d$ on $\mathbb{T}^d$ of the form, for every $u \in Q_\varepsilon^{\bar z} \subset \mathbb{T}^d$

$$\rho_t(u) = \varepsilon^{-d} \sum_{\substack{x\in\mathcal{X}_\varepsilon\\ x_{\mathsf{z}}=\bar z}} m_t^\varepsilon(x) \quad \text{and} \quad j_t(u) = \frac{1}{2\varepsilon^{d-1}} \sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon\\ x_{\mathsf{z}}=\bar z}} J_{t,u}^\varepsilon(x,y)\big(y_{\mathsf{z}} - x_{\mathsf{z}}\big),$$

where $J_{t,u}^\varepsilon \in \mathbb{R}^{\mathcal{E}_\varepsilon}$ is a convex combination of $\Big\{ J_t^\varepsilon\Big(\,\cdot - \varepsilon z\Big) \,:\, z \in \mathbb{Z}_\varepsilon^d\,,\ |z|_\infty \leq R_0 + 1 \Big\}$.

As we are going to estimate the discrete energies at any time $t \in (0,1)$, for simplicity we drop the time dependence and write $\rho = \rho_t$, $j = j_t$, $m^\varepsilon = m_t^\varepsilon$, $J^\varepsilon = J_t^\varepsilon$, $J_u^\varepsilon = J_{t,u}^\varepsilon$.

The main goal is to construct, for every $u \in Q_\varepsilon^{\bar z}$, a representative

$$\left(\frac{\widehat{m}_u}{\varepsilon^d}, \frac{\widehat{J}_u}{\varepsilon^{d-1}} \right) \in \mathsf{Rep}\,\Big(\rho(u), j(u)\Big) \tag{3.7}$$

which is approximately equal to the values of $(m^\varepsilon, J^\varepsilon)$ close to $\mathcal{X} \cap \{x_{\mathsf{z}} = \bar z\}$. The lower bound (3.6) would then follow by integrating in time the static estimate

$$\mathcal{F}_\varepsilon(m,J) \gtrsim \varepsilon^d \sum_{\bar z \in \mathbb{Z}_\varepsilon^d} \mathcal{F}\left(\frac{\widehat{m}_{\varepsilon\bar z}}{\varepsilon^d}, \frac{\widehat{J}_{\varepsilon \bar z}}{\varepsilon^{d-1}} \right) \gtrsim \int_{\mathbb{T}^d} f_{\mathrm{hom}}\Big(\rho(u), j(u)\Big)\,\mathrm{d}u = \mathbb{F}_{\mathrm{hom}}(\iota_\varepsilon m, \iota_\varepsilon J)\,, \tag{3.8}$$

and using the lower semicontinuity of $\mathbb{A}_{\mathrm{hom}}$, where in the last inequality we used the very definition of the homogenised density $f_{\mathrm{hom}}(\rho(u), j(u))$, which corresponds to the minimal microscopic cost with total mass $\rho(u)$ and flux $j(u)$.

In order to find the sought representatives in (3.7), the natural choice is to define $\widehat{m}_u \in \mathbb{R}_+^{\mathcal{X}}$ and $\tilde{J}_u \in \mathbb{R}_a^{\mathcal{E}}$ by taking the values of m and J_u in the ε-cube at $\bar z$, and insert these values at every cube in $(\mathcal{X}, \mathcal{E})$, so that the result is $\mathbb{Z}^d$-periodic. Precisely:

$$\widehat{m}_u(x) := m(\varepsilon \bar x)\,, \quad \tilde{J}_u(x,y) := J_u(\varepsilon\bar x, \varepsilon(y - x_{\mathsf{z}} + \bar z))\,, \quad \text{for } (x,y) \in \mathcal{E},$$

where $\bar x := x - x_{\mathsf{z}} + \bar z$. This would ensure that $\varepsilon^{-d}\widehat{m}_u \in \mathsf{Rep}\,\Big(\rho(u)\Big)$. Unfortunately, this construction would produce a vector field $\varepsilon^{-(d-1)}\tilde{J}_u$ which (in general) does not belong to $\mathsf{Rep}\,\Big(j(u)\Big)$: indeed, while $\tilde{J}_u$ has the desired effective flux (i.e., $\mathsf{Eff}(\varepsilon^{-(d-1)}\tilde{J}_u) = j(u)$, as given in (3.5)), it would not be (in general) divergence-free.

In order to deal with this complication, we shall introduce a *corrector field* $\bar{J}_u$, i.e., an anti-symmetric and $\mathbb{Z}^d$-periodic function $\bar{J}_u : \mathcal{E} \to \mathbb{R}$ satisfying

$$\mathsf{div}\,\bar{J}_u = -\,\mathsf{div}\,\tilde{J}_u, \quad \mathsf{Eff}(\bar{J}_u) = 0, \quad \text{and} \quad \big\|\bar{J}_u\big\|_{\ell^\infty(\mathcal{E}^Q)} \leq \tfrac{1}{2}\big\|\,\mathsf{div}\,\tilde{J}_u\big\|_{\ell^1(\mathcal{X}^Q)}, \tag{3.9}$$

whose existence we prove in Lemma 3.7.3.

It is clear that if we set $\widehat{J}_u := \tilde{J}_u + \bar{J}_u$ by construction we have div $\widehat{J}_u = 0$ and Eff $\left(\varepsilon^{-(d-1)}\widehat{J}_u\right) = j(u)$, thus

$$\frac{\widehat{J}_u}{\varepsilon^{d-1}} := \frac{\tilde{J}_u + \bar{J}_u}{\varepsilon^{d-1}} \in \mathsf{Rep}\left(j_u\right).$$

To carry out this program and prove a lower bound of the form (3.8), we need to quantify the error we perform passing from $(m^\varepsilon, J^\varepsilon)$ to $\left\{(\widehat{m}_u, \widehat{J}_u) \ : \ u \in \mathbb{T}^d\right\}$. It is evident by construction and from (3.9) that spatial and time regularity of $(m^\varepsilon, J^\varepsilon)$ are crucial to this purpose. For example, an ℓ^∞-bound on the time derivative of the form $\|\partial_t m_t^\varepsilon\|_\infty \leq C\varepsilon^d$ (or, in other words, a Lipschitz bound in time for ρ_t) together with $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon$ would imply a control on div J and thus a control of the error in (3.9) of the form $\|\varepsilon^{1-d}\bar{J}_u\|_\infty \leq C\varepsilon$.

This is why a key, first step in our proof is a regularisation procedure at the discrete level: for any given sequence of curves $\left\{(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon \ : \ \varepsilon > 0\right\}$ of (uniformly) bounded energy $\mathcal{A}_\varepsilon$, we can exihibit another sequence $\left\{(\tilde{\mathbf{m}}^\varepsilon, \tilde{\mathbf{J}}^\varepsilon) \in \mathcal{CE}_\varepsilon \ : \ \varepsilon > 0\right\}$, quantitatively close as measures and in energy $\mathcal{A}_\varepsilon$ to the first one, which enjoy good Lipschitz and l^∞ properties and for which the above explained program can be carried out.

This result is the content of Proposition 3.7.1 and it is based on a three-fold regularisation, that is in energy, in time, and in space.

Sketch of the Γ-limsup inequality. The goal is to show that, for every $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}$, we can find $\mathbf{m}^\varepsilon \in \mathcal{M}_+((0,1) \times \mathcal{X}_\varepsilon)$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+((0,1) \times \mathbb{T}^d)$ and

$$\limsup_{\varepsilon \to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) \leq \mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu}, \boldsymbol{\nu})\,. \tag{3.10}$$

In a similar fashion as in the the proof of the Γ-liminf inequality, the first step is a regularisation procedure, this time at the continuous level (Proposition 3.8.26). Thanks to this approximation result, in the sketch we can without loss of generality assume that

$$\mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu}, \boldsymbol{\nu}) < \infty \quad \text{and} \quad \left\{\Big(\rho_t(x), j_t(x)\Big) \ : \ (t,x) \in [0,1] \times \mathbb{T}^d\right\} \Subset \mathsf{D}(f_{\mathrm{hom}})^\circ\,, \tag{3.11}$$

where $(\rho_t, j_t)_t$ are the *smooth* densities of $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}$ with respect to $\mathscr{L}^{d+1}$ on $(0,1) \times \mathbb{T}^d$.

Note that the convexity of f_{hom} ensures its Lipschitz-continuity on every compact set $K \Subset \mathsf{D}(f_{\mathrm{hom}})^\circ$, hence the assumption (3.11) allows us to assume such regularity for the rest of the proof.

The idea is to split the proof of the upper bound into several steps. In short, we first discretise the continuous measures $(\boldsymbol{\mu}, \boldsymbol{\nu})$ and identify *optimal discrete microstructures*, defined through the minimisation of the cell problem f_{hom}, on each ε-cube Q_ε^z, $z \in \mathbb{Z}_\varepsilon^d$. A key difficult at this stage is that the optimal selection has the flaw of not preserving the continuity equation, hence an additional *correction* is needed. To this purpose, we first apply the discrete regularisation result Proposition 3.7.1 to obtain regular discrete curves and then find suitable *small* correctors that provide discrete competitors for $\mathcal{A}_\varepsilon$, that is solutions to $\mathcal{CE}_\varepsilon$ which are *close* to the optimal selection.

Let us explain these steps in more detail.

Step 1: For every $z \in \mathbb{Z}_\varepsilon^d$, $t \in (0,1)$, and each cube Q_ε^z we consider the natural discretisation of $(\boldsymbol{\mu}, \boldsymbol{\nu})$, that we denote by $\big(\mathsf{P}_\varepsilon \mu_t(z), \mathsf{P}_\varepsilon \nu_t(z)\big)_{t,z} \subset \mathbb{R}_+ \times \mathbb{R}^d$, given by

$$\mathsf{P}_\varepsilon \mu_t(z) := \mu_t(Q_\varepsilon^z)\,, \quad \mathsf{P}_\varepsilon \nu_t(z) := \left(\int_{\partial Q_\varepsilon^z \cap \partial Q_\varepsilon^{z+e_i}} j_t \cdot e_i \, \mathrm{d}\mathcal{H}^{d-1} \right)_{i=1}^d .$$

An important feature of the operator P_ε is that it preserves the continuity equation from $\mathbb{T}^d$ to $\mathbb{Z}_\varepsilon^d$, in the sense that for $t \in (0,1)$ and $z \in \mathbb{Z}_\varepsilon^d$

$$\partial_t \mathsf{P}_\varepsilon \mu_t(z) + \sum_{i=1}^d \Big(\mathsf{P}_\varepsilon \nu_t(z) - \mathsf{P}_\varepsilon \nu_t(z - e_i) \Big) \cdot e_i = 0\,.$$

Step 2: We build the associated *optimal discrete microstructure* for the cell problem for each cube Q_ε^z, meaning we select $(\mathbf{m}, \mathbf{J}) = \big(m_t^z, J_t^z\big)_{t \in (0,1), z \in \mathbb{Z}_\varepsilon^d}$ such that

$$\left(\frac{m_t^z}{\varepsilon^d}, \frac{J_t^z}{\varepsilon^{d-1}} \right) \in \underset{o}{\mathsf{Rep}} \left(\frac{\mathsf{P}_\varepsilon \mu_t(z)}{\varepsilon^d}, \frac{\mathsf{P}_\varepsilon \nu_t(z)}{\varepsilon^{d-1}} \right),$$

where Rep_o denotes the set of optimal representatives in the definition of the cell-formula (3.4). Using the smoothness of $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, one can in particular show that

$$\sum_{z \in \mathbb{Z}_\varepsilon^d} \varepsilon^d \mathcal{F}_1 \left(\frac{m_t^z}{\varepsilon^d}, \frac{J_t^z}{\varepsilon^{d-1}} \right) \lesssim \mathbb{F}_{\mathrm{hom}}(\mu_t, \nu_t)\,. \tag{3.12}$$

Step 3: The next step is to glue together the microstructures $(\mathbf{m}, \mathbf{J})$ defined for every $z \in \mathbb{Z}_\varepsilon^d$ via a *gluing operator* $\mathcal{G}_\varepsilon$ (Definition 3.8.4) to produce a global one $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon) \in \mathcal{M}_+((0,1) \times \mathcal{X}_\varepsilon) \times \mathcal{M}_a((0,1) \times \mathcal{E}_\varepsilon)$.

Thanks to the fact the gluing operators are mass preserving and that $m_t^z \in \mathsf{Rep}(\mathsf{P}_\varepsilon \mu_t(z))$, it is not hard to see that $\iota_\varepsilon \widehat{\mathbf{m}}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+((0,1) \times \mathbb{T}^d)$ as $\varepsilon \to 0$.

Step 4: In contrast to P_ε, the latter operation produces curves $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$ which would (in general) *not be a solution* to the discrete continuity equation $\mathcal{CE}_\varepsilon$. Therefore, we seek to find suitable *corrector vector fields* in order to obtain a discrete solution, and thus a candidate for $\mathcal{A}_\varepsilon(\widehat{\mathbf{m}}^\varepsilon)$.

To this purpose, the next step is to regularise $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$ by applying Proposition 3.7.1 and obtaining a regular curve which is quantitatively close as measures and in energy to the first one. Note that no discrete regularity is (in general) guaranteed to $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$, despite the smoothness assumption on $(\boldsymbol{\mu}, \boldsymbol{\nu})$, due to possible singularities of F_{xy}.

For the sake of the exposition, we shall discuss the last steps of the proof assuming that $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$ already enjoy the Lipschitz and ℓ^∞–regularity properties ensured by Proposition 3.7.1.

Step 5: For sufficiently regular $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$, we seek a discrete competitor for $\mathcal{A}_\varepsilon(\widehat{\mathbf{m}}^\varepsilon)$ which is close to $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$. As the latter does not necessary belong to $\mathcal{CE}_\varepsilon$, we find suitable correctors

$\mathbf{V}^\varepsilon$ such that the corrected curves $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon + \mathbf{V}^\varepsilon)$ belong to $\mathcal{CE}_\varepsilon$, with $\mathbf{V}^\varepsilon$ *quantitative small*, i.e. satisfying a bound of the form

$$\sup_{t\in(0,1)} \left\| \varepsilon^{1-d} V_t^\varepsilon \right\|_{\ell_\infty(\mathcal{E}_\varepsilon)} \leq C\varepsilon \,. \tag{3.13}$$

The existence of the corrector $\mathbf{V}^\varepsilon$, together with the quantitative bound, is quite involved and possibly the most difficult part of the proof. It is based on a localisation argument (Lemma 3.8.22) and the study of the divergence equation on periodic graphs (Lemma 3.8.16), performed at the level of each cube Q_ε^z, for every $z \in \mathbb{Z}_\varepsilon^d$.

The regularity of $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$ is crucial in order to obtain the estimate (3.13).

Step 6: The final step consists in estimating the energy of the measures defined as $\mathbf{m}^\varepsilon := \widehat{\mathbf{m}}^\varepsilon \to \boldsymbol{\mu}$ weakly as $\varepsilon \to 0$, and the vector fields $\mathbf{J}^\varepsilon := \widehat{\mathbf{J}}^\varepsilon + \mathbf{V}^\varepsilon$.

Using the regularity assumption on $(\widehat{\mathbf{m}}^\varepsilon, \widehat{\mathbf{J}}^\varepsilon)$, the smoothness (3.11) of $(\boldsymbol{\mu}, \boldsymbol{\nu})$, and the convexity of f_{hom}, together with the quantitative bound (3.13) and (3.12) for the corrector we obtain

$$\mathcal{F}_\varepsilon(m_t^\varepsilon, J_t^\varepsilon) \lesssim \mathcal{F}_\varepsilon(\widehat{m}_t^\varepsilon, \widehat{J}_t^\varepsilon) \lesssim \sum_{z\in\mathbb{Z}_\varepsilon^d} \varepsilon^d \mathcal{F}_1\left(\frac{m_t^z}{\varepsilon^d}, \frac{J_t^z}{\varepsilon^{d-1}}\right) \lesssim \mathbb{F}_{\text{hom}}(\mu_t, \nu_t) \,.$$

Using this bound and that $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon$, we integrate in time and get

$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) \leq \limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \leq \mathbb{A}_{\text{hom}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \,,$$

which is the sought upper bound (3.10).

3.2 Discrete dynamical optimal transport on $\mathbb{Z}^d$-periodic graphs

This section contains the definition of the transport problem in the discrete periodic setting. In Section 3.2.1 we introduce the basic objects: a $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$ and an admissible cost function F. Given a triple $(\mathcal{X}, \mathcal{E}, F)$, we introduce a family of transport problems on rescaled graphs $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$ in Section 3.2.2.

3.2.1 Discrete $\mathbb{Z}^d$-periodic setting

Our setup consists of the following data:

Assumption 3.2.1. $(\mathcal{X}, \mathcal{E})$ is a locally finite and $\mathbb{Z}^d$-periodic connected graph of bounded degree.

More precisely, we assume that

$$\mathcal{X} = \mathbb{Z}^d \times \mathbb{V},$$

where $\mathbb{V}$ is a finite set. The coordinates of $x = (z, v) \in \mathcal{X}$ will be denoted by

$$x_{\mathsf{z}} := z, \qquad x_{\mathsf{v}} := v.$$

The set of edges $\mathcal{E} \subseteq \mathcal{X} \times \mathcal{X}$ is symmetric and $\mathbb{Z}^d$-periodic, in the sense that

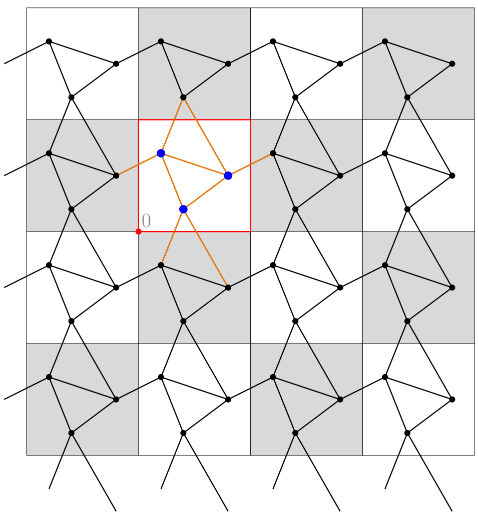

Figure 3.2: A fragment of a $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$. The blue nodes represent $\mathcal{X}^Q$ and the orange edges represent $\mathcal{E}^Q$.

$$(x, y) \in \mathcal{E} \quad \text{iff} \quad \big(S^z(x), S^z(y)\big) \in \mathcal{E} \text{ for all } z \in \mathbb{Z}^d.$$

Here, $S^{\bar z} : \mathcal{X} \to \mathcal{X}$ is the *shift operator* defined by

$$S^{\bar z}(x) = (\bar z + z, v) \quad \text{for } x = (z, v) \in \mathcal{X}.$$

We write $x \sim y$ whenever $(x, y) \in \mathcal{E}$.

Let $R_0 := \max_{(x,y)\in\mathcal{E}} |x_\mathsf{z} - y_\mathsf{z}|_{\ell_\infty^d}$ be the maximal edge length, measured with respect to the supremum norm $|\cdot|_{\ell_\infty^d}$ on $\mathbb{R}^d$. It will be convenient to use the notation

$$\mathcal{X}^Q := \{x \in \mathcal{X} \ : \ x_\mathsf{z} = 0\} \quad \text{and} \quad \mathcal{E}^Q := \Big\{(x, y) \in \mathcal{E} \ : \ x_\mathsf{z} = 0\Big\}.$$

Remark 3.2.2 (Abstract vs. embedded graphs). Rather than working with abstract $\mathbb{Z}^d$-periodic graphs, it is possible to regard $\mathcal{X}$ as a $\mathbb{Z}^d$-periodic subset of $\mathbb{R}^d$, by choosing $\mathbb{V}$ to be a subset of $[0, 1)^d$ and using the identification $(z, v) \equiv z + v$, see Figure 3.2. Since the embedding plays no role in the formulation of the discrete problem, we work with the abstract setup.

Assumption 3.2.3 (Admissible cost function)**.** The function $F : \mathbb{R}_+^{\mathcal{X}} \times \mathbb{R}_a^{\mathcal{E}} \to \mathbb{R} \cup \{+\infty\}$ is assumed to have the following properties:

(a) F is convex and lower semicontinuous.

(b) F is *local* in the sense that there exists $R_1 < \infty$ such that $F(m, J) = F(m', J')$ whenever $m, m' \in \mathbb{R}_+^{\mathcal{X}}$ and $J, J' \in \mathbb{R}_a^{\mathcal{E}}$ agree within a ball of radius R_1, i.e.,

$$\begin{aligned} m(x) &= m'(x) && \text{for all } x \in \mathcal{X} \text{ with } |x_\mathsf{z}|_{\ell_\infty^d} \le R_1, \quad \textit{and} \\ J(x, y) &= J'(x, y) && \text{for all } (x, y) \in \mathcal{E} \text{ with } |x_\mathsf{z}|_{\ell_\infty^d}, |y_\mathsf{z}|_{\ell_\infty^d} \le R_1. \end{aligned}$$

(c) F is of at least *linear growth*, i.e., there exist $c > 0$ and $C < \infty$ such that

$$F(m, J) \ge c \sum_{(x,y)\in\mathcal{E}^Q} |J(x, y)| - C\left(1 + \sum_{\substack{x\in\mathcal{X} \\ |x|_{\ell_\infty^d} \le R}} m(x)\right) \tag{3.14}$$

for any $m \in \mathbb{R}_+^{\mathcal{X}}$ and $J \in \mathbb{R}_a^{\mathcal{E}}$. Here, $R := \max\{R_0, R_1\}$.

(d) There exist a $\mathbb{Z}^d$-periodic function $m^\circ \in \mathbb{R}_+^{\mathcal{X}}$ and a $\mathbb{Z}^d$-periodic and divergence-free vector field $J^\circ \in \mathbb{R}_a^{\mathcal{E}}$ such that

$$(m^\circ, J^\circ) \in \mathsf{D}(F)^\circ. \tag{3.15}$$

Remark 3.2.4. As F is local, it depends on finitely many parameters. Therefore, $\mathsf{D}(F)^\circ$, the topological interior of its domain $\mathsf{D}(F)$ is defined unambiguously.

Remark 3.2.5. In many examples, the function F takes one of the following forms, for suitable functions F_x and F_{xy}:

$$F(m, J) = \sum_{x\in\mathcal{X}^Q} F_x\Big(m(x), \big(J(x, y)\big)_{y\sim x}\Big), \quad F(m, J) = \sum_{(x,y)\in\mathcal{E}^Q} F_{xy}\Big(m(x), m(y), J(x, y)\Big).$$

We then say that F is vertex-based (respectively, edge-based).

Remark 3.2.6. Of particular interest are edge-based functions of the form

$$F(m,J)=\frac{1}{p}\sum_{(x,y)\in\mathcal{E}^Q}\frac{|J(x,y)|^p}{\Lambda\Big(q_{xy}m(x),q_{yx}m(y)\Big)^{p-1}},\tag{3.16}$$

where $1\le p<\infty$, the constants $q_{xy},q_{yx}>0$ are fixed parameters defined for $(x,y)\in\mathcal{E}^Q$, and Λ is a suitable mean (i.e., $\Lambda:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+$ is a jointly concave and 1-homogeneous function satisfying $\Lambda(1,1)=1$). Functions of this type arise naturally in discretisations of Wasserstein gradient-flow structures [Maa11, Mie11, CHLZ12].

We claim that these cost function satisfy the growth condition (3.14). Indeed, using Young's inequality $|J|\le\frac{1}{p}\frac{|J|^p}{\Lambda^{p-1}}+\frac{p-1}{p}\Lambda$ we infer that

$$\begin{aligned}\sum_{(x,y)\in\mathcal{E}^Q}|J(x,y)|&\le\frac{1}{p}\sum_{(x,y)\in\mathcal{E}^Q}\frac{|J(x,y)|^p}{\Lambda\Big(q_{xy}m(x),q_{yx}m(y)\Big)^{p-1}}\\&\qquad+\frac{p-1}{p}\sum_{(x,y)\in\mathcal{E}^Q}\Lambda\Big(q_{xy}m(x),q_{yx}m(y)\Big)\\&\le F(m,J)+C\sum_{x\in\mathcal{X},|x|_{\ell_\infty^d}\le R_0}m(x),\end{aligned}$$

with constant $C>0$ depending on $\max_{x,y}(q_{xy}+q_{yx})$. This shows that (3.14) is satisfied.

3.2.2 Rescaled setting

Let $(\mathcal{X},\mathcal{E})$ be a locally finite and $\mathbb{Z}^d$-periodic graph as above. Fix $\varepsilon>0$ such that $\frac{1}{\varepsilon}\in\mathbb{N}$. *The assumption that* $\frac{1}{\varepsilon}\in\mathbb{N}$ *remains in force throughout the paper.*

The rescaled graph. Let $\mathbb{T}^d_\varepsilon=(\varepsilon\mathbb{Z}/\mathbb{Z})^d$ be the discrete torus of mesh size ε. The corresponding equivalence classes are denoted by $[\varepsilon z]$ for $z\in\mathbb{Z}^d$. To improve readability, we occasionally omit the brackets. Alternatively, we may write $\mathbb{T}^d_\varepsilon=\varepsilon\mathbb{Z}^d_\varepsilon$ where $\mathbb{Z}^d_\varepsilon=\Big(\mathbb{Z}/\frac{1}{\varepsilon}\mathbb{Z}\Big)^d$.

The rescaled graph $(\mathcal{X}_\varepsilon,\mathcal{E}_\varepsilon)$ is constructed by rescaling the $\mathbb{Z}^d$-periodic graph $(\mathcal{X},\mathcal{E})$ and wrapping it around the torus. More formally, we consider the finite sets

$$\mathcal{X}_\varepsilon:=\mathbb{T}^d_\varepsilon\times\mathbb{V}\quad\text{and}\quad\mathcal{E}_\varepsilon:=\Big\{\Big(T^0_\varepsilon(x),T^0_\varepsilon(y)\Big)\ :\ (x,y)\in\mathcal{E}\Big\}$$

where, for $\bar z\in\mathbb{Z}^d_\varepsilon$,

$$T^{\bar z}_\varepsilon:\mathcal{X}\to\mathcal{X}_\varepsilon,\qquad(z,v)\mapsto\Big([\varepsilon(\bar z+z)],v\Big).\tag{3.17}$$

Throughout the paper we always assume that $\varepsilon R_0<\frac{1}{2}$, to avoid that edges in $\mathcal{E}$ "bite themselves in the tail" when wrapped around the torus. For $x=\Big([\varepsilon z],v\Big)\in\mathcal{X}_\varepsilon$ we will write

$$x_{\mathsf{z}}:=z\in\mathbb{Z}^d_\varepsilon,\qquad x_{\mathsf{v}}:=v\in\mathbb{V}.$$

The rescaled cost function. Let $F:\mathbb{R}^{\mathcal{X}}_+\times\mathbb{R}^{\mathcal{E}}_a\to\mathbb{R}\cup\{+\infty\}$ be a cost function satisfying Assumption 3.2.3. For $\varepsilon>0$ satisfying the conditions above, we shall define a corresponding energy functional $\mathcal{F}_\varepsilon$ in the rescaled periodic setting.

First we introduce some notation, which we use to transfer functions defined on $\mathcal{X}_\varepsilon$ to $\mathcal{X}$ (and from $\mathcal{E}_\varepsilon$ to $\mathcal{E}$). Let $\bar{z} \in \mathbb{Z}^d_\varepsilon$. Each function $\psi : \mathcal{X}_\varepsilon \to \mathbb{R}$ induces a $\frac{1}{\varepsilon}\mathbb{Z}^d$-periodic function

$$\tau^{\bar{z}}_\varepsilon \psi : \mathcal{X} \to \mathbb{R}, \qquad \big(\tau^{\bar{z}}_\varepsilon \psi\big)(x) := \psi\big(T^{\bar{z}}_\varepsilon(x)\big) \quad \text{for } x \in \mathcal{X}.$$

see Figure 3.3. Similarly, each function $J : \mathcal{E}_\varepsilon \to \mathbb{R}$ induces a $\frac{1}{\varepsilon}\mathbb{Z}^d$-periodic function

$$\tau^{\bar{z}}_\varepsilon J : \mathcal{E} \to \mathbb{R}, \qquad \big(\tau^{\bar{z}}_\varepsilon J\big)(x,y) := J\big(T^{\bar{z}}_\varepsilon(x), T^{\bar{z}}_\varepsilon(y)\big) \quad \text{for } (x,y) \in \mathcal{E}.$$

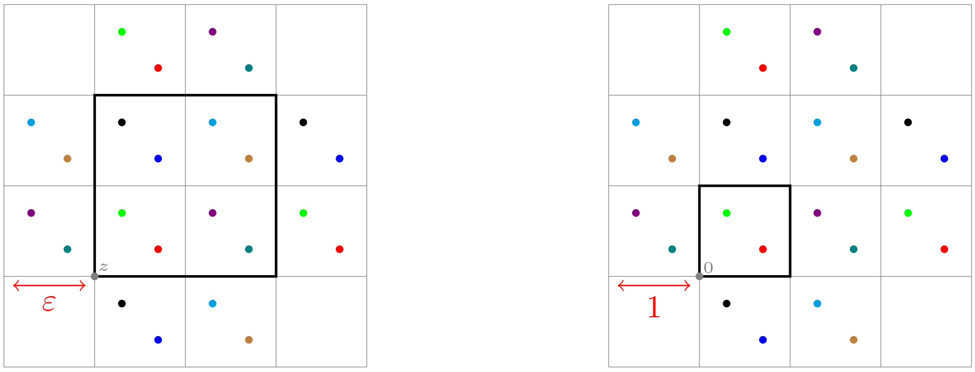

Figure 3.3: On the left, the value of a function $\psi : \mathcal{X}_\varepsilon \to \mathbb{R}$ correspond to different colors over the nodes. On the right, the corresponding values of $\tau^z_\varepsilon \psi : \mathcal{X} \to \mathbb{R}$.

Definition 3.2.7 (Discrete energy functional)**.** The rescaled cost function is defined by

$$\mathcal{F}_\varepsilon : \mathbb{R}^{\mathcal{X}_\varepsilon}_+ \times \mathbb{R}^{\mathcal{E}_\varepsilon}_a \to \mathbb{R} \cup \{+\infty\}, \qquad (m,J) \mapsto \sum_{z \in \mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{\tau^z_\varepsilon m}{\varepsilon^d}, \frac{\tau^z_\varepsilon J}{\varepsilon^{d-1}}\right).$$

Remark 3.2.8. We note that $\mathcal{F}_\varepsilon(m,J)$ is well-defined as an element in $\mathbb{R} \cup \{+\infty\}$. Indeed, the (at least) linear growth condition (3.14) yields

$$\begin{aligned}\mathcal{F}_\varepsilon(m,J) = \sum_{z \in \mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{\tau^z_\varepsilon m}{\varepsilon^d}, \frac{\tau^z_\varepsilon J}{\varepsilon^{d-1}}\right) &\geq -C \sum_{z \in \mathbb{Z}^d_\varepsilon} \varepsilon^d \left(1 + \sum_{\substack{x \in \mathcal{X} \\ |x|_{\ell^d_\infty} \leq R}} \frac{\tau^z_\varepsilon m(x)}{\varepsilon^d}\right) \\ &\geq -C\left(1 + (2R+1)^d \sum_{x \in \mathcal{X}_\varepsilon} m(x)\right) > -\infty.\end{aligned}$$

For $\bar{z} \in \mathbb{Z}^d_\varepsilon$ it will be useful to consider the *shift operator* $S^{\bar{z}}_\varepsilon : \mathcal{X}_\varepsilon \to \mathcal{X}_\varepsilon$ and $S^{\bar{z}}_\varepsilon : \mathcal{E}_\varepsilon \to \mathcal{E}_\varepsilon$ defined by

$$\begin{aligned} S^{\bar{z}}_\varepsilon(x) &= \big([\varepsilon(\bar{z} + z)], v\big) && \text{for } x = ([\varepsilon z], v) \in \mathcal{X}_\varepsilon, \\ S^{\bar{z}}_\varepsilon(x,y) &= \big(S^{\bar{z}}_\varepsilon(x), S^{\bar{z}}_\varepsilon(y)\big) && \text{for } (x,y) \in \mathcal{E}_\varepsilon. \end{aligned}$$

Moreover, for $\psi : \mathcal{X}_\varepsilon \to \mathbb{R}$ and $J : \mathcal{E}_\varepsilon \to \mathbb{R}$ we define

$$\begin{aligned} &\sigma^{\bar{z}}_\varepsilon \psi : \mathcal{X}_\varepsilon \to \mathbb{R}, && \big(\sigma^{\bar{z}}_\varepsilon \psi\big)(x) := \psi(S^{\bar{z}}_\varepsilon(x)) && \text{for } x \in \mathcal{X}_\varepsilon, \\ &\sigma^{\bar{z}}_\varepsilon J : \mathcal{E}_\varepsilon \to \mathbb{R}, && \big(\sigma^{\bar{z}}_\varepsilon J\big)(x,y) := J(S^{\bar{z}}_\varepsilon(x,y)) && \text{for } (x,y) \in \mathcal{E}_\varepsilon. \end{aligned} \tag{3.18}$$

Definition 3.2.9 (Discrete continuity equation)**.** A pair $(\mathbf{m}, \mathbf{J})$ is said to be a solution to the discrete continuity equation if $\mathbf{m} : \mathcal{I} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ is continuous, $\mathbf{J} : \mathcal{I} \to \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ is Borel measurable, and

$$\partial_t m_t(x) + \sum_{y \sim x} J_t(x, y) = 0 \tag{3.19}$$

for all $x \in \mathcal{X}_\varepsilon$ in the sense of distributions. We use the notation

$$(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^T.$$

Remark 3.2.10. We may write (3.20) as $\partial_t m_t + \operatorname{div} J_t = 0$ using the notation (3.137).

Lemma 3.2.11 (Mass preservation)**.** *Let* $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$*. Then we have* $m_s(\mathcal{X}_\varepsilon) = m_t(\mathcal{X}_\varepsilon)$ *for all* $s, t \in \mathcal{I}$*.*

Proof. Without loss of generality, suppose that $s, t \in \mathcal{I}$ with $s < t$. Approximating the characteristic function $\chi_{[s,t]}$ by smooth test functions, we obtain, for all $x \in \mathcal{X}_\varepsilon$,

$$m_t(x) - m_s(x) = \int_s^t \sum_{y \sim x} J_r(x, y) \, \mathrm{d}r.$$

Summing (3.20) over $x \in \mathcal{X}_\varepsilon$ and using the anti-symmetry of $\mathbf{J}$, the result follows. □

We are now ready to define one of the main objects in this paper.

Definition 3.2.12 (Discrete action functional)**.** For any continuous function $\mathbf{m} : \mathcal{I} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $t \mapsto \sum_{x \in \mathcal{X}_\varepsilon} m_t(x) \in L^1(\mathcal{I})$ and any Borel measurable function $\mathbf{J} : \mathcal{I} \to \mathbb{R}_a^{\mathcal{E}_\varepsilon}$, we define

$$\mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J}) := \int_{\mathcal{I}} \mathcal{F}_\varepsilon(m_t, J_t) \, \mathrm{d}t \in \mathbb{R} \cup \{+\infty\}.$$

Furthermore, we set

$$\mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}) := \inf \Big\{ \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J}) \ : \ (\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}} \Big\}.$$

Remark 3.2.13. We claim that $\mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J})$ is well-defined as an element in $\mathbb{R} \cup \{+\infty\}$. Indeed, the (at least) linear growth condition (3.14) yields as in Remark 3.2.8

$$\mathcal{F}_\varepsilon(m_t, J_t) \geq -C \bigg(1 + (2R+1)^d \sum_{x \in \mathcal{X}_\varepsilon} m_t(x) \bigg).$$

for any $t \in \mathcal{I}$. Since $t \mapsto \sum_{x \in \mathcal{X}_\varepsilon} m_t(x) \in L^1(\mathcal{I})$, the claim follows.

In particular, $\mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J})$ is well-defined whenever $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$, since $t \mapsto \sum_{x \in \mathcal{X}_\varepsilon} m_t(x)$ is constant by Lemma 3.2.11.

Remark 3.2.14. If the time interval is clear from the context, we often simply write $\mathcal{CE}_\varepsilon$ and $\mathcal{A}_\varepsilon$.

The aim of this work is to study the asymptotic behavior of the energies $\mathcal{A}_\varepsilon^{\mathcal{I}}$ as $\varepsilon \to 0$.

3.3 Dynamical optimal transport in the continuous setting

We shall now define a corresponding class of dynamical optimal transport problems on the continuous torus $\mathbb{T}^d$. We start in Section 3.3.1 by defining the natural continuous analogues of the discrete objects from Section 3.2. In Section 3.3.2 we define generalisations of these objects that have better compactness properties.

3.3.1 Continuous continuity equation and action functional

First we define solutions to the continuity equation on a bounded open time interval $\mathcal{I}$.

Definition 3.3.1 (Continuity equation)**.** A pair $(\boldsymbol{\mu}, \boldsymbol{\nu})$ is said to be a solution to the continuity equation $\partial_t \boldsymbol{\mu} + \nabla \cdot \boldsymbol{\nu} = 0$ if the following conditions holds:

(i) $\boldsymbol{\mu} : \mathcal{I} \to \mathcal{M}_+(\mathbb{T}^d)$ is vaguely continuous;

(ii) $\boldsymbol{\nu} : \mathcal{I} \to \mathcal{M}^d(\mathbb{T}^d)$ is a Borel family satisfying $\int_{\mathcal{I}} |\nu_t|(\mathbb{T}^d)\,\mathrm{d}t < \infty$;

(iii) The equation

$$\partial_t \mu_t(x) + \nabla \cdot \nu_t(x) = 0 \tag{3.20}$$

holds in the sense of distributions, i.e., for all $\phi \in \mathcal{C}_c\big(\mathcal{I} \times \mathbb{T}^d\big)$,

$$\int_{\mathcal{I}} \int_{\mathbb{T}^d} \partial_t \phi_t(x)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t + \int_{\mathcal{I}} \int_{\mathbb{T}^d} \nabla \phi_t(x) \cdot \mathrm{d}\nu_t(x)\,\mathrm{d}t = 0.$$

We use the notation

$$(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{CE}^{\mathcal{I}}.$$

We will consider the energy densities f with the following properties.

Assumption 3.3.2. Let $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous and convex function, whose domain has nonempty interior. We assume that there exist constants $c > 0$ and $C < \infty$ such that the (at least) linear growth condition

$$f(\rho, j) \geq c|j| - C(\rho + 1) \tag{3.21}$$

holds for all $\rho \in \mathbb{R}_+$ and $j \in \mathbb{R}^d$.

The corresponding *recession function* $f^\infty : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$f^\infty(\rho, j) := \lim_{t \to +\infty} \frac{f(\rho_0 + t\rho, j_0 + tj)}{t},$$

where $(\rho_0, j_0) \in \mathsf{D}(f)$ is arbitrary. It is well known that the function f^∞ is lower semicontinuous and convex, and it satisfies

$$f^\infty(\rho, j) \geq c|j| - C\rho. \tag{3.22}$$

We refer to [AFP00, Section 2.6] for a proof of these facts.

Let $\mathscr{L}^d$ denote the Lebesgue measure on $\mathbb{T}^d$. For $\mu \in \mathcal{M}_+(\mathbb{T}^d)$ and $\nu \in \mathcal{M}^d(\mathbb{T}^d)$ we consider the Lebesgue decompositions given by

$$\mu = \rho \mathscr{L}^d + \mu^\perp, \qquad \nu = j \mathscr{L}^d + \nu^\perp$$

for some $\rho \in L^1_+(\mathbb{T}^d)$ and $j \in L^1(\mathbb{T}^d;\mathbb{R}^d)$. It is always possible to introduce a measure $\sigma \in \mathcal{M}_+(\mathbb{T}^d)$ such that

$$\mu^\perp = \rho^\perp \sigma, \qquad \nu^\perp = j^\perp \sigma,$$

for some $\rho^\perp \in L^1_+(\sigma)$ and $j^\perp \in L^1(\sigma;\mathbb{R}^d)$. (Take, for instance, $\sigma = \mu^\perp + |\nu^\perp|$.) Using this notation we define the continuous energy as follows.

Definition 3.3.3 (Continuous energy functional)**.** Let $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ satisfy Assumption 3.3.2. We define the continuous energy functional by

$$\begin{aligned} &\mathbb{F} : \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d) \to \mathbb{R} \cup \{+\infty\}, \\ &\mathbb{F}(\mu,\nu) := \int_{\mathbb{T}^d} f\big(\rho(x), j(x)\big)\,\mathrm{d}x + \int_{\mathbb{T}^d} f^\infty\big(\rho^\perp(x), j^\perp(x)\big)\,\mathrm{d}\sigma(x). \end{aligned}$$

Remark 3.3.4. By 1-homogeneity of f^∞, this definition does not depend on the choice of the measure $\sigma \in \mathcal{M}_+(\mathbb{T}^d)$.

Definition 3.3.5 (Action functional)**.** For any curve $\boldsymbol{\mu} : \mathcal{I} \to \mathcal{M}_+(\mathbb{T}^d)$ with $\int_\mathcal{I} \mu_t(\mathbb{T}^d)\,\mathrm{d}t < \infty$ and any Borel measurable curve $\boldsymbol{\nu} : \mathcal{I} \to \mathcal{M}^d(\mathbb{T}^d)$ we define

$$\mathbb{A}^\mathcal{I}(\boldsymbol{\mu},\boldsymbol{\nu}) := \int_\mathcal{I} \mathbb{F}(\mu_t,\nu_t)\,\mathrm{d}t.$$

Furthermore, we set

$$\mathbb{A}^\mathcal{I}(\boldsymbol{\mu}) := \inf_{\boldsymbol{\nu}} \Big\{ \mathbb{A}^\mathcal{I}(\boldsymbol{\mu},\boldsymbol{\nu}) \ : \ (\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^\mathcal{I} \Big\}.$$

Remark 3.3.6. As $f(\rho,j) \geq -C(1+\rho)$ by (3.21), the assumption $\int_\mathcal{I} \mu_t(\mathbb{T}^d)\,\mathrm{d}t < \infty$ ensures that $\mathbb{A}^\mathcal{I}(\boldsymbol{\mu},\boldsymbol{\nu})$ is well-defined in $\mathbb{R} \cup \{+\infty\}$.

Remark 3.3.7 (Dependence on time intervals). Remark 3.2.14 applies in the continuous setting as well. If the time interval is clear from the context, we often simply write $\mathbb{CE}$ and $\mathbb{A}$.

Under additional assumptions on the function f, it is possible to prove compactness for families of solutions to the continuity equation with bounded action; see [DNS09, Corollary 4.10]. However, in our general setting, such a compactness result fails to hold, as the following example shows.

Example 3.3.8 (Lack of compactness)**.** To see this, let $y^\varepsilon(t)$ be the position of a particle of mass m that moves from 0 to $\bar{y} \in [0,\frac{1}{2}]^d$ in the time interval $(a_\varepsilon, b_\varepsilon) := \big(\frac{1-\varepsilon}{2}, \frac{1+\varepsilon}{2}\big)$ with constant speed $\frac{|\bar{y}|}{\varepsilon}$. At all other times in the time interval $\mathcal{I} = (0,1)$ the particle is at rest:

$$y^\varepsilon(t) = \begin{cases} 0, & t \in [0, a_\varepsilon], \\ \big(t - \frac{1}{2}(1-\varepsilon)\big)\varepsilon^{-1}\bar{y}, & t \in \big(a_\varepsilon, b_\varepsilon\big), \\ \bar{y} & t \in [b_\varepsilon, 1]. \end{cases}$$

The associated solution $(\boldsymbol{\mu}^\varepsilon, \boldsymbol{\nu}^\varepsilon)$ to the continuity equation $\partial_t \boldsymbol{\mu}^\varepsilon + \nabla \cdot \boldsymbol{\nu}^\varepsilon = 0$ is given by

$$\mu_t^\varepsilon(\mathrm{d}x) := m\delta_{y^\varepsilon(t)}(\mathrm{d}x), \qquad \nu_t^\varepsilon(\mathrm{d}x) := \frac{m|\bar{y}|}{\varepsilon}\chi_{(a_\varepsilon, b_\varepsilon)}(t)\delta_{y^\varepsilon(t)}(\mathrm{d}x).$$

Let $f(\rho, j) = |j|$ be the total momentum, which satisfies Assumption 3.3.2. We then have $\mathbb{F}(\mu_t^\varepsilon, \nu_t^\varepsilon) = \frac{m|\bar{y}|}{\varepsilon}\mathbb{1}_{(a_\varepsilon, b_\varepsilon)}(t)$, hence $\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}^\varepsilon, \boldsymbol{\nu}^\varepsilon) = m\bar{y}$, independently of ε.

However, as $\varepsilon \to 0$, the motion converges to the discontinous curve given by $\mu_t = \delta_0$ for $t \in [0, \frac{1}{2})$ and $\mu_t = \delta_{\bar{y}}$ for $t \in (\frac{1}{2}, 1]$. In particular, it does not satisfy the continuity equation in the sense above.

3.3.2 Generalised continuity equation and action functional

In view of this lack of compactness, we will extend the definition of the continuity equation and the action functional to more general objects.

Definition 3.3.9 (Continuity equation)**.** A pair of measures $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{M}_+\big(\mathcal{I} \times \mathbb{T}^d\big) \times \mathcal{M}^d\big(\mathcal{I} \times \mathbb{T}^d\big)$ is said to be a solution to the continuity equation

$$\partial_t \boldsymbol{\mu} + \nabla \cdot \boldsymbol{\nu} = 0 \tag{3.23}$$

if, for all $\phi \in \mathcal{C}_c^1\big(\mathcal{I} \times \mathbb{T}^d\big)$, we have

$$\int_{\mathcal{I} \times \mathbb{T}^d} \partial_t \phi \, \mathrm{d}\boldsymbol{\mu} + \int_{\mathcal{I} \times \mathbb{T}^d} \nabla \phi \cdot \mathrm{d}\boldsymbol{\nu} = 0.$$

As above, we use the notation $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$.

Clearly, this definition is consistent with Definition 3.3.5.

Let us now extend the action functional $\mathbb{A}^{\mathcal{I}}$ as well. For this purpose, let $\mathscr{L}^{d+1}$ denote the Lebesgue measure on $\mathcal{I} \times \mathbb{T}^d$. For $\boldsymbol{\mu} \in \mathcal{M}_+\big(\mathcal{I} \times \mathbb{T}^d\big)$ and $\boldsymbol{\nu} \in \mathcal{M}^d\big(\mathcal{I} \times \mathbb{T}^d\big)$ we consider the Lebesgue decompositions given by

$$\boldsymbol{\mu} = \rho \mathscr{L}^{d+1} + \boldsymbol{\mu}^\perp, \qquad \boldsymbol{\nu} = j \mathscr{L}^{d+1} + \boldsymbol{\nu}^\perp$$

for some $\rho \in L_+^1\big(\mathcal{I} \times \mathbb{T}^d\big)$ and $j \in L^1\big(\mathcal{I} \times \mathbb{T}^d; \mathbb{R}^d\big)$. As above, it is always possible to introduce a measure $\boldsymbol{\sigma} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ such that

$$\boldsymbol{\mu}^\perp = \rho^\perp \boldsymbol{\sigma}, \qquad \boldsymbol{\nu}^\perp = j^\perp \boldsymbol{\sigma}, \tag{3.24}$$

for some $\rho^\perp \in L_+^1(\boldsymbol{\sigma})$ and $j^\perp \in L^1(\boldsymbol{\sigma}; \mathbb{R}^d)$.

Definition 3.3.10 (Action functional)**.** We define the action by

$$\begin{aligned}
&\mathbb{A}^{\mathcal{I}} : \mathcal{M}_+\big(\mathcal{I} \times \mathbb{T}^d\big) \times \mathcal{M}^d\big(\mathcal{I} \times \mathbb{T}^d\big) \to \mathbb{R} \cup \{+\infty\}, \\
&\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) := \int_{\mathcal{I} \times \mathbb{T}^d} f\Big(\rho_t(x), j_t(x)\Big) \, \mathrm{d}x \, \mathrm{d}t + \int_{\mathcal{I} \times \mathbb{T}^d} f^\infty\Big(\rho_t^\perp(x), j_t^\perp(x)\Big) \, \mathrm{d}\boldsymbol{\sigma}(t, x).
\end{aligned}$$

Furthermore, we set

$$\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}) := \inf\{\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) : (\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}\}.$$

Remark 3.3.11. This definition does not depend on the choice of $\boldsymbol{\sigma}$, in view of the 1-homogeneity of f^∞. As $f(\rho, j) \geq -C(1+\rho)$ and $f_\infty(\rho, j) \geq -C\rho$ from (3.21) and (3.22), the fact that $\boldsymbol{\mu}(\mathcal{I} \times \mathbb{T}^d) < \infty$ ensures that $\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu})$ is well-defined in $\mathbb{R} \cup \{+\infty\}$.

Example 3.3.12 (Lack of compactness)**.** Continuing Example 3.3.8, we can now describe the limiting jump process as a solution to the generalised continuity equation. Consider the measures $\boldsymbol{\mu}^\varepsilon \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ and $\boldsymbol{\nu}^\varepsilon \in \mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ defined by

$$\boldsymbol{\mu}^\varepsilon(\mathrm{d}x, \mathrm{d}t) = \mu_t^\varepsilon(\mathrm{d}x)\,\mathrm{d}t, \qquad \boldsymbol{\nu}^\varepsilon(\mathrm{d}x, \mathrm{d}t) = \nu_t^\varepsilon(\mathrm{d}x)\,\mathrm{d}t.$$

Then we have $\boldsymbol{\mu}^\varepsilon \to \boldsymbol{\mu}$ and $\boldsymbol{\nu}^\varepsilon \to \boldsymbol{\nu}$ weakly, respectively, in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ and $\mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$, where $\boldsymbol{\mu}$ represents the discontinuous curve

$$\boldsymbol{\mu}(\mathrm{d}x, \mathrm{d}t) = \mathrm{d}\mu_t(x)\,\mathrm{d}t, \quad \text{where } \mu_t = \begin{cases} \delta_0, & t \in [0, \frac{1}{2}), \\ \delta_{\bar{y}}, & t \in (\frac{1}{2}, 1]. \end{cases}$$

The measure $\boldsymbol{\nu}$ does *not* admit a disintegration with respect to the Lebesgue measure on $\mathcal{I}$; in other words, it is not associated to a curve of measures on $\mathbb{T}^d$. We have

$$\boldsymbol{\nu}_t(\mathrm{d}x, \mathrm{d}t) = m|\bar{y}|\mathscr{H}^1|_{[0,\bar{y}]}(\mathrm{d}x)\delta_{1/2}(\mathrm{d}t).$$

Here $\mathscr{H}^1|_{[0,\bar{y}]}$ denotes the 1-dimensional Hausdorff measure on the (shortest) line segment connecting 0 and $\bar{y}$.

Note that $(\boldsymbol{\mu}, \boldsymbol{\nu})$ solves the continuity equation, as $\mathbb{CE}^{\mathcal{I}}$ is stable under joint weak-convergence. Furthermore, we have $\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) = m\bar{y}$.

The next result shows that any solution to the continuity equation $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ induces a (not necessarily continuous) curve of measures $(\mu_t)_t \in \mathcal{I}$. The measure $\boldsymbol{\nu}$ is not always associated to a curve of measures on $\mathcal{I}$; see Example 3.3.12. We refer to Appendix B.2 for the definition of $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$.

Lemma 3.3.13 (Disintegration of solutions to $\mathbb{CE}^{\mathcal{I}}$)**.** *Let $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$. Then $\mathrm{d}\boldsymbol{\mu}(t,x) = \mathrm{d}\mu_t(x)\,\mathrm{d}t$ for some measurable curve $t \mapsto \mu_t \in \mathcal{M}_+(\mathbb{T}^d)$ with finite constant mass. If $\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}) < \infty$, then this curve belongs to $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$ and*

$$\|\boldsymbol{\mu}\|_{\mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))} \leq |\boldsymbol{\nu}|\big(\mathcal{I} \times \mathbb{T}^d\big). \tag{3.25}$$

Proof. Let $\lambda \in \mathcal{M}_+(\mathcal{I})$ be the time-marginal of $\boldsymbol{\mu}$, i.e., $\lambda := (e_1)_\# \boldsymbol{\mu}$ where $e_1 : \mathcal{I} \times \mathbb{T}^d \to \mathcal{I}$, $e_1(t,x) = t$. We claim that λ is a constant multiple of the Lebesgue measure on $\mathcal{I}$. By the disintegration theorem (see, e.g., [AGS08, Theorem 5.3.1]), this implies the first part of the result.

To prove the claim, note that the continuity equation $\mathbb{CE}^{\mathcal{I}}$ yields

$$\int_{\mathcal{I}} \partial_t \phi(t)\,\mathrm{d}\lambda(t) = \int_{\mathcal{I} \times \mathbb{T}^d} \partial_t \phi(t)\,\mathrm{d}\boldsymbol{\mu}(t,x) = 0 \tag{3.26}$$

for all $\phi \in C_c^\infty(\mathcal{I})$.

Write $\mathcal{I} = (a,b)$, let $\psi \in C_c^\infty(\mathcal{I})$ be arbitrary, and set $\bar{\psi} := \frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} \psi(t)\,\mathrm{d}t$. We define $\phi(t) = \int_a^t \psi(s)\,\mathrm{d}s - (t-a)\bar{\psi}$. Then $\phi \in C_c^\infty(\mathcal{I})$ and $\partial_t \phi = \psi - \bar{\psi}$. Applying (3.26) we obtain $\int_{\mathcal{I}} (\psi - \bar{\psi})\,\mathrm{d}\lambda = 0$, which implies the claim, and hence the first part of the result.

To prove the second part, suppose that $\boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ has finite action, and let $\boldsymbol{\nu} \in \mathcal{M}^d\big(\mathcal{I} \times \mathbb{T}^d\big)$ be a solution to the continuity equation (3.23). Applying (3.23) to a test function $\phi \in \mathcal{C}_c^1\big(\mathcal{I}; \mathcal{C}^1(\mathbb{T}^d)\big) \subseteq \mathcal{C}_c^1\big(\mathcal{I} \times \mathbb{T}^d\big)$ such that $\max_{t \in \mathcal{I}} \|\phi_t\|_{\mathcal{C}^1(\mathbb{T}^d)} \leq 1$, we obtain

$$\int_{\mathcal{I} \times \mathbb{T}^d} \partial_t \phi_t \, \mathrm{d}\mu_t \, \mathrm{d}t = - \int_{\mathcal{I} \times \mathbb{T}^d} \nabla \phi \cdot \mathrm{d}\boldsymbol{\nu} \leq |\boldsymbol{\nu}|\big(\mathcal{I} \times \mathbb{T}^d\big) < \infty, \tag{3.27}$$

which implies the desired bound in view of (B.3). □

The next lemma deals with regularity properties for curves of measures with finite action and fine properties for the functionals $\mathbb{A}$ defined in Definition 3.3.10 with $f = f_{\mathrm{hom}}$.

Lemma 3.3.14 (Properties of $\mathbb{A}^{\mathcal{I}}$). *Let $\mathcal{I} \subset \mathbb{R}$ be a bounded open interval. The following statements hold:*

(i) The functionals $(\boldsymbol{\mu}, \boldsymbol{\nu}) \mapsto \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu})$ and $\boldsymbol{\mu} \mapsto \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu})$ are convex.

(ii) Let $\boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$. Let $\{\mathcal{I}_n\}_n$ be a sequence of bounded open intervals such that $\mathcal{I}_n \subseteq \mathcal{I}$ and $|\mathcal{I} \setminus \mathcal{I}_n| \to 0$ as $n \to \infty$. Let $\boldsymbol{\mu}^n \in \mathcal{M}_+(\mathcal{I}_n \times \mathbb{T}^d)$ be such that [1]

$$\boldsymbol{\mu}^n \to \boldsymbol{\mu} \text{ vaguely in } \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d) \quad \text{and} \quad \boldsymbol{\mu}^n(\mathcal{I}_n \times \mathbb{T}^d) \to \boldsymbol{\mu}(\mathcal{I} \times \mathbb{T}^d).$$

as $n \to \infty$. Then:

$$\liminf_{n \to \infty} \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n) \geq \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}). \tag{3.28}$$

If, additionally, $\boldsymbol{\nu} \in \mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ and $\boldsymbol{\nu}^n \in \mathcal{M}^d(\mathcal{I}_n \times \mathbb{T}^d)$ satisfy $\boldsymbol{\nu}^n \to \boldsymbol{\nu}$ vaguely in $\mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$, then we have

$$\liminf_{n \to \infty} \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \geq \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}). \tag{3.29}$$

In particular, the functionals $(\boldsymbol{\mu}, \boldsymbol{\nu}) \mapsto \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu})$ and $\boldsymbol{\mu} \mapsto \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu})$ are lower semicontinuous with respect to (joint) vague convergence.

Proof. (i): Convexity of $\mathbb{A}^{\mathcal{I}}$ follows from convexity of f, f^∞, and the linearity of the constraint (3.23).

(ii): First we show (3.29). Consider the convex energy density $g(\rho, j) := f(\rho, j) + C(\rho + 1)$, which is nonnegative by (3.14). Let $\mathbb{A}_g$ be the corresponding action functional defined using g instead of f. Using the nonnegativity of g, the fact that $|\mathcal{I} \setminus \mathcal{I}_n| \to 0$, and the lower semicontinuity result from [AFP00, Theorem 2.34], we obtain

$$\liminf_{n \to \infty} \mathbb{A}_g^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \geq \liminf_{n \to \infty} \mathbb{A}_g^{\tilde{\mathcal{I}}}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \geq \mathbb{A}_g^{\tilde{\mathcal{I}}}(\boldsymbol{\mu}, \boldsymbol{\nu}).$$

for every open interval $\tilde{\mathcal{I}} \Subset \mathcal{I}$. Taking the supremum over $\tilde{\mathcal{I}}$, we obtain

$$\liminf_{n \to \infty} \mathbb{A}_g^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \geq \mathbb{A}_g^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}). \tag{3.30}$$

[1] We regard measures on $\mathcal{I}_n \times \mathbb{T}^d$ as measures on the bigger set $\mathcal{I} \times \mathbb{T}^d$ by the canonical inclusion.

Since we have $\boldsymbol{\mu}^n\big(\mathcal{I}_n \times \mathbb{T}^d\big) \to \boldsymbol{\mu}\big(\mathcal{I} \times \mathbb{T}^d\big)$ by assumption, the desired result (3.29) follows from (3.30) and the identity

$$\mathbb{A}_g^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) = \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) + C\Big(\boldsymbol{\mu}^n(\mathcal{I}_n \times \mathbb{T}^d) + 1\Big).$$

Let us now show (3.28). Let $\{\boldsymbol{\mu}^n\}_n \subseteq \mathcal{M}_+\big(\mathcal{I}_n \times \mathbb{T}^d\big)$ be such that $\sup_n \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n) < \infty$ and $\boldsymbol{\mu}^n \to \boldsymbol{\mu}$ vaguely in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$. Let $\boldsymbol{\nu}^n \in \mathcal{M}^d(\mathcal{I}_n \times \mathbb{T}^d)$ be such that $(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \in \mathbb{CE}^{\mathcal{I}_n}$ and

$$\mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) \leq \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n) + \frac{1}{n}.$$

From Lemma 3.3.13, we infer that $\mathrm{d}\boldsymbol{\mu}^n(t,x) = \mathrm{d}\mu_t^n(x)\,\mathrm{d}t$ where $(\mu_t^n)_{t\in\mathcal{I}_n}$ is a curve of constant total mass $c_n := \mu_t^n(\mathbb{T}^d)$. Moreover, $M := \sup_n c_n < +\infty$, since $\boldsymbol{\mu}^n \to \boldsymbol{\mu}$ vaguely. The growth condition (3.21) implies that

$$\sup_n |\boldsymbol{\nu}^n|\big(\mathcal{I}^n \times \mathbb{T}^d\big) \leq \frac{1}{c} \sup_n \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_n}(\boldsymbol{\mu}^n) + \frac{C|\mathcal{I}|}{c}\big(M+1\big) < \infty.$$

Hence, by the Banach–Alaoglu theorem, there exists a subsequence of $\{\boldsymbol{\nu}^n\}_n$ (still indexed by n) such that $\boldsymbol{\nu}^n \to \boldsymbol{\nu}$ vaguely in $\mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ and $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$. Another application of Lemma 3.3.13 ensures that $\mathrm{d}\boldsymbol{\mu}(t,x) = \mathrm{d}\mu_t(x)\,\mathrm{d}t$ where $(\mu_t)_{t\in\mathcal{I}}$ is of constant mass $c := \mu_t(\mathbb{T}^d) = \lim_{n\to\infty} c_n$.

We can thus apply the first part of *(ii)* to obtain

$$\mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}) \leq \mathbb{A}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \leq \liminf_{n\to\infty} \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) = \liminf_{n\to\infty} \mathbb{A}^{\mathcal{I}_n}(\boldsymbol{\mu}^n),$$

which ends the proof. □

3.4 The homogenised transport problem

Throughout this section we assume that $(\mathcal{X}, \mathcal{E})$ safisfies Assumption 3.2.1 and F safisfies Assumption 3.2.3.

3.4.1 Discrete representation of continuous measures and vector fields

To define f_{hom}, the following definition turns out to be natural.

Definition 3.4.1 (Representation).

(i) We say that $m \in \mathbb{R}_+^{\mathcal{X}}$ represents $\rho \in \mathbb{R}_+$ if m is $\mathbb{Z}^d$-periodic and

$$\sum_{x\in\mathcal{X}^Q} m(x) = \rho.$$

(ii) We say that $J \in \mathbb{R}_a^{\mathcal{E}}$ represents a vector $j \in \mathbb{R}^d$ if

a) J is $\mathbb{Z}^d$-periodic;

b) J is divergence-free (i.e., $\operatorname{div} J(x) = 0$ for all $x \in \mathcal{X}$);

c) The effective flux of J equals j; i.e., $\mathsf{Eff}(J) = j$, where

$$\mathsf{Eff}(J) := \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} J(x,y)\big(y_\mathsf{z} - x_\mathsf{z}\big). \tag{3.31}$$

We use the (slightly abusive) notation $m \in \mathsf{Rep}(\rho)$ and $J \in \mathsf{Rep}(j)$. We will also write $\mathsf{Rep}(\rho, j) = \mathsf{Rep}(\rho) \times \mathsf{Rep}(j)$.

Remark 3.4.2. Let us remark that $x_\mathsf{z} = 0$ in the formula for $\mathsf{Eff}(J)$, since $x_\mathsf{z} \in \mathcal{X}^Q$.

Remark 3.4.3. The definition of the effective flux $\mathsf{Eff}(J)$ is natural in view of Lemmas 3.4.9 and 3.4.11 below. These results show that a solution to the continuous continuity equation can be constructed starting from a solutions to the discrete continuity equation, with a vector field of the form (3.31).

Clearly, $\mathsf{Rep}(\rho) \neq \emptyset$ for every $\rho \in \mathbb{R}_+$. It is also true, though less obvious, that $\mathsf{Rep}(j) \neq \emptyset$ for every $j \in \mathbb{R}^d$. We will show this in Lemma 3.4.5 using the $\mathbb{Z}^d$-periodicity and the connectivity of $(\mathcal{X}, \mathcal{E})$.

To prove the result, we will first introduce a natural vector field associated to each simple directed path P on $(\mathcal{X}, \mathcal{E})$, For an edge $e = (x,y) \in \mathcal{E}$, the corresponding reversed edge will be denoted by $\overline{e} = (y,x) \in \mathcal{E}$.

Definition 3.4.4 (Unit flux through a path)**.** Let $P := \{x_i\}_{i=0}^m$ be a simple path in $(\mathcal{X}, \mathcal{E})$, thus $e_i = (x_{i-1}, x_i) \in \mathcal{E}$ for $i = 1, \ldots, m$, and $x_i \neq x_k$ for $i \neq k$. The *unit flux through* P is the discrete field $J_P \in \mathbb{R}^{\mathcal{E}}_a$ given by

$$J_P(e) = \begin{cases} 1 & \text{if } e = e_i \text{ for some } i, \\ -1 & \text{if } e = \overline{e}_i \text{ for some } i, \\ 0 & \text{otherwise} \end{cases} \tag{3.32}$$

The *periodic unit flux through* P is the vector field $\tilde{J}_P \in \mathbb{R}^{\mathcal{E}}_a$ defined by

$$\tilde{J}_P(e) = \sum_{z \in \mathbb{Z}^d} J_P(T_z e) \quad \text{for } e \in \mathcal{E}. \tag{3.33}$$

In the next lemma we collect some key properties of these vector fields. Recall the definition of the discrete divergence in (3.137).

Lemma 3.4.5 (Properties of J_P)**.** *Let $P := \{x_i\}_{i=0}^m$ be a simple path in $(\mathcal{X}, \mathcal{E})$.*

(i) The discrete divergence of the associated unit flux $J_P : \mathcal{E} \to \mathbb{R}$ is given by

$$\operatorname{div} J_P = \mathbb{1}_{\{x_0\}} - \mathbb{1}_{\{x_m\}}. \tag{3.34}$$

(ii) The discrete divergence of the periodic unit flux $\tilde{J}_P : \mathcal{E} \to \mathbb{R}$ is given by

$$\operatorname{div} \tilde{J}_P(x) = \mathbb{1}_{\{(x_0)_\mathsf{v}\}}(x_\mathsf{v}) - \mathbb{1}_{\{(x_m)_\mathsf{v}\}}(x_\mathsf{v}), \quad x \in \mathcal{X}. \tag{3.35}$$

In particular, $\operatorname{div} \tilde{J}_P \equiv 0$ iff $(x_0)_\mathsf{v} = (x_m)_\mathsf{v}$.

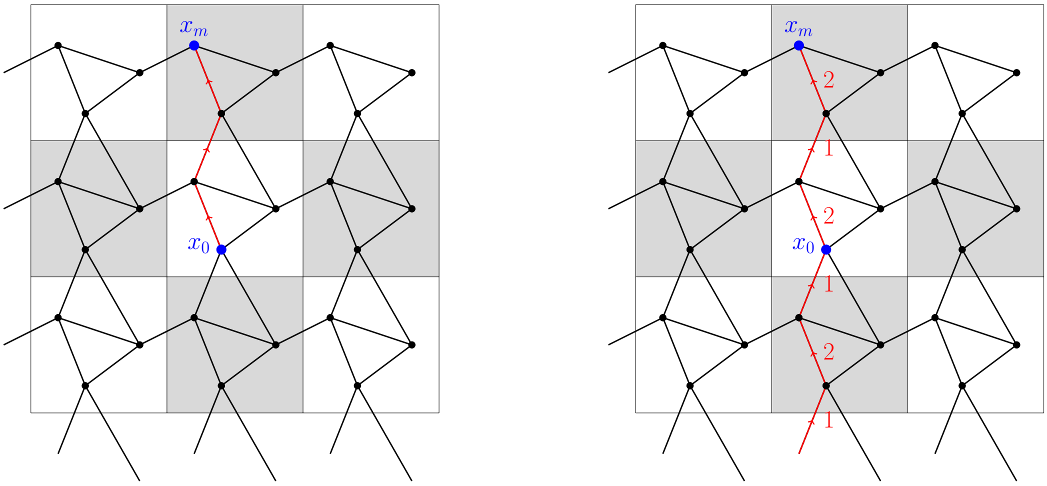

Figure 3.4: In the first figure, in red, the (directed) path P from x_0 to x_m, support of the vector field J_P. In the second one, in red, the support of the vector field $\tilde{J}_P$ and its values.

(iii) The periodic unit flux $\tilde{J}_P : \mathcal{E} \to \mathbb{R}$ *satisfies* $\mathsf{Eff}(\tilde{J}_P) = (x_m)_{\mathsf{z}} - (x_0)_{\mathsf{z}}$.

(iv) For every $j \in \mathbb{R}^d$ *we have* $\mathsf{Rep}(j) \neq \emptyset$.

Proof. (i) is straightforward to check, and *(ii)* is a direct consequence.

To prove *(iii)*, we use the definition of $\tilde{J}_{P,q}$ to obtain

$$\sum_{(x,y)\in\mathcal{E}^Q} \tilde{J}_P(x,y)\Big(y_{\mathsf{z}} - x_{\mathsf{z}}\Big) = \sum_{(x,y)\in\mathcal{E}^Q} \sum_{z\in\mathbb{Z}^d} J_P(T_z x, T_z y)\Big(y_{\mathsf{z}} - x_{\mathsf{z}}\Big)$$
$$= \sum_{(x,y)\in\mathcal{E}} J_P(x,y)\Big(y_{\mathsf{z}} - x_{\mathsf{z}}\Big).$$

By construction, we have

$$\frac{1}{2} \sum_{(x,y)\in\mathcal{E}} J_P(x,y)\Big(y_{\mathsf{z}} - x_{\mathsf{z}}\Big) = \sum_{j=1}^{m} (x_j)_{\mathsf{z}} - (x_{j-1})_{\mathsf{z}} = (x_m)_{\mathsf{z}} - (x_0)_{\mathsf{z}},$$

which yields the result.

For *(iv)*, taking $j = e_i$, we use the connectivity and nonemptyness of $(\mathcal{X}, \mathcal{E})$ to find a simple path connecting some $(v,z) \in \mathcal{X}$ to $(v, z+e_i) \in \mathcal{X}$. The resulting $\tilde{J}_P \in \mathbb{R}^{\mathcal{E}}_a$ is divergence-free by *(ii)* and $\mathsf{Eff}(\tilde{J}_P) = e_i$ by *(iii)*, so that $\tilde{J}_P \in \mathsf{Rep}(e_i)$. For a general $j = \sum_{i=1}^d j_i e_i$ we have $\mathsf{Rep}(j) \supseteq \sum_{i=1}^d j_i \,\mathsf{Rep}(e_i) \neq \emptyset$. $\square$

3.4.2 The homogenised action

We are now in a position to define the homogenised energy density.

Definition 3.4.6 (Homogenised energy density)**.** The *homogenised energy density* $f_{\hom} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is defined by the cell formula

$$f_{\hom}(\rho, j) := \inf\Big\{F(m,J) \ : \ (m,J) \in \mathsf{Rep}(\rho,j)\Big\}. \tag{3.36}$$

For $(\rho, j) \in \mathbb{R}_+ \times \mathbb{R}^d$, we say that $(m, J) \in \mathsf{Rep}(\rho, j)$ is an *optimal representative* if $F(m, J) = f_{\hom}(\rho, j)$. The set of optimal representatives is denoted by

$$\underset{o}{\mathsf{Rep}}(\rho, j).$$

In view of Lemma 3.4.5, the set of representatives $\mathsf{Rep}(\rho, j)$ is nonempty for every $(\rho, j) \in \mathbb{R}_+ \times \mathbb{R}^d$. The next result shows that $\mathsf{Rep}_o(\rho, j)$ is nonempty as well.

Lemma 3.4.7 (Properties of the cell formula)**.** *Let* $(\rho, j) \in \mathbb{R}_+ \times \mathbb{R}^d$*. If* $f_{\hom}(\rho, j) < +\infty$*, then the set of optimal representatives* $\mathsf{Rep}_o(\rho, j)$ *is nonempty, closed, and convex.*

Proof. This follows from the coercivity of F and the direct method of the calculus of variations. □

Lemma 3.4.8 (Properties of $f_{\hom}$ and $f_{\hom}^\infty$)**.** *The following properties hold:*

(i) The functions $f_{\hom}$ *and* $f_{\hom}^\infty$ *are lower semicontinuous and convex.*

(ii) There exist constants $c > 0$ *and* $C < \infty$ *such that, for all* $\rho \geq 0$ *and* $j \in \mathbb{R}^d$*,*

$$f_{\hom}(\rho, j) \geq c|j| - C(\rho + 1), \qquad f_{\hom}^\infty(\rho, j) \geq c|j| - C\rho. \tag{3.37}$$

(iii) The domain $\mathsf{D}(f_{\hom}) \subseteq \mathbb{R}_+ \times \mathbb{R}^d$ *has nonempty interior. In particular, for any pair* (m°, J°) *satisfying* (3.15)*, the element* $(\rho^\circ, j^\circ) \in (0, \infty) \times \mathbb{R}^d$ *defined by*

$$(\rho^\circ, j^\circ) := \Bigg(\sum_{x \in \mathcal{X}^Q} m^\circ(x), \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} J^\circ(x, y) \Big(y_{\mathsf{z}} - x_{\mathsf{z}} \Big) \Bigg) \tag{3.38}$$

belongs to $\mathsf{D}(f_{\hom})^\circ$*.*

Proof. (i): The convexity of $f_{\hom}$ follows from the convexity of F and the affinity of the constraints. Let us now prove lower semicontinuity of $f_{\hom}$.

Take $(\rho, j) \in \mathbb{R}_+ \times \mathbb{R}^d$ and sequences $\{\rho_n\}_n \subseteq \mathbb{R}_+$ and $\{j_n\}_n \subseteq \mathbb{R}^d$ converging to ρ and j respectively. Without loss of generality we may assume that $L := \sup_{n \to \infty} f_{\hom}(\rho_n, j_n) < \infty$. By definition of $f_{\hom}$, there exist $(m_n, J_n) \in \mathsf{Rep}(\rho_n, j_n)$ such that $F(m_n, J_n) \leq f_{\hom}(\rho_n, j_n) + \frac{1}{n}$. From the growth condition (3.14) we deduce that

$$\sup_n \sum_{x \in \mathcal{X}^Q} m_n(x) = \sup_n \rho_n < \infty \quad \text{and} \quad \sup_n \sum_{(x,y) \in \mathcal{E}^Q} |J_n(x, y)| \lesssim 1 + L + \sup_n r_n < \infty.$$

From the Bolzano–Weierstrass theorem we infer subsequential convergence of $\{(m_n, J_n)\}_n$ to some $\mathbb{Z}^d$-periodic pair $(m, J) \in \mathbb{R}_+^{\mathcal{X}} \times \mathbb{R}^{\mathcal{E}}$. Therefore, by lower semicontinuity of F, it follows that

$$F(m, J) \leq \liminf_{n \to \infty} F(m_n, J_n) \leq \liminf_{n \to \infty} f_{\hom}(\rho_n, j_n) \tag{3.39}$$

Since $(m, J) \in \mathsf{Rep}(\rho, j)$, we have $f_{\hom}(\rho, j) \leq F(m, J)$, which yields the desired result. Convexity and lower semicontinuity of $f_{\hom}^\infty$ follow from the definition, see [AFP00, Section 2.6].

(ii) Take $\rho \in \mathbb{R}_+$ and $j \in \mathbb{R}^d$. If $f_{\hom}(\rho, j) = +\infty$, the assertion is trivial, so we assume that $f_{\hom}(\rho, j) < +\infty$. Then there exists a competitor $(m, J) \in \mathsf{Rep}(\rho, j)$ such that $F(m, J) \le f_{\hom}(\rho, j) + 1$. The growth condition (3.14) asserts that

$$F(m, J) \ge c \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)| - C \sum_{x\in\mathcal{X}^Q} m(x) - C$$

Therefore, the claim follows from the fact that

$$R_0 \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)| \gtrsim |j| \quad \text{and} \quad \sum_{x\in\mathcal{X}^Q} m(x) = r,$$

where $R_0 = \max_{(x,y)\in\mathcal{E}} |x_{\mathsf{z}} - y_{\mathsf{z}}|_{\ell^d_\infty}$.

(iii): Let $(m^\circ, J^\circ) \in \mathsf{D}(F)^\circ$ satisfy Assumption 3.2.3, and define $(\rho^\circ, j^\circ) \in (0,\infty) \times \mathbb{R}^d$ by (3.38). For $i = 1, \ldots, d$, let e_i be the coordinate unit vector. Using Lemma 3.4.5 *(*iv) we take $J^i \in \mathsf{Rep}(e_i)$. For $\alpha \in \mathbb{R}$ with $|\alpha|$ sufficiently small, and $\beta = \sum_{i=1}^d \beta_i e_i \in \mathbb{R}^d$ we define

$$\begin{aligned} m_\alpha(x) &:= m^\circ(x) + \frac{\alpha}{\#(\mathcal{X}^Q)} && x \in \mathcal{X}, \\ J_\beta(x,y) &:= J^\circ(x,y) + \sum_{i=1}^d \beta_i J^i(x,y) && (x,y) \in \mathcal{E}. \end{aligned}$$

It follows that $(m_\alpha, J_\beta) \in \mathsf{Rep}(\rho^\circ + \alpha, j^\circ + \beta)$, and therefore, $f_{\hom}(\rho^\circ + \alpha, j^\circ + \beta) \le F(m_\alpha, J_\beta)$. By Assumption 3.2.3, the right-hand side is finite for $|\alpha| + |\beta|$ sufficiently small. This yields the result. □

The homogenised action $\mathbb{A}^{\mathcal{I}}_{\hom}$ can now be defined by taking $f = f_{\hom}$ in Definition 3.3.10.

3.4.3 Embedding of solutions to the discrete continuity equation

For $\varepsilon > 0$ and $z \in \mathbb{Z}$ (or more generally, for $z \in \mathbb{R}$) let $Q_\varepsilon^z := \varepsilon z + [0, \varepsilon)^d \subseteq \mathbb{T}^d$ denote the cube of side-length ε based at εz. For $m \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ and $J \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ we define $\iota_\varepsilon m \in \mathcal{M}_+(\mathbb{T}^d)$ and $\iota_\varepsilon J \in \mathcal{M}^d(\mathbb{T}^d)$ by

$$\iota_\varepsilon m := \varepsilon^{-d} \sum_{x\in\mathcal{X}_\varepsilon} m(x) \mathscr{L}^d|_{Q_\varepsilon^{x_{\mathsf{z}}}} , \tag{3.40a}$$

$$\iota_\varepsilon J := \varepsilon^{-d+1} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \frac{J(x,y)}{2} \left(\int_0^1 \mathscr{L}^d|_{Q_\varepsilon^{(1-s)x_{\mathsf{z}} + s y_{\mathsf{z}}}} \,\mathrm{d}s \right) (y_{\mathsf{z}} - x_{\mathsf{z}}) , \tag{3.40b}$$

The embeddings (3.40) are chosen to ensure that solutions to the discrete continuity equation are mapped to solutions to the continuous continuity equation, as the following result shows.

Lemma 3.4.9. *Let $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$ solve the discrete continuity equation and define $\mu_t = \iota_\varepsilon m_t$ and $\nu_t = \iota_\varepsilon J_t$. Then $(\boldsymbol{\mu}, \boldsymbol{\nu})$ solves the continuity equation (i.e., $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$).*

Proof. Let $\phi : \mathcal{I} \times \mathbb{T}^d \to \mathbb{R}$ be smooth with compact support. Then:

$$\begin{aligned}
&\int_{\mathcal{I}} \int_{\mathbb{T}^d} \nabla\phi \cdot \mathrm{d}\nu_t \,\mathrm{d}t \\
&= \frac{1}{2\varepsilon^d} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \int_{\mathcal{I}} J_t(x,y) \int_0^1 \int_{Q_\varepsilon^{(1-s)x_\mathsf{z}+sy_\mathsf{z}}} \nabla\phi(t,x) \cdot \varepsilon(y_\mathsf{z} - x_\mathsf{z}) \,\mathrm{d}\mathscr{L}^d \,\mathrm{d}s \,\mathrm{d}t \\
&= \frac{1}{2\varepsilon^d} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \int_{\mathcal{I}} J_t(x,y) \int_0^1 \partial_s \bigg(\int_{Q_\varepsilon^{(1-s)x_\mathsf{z}+sy_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d \bigg) \,\mathrm{d}s \,\mathrm{d}t \\
&= \frac{1}{2\varepsilon^d} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \int_{\mathcal{I}} J_t(x,y) \bigg(\int_{Q_\varepsilon^{y_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d - \int_{Q_\varepsilon^{x_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d \bigg) \,\mathrm{d}t.
\end{aligned}$$

On the other hand, the discrete continuity equation yields

$$\begin{aligned}
\int_{\mathcal{I}} \int_{\mathbb{T}^d} \partial_t \phi \,\mathrm{d}\mu_t \,\mathrm{d}t &= \frac{1}{\varepsilon^d} \sum_{x\in\mathcal{X}_\varepsilon} \int_{\mathcal{I}} m_t(x) \partial_t \bigg(\int_{Q_\varepsilon^{x_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d \bigg) \,\mathrm{d}t \\
&= \frac{1}{2\varepsilon^d} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \int_{\mathcal{I}} J_t(x,y) \bigg(\int_{Q_\varepsilon^{x_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d - \int_{Q_\varepsilon^{y_\mathsf{z}}} \phi \,\mathrm{d}\mathscr{L}^d \bigg) \,\mathrm{d}t.
\end{aligned}$$

Comparing both expressions, we obtain the desired identity $\partial_t \boldsymbol{\mu} + \nabla \cdot \boldsymbol{\nu} = 0$ in the sense of distributions. □

The following result provides a useful bound for the norm of the embedded flux.

Lemma 3.4.10. *For $J \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ we have*

$$|\iota_\varepsilon J|(\mathbb{T}^d) \le \frac{\varepsilon R_0 \sqrt{d}}{2} \sum_{(x,y)\in\mathcal{E}_\varepsilon} |J(x,y)|.$$

Proof. This follows immediately from (3.41), since $\mathscr{L}^d\Big(Q_\varepsilon^{(1-s)x_\mathsf{z}+sy_\mathsf{z}}\Big) = \varepsilon^d$ and $|y_\mathsf{z} - x_\mathsf{z}| \le R_0\sqrt{d}$ for $(x,y) \in \mathcal{E}_\varepsilon$. □

Note that both measures in (3.40) are absolutely continuous with respect to the Lebesgue measure. The next result provides an explicit expression for the density of the momentum field. Recall the definition of the shifting operators $\sigma_\varepsilon^{\bar{z}}$ in (3.18).

Lemma 3.4.11 (Density of the embedded flux)**.** *Fix $\varepsilon < \frac{1}{2R_0}$. For $J \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ we have $\iota_\varepsilon J = j_\varepsilon \mathscr{L}^d$ where $j_\varepsilon : \mathbb{T}^d \to \mathbb{R}^d$ is given by*

$$j_\varepsilon(u) = \varepsilon^{-d+1} \sum_{z\in\mathbb{Z}_\varepsilon^d} \chi_{Q_\varepsilon^z}(u) \Bigg(\frac{1}{2} \sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon \\ x_\mathsf{z}=z}} J_u(x,y)\big(y_\mathsf{z} - x_\mathsf{z}\big) \Bigg) \quad \textit{for } u \in \mathbb{T}^d. \tag{3.41}$$

Here, $J_u(x,y)$ is a convex combination of $\Big\{\sigma_\varepsilon^{\bar{z}} J(x,y)\Big\}_{\bar{z}\in\mathbb{Z}_\varepsilon^d}$, i.e.,

$$J_u(x,y) = \sum_{\bar{z}\in\mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,\bar{z}}(x,y) \sigma_\varepsilon^{\bar{z}} J(x,y),$$

where $\lambda_u^{\varepsilon,\bar{z}}(x,y) \ge 0$ and $\sum_{\bar{z}\in\mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,\bar{z}}(x,y) = 1$. Moreover,

$$\lambda_u^{\varepsilon,\bar{z}}(x,y) = 0 \quad \textit{whenever } u \in Q_\varepsilon^{x_\mathsf{z}},\ |\bar{z}|_\infty > R_0 + 1. \tag{3.42}$$

Proof. Fix $\varepsilon < \frac{1}{2R_0}$, let $z \in \mathbb{Z}_\varepsilon^d$ and $u \in Q_\varepsilon^z$. We have

$$\begin{aligned} j_\varepsilon(u) &= \varepsilon^{-d+1} \sum_{(x,y)\in\mathcal{E}_\varepsilon} \frac{J(x,y)}{2} \bigg(\int_0^1 \chi_{Q_\varepsilon^{(1-s)x_\mathsf{z}+sy_\mathsf{z}}}(u) \, \mathrm{d}s \bigg) (y_\mathsf{z} - x_\mathsf{z}) \\ &= \varepsilon^{-d+1} \sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon \\ x_\mathsf{z}=z}} \sum_{\bar z \in \mathbb{Z}_\varepsilon^d} \frac{\sigma_\varepsilon^{\bar z} J(x,y)}{2} \bigg(\int_0^1 \chi_{Q_\varepsilon^{\bar z+(1-s)x_\mathsf{z}+sy_\mathsf{z}}}(u) \, \mathrm{d}s \bigg) (y_\mathsf{z} - x_\mathsf{z}), \end{aligned}$$

which is the desired form (3.41) with

$$\lambda_u^{\varepsilon,\bar z}(x,y) = \bigg(\int_0^1 \chi_{Q_\varepsilon^{\bar z+(1-s)x_\mathsf{z}+sy_\mathsf{z}}}(u) \, \mathrm{d}s \bigg)$$

for $(x,y) \in \mathcal{E}_\varepsilon$ with $x_\mathsf{z} = z$. Since the family of cubes $\big\{ Q_\varepsilon^{\bar z + sy_\mathsf{z} + (1-s)x_\mathsf{z}} \big\}_{\bar z \in \mathbb{Z}_\varepsilon^d}$ is a partition of $\mathbb{T}^d$, it follows that $\sum_{\bar z\in\mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,\bar z}(x,y) = 1$.

To prove the final claim, let $(x,y) \in \mathcal{E}_\varepsilon$ with $x_\mathsf{z} = z$ as above and take $\bar z \in \mathbb{Z}_\varepsilon^d$ with $|\bar z|_\infty > R_0 + 1$. Since $|x_\mathsf{z} - y_\mathsf{z}| \le R_0$, the triangle inequality yields

$$\Big\| \Big(\bar z + s y_\mathsf{z} + (1-s) x_\mathsf{z} \Big) - x_\mathsf{z} \Big\|_\infty \ge \|\bar z\|_\infty - (1-s) \|y_\mathsf{z} - x_\mathsf{z}\|_\infty > 1 \,,$$

for $s \in [0,1]$. Therefore, $u \in Q_\varepsilon^z$ implies $\chi_{Q_\varepsilon^{\bar z+(1-s)x_\mathsf{z}+sy_\mathsf{z}}}(u) = 0$, hence $\lambda_u^{\varepsilon,\bar z}(x,y) = 0$ as desired. □

3.5 Main Results

In this section we present the main result of this paper, which asserts that the discrete action functionals $\mathcal{A}_\varepsilon$ converge to a continous action functional $\mathbb{A} = \mathbb{A}_{\mathrm{hom}}$ with the nontrivial homogenised action density function $f = f_{\mathrm{hom}}$ defined in Section 3.4.

3.5.1 Main convergence result

We are now ready to state our main result. We use the embedding $\iota_\varepsilon : \mathbb{R}_+^{\mathcal{X}_\varepsilon} \to \mathcal{M}_+(\mathbb{T}^d)$ defined in (3.40a). The proof of this result is given in Section 3.7 and 3.8.

Theorem 3.5.1 (Γ-convergence)**.** *Let $(\mathcal{X}, \mathcal{E})$ be a locally finite and $\mathbb{Z}^d$-periodic connected graph of bounded degree (see Assumption 3.2.1). Let $F : \mathbb{R}_+^{\mathcal{X}} \times \mathbb{R}_a^{\mathcal{E}} \to \mathbb{R} \cup \{+\infty\}$ be a cost function satisfying Assumption 3.2.3. Then the functionals $\mathcal{A}_\varepsilon^{\mathcal{I}}$ Γ-converge to $\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}$ as $\varepsilon \to 0$ with respect to the weak (and vague) topology. More precisely:*

*(i) (Γ-**liminf inequality**) Let $\boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$. For any sequence of curves $\{\mathbf{m}^\varepsilon\}_\varepsilon$ with $\mathbf{m}^\varepsilon = (m_t^\varepsilon)_{t\in\mathcal{I}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ vaguely in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ as $\varepsilon \to 0$, we have the lower bound*

$$\liminf_{\varepsilon\to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \ge \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}(\boldsymbol{\mu}). \tag{3.43}$$

*(ii) (Γ-**limsup inequality**) For any $\boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ there exists a sequence of curves $\{\mathbf{m}^\varepsilon\}_\varepsilon$ with $\mathbf{m}^\varepsilon = (m_t^\varepsilon)_{t\in\mathcal{I}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\mathcal{I}\times\mathbb{T}^d)$ as $\varepsilon \to 0$, and we have the upper bound*

$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}(\boldsymbol{\mu}). \tag{3.44}$$

3.5.2 Scaling limits of Wasserstein transport problems

For $1 \leq p < \infty$, recall that the energy density associated to the Wasserstein metric W_p on $\mathbb{R}^d$ is given by $f(\rho, j) = \frac{|j|^p}{\rho^{p-1}}$. This function satisfies the scaling relations $f(\lambda\rho, \lambda j) = \lambda f(\rho, j)$ and $f(\rho, \lambda j) = |\lambda|^p f(\rho, j)$ for $\lambda \in \mathbb{R}$.

In discrete approximations of W_p on a periodic graph $(\mathcal{X}, \mathcal{E})$, it is reasonable to assume analogous scaling relations for the function F, namely $F(\lambda m, \lambda J) = \lambda F(m, J)$ and $F(m, \lambda J) = |\lambda|^p F(m, J)$. The next result shows that if such scaling relations are imposed, we always obtain convergence to W_p with respect to some norm on $\mathbb{R}^d$. This norm does not have to be Hilbertian (even in the case $p = 2$) and is characterised by the cell problem (3.36).

Corollary 3.5.2. *Let $1 \leq p < \infty$, and suppose that F has the following scaling properties for $m \in \mathbb{R}_+^{\mathcal{X}}$ and $j \in \mathbb{R}_a^{\mathcal{E}}$:*

(i) $F(\lambda m, \lambda J) = \lambda F(m, J)$ for all $\lambda \geq 0$;

(ii) $F(m, \lambda J) = |\lambda|^p F(m, J)$ for all $\lambda \in \mathbb{R}$.

Then $f_{\text{hom}}(\rho, j) = \frac{\|j\|^p}{\rho^{p-1}}$ for some norm $\|\cdot\|$ on $\mathbb{R}^d$.

Proof. Fix $\rho > 0$ and $j \in \mathbb{R}^d$. The scaling assumptions imply that

$$f_{\text{hom}}(\lambda\rho, \lambda j) = \lambda f_{\text{hom}}(\rho, j) \quad \text{and} \quad f_{\text{hom}}(\rho, \lambda j) = |\lambda|^p f_{\text{hom}}(\rho, j). \tag{3.45}$$

Consequently,

$$f_{\text{hom}}(\rho, j) = \rho f_{\text{hom}}(1, j/\rho) = \frac{f_{\text{hom}}(1, j)}{\rho^{p-1}}.$$

We claim that $f_{\text{hom}}(1, j) > 0$ whenever $j \neq 0$. Indeed, it follows from (3.37) that $f_{\text{hom}}(1, j) > 0$ whenever $|j|$ is sufficiently large. By homogeneity (3.45), the same holds for every $j \neq 0$. It also follows from (3.45) that $f_{\text{hom}}(1, 0) = 0$.

We can thus define $\|j\| := f_{\text{hom}}(1, j)^{1/p} \in [0, \infty)$. In view of the previous comments, we have $\|0\| = 0$ and $\|j\| > 0$ for all $j \in \mathbb{R}^d \setminus \{0\}$. The homogeneity (3.45) implies that $\|\lambda j\| = |\lambda| \, \|j\|$ for $j \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$.

It remains to show the triangle inequality $\|j_1 + j_2\| \leq \|j_1\| + \|j_2\|$ for $j_1, j_2 \in \mathbb{R}^d$. Without loss of generality we assume that $\|j_1\| + \|j_2\| > 0$. For $\lambda \in (0, 1)$, the convexity of f_{hom} (see Lemma 3.4.8) and the homogeneity (3.45) yield

$$f_{\text{hom}}(1, j_1 + j_2) \leq (1 - \lambda) f_{\text{hom}}\Big(1, \frac{j_1}{1 - \lambda}\Big) + \lambda f_{\text{hom}}\Big(1, \frac{j_2}{\lambda}\Big) = \frac{f_{\text{hom}}(1, j_1)}{(1 - \lambda)^{p-1}} + \frac{f_{\text{hom}}(1, j_2)}{\lambda^{p-1}}.$$

Substitution of $\lambda = \frac{\|j_2\|}{\|j_1\| + \|j_2\|}$ yields the triangle inequality. □

3.5.3 Compactness results

As we frequently need to compare measures with unequal mass in this paper, it is natural to work with the the *Kantorovich–Rubinstein norm*. This metric is closely related to the transport distance W_1; see Appendix B.1.

The following compactness result holds for solutions to the continuity equation with bounded action. As usual, we use the notation $\boldsymbol{\mu}(\mathrm{d}x, \mathrm{d}t) = \mu_t(\mathrm{d}x)\,\mathrm{d}t$.

Theorem 3.5.3 (Compactness under linear growth)**.** *Let $\mathbf{m}^\varepsilon : \mathcal{I} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ be such that*

$$\sup_{\varepsilon>0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) < \infty \quad \text{and} \quad \sup_{\varepsilon>0} \mathbf{m}^\varepsilon(\mathcal{I} \times \mathcal{X}_\varepsilon) < \infty.$$

Then there exists a curve $(\mu_t)_{t\in\mathcal{I}} \in \mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$ such that, up to extracting a subsequence,

(i) $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$;

(ii) $\iota_\varepsilon m_t^\varepsilon \to \mu_t$ weakly in $\mathcal{M}_+(\mathbb{T}^d)$ for almost every $t \in \mathcal{I}$;

(iii) $t \mapsto \mu_t(\mathbb{T}^d)$ is constant.

The proof of this result is given in Section 3.6.

Under a superlinear growth condition on the cost function F, the following stronger compactness result holds.

Assumption 3.5.4 (Superlinear growth)**.** We say that F is of *superlinear growth* if there exists a function $\theta : [0, \infty) \to [0, \infty)$ with $\lim_{t\to\infty} \frac{\theta(t)}{t} = \infty$ and a constant $C \in \mathbb{R}$ such that

$$F(m, J) \geq (m_0 + 1)\theta\left(\frac{J_0}{m_0 + 1}\right) - C(m_0 + 1) \tag{3.46}$$

for all $m \in \mathbb{R}_+^{\mathcal{X}}$ and all $J \in \mathbb{R}_a^{\mathcal{E}}$, where

$$m_0 = \sum_{\substack{x\in\mathcal{X} \\ |x|_{\ell_\infty^d} \leq R}} m(x) \quad \text{and} \quad J_0 = \sum_{(x,y)\in\mathcal{E}^Q} |J(x, y)|, \tag{3.47}$$

with $R = \max\{R_0, R_1\}$ as in Assumption 3.2.3.

Remark 3.5.5. The superlinear growth condition (3.46) implies the linear growth condition (3.14). To see this, suppose that F has superlinear growth. Let $v_0 > 0$ be such that $\theta(v) \geq v$ for $v \geq v_0$. If $\frac{J_0}{m_0+1} \geq v_0$, we have

$$F(m, J) \geq (m_0 + 1)\theta\left(\frac{J_0}{m_0 + 1}\right) - C(m_0 + 1) \geq J_0 - C(m_0 + 1). \tag{3.48}$$

On the other hand, if $\frac{J_0}{m_0+1} < v_0$, the nonnegativity of θ implies that

$$F(m, J) \geq -C(m_0 + 1) \geq \frac{C}{v_0} J_0 - 2C(m_0 + 1). \tag{3.49}$$

Combining (3.48) and (3.49), we have

$$F(m,J) \geq \min\left\{1, \frac{C}{v_0}\right\} J_0 - 2C(m_0+1),$$

which is of the desired form (3.14).

Example 3.5.6. The edge-based costs

$$F(m,J) = \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)|^p$$

have superlinear growth if and only if $1 < p < \infty$ (with $\theta(t) = ct^p$ and $c = |\mathcal{E}^Q|^{1-p}$). Indeed,

$$F(m,J) = \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)|^p \geq cJ_0^p \geq c\frac{J_0^p}{(m_0+1)^{p-1}} = c(m_0+1)\theta\left(\frac{J_0}{m_0+1}\right).$$

Example 3.5.7. The functions (3.16) arising in discretisation of p-Wasserstein distances have superlinear growth if and only if $p > 1$ (with $\theta(t) = t^p$).

To see this, consider the function $G(\alpha,\beta,\gamma) := \frac{1}{p}\frac{|\gamma|^p}{\Lambda(\alpha,\beta)^{p-1}}$. Since G is convex, non increasing in (α,β), and positively one-homogeneous, we obtain

$$\begin{aligned}
F(m,J) &= \sum_{(x,y)\in\mathcal{E}^Q} G\Big(q_{xy}m(x), q_{yx}m(y), J(x,y)\Big) \\
&\geq G\left(\sum_{(x,y)\in\mathcal{E}^Q} q_{xy}m(x), \sum_{(x,y)\in\mathcal{E}^Q} q_{yx}m(y), \sum_{(x,y)\in\mathcal{E}^Q} |J(x,y)|\right) \\
&\geq cG(m_0,m_0,J_0) \geq \frac{c}{p}\frac{J_0^p}{(m_0+1)^{p-1}} = \frac{c}{p}(m_0+1)\theta\left(\frac{J_0}{m_0+1}\right),
\end{aligned}$$

where $c > 0$ depends on R, the maximum degree and the weights q_{xy}.

Theorem 3.5.8 (Compactness under superlinear growth)**.** *Suppose that Assumption 3.5.4 holds. Let $\mathbf{m}^\varepsilon : \mathcal{I} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ be such that*

$$\sup_{\varepsilon>0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) < \infty \quad \textit{and} \quad \sup_{\varepsilon>0} \mathbf{m}^\varepsilon(\mathcal{I}\times\mathcal{X}_\varepsilon) < \infty\,.$$

Then there exists a curve $(\mu_t)_{t\in\mathcal{I}} \in W^{1,1}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$ such that, up to extracting a subsequence,

(i) $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\mathcal{I}\times\mathbb{T}^d)$;

(ii) $\|\iota_\varepsilon m_t^\varepsilon - \mu_t\|_{\mathrm{KR}(\mathbb{T}^d)} \to 0$ uniformly for $t \in \mathcal{I}$;

(iii) $t \mapsto \mu_t(\mathbb{T}^d)$ is constant.

This is proven in Section 3.6.2.

Note that curve $t \mapsto \mu_t \in W^{1,1}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$ can be continuously extended to $\overline{\mathcal{I}}$. Therefore, it is meaningful to assign boundary values to these curves.

3.5.4 Result with boundary conditions

Under Assumption 3.5.4, we are able to obtain the following result on the convergence of dynamical optimal transport problems. Fix $\mathcal{I} = (a,b) \subset \mathbb{R}$ an open interval. Define for $m^a, m^b \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ with $m^a(\mathcal{X}_\varepsilon) = m^b(\mathcal{X}_\varepsilon)$ the minimal action as

$$\mathcal{MA}_\varepsilon^{\mathcal{I}}(m^a, m^b) := \inf \left\{ \mathcal{A}_\varepsilon^{\mathcal{I}}(m) \ : \ m_a = m^a, m_b = m^b) \right\}. \tag{3.50}$$

Similarly, define the minimal homogenized action for $\mu^a, \mu^b \in \mathcal{M}_+(\mathbb{T}^d)$ with $\mu^a(\mathbb{T}^d) = \mu^b(\mathbb{T}^d)$ as

$$\mathbb{MA}_{\text{hom}}^{\mathcal{I}}(\mu^a, \mu^b) := \inf \left\{ \mathbb{A}_{\text{hom}}^{\mathcal{I}}(\mu) \ : \ \mu_a = \mu^a, \mu_b = \mu^b) \right\}. \tag{3.51}$$

Note that in general, both $\mathbb{MA}_{\text{hom}}^{\mathcal{I}}$ and $\mathcal{MA}_\varepsilon^{\mathcal{I}}$ may be infinite even if the two measures have equal mass. Here, the values μ_a and μ_b are well-defined under Assumption 3.5.4 by Theorem 3.5.8. Under linear growth, μ_a and μ_b can still be defined using the trace theorem in BV, but we cannot prove the following statement in that case. We prove this in Section 3.6.3.

Theorem 3.5.9 (Γ-convergence of boundary value problems)**.** *Assume that Assumption 3.5.4 holds. Then the boundary value problems $\mathcal{MA}_\varepsilon^{\mathcal{I}}$ Γ-converge to $\mathbb{MA}_{\text{hom}}^{\mathcal{I}}$ in the weak topology of $\mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}_+(\mathbb{T}^d)$. Precisely:*

(i) For any sequences m_ε^a, $m_\varepsilon^b \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon m_\varepsilon^i \to \mu^i$ weakly in $\mathcal{M}_+(\mathbb{T}^d)$ as $\varepsilon \to 0$ for $i = a, b$, we have

$$\liminf_{\varepsilon\to 0} \mathcal{MA}_\varepsilon^{\mathcal{I}}(m_\varepsilon^a, m_\varepsilon^b) \geq \mathbb{MA}_{\text{hom}}^{\mathcal{I}}(\mu^a, \mu^b). \tag{3.52}$$

(ii) For any $(\mu^a, \mu^b) \in \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}_+(\mathbb{T}^d)$, there exist two sequences $m_\varepsilon^a, m_\varepsilon^b \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon m_\varepsilon^i \to \mu^i$ weakly in $\mathcal{M}_+(\mathbb{T}^d)$ as $\varepsilon \to 0$ for $i = a, b$ and

$$\limsup_{\varepsilon\to 0} \mathcal{MA}_\varepsilon^{\mathcal{I}}(m_\varepsilon^a, m_\varepsilon^b) \leq \mathbb{MA}_{\text{hom}}^{\mathcal{I}}(\mu^a, \mu^b). \tag{3.53}$$

3.6 Proof of compactness and convergence of minimal actions

This section is divided into three sub-parts: in the first one, we prove the general compactness result Theorem 3.5.3, which is valid under the linear growth assumption (3.2.3).

In the second and third part, we assume the stronger superlinear growth condition (3.5.4) and prove the improved compactness result Theorem 3.5.8 and the convergence results for the problems with boundary data, i.e. Theorem 3.5.9.

3.6.1 Compactness under linear growth

The only assumption here is the linear growth condition (3.2.3).

Proof of Theorem 3.5.3. For $\varepsilon > 0$, let $\mathbf{m}^\varepsilon : \mathcal{I} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ be a curve such that

$$\sup_{\varepsilon>0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) < \infty \quad \text{and} \quad \sup_{\varepsilon>0} \mathbf{m}^\varepsilon(\mathcal{I} \times \mathcal{X}_\varepsilon) < \infty. \tag{3.54}$$

We can find a solution to the discrete continuity equation $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$, such that

$$\sup_{\varepsilon>0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) < \infty.$$

Set $(\mu_t^\varepsilon, \nu_t^\varepsilon) := (\iota_\varepsilon m_t^\varepsilon, \iota_\varepsilon J_t^\varepsilon)$, where ι_ε is defined in (3.40). Lemma 3.4.9 implies that $(\boldsymbol{\mu}^\varepsilon, \boldsymbol{\nu}^\varepsilon) \in \mathbb{CE}^{\mathcal{I}}$ for every $\varepsilon > 0$.

Using Lemma 3.4.10, the growth condition (3.14), and the bounds (3.54) on the masses and the action, we infer that

$$\sup_{\varepsilon>0} |\boldsymbol{\nu}^\varepsilon|\Big(\mathcal{I} \times \mathbb{T}^d\Big) \leq \frac{R_0\sqrt{d}}{2} \sup_{\varepsilon>0} \varepsilon \int_{\mathcal{I}} \sum_{(x,y)\in\mathcal{E}_\varepsilon} |J_t^\varepsilon(x,y)| \,\mathrm{d}t < \infty. \tag{3.55}$$

Up to extraction of a subsequence, the Banach–Alaoglu Theorem yields existence of a measure $\bar{\boldsymbol{\nu}} \in \mathcal{M}^d(\overline{\mathcal{I}} \times \mathbb{T}^d)$ such that $\boldsymbol{\nu}^\varepsilon \to \bar{\boldsymbol{\nu}}$ weakly in $\mathcal{M}^d(\overline{\mathcal{I}} \times \mathbb{T}^d)$. It also follows that $|\bar{\boldsymbol{\nu}}|(\overline{\mathcal{I}} \times \mathbb{T}^d) \leq \liminf_{\varepsilon\to 0} |\boldsymbol{\nu}^\varepsilon|(\mathcal{I} \times \mathbb{T}^d) < \infty$; see, e.g., [Bog07, Theorem 8.4.7].

Furthermore, (3.56) and (3.55) imply that the BV-seminorms of $\boldsymbol{\mu}^\varepsilon$ are bounded:

$$\sup_{\varepsilon>0} \|\boldsymbol{\mu}^\varepsilon\|_{\mathrm{BV}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))} \leq \sup_{\varepsilon>0} |\boldsymbol{\nu}^\varepsilon|(\mathcal{I} \times \mathbb{T}^d) < \infty, \tag{3.56}$$

In particular, $\sup_{\varepsilon>0} \boldsymbol{\mu}^\varepsilon(\mathcal{I} \times \mathbb{T}^d) < \infty$. Thus, by another application of the Banach–Alaoglu Theorem, there exists a measure $\boldsymbol{\mu} \in \mathcal{M}_+(\overline{\mathcal{I}} \times \mathbb{T}^d)$ and a subsequence (not relabeled) such that $\boldsymbol{\mu}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\overline{\mathcal{I}} \times \mathbb{T}^d)$.

We claim that $\boldsymbol{\mu}$ does not charge the boundary $(\overline{\mathcal{I}}\setminus\mathcal{I}) \times \mathbb{T}^d$ and that $\boldsymbol{\mu}(\,\mathrm{d}x, \,\mathrm{d}t) = \mu_t(\mathrm{d}x)\,\mathrm{d}t$ for a curve $(\mu_t)_{t\in\mathcal{I}}$ of constant total mass in time. To prove the claim, write $e_1(t,x) := t$, and note that each curve $t \mapsto \mu_t^\varepsilon$ is of constant mass. Therefore, the time-marginals $(e_1)_\# \boldsymbol{\mu}^\varepsilon \in \mathcal{M}_+(\mathcal{I})$ are constant multiples of the Lebesgue measure. Since these measures are weakly-convergent to the time-marginal $(e_1)_\# \boldsymbol{\mu}$, it follows that the latter is also a constant multiple of the Lebesgue measure, which implies the claim.

By what we just proved, $\boldsymbol{\mu}$ can be identified with a measure on the open set $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$. Let $\boldsymbol{\nu}$ be the restriction of $\bar{\boldsymbol{\nu}}$ to $\mathcal{I} \times \mathbb{T}^d$. Since $\boldsymbol{\mu}^\varepsilon$ (resp. $\boldsymbol{\nu}^\varepsilon$) converges vaguely to $\boldsymbol{\mu}$ (resp. $\boldsymbol{\nu}$), it follows that $(\boldsymbol{\mu}, \boldsymbol{\nu})$ belongs to $\mathbb{CE}^{\mathcal{I}}$.

In view of (3.56), we can apply the BV-compactness theorem (see, e.g., [MR15, Theorem B.5.10]) to obtain a further subsequence such that $\|\mu_t^\varepsilon - \mu_t\|_{\mathrm{KR}(\mathbb{T}^d)} \to 0$ for almost every $t \in \mathcal{I}$, and the limiting curve $\boldsymbol{\mu}$ belongs to $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$. Proposition B.1.5 yields $\mu_t^\varepsilon \to \mu_t$ weakly in $\mathcal{M}_+(\mathbb{T}^d)$ for almost every $t \in \mathcal{I}$. □

3.6.2 Uniform compactness under superlinear growth

In the last two sections, we shall work with the stronger growth condition from Assumption 3.5.4.

Remark 3.6.1 (Property of f_{hom}, superlinear case). Let us first observe that under Assumption 3.5.4, one has superlinear growth of f_{hom}:

$$f_{\mathrm{hom}}(\rho, j) \geq \theta\Big(\frac{|j|}{\rho+1}\Big)(\rho+1) - C(\rho+1)\,, \quad \forall \rho \geq 0\,,\ j \in \mathbb{R}^d\,,$$

where we recall $\theta : [0,\infty) \to [0,\infty)$ is such that $\lim_{t\to\infty} \frac{\theta(t)}{t} = +\infty$.

In addition for all $j \neq 0$ we have

$$f_{\hom}^{\infty}(0,j) = \lim_{t\to\infty} \frac{1}{t} f_{\hom}(\rho_0, j_0 + tj) \geq \lim_{t\to\infty} \frac{\theta\left(\frac{|j_0+tj|}{\rho_0+1}\right)(\rho_0+1)}{t} = \infty. \tag{3.57}$$

In particular, if $\mathbb{A}_{\hom}^{\mathcal{I}}(\boldsymbol{\mu},\boldsymbol{\nu}) < \infty$, then $\boldsymbol{\nu} \ll \boldsymbol{\mu} + \mathscr{L}^{d+1}$. Indeed, fix $\boldsymbol{\sigma} \in \mathcal{M}_+(\mathcal{I}\times\mathbb{T}^d)$ as in (3.24) and suppose that $(\boldsymbol{\mu} + \mathscr{L}^{d+1})(A) = 0$ for some $A \subset \mathcal{I}\times\mathbb{T}^d$. By positivity of the measures, this implies that $\boldsymbol{\mu}(A) = \mathscr{L}^{d+1}(A) = 0$, thus by construction

$$\boldsymbol{\mu}^{\perp}(A) = 0 \quad \text{and} \quad \boldsymbol{\nu}(A) = \boldsymbol{\nu}^{\perp}(A)\,.$$

From the first condition and $\boldsymbol{\mu}^{\perp} = \rho^{\perp}\boldsymbol{\sigma}$, we deduce that $\rho^{\perp}(t,x) = 0$ for $\boldsymbol{\sigma}$-a.e. $(t,x)\in A$. From the assumption of finite energy and (3.57), writing $\boldsymbol{\nu}^{\perp} = j^{\perp}\boldsymbol{\sigma}$, we infer that $j^{\perp}(t,x) = 0$ for $\boldsymbol{\sigma}$-a.e. $(t,x)\in A$ as well. It follows that $\boldsymbol{\nu}(A) = \boldsymbol{\nu}^{\perp}(A) = 0$, which proves the claim.

We are ready to prove Theorem 3.5.8.

Proof of Theorem 3.5.8. Let $\{\mathbf{m}^{\varepsilon}\}_{\varepsilon}$ be a sequence of measures such that

$$M := \sup_{\varepsilon} \mathbf{m}^{\varepsilon}(\mathcal{I}\times\mathcal{X}_{\varepsilon}) + 1 < \infty \quad \text{and} \quad A := \sup_{\varepsilon} \mathcal{A}_{\varepsilon}^{\mathcal{I}}(\mathbf{m}^{\varepsilon}) < \infty. \tag{3.58}$$

Thanks to Remark (3.6.1), we have that $\boldsymbol{\nu} \ll \boldsymbol{\mu} + \mathscr{L}^{d+1}$ for all solutions $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ with $\mathbb{A}_{\hom}^{\mathcal{I}}(\boldsymbol{\mu}) < \infty$. Applying Lemma 3.3.13 we can write $\boldsymbol{\mu} = \mathrm{d}t \otimes \mu_t$ and because $\mathscr{L}^{d+1} = \mathrm{d}t\otimes\mathscr{L}^d$, we also have disintegration $\boldsymbol{\nu} = \mathrm{d}t\otimes\nu_t$ with $\nu_t \ll \mu_t + \mathscr{L}^d$ for almost every $t\in\mathcal{I}$.

Moreover, it follows from the definition of $\mathbb{CE}^{\mathcal{I}}$ that, for any test function $\phi \in \mathcal{C}_c^1(\mathcal{I};\mathcal{C}^1(\mathbb{T}^d))$ we have

$$\langle \boldsymbol{\mu}, \partial_t\phi\rangle = -\langle \boldsymbol{\nu}, \nabla\phi\rangle = -\int_{\mathcal{I}} \left\langle \frac{\mathrm{d}\nu_t}{\mathrm{d}(\mu_t + \mathscr{L}^d)}(\mu_t + \mathscr{L}^d), \nabla\phi \right\rangle \mathrm{d}t.$$

This shows that $\mathrm{d}t\otimes\mu_t \in W_{\mathrm{KR}}^{1,1}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$, with weak derivative

$$\partial_t\mu_t = \nabla\cdot\left(\frac{\mathrm{d}\nu_t}{\mathrm{d}(\mu_t+\mathscr{L}^d)}(\mu_t+\mathscr{L}^d)\right) \in \mathrm{KR}(\mathbb{T}^d) \quad \text{for a.e. } t\in\mathcal{I}\,.$$

We are left with showing uniform convergence of $\iota_{\varepsilon} m_t^{\varepsilon} \to \mu_t$ in $\mathrm{KR}(\mathbb{T}^d)$. We claim that the curves $\{t\mapsto \iota_{\varepsilon} m_t^{\varepsilon}\}_{\varepsilon}$ are equicontinuous with respect to the Kantorovich–Rubinstein norm $\|\cdot\|_{\mathrm{KR}(\mathbb{T}^d)}$.

To show the claimed equicontinuity, take $\phi\in\mathcal{C}^1(\mathbb{T}^d)$ and $s,t\in\mathcal{I}$ with $s<t$. Since $(\iota_{\varepsilon}m_t^{\varepsilon}, \iota_{\varepsilon}J_t^{\varepsilon}) \in \mathbb{CE}^{\mathcal{I}}$ we obtain using Lemma 3.4.10,

$$\begin{aligned}\left|\int_{\mathbb{T}^d}\phi\,\mathrm{d}(\iota_{\varepsilon}m_t^{\varepsilon}) - \int_{\mathbb{T}^d}\phi\,\mathrm{d}(\iota_{\varepsilon}m_s^{\varepsilon})\right| &= \left|\int_s^t\int_{\mathbb{T}^d}\nabla\phi\cdot\,\mathrm{d}(\iota_{\varepsilon}J_r^{\varepsilon})\,\mathrm{d}r\right| \\ &\leq \|\nabla\phi\|_{\mathcal{C}(\mathbb{T}^d)}\int_s^t |\iota_{\varepsilon}J_r^{\varepsilon}|(\mathbb{T}^d)\,\mathrm{d}r \\ &\leq \frac{R_0\sqrt{d}}{2}\|\nabla\phi\|_{\mathcal{C}(\mathbb{T}^d)}\int_s^t \sum_{(x,y)\in\mathcal{E}_{\varepsilon}} \varepsilon|J_r^{\varepsilon}(x,y)|\,\mathrm{d}r\,,\end{aligned} \tag{3.59}$$

To estimate the latter integral, we consider for $z \in \mathbb{Z}^d_\varepsilon$ the quantities

$$\mathsf{m}^\varepsilon_r(z) := \sum_{\substack{x\in\mathcal{X}_\varepsilon \\ |x_z - z|_{\ell^d_\infty} \le R}} m^\varepsilon_r(x) \quad \text{and} \quad \mathsf{J}^\varepsilon_r(z) := \sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon \\ x_z = z}} |J^\varepsilon_r(x,y)|.$$

We fix a "velocity threshold" $v_0 > 0$, and split $\mathbb{Z}^d_\varepsilon$ into the low velocity region $\mathcal{Z}_- := \{z \in \mathbb{Z}^d_\varepsilon \ : \ \frac{\varepsilon|\mathsf{J}^\varepsilon_r(z)|}{\mathsf{m}^\varepsilon_r(z)+\varepsilon^d} \le v_0\}$ and its complement $\mathcal{Z}_+ := \mathbb{Z}^d_\varepsilon \setminus \mathcal{Z}_-$. Then:

$$\sum_{z\in\mathcal{Z}_-} \varepsilon \mathsf{J}^\varepsilon_r(z) \le v_0 \sum_{z\in\mathcal{Z}_-} \Big(\mathsf{m}^\varepsilon_r(z) + \varepsilon^d\Big) \le C_R\Big(m^\varepsilon_r(\mathcal{X}_\varepsilon) + 1\Big)v_0, \tag{3.60}$$

where $C_R := (2R+1)^d$. For $z \in \mathcal{Z}_+$ we use the growth condition (3.46) to estimate

$$\begin{aligned}\varepsilon \mathsf{J}^\varepsilon_r(z) &\le \Big(\mathsf{m}^\varepsilon_r(z) + \varepsilon^d\Big)\theta\Big(\frac{\varepsilon \mathsf{J}^\varepsilon_r(z)}{\mathsf{m}^\varepsilon_r(z)+\varepsilon^d}\Big) \sup_{v>v_0}\frac{v}{\theta(v)} \\ &\le \varepsilon^d\bigg(F\Big(\frac{\tau^z_\varepsilon m}{\varepsilon^d}, \frac{\tau^z_\varepsilon J}{\varepsilon^{d-1}}\Big) + C\Big(\frac{\mathsf{m}^\varepsilon_r(z)}{\varepsilon^d} + 1\Big)\bigg)\sup_{v>v_0}\frac{v}{\theta(v)}.\end{aligned}$$

Since (3.46) implies non-negativity of the term in brackets, we obtain

$$\begin{aligned}\sum_{z\in\mathcal{Z}_+} \varepsilon \mathsf{J}^\varepsilon_r(z) &\le \sum_{z\in\mathbb{T}^d} \varepsilon^d\bigg(F\Big(\frac{\tau^z_\varepsilon m}{\varepsilon^d}, \frac{\tau^z_\varepsilon J}{\varepsilon^{d-1}}\Big) + C\Big(\frac{\mathsf{m}^\varepsilon_r(z)}{\varepsilon^d} + 1\Big)\bigg)\sup_{v>v_0}\frac{v}{\theta(v)} \\ &\le \mathcal{F}_\varepsilon(m^\varepsilon_r, J^\varepsilon_r) + C\Big(m^\varepsilon_r(\mathcal{X}_\varepsilon)+1\Big)\sup_{v>v_0}\frac{v}{\theta(v)}.\end{aligned} \tag{3.61}$$

Integrating in time, we combine (3.60) and (3.61) with (3.58) to obtain

$$\begin{aligned}&\int_s^t \sum_{(x,y)\in\mathcal{E}_\varepsilon} \varepsilon|J^\varepsilon_r(x,y)|\,\mathrm{d}r = \int_s^t \sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon \mathsf{J}^\varepsilon_r(z)\,\mathrm{d}r \le g(t-s), \\ &\text{where} \quad g(r) := \inf_{v_0>0}\bigg\{rC_RMv_0 + \Big(A + C(M+|\mathcal{I}|)\Big)\sup_{v>v_0}\frac{v}{\theta(v)}\bigg\}.\end{aligned} \tag{3.62}$$

Combining (3.59) and (3.62) we conclude that

$$\begin{aligned}\sup_{\varepsilon>0}\|\iota_\varepsilon m^\varepsilon_t - \iota_\varepsilon m^\varepsilon_s\|_{\mathrm{KR}(\mathbb{T}^d)} &\le \sup_{\varepsilon>0}\sup_{\|\phi\|_{C^1(\mathbb{T}^d)}\le 1}\bigg|\int_{\mathbb{T}^d}\phi\,\mathrm{d}(\iota_\varepsilon m^\varepsilon_t) - \int_{\mathbb{T}^d}\phi\,\mathrm{d}(\iota_\varepsilon m^\varepsilon_s)\bigg| \\ &\le \frac{R_0\sqrt{d}}{2}g(t-s).\end{aligned}$$

To prove the claimed equicontinuity, it suffices to show that $g(r) \to 0$ as $r \to 0$. But this follows from the growth properties of θ by picking, e.g., $v_0 := r^{-1/2}$.

Of course the masses are uniformly bounded in ε and t. The Arzela-Ascoli theorem implies that every subsequence has a subsequence converging uniformly in $\big(\mathcal{M}_+(\mathbb{T}^d), \|\cdot\|_{\mathrm{KR}}\big)$. □

3.6.3 The boundary value problems under superlinear growth

The last part of this section is devoted to the proof of the convergence of boundary value problems, under the assumption of superlinear growth, i.e. Theorem 3.5.9. The proof is a straightforward consequence of the stronger compactness result Theorem 3.5.8 (and the general convergence result Theorem 3.5.1) proved in the previous section, which ensures the stability of the boundary conditions as well. We fix $\mathcal{I} = (a,b)$.

Proof of Theorem 3.5.9. We shall prove the upper and the lower bound.

Liminf inequality. Pick any $\iota_\varepsilon m_a^\varepsilon \to \mu^a$, $\iota_\varepsilon m_b^\varepsilon \to \mu^b$ weakly in $\mathcal{M}_+(\mathbb{T}^d)$, and let $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon^\mathcal{I}$ with the same boundary data such that

$$\lim_{\varepsilon\to 0} \mathcal{A}_\varepsilon^\mathcal{I}(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) = \lim_{\varepsilon\to 0} \mathcal{MA}_\varepsilon^\mathcal{I}(m_a^\varepsilon, m_b^\varepsilon) < \infty\,.$$

By Theorem 3.5.8, there exists a (non-relabeled) subsequence of $\mathbf{m}^\varepsilon$ such that $\|\iota_\varepsilon m_t^\varepsilon - \mu_t\|_{\mathrm{KR}} \to 0$, uniformly for $t \in \overline{\mathcal{I}}$. In particular, $\mu_a = \mu^a$, $\mu_b = \mu^b$. We can then apply the lower bound of Theorem 3.5.1, and conclude

$$\mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b) \le \mathbb{A}_{\hom}^\mathcal{I}(\boldsymbol{\mu}) \le \liminf_{\varepsilon\to\infty} \mathcal{MA}_\varepsilon^\mathcal{I}(m_a^\varepsilon, m_b^\varepsilon)\,.$$

Limsup inequality. Fix $\mu^a, \mu^b \in \mathcal{M}_+(\mathbb{T}^d)$ such that $\mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b) < \infty$. By the definition of $\mathbb{MA}_{\hom}^\mathcal{I}$ and the lower semicontinuity of $\mathbb{A}_{\hom}$ (Lemma 3.3.14), there exists $\boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ with $\mathbb{A}_{\hom}^\mathcal{I}(\mu) = \mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b)$ and $\mu_a = \mu^a, \mu_b = \mu^b$.

We can then apply Theorem 3.5.1 and find a recovery sequence $(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon^\mathcal{I}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly and

$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon^\mathcal{I}(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \le \mathbb{A}_{\hom}^\mathcal{I}(\boldsymbol{\mu}) = \mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b)\,.$$

By the improved compactness result Theorem 3.5.8, $\iota_\varepsilon m_t^\varepsilon \to \mu_t$ in $\mathrm{KR}(\mathbb{T}^d)$ for every $t \in \overline{\mathcal{I}}$, in particular for $t = a, b$. This allows us to conclude

$$\limsup_{\varepsilon\to 0} \mathcal{MA}_\varepsilon^\mathcal{I}(m_a^\varepsilon, m_b^\varepsilon) \le \mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b)\,, \qquad \text{and} \qquad \iota_\varepsilon m_i^\varepsilon \to \mu^i \text{ weakly}$$

for $i = a, b$, which is sought recovery sequence for $\mathbb{MA}_{\hom}^\mathcal{I}(\mu^a, \mu^b)$. □

Remark 3.6.2. It is instructive to see that under the simple linear growth condition (3.2.3), the above written proof cannot be carried out. Indeed, by the lack of compactness in $W^{1,1}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$ (but rather only in BV by Theorem 3.5.3), we are not able to ensure stability at the level of the initial data, i.e. in general, $\mu_a \neq \mu^a$ (and similarly for $t = b$).

3.7 Proof of the Γ-liminf inequality

In this section we present the proof of the Γ-liminf inequality in our main result, Theorem 3.5.1. The proof relies on two key ingredients. The first one is a partial regularisation result for discrete measures of bounded energy, which is stated in Proposition 3.7.1 and proved in Section 3.7.1 below. The second ingredient is a lower bound of the energy under partial regularity conditions on the involved measures (Proposition 3.7.4). The proof of the Γ-liminf inequality in Theorem 3.5.1, which combines both ingredients, is given at the end of this section.

First we state the regularisation result. Recall the Kantorovich–Rubinstein norm $\|\cdot\|_{\mathrm{KR}}$. see Appendix B.1.

Proposition 3.7.1 (Discrete Regularisation). *Fix $\varepsilon < \frac{1}{2R_0}$ and let $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$ be a solution to the discrete continuity equation satisfying*

$$M := m_0(\mathcal{X}_\varepsilon) < \infty \quad \text{and} \quad A := \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J}) < \infty.$$

Then, for any $\eta > 0$ there exists an interval $\mathcal{I}^\eta \subset \mathcal{I} := (0,T)$ with $|\mathcal{I} \setminus \mathcal{I}^\eta| \le \eta$ and a solution $(\tilde{\mathbf{m}}, \tilde{\mathbf{J}}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}^\eta}$ such that:

(i) the following approximation properties hold:

$$\begin{aligned} &(\textit{measure approximation}) && \|\iota_\varepsilon(\tilde{\mathbf{m}} - \mathbf{m})\|_{\mathrm{KR}(\overline{\mathcal{I}^\eta} \times \mathbb{T}^d)} \le \eta, && (3.63a)\\ &(\textit{energy approximation}) && \mathcal{A}_\varepsilon^{\mathcal{I}^\eta}(\tilde{\mathbf{m}}, \tilde{\mathbf{J}}) \le \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}, \mathbf{J}) + \eta. && (3.63b)\end{aligned}$$

(ii) the following regularity properties hold, uniformly for any $t \in \mathcal{I}^\eta$ and any $z \in \mathbb{T}_\varepsilon^d$:

$$\begin{aligned} &(\textit{boundedness}) && \left\|\tilde{m}_t\right\|_{\ell^\infty(\mathcal{X}_\varepsilon)} + \varepsilon\left\|\tilde{J}_t\right\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le C_B \varepsilon^d, && (3.64a)\\ &(\textit{time-reg.}) && \left\|\operatorname{div} \tilde{J}_t\right\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_T \varepsilon^d, && (3.64b)\\ &(\textit{space-reg.}) && \left\|\sigma_\varepsilon^z \tilde{m}_t - \tilde{m}_t\right\|_{\ell^\infty(\mathcal{X}_\varepsilon)} + \varepsilon\left\|\sigma_\varepsilon^z \tilde{J}_t - \tilde{J}_t\right\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le C_S |z| \varepsilon^{d+1}, && (3.64c)\\ &(\textit{domain reg.}) && \left(\frac{\tau_\varepsilon^z \tilde{m}_t}{\varepsilon^d}, \frac{\tau_\varepsilon^z \tilde{J}_t}{\varepsilon^{d-1}}\right) \in K. && (3.64d)\end{aligned}$$

The constants $C_B, C_T, C_S < \infty$ and the compact set $K \subseteq \mathsf{D}(F)^\circ$ depend on η, M and A, but not on ε.

Remark 3.7.2. The ℓ^∞-bounds in (3.64a) are explicitly stated for the sake of clarity, although they are implied by the compactness of the set K in (3.64d).

Since $(\tilde{\mathbf{m}}, \tilde{\mathbf{J}}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}^\eta}$, inequality (3.64b) in effect bounds $\left\|\partial_t \tilde{m}_t\right\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_T \varepsilon^d$.

In the next result, we start by showing how to construct $\mathbb{Z}^d$-periodic solutions to the static continuity equation by superposition of unit fluxes. Additionally, we can build these solutions with vanishing effective flux and ensure good ℓ^∞-bounds.

Lemma 3.7.3 (Periodic solutions to the divergence equation). *Let $g : \mathcal{X} \to \mathbb{R}$ be a $\mathbb{Z}^d$-periodic function with $\sum_{x \in \mathcal{X}^Q} g(x) = 0$. There exists a $\mathbb{Z}^d$-periodic discrete vector field $J : \mathcal{E} \to \mathbb{R}$ satisfying*

$$\operatorname{div} J = g, \quad \mathsf{Eff}(J) = 0, \quad \text{and} \quad \|J\|_{\ell_\infty(\mathcal{E}^Q)} \le \tfrac{1}{2}\|g\|_{\ell_1(\mathcal{X}^Q)}.$$

Proof. For any $v, w \in \mathbb{V}$, fix a simple path P^{vw} in $(\mathcal{X}, \mathcal{E})$ connecting $(0, v)$ and $(0, w)$. Let $\tilde{J}_{vw} := \tilde{J}_{P^{vw}}$ be the associated periodic unit flux defined in (3.33). Since $\sum_{v\in\mathbb{V}} g(0, v) = 0$, we can pick a coupling Γ between the negative part and the positive part of g. More precisely, we may pick a function $\Gamma : \mathbb{V} \times \mathbb{V} \to \mathbb{R}_+$ with $\sum_{v,w\in\mathbb{V}} \Gamma(v, w) = \|g\|_{\ell_1(\mathcal{X}^Q)}$ such that

$$\sum_{w\in\mathbb{V}} \Gamma_{vw} = g_-(0, v) \quad \text{for } v \in \mathbb{V}, \quad \text{and} \quad \sum_{v\in\mathbb{V}} \Gamma_{vw} = g_+(0, w) \quad \text{for } w \in \mathbb{V}.$$

We then define

$$J := \sum_{v,w\in\mathbb{V}} \Gamma_{vw} \tilde{J}_{vw}.$$

It is straightforward to verify using Lemma 3.4.5 that J has the three desired properties. □

The following result states the desired relation between the functionals $\mathcal{F}_\varepsilon$ and $\mathbb{F}_{\text{hom}}$ under suitable regularity conditions for the measures involved. These regularity conditions are consistent with the regularity properties obtained in Proposition 3.7.1.

Proposition 3.7.4 (Energy lower bound for regular measures). *Let $C_B, C_T, C_S < \infty$ and let $K \subseteq \mathsf{D}(F)^\circ$ be a compact set. There exists a threshold $\varepsilon_0 > 0$ and a constant $C < \infty$ such that the following implication holds for any $\varepsilon < \varepsilon_0$: if $m \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ and $J \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ satisfy the regularity properties (3.64a)-(3.64d) then we have the energy bound*

$$\mathbb{F}_{\text{hom}}(\iota_\varepsilon m, \iota_\varepsilon J) \leq \mathcal{F}_\varepsilon(m, J) + C\varepsilon.$$

Proof. Recall from (3.41) that $\iota_\varepsilon m = \rho \mathscr{L}^d$ and $\iota_\varepsilon J = j \mathscr{L}^d$, where, for $\bar{z} \in \mathbb{Z}_\varepsilon^d$ and $u \in Q_\varepsilon^{\bar{z}}$,

$$\rho(u) := \varepsilon^{-d} \sum_{\substack{x\in\mathcal{X}_\varepsilon \\ x_{\mathsf{z}}=\bar{z}}} m(x) \quad \text{and} \quad j(u) := \frac{1}{2\varepsilon^{d-1}} \sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon \\ x_{\mathsf{z}}=\bar{z}}} J_u(x, y)\big(y_{\mathsf{z}} - x_{\mathsf{z}}\big),$$

where $J_u(x, y)$ is a convex combination of $\left\{J\big(T_\varepsilon^z x, T_\varepsilon^z y\big)\right\}_{z\in\mathbb{Z}_\varepsilon^d}$, i.e.,

$$J_u(x, y) = \sum_{z\in\mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,z}(x, y) J\big(T_\varepsilon^z x, T_\varepsilon^z y\big),$$

where $\lambda_u^{\varepsilon,\bar{z}}(x, y) \geq 0$, $\sum_{z\in\mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,z}(x, y) = 1$, and $\lambda_u^{\varepsilon,z}(x, y) = 0$ whenever $|z| > R_0$.

Step 1. Construction of a representative. Fix $\bar{z} \in \mathbb{Z}_\varepsilon^d$ and $u \in Q_\varepsilon^{\bar{z}}$. Our first goal is to construct a representative

$$\left(\frac{\widehat{m}_u}{\varepsilon^d}, \frac{\widehat{J}_u}{\varepsilon^{d-1}}\right) \in \mathsf{Rep}\,\big(\rho(u), j(u)\big).$$

For this purpose we define candidates $\widehat{m}_u \in \mathbb{R}_+^{\mathcal{X}}$ and $\tilde{J}_u \in \mathbb{R}_a^{\mathcal{E}}$ as follows. We take the values of m and J_u in the ε-cube at $\bar{z}$, and insert these values at every cube in $(\mathcal{X}, \mathcal{E})$, so that the result is $\mathbb{Z}^d$-periodic. In formulae:

$$\begin{aligned} \widehat{m}_u(z, v) &:= m(\varepsilon\bar{z}, v) && \text{for } (z, v) \in \mathcal{X} \\ \tilde{J}_u\Big((z, v), (z', v')\Big) &:= J_u\Big((\varepsilon\bar{z}, v), (\varepsilon(\bar{z} + z' - z), v')\Big) && \text{for } \Big((z, v), (z', v')\Big) \in \mathcal{E}. \end{aligned}$$

see Figure 3.5.

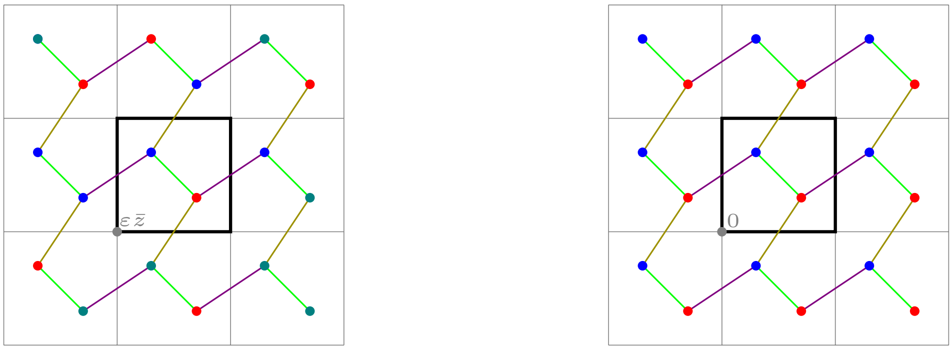

Figure 3.5: On the left, using different colors for different values, the measures m and J_u. On the right, the corresponding $\widehat{m}_u$ and $\tilde{J}_u$, for $u \in Q_\varepsilon^{\bar{z}}$.

We emphasize that the right-hand side does not depend on z, hence m_u and $\tilde{J}_u$ are $\mathbb{Z}^d$-periodic. Our construction also ensures that

$$\varepsilon^{-d} \sum_{x \in \mathcal{X}^Q} \widehat{m}_u(x) = \rho(u),$$

hence $\varepsilon^{-d}\widehat{m}_u \in \mathsf{Rep}\big(\rho(u)\big)$. However, the vector field $\varepsilon^{-(d-1)}\tilde{J}_u$ does (in general) not belong to $\mathsf{Rep}\big(j(u)\big)$: indeed, while $\tilde{J}_u$ has the desired effective flux (i.e., $\mathsf{Eff}\big(\varepsilon^{-(d-1)}\tilde{J}_u\big) = j(u)$), $\tilde{J}_u$ is not (in general) divergence-free.

To remedy this issue, we introduce a *corrector field* $\bar{J}_u$, i.e., an anti-symmetric and $\mathbb{Z}^d$-periodic function $\bar{J}_u : \mathcal{E} \to \mathbb{R}$ satisfying

$$\mathsf{div}\, \bar{J}_u = -\,\mathsf{div}\, \tilde{J}_u, \quad \mathsf{Eff}(\bar{J}_u) = 0, \quad \text{and} \quad \big\|\bar{J}_u\big\|_{\ell^\infty(\mathcal{E}^Q)} \le \tfrac{1}{2}\big\|\mathsf{div}\, \tilde{J}_u\big\|_{\ell^1(\mathcal{X}^Q)}. \tag{3.65}$$

The existence of such a vector field is guaranteed by Lemma 3.7.3. It immediately follows that $\widehat{J}_u := \tilde{J}_u + \bar{J}_u$ satisfies $\mathsf{div}\, \widehat{J}_u = 0$ and $\mathsf{Eff}\big(\varepsilon^{-(d-1)}\widehat{J}_u\big) = j(u)$, thus

$$\frac{\widehat{J}_u}{\varepsilon^{d-1}} := \frac{\tilde{J}_u + \bar{J}_u}{\varepsilon^{d-1}} \in \mathsf{Rep}\big(j_u\big).$$

Step 2. Density comparison. We will now use the regularity assumptions (3.64a)-(3.64d) to show that the representative $(\widehat{m}_u, \widehat{J}_u)$ is not too different from the shifted density $(\tau_{\bar{z}} m, \tau_{\bar{z}} J)$. Indeed, for $x = (z, v) \in \mathcal{X}$ with $|z| \le R_0$ we obtain using (3.64c),

$$|\tau_\varepsilon^{\bar{z}} m(x) - \widehat{m}_u(x)| = \Big|m\big(\varepsilon(\bar{z}+z), v\big) - m\big(\varepsilon\bar{z}, v\big)\Big| \le C_S \varepsilon^{d+1}|z|. \tag{3.66}$$

Let us now turn to the momentum field. For $(x,y) = \big((z,v),(z',v')\big) \in \mathcal{E}$ with $|z|, |z'| \le R_1$, we have, using (3.64c),

$$\begin{aligned}
&\Big|\tau_\varepsilon^{\bar{z}} J(x,y) - \tilde{J}_u(x,y)\Big| \\
&= \bigg|J\Big(\big(\varepsilon(\bar{z}+z), v\big), \big(\varepsilon(\bar{z}+z'), v'\big)\Big) - J_u\Big(\big(\varepsilon\bar{z}, v\big), \big(\varepsilon(\bar{z}+z'-z), v'\big)\Big)\bigg| \\
&= \bigg|\sum_{\tilde{z} \in \mathbb{Z}_\varepsilon^d} \lambda_u^{\varepsilon,\tilde{z}}(x,y) \Big\{ J\Big(\big(\varepsilon(\bar{z}+z), v\big), \big(\varepsilon(\bar{z}+z'), v'\big)\Big) \\
&\qquad\qquad - J\Big(\big(\varepsilon(\bar{z}+\tilde{z}), v\big), \big(\varepsilon(\bar{z}+\tilde{z}+z'-z), v'\big)\Big)\Big\}\bigg| \\
&\le C_S \varepsilon^d |z - \tilde{z}| \le R_0 C_S \varepsilon^d.
\end{aligned}$$

Moreover, using (3.65), (3.64c), and (3.64b), we obtain

$$|\bar{J}_u(x,y)| \le \tfrac{1}{2}\|\operatorname{div}\tilde{J}_u\|_{\ell^1(\mathcal{X}^Q)} \le C_T\Big(\|\operatorname{div} J\|_{\ell^\infty(\mathcal{E}_\varepsilon)} + \varepsilon^d\Big) \le C\varepsilon^d,$$

for some $C<\infty$ not depending on ε. Combining these bounds we obtain

$$|\tau_\varepsilon^{\bar{z}} J(x,y) - \widehat{J}_u(x,y)| \le |\tau_\varepsilon^{\bar{z}} J(x,y) - \tilde{J}_u(x,y)| + |\bar{J}_u(x,y)| \le C\varepsilon^d. \tag{3.67}$$

Step 3. Energy comparison. Since $\left(\frac{\tau_\varepsilon^{\bar{z}} m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar{z}} J}{\varepsilon^{d-1}}\right) \in K$ by assumption, it follows from (3.66) and (3.67) that $\left(\frac{\widehat{m}_u}{\varepsilon^d}, \frac{\widehat{J}_u}{\varepsilon^{d-1}}\right) \in K'$ for $\varepsilon > 0$ sufficiently small. Here K is a compact set, possibly slightly larger than K, contained in $\mathsf{D}(\mathcal{F})^\circ$.

Since F is convex, it is Lipschitz continuous on compact subsets in the interior of its domain. In particular, it is Lipschitz continuous on K'. Therefore, there exists a constant $C_L<\infty$ depending on $\mathcal{F}$ and K' such that

$$\begin{aligned}
\mathcal{F}\left(\frac{\tau_\varepsilon^{\bar{z}} m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar{z}} J}{\varepsilon^{d-1}}\right) &\ge \mathcal{F}\left(\frac{\widehat{m}_u}{\varepsilon^d}, \frac{\widehat{J}_u}{\varepsilon^{d-1}}\right) - C_L\left(\left\|\frac{\tau_\varepsilon^{\bar{z}} m - \widehat{m}_u}{\varepsilon^d}\right\|_{\ell^\infty_{R_0}(\mathcal{X})} + \left\|\frac{\tau_\varepsilon^{\bar{z}} J - \widehat{J}_u}{\varepsilon^{d-1}}\right\|_{\ell^\infty_{R_0}(\mathcal{E})}\right) \\
&\ge \mathcal{F}\left(\frac{\widehat{m}_u}{\varepsilon^d}, \frac{\widehat{J}_u}{\varepsilon^{d-1}}\right) - C\varepsilon \\
&\ge f_{\hom}\big(\rho(u), j(u)\big) - C\varepsilon,
\end{aligned}$$

with $C<\infty$ depending on C_L, C_S, C_T, and R_1, but not on ε. Here, the subscript R_0 in $\ell^\infty_{R_0}(\mathcal{E})$ and $\ell^\infty_{R_0}(\mathcal{X})$ indicates that only elements with $|x_{\mathsf{z}}| \le R_1$ are considered.

Integration over $Q_\varepsilon^{\bar{z}}$ followed by summation over $\bar{z} \in \mathbb{Z}^d_\varepsilon$ yields

$$\begin{aligned}
\mathcal{F}_\varepsilon(m,J) = \varepsilon^d \sum_{\bar{z}\in\mathbb{Z}^d_\varepsilon} \mathcal{F}\left(\frac{\tau_\varepsilon^{\bar{z}} m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar{z}} J}{\varepsilon^{d-1}}\right) &\ge \sum_{\bar{z}\in\mathbb{Z}^d_\varepsilon} \int_{Q_\varepsilon^{\bar{z}}} \Big(f_{\hom}\big(\rho(u), j(u)\big) - C\varepsilon\Big)\,\mathrm{d}u \\
&= \int_{\mathbb{T}^d} f_{\hom}\big(\rho(u), j(u)\big)\,\mathrm{d}u - C\varepsilon = \mathbb{F}_{\hom}(\iota_\varepsilon m, \iota_\varepsilon J) - C\varepsilon,
\end{aligned}$$

which is the desired result. □

We are now ready to give the proof of the Γ-liminf inequality in our main result, Theorem 3.5.1.

Proof of Theorem 3.5.1 (liminf inequality). Let $\boldsymbol{\mu} \in \mathcal{M}_+\big(\mathcal{I}\times\mathbb{T}^d\big)$ and let $(m_t^\varepsilon)_{t\in\mathcal{I}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ be such that the induced measures $\mathbf{m}^\varepsilon \in \mathcal{M}_+\big(\mathcal{I}\times\mathcal{X}_\varepsilon\big)$ defined by $\mathrm{d}\mathbf{m}^\varepsilon(t,x) = \mathrm{d}m_t^\varepsilon(x)\,\mathrm{d}t$ satisfy $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ vaguely in $\mathcal{M}_+(\mathcal{I}\times\mathbb{T}^d)$ as $\varepsilon\to 0$. Observe that

$$M := \sup_{\varepsilon>0} \mathbf{m}^\varepsilon\big(\mathcal{I}\times\mathcal{X}_\varepsilon\big) < \infty.$$

Without loss of generality, we may assume that

$$A := \sup_{\varepsilon>0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) < \infty.$$

Step 1 (Regularisation): Fix $\eta > 0$. Let $(J_t^\varepsilon)_{t\in\mathcal{I}} \subseteq \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ be an approximately optimal discrete vector field, i.e.,

$$(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \in \mathcal{CE}_\varepsilon^{\mathcal{I}} \quad \text{and} \quad \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) \le \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) + \eta. \tag{3.68}$$

Using Proposition 3.7.1 we take an interval $\mathcal{I}^\eta \subset \mathcal{I} := (0,T)$, $|\mathcal{I}\backslash\mathcal{I}^\eta| \le \eta$ and an approximating pair $(\tilde{\mathbf{m}}^\varepsilon, \tilde{\mathbf{J}}^\varepsilon) \in \mathcal{CE}_\varepsilon^{I_\eta}$ satisfying

$$\|\iota_\varepsilon(\tilde{\mathbf{m}}^\varepsilon - \mathbf{m}^\varepsilon)\|_{\mathrm{KR}(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d)} \le \eta \quad \text{and} \quad \mathcal{A}_\varepsilon^{\mathcal{I}^\eta}(\tilde{\mathbf{m}}^\varepsilon, \tilde{\mathbf{J}}^\varepsilon) \le \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon, \mathbf{J}^\varepsilon) + \eta, \tag{3.69}$$

together with the regularity properties (3.64) for some constants $C_B, C_T, C_S < \infty$ and a compact set $K \subseteq \mathsf{D}(F)^\circ$ depending on η, but not on ε. By virtue of these regularity properties, we may apply Proposition 3.7.4 to $(\tilde{\mathbf{m}}^\varepsilon, \tilde{\mathbf{J}}^\varepsilon)$. This yields

$$\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}^\eta}(\iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon, \iota_\varepsilon\tilde{\mathbf{J}}^\varepsilon) = \int_{\mathcal{I}^\eta} \mathbb{F}_{\mathrm{hom}}(\iota_\varepsilon\tilde{m}_t^\varepsilon, \iota_\varepsilon\tilde{J}_t^\varepsilon)\,\mathrm{d}t \le \int_{\mathcal{I}^\eta} \mathcal{F}_\varepsilon(\tilde{m}_t^\varepsilon, \tilde{J}_t^\varepsilon)\,\mathrm{d}t + C\varepsilon, \tag{3.70}$$

with $C < \infty$ depending on η, but not on ε.

Step 2 (Limit passage $\varepsilon \to 0$): It follows by definition of the Kantorovich–Rubinstein norm that

$$\begin{aligned}\sup_\varepsilon \iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon\big(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d\big) &\le \sup_\varepsilon\Big(\iota_\varepsilon\mathbf{m}^\varepsilon\big(\mathcal{I}\times\mathbb{T}^d\big) + \|\iota_\varepsilon(\tilde{\mathbf{m}}^\varepsilon - \mathbf{m}^\varepsilon)\|_{\mathrm{KR}(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d)}\Big)\\ &\le M + \eta.\end{aligned}$$

It follows from the growth condition (3.14) and (3.69) that

$$\begin{aligned}\sup_\varepsilon \big|\iota_\varepsilon\tilde{\mathbf{J}}^\varepsilon\big|\big(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d\big) &\lesssim \sup_\varepsilon \int_{\mathcal{I}^\eta} \varepsilon\|\tilde{J}_t^\varepsilon\|_{\ell^1(\mathcal{E}_\varepsilon)}\,\mathrm{d}t\\ &\lesssim \sup_\varepsilon \int_{\mathcal{I}^\eta}\Big(1 + \|\tilde{m}_t^\varepsilon\|_{\ell^1(\mathcal{X}_\varepsilon)} + \mathcal{F}_\varepsilon(\tilde{m}_t^\varepsilon, \tilde{J}_t^\varepsilon)\Big)\,\mathrm{d}t\\ &\le \sup_\varepsilon\Big(T + \iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon\big(\mathcal{I}^\eta\times\mathbb{T}^d\big) + \mathcal{A}_\varepsilon^{\mathcal{I}^\eta}(\tilde{\mathbf{m}}^\varepsilon, \tilde{\mathbf{J}}^\varepsilon)\Big)\\ &\le T + (M+\eta) + (A + 2\eta).\end{aligned} \tag{3.71}$$

Therefore, there exist measures $\boldsymbol{\mu}_\eta \in \mathcal{M}_+\big(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d\big)$ and $\boldsymbol{\nu}_\eta \in \mathcal{M}^d\big(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d\big)$ and convergent subsequences satisfying

$$\iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon \to \boldsymbol{\mu}_\eta \text{ and } \iota_\varepsilon\tilde{\mathbf{J}}^\varepsilon \to \boldsymbol{\nu}_\eta \text{ weakly in } \mathcal{M}_+(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d) \text{ and } \mathcal{M}^d(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d) \text{ as } \varepsilon\to 0. \tag{3.72}$$

The vague lower semicontinuity of the limiting functional (see Lemma 3.3.14), combined with (3.68), (3.69), and (3.70) thus yields

$$\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}^\eta}(\boldsymbol{\mu}_\eta, \boldsymbol{\nu}_\eta) \le \liminf_{\varepsilon\to 0} \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}^\eta}(\iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon, \iota_\varepsilon\tilde{\mathbf{J}}^\varepsilon) \le \liminf_{\varepsilon\to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon) + 2\eta. \tag{3.73}$$

Step 3 (Limit passage $\eta \to 0$): Let $\phi \in \mathrm{Lip}_1\big(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d\big)$, $\|\phi\|_\infty \le 1$. For brevity, write $\langle\phi, \boldsymbol{\mu}\rangle = \int_{\mathcal{I}^\eta\times\mathbb{T}^d}\phi\,\mathrm{d}\boldsymbol{\mu}$. Since from (3.72) $\iota_\varepsilon\mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ and $\iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon \to \boldsymbol{\mu}_\eta$ weakly, and $\|\iota_\varepsilon(\tilde{\mathbf{m}}^\varepsilon - \mathbf{m}^\varepsilon)\|_{\mathrm{KR}(\overline{\mathcal{I}^\eta}\times\mathbb{T}^d)} \le \eta$ we obtain

$$\begin{aligned}\langle\phi, \boldsymbol{\mu}_\eta - \boldsymbol{\mu}\rangle &\le \limsup_{\varepsilon\to 0}\Big(\big|\big\langle\phi, \boldsymbol{\mu}_\eta - \iota_\varepsilon\tilde{\mathbf{m}}^\varepsilon\big\rangle\big| + \big|\big\langle\phi, \iota_\varepsilon(\tilde{\mathbf{m}}^\varepsilon - \mathbf{m}^\varepsilon)\big\rangle\big| + \big|\big\langle\phi, \iota_\varepsilon\mathbf{m}^\varepsilon - \boldsymbol{\mu}\big\rangle\big|\Big)\\ &\le 0 + \eta + 0.\end{aligned}$$

It follows that $\|\boldsymbol{\mu}_\eta - \boldsymbol{\mu}\|_{\mathrm{KR}(\overline{\mathcal{I}^\eta} \times \mathbb{T}^d)} \leq 2\eta$, which together with $|\mathcal{I} \setminus \mathcal{I}^\eta| \leq \eta$ implies $\boldsymbol{\mu}_\eta \to \boldsymbol{\mu} \in \mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ vaguely as $\eta \to 0$.

Furthermore, (3.71) implies that $\sup_\eta \left|\boldsymbol{\nu}^\eta\right|\left(\mathcal{I}^\eta \times \mathbb{T}^d\right) < \infty$. Therefore, we may extract a subsequence so that $\boldsymbol{\nu}_\eta \to \boldsymbol{\nu}$ vaguely in $\mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ as $\eta \to 0$. It thus follows from (3.73) and the joint vague-lower semicontinuity of $\mathbb{A}_{\mathrm{hom}}$ (see Lemma 3.3.14) that

$$\mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \leq \liminf_{\varepsilon \to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon).$$

To conclude the desired estimate $\mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu}) \leq \liminf_{\varepsilon \to 0} \mathcal{A}_\varepsilon(\mathbf{m}^\varepsilon)$, it remains to show that $(\boldsymbol{\mu}, \boldsymbol{\nu})$ solves the continuity equation. To show this, we first note that $(\iota_\varepsilon \tilde{\mathbf{m}}^\varepsilon, \iota_\varepsilon \tilde{\mathbf{J}}^\varepsilon) \in \mathbb{CE}^{\mathcal{I}^\eta}$ in view of Lemma 3.4.9. It then follows from the weak convergence in (3.72) that $(\boldsymbol{\mu}_\eta, \boldsymbol{\nu}_\eta) \in \mathbb{CE}^{\mathcal{I}^\eta}$. Since $\boldsymbol{\mu}_\eta \to \boldsymbol{\mu}$, $\boldsymbol{\nu}_\eta \to \boldsymbol{\nu}$ vaguely, and $|\mathcal{I} - \mathcal{I}^\eta| \leq \eta$ it holds $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$, which completes the proof. □

3.7.1 Proof of the discrete regularisation result

This section is devoted to the proof of main discrete regularisation result, Proposition 3.7.1.

The regularised approximations are constructed by a three-fold regularisation: in time, space, and energy. Let us now describe the relevant operators.

3.7.2 Energy regularisation

First we embed m° and J° into the graph $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$. We thus define $m_\varepsilon^\circ \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ and $J_\varepsilon^\circ \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$ by

$$m_\varepsilon^\circ(\varepsilon z, v) := \varepsilon^d m^\circ(0, v) \qquad J_\varepsilon^\circ(\varepsilon z, v) := \varepsilon^{d-1} J^\circ(0, v).$$

It follows that $(m_\varepsilon^\circ, J_\varepsilon^\circ) \in \mathsf{D}(\mathcal{F}_\varepsilon)^\circ$ (by continuity of τ_ε^z, $z \in \mathbb{Z}_\varepsilon^d$) and

$$\mathcal{F}_\varepsilon(m_\varepsilon^\circ, J_\varepsilon^\circ) = F(m^\circ, J^\circ).$$

We then consider the energy regularisation operators defined by

$$\begin{aligned} R_\delta &: \mathbb{R}_+^{\mathcal{X}_\varepsilon} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}, & R_\delta m &:= (1-\delta) m + \delta m_\varepsilon^0, \\ R_\delta &: \mathbb{R}_a^{\mathcal{E}_\varepsilon} \to \mathbb{R}_a^{\mathcal{E}_\varepsilon}, & R_\delta J &:= (1-\delta) J + \delta J_\varepsilon^0. \end{aligned}$$

Lemma 3.7.5 (Energy regularisation)**.** *Let $\delta \in (0,1)$. The following inequalities hold for any $\varepsilon < \frac{1}{2R_0}$, $m \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$, and $J \in \mathbb{R}_a^{\mathcal{E}_\varepsilon}$:*

$$\begin{aligned} \mathcal{F}_\varepsilon(R_\delta m, R_\delta J) &\le (1-\delta)\mathcal{F}_\varepsilon(m, J) + \delta \mathcal{F}_\varepsilon(m_\varepsilon^\circ, J_\varepsilon^\circ), \\ \|R_\delta m\|_{\ell^\infty(\mathcal{X}_\varepsilon)} &\le (1-\delta)\|m\|_{\ell^\infty(\mathcal{X}_\varepsilon)} + \delta \varepsilon^d \|m^\circ\|_{\ell^\infty(\mathcal{X})}, \\ \|R_\delta J\|_{\ell^\infty(\mathcal{E}_\varepsilon)} &\le (1-\delta)\|J\|_{\ell^\infty(\mathcal{E}_\varepsilon)} + \delta \varepsilon^{d-1} \|J^\circ\|_{\ell^\infty(\mathcal{E})}. \end{aligned}$$

Proof. The proof is straightforward consequence of the convexity of F and the periodicity of m° and J°. □

3.7.3 Space regularisation

Our space regularisation is a convolution in the z-variable with the discretised heat kernel. It is of crucial importance that the regularisation is performed in the z-variable only. Smoothness in the v-variable is not expected.

For $\lambda > 0$ and $x \in \mathbb{T}^d$, let $h_\lambda(x)$ be the heat kernel on $\mathbb{T}^d$. We consider the discrete version

$$H_\lambda^\varepsilon : \mathbb{Z}_\varepsilon^d \to \mathbb{R}, \qquad H_\lambda^\varepsilon\big([z]\big) := \int_{Q_\varepsilon^z} h_\lambda(x) \, \mathrm{d}x,$$

where the integration ranges over the cube $Q_\varepsilon^z := \varepsilon z + [0, \varepsilon)^d \subseteq \mathbb{T}^d$. Using the boundedness and Lipschitz properties of h_δ, we infer that for $z \in \mathbb{Z}_\varepsilon^d$,

$$\inf_{\mathbb{Z}_\varepsilon^d} H_\lambda^\varepsilon \ge c_\lambda \varepsilon^d, \qquad \|H_\lambda^\varepsilon\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d)} \le C_\lambda \varepsilon^d, \tag{3.74}$$

$$\|H_\lambda^\varepsilon\|_{\ell^1(\mathbb{Z}_\varepsilon^d)} = 1, \qquad \Big\|H_\lambda^\varepsilon(\cdot + \varepsilon z) - H_\lambda^\varepsilon\Big\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d)} \le C_\lambda \varepsilon^{d+1}|z| \tag{3.75}$$

for some non-negative constant $C_\lambda < \infty$ depending only on $\lambda > 0$. We then define

$$S_\lambda : \mathbb{R}_+^{\mathcal{X}_\varepsilon} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}, \qquad S_\lambda m := \sum_{z \in \mathbb{Z}_\varepsilon^d} H_\lambda^\varepsilon(z) \sigma_\varepsilon^z m,$$
$$S_\lambda : \mathbb{R}_a^{\mathcal{E}_\varepsilon} \to \mathbb{R}_a^{\mathcal{E}_\varepsilon}, \qquad S_\lambda J := \sum_{z \in \mathbb{Z}_\varepsilon^d} H_\lambda^\varepsilon(z) \sigma_\varepsilon^z J,$$

where σ_ε^z is defined in (3.18).

Lemma 3.7.6 (Regularisation in space)**.** *Let $\lambda > 0$. There exist constants $c_\lambda > 0$ and $C_\lambda < \infty$ such that the following estimates hold, for any $\varepsilon < \frac{1}{2R_0}$, $m \in \mathbb{R}_+^{\mathcal{X}_\varepsilon}$, $J \in \mathcal{M}^d(\mathcal{E}_\varepsilon)$, and $z \in \mathbb{Z}_\varepsilon^d$:*

(i) Energy bound: $\mathcal{F}_\varepsilon(S_\lambda m, S_\lambda J) \le \mathcal{F}_\varepsilon(m, J)$.

(ii) Gain of integrability:

$$\|S_\lambda m\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_\lambda \varepsilon^d \|m\|_{\ell^1(\mathcal{X}_\varepsilon)} \quad \text{and} \quad \|S_\lambda J\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le C_\lambda \varepsilon^d \|J\|_{\ell^1(\mathcal{E}_\varepsilon)}.$$

(iii) Density lower bound: $\inf_{x \in \mathcal{X}_\varepsilon} S_\lambda m(x) \ge c_\lambda \varepsilon^d \|m\|_{\ell^1(\mathcal{X})}$.

(iv) Spatial regularisation:

$$\left\|\tau_\varepsilon^z S_\lambda m - S_\lambda m\right\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_\lambda \varepsilon^{d+1} |z| \|m\|_{\ell^1(\mathcal{X}_\varepsilon)} \quad \text{and}$$
$$\left\|\tau_\varepsilon^z S_\lambda J - S_\lambda J\right\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le C_\lambda \varepsilon^{d+1} |z| \|J\|_{\ell^1(\mathcal{E}_\varepsilon)}.$$

Proof. Using the convexity of F and the identity $\sum_z H_\lambda^\varepsilon(z) = 1$ we obtain

$$\begin{aligned}
\mathcal{F}_\varepsilon(S_\lambda m, S_\lambda J) &= \sum_{z \in \mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{\tau_\varepsilon^z S_\lambda m}{\varepsilon^d}, \frac{\tau_\varepsilon^z S_\lambda J}{\varepsilon^{d-1}}\right) \\
&\le \sum_{z \in \mathbb{Z}_\varepsilon^d} \sum_{z' \in \mathbb{Z}_\varepsilon^d} \varepsilon^d H_\lambda^\varepsilon(z') F\left(\frac{\tau_\varepsilon^{z+z'} m}{\varepsilon^d}, \frac{\tau_\varepsilon^{z+z'} J}{\varepsilon^{d-1}}\right) \\
&= \sum_{z \in \mathbb{Z}_\varepsilon^d} \left(\sum_{z' \in \mathbb{Z}_\varepsilon^d} H_\lambda^\varepsilon(z - z')\right) \varepsilon^d F\left(\frac{\tau_\varepsilon^z m}{\varepsilon^d}, \frac{\tau_\varepsilon^z J}{\varepsilon^{d-1}}\right) = F(M, J),
\end{aligned}$$

where in the last equality we used (3.75). This shows (i). Properties (ii), (iii), and (iv) are straightforward consequence of the uniform bounds (3.74), (3.75) for the discrete kernels H_λ^ε. □

3.7.4 Time regularisation

Fix an interval $\mathcal{I} = (a, b) \subset \mathbb{R}$ and a regularisation parameter $\tau > 0$. For $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$, we define for $t \in \mathcal{I}_\tau := (a + \tau, b - \tau)$

$$(T_\tau \mathbf{m})_t := \fint_{t-\tau}^{t+\tau} m_s \, \mathrm{d}s, \qquad (T_\tau \mathbf{J})_t := \fint_{t-\tau}^{t+\tau} J_s \, \mathrm{d}s\,.$$

Note that, thanks to the linearity of the continuity equation we get $(T_\tau \mathbf{m}, T_\tau \mathbf{J}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}_\tau}$.

We have the following regularisation properties for the operator T_τ.

Lemma 3.7.7 (Regularisation in time). *Let $\tau \in (0, \frac{b-a}{2})$. The following estimates hold for all $\varepsilon < \frac{1}{2R_0}$ and all Borel curves $\mathbf{m} = (m_t)_{t\in\mathcal{I}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ and $\mathbf{J} = (J_t)_{t\in\mathcal{I}} \subseteq \mathcal{M}^d(\mathcal{E}_\varepsilon)$:*

(i) Energy estimate: for some $0 \le C < \infty$ depending only on (3.14) we have

$$\mathcal{A}_\varepsilon^{\mathcal{I}_\tau}(T_\tau \mathbf{m}, T_\tau \mathbf{J}) \le \mathcal{A}_\varepsilon(\mathbf{m}, \mathbf{J}) + C\tau\Big(\mathbf{m}(\mathcal{I}\times\mathcal{X}_\varepsilon) + 1\Big)\,.$$

(ii) Mass estimate: $\sup_{t\in\mathcal{I}_\tau} \|(T_\tau m)_t\|_{\ell^p(\mathcal{X}_\varepsilon)} \le \sup_{t\in\mathcal{I}} \|m_t\|_{\ell^p(\mathcal{X}_\varepsilon)}$.

(iii) Momentum estimate: $\sup_{t\in\mathcal{I}_\tau} \|(T_\tau J)_t\|_{\ell^p(\mathcal{X}_\varepsilon)} \le \dfrac{1}{\tau}\displaystyle\int_{\mathcal{I}} \|J_t\|_{\ell^p(\mathcal{X}_\varepsilon)}\,\mathrm{d}t$.

(iv) Time regularity: $\sup_{t\in\mathcal{I}_\tau} \Big\|\partial_t (T_\tau m)_t\Big\|_{\ell^p(\mathcal{X}_\varepsilon)} \le \dfrac{1}{\tau}\sup_{t\in\mathcal{I}} \|m_t\|_{\ell^p(\mathcal{X}_\varepsilon)}$.

Proof. Set $w_\tau(s) := (2\tau)^{-1}\Big|[(s-\tau)\vee a, (s+\tau)\wedge b]\Big|$ for $s\in\mathcal{I}$. Then we have

$$\mathcal{A}_\varepsilon^{\mathcal{I}_\tau}(T_\tau \mathbf{m}, T_\tau \mathbf{J}) \le \int_{\mathcal{I}_\tau} \fint_{t-\tau}^{t+\tau} \mathcal{F}_\varepsilon(m_s, J_s)\,\mathrm{d}s\,\mathrm{d}t = \int_{\mathcal{I}} w(s)\mathcal{F}_\varepsilon(m_s,J_s)\,\mathrm{d}s\,, \tag{3.76}$$

as a consequence of Jensen's inequality and Fubini's theorem. Using that $0 \le w_\tau \le 1$, $\int_{\mathcal{I}}(1-w_\tau(s))\,\mathrm{d}s = 2\tau$, and the growth condition (3.14) we infer

$$\int_{\mathcal{I}} (1-w_\tau(s))\mathcal{F}_\varepsilon(m_s,J_s)\,\mathrm{d}s \ge -C\tau\Big(\mathbf{m}(\mathcal{I}\times\mathcal{X}_\varepsilon)+1\Big)\,,$$

which together with (3.76) shows (i).

Properties (ii), (iii) follow directly from the convexity of the ℓ_p-norms and the subadditivity of the integral.

Finally, (iv) follows from the direct computation $\partial_t(T_\tau m)_t = \frac{1}{2\tau}(m_{t+\tau} - m_{t-\tau})$. □

3.7.5 Proof of the regularisation result

We start with a lemma that shows that the effect of the three regularising operators is small if the parameters are small.

Recall the definition of the Kantorovich-Rubinstein norm as given in Appendix B.

Lemma 3.7.8 (Bounds in KR-norm). *Let $\mathcal{I}\subset\mathbb{R}$ and interval and $(m_t)_{t\in\mathcal{I}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ be a Borel measurable curve of constant total mass (i.e., $t\mapsto m_t(\mathcal{X}_\varepsilon)$ is constant), and let $\mathbf{m}\in\mathcal{M}_+(\mathcal{I}\times\mathcal{X}_\varepsilon)$ be the associated measure on space-time defined by $\mathbf{m} := \mathrm{d}t\otimes m_t$. Then there exists a constant $C<\infty$ depending on $|\mathcal{I}|$ such that:*

(i) $\|\iota_\varepsilon T_\tau \mathbf{m} - \iota_\varepsilon \mathbf{m}\|_{\mathrm{KR}(\overline{\mathcal{I}_\tau}\times\mathbb{T}^d)} \le C\tau \sup_{t\in\mathcal{I}} \Big\|m_t\Big\|_{\ell^1(\mathcal{X}_\varepsilon)}$ *for any $\tau < |\mathcal{I}|/2$.*

(ii) $\|\iota_\varepsilon S_\lambda \mathbf{m} - \iota_\varepsilon \mathbf{m}\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} \le C\sqrt{\lambda} \sup_{t\in\mathcal{I}} \Big\|m_t\Big\|_{\ell^1(\mathcal{X}_\varepsilon)}$ *for any $\lambda > 0$.*

(iii) $\|\iota_\varepsilon R_\delta \mathbf{m} - \iota_\varepsilon \mathbf{m}\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} \leq C\delta\Big(m^\circ(\mathcal{X}^Q) + \sup_{t\in\mathcal{I}} \big\|m_t\big\|_{\ell^1(\mathcal{X}_\varepsilon)}\Big)$ *for any* $\delta \in (0,1)$.

Proof. (i): For any $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{I}\times\mathbb{T}^d)$ and any Lipschitz function $\phi : \overline{\mathcal{I}_\tau}\times\mathbb{T}^d \to \mathbb{R}$ (and, in fact, for any temporally Lipschitz function) we have

$$\begin{aligned}&\bigg|\int_{\overline{\mathcal{I}_\tau}\times\mathbb{T}^d} \phi(t,x)\,\mathrm{d}\boldsymbol{\mu}(t,x) - \int_{\overline{\mathcal{I}_\tau}\times\mathbb{T}^d} \phi(t,x)\,\mathrm{d}(T_\tau\boldsymbol{\mu})(t,x)\bigg|\\ &= \bigg|\int_{\overline{\mathcal{I}_\tau}\times\mathbb{T}^d} \fint_{t-\tau}^{t+\tau} \phi(s,x) - \phi(t,x)\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\mu}(t,x)\bigg| \leq \tau[\phi]_{\mathrm{Lip}}\boldsymbol{\mu}\Big(\mathcal{I}\times\mathbb{T}^d\Big).\end{aligned}$$

Since $\iota_\varepsilon\mathbf{m}\Big(\mathcal{I}\times\mathbb{T}^d\Big) \leq |\mathcal{I}|\sup_{t\in\mathcal{I}}\big\|m_t\big\|_{\ell^1(\mathcal{X}_\varepsilon)}$ we obtain the result.

(ii): In view of mass-preservation, we have

$$\begin{aligned}\|\iota_\varepsilon S_\lambda\mathbf{m} - \iota_\varepsilon\mathbf{m}\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} &\leq \int_{\mathcal{I}} \big\|\iota_\varepsilon S_\lambda m_t - \iota_\varepsilon m_t\big\|_{\mathrm{KR}(\mathbb{T}^d)}\,\mathrm{d}t\\ &\leq \sup_{t\in\mathcal{I}} m_t(\mathcal{X}_\varepsilon)\int_{\mathcal{I}} \big\|\iota_\varepsilon H_\lambda - \iota_\varepsilon H_0\big\|_{\mathrm{KR}(\mathbb{T}^d)}\,\mathrm{d}t\\ &\leq C\sqrt{\lambda}\sup_{t\in\mathcal{I}} m_t(\mathcal{X}_\varepsilon).\end{aligned}$$

Here in the last inequality we used scaling law of the heat kernel.

(iii): Let us write $\mathbf{m}_\varepsilon^\circ := \mathrm{d}t\otimes m_\varepsilon^\circ$ for brevity. By linearity, we have

$$\begin{aligned}\|\iota_\varepsilon(R_\delta\mathbf{m} - \mathbf{m})\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} &= \delta\|\iota_\varepsilon(\mathbf{m}_\varepsilon^\circ - \mathbf{m})\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)}\\ &\leq \delta(1+|\mathcal{I}|)\Big(\mathbf{m}_\varepsilon^\circ\Big(\mathcal{I}\times\mathbb{T}_\varepsilon^d\Big) + \mathbf{m}\Big(\mathcal{I}\times\mathbb{T}_\varepsilon^d\Big)\Big)\\ &\leq \delta|\mathcal{I}|(1+|\mathcal{I}|)\Big(m^\circ(\mathcal{X}^Q) + \sup_{t\in\mathcal{I}} m_t(\mathcal{X}_\varepsilon)\Big).\end{aligned}$$

□

Proof of Proposition 3.7.1. We define

$$\tilde{\mathbf{m}} := \Big(R_\delta\circ S_\lambda\circ T_\tau\Big)\mathbf{m} \quad\text{and}\quad \tilde{\mathbf{J}} := \Big(R_\delta\circ S_\lambda\circ T_\tau\Big)\mathbf{J}.$$

We will show that the desired inequalities hold if $\delta,\lambda,\tau > 0$ are chosen to be sufficiently small, depending on the desired accuracy $\eta > 0$. Set $\mathcal{I}_\tau := (\tau, T-\tau)$.

(i): We use the shorthand notation $\mathrm{KR}_\tau := \mathrm{KR}(\overline{\mathcal{I}_\tau}\times\mathbb{T}^d)$. Using Lemma 3.7.8 we obtain

$$\begin{aligned}\|\iota_\varepsilon\mathbf{m} - \iota_\varepsilon\tilde{\mathbf{m}}\|_{\mathrm{KR}_\tau} &\leq \|\iota_\varepsilon\mathbf{m} - \iota_\varepsilon T_\tau\mathbf{m}\|_{\mathrm{KR}_\tau} + \|\iota_\varepsilon T_\tau\mathbf{m} - \iota_\varepsilon(S_\lambda T_\tau)\mathbf{m}\|_{\mathrm{KR}_\tau}\\ &\quad + \|\iota_\varepsilon(S_\lambda T_\tau)\mathbf{m} - \iota_\varepsilon(R_\delta S_\lambda T_\tau)\mathbf{m}\|_{\mathrm{KR}_\tau}\\ &\lesssim M(\tau + \sqrt{\lambda} + \delta) + m^\circ(\mathcal{X}^Q)\delta.\end{aligned} \tag{3.77}$$

Furthermore, using Lemma 3.7.5, Lemma 3.7.6(i), and Lemma 3.7.7(i) we obtain the energy bound

$$\begin{aligned}\mathcal{A}_\varepsilon^{\mathcal{I}_\tau}(\tilde{\mathbf{m}},\tilde{\mathbf{J}}) &= \mathcal{E}_\varepsilon\Big((R_\delta\circ S_\lambda\circ T_\tau)\mathbf{m}, (R_\delta\circ S_\lambda\circ T_\tau)\mathbf{J}\Big)\\ &\leq (1-\delta)\mathcal{A}_\varepsilon\Big((S_\lambda\circ T_\tau)\mathbf{m}, (S_\lambda\circ T_\tau)\mathbf{J}\Big) + \delta T\mathcal{F}_\varepsilon(m_\varepsilon^\circ, J_\varepsilon^\circ)\\ &\leq (1-\delta)\mathcal{A}_\varepsilon(\mathbf{m},\mathbf{J}) + \delta T\mathcal{F}(m^\circ, J^\circ) + C\tau(M+1)\,.\end{aligned} \tag{3.78}$$

The desired inequalities (3.63) follow by choosing δ, λ, and τ sufficiently small.

(ii): We will show that all the estimates hold with constants depending on η through the parameters δ, λ, and τ.

Boundedness:

$$\begin{aligned}\sup_{t\in\mathcal{I}_\tau}\|\tilde{m}_t\|_{\ell^\infty(\mathcal{X}_\varepsilon)} &\le \varepsilon^d\bigg((1-\delta)C_\lambda \sup_{t\in[0,T]}\|m_t\|_{\ell^1(\mathcal{X}_\varepsilon)} + \delta\|m^\circ\|_{\ell^\infty(\mathcal{X}_\varepsilon)}\bigg),\\ &\le \varepsilon^d\bigg(C_\lambda M + \delta\|m^\circ\|_{\ell^\infty(\mathcal{X}^Q)}\bigg).\end{aligned} \tag{3.79}$$

$$\begin{aligned}\sup_{t\in\mathcal{I}_\tau}\|\tilde{J}_t\|_{\ell^\infty(\mathcal{X}_\varepsilon)} &\le \varepsilon^{d-1}\bigg(\frac{1-\delta}{\tau}C_\lambda \sup_{t\in[0,T]}\int_{\mathcal{I}}\varepsilon\|J_t\|_{\ell^1(\mathcal{X}_\varepsilon)}\,\mathrm{d}t + \delta\|J^\circ\|_{\ell^\infty(\mathcal{X}_\varepsilon)}\bigg),\\ &\lesssim \varepsilon^{d-1}\bigg(\frac{C_\lambda}{\tau}\Big(T(1+M)+E\Big) + \delta\|J^\circ\|_{\ell^\infty(\mathcal{E}^Q)}\bigg).\end{aligned} \tag{3.80}$$

Time-regularity:

$$\begin{aligned}\sup_{t\in\mathcal{I}_\tau}\|\partial_t\tilde{m}_t\|_{\ell^\infty(\mathcal{X}_\varepsilon)} &\le \varepsilon^d\bigg(2\frac{1-\delta}{\tau}C_\lambda \sup_{t\in[0,T]}\|m_t\|_{\ell^1(\mathcal{X}_\varepsilon)} + \delta\|m^\circ\|_{\ell^\infty(\mathcal{X}_\varepsilon)}\bigg),\\ &\le \varepsilon^d\bigg(2\frac{C_\lambda}{\tau}M + \delta\|m^\circ\|_{\ell^\infty(\mathcal{X}^Q)}\bigg).\end{aligned} \tag{3.81}$$

Space-regularity: For $z, z' \in \mathbb{Z}^d_\varepsilon$ and $v \in \mathbb{V}$ we have

$$\begin{aligned}|\tilde{m}_t(z,v) - \tilde{m}_t(z',v)| &\le (1-\delta)\Big|\Big(S_\lambda\circ T_\tau\Big)m_t(z,v) - \Big(S_\lambda\circ T_\tau\Big)m_t(z',v)\Big|\\ &\le C_\lambda\varepsilon^{d-1}|z-z'|\Big\|T_\tau m_t\Big\|_{\ell^1(\mathcal{X}_\varepsilon)}\\ &\le C_\lambda\varepsilon^{d+1}|z-z'|\sup_{t\in[0,T]}\Big\|m_t\Big\|_{\ell^1(\mathcal{X}_\varepsilon)},\end{aligned}$$

which shows that

$$\sup_{t\in\mathcal{I}_\tau}\|\sigma^z_\varepsilon\tilde{m}_t - \tilde{m}_t\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_\lambda\varepsilon^{d+1}|z|\sup_{t\in[0,T]}\Big\|m_t\Big\|_{\ell^1(\mathcal{X}_\varepsilon)} \le C_\lambda\varepsilon^{d+1}|z|M. \tag{3.82}$$

Similarly,

$$\begin{aligned}\sup_{t\in\mathcal{I}_\tau}\|\sigma^z_\varepsilon\tilde{J}_t - \tilde{J}_t\|_{\ell^\infty(\mathcal{E}_\varepsilon)} &\le \frac{C_\lambda}{\tau}\varepsilon^{d+1}|z|\int_{\mathcal{I}}\Big\|J_s\Big\|_{\ell^1(\mathcal{E}_\varepsilon)}\,\mathrm{d}s\\ &\le \frac{C_\lambda}{\tau}\varepsilon^d|z|\Big(T(1+M)+E\Big).\end{aligned} \tag{3.83}$$

Domain-regularity: For all $t \in \mathcal{I}_\tau$ we observe that

$$\varepsilon^{-d}\|(S_\lambda T_\tau m)_t\|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C_\lambda\|(T_\tau m)_t\|_{\ell^1(\mathcal{X}_\varepsilon)} \le C_\lambda\sup_{t\in[0,T]}\|m_t\|_{\ell^1(\mathcal{X}_\varepsilon)} \le C_\lambda M,$$

$$\varepsilon^{-d}\|(S_\lambda T_\tau J)_t\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le C_\lambda\|(T_\tau m)_t\|_{\ell^1(\mathcal{E}_\varepsilon)} \le \frac{C_\lambda}{\tau}\int_{\mathcal{I}}\|J_t\|_{\ell^1(\mathcal{E}_\varepsilon)}\,\mathrm{d}t \le \frac{C_\lambda}{\tau\varepsilon}\Big(T(1+M)+E\Big).$$

We infer that

$$\left\| \frac{\tau_\varepsilon^z (S_\lambda T_\tau m)_t}{\varepsilon^d} \right\|_{\ell^\infty(\mathcal{X})} \leq C_\lambda M \quad \text{and} \quad \left\| \frac{\tau_\varepsilon^z (S_\lambda T_\tau J)_t}{\varepsilon^{d-1}} \right\|_{\ell^\infty(\mathcal{E})} \leq \frac{C_\lambda}{\tau} \Big(T(1+M) + E \Big)$$

Since

$$\left(\frac{\tau_\varepsilon^z \tilde{m}_t}{\varepsilon^d}, \frac{\tau_\varepsilon^z \tilde{J}_t}{\varepsilon^{d-1}} \right) = (1-\delta) \left(\frac{\tau_\varepsilon^z (S_\lambda T_\tau m)_t}{\varepsilon^d}, \frac{\tau_\varepsilon^z (S_\lambda T_\tau J)_t}{\varepsilon^{d-1}} \right) + \delta(m^\circ, J^\circ),$$

the claim follows by an application of Lemma B.3.1 to the product of balls in $\ell^\infty(\mathcal{X})$ and $\ell^\infty(\mathcal{E})$, taking into account that F is defined on a finite-dimensional subspace by the locality assumption.

The result follows from the inequalities (3.77)–(3.83), by choosing δ, λ, and τ sufficiently small. □

3.8 Proof of the Γ-limsup inequality

In this section we present the proof of the Γ-limsup inequality in Theorem 3.5.1. The first step is to introduce the notion of *optimal microstructures*.

3.8.1 The optimal discrete microstructures

Let $\mathcal{I}$ be an open interval in $\mathbb{R}$. We will make use of the following canonical discretisation of measures and vector fields on the cartesian grid $\mathbb{Z}^d_\varepsilon$.

Definition 3.8.1 ($\mathbb{Z}^d_\varepsilon$-discretisation of measures)**.** Let $\mu \in \mathcal{M}_+(\mathbb{T}^d)$ and $\nu \in \mathcal{M}^d(\mathbb{T}^d)$ have continuous densities ρ and j, respectively, with respect to the Lebesgue measure. Their $\mathbb{Z}^d_\varepsilon$-discretisations $P_\varepsilon\mu : \mathbb{Z}^d_\varepsilon \to \mathbb{R}_+$ and $P_\varepsilon\nu : \mathbb{Z}^d_\varepsilon \to \mathbb{R}^d$ are defined by

$$P_\varepsilon\mu(z) := \mu(Q^z_\varepsilon)\,, \quad P_\varepsilon\nu(z) := \left(\int_{\partial Q^z_\varepsilon \cap \partial Q^{z+e_i}_\varepsilon} j \cdot e_i \,\mathrm{d}\mathcal{H}^{d-1}\right)_{i=1}^d .$$

An important feature of this discretisation is the preservation of the continuity equation, in the following sense.

Definition 3.8.2 (Continuity equation on $\mathbb{Z}^d_\varepsilon$)**.** Fix $\mathcal{I} \subset \mathbb{R}$ an open interval. We say that $\mathbf{r} : \mathcal{I} \times \mathbb{Z}^d_\varepsilon \to \mathbb{R}_+$ and $\mathbf{u} : \mathcal{I} \times \mathbb{Z}^d_\varepsilon \to \mathbb{R}^d$ satisfy the continuity equation on $\mathbb{Z}^d_\varepsilon$, and write $(\mathbf{r}, \mathbf{u}) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$, if $\mathbf{r}$ is continuous, $\mathbf{u}$ is Borel measurable, and the following discrete continuity equation is satisfied in the sense of distributions:

$$\partial_t r_t(z) + \sum_{i=1}^d \Big(u_t(z) - u_t(z - e_i)\Big) \cdot e_i = 0, \quad \text{for } z \in \mathbb{Z}^d_\varepsilon . \tag{3.84}$$

Lemma 3.8.3 (Discrete continuity equation on $\mathbb{Z}^d_\varepsilon$)**.** *Let $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ have continuous densities with respect to the space-time Lebesgue measure on $\mathcal{I} \times \mathbb{T}^d$. Then $(\mathsf{P}_\varepsilon\boldsymbol{\mu}, \mathsf{P}_\varepsilon\boldsymbol{\nu}) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$.*

Proof. This follows readily from the Gauß divergence theorem. □

The key idea of the proof of the Γ-limsup inequality in Theorem 3.5.1 is to start from a (smooth) solution to the continuous equation $\mathbb{CE}^{\mathcal{I}}$, and to consider the optimal discrete microstructure of the mass and the flux in each cube Q^z_ε. The global candidate is then obtained by gluing together the optimal microstructures *cube by cube*.

We start defining the *gluing operator*. Recall the operator T^0_ε defined in (3.17).

Definition 3.8.4 (Gluing operator)**.** Fix $\varepsilon > 0$. For each $z \in \mathbb{Z}^d_\varepsilon$, let

$$m^z \in \mathbb{R}^{\mathcal{X}}_+ \quad \text{and} \quad J^z \in \mathbb{R}^{\mathcal{E}}_a$$

be $\mathbb{Z}^d$-periodic. The *gluings* of $m = (m^z)_{z \in \mathbb{Z}^d_\varepsilon}$ and $J = (J^z)_{z \in \mathbb{Z}^d_\varepsilon}$ are the functions $\mathcal{G}_\varepsilon m \in \mathbb{R}^{\mathcal{X}_\varepsilon}_+$ and $\mathcal{G}_\varepsilon J \in \mathbb{R}^{\mathcal{E}_\varepsilon}_a$ defined by

$$\begin{aligned} \mathcal{G}_\varepsilon m\Big(T^0_\varepsilon(x)\Big) &:= m^{x_z}(x) && \text{for } x \in \mathcal{X}, \\ \mathcal{G}_\varepsilon J\Big(T^0_\varepsilon(x), T^0_\varepsilon(y)\Big) &:= \frac{1}{2}\Big(J^{x_z}(x,y) + J^{y_z}(x,y)\Big) && \text{for } (x,y) \in \mathcal{E}. \end{aligned} \tag{3.85}$$

Remark 3.8.5 (Well-posedness). Note that $\mathcal{G}_\varepsilon m$ and $\mathcal{G}_\varepsilon J$ are well-defined thanks to the $\mathbb{Z}^d_\varepsilon$-periodicity of the functions m^z and J^z.

Remark 3.8.6. (Mass preservation and KR-bounds) The gluing operation is locally mass-preserving in the following sense. Let $\mu \in \mathcal{M}_+(\mathbb{T}^d)$ and consider a family of measures $m = (m^z)_{z\in\mathbb{Z}^d_\varepsilon} \subseteq \mathbb{R}^{\mathcal{X}}_+$ satisfying $m^z \in \mathsf{Rep}\Big(P_\varepsilon\mu(z)\Big)$ for some $z \in \mathbb{Z}^d_\varepsilon$. Then:

$$\mathcal{G}_\varepsilon m\Big(\mathcal{X}_\varepsilon \cap \{x_{\mathsf{z}} = z\}\Big) = \mu(Q^z_\varepsilon)$$

for every $\varepsilon > 0$. Consequently,

$$\|\iota_\varepsilon\mathcal{G}_\varepsilon\mathbf{m} - \boldsymbol{\mu}\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} \le \boldsymbol{\mu}\Big(\overline{\mathcal{I}} \times \mathbb{T}^d\Big)\sqrt{d}\varepsilon \tag{3.86}$$

for all weakly continuous curves $\boldsymbol{\mu} = (\mu_t)_{t\in\overline{\mathcal{I}}} \subseteq \mathcal{M}_+(\mathbb{T}^d)$ and all $\mathbf{m} = (m^z_t)_{t\in\overline{\mathcal{I}},z\in\mathbb{Z}^d_\varepsilon}$ such that $m^z_t \in \mathsf{Rep}\Big(\mathsf{P}_\varepsilon\mu_t(z)\Big)$ for all $t \in \overline{\mathcal{I}}$ and $z \in \mathbb{Z}^d_\varepsilon$.

Energy estimates for Lipschitz microstructures

The next lemma shows that the energy of glued measures can be controlled under suitable regularity assumptions.

Lemma 3.8.7 (Energy estimates under regularity)**.** *Fix $\varepsilon > 0$. For each $z \in \mathbb{Z}^d_\varepsilon$, let $m^z \in \mathbb{R}^{\mathcal{X}}_+$ and $J^z \in \mathbb{R}^{\mathcal{E}}_a$ be $\mathbb{Z}^d$-periodic functions satisfying:*

(i) (Lipschitz dependence): For all $z, \tilde z \in \mathbb{Z}^d_\varepsilon$

$$\Big\|m^z - m^{\tilde z}\Big\|_{\ell^\infty(\mathcal{X})} + \varepsilon\Big\|J^z - J^{\tilde z}\Big\|_{\ell^\infty(\mathcal{E})} \le L|z - \tilde z|\varepsilon^{d+1}.$$

(ii) (Domain regularity): There exists a compact and convex set $K \Subset \mathsf{D}(F)^\circ$ such that, for all $z \in \mathbb{Z}^d_\varepsilon$,

$$\left(\frac{m^z}{\varepsilon^d}, \frac{J^z}{\varepsilon^{d-1}}\right) \in K. \tag{3.87}$$

Then there exists $\varepsilon_0 > 0$ depending only on K, F such that for $\varepsilon \le \varepsilon_0$

$$\mathcal{F}_\varepsilon\Big(\mathcal{G}_\varepsilon m, \mathcal{G}_\varepsilon J\Big) \le \sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{m^z}{\varepsilon^d}, \frac{J^z}{\varepsilon^{d-1}}\right) + c\varepsilon, \tag{3.88}$$

where $c < \infty$ depends only on L, the (finite) Lipschitz constant $\mathrm{Lip}(F; K)$, and the locality radius R_1.

Proof. Fix $\bar z \in \mathbb{Z}^d_\varepsilon$. As m is $\mathbb{Z}^d$-periodic, (i) yields for $x = (z, v) \in \mathcal{X}_{R_1}$,

$$|\tau^{\bar z}_\varepsilon \mathcal{G}_\varepsilon m(x) - m^{\bar z}(x)| = |m^{\bar z+z}(x) - m^{\bar z}(x)| \le LR_1\varepsilon^{d+1}, \tag{3.89}$$

Similarly, using the $\mathbb{Z}^d$-periodicity of J, (i) yields for $(x, y) \in \mathcal{E}$ with $x = (z, v) \in \mathcal{X}_{R_1}$ and $y = (\tilde z, \tilde v) \in \mathcal{X}_{R_1}$,

$$|\tau^{\bar z}_\varepsilon \mathcal{G}_\varepsilon J(x, y) - J^{\bar z}(x, y)| = \left|\Big(\tfrac{1}{2}J^{\bar z+z} + \tfrac{1}{2}J^{\bar z+\tilde z} - J^{\bar z}\Big)(x, y)\right| \le LR_1\varepsilon^d. \tag{3.90}$$

Hence the domain regularity assumption (ii) imply a domain regularity property for the glued measures, namely

$$\left(\frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon J}{\varepsilon^{d-1}} \right) \in \tilde K$$

for all $\bar z \in \mathbb{Z}_\varepsilon^d$ and $\varepsilon \le \varepsilon_0 := \frac{1}{2}\mathrm{dist}(K, \partial\,\mathsf{D}(F)) \in (0, +\infty)$, where $\tilde K \Subset \mathsf{D}(F)^\circ$ is a slightly bigger compact set than K.

Consequently, we can use the Lipschitzianity of F on the compact set $\tilde K$ and its locality to estimate the energy as

$$\begin{aligned} &\left| F\left(\frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon J}{\varepsilon^{d-1}} \right) - F\left(\frac{M^{\bar z}}{\varepsilon^d}, \frac{J^{\bar z}}{\varepsilon^{d-1}} \right) \right| \\ &\le \mathrm{Lip}(F; \tilde K) \left(\frac{\| \tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon m - m^{\bar z} \|_{\ell^\infty(\mathcal{X}_{R_1})}}{\varepsilon^d} + \frac{\| \tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon J - J^{\bar z} \|_{\ell^\infty(\mathcal{E}_{R_1})}}{\varepsilon^{d-1}} \right), \end{aligned}$$

where $\mathcal{X}_R := \{ x \in \mathcal{X} \ : \ |x|_{\ell_\infty^d} \le R \}$ and $\mathcal{E}_R := \{ (x,y) \in \mathcal{E} \ : \ |x|_{\ell_\infty^d}, |y|_{\ell_\infty^d} \le R \}$.

Combining the estimate above with (3.89) and (3.90), we conclude that

$$\left| F\left(\frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon m}{\varepsilon^d}, \frac{\tau_\varepsilon^{\bar z} \mathcal{G}_\varepsilon J}{\varepsilon^{d-1}} \right) - F\left(\frac{M^{\bar z}}{\varepsilon^d}, \frac{J^{\bar z}}{\varepsilon^{d-1}} \right) \right| \le 2LR_1 \, \mathrm{Lip}(F; \tilde K)\varepsilon \, .$$

for $\varepsilon \le \varepsilon_0$. Summation over $\bar z \in \mathbb{Z}_\varepsilon^d$ yields the desired estimate (3.88). □

We now introduce the notion of *optimal microstructure* associated with a pair of measures $(\mu, \nu) \in \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$. First, let us define regular measures.

Definition 3.8.8 (Regular measures)**.** We say that $(\mu, \nu) \in \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$ is a *regular pair of measures* if the following properties hold:

(i) (Lipschitz regularity): With respect to the Lebesgue measure on $\mathbb{T}^d$, the measures μ and ν have Lipschitz continuous densities ρ and j respectively.

(ii) (Compact inclusion): There exists a compact set $\tilde K \Subset \mathsf{D}(f_{\mathrm{hom}})^\circ$ such that

$$\big(\rho(x), j(x) \big) \in \tilde K \quad \text{for all } x \in \mathbb{T}^d.$$

We say that $(\mu_t, \nu_t)_{t \in \mathcal{I}} \subseteq \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$ is a *regular curve of measures* if (μ_t, ν_t) are regular measures for every $t \in \mathcal{I}$ and $t \mapsto (\rho_t(x), j_t(x))$ is measurable for every $x \in \mathbb{T}^d$.

Definition 3.8.9 (Optimal microstructure)**.** Let $(\mu, \nu) \in \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$ be a regular pair of measures.

(i) We say that $(m^z, J^z)_{z \in \mathbb{Z}_\varepsilon^d} \subseteq \mathbb{R}_+^{\mathcal{X}} \times \mathbb{R}_a^{\mathcal{E}}$ is an *admissible microstructure* for (μ, ν) if

$$(m^z, J^z) \in \mathsf{Rep} \left(\frac{P_\varepsilon \mu(z)}{\varepsilon^d}, \frac{P_\varepsilon \nu(z)}{\varepsilon^{d-1}} \right)$$

for every $z \in \mathbb{Z}_\varepsilon^d$.

(ii) If, additionally, $(m^z, J^z) \in \mathsf{Rep}_o\left(\frac{P_\varepsilon\mu(z)}{\varepsilon^d}, \frac{P_\varepsilon\nu(z)}{\varepsilon^{d-1}}\right)$ for every $z \in \mathbb{Z}^d_\varepsilon$, we say that $(m^z, J^z)_{z\in\mathbb{Z}^d_\varepsilon}$ is an *optimal microstructure* for (μ, ν).

Remark 3.8.10 (Measurable dependence). If $t \mapsto (\mu_t, \nu_t)$ is a measurable curve in $\mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$, it is possible to select a collection of admissible (resp. optimal) microstructures that depend measurably on t. This follows from Lemma 3.4.7; see e.g. [RW98, Theorem 14.37]. In the sequel, we will always work with measurable selections.

The next proposition shows that each optimal microstructures associated with a regular pair of measures (μ, ν) has discrete energy which can be controlled by the homogenised continuous energy $\mathbb{F}_{\text{hom}}(\mu, \nu)$.

Proposition 3.8.11 (Energy bound for optimal microstructures). *Let $(m^z, J^z)_{z\in\mathbb{Z}^d_\varepsilon} \subseteq \mathbb{R}^{\mathcal{X}}_+ \times \mathbb{R}^{\mathcal{E}}_a$ be an optimal microstructure for a regular pair of measures $(\mu, \nu) \in \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$. Then:*

$$\sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{m^z}{\varepsilon^d}, \frac{J^z}{\varepsilon^{d-1}}\right) \le \mathbb{F}_{\text{hom}}(\mu, \nu) + C\varepsilon,$$

where $C < \infty$ depends only on $\operatorname{Lip}(f_{\text{hom}}; \tilde{K})$ and the modulus of continuity of the densities ρ and j of μ and ν.

Proof. Let us denote the densities of μ and ν by ρ and j respectively. Using the regularity of $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, and the fact that f_{hom} is Lipschitz on $\tilde{K}$, we obtain

$$\sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{m^z}{\varepsilon^d}, \frac{J^z}{\varepsilon^{d-1}}\right) = \sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon^d f_{\text{hom}}\left(\frac{P_\varepsilon\mu(z)}{\varepsilon^d}, \frac{P_\varepsilon\nu(z)}{\varepsilon^{d-1}}\right) \le \int_{\mathbb{T}^d} f_{\text{hom}}(\rho_t(a), j_t(a)) \, \mathrm{d}a + C\varepsilon,$$

which is the desired estimate. □

Remark 3.8.12 (Lack of regularity). Suppose that $\widehat{m} := \mathcal{G}_\varepsilon m$ and $\widehat{J} := \mathcal{G}_\varepsilon J$ are constructed by gluing the optimal microstructure $(m, J) = (m^z, J^z)_{z\in\mathbb{Z}^d_\varepsilon}$ from the previous lemma. It is then tempting to seek for an estimate of the form

$$\mathcal{F}_\varepsilon(\widehat{m}, \widehat{J}) \le \sum_{z\in\mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{m^z}{\varepsilon^d}, \frac{J^z}{\varepsilon^{d-1}}\right) + \{\text{small error}\}.$$

However, (m, J) does not have the required *a priori* regularity estimates to obtain such a bound. Moreover, the gluing procedure does not necessarily produce solutions to the discrete continuity equation if we start with solutions to the continuous continuity equation.

We conclude the subsection with the following L^1 and L^∞ estimates.

Lemma 3.8.13 (L^1 and L^∞ estimates). *Let $(\mu_t, \nu_t)_{t\in\mathcal{I}} \subset \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$ be a regular curve of measures satisfying*

$$M := \sup_{t\in\mathcal{I}} \mu_t(\mathbb{T}^d) < \infty \quad \text{and} \quad A := \mathbb{A}^{\mathcal{I}}_{\text{hom}}(\boldsymbol{\mu}, \boldsymbol{\nu}) < \infty. \tag{3.91}$$

Let $(m^z_t, J^z_t)_{z\in\mathbb{Z}^d_\varepsilon} \subseteq \mathcal{M}_+(\mathbb{T}^d) \times \mathcal{M}^d(\mathbb{T}^d)$ be corresponding optimal microstructures. Then:

(i) $(\mathsf{P}_\varepsilon\mu,\mathsf{P}_\varepsilon\nu)$ *satisfies the uniform estimate*

$$\sup_{\varepsilon>0}\sup_{t\in\mathcal{I}}\|\mathsf{P}_\varepsilon\mu_t\|_{\ell^1(\mathbb{Z}^d_\varepsilon)} = M. \tag{3.92}$$

(ii) $(m_t, J_t)_{t\in\mathcal{I}}$ *satisfies the uniform estimate*

$$\sup_{\varepsilon>0}\sup_{(t,x)\in\mathcal{I}\times\mathcal{X}}\sum_{z\in\mathbb{Z}^d_\varepsilon} m_t^z(x) \le M \tag{3.93}$$

$$\sup_{\varepsilon>0}\sup_{(x,y)\in\mathcal{E}} \varepsilon\int_{\mathcal{I}}\sum_{z\in\mathbb{Z}^d_\varepsilon}\left|J_t^z(x,y)\right|\,\mathrm{d}t \lesssim A+M. \tag{3.94}$$

Proof. The first claim follows since $\|P_\varepsilon\mu_t\|_{\ell^1(\mathbb{Z}^d_\varepsilon)} = \mu_t(\mathbb{T}^d)$ by construction.

To prove (ii), note that

$$\sum_{z\in\mathbb{Z}^d_\varepsilon}\sum_{x\in\mathcal{X}^Q} m_t^z(x) = \sum_{z\in\mathbb{Z}^d_\varepsilon} P_\varepsilon\mu(z) = \mu_t(\mathbb{T}^d),$$

which yields (3.93).

To prove (3.94), we use the growth condition on F, the periodicity of J_t^z, and (i) to obtain for $(x,y)\in\mathcal{E}$ and $t\in\mathcal{I}$:

$$\begin{aligned}\varepsilon\sum_{z\in\mathbb{Z}^d_\varepsilon}\left|J_t^z(x,y)\right| \le \sum_{z\in\mathbb{Z}^d_\varepsilon}\varepsilon^d\sum_{(\tilde{x},\tilde{y})\in\mathcal{E}^Q}\left|\frac{J_t^z(\tilde{x},\tilde{y})}{\varepsilon^{d-1}}\right| &\lesssim \sum_{z\in\mathbb{Z}^d_\varepsilon}\varepsilon^d F\left(\frac{m_t^z}{\varepsilon^d},\frac{J_t^z}{\varepsilon^{d-1}}\right) + M\\ &\lesssim \int_{\mathbb{T}^d} f_{\mathrm{hom}}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}x},\frac{\mathrm{d}j_t}{\mathrm{d}x}\right)\mathrm{d}x + M\,,\end{aligned}$$

where in the last inequality we applied Proposition 3.8.11. Integrating in time and taking the supremum in space and $\varepsilon>0$, we obtain (3.94). □

3.8.2 Approximation result

The goal of this subsection is to show that despite the issues of Remark 3.8.12, we can find a solution to $\mathcal{CE}_\varepsilon^{\mathcal{I}}$ with almost optimal energy that is $\|\cdot\|_{\mathrm{KR}}$-close to a glued optimal microstructure.

In the following result, $\mathcal{I}_\eta = (a-\eta, b+\eta)$ denotes the η-extension of the open interval $\mathcal{I}=(a,b)$ for $\eta>0$.

Proposition 3.8.14 (Approximation of optimal microstructures)**.** *Let $(\boldsymbol{\mu},\boldsymbol{\nu})\in\mathbb{CE}^{\mathcal{I}_\eta}$ be a regular curve of measures sastisfying*

$$M := \mu_0(\mathbb{T}^d) < \infty \quad \text{and} \quad A := \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_\eta}(\boldsymbol{\mu},\boldsymbol{\nu}) < \infty.$$

Let $(m_t^z, J_t^z)_{t\in\mathcal{I},z\in\mathbb{Z}^d_\varepsilon}\subseteq\mathbb{R}_+^{\mathcal{X}}\times\mathbb{R}_a^{\mathcal{E}}$ be a measurable family of optimal microstructures associated to $(\mu_t,\nu_t)_{t\in\mathcal{I}}$ and consider their gluing $(\widehat{m}_t,\widehat{J}_t)_{t\in\mathcal{I}}\subseteq\mathbb{R}_+^{\mathcal{X}_\varepsilon}\times\mathbb{R}_a^{\mathcal{E}_\varepsilon}$. Then, for every $\eta'>0$, there exists $\varepsilon_0>0$ such that the following holds for all $0<\varepsilon\le\varepsilon_0$: there exists a solution $(\mathbf{m}^*,\mathbf{J}^*)\in\mathcal{CE}_\varepsilon^{\mathcal{I}}$ satisfying the bounds

$$\text{(measure approximation)} \qquad \|\iota_\varepsilon(\widehat{\mathbf{m}}-\mathbf{m}^*)\|_{\mathrm{KR}(\overline{\mathcal{I}}\times\mathbb{T}^d)} \le \eta', \tag{3.95a}$$

$$\text{(energy approximation)} \qquad \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^*,\mathbf{J}^*) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}(\boldsymbol{\mu},\boldsymbol{\nu}) + \eta' + C\varepsilon, \tag{3.95b}$$

where $C<\infty$ depends on M, A, $|\mathcal{I}|$, and η', but not on ε.

Remark 3.8.15. It is also true that

$$\mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^*, \mathbf{J}^*) \leq \mathcal{A}_\varepsilon^{\mathcal{I}}(\widehat{\mathbf{m}}, \widehat{\mathbf{J}}) + \eta' + C\varepsilon\,,$$

but this information is not "useful", as we do not expect to be able to control $\mathcal{A}_\varepsilon^{\mathcal{I}}(\widehat{\mathbf{m}}, \widehat{\mathbf{J}})$ in terms of $\mathbb{A}_{\text{hom}}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu})$; see also Remark 3.8.12.

The proof consists of four steps: the first one is to consider optimal microstructures associated with $(\boldsymbol{\mu}, \boldsymbol{\nu})$ on every scale $\varepsilon > 0$ and glue them together to obtain a discrete curves $(\mathbf{m}^*, \mathbf{J}^*)$ (we omit the ε-dependence for simplicity). The second step is the space-time regularisation of such measures in the same spirit as done in the proof of Proposition 3.7.1. Subsequently, we aim at finding suitable correctors in order to obtain a solution to the continuity equation and thus a discrete competitor (in the definition of $\mathcal{A}_\varepsilon$). Finally, the energy estimates conclude the proof of Proposition 3.8.14.

Let us first discuss the third step, i.e. how to find small correctors for $(\mathbf{m}^*, \mathbf{J}^*)$ in order to obtain discrete solutions to $\mathcal{CE}_\varepsilon^{\mathcal{I}}$ which are close to the first ones. Suppose for a moment that $(\mathbf{m}^*, \mathbf{J}^*)$ are "regular", as in the outcome of Proposition 3.7.1. Then the idea is to consider how far they are from solving the continuity equation, i.e. to study the error in the continuity equation

$$g_t(x) := \partial_t m_t^*(x) + \mathsf{div}\, J_t^*(x)\,, \quad x \in \mathcal{X}_\varepsilon\,,$$

and find suitable (small) correctors $\tilde{\mathbf{J}}$ to $\mathbf{J}^*$ in such a way that $(\mathbf{m}^*, \mathbf{J}^* + \tilde{\mathbf{J}}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$.

This is based on the next result, which is obtained on the same spirit of Lemma 3.7.3 in a non-periodic setting. In this case, we are able to ensure good ℓ^∞-bounds and support properties.

Lemma 3.8.16 (Bounds for the divergence equation)**.** *Let $g : \mathcal{X}_\varepsilon \to \mathbb{R}$ with $\sum_{x\in\mathcal{X}_\varepsilon} g(x) = 0$. There exists a vector field $J : \mathcal{E}_\varepsilon \to \mathbb{R}$ such that*

$$\mathsf{div}\, J = g \quad \textit{and} \quad \|J\|_{\ell^\infty(\mathcal{E}_\varepsilon)} \leq \tfrac{1}{2}\|g\|_{\ell^1(\mathcal{X}_\varepsilon)}. \tag{3.96}$$

Moreover, $\operatorname{supp} V \subseteq \operatorname{conv}\operatorname{supp} g + B_{C\varepsilon}$ *with C depending only on $\mathcal{X}$.*

Proof. Let g_+ be the positive part of g, and let g_- be the negative part. By assumption, these functions have the same ℓ^1-norm $N := \|g_-\|_{\ell^1(\mathcal{X}_\varepsilon)} = \|g_+\|_{\ell^1(\mathcal{X}_\varepsilon)}$. Let Γ be an arbitrary coupling between the discrete probability measures g_-/N and g_+/N.

For any $x, y \in \operatorname{supp} g$: take an arbitrary path P_{xy} connecting these two points. Let $J_x y$ be the unit flux field constructed in Definition 3.4.4. Then the vector field $J := \sum_{x,y} \Gamma(x,y) J_{xy}$ has the desired properties. □

Remark 3.8.17 (Measurability). It is clear from the previous proof that one can choose the vector field $J : \mathcal{E}_\varepsilon \to \mathbb{R}$ in such a way that the function $g \mapsto J$ is a measurable map.

The plan is to apply Lemma 3.8.16 to a suitable localisation of g_t, in each cube Q_ε^z, for every $z \in \mathbb{Z}_\varepsilon^d$. Precisely, the goal is to find $g_t(z;\cdot)$ for every $z \in \mathbb{Z}_\varepsilon^d$ such that

$$\sum_{z\in\mathbb{Z}_\varepsilon^d} g_t(z;x) = g_t(x)\,, \quad \sum_{x\in\mathcal{X}_\varepsilon} g_t(z;x) = 0\,, \tag{3.97}$$

which is small on the right scale, meaning

$$\operatorname{supp} g_t(z;\cdot) \subset B_\infty(z, R\varepsilon)\,, \quad \|g_t(z;\cdot)\|_\infty \leq C\varepsilon^d\,. \tag{3.98}$$

Remark 3.8.18. Note that $\sum_{x\in\mathcal{X}_\varepsilon} g_t(x) = 0$ for all $t \in \mathcal{I}$, since $\mathbf{m}^*$ has constant mass in time and $\mathbf{J}^*$ is skew-symmetric. However, an application of Lemma 3.8.16 without localisation would not ensure a uniform bound on the corrector field, as we are not able to control the ℓ^1-norm of g_t *a priori*.

Remark 3.8.19. A seemingly natural attempt would be to define $g_t(z;x) := g_t(x)\mathbb{1}_{\{z\}}(x_{\mathsf{z}})$. However, this choice is not of zero-mass, due to the flow of mass across the boundary of the cubes.

Recall that we use the notation $(\mathbf{r},\mathbf{u}) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$ to denote solutions to the continuity equation on $\mathbb{Z}^d_\varepsilon$ in the sense of Definition 3.8.2. We shall later apply Lemma 3.8.22 to the pair $(\mathbf{r},\mathbf{u}) = (\mathsf{P}_\varepsilon\boldsymbol{\mu}, \mathsf{P}_\varepsilon\boldsymbol{\nu}) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$, thanks to Lemma 3.8.3.

The notion of *shortest path* in the next definition refers to the ℓ_1-distance on $\mathbb{Z}^d_\varepsilon$.

Definition 3.8.20. For all $z', z'' \in \mathbb{Z}^d_\varepsilon$, we choose simultaneously a shortest path $p(z',z'') := (z_0,\dots,z_N)$ of nearest neighbors in $\mathbb{Z}^d_\varepsilon$ connecting $z_0 = z'$ to $z_N = z''$ such that $p(z'+\tilde z, z''+\tilde z) = p(z',z'') + \tilde z$ for all $\tilde z \in \mathbb{Z}^d_\varepsilon$. Then define for $z, z', z'' \in \mathbb{Z}^d_\varepsilon$ and $i = 1,\dots,d$ the signs $\sigma_i^{z;z',z''} \in \{-1,0,1\}$ as

$$\sigma_i^{z;z',z''} := \begin{cases} -1 & \text{if } (z_{k-1},z_k) = (z, z-e_i) \text{ for some } k \text{ within } p(z',z''),\\ 1 & \text{if } (z_{k-1},z_k) = (z-e_i, z) \text{ for some } k \text{ within } p(z',z''),\\ 0 & \text{otherwise.}\end{cases}$$

Note that since the paths $p(z',z'')$ are simple, each pair of nearest neighbours appears at most once in any order, so that $\sigma_i^{z;z',z''}$ is well-defined.

It follows readily from Definition 3.8.20 that

$$\sum_{z\in\mathbb{Z}^d_\varepsilon} \sigma_i^{z;z',z''} = (z''-z')\cdot e_i \tag{3.99}$$

for all $z', z'' \in \mathbb{Z}^d_\varepsilon$ and $i = 1,\dots,d$.

Remark 3.8.21. A canonical choice of the paths $p(z',z'')$ is to interpolate first between $z'_1 \in \mathbb{Z}^1_\varepsilon$ and $z''_1 \in \mathbb{Z}^1_\varepsilon$ one step at a time, then between z'_2 and z''_2, and so on. The precise choice of path is irrelevant to our analysis as long as paths are short and satisfy $p(z'+\tilde z, z''+\tilde z) = p(z',z'')+\tilde z$. Since the paths are invariant under translations, so are the signs, i.e.

$$\sigma_i^{z;z'+\tilde z,z''+\tilde z} = \sigma_i^{z-\tilde z;z',z''} \tag{3.100}$$

for all $z, \tilde z, z', z'' \in \mathbb{Z}^d_\varepsilon$, which is used in the prof of Lemma 3.8.22 below.

Lemma 3.8.22 shows that if we start from a solution to the continuity equation $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ and consider an admissible microstructure $(\mathbf{m},\mathbf{J}) = (m_t^z, J_t^z)_{t\in\mathcal{I}, z\in\mathbb{Z}^d_\varepsilon}$ associated to $(\mathsf{P}_\varepsilon\boldsymbol{\mu}, \mathsf{P}_\varepsilon\boldsymbol{\nu})$, then it is possible to localise the error in the continuity equation arising from the gluing $(\mathcal{G}_\varepsilon\mathbf{M}, \mathcal{G}_\varepsilon\mathbf{U})$ as in (3.97).

Lemma 3.8.22 (Localisation of the error to $\mathcal{CE}^{\mathcal{I}}_\varepsilon$)**.** *Let $(\mathbf{r},\mathbf{u}) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$ and suppose that $m_t := (m_t^z)_{z\in\mathbb{Z}^d_\varepsilon} \subseteq \mathbb{R}^{\mathcal{X}}_+$ and $J_t := (J_t^z)_{z\in\mathbb{Z}^d_\varepsilon} \subseteq \mathbb{R}^{\mathcal{E}}_a$ satisfy*

$$(m_t^z, J_t^z) \in \mathsf{Rep}\Big(r_t(z), u_t(z)\Big)$$

for every $t \in \mathcal{I}$ and $z \in \mathbb{Z}^d_\varepsilon$. Consider their gluings $\widehat{m}_t := \mathcal{G}_\varepsilon m_t$ and $\widehat{J}_t := \mathcal{G}_\varepsilon J_t$ and define, for $z \in \mathbb{Z}^d_\varepsilon$ and $x \in \mathcal{X}_\varepsilon$,

$$g_t(x) := \partial_t \widehat{m}_t(x) + \operatorname{div} \widehat{J}_t(x), \tag{3.101}$$

$$g_t(z;x) := \partial_t \widehat{m}_t(x) \mathbb{1}_{\{z\}}(x_\mathsf{z}) + \frac{1}{2} \sum_{y \sim x} \sum_{i=1}^d \sigma_i^{z;x_\mathsf{z},y_\mathsf{z}} \Big(\tilde{J}_t(z;x,y) - \tilde{J}_t(z-e_i;x,y) \Big), \tag{3.102}$$

where $\tilde{J}_t(z;\cdot) : \mathcal{E}_\varepsilon \to \mathbb{R}$ is the $\mathbb{T}^d_\varepsilon$-periodic map satisfying $\tilde{J}_t\Big(z; T^0_\varepsilon(x'), T^0_\varepsilon(y')\Big) = J^z_t(x',y')$ for all $(x',y') \in \mathcal{E}$. Then the following statements hold for every $t \in \mathcal{I}$:

(i) $g_t(z;x)$ is a localisation of the error $g_t(x)$ of $(\widehat{m}, \widehat{J})$ from solving $\mathcal{CE}^{\mathcal{I}}_\varepsilon$, i.e.,

$$\sum_{z \in \mathbb{Z}^d_\varepsilon} g_t(z;x) = g_t(x) \quad \text{for all } x \in \mathcal{X}_\varepsilon.$$

(ii) Each localised error $g_t(z;\cdot)$ has zero mass, i.e.,

$$\sum_{x \in \mathcal{X}_\varepsilon} g_t(z;x) = 0 \quad \text{for all } z \in \mathbb{Z}^d_\varepsilon.$$

Proof. (i): For $(x,y) \in \mathcal{E}_\varepsilon$, consider the path $p(x_\mathsf{z}, y_\mathsf{z}) = (z_0, \ldots, z_N)$ constructed in Definition 3.8.20. For all $t \in \mathcal{I}$ we have

$$\begin{aligned} &\sum_{z \in \mathbb{Z}^d_\varepsilon} \sum_{i=1}^d \sigma_i^{z;x_\mathsf{z},y_\mathsf{z}} \Big(\tilde{J}_t(z;x,y) - \tilde{J}_t(z-e_i;x,y) \Big) \\ &= \sum_{k=1}^N \Big(\tilde{J}_t(z_k;x,y) - \tilde{J}_t(z_{k-1};x,y) \Big) = \tilde{J}_t(y_\mathsf{z};x,y) - \tilde{J}_t(x_\mathsf{z};x,y). \end{aligned}$$

Summation over all neighbours of $x \in \mathcal{X}_\varepsilon$ yields, for all $t \in \mathcal{I}$,

$$\begin{aligned} \sum_{z \in \mathbb{Z}^d_\varepsilon} g_t(z;x) &= \partial_t m_t(x) + \frac{1}{2} \sum_{y \sim x} \sum_{z \in \mathbb{Z}^d_\varepsilon} \sum_{i=1}^d \sigma_i^{z;x_\mathsf{z},y_\mathsf{z}} \Big(\tilde{J}_t(z;x,y) - \tilde{J}_t(z-e_i;x,y) \Big) \\ &= \partial_t m_t(x) + \frac{1}{2} \sum_{y \sim x} \Big(\tilde{J}_t(y_\mathsf{z};x,y) - \tilde{J}_t(x_\mathsf{z};x,y) \Big) \\ &= \partial_t m_t(x) + \frac{1}{2} \sum_{y \sim x} \Big(\tilde{J}_t(y_\mathsf{z};x,y) + \tilde{J}_t(x_\mathsf{z};x,y) \Big) = g_t(x), \end{aligned}$$

where we used the $\mathbb{Z}^d$-periodicity of $(\mathcal{X}, \mathcal{E})$ and the vanishing divergence of $J^{x_\mathsf{z}}_t$.

(ii): Fix $z \in \mathbb{Z}^d_\varepsilon$ and $t \in \mathcal{I}$. Using the periodicity of $\tilde{J}_t(z;\cdot)$, the identity (3.100), the group structure of $\mathbb{Z}^d_\varepsilon$, the relation between $\tilde{J}$ and J, the fact that $J^z_t \in \operatorname{Rep}\Big(u_t(z)\Big)$, and the identity

(3.99), we obtain

$$\begin{aligned}
&\sum_{(x,y)\in\mathcal{E}_\varepsilon}\sum_{i=1}^d \sigma_i^{z;x_z,y_z}\Big(\tilde{J}_t(z;x,y)-\tilde{J}_t(z-e_i;x,y)\Big)\\
&=\sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon\\ x_z=z}}\sum_{\tilde z\in\mathbb{Z}_\varepsilon^d}\sum_{i=1}^d \sigma_i^{z;x_z+\tilde z,y_z+\tilde z}\Big(\tilde{J}_t(z;x,y)-\tilde{J}_t(z-e_i;x,y)\Big)\\
&=\sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon\\ x_z=z}}\sum_{\tilde z\in\mathbb{Z}_\varepsilon^d}\sum_{i=1}^d \sigma_i^{z-\tilde z;x_z,y_z}\Big(\tilde{J}_t(z;x,y)-\tilde{J}_t(z-e_i;x,y)\Big)\\
&=\sum_{\substack{(x,y)\in\mathcal{E}_\varepsilon\\ x_z=z}}\sum_{i=1}^d \Big(\tilde{J}_t(z;x,y)-\tilde{J}_t(z-e_i;x,y)\Big)\Bigg(\sum_{\tilde z\in\mathbb{Z}_\varepsilon^d}\sigma_i^{\tilde z;x_z,y_z}\Bigg)\\
&=\sum_{(x',y')\in\mathcal{E}^Q}\sum_{i=1}^d \Big(J_t^z(x',y')-J_t^{z-e_i}(x',y')\Big)(y'_z-x'_z)\cdot e_i\\
&=2\sum_{i=1}^d\Big(u_t(z)-u_t(z-e_i)\Big)\cdot e_i .
\end{aligned}$$

By definition of $g_t(z;\cdot)$ we obtain

$$\begin{aligned}
\sum_{x\in\mathcal{X}_\varepsilon} g_t(z;x) &= \sum_{\substack{x\in\mathcal{X}_\varepsilon\\ x_z=z}}\partial_t m_t(x)+\frac12\sum_{i=1}^d\sum_{(x,y)\in\mathcal{E}_\varepsilon}\sigma_i^{z;x_z,y_z}\Big(\tilde{J}_t(z;x,y)-\tilde{J}_t(z-e_i;x,y)\Big)\\
&=\partial_t r_t(z)+\sum_{i=1}^d\Big(u_t(z)-u_t(z-e_i)\Big)\cdot e_i=0,
\end{aligned}$$

where we used that $m_t^z\in \mathsf{Rep}\big(r_t(z)\big)$ and eventually that $(\mathbf{r},\mathbf{u})\in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$. □

Now we are ready to prove Proposition 3.8.14.

Proof of Proposition 3.8.14. The proof consists of four steps. For simplicity: $\mathcal{I}:=\mathcal{I}_\eta$.

Step 1: Regularisation. Recall the operators R_δ, S_λ, and T_τ as defined in Section 3.7.1. We define

$$\mathbf{m}^*:=\Big(R_\delta\circ S_\lambda\circ T_\tau\Big)\widehat{\mathbf{m}}\quad\text{and}\quad \bar{\mathbf{J}}^*:=\Big(R_\delta\circ S_\lambda\circ T_\tau\Big)\hat{\mathbf{J}}\,,$$

where $\delta,\lambda>0$, $0<\tau<\eta$ will be chosen sufficiently small, depending on the desired accuracy $\eta'>0$. Due to special linear structure of the gluing operator $\mathcal{G}_\varepsilon$, it is clear that

$$\mathbf{m}^*=\mathcal{G}_\varepsilon\overline{\mathbf{m}}\quad\text{and}\quad \bar{\mathbf{J}}^*=\mathcal{G}_\varepsilon\bar{\mathbf{J}}\,,$$

for some $\big(\overline{\mathbf{m}},\bar{\mathbf{J}}\big)=(\overline{m}_t^z,\bar{J}_t^z)_{t\in\mathcal{I},z\in\mathbb{Z}_\varepsilon^d}$. More precisely, they correspond to the regularised version of the measures $(m_t^z,J_t^z)_{t\in\mathcal{I},z\in\mathbb{Z}_\varepsilon^d}$ with respect to the graph structure of $\mathbb{Z}_\varepsilon^d$. In particular, an application[2] of Lemma 3.8.13, Lemma 3.7.6, and Lemma 3.7.7 yields

$$\begin{aligned}
\sup_{t\in\mathcal{I}}\Big\|\overline{m}_t^{\cdot+z}-\overline{m}_t\Big\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d\times\mathcal{X})}+\varepsilon\Big\|\bar{J}_t^{\cdot+z}-\bar{J}_t\Big\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d\times\mathcal{E})}&\le C|z|\varepsilon^{d+1}\,,\\
\sup_{t\in\mathcal{I}}\Big\|\partial_t\overline{m}_t\Big\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d\times\mathcal{X})}&\le C\varepsilon^d\,,
\end{aligned}\tag{3.103}$$

[2] To be precise, this is an application of these lemmas to the case of $V:=\{v\}$, thus $\mathcal{X}_\varepsilon\simeq\mathbb{Z}_\varepsilon^d$.

for any $z \in \mathbb{Z}^d_\varepsilon$, as well as the domain regularity

$$\left\{ \left(\frac{\overline{m}^z_t}{\varepsilon^d}, \frac{\bar{J}^z_t}{\varepsilon^{d-1}} \right) \ : \ z \in \mathbb{Z}^d_\varepsilon, \, t \in \mathcal{I} \right\} \subset K \Subset (\mathsf{D}\, F)^\circ \,, \tag{3.104}$$

for a constant C and a compact set K depending only on M, A, δ, λ, and τ. We can then apply Lemma 3.8.7 and deduce that for every $t \in \mathcal{I}$, $\varepsilon \le \varepsilon_0$ (depending on K and F),

$$\mathcal{F}_\varepsilon\Big(m^*_t, \bar{J}^*_t\Big) \le \sum_{z \in \mathbb{Z}^d_\varepsilon} \varepsilon^d F\left(\frac{\overline{m}^z_t}{\varepsilon^d}, \frac{\bar{J}^z_t}{\varepsilon^{d-1}} \right) + c\varepsilon \,, \tag{3.105}$$

for a $c \in \mathbb{R}^+$ depending on the same set of parameters (via C and $\mathrm{Lip}(F;K)$) and R_1.

Step 2: Construction of a solution to $\mathcal{CE}^{\mathcal{I}}_\varepsilon$. From now on, the constants C appearing in the estimates might change line by line, but it always depends on the same set of parameters as the constant C in Step 1, and possibly on the size of the time interval $|\mathcal{I}|$.

The next step is to find a quantitative small corrector $\mathbf{V}$ in such a way that $(\mathbf{m}^*, \bar{\mathbf{J}}^* + \mathbf{V}) \in \mathcal{CE}^{\mathcal{I}}_\varepsilon$. To do so, we observe that by construction we have for every $t \in \mathcal{I}$

$$\Big(\overline{m}^z_t, \bar{J}^z_t \Big) \in \mathsf{Rep}\Big(r^*_t(z), u^*_t(z) \Big) \,,$$

where $(\mathbf{r}^*, \mathbf{u}^*) \in \mathsf{CE}^{\mathcal{I}}_{\varepsilon,d}$ (by the linearity of equation (3.84)). Consider the corresponding error functions, for $(x,y) \in \mathcal{E}_\varepsilon$, $t \in \mathcal{I}$, $z \in \mathbb{Z}^d_\varepsilon$ given by (3.101) and (3.102),

$$\begin{aligned} g_t(x) &:= \partial_t m^*_t(x) + \mathsf{div}\, \bar{J}^*_t(x) \,, \\ g_t(z;x) &:= \partial_t m^*_t(x) \mathbb{1}_{\{x_z = z\}}(x) + \frac{1}{2} \sum_{y \sim x} \sum_{i=1}^d \sigma_i^{z;x_z,y_z} \big(\tilde{J}(z;x,y) - \tilde{J}(z - e_i; x,y) \big) \,, \end{aligned}$$

where $\tilde{J}(z;\cdot) : \mathcal{E}_\varepsilon \to \mathbb{R}$ is the $\mathbb{T}^d_\varepsilon$-periodic map satisfying $\tilde{J}(z; T^0_\varepsilon(x'), T^0_\varepsilon(y')) = \bar{J}^z_t(x',y')$, for any $(x',y') \in \mathcal{E}$. Thanks to Lemma 3.8.22, we know that

$$\sum_{x \in \mathcal{X}_\varepsilon} g_t(z;x) = 0 \,, \qquad \sum_{z' \in \mathbb{Z}^d_\varepsilon} g_t(z';x) = g_t(x) \,, \qquad \forall x \in \mathcal{X}_\varepsilon \,, \, z \in \mathbb{Z}^d_\varepsilon \,.$$

Moreover, from the regularity estimates (3.103) and the local finiteness of the graph $(\mathcal{X}, \mathcal{E})$, we infer for every $z \in \mathbb{Z}^d_\varepsilon$

$$\| g_t(z;\cdot) \|_{\ell^\infty(\mathcal{X}_\varepsilon)} \le C\varepsilon^d \,, \quad \operatorname{supp} g_t(z;\cdot) \subset \{ x \in \mathcal{X}_\varepsilon \ : \ \| x_\mathsf{z} - z \|_{\ell^\infty(\mathbb{Z}^d_\varepsilon)} \le C' \} \,, \tag{3.106}$$

where C' only depends on $(\mathcal{X}, \mathcal{E})$. Hence, as an application of Lemma 3.8.16, we deduce the existence of corrector vector fields $V_t \in \mathbb{R}_a^{\mathbb{Z}^d_\varepsilon \times \mathcal{E}_\varepsilon}$ such that

$$\begin{gathered} \mathsf{div}\, V_t(z;\cdot) = g_t(z;\cdot) \,, \quad \operatorname{supp} V_t(z;\cdot) \subset \{ (x,y) \in \mathcal{E}_\varepsilon \ : \ \| x_\mathsf{z} - z \|_{\ell^\infty(\mathbb{Z}^d_\varepsilon)} \le \tilde{C}' \} \,, \\ \| V_t(z;\cdot) \|_{\ell^\infty(\mathcal{E}_\varepsilon)} \le \tfrac{1}{2} \| g_t(z;\cdot) \|_{\ell^1(\mathcal{X}_\varepsilon)} \le C\varepsilon^d \,, \end{gathered} \tag{3.107}$$

for every $t \in \mathcal{I}$, $z \in \mathbb{Z}^d_\varepsilon$. The existence of a measurable (in $t \in \mathcal{I}$ and $z \in \mathbb{Z}^d_\varepsilon$) map $V_t(z;\cdot)$ follows from the measurability of $g_t(z;\cdot)$ and Remark 3.8.17.

We then define $\mathbf{V} : \mathcal{I} \to \mathbb{R}^{\mathcal{E}_\varepsilon}_a$ and $\mathbf{J}^* : \mathcal{I} \to \mathbb{R}^{\mathcal{E}_\varepsilon}_a$ as

$$\mathbf{V} := \sum_{z \in \mathbb{Z}^d_\varepsilon} \mathbf{V}(z;\cdot) \,, \qquad \mathbf{J}^* := \bar{\mathbf{J}}^* + \mathbf{V} \,,$$

and obtain a solution to the discrete continuity equation $(\mathbf{m}^*, \mathbf{J}^*) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$.

Step 3: Energy estimates. The locality property (3.107) of $V_t(z;\cdot)$ and local finiteness of the graph $(\mathcal{X}, \mathcal{E})$ allow us to deduce the same uniform estimates on the global corrector as well. Indeed for every $t \in \mathcal{I}$, $x \in \mathcal{X}_\varepsilon$ we have

$$V_t(x,y) := \sum_{z \in B_\infty(x_{\mathsf{z}}; \tilde{C}')} V(z;x,y)\,, \quad B_\infty(x_{\mathsf{z}}; \tilde{C}') := \left\{ z \in \mathbb{Z}_\varepsilon^d \ : \ \|z - x_{\mathsf{z}}\|_{\ell^\infty(\mathbb{Z}_\varepsilon^d)} \le \tilde{C}' \right\}\,,$$

and hence from the estimate (3.107) we also deduce $\|\mathbf{V}\|_{\ell^\infty(\mathcal{I}\times\mathcal{E}_\varepsilon)} \le C\varepsilon^d$.

Since (3.104) implies that $\left(\frac{\tau_\varepsilon^z m_t^*}{\varepsilon^d}, \frac{\tau_\varepsilon^z \bar{J}_t^*}{\varepsilon^{d-1}}\right) \in K$, it then follows that $\left(\frac{\tau_\varepsilon^z m_t^*}{\varepsilon^d}, \frac{\tau_\varepsilon^z J_t^*}{\varepsilon^{d-1}}\right) \in K'$ for $0 < \varepsilon \le \varepsilon_0$ sufficiently small, where ε_0 depends on K and C. Here K' is a compact set, possibly slightly larger than K, contained in $\mathsf{D}(F)^\circ$.

Therefore, we can estimate the energy

$$\sup_{t\in\mathcal{I}} \sup_{z \in \mathbb{Z}_\varepsilon^d} \left| F\left(\frac{\tau_\varepsilon^z m_t^*}{\varepsilon^d}, \frac{\tau_\varepsilon^z \bar{J}_t^*}{\varepsilon^{d-1}}\right) - F\left(\frac{\tau_\varepsilon^z m_t^*}{\varepsilon^d}, \frac{\tau_\varepsilon^z J_t^*}{\varepsilon^{d-1}}\right) \right| \le \mathrm{Lip}(F;K') \frac{1}{\varepsilon^{d-1}} \|\mathbf{V}\|_{\ell^\infty(\mathcal{I}\times\mathcal{E}_\varepsilon)} \le C\varepsilon\,,$$

and hence $\mathcal{A}_\varepsilon^{\mathcal{I}}\big(\mathbf{m}^*, \mathbf{J}^*\big) \le \mathcal{A}_\varepsilon^{\mathcal{I}}\big(\mathbf{m}^*, \bar{\mathbf{J}}^*\big) + C\varepsilon$. Together with (3.105), this yields

$$\mathcal{A}_\varepsilon^{\mathcal{I}}\big(\mathbf{m}^*, \mathbf{J}^*\big) \le \int_{\mathcal{I}} \sum_{z\in\mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{\overline{m}_t^z}{\varepsilon^d}, \frac{\bar{J}_t^z}{\varepsilon^{d-1}}\right) \mathrm{d}t + C\varepsilon\,.$$

Finally, to control the energy of the regularised microstructures $(\bar{\mathbf{m}}, \bar{\mathbf{J}})$, we take advantage (as in (3.78)) of Lemma 3.7.5, Lemma 3.7.6 (i), and Lemma 3.7.7 (i) to obtain[3]

$$\begin{aligned}
\int_{\mathcal{I}} \sum_{z\in\mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{\overline{m}_t^z}{\varepsilon^d}, \frac{\bar{J}_t^z}{\varepsilon^{d-1}}\right) \mathrm{d}t &\le \int_{\mathcal{I}} \fint_{t-\tau}^{t+\tau} \sum_{z\in\mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{m_s^z}{\varepsilon^d}, \frac{J_s^z}{\varepsilon^{d-1}}\right) \mathrm{d}s\,\mathrm{d}t + \delta|\mathcal{I}|F(m^\circ, J^\circ) \\
&\le \int_{\mathcal{I}} \fint_{t-\tau}^{t+\tau} \mathbb{F}_{\mathrm{hom}}(\mu_s, \nu_s)\,\mathrm{d}s\,\mathrm{d}t + \delta|\mathcal{I}|F(m^\circ, J^\circ) + c'\varepsilon \\
&\le \int_{\mathcal{I}} \mathbb{F}_{\mathrm{hom}}(\mu_t, \nu_t)\,\mathrm{d}t + \delta|\mathcal{I}|F(m^\circ, J^\circ) + c'(\varepsilon + \tau)\,,
\end{aligned}$$

for a $c' < \infty$, where at last we used Proposition 3.8.11 and that f_{hom} is Lipschitz on $\tilde{K}$.

For every given $\eta' > 0$, the energy bound (3.95b) then follows choosing $\tau, \delta > 0$ small enough.

Step 4: Measures comparison. We have seen in (3.77) that Lemma 3.7.8 implies

$$\|\iota_\varepsilon \mathbf{m}^* - \iota_\varepsilon \widehat{\mathbf{m}}\|_{\mathrm{KR}([0,T]\times\mathbb{T}^d)} \lesssim M(\tau + \sqrt{\lambda} + \delta) + m^\circ(\mathcal{X}^Q)\delta\,,$$

where we also used that mass preservation of the gluing operator, see Remark 3.8.6. For every $\eta' > 0$, the distance bound (3.95a) can be then obtained choosing τ, λ, δ sufficiently small. □

[3] As before, it's an application of these lemmas on $\mathbb{Z}_\varepsilon^d$ (corresponding to $\mathbb{V} = \{v\}$).

3.8.3 Proof of the Γ-limsup inequality

This subsection is devoted to the proof of the Γ-limsup inequality in Theorem 3.5.1. First we formulate the existence of a recovery sequence in the smooth case.

Proposition 3.8.23 (Existence of a recovery sequence, smooth case)**.** *Fix $\mathcal{I} = (a,b)$, $a < b$, $\eta > 0$, and set $\mathcal{I}_\eta := (a-\eta, b+\eta)$. Let $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}_\eta}$ be a solution to the continuity equation with smooth densities $(\rho_t, j_t)_{t\in\mathcal{I}_\eta}$ and such that*

$$\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_\eta}(\boldsymbol{\mu},\boldsymbol{\nu}) < \infty \quad \text{and} \quad \Big\{\Big(\rho_t(x), j_t(x)\Big) \ : \ (t,x) \in \mathcal{I}_\eta \times \mathbb{T}^d\Big\} \Subset \mathsf{D}(f_{\mathrm{hom}})^\circ \,. \tag{3.108}$$

Then there exists a sequence of curves $(m_t^\varepsilon)_{t\in\overline{\mathcal{I}}} \subseteq \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}|_{\overline{\mathcal{I}}\times\mathbb{T}^d}$ weakly in $\mathcal{M}_+(\overline{\mathcal{I}} \times \mathbb{T}^d)$ as $\varepsilon \to 0$ and

$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_\eta}(\boldsymbol{\mu},\boldsymbol{\nu}) + C\eta|\mathcal{I}|\Big(\mu_0(\mathbb{T}^d)+1\Big)\,, \tag{3.109}$$

for some $C < \infty$.

Proof. We write $\mathrm{KR}_{\mathcal{I}} := \mathrm{KR}(\overline{\mathcal{I}} \times \mathbb{T}^d)$. Let $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}_\eta}$ be smooth curves of measures satisfying the assumptions (3.108). Let $(\widehat{\mathbf{m}}, \widehat{\mathbf{J}})$ be the gluing of a measurable family of optimal microstructure associated with $(\boldsymbol{\mu},\boldsymbol{\nu})$, for every $\varepsilon > 0$. For every $\eta' > 0$, Proposition 3.8.14 yields the existence of $(\mathbf{m}^{\eta'}, \mathbf{J}^{\eta'}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$, a constant $C_{\eta'}$, and $\varepsilon_0 = \varepsilon_0(\eta')$ depending on η' such that

$$\|\iota_\varepsilon(\mathbf{m}^{\eta'} - \widehat{\mathbf{m}})\|_{\mathrm{KR}_{\mathcal{I}}} \le \eta'\,, \quad \mathcal{A}_\varepsilon(\mathbf{m}^{\eta'}, \mathbf{J}^{\eta'}) \le \mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu},\boldsymbol{\nu}) + \eta' + \varepsilon C_{\eta'}\,,$$

for every $\varepsilon \le \varepsilon_0$.

Using Remark (3.8.6), in particular (3.86), and that $(\mathbf{m}^{\eta'}, \mathbf{J}^{\eta'}) \in \mathcal{CE}_\varepsilon^{\mathcal{I}}$, we infer

$$\|\iota_\varepsilon(\mathbf{m}^{\eta'}) - \boldsymbol{\mu}\|_{\mathrm{KR}_{\mathcal{I}}} \le \eta' + \boldsymbol{\mu}(\overline{\mathcal{I}} \times \mathbb{T}^d)\varepsilon^d\,, \quad \mathcal{A}_\varepsilon(\mathbf{m}^{\eta'}) \le \mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu},\boldsymbol{\nu}) + \eta' + \varepsilon C_{\eta'}\,.$$

for every $\varepsilon \le \varepsilon_0$. Therefore, we can find a diagunal sequence $\eta' = \eta'(\varepsilon) \to 0$ as $\varepsilon \to 0$ such that, if we set $\mathbf{m}^\varepsilon := \mathbf{m}^{\eta'(\varepsilon)}$, we obtain

$$\lim_{\varepsilon\to 0} \|\iota_\varepsilon(\mathbf{m}^\varepsilon) - \boldsymbol{\mu}\|_{\mathrm{KR}_{\mathcal{I}}} = 0\,,$$
$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}(\boldsymbol{\mu},\boldsymbol{\nu}) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_\eta}(\boldsymbol{\mu},\boldsymbol{\nu}) + C\eta|\mathcal{I}|(\mu_0(\mathbb{T}^d)+1)\,,$$

where at last we used the growth condition (3.21). □

In order to apply Proposition 3.8.23 for the existence of the recovery sequence in Theorem 3.5.1 we prove that the set of solutions to the continuity equation (3.23) with smooth densities are dense-in-energy for $\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}$.

Definition 3.8.24 (Affine change of variable in time)**.** Fix $\mathcal{I} = (a,b)$. For any $\eta > 0$, we consider the unique bijective increasing affine map $S^\eta : \mathcal{I} \to (a-2\eta, b+2\eta)$. For every interval $\tilde{\mathcal{I}} \subseteq \mathcal{I}$ and every vector-valued measure $\boldsymbol{\xi} \in \mathcal{M}^n(\tilde{\mathcal{I}} \times \mathbb{T}^d)$, $n \in \mathbb{N}$, we define the changed-variable measure

$$S^\eta[\boldsymbol{\xi}] \in \mathcal{M}^n(S^\eta(\tilde{\mathcal{I}}) \times \mathbb{T}^d)\,, \quad S^\eta[\boldsymbol{\xi}] := \frac{|\mathcal{I}| + 4\eta}{|\mathcal{I}|}\Big(S^\eta, id\Big)_{\#}\boldsymbol{\xi}\,. \tag{3.110}$$

Remark 3.8.25 (Properties of S^η). The scaling factor of $S^\eta[\boldsymbol{\xi}]$ is chosen so that if $\boldsymbol{\xi} \ll \mathscr{L}^{d+1}$, then $\mathsf{S}^\eta[\boldsymbol{\xi}] \ll \mathscr{L}^{d+1}$ and we have for $(t,x) \in S^\eta(\tilde{\mathcal{I}}) \times \mathbb{T}^d$ the equality of densities

$$\frac{\mathrm{d}\mathsf{S}^\eta[\boldsymbol{\xi}]}{\mathrm{d}\mathscr{L}^{d+1}}(t,x) = \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}\mathscr{L}^{d+1}}((S^\eta)^{-1}(t),x). \tag{3.111}$$

Moreover, if $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ then $\left(\frac{|\mathcal{I}|+4\eta}{|\mathcal{I}|}\mathsf{S}^\eta[\boldsymbol{\mu}], \mathsf{S}^\eta[\boldsymbol{\nu}]\right) \in \mathbb{CE}^{S^\eta(\mathcal{I})}$.

We are ready to state and prove the last result of this section.

Proposition 3.8.26 (Smooth approximation of finite energy solutions to $\mathbb{CE}^{\mathcal{I}}$)**.** *Fix $\mathcal{I} := (a,b)$ and fix $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{\mathcal{I}}$ with $\mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu},\boldsymbol{\nu}) < \infty$. Then there exists a sequence $\{\eta_k\}_k \subset \mathbb{R}^+$ such that $\eta_k \to 0$ as $k \to \infty$ and measures $(\boldsymbol{\mu}^k,\boldsymbol{\nu}^k) \in \mathbb{CE}^{\mathcal{I}_k}$ for $\mathcal{I}_k := (a-\eta_k, b+\eta_k)$ so that as $k \to \infty$*

$$(\boldsymbol{\mu}^k,\boldsymbol{\nu}^k) \to (\boldsymbol{\mu},\boldsymbol{\nu}) \text{ weakly in } \mathcal{M}_+\big(\mathcal{I}\times\mathbb{T}^d\big) \times \mathcal{M}^d\big(\mathcal{I}\times\mathbb{T}^d\big), \tag{3.112}$$

$$\frac{\mathrm{d}\boldsymbol{\mu}^k}{\mathrm{d}\mathscr{L}^{d+1}} \in \mathcal{C}_b^\infty\big(\mathcal{I}_k\times\mathbb{T}^d\big), \qquad \frac{\mathrm{d}\boldsymbol{\nu}^k}{\mathrm{d}\mathscr{L}^{d+1}} \in \mathcal{C}_b^\infty\big(\mathcal{I}_k\times\mathbb{T}^d;\mathbb{R}^d\big), \tag{3.113}$$

and such that the following energy bound holds true:

$$\limsup_{k\to\infty} \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}_k}(\boldsymbol{\mu}^k,\boldsymbol{\nu}^k) \le \mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}(\boldsymbol{\mu},\boldsymbol{\nu}). \tag{3.114}$$

Moreover, for any given $k \in \mathbb{N}$ we have the inclusion

$$\left\{\left(\frac{\mathrm{d}\boldsymbol{\mu}^k}{\mathrm{d}\mathscr{L}^{d+1}}(t,x), \frac{\mathrm{d}\boldsymbol{\nu}^k}{\mathrm{d}\mathscr{L}^{d+1}}(t,x)\right) : (t,x) \in \mathcal{I}_k \times \mathbb{T}^d\right\} \Subset (\mathsf{D}\, f_{\mathrm{hom}})^\circ. \tag{3.115}$$

Proof. Without loss of generality we can assume $f_{\mathrm{hom}} \ge 0$, if not we simply consider $g(\rho,j) = f_{\mathrm{hom}}(\rho,j) + C\rho + C$ for $C \in \mathbb{R}_+$ as in Lemma 3.3.14. For simplicity, we also assume $\mathcal{I} := (0,T)$, the extension to a generic interval is straightforward.

Fix $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^T$ with $\mathbb{A}_{\mathrm{hom}}(\boldsymbol{\mu},\boldsymbol{\nu}) < \infty$.

Step 1: regularisation. The first step is to regularise in time and space. To do so, we consider two sequences of smooth mollifiers $\phi_1^k : \mathbb{R} \to \mathbb{R}_+$, $\phi_2^k : \mathbb{T}^d \to \mathbb{R}$ for $k \in \mathbb{N}$ of integral 1, where $\operatorname{supp}\phi_1^k = [-\alpha_k,\alpha_k]$, $\operatorname{supp}\phi_2^k = B_{\frac{1}{k}}(0) \subset \mathbb{T}^d$ with $\alpha_k \to 0$ as $k \to \infty$ to be suitably chosen. We then set $\phi^k : \mathbb{R}\times\mathbb{T}^d \to \mathbb{R}_+$ as $\phi^k(t,x) := \phi_1^k(t)\phi_2^k(x)$.

We define space-time regular solutions to the continuity equation as

$$(\tilde{\boldsymbol{\mu}}^k,\tilde{\boldsymbol{\nu}}^k) := \phi^k * (\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathbb{CE}^{(\alpha_k, T-\alpha_k)},$$
$$(\widehat{\boldsymbol{\mu}}^k,\widehat{\boldsymbol{\nu}}^k) := \left(\frac{T+4\eta_k}{T}\mathsf{S}^{\eta_k}[\tilde{\boldsymbol{\mu}}^k], \mathsf{S}^{\eta_k}[\tilde{\boldsymbol{\nu}}^k]\right) \in \mathbb{CE}^{\mathcal{I}_k},$$

where $\mathcal{I}_k := S^{\eta_k}\big((\alpha_k, T-\alpha_k)\big)$. Note that the mollified measures are defined only We choose $\alpha_k := \frac{T\eta_k}{T+4\eta_k}$, so that $\mathcal{I}_k = (-\eta_k, T+\eta_k)$.

Finally, for (ρ°, j°) as given in (3.38), we define

$$(\boldsymbol{\mu}^k,\boldsymbol{\nu}^k) := (1-\delta_k)(\widehat{\boldsymbol{\mu}}^k,\widehat{\boldsymbol{\nu}}^k) + \delta_k(\rho^\circ,j^\circ)\mathscr{L}^{d+1} \in \mathbb{CE}^{\mathcal{I}_k}, \tag{3.116}$$

for some suitable choice of $\eta_k, \delta_k \to 0$.

Step 2: Properties of the regularised measures. First of all, we observe that $(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k) \ll \mathscr{L}^{d+1}$ with smooth densities for every $k \in \mathbb{N}$, so that (3.113) is satisfied. Secondly, the convergence (3.112) easily follows by the properties of the mollifiers and the fact that $S^\eta \to \mathrm{id}$ uniformly in $(0,T)$ as $\eta \to 0$.

Moreover, we note that for $t > 0$, using that $\mu_t(\mathbb{T}^d)$ is constant on $(0,T)$ one gets

$$\begin{aligned} \sup_{t\in(\alpha_k, T-\alpha_k)} \left\| \frac{\mathrm{d}\tilde{\mu}_t^k}{\mathrm{d}x} \right\|_\infty &\le \|\phi_2^k\|_\infty \boldsymbol{\mu}\Big((0,T)\times\mathbb{T}^d\Big) =: C_k < +\infty, \\ \left\| \frac{\mathrm{d}\tilde{\boldsymbol{\nu}}^k}{\mathrm{d}\mathscr{L}^{d+1}} \right\|_\infty &\le \|\phi^k\|_\infty |\boldsymbol{\nu}|\Big((0,T)\times\mathbb{T}^d\Big) < \infty\,, \end{aligned} \tag{3.117}$$

and thanks to (3.111) an analogous uniform estimate holds true for $(\widehat{\boldsymbol{\mu}}^k, \widehat{\boldsymbol{\nu}}^k)$ too. We can then apply Lemma B.3.1 and find convex compact sets $K_k \subset (\mathrm{D}\, f_{\mathrm{hom}})^\circ$ such that $\left\{\left(\frac{\mathrm{d}\boldsymbol{\mu}^k}{\mathrm{d}\mathscr{L}^{d+1}}(\cdot), \frac{\mathrm{d}\boldsymbol{\nu}^k}{\mathrm{d}\mathscr{L}^{d+1}}(\cdot)\right)\right\} \subset K_k$, so that (3.115) follows.

Additionally, pick $\theta > 0$ such that $B^\circ := B((\rho^\circ, j^\circ), \theta) \subset (\mathrm{D}\, f_{\mathrm{hom}})^\circ$. From (3.111), if one sets $S_k := S^{\eta_k}$, we see that

$$\left(\frac{\mathrm{d}\boldsymbol{\mu}^k}{\mathrm{d}\mathscr{L}^{d+1}}, \frac{\mathrm{d}\boldsymbol{\nu}^k}{\mathrm{d}\mathscr{L}^{d+1}}\right)(t,x) = (1-\delta_k)\left(\frac{\mathrm{d}\tilde{\boldsymbol{\mu}}^k}{\mathrm{d}\mathscr{L}^{d+1}}, \frac{\mathrm{d}\tilde{\boldsymbol{\nu}}^k}{\mathrm{d}\mathscr{L}^{d+1}}\right)(S_k^{-1}(t), x) + \delta_k(\tilde{\rho}_t^k(x), j^\circ) \tag{3.118}$$

for $t \in \mathcal{I}_k$ and $x \in \mathbb{T}^d$, where the functions $\tilde{\rho}^k$ are given by

$$\tilde{\rho}_t^k(x) := \rho^\circ + \frac{1-\delta_k}{\delta_k} 2\eta_k \frac{\mathrm{d}\tilde{\boldsymbol{\mu}}^k}{\mathrm{d}\mathscr{L}^{d+1}}(S_k^{-1}(t), x).$$

We choose δ_k such that $\theta\delta_k > 2\eta_k C_k$ and from (3.117) we get that

$$(\tilde{\rho}_t^k(x), j^\circ) \in B^\circ\,, \quad \forall t \in \mathcal{I}_k\,,\ x \in \mathbb{T}^d\,,\ k \in \mathbb{N}\,. \tag{3.119}$$

For example we can pick $\eta_k := (4kC_k)^{-1}$ and $\theta\delta_k = k^{-1}$, both going to zero when $k \to +\infty$.

Step 3: energy estimation. As the next step we show that

$$\mathbb{A}_{\mathrm{hom}}^{(\alpha_k, T-\alpha_k)}\Big(\tilde{\boldsymbol{\mu}}^k, \tilde{\boldsymbol{\nu}}^k\Big) \le \mathbb{A}_{\mathrm{hom}}^T(\boldsymbol{\mu}, \boldsymbol{\nu}), \quad \forall k \in \mathbb{N}\,. \tag{3.120}$$

One can prove (3.120) using e.g. the fact [BF91] that for every interval $\mathcal{I}$ the energy $\mathbb{A}_{\mathrm{hom}}^{\mathcal{I}}$ is the relaxation of the functional

$$(\boldsymbol{\mu}, \boldsymbol{\nu}) \mapsto \begin{cases} \displaystyle\int_{\mathcal{I}\times\mathbb{T}^d} f_{\mathrm{hom}}\left(\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\mathscr{L}^{d+1}}, \frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}\mathscr{L}^{d+1}}\right) \mathrm{d}\mathscr{L}^{d+1}, & \text{if } (\boldsymbol{\mu}, \boldsymbol{\nu}) \ll \mathrm{d}\mathscr{L}^{d+1}, \\ +\infty, & \text{otherwise}\,, \end{cases}$$

for which (3.120) follows from the convexity and nonnegativity of f_{hom} and the properties of the mollifiers ϕ^k.

We shall then estimate the energy of $(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k)$. From (3.118) and (3.119), using the convexity of f_{hom} and the definition of the map S^η, we obtain

$$\begin{aligned}
&\mathbb{A}_{\text{hom}}^{\mathcal{I}_k}(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k) - (1+2\eta_k)\delta_k \sup_{B^\circ} f_{\text{hom}} \\
&\qquad \leq (1-\delta_k) \int_{\mathcal{I}_k \times \mathbb{T}^d} f_{\text{hom}}\Big(\frac{\mathrm{d}\tilde{\boldsymbol{\mu}}^k}{\mathrm{d}\mathscr{L}^{d+1}}(S_k^{-1}(t), x), \frac{\mathrm{d}\tilde{\boldsymbol{\nu}}^k}{\mathrm{d}\mathscr{L}^{d+1}}(S_k^{-1}(t), x)\Big) \,\mathrm{d}\mathscr{L}^{d+1} \\
&\qquad \leq (1-\delta_k)(1+4\eta_k)\mathbb{A}_{\text{hom}}^{(\alpha_k, T-\alpha_k)}(\tilde{\boldsymbol{\mu}}^k, \tilde{\boldsymbol{\nu}}^k) \leq (1-\delta_k)(1+4\eta_k)\mathbb{A}_{\text{hom}}^{T}(\boldsymbol{\mu}, \boldsymbol{\nu}) \,,
\end{aligned}$$

where in the last inequality we used (3.120). Taking the limsup in $k \to \infty$

$$\limsup_{k \to +\infty} \mathbb{A}_{\text{hom}}^{\mathcal{I}_k}(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k) \leq \mathbb{A}_{\text{hom}}^{T}(\boldsymbol{\mu}, \boldsymbol{\nu}) \tag{3.121}$$

which concludes the proof of (3.114). □

Now we are ready to prove the Γ-limsup inequality (3.44) in Theorem 3.5.1.

Proof of Theorem 3.5.1 (limsup inequality). Fix $\boldsymbol{\mu} \in \mathcal{M}_+\big(\mathcal{I} \times \mathbb{T}^d\big)$. By definition of $\mathbb{A}_{\text{hom}}^{\mathcal{I}}(\boldsymbol{\mu})$, it suffices to prove that for every $\boldsymbol{\nu} \in \mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ such that $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}^T$ and $\mathbb{A}_{\text{hom}}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) < +\infty$, we can find $\mathbf{m}^\varepsilon : \overline{\mathcal{I}} \to \mathbb{R}_+^{\mathcal{X}_\varepsilon}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ and $\limsup_\varepsilon \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \leq \mathbb{A}_{\text{hom}}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu})$.

For any such $(\boldsymbol{\mu}, \boldsymbol{\nu})$, we apply Proposition 3.8.26 and find a smooth sequence $(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k)_k \in \mathbb{CE}^{\mathcal{I}(k)}$ where $\mathcal{I}(k) = (-\eta_k, T + \eta_k)$, where $\eta_k \to 0$ and such that (3.114) and (3.115) hold with $(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k) \to (\boldsymbol{\mu}, \boldsymbol{\nu})$ weakly in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d) \times \mathcal{M}^d(\mathcal{I} \times \mathbb{T}^d)$ as $k \to +\infty$. In particular

$$\sup_{k \in \mathbb{N}} \sup_{t \in \mathcal{I}} \mu_t^k(\mathbb{T}^d) = \sup_{k \in \mathbb{N}} \mu_0^k(\mathbb{T}^d) < \infty \,. \tag{3.122}$$

Hence we can apply Proposition 3.8.23 and find $\mathbf{m}^{\varepsilon,k} \in \mathcal{M}_+(\overline{\mathcal{I}} \times \mathbb{T}^d)$ such that $\iota_\varepsilon \mathbf{m}^{\varepsilon,k} \to \boldsymbol{\mu}^k$ weakly in $\mathcal{M}_+(\overline{\mathcal{I}} \times \mathbb{T}^d)$ and for each $k \in \mathbb{N}$,

$$\limsup_{\varepsilon \to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^{\varepsilon,k}) \leq \mathbb{A}_{\text{hom}}^{\mathcal{I}(k)}(\boldsymbol{\mu}^k, \boldsymbol{\nu}^k) + C\eta_k |\mathcal{I}| \Big(\mu_0^k(\mathbb{T}^d) + 1 \Big). \tag{3.123}$$

We conclude by extracting a subsequence $\mathbf{m}^\varepsilon := \mathbf{m}^{\varepsilon, k(\varepsilon)}$ such that $\iota_\varepsilon \mathbf{m}^\varepsilon \to \boldsymbol{\mu}$ weakly in $\mathcal{M}_+(\mathcal{I} \times \mathbb{T}^d)$ as $\varepsilon \to 0$ and from (3.122), (3.123), (3.114) we have

$$\limsup_{\varepsilon \to 0} \mathcal{A}_\varepsilon^{\mathcal{I}}(\mathbf{m}^\varepsilon) \leq \mathbb{A}_{\text{hom}}^{\mathcal{I}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \,,$$

which concludes the proof. □

3.9 Analysis of the cell-problem

In the last section of this work, we discuss some properties of the limit functional $\mathbb{A}_{\text{hom}}$ and analyse examples where explicit computations can be performed. For $\rho \in \mathbb{R}_+$ and $j \in \mathbb{R}^d$, recall that

$$f_{\text{hom}}(\rho, j) := \inf \Big\{ F(m, J) \; : \; (m, J) \in \mathsf{Rep}(\rho, j) \Big\} \,,$$

where $\mathsf{Rep}(\rho, j)$ denotes the set of representatives of (ρ, j), i.e., all $\mathbb{Z}^d$-periodic functions $m \in \mathbb{R}_+^{\mathcal{X}}$ and $J \in \mathbb{R}_a^{\mathcal{E}}$ satisfying

$$\sum_{x \in \mathcal{X}^Q} m(x) = \rho, \quad \mathsf{Eff}(J) = \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} J(x,y)(y_{\mathsf{z}} - x_{\mathsf{z}}) = j, \quad \text{and} \quad \operatorname{div} J \equiv 0.$$

Invariance by rescaling. We start with an invariance property of the cell-problem. Fix a $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$ as defined in Assumption 3.2.1. For every $\varepsilon > 0$, we consider the corresponding rescaled graph $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$. Using the fact that $\mathbb{T}_\varepsilon^d \subset \mathbb{T}^d$ and keeping in mind the considerations in Remark 3.2.2, the rescaled graph corresponds to a $\mathbb{Z}^d$-periodic graph $(\tilde{\mathcal{X}}, \tilde{\mathcal{E}})$, where the corresponding $\tilde{\mathbb{V}}$ is identified with the points of $\mathbb{T}_\varepsilon^d$.

We are thus considering the rescaled graph as a new initial graph, with a new cost function $\tilde{F} := \mathcal{F}_\varepsilon$. Denote by $\tilde{f}_{\hom}$ the corresponding limit density. In view of our convergence result, we then expect the corresponding cell-formula to be *invariant*, namely $f_{\hom} = \tilde{f}_{\hom}$. This is indeed the case, as we are going to see, and it is a consequence of the convexity of F.

One inequality follows from the natural inclusion of representatives

$$\mathsf{Rep}(\rho) \hookrightarrow \varepsilon^d \widetilde{\mathsf{Rep}}(\rho), \quad \mathsf{Rep}(j) \hookrightarrow \varepsilon^{d-1} \widetilde{\mathsf{Rep}}(j)\,, \tag{3.124}$$

which is obtained as $\tilde{m} := \varepsilon^d (\tau_\varepsilon^0)^{-1}(m)$ and $\tilde{J} := \varepsilon^{d-1} (\tau_\varepsilon^0)^{-1}(J)$, for every $(m, J) \in \mathsf{Rep}(\rho, j)$ (note that the inverse of τ_ε^0 is well-defined on $\mathbb{Z}^d$-periodic maps). In particular we have

$$\sum_{x \in \tilde{\mathcal{X}}_Q} \tilde{m}(x) = \sum_{x \in \mathcal{X}_Q} m(x) = \rho, \quad \mathsf{Eff}(\tilde{J}) = \mathsf{Eff}(J), \quad \tilde{F}(\tilde{m}, \tilde{J}) = F(m, J)\,,$$

which implies $f_{\hom} \geq \tilde{f}_{\hom}$.

The opposite inequality is where the convexity of F comes into play. Pick any $(\tilde{m}, \tilde{J}) \in \widetilde{\mathsf{Rep}}(\rho, j)$. A first attempt to define a couple in $\mathsf{Rep}(\rho, j)$ would be to consider the inverse map of what we did in (3.124), but this would not give us $\mathbb{Z}^d$-periodic maps (but only $\frac{1}{\varepsilon}\mathbb{Z}^d$-periodic). What we can do is to consider a convex combination of such values. Precisely, we define

$$m(x) := \varepsilon^d \sum_{z \in \mathbb{Z}_\varepsilon^d} \frac{\tau_\varepsilon^z \tilde{m}(x)}{\varepsilon^d}, \quad J(x,y) := \varepsilon^d \sum_{z \in \mathbb{Z}_\varepsilon^d} \frac{\tau_\varepsilon^z \tilde{J}(x,y)}{\varepsilon^{d-1}}, \quad \forall (x,y) \in \mathcal{X}_Q\,.$$

It is not difficult to see that $(m, J) \in \mathsf{Rep}(\rho, j)$ (by linearity of the constraints) and using the convexity of F we obtain

$$F(m, J) = F\left(\varepsilon^d \sum_{z \in \mathbb{Z}_\varepsilon^d} \left(\frac{\tau_\varepsilon^z \tilde{m}}{\varepsilon^d}, \frac{\tau_\varepsilon^z \tilde{J}}{\varepsilon^{d-1}} \right)\right) \leq \sum_{z \in \mathbb{Z}_\varepsilon^d} \varepsilon^d F\left(\frac{\tau_\varepsilon^z m}{\varepsilon^d}, \frac{\tau_\varepsilon^z J}{\varepsilon^{d-1}} \right) = \tilde{F}(\tilde{m}, \tilde{J})\,,$$

which in particular proves $f_{\hom} \leq \tilde{f}_{\hom}$.

The simplest case: $V = \{v\}$ and nearest-neighbor interaction. The easiest example we can consider is the one where the set V consists of only one element $v \in V$. In other words, we focus on the case when $\mathcal{X} \simeq \mathbb{Z}^d$ and thus $\mathcal{X}_\varepsilon \simeq \mathbb{Z}_\varepsilon^d$. We also consider the graph structure defined via the *nearest-neighbor interaction*, meaning that $\mathcal{E}$ consists of the elements of $(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d$ such that $|x-y|_\infty = 1$.

In this setting, we can identify $\mathcal{X}_Q \simeq V$ (in particular it consists of only one element) and $\mathcal{E}_Q \simeq \{(v, v \pm e_i) \, : \, i = 1, \dots, d\}$ of cardinality $2d$. In particular, for every $\rho \in \mathbb{R}_+$, $j \in \mathbb{R}^d$, the set Rep(ρ, j) consists of only one element $(\underline{m}, \underline{J})$ given by

$$\underline{m}(x) = \rho, \quad \underline{J}(v, v \pm e_i) = \pm j_i, \quad \forall (x, y) \in \mathcal{E}, \; i = 1, \dots, d \, .$$

Consequently, the limit problem is explicitly computable as $f_{\text{hom}}(\rho, j) = F(\underline{m}, \underline{J})$.

For example, if F is edge-based (Remark 3.2.5) with edge-energies $\{F_{xy}\}$, then we obtain

$$f_{\text{hom}}(\rho, j) = 2 \sum_{i=1}^{d} F_{xy}(\rho, \rho, j_i) \, , \quad \forall \rho \in \mathbb{R}_+, \; j \in \mathbb{R}^d \, .$$

In the even more special case of the discretised p-Wasserstein distances as described in (3.16), using the properties of the mean Λ, we end up with

$$f_{\text{hom}}(\rho, j) = \frac{1}{p} \frac{|j|_p^p}{\rho^{p-1}} \, , \quad \forall \rho \in \mathbb{R}_+, \, j \in \mathbb{R}^d \, ,$$

which correspond to the p-Wasserstein distance with underlined metric given by the ℓ_p-distance $| \cdot |_p$. The case $p = 2$ corresponds to the framework studied in [GM13].

As we are going to discuss in the last section, this is just a special situation of a more general framework, which is the one of isotropic finite-volume partition of $\mathbb{T}^d$.

Finite-volume partitions of $\mathbb{T}^d$. The next class of examples are the graph structures associated with $\mathbb{Z}^d$-periodic *finite volume partitions* (FVPs) $\tilde{\mathcal{T}}$ of $\mathbb{R}^d$.

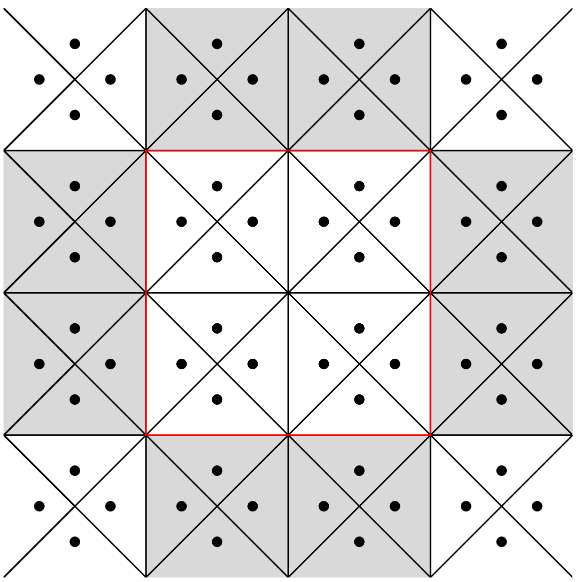

Figure 3.6: A $\mathbb{Z}^2$-periodic finite volume partition of $\mathbb{R}^2$. In red, the unitary cube $[0, 1]^2 \subset \mathbb{R}^2$.

Definition 3.9.1 (Finite-volume on $\mathbb{T}^d$)**.** A $\mathbb{Z}^d$-periodic finite-volume partitions $\tilde{\mathcal{T}}$ of $\mathbb{R}^d$ is a locally finite family of points $x \in \tilde{\mathcal{X}}$ and convex sets $K_x \subset \mathbb{R}^d$

$$\tilde{\mathcal{T}} := \Big\{(x, K_x) \ : \ \dot{\bigcup_{x\in\mathcal{X}}} K_x = \mathbb{R}^d, \ x \in K_x\Big\}$$

which satisfies the following properties:

(i) $\tilde{\mathcal{T}}$ is $\mathbb{Z}^d$-periodic, i.e. $(x+z, K_x+z) \in \tilde{\mathcal{T}}$, for every $(x, K_x) \in \mathcal{T}$, $z \in \mathbb{Z}^d$.

(ii) $\tilde{\mathcal{T}}$ is *admissible*, i.e. $y - x \perp \partial K_x \cap \partial K_y$ for any neighbouring cells, thus satisfying $\mathscr{H}^{d-1}(\partial K_x \cap \partial K_y) \neq 0$.

A finite-volume partition $\mathcal{T}$ on $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is obtained from any $\mathbb{Z}^d$-periodic FVP of $\mathbb{R}^d$ $\tilde{\mathcal{T}}$ by taking the quotient by the action of $\mathbb{Z}^d$. We set $\mathcal{X} := \tilde{\mathcal{X}}/\mathbb{Z}^d \subset \mathbb{T}^d$.

For any $\mathbb{Z}^d$-periodic finite volume partition on $\mathbb{R}^d$, we can associate the embedded $\mathbb{Z}^d$-periodic graph $(\mathcal{X}, \mathcal{E})$ where the edges $\mathcal{E}$ are given by every (x, y) such that $\mathscr{H}^{d-1}(\partial K_x \cap \partial K_y) \neq 0$. In this case we write the usual notation $y \sim x$.

Throughout the whole section, we use the notation

$$\pi(x) := \mathscr{L}^d(K_x), \quad d_{xy} := |y - x|, \quad \tau_{xy} := \frac{y - x}{d_{xy}} \in \mathcal{S}^{d-1},$$
$$|(x|y)| := \mathscr{H}^{d-1}(\partial K_x \cap \partial K_y), \quad \omega_{xy} := \frac{|(x|y)|}{d_{xy}},$$

for $x, y \in \mathcal{X}$.

Geometric expression for the effective flux in embedded graphs

If $(\mathcal{X}, \mathcal{E})$ is an embedded $\mathbb{Z}^d$-periodic graph in $\mathbb{R}^d$ in the sense of Remark 3.2.2, it is possible to give an equivalent *geometric* definition of the effective flux. We thus consider the situation where V is a subset of $[0,1)^d$ and use the identification $(z, v) \equiv z + v$, so that $\mathcal{X}$ can be identified with a $\mathbb{Z}^d$-periodic subset of $\mathbb{R}^d$. Let us define

$$\mathsf{Eff}_{\mathrm{geo}}(J) := \frac{1}{2} \sum_{(x,y)\in\mathcal{E}^Q} J(x,y)\Big(y - x\Big).$$

Note that we simply replaced $y_{\mathsf{z}} - x_{\mathsf{z}}$ by $y - x$ in the definition of $\mathsf{Eff}(J)$. Remarkably, the following result shows that $\mathsf{Eff}_{\mathrm{geo}}(J) = \mathsf{Eff}(J)$ for any periodic divergence-free vector field J. In particular, $\mathsf{Eff}_{\mathrm{geo}}(J)$ does *not* depend on the choice of the embedding into $\mathbb{R}^d$. As a consequence, one can equivalently define $\mathsf{Rep}(j)$, and hence the renormalised energy density $f_{\mathrm{hom}}(\rho, j)$, in terms of $\mathsf{Eff}_{\mathrm{geo}}(J)$ instead of $\mathsf{Eff}(J)$.

Proposition 3.9.2. *For every periodic and divergence-free vector field $J \in \mathbb{R}^{\mathcal{E}}_a$ we have* $\mathsf{Eff}(J) = \mathsf{Eff}_{\mathrm{geo}}(J)$.

Proof. Without loss of generality we assume that $\mathcal{X}^Q \subseteq (0,1)^d$. The general case then follows by continuity.

Fix a vertex $x_0 \in \mathcal{X}^Q$. For $t > 0$ sufficiently small and $v \in \mathbb{R}^d$, consider the modified embedded $\mathbb{Z}^d$-periodic graph $(\mathcal{X}(t), \mathcal{E}(t))$ in $\mathbb{R}^d$ obtained from $\mathcal{X}$ by *shifting* the nodes $x_0 + \mathbb{Z}^d$ by $tv \in \mathbb{R}^d$, i.e., we consider the shifted nodes $x_0(t) + \mathbb{Z}^d := x_0 + tv + \mathbb{Z}^d$ (and with them, the associated edges).

Let $J \in \mathbb{R}_a^{\mathcal{E}} \simeq \mathbb{R}_a^{\mathcal{E}(t)}$ be a $\mathbb{Z}^d$-periodic divergence-free discrete vector field and consider the corresponding effective fluxes

$$\mathsf{Eff}_{\mathrm{geo}}(t, J) := \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q(t)} J(x,y)\big(y - x\big)$$

for $t > 0$ small enough. As J is periodic and divergence-free, we have

$$\begin{aligned} 2\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \mathsf{Eff}_{\mathrm{geo}}(t, J) &= \sum_{x:(x,x_0)\in\mathcal{E}^Q} J(x, x_0)v - \sum_{y:(x_0,y)\in\mathcal{E}^Q} J(x_0, y)v \\ &= -\sum_{y\sim x_0} J(x_0, y)v + \sum_{z\in\mathbb{Z}^d} \sum_{y\in\mathcal{X}^Q,\, y\sim x_0+z} J(y, x_0 + z)v \\ &\overset{J \text{ per.}}{=} -\operatorname{div} J(x_0)v + \sum_{z\in\mathbb{Z}^d} \sum_{y\in\mathcal{X}^Q,\, y-z\sim x_0} J(y - z, x_0)v \\ &= -\operatorname{div} J(x_0)v + \sum_{y'\sim x_0} J(y', x_0)v = -2\operatorname{div} J(x_0)v = 0, \end{aligned}$$

hence $t \mapsto \mathsf{Eff}_{\mathrm{geo}}(t, J)$ is constant. It readily follows that $\mathsf{Eff}_{\mathrm{geo}}(J)$ does not depend on the position of the embedded vertices. We also obtain the sought equality $\mathsf{Eff}(J) = \mathsf{Eff}_{\mathrm{geo}}(J)$, as $\mathsf{Eff}(J)$ corresponds to the limiting case where all the elements of V "collapse" into a single point of $[0, 1)^d$. □

A natural class of energies to consider are the edge-based ones given by

$$F_f(m, J) := \frac{1}{2} \sum_{(x,y)\in\mathcal{X}_Q} \frac{1}{\omega_{xy}} f(\hat{r}_{xy}, J(x, y)), \quad \text{where} \quad \hat{r}_{xy} := \Lambda\left(\frac{m(x)}{\pi(x)}, \frac{m(y)}{\pi(y)}\right), \tag{3.125}$$

where $\{\Lambda_{xy}\}_{xy}$ is a family of admissible means (in the sense of [GKM20, Definition 2.2]) and $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is a convex, lower-semicontinuous function such that F_f satisfies Assumption 3.2.3.

Example 3.9.3. In many interesting examples, the energy f is chosen of the particular form

$$f_\psi^{\mathfrak{m}}(r, J) := \mathfrak{m}(r)\psi\left(\frac{2J}{\mathfrak{m}(r)}\right), \quad \begin{cases} \psi : \mathbb{R} \to [0, \infty], & \psi \text{ convex}, \ \psi(0) = 0, \\ \mathfrak{m} : \mathbb{R}_+ \to [0, +\infty), & \mathfrak{m} \text{ concave, non-decreasing}. \end{cases} \tag{3.126}$$

The discretised p-Wasserstein distances (3.16) are a special, p-homogeneous subcase of this. An interesting, different choice for ψ would be to consider $\psi(J) = \cosh(J) - 1$, which has been studied in [MPR14] in connection to the theory of Large Deviations in the setting of discrete Markov chains.

The framework of finite-volume partitions of euclidean domains have been extensively studied in [GKM20], in the special case of discretisation of the 2-Wasserstein distance on convex and bounded domains of $\mathbb{R}^d$.

It has been showed in [GKM20] that the limit of the discrete distances $\mathcal{W}_\varepsilon$ (in the Gromow-Hausdorff sense) as $\varepsilon \to 0$ coincides with the 2-Wasserstein distance $\mathbb{W}_2$ on $\mathcal{P}(\mathbb{T}^d)$ if and only if an asymptotic local isotropy condition is satisfied.

The goal of this section is to discuss the role of isotropy in the periodic setting. In this framework, we have an easy equivalent way to formulate such a condition.

Definition 3.9.4. (Isotropic meshes) Given a $\mathbb{Z}^d$-periodic finite volume partition $\mathcal{T}$ on $\mathbb{R}^d$, we say the isotropy condition holds with parameters $\{\lambda_{xy}\}_{x,y\in\mathcal{X}}$ whenever

$$\sum_{y\sim x} \lambda_{xy} d_{xy} |(x|y)| \tau_{xy} \otimes \tau_{xy} = \pi(x) \mathsf{Id}, \quad \forall x \in \mathcal{X} \tag{3.127}$$

where $\lambda_{xy} + \lambda_{yx} = 1$ for every $(x, y) \in \mathcal{E}$.

It is possible to show that the previous definition is equivalent, in the periodic setting, to the asymptotic condition introduced in [GKM20], see [GKMP20] for a proof of this fact in the one-dimensional case.

Let us start with the simpler one-dimensional picture, that is $d = 1$.

The one-dimensional case: We consider the graph structure $(\mathcal{X}, \mathcal{E})$ associated with a finite-volume partition of the circle $\mathbb{T}^1 = \mathcal{S}^1$ with energy of the form F_f, as introduced in (3.125). In the special case of $d = 1$, we can identify $\mathcal{E}_Q$ with a set of indexes $i = 1, \dots, M$, where M is the cardinality of $\mathcal{X}_Q$. We can write the energy in the form

$$F_f(m, J) = \frac{1}{2} \sum_{i=1}^{M} d_i \Big(f(\widehat{r}_i, J_{i,i+1}) + f(\widehat{r}_i, -J_{i,i+1}) \Big),$$

where $r_i \pi_i = m_i$ and $\widehat{r}_i = \Lambda(r_i, r_{i+1})$. In the one-dimensional case, the cell-formula drastically simplifies. In particular, for every $j \in \mathbb{R}$, the set $\mathsf{Eff}(j)$ only consists of constant vector fields, i.e. satisfying $J_{i,i+1} = -J_{i+1,i} = j$, for every $i = 1, \dots, M$. Hence the limit problem can be equivalently recast as

$$f_{\hom}(\rho, j) = \inf \left\{ \frac{1}{2} \sum_{i=1}^{M} d_i \Big(f(\widehat{r}_i, j) + f(\widehat{r}_i, -j) \Big) \ : \ \sum_{i=1}^{M} r_i \pi_i = \rho\,, \ r_{M+1} := r_1 \right\}.$$

This problem can be explicitly solved under some additional geometric assumptions.

Definition 3.9.5. A family of means $\{\Lambda_{xy}\}_{xy}$ are adapted to $\{\lambda_{xy}\}_{xy} \subset [0, 1]$ when

$$\Lambda_{xy}(a, b) = \Lambda_{yx}(b, a), \quad \Lambda_{xy}(a, b) \le \lambda_{xy} a + \lambda_{yx} b$$

for any $(x, y) \in \mathcal{E}$, $a, b \in \mathbb{R}_+$. Moreover we assume $\lambda_{xy} = 1 - \lambda_{yx} \in [0, 1]$.

Remark 3.9.6. Each continuously differentiable mean Λ is λ-balanced for exactly one value of $\lambda \in [0, 1]$, namely

$$\lambda = \partial_1 \Lambda(1, 1) . \tag{3.128}$$

We claim that under the additional assumption that $\{\Lambda_{xy}\}_{xy}$ are adapted to $\{\lambda_{xy}\}_{xy}$ and the mesh is isotropic with same parameters (in the sense of Definition 3.9.4), then we have

$$f_{\hom}(\rho, j) = \frac{1}{2} \Big(f(\rho, j) + f(\rho, -j) \Big), \quad \forall \rho, j \in \mathbb{R}_+ \times \mathbb{R},$$

for every f which is not increasing in the first variable.

The key observation is that $\sum_{i=1}^{M} d_i = |\mathcal{S}^1| = 1$. Using the trivial competitor $r_i \equiv \rho$, we obtain the upper bound (note that no isotropy is needed for this step). To obtain the lower bound, we first observe that, by adaptedness of Λ and the isotropy[4] of the mesh, we have:

$$\sum_{i=1}^{M} d_i \widehat{r}_i \leq \sum_{i=1}^{M} d_i \Big(\lambda_{i,i+1} r_i + \lambda_{i+1,i} r_{i+1} \Big) = \sum_{i=1}^{M} r_i \Big(d_i \lambda_{i,i+1} + d_{i-1} \lambda_{i,i-1} \Big) = \sum_{i=1}^{M} r_i \pi_i = \rho \,,$$

for every competitor $m = \pi r \in \mathsf{Rep}(\rho)$. We can then apply this bound, together with the convexity of f and its monotonicity to show that

$$\sum_{i=1}^{M} d_i \Big(f(\widehat{r}_i, j) + f(\widehat{r}_i, -j) \Big) \geq f\Big(\sum_{i=1}^{M} d_i \widehat{r}_i, j \Big) + f\Big(\sum_{i=1}^{M} d_i \widehat{r}_i, -j \Big) \geq \Big(f(\rho, j) + f(\rho, -j) \Big) \,,$$

for every competitor $m = \pi r \in \mathsf{Rep}(\rho)$, which concludes the proof of the claim.

Arbitary dimension, quadratic energies. Finally, we discuss a particular class of energies associated to finite-volume partitions in arbitrary dimension. We focus on quadratic energies, where f in (3.125) is of the form

$$f^{\mathfrak{m}}(\rho, J) = \begin{cases} \frac{|J|^2}{\mathfrak{m}(\rho)} & \text{if } \rho > 0, \\ 0 & \text{if } \rho = 0 = J, \\ +\infty & \text{otherwise} \end{cases} \tag{3.129}$$

where $\mathfrak{m} : \mathbb{R}_+ \to \mathbb{R}_+$ is a concave, non-decreasing mobility function. They represent the discrete counterparts of generalised Wasserstein distances, as introduced in [DNS09], given by

$$\mathbb{W}_{\mathfrak{m}}^2(\mu_0, \mu_1) := \inf \left\{ \mathbb{A}^{\mathfrak{m}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \ : \ (\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}(\mu_0, \mu_1) \right\} = \mathbb{MA}^{\mathfrak{m}}(\mu_0, \mu_1) \,,$$

where $\mathbb{A}^{\mathfrak{m}}$ is the continuous energy functional associated with $f^{\mathfrak{m}}$, as in Definition 3.3.10.

The linear mobility case $\mathfrak{m}(\rho) = \rho$ corresponds to the study of the discrete energies as introduced in [GKM20], where the authors deal with the limit behaviour as $\varepsilon \to 0$ of the associated Riemannian distances on $\mathcal{P}(\mathcal{X}_\varepsilon)$ defined via the minimisation

$$\mathcal{W}_\varepsilon(m^0, m^1) := \inf \Big\{ \mathcal{A}_\varepsilon^{\mathfrak{m}}(\mathbf{m}, \mathbf{J}) \ : \ (\mathbf{m}, \mathbf{J}) \in \mathcal{CE}_\varepsilon(m_0, m_1) \Big\} = \mathcal{MA}_\varepsilon^{\mathfrak{m}}(m_0, m_1) \,,$$

although the authors do not work in a periodic setting, but rather with general convex and bounded domains of $\mathbb{R}^d$.

In this final section of our work, we apply the homogenisation result Theorem 3.5.1 and show that the isotropy condition (3.9.4) is equivalent to the fact that the discrete energies $\mathcal{A}_\varepsilon^{\mathfrak{m}}$ converge to the continuous ones $\mathbb{A}^{\mathfrak{m}}$, namely that $f_{\hom}^{\mathfrak{m}}(\rho, j) = f^{\mathfrak{m}}(\rho, |j|)$ (coherently with what shown in [GKM20] in the linear mobility case). We start by showing that, regardless of any additional condition, the limit density is always smaller than $f^{\mathfrak{m}}$.

Lemma 3.9.7. *Consider the periodic graph structure of $(\mathcal{X}_\varepsilon, \mathcal{E}_\varepsilon)$ induced by a finite volume partition $\mathcal{T}$ and let $\mathcal{A}_\varepsilon^{\mathfrak{m}}$ be the quadratic energies associated to $f^{\mathfrak{m}}$ as defined in (3.129). Let $f_{\hom}^{\mathfrak{m}}$ be corresponding limit density as $\varepsilon \to 0$ as given by Theorem 3.5.1. Then we have*

$$f_{\hom}^{\mathfrak{m}}(r, j) \leq f^{\mathfrak{m}}(r, |j|), \quad \forall r > 0, \ j \in \mathbb{R}^d.$$

[4] In $d = 1$, it means $\pi_i = d_i \lambda_{i,i+1} + d_{i-1} \lambda_{i,i-1}$, for every $i = 1, \ldots, M$, see [GKMP20, Definition 4.3].

Proof. Pick any $j \in \mathbb{R}^d$ and define the discrete vector field J^* as

$$J^*(x,y) := \langle j, \tau_{xy} |(x|y)| \rangle, \quad x, y \in \mathcal{X}. \tag{3.130}$$

We claim that such a vector field is a competitor for the cell problem $f_{\text{hom}}^{\mathfrak{m}}(r,j)$. Indeed, for any $x \in \mathcal{X}$

$$\sum_{y \sim x} J^*(x,y) = \left\langle j, \sum_{y \sim x} \tau_{xy} |(x|y)| \right\rangle = \left\langle j, \int_{\partial K_x} \nu_{\text{ext}} \, \mathrm{d}\mathscr{H}^{d-1} \right\rangle = 0 \tag{3.131}$$

where in the last step we used Stokes' theorem. This shows that J^* is divergence-free. We now compute the effective flux of J^* is given by j. By definition of J^*

$$\mathsf{Eff}(J^*) = \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} J^*(x,y)(y-x) = \left(\frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \tau_{xy} \otimes \tau_{xy} \right) j = j \,, \tag{3.132}$$

where we used the fact that

$$\frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \tau_{xy} \otimes \tau_{xy} = \mathsf{Id}. \tag{3.133}$$

This can be proved, in the periodic setting, in the same way as in [GKM20, Lemma 5.4].

Finally, we prove the theorem choosing as competitors

$$m(x) = m^*(x) := r\pi_x, \quad J(x,y) = J^*(x,y). \tag{3.134}$$

A direct computation shows

$$\begin{aligned}
2F_{f^{\mathfrak{m}}}(m^*, J^*) &= \frac{1}{\mathfrak{m}(r)} \sum_{(x,y) \in \mathcal{E}^Q} \frac{d_{xy} |J^*(x,y)|^2}{|(x|y)|} \\
&= \frac{1}{\mathfrak{m}(r)} \sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 \\
&= \frac{1}{\mathfrak{m}(r)} \left\langle \left(\sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \tau_{xy} \otimes \tau_{xy} \right) j, j \right\rangle \\
&= \frac{2|j|^2}{\mathfrak{m}(r)} \,,
\end{aligned}$$

where we used (3.133) once again. This ends the proof. □

Next, we show that (m^*, J^*) as defined in (3.134) is an optimal competitor if and only if the isotropy condition holds, in particular condition (3.9.4) is equivalent to $\mathbb{A}_{\text{hom}}^{\mathfrak{m}} = \mathbb{A}^{\mathfrak{m}}$.

Theorem 3.9.8 (Isotropy is equivalent to $f_{\text{hom}}^{\mathfrak{m}} = f^{\mathfrak{m}}$)**.** *Let $\{\Lambda_{xy}\}_{xy}$ be a family of means that are adapted to $\{\lambda_{xy}\}_{xy}$. Consider $f_{\text{hom}}^{\mathfrak{m}}$, $f^{\mathfrak{m}}$ as before.*

1. *If $\mathcal{T}$ satisfies the isotropy condition with parameters $\{\lambda_{xy}\}$, then $f_{\text{hom}}^{\mathfrak{m}} = f^{\mathfrak{m}}$.*
2. *Assume that each mean Λ_{xy} and $\mathfrak{m}$ are continuously differentiable. If $f_{\text{hom}}^{\mathfrak{m}} = f^{\mathfrak{m}}$, then $\mathcal{T}$ satisfies the isotropy condition with parameters $\{\lambda_{xy}\}$.*

Proof of (1). By homogeneity, it is enough to consider $j \in \mathcal{S}^{d-1}$, and to show that (m^*, J^*) as defined in (3.134) is a minimizer for the cell problem $f_{\text{hom}}^{\mathfrak{m}}(r, j)$. We introduce the constraint functions $g_1 : \mathbb{R}^{\mathcal{X}} \to \mathbb{R}_+$, $g_2 : \mathbb{R}^{\mathcal{E}} \to \mathbb{R}^{\mathcal{X}} \times \mathbb{R}^d$ given by

$$g_1(m) = \sum_{x \in \mathcal{X}} m(x), \quad g_2(J) = \Big(\operatorname{div} J, \operatorname{Eff}(J) \Big).$$

Using these notations, the cell problem reads

$$f_{\text{hom}}^{\mathfrak{m}}(r, j) = \inf_{(m,J)} \Big\{ F^{\mathfrak{m}}(m, J) \ : \ g_1(m) = r, \quad g_2(J) = (0, j) \Big\} ,$$

for $F^{\mathfrak{m}} := F_{f^{\mathfrak{m}}}$ as defined in (3.125). Note that the minimisation is described by two linear constraints in m and J which depend on just one of the two variables. Therefore by convexity of $F^{\mathfrak{m}}$, to show (1) it is enough to prove stationarity of (m^*, J^*) along the two corresponding affine subspaces.

Step 1: let m be such that $g_1(m) = r$. Note that (3.133) implies

$$\frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 = 1, \quad j \in \mathcal{S}^{d-1} , \tag{3.135}$$

whereas the very definition of J^* (3.130) shows that

$$|J^*(x, y)| = |(x|y)| \, |\langle j, \tau_{xy} \rangle| , \quad \forall x, y \in \mathcal{T} .$$

These equations, together with the weighted arithmetic-harmonic mean inequality, the adaptedness, the monotonicity of $\mathfrak{m}$, and the isotropy condition, gives us the lower bound

$$\begin{aligned}
F^{\mathfrak{m}}(m, J^*) &= \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 \, \frac{1}{\mathfrak{m} \circ \Lambda_{xy}(r(x), r(y))} \\
&\geq \left(\sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 \, \mathfrak{m} \circ \Lambda_{xy}(r(x), r(y)) \right)^{-1} \\
&\geq \left(\sum_{(x,y) \in \mathcal{E}^Q} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 \, \mathfrak{m} \left(\lambda_{xy} r(x) + \lambda_{yx} r(y) \right) \right)^{-1} \\
&\geq \left[\mathfrak{m} \left(\sum_{x \in \mathcal{X}} r(x) \sum_{y \sim x} \lambda_{xy} d_{xy} |(x|y)| \, |\langle j, \tau_{xy} \rangle|^2 \right) \right]^{-1} \\
&= \left[\mathfrak{m} \left(\sum_{x \in \mathcal{X}} r(x) \pi(x) \right) \right]^{-1} = \frac{1}{\mathfrak{m}(r)} = F^{\mathfrak{m}}(m^*, J^*) ,
\end{aligned}$$

where we used the notation $m(x) = r(x)\pi(x)$, so that r denotes the density of m with respect to π. In the last inequality we also used (3.135) and the concavity of $\mathfrak{m}$.

Step 2: consider the following minimization problem

$$\min_J \left\{ F^{\mathfrak{m}}(m^*, J) = \frac{1}{2} \sum_{(x,y) \in \mathcal{E}^Q} \mu_{xy} |J(x, y)|^2 \ : \ g_2(J) = (0, j) \right\} ,$$

where we set $\mu_{xy} = d_{xy}(\mathfrak{m}(r)|(x|y)|)^{-1}$ and let J_0 be the corresponding minimizer. We claim that $J_0 = J^*$. To prove it, observe that the minimizer J_0 must satisfy

$$\mu_{xy} J_0(x,y) = \lambda_0(x) + \sum_{i=1}^{d} \lambda_i \langle y - x, \partial_i \rangle = \lambda_0(x) + \left\langle y - x, \sum_{i=1}^{d} \lambda_i \partial_i \right\rangle ,$$

for some Lagrangian multipliers $\lambda_0 \in \mathbb{R}^{\mathcal{X}}$, $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$, where ∂_i just denotes the i-th element of the canonical basis of $\mathbb{R}^d$. In particular one can equivalently express the latter equation as follows

$$J_0(x,y) = \frac{\lambda_0(x)}{\mu_{xy}} + \langle \tau_{xy}|(x|y)|, w_0 \rangle , \quad w_0 = \sum_{i=1}^{d} r \lambda_i \partial_i .$$

In other words, J_0 is a perturbation of the divergence-free vector field $\langle \tau_{xy}|(x|y)|, w_0 \rangle$ (basically of J^* where the * is made with respect to w_0) which is divergence-free itself, as shown in (3.131). This forces $\lambda_0 \equiv 0$. Finally, using (3.133) as in (3.132), we see that $g_2(J_0) = (0, j)$, which implies $w_0 = j$, thus $J_0 = J^*$.

We then just proved that $F^{\mathfrak{m}}(m^*, J) \geq F^{\mathfrak{m}}(m^*, J^*)$ for any J such that $g_2(J) = (0, j)$, which ends the proof of (1). □

Proof of (2). By hypothesis, for any $j \in \mathcal{S}^{d-1}$ the optimal competitor for the cell problem is given by (m^*, J^*) as defined in (3.130), (3.134), which means

$$f^{\mathfrak{m}}_{\hom}(r, j) = F^{\mathfrak{m}}(m^*, J^*), \quad \forall j \in \mathcal{S}^{d-1}.$$

For any $x \sim y$, let us consider the following variation of m^*, given by

$$m_\alpha(s) = m^*(s) + \alpha \delta_x(s) - \alpha \delta_y(s), \quad s \in \mathcal{X}.$$

Clearly m_α is an admissible competitor for α small enough and by optimality

$$F^{\mathfrak{m}}(m^*, J^*) \leq F^{\mathfrak{m}}(m_\alpha, J^*), \quad \forall \alpha << 1, \quad \forall x \sim y\,.$$

In particular, we infer

$$\frac{\partial}{\partial \alpha}\bigg|_{\alpha=0} F^{\mathfrak{m}}(m_\alpha, J^*) = 0\,. \tag{3.136}$$

In order to simplify a bit the notation, we set $\epsilon_{xy} := d_{xy}|(x|y)|\,|\langle j, \tau_{xy} \rangle|^2$ so that

$$F^{\mathfrak{m}}(m_\alpha, J^*) =$$

$$\frac{1}{2}\left(\sum_{s \sim x,\, s \neq y} \frac{\epsilon_{xs}}{\mathfrak{m} \circ \Lambda_{xs}\left(r + \frac{\alpha}{\pi_x}, r\right)} + \sum_{s \sim y,\, s \neq x} \frac{\epsilon_{ys}}{\mathfrak{m} \circ \Lambda_{ys}\left(r - \frac{\alpha}{\pi_y}, r\right)} + \frac{\epsilon_{xy}}{\mathfrak{m} \circ \Lambda_{xy}\left(r + \frac{\alpha}{\pi_x}, r - \frac{\alpha}{\pi_y}\right)} \right)$$
$$+ \text{ terms independent of } \alpha.$$

Assuming that Λ_{xs} is smooth, the adaptness and the homogeneity of the means yield

$$\partial_2 \Lambda_{sx}(r,r) = \partial_1 \Lambda_{xs}(r,r) = \partial_1 \Lambda_{xs}(1,1) = \lambda_{xs}, \quad \forall r \geq 0,\ \forall x, s \in \mathcal{X}.$$

A straightforward computation then shows

$$\frac{\partial}{\partial \alpha}\bigg|_{\alpha=0} F^{\mathfrak{m}}(m_\alpha, J^*) = \frac{\mathfrak{m}'(r)}{\mathfrak{m}^2(r)}\Big(A_y - A_x\Big)\,,$$

where

$$A_x := \sum_{s\sim x} \frac{\epsilon_{xs}}{\pi_x} \lambda_{xs} \, .$$

Reasoning as above for every $x \sim y$, we deduce from the optimality conditions (3.136) that there must exist a $\beta \in \mathbb{R}_+$ such that $A_x = \beta$ for all $x \in \mathcal{X}$. In particular, this means that

$$\sum_{s\sim x} \lambda_{xs} d_{xs} |(x|s)| \tau_{xs} \otimes \tau_{xs}(j,j) = \pi_x \beta, \quad \forall x \in \mathcal{X}, \quad \forall j \in \mathcal{S}^{d-1}.$$

We conclude once again thanks to (3.133) which implies $\beta = 1$ (and hence the isotropy condition). $\square$

3.10 Notation

For the convenience of the reader we collect some notation used in this paper.

A°	topological interior of a set A
$\mathsf{D}(F)$	domain of a functional F: $\mathsf{D}(F) = \{x \in \mathcal{X} : F(x) < \infty\}$.
$\mathcal{I}$	bounded open time interval.
$\mathcal{M}^d(A)$	the space of finite $\mathbb{R}^d$-valued Radon measures on A.
$\mathcal{M}_+(A)$	the space of finite (positive) Radon measures on A.
$\mathcal{X}^Q$	the set of all $x \in \mathcal{X}$ with $x_\mathsf{z} = 0$.
$\mathcal{E}^Q$	the set of all $(x, y) \in \mathcal{E}$ with $x_\mathsf{z} = 0$.
$\mathbb{R}^\mathcal{E}_a$	the set of anti-symmetric real functions on $\mathcal{E}$.
$\mathbb{T}^d_\varepsilon$, $\mathbb{Z}^d_\varepsilon$	the discrete torus of mesh size $\varepsilon > 0$: $\mathbb{T}^d_\varepsilon = (\varepsilon\mathbb{Z}/\mathbb{Z})^d = \varepsilon\mathbb{Z}^d_\varepsilon$.
$\mathsf{Eff}(J)$	the effective flux of J: $\mathsf{Eff}(J) = \frac{1}{2}\sum_{(x,y)\in\mathcal{E}^Q} J(x, y)(y_\mathsf{z} - x_\mathsf{z})$.
$\mathsf{Rep}(\rho)$	the set of representatives of $\rho \in \mathbb{R}_+$, i.e, all $m \in \mathbb{R}^\mathcal{X}_+$ s.t. $\sum_{x\in\mathcal{X}^Q} m(x) = \rho$.
$\mathsf{Rep}(j)$	the set of representatives of $j \in \mathbb{R}^d$, i.e, all $J \in \mathbb{R}^\mathcal{E}_a$ divergence-free and s.t. $\frac{1}{2}\sum_{(x,y)\in\mathcal{X}^Q} J(x, y)(y_\mathsf{z} - x_\mathsf{z}) = j$.
$\mathsf{Rep}(\rho, j)$	the set of representatives of $\rho \in \mathbb{R}_+$, $j \in \mathbb{R}^d$: $\mathsf{Rep}(\rho, j) = \mathsf{Rep}(\rho) \times \mathsf{Rep}(j)$.
Q^z_ε	the cube of size $\varepsilon > 0$ centered in $\varepsilon z \in \mathbb{T}^d$: for $z \in \mathbb{Z}^d_\varepsilon$, $Q^z_\varepsilon := [0, \varepsilon)^d + \varepsilon z$.
$S^{\bar z}$	shift operator: $S^{\bar z}_\varepsilon : \mathcal{X} \to \mathcal{X}$, $S^{\bar z}_\varepsilon(x) = (\bar z + z, v)$ for $x = (z, v) \in \mathcal{X}$.
	shift operator: $S^{\bar z}_\varepsilon : \mathcal{E} \to \mathcal{E}$, $S^{\bar z}_\varepsilon(x, y) := \big(S^{\bar z}_\varepsilon(x), S^{\bar z}_\varepsilon(y)\big)$ for $(x, y) \in \mathcal{E}_\varepsilon$
σ^z	$\sigma^{\bar z}_\varepsilon\psi : \mathcal{X}_\varepsilon \to \mathbb{R}$, $(\sigma^{\bar z}_\varepsilon\psi)(x) := \psi(S^{\bar z}_\varepsilon(x))$ for $x \in \mathcal{X}_\varepsilon$.
	$\sigma^{\bar z}_\varepsilon J : \mathcal{E}_\varepsilon \to \mathbb{R}$, $(\sigma^{\bar z}_\varepsilon J)(x, y) := J(S^{\bar z}_\varepsilon(x, y))$ for $(x, y) \in \mathcal{E}_\varepsilon$.
$T^{\bar z}_\varepsilon$	rescaling operator: $T^{\bar z}_\varepsilon : \mathcal{X} \to \mathcal{X}_\varepsilon$: $T^{\bar z}_\varepsilon(x) = (\varepsilon(\bar z + z), v)$ for $x = (z, v) \in \mathcal{X}$.
τ^z_ε	$\tau^{\bar z}_\varepsilon\psi : \mathcal{X} \to \mathbb{R}$, $\big(\tau^{\bar z}_\varepsilon\psi\big)(x) := \psi\big(T^{\bar z}_\varepsilon(x)\big)$ for $x \in \mathcal{X}$.
	$\tau^{\bar z}_\varepsilon J : \mathcal{E} \to \mathbb{R}$, $\big(\tau^{\bar z}_\varepsilon J\big)(x, y) := J\big(T^{\bar z}_\varepsilon(x), T^{\bar z}_\varepsilon(y)\big)$ for $(x, y) \in \mathcal{E}$.
$\mathcal{CE}$	discrete continuity equation: $(\mathbf{m}, \mathbf{J}) \in \mathcal{CE}$ iff $\partial_t m_t + \mathsf{div}\, J = 0$ on $(\mathcal{X}, \mathcal{E})$.
$\mathbb{CE}$	continuous continuity equation: $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{CE}$ iff $\partial_t\mu_t + \nabla \cdot \boldsymbol{\nu} = 0$ on $\mathbb{T}^d$.
BV	more precisely $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$: the space of time-dependent curves of (positive) measures with bounded variation with respect to the KR norm (Kantorovich-Rubenstein) on $\mathcal{M}_+(\mathbb{T}^d)$.
$W^{1.1}$	more precisely $W^{1,1}_{\mathrm{KR}}(\mathcal{I}; \mathcal{M}_+(\mathbb{T}^d))$: the space of time-dependent curves of (positive) measures belonging to the Banach space $W^{1,1}\big(\mathcal{I}; (\mathcal{C}^1(\mathbb{T}^d))^*\big)$.
$\mathsf{P}_\varepsilon\mu, \mathsf{P}_\varepsilon\nu$	discretisation of $\mu \in \mathcal{M}_+(\mathbb{T}^d)$, $\nu \in \mathcal{M}^d(\mathbb{T}^d)$: for $z \in \mathbb{Z}^d_\varepsilon$, $(\mathsf{P}_\varepsilon\mu(z), \mathsf{P}_\varepsilon\nu(z)) \in \mathbb{R}_+ \times \mathbb{R}^d$, given by $\mathsf{P}_\varepsilon\mu(z) = \mu(Q^z_\varepsilon)$, $\mathsf{P}_\varepsilon\nu(z) = \Big((\nu \cdot e_i)(\partial Q^z_\varepsilon \cap \partial Q^{z+e_i}_\varepsilon)\Big)_i$.

In the paper we use some standard terminology from graph theory. Let $(\mathcal{X}, \mathcal{E})$ be a locally finite graph. A *discrete vector field* is an anti-symmetric function $J : \mathcal{E} \to \mathbb{R}$. Its *discrete divergence* is the function $\mathsf{div}\, J : \mathcal{X} \to \mathbb{R}$ defined by

$$\mathsf{div}\, J(x) := \sum_{y\sim x} J(x, y). \tag{3.137}$$

We say that J is *divergence-free* if $\mathsf{div}\, J = 0$.

CHAPTER 4

Evolutionary Γ-convergence of entropic gradient flow structures for Fokker–Planck equations in multiple dimensions

In this chapter we present a discrete-to-continuum convergence result for gradient-flow structures for Fokker–Planck equation in arbitrary dimension. This is the content of the work [FMP20], obtained in collaboration with Dominik Forkert and Jan Maas.

More in detail, we consider finite-volume approximations of Fokker-Planck equations on bounded convex domains in $\mathbb{R}^d$ and study the corresponding gradient flow structures. We reprove the convergence of the discrete to continuous Fokker-Planck equation via the method of Evolutionary Γ-convergence, i.e., we pass to the limit at the level of the gradient flow structures, generalising the one-dimensional result obtained by Disser and Liero. The proof is of variational nature and relies on a Mosco convergence result for functionals in the discrete-to-continuum limit that is of independent interest. Our results apply to arbitrary regular meshes, even though the associated discrete transport distances may fail to converge to the Wasserstein distance in this generality.

4.1 Introduction

This paper deals with the convergence of discrete gradient flow structures arising from finite volume discretisations of Fokker-Planck equations on bounded convex domains $\Omega \subset \mathbb{R}^d$. For a given potential $V \in C^1(\Omega) \cap C(\overline{\Omega})$ we consider the Fokker-Planck equation

$$\partial_t \mu = \Delta \mu + \nabla \cdot (\mu \nabla V) \quad \text{on } (0,T) \times \Omega, \quad \mu|_{t=0} = \mu_0 \tag{4.1}$$

with no-flux boundary conditions, for $T \in (0, +\infty)$. Since the seminal works of Jordan, Kinderlehrer, and Otto [JKO98, Ott01] it is known that (4.1) can be formulated as a gradient flow in the space of probability measures $\mathcal{P}(\overline{\Omega})$ endowed with the 2-Wasserstein distance $\mathbb{W}$ from optimal transport. The driving functional is the relative entropy with respect to the invariant measure $\mathbf{m}(\mathrm{d}x) := \frac{1}{Z_V} \exp(-V(x))\,\mathrm{d}x$, where Z_V is a normalising constant. Here we consider spatial discretisations of (4.1) obtained by finite volume methods for a general class of admissible meshes. In this setting it is very well known that solutions to the discrete equations converge to solutions of (4.1); see, e.g., [EGH00], [BHO18] for results in dimension 2 and 3, and [DEG+18] for more general situations.

The discretised Fokker-Planck equation can also be formulated as gradient flow, with respect to a suitable discrete dynamical transport distance $\mathcal{W}_{\mathcal{T}}$; see the independent works [CHLZ12, Maa11, Mie11]. Here we exploit this gradient flow structure to reprove the convergence of discrete to continuous Fokker-Planck equations via the method of *evolutionary Γ-convergence*; i.e., rather than directly passing to the limit at the level of the gradient flow equation, we pass to the limit in the *energy-dissipation inequality* that characterises the gradient flow structure.

This yields a new proof of convergence for the associated gradient flow equations, which does not rely on specific properties such as linearity or second order. Instead, the method is based on properties of functionals and tools such as Γ- and Mosco convergence.

The method of evolutionary Γ-convergence was pioneered by Sandier and Serfaty [SS04]; see [Mie16b] for a survey on the topic and [MMP21] for important refinements. It has recently been applied to gradient system with a wiggly energy [DFM19, MMP21], coarse graining in linear fast-slow reaction systems [MS19], diffusion in thin structures [FL18], chemical reaction systems [MM20], and various other situations.

For Fokker-Planck equations in dimension $d = 1$, evolutionary Γ-convergence of the discrete gradient flow structures was proved by Disser and Liero [DL15], for a specific class of finite-volume discretisations (cf. Section 4.3.3). Their proof relies on interpolation techniques which do not easily extend to multiple dimensions. Our proof is different and relies on compactness and representation theorems, in particular [BFLM02, Theorem 2], adapting ideas from [AC04]. Our variational proof suggests the possibility of extending these techniques to more general settings, e.g., to higher order and/or nonlinear PDEs.

The fact that the method of evolutionary Γ-convergence of the gradient structures works on arbitrary admissible meshes is remarkable in view of recent work on the discrete-to-continuous limit of the associated transport distances. In fact, it was shown in [GKM20] that the convergence of the discrete transport distances to the Wasserstein distance $\mathbb{W}$ (in the limit of vanishing mesh size) requires a restrictive isotropy condition on the meshes; see [GKMP20] for explicit examples. This discrepancy in convergence behaviour can be explained by a difference in regularity: to prove Γ-convergence of the discrete gradient flow structures one may exploit spatial smoothness assumptions on the discrete dynamics (in view of regularity results for the discrete gradient flows). By contrast, the transport costs on anisotropic meshes are minimised along highly oscillatory curves.

Organisation of the paper. In Section 4.2 we discuss gradient flow structures for continuous and discretised Fokker-Planck equations. Section 4.3 contains the main result of this paper, namely, the evolutionary Γ-convergence of discrete to continuous gradient flow structures (Theorem 4.3.7). This result relies on energy bounds (Theorem 4.3.3) which are proved using Mosco convergence results in the discrete-to-continuum limit that are of independent interest (Theorem 4.3.9). In Section 4.3.3 we discuss related work. Section 4.4 contains the proofs of Theorem 4.3.3 and Theorem 4.3.7. The proof of Theorem 4.3.9 is contained in Sections 4.5, 4.6, and 4.7.

4.2 Finite-volume discretisation of Wasserstein gradient flows

In this section we describe the formulation of the Fokker-Planck equations as gradient flow in the space of probability measures, both at the continuous and at the discrete level. For the sake of clarity, our discussion will be informal. We refer to Section 4.3 below for rigorous statements of the main results.

4.2.1 Fokker-Planck equations as Wasserstein gradient flows

On a bounded convex domain $\Omega \subset \mathbb{R}^d$ we consider the Fokker-Planck equation

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot (\mu_t \nabla V) \tag{4.2}$$

with no-flux boundary conditions. From now on, we assume that $V \in C(\overline{\Omega}) \cap C^2(\Omega)$ is a driving potential with bounded second derivative. This ensures the equivalence between different formulation of the gradient-flow evolutions.

This equation describes the time-evolution of the law of a Brownian particle in a potential field. The steady state is given by the probability measure

$$\mathbf{m} \in \mathcal{P}(\overline{\Omega}) \quad \text{with density} \quad \sigma(x) = \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}x} = \frac{1}{Z_V} e^{-V(x)}, \tag{4.3}$$

where $Z_V \in \mathbb{R}_+$ is a normalising constant.

Since the seminal work of Jordan, Kinderlehrer and Otto [JKO98] it is known that (4.2) is a gradient flow with respect to the Wasserstein distance $\mathbb{W}$ from optimal transport. In its dynamical formulation, $\mathbb{W}$ is given by the *Benamou–Brenier formula*

$$\mathbb{W}(\mu_0, \mu_1)^2 = \inf \left\{ \int_0^1 \int_{\overline{\Omega}} |v_t(x)|^2 \,\mathrm{d}\mu_t(x) \,\mathrm{d}t \right\}, \tag{4.4}$$

where the infimum runs over all curves $(\mu_t)_t$ in the space of probability measures and all vector fields $(v_t)_t$ satisfying the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\mu_t v_t) = 0$$

in the sense of distributions, with boundary conditions $\mu_t|_{t=0} = \mu_0$ and $\mu_t|_{t=1} = \mu_1$. The driving functional in this gradient flow formulation is the relative entropy $\mathbf{H} : \mathcal{P}(\overline{\Omega}) \to [0, +\infty]$ given by

$$\mathbf{H}(\mu) := \begin{cases} \int_{\overline{\Omega}} \rho(x) \log \rho(x) \,\mathrm{d}\mathbf{m} & \text{if } \mathrm{d}\mu = \rho \,\mathrm{d}\mathbf{m}, \\ +\infty & \text{otherwise.} \end{cases}$$

The gradient flow structure can be interpreted at various levels: the original formulation in [JKO98] was given in terms of a time-discrete minimising movement scheme. Another interpretation is in terms of Otto's formal infinite-dimensional Riemannian calculus on the Wasserstein space [Ott01]. Yet another approach relies on the metric formulation of gradient flows in terms of the *energy dissipation inequality* (EDI)

$$\mathbf{H}(\mu_t) + \frac{1}{2} \int_0^T |\dot{\mu}_t|^2_{\mathbb{W}} + |\partial_{\mathbb{W}} \mathbf{H}(\mu_t)|^2 \,\mathrm{d}t \le \mathbf{H}(\mu_0), \tag{4.5}$$

where $|\dot{\mu}_t|$ denotes the $\mathbb{W}$-metric derivative of the curve μ_t and $\partial_{\mathbb{W}}\mathbf{H}(\mu)$ the slope of the relative entropy, namely

$$|\dot{\mu}_t|_{\mathbb{W}} := \lim_{h\to 0} \frac{1}{h}\mathbb{W}(\mu_{t+h}, \mu_t), \qquad |\partial_{\mathbb{W}}\mathbf{H}(\mu)| := \limsup_{\nu\to\mu} \frac{[\mathbf{H}(\mu) - \mathbf{H}(\nu)]_-}{\mathbb{W}(\mu,\nu)},$$

where $[a]_- = \max\{0, -a\}$. Writing $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{m}}$, we have the identity

$$|\partial_{\mathbb{W}}\mathbf{H}(\mu)|^2 = \mathbf{I}(\mu), \quad \text{where} \quad \mathbf{I}(\mu) := \int_\Omega |\nabla \log\rho|^2 \rho \,\mathrm{d}\mathbf{m} = 4\int_\Omega |\nabla\sqrt{\rho}|^2 \,\mathrm{d}\mathbf{m} \tag{4.6}$$

is the *relative Fisher information* with respect to $\mathbf{m}$. The equivalence between the notion of the EDI and the Fokker–Planck solutions is consequence of our regularity assumption on the driving potential, which ensure the λ-convexity of the entropy functionals, see also [AGS08].

$\mathbb{A}$-$\mathbb{A}^*$ formalism of gradient flows

One can recast (4.5) in terms of a suitable weighted Dirichlet energy $\mathbb{A}$ and its Legendre transform $\mathbb{A}^*$. Let us consider the energy functional

$$\mathbb{A}(\mu,\varphi) := \frac{1}{2}\int_\Omega |\nabla\varphi|^2 \,\mathrm{d}\mu, \quad \varphi \in C_c^\infty(\mathbb{R}^d),\ \mu \in \mathcal{P}(\overline{\Omega}), \tag{4.7}$$

and its Legendre dual of $\mathbb{A}$ with respect to the second variable:

$$\mathbb{A}^*(\mu,\eta) = \sup_{\varphi\in C_c^\infty(\mathbb{R}^d)} \{\langle\varphi,\eta\rangle - \mathbb{A}(\mu,\varphi)\}$$

for any distribution $\eta \in \mathcal{D}'(\mathbb{R}^d)$. Note that $\mathbb{A}^*(\mu, w) = \mathbb{A}(\mu,\varphi)$ whenever $w = -\nabla\cdot(\mu\nabla\varphi)$. The connection to Wasserstein geometry is given by the infinitesimal Benamou–Brenier formula

$$\frac{1}{2}|\dot{\mu}_t|^2_{\mathbb{W}} = \mathbb{A}^*(\mu_t, \partial_t\mu_t).$$

Moreover, the relative Fisher information can be written as

$$\mathbf{I}(\mu) = 2\mathbb{A}\Big(\mu, -D\mathbf{H}(\mu)\Big), \tag{4.8}$$

where $D\mathbf{H}(\mu) = \log\rho$ is the $L^2(\mathbf{m})$-differential of $\mathbf{H}$. Hence, it follows that (4.5) can be stated equivalently as

$$\mathbf{H}(\mu_T) + \int_0^T \mathbb{A}^*(\mu_t, \dot{\mu}_t) + \mathbb{A}\Big(\mu_t, -D\mathbf{H}(\mu_t)\Big)\,\mathrm{d}t \le \mathbf{H}(\mu_0). \tag{4.9}$$

This formulation is particularly convenient to relate the discrete framework to the continuous one, as we discuss in the next subsection.

4.2.2 The discrete Fokker-Planck equation as gradient flow

We consider a finite volume discretisation of $\overline{\Omega}$, closely following the setup in [EGH00]. We thus consider finite partition $\mathcal{T}$ of $\overline{\Omega}$ into sets (called cells) with nonempty and convex interior. Note that all interior cells are polytopes. We assume that $\mathcal{T}$ is *admissible*, in the sense that each of

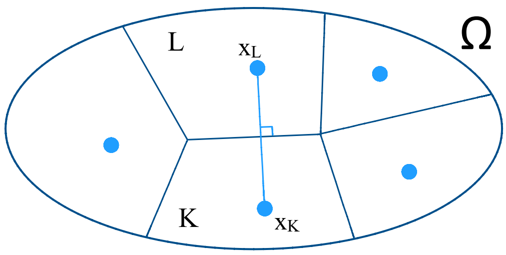

Figure 4.1: An admissible mesh with cells $K, L, \dots$ on a domain $\Omega \subset \mathbb{R}^d$.

the cells $K \in \mathcal{T}$ contains a point $x_K \in \overline{K}$ such that $x_K - x_L$ is orthogonal to the boundary surface $\Gamma_{KL} := \partial\overline{K} \cap \partial\overline{L}$, whenever K and L are *neighbouring cells*, i.e., $\mathscr{H}^{d-1}(\Gamma_{KL}) > 0$. In this case we write $K \sim L$. This in a standard finite-volume setup.

An admissible mesh is said to be ζ-*regular* for some $\zeta \in (0,1]$, if the following conditions hold:

$$\begin{aligned} &(\textit{inner ball}) \quad && B\Big(x_K, \zeta[\mathcal{T}]\Big) \subseteq K && \text{for all } K \in \mathcal{T}, \\ &(\textit{area bound}) \quad && \mathscr{H}^{d-1}(\Gamma_{KL}) \geq \zeta[\mathcal{T}]^{d-1} && \text{for all } K, L \in \mathcal{T} \text{ with } K \sim L, \end{aligned}$$

where $[\mathcal{T}] := \max\Big\{ \operatorname{diam}(K) \ : \ K \in \mathcal{T} \Big\}$ denotes the size of the mesh.

Discrete Fokker-Plank equations

We consider discrete Fokker-Planck equations of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(K) = \sum_{L \sim K} w_{KL} \left(\frac{m_t(L)}{\pi_{\mathcal{T}}(L)} - \frac{m_t(K)}{\pi_{\mathcal{T}}(K)} \right). \tag{4.10}$$

Here, the probability measure $\pi_{\mathcal{T}} \in \mathcal{P}(\mathcal{T})$ is the canonical discretisation of $\mathbf{m}$, and the coefficients w_{KL} are defined using the geometry of the mesh:

$$\pi_{\mathcal{T}}(\{K\}) := \mathbf{m}(K), \qquad w_{KL} := \frac{\mathscr{H}^{d-1}(\Gamma_{KL})}{|x_K - x_L|} S_{KL} \quad \text{for } K \sim L. \tag{4.11}$$

where S_{KL} is a suitable average of the stationary density σ on K and L, i.e., $S_{KL} := \theta\Big(\sigma(x_K), \sigma(x_L)\Big)$ for a fixed function $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\min\{a,b\} \leq \theta(a,b) \leq \max\{a,b\}$.

As (4.10) is the forward equation for a reversible Markov chain on $\mathcal{T}$, it follows from the theory in [Maa11] and [Mie11] that this equation is the gradient flow of the relative entropy $\mathcal{H}_{\mathcal{T}} : \mathcal{P}(\mathcal{T}) \to \mathbb{R}_+$ given by

$$\mathcal{H}_{\mathcal{T}}(m) := \sum_{K \in \mathcal{T}} m(K) \log \frac{m(K)}{\pi_{\mathcal{T}}(K)}.$$

The discrete analogue of (4.7) is given by the operator $\mathcal{A}_{\mathcal{T}} : \mathcal{P}(\mathcal{T}) \times \mathbb{R}^{\mathcal{T}} \to \mathbb{R}_+$ defined by

$$\mathcal{A}_{\mathcal{T}}(m, f) = \frac{1}{4} \sum_{K,L \in \mathcal{T}} \Big(f(K) - f(L) \Big)^2 \theta_{\log} \left(\frac{m(K)}{\pi_{\mathcal{T}}(K)}, \frac{m(L)}{\pi_{\mathcal{T}}(L)} \right) w_{KL}, \tag{4.12}$$

where $\theta_{\log}(a,b) = \frac{a-b}{\log a - \log b}$ denotes the logarithmic mean. Its Legendre transform $\mathcal{A}^*_{\mathcal{T}} : \mathcal{P}(\mathcal{T}) \times \mathbb{R}^{\mathcal{T}} \to \mathbb{R}$ with respect to the second variable is given by

$$\mathcal{A}^*_{\mathcal{T}}(m,\sigma) = \sup_{f \in \mathbb{R}^{\mathcal{T}}} \left\{ \sum_{K \in \mathcal{T}} \sigma(K) f(K) - \mathcal{A}_{\mathcal{T}}(m,f) \right\}.$$

In analogy to (4.9), we can formulate the gradient flow structure for the discrete Fokker-Planck equation (4.10) in terms of the *discrete EDI*

$$\mathcal{H}_{\mathcal{T}}(m_T) + \int_0^T \mathcal{A}^*_{\mathcal{T}}(m_t, \dot{m}_t) + \mathcal{A}_{\mathcal{T}}\Big(m_t, -D\mathcal{H}_{\mathcal{T}}(m_t)\Big)\,\mathrm{d}t \le \mathcal{H}_{\mathcal{T}}(m_0). \tag{4.13}$$

The discrete counterpart of (4.8) is the *discrete Fisher information* $\mathcal{I}_{\mathcal{T}}(m)$ given by

$$\mathcal{I}_{\mathcal{T}}(m) := 2\mathcal{A}_{\mathcal{T}}\Big(m, -D\mathcal{H}_{\mathcal{T}}(m)\Big), \qquad m \in \mathcal{P}(\mathcal{T}).$$

4.3 Statement of the results

In this section we present our main result, the evolutionary Γ-convergence of the gradient flow structures in the discrete-to-continuum limit for Fokker-Planck equations on a bounded convex domain $\Omega \subset \mathbb{R}^d$.

Let $\mathcal{T}$ be an admissible mesh on Ω. To compare measures on different spaces we introduce the canonical projection and embedding operators $\mathsf{P}_{\mathcal{T}}$ and $\mathsf{Q}_{\mathcal{T}}$ defined by

$$\begin{aligned} &\mathsf{P}_{\mathcal{T}} : \mathcal{M}(\overline{\Omega}) \to \mathcal{M}(\mathcal{T}) && \Big(\mathsf{P}_{\mathcal{T}}\mu\Big)(K) = \mu(K) && \text{for } K \in \mathcal{T}, \\ &\mathsf{Q}_{\mathcal{T}} : \mathcal{M}(\mathcal{T}) \to \mathcal{M}(\overline{\Omega}) && \mathsf{Q}_{\mathcal{T}} m = \sum_{K \in \mathcal{T}} m(K)\,\mathcal{U}_K && \text{for } m \in \mathcal{P}(\mathcal{T}). \end{aligned} \tag{4.14}$$

Here, $\mathcal{U}_K$ denotes the uniform probability measure on $K \subset \overline{\Omega}$, and $\mathcal{M}(\mathcal{X})$ denotes the set of finite measures on the space $\mathcal{X}$. In particular, $\mathsf{Q}_{\mathcal{T}}$ is a right-inverse of $\mathsf{P}_{\mathcal{T}}$ and both mappings are mass and positivity preserving. By construction we have $\pi_{\mathcal{T}} := \mathsf{P}_{\mathcal{T}}\mathbf{m}$.

It is also useful to introduce an operator for the piecewise constant embedding of functions:

$$Q_{\mathcal{T}} : \mathbb{R}^{\mathcal{T}} \to L^\infty(\overline{\Omega}), \qquad \Big(Q_{\mathcal{T}} f\Big)(x) = f(K) \quad \text{for } x \in K,\ K \in \mathcal{T}.$$

Let us now consider a sequence of admissible, ζ-regular meshes $\mathcal{T}_N$ with mesh size $[\mathcal{T}_N] \to 0$ as $N \to \infty$. To avoid towers of subscripts, we simply write $\mathcal{A}_N := \mathcal{A}_{\mathcal{T}_N}$, $\mathsf{P}_N := \mathsf{P}_{\mathcal{T}_N}$, etc.

4.3.1 Evolutionary Γ-convergence of discrete Fokker-Planck equations

In this subsection we fix a reference probability $\mathbf{m} \in \mathcal{P}(\overline{\Omega})$ with density $\sigma(x) = \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}x} = \frac{1}{Z_V} e^{-V(x)}$ as in (4.3). For neighbouring cells $K, L \in \mathcal{T}_N$ we fix $S_{KL} > 0$ such that

$$\min\Big\{\sigma(x_K), \sigma(x_L)\Big\} \le S_{KL} \le \max\Big\{\sigma(x_K), \sigma(x_L)\Big\} \tag{4.15}$$

as in Section 4.2.

We start by collecting some conditions of the densities that will be imposed in the sequel.

Definition 4.3.1 (Assumptions on approximating sequences)**.** Let $(\mathcal{T}_N)_N$ be a vanishing sequence of ζ-regular meshes for some $\zeta > 0$. For a sequence of measures $m_N \in \mathcal{P}(\mathcal{T}_N)$ with densities $r_N = \frac{\mathrm{d}m_N}{\mathrm{d}\pi_N}$, we consider the following conditions:

(i) The *density lower bound* holds if, for some $\underline{k} > 0$,

$$\inf_{K \in \mathcal{T}_N} r_N(K) \geq \underline{k} > 0 \quad \forall N \in \mathbb{N}. \tag{lb}$$

(ii) The *density upper bound* holds if, for some $\bar{k} < \infty$,

$$\sup_{K \in \mathcal{T}_N} r_N(K) \leq \bar{k} < +\infty \quad \forall N \in \mathbb{N}. \tag{ub}$$

(iii) The *neighbour continuity bound* holds if

$$\lim_{N \to \infty} \sup_{\substack{K, L \in \mathcal{T}_N \\ K \sim L}} |r_N(K) - r_N(L)| = 0. \tag{nc}$$

(iv) The *pointwise condition* holds if there exists a measure $\mu \in \mathcal{P}(\overline{\Omega})$ with density $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{m}}$ such that $\mu_N := \mathsf{Q}_N m_N \rightharpoonup \mu$ and, for a.e. $x_0 \in \Omega$:

$$\lim_{\varepsilon \to 0} \liminf_{N \to \infty} \sup_{x \in Q_\varepsilon(x_0)} \rho_N(x) \leq \rho(x_0) \leq \lim_{\varepsilon \to 0} \limsup_{N \to \infty} \inf_{x \in Q_\varepsilon(x_0)} \rho_N(x). \tag{pc}$$

Here, $Q_\varepsilon(x_0)$ denotes the open cube of side-length $\varepsilon > 0$ centered at x_0, and $\rho_N(x) := r_N(K)$ for $x \in K$.

Remark 4.3.2. These conditions do not depend on the reference measure $\mathbf{m}$, except for the value of the constants $\underline{k}$ and $\bar{k}$. Clearly, the pointwise condition holds if ρ belongs to $C(\overline{\Omega})$ and ρ_n converges uniformly to ρ. Moreover, this condition implies subsequential pointwise convergence of ρ_N to ρ.

We now present the crucial Γ-liminf inequalities for the functionals in the EDI (4.13).

Theorem 4.3.3 (Lower bounds for functionals)**.** *Let $(\mathcal{T}_N)_N$ be a vanishing sequence of ζ-regular meshes for some $\zeta > 0$. The following assertions hold for any $\mu \in \mathcal{P}(\overline{\Omega})$ and $m_N \in \mathcal{P}(\mathcal{T}_N)$ such that $\mathsf{Q}_N m_N \rightharpoonup \mu$ as $N \to \infty$:*

(i) The relative entropy functionals satisfy the liminf inequality

$$\liminf_{N \to \infty} \mathcal{H}_N(m_N) \geq \mathbf{H}(\mu). \tag{4.16}$$

(ii) Assume (nc). The Fisher information functionals satisfy the liminf inequality

$$\liminf_{N \to \infty} \mathcal{I}_N(m_N) \geq \mathbf{I}(\mu). \tag{4.17}$$

(iii) Assume (lb), (ub), and (pc). For any $\eta \in L^2(\Omega)$ and any $e_N \in \mathbb{R}^{\mathcal{T}_N}$ such that $Q_N e_N \rightharpoonup \eta$ in $L^2(\Omega)$ we have

$$\liminf_{N \to \infty} \mathcal{A}_N^*(m_N, e_N) \geq \mathbb{A}^*(\mu, \eta). \tag{4.18}$$

The same bound holds without assuming (lb) if $(e_N)_N$ satisfies the additional assumption $\limsup_{N \to \infty} \mathcal{A}_N^*(\pi_N, e_N) < +\infty$.

Remark 4.3.4. We emphasize that the lower bound (lb) is not required to obtain (4.16) and (4.17).

Remark 4.3.5. The bound (4.18) can be obtained without assuming (ub) and (pc) if the mesh satisfies the so-called asymptotic isotropy condition (4.27); cf. Definition 4.3.11 and Proposition 4.3.12 below.

Remark 4.3.6. If $\mu \in \mathcal{P}(\overline{\Omega})$ is absolutely continuous with respect to the Lebesgue measure and $m_N = \mathsf{P}_N\mu$, (4.18) can be proved under the assumptions that $\eta \in \mathcal{M}(\Omega)$ and $e_N \in \mathbb{R}^{\mathcal{T}_N}$ satisfies $\mathsf{Q}_N e_N \rightharpoonup \eta$ in $\mathcal{D}^{'}(\Omega)$. This is a consequence of an explicit construction of a recovery sequence for the action $\mathcal{A}_N(m_N, \cdot)$ (as in the isotropic case in Proposition 4.3.12); cf. Remark 4.6.7.

Using Theorem 4.3.3 we obtain our main result, the evolutionary Γ-convergence of the discrete gradient flow structures. The following result shows that one can pass to the limit in each of the terms of the discrete gradient flow formulation (4.13) and recover the Wasserstein gradient flow structure as a consequence.

Theorem 4.3.7 (Evolutionary Γ-convergence)**.** *Let $T > 0$ and consider a vanishing sequence of ζ-admissible meshes $(\mathcal{T}_N)_N$. Fix an initial measure $\mu_0 \in \mathcal{P}(\overline{\Omega})$ such that $\mathbf{H}(\mu_0) < +\infty$, together with measures $m_0^N \in \mathcal{P}(\mathcal{T}_N)$ for $N \geq 1$, that are well-prepared in the sense that*

$$\mathsf{Q}_N m_0^N \rightharpoonup \mu_0 \quad \textit{and} \quad \lim_{N\to\infty} \mathcal{H}_N(m_0^N) = \mathbf{H}(\mu_0).$$

For each $N \geq 1$, let $(m_t^N)_{t\in[0,T]}$ be the solution to the discrete Fokker-Planck equation (4.10) with initial datum m_0^N, which satisfies the EDI

$$\mathcal{H}_N(m_t^N) + \int_0^T \mathcal{A}_N^*(m_t^N, \dot{m}_t^N) + \mathcal{A}_N\Big(m_t^N, -D\mathcal{H}_N(m_t^N)\Big)\,\mathrm{d}t \leq \mathcal{H}_N(m_0^N).$$

Then:

(i) The sequence of curves $(\mu^N)_N$ defined by $\mu_t^N := \mathsf{Q}_N m_t^N$ is compact in the space $C\Big([0,T];(\mathcal{P}(\overline{\Omega}), \mathbb{W})\Big)$. Thus, up to a subsequence, we have

$$\sup_{t\in[0,T]} \mathbb{W}\Big(\mu_t^N, \mu_t\Big) \to 0 \quad \textit{as } N \to \infty. \tag{4.19}$$

(ii) The following estimates hold:

$$\textit{Entropy:} \qquad \liminf_{N\to\infty} \mathcal{H}_N(m_t^N) \geq \mathbf{H}(\mu_t) \quad \forall t \in [0,T], \tag{4.20a}$$

$$\textit{Fisher I.:} \quad \liminf_{N\to\infty} \int_0^T \mathcal{A}_N\Big(m_t^N, -D\mathcal{H}_N(m_t^N)\Big)\,\mathrm{d}t \geq \int_0^T \mathbb{A}\Big(\mu_t, -D\mathbf{H}(\mu_t)\Big)\,\mathrm{d}t, \tag{4.20b}$$

$$\textit{Speed:} \qquad \liminf_{N\to\infty} \int_0^T \mathcal{A}_N^*(m_t^N, \dot{m}_t^N)\,\mathrm{d}t \geq \int_0^T \mathbb{A}^*(\mu_t, \dot{\mu}_t)\,\mathrm{d}t. \tag{4.20c}$$

(iii) The curve (μ_t) solves the EDI (4.9), and hence, the Fokker-Planck equation (4.1).

Remark 4.3.8. The well-preparedness assumption holds in the special case where the discrete measures are defined by $m_0^N := \mathsf{P}_N\mu_0$ as in (4.14). Indeed, in that case we have $\mathcal{H}_N(m_0^N) = \mathbf{H}(\mathsf{Q}_N\mathsf{P}_N\mu_0)$, so that the convergence of the relative entropy functionals follows from Jensen's inequality.

The proofs of Theorem 4.3.3 and Theorem 4.3.7 appear in Section 4.4. They rely on a Mosco convergence result for discrete energy functionals of independent interest, which we will now describe.

4.3.2 Mosco convergence of Dirichlet energy functionals

Fix an absolutely continuous probability measure $\mu \in \mathcal{P}(\overline{\Omega})$, and assume that its density υ with respect to the Lebesgue measure on $\overline{\Omega}$ satisfies the two-sided bounds

$$0 < \underline{c} \le \upsilon(x) \le \overline{c} < \infty \text{ for all } x \in \overline{\Omega}.$$

We consider the *continuous Dirichlet energy* $\mathbb{F}_\mu : L^2(\Omega) \to \mathbb{R}_+ \cup \{+\infty\}$ given by

$$\mathbb{F}_\mu(\varphi) := \mathbb{A}(\mu, \varphi) = \begin{cases} \frac{1}{2}\int_\Omega |\nabla\varphi|^2 \,\mathrm{d}\mu & \text{if } \varphi \in H^1(\Omega), \\ +\infty & \text{otherwise} \end{cases} \tag{4.21}$$

where $\mathbb{A}$ is defined in (4.7).

Similarly, for a ζ-regular mesh $\mathcal{T}$ and a probability measure $m \in \mathcal{P}(\mathcal{T})$, we consider the *discrete Dirichlet energy* $\mathcal{F}_\mathcal{T} : \mathbb{R}^\mathcal{T} \to \mathbb{R}_+$ defined by

$$\mathcal{F}_\mathcal{T}(f) = \frac{1}{4}\sum_{K,L\in\mathcal{T}} \Big(f(K) - f(L)\Big)^2 U_{KL} \frac{\mathscr{H}^{d-1}(\Gamma_{KL})}{|x_K - x_L|} \tag{4.22}$$

where $\min\left\{\frac{m(K)}{|K|}, \frac{m(L)}{|L|}\right\} \le U_{KL} \le \max\left\{\frac{m(K)}{|K|}, \frac{m(L)}{|L|}\right\}$. In the special case where U_{KL} is defined in terms of the logarithmic mean of r_K and r_L, namely, $U_{KL} = \frac{r_K - r_L}{\log r_K - \log r_L} S_{KL}$ with $r_K = \frac{m(K)}{\pi_\mathcal{T}(K)}$, this functional is related to the functional $\mathcal{A}_\mathcal{T}$ by

$$\mathcal{F}_\mathcal{T}(f) := \mathcal{A}_\mathcal{T}(m, f). \tag{4.23}$$

To compare the discrete and the continuous functionals we consider the embedded funtionals $\tilde{\mathbb{F}}_\mathcal{T} : L^2(\Omega) \to \mathbb{R}_+ \cup \{+\infty\}$ defined by

$$\tilde{\mathbb{F}}_\mathcal{T}(\varphi) := \begin{cases} \mathcal{F}_\mathcal{T}\Big(P_\mathcal{T}\varphi\Big) & \text{if } \varphi \in \mathsf{PC}_\mathcal{T}, \\ +\infty & \text{otherwise,} \end{cases} \tag{4.24}$$

where $\mathsf{PC}_\mathcal{T}$ denotes the space of all functions in $L^2(\Omega)$ that are constant a.e. on each cell $K \in \mathcal{T}$, and

$$\Big(P_\mathcal{T}\varphi\Big)(K) := \varphi(x_K) \quad \text{for } \varphi : \Omega \to \mathbb{R}. \tag{4.25}$$

We then obtain the following convergence result. For the definition of Mosco convergence we refer to Definition 4.5.1 below.

Theorem 4.3.9 (Mosco convergence)**.** *Let $(\mathcal{T}_N)_N$ be a vanishing sequence of ζ-regular meshes, and suppose that μ and $(m_N)_N$ satisfy* (lb)*,* (ub)*, and* (pc)*. Then we have Mosco convergence $\tilde{\mathbb{F}}_{\mathcal{T}_N} \xrightarrow{M} \mathbb{F}_\mu$ with respect to the $L^2(\Omega)$-topology.*

The proof of this result follows the strategy developed in [AC04], where similar Γ-convergence results have been obtained for more general energy functionals on a particular mesh (the cartesian grid). In that paper, the authors do not explicitly characterise the limiting functional, except in special situations, such as the periodic setting. For our application to evolutionary Γ-convergence, a characterisation of the limiting functional is crucial.

Remark 4.3.10. Mosco convergence of Dirichlet energy functionals is equivalent to strong convergence of the associated semigroups [Mos94]; see also [KS03] for a generalisation to Dirichlet forms defined on different spaces.

4.3.3 Related work

We close this section with some comments on related work.

Convergence of the discrete Fokker-Planck equations

It is well known that the discrete heat flow converges to the continuous heat flow for any sequence of admissible meshes with vanishing diameter. The authors in [EGH00], [BHO18] exploit classical Sobolev a priori estimates and pass to the limit in the weak formulation of the equation, in dimension 2 and 3 (see [BHO18, Lemma 8]). A unified framework for discretisation of partial differential equations in higher dimension can be found in [DEG+18]. Convergence results for finite-volume discretisations of Fokker-Planck equations based on different Stolarsky means have recently been obtained in [HKS20].

Entropy gradient flows in discrete settings

Entropy gradient flow structures for discrete dynamics have been intensively studied in discrete settings following the papers [CHLZ12, Maa11, Mie11]. Many subsequent works deal with connections to curvature bounds and functional inequalities [EM12, Mie13, EM14, EMT15, FM16, EMW19]. Entropy gradient flow structures have also been exploited to analyse the discrete-to-continuum limit from several perspectives, see, e.g., [CG17, CGLM19, CMRS19, BBC20].

Evolutionary Γ-convergence for Fokker-Planck in 1D

Evolutionary Γ-convergence of the discrete gradient flow structures for Fokker-Planck equations has been proved in the one-dimensional setting under additional geometric conditions using methods that do not extend straightforwardly to higher dimensions [DL15].

In particular, the authors work with meshes that satisfy the *center of mass condition*

$$\fint_{\Gamma_{KL}} x \, \mathrm{d}\mathscr{H}^{d-1} = \frac{x_K + x_L}{2}, \quad \text{for all } K \sim L \in \mathcal{T}. \tag{4.26}$$

This condition implies the Gromov-Hausdorff convergence of the associated transport metrics [GKM20]. Here, we work with more general meshes for which Gromov-Hausdorff convergence of the associated transport metrics does not always hold.

Moreover, in one dimension, it is possible to construct explicit solutions to the continuity equation from discrete vector fields using linear interpolation techniques. As such methods are not available in higher dimensions, we take a more variational approach in this paper.

Scaling limits for discrete optimal transport in any dimension.

The crucial liminf inequality (4.18) can be proved under weaker assumptions on the approximating sequence of measures if the meshes satisfy a suitable isotropy condition, which we will now recall.

Definition 4.3.11 (Asymptotic isotropy)**.** A vanishing sequence of meshes $(\mathcal{T}_N)_N$ is said to satisfy the *asymptotic isotropy condition* if, for every $N \in \mathbb{N}$,

$$\frac{1}{2} \sum_{L \in \mathcal{T}_N} w_{KL} (x_K - x_L) \otimes (x_K - x_L) \leq \pi_N(K)\Big(I_d + \eta_{\mathcal{T}_N}(K)\Big) \qquad \forall K \in \mathcal{T}_N, \tag{4.27}$$

where $\sup_{K \in \mathcal{T}_N} \|\eta_{\mathcal{T}}(K)\| \to 0$ as $N \to \infty$.

Under this condition, the following following version of (4.18) has been proved in [GKM20, Proposition 6.6]. In that paper the reference measure is the Lebesgue measure. Here we formulate a slight generalisation with the reference measure $\mathbf{m}$.

Proposition 4.3.12 (Action bounds)**.** *Let $(\mathcal{T}_N)_N$ be a vanishing sequence of meshes satisfying the asymptotic isotropy condition (4.27). Let $\mu \in \mathcal{P}(\overline{\Omega})$ and $\eta \in \mathcal{M}_0(\overline{\Omega})$, and suppose that $m_N \in \mathcal{P}(\mathcal{T}_N)$ and $e_N \in \mathcal{M}_0(\mathcal{T}_N)$ satisfy $\mathsf{Q}_N m_N \rightharpoonup \mu$ and $\mathsf{Q}_N e_N \rightharpoonup \eta$ as $N \to \infty$. Then we have the lower bound*

$$\liminf_{N \to \infty} \mathcal{A}_N^*(m_N, e_N) \geq \mathbb{A}^*(\mu, \eta). \tag{4.28}$$

It has also been shown in [GKM20] that Gromov–Hausdorff convergence of the associated transport distances holds under the asymptotic isotropy condition; see also [GKMP20] for a study of the limiting metric in the one-dimensional periodic setting. In the current paper we do not assume that the discrete meshes satisfy an isotropy condition.

Notation

Throughout the paper we use the notation $a \lesssim b$ (or $b \gtrsim a$) if $a \leq Cb$ with $C < \infty$ depending only on Ω, ζ, and $\mathbf{m}$. We write $a \eqsim b$ if $a \lesssim b$ and $a \gtrsim b$.

4.4 Proof of the main result: the Wasserstein evolutionary Γ-convergence

In this section we prove our main result, the evolutionary Γ-convergence of the discrete gradient flow structures (Theorem 4.3.7). The section is divided into three parts: the first subsection concerns the proof of Theorem 4.3.3, which relies on Theorem 4.3.9. The second subsection contains a proof of compactness for the continuously embedded discrete solutions. In the third and final part we complete the proof of Theorem 4.3.7.

4.4.1 Asymptotic lower bounds for the functionals

Let μ and m_N be as in the statement of Theorem 4.3.3. Write $\mu_N := \mathsf{Q}_N m_N$ and let ρ_N be the density of μ_N with respect to $\mathbf{m}$.

Proof of Theorem 4.3.3. The proof consists of three parts.

(i) Lower bound for the entropy. Note that $\mathcal{H}_N(m_N) = \mathsf{Ent}(\mu_N|\mathsf{Q}_N\pi_N)$ and $\mathbf{H}(\mu) = \mathsf{Ent}(\mu|\mathbf{m})$, where $\mathsf{Ent}(\cdot|\cdot)$ denotes the relative entropy. Since $\mu_N \rightharpoonup \mu$ and $\mathsf{Q}_N\pi_N \rightharpoonup \mathbf{m}$, the result follows immediately from the joint weak lower semicontinuity of $\mathsf{Ent}(\cdot|\cdot)$, see, e.g., [AGS08, Lemma 9.4.3].

(ii) Lower bound for the Fisher information. Assume that (nc) holds. We first prove the lower bound (4.17) under the additional assumption (lb). This assumption will be removed at the end of the proof. The key identity for the Fisher information is

$$\tilde{\mathcal{A}}_N\Big(m_N, -D\mathcal{H}_N(m_N)\Big) = 4\mathcal{E}_N\Big(\sqrt{r_N}\Big), \tag{4.29}$$

where $\mathcal{E}_N(f) := \mathcal{A}_N(\pi_N, f)$ is the discrete Dirichlet energy with reference measure π_N, and $\tilde{\mathcal{A}}_N$ is defined by replacing the logarithmic mean $\theta_{\log}$ in the definition of $\mathcal{A}_N$ by $\tilde{\theta}(a,b) := \theta_{\log}(\sqrt{a},\sqrt{b})^2$. Since $\min\{a,b\} \le \tilde{\theta}(a,b), \theta_{\log}(a,b) \le \max\{a,b\}$, we have

$$|\theta_{\log}(a,b) - \tilde{\theta}(a,b)| \le |a-b| \le \frac{|a-b|}{\min\{a,b\}}\tilde{\theta}(a,b).$$

The assumptions (nc) and (lb) yield

$$\varepsilon_N := \sup_{\substack{K,L\in\mathcal{T}_N\\ K\sim L}} |r_N(K) - r_N(L)| \to 0 \quad \text{and} \quad \inf_{K\in\mathcal{T}_N} r_N(K) \ge \underline{k}, \tag{4.30}$$

Using these estimates and the identity $(\log a - \log b)^2\tilde{\theta}(a,b) = 4\Big(\sqrt{a}-\sqrt{b}\Big)^2$ we obtain

$$\begin{aligned}
\Big|\tfrac{1}{2}\mathcal{I}_N(m_N) - 4\mathcal{E}_N(\sqrt{r_N})\Big| &= \Big|\Big(\mathcal{A}_N - \tilde{\mathcal{A}}_N\Big)\Big(m_N, -D\mathcal{H}_N(m_N)\Big)\Big| \\
&= \frac{1}{4}\sum_{K,L\in\mathcal{T}_N} w_{KL}\Big(\log r_N(K) - \log r_N(L)\Big)^2 \\
&\qquad \times \Big(\theta_{\log}\Big(r_N(K), r_N(L)\Big) - \tilde{\theta}_{\log}\Big(r_N(K), r_N(L)\Big)\Big) \\
&\le \frac{4\varepsilon_N}{\underline{k}}\mathcal{E}_N\Big(\sqrt{r_N}\Big).
\end{aligned} \tag{4.31}$$

Let us now assume that $\sup_N \mathcal{I}_N(m_N) < \infty$ along a subsequence; if this were not the case, the result holds trivially. The previous bound implies that also $\sup_N \mathcal{E}_N\Big(\sqrt{r_N}\Big) < \infty$, hence $\Big(\sqrt{\rho_N}\Big)_N$ has a subsequence that converges strongly in $L^2(\Omega,\mathbf{m})$ by Proposition 4.6.5 below. Let $g \in L^2(\Omega,\mathbf{m})$ be its limit. Since $\|\rho_N - g^2\|_{L^1} \le \|\sqrt{\rho_N} - g\|_{L^2}\|\sqrt{\rho_N}+g\|_{L^2}$, we infer that $\rho_N \to g^2$ in $L^1(\Omega,\mathbf{m})$. As $\mu_N = \rho_N\mathbf{m} \rightharpoonup \mu$ in $\mathcal{P}(\overline{\Omega})$ by assumption, we infer that $\mu = \rho\mathbf{m}$ with $\rho := g^2$. Now we apply (4.31) and the Mosco convergence from Theorem 4.3.9 to obtain

$$\liminf_{N\to\infty} \mathcal{I}_N(m_N) \ge \liminf_{N\to\infty} 8\mathcal{E}_N\Big(\sqrt{r_N}\Big) \ge 8\mathbb{A}\Big(\mathbf{m}, \sqrt{\rho}\Big) = \mathbf{I}(\mu),$$

which concludes the proof.

Let us now show how to remove the assumption (lb). The argument is based on the convexity of $m \mapsto \mathcal{I}_N(m)$, which is a consequence of the joint convexity of the map $(a,b) \mapsto (a-b)(\log a - \log b)$ on $(0,\infty)\times(0,\infty)$.

Pick $\delta \in (0,1)$ and set $m_N^\delta := (1-\delta)m_N + \delta\pi_N$. Note that m_N^δ satisfies (lb) with $\underline{k} = \delta$. Moreover, $\mathsf{Q}_N m_N^\delta \rightharpoonup \mu^\delta := (1-\delta)\mu + \delta\mathbf{m}$. Applying the first part of the result we obtain

$$\mathbf{I}(\mu^\delta) \le \liminf_{N\to\infty} \mathcal{I}_N(m_N^\delta) \le (1-\delta) \liminf_{N\to\infty} \mathcal{I}_N(m_N)$$

for every $\delta \in (0,1]$, where the last inequality uses the convexity of $\mathcal{I}_N$ and the fact that $\mathcal{I}_N(\pi_N) = 0$. Since $\mu^\delta \rightharpoonup \mu$, the result follows from the lower semicontinuity of $\mathbf{I}$ with respect to the weak convergence in $\mathcal{P}(\overline{\Omega})$; see [GST09, Lemma 2.2].

(iii) Lower bound for $\mathcal{A}_N^*$. Assume first that (lb), (ub), and (pc) hold. Fix $\eta \in L^2(\Omega,\mathbf{m})$ and let $e_N \in L^2(\mathcal{T}_N, \pi_N)$ be such that $\mathsf{Q}_N e_N \rightharpoonup \eta$ in $L^2(\Omega,\mathbf{m})$. Theorem 4.3.9 (in particular, the existence of a recovery sequence) implies that for every $\varphi \in C_c(\Omega)$ there exist $f_N \in L^2(\mathcal{T}_N,\pi_N)$ such that $\mathsf{Q}_N f_N \to \varphi$ in $L^2(\Omega,\mathbf{m})$ and

$$\limsup_{N\to\infty} \mathcal{A}_N(m_N, f_N) \le \mathbb{A}(\mu,\varphi).$$

Since $\mathsf{Q}_N e_N \rightharpoonup \eta$ in $L^2(\Omega,\mathbf{m})$, it follows that $\langle e_N, f_N\rangle_{L^2(\mathcal{T}_N,\pi_N)} \to \langle \eta,\varphi\rangle_{L^2(\Omega,\mathbf{m})}$ and

$$\begin{aligned}\langle \eta,\varphi\rangle_{L^2(\Omega,\mathbf{m})} - \mathbb{A}(\mu,\varphi) &\le \liminf_{N\to\infty} \langle e_N, f_N\rangle_{L^2(\mathcal{T}_N,\pi_N)} - \mathcal{A}_N(m_N,f_N)\\ &\le \liminf_{N\to\infty} \mathcal{A}_N^*(m_N,e_N).\end{aligned}$$

Taking the supremum over φ, we infer that $\mathbb{A}^*(\mu,\eta) \le \liminf_{N\to\infty} \mathcal{A}_N^*(m_N,e_N)$, as desired.

Assume now that (ub), (pc) hold, and that $E := \limsup_{N\to\infty} \mathcal{A}_N^*(\pi_N,e_N) < +\infty$, instead of (lb). The key observation is that the map $m_N \mapsto \mathcal{A}_N^*(m_N,e_N)$ is convex: indeed, the concavity of $\theta_{\log}$ implies the concavity of $m_N \mapsto \mathcal{A}(m_N,f_N)$, and thus the convexity of its Legendre dual as a supremum of convex maps. To take advantage of this fact, we fix $\delta\in(0,1)$ and define $m_N^\delta := (1-\delta)m_N + \delta\pi_N$. Note that $\mathsf{Q}_N m_N^\delta \rightharpoonup \mu^\delta := (1-\delta)\mu + \delta\mathbf{m}$ and m_N^δ satisfies (lb) with $\underline{k} = \delta$. We may thus apply the first part of the result and the convexity to obtain

$$\begin{aligned}\mathbb{A}^*(\mu^\delta,\eta) &\le \liminf_{N\to\infty} \mathcal{A}_N^*(m_N^\delta,e_N)\\ &\le \liminf_{N\to\infty} (1-\delta)\mathcal{A}_N^*(m_N,e_N) + \delta\mathcal{A}_N^*(\pi_N,e_N)\\ &\le (1-\delta)\Big(\liminf_{N\to\infty} \mathcal{A}_N^*(m_N,e_N)\Big) + \delta E.\end{aligned}$$

Using the weak lower semicontinuity of $\mathbb{A}^*(\cdot,\eta)$, we obtain the desired inequality (4.18) by passing to the limit $\delta\to 0$. □

4.4.2 Compactness and space-time regularity

In this section we prove the compactness of the family of embedded discrete gradient flow curves $(t\mapsto\mu_t^N)_N$ in the space $C\Big([0,T];(\mathcal{P}(\overline{\Omega}),\mathbb{W})\Big)$. We follow the strategy of [LMPR17, Theorem 3.1], which is based on a metric Ascoli-Arzelá theorem. The corresponding one-dimensional result has been obtained in [DL15] using explicit interpolation formulas that are not available in the multi-dimensional setting.

Our proof is based on the following coarse energy bound from [GKM20, Lemma 3.4]. Here and below, $(H_t)_{t\geq 0}$ denotes the Neumann heat semigroup on Ω. Moreover, $\mathcal{M}_0(\mathcal{T})$ denotes the space of signed measure on $\mathcal{T}$ with zero total mass.

Lemma 4.4.1 (Coarse energy bound). *Fix $\zeta \in (0,1]$. There exists a constant $C < \infty$ such that for any ζ-regular mesh $\mathcal{T}$, for any $m \in \mathcal{P}(\mathcal{T})$ and any $\sigma \in \mathcal{M}_0(\mathcal{T})$, we have*

$$\mathbb{A}^*\Big(H_{[\mathcal{T}]}\mathsf{Q}_{\mathcal{T}} m, H_{[\mathcal{T}]}\mathsf{Q}_{\mathcal{T}}\sigma\Big) \leq C\mathcal{A}^*_{\mathcal{T}}(m,\sigma). \tag{4.32}$$

Let us stress that this result holds without any isotropy assumption on the mesh.

Lemma 4.4.2 ($\mathbb{W}$-Equi-continuity). *Let $\{\mathcal{T}_N\}_N$ be a vanishing sequence of ζ-regular meshes. For each $N \in \mathbb{N}$, let $(m_t^N)_{t\in[0,T]}$ be a continuous curve in $\mathcal{P}(\mathcal{T}_N)$, and suppose that the following uniform energy bound holds:*

$$A := \sup_N \int_0^T \mathcal{A}^*_N\Big(m_t^N, \dot m_t^N\Big)\,\mathrm{d}t < +\infty. \tag{4.33}$$

Then the curves $\tilde\mu^N : [0,T] \to \Big(\mathcal{P}(\overline{\Omega}), \mathbb{W}\Big)$ defined by $\tilde\mu_t^N := H_{[\mathcal{T}_N]}\mathsf{Q}_N m_t^N$ are equi-$\frac{1}{2}$-Hölder continuous, i.e., for $0 \leq s < t \leq T$ we have

$$\mathbb{W}\Big(\tilde\mu_t^N, \tilde\mu_s^N\Big) \lesssim \sqrt{A(t-s)}. \tag{4.34}$$

Proof. For $0 \leq s \leq t \leq T$ we invoke the Benamou-Brenier formula (4.4) and Lemma 4.4.1 to obtain

$$\begin{aligned}\mathbb{W}^2\Big(\tilde\mu_t^N, \tilde\mu_s^N\Big) &\leq (t-s)\int_s^t \mathbb{A}^*\Big(\tilde\mu_h^N, \partial_h\tilde\mu_h^N\Big)\,\mathrm{d}h\\ &\lesssim (t-s)\sup_N \int_0^T \mathcal{A}^*_N\Big(m_h^N, \partial_h m_h^N\Big)\,\mathrm{d}h \leq A(t-s),\end{aligned}$$

which concludes the proof. □

A corollary of Lemma 4.4.2 is the following compactness and regularity result.

Proposition 4.4.3 (Compactness and regularity). *For $t \in [0,T]$ and $N \geq 1$, let $\mu_t^N := \mathsf{Q}_N m_t^N \in \mathcal{P}(\overline{\Omega})$ be defined as in Theorem 4.3.7, and let ρ_t^N be the density of μ_t^N with respect to $\mathbf{m}$. There exists a $\mathbb{W}$-continuous curve $t \mapsto \mu_t \in \mathcal{P}(\overline{\Omega})$ satisfying, up to a subsequence,*

$$\sup_{t\in[0,T]} \mathbb{W}\Big(\mu_t^N, \mu_t\Big) \to 0 \quad \textit{as } N \to +\infty.$$

Proof. We apply Lemma 4.4.2 to the family of discrete gradient flow solutions $(t \mapsto m_t^N)_N$. In this case, the required estimate (4.33) follows directly from the discrete EDI (4.13) and the well-preparedness of the initial conditions $(m_0^N)_N$. Thus, Lemma 4.4.2 implies the $\mathbb{W}$-equi-continuity of the curves $(\tilde\mu^N)_N$ defined by $\tilde\mu_t^N := H_{\varepsilon_N}\mathsf{Q}_N m_t^N$, where $\varepsilon_N := [\mathcal{T}_N]$. The metric Arzelá-Ascoli Theorem [AGS08, Proposition 3.3.1] yields the existence of a limiting curve $t \mapsto \mu_t$ satisfying $\sup_t \mathbb{W}(\tilde\mu_t^N, \mu_t) \to 0$ as $N \to \infty$ (note that the compactness of μ_t^N for a fixed time t follows from the compactness of $\overline{\Omega}$). Using the well-known heat flow bound $\mathbb{W}(\tilde\mu_t^N, \mu_t^N) \leq C\sqrt{\varepsilon_N}$ (see, e.g., [GKM20, Lemma 2.2(iii)] for a proof), we obtain the desired result. □

4.4.3 Proof of the Wasserstein evolutionary Γ-convergence

In the proof of the main theorem we use the following regularity result for the discrete Fokker-Planck equation.

Proposition 4.4.4 (Regularity of the discrete flows)**.** *Let $\mathcal{T}$ be a ζ-regular mesh, let $(m_t)_t \subset \mathcal{P}(\mathcal{T})$ be a solution to the discrete Fokker-Planck equation, and set $r_t := \frac{\mathrm{d}m_t}{\mathrm{d}\pi}$.*

(i) For any $t > 0$ there exist $C = C(\Omega, \mathbf{m}, \zeta, t) < \infty$ and $\lambda = \lambda(\Omega, \mathbf{m}, \zeta) > 0$ such that the following Hölder estimate holds:

$$|r_t(K) - r_t(L)| \leq C|x_K - x_L|^\lambda \sup_{K' \in \mathcal{T}} |r_{t/2}(K')| \quad \forall K, L \in \mathcal{T}. \tag{4.35}$$

(ii) For any $t > 0$ the ultracontractivity estimate

$$\|r_t\|_{L^\infty(\pi_\mathcal{T})} \leq C\Big(1 \vee t^{-\frac{d}{2}}\Big)\|r_0\|_{L^1(\pi_\mathcal{T})} \tag{4.36}$$

holds with $C = C(\Omega, \mathbf{m}, \zeta) < \infty$.

We stress that the constants depend only on the aforementioned parameters. The proof of this result is based on standard arguments using volume doubling and a weak Poincaré inequality. We refer to Appendix A.2 for a sketch of the proof.

We will also use the following auxiliary result from [Ste08, Corollary 4.4].

Proposition 4.4.5 (Evolutionary Γ-liminf inequality)**.** *Let $\mathcal{X}$ be a separable Hilbert space and fix $T > 0$. Let $g_N, g_\infty : (0,T) \times \mathcal{X} \to [0, +\infty]$ be such that, for a.e. $t \in (0,T)$,*

(i) $g_N(t, \cdot), g_\infty(t, \cdot)$ are convex and lower semicontinuous;

(ii) $g_\infty(t, \varphi) \leq \inf\Big\{ \liminf_{N\to\infty} g_N(t, \varphi_N) \ : \ \varphi_N \rightharpoonup \varphi \text{ in } \mathcal{X}\Big\}$ for all $\varphi \in \mathcal{X}$.

Then, for any $\varphi_N, \varphi \in L^2(0,T;\mathcal{X})$ with $\varphi_N \rightharpoonup \varphi$ in $L^2(0,T;X)$, we have

$$\int_0^T g_\infty\Big(t, \varphi(t)\Big)\,\mathrm{d}t \leq \liminf_{N\to\infty} \int_0^T g_N\Big(t, \varphi_N(t)\Big)\,\mathrm{d}t. \tag{4.37}$$

Proof of Theorem 4.3.7. (i): The compactness of (μ^N) in $C\Big([0,T]; (\mathcal{P}(\overline{\Omega}), \mathbb{W})\Big)$ follows from Proposition 4.4.3.

(ii): We prove the inequalities in (4.20). The inequalities (4.20a) and (4.20b) follow straightforwardly from the bounds of Theorem 4.3.3. More work is required to prove (4.20c), as we only have time-averaged bounds on $\mathcal{A}_N^*(m_t^N, \dot{m}_t^N)$ along the discrete flows. Here we proceed using Proposition 4.4.5.

Evolutionary lower bound for the relative entropy (4.20a). In view of the weak convergence $\mathsf{Q}_N m_t^N \rightharpoonup \mu_t$, this bound follows from the liminf inequality for the entropies (4.16) obtained in Theorem 4.3.3.

Evolutionary lower bound for the Fisher information (4.20b). It follows from the Hölder regularity result in Proposition 4.4.4 that the sequence of discrete measures $(m_t^N)_N$ satisfies (nc) for any $t \in (0,T]$. Consequently, $\liminf_{N\to\infty} \mathcal{I}_N(m_t^N) \geq \mathbf{I}(\mu_t)$ by the liminf inequality for the relative Fisher information (4.17) obtained in Theorem 4.3.3. Therefore, the desired inequality (4.20b) follows from Fatou's Lemma.

Evolutionary lower bound for the metric derivative (4.20c). To ensure that our densities are bounded away from 0, we set

$$m_t^{N,\alpha} := (1-\alpha) m_t^N + \alpha \pi_N \quad \text{and} \quad \mu_t^\alpha := (1-\alpha)\mu_t + \alpha \mathbf{m}$$

for $\alpha \in (0,1)$. Fix $0 < \delta < (1 \wedge T)$ and define $g_N, g_\infty : (\delta, T) \times L^2(\Omega, \mathbf{m}) \to [0, +\infty]$ by

$$g_N^\alpha(t,\varphi) := \begin{cases} \mathcal{A}_N^*\Big(m_t^{N,\alpha}, (P_N\varphi)\pi_N\Big) & \text{if } \varphi \in \mathsf{PC}_N \\ +\infty & \text{otherwise} \end{cases}, \qquad g_\infty^\alpha(t,\varphi) := \mathbb{A}^*(\mu_t^\alpha, \varphi\mathbf{m}).$$

We will check that the conditions (i) and (ii) of Proposition 4.4.5 are satisfied.

Step 1. *Verification of conditions (i) and (ii).*

Clearly, the maps $g_N^\alpha(t,\cdot)$ are convex and lower semicontinuous in $L^2(\Omega,\mathbf{m})$ for every $t \in (\delta, T)$, which shows that condition (i) holds.

To verify condition (ii), we pick $\eta \in L^2(\Omega)$ and $e_N \in \mathbb{R}^{\mathcal{T}_N}$ such that $Q_N e_N \rightharpoonup \eta$ in $L^2(\Omega)$. We have to show that $\liminf_{N\to\infty} \mathcal{A}_N^*(m_t^{N,\alpha}, e_N) \geq \mathbb{A}^*(\mu_t^\alpha, \eta)$. To show this, we will check the conditions (ub), (lb), and (pc) of Theorem 4.3.3(iii).

Step 2. *Verification of* (ub), (lb), *and* (pc).

By construction, $(m_t^{N,\alpha})_N$ clearly satisfies (lb). Moreover, the hypercontractivity estimate from Proposition 4.4.4 implies that $(m_t^{N,\alpha})_N$ satisfies (ub). Therefore, it remains to show that $(m_t^{N,\alpha})_N$ satisfies (pc). Clearly, it suffices to prove that this property holds for (m_t^N).

To show this, we fix $x_0 \in \Omega$ and $\varepsilon > 0$. Let ρ_t^N be the density of μ_t^N with respect to $\mathbf{m}$. Using the Hölder regularity and the hypercontractivity result from Proposition 4.4.4, we infer that

$$|\rho_t^N(x) - \rho_t^N(y)| \leq C_t\Big(\varepsilon\sqrt{d} + 2[\mathcal{T}_N]\Big)^\lambda =: E_t^N(\varepsilon)$$

for any $x, y \in Q_\varepsilon(x_0)$, for a suitable t-dependent constant $C_t < \infty$ and $\lambda \in (0,1]$. It follows that

$$\left(\sup_{Q_\varepsilon(x_0)} \rho_t^N\right) - E_t^N(\varepsilon) \leq \fint_{Q_\varepsilon(x_0)} \rho_t^N \,\mathrm{d}\mathbf{m} \leq \left(\inf_{Q_\varepsilon(x_0)} \rho_t^N\right) + E_t^N(\varepsilon). \tag{4.38}$$

Taking into account that r_t^N is a probability density, it follows from the Hölder bound (4.35) that the familiy $(\rho_t^N)_{N\geq 1}$ is uniformly bounded in $L^\infty(\Omega,\mathbf{m})$. Hence, the Banach-Alaoglu Theorem yields the existence of a subsequential weak*-limit $\rho_t \in L^\infty(\Omega,\mathbf{m})$. Since $\mathbb{W}(\mu_t^N, \mu_t) \to 0$, we infer that $\mu_t = \rho_t \mathbf{m}$ and $\int_{Q_\varepsilon(x_0)} \rho_t^N \,\mathrm{d}\mathbf{m} \to \mu_t(Q_\varepsilon(x_0))$. Therefore, (4.38) yields

$$\left(\limsup_{N\to\infty} \sup_{Q_\varepsilon(x_0)} \rho_t^N\right) - C_t\Big(\varepsilon\sqrt{d}\Big)^\lambda \leq \frac{\mu_t(Q_\varepsilon(x_0))}{\mathbf{m}(Q_\varepsilon(x_0))} \leq \left(\liminf_{N\to\infty} \inf_{Q_\varepsilon(x_0)} \rho_t^N\right) + C_t\Big(\varepsilon\sqrt{d}\Big)^\lambda.$$

Passing to the limit $\varepsilon \to 0$ we obtain

$$\lim_{\varepsilon\to 0}\liminf_{N\to\infty}\sup_{x\in Q_\varepsilon(x_0)}\rho_t^N(x) \le \rho_t(x_0) \le \lim_{\varepsilon\to 0}\limsup_{N\to\infty}\inf_{x\in Q_\varepsilon(x_0)}\rho_t^N(x),$$

which is the desired result (pc).

Therefore, we can apply Theorem 4.3.3(iii) to obtain the desired inequality

$$\liminf_{N\to\infty}\mathcal{A}_N^*(m_t^{N,\alpha}, e_N) \ge \mathbb{A}^*(\mu_t^\alpha, \eta),$$

which implies that condition (ii) of Proposition 4.4.5 is satisfied.

Step 3. *Weak convergence of the time derivatives.*

In order to apply Proposition 4.4.5 we will now show that the sequence of time derivatives $\dot{m}^N$ is weakly convergent in $L^2\big((\delta,T); L^2(\Omega,\mathbf{m})\big)$.

Indeed, by self-adjointness of the discrete generator $\mathcal{L}_N$ in $L^2(\mathcal{T}_N,\pi_N)$ we have

$$\|\dot{r}_t^N\|_{L^2(\mathcal{T}_N,\pi_N)} = \|\mathcal{L}_N r_t^N\|_{L^2(\mathcal{T}_N,\pi_N)} \le (t-\delta/2)^{-1}\|r_{\delta/2}^N\|_{L^2(\mathcal{T}_N,\pi_N)}$$

for any $t > \delta/2$; see, e.g., [Bre10, Theorem 7.7]. Moreover, from (4.36) we infer that

$$\|r_t^N\|_{L^\infty(\mathcal{T}_N,\pi_N)} \lesssim 1 \vee t^{-\frac{d}{2}}$$

for $t > 0$. As $\delta < 1$, it follows from these bounds that

$$\int_\delta^T \|\rho_t^N\|_{L^2(\Omega,\mathbf{m})}^2 \,\mathrm{d}t \lesssim T\delta^{-d} \quad \text{and} \quad \int_\delta^T \|\dot{\rho}_t^N\|_{L^2(\Omega,\mathbf{m})}^2 \,\mathrm{d}t \lesssim T\delta^{-(d+1)}.$$

The Banach-Alaoglu theorem implies that any subsequence of $(\rho^N)_N$ has a subsequence converging weakly in $H^1\big((\delta,T); L^2(\Omega,\mathbf{m})\big)$. Since $\mathbb{W}(\mu_t^N,\mu_t) \to 0$, we infer that $\rho^N \rightharpoonup \rho$ in $H^1\big((\delta,T); L^2(\Omega,\mathbf{m})\big)$, and $\rho_t = \frac{\mathrm{d}\mu_t}{\mathrm{d}\mathfrak{m}}$, as desired.

Applying Proposition 4.4.5 with $\varphi_N(t) := \dot{\rho}_t^N$ and $\varphi(t) := \dot{\rho}_t$, we obtain

$$\int_\delta^T \mathbb{A}^*(\mu_t^\alpha, \dot{\mu}_t)\,\mathrm{d}t \le \liminf_{N\to\infty}\int_\delta^T \mathcal{A}_N^*\Big(m_t^{N,\alpha}, \dot{m}_t^N\Big)\,\mathrm{d}t.$$

Step 4. *Removal of the regularisation.*

Using the weak convergence $\mu_t^\alpha \rightharpoonup \mu_t$ as $\alpha \to 0$ and the weak lower-semicontinuity of $\mathbb{A}^*(\cdot,\dot{\mu}_t)$, an application of Fatou's lemma yields

$$\int_\delta^T \mathbb{A}^*(\mu_t, \dot{\mu}_t)\,\mathrm{d}t \le \liminf_{\alpha\to 0}\liminf_{N\to\infty}\int_\delta^T \mathcal{A}_N^*\Big(m_t^{N,\alpha}, \dot{m}_t^N\Big)\,\mathrm{d}t.$$

By convexity, we obtain

$$\mathcal{A}_N^*\Big(m_t^{N,\alpha}, \dot{m}_t^N\Big) \le (1-\alpha)\mathcal{A}_N^*\Big(m_t^N, \dot{m}_t^N\Big) + \alpha\mathcal{A}_N^*\Big(\pi_N, \dot{m}_t^N\Big).$$

We claim that $A := \sup_N \sup_{t\geq\delta} \mathcal{A}_N^*\big(\pi_N, \dot{m}_t^N\big) < \infty$. Indeed, in view of the self-adjointness of the discrete generator $\mathcal{L}_N$ and the ultracontractivity bound (4.36), we infer that

$$\mathcal{A}_N^*(\pi_N, \dot{m}_t^N) = \mathcal{A}(\pi_N, r_t^N) = \mathcal{E}_N(r_t^N) \leq t^{-1} \|r_t^N\|^2_{L^2(\mathcal{T}_N, \pi_N)} \leq Ct^{-1}\big(1 \vee t^{-d}\big),$$

which yields the claim. Consequently, we obtain

$$\int_\delta^T \mathbb{A}^*(\mu_t, \dot{\mu}_t)\,\mathrm{d}t \leq \liminf_{N\to\infty} \int_\delta^T \mathcal{A}_N^*\Big(m_t^N, \dot{m}_t^N\Big)\,\mathrm{d}t.$$

The final result follows by passing to the limit $\delta \to 0$.

(iii): This follows immediately by combining the inequalities from *(ii)*. □

4.5 Mosco convergence of discrete energies: proof strategy

In this section we give a sketch of the proof of the Mosco convergence of the discrete energy functionals (Theorem 4.3.9). This result is a key tool in the proof of evolutionary Γ-convergence; cf. Section 4.4. Let us first recall the relevant definitions.

Definition 4.5.1 (Γ- and Mosco convergence)**.** Let $\mathcal{F}, \mathcal{F}_N : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be functionals defined on a complete metric space $\mathcal{X}$. The sequence $(\mathcal{F}_N)_N$ is said to be Γ-convergent to $\mathcal{F}$ if the following conditions hold:

(i) For every sequence $(x_N)_N \subseteq \mathcal{X}$ such that $x_N \to x \in \mathcal{X}$ we have the *liminf inequality*

$$\liminf_{N\to\infty} \mathcal{F}_N(x_N) \geq \mathcal{F}(x). \tag{4.39}$$

(ii) For every $\bar{x} \in \mathcal{X}$ there exists a *recovery sequence* $(\bar{x}_N)_N \subseteq \mathcal{X}$, i.e., $x_N \to x$ and

$$\limsup_{N\to\infty} \mathcal{F}_N(x_N) \leq \mathcal{F}(x). \tag{4.40}$$

If $\mathcal{X}$ is a complete topological vector space, we say that $(\mathcal{F}_N)_N$ is Mosco convergent to $\mathcal{F}$ if the same conditions hold, with the modification that *weakly* convergent sequences are considered in the liminf inequality.

We use the notation $\mathcal{F}_N \xrightarrow{\Gamma} \mathcal{F}$ and $\mathcal{F}_N \xrightarrow{M} \mathcal{F}$ to denote Γ- and Mosco convergence.

Let us now fix the setup, which remains in force throughout Sections 4.5, 4.6, and 4.7. Consider a family of ζ-regular meshes $(\mathcal{T}_N)_N$ with $[\mathcal{T}_N] \to 0$ as $N \to \infty$. We then consider a measure $\mu \in \mathcal{P}(\overline{\Omega})$, and let $v \in L^1(\Omega)$ be its density with respect to the Lebesgue measure. At the discrete level we consider measures $m_N \in \mathcal{P}(\mathcal{T}_N)$. We define the corresponding energy functionals $\mathcal{F}_N$, $\tilde{\mathbb{F}}_N$, and $\mathbb{F}_\mu$ as in Section 4.3. The goal is to prove the Mosco convergence in $L^2(\Omega)$ of $\tilde{\mathbb{F}}_N$ to $\mathbb{F}_\mu$ as $N \to \infty$ under the assumptions (lb), (ub), and (pc).

Our strategy is based on a compactness and representation procedure, following ideas from [AC04]. A key ingredient in the proof is a representation result from [BFLM02, Theorem 2], for which we need to perform a localisation procedure. Let $\mathcal{O}(\Omega)$ be the collection of all open subsets of Ω. For $A \in \mathcal{O}(\Omega)$ we then introduce the functionals $\mathcal{F}_{\mathcal{T}} : L^2(\mathcal{T}, \pi_{\mathcal{T}}) \times \mathcal{O}(\Omega) \to [0, +\infty)$ by

$$\mathcal{F}_{\mathcal{T}}(f, A) := \frac{1}{4} \sum_{K,L \in \mathcal{T}|_A} \big(f(K) - f(L)\big)^2 U_{KL} \frac{|\Gamma_{KL}|}{d_{KL}},$$

where, for any subset $A \subseteq \Omega$,

$$\mathcal{T}|_A := \Big\{K \in \mathcal{T} \ : \ \overline{K} \cap A \neq \emptyset\Big\} \tag{4.41}$$

and U_{KL} is as in Section 4.3. The corresponding embedded functional $\tilde{\mathbb{F}}_{\mathcal{T}} : L^2(\Omega) \times \mathcal{O}(\Omega) \to [0, +\infty]$ is given by

$$\tilde{\mathbb{F}}_{\mathcal{T}}(\varphi, A) := \begin{cases} \mathcal{F}_{\mathcal{T}}(P_{\mathcal{T}}\varphi, A) & \text{if } \varphi \in \mathsf{PC}_{\mathcal{T}} \\ +\infty & \text{otherwise,} \end{cases}$$

where $P_{\mathcal{T}}$ is the projection defined in (4.25).

The proof of Theorem 4.3.9 consists of the following steps:

(Step 1) We show first, as in [AC04, Proposition 3.4], that any subsequential Γ-limit point $\mathbb{F}(\cdot, A)$ of the sequence $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ is only finite on $H^1(\Omega)$. This result is a prerequisite for performing Step 3. We also show that Γ-convergence implies Mosco convergence in this situation.

(Step 2) For any subsequential Γ-limit point $\mathbb{F}(\cdot, A)$, we prove an inner regularity result. Using this result, we can apply a compactness result [BD98, Theorem 10.3] to infer that there exists a subsequence, such that, for any $A \in \mathcal{O}(\Omega)$, the functionals $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ Γ-converge to a limiting functional $\mathbb{F}(\cdot, A)$.

(Step 3) We prove the applicability of a representation theorem [BFLM02, Theorem 2], which allows us to deduce the following expression

$$\mathbb{F}(\varphi) = \int_\Omega F(x, \varphi, \nabla\varphi)\, \mathrm{d}x. \tag{4.42}$$

(Step 4) In view of the previous steps, it remains to show that $F(x, u, \xi) = v(x)|\xi|^2$.

Steps 1 and 2 will be carried out in Section 4.6, while Steps 3 and 4 will be performed in Section 4.7.

4.6 Mosco convergence of the localised functionals

In this section we perform Steps 1 and 2 of the proof strategy described above. As before, we consider a vanishing sequence of ζ-regular meshes $(\mathcal{T}_N)_N$ and a sequence of discrete measures $m_N \in \mathcal{P}(\mathcal{T}_N)$. We will prove the following results.

Theorem 4.6.1 (Regularity of Γ-limits)**.** *Assume (Ib). For $A \in \mathcal{O}(\Omega)$, let $\mathbb{F}(\cdot, A)$ be a subsequential Γ-limit of the sequence $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ in the $L^2(\Omega)$-topology. Then $\mathbb{F}(\varphi, A) = +\infty$ for any $\varphi \notin H^1(\Omega)$. Moreover, the subsequence is also convergent in the Mosco sense.*

The proof of this result is contained in Section 4.6.1 and relies on an L^2-Hölder continuity result (Proposition 4.6.5).

Theorem 4.6.2 (Localised Mosco compactness)**.** *Assume (lb) and (ub). There exists a subsequence of $(\tilde{\mathbb{F}}_N)_N$ such that, for any $A \in \mathcal{O}(\Omega)$, the sequence $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ is Mosco convergent in $L^2(\Omega)$-topology.*

The proof of this result is contained in Section 4.6.2 and relies on an inner regularity result (Proposition 4.6.8). The latter result will be proved using a Sobolev upper bound (Proposition 4.6.6).

4.6.1 Regularity of finite energy sequences

In this subsection we prove that any subsequential Γ-limit $\mathbb{F}$ of the sequence $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ is only finite on Sobolev maps, which allows us to work with Theorem 4.7.3. A corresponding result was proved on the cartesian grid in [AC04, Proposition 3.4], using affine interpolations of vector fields that are not available in the present context.

For $h \in \mathbb{R}^d$ we write $K \overset{h}{\sim} L$ if $\overline{K} \cap (\overline{L} + h) \neq \emptyset$.

Lemma 4.6.3 (Existence of good paths)**.** *Let $\mathcal{T}$ be a ζ-regular mesh. Then there exists a family of paths $\{\gamma_{KL}\}_{K,L\in\mathcal{T}}$, where*

$$\gamma_{KL} = \{\gamma_{KL}(i) \ : \ i = 0, \dots, n_{KL}\}, \quad K = \gamma_{KL}(0) \sim \gamma_{KL}(1) \sim \dots \sim \gamma_{KL}(n_{KL}) = L,$$

such that the following properties hold:

1. *For all $K, L \in \mathcal{T}$ we have*

$$n_{KL} \lesssim \frac{|x_K - x_L|}{[\mathcal{T}]} \quad \textit{and} \quad \sum_{i=0}^{n_{KL}} |x_{\gamma_{KL}(i)} - x_{\gamma_{KL}(i+1)}| \lesssim |x_K - x_L|; \tag{4.43}$$

2. *For any $h \in \mathbb{R}^d$ and $M, N \in \mathcal{T}$ with $M \sim N$ we have*

$$\#\Big\{(K, L) \in \mathcal{T}^2 \ : \ K \overset{h}{\sim} L, \ \{M, N\} \subset \gamma_{KL}\Big\} \lesssim 1 \vee \frac{|h|}{[\mathcal{T}]}. \tag{4.44}$$

The implied constants depend only on Ω and ζ.

Proof. Part (1) has been proved in [GKM20, Lemma 2.12], so we focus on (2).

Fix $h \in \mathbb{R}^d$ and $M, N \in \mathcal{T}$ with $M \sim N$. Without loss of generality we may assume that $x_M = 0$ and h is parallel to the d-th unit vector in $\mathbb{R}^d$. Let $\mathcal{S}$ be the set whose cardinality we would like to bound, and let $\mathcal{S}_1$ be the collection of all $K \in \mathcal{T}$ such that $(K, L) \in \mathcal{S}$ for some $L \in \mathcal{T}$.

We claim that

$$\bigcup_{K\in\mathcal{S}_1} K \subset \mathsf{Cyl}(r, \ell) \tag{4.45}$$

for some $r \lesssim [\mathcal{T}]$ and $\ell \lesssim |h| + [\mathcal{T}]$. Here, $\mathsf{Cyl}(r,\ell)$ denotes the cylinder of radius $r > 0$ and height $2\ell > 0$, i.e.,

$$\mathsf{Cyl}(r,\ell) := \Big\{ v \in \mathbb{R}^d \ : \ v^* \in B_r^{d-1}, \ v_d \in [-\ell, \ell] \Big\}.$$

where B_r^{d-1} denotes the closed ball of radius r around the origin in $\mathbb{R}^{d-1}$.

Indeed, by the construction in [GKM20], $M \cup N$ is contained in the cylinder of radius $2[\mathcal{T}]$, whose central axis is obtained by extending the line segment between x_K and x_L by a distance $[\mathcal{T}]$ in both directions, for all cells $K, L \in \mathcal{T}$. The claim follows using the fact that $K \overset{h}{\sim} L$.

Next we use a simple volume comparison. Using ζ-regularity, it follows that

$$\mathscr{L}^d\Big(\bigcup_{K \in \mathcal{S}_1} K \Big) = \sum_{K \in \mathcal{S}_1} \mathscr{L}^d(K) \gtrsim [\mathcal{T}]^d (\#\mathcal{S}_1), \tag{4.46}$$

where $\#\mathcal{S}_1$ denotes the cardinality of $\mathcal{S}_1$. Combining (4.45) and (4.46) we infer that $\#\mathcal{S}_1 \lesssim 1 \vee \frac{|h|}{[\mathcal{T}]}$.

To conclude the proof, it remains to show that $\#\mathcal{S} \lesssim \#\mathcal{S}_1$. To see this, note that for every $K \in \mathcal{S}_1$, there exists a universally bounded number of cells $L \in \mathcal{T}$ such that $(K, L) \in \mathcal{S}$. This is due to the fact that if $L, L' \in \mathcal{T}$ are such that $(K, L), (K, L') \in \mathcal{S}$, we deduce that $d_{L,L'} \lesssim [\mathcal{T}]$ by the triangle inequality. The desired result follows from this observation by ζ-regularity. □

The following lemma is the crucial estimate needed to deduce L^2-strong compactness of sequences with bounded energy. A similar result has been obtained in dimension $d = 2, 3$ in [EGH00, Lemma 3.3] with bounds in terms of discrete Sobolev norms.

Lemma 4.6.4 (L^2-Hölder continuity). *Assume* (lb)*. Fix $A \in \mathcal{O}(\Omega)$ and set $A_\delta := \{x \in A \ : \ \mathsf{dist}(x, \partial\Omega) > \delta\}$ for $\delta > 0$. Let $\mathcal{T}$ be a ζ-regular mesh, let $f \in L^2(\mathcal{T}|_A)$ and define $\varphi := \mathsf{Q}_\mathcal{T} f \in L^2(A)$. For any $h \in \mathbb{R}^d$ we have the L^2-bound*

$$\|\tau_h \varphi - \varphi\|^2_{L^2(A_{|h|})} \lesssim \frac{|h|}{\underline{k}} \Big(|h| \vee [\mathcal{T}] \Big) \mathcal{F}_\mathcal{T}(f, A), \tag{4.47}$$

where $\tau_h \varphi(\cdot) := \varphi(\cdot - h)$, and $\underline{k} > 0$ is the lower bound in (lb)*.*

Proof. For any $h \in \mathbb{R}^d$ we have

$$\|\tau_h \varphi - \varphi\|^2_{L^2(A_{|h|})} = \int_{A_{|h|}} \Big(\varphi(x - h) - \varphi(x) \Big)^2 \, \mathrm{d}x \leq \sum_{K, L \in \mathcal{T}|_A} |C_{KL}| \Big(f(L) - f(K) \Big)^2, \tag{4.48}$$

where $C_{KL} = \{x \in K \, : \, x - h \in L\}$. For $K, L \in \mathcal{T}|_A$ we use Lemma 4.6.3 and the Cauchy-Schwarz inequality to write

$$\Big(f_N(K) - f_N(L) \Big)^2 \leq n_{KL} \sum_{i=1}^{n_{KL}} \Big(f_N(K_{i-1}) - f_N(K_i) \Big)^2, \tag{4.49}$$

where $K = K_0 \sim K_1 \sim \ldots \sim K_{n_{KL}} = L$, and $n_{KL} \lesssim \frac{d_{KL}}{[\mathcal{T}]}$. Observe that $d_{KL} \lesssim [\mathcal{T}] \vee |h|$ whenever $C_{KL} \neq \emptyset$.

To estimate the measure of C_{KL}, we pick a hyperplane H that separates K and L (which exists by the Hahn-Banach theorem, in view of the convexity of the cells). By construction, C_{KL} is contained in the strip between H and $H+h$. Moreover, we have $C_{KL} \subseteq K$, which means that C_{KL} is contained in a ball of radius $\lesssim [\mathcal{T}]$. Combining these two facts, we infer that $|C_{KL}| \lesssim [\mathcal{T}]^{d-1}|h|$, hence $|C_{KL}| \lesssim [\mathcal{T}]^{d-1}\big(|h| \wedge [\mathcal{T}]\big)$ by ζ-regularity.

Putting these estimates together, we obtain

$$|C_{KL}|\Big(f(K) - f(L)\Big)^2 \lesssim [\mathcal{T}]^{d-1}|h| \sum_{i=1}^{n_{KL}} \Big(f(K_{i-1}) - f(K_i)\Big)^2. \tag{4.50}$$

Let α_{KL} denote the left-hand side in (4.44). Using (4.48) and (4.50) we find that

$$\|\tau_h\varphi - \varphi\|^2_{L^2(A_{|h|})} \lesssim [\mathcal{T}]^{d-1}|h| \sum_{\substack{K,L\in\mathcal{T}|_A \\ L\sim K}} \alpha_{KL}\Big(f(L) - f(K)\Big)^2.$$

On the other hand, in view of the ζ-regularity and the assumption (lb), we have

$$\mathcal{F}_{\mathcal{T}}(f, A) \gtrsim \underline{k}[\mathcal{T}]^{d-2} \sum_{\substack{K,L\in\mathcal{T}|_A \\ L\sim K}} \Big(f(K) - f(L)\Big)^2.$$

The desired result follows, since $\alpha_{KL} \leq 1 \vee \frac{|h|}{[\mathcal{T}]}$ by Lemma 4.6.3. □

The compactness result now follows easily.

Proposition 4.6.5 (Compactness)**.** *Fix $A \in \mathcal{O}(\Omega)$ and assume that the lower bound (lb) holds. Let $(\mathcal{T}_N)_N$ be a vanishing sequence of ζ-regular meshes. Let $f_N \in L^2(\mathcal{T}_N|_A)$ be such that*

$$\alpha := \sup_{N\in\mathbb{N}} \mathcal{F}_N(f_N, A) < +\infty,$$

and define $\varphi_N := \mathsf{Q}_N f_N \in L^2(A)$. Then the sequence $(\varphi_N)_N$ is relatively compact in $L^2(A)$. Moreover, any subsequential limit φ belongs to $H^1(A)$ and satisfies

$$\|\nabla\varphi\|_{L^2(A)} \lesssim \sqrt{\frac{\alpha}{\underline{k}}}.$$

Proof. The L^2-compactness follows from (4.47) in view of the Kolmogorov-Riesz-Frechét theorem [Bre10, Theorem 4.26]. Let φ be any subsequential limit point of φ_N as $[\mathcal{T}_N] \to 0$. Another application of (4.47) yields, for any $h \in \mathbb{R}^d$ and $\delta > 0$,

$$\|\tau_h\varphi - \varphi\|^2_{L^2(A_\delta)} = \lim_{N\to\infty} \|\tau_h\varphi_N - \varphi_N\|^2_{L^2(A_\delta)} \lesssim \frac{\alpha}{\underline{k}}|h|^2,$$

which implies that $\varphi \in H^1(A)$ by the characterisation of $H^1(A)$ as the space of functions which are Lipschitz continuous in L^2-norm (see, e.g., [Bre10, Proposition 9.3]). □

Proof of Theorem 4.6.1. Proposition 4.6.5 shows that $\varphi \in H^1(\Omega)$ whenever $\mathbb{F}(\varphi) < \infty$. It also follows from Proposition 4.6.5 that every L^2-weakly convergent sequence $\varphi_N = \mathsf{Q}_N f_N$ with bounded energy $\sup_N \mathcal{F}_N(f_N, A) < +\infty$ converges strongly in L^2. Therefore, Mosco and Γ-convergence are equivalent in this situation. □

4.6.2 Sobolev bound and inner regularity

This part focuses on a Sobolev upper bound for subsequential Γ-limit functionals, which will be useful in Proposition 4.6.8 and in Theorem 4.7.3 below.

Proposition 4.6.6 (Sobolev upper bound). *Assume* (ub) *and let* $A \in \mathcal{O}(\Omega)$. *For any subsequential* Γ*-limit* $\mathbb{F}(\cdot, A)$ *of the sequence* $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ *in the* $L^2(\Omega)$*-topology, we have the Sobolev upper bound*

$$\mathbb{F}(\varphi, A) \lesssim \bar{k} \int_A |\nabla\varphi|^2 \, \mathrm{d}x \tag{4.51}$$

for any $\varphi \in H^1(\Omega)$.

Here and in the proof, the implied constants depend only on Ω and the regularity parameter ζ.

Proof. Let us first prove (4.51) for $\varphi \in C_c^\infty(\mathbb{R}^d)$. For $N \in \mathbb{N}$, define $f_N : \mathcal{T}_N \to \mathbb{R}$ by $K \in \mathcal{T}_N$, define

$$f_N(K) := \varphi(x_K) \quad \text{for } K \in \mathcal{T}_N.$$

Write $\nu_{KL} := \frac{x_K - x_L}{d_{KL}}$. By smoothness of φ and σ, we have

$$\varepsilon_N := \sup_{K,L\in\mathcal{T}_N} \left| \left(\frac{f_N(K) - f_N(L)}{d_{KL}} \right)^2 - \big(\nabla\varphi(x_K)\cdot\nu_{KL}\big)^2 \right| \to 0.$$

Using this estimate, assumption (ub), and the ζ-regularity, we obtain

$$\begin{aligned}
\tilde{\mathbb{F}}_N(Q_N f_N, A) &= \frac{1}{4} \sum_{K,L\in\mathcal{T}_N|_A} \left(\frac{f_N(K) - f_N(L)}{d_{KL}} \right)^2 U_{KL} d_{KL} |\Gamma_{KL}| \\
&\lesssim \bar{k} \sum_{K\in\mathcal{T}_N|_A} \Big(\big(\nabla\varphi(x_K)\cdot\nu_{KL}\big)^2 + \varepsilon_N \Big) \Big(\sum_{L:L\sim K} d_{KL} |\Gamma_{KL} \Big) \\
&\lesssim \bar{k} \sum_{K\in\mathcal{T}_N|_A} \Big(|\nabla\varphi(x_K)|^2 + \varepsilon_N \Big) |K|.
\end{aligned}$$

The smoothness of the function $|\nabla\varphi|^2$ and the identity $\sum_{K\in\mathcal{T}_N} |K| = |\Omega|$ now yield

$$\limsup_{N\to\infty} \tilde{\mathbb{F}}_N(Q_N f_N, A) \lesssim \bar{k} \int_A |\nabla\varphi|^2 \, \mathrm{d}x.$$

Since $Q_N f_N$ converges to φ in $L^2(A)$, the Γ-convergence of $\tilde{\mathbb{F}}_N(\cdot, A)$ to $\mathbb{F}(\cdot, A)$ yields the desired bound (4.51).

It remains to extend the result to $H^1(\Omega)$ by a density argument. Indeed, for any $\varphi \in H^1(\Omega)$ there exists a sequence $(\varphi_i)_i \subseteq C_c^\infty(\mathbb{R}^d)$ such that $\varphi^i \to \varphi$ in $H^1(\Omega)$. As $\mathbb{F}(\cdot, A)$ is lower semicontinuous in $L^2(\Omega)$, we can apply (4.51) to φ_i to obtain

$$\mathbb{F}(\varphi, A) \le \liminf_{i\to\infty} \mathbb{F}(\varphi_i, A) \lesssim \bar{k} \liminf_{i\to\infty} \int_A |\nabla\varphi_i|^2 \, \mathrm{d}x = \bar{k} \int_A |\nabla\varphi|^2 \, \mathrm{d}x,$$

which shows (4.51) for $\varphi \in H^1(\Omega)$. □

Remark 4.6.7. In the case where $m_N = \mathsf{P}_N(\rho\,\mathrm{d}x)$ for a continuous density ρ, it is possible to prove the sharp upper bound $\mathbb{F} \le \mathbb{F}_\mu$ by a similar argument with a bit more effort. However, we are not aware of a simple argument for the corresponding liminf inequality. Therefore, we pass through the compactness and representation scheme, which yields the sharp upper bound as a byproduct.

We now focus on the inner regularity of subsequential Γ-limit functionals. We will prove something slightly stronger than the classical inner regularity, namely, an inner approximation result with sets of Lebesgue measure 0. This sharpening will be useful in the proof of the locality in Proposition 4.7.5 below.

For any $A, B \subset \Omega$, we write $A \Subset B$ as a shorthand for A being relatively compact in B.

Proposition 4.6.8 (Inner regularity)**.** *Assume* (ub)*. For $A \in \mathcal{O}(\Omega)$, let $\mathbb{F}(\cdot, A)$ be a subsequential Γ-limit of the sequence $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ in the $L^2(\Omega)$-topology. Then the function $A \mapsto \mathbb{F}(\varphi, A)$ is inner regular on $\mathcal{O}(\Omega)$, i.e.,*

$$\sup_{\substack{A' \Subset A \\ \mathscr{L}^d(\partial A')=0}} \mathbb{F}(\varphi, A') = \sup_{A' \Subset A} \mathbb{F}(\varphi, A') = \mathbb{F}(\varphi, A). \tag{4.52}$$

for any $\varphi \in H^1(\Omega)$ and $A \in \mathcal{O}(\Omega)$.

Proof. Fix $\varphi \in H^1(\Omega)$ and $A \in \mathcal{O}(\Omega)$. It immediately follows from the definitions that (4.52) holds with "$\le$" (twice) instead of "$=$". It thus suffices to prove that

$$\mathbb{F}(\varphi, A) \le \sup_{\substack{A' \Subset A \\ \mathscr{L}^d(\partial A')=0}} \mathbb{F}(\varphi, A').$$

We adapt the proof for the cartesian grid as given in [AC04, Proposition 3.9].

Fix $\delta > 0$ and consider a non-empty set $A'' \in \mathcal{O}(\Omega)$ such that $A'' \Subset A$ and

$$\int_{A\setminus\overline{A''}} |\nabla\varphi|^2 \,\mathrm{d}x < \delta.$$

Let $\varepsilon_N := \mathsf{Q}_N e_N$ be a recovery sequence for $\mathbb{F}(\varphi, A \setminus \overline{A''})$, i.e.,

$$\varepsilon_N \to \varphi \text{ in } L^2(\Omega) \quad \text{and} \quad \limsup_{N\to\infty} \mathcal{F}_N(e_N, A \setminus \overline{A''}) \le \mathbb{F}(\varphi, A \setminus \overline{A''}) \lesssim \bar{k}\delta, \tag{4.53}$$

where the last bound is a consequence of Proposition 4.6.6.

Take $A' \in \mathcal{O}(\Omega)$ such that $A'' \Subset A' \Subset A$ and $\mathscr{L}^d(\partial A') = 0$. Note that this can always be done, since one can pick a compact set K satisfying $A'' \subset K \Subset A$, and then choose A' as the union of any finite open cover of K by balls whose closures are contained in A. Let $\varphi_N := Q_N f_N$ be a recovery sequence for $\mathbb{F}(\varphi, A')$, so that

$$\varphi_N \to \varphi \text{ in } L^2(\Omega) \quad \text{and} \quad \limsup_{N\to\infty} \mathcal{F}_N(f_N, A') \le \mathbb{F}(\varphi, A'). \tag{4.54}$$

Fix $M \in \mathbb{N}$ and suppose that $[\mathcal{T}_N] < \frac{1}{5(M+1)}$. Define $A'' \subset A_1 \subset A_2 \subset \ldots \subset A_{5(M+1)} \subset A'$ by

$$A_j := \left\{ x \in A' \ : \ d(x, A'') < \frac{j}{5(M+1)} d\Big((A')^c, A''\Big) \right\}.$$

Moreover, for $i \in \{1, \dots, M\}$ we consider a cutoff function $\rho_i \in C^\infty(\mathbb{R}^d)$ satisfying

$$\rho_i|_{A_{5i+2}} = 1, \quad \rho_i|_{\Omega \setminus A_{5i+3}} = 0, \quad 0 \le \rho_i \le 1, \quad |\nabla \rho_i| \lesssim M. \tag{4.55}$$

Set $r_N^i(K) := \rho_i(x_K)$ for $K \in \mathcal{T}_N$, and define

$$f_N^i := r_N^i f_N + (1 - r_N^i) e_N, \quad \text{so that } \varphi_N^i := \mathsf{Q}_N f_N^i \to \varphi$$

as $N \to \infty$, uniformly for $i \in \{1, \dots, M\}$. As $[\mathcal{T}_N] < \frac{1}{5(M+1)}$, we have by (4.55),

$$f_N^i \equiv f_N \text{ on } \mathcal{T}_N|_{A_{5i+1}}, \quad f_N^i \equiv e_N \text{ on } \mathcal{T}_N|_{(A \setminus \overline{A_{5i+4}})}. \tag{4.56}$$

Using these identities and the inclusions $A_{5i+1} \subset A'$ and $A'' \subset A_{5i+4}$ we obtain

$$\begin{aligned}\mathcal{F}_N(f_N^i, A) &\le \mathcal{F}_N(f_N^i, A_{5i+1}) + \mathcal{F}_N(f_N^i, A_{5(i+1)} \setminus \overline{A_{5i}}) + \mathcal{F}_N(f_N^i, A \setminus \overline{A_{5i+4}}) \\ &\le \mathcal{F}_N(f_N, A') + \mathcal{F}_N(f_N^i, A_{5(i+1)} \setminus \overline{A_{5i}}) + \mathcal{F}_N(e_N, A \setminus \overline{A''}).\end{aligned} \tag{4.57}$$

To estimate the middle term, let $\nabla g(K, L) := g(L) - g(K)$ denote the discrete derivative and observe that

$$\begin{aligned}\nabla f_N^i(K, L) = r_N^i(L) \nabla f_N(K, L) &+ \big(1 - r_N^i(L)\big) \nabla e_N(K, L) \\ &+ \big(f_N(K) - e_N(K)\big) \nabla r_N^i(K, L)\end{aligned}$$

for any $K, L \in \mathcal{T}_N$. Consequently,

$$|\nabla f_N^i(K, L)|^2 \lesssim |\nabla f_N(K, L)|^2 + |\nabla e_N(K, L)|^2 + M^2 d_{KL}^2 |f_N(K) - e_N(K)|^2.$$

Using this bound and the ζ-regularity of the mesh, we obtain

$$\begin{aligned}&\sum_{i=1}^M \mathcal{F}_N(f_N^i, A_{5(i+1)} \setminus \overline{A_{5i}}) \\ &\lesssim \sum_{i=1}^M \Big(\mathcal{F}_N(f_N, A_{5(i+1)} \setminus \overline{A_{5i}}) + \mathcal{F}_N(e_N, A_{5(i+1)} \setminus \overline{A_{5i}}) + \bar{k} M^2 \|\varphi_N - \varepsilon_N\|_{L^2(\Omega)}^2 \Big) \\ &\le 2 \Big(\mathcal{F}_N(f_N, A' \setminus \overline{A''}) + \mathcal{F}_N(e_N, A' \setminus \overline{A''}) \Big) + \bar{k} M^3 \|\varphi_N - \varepsilon_N\|_{L^2(\Omega)}^2.\end{aligned}$$

Taking into account that that $\varphi_N, \varepsilon_N \to \varphi$ in L^2, we can pass to the limsup as $N \to \infty$, using (4.53), (4.54), and Proposition 4.6.6, to obtain

$$\begin{aligned}\limsup_{N \to \infty} \sum_{i=1}^M \mathcal{F}_N(f_N^i, A_{5(i+1)} \setminus \overline{A_{5i}}) &\lesssim \limsup_{N \to \infty} \mathcal{F}_N(f_N, A') + \limsup_{N \to \infty} \mathcal{F}_N(e_N, A \setminus \overline{A''}) \\ &\lesssim \mathbb{F}(\varphi, A') + \mathbb{F}(\varphi, A \setminus \overline{A''}) \\ &\lesssim \bar{k} \int_A |\nabla \varphi|^2 \,\mathrm{d}x.\end{aligned}$$

Using this bound and (4.53), (4.54) once more, it follows from (4.57) that

$$\limsup_{N \to \infty} \bigg(\frac{1}{M} \sum_{i=1}^M \mathcal{F}_N(f_N^i, A) \bigg) \le \mathbb{F}(\varphi, A') + C\bar{k} \bigg(\frac{1}{M} \int_A |\nabla \varphi|^2 \,\mathrm{d}x + \delta \bigg).$$

where $C < \infty$ depends only on Ω, ζ.

Clearly, for each N, there exists $i_N \in \{1, \dots, M\}$ such that

$$\mathcal{F}_N(f_N^{i_N}, A) \leq \frac{1}{M} \sum_{i=1}^{M} \mathcal{F}_N(f_N^i, A),$$

Since $\sup_{1 \leq i \leq M} \|\varphi_N^i - \varphi\|_{L^2(\Omega)} \to 0$ as $N \to \infty$, we have $\varphi_N^{i_N} \to \varphi$ in $L^2(\Omega)$. Therefore, using the Γ-convergence we obtain

$$\mathbb{F}(\varphi, A) \leq \liminf_{N \to \infty} \mathcal{F}_N(f_N^{i_N}, A) \leq \mathbb{F}(\varphi, A') + C\overline{k}\biggl(\frac{1}{M}\int_A |\nabla\varphi|^2 \,\mathrm{d}x + \delta\biggr).$$

As $\delta > 0$ and $M < \infty$ are arbitrary, this is the desired result. □

Proof of Theorem 4.6.2. By Proposition 4.6.8 and [BD98, Theorem 10.3], there exists a subsequence such that, for any $A \in \mathcal{O}(\Omega)$, the functionals $\bigl(\tilde{\mathbb{F}}_N(\cdot, A)\bigr)_N$ are Γ-converging in $L^2(\Omega)$-topology to a limit functional $\mathbb{F}(\cdot, A)$. The fact that Γ-convergence implies Mosco convergence has already been observed in Theorem 4.6.1. □

4.7 Representation and characterisation of the limit

We fix the same setup as in Section 4.6. We thus consider a vanishing sequence of ζ-regular meshes $(\mathcal{T}_N)_N$ and a sequence of discrete measures $m_N \in \mathcal{P}(\mathcal{T}_N)$.

We show the following representation formula for the Γ-limits from Section 4.6:

Theorem 4.7.1 (Representation of the Γ-limit). *Assume* (lb) *and* (ub)*, and suppose that, for every $A \in \mathcal{O}(\Omega)$, the functionals $\bigl(\tilde{\mathbb{F}}_N(\cdot, A)\bigr)_N$ are $L^2(\Omega)$-Mosco convergent to a functional $\mathbb{F}(\cdot, A)$. Then the functional $\mathbb{F}$ can be represented as*

$$\mathbb{F}(\varphi, A) = \begin{cases} \displaystyle\int_A F(x, \varphi, \nabla\varphi) \,\mathrm{d}x & \text{for } \varphi \in H^1(\Omega), \\ +\infty & \text{for } \varphi \in L^2(\Omega) \setminus H^1(\Omega), \end{cases} \tag{4.58}$$

for some measurable function $F : \Omega \times \mathbb{R} \times \mathbb{R}^d \to [0, +\infty)$.

Combined with the following result, this will complete the proof of Theorem 4.3.9.

Theorem 4.7.2 (Characterisation of F). *Assume* (lb)*,* (ub)*, and* (pc)*. Then the function $F : \Omega \times \mathbb{R}^d \to [0, +\infty)$ defined in Theorem 4.7.1 is given by*

$$F(x, u, \xi) = |\xi|^2 \upsilon(x) \quad \forall x \in \Omega,\ u \in \mathbb{R},\ \xi \in \mathbb{R}^d.$$

In particular, the sequence $\bigl(\tilde{\mathbb{F}}_N(\cdot, A)\bigr)_N$ is $L^2(\Omega)$-Mosco convergent to $\mathbb{F}_\mu(\cdot, A)$.

To prove Theorem 4.7.1, we use a representation result for functionals on Sobolev spaces [BFLM02]. In our application, we have $\mathbb{E}(\cdot, A) := \mathbb{F}(\cdot, A)$, where $\mathbb{F}(\cdot, A)$ is a subsequential Γ-limit point of $\bigl(\tilde{\mathbb{F}}_N(\cdot, A)\bigr)_N$.

Theorem 4.7.3. *Let $\mathbb{E} : H^1(\Omega) \times \mathcal{O}(\Omega) \to [0, +\infty]$ be a functional satisfying the following conditions:*

(i) locality*: $\mathbb{E}$ is local, i.e., for all $A \in \mathcal{O}(\Omega)$ we have $\mathbb{E}(\varphi, A) = \mathbb{E}(\psi, A)$ if $\varphi = \psi$ a.e. on A.*

(ii) measure property*: For every $\varphi \in H^1(\Omega)$ the set map $\mathbb{E}(\varphi, \cdot)$ is the restriction of a Borel measure to $\mathcal{O}(\Omega)$.*

(iii) Sobolev bound*: There exists a constant $c > 0$ and $a \in L^1(\Omega)$ such that*

$$\frac{1}{c} \int_A |\nabla\varphi|^2 \,\mathrm{d}x \leq \mathbb{E}(\varphi, A) \leq c \int_A \left(a(x) + |\nabla\varphi|^2\right) \mathrm{d}x$$

for all $\varphi \in H^1(\Omega)$ and $A \in \mathcal{O}(\Omega)$.

(iv) lower semicontinuity*: $\mathbb{E}(\cdot, A)$ is weakly sequentially lower semicontinuous in $H^1(\Omega)$.*

Then $\mathbb{E}$ can be represented in integral form

$$\mathbb{E}(\varphi, A) = \int_A f(x, \varphi, \nabla\varphi) \,\mathrm{d}x,$$

where the measurable function $f : \Omega \times \mathbb{R} \times \mathbb{R}^d \to [0, +\infty)$ satisfies the self-consistent formula

$$f(x, u, \xi) := \limsup_{\varepsilon \to 0^+} \frac{M\Big(u + \xi(\cdot - x), Q_\varepsilon(x)\Big)}{\varepsilon^d}, \tag{4.59}$$

where $Q_\varepsilon(x)$ is the open cube of side-length $\varepsilon > 0$ centred at x, and

$$M(\psi, A) := \inf\Big\{\mathbb{E}(\varphi, A) \ : \ \varphi \in H^1(\Omega), \ \varphi - \psi \in H_0^1(A)\Big\} \tag{4.60}$$

for any $\psi \in H^1(\Omega)$ and any open cube $A \subseteq \Omega$.

Remark 4.7.4 (Equivalence of definitions)*.* The paper [BFLM02] contains the statement of Theorem 4.7.3 with $M(\psi, A)$ replaced by

$$\bar{M}(\psi, A) := \inf\Big\{\mathbb{E}(\varphi, A) \ : \ \varphi \in H^1(\Omega), \ \varphi = \psi \text{ in a neighbourhood of } A\Big\}.$$

We claim that $M = \bar{M}$. As any competitor φ for M is a competitor for $\bar{M}$, it is clear that $M \geq \bar{M}$. To show the opposite inequality, we fix $\varepsilon > 0$ and take $\varphi \in H^1(A)$ such that $\mathbb{E}(\varphi, A) \leq \bar{M}(\psi, A) + \varepsilon$. It follows that $\varphi - \psi \in H_0^1(A)$, and there exists a sequence $(\eta_n)_n \subseteq C_c^\infty(A)$ such that $\eta_n \to \varphi - \psi$ in $H^1(\Omega)$ as $n \to \infty$. Set $\varphi_n := \psi + \eta_n$, so that $\varphi_n \to \varphi$ in $H^1(\Omega)$. Note that φ_n is a competitor for $M(\psi, A)$, as it coincides with ψ on $A \setminus \mathrm{spt}(\eta_n)$, hence $M(\psi, A) \leq \mathbb{E}(\varphi_n, A)$ for all $n \in \mathbb{N}$. Using continuity of $\mathbb{E}(\cdot, A)$ with respect to the strong $H^1(\Omega)$ convergence (as follows from (iii)), we may pass to the limit to obtain

$$M(\psi, A) \leq \lim_{n \to \infty} \mathbb{E}(\varphi_n, A) = \mathbb{E}(\varphi, A) \leq \bar{M}(\psi, A) + \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, the claim follows.

In the remainder of this section we will verify that the functional $\mathbb{F}$ from Theorem 4.6.2 satisfies the conditions of Theorem 4.7.3. In particular, we will prove the locality (Section 4.7.1) and the subadditivity (Section 4.7.2). The proof of Theorem 4.7.1 will be completed at the end of Section 4.7.2. The proof of Theorem 4.7.2 is contained in Section 4.7.3.

4.7.1 Locality

A consequence of the inner regularity result from Proposition 4.6.8 is a simple proof of the locality of $\mathbb{F}$. An analogous result appears in [AC04, Proposition 3.9] on the cartesian grid. The proof in our setting is much simpler due to the short range of interactions.

Proposition 4.7.5 (Locality). *Assume that (ub) holds. Suppose that $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ is $L^2(\Omega)$-Mosco convergent to some functional $\mathbb{F}(\cdot, A)$ for every $A \in \mathcal{O}(\Omega)$. Then $\mathbb{F}$ is local, i.e., for any $A \in \mathcal{O}(\Omega)$ and $\varphi, \psi \in L^2(\Omega)$ such that $\varphi = \psi$ a.e. on A, we have $\mathbb{F}(\varphi, A) = \mathbb{F}(\psi, A)$.*

Proof. Let $A \in \mathcal{O}(\Omega)$ and take $\varphi, \psi \in L^2(\Omega)$ such that $\varphi = \psi$ a.e. on A. In view of the inner regularity result from Proposition 4.6.8 we may assume that $\mathscr{L}^d(\partial A) = 0$. By symmetry, it suffices to prove that $\mathbb{F}(\varphi, A) \geq \mathbb{F}(\psi, A)$.

Define $C_N := \bigcup\{K : K \in \mathcal{T}_N|_A\}$ and $C := \bigcup_N C_N$, so that $C \supseteq A$. We claim that

$$C \setminus A \subseteq B_N, \quad \text{where } B_N := \Big\{x \in \Omega \ : \ d(x, \partial A) < 2[\mathcal{T}_N]\Big\}. \tag{4.61}$$

Indeed, for every $x \in C \setminus A$ there exists $N \geq 1$ and $K \in \mathcal{T}_N$ such that $x \in K \setminus A$ and $\overline{K} \cap A \neq \emptyset$. Therefore, $d(x, \partial A) = d(x, A) \leq \operatorname{diam}(K) \leq [\mathcal{T}_N]$, which implies (4.61).

Let $(\varphi_N)_N$ be a recovery sequence for $\mathbb{F}(\varphi, A)$, i.e., $\varphi_N \to \varphi$ in $L^2(\Omega)$ and

$$\lim_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N, A) = \mathbb{F}(\varphi, A). \tag{4.62}$$

Fix $\widehat{\psi}_N \in \mathsf{PC}_N$ such that $\widehat{\psi}_N \to \psi$ in $L^2(\Omega)$ as $N \to \infty$, and define $\psi_N : \Omega \to \mathbb{R}$ by

$$\psi_N(x) := \begin{cases} \varphi_N(x) & \text{if } x \in C, \\ \widehat{\psi}_N(x) & \text{if } x \in \Omega \setminus C. \end{cases}$$

We claim that $\psi_N \to \psi$ in $L^2(\Omega)$ as $N \to \infty$. Indeed, since $\varphi = \psi$ a.e. on A, we have

$$\|\psi_N - \psi\|^2_{L^2(\Omega)} = \|\widehat{\psi}_N - \psi\|^2_{L^2(\Omega\setminus C)} + \|\varphi_N - \psi\|^2_{L^2(C\setminus A)} + \|\varphi_N - \varphi\|^2_{L^2(A)}. \tag{4.63}$$

The first and the last term on the right-hand side vanish as $N \to \infty$, since $\varphi_N \to \varphi$ and $\widehat{\psi}_N \to \psi$ in $L^2(\Omega)$. On the other hand, (4.61) yields

$$\begin{aligned}\limsup_{N\to\infty} \|\varphi_N - \psi\|_{L^2(C\setminus A)} &\leq \limsup_{N\to\infty} \Big(\|\varphi\|_{L^2(B_N)} + \|\psi\|_{L^2(B_N)}\Big) \\ &= \|\varphi\|_{L^2(\partial A)} + \|\psi\|_{L^2(\partial A)} = 0,\end{aligned}$$

since $\mathscr{L}^d(\partial A) = 0$. Therefore, using (4.63) we infer that $\psi_N \to \psi$ in $L^2(\Omega)$ as $N \to \infty$. Using this fact, the Γ-convergence of $\tilde{\mathbb{F}}_N$ in L^2, the fact that $\varphi_N = \psi_N$ a.e. on C, and (4.62), we obtain

$$\mathbb{F}(\psi, A) \leq \limsup_{N\to\infty} \tilde{\mathbb{F}}_N(\psi_N, A) = \limsup_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N, A) = \mathbb{F}(\varphi, A),$$

which concludes the proof. □

4.7.2 Subadditivity

We now prove subadditivity of the functional $A \mapsto \mathbb{F}(\varphi, A)$ for any $\varphi \in H^1(\Omega)$. This is the first step towards the verification of (ii) in Theorem 4.7.3. The proof is inspired by [AC04, Proposition 3.7] and follows similar ideas as in the proof of Proposition 4.6.8.

Proposition 4.7.6 (Subadditivity)**.** *Assume* (ub). *Suppose that* $\big(\tilde{\mathbb{F}}_N(\cdot, A)\big)_N$ *is* $L^2(\Omega)$*-Mosco convergent to some functional* $\mathbb{F}(\cdot, A)$ *for every* $A \in \mathcal{O}(\Omega)$*. Then the functional* $\mathbb{F}(\varphi, \cdot)$ *is subadditive for any* $\varphi \in H^1(\Omega)$*, in the sense that*

$$\mathbb{F}(\varphi, A \cup B) \le \mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B) \qquad \textit{for all } A, B \in \mathcal{O}(\Omega). \tag{4.64}$$

Proof. Fix $A, B \in \mathcal{O}(\Omega)$. For all $A' \Subset A$, $B' \Subset B$, and $\varphi \in H^1(\Omega)$ we will prove that

$$\mathbb{F}(\varphi, A' \cup B') \le \mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B).$$

In view of the the inner regularity (Proposition 4.6.8), this implies (4.64).

Pick $A' \Subset A$ and $B' \Subset B$ and let $(\psi_N)_N$, $(\varphi_N)_N$ be recovery sequences for $\mathbb{F}(\varphi, A)$ and $\mathbb{F}(\varphi, B)$ respectively, which we can assume to be finite. Fix $M \in \mathbb{N}$ and suppose that $[\mathcal{T}_N] < \frac{1}{5(M+1)}$. We define the sets

$$A_j := \left\{ x \in A \ : \ d(x, A') < \frac{j}{5(M+1)} d(A', A^c) \right\} \subset A$$

for $j \in \{1, \dots, 5(M+1)\}$. Moreover, for $i \in \{1, ..., M\}$ let ρ_i be a cutoff function $\rho_i \in C^\infty(\mathbb{R}^d)$ satisfying

$$\rho_i|_{A_{5i+2}} = 1, \quad \rho_i|_{\Omega \setminus A_{5i+3}} = 0, \quad 0 \le \rho_i \le 1, \quad |\nabla \rho_i| \lesssim M.$$

We then consider the $L^2(\Omega)$-convergent sequences

$$\varphi_N^i := \mathsf{Q}_N P_N \Big(\rho_i \psi_N + (1 - \rho_i) \varphi_N \Big) \xrightarrow[N \to \infty]{} \varphi, \quad \forall i \in \{1, \dots, M\}.$$

By definition, we have $\varphi_N^i \equiv \psi_N$ in A_{5i+1} and $\varphi_N^i \equiv \varphi_N$ in $\Omega \setminus \overline{A_{5i+4}}$. Arguing as in the proof of Proposition 4.6.8, one deduces the bound

$$\tilde{\mathbb{F}}_N(\varphi_N^i, A' \cup B') \le \tilde{\mathbb{F}}_N(\psi_N, A) + \tilde{\mathbb{F}}_N\Big(\varphi_N^i, (A_{5(i+1)} \setminus \overline{A_{5i}}) \cap B'\Big) + \tilde{\mathbb{F}}_N(\varphi_N, B) \tag{4.65}$$

for $i \in \{1, \dots, M\}$, as well as the bound

$$\frac{1}{M} \sum_{i=1}^{M} \tilde{\mathbb{F}}_N\Big(\varphi_N^i, (A_{5(i+1)} \setminus \overline{A_{5i}}) \cap B'\Big) \lesssim \frac{E}{M} + \bar{k} M^2 \|\psi_N - \varphi_N\|_{L^2(\Omega)}^2,$$

where we used that $(A_{5(i+1)} \setminus \overline{A_{5i}}) \cap B' \subset A \cap B$ and that the energy of the recovery sequences ψ_N and φ_N is bounded from above, thus

$$\sup_{N \in \mathbb{N}} \tilde{\mathbb{F}}_N(\psi_N, A) \vee \sup_{N \in \mathbb{N}} \tilde{\mathbb{F}}_N(\varphi_N, B) \le E = E(A, B) < +\infty.$$

We then plug the error estimates above into (4.65) and deduce

$$\frac{1}{M} \sum_{i=1}^{M} \tilde{\mathbb{F}}_N(\varphi_N^i, A' \cup B') - \tilde{\mathbb{F}}_N(\psi_N, A) - \tilde{\mathbb{F}}_N(\varphi_N, B) \lesssim \frac{E}{M} + \bar{k} M^2 \|\psi_N - \varphi_N\|_{L^2(\Omega)}^2.$$

Using the fact that $\psi_N, \varphi_N \to \varphi$ are recovery sequences, we may pass to the limit $N \to \infty$ in the previous bound and obtain, for fixed $M \in \mathbb{N}$,

$$\limsup_{N\to\infty} \frac{1}{M} \sum_{i=1}^{M} \tilde{\mathbb{F}}_N(\varphi_N^i, A' \cup B') - \mathbb{F}(\varphi, A) - \mathbb{F}(\varphi, B) \lesssim \frac{E}{M}. \tag{4.66}$$

Arguing again as in the proof of Proposition 4.6.8, we note that, for fixed $M \in \mathbb{N}$, there exists a sequence $\varphi_N^{i_N}$ satisfying $\varphi_N^{i_N} \to \varphi$ in $L^2(\Omega)$ as $N \to \infty$ and

$$\tilde{\mathbb{F}}_N(\varphi_N^{i_N}, A' \cup B') \le \frac{1}{M} \sum_{i=1}^{M} \tilde{\mathbb{F}}_N(\varphi_N^i, A' \cup B').$$

Together with (4.66), this yields

$$\mathbb{F}(\varphi, A' \cup B') \le \limsup_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N^{i_N}, A' \cup B') \le \mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B) + C\frac{E}{M}$$

for every $M \in \mathbb{N}$, for some $C = C(d, \zeta)$ and $E = E(A, B) \in \mathbb{R}_+$. Taking the limit $M \to \infty$, we infer that

$$\mathbb{F}(\varphi, A' \cup B') \le \mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B)$$

and the proof is complete. □

The following additivity property turns out to be much easier to prove than the corresponding result on the grid in [AC04], due to inner regularity in combination with the very short range of interaction (nearest neighbours on a scale of order $[\mathcal{T}_N]$).

Proposition 4.7.7 (Additivity on disjoint sets)**.** *Assume* (ub)*. For any $\varphi \in H^1(\Omega)$ the function $\mathbb{F}(\varphi, \cdot)$ is additive on disjoint sets, i.e.,*

$$\mathbb{F}(\varphi, A \cup B) = \mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B) \tag{4.67}$$

for all $A, B \in \mathcal{O}(\Omega)$ such that $A \cap B = \emptyset$.

Proof. In view of the subadditivity result from Proposition 4.7.6, it remains to show superadditivity on disjoint sets. Fix $A, B \in \mathcal{O}(\Omega)$ with $A \cap B = \emptyset$. By inner regularity (Proposition 4.6.8) we may assume that $d(A, B) > 0$. Consequently, for N sufficiently large we have

$$\tilde{\mathbb{F}}_N(\varphi, A \cup B) = \tilde{\mathbb{F}}_N(\varphi, A) + \tilde{\mathbb{F}}_N(\varphi, B) \quad \forall \varphi \in H^1(\Omega).$$

Fix $\varphi \in H^1(\Omega)$ and let $(\varphi_N)_N$ be a recovery sequence for $\mathbb{F}(\varphi, A \cup B)$. Using the previous identity we obtain

$$\begin{aligned}
\mathbb{F}(\varphi, A) + \mathbb{F}(\varphi, B) &\le \liminf_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N, A) + \liminf_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N, B) \\
&\le \liminf_{N\to\infty} \Big(\tilde{\mathbb{F}}_N(\varphi_N, A) + \tilde{\mathbb{F}}_N(\varphi_N, B) \Big) \\
&= \liminf_{N\to\infty} \tilde{\mathbb{F}}_N(\varphi_N, A \cup B) \\
&= \mathbb{F}(\varphi, A \cup B),
\end{aligned}$$

which is the desired superadditivity inequality. □

We are now in a position to collect the pieces for the proof of Theorem 4.7.1.

Proof of Theorem 4.7.1. In view of Theorem 4.6.1, we know that $\mathbb{F} = +\infty$ outside of $H^1(\Omega)$. To obtain the desired result on $H^1(\Omega)$ we check that $\mathbb{F}(\cdot, A)$ satisfies the conditions of Theorem 4.7.3.

The locality (i) has been shown in Proposition 4.7.5.

To prove (ii), we apply the De Giorgi-Letta criterion, cf. [DGL77], [BD98]. For any $\varphi \in H^1(\Omega)$, it follows from Propositions 4.6.8, 4.7.6, and 4.7.7 that $\mathbb{F}(\varphi, \cdot)$ is the restriction of a Borel measure to $\mathcal{O}(\Omega)$.

The Sobolev upper bound (iii) has been proved in Proposition 4.6.6, whereas the corresponding lower bound follows from Proposition 4.6.5.

Finally, to prove (iv) we note that lower semicontinuity with respect to strong $L^2(\Omega)$-convergence follows from the fact any Γ-limit is lower semicontinuous; see [Bra02, Proposition 1.28]. Since $H^1(\Omega)$ is compactly embedded in $L^2(\Omega)$, the result follows. □

4.7.3 The characterisation of the Γ-limit

To prove Theorem 4.3.9 it remains to characterise the Γ-limit $\mathbb{F}$ obtained in Theorem 4.7.1. It thus remains to compute the function F appearing in Theorem 4.7.1. From (4.59) it follows that for $x \in \Omega$, $u \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$,

$$F(x, u, \xi) = \limsup_{\varepsilon \to 0^+} \frac{\mathbb{M}\Big(u + \xi(\cdot - x); Q_\varepsilon(x)\Big)}{\varepsilon^d}, \tag{4.68}$$

where $Q_\varepsilon(x)$ denotes the open cube of side-length ε centred at x and

$$\mathbb{M}(\varphi, A) := \inf_{\psi} \Big\{ \mathbb{F}(\psi, A) \ : \ \psi \in H^1(\Omega) \text{ s.t. } \psi - \varphi \in H^1_0(A) \Big\}$$

for any Lipschitz function $\varphi : \Omega \to \mathbb{R}$ and any open set $A \subseteq \Omega$ with Lipschitz boundary. As we will compute $\mathbb{M}$ by discrete approximation, we consider its discrete counterpart $\mathcal{M}_\mathcal{T}$ defined by

$$\mathcal{M}_\mathcal{T}(f, A) := \inf_g \{ \mathcal{F}_\mathcal{T}(g, A) \ : \ g \in \mathbb{R}^\mathcal{T} \text{ s.t. } f = g \text{ on } \mathcal{T}|_{A^c} \}$$

for $f : \mathcal{T} \to \mathbb{R}$, where $\mathcal{T}|_A$ for $A \subset \Omega$ is defined in (4.41).

Remark 4.7.8 (Strong continuity of $\mathbb{F}(\cdot, A)$ in $H^1(\Omega)$). The quadratic nature of the discrete problems allows us to infer more information about the limit density. In fact, it follows that $F(x, u, \xi) = \langle a(x)\xi, \xi \rangle$ for some bounded matrix-valued function a; see [AC04, Remark 3.2]. Consequently, for every $A \in \mathcal{O}(\Omega)$, the Γ-limit $\mathbb{F}(\cdot, A)$ is continuous for the strong topology of $H^1(\Omega)$. This fact will be used in the proof of Lemma 4.7.9 below.

The following result is crucial in the proof of Theorem 4.7.2.

Lemma 4.7.9. *Assume (ub), and suppose that $\tilde{\mathbb{F}}_N(\cdot, B) \xrightarrow{\Gamma} \mathbb{F}(\cdot, B)$ in $L^2(\Omega)$ as $N \to \infty$ for any $B \in \mathcal{O}(\Omega)$. Then, for any $A \in \mathcal{O}(\Omega)$ with Lipschitz boundary and any Lipschitz function $\varphi : \Omega \to \mathbb{R}$, we have*

$$\mathcal{M}_N(P_N \varphi, A) \to \mathbb{M}(\varphi, A). \tag{4.69}$$

Proof. First we embed the discrete functionals in the continuous setting. For any Lipschitz function $\varphi : \overline{\Omega} \to \mathbb{R}$ and any open set $A \subseteq \Omega$ we set

$$\mathsf{PC}_N(\varphi, A) := \{\psi \in \mathsf{PC}_N \ : \ \psi(x_K) = \varphi(x_K) \ \forall K \in \mathcal{T}_N|_{A^c}\}. \tag{4.70}$$

We consider the embedded discrete energies $\tilde{\mathbb{F}}_N^\varphi : L^2(\Omega) \to [0, +\infty]$ defined by

$$\tilde{\mathbb{F}}_N^\varphi(\psi, A) := \begin{cases} \mathcal{F}_N(P_N\psi, A) & \text{if } \psi \in \mathsf{PC}_N(\varphi, A), \\ +\infty & \text{otherwise,} \end{cases}$$

and their continuous counterpart $\mathbb{F}^\varphi : L^2(\Omega) \to [0, +\infty]$ defined by

$$\mathbb{F}^\varphi(\psi, A) := \begin{cases} \mathbb{F}(\psi, A) & \text{if } \psi - \varphi \in H_0^1(A), \\ +\infty & \text{otherwise.} \end{cases}$$

We claim that

$$\tilde{\mathbb{F}}_N^\varphi(\cdot, A) \xrightarrow{\Gamma} \mathbb{F}^\varphi(\cdot, A), \quad \forall A \subseteq \Omega \text{ with Lipschitz boundary, } \varphi \in \mathrm{Lip}(\mathbb{R}^d),$$

which implies, together with Proposition 4.6.5 and by a basic result from the theory of Γ-convergence [Bra02, Theorem 1.21], the desired convergence of the minima in (4.69). To prove the claim, we argue as in [AC04, Theorem 3.10].

To prove the liminf inequality, we consider a sequence $\psi_N \rightharpoonup \psi$ in $L^2(\Omega)$ satisfying $\sup_N \tilde{\mathbb{F}}_N^\varphi(\psi_N, A) < +\infty$. In particular, this implies that $\psi_N \in \mathsf{PC}_N(\varphi, A)$ and $\tilde{\mathbb{F}}_N^\varphi(\psi_N, A) = \tilde{\mathbb{F}}_N(\psi_N, A)$. Since $\tilde{\mathbb{F}}_N(\cdot, A) \xrightarrow{\Gamma} \mathbb{F}(\cdot, A)$, it remains to prove that $\psi - \varphi \in H_0^1(A)$. In view of the boundary condition and the fact that $\varphi \in \mathrm{Lip}(\mathbb{R}^d)$, we have

$$\tilde{\mathbb{F}}_N(\psi_N, \Omega) \le \tilde{\mathbb{F}}_N(\psi_N, A) + \tilde{\mathbb{F}}_N(\varphi, \Omega) \lesssim \tilde{\mathbb{F}}_N(\psi_N, A) + \overline{k}\, \mathrm{Lip}(\varphi)^2.$$

It follows from this bound and Proposition 4.6.5 that $\psi_N \to \psi$ strongly in $L^2(\Omega)$ and $\psi \in H^1(\Omega)$. Moreover, by construction we have $\psi_N \to \varphi$ in $L^2(\Omega \setminus A)$. Since A has Lipschitz boundary, we conclude that $\psi - \varphi \in H_0^1(A)$.

Let us now prove the limsup inequality. Pick $\psi \in L^2(\Omega)$ such that $\mathbb{F}^\varphi(\psi, A) < +\infty$. In particular, $\psi - \varphi \in H_0^1(A)$. Without loss of generality, we may assume that $\mathrm{supp}(\psi - \varphi) \Subset A$, as the general case follows from this by a density argument using the continuity of $\mathbb{F}$ in the strong $H^1(\Omega)$-topology; see Remark 4.7.8. Consider a recovery sequence $\psi_N \to \psi$ in $L^2(\Omega)$ such that $\tilde{\mathbb{F}}_N(\psi_N, A) \to \mathbb{F}(\psi, A) = \mathbb{F}^\varphi(\psi, A)$ as $N \to \infty$. Now we argue as in the proof of Proposition 4.6.8. For any $\delta > 0$ there exists a cutoff function ζ_δ with the following properties:

(i) $\mathrm{supp}(\psi - \varphi) \Subset \mathrm{supp}\, \zeta_\delta \Subset A$;

(ii) the functions $\psi_N^\delta := Q_N \circ P_N\Big(\zeta_\delta \psi_N + (1 - \zeta_\delta)\varphi\Big)$ satisfy

$$\begin{aligned}\limsup_{N\to\infty} \tilde{\mathbb{F}}_N^\varphi(\psi_N^\delta, A) &= \limsup_{N\to\infty} \tilde{\mathbb{F}}_N(\psi_N^\delta, A) \\ &\le \limsup_{N\to\infty} \tilde{\mathbb{F}}_N(\psi_N, A) + \delta = \mathbb{F}^\varphi(\psi, A) + \delta.\end{aligned}$$

Passing to the limit $\delta \to 0$ using a diagonal subsequence $\psi_N^{\delta(N)} \to \psi$ in $L^2(\Omega)$, the result follows. □

Proof of Theorem 4.7.2. We split the proof into two parts.

Step 1. We first suppose that μ is the normalised Lebesgue measure and $m_N(K) = \pi_N(K) = \frac{|K|}{|\Omega|}$, and we fix $\varepsilon > 0$. For fixed $b \in \mathbb{R}$, $z \in \Omega$, and $\xi \in \mathbb{R}^d$ we will compute

$$\mathcal{M}_N\Big(f_N, Q_\varepsilon(z)\Big), \quad \text{where } f_N(K) := \varphi^\xi_{b,z}(x_K) \text{ and } \varphi^\xi_{b,z}(\cdot) := u + \xi(\cdot - z)$$

As a shorthand we write $Q_\varepsilon := Q_\varepsilon(z)$. Recall that

$$\mathcal{M}_N(f, Q_\varepsilon) = \inf_g \Big\{\mathcal{F}_N(g, Q_\varepsilon) \ : \ g \in \mathbb{R}^{\mathcal{T}_N} \text{ and } g(K) = f(K) \text{ for } K \in \mathcal{T}_N|_{Q^c_\varepsilon}\Big\}.$$

In other words, we minimise the discrete Dirichlet energy localised on Q_ε with Dirichlet boundary conditions given by the discretised affine function f. As follows by computing the first variation of the action, the unique minimiser is given by the solution $h : \mathcal{T}_N \to \mathbb{R}$ of the corresponding discrete Laplace equation

$$\begin{cases} \mathcal{L}_N h(K) = 0 & \text{for } K \in \mathcal{T}_N \setminus \mathcal{T}_N|_{Q^c_\varepsilon}, \\ h(K) = f_N(K) & \text{for } K \in \mathcal{T}_N|_{Q^c_\varepsilon}. \end{cases} \tag{4.71}$$

We claim that the function f_N solves (4.71). Indeed, the boundary conditions hold trivially. Moreover, writing $\tau_{KL} := \frac{x_K - x_L}{|x_K - x_L|}$ we obtain for any $K \in \mathcal{T}_N \setminus \mathcal{T}_N|_{Q^c_\varepsilon}$,

$$\begin{aligned} \pi_N(K)\mathcal{L}_N f_N(K) &= \sum_{L\sim K} \frac{|\Gamma_{KL}|}{d_{KL}}\Big(f_N(L) - f_N(K)\Big) = -\sum_{L\sim K} |\Gamma_{KL}|\langle \xi, \tau_{KL}\rangle \\ &= \int_{\partial K} \langle \xi, \nu_{\text{ext}}\rangle \, \mathrm{d}\mathscr{H}^{d-1} = 0, \end{aligned}$$

where ν_{ext} denotes the outward normal unit normal and in the last step we used Stokes' theorem. This computation shows the optimality of f and hence

$$\mathcal{M}_N(f_N, Q_\varepsilon) = \mathcal{F}_N(f_N, Q_\varepsilon).$$

For the asymptotic computation of $\mathcal{F}_N(f_N, Q_\varepsilon)$ we use the average isotropy property of any regular mesh (see [GKM20, Lemma 5.4]) to obtain

$$\begin{aligned} \Big|\mathcal{F}_N(f_N, Q_\varepsilon) - \varepsilon^d|\xi|^2\Big| &= \left|\left(\frac{1}{4}\sum_{\substack{K,L\in\mathcal{T}_N \\ \overline{K},\overline{L}\cap Q_\varepsilon \neq \emptyset}} d_{KL}|\Gamma_{KL}|\langle\xi, \tau_{KL}\rangle^2\right) - |\xi|^2|Q_\varepsilon|\right| \\ &\le \Big|B\Big(\partial Q_\varepsilon, 5[\mathcal{T}_N]\Big)\Big| \to 0 \text{ as } N \to \infty, \end{aligned}$$

where $B(C, r) := \{x \in \Omega \ : \ d(x, C) < r\}$. Note that we get $|B(\partial Q_\varepsilon, 5[\mathcal{T}_N])|$ instead of $|B(\partial Q_\varepsilon, 4[\mathcal{T}_N])|$ as in [GKM20, Lemma 5.4] because we take into account all the cells whose closure intersects the cube Q_ε and not only the ones contained in it. Together with Lemma 4.7.9, we obtain, for all $\xi \in \mathbb{R}^d$ and $\varepsilon > 0$,

$$\mathbb{M}(\varphi^\xi_{b,z}, Q_\varepsilon) = \lim_{N\to\infty} \mathcal{M}_N(f, Q_\varepsilon) = \lim_{N\to\infty} \mathcal{F}_N(f, Q_\varepsilon) = \varepsilon^d|\xi|^2, \tag{4.72}$$

hence

$$F(x, u, \xi) = \limsup_{\varepsilon\to 0^+} \frac{\mathbb{M}(\varphi^\xi_{b,z}, Q_\varepsilon)}{\varepsilon^d} = |\xi|^2,$$

which concludes the proof in the special case σ, $\rho \equiv 1$, $m_N = \pi_N$.

Step 2. Let us now consider the general case where m_N and μ satisfy (lb), (ub), and (pc). We write $\bar{\mathcal{F}}_N, \bar{\mathcal{M}}_N$ for the analogues of $\mathcal{F}_N, \mathcal{M}_N$ in the special case where μ is the normalised Lebesgue measure and $m_N = \pi_N$, which we considered in Step 1.

Fix $b \in \mathbb{R}$, $z \in \Omega$, and $\xi \in \mathbb{R}^d$, and let Q_ε, $\varphi^\xi_{b,z}$, and f be as above. Furthermore, let υ_N be the density of $\mathsf{Q}_N m_N$ with respect to the Lebesgue measure. For all $g : \mathcal{T}_N \to \mathbb{R}$ we have by construction,

$$\Big(\inf_{Q_{2\varepsilon}} \upsilon_N\Big) \bar{\mathcal{F}}_N(g, Q_\varepsilon) \le \mathcal{F}_N(g, Q_\varepsilon) \le \Big(\sup_{Q_{2\varepsilon}} \upsilon_N\Big) \bar{\mathcal{F}}_N(g, Q_\varepsilon),$$

hence, in particular,

$$\Big(\inf_{Q_{2\varepsilon}} \upsilon_N\Big) \bar{\mathcal{M}}_N(f, Q_\varepsilon) \le \mathcal{M}_N(f, Q_\varepsilon) \le \Big(\sup_{Q_{2\varepsilon}} \upsilon_N\Big) \bar{\mathcal{M}}_N(f, Q_\varepsilon).$$

As a consequence of the first part of the proof and (4.72), taking the limit as $N \to \infty$ and applying (4.69), we deduce

$$\Big(\limsup_{N\to\infty} \inf_{Q_{2\varepsilon}} \rho_N\Big) |\xi|^2 \varepsilon^d \le \mathbb{M}(\varphi^\xi_{b,z}, Q_\varepsilon) \le \Big(\liminf_{N\to\infty} \sup_{Q_{2\varepsilon}} \rho_N\Big) |\xi|^2 \varepsilon^d.$$

Taking the limsup as $\varepsilon \to 0$, we deduce from (4.68) and the condition (pc),

$$F(x, u, \xi) = \limsup_{\varepsilon\to 0} \frac{\mathbb{M}(\varphi^\xi_{b,z}, Q_\varepsilon)}{\varepsilon^d} = |\xi|^2 \upsilon(x) \quad \text{for a.e. } z \in \Omega,$$

which concludes the proof. □

CHAPTER 5

A non-commutative entropic optimal transport approach to quantum composite systems at positive temperature

In the last chapter, we present a duality result and the convergence of a Sinkhorn algorithm for multimarginal, non-commutative optimal transport problems in finite dimension. This is the content of the work [FGP21] in collaboration with Dario Feliciangeli and Augusto Gerolin.

More in detail, we study a multimarginal, non-commutative analogue of the classical Schrödinger problem, seen as a entropic regularisation of an optimal transport problem between density matrices on finite dimensional Hilbert spaces. From a physical perspective, this describes a composite quantum system at positive temperature conditional to the knowledge of the states of all its subsystems. As a particular case, we discuss applications to the one-body reduced density matrix functional theory (1RDMFT), both in the bosonic and in the fermionic setting. Moreover, we introduce a non-commutative analogue of the (multimarginal) Sinkhorn algorithm and proves its convergence to the optimal states. Our results are based on a novel, non-commutative notion of $(\mathsf{H}, \varepsilon)$-transform, which takes inspiration from the recent contribution of Di Marino and Gerolin [DMG20a] in the classical setting.

5.1 Introduction

In this work we are interested in studying the ground state energy of a finite dimensional composite quantum system at positive temperature. In particular, we focus on the problem of minimizing the energy of the composite system *conditionally* to the knowledge of the states of all its subsystems.

The first motivation for this study is physical: it is useful to understand how one could infer the state of a composite system when one only has experimental access to the measurement of the states of its subsystems. The second motivation is mathematical: indeed this problem can be cast as a non-commutative optimal transport problem, therefore showcasing how several ideas and concepts introduced in the commutative setting carry through to the non-commutative framework. Finally, a third motivation comes from the fact that one-body reduced density matrix functional theory, which is of interest on its own, can be framed as a special case of our setting.

Let us consider a composite system with N subsystems, each with state space given by the complex Hilbert space $\mathfrak{h}_j$ of dimension $d_j < \infty$, for $j = 1, \dots, N$, and denote the state space of the composite system $\mathfrak{h} := \mathfrak{h}_1 \otimes \mathfrak{h}_2 \otimes \cdots \otimes \mathfrak{h}_N$ (with dimension $d = d_1 \cdot d_2 \cdot \ldots d_N$). Further denote by H the Hamiltonian to which the whole system is subject and suppose that $\mathrm{H} = \mathrm{H}_0 + \mathrm{H}_{\mathsf{int}}$, where H_0 is the non-interacting part of the Hamiltonian, i.e. $\mathrm{H}_0 = \bigoplus_{j=1}^N \mathrm{H}_j := \mathrm{H}_1 \otimes \mathbb{1} \cdots \otimes \mathbb{1} + \mathbb{1} \otimes \mathrm{H}_2 \otimes \mathbb{1} \cdots \otimes \mathbb{1} + \cdots + \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \mathrm{H}_N$ with H_j acting on $\mathfrak{h}_j$, and $\mathrm{H}_{\mathsf{int}}$ is its interacting part. Finally, suppose to have knowledge of the states $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)$ of the N subsystems, where each γ_j is a density matrix over $\mathfrak{h}_j$.

Then the energy of the composite system at temperature $\varepsilon > 0$ is given by

$$\begin{aligned} \inf_{\Gamma \mapsto \boldsymbol{\gamma}} \{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon S(\Gamma)\} &= \sum_{j=1}^N \mathrm{Tr}(\mathrm{H}_j\,\gamma_j) + \mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) \\ &:= \sum_{j=1}^N \mathrm{Tr}(\mathrm{H}_j\,\gamma_j) + \inf_{\Gamma \mapsto \boldsymbol{\gamma}} \{\mathrm{Tr}(\mathrm{H}_{\mathsf{int}}\,\Gamma) + \varepsilon S(\Gamma)\}\,, \end{aligned} \tag{5.1}$$

where the shorthand notation $\Gamma \mapsto \boldsymbol{\gamma}$ denotes the set of density matrices over $\mathfrak{h}$ with j-th marginal equal to γ_j, and $S(\Gamma) := \mathrm{Tr}\,(\Gamma \log(\Gamma))$ is the opposite of the Von Neumann entropy of Γ (note that we prefer to adopt the mathematical sign convention).

Our approach for the study of $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$ borrows ideas from optimal transport and convex analysis, and takes the following observation as a starting point: the minimization appearing in $\mathfrak{F}^\varepsilon$ can be cast as a non-commutative entropic optimal transport problem. Indeed, one looks for an optimal non-commutative coupling Γ, with fixed non-commutative marginals (i.e. partial traces) $\boldsymbol{\gamma}$, which minimizes the sum of a transport cost (given by $\mathrm{Tr}(\mathrm{H}_{\mathsf{int}}\,\Gamma)$) and an entropic term. In light of this interpretation, setting the quantum problem at positive temperature ε corresponds to consider an entropic optimal transport problem with parameter ε.

Guided by this viewpoint, we first show that $\mathfrak{F}^\varepsilon$ has a dual formulation (see Theorem 5.2.1 (i)), i.e. that the constrained minimization appearing in its definition is in duality with an unconstrained maximization problem (defined in (5.7)). We can then consider any vector $(U_1^\varepsilon, \dots, U_N^\varepsilon)$ of self-adjoint matrices which is a maximizer in the dual functional of $\mathfrak{F}^\varepsilon$, whose existence and uniqueness up to trivial transformations we prove in Theorem 5.2.1(ii). We refer to such U_i^ε-s as *Kantorovich potentials* and show in Theorem 5.2.1(iii) that the unique minimizer Γ^ε realizing $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$ can be written in terms of them as

$$\Gamma^\varepsilon = \exp\left(\frac{\bigoplus_{i=1}^N U_i^\varepsilon - \mathrm{H}_{\mathsf{int}}}{\varepsilon}\right), \tag{5.2}$$

in the case of all the γ_j-s having trivial kernels (in the general case a very similar formula holds). In this setting, $\mathfrak{F}^\varepsilon$ is continuous and its functional derivative can be computed in terms of the *Kantorovich potentials* as

$$\frac{\mathrm{d}\mathfrak{F}^\varepsilon}{\mathrm{d}\gamma_i}(\boldsymbol{\gamma}) = U_i^\varepsilon, \quad \text{for all} \quad i = 1, \dots, N, \tag{5.3}$$

as we show in Proposition 5.2.2.

Furthermore, we introduce the Non-Commutative Sinkhorn algorithm to compute the optimizer realizing $\mathfrak{F}^\varepsilon(\gamma)$. This algorithm exploits the shape of the minimizer obtained in (5.2), in order to construct a sequence $\Gamma^{(k)}$ of density matrices converging to Γ^ε of the form

$$\Gamma^{(k)} = \exp\left(\frac{\bigoplus_{i=1}^N U_i^{(k)} - \mathrm{H}_{\mathsf{int}}}{\varepsilon}\right), \tag{5.4}$$

where the vector $(U_1^{(k)}, \dots, U_N^{(k)})$ is iteratively updated by progressively imposing that $\Gamma^{(k)}$ has at least one correct marginal. We prove the convergence and the robustness of this algorithm in Section 5.5.

It is important to note that studying $\mathfrak{F}^\varepsilon(\gamma)$, i.e. the constrained minimization at fixed marginals, can also help solving the unconstrained minimization of the Hamiltonian H at positive temperature ε. Indeed, denoting by $\mathfrak{P}(\mathfrak{h})$ the set of density matrices over $\mathfrak{h}$, then

$$E^\varepsilon(\mathrm{H}) := \inf_{\Gamma\in\mathfrak{P}(\mathfrak{h})} \{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon S(\Gamma)\} = \inf_\gamma \left\{\sum_{j=1}^N \mathrm{Tr}(\mathrm{H}_j\,\gamma_j) + \mathfrak{F}^\varepsilon(\gamma)\right\}. \tag{5.5}$$

Combining (5.3) and (5.5) allows to write down the Euler–Lagrange equation of (5.5) recovering its optimizer, i.e. the Gibbs state constructed with H at temperature ε.

Our work is not the first to try to extend the theory of optimal transport to the non-commutative setting. One of the first attempts was carried out by E. Carlen and J. Maas [CM14], followed by many others (e.g. [BV20, CGP18, CGP20, CGGT19, CGT18, DPT19, DPTGA18, GP15, MM17, MV20, PCVS19]). There is an important distinction to be made here. Commutative optimal transport can be cast *equivalently* as a static coupling problem or as a dynamical optimization problem. On the other hand, in the non-commutative setting it is not clear what is the relation (if any) between the two interpretations. This singles out a big difference between works that consider the dynamical formulation of commutative optimal transport as a starting point (e.g. [BV20, CM14, CGGT19, CGT18, MM17, MV20, PCVS19]) and the ones which instead focus on its static formulation (e.g. [Cut13, GS10, L1́4, Sch31, Zam86]).

This paper adopts an even different approach. We consider as a starting point the Entropic regularization of optimal transport (which is to be considered as an extension of static optimal transport, see e.g. the survey [L1́4] and references therein) and introduce its non-commutative counterpart. We carry out this program by extending the method developed in [DMG20a, DMG20b, GKR20]. See also Section 5.5 for a detailed explanation of the multimarginal Sinkhorn algorithm in the commutative setting, as studied in [DMG20a].

In the work [CGP18], the authors study the case of $\varepsilon = 0$ temperature and prove a duality result for the non-commutative problem in the very same spirit of the Kantorovich duality for the classical Monge problem. The recent work [Wir18] studies the entropic quantum optimal transport problem as well, adopting, in constrast to our static approach, a dynamical formulation. Therein, the author proves a dynamical duality result at positive and zero temperature. To the best of our knowledge, the present work is the first complete analysis of the quantum entropic transport problem in the static framework.

As for the Sinkhorn Algorithm, another concept which we borrow from the commutative setting and extend to the quantum one, its convergence in the commutative setting was first established in the $N = 2$ marginal case [FL89, Sin64] for discrete measures and in [Rus95] for continuous measures (see also [CGP16]). In the multi-marginal setting, convergence guarantees were obtained for the discrete case in [CPSV18, KR17] and for continuous measures in [DMG20a, DMG20b]. Other variants of the Sinkhorn algorithm for (unbalanced) tensor-valued measures or matrix optimal mass transport have been studied in [PCVS19, RCLO18] and do not apply to our setting. In the context of Computational Optimal Transport, the entropic regularization and the Sinkhorn algorithm was introduced in [Cut13, GS10].

Enforcing symmetry constraints: One-body Reduced Density Matrix Functional Theory

We conclude this introduction by briefly discussing the case in which symmetry conditions are enforced on the problem, either bosonic or fermionic, which we can also treat (see Section 5.2.3). In this case, (5.1) makes sense only for $\mathfrak{h}_j = \mathfrak{h}_0$ for all $j = 1, \ldots, N$ and $\boldsymbol{\gamma} = (\gamma, \ldots, \gamma)$ (i.e. the underlying Hilbert spaces and the marginals must all be the same) and its study can be framed in the context of One-body Reduced Density Matrix Functional Theory (1RDMFT), introduced in 1975 by Gilbert [Gil75] as an extension of the Hohenberg-Kohn (Levy-Lieb) formulation of Density Functional Theory (DFT) [HK64, Lev76, Lie02]. In the last decades, DFT and 1RDMFT have been standard methods for numerical electronic structure calculations and are to be considered a major breakthrough in fields ranging from materials science to chemistry and biochemistry.
In both these theories one tries to approximate a complicated N-particle quantum system by studying one-particle objects, namely one-body densities in the case of DFT and one-body reduced density matrices in the case of 1RDMFT, by using a two-steps minimization analogous to the one introduced in (5.5).

It is interesting to see that the well-known Pauli principle (see e.g. [LS10, Theorem 3.2]) , which provides necessary and sufficient conditions for γ to be the one-body reduced density matrix of an N-body antisymmetric density matrix, finds a variational interpretation in our discussion. Indeed, in the antisymmetric case we show (see Proposition 5.2.8) that γ satisfies the Pauli principle (resp. satisfies the Pauli principle *strictly*) if and only if the supremum of the dual functional of $\mathfrak{F}^{\varepsilon}$ is finite (resp. is attained), as it is to be expected.

Other extensions of DFT have been considered, including Mermin's Thermal Density Functional Theory [Mer65], Spin DFT [vBH72], and Current DFT [VR87]. Physical and computational aspects of 1RDMFT have been investigated in [AL05, BCG15, BEG12, BG12, BB02, Men15, Mül84, Per05, PG15, RP08, Sch19, vL07]. A framework for 1RDMFT for Bosons at zero temperature was recently introduced in [BRWMS20] (see also [GR19] and references therein for a recent review). In particular, the first exchange-correlation energy in density-matrix functional theory was introduced by Müller [Mül84], leading to mathematical results [FLSS07, FNVDB18].

Organisation of the paper

The paper is divided as follows: in Section 5.2 we introduce the framework, the main definitions, and present our main results Theorem 5.2.1, Theorem 5.2.3, and Theorem 5.2.9. In Section 5.3 we introduce and develop the technical tools needed to prove our main results, in particular we define the notion of non-commutative $(\mathrm{H}, \varepsilon)$-transform (see Section 5.3.1) and prove a stability and differentiability result for the primal problem in Proposition 5.2.2. In Section 5.4, Section 5.5, and Section 5.6 we build upon Section 5.3 and prove our main results, respectively, Theorem 5.2.1, Theorem 5.2.3, and Theorem 5.2.9.

5.2 Contributions and statements of the main results

The main contributions of this work consist in

- Theorem 5.2.1, which represents a duality result for the functional $\mathfrak{F}^\varepsilon$ (whose definition is recalled below in equation (5.6)). Theorem 5.2.1 also includes the characterization of the optimizers of $\mathfrak{F}^\varepsilon$ (and of its dual functional).
- The introduction of a non-commutative Sinkhorn algorithm, which can be used to compute the aforementioned optimizers. We also prove convergence and robustness of this algorithm in Theorem 5.2.3.
- The introduction of a non-commutative notion of $(\mathrm{H}, \varepsilon)$-transform and the proof of suitable a priori estimates in Section 5.3, which turn out to be crucial in the proof of Theorem 5.2.1 and the convergence of the Sinkhorn algorithm (Theorem 5.2.3). Consequently, we are also able to show stability and differentiability of $\mathfrak{F}^\varepsilon(\cdot)$ in Proposition 5.2.2.
- The generalization of Theorem 5.2.1 to the case of bosonic or fermionic systems, stated in Theorem 5.2.9. This also allows to give an interesting variational characterization of the Pauli exclusion principle (see Proposition 5.2.8).

We now proceed to introduce our setting and state our main contributions.

5.2.1 Duality and minimization of $\mathfrak{F}^\varepsilon$

We recall that in this case we simply work with a general composite system, with no symmetry constraints enforced. For $d \in \mathbb{N}$, we shall denote by $\mathcal{M}^d = \mathcal{M}^d(\mathbb{C})$ the set of all $d \times d$ complex matrices, by $\mathcal{S}^d$ the hermitian elements of $\mathcal{M}^d$, and by $\mathcal{S}^d_{\geq}$ (respectively $\mathcal{S}^d_{>}$) the set of all the positive semidefinite (resp. positive definite) elements of $\mathcal{S}^d$. With a slight abuse of notation, we denote by Tr the trace operator on $\mathcal{M}_d$ for any dimension d. Furthemore, for any Hilbert space $\mathfrak{h}$, we denote by $\mathfrak{P}(\mathfrak{h})$ the set of *density matrices* over $\mathfrak{h}$, namely the positive self-adjoint operators with trace one. For simplicity, we shall also use the notation $\mathfrak{P}^d = \mathfrak{P}(\mathbb{C}^d)$. For every $N \in \mathbb{N}$ we adopt the notation $[N] := \{1, ..., N\}$.

Our main object of study is the minimisation problem for $N \in \mathbb{N}$, $i \in [N]$, $\gamma_i \in \mathfrak{P}^{d_i}$, $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$

$$\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \inf \left\{ \mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon\,\mathrm{Tr}(\Gamma \log \Gamma) \;:\; \Gamma \in \mathfrak{P}^{\boldsymbol{d}} \text{ and } \Gamma \mapsto \boldsymbol{\gamma} \right\}, \tag{5.6}$$

where $d_i \in \mathbb{N}$, $\boldsymbol{d} := \prod_{i=1}^N d_i$, $\boldsymbol{\gamma} := (\gamma_i)_{i \in [N]}$, and $\Gamma \mapsto \boldsymbol{\gamma}$ means that the i-th marginal (5.23) of Γ is equal to γ_i. This coincides with the Definition of $\mathfrak{F}^\varepsilon$ given in (5.1).

The natural space to work with is given by $\mathcal{O} := \bigotimes_{i=1}^N \left(\ker \gamma_i \right)^\perp$ where for simplicity we set $\widehat{d}_i := (d_i - \dim \ker \gamma_i)$ and $\widehat{\mathbf{d}} := \prod_{i=1}^N \widehat{d}_i$. We also denote by $\mathrm{H}_\mathcal{O}$ the restriction of H to the subspace $\mathcal{O}$. The corresponding dual problem is defined as

$$\mathfrak{D}^\varepsilon(\boldsymbol{\gamma}) = \sup \left\{ \sum_{i=1}^N \mathrm{Tr}(U_i \gamma_i) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{\bigoplus_{i=1}^N U_i - \mathrm{H}_\mathcal{O}}{\varepsilon} \right] \right) \;:\; U_i \in \mathcal{S}^{\widehat{d}_i} \right\} + \varepsilon\,, \tag{5.7}$$

where $\bigoplus$ denotes the Kronecker sum (5.24).

Our first result is a duality result and serves also as a characterization of the minimizers in (5.6). Note that, throughout the whole paper, when no confusion can arise, we shall use the slightly imprecise notation $\alpha\mathbb{1} = \alpha$ for $\alpha \in \mathbb{C}$.

Theorem 5.2.1 (Duality). *Let $\varepsilon > 0$, $N \in \mathbb{N}$, and $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$. For fixed $\boldsymbol{\gamma} = (\gamma_i \in \mathfrak{P}^{d_i})_{i\in[N]}$, consider the primal and dual problems $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$, $\mathfrak{D}^\varepsilon(\boldsymbol{\gamma})$ as in (5.6), (5.7) respectively. We then have that*

(i) the primal and dual problems coincide, i.e. $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \mathfrak{D}^\varepsilon(\boldsymbol{\gamma})$.

(ii) $\mathfrak{D}^\varepsilon(\boldsymbol{\gamma})$ admits a maximizer $\{U_i^\varepsilon \in \mathcal{S}^{\widehat{d_i}}\}_{i=1}^N$, which is unique up to trival translation. Precisely, if $\{\tilde{U}_i^\varepsilon \in \mathcal{S}^{\widehat{d_i}}\}_{i=1}^N$ is another maximizer, then $\tilde{U}_i^\varepsilon - U_i^\varepsilon = \alpha_i \in \mathbb{R}$ with $\sum_i \alpha_i = 0$.

(iii) There exists a unique $\Gamma^\varepsilon \in \mathfrak{P}^{\boldsymbol{d}}$ with $\Gamma^\varepsilon \mapsto \boldsymbol{\gamma}$ which minimizes the functional $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$. Moreover, Γ^ε and $\{U_i^\varepsilon\}$ are related via the formula

$$\Gamma^\varepsilon = \exp\left(\frac{\bigoplus_{i=1}^N U_i^\varepsilon - \mathrm{H}_{\mathcal{O}}}{\varepsilon}\right) \quad \text{on } \mathcal{O} \tag{5.8}$$

and $\Gamma^\varepsilon = 0$ on $\mathcal{O}^\perp$.

The proof of the existence of maximizers for the dual problem follows the direct method of Calculus of Variations. In analogy with [DMG20a, DMG20b], where the notion of commutative (c,ε)-trasform is introduced, we define the non-commutative (H,ε)-transform (see Section 5.3.1). We use this tool to obtain a priori estimates on U and infer compactness of the maximizing sequences of Kantorovich potentials. Although this approach is not strictly necessary in our finite dimensional setting to prove (i), we believe these estimates to have independent interests and, in particular, they are fundamental to prove the convergence of the so-called non-commutative Sinkhorn algorithm, the second contribution of this work.

As a byproduct of the a priori estimates obtained in Section 5.3.1, it is possible to prove a stability result (with respect to the marginals) for the Kantorovich potentials and compute the Frechét derivative of $\mathfrak{F}^\varepsilon(\cdot)$. This is the content of the following proposition, which is proved in Section 5.4. For simplicity, we here assume that the marginals have trivial kernel. With a bit more effort, and arguing as in Theorem 5.2.1 (see also Remark 5.3.9), one can obtain a similar result in the general setting as well.

Proposition 5.2.2 (Stability and differentiability of $\mathfrak{F}^\varepsilon$). *Fix $\varepsilon > 0$ and assume $\ker(\gamma_i) = \{0\}$.*

(i) Stability*: if $\boldsymbol{\gamma}^n = (\gamma_i^n)_{n\in\mathbb{N}}$, $\gamma_i^n \subset \mathfrak{P}^{d_i}$ is a sequence of density matrices converging to $\boldsymbol{\gamma} = (\gamma_i)_{n\in\mathbb{N}}$ as $n \to \infty$, then any sequence of Kantorovich potentials $\boldsymbol{U}^{\varepsilon,n}$ converges, up to subsequences and renormalisation, to a Kantorovich potential $\boldsymbol{U}^\varepsilon$ for $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$. Therefore, the functional $\mathfrak{F}^\varepsilon(\cdot)$ is continuous.*

(ii) Frechét differential*: $\mathfrak{F}^\varepsilon(\cdot)$ is Fréchet differentiable and for every $i \in [N]$ it holds*

$$\left(\frac{d\mathfrak{F}^\varepsilon}{d\gamma_i}\right)_{\boldsymbol{\gamma}}(\sigma) = \mathrm{Tr}\left(\boldsymbol{U}_i^\varepsilon \sigma\right), \quad \forall \sigma \in \mathcal{S}^{d_i},\ \mathrm{Tr}(\sigma) = 0\,, \tag{5.9}$$

where $\boldsymbol{U}^\varepsilon$ is a Kantorovich potential for $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$.

As derived in [Gil75] and explained, for instance, in [Per05], the relevance of the functional derivative in the 1RDMFT case is to find an eigenvalue equation to find an efficient optimization for the one-particle eigenvalue equations.

5.2.2 Non-commutative Sinkhorn algorithm

The second contribution of this work is to introduce and prove the convergence of a non-commutative Sinkhorn algorithm (see Section 5.5), aimed at computing numerically the optimal density matrix Γ^ε and the corresponding Kantorovich potentials $\{U_i^\varepsilon\}_i$.

For this purpose, we define non-commutative (H,ε)-transform operators, which extend the notion of $(c,\varepsilon)-$transforms as introduced in [DMG20a] (see also Section 5.5 for a detailed explanation). Note that the (H,ε)-transform also depends on $\boldsymbol{\gamma}$, but we omit this dependence as $\boldsymbol{\gamma}$ is a fixed parameter of the problem.

For $i \in [N]$ and $\varepsilon > 0$, we consider the operators $\mathcal{T}_i^\varepsilon : \bigtimes_{j=1}^N \mathcal{S}^{\widehat{d}_j} \to \bigtimes_{j=1}^N \mathcal{S}^{\widehat{d}_j}$ of the form

$$\boldsymbol{U} := (U_1,\dots,U_N), \quad \left(\mathcal{T}_i^\varepsilon(\boldsymbol{U})\right)_j = \begin{cases} U_j & \text{if } j \neq i, \\ \mathfrak{T}_i^\varepsilon(U_1,\dots,U_{i-1},U_{i+1},\dots U_N) & \text{if } j = i \end{cases}$$

where each $\mathfrak{T}_i^\varepsilon$ is defined implicitly via

$$\mathsf{P}_i\left[\exp\left(\frac{\bigoplus_{j=1}^N \left(\mathcal{T}_i^\varepsilon(\boldsymbol{U})\right)_j - \mathrm{H}_{\mathcal{O}}}{\varepsilon}\right)\right] = \gamma_i \tag{5.10}$$

and P_i denotes the i-th marginal operator, obtained by tracing out all but the i-th coordinate, see (5.23). In Section 5.5, we show that the maps $\mathcal{T}_i^\varepsilon$ are well-defined, i.e. the equation (5.10) admits a unique solution $\mathcal{T}_i^\varepsilon(\boldsymbol{U})$.

Note that, by construction, the matrix $\exp\left(\bigoplus_{i=1}^N \left((\mathcal{T}_i^\varepsilon(\boldsymbol{U}))_j - \mathrm{H}_{\mathcal{O}}\right)/\varepsilon\right) \in \mathfrak{P}^{\widehat{\mathbf{d}}}$ and it has the i-th marginal equal to γ_i. The non-commutative Sinkhorn algorithm is then defined by iterating this procedure for every $i \in [N]$. We define the one-step Sinkhorn map as

$$\begin{aligned} \tau : \bigtimes_{j=1}^N \mathcal{S}^{\widehat{d}_j} &\to \bigtimes_{j=1}^N \mathcal{S}^{\widehat{d}_j}, \\ \tau(\boldsymbol{U}) &:= (\mathcal{T}_N^\varepsilon \circ \dots \circ \mathcal{T}_1^\varepsilon)(\boldsymbol{U}). \end{aligned}$$

The Sinkhorn algorithm is obtained by iteration of the map τ and this is sufficient to guarantee that the limit point of the resulting sequence is an optimizer for the dual problem (5.7), as stated in the following Theorem.

Theorem 5.2.3 (Convergence of the non-commutative Sinkhorn algorithm)**.** *Fix $\varepsilon > 0$. The definition (5.10) of the operators $\mathcal{T}_i^\varepsilon$ is well-posed. Additionally, for any initial matrix $\boldsymbol{U}^{(0)} = (U_1,\dots,U_N) \in \bigtimes_{j=1}^N \mathcal{S}^{\widehat{d}_j}$, there exist $\boldsymbol{\alpha}^k \in \mathbb{R}^N$ with $\sum_{i=1}^N \boldsymbol{\alpha}_i^k = 0$ such that*

$$\boldsymbol{U}^{(k)} := \tau^k(\boldsymbol{U}^{(0)}) + \boldsymbol{\alpha}^k \to \boldsymbol{U}^\varepsilon \quad \textit{as } k \to +\infty, \tag{5.11}$$

where $\boldsymbol{U}^\varepsilon = (U_1^\varepsilon,\dots,U_N^\varepsilon)$ is optimal for the dual problem and $\tau^k := \underbrace{\tau \circ \dots \circ \tau}_{k\text{-times}}$.

Consequently, if one defines for $k \in \mathbb{N}$

$$\Gamma^{(k)} := \exp\left(\frac{\bigoplus_{i=1}^N (\boldsymbol{U}^{(k)})_i - \mathrm{H}_{\mathcal{O}}}{\varepsilon}\right) \quad \text{on } \mathcal{O}, \tag{5.12}$$

and 0 on $\mathcal{O}^\perp$, then $\Gamma^{(k)} \to \Gamma^\varepsilon$ as $k \to +\infty$ where Γ^ε, $\boldsymbol{U}^\varepsilon$ satisfy (5.8). In particular, Γ^ε is optimal for $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$.

Remark 5.2.4 (Renormalisation). In the previous theorem, a renormalisation procedure is needed in order to obtain compactness for the dual potentials $\boldsymbol{U}^k$. Nonetheless, due to the fact that $\sum_{i=1}^N (\boldsymbol{\alpha}^k)_i = 0$ and by the properties of the operator $\bigoplus$, we observe that for $k \in \mathbb{N}$, the equality

$$\Gamma^{(k)} = \exp\left(\frac{\bigoplus_{i=1}^N \tau^k(\boldsymbol{U})_i - \mathrm{H}_{\mathcal{O}}}{\varepsilon}\right) \quad \text{on } \mathcal{O}$$

is also satisfied. In fact, this shows that no renormalisation procedure is needed at the level of the primal problem, i.e. for the density matrices $\Gamma^{(k)}$.

Remark 5.2.5. (Umegaki relative entropies) Similar results can be obtained if instead of the Von Neumann entropy one uses the quantum Umegaki relative entropy with respect to a reference density matrix with trivial kernel. Specifically, suppose that $m_i \in \mathcal{S}^{d_i}$ with $\ker m_i = \{0\}$. Then one can consider the minimisation problem

$$\mathfrak{F}^\varepsilon_{\boldsymbol{m}}(\boldsymbol{\gamma}) = \inf\left\{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon S(\Gamma|\boldsymbol{m}) \ : \ \Gamma \in \mathfrak{P}^{\boldsymbol{d}} \text{ and } \Gamma \mapsto \boldsymbol{\gamma}\right\},$$

where we set $\boldsymbol{m} := \bigotimes_{i=1}^N m_i$ and $S(\Gamma|\boldsymbol{m}) := \mathrm{Tr}(\Gamma(\log\Gamma - \log\boldsymbol{m})$ denotes the relative entropy of Γ with respect to $\boldsymbol{m}$. The functional $\mathfrak{F}^\varepsilon$ defined in (5.6) corresponds to the case $\boldsymbol{m}$ equals the identity matrix. The corresponding dual functional $\mathfrak{D}^\varepsilon$ as defined in (5.7) is replaced by

$$\mathfrak{D}^\varepsilon_{\boldsymbol{m}}(\boldsymbol{\gamma}) = \sup\left\{\sum_{i=1}^N \mathrm{Tr}(U_i\gamma_i) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{\bigoplus_{i=1}^N U_i - \mathrm{H}^\varepsilon_{\boldsymbol{m}}}{\varepsilon}\right]\right) \ : \ U_i \in \mathcal{S}^{d_i}\right\} + \varepsilon,$$

for a modified matrix $\mathrm{H}^\varepsilon_{\boldsymbol{m}} := \mathrm{H} - \varepsilon\log\boldsymbol{m}$ (restricted to $\mathcal{O}$ in the case of non-trival kernels). It is easy to see that our approach can also be used in this case. In particular, performing a change of variables in the dual potentials of the form $\tilde{\boldsymbol{U}} = \boldsymbol{U} + \varepsilon\log\boldsymbol{m}$ and using that $S(\Gamma|\mathrm{Id}) = S(\Gamma|\boldsymbol{m}) + \sum_{i=1}^N [S(\gamma_i) - S(\gamma_i|m_i)]$, one readily derives the validity of the same results obtained in Theorem 5.2.1 and Theorem 5.2.3, with the substitution of H with $\mathrm{H}^\varepsilon_{\boldsymbol{m}}$.

5.2.3 The symmetric case: one-body reduced density matrix functional theory

We are able to obtain the duality results stated above also in the symmetric cases (either bosonic or fermionic). For given $d, N \in \mathbb{N}$, we set $\boldsymbol{d} = d^N$. We consider the bosonic (resp. fermionic) projection operator Π_+ (resp. Π_-)

$$\Pi_+ : \bigotimes_{i=1}^N \mathbb{C}^d \to \bigodot_{i=1}^N \mathbb{C}^d, \qquad \Pi_- : \bigotimes_{i=1}^N \mathbb{C}^d \to \bigwedge_{i=1}^N \mathbb{C}^d, \tag{5.13}$$

where $\odot$ (resp. $\wedge$) denotes the symmetric (resp. antisymmetric) tensor product. Note that the cardinality of $\bigwedge_{i=1}^N \mathbb{C}^d$ is $\binom{d}{N}$, therefore $\bigwedge_{i=1}^N \mathbb{C}^d \neq \{0\}$ if and only if $N \leq d$. We denote by

$$\mathfrak{P}_+^d := \mathfrak{P}\left(\bigodot_{i=1}^N \mathbb{C}^d\right), \quad \mathfrak{P}_-^d := \mathfrak{P}\left(\bigwedge_{i=1}^N \mathbb{C}^d\right), \tag{5.14}$$

the set of bosonic and fermionic density matrices. We fix $\mathrm{H} \in \mathcal{S}^d$ such that

$$\mathrm{S}_i \circ \mathrm{H} \circ \mathrm{S}_i = \mathrm{H}, \quad \forall i = 1, \dots, N, \tag{5.15}$$

where the S_i are the permutation operators in Definition 5.3.3. It is well-known that there exists $\Gamma \in \mathfrak{P}_-^d$ such that $\Gamma \mapsto \gamma$ (where $\Gamma \mapsto \gamma$ means that Γ has all marginals equal to γ) if and only if γ satisfies the *Pauli exclusion principle*, i.e. if and only if $\gamma \in \mathfrak{P}^d$ and $\gamma \leq 1/N$ (see for example [LS10, Theorem 3.2]).

Definition 5.2.6 (Bosonic and fermionic primal problems)**.** For any $\gamma \in \mathfrak{P}^d$, we define the *bosonic* primal problem as

$$\mathfrak{F}_+^\varepsilon(\gamma) := \inf\left\{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon\,\mathrm{Tr}(\Gamma\log\Gamma) \ : \ \Gamma \in \mathfrak{P}_+^d \text{ and } \Gamma \mapsto \gamma\right\}. \tag{5.16}$$

For any $\gamma \in \mathfrak{P}^d$ such that $\gamma \leq 1/N$, we define the *fermionic* primal problem as

$$\mathfrak{F}_-^\varepsilon(\gamma) := \inf\left\{\mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon\,\mathrm{Tr}(\Gamma\log\Gamma) \ : \ \Gamma \in \mathfrak{P}_-^d \text{ and } \Gamma \mapsto \gamma\right\}. \tag{5.17}$$

An analysis of the extremal points and the existence of the minimizer in (5.16) and (5.17) have been carried out in [Col63] for the zero temperature case, and in [GR19] in the positive temperature case. As in the non-symmetric case, we consider the associated bosonic and fermionic dual problems. For any given operator $A \in \mathcal{S}^d$, we denote by $A_\pm$ the corresponding projection onto the symmetric space, obtained as $A_\pm := \Pi_\pm \circ A \circ \Pi_\pm$.

Definition 5.2.7 (Bosonic and fermionic dual problems)**.** For any $\gamma \in \mathfrak{P}^d$, we define the *bosonic* dual functional $\mathrm{D}_\gamma^{+,\varepsilon}$ and the *fermionic* dual functional $\mathrm{D}_\gamma^{-,\varepsilon}$ as

$$\mathrm{D}_\gamma^{\pm,\varepsilon} : \mathcal{S}^d \to \mathbb{R}, \quad \mathrm{D}_\gamma^{\pm,\varepsilon}(U) := \mathrm{Tr}(U\gamma) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{1}{\varepsilon}\left(\frac{1}{N}\bigoplus_{i=1}^N U - H\right)_\pm\right]\right) + \varepsilon\,. \tag{5.18}$$

The corresponding dual problems are given by

$$\mathfrak{D}_\pm^\varepsilon(\gamma) := \sup\left\{\mathrm{D}_\gamma^{\pm,\varepsilon}(U) \ : \ U \in \mathcal{S}^d\right\}. \tag{5.19}$$

We note that a priori $\mathfrak{D}_\varepsilon^-(\gamma)$ can be defined for any $\gamma \in \mathfrak{P}^d$, whereas $\mathfrak{F}_-^\varepsilon(\gamma)$ is only well defined for $\gamma \in \mathfrak{P}^d$ such that $\gamma \leq 1/N$. This constraint on the primal problem naturally translates to an admissibility condition in order to have $\mathfrak{D}_\varepsilon^-(\gamma) < \infty$. To ensure the existence of a maximizer for $\mathrm{D}_\gamma^{-,\varepsilon}$ we further need to impose $\gamma < 1/N$. The following proposition gives an interesting and variational point of view of the Pauli principle, and it is proved in Section 5.6.1.

Proposition 5.2.8 (Pauli's principle and duality)**.** *We have the following equivalences:*

1. $\mathfrak{D}_\varepsilon^-(\gamma) < \infty$ *if and only if* $\gamma \in \mathfrak{P}^d$ *and* $\gamma \leq \frac{1}{N}$,

2. *There exists a maximiser* $U_0 \in \mathcal{S}^d$ *of* $\mathsf{D}^{-,\varepsilon}_\gamma$ *if and only if* $\gamma \in \mathfrak{P}^d$ *and* $0 < \gamma < \frac{1}{N}$.

Finally we state the duality result in the fermionic and bosonic setting.

Theorem 5.2.9 (Fermionic and bosonic duality)**.** *Let* $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$ *satisfying* (5.15).

(i) For any given $\gamma \in \mathfrak{P}^d$, *such that* $\gamma \leq \frac{1}{N}$, *the fermionic primal and dual problems coincide, thus* $\mathfrak{F}^\varepsilon_-(\gamma) = \mathfrak{D}^-_\varepsilon(\gamma)$. *Moreover, if* $0 < \gamma < \frac{1}{N}$ *then* $\mathsf{D}^{-,\varepsilon}_\gamma$ *admits a unique maximizer* U^ε_- *such that*

$$\Gamma^\varepsilon_- = \exp\left(\frac{1}{\varepsilon}\left[\frac{1}{N}\bigoplus_{i=1}^N U^\varepsilon_- - \mathrm{H}\right]_-\right) \tag{5.20}$$

is the unique optimal fermionic solution to the primal problem $\mathfrak{F}^\varepsilon_-(\gamma)$.

(ii) For any given $\gamma \in \mathfrak{P}^d$, *the bosonic primal and dual problems coincide, thus* $\mathfrak{F}^\varepsilon_+(\gamma) = \mathfrak{D}^+_\varepsilon(\gamma)$. *Moreover, if* $\gamma > 0$, $\mathsf{D}^{+,\varepsilon}_\gamma$ *admits a unique maximizer* U^ε_+ *such that*

$$\Gamma^\varepsilon_+ = \exp\left(\frac{1}{\varepsilon}\left[\frac{1}{N}\bigoplus_{i=1}^N U^\varepsilon_+ - \mathrm{H}\right]_+\right) \tag{5.21}$$

is the unique optimal bosonic solution to the primal problem $\mathfrak{F}^\varepsilon_+(\gamma)$.

5.3 Preliminaries and a priori estimates

We start this section by recalling the setting and the notation. For $d \in \mathbb{N}$, we denote by $\mathcal{M}^d = \mathcal{M}^d(\mathbb{C})$ the set of all $d \times d$ complex matrices, by $\mathcal{S}^d$ the hermitian elements of $\mathcal{M}^d$, and by $\mathcal{S}^d_{\geq}$ (respectively $\mathcal{S}^d_{>}$) the set of all the positive semidefinite (positive definite) elements of $\mathcal{S}^d$. With a slight abuse of notation, we denote by Tr the trace operator on $\mathcal{M}_d$ for any dimension d. Furthemore, we denote by $\mathfrak{P}^d$ the set of $d \times d$ *density matrices*, namely the matrices in $\mathcal{S}^d_{\geq}$ with trace one. For the sake of notation, for every $N \in N$ we denote by $[N] := \{1, ..., N\}$.

For a given $N \in \mathbb{N}$ and $(d_i)_{i=1}^N \subset \mathbb{N}$, we consider for any $i \in [N]$ the injective maps

$$\begin{gathered} Q_i : \mathcal{M}^{d_i} \to \mathcal{M}^{\boldsymbol{d}} = \bigotimes_{j=1}^N \mathcal{M}^{d_j}, \quad \boldsymbol{d} := \prod_{j=1}^N d_j, \\ \forall A \in \mathcal{M}^{d_i}, \quad \mathsf{Q}_i(A) := \bigotimes_{j=1}^N A_j, \quad A_j = \begin{cases} A & \text{if } j = i, \\ \mathbb{1} & \text{if } j \neq i. \end{cases} \end{gathered} \tag{5.22}$$

We shall use the same notation also for subsets of $\mathbb{C}^d$. I.e., we also denote by Q_i the map $\mathsf{Q}_i : \mathbb{C}^{d_i} \to \mathbb{C}^{\boldsymbol{d}}$ defined as

$$\forall K \subset \mathbb{C}^{d_i}, \quad \mathsf{Q}_i(K) := \bigotimes_{j=1}^N K_j \subset \mathbb{C}^{\boldsymbol{d}}, \quad K_j = \begin{cases} K & \text{if } j = i, \\ \mathbb{C}^{d_j} & \text{if } j \neq i. \end{cases}$$

The *marginal* operators are the adjoints of the Q_i, namely $\mathsf{P}_i : \mathcal{M}^{\boldsymbol{d}} \to \mathcal{M}^{d_i}$, where for every $\Gamma \in \mathcal{M}^{\boldsymbol{d}}$, $\mathsf{P}_i(\Gamma) \in \mathcal{M}^{d_i}$ is defined by duality as

$$\mathrm{Tr}(\mathsf{P}_i(\Gamma)A) = \mathrm{Tr}\Big(\Gamma\, \mathsf{Q}_i(A)\Big), \quad \forall A \in \mathcal{M}^{d_i}. \tag{5.23}$$

Remark 5.3.1. Observe that $\mathrm{Tr}(\mathsf{P}_i(A)) = \mathrm{Tr}(A)$ for every $i = 1, \dots, N$ and $A \in \mathcal{M}^{\boldsymbol{d}}$. Furthermore, if $A = \bigotimes_{i=1}^N A_i$ with $\mathrm{Tr}(A_i) = 1$, then $\mathsf{P}_i(A) = A_i$.

For a given family of density matrices $\gamma_i \in \mathfrak{P}^{d_i}$, we use the notation $\boldsymbol{\gamma} := (\gamma_i)_{i\in[N]}$ and we write $\Gamma \mapsto \boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)$ whenever $\Gamma \in \mathfrak{P}^{\boldsymbol{d}}$ and $\mathsf{P}_i(\Gamma) = \gamma_i$ for every $i = [N]$. With the next definitions, we introduce the Kronecker sum and Permutation operators.

Definition 5.3.2 (Kronecker sum)**.** For $A_i \in \mathcal{M}^{d_i}$, we call their *Kronecker sum* the matrix

$$\bigoplus_{i=1}^N A_i := \sum_{i=1}^N \mathrm{Q}_i(A_i) \in \mathcal{M}^{\boldsymbol{d}} \tag{5.24}$$

where Q_i is defined in (5.22).

Definition 5.3.3 (Permutation operators)**.** For any $i \in [N]$, we introduce the *permutation operator* $\mathrm{S}_i : \mathcal{M}^{\boldsymbol{d}} \approx \bigotimes_{j=1}^N \mathcal{M}^{d_j} \to \mathcal{M}^{\boldsymbol{d}}$ as the map definitioned by

$$\mathrm{S}_i\left(\bigotimes_{j=1}^N A_j\right) = A_1 \otimes \cdots \otimes A_{i-1} \otimes A_{i+1} \otimes \cdots \otimes A_N \otimes A_i,$$

for any $A_i \in \mathcal{M}^{d_i}$ and extended to the whole $\mathcal{M}^{\boldsymbol{d}}$ by linearity.

Remark 5.3.4. The permutation operators preserve the spectral properties of any operator. Precisely, $\sigma(\mathrm{S}_i(A)) = \sigma(A)$ for every $i \in [N]$, $A \in \mathcal{S}^{\boldsymbol{d}}$, where $\sigma(A)$ denotes the spectrum of A. In particular, for every continuous function $f : \mathbb{R} \to \mathbb{R}$, we have that $\mathrm{Tr}(f(\mathrm{S}_i(A))) = \mathrm{Tr}(f(A))$, for every $A \in \mathcal{S}^{\boldsymbol{d}}$.

5.3.1 Non-commutative $(\mathrm{H}, \varepsilon)$-transforms

For this section, we specify to the simply case of a two-fold tensor product and introduce the notion of non-commutative $(\mathrm{H}, \varepsilon)$-transform, which is a central object in our discussion. We shall see in Section 5.3.2 how it is then easy to extend this notion to a general N-fold tensor product. We fix $d, d' \in \mathbb{N}$, $0 < \alpha \in \mathfrak{P}^{d'}$, $\mathrm{H} \in \mathcal{S}^{dd'}$ and $\varepsilon > 0$ and define the map $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}} : \mathcal{S}^d \times \mathcal{S}^{d'} \to \mathbb{R}$ as

$$\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, V) := \mathrm{Tr}(V\alpha) - \varepsilon \,\mathrm{Tr}\left(\exp\left[\frac{U \oplus V - \mathrm{H}}{\varepsilon}\right]\right). \tag{5.25}$$

The $(\mathrm{H}, \varepsilon)$-transform of any $U \in \mathcal{S}^d$ is obtained as the maximiser of the map $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, \cdot)$.

Definition 5.3.5 ($(\mathrm{H}, \varepsilon)$-transform)**.** We call the unique maximizer of $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, \cdot)$ the $(\mathrm{H}, \varepsilon)$-*transform* of $U \in \mathcal{S}^d$. We use the notation

$$\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}} : \mathcal{S}^d \to \mathcal{S}^{d'}, \quad \mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U) = \text{arg max}\{\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, V) \ : \ V \in \mathcal{S}^{d'}\}. \tag{5.26}$$

The following lemma shows that the definition of $(\mathrm{H}, \varepsilon)$-transform is indeed well-posed.

Lemma 5.3.6. *Let $U \in \mathcal{S}^d$. Then there exists a unique maximizer $\bar{V} \in \mathcal{S}^{d'}$ of $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, \cdot)$.*

Proof. Fix $U \in \mathcal{S}^d$. For every $V \in \mathcal{S}^{d'}$, we write $V = V_+ - V_-$ where $V_+, V_- \in \mathcal{S}^{d'}$ denote respectively the positive and the negative part of V (with respect to its spectrum). We begin by observing that

$$\operatorname{Tr}\left(\exp\left[\frac{U \oplus V - \mathrm{H}}{\varepsilon}\right]\right) \geq \operatorname{Tr}\left(\exp\left[\frac{U \oplus V - \|\mathrm{H}\|_\infty}{\varepsilon}\right]\right)$$
$$= \operatorname{Tr}\left(\exp\left(\frac{V}{\varepsilon}\right)\right)\operatorname{Tr}\left(\exp\left(\frac{U}{\varepsilon}\right)\right)\exp\left[\frac{-\|\mathrm{H}\|_\infty}{\varepsilon}\right] =: \kappa \operatorname{Tr}\left(\exp\left(\frac{V}{\varepsilon}\right)\right) \geq \kappa e^{\varepsilon^{-1}\|V_+\|_\infty}, \tag{5.27}$$

where in the second step we used that $\exp(U \oplus V) = \exp(U) \otimes \exp(V)$ and $\kappa = \kappa(U, \varepsilon, \mathrm{H})$ is a *finite* constant depending on U, ε, and H. On the other hand, it clearly holds $\operatorname{Tr}(V\alpha) \leq \|V_+\|_\infty$ which combined with (5.27) yields for every $V \in \mathcal{S}^{d'}$

$$\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, V) \leq \|V_+\|_\infty - \kappa e^{\varepsilon^{-1}\|V_+\|_\infty}. \tag{5.28}$$

Moreover, it is immediate to obtain that

$$\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, V) \leq \operatorname{Tr}(V\alpha) = \operatorname{Tr}(V_+\alpha) - \operatorname{Tr}(V_-\alpha) \leq \|V_+\|_\infty - \sigma_{\min}(\alpha)\|V_-\|_\infty, \tag{5.29}$$

where $\sigma_{\min}(\alpha)$ is the spectral gap of α, which is strictly positive by assumption. Let V_n be a maximizing sequence for $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, \cdot)$, then the bounds (5.28) and (5.29) imply that $(V_n)_+$, $(V_n)_-$ (and hence V_n) are uniformly bounded. Therefore, we can obtain a subsequence (which we do not relabel) such that $V_n \to \bar{V} \in \mathcal{S}^{d'}$. The optimality of $\bar{V}$ follows from the fact that $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U, \cdot)$ is continuous and strictly concave (see for example [Car10]), which also implies uniqueness. □

In the following lemma we use the fact that the $(\mathrm{H}, \varepsilon)$-transform is obtained through a maximization to show that it can be characterized as the solution of the associated Euler–Lagrange equation. This property is crucial for the proof of our main results.

Lemma 5.3.7 (Optimality conditions for the $(\mathrm{H}, \varepsilon)$-transforms). *Given $d, d' \in \mathbb{N}$, $0 < \alpha \in \mathfrak{P}^{d'}$, $\mathrm{H} \in \mathcal{S}^{dd'}$, $\varepsilon > 0$, the operator $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}$ can be characterized implicitly by the fact that, for any $U \in \mathcal{S}^d$, $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)$ is the unique solution of*

$$\alpha = \mathsf{P}_2\left(\exp\left[\frac{U \oplus \mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U) - \mathrm{H}}{\varepsilon}\right]\right). \tag{5.30}$$

Proof. Let us pick any $\Lambda \in \mathcal{S}^d$ and define $V_s := \mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U) + s\Lambda$. By construction, due to the optimality of $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)$, the map

$$s \mapsto g(s) := \operatorname{Tr}(V_s\alpha) - \varepsilon \operatorname{Tr}\left(\exp\left[\frac{U \oplus V_s - \mathrm{H}}{\varepsilon}\right]\right)$$

must have vanishing derivative at $s = 0$. This can be computed [Car10, Section 2.2] as

$$g'(0) = \operatorname{Tr}(\Lambda\alpha) - \varepsilon \operatorname{Tr}\left((I \otimes \Lambda)\exp\left[\frac{U \oplus \mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U) - \mathrm{H}}{\varepsilon}\right]\right). \tag{5.31}$$

Using the definition of partial trace and the previous formula, we infer

$$\mathrm{Tr}\left(\Lambda\Big(\alpha - \mathsf{P}_2\Big(\exp\left[\frac{U\oplus\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)-\mathrm{H}}{\varepsilon}\right]\Big)\Big)\right)=0 \tag{5.32}$$

for every $\Lambda\in\mathcal{S}^d$. Note that $\alpha, U, V_s, \mathrm{H}$ being self-adjoint, it follows that the operator

$$\alpha-\mathsf{P}_2\left(\exp\left[\frac{U\oplus\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)-\mathrm{H}}{\varepsilon}\right]\right)$$

is self-adjoint as well. Together with (5.32), this shows (5.30). On the other hand, since (5.32) is the Euler Lagrange equation associated to the maximization of the strictly concave functional $\mathrm{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U,\cdot)$, any solution of (5.32) is necessarily a maximizer and hence coincides with $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)$, by uniqueness (see Lemma 5.3.6).

□

The next step is to obtain some regularity estimates on $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)$. To do so, we extract information from the optimality conditions proved in Lemma 5.3.7.

Proposition 5.3.8 (Regularity of the (H,ε)-transform). *Given $d,d'\in\mathbb{N}$, $0<\alpha\in\mathfrak{P}^{d'}$, $\mathrm{H}\in\mathcal{S}^{dd'}$, $\varepsilon>0$, we define for all $A\in\mathcal{S}^d$ (or $A\in\mathcal{S}^{d'}$)*

$$\lambda_\varepsilon(A):=\varepsilon\log\left(\mathrm{Tr}\left[\exp\left(\frac{A}{\varepsilon}\right)\right]\right). \tag{5.33}$$

Then for every $U\in\mathcal{S}^d$ it holds

$$\left|\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)-\varepsilon\log\alpha+\lambda_\varepsilon(U)\mathbb{1}\right|\le\|H\|_\infty\mathbb{1}, \tag{5.34}$$

$$\left|\lambda_\varepsilon(U)+\lambda_\varepsilon\Big(\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)\Big)\right|\le\|H\|_\infty, \tag{5.35}$$

$$\left|\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)-\varepsilon\log\alpha-\lambda_\varepsilon(\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U))\mathbb{1}\right|\le 2\|H\|_\infty\mathbb{1}. \tag{5.36}$$

where the inequalities are understood as two-sided quadratic forms bounds.

Proof. Note that (5.36) is an immediate consequence of (5.34) and (5.35) and we shall therefore only prove the latter two. Let us start with the proof of (5.34). We know from Lemma 5.3.7 that for every $U\in\mathcal{S}^d$, $\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)$ satisfies equation (5.30). By the properties of the partial trace (Remark 5.3.1) and $\mathrm{H}\le\|\mathrm{H}\|_\infty\mathbb{1}$, it follows that

$$\begin{aligned}\alpha&\le e^{\frac{\|\mathrm{H}\|_\infty}{\varepsilon}}\mathsf{P}_2\left(\exp\left[\frac{U\oplus\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right]\right)\\&=e^{\frac{\|\mathrm{H}\|_\infty}{\varepsilon}}\mathsf{P}_2\left(\exp\left(\frac{U}{\varepsilon}\right)\otimes\exp\left(\frac{\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right)\right)\\&=e^{\frac{\|\mathrm{H}\|_\infty}{\varepsilon}}\mathrm{Tr}\left(\exp\left(\frac{U}{\varepsilon}\right)\right)\exp\left(\frac{\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right),\end{aligned} \tag{5.37}$$

where in the first equality we used that $\exp(A\oplus B)=\exp A\otimes\exp B$. Similarly, using instead the lower bound $\mathrm{H}\ge-\|\mathrm{H}\|_\infty\mathbb{1}$, from (5.30) we can also obtain

$$\alpha\ge e^{\frac{-\|\mathrm{H}\|_\infty}{\varepsilon}}\mathrm{Tr}\left(\exp\left(\frac{U}{\varepsilon}\right)\right)\exp\left(\frac{\mathfrak{T}^{\varepsilon}_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right). \tag{5.38}$$

We can put together the two bounds in (5.37), (5.38) to obtain

$$\alpha e^{\frac{-\|\mathrm{H}\|_\infty}{\varepsilon}} \leq \mathrm{Tr}\left(\exp\left(\frac{U}{\varepsilon}\right)\right)\exp\left(\frac{\mathfrak{T}^\varepsilon_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right) \leq \alpha e^{\frac{\|\mathrm{H}\|_\infty}{\varepsilon}}. \tag{5.39}$$

Taking the log in the latter inequalities we conclude the proof of (5.34). If we instead take the trace of both sides in (5.39), we obtain

$$e^{\frac{-\|\mathrm{H}\|_\infty}{\varepsilon}} \leq \mathrm{Tr}\left(\exp\left(\frac{U}{\varepsilon}\right)\right)\mathrm{Tr}\left(\exp\left(\frac{\mathfrak{T}^\varepsilon_{\alpha,\mathrm{H}}(U)}{\varepsilon}\right)\right) \leq e^{\frac{\|\mathrm{H}\|_\infty}{\varepsilon}},$$

and then applying the log, we conclude the proof (5.35). □

5.3.2 Vectorial (H,ε)-transforms

In this section, we consider a vectorial version of the (H,ε)-transforms introduced in the previous section. This turns out to be a key object in the proof of Theorem 5.2.1 and Theorem 5.2.3, necessary to deal with the multi-marginal setting.

Let us first introduce the general framework, which remains in force throughout the section. Let $N \in \mathbb{N}$ and $[N]$ be a index set of N elements. For all $i \in [N]$, let $d_i \in \mathbb{N}$ and $\gamma_i \in \mathfrak{P}^{d_i}$ be density matrices. Set $\boldsymbol{\gamma} := (\gamma_i)_{i\in[N]}$, $\boldsymbol{d} = \prod_{j=1}^N d_i$. Finally, consider a Hamiltonian $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$.

Remark 5.3.9. (Kernels) Without loss of generality, we can assume $\ker\gamma_i = \{0\}$, for every $i \in [N]$. In the general case, it suffices to consider the restriction to the set $\mathcal{O} := \bigotimes_{i=1}^N \left(\ker\gamma_i\right)^\perp$ and consider the matrix $\mathrm{H}_{\mathcal{O}} = \Pi_{\mathcal{O}}\,\mathrm{H}\,\Pi_{\mathcal{O}}$, where $\Pi_{\mathcal{O}}$ is the projector onto $\mathcal{O}$.

We therefore assume that $\ker\gamma_i = \{0\}$ for all $i \in [N]$. In this section we extend the notion of (H,ε)-transform as introduced in previous section 5.3.1 to the multi-marginal setting, and we apply it to our specific setting. We are interested in the maximization (5.7) of the dual functional, that we introduce below.

Definition 5.3.10 (Dual Functional)**.** For any $\boldsymbol{U} = (U_1,\dots,U_N) \in \bigtimes_{j=1}^N \mathcal{S}^{d_j}$, we define

$$\mathrm{D}^\varepsilon_{\boldsymbol{\gamma}}(\boldsymbol{U}) = \sum_{i=1}^N \mathrm{Tr}(U_i\gamma_i) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{\bigoplus_{i=1}^N U_i - \mathrm{H}}{\varepsilon}\right]\right) + \varepsilon.$$

Remark 5.3.11. Note that $\mathrm{D}^\varepsilon_{\boldsymbol{\gamma}}$ is invariant by translation for any vector $\boldsymbol{a} = (a_1,\dots,a_N) \in \mathbb{R}^N$ such that $\sum_{k=1}^N a_k = 0$, i.e.

$$\mathrm{D}^\varepsilon_{\boldsymbol{\gamma}}(\boldsymbol{U}+\boldsymbol{a}) = \mathrm{D}^\varepsilon_{\boldsymbol{\gamma}}(\boldsymbol{U}).$$

As a consequence of this property, we see in Section 5.5 that the set of maximizers is invariant by such transformations (Lemma 5.4.4).

With the following definition, we introduce the vectorial (H,ε)-transforms.

Definition 5.3.12 (Vectorial (H,ε)-transform)**.** For any $i \in [N]$, we define the *i-th vectorial (H,ε)-transform* $\mathfrak{T}^\varepsilon_i$ as the map

$$\mathfrak{T}^\varepsilon_i : \bigtimes_{j=1, j\neq i}^N \mathcal{S}^{d_j} \to \mathcal{S}^{d_i},$$

$$\mathfrak{T}^\varepsilon_i(\widehat{\boldsymbol{U}}_i) = \underset{V\in\mathcal{S}^{d_i}}{\mathrm{argmax}}\left\{\mathrm{Tr}(V\gamma_i) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{1}{\varepsilon}\left(U_1\oplus\cdots\oplus U_{i-1}\oplus V\oplus U_{i+1}\oplus\cdots\oplus U_N - \mathrm{H}\right)\right]\right)\right\},$$

where for $\boldsymbol{U} \in \times_{j=1}^{N} \mathcal{S}^{d_j}$, we set $\widehat{\boldsymbol{U}}_i$ to be the product of all the U_j but the i-th one, namely

$$\widehat{\boldsymbol{U}}_i := \Big(U_1, \dots, U_{j-1}, U_{j+1}, \dots, U_N\Big) \in \bigtimes_{j=1,\, j\neq i}^{N} \mathcal{S}^{d_j}. \tag{5.40}$$

Remark 5.3.13. Observe that we can identify the i-th vectorial $(\mathrm{H}, \varepsilon)$-transforms with a particular case of the operators $\mathfrak{T}_{\varepsilon,\mathrm{H},\alpha}$ as introduced in Section 5.3.1. Indeed, as a consequence of Remark 5.3.4 it is straightforward to see that for $i \in [N]$

$$\mathfrak{T}_i^{\varepsilon}(\widehat{\boldsymbol{U}}_i) = \mathfrak{T}^{\varepsilon}_{\gamma_i, \mathrm{S}_i(\mathrm{H})}\left(\bigoplus_{j=1,\, j\neq i}^{N} U_i\right), \quad \mathfrak{T}^{\varepsilon}_{\gamma_i, \mathrm{S}_i(\mathrm{H})} : \bigotimes_{j=1,\, j\neq i}^{N} \mathcal{S}^{d_j} \approx \mathcal{S}^{\tilde{d}_i} \to \mathcal{S}^{d_i}, \tag{5.41}$$

where we set $\tilde{d}_i := \prod_{j\neq i} d_j$ and the S_i are the permutation operators in Definition 5.3.3. This shows that the definition is well posed (i.e. that the argmax appearing in the definition exists and is unique). Moreover it allows us to extend the validity of the properties of the $(\mathrm{H}, \varepsilon)$-transform shown in Section 5.3.1 to the operators $\mathfrak{T}_i^{\varepsilon}$, as we shall see in Lemma 5.3.18 and Proposition 5.3.19 below. Note that the dependence on the specific entry i is reflected in both the use of γ_i and in the fact that the transform is performed w.r.t. $\mathrm{S}_i(\mathrm{H})$.

5.3.3 One-step and Sinkhorn operators

We use the vectorial $(\mathrm{H}, \varepsilon)$-transforms to define what we call *one-step operators* and *Sinkhorn operators*. The first ones map a vector of N potentials into a vector of N potentials, exchanging its i-th entry with the i-th vectorial $(\mathrm{H}, \varepsilon)$-transform applied to the other $N-1$. The second is simply obtained by composing all the different N one-step operators.

Definition 5.3.14 (One-step operators)**.** For $i \in [N]$, we introduce the *one-step operators* $\mathcal{T}_i^{\varepsilon}$, which are defined by

$$\mathcal{T}_i^{\varepsilon} : \bigtimes_{j=1}^{N} \mathcal{S}^{d_j} \to \bigtimes_{j=1}^{N} \mathcal{S}^{d_j}$$
$$\boldsymbol{U} := (U_1, \dots, U_N) \longmapsto (U_1, \dots, U_{i-1}, \mathfrak{T}_i^{\varepsilon}(\widehat{\boldsymbol{U}}_i), U_{i+1}, \dots, U_N) =: \mathcal{T}_i^{\varepsilon}(\boldsymbol{U}).$$

The Sinkhorn operator is simply the composition of the N one-step operators $\mathcal{T}_i^{\varepsilon}, i \in [N]$.

Definition 5.3.15 (Sinkhorn Operator)**.** We introduce the *Sinkhorn operator* τ, defined by

$$\tau : \bigtimes_{j=1}^{N} \mathcal{S}^{d_j} \to \bigtimes_{j=1}^{N} \mathcal{S}^{d_j},$$
$$\tau(\boldsymbol{U}) := (\mathcal{T}_N^{\varepsilon} \circ \dots \circ \mathcal{T}_1^{\varepsilon})(\boldsymbol{U}).$$

Remark 5.3.16. Note that, by definition of τ, it follows immediately that, for any $\boldsymbol{U} \in \times_{j=1}^{N} \mathcal{S}^{d_j}$

$$\mathrm{D}_{\gamma}^{\varepsilon}(\tau(\boldsymbol{U})) \geq \mathrm{D}_{\gamma}^{\varepsilon}(\boldsymbol{U}),$$

i.e. applying τ to any vector increases its energy. Moreover, any maximizer of $\mathrm{D}_{\gamma}^{\varepsilon}$ is a fixed point of τ (as a consequence of the uniqueness proved in Lemma 5.3.6). The converse is also true and implies that the set of maximizers of $\mathrm{D}_{\gamma}^{\varepsilon}$ coincides with the set of fixed points of τ, see Remark 5.4.3.

Remark 5.3.17. Note that for any vector $\boldsymbol{a} \in \mathbb{R}^N$ such that $\sum_{k=1}^N \boldsymbol{a}_k = 0$, one has

$$\mathcal{T}_i^\varepsilon(\boldsymbol{U} + \boldsymbol{a}) = \mathcal{T}_i^\varepsilon(\boldsymbol{U}) + \boldsymbol{a},$$

i.e. $\mathcal{T}_i^\varepsilon$ commutes with translations by vectors whose coordinates sum up to zero (notice that this fact is particularly interesting in light of Remark 5.3.11). This is a straightforward consequence of the fact that

$$\mathfrak{T}_i^\varepsilon\left((\widehat{\boldsymbol{U} + \boldsymbol{a}})_i\right) = \mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) + a_i,$$

which can be readily verified from the definitions. Trivially, this also implies

$$\tau(\boldsymbol{U} + \boldsymbol{a}) = \tau(\boldsymbol{U}) + \boldsymbol{a}.$$

We now take advantage of the observations in Remark 5.3.13 to deduce properties for the vectorial $(\mathrm{H}, \varepsilon)$-transforms, the one-step operators, and the Sinkhorn operator. First of all, as a corollary of Lemma 5.3.7, we characterize the vectorial $(\mathrm{H}, \varepsilon)$-transforms as solutions of implicit equations.

Lemma 5.3.18 (Optimality conditions for vectorial $(\mathrm{H}, \varepsilon)$-transforms)**.** *Let $i \in [N]$, $\varepsilon > 0$, $\gamma_i \in \mathfrak{P}^{d_i}$, $\mathrm{H} \in \mathcal{S}^d$, with $\ker \gamma_i = \{0\}$. For any $\boldsymbol{U} \in \times_{j=1}^N \mathcal{S}^{d_j}$, the one step-operator $\mathcal{T}_i^\varepsilon(\boldsymbol{U})$ (or equivalently the i-th vectorial $(\mathrm{H}, \varepsilon)$-transform $\mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i)$) is implicitly characterized as the unique solution of the equation*

$$\gamma_i = P_i\left(\exp\left[\frac{1}{\varepsilon}\left(\bigoplus_{j=1}^N (\mathcal{T}_i^\varepsilon(\boldsymbol{U}))_j - \mathrm{H}\right)\right]\right). \tag{5.42}$$

Proof. As a consequence of (5.41), we can apply Lemma 5.3.7 and deduce

$$\begin{aligned}\gamma_i &= \mathrm{P}_N\left(\exp\left[\frac{1}{\varepsilon}\left(\left(\bigoplus_{j=1, j\neq i}^N U_i\right) \oplus \mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) - \mathrm{S}_i(\mathrm{H})\right)\right]\right)\\ &= \mathrm{P}_i\left(\exp\left[\frac{1}{\varepsilon}\left(\bigoplus_{j=1}^N (\mathcal{T}_i^\varepsilon(\boldsymbol{U}))_j - \mathrm{H}\right)\right]\right),\end{aligned}$$

where S_i is the i-th permutation operator, as defined in 5.3.3, and in the last equality we used Remark 5.3.4 and that

$$\left(\bigoplus_{j=1, j\neq i}^N U_i\right) \oplus \mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) = \mathrm{S}_i\left(\bigoplus_{j=1}^N (\mathcal{T}_i^\varepsilon(\boldsymbol{U}))_j\right)$$

for every $i \in [N]$ and $\boldsymbol{U} \in \times_{j=1}^N \mathcal{S}^{d_j}$. □

The next proposition collects the regularity properties of the $(\mathrm{H}, \varepsilon)$-transforms. Once again, they are direct consequence of the properties proved in the two marginals case, in particular in Proposition 5.3.8.

Proposition 5.3.19 (Regularity of the (H,ε)-transforms)**.** *Let $i \in [N]$, $\varepsilon > 0$, $\gamma_i \in \mathfrak{P}^{d_i}$, $\mathrm{H} \in \mathcal{S}^d$, with $\ker \gamma_i = \{0\}$. Then for every $\boldsymbol{U} \in \times_{j=1}^N \mathcal{S}^{d_j}$, for every $i \in [N]$ it holds*

$$\left| \mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) - \varepsilon \log \gamma_i + \sum_{j=1, j\neq i}^N \lambda_\varepsilon(U_j) \mathbb{1} \right| \leq \|H\|_\infty \mathbb{1}\,, \tag{5.43}$$

$$\left| \sum_{j=1, j\neq i}^N \lambda_\varepsilon(U_j) + \lambda_\varepsilon\left(\mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i)\right) \right| \leq \|H\|_\infty\,, \tag{5.44}$$

$$\left| \mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) - \varepsilon \log \gamma_i - \lambda_\varepsilon\left(\mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i)\right) \mathbb{1} \right| \leq (2\|H\|_\infty) \mathbb{1}, \tag{5.45}$$

where λ_ε is defined in (5.33).

Proof. The proof is a direct application of Proposition 5.3.8 and the considerations in Remark 5.3.13. Precisely, the estimate (5.43) follows from (5.34), (5.44) follows from (5.35) and (5.45) follows from (5.36), together with the fact that

$$\lambda_\varepsilon\left(\bigoplus_{j=1, j\neq i}^N U_j \right) = \sum_{j=1, j\neq i}^N \lambda_\varepsilon(U_j)\,, \quad \forall \boldsymbol{U} \in \bigtimes_{j=1}^N \mathcal{S}^{d_j}.$$

□

In light of Remark 5.3.16, it is reasonable to check whether sequences of the form $\tau^k(\boldsymbol{U}_0)$ are maximizing for $\mathrm{D}_\gamma^\varepsilon$ and compact. On the other hand, a priori it is not clear how to obtain compactness for such sequences and Remark 5.3.11 shows that there could even exist sequences 'converging' to the set of maximizers which are not compact. It is therefore natural to introduce a suitable renormalization operator, aimed at retrieving compactness. Note that any such operator should increase or leave invariant the value of $\mathrm{D}_\gamma^\varepsilon$ and therefore, by Remark 5.3.11, any translation by vectors whose coordinates sum up to zero is a good candidate.

Definition 5.3.20 (Renormalisation)**.** Let λ_ε be defined as in (5.33). We define the *renormalisation map* $\mathrm{Ren} : \times_{i=1}^N \mathcal{S}^{d_i} \to \times_{i=1}^N \mathcal{S}^{d_i}$ as the function

$$\mathrm{Ren}(\boldsymbol{U})_i = \begin{cases} U_i - \lambda_\varepsilon(U_i), & \text{if } i \in \{1, \ldots, N-1\} \\ U_N + \sum_{j=1}^{N-1} \lambda_\varepsilon(U_j), & \text{if } i = N. \end{cases}$$

Note the choice of the N-component of the renormalisation operator is chosen in such a way that the translation is associated to a vector whose coordinates sum up to zero.

In the following proposition we show that $\mathrm{Ren}\left(\tau\left(\times_{i=1}^N \mathcal{S}^{d_i}\right)\right)$ is bounded and therefore compact. This shows that the map Ren is indeed a reasonable renormalization operator for our purposes.

Proposition 5.3.21 (Renormalisation of (H,ε)-transforms and uniform bounds)**.** *Let $i \in [N]$, $\varepsilon > 0$, $\gamma_i \in \mathfrak{P}^{d_i}$, $\mathrm{H} \in \mathcal{S}^d$, with $\ker \gamma_i = \{0\}$. Then, for any $\boldsymbol{U} \in \mathcal{S}^d$, one has that $\mathrm{D}_\gamma^\varepsilon(\mathrm{Ren}\,\tau(\boldsymbol{U})) \geq \mathrm{D}_\gamma^\varepsilon(\boldsymbol{U})$, and the following bounds hold true:*

$$\left| (\mathrm{Ren}\,\tau(\boldsymbol{U}))_i - \varepsilon \log \gamma_i \right| \leq 2\|H\|_\infty \mathbb{1}, \quad \forall i \in [N]. \tag{5.46}$$

Proof. First of all, Remark 5.3.11 and Remark 5.3.16 trivially yield $\mathrm{D}^\varepsilon_\gamma(\mathrm{Ren}\,\tau(\boldsymbol{U})) \geq \mathrm{D}^\varepsilon_\gamma(\boldsymbol{U})$. To show (5.46), note that for any $i \in [N]$, $(\tau(\boldsymbol{U}))_i$ is obtained applying $\mathfrak{T}^\varepsilon_i$ to some element of $\bigtimes_{j=1, j\neq i}^N \mathcal{S}^{d_j}$. Therefore, applying (5.45) from Proposition 5.3.19, we obtain

$$\|(\mathrm{Ren}\,\tau(\boldsymbol{U}))_i - \varepsilon \log \gamma_i\|_\infty = \|(\tau(\boldsymbol{U}))_i - \varepsilon \log \gamma_i - \lambda_\varepsilon((\tau(\boldsymbol{U}))_i)\|_\infty \leq 2\|\,\mathrm{H}\,\|_\infty$$

for every $i \in [N-1]$. Moreover, $(\tau(\boldsymbol{U}))_N = \mathfrak{T}^N_\varepsilon(\widehat{\tau(\boldsymbol{U})}_N)$ and hence, applying (5.43) from Proposition 5.3.19, we arrive at

$$\|(\mathrm{Ren}\,\tau(\boldsymbol{U}))_N - \varepsilon \log \gamma_N\|_\infty = \left\|(\tau(\boldsymbol{U}))_N - \varepsilon \log \gamma_N + \sum_{j=1}^{N-1} \lambda_\varepsilon((\tau(\boldsymbol{U}))_j)\right\|_\infty \leq \|\,\mathrm{H}\,\|_\infty,$$

which completes the proof. □

5.4 Non-commutative multi-marginal optimal transport

In this section we prove Theorem 5.2.1, our first main result stated in Section 5.2, exploiting the tools developed in Section 5.3. Again, we fix the setup, which remains in force throughout the whole Section 5.4 and Section 5.5. Let $N \in \mathbb{N}$, and for $i \in [N]$ we consider density matrices $\gamma_i \in \mathfrak{P}^{d_i}$. Set $\boldsymbol{\gamma} := (\gamma_i)_{i\in[N]}$, $\boldsymbol{d} = \prod_{j=1}^N d_i$, and assume that $\ker \gamma_i = \{0\}$ (see Remark 5.3.9). We also fix $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$. In this section, we prove the Theorem 5.2.1.

We begin by introducing the primal functional, which appears in the minimisation (5.6).

Definition 5.4.1 (Primal Functional)**.** For $\Gamma \in \mathfrak{P}^{\mathbf{d}}$, the primal functional is defined by

$$\mathrm{F}^\varepsilon(\Gamma) = \mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon S(\Gamma) = \mathrm{Tr}(\mathrm{H}\,\Gamma) + \varepsilon\,\mathrm{Tr}(\Gamma \log \Gamma). \tag{5.47}$$

We also recall the definitions of the primal and the dual problem

$$\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \inf\left\{\mathrm{F}^\varepsilon(\Gamma) \;:\; \Gamma \in \mathfrak{P}^{\mathbf{d}} \text{ and } \Gamma \mapsto (\gamma_1, \dots, \gamma_N)\right\}, \tag{5.48}$$

$$\mathfrak{D}^\varepsilon(\boldsymbol{\gamma}) = \sup\left\{\mathrm{D}^\varepsilon_\gamma(\boldsymbol{U}) \;:\; \boldsymbol{U} \in \bigtimes_{i=1}^N \mathcal{S}^{d_i}\right\}, \tag{5.49}$$

where the dual functional $\mathrm{D}^\varepsilon_\gamma$ is given in Definition 5.3.10.

5.4.1 Primal and dual functionals: lower bound and structure of the optimizers

We begin with the proof of the lower bound for the primal functional (5.47), in terms of the dual functional (5.3.10).

Proposition 5.4.2 (Lower bound)**.** *Fix $N \in \mathbb{N}$ and $\varepsilon > 0$. For all $i \in [N]$, let $\gamma_i \in \mathfrak{P}^{d_i}$ be density matrices, $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$. Then, for all $\boldsymbol{U} \in \bigtimes_{i=1}^N \mathcal{S}^{d_i}$ and every $\Gamma \in \mathfrak{P}^{\boldsymbol{d}}$, $\Gamma \mapsto \boldsymbol{\gamma}$ we have that*

$$\mathrm{F}^\varepsilon(\Gamma) \geq \mathrm{D}^\varepsilon_\gamma(\boldsymbol{U})\,.$$

Proof. For any $\boldsymbol{U} \in \bigtimes_{i=1}^{N} \mathcal{S}^{d_i}$ and any admissible $\Gamma \in \mathfrak{P}^{\boldsymbol{d}}$, $\Gamma \mapsto \boldsymbol{\gamma}$, we can write

$$\begin{aligned}\mathrm{F}^{\varepsilon}(\Gamma) &= \mathrm{F}^{\varepsilon}(\Gamma) + \sum_{j=1}^{N} \mathrm{Tr}(U_j\gamma_j) - \mathrm{Tr}\left(\left(\bigoplus_{j=1}^{N} U_j\right)\Gamma\right) \\ &= \sum_{j=1}^{N} \mathrm{Tr}(U_j\gamma_j) + \varepsilon S(\Gamma) - \mathrm{Tr}\left(\Gamma\left(\bigoplus_{j=1}^{N} U_j - \mathrm{H}\right)\right).\end{aligned}$$

Let us denote the Hilbert-Schmidt scalar product (on $\mathcal{M}^{\boldsymbol{d}}$) by $\langle\cdot,\cdot\rangle_{HS}$. It follows that

$$\mathrm{F}^{\varepsilon}(\Gamma) = \sum_{j=1}^{N} \mathrm{Tr}(U_j\gamma_j) + \varepsilon\left[S(\Gamma) - \langle \Gamma, \overline{Y}\rangle_{HS}\right] \geq \sum_{j=1}^{N} \mathrm{Tr}(U_j\gamma_j) - \varepsilon S^*(\overline{Y}), \tag{5.50}$$

where $\overline{Y} = \varepsilon^{-1}\left(\bigoplus_{j=1}^{N} U_j - \mathrm{H}\right) \in \mathcal{S}^{\boldsymbol{d}}$ and, for any $Y \in \mathcal{S}^{\boldsymbol{d}}$

$$S^*(Y) := \sup_{\Gamma\in\mathcal{S}^{\boldsymbol{d}}_{\geq}} \{\langle Y, \Gamma\rangle_{HS} - S(\Gamma)\}$$

denotes the Legendre transform of S on the subspace $\mathcal{S}^{\boldsymbol{d}}_{\geq}$. This can be explicitly computed as

$$S^*(Y) = \mathrm{Tr}\left[\exp(Y-1)\right], \quad \forall Y \in \mathcal{S}^{\boldsymbol{d}}. \tag{5.51}$$

For the sake of completeness, let us explain how to prove (5.51). First of all we show that for any $Y \in \mathcal{S}^{\boldsymbol{d}}$ the supremum appearing in the definition of $S^*(Y)$ is attained at some $\bar{\Gamma} \in \mathcal{S}^{\boldsymbol{d}}_{>}$. Indeed, for any $\Gamma \geq 0$ define σ_+ to be the maximum of its spectrum, then it holds

$$\langle Y, \Gamma\rangle_{HS} - S(\Gamma) \leq \boldsymbol{d}^2\|Y\|_\infty\sigma_+ - \sigma_+\log\sigma_+ - \min_{\mathbb{R}_+}\{x\log x\}(\boldsymbol{d}^2-1) \xrightarrow{\sigma_+\to\infty} -\infty.$$

This implies that the super-levels of $\langle y, \Gamma\rangle_{HS} - f(\Gamma)$ are bounded and hence pre-compact and allows us to conclude the existence of a maximizer $\bar{\Gamma}$. Moreover, it is straightforward to show that $\bar{\Gamma} > 0$, otherwise one would have a contradiction by perturbing $\bar{\Gamma}$ with $\Pi_{\ker\bar{\Gamma}}$ (the projector onto $\ker\bar{\Gamma}$).

Let us derive the optimality conditions for $\bar{\Gamma}$. Define $\Gamma_s := \bar{\Gamma} + s\Gamma'$ with $\Gamma' \in \mathcal{S}^{\boldsymbol{d}}$ (note that for any $\Gamma' \in \mathcal{S}^{\boldsymbol{d}}$ for s sufficiently small Γ_s is positive since $\bar{\Gamma} > 0$), then the Euler-Lagrange equation for the maximization problem reads

$$0 = \left.\frac{d}{ds}\right|_{s=0}(\langle Y, \Gamma_s\rangle_{HS} - S(\Gamma_s)) = \langle Y, \Gamma'\rangle_{HS} - \mathrm{Tr}\left[\Gamma'(\log\bar{\Gamma}+1)\right].$$

This yields $\bar{\Gamma} = \exp(Y-1)$. Substituting in the expression for S^*, we arrive at (5.51).

Plugging this into (5.50) with $Y = \overline{Y}$ and recalling the definition of Y, we obtain

$$\mathrm{F}^{\varepsilon}(\Gamma) \geq \sum_{j=1}^{N} \mathrm{Tr}(U_j\gamma_j) - \varepsilon\,\mathrm{Tr}\left(\exp\left(\frac{\bigoplus_{j=1}^{N} U_j - \mathrm{H} - \varepsilon}{\varepsilon}\right)\right).$$

Changing the variable U_1 to $\tilde{U}_1 := U_1 + \varepsilon$, we conclude the proof. □

Remark 5.4.3 (The non-commutative Schrödinger problem). Suppose that $\boldsymbol{U} \in \bigtimes_{i=1}^N \mathcal{S}^{d_i}$ is a fixed point for τ, namely $\tau(\boldsymbol{U}) = \boldsymbol{U}$. This can be equivalently recast as $\mathfrak{T}_i^\varepsilon(\widehat{\boldsymbol{U}}_i) = U_i$, $\forall i \in [N]$. Then Lemma 5.3.18, (5.42) imply that the density matrix defined by

$$\Gamma := \exp\left(\frac{\bigoplus_{i=1}^N U_i - \mathrm{H}}{\varepsilon}\right) \tag{5.52}$$

has the correct marginals $\Gamma \mapsto (\gamma_1, \dots, \gamma_N)$ and thus it is admissible for the primal problem. In particular, it has trace 1 and we have

$$\mathrm{D}_\gamma^\varepsilon(U_1, \dots, U_N) = \mathrm{D}_\gamma^\varepsilon(\boldsymbol{U}) = \sum_{i=1}^N \mathrm{Tr}(U_i\gamma_i) = \mathrm{Tr}\left(\left(\bigoplus_{i=1}^N U_i\right)\Gamma\right).$$

On the other hand, directly from formula (5.52), we compute

$$\Gamma\,\mathrm{H} + \varepsilon\,\Gamma\log\Gamma = \Gamma\,\mathrm{H} + \Gamma\left(\bigoplus_{i=1}^N U_i - \mathrm{H}\right) = \Gamma\left(\bigoplus_{i=1}^N U_i\right)$$

and thus

$$\mathrm{F}^\varepsilon(\Gamma) = \mathrm{Tr}\left(\left(\bigoplus_{i=1}^N U_i\right)\Gamma\right) = \mathrm{D}_\gamma^\varepsilon(U_1, \dots, U_N). \tag{5.53}$$

In light of Proposition 5.4.2, this shows that if we are able to find a fixed point of τ, then this must be optimal for the dual problem (note that any maximizer is also a fixed point for τ as discussed in Remark 5.3.16) and the corresponding Γ as obtained in (5.52) must be optimal for the primal problem.

Another consequence of the above observations is that the set of maximizers for the dual problem is invariant under translations.

Lemma 5.4.4 (Structure of the maximizers)**.** *Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be two maximizers of $\mathrm{D}_\gamma^\varepsilon$, then there exists $\boldsymbol{\alpha} \in \mathbb{R}^N$ such that $\sum_{i=1}^N \boldsymbol{\alpha}_i = 0$ and $\boldsymbol{U} = \boldsymbol{V} + \boldsymbol{\alpha}$.*

Proof. Thanks to Remark 5.4.3 and using that the primal functional admits an unique minimizer by strict convexity, we find

$$\exp\left(\frac{\bigoplus_{i=1}^N (\boldsymbol{U})_i - \mathrm{H}}{\varepsilon}\right) = \exp\left(\frac{\bigoplus_{i=1}^N (\boldsymbol{V})_i - \mathrm{H}}{\varepsilon}\right) \implies \bigoplus_{i=1}^N (\boldsymbol{U})_i = \bigoplus_{i=1}^N (\boldsymbol{V})_i\,. \tag{5.54}$$

Applying the partial traces to the latter equality, we obtain

$$(\boldsymbol{U})_i = (\boldsymbol{V})_i + \sum_{j=1, j\neq i}^N \mathrm{Tr}(\boldsymbol{V})_j - \mathrm{Tr}(\boldsymbol{U})_j =: (\boldsymbol{V})_i + \boldsymbol{\alpha}_i.$$

Using (5.54) once again, one sees that

$$\sum_{i=1}^N \boldsymbol{\alpha}_i = (N-1)\left(\mathrm{Tr}\left(\bigoplus_{i=1}^N (\boldsymbol{U})_i\right) - \mathrm{Tr}\left(\bigoplus_{i=1}^N (\boldsymbol{V})_i\right)\right) = 0,$$

which concludes the proof. ∎

5.4.2 Proof of Theorem 5.2.1

We are finally ready to prove the equivalence between dual and primal problem, and to characterise the optimisers of the two problems. For the sake of clarity, recall that

$$\mathrm{Ren}(\boldsymbol{U})_i = \begin{cases} U_i - \lambda_\varepsilon(U_i), & \text{if } i \in \{1, \dots, N-1\} \\ U_N + \sum_{j=1}^{N-1} \lambda_\varepsilon(U_i), & \text{if } i = N, \end{cases}$$

as in Definition 5.3.20 and λ_ε is defined in (5.33) as $\lambda_\varepsilon(A) := \varepsilon \log\left(\mathrm{Tr}\left[\exp\left(\frac{A}{\varepsilon}\right)\right]\right)$, for every $A \in \mathcal{S}^d$, $d \in \mathbb{N}$.

Proof of Theorem 5.2.1. (ii). Take a maximizing sequence $\boldsymbol{U}_n$ for the dual problem and consider $\tilde{\boldsymbol{U}}_n := \mathrm{Ren}\,\tau(\boldsymbol{U}_n)$, where $\tau = \mathcal{T}_N^\varepsilon \circ \cdots \circ \mathcal{T}_1^\varepsilon$ is the Sinkhorn operator as introduced in Definition 5.3.15. Thanks to Proposition 5.3.21, $\tilde{\boldsymbol{U}}_n$ is again a maximizing sequence that satisfies

$$\left\|\tilde{\boldsymbol{U}}_n\right\|_\infty \leq 2\|\,\mathrm{H}\,\|_\infty + \varepsilon \sup_{i\in[N]} \|\log \gamma_i\|_\infty < \infty, \quad \forall n \in \mathbb{N},$$

and it is therefore compact. Pick any $\boldsymbol{U}^\varepsilon \in \times_{i=1}^N \mathcal{S}^{d_i}$ limit point of $\tilde{\boldsymbol{U}}_n$. By continuity of the dual functional we infer

$$\mathfrak{D}^\varepsilon(\boldsymbol{\gamma}) = \lim_{N\to\infty} \mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\tilde{\boldsymbol{U}}_n) = \mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\boldsymbol{U}^\varepsilon)$$

which shows that $\boldsymbol{U}^\varepsilon$ is a maximizer for $\mathfrak{D}^\varepsilon(\boldsymbol{\gamma})$. The fact that any other maximizer must coincide with $\boldsymbol{U}^\varepsilon$ follows from Lemma 5.4.4.

$(i)\&(iii)$ Proposition 5.4.2 proves one of the inequalities. To show the other inequality, we take any maximizer $\boldsymbol{U}^\varepsilon$ (which exists by the previous proof of (ii)). By construction of the Sinkhorn map, $\boldsymbol{U}^\varepsilon$ must be a fixed point of τ. Thanks to Remark 5.4.3, we conclude that

$$\Gamma^\varepsilon = \exp\left(\frac{\bigoplus_{i=1}^N \boldsymbol{U}_i^\varepsilon - \mathrm{H}}{\varepsilon}\right)$$

satisfies $\mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\boldsymbol{U}^\varepsilon) = \mathrm{F}^\varepsilon(\Gamma^\varepsilon) \geq \mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$. Hence Γ^ε is optimal for F^ε and $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma}) = \mathfrak{D}^\varepsilon(\boldsymbol{\gamma})$. □

5.4.3 Stability and the functional derivative of $\mathfrak{F}^\varepsilon$

In this last section, we show stability of the Kantorovich potentials with respect to the marginals $\boldsymbol{\gamma}$ and compute the Fréchet differential of $\mathfrak{F}^\varepsilon(\boldsymbol{\gamma})$ (or simply the differential in our finite dimensional setting). A similar result was first obtained by Pernal in [Per05] at zero temperature and in [GR19] in the positive temperature 1RDMFT case, i.e. considering also the fermionic and bosonic symmetry constraints. In [Per05], the result follows by a direct computation via chain rule, by taking the partial derivatives with respect to the eigenvalues and eigenvectors of a density matrix Γ. On the other hand, [GR19] uses tools from convex analysis and exploits the regularity of $\mathfrak{F}^\varepsilon$.

Our strategy is based on the Kantorovich formulation of (5.6) and follows ideas contained in [DMG20b].

Proof of Proposition 5.2.2. Consider $\gamma^n \xrightarrow{n\to\infty} \gamma$ and pick any sequence of Kantorovich potentials $\boldsymbol{U}^{\varepsilon,n}$ for $\mathfrak{F}^\varepsilon(\gamma^n)$. By optimality, they must be a fixed point for τ and hence, thanks to Proposition 5.3.21, $\mathrm{Ren}(\boldsymbol{U}^{\varepsilon,n})$ is uniformly bounded. Note that $\mathrm{Ren}(\boldsymbol{U}^{\varepsilon,n})$ are also maximizers for $\mathfrak{D}^\varepsilon(\gamma^n)$. This implies that any limit point of $\mathrm{Ren}(\boldsymbol{U}^{\varepsilon,n})$ must be a maximizer for $\mathfrak{D}^\varepsilon(\gamma)$. The continuity of $\mathfrak{F}^\varepsilon(\cdot)$ directly follows from this stability property.

Let us prove the differentiability. Fix $\sigma \in \mathcal{S}^{d_i}$, with $\mathrm{Tr}(\sigma) = 0$, and denote by γ^h the pertubation of γ with $+h\sigma$ in the ith entry. Denote by $\boldsymbol{U}^\varepsilon$ any Kantorovich potential for $\mathfrak{F}^\varepsilon(\gamma)$. From duality (Theorem 5.2.1) we can estimate

$$\frac{1}{h}\left(\mathfrak{F}^\varepsilon(\gamma^h) - \mathfrak{F}^\varepsilon(\gamma)\right) \geq \frac{1}{h}\left(\sum_{i=1}^N \mathrm{Tr}\left(\boldsymbol{U}_i^\varepsilon \gamma_i^h - \boldsymbol{U}_i^\varepsilon \gamma_i\right)\right) = \mathrm{Tr}(\boldsymbol{U}_i^\varepsilon \sigma) \tag{5.55}$$

for every $h \in \mathbb{R}$. Reversely, denote by $\boldsymbol{U}^{\varepsilon,h}$ any sequence of Kantorovich potentials for $\mathfrak{F}^\varepsilon(\gamma^h)$. Then for every $h > 0$ we obtain

$$\frac{1}{h}\left(\mathfrak{F}^\varepsilon(\gamma^h) - \mathfrak{F}^\varepsilon(\gamma)\right) \leq \frac{1}{h}\left(\sum_{i=1}^N \mathrm{Tr}\left(\boldsymbol{U}_i^{\varepsilon,h} \gamma_i^h - \boldsymbol{U}_i^{\varepsilon,h} \gamma_i\right)\right) = \mathrm{Tr}(\boldsymbol{U}_i^{\varepsilon,h} \sigma). \tag{5.56}$$

From the first part of the proof, we know that any limit point of $\mathrm{Ren}(\boldsymbol{U}_i^{\varepsilon,h})$ is a Kantorovich potential, which up to translation (Lemma 5.4.4) must coincide with $\boldsymbol{U}_i^\varepsilon$. Therefore, passing to the limit in (5.55) and (5.56), we obtain (5.9). □

5.5 Non-commutative Sinkhorn algorithm

In this section we introduce and prove convergences guarantees (Theorem 5.2.3) of the non-commutative version of the Sinkhorn algorithm, allowing us to compute numerically the minimiser (5.8) of the non-commutative multi-marginal optimal transport problem (5.47).

The idea of the Sinkhorn algorithm is to fix the shape of an ansatz

$$\Gamma^{(k)} = \exp\left(\frac{\bigoplus_{i=1}^N U_i^{(k)} - \mathrm{H}}{\varepsilon}\right),$$

since it is the actual shape of the minimizer in (5.8), and alternately project the Kantorovich potentials $U_i^{(k)}$ via the (H,ε)-transforms (Definition 5.3.12) to approximately reach the constraints $\Gamma^{(k)} \mapsto (\gamma_1, \dots, \gamma_N)$. Recall that for $i \in [N]$, the one-step operators $\mathcal{T}_i^\varepsilon : \bigtimes_{i=1}^N \mathcal{S}^{d_j} \to \bigtimes_{i=1}^N \mathcal{S}^{d_j}$ are given by

$$\boldsymbol{U} := (U_1, \dots, U_N), \quad \left(\mathcal{T}_i^\varepsilon(\boldsymbol{U})\right)_j = \begin{cases} U_j & \text{if } j \neq i, \\ \mathfrak{T}_i^\varepsilon(U_1, \dots, U_{i-1}, U_{i+1}, \dots U_N) & \text{if } j = i \end{cases}$$

where $\mathfrak{T}_i^\varepsilon$ can be implicitly defined (Lemma 5.3.18) solving the equation

$$\mathsf{P}_i\left[\exp\left(\frac{\bigoplus_{j=1}^N \left(\mathcal{T}_i^\varepsilon(\boldsymbol{U})\right)_j - \mathrm{H}}{\varepsilon}\right)\right] = \gamma. \tag{5.57}$$

Connection with the multi-marginal Sinkhorn algorithm: let us shortly describe what is the corresponding picture in the commutative setting [DMG20a, DMG20b]. For every $i \in [N]$, let

X_i be Polish Spaces, $\rho_i\mathfrak{m}_i \in \mathcal{P}(X_i)$ be probability measures with reference measures $\mathfrak{m}_i$. The Hamiltonian H corresponds to a bounded cost function $c : X_1 \times \cdots \times X_N \to \mathbb{R}$.

The Sinkhorn iterates define recursively the sequences $(a_j^n)_{n\in\mathbb{N}}, j \in [N]$ by

$$\begin{aligned} a_j^0(x_j) &= \rho_j(x_j), \quad j \in \{2, \dots, N\}, \\ a_j^n(x_j) &= \frac{\rho_j(x_j)}{\int \otimes_{i<j}^N a_i^n(x_i) \otimes_{i>j}^N a_i^{n-1}(x_i) e^{-c(x_1,\dots,x_N)/\varepsilon} \mathsf{d}(\otimes_{i\neq j}^N \mathfrak{m}_i)}, \ \forall n \in \mathbb{N} \text{ and } j \in [N]. \end{aligned} \tag{5.58}$$

Via the new variables $u_j^n = \varepsilon \ln(a_j^n)$, $j \in [N]$, one can rewrite the Sinkhorn sequences (5.58) as

$$\begin{aligned} u_j^n(x_j) &= -\varepsilon \log \left(\int_{\Pi_{i\neq j} X_i} \exp\left(\frac{\sum_{i\neq j} u_i^n(x_i) - c(x_1, \dots, x_N)}{\varepsilon} \right) \mathsf{d}\left(\otimes_{i\neq j}^N \mathfrak{m}_i\right) \right) + \varepsilon \log(\rho_j) \\ &= (\widehat{u_j^n})^{(N,c,\varepsilon)}(x_j). \end{aligned}$$

Or, more generally, for every $j \in [N]$, $u_j^n(x_j)$ corresponds to the solution of the maximisation

$$\operatorname*{argmax}_{u_i \in L^\infty(X_i)} \left\{ \sum_{i=1}^N \int_{X_j} u_i \rho_j \mathsf{d}\mathfrak{m}_j - \varepsilon \int_{\Pi_{i\neq j}^N X_i} \exp\left(\frac{\sum_{i\neq j} u_i^n + u - c}{\varepsilon} \right) \mathsf{d}\left(\otimes_{i\neq j}^N \mathfrak{m}_i\right) \right\} + \varepsilon \log(\rho_j)$$

which corresponds to the commutative counterpart of the i-th vectorial $(\mathrm{H}, \varepsilon)$-transform in Definition 5.3.12.

5.5.1 Definition of the algorithm

The non-commutative Sinkhorn algorithm is then defined iterating the $(\mathrm{H}, \varepsilon)$-transforms as in (5.57) for every $i \in [N]$. Note that, by construction, the matrix $\exp\left(\bigoplus_{i=1}^N \left((\mathcal{T}_i^\varepsilon(\boldsymbol{U}))_j - \mathrm{H}\right)/\varepsilon\right) \in \mathfrak{P}^{\boldsymbol{d}}$ and its i-th marginal coincide with γ_i. We define the one-step Sinkhorn map as

$$\begin{aligned} \tau : \bigtimes_{j=1}^N \mathcal{S}^{d_j} &\to \bigtimes_{j=1}^N \mathcal{S}^{d_j}, \\ \tau(\boldsymbol{U}) &:= (\mathcal{T}_N^\varepsilon \circ \cdots \circ \mathcal{T}_1^\varepsilon)(\boldsymbol{U}). \end{aligned}$$

Note that this is the non-commutative counterpart of the iteration defined in (5.58). The Sinkhorn algorithm is obtained iterating the map τ in the following way.

Step 0. We fix $\boldsymbol{U}^{(0)} \in \times_{i=1}^N \mathcal{S}^{d_i}$ an initial vector of potentials and define the density matrix

$$\Gamma^{(0)} := \exp\left(\frac{\bigoplus_{i=1}^N \boldsymbol{U}_i^{(0)} - \mathrm{H}}{\varepsilon} \right) \in \mathfrak{P}^{\boldsymbol{d}}.$$

Step k. For every $k \in \mathbb{N}$, we define the k-th density matrix via the formula

$$\Gamma^{(k)} := \exp\left(\frac{\bigoplus_{i=1}^N \tau^k(\boldsymbol{U}^{(0)})_i - \mathrm{H}}{\varepsilon} \right) \in \mathfrak{P}^{\boldsymbol{d}}, \tag{5.59}$$

where we write $\tau^k := \tau \circ \cdots \circ \tau$ the composition of τ for k-times.

Our goal is to prove the convergence $\Gamma^{(k)} \to \Gamma^\varepsilon$ where Γ^ε is optimal for $\mathfrak{F}^\varepsilon(\gamma)$. To do so, our plan is to obtain compactness at the level of the corresponding dual potentials. Nonetheless, the vectors $\tau^k(\boldsymbol{U}^{(0)})$ do not enjoy good a priori estimates and a renormalisation procedure is needed. For any given sequence $(\boldsymbol{\alpha}^k)_{k\in\mathbb{N}} \subset \mathbb{R}^N$ such that $\sum_{i=1}^N \boldsymbol{\alpha}_i^k = 0$, we define

$$\boldsymbol{U}^{(k)} := \tau^k(\boldsymbol{U}^{(0)}) + \boldsymbol{\alpha}^k, \quad k \in \mathbb{N}, \tag{5.60}$$

and observe that, by the properties of $\bigoplus$, the correspond density matrix does not change, thus

$$\Gamma^{(k)} = \exp\left(\frac{\bigoplus_{i=1}^N \boldsymbol{U}_i^{(k)} - \mathrm{H}}{\varepsilon}\right) \in \mathfrak{P}^{\boldsymbol{d}}, \quad \forall k \in \mathbb{N}. \tag{5.61}$$

Thanks to the good property of the renormalisation map and the Sinkhorn operator, we claim we can find a sequence $\boldsymbol{\alpha}^k$ such that the corresponding potentials $\boldsymbol{U}^{(k)}$ as defined in (5.60) do enjoy good a priori estimates and they can be used to prove the convergence of the algorithm, as we see in the next section.

5.5.2 Convergence guarantees: proof of Theorem 5.2.3

We are ready to prove our main result Theorem 5.2.3, which follows from the next Proposition.

Proposition 5.5.1 (Convergence of non-commutative Sinkhorn algorithm). *Fix $N \in \mathbb{N}$ and $\varepsilon > 0$. For all $i \in [N]$, let $\gamma_i \in \mathfrak{P}^{d_i}$ be density matrices, $\mathrm{H} \in \mathcal{S}^{\boldsymbol{d}}$, with $\ker \gamma_i = \{0\}$. For any initial potential $\boldsymbol{U}^{(0)} \in \times_{i=1}^N \mathcal{S}^d$, we consider the sequence $\Gamma^{(k)} \in \mathfrak{P}^{\boldsymbol{d}}$ as defined in (5.59).*

1. *There exist $\boldsymbol{\alpha}^k \in \mathbb{R}^N$ with $\sum_{i=1}^N \alpha_i^k = 0$ such that*
$$\boldsymbol{U}^{(k)} = \tau^k(\boldsymbol{U}) + \boldsymbol{\alpha}^k \to \boldsymbol{U}^\varepsilon \quad \text{as } k \to +\infty. \tag{5.62}$$
2. *$\boldsymbol{U}^\varepsilon = (\boldsymbol{U}_1^\varepsilon, \dots, \boldsymbol{U}_N^\varepsilon)$ is optimal for the dual problem $\mathfrak{D}^\varepsilon(\gamma)$, as defined in (5.49).*
3. *$\Gamma^{(k)}$ converges as $k \to \infty$ to some $\Gamma^\varepsilon \in \mathfrak{P}^{\boldsymbol{d}}$ which is optimal for the primal problem $\mathfrak{F}^\varepsilon(\gamma)$, as defined in (5.48). In particular, it holds*
$$\Gamma^\varepsilon = \exp\left(\frac{\bigoplus_{i=1}^N \boldsymbol{U}_i^\varepsilon - \mathrm{H}}{\varepsilon}\right). \tag{5.63}$$

Proof. For any $\boldsymbol{U}^{(0)} \in \times_{i=1}^N \mathcal{S}^{d_i}$, we define the sequence $\boldsymbol{U}_k := \mathrm{Ren}\, \tau^k(\boldsymbol{U}^{(0)})$. Note that $\boldsymbol{U}_k$ is of the form (5.60), for some $\boldsymbol{\alpha}^k$. Thanks to Proposition 5.3.21, we infer that $\boldsymbol{U}_k$ is uniformly bounded and hence compact. Therefore, there exists a subsequence $\boldsymbol{U}_{k_j} \to \boldsymbol{U}^\varepsilon$. We first show that $\boldsymbol{U}^\varepsilon$ is a maximizer for the dual problem. Indeed, using the properties of Ren and τ, it holds

$$\mathrm{D}_\gamma^\varepsilon(\tau(\boldsymbol{U}_{k_j})) = \mathrm{D}_\gamma^\varepsilon(\tau^{k_j+1}(\boldsymbol{U}^{(0)})) \leq \mathrm{D}_\gamma^\varepsilon(\tau^{k_{j+1}}(\boldsymbol{U}^{(0)})) = \mathrm{D}_\gamma^\varepsilon(\boldsymbol{U}_{k_{j+1}}).$$

Passing to the limit the previous inequality, using the continuity of $\mathrm{D}_\gamma^\varepsilon$ and τ and recalling that for any $\boldsymbol{U}$ we have $\mathrm{D}_\gamma^\varepsilon(\tau(\boldsymbol{U})) \geq \mathrm{D}_\gamma^\varepsilon(\boldsymbol{U})$, we obtain

$$\mathrm{D}_\gamma^\varepsilon(\tau(\boldsymbol{U}^\varepsilon)) = \mathrm{D}_\gamma^\varepsilon(\boldsymbol{U}^\varepsilon).$$

By definition, this means that $\boldsymbol{U}^\varepsilon$ is a fixed point for τ and therefore a maximizer (Remark 5.4.3).

In order to prove (1), we show there exists a choice $\boldsymbol{\alpha}^k$ such that $\boldsymbol{U}_k + \boldsymbol{\alpha}^k \to \boldsymbol{U}^\varepsilon$. For $k = k_j$ for some j, we pick $\boldsymbol{\alpha}^k = 0$, for all the others k, we instead pick $\boldsymbol{\alpha}^k$ defined by

$$\boldsymbol{\alpha}^k = \operatorname{argmin}_{\boldsymbol{\alpha}} \left\{ \|\boldsymbol{U}_k + \boldsymbol{\alpha} - \boldsymbol{U}^\varepsilon\|_\infty \ : \ \sum_{i=1}^N \boldsymbol{\alpha}_i = 0 \right\}.$$

Note that, by Lemma 5.4.4, this is equivalent to picking $\boldsymbol{\alpha}^k$ such that $\boldsymbol{U}^\varepsilon$ is the closest maximizer to $\boldsymbol{U}_k + \boldsymbol{\alpha}^k$. We claim this is the right choice. Suppose indeed by contradiction that there exists a subsequence $\boldsymbol{U}_{k'_j}$ such that $\|\boldsymbol{U}_{k'_j} + \boldsymbol{\alpha}_{k'_j} - \boldsymbol{U}^\varepsilon\|_\infty \geq \delta > 0$, then by construction $\|\boldsymbol{U}_{k'_j} + \boldsymbol{\alpha}_{k'_j} - \boldsymbol{U}'\|_\infty \geq \delta$ for any other maximizer $\boldsymbol{U}'$. By compactness, this is a contradiction, since there exists a further subsequence $\boldsymbol{U}_{k''_j}$ of $\boldsymbol{U}_{k'_j}$ converging to a maximizer $\boldsymbol{U}'$ (by the same reasoning carried out above). This proves (1) and by optimality of $\boldsymbol{U}^\varepsilon$, (2) as well. The convergence of $\Gamma^{(k)}$ follows from the compactness of $\boldsymbol{U}^{(k)}$ and (5.61), whereas the optimality of the limit point Γ^ε and (5.63) are consequence of the optimality of $\boldsymbol{U}^\varepsilon$ and Remark 5.4.3. □

5.6 One-body reduced density matrix functional theory

In this last section, we prove Proposition 5.2.8 and consequently Theorem 5.2.9.

For given $d, N \in \mathbb{N}$, we set $\boldsymbol{d} = d^N$ and consider the space of bosonic (resp. fermionic) density matrices $\mathfrak{P}^d_+$ (resp. $\mathfrak{P}^d_-$) as introduced in (5.14). Recall as well that for any given operator $A \in \mathcal{S}^{\boldsymbol{d}}$, we denote by $A_\pm$ the corresponding projection onto the symmetric space, obtained as $A_\pm := \Pi_\pm \circ A \circ \Pi_\pm$, where $\Pi_\pm$ are defined in (5.13).

The universal functional in the bosonic and in the fermionic case is then given as in Definition 5.2.6, which we recall here for simplicity is given by

$$\mathfrak{F}^\varepsilon_\pm(\gamma) := \inf \left\{ \operatorname{Tr}(\mathrm{H}\,\Gamma) + \varepsilon \operatorname{Tr}(\Gamma \log \Gamma) \ : \ \Gamma \in \mathfrak{P}^{\boldsymbol{d}}_\pm \text{ and } \Gamma \mapsto \gamma \right\},$$

whereas the corresponding dual functional and problem (see Definition 5.2.7) are given by

$$\mathsf{D}^{\pm,\varepsilon}_\gamma(U) := \operatorname{Tr}(U\gamma) - \varepsilon \operatorname{Tr}\left(\exp\left[\frac{1}{\varepsilon} \left(\frac{1}{N} \bigotimes_{i=1}^N U - \mathrm{H} \right)_\pm \right] \right) + \varepsilon\,,$$
$$\mathfrak{D}^\varepsilon_\pm(\gamma) := \sup \left\{ \mathsf{D}^{\pm,\varepsilon}_\gamma(U) \ : \ U \in \mathcal{S}^d \right\}.$$

We are interested in fully characterizing the existence of the optimizers in the primal and the dual problems, for both bosonic and fermionic cases. Proceeding in a similar way as in the proof of Lemma 5.3.7, one can prove that every maximizer $U^\varepsilon_\pm$ of the dual functional $\mathsf{D}^{\pm,\varepsilon}_\gamma(\cdot)$ must satisfy the corresponding Euler-Lagrange equation given by

$$\gamma = \mathsf{P}_1 \left(\exp\left[\frac{1}{\varepsilon} \left(\frac{1}{N} \bigoplus_{i=1}^N U^\varepsilon_\pm - \mathrm{H} \right)_\pm \right] \right). \tag{5.64}$$

5.6.1 Fermionic dual problem and Pauli's exclusion principle

The aim of this section is to prove Proposition 5.2.8. For simplicity we assume, with no loss of generality, that $\varepsilon = 1$ and set $\mathsf{D}^-_\gamma := \mathsf{D}^{-,1}_\gamma$.

For any $U \in \mathcal{S}^d$, we fix a basis of normalized eigenvectors of U, denoted by $\{\psi_j\}_j$, and consider the decomposition

$$U = \sum_{j=1}^{d} u_j |\psi_j\rangle\langle\psi_j|, \quad u_j \in \mathbb{R} \quad \text{(eigenvalues)}\,. \tag{5.65}$$

We also denote by $\gamma_j := \langle\psi_j|\gamma|\psi_j\rangle$. In particular, the linear terms read

$$\mathrm{Tr}(U\gamma) = \sum_{j=1}^{d} \gamma_j u_j\,.$$

For any such basis $\{\psi_i\}_i$, we obtain a basis of the fermionic tensor product

$$\psi^{as}_{\boldsymbol{j}} := \bigwedge_{i=1}^{N} \psi_{j_i}\,, \quad \boldsymbol{j} = (j_i)_{i=1}^{N} \in \Theta_-\,,$$
$$\Theta_- := \Big\{(j_1,\dots,j_N)\ :\ j_i \in \{1,\dots,d\},\ j_i \neq j_k,\ \text{if } i \neq k\Big\}/\S_N\,,$$

where $\S_N$ denotes the set of permutations of N elements. With respect to this basis, we can write

$$\frac{1}{N}\left(\bigoplus_{i=1}^{N} U\right)_- = \sum_{\boldsymbol{j}\in\Theta_-}\left(\frac{1}{N}\sum_{i=1}^{N} u_{j_i}\right)|\psi^{as}_{\boldsymbol{j}}\rangle\langle\psi^{as}_{\boldsymbol{j}}|\,. \tag{5.66}$$

Using the monotonicity of the exponential and the trace, we obtain the following result.

Lemma 5.6.1 (Bounds for $\mathrm{D}^-_\gamma(U)$)**.** *Fix $U \in \mathcal{S}^d$ with eigenvalues u_j and eigenvectors $\{\psi_j\}_j$. For $\gamma \in \mathfrak{P}(d)$, set $\gamma_j := \langle\psi_j|\gamma|\psi_j\rangle$. Then one has*

$$\begin{aligned}\sum_{j=1}^{d}\gamma_j u_j - C\sum_{\boldsymbol{j}\in\Theta_-}\exp\left(\frac{1}{N}\sum_{i=1}^{N} u_{j_i}\right) &\leq \mathrm{D}^-_\gamma(U) - 1\\ &\leq \sum_{j=1}^{d}\gamma_j u_j - \frac{1}{C}\sum_{\boldsymbol{j}\in\Theta_-}\exp\left(\frac{1}{N}\sum_{i=1}^{N} u_{j_i}\right),\end{aligned} \tag{5.67}$$

where $C = \exp\Big(\|\,\mathrm{H}\,\|_\infty\Big) \in (0,+\infty)$.

Before moving to the proof of Proposition 5.2.8, we need the following technical lemma.

Lemma 5.6.2 (Linear term estimates)**.** *Consider $\{u_j\}_{j=1}^d \subset \mathbb{R}$ and $\{\gamma_j\}_{j=1}^d$ such that*

$$\gamma_j \in \left(\delta, \frac{1}{N} - \delta\right), \quad \sum_{j=1}^{d}\gamma_j = 1\,, \tag{5.68}$$

for some $\delta \in \left[0, \dfrac{1}{2N}\right)$. Suppose that $u_j \leq u_k$ if $j \leq k$. Then we have

$$\sum_{j=1}^{d}\gamma_j u_j \leq \frac{1}{N}\sum_{j=1}^{N} u_j - \delta(u_1 - u_d)\,. \tag{5.69}$$

Proof. Thanks to the fact the u_j are ordered, we have the inequality

$$\sum_{j=1}^{d} \bar{\gamma}_j u_j \leq \frac{1}{N} \sum_{j=1}^{N} u_j, \quad \forall\, 0 \leq \bar{\gamma}_j \leq \frac{1}{N}, \quad \sum_{j=1}^{d} \bar{\gamma}_j = 1\,.$$

Then (5.69) follows applying the above inequality to

$$\bar{\gamma}_1 := \gamma_1 + \delta \in \left(0, \frac{1}{N}\right), \quad \bar{\gamma}_j := \gamma_j, \quad \bar{\gamma}_d := \gamma_d - \delta \in \left(0, \frac{1}{N}\right),$$

for every $j \in \{2, \dots, d-1\}$. □

We are ready to prove Proposition 5.2.8.

Proof. ($\gamma \leq 1/N \Rightarrow \sup \mathsf{D}_\gamma^- < \infty$). This is consequence of Proposition 5.6.2 with $\delta = 0$. More precisely, pick $U \in \mathcal{S}^d$ and consider a decomposition in eigenfunctions as in (5.65). Assume that $\{u_j\}_j$ are non increasing in j (with no loss of generality). We can then apply Proposition 5.6.2 with $\delta = 0$ and from (5.69) and (5.67) we deduce

$$\mathsf{D}_\gamma^-(U) - 1 \leq \frac{1}{N} \sum_{i=1}^{N} u_j - \frac{1}{C} \sum_{j \in \Theta_-} \exp\left(\frac{1}{N} \sum_{i=1}^{N} u_{j_i}\right) \leq \frac{1}{N} \sum_{i=1}^{N} u_j - \frac{1}{C} \exp\left(\frac{1}{N} \sum_{i=1}^{N} u_j\right),$$

where in the last inequality we used the positivity of the exponential. Therefore

$$\sup_{U \in \mathcal{S}^d} \mathsf{D}_\gamma^-(U) \leq \sup_{x \in \mathbb{R}} (x - \frac{1}{C} e^x) + 1 = \log C < \infty\,.$$

($\sup \mathsf{D}_\gamma^- < \infty \Rightarrow \gamma \leq 1/N$). Suppose by contradiction that the Pauli's principle is not satisfied. With no loss of generality, we can assume that

$$\gamma = \sum_{i=1}^{d} \gamma_i |\psi_i\rangle\langle\psi_i|, \quad \gamma_1 > \frac{1}{N}\,.$$

We build the sequence of bounded operators $U^n \in \mathcal{S}^d$ given by

$$U^n := \sum_{i=1}^{d} u_i^n |\psi_i\rangle\langle\psi_i|, \quad u_1^n := n\,, \quad u_j^n := -\frac{n}{N-1}, \quad \forall j \geq 2\,. \tag{5.70}$$

Observe that by construction, we can estimate the non-linear part of $\mathsf{D}_\gamma^-(U)$ as

$$\forall \boldsymbol{j} \in \Theta_-, \quad \exp\left(\frac{1}{N} \sum_{i=1}^{N} u_{j_i}\right) \begin{cases} = 1 & \text{if } j_i = 1 \text{ for some } i\,, \\ \leq 1 & \text{otherwise}\,. \end{cases}$$

It follows that we can bound from below $\mathsf{D}_\gamma^-(U^n)$ as

$$\mathsf{D}_\gamma^-(U^n) \geq \sum_{j=1}^{d} \gamma_j u_j^n - C \binom{d}{N}\,. \tag{5.71}$$

We claim that the linear contribution goes to $+\infty$ as $n \to +\infty$. To see that, note that

$$\sum_{j=1}^{d} \gamma_j u_j^n = n\Big(\gamma_1 - \frac{1}{N-1}\sum_{i=2}^{d}\gamma_2\Big) = \frac{n}{N-1}\Big(N\gamma_1 - 1\Big), \tag{5.72}$$

where we used that $\sum_i \gamma_i = 1$. From this, using $\gamma_1 > \frac{1}{N}$ and (5.71) we deduce $\mathsf{D}_\gamma^-(U^n) \to +\infty$ as $n \to +\infty$, thus a contradiction.

(*Equation for the maximizer and uniqueness*). If a maximizer exists, then it solves the equation (5.64). Thanks to the Peierls inequality, we also know that D_γ^- is strictly concave (because the exponential is strictly convex), hence the uniqueness of the maximizer.

(*Existence of* $\operatorname{argmax} \mathsf{D}_\gamma^- \Rightarrow 0 < \gamma < 1/N$). We proceed as in the latter proof. By contradiction, assume that

$$\gamma = \sum_{j=1}^{d} \gamma_j |\psi_j\rangle\langle\psi_j|, \quad \gamma_1 = \frac{1}{N}, \quad \gamma_j \in \Big(0, \frac{1}{N}\Big), \quad \forall j \geq 2\,.$$

The case $\gamma_j = 0$ can be directly ruled out from the Euler-Lagrange equation for a maximizer (5.64). We can then consider the very same sequence U^n as defined in (5.70). From (5.71), (5.72), and the first part of Theorem 5.2.8, on one hand we deduce

$$-C\binom{d}{N} \leq \mathsf{D}_\gamma^-(U^n) \leq \log C\,, \quad \forall n \in \mathbb{N}\,.$$

On the other hand, $\|U^n\|_\infty \to +\infty$ as $n \to \infty$, which means that D_γ^- is not coercive. Thanks to Peierls inequality, we also know that D_γ^- is strictly concave, which implies that D_γ^- can not attain its maximum.

($0 < \gamma < 1/N \Rightarrow$ *existence of* $\operatorname{argmax} \mathsf{D}_\gamma^-$). Let $U \in \mathcal{S}^d$ and consider a decomposition in eigenfunctions as in (5.65). Assume that $\{u_j\}_j$ are non increasing in j (with no loss of generality) and denote by $\gamma_j := \langle\psi_j|\gamma|\psi_j\rangle$. By assumption, there exists $\delta \in \big(0, \frac{1}{N}\big)$ such that

$$\sum_{j=1}^{d} \gamma_j = 1, \quad \gamma_j \in \Big(\delta, \frac{1}{N} - \delta\Big), \quad \forall j \in \{1, \dots, d\}\,. \tag{5.73}$$

We can then apply Proposition 5.6.2 and (5.67) to obtain

$$\mathsf{D}_\gamma^-(U) - 1 \leq \frac{1}{N}\sum_{i=1}^{N} u_j - \frac{1}{C}\sum_{j\in\Theta_-} \exp\Big(\frac{1}{N}\sum_{i=1}^{N} u_{j_i}\Big) - \delta(u_1 - u_d) \tag{5.74}$$

$$\leq \frac{1}{N}\sum_{i=1}^{N} u_j - \frac{1}{C}\exp\Big(\frac{1}{N}\sum_{i=1}^{N} u_j\Big) - \delta(u_1 - u_d), \tag{5.75}$$

where we used the positivity of the exponential. Set $S := \sup_x(x - \frac{e^x}{C}) + 1 < \infty$, and infer

$$\mathsf{D}_\gamma^-(U) \leq S - \delta(u_{max} - u_{min}), \quad \forall U \in \mathcal{S}^d, \quad U = \sum_{j=1}^{d} u_j|\psi_j\rangle\langle\psi_j|\,, \tag{5.76}$$

where u_{max} and u_{min} denotes respectively the maximum/minimum eigenvalue of U. Let us use this estimate to prove to existence of a maximizer for D_γ^-. Consider a maximizing sequence U^n of bounded operators. In particular, we can assume that $-I := \inf_n \mathsf{D}_\gamma^-(U^n) \geq -\infty$. If the sequence $\{U^n\}_n$ is bounded in $\mathcal{S}^d$, then any limit point is a maximum for D_γ^-, by concavity and continuity of D_γ^-, and the proof is complete. Suppose by contradiction that $\|U^n\|_\infty \to +\infty$ as $n \to +\infty$. Note that from (5.76) we deduce

$$\sup_{n\in\mathbb{N}} \left(u^n_{max} - u^n_{min}\right) \leq \frac{S+I}{\delta} < \infty\,, \tag{5.77}$$

therefore we deduce that either $u_j^n \to -\infty$ or $u_j^n \to +\infty$ for every $j \in \{1,\dots,d\}$. In the first case, we would have a contradiction, because

$$-I \leq \mathsf{D}_\gamma^-(U^n) \leq \sum_{j=1}^d \gamma_j u_j^n + 1 \to -\infty \quad \text{as } n \to +\infty\,.$$

In the second case, we can use (5.74) to find a contradiction, because

$$-I \leq \mathsf{D}_\gamma^-(U^n) \leq \frac{1}{N}\sum_{i=1}^N u_j^n - \frac{1}{C}\exp\left(\frac{1}{N}\sum_{i=1}^N u_j^n\right) + 1 \to -\infty\,,$$

where we used that $\lim_{x\to+\infty}(x - C^{-1}e^x) = -\infty$. The proof is complete. □

5.6.2 Duality theorem for fermionic and bosonic systems

In this section we prove Theorem 5.2.9. The proof relies on the use of Theorem 5.2.1 and the existence of maximizers for $\mathsf{D}_\gamma^{-,\varepsilon}$, proved in Proposition 5.2.8, and $\mathsf{D}_\gamma^{+,\varepsilon}$. The latter can be proven easily by noting that the spectrum of $\left(\bigoplus_{j=1}^N U_j\right)_+$ contains the spectrum of U and, applying similar computations to the ones used in the case of $\mathsf{D}_\gamma^{-,\varepsilon}$, deducing the coercivity of $\mathsf{D}_\gamma^{+,\varepsilon}$. We also need the following observation.

Remark 5.6.3. If H satisfies (5.15) and $\boldsymbol{\gamma} = (\gamma_i)_i$, $\gamma_i = \gamma$, then the minimizers of $\mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon$ (the dual functional without symmetry constraints) can be taken to satisfy $U_i \equiv U$, for some $U \in \mathcal{S}^d$. In particular

$$\mathfrak{D}^\varepsilon(\boldsymbol{\gamma}) = \sup_{\boldsymbol{U}\in(\mathcal{S}^d)^N} \mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\boldsymbol{U}) = \sup_{U\in\mathcal{S}^d}\left\{\mathrm{Tr}(U\gamma) - \varepsilon\,\mathrm{Tr}\left(\exp\left[\frac{1}{\varepsilon}\left(\frac{1}{N}\bigotimes_{i=1}^N U - \mathrm{H}\right)\right]\right)\right\} + \varepsilon\,.$$

This follows from the observation that if $\boldsymbol{U} \in (\mathcal{S}^d)^N$, then we obtain a symmetric competitor $\tilde{\boldsymbol{U}}$

$$(\tilde{\boldsymbol{U}})_i = \frac{1}{N}\sum_{j=1}^N U_j\,, \quad \text{such that} \quad \mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\tilde{\boldsymbol{U}}) = \mathrm{D}_{\boldsymbol{\gamma}}^\varepsilon(\boldsymbol{U})\,.$$

Proof of Theorem 5.2.9. Let us assume that $\gamma > 0$ in the bosonic case ($0 < \gamma < \frac{1}{N}$ in the fermionic case). The general duality result (including the case γ in which does not satisfy the above strict inequalities) can be handled by decomposition of the space, in the same way as in Remark 5.3.9.

Under these assumptions, thanks to Proposition 5.2.8, we know that a maximizer $U_\pm^\varepsilon$ exists and satisfies (5.64).

We then define the N-particle density matrix

$$\tilde{\Gamma}^{\varepsilon}_{\pm} := \exp\left[\frac{1}{\varepsilon}\left(\frac{1}{N}\bigoplus_{i=1}^{N} U^{\varepsilon}_{\pm} - \mathrm{H}\right)_{\pm}\right] \in \mathcal{S}^{\boldsymbol{d}}\,,$$

and thanks to Remark 5.4.3, we know that $\tilde{\Gamma}^{\varepsilon}_{\pm}$ is optimal for the problem $\mathfrak{F}^{\varepsilon}\Big(\mathsf{P}_1(\tilde{\Gamma}^{\varepsilon}_{\pm})\Big)$ without symmetry contraints. Observing that $(\tilde{\Gamma}^{\varepsilon}_{\pm})_{\pm} = \Gamma^{\varepsilon}_{\pm}$ (defined in (5.20),(5.21)), we deduce that $\Gamma^{\varepsilon}_{\pm}$ must be optimal for the primal problem $\mathfrak{F}^{\varepsilon}_{\pm}(\gamma)$ with symmetry constraints. This also proves the equality between primal and dual problems and concludes the proof. □

Acknowledgments

This work started when A.G. was visiting the Erwin Schrödinger Institute and then continued when D.F. and L.P visited the Theoretical Chemistry Department of the Vrije Universiteit Amsterdam. The authors thanks the hospitality of both places and, especially, P. Gori-Giorgi and K. Giesbertz for fruitful discussions and literature suggestions in the early state of the project. Finally, the authors also thanks J. Maas and R. Seiringer for their feedback and useful comments to a first draft of the article.

L.P. acknowledges support by the Austrian Science Fund (FWF), grants No W1245 and grant No F65. D.F acknowledges support by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 716117). A.G. acknowledges funding by the European Research Council under H2020/MSCA-IF "OTmeetsDFT" [grant ID: 795942].

Bibliography

[AC04] Roberto Alicandro and Marco Cicalese. A general integral representation result for continuum limits of discrete energies with superlinear growth. *SIAM J. Math. Anal.*, 36(1):1–37, 2004.

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[AL05] Paul W Ayers and Mel Levy. Generalized density-functional theory: Conquering the n-representability problem with exact functionals for the electron pair density and the second-order reduced density matrix. *Journal of Chemical Sciences*, 117(5):507–514, 2005.

[ARM17] F. Al Reda and B. Maury. Interpretation of finite volume discretization schemes for the Fokker-Planck equation as gradient flows for the discrete Wasserstein distance. In *Topological optimization and optimal transport*, volume 17 of *Radon Ser. Comput. Appl. Math.*, pages 400–416. De Gruyter, Berlin, 2017.

[Bar17] Martin T. Barlow. *Random walks and heat kernels on graphs*, volume 438 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2017.

[BB00] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[BB02] MA Buijse and EJ Baerends. An approximate exchange-correlation hole density as a functional of the natural orbitals. *Molecular Physics*, 100(4):401–421, 2002.

[BBC20] Maria Bruna, Martin Burger, and José Antonio Carrillo. Coarse graining of a Fokker–Planck equation with excluded volume effects preserving the gradient flow structure. *European Journal of Applied Mathematics*, pages 1–35, 2020.

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[BBK09] Martin T. Barlow, Richard F. Bass, and Takashi Kumagai. Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. *Math. Z.*, 261(2):297–320, 2009.

[BCG15] Tim Baldsiefen, Attila Cangi, and Eberhard KU Gross. Reduced-density-matrix-functional theory at finite temperature: Theoretical foundations. *Physical Review A*, 92(5):052514, 2015.

[BD98] Andrea Braides and Anneliese Defranceschi. *Homogenization of multiple integrals*, volume 12 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1998.

[BEG12] Tim Baldsiefen, Florian G Eich, and Eberhard KU Gross. Reduced density matrix functional theory at finite temperature. II. Application to the electron gas: Exchange only. *arXiv preprint arXiv:1208.4705*, 2012.

[BF91] Giuseppe Buttazzo and Lorenzo Freddi. Functionals defined on measures and applications to non-equi-uniformly elliptic problems. *Ann. Mat. Pura Appl. (4)*, 159:133–149, 1991.

[BFLM02] Guy Bouchitté, Irene Fonseca, Giovanni Leoni, and Luísa Mascarenhas. A global method for relaxation in $W^{1,p}$ and in SBV_p. *Arch. Ration. Mech. Anal.*, 165(3):187–242, 2002.

[BG12] Tim Baldsiefen and Eberhard KU Gross. Reduced density matrix functional theory at finite temperature. III. Application to the electron gas: Correlation effects. *arXiv preprint arXiv:1208.4707*, 2012.

[BHO18] Timothy Barth, Raphaèle Herbin, and Mario Ohlberger. Finite volume methods: foundation and analysis. *Encyclopedia of Computational Mechanics Second Edition*, pages 1–60, 2018.

[Bog07] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[Bra02] Andrea Braides. *Gamma-convergence for Beginners*, volume 22. Clarendon Press, 2002.

[Bre87] Yann Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math.*, 305:805–808, 1987.

[Bre10] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.

[BRWMS20] Carlos L Benavides-Riveros, Jakob Wolff, Miguel A L Marques, and Christian Schilling. Reduced density matrix functional theory for bosons. *Physical Review Letters*, 124(18):180603, 2020.

[BV20] Yann Brenier and Dmitry Vorotnikov. On optimal transport of matrix-valued measures. *SIAM Journal on Mathematical Analysis*, 52(3):2849–2873, 2020.

[Car10] Eric Carlen. Trace inequalities and quantum entropy: an introductory course. *Entropy and the quantum*, 529:73–140, 2010.

[CG17] Clément Cancès and Cindy Guichard. Numerical analysis of a robust free energy diminishing finite volume scheme for parabolic equations with gradient structure. *Found. Comput. Math.*, 17(6):1525–1584, 2017.

[CGGT19] Yongxin Chen, Wilfrid Gangbo, Tryphon Georgiou, and Allen Tanenbaum. On the matrix Monge–Kantorovich problem. *European Journal of Applied Mathematics*, pages 1–27, 2019.

[CGLM19] Clément Cancès, Thomas Gallouët, Maxime Laborde, and Léonard Monsaingeon. Simulation of multiphase porous media flows with minimising movement and finite volume schemes. *European J. Appl. Math.*, 30(6):1123–1152, 2019.

[CGP16] Yongxin Chen, Tryphon Georgiou, and Michele Pavon. Entropic and displacement interpolation: a computational approach using the hilbert metric. *SIAM Journal on Applied Mathematics*, 76(6):2375–2396, 2016.

[CGP18] Emanuele Caglioti, François Golse, and Thierry Paul. Towards optimal transport for quantum densities. *Available on https://hal.archives-ouvertes.fr/hal-01963667/*, 2018.

[CGP20] Emanuele Caglioti, François Golse, and Thierry Paul. Quantum optimal transport is cheaper. *Journal of Statistical Physics (2020)*, 181:149–162, 2020.

[CGT18] Y. Chen, T. T. Georgiou, and A. Tannenbaum. Matrix Optimal Mass Transport: A Quantum Mechanical Approach. *IEEE Transactions on Automatic Control*, 63(8):2612–2619, 2018.

[CHLZ12] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker–Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.*, 203(3):969–1008, 2012.

[CKS87] Eric Anders Carlen, Shigeo Kusuoka, and Daniel W Stroock. Upper bounds for symmetric Markov transition functions. In *Annales de l'IHP Probabilités et statistiques*, volume 23, pages 245–287, 1987.

[CM14] Eric A. Carlen and Jan Maas. An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker–Planck equation is gradient flow for the entropy. *Comm. Math. Phys.*, 331(3):887–926, 2014.

[CM17] Eric A. Carlen and Jan Maas. Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance. *J. Funct. Anal.*, 273(5):1810–1869, 2017.

[CMRS19] Fabio A. C. C. Chalub, Léonard Monsaingeon, Ana Margarida Ribeiro, and Max O. Souza. Gradient flow formulations of discrete and continuous evolutionary models: a unifying perspective, 2019.

[Col63] A John Coleman. Structure of fermion density matrices. *Reviews of modern Physics*, 35(3):668, 1963.

[CP19] Marco Cuturi and Gabriel Peyré. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[CPSV18] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609, 2018.

[Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.

[DEG+18] Jérôme Droniou, Robert Eymard, Thierry Gallouët, Cindy Guichard, and Raphaele Herbin. *The gradient discretisation method*, volume 82. Springer, 2018.

[Del99] Thierry Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana*, 15(1):181–232, 1999.

[DFDMPR21] Marco Di Francesco, Simone Di Marino, Lorenzo Portinale, and Emanuela Radici. Transport problems with non linear mobilities: a particle approximation result. *In preparation*, 2021.

[DFM19] Patrick Dondl, Thomas Frenzel, and Alexander Mielke. A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM Control Optim. Calc. Var.*, 25:Paper No. 68, 45, 2019.

[DGL77] E. De Giorgi and G. Letta. Une notion générale de convergence faible pour des fonctions croissantes d'ensemble. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 4(1):61–99, 1977.

[DL15] Karoline Disser and Matthias Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Netw. Heterog. Media*, 10(2):233–253, 2015.

[DMG20a] Simone Di Marino and Augusto Gerolin. An Optimal Transport approach for the Schrödinger bridge problem and convergence of Sinkhorn algorithm. *Journal of Scientific Computing*, 85(27), 2020.

[DMG20b] Simone Di Marino and Augusto Gerolin. Optimal Transport losses and Sinkhorn algorithm with general convex regularization. *arXiv preprint arXiv:2007.00976*, 2020.

[DNS09] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

[DNS12] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. From Poincaré to logarithmic Sobolev inequalities: a gradient flow approach. *SIAM J. Math. Anal.*, 44(5):3186–3216, 2012.

[DPT19] Giacomo De Palma and Dario Trevisan. Quantum optimal transport with quantum channels. *arXiv preprint arXiv:1911.00803*, 2019.

[DPTGA18] Giacomo De Palma, Dario Trevisan, Vittorio Giovannetti, and Luigi Ambrosio. Gaussian optimizers for entropic inequalities in quantum information. *Journal of Mathematical Physics*, 59(8):081101, 2018.

[EF18] Matthias Erbar and Max Fathi. Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. *J. Funct. Anal.*, 274(11):3056–3089, 2018.

[EGH00] Robert Eymard, Thierry Gallouët, and Raphaèle Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.

[EHMT17] Matthias Erbar, Christopher Henderson, Georg Menz, and Prasad Tetali. Ricci curvature bounds for weakly interacting Markov chains. *Electron. J. Probab.*, 22:Paper No. 40, 23, 2017.

[EM12] Matthias Erbar and Jan Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.*, 206(3):997–1038, 2012.

[EM14] Matthias Erbar and Jan Maas. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.*, 34(4):1355–1374, 2014.

[EMT15] Matthias Erbar, Jan Maas, and Prasad Tetali. Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):781–800, 2015.

[EMW19] Matthias Erbar, Jan Maas, and Melchior Wirth. On the geometry of geodesics in discrete optimal transport. *Calc. Var. Partial Differential Equations*, 58(1):Art. 19, 19, 2019.

[FGP21] Dario Feliciangeli, Augusto Gerolin, and Lorenzo Portinale. A non-commutative entropic optimal transport approach to quantum composite systems at positive temperature. *Submitted, arxiv:2106.11217*, 2021.

[FL89] Joel Franklin and Jens Lorenz. On the scaling of multidimensional matrices. *Linear Algebra and its applications*, 114:717–735, 1989.

[FL18] Thomas Frenzel and Matthias Liero. Effective diffusion in thin structures via generalized gradient systems and EDP-convergence. *Discrete & Continuous Dynamical Systems-S*, page 0, 2018.

[FLSS07] Rupert L Frank, Elliott H Lieb, Robert Seiringer, and Heinz Siedentop. Müller's exchange-correlation energy in density-matrix-functional theory. *Physical Review A*, 76(5):052517, 2007.

[FM16] Max Fathi and Jan Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806, 2016.

[FMP20] Dominik Forkert, Jan Maas, and Lorenzo Portinale. Evolutionary Γ-convergence of entropic gradient flow structures for Fokker–Planck equations in multiple dimensions. *Submitted*, arXiv:2008.10962, 2020.

[FNVDB18] Rupert L Frank, Phan Thành Nam, and Hanne Van Den Bosch. The maximal excess charge in Müller density-matrix-functional theory. In *Annales Henri Poincaré*, volume 19, pages 2839–2867. Springer, 2018.

[Gig11] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(\mathrm{P}_2(M), W_2)$. *Methods Appl. Anal.*, 18(2):127–158, 2011.

[Gil75] Thomas L Gilbert. Hohenberg–Kohn theorem for nonlocal external potentials. *Physical Review B*, 12(6):2111, 1975.

[GKM20] Peter Gladbach, Eva Kopfer, and Jan Maas. Scaling limits of discrete optimal transport. *SIAM J. Math. Anal.*, 52(3):2759–2802, 2020.

[GKMP20] Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Homogenisation of one-dimensional discrete optimal transport. *J. Math. Pures Appl. (9)*, 139:204–234, 2020.

[GKMP21] Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Discrete-to-continuum limits of dynamical transport problems on periodic graphs. *In preparation*, 2021.

[GKR20] Augusto Gerolin, Anna Kausamo, and Tapio Rajala. Multi-marginal Entropy-Transport with repulsive cost. *Calc. Var.*, 59(90), 2020.

[GM13] Nicola Gigli and Jan Maas. Gromov-Hausdorff convergence of discrete transportation metrics. *SIAM J. Math. Anal.*, 45(2):879–899, 2013.

[GP15] Tryphon T Georgiou and Michele Pavon. Positive contraction mappings for classical and quantum schrödinger systems. *Journal of Mathematical Physics*, 56(3):033301, 2015.

[GR19] Klaas JH Giesbertz and Michael Ruggenthaler. One-body reduced density-matrix functional theory in finite basis sets at elevated temperatures. *Physics Reports*, 806:1–47, 2019.

[Gri09] Alexander Grigor'yan. *Heat kernel and analysis on manifolds*, volume 47 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.

[GS10] Alfred Galichon and Bernard Salanié. Matching with trade-offs: Revealed preferences over competing characteristics. *CEPR Discussion Paper No. DP7858*, 2010.

[GST09] Ugo Gianazza, Giuseppe Savaré, and Giuseppe Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Ration. Mech. Anal.*, 194(1):133–220, 2009.

[GT20] Nicolás García Trillos. Gromov–Hausdorff limit of Wasserstein spaces on point clouds. *Calc. Var. Partial Differential Equations*, 59(2):Paper No. 73, 43, 2020.

[HK64] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(B864), 1964.

[HKS20] Martin Heida, Markus Kantner, and Artur Stephan. Consistency and convergence for a family of finite volume discretizations of the Fokker–Planck operator. *arXiv preprint arXiv:2002.09385*, 2020.

[HPR19] Martin Heida, Robert I. A. Patterson, and D. R. Michiel Renger. Topologies and measures on the space of functions of bounded variation taking values in a Banach or metric space. *J. Evol. Equ.*, 19(1):111–152, 2019.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[Kan42] L. Kantorovitch. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.

[Kas07] Moritz Kassmann. Harnack inequalities: an introduction. *Bound. Value Probl.*, pages Art. ID 81415, 21, 2007.

[KR17] Johan Karlsson and Axel Ringh. Generalized Sinkhorn iterations for regularizing inverse problems using optimal mass transport. *SIAM Journal on Imaging Sciences*, 10(4):1935–1962, 2017.

[KS03] Kazuhiro Kuwae and Takashi Shioya. Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. *Comm. Anal. Geom.*, 11(4):599–673, 2003.

[LÍ4] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[Lev76] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 12:6062–6065, 76.

[Lie02] Elliott H Lieb. Density functionals for Coulomb systems. In *Inequalities*, pages 269–303. Springer, 2002.

[LM10] Stefano Lisini and Antonio Marigonda. On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscripta Math.*, 133(1-2):197–224, 2010.

[LMPR17] Matthias Liero, Alexander Mielke, Mark A. Peletier, and D. R. Michiel Renger. On microscopic origins of generalized gradient structures. *Discrete & Continuous Dynamical Systems - S*, 10, 2017.

[LS10] Elliott H. Lieb and Robert Seiringer. *The stability of matter in quantum mechanics*. Cambridge University Press, Cambridge, 2010.

[LV07] John Lott and Cédric Villani. Weak curvature conditions and functional inequalities. *Journal of Functional Analysis*, 245(1):311–333, 2007.

[LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[Maa11] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[McC95] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.

[Men15] LM Mentel. *Reduced Density Matrix inspired approaches to electronic structure theory*. PhD thesis, Vrije Universiteit Amsterdam:, 2015.

[Mer65] N David Mermin. Thermal properties of the inhomogeneous electron gas. *Physical Review*, 137(5A):A1441, 1965.

[Mie11] Alexander Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.

[Mie13] Alexander Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations*, 48(1-2):1–31, 2013.

[Mie16a] Alexander Mielke. Deriving effective models for multiscale systems via evolutionary Γ-convergence. In *Control of self-organizing nonlinear systems*, Underst. Complex Syst., pages 235–251. Springer, [Cham], 2016.

[Mie16b] Alexander Mielke. On evolutionary Γ-convergence for gradient systems. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 187–249. Springer, [Cham], 2016.

[MM17] Markus Mittnenzweig and Alexander Mielke. An entropic gradient structure for Lindblad equations and couplings of quantum systems to macroscopic models. *Journal of Statistical Physics*, 167(2):205–233, 2017.

[MM20] Jan Maas and Alexander Mielke. Modeling of chemical reaction systems with detailed balance using gradient structures. *J. Stat. Phys.*, 181(6):2257–2303, 2020.

[MMP21] Alexander Mielke, Alberto Montefusco, and Mark A Peletier. Exploring families of energy-dissipation landscapes via tilting: three types of edp convergence. *Continuum Mechanics and Thermodynamics*, pages 1–27, 2021.

[Mon81] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*, 1781.

[Mos61] Jürgen Moser. On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.*, 14:577–591, 1961.

[Mos64] Jürgen Moser. A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.*, 17:101–134, 1964.

[Mos94] Umberto Mosco. Composite media and asymptotic Dirichlet forms. *J. Funct. Anal.*, 123(2):368–421, 1994.

[MPR14] Alexander Mielke, Mark A Peletier, and DR Michiel Renger. On the relation between gradient flows and the large-deviation principle, with applications to markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

[MR15] Alexander Mielke and Tomáš Roubíček. *Rate-independent systems*, volume 193 of *Applied Mathematical Sciences*. Springer, New York, 2015. Theory and application.

[MS19] Alexander Mielke and Artur Stephan. Coarse-graining via EDP-convergence for linear fast-slow reaction systems. *arXiv preprint arXiv:1911.06234*, 2019.

[Mül84] AMK Müller. Explicit approximate relation between reduced two-and one-particle density matrices. *Physics Letters A*, 105(9):446–452, 1984.

[MV20] Léonard Monsaingeon and Dmitry Vorotnikov. The schrödinger problem on the non-commutative Fisher-Rao space. *arXiv preprint arXiv:2007.09042*, 2020.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[OV00] Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[PC19] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[PCVS19] Gabriel Peyré, Lénaïc Chizat, François-Xavier Vialard, and Justin Solomon. Quantum entropic regularization of matrix-valued optimal transport. *European Journal of Applied Mathematics*, 30(6):1079–1102, 2019.

[Per05] Katarzyna Pernal. Effective potential for natural spin orbitals. *Physical review letters*, 94(23):233002, 2005.

[PG15] Katarzyna Pernal and Klaas JH Giesbertz. Reduced density matrix functional theory (RDMFT) and linear response time-dependent rdmft (TD-RDMFT). In *Density-Functional Methods for Excited States*, pages 125–183. Springer, 2015.

[PS19] Lorenzo Portinale and Ulisse Stefanelli. Penalization via global functionals of optimal-control problems for dissipative evolution. *Adv. Math. Sci. Appl., 28*, pages 425–447, 2019.

[RCLO18] Ernest K Ryu, Yongxin Chen, Wuchen Li, and Stanley Osher. Vector and matrix optimal mass transport: theory, algorithm, and applications. *SIAM Journal on Scientific Computing*, 40(5):A3675–A3698, 2018.

[RP08] Ryan Requist and Oleg Pankratov. Generalized Kohn-Sham system in one-matrix functional theory. *Physical Review B*, 77(23):235121, 2008.

[Rus95] Ludger Ruschendorf. Convergence of the iterative proportional fitting procedure. *The Annals of Statistics*, 23(4):1160–1174, 1995.

[RW98] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[San15] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[SC02] Laurent Saloff-Coste. *Aspects of Sobolev-type inequalities*, volume 289 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2002.

[Sch31] Erwin Schrödinger. *Über die umkehrung der naturgesetze*. Verlag Akademie der wissenschaften in kommission bei Walter de Gruyter u. Company, 1931.

[Sch19] Robert Schade. *New Methods for the Ab-initio Simulation of Correlated Systems*. PhD thesis, Niedersächsische Staats-und Universitätsbibliothek Göttingen, 2019.

[Sin64] Richard Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The annals of mathematical statistics*, 35(2):876–879, 1964.

[SS04] Etienne Sandier and Sylvia Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[Ste08] Ulisse Stefanelli. The Brezis–Ekeland principle for doubly nonlinear equations. *SIAM Journal on Control and Optimization*, 47(3):1615–1642, 2008.

[Stu96] K. T. Sturm. Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pures Appl. (9)*, 75(3):273–297, 1996.

[Stu06] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[SW19] Bernhard Schmitzer and Benedikt Wirth. Dynamic models of Wasserstein-1-type unbalanced transport. *ESAIM Control Optim. Calc. Var.*, 25:Paper No. 23, 54, 2019.

[vBH72] Ulf von Barth and Lars Hedin. A local exchange-correlation potential for the spin polarized case. i. *Journal of Physics C: Solid State Physics*, 5(13):1629, 1972.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil08] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[vL07] R van Leeuwen. Functionals of the one-particle density matrix: Uniqueness theorems and v-representability in equilibrium and time-dependent systems. *University of Jyväskylä, Jyväskylä, Finland, private communication (unpublished)*, 2007.

[VR87] G Vignale and Mark Rasolt. Density-functional theory in strong magnetic fields. *Physical review letters*, 59(20):2360, 1987.

[Wir18] Melchior Wirth. A noncommutative transport metric and symmetric quantum markov semigroups as gradient flows of the entropy. *arXiv preprint arXiv:1808.05419*, 2018.

[Zam86] JC Zambrini. Stochastic mechanics according to E. Schrödinger. *Physical review A*, 33(3):1532, 1986.

APPENDIX A

Harnack inequalities on graphs

In this appendix, we prove Harnack inequalities for diffusion equations on finite-state Markov chains satisfying a suitable ellipticity condition. This is a generalisation of the classical result from Delmotte [Del99] and the proof follows similar ideas, based on the Moser's iteration technique. We here sketch the main steps of the proof and discuss an application to the finite-volume framework. This allows us to show Hölder regularity for solutions to the discrete Fokker–Planck equations, property that plays a key role in the proof of the evolutionary Γ-convergence result of discrete gradient-flow structures presented in Chapter 4 (that is, we prove Proposition 4.4.4).

A.1 Harnack inequality for diffusion equations on finite Markov chains

The importance of parabolic Harnack inequalities for diffusion processes is well-established in the literature and its origins go back to Carl Gustav Axel von Harnack in the 19th century, in the context of harmonic functions in Euclidean domains. One particularly significant application of Harnack inequalities is the Hölder continuity of the correspondent solution – a parabolic version of the Harnack inequality even implies Hölder regularity of the associated flow. We refer the reader to [Kas07] for a general introduction to the topic.

Given their fundamental impact, Harnack inequalities have been widely studied. Let us recall the works of Grigor'yan [Gri09] and Saloff-Coste [SC02] for Laplace–Beltrami operators on Riemaniann manifolds, where equivalent characterisations for parabolic Harnack inequalities have been investigated. In particular, they showed that a parabolic Harnack inequality is equivalent to a Poincaré inequality together with a doubling condition for the volume measure; thus, highlighting a deep connection between properties of solutions to the heat flow on a manifold and more geometric and analytic aspects of the space itself. Similar results have been extended to the case of symmetric diffusions on metric measure spaces by Sturm [Stu96] and to random walks on graphs by Delmotte [Del99]. All these results concern a classical linear diffusion regime. Other regimes have been considered in [BBK09].

Albeit the existence of an involved history of works, the particular case of linear diffusions on locally finite graphs appeared, to our knowledge, slightly incomplete. In particular, the main reference work in this setting, given by [Del99], deals with parabolic Harnack inequalities for diffusions where the reference measure μ and the jump kernel j are related by the condition

$$\mu(x) = \sum_y j(x,y). \tag{A.1}$$

The goal of this appendix is to show that the results holds in a slightly more general setting, extending the result of Delmotte.

We consider a Markov chain on a finite-state space $\mathcal{X}$ with rates $w(x,y)$, invariant measure $\pi(x)$, and generator

$$\mathcal{L}f(x) := \frac{1}{\pi(x)} \sum_{y\in\mathcal{X}} w(x,y)\big(f(y) - f(x)\big), \quad x \in \mathcal{X}\,.$$

We denote by $X_t : (\Omega, \mathbb{P}) \to \mathcal{X}$ the associated time-continuous Markov process and denote by $m_t := \mathsf{law}(X_t) = (X_t)_\# \mathbb{P} \in \mathscr{P}(\mathcal{X})$. The corresponding density solves

$$r_t := \frac{\mathrm{d}m_t}{\mathrm{d}\pi}, \quad \dot{r}_t(x) = \mathcal{L}r_t(x), \quad t > 0,\ x \in \mathcal{X}. \tag{A.2}$$

Assumption (ellipticity): there exists $1 \le C < \infty$ such that for every $x \in \mathcal{X}$.

$$w(x,y) \ge C\mu(y), \quad \forall x \sim y\,. \tag{A.3}$$

Notation. We introduce the finite measures

$$J(\,\mathrm{d}x,\,\mathrm{d}y) := \omega(x,y)C(\,\mathrm{d}x,\,\mathrm{d}y), \quad \mu(\,\mathrm{d}x) = \pi(x)C(\,\mathrm{d}x)\,,$$

where C denotes the counting measure (respectively) on $\mathcal{X} \times \mathcal{X}$ and $\mathcal{X}$. Denote by $\mathsf{B}(x,R)$ the ball of radius R and center $x \in \mathcal{X}$, with respect to the graph distance $\mathsf{d} = d_{\mathsf{gra}}$ on $(\mathcal{X}, w)$, and let $\mathscr{V}(x,r) := \mu\big(\mathsf{B}(x,r)\big)$ be its volume.

Definition A.1.1 (Volume doubling)**.** We say $(\mathcal{X}, \omega, \pi)$ satisfies the volume doubling condition with constant $c_D \in (0, +\infty)$ if for any $x \in \mathcal{X}$ and $r > 0$ we have the doubling property $\mathscr{V}(K, 2r) \le c_D \mathscr{V}(K, r)$.

Definition A.1.2 (Weak Poincaré inequality (PI))**.** We say $(\mathcal{X}, w, \pi)$ satisfies a *weak Poincaré inequality* if there exists a constant $c_P < \infty$, such that, for any ball $\mathsf{B}_r = \mathsf{B}(K_0, r)$ and any $f : \mathcal{X} \to \mathbb{R}$,

$$\int_{\mathsf{B}_r} \Big(f(x) - \overline{f}_{\mathsf{B}_r}\Big)^2 \mathrm{d}\mu(x) \le c_P r^2 \int_{\mathsf{B}_{cr}^2} \Big(f(x) - f(y)\Big)^2 \mathrm{d}J(x,y),$$

where $\overline{f}_{\mathsf{B}_r} = \fint_{\mathsf{B}_r} f \,\mathrm{d}\mu$.

Definition A.1.3 (Harnack inequality)**.** We say that $q : \mathbb{R}^+ \times \mathcal{X} \to \mathbb{R}_+$ satisfies the *Harnack inequality* if there exists constants $C_H \in \mathbb{R}_+$ and $1 < \theta_1 < \theta_2 < \theta_3 < +\infty$ such that for every $t_0 > 0$, $x_0 \in \mathcal{X}$, $R > 0$ it holds

$$\sup_{Q_-} q \le C_H \inf_{Q_+} q$$

where Q_-, Q_+ denotes the cylinders given by

$$Q_- := [t_0, t_0 + \theta_1 R^2] \times \mathsf{B}(x_0, R), \quad Q_+ := [t_0 + \theta_2 R^2, t_0 + \theta_3 R^2] \times \mathsf{B}(x_0, R).$$

It is possible to show that if the Harnack inequality holds for C_H, θ_1, θ_2, and θ_3, then it also holds for any other choice $0 < \theta_1' < \theta_2' < \theta_3' < \infty$ with a constant C_H' depending only on C_H, θ_i, and θ_i'.

Theorem A.1.4 (Poincaré + Doubling implies Harnack)**.** *Suppose that (X, w, π) satisfies the assumption (A.3), volume doubling, and the weak Poincaré inequality. Then the solutions r_t of the equation (A.2) satisfy the Harnack inequality as in Definition A.1.3 with a constant C_H depending only on C, c_D, and c_P.*

This theorem has been proved by T. Delmotte in [Del99, Theorem 2.1] under the additional assumption that

$$\mu(x) = \sum_{y \in \mathcal{X}} w(x, y), \quad \forall x \in \mathcal{X}. \tag{A.4}$$

In this short note we shortly discuss the main strategy and check that the same result holds true without assuming (A.4).

Remark A.1.5. The volume doubling condition, the assumption (A.3), and the weak Poincaré inequality imply the strong Poincaré inequality ($c = 1$), see [Bar17, Corollary A.51].

The main strategy of Delmotte follows the one of Moser for parabolic [Mos64] and elliptic [Mos61] Harnack inequalities for elliptic operators in $\mathbb{R}^d$. In particular, the proof proceeds as follows:

1. One proves a Sobolev's inequality as a consequence of Poincaré, volume doubling, and the ellipticity assumption, see Theorem A.1.6.
2. Any super/subsolution of equation (A.2) are proved to satisfy Caccioppoli-type estimates, see Theorem A.1.9.
3. A weighted Poincaré inequality is proved from the weak Poincaré inequality, volume doubling, and the ellipticity condition, see Theorem A.1.7.
4. Sobolev and Caccioppoli allows us to run the Moser's iterations for both super/subsolutions to (A.2), see Theorem A.1.10 and Theorem A.1.12.
5. The weighted Poincaré inequality links the L^p-norms with positive exponent with the ones with negative exponent, and the proof of Harnack is complete.

Let us explain the details of this plan. We start with Sobolev and weighted Poincaré. From now on, we work under the assumptions of Theorem A.1.4.

For simplicity, we write $\lesssim$, $\approx$ whenever the inequalities holds up to multiplication for a positive constant that depends only on C, c_P, and c_D.

Theorem A.1.6 (Sobolev's inequality)**.** *There exists $\theta > 1$ depending only on c_D such that for every $x_0 \in \mathcal{X}$, $r > 0$, $f \in \mathbb{R}^{\mathsf{B}}$, $\mathsf{B} := \mathsf{B}(x_0, r)$ we have*

$$\left(\fint_{\mathsf{B}} f^{2\theta} \, \mathrm{d}\mu \right)^{\frac{1}{\theta}} \lesssim \frac{1}{\mathscr{V}(\mathsf{B})} \left(r^2 \int_{\mathsf{B}^2} \big(f(y) - f(x)\big)^2 \, \mathrm{d}J(x, y) + \int_{\mathsf{B}} f^2 \, \mathrm{d}\mu \right). \tag{Sob}$$

Delmotte does not provide the proof of this result, but it can be proved following the same ideas as in the proof of Saloff-Coste [SC02].

Theorem A.1.7 (Weighted Poincaré inequality)**.** *For any given* $x_0 \in \mathcal{X}$, $r \in \mathbb{N}$, *set*

$$\psi(x) := 1 - \frac{\mathsf{d}(x_0, x)}{r} \in (0, 1] \quad \text{on } \mathsf{B}(x_0, r)\,.$$

Then for every $f \in \mathbb{R}^{\mathsf{B}}$, $\mathsf{B} := \mathsf{B}(x_0, r)$ *we have*

$$\int_{\mathsf{B}} \psi^2 \big(f - f_{\mathsf{B}_\psi}\big)^2 \,\mathrm{d}\mu \lesssim r^2 \int_{\mathsf{B}^2} \big(\psi(x) \wedge \psi(y)\big)\big(f(y) - f(x)\big)^2 \,\mathrm{d}J(x, y)\,, \tag{WP}$$

where f_{B_ψ} *is weighted average that minimises the LHS, i.e.*

$$f_{\mathsf{B}_\psi} := \frac{\displaystyle\int_{\mathsf{B}} \psi^2 f^2 \,\mathrm{d}\mu}{\displaystyle\int_{\mathsf{B}} \psi^2 \,\mathrm{d}\mu}\,.$$

The proof of this estimate can be found in [Del99, Proposition 2.2]. The proof of Moser is even for more general weights, see [Mos64, Lemma 3], and it is consequence of the strong Poincaré inequality, see Remark A.1.5.

The next step is to prove the Caccioppoli-type estimates for sub-supersolutions to (A.2). We write $u_t := u(t, \cdot)$.

Definition A.1.8. u is a positive subsolution on $Q := I \times \mathsf{B}(x_0, r)$ if $u \geq 0$ and

$$m(x)\partial_t u_t(x) \leq \sum_{y \in \mathcal{X}} w(x, y)\big(u_t(y) - u_t(x)\big)\,, \quad \forall (t, x) \in I \times \mathsf{B}(x_0, r - 1)\,.$$

r is a positive supersolution on $Q := I \times \mathsf{B}(x_0, r)$ if $u \geq 0$ and

$$m(x)\partial_t u_t(x) \geq \sum_{y \in \mathsf{B}} w(x, y)\big(u_t(y) - u_t(x)\big)\,, \quad \forall (t, x) \in Q\,,$$

where we set $\mathsf{B} = \mathsf{B}(x_0, r)$.

Notation: for $s_1, s_2 \in \mathbb{R}$, $\mathsf{B} = \mathsf{B}(x, r)$, $Q = I \times \mathsf{B}$, and $\sigma \in (0, 1/2)$, we define:

$$\begin{gathered} \mathsf{B}_\sigma := (1 - \sigma)\mathsf{B}, \quad \mathsf{B}_\sigma \subset \mathsf{B}\,, \\ I_\sigma := \Big[(1 - \sigma^2)s_1 + \sigma^2 s_2, s_2\Big]\,, \quad I'_\sigma := \Big[s_1, \sigma^2 s_2 + (1 - \sigma^2)s_2,\Big]\,, \\ I''_\sigma := \Big[(1 - \sigma^2)s_1 + \sigma^2 s_2, \sigma^2 s_2 + (1 - \sigma^2)s_2\Big]\,, \quad I_\sigma\,, I'_\sigma\, I''_\sigma \subset I\,, \\ Q_\sigma := I_\sigma \times \mathsf{B}_\sigma\,, \quad Q'_\sigma := I'_\sigma \times \mathsf{B}_\sigma\,, \quad Q''_\sigma := I''_\sigma \times \mathsf{B}_\sigma\,, \quad Q_\sigma\,, Q'_\sigma\,, Q''_\sigma \subset Q\,. \end{gathered}$$

Theorem A.1.9 (Caccioppoli-type estimates)**.** *Let* u *be a positive subsolution on* Q*. Then we have*

$$\sup_{t \in I_\sigma} \int_{\mathsf{B}_\sigma} u_t^2 \,\mathrm{d}\mu \lesssim \frac{1}{\sigma^2 r^2} \int_Q u^2 \,\mathrm{d}\mu \,\mathrm{d}t\,, \tag{C1}$$

$$\frac{1}{2} \int_{Q_\sigma} \big(u_t(y) - u_t(x)\big)^2 \,\mathrm{d}\mu \,\mathrm{d}t \lesssim \frac{1}{\sigma^2 r^2} \int_Q u^2 \,\mathrm{d}\mu \,\mathrm{d}t\,. \tag{C2}$$

If u is a positive supersolution on Q*, the same holds but with* Q'_σ *instead of* Q_σ*.*

This is the discrete counterpart of the classical continuous estimate for $\mathcal{L} = \Delta$

$$\int \psi^2 \partial_t (u^2) \,\mathrm{d}x \,\mathrm{d}t + \int \psi^2 |\nabla u|^2 \,\mathrm{d}x \,\mathrm{d}t \le 4 \int |\nabla \psi|^2 u^2 \,\mathrm{d}x \,\mathrm{d}t \,, \quad \psi \in C_c^\infty \,.$$

The proof of this estimates can be found in [Del99] inside the proof of Lemma 2.5 and it follows directly from the definition of sub-supersolution, using suitable test functions. Only the symmetry of w is needed.

We proceed defining the L^p averages of $u \in \mathbb{R}^Q$, $Q \subset \mathbb{R} \times \mathcal{X}$ as

$$\mathcal{M}(u, p, Q) := \left(\fint_Q u^{2p} \,\mathrm{d}\mu \,\mathrm{d}t \right)^{\frac{1}{p}} .$$

The Moser's iterations technique consists in estimating the above defined averages using (Sob) and (C1), (C2) and iterate them. Note that $p = +\infty$ corresponds to the $\sup_Q u$ and $p = -\infty$ to $\inf_Q u$.

Theorem A.1.10 (Moser's fundamental estimates). *Set $k := 2 - 1/\theta > 1$, where θ is the one of Theorem A.1.6. Set $Q := [0, r^2] \times \mathsf{B}$, for $r > 0$.*

(a) If v is a positive subsolution on Q and $\frac{1}{r} \le \sigma \le \frac{1}{2}$, then

$$\mathcal{M}(v, k, Q_\sigma) \le \left(\frac{A}{\sigma^2} \right)^{\frac{1}{k}} \mathcal{M}(v, 1, Q) \,. \tag{M1}$$

(b) If v is a positive supersolution on Q and $\frac{1}{r} \le \sigma \le \frac{1}{2}$, then

$$\mathcal{M}(v, k, Q'_\sigma) \le \left(\frac{A}{\sigma^2} \right)^{\frac{1}{k}} \mathcal{M}(v, 1, Q) \,. \tag{M2}$$

In both cases, A is a constant such that $A \lesssim 1$.

The Moser's fundamental estimates are direct consequence of Theorem A.1.6 and Theorem A.1.9. The proof, under the additional assumption (A.4), can be found in [Del99, Lemma 2.5]. The Theorem says that along sub-supersolutions one can estimate L^q norms with exponent q with Lebesgue norms with a smaller exponent. The key observation is that powers of solutions are sub-supersolutions as well and this allows us to iterate these estimates up to $p = +\infty$ and $p = -\infty$.

Lemma A.1.11 (Powers of solutions). *Let u be a positive solution on $Q = I \times \mathsf{B}$.*

(a) u^p is a subsolution on Q for every $p \le 0$ and $p \ge 1$.

(b) u^p is a supersolution on Q for every $0 \le p \le 1$.

The proof is straightforward consequence of convexity/concavity of $f(t) = t^p$, see [Del99, Lemma 2.6]. Lemma A.1.11 together with iterations of Theorem A.1.10 yields the following result.

Theorem A.1.12 (Moser's iterations). *Let $Q = [0, r^2] \times \mathsf{B}$, u be a positive solution on Q, and $0 \le \delta \le \frac{1}{2}$. Then for all $p > 0$*

$$\mathcal{M}(u, -p, Q) \lesssim \left(D\delta^{-\gamma}\right)^{\frac{1}{p}} \inf_{Q_\delta} u^2 , \qquad (\mathsf{M}_{-\infty})$$

$$\sup_{Q''_\delta} u^2 \lesssim \left(D\delta^{-\gamma}\right)^{\frac{1}{p}} \mathcal{M}(u, p, Q) . \qquad (\mathsf{M}_{+\infty})$$

for some D, $\gamma \lesssim 1$.

The proof assuming (A.4) can be found in [Del99, Lemma 2.7]. It follows by iterations of Theorem A.1.10 to $v = u^q$, which are sub-supersolutions thanks to Lemma A.1.11.

The last step consists in linking negative and positive exponents, in the form

$$\mathcal{M}(u, -p, Q) \lesssim \mathcal{M}(u, p, Q), \quad \text{for small } p > 0 . \qquad (\text{A.5})$$

Delmotte follows the original ideas of Moser, which is to study the function $v = -\log u$. The reason why this is the right thing to do can be intuitively understood by the fact that

$$\lim_{p \to 0} \left(\int_Q u^p \, \mathrm{d}\mu \, \mathrm{d}t \right)^{\frac{1}{p}} = \exp \int_Q \log u \, \mathrm{d}\mu \, \mathrm{d}t , \quad u \ge 0 .$$

The key lemma is [Del99, Lemma 2.8], which Delmotte describes as "BMO-type estimates" and studies the oscillations of $\log u$, for u solution to (A.2). The proof is quite involved but it does only rely on volume doubling and Poincaré. [Del99, Lemma 2.8] together with Theorem A.1.12 yields the Harnack inequality and the proof of Theorem A.1.4 is complete. See [Del99, Section 2.5, p. 210] for the details.

A.2 Proof of Proposition 4.4.4

We present a proof of Proposition 4.4.4 obtained as an application of the result showed in the previous section, that is Theorem A.1.4. To this purpose, we prove a *volume doubling property* and a *weak Poincaré inequality* for the discrete Fokker–Planck equation on ζ-regular meshes, which hold uniformly in the mesh size.

More precisely, for a ζ-regular mesh $\mathcal{T}$ we consider the weighted graph (Γ, μ) given by

$$\Gamma := \mathcal{T}, \qquad \mu_{KL} = \mu_{LK} := [\mathcal{T}]^{2-d} w_{KL} \approx 1 .$$

The associated generator $\mathcal{L}$ and the invariant measure m are given by

$$\left(\mathcal{L}f\right)(K) := \frac{1}{m(K)} \sum_{L \sim K} \mu_{KL}\left(f(L) - f(K)\right), \qquad m(K) := [\mathcal{T}]^{-d} \pi_{\mathcal{T}}(K) \approx 1 . \qquad (\text{A.6})$$

Let $\mathsf{d} := d_{\mathsf{gra}}$ be the graph distance on Γ (defined as the length of a minimal path connecting cells using the discrete metric). For $K \in \mathcal{T}$, let $\mathsf{B}(K, r)$ denote the closed ball of radius $r \ge 0$ in (Γ, d), and let $\mathscr{V}(K, r) := m\big(\mathsf{B}(K, r)\big)$ be its volume. Closed balls in euclidean space will be denoted by $B(x, r)$.

First we show that (Γ, μ, m) satisfies a volume doubling property, with proportionality constants depending only on Ω, ζ, and $\mathbf{m}$.

Proposition A.2.1 (Distance comparison and volume doubling)**.** *There exists a constant $c_1 < \infty$, depending only on depending only on Ω, ζ, such that the following properties hold:*

(i) For any $K, L \in \mathcal{T}$ we have the distance comparison

$$\frac{1}{2}|x_K - x_L| \leq \mathsf{d}(K, L)[\mathcal{T}] \leq c_1|x_K - x_L|. \tag{A.7}$$

(ii) For any $K \in \mathcal{T}$ and $r > 0$ we have the volume bounds

$$1 \vee \Big(r - c_1\Big)_+^d \lesssim \mathscr{V}(K, r) \lesssim \Big(r + 1\Big)^d. \tag{A.8}$$

(iii) For any $K \in \mathcal{T}$ and $r > 0$ we have the doubling property $\mathscr{V}(K, 2r) \lesssim \mathscr{V}(K, r)$.

Proof. (i): The upper bound has been proved in [GKM20, Lemma 1.12]. To prove the lower bound, fix $K, L \in \mathcal{T}$. By definition of the graph distance, there exists a path $K = K_0 \sim K_1 \sim \ldots \sim K_{n-1} \sim K_n = L$ of length $n \leq \mathsf{d}(K, L)$ connecting K and L. Consequently,

$$|x_K - x_L| \leq \sum_{i=1}^n |x_i - x_{i-1}| \leq 2n[\mathcal{T}] \leq 2\mathsf{d}(K, L)[\mathcal{T}].$$

(ii): Fix $K \in \mathcal{T}$ and suppose that $\mathsf{d}(K, L) \leq r$. By (i), we have $|x_K - x_L| \leq 2r[\mathcal{T}]$, which implies that $L \subseteq B\Big(x_K, (2r + 1)[\mathcal{T}]\Big)$. Consequently,

$$\mathscr{V}(K, r) = \sum_{L:\mathsf{d}(K,L)\leq r} m(L) \leq [\mathcal{T}]^{-d}\,\mathbf{m}\Big(B(x_K, (2r + 1)[\mathcal{T}])\Big) \lesssim \Big(r + 1\Big)^d,$$

which is the desired upper bound.

Since $K \in \mathsf{B}(K, r)$ and $m(K) \gtrsim 1$, the lower bound follows immediately if $r \leq c_1$. We thus suppose that $r > c_1$, and observe that any cell $L \in \mathcal{T}$ with $|x_K - x_L| \leq [\mathcal{T}]r/c_1$ satisfies $\mathsf{d}(K, L) \leq r$. As the cells $L \in \mathcal{T}$ with $|x_K - x_L| \leq [\mathcal{T}]r/c_1$ cover the ball $B(x_K, \Big(r/c_1 - 1\Big)[\mathcal{T}])$, and each of these cells has euclidean volume $\lesssim [\mathcal{T}]^d$, there must be at least $\gtrsim (r - c_1)^d$ of such cells. Since $m(L) \gtrsim 1$, we infer that $\mathscr{V}(K, r) \gtrsim (r - c_1)^d$.

(iii): This follows immediately from (ii). □

Next we show a weak Poincaré inequality with constant depending only on Ω, ζ, and $\mathbf{m}$.

Proposition A.2.2 (Weak Poincaré inequality)**.** *There exists a constant $c < \infty$, depending only on Ω and ζ, such that, for any ball $\mathsf{B}_r = \mathsf{B}(K_0, r)$ and any $f : \Gamma \to \mathbb{R}$,*

$$\sum_{K\in\mathsf{B}_r} \Big(f(K) - \overline{f}_{\mathsf{B}_r}\Big)^2 m(K) \lesssim r^2 \sum_{K,L\in\mathsf{B}_{cr}} \Big(f(K) - f(L)\Big)^2 \mu_{KL},$$

where $\overline{f}_{\mathsf{B}_r} = \Big(m(\mathsf{B}_r)\Big)^{-1} \int_{\mathsf{B}_r} f \,\mathrm{d}m$.

Proof. We adapt the argument from [EGH00, Lemma 3.7]; see also [GKM20, Proposition 4.5], taking into account that the balls B_r do not correspond to convex subsets in euclidean space.

Fix $K_0 \in \mathcal{T}$ and $r > 0$. The desired bound can equivalently be stated as

$$\sum_{K,L \in \mathsf{B}_r} \Big(f(K) - f(L)\Big)^2 m(K) m(L) \lesssim r^2 m(\mathsf{B}_r) \sum_{K,L \in \mathsf{B}_{cr}} \Big(f(K) - f(L)\Big)^2 \mu_{KL}. \tag{A.9}$$

If $r < 1$, we have $\mathsf{B}_r = \{K_0\}$, hence the left-hand side vanishes and the claim is trivial.

If $1 \le r < 2c_1$ (with c_1 being the constant from (A.7)), the estimate (A.9) (with $c = 1$) follows immediately by ζ-regularity.

We thus assume from now on that $r \ge 2c_1$, so that $m(\mathsf{B}_r) \eqsim r^d$ by Proposition A.2.1. Define $\psi : \Omega \to \mathbb{R}$ by $\psi(x) = f(K)$ for $x \in K$. Let us write $\tilde{B}_r := \{x \in \Omega \ : \ x \in K \text{ for some } K \in \mathsf{B}_r\}$. For $K, L \in \mathcal{T}$ and $x, y \in \mathbb{R}^d$, set $\chi_{KL}(x,y) = 1$ if $x, y \in \tilde{B}_r$, $K \sim L$, the interface Γ_{KL} intersects the line segment from x to y, and $(y - x) \cdot (x_L - x_K) > 0$. Otherwise, set $\chi_{KL}(x,y) = 0$.

For $x, y \in \tilde{B}_r$, the volume comparison bounds in (A.8) imply that the line segment from x to y is contained in a ball $\tilde{B}_{cr}$ for some constant $c \ge 1$ depending only on Ω and ζ. Hence, for a.e. $x, y \in \tilde{B}_r$,

$$|\psi(x) - \psi(y)| \le \sum_{K,L \in \mathsf{B}_{cr}} |f(K) - f(L)| \chi_{KL}(x,y).$$

For $K, L \in \mathcal{T}$ and $z \in \mathbb{R}^d$, set $\alpha_{KL}(z) := \frac{z}{|z|} \cdot \frac{x_L - x_K}{|x_K - x_L|}$. Using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\Big(\psi(x) - \psi(y)\Big)^2 \le \Bigg(\sum_{K,L \in \mathsf{B}_{cr}} & \frac{\Big(f(L) - f(K)\Big)^2}{\alpha_{KL}(y-x)} \frac{\chi_{KL}(x,y)}{|x_K - x_L|} \Bigg) \\ & \times \Bigg(\sum_{K,L \in \mathsf{B}_{cr}} \alpha_{KL}(y-x) |x_K - x_L| \chi_{KL}(x,y) \Bigg).\end{aligned}$$

For fixed x and y, let $L_0, L_1, \ldots, L_N$ be the cells that subsequently intersect the line segment from x to y. We have

$$\begin{aligned}\sum_{K,L \in \mathsf{B}_{cr}} \alpha_{KL}(y-x) |x_K - x_L| \chi_{KL}(x,y) &= \sum_{i=1}^N \alpha_{L_{i-1},L_i}(y-x) |x_{L_{i-1}} - x_{L_i}| \\ = \frac{y-x}{|y-x|} \cdot \sum_{i=1}^N (x_{L_i} - x_{L_{i-1}}) &= \frac{y-x}{|y-x|} \cdot (x_{L_N} - x_{L_0}) \lesssim r[\mathcal{T}]\,.\end{aligned}$$

Using the change of variables $z = y - x$ we estimate, for some $c' < \infty$,

$$\begin{aligned}\int_{\tilde{B}_{cr}} \int_{\tilde{B}_{cr}} \frac{\chi_{KL}(x,y)}{\alpha_{KL}(y-x)} \, \mathrm{d}x \, \mathrm{d}y &\le \int_{B(0,c'r)} \frac{1}{\alpha_{KL}(z)} \int_{\mathbb{R}^d} \chi_{KL}(x, x+z) \, \mathrm{d}x \, \mathrm{d}z \\ &\lesssim \mathcal{H}^{d-1}(\Gamma_{KL}) \int_{B(0,c'r)} |z| \, \mathrm{d}z \lesssim (r[\mathcal{T}])^{d+1} \mathcal{H}^{d-1}(\Gamma_{KL})\,.\end{aligned}$$

Using ζ-regularity and the volume bounds (A.8), we combine these estimates to obtain

$$\begin{aligned}\sum_{K,L\in \mathsf{B}_r} \Big(f(K)-f(L)\Big)^2 m(K)m(L) &= [\mathcal{T}]^{-2d} \int_{\tilde{B}_r}\int_{\tilde{B}_r} \Big(\psi(x)-\psi(y)\Big)^2 \,\mathrm{d}\mathbf{m}(x)\,\mathrm{d}\mathbf{m}(y)\\ &\lesssim r^{d+2}[\mathcal{T}]^{2-d}\mathcal{H}^{d-1}(\Gamma_{KL}) \sum_{K,L\in \mathsf{B}_{cr}} \frac{\Big(f(K)-f(L)\Big)^2}{|x_K-x_L|}\\ &\lesssim r^2 m(\mathsf{B}_r) \sum_{K,L\in\mathsf{B}_{cr}} \Big(f(K)-f(L)\Big)^2 \mu_{KL}\,,\end{aligned}$$

as desired. □

The regularity of the discrete flows can now be shown using the volume doubling property (Proposition A.2.1) and the weak Poincaré inequality (Proposition A.2.2).

Proof of Proposition 4.4.4. First we note that (Γ,μ,m) satisfies a uniform ellipticity property, i.e., $\mu_{KL}\gtrsim m(K)$ for every $K\sim L$. Combined with Proposition A.2.1 (volume doubling) and Proposition A.2.2 (Poincaré inequality), this allows us to apply Theorem A.1.4 and prove a Harnack inequality (with unviersal constant $C_H\lesssim 1$) for the solutions to the equation $\partial_t u = \mathcal{L}u$, where $\mathcal{L}$ is the generator introduced in (A.6). Finally, the parabolic Harnack inequality implies Hölder continuity of the associated evolution equation [Del99, Proposition 4.1], which concludes the proof of (4.35).

The ultracontractivity estimates also follows from Proposition A.2.1 and Proposition A.2.2. More precisely, it is well-known (see e.g. [SC02]) that volume doubling and the Poincaré inequality implies the Nash inequality

$$\|f\|_2^{2+\frac{4}{d}} \lesssim \mathcal{E}(f,f)\|f\|_1^{\frac{4}{d}}\,,$$

where $\mathcal{E}$ is the Dirichlet form associated to $\mathcal{L}$. The ultracontractivity property for parabolic evolutions u_t associated to $\mathcal{L}$ then easily follows from the Nash inequality (see e.g. [CKS87, Theorem 2.1] for a proof), that is

$$\|u_t\|_{L^\infty(m)} \lesssim \Big(1\vee t^{-\frac{d}{2}}\Big)\|u_0\|_{L^1(m)}\,.$$

Performing the time rescaling $r_t := u_{t[\mathcal{T}]^{-2}}$, we obtain the ultracontractivity estimate (4.36) for r_t and conclude the proof. □

APPENDIX B

Some additional material

In this last part of the work, we include some results concerning metrics on the space of measures and property of convex functions, which find applications in particular in 3.

B.1 The Kantorovich–Rubinstein metric on signed measures

We collect some facts on the Kantorovich–Rubinstein metric that are used in this thesis, in particular Chpater 3. We refer to [Bog07, Section 8.10(viii)] for more details.

Let (X, d) be a metric space. Let $\mathcal{M}(X)$ denote the space of finite signed Borel measures on X. For $\mu \in \mathcal{M}(X)$, let $\mu^+, \mu^- \in \mathcal{M}_+(X)$ be the positive and negative parts, respectively. Let $|\mu| = \mu^+ + \mu^-$ be its variation, and $\|\mu\|_{\mathrm{TV}} := |\mu|(X)$ be its total variation.

Definition B.1.1 (Weak and vague convergence)**.** Let $\mu, \mu_n \in \mathcal{M}(X)$ for $n = 1, 2, \ldots$.

(i) We say that $\mu_n \to \mu$ weakly in $\mathcal{M}(X)$ if $\int_X \psi \,\mathrm{d}\mu_n \to \int_X \psi \,\mathrm{d}\mu$ for every $\psi \in \mathcal{C}_b(X)$.

(ii) We say that $\mu_n \to \mu$ vaguely in $\mathcal{M}(X)$ if $\int_X \psi \,\mathrm{d}\mu_n \to \int_X \psi \,\mathrm{d}\mu$ for every $\psi \in \mathcal{C}_c(X)$.

If (X, d) is compact, $\mathcal{M}(X)$ is a Banach space endowed with the norm $\|\mu\|_{\mathrm{TV}}$. By the Riesz–Markov theorem, it is the dual space of the Banach space $\mathcal{C}(X)$ of all continuous functions $\psi : X \to \mathbb{R}$ endowed with the supremum norm $\|\psi\|_\infty = \sup_{x\in X} |\psi(x)|$.

For $\psi : X \to \mathbb{R}$ let $\mathrm{Lip}(\psi) := \sup_{x\neq y} \frac{|\psi(x)-\psi(y)|}{d(x,y)}$ be its Lipschitz constant.

Definition B.1.2. Let (X, d) be a compact metric space. The *Kantorovich–Rubinstein norm* on $\mathcal{M}(X)$ is defined by

$$\|\mu\|_{\mathrm{KR}(X)} := \sup\left\{ \int_X \psi \,\mathrm{d}\mu \;:\; \psi \in \mathcal{C}(X),\ \|\psi\|_\infty \le 1,\ \mathrm{Lip}(\psi) \le 1 \right\}. \tag{B.1}$$

In non-trivial situations (i.e., when X contains an infinite convergent sequence), the norms $\|\cdot\|_{\mathrm{KR}}$ and $\|\cdot\|_{\mathrm{TV}}$ are not equivalent. Thus, by the open mapping theorem, $(\mathcal{M}(X), \|\cdot\|_{\mathrm{KR}})$ is not a complete space.

A closely related norm on $\mathcal{M}(X)$ that is often considered is

$$\|\mu\|_{\widetilde{\mathrm{KR}}(X)} := |\mu(X)| + \sup\left\{ \int_X \psi \, \mathrm{d}\mu \ : \ \psi \in \mathcal{C}(X), \ \psi(x_0) = 0, \ \mathrm{Lip}(\psi) \leq 1 \right\},$$

for some fixed $x_0 \in X$; see [Bog07, Section 8.10(viii)]. The next result shows that these norms are equivalent.

Proposition B.1.3. *Let (X, d) be a compact metric space. For $\mu \in \mathcal{M}(X)$ we have*

$$\|\mu\|_{\mathrm{KR}(X)} \leq \|\mu\|_{\widetilde{\mathrm{KR}}(X)} \leq c_X \|\mu\|_{\mathrm{KR}(X)},$$

where $c_X < \infty$ depends only on $\mathrm{diam}(X)$.

Proof. We start with the first inequality. Let $\psi \in \mathcal{C}(X)$ with $\|\psi\|_\infty \leq 1$ and $\mathrm{Lip}(\psi) \leq 1$. Define $\phi := \psi - \psi(x_0)$, so that $\phi(x_0) = 0$ and $\mathrm{Lip}(\phi) = \mathrm{Lip}(\psi) \leq 1$. Then

$$\int \psi \, \mathrm{d}\mu = \int \psi(x_0) + \phi \, \mathrm{d}\mu = \psi(x_0)\mu(X) + \int \phi \, \mathrm{d}\mu \leq |\mu(X)| + \int \phi \, \mathrm{d}\mu \leq \|\mu\|_{\widetilde{\mathrm{KR}}}.$$

Taking the supremum over ψ yields the desired bound.

Let us now prove the second inequality. Set $\Delta := 1 \vee \mathrm{diam}(X)$. Take $\psi \in \mathcal{C}(X)$ with $\psi(x_0) = 0$ and $\mathrm{Lip}(\psi) \leq 1$. Then $|\psi(x)| = |\psi(x) - \psi(x_0)| \leq d(x, x_0) \leq \mathrm{diam}(X) \leq \Delta$ for all $x \in X$, so that $\|\frac{\psi}{\Delta}\|_\infty \leq 1$ and $\mathrm{Lip}(\frac{\psi}{\Delta}) \leq 1$. We obtain

$$\int \psi \, \mathrm{d}\mu = \Delta \int \frac{\psi}{\Delta} \, \mathrm{d}\mu \leq \Delta \|\mu\|_{\mathrm{KR}}.$$

Moreover, $|\mu(X)| \leq \|\mu\|_{\mathrm{KR}}$ as can be seen by taking $\psi = \pm 1$ in (B.1) It follows that

$$\|\mu\|_{\widetilde{\mathrm{KR}}} \leq (1 + \Delta)\|\mu\|_{\mathrm{KR}},$$

as desired. □

Proposition B.1.4 (Relation to $\mathbb{W}_1$). *Let (X, d) be a compact metric space. If $\mu_1, \mu_2 \in \mathcal{M}_+(X)$ are nonnegative measures of equal total mass, we have $\|\mu_1 - \mu_2\|_{\widetilde{\mathrm{KR}}} = \mathbb{W}_1(\mu_1, \mu_2)$.*

Proof. This follows from the Kantorovich duality for the distance $\mathbb{W}_1$. □

On the subset of *nonnegative* measures, the KR-norm induces the weak* topology:

Proposition B.1.5 (Relation to weak*-convergence). *Let (X, d) be a compact metric space. For $\mu_n, \mu \in \mathcal{M}_+(X)$ we have*

$$\mu_n \to \mu \text{ weakly} \quad \text{if and only if} \quad \|\mu_n - \mu\|_{\mathrm{KR}} \to 0.$$

Proof. See [Bog07, Theorem 8.3.2]. □

Remark B.1.6 (Testing against smooth functions). If $X = \mathbb{T}^d$, the space of $\mathcal{C}^1$ functions ψ with $\mathrm{Lip}(\psi) \leq 1$ is dense in the set of Lipschitz functions with $\mathrm{Lip}(\psi) \leq 1$; see, e.g., [SW19, Proposition A.5]. Consequently,

$$\|\mu\|_{\mathrm{KR}(X)} = \sup\left\{ \int_X \psi \, \mathrm{d}\mu \ : \ \psi \in \mathcal{C}^1(\mathbb{T}^d), \ \|\psi\|_\infty \leq 1, \ \|\nabla\psi\|_\infty \leq 1 \right\}. \tag{B.2}$$

Remark B.1.7. The identity (B.2) shows that $\|\cdot\|_{\mathrm{KR}}$ is the dual norm of the separable Banach space $\mathcal{C}^1(Q)$. The dual space of $\mathcal{C}^1(Q)$ is a strict superset of the finite Borel measures.

B.2 Norms on curves in the space of measures

We work with curves of bounded variation taking values in the space $\mathcal{M}_+(\mathbb{T}^d)$.

Definition B.2.1 (Curves of bounded variation)**.** The space $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$ consists of all curves of measures $\boldsymbol{\mu}:\mathcal{I}\to\mathcal{M}_+(\mathbb{T}^d)$ such that the BV-seminorm

$$\|\boldsymbol{\mu}\|_{\mathrm{BV}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))} := \sup\left\{\int_{\mathcal{I}}\int_{\mathbb{T}^d}\partial_t\phi_t\,\mathrm{d}\mu_t\,\mathrm{d}t \;:\; \phi\in\mathcal{C}^1_c\Big(\mathcal{I};\mathcal{C}^1(\mathbb{T}^d)\Big),\ \max_{t\in\mathcal{I}}\|\phi\|_{\mathcal{C}^1(\mathbb{T}^d)}\le 1\right\} \tag{B.3}$$

is finite.

Remark B.2.2. The space $\mathrm{BV}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$ is a (non-closed) subset of the space $\mathrm{BV}(\mathcal{I};X^*)$, where X is the separable Banach space $\mathcal{C}^1(\mathbb{T}^d)$. We refer to [HPR19, Section 2] for the equivalence of several definitions of $\mathrm{BV}\big(\mathcal{I};X^*\big)$.

Definition B.2.3. The space $W^{1,1}_{\mathrm{KR}}(\mathcal{I};\mathcal{M}_+(\mathbb{T}^d))$ consists of all curves $(\mu_t)_{t\in\mathcal{I}}$ in the Banach space-valued Sobolev space $W^{1,1}\Big(\mathcal{I};(\mathcal{C}^1(\mathbb{T}^d))^*\Big)$ such that $\mu_t\in\mathcal{M}_+(\mathbb{T}^d)$ for a.e. $t\in\mathcal{I}$.

B.3 Domain properties of convex functions

Lemma B.3.1 (Domain properties of convex functions)**.** *Let $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ be convex, and let $x^\circ\in \mathsf{D}(f)^\circ$. For every $\lambda\in(0,1)$ and every bounded set $K\subseteq\mathsf{D}(f)$, there exists a compact convex set $K_\lambda\subseteq\mathsf{D}(f)^\circ$ such that*

$$(1-\lambda)K+\lambda x^\circ\subseteq K_\lambda.$$

Proof. Let $K\subseteq\mathsf{D}(f)$ be bounded and $\lambda\in(0,1)$. Since $x^\circ\in\mathsf{D}(f)^\circ$, we can pick $r>0$ such that $B(x^\circ,r)\subseteq\mathsf{D}(f)^\circ$. Fix $y\in\bar{K}$ and set $y_\lambda:=(1-\lambda)y+\lambda x^\circ$. We claim that $B(y_\lambda,\lambda r)\subseteq\mathsf{D}(f)^\circ$.

To prove the claim, it suffices to show that $B(y_\lambda,\lambda r)\subseteq\mathsf{D}(f)$, since $B(y_\lambda,\lambda r)$ is open. Take $z\in B(y_\lambda,\lambda r)$ and pick a sequence $(y_n)_n\subset K$ such that $y_n\to y$. Observe that $z=(1-\lambda)y_n+\lambda\tilde{x}_n$ with $\tilde{x}_n\in B(x^\circ,r)$ if n is large enough (indeed, $\tilde{x}_n-x^\circ=\frac{1}{\lambda}(z-y_\lambda)+\frac{1-\lambda}{\lambda}(y-y_n)$ and $|z-y_\lambda|<\lambda r$). Since $y_n,\tilde{x}_n\in\mathsf{D}(f)$, the claim follows by convexity of f.

We now define

$$C_\lambda:=\bigcup_{y\in K}B\Big(y_\lambda,\frac{\lambda r}{3}\Big)\quad\text{and}\quad K_\lambda:=\mathrm{Conv}(\overline{C_\lambda}).$$

By construction, K_λ is convex, bounded, and closed, thus compact. Let us show that $K_\lambda\subseteq\mathsf{D}(f)^\circ$.

By convexity of f, it suffices to show that $\overline{C_\lambda}\subseteq\mathsf{D}(f)^\circ$. Pick $z\in\overline{C_\lambda}$ and $\{z_n\}_n\subseteq C_\lambda$ such that $z_n\to z$. Then there exists $y_n\in K$ such that $z_n\in B\Big((y_n)_\lambda,\frac{\lambda r}{3}\Big)$. Passing to a subsequence, we may assume that $y_n\to\bar{y}$ for some $\bar{y}\in\bar{K}$ and $z_n\in B\Big(\bar{y}_\lambda,\frac{\lambda r}{2}\Big)$ for $n\ge\bar{n}\in\mathbb{N}$. Taking the limit as $n\to+\infty$ we infer that $z\in\overline{B\Big(\bar{y}_\lambda,\frac{\lambda r}{2}\Big)}$. Since $B\Big(\bar{y}_\lambda,\lambda r\Big)\subseteq\mathsf{D}(f)^\circ$, it follows that $z\in\mathsf{D}(f)^\circ$. □