

## MULTILEVEL MONTE CARLO METHODS FOR THE DEAN–KAWASAKI EQUATION FROM FLUCTUATING HYDRODYNAMICS\*

FEDERICO CORNALBA<sup>†</sup> AND JULIAN FISCHER<sup>‡</sup>

**Abstract.** Stochastic PDEs of fluctuating hydrodynamics are a powerful tool for the description of fluctuations in many-particle systems. In this paper, we develop and analyze a multilevel Monte Carlo (MLMC) scheme for the Dean–Kawasaki equation, a pivotal representative of this class of SPDEs. We prove analytically and demonstrate numerically that our MLMC scheme provides a significant reduction in computational cost (with respect to a standard Monte Carlo method) in the simulation of the Dean–Kawasaki equation. Specifically, we link this reduction in cost to having a sufficiently large average particle density and show that sizeable cost reductions can be obtained even when we have solutions with regions of low density. Numerical simulations are provided in the two-dimensional case, confirming our theoretical predictions. Our results are formulated entirely in terms of the law of distributions rather than in terms of strong spatial norms: this crucially allows for MLMC speed-ups altogether despite the Dean–Kawasaki equation being highly singular.

**Key words.** multilevel Monte Carlo methods, Dean–Kawasaki equation, fluctuating hydrodynamics, many-particle systems, noise coupling

**MSC codes.** 65C05, 60H15, 35R60, 65N06, 82M36, 82C22

**DOI.** 10.1137/23M1617345

**1. Introduction.** In the regime of large particle numbers, the behavior of interacting particle systems often gives rise to a classical PDE description, for instance, in form of the classical equations of continuum mechanics. For medium-sized particle systems consisting of only, e.g.,  $10^5$ – $10^9$  particles, this idealized description often becomes insufficient, as thermal fluctuations may begin to impact the behavior. The theory of fluctuating hydrodynamics augments the classical PDEs of continuum mechanics with suitable noise terms to account for thermal fluctuations, giving rise to SPDE models; we refer the reader to Landau and Lifshitz [22], Spohn [25], and te Vrugt, Löwen, and Wittkowski [26].

In this work, we are concerned with the numerical approximation of one of the most basic equations of fluctuating hydrodynamics, the *Dean–Kawasaki equation*

$$(1.1) \quad \partial_t \rho = \frac{1}{2} \Delta \rho + \nabla \cdot (\rho (\nabla V * \rho)) + N^{-1/2} \nabla \cdot (\sqrt{\rho} \xi).$$

It describes the effective behavior of the density  $\rho$  of a system of  $N$  weakly interacting diffusing particles  $\{X_i(t)\}_{i=1}^N$  in the regime of large particle numbers  $N \gg 1$ . Here,  $V$  is a (sufficiently regular) interaction potential and  $\xi$  denotes vector-valued space-time white noise. While the McKean–Vlasov equation  $\partial_t \bar{\rho} = \frac{1}{2} \Delta \bar{\rho} + \nabla \cdot (\bar{\rho} (\nabla V * \bar{\rho}))$ ,

\*Received by the editors November 15, 2023; accepted for publication (in revised form) October 25, 2024; published electronically January 31, 2025.

<https://doi.org/10.1137/23M1617345>

**Funding:** The work of the authors was supported by the Austrian Science Fund (FWF) project F65.

<sup>†</sup>University of Bath, Bath, UK (fc402@bath.ac.uk, <https://sites.google.com/view/federicocornalba/home>).

<sup>‡</sup>Institute of Science and Technology Austria, Klosterneuburg, Austria (julian.fischer@ista.ac.at, <http://j-fischer.eu/>).

obtained formally in the limit  $N \rightarrow \infty$ , describes the mean-field limit profile  $\bar{\rho}$  of the particle system, the Dean–Kawasaki equation (1.1) for  $N \gg 1$  in addition captures the law of the density fluctuations due to the finite number of particles.

The Dean–Kawasaki equation (1.1) is a highly singular SPDE, namely, it is even too singular to be renormalized by approaches like regularity structures or paracontrolled calculus [17, 16]. As shown in [20, 21], the only martingale solutions to (1.1) are in fact given by the empirical measures of the underlying particle system, that is,  $\rho(x, t) \equiv \mu_N^t(x) := N^{-1} \sum_{i=1}^N \delta(x - X_i(t))$ . Nevertheless, a justification of SPDEs related to (1.1) has recently been given in [12, 10] in terms of large-deviation principles. Justifications for Dean–Kawasaki type models with regularized noise have been developed in [10, 11].

Recently, the authors and collaborators [7, 8] have shown that the Dean–Kawasaki equation may be viewed as a recipe for accurate and efficient simulations of the density fluctuations in the interacting particle system: When applying a formal spatial semidiscretization to (1.1), the law of density fluctuations predicted by the discretization accurately describes the law of density fluctuations in the underlying particle system, as long as the grid size  $h$  is such that  $Nh^d \gg 1$  ( $d$  is the spatial dimension)—that is, as long as on average there is substantially more than one particle per grid cell.

Nevertheless, computing statistical properties of particle density fluctuations  $\bar{\rho}$  such as variances  $\mathbb{E}[|N^{1/2} \int (\mu_N^T - \bar{\rho}^T)(x) \varphi(x) dx|^2]$  (or, more generally, of random variables of the form

$$(1.2) \quad Q = \psi \left( N^{1/2} \int (\mu_N^T - \bar{\rho}^T)(x) \varphi(x) dx \right)$$

for sufficiently regular test functions  $\psi, \varphi$ ) via discretizations of (1.1) is still a computationally demanding task, as in addition to solving an SPDE one needs to sample over many realizations of the noise. Multilevel Monte Carlo (MLMC) methods are a powerful numerical tool that often allows one to offset a large amount of the sampling cost by performing the computation for most samples only on a much coarser numerical grid. Some of the first applications of MLMC methods have been in the context of parametric integration [18] and stochastic differential equations in mathematical finance [13];<sup>1</sup> MLMC methods have since found widespread applications for random and stochastic PDEs (see, for instance, [1, 2, 4, 6, 9, 14, 15, 19]).

In this work, we develop and analyze MLMC methods to approximate  $\mathbb{E}[Q]$  for  $Q$  of type (1.2), using discretizations of the Dean–Kawasaki equation (1.1). To the best of our knowledge, our present work is one of the first mathematical results on MLMC methods both for singular stochastic PDEs and for stochastic PDEs involving multiplicative space-time white noise.

**1.1. Summary of main results.** In order to approximate  $\mathbb{E}[Q]$  (see (1.2)), we set up a sequence of random variables  $(P_\ell)_\ell$  associated with space-time discretizations of (1.1) over levels  $\ell = 0, \dots, L$  (the higher the  $\ell$ , the smaller the spatial grid size  $h_\ell$  and time-step<sup>2</sup>  $\tau_\ell$ , and the more accurately  $P_\ell$  approximates  $Q$ ). A key tool of our analysis is to couple the noises of discretized versions of (1.1) on consecutive levels  $\ell - 1, \ell$ . We provide two different ways of doing this. The first way is to couple the

<sup>1</sup>Earlier applications of related multigrid algorithms can be found, for instance, in the context of statistical physics [3].

<sup>2</sup>Since we will link  $h_\ell$  and  $\tau_\ell$  via a standard Courant–Friedrichs–Lewy condition, many of our considerations will be stated—for notational convenience—just in terms of the spatial discretization parameters  $(h_\ell)_\ell$ .

noises over the Fourier frequencies, thus having the noise on level  $\ell - 1$  share all of its frequencies with the noise on level  $\ell$  (“Fourier coupling”; cf. section 3.1). The second way is to construct the noise on level  $\ell - 1$  using only local spatial information coming from the noise on level  $\ell$  (“rightmost nearest neighbors (NN) coupling”; cf. section 3.2).

We can now informally state our main result.

**THEOREM 1.1.** *Let  $\varepsilon > 0$  be a given accuracy. Consider a sequence of levels for which the finest grid size  $h_{\min}$  satisfies  $h_{\min}^2 \lesssim \varepsilon$ . Under the technical Assumptions 1, 2, 3, 4, and 5, which in particular require the average particle density  $Nh_{\min}^d$  to satisfy*

$$(1.3) \quad Nh_{\min}^d \gtrsim h_{\min}^{-\beta}, \quad \text{where } \beta := \begin{cases} 4 & \text{for Fourier coupling; cf. section 3.1} \\ 2 & \text{for NN coupling; cf. section 3.2,} \end{cases}$$

*we can set up a MLMC estimator  $\mu_{MLMC}$  for the simulation of (1.1) which achieves a mean-square error  $\mathbb{E}[|\mu_{MLMC} - \mathbb{E}[Q]|^2]$  of order  $\varepsilon^2$ , while carrying a computational cost  $C_{MLMC}$  bounded by*

$$(1.4) \quad \mathbb{E}[C_{MLMC}] \lesssim \begin{cases} \varepsilon^{-2} & \text{for } d=1 \text{ and Fourier coupling; cf. section 3.1,} \\ \varepsilon^{-2} \cdot (\log \varepsilon)^2 & \text{for } d=2 \text{ and Fourier coupling; cf. section 3.1,} \\ \varepsilon^{-2} \cdot \varepsilon^{-(d-2)/2} & \text{for } d \geq 3 \text{ and Fourier coupling; cf. section 3.1,} \\ \varepsilon^{-2} \cdot \varepsilon^{-d/2} & \text{for } d \geq 1 \text{ and NN coupling; cf. section 3.2,} \end{cases}$$

*as opposed to the standard MC method, whose cost bound is*

$$\mathbb{E}[C_{MC}] \lesssim \varepsilon^{-2} \cdot \varepsilon^{-d/2-1}.$$

As a key challenge for our analysis, unlike in MLMC approaches for parabolic SPDEs in the literature [1, 15], the highly singular nature of (1.1) prevents strong (pathwise) convergence of numerical solutions in the limit of arbitrarily fine discretizations. Instead, convergence takes place only up to a minimal  $h \gg N^{-1/d}$ , below which the sequence begins to diverge in  $C^0$ . For an illustration of this lack of regularity and convergence of solutions, we refer to Figures 2 and 3. Moreover, while microscopic fluctuations in (1.1) may be drastic, the macroscopic impact of fluctuations in (1.1)—i.e., the impact of fluctuations on weighted spatial averages like  $\int \rho^t(x) \varphi(x) dx$ —is rather small due to the  $N^{-1/2}$  prefactor of the noise. This discrepancy is directly related to the singular nature of the SPDE (1.1). In particular, this forces us to work with stochastically weak convergence estimates—i.e., convergence estimates for the law of distributions—to achieve the MLMC cost bound (1.4).

As with any MLMC method, we need to provide suitable bounds for the systematic error  $|\mathbb{E}[P_\ell - Q]|$  and the cross-level variance  $\text{Var}[P_\ell - P_{\ell-1}]$  in order to have a cost reduction. Specifically, we have the systematic error bound

$$(1.5) \quad |\mathbb{E}[P_\ell - Q]| \lesssim (Nh_\ell^d)^{-1} + h_\ell^2$$

(which is a relatively straightforward consequence of results in [7]) and a nontrivial, similar looking bound for  $\text{Var}[P_\ell - P_{\ell-1}]$  (cf. Proposition 5.1), which reads

$$(1.6) \quad \text{Var}[P_\ell - P_{\ell-1}] \lesssim (Nh_\ell^d)^{-1} + h_\ell^\beta,$$

where  $\beta$  is defined in (1.3). Apart from some polynomial numerical error in  $h_\ell$ , both of the bounds (1.5)–(1.6) feature the inverse of the average particle density  $Nh_\ell^d$ : crucially, these bounds are geometrically decaying in the level  $\ell$ —and therefore conform to the assumptions of the general MLMC complexity theorem [14, Theorem 2.1]—as long as the average particle density is sufficiently large on all levels required to achieve the systematic error  $O(\varepsilon)$ . The presence of the term  $(Nh_\ell^d)^{-1}$  in (1.5)–(1.6) is rooted in the singular nature of the Dean–Kawasaki equation, as discussed above. However, and most importantly, having a large density requirement is not a limitation but is in fact absolutely natural: it corresponds to a regime in which our SPDE models are computationally more efficient to simulate than the underlying particle systems.

Finally, in addition to the general complexity bound (1.4), we also show in Proposition 5.4 that our MLMC method achieves a variance reduction factor  $F$  (with respect to the standard MC method) given by

$$(1.7) \quad F \propto \begin{cases} \left[ (Nh_{\min}^d)^{-1} + h_{\min}^{(d+2) \wedge 4} L^* \right]^{-1} |\log(h_{\min})|^{-1} & \text{for Fourier coupling; cf. section 3.1} \\ \left[ (Nh_{\min}^d)^{-1} + h_{\min}^2 \right]^{-1} |\log(h_{\min})|^{-1} & \text{for NN coupling; cf. section 3.2,} \end{cases}$$

where we have set  $L^* = 1$  if  $d \neq 2$ , and  $L^* = |\log(h_{\min})|$  if  $d = 2$ .

*Remark 1.2.* While the bound (1.7) is somewhat tied to the results of Theorem 1.1, we choose to present it on its own: the reason for this is that (1.7) shows an explicit and interpretable dependence on the average particle density  $Nh_{\min}^d$ .

*Remark 1.3.* For simplicity, we focus on the noninteracting case  $V \equiv 0$  in (1.1) (in this case,  $\bar{\rho}^t \equiv \mathbb{E}[\mu_N^t]$ ), although we expect that with the ideas of [8] one could generalize the result to interacting particles. Moreover, we limit ourselves to the analysis of a finite difference scheme; due to the lack of regularity of solutions (see Figure 3 for a plot of a sample path), a finite element scheme would not offer compelling advantages.

**1.2. Structure of the paper.** Notation is introduced in section 2 (discrete Dean–Kawasaki model) and section 3 (MLMC setup). Assumptions, statements of main results, and proofs of all main results are given in sections 4, 5, and 6, respectively. Numerical results are discussed in section 7. Final considerations are given in section 8.

**2. Basic notation.** We work with uniformly spaced grids on the spatial domain  $\mathbb{T}^d$  of the type  $G_{h,d} := \{-\pi; -\pi + h; \dots; \pi - h\}^d$  for any compatible mesh-size  $h > 0$ . We denote by  $(f_h, g_h)_h := h^d \sum_{x \in G_{h,d}} f_h(x) \cdot g_h(x)$  (respectively,  $\|\cdot\|_h$ ) the inner product (respectively, the norm) on  $L^2(G_{h,d})$ . We denote by  $\mathcal{I}_h: C(\mathbb{T}^d) \rightarrow L^2(G_{h,d})$  the pointwise interpolation operator. Furthermore,  $\Delta_h, \nabla_h$  denote discrete counterparts of Laplace and gradient operators (precise details are given later on). As for time, we work on a fixed interval  $[0, T]$ , and consider uniform partitions of the type  $\mathcal{S}_\tau = \{0, \tau, \dots, T\}$ , for a compatible time-step  $\tau > 0$ .

Whenever we want to provide a lower (respectively, upper) bound which holds up to a constant depending on a specific parameter  $a$ , we write  $\gtrsim_a$  (respectively,  $\lesssim_a$ ).

**DEFINITION 2.1** (space-time discretized Dean–Kawasaki model approximating an underlying system of  $N$  noninteracting Brownian particles in  $\mathbb{T}^d$ ). *We denote by  $\rho_{h,\tau}$  the solution to the following finite difference Dean–Kawasaki equation,*

$$(2.1) \quad \rho_{h,\tau}^{m\tau} = \rho_{h,\tau}^{(m-1)\tau} + \tau \sum_{j=0}^1 b_j \frac{1}{2} \Delta_h \rho_{h,\tau}^{(m-j)\tau} + N^{-1/2} \nabla_h \cdot \left( \sum_{\substack{(y,r) \in \\ (G_{h,d}, \{1, \dots, d\})}} \sqrt{[\rho_{h,\tau}^{(m-1)\tau}]^+} e_{h,y,r}^d \Delta \beta_{(y,r)}^{m-1} \right),$$

started from  $\rho_{h,\tau}^0 = \rho_{h,0}$ , where  $[y]^+ := \max\{y; 0\}$ , and where

- $\Delta \beta_{(y,r)}^{m-1} := \beta_{(y,r)}(m\tau) - \beta_{(y,r)}((m-1)\tau)$  are increments for a family of independent Brownian motions  $\{\beta_{(y,r)}\}_{(y,r) \in (G_{h,d}, \{1, \dots, d\})}$ ;
- $\rho_{h,0}$  is a suitable—possibly random—initial profile;
- $\{e_{h,y,r}^d\}_{(y,r) \in (G_{h,d}, \{1, \dots, d\})}$  is the orthonormal basis of  $[L^2(G_{h,d})]^d$  given by  $e_{h,y,r}^d(z) := h^{-d/2} \mathbf{1}_{y=z} f_r$ , where  $\{f_r\}_{r=1}^d$  is the canonical basis of  $\mathbb{R}^d$ ;
- $b_0, b_1$  are fixed, nonnegative weights such that  $b_0 + b_1 = 1$ . These weights determine the specific nature of the (one-step) time discretization scheme.

**3. MLMC framework.** The goal is to compute statistical properties of the particle density fluctuations. Specifically, we aim to compute the expected value of random variables of the form (1.2)

$$Q = \psi \left( N^{1/2} \int (\mu_N^T - \bar{\rho}^T)(x) \varphi(x) dx \right)$$

for sufficiently regular test functions  $\psi, \varphi$ . In order to approximate  $\mathbb{E}[Q]$ , we can compute the expected value of the quantity

$$(3.1) \quad P := \psi \left( N^{1/2} (\rho_{h,\tau}^T - \mathbb{E}[\rho_{h,\tau}^T], \mathcal{I}_h \varphi)_h \right),$$

where  $\rho_{h,\tau}$  is the solution to the discrete Dean–Kawasaki model (2.1) with space (respectively, time) discretization parameter  $h$  (respectively,  $\tau$ ). We think of the parameters  $(h, \tau)$  in the random variable  $P$  introduced in (3.1) as being the smallest in a sequence of space/time parameters (so, effectively,  $h = h_{\min}, \tau = \tau_{\min}$ ): therefore, we construct a MLMC scheme for approximating  $\mathbb{E}[P]$  by using the auxiliary quantities

$$(3.2) \quad P_\ell := \psi \left( N^{1/2} (\rho_{h_\ell, \tau_\ell}^T - \bar{\rho}_{h_\ell, \tau_\ell}^T, \mathcal{I}_{h_\ell} \varphi)_{h_\ell} \right),$$

on levels  $\ell = 0, \dots, L$ , and for some  $\{(h_\ell, \tau_\ell)\}_{\ell=0}^L$ . Since the deterministic dynamics of (2.1) is just a finite difference approximation of the Laplacian operator, it is natural to define the sequences  $\{h_\ell\}_{\ell=0}^L$  and  $\{\tau_\ell\}_{\ell=0}^L$  in terms of a Courant–Friedrichs–Lewy (CFL) ratio. Namely, we choose  $\{h_\ell\}_{\ell=0}^L$  and  $\{\tau_\ell\}_{\ell=0}^L$  to be geometric progressions with fixed CFL ratio across levels. That is, we assume

$$(3.3) \quad \tau_\ell / h_\ell^2 = \mu \quad \forall \ell \in \{0, \dots, L\}$$

for some fixed  $\mu > 0$ . The ratio of the sequence  $\{h_\ell\}_{\ell=0}^L$  (which, by (3.3), also uniquely determines the ratio of the sequence  $\{\tau_\ell\}_{\ell=0}^L$ ) will depend on the specific choice of coupling across levels, as discussed in subsections 3.1 and 3.2 below.

*Remark 3.1.* Upon imposing sensible assumptions, there would be no harm in also letting the ratios  $\tau_\ell / h_\ell^2$  in (3.3) depend on the level  $\ell$ . We do not do this, as this would entail further notational burden.

With this notation,  $P_L = P$ ,  $(h_{\min}, \tau_{\min}) = (h_L, \tau_L)$ , and  $(h_{\max}, \tau_{\max}) = (h_0, \tau_0)$ . When there is no ambiguity, we may choose to simplify the notation related to the level  $\ell$ : for instance, we abbreviate  $\psi(N^{1/2}(\rho_\ell^T - \bar{\rho}_\ell^T, \varphi_\ell)_\ell) \equiv \psi(N^{1/2}(\rho_{h_\ell, \tau_\ell}^T - \bar{\rho}_{h_\ell, \tau_\ell}^T, \mathcal{I}_{h_\ell} \varphi)_{h_\ell})$ .

We consider two alternative ways to couple the noise on consecutive levels  $\ell - 1, \ell$ .

**3.1. First coupling: Fourier coupling.** We split the spatial grid into three at each level increment (i.e.,  $h_{\ell-1} = 3h_\ell$ ). We write the stochastic noise of (2.1) using a Fourier expansion in space and standard coupling in time. In other words, at the coarse level  $\ell - 1$ ,

- all Fourier frequencies coming from the finer level  $\ell$  are reused, and
- noise increments are obtained by summing up noise increments at level  $\ell$ .

Explicitly, denoting the set of frequencies  $I_\ell := \{-\pi h_\ell^{-1}; -\pi h_\ell^{-1} + 1; \dots; \pi h_\ell^{-1} - 1\}^d$ , and setting  $A := \{-\pi h_\ell^{-1}; 0\}^d$ , the noise increments at levels  $\ell - 1$  and  $\ell$  are given by

$$\begin{aligned} & W_\ell(x, t + \tau_\ell) - W_\ell(x, t) \\ & := \sum_{(\xi, r) \in (I_\ell, \{1, \dots, d\})} e^{ix \cdot \xi} f_r(\beta_{\xi, r}(t + \tau_\ell) - \beta_{\xi, r}(t)), \quad \text{for } (x, t) \in G_{\ell, d} \times \mathcal{S}_\ell, \\ & W_{\ell-1}(x, t + \tau_{\ell-1}) - W_{\ell-1}(x, t) \\ & := \sum_{j=0}^{\kappa_t-1} \sum_{(\xi, r) \in (I_{\ell-1}, \{1, \dots, d\})} e^{ix \cdot \xi} f_r[\beta_{\xi, r}(t + (j+1)\tau_\ell) - \beta_{\xi, r}(t + j\tau_\ell)] \\ & \quad \text{for } (x, t) \in G_{\ell-1, d} \times \mathcal{S}_{\ell-1}, \end{aligned}$$

where  $\kappa_t = 3^2$  is the ratio of the sequence  $\{\tau_\ell\}_{\ell=0}^L$  (see (3.3)), and

$$\{(\beta_{\xi, r}, \beta_{-\xi, r})\}_{(\xi, r) \in (I_{\ell-1} \setminus A, \{1, \dots, d\})} \cup \{\beta_{\xi, r}\}_{(\xi, r) \in (A, \{1, \dots, d\})} =: \mathcal{B}_{I_{\ell-1} \setminus A} \cup \mathcal{B}_A$$

is an independent family with complex conjugate couples of Brownian motions for the set  $\mathcal{B}_{I_{\ell-1} \setminus A}$  and standard real-valued Brownian motions for the set  $\mathcal{B}_A$ .

*Remark 3.2.* Note that, for each  $\ell$ , the noise increments  $W_\ell$  have the same distribution as the noise increments of the noise prescribed in (2.1) for corresponding  $h_\ell, \tau_\ell$ , irrespective of whether  $W_\ell$  is associated with the “fine” or the “coarse” discretization in a given pair of discretizations. In particular, this ensures that the MLMC estimator for  $P$  is unbiased: specifically, we have that  $\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{\ell=1}^L \mathbb{E}[P_\ell - P_{\ell-1}]$ .

**3.2. Second coupling: Rightmost nearest neighbors coupling.** While the coupling in time is the same as in the previous case, coupling in space is obtained by summing up noise increments in a “rightmost” neighborhood of each spatial point. Specifically, splitting the spatial grid in two at each level increment (i.e.,  $h_{\ell-1} = 2h_\ell$ ), if we set  $B_{\ell-1}^{\rightarrow}(x) := \{x + h_\ell v : v \in \{0; 1\}^d\}$  for each  $x \in G_{\ell-1, d}$ , we define

$$\begin{aligned} & W_\ell(x, t + \tau_\ell) - W_\ell(x, t) \\ & := \sum_{(y, r) \in (G_{\ell-1, d}, \{1, \dots, d\})} e_{h_\ell, y, r}^d(x) (\beta_{y, r}(t + \tau_\ell) - \beta_{y, r}(t)) \quad \text{for } (x, t) \in G_{\ell, d} \times \mathcal{S}_\ell, \\ & W_{\ell-1}(x, t + \tau_{\ell-1}) - W_{\ell-1}(x, t) \\ & := \sum_{(y, r) \in (G_{\ell-1, d}, \{1, \dots, d\})} e_{h_{\ell-1}, y, r}^d(x) \sum_{j=0}^{\kappa_t-1} (\tilde{\beta}_{y, r}(t + (j+1)\tau_\ell) - \tilde{\beta}_{y, r}(t + j\tau_\ell)) \\ & \quad \text{for } (x, t) \in G_{\ell-1, d} \times \mathcal{S}_{\ell-1}, \end{aligned}$$

where  $\kappa_t = 2^2$  is the ratio of the sequence  $\{\tau_\ell\}_{\ell=0}^L$  (see (3.3)), and where  $\tilde{\beta}_{y,r} := 2^{-d/2} \sum_{z \in B_{\ell-1}^{\rightarrow}(y)} \beta_{z,r}$ , with  $\{\beta_{\ell,z,r}\}_{(z,r) \in (G_{\ell-1,d}, \{1, \dots, d\})}$  being standard independent Brownian motions.

The contents of Remark 3.2 apply verbatim also for this coupling.

*Remark 3.3.* As suggested by one of the referees, the coupling in subsection 3.2 could, alternatively, be set up in a symmetric way: this would entail keeping the coupling as is, but using the shifted grid points  $\{-\pi + h/2; -\pi + (3/2)h; \dots; \pi - h/2\}^d$ .

**4. Assumptions.** We work under the following assumptions.

*Assumption 1* (parameter scaling). Let  $L \in \mathbb{N}$  be the number of levels being considered, and let  $h_{\min}$  be the smallest associated grid size. We choose  $N$  to be sufficiently large so that the following scaling holds:

$$(4.1) \quad N h_\ell^d \gtrsim h_\ell^{-\beta} (1 \vee \mu^2) |\log h_\ell|, \quad \ell \in \{0, \dots, L\},$$

where  $\mu := \tau_\ell / h_\ell^2$  and where

$$\beta := \begin{cases} 4 & \text{for Fourier coupling; cf. section 3.1,} \\ 2 & \text{for NN coupling; cf. section 3.2.} \end{cases}$$

*Assumption 2* (continuous mean field limit). We assume the continuous mean-field limit  $\bar{\rho}^t$  to be strictly positive and bounded ( $0 < \rho_{\min} \leq \bar{\rho}^t(x) \leq \rho_{\max}$ ), and satisfying  $\bar{\rho}^t \in C^{R+3}$  for some  $R > 4 + d/2$  in the case of Fourier coupling (subsection 3.1) or  $\bar{\rho}^t \in C^2$  in the case of rightmost NN coupling (subsection 3.2).

*Assumption 3* (discrete Dean–Kawasaki dynamics). With regard to (2.1), (i) the operators  $\Delta_h$  and  $\nabla_h$  are second-order finite difference operators in the spatial discretization; (ii) the couples  $\{h_\ell, \tau_\ell\}_\ell$  are such that the noiseless version of the scheme (2.1) (i.e., the discrete mean field limit) is stable for all  $\ell$  and satisfies the discrete maximum principle.

*Assumption 4* (test functions). For  $\psi, \varphi$  introduced in (3.1), we require that

$$(4.2) \quad \sup_{z \in \mathbb{R}^K} |\nabla \psi(z)| / (1 + |z|^r) \leq C \quad \text{for some } r > 0$$

and that  $\varphi \in C^{R+4}$  for some  $R > 4 + d/2$  in the case of Fourier coupling (subsection 3.1), or  $\varphi \in C^3$  in the case of rightmost NN coupling (subsection 3.2). Furthermore, we use the notation  $\phi^t$  (respectively,  $\phi_\ell^t$ ) to indicate the solution to the continuous (respectively, discrete) backward heat equation ending at  $\phi^T = \varphi$  (respectively, at  $\phi_\ell^T = \mathcal{I}_\ell \varphi$ ).

*Assumption 5* (initial datum). We set the initial condition of the discrete mean-field limit  $\bar{\rho}_\ell^0 := \mathcal{I}_{h_\ell} \bar{\rho}^0$ . We require the following bound for large fluctuations,

$$(4.3) \quad \mathbb{E} [\|\rho_\ell^0 - \bar{\rho}_\ell^0\|_\infty^j] \lesssim j^j \left( N^{-1/2} h_\ell^{-d/2} \right)^j, \quad j \in \mathbb{N}$$

(see, for instance, [5, Theorems 3.4–3.5] and [23]), as well as the fluctuation bound

$$(4.4) \quad \mathbb{E} \left[ \left| N^{1/2} (\rho_\ell^0 - \bar{\rho}_\ell^0, \phi_\ell^0)_\ell - N^{1/2} \int (\mu_N^0 - \bar{\rho}^0)(x) \phi^0(x) dx \right|^4 \right]^{1/4} \lesssim C(\varphi) h_\ell^2$$

(see also [8, Bound (3.10)]).

*Remark 4.1.* The existence of  $\rho_{\min}$  and  $\rho_{\max}$  in Assumption 2 is guaranteed by the maximum principle for the continuous heat equation, provided we start from a strictly positive initial datum  $\bar{\rho}_0$ . The same  $\rho_{\min}$  and  $\rho_{\max}$  then also bound the discrete mean field limit  $\rho_{h,\tau}$  thanks to Assumption 3.

*Remark 4.2.* Assumption 5 is related to a fairly standard concentration inequality; see also similar discussions in [7, 8].

*Remark 4.3.* If we considered the case  $V \neq 0$  (which is not treated in this paper), then we would need to require strict positivity of the continuous mean-field limit (Assumption 2) and guarantee strict positivity of the discrete mean-field limit (Assumption 3).

**5. Cross-level variance bound, and variance reduction result.** The main result of this paper (MLMC complexity, Theorem 1.1) and the variance reduction results (Proposition 5.4) rely heavily on the following proposition.

**PROPOSITION 5.1** (bounding  $\text{Var}\{P_\ell - P_{\ell-1}\}$ ). *Assume the validity of Assumptions 1, 3, 2, 4, 5, and set  $\mu := \tau_\ell/h_\ell^2$ . We have the estimate*

(5.1)

$$\begin{aligned} & \text{Var}[P_\ell - P_{\ell-1}] \\ & \lesssim_{\varphi} \min \left\{ \exp \left( -\frac{C\rho_{\min}}{\rho_{\max}^{1/2}} \left[ \frac{Nh_\ell^d}{\mu^2 \vee 1} \right]^{\frac{1}{2}} \right) \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right]^{\frac{1}{2}} + \rho_{\min}^{-1} \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right]; \right. \\ & \quad \left. \times \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right]^{\frac{1}{2}} \right\} + \text{Err}_{num} := \text{Err}_{mod} + \text{Err}_{num}, \end{aligned}$$

where

$$\text{Err}_{num} \lesssim \begin{cases} \min[\rho_{\min}^{-1}(h_\ell^4 + \tau_\ell^2); h_\ell^2 + \tau_\ell] + C(\rho_{\min})[h_\ell^4 + \tau_\ell^2] & \text{for Fourier coupling,} \\ \min\{\rho_{\min}^{-1}(h_\ell^2 + \tau_\ell); h_\ell + \tau_\ell\} & \text{for R-M coupling.} \end{cases}$$

*Remark 5.2.* The name  $\text{Err}_{mod}$  above simply reflects the fact that such an error is—morally—of the modelling type (as it is directly related to the average particle density). On the other hand,  $\text{Err}_{num}$  is a purely numerical error.

*Remark 5.3.* Bound (5.1) takes different forms, depending on the choice of coupling and on which of the arguments in the minima is smaller. As we do not want to unnecessarily overcomplicate the statement of our main Theorem 1.1 (which stems directly from Proposition 5.1), we limit ourselves to stating Proposition 5.4 in two relevant subset of cases, one for each coupling.

**PROPOSITION 5.4** (variance reduction). *Assume the validity of Assumptions 2, 3, 4, and 5. Let  $M_\ell$ ,  $\ell \in \{0, \dots, L\}$  be the number of samples of  $P_\ell - P_{\ell-1}$  (we understand  $P_{-1} \equiv 0$ ), and let  $M_0$  be the number of samples of  $P_0$ . Choose*

$$(5.2) \quad M_\ell := (h_\ell/h_{\min})^d \cdot (\tau_\ell/\tau_{\min}) = (h_\ell/h_{\min})^{d+2}$$

and  $L \propto |\log(h_{\min})|$ . *Assume  $\bar{\rho}$  to be regular enough so that the order in the  $h$ -bound for  $\text{Err}_{num}$  in (5.1) is the highest of the two available (i.e., 4 for Fourier coupling and 2 for rightmost NN coupling). Then the variance of the MLMC estimator*

$$(5.3) \quad \mu_{MLMC} := M_0^{-1} \sum_{i=1}^{M_0} P_{0,(i)} + \sum_{\ell=1}^L M_\ell^{-1} \sum_{i=1}^{M_\ell} (P_{\ell,(i)} - P_{\ell-1,(i)})$$

satisfies the bound

$$(5.4) \quad \text{Var}[\mu_{MLMC}] \lesssim \begin{cases} (Nh_{\min}^d)^{-1} + h_{\min}^{(d+2)\wedge 4} L^* & \text{for Fourier coupling; cf. section 3.1,} \\ (Nh_{\min}^d)^{-1} + h_{\min}^2 & \text{for NN coupling; cf. section 3.2,} \end{cases}$$

where we have set

$$(5.5) \quad L^* := \begin{cases} 1 & \text{if } d \neq 2, \\ |\log(h_{\min})| & \text{if } d = 2, \end{cases}$$

while carrying a total computational cost,

$$(5.6) \quad \text{Cost}_{\text{tot}}(\mu_{MLMC}) \propto h_{\min}^{-(d+2)} \cdot |\log(h_{\min})|.$$

As a result of (5.4)–(5.6), using the MLMC estimator  $\mu_{MLMC}$  defined in (5.3) to approximate  $\mathbb{E}[\psi(N^{1/2} \int (\mu_N^T - \bar{\rho}^T)(x) \varphi(x) dx)]$  (see (1.2))—as opposed to a standard MC estimator on the finest scale  $(h_{\min}, \tau_{\min})$ —grants a variance reduction factor

$$(5.7) \quad F \propto \begin{cases} \left[ (Nh_{\min}^d)^{-1} + h_{\min}^{(d+2)\wedge 4} L^* \right]^{-1} |\log(h_{\min})|^{-1} & \text{for Fourier coupling; cf. section 3.1,} \\ \left[ (Nh_{\min}^d)^{-1} + h_{\min}^2 \right]^{-1} |\log(h_{\min})|^{-1} & \text{for NN coupling; cf. section 3.2.} \end{cases}$$

*Remark 5.5.* Although the choice of samples in (5.2) amounts to a slightly less efficient MLMC estimator than what we would otherwise have if we followed standard cost optimization (see Algorithm 7.1 below), it, however, enables us to obtain the variance reduction factor (5.7) with a short and intuitive proof. The standard cost optimization procedure is instead followed when producing the computational results associated with Theorem 1.1 (see [14] and Algorithm 7.1 below). In contrast to Theorem 1.1, Proposition 5.4 is solely concerned with variance reduction, and only features  $h_{\min}$  and the inverse particle density  $(Nh_{\min}^d)^{-1}$ .

**6. Proofs of Proposition 5.1, Theorem 1.1, and Proposition 5.4.** The following lemma—whose proof is deferred to Appendix B—is needed for proving Proposition 5.1.

LEMMA 6.1. *Under Assumptions 1 and 5 we have*

$$(6.1) \quad \mathbb{P} \left[ \sup_{t \in \mathcal{S}_\tau} \|\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t\|_\infty \geq \frac{B\rho_{\min}}{4} \right] \lesssim \exp \left( -\frac{C\rho_{\min}}{\rho_{\max}^{1/2}} \left[ \frac{Nh^d}{\mu^2 \vee 1} \right]^{1/2} \frac{B}{\sqrt{B+1}} \right)$$

for any  $B \geq 0$ , and where  $\mu = \tau/h^2$ . Furthermore, we have the estimates

$$(6.2) \quad \mathbb{E} \left[ \sup_{t \in \mathcal{S}_\tau} \|\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t\|_\infty^j \right]^{1/j} \lesssim C(j) \rho_{\max}^{1/2} (N^{-1}h^{-d})^{1/2} (\mu^2 \vee 1)^{1/2},$$

$$(6.3) \quad \mathbb{E} \left[ \left| (\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t, \mathcal{I}_h \varphi)_\ell \right|^j \right]^{1/j} \lesssim C(j) \rho_{\max}^{1/2} N^{-1/2} \|\varphi\|_{C^1}.$$

**6.1. Proof of Proposition 5.1 with Fourier coupling.** For notational convenience, we occasionally drop the time dependence and the vectorial notation over the  $d$  components of the noise (i.e., we write  $e^{ix \cdot \xi}$  instead of  $\{e^{ix \cdot \xi} f_r\}_{r=1}^d$ ).

*Step 1: Rewriting  $\text{Var}[P_\ell - P_{\ell-1}]$ .* A first-order Taylor expansion on  $\psi$  gives

$$\begin{aligned} \text{Var}[P_\ell - P_{\ell-1}] &\leq \mathbb{E} \left[ |P_\ell - P_{\ell-1}|^2 \right] \\ &= \mathbb{E} \left[ \left| \psi \left( N^{1/2} (\rho_\ell^T - \bar{\rho}_\ell^T, \varphi_\ell)_\ell \right) - \psi \left( N^{1/2} (\rho_{\ell-1}^T - \bar{\rho}_{\ell-1}^T, \varphi_{\ell-1})_{\ell-1} \right) \right|^2 \right] \\ (6.4) \quad &\leq \mathbb{E} \left[ |\nabla \psi(z)|^4 \right]^{\frac{1}{2}} \mathbb{E} \left[ \left| N^{\frac{1}{2}} \left[ (\rho_\ell^T - \bar{\rho}_\ell^T, \varphi_\ell)_\ell - (\rho_{\ell-1}^T - \bar{\rho}_{\ell-1}^T, \varphi_{\ell-1})_{\ell-1} \right] \right|^4 \right]^{\frac{1}{2}} \end{aligned}$$

for some random  $z$  such that

$$(6.5) \quad |z| \leq \max \left\{ \left| N^{1/2} (\rho_\ell^T - \bar{\rho}_\ell^T, \varphi_\ell)_\ell \right|; \left| N^{1/2} (\rho_{\ell-1}^T - \bar{\rho}_{\ell-1}^T, \varphi_{\ell-1})_{\ell-1} \right| \right\}.$$

*Step 2: Bounding  $\mathbb{E}[|\nabla \psi(z)|^4]$  in (6.4).* Using (4.2), (6.5), and (6.3), we get

$$(6.6) \quad \mathbb{E} [|\nabla \psi(z)|^4] \leq C(\rho_{\max}, r) \|\varphi\|_{C^1}^{4r}.$$

*Step 3: Itô formula for second term in right-hand side of (6.4).* Let the test functions  $\phi_\ell$  (respectively,  $\phi_{\ell-1}$ ) satisfy the backward evolution (A.1) ending in  $\varphi_\ell$  (respectively, in  $\varphi_{\ell-1}$ ). Thanks to the discrete martingale property in Lemma A.5 and a simple interpolation argument on the noise, we can define a continuous-time martingale  $D_{\ell, \ell-1}^t$  such that, crucially,

$$D_{\ell, \ell-1}^t = N^{1/2} (\rho_\ell^t - \bar{\rho}_\ell^t, \phi_\ell^t)_\ell - N^{1/2} (\rho_{\ell-1}^t - \bar{\rho}_{\ell-1}^t, \phi_{\ell-1}^t)_{\ell-1} \quad \forall t \in \mathcal{S}_{\ell-1}.$$

Now abbreviate  $t_{\leftarrow}^\ell := \max\{m \in \mathcal{S}_{\tau_\ell} : m < t\}$ . Using the continuous Itô lemma, the noise coupling as stated in subsection 3.1, and the error bound (A.4), we deduce

$$\begin{aligned} (6.7) \quad &\mathbb{E} [(D_{\ell, \ell-1}^z)^4] \\ &\lesssim \mathbb{E} [(D_{\ell, \ell-1}^0)^4] + \int_0^z \mathbb{E} \left[ (D_{\ell, \ell-1}^t)^2 \times \tau_\ell^2 + (D_{\ell, \ell-1}^t)^2 \right. \\ &\quad \times \left. \left\{ \sum_{\xi \in I_{\ell-1}} \left| \left( \sqrt{\rho_\ell^{t_{\leftarrow}^\ell, +}} e^{i\xi(\cdot)}, \nabla_\ell \phi_\ell^{t_{\leftarrow}^\ell} \right)_\ell - \left( \sqrt{\rho_{\ell-1}^{t_{\leftarrow}^{\ell-1}, +}} e^{i\xi(\cdot)}, \nabla_{\ell-1} \phi_{\ell-1}^{t_{\leftarrow}^{\ell-1}} \right)_{\ell-1} \right|^2 \right. \right. \\ &\quad \left. \left. + \sum_{\xi \in I_\ell \setminus I_{\ell-1}} \left| \left( \sqrt{\rho_\ell^{t_{\leftarrow}^\ell, +}} e^{i\xi(\cdot)}, \nabla_\ell \phi_\ell^{t_{\leftarrow}^\ell} \right)_\ell \right|^2 \right\} \right] dt \\ &=: \mathbb{E} [(D_{\ell, \ell-1}^0)^4] + \tau_\ell^4 + \int_0^z \mathbb{E} [(D_{\ell, \ell-1}^t)^2 \times \{A_{I_{\ell-1}} + A_{I_\ell \setminus I_{\ell-1}}\}] dt, \end{aligned}$$

where  $(\cdot)$  in  $e^{i\xi(\cdot)}$  is reserved for the spatial variable  $x$ . The inequality (6.7) implies

$$(6.8) \quad \begin{aligned} &\mathbb{E} [(D_{\ell, \ell-1}^z)^4] \\ &\lesssim \mathbb{E} [(D_{\ell, \ell-1}^0)^4] + \tau_\ell^4 + \int_0^z \mathbb{E} [(D_{\ell, \ell-1}^t)^4] dt + \int_0^z \mathbb{E} [A_{I_{\ell-1}}^2 + A_{I_\ell \setminus I_{\ell-1}}^2] dt. \end{aligned}$$

Step 4: Bounding  $A_{I_{\ell-1}}^2$  in (6.8). By adding and subtracting zero we get

$$\begin{aligned}
 A_{I_{\ell-1}} &\lesssim \sum_{\xi \in I_{\ell-1}} \left| \left( \sqrt{\rho_{\ell-1}^{t_{\ell-1}^{\ell-1},+}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell} - \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell} \right|^2 \\
 &\quad + \sum_{\xi \in I_{\ell}} \left| \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell} - \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell} \right|^2 \\
 &\quad + \sum_{\xi \in I_{\ell-1}} \left| \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell} - \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell-1} \right|^2 \\
 &\quad + \sum_{\xi \in I_{\ell-1}} \left| \left( \sqrt{\rho_{\ell-1}^{t_{\ell-1}^{\ell-1},+}} e^{i\xi(\cdot)}, \nabla_{\ell-1} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell-1} - \left( \sqrt{\bar{\rho}^{t_{\ell-1}^{\ell-1}}} e^{i\xi(\cdot)}, \nabla_{\ell-1} \phi_{\ell-1}^{t_{\ell-1}^{\ell-1}} \right)_{\ell-1} \right|^2
 \end{aligned}$$

(6.9)  $\quad =: R_{\ell} + R_{t_{\ell-1}^{\ell-1}, t_{\ell-1}^{\ell-1}} + R_{\ell, \ell-1} + R_{\ell-1}.$

Adding and subtracting zero (and dropping the dependence on  $t_{\ell-1}^{\ell-1}$ ), we obtain

$$\begin{aligned}
 (6.10) \quad R_{\ell} &\lesssim \sum_{\xi \in I_{\ell-1}} \left| \left( \left[ \sqrt{\rho_{\ell}^+} - \sqrt{\bar{\rho}} \right] e^{i\xi(\cdot)}, \nabla_{\ell} \phi_{\ell} \right)_{\ell} - \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, [\nabla \phi - \nabla_{\ell} \phi_{\ell}] \right)_{\ell} \right|^2 \\
 &\lesssim \left\| \left[ \sqrt{\rho_{\ell}^+} - \sqrt{\bar{\rho}} \right] \nabla_{\ell} \phi_{\ell} \right\|_{\ell}^2 + \left\| \sqrt{\bar{\rho}} [\nabla \phi - \nabla_{\ell} \phi_{\ell}] \right\|_{\ell}^2 \\
 &\lesssim \mathbf{1}_{\inf \rho_{\ell} \leq K} \left\| \left[ \sqrt{\rho_{\ell}^+} - \sqrt{\bar{\rho}} \right] \nabla_{\ell} \phi_{\ell} \right\|_{\ell}^2 + \mathbf{1}_{\inf \rho_{\ell} > K} \cdot \rho_{\min}^{-1} \left\| [\rho_{\ell} - \bar{\rho}] \nabla_{\ell} \phi_{\ell} \right\|_{\ell}^2 \\
 &\quad + \left\| \sqrt{\bar{\rho}} [\nabla \phi - \nabla_{\ell} \phi_{\ell}] \right\|_{\ell}^2 =: R_{\ell,1} + R_{\ell,2} + R_{\ell,3}
 \end{aligned}$$

with  $K$  to be specified later. An analogous bound holds for  $R_{\ell-1}$ .

In order to treat  $R_{\ell, \ell-1}$  (we drop the time dependence  $t_{\ell-1}^{\ell-1}$  for this term), we abbreviate  $g := \sqrt{\bar{\rho}} \nabla \phi$  and perform the rewriting

$$\begin{aligned}
 (6.11) \quad &\left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_{\ell} - \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_{\ell-1} \\
 &= \sum_{x \in G_{\ell-1,d}} h_{\ell-1}^d \left( e^{ix \cdot \xi} g(x) - \sum_{G_{\ell,d} \ni y \sim x} 3^{-d} e^{iy \cdot \xi} g(y) \right),
 \end{aligned}$$

where  $y \sim x$  indicates all points  $y \in G_{\ell,d}$  such that  $|y - x| \leq h_{\ell}$ . In order to sum up over the frequencies  $\xi$ , we integrate by parts on the lattice  $G_{\ell-1,d}$ . More specifically, we rely on the increment relation

$$(6.12) \quad e^{i\xi \cdot (x+h e_j)} - e^{i\xi \cdot x} = e^{i\xi \cdot x} \underbrace{(e^{i\xi_j h} - 1)}_{=: P(\xi_j h)}, \quad j \in \{1, \dots, d\}.$$

Take  $\xi \neq 0$ . Using (6.12) with the same increment  $h = h_{\ell-1}$  in all sums of (6.11), we perform discrete integration by parts  $R$  times in direction  $j$  such that  $\xi_j \neq 0$ , and get

$$\begin{aligned}
 (6.13) \quad &\left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_{\ell} - \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_{\ell-1} \\
 &= \left[ \frac{-1}{P(\xi_j h)} \right]^R \sum_{x \in G_{\ell-1,d}} h_{\ell-1}^d \left( e^{ix \cdot \xi} \partial_{h,j}^R g(x) - \sum_{G_{\ell,d} \ni y \sim x} 3^{-d} e^{iy \cdot \xi} (y) \partial_{h,j}^R g(y) \right),
 \end{aligned}$$

where  $\partial_{h,j}^R f(x) := \sum_{i=0}^R (-1)^i \binom{R}{i} f(x - i h e_j)$  is the standard backward finite difference operator satisfying the relation

$$(6.14) \quad \partial_{h,j}^R f(x) = h^R \partial_j^R f(x) + h^{R+1} C(R) \partial^{R+1} f(\zeta) \quad \text{for some } \zeta = \zeta(x) \in \mathbb{T}^d.$$

By using the order-two bound  $|f(x + h e_j) - 2f(x) + f(x - h e_j)| \lesssim h^2 \|f\|_{C^2}$  in the round bracket of (6.13) (with  $f(x) := e^{i\xi \cdot x} \partial_{h,j}^R g(x)$ ), and the equivalence  $\sin(z) \propto z$  for  $z \in [-\pi/2, \pi/2]$ , we rely on Assumptions 2–4 to integrate by parts  $R$  times in all directions  $j$  with  $\xi_j \neq 0$  (with  $R$  to be determined) and deduce

$$(6.15) \quad \begin{aligned} & \left| \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_\ell - \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_{\ell-1} \right| \\ & \leq \frac{1}{\left( \sum_{j=1}^d |P(\xi_j h_{\ell-1})| \right)^R} \sum_{y \in G_{\ell,d}} h_{\ell-1}^d \|g\|_{C^{R+3}} h_{\ell-1}^R h_{\ell-1}^2 |\xi|^4 \\ & \leq |\xi|^{4-R} \|g\|_{C^{R+3}} h_{\ell-1}^2, \end{aligned}$$

where we have also used (6.14) in the second-to-last inequality and the fact that  $|P(\xi_j h)| = |\sin(\xi_j h/2)|$  for all  $j \in \{1, \dots, d\}$  in the last inequality. The estimate for  $\xi = 0$  is even simpler and requires no integration by parts at all. Provided  $R > 4 + d/2$ , (6.15) implies that we can sum up the terms making up  $R_{\ell, \ell-1}$  in (6.9) and get

$$(6.16) \quad R_{\ell, \ell-1} \lesssim \|g\|_{C^{R+3}}^2 h_{\ell-1}^4.$$

*Step 5: Bounding  $A_{I_\ell \setminus I_{\ell-1}}^2$  in (6.8).* Dropping the dependence on  $t_{\ell-1}^\ell$ , we write

$$(6.17) \quad \begin{aligned} A_{I_\ell \setminus I_{\ell-1}} & \lesssim \sum_{\xi \in I_\ell \setminus I_{\ell-1}} \left| \left( \sqrt{\rho_\ell^+} e^{i\xi(\cdot)}, \nabla \phi \right)_\ell - \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_\ell \right|^2 \\ & + \sum_{\xi \in I_\ell \setminus I_{\ell-1}} \left| \left( \sqrt{\bar{\rho}} e^{i\xi(\cdot)}, \nabla \phi \right)_\ell \right|^2 =: \bar{R}_\ell + \bar{R}_{I_\ell \setminus I_{\ell-1}}. \end{aligned}$$

The term  $\bar{R}_\ell$  obviously has the same bound as the previously treated  $R_\ell$ . Furthermore, as  $R > 4 + d/2$ , and reusing computations from (6.15), we get

$$(6.18) \quad \begin{aligned} \bar{R}_{I_\ell \setminus I_{\ell-1}} & \leq \sum_{\xi \in I_\ell \setminus I_{\ell-1}} \frac{1}{|P(\xi h_{\ell-1})|^{2R}} \|g\|_{C^R}^2 h_{\ell-1}^{2R} \\ & \leq \sum_{\xi \in I_\ell \setminus I_{\ell-1}} \frac{1}{|\xi|^{2R}} \|g\|_{C^R}^2 \leq \|g\|_{C^R}^2 h_{\ell-1}^4. \end{aligned}$$

*Step 6: Taking the expectations in (6.8).* In order to bound  $\mathbb{E}[R_{\ell,1}^2 + R_{\ell,2}^2]$ , we consider the best estimate originating from picking either  $K = \rho_{\min}/2$  or  $K = \infty$  in (6.10): More precisely, using Hölder's inequality, Lemma 6.1, and the simple estimate  $\|\bar{\rho} - \bar{\rho}_\ell\|_\ell \propto h_\ell^2 + \tau_\ell$ , we obtain

$$\begin{aligned} & \mathbb{E}[R_{\ell,1}^2 + R_{\ell,2}^2] \\ & \lesssim_\varphi \min \left\{ \exp \left( -\frac{C \rho_{\min}}{\rho_{\max}^{1/2}} \left[ \frac{N h_\ell^d}{\mu^2 \vee 1} \right]^{\frac{1}{2}} \right) \left[ N^{-1} h_\ell^{-d} (\mu^2 \vee 1) \right] + \rho_{\min}^{-1} \left[ N^{-1} h_\ell^{-d} (\mu^2 \vee 1) \right]^2 \right. \\ & \quad \left. \times \left[ N^{-1} h_\ell^{-d} (\mu^2 \vee 1) \right] \right\} + \min \{ \rho_{\min}^{-2} (h_\ell^8 + \tau_\ell^4); h_\ell^4 + \tau_\ell^2 \}. \end{aligned}$$

As  $\|\nabla\phi - \nabla\ell\phi_\ell\| \propto h_\ell^2 + \tau_\ell$  (cf. Lemma A.3), we get  $\mathbb{E}[R_{\ell,3}^2] \leq C_\varphi(h_\ell^8 + \tau_\ell^4)$ , and an analogous bound holds for  $R_{\ell-1}$ . Similar arguments (involving the same thresholds  $K = \rho_{\min}/2$  or  $K = \infty$ ), a Taylor expansion, and the discrete Parseval identity grant the estimate  $R_{\ell-1} \lesssim_{\varphi} \min\{\rho_{\min}^{-1}\tau_\ell^2, \tau_\ell\}$ .

*Step 7: Concluding the argument.* Combining the estimates in Steps 4–6, we get

$$\begin{aligned} & \mathbb{E} \left[ A_{I_{\ell-1}}^2 + A_{I_\ell \setminus I_{\ell-1}}^2 \right] \\ & \lesssim_{\varphi} \min \left\{ \exp \left( -\frac{C\rho_{\min}}{\rho_{\max}^{1/2}} \left[ \frac{Nh_\ell^d}{\mu^2 \vee 1} \right]^{\frac{1}{2}} \right) \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right] + \rho_{\min}^{-1} \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right]^2 \right. \\ & \quad \left. \times \left[ N^{-1}h_\ell^{-d}(\mu^2 \vee 1) \right] \right\} \\ & \quad + \min\{\rho_{\min}^{-2}(h_\ell^8 + \tau_\ell^4); h_\ell^4 + \tau_\ell^2\} + C(\rho_{\min})\{h_\ell^8 + \tau_\ell^4\}. \end{aligned}$$

Using the above inequality, (4.4), (6.4), and (6.6), we apply Gronwall’s lemma in (6.8) and conclude the proof.

*Remark 6.2.* The proof of Proposition 5.1 with rightmost NN coupling is similar to the one we have provided for the Fourier coupling. The main adjustment concerns adapting the cross-variation structure (i.e., replacing the terms  $A_{I_{\ell-1}} + A_{I_\ell \setminus I_{\ell-1}}$  in (6.7)). It is easy to see that the term  $A_{I_{\ell-1}} + A_{I_\ell \setminus I_{\ell-1}}$  in (6.7) is replaced by

$$\begin{aligned} & \sum_{x \in G_{\ell,d}} h_\ell^d \rho_{h_\ell}^+(x) |\nabla_{h_\ell} \phi_{h_\ell}(x)|^2 + \sum_{x \in G_{\ell-1,d}} h_{\ell-1}^d \rho_{h_{\ell-1}}^+(x) |\nabla_{h_{\ell-1}} \phi_{h_{\ell-1}}(x)|^2 \\ & \quad - 2 \sum_{y \in B_{\ell-1}^+(x)} \sum_{x \in G_{\ell-1,d}} h_\ell^d \sqrt{\rho_{h_\ell}^+(y)} \nabla_{h_\ell} \phi_{h_\ell}(y) \cdot \sqrt{\rho_{h_{\ell-1}}^+(x)} \nabla_{h_{\ell-1}} \phi_{h_{\ell-1}}(x) \\ (6.19) \quad & = \sum_{y \in B_{\ell-1}^+(x)} \sum_{x \in G_{\ell-1,d}} h_\ell^d \left| \sqrt{\rho_{h_\ell}^+(y)} \nabla_{h_\ell} \phi_{h_\ell}(y) - \sqrt{\rho_{h_{\ell-1}}^+(x)} \nabla_{h_{\ell-1}} \phi_{h_{\ell-1}}(x) \right|^2. \end{aligned}$$

From (6.19), it is relatively straightforward to get the  $Err_{num}$  contribution using similar techniques to those deployed in (6.10) (i.e., choosing the best cutoff for the square-root between  $K = \rho_{\min}/2$  and  $K = \infty$ ). All other components of the estimate do not change with respect to the proof in the Fourier case.

*Remark 6.3.* Having two separate estimates originating from two cutoff values  $K = \rho_{\min}/2$  or  $K = \infty$  helps in providing a reliable estimate for  $Err_{num}$  in (5.1): However, these two estimates can—in many cases—still be substantially suboptimal, as they do not rely on the local structure of the mean-field limit  $\bar{\rho}$ . A relevant example is discussed in our simulations in section 7.

**6.2. Proof of Theorem 1.1.** Let  $\kappa$  be the common ratio of the geometric sequence  $(h_\ell)_\ell$ . Assumption 1 implies that, for all levels needed to get to the minimum grid size  $h_{\min}^2 \lesssim \varepsilon$ , we have the bound

$$Var[P_\ell - P_{\ell-1}] \lesssim h_\ell^\beta \propto \kappa^{-\beta\ell}, \quad \beta = \begin{cases} 4 & \text{for Fourier coupling; cf. section 3.1,} \\ 2 & \text{for NN coupling; cf. section 3.2,} \end{cases}$$

while a simple extension of the results in [7] give

$$(6.20) \quad |\mathbb{E}[P_\ell - Q]| \lesssim h_\ell^2 \propto \kappa^{-\alpha\ell}, \quad \alpha = 2.$$

Additionally, the computational cost for simulating a single sample of  $P_\ell$  grows like

$$(6.21) \quad \text{Cost}(P_\ell) \lesssim h_\ell^{-d-2} \propto \kappa^{(d+2)\ell}, \quad \gamma = d + 2.$$

Then (1.4) follows from applying the general MLMC complexity theorem [14, Theorem 2.1].

**6.3. Proof of Proposition 5.4.** Since  $\bar{\rho}$  is assumed to be regular, we have from Proposition 5.1 that  $\text{Var}[P_\ell - P_{\ell-1}] \lesssim (Nh_\ell^d)^{-1} + h_\ell^\beta$ , with  $\beta = 4$  for the Fourier coupling and  $\beta = 2$  for the NN coupling. Using (5.2), and denoting by  $\kappa$  the common ratio of the geometric sequence  $(h_\ell)_\ell$ , we deduce

$$(6.22) \quad \begin{aligned} \text{Var}[\mu_{MLMC}] &\lesssim M_0^{-1} + \sum_{\ell=1}^L M_\ell^{-1} \left[ (Nh_\ell^d)^{-1} + h_\ell^\beta \right] \\ &\lesssim M_0^{-1} + \sum_{\ell=1}^L \left( \frac{h_{\min}}{h_\ell} \right)^{d+2} \left[ (Nh_\ell^d)^{-1} + h_\ell^\beta \right] \\ &\lesssim M_0^{-1} + (Nh_{\min}^d)^{-1} + h_{\min}^{d+2} \sum_{\ell=1}^L (h_{\min} \kappa^{L-\ell})^{\beta-d-2}. \end{aligned}$$

The choice  $L \propto |\log(h_{\min})|$  and (6.22) readily imply the variance bound (5.4). Proving (5.6) is immediate. Finally, the variance reduction factor (1.7) can be readily deduced by dividing the MLMC variance (6.22) with the variance of the standard MC estimator with the same cost as (5.6). Proposition 5.4 is thus proved.

**7. Numerical simulations.** We provide numerical simulations demonstrating the validity of our theoretical results in the two dimensional (2D) case ( $d = 2$ ), and with the NN coupling.

**7.1. Types of experiments.** The experiments conducted are of two types:

*Type 1.* In this type of experiment, we provide computational results for Theorem 1.1 by running the MLMC implementation as described in [14, Algorithm 1]. A minimal version of the pseudocode of this implementation is given in Algorithm 7.1.

In essence, to reach a given accuracy  $\varepsilon$ , Algorithm 7.1 recursively keeps on adding levels and computing the optimal number of samples

$$(7.1) \quad M_\ell = \left\lceil 2\varepsilon^{-2} \sqrt{V_\ell/C_\ell} \left( \sum_{\ell=0}^L \sqrt{V_\ell C_\ell} \right) \right\rceil$$

for all  $\ell = 0, 1, \dots, L$ , where  $L$  is the current maximum level,  $V_\ell := \text{Var}[P_\ell - P_{\ell-1}]$ ,  $\ell = 0, \dots, L$  (for notational simplicity, we always understand the notation  $P_{-1}$  to mean  $P_{-1} \equiv 0$ ), and  $C_\ell$  is the cost of computing one sample on level  $\ell$ . The convergence criterion identifying the level  $L$  at which the algorithm stops is given by the following robust check for the systematic error:<sup>3</sup>

$$(7.2) \quad \text{stop on level } L \text{ if } \max_{\tilde{\ell} \in \{0;1;2\}} \{2^{-\tilde{\ell}\alpha} |\mathbb{E}[P_{L-\tilde{\ell}} - P_{L-\tilde{\ell}-1}]|\} / (2^\alpha - 1) < \frac{\varepsilon}{\sqrt{2}}$$

for  $\alpha$  as in (6.20).

<sup>3</sup>The criterion (7.2) is the same as the one in [14], which is derived under the power law decay assumption for the systematic error on all  $\ell \in \mathbb{N}$ . As we do not have power law decay on all  $\ell \in \mathbb{N}$ , we heuristically expect good performance of this convergence criterion under the—usual—assumption of large average particle density.

---

**Algorithm 7.1.** (see [14, Algorithm 1]).

---

Choose accuracy  $\varepsilon > 0$  and  $\overline{M} \in \mathbb{N}$ .

Start with  $L = 2$ , initial target of  $\overline{M}$  samples on levels  $\ell = 0, 1, 2$ .

**while** extra samples need to be evaluated **do**

    evaluate extra samples on each level.

    compute/update estimates for  $V_\ell := \text{Var}[P_\ell - P_{\ell-1}]$ ,  $\ell = 0, \dots, L$ .

    define optimal  $M_\ell$ ,  $\ell = 0, \dots, L$ , according to (7.1).

    test for weak convergence using (7.2).

**if** not converged, **then**

        set  $L \leftarrow L + 1$

        initialize target  $M_L$ .

**end if**

**end while**

---

*Type 2.* for this type of experiment, the number of samples  $M_\ell$  is chosen to be proportional to a given, specified geometric progression. The variance of the resulting MLMC estimator is then evaluated against the variance of the corresponding standard MC estimator with the same computational time. A computational estimate of the variation reduction factor in Proposition 5.4 is thus produced.

**7.2. Setting and specifics.** The precise details of our setting are as follows:

- Space domain: We work on discretizations of  $\mathbb{T}^2 = (0, 2\pi)^2$ .
- Initial condition for mean-field dynamics: We use two different initial conditions. The first one

$$(7.3) \quad \bar{\rho}_{0,reg} = Z_{reg}^{-1} \left( 1 + e^{-(\sin^2(x-\pi/2) + \sin^2(y-3\pi/2))/2} / (\sqrt{2\pi}) \right)$$

—where  $Z_{reg}$  is the normalizing constant—is bounded away from zero, and with relatively low ratio  $\sqrt{\rho_{\max}}/\rho_{\min} \approx 13.4$ . The second initial condition

$$(7.4) \quad \bar{\rho}_{0,irreg} = Z_{irreg}^{-1} e^{-(\sin^2(x-\pi/2) + \sin^2(y-3\pi/2))/(2 \cdot 0.1)}$$

has ultra-low density regions and much larger ratio  $\sqrt{\rho_{\max}}/\rho_{\min} \gg 10^6$ . The two initial conditions are shown in Figure 1.

- Time-stepping: We stick to a simple explicit Euler–Maruyama scheme, i.e., we pick  $b_0 = 0, b_1 = 1$  in (2.1).
- MLMC discretization: We consider at most  $L = 5$  and use rightmost NN coupling (see subsection 3.2). When  $L = 5$ , for the finest level in space, each axis is split into  $2^7 = 128$  parts, i.e.,  $h_{\min} = 2\pi \cdot 2^{-7} \approx 0.05$ . Furthermore, the finest level in time has time-step  $t_{\min} = 10^{-3}$ . Consecutive levels are designed to preserve the CFL condition: specifically, we choose  $h_{\ell-1} = 2h_\ell$  and  $\tau_{\ell-1} = 4\tau_\ell$ , so that  $\mu = \tau_\ell/h_\ell^2$  is constant over  $\ell$ . Putting all together, we have

$$h_\ell = 2\pi \cdot 2^{-(2+\ell)}, \quad \tau_\ell = 10^{-3} \cdot 4^{5-\ell}.$$

Concerning the number of samples on each level, the following hold:

- For the purpose of estimating the computational gain of the MLMC method for given accuracy  $\varepsilon$  (so for the experiments of Type 1), the number of samples is determined by Algorithm 7.1 (i.e., by (7.1)).

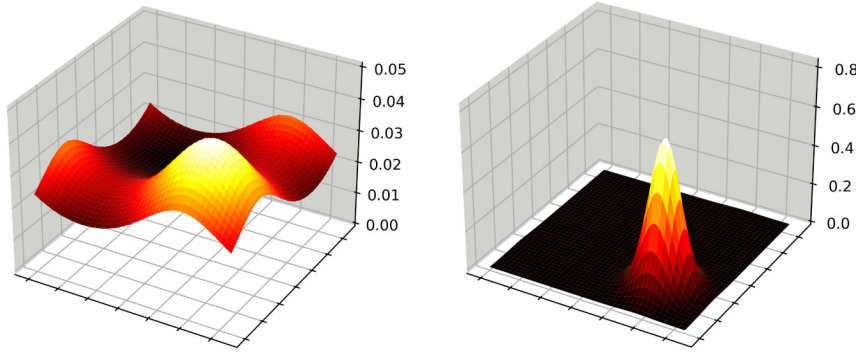


FIG. 1. Left: 3D heat map of  $\bar{\rho}_{0,reg}$ . Right: 3D heat map of  $\bar{\rho}_{0,irreg}$ .

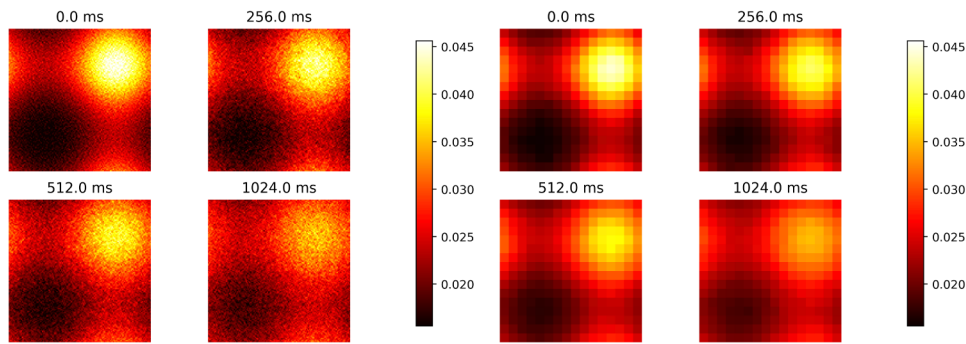


FIG. 2. 2D snapshots of a trajectory of the discrete Dean–Kawasaki equation (2.1) started from  $\bar{\rho}_{0,reg}$  and with  $N = 2 \cdot 10^6$  particles. Left plot: Finest discretization level ( $h = 2\pi \cdot 2^{-7}, \tau = 0.001$ ). Right plot: Coarser discretization level ( $h = 2\pi \cdot 2^{-4}, \tau = 0.001 \cdot 4^3$ ).

- For the purpose of validating the variance reduction factor estimate (1.7) (so for the experiments of Type 2), we instead run the MLMC with the predetermined geometric progression  $M_{\ell-1} = 4M_{\ell}$ .

For qualitative illustration purposes, we show snapshots of a given trajectory of (2.1) in Figure 2 (2D heat maps) and Figure 3 (3D heat maps).<sup>5</sup>

- Time domain: We run simulations on the interval  $[0, 2^{10} \cdot t_{min}] = [0, 1.024]$ .
- Differential operators: The second-order operators used in (2.1) are

$$\Delta_h a_h(x) := \frac{-4a_h(x) + \sum_{y \sim x} a_h(y)}{h^2},$$

$$\nabla_h \cdot b_h(x) := \sum_{r=1}^d \frac{[b_h]_r(x + hf_r) - [b_h]_r(x - hf_r)}{2h},$$

where  $y \sim x$  indicates that  $x, y$  are adjacent grid points, the brackets  $[b_h]_r$  indicate the  $r$ th component of the vector  $b_h$ , and  $\{f_r\}_{r=1}^d$  is the canonical basis of  $\mathbb{R}^d$ .

- Test functions: We use  $\psi(x) := x^2$  and  $\phi(x) := \sin(x) + \sin(y)$ .

<sup>4</sup>We chose this geometric progression for  $M_{\ell}$ —which is slightly different than that indicated in (5.2)—for purely practical implementation reasons. Note that this choice does not affect the scaling of the variance reduction factor (1.7), and this is also verified numerically.

<sup>5</sup>The graphics commands used to generate Figure 2 are based on the code available at <https://scipython.com/book/chapter-7-matplotlib/examples/the-two-dimensional-diffusion-equation/>.

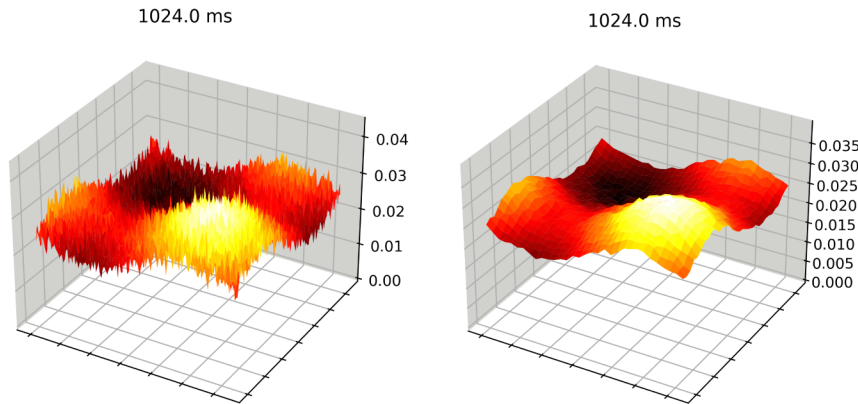


FIG. 3. 3D snapshots at time  $T = 1024$  ms of a trajectory of the discrete Dean-Kawasaki equation (2.1) started from  $\bar{\rho}_{0,reg}$ , and with  $N = 2 \cdot 10^6$  particles, on different levels ( $h = 2\pi \cdot 2^{-7}$ ,  $\tau = 0.001$  for left plot,  $h = 2\pi \cdot 2^{-5}$ ,  $\tau = 0.001 \cdot 4^2$  for right plot).

**7.3. Numerical results.** We report quantitative results of the simulations for

- $N = 2 \cdot 10^9$  and  $\rho_0 = \bar{\rho}_{0,reg}$  (Figure 4 and Table 1);
- $N = 2 \cdot 10^5$  and  $\rho_0 = \bar{\rho}_{0,reg}$  (Figure 5 and Table 1);
- $N = 2 \cdot 10^9$  and  $\rho_0 = \bar{\rho}_{0,irreg}$  (Figure 6 and Table 1).

Figures 4, 6, 5 are made of six plots, namely, variance estimate across levels (top-left), expected value estimate across levels (top-right), time-dependent estimate of variance for MC (middle-left), variance reduction factor between MC and MLMC for various varying  $h_{\min}$  (middle-right), MLMC and MC computational cost versus accuracy  $\varepsilon$  (bottom-left), and  $\varepsilon^2$  times the MLMC and MC computational cost versus accuracy  $\varepsilon$  (bottom-right).

We summarize our findings.

**7.3.1. Variance estimates.** The quadratic decay  $O(h_\ell^2)$  for  $\text{Var}[P_\ell - P_{\ell-1}]$  in (5.1) is clearly visible for the simulation with  $N = 2 \cdot 10^9$  and  $\rho_0 = \bar{\rho}_{0,reg}$  (Figure 4). The same clear decay also holds in the case  $N = 2 \cdot 10^9$  and  $\rho_0 = \bar{\rho}_{0,irreg}$  (Figure 6), even though (5.1) would only predict linear decay  $O(h_\ell)$  in this case. The reason for the better-than-expected result (cf. Remark 6.3) is likely to be the specific nature of the mean-field limit, in which areas of low density coincide with areas of low gradient values, thus making the constant multiplying the quadratic contributions sufficiently low to win over linear decay. When  $N$  is decreased to  $N = 2 \cdot 10^5$  (with  $\rho_0 = \bar{\rho}_{0,reg}$ ), the quadratic decay breaks down for the lowest value of  $h$  (Figure 5): this is to be expected from (5.1), as the contribution  $\propto N^{-1}h_\ell^{-2}$  starts dominating over the numerical error  $\propto h_\ell^2$  at this point.

**7.3.2. Expected values estimates.** Quadratic decay  $O(h_\ell^2)$  for  $\mathbb{E}[P_\ell - P_{\ell-1}]$  is clearly visible in Figures 4, 6, where  $N = 2 \cdot 10^9$ . When  $N = 2 \cdot 10^5$  (Figure 5), the bound saturates its  $h$ -accuracy and breaks down just like the one for the variance. This is consistent with the results shown in [7] and validates the systematic error of our method.

**7.3.3. Variance decay for MC.** We plot the variance of the standard Monte Carlo estimator on level  $L = 5$  as a function of the computational time.

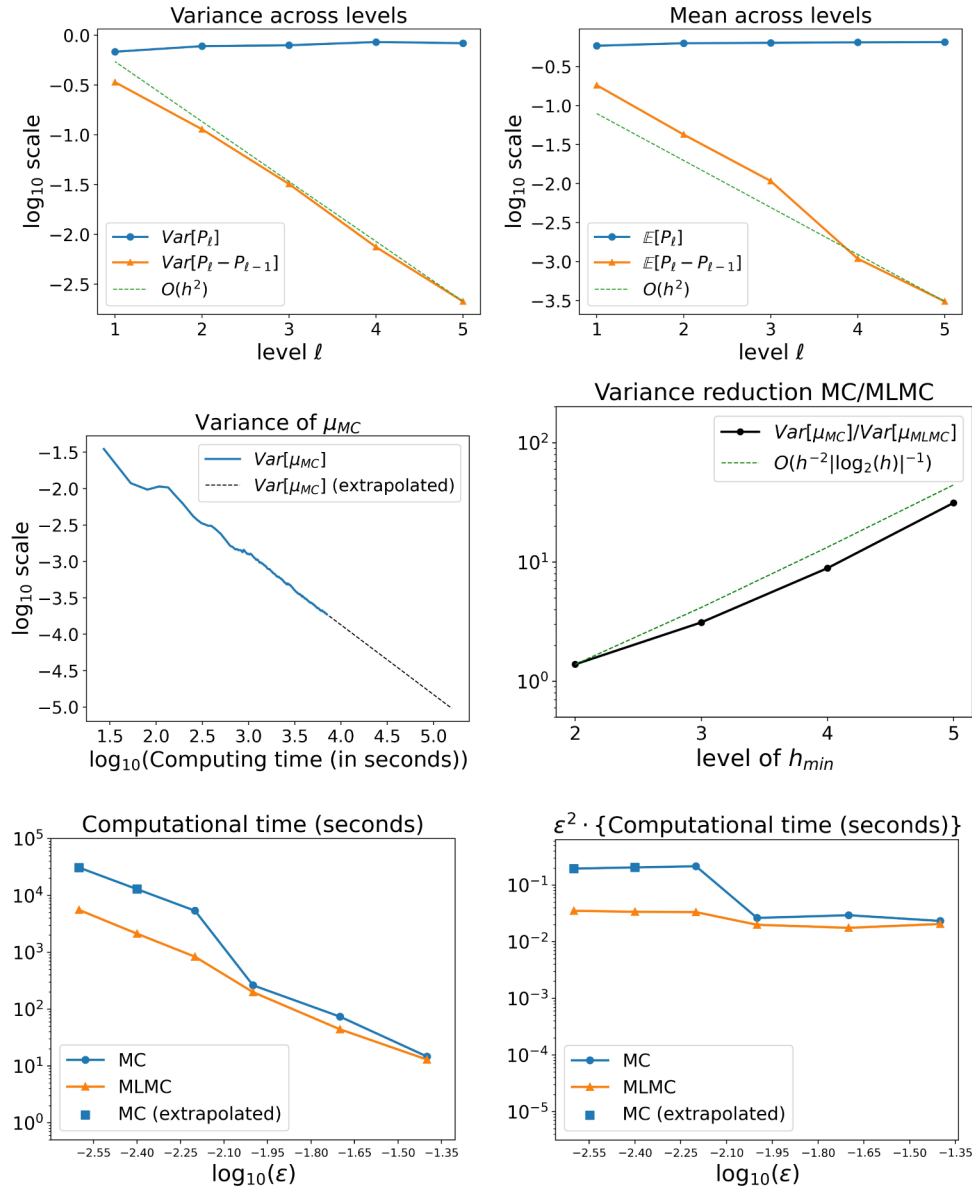


FIG. 4. Computational results ( $N = 2 \cdot 10^9$ ,  $\bar{\rho}_0 = \bar{\rho}_{0,reg}$ ).

**7.3.4. Variance reduction of MLMC over MC.** In all cases with the exception of the simulation with  $N = 2 \cdot 10^5$ , there is clear agreement between the variance reduction between MC/MLMC methods and its ideal predicted growth  $O(h_{min}^{-2}|\log_2(h_{min})|^{-1})$ . For  $N = 2 \cdot 10^5$ , the variance reduction factor stops increasing once the smallest value of  $h_{min}$  is reached. This is consistent with (5.7).

**7.3.5. Accuracy  $\epsilon$  versus computational time (and  $\epsilon^2 \times$  computational time).** For all experiments involving the regular initial condition  $\bar{\rho}_0 = \bar{\rho}_{0,reg}$ , we run Algorithm 7.1 for these values for the accuracy  $\epsilon$ :

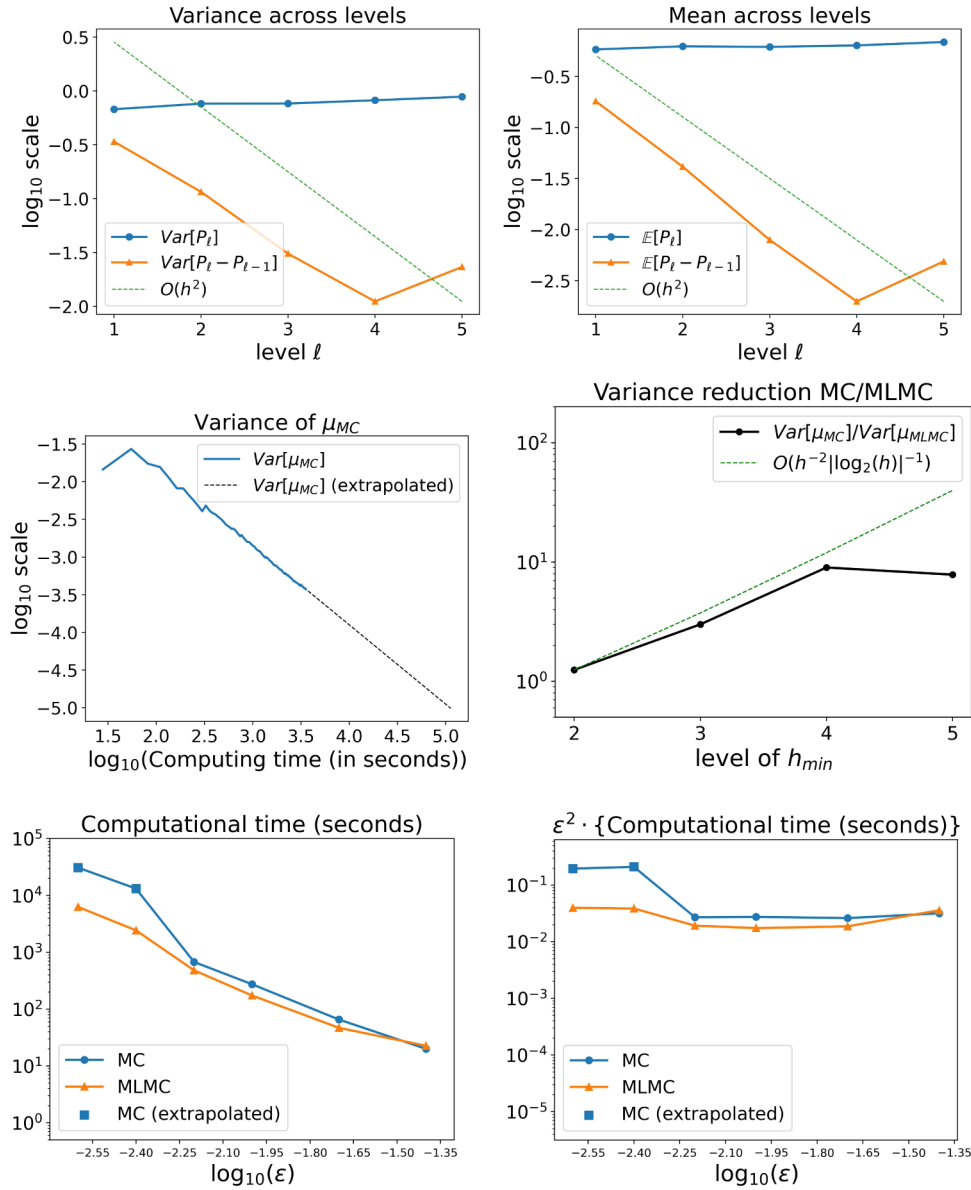


FIG. 5. Computational results ( $N = 2 \cdot 10^5$ ,  $\bar{\rho}_0 = \bar{\rho}_{0,reg}$ ).

$$(7.5) \quad \epsilon = 10^{-1.4}, 10^{-1.7}, 10^{-2}, 10^{-2.2}, 10^{-2.4}, 10^{-2.6}.$$

For the experiment involving the irregular initial condition  $\bar{\rho}_0 = \bar{\rho}_{0,irreg}$ , we run Algorithm 7.1 for these values for the accuracy  $\epsilon$ :

$$(7.6) \quad \epsilon = 10^{-1.4}, 10^{-1.7}, 10^{-2}, 10^{-2.2}, 10^{-2.4}, 10^{-2.75}.$$

Concurrently, for each one of the accuracy values, we also run the standard MC method on the finest level reached by the MLMC method on that very accuracy value. Due to high computational cost, the running time for the MC method for the smallest values of  $\epsilon$  is extrapolated based on the MC variance trend.

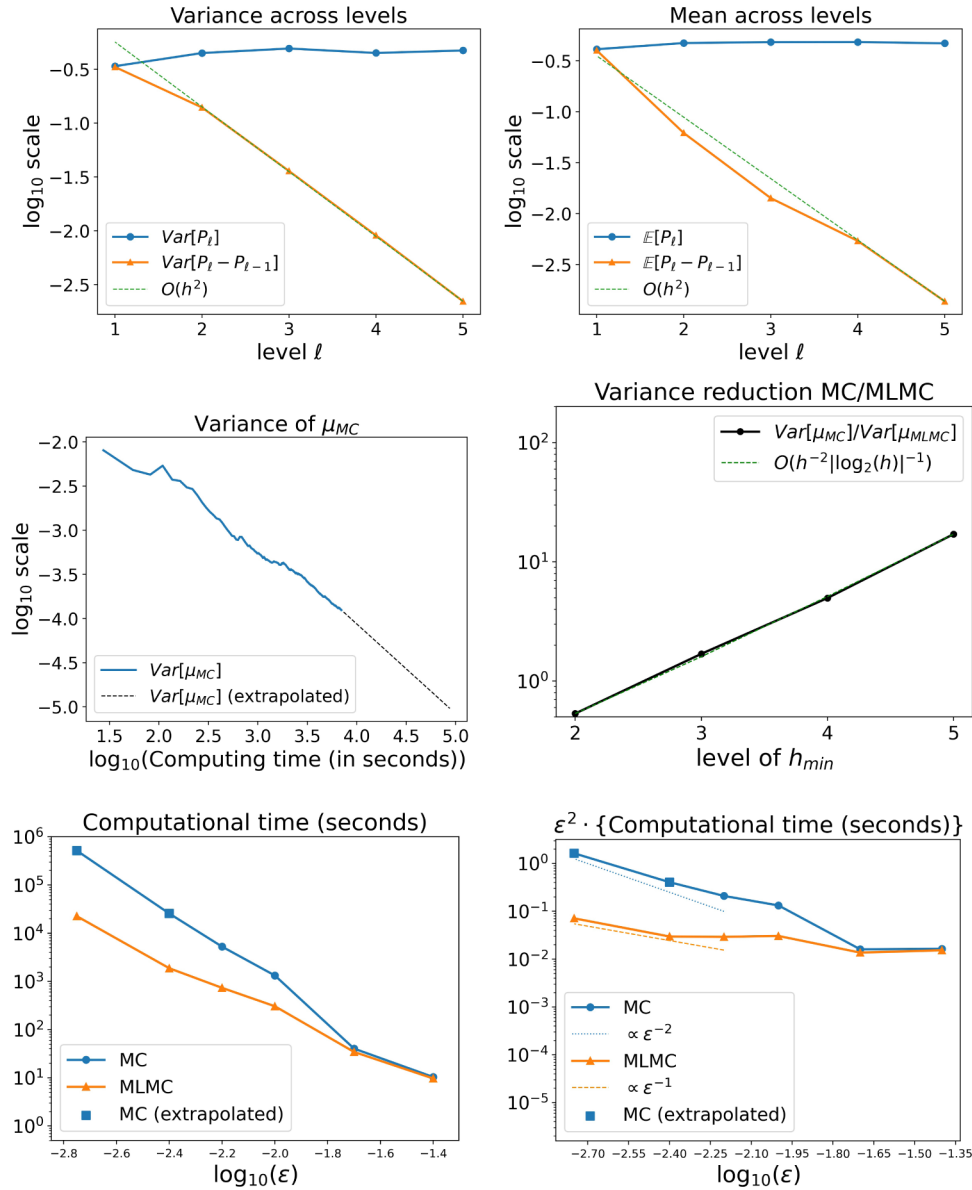


FIG. 6. Computational results ( $N = 2 \cdot 10^9$ ,  $\bar{\rho}_0 = \bar{\rho}_{0,irreg}$ ).

We see that the MLMC provides a tangible computational gain across all simulations. This is especially visible for the smallest accuracy value across all simulations (see Table 1). In the experiment with  $\bar{\rho}_0 = \bar{\rho}_{0,irreg}$ , the MLMC method reaches level  $L = 5$ , while in all other simulations the maximum level reached is level  $L = 4$  (this is reflected with a higher MLMC gain; see Table 1). This is simply due to the systematic error decay having lower multiplicative constants for the experiments with regular datum  $\bar{\rho}_0 = \bar{\rho}_{0,reg}$ . Therefore, in order to see the level change  $L = 5$  also for these simulations, we would have had to set a much lower accuracy (even lower than  $\epsilon = 10^{-2.75}$ ), and this was beyond our computing capabilities.

TABLE 1  
MLMC/MC speed-up factors for smallest accuracies considered across experiments.

	Smallest $\varepsilon$	$L$ reached	Speed-up: $\frac{\text{time MC}}{\text{time MLMC}}$
$N = 2 \cdot 10^9, \bar{\rho}_0 = \bar{\rho}_{0,reg}$	$10^{-2.6}$	4	$\approx 5.6$
$N = 2 \cdot 10^5, \bar{\rho}_0 = \bar{\rho}_{0,reg}$	$10^{-2.6}$	4	$\approx 5.4$
$N = 2 \cdot 10^9, \bar{\rho}_0 = \bar{\rho}_{0,irreg}$	$10^{-2.75}$	5	$\approx 22.8$

In the experiment with  $\bar{\rho}_0 = \bar{\rho}_{0,irreg}$ , where level  $L = 5$  is reached, we see the beginning of the asymptotic region in  $\varepsilon$ , with MLMC and MC rates aligning with those given by Theorem 1.1. This is visible in Figure 6, where the trend of  $\varepsilon^2 \times$  (computational time) starts behaving like  $O(\varepsilon^2 \cdot \varepsilon^{-2} \cdot \varepsilon^{-2/2}) = O(\varepsilon^{-1})$  for the MLMC algorithm, and like  $O(\varepsilon^2 \cdot \varepsilon^{-2} \cdot \varepsilon^{-2/2-1}) = O(\varepsilon^{-2})$  for the standard MC method.

*Remark 7.1.* Additionally, note that the performance of the MLMC algorithm with  $N = 2 \cdot 10^5$  and regular datum is similar to that of those with same datum and larger number of particles  $N = 2 \cdot 10^9$ . This is because, for the accuracies considered, level  $L = 5$ —at which the variance estimate breaks down—is never reached.

**8. Summary of work done and final considerations.** In this paper, we have proved analytically and verified numerically that MLMC methods provide a speed-up in the computation of discretizations of the Dean–Kawasaki equation (1.1), despite this being a highly singular SPDE. We have proposed two different types of noise couplings (“Fourier coupling,” subsection 3.1, and “nearest neighbors (NN) coupling,” subsection 3.2) which grant bounds decaying like a power law in  $2^\ell$  for cross-level variance and systematic error on all levels  $\ell$  for which the average particle density  $Nh_\ell^d$  is sufficiently large (see (1.3)). This has allowed us to get the MLMC complexity result (1.4) and the MLMC/MC variance reduction result (1.7).

While the simulations conducted so far provide, in our opinion, a clear indication of the effectiveness of our results (at least in the cases we have analyzed), we also believe that further testing is needed to consolidate the results. In this respect, the following hold:

- It would be interesting to implement the Fourier coupling and assess its competitive advantage against the NN coupling.
- With the exception of the experiment started from  $\bar{\rho}_0 = \bar{\rho}_{0,irreg}$ , it is difficult to see whether the predicted rate of growth of the MLMC cost is indeed behaving like  $O(\varepsilon^{-2} \cdot \varepsilon^{-1})$ , as it is given in (1.4). This is mainly due to the MLMC algorithm remaining in the  $\varepsilon$ -preasymptotic range.<sup>6</sup> A thorough and comprehensive validation of the  $\varepsilon$ -dependence of the bound (1.4) goes beyond the scope of this paper.

**Appendix A. Estimates for fully discrete setting.** In what follows, it is convenient to use the shorthand notation  $\mu := \tau/h^2$  for the standard CFL quotient.

LEMMA A.1 (discrete energy estimates). *We define the test function dynamics*

$$(A.1) \quad \phi_{h,\tau}^{m\tau} = A_h \phi_{h,\tau}^{(m+1)\tau}$$

<sup>6</sup>This appears to be due—for the case of regular initial datum  $\rho_0 = \bar{\rho}_{0,reg}$ —to systematic error estimates having very low leading constants for the setup considered, and thus preventing the MLMC algorithm from adding too many levels.

with  $A_h := (I_d - \tau b_0 \frac{1}{2} \Delta_h)^{-1} (I_d + \tau b_1 \frac{1}{2} \Delta_h)$ , and where  $b_0, b_1$  are the coefficients from (2.1) (see also Assumption 3). If  $b_1 > b_0$ , the energy estimate for  $\phi_{h,\tau}$

$$(A.2) \quad \tau \sum_m \|\nabla_h \phi_{h,\tau}^{m\tau}\|_h^2 \lesssim (\mu^2 \vee 1) \|\phi_{h,\tau}^T\|_h^2$$

holds subject to  $\mu$  being sufficiently small. If  $b_1 \leq b_0$ , then (A.2) holds with no smallness requirement on  $\mu$ .

*Proof.* We multiply both sides in (A.1) by  $(I_d - \tau \frac{b_0}{2} \Delta_h)$ , compute the  $\|\cdot\|_h^2$  norm of both resulting sides, reshuffle the terms, use the integration-by-parts formula  $(f_h, \Delta_h f_h)_h = -\|\nabla_h f_h\|_h^2$  for all  $f_h \in L^2(G_{h,d})$ , and obtain

$$\begin{aligned} & b_0 \tau \|\nabla_h \phi_{h,\tau}^{m\tau}\|_h^2 + b_1 \tau \|\nabla_h \phi_{h,\tau}^{(m+1)\tau}\|_h^2 \\ &= \|\phi_{h,\tau}^{(m+1)\tau}\|_h^2 - \|\phi_{h,\tau}^{m\tau}\|_h^2 + \tau^2 \frac{b_1^2}{4} \|\Delta_h \phi_{h,\tau}^{(m+1)\tau}\|_h^2 - \tau^2 \frac{b_0^2}{4} \|\Delta_h \phi_{h,\tau}^{m\tau}\|_h^2. \end{aligned}$$

Summing over  $m$ , we control the sum of the contributions  $\tau^2 (b_1^2/4) \|\Delta_h \phi_{h,\tau}^{(m+1)\tau}\|_h^2 - \tau^2 (b_0^2/4) \|\Delta_h \phi_{h,\tau}^{m\tau}\|_h^2$  by using a telescopic argument if  $b_1 \leq b_0$ , or by absorbing the term  $\tau^2 (b_1^2/4) \|\Delta_h \phi_{h,\tau}^{(m+1)\tau}\|_h^2$  in the term  $b_1 \tau \|\nabla_h \phi_{h,\tau}^{(m+1)\tau}\|_h^2$  (subject to  $\mu$  being sufficiently small) if instead  $b_1 > b_0$ . In all cases, (A.2) follows by additionally using that  $b_1^2 \tau^2 \|\Delta_h \phi_{h,\tau}^T\|_h \lesssim \mu^2 \|\phi_{h,\tau}^T\|_h^2$ .  $\square$

*Remark A.2.* As this is not crucial in this paper, we do not seek to obtain the optimal dependency on  $\mu$  in Lemma A.1.

LEMMA A.3 (error estimates for standard heat equation; cf. [24]). *Let  $\phi$  satisfy the continuous backward heat equation ending in  $\phi^T = \varphi \in C^2$ , and let  $\phi_{h,\tau}$  satisfy the discrete backward heat equation (A.1) ending in  $\mathcal{I}_h \varphi$ . Then, for all  $m$ , we have*

$$(A.3) \quad \|\phi(\cdot, m\tau) - \phi_{h,\tau}^{m\tau}\|_\infty \lesssim h^2 + \tau.$$

LEMMA A.4. *For any discrete function  $\phi_h \in H_h^2$ , we have*

$$(A.4) \quad \|\|\nabla_h [(I_d - b_0 \tau \Delta_h)^{-1} \phi_h - \phi_h]\|_h\|_h^2 \lesssim b_0^2 \tau^2 \|\phi_h\|_{H_h^2}^2,$$

where  $H_h^2$  is the discrete Sobolev space of order two.

*Proof.* This follows by taking the Fourier transform of  $y_h := (I_d - b_0 \tau \Delta_h)^{-1} \phi_h$ , which gives  $\hat{y}_h(\xi) - \hat{\phi}_h(\xi) \propto (\tau P(\xi) \hat{\phi}_h(\xi)) / (1 + \tau P(\xi))$  for some  $P(\xi) \propto |\xi|^2$ .  $\square$

LEMMA A.5 (martingale property for discrete fluctuations). *Let (A.1) be the dynamics for the test function  $\phi_{h,\tau}$ . Let  $\rho_{h,\tau}$  be the solution to (2.1). Then  $(\rho_{h,\tau}, \phi_{h,\tau})_h$  is a discrete time martingale.*

*Proof.* The definition of  $A_h$  entails that  $\rho_{h,\tau}^{(m+1)\tau} = A_h \rho_{h,\tau}^{m\tau} + \mathcal{M}_m$ , where  $\mathcal{M}_m$  encodes the noise. Using the symmetry of  $A_h$  and (A.1), we end the proof by writing

$$\begin{aligned} (\rho_{h,\tau}^{(m+1)\tau}, \phi_{h,\tau}^{(m+1)\tau})_h &= (A_h \rho_{h,\tau}^{m\tau} + \mathcal{M}_m, \phi_{h,\tau}^{(m+1)\tau})_h \\ &= (\rho_{h,\tau}^{m\tau}, A_h \phi_{h,\tau}^{(m+1)\tau})_h + (\mathcal{M}_m, \phi_{h,\tau}^{(m+1)\tau})_h \\ &= (\rho_{h,\tau}^{m\tau}, \phi_{h,\tau}^{m\tau})_h + (\mathcal{M}_m, \phi_{h,\tau}^{(m+1)\tau})_h. \end{aligned} \quad \square$$

**Appendix B. Proof of Lemma 6.1.** We need the following simple estimate.

LEMMA B.1. *For any  $\beta > 0$ , there exists a constant  $C > 0$  such that for any  $\alpha \geq 1$  we have the inequality  $\int_0^\infty \exp\{-\alpha x^{1/\beta}/\sqrt{1+x^{1/\beta}}\} dx \leq C\alpha^{-\beta}$ .*

*Proof.* Using the inequality  $1 + x^{1/\beta} \leq x^{1/\beta}(1 + \alpha)$  (valid for  $x \geq \alpha^{-\beta}$ ) we write

$$\begin{aligned} & \int_0^{\alpha^{-\beta}} \exp\left\{-\alpha \frac{x^{1/\beta}}{\sqrt{1+x^{1/\beta}}}\right\} dx + \int_{\alpha^{-\beta}}^\infty \exp\left\{-\alpha \frac{x^{1/\beta}}{\sqrt{1+x^{1/\beta}}}\right\} dx \\ & \lesssim_\beta \alpha^{-\beta} + \int_0^\infty \exp\left\{-\frac{\sqrt{\alpha}y^2}{\sqrt{2}(4\beta)^2}\right\} y^{4\beta-1} dy \lesssim_\beta \alpha^{-\beta}, \end{aligned}$$

where we used the bound  $\alpha/\sqrt{1+\alpha} \geq \sqrt{\alpha}/\sqrt{2}$  (which is valid for  $\alpha \geq 1$ ), the change of variables  $y = 4\beta x^{1/(4\beta)}$  in the second-to-last inequality, and Gaussian moment bounds in the last inequality.  $\square$

*Proof of Lemma 6.1.* Let  $x \in G_{h,d}$  and  $t \in \mathcal{S}_\tau$  be fixed. Let  $\varphi_h$  be the discrete Dirac delta in  $x$ , namely,  $\delta_h := \mathbf{1}_{y=x} h^{-d}$ . Moreover, let  $\phi_{h,\tau}$  be given by the discrete backward heat equation (A.1) and with final datum  $\phi_{h,\tau}^t = \delta_h$ . We define the stopping time  $T_\circ := \min\{s \in \mathcal{S}_\tau : \|\rho_{h,\tau}^s - \bar{\rho}_{h,\tau}^s\|_\infty \geq B\rho_{\min}/2\}$ . Let  $j \in \mathbb{N}$  be even. The same exact reasoning as in Step 3 from the proof of Proposition 5.1 allows one to define a continuous-time martingale  $D^s$  such that  $D^s = (\rho_{h,\tau}^s - \bar{\rho}_{h,\tau}^s, \phi_{h,\tau}^s)_h$  for all  $\mathcal{S}_\tau \ni s \leq t$ . Using the continuous Itô formula, the bound  $\lambda_{\max}(I_d - b_0\tau\Delta_h)^{-1} \leq 1$ , the definition of  $T_\circ$ , and the notation  $s_\leftarrow := \max\{m \in \mathcal{S}_\tau : m < s\}$ , we obtain

$$\begin{aligned} \mathbb{E}[(D^{t \wedge T_\circ})^j] & \leq \mathbb{E}[(D^0)^j] + \frac{j^2}{N} \mathbb{E}\left[\int_0^{t \wedge T_\circ} (D^s)^{j-2} ([\rho_{h,\tau}^{s_\leftarrow}]^+, |\nabla_h \phi_{h,\tau}^{s_\leftarrow}|^2)_h ds\right] \\ & \lesssim \mathbb{E}[\|\rho_{h,\tau}^0 - \bar{\rho}_{h,\tau}^0\|_\infty^j] \|\phi_{h,\tau}^0\|_{L^1}^j \\ & \quad + \frac{j^2}{N} \mathbb{E}\left[\sup_{s \leq t \wedge T_\circ} (D^s)^{j-2}\right] \rho_{\max}(B+1) \sum_{s \in \mathcal{S}_\tau, s \leq t} \|\nabla_h \phi_{h,\tau}^s\|_h^2 \tau. \end{aligned}$$

Lemma A.1 with the fact that  $\|\phi_{h,\tau}^t\|^2 \propto h^{-d}$  and Doob's martingale inequality imply

$$\begin{aligned} \mathbb{E}\left[\sup_{s \leq t \wedge T_\circ} (D^s)^j\right] & \lesssim \mathbb{E}[(D^{t \wedge T_\circ})^j] \\ & \leq (N^{-1}h^{-d})^{\frac{j}{2}} + N^{-1}j^2 \mathbb{E}\left[\sup_{s \leq t \wedge T_\circ} (D^s)^{j-2}\right] \rho_{\max}(B+1)(\mu^2 \vee 1)h^{-d}. \end{aligned}$$

Using Hölder and Young inequalities with conjugate exponents  $j/2, j/(j-2)$  and absorbing the term  $\mathbb{E}[\sup_{s \leq t \wedge T_\circ} (D^s)^j]$ , we arrive at

$$(B.1) \quad \mathbb{E}\left[\sup_{s \leq t \wedge T_\circ} (D^s)^j\right] \lesssim \left(Cj^2 \frac{(B+1)\rho_{\max}(\mu^2 \vee 1)}{Nh^d}\right)^{j/2}.$$

Using Chebyshev's inequality and (B.1), we get

$$(B.2) \quad \begin{aligned} \mathbb{P}\left[\sup_{s \leq t \wedge T_\circ} |D^s| \geq \frac{B\rho_{\min}}{4}\right] & \lesssim \frac{\mathbb{E}\left[\sup_{s \leq t \wedge T_\circ} (D^s)^j\right]}{(B\rho_{\min})^j} \\ & \lesssim \left(\frac{\sqrt{C}\sqrt{B+1}\sqrt{\rho_{\max}}}{B\rho_{\min}} \left\{\frac{\mu^2 \vee 1}{Nh^d}\right\}^{1/2}\right)^j \cdot j^j =: \eta^j \cdot j^j, \end{aligned}$$

where we have set

$$(B.3) \quad \eta := \alpha^{-1} \cdot \frac{\sqrt{B+1}}{B}, \quad \alpha =: \left( \sqrt{C} \frac{\sqrt{\rho_{\max}}}{\rho_{\min}} \left\{ \frac{\mu^2 \vee 1}{Nh^d} \right\}^{1/2} \right)^{-1}.$$

If  $\eta \lesssim 1$ , we can perform an optimization argument over  $j \geq 2$  for the right-hand side of (B.2). Thus the optimal  $j$  is chosen as either  $\lceil e^{-1}\eta^{-1} \rceil$  or  $\lfloor e^{-1}\eta^{-1} \rfloor$ . In either case, this optimization procedure allows one to carry on in (B.2) and obtain (for a different constant  $C$  rescaled by  $e^{-1}$ )

$$(B.4) \quad \mathbb{P} \left[ \sup_{s \leq t \wedge T_{\mathcal{O}}} |D^s| \leq \frac{B\rho_{\min}}{4} \right] \lesssim \exp \left( -\frac{C\rho_{\min}}{\rho_{\max}^{1/2}} \cdot \left\{ \frac{Nh^d}{\mu^2 \vee 1} \right\}^{1/2} \cdot \frac{B}{\sqrt{B+1}} \right).$$

Equation (B.4) holds also in the case  $\eta \gtrsim 1$ , thanks to the trivial bounds

$$\mathbb{P} \left[ \sup_{s \leq t \wedge T_{\mathcal{O}}} |D^s| \geq \frac{B\rho_{\min}}{4} \right] \leq 1 \lesssim e^{-1} \leq e^{-\eta^{-1}}.$$

In particular, (B.4) holds for all  $B \geq 0$ . Using (B.4) and the fact that the test function  $\delta_h := \mathbf{1}_{y=x} h^{-d}$  (being a discrete Dirac delta at point  $x$ ) recovers the value  $\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t$ , we obtain

$$(B.5) \quad \begin{aligned} & \mathbb{P} \left[ t \leq T_{\mathcal{O}}, |\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t|(x) \geq \frac{B\rho_{\min}}{4} \right] \\ & \lesssim \exp \left( -\frac{C\rho_{\min}}{\rho_{\max}^{1/2}} \cdot \left\{ \frac{Nh^d}{\mu^2 \vee 1} \right\}^{1/2} \cdot \frac{B}{\sqrt{B+1}} \right) \quad \forall B \geq 0. \end{aligned}$$

Now (6.1) follows from applying (B.5) over the union of all  $(t, x) \in \mathcal{S}_{\tau} \times G_{h,d}$ .

We now move to proving (6.2). Assume that  $\alpha \geq 1$ , where  $\alpha$  is defined in (B.3). Then the inequality (6.2) is deduced from (6.1) using the equality  $\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X \geq x) dx$  (which is valid for every nonnegative real-valued random variable  $X$ ), a simple change of variables, and Lemma B.1 with the aforementioned  $\alpha$ . Namely,

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in \mathcal{S}_{\tau}} \|\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t\|_{\infty}^j \right] &= \int_0^{\infty} \mathbb{P} \left( \sup_{t \in \mathcal{S}_{\tau}} \|\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t\|_{\infty}^j \geq x \right) dx \\ &= \int_0^{\infty} \mathbb{P} \left( \sup_{t \in \mathcal{S}_{\tau}} \|\rho_{h,\tau}^t - \bar{\rho}_{h,\tau}^t\|_{\infty} \geq B^{1/j} \rho_{\min} \right) \rho_{\min}^j dB \\ &\lesssim \left\{ C^{-1/2} (\rho_{\min}/\rho_{\max}^{1/2}) [Nh^d/(\mu^2 \vee 1)]^{1/2} \right\}^{-j} \rho_{\min}^j, \end{aligned}$$

and (6.2) is proven. Adapting the proof of (6.2) in the case  $\alpha < 1$  to get a bound which is independent of  $\rho_{\min}$  can be done by replicating the analysis leading up to (B.5) by replacing  $\rho_{\min}$  with  $\rho_{\max}$  in the definition of the stopping time  $T_{\mathcal{O}}$  and all relevant thresholds (we omit the details). Finally, (6.3) is easily proved along the lines of [7, Lemma 16].  $\square$

**Data availability statement.** The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

**Conflict of interest.** The authors have no relevant financial or nonfinancial interest to disclose.

**Acknowledgments.** The authors wish to thank the anonymous referees for their very careful reading of the manuscript and for their valuable suggestions and comments, which led to a substantially improved presentation of the paper. The authors also wish to thank Quinn Winters for useful preliminary discussions on the subject of this paper.

## REFERENCES

- [1] A. BARTH, A. LANG, AND C. SCHWAB, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT, 53 (2013), pp. 3–27, <https://doi.org/10.1007/s10543-012-0401-5>.
- [2] A. BARTH, C. SCHWAB, AND N. ZOLLINGER, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161, <https://doi.org/10.1007/s00211-011-0377-0>.
- [3] A. BRANDT, M. GALUN, AND D. RON, *Optimal multigrid algorithms for calculating thermodynamic limits*, J. Stat. Phys., 74 (1994), pp. 313–348, <https://doi.org/10.1007/BF02186816>.
- [4] J. CHARRIER, R. SCHEICHL, AND A. L. TECKENTRUP, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM J. Numer. Anal., 51 (2013), pp. 322–352, <https://doi.org/10.1137/110853054>.
- [5] F. CHUNG AND L. LU, *Concentration inequalities and martingale inequalities: A survey*, Internet Math., 3 (2006), pp. 79–127, <https://doi.org/10.1080/15427951.2006.10129115>.
- [6] K. A. CLIFFE, M. B. GILES, R. SCHEICHL, AND A. L. TECKENTRUP, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14 (2011), pp. 3–15, <https://doi.org/10.1007/s00791-011-0160-x>.
- [7] F. CORNALBA AND J. FISCHER, *The Dean–Kawasaki equation and the structure of density fluctuations in systems of diffusing particles*, Arch. Ration. Mech. Anal., 247 (2023), <https://doi.org/10.1007/s00205-023-01903-7>.
- [8] F. CORNALBA, J. FISCHER, J. INGMANN, AND C. RAITHEL, *Density fluctuations in weakly interacting particle systems via the Dean–Kawasaki equation*, Ann. Probab., to appear.
- [9] M. CROCI, M. B. GILES, M. E. ROGNES, AND P. E. FARRELL, *Efficient white noise sampling and coupling for multilevel Monte Carlo with nonnested meshes*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 1630–1655, <https://doi.org/10.1137/18M1175239>.
- [10] N. DIRR, B. FEHRMAN, AND B. GESS, *Conservative Stochastic PDE and Fluctuations of the Symmetric Simple Exclusion Process*, preprint, 2020, <https://arxiv.org/abs/2012.02126>.
- [11] A. DJURDJEVAC, H. KREMP, AND N. PERKOWSKI, *Weak error analysis for a nonlinear SPDE approximation of the Dean–Kawasaki equation*, Stoch. Partial Differ. Equ. Anal. Comput., 12 (2024), pp. 1–26, <https://doi.org/10.1007/s40072-024-00324-1>.
- [12] B. FEHRMAN AND B. GESS, *Large Deviations for Conservative Stochastic PDE and Non-Equilibrium Fluctuations*, preprint, 2019, <https://arxiv.org/abs/1910.11860>.
- [13] M. B. GILES, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617, <https://doi.org/10.1287/opre.1070.0496>.
- [14] M. B. GILES, *Multilevel Monte Carlo methods*, Acta Numer., 24 (2015), pp. 259–328, <https://doi.org/10.1017/S096249291500001X>.
- [15] M. B. GILES AND C. REISINGER, *Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance*, SIAM J. Financial Math., 3 (2012), pp. 572–592, <https://doi.org/10.1137/110841916>.
- [16] M. GUBINELLI, P. IMKELLER, AND N. PERKOWSKI, *Paracontrolled distributions and singular PDEs*, Forum Math. Pi, 3 (2015), e6, <https://doi.org/10.1017/fmp.2015.2>.
- [17] M. HAIRER, *A theory of regularity structures*, Invent. Math., 198 (2014), pp. 269–504, <https://doi.org/10.1007/s00222-014-0505-4>.
- [18] S. HEINRICH, *Multilevel Monte Carlo methods*, in Large-Scale Scientific Computing, S. Margenov, J. Waśniewski, and P. Yalamov, eds., Springer, Berlin, 2001, pp. 58–67.
- [19] A. KHODADADIAN, M. PARVIZI, M. ABBASZADEH, M. DEGHAN, AND C. HEITZINGER, *A multi-level Monte Carlo finite element method for the stochastic Cahn–Hilliard–Cook equation*, Comput. Mech., 64 (2019), pp. 937–949, <https://doi.org/10.1007/s00466-019-01688-1>.
- [20] V. KONAROVSKIY, T. LEHMANN, AND M.-K. VON RENESSE, *Dean–Kawasaki dynamics: Ill-posedness vs. triviviality*, Electron. Commun. Probab., 24 (2019), 8, <https://doi.org/10.1214/19-ECP208>.

- [21] V. KONAROVSKIY, T. LEHMANN, AND M.-K. VON RENESSE, *On Dean–Kawasaki dynamics with smooth drift potential*, J. Stat. Phys., 178 (2020), pp. 666–681, <https://doi.org/10.1007/s10955-019-02449-3>.
- [22] L. D. LANDAU AND E. M. LIFSHITZ, *Fluid Mechanics: Course of Theoretical Physics*, Vol. 6, 2nd ed., Pergamon Press, Oxford, 1987.
- [23] C. MCDIARMID, *Concentration*, in Probabilistic Methods for Algorithmic Discrete Mathematics, Algorithms Combin., 16 (1998), pp. 195–248, <https://doi.org/10.1007/978-3-662-12788-9>.
- [24] K. W. MORTON AND D. F. MAYERS, *Numerical Solution of Partial Differential Equations: An Introduction*, Cambridge University Press, Cambridge, 2005.
- [25] H. SPOHN, *Large Scale Dynamics of Interacting Particles*, Springer, New York, 2012.
- [26] M. TE VRUGT, H. LÖWEN, AND R. WITTKOWSKI, *Classical dynamical density functional theory: From fundamentals to applications*, Adv. Phys., 69 (2020), pp. 121–247, <https://doi.org/10.1080/00018732.2020.1854965>.