# Structural Properties of Games on Graphs

by

**Jakub Svoboda**

July, 2025

*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:

Francesco Locatello, Chair
Krishnendu Chatterjee
Matjaž Perc
Krzysztof Pietrzak

The thesis of Jakub Svoboda, titled *Structural Properties of Games on Graphs*, is approved by:

**Supervisor**: Krishnendu Chatterjee, ISTA, Klosterneuburg, Austria

Signature: _____

**Committee Member**: Matjaž Perc, University of Maribor, Maribor, Slovenia

Signature: _____

**Committee Member**: Krzysztof Pietrzak, ISTA, Klosterneuburg, Austria

Signature: _____

**Defense Chair**: Francesco Locatello, ISTA, Klosterneuburg, Austria

Signature: _____

Signed page is on file

ISTA Thesis, ISSN: 2663-337X

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 accept full responsibility for the content and factual accuracy of this work, including the data and their analysis and presentation, and the text and citation of other work.

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: _____

Jakub Svoboda
July, 2025

Signed page is on file

# Abstract

The evolution shapes the world around us. Not only in biology, where the fittest individuals spread their genes but also in physics and social dynamics, the evolutionary forces determine the development of a state of matter or public opinions. Many models describe these dynamics. This thesis examines the role of the structure in the models of selection. The population structure is represented as a graph or a network, and each vertex is occupied by one individual. Every individual has a type and fitness that represents the reproductive potential and depends on the type, occupied vertex, and the arrangement of the neighbors. The evolution is modeled in discrete steps; in one step, one individual is replaced by a neighbor selected randomly with the influence of fitness.

The role of the networks is widely examined in the literature. The structures that promote the spread of the desired type compared to the structureless case are called amplifiers. The existence of amplifiers in various settings is an intensively studied topic, and in some settings, the amplifiers have been identified. Moreover, there are other important questions about the number of steps until one type spreads over the whole network (fixation time), the computational complexity, and the questions about the robustness of these processes.

This thesis explores the role of structure in evolution from many perspectives. First, it introduces different models and various choices that can be made in the models of evolution. It highlights the role of the structure in the real world and how this is reflected in these models. Then, it describes the previous results and open problems. Second, the thesis describes an amplifier for two variants of the Moran process: one with a constant birth rate and the other with a constant death rate. This is an important contribution to the robustness of the amplification. Third, the thesis determines the complexity of spatial games. These are processes where the fitness comes from a game, and the strength of selection is high. It shows that determining the fate of cooperation in these games is a PSPACE-complete problem. Fourth, the thesis describes the amplifier of cooperation for spatial games. This is the first amplifier in this setting. Fifth, the thesis examines the coexistence in the Moran process with environmental heterogeneity. In this setting, the fitness depends not only on the type of the individual but also on the occupied vertex. The chapter determines the relationship between the interactions of vertices of different types and the coexistence time. Sixth, the thesis examines the social balance on networks and proposes a stochastic dynamic partially aware of the state of the graph, which reaches a balanced position quickly. Finally, the thesis presents conclusions and outlines the directions for future work.

# Acknowledgements

I would like to thank my supervisor, Krishnendu Chatterjee, first for accepting me for an internship and then for leading me during my whole doctoral studies. I am especially grateful for the freedom to explore different corners of computer science and mathematics.

I thank all my collaborators who helped me with the work presented in this thesis. They are: Josef Tkadlec, Djordje Zikelic, Andreas Pavlogiannis, Ismael Jecker, Kamran Kaveh, Rasmus Ibsen-Jensen, and Soham Joshi. Without them, the thesis would be half as good and twice as painful or even impossible. I especially thank Josef Tkadlec for so many exciting mathematical discussions and for being the pessimist who makes sure all ambitious ideas are mathematically sound.

I would like to thank other professors I had the privilege to collaborate with: Martin Nowak, Seth Gilbert, Stefan Schmid, and Suguman Bansal. In every project, I learned a lot.

I would like to thank all my collaborators who explored with me stochastic games: Raimundo Saona, Ali Asadi, Ruichen Luo, Tobias Meggendorfer, and Mahdi Jafari. Thinking about algorithms for stochastic games was a lot of fun. I would like to thank all who explored distributed algorithms and problems on blockchains with me. Especially Michelle Yeo, who can identify interesting problems and generously share not only the problems but also the tools. I would like to thank Matyáš Křišťan, with whom we explored fun combinatorial problems.

I would like to thank all the people in my group for group dinners, interesting lunch discussions, and playing exciting games of football. Except for the previously mentioned, they are Valentin Hübner (who importantly helped me to clear the final hurdle of converting this document to PDF/A), Mehrdad Karrabi, Ehsan Goharshady, and all the people from the Henzinger group. Moreover, many interns and visitors always contributed to making the group a great place not only for research but also a nice place to be. I would like to thank Ksenja Harpprecht for making sure my reimbursement papers are in order, so I could visit many conferences and collaborators.

I would like to thank my family and friends who supported me on my way.

**Funding.** This work was supported by the European Research Council CoG 863818 (ForM-SMArt) and Austrian Science Fund 10.55776/COE12.

# About the Author

Jakub Svoboda obtained a bachelor's degree in computer science and a master's in theoretical computer science from Charles University in Prague. He joined IST Austria in September 2019. He is interested in problems from evolutionary dynamics, through games on graphs and artificial inteligence, to distributed algorithms. The connecting thread in all of his work is the analysis of stochastic processes. Jakub's work was published in journals such as PNAS, PRE, and PLOS CB, and in conferences such as SODA, LICS, AAAI, ICML, FC, PODC, and with best paper in SIROCCO.

# List of Collaborators and Publications

1. Chapter 2 is based on Jakub Svoboda, Soham Joshi, Josef Tkadlec, and Krishnendu Chatterjee. Amplifiers of selection for the moran process with both birth-death and death-birth updating. *PLOS Computational Biology*, 20(3):1012008, 2024
2. Chapter 3 is based on Krishnendu Chatterjee, Rasmus Ibsen-Jensen, Ismaël Jecker, and Jakub Svoboda. Complexity of spatial games. In Anuj Dawar and Venkatesan Guruswami, editors, *42nd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2022, December 18-20, 2022, IIT Madras, Chennai, India*, volume 250 of *LIPIcs*, pages 11:1–11:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022
3. Chapter 4 is based on Jakub Svoboda and Krishnendu Chatterjee. Density amplifiers of cooperation for spatial games. *Proceedings of the National Academy of Sciences*, 121(50):e2405605121, 2024
4. Chapter 5 is based on Jakub Svoboda, Josef Tkadlec, Kamran Kaveh, and Krishnendu Chatterjee. Coexistence times in the moran process with environmental heterogeneity. *Proceedings of the Royal Society A*, 479(2271):20220685, 2023
5. Chapter 6 is based on Krishnendu Chatterjee, Jakub Svoboda, Djordje Žikelić, Andreas Pavlogiannis, and Josef Tkadlec. Social balance on networks: Local minima and best-edge dynamics. *Physical Review E*, 106(3):034321, 2022

The following list contains works written during the PhD but not included in the thesis. Some works are discussed in Chapter 7.

6. Jakub Svoboda, Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A Nowak. Infection dynamics of covid-19 virus under lockdown and reopening. *Scientific reports*, 12(1):1526, 2022
7. Krishnendu Chatterjee, Tobias Meggendorfer, Raimundo Saona, and Jakub Svoboda. Faster algorithm for turn-based stochastic games with bounded treewidth. In Nikhil Bansal and Viswanath Nagarajan, editors, *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 4590–4605. SIAM, 2023
8. Mahsa Bastankhah, Krishnendu Chatterjee, Mohammad Ali Maddah-Ali, Stefan Schmid, Jakub Svoboda, and Michelle Yeo. R2: boosting liquidity in payment channel networks with online admission control. In Foteini Baldimtsi and Christian Cachin, editors, *Financial Cryptography and Data Security - 27th International Conference, FC 2023, Bol, Brač, Croatia, May 1-5, 2023, Revised Selected Papers, Part I*, volume 13950 of *Lecture Notes in Computer Science*, pages 309–325. Springer, 2023

9. Jan Matyáš Křistan and Jakub Svoboda. Shortest dominating set reconfiguration under token sliding. In Henning Fernau and Klaus Jansen, editors, *Fundamentals of Computation Theory - 24th International Symposium, FCT 2023, Trier, Germany, September 18-21, 2023, Proceedings*, volume 14292 of *Lecture Notes in Computer Science*, pages 333–347. Springer, 2023

10. Krishnendu Chatterjee, Rasmus Ibsen-Jensen, Ismaël Jecker, and Jakub Svoboda. Simplified game of life: Algorithms and complexity. In Javier Esparza and Daniel Král', editors, *45th International Symposium on Mathematical Foundations of Computer Science, MFCS 2020, August 24-28, 2020, Prague, Czech Republic*, volume 170 of *LIPIcs*, pages 22:1–22:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020

11. Ali Asadi, Krishnendu Chatterjee, Jakub Svoboda, and Raimundo Saona Urmeneta. Deterministic sub-exponential algorithm for discounted-sum games with unary weights. In Pawel Sobocinski, Ugo Dal Lago, and Javier Esparza, editors, *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2024, Tallinn, Estonia, July 8-11, 2024*, pages 6:1–6:12. ACM, 2024

12. Jakub Svoboda, Suguman Bansal, and Krishnendu Chatterjee. Reinforcement learning from reachability specifications: PAC guarantees with expected conditional distance. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024

13. Esra Ceylan, Krishnendu Chatterjee, Stefan Schmid, and Jakub Svoboda. Congestion-free rerouting of network flows: Hardness and an FPT algorithm. In *NOMS 2024 IEEE Network Operations and Management Symposium, Seoul, Republic of Korea, May 6-10, 2024*, pages 1–7. IEEE, 2024

14. Ali Asadi, Krishnendu Chatterjee, Raimundo Saona, and Jakub Svoboda. Concurrent stochastic games with stateful-discounted and parity objectives: Complexity and algorithms. In Siddharth Barman and Slawomir Lasota, editors, *44th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2024, December 16-18, 2024, Gandhinagar, Gujarat, India*, volume 323 of *LIPIcs*, pages 5:1–5:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024

15. Zeta Avarikioti, Mahsa Bastankhah, Mohammad Ali Maddah-Ali, Krzysztof Pietrzak, Jakub Svoboda, and Michelle Yeo. Route discovery in private payment channel networks. In *European Symposium on Research in Computer Security*, pages 207–223. Springer, 2024

16. Stefan Schmid, Jakub Svoboda, and Michelle Yeo. Weighted packet selection for rechargeable links in cryptocurrency networks: Complexity and approximation. *Theor. Comput. Sci.*, 989:114353, 2024

17. Jan Matyás Kristan and Jakub Svoboda. Reconfiguration using generalized token jumping. In Shin-ichi Nakano and Mingyu Xiao, editors, *WALCOM: Algorithms and Computation - 19th International Conference and Workshops on Algorithms and Computation, WALCOM 2025, Chengdu, China, February 28 - March 2, 2025, Proceedings*, volume 15411 of *Lecture Notes in Computer Science*, pages 244–265. Springer, 2025

18. Krishnendu Chatterjee, Ruichen Luo, Raimundo Saona, and Jakub Svoboda. Linear equations with min and max operators: Computational complexity. In Toby Walsh,

Julie Shah, and Zico Kolter, editors, *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, pages 11150–11157. AAAI Press, 2025

19. Krishnendu Chatterjee, Mahdi JafariRaviz, Raimundo Saona, and Jakub Svoboda. Value iteration with guessing for markov chains and markov decision processes. In Arie Gurfinkel and Marijn Heule, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 31st International Conference, TACAS 2025, Held as Part of the International Joint Conferences on Theory and Practice of Software, ETAPS 2025, Hamilton, ON, Canada, May 3-8, 2025, Proceedings, Part II*, volume 15697 of *Lecture Notes in Computer Science*, pages 217–236. Springer, 2025

20. David A Brewster, Jakub Svoboda, Dylan Roscow, Krishnendu Chatterjee, Josef Tkadlec, and Martin A Nowak. Maintaining diversity in structured populations. *arXiv preprint arXiv:2503.09841*, 2025

21. Krishnendu Chatterjee, Seth Gilbert, Stefan Schmid, Jakub Svoboda, and Michelle Yeo. When is liquid democracy possible? on the manipulation of variance. In *Proceedings of the 44rd ACM Symposium on Principles of Distributed Computing*, 2025

# Table of Contents

# List of Figures

# List of Tables

CHAPTER 1

# Introduction

The world is complex and ever-changing, and science aims to understand all the rules governing it. Physics examines how particles change their states, how they create attachments, and how systems settle into positions with minimal energy. Biology explores how cells form organs and whole bodies, how bacteria spread in a beneficial environment, and how animals compete for resources. Social sciences and game theory examine how humans interact with each other, how they cooperate, and how they spread ideas.

For many behaviors in nature, the structure or spatial arrangement plays a crucial role. The states of particles in physics are influenced by the particles nearby. The spread of infection is determined by the population structure. The likelihood that the population consists of altruistic individuals depends on the network of friendships among people in the population.

Another important force is randomness. In physics, the next state of a particle is influenced by particles nearby, but they do not completely determine the next state; they only change the probability. In biology, the success of a particular animal depends on its fitness. However, even the most fit individual can perish in an unexpected event and not produce any offspring. In social dynamics, the idea spreads not only if it is correct but also if it becomes fashionable, and there, randomness plays a crucial role.

This thesis explores the fascinating topic of the role of structure and randomness in many models from physics, biology, and computer science. This research direction is extremely diverse. There are a plethora of questions to be asked. However, these models share many commonalities, which means similar theoretical techniques and approaches can be applied to answer these important questions.

## 1.1 Evolutionary dynamics in many fields

### 1.1.1 Biology

The field of evolutionary biology examines the emergence and the spread of a mutation. There are many models in evolutionary biology, each fitting a particular reality. This work focuses on models of selection. This means that at the beginning of the process, all different types of individuals are already present, and the process continues until one of the types spreads over the whole population. These models suppose that mutation is rare and the population spends most of the time in the homogenous state.

Many important choices reflect particular realities of given models. First, one choice in the models is the size of the population. The size of the population can be either finite or infinite. Variable population size describes the introduction of a new species into an environment. The constant population size describes the fact that a new mutation competes against relatively well-adapted opponents.

Second, in the models, the individuals can be treated as discrete or continuous. The models can either describe the population as consisting of different discrete individuals or they can treat the composition of the population as ratios between different types. In a continuous population, the more fit individuals surely proliferate, and the question is, what is the ratio between the mutant and the wild-type or resident? In the discrete population, usually, only one individual is the carrier of a gene or a mutation. In this stochastic process, it can die and be replaced by another, while the mutation dies with the individual.

Third, another aspect of the models is how they treat generations of individuals. The new offspring can be either incorporated into the population one by one (overlapping generations) or the new offspring form separate generations. In the models with separate generations, all individuals create a pool of offspring, and then from this pool, a new generation is selected. With overlapping generations, only a subset of the population reproduces and potentially replaces another subset of the population.

Fourth, the decision also lies in how to treat time. Time is either discrete or continuous. In discrete time, every reproduction happens individually at one time point, and reproduction happens step-by-step. Continuous time is usually combined with an infinite continuous population. There, the equations based on time describe the composition of the population.

These are only simple choices for basic models. There are various possible extensions, such as considering mutation or sexual reproduction. The choices are used in various models, some of the most famous and widely used are introduced below.

**Moran process.** The Moran process, developed by Patric Moran [Mor58], describes a finite population of discrete individuals with overlapping generations in a discrete time. The population consists of $n$ individuals, one of them is a mutant with reproductive fitness $r$ (usually $r > 1$), the rest of the individuals are called *residents* with reproductive fitness $1$. In one step of the process, one random individual is selected according to its fitness, and its copy reproduces a random individual from the population.

The population consists of $n$ individuals, and the composition of the population is described by the number of mutants ($a$) and the number of residents ($n-a$). This means a one-dimensional Markov Chain with $n+1$ states can describe the whole dynamic. The Markov Chain has two absorbing states for $a = 0$ and $a = n$, and the transitions connect only states that represent configurations where the number of mutants differs by $1$. In one state of the Markov Chain, the probability of increasing the number of mutants by $1$ is $r$ times the probability of decreasing the number of mutants. The process starts in the state with only one mutant, and mutants spread over the whole network if they reach a state with $a = n$. This event is called *fixation*. Since the fate is described by a one-dimensional Markov Chain, the fixation probability of mutants can be computed in a closed form as:

$$\frac{1-\frac{1}{r}}{1-\frac{1}{r^{n-1}}}.$$

**Wright-Fisher model.** Wright-Fisher model [Wri31, Fis99] describes the evolution in a finite population with discrete individuals and separate generations in discrete time. The population consists of $n$ individuals of different types. In one generation, $n$ individuals are selected with replacement from the previous generation. The number of selections of one individual denotes the number of offspring in the next generation. Compared to the Moran process, there are no beneficial mutations; all individuals have the same fitness, which represents the genetic drift. This means that the fixation probability of one individual is

$$\frac{1}{n}.$$

**Lotka–Volterra equations.** Lotka-Volterra equations [Lot25, Lot27, HS98] describe the dynamics of two types of individuals, usually denoted by predator and prey. Individuals as well as time are continuous. The evolution of the population is described by two first-order nonlinear differential equations.

**Replicator dynamics.** A more general version of the Lotka-Volterra equations is replicator dynamics [TJ78, HS98, HS88]. They usually consider an infinite population with continuous individuals in continuous time. The population state is given by a vector of frequencies of different types, and the dynamics is given by a differential equation that determines the frequencies of individuals after time $t$.

### 1.1.2 Physics

Physics also studies models similar to evolutionary biology. However, these models try to explain a different reality. Usually, they describe an inner state of charged particles. Configurations of the population of particles have different energies. The process then describes the evolution of the population towards the state with the lowest energy or towards the local minima.

There are a lot of similarities: population can be finite or infinite; the particles can be represented by the number or the proportion; the generations can be either overlapping or non-overlapping; and the time can be treated either discretely or continuously. However, there are some differences: the most important difference is that during the process, new types (or charges) can stochastically appear.

**Ising model.** The Ising model [Isi25, Ons44, Yeo92] represents ferromagnetism. The population consists of particles that can have a spin or a charge, either $+1$ or $-1$. The particles are arranged on a lattice, possibly of a higher dimension. Two neighboring particles with the same spin have lower energy than two particles with different spins. The particles randomly switch their spins to minimize the energy, however, heat (or an incoming energy) can disturb this process. The most interesting questions in these models are phase transitions based on the amount of incoming energy.

**Game of life.** The Game of Life [BCG04] and, in general, cellular automata [WGeH03] are a general model in physics and computer science. The individuals are arranged on a grid, and they have two states. In one generation, every individual changes its state based on the states that it sees in the neighborhood. The model exhibits very complex behavior, and problems are hard to solve [CIJS20] even for simple rules.

### 1.1.3 Social Dynamics

Social dynamics models the spread of strategies, opinions, or infection in a society. The individuals represent people, and their state is the strategy, opinion, or health status. The choices that can be made in these models are very similar to those in biology, however, the individuals can be more complicated. For instance, the individual can deterministically change its state when at least half of the population is in a particular state.

**Voter model.** The voter model [CS73, Lig13, PJR+17, Per16] describes the population of individuals, each holding a binary opinion. One individual is randomly selected to potentially update their opinion. Based on the neighbors that represent friends or people the individual interacts with, the new opinion is selected.

**Contact Process.** The contact process [Har74, Lig12] describes the spread of disease in the population. Individuals represent people, and the state represents health status. Infected individuals become healthy at a constant rate, however, the infected individuals can spread the disease to their neighbors.

## 1.2 The role of the networks

Some of the models, mainly from physics and social dynamics, suppose that there is some population structure, usually a grid. In general, the population structure influences the trajectory in all evolutionary dynamics [Now06a, SF07, PJR+17]. Some of the models are better suited to consider the population structure. The models that consider the population structure should consider a finite population, and the individuals should be treated discretely, such that their neighbors can be identified. This section first defines graphs, which formally describe the structure, and then introduces ways to add the population structure to the two models examined later.

**Graphs.** Graph $G = (V, E)$ consists of a set of vertices (or nodes) $V$ and a set $E$ of unordered pairs of vertices. Two vertices connected by an edge are called *neighbors*. A weighted graph consits of a triple $G = (V, E, w)$ where $w$ is a weight function that for every edge $e \in E$ returns a (positive) number denoting the weight. The weight of an edge signifies the strength of the connection. The graph can be directed, in that case $E$ consits of ordered pairs of vertices. The orientation of an edge means that only one vertex can influence another, not the other way around. A complete graph is a graph where $E$ contains all pairs of vertices and is denoted by $K_n$. A graph is $d$-regular if every vertex is contained exactly in $d$ edges.

### 1.2.1 Moran process

In the Moran process, the structure was first considered in [LHN05]. The structure is represented by a graph with $n$ vertices, and every vertex is occupied by one individual. With structure, we recover two different update rules that nearly coincide in the structureless case. They are Birth-death and death-Birth updating. The name signifies the order of events, the Birth is capitalized to highlight that fitness plays a role in this step.

**Birth-death.** In the Birth-death process, first, an individual $v$ from the population is selected proportionally to its fitness. Second, a random neighbor $u$ is selected proportionally to the weight of the edge coming from $v$. Then, the copy $v$ replaces the individual occupying vertex $u$.

**death-Birth.** In the death-Birth process, death precedes birth. First, an individual $v$ is selected to die. Second, among its neighbors, $u$ is selected proportionally to the fitness times the weight of an edge coming to $v$. Then, $v$ is replaced by a copy of $u$.

### 1.2.2 Replicator dynamic

Replicator dynamics are used in physics and social dynamics [SF07, PS10a, NM92a, SP05] and also consider a population structure expressed as a network. As opposed to the Moran process, in replicator dynamics, one edge between $u$ and $v$ is selected randomly. Then, either $u$ replaces $v$ or $v$ replaces $u$. With higher probability individual with a higher fitness replaces an individual with a lower fitness. This parameter can be tuned and is known as the strength of selection. The strength of selection can be weak, where the fitness plays a very small role in the replacement, to the limit of strong selection, where the more fit individual surely replaces the less fit one.

## 1.3 Constant and Frequency Dependent Selection

This section discusses the fitness of individuals. In general, the fitness can depend on the whole configuration of the network and the spatial arrangement of the types. However, this is untractable and some simplifications have to be made. In general, models consider into the fitness the following:

- the type of an individual;
- the inhabited vertex; and
- the number, types, and arrangement of the neighbors.

Considering these parameters, for the computation of the fitness, the knowledge of the local neighborhood is needed.

### 1.3.1 Constant Selection

In the constant selection, the fitness of the individual depends only on the type of the individual and potentially the vertex it inhabits. This represents a simple beneficial mutation that provides a higher fitness compared to the original type. In these settings, the more fit type has fitness $r > 1$ and is called a (beneficial) mutant, and the less fit type with fitness $1$ is called a resident. Usually, the dynamic starts with only one mutant in a graph inhabited by residents. The constant selection is usually examined in the context of the Moran process, either Birth-death or death-Birth. The most studied question about the process is: What is the probability that the mutant spreads over the whole population (fixation probability)?

Potentially, not only the type, but also the inhabited vertex and type can be considered for the computation of the fitness. If one type has always higher or equal fitness than the other type, for instance, in the case where the fitness advantage is felt only in some

vertices [BKP+22, EDS07], the advantageous type is still called mutant, and the question still regards the fixation probability of a mutant. However, in general, there might exist vertices where one type has a fitness advantage, together with vertices where the other type has a fitness advantage. In this setting, the main question changes. The main interest is in the number of steps until one of the types fixates.

### 1.3.2 Frequency dependent selection

The frequency-dependent selection considers the type of the individual together with the arrangement of the neighborhood. In general, any arrangement of the neighbors can translate to any fitness. However, in practice, the games are used for determining fitness.

**Games.** Game theory [VNM47, OR94] is an important field that studies optimal decision making. Two players independently and simultaneously can choose an action (strategy), and then they receive a payoff based on the actions selected. The payoff is determined by a payoff matrix. In evolutionary dynamics on graphs, the type is usually equated with a pure strategy (an individual selects the same action every time). All individuals play the game with all of their neighbors and collect the payoff. The payoff can be either summed from all interactions or averaged.

**Fitness from games.** The payoff is then translated to the fitness (or, in the replicator dynamics, the payoffs of two selected individuals are compared by a function that corresponds to the fitness). In the weak selection, the fitness is multiplied by a small constant and then added to $1$. In medium and strong selection, the fitness is an exponential function of the payoff [MRH21].

**Important games.** Any matrix can represent a game. However, in practice, the most interesting games are social dilemmas. Where either the global welfare is maximized by one strategy, however, selfish individuals benefit from switching away from that strategy, or where coordination is needed.

The game is described by the following general matrix of payoffs for the first player:

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}$$

The first player chooses the row of the matrix, and the second player chooses the column of the matrix. The first player receives the payoff equal to the selected cell of the matrix, and the second player receives the payoff equal to the selected cell of the transposed matrix. In many settings, the first action (first row or first column) is treated as a cooperation where both players receive $R$, the reward for cooperation. The first player can change to the second action (called defection) and then receives the reward $T$, the temptation for defection against cooperation. If the first player cooperates and the second player defects, the first player recieves the sucker's payoff. When both players defect (play the second action), they receive the punishment $P$.

Based on the ordering of the payoffs $R$, $S$, $T$, and $P$, we have different games and dilemmas. If $R \geq T, S \geq P$, both players individually prefer to cooperate, there is no dilemma. If $T \geq R \geq S \geq P$, the game is called the snowdrift game (or hawk-dove game) [HD04]. In the game, both players are tempted to defect, however, it is better to cooperate against defection

than to defect. If $R \geq T \geq P \geq S$, the game is called the stag hunt game [Cam11]. When both players are cooperating, there is no incentive to change one's strategy. However, it is also impossible to escape mutual defection, since $P$ is higher than $S$ for unilateral cooperation. Finally, $T \geq R \geq P \geq S$, gives the Prisoner's dilemma [Nas51]. Any player in any situation is tempted to defect to obtain a higher payoff: If the other player cooperates, defecting gives $T$, which is bigger than $R$. If the other player defects, defecting gives $P$, which is bigger than $S$. The prisoner's dilemma is one of the most widely studied games because it captures the tension between cooperation and defection.

The matrix games are very general; some predator-prey dynamics can be expressed as a matrix game (then $T \geq P \geq R \geq S$). The preference for the opposite type can be expressed by $R = 0$, $S = 1$, $T = 1$, and $P = 0$, and the preference for the same type as $R = P = 1$ and $S = T = 0$.

## 1.4 Previous results

### 1.4.1 Constant selection

The exploration of the constant selection mainly considers Birth-death updating, however, there are also some results for the death-Birth updating.

**Amplifiers.** In the Moran process with constant selection, the well-mixed population is considered a baseline. There, one mutant with fitness $r > 1$ spreads over the whole population with probability $1 - \frac{1}{r}$ in the limit of a large population [Mor58]. In the presence of structure, the fixation of one randomly placed mutant can increase. The graph that increases the fixation probability for a randomly placed beneficial mutant is called an *amplifier*; on the other hand, the graph that decreases the fixation probability compared to the well-mixed case is called *suppressor*. Moreover, with the structure, it matters whether the birth event is followed by death or vice versa.

**Isothermal theorem.** Isothermal theorem [Now06a, ALN19] examines the fixation probability of mutants on regular graphs, which are graphs where all vertices have the same degree and contain, for instance, grids of any dimension. The theorem states that the fixation probability on the regular graph is the same as in the complete graph. This means that the structure alone is not enough, also, the imbalance between the number of neighbors and vertices is needed to achieve amplification.

**Amplification in Birth-death updating.** The study of the Moran process starts with an observation that the structure can increase the fixation probability of a beneficial mutant in the Birth-death process. On the star, very simple graph, the fixation probability of a randomly placed mutant with fitness $r > 1$ is $1 - \frac{1}{r^2}$, see [LHN05]. The structure has a similar effect as increasing the fitness for the mutant from $r$ to $r^2$. The reason for amplification is that the central vertex of the star is replaced very often; in essence, one vertex needs to be selected twice in a row to reproduce. One resident reproduces twice with probability proportional to $1 \cdot 1$, however, one mutant reproduces with probability proportional to $r \cdot r$.

In Birth-death updating, there exist even stronger amplifiers than the star. There are so-called *superamplifiers* that for all $r > 1$ increase the fixation probability to $1$ in the limit of large population structure, see [Gia16, GLL$^+$19, GGG$^+$17, PTCN17, TPCN19]. This is the strongest

possible amplification that guarantees almost sure fixation probability to any beneficial mutant. The main idea for amplification is to have many vertices with a small incoming degree. There, one mutant can stay and potentially spread to more connected vertices. Since the mutant has a fitness advantage, it is more likely to spread further before it gets replaced.

There is also an interesting relationship between the fixation probability and fixation time [TPCN21]. Any superamplifier increases the number of steps until fixation. The reasoning behind this is that in the process that is fast, the benefits of the mutation are harder to realize.

Besides the amplifiers, for Birth-death updating, there also exist suppressors, see [Gia16]. These graphs ensure that the mutant reproduces into a vertex that is soon replaced by a resident.

**Amplification in death-Birth updating.** In the death-Birth updating, the amplification is much harder to achieve. The best-known amplifier requires weighted edges or amplifies with extremely low advantage, see [Ric21, Ric23]. Moreover, there are limits on amplification in death-Birth updating, see [TPCN20]. Since death is followed by reproduction, for large fitness $r$, the mutant cannot use large fitness if it dies first or its neighbor does not die. In graphs that are not complete, this means the fitness advantage is used with smaller probability, and these graphs are suppressors.

**Approximating the fixation probability.** Diaz et al. [DGRS16] shows that the process with constant selection finishes in time that is polynomial with the size of the network. This means that approximating the fixation probability of a given time can also be done in polynomial time. The approach is to simulate the process many times until fixation. The empirical probability is then a good approximation of the real fixation probability.

**Heterogenous networks.** The networks are heterogeneous in practice; a vertex of a graph can have an influence on the fitness of an individual inhabiting the vertex [KMN19]. There are many research directions in heterogeneous networks, for instance, how to select the vertices where the mutant has an advantage to maximize the fixation probability [BKP+22].

### 1.4.2 Frequency dependent selection

Frequency-dependent selection is studied from more sides. It is studied not only from a biological perspective, but also from the perspective of physics and social dynamics. The literature is more extensive than in the constant selection. This means only selected results are highlighted here. The works mostly focus on the prisoner's dilemma, since this is a simple game that nevertheless captures the essence of the most important question.

**Supporting cooperation.** Rational agents should never cooperate in one game of Prisoner's Dilemma. However, cooperation is not hard to find in the real world. There are many mechanisms that support the cooperation [Now06b], such as reciprocity, reputation, or the population structure.

**Updating rules.** In the constant selection on graphs, there are some similarities between Birth-death and death-Birth updating. In the frequency-dependent selection, the choice of the updating rule is even more important. The work of Ohtsuki et al. [OHLN06] shows that in the death-Birth updating, the cooperation has a chance to spread in the donation games

if the benefit $b$ divided by cost $c$ (also parametrized as benefit-to-cost ratio $\beta$) is above the degree $d$. On the other hand, the cooperation never spreads in the Birth-death updating.

**Weak selection.** Moreover, [OHLN06] introduces an important technique for computing the fixation probability in evolutionary games with weak selection. This technique is then used to discover graphs that promote cooperation for the benefit-to-cost ratio $\beta > 1.5$ [ALC$^+$17, FMAN18].

**Payoff-to-fitness function.** The techniques for the weak selection consider a linear payoff-to-fitness function. In the weak selection, this is a sensible choice, since the effect of the payoff on fitness is small. McAvoy et al. [MRH21] show that the only function that satisfies all natural requirements to translate the payoff to fitness in games with medium or strong selection is the exponential function.

**Hardness of computation.** An important theoretical question in computer science is the computational complexity of a problem. For the frequency-dependent selection, the problem can be formulated as: given a configuration of different types on the graph, compute the probability that one given type spreads over the whole network. Ibsen-Jensen [IJCN15] shows the hardness of the problem for a specific frequency-dependent selection.

**Scale-free networks.** In the evolutionary games, any analytical result is hard to obtain. This gives rise to many empirical approaches that simulate the process for different graphs and different settings [WSP13a, WSP13b, JP13, PS10a, SP05]. The main result is that the scale-free networks that mimic the social networks are better conducive to cooperation than other graphs.

## 1.5 Open questions

Evolutionary dynamics on graphs is a wide research topic. There are many questions and open problems. This section highlights some of the problems that answer natural questions about the evolutionary dynamics. Moreover, these questions deal with various aspects of these dynamics, not only amplification and the fixation times, but also robustness and computational complexity.

1. **Amplification for Birth-death and death-Birth updating.** In the constant selection, the existence of amplifiers, graphs that increase the fixation probability of a randomly placed advantageous mutant, is a widely studied topic. There are known amplifiers for both the Birth-death and death-Birth processes. However, these amplifiers are not robust with respect to the updating rule. An important question is: Do there exist graphs that are amplifiers for both Birth-death and death-Birth processes?

2. **Complexity of spatial games.** The work of [IJCN15] shows the computational hardness of a version of frequency-dependent selection. However, the case of spatial games on the graphs is not covered by the proof. The main question is: What is the computational complexity of approximating the fixation probability of cooperators?

3. **Existence of general amplifiers for spatial games.** The notion of amplifiers is well established in the Moran process and the constant selection regime. For some

range of parameters, there exist graphs that promote cooperation, such as graphs from [ALC+17, FMAN18]. However, these graphs promote the cooperation in regimes where the cooperation is comparably a good choice against defection (high benefit-to-cost ratio). In the games, where the temptation for defection $T$ is high compared to the payoff for mutual cooperation $P$, the existence of any graphs that can promote cooperation is an open question of great importance.

4. **Heterogenous networks.** In the heterogeneous networks, the fitness of one type also depends on the vertex it occupies. The amplification is not the only question, it is also important to know what graphs can maintain a diversity of the population. Or in other words, how long can multiple types coexist, and what are the necessary parameters for that coexistence? Note that in this setting, the bounds on the number of steps from [DGM+13] no longer hold.

5. **Changing the population structure.** In most of the literature, the underlying population structure is fixed. However, this is not the case in practice, where the individuals can break friendships or the structure can change over time. The questions about the dynamic structure are underexplored.

## 1.6 Thesis Outline and Contributions

The thesis presents five scientific peer-reviewed articles [SJTC24, CIJS22, SC24, STKC23, CSŽ+22] dealing with the questions relating to the role of networks in evolutionary processes.

**Amplification for Birth-death and death-Birth updating.** Chapter 2 presents the first amplifier for both processes: Birth-death and death-Birth, and by this solves Problem 1. This is the first such graph that amplifies the advantage of mutants. The graph is an amplifier only for small mutant advantage $r$, however, it is known that no graph exists that can amplify the advantage for mutants in death-Birth updating.

**Complexity of spatial games.** Chapter 3 examines the spatial games with synchronous updating. In spatial games with synchronous updating, which represent the limit of strong selection, in one time step, every individual adopts the strategy of the neighbor with the highest payoff. The chapter presents a proof that the problem of approximating the fixation probability of cooperators is PSPACE-complete, which solves Problem 2. The main contribution is the construction of a graph and configuration of cooperators and defectors such that the cooperators fixate only if a given Turing machine and configuration of the Turing machine that is accepted. Moreover, the construction is a very simple graph. It is a subset of a grid, which means that the complexity result holds for many restricted classes of graphs.

**Amplifiers for spatial games.** The spatial games with synchronous or asynchronous updating are important models of the spread of cooperation. Chapter 4 presents the construction of the first structure that is an amplifier, that means: first, one randomply placed cooperator in a graph full of defectors is more likely to spread over the whole graph; and second, the cooperators initialized with some positive density in the graph full of defectors spread over the whole graph with high probability. The most important contribution of the chapter is the construction of such a graph, which is the first graph that amplifies the advantage for small differences between the temptation $T$ and payoff $P$. This chapter answers Problem 3 in the

affirmative. Moreover, the work also presents the theoretical analysis that can be used to argue about the amplification property for other graphs.

**Coexistence in heterogeneous networks.** Chapter 5 examines the fixation times (or coexistence times) for a Birth-death process on a complete graph where in one half of the vertices, one of the types has an advantage and in the other half, the other type has an advantage. The work finds the threshold for the interaction between two halves of the graph, such that the fixation time is either polynomial or exponential. The main contribution is the analytical results for the process on heterogeneous networks. This chapter describes the parameters for the coexistence and thus solves Problem 4 Moreover, the work also contains an important lemma about so-called *potent* Markov chains. These are Markov chains where a special potential can express their change. The lemma bounds the number of steps in these Markov chains until the absorbing state.

**Signed graphs.** Chapter 6 examines the dynamic inspired by physics and social dynamics, where triads of individuals are trying to minimize the energy between them by creating balanced triangles. The dynamic happens on a complete graph where every edge has a label $+$ or $-$, and a triangle is balanced when it contains an even number of $-$ signs. The bounds on the convergence of local dynamics, where one random unbalanced triangle is selected and a sign of one edge is flipped, are an open question. The main contribution of the work is the new dynamic, together with the bounds on the number of steps, which is the foundation to answer Problem 5.

CHAPTER 2

# Amplifiers for both Birth-death and death-Birth updating

This chapter appears in full in [SJTC24].

## Abstract

Populations evolve by accumulating advantageous mutations. Every population has some spatial structure that can be modeled by an underlying network. The network then influences the probability that new advantageous mutations fixate. Amplifiers of selection are networks that increase the fixation probability of advantageous mutants, as compared to the unstructured fully-connected network. Whether or not a network is an amplifier depends on the choice of the random process that governs the evolutionary dynamics. Two popular choices are Moran process with Birth-death updating and Moran process with death-Birth updating. Interestingly, while some networks are amplifiers under Birth-death updating and other networks are amplifiers under death-Birth updating, so far no spatial structures have been found that function as an amplifier under both types of updating simultaneously. In this work, we identify networks that act as amplifiers of selection under both versions of the Moran process. The amplifiers are robust, modular, and increase fixation probability for any mutant fitness advantage in a range $r \in (1, 1.2)$. To complement this positive result, we also prove that for certain quantities closely related to fixation probability, it is impossible to improve them simultaneously for both versions of the Moran process. Together, our results highlight how the two versions of the Moran process differ and what they have in common.

### Author summary

The long-term fate of an evolving population depends on its spatial structure. Amplifiers of selection are spatial structures that enhance the probability that a new advantageous mutation propagates through the whole population, as opposed to going extinct. Many amplifiers of selection are known when the population evolves according to the Moran Birth-death updating, and several amplifiers are known for the Moran death-Birth updating. Interestingly, none of the spatial structures that work for one updating seem to work for the other one. Nevertheless, in this work we identify spatial structures that function as amplifiers of selection for both types

of updating. We also prove two negative results that suggest that stumbling upon such spatial structures by pure chance is unlikely.

## 2.1 Introduction

Moran process is a classic stochastic process that models natural selection in populations of asexually reproducing individuals, especially when new mutations are rare [Mor58, Ewe04]. It is commonly used to understand the fate of a single new mutant, as it attempts to invade a population of indistinguishable residents. Eventually, the new mutation will either fixate on the whole population, or it will go extinct. It is known that when the invading mutant has relative fitness advantage $r > 1$ as compared to the residents, this fixation probability tends to a positive constant $1 - 1/r$ as the population size $N$ grows large.

On spatially structured populations, fixation probability of an invading mutant can both increase or decrease. In the framework of evolutionary graph theory [LHN05, Now06a], the spatial structure is represented by a graph (network) in which nodes (vertices) correspond to individual sites, and edges (connections) correspond to possible migration patterns. Each edge is assigned a weight that represents the strength of the connection. Such network-based spatial structures can represent island models, metapopulations, lattices, as well as other arbitrarily complex structures [YT21, MLB21, STKC23, YST23, TKCN23]. Spatial structures that increase the fixation probability of a randomly occurring advantageous mutant beyond the constant $1 - 1/r$ are called amplifiers of selection [ACN15]. The logic behind the name is that living on such a structure effectively amplifies the fitness advantage that the mutants has, as compared to living on the unstructured (well-mixed) population. Identifying amplifiers is desirable, since they could potentially serve as tools in accelerating the evolutionary search, especially when new mutations are rare [FRT13, TPCN19].

When run on a spatial structure, Moran process can be implemented in two distinct versions. They are called Moran Birth-death process and Moran death-Birth process. In the Moran Birth-death process, first an individual is selected for reproduction with probability proportional to its fitness, and the offspring then replaces a random neighbor. In contrast, in the Moran death-Birth process, first a random individual dies and then its neighbors compete to fill up the vacant site (see Figure 2.1). Both the Moran Bd-updating [Mor58, LHN05, BKP$^+$22] and the Moran dB-updating [CS73, Kom06, ALC$^+$17, Ric21] have been studied extensively. While essentially identical on the unstructured population, the two versions of the process yield different results when run on most spatial structures [ARS06, BR08, KKK15].

In the world of the Bd-updating, amplifiers are ubiquitous [HT15, PTCN18, TPCN19, MHT19, PTCN17]. Almost all small spatial structures function as amplifiers of selection [HT15]. A prime example of an amplifier under the Bd-updating is the Star graph, which improves the mutant fixation probability to roughly $1 - 1/r^2$ [BR08, HBR11, MGP14, Cha14]. In particular, when $r = 1 + \varepsilon$, this is approximately a two-fold increase over the baseline value $1 - 1/r$ given by the unstructured population. Moreover, certain large spatial structures function as so-called superamplifiers, that is, they increase the mutant fixation probability arbitrarily close to 1, even when the mutant has only negligible fitness advantage $r = 1 + \varepsilon$ [GGG$^+$17]. Many other superamplifiers are known, including Incubators [GLL$^+$19], or Selection Reactors [TPCN21].

In contrast, in the world of dB-updating, only a handful of amplifiers are known [Ric23]. Perhaps the most prominent examples are the Fan graphs (see Fig. 2.2) that increase the fixation probability of near-neutral mutants by a factor of up to $1.5$ [ASJ$^+$20]. Interestingly,

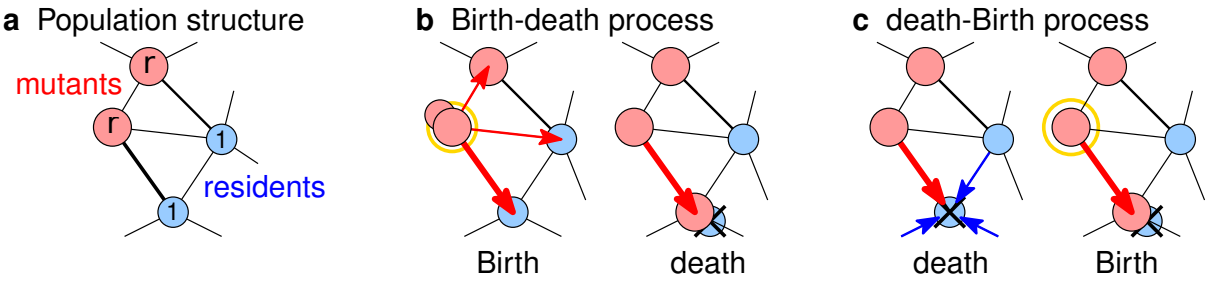



Figure 2.1: **Moran Birth-death and death-Birth processes on a population structure. a,** Each node is occupied by a resident with fitness 1 (blue), or a mutant with fitness $r \geq 1$ (red). Thicker edges denote higher edge weights (stronger interactions). **b,** In Moran Birth-death process, a random individual reproduces, and the produced offspring migrates along a random edge. **c,** In Moran death-Birth process, a random individual dies, and the vacancy is filled by a random neighbor. In both cases, edges with higher weight are selected more often, and fitness plays a role in the Birth step but not in the death step.

all dB-amplifiers are necessarily transient, meaning that the provided amplification effect disappears when the mutant fitness advantage exceeds a certain threshold [TPCN20]. In particular, large Fan graphs increase the fixation probability of the invading mutants for $r \in (1, \varphi)$, where $\varphi \approx 1.618$ is the golden ratio, but decrease it when $r > \varphi$ [ASJ+20].

Unfortunately, the Fan graphs do not function as amplifiers when we instead consider them under Bd-updating (see Fig. 2.2). This is unexpected, since amplification in the Bd-world is so pervasive. And it begs a question. Do there exist spatial structures that function as amplifiers both under the Bd-updating and under the dB-updating? That is, do there exist structures for which the amplification effect is robust with respect to the seemingly arbitrary choice of which version of the Moran process we decide to run?

In this work, we first show three negative results that indicate that the requirements for Bd-amplification and dB-amplification are often conflicting. First, we show that known amplifiers of selection under the Bd-updating are suppressors of selection for the dB-updating and vice versa. Second, we prove that simultaneous Bd- and dB-amplification is impossible under neutral drift ($r = 1$) when the initial mutant location is fixed to a specific starting node. Third, we define a quantity that corresponds to the probability of "mutants going extinct immediately". We then prove that, roughly speaking, no graph improves this quantity as compared to the complete graph under both Bd- and dB-updating. Thus, improving fixation probability under both Bd- and dB-udpating as compared to the complete graph might seem unlikely.

Despite those negative results, we identify a class of population structures that function as amplifiers of selection under both Birth-death and death-Birth updating, for any mutation that grants a relative fitness advantage $r \in (1, 1.2)$. We also present numerical computation that illustrates that the amplification strength is substantial.

## 2.2 Model

Here we formally introduce the terms and notation that we use later, such as the evolutionary dynamics of Moran Birth-death and Moran death-Birth process, the fixation probability, and the notion of an amplifier.

### 2.2.1 Population structure

The spatial structure of the population is represented as a graph (network), denoted $G_N = (V, E)$, where $V$ is a set of $N$ nodes (vertices) of $G_N$ that represent individual sites, and $E$ is a set of edges (connections) that represent possible migration patterns for the offspring. The edges are undirected (two-way) and may be weighted to distinguish stronger interactions from the weaker ones, see Fig. 2.1a. The weight of an edge between nodes $u$ and $v$ is denoted $w(u, v)$. If all edge weights are equal to 1 we say that the graph is *unweighted*. At any given time, each site is occupied by a single individual, who is either a resident with fitness 1, or a mutant with fitness $r \geq 1$. The fitness of an individual at node $u$ is denoted $f(u)$.

### 2.2.2 Moran process

Moran process is a classic discrete-time stochastic process that models the evolutionary dynamics of selection in a population of asexually reproducing individuals. Initially, each node is occupied either by a resident or by a mutant. As long as both mutants and residents co-exist in the population, we perform discrete time steps that change the state of (at most) one node at a time.

There are two versions of the Moran process (see Fig. 2.1). In the *Moran Birth-death process*, we first select an individual to reproduce (randomly, proportionally to the fitness of the individual), and then the offspring migrates along one adjacent edge (randomly, proportionally to the weight of that edge) to replace the neighbor. Formally, denoting by $F = \sum_u f(u)$ the total fitness of the population, node $u$ gets selected for reproduction with probability $f(u)/F$, and then it replaces a neighbor $v$ with probability $p_{u\to v} = w(u, v)/\sum_{v'} w(u, v')$.

In contrast, in the *Moran death-Birth process*, we first select an individual to die (uniformly at random), and then the neighbors compete to fill in the vacancy (randomly, proportionally to the edge weight and the fitness of the neighbor). Formally, node $v$ dies with probability $1/N$ and it gets replaced by a node $u$ with probability $p_{u\to v} = f(u) \cdot w(u, v)/(\sum_{u'} f(u') \cdot w(u', v))$. We note that in both versions we capitalize the word "Birth" to signify that fitness plays a role in the birth step (and not in the death step).

### 2.2.3 Fixation probability and Amplifiers

If the graph $G_N$ that represents the population structure is connected then the Moran process eventually reaches a "homogeneous state", where either all nodes are occupied by mutants (we say that mutants *fixated*), or all nodes are occupied by residents (we say that mutants *went extinct*). Given a graph $G_N$, a mutant fitness advantage $r \geq 1$, and a set $S \subseteq V$ of nodes initially occupied by mutants, we denote by $\rho_r^{\mathrm{Bd}}(G_N, S)$ the *fixation probability*, that is, the probability that mutants eventually reach fixation, under Moran Birth-death process. We are particularly interested in the fixation probability of a single mutant who appears at a node selected uniformly at random. We denote this *fixation probability under uniform initialization* by $\rho_r^{\mathrm{Bd}}(G_N) = \frac{1}{N}\sum_{v\in V} \rho_r^{\mathrm{Bd}}(G_N, \{v\})$. We define $\rho_r^{\mathrm{dB}}(G_N, S)$ and $\rho_r^{\mathrm{dB}}(G_N)$ analogously.

In this work we focus on population structures that increase the fixation probability of invading mutants. The base case is given by an unweighted complete graph $K_N$ that includes all edges and represents an unstructured, well-mixed population. It is known [Now06a, HT15, KKK15] that

$$\rho_r^{\mathrm{Bd}}(K_N) = \frac{1-\frac{1}{r}}{1-\frac{1}{r^N}} \quad \text{and} \quad \rho_r^{\mathrm{dB}}(K_N) = \frac{N-1}{N} \cdot \frac{1-\frac{1}{r}}{1-\frac{1}{r^{N-1}}}.$$

Given a graph $G_N$ and a mutant fitness advantage $r \geq 1$, we say that $G_N$ is a *Bd$_r$-amplifier* if $\rho_r^{\mathrm{Bd}}(G_N) > \rho_r^{\mathrm{Bd}}(K_N)$. We define dB$_r$ amplifiers analogously, that is, as those graphs $G_N$ that satisfy $\rho_r^{\mathrm{dB}}(G_N) > \rho_r^{\mathrm{dB}}(K_N)$. Similarly, *suppressors* are graphs that decrease the fixation probability as compared to the complete graph.

## 2.3 Results

First, we present three negative results that illustrate that the two worlds of Birth-death and death-Birth updating often present contradictory requirements when it comes to enhancing the fixation probability of a single newly occurring mutant. Nevertheless, as our main contribution in the positive direction, we then present population structures that are both Bd$_r$-amplifiers and dB$_r$-amplifiers for a range of mutant fitness advantages $r \in (1, 1.2)$.

### 2.3.1 Negative results

In this section, we present results that suggest that finding simultaneous Bd$_r$- and dB$_r$-amplifiers is not easy. First, we show empirically that known amplifiers for one process are suppressors for the other process. Second, we show that in the neutral regime ($r = 1$), any fixed vertex is a "good" starting vertex for the mutant in at most one of the two processes. Finally, we show that for any starting vertex, the chance of not dying immediately can be enhanced in at most one of the two processes (see below for details).

#### Known amplifiers for one process

In this section we examine spatial structures that are known to amplify under one of the two versions of the Moran process, in order to see whether they amplify under the other version of the Moran process.

First, we consider the smallest known unweighted dB-amplifier [Ric21], which is a certain graph on $N = 11$ nodes (see Fig. 2.2). We call the graph $D_{11}$. The graph $D_{11}$ is an extremely weak dB$_r$-amplifier in a range of approximately $r \in (1, 1.00075)$, where it increases the fixation probability by a factor less than $1.0000001\times$ (see [Ric21, Fig 1]). For $r \in (1.01, 1.1)$ the graph $D_{11}$ appears to function as a very slight suppressor under both dB-updating and Bd-updating. In particular, at $r = 1.1$ we obtain $\rho_r^{\mathrm{Bd}}(D_{11})/\rho_r^{\mathrm{Bd}}(K_{11}) \doteq 0.996$ and $\rho_r^{\mathrm{dB}}(D_{11})/\rho_r^{\mathrm{dB}}(K_{11}) \doteq 0.997$.

Next, we examine the star graph $S_{11}$ on 11 vertices which, to our knowledge, is the strongest unweighted amplifier for Bd-updating at this population size. The Star graph is a clear Bd$_r$-amplifier for $r \in (1.01, 1.1)$, but an equally clear dB$_r$-suppressor in that range.

The situation is reversed for the Fan graph $F_{11}$ [ASJ+20]. While the Fan graph clearly functions as an amplifier under the dB-updating when $r \in (1.01, 1.1)$, it lags behind the baseline given by the complete graph under the Bd-updating.

#### Neutral regime ($r = 1$)

The second negative result pertains to the case of neutral mutations ($r = 1$). Recall that $\rho_r^{\mathrm{Bd}}(G_N, v)$ and $\rho_r^{\mathrm{dB}}(G_N, v)$ denote the fixation probabilities when the initial mutant appears at node $v$. The following theorem states that for neutral mutations ($r = 1$), no initial mutant node increases the fixation probability both for Birth-death and death-Birth updating.

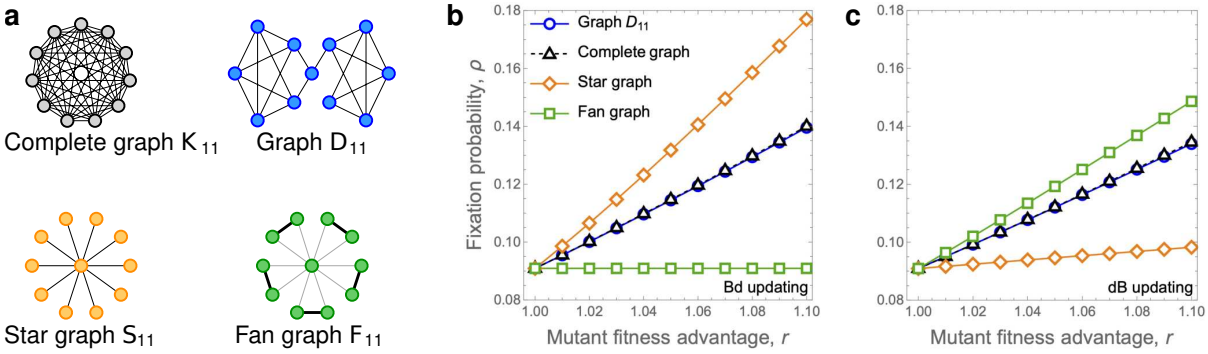



Figure 2.2: **Known amplifiers are suppressors for the other process. a,** We consider four graphs on $N = 11$ nodes, namely the Complete graph $K_{11}$, the star graph $S_{11}$, the Fan graph $F_{11}$, and the smallest known undirected amplifier $D_{11}$ (see [Ric21]). **b,** Under Bd-updating, the only amplifier for $r \in \{1.01, \dots, 1.1\}$ is the Star graph $S_{11}$. **c,** Under dB-updating, the only amplifier for $r \in \{1.01, \dots, 1.1\}$ is the Fan graph $F_{11}$. Values computed by numerically solving the underlying Markov chains.

**Theorem 1.** *Let $G_N$ be a graph and $v$ an initial mutant node. Then at least one of the following is true:*

1. $\rho^{\mathrm{Bd}}_{r=1}(G_N, v) < \rho^{\mathrm{Bd}}_{r=1}(K_N)$; *or*
2. $\rho^{\mathrm{dB}}_{r=1}(G_N, v) < \rho^{\mathrm{dB}}_{r=1}(K_N)$; *or*
3. $\rho^{\mathrm{Bd}}_{r=1}(G_N, v) = \rho^{\mathrm{Bd}}_{r=1}(K_N)$ *and* $\rho^{\mathrm{dB}}_{r=1}(G_N, v) = \rho^{\mathrm{dB}}_{r=1}(K_N)$.

The idea behind the proof is that for neutral evolution there are explicit formulas for fixation probabilities $\rho^{\mathrm{Bd}}_r(G_N, v)$ and $\rho^{\mathrm{dB}}_r(G_N, v)$ on any undirected graph $G_N$ [BHRS10, Mac14]. The result then follows by applying Cauchy-Schwarz inequality. See Section 2.5 for details. In Section 2.5, we also note that Theorem 1 does not generalize to the case when instead of having one initial mutant node we start with an initial subset $S$ of $k \geq 2$ nodes occupied by mutants.

### Immediate extinction and forward bias

In order to present our third and final negative result, we need to introduce additional notions and notation. When tracking the evolutionary dynamics on a given graph $G_N$ with a given mutant fitness advantage $r \geq 1$, it is often useful to disregard the exact configuration of which nodes are currently occupied by mutants, and only look at *how many* nodes are occupied by mutants.

One example of this is the celebrated Isothermal theorem [LHN05] which states that once $N$ and $r$ are fixed, the fixation probability under the Moran Birth-death process on any regular graph is the same. Here, a graph is *regular* if each node has the same total weight of adjacent edges. Examples of regular graphs include the complete graph, the cycle graph, or any grid graph with periodic boundary condition.

The intuition behind the proof of the Isothermal theorem is that for any regular graph $R_N$, the Moran Birth-death process can be mapped to a random walk that tracks just the number of mutants, instead of their exact positions on the graph. It can be shown that this random walk has a constant forward bias, that is, the probabilities $p^+$ (resp. $p^-$) that the size of the mutant

subpopulation increases (resp. decreases) satisfy $p^+/p^- = r$, for any number of mutants in any particular mutant-resident configuration. A natural approach to construct amplifiers is thus to construct graphs for which this forward bias satisfies an inequality $p^+/p^- \geq r$ for the Moran Birth-death process and an analogous inequality for the Moran death-Birth process. Our final negative result shows that this goal can not be achieved already in the first step.

Formally, consider the Moran Birth-death process on a graph $G_N$ with a single initial mutant placed at node $u$. Let $\gamma_r^{\mathrm{Bd}}(G_N, u)$ be the probability that the first reproduction event that changes the size of the mutant subpopulation is the initial mutant reproducing (as opposed to the initial mutant being replaced by one of its neighbors). In other words, $\gamma_r^{\mathrm{Bd}}(G_N, u)$ is the probability that the first step that changes the configuration of the mutants does *not* eliminate the initial mutant, leaving the options of later mutant extinction or mutant fixation.

For the complete graph $K_N$ (and any single mutant node) it is not hard to show that $\gamma_r^{\mathrm{Bd}}(K_N) = \gamma_r^{\mathrm{Bd}}(K_N, u) = r/(r+1)$ for any node $u$. Moreover, by a slight extension of the Isothermal theorem, we have $\gamma_r^{\mathrm{Bd}}(R_N, u) = r/(r+1)$ for any regular graph $R_N$ and any node $u$. For Moran death-Birth process, we define $\gamma_r^{\mathrm{dB}}(G_N, u)$ and $\gamma_r^{\mathrm{dB}}(K_N)$ analogously. To construct a graph that is both a Bd- and a dB-amplifier, a natural approach is to look for a graph and an initial mutant node $u$ such that $\gamma_r^{\mathrm{Bd}}(G_N, u) > \gamma_r^{\mathrm{Bd}}(K_N)$ and $\gamma_r^{\mathrm{dB}}(G_N, u) > \gamma_r^{\mathrm{dB}}(K_N)$. However, the following theorem states that no such graphs exist.

**Theorem 2.** *Let $G_N$ be a graph, $u$ an initial mutant node, and $r \geq 1$. Then at least one of the following is true:*

1. $\gamma_r^{\mathrm{Bd}}(G_N, u) < \gamma_r^{\mathrm{Bd}}(K_N)$*; or*
2. $\gamma_r^{\mathrm{dB}}(G_N, u) < \gamma_r^{\mathrm{dB}}(K_N)$*; or*
3. $\gamma_r^{\mathrm{Bd}}(G_N, u) = \gamma_r^{\mathrm{Bd}}(K_N)$ *and* $\gamma_r^{\mathrm{dB}}(G_N, u) = \gamma_r^{\mathrm{dB}}(K_N)$.

The proof relies on the notion of the temperature of a node. Formally, given a graph $G_N = (V, E)$ we first define a (weighted) *degree* of a node $v$ as $\deg(v) = \sum_{v':(v,v')\in E} w(v, v')$. Then, given a node $u$, we define its *temperature* $T(u)$ as

$$T(u) = \sum_{v:\ (v,u)\in E} \frac{w(v,u)}{\deg(v)}.$$

The temperature of a node represents the rate at which the node is being replaced by its neighbors in the Moran Birth-death process when $r = 1$. Nodes with high temperature are replaced often, whereas nodes with low temperature are replaced less frequently. Building on this, it is straightforward to show that if a node $u$ has above-average temperature, then $\gamma_r^{\mathrm{Bd}}(G_N, u) < \gamma_r^{\mathrm{Bd}}(K_N)$, that is, in Moran Birth-death process with a single mutant at $u$ the forward bias is lower than the forward bias on a complete graph. To complete the proof, we then show that for any node $u$ with below-average temperature, we have $\gamma_r^{\mathrm{dB}}(G_N, u) < \gamma_r^{\mathrm{dB}}(K_N)$. Our proof of the latter claim uses Jensen's inequality for a certain concave function. See Section 2.5 for details.

### 2.3.2 Positive result

Despite the above negative results, in this section we identify population structures $A_N$ that substantially amplify the fixation probability under both Birth-death updating and death-Birth updating when the number $N$ of nodes is sufficiently large.

The structures $A_N$ are composed of two large chunks $A^{\mathrm{Bd}}$ and $A^{\mathrm{dB}}$ that are connected by a single edge, see Fig. 2.3a for an illustration. The chunk $A^{\mathrm{dB}}$ is a Fan graph [ASJ+20], which is to our knowledge the strongest currently known dB-amplifier. The chunk $A^{\mathrm{Bd}}$ could be any of the many strong Bd-amplifiers. For definiteness, in Fig. 2.3a we use a Fan-like structure with $a$ nodes in a central hub and $b$ blades of two nodes each surrounding it. The single connecting edge has a very low edge weight so that the two chunks interact only rarely. For population size $N = 1001$, the resulting weighted graph is both a Bd$_r$-amplifier and a dB$_r$-amplifier for any $r \in (1, 1.09)$, see Fig. 2.3b.

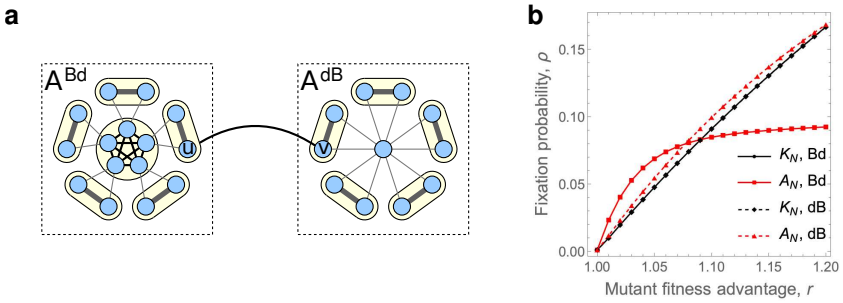



Figure 2.3: **Simultaneous Bd- and dB-amplifier $A_N$. a,** The graph $A_N$ is composed of two large chunks $A^{\mathrm{Bd}}$ and $A^{\mathrm{dB}}$ that are connected by a single edge. The chunk $A^{\mathrm{dB}}$ is a Fan graph with $f$ nodes. The chunk $A^{\mathrm{Bd}}$ is a fan-like graph with $a$ vertices in a central hub and $b$ blades of two nodes each. The total population size is $N = a + 2b + f$ (here $a = b = 5$, $f = 11$, and $N = 26$). The edge weights are defined such that different circled units within the chunks interact only rarely, and the chunks themselves interact even more rarely. **b,** Here we consider graph $A_N$ with population size $N = 1001$ and $(a, b, f) = (30, 85, 801)$. The fixation probabilities under Bd- and dB-updating are computed by numerically solving the underlying Markov chain. We find that the inequality $\rho_r^{\mathrm{Bd}}(A_N) > \rho_r^{\mathrm{Bd}}(K_N)$ is satisfied for $r \in (1, 1.09)$ and the inequality $\rho_r^{\mathrm{dB}}(A_N) > \rho_r^{\mathrm{dB}}(K_N)$ is satisfied for $r \in (1, 1.2)$. In particular, at $r = 1.05$ the ratios satisfy $\rho_r^{\mathrm{Bd}}(A_N)/\rho_r^{\mathrm{Bd}}(K_N) > 1.44$ and $\rho_r^{\mathrm{dB}}(A_N)/\rho_r^{\mathrm{dB}}(K_N) > 1.14$.

Similarly, we identify large population structures that serve as both Bd$_r$-amplifiers and dB$_r$-amplifiers for any $r \in (1, 1.2)$.

**Theorem 3** (Simultaneous Bd- and dB-amplifier)**.** *For every large enough population size $N$ there exists a graph $A_N$ such that for all $r \in (1, 1.2)$ we have*

$$\rho_r^{\mathrm{Bd}}(A_N) > \rho_r^{\mathrm{Bd}}(K_N) \quad \textit{and} \quad \rho_r^{\mathrm{dB}}(A_N) > \rho_r^{\mathrm{dB}}(K_N).$$

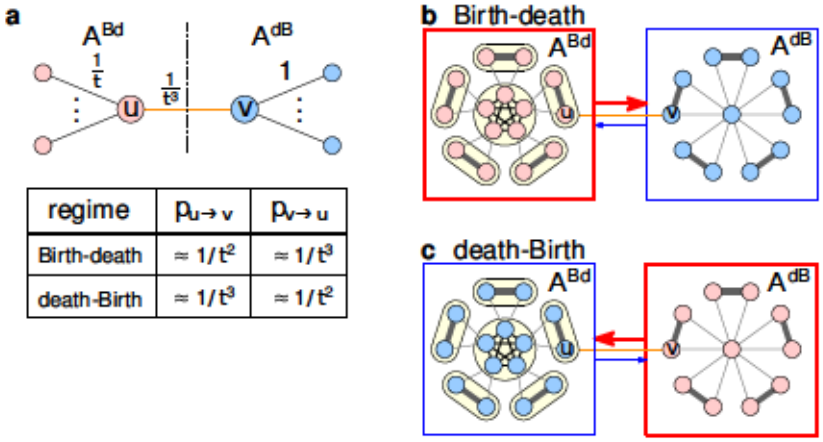
In what follows we provide intuition about the proof of Theorem 3.The fully rigorous proof is relegated to Section 2.5. Let $e$ be the edge connecting the two chunks, $u$ its endpoint in $A^{\mathrm{Bd}}$, and $v$ its endpoint in $A^{\mathrm{dB}}$.

First, observe that since $e$ has a low weight, the two chunks evolve mostly independently. This means that, with high probability, each chunk resolves to a homogeneous state in between any two interactions across the chunks. In particular, if the initial mutant appears in the chunk where it is favored (e.g. if it appears in the chunk $A^{\mathrm{Bd}}$ when Bd-updating is run), the mutants fixate on that chunk with reasonable probability. If that occurs, we say that mutants are "half done".

Once the mutants are half done, the next relevant step occurs when the two chunks interact. There are two cases. Either a mutant at $u$ reproduces and the offspring migrates along $e$ to $v$,

or a resident at $v$ reproduces and the offspring migrates along $e$ to replace the mutant at $u$. In both cases, the individual (mutant or resident) who "invades" the other half eventually either succeeds in spreading through that half, or they fail at doing that. If the latter occurs, we are back at the situation in which mutants are half done and the situation repeats. By bounding all the relevant probabilities, we show that once half done, mutants are overwhelmingly likely to fixate, as opposed to going extinct.

We highlight an interesting phenomenon that occurs in our proof. As we run the evolutionary dynamics, we can look at the flow along the connecting edge $e$. Thanks to the edge weights, it turns out that the direction of the flow along $e$ flips depending on whether we run the Moran Birth-death process or the Moran death-Birth process. In particular, under the Bd-updating the edge $e$ is used mostly in the direction from $u$ to $v$. That is, many individuals migrate from $u$ to $v$, whereas few individuals migrate from $v$ to $u$. Under dB-updating the situation reverses. That is, many individuals migrate from $v$ to $u$, whereas few of them migrate from $u$ to $v$. Thus, under the Bd-updating the $A^{\mathrm{Bd}}$ chunk is effectively upstream of the chunk $A^{\mathrm{dB}}$, whereas under the dB-updating the $A^{\mathrm{dB}}$ chunk is effectively upstream of the chunk $A^{\mathrm{Bd}}$. This asymmetry is a key factor that contributes to the fact that once the mutants are half done, they are likely to fixate on the whole graph (see Fig. 2.4).




| regime | $p_{u\to v}$ | $p_{v\to u}$ |
|---|---|---|
| Birth-death | $\approx 1/t^2$ | $\approx 1/t^3$ |
| death-Birth | $\approx 1/t^3$ | $\approx 1/t^2$ |

Figure 2.4: **Interactions between $A^{\mathrm{Bd}}$ and $A^{\mathrm{dB}}$. a,** The edge weights in the chunks $A^{\mathrm{Bd}}$ (red) and $A^{\mathrm{dB}}$ (blue) are shown as a function of $t$ (here $t \gg 1$ is large). The connecting edge has weight $1/t^3$, all other edges with endpoint $u$ have total weight $1/t$ and all other edges with endpoint $v$ have total weight 1. For each of two versions of the Moran process, the rates at which the offspring migrate from $u$ to $v$ and from $v$ to $u$ can be calculated and are listed in the table. **b,** Under Birth-death updating, the migration rate $p_{u\to v}$ from $u$ to $v$ is roughly $t\times$ larger than the migration rate $p_{v\to u}$ from $v$ to $u$, so the chunk $A^{\mathrm{Bd}}$ is upstream of the chunk $A^{\mathrm{dB}}$, and a mutant who has fixated over $A^{\mathrm{dB}}$ is likely to fixate over $A^{\mathrm{Bd}}$ too. **c,** In contrast, under death-Birth updating we have $p_{v\to u} \approx t \cdot p_{u\to v}$, hence the chunk $A^{\mathrm{dB}}$ is upstream of $A^{\mathrm{Bd}}$.

What remains in the proof is to balance out the sizes of the two chunks. For small $r > 1$, the strongest known dB-amplifiers are roughly $\frac{3}{2}\times$ stronger than the Complete graph (in terms of the fixation probability). Thus, in order to achieve amplification under dB-updating, we need the chunk $A^{\mathrm{dB}}$ to take up at least $2/3$ of the total population size. The chunk $A^{\mathrm{Bd}}$ then takes up at most $1/3$ of the total population size. In order to achieve Bd-amplification, fixation probability on $A^{\mathrm{Bd}}$ under Bd-updating must therefore be at least $3\times$ larger than that on the Complete graph. Interestingly, a Star graph is not strong enough to do that (for $r \approx 1$

and large population size $N$ it is only roughly $2\times$ stronger than the Complete graph), but sufficiently strong Bd-amplifiers do exist (e.g. any superamplifier).

## 2.4 Discussion

Population structure has a profound impact on the outcomes of evolutionary processes and, in particular, on the probability that a novel mutation achieves fixation [DL94a, LHN05]. Population structures that increase the fixation probability of beneficial mutants, when compared to the case of a well-mixed population, are known as amplifiers of selection.

Somewhat surprisingly, to tell whether a specific spatial structure is an amplifier or not, one needs to specify seemingly minor details of the evolutionary dynamics. The well-studied Moran process comes in two versions, namely Moran process with Birth-death updating and Moran process with death-Birth updating. While many spatial structures are amplifiers under the Bd-updating [HT15], only a handful of amplifiers under the dB-updating are known [Ric23]. Moreover, none of the dB-amplifiers that we checked amplify under the Bd-updating.

In this work we help explain this phenomenon by proving mathematical results which illustrate that the two objectives of amplifying under the Bd-updating and amplifying under the dB-updating are often contradictory. Thus, one might be tempted to conclude that perhaps there are no population structures that amplify in both worlds, that is, regardless of the choice of the underlying dynamics (Bd or dB). Nevertheless, we proceed to identify population structures that serve as amplifiers of selection under both Bd-updating and dB-updating.

The amplifiers we identify in this work have several interesting features. First, they are robust in the sense that they amplify selection under both the Bd-updating and the dB-updating. Second, they provide amplification for any mutant fitness advantage $r$ in a range $r \in (1, 1.2)$, which covers many realistic values of the mutant fitness advantage, and the amplification is non-negligible (for instance, for $r = 1.05$ the fixation probability increases by 14% and 44%, respectively. see Fig. 2.3). Third, the amplifiers are modular. That is, they consist of two large chunks that serve as building blocks and that interact rarely. For definiteness, in this work we specified the two chunks and their relative sizes, but each chunk can be replaced by an alternative building block and the relative sizes can be altered. For example, the best currently known dB-amplifiers amplify by a factor of $1.5\times$ for $r \approx 1$ and continue to amplify for $r$ in a range $r \in (1, \varphi)$, where $\varphi = \frac{1}{2}(\sqrt{5}+1) \approx 1.618$ is the golden ratio [ASJ+20]. If better dB-amplifiers are found, they can be used as a building block in place of one of the chunks to improve the range $r \in (1, 1.2)$ for which the resulting structure amplifies in both worlds.

In this work, our objective was to increase the fixation probability of an invading mutant in both worlds (Bd-updating and dB-updating). An interesting direction for future work is to optimize other quantities in both worlds.

One such quantity is the duration of the process until fixation occurs [DGRS16, MvS20, MvS21]. For example, achieving short fixation times in combination with increasing the fixation probability does not appear to be easy. Our proofs rely on the existence of small edge weights to separate the time scales at which different stages of the process happen. While using more uniform edge weights might still lead to the same outcome, the proofs would need to become more delicate. A possible approach to identify structures that serve as fast amplifiers in both worlds would be to find unweighted amplifiers, because then the time would be guaranteed to be at most polynomial [DGM+14, DKPT22]. The first step in this direction would be to

identify large and substantially strong unweighted dB-amplifiers. There are promising recent results in this direction [Ric23].

Looking beyond fixation time, there are other relevant quantities such as the recently introduced rate at which beneficial mutations accumulate [ST22]. Existing research suggests that the two versions of the Moran process behave quite differently in terms of the fixation probability [HT15], but quite similarly in terms of the fixation time [DGM$^+$14, DKPT22]. Which of those two cases occurs for other relevant quantities remains to be seen.

### Data and code availability

Code for the figures and the computational experiments is available from the Figshare repository: `https://figshare.com/s/4e08d78c892749f84201`.

## 2.5 Additional proofs

### 2.5.1 Preliminaries

Given an undirected graph $G_N = (V, E)$ on $N$ nodes, the *degree* of a node $u$, denoted $\deg(u)$, is the number of neighbors of $u$ in $G_N$. When the edges are weighted, we define the degree $\deg(u) = \sum_{v:(u,v)\in E} w(u, v)$ as the sum of the weights of all the adjacent edges. As a direct extension of [ARS06, BHRS10, Mac14] and as noted in [AM19] we obtain the following formula for fixation probability under neutral drift ($r = 1$). For completeness, we include a proof.

**Lemma 1** (Fixation probability on edge-weighted undirected graphs when $r = 1$)**.** *Let $G_N = (V, E)$ be an edge-weighted undirected graph on $N$ nodes and $S \subset V$ any set of vertices occupied by mutants. Then*

$$\rho^{\mathrm{Bd}}_{r=1}(G_N, S) = \frac{\sum_{u\in S} 1/\deg(u)}{\sum_{v\in V} 1/\deg(v)} \qquad \textit{and} \qquad \rho^{\mathrm{dB}}_{r=1}(G_N, S) = \frac{\sum_{u\in S} \deg(u)}{\sum_{v\in V} \deg(v)}.$$

*Proof.* Let $p_{u\to v}$ be the probability that, in a single step, an individual at node $u$ produces an offspring that replaces an individual at node $v$. For Birth-death updating, it suffices to check that for any subset $S \subset V$ of mutant nodes and any edge $(u, v)$ connecting a mutant node $u \in S$ and a non-mutant node $v \notin S$ we have

$$p_{u\to v} \cdot \frac{1/\deg(v)}{\sum_{v'\in V} 1/\deg(v')} = p_{v\to u} \cdot \frac{1/\deg(u)}{\sum_{v'\in V} 1/\deg(v')}.$$

Since for Birth-death updating and $r = 1$ we have $p_{u\to v} = \frac{1}{N} \cdot \frac{w(u,v)}{\deg(u)}$, both sides rewrite as

$$\frac{\frac{1}{N} \cdot \frac{w(u,v)}{\deg(u)\deg(v)}}{\sum_{v'\in V} 1/\deg(v')},$$

and so the claim is proved. Likewise, for death-Birth updating it suffices to check that

$$p_{u\to v} \cdot \frac{\deg(v)}{\sum_{v'\in V} \deg(v')} = p_{v\to u} \cdot \frac{\deg(u)}{\sum_{v'\in V} \deg(v')}.$$

Since for death-Birth updating and $r = 1$ we have $p_{u\to v} = \frac{1}{N} \cdot \frac{w(u,v)}{\deg(v)}$, this time both sides rewrite as

$$p_{u\to v} \cdot \frac{\frac{w(u,v)}{N}}{\sum_{v'\in V} \deg(v')}.$$

□

The proof of our positive result relies on three existing results. For convenience, we list them here. First, there exist unweighted graphs called *Incubators* that are strong amplifiers under Birth-death updating [GLL+19, Theorem 2].

**Lemma 2.** *There exists a family of graphs $\mathcal{A}_N^{Bd}$ such that for all $r > 1$, we have*

$$\rho_r^{\mathrm{Bd}}(\mathcal{A}_N^{Bd}) \geq 1 - \mathcal{O}(N^{-1/3}).$$

Second, there exist edge-weighted graphs called *Separated Hubs* that are substantial amplifiers under death-Birth updating [ASJ+20, Theorem 3].

**Lemma 3.** *There exists a family of graphs $\mathcal{A}_N^{dB}$ such that for all $r > 1$, we have*

$$\rho_r^{\mathrm{dB}}(G_N) = \frac{N}{2N+1} \cdot \frac{1 - \frac{1}{r^3}}{1 - \frac{1}{r^{3N}}}.$$

Third, the evolutionary dynamics terminates polynomially quickly in terms of the population size $N$, under both the Birth-death updating [DGM+14, Theorem 9] and the death-Birth updating [DKPT22, Theorem 1].

**Lemma 4.** *Fix $r > 1$. For Bd and dB process on an undirected graph with $N$ vertices with the highest ratio between edge weights $\frac{1}{\varepsilon}$, the expected fixation time is in $\mathcal{O}(\frac{N^4}{\varepsilon})$.*

### 2.5.2 Negative result 2

In this section, we show that one fixed neutral mutant cannot have a better fixation probability in both processes than on a complete graph. This means that even if we can choose the starting position, we are not guaranteed to increase the fixation probability for both processes.

**Theorem 4.** *Let $G_N$ be a graph and $v$ an initial mutant node. Then at least one of the following is true:*

1. $\rho_{r=1}^{\mathrm{Bd}}(G_N, v) < \rho_{r=1}^{\mathrm{Bd}}(K_N)$;
2. $\rho_{r=1}^{\mathrm{dB}}(G_N, v) < \rho_{r=1}^{\mathrm{dB}}(K_N)$;
3. $\rho_{r=1}^{\mathrm{Bd}}(G_N, v) = \rho_{r=1}^{\mathrm{Bd}}(K_N)$ *and* $\rho_{r=1}^{\mathrm{dB}}(G_N, v) = \rho_{r=1}^{\mathrm{dB}}(K_N)$.

*Proof.* First, note that $\rho_{r=1}^{\mathrm{Bd}}(K_N) = \rho_{r=1}^{\mathrm{dB}}(K_N) = 1/N$. Next, recall the known formulas for the fixation probability on undirected graphs under neutral drift (see Lemma 1 and [BHRS10, Mac14]], namely:

$$\rho_{r=1}^{\mathrm{Bd}}(G_N, v) = \frac{1/\deg(v)}{\sum_{u\in V} 1/\deg(u)} \quad \text{and} \quad \rho_{r=1}^{\mathrm{dB}}(G_N, v) = \frac{\deg(v)}{\sum_{u\in V} \deg(u)}.$$

As the final ingredient, note that for any $N$ non-negative numbers $x_1, \dots, x_N$ we have a bound

$$\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_N}\right) \cdot (x_1 + x_2 + \cdots + x_N) \geq N^2.$$

This follows e.g. from the inequality between the arithmetic and harmonic mean of numbers $x_1, \dots, x_N$ (called AM-HM), or from Cauchy-Schwarz inequality. Moreover, the equality occurs if and only if $x_1 = x_2 = \cdots = x_N$. Applying this bound to $x_i = \deg(v_i)$ we obtain

$$\rho^{\mathrm{Bd}}_{r=1}(G_N, v) \cdot \rho^{\mathrm{dB}}_{r=1}(G_N, v) = \frac{1/\deg(v)}{\sum_{u \in V} 1/\deg(u)} \cdot \frac{\deg(v)}{\sum_{u \in V} \deg(u)} = \frac{1}{(\sum_{u \in V} 1/\deg(u)) \cdot (\sum_{u \in V} \deg(u))} \leq \frac{1}{N^2}.$$

If equalities occur everywhere then $\deg(v_1) = \cdots = \deg(v_N)$, thus $\rho^{\mathrm{Bd}}_{r=1}(G_N, v) = \rho^{\mathrm{dB}}_{r=1}(G_N, v) = 1/N$. Otherwise, the product is strictly less than $1/N^2$, thus at least one of $\rho^{\mathrm{Bd}}_{r=1}(G_N, v)$ and $\rho^{\mathrm{dB}}_{r=1}(G_N, v)$ is strictly less than $1/N$. □

The following example illustrates that there exists a graph and a subset $S = \{u, v\}$ of $k = 2$ nodes, such that the fixation probability starting from mutants at both $u$ and $v$ is strictly greater than fixation probability starting from $k = 2$ mutant nodes on a well-mixed population, both for the Birth-death and for the death-Birth updating.

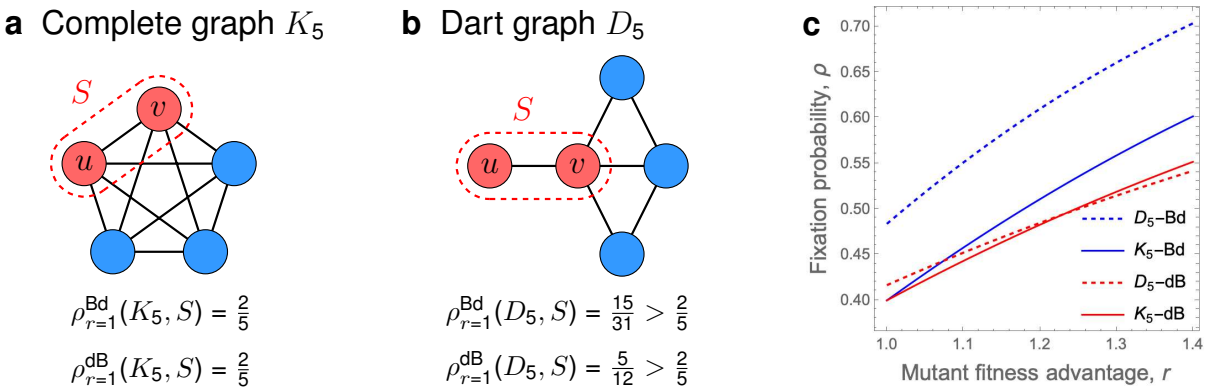



Figure 2.5: **Mutant subset that amplifies for both Bd and dB. a,** With two neutral mutants ($r = 1$) on a complete graph $K_N$, the fixation probability is equal to $2/N$ under both Birth-death and death-Birth updating. **b,** When two neutral mutants initially occupy vertices $u$ and $v$ of the so-called dart graph $D_5$, the fixation probability under both Birth-death and death-Birth updating is increased. **c,** As $r$ increases above roughly $r \approx 1.24$, the fixation probability on the Dart graph under death-Birth updating drops below the reference value of two mutants on a complete graph $K_5$. Under Birth-death updating, the effect persists for $r \geq 1$. (Data obtained by numerically solving the underlying Markov chains.)

The intuition behind the result is that node $u$ is a really good initial mutant node for Birth-death updating, and node $v$ is a really good initial mutant node for death-Birth updating. Together, they form an above-average set of two mutant nodes, even when compared to a complete graph with two initial mutants.

### 2.5.3 Negative result 3

In this section, we prove that for any fixed vertex, in the first step, the ratio between increasing and decreasing the number of mutants cannot be better than in the complete graph in both processes. This means we cannot find a vertex from which both processes spread better than in

the complete graph. To achieve amplification for both processes, we know that some vertices will be better for Bd and some for dB amplification.

Recall that given Moran Birth-death process run on graph $G_N$ with an initial mutant node $u$, the quantity $\gamma_r^{\mathrm{Bd}}(G_N, u)$ is the probability that the first reproduction event that changes the size of the mutant subpopulation is the initial mutant reproducing (as opposed to the initial mutant being replaced by one of its neighbors). Similarly, we define $\gamma_r^{\mathrm{dB}}(G_N, u)$ for the death-Birth process.

**Theorem 5.** *Let $G_N$ be a graph, $u$ an initial mutant node, and $r \geq 1$. Then at least one of the following is true:*

1. $\gamma_r^{\mathrm{Bd}}(G_N, u) < \gamma_r^{\mathrm{Bd}}(K_N)$;
2. $\gamma_r^{\mathrm{dB}}(G_N, u) < \gamma_r^{\mathrm{dB}}(K_N)$;
3. $\gamma_r^{\mathrm{Bd}}(G_N, u) = \gamma_r^{\mathrm{Bd}}(K_N)$ *and* $\gamma_r^{\mathrm{dB}}(G_N, u) = \gamma_r^{\mathrm{dB}}(K_N)$.

*Proof.* Denote by $T(u) = \sum_{v\colon (u,v)\in E} \frac{w(u,v)}{\deg v}$ the so-called *temperature* of node $u$, that is, the rate at which node $u$ is replaced by its neighbors in the neutral case.

Denote by $p^+_{\mathrm{Bd},r} = p^+_{\mathrm{Bd},r}(G_N, u)$ the probability that in a single step of the Moran Birth-death process the mutant reproduces, and by $p^-_{\mathrm{Bd},r} = p^-_{\mathrm{Bd},r}(G_N, u)$ the probability that it gets replaced by a resident. Denoting the total fitness by $F = N + (r-1)$ we have

$$p^+_{\mathrm{Bd},r} = \frac{r}{F} \quad \text{and} \quad p^-_{\mathrm{Bd},r} = \sum_{v\colon (u,v)\in E} \frac{1}{F} \cdot \frac{w(u,v)}{\deg v},$$

and thus

$$\gamma_r^{\mathrm{Bd}}(G_N, u) = \frac{p^+_{\mathrm{Bd},r}}{p^+_{\mathrm{Bd},r} + p^-_{\mathrm{Bd},r}} = \frac{r}{r + \sum_{v\colon (u,v)\in E} \frac{w(u,v)}{\deg v}} = \frac{r}{r + T(u)}.$$

In particular, in the complete graph $K_N$ each node has temperature $1$, and thus

$$\gamma_r^{\mathrm{Bd}}(K_N) = \frac{r}{r+1}.$$

If $T(u) \geq 1$ then $r/(r+T(u)) \leq r/(r+1)$ and hence $\gamma_r^{\mathrm{Bd}}(G_N, u) \leq \gamma_r^{\mathrm{Bd}}(K_N)$ with equality if and only if $T(u) = 1$. From now on, suppose $T(u) \leq 1$.

Consider Moran death-Birth process and define the quantities $p^+_{\mathrm{dB},r} = p^+_{\mathrm{dB},r}(G_N, u)$ and $p^-_{\mathrm{dB},r} = p^-_{\mathrm{dB},r}(G_N, u)$ as above. Then

$$p^+_{\mathrm{dB},r} = \sum_{v\colon (u,v)\in E} \frac{1}{N} \cdot \frac{r \cdot w(u,v)}{(r-1)w(u,v) + \deg(v)} \quad \text{and} \quad p^-_{\mathrm{dB},r} = \frac{1}{N},$$

therefore

$$\gamma_r^{\mathrm{dB}}(G_N, u) = \frac{\sum_{v\colon (u,v)\in E} \frac{r\cdot w(u,v)}{(r-1)w(u,v)+\deg(v)}}{1 + \sum_{v\colon (u,v)\in E} \frac{r\cdot w(u,v)}{(r-1)w(u,v)+\deg(v)}}.$$

In particular, for the complete graph $K_N$ and any its node $u$ we have

$$\sum_{v\colon (u,v)\in E} \frac{r \cdot w(u,v)}{(r-1)w(u,v) + \deg(v)} = (N-1) \cdot \frac{r}{(r-1) + (N-1)}.$$

Hence in order to prove $\gamma_r^{\mathrm{dB}}(G_N, u) \leq \gamma_r^{\mathrm{dB}}(K_N)$, it suffices to prove

$$\sum_{v\colon (u,v)\in E} \frac{r \cdot w(u,v)}{(r-1)w(u,v) + \deg(v)} \leq \frac{(N-1)r}{(r-1)+(N-1)}.$$

We rearrange

$$\sum_{v\colon (u,v)\in E} \frac{\frac{w(u,v)}{\deg(v)}}{(r-1)\frac{w(u,v)}{\deg(v)} + 1} \leq \frac{1}{\frac{r-1}{N-1} + 1}.$$

When $r = 1$, the desired claim reduces precisely to $T(u) \leq 1$. Suppose $r > 1$, that is $r-1 > 0$, and consider a function $f\colon (0,\infty) \to (0,\infty)$ defined by $f(x) = \frac{x}{(r-1)x+1}$. Then $f$ is concave and increasing, therefore by Jensen's inequality we have

$$\sum_{v\colon (u,v)\in E} \frac{\frac{w(u,v)}{\deg(v)}}{(r-1)\frac{w(u,v)}{\deg(v)} + 1} \leq |N(u)| \cdot \frac{\frac{1}{|N(u)|}\sum_{v\colon (u,v)\in E} \frac{w(u,v)}{\deg(v)}}{(r-1)\frac{1}{|N(u)|}\sum_{v\colon (u,v)\in E} \frac{w(u,v)}{\deg(v)} + 1} = \frac{T(u)}{\frac{r-1}{|N(u)|}\cdot T(u) + 1},$$

where $|N(u)| = |\{v\colon (u,v) \in E\}|$ is the number of neighbors of $u$ in $G$.

Finally, since the function $f$ is increasing, using bounds $T(u) \leq 1$ and $|N(u)| \leq N-1$, the right-hand side is at most

$$\frac{T(u)}{\frac{r-1}{|N(u)|}\cdot T(u) + 1} \leq \frac{1}{\frac{r-1}{|N(u)|} + 1} \leq \frac{1}{\frac{r-1}{N-1} + 1}$$

as desired. For the equality to occur in the first step, we must in particular have $T(u) = 1$, in which case the other equality $\gamma_r^{\mathrm{Bd}}(G_N, u) = \gamma_r^{\mathrm{Bd}}(K_N)$ holds too.

□

### 2.5.4 Positive result

In this section, we prove the main positive result which states that there exists an undirected, edge-weighted graph that is simultaneously an amplifier of selection for Birth-death Moran process and for death-Birth Moran process (under uniform mutant initialization). We first bound the number of steps until fixation or extinction for both processes (Bd and dB) and any graph. Second, we show that for any graph, there is a good starting vertex where a mutant has fixation probability at least $\frac{1}{N}$. Then we construct the graph and we prove that it is indeed an amplifier for both processes.

#### Auxiliary statements

**Lemma 5.** *For Bd and dB process for any $r$ on an undirected graph with $N$ vertices with the ratio between edge weights at most $\frac{1}{\varepsilon}$, the probability that the process is not completed after $\mathcal{O}(N^5/\varepsilon)$ steps is in $\mathcal{O}(\frac{1}{2^N})$.*

*Proof.* From Lemma 4, we can take constant $c$ such that for both processes and all graphs with $N$ vertices, the expected time is at most $cN^4/\varepsilon$. From Markov's inequality [Lin10], the probability that the process takes more than $2cN^4/\varepsilon$ steps is at most $\frac{1}{2}$. If the process does not finish, the expected time is again $cN^4/\varepsilon$. That means we can take $N$ epochs of size $2cN^4/\varepsilon$ each, and the probability that the process does not finish in any epoch is at most $\frac{1}{2^N}$. □

**Lemma 6.** *For any graph $G_N$ with $N$ vertices any $r \geq 1$, and a process $p \in \{Bd, dB\}$ there exists a vertex $v$ such that $\rho_r^{\mathrm{p}}(G_N, v) \geq \frac{1}{N}$.*

*Proof.* It suffices to prove the statement for $r = 1$, since increasing the mutant fitness advantage $r$ increases its fixation probability [DGRS16, Theorem 6].

In the neutral case ($r = 1$), we have $\sum_{v \in V} \rho_{r=1}(G_N, v) = 1$, thus there exists at least one vertex with fixation probability at least $\frac{1}{N}$. □

Note that in some cases, no starting vertex $v$ satisfies both $\rho_r^{\mathrm{Bd}}(G_N, v) \geq 1/N$ and $\rho_r^{\mathrm{dB}}(G_N, v) \geq 1/N$ simultaneously. An example is a Star graph $S_3$ on $N = 3$ vertices with center $c$ and leaves $l_1$, $l_2$ when $r = 1$. Then $\rho_r^{\mathrm{Bd}}(S_3, c) = 1/5 < 1/3$ and $\rho_r^{\mathrm{dB}}(S_3, l_1) = \rho_r^{\mathrm{dB}}(S_3, l_2) = 1/4 < 1/3$.

### Construction

For given $N$ and $\gamma \in (0, 1)$, we describe how to construct graph $A_{N,\gamma}$. We show that for some $\gamma$, this graph is an amplifier for both processes for $r \in (1, 1.2)$. The graph $A_{N,\gamma}$ has two parts. The first part is a graph $\mathcal{A}^{Bd}_{(1-\gamma)N}$ (from Lemma 2) on $(1-\gamma)N$ vertices, the second part is a graph $\mathcal{A}^{dB}_{\gamma N}$ (from Lemma 3) on $\gamma N$ vertices. Let $\varepsilon$ be the smallest weight among edges when both graphs are independently scaled such that the largest edge weight is $1$.

We will connect the two parts by a single edge. To that end, we select a vertex $v$ from $\mathcal{A}^{Bd}_{(1-\gamma)N}$ such that the fixation probability starting from $v$ in $\mathcal{A}^{Bd}_{(1-\gamma)N}$ in dB-process is at least $\frac{1}{N}$, (such vertex exists from Lemma 6). Similarly, we select a vertex $v'$ from $\mathcal{A}^{dB}_{\gamma N}$ such that the fixation probability starting from $v'$ in the graph under Bd-process is at least $\frac{1}{N}$, (existence follows from Lemma 6). Then, we connect $v$ and $v'$ by an edge of weight $w = \frac{\varepsilon^3}{N^9}$.

Finally, we scale all edges in the first part $\mathcal{A}^{Bd}_{(1-\gamma)N}$ by a factor of $\frac{\varepsilon}{N^3}$. That is, the heaviest edge in $\mathcal{A}^{dB}_{\gamma N}$ has weight $1$, and the heaviest edge in $\mathcal{A}^{Bd}_{(1-\gamma)N}$ has weight $\frac{\varepsilon}{N^3}$. Observe that the scaling of edges in $\mathcal{A}^{Bd}_{(1-\gamma)N}$ does not influence the fixation time.

Before we turn to the main proof, we show several properties of the graph we $A_{N,\gamma}$ we have just constructed. The first property is that the two parts $\mathcal{A}^{Bd}_{(1-\gamma)N}$ and $\mathcal{A}^{dB}_{\gamma N}$ interact so rarely that most of the time they interact, the population on either part is already homogeneous (all mutants or all residents). Then we show Lemma 8 and Lemma 9. The lemmas show that in both processes, the probability of an individual reproducing over the edge between $v$ and $v'$ is unbalanced and in both processes, the individual in the respective amplifier is more likely to spread to the other graph.

**Lemma 7.** *For any $N$, $\gamma$, both processes, and a randomly placed mutant in $A_{N,\gamma}$, the probability that mutants become extinct or fixate on their part of $A_{N,\gamma}$ before any reproduction over edge $v, v'$ is at least*

$$1 - \mathcal{O}(1/N^2).$$

*Proof.* First, we bound the probability that edge $v, v'$ is selected in both processes and then we use the union bound.

For Bd, the edge $v, v'$ is used either by (i) selecting the individual at $v$ and spreading over $v, v'$, or (ii) selecting the individual at $v'$ and spreading over $v', v$. Event (i) happens with

probability at most $\frac{r}{N+(r-1)} \cdot \frac{\varepsilon^3/N^9}{\varepsilon^3/N^9+\varepsilon^2/N^3} < \frac{r\varepsilon}{N^7}$. Event (ii) happens with probability at most $\frac{r}{N+(r-1)} \cdot \frac{\varepsilon^3/N^9}{\varepsilon^3/N^9+\varepsilon} < \frac{r\varepsilon^2}{N^{10}}$. The sum of these probabilities is at most $\frac{2r\varepsilon}{N^7}$.

For dB, the edge $v, v'$ is used either if (i) individual at $v$ dies and is replaced individual at $v'$, or (ii) individual at $v'$ dies and is replaced by individual at $v$. Event (i) happens with probability at most $\frac{1}{N} \cdot \frac{r\varepsilon^3/N^9}{\varepsilon^3/N^9+\varepsilon^2/N^3} < \frac{r\varepsilon}{N^7}$. Event (ii) happens with probability at most $\frac{1}{N} \cdot \frac{r\varepsilon^3/N^9}{\varepsilon^3/N^9+\varepsilon} < \frac{r\varepsilon^2}{N^{10}}$. The sum of these probabilities is at most $\frac{2r\varepsilon}{N^7}$.

From Lemma 5, we know that with high probability the process ends in $\mathcal{O}(N^5/\varepsilon)$ steps. In every step the probability of using edge $v, v'$ is at most $\frac{2r\varepsilon}{N^7}$, that gives probability of using $v, v'$ at most $\mathcal{O}(\frac{1}{N^2})$ at first $N^5/\varepsilon$ steps from union bound. Since the probability that the process does not end during these steps is also in $\mathcal{O}(\frac{1}{N^2})$, we have that the randomly placed mutant resolves on one part of the graph before using edge $v, v'$ with a probability at least $1 - \mathcal{O}(\frac{1}{N^2})$ □

**Lemma 8.** *In the graph $A_{N,\gamma}$ under the Bd process, if edge $v, v'$ is used, then with probability at least $1 - \frac{r^2}{N^2}$ occupant of $v$ spreads to $v'$.*

*Proof.* At one step, individual at $v$ spreads to $v'$ with probability at least $\frac{1}{rN} \cdot \frac{\varepsilon^3/N^9}{\varepsilon^3/N^9+(N-1)\cdot\varepsilon/N^3} > \frac{\varepsilon^2}{rN^8}$. Individual at $v'$ spreads to $v$ with probability at most $\frac{r}{N+(r-1)} \cdot \frac{\varepsilon^3/N^9}{\varepsilon^3/N^9+\varepsilon} < \frac{r\varepsilon^2}{N^{10}}$. Conditioned that the spread over $v, v'$ happens, it is from $v'$ to $v$ with probability at most

$$\frac{\frac{r\varepsilon^2}{N^{10}}}{\frac{\varepsilon^2}{rN^8} + \frac{r\varepsilon^2}{N^{10}}} < \frac{r^2}{N^2}\,.$$

The opposite event, $v$ spreading to $v'$ happens with probability at least $1 - \frac{r^2}{N^2}$. □

**Lemma 9.** *In the graph $A_{N,\gamma}$ under the dB process, if edge $v, v'$ is used, then with probability at least $1 - \frac{r^2}{N^2}$ occupant of $v'$ spreads to $v$.*

*Proof.* At one step, individual at $v'$ spreads to $v$ with probability at least $\frac{1}{N} \cdot \frac{\varepsilon^3/N^9}{\varepsilon^3/N^9+r(N-1)\cdot\varepsilon/N^3} > \frac{\varepsilon^2}{rN^8}$. Individual at $v'$ spreads to $v$ with probability at most $\frac{1}{N} \cdot \frac{r\varepsilon^3/N^9}{r\varepsilon^3/N^9+\varepsilon} < \frac{r\varepsilon^2}{N^{10}}$. Conditioned that the spread over $v, v'$ happens, it is from $v$ to $v'$ with probability at most

$$\frac{\frac{r\varepsilon^2}{N^{10}}}{\frac{r\varepsilon^2}{N^8} + \frac{r\varepsilon^2}{N^{10}}} < \frac{r^2}{N^2}\,.$$

The opposite event, $v$ spreading to $v'$ happens with probability at least $1 - \frac{r^2}{N^2}$. □

### Proof of Amplification

**Lemma 10** (Amplification under Bd)**.** *For every $r$ and Bd updating, the fixation probability on $A_{N,\gamma}$ is at least*

$$1 - \gamma - \mathcal{O}(N^{-1/3})\,.$$

*Proof.* For Bd, we first bound the probability that mutants conquer $\mathcal{A}^{Bd}_{(1-\gamma)N}$. For this to happen, it suffices if:

1. The initial mutant appears at the correct part of the graph (that is, $\mathcal{A}^{Bd}_{(1-\gamma)N}$).
2. In the next $N^5/\varepsilon$ steps, the process ends in $\mathcal{A}^{Bd}_{(1-\gamma)N}$ without edge $v, v'$ being used.
3. Mutants conquer $\mathcal{A}^{Bd}_{(1-\gamma)N}$ within $N^5/\varepsilon$ steps.

The first condition is fulfilled with probability $1-\gamma$ since the initialization is uniformly random. The process resolves on $\mathcal{A}^{Bd}_{(1-\gamma)N}$ without edges $v, v'$ interference with probability $1-\mathcal{O}(\frac{1}{N^2})$, from Lemma 7. If the process is finished, the mutants spread with probability at least $1-\mathcal{O}(N^{-1/3})$, from Lemma 2.

Putting these probabilities together gives a probability at least

$$(1-\gamma)\cdot\left(1-\mathcal{O}\left(N^{-2}\right)\right)\cdot\left(1-\mathcal{O}(N^{-1/3})\right) > 1-\gamma-\mathcal{O}(N^{-1/3})$$

that the mutants conquer $\mathcal{A}^{Bd}_{(1-\gamma)N}$.

After the graph $\mathcal{A}^{Bd}_{(1-\gamma)N}$ is occupied by mutants, we bound from below the probability that the mutants fixate in the rest of the graph. We wait until the edge $(v, v')$ is used for reproduction, in one of its two directions. For fixation on the whole graph to occur, it suffices if:

1. The edge $v, v'$ was used in the right direction (from $v$ to $v'$).
2. In the next $N^5/\varepsilon$ steps, edge $v, v'$ is not used for reproduction.
3. Mutants fixate on $\mathcal{A}^{dB}_{\gamma N}$ within $N^5/\varepsilon$ steps.

The first condition happens with probability at least $1-\frac{r^2}{N^2}$, from Lemma 8. The edge $v, v'$ is not used within $N^5/\varepsilon$ steps with probability at least $1-\mathcal{O}(N^{-2})$, again from Lemma 7. If both of those occur, the mutants fixate with probability at least $\frac{1}{N}$, from Lemma 6 and since the process finishes within $N^5/\varepsilon$ steps with probability at least $1-2^{-N}$, by Union Bound the fixation probability is at least $\frac{1}{N}-2^{-N}$.

This gives the probability at least

$$\left(1-\frac{r^2}{N^2}\right)\cdot\left(1-\mathcal{O}(N^{-2})\right)\cdot\left(\frac{1}{N}-2^{-N}\right) > \frac{1}{N}-\frac{1}{N^2}$$

that if the edge $(v, v')$ is used, the process finishes with mutant fixation on the whole graph without edge $(v, v')$ being used again.

In contrast, if condition 1. fails, that is, the edge $(v, v')$ is instead used in the wrong direction (from $v'$ to $v$), we declare a failure (even though some of those evolutionary trajectories might eventually lead to mutant fixation). By Lemma 8, this happens with probability at most $\frac{r^2}{N^2}$. Similarly, we declare a failure if condition 2. fails, that is, when the edge $(v, v')$ is used (in either direction) during the $N^5/\varepsilon$ steps, potentially interrupting the process. Note that when condition 3. fails, that is, the mutants do not fixate in $\mathcal{A}^{dB}_{\gamma N}$ (but the edge $v, v'$ is not used), we are in the same state as before, where we can compute the fixation versus failure probability.

The failure probability is in $\mathcal{O}(N^{-2})$, the immediate fixation probability is at least $\frac{1}{N}-\frac{1}{N^2}$, otherwise, we can retry. This gives the fixation probability at least

$$\frac{\frac{1}{N}-\frac{1}{N^2}}{\mathcal{O}(N^{-2})+\left(\frac{1}{N}-\frac{1}{N^2}\right)} = 1-\mathcal{O}\left(\frac{1}{N}\right).$$

Overall, the fixation probability of a randomly placed mutant on $A_{N,\gamma}$ is thus at least

$$\left(1-\gamma-\mathcal{O}(N^{-1/3})\right)\cdot\left(1-\mathcal{O}\left(\frac{1}{N}\right)\right)\geq 1-\gamma-\mathcal{O}(N^{-1/3}).$$

□

**Lemma 11** (Amplification under dB)**.** *For every $r$ and dB updating, the fixation probability on $A_{N,\gamma}$ is at least*

$$\left(\frac{1}{2}\gamma(1-r^{-3})-\mathcal{O}(N^{-1})\right).$$

*Proof.* For dB, we proceed similarly as in the previous lemma. First, we again bound the probability that mutants conquer $\mathcal{A}^{dB}_{\gamma N}$. For this to happen, it suffices if:

1. The initial mutant appears at the correct part of the graph: $\mathcal{A}^{dB}_{\gamma N}$.
2. In the next $N^5/\varepsilon$ steps, the process ends in $\mathcal{A}^{dB}_{\gamma N}$ without edge $v,v'$ being used.
3. Mutants conquer $\mathcal{A}^{dB}_{\gamma N}$ within $N^5/\varepsilon$ steps.

The first condition is fulfilled with probability $\gamma$ since the initialization is uniformly random. The process resolves on $\mathcal{A}^{dB}_{\gamma N}$ without edges $v,v'$ interference with probability $1-\mathcal{O}(\frac{1}{N^2})$, from Lemma 7. If the process has resolved on $\mathcal{A}^{dB}_{\gamma N}$, the mutants conquer it with probability at least $\frac{N}{2N+1}\cdot\frac{1-\frac{1}{r^3}}{1-\frac{1}{r^{3N}}}-\mathcal{O}(2^{-N})$, from Lemma 3 and Union Bound.

Putting these probabilities together gives a probability at least

$$\gamma\cdot\left(1-\mathcal{O}\left(N^{-2}\right)\right)\cdot\left(\frac{N}{2N+1}\cdot\frac{1-\frac{1}{r^3}}{1-\frac{1}{r^{3N}}}-\mathcal{O}(2^{-N})\right)>\frac{1}{2}\gamma(1-r^{-3})-\mathcal{O}(N^{-1})$$

that the mutants conquer $\mathcal{A}^{dB}_{\gamma N}$.

After the graph $\mathcal{A}^{dB}_{\gamma N}$ is occupied by mutants, we bound the probability that the mutants fixate in the rest of the graph. Conditioned on the fact that the edge $v,v'$ is used, it happens when

1. The edge $v,v'$ was used in the right direction (from $v'$ to $v$).
2. In the next $N^5/\varepsilon$ steps, edge $v,v'$ is not used.
3. Mutants fixate on $\mathcal{A}^{Bd}_{(1-\gamma)N}$ within $N^5/\varepsilon$ steps.

If the edge $v,v'$ is used in the wrong direction, we call it a fail, this happens with probability at most $\frac{r^2}{N^2}$, from Lemma 9 (if the edge is used).

The first condition happens with probability at least $1-\frac{r^2}{N^2}$, from Lemma 9. The edge $v,v'$ is not used within $N^5/\varepsilon$ steps with probability at least $1-\mathcal{O}(N^{-2})$, again from Lemma 7. The mutants fixate with probability at least $\frac{1}{N}$, from Lemma 6 and since process finishes within $N^5/\varepsilon$ steps with probability at least $1-2^{-N}$, the fixation is at least $\frac{1}{N}-2^{-N}$.

This gives the probability at least

$$\left(1-\frac{r^2}{N^2}\right)\cdot\left(1-\mathcal{O}(N^{-2})\right)\cdot\left(\frac{1}{N}-2^{-N}\right)>\frac{1}{N}-\frac{1}{N^2}$$

that if the edge $v, v'$ is used, the process finishes without edge $v, v'$ being used again.

However, when the mutants do not fixate in $\mathcal{A}^{Bd}_{(1-\gamma)N}$ (but the edge $v, v'$ is not used), we are in the same state as before, where we can compute the fixation versus fail probability.

The fail probability is in $\mathcal{O}(N^{-2})$, the immediate fixation probability is at least $\frac{1}{N} - \frac{1}{N^2}$, otherwise, we can retry. This gives total fixation $1 - \mathcal{O}(\frac{1}{N})$.

Overall, the fixation probability of a randomly placed mutant on $A_{N,r}$ is at least

$$\left(\frac{1}{2}\gamma(1-r^{-3}) - \mathcal{O}(N^{-1})\right) \cdot \left(1 - \mathcal{O}(\frac{1}{N})\right) \geq \left(\frac{1}{2}\gamma(1-r^{-3}) - \mathcal{O}(N^{-1})\right) .$$

□

The following theorem shows that for a particular $\gamma$, our construction is an amplifier for both processes for $r \in (1, 1.2)$.

**Theorem 6** (Simultaneous Bd- and dB-amplifier)**.** *For every large enough population size $N$, for graph $A_{N,\gamma}$, where $\gamma = \frac{2 \cdot 1.2^3}{3 \cdot 1.2^3 - 1} = 0.826004$ we have*

$$\rho^{\mathrm{Bd}}_r(A_{N,\gamma}) > \rho^{\mathrm{Bd}}_r(K_N) \quad \textit{and} \quad \rho^{\mathrm{dB}}_r(A_{N,\gamma}) > \rho^{\mathrm{dB}}_r(K_N)$$

*for every $r \in (1, 1.2)$.*

*Proof.* We know that $\rho^{\mathrm{Bd}}_r(K_N) = \frac{1-r^{-1}}{1-r^{-N}}$ and $\rho^{\mathrm{dB}}_r(K_N) = \frac{N-1}{N}\frac{1-r^{-1}}{1-r^{-N+1}}$. Setting $N$ so big that $\frac{1}{1-r^{-N+1}} < 1.00001$, we have that $\rho^{\mathrm{Bd}}_r(K_N) < (1-r^{-1}) \cdot 1.00001 < 0.16667$ and $\rho^{\mathrm{dB}}_r(K_N) < (1-r^{-1}) \cdot 1.00001$.

From Lemma 10, plugging $\gamma$, we have that the fixation probability is at least $1 - 0.826004 - \mathcal{O}(N^{-1/3}) = 0.173996 - \mathcal{O}(N^{-1/3})$ for the Birth-death process which is bigger than the maximal fixation probability for $K_N$ $(0.16667)$.

From Lemma 11, plugging $\gamma$, we have that the fixation probability is at least $\left(\frac{1}{2}0.826004(1-r^{-3}) - \mathcal{O}(N^{-1})\right)$ for the death-Birth process. We have

$$\begin{aligned}
(1-r^{-1}) \cdot 1.00001 &< \left(\frac{1}{2}0.826004(1-r^{-3}) - \mathcal{O}(N^{-1})\right) \\
1.00001 &< \left(\frac{1}{2}0.826004(1+r^{-1}+r^{-2})\right) - \mathcal{O}(N^{-1}) \\
1.00001 &< \left(\frac{1}{2}0.826004 \cdot 2.52778\right) - \mathcal{O}(N^{-1}) \\
1.00001 &< 1.04 - \mathcal{O}(N^{-1}) ,
\end{aligned}$$

which proves the theorem. □

The following theorem shows how to choose $\gamma$ to achieve the best amplification so that the fixation probability of the amplifier is at least $1.04$ times better than the complete graph for both processes.

**Theorem 7** (Optimal Bd- and dB-amplifier)**.** *For any $r \in (1, 1.2)$, and every large enough population size $N$, for graph $A_{N,\gamma}$, where $\gamma = \frac{2r^3}{3r^3-1}$ we have*

$$\rho_r^{\mathrm{Bd}}(A_{N,\gamma}) > X \cdot \rho_r^{\mathrm{Bd}}(K_N) \quad \textit{and} \quad \rho_r^{\mathrm{dB}}(A_{N,\gamma}) > X \cdot \rho_r^{\mathrm{dB}}(K_N).$$

*for $X = 1.04$.*

*Proof.* Again, $\rho_r^{\mathrm{Bd}}(K_N) = \frac{1-r^{-1}}{1-r^{-N}}$ and $\rho_r^{\mathrm{dB}}(K_N) = \frac{N-1}{N}\frac{1-r^{-1}}{1-r^{-N+1}}$. Setting $N$ so big that $\frac{1}{1-r^{-N+1}} < 1.00001$, we have that $\rho_r^{\mathrm{Bd}}(K_N) < (1-r^{-1}) \cdot 1.00001$ and $\rho_r^{\mathrm{dB}}(K_N) < (1-r^{-1}) \cdot 1.00001$.

From Lemma 10, plugging $\gamma$, we have that the fixation probability is at least $\frac{r^3-1}{3r^3-1} - \mathcal{O}(N^{-1/3})$ for the Birth-death process. We have

$$\begin{aligned}
(1-r^{-1}) \cdot 1.00001 \cdot X &< \frac{r^3-1}{3r^3-1} - \mathcal{O}(N^{-1/3}) \\
1.00001 \cdot X &< \frac{r(1+r+r^2)}{3r^3-1} - \mathcal{O}(N^{-1/3}) \\
1.00001 \cdot X &< 1.04398 - \mathcal{O}(N^{-1/3}).
\end{aligned}$$

From Lemma 11, plugging $\gamma$, we have that the fixation probability is at least $\left(\frac{1}{2}\frac{2r^3}{3r^3-1}(1-r^{-3}) - \mathcal{O}(N^{-1})\right)$ for the death-Birth process. We have

$$\begin{aligned}
(1-r^{-1}) \cdot 1.00001 \cdot X &< \left(\frac{1}{2}\frac{2r^3}{3r^3-1}(1-r^{-3}) - \mathcal{O}(N^{-1})\right) \\
1.00001 \cdot X &< \left(\frac{1}{2}\frac{2r^3}{3r^3-1}(1+r^{-1}+r^{-2}) - \mathcal{O}(N^{-1})\right) \\
1.00001 \cdot X &< \left(\frac{r(1+r+r^2)}{3r^3-1} - \mathcal{O}(N^{-1})\right) \\
1.00001 \cdot X &< 1.04398 - \mathcal{O}(N^{-1}).
\end{aligned}$$

□

CHAPTER 3

# Complexity of Spatial Games

This chapter appears in full in [CIJS22].


## Abstract

Spatial games form a widely-studied class of games from biology and physics modeling the evolution of social behavior. Formally, such a game is defined by a square ($d$ by $d$) payoff matrix $M$ and an undirected graph $G$. Each vertex of $G$ represents an individual, that initially follows some strategy $i \in \{1, 2, \ldots, d\}$. In each round of the game, every individual plays the matrix game with each of its neighbors: An individual following strategy $i$ meeting a neighbor following strategy $j$ receives a payoff equal to the entry $(i, j)$ of $M$. Then, each individual updates its strategy to its neighbors' strategy with the highest sum of payoffs, and the next round starts. The basic computational problems consist of reachability between configurations and the average frequency of a strategy. For general spatial games and graphs, these problems are in PSPACE. In this paper, we examine restricted setting: the game is a prisoner's dilemma; and $G$ is a subgraph of grid. We prove that basic computational problems for spatial games with prisoner's dilemma on a subgraph of a grid are PSPACE-hard.


## 3.1 Introduction

Spatial evolutionary games is a classic and well-studied model of evolutionary dynamics on graphs, which has been studied across fields, e.g., biology [OHLN06, NM92a], physics [ONP07, RCS06], and computer science [DGP09, CDT09].

While computer science studies games with few players and a large number of actions, evolutionary game theory studies games with few actions and strategies but with many players (see the survey [SF07]). Specifically, each spatial evolutionary game consists of a square, skew-symmetric, bimatrix game (i.e. the outcome in entry $(i, j)$ for player 1 is the same as the outcome for player $2$ in $(j, i)$ for all $i, j$) and a finite graph. The game is played over a number of rounds. Each node of the graph corresponds to a player. Each node/player is associated with a current row and corresponding column. In each round, each player plays the matrix game against each of their neighbors, by playing their row against their neighbor's column (because of the skew-symmetry, who plays rows and who plays columns does not matter) and gets a payoff assigned, which is the sum of outcomes of the games they played in that round.

Each player then switches to the row played by their neighbor that had the highest payoff (or keeps their strategy). Since this is a deterministic dynamic, whenever we reach a round such that there is a previous round in which each player had the same row as now, the game "loops". All spatial evolutionary games will therefore loop after at most $d^n$ rounds when the bi-matrix is $d$ by $d$ and there are $n$ nodes.

**Most studied setup: Grids and prisoner's dilemma.** The standard study of spatial evolutionary games focuses on small (2 by 2 or 3 by 3) bi-matrices and on grids. A good understanding of the dynamics is known for a number of such setups (including all 2 by 2 bi-matrices on 2-dimensional grids). Grids are typically chosen since in biology they naturally model how cells interact with nearby cells. In particular, the focus has been on prisoner's dilemma (PD) matrices: A prisoner's dilemma bi-matrix is a 2 by 2 bi-matrix in which the two rows are called cooperation and defection, such that it is an advantage to defect if the other player cooperates, but it is better for both players if they both cooperate as compared to the mutual defection. We study these games to better understand why cooperation develops as we see in humans and many animal species. Indeed, this was the focus of the original paper [NM92b] and later works extended this basic model in myriad ways:

1. What happens if you add in mobile agents [VTA07, HY09]?
2. What about age[PS08, WWZA12]?
3. What if nodes/players do not have the same objectives [SS07, PW10]?
4. What if there were different timescales [RCS06, WRH09]?

See also the survey [PS10b]. A common generalization considers more general graph types. We mention a few examples of the papers pursuing this direction:

1. Kabir et al. showed how increasing the network reciprocity changes the likelihood of cooperation in PD [KTW18].
2. Hassell et al. considered how multiple different species on models with islands influence the outcome [HCM94].
3. Santos and Pacheco showed how scale-free networks (when generated following specific paradigms) promote cooperation [SP05].
4. Yamauchi et al. showed how, if both the neighbors and strategies can change over time, cooperation can evolve [SPL06a].
5. Ohtsuki et al. gave a simple heuristic for when cooperation can evolve on a variety of different networks, including social networks [OHLN06].

**Computational problems:** The works mentioned previously usually examine how cooperation spreads when the process is applied many times. The question they are trying to decide is: Given a starting position, what is the average number of cooperators in the long run?

**Open questions:** The general spatial games problem is in PSPACE. This is because a configuration consists of a strategy for every vertex, which can be stored in polynomial space, and the update of the configuration according to the rules can also be achieved in PSPACE. The most well-studied problem for spatial games is the prisoner's dilemma, which has only two strategies, namely, cooperation and defection. Moreover, such games have been studied for special classes of graphs. The main open question is whether efficient algorithms can be obtained for prisoner's dilemma on graphs like grids.

**Our results:** In this paper, we consider spatial evolutionary games with a prisoner's dilemma matrix on *subsets*[1] of 2-dimensional grids. The subset models the situation when locations of e.g. cells or connections between them have been destroyed or are otherwise inaccessible.

Our main result is that for subgraphs of a two-dimensional grid and prisoners dilemma the reachability and average cooperation problems are PSPACE-hard. Additionally, we show that induced subgraphs can loop in loops of exponential length. Subsets of grids are simpler than scale-free or evolving networks, so our hardness result holds for more general graphs.

## 3.2 Model and definitions

**Graphs and grids** A graph $G = (V, E)$ consists of a set of vertices $V$ and a set of undirected edges $E$. Every vertex is occupied by one individual. Edges determine pairs of individuals that interact. Two-dimensional grids are specific graphs where each vertex is assigned a unique pair of integers $(i, j)$, and has (at most) $8$ neighbors: vertices whose pairs differ by at most $1$ in both coordinates. In our construction, we use graphs derived from grids: Given a grid $(V, E)$, a *induced subgraph* of $(V, E)$ is a graph $(V', (V' \times V') \cap E)$ with $V' \subseteq V$. An *subgraph* of $(V, E)$ is a graph $(V', E')$ with $V' \subseteq V$ and $E' \subseteq (V' \times V') \cap E$.

**Games and individuals** An individual occupying a vertex has one of two types: cooperator if it plays $C$ or defector if it plays $D$. In the figures, we use black for cooperators and white for defectors. We call *configuration* of a graph an assignment of each vertex to a strategy (cooperator or defector).

From all matrix games, we focus on prisoner's dilemma. The game is denoted by a matrix

| | C | D |
|---|---|---|
| C | 1 | 0 |
| D | b | 0 |

where $b > 1$. It means that $C$ gets $1$ for interacting with $C$, $D$ gets $b$ for interacting with $C$, and everyone gets $0$ for interacting with $D$. We denote this game $M^b$.

**Steps and updates** The evolution is simulated in rounds. In one round, every individual interacts with all neighbors and collects the total payoff. Then every individual compares the received payoff with the payoffs of neighbors. The individual keeps the strategy if it is the highest or changes the strategy to the strategy of a neighbor with the highest payoff (in a tie, defection is preferred).

We denote the proccess starting from position $S$ on graph $G$ with a game matrix $M^b$ as $\mathcal{S}(G, M^b, S)$.

[1] Either subgraphs or induced subgraphs

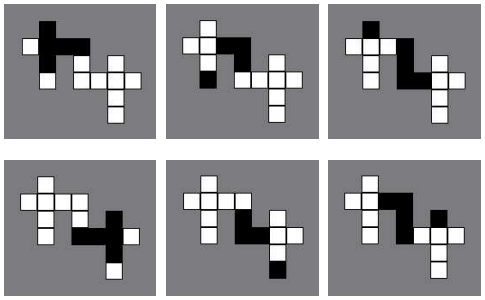


Figure 3.1: All configurations of the gadget $c_3$ initiated by the top left configuration. These configurations (in rows) show period 6.

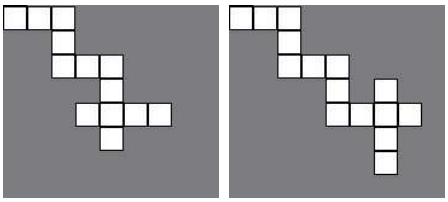


Figure 3.2: Extension of the gadget $c_k$ by two cells. We extend the gadget to the lower right, the gadget itself can be as long as needed connected to the upper left which increases period by 2.

**Complexity problems** We consider two complexity problems.

Reach : Given starting configuration, does the process reach a given configuration?

Avg : For a given starting configuration, what is the average number of cooperators (black vertices) in all succeeding configurations?

Note that for one configuration, there is only one possible succeeding configuration. This restricts the configuration graph. That means eventually the configuration graph creates a loop (as was noted before [Vir08]).

**Ranges of $b$** There is a reasonable range for $b$ in $M^b$. If $b < 1$, then the game is not a prisoner's dilemma. If $b$ is larger than the maximal degree in a graph, the dynamic is trivial, a defector cannot become a cooperator. For our constructions in this paper, we suppose that $b \in (\frac{3}{2}, 2)$. In Section 3.6, we show ideas explaining how our construction can be adapted to other values of $b$.

## 3.3 Exponential cycle

In this section, we show that even on an induced subgraph of a square grid, we observe a complex behavior. Namely, there exists a graph and a configuration that returns back to the starting configuration only after an exponential number of steps, we use $\tilde{(}\Omega)$ and $\tilde{(}\mathcal{O})$ which hide logarithmic factors.

**Theorem 8.** *For $b \in (\frac{3}{2}, 2)$, there exists a graph $G$, induced subgraph of a square grid, with $n$ vertices and a starting configuration $S$, such that it takes $2^{\tilde{\Omega}(\sqrt{n})}$ steps until $\mathcal{S}(G, M^b, S)$ reaches $S$ again.*

*Proof.* We describe a family of gadgets and starting configurations with different periods, where *period* is the number of steps the gadget needs to return to the starting position again. Then we combine some of them to create a graph with a period equal to the lowest common multiple of the periods of its components.

For $k \geq 3$, we construct inductively $c_k$, a gadget that is a induced subgraph of a square grid with size $9 + 2k$ and period $2k$. The gadget $c_3$ in its starting configuration is depicted on

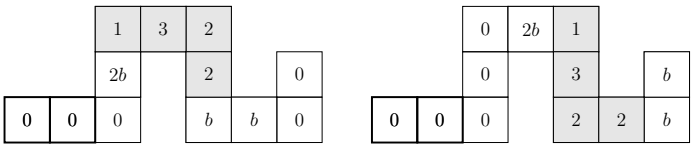



Figure 3.3: Sending signal through the wire with explicit payoffs with cooperators denoted by gray. Thicker vertices are input vertices.

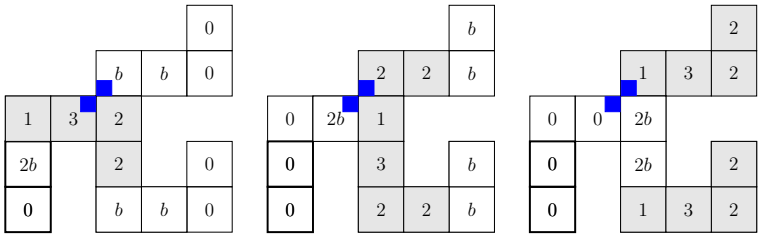



Figure 3.4: Splitting the signal in two. Blue boxes denote a deleted edge with cooperators denoted in black. Thicker vertices are input vertices.

Figure 3.1. By rearranging and adding two squares, we create the gadget $c_{k+1}$ from $c_k$, as depicted on Figure 3.2.

For every integer $g > 1$, let us now define $G_g$ as the disjoint union of the gadgets $c_{p_2}$, $c_{p_3}$, ..., $c_{p_g}$ where $p_i$ is the $i$-th prime number. By the Chinese remainder theorem, the number of steps needed so that all gadgets are in the same state is the smallest common multiple of their periods, which is the product of the first $g$ primes. We know that this number is $\Theta(e^{g \log g})$ from [MR97] and the Prime Number Theorem. Moreover, the number of squares (vertices) of $G_g$ linear in the sum of the first $g$ primes which is $\tilde{\mathcal{O}}(g^2)$.

Therefore, by using a induced subgraph of the square grid of size $n \in \mathbb{N}$, we can create a union of gadgets that has a period of size $2^{\tilde{\Omega}(\sqrt{n})}$. □

## 3.4 Construction of a Turing machine

We show that both REACH and AVG are P-SPACE hard. For a polynomially bounded Turing machine and its input, we create a graph of a polynomial-size that is a subgraph of the square grid and an initial configuration of cooperators and defectors such that the process reaches a predefined configuration if and only if the Turing machine accepts the given input.

Since both problems, REACH and AVG , can be easily solved in P-SPACE by a simulation, that means these problems are P-SPACE complete.

**Theorem 9.** *For $b \in (\frac{3}{2}, 2)$ holds:*

*For any Turing machine $T$ polynomially bounded by $n$ and its input, there exists a subgraph of a square grid $G$ with poly$(n)$ vertices, a starting configuration $S$, and a target configuration $Q$, such that $\mathcal{S}(G, M^b, S)$ reaches $Q$ if and only if $T$ accepts the given input.*

*Moreover, for any Turing machine $T$ and its input $I$, there exists a subgraph of an infinite square grid $G$, a starting configuration $S$ with poly$(|I|)$ cooperators, and a target configuration $Q$, such that $\mathcal{S}(G, M^b, S)$ reaches $Q$ if and only if $T$ accepts the given input.*

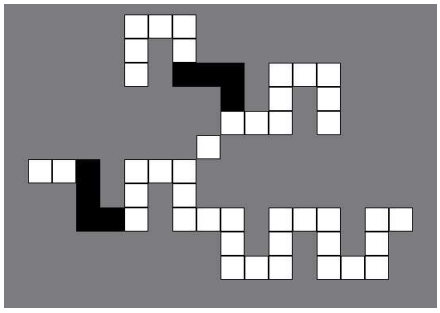


Figure 3.5: NOT gate: the upper signal is negated, the lower signal is a clock signal.

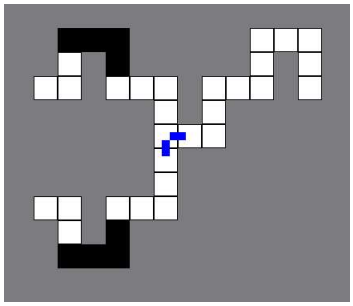


Figure 3.6: AND gate: both signals are essential to tunnel through. There are two deleted edges denoted by blue boxes.

We construct $G$ as a white (defector) graph with a few black (cooperator) vertices that carry signals and store data to simulate the behavior of $T$. On our figures, we use white for defectors, black for cooperators, gray for deleted vertices, and blue for *deleted edges*: To visualize deleted edges, we subdivide each square denoting some vertex $v$ into $9$ parts. The middle square corresponds to $v$ itself, and the other squares denote the neighbors in relative position to $v$ (upper-left, upper-middle, ...). To represent that an edge between two vertices $v$ and $w$ is deleted, we color in blue the subsquare corresponding to $w$ in $v$'s square and the subsquare corresponding to $v$ in $w$'s square.

First, we describe wires and basic logic gates. With that, we use a construction described in [CIJS20]. We use Lemma 7 and 9 from the paper, but explain technicalities emerging from more restrictive construction (the graph is a subgraph of a square grid). Both lemmas use simple gadgets to create functions and then a whole Turing machine.

### 3.4.1 Basic gadgets

We describe the computational gadgets used in our construction. Every gadget $g$ has one or two inputs ($I_1$, $I_2$) and one or two outputs ($O_1$, $O_2$). We imagine the inputs being on the left and bottom and the outputs on the right and top. Every gadget fits into a constantly sized rectangle of the grid.

If at time $t$ at least one input of $g$ is true (input vertices are cooperators and other vertices are defectors), then at time $t + t_g$ it outputs true or false based on a function $g$ computes.

One input (or output) consists of two connected vertices. We say that the input is true if both vertices are cooperators at time $t$ and $t + 1$, and in the first step, they can convert all neighbors to cooperators. The previous gadget (connected by its output to the given input) is responsible for turning the input vertices back to defectors at time $t + 2$.

Here are the gadgets, some of the gadgets need to cross signals, we describe how to do it later:

**Wire:** it transmits a signal (See Figure 3.3). The wire has one input, one output, and if $I_1$ is true at time $t$, then $O_1$ is true at time $t + 2$.

**Splitter:** it splits one signal in two (See Figure 3.4). The splitter has one input and two outputs. If $I_1$ is true at time $T$, then $O_1$ and $O_2$ are true at time $t + 4$. An interesting

property of the splitter is that it does not matter if the signal arrives from $I_1$ or $O_1$. From the graph perspective, these two are symmetric.

**NOT gate:** it computes the logical negation, but necessitates a clock signal (See Figure 3.5). The NOT gate has two inputs and one output. Input $I_2$ is a clock input: if $I_1$ and $I_2$ are true at time $t$, then $O_1$ is false at time $t+8$, if only $I_2$ is true at time $t$, then $O_1$ is true at time $t+8$. Note that the NOT gate actually computes the function $\neg I_1 \mathsf{AND} I_2$, also the gates $I_1$ and $O_1$ are interchangeable.

**AND gate:** it computes the logical conjunction (See Figure 3.6). The AND gate has two inputs and one output. If both $I_1$ and $I_2$ are true at time $t$, then $O_1$ is true at time $t+8$. Otherwise, $O_1$ is false.

**XOR gate:** it computes logical exclusive disjunction. We can create it from other gadgets, but it makes our construction easier.

We compute XOR using connection of splitter and NOT gates (see Figure 3.7). It has two inputs $I_1$ and $I_2$ and one output $O_1$. First every input is split to get $I_{1,1}, I_{1,2}, I_{2,1}$ and $I_{2,2}$, then $I_{1,1}$ is sent to $I_1$ of NOT gate and $I_{2,1}$ is sent to $O_1$ of a splitter $S$. We already know that if only one signal equals true, then $S$ splits the signal. If both are positive, signals annihilate each other. So $O_2$ of $S$ has value $I_1 \mathsf{XOR} I_2$. The only problem is when $I_1$ is true, and $I_2$ is not (or symmetric). Then the signal travels back from the splitter through wire $I_{2,1}$, then we use the NOT gate with clock signal $I_{2,1}$ and signal $I_{1,2}$ that stops it.

**OR gate:** it computes logical disjunction. It has two inputs $I_1$ and $I_2$, and the output $O_1$ satisfies the function $(I_1 \mathsf{AND} I_2) \mathsf{XOR} (I_1 \mathsf{XOR} I_2)$, it consists of gadgets described above. Note that we don't need the NOT gate directly, so clock signal is not necessary for that gadget.

Another gadget that the construction needs is a storage unit. It has an inner state $S \in \{0, 1\}$, two inputs and one output, see Figure 3.8. The storage unit consists of a big cyclic wire where a signal loops if the stored value $S$ is $1$. If a signal is sent via $I_1$, then the storage unit changes state: $S' = \neg S$, this is ensured by a XOR gate. On the cycle, there is a splitter that splits the signal towards an AND gate. If a signal is sent via $I_2$, $S$ does not change, but the signal reaches AND and is sent to $O_2$ if and only if $S = 1$. The storage unit requires synchronicity, the signal from the input has to reach the splitter or XOR at the right time. But this is not hard to ensure by longer wires.

### 3.4.2 Connecting the graph

To make the graph a subgraph of a square grid, we need to prove two things: we can cross two signals, and we can ensure that the signals meet at the right place, at the right time.

**Crossing** We use the structure described in Figure 3.13. Crossing accepts only one signal at a time and supposes that no other signal arrives for $10$ steps afterward. Crossing ensures that in a square subset of a grid, the signal can travel only from upper left to lower right corner and from lower left to upper right without spreading or dying out. If a signal arrives, it is spreading to the other input and all the outputs, but the wires close to the active input get stopped, so only the wire that is not adjacent to the input wire continues to carry the signal.

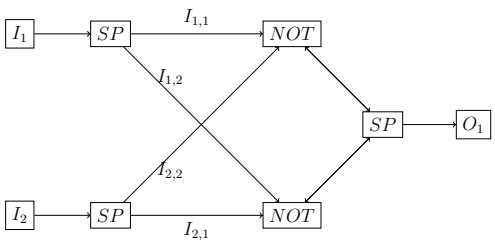



Figure 3.7: XOR gate: Combination of previous gates. Signals $I_{1,2}$ and $I_{2,2}$ are delayed such that they arrive at the NOT gate when signal $I_{1,1}$ or $I_{2,1}$ from the central splitter would.

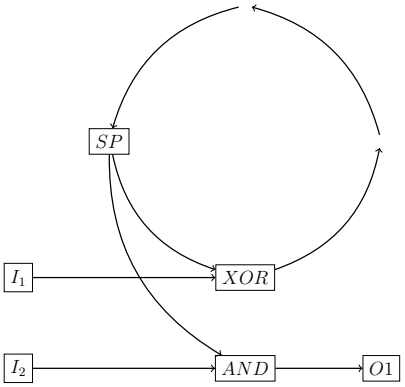



Figure 3.8: Storage unit

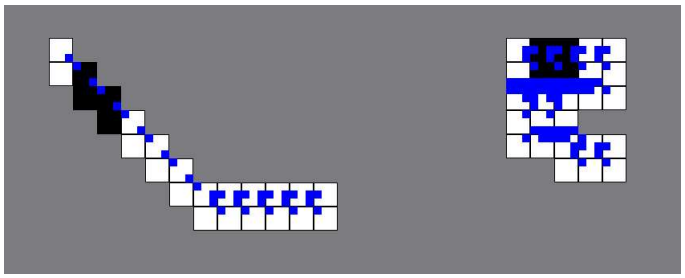


Figure 3.9: Two wires of different lengths connecting two points.

**Making the construction on square grid** On the Figure 3.9, we see that a signal starting at position $(0, 0)$ going to position $(x, y)$ can take between $\max(x, y)$ and $\frac{1}{8} \cdot xy$ steps. The signal also does not leave the rectangle dentoted by points $(0, 0)$ and $(x, y)$.

Every gadget that was described above can be padded by wires of different densities such that the gadget fits into a rectangle with predetermined width and height. Moreover, all inputs are in the same position (relative to the rectangle) and the time of evaluation is the same for all (constant).

Then everything in the following construction can be viewed as placing a column of tiles (gadgets) one after another, where columns are connected by short wires.

### 3.4.3 Function construction

Here we construct a function using previously described gadgets. We bound the number of vertices and steps needed for the construction and the function evaluation.

We imagine a function as signals going from left to right. The input and output signals have constant horizontal distance.

**Definition 1.** *We say that a function $f : \{0,1\}^k \to \{0,1\}^l$ is computed by a graph $G$ (subgraph of a grid) in time $g$ and space $h$ if $G$ has $h$ vertices, $k$ inputs $I_1$, $I_2$,..., $I_k$ and $l$ outputs $O_1$, $O_2$, ..., $O_l$ spatially arranged, and such that for all $(x_1, x_2, \ldots, x_k) \in \{0,1\}^k$, if at the time $t$ we have $I_j = x_j$ for all $1 \leq j \leq l$ (and all the other vertices are defectors), then at the time $t + g$ we have $O_j = y_j$ for all $1 \leq j \leq l$, where $f(x_1, x_2, \ldots, x_k) = (y_1, y_2, \ldots, y_l)$.*

Note that all our basic gadgets have a constant size. So, when analyzing the asymptotic space complexity, we can consider these gadgets and single vertices interchangeably.

**Lemma 12.** *Computing the function $f(x_1, x_2, \ldots, x_c) = (x_1, x_2, \ldots, x_c, x_1, x_2, \ldots, x_c)$ takes time $\mathcal{O}(c)$ and space $\mathcal{O}(c^2)$.*

*Proof.* We describe the gadget computing $f$. First, we split every signal and then cross every copy of it with others to the right place. One signal needs to cross $\mathcal{O}(c)$ others, so that is the number of steps and we have $c$ signals going through a wire of length $c$, that makes $\mathcal{O}(c^2)$ gadgets. □

**Lemma 13.** *Let $c \in \mathbb{N}$. Every function $f : \{0,1\}^c \to \{0,1\}$ mapping to $0$ every tuple whose first component is $0$ can be computed in space $\mathcal{O}(3^c)$ and time $\mathcal{O}(c^2)$.*

*Proof.* We use an induction on the dimension $c$ of the domain. If $c = 1$, since $f$ satisfies $f(0) = 0$ by supposition, either $f$ is the identity, which is realised by a wire, or $f$ is the zero function, which is realised by simply disconnecting the input and output.

Now, suppose that $c > 1$, and we can realise any function $f : \{0,1\}^{c-1} \to \{0,1\}$ that satisfies the condition. In particular, there exist two gadgets $g_0$ and $g_1$ computing

$$\begin{array}{rccc} f_0: & \{0,1\}^{c-1} & \to & \{0,1\}, \\ & (x_1, x_2, \ldots, x_{c-1}) & \mapsto & f(x_1, x_2, \ldots, x_{c-1}, 0), \\ f_1: & \{0,1\}^{c-1} & \to & \{0,1\}, \\ & (x_1, x_2, \ldots, x_{c-1}) & \mapsto & f(x_1, x_2, \ldots, x_{c-1}, 1). \end{array}$$

Having these values, the computation is straightforward as the value $f(x_1, x_2, \ldots, x_c)$ is given by the formula

$$(f_0(x_1, x_2, \ldots, x_{c-1}) \wedge \neg x_c) \vee (f_1(x_1, x_2, \ldots, x_{c-1}) \wedge x_c).$$

To construct a gadget $g$ that is a subgraph of a grid $g$ and computes $f$, we proceed as follows.

**Computing the function** We apply the gadget described in Lemma 12 to the input, and we use a splitter to get one more signal $x_1$. At this point, we have the arranged signals $x_1,\ x_1,\ x_2,\ x_3,\ \ldots,\ x_c,\ x_1,\ x_2,\ x_3,\ \ldots,\ x_c$. We apply the gadgets computing $f_0$ and $f_1$ to get the signals $f_0(x_1, x_2, \ldots, x_c)$ and $f_1(x_1, x_2, \ldots, x_c)$, now the signals are arranged as $x_1, f_0(x_1, x_2, \ldots, x_{c-1}), x_c, f_1(x_1, x_2, \ldots, x_{c-1}), x_c$. We cross the signals $x_1$ and $f_0(x_1, x_2, \ldots, x_{c-1})$, and then we use $x_1$ as clock signal towards a NOT gate with $x_c$. By this, we either get the signal $\neg x_c$, or $x_1$ is zero, then the result should be zero anyway.

Now we send $f_0(x_1, x_2, \ldots, x_{c-1})$ and $\neg x_c$ towards an AND gate, similarly we send $f_1(x_1, x_2, \ldots, x_{c-1})$ and $x_c$ towards another AND gate, and finally we send both results towards an OR gate.

**Size of the gadget** Now, we show that the computation is fast and requires a reasonable number of vertices. Let there be a recurrent formula $R_s$ that maps any integer $c$ to the maximal number of gadgets needed to compute the function $f$. Using induction, we see that it satisfies:

$$R_s(c) = \mathcal{O}(c^2) + 2R_s(c-1) + \mathcal{O}(1).$$

Therefore, $R(c) \in \mathcal{O}(3^c)$. Similarly, we use recurrent formula for the time needed

$$R_t(c) = \mathcal{O}(c) + R_t(c-1) + \mathcal{O}(1)$$

and the solution for this recurrence is $\mathcal{O}(c^2)$. □

**Lemma 14.** *Let $c, d \in \mathbb{N}$. Every function $f : \{0,1\}^c \to \{0,1\}^d$ mapping to $(0, 0, \dots)$ every tuple whose first component is $0$ can be computed in space $\mathcal{O}(d3^c)$ and time $\mathcal{O}(c^2 + c\log(d))$.*

*Proof.* The idea is easy, we split the inputs $d$ times using Lemma 12 and then we use Lemma 13 for every copy.

Multiplying the inputs needs $\mathcal{O}(d)$ described in Lemma 12. We can arrange them in layers where every layer doubles the input, so the whole multiplying takes time $\mathcal{O}(\log(d) \cdot c)$.

Then we use $d$ gadgets described in Lemma 13 in parallel which gives time $\mathcal{O}(c^2 + c\log(d))$ and space $\mathcal{O}(d3^c)$. □

### 3.4.4 Blob and connections

**Lemma 15.** *Let $T$ be a Turing machine. For every input $u$ evaluated by $T$ using $C \in \mathbb{N}$ cells of the tape, there exists a subgraph of a grid $G$ on $\mathcal{O}(C)$ vertices and an initial configuration $c_0$ of $G$ such that $T$ stops over the input $u$ if and only if $\mathcal{S}(G, M^b, c_0)$ for $b \in (\frac{3}{2}, 2)$ eventually reaches a configuration without cooperators.*

*Proof.* We suppose that the Turing machine $T$ has a single final state, which can only be accessed after clearing the tape. We present the construction of the graph $G$ simulating $T$ through the following steps. First, we encode the states of $T$, the tape alphabet, and the transition function in binary. Then, we introduce the notion of a blob, the building block of $G$, and we show that blobs accurately simulate the transition function of $T$. Afterward, we approximate the size of a blob, and finally, we define $G$ as a composition of blobs.

**Binary encoding** Let $T_s \in \mathbb{N}$ be the number of states of $T$, and $T_a \in \mathbb{N}$ be the size of its tape alphabet. We pick two small integers $s$ and $n$ satisfying $T_s \leq 2^{s-1}$ and $T_a \leq 2^{n-1}$. We encode the states of $T$ as elements of $\{0,1\}^s$, and the alphabet symbols as elements of $\{0,1\}^n$, while respecting the following three conditions: the blank symbol maps to $0^n$, the final state of $T$ maps to $0^s$, and all the others map to strings starting with $1$. Then, for these mappings, we modify the transition function of $T$ to:

$$F : \{0,1\}^s \times \{0,1\}^n \to \{0,1\}^s \times \{0,1\}^s \times \{0,1\}^n.$$

Instead of using one bit to denote if the head is going left or right, we use $2s$ bits to store the state and signify the movement: if the first $s$ bits are zero, the head is moving right; if the second $s$ bits are zero, it is moving left; if the first $2s$ bits are zero, the computation ended. Moreover, the last $n$ bits of the image of $F$ do not encode the new symbol, but the symmetric difference between the previous and the next symbol: if the $i$-th bit of the tape symbol goes from $y_i$ to $z_i$, then $F$ outputs $d_i = y_i \oplus z_i$ (XOR of these two).

**Constructing blobs** We construct the graph $G$ by simulating each cell of the tape with a blob. Blob stores a tape symbol, and after receiving a signal corresponding to a state of $T$ it computes the transition function. The main components of a blob are as follows.

- Memory: $n$ storage units ($s_1$, $s_2$, …, $s_n$) are used to keep in memory a tape symbol $a \in \{0,1\}^n$ of $T$.

- Receptor: $2s$ inputs ($I_1$, $I_2$, ..., $I_{2s}$) are used to receive states $q \in \{0,1\}^s$ of $T$ either from the left or from the right.
- Transmitter: $2s$ outputs ($O_1$, $O_2$, ..., $O_{2s}$) are used to send states $q \in \{0,1\}^s$ of $T$ either to the right or to the left.
- Transition gadget: We use gadget from Lemma 14, it needs $\mathcal{O}((n+s)3^{n+s})$ space and $\mathcal{O}((n+s)^2)$ time.

Blobs are connected in a row to act as a tape: for every $1 \leq i \leq s$, the output $O_i$ of each blob connects to the input $I_i$ of the blob to its right, and the output $O_{s+i}$ of each blob connects to the input $I_{s+i}$ of the blob to its left. When receiving a signal, the blob transmits the received state and the tape symbol stored in memory to the transition gadget $g_F$, which computes the corresponding transition, and then apply its results. We now detail this inner behavior. Note that when a gadget is supposed to receive simultaneously a set of signals coming from different sources, it is always possible to add wires of adapted length to ensure that all of them end up synchronized.

**Simulating the transition function** To simulate the transition function of $T$, a blob acts according to the three following steps:

**1.** Transmission of the state. A blob can receive a state either from the left (through inputs $I_1, I_2, \ldots, I_s$) or from the right (through inputs $I_{s+1}, I_{s+2}, \ldots, I_{2s}$), but not from both sides at the same time, since at every point in time there is at most one active state. Therefore, if for every $1 \leq i \leq s$ we denote by $x_i$ the disjunction of the signals received by $I_i$ and $I_{s+i}$, then the resulting tuple $(x_1, x_2, \ldots, x_s)$ is equal to the state received as signal (either from the left or the right), which can be fed to the gadget $g_F$. Formally, the blob connects, for all $1 \leq i \leq s$, the pair $I_i$, $I_{s+i}$ to an OR gate whose output is linked to the input $I_i$ of $g_F$.

**2.** Transmission of the tape symbol. Since the first component of any state apart from the final state is always $1$, whenever a blob receives a state, the component $x_1$ defined in the previous paragraph has value $1$. The tape symbol $(y_1, y_2, \ldots, y_n)$ currently stored in the blob can be obtained by sending, for every $1 \leq i \leq n$, a copy of $x_1$ to the input $I_2$ of the storage unit $s_i$, causing it to broadcast its stored state $y_i$. The tuple continues to the gadget $g_F$. Formally, the blob uses $n$ splitters to transmit the result of the OR gate between $I_1$ and $I_{s+1}$ to the input $I_1$ of each storage unit. Then, for every $1 \leq i \leq n$, the output $O_1$ of the storage unit $s_i$ is connected to the input $I_{s+i}$ of $g_F$.

**3.** Application of the transition. Upon receiving a state and a tape symbol, $g_F$ computes the result of the transition function, yielding a tuple $(r_1, r_2, \ldots, r_{s+n})$. The blob now needs to do two things: send a state to the successor blob and update the element of the tape.

Connecting the output $O_i$ of $g_F$ to the output $O_i$ of the blob for every $1 \leq i \leq 2s$ ensures that the state is sent to the correct neighbor: the values $(r_1, r_2, \ldots, r_s)$ are nonzero if the head is supposed to move to the next block on the right (outputs $O_1, O_2, \ldots, O_s$ are connected that blob). Conversely, $(r_{s+1}, r_{s+2}, \ldots, r_{2s})$ are nonzero if the head is supposed to move to the left, (outputs $O_{s+1}, O_{s+2}, \ldots, O_{2s}$ are connected to the left blob).

Finally, connecting the output $O_{2s+i}$ of $g_F$ to the input $I_1$ of $s_i$ for all $1 \leq i \leq n$ ensures that the state is correctly updated: this sends the signal $d_i$ to the input $I_1$ of the storage unit $s_i$. Since $d_i$ is the difference between the current bit and the next, the state of $s_i$ will change if it has to.

**The size of blob** The size of the blob is determined by the size of the transition gadget $g_F$: one blob is composed of $\mathcal{O}(3^{n+s})$ vertices, and evaluating a transition requires $\mathcal{O}((n+s)^2)$ steps by Lemma 14. Since $n$ and $s$ are constants (they depend on $T$, and not on the input $u$), the blob has constant size. Crossing wires to get them to the right place also takes space $\mathcal{O}((n+s)^2)$.

**Constructing $G$** Now that we have blobs that accurately simulate the transition function of $T$, constructing the graph $G$ mimicking the behavior of $T$ over the input $u$ is straightforward: we take a row of $C$ blobs (remember that $C \in \mathbb{N}$ is the number of tape cells used by $T$ to process $u$). Since the size of a blob is constant, the size of $G$ is polynomial in $C$. We define the initial configuration of $G$ by setting the states of the $|u|$ blobs on the left of the row to the letters of $u$, and setting the inputs $I_1$ to $I_s$ of the leftmost blob to the signal corresponding to the initial state of $T$ as if it was already in the process. As explained earlier, the blobs then evolve by following the run of $T$. If the Turing machine stops, then its tape is empty, the final state is encoded by $0^s$, and the blank symbol is encoded by $0^n$. So $G$ reaches the configuration where all vertices are defectors. Conversely, if $T$ runs forever starting from the input $u$, there will always be some cooperators vertices in $G$ to transmit the signal corresponding to the state of $T$. □

*Proof of Theorem 9,part 1.* By Lemma 15, we can reduce any problem solvable by a polynomially bounded Turing machine into $\textsc{Reach}$ , asking whether we reach the configuration with only defectors. or into $\textsc{Avg}$ , asking whether the long-run average is strictly above $0$. □

*Proof of Theorem 9,part 2.* Again, by Lemma 15, we can create an infinite chain of blobs and then we can reduce any problem solvable by any Turing machine into $\textsc{Reach}$ , asking whether we reach the configuration with only defectors, or into $\textsc{Avg}$ , asking whether the long-run average is strictly above $0$. □

## 3.5 Conclusion

Our result shows that going beyond the basic grid model even slightly makes spatial games PSPACE-hard. It ensures that there is no efficient (polynomial-time) algorithm for $\textsc{Reach}$ or $\textsc{Avg}$ , even for the simplest game of prisoner's dilemma.

We studied games with synchronous and deterministic updating. It poses a question: does lifting any restriction permit an efficient algorithm? Moreover, we can ask whether it is possible to construct graphs on which the game has certain properties, such as: are there graphs where the cooperative behavior is favored?

## 3.6 Construction for a more general b

We can make a similar construction for different $b$'s than those from the range $(\frac{3}{2}, 2)$. By selecting different $b$, the construction is not a subgraph of a square grid.

We can interpret the wire from Figure 3.3 as a path (lower row of vertices) with auxiliary vertices (upper row) that empower the tip of the cooperator signal and also help the first defector behind it. For different $b$'s, we can add more auxiliary vertices. Details are on Figure 3.10.

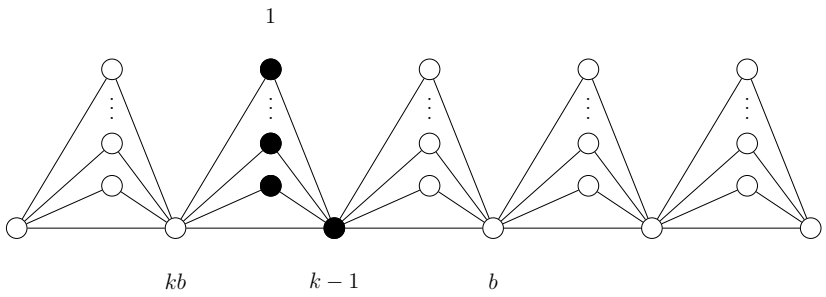



Figure 3.10: Schematic wire construction for general $b$ with the slices of size $k$. The numbers are payoffs of important vertices.

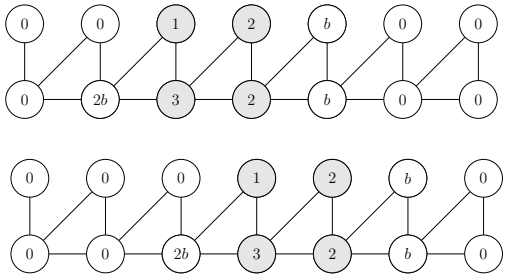



Figure 3.11: Sending signal throught the wire with explicit payoffs.

For setting $b \in (\frac{3}{2}, 2)$, the path and auxiliary vertices interchangeable, it's not the case for different $b$s. Our construction uses this, but only to preserve the topology properties. When we add $k$ auxiliary vertices, we need to ensure that $kb > k - 1$ and $k - 1 > b$. For small $b$, we change the signal that it does not span two consecutive slices, but one (as on Figure 3.10.

## 3.7 More gadgets

In this section, we provide more detailed figures for some gadgets. On Figure 3.11 and Figure 3.12 we see explicit graph shown on Figure 3.3 and Figure 3.4.

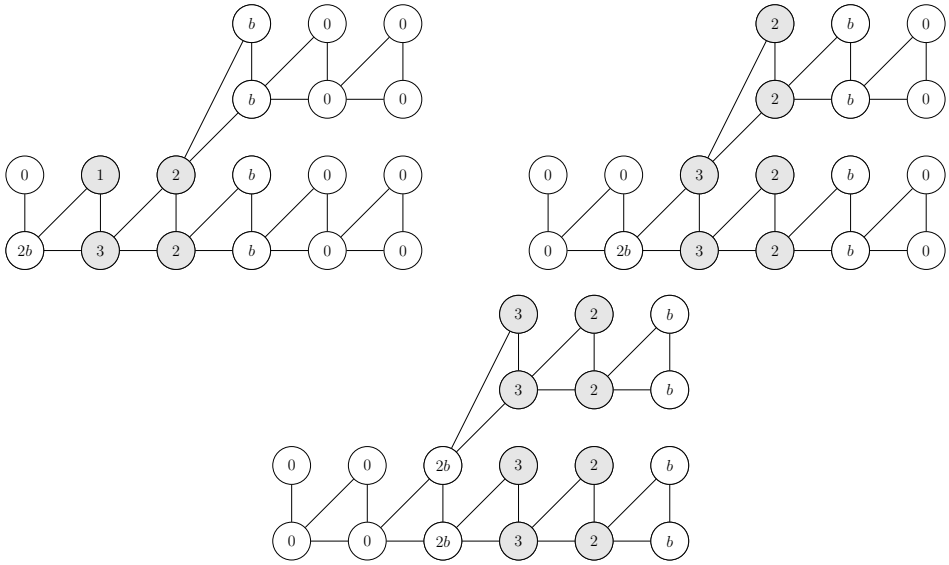



Figure 3.12: Splitting the signal in two. Note that the splitter does not have a direction (a signal from any input/output is split and sent by the other two inputs/outputs).

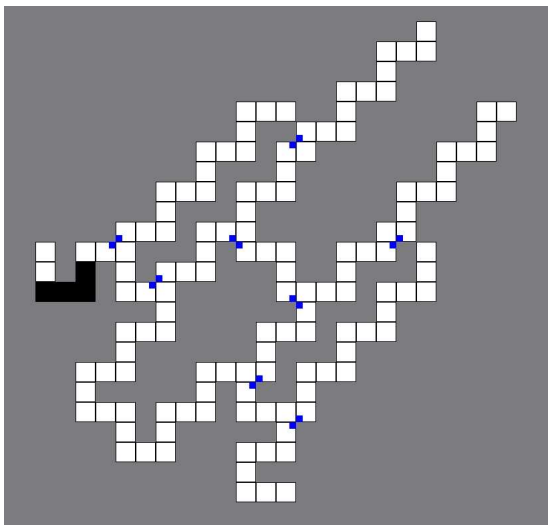


Figure 3.13: Crossing of two wires. Signal coming from left leaves the gadget by the right wire and signal coming from bottom leaves the gadget by the top wire.

CHAPTER 4

# Amplifiers of Cooperation for Spatial Games



## Abstract

Spatial games provide a simple and elegant mathematical model to study the evolution of cooperation in networks. In spatial games, individuals reside in vertices, adopt simple strategies, and interact with neighbors to receive a payoff. Depending on their own and neighbors' payoffs, individuals can change their strategy. The payoff is determined by the Prisoners' Dilemma, a classical matrix game, where players cooperate or defect. While cooperation is the desired behavior, defection provides a higher payoff for a selfish individual. There are many theoretical and empirical studies related to the role of the network in the evolution of cooperation. However, the fundamental question of whether there exist networks that for low initial cooperation rate ensure a high chance of fixation, i.e., cooperation spreads across the whole population, has remained elusive for spatial games with strong selection. In this work, we answer this fundamental question in the affirmative by presenting the first network structures that ensure high fixation probability for cooperators in the strong selection regime. Besides, our structures have many desirable properties: (a) they ensure the spread of cooperation even for a low initial density of cooperation and high temptation of defection, (b) they have constant degrees, and (c) the number of steps, until cooperation spreads, is at most quadratic in the size of the network.

### Significance statement

Spatial games provide a broad framework to study the spread of strategies over networks. Individuals residing in vertices play a matrix game with neighbors and update strategies based on the payoff. We address a long-standing problem: Do there exist networks that promote the spread of cooperation for Prisoners' dilemma? We present the first networks that ensure the spread of cooperation with high probability. We also establish several robustness properties of our networks: they promote cooperation even with a low initial cooperation rate and high temptation and ensure that cooperation spreads quickly. Due to the broad connection of

spatial games in several applications ranging from biology to physics, our new structures are significant and relevant in all these domains.

## 4.1 Introduction

Game theory is a broad field that provides the mathematical foundations related to decision-making, which has many applications in economics [Owe13, KMR93], computer science and artificial intelligence [NRTV07], evolutionary dynamics [Smi82, Now07, BR22, TK97, Ewe04], and physics [PJR$^+$17, CFL09, JHK$^+$22]. For example: one-shot or matrix games have been studied in [VNM47, Nas51]; their generalization as normal and extensive form games have wide applications in economics [Owe13, FM99]; and evolutionary games study competitive and collaborative behavior in biology [HS98, AFC15, GYVO09, KSW$^+$14, JMG$^+$13, Tar17, NDF16, IS01, Dug97, HG04].

An important class of games that arise in many different contexts is games on graphs. The graph represents a network or a population structure. Individuals reside in vertices, and the interactions happen over the edges of the graph. Well-known examples of games over graphs are: game of life [BCG04], games with cellular automata [Wol83]; evolutionary games on graphs [LHN05, ARS06, SAR08]; and spatial games [NM92a, ANT09, TOA$^+$09, DHD14, HD04, MRH21].

One of the classical examples in game theory to represent social tension by a matrix game is Prisoners' Dilemma (PD) [Nas51], which has been studied both theoretically [NRTV07] as well as experimentally [AH81, Axe97]. In PD, individuals can choose to cooperate or defect. If they mutually cooperate, they both receive payoff $R$. If they mutually defect, they receive payoff $P$. A cooperator facing a defector obtains payoff $S$, while a defector facing a cooperator gets $T$. The payoffs follow inequality $T > R > P > S$, which represents the dilemma between altruistic behavior and self-interest. Mutual cooperation is the desired behavior; however, defection is the dominant strategy, which achieves the best payoff for a selfish individual no matter the strategy of the other player. Hence mutual defection is the only Nash equilibrium. The population structure can help the cooperation to overcome this trap [NSTF04, May87, FF03, PWAT16, SSP08, SAP22, SMMP22].

Spatial games provide the framework to study PD games over structured populations (graphs) played over multiple rounds. Individuals residing in vertices can adopt only simple strategies, always cooperate $C$, or always defect $D$ in every round. In one round, the payoff of an individual is obtained by interacting with all neighbors by playing pairwise PD. Based on the individual's payoff and the payoffs of neighbors, the strategy is updated by a variant of a replicator dynamic [NM92a, HS98, Now07]. The strategy can be updated synchronously, where every individual can update its strategy; or asynchronously, where only one individual from a random pair of neighbors can update its strategy. The strategy update can be either deterministic: the individual with a lower payoff adopts the strategy of a more successful individual; or randomized: with some probability (given by the Fermi function) even the strategy with a lower payoff can replace a higher payoff strategy. This represents the specific regime of spatial games with strong selection, i.e., in the deterministic setting we have the best-response dynamics, and in the randomized setting replacement probability depends exponentially on payoff difference.

The study of spatial games was initiated in the seminal work of Nowak and May [NM92a] and since then, it received a lot of attention from several research perspectives. Oht-

suki et al. [OHLN06] theoretically examine the graphs conducive to cooperation. Traulsen et al. [TSS+10] examine how humans update their strategies in the real world. Many works [SPL06b, SP05, GGCFM07] simulate the process on grids or natural graphs that approximate the social networks (scale-free graphs). On scale-free graphs, cooperation is promoted in comparison to highly regular graphs. The reason is that scale-free graphs have large differences between the degrees of the vertices. Highly connected cooperators can tolerate a lot of defection in the neighborhood while spreading cooperation.

A fundamental question in spatial games is the role of the graph structure in boosting or amplifying the desired cooperative behavior. There are two notions of cooperation amplification: (1) an increase of cooperation in the stationary state during coexistence; and (2) an increase in the probability that the cooperators replace all defectors, this event is called cooperator fixation. Most works on spatial games [PJR+17, CFL09, JHK+22] study mainly the aspect that cooperation increases in the stationary state with results on scale-free networks [SPL06b, SP05, GGCFM07]. Moreover, these works focus on empirical results and do not provide theoretical guarantees with analytical results. Our work is distinguished in two aspects. First, as compared to increasing cooperation in the stationary state, we consider fixation, which is more desirable as it ensures complete cooperation in the stationary state. However, note that ensuring fixation is more challenging and usually takes a long time to achieve in simulations. Second, our goal is to establish analytical results that can provide theoretical guarantees on fixation probability. By amplifying cooperation, we mean increasing the fixation probability. This type of amplifying has been studied in many contexts: such as amplification for Moran processes [LHN05, TPCN21, ACN15] where the mutant has constant advantage, or amplification for evolutionary games for weak selection regimes [ALC+17]. Despite the importance of amplifying cooperation, in the context of spatial games in strong selection, the existence of graphs that amplify cooperation has remained elusive even after two decades of active research, which is the question we address.

In this work, we present a classification of density amplifiers in spatial games. A weak density amplifier ensures fixation with a positive probability for a low initial cooperation rate. A strong density amplifier guarantees the fixation probability close to $1$ for a low initial cooperation rate. We show that previously studied structures, such as grids and regular graphs are not even weak density amplifiers. In contrast, we present graph structures that are not only weak but even strong density amplifiers. We demonstrate this amplifying effect across three replicator dynamics: deterministic synchronous, deterministic asynchronous, and randomized asynchronous. Thus, our construction answers the open question in the affirmative for the existence of density amplifiers in spatial games.

Besides answering the open question, we present several other results related to our construction. First, we present two robustness results: our structure ensures high fixation probability even with: (a) a low rate of initial cooperation; and (b) high temptation-reward ratio $T/R$ above $2$. Previous literature in the context of spatial games [NSTF04, ALC+17, SP05, GGCFM07] requires a very low temptation-reward ratio $T/R$ and shows that with a high initial cooperation rate (around 50%) structures can ensure a steady-state of cooperation rate (around 50%). These results neither start with a low cooperation rate (e.g., 5%) and ensure a steady-state cooperation rate of 50%, nor start with a high initial cooperation rate and ensure fixation. In contrast, our structures with a low initial cooperation rate of 5% ensure fixation, which is a significant advancement in the study of spatial games. Second, we show that our structure ensures the fixation quickly. In contrast to many structures in the literature where fixation requires exponentially many steps, see [NM92a, CIJS22], we show that for our structures

the fixation happens within quadratic steps. Third, our construction has degree variations between neighbors, similar to scale-free networks, however, we ensure that the maximal degree is constant (proportional to $T/R$). Finally, we supplement our theoretical guarantees for large population limit by simulation results to show the effectiveness of our structures even for small population sizes. The simulation results consider small population sizes with various ratios of $T/R$ and show that the fixation probability is high even for a small initial density of cooperation and large $T/R$.

The framework of spatial games is quite general, with several regimes, e.g., weak and strong selection and deterministic and stochastic dynamics. Our work provides density amplification in the strong selection regime. The extension of the density amplification for all selection regimes and dynamics are an interesting direction for future work.

## 4.2 Results

### 4.2.1 Model

As in the classical literature [NM92a, SP05], we focus on PD with normalized payoffs, where $R = 1$, $P = \varepsilon$, $S = 0$ and $T = b > 1$. The individuals follow the replicator dynamics. The detailed explanation is in Figure 4.1. The individuals play PD with all neighbors and collect the payoff, then change their strategies based on their and the neighbor's payoffs. We examine the three main variants in replicator dynamics: (1) *Synchronous deterministic*: Every individual that has a neighbor with a higher payoff adopts a strategy of the neighbor with the highest payoff; (2) *Asynchronous deterministic*: One edge is selected randomly, and the individual with a smaller payoff adopts its neighbor's strategy; and (3) *Asynchronous randomized*: One edge is selected randomly. Let $x$ and $y$ be the selected neighbors. If the payoffs are $P_x$ and $P_y$ then $y$ changes its strategy to $x$'s strategy with probability $\frac{1}{1+e^{(P_y-P_x)/K}}$, where $K$ (selection intensity) is the noise parameter, otherwise $y$ adopts $x$'s strategy. See Figure 4.1 (panels d,e,f) for details on replicator dynamics.

The dynamic starts with some cooperators already present in a graph, we denote the initial density of cooperation by $p$. Usually, simulation studies start with large $p$ (around $\frac{1}{2}$). However, ensuring initial configurations with such a large fraction of cooperators is unrealistic. Here, we consider $p$ to be small, which goes to zero in the limit of a large population. This requires the spread of cooperation even from a disadvantageous initial position. Spatial games have rich and complex dynamics. Deciding whether the cooperation survives on a general graph is computationally hard [CIJS22]. They give rise to beautiful patterns even for grids, see Figure 4.2. Given these complicated dynamics, analytical results are challenging to achieve. The landscape of results is dominated by simulations and case studies. The main challenge is to establish results with provable guarantees, which have been elusive.

We say that cooperators *fixate* if they replace all defectors and the probability of this event is called the *fixation probability*. Traditionally, previous works on spatial games consider the density of cooperators in the steady-state of coexistence of cooperators and defectors. The fixation probability is much more desirable as it ensures only cooperators in the steady state.

Now, we define the notion of *density amplifiers*. Given temptation $b > 1$, the initial density of cooperators $p > 0$ and number of vertices $n$, we call a family of graphs: (a) *weak density amplifiers* if the fixation probability for all graphs in the family is above $0$, even in the limit of large $n$; (b) *mild density amplifiers* if the fixation probability for all graphs in the family is

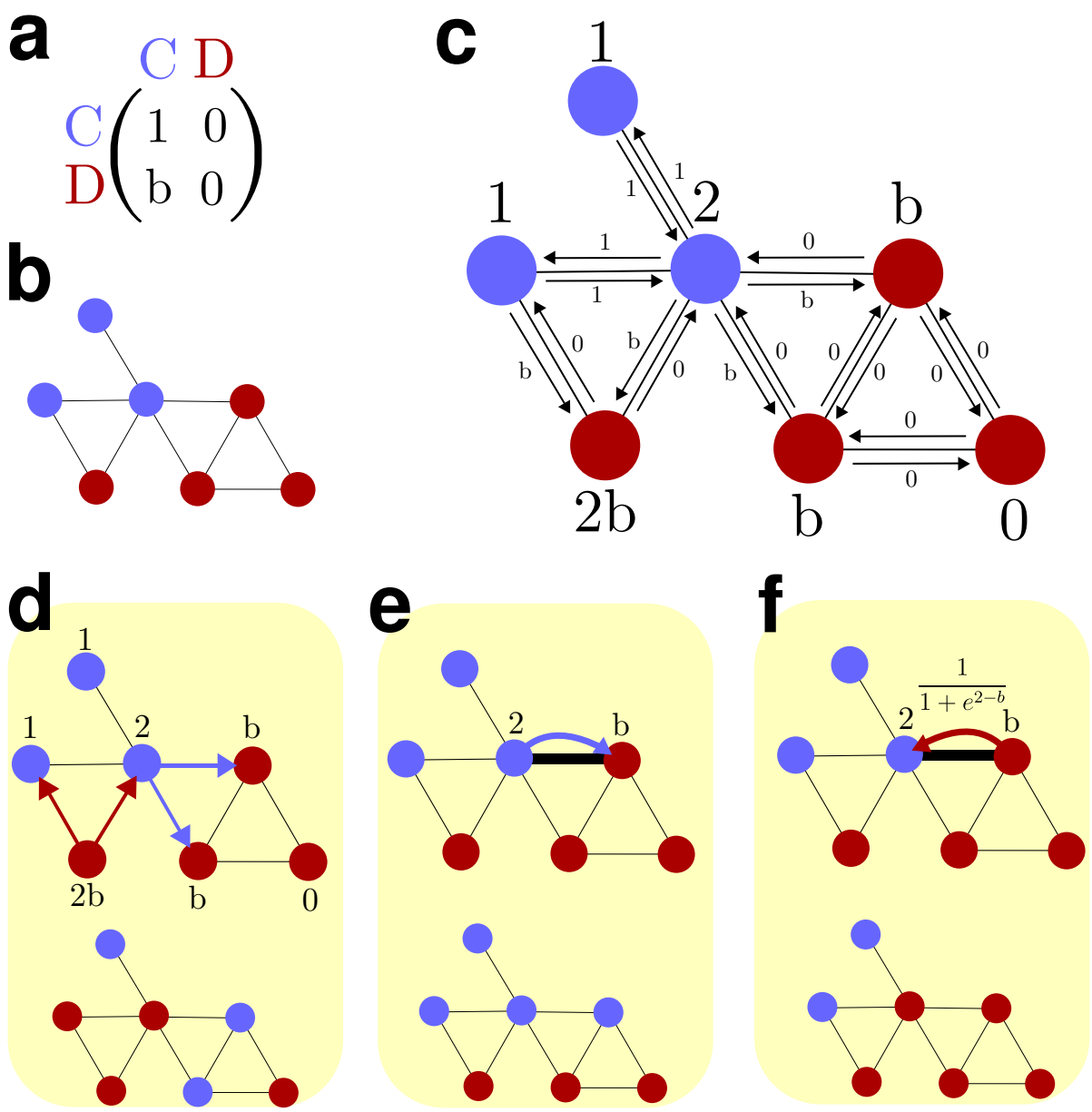



Figure 4.1: **The description of the Prisoner's dilemma on graphs.** **a** The payoff matrix of Prisoner's dilemma, we have $b > 1$ and $\varepsilon = 0$. **b** The initial configuration on a graph, with cooperators being (pale) blue and defectors (dark) red. **c** Payoffs of all players in one turn. The payoff of every cooperator is equal to the number of cooperating neighbors, and the payoff of every defector is $b$ times the number of cooperating neighbors. **d** Synchronous deterministic updating: every individual updates its strategy to the strategy of a neighbor with the highest payoff. The arrows denote the spread of cooperation/defection. **e** Asynchronous deterministic updating: One random edge is selected (highlighted in the figure), and the individual with a higher payoff replaces the individual with a lower payoff. **f** Asynchronous randomized updating: One random edge is selected, and one individual replaces the other with the probability given the Fermi function of the payoffs. In the example, the defector replaces the cooperator with probability $\frac{1}{1+e^{2-b}}$.

higher than the probability that cooperators will become extinct; (c) *strong density amplifiers* if the fixation probability tends to $1$ in the limit of large $n$.

Note that complete graphs (unstructured populations) are not even weak density amplifiers. By design of the Prisoner's dilemma, every defector in the population has a bigger payoff than any cooperator. Moreover, results of [OHLN06] imply that regular graphs are also not even weak density amplifiers for all $b$. Some graphs, for example, stars, are weak density amplifiers: if the central vertex of the star starts as a cooperator, then the cooperation has a chance to replace all defectors. However, if the central vertex starts as a defector, then cooperation disappears. Hence stars are not mild density amplifiers for $p \leq \frac{1}{2}$. The two fundamental open questions are: (a) The existence of mild density amplifiers; and (b) the existence of strong density amplifiers.

### 4.2.2 Analytical Results

To answer positively to both questions about the existence of density amplifiers, we describe two graph families parametrized by $n$, the number of nodes, and $b$, the temptation. Later, we show that graphs $\mathcal{A}^d$ are strong density amplifiers for deterministic (synchronous and asynchronous) settings and graphs $\mathcal{A}^r$ are strong density amplifiers for randomized settings. To simplify the construction we first increase the defectors' payoff against the cooperator by the highest degree times $\varepsilon$ and then round up the payoff to the nearest higher integer. It allows us to treat the payoff of a defector interacting with a defector as $0$. This change gives the defector an advantage by giving them a higher payoff. Finally, if there is a tie in the accumulated payoff, then cooperators win ties. However, note that since we have increased the defectors' payoff, a tie implies a higher payoff for cooperators in the original setting. Thus this procedure still only provides an advantage to defectors.

The structure of $\mathcal{A}^d(n, b)$ is as follows: There are four types of vertices: *big*, *small*, *bridge*, and *leaf*. The graph consists of a path on which big and small vertices alternate. Between two neighboring big and small vertices there are $b - 1$ bridge vertices connected to both of them. Moreover, to every small vertex, $b$ leaf vertices connect, and to every big vertex, $10b^2 - 2b$ leaf vertices connect. The number $n$ bounds the number of big vertices that can be in the graph (around $\frac{n}{10b^2}$). Figure 4.3 shows the structure of $\mathcal{A}^d$ together with the spread of cooperation.

The structure $\mathcal{A}^r(n, b)$ is similar. The graph contains *big*, *small*, and *leaf* vertices. Again, big and small vertices alternate on a path. Every big vertex connects to $10b^2 - 2$ leaf vertices. This time, small vertices neighbor only two big vertices.

We describe the spread of cooperation on $\mathcal{A}^d$ in a deterministic (synchronous or asynchronous) setting. We suppose that $b$ is arbitrary and $n$ sufficiently large. Since the initial density $p$ is above $0$, we know that in the graph, there are a lot of *seeded* vertices. Seeded vertices are big vertices that are cooperators and have at least $b$ cooperating neighbors. Seeded vertices can convert other leaf vertices to cooperation, see Figure 4.3 (panel e,f,g) for a detailed description of the neighborhood of the seeded vertex. With substantial probability, the seeded vertex converts at least $3b^2$ of leaf vertices to cooperators. This cooperator cannot be converted: no neighbor can have a higher payoff than $3b^2$ since no neighbor has more than $3b$ neighbors. We call such a vertex *invincible*. After an invincible vertex appears, the cooperation cannot die out. Moreover, from every position, the probability that the cooperation fixates is nonzero and happens quickly.

Observe that the probability that a big vertex becomes invincible depends only on the neighborhood of the big vertex and is also independent from all events happening at higher

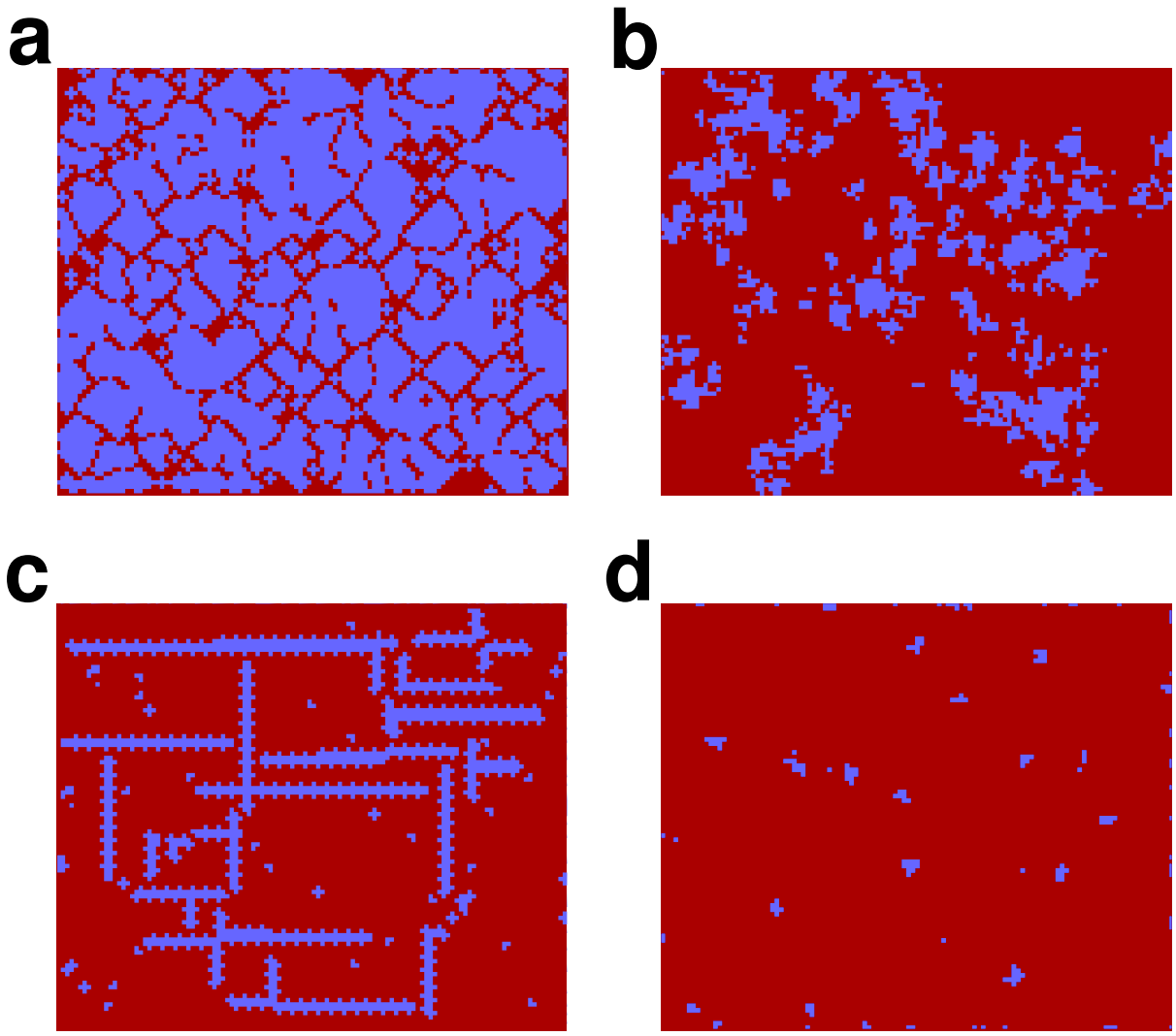



Figure 4.2: **Simulations of the spatial games on the lattice** with unchaengable boundary and size $100 \times 120$. Cooperators are denoted blue, and defectors are red. The initial density of cooperators is $p = \frac{1}{2}$, and the position is recorded after reaching equilibria. **a** Process for $b = 1.3$ and synchronous deterministic updating. **b** Process for $b = 1.99$ and synchronous deterministic updating. Synchronicity and determinism ensure that the cooperators create stable structures that are impossible to invade. **c** Process for $1.03$ and deterministic asynchronous updating. **d** Process for $b = 1.03$ and randomized asynchronous updating. In the randomized updating, cooperators cannot create structures, therefore defectors slowly erode the cooperation.

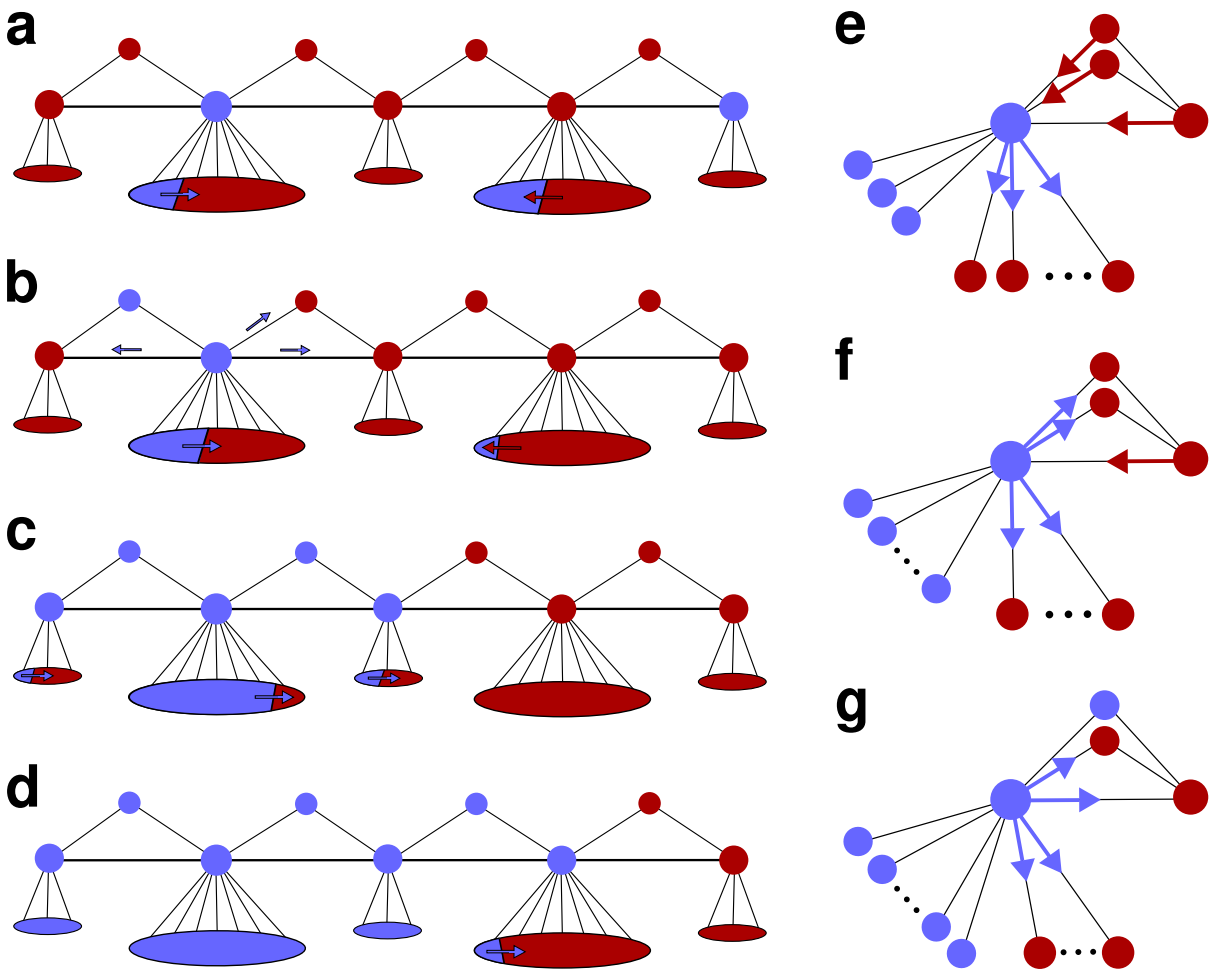



Figure 4.3: **The spread of cooperation in $\mathcal{A}^d$. a** Big and small vertices alternate on a line. One bridge vertex is connected to neighboring small and big vertex, leaf vertices only connect to small and big vertices. More leaf vertices connect to the big vertex. A big vertex that started as a cooperator with a lot of cooperating neighbors spreads the cooperation while the cooperation recedes everywhere else. **b** The cooperating big vertex becomes invincible. Cooperation still recedes everywhere else. **c** The invincible vertex spreads the cooperation to a small vertex, which in turn starts spreading cooperation among the bridge neighbors. **d** The small vertex can convert the big neighbor, and the spread of cooperation continues. **e** The neighborhood of a seeded vertex for $b = 3$. The seeded vertex can convert only leaf neighbors, bridge and small vertices can have a higher payoff. **f** After a big vertex has more than $2b$ cooperating neighbors, it can convert bridge vertices. **g** Invincible vertex has at least $3b^2$ neighbors and can spread the cooperation to any neighbor.

distances in the graph. That means increasing the size of the graph from $n$ vertices to $2n$ vertices increases the probability of invincible vertex appearing (and thus fixation probability) roughly from $\rho$ to $1 - (1 - \rho)(1 - \rho) = 2\rho - \rho^2$. Also, for setting $p = \frac{1}{2}$, the expected number of cooperating neighbors of a big vertex is around $5b^2$, which means that a big vertex initialized as a cooperator is invincible from the start, so the probability of an invincible vertex not appearing is exponentially small.

Similar reasoning holds for randomized asynchronous settings. Some big vertices have a cooperator in a neighborhood. This cooperator has a small but positive chance to spread to the big vertex and then convert all leaf vertices to cooperators. This vertex, so-called

*stronghold*, is more likely to convert another neighboring vertex to a stronghold than to turn to a defector. The spread of cooperators is a walk on a biased Markov Chain. Since the bias is significant, the cooperators spread over the whole graph with a large probability.

In summary, we show that the family of graphs $\mathcal{A}^d(n,b)$ ensures high fixation probability for synchronous deterministic and asynchronous deterministic processes, and family of graphs $\mathcal{A}^r(n,b)$ ensures the same for asynchronous randomized setting, see Section 4.4 for details. In other words, we establish the existence of strong density amplifiers for all three replicator dynamics. Along with the main result proving the existence of strong density amplifiers, our construction has several other desirable properties. First, (a) in contrast to existing simulation studies that consider high $p = 1/2$ [SF07, JHK$^+$22, SP05], our structure ensures high fixation probability even for low $p$, for instance, for $p$ of order of $n^{-\frac{1}{3(b+1)}}$, that goes to 0 as $n$ goes to $\infty$, the fixation probability is at least $1 - e^{-0.2n^{\frac{1}{3}}}$ in both density amplifier families; (b) in contrast to previous studies that consider low $b$, our structure ensures high fixation probability for high $b$, for instance for $b$ that is of order of $\sqrt{\log n}$ that goes to $\infty$ as $n$ goes to $\infty$, the fixation probability is again at least $1 - e^{-0.2n^{\frac{1}{3}}}$ in both density amplifier families. It shows the robustness of our result with respect to low initial cooperation and high temptation. Second, the number of steps to achieve fixation is asymptotically quadratic in the size of the graph for both $\mathcal{A}^d$ and $\mathcal{A}^r$ (see Section 4.4 for details). It means that the fixation is achieved quickly with respect to the network size. Third, our graph structures have constant degree, i.e., the degree is bounded by $10b^2$ even when the population size increases to $\infty$. The constant-degree property is desirable as it ensures that every individual finishes all interactions in constant time in every round.

**Significance of results.** We further emphasize the strength and significance of our results in the following ways:

*Initial cooperation rate.* First, while previous literature for constant or weak selection, considers the probability that a single cooperator fixates, we show in Section 4.4 that for spatial games a lower bound on the initial cooperation rate that is required for the deterministic setting. This complements our initial cooperation density requirement.

*Extinction of defectors.* Our strong density amplifiers ensure a high fixation probability of cooperation even with a low initial cooperation rate. This also ensures that even if there is a relatively high initial density of defectors (say $\frac{1}{2}$) still the defectors become extinct. In other words, our strong density amplifiers ensure high fixation for cooperators and low fixation for defectors.

*Payoff matrix.* While we focus our results on the classic matrix from [Now07, NM92a], we also show the condition under which we obtain strong density amplifiers for the general Prisoners' dilemma payoff matrix. We show that our results about the existence of strong density amplifiers hold in the randomized update for all parameters, and for the deterministic update when $R^2 > T \cdot P$.

*Initialization.* We focus on the randomized initialization of cooperators as this is one of the most difficult conditions where cooperators are not clustered. We argue (in Section 4.4) that our results hold for other initialization, e.g., temperature and correlated initialization.

*Mutation rate.* We present simulation results that consider the cooperation rate over time for various mutation rates, and our structure significantly outperforms star and grid.

**Fixation probability bounds for single mutant.** As mentioned above, in the deterministic setting we show that the initial density of cooperators is required. We also consider the bounds

on fixation probability of a single mutant: a single cooperator among defectors (fixation probability denoted as $\rho_C^n$ in population of size $n$) and a single defector among cooperators (fixation probability of defector denoted as $\rho_D^n$). For $\mathcal{A}^r$ we show that $\rho_C^n$ is constant and $\rho_D^n$ is exponentially small in $n$. In contrast, for complete graph and grid, $\rho_C^n$ is exponentially small in $n$ whereas $\rho_D^n$ is constant, and for star both $\rho_C^n$ and $\rho_D^n$ are proportional to $1/n$. The results are summarized in Table 4.1.

| | $\rho_C^n$ | $\rho_D^n$ | $\rho_C^n \bowtie \rho_D^n$ | $A_n = \frac{\rho_C^n}{\rho_D^n}$ | $\lim_{n\to\infty} A_n$ | $B_n = \frac{\rho_C^n}{\rho_n}$ | $\lim_{n\to\infty} B_n$ |
|---|---|---|---|---|---|---|---|
| Complete Graph | $2^{-c_1 n}$ | $c_2$ | $<<$ | $2^{-c_1 n}/c_2$ | $0$ | $n2^{-c_1 n}$ | $0$ |
| Grid | $2^{-c_3 n}$ | $c_4$ | $<<$ | $2^{-c_3 n}/c_4$ | $0$ | $n2^{-c_3 n}$ | $0$ |
| Star | $c_5 n^{-1}$ | $c_6 n^{-1}$ | $\approx$ | $c_5/c_6$ | $c_5/c_6$ | $c_5$ | $c_5$ |
| $\mathcal{A}^r$ | $c_7$ | $2^{-c_8 n}$ | $>>$ | $c_7 2^{c_8 n}$ | $\infty$ | $c_7 n$ | $\infty$ |

Table 4.1: Fixation probabilities for a single random mutant in graph occupied by individuals of the other type: The first column describes the graph type, the second and third column represents the fixation probability of a single cooperator and defector, respectively. In the third column, we compare the two fixation probabilities where $\bowtie$ is the comparison operator, which can be very small $<<$, comparable $\approx$, or very large $>>$. The fourth and fifth columns consider the ratio $A_n$ of the two fixation probabilities and the respective large-population limit. The sixth and seventh columns consider the ratio $B_n$ of the fixation probability of a cooperator to the fixation probability of a neutral mutant, denoted $\rho_n$, which is $1/n$, and the respective large-population limit. In the table, $c_1$, $c_2$, ..., $c_8$ denote constants that are independent of $n$. The table summarizes the following: (a) for complete graph and grid, the fixation probability of cooperators is exponentially small in $n$ whereas for defectors it is constant; (b) for star, the fixation probabilities are proportional to $1/n$; and (c) for $\mathcal{A}^r$ the fixation probability of cooperators is constant, whereas for defectors it is exponentially small. The large-population limit of the desired comparison ratios vanishes for the complete graph and grid, is constant for the star, and goes to $\infty$ for $\mathcal{A}^r$. Observe that only $\mathcal{A}^r$ yields that the desired ratio goes to $\infty$ in the large-population limit.

### 4.2.3 Simulation results

Finally, we supplement our theoretical findings with simulation results. While the theory provides guarantees for large population sizes, the simulation results demonstrate the effectiveness of our structure even for small population sizes. We examine three baselines: complete graph, grid, and star. No matter the process and the temptation ($b$), the fixation probability on these graphs does not change much. On a complete graph, every defector has a higher payoff than every cooperator, which means the fixation probability tends to $0$. Similarly, on a grid, the last defector has a higher payoff than all neighbors, so it cannot be converted in a deterministic setting. In a randomized setting, the fixation probability is also $0$. On a star, the probability that the center is a cooperator is proportional to $p$. If at the same time $p \geq \frac{b}{n-1}$, the central cooperator has a higher payoff than neighbors and can spread. This means the fixation probability is proportional to $p$.

We run the asynchronous deterministic setting for three temptations and graphs: $\mathcal{A}^d(10^4, 2)$, $\mathcal{A}^d(10^4, 4)$, and $\mathcal{A}^d(10^4, 6)$. We examine how the initial cooperator density $p$ influences the fixation of cooperation in $\mathcal{A}^d$ against the baseline graphs. For every graph and value of $p$, we report the fixation probability by averaging over $5 \cdot 10^4$ runs. Every run is simulated until one of

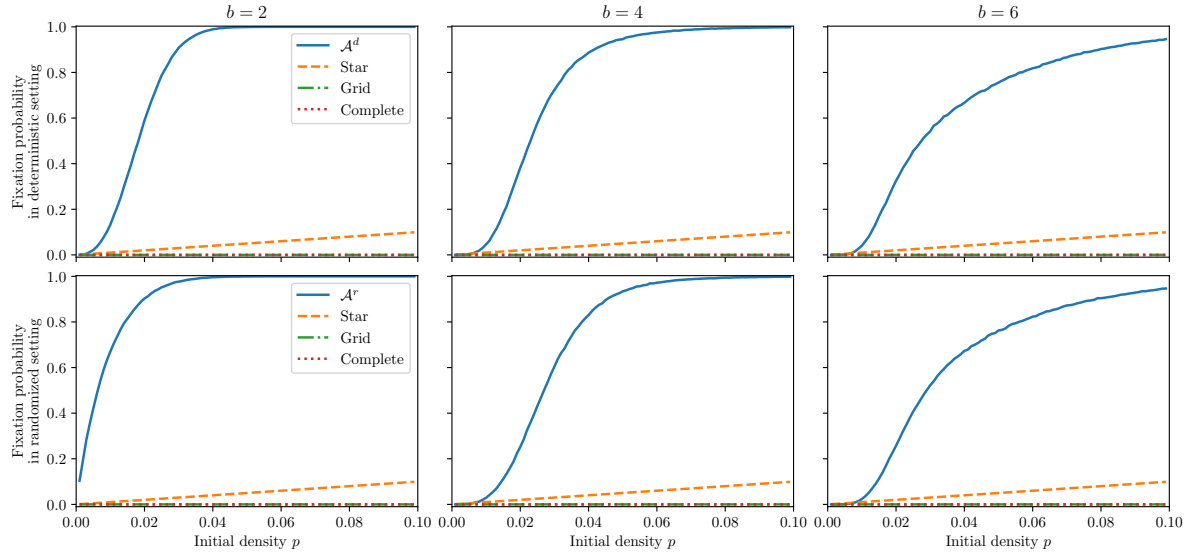



Figure 4.4: **Simulation results.** We examine the fixation probability (y-axis) with respect to the starting density of cooperation $p$ (x-axis) for four graphs: a density amplifier, a star, a grid, and a complete graph. The rows examine asynchronous deterministic and asynchronous randomized setting, the collums examine different temptation $b \in \{2, 4, 6\}$. All graphs have $10^4$ vertices and the results are averaged over $5 \cdot 10^4$ runs. Cooperators on the grid and complete graph have fixation probability $0$. On star, the fixation probability is proportional $p$. The fixation probability on both density amplifiers quickly reaches $1$ for $p = 0.04$ already.

the two things happens: cooperators become extinct, or cooperators create an invincible vertex (this is a big vertex with $3b^2$ cooperating neighbors). Since the invincible vertex cannot be converted and cooperators eventually fixate, these two conditions are equivalent to simulating until cooperators or defectors spread over the whole graph. The first row of Figure 4.4 shows the results. We see that for all settings, the fixation probability increases with increasing initial probability. At $p = 0.04$, the fixation probability on $\mathcal{A}^d(10^4, 2)$ is already almost $1$. The fixation probability for other graphs rises more slowly since the temptation to defect is larger, and the graphs $\mathcal{A}^d(10^4, 4)$ and $\mathcal{A}^d(10^4, 6)$ have fewer big vertices than the graph $\mathcal{A}^d(10^4, 2)$.

For the asynchronous randomized setting, we again examine three graphs: $\mathcal{A}^r(10^4, 2)$, $\mathcal{A}^r(10^4, 4)$, and $\mathcal{A}^r(10^4, 6)$. We examine how the initial cooperator density $p$ influences the fixation of cooperation in $\mathcal{A}^r$ against the baseline graphs. We run the process for every combination of parameters $5 \cdot 10^4$ times until either cooperation or defection fixate and report the fixation probability. In the second row of Figure 4.4, as our result suggests, we see that cooperation increases even more steeply than in the deterministic setting. At $p = 0.02$, the fixation probability on $\mathcal{A}^r(10^4, 2)$ is already almost $1$.

## 4.3 Discussion

In this work, we present $\mathcal{A}^d$ and $\mathcal{A}^r$, the first graph families that increase the fixation probability of cooperation in spatial games with Prisoners' dilemma with strong selection. Moreover, the graph families ensure the fixation with several desirable properties: (a) they are robust with respect to a low initial density of cooperation and high temptation; (b) they have a constant

degree; and (c) they ensure a fast spread of cooperation within quadratic steps in the size of the network.

First, the study of evolutionary graph theory for constant (or frequency-independent) fitness has been widely studied in the context of the Moran process [LHN05]. The role of graphs that can amplify the fixation probability has been a key topic of interest [GGG+17, PTCN18]. Moreover, the time to fixation is another important aspect [FRT13], and there is a very interesting tradeoff [TPCN19]. The existence of strong density amplifiers with fast fixation time has been established in [TPCN21]. However, the techniques of these works do not extend to evolutionary games, which represent frequency-dependent selection.

Second, the study of evolutionary games on graphs, which represent frequency-dependent selection, also received broad attention [OHLN06, NM92a]. The computational hardness of such games has been established in [IJCN15, CIJS22]. In the regime of weak selection, the role of graphs that help in increasing cooperative behavior has been considered by Allen et al. [ALC+17], and by Fotouhi et al. [FMAN18]. While Allen et al. give the theoretical algorithms and guarantee, Fotouhi et al. present natural graphs and empirical results. However, these results do not extend to strong selection, where the existence of density amplifiers is an interesting open question. Spatial games intuitively represent the strong selection limit, and our results complement the existing results in the literature establishing the existence of strong density amplifiers in this regime.

Finally, we believe that the structures we present have wider applicability. The reasoning in our proof can be extended to other two-player matrix games (dilemmas), such as snowdrift and stag-hunt games. Even in these cases, pro-social behavior creates invincible parts of the graph. Our structures will be useful in problems where there is a hard-to-reach absorbing state that is desired. Exploring the role of our structures in specific applications is an interesting direction for future work.

In this section, we describe the main ingredients for both proofs.

### 4.3.1 Proof sketch for asynchronous deterministic updating

The graph family $\mathcal{A}^d$ are strong density amplifiers. The proof has three main ingredients. First, we show that with substantial probability there are at least $n^{\frac{1}{3}}$ big vertices that are cooperators and have at least $b$ cooperating neighbors (we call them seeded vertices). Second, we prove that one seeded vertex becomes invincible with a large probability. Third, we describe how one invincible vertex can convert the whole graph to cooperation.

A vertex starts as a cooperator with probability $p$ and we suppose that

$$p \geq \frac{1}{n^{1/(3b+3)}}\,. \tag{4.1}$$

Let us examine the probability that a big vertex is a cooperator and has at least $b$ cooperating leaf neighbors. With probability $p$, the big vertex is a cooperator. It has $10b^2 - 2b$ leaf neighbors and the number of cooperating neighbors follows a binomial distribution, which gives the probability

$$p \cdot \sum_{x \geq b} p^x (1-p)^{10b^2-2b-x} \binom{10b^2-2b}{x}\,. \tag{4.2}$$

We split the graph to $n^{\frac{1}{3}}$ parts. In one part, using $b \geq 2$, Equation (4.1), and Equation (4.2), we have that there is no seeded vertex with probability at most

$$\left(1 - \frac{10b^2 + b + 2}{n^{1/3}}\right)^{\frac{n^{2/3}}{10b^2+b+2}} < e^{-2n^{1/3}}. \tag{4.3}$$

From union bound and Equation (4.3), we get that in every part of the graph, there is at least one seeded vertex (that means at least $n^{1/3}$ seeded vertices in the whole graph) with probability at least

$$1 - e^{-n^{1/3}}. \tag{4.4}$$

The seeded vertex has a payoff big enough to convert any leaf to cooperation. In the worst case, the neighboring small and bridge vertices can revert the seeded vertex to defection. In one active step, the probability of the big vertex being turned into defection is at most $\frac{2b}{10b^2-4b} < \frac{1}{4b}$. When the seeded vertex has at least $2b$ cooperating leaf neighbors, the bridge vertices stop being threatening. The failure probability in one active step decreases to $\frac{2}{9b^2-2(b-1)} < \frac{1}{4b^2}$ until the big vertex has $b^2$ cooperating neighbors. Finally, the failure probability is at most $\frac{2}{7b^2-2(b-1)} \leq \frac{1}{3b^2}$ in one step until the big vertex has at least $3b^2$ cooperating neighbors. At that point, it becomes invincible, no neighbor can have a higher payoff, so it cannot be converted. From union bound, the seeded vertex becomes invincible with probability at least

$$\left(1 - b \cdot \frac{1}{4b}\right)\left(1 - b^2 \cdot \frac{1}{4b^2}\right)\left(1 - 2b^2 \cdot \frac{1}{3b^2}\right) \geq \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{3} \geq \frac{3}{16}. \tag{4.5}$$

One invincible vertex can convert the rest of the graph to cooperation. It first converts all bridge vertices and neighboring small vertex. This small vertex converts its leaf neighbors, then the neighboring big vertex and remaining bridge vertices. By this, the big vertex has at least $b$ cooperating neighbors and can start converting other vertices and becomes invincible. The probability of everything happening is nonzero and the invincible vertex cannot be converted, which means it happens eventually. To increase the number of invincible vertices by one, we need the number of steps around the newly invincible vertex to be around $\mathcal{O}((20b)^{2b})$. This means the number of steps until all vertices are converted to cooperators is

$$\mathcal{O}\left((20b)^{2b} n^2\right). \tag{4.6}$$

That means, for initial rate of cooperation at least $p \geq \frac{1}{n^{1/(3b+3)}}$ (Equation (4.1)), from Equation (4.4) and Equation (4.5), the fixation probability is at least

$$\left(1 - e^{-n^{1/3}}\right) \cdot \left(1 - \left(\frac{3}{16}\right)^{n^{1/3}}\right) \geq 1 - e^{-0.2n^{1/3}}. \tag{4.7}$$

Note that for $b$ around $\sqrt{\log n}$ (which is unbounded if $n$ grows), the initial density of cooperation is around $e^{-\sqrt{\log n}}$, which tends to $0$ as $n$ grows. However, the theorem still holds and guarantees fixation probability that goes to $1$ with growing $n$.

### 4.3.2 Proof sketch for asynchronous randomized updating

The graph family $\mathcal{A}^r$ are strong density amplifiers. The proof consists of two important steps. First, we show that a big vertex that is seeded (has one cooperator in the neighborhood)

becomes a stronghold (all its leaf neighbors are cooperators) with a large probability. Second, we estimate the probability that a stronghold vertex converts the rest of the graph.

We split the spread of cooperation into several stages. In the first stage, the big vertex $v$ becomes a cooperator. Let $i$ denote the number of cooperating neighbors of $v$. In the second stage, we have $i \leq b$, in the third $b < i \leq 3b$, and in the fourth $3b < i$.

The first stage succeeds with probability

$$\frac{1}{1+e^{b-0}}\,. \tag{4.8}$$

During the second and third stages, we suppose that the edge between the small and big vertex was not selected. This happens with probability at most

$$\left(1-\frac{2}{10b^2-3b}\right)^{3b} > 1-\frac{6b}{10b^2-3b} > 1-\frac{1}{\frac{17}{12}b}\,. \tag{4.9}$$

With probability $\frac{1}{1+e^{b-i}}$ after selecting the edge between leaf and big vertex, the cooperation spreads, that means the second phase succeeds with the probability

$$\prod_{i=0}^{b}\frac{1}{1+e^{b-i}}\,. \tag{4.10}$$

By the same reasoning, the third phase succeeds with probability

$$\prod_{i=b+1}^{3b}\frac{1}{1+e^{b-i}} > \prod_{i=1}^{2b}\frac{1}{1+e^{-i}}\,, \tag{4.11}$$

and fourth with

$$\prod_{i=3b+1}^{10b^2}\frac{1}{1+e^{2b-i}} = \prod_{i=0}^{10b^2}\frac{1}{1+e^{-b-1-i}} > \prod_{i=0}^{\infty}\frac{e^{b+1+i}}{1+e^{b+1+i}}\,. \tag{4.12}$$

Combining Equation (4.9), Equation (4.10), Equation (4.11), and Equation (4.12), we get the success probability

$$2^{-5}e^{-b^2}\,. \tag{4.13}$$

Having a stronghold in the graph and selecting an edge between the stronghold and neighboring small vertex means that either stronghold becomes a defector (with probability roughly $\frac{1}{1+e^{10b^2-2}}$) or the neighboring big vertex becomes seeded and it might become a stronghold with probability $2^{-5}e^{-b^2}$. If we observe the number of strongholds, their expansion creates a Markov Chain with the ratio between increasing and decreasing cooperation is above

$$e^{6b^2}\,. \tag{4.14}$$

Suppose that $p \geq \frac{e^{b^2}2^3\cdot 10b^2}{n^2 3}$, then a big vertex is seeded with probability at least $2^5 p$ and from Equation (4.13), we have that a big vertex becomes a stronghold with probability at least $e^{-b^2}p$. With probability at least $1-e^{\frac{49}{16}n^{1/3}}$, at least $n^{1/3}$ seeded vertices turns into stronghold. In the Markov Chain tracking the number of strongholds (with the ratio Equation (4.14)), the strongholds disappear with probability at most $e^{-6b^2n^{1/3}}$. Therefore the fixation probability is at least

$$1-e^{-3n^{1/3}}\,. \tag{4.15}$$

Again, note that we can set $b=(\log n)^{1/3}$ which grows with $n$, and then $p$ tends to $0$ with large $n$ and the fixation probability tends to $1$.

## 4.4 Additional proofs

This is supporting information to manuscript *Density Amplifiers of Cooperation for Spatial Games*. The organization of the text is as follows: Section 4.4.1 introduces the model of spatial games of social dilemma. Section 4.4.2 describes the structure of strong density amplifiers for deterministic and randomized settings. Section 4.4.3 proves that $\mathcal{A}^d$ is a family of strong density amplifiers for the deterministic setting. Section 4.4.4 shows that $\mathcal{A}^r$ is a family of strong density amplifiers for the randomized setting. Section 4.4.5 discusses the smallest viable initial density of cooperators in the deterministic setting. Section 4.4.6 determine the range of parameters of Prisoners' dilemma for which our structures support cooperation. Section 4.4.7 discusses changes of fixation probabilities with different initializations. Section 4.4.8 examines the fixation probability of one mutant (cooperator or defector) inside a graph occupied by the other type. Section 4.4.9 explores the model with mutation.

### 4.4.1 Model of Spatial Games of Social Dilemma

In this section, we introduce the model of spatial games of social dilemmas. The deterministic update follows definitions introduced by Nowak [NM92a]. The asynchronous updates are similar to ones introduced in [Now07, HS98].

#### Preliminaries

**Spatial Games.** Spatial games track the spread of strategies in a structured population. Individuals are arranged on a graph. Two individuals are connected by an edge if they interact with each other. In one step of the process, every individual plays a game with all neighbors and receives a payoff based on his and the neighbor's strategy. Based on the payoff, some individuals update their strategies.

**Prisoners dilemma.** The game that introduces tension between selfishness and global optimality is The Prisoner's Dilemma. There are two strategies: cooperation (C) and defection (D). The game is given by the following matrix:

| | C | D |
|---|---|---|
| C | $1$ | $0$ |
| D | $b$ | $\varepsilon$ |

The parameter $b$ ($b > 1$) can be arbitrarily large, and we call it temptation. We suppose that $\varepsilon$ is a small positive constant.

The payoff of a cooperator equals the number of cooperating neighbors, and the payoff of a defector is $b$ times the number of cooperating neighbors plus $\varepsilon$ times the number of neighboring defectors (this corresponds to pp-goods from [MRH21]).

**Initialization.** The process is parametrized by $p$: initial density of cooperators. At the beginning of the process, every individual is a cooperator with probability $p$ and a defector with probability $1 - p$.

**Update Notions.** There are various notions [SF07, HS98] of how an individual can update its strategy in one step of the process.

- *Synchronous deterministic*: Every individual who has a neighbor with a higher payoff updates its strategy to the strategy of the highest neighbor.
- *Asynchronous deterministic*: One edge is selected randomly, and the individual with a smaller payoff adopts a strategy of its neighbor.
- *Asynchronous randomized*: One edge is selected randomly. Let it be between individuals $x$ and $y$. Let the edge be between $x$ and $y$ where individuals have payoffs $P_x$ and $P_y$ respectively. Then, the individual at $y$ changes its strategy to $x$'s strategy with probability

$$\frac{1}{1+e^{\frac{P_y-P_x}{K}}},$$

where $K$ is the noise parameter (with probability $\frac{1}{1+e^{\frac{P_x-P_y}{K}}}$, $x$ changes its strategy to $y$'s strategy)

Note that when $K \to 0$, the asynchronous random model changes to deterministic. For ease of notation, we use $K = 1$ in our computations, different values of $K$ do not change our result significantly.

**Fixation probability.** We say a type *fixates* if it replaces all types (i.e., eventually all individuals are of the type). Given a graph $G$, strength of defection $b$ and initial cooperator density $p$, we denote $\rho_C(G, b, p)$ the *fixation probability* i.e. the probability that the cooperators fixate. Similarly, we denote $\rho_D(G, b, p)$, the probability that defectors fixate. We call it *extinction probability* of cooperators. We can view $\rho_C(G, b, p)$ (resp., $\rho_D(G, b, p)$) as the final expected density of cooperators (resp., defectors) given the initial density of cooperators is $p$.

**Density amplifiers.** Given $b > 1$ and $p > 0$, we call a family of graphs $\mathcal{G}$ parametrized by $n$, number of vertices

- *Weak density amplifiers* if the fixation probability for all graphs in $\mathcal{G}$ is higher than $0$ even for large $n$, i.e., for all $G_n \in \mathcal{G}$ holds $\lim_{n\to\infty} \rho_C(G_n, b, p) > 0$.
- *Mild density amplifiers* if for all graphs in $\mathcal{G}$ the fixation probability of cooperation is higher than the extinction probability, i.e., for all $G_n \in \mathcal{G}$ holds $\rho_C(G_n, b, p) > \rho_D(G_n, b, p)$.
- *Strong density amplifiers* if the fixation probability of graphs in $\mathcal{G}$ tends to $1$ with increasing $n$, i.e., for all $G_n \in \mathcal{G}$ holds $\lim_{n\to\infty} \rho_C(G_n, b, p) = 1$.

Note that for weak density amplifiers, we can have two baseline comparisons for density amplification: fixation probability at least $1/N$, which is the neutral drift, or exponentially small in $N$, which is the fixation probability for complete graphs in games. Our definition is stronger in the sense that it requires the fixation to be higher (i.e., constant) in the limit of a large population.

### Existence of Density Amplifiers and Open Questions

In this section, we state some negative results and open questions about the existence of density amplifiers.

**Complete and regular graphs.** We discuss the amplifying properties of complete and regular graphs.

**Complete graphs:** Complete graphs are not weak density amplifiers for all $b > 1$, $p < 1$, and any process. A defector has a higher payoff than all the cooperators. In a deterministic setting, this defector cannot be converted. In a randomized setting spreading defection is always more probable than spreading cooperation. That means $\lim_{n\to\infty} \rho_C(K_n, b, p) = 0$.

**Regular graphs:** Regular graphs are not weak density amplifiers for $b > 1$ and $p < 1$ for the asynchronous deterministic process. For the sake of contradiction, suppose that the cooperators fixate, that means the last defector is replaced in the graph. In the $d$-regular graph, this defector has a payoff $d \cdot b$, and all neighboring cooperators have a payoff $d - 1$. The payoff of the defector is bigger, hence it cannot be converted, which means $\rho_C(G, b, p) = 0$.

**Discussion on a star.** Stars are graphs with one central vertex connected to all other vertices called leaves. Leaves are connected just to the central vertex. Proposition 1 shows that stars are weak density amplifiers, but not mild density amplifiers.

**Proposition 1.** *For synchronous and asynchronous deterministic updating, stars are weak density amplifiers for all $b > 1$ and $p > 0$, but not mild density amplifiers for $p < \frac{1}{2}$.*

*Proof.* On average, at the start, $pn$ cooperators are initialized. If the center is occupied by a cooperator and the number of cooperators overall is at least $b + 1$, then the central cooperator has a bigger payoff than defectors in the leaves. In the deterministic setting, this means this cooperator eventually converts all defectors. In the randomized setting, after selecting an edge between the center and a defector, the failure probability is $1/(1 + e^{i-b})$ where $i$ is the number of the cooperator neighbors. That means the probability that the cooperators lose the center is

$$\sum_{j=i}^{n} \frac{1}{1 + e^{i-b}} < 1 - e^{-i+b}.$$

Here $i$ is the number of initialized cooperators in the leaves. The center is initialized as a cooperator with probability $p$, if $pn = i > b + 1$, then the fixation probability is $p$, which makes stars weak density amplifiers.

On the other hand, if $p < \frac{1}{2}$, with probability $1 - p$, the central vertex is a defector. All cooperators have payoff $0$, they reside in leaves and neighbor only a defector. From the same argument as above, the center is hard to replace. That means the defector converts all cooperators, therefore $\rho_C(G, b, p) < \rho_D(G, b, p)$ for $G$ star, any $b$ (even $b > 0$) and $p < \frac{1}{2}$. □

In a randomized setting, one can show similar results.

**Open questions.** We know that there are weak density amplifiers, which are not mild density amplifiers. The remaining questions are:

- Does there exist a family of mild density amplifiers for all $b > 1$ and $p > 0$?
- Does there exist a family of strong density amplifiers for all $b > 1$ and $p > 0$?

**Organization.** In Section 4.4.2, we show two families of strong density amplifiers for asynchronous deterministic and randomized updating. In Section 4.4.3 and Section 4.4.4, we prove that these graphs are indeed strong density amplifiers.

### 4.4.2 Construction of Strong Density Amplifiers

In this section, we construct strong density amplifiers. We use some properties of stars. In Section 4.4.2, we describe strong density amplifiers for synchronous deterministic updating. In Section 4.4.2, we describe strong density amplifiers for asynchronous deterministic updating.

**Note on temptation.** To simplify the construction and deal with $\varepsilon > 0$, we increase the defector's payoff against cooperators by $d\varepsilon$, where $d$ is the highest degree and then round up the number to the nearest higher integer. This gives the defector an advantage by giving them a higher payoff and allows us to treat $\varepsilon$, the payoff of the defector against the defector as $0$. Finally, if there is a tie in the accumulated payoff, then cooperators win ties. However, note that since we have increased the defectors' payoff, a tie implies a higher payoff for cooperators in the original games. Thus this procedure still only provides an advantage to defectors. In other words, we show that our structures ensure strong density amplification even giving further advantage to defectors, and this further emphasizes the robustness of our construction.

#### Strong Density Amplifiers for Deterministic Setting

We construct a family of graphs $\mathcal{A}^d(n, b)$ with parameters: $n$, number of vertices; and $b$, the temptation, see Figure 4.5. The construction consists of multiple stars of alternating sizes connected in a path. Moreover, special *bridge* vertices ensure that cooperation spreads between the centers of stars.

We have four types of vertices in $\mathcal{A}^d(n, b)$. They can be classified by the number of neighbors.

**Big vertices:** They have $10b^2$ neighbors and are the basic building block in the spread cooperation.

**Small vertices:** They have $3b$ neighbors and create a buffer between two big vertices while being able to spread cooperation.

**Bridge vertices:** They have $2$ neighbors, they are one big and one small vertex that are neighbors. They ensure that a new cooperator in these vertices can start converting its neighbors.

**Leaves:** They have one neighbor and support the cooperators in big or small vertices.

We call big and small vertices *central* to refer to them in a group.

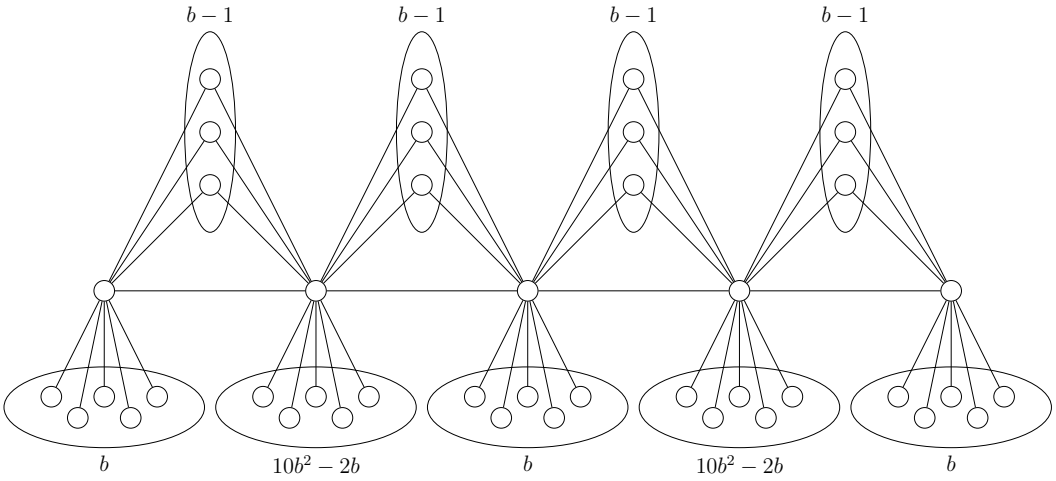



Figure 4.5: The structure of $\mathcal{A}^d$, the density amplifier for the deterministic setting.

We construct the graph as follows: There is a path of alternating big and small vertices. There are $b-1$ bridge vertices connected to every two neighboring central vertices. Leaves are added to the central vertices to ensure the degree constraints ($10b^2$ for big and $3b$ for small).

**Theorem 10.** *For synchronous and asynchronous deterministic updating, all $b > 1$, and $p > 0$, graphs $\mathcal{A}^d(n, b)$ are strong density amplifiers.*

Section 4.4.3 is dedicated to prove Theorem 10.

#### Strong Density Amplifiers for Randomized Setting

We construct a family of graphs $\mathcal{A}^r(n, b)$ with parameters: $n$, number of vertices; and $b$, the temptation, see Figure 4.6. Density amplifier $\mathcal{A}^r(n, b)$ has three types of vertices classified by the number of neighbors. Since some terminology translates from deterministic updating, we name them similarly.

**Big vertices:** They have $10b^2$ neighbors and are the basic building block in the spread the cooperation.

**Small vertices:** They have $2$ neighbours, both of them are big vertices. They allow the spread of cooperation between big vertices.

**Leaves:** They are connected only to a big vertex and support the cooperator.

We construct the graph as follows: There is a path of alternating big and small vertices. Leaves are added to the big vertices to ensure degree $10b^2$.

**Theorem 11.** *For asynchronous randomized updating, all $b > 1$, and $p > 0$, graphs $\mathcal{A}^r(n, b)$ are strong density amplifiers.*

Section 4.4.4 is dedicated to prove Theorem 11.

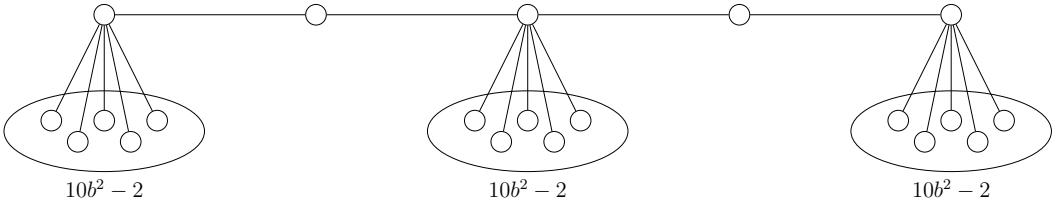



Figure 4.6: The structure of $\mathcal{A}^r$, the density amplifier in the randomized setting.

### 4.4.3 Proof of Theorem 10

**Seeded and invincible vertex.** Given a graph $\mathcal{A}^d(n, b)$, we say that a big vertex is *seeded* if it is a cooperator and has at least $b$ cooperating neighbors among leaves. We say that a big vertex is *invincible* if it is a cooperator and at least $3b^2$ leaves are cooperators.

Observe that an invincible vertex $v$ cannot become a defector. All neighbors of $v$ have a smaller degree than $3b$, so no matter the arrangement of cooperators and defectors, no neighbor of $v$ has a higher payoff than $3b^2$.

**Spread of cooperators.** Here, we describe the spread of cooperation in $\mathcal{A}^d(n, b)$ on a high level. Later, we compute the exact probabilities.

1. With a substantial probability, many big vertices are seeded (Lemma 16).
2. These seeded vertices convert leaves until at least one seeded vertex becomes invincible. One seeded vertex becomes invincible with high probability (Lemma 17).
3. Since the invincible vertex cannot convert to defection and construction of $\mathcal{A}^d(n, b)$, the cooperation spreads over the whole graph quickly (Lemma 18).

The lemmas connect in Lemma 19, which is used to prove Theorem 10.

**Lemma 16.** *For graph $\mathcal{A}^d(n, b)$ and $p \geq \frac{1}{\left(\frac{n}{t^2}\right)^{\frac{1}{b+1}}}$ where $t \geq 1$, the probability that at least $t$ big vertices are seeded is at least*

$$1 - e^{-t}.$$

*Proof.* First, we split the graph into $t$ equal parts and bound the probability that no vertex in a given part is seeded.

A big vertex is seeded if it is a cooperator (probability $p$), and among its leaf neighbors, there are at least $b$ cooperators. The overall probability is:

$$p \cdot \sum_{x \geq b} p^x (1-p)^{10b^2 - 2b - x} \binom{10b^2 - 2b}{x}$$

Rearranging and forgetting summands with $x > b$, we have

$$p^{b+1} \cdot \binom{10b^2 - 2b}{b} \cdot (1-p)^{10b^2 - 3b} > p^{b+1} \cdot (1-p)^{10b^2 - 3b} \cdot \left(\frac{10b^2 - 2b}{b}\right)^b$$

Since $b \geq 2$ and $p$ is small relative to $b$, we can bound $(1-p)^{10b^2-3b} \cdot \left(\frac{10b^2-2b}{b}\right)^b < 2(10b^2 + b + 2)$.

There are at least $\frac{n}{t(10b^2+b+2)}$ big vertices in one part. The probability that no big vertex is seeded is at least

$$\left(1 - p^{b+1} \cdot 2(10b^2 + b + 2)\right)^{\frac{n}{t(10b^2+b+2)}}.$$

After plugging in $p \geq \frac{1}{\left(\frac{n}{t^2}\right)^{\frac{1}{b+1}}}$, we have a lower bound on the probability that no big vertex is seeded. It is at most

$$\left(1 - \frac{t^2 \cdot (10b^2 + b + 2)}{n}\right)^{\frac{n}{t(10b^2+b+2)}} < e^{-2t}.$$

The probability that no big vertex in one part is seeded is $e^{-2t}$. From union bound, we know that there is one part where no big vertex is seeded with probability at most $te^{-2t}$. Since $t \geq 1$, we have $te^{-2t} < e^{-t}$. That means in every part there is at least one seeded big vertex with probability at least $1 - e^{-t}$. □

**Lemma 17.** *For graph $\mathcal{A}^d(n, b)$ and any position that contains a seeded vertex $v$, the probability that $v$ becomes invincible is at least $\frac{3}{16}$.*

*Proof.* At the beginning, the seeded vertex can be replaced by a defector in a bridge vertex or a small vertex. After the seeded vertex has at least $2b$ cooperating leaf neighbors, it can be replaced only by a defector in a small vertex. We split the proof into two parts. First, we bound the probability that $v$ is replaced by a bridge or small vertex before it has $2b$ cooperating leaf neighbors. Second, we bound the probability that $v$ is replaced by a small vertex before it becomes invincible. During the process, we focus on events around $v$. In every step, we suppose the worst configuration in the neighborhood of $v$.

Before $v$ has $2b$ cooperator leaf neighbors, the success is choosing an edge between a defector leaf and $v$. This happens with probability at least $\frac{10b^2-2b-2b}{10b^2}$. The failure is to select an edge between the bridge or small vertex and $v$. This happens with probability $\frac{2b}{10b^2}$. (With probability at most $\frac{2b}{10b^2}$ edge between two cooperators is selected and nothing changes.) In one active step, the probability of failure is at most $\frac{2b}{10b^2-4b} = \frac{1}{5b-2} < \frac{1}{4b}$ since $b \geq 2$. The probability of failure in $b$ active steps is at most $\frac{1}{4}$ from the union bound.

In the next steps, the success is selecting an edge between the defector leaf and $v$. The failure is to select an edge between a small vertex and $v$. If $v$ has $i$ cooperating neighbors, the success happens with probability $\frac{10b^2-2b-i}{10b^2-2(b-1)-i}$, the failure with probability $\frac{2}{10b^2-2(b-1)-i}$.

While $2b < i \leq b^2$, in one step, the failure happens with probability at most $\frac{2}{9b^2-2(b-1)} < \frac{1}{4b^2}$ since $b \geq 2$. There are $b^2$ steps, and the probability of failure is at most $\frac{1}{4}$ from union bound.

While $b^2 < i \leq 3b^2$, the failure probability in one step is at most

$$\frac{2}{7b^2 - 2(b-1)} < \frac{1}{3b^2}.$$

Again, from the union bound, the probability of failure in $2b^2$ steps is at most $\frac{2}{3}$.

That means the probability that a seeded vertex is converted to the invincible vertex is at least

$$\frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{3}{16}.$$

□

**Lemma 18.** *In a graph $\mathcal{A}^d(n, b)$, for any configuration where at least one vertex is invincible, the cooperators spread over the whole graph and the spread happens in the expected time $\mathcal{O}((20b)^{2b}n^2)$.*

*Proof.* We already know that the invincible vertex cannot be converted. We show that from any position, the fixation has a positive probability and we bound the expected time for fixation.

In any configuration, we show the number of invincible vertices increases with positive probability. When all big vertices are invincible, we show that the number of small vertices that are cooperators and have at least $2b$ cooperating neighbors increases with positive probability. Moreover, such cooperators cannot be converted. When all big and small vertices are cooperators with at least $2b$ cooperating neighbors, no cooperator can be converted. Any defector can be converted with positive probability.

Let $D$ be a constant, approximately double the maximal degree of the graph. Every step that we examine is local (the changes are happening in distance at most $4$ from an invincible neighbor). The value $D$ is the number of edges in the neighborhood that matter.

In any configuration, if a big vertex is seeded, Lemma 17 shows that the probability that the seeded vertex becomes invincible is positive. Moreover, any big vertex that has at least $b$ cooperating neighbors can become seeded with probability at least $\frac{1}{D^b}$ (it converts $b$ leaves before its other cooperating neighbors are converted). Now, we consider that no big cooperator has $b$ cooperating neighbors.

In any configuration, we fix an invincible vertex $v$ and its small neighbor $u$ and its big neighbor $w$ (that has less than $b$ cooperating neighbors). First, $w$ has to be convertible to a cooperator. This happens when $w$ becomes the defector and all its neighbors become defectors (probability $\frac{1}{D^{b+1}}$). Then, $u$ is converted to cooperator and all bridge vertices between $u$ and $v$ too (probability $\frac{1}{D^b}$). Then, all leaves of $u$ are converted (probability $\frac{1}{D^b}$). Finally, all bridge vertices between $u$ and $w$ and $w$ itself are converted to cooperators (probability $\frac{1}{D^b}$). After converting $b$ leaves to cooperation (probability $\frac{1}{D^b}$), $w$ is seeded and we can use Lemma 17.

This means the big vertex $w$, a neighbor of $v$, becomes invincible with probability at least $\frac{1}{3D^{5b}}$ from any position in at most $4b^2$ steps. That means the expected number of steps around $w$ is $\mathcal{O}((20b)^{5b})$ to make $w$ invincible.

Now, when all big vertices are invincible, we can look at small vertex $u$ between $v$ and $w$. The vertex $u$ itself and all bridge vertices can be converted by $v$ or $w$, this happens with probability at least $\frac{1}{D^{2b}}$. The number of steps around $u$ needed is $\mathcal{O}((20b)^{2b})$.

After all big vertices are invincible and all bridge vertices and small vertices are cooperators, selecting any edge leads to the spread of cooperation. There are less than $2n$ edges in the graph, the time until all of them are selected at least once is $\mathcal{O}(n \log n)$.

To make $s$ steps around one vertex, the process needs to make $s \cdot \frac{n}{10b^2}$ steps in the whole graph on average. To spread the cooperation over all big vertices, we need to increase the number of vertices $\mathcal{O}(n)$ times. That means the number of steps to convert all big vertices to invincible is $\mathcal{O}((20b)^{2b}n^2)$. □

**Lemma 19.** *For $p \geq \frac{1}{\left(\frac{n}{t^2}\right)^{\frac{1}{b+1}}}$ where $t \geq 1$, on the graph $\mathcal{A}^d(n,b)$, the probability that cooperators fixate is at least*

$$1 - e^{-0.2t}\,,$$

*and the expected time to fixate is $\mathcal{O}((20b)^{2b}n^2)$.*

*Proof.* From Lemma 16, with probability $1 - e^{-t}$, there are at least $t$ seeded vertices. No matter the neighborhood, from Lemma 17, a seeded vertex becomes invincible with probability at least $\frac{3}{16}$. So the vertex fails to become invincible with probability at most $\frac{13}{16}$. The probability that among $t$ seeded vertices no one turns invincible is $\left(\frac{13}{16}\right)^t$. That means the probability that no vertex became invincible is at most $e^{-t} + \left(\frac{13}{16}\right)^t < e^{-0.2t}$ for $t \geq 1$, therefore some vertex becomes invincible with probability at least $1 - e^{-0.2t}$.

From Lemma 18, we know that with probability $1$, one invincible vertex converts all defectors.

The seeded vertex becomes invincible in $\mathcal{O}(n \cdot 10b^2)$ steps in expectation. All the steps have to spread cooperation. From Lemma 18, we know that the expected number of steps for cooperation to fixate is $\mathcal{O}((20b)^{2b}n^2)$. □

*Proof of Theorem 10.* For $b > 1$, we have Lemma 19, setting $t$ to $\sqrt{np^{b+1}}$, yields $\rho_C(\mathcal{A}^d(n,b),b,p) = 1 - e^{-0.2\sqrt{np^{b+1}}}$, therefore

$$\lim_{n\to\infty} \rho_C(\mathcal{A}^d(n,b),b,p) = 1\,.$$

□

### 4.4.4 Proof of Theorem 11

This section follows a similar outline as Section 4.4.3. We use similar terminology because a lot of concepts in the spread of cooperation are the same.

**Seeded and stronghold vertex.** In $\mathcal{A}^r(n,b)$, we say that a big vertex is *seeded* if it is a cooperator or has at least one cooperating neighbor. We call a big vertex a *stronghold* if all its leaves are cooperators.

**Spread of cooperators.** Here, we describe the spread of cooperation in $\mathcal{A}^r(n,b)$ on the high level. Later, we compute the exact probabilities.

1. Every seeded vertex has a substantial probability of turning into a stronghold (Lemma 20).
2. The stronghold vertex is more likely to make neighboring big vertex seeded and eventually stronghold than to convert to defection. The ratio between these two events is substantial (Lemma 21).

The lemmas connect in Lemma 22, which is used to prove Theorem 11.

**Lemma 20.** *For graph $\mathcal{A}^r(n,b)$ and any position that contains a seeded vertex $v$, the probability that $v$ becomes stronghold is at least*

$$2^{-5}e^{-b^2}\,.$$

*Proof.* We estimate the probability that the seeded vertex $v$ turns into a stronghold. First, the vertex $v$ needs to become a cooperator. Then, it needs to spread cooperation to its neighbors. Until the number of cooperating neighbors is below $b$, the spread is unlikely: any defector has a higher payoff than $v$. Until the number of cooperators is below $2b$, the edge between $v$ and a small vertex is dangerous. This defector can have two cooperating neighbors. Afterwards, the spread of cooperation is likely.

We split the spread of cooperation into several phases. Let $i$ be the number of cooperating neighbors of $v$.

In the first phase, the vertex $v$ becomes a cooperator. The second phase is for $i \leq b$. The third phase is for $b < i \leq 3b$. The fourth phase is for $3b < i$.

**First phase.** If the big vertex $v$ is a defector and it has more than $b$ cooperating neighbors, we wait until one happens: $v$ becomes cooperator, or the number of cooperating neighbors is at most $1$. One or the other happens with probability $1$, and we focus on the more pessimistic case where $v$ is a defector.

If an edge is selected between $v$ and a cooperator, the spread of cooperation is successful with the probability of at least

$$\frac{1}{1+e^{b-0}}\,.$$

It is since $v$'s payoff is at most $b$ and the payoff of a lonely cooperator is $0$.

**Avoiding edges to small vertex.** To simplify computation, we bound the probability that an edge between a small vertex and $v$ is selected in the second or third phase. We call it a bad edge, and if a bad edge is selected, we suppose that $v$ did not turn into a stronghold. Having $i$ cooperating neighbors and if $v$ is a cooperator, there are $10b^2 - i$ edges that can be selected and $v$ or its neighbor change. Two edges from this are bad, so with probability $1 - \frac{2}{10b^2-i}$ bad edge is not selected. There are at most $3b$ steps in the second and third phases (until the number of cooperators is at least $3b$). That means we can bound the probability that no bad edge is selected by

$$\left(1 - \frac{2}{10b^2 - 3b}\right)^{3b} > 1 - \frac{6b}{10b^2 - 3b} > 1 - \frac{1}{\frac{17}{12}b}\,.$$

**Second phase.** When $v$ is a cooperator with $i \leq b$ cooperating leaves, the probability that the number of cooperating leaves increases conditioned on the fact that $v$ or its neighbors change is at least

$$\frac{1}{1+e^{b-i}}\,.$$

because we know that no bad edge is selected.

The probability that the second phase succeeds is

$$\prod_{i=0}^{b} \frac{1}{1+e^{b-i}}\,.$$

Now, we show that the probability of success in the first and second phases is at least

$$2^{-4}e^{-b^2}\,.$$

The probability of success is $\frac{1}{1+e^b} \cdot \prod_{i=0}^{b} \frac{1}{1+e^{b-i}}$. Since $b$ is an integer, we use the induction over $b$, for $b = 2$, we have $\frac{1}{1+e^2} \cdot \prod_{i=0}^{2} \frac{1}{1+e^{2-i}} > 2^{-4}e^{-4}$. Increasing $b$ from $k$ to $k+1$ multiplies the left side by $\frac{1+e^k}{1+e^{k+1}} \cdot \frac{1}{1+e^{k+1}}$. The right side is multiplied by $e^{-2(k+1)-1}$, which is smaller. Therefore, the inequality holds.

**Third phase.** In the third phase, the probabilities are the same as in the second phase. Now, every step succeeds with high probability. The success probability of the third phase is

$$\prod_{i=b+1}^{3b} \frac{1}{1+e^{b-i}} > \prod_{i=1}^{2b} \frac{1}{1+e^{-i}} .$$

**Fourth phase.** In the fourth phase, an edge between a small vertex and $v$ can be selected, the defector can have a payoff of $2b$. It gives the probability of success of the fourth phase:

$$\prod_{i=3b+1}^{10b^2} \frac{1}{1+e^{2b-i}} = \prod_{i=0}^{10b^2} \frac{1}{1+e^{-b-1-i}} > \prod_{i=0}^{\infty} \frac{e^{b+1+i}}{1+e^{b+1+i}} .$$

Even for the smallest $b = 2$, we have the success probability of the third and fourth phases together with the probability of avoiding small vertex is at least $\frac{1}{2}$. So, the probability that one seeded vertex turns into a stronghold is at least

$$2^{-5}e^{-b^2} .$$

□

**Lemma 21.** *In a graph $\mathcal{A}^r(n, b)$, for any configuration where at least one vertex is a stronghold, the ratio between increasing and decreasing the number of strongholds is at least*

$$e^{6b^2} ,$$

*and the expected time until the number of strongholds changes is $\mathcal{O}(e^{6b^2}n)$.*

*Proof.* Let $v$ be a stronghold vertex and $u$ its neighbor, and let $w$ be its big neighbor. If the edge between $u$ and $v$ is selected, then with probability at least $\frac{1}{1+e^{-10b^2+2+2b}}$, vertex $w$ is seeded ($u$ becomes cooperator). From Lemma 20, we know that with probability $2^{-5}e^{-b^2}$, vertex $w$ turns into a stronghold. Note that during the whole process of turning $w$ into a stronghold, vertex $u$ is a cooperator, so it cannot turn $v$ into a defector.

With probability $\frac{1}{1+e^{10b^2-2-2b}}$, vertex $v$ changes into defector, which means the number of strongholds decreases.

That means the ratio between increasing and decreasing the number of strongholds is

$$\frac{2^{-5}e^{-b^2} \cdot \frac{1}{1+e^{-10b^2+2+2b}}}{\frac{1}{1+e^{10b^2-2-2b}}} > 2^{-5}e^{-b^2} \cdot e^{10b^2-2-2b} > e^{10b^2-2-2b-6-b^2} \geq e^{6b^2} .$$

At worst, there is only one pair $v, w$, where $v$ is the stronghold and $w$ is the neighboring big vertex. For one step to happen around $w$, there need to be $\mathcal{O}(n)$ steps in the whole graph. If $w$ is seeded the expected time (around $w$) until it becomes stronghold is $\mathcal{O}(30b^2 \log 30b^2)$ (which

is smaller than $e^{6b^3}$) by that time all edges are selected at least three times on average. That means the expected number of steps around $w$ required to change the number of strongholds is at most $\mathcal{O}(e^{6b^2} + 30b^2 \log 30b^2)$. Which gives $\mathcal{O}(e^{6b^2} n)$ steps overall.

□

**Lemma 22.** *For $p \geq t \cdot \frac{e^{b^2} 2^3 \cdot 10b^2}{n}$, on the graph $\mathcal{A}^r(n, b)$, the probability that cooperators fixate is at least*

$$1 - e^{-3t}\,,$$

*and the expected time to fixation is $\mathcal{O}(e^{6b^2} n^2)$.*

*Proof.* At the start of the process, the probability that one big vertex is not seeded is $(1-p)^{10b^2}$, which for $b = 2$ is at most $1 - 2^5 \cdot p$. So, the probability that a big vertex is seeded is at least $2^5 \cdot p$. The probability that one seeded vertex turns into a stronghold is $2^{-5} e^{-b^2}$, therefore from the start, we have that a big vertex becomes a stronghold with probability $e^{-b^2} p$.

Let $X_i$ be the random variable that for big vertex $i$ is one if it becomes seeded at the beginning and then turns to the stronghold and is zero otherwise. For every big vertex (there are $\frac{n}{10b^2}$ of them), the probabilities are independent. (Strictly speaking, they are not independent but the lower bounds expect the worst position, which is independent.) Now, we use Chernoff bound for $X = \sum X_i$ (note that $\mu = e^{-b^2} p \cdot \frac{n}{10b^2}$), which gives:

$$\begin{aligned} \mathbb{P}[X \leq (1-\delta)\mu] &\leq e^{\frac{\delta^2 \mu}{2}} \\ \mathbb{P}[X \leq \left(1 - \frac{7}{8}\right) e^{-b^2} p \cdot \frac{n}{10b^2}] &\leq e^{\frac{7^2 e^{-b^2} p \cdot n}{2 \cdot 10 \cdot 8^2 \cdot b^2}} \\ \mathbb{P}[X \leq t] &\leq e^{\frac{49t}{16}} \end{aligned}$$

after setting $p = t \cdot \frac{e^{b^2} 2^3 \cdot 10b^2}{n}$. That means with probability at least $1 - e^{\frac{49}{16}t}$, there are at least $t$ stronghold vertices.

We know that stronghold vertices have spread ratio $e^{6b^2}$, they form a dimensional Markov Chain with constant bias and two absorbing ends. The probability that we reach the end with all big vertices being strongholds while starting with $t$ strongholds is at least

$$1 - e^{-6b^2 t}\,.$$

So the probability that the cooperators fixate is at least $1 - e^{-3t}$.

Since the Markov Chain is biased, the expected number of steps (changes in the number of strongholds) is $\mathcal{O}(n)$. One change requires $\mathcal{O}(e^{6b^2} n)$ steps, therefore we have the total expected number of steps $\mathcal{O}(e^{6b^2} n^2)$. □

*Proof of Theorem 11.* For all $b > 1$, we have Lemma 22, setting $t$ to $\sqrt{n}$, we have that $\rho_C(\mathcal{A}^r(n, b), b, p) = 1 - e^{-3\sqrt{n}}$, therefore

$$\lim_{n \to \infty} \rho_C(\mathcal{A}^r(n, b), b, p) = 1\,.$$

□

**Remark 1** (Extension to general $K$)**.** *In the proof, we used $K = 1$. All the proofs can be easily extended to general $K$ which we discuss below. (a) Lemma 20: In the case of general $K$, Lemma 20 shows the ratio $2^{-5/K}e^{-b^2/K}$. In the proof of the lemma, it is straightforward to add $K$ in all desired inequalities. The only change needs to be done for* Avoiding edges to small vertex *for large $K$. If $K$ is large, we need to consider that the small vertex was converted to cooperation if the edge between the small and big vertex was selected. (b) Lemma 21: This lemma gives the ratio $e^{6b^2/K}$ by straightforward calculation. (c) Lemma 19: Finally, having the previous lemmas, in Lemma 22 we replace factor $e^{b^2}$ by $e^{b^2/K}$ and the time decreases from $\mathcal{O}(e^{6b^2}n)$ to $\mathcal{O}(e^{6b^2/K}n)$.*

*In summary, we have that for large $K$, the process is faster and requires a slightly smaller initial rate of cooperation. For small $K$, we still have the strong density amplifier property, but the process is slower and requires a higher initial rate of cooperation.*

### 4.4.5 Lower Bound on Initial Cooperation Density in Deterministic Setting

In several previous settings with constant fitness or weak selection, previous literature considers the fixation probability of a single mutant [OHLN06, LHN05], or of some cooperator density [PJR+17, ARS06, HD04, SAP22]. However, note that for spatial games with deterministic update, which are the limit of strong selection, if there are less than $b+1$ cooperators initially, then no graph can achieve fixation. Hence in the setting of spatial games with deterministic update, a single initial cooperator would lead to extinction for every graph. Below we present an even stronger requirement (i.e., lower bound required) for fixation for the initial cooperation density parametrized by the highest degree. We determine a value that is a function of the maximal degree of the graph such that if $p$ is smaller than the value, then with a high probability no cooperator has $b$ or more neighbors, which implies extinction.

**Theorem 12.** *For all graphs with $n$ vertices and highest degree $d$, for*

$$p \leq \frac{1}{\left((ed)^b n\right)^{\frac{1}{b+1}}},$$

*the extinction probability of cooperators is at least $\frac{1}{2}$.*

*Proof.* In the deterministic setting, if a cooperator has fewer than $b$ cooperating neighbors, then it cannot spread. We show that if $p \leq \frac{1}{\left((ed)^b n\right)^{\frac{1}{b+1}}}$, then the probability that no cooperator has $b$ or more cooperating neighbors is at most $\frac{1}{2}$, which ensures extinction probability is at least $\frac{1}{2}$.

For one given vertex, the probability that this vertex is a cooperator and has at least $b$ cooperating neighbors can be bounded by the following

$$p \cdot \sum_{x \geq b} p^x (1-p)^x \binom{d}{x} \leq \sum_{x \geq b} p^{x+1} \left(\frac{de}{x}\right)^x.$$

Since $\frac{de}{b} \cdot p \leq \frac{1}{2}$, we know that the previous summand is at least twice as high as the next. That means it sums to at most

$$2p^{b+1}\left(\frac{de}{b}\right)^b \leq 2b^{-b}n^{-1}.$$

There are at most $n$ vertices, from the union bound, the probability that no cooperator has $b$ or more neighbors is at most

$$n \cdot 2b^{-b}n^{-1} .$$

For $b \geq 2$, this gives $\frac{1}{2}$. This completes the desired proof. □

Note that our construction from Lemma 19 requires a bound of $p$ which is close to the above lower bound.

### 4.4.6 Parameters of Prisoners' dilemma

In the main article, we consider the classic parametrization of the payoff for spatial games from [NM92a, Now07]. The payoff matrix allows defectors to be as strong as possible while all possible payoffs are nonnegative. In this section, we examine the range of parameters of the Prisoners' dilemma for which our result holds. In general Prisoner's dilemma is given by the following matrix:

| | C | D |
|---|---|---|
| C | $R$ | $S$ |
| D | $T$ | $P$ |

and the payoffs follow inequality $T > R > P > S$. If $S < 0$, we add $|S|$ to all entries of the payoff matrix to avoid the possibility of negative fitness. Then we have $T > R > P > S \geq 0$.

We argue two aspects: (a) for the deterministic setting we present the range of parameters for which our graph structures are strong density amplifiers; and (b) for the randomized setting we show that for all parameters our graph structures are strong density amplifiers.

#### Generalization for deterministic setting

We show that our results hold in the deterministic setting if the following inequality holds: $R^2 > T \cdot P$. We parametrize our construction as follows: (a) instead of $b-1$ bridge vertices, there are $n_b$ bridge vertices; (b) the number of support vertices next to the small vertex is $s_s$; and (c) the number of support vertices next to a big vertex is $s_b$. In the proof for strong density amplifiers, we require the following three aspects:

1. notion of invincibility,
2. initial probability such that one big vertex becomes invincible,
3. possibility to spread from the small vertex to the big vertex and make it invincible.

We argue about each item below:

1. For the notion of invincibility, we need the big vertex to have enough neighbors (that are cooperators) such that no neighbor (small vertex) can have a higher payoff:

$$s_b \cdot R > (2n_b + s_s) \cdot T . \tag{4.16}$$

2. We want to have $s_b$ significantly bigger than $n_b$, such that the seeded vertex turns invincible easily. We need to be able to convert a small vertex from a bigger one, but this condition is weaker than Equation (4.16).

3. We need the small vertex to be able to convert its support vertices if it has only $n_b$ cooperating neighbors, therefore $n_b \cdot R > T$. Moreover, for the possibility of the spread from small to big vertex, a small vertex needs to have enough neighbors to overweight the advantage of a defector in a big vertex getting $P$ from all neighbors:

$$(n_b + s_s) \cdot R > (s_b + 2n_b) \cdot P\,. \tag{4.17}$$

Except for $n_b \cdot R > T$, we do not have any condition that forces $n_b$ to be large. In other conditions, we want $n_b$ to be small. So in the following computations, we consider that $n_b$ is small compared to $s_s$ and $s_b$. Multiplying Equation (4.16) and Equation (4.17), we get

$$R^2 > \frac{s_s + 2n_b}{s_s + n_b} \cdot \frac{s_b + 2n_b}{s_b} \cdot T \cdot P\,. \tag{4.18}$$

Since we consider $n_b$ to be small, we have

$$R^2 > T \cdot P\,.$$

If we have the inequality, we can construct the graph with parameters $n_b$, $s_s$, and $s_b$. Moreover, since $R^2 > T \cdot P + \varepsilon$, we have that $\frac{n_b}{s_b}$ is proportional to $\varepsilon$. We justify our result:

- *Seeded and invincible vertex.* Vertex is seeded if it is a cooperator with enough cooperator neighbors to be able to convert defector leaves. Vertex is invincible if it satisfies Equation (4.16).

- *Lemma 16* uses combinatorial reasoning and does not consider $T$ or $S$, so the arguments direclty translate for $b = \frac{T}{R}$.

- *Lemma 17* is also about combinatorial reasoning. In the proof, we use the fact that it is enough to have $3b^2$ cooperator neighbors out of $10b^2$ to become invincible. Now, the big vertex needs a higher proportion of neighbors to be cooperators to become invincible. It is more likely that the edge between the big and small vertex is selected before the big vertex is invincible. However, we can track the payoff of the small vertex to show that the big vertex is not converted. In expectation, the defector in a small vertex converts at least half of its neighbors to defection until the edge between the big and small vertex is selected.

- *Lemma 18* describes how the cooperation spreads from the invincible vertex to the next small and next big vertex, the spread works the same.

The above arguments imply that our structures are strong density amplifiers for $R^2 > T \cdot P$.

**Generalization for randomized setting**

We now argue that in the randomized setting the only requirement is $S \geq 0$. For the graph, we just need one parameter, the number of support vertices of a big vertex, $s_b$. In the initialization, we consider a vertex seeded only when the big vertex is cooperator itself (this

requires increasing $p$, but not too much). Then, the proof of Lemma 20 works the same. We call the big vertex a stronghold if it has $x$ cooperating neighbors and the following holds:

$$\frac{1}{1+e^{x\cdot R}} < \frac{1}{1+e^{P\cdot s_b}} \cdot \left(\frac{1}{1+e^{T}}\right)^{\lceil T/R \rceil},$$

which means that the stronghold is more likely to convert a neighboring vertex to a stronghold than to be converted itself, similarly as Lemma 21. Note that since $s_b$ is the only parameter and $R > P$, there is a large enough $s_b$ that satisfies the condition. This means the argument similar to the proof of Lemma 21 also applies for the general setting of Prisoners' dilemma with $S \geq 0$.

#### Donation games

We consider parametrization from [SP13, OHLN06, ALC$^+$17] called donation games. There are two parameters $b$ and $c$ with $b > c > 0$, and we have

| | C | D |
|---|---|---|
| C | $b-c$ | $-c$ |
| D | $b$ | $0$ |

To ensure non-negativity adding $c$ to all entries, we obtain

$$\begin{matrix} b & 0 \\ b+c & c \end{matrix}$$

Let $\beta = \frac{b}{c}$ be the benefit-to-cost ratio. From Section 4.4.6, in the deterministic setting, we have that our construction is a strong amplifier for $\beta^2 > \beta + 1$, which gives $\beta > \frac{1+\sqrt{5}}{2} \approx 1.62$. The ratio we require $\beta > 1.62$, is in contrast to other works [ALC$^+$17] that in graphs for survival of cooperation require $\beta$ to be at least average degree (e.g., $\beta > 4$ for grids and $\beta > 2$ for all connected graphs). From Section 4.4.6, in the randomized setting, our results hold for all $\beta > 1$.

### 4.4.7 Other Initializations

Our main theorems consider the random initialization where every vertex is independently a cooperator with probability $p$. This corresponds to cooperators arising spontaneously. In this section, we discuss the influence of other natural initializations on the fixation probability. We consider mainly deterministic setting, since Section 4.4.9 show that even one cooperator in the randomized updating fixates with a substantial probability. A key property for the strong density amplification property of our graph structures is the requirement that a seeded vertices appear in the initialization. We argue about this seeded property for other initializations.

#### Temperature initization

Along with random initialization (that corresponds to spontaneously arising mutation) another natural initialization is proportional to the in-degree (or temperature), which corresponds to mutations arising during reproduction. First, our main result shows that our graph structures are strong density amplifiers as $n$ goes to $\infty$ and $p$ goes to $0$. Second, if for random initialization a vertex is a cooperator with probability $p$, then under temperature initialization,

it is a cooperator with probability at least $\frac{p}{d^2}$, where $d$ is the maximum degree. Third, the above implies that if we consider an initialization probability $p \cdot d^2$ for temperature initialization, then the seeded property is satisfied. Finally, from the third and first item above we conclude that since we have a constant-degree graph family with initialization probability multiplied by $d^2$, we still obtain strong density amplifiers where $p$ still goes to $0$ as $n$ goes to $\infty$ for temperature initialization.

#### Correlated initialization

Another natural initialization is a correlation between cooperators. In this setting a pair of cooperating neighbors is more likely, e.g., the mutation comes as multiple offsprings from the same parent. The correlated initialization increases the fixation probability. The vertex is seeded with higher probability because if a big vertex is a cooperator, it is more likely that the neighbors are also cooperators, i.e., if with initialization probability $p$ at random, we satisfy the seeded property, we also satisfy the property for correlated initialization. Hence our graph structures are also strong density amplifiers for correlated initialization.

### 4.4.8 Fixation probability of Single Mutant in Randomized Setting

In this section, we study the fixation probability of a single mutant (i.e., a single cooperator among all defectors, or a single defector among all cooperators). Theorem 12 shows that in the deterministic setting a single cooperator always goes extinct, and hence we focus only on the randomized setting for fixation of a single mutant.

*Notation.* We denote by $\rho_C(G, b)$ (resp., $\rho_D(G, b)$) the fixation probability of one randomly placed cooperator (resp., defector) in a graph $G$ where all other individuals are defectors (resp., cooperators).

#### Bounds on $\rho_C(G, b)$ and $\rho_D(G, b)$ for $\mathcal{A}^r$

We present bounds on the fixation probability of a single mutant for $\mathcal{A}^r$ in the following two results.

**Theorem 13.** *For asynchronous randomized updating and all $b$, $n$, and graph $G = \mathcal{A}^r(n, b)$ the following assertion holds*

$$\rho_C(G, b) \geq 2^{-5} e^{-b^2} \left(1 - e^{-6b^2}\right) .$$

*Proof.* If there is one cooperator in the graph, at least one big vertex is seeded. From Lemma 20, with probability $2^{-5}e^{-b^2}$ it turns into a stronghold.

From Lemma 21, we have that the ratio between increasing the number of strongholds and decreasing it is at least $e^{6b^2}$. That means the probability that one stronghold turns all other big vertices into strongholds is at least

$$1 - e^{6b^2} .$$

So the overall fixation probability is $2^{-5} e^{-b^2} \left(1 - e^{-6b^2}\right)$. □

**Theorem 14.** *For asynchronous randomized updating and all $b$, $n$, and graph $G = \mathcal{A}^r(n, b)$ the following assertion holds*

$$\rho_D(G, b) \leq 2^{-c \cdot n}$$

*for some constant $c > 0$.*

*Proof.* One randomly dropped defector might make at most one stronghold to cease to be a stronghold. Then, cooperators still occupy at least $\frac{n}{10b^2} - 1$ strongholds.

From Lemma 21, the ratio between increasing the number of strongholds and decreasing it is $e^{6b^2}$. That means the cooperators fixate with probability at least

$$1 - \left(e^{6b^2}\right)^{\frac{n}{10b^2}-1},$$

which gives the fixation probability to defectors at most

$$\left(e^{6b^2}\right)^{\frac{n}{10b^2}-1} = 2^{-cn},$$

for some constant $c > 0$. The desired result follows. □

**Remark 2.** *The above two results for $\mathcal{A}^r$ establish that the fixation probability of a single cooperator among defectors is constant (independent of the population size $n$), whereas the fixation probability of a single defector among cooperators is exponentially small in the population size $n$. This establishes the last row of Table $1$ in the main article.*

**Bounds on $\rho_C(G, b)$ and $\rho_D(G, b)$ for other graphs**

We now discuss the relevant bounds for other classical graphs, i.e., complete graph, grid, and star. Before we reason about the graphs, we recall a simple property from the literature on constant selection. If a random mutant is introduced in a population of residents and the relative fitness advantage is $r$, then the fixation probability is constant for $r > 1$ and is exponentially small for $r < 1$ (see [Now07], Equation 6.13).

**Complete graph.** In the complete graph, at any configuration, every defector has a higher payoff than every cooperator, and the relative fitness advantage of defectors over cooperators is always at least $r > 1$, which implies the relative fitness advantage of cooperators over defectors is at most $1/r < 1$. The simple property above implies that the fixation probability of a single defector is constant, whereas the fixation probability of a single cooperator is exponentially small. This establishes the first row of Table $1$ in the main article.

**Grid.** The grid is a regular graph (i.e., an isothermal graph) that has similar properties in the constant selection regime as the complete graph. For $b \geq 3$, the same arguments as in the case of the complete graph apply for the relative fitness advantage and we have the same conclusion related to the fixation probability bounds for a single mutant. This establishes the second row of Table $1$ in the main article.

**Star.** For a star, one randomly placed defector into the star spreads with probability close to $1$ only if it is placed in the center of the star. Hence the fixation probability of a single defector is proportional to $\frac{1}{n}$. For a single cooperator, there are two cases: (a) When the initial cooperator is in the center, which happens with probability $1/n$, the fixation probability is at most a constant (upper bound 1); and (b) When the initial cooperator is placed on a leaf of the star, we need to consider $\varepsilon$, which is a small constant that does not matter in graphs with constant degrees, but here, the payoff of the central defector is $(n-2)\varepsilon + b$. That means the cooperator in the leaf spreads with probability at most $\frac{1}{1+e^{(n-2)\varepsilon+b}}$. Combining the above two cases, we have that the fixation probability of a randomly placed cooperator is again proportional to $\frac{1}{n}$. This establishes the third row of Table $1$ in the main article.

### 4.4.9 Mutation

In this section, we examine the robustness of our construction against mutation in the randomized setting. We consider a graph initially inhabited only by defectors and a small mutation rate $\mu$. During the replacement, one of the two individuals that participated in the replacement has a probability $\mu$ of changing to the other type. This represents a mutation of offspring where the offspring or parent moved to a new location.

In Figure 4.7, we see that for our strong density amplifiers, the cooperator density increases with time until it reaches $1$. After any mutation event, there is a high probability that a big vertex is seeded, as Lemma 20 shows the seeded vertex becomes stronghold with high probability. In a grid, we do not see any cooperation evolving, the spread of cooperation is limited to the neighborhood. The star can support some cooperation. We see the cooperation density hovering below $0.2$, but the star is very unstable. One cooperator can spread, but it is prone to being conquered by one defector, so the star alternates between full cooperation and full defection.

The simulation results for $\mathcal{A}^r(10^4, 1.5)$ in Figure 4.7 show the two following facts: (a) the ratio between one random cooperator fixating inside the population of defectors is significantly higher than the probability of one defector fixating among the cooperators, because the cooperator density increases above $\frac{1}{2}$; and (b) the fixation probability of one cooperator among defectors is above $\frac{1}{N}$. We can observe this as follows: if the probability of one cooperator fixating is around $1/N$, then this requires in expectation $N \cdot 1/\mu$ rounds for substantial increase in fixation, which is $10^{11}$ rounds. Instead, we see the increase of cooperation by orders of magnitude faster than $10^{11}$.

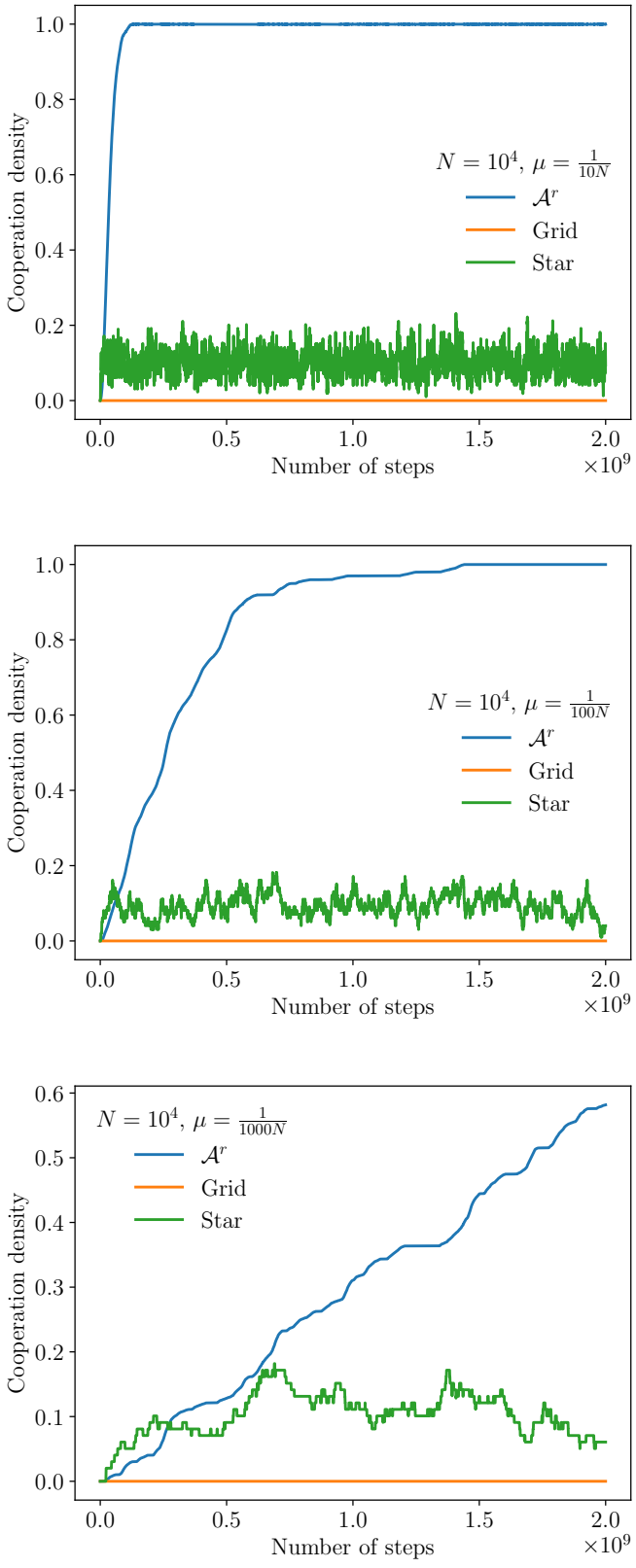



Figure 4.7: The density of cooperation as a function of the number of steps with small mutation rate $\mu \in \{10^{-5}, 10^{-6}, 10^{-7}\}$ in the randomized setting with $K = 1$. The density is an average over 100 runs. The graph size is $N = 10^4$ and the graphs $\mathcal{A}^r(10^4, 2)$, star, and grid are considered. The value of $b$ is set to 1.5.

CHAPTER 5

# Coexistence times in Moran process with environmental heterogeneity

This chapter appears in full in [STKC23].


## Abstract

Populations evolve in spatially heterogeneous environments. While a certain trait might bring a fitness advantage in some patch of the environment, a different trait might be advantageous in another patch. Here we study Moran Birth-death process with two types of individuals in a population stretched across two patches of size $N$, each patch favoring one of the two types. We show that the long-term fate of such populations crucially depends on the migration rate $\mu$ between the patches. To classify the possible fates, we use the distinction between polynomial (short) and exponential (long) timescale. We show that when $\mu$ is high then one of the two types fixates on the whole population after a number of steps that is only polynomial in $N$. In contrast, when $\mu$ is low then each type holds majority in the patch where it is favored for a number of steps that is at least exponential in $N$. Moreover, we precisely identify the threshold migration rate $\mu^\star$ that separates those two scenarios, thereby exactly delineating the situations that support long-term coexistence of the two types.


## 5.1 Introduction

Evolution is a stochastic process that acts on populations of reproducing individuals. Each individual has a fitness that determines its reproductive rate. In the absence of mutation, one individual eventually produces a lineage of offspring that takes over the whole population. This event is called a *fixation*. The key quantities of the evolutionary process are the fixation probabilities of the respective individuals and the (expected) time until the fixation happens [Mor62, Ewe04, Now06a].

Population structure is known to substantially affect those quantities, thereby altering the typical fate of a population [Sla81, DL94b, Whi03, HV11, FRT13]. The effects of various population structures are conveniently studied within the framework of evolutionary graph theory [LHN05, SF07]. The spatial structure of the population is described by a graph (network) whose nodes correspond to sites. Each site is occupied by a single individual and the

edges (possibly weighted) represent migration rates between pairs of sites. The population of asexually reproducing individuals then evolves according to a discrete-time stochastic process called *Moran Birth-death process*. [Mor58] In each step, first ("Birth") an individual is selected for reproduction with probability proportional to its fitness, and then ("death") the offspring migrates to a neighboring site and replaces its initial occupant. Thus the population size $N$ remains constant. The case of a perfectly well-mixed population consisting of $N$ individuals is recovered by taking a complete graph $K_N$ with all edges having unit weight.

Of special interest is the case of a single beneficial mutant with relative fitness advantage $r > 1$ invading a background population of $N - 1$ indistinguishable residents, each with a fitness $1$. For the well-mixed population, the *fixation probability* $\rho(K_N, r)$ of the mutant is equal [LHN05] to $(1 - 1/r)/(1 - 1/r^N)$, which tends to $1 - 1/r$ as $N \to \infty$ and, on average, the process terminates after approximately $(1 + 1/r) \cdot N \log N$ steps. [AT09, DGRS16] That is, for fixed $r > 1$, the *fixation time* $\mathrm{FT}(K_N, r)$ is proportional to $N \log N$.

An important driving question in the field over the past decade has been the hunt for population structures, so-called *amplifiers of selection*, that enhance the fixation probability of a single beneficial mutant invading a background population of indistinguishable residents, as compared to the well-mixed population. [ACN15, BRS11, HT15, ALC+17, PTCN17, TPCN20, ASJ+20] A prime example of an amplifier is a Star graph $S_N$ consisting of $N - 1$ leaf nodes, all of them connected to a single central node (but not to each other). It is known that, in the limit $N \to \infty$, we have [BR08, MGP14, Cha16] $\rho(S_N, r) \to 1 - 1/r^2$. When $r = 1 + s$ for $s$ small, this is roughly a two-fold increase compared to the baseline given by $\rho(K_N, r)$. Even more strongly, there exist structures (so-called *superamplifiers*) that guarantee fixation of the mutant in the limit $N \to \infty$, no matter how small its fitness advantage $r > 1$ is [Gia16, GGG+17, PTCN18, GLL+19].

Regarding the timescale of the process, formal results concerning Moran Birth-death process on structured populations are comparably scarcer and those that exist focus on identifying population structure with short evolutionary timescales. The reason is that such structures, especially when they amplify, could potentially speed up the rate of evolution. For example, it is known that for Star graphs, the process terminates after roughly $N^2 \log N$ steps. [BRS11] More generally, when the underlying graph is undirected (that is, all edges are two-way) then the process terminates after a number of steps that is polynomial in $N$. [DGM+14, AGLR20] In contrast, there exist directed and weighted graphs for which the process takes exponentially many steps. [DGRS16] Also, it has been empirically observed that structures that enhance the fixation probability of a mutant tend to increase the fixation time. [TPCN19, MHT19, TPCN21] However, an important limitation of fixation time as a quantity is that it relates only to the overall duration of the process. In other words, it is oblivious to what is actually happening during the process before one of the types fixates.

More recently, the framework of evolutionary graph theory has been enriched with environmental heterogeneity. [MP14] This is done by partitioning the sites into patches. The fitness of each individual is constant within each patch but it can vary across different patches (see Fig. 5.1b). Concerning the fixation probability, analytical results are known for large well-mixed populations and any number of patches [KMN19], and for certain special families of regular graphs and two patches. [KMCN20] Concerning the duration of the process, to our knowledge no results are known.

In this work, we study the timescale of the Moran Birth-death process in populations that are spatially structured and environmentally heterogeneous. This differs from the earlier research focus in four regards. First, for populations that are environmentally heterogeneous, to our

knowledge no results on evolutionary timescales are known (the previous research focused on fixation probabilities of the respective types). Second, for homogeneous populations, the past work related to the duration of the process has focused on identifying population structures with short fixation times (and high fixation probability), since such structures could potentially be used to speed up the evolutionary process. In contrast, our goal here is to characterize structures that support long-term coexistence of the two competing types. Such structures are important in population genetics where they correspond to multiple-niche ecosystems with protected polymorphism.[Lev53, Hed86]. As another application, consider a single-species biofilms: Although very small in scale, biofilms typically consist of several microenvironments and show substantial heterogeneity, both genetic and phenotypic.[DLL10, SF08]. Third, while the past research used the notion of a fixation time, here we introduce a refined notion of a coexistence time. Our results thus provide stronger guarantees about the state of the population throughout the process. Fourth, we utilize a distinction between polynomial (short) and exponential (long) timescales. As an illustration, consider a population of size $N = 100$, a polynomial function $N^2$ and an exponential function $2^N$. Then $N^2 = 10^4$ steps correspond to $100$ generations, which is a moderate number. In contrast, $2^N \approx 10^{30}$ steps correspond to $10^{28}$ generations which is effectively infinite for all practical purposes. As a consequence, when evolutionary timescales are exponential, quantities such as fixation probability are largely irrelevant.

In order to present our results, we define certain natural two-patch structures $\mathrm{Isl}_{N,\varepsilon}(\mu)$ defined by three numbers $N$, $\varepsilon$, $\mu$ as follows: Two types of individuals are spread over two well-mixed patches of size $N$ each, thus the total population size is $2 \cdot N$. Each type of individuals is favored in one patch, by having relative fitness advantage $1 + \varepsilon$ rather than $1$, for some fixed $\varepsilon > 0$. Finally, whenever an individual reproduces, its offspring migrates to the other patch with probability $\mu$ and replaces a random individual there (otherwise it replaces a random individual within its patch). From the perspective of population genetics, the structure $\mathrm{Isl}_{N,\varepsilon}(\mu)$ thus corresponds to an island model [Wri31, Wri43] with two islands of equal size $N$ and a bi-directional migration rate $\mu$.

Here we present two results on coexistence times of environmentally heterogeneous populations. First, we show that on well-mixed populations, the process terminates after a number of steps that is of the order of at least $N^2$ and at most $N^3$. When compared to the $N \log N$ steps on homogeneous well-mixed populations, the process is thus slowed down but the expected number of steps is still only polynomial. Second, we show that the long-term evolution of a Moran Birth-death process on the two-island population structure $\mathrm{Isl}_{N,\varepsilon}(\mu)$ crucially depends on the migration rate $\mu$ between the two patches. Specifically, we show that when $\mu \geq 1/2$ then, with high probability, the stochastic process terminates after a number of steps that is only polynomial in $N$, regardless of $\varepsilon > 0$. In sharp contrast, when $\mu < 1/2$ then, with high probability, each type constitutes a majority in the patch where it is favored for a number of steps that is at least exponential in $N$. When combined, those two results present a strong dichotomy and precisely delineate the scenarios that support long-term coexistence of the two types.

## 5.2 Model

We consider Moran Birth-death process acting on structured populations consisting of two types of individuals $T_1$ and $T_2$. First we recall the general framework of evolutionary graph theory [LHN05], including the environmental heterogeneity [MP14].

**Spatial structure.** The spatial structure of a population is described by a connected graph (network) $G = (V, E)$ whose nodes $u \in V$ correspond to sites. Each site is occupied by a single individual and the edges $(u, v) \in E$ (including self-loops) represent where an individual can place an offspring. Moreover, each edge $(u, v) \in E$ is assigned a weight $w_{u,v} \in (0, 1]$ that represents the strength of the connection. The well-mixed population is represented by a complete graph $K_N$ where all edges and self-loops have unit weight.

**Environmental heterogeneity.** On top of that, each node $u$ is assigned a *signature* $\mathrm{sg}(u) = (f(u)_1, f(u)_2)$, where $f(u)_i$ denotes the fitness of a type $T_i$ individual when it occupies the site $u$ (for $i \in \{1, 2\}$). A set of nodes that all have the same signature is called a *patch*. In this work, we consider populations formed by two patches $P_1$, $P_2$, each patch favoring the corresponding type by the same margin $\varepsilon$. In other words, for $u \in P_1$ we have $\mathrm{sg}(u) = (1 + \varepsilon, 1)$, whereas for $u \in P_2$ we have $\mathrm{sg}(u) = (1, 1 + \varepsilon)$, for some fixed $\varepsilon > 0$. See Fig. 5.1a.

**Moran process.** The population evolves according to a Moran Birth-death process adapted to a population structure. We assume that initially each node in a patch $P_i$ is occupied by an individual of type $T_i$. Moran Birth-death process is a stochastic (random) process that proceeds in discrete time-steps as follows:

1. Birth: Select an individual randomly, with probability proportional to its fitness. (That is, denoting the total fitness of the population by $F$, an individual with fitness $f$ is selected with probability $f/F$.) That individual, say at node $u$, produces an offspring which is a copy of itself.
2. death: Select a node adjacent to $u$ randomly, with probability proportional to the edge weight $w_{u,v}$. (That is, a node $u'$ is selected with probability $w_{u,u'} / \sum_{v \in V} w_{u,v}$.) The offspring then migrates to site $u'$ and replaces its original inhabitant.

Note that throughout the process, the population size and structure remain constant. See Fig. 5.1b.

**Fixation time and $c$-coexistence time.** When the underlying graph $G$ is connected, Moran process eventually terminates with one type having spread over all nodes. This event is called *fixation*. For $i \in \{1, 2\}$ we denote by $\rho_i(G)$ the *fixation probability* of type $T_i$ and by $\mathrm{FT}(G)$ the *fixation time*, that is, the (expected) number of steps until the process terminates. Note that the fixation time ignores how the composition of the population fluctuates before the process terminates. To capture those fluctuations, we define a quantity which we call the coexistence time. Formally, given a constant $c \in [0, 1]$, we denote by $\mathrm{CT}^c(G)$ the *$c$-coexistence time*, that is, the (expected) number of steps until one of the types ceases to occupy a $c$-portion of the sites within the patch where it is favored. For instance, the $1/2$-coexistence time is the (expected) number of steps for which each type holds majority within its patch. By definition, a $c$-coexistence time can never be longer than a fixation time (but it can be substantially shorter).

**Population structures $\mathrm{Isl}_{N,\varepsilon}(\mu)$, $R^{\mathrm{alt}}_{N,\varepsilon}$, $R^{\mathrm{split}}_{N,\varepsilon}$.** Our main results apply to certain two-patch population structures $\mathrm{Isl}_{N,\varepsilon}(\mu)$ characterized by three parameters $N$, $\varepsilon > 0$, $\mu \in (0, 1)$ as follows: The population structure is a complete graph on $2N$ nodes split into two patches $P_1$, $P_2$ of size $N$ each. The edges within each patch all have unit weight, the edges connecting nodes in opposite patches all have weight $w = \mu/(1 - \mu)$. This choice of $w$ guarantees that an offspring of a reproducing individual migrates to the opposite patch with probability

$Nw/(Nw+N) = \mu$. Finally, each patch increases the fitness of one type from 1 to $1+\varepsilon$. That is, for $u \in P_1$ we have $\mathrm{sg}(u) = (1+\varepsilon, 1)$, whereas for $u \in P_2$ we have $\mathrm{sg}(u) = (1, 1+\varepsilon)$. Specifically, for migration rate $\mu = 1/2$ we recover the case of a well-mixed population $K_{N,\varepsilon}$. See Fig. 5.1c.

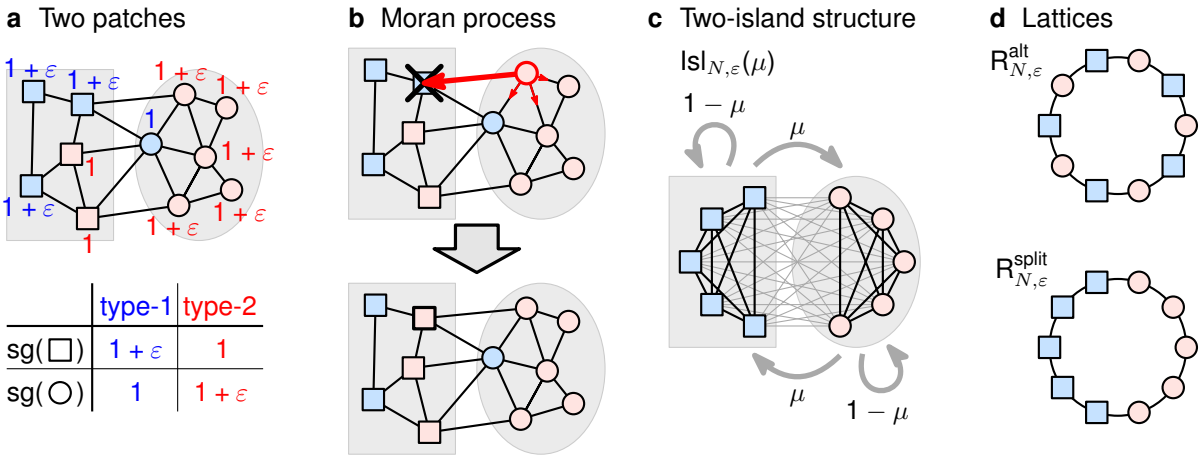



Figure 5.1: **Moran process on a population structure**. **a,** In the population structure, the $2N$ nodes (sites) are split into two patches of size $N$ (boxes, circles), each patch giving a relative fitness advantage $1+\varepsilon$ to one of two possible types of individuals (blue, red, respectively). **b,** In each step of the Moran process, first an individual is selected for reproduction proportionally to its fitness, and then the offspring replaces a random neighbor. **c,** In the two-island structure $\mathrm{Isl}_{N,\varepsilon}(\mu)$ each patch is a well-mixed population (island) and the offspring migrates to the other island with probability $\mu$. **d,** We also consider the 1-dimensional lattice, and two special decompositions of its nodes into patches: The nodes either alternate ($R^{\mathrm{alt}}_{N,\varepsilon}$), or they form two large blocks of $N$ consecutive nodes ($R^{\mathrm{split}}_{N,\varepsilon}$).

Later, we also study 1-dimensional lattices and we investigate how the coexistence time depends on the relative layout of the two patches. The underlying graph is a cycle $C_{2N}$ where the $2N$ nodes are arranged circularly and each node is connected to its two neighbors by an edge with unit weight. We consider two ways to partition the nodes into patches $P_1$, $P_2$: In $R^{\mathrm{alt}}_{N,\varepsilon}$, the nodes alternately belong to $P_1$ and $P_2$. In $R^{\mathrm{split}}_{N,\varepsilon}$, both $P_1$ and $P_2$ consist of a chunk of $N$ consecutive nodes. See Fig. 5.1d.

**Asymptotic notation.** In order to compare the relative growth rate of fixation and coexistence times in the limit of large population size $N \to \infty$, we briefly recall a standard mathematical notation $\Theta(\cdot)$, $\mathcal{O}(\cdot)$ and $\Omega(\cdot)$ for asymptotic tight bound, upper bound and lower bound, respectively. For example, we write $\frac{1}{2}N(N+1) \in \Theta(N^2)$ and $N^2 \in \mathcal{O}(N^3)$ and $2^N \in \Omega(N^3)$ to denote that, up to constant factors, for large $N$ we have $\frac{1}{2}N(N+1) \approx N^2 \ll N^3 \ll 2^N$. Moreover, we say that a function $f(N)$ is (at most) *polynomial* if $f(N) \in N^{\mathcal{O}(1)}$ and it is (at least) *exponential* if $f(N) \in 2^{\Omega(N)}$. For detailed treatment see [Cor09, Section 1.3]. We note that the distinction between polynomial and exponential growth rate is fundamental. For example, problems in computer science whose solution can be found in polynomial time are considered tractable in practice, whereas problems requiring exponential time are considered intractable.

## 5.3 Results

Here we state our analytical results, give intuition about their proofs, and illustrate them with numerical computations and computer simulations. The fully rigorous mathematical proofs are deferred to Section 5.5.

Recall that in all instances, we consider populations of $2N$ individuals split into two patches $P_1$, $P_2$ of size $N$ each, each patch increasing the fitness of the respective type of individual from 1 to $1+\varepsilon$, for some fixed $\varepsilon > 0$.

Our contribution is two-fold: First, we analyze the process on three different types of natural population structures, namely the well-mixed populations, the two-island graphs, and different one-dimensional lattices. Second, we prove a general upper bound on the timescale of coexistence.

**Complete graphs $K_{N,\varepsilon}$.** First, we consider the case of a well-mixed population $K_{N,\varepsilon}$ spanning two different patches of size $N$ each. Formally, $K_{N,\varepsilon}$ is recovered from the two-island population structure $\mathrm{Isl}_{N,\varepsilon}(\mu)$ by setting the migration rate equal to $\mu = 1/2$.

**Theorem 15** (Well-mixed populations)**.** *Fix $\varepsilon > 0$. Then*

$$\mathrm{FT}(K_{N,\varepsilon}) \in \Omega(N^2) \quad \textit{and} \quad \mathrm{FT}(K_{N,\varepsilon}) \in \mathcal{O}(N^3).$$

It is known that in the environmentally homogeneous regime where one type has fitness advantage $1+\varepsilon$ at all nodes, the fixation time on a well-mixed population is of the order of $\Theta(N \log N)$ steps. [AT09, DGRS16] The asymptotic lower bound $\mathrm{FT}(K_{N,\varepsilon}) \in \Omega(N^2)$ in Theorem 15 thus implies that with environmental heterogeneity, the fixation time is increased. However, the asymptotic upper bound $\mathrm{FT}(K_{N,\varepsilon}) \in \mathcal{O}(N^3)$ implies that the process still terminates after a number of steps that is only polynomial in the population size, and thus long-term coexistence is not supported. Numerical computation suggests that for any $\varepsilon > 0$ the fixation time in fact scales as $\Theta(N^2)$, see Fig. 5.2a.

The idea behind the proof is as follows: We represent the stochastic process as a Markov chain that has a state for every possible configuration of mutants and residents. We then define a carefully chosen "potential function" that assigns a real number to each possible configuration and we prove that, in each step of the process, this potential function changes in a controlled way, in expectation. Namely, we prove that it increases at most by $c_1$ and at least by $c_2$, where $c_1 > c_2 > 0$ are two real constants. Since we can also compute the initial and the final value of this potential function, this allows us to bound the expected number of steps that happen until fixation occurs. See Section 5.5 for a full proof.

**Two-island graphs $\mathrm{Isl}_{N,\varepsilon}(\mu)$.** Second, we show that for the two-island population structure $\mathrm{Isl}_{N,\varepsilon}(\mu)$, the migration rate $\mu^\star = 1/2$ which corresponds to the well-mixed population is in fact a threshold value. Recall that for a given population structure $G_N$, the quantity $\mathrm{CT}^{1/2}(G_N)$ is the (expected) number of steps until either of the types ceases to hold majority in the patch where it is favored.

**Theorem 16** (High migration rate)**.** *Fix $\varepsilon > 0$ and $\mu \geq 1/2$. Then* $\mathrm{FT}(\mathrm{Isl}_{N,\varepsilon}(\mu)) \in \mathcal{O}(N^3)$.

**Theorem 17** (Low migration rate)**.** *Fix $\varepsilon > 0$ and $\mu < 1/2$. Then* $\mathrm{CT}^{1/2}(\mathrm{Isl}_{N,\varepsilon}(\mu)) \in 2^{\Omega(N)}$.

Theorem 16 states that when the migration rate exceeds the threshold value $\mu^\star = 1/2$ or is equal to it, then fixation time is still only polynomial in the population size, thus long-term

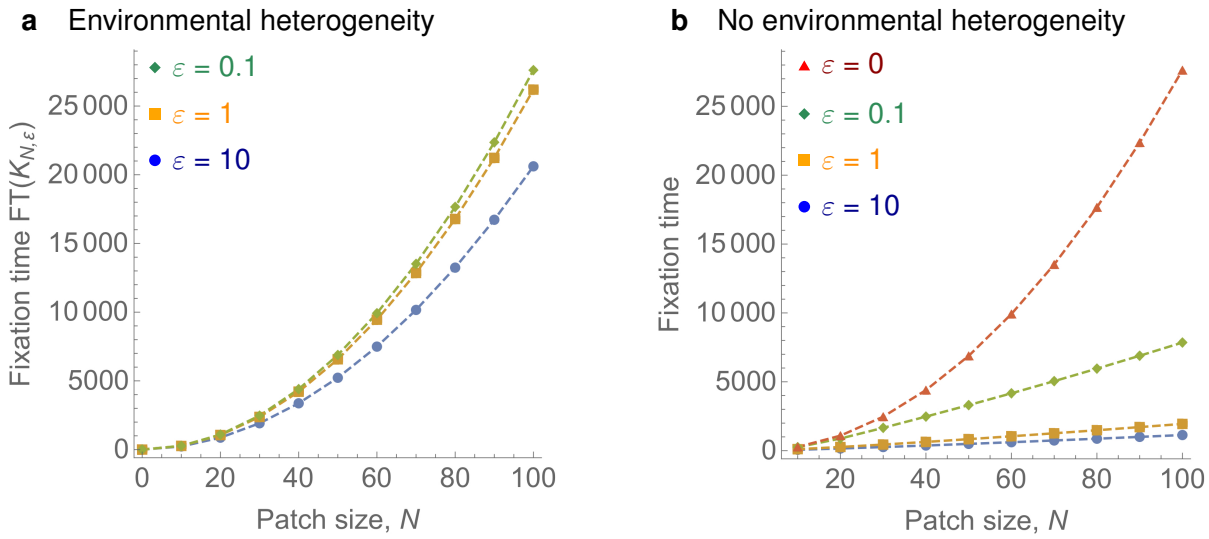



Figure 5.2: **Fixation time on a Complete graph $K_{N,\varepsilon}$. a,** Numerical computation shows that for a fixed $\varepsilon > 0$ the fixation time $\mathrm{FT}(K_{N,\varepsilon})$ scales as $c_\varepsilon \cdot N^2$. Thus the lower bound $\mathrm{FT}(K_{N,\varepsilon}) \in \Omega(N^2)$ from Theorem 15 is tight. Specifically, we obtain $c_{10} \doteq 2.07$, $c_{0.1} \doteq 2.62$ and $c_{0.01} \doteq 2.76$. **b,** In the regime without environmental heterogeneity (that is, when the same type is favored in both patches), the fixation time is known to scale as $\Theta(N \log N)$ when $\varepsilon > 0$ and as $\Theta(N^2)$ when $\varepsilon = 0$. Here $N = 10, 20, \ldots, 100$.

coexistence is not supported. In contrast, Theorem 17 shows that for migration rates $\mu < \mu^\star$, each type maintains a majority in the patch where it is favored for a number of generations that is exponential in the population size, see Fig. 5.3.

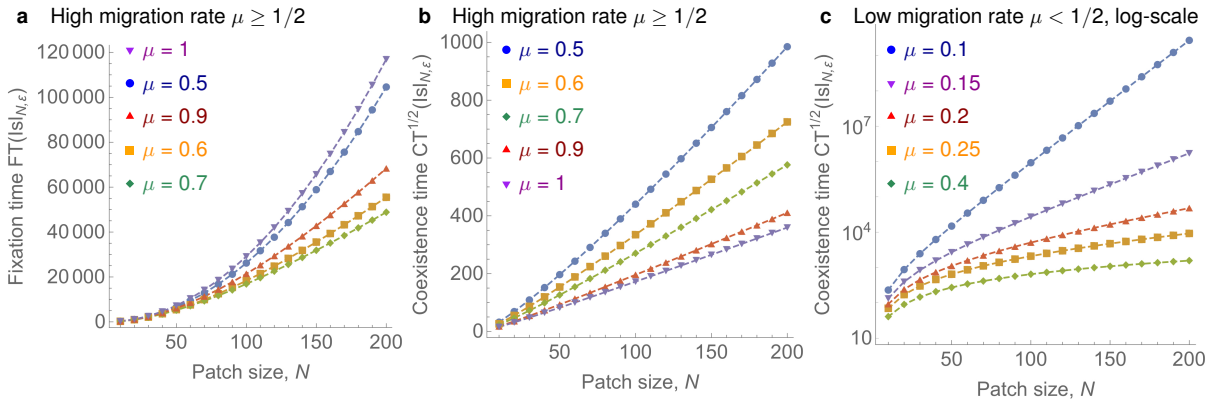



Figure 5.3: **Fixation time on a two-island graph $\mathrm{Isl}_{N,\varepsilon}(\mu)$. a,** When $\mu \geq 1/2$, the fixation time $\mathrm{FT}(\mathrm{Isl}_{N,\varepsilon}(\mu))$ scales as $N^2$ (when $\mu = 0.5$ or $\mu = 1$), or even slower than that (when $0.5 < \mu < 1$). This is in agreement with the upper bound $\mathrm{FT}(\mathrm{Isl}_{N,\varepsilon}(\mu)) \in \mathcal{O}(N^3)$ from Theorem 16. **b,** The coexistence time $\mathrm{CT}^{1/2}(\mathrm{Isl}_{N,\varepsilon}(\mu))$ is substantially shorter, scaling roughly linearly with the population size $N$. **c,** In contrast, when $\mu < 1/2$, the coexistence time $\mathrm{CT}^{1/2}(\mathrm{Isl}_{N,\varepsilon}(\mu))$, and thus also the fixation time, is at least exponential in the population size $N$ (here the $y$-axis is log-scale). This is in agreement with Theorem 17. In all panels, we consider $\varepsilon = 1$ and $N = 10, 20, \ldots, 200$.

The proof of Theorem 16 is an extension of the argument used to derive Theorem 15. As for the argument behind the proof of Theorem 17, intuitively the idea is to show that in order for one type to lose majority in the patch where it is favored, the random evolutionary trajectory would have to cross one of three "barriers", each with a "thickness" that is linear

in $N$. Using standard results on the absorption time of one-dimensional random walks with constant forward bias, this allows us to conclude that for a number of steps that is exponential in $N$, with high probability none of the three barriers will be crossed and thus coexistence will be maintained. See Section 5.5 for a full proof.

**General upper bound.** Next we show that within a certain broad class of population structures, a coexistence on a substantially longer than exponential timescale is impossible. Thus the exponential coexistence that occurs for two-island structures with low migration rates is close to optimal. Namely, given a real number $w_{\min} > 0$ we denote by $\mathcal{G}(w_{\min})$ the class of all connected graphs in which each edge is assigned a weight at least $w_{\min}$.

**Theorem 18** (General upper bound). *Fix $w_{\min} > 0$ and $\varepsilon > 0$. Then for any population structure $G_{N,\varepsilon} \in \mathcal{G}(w_{\min})$ on $2N$ nodes we have $\mathrm{FT}(G_{N,\varepsilon}) \in 2^{\mathcal{O}(N \cdot \log N)}$.*

Note that Theorem 18 gives an upper bound on the fixation time, and thus also on the $c$-coexistence time for any $c > 0$. The proof is based on a simple idea that if, from any configuration of individuals, a fixation occurs with probability at least $p_{\mathrm{fix}}$ within the next $s$ steps, then the fixation time is at most $s/p_{\mathrm{fix}}$. Moreover, the statement applies more generally to any number of patches of arbitrary sizes and with arbitrary (but fixed) signatures. See Section 5.5 for a full proof.

**One-dimensional lattices.** Finally, we show that even if we fix the underlying graph and the effects of the patches on the fitness, the coexistence time critically depends on the relative layout of the two patches. To that end, we consider large one-dimensional lattices $R^{\mathrm{alt}}(N, \varepsilon)$ and $R^{\mathrm{split}}(N, \varepsilon)$ (see Fig. 5.1d).

**Theorem 19** (One-dimensional lattices). *Fix $\varepsilon > 0$. Then*

$$\mathrm{FT}(R^{\mathrm{alt}}(N, \varepsilon)) \in \mathcal{O}(N^3) \quad \textit{and} \quad \mathrm{CT}^{1/2}(R^{\mathrm{split}}(N, \varepsilon)) \in 2^{\Omega(N)}.$$

In other words, when the nodes of a long one-dimensional lattice alternately belong to patches $P_1$ and $P_2$ then the process terminates in polynomial time. In contrast, when each patch forms a contiguous block of $N$ nodes, each type holds majority in its patch for a number of steps that is exponential in $N$, see Fig. 5.4.

The argument behind the first claim makes use of the fact that the evolution on $R^{\mathrm{alt}}(N, \varepsilon)$ can be efficiently mapped to an evolution on a well-mixed population. [MP14, KMCN20] The idea behind the proof of the second claim is that each patch contains a "core" – a large set of nodes in its middle which is well protected from the invasion by the other type, and which is maintained over an exponential timescale. See Section 5.5 for a full proof.

## 5.4 Discussion

In this work, we used the framework of evolutionary graph theory to study the evolutionary timescales of populations that are both spatially structured and environmentally heterogeneous. To our knowledge, the previous research in this setting focused either on computing the fixation probabilities, or on identifying population structures with short evolutionary timescales. In contrast, our main focus here was to charaterize structures that support long-term coexistence of two competing types.

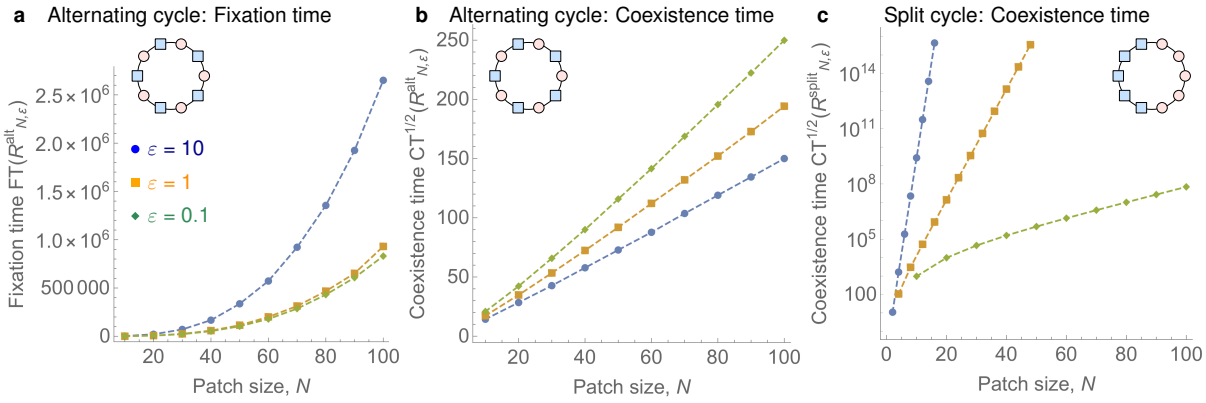



Figure 5.4: **Fixation time on one-dimensional lattices. a,** Computer simulations ($10^4$ repetitions) show that on the Alternating cycle $R^{\mathrm{alt}}(N,\varepsilon)$ the fixation time $\mathrm{FT}(R^{\mathrm{alt}}(N,\varepsilon))$ scales as $\Theta(N^3)$, for any $\varepsilon > 0$. This is in perfect agreement with the upper bound $\mathcal{O}(N^3)$ from Theorem 19. **b,** The coexistence time $\mathrm{CT}^{1/2}(R^{\mathrm{alt}}(N,\varepsilon))$ is even shorter – it scales roughly as $\Theta(N)$, for any $\varepsilon > 0$. **c,** In contrast, on the Split cycle $R^{\mathrm{split}}(N,\varepsilon)$, the $\frac{1}{2}$-coexistence time $\mathrm{CT}^{1/2}(R^{\mathrm{split}}(N,\varepsilon))$ is exponential in the population size $N$ (note that the $y$-axis is log-scale). In all panels, we consider $\varepsilon \in \{10, 1, 0.1\}$ and $N$ up to 100.

To that end, we considered the Moran Birth-death process acting on populations stretched across two patches, where each of the two competing types has a fixed selective advantage in one patch. To address the question of long-term coexistence of the two types, we refined the classical notion of a fixation time and defined the coexistence time as the (expected) number of steps during which each type constitutes a majority in the patch where it is favored. For certain natural two-island population structures characterized by a migration rate $\mu$, we then showed that the long-term behavior of the system exhibits a threshold behavior in parameter $\mu$: When the migration rate is high ($\mu \geq 1/2$), the process terminates with one type fixating on the whole population after a number of steps that is only polynomial in the population size $N$. In sharp contrast, when the migration rate is low ($\mu < 1/2$), the two types coexist for a number of steps that is (at least) exponential in $N$. We also showed that on a fixed population structure given by a 1-dimensional lattice, the coexistence time can be both polynomial and exponential, depending on the relative layout of the two patches.

Coexistence of multiple types of individuals in structured populations has been extensively studied in various fields. Below we list some results for related models, highlighting the differences to our setup. In the terminology of population genetics, the analogue of our question is whether polymorphism is protected in a one-locus, two-allele population inhabiting a two-niche ecosystem. Following the seminal work of Levene [Lev53], a now classical line of research identified several necessary and sufficient conditions for the polymorphism to be maintained [Dea66, LM66, MS70, KM72, Chr74, GS75, Gil75]. However, those results are derived for deterministic models under the assumption of infinite population size, see also [CF75, Fel76] for reviews. In contrast, our model accounts for stochasticity inherent to the evolutionary process [DL94a], it deals with populations of finite (and arbitrarily large) size, and it classifies the fate of the population in terms of the polynomial-exponential dichotomy. In statistical physics, similar models are studied under the name of interacting particle systems [Lig12]. Those models are stochastic and spatial but the spatial structure is typically assumed to be an integer lattice. Again, conditions characterizing coexistence are known [Dur02, D$^+$09].

The most closely related results to ours are those that study stochastic models on spatial structures that correspond to island models of finite size. As with the simpler regime without any environmental heterogeneity, fixation probability is a quantity that is relatively approachable. [Ave78, TI91, GG02, WG05]. Regarding fixation time, building on the work of Bulmer [Bul72], Yeaman and Otto [YO11] use computer simulations to observe that polymorphism is maintained when the migration rate exceeds a certain critical threshold and they approximate the threshold by splicing the predictions of the deterministic models with a diffusion approximation for finite populations. While their results are qualitatively similar to ours, there are important differences. First, we work with a purely stochastic model and we directly analyze the underlying Markov chain which allows us to obtain exact rigorous mathematical results without resorting to approximations. Second, we highlight the fundamental distinction between polynomial and exponential timescale.

Within the field of evolutionary graph theory, the past research agenda related to evolutionary timescale focused mostly on understanding which population structures lead to short timescales. Here our objective is the exact opposite – we study population structures that lead to long evolutionary timescales. Furthermore, using the refined notion of a coexistence time (rather than simply a fixation time) we are able to guarantee that not only will the evolutionary process run for exponentially many steps, but also that throughout that timeframe each type will constitute a healthy portion of the population.

In this work we focus on the simple population structures, such as the well-mixed populations, one-dimensional lattices, and two-island populations. This allows us to cleanly introduce the key notions and derive exact analytical results. It is our hope that a subsequent work on more complex population structures, possibly with multiple patches and with multiple competing types, will lead towards a better understanding of the role of diversity and its maintenance in populations at large.

## 5.5 Additional proofs

### 5.5.1 Preliminaries

Recall that we consider Moran Birth-death process adapted to a population structure, where each node belongs to one of two *patches* $P_1$, $P_2$ and a patch $P_i$ confers a fitness advantage $1+\varepsilon$ onto an individual of type $T_i$. In other words, the nodes in $P_1$ have signature $(1+\varepsilon, 1)$ whereas the nodes in $P_2$ have signature $(1, 1+\varepsilon)$. For details, see Section Model in the main text.

Next we introduce several notions that repeatedly appear in the proofs of our theorems.

We denote by $N$ the population size, by $[N] = \{0, 1, \ldots, N\}$ the set of non-negative integers up to $N$, and by $[N^+] = \{1, \ldots, N\}$ the set of positive integers up to $N$. At any given time-point, a *configuration* describes which vertices are occupied by which individuals. (Formally, it is a mapping from the vertices to the types.) Given a current configuration $X$, the total fitness of the population is denoted by $F = F(X)$. Note that $2N \leq F(X) \leq 2N \cdot (1+\varepsilon)$, thus $F(X) \in \Theta(N)$ for all configurations $X$.

We say that an individual is *home* (resp. *not-home*) if it occupies a vertex in a patch where it is (resp. is not) favored. In the initial configuration, all individuals are home.

We say that an edge is *active* at a given time-point if its two endpoints are occupied by individuals of different types. A single step of the process is *active* if the offspring migrates

along an active edge. Note that steps that are not active do not change the configuration (but they are counted when computing the fixation time).

### 5.5.2 General potential

In this section, we prove a lemma that bounds the absorption time for certain families of two-dimensional Markov chains. We call such families of Markov chains *potent*. (Later, we show that the Markov chains that describe the evolution on the complete graphs and on the two-island population structure $\mathrm{ISL}_{N,\varepsilon}(\mu)$ are in fact potent.)

First we introduce some notation. We consider Markov chains $M_N$ parametrized by an integer $N$ such that the states form a set $S_N = \{(a,b) \mid 0 \le a,b \le N\}$, the initial state is $(N,N)$, there are two absorbing states $(0,N)$ and $(N,0)$, and any transition changes at most one coordinate by at most one. (Later, state $(a,b) \in S_N$ will correspond to a configuration with $a$ home type-1 individuals and $b$ home type-2 individuals in a population of $2N$ individuals on a certain population structure.) Given a state $(a,b) \in S_N$, we denote the transition probabilities by

$$\begin{aligned} a_N^+(a,b) &= \mathbb{P}[(a,b) \to (a+1,b)], & a_N^-(a,b) &= \mathbb{P}[(a,b) \to (a-1,b)], \\ b_N^+(a,b) &= \mathbb{P}[(a,b) \to (a,b+1)], & b_N^-(a,b) &= \mathbb{P}[(a,b) \to (a,b-1)], \\ p_N^{\mathrm{loop}}(a,b) &= \mathbb{P}[(a,b) \to (a,b)], & p_N^{\mathrm{act}}(a,b) &= 1 - p_N^{\mathrm{loop}}(a,b). \end{aligned}$$

The quantity $p_N^{\mathrm{act}}(a,b)$ is the probability that the Markov chain makes a step that changes the state (this corresponds to the process making an active step that changes the current configuration).

We note that the quantities such as $a_N^+(a,b)$ depend on both $N$ and the pair $(a,b)$. In what we write below, the values of $N$ and/or $(a,b)$ are often fixed and understood from the context. For ease of notation, in those cases we omit the explicit dependence and we write e.g. $a^+ = a_N^+(a,b)$ and so on.

**Definition 2** (Potent Markov chains)**.** *Given $c > 0$, we say that a sequence $(M_N)_{N \ge 1}$ of Markov chains of the above form is* $c$-potent *if the following conditions hold for each $N \ge 1$:*

1. *(Symmetry) For any $(a,b) \in S_N$ we have $b_N^+(a,b) = a_N^+(b,a)$ and $b_N^-(a,b) = a_N^-(b,a)$.*
2. *(Special form) There exist functions $g_N, h_N \colon S_N \to \mathbb{R}$ such that for all $(a,b) \in S_N$ we have $g_N(a,b) = g_N(b,a)$, $h_N(a,b) = h_N(b,a) \le 0$, and*

$$a_N^+(a,b) - a_N^-(a,b) = (N + c \cdot a) \cdot g_N(a,b) + h_N(a,b).$$

An example of a 1-potent family of Markov chains is a family, where $a_N^+(a,b) = \frac{N^2+N+2a+b+2}{N^3}$, $a_N^-(a,b) = \frac{(N^2+a+1)(N^2+b+1)}{N^5}$, $b_N^+(a,b) = \frac{N^2+N+2b+a+2}{N^3}$, $b_N^-(a,b) = \frac{(N^2+b+1)(N^2+a+1)}{N^5}$, for each $N \ge 1$. Indeed, this is witnessed by functions $g_N(a,b) = \frac{1}{N^3}$, and $h_N(a,b) = \frac{-(a+1)(b+1)}{N^5}$. For other examples, see the Markov chains corresponding to the complete graph (Section 5.5.3) and the two-island graph (Section 5.5.4).

The following lemma is our key technical contribution. It bounds the absorption time $\mathrm{AT}(M_N)$ of a $c$-potent family of Markov chains in terms of $c$ and the bounds on the functions $h_N$ and the probabilities $p_N^{\mathrm{act}}$ of making an active step.

**Lemma 23.** *Fix $c > 0$ and suppose that $(M_N)_{N\geq 1}$ is a $c$-potent family of Markov chains characterized by functions $g_N, h_N \colon S_N \to \mathbb{R}$. Suppose that for each $N \geq 1$ and each $(a,b) \in S_N$ we have bounds $p_N^{\min} \leq p_N^{\mathrm{act}}(a,b) \leq p_N^{\max}$ and $|h_N(a,b)| \leq h_N^{\max}$. Then the expected absorption time $\mathrm{AT}(M_N)$ satisfies*

$$\mathrm{AT}(M_N) \in \mathcal{O}\left(\frac{N^2}{p_N^{\min}}\right) \quad \textit{and} \quad \mathrm{AT}(M_N) \in \Omega\left(\min\left\{\frac{N^2}{p_N^{\max}}, \frac{N}{h_N^{\max}}\right\}\right).$$

*Proof.* Fix $N \geq 1$ and denote by $S = S_N$ the state space of the Markov chain $M = M_N$. Let $f(t) = N + c \cdot t$ and consider a function $\varphi \colon S \to \mathbb{R}$ defined by

$$\varphi(a,b) = \frac{f(a)}{f(b)} + \frac{f(b)}{f(a)} = \frac{N+ca}{N+cb} + \frac{N+cb}{N+ca}.$$

Below we show that the value $\varphi(a,b)$ changes in a controlled way as we run the process. Namely, we show that for any $(a,b) \in S \setminus \{(0,N),(N,0)\}$, the expected increase $\Delta\varphi(a,b)$ upon performing a single step of the Moran process is sandwiched between two positive constants $\Delta_{\min}$, $\Delta_{\max}$. Since $\varphi(a,b) > 0$ for any $(a,b) \in S$, this allows us to bound the absorption time $\mathrm{AT}(M)$ using a standard drift analysis and martingale machinery, see e.g. Theorem 1 in [Len19] and other references therein. Namely, denoting by $\varphi_{\max} = \max_{(a,b)\in S}\{\varphi(a,b)\}$ the maximum attainable potential value, we obtain

$$\frac{\varphi_{\max}}{\Delta_{\max}} \leq \mathrm{AT}(M) \leq \frac{\varphi_{\max}}{\Delta_{\min}}.$$

It remains to compute the asymptotics of $\varphi_{\max}$, $\Delta_{\min}$, and $\Delta_{\max}$. Regarding $\varphi_{\max}$, we clearly have $\varphi_{\max} \geq 1 + 1 = 2$ and $\varphi_{\max} \leq 2 \cdot (N + cN)/N = 2(1+c)$, thus $\varphi_{\max} \in \Theta(1)$ is a constant. The computation for $\Delta_{\min}$, and $\Delta_{\max}$ is conceptually straightforward, but technically much more demanding. In particular, below we show that

$$\Delta_{\min} \in \Omega\left(\frac{p_N^{\min}}{N^2}\right) \quad \text{and} \quad \Delta_{\max} \in \mathcal{O}\left(\frac{p_N^{\max}}{N^2} + \frac{h_N^{\max}}{N}\right)$$

from which the statement of the lemma follows.

Fix $(a,b) \in S \setminus \{(0,N),(N,0)\}$. The expected change of the potential in a single step can be expressed as

$$\begin{aligned}\Delta\varphi(a,b) = a^+ \cdot (\varphi(a+1,b) - \varphi(a,b)) + a^- \cdot (\varphi(a-1,b) - \varphi(a,b)) \\ + b^+ \cdot (\varphi(a,b+1) - \varphi(a,b)) + b^- \cdot (\varphi(a,b-1) - \varphi(a,b)).\end{aligned}$$

The potential $\varphi(a,b)$ is itself a sum of two parts, $\frac{f(a)}{f(b)}$ and $\frac{f(b)}{f(a)}$. First, we look at the first part only, we substitute for the difference $f(t+1) - f(t) = c$, and we collect terms with the same

power of $c$. We get

$$\begin{aligned}
\Delta_1 &= a^+\frac{f(a+1)-f(a)}{f(b)} + a^-\frac{f(a-1)-f(a)}{f(b)} + \\
&\quad b^+f(a)\left(\frac{1}{f(b+1)}-\frac{1}{f(b)}\right) + b^-f(a)\left(\frac{1}{f(b-1)}-\frac{1}{f(b)}\right) \\
&= a^+\frac{c}{f(b)} - a^-\frac{c}{f(b)} - b^+f(a)\frac{c}{f(b+1)f(b)} + b^-f(a)\frac{c}{f(b-1)f(b)} \\
&= \frac{c}{f(b)}\left(a^+ - a^- - \frac{b^+f(a)}{f(b)+c} + \frac{b^-f(a)}{f(b)-c}\right) \\
&= \frac{c}{f(b)}\left(a^+ - a^- - \frac{f(a)f(b)(b^+-b^-) - cf(a)(b^++b^-)}{f(b)^2-\delta^2}\right) \\
&= \frac{c}{f(b)}\left(a^+ - a^- - \frac{f(a)f(b)(b^+-b^-)}{f(b)^2-\delta^2}\right) + \frac{c^2f(a)(b^++b^-)}{f(b)\left(f(b)^2-c^2\right)} \\
&= \frac{c}{f(b)\left(f(b)^2-c^2\right)}\Big((a^+-a^-)(f(b)^2-c^2) - f(a)f(b)(b^+-b^-)\Big) + \frac{c^2f(a)(b^++b^-)}{f(b)\left(f(b)^2-c^2\right)} \\
&= \frac{c}{f(b)\left(f(b)^2-c^2\right)}\Big((a^+-a^-)f(b)^2 - f(a)f(b)(b^+-b^-)\Big) \\
&+ \frac{c^2f(a)(b^++b^-)}{f(b)\left(f(b)^2-c^2\right)} - \frac{c^3(a^+-a^-)}{f(b)\left(f(b)^2-c^2\right)}.
\end{aligned} \tag{5.1}$$

For now, we focus on the first term $L_1$ which is linear in $c$ (we get back to the quadratic and the cubic terms later). We cancel $f(b)$ and rewrite $L_1$ further using the relation $a^+ - a^- = f(a)g(a,b) + h(a,b)$ as follows:

$$\begin{aligned}
L_1 &= \frac{c\left((a^+-a^-)f(b) - f(a)(b^+-b^-)\right)}{f(b)^2-c^2} \\
&= \frac{c\left((f(a)g(a,b)+h(a,b))f(b) - f(a)(f(b)g(b,a)+h(b,a))\right)}{f(b)^2-c^2} \\
&= \frac{c\left(f(a)g(a,b)f(b) - f(a)f(b)g(b,a)\right)}{f(b)^2-c^2} + \frac{c\left(f(b)h(a,b) - f(a)h(b,a)\right)}{f(b)^2-c^2} \\
&= \frac{cf(a)f(b)\left(g(a,b)-g(b,a)\right)}{f(b)^2-c^2} + \frac{c\left(f(b)h(a,b) - f(a)h(b,a)\right)}{f(b)^2-c^2}.
\end{aligned}$$

We know that $g(a,b) = g(b,a)$, so the first term $\frac{cf(a)f(b)(g(a,b)-g(b,a))}{f(b)^2-c^2}$ disappears.

Next we use the fact that $h(a,b) = h(b,a)$ and we sum whatever is left of $L_1$ with the corresponding term $L_2$ coming from the part $\frac{f(b)}{f(a)}$ of the potential. (This is the same

expression, just with exchanged $a$ and $b$.). We get

$$\begin{aligned} L_1 + L_2 &= \frac{c\,(f(b)h(a,b) - f(a)h(b,a))}{f(b)^2 - c^2} + \frac{c\cdot(f(a)h(b,a) - f(b)h(a,b))}{f(a)^2 - c^2} \\ &= \frac{c\cdot h(a,b)\,(f(b) - f(a))}{f(b)^2 - c^2} - \frac{c\cdot h(a,b)\,(f(b) - f(a))}{f(a)^2 - c^2} \\ &= \frac{c\cdot h(a,b)\,(f(b) - f(a))\,(f(a)^2 - f(b)^2)}{(f(b)^2 - c^2)(f(a)^2 - c^2)} \\ &= -\frac{c\cdot h(a,b)\,(f(a) - f(b))^2\,(f(a) + f(b))}{(f(b)^2 - c^2)(f(a)^2 - c^2)}. \end{aligned}$$

Note that since $f(t) \geq 0$, $h(a,b) \leq 0$, and the inequality $f(a), f(b) \geq N > c$ holds for all large enough $N$, the combined linear term $L_1 + L_2$ is positive. On the other hand, since $f(t) \in \Theta(N)$ and $c \in \Theta(1)$, we can bound it from above as

$$L_1+L_2 = -\frac{c\cdot h(a,b)\,(f(a) - f(b))^2\,(f(a) + f(b))}{(f(b)^2 - c^2)(f(a)^2 - c^2)} \in \mathcal{O}\left(\frac{\max_{a,b}\{|h(a,b)|\}\cdot N^2\cdot N}{N^2\cdot N^2}\right) = \mathcal{O}\left(\frac{h^{\max}}{N}\right).$$

Now we get back to the higher order (quadratic and cubic) terms from Eq. (5.1). We again look at them together with the corresponding terms coming from the part $\frac{f(b)}{f(a)}$ of the potential. We get

$$H = \frac{c^2 f(a)(b^+ + b^-)}{f(b)\,(f(b)^2 - c^2)} - \frac{c^3(a^+ - a^-)}{f(b)\,(f(b)^2 - c^2)} + \frac{c^2 f(b)(a^+ + a^-)}{f(a)\,(f(a)^2 - c^2)} - \frac{c^3(b^+ - b^-)}{f(a)\,(f(a)^2 - c^2)}.$$

Finally, it remains to show that $H > 0$ and $H \in \Theta(p^{\mathrm{act}}/N^2)$. We rewrite $H = A^+ + B^+ + A^- + B^-$, where $A^+$ collects all the terms with $a^+$ and likewise for $b^+$, $a^-$, and $b^-$. We have

$$A^+ = a^+ \cdot \left(\frac{c^2 f(b)}{f(a)(f(a)^2 - c^2)} - \frac{c^3}{f(b)(f(b)^2 - c^2)}\right)$$

and since $f(t) \in \Theta(N)$ and $c \in \Theta(1)$, we can further rewrite this as

$$A^+ = a^+ \cdot \Theta\left(\frac{N}{N\cdot N^2} - \frac{1}{N\cdot N^2}\right) = a^+ \cdot \Theta\left(\frac{1}{N^2}\right).$$

Likewise, we get $B^+ = b^+ \cdot \Theta(1/N^2)$ by switching the roles of $a$ and $b$. For $A^-$ we proceed completely analogously (this time having a plus instead of a minus), again leading to

$$A^- = a^- \cdot \left(\frac{c^2 f(b)}{f(a)(f(a)^2 - c^2)} + \frac{c^3}{f(b)(f(b)^2 - c^2)}\right) = a^- \cdot \Theta\left(\frac{N}{N^3} + \frac{1}{N^3}\right) = a^- \cdot \Theta\left(\frac{1}{N^2}\right).$$

Similarly, we get $B^- = b^+ \cdot \Theta(1/N^2)$. Therefore we have

$$H = A^+ + B^+ + A^- + B^- = (a^+ + b^+ + a^- + b^-)\cdot\Theta(1/N^2) \in \Theta\left(\frac{p^{\mathrm{act}}}{N^2}\right),$$

which finishes the proof.

□

### 5.5.3 Complete graph

In this section, we prove Theorem 20 from the main text. Consider the complete graph $K_{N,\varepsilon}$ spanning two patches of size $N$ each. At any given time-point, the configuration can be described by a pair $(a, b)$ (with $0 \leq a, b \leq N$) where $a$ counts the type-1 individuals that are home and $b$ counts the type-2 individuals that are home. Thus the initial configuration is $(N, N)$ and fixation of types 1 and 2 occurs at configurations $(N, 0)$, $(0, N)$, respectively.

**Theorem 20** (Well-mixed populations)**.** *Fix $\varepsilon > 0$. Then*

$$\mathrm{FT}(K_{N,\varepsilon}) \in \Omega(N^2) \quad \textit{and} \quad \mathrm{FT}(K_{N,\varepsilon}) \in \mathcal{O}(N^3).$$

*Proof.* We aim to use Lemma 23. We show that the Markov chain corresponding to the process is $\varepsilon$-potent.

Consider any time-point and suppose the current configuration is $(a, b)$, that is, $a$ type-1 individuals and $b$ type-2 individuals are home. Let

$$F = (1+\varepsilon)a + N - a + (1+\varepsilon)b + N - b = 2N + \varepsilon(a+b) = \Theta(N)$$

be the total fitness of the population. Denote by $a^+$, $a^-$, $b^+$, and $b^-$ the probabilities that, after a single step of the process, the configuration becomes $(a+1, b)$, $(a-1, b)$, $(a, b+1)$ and $(a, b-1)$, respectively. Then

$$\begin{aligned}
a^+ &= \frac{(1+\varepsilon)a(N-a) + (N-b)(N-a)}{F \cdot (2N-1)}, \\
a^- &= \frac{(1+\varepsilon)ba + (N-a)a}{F \cdot (2N-1)}, \\
b^+ &= \frac{(1+\varepsilon)b(N-b) + (N-a)(N-b)}{F \cdot (2N-1)}, \\
b^- &= \frac{(1+\varepsilon)ba + (N-b)b}{F \cdot (2N-1)}.
\end{aligned}$$

This yields

$$a^+ - a^- = (N + \varepsilon a) \cdot \frac{N-a-b}{F \cdot (2N-1)},$$

thus by setting $g(a, b) = \frac{N-a-b}{F(N-1)}$ and $h(a, b) = 0$ we conclude that the Markov chain is indeed $\varepsilon$-potent. Thus, Lemma 23 applies and since $h(a, b) = 0$, it remains to find bounds $p^{\min}$, $p^{\max}$ on $p^{\mathrm{act}} = a^+ + a^- + b^+ + b^-$.

We clearly have $p^{\mathrm{act}} \leq 1$, so we can set $p^{\max} = 1$ and get

$$\mathrm{FT}(K_{N,\varepsilon}) \in \Omega\left(N^2/p^{\max}\right) = \Omega(N^2).$$

For the upper bound, note that until fixation we have $|a-b| \leq N-1$, hence by using $1+\varepsilon \geq 1$ we can bound

$$a^+ + a^- + b^+ + b^- \geq \frac{2}{F(2N-1)} \cdot (a + (N-b))(b + (N-a)) = \frac{2}{F(2N-1)} \cdot (N^2 - (a-b)^2) \geq \frac{2}{F} \in \Omega(1/N).$$

Therefore

$$\mathrm{FT}(K_{N,\varepsilon}) \in \mathcal{O}(N^2/p^{\min}) = \mathcal{O}(N^2/(1/N)) = \mathcal{O}(N^3)$$

as claimed. □

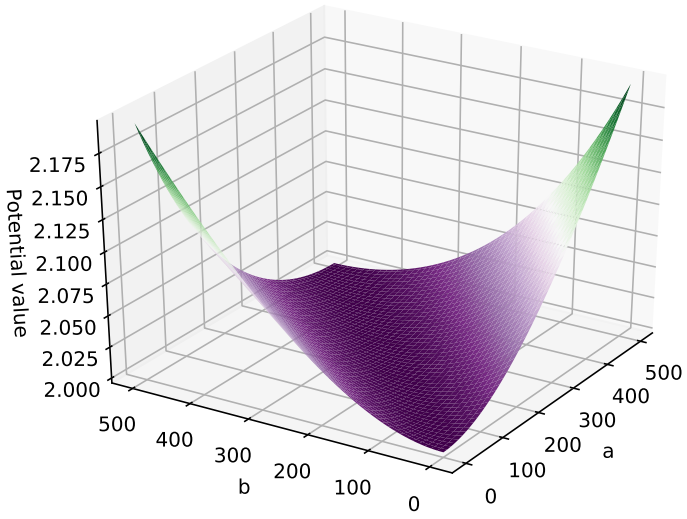



Figure 5.5: The surface of the potential $\varphi$ from Lemma 23 in the special case of $K_{N,\varepsilon}$ with $N = 500$, and $\varepsilon = 1.1$.

As an illustration, we plot the potential used in the proof of Lemma 23 in the special case of a Markov chain arising from the complete graph, see Fig. 5.5.

### 5.5.4 Two-island graphs

In this section, we prove Theorems 21 and 22 from the main text. Consider the two-island graphs $\mathrm{ISL}_{N,\varepsilon}(\mu)$, where each island constitutes one patch and the offspring migrates to the other island with probability $p$. As before, the current configuration can be described by a pair $(a, b)$ (with $0 \leq a, b \leq N$) where $a$ counts the type $T_1$ individuals that are home and $b$ counts the type $T_2$ individuals that are home. Again we denote by $a^+ = a^+_N(a, b)$, $a^-$, $b^+$, and $b^-$ the probabilities that, after a single step of the process, the configuration becomes $(a + 1, b)$, $(a - 1, b)$, $(a, b + 1)$ and $(a, b - 1)$, respectively, and by $F = 2N + \varepsilon(a + b) \in \Theta(N)$ the total fitness of the population. Direct computation gives

$$a^+ = \frac{(1-\mu)(1+\varepsilon)a(N-a) + \mu(N-b)(N-a)}{FN}$$
$$a^- = \frac{\mu(1+\varepsilon)ab + (1-\mu)(N-a)a}{FN}$$

and $b^+$ (resp. $b^-$) is obtained from $a^+$ (resp. $a^-$) by swapping $a$ and $b$.

#### High-migration regime

First we prove that when the migration rate is high, the fixation time is only polynomial, namely at most cubic.

**Theorem 21** (High migration rate)**.** *Fix $\varepsilon > 0$ and $\mu \geq 1/2$. Then* $\mathrm{FT}(\mathrm{ISL}_{N,\varepsilon}(\mu)) \in \mathcal{O}(N^3)$.

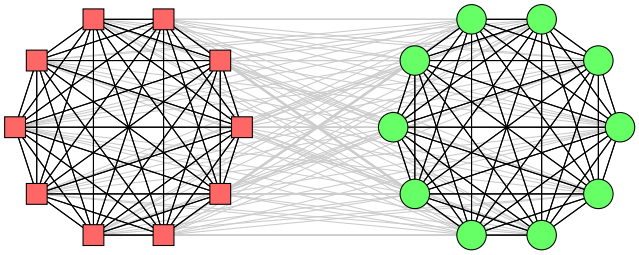


Figure 5.6: Two-island graph. Vertices have two types and are more strongly connected to vertices with the same signature then to vertices with a different signature.

*Proof.* Again we plan to use Lemma 23. We have

$$\begin{aligned}
a^+ - a^- &= \frac{(1-\mu)(1+\varepsilon)a(N-a) + \mu(N-b)(N-a)}{FN} - \frac{\mu(1+\varepsilon)ab + (1-\mu)(N-a)a}{FN} \\
&= \frac{1}{FN}((1-\mu)(1+\varepsilon)a(N-a) + \mu(N-b)(N-a) - \mu(1+\varepsilon)ab - (1-\mu)(N-a)a) \\
&= \frac{1}{FN}((\mu N + \varepsilon(1-\mu)a)(N-a-b) + (1-2\mu)\varepsilon ab) \\
&= \left(N + \frac{\varepsilon(1-\mu)}{\mu}a\right) \cdot \frac{\mu(N-a-b)}{FN} + \frac{(1-2\mu)\varepsilon ab}{FN},
\end{aligned}$$

so we can set $g = \frac{N-a-b}{FN}$ and $h = \frac{(1-2\mu)\varepsilon ab}{FN}$. Then $g$ and $h$ are symmetric in $a$ and $b$ and since $\mu \geq \frac{1}{2}$, we know that $h(a,b) \leq 0$. Therefore the corresponding Markov chains are $\frac{\varepsilon(1-\mu)}{\mu}$-potent, so Lemma 23 applies and it remains to show that $p_N^{\min} \in \Omega(1/N)$.

Using $1-\mu \geq 0$ and $1+\varepsilon \geq 1$ we have

$$p_N^{\mathrm{act}}(a,b) \geq a^+ + a^- \geq \frac{1}{FN}(ab + (N-a)(N-b)) \geq \frac{1}{FN} \cdot N \in \Theta(1/N),$$

where the last inequality holds for all pairs $(a,b) \in S_N \setminus \{(N,0),(0,N)\}$: Indeed, if $a = b \in \{0,N\}$ then $ab + (N-a)(N-b) = N^2 \geq N$, otherwise without loss of generality $a \notin \{0,N\}$. Then $a \geq 1$ and $N - a \geq 1$, hence $ab + (N-a)(N-b) \geq b + (N-b) \geq N$. □

### Low-migration regime

Here we prove that when the migration rate is low, the coexistence time (and thus the fixation time) is at least exponential.

**Theorem 22** (Low migration rate)**.** *Fix* $\varepsilon > 0$ *and* $\mu < 1/2$*. Then* $\mathrm{CT}(\mathrm{ISL}_{N,\varepsilon}(\mu)) \in 2^{\Omega(N)}$.

*Proof.* We use the same Markov chain as in the previous section. A state is determined by $a, b$. We identify "obstacles" between the starting position and both end positions. The obstacles will have thickness $\Omega(N)$, and we will show that it takes a long time to pass through them. There are three obstacles. The first and the third obstacle are symmetric in $a$ and $b$.

We set $w = \frac{(1-2\mu)\min(\varepsilon,1)}{8}$, $\gamma = \frac{w}{4}$, $x_1 = \frac{\varepsilon^3(1-2\mu)^2}{350(2+\varepsilon)}$, and $x_2 = \frac{\varepsilon(1-2\mu)}{64(1+\varepsilon)}$. Observe that $w$, $\gamma$, $x_1$, and $x_2$ are constants (they do not depend on $N$). Using these values, the obstacles are created as seen in Figure 5.7.

**First Obstacle** For $a \in [(\frac{1}{2}+w-\gamma)n, n]$ and $b \in [-wn+a, (\gamma-w)n+a]$ we show that $\frac{a^-+b^+}{a^++b^-} \geq 1+x_1$.

We express the probabilities:

$$\begin{aligned} -a^+ + a^- + b^+ - b^- &\geq x_1(a^+ + b^-) \\ -(1-\mu)\varepsilon(a-b)(n-a-b) &\geq x_1((1-\mu)((1+\varepsilon)a(n-a)+b(n-b)) \\ &\quad + \mu((n-b)(n-a)+(1+\varepsilon)ba)). \end{aligned}$$

Since for $a, b$ in the given ranges, we have $a+b>n$ and $a>b$. We plug extremal values of $b$ and $a$ to the expression on the left side and using $4\gamma = w$ and $\mu < \frac{1}{2}$, we get:

$$\begin{aligned} (1-\mu)\varepsilon(a-b)(a+b-n) \geq& (1-\mu)\varepsilon(a-(a+(\gamma-w)n))(a+a-wn-n) \\ \geq& (1-\mu)\varepsilon(w-\gamma)n^2(1+2w-2\gamma-w-1) \\ \geq& (1-\mu)\varepsilon(w-\gamma)n^2(w-2\gamma) \\ \geq& (1-\mu)\varepsilon(4\gamma-\gamma)n^2(4\gamma-2\gamma) \\ \geq& 3\varepsilon\gamma^2n^2 \end{aligned}$$

On the right side, using $0 \leq a, b \leq n$, we have:

$$\begin{aligned} x_1((1-\mu)((1+\varepsilon)a(n-a)+b(n-b))+&\mu((n-b)(n-a)+(1+\varepsilon)ba)) \\ \leq& x_1((1-\mu)((1+\varepsilon)n^2+n^2)+\mu(n^2+(1+\varepsilon)n^2)) \\ \leq& x_1(2+\varepsilon)n^2 \end{aligned}$$

Using $\gamma = \frac{\min(1,\varepsilon)(1-2\mu)}{32}$, we evaluate:

$$\begin{aligned} 3\varepsilon\gamma^2n^2 &\geq x_1(2+\varepsilon)n^2 \\ \frac{3\varepsilon c^2}{2+\varepsilon} &\geq x_1 \\ \frac{3\varepsilon^3(1-2\mu)^2}{1024(2+\varepsilon)} &\geq x_1 \\ \frac{\varepsilon^3(1-2\mu)^2}{350(2+\varepsilon)} &\geq x_1 \end{aligned}$$

which proves the statement.

**Second obstacle** For $a \in [\frac{1}{2}N, (\frac{1}{2}+w)N]$ and $b \in [(1+w)N-a, (1+w+\gamma)N-a]$ we show that $\frac{a^++b^+}{a^-+b^-} \geq 1+x_2$. Since the statement is symmetric in $a$ and $b$, it holds also for obstacle defined by switching $a$ and $b$. The second obstacle is defined as union of $a \in [\frac{1}{2}n, (\frac{1}{2}+w)n]$ and $b \in [(1+w)n-a, (1+w+\gamma)n-a]$ with $b \in [\frac{1}{2}n, (\frac{1}{2}+w)n]$ and $a \in [(1+w)n-b, (1+w+\gamma)n-b]$.

We express the probabilities (for first part of the obstacle):

$$\begin{aligned}
&a^+ - a^- + b^+ - b^- \geq x_2(a^- + b^-)\\
&(n-a-b)(\varepsilon(1-\mu)(a+b) + 2\mu n) + 2(1-2\mu)\varepsilon ab\\
&\qquad \geq x_2(2\mu(1+\varepsilon)ab + (1-\mu)(n-a)a + (1-\mu)(n-b)b)
\end{aligned}$$

For the right side and $0 \leq a, b \leq n$, we have:

$$\begin{aligned}
x_2(2\mu(1+\varepsilon)ab + (1-\mu)(n-a)a + (1-\mu)(n-b)b) \leq& x_2(2\mu(1+\varepsilon)n^2 + (1-\mu)n^2 + (1-\mu)n^2)\\
\leq& x_2(1+\varepsilon)(2\mu n^2 + 2(1-\mu)n^2)\\
\leq& x_2 2(1+\varepsilon)n^2
\end{aligned}$$

Again $a + b > n$, thus we have

$$\begin{aligned}
&(n-a-b)(\varepsilon(1-\mu)(a+b) + 2\mu n) + 2(1-2\mu)\varepsilon ab \geq\\
&(-w-\gamma)n(\varepsilon(1-\mu)(1+w+\gamma)n + 2\mu n) + 2(1-2\mu)\varepsilon\frac{1}{2}n((1+w)n - a) \geq\\
&(-w-\gamma)n^2(\varepsilon(1-\mu)(1+w+\gamma) + 2\mu) + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq\\
&-5\gamma n^2(\varepsilon(1-\mu)(1+5\gamma) + 2\mu) + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq
\end{aligned}$$

By the choice of $\gamma$, we know that $\varepsilon(1-\mu)(1+5\gamma) < \varepsilon(1+\min(1,\varepsilon)) < 2\varepsilon$, moreover $\min(1,\varepsilon)\cdot(2\varepsilon+1) \leq 3\varepsilon$, this gives

$$\begin{aligned}
&-5\gamma n^2(\varepsilon(1-\mu)(1+5\gamma) + 2\mu) + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq\\
&-5\gamma n^2(2\varepsilon+1) + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq\\
&-5\frac{(1-2\mu)}{32}n^2\min(1,\varepsilon)(2\varepsilon+1) + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq\\
&-\frac{(1-2\mu)15\varepsilon}{32}n^2 + (1-2\mu)\varepsilon n^2\frac{1}{2} \geq\\
&\varepsilon(1-2\mu)n^2(\frac{1}{2} - \frac{15}{32}) \geq\\
&\frac{\varepsilon(1-2\mu)n^2}{32}
\end{aligned}$$

Comparing both sides gives

$$\begin{aligned}
\frac{\varepsilon(1-2\mu)n^2}{32} \geq& x_2 2(1+\varepsilon)n^2\\
\frac{\varepsilon(1-2\mu)}{64(1+\varepsilon)} \geq& x_2
\end{aligned}$$

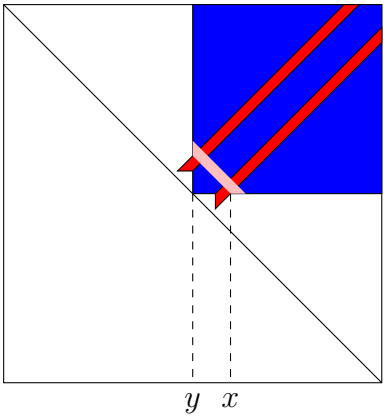



Figure 5.7: Obstacles: blue region highlights permissible values of $(a, b)$. The first obstacle is in red and the second is in pink (note that it is union of two obstacles).

**Third obstacle** For $b \in [(\frac{1}{2} + w - \gamma)N, N]$ and $a \in [-wN + b, (\gamma - w)N + b]$ we know that $\frac{a^- + b^+}{a^+ + b^-} \geq 1 + x_1$ since $a$ and $b$ are symmetric.

**Passing through obstacle** To get from the starting state $(N, N)$ to one of the absorbing states $(0, N)$ or $(N, 0)$, the process needs to pass through one entire obstacle.

Suppose that it is the first obstacle. The number of steps that increase $a$ or decrease $b$ needs to be larger by at least $cn$ than the number of steps that decrease $a$ or increase $b$. If the process is in the obstacle, we collapse the two dimensional Markov chain to one dimensional. The position is denoted by $a-b$. The obstacle starts at position $a-b = wn$ and if $a-b < (w-\gamma)n$ the obstacle is passed.

We know that inside the obstacle holds $\frac{a^- + b^+}{a^+ + b^-} \geq 1 + x_1$. This means the Markov chain has a constant bias, by result It takes $(1 + x_1)^{\gamma n}$ (which is $2^{\mathcal{O}(n)}$) steps on average to get through the obstacle.

The statement holds for the third obstacle symmetrically and for the second obstacle with time $((1 + x_2)^{\gamma n})$ by similar reasoning.

In the intersection of the first and the second obstacle, $\frac{1}{n}n \leq a \leq b \leq n$, which implies $a^+ \geq b^+$ $((1 + \varepsilon)(n - a)a \geq (1 + \varepsilon)(n - b)b)$ and trivially $a^+ \geq b^-$. This means passing through both obstacles takes also exponential time.

That means leaving the coexistence regime needs at least exponential time.

□

### 5.5.5 Cycles

In this section, we prove Theorem 23 from the main text. Among other implications, this illustrates that even if we fix the underlying graph and the effects of the patches on the fitness, the coexistence time critically depends on the relative layout of the two patches. In particular, we show that when the nodes of a long one-dimensional lattice alternately belong to patches $P_1$ and $P_2$ then the process terminates in polynomial time. In contrast, when each patch forms a contiguous block of $N$ nodes, each type holds majority in its patch for a number of steps that is exponential in $N$.

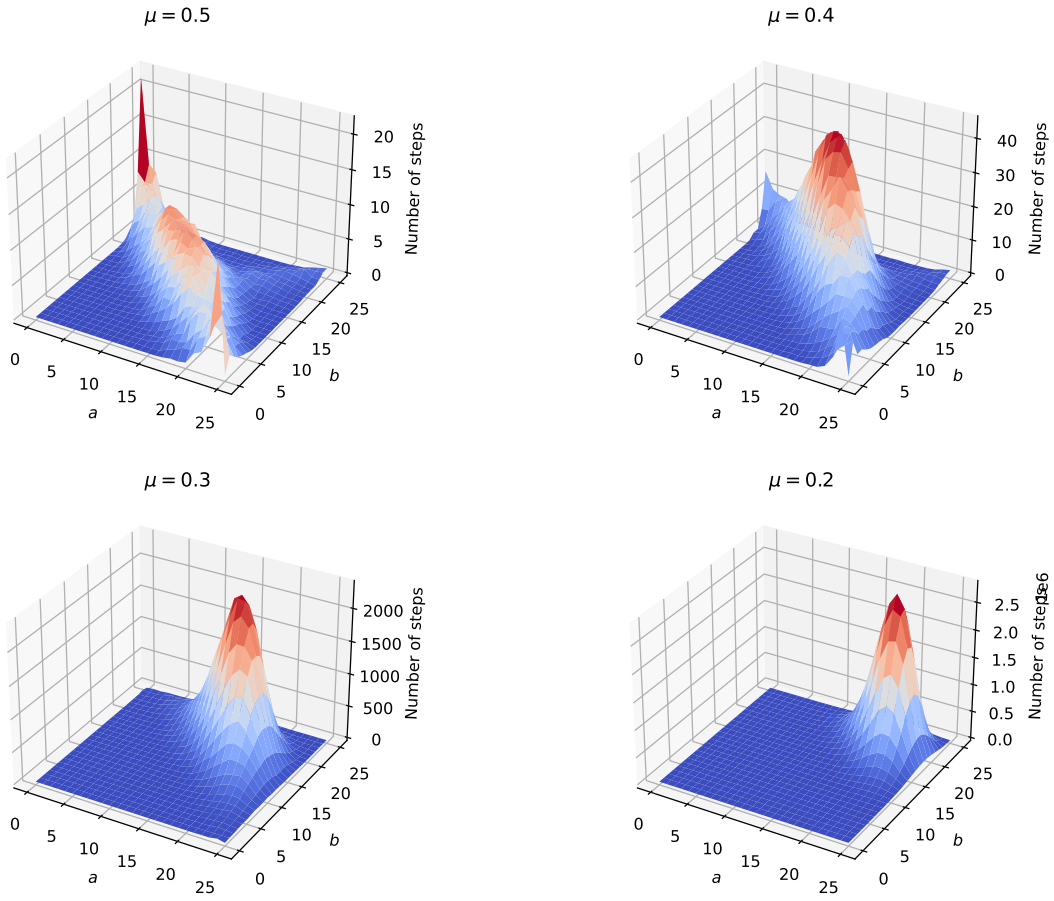



Figure 5.8: The simulation of the process for different $\mu$. We see that the lower $\mu$ is the more is spent near the beginning. Also the lower the $\mu$ is the number of steps is higher.

**Theorem 23** (One-dimensional lattices)**.** *Fix $\varepsilon > 0$. Then*

$$\mathrm{FT}(R^{\mathrm{alt}}(N,\varepsilon)) \in \mathcal{O}(N^3) \quad \textit{and} \quad \mathrm{CT}\,1/2(R^{\mathrm{split}}(N,\varepsilon)) \in 2^{\Omega(N)}.$$

*Proof.* Regarding the first claim, we build on the fact that the process can be mapped to an unbiased random walk, see Theorem 1 in [MP14].

Note that for any active edge, the individuals occupying its endpoints have the same fitness: If they are both home then they both have fitness $1+\varepsilon$, otherwise they both have fitness $1$.

Consider any time-point and any active edge. Since its endpoints both have degree 2, the edge is equally likely to be used for a reproduction event in either direction. Summing over all active edges we learn that, at each time-point, the probabilities $p^+$ (resp. $p^-$) that the number of T1 individuals increases (resp. decreases) are equal. Tracking only steps that happen along active edges, we thus obtain an unbiased random walk $(x_t)_{t\geq 0}$ which:

1. starts at $x_0 = N$,

2. when at $x_t$, it jumps to either $x_{t+1} = x_t + 1$ or $x_{t+1} = x_t - 1$ with probability $1/2$ each, and

3. it terminates once it reaches $x_t \in \{0, 2N\}$.

The absorption time of such a random walk is $\tau = x_0(2N - x_0) = N^2$.

Finally, note that at each time-point until the process terminates, there exist at least 2 active edges, hence the probability that the next reproduction event happens along an active edge is at least $p_a \geq 2/F \in \Omega(1/N)$, where $F \in [2N, (1+\varepsilon)2N] \in \Theta(N)$ is the total fitness of the population. Since other reproduction events do not change the configuration of the individuals, we obtain that the total fixation time is at most $\mathrm{FT}(R^{\mathrm{alt}}(N, \varepsilon)) \leq T/p_a \in \mathcal{O}(N^3)$ as desired.

Regarding the second claim, we again track only the active steps and we map the process to a random walk with a constant backward bias.

Consider the set $S'$ of 4 nodes that are the endpoints of the two edges that connect the two patches. Define $S = V \setminus S'$ to be the set of the remaining $2N - 4$ nodes. Given any fixed time-point and $i \in \{1, 2\}$, denote by $a_i$ the number of T$i$ individuals who are not home and in $S$. Initially, we have $a_1 = a_2 = 0$ and for one type to lose majority in its patch we must have $a_i \geq \frac{1}{2}N - 2$ for some $i \in \{1, 2\}$. Also, if $1 \leq a_i \leq \frac{1}{2}N - 2$ then, in a single step, $a_i$ is $(1+\varepsilon)$-times more likely to decrease then to increase: Indeed, for $a_i$ to increase, a not-home individual must replace a neighboring individual who is home. But any such edge is $(1+\varepsilon)$-times more likely to be used for reproduction in the opposite direction, since the individual who is home has fitness $1+\varepsilon$, whereas the other individual has fitness 1 (and both endpoints of the edge have the same degree 2). Thus the $\frac{1}{2}$-coexistence time is at least as large as the absorption time of a biased random walk $(x_t)_{t \geq 0}$ which:

1. starts at $x_0 = 0$,

2. when $x_t = 0$ then $x_{t+1} = 1$, otherwise it jumps to $x_{t+1} = x_t + 1$ with probability $p^+ = 1/(2+\varepsilon)$ and to $x_{t+1} = x_t - 1$ with probability $p^- = (1+\varepsilon)/(1+2\varepsilon)$, and

3. it terminates once it reaches $x_t = \frac{1}{2}N - 2$.

For fixed $\varepsilon > 0$ and $N \to \infty$, the absorption time of such a random walk is $\Theta((p^-/p^+)^{\frac{1}{2}N-2}) = 2^{\Theta(N)}$, thus $\mathrm{CT}\, 1/2(R^{\mathrm{split}}(N, \varepsilon)) \in 2^{\Omega(N)}$ as claimed. □

### 5.5.6 General upper bound

In this section, we prove Theorem 5 from the main text. That is, we show that within a certain broad class of population structures, a coexistence on a substantially longer than exponential timescale is impossible. Thus, the exponential coexistence that occurs for two-island structures with low migration rates is close to optimal. Namely, given a real number $w^{\min} > 0$ we denote by $\mathcal{G}(w^{\min})$ the class of all connected graphs in which each edge is assigned a weight at least $w^{\min}$.

**Theorem 5** (General upper bound)**.** *Fix $w^{\min} > 0$ and $\varepsilon > 0$. Then for any population structure $G_{N,\varepsilon} \in \mathcal{G}(w^{\min})$ on $2N$ nodes we have $\mathrm{FT}(G_{N,\varepsilon}) \in 2^{\mathcal{O}(N \cdot \log N)}$.*

*Proof.* Consider any time-point and an arbitrary ("focal") individual. We bound from below the probability $p$ that in the next $2N-1$ steps this focal individual produces a lineage that takes over the whole population. By summing up a geometric series, the fixation time is then at most $\mathrm{FT}(G_{N,\varepsilon}) \leq (2N-1)/p$.

Note that, at any time-point, the total fitness $F$ of the population satisfies $2N \leq F \leq 2N \cdot (1+\varepsilon)$, hence $F \in \Theta(N)$. Also, until fixation occurs, there exists an edge connecting a node $a$ occupied by an offspring of the focal individual and a node $b$ occupied by its non-offspring. With probability $p_a \geq 1/F \in \Omega(1/N)$ the individual at node $a$ is selected for reproduction and with probability $p_{a\to b} \geq w/((N-1)\cdot 1) \in \Omega(1/N)$ the produced offspring migrates to node $b$. Thus we get $p \geq (p_a \cdot p_{a\to b})^{2N-1} \in \Omega(1/N^{4N-2})$ and then the desired

$$\mathrm{FT}(G_{N,\varepsilon}) \leq (2N-1)/p \in \mathcal{O}\left((2N-1)\cdot N^{4N-2}\right) = 2^{\mathcal{O}(N\log N)}. \qquad \square$$

Note that Theorem 5 gives an upper bound on the fixation time, and thus also on the $c$-coexistence time for any $c > 0$. Also, the same proof clearly applies to populations consisting of multiple patches and with multiple competing types, so long as the fitness of each type in each patch is bounded between two constants $1+\varepsilon_{\min}$ and $1+\varepsilon_{\max}$.

CHAPTER 6

# Social Balance on Networks: Local Minima and Best Edge Dynamics

This chapter appears in full in [CSŽ+22]. Reprinted from Krishnendu Chatterjee, Jakub Svoboda, Djordje Žikelić, Andreas Pavlogiannis, and Josef Tkadlec. Social balance on networks: Local minima and best-edge dynamics. *Physical Review E*, 106(3):034321, 2022 with the permission of American Physical Society. 

## Abstract

Structural balance theory is an established framework for studying social relationships of friendship and enmity. These relationships are modeled by a signed network whose energy potential measures the level of imbalance, while stochastic dynamics drives the network towards a state of minimum energy that captures social balance. It is known that this energy landscape has local minima that can trap socially-aware dynamics, preventing it from reaching balance. Here we first study the robustness and attractor properties of these local minima. We show that a stochastic process can reach them from an abundance of initial states, and that some local minima cannot be escaped by mild perturbations of the network. Motivated by these anomalies, we introduce Best Edge Dynamics (BED), a new plausible stochastic process. We prove that BED always reaches balance, and that it does so fast in various interesting settings.

## 6.1 Introduction

The formation of social relationships is a complex process that has long fascinated researchers. It is well-understood that, besides pairwise interactions, friendships and rivalries are affected by social context. The study of such phenomena dates back to Heider's theory of *social balance* [Hei44, Hei46, Hei58], which can be seen as a rigorous realization of the proverb "the enemy of my enemy is my friend". The theory classifies a social state as *balanced* whenever every group of three entities (a *triad*) is balanced: it consists of either three mutual friendships, or one friendship whose both parties have a mutual enemy. The other types of triads create social unrest that eventually gets resolved by changing the relationship between two parties. For example, a triad with three mutual enmities will eventually lead to two entities forming an alliance against the common enemy. A triad with exactly one enmity will either see

a reconciliation to a friendship under the uniting influence of the common friend, or lead to the break of one friendship following the social axiom "the friend of my enemy is my enemy" [Sch10].

Cartwright and Harary developed a graph-theoretic model of Heider's theory [CH56], and showed that any balanced state is either a utopia without any enmities, or it consist of two mutually antagonistic groups [H$^+$53, Dav67]. This structural theory of social balance has seen applications across various fields ranging from philosophy, sociology, or political science [BGS$^+$17, Tay70, FA21, She71, Moo79] all the way to fields such as neuroscience or computer science [CCC$^+$20, MPKHT20, LCWZ13, Alt12], It has also been supported by empirical evidence [RF17, KCN19, ASBF20], see also [ZZW15] for a review. The setting is attractive to physicists due to its intimate connection to the Ising model and spin glasses [FIA11], and indeed tools and techniques from statistical physics have proved to be instrumental in improving our understanding of such systems [BHR$^+$17, SATJK17, MM21], see also [CFL09] for a review.

It is natural to associate each network state with a *potential energy* that counts the difference of imbalanced minus balanced triads; hence the perfectly balanced states are those that minimize the energy of the network [MSK09]. Understanding how energy is minimized in a system is a fundamental problem studied across different physics fields, and signed graphs present a clean theoretical framework to study this problem in a setting with a population structure. It is well known that the energy landscape over signed graphs has local minima (also known as *jammed states*) [AKR05], that is, states from which all paths to social balance must temporarily increase the number of imbalanced triads.

When the network state is imbalanced, we expect that a social process will perturb it until balance is reached. The seminal work [AKR06] introduced a stochastic process known as *Local Triad Dynamics (LTD)*, according to which imbalanced triads are sampled at random, and the sampled triad is balanced by flipping the relationship of two of its entities. This step is called an *edge flip*. The same work also introduced *Constrained Triad Dynamics (CTD)*, a socially-aware variant of LTD under which an edge flip is only possible if it reduces the number of imbalanced triads. Unfortunately, the existence of local minima in the energy landscape implies that CTD can get stuck in jammed states and thus remain permanently imbalanced.

Although the existence of jammed states is well understood in terms of the energy landscape, little is known about them from the perspective of the stochastic process, that is, about their reachability properties. For example, from which initial states is it possible to reach a jammed state? Moreover, if a jammed state is reached, can the process escape if we slightly perturb the network? Finally, is there a plausible, socially-aware stochastic dynamics (like CTD) that always reaches balance (unlike CTD)? We tackle these questions in this work.

First, we study the robustness and attractor properties of the local minima of the energy landscape. We show that the number of jammed states is super-exponential, compared to the previously known exponential lower-bound, and that jammed states are reachable from any initial state that is not too friendship-dense. Moreover, we show that some of those jammed states are strongly attracting: even when perturbing a constant portion of edges adjacent to each vertex, the same jammed state is subsequently reached with probability $1$. As a byproduct, our results resolve an open problem from [MSK09].

Second, we propose a new plausible dynamics called *Best Edge Dynamics (BED)*. Like CTD, BED is a stochastic process in which edge flips are socially-aware, in the sense that they maximize the number of newly balanced triads (see below for details). We prove that, unlike

CTD, BED always reaches a balanced state from any initial state. Moreover, we show that BED converges faster to a balanced state than CTD in various interesting settings, such as when started from a state that is already close to being balanced.

Finally, we complement our analytical results with computer simulations in the cases when the initial friendship edges form a random Erdős-Rényi network or a random scale-free networks.

## 6.2 Triad Dynamics in Social Networks

Balance on social networks is studied in terms of signed graphs. A *signed graph* $G = (V, E, s)$ consists of a finite complete graph $(V, E)$ on $|V| = n$ vertices together with an edge labeling

$$s\colon E \to \{-1, +1\}.$$

The labeling $s$ assigns to each edge one of the two signs; the edges labeled by $+1$ are *friendships* and those labeled $-1$ are *enmities*. Thus each pair of individuals (modeled by vertices) has a defined relationship: either they are friends, or they are enemies.

Given a signed graph $G = (V, E, s)$, a *triad* is a subgraph of $G$ defined by any three of its vertices. A triad is of *type* $\Delta_k$ for $k = 0, 1, 2, 3$ if it contains exactly $k$ edges labeled $-1$. A triad is *balanced* if its type is $\Delta_0$ or $\Delta_2$. Intuitively, a triad is balanced if it satisfies the known proverb "the enemy of my enemy is my friend". For an edge $e$ in $G$, its *rank* $r_e$ is the number of imbalanced triads containing $e$. Finally, a signed graph $G$ is *balanced* if each triad in $G$ is balanced, see Fig. 6.1.

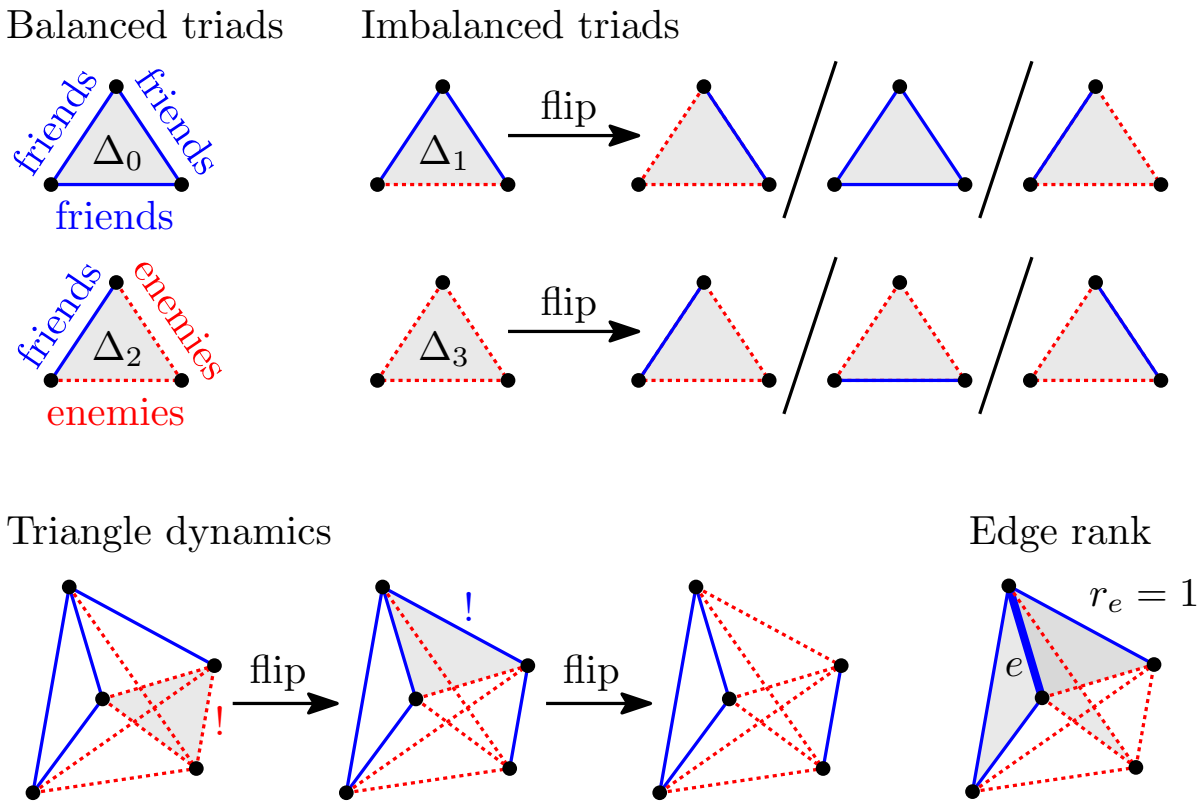



Figure 6.1: A triad of type $\Delta_k$ contains $k$ enmity edges. The imbalanced triads $\Delta_1$ and $\Delta_3$ can be made balanced by flipping any one edge. A sequence of flips typically reaches a state where all triads are balanced. The rank $r_e$ of edge $e$ is the number of imbalanced triads containing $e$.

It is known that a signed graph is balanced if and only if its enmity edges form a complete bipartite graph over the vertex set [CH56]. This means that we may partition the vertices of a balanced signed graph into two vertex classes, such that all pairs of vertices from the same class form friendship edges, and all pairs of vertices from different classes form enmity edges. Moreover, every signed graph which admits such a partitioning is clearly balanced. In the special case where one of the vertex classes in this partitioning is empty, each pair of vertices forms a friendship edge and we refer to this balanced signed graph as *utopia*.

The main interest in the study of social networks modeled by signed graphs is the evolution of the network according to some pre-specified dynamics, and the time until the balance is reached. The goal is to understand which simple dynamics ensure fast convergence to balance. Following the work of [AKR06], we focus on those dynamics that, at each time step, select one edge $e$ according to some rule and then flip its sign. We then say that "$e$ is flipped". In [AKR06], two such dynamics on signed graphs were introduced: *Local Triad Dynamics (LTD)*, and *Constrained Triad Dynamics (CTD)*. In the rest of this section, we define these two dynamics and discuss their advantages and limitations.

### 6.2.1 Local Triad Dynamics

Let $G$ be a signed graph modeling a social network with friendships and enmities.

The *Local Triad Dynamics (LTD)* with parameter $p \in [0, 1]$ is a discrete-time random process that starts in $G$ and repeats the following procedure until there are no imbalanced triads in $G$:

1. Select an imbalanced triad $\Delta$ uniformly at random.
2. If $\Delta$ is of type $\Delta_3$, then an edge of $T$ is chosen to be flipped uniformly at random. If $\Delta$ is of type $\Delta_1$, then the unique edge with sign $-1$ is chosen to be flipped with probability $p$ and each of the two other edges is chosen with probability $\frac{1-p}{2}$.

We refer to distinct signed graphs as *states*.

LTD is socially oblivious in the sense that once an imbalanced triad is selected, the edge to be flipped is chosen according to a (stochastic) rule that disregards the rest of the network. Moreover, the guarantees of LTD on the expected time to reach a balanced state are not very plausible: it was shown in [AKR06] that if $p < \frac{1}{2}$, then the expected time grows exponentially with the size of the signed graph. On the other hand, if $p > \frac{1}{2}$, then the dynamics is more likely to create rather than remove friendships and it reaches utopia with high probability. This means that the eventual balanced state is essentially pre-determined.

### 6.2.2 Constrained Triad Dynamics

The *Constrained Triad Dynamics (CTD)* is another dynamics on signed graphs. Given a signed graph $G$ on $n$ vertices, CTD is a random process that starts in $G$ and repeats the following procedure until there are no imbalanced triads in $G$:

1. Select an imbalanced triad $\Delta$ uniformly at random.
2. Select an edge $e$ of $\Delta$ uniformly at random.

3. Flip $e$ if $r_e \geq \frac{1}{2}n - 1$, that is, if the number of imbalanced triads in $G$ does not increase upon the flip (in the case of equality, the flip happens with probability $1/2$), otherwise do nothing.

Note that CTD introduces a non-local, socially-aware rule: When deciding whether a selected edge should be flipped, we take into account all triads that contain it. In [AKR06] it was claimed that, starting from any initial signed graph $G$, CTD converges to a balanced state and that balance is reached fast – in a time that scales logarithmically with the size $n$ of the signed graph. If true, this would imply that CTD overcomes the limitation of LTD in which the expected convergence time could be exponential in $n$. However the claim, which was supported by an informal argument, is not quite true, since the energy landscape is rugged: There are states, called *jammed states*, that are not balanced but where CTD can not make a move, since any flip would (temporarily) increase the number of imbalanced triads [AKR05]. Moreover, it is known that there are at least roughly $3^n$ jammed states (compared to roughly $2^n$ balanced states) and that some of the jammed states have zero energy [MSK09].

## 6.3 Reaching and Escaping the Jammed States

Even though there are exponentially many jammed states, computer simulations on small populations were used to suggest that they can effectively be ignored [AKR05]. In contrast, in this section we present three results which indicate that for large population sizes the jammed states are important.

First ("counting"), we study the number of jammed states. It is known [AKR05], that there are at least $3^n$ jammed states, that is, at least exponentially many. Here we construct a family of simple, previously unreported jammed states and we show that the total number of jammed states on $n$ labeled vertices is super-exponential, namely at least $2^{\Omega(n \log n)}$. This shows that even on the logarithmic scale, the jammed states are substantially more numerous than the balanced states (of which there are "only" $2^{n-1}$).

Our second result ("reaching") shows that, starting from any initial signed graph that is not too friendship-dense, a specific jammed state $J$, which we construct below, is reached with positive probability. In particular this implies that the expected time to balance in this stochastic process is formally infinite, even for signed graphs that have a constant positive density of friendship edges.

Our third result ("escaping") shows that this specific jammed state $J$ forms a deep well in the energy landscape: Once it is reached, it can not be escaped even if we perturb a constant portion of edges incident to each vertex.

In the rest of this section, we sketch the intuition behind these results. For the formal statements and proofs, see Theorems 7 to 9 in Section 6.7.1, respectively.

Regarding the first result ("counting"), the new jammed states are defined in terms of an integer parameter $d$. We partition the population into $4d + 2$ clusters ($4d + 1$ or $4d + 3$ would work too), arrange the clusters along a circle, and assign a sign $+1$ to those (and only those) edges that connect individuals who live in clusters that are at most $d$ steps apart, see Fig. 6.2. We then show that, for each friendship or enmity edge $(u, v)$, a strict majority $2(d + 1) > \frac{1}{2}(4d + 3)$ of clusters have the property that any vertex $w$ in that cluster forms a balanced triad $(u, v, w)$ with $(u, v)$. Thus, flipping $(u, v)$ would increase the number of imbalanced triads. When all the clusters are roughly equal in size, which is the typical

behavior for large population sizes, the state is thus jammed. Our construction also resolves in affirmative an open question [MSK09] which asks whether there exist jammed states with an even number of friendship cliques (here this number is $4d + 2$).

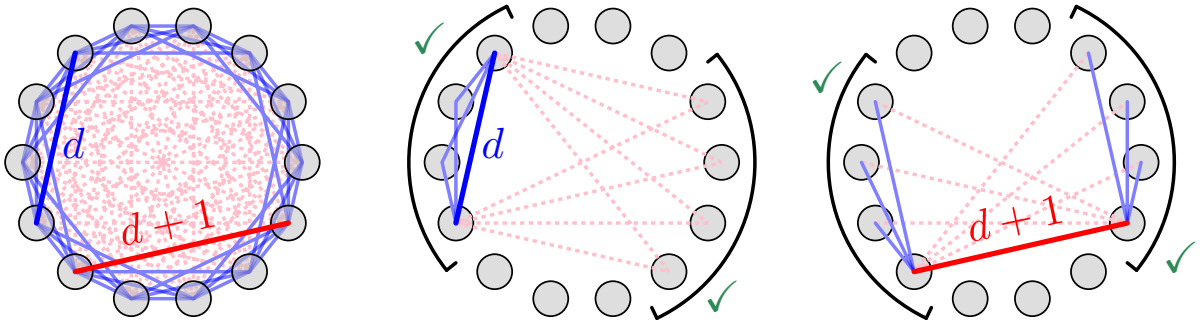



Figure 6.2: A jammed state consisting of $4d + 2$ roughly equal clusters, each connected by friendships to the clusters at most $d$ steps apart.

Regarding the second result ("reaching"), we consider any initial state $I_n$ on $n$ vertices in which each vertex is incident to at most $n/12 - 1$ edges labeled $+1$. We define a jammed state $J_n$ as follows: We partition the vertices into three clusters $V_1$, $V_2$, $V_3$ of roughly equal size, label all edges within each set $+1$ and all other edges $-1$. Then we exhibit a sequence of flips that transforms $I_n$ into $J_n$. This is done in two phases: First, one can verify that any time we select an imbalanced triangle that contains an edge labeled $-1$ within one cluster $V_i$, this edge can be flipped. Hence, we may flip all enmity edges within the three clusters to reach a state in which all edges within each $V_i$ are labeled $+1$. After that, one can similarly verify that all edges labeled $+1$ that connect vertices in two different parts $V_i$, $V_j$ can be flipped one by one, thereby reaching the jammed state $J_n$, see Fig. 6.3.

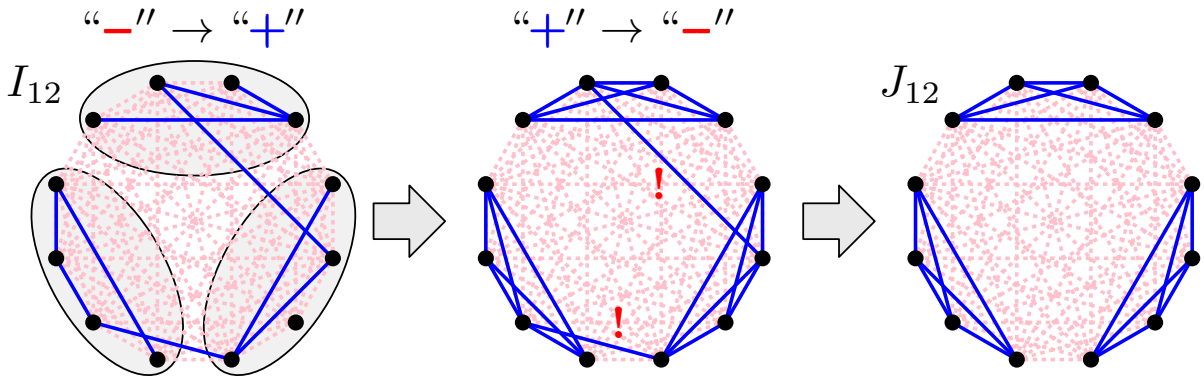



Figure 6.3: A jammed state $J_n$ (right) on $n$ vertices can be reached from any not too friendship-dense initial state $I_n$ (left). Moreover, once it is reached, it can not be escaped, even if a substantial portion of edges around each vertex are perturbed.

Regarding the third result ("escaping"), we consider any state $S_n$ that can be obtained from $J_n$ by flipping a set $E_0$ of edges such that each vertex is incident to at most $n/12 - 1$ edges of $E_0$. We then show that edges that do *not* belong to $E_0$ can never be flipped. On the other hand, each edge that does belong to $E_0$ can be flipped. Thus all edges in $E_0$ are eventually flipped back and the jammed state $J_n$ is reached again.

## 6.4 Best Edge Dynamics

Our results in the previous section show that the jammed states are a profound feature of the energy landscape: They are reachable from many conceivable initial states, and some of them trap stochastic dynamics such as CTD forever even if we allow substantial perturbations.

This brings up a question of whether there exists a simple socially-aware dynamics with all the desirable properties of CTD, but one that can not get stuck in a jammed state. To address this question, we propose the *Best Edge Dynamics (BED)*, a modification of CTD that, unlike CTD, reaches a balanced state with probability $1$ from every initial state. Moreover, we prove that BED converges fast to a balanced state in several important cases (see Propositions 2 and 3), and our empirical evaluation of both BED and CTD in Section 6.5 shows that in general the convergence times are comparable (after we exclude the runs where CTD does not terminate).

Let $G$ be a signed graph. Then the *Best Edge Dynamics (BED)* is a discrete-time random process that starts in $G$ and repeats the following procedure until there are no imbalanced triads in $G$:

1. Select an imbalanced triad $\Delta$ uniformly at random.
2. Select an edge $e$ from $\Delta$ with the highest rank $r_e$ (in case of ties, pick one such edge uniformly at random).
3. Flip $e$.

Note that, in contrast to CTD, we flip $e$ even when its rank $r_e$ satisfies $r_e < \frac{1}{2}|V| - 1$, that is, when flipping the best edge creates more imbalanced triads than it removes. In particular, whenever we reach a jammed state, we still make a flip. In principle, it could still happen that BED remains trapped in a subset of imbalanced states, toggling edges back and forth unable to escape it, but in fact we prove that this event occurs with probability 0.

**Theorem 6.** *For any initial signed graph on $n$ vertices, BED reaches a balanced state with probability $1$ and in finite expected time.*

To prove Theorem 6, we will show that any signed graph on $n$ vertices can become balanced upon $\mathcal{O}(n^3)$ flips. This suffices since BED induces a finite Markov chain over the set of all states, and the absorbing states of the Markov chain are precisely the balanced signed graphs.

Fix an edge $(v_1, v_2)$ in $G$ of the lowest rank, and set $B = \{v_1, v_2\}$. Note that it is possible to flip one edge in each imbalanced triad containing $(v_1, v_2)$ without flipping $(v_1, v_2)$ itself, to make all triads containing $(v_1, v_2)$ balanced: Indeed, if we consider an imbalanced triad containing $v_1$, $v_2$ and some third vertex $w$, as BED flips an edge of the highest rank, it can flip either $(v_1, w)$ or $(v_2, w)$. This makes the triad balanced and decreases $r_{(v_1,v_2)}$ by $1$, while decreasing the rank of any edge that hasn't been flipped by at most $1$. Hence, $(v_1, v_2)$ will still be of the lowest rank in all imbalanced triads containing it, so we can flip one edge in each such triad until all triads containing $(v_1, v_2)$ become balanced.

The rest of the construction proceeds inductively by adding a new vertex to $B$ in each step and making all triads which contain an edge with endpoints in $B$ balanced. The process ends when all vertices of $G$ have been added to $B$. For the inductive step, suppose that every triad containing at least two vertices in $B$ is balanced. Let $(v, w)$ be an edge in $G$ which is

of lowest rank among all edges with $v \in B$ and $w \notin B$. Each triad containing $w$ and two vertices in $B$ is balanced. On the other hand, since $(v, w)$ is an edge of lowest rank among all edges with $v \in B$ and $w \notin B$, it follows that for each imbalanced triad $\Delta$ containing $v$, $w$ and a third vertex $u$, BED can flip either $(w, u)$ or each edge $(v', u)$ with $v' \in B$ to make $\Delta$ and all other triads containing $u$ and two vertices in $B$ balanced. By doing this for each imbalanced triad containing $(v, w)$, we modify the signs in the graph in such a way that all triads containing at least two vertices in $B \cup \{w\}$ become balanced. Thus we can add $w$ to $B$. By induction on the size of $B$, this way we eventually reach a balanced state.

Notice that in each iteration of the above construction, at most $n \cdot |B|$ edges are flipped. Hence the total number of edge flips is at most $n \sum_{i=1}^{n} i = \mathcal{O}(n^3)$ as claimed.

### 6.4.1 Fast convergence and red-black graphs

So far we have shown that, unlike CTD, BED ensures convergence to balance with probability $1$ and in finite expected time. In the rest of this section we show that BED also provides theoretical guarantees on fast convergence when started in certain states that are either "close" to being balanced (Proposition 2) or jammed (Proposition 3), showing that this new dynamics is robust.

We start by introducing the *red-black graphs*, a new concept that allows neat reasoning about signed graphs that are close to being balanced. Given a signed graph $G$, let $C$ be a balanced signed graph on the same number of vertices which differs from $G$ in the smallest number of edge signs. We refer to $C$ as a *closest* balanced state to $G$. Then the red-black graph $R$ associated to $G$ and $C$ is obtained from $G$ by coloring each edge of $G$ in black if the signs of the edge in $G$ and $C$ agree, and in red otherwise. Thus, red edges are precisely those edges whose signs in $G$ and $C$ are misaligned. Figure 6.4 shows an example of a signed graph and the corresponding red-black graph.

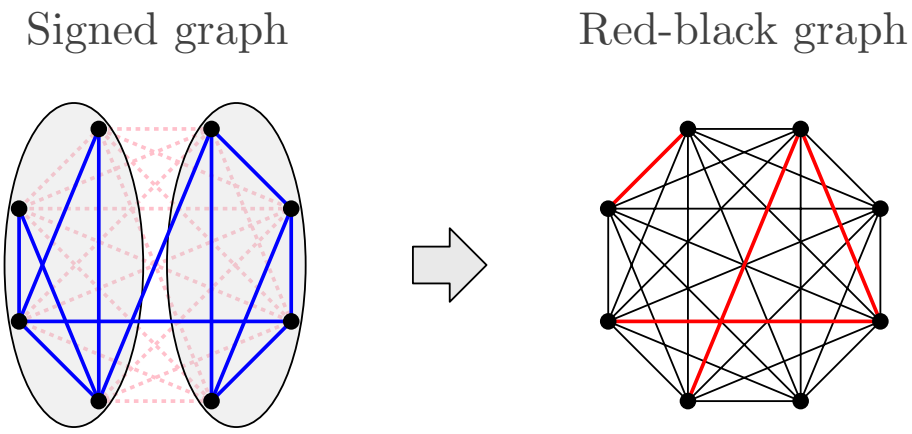



Figure 6.4: An example of a signed graph (left) and the corresponding red-black graph (right).

The key property of red-black graphs is that red and black edges can be viewed as enmities and friendships when reasoning about balanced triads in the following sense: A triad in $G$ is imbalanced if and only if exactly $1$ or $3$ of its edges are red in $R$. The proof of this claim is by casework and is deferred to Section 6.7.2. This also implies that the rank of an edge in $G$ is equal to its rank in $R$ if we treat red edges in $R$ as enmities and black edges as friendships.

### 6.4.2 Fast convergence around balanced states

We are now ready to study the convergence of BED when started in a signed graph which is close to being balanced.

**Proposition 2.** *Consider a signed graph $G$ whose red-black graph $R$ satisfies one of the following two conditions:*

1. *Each vertex is incident to at most $\frac{1}{4}n - 1$ red edges.*
2. *There are at most $\frac{1}{2}n$ vertices incident to a red edge.*

*Then BED reaches a balanced state in $\mathcal{O}(n^2)$ steps in the worst case.*

To prove the proposition, it suffices to show that BED flips only red edges: Since in total there are $\mathcal{O}(n^2)$ red edges, BED reaches a balanced state in $\mathcal{O}(n^2)$ steps.

Consider the imbalanced triad $(u, v, w)$ selected by BED. If the triad contains $3$ red edges, clearly BED flips a red edge. Otherwise, from the key property of red-black graphs (Lemma 24) we know that the triad contains exactly $1$ red edge. Without loss of generality suppose $e = (u, v)$ is the red edge. For each of the two conditions in Proposition 2 we argue separately:

1. Since $(v, w)$ is a black edge, any triad containing $(v, w)$ is imbalanced if it contains precisely $1$ red edge. There are at most $\frac{1}{2}n - 2$ such triads, since there are at most $\frac{1}{4}n - 1$ red edges incident to $v$ and similarly at most $\frac{1}{4}n - 1$ to $w$. Thus, $r_{(v,w)} \leq \frac{1}{2}n - 2$. Analogously $r_{(u,w)} \leq \frac{1}{2}n - 2$. On the other hand, as $(u, v)$ is red, a triad containing $(u, v)$ which is balanced has to contain exactly two red edges. So $(u, v)$ is contained in at most $\frac{1}{2}n - 2$ balanced triads by the same argument as above, and $r_{(u,v)} \geq n - 2 - (\frac{1}{2}n - 2) = \frac{1}{2}n$. Thus
$$r_{(u,v)} > \max\{r_{(u,w)}, r_{(v,w)}\},$$
so BED will flip the red edge.

2. By assumption, we can partition the vertices of $G$ into two sets $V_1$ and $V_2$ such that $|V_1| \geq |V_2|$ and all red edges have both endpoints in $V_2$. Then any triad $(u, v, w')$ with $w' \in V_1$ contains exactly $1$ red edge and is imbalanced, so $r_e \geq |V_1|$. On the other hand, as $(u, w)$ is black, for any third vertex contained in $V_1$ the triad is balanced, so $r_{(u,w)} \leq |V| - 2 - (|V_1| - 1) = |V_2| - 1 < |V_1|$. Analogously $r_{(v,w)} < |V_1|$, so $e$ has the highest rank in $(u, v, w)$ and BED flips $e$.

### 6.4.3 Fast convergence from jammed states

Recall that a state is jammed if it is not balanced but CTD cannot flip an edge in any imbalanced triad. Here we show that from certain jammed states, BED converges to a balanced state after $\mathcal{O}(n^2)$ edge flips, in expectation. Thus, BED ensures fast convergence even when the convergence time of CTD is infinite. (For details, see Section 6.7.3 and Proposition 3.)

As before, consider the jammed state $J_n$ consisting of three large roughly equal clusters of friends on $n$ vertices in total. Fig. 6.5 illustrates that, started from $J_n$, BED converges to balance in $\mathcal{O}(n^2)$ time. Initially, BED keeps adding friendship edges connecting different

clusters. Due to random fluctuations, the symmetry among the three clusters breaks and one pair of clusters becomes more densely connected than the other pairs. This difference is exaggerated over time and eventually that pair of clusters merges.

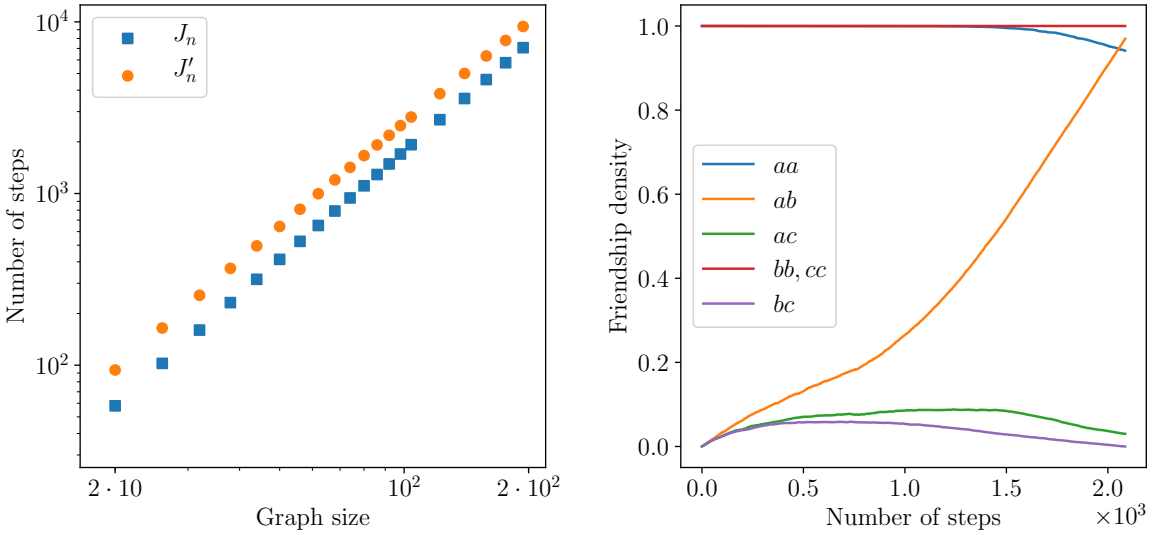



Figure 6.5: Left: Started from $J_n$, BED reaches balance in $\mathcal{O}(n^2)$ expected steps. Right: Friendship densities in different portions of the signed graph $J_{100}$, in a single run of BED. Apart from possibly the very end, the friendships within clusters ($aa$, $bb$, $cc$) are never flipped. Eventually, one pair of clusters (here $a$ and $b$) merges.

Next, we consider a state $J'_n$ whose $n$ vertices are split into 6 clusters arranged along a circle with relative sizes roughly $2:1:1:2:1:1$. Two vertices are connected by a friendship edge if they belong to the same cluster or to adjacent clusters (see Fig. 6.6, left). The different cluster sizes ensure that the symmetry is broken from the very beginning and allow for a simpler formal argument.

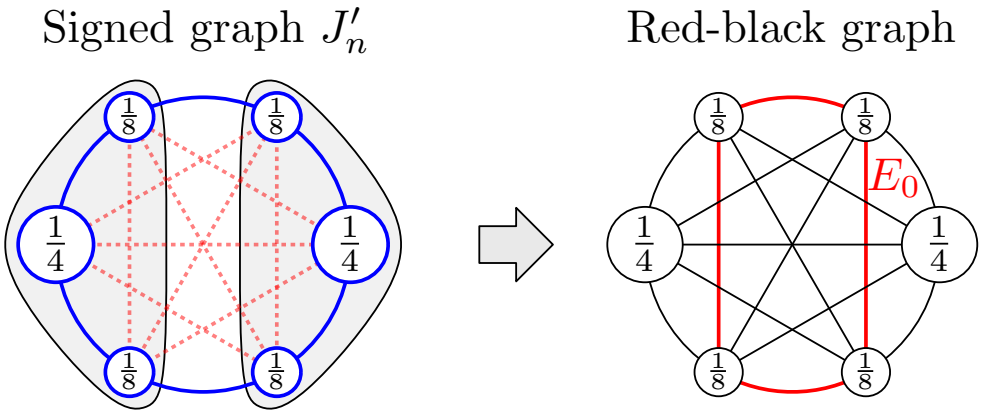



Figure 6.6: A jammed state $J'_n$ from which BED reaches a balanced state in $\mathcal{O}(n^2)$ expected steps.

It is straightforward to check that the state $J'_n$ is jammed and that the closest balanced state $C$ is the one depicted in the left figure. The corresponding red-black graph is shown on the right. Let $E_0$ be the set of edges that are initially red in the red-black graph. We show that the process always flips an edge $e \in E_0$ and that, at each point in time, we are a constant-factor

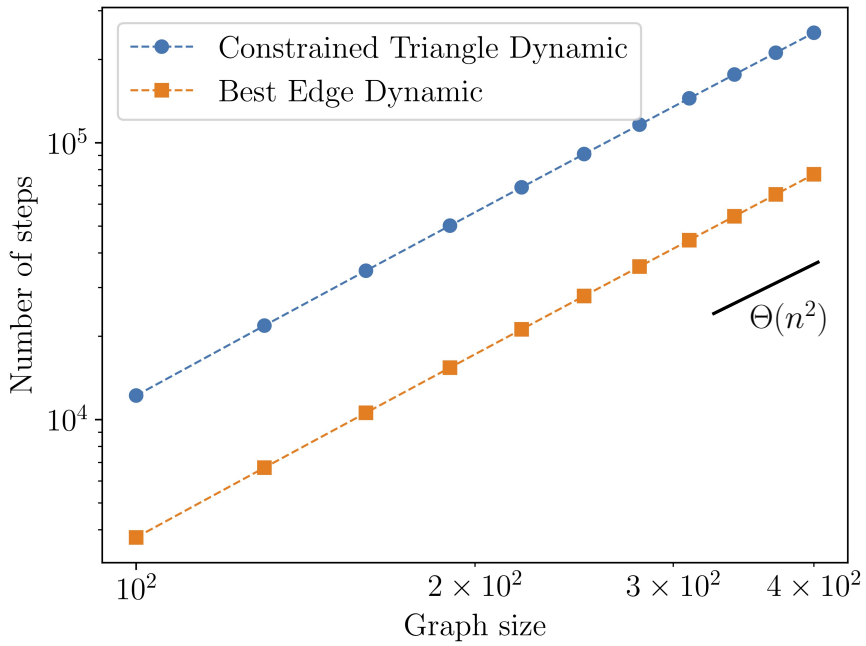



Figure 6.7: Average number of steps until balance for CTD (excluding the runs that get jammed) and BED, over $10^5$ runs. The friendships in the initial signed graphs form the Erdős-Rényi graph with edge probability $p = \frac{1}{2}$ and size $n \leq 400$. Both quantities scale as $\Theta(n^2)$.

more likely to turn a red edge into a black one rather than the other way around. Thus the stochastic process can be projected onto a random walk with a constant forward bias. Since such a random walk terminates in the number of steps that is linear in its length, this proves that the process finishes in $\mathcal{O}(|E_0|) = \mathcal{O}(n^2)$ steps in expectation.

## 6.5 Computer Simulations

In this section, we compare the two dynamics CTD and BED by means of computer simulations. In each simulation, we generate a network (possibly randomly) and assign "+" to each its edge. All other edges are assigned "−", so the underlying network is always a complete graph. Then we simulate each dynamics to determine the quantities such as the typical outcome and the number of steps until it reaches balance.

### 6.5.1 Erdős-Rényi graphs

First we consider random Erdős-Rényi graphs $\mathrm{ER}(n, p)$, where each two of the $n$ vertices are connected by a friendship edge with probability $p$, independently of each other.

Since CTD can get jammed, the average number of steps until it reaches balance is infinite for many initial states. To have a meaningful comparison, we first exclude runs in which CTD gets jammed (later we report their proportion). Upon this exclusion, the two dynamics are comparable. First, the number of steps until balance for both dynamics scales as $\Theta(n^2)$ with the population size $n$, see Fig. 6.7.

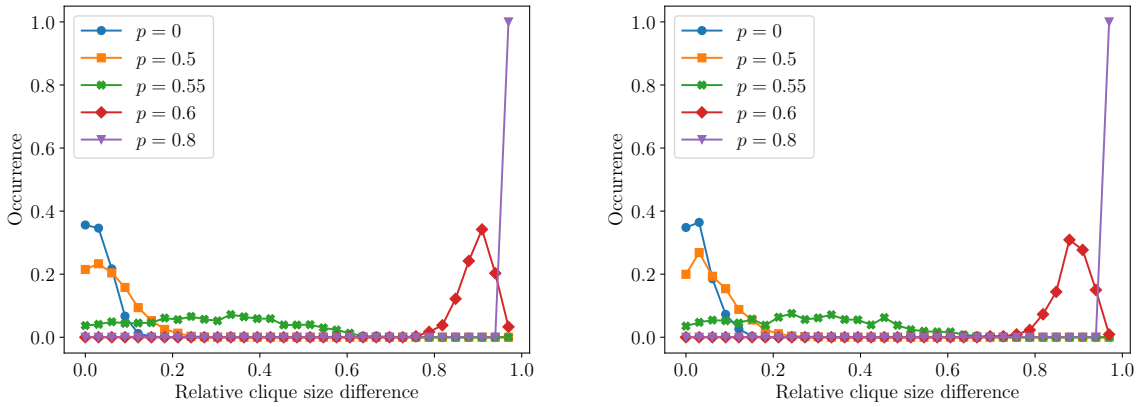



Figure 6.8: Distribution of the relative difference of the clique sizes once balance is reached, for BED (left) and CTD (right) over $10^5$ runs. Here $n = 128$.

Second, the final configuration does not depend on the choice of the dynamics but it is strongly dependent on the parameter $p$. Namely, for $p \leq 0.5$ the two cliques are almost always roughly equal in size, whereas for $p \geq 0.6$ the larger clique contains almost all the vertices, see Fig. 6.8. (See also Section 6.7.4 for tables showing several network descriptors before and after the network becomes balanced.)

Next we focus on the probability that CTD gets jammed. For fixed $n$, this probability exhibits a threshold behavior as a function of the friendship density $p$, see Fig. 6.9. The intuition is as follows: When the initial friendship density is large (here $p \geq 0.6$), then the initial state is close to utopia (the balanced state that consists of only friendships). Utopia is then reached quickly and with high probability (cf. Section 6.4.2). When the initial friendship density is small (here $p \leq 0.5$), the jamming probability is nonzero (cf. Section 6.3). Most imbalanced triads are of type $\Delta_3$ (all enmities). The dynamics thus keeps adding friendship edges and the jamming probability is mostly independent of $p$. The same phenomenon occurs for other sizes $n$, see Fig. 6.10.

### 6.5.2 Scale-free networks

Apart from Erdős-Rényi graphs we also consider the Barabási-Albert model $\mathrm{BA}(n, d)$ for scale-free networks [BA99]. This model creates a scale-free graph with edge density $d \in [0, 1]$. In particular, we start with a path on $11$ vertices and then process the remaining $n - 11$ vertices one by one. When processing a vertex $v$, we randomly connect it to a subset of vertices already present in the network in such a way that the probability that a pair $uv$ forms a (friendship) edge is proportional to the degree of $u$ ("preferential attachment") and the resulting (expected) edge density equals $d$.

Compared to Erdős-Rényi graphs, the degrees of the vertices are unequally distributed. Despite this difference, the number of steps until balance for both CTD and BED still scales as $\Theta(n^2)$ with the size $n$, see Fig. 6.11. Moreover, the jamming probability for CTD still exhibits a similar threshold behavior, see Figs. 6.12 and 6.13.

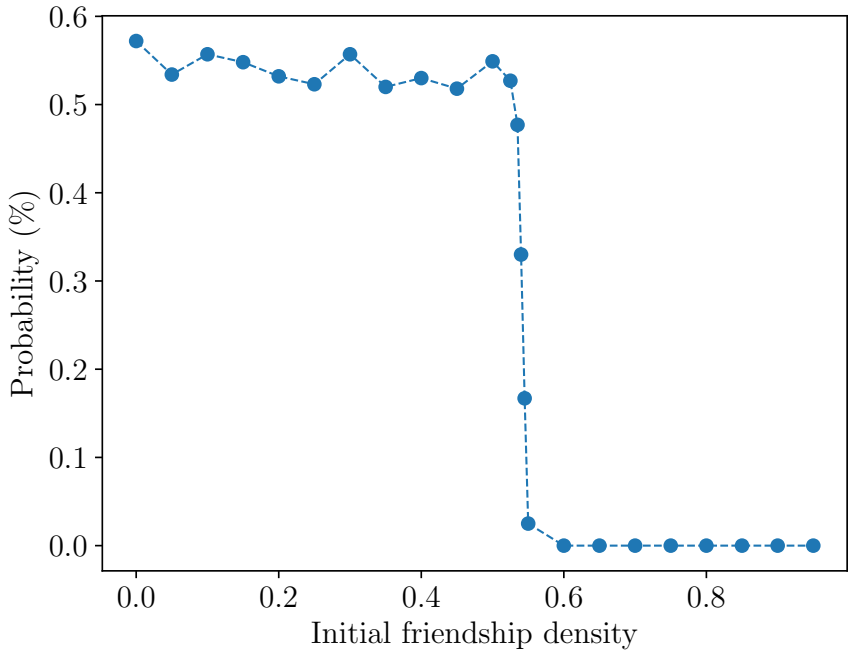



Figure 6.9: The jamming probability for CTD, when friendships form an Erdős-Rényi graph with size $n = 250$ and edge density $p \in [0, 1]$ exhibits a threshold behavior.

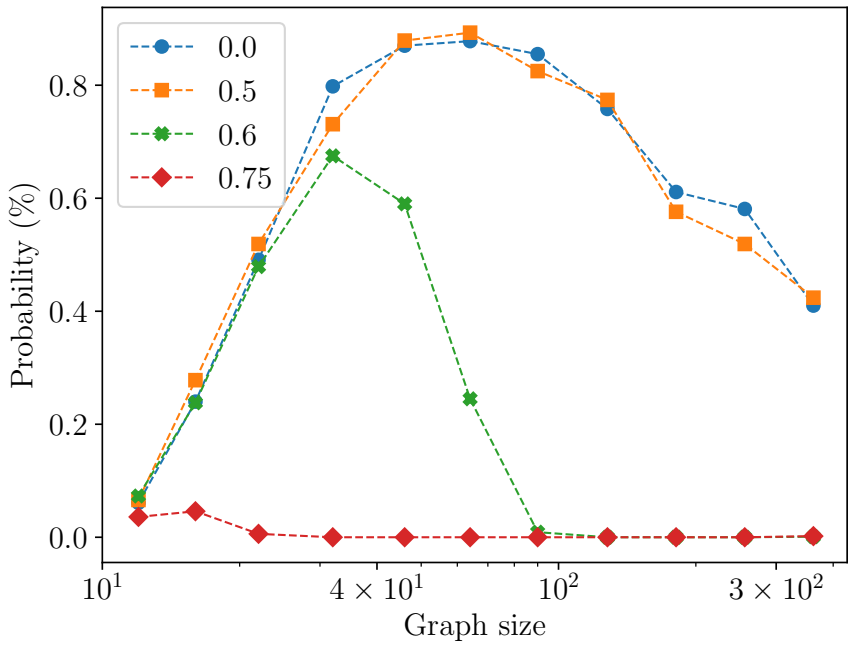



Figure 6.10: The jamming probability for CTD, when friendships form an Erdős-Rényi graph with size $n \leq 400$ and edge density $p \in \{0, 0.5, 0.6, 0.75\}$. When $p \geq 0.75$ (or $p \geq 0.6$ and $n$ large), the dynamics typically reaches utopia, otherwise there is a non-negligible probability of reaching a jammed state.

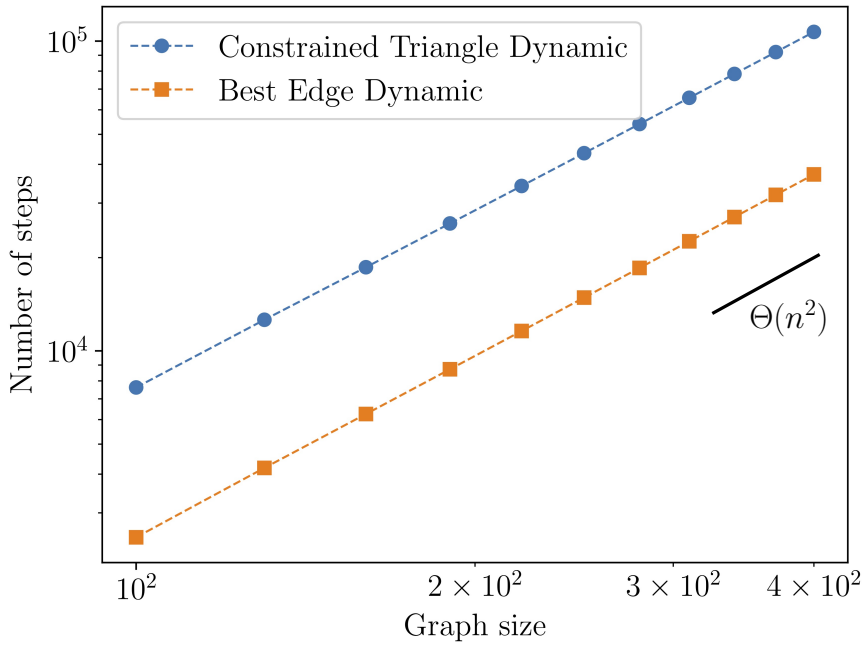



Figure 6.11: Average number of steps until balance for CTD (excluding the runs that get jammed) and BED, over $10^5$ runs. The initial signed graphs are Barabási-Albert with degree parameter $d = 0.5$ and size $n \leq 400$. Both quantities scale as $\Theta(n^2)$.

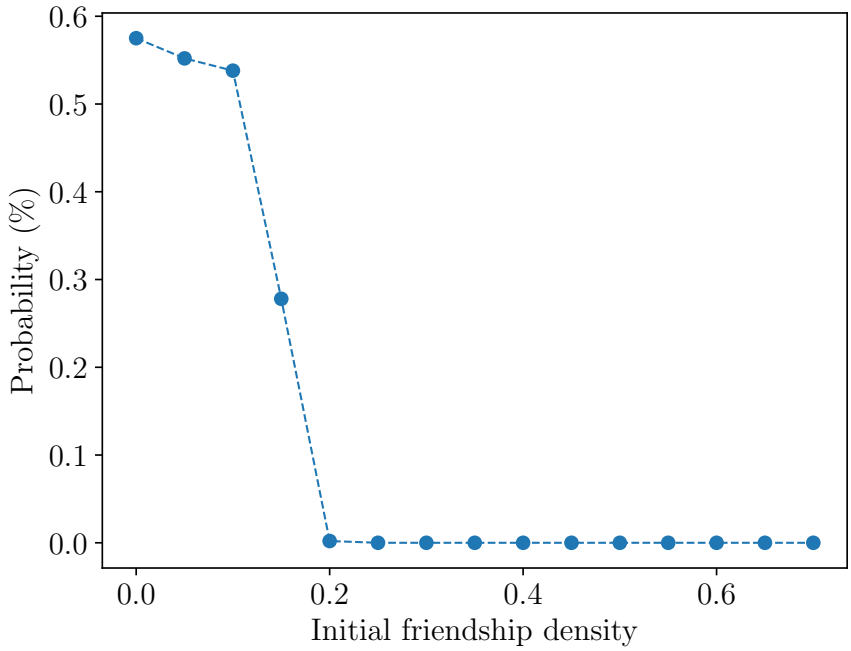



Figure 6.12: The jamming probability in CTD, for Barabási-Albert networks with size $n = 250$ and parameter $d \in [0, 0.7]$ exhibits a threshold behavior comparable to Erdős-Rényi graphs, but with significantly lower edge density.

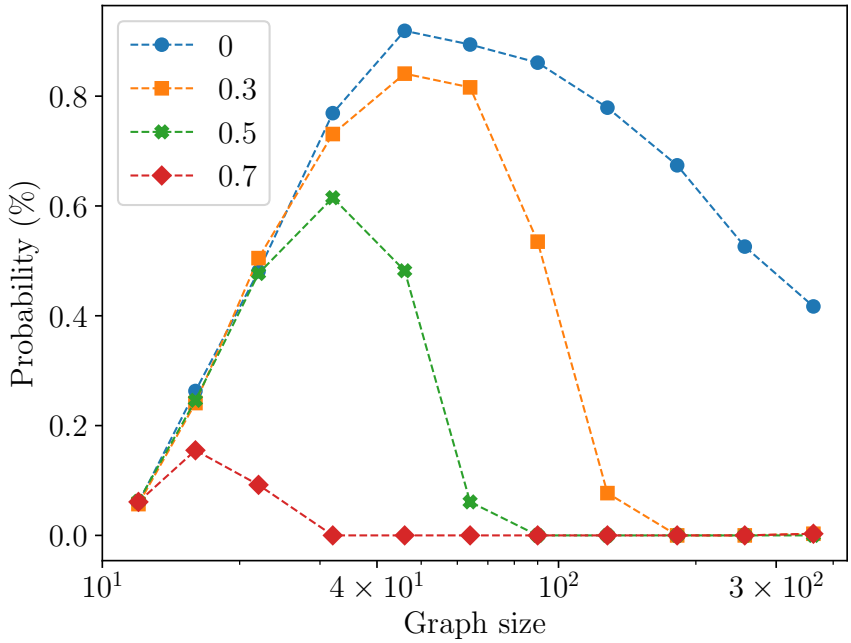



Figure 6.13: The jamming probability in CTD, for Barabási-Albert networks with size $n \leq 400$ and parameters $d \in \{0.0, 0.3, 0.5, 0.7\}$.

## 6.6 Summary and Discussion

The theory of structural balance provides a rigorous framework for the study of friendships and enmities in a population. A central concept in this theory has been the energy landscape of networks, and particularly the energy properties of its local minima. In this paper we have taken a closer look at the properties of these local minima with respect to the stochastic process, addressing questions regarding their reachability and attractor properties. We have shown that there are super-exponentially many jammed states, as opposed to the exponentially many balanced states, and that any initial state that is not too friendship-dense can reach a jammed state. Moreover, such jammed states are attractors, and hence cannot be escaped by random perturbations of the network. These findings have strong implications for the socially-aware CTD process, which in fact gets stuck in such jammed states.

Motivated by these rich reachability and attractor properties of jammed states, we have introduced the plausible socially-aware dynamics BED. We have shown that BED does not get stuck in jammed states and that it always reaches balance. Moreover, we have seen that BED converges fast from many interesting states, such as those that are not too far from balance.

The new BED dynamics spawns some natural questions regarding its asymptotic behavior. Although we have shown that BED converges fast (in $\mathcal{O}(n^2)$ time) to balance from any state that is suitably close to balance, the general convergence rate remains open.

An assumption made throughout our work is that the underlying network is complete. That is, at each point in time, every two individuals have a defined relationship (they are either friends or enemies). It is natural to consider non-complete underlying networks $U$, where only those pairs of individuals who are connected by an edge $e \in U$ have a defined relationship. We note that this generalized setting is considerably more complicated. First, one needs to adapt

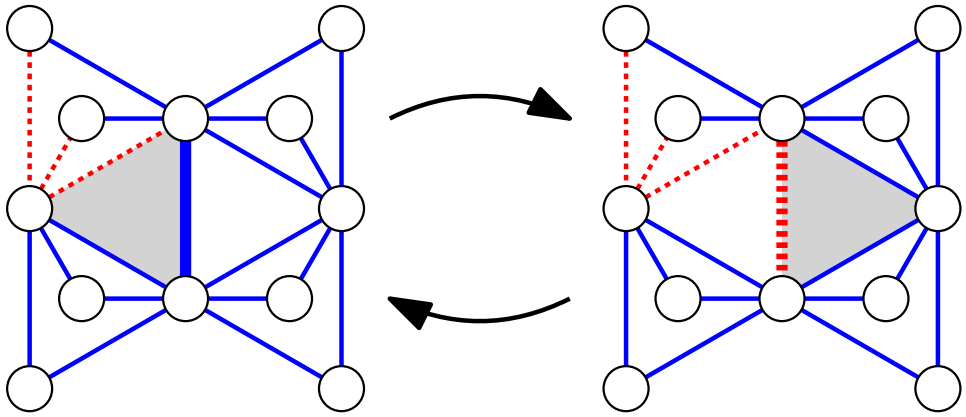


Figure 6.14: A blinker. Under both CTD and BED, the thick edge in the middle keeps toggling between friendship (blue) and enmity (red) indefinitely. There is always only one imbalanced triangle (shaded) and flipping any other its edge would create more imbalanced triangles.

the notion of balance accordingly. One way to do this is to say that a state is balanced if all cycles are balanced, where a cycle in $U$ is balanced if it contains an even number of edges labeled "$-$". While checking whether a current state is balanced can be done efficiently [HK80], several fundamental problems remain. For instance, computing the distance to the closest balanced state is known to be intractable [FIA11]. As another example, to our knowledge the balanced states do not have any simple structure and even the complexity of computing their number (for a given non-complete underlying network $U$) is open. As a final illustration, we note that there exist "blinkers" [AKR05], that is, states where CTD and BED get stuck repeating moves back and forth (rather than getting stuck being unable to make a move), see Fig. 6.14. (Note that when the underlying network is complete, there are no blinkers for BED due to Theorem 6.) Investigating the properties of BED adapted to such generalized settings is thus left as an interesting direction for future research.

## 6.7 Additional proofs

### 6.7.1 Proofs for Reaching and Escaping the Jammed States

Here we formally state and prove our results on jammed states: We present a family of new jammed states, we show that the jammed states vastly outnumber the balanced ones, and we establish the reachability properties stated in the main text.

**Definition 3** (Circular graph $S_k(n_0, \ldots, n_{d-1})$)**.** *Given an integer $k \geq 0$ and a partition $n = n_0 + \cdots + n_{d-1}$ of $n$ into $d$ parts, the* circular graph $S_k(n_0, \ldots, n_{d-1})$ *is a signed graph consisting of $d$ clusters $V_0, \ldots, V_{d-1}$ of sizes $n_0, \ldots, n_{d-1}$, respectively, arranged along a circle in this order, such that the edge $(u, v)$ with $u \in V_i$, $v \in V_j$ is assigned a sign "$+1$" if and only if $V_i$ and $V_j$ are at most $k$ steps apart.*

In particular, when $n$ is divisible by 3 then the circular graph $S_0(n/3, n/3, n/3)$ corresponds to the jammed state $J_n$ from Section 6.3.

**Theorem 7** (Counting jammed states)**.** *There are at least $2^{\Omega(n \log n)}$ jammed states on $n$ labeled vertices.*

*Proof.* The proof proceeds in two steps: First, we show that for any signed graph $S_d(n_0, \dots, n_{4d+1})$ and an edge $(u, v)$ there are at least $2d + 2$ clusters with the property that all the vertices from those clusters form a balanced triad with $(u, v)$. In particular, this immediately implies that the signed graph $S_d^2$ with $n_0 = \cdots = n_{4d+1} = 2$ is jammed: Indeed, any edge $(u, v)$ is contained in at least $2(2d + 2) - 2 = 4d + 2$ balanced triads (the $-2$ comes from omitting the vertices $u$, $v$ themselves) and in at most $2 \cdot 2d = 4d$ imbalanced triads. Second, we show that there are $2^{\Omega(n \log n)}$ ways to draw the signed graph $S_d^2$ over the $n$ labelled vertices, hence at least $2^{\Omega(n \log n)}$ jammed states.

For the first part, suppose that the vertices $u$, $v$ belong to clusters that are $i$ steps apart. We distinguish two cases.

1. $i \leq d$ (that is, $(u, v)$ is labeled "+"): Then there are $2d + 1 - i$ "nearby" clusters whose vertices $w$ form triads $(u, v, w)$ of type $\Delta_0$, and similarly $2d + 1 - i$ "far-away" clusters whose vertices $w$ form triads $(u, v, w)$ of type $\Delta_2$. In total, this is $4d + 2 - 2i \geq 2d + 2$ clusters with the desired property.

2. $i > d$ (that is, $(u, v)$ is labeled "−"): Then there are $2i \geq 2d + 2$ clusters "nearby" either $u$ or $v$ and "far" from the other vertex. All vertices $w$ from those clusters form triads $(u, v, w)$ of type $\Delta_2$.

For the second part, we count only those jammed states in which each cluster has size 2. Note that there are $n - 1$ ways to pick a vertex to join the cluster of vertex 0. Then there are $\binom{n-2}{2}$ ways to select two vertices for the next (clockwise) cluster, then $\binom{n-4}{2}$ ways for the next cluster, and so on. Finally, we must divide by 2, since the same signed graph would be obtained by selecting the vertices in the reverse order (or going counter-clockwise). In total, using the Stirling approximation $n! \geq (n/e)^n$ and a trivial inequality $e\sqrt{2} < 4$, we obtain that the number of different jammed states is at least

$$\begin{aligned} \frac{(n-1)!}{2^{n/2}} &\geq \frac{1}{n} \frac{n!}{(\sqrt{2})^n} \geq \frac{1}{n} (n/4)^n \\ &= 2^{n \log_2 n - 2n - \log_2 n} = 2^{\Omega(n \log n)}. \end{aligned}$$

□

Note that in comparison there are $2^{n-1}$ balanced states, since each balanced state is characterized by a subset of vertices of $\{1, \dots, n-1\}$ which are connected to vertex 0 by a friendship edge. On the other hand, the total number of signed graphs is $2^{\binom{n}{2}} = 2^{\Theta(n^2)}$.

Also, note that each of the $4d + 2$ clusters of the jammed state $S_d^2$ constitutes a balanced clique, in the sense of [MSK09]. This answers in affirmative an open question posed there: For any $m \equiv 2 \pmod 4$ there exists a jammed state with $m$ balanced cliques.

To prove the reachability properties, we define a specific jammed state $J_n$ on $n$ vertices labeled $1, \dots, n$: The edges labeled "+" in $J$ form three roughly equal clusters: One on vertices labeled $1, \dots, \lfloor n/3 \rfloor$, one on vertices labeled $\lfloor n/3 \rfloor + 1, \dots, 2\lfloor n/3 \rfloor$, and one on the remaining vertices $2\lfloor n/3 \rfloor + 1, \dots, n$. It is easy to verify that for $n \geq 11$ the state $J_n$ is jammed.

**Theorem 8** (Reaching a jammed state)**.** *Let $G$ be a signed graph on $n \geq 11$ vertices such that each vertex is incident to at most $\frac{n}{12} - 1$ friendship edges. Then CTD reaches $J_n$ with positive probability.*

Since $J_n$ is jammed, as a corollary we obtain that for any such $G$ the expected time to reach a balanced state is infinite.

*Proof.* For brevity, assume $n \equiv 0 \pmod 3$ (the other cases are completely analogous). We describe a finite sequence of selected imbalanced triads and edge flips that results in $J_n$. Denote the three clusters of $J_n$ by $V_1$, $V_2$ and $V_3$, respectively.

First, we show that one by one, all the enmity edges within each cluster may be flipped into friendship edges. Fix an enmity edge $e = (u, v)$ where $u, v \in V_i$. It suffices to show that, throughout this phase, $e$ belongs to at most $n/2 - 2$ balanced triads. Each balanced triad must contain a friendship edge incident to $u$ or to $v$. Initially, there were $\frac{n}{12} - 1$ friendship edges incident to $u$, and that many to $v$. Moreover, throughout this phase, friendship edges to the other $n/3 - 2$ vertices within the cluster $V_i$ might have been added. In total, this is at most $n/2 - 4$ vertices $w$ connected to one of $u$, $v$ by a friendship edge, thus at most $n/2 - 4$ balanced triads containing $e$. Hence $e$ can be flipped.

Second, we show that one by one, all friendship edges connecting vertices from different clusters may be flipped into enmities. Suppose that $e = (u, v)$ is such a friendship edge. It suffices to find $n/2$ imbalanced triads containing $(u, v)$. Consider the other vertices in the cluster containing $u$. They are all friends with $u$, but at most $n/12 - 1$ of them are friends with $v$ (since we never add friendship edges leading across clusters). Thus there are at least $(n/3 - 1) - (n/12 - 1) = n/4$ imbalanced triads (of type $\Delta_1$) containing $e$ and another vertex in the cluster of $v$. Similarly, there are at least $\frac{n}{4}$ triads of type $\Delta_1$ defined by $e$ and another vertex in the cluster of $u$. Hence $r_e \geq \frac{1}{2}n$, as claimed.

By flipping all friendship edges between different clusters to enmity edges, we reach a jammed state $J_n$ as claimed. □

**Theorem 9** (Escaping a jammed state)**.** *Let $E_0$ be any set of edges such that each vertex is incident to at most $n/12 - 1$ edges of $E_0$. Let $S_n$ be a state obtained from $J_n$ by flipping the edges of $E_0$. Then the CTD run from $S_n$ reaches $J_n$.*

*Proof.* Again, without loss of generality, we assume that $n \equiv 0 \pmod 3$. We first show that no edge $e = (u, v) \notin E_0$ can ever be flipped. In $J_n$, any enmity edge belongs to $2n/3 - 2$ balanced triads (and the friendship edges belong to even $n - 2$ balanced triads). Since $S_n$ differs from $J_n$ by at most $n/12 - 1$ edges incident to each vertex, each edge $e \notin E_0$ belongs to at least $2n/3 - 3 - 2(n/12 - 1) = n/2$ balanced triads in $S_n$ and thus can not be flipped.

On the other hand, any edge $e \in E_0$ belonged to at least $2n/3 - 2$ balanced triads in $J_n$, thus it belongs to at least $2n/3 - 2 - 2(n/12 - 1) = n/2$ imbalanced triads in $S_n$, and as such can be flipped. Moreover, once such an edge has been flipped, by the above argument it cannot be flipped again. Hence CTD will flip each edge in $E_0$ once and return to the jammed state $J_n$. □

### 6.7.2 Red-black graphs

The following lemma formalizes the key property of the red-black graphs.

**Lemma 24.** *Let $G$ be a signed graph, let $C$ be a balanced state closest to $G$, let $R$ be the red-black graph associated to $G$ and $C$. Then a triad in $G$ is imbalanced if and only if exactly $1$ or $3$ of its edges are red in $R$.*

*Proof.* To prove the lemma, we pick a triad in $R$ and check each of the $4$ possible cases:

- If a triad contains $0$ red edges, then all edge signs in $G$ agree with those in $C$ thus the triad is balanced in $G$ as $C$ is a balanced state.
- If a triad contains $1$ red edge, we distinguish two cases. If both vertices of the red edge are in the same vertex class of $C$ (when treated as a bipartite graph w.r.t. the friendship edges), then the triad is of type $\Delta_1$ in $G$. If vertices are in different vertex classes of $C$, then the triad is again of type $\Delta_1$ in $G$. Thus the triad is imbalanced in $G$.
- If a triad contains $2$ red edges, we distinguish two cases. If all $3$ vertices of the triad are in the same vertex class of $C$, then the triad is of type $\Delta_2$ in $G$. If $2$ vertices are in one class and the third vertex is in the other, then depending on which two edges are red the triad is either of type $\Delta_0$ or $\Delta_2$ in $G$. Thus the triad is balanced in $G$.
- If a triad contains $3$ red edges, we distinguish two cases. If all $3$ vertices of the triad are in the same vertex class of $C$, the triad is of type $\Delta_3$ in $G$. If $2$ vertices are in one class and the third vertex is in the other, the triad is of type $\Delta_1$ in $G$. Thus the triad is balanced in $G$. □

### 6.7.3 Fast convergence of BED from a jammed state

**Proposition 3.** *There exists a family of jammed states of increasing size $n$ such that BED starting in those states reaches a balanced state in $\mathcal{O}(n^2)$ expected steps.*

*Proof.* Let $n \geq 72$ be a positive integer divisible by $8$. Consider a circular graph $J'_n = S_1(x, y, y, x, y, y)$, where $x = |V_0| = |V_3| = \frac{1}{4}n - 2$ and $y = |V_i| = \frac{1}{8}n + 1$ for $i \in \{1, 2, 4, 5\}$ (see Definition 3).

First we show that $J'_n$ is jammed. Consider a balanced state $B_n$ with parts $V_0 \cup V_1 \cup V_5$ and $V_2 \cup V_3 \cup V_4$ and denote by $E_0$ the set of red edges in the corresponding red-black graph. Note that there are no triads with all edges red, hence the rank of an edge is the number of triads that contain it and contain precisely one red edge. For any red edge $(u, v)$ we have $r_{(u,v)} = 2x$ due to $V_1$ and $V_4$. Similarly, for any black edge $(u, v)$ such that $|\{u, v\} \cap (V_1 \cup V_2 \cup V_4 \cup V_5)| = 1$ we have $r_{(u,v)} = 2y$ and for other black edges we have $r_{(u,v)} = 0$. Since all ranks are less than $\frac{1}{2}n$, the state $J'_n$ is indeed jammed.

Next, denote by $E_t$ the set of red edges (with respect to the same balanced state $B_n$) after $t$ steps of BED. We will show that:

1. $E_t \subseteq E_0$, and that
2. at each point $t$ in time, $\mathbb{P}[|E_{t+1}| < |E_t|] \geq 2 \cdot \mathbb{P}[|E_{t+1}| > |E_t|]$.

Mapping the evolutionary dynamics to a one-dimensional random walk with a constant forward bias and an absorbing barrier corresponding to $|E_t| = 0$, we thus conclude that the expected number of steps till balance is $\mathcal{O}(|E_0|) = \mathcal{O}(n^2)$.

To prove Item 1, we proceed by induction. Consider $E_t \subseteq E_0$. Note that, as before, there are no triads with all edges red. Also:

1. When $(u, v) \in E_t$ (that is, $(u, v)$ is red) then as before $r_{(u,v)} \geq 2x$ due to triads $(u, v, w)$ with $w \in V_0 \cup V_3$.

2. When $(u, v) \notin E_0$ then $(u, v)$ is black and as before we have $r_{(u,v)} \leq 2y$.

Now consider any imbalanced triad. It contains a red edge. Since $2x > 2y$, we always flip that red edge rather than any edge outside of $E_0$, thus $E_{t+1} \subseteq E_0$ as desired. (Note that it is possible that we flip a black edge in $E_0$.)

To prove Item 2, consider any time point $t$ and any red edge $(u, v)$. We say that an imbalanced triad is *good* if its red edge has a strictly higher rank than its other two edges, and *bad* otherwise. It suffices to show that $(u, v)$ belongs to twice as many good triads as bad triads. Recall that for $w \in V_0 \cup V_3$ the triad $(u, v, w)$ is good, hence $(u, v)$ belongs to at least $2x = 2 \cdot (\frac{1}{4}n - 2) \geq 32$ good triads (here we use $n \geq 72$).

On the other hand, suppose that $(u, v, w)$ is a bad triad and without loss of generality, $(u, w)$ is the (black) edge with rank at least $2x$. Note that $(u, w)$ belongs to $E_0 \setminus E_t$ (other black edges have rank at most $2y$). Denote by $d_i$ the *red degree* of vertex $i$, that is, the number of red edges incident to $i$. Then $2x \leq r_{(u,w)} = d_u + d_w \leq d_u + 2y$, thus $d_u \geq 2x - 2y$. Vertex $u$ is therefore connected to at most $2y - (2x - 2y) = 4y - 2x = 8$ vertices in $E_0$ by a black edge. Each such edge gives rise to at most one bad triad and likewise for the edge $(v, w)$, so in total $(u, v)$ belongs to at most $2 \cdot 8 = 16$ bad triads, concluding the proof. □

### 6.7.4 Network descriptors for BED and CTD on Erdős-Rényi graphs

Here we present the network descriptors when BED and CTD are run on Erdős-Rényi graphs with $n = 128$ and $p \in \{0, 0.4, 0.5, 0.6, 0.7\}$, both before the process starts and after it finishes.

Before

| $p$ | $\overline{d}$ | $C$ | $\mathbb{E}[S]$ | $\mathrm{Var}[S]$ |
|---|---|---|---|---|
| 0 | 0 | 0 | - | - |
| 0.4 | 50.8 | 0.064 | - | - |
| 0.5 | 63.5 | 0.125 | - | - |
| 0.6 | 76.2 | 0.216 | - | - |
| 0.7 | 88.9 | 0.343 | - | - |

After BED

| $p$ | $\overline{d}$ | $C$ | $\mathbb{E}[S]$ | $\mathrm{Var}[S]$ |
|---|---|---|---|---|
| 0 | 63.160 | 0.246 | 61.468 | 3.820 |
| 0.4 | 68.294 | 0.307 | 45.818 | 8.244 |
| 0.5 | 78.144 | 0.423 | 32.990 | 7.596 |
| 0.6 | 103.317 | 0.720 | 13.264 | 6.127 |
| 0.7 | 125.260 | 0.979 | 0.883 | 0.815 |

After CTD

| $p$ | $\overline{d}$ | $C$ | $\mathbb{E}[S]$ | $\mathrm{Var}[S]$ |
|---|---|---|---|---|
| 0 | 63.179 | 0.246 | 61.321 | 4.279 |
| 0.4 | 68.274 | 0.306 | 45.866 | 8.716 |
| 0.5 | 78.116 | 0.423 | 33.025 | 7.967 |
| 0.6 | 103.185 | 0.719 | 13.348 | 6.236 |
| 0.7 | 125.144 | 0.978 | 0.942 | 0.867 |

As in the main text, we average over $10^5$ runs and exclude the runs of CTD that got jammed. Here $\overline{d}$ is the average degree, $C$ is the clustering coefficient, $S$ is the size of the smaller clique once the process finishes, $\mathbb{E}[S]$ is its mean, and $\mathrm{Var}[S]$ its variance. The two dynamics match almost perfectly.

CHAPTER 7

# Future directions and conclusion

This chapter first introduces related work. It consits of games on graphs, distributed computing, and learning and optimization. The motivation for the work is different, the questions come from theoretical computer science and financial cryptography. However, the techniques, especially arguments about stochastic processes, are similar. Second, this chapter presents future directions. The possible directions are: diversity in the models of selection, robust amplification, and considering the mutation. These are underexplored topics in these models. Finally, the chapter presents the conclusion that discusses the mathematical tools developed in the previous chapters.

## 7.1 Related work

### 7.1.1 Games on graphs

Games on graphs model the behavior of a two-player stochastic system. A token is placed on a graph, its movement is influenced by randomness, and two players try to satisfy competing objectives. There are many variants of the problem. In turn-based stochastic games, the vertices are divided between the maximizing player, the minimizing player, and the randomness. If a token lands on a vertex, the owner decides the next edge used, or the randomness chooses the edge according to a determined distribution. Chatterjee et al. [CMSS23] describe a state-of-the-art algorithm for the reachability problem parametrized by treewidth. Asadi et al. [ACSU24] present a novel algorithm for the discounted sum problem with unary weights. In concurrent games, two players choose an action simultaneously, and the token moves according to the pair of actions selected. This setting is explored in [ACSS24], which determines the complexity of the problem.

The games on graphs are relevant to the research on evolutionary graph theory in two ways. First, in both models, the stochasticity poses the biggest problem in finding the correct strategy or the fixation probability. Second, Markov chains that describe the state of the graph in evolution are a special case of the games on graphs.

### 7.1.2 Distributed computing

Distributed computing describes a network of computational units. This model is also general and powerful. For instance, the computational units can represent voters who are allowed

to delegate their votes. Chatterjee et al. [CGS+25] examine the structures for which good delegating rules are possible.

Distributed computing also represents the computational nodes of a blockchain that want to transact with each other. In works [BCM+23, SSY23, ABMA+24], the algorithms for problems that arise in the payment channel networks (PCNs) are explored. In PCNs, participants (nodes) can lock funds into a channel (edge) to facilitate transactions between other nodes of the network and collect fees for enabling these transactions. This creates an online optimization problem that can be approximated by randomized algorithms.

### 7.1.3 Learning and optimization

Learning and optimization can be represented as a movement in an unknown stochastic environment and extracting information from that movement. The movement can be represented by Markov Chains or the generalized Markov decision processes. In [SBC24], a learning algorithm is presented that describes the algorithm with guarantees for a reachability problem parametrized by the length of the solution. Chatterjee et al. [CLSUS25] explore a generalization of games on graphs and show the complexity of a natural extension of the problem.

## 7.2 Future directions and open problems

### 7.2.1 Diversity in populations

A population with multiple genetic variants is more resistant to disease [KL12]. Genetic diversity increases the long-term evolutionary potential [BS08] and makes the population resilient towards change in the environment [HS11]. The structure of the population influences the speed with which different variants disappear from the population. In diversity (or coexistence), the goal is to maintain many different types for as long as possible.

The model for diversity follows the Moran process. The graph describes a population structure, and initially, different types reside in every vertex. In general, the fitness of an individual can depend on the type, the neighborhood arrangement, or the inhabited vertex. Then, the process (Birth-death or death-Birth) is simulated on the network. Possibly, the process is augmented by mutation, where, with a small probability, the offspring differs from the parent. The diversity of a population can be measured as the time until all individuals have only one type.

Work of Brewster et al. [BSR+25] provides the initial exploration of the diversity in the neutral drift. However, there are many unexplored questions: what are the exact bounds for the time, or what are the structures that maximize the number of edges with different endpoints in time?

### 7.2.2 Robust amplification

The amplification is an important concept in both constant selection [TPCN21, GLL+19] and frequency-dependent selection [SC24, FMAN18]. The graphs that promote the advantageous mutation or cooperative behaviour usually do not promote the desired behavior in slightly different settings.

The work [SJTC24] introduces an amplifier for both processes, however, there are no guarantees for the process when some Birth-death steps can be followed by a death-Birth step, for instance, when the Birth-death step happens with probability $\delta$ and the death-Birth step happens with the probability $1 - \delta$. Then, there are no known amplifiers.

Moreover, the known amplifiers for cooperation give guarantees only for weak or very strong selection. There are no amplifiers that promote cooperation in all strengths of selection, and the known amplifiers are not very robust against the parameters of the game.

### 7.2.3 Mutation

The models of selection disregard mutation. This is a well-founded assumption in the evolution of small populations; however, looking at populations of larger size, the mutation should be considered. In the presence of mutation, there is a small probability $\mu$ such that the offspring is of a different type than the parent.

The mutation poses many questions to consider. First, the mutation can produce individuals with different fitnesses: what is the structure that maximizes the long-run average of fitnesses? Second, the mutation introduces diversity: what structures increase the diversity the most in the presence of mutation? Finally, different lineages of mutants can have different fitnesses. For instance, a lineage has to pass through an intermediate state with lower fitness. What is the best structure for these mutation lineages?

## 7.3 Conclusion

This thesis introduces important concepts related to the role of structure in the models of evolution. It identifies problems inspired by multiple fields, such as biology, social dynamics, and physics. The thesis develops new mathematical tools and solves the proposed problems.

The main problems are highlighted in the Chapter 1. This section is dedicated to the discussion of the development of the mathematical tools used to solve the problems. Moreover, the important concepts are also highlighted.

Chapter 2 presents the first structures that are amplifiers for the Birth-death and death-Birth updating. One of the most important notions of the work is the question of robustness of the Moran process with constant selection. The important tool to prove the amplification process is the careful treatment of the weights of the graph. With carefully selecting the weights, an edge between two vertices $u$ and $v$ can be created, such that in the Birth-death process, the individuals are significantly more likely to reproduces from $u$ to $v$ and in the death-Birth process, the individuals are significantly more likely to reproduce from $v$ to $u$. This is an important observation that can be used to argue about the combination of Birth-death and death-Birth processes.

Chapter 3 presents the complexity proof for the spatial games on graphs for a very simple graph. It adapts and extends the ideas from [CIJS20], which shows that if some logical gates can be created in some process, then this process can simulate the Turing machine. This is an important simplification for the complexity proofs in these processes.

Chapter 4 introduces the first amplifiers for spatial games. The main theoretical contribution is the analysis of the process on the proposed graphs. The analysis computes the probability of one cooperator spreading in a small subgraph of the proposed graph. The main techniques

previously focused on the regime with weak selection, and this work complements this with a focus on the strong selection.

Chapter 5 proves the exact bounds for the coexistence (diversity) time on heterogeneous networks. The main theoretical contribution is Lemma 23 that bounds the absorption times for potent Markov Chains. Potent Markov Chains are a special class of multidimensional Markov Chains with symmetry between some states. The symmetry allows bounding the absorption time.

Chapter 6 examines a dynamic of signed graphs and presents a rule with guarantees for the convergence. An interesting theoretical tool in these dynamics is a red-black graph, which describes the difference from the desired state. This notion simplifies many proofs and can find wider applicability.

The work presented in the thesis not only solves important problems but also presents useful mathematical tools that have wider application in the analysis of evolutionary or stochastic processes.

# Bibliography


[ABMA+24] Zeta Avarikioti, Mahsa Bastankhah, Mohammad Ali Maddah-Ali, Krzysztof Pietrzak, Jakub Svoboda, and Michelle Yeo. Route discovery in private payment channel networks. In *European Symposium on Research in Computer Security*, pages 207–223. Springer, 2024.

[ACN15] Ben Adlam, Krishnendu Chatterjee, and Martin A Nowak. Amplifiers of selection. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150114, 2015.

[ACSS24] Ali Asadi, Krishnendu Chatterjee, Raimundo Saona, and Jakub Svoboda. Concurrent stochastic games with stateful-discounted and parity objectives: Complexity and algorithms. In Siddharth Barman and Slawomir Lasota, editors, *44th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2024, December 16-18, 2024, Gandhinagar, Gujarat, India*, volume 323 of *LIPIcs*, pages 5:1–5:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[ACSU24] Ali Asadi, Krishnendu Chatterjee, Jakub Svoboda, and Raimundo Saona Urmeneta. Deterministic sub-exponential algorithm for discounted-sum games with unary weights. In Pawel Sobocinski, Ugo Dal Lago, and Javier Esparza, editors, *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2024, Tallinn, Estonia, July 8-11, 2024*, pages 6:1–6:12. ACM, 2024.

[AFC15] Marco Archetti, Daniela A. Ferraro, and Gerhard Christofori. Heterogeneity for igf-ii production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *Proceedings of the National Academy of Sciences*, 112(6):1833–1838, 2015.

[AGLR20] Leslie Ann Goldberg, John Lapinskas, and David Richerby. Phase transitions of the moran process and algorithmic consequences. *Random Structures & Algorithms*, 56(3):597–647, 2020.

[AH81] Robert Axelrod and William D Hamilton. The evolution of cooperation. *science*, 211(4489):1390–1396, 1981.

[AKR05] T. Antal, P. L. Krapivsky, and S. Redner. Dynamics of social balance on networks. *Phys. Rev. E*, 72:036121, Sep 2005.

[AKR06] T. Antal, P.L. Krapivsky, and S. Redner. Social balance on networks: The dynamics of friendship and enmity. *Physica D: Nonlinear Phenomena*, 224(1):130 – 136, 2006. Dynamics on Complex Networks and Applications.

[ALC+17] Benjamin Allen, Gabor Lippner, Yu-Ting Chen, Babak Fotouhi, Naghmeh Momeni, Shing-Tung Yau, and Martin A Nowak. Evolutionary dynamics on any population structure. *Nature*, 544(7649):227–230, 2017.

[ALN19] Benjamin Allen, Gabor Lippner, and Martin A Nowak. Evolutionary games on isothermal graphs. *Nature communications*, 10(1):1–9, 2019.

[Alt12] Claudio Altafini. Consensus problems on networks with antagonistic interactions. *IEEE transactions on automatic control*, 58(4):935–946, 2012.

[AM19] Benjamin Allen and Alex McAvoy. A mathematical formalism for natural selection with arbitrary spatial and genetic structure. *Journal of mathematical biology*, 78(4):1147–1210, 2019.

[ANT09] Tibor Antal, Martin A Nowak, and Arne Traulsen. Strategy abundance in $2\times 2$ games for arbitrary mutation rates. *Journal of theoretical biology*, 257(2):340–344, 2009.

[ARS06] Tibor Antal, Sidney Redner, and Vishal Sood. Evolutionary dynamics on degree-heterogeneous graphs. *Physical review letters*, 96(18):188104, 2006.

[ASBF20] Omid Askarisichani, Ambuj K Singh, Francesco Bullo, and Noah E Friedkin. The 1995-2018 global evolution of the network of amicable and hostile relations among nation-states. *Communications Physics*, 3(1):1–11, 2020.

[ASJ+20] Benjamin Allen, Christine Sample, Robert Jencks, James Withers, Patricia Steinhagen, Lori Brizuela, Joshua Kolodny, Darren Parke, Gabor Lippner, and Yulia A Dementieva. Transient amplifiers of selection and reducers of fixation for death-birth updating on graphs. *PLoS computational biology*, 16(1):e1007529, 2020.

[AT09] Philipp M Altrock and Arne Traulsen. Fixation times in evolutionary games under weak selection. *New Journal of Physics*, 11(1):013012, 2009.

[Ave78] PJ Avery. Selection effects in a model of two intermigrating colonies of finite size. *Theoretical population biology*, 13(1):24–39, 1978.

[Axe97] Robert Axelrod. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration: Agent-Based Models of Competition and Collaboration*. Princeton university press, 1997.

[BA99] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[BCG04] Elwyn R Berlekamp, John H Conway, and Richard K Guy. *Winning Ways for Your Mathematical Plays, Volume 4*. AK Peters/CRC Press, 2004.

[BCM+23] Mahsa Bastankhah, Krishnendu Chatterjee, Mohammad Ali Maddah-Ali, Stefan Schmid, Jakub Svoboda, and Michelle Yeo. R2: boosting liquidity in payment channel networks with online admission control. In Foteini Baldimtsi and Christian Cachin, editors, *Financial Cryptography and Data Security - 27th International Conference, FC 2023, Bol, Brač, Croatia, May 1-5, 2023, Revised Selected Papers, Part I*, volume 13950 of *Lecture Notes in Computer Science*, pages 309–325. Springer, 2023.

[BGS+17] Aaron Bramson, Patrick Grim, Daniel J Singer, William J Berger, Graham Sack, Steven Fisher, Carissa Flocken, and Bennett Holman. Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of science*, 84(1):115–159, 2017.

[BHR+17] Andres M Belaza, Kevin Hoefman, Jan Ryckebusch, Aaron Bramson, Milan van den Heuvel, and Koen Schoors. Statistical physics of balance theory. *PLoS one*, 12(8):e0183696, 2017.

[BHRS10] M Broom, C Hadjichrysanthou, J Rychtář, and BT Stadler. Two results on evolutionary processes on general non-directed graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2121):2795–2798, 2010.

[BKP+22] Joachim Brendborg, Panagiotis Karras, Andreas Pavlogiannis, Asger Ullersted Rasmussen, and Josef Tkadlec. Fixation maximization in the positional moran process. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 9304–9312, 2022.

[BR08] Mark Broom and Jan Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2098):2609–2627, 2008.

[BR22] Mark Broom and Jan Rychtář. *Game-theoretical models in biology*. Chapman and Hall/CRC, 2022.

[BRS11] M. Broom, J. Rychtář, and B.T. Stadler. Evolutionary dynamics on graphs - the effect of graph structure and initial placement on mutant spread. *J Stat. Theory Pract.*, 5(3):369–381, 2011.

[BS08] Rowan DH Barrett and Dolph Schluter. Adaptation from standing genetic variation. *Trends in ecology & evolution*, 23(1):38–44, 2008.

[BSR+25] David A Brewster, Jakub Svoboda, Dylan Roscow, Krishnendu Chatterjee, Josef Tkadlec, and Martin A Nowak. Maintaining diversity in structured populations. *arXiv preprint arXiv:2503.09841*, 2025.

[Bul72] MG Bulmer. Multiple niche polymorphism. *The American Naturalist*, 106(948):254–257, 1972.

[Cam11] Colin F Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton university press, 2011.

[CCC+20] Yen-Sheng Chiang, Yen-Wen Chen, Wen-Chi Chuang, Chyi-In Wu, and Chien-Te Wu. Triadic balance in the brain: Seeking brain evidence for heider's structural balance theory. *Social Networks*, 63:80–90, 2020.

[CCSS24] Esra Ceylan, Krishnendu Chatterjee, Stefan Schmid, and Jakub Svoboda. Congestion-free rerouting of network flows: Hardness and an FPT algorithm. In *NOMS 2024 IEEE Network Operations and Management Symposium, Seoul, Republic of Korea, May 6-10, 2024*, pages 1–7. IEEE, 2024.

[CDT09] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the ACM (JACM)*, 56(3):1–57, 2009.

[CF75] Freddy Bugge Christiansen and Marcus W Feldman. Subdivided populations: a review of the one-and two-locus deterministic theory. *Theoretical Population Biology*, 7(1):13–38, 1975.

[CFL09] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591–646, 2009.

[CGS+25] Krishnendu Chatterjee, Seth Gilbert, Stefan Schmid, Jakub Svoboda, and Michelle Yeo. When is liquid democracy possible? on the manipulation of variance. In *Proceedings of the 44rd ACM Symposium on Principles of Distributed Computing*, 2025.

[CH56] D Cartwright and F Harary. Structural balance: a generalization of heider's theory. *Psychological review*, pages 277–293, 1956.

[Cha14] Fabio ACC Chalub. Asymptotic expression for the fixation probability of a mutant in star graphs. *arXiv preprint arXiv:1404.3944*, 2014.

[Cha16] Fabio A. C. C. Chalub. An asymptotic expression for the fixation probability of a mutant in star graphs. *Journal of Dynamics and Games*, 3:217–223, 2016.

[Chr74] Freddy Bugge Christiansen. Sufficient conditions for protected polymorphism in a subdivided population. *The American Naturalist*, 108(960):157–166, 1974.

[CIJS20] Krishnendu Chatterjee, Rasmus Ibsen-Jensen, Ismaël Jecker, and Jakub Svoboda. Simplified game of life: Algorithms and complexity. In Javier Esparza and Daniel Král', editors, *45th International Symposium on Mathematical Foundations of Computer Science, MFCS 2020, August 24-28, 2020, Prague, Czech Republic*, volume 170 of *LIPIcs*, pages 22:1–22:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[CIJS22] Krishnendu Chatterjee, Rasmus Ibsen-Jensen, Ismaël Jecker, and Jakub Svoboda. Complexity of spatial games. In Anuj Dawar and Venkatesan Guruswami, editors, *42nd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2022, December 18-20, 2022, IIT Madras, Chennai, India*, volume 250 of *LIPIcs*, pages 11:1–11:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[CJSS25] Krishnendu Chatterjee, Mahdi JafariRaviz, Raimundo Saona, and Jakub Svoboda. Value iteration with guessing for markov chains and markov decision processes. In Arie Gurfinkel and Marijn Heule, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 31st International Conference, TACAS 2025, Held as Part of the International Joint Conferences on Theory and Practice of Software, ETAPS 2025, Hamilton, ON, Canada, May 3-8, 2025, Proceedings, Part II*, volume 15697 of *Lecture Notes in Computer Science*, pages 217–236. Springer, 2025.

[CLSS25] Krishnendu Chatterjee, Ruichen Luo, Raimundo Saona, and Jakub Svoboda. Linear equations with min and max operators: Computational complexity. In Toby Walsh, Julie Shah, and Zico Kolter, editors, *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, pages 11150–11157. AAAI Press, 2025.

[CLSUS25] Krishnendu Chatterjee, Ruichen Luo, Raimundo J Saona Urmeneta, and Jakub Svoboda. Linear equations with min and max operators: Computational complexity. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence*, volume 39, 2025.

[CMSS23] Krishnendu Chatterjee, Tobias Meggendorfer, Raimundo Saona, and Jakub Svoboda. Faster algorithm for turn-based stochastic games with bounded treewidth. In Nikhil Bansal and Viswanath Nagarajan, editors, *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 4590–4605. SIAM, 2023.

[Cor09] Thomas H Cormen. *Introduction to algorithms*. MIT press, 2009.

[CS73] Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

[CSŽ+22] Krishnendu Chatterjee, Jakub Svoboda, Djordje Žikelić, Andreas Pavlogiannis, and Josef Tkadlec. Social balance on networks: Local minima and best-edge dynamics. *Physical Review E*, 106(3):034321, 2022.

[D+09] Rick Durrett et al. Coexistence in stochastic spatial models. *The Annals of Applied Probability*, 19(2):477–496, 2009.

[Dav67] James A Davis. Clustering and structural balance in graphs. *Human relations*, 20(2):181–187, 1967.

[Dea66] Michael AB Deakin. Sufficent conditions for genetic polymorphism. *The American Naturalist*, 100(916):690–692, 1966.

[DGM+13] Josep Díaz, Leslie Ann Goldberg, George B Mertzios, David Richerby, Maria Serna, and Paul G Spirakis. On the fixation probability of superstars. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 469(2156):20130193, 2013.

[DGM+14] Josep Díaz, Leslie Ann Goldberg, George B Mertzios, David Richerby, Maria Serna, and Paul G Spirakis. Approximating fixation probabilities in the generalized moran process. *Algorithmica*, 69:78–91, 2014.

[DGP09] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[DGRS16] Josep Díaz, Leslie Ann Goldberg, David Richerby, and Maria Serna. Absorption time of the moran process. *Random Structures & Algorithms*, 49(1):137–159, 2016.

[DHD14] Florence Débarre, Christoph Hauert, and Michael Doebeli. Social evolution in structured populations. *Nature Communications*, 5(1):3409, 2014.

[DKPT22] Loke Durocher, Panagiotis Karras, Andreas Pavlogiannis, and Josef Tkadlec. Invasion dynamics in the biased voter process. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pages 265–271, 2022.

[DL94a] Richard Durrett and Simon Levin. The importance of being discrete (and spatial). *Theoretical population biology*, 46(3):363–394, 1994.

[DL94b] Richard Durrett and Simon A Levin. Stochastic spatial models: a user's guide to ecological applications. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.*, 343(1305):329–350, 1994.

[DLL10] Delphine Dufour, Vincent Leung, and Céline M Lévesque. Bacterial biofilm: structure, function, and antimicrobial resistance. *Endodontic Topics*, 22(1):2–16, 2010.

[Dug97] Lee Alan Dugatkin. *Cooperation among animals: an evolutionary perspective*. Oxford University Press, USA, 1997.

[Dur02] Richard Durrett. *Mutual invadability implies coexistence in spatial models*, volume 740. American Mathematical Soc., 2002.

[EDS07] Eyal Even-Dar and Asaf Shapira. A note on maximizing the spread of influence in social networks. In *International Workshop on Web and Internet Economics*, pages 281–286. Springer, 2007.

[Ewe04] Warren John Ewens. *Mathematical population genetics: theoretical introduction*, volume 27. Springer, 2004.

[FA21] Angela Fontan and Claudio Altafini. A signed network perspective on the government formation process in parliamentary democracies. *Scientific reports*, 11(1):1–17, 2021.

[Fel76] Joseph Felsenstein. The theoretical population genetics of variable selection and migration. *Annual review of genetics*, 10(1):253–280, 1976.

[FF03] Ernst Fehr and Urs Fischbacher. The nature of human altruism. *Nature*, 425(6960):785–791, 2003.

[FIA11] Giuseppe Facchetti, Giovanni Iacono, and Claudio Altafini. Computing global structural balance in large-scale signed social networks. *Proceedings of the National Academy of Sciences*, 108(52):20953–20958, 2011.

[Fis99] Ronald Aylmer Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1999.

[FM99] D Fudenberg and E Maskin. Evolution and cooperation in noisy repeated games. *International Library of Critical Writings in Economics*, 109:339–344, 1999.

[FMAN18] Babak Fotouhi, Naghmeh Momeni, Benjamin Allen, and Martin A Nowak. Conjoining uncooperative societies facilitates evolution of cooperation. *Nature human behaviour*, 2(7):492–499, 2018.

[FRT13] Marcus Frean, Paul B Rainey, and Arne Traulsen. The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B: Biological Sciences*, 280(1762):20130211, 2013.

[GG02] Sergey Gavrilets and Nathan Gibson. Fixation probabilities in a spatially heterogeneous environment. *Population Ecology*, 44(2):51–58, 2002.

[GGCFM07] Jesús Gómez-Gardenes, Michel Campillo, Luis Mario Floría, and Yamir Moreno. Dynamical organization of cooperation in complex topologies. *Physical Review Letters*, 98(10):108103, 2007.

[GGG+17] Andreas Galanis, Andreas Göbel, Leslie Ann Goldberg, John Lapinskas, and David Richerby. Amplifiers for the moran process. *Journal of the ACM (JACM)*, 64(1):1–90, 2017.

[Gia16] George Giakkoupis. Amplifiers and suppressors of selection for the moran process on undirected graphs. *arXiv preprint arXiv:1611.01585*, 2016.

[Gil75] John H Gillespie. The role of migration in the genetic structure of populations in temporarily and spatially varying environments. i. conditions for polymorphism. *The American Naturalist*, 109(966):127–136, 1975.

[GLL+19] Leslie Ann Goldberg, John Lapinskas, Johannes Lengler, Florian Meier, Konstantinos Panagiotou, and Pascal Pfister. Asymptotically optimal amplifiers for the moran process. *Theoretical Computer Science*, 758:73–93, 2019.

[GS75] Chris Gliddon and Curtis Strobeck. Necessary and sufficient conditions for multiple-niche polymorphism in haploids. *The American Naturalist*, 109(966):233–235, 1975.

[GYVO09] Jeff Gore, Hyun Youk, and Alexander Van Oudenaarden. Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459(7244):253–256, 2009.

[H+53] Frank Harary et al. On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2(2):143–146, 1953.

[Har74] Theodore E Harris. Contact interactions on a lattice. *The Annals of Probability*, 2(6):969–988, 1974.

[HBR11] Christophoros Hadjichrysanthou, Mark Broom, and Jan Rychtár. Evolutionary games on star graphs under various updating rules. *Dynamic Games and Applications*, 1(3):386–407, 2011.

[HCM94] Michael P. Hassell, Hugh N. Comins, and Robert M. May. Species coexistence and self-organizing spatial dynamics. *Nature*, 370:290–292, 1994.

[HD04] Christoph Hauert and Michael Doebeli. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428(6983):643–646, 2004.

[Hed86] Philip W Hedrick. Genetic polymorphism in heterogeneous environments: a decade later. *Annual review of ecology and systematics*, 17(1):535–566, 1986.

[Hei44] F. Heider. Social perception and phenomenal causality. *Psychological Review*, 51(6):358–374, 1944.

[Hei46] Fritz Heider. Attitudes and cognitive organization. *The Journal of Psychology*, 21(1):107–112, 1946. PMID: 21010780.

[Hei58] Fritz Heider. *The psychology of interpersonal relations.* The psychology of interpersonal relations. John Wiley & Sons Inc, Hoboken, NJ, US, 1958.

[HG04] Ilkka A Hanski and Oscar E Gaggiotti. *Ecology, genetics and evolution of metapopulations*. Academic Press, 2004.

[HK80] Frank Harary and Jerald A Kabell. A simple algorithm to detect balance in signed graphs. *Mathematical Social Sciences*, 1(1):131–136, 1980.

[HS88] Josef Hofbauer and Karl Sigmund. The theory of evolution and dynamical systems: mathematical aspects of selection. *(No Title)*, 1988.

[HS98] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

[HS11] Ary A Hoffmann and Carla M Sgrò. Climate change and evolutionary adaptation. *Nature*, 470(7335):479–485, 2011.

[HT15] Laura Hindersin and Arne Traulsen. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. *PLoS computational biology*, 11(11):e1004437, 2015.

[HV11] Bahram Houchmandzadeh and Marcel Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New J. Phys.*, 13:073020, July 2011.

[HY09] Dirk Helbing and Wenjian Yu. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences*, 106(10):3680–3685, 2009.

[IJCN15] Rasmus Ibsen-Jensen, Krishnendu Chatterjee, and Martin A Nowak. Computational complexity of ecological and evolutionary spatial dynamics. *Proceedings of the National Academy of Sciences*, 112(51):15636–15641, 2015.

[IS01] Marianne Imhof and Christian Schlötterer. Fitness effects of advantageous mutations in evolving escherichia coli populations. *Proceedings of the National Academy of Sciences*, 98(3):1113–1117, 2001.

[Isi25] Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[JHK+22] Marko Jusup, Petter Holme, Kiyoshi Kanazawa, Misako Takayasu, Ivan Romić, Zhen Wang, Sunčana Geček, Tomislav Lipić, Boris Podobnik, Lin Wang, et al. Social physics. *Physics Reports*, 948:1–148, 2022.

[JMG+13] Thomas Julou, Thierry Mora, Laurent Guillon, Vincent Croquette, Isabelle J Schalk, David Bensimon, and Nicolas Desprat. Cell–cell contacts confine public goods diffusion inside pseudomonas aeruginosa clonal microcolonies. *Proceedings of the National Academy of Sciences*, 110(31):12577–12582, 2013.

[JP13] Luo-Luo Jiang and Matjaž Perc. Spreading of cooperative behaviour across interdependent groups. *Scientific reports*, 3(1):2483, 2013.

[KCN19] Alec Kirkley, George T. Cantwell, and M. E. J. Newman. Balance in signed networks. *Phys. Rev. E*, 99:012320, Jan 2019.

[KKK15] Kamran Kaveh, Natalia L Komarova, and Mohammad Kohandel. The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. *Royal Society open science*, 2(4):140465, 2015.

[KL12] KC King and CM Lively. Does genetic diversity limit disease spread in natural host populations? *Heredity*, 109(4):199–203, 2012.

[KM72] Samuel Karlin and James McGregor. Application of method of small parameters to multi-niche population genetic models. *Theoretical population biology*, 3(2):186–209, 1972.

[KMCN20] Kamran Kaveh, Alex McAvoy, Krishnendu Chatterjee, and Martin A Nowak. The moran process on 2-chromatic graphs. *PLOS Computational Biology*, 16(11):e1008402, 2020.

[KMN19] Kamran Kaveh, Alex McAvoy, and Martin A Nowak. Environmental fitness heterogeneity in the moran process. *Royal Society open science*, 6(1):181661, 2019.

[KMR93] Michihiro Kandori, George J Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pages 29–56, 1993.

[Kom06] Natalia L Komarova. Spatial stochastic models for cancer initiation and progression. *Bulletin of mathematical biology*, 68:1573–1599, 2006.

[KS23] Jan Matyáš Křistan and Jakub Svoboda. Shortest dominating set reconfiguration under token sliding. In Henning Fernau and Klaus Jansen, editors, *Fundamentals of Computation Theory - 24th International Symposium, FCT 2023, Trier, Germany, September 18-21, 2023, Proceedings*, volume 14292 of *Lecture Notes in Computer Science*, pages 333–347. Springer, 2023.

[KS25] Jan Matyás Kristan and Jakub Svoboda. Reconfiguration using generalized token jumping. In Shin-ichi Nakano and Mingyu Xiao, editors, *WALCOM: Algorithms and Computation - 19th International Conference and Workshops on Algorithms and Computation, WALCOM 2025, Chengdu, China, February 28 - March 2, 2025, Proceedings*, volume 15411 of *Lecture Notes in Computer Science*, pages 244–265. Springer, 2025.

[KSW+14] Rolf Kümmerli, Konstanze T Schiessl, Tuija Waldvogel, Kristopher McNeill, and Martin Ackermann. Habitat structure and the evolution of diffusible siderophores in bacteria. *Ecology letters*, 17(12):1536–1544, 2014.

[KTW18] KM Ariful Kabir, Jun Tanimoto, and Zhen Wang. Influence of bolstering network reciprocity in the evolutionary spatial prisoner's dilemma game: a perspective. *The European Physical Journal B*, 91(12), December 2018.

[LCWZ13] Yanhua Li, Wei Chen, Yajun Wang, and Zhi-Li Zhang. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 657–666, 2013.

[Len19] Johannes Lengler. Drift analysis. In *Theory of evolutionary computation: Recent developments in discrete optimization*, pages 89–131. Springer, 2019.

[Lev53] Howard Levene. Genetic equilibrium when more than one ecological niche is available. *The American Naturalist*, 87(836):331–333, 1953.

[LHN05] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005.

[Lig12] Thomas Milton Liggett. *Interacting particle systems*, volume 276. Springer Science & Business Media, 2012.

[Lig13] Thomas M Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324. springer science & Business Media, 2013.

[Lin10] Zhengyan Lin. *Probability inequalities*. Springer, 2010.

[LM66] Richard Levins and Robert MacArthur. The maintenance of genetic polymorphism in a spatially heterogeneous environment: variations on a theme by howard levene. *The American Naturalist*, 100(916):585–589, 1966.

[Lot25] Alfred James Lotka. *Elements of physical biology*. Williams & Wilkins, 1925.

[Lot27] Alfred J Lotka. Fluctuations in the abundance of a species considered mathematically. *Nature*, 119(2983):12–12, 1927.

[Mac14] Wes Maciejewski. Reproductive value in graph-structured populations. *Journal of Theoretical Biology*, 340:285–293, 2014.

[May87] Robert M May. More evolution of cooperation. *Nature*, 327(6117):15–17, 1987.

[MGP14] Travis Monk, Peter Green, and Mike Paulin. Martingales and fixation probabilities of evolutionary graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 470(2165):20130730, 2014.

[MHT19] Marius Möller, Laura Hindersin, and Arne Traulsen. Exploring and mapping the universe of evolutionary graphs identifies structural properties affecting fixation probability and time. *Communications biology*, 2(1):137, 2019.

[MLB21] Loïc Marrec, Irene Lamberti, and Anne-Florence Bitbol. Toward a universal model for spatially structured populations. *Physical review letters*, 127(21):218102, 2021.

[MM21] Pouya Manshour and Afshin Montakhab. Dynamics of social balance on networks: The emergence of multipolar societies. *Physical Review E*, 104(3):034303, 2021.

[Moo79] Michael Moore. Structural balance and international relations. *European Journal of Social Psychology*, 9(3):323–326, 1979.

[Mor58] Patrick Alfred Pierce Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.

[Mor62] Patrick Alfred Pierce Moran. *The statistical processes of evolutionary theory*. Oxford University Press, 1962.

[MP14] Wes Maciejewski and Gregory J Puleo. Environmental evolutionary graph theory. *Journal of theoretical biology*, 360:117–128, 2014.

[MPKHT20] Tuan Minh Pham, Imre Kondor, Rudolf Hanel, and Stefan Thurner. The effect of social balance on social fragmentation. *Journal of The Royal Society Interface*, 17(172):20200752, 2020.

[MR97] Sebastián M. Ruiz. 81.27 a result on prime numbers. *The Mathematical Gazette*, 81:269, 07 1997.

[MRH21] Alex McAvoy, Andrew Rao, and Christoph Hauert. Intriguing effects of selection intensity on the evolution of prosocial behaviors. *PLoS Computational Biology*, 17(11):e1009611, 2021.

[MS70] J Maynard Smith. Genetic polymorphism in a varied environment. *The American Naturalist*, 104(939):487–490, 1970.

[MSK09] Seth A. Marvel, Steven H. Strogatz, and Jon M. Kleinberg. Energy landscape of social balance. *Phys. Rev. Lett.*, 103:198701, Nov 2009.

[MvS20] Travis Monk and André van Schaik. Wald's martingale and the conditional distributions of absorption time in the moran process. *Proceedings of the Royal Society A*, 476(2241):20200135, 2020.

[MvS21] Travis Monk and André van Schaik. Martingales and the characteristic functions of absorption time on bipartite graphs. *Royal Society Open Science*, 8(10):210657, 2021.

[Nas51] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[NDF16] Carey D Nadell, Knut Drescher, and Kevin R Foster. Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology*, 14(9):589–600, 2016.

[NM92a] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 10 1992.

[NM92b] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, October 1992.

[Now06a] Martin A Nowak. *Evolutionary dynamics: exploring the equations of life*. Harvard University Press, 2006.

[Now06b] Martin A Nowak. Five rules for the evolution of cooperation. *science*, 314(5805):1560–1563, 2006.

[Now07] Martin Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*, volume 82. Belknap Press, 01 2007.

[NRTV07] Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[NSTF04] Martin A Nowak, Akira Sasaki, Christine Taylor, and Drew Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428(6983):646–650, 2004.

[OHLN06] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.

[ONP07] Hisashi Ohtsuki, Martin A. Nowak, and Jorge M. Pacheco. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.*, 98:108106, Mar 2007.

[Ons44] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical review*, 65(3-4):117, 1944.

[OR94] Martin J Osborne and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.

[Owe13] Guillermo Owen. *Game theory*. Emerald Group Publishing, 2013.

[Per16] Matjaž Perc. Phase transitions in models of human cooperation. *Physics Letters A*, 380(36):2803–2808, 2016.

[PJR+17] Matjaž Perc, Jillian J. Jordan, David G. Rand, Zhen Wang, Stefano Boccaletti, and Attila Szolnoki. Statistical physics of human cooperation. *Physics Reports*, 687:1–51, 2017. Statistical physics of human cooperation.

[PS08] Matjaž Perc and Attila Szolnoki. Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. *Phys. Rev. E*, 77:011904, Jan 2008.

[PS10a] Matjaž Perc and Attila Szolnoki. Coevolutionary games—a mini review. *BioSystems*, 99(2):109–125, 2010.

[PS10b] Matjaž Perc and Attila Szolnoki. Coevolutionary games—a mini review. *Biosystems*, 99(2):109–125, 2010.

[PTCN17] Andreas Pavlogiannis, Josef Tkadlec, Krishnendu Chatterjee, and Martin A Nowak. Amplification on undirected population structures: comets beat stars. *Scientific reports*, 7(1):82, 2017.

[PTCN18] Andreas Pavlogiannis, Josef Tkadlec, Krishnendu Chatterjee, and Martin A Nowak. Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory. *Communications biology*, 1(1):71, 2018.

[PW10] Matjaž Perc and Zhen Wang. Heterogeneous aspirations promote cooperation in the prisoner's dilemma game. *PLOS ONE*, 5(12):1–8, 12 2010.

[PWAT16] Jorge Peña, Bin Wu, Jordi Arranz, and Arne Traulsen. Evolutionary games of multiplayer cooperation on graphs. *PLoS computational biology*, 12(8):e1005059, 2016.

[RCS06] Carlos P. Roca, José A. Cuesta, and Angel Sánchez. Time scales in evolutionary dynamics. *Phys. Rev. Lett.*, 97:158701, Oct 2006.

[RF17] Craig M Rawlings and Noah E Friedkin. The structural balance theory of sentiment networks: Elaboration and test. *American Journal of Sociology*, 123(2):510–548, 2017.

[Ric21] Hendrik Richter. Spectral analysis of transient amplifiers for death–birth updating constructed from regular graphs. *Journal of Mathematical Biology*, 82(7):61, 2021.

[Ric23] Hendrik Richter. Spectral dynamics of guided edge removals and identifying transient amplifiers for death–birth updating. *Journal of Mathematical Biology*, 87(1):3, 2023.

[SAP22] Qi Su, Benjamin Allen, and Joshua B Plotkin. Evolution of cooperation with asymmetric social interactions. *Proceedings of the National Academy of Sciences*, 119(1):e2113468118, 2022.

[SAR08] Vishal Sood, Tibor Antal, and Sidney Redner. Voter models on heterogeneous networks. *Physical Review E*, 77(4):041121, 2008.

[SATJK17] M Saeedian, N Azimi-Tafreshi, GR Jafari, and J Kertesz. Epidemic spreading on evolving signed networks. *Physical Review E*, 95(2):022314, 2017.

[SBC24] Jakub Svoboda, Suguman Bansal, and Krishnendu Chatterjee. Reinforcement learning from reachability specifications: PAC guarantees with expected conditional distance. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024.

[SC24] Jakub Svoboda and Krishnendu Chatterjee. Density amplifiers of cooperation for spatial games. *Proceedings of the National Academy of Sciences*, 121(50):e2405605121, 2024.

[Sch10] Thomas Schwartz. The friend of my enemy is my enemy, the enemy of my enemy is my friend: Axioms for structural balance and bi-polarity. *Mathematical Social Sciences*, 60(1):39–45, 2010.

[SF07] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.

[SF08] Philip S Stewart and Michael J Franklin. Physiological heterogeneity in biofilms. *Nature Reviews Microbiology*, 6(3):199–210, 2008.

[She71] Ronald G Sherwin. Introduction to the graph theory and structural balance approaches to international relations. Technical report, University of Southern California Los Angeles, 1971.

[SJTC24] Jakub Svoboda, Soham Joshi, Josef Tkadlec, and Krishnendu Chatterjee. Amplifiers of selection for the moran process with both birth-death and death-birth updating. *PLOS Computational Biology*, 20(3):1012008, 2024.

[Sla81] Montgomery Slatkin. Fixation probabilities and fixation times in a subdivided population. *Evolution*, pages 477–488, 1981.

[Smi82] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.

[SMMP22] Qi Su, Alex McAvoy, Yoichiro Mori, and Joshua B Plotkin. Evolution of prosocial behaviours in multilayer populations. *Nature Human Behaviour*, 6(3):338–348, 2022.

[SP05] Francisco C Santos and Jorge M Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical review letters*, 95(9):098104, 2005.

[SP13] Alexander J Stewart and Joshua B Plotkin. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38):15348–15353, 2013.

[SPL06a] Francisco C. Santos, Jorge M. Pacheco, and Tom Lenaerts. Cooperation prevails when individuals adjust their social ties. *PLOS Computational Biology*, 2(10):1–8, 10 2006.

[SPL06b] Francisco C Santos, Jorge M Pacheco, and Tom Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences*, 103(9):3490–3494, 2006.

[SS07] Attila Szolnoki and Gyorgy Szabó. Cooperation enhanced by inhomogeneous activity of teaching for evolutionary prisoner's dilemma games. *Europhysics Letters (EPL)*, 77(3):30004, jan 2007.

[SSP08] Francisco C Santos, Marta D Santos, and Jorge M Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454(7201):213–216, 2008.

[SSY23] Stefan Schmid, Jakub Svoboda, and Michelle Yeo. Weighted packet selection for rechargeable links in cryptocurrency networks: Complexity and approximation. In *International Colloquium on Structural Information and Communication Complexity*, pages 576–594. Springer, 2023.

[SSY24] Stefan Schmid, Jakub Svoboda, and Michelle Yeo. Weighted packet selection for rechargeable links in cryptocurrency networks: Complexity and approximation. *Theor. Comput. Sci.*, 989:114353, 2024.

[ST22] Nikhil Sharma and Arne Traulsen. Suppressors of fixation can increase average fitness beyond amplifiers of selection. *Proceedings of the National Academy of Sciences*, 119(37):e2205424119, 2022.

[STKC23] Jakub Svoboda, Josef Tkadlec, Kamran Kaveh, and Krishnendu Chatterjee. Coexistence times in the moran process with environmental heterogeneity. *Proceedings of the Royal Society A*, 479(2271):20220685, 2023.

[STP+22] Jakub Svoboda, Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A Nowak. Infection dynamics of covid-19 virus under lockdown and reopening. *Scientific reports*, 12(1):1526, 2022.

[Tar17] Corina E Tarnita. The ecology and evolution of social behavior in microbes. *Journal of Experimental Biology*, 220(1):18–24, 2017.

[Tay70] Howard Francis Taylor. *Balance in small groups*. New York: Van Nostrand Reinhold Company, 1970.

[TI91] Hidenori Tachida and Masaru Iizuka. Fixation probability in spatially changing environments. *Genetics Research*, 58(3):243–251, 1991.

[TJ78] Peter D Taylor and Leo B Jonker. Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1-2):145–156, 1978.

[TK97] David Tilman and Peter Kareiva. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, 1997.

[TKCN23] Josef Tkadlec, Kamran Kaveh, Krishnendu Chatterjee, and Martin A Nowak. Evolutionary dynamics of mutants that modify population structure. *Journal of the Royal Society Interface*, 20(208):20230355, 2023.

[TOA+09] Corina E Tarnita, Hisashi Ohtsuki, Tibor Antal, Feng Fu, and Martin A Nowak. Strategy selection in structured populations. *Journal of theoretical biology*, 259(3):570–581, 2009.

[TPCN19] Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A Nowak. Population structure determines the tradeoff between fixation probability and fixation time. *Communications biology*, 2(1):138, 2019.

[TPCN20] Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A Nowak. Limits on amplifiers of natural selection under death-birth updating. *PLoS computational biology*, 16(1):e1007494, 2020.

[TPCN21] Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A Nowak. Fast and strong amplifiers of natural selection. *Nature Communications*, 12(1):4009, 2021.

[TSS+10] Arne Traulsen, Dirk Semmann, Ralf D. Sommerfeld, Hans-Jürgen Krambeck, and Manfred Milinski. Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences*, 107(7):2962–2966, 2010.

[Vir08] Virgil. *Eclogue IV*. Cambridge University Press, 2008.

[VNM47] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1947.

[VTA07] Mendeli H. Vainstein, Ana T.C. Silva, and Jeferson J. Arenzon. Does mobility decrease cooperation? *Journal of Theoretical Biology*, 244(4):722–728, 2007.

[WG05] Michael C Whitlock and Richard Gomulkiewicz. Probability of fixation in a heterogeneous environment. *Genetics*, 171(3):1407–1417, 2005.

[WGeH03] Stephen Wolfram and M Gad-el Hak. A new kind of science. *Appl. Mech. Rev.*, 56(2):B18–B19, 2003.

[Whi03] Michael C Whitlock. Fixation probability and time in subdivided populations. *Genetics*, 164(2):767–779, 2003.

[Wol83] Stephen Wolfram. Statistical mechanics of cellular automata. *Reviews of modern physics*, 55(3):601, 1983.

[WRH09] Zhi-Xi Wu, Zhihai Rong, and Petter Holme. Diversity of reproduction time scale promotes cooperation in spatial prisoner's dilemma games. *Phys. Rev. E*, 80:036106, Sep 2009.

[Wri31] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97–159, 03 1931.

[Wri43] Sewall Wright. Isolation by distance. *Genetics*, 28(2):114, 1943.

[WSP13a] Zhen Wang, Attila Szolnoki, and Matjaž Perc. Interdependent network reciprocity in evolutionary games. *Scientific reports*, 3(1):1183, 2013.

[WSP13b] Zhen Wang, Attila Szolnoki, and Matjaž Perc. Optimal interdependence between networks for the evolution of cooperation. *Scientific reports*, 3(1):2470, 2013.

[WWZA12] Zhen Wang, Zhen Wang, Xiaodan Zhu, and Jeferson J. Arenzon. Cooperation and age structure in spatial games. *Phys. Rev. E*, 85:011149, Jan 2012.

[Yeo92] Julia M Yeomans. *Statistical mechanics of phase transitions*. Clarendon Press, 1992.

[YO11] Sam Yeaman and Sarah P Otto. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution: International Journal of Organic Evolution*, 65(7):2123–2129, 2011.

[YST23] Sedigheh Yagoobi, Nikhil Sharma, and Arne Traulsen. Categorizing update mechanisms for graph-structured metapopulations. *Journal of the Royal Society Interface*, 20(200):20220769, 2023.

[YT21] Sedigheh Yagoobi and Arne Traulsen. Fixation probabilities in network structured meta-populations. *Scientific Reports*, 11(1):17979, 2021.

[ZZW15] Xiaolong Zheng, Daniel Zeng, and Fei-Yue Wang. Social balance in signed networks. *Information Systems Frontiers*, 17(5):1077–1095, 2015.