Spatio-temporal dynamics of random-access networks : an interacting particle approach

Zocca, A.

Published: 17/12/2015

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

SPATIO-TEMPORAL DYNAMICS OF RANDOM-ACCESS NETWORKS: AN INTERACTING PARTICLE APPROACH

This work was financially supported by The Netherlands Organization for Scientific Research (NWO) through the TOP-GO grant 613.001.012

Nederlandse Organisatie voor Wetenschappelijk Onderzoek

Spatio-Temporal Dynamics of Random-Access Networks: An Interacting Particle Approach / by A. Zocca - Eindhoven University of Technology, 2015

Mathematics Subject Classification (2010): 60K35, 68M20, 90B15, 60K25

A catalogue record is available from the Eindhoven University of Technology library
ISBN: 978-90-386-3981-9

Printed by Gildeprint Drukkerijen, Enschede

SPATIO-TEMPORAL DYNAMICS OF RANDOM-ACCESS NETWORKS: AN INTERACTING PARTICLE APPROACH

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Technische Universiteit Eindhoven, op gezag van
de rector magnificus, prof.dr.ir. F.P.T. Baaijens, voor
een commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen op
donderdag 17 december 2015 om 16:00 uur

door

Alessandro Zocca

geboren te Bussolengo, Italië

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. B. Koren
1e promotor:	prof.dr.ir. S.C. Borst
2e promotor:	prof.dr. J.S.H. van Leeuwaarden
copromotor:	dr. F.R. Nardi
leden:	prof.dr. W.T.F. den Hollander (Universiteit Leiden)
	prof.dr. M.R.H. Mandjes (Universiteit van Amsterdam)
	prof.dr.ir. D.H.J. Epema
	prof.dr. M.A. Peletier

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

ACKNOWLEDGMENTS

This thesis would have never seen the light without the help and the support, either direct or indirect, of many people and I would like to use this opportunity to express my gratitude to them.

First and foremost, I am deeply indebted to my supervisors, Sem and Johan. Thanks for giving me the opportunity to pursue my PhD under your guidance and for having taught me so much over these four years. I benefited immensely from your deep insight and your constructive advice. Always with a positive outlook, you have constantly pushed me to go one step further or one level deeper with my research. I hope to have made this mindset mine, as well as your thoroughness and preciseness. Sem, I am truly grateful for your massive support and helpfulness, and, Johan, I wish I took some of your eagerness and frankness. I will always look up to you, both as researchers and as supervisors.

I am also immensely grateful to Francesca, who kindly invited me in when I knocked at her door three years ago. Since then, she has enthusiastically shared with me her knowledge and ideas; I am glad that those discussions matured in a fruitful collaboration.

I would like to express my sincere gratitude also to Frank den Hollander, Michael Mandjes, Mark Peletier and Dick Epema for agreeing to serve on my doctoral committee and for reading and commenting on this thesis.

A special thanks should be given to my co-authors, Boris and Alessandro, for involving me in your projects to which I was glad to contribute, and to Julien for the helpful discussions.

My journey into the probability realm would have never been undertaken without the guidance my bachelor supervisor, Paolo Dai Pra: It was during his lectures that I fell in love with this branch of mathematics. The dynamical and motivating atmosphere I found later during my masters in Cambridge, together with the energy and the drive of the people I met there (in particular Hanna, Jann, Hanns Hagen, and Marius) gave me a tremendous momentum, which ultimately drove my decision to pursue a PhD in mathematics.

The Stochastics group at TU/e has been a great and lively place to work, here I found many colleagues willing to share their knowledge and venture in fruitful discussions. Thank you to you all, and, in particular, to its two pillars, Onno and Remco, who are probably the nicest and most helpful bosses I have ever met.

Moreover, I have been extremely lucky to have an office just few meters away from EURANDOM, whose countless seminars and workshops creates a marvelous flow of scientists and knowledge in which I delightedly dived in.

Across the last four years I have witnessed several generations of PhD students in the Stochastics group and I have had good time with all of them,

during and after work; thank you all for your friendship. Among these many fellow PhDs I should probably mention Serban and Jaron, with whom I shared my entire four-year path at TU/e, and the several office mates I had. I want to thank in particular Gianmarco, Jori, and Marta for bearing with me in the most difficult period of my PhD: Both your support and your cheerfulness have been crucial in the last year.

The long hours at the office would not have been bearable without a daily dosage of endorphin: I wish to thank you all my volleyball teammates at Hajraa and my climbing friends at Monk for the countless hours spent training and having fun together. A special thought goes to Stefano, who accompanied me in both my volleyball and climbing adventures here in the Netherlands.

These four years in the Netherlands would have not been the same without the huge Italian PhD community at TU/e, a "Little Italy" here in Eindhoven. Feeling homesick has been nearly impossible having you all around and, ironically, I got to know more "Italy" in these four years in Eindhoven than when I was living there. Among them, Carlo and Gianmarco deserve a special mention for having been lovely housemates. A big thank you goes also to all the friends from Verona that are still part of my life, regardless of the distance between us.

Beatrice, Enzo, Federica, Pietro, and Leonardo, your constant encouragement and affection despite my intermittent presence in Italy means a lot to me; I always feel at home when I am around you, thank you for this.

I own a great debt of gratitude to my parents, Lucia and Giuseppe, and my brother Pietro for their love and unconditional support. You cannot imagine how rewarding it is to see you so proud of me. I hope you could consider my achievements, such as this doctorate degree, also a bit as your own.

My last thought goes to you, Chiara, thanks for being with me all the way. You, more than anyone else, shared with me the joys and the challenges of this PhD experience abroad, I would probably would have not gotten this far without you by my side. Your are my smile and I am looking forward to our future together.

CONTENTS

1 INTRODUCTION

1.1 PREVIEW

In this thesis we study mathematical models that capture the collective behavior of devices sharing a wireless medium in a distributed fashion.

Our research is motivated by fundamental challenges in wireless networks, which typically are very large and lack centralized control. Instead these networks vitally rely on a distributed mechanism for regulating the access of the various devices to the shared medium. Randomized algorithms provide a popular mechanism for distributed medium access control, thanks to their low implementation complexity.

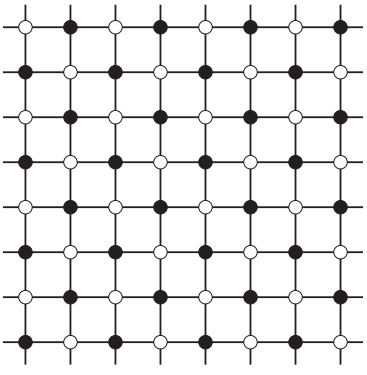
The idea behind random access is simple: Each device operates a random clock, independent of all other devices, which determines when the device attempts to access the medium. So from the viewpoint of the device and the implementation, this randomized algorithm is simple as it only requires local information. However, the macroscopic behavior of this random-access algorithm in large networks tends to be complex and critically depends on global spatial characteristics of the network, in particular when devices have sensing capabilities. Indeed, for such a class of algorithms nearby devices are prevented from simultaneous activity, since otherwise they would interfere and disturb each others signals.

In our work we focus on random-access algorithms. We consider stylized stochastic models to understand how the spatial deployment of the various transmitter-receiver pairs affects the global performance of the network. Specifically, we model the random-access network as an interacting particle system on a graph, which captures the interplay of conflicting transmissions due to interference and to which we will refer as *conflict graph*. Every transmitter-receiver pair is represented by a particle, which is active if there is an ongoing transmission or inactive otherwise. The graph encodes the spatial structure of the network, since neighboring particles are prevented from simultaneous activity. Each particle is endowed with a random clock, which is independent of the clocks of all other particles. Such random clocks create the dynamics of the system: When one of these random clocks rings, the corresponding particle changes its state, from active to inactive or vice-versa. However, a particle can become active only if all its neighboring particles are inactive. The global evolution of the particle system is then described by a continuous-time Markov process, which exhibits fascinating connections with the hard-core interaction between gas particles studied in chemistry and statistical mechanics.

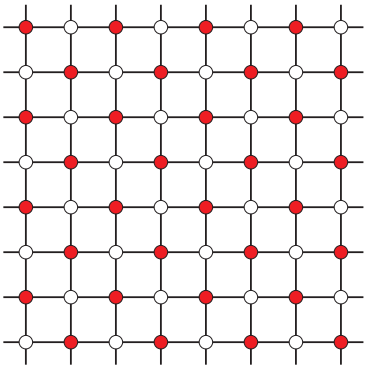
Specific attention is paid to scenarios in which particles become more aggressive in trying to activate, with higher clock rates for activation compared

to deactivation. This is relevant for networks in high-load regimes, where one cannot afford to leave network resources unutilized. In these scenarios, the most likely activity states for the network are those with a maximum number of active particles, to which we refer as *dominant states*. A major complication then for network behavior is that in this high-load regime the dominant states become extremely rigid, by which we mean that the transitions between dominant states can be extremely slow, causing starvation for the particles not in the currently active dominant state.

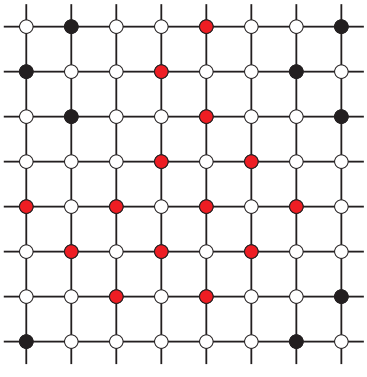
As an illustration, consider a symmetric grid network of particles equipped with random clocks with the same activation rate. This network has two dominant states, corresponding to the two chessboard patterns of the grid, see Figures 1.1a and 1.1b. When the activation rate grows large, the system spends roughly half of the time in each of these two dominant states. However, it takes a long time for the activity process to move from one dominant state to the other, since such a transition involves the occurrence of rare events. In-



(a) First dominant state



(b) Second dominant state



(c) Intermediate state

Figure 1.1: Three activity states on an 8×8 grid network with $N = 64$ particles

deed, intuitively, the activity process must follow a transition path through some highly unlikely states with simultaneous activity on the even and on the odd chessboards, such as the one in Figure 1.1c, and the time to reach such activity states is correspondingly long. These mixed activity patterns are extremely low-likelihood states because more inactive particles than in the dominant states are needed to create the two-layer interface that must separate clusters of active particles on different chessboards. Thus each particle basically experiences long sequences of activations in rapid succession, interspersed with extended periods of starvation. Hence, significant temporal starvation effects can arise due to these slow transitions between dominant states.

Temporal starvation can have a huge impact on the network performance. Suppose that each particle is equipped with a buffer and let the data packets waiting for transmission be stored in the buffer of the corresponding particle, as shown by Figure 1.2. During a long period of inactivity of a given particle the number of data packets in the buffer builds up, causing delay in their transmission. In this thesis, we exploit the particle system description of the network to investigate the long transition times between dominant activity states and the temporal starvation and delay that this causes.

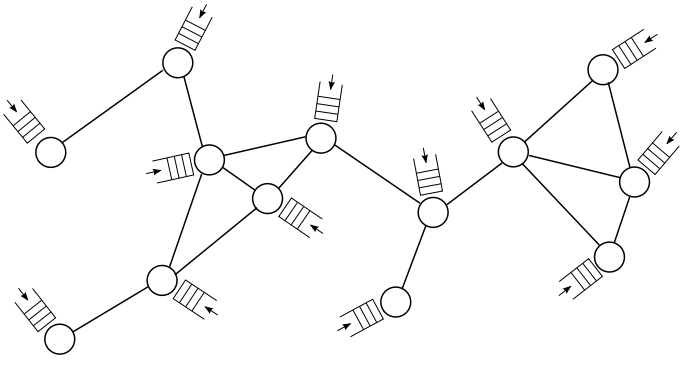


Figure 1.2: Network of particles corresponding to conflicting transmission links with their data packet buffers

In order to quantify the temporal starvation of random-access networks, we study the transient behavior of the corresponding interacting particle systems. In particular, we analyze the asymptotic behavior of first hitting times between dominant states and obtain the convergence rate to equilibrium of the Markov process in terms of mixing times. In our study we focus on extreme network topologies as prototypical scenarios. We first consider complete partite graphs, which are graphs that are useful for modeling dense networks, and therefore provide a "worst-case" perspective. We then turn attention to non-complete partite graphs; in particular regular meshes such as grid graphs and triangular grid graphs. For these topologies, we develop a novel combinatorial method to analyze the structure and the features of the typical transition

paths between dominant states. As a result, we then quantify how the order of magnitude of the transition times between dominant states depends on the size and properties of the network topology.

We also consider extensions of the interacting particle system described above, in which the particles are not all equal, but can be of different types. In particular, we assume that the repelling effects continue to exist only between particles of different types, while neighboring particles of homogeneous type can be active simultaneously. This class of interacting particle systems can represent random-access algorithms used in wireless networks in which nearby devices can transmit simultaneously without interfering by transmitting on different frequency channels. Multi-channel random-access networks and their performance are further discussed in Section 1.5.

1.2 WIRELESS NETWORKS

1.2.1 *Medium access schemes*

Wireless signals typically propagate in all directions and are often overheard by non-intended receivers. As a consequence, data packets may not be received correctly if the intended receiver overhears too many conflicting signals on the same frequency channel. In this case, we say that a *collision* occurs due to the *interference* of nearby ongoing conflicting transmissions. Hence, a medium access control mechanism is required in order to reduce collisions and improve network performance. Many such mechanisms have been proposed and studied in the literature, aiming either to detect such collisions when they occur or to avoid them altogether. Within the class of collision-avoidance medium access algorithms, there exist two main sub-classes: *Centralized algorithms* and *distributed algorithms*.

Centralized algorithms require a global control entity which has perfect information of all interference constraints and can coordinate all transmissions by prescribing a certain schedule to the nodes of the network, i.e. the transmitter-receiver pairs. The *MaxWeight scheduling algorithm* [134] is an example of a centralized scheme that operates in slotted time. In each time slot, a set of non-conflicting nodes (an independent set) in the conflict graph that has the maximum weight in terms of aggregated buffer content sizes is scheduled. However, implementing MaxWeight is challenging because finding a maximum-weighted independent set (required at each time slot) is an NP-hard problem.

In contrast, distributed algorithms do not require a central controller, since nodes decide autonomously when to start a transmission using only local information. These schemes often rely on randomness to both avoid simultaneous transmissions and share the medium in the most efficient way: For this reason they are called *random-access algorithms*.

The ALOHA protocol [1, 124], developed in the 1970's, was one of the first schemes that exploit randomness for medium access. According to ALOHA, every node is inactive for a random amount of time after every transmission attempt. This *back-off mechanism* attenuates simultaneous activity of nearby nodes and thus reduces the chances of collisions, although collisions can still occur.

The Carrier-Sense Multiple-Access (CSMA) algorithm refines ALOHA by combining the random back-off mechanism with interference sensing [86]. In the CSMA algorithm, each node attempts to access the medium after a random back-off time, but nodes that sense activity of interfering nodes freeze their back-off timer until the medium is sensed idle again. For this reason, the CSMA mechanism is collision-free, at least in idealized scenarios. The CSMA algorithm is a popular distributed medium-access control mechanism and, in fact, various of its incarnations are currently implemented in IEEE 802.11 WiFi networks.

In this thesis we consider stylized models for CSMA-type algorithms to investigate the impact of the network structure on the network performance.

1.2.2 *Saturated CSMA model*

We consider a network of transmitter-receiver pairs sharing a wireless medium according to a CSMA-type algorithm. A *node* indicates potential data transmission between a transmitter and a receiver. In this thesis we assume that every transmitter has exactly one intended receiver, and we consider only the links that correspond to these intended transmissions. We further assume that every receiver needs to receive data from only one transmitter. These assumptions are not particularly crucial, and results can be generalized to other scenarios. Every node can either be active or inactive, depending on whether the data transmission is ongoing or not. We assume that the network consists of N such nodes, so that the network activity state can then be described by an N-dimensional vector x, where $x_i = 1$ if node i is active and $x_i = 0$ otherwise.

We assume that the network structure and interference conditions can be described by means of an undirected finite graph $G = (V, E)$, called *conflict graph*, where the set of vertices $V = \{1, \ldots, N\}$ represents the nodes of the network and the set of edges $E \subseteq V \times V$ indicate which pairs of nodes cannot be active simultaneously. Therefore, neighboring nodes in the conflict graph are prevented from simultaneous activity by the carrier-sensing mechanism.

We first focus on scenarios where nodes are saturated, which means that nodes always have packets available for transmission. The transmission times of node i are independent and exponentially distributed with mean $1/\mu_i$. When the transmission of a packet is completed, node i releases the medium and starts a back-off period. The back-off periods of node i are independent and exponentially distributed with mean $1/\nu_i$.

Let $\mathcal{X} \subseteq \{0,1\}^N$ be the set of all feasible joint activity states of the network. Since the interference is modeled by the conflict graph G, the set $\mathcal{X}$ consists of the incidence vectors of all independent sets of the conflict graph G:

$$\mathcal{X} := \Big\{ x \in \{0,1\}^N \ : \ x_i x_j = 0 \ \forall \, (i,j) \in E \Big\}. \tag{1.1}$$

If we let $X(t)$ denote the network activity state at time t as defined earlier, then $\{X(t)\}_{t \geq 0}$ is a continuous-time Markov process on the state space $\mathcal{X}$ with transition rates

$$q(x,y) := \begin{cases} \nu_i & \text{if } y = x + e_i \in \mathcal{X}, \\ \mu_i & \text{if } y = x - e_i \in \mathcal{X}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.2}$$

where $e_i \in \{0,1\}^N$ is the vector with all zeros except for a 1 in position i. The Markov process $\{X(t)\}_{t \geq 0}$ is reversible [18] and has a product-form stationary distribution

$$\pi(x) := Z^{-1} \prod_{i=1}^{N} \Big(\frac{\nu_i}{\mu_i} \Big)^{x_i}, \quad x \in \mathcal{X}, \tag{1.3}$$

where Z is the normalizing constant

$$Z := \sum_{x \in \mathcal{X}} \prod_{i=1}^{N} \Big(\frac{\nu_i}{\mu_i} \Big)^{x_i}.$$

The stationary distribution (1.3) is insensitive to the distributions of back-off periods and transmission times, in the sense that it depends on these only through their averages $1/\nu_i$ and $1/\mu_i$, as proved in [136]. Hence (1.3) holds in fact for general back-off and transmission time distributions.

Define the *throughput* θ_i of node i as the expected number of successful transmissions of node i per unit of time, multiplied by the average length of a transmission period. Since nodes are saturated, θ_i equals the long-run fraction of time that node i is active, i.e.

$$\theta_i := \sum_{x \in \mathcal{X}} \pi(x) x_i. \tag{1.4}$$

Being a function of the stationary distribution (1.3), the throughput is also insensitive to the distributions of the back-off and transmission times.

This CSMA model was originally considered in the context of packet-radio networks [17, 18, 81, 84, 119, 140] and later investigated in the context of IEEE 802.11 systems, first in [138] and then in [43, 44, 45, 46, 55]. Although such a representation of the IEEE 802.11 back-off mechanism is not as detailed as in the landmark work of [16], the general conflict graph provides far greater versatility in describing a broad range of topologies. Experimental

results in [100] demonstrate that these models, while idealized, yield throughput estimates that match remarkably well with measurements in actual IEEE 802.11 networks.

This saturated CSMA model has strong connections with Markov random fields and can be seen as a special instance of a loss network [80, 82, 83, 133, 142]. Several models in this class have been shown to oscillate on very large timescales between two (or more) stable states. In particular, *bi-stability* has been observed in loss networks, both with dynamic routing [60, 107] and fixed routing [144]. In [7] and [8] similar network models have been shown to have multiple stable points between which the systems move on long time scales. The metastability of loss networks is closely related to phase transitions of the corresponding models in statistical physics [52, 105, 122].

1.2.3 *Unsaturated CSMA model*

We can also consider CSMA networks in unsaturated traffic conditions. Data packets are then assumed to arrive at each node according to an exogenous stochastic process and are temporarily stored in a FIFO manner in a buffer at the corresponding node pending transmission (more precisely at the buffer of the transmitter corresponding to that node). We assume that the packet arrival processes at different nodes are independent Poisson processes and that the packet arrival rate at node i is λ_i. The *traffic intensity* or *load* at node i is defined as

$$\rho_i := \frac{\lambda_i}{\mu_i}, \tag{1.5}$$

representing the long-run fraction of time that node i must be active to transmit all arriving packets.

Upon activation, a single packet is transmitted and the node starts a backoff period before the next packet transmission. The activation process and transmission times at each node are assumed to be independent of the corresponding packet arrival process. We assume that a node can be active, even when there are no data packets waiting to be transmitted in its buffer. In that case a node transmits dummy packets. The transmission of a dummy packet is stopped upon arrival of a new packet at that node, whose transmission is then started immediately. It would be interesting to extend the analysis to the case where nodes with empty buffers refrain from transmission activity, in which we expect many of our results to still hold, especially when the network is highly loaded (since in this regime buffers are rarely empty). However, this extension would be challenging from the mathematical standpoint, since the behavior of the activity process then does depend on the queue-length process.

Let $\{Q(t)\}_{t\geq 0}$ be the joint queue-length process, with $Q_i(t)$ the number of packets waiting at node i at time t (excluding the packet that may be in the process of being transmitted). Under the random-access algorithm, the process

$\{(X(t), Q(t))\}_{t\geq 0}$ evolves as a continuous-time Markov process on the state space $\mathcal{X} \times \mathbb{N}^N$. The activity process $\{X(t)\}_{t\geq 0}$ does not depend on the process $\{Q(t)\}_{t\geq 0}$, thanks to the dummy packet assumption. In contrast, the queue-length process $\{Q(t)\}_{t\geq 0}$ strongly depends on the activity process $\{X(t)\}_{t\geq 0}$, and is considerably harder to analyze. Since the evolution of $\{Q(t)\}_{t\geq 0}$ is modulated by that of $\{X(t)\}_{t\geq 0}$, the former process can be viewed as a queueing network in a random environment. A simple necessary and sufficient condition for the all queue-length processes $\{Q_i(t)\}_{t\geq 0}$, $i = 1, \ldots, N$, to be positive recurrent is $\rho_i < \theta_i$, for all $i = 1, \ldots, N$, see Section 2.4.

If we assume that nodes refrain from competition for the medium when their buffer is empty, the interaction between the activity states $\{X(t)\}_{t\geq 0}$ and the buffer contents $\{Q(t)\}_{t\geq 0}$ becomes bidirectional and gives rise to even more complicated behavior. In fact, in this case the stationary distribution of the Markov process $\{(X(t), Q(t))\}_{t\geq 0}$ in general does not admit a closed-form expression and even the basic throughput characteristics and stability conditions are not known. An explicit description of the stability condition is very difficult for general network topologies, as illustrated in [136], and only structural representations or asymptotic results are known [32, 92].

Inspired by the work in [71], powerful algorithms have been proposed for adapting back-off parameters based on queue lengths [71, 69, 121, 128]. Under mild assumptions, these algorithms have been shown to achieve throughput optimality. The case of fixed activation and deactivation rates on which this thesis focuses is nevertheless highly relevant since it is a canonical model for gaining insight in the deep interplay between the structure of the network and its performance.

1.3 INTERACTING PARTICLE SYSTEMS

In this section we illustrate how the CSMA models described in Section 1.2 can be viewed as an interacting particle system.

1.3.1 *Hard-core model on finite graphs*

Concepts from statistical physics have often been applied to analyze large-scale distributed resource sharing mechanisms [43, 45, 50, 85, 105, 122, 140]. In the same spirit, the saturated CSMA model described in Subsection 1.2.2 can equivalently be described as an interacting particle system. More precisely, it corresponds to a certain stochastic model where particles dynamically interact in a finite volume subject to hard-core constraints. This model was introduced in the chemistry and physics literature under the name *hard-core lattice gas model*, and serves to describe the behavior of a gas whose particles have non-negligible radii and cannot overlap [57, 137].

We describe the spatial structure of the gas volume in terms of a finite undirected graph $G = (V, E)$, which is the physical counterpart of the wireless

conflict graph introduced in Section 1.2. The N vertices of the graph G represent the possible sites where particles can reside. The hard-core constraints are represented by edges connecting the pairs of sites where particles cannot be active simultaneously. The *state* of the system, i.e. the particle configuration on G, is described by a vector $x \in \{0,1\}^N$, where $x_i = 1$ if the particle in site i is active and $x_i = 0$ otherwise. We say that a state x is *admissible* if it does not violate the hard-core constraints, i.e. if $x_i x_j = 0$ for every $(i,j) \in E$. In other words, the state x is admissible if the subset $\{i \in V \ : \ x_i = 1\}$ of sites where the active particles reside in state x is an independent set of the graph G. The set $\mathcal{X}$ of admissible states thus corresponds to the collection of independent sets of G and is the same as in (1.1).

The potential activation epochs of an inactive particle in site i are determined by a random Poissonian clock that rings at rate ν_i. These N clocks are assumed to be independent of each other. If a particle is inactive and its random clock rings, it activates only if all its neighboring particles on G are currently inactive. A newly activated particle stays active for a random time, which we assume to be exponentially distributed with mean $1/\mu_i$. In this way, the global activity of particles on the graph G is described by a continuous-time Markov process $\{X(t)\}_{t\geq 0}$ with transition rates (1.2), which is ergodic and reversible with respect to the stationary distribution (1.3).

Denote the number of active particles in state $x \in \mathcal{X}$ by $\|x\| := \sum_{i=1}^N |x_i|$. In the case of homogeneous activation and deactivation rates, we denote $\nu = \nu_i/\mu_i$ and the stationary distribution (1.3) can be rewritten as

$$\pi_\nu(x) := \frac{\nu^{\|x\|}}{Z_\nu(G)}, \quad x \in \mathcal{X}, \tag{1.6}$$

where $Z_\nu(G)$ is the appropriate normalizing constant, also called *partition function*. The stationary distribution (1.6) is usually called *hard-core measure with activity* (or *fugacity*) ν.

In this thesis, we focus on the high-activity regime where $\nu \to \infty$, which is the mode in which the corresponding random-access network should operate to achieve a high throughput. Indeed, the hard-core measure (1.6) with a large activity ν favors states with a maximum number of active particles, to which we refer as *dominant states*. In the next subsection we will show how this regime corresponds to the low-temperature limit for the interacting particle system described as a Freidlin-Wentzell Markov chain.

1.3.2 *Uniformization and Metropolis Markov chains*

First hitting times will be one of the central objects of our study of the hard-core particle dynamics. There exists a large literature devoted to the asymptotic behavior of first hitting times for a wide range of models [35, 36, 39, 41, 42, 87, 111, 113], mostly to mathematically study the metastability phenomena they exhibit. It turns out to be convenient to study the hard-core dynam-

ics in discrete time, by looking at the uniformized version of Markov process $\{X(t)\}_{t\geq 0}$, since most of this literature uses discrete-time Markov chains.

In this way, after having identified $\nu = e^{\beta}$, we obtain a family of Markov chains $\{X_t^{\beta}\}_{t\in\mathbb{N}}$ on $\mathcal{X}$ parametrized by the *inverse temperature* $\beta > 0$ with transition probabilities

$$P_\beta(x,y) := \begin{cases} c(x,y)e^{-\beta[H(y)-H(x)]^+}, & \text{if } x \neq y, \\ 1 - \sum_{z\neq x} P_\beta(x,z), & \text{if } x = y, \end{cases} \tag{1.7}$$

where $H : \mathcal{X} \to \mathbb{R}$ is the *energy function* that counts the number of active particles, i.e.

$$H(x) := -\|x\|, \quad x \in \mathcal{X}, \tag{1.8}$$

and $c : \mathcal{X} \times \mathcal{X} \to [0,1]$ is the *connectivity function* that allows only single-site updates

$$c(x,y) := \begin{cases} \frac{1}{N}, & \text{if } \|x-y\| = 1, \\ 0, & \text{if } \|x-y\| > 1, \\ 1 - \sum_{z\neq x} c(x,z), & \text{if } x = y. \end{cases} \tag{1.9}$$

The Markov chain $\{X_t^{\beta}\}_{t\in\mathbb{N}}$ embedded at jump epochs of the original hard-core particle process is a Freidlin-Wentzell chain with Metropolis transition probabilities. This is a class of reversible Markov chains with transition probabilities of the form (1.7) uniquely characterized by an *energy landscape* $(\mathcal{X}, H, c)$ and by a positive parameter β, which represents the *inverse temperature*. Such a Markov chain is reversible with respect to the *Gibbs measure*

$$\pi_\beta(x) := \frac{e^{-\beta H(x)}}{\sum_{y\in\mathcal{X}} e^{-\beta H(y)}}, \quad x \in \mathcal{X}. \tag{1.10}$$

In this framework the dominant states are then the global minima of the energy function (1.8) and they are often called *stable states*. The spatial structure of the volume in which the particles interact, encoded in the graph $G = (V, E)$, plays a crucial role, since it determines the structure and the features of such an energy landscape $(\mathcal{X}, H, c)$.

1.3.3 *Hitting times between dominant states*

To understand the transient behavior of the hard-core model on a graph G, we study the *first hitting times* of the corresponding Metropolis Markov chain $\{X_t^{\beta}\}_{t\in\mathbb{N}}$ in the low-temperature limit $\beta \to \infty$. Given a target subset of states $A \subset \mathcal{X}$ and an initial state $x \in \mathcal{X} \setminus A$, the *first hitting time* τ_A^x of the subset A

for the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ with initial state x at time $t = 0$ is defined as the random variable

$$\tau_A^x := \inf\Big\{t > 0 \ : \ X_t^\beta \in A, \mid X_0^\beta = x\Big\}.$$

We focus on the asymptotic behavior of the hitting times between dominant states, which tell us how rigid the dominant states are and how long it takes for the particle system to "switch" between them. For example, the hard-core model on rectangular grids has two dominant states, **e** and **o**, which correspond to full particle activity on the even chessboard and the odd chessboard (see Figures 1.1a and 1.1b, respectively). We are interested in the asymptotic behavior of the hitting time $\tau_{\mathbf{o}}^{\mathbf{e}}$.

In this thesis, we follow the *pathwise approach*, which has been used to study the metastability problem for many finite-volume models in a low-temperature regime, see [35, 36, 39, 41, 42, 87, 111, 113]. The crucial idea behind this method is to understand which paths the Markov chain most likely follows in the low-temperature regime. The characterization of these most likely paths then leads to asymptotic results for hitting times. For Freidlin-Wentzell Markov chains this can be done by analyzing the energy landscape to find the paths between the initial and the target states with a minimum energy barrier. In the case of the hitting time between dominant states of the hard-core model, this problem reduces to identifying the most efficient way, starting from a dominant state, to activate particles of the target dominant state.

Consider first the case of a complete K-partite conflict graph, which is a graph consisting of K components where particles are prevented from simultaneous activity if and only if they belong to different components. If the spatial structure is described by such a complete partite graph, it is clear that the dominant states correspond to its largest components and that any trajectory between them needs to pass through the state with no active particles. Indeed a particle can become active only when all particles in the other components are inactive. Hence, in this case the orders of magnitude of the hitting times are related to the sizes of the largest partitions of the graph. Figure 1.3 illustrates the state space $\mathcal{X}$ corresponding to the hard-core model evolving on the complete bipartite graph $K_{8,8}$, which has 8 sites in each of the two components. The leftmost (rightmost) vertex corresponds to the network state

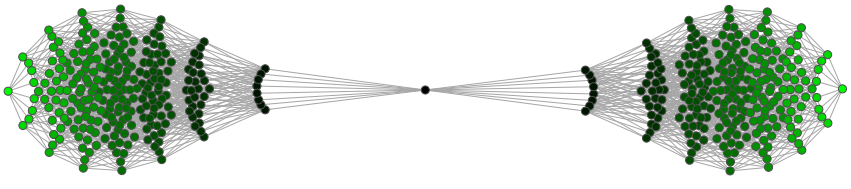


Figure 1.3: State space for the hard-core dynamics on the complete bipartite graph $K_{8,8}$

with full activity in the first component (second component). The central vertex corresponds to the network state where there are no active particles and is clearly a bottleneck for the transition between the two dominant states.

For non-complete partite graphs, such as grid graphs or triangular grid graphs, the various partitions of the graph still correspond to the dominant states, but the transitions between them do not necessarily require to visit a state with no active particles. In fact, there are paths that visit mixed-activity states, where a layer of inactive particles separate clusters of active particles belonging to different partitions. Figure 1.4 illustrates the state space $\mathcal{X}$ corresponding to the hard-core model evolving on a 4×4 grid with open boundary conditions, similar to the one displayed in Figure 1.1. The leftmost state corresponds to the dominant state with particles active on the even chessboard, while the rightmost state corresponds to the state with all active particles on the odd chessboard. The crucial difference with the complete partite case is that here the bottleneck is not a single state and is much harder to identify.

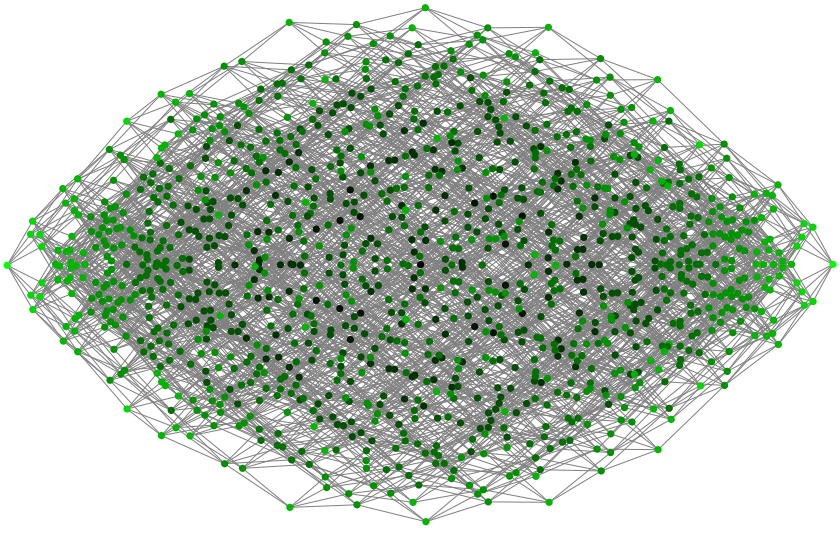
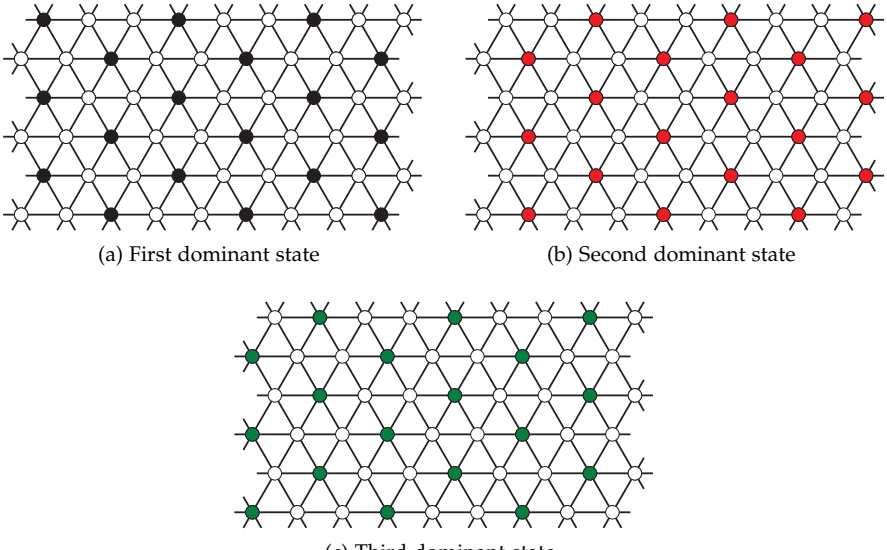


Figure 1.4: State space for the hard-core dynamics on the 4×4 grid graph with open boundary conditions

Another example of non-complete partite graphs are triangular grid graphs, which have three dominant states. Figure 1.5 displays these three dominant states on a 6×9 triangular grid network. The three dominant states are clearly distinguishable as the three "corners" of the state space $\mathcal{X}$ corresponding to the hard-core model on the 4×6 triangular grid graph with periodic boundary conditions displayed in Figure 1.6.



(a) First dominant state

(b) Second dominant state

(c) Third dominant state

Figure 1.5: The three dominant activity states on a 6×9 triangular grid network

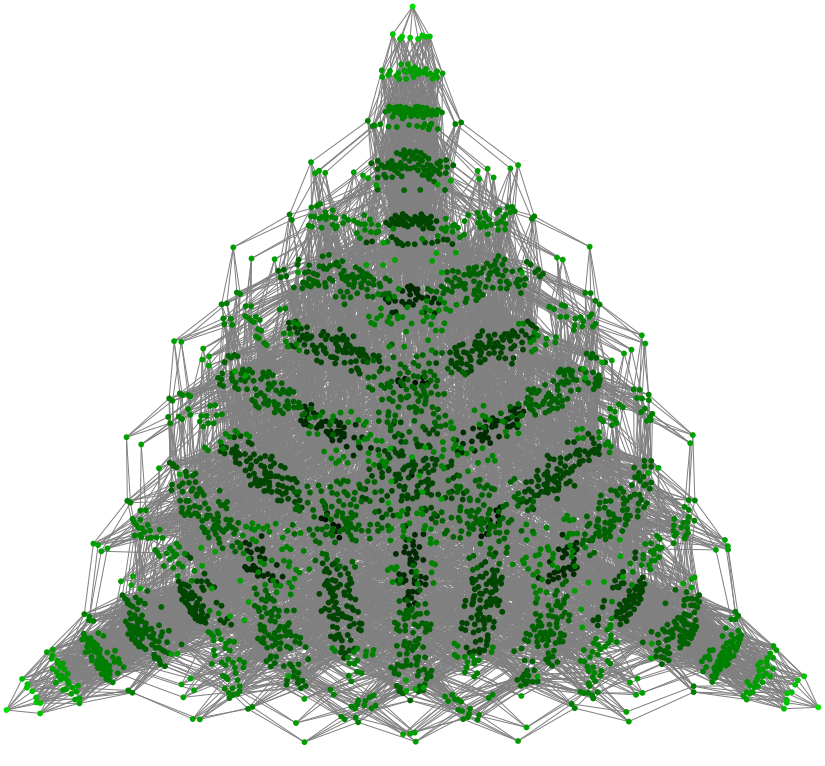


Figure 1.6: State space for the hard-core dynamics on the 4×6 triangular grid graph with periodic boundary conditions

For all these non-complete partite graphs, by exploring detailed geometric properties of these mixed-activity states, we develop a novel combinatorial method to quantify their "inefficiency" and obtain in this way the minimum energy barrier $\Gamma(G)$ that has to be overcome in the energy landscape (corresponding to the spatial structure encoded by the graph G) for the required transition to occur. The energy barrier $\Gamma(G)$ is precisely the quantity that appears in the asymptotic results for the transition times between dominant states and, in particular, identifies the timescale $e^{\beta\Gamma(G)}$ at which transactions most likely occur. The energy barrier $\Gamma(G)$ is a positive number that in many cases grows with the size of the conflict graph G, which means that the transition between dominant states becomes very slow in the regime where the activation rate ν (or equivalently the inverse temperature β) becomes large. As argued before, the stationary distribution (1.10) concentrates in the dominant states of G as $\beta \to \infty$. Hence in this low-temperature regime, for a conflict graph G with two or more dominant states, the timescale at which the transitions between dominant states occur is intimately related to the order of magnitude of the *mixing time* of the hard-core dynamics on G. Indeed, the mixing time of a Markov chain quantifies the time it takes for this stochastic process to become "close" to its steady-state distribution. Using structural properties of the corresponding energy landscapes and classical results [29, 109] for Friendlin-Wentzell Markov chains, we make this intuition rigorous and show that the hard-core dynamics exhibit indeed *slow mixing* on the highly symmetric conflict graphs considered in this thesis.

1.4 NETWORK DELAY PERFORMANCE

In the previous section we have illustrated how the random-access CSMA networks presented in Subsection 1.2.2 can be viewed as interacting particle systems with hard-core interaction. This connection allows us to study the asymptotic behavior of first hitting times between dominant states using the convenient framework of Metropolis Markov chains. In particular, we expect long transition times between dominant states (and thus slow mixing, if there are more than two dominant states) in the regime where the nodes activate more aggressively. In this section we take this analysis one step further, studying the impact of such long transition times and slow mixing on temporal starvation. In this way we show how the average long-run packet delay grows in high-load scenarios and demonstrate how its order of magnitude depends on the network structure.

Before illustrating our approach, we briefly review the literature on the performance of CSMA algorithms and the deep connections with other related areas. Several different models have been introduced in the literature to study the performance of CSMA-like algorithms; see [141] for a survey. Despite their asynchronous and distributed nature, these algorithms have been shown to be *throughput-optimal*, which means that they offer the capability of achiev-

ing the full capacity region and thus match the optimal throughput performance of centralized scheduling mechanisms operating in slotted time (such as MaxWeight), see for instance [71, 69, 102]. In addition to this "first-order" metric, the delay performance of CSMA networks has been studied in the literature as well. Simulation results indicate that the delay performance of these algorithms can be rather poor and much worse than for other mechanisms, such as MaxWeight. This fact has triggered a strong interest in investigating the causes of such poor delay performance and in developing approaches to mitigate them, see [58, 68, 89, 90, 103, 115, 127].

This research direction is intimately related with the problem of designing efficient randomized algorithms to sample or approximate combinatorial quantities on graphs. Indeed, the random-access algorithms described in the previous sections can also be interpreted as randomized schemes with local updates to find maximum independent sets, which is an NP-hard problem. In particular, the hard-core particle dynamics on a finite graph $G = (V, E)$ described in Subsection 1.3.1 can be used to sample weighted independent sets, with weight $\nu^{|I|}$ for $I \in \mathcal{I}(G)$. Taking for instance the activity rate equal to 1, we are sampling uniformly independent sets of G, since the stationary distribution (1.6) is uniform over $\mathcal{I}(G)$. Hence, when one is interested in finding or sampling maximum independent sets, one could consider a large activity rate ν, which increases their likelihood in view of (1.6). However, the resulting Markov process can then take a very long time to converge to stationarity, exhibiting so-called *slow/torpid mixing*. Several papers [19, 47, 51, 53, 54, 61, 104, 108, 123] investigate (mostly for Glauber dynamics in discrete time) how the mixing times of these processes scale with the size N of the graph, depending on the type of graph and/or its maximum vertex degree.

It is natural to expect slow mixing of the hard-core dynamics for a large activity rate ν, since fast mixing in this regime would imply that the NP-hard problem of finding maximum independent sets could be solved or approximated efficiently. By exploiting this relationship between random-access algorithms and independent sets, it has been shown in [129] that low-complexity schemes (such as CSMA) cannot be expected to achieve low delay in arbitrary topologies (unless P equals NP).

For CSMA networks lower bounds for the average steady-state delay were established in [20] using mixing-time results, showing that delays can grow dramatically with the load of the system. Upper bounds for the average steady-state delay based on mixing time results for Glauber dynamics were derived in [72, 68], but they are valid only for sufficiently low loads. Furthermore, it was shown in [132] that mixing-time based approaches may not be the right way to capture delay dynamics, even in an asymptotic sense. In [98] it was shown by using stochastic majorization techniques that CSMA algorithms that reduce asymptotic variance have better delay performance.

In this thesis we follow a different approach to estimate the average steady-state delay in the scenario where the whole network is stable, that is when

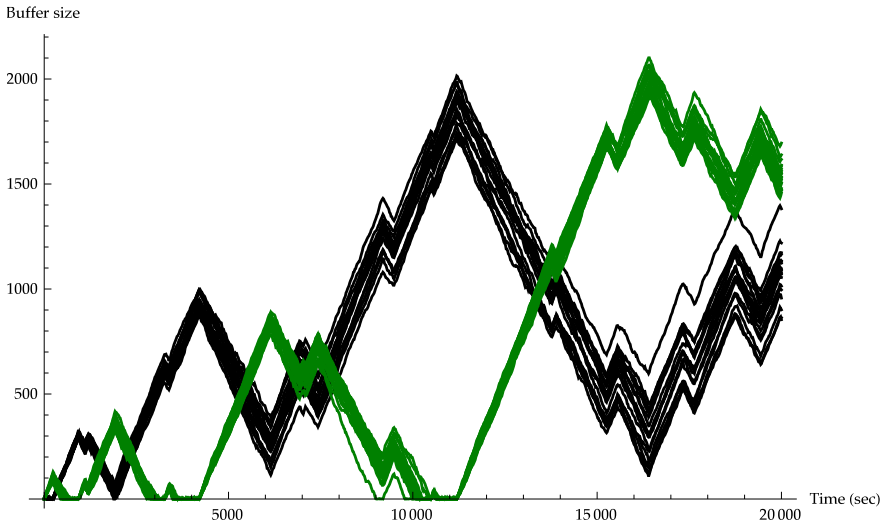


Figure 1.7: Simulated buffer content evolution over time for a 6×6 grid network with homogeneous rates $\lambda = 0.45$, $\nu = 5$ and $\mu = 1$

the queue-length process at each node is positive recurrent. In particular, we investigate how it is related to temporal starvation, which for the graphs we consider in this thesis is the root cause for the poor delay performance. Indeed, intuitively (i) high load requires high activation rates, (ii) high activation rates imply extremely slow transitions between dominant activity states, and (iii) slow transitions cause starvation and hence excessively long queues and delays. Therefore, we estimate the average steady-state delay by studying the asymptotic behavior of transition times between dominant states of the Markov process $\{X(t)\}_{t\geq 0}$ in high-load scenarios. Figure 1.7 shows the simulation of the evolution over time of the buffer contents for a homogeneous network of 36 nodes placed on a 6×6 grid (with periodic boundary conditions). This simulation corroborates the intuition that the buffer contents of the odd nodes build up packets while the network is active mostly in the even nodes and vice versa.

In a high-load scenario the average steady-state queue length is then proportional to the average transition time between the two dominant states of the grid network. These arguments are valid for other symmetric networks with two or more dominant states as well, as illustrated by Figure 1.8 in the case of a triangular grid graph, which has three dominant states.

By making these arguments rigorous, we will derive lower bounds for the average stationary queue length in terms of the average transition time between dominant states. By virtue of Little's law, these also provide lower bounds for the expected stationary delay. Furthermore, these delay lower bounds become explicit for many symmetric partite networks of interest, since we can calculate the asymptotic order of magnitude of the average transition

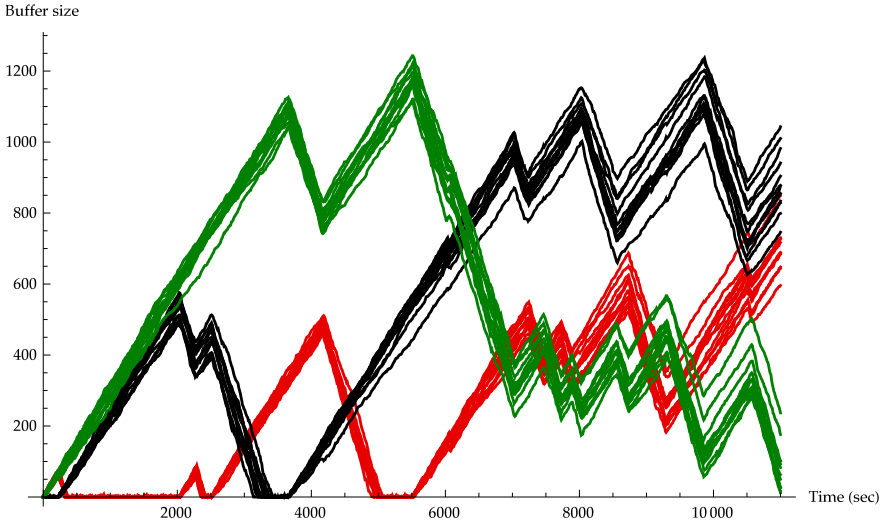


Figure 1.8: Simulated buffer contents evolution over time for a 4×6 triangular grid network with homogeneous rates $\lambda = 0.3$, $\nu = 15$ and $\mu = 1$

times between their dominant states. Suppose that for a given conflict graph G we can prove, as outlined in Subsection 1.3.3, that the average transition time between dominant states is of the order $e^{\beta\Gamma(G)}$ as $\beta \to \infty$ (that is the regime where the activation rate ν grows large). Then such a result for hitting times can be translated into an asymptotic lower bound for the long-term average delay $\mathbb{E}W$ (waiting time plus service time) at every node. More specifically, we show that $\mathbb{E}W$ scales asymptotically at least as fast as

$$\left(\frac{1}{1-\rho}\right)^{\Gamma(G)}, \quad \text{as } \rho \uparrow 1. \tag{1.11}$$

This delay scaling is to be contrasted with the usual linear scaling in $\frac{1}{1-\rho}$ in conventional queueing systems at high load. Specifically, in the case of the complete K-partite graph G on N nodes with all components of the same size N/K (assuming that N multiple of K) we obtain that

$$\Gamma(G) = \frac{N}{K}.$$

Such complete partite graphs are arguably the worst possible networks in terms of transition times and starvation effects, given the size of the network and the size of the components. Indeed, the fact that the nodes are grouped into components, with no interference within components and full interference between components, turns out to be a key element for starvation to occur. This is reflected in the fact that the transition times exhibit exponential growth in the component size and that the starvation effects are very pronounced.

Graphs that are partite but not complete have a less extreme tendency for starvation, although the issue may still arise. The lack of interference between some nodes in different components results in bottleneck states where some of the nodes may be active, and increases the likelihood of the bottleneck state relative to the dominant activity states. This is the case for the grid conflict graphs analyzed in Chapters 5. The results therein show that the transition times, while still severe, are of a lower order than for the complete partite graphs of the same size. More precisely, for a square grid network G with with N nodes, with conflict graph the $\sqrt{N} \times \sqrt{N}$ grid (with $\sqrt{N}$ an even integer),

$$\Gamma(G) = \sqrt{N} \quad \text{or} \quad \Gamma(G) = \frac{\sqrt{N}}{2},$$

depending on whether one considers periodic boundary conditions or not, respectively. In fact, for all the extreme network structures studied in this thesis the exponent $\Gamma(G)$ increases with the network size. This observation, together with (1.11), shows that the average delay can grow dramatically with the dimension of the network.

In the case of complete partite conflict graphs, the connection between transition times and starvation is even stronger and yields an asymptotic lower bound for the probability of starvation and a complementary result which indicates over what time scales near-saturation of throughput occurs.

1.5 MULTI-CHANNEL CSMA NETWORKS

The CSMA models presented in Subsections 1.2.2 and 1.2.3 describe networks with nodes that all compete for the same wireless channel. As noted earlier, the total available wireless spectrum could be divided into C orthogonal channels on which nearby nodes can transmit without interfering with each other. CSMA-like algorithms have been been proposed as efficient distributed schemes for these multi-channel networks and various models have been studied in the literature to investigate their performance [6, 91, 101, 120].

1.5.1 *Model description and stationary distribution*

We consider an extension of the saturated CSMA model where each node can sense the interference and transmit on any of the C available channels. We assume that for every $c = 1, \ldots, C$ all possible conflicts between nodes on channel c are described by a conflict graph $G_c = (V, E_c)$. Node i in the network has a different back-off timer for each of the C available channels and we model these timers as C independent Poissonian clocks, ticking at rates $\nu_{i,1}, \ldots, \nu_{i,C}$. When the first of these C clocks rings for an inactive node, it activates on the corresponding channel, say c, if and only if the neighboring nodes of i in G_c are not active on the same channel. The transmission times of

node i on channel c are independent and exponentially distributed with mean $1/\mu_{i,c}$.

A network activity state is described by a vector $x \in \{0,1,\dots,C\}^N$ where $x_i = 0$ if node i is inactive and $x_i = c$ if node i is active on channel c with $1 \le c \le C$. Let $\mathcal{X}$ be the collection of feasible network activity states, which consists of all the vectors $x \in \{0,1,\dots,C\}^N$ such that for every $c = 1,\dots,N$ two neighboring nodes in G_c are not simultaneously active on channel c. The activity state of the network at time t is described by the vector $X(t)$, where the i-th entry $X_i(t)$ indicates the activity state of node i. Then $\{X(t)\}_{t\ge 0}$ is a Markov process on the state space $\mathcal{X}$ with transition rates

$$q(x,y) = \begin{cases} \nu_{i,c} & \text{if } y = x + c\cdot e_i \in \mathcal{X}, \\ \mu_{i,c} & \text{if } y = x - c\cdot e_i \in \mathcal{X}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.12}$$

where $e_i \in \{0,1\}^N$ is the vector with all zeros except for a 1 at position i.

We now argue that the evolution of a network using a multi-channel CSMA algorithm can be equivalently described as a *virtual network* operating under the single-channel CSMA algorithm on a modified conflict graph G^*. Consider the undirected graph $G^* = (V^*, E^*)$ with vertex set $V^* := V \times \{1,\dots,C\}$ where two vertices (i,c) and (i',c') are adjacent if and only if

$$\begin{cases} c \ne c', \\ i = i', \end{cases} \qquad \text{or} \qquad \begin{cases} c = c', \\ (i,i') \in E_c. \end{cases}$$

To avoid confusion with the original nodes, we refer to the nodes of G^* as *virtual nodes*. The activity state of the virtual network is then represented by a 0-1 vector of length $C\cdot N$ and we denote by $\mathcal{X}^* \subset \{0,1\}^{C\cdot N}$ the set of admissible virtual activity states on G^*. Suppose further that the $C\cdot N$ virtual nodes operate using the single-channel CSMA algorithm described in Subsection 1.2.2, assuming that the virtual node (i,c) has back-off periods that are exponentially distributed with mean $1/\nu_{i,c}$ and transmission periods that are exponentially distributed with mean $1/\mu_{i,c}$.

The interference constraints in the virtual network are such that the collection $\mathcal{X}^*$ of admissible activity states on the virtual network G^* is in one-to-one correspondence with the collection $\mathcal{X}$ of admissible multi-channel activity states on G. Furthermore, the transition rates according to which the virtual nodes evolve are chosen in such a way that the evolution of the two networks is identical in law. Hence, we deduce that the multi-channel $\{X(t)\}_{t\ge 0}$ is also a reversible Markov process and that it has stationary distribution

$$\pi(x) = Z^{-1} \prod_{i=1}^{N} \prod_{c=1}^{C} \left(\frac{\nu_{i,c}}{\mu_{i,c}}\right)^{\mathbb{1}_{\{x_i = c\}}}, \quad x \in \mathcal{X}, \tag{1.13}$$

where Z is the appropriate normalizing constant. Furthermore, the insensitivity property carries over to (1.13), which thus holds for general back-off and transmission time distributions.

We remark that the fact that the multi-channel CSMA dynamics have a product-form distribution was already shown in [120] in the special case where the interference is described by the same graph on every channel, namely $E_1 = \cdots = E_C$. Our proof approach not only recovers the same result in this special case, but proves the insensitivity property with respect to the back-off and transmission time distributions and shows how a multi-channel CSMA network can be studied as single-channel CSMA network.

1.5.2 *Dominant states and performance trade-offs*

We now provide some insight as to how the number of channels C affects the performance in terms of aggregate throughput. While increasing the number of channels evidently provides greater transmission opportunities, the net impact on throughout performance is non-obvious since the transmission capacity per channel is inversely proportional to the number of channels.

Consider a multi-channel CSMA network where the same conflict graph $G = (V, E)$ describes the interference on all C channels and assume that each of these channels has capacity $1/C$. Assume further that all the nodes have homogeneous activation and transmission rates, namely

$$\nu_{i,c} \equiv \nu \quad \text{and} \quad \mu_{i,c} \equiv \mu \quad \forall\, c = 1, \dots, C, \quad \forall\, i = 1, \dots, N. \tag{1.14}$$

The stationary distribution of the activity process $\{X(t)\}_{t\geq 0}$ then reads

$$\pi(x) = Z^{-1} \prod_{i=1}^{N} \sigma^{\mathbb{1}_{\{x_i \neq 0\}}} = Z^{-1} \sigma^{\sum_{i=1}^{N} \mathbb{1}_{\{x_i \neq 0\}}}, \quad x \in \mathcal{X}. \tag{1.15}$$

We analyze the *aggregate throughput* $\Theta(C)$ of this network in the asymptotic regime where the activity factor $\sigma := \frac{\nu}{\mu}$ grows large. The aggregate throughput when there are C available channels is defined as

$$\Theta(C) := \lim_{\sigma \to \infty} \sum_{i \in V} \frac{1}{C} \sum_{x \in \mathcal{X}} \pi(x) \mathbb{1}_{\{x_i \neq 0\}},$$

where π is the stationary distribution (1.15) of the multi-channel activity process $\{X(t)\}_{t\geq 0}$ with transition rates (1.14).

As in the single-channel scenario, this random-access network should also operate with a large activity factor σ to achieve a high throughput. In this regime, the stationary distribution (1.15) favors the activity states with the largest number of active nodes, regardless of the channels they are active on. We refer to such activity states as *dominant states* also in this multi-channel context.

In order to characterize the dominant states of a multi-channel CSMA network, we need some further definitions. A *vertex coloring* of the graph G is a labelling of the graph's vertices with colors such that no two vertices sharing the same edge have the same color. A *(proper) q-coloring* is a vertex coloring using at most q colors, see e.g. [15, 25]. The smallest number of colors needed to color a graph G is called its *chromatic number*, and is often denoted as $\chi(G)$. Define $d(C)$ to be the maximum number of active nodes in conflict graph G when C channels are available, i.e.

$$d(C) := \max_{x \in \mathcal{X}} \sum_{i \in V} \mathbb{1}_{\{x_i \neq 0\}}.$$

If $C = 1$, we are back in the single-channel scenario and clearly $d(1) = \alpha(G)$, where $\alpha(G)$ denotes the *independence number* of the graph G, that is the cardinality of its maximum independent set. For $C \geq 2$, the multi-channel activity on G is equivalent to the single-channel activity on the virtual network G^* introduced in Subsection 1.5.1. In the scenario considered in this subsection, the virtual network G^* is the Cartesian product of G and the complete graph on C vertices and, in particular, $d(C)$ is equal to the independence number $\alpha(G^*)$ of the graph G^*, as proved in [15, Lemma 1].

If the number of available channels in a CSMA network is larger than or equal to the chromatic number of the corresponding conflict graph, i.e. $C \geq \chi(G)$, then every proper C-coloring of the graph G corresponds to a dominant state for the network dynamics and, in particular, $d(C) = N$. If instead $C < \chi(G)$, by definition of the chromatic number, a proper C-coloring of the graph G does not exist and, in particular, there are no admissible activity states where all nodes are active. All admissible activity states correspond to *partial C-colorings* of the graph G. However, the problem of finding the maximum number of nodes that can be active simultaneously in a general conflict graph G with C available channels is non-trivial, since it is at least as hard as finding the maximum independent sets of G^*. Nonetheless, some conclusions for the aggregate throughput $\Theta(C)$ of a network using C channels can be drawn.

Let us start by noticing that the function $d(C)/C$ is a non-increasing function of the number of channels C, i.e.

$$\frac{d(C+1)}{C+1} \leq \frac{d(C)}{C}, \quad \forall\, C \geq 1, \tag{1.16}$$

and that $d(C) = C \cdot \alpha(G)$ for every $C \leq C^*(G)$, where $C^*(G)$ represents the number of disjoint maximum independent sets of the graph G, and $d(C) = N$ for every $C \geq \chi(G)$.

Inequality (1.16) trivially holds when $C \geq \chi(G)$, since $d(C) = N$, and the proof in the case $C \leq \chi(G)$ readily follows by considering a $(C+1)$-coloring that yields $d(C+1)$, pick the color that is used the least and discolor the corresponding nodes, which are at most $\lfloor \frac{d(C+1)}{C+1} \rfloor$, obtaining a C-coloring with $d(C+1) - \lfloor \frac{d(C+1)}{C+1} \rfloor \geq \frac{C}{C+1} d(C+1)$ colored nodes. The second statement can be easily proved by induction on C.

Denote by $\mathcal{D}$ the collection of the dominant states of the network G. Inequality (1.16) implies that the aggregate throughput $\Theta(C)$ is a non-increasing function of the number of channels C as well, since the definition of dominant state and (1.15) yield that $\sum_{x\in\mathcal{D}} \pi(x) \to 1$ as $\sigma \to \infty$, and thus

$$\begin{aligned}\Theta(C) &= \lim_{\sigma\to\infty} \sum_{i\in V} \frac{1}{C} \sum_{x\in\mathcal{X}} \pi(x) \mathbb{1}_{\{x_i\neq 0\}} = \frac{1}{C} \lim_{\sigma\to\infty} \sum_{x\in\mathcal{X}} \pi(x) \sum_{i\in V} \mathbb{1}_{\{x_i\neq 0\}} \\ &= \frac{d(C)}{C} \lim_{\sigma\to\infty} \sum_{x\in\mathcal{D}} \pi(x) = \frac{d(C)}{C}.\end{aligned}$$

Furthermore, $\Theta(C)$ is constant and equal to $\alpha(G)$ for C up to $C^*(G)$, where $C^*(G)$ represents the number of disjoint maximum independent sets of the graph G, and becomes a strictly decreasing function for $C \geq \chi(G)$, since

$$\Theta(C) = \frac{N}{C}, \quad \forall\, C \geq \chi(G).$$

From these considerations, it seems that the network should be operated using a single channel, but this is also the scenario where the temporal starvation is more pronounced. We expect that by exploiting more channels the temporal starvation effects can be mitigated, so that the packet delay should also be a non-increasing function of the number C of available channels. However, making this dependence more explicit is a challenging task, since the structure of the conflict graph also plays a crucial role. Indeed, in some conflict graphs the packet delay is already $O(1)$ for just a single channel, e.g. complete graphs, while in others it only becomes $O(1)$ when the number of channels matches the chromatic number, e.g. complete K-partite graphs. Thus, in general, a non-trivial trade-off emerges between aggregate throughput and packet delay that the network can achieve and a balance can be found using a number of channels in the range between 1 and $\chi(G)$.

The focus on this thesis is on single-channel scenarios, but our framework can be applied for the analysis of transition times in multi-channel networks in view of the equivalence described in Subsection 1.5.1. Furthermore, multi-channel networks can in some cases be described directly as interacting particle systems with two or more types of particles subject to repelling forces, as we will show in Chapter 7. However, a comprehensive translation of the results for packet delay in this multi-channel scenario is beyond the scope of this thesis.

In practical deployments, the number of channels C is normally given, and *channel bonding* may then be an interesting strategy to improve the network performance (e.g. if C lies outside the indicated range). The key idea behind this mechanism is that more channels can be aggregated to obtain wider channels. We show in [10] that a modified version of the single-channel CSMA model described in Subsection 1.2.2 provides a good theoretical framework for this different random-access algorithm and we investigate numerically this non-intuitive trade-off between using wider channels to increase the throughput and higher chance of interference.

1.6 FINAL CONSIDERATIONS AND OPEN PROBLEMS

In this thesis we build a versatile framework to investigate delay performance of random-access networks, which combines techniques and concepts from applied probability and statistical physics. More specifically, we describe conflicting transmissions in a random-access network as an interacting particle system with hard-core constraints and investigate temporal starvation phenomena by analyzing hitting times of suitable Markov chains.

It has been observed in many simulations that CSMA algorithms can induce severe temporal starvation effects [74, 91] and long packet delays [59, 66, 89, 90, 103]. In this thesis we derive analytical results for average packet delay performance and, in particular, show how its asymptotic order of magnitude depends on the network structure and size in high-load scenarios, as briefly illustrated in Section 1.4.

In the model under consideration potentially conflicting transmissions are modeled by means of a conflict graph, which is a common paradigm in the literature, see e.g. [18, 55, 68, 67, 70, 103, 121], and has in particular been used and validated for estimation of stationary throughputs [100]. Various methods have been proposed to construct conflict graphs that represent interference constraints more accurately and effectively, see e.g. [63, 143].

In this work we choose to focus on networks with a symmetric structure, in order to remove all disparities between nodes, which in particular will then have the same long-term stationary throughput. This choice allows us to investigate solely transient starvation effects and, at the same time, obtain more explicit analytical results. We expect our results to extend to more general topologies as well, but heterogeneity makes temporal starvation effects more case-specific and thus harder to treat mathematically in full generality.

Several solutions have been proposed in the literature to mitigate temporal starvation effects, for instance introducing a periodical reset of all ongoing transmissions, as proposed by [103], but the practical implementation is not straightforward. Another possible solution could be the usage of multiple frequencies in [91], as briefly illustrated in Section 1.5 and future work could investigate delay performance and possibly extend our results in this multi-channel scenario.

Our analytical results for hitting times and average delay are asymptotic in nature, either when the normalized load approaches unity or the activation rate grows large. There is numerical evidence that long starvation periods occur already for relatively small values of these parameters, see for instance Figures 1.7 and 1.8. It would be interesting to characterize how large the system parameters should be so that the network performance can be well-approximated by means of our asymptotic results. Another intriguing yet challenging research direction would be investigate whether our results for transition times and delays extend when the durations of back-off and transmission periods have non-exponential distributions or when the dummy packet assumption is dropped (see Subsection 1.2.3).

From a methodological standpoint, we expect that the framework we developed to analyze hitting times can be extended to more general probability kernels than the Metropolis one, see (1.7), and possibly to irreversible Markov chains, similarly to what has been done in the context of metastability in [38]. Such a more general framework would be very helpful for instance to study the behavior of CSMA protocol using dynamic channel bonding, which can be modeled by means of irreversible Markov chains [48].

To complete the study of the hard-core model on grid and triangular grid graphs from a statistical physics standpoint, the next step would be to understand the geometric features of the critical configurations visited along the transition between dominant states and describe the tube of typical paths. Lastly, the ideas behind our combinatorial approach can probably be adopted to understand transition times between maximum activity states on other graphs with symmetric structures, e.g. hexagonal or Kagome lattices.

1.7 OUTLINE OF THE THESIS

This thesis is organized as follows. We examine the impact of slow transitions between dominant states on the delay performance in random-access networks in Chapter 2. Focusing on highly symmetric networks, among which regular meshes like grid and triangular grid graphs, we show how delays in symmetric partite graphs are intimately connected with the transition times between dominant states in a high-load scenario. In particular, we prove an asymptotic lower bound for the average steady-state delay for such networks. Chapter 2 expands the ideas first published in [149].

In Chapter 3 we study hitting times and mixing properties of *dense* random-access networks, which we model by complete partite graphs, focusing on the regime where the heterogeneous activity rates become large. In particular, we derive asymptotic results for the transition time between any pair of network activity states, both in expectation and in distribution (after scaling), and show how they depend on the activity rates and network structure. In the special case of homogeneous activity rates, we show that there are only two possible asymptotic distributions for such scaled transition times, one of them being the exponential distribution, and derive the asymptotic order of magnitude of the mixing time. Lastly, we identify the time scales at which full throughput starvation and near-saturation occur for each of the components of the network. Chapter 3 is based on [147, 148].

In Chapter 4 we introduce the framework of Metropolis Markov chains, in which the uniformized version of the saturated CSMA model can be put. We extend the model-independent framework in [106] for first hitting times τ_A^x to allow for a more general initial state x and target subset A. More specifically, we show how to derive asymptotic bounds in probability for such hitting times by analyzing the most likely paths that the process follows in the low-temperature regime. Furthermore, we give a sufficient condition that guaran-

tees that the scaled hitting times are exponentially distributed. The content of this chapter is based on Sections 3 and 4 of [112].

In Chapter 5 we apply these model-independent results to the hard-core model on rectangular grid graphs to understand the asymptotic behavior of the transition time between the two dominant states, where the active particles are arranged in a checkerboard fashion on even and odd sites, respectively. Using a novel combinatorial method, we identify the minimum energy barrier between these two dominant states and obtain in this way sharp bounds in probability for such a transition time and find its order of magnitude on a logarithmic scale. In addition, our analysis of the energy landscape shows that the scaled transition time is exponentially distributed in the low-temperature regime and yields the order of magnitude of the mixing time of the Markov chain. Chapter 5 is based on [112].

In Chapter 6 we turn to the hard-core model on triangular grid graphs, which has three dominant states. By identifying the minimum energy barrier between any pair of dominant states and by proving the absence of deep cycles in the corresponding energy landscape, we derive several results for the transition times between them in the low-temperature regime. The content of this chapter is based on [145].

Lastly, in Chapter 7 we study the Widom-Rowlison model with two types of particles, which we show to be the discrete-time counterpart of a multi-channel CSMA network when there are two available channels and the conflict graph is bipartite. This interacting particle system can also be put in the framework of Metropolis Markov chain presented in Chapter 4. Focusing on grid graphs, we study the asymptotic behavior of the transition time between its two dominant states in the low-temperature regime. In the analysis of the corresponding energy landscape, we exploit a combinatorial method similar to that developed in Chapter 5 for the hard-core model, to which the Widom-Rowlison model is intimately related. This chapter is based on [146].

2 TEMPORAL STARVATION AND DELAYS

In this chapter we investigate the interplay between temporal starvation and delays in networks operating under the single-channel CSMA algorithm described in Section 1.2. In particular, we derive lower bounds for the long-term average delay in random-access networks. Our proof approach revolves around three simple observations:

(i) high load requires high activity factors for stability;

(ii) high activity factors cause slow transitions between dominant activity states;

(iii) slow transitions between dominant states imply long starvation periods, and hence large queue-lengths and delays.

In Section 2.1 we first review the model, first introduced in Subsection 1.2.3, and explain in detail how the buffer dynamics depend on the underlying activity process. Our analysis focuses on a specific class of highly symmetric networks for which explicit results can be obtained. The features of the corresponding conflict graphs are presented in Section 2.2 together with some relevant examples. In Section 2.3 we derive some key properties for transition times between the dominant states for these networks, while the network stability conditions are established in Section 2.4. In Section 2.5 we derive lower bounds for the average stationary delay in terms of the expected transition times between dominant states, which we make explicit in Section 2.6 for some specific networks of interest.

2.1 UNSATURATED CSMA MODEL DYNAMICS

In this section we revisit the unsaturated CSMA model described in Subsection 1.2.3 and introduce some further assumptions.

The network structure is described in terms of an undirected graph $G = (V, E)$ of N nodes, where the set of edges $E \subset V \times V$ indicates which pairs of nodes are prevented from simultaneous activity. We denote by $\mathcal{X}$ the set of admissible network states on G, as defined in (1.1).

Packets arrive at node i according to a Poisson process of rate λ_i, independently of the other nodes and of the state of the network. Upon arrival packets are stored in a buffer at the corresponding node until they eventually get transmitted. The nodes share the medium in a decentralized fashion according to a CSMA algorithm, similar to that described in Subsections 1.2.2 and 1.2.3. More specifically, we assume that, when inactive and unblocked,

node i may start transmitting at the instants of a Poisson process of intensity ν_i and that the packet transmission times at node i are independent of each other and of the arrival process, and are exponentially distributed with mean $1/\mu_i$. The only difference with the model presented earlier is that nodes do not necessarily release the wireless medium when they complete a packet transmission. Indeed, we assume that node i either releases the medium and starts a back-off period with probability p_i before the next packet transmission, or immediately starts the next packet transmission otherwise. The dynamics described in Subsection 1.2.2 correspond to the case $p_i = 1$ for $i = 1, \ldots, N$.

Recall that $X(t) \in \mathcal{X}$ represents the joint activity state of the network at time t, with $X_i(t)$ indicating whether node i is active at time t or not, and $\{Q(t)\}_{t\geq 0}$ is the joint queue-length process, with $Q_i(t)$ the number of packets waiting at node i at time t (excluding the packet that may be in the process of being transmitted). Under the CSMA algorithm described above, the process $\{(X(t), Q(t))\}_{t\geq 0}$ evolves as a continuous-time Markov process with state space $\mathcal{X} \times \mathbb{N}^N$. Transitions due to arrivals from a state (X, Q) to $(X, Q + e_i)$ occur at rate λ_i, transitions due to activations from a state (X, Q) with $X_i = 0$ and $X_j = 0$ for all neighbors j of node i to $(X + e_i, Q - e_i \mathbb{1}_{\{Q_i>0\}})$ occur at rate ν_i, transitions due to transmission completions followed by a back-off period from a state (X, Q) with $X_i = 1$ to $(X - e_i, Q)$ occur at rate $p_i\mu_i$, and transitions due to transmission completions immediately followed by another transmission (without an intermediate back-off period) from a state (X, Q) with $X_i = 1$ to $(X, Q - e_i \mathbb{1}_{\{Q_i>0\}})$ occur at rate $(1 - p_i)\mu_i$.

Even with possible back-to-back transmissions, the activity process remains a reversible Markov process with product-form stationary distribution given by (1.3), where the factors ν_i/μ_i, $i = 1, \ldots, N$, should be replaced with

$$\sigma_i := \frac{\nu_i}{p_i\mu_i}, \quad i = 1, \ldots, N. \tag{2.1}$$

In the rest of the chapter we will refer to σ_i as the *activity factor of node i*.

Remark. We have implicitly assumed that a node can be active even when it has no packets waiting for transmission. In that case a node transmits dummy packets. The transmission of a dummy packet is stopped upon arrival of a new packet at that node, whose transmission is then started immediately. It would be interesting to extend the analysis to the case where nodes with empty buffers refrain from transmission activity, but this is challenging since the behavior of the activity process then does depend on the queue-length process, as discussed in Subsection 1.2.3.

2.2 SYMMETRIC PARTITE NETWORKS

As alluded to in Chapter 1, transition times, and thus starvation effects and delay asymptotics, strongly depend on the specific structure of the underlying conflict graph. In order to allow for a systematic treatment of a family

of graphs, in this thesis we therefore focus on symmetric scenarios. In this section we specify our assumptions for the conflict graphs and present a few important examples that are analyzed in the subsequent chapters. We remark that slow transitions and excessive delays are not a peculiarity of these highly symmetric networks and may very well occur in asymmetric topologies as well.

To obtain more transparent results, we henceforth assume homogeneous arrival, transmission and activation rates, as well as equal back-off probabilities at all nodes:

$$\lambda_i \equiv \lambda, \quad \mu_i \equiv \mu, \quad \nu_i \equiv \nu, \quad p_i \equiv p, \quad i = 1, \ldots, N.$$

All nodes have then the same activity factor $\sigma_i \equiv \sigma$, and the stationary distribution (1.10) of the activity process $\{X(t)\}_{t\geq 0}$ simplifies to

$$\pi(x) = Z^{-1}\sigma^{\|x\|}, \quad x \in \mathcal{X}, \tag{2.2}$$

where $Z := \sum_{y\in\mathcal{X}} \sigma^{\|y\|}$ is the normalizing constant. As (2.2) shows, when $\sigma > 1$, the stationary probability of a network state $x \in \mathcal{X}$ increases exponentially with its cardinality $\|x\|$.

We first give some definitions and recall several notions from graph theory. Recall that the set $\mathcal{X}$ of admissible network states on the conflict graph G is in one-to-one correspondence with the collection of independent sets of G. In particular, the *dominant states* of the conflict graph G, which we define as the admissible activity states with the largest number of active nodes, correspond to the maximum independent sets of G. Let $\mathcal{D} = \mathcal{D}(G)$ be the collection of all dominant states of the conflict graph G, i.e.

$$\mathcal{D} := \{x \in \mathcal{X} \ : \ \|x\| = \max_{y\in\mathcal{X}} \|y\|\}. \tag{2.3}$$

An *automorphism* of a graph $G = (V, E)$ is a permutation ξ of its vertex set V, such that for all $i, j \in V$

$$(i, j) \in E \iff (\xi(i), \xi(j)) \in E. \tag{2.4}$$

It is immediate from the definition that, given an automorphism ξ of a graph G, a subset $I \subset V$ of nodes is an independent set of G if and only if the subset $\xi(I) = \{\xi(i) \ : \ i \in I\} \subset V$ is. In particular, any automorphism of the graph G induces a permutation on the collection of its maximum independent sets (and thus on the collection of the dominant states of G).

We call a network represented by a conflict graph $G = (V, E)$ *symmetric K-partite* if:

(P1) G is *K-partite*: its nodes can be partitioned into K disjoint independent sets $V_1, \ldots, V_K$;

(P2) For any $k, k' \in \{1, \ldots, K\}$, $k \neq k'$, there exists an automorphism ξ of the graph G such that $\xi(V_k) = V_{k'}$ and $\xi(V_{k'}) = V_k$, while $\xi(V_l) = V_l$ for every $l \neq k, k'$;

(P3) The subsets $V_1, \dots, V_K$ are the only maximum independent sets of G.

In the rest of the chapter, we will tacitly assume that $K \geq 2$. The case $K = 1$ is not interesting, since in view of property (P1) the network G would consist of N isolated nodes. Furthermore, we can always assume that the graph G is connected, since otherwise the connected components of the network can be analyzed separately: Indeed, in this case the evolution of the activity process on each of them would be independent.

For every $k = 1, \dots, K$, denote by $x^{(k)} \in \{0,1\}^N$ the vector describing the network state with full activity in component V_k and no activity elsewhere, i.e.

$$x_i^{(k)} := \begin{cases} 1 & \text{if } i \in V_k, \\ 0 & \text{otherwise.} \end{cases}$$

For every subset $I \subset V$ of nodes, denote by $|I|$ its cardinality. Thanks to property (P1), $x^{(1)}, \dots, x^{(K)}$ are all admissible states and clearly $\|x^{(k)}\| = |V_k|$. Property (P2) yields that $|V_i| = |V_j|$ for any $i, j \in \{1, \dots, K\}$ and therefore $|V_j| = N/K$ for every $j = 1, \dots, K$, which means that

$$\|x^{(1)}\| = \cdots = \|x^{(K)}\| = \frac{N}{K}.$$

In particular, note that in a symmetric K-partite network the total number N of nodes must be a multiple of K. Property (P3) implies that $x^{(1)}, \dots, x^{(K)}$ are the only dominant states for a symmetric K-partite network G. In the regime where σ grows large, the stationary distribution (2.2) concentrates in the dominant states of the network G. In particular,

$$\sum_{x \in \mathcal{D}} \pi(x) \uparrow 1, \quad \text{as } \sigma \to \infty,$$

and for every dominant state $x \in \mathcal{D}$ we have

$$\pi(x) \uparrow \frac{1}{|\mathcal{D}|} = \frac{1}{K}, \quad \text{as } \sigma \to \infty. \tag{2.5}$$

Recall that θ_i has been defined in (1.4) as the *throughput of node i*, i.e. $\theta_i := \sum_{x \in \mathcal{X}} x_i \pi(x)$. In a symmetric K-partite network each node is active in exactly one dominant state and inactive in all the other $K-1$ dominant states, thanks to properties (P1) and (P3). This means that as $\sigma \to \infty$ all nodes have asymptotically the same throughput by virtue of (2.5), since for every $i = 1, \dots, N$

$$\theta_i \uparrow \frac{1}{|\mathcal{D}|} = \frac{1}{K}, \quad \text{as } \sigma \to \infty.$$

In view of these observations, it is natural to define the normalized load as $\rho := K\lambda/\mu$ and to assume $\rho < 1$, since otherwise all the queues will be unstable. The latter condition is not only necessary for stability, but in fact also sufficient for the queue-length process at all nodes to be positive recurrent for large enough values of σ, as we will prove in Section 2.4.

Remark. In this chapter we focus exclusively on CSMA networks with a single transmitting channel available. However, consider for a moment the multi-channel CSMA model described in Section 1.5 and the notation introduced therein. If the interference constraints are described by the same symmetric K-partite graph G on all the available C channels, then we can calculate explicitly the maximum number $d(C)$ of active nodes when C channels are available as

$$d(C) = \begin{cases} \frac{NC}{K} & \text{if } C < K, \\ N & \text{if } C \geq K. \end{cases}$$

and in this way obtain that the aggregate throughput $\Theta(C)$ in the limit $\sigma \to \infty$ of the corresponding network is

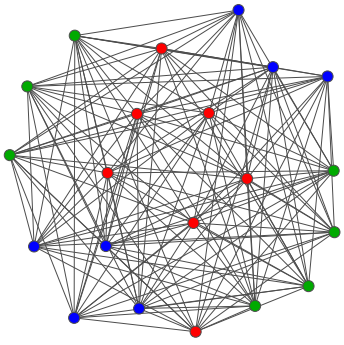
$$\Theta(C) = \begin{cases} \frac{N}{K} & \text{if } C < K, \\ \frac{N}{C} & \text{if } C \geq K. \end{cases}$$

These results follows immediately from the considerations made in Section 1.5 and from the fact that if G is a symmetric K-partite graph, then $\alpha(G) = \frac{N}{K}$ and $C^*(G) = K = \chi(G)$.

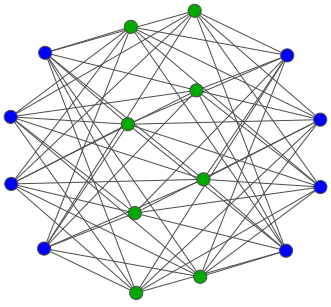
We now present a few key examples of symmetric partite networks.

Turán graphs

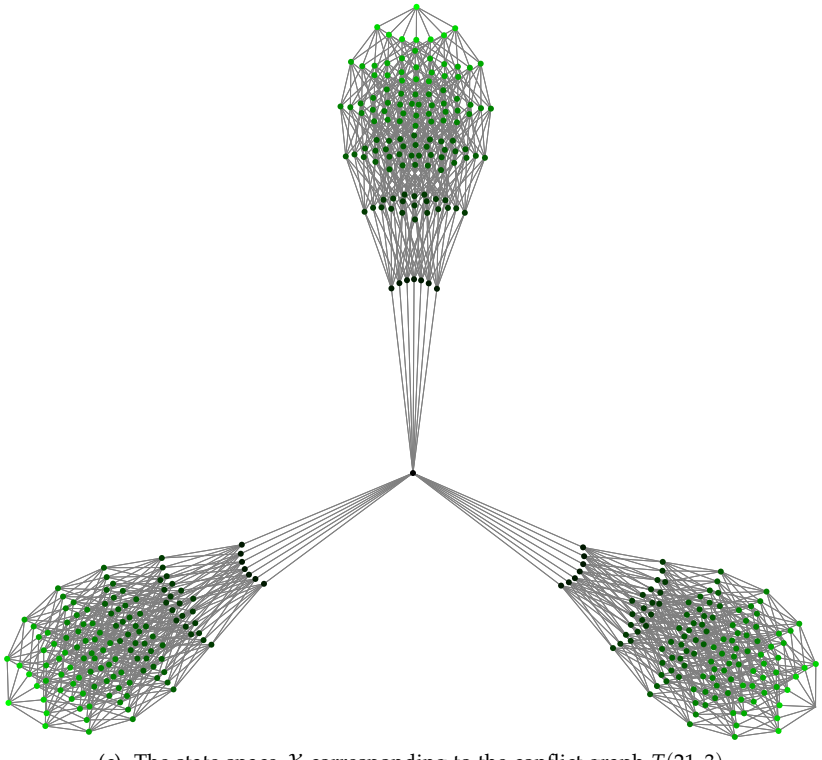
The first example of a symmetric partite network is the *Turán graph* $T(KL, K)$, with $K, L \in \mathbb{N}$ and $K \geq 2$, which is a complete K-partite graph $G = (V, E)$ with KL nodes whose components $V_1, \ldots, V_k$ have all equal size L. More specifically, the vertex set of $T(KL, K)$ is $V = \{1, \ldots, KL\}$ and component V_j is defined as $V_j = \{v \in V \ : \ \lceil v/L \rceil = j\}$. The fact that $T(KL, K)$ is a *complete* partite graph means that there is an edge between two nodes if and only if they belong to different components, i.e. the edge set is $E = \{(v, w) \in V \times V \ : \ \lceil v/L \rceil \neq \lceil w/L \rceil\}$. Two examples of Turán graphs are displayed in Figures 2.1a and 2.1b, where the different components are visualized using different colors. Properties (P1)-(P3) are trivially satisfied by $T(KL, K)$, and therefore the activity process $\{X(t)\}_{t \geq 0}$ on $T(KL, K)$ has K dominant states. Figures 2.1c and 2.1d display the two state space diagrams corresponding to the conflict graphs in Figures 2.1a and 2.1b, respectively. The more active nodes an activity state has, the lighter green is used to color the corresponding vertex in the state space diagram. It is then clear that the outermost nodes in Figure 2.1c (respectively Figure 2.1d) correspond to the 3 (respectively 2) dominant states of the conflict graph $T(21, 3)$ (respectively $T(16, 2)$). In Section 2.3 we will argue that any automorphism of the conflict graph induces an automorphism on the corresponding state space $\mathcal{X}$. Such induced symmetries of the state space are particularly evident in the case of Turán conflict graphs, as illustrated in Figures 2.1c and 2.1d.



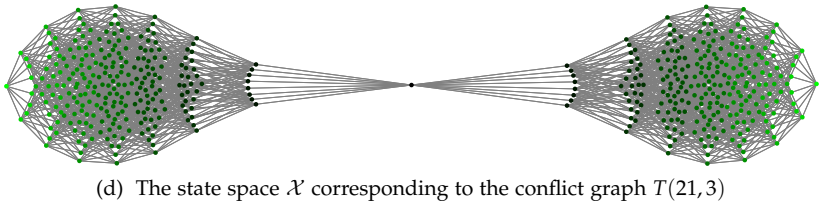
(a) The Turán graph $T(21,3)$



(b) The Turán graph $T(16,2)$



(c) The state space $\mathcal{X}$ corresponding to the conflict graph $T(21,3)$

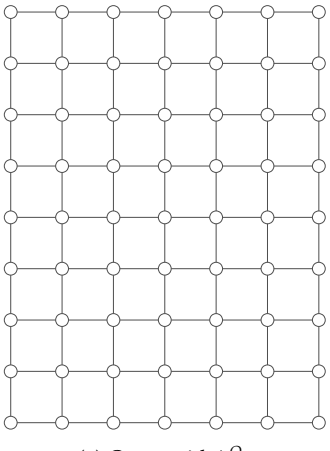


(d) The state space $\mathcal{X}$ corresponding to the conflict graph $T(21,3)$

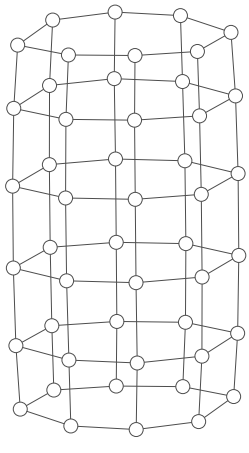
Figure 2.1: The examples of Turán graphs and and the corresponding state space diagrams

Grid graphs

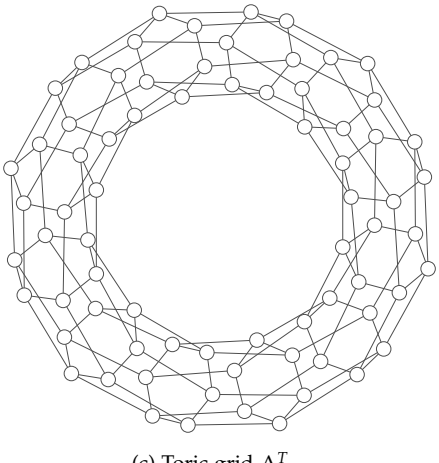
The second example of symmetric partite networks are grid graphs, i.e. finite two-dimensional rectangular square lattices. More precisely, given two integers $K, L \geq 2$, we consider the $K \times L$ grid graph $\Lambda = \Lambda_{K,L}$ with three possible boundary conditions: Toroidal (periodic), cylindrical (semiperiodic) and open. We denote them by $\Lambda^T_{K,L}$, $\Lambda^C_{K,L}$ and $\Lambda^O_{K,L}$, respectively. Figure 2.2 shows examples of the three possible types of boundary conditions. The vertex set of such grid graphs is $\{0, \ldots, L-1\} \times \{0, \ldots, K-1\}$ and thus there are $N = KL$ nodes in total. Every node $v \in \Lambda$ is characterized by its coordinates (v_1, v_2) and is called *even* (*odd*) if the sum of its two coordinates is even (odd).



(a) Open grid $\Lambda^O_{9,7}$



(b) Cylindrical grid $\Lambda^C_{6,8}$



(c) Toric grid $\Lambda^T_{6,12}$

Figure 2.2: Examples of grid graphs

We denote by V_e and V_o the collection of even nodes and that of odd nodes of Λ, respectively. In order for these grid graphs to be symmetric partite networks, we need some further assumptions. The open grid $\Lambda^O_{K,L}$ satisfies property (P1), since it is a bipartite graph (all the neighbors in $\Lambda^O_{K,L}$ of an even node are odd nodes and vice versa). We assume further that the product KL is even, since otherwise the two components V_e and V_o would have different sizes (respectively $\lceil KL/2 \rceil$ and $\lfloor KL/2 \rfloor$) and an automorphism mapping V_e into V_o cannot exist; see property (P2). If $KL \equiv 0 \pmod 2$, then such an automorphism exists, for instance the axial symmetry along the (vertical or horizontal) line that splits the nodes in half, as shown in Figure 2.3.

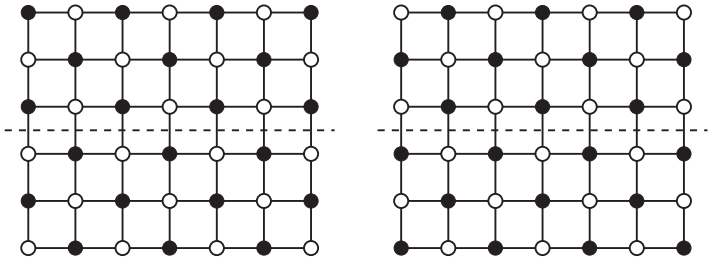


Figure 2.3: The automorphism ξ of the open grid $\Lambda^O_{6,7}$ induced by axial symmetry along the dashed line

In contrast, the cylindrical and toric grids may not be bipartite. In order for them to satisfy property (P1), we further assume that K is an even integer for the cylindrical grid $\Lambda^C_{K,L}$ and that both K and L are even integers for the toric grid $\Lambda^T_{K,L}$. Under these additional assumptions, both these grid types satisfy property (P2) as well, since also for them we can define an automorphism ξ that maps V_e into V_o and vice versa (using an axial symmetry as before). Figure 2.4 illustrates the case of the 6×8 toric grid $\Lambda^T_{6,8}$.

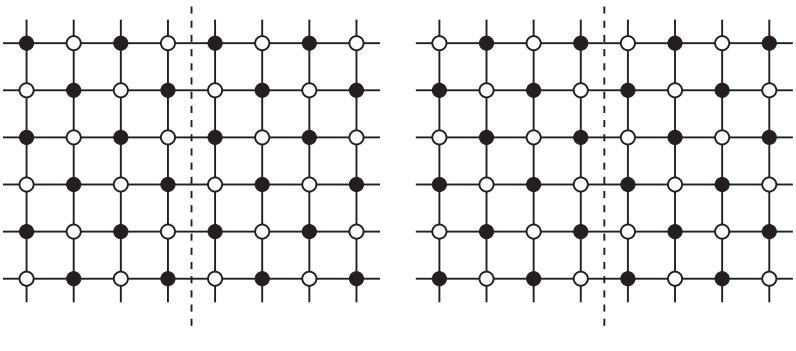


Figure 2.4: The automorphism ξ of the toric grid $\Lambda^T_{6,8}$ induced by axial symmetry along the dashed line

For all three types of boundary conditions, we will prove in Chapter 5 that V_e and V_o are the only two maximum independent sets of $\Lambda^O_{K,L}$ provided that $KL \equiv 0 \pmod 2$ and thus they all satisfy property (P3). Any automorphism ξ of a grid graph $\Lambda_{K,L}$ induces an automorphism $\overline{\xi}$ of the corresponding state space $\mathcal{X}$, see Section 2.3 for further details. Figure 2.5 below displays the state space $\mathcal{X}$ corresponding to the conflict graph $\Lambda^T_{4,4}$. The axial symmetry of $\Lambda^T_{4,4}$ (analogous to that presented in Figure 2.4) induces an automorphism of this state space diagram, which is clearly visible as left-right symmetry.

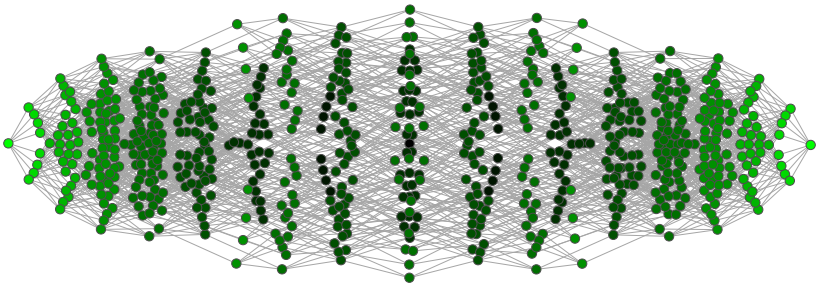


Figure 2.5: The state space diagram $\mathcal{X}$ corresponding to the activity process on the 4×4 grid graph $\Lambda^T_{4,4}$

Triangular grid graphs

The last example of a symmetric partite network is a finite rectangular region (with periodic boundary conditions) of the triangular lattice. More precisely, given two integers $K \geq 2$ and $L \geq 1$, we consider the *triangular grid graph* $\mathcal{T}_{K,L} = (V, E)$ which consists of $N = |V| = 6KL$ nodes of the triangular lattice, which are placed on $2K$ rows of $3L$ nodes each (or equivalently on $6L$ columns with K nodes each), see Figure 2.6. The graph $\mathcal{T}_{K,L} = (V, E)$ satisfies

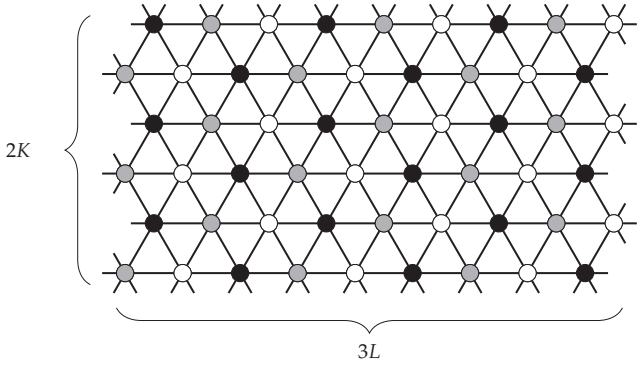


Figure 2.6: The triangular grid graph $\mathcal{T}_{3,3}$

property (P1), since it has a natural tri-partition $V = V_a \cup V_b \cup V_c$, illustrated in Figure 2.6 by coloring the three components respectively in black, white, and gray. The vertical and horizontal dimensions of the triangular grid graph $\mathcal{T}_{K,L}$ are chosen in such a way that the three components have the cardinality, which is required for property (P2) to hold. By considering the axial symmetries along vertical lines, we can easily construct automorphisms that swap two components while mapping the third one to itself. For instance, the axial symmetry displayed in Figure 2.3 maps component V_a ("the gray nodes") into component V_b ("the black nodes"), while the nodes in V_c ("the white nodes") are only permuted. Also in this case, the automorphisms induced by axial symmetries of the triangular grid graph $\mathcal{T}_{3,3}$ on its state space $\mathcal{X}$ are clearly visible in the state space diagram in Figure 1.6, which has axial symmetries as well. Finally, we prove that triangular grid graphs satisfy property (P3) in Chapter 6 by showing that that V_a, V_b and V_c are the only three maximum independent sets of these graphs.

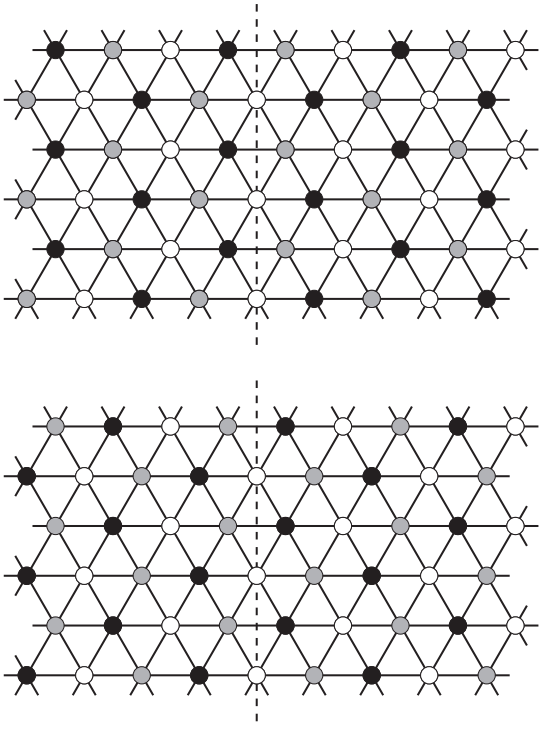


Figure 2.7: The automorphism ξ of the triangular grid graph $\mathcal{T}_{3,3}$ induced by axial symmetry along the dashed line

2.3 TRANSITION TIMES BETWEEN DOMINANT STATES

Given a symmetric partite conflict graph G, it is intuitive that the corresponding state space $\mathcal{X}$ on which the activity process $\{X(t)\}_{t\geq 0}$ evolves is highly symmetric. In this section, we illustrate the consequences of properties (P1)-(P3) for the structure of $\mathcal{X}$ and leverage this symmetry to derive results for the transition times between its dominant states, see Proposition 2.3.1 and Corollary 2.3.2. Furthermore, as already illustrated by (2.5), the stationary distribution of the activity process concentrates in the dominant states as $\sigma \to \infty$. Proposition 2.3.3 complements this fact, by showing that the activity process started in a dominant state spends most of its time there until another dominant state is hit.

For compactness we henceforth we suppress from our notation the dependence on the activity factor σ of all the random variables.

The first result exploits the automorphisms of the conflict graph G, whose existence is guaranteed by property (P2), to construct a coupling between different copies of the activity process and prove in this way properties of the activity process and the first hitting times $T^x_{\mathcal{D}\setminus\{x\}}$ with $x \in \mathcal{D}$.

Proposition 2.3.1 (Transition time properties)**.** *Let $\{X(t)\}_{t\geq 0}$ be the activity process on a symmetric K-partite network G. Then at stationarity, for every dominant state $x \in \mathcal{D}$ and every activity factor σ,*

(i) *The random variable $X(T^x_{\mathcal{D}\setminus\{x\}})$ has a uniform distribution over $\mathcal{D}\setminus\{x\}$;*

(ii) *The random variable $T^x_{\mathcal{D}\setminus\{x\}}$ does not depend on x;*

(iii) *The random variables $T^x_{\mathcal{D}\setminus\{x\}}$ and $X(T^x_{\mathcal{D}\setminus\{x\}})$ are independent.*

Proof. Consider four dominant states $x, y, v, w \in \mathcal{D}$ such that $x \neq v$ and $y \neq w$, and the four corresponding components V_x, V_y, V_v and V_w. Since G is a symmetric K-partite conflict graph, by property (P2) there exist two automorphisms ξ_1 and ξ_2 of the graph G such that:

- $\xi_1(V_x) = V_y$, while all the other components are mapped to themselves (in particular $\xi_1(V_v) = V_v$ and $\xi_1(V_w) = V_w$);
- $\xi_2(V_v) = V_w$, while all the other components are mapped to themselves (in particular $\xi_2(V_x) = V_x$ and $\xi_2(V_y) = V_y$).

In particular, ξ_1 maps the active nodes of x into those of y, and ξ_2 maps the active nodes of v into those of w. Note that since we did not assume that $x \neq y$ and $v \neq w$, we may have to consider the trivial automorphism in place of ξ_1 or ξ_2. Consider the composition $\xi = \xi_1 \circ \xi_2$ of these two automorphisms (which is an automorphism itself). Such an automorphism ξ induces a permutation $\overline{\xi}$ of the corresponding set $\mathcal{X}$ of admissible states on G, defined as

$$\overline{\xi}(z) = (z_{\xi(1)}, \dots, z_{\xi(N)}), \quad z \in \mathcal{X}.$$

In fact, $\overline{\xi}$ is an automorphism of the state space diagram, seen as a graph with vertex set $\mathcal{X}$ and such that any pair of admissible states $z, z' \in \mathcal{X}$ is connected by an edge if and only if $\|z - z'\| \leq 1$. By construction, the automorphism $\overline{\xi}$ maps x to y and v to w simultaneously, i.e.

$$\overline{\xi}(x) = y, \quad \overline{\xi}(\mathcal{D} \setminus \{x\}) = \mathcal{D} \setminus \{y\}, \quad \text{and} \quad \overline{\xi}(v) = w. \tag{2.6}$$

Assume the activity process $\{X(t)\}_{t\geq 0}$ on the conflict graph G to start in the dominant state x at time 0. Let $\{Y(t)\}_{t\geq 0}$ be the process that mimics the moves of the process $X(t)$ via the automorphism $\overline{\xi}$, i.e. set

$$Y(t) := \overline{\xi}(X(t)), \quad \forall\, t \geq 0.$$

For any pair of admissible activity states $z, z' \in \mathcal{X}$, any transition of the process $Y(t)$ from $u = \overline{\xi}(z)$ to $u' = \overline{\xi}(z')$ is feasible and occurs at the same rate as the transition from z to z', since $\overline{\xi}$ is an automorphism. Therefore, the processes $\{X(t)\}_{t\geq 0}$ and $\{Y(t)\}_{t\geq 0}$ are two copies of the activity process on the conflict graph G living in the same probability space, and we have then defined in this way a coupling between them. In view of (2.6), this coupling immediately implies that the activity process $\{X(t)\}_{t\geq 0}$ started at x hits a dominant state in $\mathcal{D} \setminus \{x\}$ precisely when the activity process $\{Y(t)\}_{t\geq 0}$ started at $y = \overline{\xi}(x)$ hits a dominant state in $\mathcal{D} \setminus \{y\}$. We claim that for every $x, y, v, w \in \mathcal{D}$ such that $x \neq v$ and $y \neq w$, we get

$$\mathbb{P}\Big(X(T^x_{\mathcal{D}\setminus\{x\}}) = v,\, T^x_{\mathcal{D}\setminus\{x\}} \leq t\Big) = \mathbb{P}\Big(Y(T^y_{\mathcal{D}\setminus\{y\}}) = w,\, T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big). \tag{2.7}$$

Indeed,

$$\begin{aligned}\mathbb{P}\Big(X(T^x_{\mathcal{D}\setminus\{x\}}) = v, T^x_{\mathcal{D}\setminus\{x\}} \leq t\Big) &= \mathbb{P}\Big(\overline{\xi}(X(T^x_{\mathcal{D}\setminus\{x\}})) = \overline{\xi}(v), T^{\overline{\xi}(x)}_{\overline{\xi}(\mathcal{D}\setminus\{x\})} \leq t\Big)\\ &= \mathbb{P}\Big(Y(T^y_{\mathcal{D}\setminus\{y\}}) = w,\, T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big).\end{aligned}$$

Taking $x = y$ and the limit $t \to \infty$ in (2.7), we obtain that for every $v, w \in \mathcal{D} \setminus \{x\}$

$$\mathbb{P}\Big(X(T^x_{\mathcal{D}\setminus\{x\}}) = v\Big) = \mathbb{P}\Big(Y(T^x_{\mathcal{D}\setminus\{x\}}) = w\Big).$$

Using the fact that $\{X(t)\}_{t\geq 0}$ and $\{Y(t)\}_{t\geq 0}$ have the same statistical law, being two copies of the activity process, it then follows that the random variable $X(T^x_{\mathcal{D}\setminus\{x\}})$ has a uniform distribution over $\mathcal{D} \setminus \{x\}$, that is property (i). In particular,

$$\mathbb{P}\Big(X(T^x_{\mathcal{D}\setminus\{x\}}) = y\Big) = \frac{1}{|\mathcal{D} \setminus \{x\}|} = \frac{1}{K-1}. \tag{2.8}$$

By summing over $v \in \mathcal{D} \setminus \{x\}$ in (2.7), we have that for every $w \in \mathcal{D} \setminus \{y\}$ and every $t \geq 0$

$$\mathbb{P}\Big(T^x_{\mathcal{D}\setminus\{x\}} \leq t\Big) = (K-1) \cdot \mathbb{P}\Big(X(T^y_{\mathcal{D}\setminus\{y\}}) = w,\, T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big). \tag{2.9}$$

By summing over $w \in \mathcal{D} \setminus \{y\}$ in (2.9), we obtain that for every $t \geq 0$

$$\mathbb{P}\Big(T^x_{\mathcal{D}\setminus\{x\}} \leq t\Big) = \mathbb{P}\Big(T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big), \tag{2.10}$$

proving property (ii). Substituting (2.10) into identity (2.9) and using (2.8), we deduce that every $y, w \in \mathcal{D}$ with $y \neq w$ and for every $t \geq 0$,

$$\mathbb{P}\Big(Y(T^y_{\mathcal{D}\setminus\{y\}}) = w,\, T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big) = \mathbb{P}\Big(Y(T^y_{\mathcal{D}\setminus\{y\}}) = w\Big)\mathbb{P}\Big(T^y_{\mathcal{D}\setminus\{y\}} \leq t\Big),$$

that is property (iii). □

Proposition 2.3.1 can be used to obtain the following stochastic representation for the transition time T^x_y between two dominant states $x, y \in \mathcal{D}$:

Corollary 2.3.2 (Stochastic representation of the transition time T^x_y)**.** *For $x, y \in \mathcal{D}$, $x \neq y$:*

$$T^x_y \overset{d}{=} \sum_{i=1}^{N_{\mathcal{D}}} T^{(i)}, \tag{2.11}$$

where $\{T^{(i)}\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables distributed as $T^x_{\mathcal{D}\setminus\{x\}}$ and $N_{\mathcal{D}}$ is an independent geometric random variable with success probability $1/(K-1)$,

$$\mathbb{P}(N_{\mathcal{D}} = m) = \left(\frac{K-2}{K-1}\right)^{m-1} \frac{1}{K-1}, \quad m \geq 1.$$

In particular,

$$\mathbb{E}T^x_y = (K-1) \cdot \mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}. \tag{2.12}$$

Furthermore, if additionally there exists a non-negative random variable Y such that that

$$\frac{T^x_{\mathcal{D}\setminus\{x\}}}{\mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}} \xrightarrow{d} Y, \quad \text{as } \sigma \to \infty,$$

then

$$\frac{T^x_y}{\mathbb{E}T^x_y} \xrightarrow{d} \frac{1}{\mathbb{E}N_{\mathcal{D}}} \sum_{i=1}^{N_{\mathcal{D}}} Y^{(i)}, \quad \text{as } \sigma \to \infty, \tag{2.13}$$

where $\{Y^{(i)}\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables distributed as Y.

Proof. Let $N_{\mathcal{D}}$ be the random variable counting the number of non-consecutive visits of the activity process to dominant states until the dominant state y is hit. In view of Proposition 2.3.1(i), the random variable $N_{\mathcal{D}}$ is geometrically distributed with success probability $\frac{1}{K-1}$, i.e.

$$\mathbb{P}(N_{\mathcal{D}} = m) = \left(\frac{K-2}{K-1}\right)^{m-1} \frac{1}{K-1}, \quad m = 1, 2, \ldots.$$

Therefore $N_{\mathcal{D}}$ depends only on the number of dominant states and, in particular, it does not depend on the activity factor σ. The amount of time $T^x_{\mathcal{D}\setminus\{x\}}$ it takes for the activity process started in a dominant state x to hit a different dominant state in $\mathcal{D}\setminus\{x\}$ does not depend on the starting dominant state, by virtue of Proposition 2.3.1(ii). In view of these considerations and using the independence property in Proposition 2.3.1(iii), we deduce (2.11). Identity (2.12) then immediately follows from Wald's identity, since both $N_{\mathcal{D}}$ and $T^x_{\mathcal{D}\setminus\{x\}}$ have finite expectation.

Lastly, we turn to the proof of the limit in distribution (2.13). Denoting by $\mathcal{L}_A(s) = \mathbb{E}(e^{-sA})$, with $s \geq 0$, the Laplace transform of a random variable A, the stochastic representation (2.11) yields

$$\mathcal{L}_{T^x_y} = G_{N_{\mathcal{D}}}\left(\mathcal{L}_{T^x_{\mathcal{D}\setminus\{x\}}}(s)\right),$$

where $G_{N_{\mathcal{D}}}(\cdot)$ is the probability generating function of the random variable $N_{\mathcal{D}}$, i.e. $G_{N_{\mathcal{D}}}(z) = \mathbb{E}(z^{N_{\mathcal{D}}})$ for every $z \in [0,1]$. By assumption we have that for any $x \in \mathcal{D}$,

$$\mathcal{L}_{T^x_{\mathcal{D}\setminus\{x\}}/\mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}}(s) \to \mathcal{L}_Y(s), \quad \text{as } \sigma \to \infty.$$

Using the fact that $\mathbb{E}T^x_y = \mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}\mathbb{E}N_{\mathcal{D}}$ and (2.12) we obtain

$$\mathcal{L}_{T^x_y/\mathbb{E}T^x_y} = G_{N_{\mathcal{D}}}\left(\mathcal{L}_{T^x_{\mathcal{D}\setminus\{x\}}/\mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}}(s/\mathbb{E}N_{\mathcal{D}})\right) \overset{\sigma\to\infty}{\longrightarrow} G_{N_{\mathcal{D}}}\left(\mathcal{L}_Y(s/\mathbb{E}N_{\mathcal{D}})\right),$$

and the continuity theorem for Laplace transforms yields the conclusion. □

For any subset of states $A \subseteq \mathcal{X}$, let $O_A(t)$ denote the *occupation time* up to time t of the subset A, which is the stochastic process counting the amount of time that the activity process spends in the subset A in the time interval $[0,t]$,

$$O_A(t) := \int_{s=0}^{t} \mathbb{1}_{\{X(s)\in A\}}\,\mathrm{d}s, \tag{2.14}$$

and let $U_A(t)$ denote the stochastic process counting the amount of time that the activity process does *not* spend in the subset A in the time interval $[0,t]$,

$$U_A(t) := t - O_A(t) = \int_{s=0}^{t} \mathbb{1}_{\{X(s)\notin A\}}\,\mathrm{d}s. \tag{2.15}$$

Note that the processes $\{O_A(t)\}_{t\geq 0}$ and $\{U_A(t)\}_{t\geq 0}$ both depend on the activity factor σ, since the activity process does, but we suppress the explicit dependence for compactness.

Assume that the activity process starts at time 0 in the dominant state $x \in \mathcal{D}$. The next result illustrates the intuitive fact that in the limit $\sigma \to \infty$ the activity process started in a state $x \in \mathcal{D}\setminus\{y\}$ at time 0 spends most of its time in dominant states in $\mathcal{D}\setminus\{y\}$ before hitting the dominant state y.

Proposition 2.3.3 (Average time spent in dominant states). *Let G be a symmetric K-partite network. Then, for every pair of dominant states $x, y \in \mathcal{D}$ with $x \neq y$, the activity process on G started in x at time 0 satisfies*

$$\frac{\mathbb{E}U_{\mathcal{D}\setminus\{y\}}(T_y^x)}{\mathbb{E}T_y^x} \downarrow 0, \quad \text{as } \sigma \to \infty, \tag{2.16}$$

and

$$\frac{\mathbb{E}O_x(T^x_{\mathcal{D}\setminus\{x\}})}{\mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}} \uparrow 1, \quad \text{as } \sigma \to \infty. \tag{2.17}$$

Proof. First notice that the quantity $\mathbb{E}U_{\mathcal{D}\setminus\{y\}}(T_y^x)$ is finite as long as $\mathbb{E}T_y^x$ is, since

$$\mathbb{E}U_{\mathcal{D}\setminus\{y\}}(T_y^x) = \mathbb{E}\int_{s=0}^{T_y^x} \mathbb{1}_{\{X(s)\notin\mathcal{D}\setminus\{y\}\}}\,\mathrm{d}s \leq \mathbb{E}\int_{s=0}^{T_y^x}\mathrm{d}s = \mathbb{E}T_y^x.$$

Note that $O_{\mathcal{D}\setminus\{y\}}(T_y^x) \stackrel{d}{=} O_x(T_y^x) + O_{\mathcal{D}\setminus\{x,y\}}(T_y^x)$ and thus

$$U_{\mathcal{D}\setminus\{y\}}(T_y^x) \stackrel{d}{=} T_y^x - (O_{\mathcal{D}\setminus\{x,y\}}(T_y^x) + O_x(T_y^x)),$$

and

$$\frac{\mathbb{E}U_{\mathcal{D}\setminus\{y\}}(T_y^x)}{\mathbb{E}T_y^x} = 1 - \frac{\mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T_y^x)}{\mathbb{E}T_y^x} - \frac{\mathbb{E}O_x(T_y^x)}{\mathbb{E}T_y^x}. \tag{2.18}$$

Consider the random time T^* at which the activity process started in x at time 0 returns to state x after having visited state y. The random time T^* is a regeneration point for the activity process and $T^* \stackrel{d}{=} T_y^x + T_x^y$. Since $T_y^x \stackrel{d}{=} T_x^y$ by virtue of Corollary 2.3.2, we derive

$$\mathbb{E}T^* = 2\mathbb{E}T_y^x. \tag{2.19}$$

The renewal-reward theorem (see for instance [130]) yields

$$\frac{\mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T^*)}{\mathbb{E}T^*} = \pi(\mathcal{D}\setminus\{x,y\}). \tag{2.20}$$

Furthermore,

$$\mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T^*) = \mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T_y^x) + \mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T_x^y) = 2\mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T_y^x). \tag{2.21}$$

Combining (2.19)-(2.21) gives

$$\frac{\mathbb{E}O_{\mathcal{D}\setminus\{x,y\}}(T_y^x)}{\mathbb{E}T_y^x} = \pi(\mathcal{D}\setminus\{x,y\}) \uparrow \frac{K-2}{K}, \quad \text{as } \sigma \to \infty. \tag{2.22}$$

Using again the renewal-reward theorem, we get that

$$\frac{\mathbb{E}O_x(T^*)}{\mathbb{E}T^*} = \pi(x),$$

and, from (2.19) and the fact that $\mathbb{E}O_x(T^*) = \mathbb{E}O_x(T_y^x)$ it follows that

$$\frac{\mathbb{E}O_x(T_y^x)}{\mathbb{E}T_y^x} = 2 \cdot \pi(x) \uparrow \frac{2}{K}, \quad \text{as } \sigma \to \infty. \tag{2.23}$$

Combining (2.18), (2.22) and (2.23) gives

$$\frac{\mathbb{E}U_{\mathcal{D}\setminus\{y\}}(T_y^x)}{\mathbb{E}T_y^x} \downarrow 1 - \frac{K-2}{K} - \frac{2}{K} = 0, \quad \text{as } \sigma \to \infty.$$

Consider the random time T' at which the activity process started in x at time 0 returns to state x after having visited a state in the subset $\mathcal{D} \setminus \{x\}$. The random time T' is a regeneration epoch for the activity process and, invoking (2.12), its expected value satisfies the identity

$$\begin{aligned}\mathbb{E}T' &= \mathbb{E}T^x_{\mathcal{D}\setminus\{x\}} + \sum_{z\in\mathcal{D}\setminus\{x\}} \mathbb{P}\Big(X(T^x_{\mathcal{D}\setminus\{x\}}) = z\Big)\mathbb{E}T_x^z \\ &= (K-1)\cdot ET_y^x + \mathbb{E}T_x^y = K \cdot \mathbb{E}T^x_{\mathcal{D}\setminus\{x\}},\end{aligned} \tag{2.24}$$

where we have used Proposition 2.3.1 and (2.12). The renewal-reward theorem says that $\frac{\mathbb{E}O_x(T')}{\mathbb{E}T'} = \pi(x)$, and, from (2.24) and the fact that $\mathbb{E}O_x(T') = \mathbb{E}O_x(T^x_{\mathcal{D}\setminus\{x\}})$ we get

$$\frac{\mathbb{E}O_x(T^x_{\mathcal{D}\setminus\{x\}})}{\mathbb{E}T^x_{\mathcal{D}\setminus\{x\}}} = K \cdot \pi(x) \uparrow 1, \quad \text{as } \sigma \to \infty. \qquad \square$$

2.4 STABILITY OF THE QUEUE-LENGTH PROCESS

In this section we prove that for a sufficiently large activity factor σ all queue-length processes $\{Q_i(t)\}_{t\geq 0}$, $i = 1, \ldots, N$, on symmetric K-partite networks are positive recurrent when the normalized load $\rho = K\lambda/\mu$ is strictly smaller than 1. The crucial idea is that for sufficiently large σ the activity process spends a sufficiently large fraction of time in dominant states, so that the buffer content at each node has a negative expected drift. Furthermore, in Proposition 2.4.3 we show specifically how large the activity factor should be in order for positive recurrence to be feasible.

Although the aim of this section is to prove stability results for symmetric K-partite networks, the next two auxiliary results hold in more generality. Consider the activity process $\{X(t)\}_{t\geq 0}$ on a general conflict graph $G = (V, E)$ of N nodes with heterogeneous rates. In particular, node i has back-off rate ν_i,

transmission rate μ_i, probability p_i of entering back-off after a transmission is completed, and packet arrival rate λ_i, so that its load is $\rho_i = \frac{\lambda_i}{\mu_i}$.

The first auxiliary result relates the average time that a given node is active in the time interval $[0, t]$ and the average number of packet transmissions of a node up to time t. Consider a node i and assume that the activity process starts at time 0 in a state $x \in \mathcal{X}$ such that $x_i = 0$. Denote by $S_i(t)$ the stochastic process counting the number of transmission completions at node i in the time interval $[0, t]$ and by $A_i(t)$ the cumulative amount of time that node i has been active in the time interval $[0, t]$,

$$A_i(t) := \int_{s=0}^{t} X_i(s)\,\mathrm{d}s, \quad t \geq 0.$$

It is implicit in these definitions that $S_i(0) = 0$ and $A_i(0) = 0$. Define the continuous-time process

$$M_i(t) := S_i(t) - \mu_i A_i(t), \quad t \geq 0.$$

Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the natural filtration of the Markov process $\{X(t), Q(t)\}_{t\geq 0}$.

Proposition 2.4.1 ($M(t)$ is a martingale)**.** *The stochastic process $\{M_i(t)\}_{t\geq 0}$ is a $\mathcal{F}_t$-martingale. In particular, for every $t \geq 0$,*

$$\mathbb{E}S_i(t) = \mu_i \mathbb{E}A_i(t). \tag{2.25}$$

Proof. For simplicity, consider first the case in which $p_i = 1$ for every $i = 1, \ldots, N$, so that node i always enters back-off after a transmission is completed. The process $\{M_i(t)\}_{t\geq 0}$ is clearly adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, since the processes $S_i(t)$ and $A_i(t)$ are $\mathcal{F}_t$-measurable for every $t \geq 0$. Furthermore, for any $t \geq 0$,

$$-\mu_i A_i(t) \leq_{\mathrm{st}} M_i(t) \leq_{\mathrm{st}} S_i(t),$$

and thus $\mathbb{E}|M_i(t)| \leq \max\{\mu_i \mathbb{E}A_i(t), \mathbb{E}S_i(t)\} \leq \mu_i t$. Finally, the identity

$$\mathbb{E}(S_i(t) - S_i(s) \mid \mathcal{F}_s) = \mu_i \mathbb{E}(A_i(t) - A_i(s) \mid \mathcal{F}_s)$$

holds, since the number of completed transmissions during a time interval of length l where node i is continuously active has a Poisson distribution with rate $\mu_i l$.

It is easy to see that the same arguments hold also in the case where the probability p_i for node i to enter back-off after a transmission is completed is strictly smaller than 1. Indeed, having $p_i < 1$ influences only the back-off periods (in particular how often they occur), but it has no influence on the relation between $S_i(t)$ and $A_i(t)$, as long as the transmission times are exponentially distributed. □

The second auxiliary result proves that $\rho_i < \theta_i$ for all $i = 1, \ldots, N$ is a sufficient and necessary condition for the queue-length process at each node to be positive recurrent, as claimed in Subsection 1.2.3.

Proposition 2.4.2 (Queue-length process is positive recurrent)**.** *Consider a network described by a conflict graph G of N nodes. Then all queue-length processes $\{Q_i(t)\}_{t\geq 0}$, $i = 1, \ldots, N$, are positive recurrent if and only if*

$$\rho_i < \theta_i, \quad \forall\, i = 1, \ldots, N. \tag{2.26}$$

Proof. The activity process on G is an irreducible Markov process on $\mathcal{X}$, which is a finite state space, and therefore any activity state can be seen as renewal point. Consider the activity state where all nodes are inactive, which we denote by $\mathbf{0} \in \mathcal{X}$. The moments at which the activity process $\{X(t)\}_{t\geq 0}$ hits $\mathbf{0}$ are regeneration epochs with regeneration time $\tau := T_{\mathbf{0}}^{\mathbf{0}}$. Consider node i and define $\{\hat{Q}_n\}_{n\in\mathbb{N}}$ to be the queue-length process at node i embedded at the epochs when the activity process hits the state $\mathbf{0}$. Thanks to the strong Markov property, this latter process is an irreducible time-homogeneous Markov chain on $\mathbb{N}$. Denote by S_n the number of (potential) packet transmissions during the n-th cycle (which counts also the transmissions of dummy packets) and by R_n the total number of packets which arrived during the n-th cycle. By construction, these two random variables are independent of each other and independent from those of different cycles. We claim that the following inequality holds for every $n \in \mathbb{N}$:

$$\hat{Q}_{n+1} \leq (\hat{Q}_n - S_n)^+ + R_n.$$

Indeed, we underestimate the total number of packets served by allowing transmission only of the awaiting packets which arrived in the previous cycles. Consider now another fictitious Markov chain $\{\overline{Q}_n\}_{n\in\mathbb{N}}$ defined by the initial condition $\overline{Q}(0) = \hat{Q}(0)$ and by the classical Lindley recursion $\overline{Q}_{n+1} = (\overline{Q}_n - S_n + R_n)^+$. Clearly,

$$\hat{Q}_n \leq \overline{Q}_n + R_n, \quad \forall\, n \in \mathbb{N},$$

hence if 0 is a positive recurrent state for $\{\overline{Q}_n\}_{n\in\mathbb{N}}$, then it is so for $\{\hat{Q}_n\}_{n\in\mathbb{N}}$ as well, since there is a positive probability that $R_n = 0$. Let Y be a random variable distributed as $R - S$. It is well known that 0 is a positive recurrent state for the Markov chain $\{\overline{Q}_n\}_{n\in\mathbb{N}}$ if $\mathbb{E}Y < 0$. We will show that this is indeed the case. The average number $\mathbb{E}S$ of packet transmissions during a cycle at node i is equal to the transmission rate μ_i multiplied by the average cumulative amount of time that node i is active in such a cycle, in view of Proposition 2.4.1, and thus

$$\mathbb{E}S = \mathbb{E}S_i(\tau) = \mu_i \mathbb{E}A_i(\tau).$$

Therefore,

$$\mathbb{E}Y = \mathbb{E}R - \mathbb{E}S = \lambda_i \mathbb{E}\tau - \mu_i \mathbb{E}A_i(\tau) = \lambda_i \mathbb{E}\tau - \mu_i \frac{\mathbb{E}A_i(\tau)}{\mathbb{E}\tau}\mathbb{E}\tau = (\lambda_i - \mu_i\theta_i)\mathbb{E}\tau,$$

where the last equality follows by the renewal-reward theorem, since

$$\frac{\mathbb{E}A_i(\tau)}{\mathbb{E}\tau} = \lim_{t\to\infty} \frac{1}{t}\int_0^t X_i(s)\,\mathrm{d}s = \sum_{x\in\mathcal{X}} \pi(x)x_i = \theta_i.$$

Since $\mathbb{E}\tau > 0$, assumption (2.26) immediately yields that $\mathbb{E}Y < 0$. Therefore, also for the chain $\{\hat{Q}_n\}_{n\in\mathbb{N}}$ the state 0 is positive recurrent and, since the cycle time has finite expectation (the Markov process $\{X(t)\}_{t\geq 0}$ is irreducible on a finite state space, hence positive recurrent), it follows that the original continuous-time process $\{Q_i(t)\}_{t\geq 0}$ is positive recurrent as well. For the converse implication, it is easy to show that if $\rho_i \geq \theta_i$ for some $i = 1, \ldots, N$, then the queue-length process $\{Q_i(t)\}_{t\geq 0}$ is not positive recurrent. □

We now return to the case of symmetric K-partite graphs with homogeneous rates. As mentioned earlier in this section, the fraction of time θ that each individual node is active will approach $1/K$ arbitrarily closely for sufficiently large values of σ. This suggests that $\rho = \frac{K\lambda}{\mu} < 1$ is sufficient for stability of the queue-length process at each node for large enough values of the activity factor σ. We now show that this is indeed the case. If $\rho < 1$, there exists an $\epsilon > 0$ such that

$$\frac{\lambda}{\mu} < \frac{1}{K} - \epsilon.$$

Since for every node $\theta \uparrow \frac{1}{K}$ as $\sigma \to \infty$, there exists an activity factor σ_ϵ such that $\theta > \frac{1}{K} - \epsilon$. This fact implies that for a sufficiently large activity factor $\sigma > \sigma_\epsilon$, for every node the inequality $\frac{\lambda}{\mu} < \theta$ holds, and thus all queue-length processes $\{Q(t)_i\}_{t\geq 0}$, $i = 1, \ldots, N$, are positive recurrent by virtue of Proposition 2.4.2.

Furthermore, for symmetric K-partite networks we can obtain an explicit lower bound in terms of the normalized load ρ for the value of the activity factor σ that is required for all the queues to be stable. The lower bound, presented in the next proposition, implies that σ should be of the same order as $(1-\rho)^{-1}$, corroborating the intuition that a high activity factor is necessary to cope with a high load.

Proposition 2.4.3 (Necessary condition on the activity factor for stability)**.** *Consider a symmetric K-partite network $G = (V, E)$ on N nodes with $N/K > 1$. Then, in order to for all queues to be stable, it is required that*

$$\sigma > \frac{1}{1-\rho} - \frac{N}{K}.$$

Proof. We have already proved in Proposition 2.4.2 that

$$\rho_i < \theta_i, \quad \forall\, i = 1, \ldots, N,$$

is a necessary condition for all queue-length processes to be positive recurrent. These inequalities imply that the aggregate load must be smaller than the mean total number of active nodes, yielding

$$\begin{aligned} N\frac{\lambda}{\mu} = \sum_{i\in V} \rho_i < \sum_{i\in V} \theta_i = \sum_{i\in V} \sum_{x\in\mathcal{X}} \pi(x) x_i = \sum_{x\in\mathcal{X}} \pi(x) \sum_{i\in V} x_i \\ \leq \frac{N}{K} \sum_{x\in\mathcal{D}} \pi(x) + \left(\frac{N}{K} - 1\right) \sum_{x\in\mathcal{X}\setminus\mathcal{D}} \pi(x) = \frac{N}{K} - \sum_{x\in\mathcal{X}\setminus\mathcal{D}} \pi(x). \end{aligned} \tag{2.27}$$

We have implicitly used the fact that all the components have $N/K > 1$ nodes and that any non-dominant states have at most $N/K - 1$ active nodes. By definition the normalized load ρ is equal to $K\lambda/\mu$, so it follows from (2.27) that

$$\rho < 1 - \frac{K}{N}\pi(\mathcal{X} \setminus \mathcal{D})$$

or, equivalently,

$$\pi(\mathcal{D}) > 1 - \frac{N}{K}(1-\rho). \tag{2.28}$$

On the other hand, the fraction of time that the system spends in the subset $\mathcal{D}$ of dominant states in the long-run obeys

$$\begin{aligned}\pi(\mathcal{D}) &= \sum_{x\in\mathcal{D}} \pi(x) = \sum_{x\in\mathcal{D}} Z^{-1}\sigma^{N/K} = Z^{-1}K\sigma^{N/K} \\ &\leq \frac{K\sigma^{N/K}}{K\sigma^{N/K} + N/K \cdot K\sigma^{N/K-1}} = \frac{K\sigma}{K\sigma + N}.\end{aligned} \tag{2.29}$$

Combining inequalities (2.28) and (2.29), we obtain

$$\sigma > \frac{1 - \frac{N}{K}(1-\rho)}{1-\rho} = \frac{1}{1-\rho} - \frac{N}{K}.$$

□

2.5 AVERAGE STEADY-STATE DELAY LOWER BOUNDS

In this section we derive lower bounds for the average steady-state packet delay in symmetric partite networks. In the prologue of this chapter we have highlighted the three crucial ideas behind the proof. The first of these three observations has already been formalized by Proposition 2.4.3 in the previous section. As far as the second observation is concerned, an intuitive explanation is given in Chapter 1 and the precise results will be provided in Chapter 4. The third observation states that the long-term average packet delay at any node in the network is intimately related to the average transition time between dominant states. This section is devoted this latter observation, which is formalized by Theorem 2.5.1 below.

We henceforth assume μ to be fixed, and when we write $\sigma \to \infty$, we allow for either $\nu \to \infty$, $p \downarrow 0$, or both. For compactness we only attach σ in brackets to various quantities to reflect the dependence on both ν and p in limit statements for $\sigma \to \infty$.

Let $\mathbb{E}W(\sigma)$ be the long-term average packet delay (waiting time plus service time) of a given node, where again the attached σ in brackets describes the dependence on both ν and p.

Theorem 2.5.1 (Long-term Average Delay)**.** *Consider a symmetric K-partite network G. For any dominant states $x, y \in \mathcal{D}$ with $x \neq y$,*

$$\liminf_{\sigma\to\infty} \frac{\mathbb{E}W(\sigma)}{\mathbb{E}T_y^x(\sigma)} \geq \frac{K-1}{2(K-\rho)}.$$

Proof. Assume that the activity process starts in a dominant state x and consider a node, say node i, that is active in a different dominant state, say $y \neq x$. Recall that properties (P1) and (P3) imply that node i is inactive in all the other dominant states in $\mathcal{D}^* := \mathcal{D} \setminus \{y\}$. Let

$$\mathbb{E}L = \lim_{T\to\infty} \frac{1}{T} \int_{t=0}^{T} L(t)\mathrm{d}t, \tag{2.30}$$

with $L(t)$ denoting the queue length at node i at time t, i.e. the total number of packets, including the packet that may possibly be in the process of being transmitted. Assuming that $\mathbb{E}W$ exists, Little's law can then be invoked to conclude that $\mathbb{E}L = \lambda\mathbb{E}W$ exists as well, so it suffices to establish a lower bound for $\mathbb{E}L$.

The fact that the quantity $\mathbb{E}L$ in (2.30) is well defined is guaranteed by the renewal-reward theorem, which is valid for $\{L(t)\}_{t\geq 0}$, since it is the projection (which is a measurable function) of a positive recurrent regenerative process $\{X(t), Q(t)\}$, with the same regeneration time τ^* of the original process $\{X(t), Q(t)\}_{t\geq 0}$, with $\mathbb{E}\tau^* < \infty$ and $\mathbb{E}\int_0^{\tau^*} L(t)\,\mathrm{d}t < \infty$. Furthermore [130, Theorem 2.1] states that

$$\mathbb{E}L = \lim_{T\to\infty} \frac{1}{T} \int_0^T L(t)\,\mathrm{d}t \overset{\text{w.p.1}}{=} \frac{\mathbb{E}\left[\int_0^{\tau^*} L(t)\,\mathrm{d}t\right]}{\mathbb{E}\tau^*} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}L(t)\,\mathrm{d}t. \tag{2.31}$$

By definition, $L(t)$ increases whenever a packet arrives at the node, as governed by a Poisson process of rate λ. Further observe that $L(t)$ cannot decrease when the activity process resides in any of the $K-1$ dominant states different from y, since they preclude the activity of node i. When the activity process does not reside in a dominant state in $\mathcal{D}^*$, node i could potentially be active, and $L(t)$ decreases whenever a transmission is completed and there are packets in the node ($L(t) \neq 0$).

In order to derive a lower bound for $\mathbb{E}L$, we consider an alternating renewal process. Let $\{Z(t)\}_{t\geq 0}$ be the process with $Z(0) = L(0)$ that increases at rate λ when the activity process resides in the subset of dominant states $\mathcal{D}^*$ and decreases at rate $\lambda - \mu$ at all other times, as long as $Z(t)$ is positive. We will show later that $\{Z(t)\}_{t\geq 0}$ is a regenerative process and that the quantity

$$\mathbb{E}Z = \lim_{T\to\infty} \frac{1}{T} \int_{t=0}^{T} Z(t)\,\mathrm{d}t \tag{2.32}$$

is well defined. Let us show first that $\mathbb{E}L \geq \mathbb{E}Z$ almost surely. Denote by $\{R(t)\}_{t\geq 0}$ the packet arrival process at node i, which is a Poisson process of rate λ. Let $S^{\text{true}}(t)$ be the integer-valued random variable counting the total number of completed transmissions up to time t, excluding the transmissions of dummy packets. By definition, we have that for every $t \geq 0$,

$$\mathbb{E}L(t) = \mathbb{E}R(t) - \mathbb{E}S^{\text{true}}(t),$$

and, since the left-hand side is always non-negative, so is the right-hand side. Moreover, since the process $S(t)$ counts *all* completed transmissions up to time t, the following stochastic inequality holds

$$S(t) \geq_{\text{st}} S^{\text{true}}(t), \quad \forall\, t \geq 0. \tag{2.33}$$

Combining these facts, we obtain that for every $t \geq 0$,

$$\begin{aligned}\mathbb{E}L(t) &= \mathbb{E}R(t) - \mathbb{E}S^{\text{true}}(t) \geq [\mathbb{E}R(t) - \mathbb{E}S(t)]^+ \\ &\overset{(2.25)}{=} [\lambda t - \mu \mathbb{E}A(t)]^+ \geq [\lambda t - \mu \mathbb{E}U_{\mathcal{D}^*}(t)]^+ = \mathbb{E}Z(t).\end{aligned} \tag{2.34}$$

Hence, it follows from (2.31) and (2.34) that

$$\begin{aligned}\mathbb{E}L &= \lim_{T\to\infty} \frac{1}{T}\int_0^T L(t)\,\mathrm{d}t \overset{\text{w.p.1}}{=} \lim_{T\to\infty} \frac{1}{T}\int_{t=0}^T \mathbb{E}L(t)\mathrm{d}t \\ &\overset{(2.34)}{\geq} \lim_{T\to\infty} \frac{1}{T}\int_{t=0}^T \mathbb{E}Z(t)\mathrm{d}t \overset{\text{w.p.1}}{=} \lim_{T\to\infty} \frac{1}{T}\int_0^T Z(t)\,\mathrm{d}t.\end{aligned} \tag{2.35}$$

We will now show that $\{Z(t)\}_{t\geq 0}$ is a regenerative process and derive a lower bound for the quantity $\mathbb{E}Z$. We say that a *bad renewal period* starts each time we observe a first entrance into a dominant state in $\mathcal{D}^* = \mathcal{D} \setminus \{y\}$ after a visit to the dominant state y, and likewise a *good renewal period* starts each time the process enters the dominant state y for the first time after a visit to another dominant state in $\mathcal{D}^*$. Thus the lengths of the good and bad renewal periods correspond to the transition times $T^y_{\mathcal{D}^*}$ and T^x_y (for any $x \in \mathcal{D}^*$ thanks to the symmetry), respectively. We define a *cycle* as the period consisting of a bad and subsequent good renewal period.

Let $U_{\mathcal{D}^*}$ be the random variable representing the amount of time during a bad renewal period that the activity process $\{X(t)\}_{t\geq 0}$ does not reside in the subset $\mathcal{D}^*$. Since the dominant state y cannot be visited during a bad renewal period, the random variable $U_{\mathcal{D}^*}$ counts the total amount of time spent in non-dominant states during a bad renewal period. Thanks to the regenerative structure of the process $\{Z(t)\}_{t\geq 0}$, if $X(0) = x$ for any $x \in \mathcal{D}^*$ we have

$$U_{\mathcal{D}^*} = \int_{s=0}^{T^x_y} \mathbb{1}_{\{X(s)\notin\mathcal{D}^*\}}\,\mathrm{d}s.$$

Using the notation introduced in Section 2.3, we have $U_{\mathcal{D}^*} \overset{d}{=} U_{\mathcal{D}^*}(T^x_y)$ and, in particular, Proposition 2.3.3 applies.

In order to evaluate the quantity (2.32), let T_k and V_k be the durations of the k-th bad and k-th good renewal periods, respectively. Let U_k be the amount of time that the activity process does not reside in $\mathcal{D}^*$ during the k-th bad renewal period, and let $S_m = \sum_{k=1}^m (T_k + V_k)$ be the duration of the first m cycles. Note that T_k and V_k are i.i.d. copies of the random variables T^x_y and $T^y_{\mathcal{D}^*}$, respectively. Assume that a bad renewal period starts at time 0 with $Z(0) = 0$, and define the random variable $M = \inf\{m \geq 1 : Z(S_m) = 0\}$. Observe that

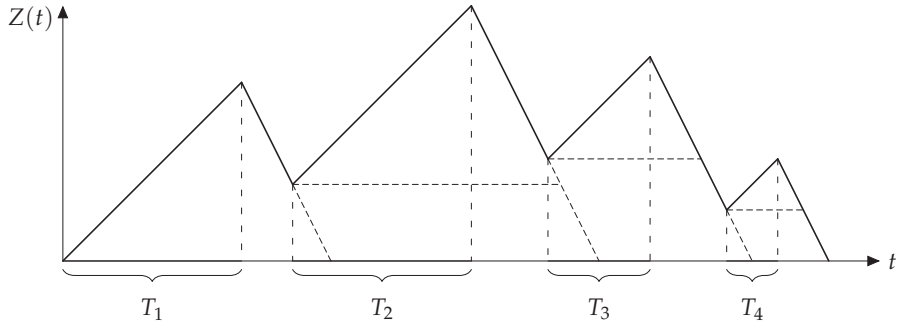


Figure 2.8: Sample path of the process $Z(t)$, with $M = 4$

S_M is a regeneration epoch for the process $\{Z(t)\}_{t\geq 0}$, and hence the renewal-reward theorem implies that

$$\lim_{T\to\infty} \frac{1}{T}\int_{t=0}^{T} Z(t)\mathrm{d}t = \frac{\mathbb{E}\int_{t=0}^{S_M} Z(t)\mathrm{d}t}{\mathbb{E}S_M}. \tag{2.36}$$

Considering the denominator, it follows from Wald's equation, symmetry considerations and (2.12) that

$$\mathbb{E}S_M = \mathbb{E}M(\mathbb{E}T_y^x + \mathbb{E}T_{\mathcal{D}^*}^y) = K\cdot\mathbb{E}M\cdot\mathbb{E}T_{\mathcal{D}^*}^y = \frac{K}{K-1}\mathbb{E}M\cdot\mathbb{E}T_y^x. \tag{2.37}$$

Turning attention to the numerator in (2.36), we first condition on the number of cycles M that have elapsed by the regeneration epoch, the durations $T_1, T_2, \ldots, T_M$ of the bad renewal periods involved, and the amounts of time $U_1, U_2, \ldots, U_M$ that the activity process does not reside in $\mathcal{D}^*$. When $U_k = 0$, inspection of Figure 2.8 shows that the area of the triangle associated with the k-th cycle is

$$\frac{1}{2}\lambda T_k^2 + \frac{1}{2}\frac{\lambda^2}{\mu-\lambda}T_k^2 = \frac{1}{2}\frac{\lambda\mu}{\mu-\lambda}T_k^2,$$

while the area of the parallelogram associated with the k-th cycle is

$$Z_k\Big(T_k + \frac{\lambda}{\mu-\lambda}T_k\Big) = \frac{\mu}{\mu-\lambda}Z_kT_k,$$

with $Z_k = Z(S_k)$ the value of the process $Z(\cdot)$ at the start of the k-th cycle.

In general, when U_k may not be zero, a similar geometric construction leads to the conclusion that the two areas are

$$\frac{1}{2}\lambda T_k\left(T_k - \frac{\mu}{\lambda}U_k\right) + \frac{1}{2}\frac{\lambda^2}{\mu-\lambda}\left(T_k - \frac{\mu}{\lambda}U_k\right)^2 \geq \frac{1}{2}\frac{\lambda\mu}{\mu-\lambda}\left(T_k - \frac{\mu}{\lambda}U_k\right)^2,$$

and

$$Z_k\Big(T_k + \frac{\lambda}{\mu-\lambda}\left(T_k - \frac{\mu}{\lambda}U_k\right)\Big) = \frac{\mu}{\mu-\lambda}Z_k\left(T_k - U_k\right),$$

respectively. Unconditioning, i.e. taking expectations with respect to the number of cycles M and the random variables T_k and U_k, $k = 1, \ldots, M$, we find

$$\mathbb{E}\int_{t=0}^{S_M} Z(t)\mathrm{d}t \geq \frac{1}{2}\frac{\lambda\mu}{\mu-\lambda}\mathbb{E}\Big(\sum_{k=1}^{M}(T_k - \frac{\mu}{\lambda}U_k)^2\Big) + \frac{\mu}{\mu-\lambda}\mathbb{E}\Big(\sum_{k=1}^{M} Z_k(T_k - U_k)\Big).$$

Since $\mathbb{E}\Big(\sum_{k=1}^{M} Z_k(T_k - U_k)\Big) \geq 0$, this inequality, together with (2.37), yields

$$\begin{aligned}\frac{\mathbb{E}\int_{t=0}^{S_M} Z(t)\mathrm{d}t}{\mathbb{E}S_M} &\geq \frac{K-1}{2K}\frac{\lambda\mu}{(\mu-\lambda)\mathbb{E}T_y^x}\frac{\mathbb{E}\left(\sum_{k=1}^{M}(T_k - \frac{\mu}{\lambda}U_k)^2\right)}{\mathbb{E}M}\\ &= \frac{K-1}{2K}\frac{\lambda\mu}{(\mu-\lambda)\mathbb{E}T_y^x}\mathbb{E}\left(T_y^x - \frac{\mu}{\lambda}U_{\mathcal{D}^*}\right)^2,\end{aligned}$$

where the random variables T_y^x and $U_{\mathcal{D}^*}$ have the joint distribution of the duration of a bad renewal period and the amount of time that the activity process does not reside in $\mathcal{D}^*$ during that period. Applying Jensen's inequality, we obtain

$$\begin{aligned}\frac{\mathbb{E}\int_{t=0}^{S_M} Z(t)\mathrm{d}t}{\mathbb{E}S_M} &\geq \frac{K-1}{2K}\frac{\lambda\mu}{(\mu-\lambda)\mathbb{E}T_y^x}\left(\mathbb{E}T_y^x - \frac{\mu}{\lambda}\mathbb{E}U_{\mathcal{D}^*}\right)^2\\ &\geq \frac{K-1}{2K}\frac{\lambda\mu}{\mu-\lambda}\mathbb{E}T_y^x\Big(1 - \frac{\mu}{\lambda}\frac{\mathbb{E}U_{\mathcal{D}^*}}{\mathbb{E}T_y^x}\Big)^2.\end{aligned}$$

In view of (2.16), we find that for any $\delta > 0$, the liminf of the latter expression (for large values of σ) is bounded from below by

$$\frac{K-1}{2K}\frac{\lambda\mu}{\mu-\lambda}\mathbb{E}T_y^x\left(1 - \delta\frac{\mu}{\lambda}\right)^2.$$

Since $\delta > 0$ is arbitrary, we deduce

$$\liminf_{\sigma\to\infty}\frac{\mathbb{E}L}{\mathbb{E}T_y^x} \geq \frac{(K-1)\lambda}{2K}\frac{\mu}{\mu-\lambda} = \frac{(K-1)\lambda}{2(K-\rho)},$$

and the statement of the theorem then follows by applying Little's law. □

Remark. As mentioned in Chapter 1, the stationary distribution of the activity process is insensitive with respect to the distribution of the back-off and transmission times. However, this insensitivity property does not apply to the transient behavior of the activity process and, in particular, to transition times. In fact, for non-exponential distributions, the activity process is no longer Markovian, unless one adopts a more elaborate state space description. Nevertheless, we expect that some of the results extend to non-exponential distributions as well, in view of the fact that the order of magnitude of transition times is closely related with (ratios of) stationary probabilities. A detailed investigation of this more general scenario is beyond the scope of this thesis and could be an interesting topic for further research.

2.6 DELAY SCALING IN SYMMETRIC PARTITE NETWORKS

In this section we derive the average long-term delay scaling for symmetric partite networks and then illustrate in more detail the results for the specific networks introduced in Section 2.2. The main result of this section shows that, in order to determine the asymptotic scaling of the average delay in symmetric partite networks, the key quantity to study is the asymptotic order of magnitude $\Gamma(G)$ of the average transition time between dominant states when σ grows large. The proof combines the lower bound for the activity factor σ given in Proposition 2.4.3 and the deep connection between delay and transition times established in Theorem 2.5.1.

Corollary 2.6.1 (Asymptotic delay scaling)**.** *Consider a symmetric K-partite network G. Assume further that there exists a constant $\Gamma(G) > 0$ such that for any pair of dominant states $x, y \in \mathcal{D}$*

$$\lim_{\sigma\to\infty} \frac{\log \mathbb{E}T_y^x(\sigma)}{\log \sigma} = \Gamma(G) > 0. \tag{2.38}$$

If all queues are stable, then

$$\liminf_{\rho\uparrow 1} \log_{\frac{1}{1-\rho}} \mathbb{E}W(\rho) = \liminf_{\rho\uparrow 1} \frac{\log \mathbb{E}W(\rho)}{\log\left(\frac{1}{1-\rho}\right)} \geq \Gamma(G). \tag{2.39}$$

Proof. By inspecting the final part of the proof of Theorem 2.5.1, we see that

$$\begin{aligned}\liminf_{\sigma\to\infty} \log \mathbb{E}W(\rho) &\geq \liminf_{\sigma\to\infty} \log\left(\frac{K-1}{2(K-\rho)} \mathbb{E}T_y^x(\sigma)\right) \\ &\geq \log\left(\frac{1}{4}\right) + \liminf_{\sigma\to\infty} \log \mathbb{E}T_y^x(\sigma). \end{aligned} \tag{2.40}$$

In order for each of the queues to be stable, the activity factor σ must diverge to infinity when $\rho \uparrow 1$ by virtue of Proposition 2.4.3. Thus, the previous inequality yields

$$\liminf_{\rho\uparrow 1} \log \mathbb{E}W(\rho) \geq \liminf_{\sigma\to\infty} \log \mathbb{E}T_y^x(\sigma) - \log 4.$$

Using again the fact that $\sigma > \frac{1}{1-\rho} - \frac{N}{K}$, we obtain

$$\begin{aligned}\liminf_{\rho\uparrow 1} \frac{\log \mathbb{E}W(\rho)}{\log(\frac{1}{1-\rho})} &= \liminf_{\rho\uparrow 1} \frac{\log \mathbb{E}W(\rho)}{\log(\frac{1}{1-\rho})} \frac{\log(\frac{1}{1-\rho})}{\log(\frac{1}{1-\rho} - \frac{N}{K})} \\ &= \liminf_{\rho\uparrow 1} \frac{\log \mathbb{E}W(\rho)}{\log(\frac{1}{1-\rho} - \frac{N}{K})} \\ &\geq \liminf_{\rho\uparrow 1} \frac{\log \mathbb{E}T_y^x(\sigma) - \log 4}{\log \sigma} \\ &= \liminf_{\sigma\to\infty} \frac{\log \mathbb{E}T_y^x(\sigma)}{\log \sigma} \overset{(2.38)}{\geq} \Gamma(G). \end{aligned}$$

□

Corollary 2.6.1 says that the long-run average delay $\mathbb{E}W(\rho)$ in a symmetric partite network G grows dramatically as the load ρ approaches 1, namely in view of (2.39) it must roughly scale at least as

$$\left(\frac{1}{1-\rho}\right)^{\Gamma(G)}, \quad \text{as } \rho \uparrow 1,$$

with $\Gamma(G) \geq 1$. This is to be contrasted with the usual linear scaling $\frac{1}{1-\rho}$ in conventional queueing networks. Furthermore, as we will see later in this section, the exponential rate of growth $\Gamma(G)$ usually increases with the size of the network G.

In order to obtain the asymptotic order of magnitude of the transition time T_y^x as in (2.38) and in this way find the value of the constant $\Gamma(G)$, it is convenient to look at the uniformized version of the activity process that can be put in the framework of Metropolis Markov chains, as we illustrated in Section 1.3. In Subsection 2.6.1 we show how the result for first hitting times can be translated back from discrete time to continuous time. In Subsection 2.6.2 we highlight the delay scaling results for the symmetric partite graphs described in Section 2.2.

2.6.1 *Uniformization of the activity process*

Consider the activity process $\{X(t)\}_{t\geq 0}$ with homogeneous rates on a general conflict graph G of N nodes. Recall that we assumed μ to be fixed, and when we write $\sigma \to \infty$ we allow for either $\nu \to \infty$, $p \downarrow 0$, or both. Define

$$\alpha := \liminf_{\sigma\to\infty} \frac{\log p}{\log \sigma} \geq 0.$$

Since we are interested in the regime $\sigma \to \infty$, we can assume without loss of generality that the activity factor σ is bigger than 1 and thus $\nu \geq p\mu$, which implies that

$$\max_{x,y\in\mathcal{X}} q(x,y) = N\nu.$$

Therefore we can consider the uniformization at rate $q_{\max} = N\nu$ for the activity process $\{X(t)\}_{t\geq 0}$. The Markov chain $\{\widetilde{X}(n)\}_{n\in\mathbb{N}}$ obtained in this way is in fact a Metropolis Markov chain (see Subsection 1.3.2 and Chapter 4), after the identification $\sigma = e^{\beta}$. For any state $x \in \mathcal{X}$ and subset $A \subseteq \mathcal{X}$, let τ_A^x be the first hitting time of the set A for the Markov chain $\{\widetilde{X}(n)\}_{n\in\mathbb{N}}$ started in x at time $n = 0$, i.e.

$$\tau_A^x(\sigma) := \inf\{n > 0 : \widetilde{X}(n) \in A\}.$$

Recall that $T_A^x(\sigma)$ denotes its continuous-time counterpart, namely the first hitting time of the subset A for the Markov process $\{X(t)\}_{t\geq 0}$ starting in x at

$t = 0$. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. exponential random variables with mean $\frac{1}{q_{\max}} = \frac{1}{N\nu}$. If $\widetilde{X}(0) = X(0)$ and $n(t) := \sup\{m : \sum_{n=1}^{m} Y_n \le t\}$, then

$$\{X(t)\}_{t\ge 0} \overset{d}{=} \{\widetilde{X}(n(t))\}_{t\ge 0}. \tag{2.41}$$

Thus the hitting times T_A^x and τ_A^x are closely related, namely

$$T_A^x \overset{d}{=} \sum_{n=1}^{\tau_A^x} Y_n,$$

and hence

$$\mathbb{E}T_A^x(\sigma) = \frac{1}{q_{\max}}\mathbb{E}\tau_A^x(\sigma) = \frac{1}{N\sigma p\mu}\mathbb{E}\tau_A^x(\sigma). \tag{2.42}$$

The next lemma shows we can deduce the asymptotic order of magnitude of the expected hitting time $\mathbb{E}T_A^x$ of the activity process $\{X(t)\}_{t\ge 0}$ if we know how to calculate that of the corresponding uniformized Markov chain $\{\widetilde{X}(n)\}_{n\in\mathbb{N}}$.

Lemma 2.6.2 (Hitting time scaling from discrete time to continuous time)**.** *Assume that there exists a positive constant* $\Gamma(x, A) > 0$ *such that*

$$\lim_{\sigma\to\infty} \frac{\log \mathbb{E}\tau_A^x(\sigma)}{\log \sigma} = \Gamma(x, A). \tag{2.43}$$

Then

$$\lim_{\sigma\to\infty} \frac{\log \mathbb{E}T_A^x(\sigma)}{\log \sigma} = \Gamma(x, A) - 1 + \alpha.$$

Proof. Equation (2.42) implies that

$$\log\left(\mu\mathbb{E}T_A^x(\sigma)\right) = \log \mathbb{E}\tau_A^x(\sigma) - \log\left(N\sigma p\right).$$

Since

$$\lim_{\sigma\to\infty} \frac{N\sigma p}{\log \sigma} = \lim_{\sigma\to\infty} \frac{\log N}{\log \sigma} + \lim_{\sigma\to\infty} \frac{\log \sigma}{\log \sigma} + \lim_{\sigma\to\infty} \frac{\log p}{\log \sigma} = 1 - \alpha,$$

we have

$$\begin{aligned}
\lim_{\sigma\to\infty} \frac{\log\left(\mu\mathbb{E}T_A^x(\sigma)\right)}{\log \sigma} &= \lim_{\sigma\to\infty} \left[\frac{\log \mathbb{E}\tau_A^x(\sigma)}{\log \sigma} - \frac{\log N\sigma p}{\log \sigma}\right] \\
&= \lim_{\sigma\to\infty} \frac{\log \mathbb{E}\tau_A^x(\sigma)}{\log \sigma} + \alpha - 1 \\
&\overset{(2.43)}{=} \Gamma(x, A) - 1 + \alpha.
\end{aligned}$$

□

2.6.2 *Average delay scaling results*

In this last subsection, we exhibit the values that the coefficient $\Gamma(G)$ takes when G is one of the symmetric partite graphs introduced in Section 2.2 and show how the average delay in the corresponding network scales when the load ρ approaches 1.

If G is a $2K \times 2L$ grid graph with periodic boundary conditions, we prove in Chapter 5 that $\Gamma(G) = 2\min\{K, L\}$, which means that for a square $\sqrt{N} \times \sqrt{N}$ grid graph of N nodes, with $\sqrt{N}$ an even integer, the long-run average delay $\mathbb{E}W(\rho)$ at every node scales as $\rho \uparrow 1$ roughly at least as

$$\left(\frac{1}{1-\rho}\right)^{\sqrt{N}}.$$

The value of the constant $\Gamma(G)$ is $\sqrt{N}/2$ if the grid graph has *open* boundary conditions. Indeed, we prove that $\Gamma(G) = \min\{K, L\}$ in the case of a $2K \times 2L$ grid graph with open boundary conditions.

In Chapter 6 we establish that $\Gamma(G) = \min\{K, 2L\}$ when G is a $2K \times 3L$ triangular grid graph with periodic boundary conditions. Therefore, the asymptotic scaling for $\mathbb{E}W(\rho)$ as $\rho \uparrow 1$ in this case reads

$$\left(\frac{1}{1-\rho}\right)^{\min\{K,2L\}}.$$

In the case of Turán graphs with N nodes and components of equal size N/K, we prove in Chapter 3 that $\Gamma(G) = N/K$. In view of Corollary 2.6.1, this fact implies that at every node the long-run average delay $\mathbb{E}W(\rho)$ scales as $\rho \uparrow 1$ roughly at least as

$$\left(\frac{1}{1-\rho}\right)^{N/K-1}.$$

We remark that the structure of the state space $\mathcal{X}$ corresponding to complete partite graphs allows for more detailed asymptotic analysis of transition times, to which Chapter 3 is devoted. This leads to sharper delay bounds for Turán graphs, which are presented in Section 3.4 together with a comparison with existing results, see [20, Theorem 1], for the same type of networks.

3 COMPLETE PARTITE GRAPHS

This chapter is devoted to the activity process on complete partite graphs, which provide a prototypical worst-case scenario for dense networks. In particular, we study the asymptotic behavior of hitting times as well as mixing properties in the regime where the activation rates grow large. This asymptotic regime in which nodes activate aggressively is relevant in highly loaded networks and gives rise to the pronounced starvation effects.

Section 3.1 presents a detailed model description and Section 3.2 gives an overview of the main results. In Section 3.3 we study the asymptotic behavior of the activity process within a single component of the conflict graph, which will be leveraged in Section 3.4 to prove the main results for the hitting times in conjunction with a geometric-sum representation of these latter. In Section 3.5 we briefly discuss the lower bounds for the average packet delay that follow from the framework developed in Chapter 2 and compare them with the existing literature. Section 3.6 is devoted to the analysis of the throughput starvation phenomenon, while in Section 3.7 we study the mixing time of the activity process. In Section 3.8 we sketch how the approach extends to scenarios where some of the nodes within the same component may interfere as well, relying on the same geometric-sum representation, but using more general asymptotic exponentiality results in [62] for the single-component behavior.

3.1 MODEL DESCRIPTION

We consider the saturated CSMA model described in Subsection 1.2.2 on *complete partite* conflict graphs. In such networks the nodes can be partitioned into K disjoint sets called *components*, such that two nodes interfere if and only if they belong to different components. More precisely, nodes in the same component do not share an edge, while each node has an edge with all nodes in all other components. Denote by $V_1, \dots, V_K$ the K components of G, with $K \in \mathbb{N}$ finite, and define $L_k := |V_k|$ as the size of component V_k. Note that the components $V_1, \dots, V_K$ are the K *maximal* independent sets of the graph G. Moreover, component V_k corresponds to a *maximum* independent set if and only if $L_k \geq L_j$ for all $j = 1, \dots, K$. Figure 3.1 shows an example of such a dense conflict graph, where $K = 5$ and the components have sizes $(L_1, L_2, L_3, L_4, L_5) = (3, 4, 5, 2, 6)$.

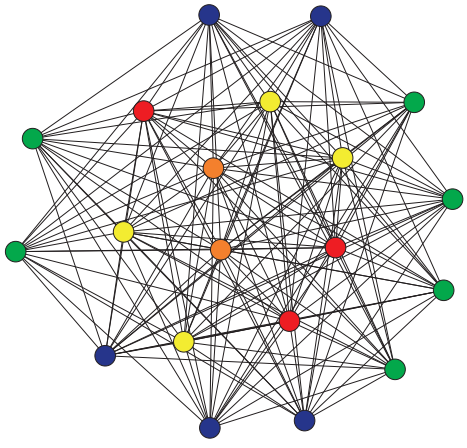


Figure 3.1: Example of complete K-partite conflict graph with $K = 5$

Figure 3.2a shows the state space diagram corresponding to the conflict graph in Figure 3.1. The network activity is described by the reversible Markov process $\{X(t)\}_{t\geq 0}$ defined in Subsection 1.2.2, with the additional assumption that the average transmission time at every node is equal to 1. In other words, we consider the Markov process $\{X(t)\}_{t\geq 0}$ on $\mathcal{X}$ with transition rates (1.2) where we take $\mu_i = 1$ for every $i = 1, \ldots, N$. The stationary distribution (1.3) then reads

$$\pi(x) = \lim_{t\to\infty} \mathbb{P}\Big(X(t) = x\Big) = \frac{\prod_{i\in V} \nu_i^{x_i}}{\sum_{y\in\mathcal{X}} \prod_{i\in V} \nu_i^{y_i}}, \quad x \in \mathcal{X}. \tag{3.1}$$

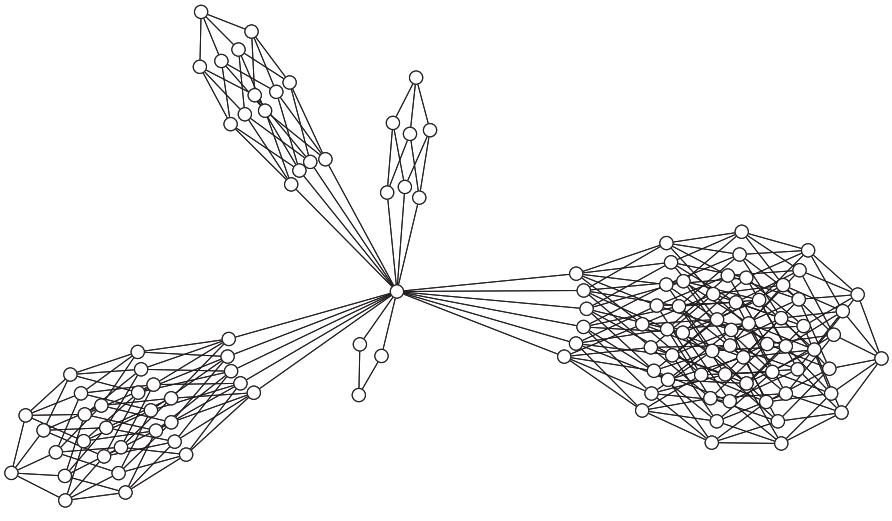
We further assume that the exponential rate at which a node activates depends only on a global aggressiveness parameter ν and on the component it belongs to, namely

$$\nu_i = f_k(\nu) \quad \text{if } i \in V_k,$$

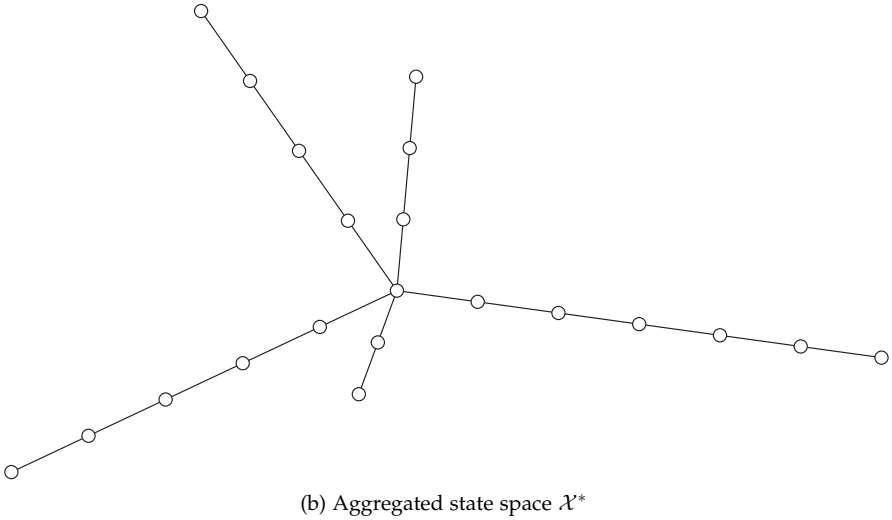
for some monotone function $f_k : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{\nu\to\infty} f_k(\nu) = \infty$. We will refer to the function $f_k(\cdot)$ as the *activation rate* of component V_k, for $k = 1, \ldots, K$.

In view of symmetry, all states with the same number of active nodes in a given component can be aggregated, and we only need to keep track of the number l of active nodes, if any, and the index k of the component V_k they belong to. More precisely, for each $k = 1, \ldots, K$ and $l = 1, \ldots, L_k$, we aggregate all the admissible joint activity states that have l active nodes in component V_k into a new state $(k, l) \in \mathcal{X}^*$. The unique state where no node is active will be denoted by $0 \in \mathcal{X}^*$. This state aggregation yields a new Markov process $\{X^*(t)\}_{t\geq 0}$ still indexed by the aggressiveness parameter ν on a star-shaped state space $\mathcal{X}^*$ with K "branches", where each branch emanates from a common root node and describes one of the components of the conflict graph.

Figure 3.2b shows the aggregated state space corresponding to the previous example.



(a) State space $\mathcal{X}$



(b) Aggregated state space $\mathcal{X}^*$

Figure 3.2: State space $\mathcal{X}$ and aggregated state space $\mathcal{X}^*$, for the conflict graph in Figure 3.1

For $k = 1, \ldots, K$, let $\mathcal{B}_k := \{(k, l) : 1 \leq l \leq L_k\}$ denote the *branch* of the state space $\mathcal{X}^*$ that corresponds to activity inside component V_k, where state (k, l) indicates that l nodes are active in component V_k. Then

$$\mathcal{X}^* = \bigcup_{k=1}^{K} \mathcal{B}_k \cup \{0\},$$

where 0 is the state in which all nodes are inactive. The transition rates of the process $\{X^*(t)\}_{t\geq 0}$ then read, for $k = 1, \ldots, K$

$$\begin{aligned} q(0,(k,1)) &= L_k f_k(\nu), \\ q((k,1),0) &= 1, \\ q((k,l),(k,l+1)) &= (L_k - l) f_k(\nu), \quad l = 1, \ldots, L_k - 1, \\ q((k,l),(k,l-1)) &= l, \quad l = 2, \ldots, L_k. \end{aligned}$$

The stationary distribution of the process $\{X^*(t)\}_{t\geq 0}$ can be easily derived from the original process $\{X(t)\}_{t\geq 0}$ and reads

$$\begin{aligned} \pi_0(\nu) &= \left(1 + \sum_{k=1}^{K} \sum_{l=1}^{L_k} \binom{L_k}{l} f_k(\nu)^l\right)^{-1}, \\ \pi_{(k,l)}(\nu) &= \pi_0(\nu) \binom{L_k}{l} f_k(\nu)^l, \quad l = 1, \ldots, L_k,\ k = 1, \ldots, K. \end{aligned} \tag{3.2}$$

We choose to use a different notation for the stationary distribution, putting the state as subscript and writing explicitly the dependence on ν, to avoid confusion with the notation of the original process $\{X(t)\}_{t\geq 0}$. The state (k, L_k) corresponds to the maximum activity state inside component V_k, which becomes the most likely state within the branch $\mathcal{B}_k$ as $\nu \to \infty$, in view of (3.2).

For any pair of states $x, y \in \mathcal{X}^*$, with a minor abuse of notation we denote also by $T_y^x(\nu)$ the *transition time* for the aggregated process $\{X^*(t)\}_{t\geq 0}$ with aggressiveness ν from state x to state y, i.e.

$$T_y^x(\nu) := \inf\{t > 0 : X^*(t) = x \mid X^*(0) = y\}.$$

Consider two different branches, say $\mathcal{B}_{k_1}$ and $\mathcal{B}_{k_2}$ with $k_1 \neq k_2$, and two activity states, $(k_1, l_1) \in \mathcal{B}_{k_1}$ and $(k_2, l_2) \in \mathcal{B}_{k_2}$. If the activity process $\{X^*(t)\}_{t\geq 0}$ starts in a state in branch $\mathcal{B}_{k_1}$, all nodes in V_{k_2} starve until branch $\mathcal{B}_{k_2}$ is hit. It is then natural to investigate the asymptotic behavior of transition time $T_{(k_2,l_2)}^{(k_1,l_1)}(\nu)$ as $\nu \to \infty$.

We now introduce a few parameters that will turn out to play a key role in the asymptotic distribution of the transition time. Define for $k \neq k_2$,

$$\gamma_k := \lim_{\nu\to\infty} \frac{f_k(\nu)^{L_k}}{\sum_{j\neq k_2} f_j(\nu)^{L_j}}. \tag{3.3}$$

To avoid technicalities, we assume throughout that all parameters γ_k are well defined. In view of (3.2), γ_k may be interpreted as the stationary fraction of time that the activity process spends in branch $\mathcal{B}_k$ as $\nu \to \infty$, excluding the target branch $\mathcal{B}_{k_2}$. As it turns out, γ_k also equals the fraction of time that the activity process spends in branch $\mathcal{B}_k$ during the transition from 0 to (k_2, l_2) as $\nu \to \infty$.

Branch $\mathcal{B}_k$ is called *dominant* if $\gamma_k > 0$ and we let $K_* := \{k \neq k_2 : \gamma_k > 0\}$ be the index set of all *dominant branches*. Note that, by construction, the set

K_* is never empty and thus there is always at least one dominant branch. To avoid confusion, we remark that the notion of dominant branch is slightly different from that of dominant state given in Chapter 2. Indeed, due to the heterogeneous activation rates, a branch $\mathcal{B}_k$ could be dominant even if the corresponding component V_k does not have maximum size (and vice versa).

3.2 MAIN RESULTS

In this section we present our main results, which are all related to the asymptotic behavior of the transition time $T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)$ from a state (k_1, l_1) to a state (k_2, l_2), with $k_1 \neq k_2$, in the asymptotic regime of a large activation rate ν.

Our first result characterizes the asymptotic order of magnitude of the mean transition time in terms of the activation rates and the network structure. For any two real-valued functions $f(\cdot)$ and $g(\cdot)$, let $f(\nu) \sim g(\nu)$ as $\nu \to \infty$ indicate that $\lim_{\nu\to\infty} f(\nu)/g(\nu) = 1$ and, similarly, let $f(\nu) \gtrsim g(\nu)$ as $\nu \to \infty$ indicate that $\lim_{\nu\to\infty} f(\nu)/g(\nu) \geq 1$.

Theorem 3.2.1 (Asymptotic average transition time)**.** *If $k_1 \neq k_2$, then*

$$\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu) \sim \frac{1}{L_{k_1} f_{k_1}(\nu)^{L_{k_1}-1}} + \frac{1}{L_{k_2} f_{k_2}(\nu)} \sum_{k\in K_*} f_k(\nu)^{L_k}, \quad \text{as } \nu \to \infty. \tag{3.4}$$

The first term on the right-hand side of (3.4) corresponds to the asymptotic mean *escape time* $\mathbb{E}T^{(k_1,l_1)}_0(\nu)$ from the initial branch $\mathcal{B}_{k_1}$, while the second term describes the contribution of the mean time spent visiting dominant branches, possibly including branch $\mathcal{B}_{k_1}$ as well. Let

$$\alpha := \lim_{\nu\to\infty} \frac{\mathbb{E}T^{(k_1,l_1)}_0(\nu)}{\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)} \in [0,1] \tag{3.5}$$

denote the relative weight of $\mathcal{B}_{k_1}$.

Our second result gives the asymptotic distribution of the transition time $T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)$ scaled by its mean as $\nu \to \infty$.

Theorem 3.2.2 (Asymptotic distribution of the scaled transition time)**.** *If $k_1 \neq k_2$, then*

$$\frac{T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)}{\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)} \xrightarrow{d} Z, \quad \text{as } \nu \to \infty.$$

The random variable Z can be expressed as

$$Z \overset{d}{=} \alpha Y + (1-\alpha)W,$$

where the random variable Y is exponentially distributed with unit mean and the random variable W is independent of Y with Laplace transform (3.21).

The random variable W has a complicated distribution presented later, which depends on the sizes and activation rates of the dominant branches only. The possible distributions of Z are summarized in Table 3.1 in Section 3.4. In several cases the distribution of Z is exponential, which may be expected in view of the connection with many exponentiality results for the occurrence of rare events [2, 3, 4, 62, 77, 79]. In addition, we identify several cases that lead to *non*-exponentiality, typically due to the fact that the activity process spends a substantial period in branches other than k_1 and k_2.

Our third result concerns the starvation phenomenon and in particular gives an asymptotic lower bound on the probability of throughput starvation. For $k = 1, \ldots, K$, define the random variable

$$\tau_k(t) := \int_0^t \mathbb{1}_{\{X^*(s) \in \mathcal{B}_k\}} \, \mathrm{d}s,$$

which measures how much time the activity process $\{X^*(t)\}_{t \geq 0}$ spends in branch $\mathcal{B}_k$ during the interval $[0, t]$. We can think of $\tau_k(t)$ as a measure of the *throughput* of component V_k over the time interval $[0, t]$. We speak of *complete starvation* or *zero throughput* of component V_k in $[0, t]$ when $\tau_k(t) = 0$. The next theorem provides insight into the time scales at which throughput starvation occurs for a component of the network.

Theorem 3.2.3 (Throughput starvation)**.** *Assume $X(0) = (k_1, l_1)$ and $k_2 \neq k_1$. If $t(\nu) \sim \omega \mathbb{E}T_{(k_2,1)}^{(k_1,l_1)}(\nu)$, with $\omega \in \mathbb{R} \cup \{0\}$, then*

$$\lim_{\nu \to \infty} \mathbb{P}(\tau_{k_2}(t(\nu)) = 0) = \mathbb{P}(Z \geq \omega), \tag{3.6}$$

where Z is the asymptotic distribution of the scaled transition time introduced in Theorem 3.2.1. In particular, if $t(\nu) = o(\mathbb{E}T_{(k_2,1)}^{(k_1,l_1)}(\nu))$, then

$$\lim_{\nu \to \infty} \mathbb{P}(\tau_{k_2}(t(\nu)) = 0) = 1,$$

i.e. all nodes in V_{k_2} have zero throughput for a period of length $t(\nu)$ with probability one as $\nu \to \infty$.

The limit in (3.6) says that, even if there is long-term fairness among the components, for large values of ν all nodes in V_{k_2} will face starvation on all time scales smaller than the mean transition time from the initial component to V_{k_2}.

For the Markov process at hand, slow transitions and starvation effects are intimately related with the mixing time. Indeed, due to the complete partite structure of the conflict graph, the process is bound to be stuck in one of the dominant branches, leading to slow convergence to equilibrium. In Section 3.7 we define the mixing time in terms of the total variation distance from stationarity, and prove a lower bound for a large enough activation parameter ν. This lower bound (see Proposition 3.7.1) indicates that the mixing time of the process is at least as large as the mean escape time from the dominant branch, which establishes a direct connection between transition times and mixing times.

3.3 HITTING TIMES WITHIN A SINGLE BRANCH

In this section we study the activity process within a single component, where it behaves as a birth-and-death process, bringing the asymptotic behavior within the realm of classical results. In Section 3.4 we will leverage these results in conjunction with a geometric-sum representation to prove both Theorems 3.2.1 and 3.2.2.

We start by presenting a few results for the case where the two states (k_1, l_1) and (k_2, l_2) belong to the same branch, i.e. $k_1 = k_2$, and $l_1 > l_2$. In this case, the presence of the other components does not affect the transition time, and hence we focus on a single branch, dropping the component index until further notice. Within a single component of size L, the process $\{X^*(t)\}_{t\geq 0}$ evolves as an elementary birth-and-death process on the state space $\{L, L-1, \dots, 1, 0\}$, so we can exploit several classical results for such processes. If we denote by $f(\nu)$ the activation rate for this component as a function of ν, then the transition rates read

$$q(l, l+1) = (L-l)f(\nu), \qquad l = 0, \dots, L-1,$$
$$q(l, l-1) = l, \quad l = 1, \dots, L.$$

3.3.1 *Asymptotic growth rate*

We first show how the mean transition time scales with the aggressiveness parameter ν.

Proposition 3.3.1 (Asymptotic transition time average within a single branch)**.** *For $L \geq l_1 > l_2 \geq 0$,*

$$\mathbb{E}T_{l_2}^{l_1}(\nu) \sim \frac{l_2!(L-l_2-1)!}{L!} f(\nu)^{L-l_2-1}, \quad \text{as } \nu \to \infty.$$

Proof. Observe that $\mathbb{E}T_{l_2}^{l_1}(\nu) = \sum_{l=l_1}^{l_2+1} \mathbb{E}T_{l-1}^{l}(\nu)$, so we can exploit a general result for birth-and-death processes [76], which in the present case says that, for $0 < l \leq L$,

$$\mathbb{E}T_{l-1}^{l}(\nu) = \frac{1}{l} \sum_{n=l}^{L} \frac{\pi_n(\nu)}{\pi_l(\nu)}.$$

Now (3.2) implies that $\pi_n(\nu) = o(\pi_L(\nu))$ as $\nu \to \infty$ for all $n = l, \dots, L-1$, so that

$$\mathbb{E}T_{l-1}^{l}(\nu) \sim \frac{1}{l} \frac{\pi_L(\nu)}{\pi_l(\nu)} = \frac{(l-1)!(L-l)!}{L!} f(\nu)^{L-l}, \quad \text{as } \nu \to \infty.$$

Thus $\mathbb{E}T_{l-1}^{l}(\nu) = o(\mathbb{E}T_{l_2}^{l_2+1}(\nu))$ as $\nu \to \infty$ for all $l = l_1, \dots, l_2$, and hence $\mathbb{E}T_{l_2}^{l_1}(\nu) \sim \mathbb{E}T_{l_2}^{l_2+1}(\nu)$ as $\nu \to \infty$ and the result follows. □

In order to gain insight in starvation effects, we are particularly interested in the time for the activity process to reach the center state 0, referred to as *escape time*, because at such points in time nodes in other components have an opportunity to activate. Proposition 3.3.1 shows that

$$\mathbb{E}T_0^{l_1}(\nu) \sim \frac{1}{L} f(\nu)^{L-1}, \quad \text{as } \nu \to \infty. \tag{3.7}$$

Hence, the mean escape time grows asymptotically as a power of $f(\nu)$, with the exponent equal to the component size minus one, and independent of the starting state l_1.

3.3.2 *Asymptotic exponentiality*

We now turn to the scaled escape time, and show that it has an asymptotically exponential distribution. We will leverage the following well-known result for birth-and-death processes, which is commonly attributed to Keilson [78] or Karlin and McGregor [75].

Theorem 3.3.2 (Absorption time distribution for birth-and-death processes)**.** *Consider a birth-and-death process with generator matrix Q on $\{0, \dots, L\}$ started at state L. Assume that 0 is an absorbing state, and that the birth rates $\{\lambda_i\}_{i=1,\dots,L-1}$ and death rates $\{\mu_i\}_{i=1,\dots,L}$ are positive. Then the absorption time in state 0 is distributed as the sum of L independent exponential random variables whose rate parameters are the L nonzero eigenvalues of $-Q$.*

Let $Q(\nu)$ be the generator matrix of the birth-and-death process $\{X^*(t)\}_{t\geq 0}$ on the state space $\{L, L-1, \dots, 1, 0\}$, with 0 an absorbing state. Denote by $\{\theta_i(\nu)\}_{i=1,\dots,L}$ the non-zero eigenvalues of $-Q(\nu)$. It is known [97] that these eigenvalues are distinct, real and strictly positive, so we denote $0 < \theta_1(\nu) < \theta_2(\nu) < \cdots < \theta_L(\nu)$. Theorem 3.3.2 gives

$$T_0^L(\nu) \stackrel{d}{=} \sum_{i=1}^{L} Y_i(\nu), \tag{3.8}$$

with $Y_1(\nu), \dots, Y_L(\nu)$ independent and exponentially distributed random variables with $\mathbb{E}Y_i(\nu) = 1/\theta_i(\nu)$.

The following lemma relates the growth rates of the eigenvalues as $\nu \to \infty$ to the mean escape time $\mathbb{E}T_0^L(\nu)$.

Lemma 3.3.3. $\lim_{\nu\to\infty} \theta_i(\nu) \cdot \mathbb{E}T_0^L(\nu) = 1$ *if* $i = 1$ *and* ∞ *if* $i = 2, \dots, L$.

The proof of Lemma 3.3.3 is presented in Appendix 3.A, and exploits detailed information about the growth rates of the eigenvalues obtained via symmetrization and the Gershgorin circle theorem. Lemma 3.3.3 shows that the smallest eigenvalue $\theta_1(\nu)$ becomes dominant as $\nu \to \infty$, but also proves the asymptotic exponentiality of the escape time. Indeed, denoting by $\mathcal{L}_X(s) =$

$\mathbb{E}(e^{-sX})$, with $\mathrm{Re}(s) > 0$, the Laplace transform of a random variable X, (3.8) gives

$$\mathcal{L}_{T_0^L(\nu)/\mathbb{E}T_0^L(\nu)}(s) = \prod_{i=1}^{L} \Big(1 + \frac{s}{\theta_i(\nu) \cdot \mathbb{E}T_0^L(\nu)}\Big)^{-1}.$$

Lemma 3.3.3 implies that

$$\lim_{\nu\to\infty} \mathcal{L}_{T_0^L(\nu)/\mathbb{E}T_0^L(\nu)}(s) = \frac{1}{1+s}.$$

The continuity theorem for Laplace transforms then yields that the scaled escape time has an asymptotically exponential distribution as stated in the next theorem, where $\mathrm{Exp}(\lambda)$ denotes an exponentially distributed random variable with mean $1/\lambda$.

Theorem 3.3.4 (Asymptotic exponentiality of the scaled escape time)**.**

$$\frac{T_0^L(\nu)}{\mathbb{E}T_0^L(\nu)} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \nu \to \infty.$$

This result can be understood as follows. For large ν, the probability of hitting state 0 before the first return to state L becomes small. So the time $T_0^L(\nu)$ consists of a geometrically distributed number of excursions from L which return to L without hitting 0, followed by the remaining part of the excursion that hits 0. Hence, apart from this final part, $T_0^L(\nu)$ is the sum of a large geometrically distributed number of i.i.d. random variables, which indeed is expected to be exponential.

The fact that the time until the first occurrence of a rare event is asymptotically exponential, is a widely observed phenomenon [79]. Exponentiality of the hitting time of some subset B of the state space typically arises when the probability of hitting B in a single regenerative cycle is 'small', and the cycle lengths are 'not too heavy tailed' [62, 79]. This is also true for our situation, and hence an alternative proof of Theorem 3.3.4 can be obtained using [62, Thm. 1] (which is a generalized version of [77]). We do not use the probabilistic approach in [62] here, because the special case of a birth-and-death process allows for explicit analysis. Let us finally remark that for reversible Markov processes similar exponentiality results were established in [2, 3, 4]. Aldous [2] showed that a result like Theorem 3.3.4 can be expected when the underlying Markov process converges rapidly to stationarity. This is indeed the case for the Markov process $\{X^*(t)\}_{t\geq 0}$ restricted to a single branch.

To extend Theorem 3.3.4 to the case of a general starting state $0 < l \leq L$, we need the following technical lemma, whose proof is given in Appendix 3.B.

Lemma 3.3.5. *Let $T(\nu), U(\nu), V(\nu), W(\nu)$ be non-negative random variables. Consider the properties*

(1) $\lim_{\nu\to\infty} \mathbb{E}V(\nu)/\mathbb{E}U(\nu) = \lim_{\nu\to\infty} \mathbb{E}W(\nu)/\mathbb{E}U(\nu) = 0.$

(2) *For every* $\nu > 0$, $U - V \leq_{\mathrm{st}} T \leq_{\mathrm{st}} U + W$, *i.e. for every* $t \geq 0$,

$$\mathbb{P}\Big(U - V > t\Big) \leq \mathbb{P}\Big(T > t\Big) \leq \mathbb{P}\Big(U + W > t\Big).$$

(3) $U(\nu)/\mathbb{E}U(\nu) \xrightarrow{d} Z$ *as* $\nu \to \infty$, *where* Z *is a continuous random variable independent of* ν.

Then,

(i) *If properties* (1) *and* (2) *hold, then* $\lim_{\nu\to\infty} \mathbb{E}T(\nu)/\mathbb{E}U(\nu) = 1$.

(ii) *If properties* (1), (2) *and* (3) *hold, then* $T(\nu)/\mathbb{E}T(\nu) \xrightarrow{d} Z$, *as* $\nu \to \infty$.

Proposition 3.3.6 (Asymptotic exponentiality of the scaled escape time)**.** *For any* $0 < l \leq L$,

$$\frac{T_0^l(\nu)}{\mathbb{E}T_0^l(\nu)} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \nu \to \infty.$$

Proof. The birth-and-death structure of the process and the strong Markov property yield the stochastic identity $T_0^L(\nu) \overset{d}{=} T_l^L(\nu) + T_0^l(\nu)$, which gives the stochastic bounds $T_0^L(\nu) - T_l^L(\nu) \leq_{\mathrm{st}} T_0^l(\nu) \leq_{\mathrm{st}} T_0^L(\nu)$ (the two terms in the lower bound being dependent). By virtue of Theorem 3.3.4 the limit in distribution $T_0^L(\nu)/\mathbb{E}T_0^L(\nu) \xrightarrow{d} \mathrm{Exp}(1)$ holds as $\nu \to \infty$. In order to complete the proof, we can then use Lemma 3.3.5, taking $U(\nu) = T_0^L(\nu)$, $V(\nu) = T_l^L(\nu)$ and $W(\nu) = 0$. The condition which needs to be checked is $\lim_{\nu\to\infty} \mathbb{E}V(\nu)/\mathbb{E}U(\nu) = 0$, which follows directly from Proposition 3.3.1. □

3.3.3 *More general coefficients and applications*

We can extend our analysis to more general activation and deactivation dynamics inside a single branch, described by

$$q(l, l+1) = a_l f(\nu), \quad l = 1, \dots, L-1,$$
$$q(l, l-1) = d_l, \quad l = 2, \dots, L,$$

where a_l, d_l are positive real coefficients. Specifically, Proposition 3.3.1 can be generalized to the following result. For $L \geq l_1 > l_2 \geq 0$,

$$\mathbb{E}T_{l_2}^{l_1}(\nu) \sim \frac{1}{d_{l_2+1}} \Big(\prod_{i=l_2+1}^{L-1} \frac{a_i}{d_{i+1}} \Big) f(\nu)^{L-l_2-1}, \quad \text{as } \nu \to \infty. \tag{3.9}$$

Also Lemma 3.3.3 and thus Proposition 3.3.6 can be shown to hold for these more general rates (see Appendix 3.A).

These results for general coefficients have some interesting applications, beyond the model considered in this chapter. We present here an application to

an $M/M/c/c$ model in two extreme regimes: Heavy-traffic and light-traffic conditions. Let $\{M(t)\}_{t\geq 0}$ be the continuous-time Markov process on the state space $\{0,1,\ldots,c\}$, describing the number of busy servers at time t. Suppose that the service rate of each server is 1 and the arrival rate is ν, which will grow large in a heavy-traffic regime. The escape time $T_0^s(\nu)$, choosing $a_n = 1$ and $d_n = n$, $n = 1,\ldots,c$, then describes the time it takes for this system to *drain* (i.e. to have all the servers idle) when starting with $s \geq 1$ busy servers. In other words,

$$T_0^s(\nu) := \inf\{t > 0 \ : \ M(t) = 0 \mid M(0) = s\}.$$

Then (3.9) gives

$$\mathbb{E}T_0^s(\nu) \sim \frac{\nu^{c-1}}{c!}, \quad \text{as } \nu \to \infty,$$

which does not depend on the starting state $s \geq 1$. Furthermore, the scaled drain time obeys

$$\frac{T_0^s(\nu)}{\mathbb{E}T_0^s(\nu)} \xrightarrow{d} \text{Exp}(1), \quad \text{as } \nu \to \infty.$$

We now turn to a light-traffic regime where the arrival rate λ becomes small, and let $s \in \{0,1,\ldots,c\}$ count instead the number of idle servers. The quantity of interest in this case is the time $T_b(\lambda)$ it takes for the system to fill up, i.e. to reach the situation where none of the servers is idle, having started with $s \geq 1$ idle servers. This hitting time, measured in time units $1/\lambda$, corresponds again to the absorption time $T_0^s(\nu)$, with $f(\nu) = 1/\lambda$, $a_n = c - n$ and $d_n = 1$, $n = 1,\ldots,c$. The light-traffic regime is characterized by $\lambda \downarrow 0$, which corresponds to the limiting regime $\nu \to \infty$. Then (3.9) gives

$$\mathbb{E}T_b\left(\frac{1}{\nu}\right) = \nu \cdot \mathbb{E}T_0^s(\nu) \sim (c-1)!\,\nu^{c-1}, \quad \text{as } \nu \to \infty,$$

which does not depend on the starting state $s \geq 1$. The asymptotic scaled fill time obeys

$$\frac{T_b(\frac{1}{\nu})}{\mathbb{E}T_b(\frac{1}{\nu})} = \frac{T_0^s(\nu)}{\mathbb{E}T_0^s(\nu)} \xrightarrow{d} \text{Exp}(1), \quad \text{as } \nu \to \infty.$$

3.4 PROOFS OF THEOREMS 3.2.1 AND 3.2.2

In this section we investigate the asymptotic behavior of the transition time $T_{(k_2,l_2)}^{(k_1,l_1)}(\nu)$ as $\nu \to \infty$ for any pair of states (k_1,l_1) and (k_2,l_2), with $k_1 \neq k_2$. In Subsection 3.4.1 we provide a stochastic representation of the transition time, which we use to derive the asymptotic mean transition time in Subsection 3.4.2 leading to Theorem 3.2.1. In Subsection 3.4.3 we will obtain the

asymptotic distribution of the scaled transition time leading to Theorem 3.2.2. In Subsection 3.4.4 we consider in detail the random variable W that occurs in Theorem 3.2.2. We give an overview of all possible forms of asymptotic behavior and the conditions under which they occur in Subsection 3.4.5.

3.4.1 *Stochastic representation of the transition time*

Consider the evolution of the process while it makes a transition from a state (k_1, l_1) to a state (k_2, l_2) and define the following random variables:

- $T_0^{(k_1,1)}(0)$: time to reach state 0 after state $(k_1, 1)$ is visited for the first time;
- $N_k(\nu)$: number of times the process makes a transition $0 \to (k, 1)$, $k \neq k_2$, before the first transition $0 \to (k_2, 1)$ occurs;
- $\hat{T}^0_{(k,1)}(i)$: time spent in state 0 before the i-th transition to state $(k, 1)$, $k \neq k_2$, $i = 1, \ldots, N_k(\nu)$;
- $\hat{T}^0_{(k_2,1)}$: time spent in state 0 before the first transition to state $(k_2, 1)$;
- $T_0^{(k,1)}(i)$: time to return to state 0 after the i-th transition to state $(k, 1)$, $k \neq k_2$, $i = 1, \ldots, N_k(\nu)$;
- $T^{(k_2,1)}_{(k_2,l_2)}$: time to reach state (k_2, l_2) after the first hitting of state $(k_2, 1)$.

Note that the dependence on the parameter ν of all the random variables representing time durations we just introduced is suppressed for compactness. With the above definitions, it is readily seen that the following stochastic representation holds.

Proposition 3.4.1 (Stochastic representation of the transition time)**.** *The transition time $T^{(k_1,l_1)}_{(k_2,l_2)}$ can be represented as*

$$T^{(k_1,l_1)}_{(k_2,l_2)} \stackrel{d}{=} T^{(k_1,l_1)}_{(k_1,1)} + T_0^{(k_1,1)}(0) + \sum_{k \neq k_2} \sum_{i=1}^{N_k(\nu)} \left(\hat{T}^0_{(k,1)}(i) + T_0^{(k,1)}(i) \right) + \hat{T}^0_{(k_2,1)} + T^{(k_2,1)}_{(k_2,l_2)}, \tag{3.10}$$

where all the random variables representing time durations are mutually independent as well as independent of the random variables $N_k(\nu)$, $k \neq k_2$.

Denote $F(\nu) = \sum_{k=1}^K L_k f_k(\nu)$. The random variables $T_0^{(k,1)}(i)$ are i.i.d. copies of $T_0^{(k,1)}$, $i = 1, \ldots, N_k(\nu)$, $k \neq k_2$, while the random variables $\hat{T}^0_{(k_2,1)}$ and $\hat{T}^0_{(k,1)}(i)$, $k \neq k_2$, $i = 1, \ldots, N_k(\nu)$, are i.i.d. copies of $\mathrm{Exp}(F(\nu))$, which is the residence time in state 0. Write $X \stackrel{d}{=} \mathrm{Geo}(p)$ when X is a random variable with geometric distribution $\mathbb{P}(X = n) = p(1-p)^n$, $n \in \mathbb{N} \cup \{0\}$. Define the

random variable $N(\nu) := \sum_{k \neq k_2} N_k(\nu)$, counting the total number of entrances in branches other than k_2 before hitting the target branch $\mathcal{B}_{k_2}$. For all $k = 1, \dots, K$, denote $p_k(\nu) := L_k f_k(\nu)/F(\nu)$. Clearly,

$$N(\nu) \stackrel{d}{=} \mathrm{Geo}(p_{k_2}(\nu)),$$

and the marginal distribution of $N_k(\nu)$ is $\mathrm{Geo}(\frac{p_{k_2}(\nu)}{p_{k_2}(\nu)+p_k(\nu)})$.

We want to distinguish the branches that significantly affect the dynamics of the process (and hence the transition time) from those that do not. The quantity $\mathbb{E}N_k(\nu) \cdot \mathbb{E}T_0^{(k,1)}(\nu)$, for $k \neq k_2$, is the mean time that the process spends in branch $\mathcal{B}_k$ along the transition $0 \to (k_2, l_2)$. Note that Proposition 3.3.1 gives

$$\mathbb{E}T_0^{(k,1)}(\nu) \sim \frac{1}{L_k} f_k(\nu)^{L_k - 1}, \quad \text{as } \nu \to \infty, \tag{3.11}$$

and

$$\mathbb{E}N_k(\nu) = \frac{p_k(\nu)}{p_{k_2}(\nu)} = \frac{L_k f_k(\nu)}{L_{k_2} f_{k_2}(\nu)}. \tag{3.12}$$

Therefore,

$$\frac{\mathbb{E}N_k(\nu) \cdot \mathbb{E}T_0^{(k,1)}(\nu)}{\mathbb{E}N_j(\nu) \cdot \mathbb{E}T_0^{(j,1)}(\nu)} \sim \frac{f_k(\nu)^{L_k}}{f_j(\nu)^{L_j}}, \quad \text{as } \nu \to \infty,$$

which shows that indeed only the visits to dominant branches asymptotically contribute to the mean transition time.

3.4.2 *Asymptotic mean transition time*

We present here the proof of Theorem 3.2.1. Consider the stochastic representation (3.10) of the transition time $T_{(k_1,l_1),(k_2,l_2)}(\nu)$. Proposition 3.3.1 implies that

$$\mathbb{E}T_{(k_1,1)}^{(k_1,l_1)}(\nu) \sim \frac{1}{L_{k_1}(L_{k_1} - 1)} f_{k_1}(\nu)^{L_{k_1} - 2}, \quad \text{as } \nu \to \infty,$$

and that

$$\mathbb{E}T_0^{(k_1,1)}(\nu) \sim \frac{1}{L_{k_1}} f_{k_1}(\nu)^{L_{k_1} - 1}, \quad \text{as } \nu \to \infty.$$

Hence $\mathbb{E}T_{(k_1,1)}^{(k_1,l_1)}(\nu) = o(\mathbb{E}T_0^{(k_1,1)}(\nu))$ as $\nu \to \infty$. Moreover $\mathbb{E}\hat{T}_{(k,1)}^0(\nu) = o(1)$. Lemma 3.4.2 below implies that $\mathbb{E}T_{(k_2,l_2)}^{(k_2,1)}(\nu) = o(\mathbb{E}T_{(k_2,l_2)}^{(k_1,l_1)}(\nu))$. The asymptotic relation (3.4) then follows using the definition of K_*, the asymptotic estimate (3.11) and the identity (3.12). Lemma 3.4.2 guarantees that once the process has entered the target branch $\mathcal{B}_{k_2}$, even if it may exit from it, the mean time it takes to reach the target state (k_2, l_2) is negligible with respect to the mean overall transition time.

Lemma 3.4.2.

$$\mathbb{E}T^{(k_2,1)}_{(k_2,l_2)}(\nu) = o\left(\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)\right), \quad \text{as } \nu \to \infty.$$

Proof. Consider the event

$$\begin{aligned}\mathcal{E}(\nu) &= \{\text{the first } l_2 - 1 \text{ transitions are all towards the state } (k_2, l_2)\} \\ &= \bigcap_{i=1}^{l_2-1} \{\text{the } i\text{-th transition is from } (k_2, i) \text{ to } (k_2, i+1)\}.\end{aligned}$$

Exploiting the fact that all these events are independent, we can compute

$$\begin{aligned}\mathbb{P}(\mathcal{E}(\nu)) &= \prod_{i=1}^{l_2-1} \mathbb{P}\Big(\text{the } i\text{-th transition is from } (k_2, i) \text{ to } (k_2, i+1)\Big) \\ &= \prod_{i=1}^{l_2-1} \frac{(L_{k_2} - i) f_{k_2}(\nu)}{(L_{k_2} - i) f_{k_2}(\nu) + i'},\end{aligned}$$

and clearly $\lim_{\nu\to\infty} \mathbb{P}(\mathcal{E}(\nu)) = 1$. We have that

$$\mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}(\nu)\} = \sum_{m=1}^{l_2-1} \frac{1}{(L_{k_2} - m) f_{k_2}(\nu) + m} =: g(\nu),$$

where $g(\nu) \downarrow 0$ as $\nu \to \infty$. For $n = 1, \dots, l_2 - 1$, consider the event

$$\mathcal{E}^c_n(\nu) = \{\text{the first transition towards state 0 is the } n\text{-th one}\}.$$

Note that the event $\mathcal{E}^c(\nu)$ can be decomposed as $\mathcal{E}^c(\nu) = \bigcup_{n=1}^{l_2-1} \mathcal{E}^c_n(\nu)$. Using the events $\mathcal{E}(\nu)$ and $\mathcal{E}^c_n(\nu)$, we can write

$$\begin{aligned}\mathbb{E}T^{(k_2,1)}_{(k_2,l_2)}(\nu) =& \mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}(\nu)\}\mathbb{P}(\mathcal{E}(\nu)) \\ &+ \sum_{n=1}^{l_2-1} \mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}^c_n(\nu)\}\mathbb{P}(\mathcal{E}^c_n(\nu)).\end{aligned} \tag{3.13}$$

If the event $\mathcal{E}^c_1(\nu)$ occurs, then the first transition is towards state 0. In this case, we have

$$\begin{aligned}\mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}^c_1(\nu)\} &\le \mathbb{E}\{T^{(k_2,1)}_0(\nu) \mid \mathcal{E}^c_1(\nu)\} + \mathbb{E}T^0_{(k_2,l_2)}(\nu) \\ &\le \frac{1}{(L_{k_2} - 1) f_{k_2}(\nu) + 1} + \mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu),\end{aligned} \tag{3.14}$$

and

$$\mathbb{P}(\mathcal{E}^c_1(\nu)) = \frac{1}{(L_{k_2} - 1) f_{k_2}(\nu) + 1}. \tag{3.15}$$

Moreover, for $n = 2, \dots, l_2 - 1$, we have

$$\begin{aligned}\mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}^c_n(\nu)\} &\le \mathbb{E}\{T^{(k_2,1)}_{(k_2,n-1)}(\nu) \mid \mathcal{E}^c_n(\nu)\} + \mathbb{E}T^{(k_2,n-1)}_{(k_2,l_2)}(\nu) \\ &\le \mathbb{E}\{T^{(k_2,1)}_{(k_2,n-1)}(\nu) \mid \mathcal{E}(\nu)\} + \mathbb{E}T^{0}_{(k_2,l_2)}(\nu) \\ &\le \mathbb{E}\{T^{(k_2,1)}_{(k_2,l_2)}(\nu) \mid \mathcal{E}(\nu)\} + \mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu). \qquad (3.16)\end{aligned}$$

From (3.13)-(3.16) it follows that

$$\begin{aligned}\mathbb{E}T^{(k_2,1)}_{(k_2,l_2)}(\nu) &\le g(\nu) + \mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)\mathbb{P}(\mathcal{E}^c(\nu)) + \mathbb{E}\{T^{(k_2,1)}_0(\nu) \mid \mathcal{E}^c_1(\nu)\} \\ &\le 2g(\nu) + \mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)\mathbb{P}(\mathcal{E}^c(\nu)).\end{aligned}$$

We divide both sides by $\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)$, which is greater than 1 for ν sufficiently large, thanks to (3.7). Since $g(\nu)$ and $\mathbb{P}(\mathcal{E}^c(\nu))$ are both $o(1)$, the proof of the lemma is complete. □

3.4.3 *Asymptotic distribution of the transition time*

We now turn to the proof of Theorem 3.2.2. It is clear that only the dominant branches that asymptotically contribute to the expected magnitude of the transition time will play a role, possibly along with the escape time from the initial branch. As we will show, the various dominant branches may play different roles, depending on whether the expected number of visits during the transition time is zero, $O(1)$ or infinite in the limit as $\nu \to \infty$. We introduce

$$A(\nu) := T^{(k_1,1)}_0 \quad \text{and} \quad B(\nu) := \sum_{k \in K_*} \sum_{i=1}^{N_k(\nu)} T^{(k,1)}_0(i), \qquad (3.17)$$

whose mean values correspond to the two terms at the right-hand side of (3.4). From the definition (3.5) of the coefficient α it follows that

$$\alpha = \lim_{\nu \to \infty} \frac{\mathbb{E}A(\nu)}{\mathbb{E}A(\nu) + \mathbb{E}B(\nu)}.$$

When $\alpha = 0$ the term $A(\nu)$ becomes asymptotically negligible compared to $B(\nu)$, while the opposite holds when $\alpha = 1$. Proposition 3.3.6 already describes the asymptotic behavior of $A(\nu)$ after scaling. We need to understand the asymptotic behavior of $B(\nu)$ and for this purpose, it will be convenient to use a slightly different representation for it.

Define $p_*(\nu) := \sum_{k \in K_*} p_k(\nu)$ and $\hat{p}(\nu) := \frac{p_*(\nu)}{p_{k_2}(\nu) + p_*(\nu)}$. Introduce the random variable $N_*(\nu) := \text{Geo}(1 - \hat{p}(\nu))$, which represents the number of visits to the dominant branches, before entering the target branch $\mathcal{B}_{k_2}$. Intro-

duce the sequence $(\tau^{(i)}(\nu))_{i\geq 1}$ of i.i.d. random variables, $\tau^{(i)}(\nu) \stackrel{d}{=} \tau(\nu)$, where $\tau(\nu) \stackrel{d}{=} T_0^{(k,1)}(\nu)$ with probability $p_k(\nu)/p_*(\nu)$ for every $k \in K_*$. Then

$$B(\nu) \stackrel{d}{=} \sum_{i=1}^{N_*(\nu)} \tau^{(i)}(\nu). \tag{3.18}$$

For $k \in K_*$ we define

$$\beta_k := \lim_{\nu\to\infty} \frac{L_k f_k(\nu)}{L_{k_2} f_{k_2}(\nu)}. \tag{3.19}$$

In view of (3.2), β_k may be interpreted as the stationary ratio between the number of visits to branch k and to branch k_2 as $\nu \to \infty$. Thanks to (3.12), β_k also represents the asymptotic mean number of visits to branch $\mathcal{B}_k$ before the first entrance in branch $\mathcal{B}_{k_2}$ as $\nu \to \infty$, when the process starts in state 0, i.e.

$$\beta_k = \lim_{\nu\to\infty} \mathbb{E}N_k(\nu). \tag{3.20}$$

To avoid technicalities, we henceforth assume that all the parameters β_k are well defined. Moreover, we introduce the parameter $\beta := \sum_{k\in K_*} \beta_k$, which is the asymptotic mean number of visits to dominant branches before hitting $\mathcal{B}_{k_2}$ as $\nu \to \infty$, i.e. $\beta = \lim_{\nu\to\infty} \mathbb{E}N_*(\nu)$. Based on the definition of the parameter β_k in (3.19), we partition the index set K_* of the dominant branches into three subsets, namely

$$K_* = \mathcal{N} \cup \mathcal{A} \cup \mathcal{S},$$

using the following rule:

- $k \in \mathcal{N}$ if $\beta_k = 0$;
- $k \in \mathcal{A}$ if $\beta_k \in \mathbb{R}_+$;
- $k \in \mathcal{S}$ if $\beta_k = \infty$.

The branches in $\mathcal{N}$, $\mathcal{A}$ and $\mathcal{S}$ will be called *non-attracting*, *attracting* and *strongly attracting*, respectively. Define moreover the coefficients $\gamma_{\mathcal{N}} := \sum_{k\in\mathcal{N}} \gamma_k$, $\gamma_{\mathcal{A}} := \sum_{k\in\mathcal{A}} \gamma_k$ and $\gamma_{\mathcal{S}} := \sum_{k\in\mathcal{S}} \gamma_k$, with the parameters γ_k as defined in (3.3).

We are now ready to present the proof of Theorem 3.2.2. Specifically, we prove that if $k_1 \neq k_2$, $1 \leq l_1 \leq L_{k_1}$ and $1 \leq l_2 \leq L_{k_2}$, then

$$\frac{T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)}{\mathbb{E}T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)} \xrightarrow{d} \alpha Y + (1-\alpha)W, \qquad \text{as } \nu \to \infty,$$

where α is the constant defined in (3.5), Y is an exponential random variable with unit mean and W is a random variable independent of Y, with Laplace transform

$$\mathcal{L}_W(s) = \frac{1}{1 + \sum_{k\in\mathcal{A}} \frac{\gamma_k s}{1+\gamma_k s/\beta_k} + s\gamma_{\mathcal{S}}}. \tag{3.21}$$

The crucial idea of the proof is to use Lemma 3.3.5 with the dominant term $U(\nu)$ defined as the sum of the two random variables introduced in (3.17), i.e.

$$U(\nu) := A(\nu) + B(\nu) = T_0^{(k_1,1)} + \sum_{k \in K_*} \sum_{i=1}^{N_k(\nu)} T_0^{(k,1)}(i).$$

Theorem 3.2.1 implies that $\mathbb{E}T_{(k_2,l_2)}^{(k_1,l_1)}(\nu) \sim \mathbb{E}U(\nu)$ as $\nu \to \infty$ and its proof shows that all the other terms present in the stochastic representation (3.10) are negligible compared to $U(\nu)$. Note that

$$\frac{U(\nu)}{\mathbb{E}U(\nu)} = \frac{A(\nu)}{\mathbb{E}U(\nu)} + \frac{B(\nu)}{\mathbb{E}U(\nu)} = \frac{\mathbb{E}A(\nu)}{\mathbb{E}U(\nu)} \frac{A(\nu)}{\mathbb{E}A(\nu)} + \frac{\mathbb{E}B(\nu)}{\mathbb{E}U(\nu)} \frac{B(\nu)}{\mathbb{E}B(\nu)}.$$

Recall that $A(\nu)$ and $B(\nu)$ are independent by construction. If we knew that there exist two random variables Y and W such that $A(\nu)/\mathbb{E}A(\nu) \xrightarrow{d} Y$ and $B(\nu)/\mathbb{E}B(\nu) \xrightarrow{d} W$ as $\nu \to \infty$, then

$$\frac{U(\nu)}{\mathbb{E}U(\nu)} \xrightarrow{d} \alpha Y + (1-\alpha)W, \quad \text{as } \nu \to \infty,$$

and Lemma 3.3.5 would imply that

$$\frac{T_{(k_2,l_2)}^{(k_1,l_1)}(\nu)}{\mathbb{E}T_{(k_2,l_2)}^{(k_1,l_1)}(\nu)} \xrightarrow{d} \alpha Y + (1-\alpha)W, \quad \text{as } \nu \to \infty.$$

Proposition 3.3.6 immediately gives that

$$\frac{A(\nu)}{\mathbb{E}A(\nu)} \xrightarrow{d} Y, \quad \text{as } \nu \to \infty,$$

where Y is an exponential random variable with mean one.

Thus it remains to establish that the random variable $B(\nu)/\mathbb{E}B(\nu)$ converges to W in distribution. For $B(\nu)$ defined in (3.17) it follows from (3.18) that $\mathbb{E}B(\nu) = \mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)$ and that

$$\mathcal{L}_{B(\nu)/\mathbb{E}B(\nu)}(s) = G_{N_*(\nu)}\left(\mathcal{L}_{\tau(\nu)/\mathbb{E}B(\nu)}(s)\right) = G_{N_*(\nu)}\left(\mathcal{L}_{\tau(\nu)/\mathbb{E}\tau(\nu)}\left(\frac{s}{\mathbb{E}N_*(\nu)}\right)\right), \tag{3.22}$$

where

$$G_{N_*(\nu)}(z) = \mathbb{E}(z^{N_*(\nu)}) = \frac{1}{1 + (1-z)\mathbb{E}N_*(\nu)}. \tag{3.23}$$

We need then to derive the asymptotic distribution of the random variable $\tau(\nu)/\mathbb{E}\tau(\nu)$. Let $T_k(\nu) = T_0^{(k,1)}(\nu)$. Then $\mathcal{L}_{\tau(\nu)}(s) = \sum_{k\in K_*} \frac{p_k(\nu)}{p_*(\nu)} \mathcal{L}_{T_k(\nu)}(s)$ and, using $\mathbb{E}N_k(\nu) = p_k(\nu)/p_{k_2}(\nu)$ from (3.12), we get

$$\begin{aligned}\mathcal{L}_{\tau(\nu)/\mathbb{E}\tau(\nu)}(s/\mathbb{E}N_*(\nu)) &= \mathcal{L}_{\tau(\nu)}\Big(\frac{s}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}\Big)\\ &= \sum_{k\in K^*} \frac{p_k(\nu)}{p_*(\nu)} \mathcal{L}_{T_k(\nu)}\Big(\frac{s}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}\Big)\\ &= \sum_{k\in K^*} \frac{\mathbb{E}N_k(\nu)}{\mathbb{E}N_*(\nu)} \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}\Big(\frac{s\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}\Big).\end{aligned}$$

From the definition of γ_k in (3.3) and the identity in distribution (3.18), by means of (3.11) and (3.12), we can see that

$$\gamma_k = \lim_{\nu\to\infty} \frac{\mathbb{E}N_k(\nu)\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}, \quad k \in K_*, \tag{3.24}$$

which shows that, as claimed at the end of Section 3.1, γ_k equals the asymptotic fraction of time that the activity process spends in the branch $\mathcal{B}_k$ during the transition from 0 to (k_2, l_2) as $\nu \to \infty$. For $k \in K_*$, define

$$h_k(\nu) := \frac{\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}, \tag{3.25}$$

and note that $\lim_{\nu\to\infty} h_k(\nu) = \gamma_k/\beta_k$. Indeed, it follows from (3.20) and (3.24) that

$$\lim_{\nu\to\infty} h_k(\nu) = \lim_{\nu\to\infty} \frac{\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)} = \lim_{\nu\to\infty} \frac{\mathbb{E}N_k(\nu)\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)} \frac{1}{\mathbb{E}N_k(\nu)} = \frac{\gamma_k}{\beta_k}.$$

Combining (3.22)-(3.25) yields

$$\begin{aligned}&\mathcal{L}_{B(\nu)/\mathbb{E}B(\nu)}(s)\\ &= \Big[1 + \Big(1 - \mathcal{L}_{\tau(\nu)/\mathbb{E}\tau(\nu)}(s/\mathbb{E}N_*(\nu))\Big)\mathbb{E}N_*(\nu)\Big]^{-1}\\ &= \left[1 + \Big(1 - \sum_{k\in K^*} \frac{\mathbb{E}N_k(\nu)}{\mathbb{E}N_*(\nu)} \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}\Big(\frac{s\mathbb{E}T_k(\nu)}{\mathbb{E}N_*(\nu)\mathbb{E}\tau(\nu)}\Big)\Big)\mathbb{E}N_*(\nu)\right]^{-1}\\ &= \Big[1 + \Big(\mathbb{E}N_*(\nu) - \sum_{k\in K^*} \mathbb{E}N_k(\nu)\mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\Big)\Big]^{-1}\\ &= \Big[1 + \sum_{k\in K^*} \mathbb{E}N_k(\nu)\Big(1 - \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\Big)\Big]^{-1}.\end{aligned} \tag{3.26}$$

In order to obtain an explicit expression for $\mathcal{L}_{B(\nu)/\mathbb{E}B(\nu)}(s)$ as $\nu \to \infty$, we need the following technical lemma, which is proved in Appendix 3.C.

Lemma 3.4.3. *The following statements hold:*

(i) *If $k \in \mathcal{N}$, then*

$$\lim_{\nu\to\infty} \mathbb{E}N_k(\nu)\left(1 - \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\right) = 0.$$

(ii) *If $k \in \mathcal{A}$, then*

$$\lim_{\nu\to\infty} \mathbb{E}N_k(\nu)\left(1 - \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\right) = \frac{\gamma_k s}{1+\gamma_k s/\beta_k}.$$

(iii) *If $k \in \mathcal{S}$, then*

$$\lim_{\nu\to\infty} \mathbb{E}N_k(\nu)\left(1 - \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\right) = \gamma_k s.$$

From Lemma 3.4.3 and (3.26) it follows that

$$\mathcal{L}_W(s) = \lim_{\nu\to\infty} \mathcal{L}_{B(\nu)/\mathbb{E}B(\nu)}(s) = \left[1 + \sum_{k\in\mathcal{A}} \frac{\gamma_k s}{1+\gamma_k s/\beta_k} + \sum_{k\in\mathcal{S}} \gamma_k s\right]^{-1}.$$

The independence of Y and W easily follows from the independence of the corresponding terms in the stochastic representation (3.10).

3.4.4 *The random variable W: Properties and interpretation*

The random variable W is defined by its Laplace transform, see (3.21). We remark that the shape of the distribution W is fully determined by the branches in $\mathcal{A}$ and $\mathcal{S}$, independently of the branches in $\mathcal{N}$. Indeed the random variable W can be represented as

$$W \stackrel{d}{=} (1-\gamma_{\mathcal{N}})\overline{W},$$

where $\overline{W}$ is a unit-mean random variable that in no way depends on the parameters of the branches in the set $\mathcal{N}$. On the other hand, the presence of the factor $(1-\gamma_{\mathcal{N}})$ reflects the fact that the branches in $\mathcal{N}$ do affect the mean of the asymptotic scaled transition time: Indeed convergence of the first moments holds if and only if $\alpha = 1$ or $\mathcal{N} = \emptyset$. Indeed,

$$\alpha\mathbb{E}Y + (1-\alpha)\mathbb{E}W = \alpha + (1-\alpha)(1-\gamma_{\mathcal{N}}),$$

and, if $\mathcal{N} \neq \emptyset$, then $\gamma_{\mathcal{N}} > 0$ and so $\alpha\,\mathbb{E}Y + (1-\alpha)\,\mathbb{E}W < 1$ when $\alpha \neq 1$.

Whenever either $\mathcal{A}$ or $\mathcal{S}$ is empty, the distribution of W is known explicitly, cf. Table 3.1. However, also in the scenario where both $\mathcal{A}$ and $\mathcal{S}$ are non-empty, it is still possible to give an interpretation of the distribution of W. If $\mathcal{A} \neq \emptyset$, define $m := |\mathcal{A}|$ and label the branches belonging to $\mathcal{A}$ as $1, 2, \ldots, m$. Let $\beta_{\mathcal{A}} := \sum_{k=1}^{m} \beta_k \in (0, \infty)$ be the asymptotic mean number of visits to attracting

branches as $\nu \to \infty$. Consider a hyper-exponentially distributed random variable H with rates β_k/γ_k and probabilities $\beta_k/\beta_{\mathcal{A}}$, $k = 1, \ldots, m$, whose Laplace transform is

$$\mathcal{L}_H(s) = \sum_{k=1}^{m} \frac{\beta_k}{\beta_{\mathcal{A}}} \frac{\beta_k/\gamma_k}{\beta_k/\gamma_k + s}.$$

Furthermore, consider a marked Poisson process with rate $\lambda = \beta_{\mathcal{A}}/\gamma_{\mathcal{S}}$ and i.i.d. marks distributed according to H. The random variable W in (3.21) corresponds to the sum of a random time $\mathcal{T}$, with $\mathcal{T}$ exponentially distributed with mean $1/\mu = \gamma_{\mathcal{S}}$, and the total size $\mathcal{W}(\mathcal{T})$ of the marks associated with all the events in interval $[0, \mathcal{T}]$. Indeed

$$\begin{aligned}
\mathcal{L}_{\mathcal{T}+\mathcal{W}(\mathcal{T})}(s) &= \int_{t=0}^{\infty} e^{-st} e^{\lambda t \left(\sum_{k=1}^{m} \frac{\beta_k}{\beta_{\mathcal{A}}} \frac{\beta_k/\gamma_k}{\beta_k/\gamma_k + s} - 1\right)} \mu e^{-\mu t} \, \mathrm{d}t \\
&= \left[1 + \frac{\lambda}{\mu} \left(\sum_{k=1}^{m} \frac{\beta_k}{\beta_{\mathcal{A}}} \frac{\beta_k/\gamma_k}{\beta_k/\gamma_k + s}\right) + \frac{s}{\mu}\right]^{-1} \\
&= \left[1 + \beta_{\mathcal{A}} \left(\sum_{k=1}^{m} \frac{\beta_k}{\beta_{\mathcal{A}}} \frac{s}{\beta_k/\gamma_k + s}\right) + s\gamma_{\mathcal{S}}\right]^{-1} \\
&= \left[1 + \sum_{k \in \mathcal{A}} \frac{\gamma_k s}{1 + \gamma_k s/\beta_k} + s\gamma_{\mathcal{S}}\right]^{-1}.
\end{aligned}$$

The stochastic equality $W \stackrel{d}{=} \mathcal{T} + \mathcal{W}(\mathcal{T})$ may be interpreted as follows. Define $p_{\mathcal{A}} := \sum_{k \in \mathcal{A}} p_k(\nu)$ and $p_{\mathcal{S}} := \sum_{k \in \mathcal{S}} p_k(\nu)$. The total number of visits during the transition time to the branches in $\mathcal{S}$ is geometrically distributed with parameter $p_{k_2}/p_{\mathcal{S}}$. Since the durations of these visits are independent and each relatively short compared to the transition time, the total normalized amount of time spent in the branches in $\mathcal{S}$ is exponentially distributed in the limit as $\nu \to \infty$ with mean $\gamma_{\mathcal{S}}$. The visits to the branches in $\mathcal{S}$ are interspersed with visits to the branches in $\mathcal{A}$. The number of visits to branches in $\mathcal{S}$ between two consecutive visits to branches in $\mathcal{A}$ is geometrically distributed with parameter $p_{\mathcal{A}}/p_{\mathcal{S}}$. The normalized durations of the visits to the branches in $\mathcal{A}$ have the hyper-exponential distribution H as specified above. By similar arguments as mentioned above, the normalized amounts of time between these visits are independent and exponentially distributed in the limit as $\nu \to \infty$ with mean $\gamma_{\mathcal{S}} \cdot p_{k_2}/p_{\mathcal{A}} = \gamma_{\mathcal{S}}/\beta_{\mathcal{A}}$. In other words, the visits to the branches in $\mathcal{A}$ occur as a Poisson process with rate $\lambda = \beta_{\mathcal{A}}/\gamma_{\mathcal{S}}$.

3.4.5 *An overview of the possible limiting distributions*

In this subsection we present an overview of all the possible limiting distributions of the scaled transition time by means of Table 3.1.

The case $\alpha = 1$ always yields asymptotic exponentiality: This happens when the escape time from branch $\mathcal{B}_{k_1}$ dominates the total transition time. As soon

as $\alpha \neq 1$, the set of dominant branches starts to play an important role. In particular, the shape of the asymptotic distribution depends only on the branches in the sets $\mathcal{A}$ and $\mathcal{S}$ and changes substantially whenever one of these two subsets (or both) are empty. In the case $\alpha = 0$ a diverse range of behaviors may occur, with asymptotic exponentiality only in a somewhat degenerate special case 1c. The behavior for $\alpha \in (0,1)$ is just a weighted combination of the extreme cases $\alpha = 0$ and $\alpha = 1$, as described in Theorem 3.2.2. It does not give rise to fundamentally different behavior, but interestingly enough, it does yield asymptotic exponentiality in some very special cases.

If all nodes have the same activation rate, then without loss of generality, we may assume $f_k(\nu) = \nu$, $k = 1, \ldots, K$. Under this homogeneity assumption, the sizes of components become crucial. Indeed, if one defines $L_* := \max_{k \neq k_2} L_k$ to be the size of the largest component, then $K_* = \{k \neq k_2 : L_k = L_*\}$. In this case the orders of magnitude of the two dominant terms defined in (3.17) of the stochastic representation (3.10) are

$$\mathbb{E}A(\nu) = \mathbb{E}T_0^{(k_1,1)}(\nu) \sim \frac{\nu^{L_{k_1}-1}}{L_{k_1}}, \quad \text{as } \nu \to \infty,$$

and

$$\mathbb{E}B(\nu) = \mathbb{E}\left(\sum_{k \in K_*} \sum_{i=1}^{N_k(\nu)} T_0^{(k,1)}(i)\right) \sim |K_*| \frac{\nu^{L_*-1}}{L_{k_2}}, \quad \text{as } \nu \to \infty,$$

and hence for $1 \leq l_1 \leq L_{k_1}$, $1 \leq l_2 \leq L_{k_2}$ and $k_1 \neq k_2$,

$$\mathbb{E}T_{(k_2,l_2)}^{(k_1,l_1)}(\nu) \sim \left(\frac{\mathbb{1}_{\{k_1 \in K_*\}}}{L_*} + \frac{|K_*|}{L_{k_2}}\right) \nu^{L_*-1}, \quad \text{as } \nu \to \infty. \tag{3.27}$$

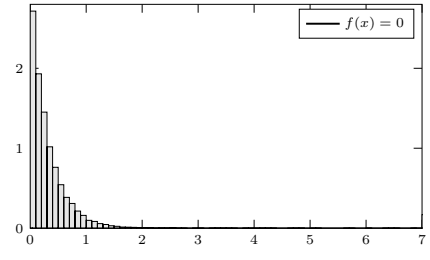
Moreover, $\beta_k/\gamma_k = (1-\alpha)/\alpha$ for every $k \in \mathcal{A}$ and thus only two possible scenarios can occur, namely 1b* and 2b***. The discriminating factor between these two scenarios is the value of α. More specifically, if $k_1 \notin K_*$, then $\alpha = 0$ and thus we are in scenario 1b*. If instead $k_1 \in K_*$, then $\alpha = L_{k_2}/(|K_*|L_*)$, which means that scenario 2b*** occurs and hence asymptotic exponentiality arises.

In Table 3.1, we denote by $(H_i)_{i \in \mathbb{N}}$ a sequence of i.i.d. hyper-exponential random variables, i.e. $H_i \overset{d}{=} H$ (see the definition of H in Subsection 3.4.4), while $\mathcal{G}$ is a geometric random variable $\mathcal{G} \overset{d}{=} \text{Geo}(\frac{1}{1+\beta_{\mathcal{A}}})$, independent of all the other random variables.

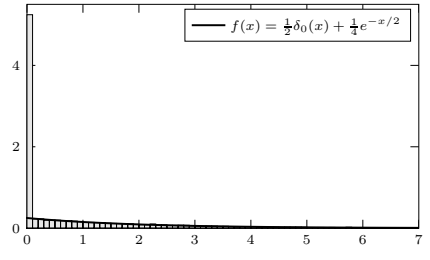
α	$\mathcal{A}$	$\mathcal{S}$	LIMITING DISTRIBUTION	ADDITIONAL CONDITIONS	CASE
0	$\emptyset$	$\emptyset$	δ_0 (trivial r.v. identical to 0)		1a
	non-empty	$\emptyset$	$\sum_{i=1}^{\mathcal{G}} H_i\left(\frac{\beta_1}{\beta_{\mathcal{A}}}, \ldots, \frac{\beta_m}{\beta_{\mathcal{A}}}, \frac{\beta_1}{\gamma_1}, \ldots, \frac{\beta_m}{\gamma_m}\right)$		1b
			$\sum_{i=1}^{\mathcal{G}} \mathrm{Exp}_i(\lambda)$	$\beta_k/\gamma_k = \lambda \quad \forall k \in \mathcal{A}$	1b*
	$\emptyset$	non-empty	$\mathrm{Exp}(1/\gamma_{\mathcal{S}})$		1c
	non-empty	non-empty	W		1d
$(0,1)$	$\emptyset$	$\emptyset$	$\mathrm{Exp}(1/\alpha)$		2a
	non-empty	$\emptyset$	$\mathrm{Exp}(1/\alpha) + \sum_{i=1}^{\mathcal{G}} H_i\left(\frac{\beta_1}{\beta_{\mathcal{A}}}, \ldots, \frac{\beta_m}{\beta_{\mathcal{A}}}, \frac{\beta_1}{(1-\alpha)\gamma_1}, \ldots, \frac{\beta_m}{(1-\alpha)\gamma_m}\right)$		2b
			$\mathrm{Exp}(1/\alpha) + \sum_{i=1}^{\mathcal{G}} \mathrm{Exp}_i\left(\frac{\lambda}{1-\alpha}\right)$	$\beta_k/\gamma_k = \lambda \quad \forall k \in \mathcal{A}$	2b*
			$\mathrm{Exp}\left(\frac{1}{\alpha(1+\sum_{k=1}^{m}\beta_k)}\right)$	$\beta_k/\gamma_k = \frac{1-\alpha}{\alpha} \quad \forall k \in \mathcal{A}$	2b**
			$\mathrm{Exp}(1)$	$\beta_k/\gamma_k = \frac{1-\alpha}{\alpha} \quad \forall k \in \mathcal{A}$ $\sum_{i=1}^{m} \beta_k = \frac{1-\alpha}{\alpha}$	2b***
	$\emptyset$	non-empty	$\mathrm{Exp}(1/\alpha) + \mathrm{Exp}(1/(1-\alpha)\gamma_{\mathcal{S}})$		2c
			$\mathrm{Erlang}(2, 1/\alpha)$	$\alpha = \gamma_{\mathcal{S}}/(1+\gamma_{\mathcal{S}})$	2c*
	non-empty	non-empty	$\mathrm{Exp}(1/\alpha) + (1-\alpha)W$		2d
1	-	-	$\mathrm{Exp}(1)$		3

Table 3.1: Overview of the possible asymptotic distributions of the scaled transition time

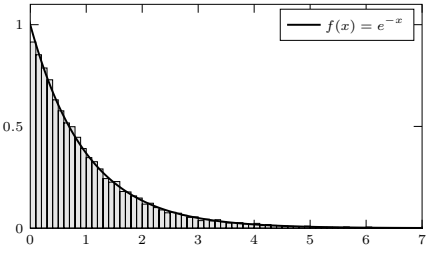
To illustrate the range of possible limiting distributions, we present some simulation results. We consider the simplest system that is sufficiently rich to show the wide range of behaviors presented in Table 3.1. Specifically, we consider a complete 3-partite network, whose three components have sizes L_1, L_2 and L_3, and assume that the process starts in state $(1, L_1)$ and the target state is $(3, L_3)$. We use activation rates of the form $f_k(\nu) = \nu^{a_k}$. For compactness, we write $\underline{a}$ for (a_1, a_2, a_3) and $\underline{L}$ for (L_1, L_2, L_3). This choice for the activation rates allows to invert the Laplace transform of W in all the cases and thus obtain a probability density function $f(x)$, which can be compared with the simulation data. All simulations have been performed choosing the parameter $\nu = 150$ and simulating the transition time for each network 20000 times with a customized code written in the programming language C. The results are shown in Figures 3.3 and 3.4.



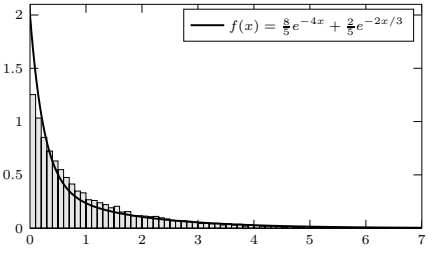
(a) Case 1a: $\underline{L} = (3,4,6)$, $\underline{a} = (1,1,5/3)$



(b) Case 1b: $\underline{L} = (3,5,5)$, $\underline{a} = (1/2,1/2,1/2)$

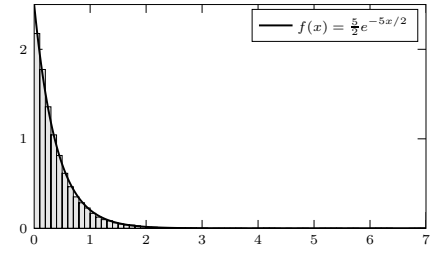


(c) Case 1c: $\underline{L} = (3,3,3)$, $\underline{a} = (4/5,3/5,3/5)$

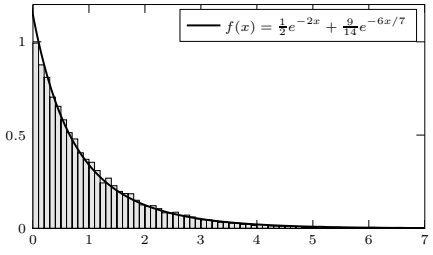


(d) Case 1d: $\underline{L} = (3,4,6)$, $\underline{a} = (1,3/4,3/4)$

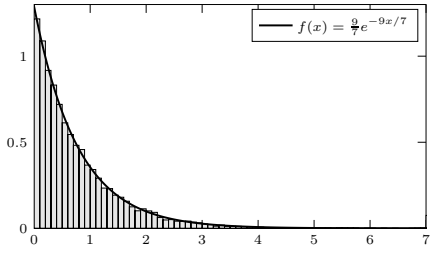
Figure 3.3: Plots of the empirical probability density function of the scaled transition times and the density $f(x)$ of $\alpha Y + (1-\alpha)W$



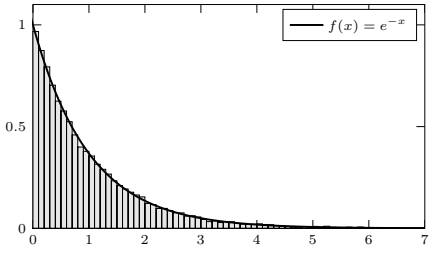
(a) Case 2a: $\underline{L} = (3,4,2)$, $\underline{a} = (9/10, 9/10, 9/5)$



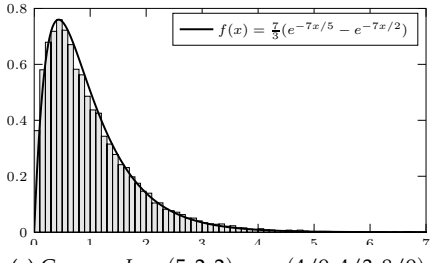
(b) Case 2b*: $\underline{L} = (4,3,4)$, $\underline{a} = (2/3, 1, 1)$



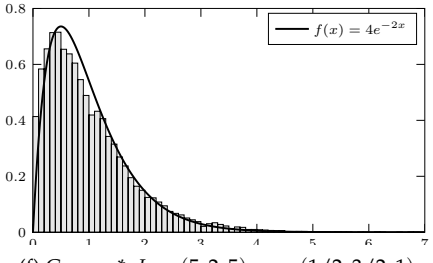
(c) Case 2b**: $\underline{L} = (2,4,5)$, $\underline{a} = (7/4, 7/8, 7/4)$



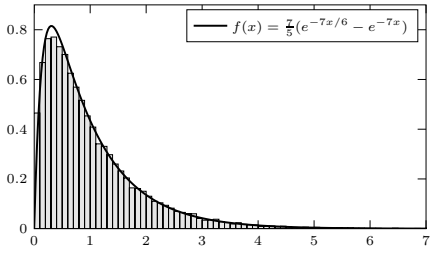
(d) Case 2b***: $\underline{L} = (3,3,5)$, $\underline{a} = (7/8, 7/8, 7/8)$



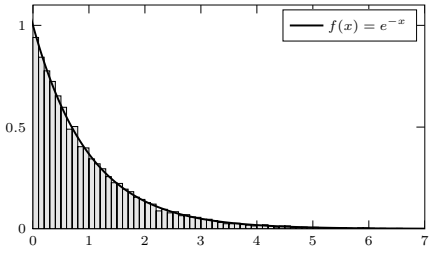
(e) Case 2c: $\underline{L} = (5,2,2)$, $\underline{a} = (4/9, 4/3, 8/9)$



(f) Case 2c*: $\underline{L} = (5,2,5)$, $\underline{a} = (1/2, 3/2, 1)$



(g) Case 2d: $\underline{L} = (4,2,6)$, $\underline{a} = (3/5, 6/5, 3/5)$



(h) Case 3: $\underline{L} = (3,3,3)$, $\underline{a} = (1, 3/4, 3/2)$

Figure 3.4: Plots of the empirical probability density function of the scaled transition times and the density $f(x)$ of $\alpha Y + (1-\alpha)W$

3.5 DELAY BOUNDS REVISITED

In the special case in which the K components $V_1, \dots, V_K$ of the conflict graph G have all equal size $|V_j| = \cdots = |V_K| = L = N/K$, G is a symmetric partite network (see the definition in Section 2.2). We further assume, only in this section, that all nodes have the same activation rate, i.e. $f_k(\nu) = \nu$ for every $k = 1, \dots, K$, so that we are back in the framework of Chapter 2.

After having identified $\nu = \sigma$, it immediately follows from (3.27) that $\Gamma(G) = N/K - 1$ and in this way we obtain the lower bound for the long-term average delay on complete K-partite graphs presented in Section 2.4, namely

$$\mathbb{E}W(\rho) \gtrsim \left(\frac{1}{1-\rho}\right)^{N/K-1}, \quad \text{as } \rho \uparrow 1.$$

However, using (3.27) and the structure of this conflict graph, we can sharpen this lower bound. Firstly, a careful inspection of the proof of Theorem 2.5.1 shows that the inequality

$$\mathbb{E}W(\rho) \geq \frac{K-1}{2(K-\rho)} \mathbb{E}T^{(k_1,L)}_{(k_2,L)}(\nu)$$

holds in fact *for any* activation rate $\nu > 0$ and not only in the limit $\nu \to \infty$. Furthermore, from (3.27) it follows that

$$\mathbb{E}T^{(k_1,L)}_{(k_2,L)}(\sigma) \sim \frac{K^2}{N} \sigma^{N/K-1}, \quad \text{as } \sigma \to \infty. \tag{3.28}$$

Theorem 2.5.1 then yields

$$\liminf_{\sigma\to\infty} \frac{\mathbb{E}W(\sigma)}{\mathbb{E}T^{(k_1,L)}_{(k_2,L)}(\sigma)} \geq \frac{K-1}{2(K-\rho)} \frac{K^2}{N} \sigma^{N/K-1},$$

and, in view of Proposition 2.4.3, the long-term average delay $\mathbb{E}W(\rho)$ at any node scales as $\rho \uparrow 1$ at least as

$$\frac{(K-1)}{2(K-\rho)} \frac{K^2}{N} \left(\frac{1}{1-\rho} - \frac{N}{K}\right)^{\frac{N}{K}-1}. \tag{3.29}$$

A similar result for the long-run average delay can be deduced from [20, Theorem 2], namely that for every $\rho \in (0,1)$,

$$\mathbb{E}W(\rho) \geq \frac{(K-1)^2 \rho^{N/K+1}}{2NK^{N/K-1}(K-(K-1)\rho)} \left(\frac{1}{1-\rho}\right)^{N/K-1}. \tag{3.30}$$

Note that the constant appearing in this inequality in front of the leading term $(\frac{1}{1-\rho})^{N/K-1}$ is smaller than the one in (3.29), so our result is sharper. However, inequality (3.30) holds for any load $\rho \in (0,1)$, while our result is meaningful only asymptotically.

3.6 THROUGHPUT STARVATION AND NEAR-SATURATION

In this section we show how the results for the asymptotics of the transition time $T^{(k_1,l_1)}_{(k_2,l_2)}(\nu)$ in Theorems 3.2.1 and 3.2.2 can be exploited to gain insight into phenomena like throughput starvation or near-saturation. More specifically, we present the proof of Theorem 3.2.3, which gives an asymptotic lower bound on the probability of throughput starvation, and in Subsection 3.6.1 we prove Proposition 3.6.3, a complementary result which indicates over what time scales throughput near-saturation occurs.

Proof of Theorem 3.2.3. Recall that in Section 3.2 we defined $\tau_k(t)$ as $\tau_k(t) = \int_0^t \mathbb{1}_{\{X^*(s)\in\mathcal{B}_k\}}\,\mathrm{d}s$. Observe that $\tau_{k_2}(t(\nu)) > 0$ if and only if $t(\nu) > T^{(k_1,l_1)}_{(k_2,1)}(\nu)$, because the throughput of branch $\mathcal{B}_{k_2}$ remains zero until the activity process enters $\mathcal{B}_{k_2}$. Hence

$$\mathbb{P}(\tau_{k_2}(t(\nu)) > 0) = \mathbb{P}\Big(T^{(k_1,l_1)}_{(k_2,1)}(\nu) < t(\nu)\Big) = \mathbb{P}\Big(\frac{T^{(k_1,l_1)}_{(k_2,1)}(\nu)}{\mathbb{E}T^{(k_1,l_1)}_{(k_2,1)}(\nu)} < \frac{t(\nu)}{\mathbb{E}T^{(k_1,l_1)}_{(k_2,1)}(\nu)}\Big).$$

Taking the limit as $\nu\to\infty$, Theorem 3.2.2, gives

$$\lim_{\nu\to\infty}\mathbb{P}(\tau_{k_2}(t(\nu)) > 0) \le \mathbb{P}(Z < \omega),$$

and (3.6) follows. □

3.6.1 *Throughput near-saturation*

Assume that at time $t = 0$ there is at least one node active in component V_k, i.e. $X(0) = (k,l) \in \mathcal{B}_k$. Define the total full-component active time in $[0,t]$ as

$$\tau_k[0,t] := \int_0^t \mathbb{1}_{\{X(s)=(k,L_k)\}}\,\mathrm{d}s,$$

the residual time in component V_k during $[0,t]$ as

$$R_k[0,t] := \int_0^t \mathbb{1}_{\{X(r)\in\mathcal{B}_k\,\forall\, r\in[0,s]\}}\,\mathrm{d}s,$$

and the full-component active time contained in the residual time in V_k during $[0,t]$ as

$$\tau^{\mathrm{res}}_k[0,t] := \int_0^t \mathbb{1}_{\{X(r)\in\mathcal{B}_k\,\forall\, r\in[0,s]\}}\mathbb{1}_{\{X(s)=(k,L_k)\}}\,\mathrm{d}s.$$

For compactness, we have suppressed the implicit dependence on the parameter ν and the initial state (k,l) in the notation. From this point onwards, we will also drop the subscript k to keep the notation light. Note that

$$R[0,t] \overset{d}{=} \min\{t, T^{(k,l)}_0\} \quad \text{and } \tau^{\mathrm{res}}[0,t] \overset{d}{=} \tau[0,R[0,t]].$$

The random variables $\tau[0,t]$, $R[0,t]$ and $\tau^{\text{res}}[0,t]$, being particular occupancy times, are non-decreasing in t on every sample path of the activity process $\{X(t)\}_{t\geq 0}$. Therefore, the random variables

$$\begin{aligned}
&\tau[0,\infty] := \lim_{t\to\infty} \tau[0,t],\\
&R[0,\infty] := \lim_{t\to\infty} R[0,t] = T_0^{(k,l)},\\
&\tau^{\text{res}}[0,\infty] := \lim_{t\to\infty} \tau^{\text{res}}[0,t]
\end{aligned}$$

are well defined. For $0 \leq s \leq t \leq \infty$, we define

$$\begin{aligned}
&\tau[s,t] := \tau[0,t] - \tau[0,s],\\
&R[s,t] := R[0,t] - R[0,s],\\
&\tau^{\text{res}}[s,t] := \tau^{\text{res}}[0,t] - \tau^{\text{res}}[0,s].
\end{aligned}$$

From the above definition, it is easily seen that for every sample path, $\tau^{\text{res}}[s,t]$ provides a lower bound for both $\tau[s,t]$ and $R[s,t]$, as stated in the next lemma.

Lemma 3.6.1. *For $0 \leq s \leq t \leq \infty$, $\tau^{\text{res}}[s,t] \leq \tau[s,t]$ and $\tau^{\text{res}}[s,t] \leq R[s,t]$.*

Proof. Rearranging terms, both the differences $\tau[s,t] - \tau^{\text{res}}[s,t]$ and $R[s,t] - \tau^{\text{res}}[s,t]$ can be written as integrals with a non-negative integrand. □

In particular, Lemma 3.6.1 implies that, for every $0 \leq s \leq t \leq \infty$

$$\mathbb{E}\tau^{\text{res}}[s,t] \leq \mathbb{E}R[s,t]. \tag{3.31}$$

However, as stated in the next lemma, in the limit as $\nu \to \infty$, the ratio of the expected values of $\tau^{\text{res}}[0,\infty]$ and $T_0^{(k,l)} = R[0,\infty]$ converges to 1.

Lemma 3.6.2. *For any initial state $X(0) = (k,l) \in \mathcal{B}_k$,*

$$\lim_{\nu\to\infty} \frac{\mathbb{E}\tau^{\text{res}}[0,\infty]}{\mathbb{E}T_0^{(k,l)}} = 1.$$

Proof. Since the ratio is clearly less than 1 by (3.31), it suffices to show that the liminf as $\nu \to \infty$ is larger than 1. Applying the result in [131] and using (3.2), one obtains that for every $1 \leq l \leq L_k$, if $X(0) = (k,l)$, then

$$\mathbb{E}\tau^{\text{res}}[0,\infty] = \mathbb{E}\Big(\int_0^{T_0^{(k,l)}} \mathbb{1}_{\{X(s)=(k,L_k)\}}\, \mathrm{d}s\Big) \geq \frac{1}{L_k} f_k(\nu)^{L_k-1},$$

and thus, involving (3.7),

$$\liminf_{\nu\to\infty} \frac{\mathbb{E}\tau^{\text{res}}[0,\infty]}{\mathbb{E}T_0^{(k,l)}} \geq \liminf_{\nu\to\infty} \frac{f_k(\nu)^{L_k-1}/L_k}{\mathbb{E}T_0^{(k,l)}} = 1.$$ □

The next proposition establishes a near-saturation property in the sense that if $X(0) = (k,l) \in B_k$, then for any time period $t(\nu) = o(\mathbb{E}T_0^{(k,l)})$ every node in component V_k will be active an arbitrarily large fraction of the time with probability one as $\nu \to \infty$.

Proposition 3.6.3 (Throughput near-saturation). *Suppose that* $X(0) = (k,l) \in \mathcal{B}_k$ *and that* $T_0^{(k,l)}/\mathbb{E}T_0^{(k,l)} \xrightarrow{d} Z$ *as* $\nu \to \infty$. *Then for every* $\omega \in [0,1]$ *and every* $\delta > 0$,

$$\liminf_{\nu\to\infty} \mathbb{P}\Big(\tau[0, \omega\mathbb{E}T_0^{(k,l)}] \geq (1-\delta)\omega\mathbb{E}T_{(k,l),0}\Big) \geq \mathbb{P}(Z \geq \omega).$$

In particular, for any $t(\nu) = o(\mathbb{E}T_0^{(k,l)}(\nu))$,

$$\liminf_{\nu\to\infty} \mathbb{P}\Big(\tau[0, t(\nu)] \geq (1-\delta)t(\nu)\Big) = 1.$$

Remark. As mentioned earlier, the hypothesis that $T_0^{(k,l)}/\mathbb{E}T_0^{(k,l)} \xrightarrow{d} Z$ is not just a convenient assumption, but something that we actually know. In particular, Proposition 3.3.6 says that $Z \stackrel{d}{=} \mathrm{Exp}(1)$. Moreover, since the result holds for every initial state in $\mathcal{B}_k$, it is true also for a random initial state in $\mathcal{B}_k$. Indeed, as seen in Section 3.3, the convergence in distribution of $T_0^{(k,l)}(\nu)/\mathbb{E}T_0^{(k,l)}(\nu)$ to Z as $\nu \to \infty$ does not depend on the initial state, as long as it belongs to $\mathcal{B}_k$.

Proof. We abbreviate in this proof the hitting time $T_0^{(k,l)}$ as T. First, Lemma 3.6.1 implies that

$$\mathbb{P}(\tau[0, \omega\mathbb{E}T] \geq (1-\delta)\omega\mathbb{E}T) \geq \mathbb{P}(\tau^{\mathrm{res}}[0, \omega\mathbb{E}T] \geq (1-\delta)\omega\mathbb{E}T).$$

Moreover, by definition of $R[0,t] = \min\{t, T\}$, we have

$$\mathbb{P}(R[0, \omega\mathbb{E}T] \geq \omega\mathbb{E}T) = \mathbb{P}(T \geq \omega\mathbb{E}T).$$

In view of the assumption that $T(\nu)/\mathbb{E}T(\nu) \xrightarrow{d} Z$ as $\nu \to \infty$, it therefore suffices to prove that for every $\omega \in [0,1]$ and every $\delta > 0$,

$$\liminf_{\nu\to\infty} \mathbb{P}(\tau^{\mathrm{res}}[0, \omega\mathbb{E}T] \geq (1-\delta)\omega\mathbb{E}T) \geq \liminf_{\nu\to\infty} \mathbb{P}(R[0, \omega\mathbb{E}T] \geq \omega\mathbb{E}T).$$

Suppose that this latter statement is false, i.e. there exist $\omega_0 \in [0,1]$, $\delta > 0$ and $\eta > 0$ such that

$$\liminf_{\nu\to\infty} \mathbb{P}(\tau^{\mathrm{res}}[0, \omega_0\mathbb{E}T] \geq (1-\delta)\omega_0\mathbb{E}T) \leq \liminf_{\nu\to\infty} \mathbb{P}(R[0, \omega_0\mathbb{E}T] \geq \omega_0\mathbb{E}T) - \eta. \tag{3.32}$$

Then it can be shown that there exists $\epsilon_{\omega_0,\delta} > 0$ such that

$$\liminf_{\nu\to\infty} \frac{\mathbb{E}\tau^{\mathrm{res}}[0, \omega_0\mathbb{E}T]}{\mathbb{E}T} \leq \liminf_{\nu\to\infty} \frac{\mathbb{E}R[0, \omega_0\mathbb{E}T]}{\mathbb{E}T} - \epsilon_{\omega_0,\delta}. \tag{3.33}$$

Indeed,

$$\begin{aligned}
&\liminf_{\nu\to\infty}\left(\frac{\mathbb{E}R[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}-\frac{\mathbb{E}\tau^{\mathrm{res}}[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\right)\\
&=\liminf_{\nu\to\infty}\int_0^{\infty}\mathbb{P}\Big(\frac{R[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)-\mathbb{P}\Big(\frac{\tau^{\mathrm{res}}[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)\,\mathrm{d}y\\
&\geq\liminf_{\nu\to\infty}\int_{(1-\delta)\omega_0}^{\omega_0}\mathbb{P}\Big(\frac{R[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)-\mathbb{P}\Big(\frac{\tau^{\mathrm{res}}[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)\,\mathrm{d}y\\
&\geq\int_{(1-\delta)\omega_0}^{\omega_0}\liminf_{\nu\to\infty}\left(\mathbb{P}\Big(\frac{R[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)-\mathbb{P}\Big(\frac{\tau^{\mathrm{res}}[0,\omega_0\mathbb{E}T]}{\mathbb{E}T}\geq y\Big)\right)\,\mathrm{d}y\\
&\geq\eta\delta\omega_0>0,
\end{aligned}$$

where the third last inequality is a consequence of Lemma 3.6.1, the second last inequality follows from the generalized Fatou's lemma, while the last inequality follows from (3.32) and is illustrated by Figure 3.5. Thus we can take $\epsilon_{\omega_0,\delta}:=\eta\delta\omega_0$. Equation (3.31) yields

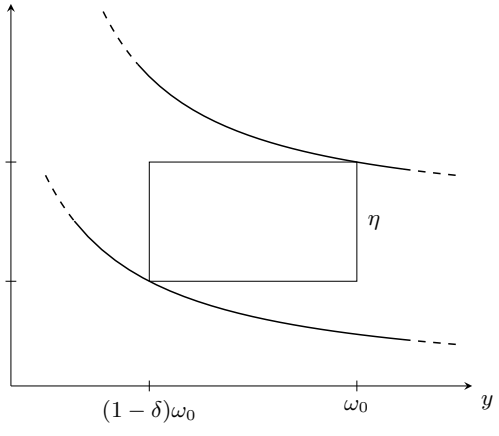


Figure 3.5: $\mathbb{P}(\tau^{\mathrm{res}}[0,\omega_0\mathbb{E}T]/\mathbb{E}T\geq y)$ (lower line) and $\mathbb{P}(R[0,\omega_0\mathbb{E}T]/\mathbb{E}T\geq y)$ (upper line)

$$\liminf_{\nu\to\infty}\frac{\mathbb{E}\tau^{\mathrm{res}}[\omega_0\mathbb{E}T,\infty]}{\mathbb{E}T}\leq\liminf_{\nu\to\infty}\frac{\mathbb{E}R[\omega_0\mathbb{E}T,\infty]}{\mathbb{E}T},\tag{3.34}$$

and thus, summing term by term (3.33) and (3.34) and using the fact that $\mathbb{E}R[0,\infty]=\mathbb{E}T$ by definition, we get

$$\liminf_{\nu\to\infty}\frac{\mathbb{E}\tau^{\mathrm{res}}[0,\infty]}{\mathbb{E}T}\leq\liminf_{\nu\to\infty}\frac{\mathbb{E}R[0,\infty]}{\mathbb{E}T}-\epsilon_{\omega_0,\delta}=1-\epsilon_{\omega_0,\delta},$$

which contradicts Lemma 3.6.2. □

3.7 MIXING TIMES

In the previous sections we have analyzed the transient behavior of the aggregated Markov process $\{X^*(t)\}_{t\geq 0}$ in terms of hitting times and we have shown how this leads to starvation of individual nodes over certain time scales. In this section we turn attention to the long-run behavior of the Markov process $\{X^*(t)\}_{t\geq 0}$ and in particular examine the rate of convergence to the stationary distribution. We measure the rate of convergence in terms of the total variation distance and the so-called *mixing time*, which describes the time required for the distance to stationarity to become small.

The mixing time becomes particularly relevant when the network has two or more dominant components which together attract the entire probability mass in the limit as $\nu \to \infty$. Indeed, in this case, the mixing time provides an indication of how long it takes the activity process to reach a certain level of fairness among the dominant components. We will prove a lower bound for the mixing time using the notion of conductance.

Define $P_\nu^t(x,y) := \mathbb{P}(X^*(t) = y \mid X^*(0) = x)$ and denote by $P_\nu^t(x,\cdot)$ the distribution at time t of the Markov process $\{X^*(t)\}_{t\geq 0}$ started at time 0 in x. We will also use the notation μP_ν^t to denote the distribution at time t of the Markov process $\{X^*(t)\}_{t\geq 0}$ with initial distribution μ at time 0. The maximal distance over $x \in \mathcal{X}^*$, measured in terms of total variation, between the distribution at time t and the stationary distribution $\pi(\nu)$ is defined as

$$d(t,\nu) := \max_{x\in\mathcal{X}^*} \|P_\nu^t(x,\cdot) - \pi_\cdot(\nu)\|_{\mathrm{TV}}.$$

Define the mixing time of the process $\{X^*(t)\}_{t\geq 0}$ as

$$t_{\mathrm{mix}}(\epsilon,\nu) := \inf\{t \geq 0 : d(t,\nu) \leq \epsilon\}.$$

For a fixed $r \in (0,1)$ consider the subset $\tilde{K}(r)$ of branches whose stationary probability is asymptotically no more than r, i.e.

$$\tilde{K}(r) := \{k \ : \ \lim_{\nu\to\infty} \pi_{\mathcal{B}_k}(\nu) \leq r\}.$$

Define $\kappa = \kappa(r)$ as the index corresponding to the branch $\mathcal{B}_\kappa$ which has asymptotically the largest mean escape time among those in $\tilde{K}(r)$, i.e. such that for every $j \in \tilde{K}(r)$,

$$\lim_{\nu\to\infty} \frac{\mathbb{E}T_0^{(\kappa,1)}(\nu)}{\mathbb{E}T_0^{(j,1)}(\nu)} = \lim_{\nu\to\infty} \frac{L_j f_\kappa(\nu)^{L_\kappa - 1}}{L_\kappa f_j(\nu)^{L_j - 1}} \geq 1. \tag{3.35}$$

The next result shows that the mixing time is asymptotically at least of the same order of magnitude as the escape time from branch $\mathcal{B}_\kappa$.

Proposition 3.7.1 (Activity process mixing time)**.** *For any $r \in (0,1)$ and $\epsilon \in (0, 1-r)$, there exists a constant $C_{\epsilon,r} > 0$ such that*

$$t_{\mathrm{mix}}(\epsilon,\nu) \gtrsim C_{\epsilon,r} \frac{f_\kappa(\nu)^{L_\kappa - 1}}{L_\kappa}, \quad \text{as } \nu \to \infty.$$

Proposition 3.7.1 shows that it can take an extremely long time for the process $\{X^*(t)\}_{t\geq 0}$ to reach stationarity, especially when ν is large. Such a long mixing time is typically due to the activity process being stuck for a considerable period in one of the components, and thus not visiting the states in the other components. In fact, the statement of the proposition can be rewritten as

$$t_{\text{mix}}(\epsilon,\nu) \gtrsim C_{\epsilon,r}\mathbb{E}T_0^{(\kappa,1)}(\nu), \quad \text{as } \nu\to\infty.$$

We will prove Proposition 3.7.1 exploiting the presence of a bottleneck in the state space and using the notion of conductance.

For any subset $S\subseteq\mathcal{X}^*$, let $\pi_S(\nu)=\sum_{s\in S}\pi_s(\nu)$ be the stationary probability of the subset S. Define the *probability flow out of* S as

$$Q_{S,S^c}(\nu) := \sum_{s\in S, s'\in S^c}\pi_s(\nu)q(s,s'),$$

where $q(s,s')$ is the transition rate between states $s,s'\in\mathcal{X}$ for the aggregated process $\{X^*(t)\}_{t\geq 0}$ introduced in Section 3.1. Define moreover the *conductance of* $S\subseteq\mathcal{X}^*$ as

$$\Phi_S(\nu) := \frac{Q_{S,S^c}(\nu)}{\pi_S(\nu)}.$$

The *conductance profile* of the process $\{X^*(t)\}_{t\geq 0}$ is defined as

$$\Phi_r(\nu) := \min_{S\subset\mathcal{X}\,:\,\pi_S(\nu)\leq r}\Phi_S(\nu).$$

The following result, valid for any Markov process on a finite state space $\mathcal{X}^*$ with conductance profile Φ_r, shows how the conductance of the process yields a lower bound on the mixing time. It is a continuous-time version of [99, Theorem 7.3] and the proof is relegated to Appendix 3.D.

Lemma 3.7.2. *For any $r\in(0,1)$ and any $\epsilon\in(0,1-r)$,*

$$t_{\text{mix}}(\epsilon,\nu)\geq\frac{1-r-\epsilon}{\Phi_r(\nu)}.$$

In order to get a sharp bound for the conductance and hence a sharp lower bound for the mixing time, we need to identify a subset S with low conductance. As proved in [93], it suffices to look at the connected subsets of the state space. Therefore the branches in $\tilde{K}(r)$ become good candidates for being the lowest-conductance subsets of $\mathcal{X}^*$. From (3.2) it follows that for any state (k,l) in branch $\mathcal{B}_k$ with $l<L_k$,

$$\frac{\pi_{(k,l)}(\nu)}{\pi_{(k,L_k)}(\nu)}=\binom{L_k}{l}f(\nu)^{l-L_k}, \quad \text{as } \nu\to\infty.$$

Using this fact, we obtain the following asymptotic estimate for the conductance of branch $\mathcal{B}_k$ as $\nu \to \infty$

$$\Phi_{\mathcal{B}_k}(\nu) = \frac{\pi_{(k,1)}(\nu) \cdot 1}{\sum_{l=1}^{L_k} \pi_{(k,l)}(\nu)} = \frac{\frac{\pi_{(k,1)}(\nu)}{\pi_{(k,L_k)}(\nu)}}{\sum_{l=1}^{L_k} \frac{\pi_{(k,l)}(\nu)}{\pi_{(k,L_k)}(\nu)}} \geq \frac{\pi_{(k,1)}(\nu)}{\pi_{(k,L_k)}(\nu)} \sim L_k f_k(\nu)^{1-L_k}.$$

Thanks to the definition (3.35) of $\kappa = \kappa(r)$, $\pi_{\mathcal{B}_k}(\nu) \leq r$. Since by definition $\Phi_r(\nu) \leq \Phi_{\mathcal{B}_k}(\nu)$, the asymptotic estimate for $\Phi_{\mathcal{B}_k}(\nu)$ and Lemma 3.7.2 imply that for every $\epsilon \in (0, \frac{1}{4})$ and for ν sufficiently large

$$t_{\text{mix}}(\epsilon, \nu) \gtrsim (1 - r - \epsilon) \frac{f_\kappa(\nu)^{L_\kappa - 1}}{L_\kappa}, \quad \text{as } \nu \to \infty,$$

which completes the proof of the lower bound claimed in Proposition 3.7.1.

3.8 MODEL EXTENSIONS

So far we have assumed that two nodes interfere if and only if they belong to different components. In this section, we continue to assume that nodes that belong to different components interfere, but we allow nodes within the same component to interfere as well. If two or more nodes within component V_k interfere with each other, there will be fewer admissible activity states of smaller size. In particular, it is not longer the case that all the L_k nodes of component V_k can be active simultaneously, which makes the transition between different components potentially faster.

The components are assumed to be *minimal*, in the sense that they cannot be split into two non-trivial components, while retaining the full interference across components. As before, each independent set of the conflict graph must be a subset of one of the components, because two nodes that belong to different components by definition interfere. However, some subsets within the same component may no longer be independent sets in the conflict graph.

In the previous sections, we further assumed all nodes within the same component to have the same activation rate, so that state aggregation could be applied to obtain an equivalent Markov process with a star-shaped state space. In this section, we allow the nodes within the same component to have possibly different activation rates. With minor abuse of notation, denote by $f_l(\nu)$ the activation rate of node l, and define $F_k(\nu) := \sum_{j \in V_k} f_l(\nu)$ as the aggregate activation rate of all nodes in the k-th component. Due to this heterogeneity, it is not possible anymore to aggregate activity states with equal number of active nodes in a component as we did in Section 3.1. Hence, we go back to the setting of Subsection 1.2.2 and consider the activity process $\{X(t)\}_{t \geq 0}$ on the state space $\mathcal{X}$. In particular, an activity state $x \in \mathcal{X}$ is a N-dimensional 0-1 vector.

For every $x \in \mathcal{X}$ define $I_x \subseteq V$ to be the subset of nodes which are active in state x, i.e.

$$I_x := \{i \in V : x_i = 1\}.$$

For every $x \in \mathcal{X}$, I_x is by construction an independent set in the conflict graph G. Define moreover $\mathcal{X}_k := \{x \in \mathcal{X} : I_x \subseteq V_k,\ x \neq 0\}$. Then

$$\mathcal{X} = \{0\} \cup \bigcup_{k \in K} \mathcal{X}_k.$$

Each component satisfies the following monotonicity property: If $I_x \in \mathcal{X}_k$, then for every nonempty $I_y \subseteq I_x$, we have $I_y \in \mathcal{X}_k$. Indeed, if x is an admissible state and it belongs to $\mathcal{X}_k$, then any state obtained from it by switching off some nodes is still admissible and belongs to $\mathcal{X}_k$ as well. The next lemma shows that between any pair of activity states corresponding to the same component, there continues to exist a path which does not visit the state $0 \in \mathcal{X}$.

Lemma 3.8.1. *If x, y are two admissible states in $\mathcal{X}_k$ both with at least one active node, then there exists a path in $\mathcal{X}_k$ from x to y that does not pass through the state $0 \in \mathcal{X}$, i.e. a finite sequence $\omega = (\omega_1, \dots, \omega_n)$ of admissible states with $\omega_1 = x$, $q(\omega_j, \omega_{j+1}) > 0$ for every $j = 1, \dots, n-1$ and $\omega_n = y$ such that $0 \notin \omega$.*

Proof. Consider the two independent sets I_x and I_y corresponding to the activity states x and y, respectively. Since $x, y \in \mathcal{X}_k$ and they both have at least one active node, it follows that $I_x, I_y \subseteq V_k$ and that $I_x, I_y \neq \emptyset$. If $I_x \cap I_y \neq \emptyset$, then the statement is trivially true. Suppose instead that $I_x \cap I_y = \emptyset$ and without loss of generality take $I_x = A$ and $I_y = B$. Since all singletons $\{a\}$, $a \in A$, and $\{b\}$, $b \in B$, are subsets of V_k, thanks to the monotonicity property of V_k we have that each one of the corresponding activity states $\mathbb{1}_{\{a\}}$, $a \in A$, and $\mathbb{1}_{\{b\}}$, $b \in B$ is admissible and can be reached starting from x and y, respectively, without passing through the state $0 \in \mathcal{X}$. If there exists two states, $a \in A$ and $b \in B$ such that a does not interfere with b, then the activity state $\mathbb{1}_{\{a,b\}}$ would be admissible and we can construct a path from x to y which does not pass through state 0, namely

$$x = \mathbb{1}_A \to \cdots \to \mathbb{1}_{\{a\}} \to \mathbb{1}_{\{a,b\}} \to \mathbb{1}_{\{b\}} \to \cdots \to \mathbb{1}_B = y.$$

Consider now the case where every node in A interferes with every node in B and denote by C the subset $V_k \setminus (A \cup B)$. Furthermore, define $C_A \subseteq C$ as the collection of the nodes $c \in C$ that interfere with all the nodes in A and define $C_B \subseteq C$ analogously. We claim that every $c \in C$ interferes either with all the nodes in A *or* with all the nodes in B, i.e.

$$C = C_A \cup C_B.$$

Indeed, if there exist $a \in A$, $b \in B$ such that both a and b do not interfere with c, then we could construct a path from x to y which does not pass through state 0, namely

$$x = \mathbb{1}_A \to \cdots \to \mathbb{1}_{\{a\}} \to \mathbb{1}_{\{a,c\}} \to \mathbb{1}_{\{c\}} \to \mathbb{1}_{\{b,c\}} \to \mathbb{1}_{\{b\}} \to \cdots \to \mathbb{1}_B = y.$$

Furthermore,

$$C \neq C_A \cap C_B,$$

since otherwise all nodes in C would interfere with all nodes in A *and* in B and V_k could be split into C and $A \cup B$, violating the minimality assumption. Hence, one of the two sets $C_A \setminus C_B$ or $C_B \setminus C_A$ is not empty, say $C_A \setminus C_B$ without loss of generality. Also the other subset $C_B \setminus C_A$ is not empty, since otherwise we could split V_k into A and $V_k \setminus A$.

If all nodes in C_A interfere with all nodes in C_B, then $A \cup C_B$ and $B \cup C_A$ would be a non-trivial split of the component V_k, thus there exists two nodes, $c_a \in C_A$ and $c_b \in C_B$, that do not interfere with each other. But then we can construct a path

$$x = \mathbb{1}_A \to \cdots \to \mathbb{1}_{\{c_b\}} \to \mathbb{1}_{\{c_a,c_b\}} \to \mathbb{1}_{\{c_a,c_b\}} \to \cdots \to \mathbb{1}_B = y,$$

which does not pass through the state $0 \in \mathcal{X}$. □

State $0 \in \mathcal{X}$ continues to be a bottleneck state which must be visited along any path between activity states corresponding to different components. For a node $l \in V$, denote by e_l the activity state in $\mathcal{X}$ where only the node l is active. Clearly, $l \in V_k$ if and only if $e_l \in \mathcal{X}_k$. Recall that for any two states $x, y \in \mathcal{X}$ we denoted by T_y^x the first hitting time of state y starting in state x for the activity process $\{X(t)\}_{t\geq 0}$.

Let $x, y \in \mathcal{X}$ be two activity states, with $I_x \subseteq V_{k_1}$ and $I_y \subseteq V_{k_2}$, $k_1 \neq k_2$. In order to give a stochastic representation of the transition time T_y^x, similar in spirit to (3.10), we define the following random variables for $l \in V \setminus (\{0\} \cup V_{k_2})$:

- $N_l(\nu)$: number of times the process makes a transition $0 \to e_l \in \mathcal{X}_k$, $k \neq k_2$, before the first transition to V_{k_2} occurs;
- $\hat{T}_{e_l}^0(i)$: time spent in state o before the i-th transition to state $e_l \in \mathcal{X}_k$, with $k \neq k_2$, $i = 1, \ldots, N_l(\nu)$;
- $T_0^{e_l}(i)$: time to return to state 0 after the i-th transition to state $e_l \in \mathcal{X}_k$, with $k \neq k_2$, $i = 1, \ldots, N_l(\nu)$.

Moreover, for $l \in V_{k_2}$, define $\hat{T}_{e_l}^0(\nu)$ as the time spent in state 0 before the first transition to state $e_l \in \mathcal{X}_{k_2}$. Lemma 3.8.1 implies that the transition time T_y^x may be represented as

$$T_y^x = T_0^x + \sum_{k\neq k_2} \sum_{l\in V_k} \sum_{i=1}^{N_l} (\hat{T}_{e_l}^0(i) + T_0^{e_l}(i)) + \sum_{l\in V_{k_2}} I_l(\hat{T}_{e_l}^0 + T_y^{e_l}), \tag{3.36}$$

where I_l, $l \in V_{k_2}$, are 0-1 variables with $\sum_{l\in V_{k_2}} I_l = 1$ and $\mathbb{P}(I_l = 1) = f_l(\nu)/F_{k_2}(\nu)$, $l \in V_{k_2}$, and the variables $T_0^{e_l}$, $l \in V_k$, are transition times when

considering the component V_k in isolation. Moreover, in the above representation the dependence on the parameter ν is suppressed for compactness and all the random variables representing time durations are mutually independent as well as independent of the random variables $N_l(\nu)$, $l \in V \setminus (\{0\} \cup V_{k_2})$.

In order to determine the asymptotic behavior of the transition time $T_y^x(\nu)$ as $\nu \to \infty$, we now proceed to analyze the asymptotic behavior of the escape times $T_0^{e_l}$, $l \in V_k$, in the stochastic representation. Unless stated otherwise, we henceforth let $z \in \mathcal{X}_k$ and focus on the Markov process $\{X(t)\}_{t\geq 0}$ restricted to the state space $\mathcal{X}_k^+ = \mathcal{X}_k \cup \{0\}$.

The steady-state probability of a state $u \in \mathcal{X}_k^+$ is

$$\pi_u(\nu) = \frac{1}{Z_k(\nu)} \prod_{l\in V_k} f_l(\nu)^{u_l},$$

with normalization constant

$$Z_k(\nu) = \sum_{u'\in\mathcal{X}_k^+} \prod_{l\in V_k} f_l(\nu)^{u'_l}.$$

Define

$$g_k(\nu) := \max_{u\in\mathcal{X}_k} \prod_{l\in V_k} f_l(\nu)^{u_l} \quad \text{and} \quad \eta_k := \min_{l\in V_k} \lim_{\nu\to\infty} \log_\nu f_l(\nu).$$

We make the mild technical assumptions that $\eta \in (0,\infty)$ and that $\psi_k = \lim_{\nu\to\infty} Z_k(\nu)/g_k(\nu)$ exists. Then the following two asymptotic properties of the escape time $T_0^z(\nu)$ can be established:

$$\mathbb{E}T_0^z(\nu) \sim \frac{\psi_k g_k(\nu)}{F_k(\nu)}, \quad \text{as } \nu \to \infty, \tag{3.37}$$

and

$$\frac{T_0^z(\nu)}{\mathbb{E}T_0^z(\nu)} \xrightarrow{d} \text{Exp}(1), \quad \text{as } \nu \to \infty. \tag{3.38}$$

In order to provide a brief sketch of the proof arguments, we first introduce some further useful notation. Let $N_{z,0}(\nu)$ be a random variable representing the number of visits to state 0 in between two consecutive visits to state z. Let $R_z(\nu)$ be the residence time in state z and $T_{z,z}^+(\nu)$ the first return time to state z. Noting that

$$\pi_z(\nu) = \frac{\mathbb{E}R_z(\nu)}{\mathbb{E}T_{z,z}^+(\nu)}, \quad \mathbb{E}R_z(\nu) \leq 1 \quad \text{and} \quad \pi_z(\nu) = \frac{1}{Z_k(\nu)} \prod_{l\in V_k} f_l(\nu)^{z_l},$$

we obtain

$$\frac{\nu^\eta \mathbb{E}T_{z,z}^+(\nu)}{\psi_k g_k(\nu)} = \frac{\nu^\eta \mathbb{E}R_z(\nu)}{\pi_z(\nu)\psi_k g_k(\nu)} \leq \frac{\nu^\eta}{\prod_{l\in V_k} f_l(\nu)^{z_l}} \frac{Z_k(\nu)}{\psi_k g_k(\nu)} = o(1), \quad \text{as } \nu \to \infty.$$

Using similar arguments as in [118], it may be shown that

$$\mathbb{P}\Big(T_z^u(\nu) > g_k(\nu)\nu^{-\eta_k/2}\Big) \leq r < 1$$

for all states u with $I_u \in \mathcal{B}_k$, implying (by the strong Markov property)

$$\mathbb{P}\Big(T_z^u(\nu) > ng_k(\nu)\nu^{-\eta_k/2}\Big) \leq r^n,$$

and that $\mathbb{P}\Big(T_z^0(\nu) < T_{0,0}^+(\nu)\Big) \geq s(\nu)$, with $s(\nu) \to 1$ as $\nu \to \infty$. This means that after a visit to state o, the number of additional visits to that state before the first visit to state z is stochastically bounded from above by a geometrically distributed random variable with parameter $1 - s(\nu)$. This implies

$$\mathbb{E}N_{z,0}(\nu) \sim \mathbb{P}\Big(N_{z,0}(\nu) \geq 1\Big) \quad \text{as } \nu \to \infty. \tag{3.39}$$

It may then be deduced that the distribution of $T_{z,z}^+(\nu)$ satisfies the uniform integrability condition in [62] and [62, Theorem 1] then yields the asymptotic exponentiality property in (3.38) and

$$\mathbb{E}T_0^z(\nu) \sim \frac{\mathbb{E}T_{z,z}^+(\nu)}{\mathbb{P}\Big(N_{z,0}(\nu) \geq 1\Big)}.$$

Observing that

$$\frac{\pi_0(\nu)}{\mathbb{E}R_0(\nu)} = \mathbb{E}N_{z,0}(\nu)\frac{\pi_z(\nu)}{\mathbb{E}R_z(\nu)},$$

and invoking (3.39), we deduce that the term in the right-hand side asymptotically behaves as

$$\frac{\mathbb{E}T_{z,z}^+(\nu)\pi_z(\nu)\mathbb{E}R_0(\nu)}{\pi_0(\nu)\mathbb{E}R_z(\nu)} = \frac{\mathbb{E}R_0(\nu)}{\pi_0(\nu)} = \frac{Z_k(\nu)}{F_k(\nu)} \sim \frac{\psi_k g_k(\nu)}{F_k(\nu)},$$

yielding (3.37) as stated.

The two asymptotic properties (3.37) and (3.38) for the order of magnitude and the scaled distribution of the escape time $T_0^z(\nu)$ mirror those stated in (3.7) and Proposition 3.3.6. Using these two properties and the stochastic representation (3.36), similar results can be established for the asymptotic behavior of the transition time $T_y^x(\nu)$ as in Theorems 3.2.1 and 3.2.2. For any $l \in V_k$, define

$$\Theta_l(\nu) = \frac{f_l(\nu)g_k(\nu)}{F_k(\nu)}.$$

In this case the set K^* needs to be defined as those $l \in \bigcup_{k \neq k_2} V_k$ such that $\lim_{\nu\to\infty} \Theta_l(\nu)/\Theta_m(\nu) > 0$ for all $m \in \bigcup_{k\neq k_2} V_k$. Also, additional conditions need to be imposed in order to ensure that

$$\sum_{l \in V_{k_2}} \frac{f_l(\nu)}{F_{k_2}(\nu)}\mathbb{E}T_y^{e_l}(\nu) = o(\mathbb{E}T_y^x(\nu)), \quad \text{as } \nu \to \infty,$$

which guarantees that the expected time to reach state y, once the process hits the target component V_{k_2}, is asymptotically small with respect to the overall transition time $T_y^x(\nu)$.

APPENDIX

3.A *Proof of Lemma 3.3.3*

We will prove a slightly more general version of Lemma 3.3.3, assuming the rates of the process are those described in Subsection 3.3.3. Order the state space as $\mathcal{X}^* = \{L, L-1, \dots, 1, 0\}$ and consider the generator matrix $Q(\nu)$ of the process $\{X(t)\}_{t\geq 0}$ with 0 an absorbing state. That is,

$$Q(\nu) = \begin{bmatrix} q_L & d_L & 0 & & \dots & & 0 \\ a_{L-1}f(\nu) & q_{L-1} & d_{L-1} & 0 & & & \\ 0 & a_{L-2}f(\nu) & q_{L-2} & d_{L-2} & 0 & & \\ \vdots & & \ddots & \ddots & \ddots & & \\ & & & 0 & a_1 f(\nu) & q_1 & d_1 \\ 0 & \dots & & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the diagonal elements are $q_l(\nu) = -a_l f(\nu) - d_l$ for $l = 1, \dots, L_1$ and $q_L(\nu) = -d_L$. The matrix $Q(\nu)$ can be written as

$$Q(\nu) = \begin{pmatrix} \mathbf{T}(\nu) & \mathbf{t}(\nu) \\ \mathbf{0}^{\mathrm{T}} & 0 \end{pmatrix},$$

where $\mathbf{T}(\nu)$ is an $L \times L$ invertible matrix. Since the characteristic polynomials of $-Q(\nu)$ and $-\mathbf{T}(\nu)$ satisfy the relation $p_{-Q(\nu)}(z) = -z\, p_{-\mathbf{T}(\nu)}(z)$, the spectrum of $-Q(\nu)$ consists of that of $-\mathbf{T}(\nu)$ plus the eigenvalue zero with multiplicity one. Denote by $D(\nu)$ the $L \times L$ diagonal matrix, whose diagonal entries are $\{\sqrt{\xi_l(\nu)}\}_{l=L,L-1,\dots,1}$, where the ξ's are the so-called *potential coefficients*, defined recursively as

$$\xi_L(\nu) = 1 \quad \text{and} \quad \xi_l(\nu) = \frac{d_{l+1}}{a_l f(\nu)} \xi_{l+1}(\nu), \quad l = L-1, \dots, 1.$$

The $L \times L$ matrix $G(\nu) = -D(\nu)^{1/2}\mathbf{T}(\nu)D(\nu)^{-1/2}$ is tridiagonal and symmetric with diagonal entries $G_{l,l}(\nu) = q_{L-l+1}(\nu)$ and $G_{l,l+1}(\nu) = g_{l+1,l}(\nu) = -\sqrt{d_l a_{l-1} f(\nu)}$. Since $G(\nu)$ is similar to $-\mathbf{T}(\nu)$, they have the same spectrum. Denote by $\mathcal{D}(p, R)$ the closed disc centered in p with radius R, i.e. $\mathcal{D}(p, R) = \{z \in \mathbb{C} \ : \ |z - p| \leq R\}$. Consider the *Gershgorin discs* $\{\mathcal{D}_l(\nu)\}_{l=1}^L$ of $G(\nu)$,

defined as $\mathcal{D}_l(\nu) := \mathcal{D}(-q_l(\nu), R_l(\nu))$, where the radius $R_l(\nu)$ is the sum of the absolute values of the non-diagonal entries in the $L-l+1$-th row, i.e. $R_l(\nu) := \sum_{m \neq L-l+1} |G_{L-l+1,m}(\nu)|$. Then

$$\begin{aligned}
\mathcal{D}_L(\nu) &= \mathcal{D}(d_L, \sqrt{d_L a_{L-1} f(\nu)}),\\
\mathcal{D}_{L-1}(\nu) &= \mathcal{D}(d_{L-1} + a_{L-1} f(\nu), \sqrt{d_L a_{L-1} f(\nu)} + \sqrt{d_{L-1} a_{L-2} f(\nu)}),\\
&\vdots\\
\mathcal{D}_2(\nu) &= \mathcal{D}(d_2 + a_2 f(\nu), \sqrt{d_3 a_2 f(\nu)} + \sqrt{d_2 a_1 f(\nu)}),\\
\mathcal{D}_1(\nu) &= \mathcal{D}(d_1 + a_1 f(\nu), \sqrt{d_2 a_1 f(\nu)}).
\end{aligned}$$

We now exploit the second Gershgorin circle theorem, which is reproduced here for completeness.

Theorem (Gershgorin circle theorem)**.** *If the union of j Gershgorin discs of a real $r \times r$ matrix A is disjoint from the union of the other $r-j$ Gershgorin discs, then the former union contains exactly j and the latter the remaining $r-j$ eigenvalues of A.*

In our case, for ν sufficiently large, the disc $\mathcal{D}_L(\nu)$ does not intersect with the union $\bigcup_{l=1}^{L-1} \mathcal{D}_l(\nu)$, thus the smallest eigenvalue $\theta_1(\nu)$ lies in $\mathcal{D}_L(\nu)$ and the other $L-1$ ones in $\bigcup_{l=1}^{L-1} \mathcal{D}_l(\nu)$. Hence, for ν sufficiently large, the following inequalities hold

$$\theta_1(\nu) \leq A + B\sqrt{f(\nu)}, \quad \text{and} \quad \theta_i(\nu) \geq Cf(\nu) - D\sqrt{f(\nu)}, \quad i = 2, \dots, L,$$

where $A, B, C, D \in \mathbb{R}_+$ and, more precisely,

$$\begin{aligned}
&A = d_L, \quad B = \sqrt{d_L a_{L-1}}, \quad C = \min_{l=1,\dots,L-1} a_l \quad \text{and}\\
&D = \max\{\sqrt{d_L a_{L-1}} + \sqrt{d_{L-1} a_{L-2}}, \dots, \sqrt{d_3 a_2} + \sqrt{d_2 a_1}, \sqrt{d_2 a_1}\}.
\end{aligned}$$

Therefore, for ν sufficiently large,

$$0 < \frac{\theta_1(\nu)}{\theta_i(\nu)} \leq \frac{A + B\sqrt{f(\nu)}}{Cf(\nu) - D\sqrt{f(\nu)}},$$

and so $\lim_{\nu \to \infty} \theta_1(\nu)/\theta_i(\nu) = 0$ for $i = 2, \dots, L$. Hence,

$$\mathbb{E}T_0^L(\nu) \cdot \theta_1(\nu) = 1 + \sum_{i=2}^{L} \frac{\theta_1(\nu)}{\theta_i(\nu)} \to 1, \quad \text{as } \nu \to \infty,$$

while for $2 \leq i \leq L$,

$$\mathbb{E}T_0^L(\nu) \cdot \theta_i(\nu) > \frac{\theta_i(\nu)}{\theta_1(\nu)} \to \infty, \quad \text{as } \nu \to \infty.$$

3.B *Proof of Lemma 3.3.5*

The proof of statement (i) consists of a lower and an upper bound which asymptotically coincide. Indeed, using the bounds in property (2), one obtains that

$$\liminf_{\nu\to\infty} \frac{\mathbb{E}T(\nu)}{\mathbb{E}U(\nu)} \geq \liminf_{\nu\to\infty} \frac{\mathbb{E}U(\nu) - \mathbb{E}V(\nu)}{\mathbb{E}U(\nu)} = 1,$$

and

$$\limsup_{\nu\to\infty} \frac{\mathbb{E}T(\nu)}{\mathbb{E}U(\nu)} \leq \limsup_{\nu\to\infty} \frac{\mathbb{E}U(\nu) + \mathbb{E}W(\nu)}{\mathbb{E}U(\nu)} = 1.$$

Once again, the proof of statement (ii) consists of a lower and an upper bound which asymptotically coincide for all the continuity points of the tail distribution of Z, which will be denoted by $F(s) = \mathbb{P}(Z > s)$.

For the lower bound, argue as follows. Property (1) implies that for any $\delta \in (0,1)$, $\mathbb{E}W(\nu) \leq \delta\mathbb{E}U(\nu)$ for ν sufficiently large. Thus, using the lower bound in property (3), for ν sufficiently large,

$$\begin{aligned}
\mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) &\geq \mathbb{P}\Big(\frac{U(\nu) - V(\nu)}{\mathbb{E}U(\nu) + \mathbb{E}W(\nu)} > t\Big) \\
&\geq \mathbb{P}\Big(U(\nu) - V(\nu) > \mathbb{E}U(\nu)(1+\delta)t\Big) \\
&\geq \mathbb{P}\Big(U(\nu) > \mathbb{E}U(\nu)(1+2\delta)t\Big) - \mathbb{P}\Big(V(\nu) > \delta\mathbb{E}U(\nu)t\Big).
\end{aligned}$$

Property (3) implies that

$$\lim_{\nu\to\infty} \mathbb{P}\Big(\frac{U(\nu)}{\mathbb{E}U(\nu)} > (1+2\delta)t\Big) = F\big((1+2\delta)\,t\big).$$

Property (1) implies that for ν sufficiently large, $\mathbb{E}V(\nu) \leq \delta^2\mathbb{E}U(\nu)$, so that

$$\mathbb{P}\Big(V(\nu) > \delta\mathbb{E}U(\nu)t\Big) \leq \frac{\mathbb{E}V(\nu)}{\delta\mathbb{E}U(\nu)t} \leq \frac{\delta}{t},$$

by Markov's inequality. Taking liminf's, we obtain

$$\liminf_{\nu\to\infty} \mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) \geq F\big((1+2\delta)t\big) - \frac{\delta}{t}.$$

Letting $\delta \downarrow 0$, we find

$$\liminf_{\nu\to\infty} \mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) \geq F\,(t)\,. \tag{3.40}$$

For the upper bound, argue as follows. Property (1) implies that for any $\delta \in (0,1)$, $\mathbb{E}V(\nu) \leq \delta \mathbb{E}U(\nu)$ for ν sufficiently large. Thus, using the upper bound in property (3), for ν sufficiently large,

$$\begin{aligned}\mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) &\leq \mathbb{P}\Big(\frac{U(\nu)+W(\nu)}{\mathbb{E}U(\nu)-\mathbb{E}V(\nu)} > t\Big)\\ &\leq \mathbb{P}\Big(U(\nu)+W(\nu) > (1-\delta)\mathbb{E}U(\nu)t\Big)\\ &\leq \mathbb{P}\Big(U(\nu) > (1-2\delta)\mathbb{E}U(\nu)t\Big) + \mathbb{P}\Big(W(\nu) > \delta\mathbb{E}U(\nu)t\Big).\end{aligned}$$

Property (3) implies that

$$\lim_{\nu\to\infty} \mathbb{P}\Big(\frac{U(\nu)}{\mathbb{E}U(\nu)} > (1-2\delta)t\Big) = F\big((1-2\delta)t\big).$$

Property (1) implies that for ν sufficiently large, $\mathbb{E}W(\nu) \leq \delta^2 \mathbb{E}U(\nu)$, so that

$$\mathbb{P}\Big(W(\nu) > \delta\mathbb{E}U(\nu)t\Big) \leq \frac{\mathbb{E}W(\nu)}{\delta\mathbb{E}U(\nu)t} \leq \frac{\delta}{t},$$

by Markov's inequality. Taking limsup's, we obtain

$$\limsup_{\nu\to\infty} \mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) \leq F\big((1-2\delta)\,t\big) + \frac{\delta}{t}.$$

Letting $\delta \downarrow 0$, we find

$$\limsup_{\nu\to\infty} \mathbb{P}\Big(\frac{T(\nu)}{\mathbb{E}T(\nu)} > t\Big) \leq F\,(t)\,. \tag{3.41}$$

Combining (3.40) and (3.41) completes the proof.

3.C *Proof of Lemma 3.4.3*

Statement (i) is trivial, since $0 \leq \mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(s) \leq 1$ for all $s \in [0,\infty[$ and $\beta_k = \lim_{\nu\to\infty} \mathbb{E}N_k(\nu) = 0$ for every $k \in \mathcal{N}$. Statement (ii) follows immediately after substituting $\beta_k \in \mathbb{R}_+$ and $\gamma_k \in (0,1]$ in the limit, since $T_k(\nu)/\mathbb{E}T_k(\nu) \xrightarrow{d} \mathrm{Exp}(1)$. The proof of claim (iii) is more involved and we need an auxiliary lemma.

Let $S_k(\nu) := \sum_{i=1}^{N_k(\nu)} T_k^{(i)}(\nu)$. From the integrability of $N_k(\nu)$ and $T_k(\nu)$ it follows that $\mathbb{E}S_k(\nu) = \mathbb{E}N_k(\nu) \cdot \mathbb{E}T_k(\nu)$. Consider the random variable

$$\tilde{S}(\nu) = h_k(\nu) \sum_{i=1}^{N_k(\nu)} \frac{T_k^{(i)}(\nu)}{\mathbb{E}T_k^{(i)}(\nu)} = h_k(\nu)\mathbb{E}N_k(\nu)\frac{S_k(\nu)}{\mathbb{E}S_k(\nu)}.$$

Since $N_k(\nu)$ has a geometric distribution, the Laplace transform of the random variable $S_k(\nu)/\mathbb{E}S_k(\nu)$ is given by

$$\begin{aligned}\mathcal{L}_{S_k(\nu)/\mathbb{E}S_k(\nu)}(s) &= G_{N_k(\nu)}\left(\mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(s/\mathbb{E}N_k(\nu))\right)\\ &= \frac{1}{1+(1-\mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(s/\mathbb{E}N_k(\nu)))\cdot\mathbb{E}N_k(\nu)},\end{aligned}$$

and, hence,

$$\begin{aligned}\mathcal{L}_{\tilde{S}(\nu)}(s) &= \mathcal{L}_{S_k(\nu)/\mathbb{E}S_k(\nu)}(sh_k(\nu)\mathbb{E}N_k(\nu))\\ &= \frac{1}{1+\left(1-\mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\right)\cdot\mathbb{E}N_k(\nu)}.\end{aligned} \tag{3.42}$$

One can check that, if $k\in\mathcal{S}$, then

(a) $N_k(\nu)/\mathbb{E}N_k(\nu)\xrightarrow{d}\mathrm{Exp}(1)$, and

(b) $\lim_{\nu\to\infty}\dfrac{\mathrm{Var}\left(T_k(\nu)/\mathbb{E}T_k(\nu)\right)}{\mathbb{E}N_k(\nu)}=0.$

Claim (a) is a standard result for geometric random variables, which uses only the fact that $\lim_{\nu\to\infty}\mathbb{E}N_k(\nu)=\infty$ for $k\in\mathcal{S}$. Moreover, since $T_k(\nu)$ is a first-passage time of a birth-and-death process, using [73, Corollary 4], we can obtain explicitly the asymptotic order of magnitude of $\mathrm{Var}(T_k(\nu))$ and prove that $\lim_{\nu\to\infty}\mathrm{Var}(T_k(\nu))/(\mathbb{E}T_k(\nu))^2=1$, from which claim (b) follows.

These two facts and the technical lemma below, see Lemma 3.C.1, imply that $S_k(\nu)/\mathbb{E}S_k(\nu)\xrightarrow{d}\mathrm{Exp}(1)$, and hence

$$\lim_{\nu\to\infty}\mathcal{L}_{\tilde{S}(\nu)}(s)=\lim_{\nu\to\infty}\mathcal{L}_{S_k(\nu)/\mathbb{E}S_k(\nu)}\left(sh_k(\nu)\mathbb{E}N_k(\nu)\right)=\frac{1}{1+s\gamma_k}.$$

This fact, together with (3.42), implies that

$$\lim_{\nu\to\infty}\left(1-\mathcal{L}_{T_k(\nu)/\mathbb{E}T_k(\nu)}(sh_k(\nu))\right)\cdot\mathbb{E}N_k(\nu)=s\gamma_k,$$

and the proof of statement (iii) is concluded. We now state and prove the technical lemma mentioned above.

Lemma 3.C.1. *For every $\nu>0$, let $(X_i(\nu))_{i\in\mathbb{N}}$ be a sequence of i.i.d. integrable random variables, $X_i(\nu)\overset{d}{=}X(\nu)$ and let $N(\nu)$ be an integer-valued random variable, independent of all the $X_i(\nu)$'s and integrable. Assume further that*

(1) $N(\nu)/\mathbb{E}N(\nu)\xrightarrow{d}Z$ *as* $\nu\to\infty$, *with* $\mathbb{P}(Z=0)=0$;

(2) $\lim_{\nu\to\infty}\dfrac{\mathrm{Var}\left(X(\nu)/\mathbb{E}X(\nu)\right)}{\mathbb{E}N(\nu)}=0.$

Define $S(\nu) := \sum_{i=1}^{N(\nu)} X_i(\nu)$. *Then,*

$$\frac{S(\nu)}{\mathbb{E}S(\nu)} \xrightarrow{d} Z, \quad \text{as } \nu \to \infty.$$

Proof. Wald's identity guarantees that $S_N(\nu)$ is integrable and that $\mathbb{E}S(\nu) = \mathbb{E}N(\nu)\mathbb{E}X(\nu)$ for every $\nu > 0$. Hence, without loss of generality, we can assume that $\mathbb{E}X(\nu) = 1$ and study the asymptotic distribution of $S(\nu)/\mathbb{E}N(\nu)$. Define $S_n(\nu) := \sum_{i=1}^{n} X_i(\nu)$. Note that we can rewrite

$$\frac{S(\nu)}{\mathbb{E}N(\nu)} = \frac{S(\nu)}{N(\nu)} \frac{N(\nu)}{\mathbb{E}N(\nu)}.$$

Now we claim that

$$\frac{S(\nu)}{N(\nu)} \xrightarrow{\mathbb{P}} 1, \quad \text{as } \nu \to \infty. \tag{3.43}$$

Indeed, for every $\epsilon > 0$ we may write

$$\begin{aligned}
&\mathbb{P}\Big(|S(\nu)/N(\nu) - \mathbb{E}X(\nu)| > \delta\Big) \\
&\quad = \sum_{n=0}^{\infty} \mathbb{P}\Big(|S(\nu)/N(\nu) - \mathbb{E}X(\nu)| > \delta,\, N(\nu) = n\Big) \\
&\quad = \sum_{n=0}^{\lfloor \epsilon \mathbb{E}N(\nu) \rfloor} \mathbb{P}(\ldots) + \sum_{n=\lfloor \epsilon \mathbb{E}N(\nu) \rfloor + 1}^{\infty} \mathbb{P}(\ldots) \\
&\quad \leq \sum_{n=0}^{\lfloor \epsilon \mathbb{E}N(\nu) \rfloor} \mathbb{P}(N(\nu) = n) \\
&\qquad + \sum_{n=\lfloor \epsilon \mathbb{E}N(\nu) \rfloor + 1}^{\infty} \mathbb{P}\Big(|S(\nu)/N(\nu) - \mathbb{E}X(\nu)| > \delta,\, N(\nu) = n\Big) \\
&\quad \leq \mathbb{P}(N(\nu) \leq \epsilon \mathbb{E}N(\nu)) \\
&\qquad + \sum_{n=\lfloor \epsilon \mathbb{E}N(\nu) \rfloor + 1}^{\infty} \mathbb{P}\Big(|S_n(\nu)/n - \mathbb{E}X(\nu)| > \delta\Big) \mathbb{P}(N(\nu) = n).
\end{aligned}$$

Using Chebyshev's inequality, we find that the second term is bounded from above by

$$\begin{aligned}
&\sum_{n=\lfloor \epsilon \mathbb{E}N(\nu) \rfloor + 1}^{\infty} \frac{\mathrm{Var}(X(\nu))}{\delta^2 n} \mathbb{P}(N(\nu) = n) \\
&\qquad \leq \frac{\mathrm{Var}(X(\nu))}{\epsilon \mathbb{E}N(\nu) \delta^2} \sum_{n=\lfloor \epsilon \mathbb{E}N(\nu) \rfloor + 1}^{\infty} \mathbb{P}(N(\nu) = n) \\
&\qquad \leq \frac{\mathrm{Var}(X(\nu))}{\epsilon \mathbb{E}N(\nu) \delta^2}.
\end{aligned}$$

For every $\xi > 0$ and every $\epsilon > 0$, there exists $\nu_{\epsilon,\xi} > 0$ such that for $\nu > \nu_{\epsilon,\xi}$,

$$\mathbb{P}(N(\nu) \leq \epsilon \mathbb{E}N(\nu)) \leq \mathbb{P}(Z \leq \epsilon) + \xi \quad \text{and} \quad \frac{\mathrm{Var}X(\nu)}{\epsilon \mathbb{E}N(\nu)} \leq \xi.$$

The first inequality follows from the fact that the c.d.f. of $N(\nu)/\mathbb{E}N(\nu)$ converges pointwise to that of Z, by assumption (1). On the other hand, the second inequality follows immediately from assumption (2). Therefore, for $\nu > \nu_{\epsilon,\xi}$,

$$\mathbb{P}\Big(|S(\nu)/N(\nu) - \mathbb{E}X(\nu)| > \delta\Big) \leq \mathbb{P}(Z \leq \epsilon) + \xi + \frac{\xi}{\delta^2}.$$

Take e.g. $\epsilon = \delta$, $\xi = \delta^3$. Then for ν sufficiently large,

$$\mathbb{P}\Big(|S(\nu)/N(\nu) - \mathbb{E}X(\nu)| > \delta\Big) \leq \mathbb{P}(Z \leq \delta) + \delta^3 + \delta.$$

Letting $\delta \downarrow 0$ and using the fact that $\mathbb{P}(Z = 0) = 0$, we obtain (3.43). Then assumption (1) and Slutsky's theorem imply the conclusion. □

3.D *Proof of Lemma 3.7.2*

In this proof we use the same notation introduced in Section 3.6, but without writing explicitly the dependence on ν. Denote by μ^S the function defined by

$$\mu^S(x) := \begin{cases} \pi_x & \text{if } x \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Note that since $\sum_{y \in \mathcal{X}^*} \mu^S(y) = \pi_S$, in general μ^S is not a probability distribution, but μ^S/π_S always is. Denote by μ_t^S the distribution on $\mathcal{X}^*$ at time t of the Markov process $\{X(t)\}_{t \geq 0}$ if the initial distribution was equal to π conditioned on being in S, i.e. $\mu_0^S := \mu^S/\pi_S$. For every $y \in \mathcal{X}^*$ we can write

$$\mu_t^S(y) = \sum_{x \in \mathcal{X}^*} \frac{\mu^S(x)}{\pi_S} P^t(x,y) = \sum_{x \in S} \frac{\pi_x}{\pi_S} P^t(x,y). \tag{3.44}$$

We now claim that

$$\pi_S ||\mu_t^S - \mu_0^S||_{\mathrm{TV}} = \sum_{x \in S} \sum_{y \in S^c} \pi_x P^t(x,y) \tag{3.45}$$

and

$$\sum_{x \in S} \sum_{y \in S^c} \pi_x P^t(x,y) \leq t \cdot Q_{S,S^c}. \tag{3.46}$$

Let us first prove equality (3.45). A well-known characterization of the total variation distance (see for instance [99, Remark 4.3]) states that if μ, μ' are two probability distributions on $\mathcal{X}^*$, then

$$||\mu - \mu'||_{\mathrm{TV}} = \sum_{y \in \mathcal{X}^* : \mu(y) \geq \mu'(y)} \left[\mu(y) - \mu'(y)\right].$$

Applying this to the distributions μ_t^S and μ_0^S, we get

$$||\mu_t^S - \mu_0^S||_{\mathrm{TV}} = \sum_{y \in \mathcal{X}^* : \mu_t^S(y) \geq \mu_0^S(y)} \left[\mu_t^S(y) - \mu_0^S(y)\right]. \tag{3.47}$$

If $y \in S$, then

$$\mu_t^S(y) = \sum_{x \in S} \frac{\pi_x}{\pi_S} P^t(x,y) \leq \sum_{x \in \mathcal{X}^*} \frac{\pi_x}{\pi_S} P^t(x,y) = \frac{\pi_y}{\pi_S} = \mu_0^S(y),$$

and therefore we can restrict the sum on the right-hand side of (3.47) to the states $y \in S^c$. Moreover, if $y \in S^c$, then by definition $\mu_0^S(y) = 0$. Therefore,

$$||\mu_t^S - \mu_0^S||_{\mathrm{TV}} = \sum_{y \in S^c} \mu_t^S(y).$$

By multiplying both sides of the last equality by π_S and using (3.44), we get

$$\pi_S ||\mu_t^S - \mu_0^S||_{\mathrm{TV}} = \pi_S \sum_{y \in S^c} \mu_t^S(y) = \pi_S \sum_{y \in S^c} \sum_{x \in S} \frac{\pi_x}{\pi_S} P^t(x,y) = \sum_{x \in S} \sum_{y \in S^c} \pi_x P^t(x,y),$$

and (3.45) is proved.

We now turn our attention to the claim in (3.46). For any $x, y \in \mathcal{X}^*$, define the random variable $N_{x \to y}(t)$ as the number of transitions from state x to state y during the time interval $[0, t]$, so that

$$\mathbb{P}\Big(X(t) \in S^c \mid X(0) = x\Big) \leq \sum_{x' \in S,\, y' \in S^c} \mathbb{P}\Big(N_{x' \to y'}(t) \geq 1 \mid X(0) = x\Big),$$

from which we get that

$$
\begin{aligned}
&\sum_{x\in S}\sum_{y\in S^c}\pi_x P^t(x,y)\\
&\quad=\sum_{x\in S}\pi_x\mathbb{P}\Big(X(t)\in S^c \mid X(0)=x\Big)\\
&\quad\le\sum_{x\in S}\pi_x\sum_{x'\in S,\,y'\in S^c}\mathbb{P}\Big(N_{x'\to y'}(t)\ge 1 \mid X(0)=x\Big)\\
&\quad=\sum_{x\in S}\pi_x\sum_{x'\in S,\,y'\in S^c}\int_{u=0}^{t}\mathbb{P}\Big(X(u)=x' \mid X(0)=x\Big)\,q(x',y')\,\mathrm{d}u\\
&\quad=\sum_{x'\in S,\,y'\in S^c}q(x',y')\int_{u=0}^{t}\left(\sum_{x\in S}\pi_x\mathbb{P}\Big(X(u)=x' \mid X(0)=x\Big)\right)\mathrm{d}u\\
&\quad\le\sum_{x'\in S,\,y'\in S^c}q(x',y')\int_{u=0}^{t}\pi_{x'}\,\mathrm{d}u\\
&\quad=t\sum_{x'\in S}\sum_{y'\in S^c}\pi_{x'}q(x',y')=t\cdot Q_{S,S^c},
\end{aligned}
$$

and (3.46) is then proved. Note that we used the fact that

$$
\begin{aligned}
\sum_{x\in S}\pi_x\mathbb{P}\Big(X(u)=x' \mid X(0)=x\Big)&\le\sum_{x\in\mathcal{X}^*}\pi_x\mathbb{P}\Big(X(u)=x' \mid X(0)=x\Big)\\
&=\sum_{x\in\mathcal{X}^*}\pi_x P^u(x,x')=\pi_{x'}.
\end{aligned}
$$

From (3.45) and (3.46) it follows that

$$||\mu_t^S-\mu_0^S||_{\mathrm{TV}}\le t\,\Phi(S).$$

Assume that the subset S is such that $\pi_S\le r$, with $r\in[0,1]$. Then, since $\mu^S(S^c)=0$,

$$||\mu_0^S-\pi||_{\mathrm{TV}}=\max_{A\subset\mathcal{X}^*}|\pi_A-\mu_0^S(A)|\ge\pi_{S^c}=1-\pi_S\ge 1-r,$$

and, using the triangular inequality, we get

$$1-r\le||\mu_0^S-\pi||_{\mathrm{TV}}\le||\mu_0^S-\mu_t^S||_{\mathrm{TV}}+||\mu_t^S-\pi||_{\mathrm{TV}}.$$

Since the state space $\mathcal{X}$ is finite, we have the following equality for the distance from stationarity:

$$d(t)=\max_{x\in\mathcal{X}^*}\|P^t(x,\cdot)-\pi(\nu)\|_{\mathrm{TV}}=\sup_{\mu\in\mathcal{P}(\mathcal{X}^*)}\|\mu P^t-\pi(\nu)\|_{\mathrm{TV}},$$

where $\mathcal{P}(\mathcal{X}^*)$ denotes the collection of all probability distributions on $\mathcal{X}^*$. By taking $\mu=\mu^S$ and time $t=t_{\mathrm{mix}}(\epsilon)$, by definition of mixing time, we have that $||\mu_t^S-\pi||_{\mathrm{TV}}\le d(t_{\mathrm{mix}}(\epsilon))=\epsilon$ and therefore

$$1-r\le t_{\mathrm{mix}}(\epsilon)\Phi(S)+\epsilon.$$

Rearranging and minimizing over S concludes the proof.

4

HITTING TIME ASYMPTOTICS FOR METROPOLIS MARKOV CHAINS

This chapter is dedicated to reversible Freidlin-Wentzel Markov chains with Metropolis transition probabilities and, in particular, to the study of the asymptotic behavior of first hitting times in the low-temperature regime. The main contribution of this chapter is the extension of the model-independent framework [106] in order to obtain asymptotic results for the hitting time τ_A^x for *any* starting state x, not necessarily metastable, and *any* target subset A, not necessarily the set of stable configurations. In particular, we identify the two crucial exponents $\Gamma_-(x,A)$ and $\Gamma_+(x,A)$ that appear in the upper and lower bounds in probability for τ_A^x in the low-temperature regime. These two exponents might be hard to derive for a given model and, in general, they are not equal. However, we obtain a sufficient condition that guarantees that they coincide and also yields the order of magnitude of the first moment of τ_A^x on a logarithmic scale. Furthermore, we give another slightly stronger condition under which the hitting time τ_A^x normalized by its mean converges in distribution to an exponential random variable. This generalization is crucial to study asymptotic properties of transition times between dominant states of the CSMA models described in Section 1.2, which fit in this framework after uniformization, see Subsection 1.3.2 and Chapter 7 for further details.

This chapter is structured as follows. In Section 4.1 we introduce the formal definitions of the above-described framework and give an overview of our approach. Such an approach will be then developed in Section 4.2, in which we present our main results. Some of the proofs are rather technical and therefore presented later in Section 4.3.

4.1 INTRODUCTION

Let $\mathcal{X}$ be a finite state space and let $H : \mathcal{X} \to \mathbb{R}$ be the *Hamiltonian*, i.e. a non-constant *energy function*. We consider the family of Markov chains $\{X_t^\beta\}_{t\in\mathbb{N}}$ on $\mathcal{X}$ with Metropolis transition probabilities P_β indexed by a positive parameter β

$$P_\beta(x,y) := \begin{cases} c(x,y)e^{-\beta[H(y)-H(x)]^+}, & \text{if } x \neq y, \\ 1 - \sum_{z\neq x} P_\beta(x,z), & \text{if } x = y, \end{cases} \tag{4.1}$$

where $c : \mathcal{X} \times \mathcal{X} \to [0,1]$ is a matrix that does not depend on β. The matrix c is the *connectivity function* and we assume it to be

- Stochastic, i.e. $\sum_{y\in\mathcal{X}} c(x,y) = 1$ for every $x \in \mathcal{X}$;

- Symmetric, i.e. $c(x,y) = c(y,x)$ for every $x, y \in \mathcal{X}$;
- Irreducible, i.e. for any $x, y \in \mathcal{X}$, $x \neq y$, there exists a finite sequence ω of states $\omega_1, \ldots, \omega_n \in \mathcal{X}$ such that $\omega_1 = x$, $\omega_n = y$ and $c(\omega_i, \omega_{i+1}) > 0$, for $i = 1, \ldots, n-1$. We will refer to such a sequence as a *path* from x to y and we will denote it by $\omega : x \to y$.

We call the triplet $(\mathcal{X}, H, c)$ an *energy landscape*. The Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ is reversible with respect to the *Gibbs measure*

$$\mu_\beta(x) := \frac{e^{-\beta H(x)}}{\sum_{y\in\mathcal{X}} e^{-\beta H(y)}}. \tag{4.2}$$

Furthermore, it is well-known (e.g. [29, Proposition 1.1]) that the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ is aperiodic and irreducible on $\mathcal{X}$. Hence, $\{X_t^\beta\}_{t\in\mathbb{N}}$ is ergodic on $\mathcal{X}$ with stationary distribution μ_β.

For a nonempty subset $A \subset \mathcal{X}$ and a state $x \in \mathcal{X}$, we denote by τ_A^x the *first hitting time* of the subset A for the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ with initial state x at time $t = 0$, i.e.

$$\tau_A^x := \inf\{t > 0 : X_t^\beta \in A \mid X_0^\beta = x\}.$$

Let $\mathcal{X}^s$ be the set of *stable states* of the energy landscape $(\mathcal{X}, H, c)$, that is the set of global minima of H on $\mathcal{X}$. Since $\mathcal{X}$ is finite, the set $\mathcal{X}^s$ is always nonempty. Define the *stability level* $\mathcal{V}_x$ of a state $x \in \mathcal{X}$ by

$$\mathcal{V}_x := \min_{y\in\mathcal{I}_x} \Phi(x,y) - H(x) = \Phi(x, \mathcal{I}_x) - H(x), \tag{4.3}$$

where $\mathcal{I}_x := \{z \in \mathcal{X} : H(z) < H(x)\}$ is the set of states with energy lower than x and $\Phi(x,y) := \min_{\omega:x\to y} \max_{z\in\omega} H(z)$ is the communication height between x and y (see Subsection 4.2.1). We set $\mathcal{V}_x := \infty$ if $\mathcal{I}_x$ is empty, i.e. when x is a stable state. Denote by $\mathcal{X}^m$ the set of *metastable states*, which are the local minima of H in $\mathcal{X} \setminus \mathcal{X}^s$ with maximum stability level. In symbols,

$$\mathcal{X}^m := \{x \in \mathcal{X} : \mathcal{V}_x = \max_{z\in\mathcal{X}\setminus\mathcal{X}^s} \mathcal{V}_z\}. \tag{4.4}$$

The first hitting time τ_A^x is often called *tunneling time* when x is a stable state and the target set is some $A \subseteq \mathcal{X}^s \setminus \{x\}$, or *transition time from metastable to stable* when $x \in \mathcal{X}^m$ and $A = \mathcal{X}^s$.

4.1.1 *Overview of the chapter*

Hitting times for Freidlin-Wentzel Markov chains are central in the mathematical study of metastability. In the literature, several different approaches have been introduced to study the time it takes for a particle system to reach a stable state starting from a metastable configuration. Two approaches have been

independently developed based on large deviations techniques: The *pathwise approach*, first introduced in [26] and then developed in [116, 117, 118], and the *graphical approach* in [27, 28, 29, 30, 31, 135]. Other approaches to metastability are the *potential theoretic approach* [21, 22, 23] and, more recently introduced, the *martingale approach* [11, 12, 14], see [37] for a more detailed review.

In this thesis we follow the pathwise approach, which has already been used to study many finite-volume models in a low-temperature regime, see [13, 34, 35, 36, 39, 40, 56, 88, 41, 42, 87, 110, 111, 113, 114], where the state space is seen as an energy landscape and the paths which the Markov chain will most likely follow are those with a minimum energy barrier. In [116, 117, 118] the authors derive general results for first hitting times for the transition from metastable to stable states, the critical configurations (or bottlenecks) that the system visits during this transition and the tube of trajectories that the system typically follows on its transition to the stable state. These are the three main problems usually tackled in the study of metastability. In [106] the results on hitting times are obtained with minimal model-dependent knowledge, since it is only necessary to find all the metastable states and the minimal energy barrier which separates them from the stable states.

In this chapter we generalize the classical pathwise approach [106] to study the first hitting time τ_A^x for a Metropolis Markov chain for any pair of starting state x and target subset A. The interest of extending these results to the tunneling time between two stable states was already mentioned in [106, 118], but our framework is even more general and we could study τ_A^x for any pair (x, A).

Our analysis relies on the classical notion of a *cycle*, which is a maximal connected subset of states lying below a given energy level. The exit time from a cycle in the low-temperature regime is well-known in the literature [29, 30, 37, 116, 118] and is characterized by the *depth* of the cycle, which is the minimum energy barrier that separates the bottom of the cycle from its external boundary. The usual strategy presented in the literature to study the first hitting time from $x \in \mathcal{X}^m$ to $A = \mathcal{X}^s$ is to look at the decomposition into maximal cycles (by inclusion) of the relevant part of the energy landscape, i.e. $\mathcal{X} \setminus \mathcal{X}^s$. The first model-dependent property one has to prove is that the starting state x is metastable, which guarantees that there are no cycles in $\mathcal{X} \setminus \mathcal{X}^s$ deeper than the maximal cycle containing the starting state x, denoted by $C_A(x)$. In this scenario, the time spent in maximal cycles different from $C_A(x)$, and hence the time it takes to reach $\mathcal{X}^s$ from the boundary of $C_A(x)$, is comparable to or negligible with respect to the exit time from $C_A(x)$, making the exit time from $C_A(x)$ and the first hitting time τ_A^x of the same order.

In contrast, for a general starting state x and target subset A *all* maximal cycles of $\mathcal{X} \setminus A$ can potentially have a non-negligible impact on the transition from x to A in the low-temperature regime. By analyzing these maximal cycles and the possible *cycle-paths*, we can establish bounds in probability

for the hitting time τ_A^x on a logarithmic scale, i.e. obtain a pair of exponents $\Gamma_-(x,A), \Gamma_+(x,A)$ such that for every $\epsilon > 0$

$$\lim_{\beta\to\infty} \mathbb{P}\left(e^{\beta(\Gamma_-(x,A)-\epsilon)} \leq \tau_A^x \leq e^{\beta(\Gamma_+(x,A)+\epsilon)}\right) = 1.$$

The sharpness of the exponents $\Gamma_-(x,A)$ and $\Gamma_+(x,A)$ crucially depends on how precisely one can determine which maximal cycles are likely to be visited and which ones are not, see Section 4.2 for further details. Furthermore, we give a sufficient condition (see Assumption A in Section 4.2), which is the *absence of deep typical cycles*, which guarantees that $\Gamma_-(x,A) = \Gamma = \Gamma_+(x,A)$, proving that the random variable $\beta^{-1}\log\tau_A^x$ converges in probability to Γ as $\beta\to\infty$, and that $\lim_{\beta\to\infty}\beta^{-1}\log\mathbb{E}\tau_A^x = \Gamma$. In many cases of interest, one could show that Assumption A holds for the pair (x,A) without detailed knowledge of the typical paths from x to A. Indeed, by proving that the model exhibits *absence of deep cycles* (see Proposition 4.2.18), similarly to [106], also in our framework the study of the hitting time τ_A^x is decoupled from an exact control of the typical paths from x to A. More precisely, one can obtain asymptotic results for the hitting time τ_A^x in probability, in expectation and in distribution without the detailed knowledge of the critical configuration or of the tube of typical paths. Proving the absence of deep cycles when $x \in \mathcal{X}^m$ and $A = \mathcal{X}^s$ corresponds precisely to identifying the set of metastable states $\mathcal{X}^m$, while, when $x \in \mathcal{X}^s$ and $A = \mathcal{X}^s \setminus \{x\}$, it is enough to show that the energy barrier that separates any state from a state with lower energy is not bigger than the energy barrier separating any two stable states.

Moreover, we give another sufficient condition (see Assumption B in Section 4.2), called *"worst initial state"* assumption, to show that the hitting time τ_A^x normalized by its mean converges in distribution to an exponential unit-mean random variable. However, checking Assumption B for a specific model can be very involved, and hence we provide a stronger condition (see Proposition 4.2.20), which includes the case of the tunneling time between stable states and the classical transition time from a metastable to a stable state.

The hard-core model on complete K-partite graphs, which is the discrete-time version of the activity process studied in Chapter 3, is used as an example to illustrate scenarios where Assumptions A or B are violated. Indeed, the energy landscape corresponding to this model is simple enough to allow for explicit calculations for the hitting times between any pair of configurations. In particular, we show in Subsection 4.2.7 that whenever Assumption A is not satisfied $\Gamma_-(x,A) \neq \Gamma_+(x,A)$ and the asymptotic result for $\mathbb{E}\tau_A^x$ of the first moment does not hold, while whenever Assumption B is not satisfied the scaled hitting time is not exponentially distributed.

Lastly, we show that by understanding the global structure of an energy landscape $(\mathcal{X}, H, c)$ and the maximum depths of its cycles, we can also derive results for the mixing time of the corresponding Metropolis Markov chains $\{X_t^\beta\}_{t\in\mathbb{N}}$, as illustrated in Subsection 4.2.8. In particular, our results show that in the special case of an energy landscape with multiple stable states

and without other deep cycles, the hitting time between any two stable states and the mixing time of the chain are of the same order of magnitude in the low-temperature regime. This is the case also for the Metropolis hard-core dynamics on grid graphs (see Theorems 5.2.1, 5.2.2 and 5.2.3 in Chapter 5) and triangular grid graphs (see Theorem 6.2.1 and 6.2.2 in Chapter 6).

4.2 ASYMPTOTIC BEHAVIOR OF HITTING TIMES FOR METROPOLIS MARKOV CHAINS

In this section we present model-independent results valid for any Markov chains with Metropolis transition probabilities (4.1) defined in Section 4.1. In Subsection 4.2.1 we introduce the classical notion of a *cycle*. If the considered model allows only for a very rough energy landscape analysis, well-known results for cycles are shown to readily yield upper and lower bounds in probability for the hitting time τ_A^x: Indeed, one can use the depth of the initial cycle $C_A(x)$ as $\Gamma_-(x,A)$ (see Proposition 4.2.4) and the maximum depth of a cycle in the partition of $\mathcal{X} \setminus A$ as $\Gamma_+(x,A)$ (see Proposition 4.2.7). If one has a good handle on the model-specific *optimal paths* from x to A, i.e. those paths along which the maximum energy is precisely the min-max energy barrier between x and A, sharper exponents can be obtained, as established in Proposition 4.2.10, by focusing on the *relevant cycle*, where the process $\{X_t^\beta\}_{t\in\mathbb{N}}$ started in x spends most of its time before hitting the subset A. We sharpen these bounds in probability for the hitting time τ_A^x even further with Proposition 4.2.15 by studying the *tube of typical paths from x to A* or *standard cascade*, a task that in general requires a very detailed but local analysis of the energy landscape. To complete the study of the hitting time in the regime $\beta \to \infty$, we prove in Subsection 4.2.5 the convergence of the first moment of the hitting time τ_A^x on a logarithmic scale under suitable assumptions (see Theorem 4.2.17) and give in Subsection 4.2.6 sufficient conditions for the scaled hitting time $\tau_A^x / \mathbb{E}\tau_A^x$ to converge in distribution as $\beta \to \infty$ to an exponential unit-mean random variable, see Theorem 4.2.19. Furthermore, we illustrate in detail two special cases which fall within our framework, namely the classical transition from a metastable state to a stable state and the tunneling between two stable states, which is the relevant one for the models considered in this thesis. In Subsection 4.2.7 we briefly present the hard-core model on a complete K-partite graph, which is an example of a model where the asymptotic exponentiality of the scaled hitting times does not always hold. Lastly, in Subsection 4.2.8 we present some results for the mixing time and the spectral gap of Metropolis Markov chains and show how they are linked with the critical depths of the energy landscape.

In the rest of this section and in Section 4.3, $\{X_t^\beta\}_{t\in\mathbb{N}}$ will denote a general Metropolis Markov chain with energy landscape $(\mathcal{X}, H, c)$ and inverse temperature β, as defined in Section 4.1.

4.2.1 *Cycles: Definitions and classical results*

We recall here the definition of a cycle and present some well-known properties.

A path $\omega : x \to y$ has been defined in Section 4.1 as a finite sequence of states $\omega_1, \dots, \omega_n \in \mathcal{X}$ such that $\omega_1 = x$, $\omega_n = y$ and $c(\omega_i, \omega_{i+1}) > 0$, for $i = 1, \dots, n-1$. A subset $A \subset \mathcal{X}$ with at least two elements is *connected* if for all $x, y \in A$ there exists a path $\omega : x \to y$, such that $\omega_i \in A$ for every $i = 1, \dots, |\omega|$. Given a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$, we define $\Omega_{x,A}$ as the collection of all paths $\omega : x \to y$ for some $y \in A$ that do not visit A before hitting y, i.e.

$$\Omega_{x,A} := \{\omega : x \to y \mid y \in A,\ \omega_i \notin A\ \forall\, i < |\omega|\}. \tag{4.5}$$

We remark that only the endpoint of each path in $\Omega_{x,A}$ belongs to A. Given a path $\omega = (\omega_1, \dots, \omega_n)$ in $\mathcal{X}$, we denote by $|\omega| := n$ its *length* and define its *height* or *elevation* by

$$\Phi_\omega := \max_{i=1,\dots,|\omega|} H(\omega_i). \tag{4.6}$$

The *communication height* between a pair $x, y \in \mathcal{X}$ is the minimum value that has to be reached by the energy in every path $\omega : x \to y$, i.e.

$$\Phi(x, y) := \min_{\omega : x \to y} \Phi_\omega. \tag{4.7}$$

Given two nonempty disjoint subsets $A, B \subset \mathcal{X}$, we define the communication height between A and B by

$$\Phi(A, B) := \min_{x \in A, y \in B} \Phi(x, y). \tag{4.8}$$

Given a nonempty set $A \subset \mathcal{X}$, we define its *external boundary* by

$$\partial A := \{y \notin A \mid \exists\, x \in A \,:\, c(x, y) > 0\}.$$

For a nonempty set $A \subset \mathcal{X}$ we define its *bottom* $\mathcal{F}(A)$ as the set of all minima of the energy function $H(\cdot)$ on A, i.e.

$$\mathcal{F}(A) := \{y \in A \mid H(y) = \min_{x \in A} H(x)\}.$$

We call a nonempty subset $C \subset \mathcal{X}$ a *cycle* if it is either a singleton or a connected set such that

$$\max_{x \in C} H(x) < H(\mathcal{F}(\partial C)). \tag{4.9}$$

In other words, cycles are connected subsets of sublevel sets of the energy function H. A cycle C for which condition (4.9) holds is called *non-trivial cycle*. If C is a non-trivial cycle, we define its *depth* as

$$\Gamma(C) := H(\mathcal{F}(\partial C)) - H(\mathcal{F}(C)). \tag{4.10}$$

Any singleton $C = \{x\}$ for which condition (4.9) does not hold is called *trivial cycle*. We set the depth of a trivial cycle C to be equal to zero, i.e. $\Gamma(C) = 0$. Given a non-trivial cycle C, we will refer to the set $\mathcal{F}(\partial C)$ of minima on its boundary as its *principal boundary*. Note that

$$\Phi(C, \mathcal{X} \setminus C) = \begin{cases} H(x) & \text{if } C = \{x\} \text{ is a trivial cycle,} \\ H(\mathcal{F}(\partial C)) & \text{if } C \text{ is a non-trivial cycle.} \end{cases}$$

In this way, we have the following alternative expression for the depth of a cycle C, which has the advantage of being valid also for trivial cycles:

$$\Gamma(C) = \Phi(C, \mathcal{X} \setminus C) - H(\mathcal{F}(C)). \tag{4.11}$$

The next lemma gives an equivalent characterization of a cycle.

Lemma 4.2.1. *A nonempty subset $C \subset \mathcal{X}$ is a cycle if and only if it is either a singleton or a connected set that satisfies*

$$\max_{x,y \in C} \Phi(x, y) < \Phi(C, \mathcal{X} \setminus C).$$

The proof easily follows from definitions (4.7), (4.8) and (4.9) and the fact that if C is not a singleton and is connected, then

$$\max_{x,y \in C} \Phi(x, y) = \max_{x \in C} H(x). \tag{4.12}$$

We remark that the equivalent characterization of a cycle given in Lemma 4.2.1 is the "correct definition" of a cycle in the case where the transition probabilities are not necessarily Metropolis but satisfy the more general *Freidlin-Wentzell condition*

$$\lim_{\beta \to \infty} -\frac{1}{\beta} \log P_\beta(x, y) = \Delta(x, y) \quad \forall\, x, y \in \mathcal{X}, \tag{4.13}$$

where $\Delta(x, y)$ is an appropriate *rate function* $\Delta : \mathcal{X}^2 \to \mathbb{R}^+ \cup \{\infty\}$. The Metropolis transition probabilities correspond to the case (see [38] for more details) where

$$\Delta(x, y) = \begin{cases} [H(y) - H(x)]^+ & \text{if } c(x, y) > 0, \\ \infty & \text{otherwise.} \end{cases}$$

The next theorem collects well-known results for the asymptotic behavior of the exit time from a cycle as β becomes large, where the depth $\Gamma(C)$ of the cycle plays a crucial role.

Theorem 4.2.2 (Properties of the exit time from a cycle)**.** *Consider a non-trivial cycle $C \subset \mathcal{X}$.*

(i) *For any $x \in C$ and for any $\epsilon > 0$, there exists $k_1 > 0$ such that for all β sufficiently large*

$$\mathbb{P}\Big(\tau_{\partial C}^x < e^{\beta(\Gamma(C)-\epsilon)}\Big) \leq e^{-k_1\beta}.$$

(ii) *For any $x \in C$ and for any $\epsilon > 0$, there exists $k_2 > 0$ such that for all β sufficiently large*

$$\mathbb{P}\Big(\tau_{\partial C}^x > e^{\beta(\Gamma(C)+\epsilon)}\Big) \leq e^{-e^{k_2\beta}}.$$

(iii) *For any $x, y \in C$, there exists $k_3 > 0$ such that for all β sufficiently large*

$$\mathbb{P}\Big(\tau_y^x > \tau_{\partial C}^x\Big) \leq e^{-k_3\beta}.$$

(iv) *There exists $k_4 > 0$ such that for all β sufficiently large*

$$\sup_{x\in C} \mathbb{P}\Big(X^\beta_{\tau_{\partial C}^x} \notin \mathcal{F}(\partial C)\Big) \leq e^{-k_4\beta}.$$

(v) *For any $x \in C$, $\epsilon > 0$ and $\epsilon' > 0$, for all β sufficiently large*

$$\mathbb{P}\Big(\tau_{\partial C}^x < e^{\beta(\Gamma(C)+\epsilon)},\, X^\beta_{\tau_{\partial C}^x} \in \mathcal{F}(\partial C)\Big) \geq e^{-\epsilon'\beta}.$$

(vi) *For any $x \in C$, any $\epsilon > 0$ and all β sufficiently large*

$$e^{\beta(\Gamma(C)-\epsilon)} < \mathbb{E}\tau_{\partial C}^x < e^{\beta(\Gamma(C)+\epsilon)}.$$

The first three properties can be found in [118, Theorem 6.23], the fourth one is [118, Corollary 6.25] and the fifth one in [106, Theorem 2.17]. The sixth property is given in [116, Proposition 3.9] and implies that

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\partial C}^x = \Gamma(C). \tag{4.14}$$

The third property states that, given that C is a cycle, for any starting state $x \in C$, the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ visits any state $y \in C$ before exiting from C with a probability exponentially close to one. This is a crucial property of the cycles and in fact can be given as alternative definition, see for instance [29, 30]. The equivalence of the two definitions has been proved in [38] in greater generality for a Markov chain satisfying the Freidlin-Wentzell condition (4.13). Leveraging this fact, many properties and results from [29] will be used or cited.

We denote by $\mathcal{C}(\mathcal{X})$ the set of cycles of $\mathcal{X}$. The next lemma, see [118, Proposition 6.8], implies that the set $\mathcal{C}(\mathcal{X})$ has a tree structure with respect to the inclusion relation, where $\mathcal{X}$ is the root and the singletons are the leaves.

Lemma 4.2.3 (Cycle tree structure). *Two cycles $C, C' \in \mathcal{C}(\mathcal{X})$ are either disjoint or comparable for the inclusion relation, i.e. $C \subseteq C'$ or $C' \subseteq C$.*

Lemma 4.2.3 also implies that the set of cycles to which a state $x \in \mathcal{X}$ belongs is totally ordered by inclusion. Furthermore, we remark that if two cycles $C, C' \in \mathcal{C}(\mathcal{X})$ are such that $C \subseteq C'$, then $\Gamma(C) \leq \Gamma(C')$; this latter inequality is strict if and only if the inclusion is strict.

4.2.2 *Classical bounds in probability for hitting time τ_A^x*

In this subsection we start investigating the first hitting time τ_A^x. Thus, we will tacitly assume that the *target set* A is a nonempty subset of $\mathcal{X}$ and the *initial state* x belongs to $\mathcal{X} \setminus A$. Moreover, without loss of generality, we will henceforth assume that

$$A = \{y \in \mathcal{X} \mid \forall \omega : x \to y \quad \omega \cap A \neq \emptyset\}, \tag{4.15}$$

which means that we add to the original target subset A all the states in $\mathcal{X}$ that cannot be reached from x without visiting the subset A. Note that this assumption does not change the distribution of the first hitting time τ_A^x, since the states which we may have added in this way could not have been visited without hitting the original subset A first.

Given a nonempty subset $A \subset \mathcal{X}$ and $x \in \mathcal{X}$, we define the *initial cycle* $C_A(x)$ by

$$C_A(x) := \{x\} \cup \{z \in \mathcal{X} : \Phi(x, z) < \Phi(x, A)\}. \tag{4.16}$$

If $x \in A$, then $C_A(x) = \{x\}$ and thus is a trivial cycle. If $x \notin A$, the subset $C_A(x)$ is either a trivial cycle (when $\Phi(x, A) = H(x)$) or a non-trivial cycle containing x, if $\Phi(x, A) > H(x)$. In any case, if $x \notin A$, then $C_A(x) \cap A = \emptyset$. For every $x \in \mathcal{X}$, we denote by $\Gamma(x, A)$ the depth of the initial cycle $C_A(x)$, i.e.

$$\Gamma(x, A) := \Gamma(C_A(x)). \tag{4.17}$$

Clearly if $C_A(x)$ is trivial (and in particular when $x \in A$), then $\Gamma(x, A) = 0$. Note that by definition the quantity $\Gamma(x, A)$ is always non-negative, and in general

$$\Gamma(x, A) = \Phi(x, A) - H(\mathcal{F}(C_A(x))) \geq \Phi(x, A) - H(x),$$

with equality if and only if $x \in \mathcal{F}(C_A(x))$.

If $x \notin A$, then the initial cycle $C_A(x)$ is, by construction, the maximal cycle (in the sense of inclusion) that contains the state x and has an empty intersection with A. Therefore, any path $\omega : x \to A$ has at some point to exit from $C_A(x)$, by overcoming an energy barrier not smaller than its depth $\Gamma(x, A)$. The next proposition gives a probabilistic bound for the hitting time τ_A^x by looking precisely at this *initial ascent* up until the boundary of $C_A(x)$.

Proposition 4.2.4 (Initial-ascent bound)**.** *Consider a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$. For any $\epsilon > 0$ there exists $\kappa > 0$ such that for β sufficiently large*

$$\mathbb{P}\left(\tau_A^x < e^{\beta(\Gamma(x,A)-\epsilon)}\right) < e^{-\kappa\beta}. \tag{4.18}$$

The proof is essentially adopted from [118] and follows easily from Theorem 4.2.2(i), since by definition of $C_A(x)$, we have that $\tau_A^x \geq_{\text{st}} \tau^x_{\partial C_A(x)}$.

Before stating an upper bound for the tail probability of the hitting time τ_A^x, we need some further definitions. Given a nonempty subset $B \subset \mathcal{X}$, we denote by $\mathcal{M}(B)$ the *collection of maximal cycles that partitions* B, i.e.

$$\mathcal{M}(B) := \{C \in \mathcal{C}(\mathcal{X}) \mid C \text{ maximal by inclusion and } C \subseteq B\}. \tag{4.19}$$

Lemma 4.2.3 implies that every nonempty subset $B \subset \mathcal{X}$ has a partition into maximal cycles and hence guarantees that $\mathcal{M}(B)$ is well defined. Note that if $C \in \mathcal{C}(\mathcal{X})$ is itself a cycle, then $\mathcal{M}(C) = \{C\}$. The following lemma shows that initial cycles can be used to obtain the partition in maximal cycles of any subset of the state space.

Lemma 4.2.5. [106, Lemma 2.26] *Given a nonempty subset $A \subset \mathcal{X}$, the collection $\{C_A(x)\}_{x\in\mathcal{X}\setminus A}$ of initial cycles is the partition into maximal cycles of $\mathcal{X} \setminus A$, i.e.*

$$\mathcal{M}(\mathcal{X} \setminus A) = \{C_A(x)\}_{x\in\mathcal{X}\setminus A}.$$

We can extend the notion of depth to subsets $B \subsetneq \mathcal{X}$ which are not necessarily cycles by using the partition of B into maximal cycles. More precisely, we define the *maximum depth* $\widetilde{\Gamma}(B)$ of a nonempty subset $B \subsetneq \mathcal{X}$ as the maximum depth of a cycle contained in B, i.e.

$$\widetilde{\Gamma}(B) := \max_{C\in\mathcal{M}(B)} \Gamma(C). \tag{4.20}$$

Trivially $\widetilde{\Gamma}(C) = \Gamma(C)$ if $C \in \mathcal{C}(\mathcal{X})$. The next lemma gives two equivalent characterizations of the maximum depth $\widetilde{\Gamma}(B)$ of a nonempty subset $B \subsetneq \mathcal{X}$.

Lemma 4.2.6 (Equivalent characterizations of the maximum depth)**.** *Given a nonempty subset $B \subsetneq \mathcal{X}$,*

$$\widetilde{\Gamma}(B) = \max_{x\in B} \Gamma(x, \mathcal{X} \setminus B) = \max_{x\in B} \Big\{ \min_{y\in\mathcal{X}\setminus B} \Phi(x,y) - H(x) \Big\}. \tag{4.21}$$

In view of Lemma 4.2.6, $\widetilde{\Gamma}(B)$ is the maximum initial energy barrier that the process started inside B possibly has to overcome to exit from B. As established by the next proposition, one can get a (super-)exponentially small upper bound for the tail probability of the hitting time τ_A^x, by looking at the maximum depth $\widetilde{\Gamma}(\mathcal{X} \setminus A)$ of the complementary set $\mathcal{X} \setminus A$, where the process resides before hitting the target subset A.

Proposition 4.2.7 (Deepest-cycle bound, [29, Proposition 4.19]). *Consider a nonempty subset $A \subsetneq \mathcal{X}$ and $x \notin A$. For any $\epsilon > 0$ there exists $\kappa' > 0$ such that for β sufficiently large*

$$\mathbb{P}\left(\tau_A^x > e^{\beta(\widetilde{\Gamma}(\mathcal{X}\setminus A)+\epsilon)}\right) < e^{-e^{\kappa'\beta}}. \tag{4.22}$$

By definition we have $\Gamma(x,A) \leq \widetilde{\Gamma}(\mathcal{X}\setminus A)$, but in general $\Gamma(x,A) \neq \widetilde{\Gamma}(\mathcal{X}\setminus A)$ and neither bound presented in this subsection is actually tight, so we will proceed to establish sharper but more involved bounds in the next subsection.

4.2.3 *Optimal paths and refined bounds in probability for hitting time τ_A^x*

The quantity $\Gamma(x,A)$ appearing in Proposition 4.2.4 only accounts for the energy barrier that has to be overcome starting from x, but there is such an energy barrier for every state $z \notin A$ and it may well be that to reach A it is inevitable to visit a state z with $\Gamma(z,A) > \Gamma(x,A)$. Similarly, also the exponent $\widetilde{\Gamma}(\mathcal{X}\setminus A)$ appearing in Proposition 4.2.7 may not be sharp in general. For instance, the maximum depth $\widetilde{\Gamma}(\mathcal{X}\setminus A)$ could be determined by a deep cycle C in $\mathcal{X}\setminus A$ that cannot be visited before hitting A or that is visited with a vanishing probability as $\beta\to\infty$. In this subsection, we refine the bounds given in Propositions 4.2.4 and 4.2.7 by using the notion of *optimal path* and identifying the subset of the state space $\mathcal{X}$ in which these optimal paths lie.

Given a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$, define the *set of optimal paths* $\Omega_{x,A}^{\mathrm{opt}}$ as the collection of all paths $\omega \in \Omega_{x,A}$ along which the maximum energy Φ_ω is equal to the communication height between x and A, i.e.

$$\Omega_{x,A}^{\mathrm{opt}} := \{\omega \in \Omega_{x,A} \mid \Phi_\omega = \Phi(x,A)\}. \tag{4.23}$$

Define the *relevant cycle* $C_A^+(x)$ as the minimal cycle in $\mathcal{C}(\mathcal{X})$ such that $C_A(x) \subsetneq C_A^+(x)$, i.e.

$$C_A^+(x) := \min\{C \in \mathcal{C}(\mathcal{X}) \mid C_A(x) \subsetneq C\}. \tag{4.24}$$

The cycle $C_A^+(x)$ is well defined, since the cycles in $\mathcal{C}(\mathcal{X})$ that contain x are totally ordered by inclusion, as remarked after Lemma 4.2.3. By construction, $C_A^+(x)\cap A \neq \emptyset$ and thus $C_A^+(x)$ contains at least two states, so it has to be a non-trivial cycle. The minimality of $C_A^+(x)$ with respect to the inclusion gives that

$$\max_{z\in C_A^+(x)} H(z) = \Phi(x,A),$$

and then, by using Lemma 4.2.1, one obtains

$$\Phi(x,A) < H(\mathcal{F}(\partial C_A^+(x))). \tag{4.25}$$

The choice of the name *relevant cycle* for $C_A^+(x)$ comes from the fact that all paths the Markov chain will follow to go from x to A will almost surely

not exit from $C_A^+(x)$ in the limit $\beta \to \infty$. Indeed, for the relevant cycle $C_A^+(x)$ Theorem 4.2.2(iii) reads

$$\lim_{\beta\to\infty} \mathbb{P}\left(\tau_A^x < \tau^x_{\partial C_A^+(x)}\right) = 1. \tag{4.26}$$

The next lemma, which is proved in Section 4.3, states that an optimal path from x to A is precisely a path from x to A that does not exit from $C_A^+(x)$.

Lemma 4.2.8 (Optimal path characterization)**.** *Consider a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$. Then*

$$\omega \in \Omega^{\mathrm{opt}}_{x,A} \iff \omega \in \Omega_{x,A} \quad \text{and} \quad \omega \subseteq C_A^+(x).$$

Lemma 4.2.8 implies that the relevant cycle $C_A^+(x)$ can be equivalently defined as

$$\begin{aligned} C_A^+(x) &= \Big\{ y \in \mathcal{X} \;:\; \Phi(x,y) \le \Phi(x,A) \Big\} \\ &= \Big\{ y \in \mathcal{X} \;:\; \Phi(x,y) < \Phi(x,A) + \delta_0/2 \Big\}, \end{aligned} \tag{4.27}$$

where δ_0 is the minimum energy gap between an optimal and a non-optimal path from x to A, i.e.

$$\delta_0 = \delta_0(x,A) := \min_{\omega\in\Omega_{x,A}\setminus\Omega^{\mathrm{opt}}_{x,A}} \Phi_\omega - \Phi(x,A).$$

In view of Lemma 4.2.8 and (4.26), the Markov chain started in x follows in the limit $\beta \to \infty$ almost surely an optimal path in $\Omega^{\mathrm{opt}}_{x,A}$ to hit A. It is then natural to define the following quantities for a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$:

$$\Psi_{\min}(x,A) := \min_{\omega\in\Omega^{\mathrm{opt}}_{x,A}} \max_{z\in\omega} \Gamma(z,A), \tag{4.28}$$

and

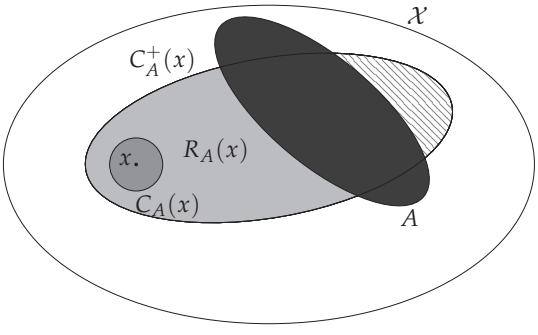
$$\Psi_{\max}(x,A) := \max_{\omega\in\Omega^{\mathrm{opt}}_{x,A}} \max_{z\in\omega} \Gamma(z,A). \tag{4.29}$$

Definition (4.28) implies that every optimal path $\omega \in \Omega^{\mathrm{opt}}_{x,A}$ has to enter at some point a cycle in $\mathcal{M}(\mathcal{X}\setminus A)$ of depth at least $\Psi_{\min}(x,A)$, while definition (4.29) means that every cycle visited by any optimal path $\omega \in \Omega^{\mathrm{opt}}_{x,A}$ has depth less than or equal to $\Psi_{\max}(x,A)$.

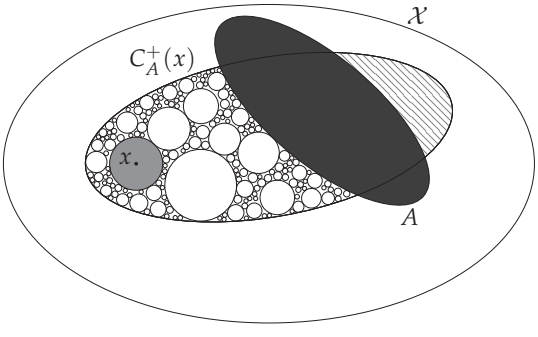
An equivalent characterization for the energy barrier $\Psi_{\max}(x,A)$ can be given, but we first need one further definition. Define $R_A(x)$ as the subset of states which belong to at least one optimal path in $\Omega^{\mathrm{opt}}_{x,A}$, i.e.

$$R_A(x) := \{ y \in \mathcal{X} \;:\; \exists\, \omega \in \Omega^{\mathrm{opt}}_{x,A} \;:\; y \in \omega \}. \tag{4.30}$$

Note that $A \cap R_A(x) \neq \emptyset$, since the endpoint of each path in $\Omega_{x,A}$ belongs to A, by definition (4.5). In view of Lemma 4.2.8, $R_A(x) \subseteq C_A^+(x)$. We remark that this latter inclusion could be strict, since in general $R_A(x) \neq C_A^+(x)$. Indeed, there could exist a state $y \in C_A^+(x)$ such that all paths $\omega : x \to y$ that do not exit from $C_A^+(x)$ always visit the target set A before reaching y, and thus they do not belong to $\Omega_{x,A}^{\text{opt}}$ (see definitions (4.5) and (4.23)), see Figure 4.1.



(a) The subset $R_A(x)$ (in light gray)



(b) The partition into maximal cycles of $R_A(x)$, including the initial cycle $C_A(x)$ (in dark gray)

Figure 4.1: Example of an energy landscape $\mathcal{X}$ with highlighted the subset A (in black), the relevant cycle $C_A^+(x)$ and the subset $C_A^+(x) \setminus (R_A(x) \cup A)$ (with diagonal mesh)

The next lemma characterizes the quantity $\Psi_{\max}(x, A)$ as the maximum depth of the subset $R_A(x) \setminus A$ (see definition (4.20)).

Lemma 4.2.9 (Equivalent characterization of $\Psi_{\max}(x, A)$)**.**

$$\Psi_{\max}(x, A) = \widetilde{\Gamma}(R_A(x) \setminus A). \tag{4.31}$$

Using the two quantities $\Psi_{\min}(x, A)$ and $\Psi_{\max}(x, A)$, we can obtain sharper bounds in probability for the hitting time τ_A^x, as stated in the next proposition, which is proved in Section 4.3.

Proposition 4.2.10 (Optimal paths depth bounds)**.** *Consider a nonempty subset $A \subset \mathcal{X}$ and $x \in \mathcal{X} \setminus A$. For any $\epsilon > 0$ there exists $\kappa > 0$ such that for β sufficiently large*

$$\mathbb{P}\left(\tau_A^x < e^{\beta(\Psi_{\min}(x,A)-\epsilon)}\right) < e^{-\kappa\beta}, \tag{4.32}$$

and

$$\mathbb{P}\left(\tau_A^x > e^{\beta(\Psi_{\max}(x,A)+\epsilon)}\right) < e^{-\kappa\beta}. \tag{4.33}$$

This proposition is in fact a sharper result than Propositions 4.2.4 and 4.2.7, since

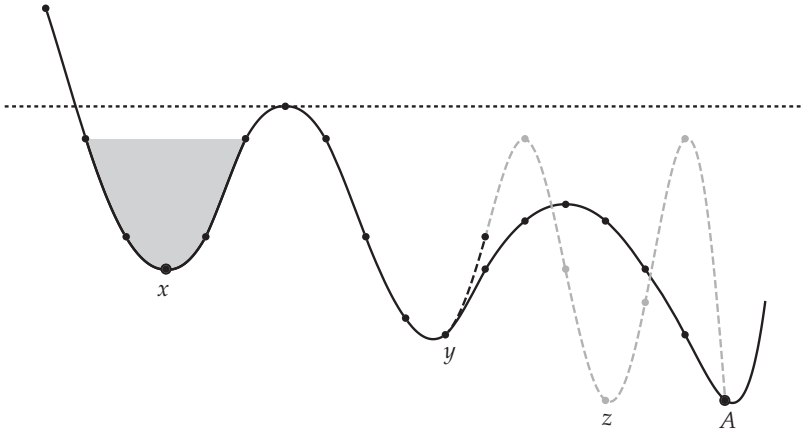
$$\Gamma(x,A) \le \Psi_{\min}(x,A) \le \Psi_{\max}(x,A) \le \widetilde{\Gamma}(\mathcal{X} \setminus A). \tag{4.34}$$

Indeed, since the starting state x trivially belongs to every path in $\Omega_{x,A}^{\mathrm{opt}}$, we have that $\Gamma(x,A) \le \max_{z\in\omega} \Gamma(z,A)$ for every $\omega \in \Omega_{x,A}^{\mathrm{opt}}$ and thus $\Gamma(x,A) \le \Psi_{\min}(x,A)$. Moreover, since by definition $C_A^+(x) \setminus A \subseteq \mathcal{X} \setminus A$, Lemma 4.2.9 yields that $\Psi_{\max}(x,A) \le \widetilde{\Gamma}(\mathcal{X} \setminus A)$.

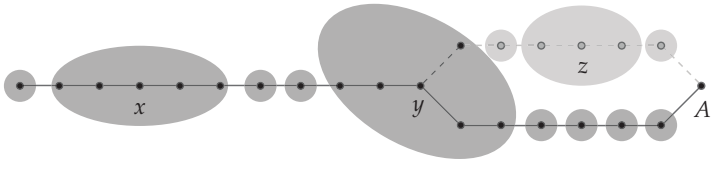
If $\Gamma(x,A) = \widetilde{\Gamma}(\mathcal{X} \setminus A)$, it follows from (4.34) that $\Psi_{\min}(x,A) = \Psi_{\max}(x,A)$. However, in general, the exponents $\Psi_{\min}(x,A)$ and $\Psi_{\max}(x,A)$ are not equal and may not be sharp either, as illustrated by the energy landscape in Figure 4.2.

In this example, there are two paths to go from x to A: The path ω which goes from x to y and then follows the solid path until A, and the path ω', which goes from x to y and then follows the dashed path through z and eventually hitting A. Note that $\Phi_\omega = \Phi_{\omega'} = \Phi(x,A)$, so both ω and ω' are optimal paths from x to A. By inspection, we get that $\Psi_{\max}(x,A) = \Gamma(z,A)$. However, the path ω' does not exit the cycle $C_A(y)$ passing by its principal boundary and, in view of Theorem 4.2.2(iv), it becomes less likely than the other path as $\beta \to \infty$. In fact, the transition from x to A is likely to occur on a smaller time-scale than suggested by the upper bounds in Proposition 4.2.10 and in particular the exponent $\Psi_{\max}(x,A)$ is not sharp in this example.

In the next subsection, we will show that a more precise control in probability of the hitting time τ_A^x is possible, at the expense of a more involved analysis of the energy landscape.



(a) Energy profile of the energy landscape with the initial cycle $C_A(x)$ (in grey) and the relevant cycle $C_A^+(x)$ (below the dashed black line)



(b) Partition into maximal cycles of $\mathcal{X} \setminus A$ for the same energy landscape

Figure 4.2: An example energy landscape for which $\Psi_{\max}(x, A)$ is not sharp

4.2.4 *Sharp bounds for hitting time τ_A^x using typical paths*

As illustrated at the end of the previous subsection, the exponents $\Psi_{\min}(x, A)$ and $\Psi_{\max}(x, A)$ appearing in the probability bounds (4.32) and (4.33) for the hitting time τ_A^x may not be sharp in general. In this subsection we obtain exponents that are potentially sharper than $\Psi_{\min}(x, A)$ and $\Psi_{\max}(x, A)$ by looking in more detail at the cycle decomposition of $C_A^+(x) \setminus A$ and by identifying inside it the *tube of typical paths from x to A*. In particular, we focus on how the process moves from two maximal cycles in the partition of $C_A^+(x) \setminus A$ and determine which of these transitions between maximal cycles are the most likely ones.

Some further definitions are needed. We introduce the notion of *cycle-path* and a way of mapping every path $\omega \in \Omega_{x,A}$ into a cycle-path $\mathcal{C}_\omega$. Recall that for a nonempty subset $A \subset \mathcal{X}$, ∂A is its external boundary and $\mathcal{F}(A)$ is its bottom, i.e. the set of the minima of the energy function H in A. A *cycle-path* is a finite sequence $(C_1, \dots, C_m)$ of (trivial and non-trivial) cycles $C_1, \dots, C_m \in \mathcal{C}(\mathcal{X})$ such that

$$C_i \cap C_{i+1} = \emptyset \quad \text{and} \quad \partial C_i \cap C_{i+1} \neq \emptyset, \quad \text{for every } i = 1, \dots, m-1.$$

It can be easily proved that, in a cycle-path $(C_1, \dots, C_m)$, if C_i is a non-trivial cycle for some $i = 1, \dots, m$, then its predecessor C_{i-1} and successor C_{i+1} (if any) are trivial cycles, see [37, Lemma 2.5]. We can consider the collection $\mathcal{P}_{x,A}$ of cycle-paths that lead from x to A and consist of maximal cycles in $\mathcal{X} \setminus A$ only, namely

$$\begin{aligned}\mathcal{P}_{x,A} := \{&\text{cycle-path } (C_1, \dots, C_m) \mid \\ &C_1, \dots, C_m \in \mathcal{M}(\mathcal{X} \setminus A),\, x \in C_1,\, \partial C_m \cap A \neq \emptyset\}.\end{aligned} \tag{4.35}$$

Recall that the collection of cycles $\mathcal{M}(\mathcal{X} \setminus A)$ can be constructed using initial cycles, as established by Lemma 4.2.5.
We define a mapping $\Omega_{x,A} \to \mathcal{P}_{x,A}$ by assigning to a path $\omega = (\omega_1, \dots, \omega_n) \in \Omega_{x,A}$ the cycle-path $\mathcal{C}_\omega = (C_1, \dots, C_{m(\omega)}) \in \mathcal{P}_{x,A}$ as follows. Set $t_0 = 1$, $C_1 = C_A(x)$, and then define recursively

$$t_i := \min\{k > t_{i-1} \mid \omega_k \notin C_i\} \quad \text{and} \quad C_{i+1} := C_A(\omega_{t_i}).$$

The path ω is a finite sequence and $\omega_n \in A$, so there exists an index $m(\omega) \in \mathbb{N}$ such that $\omega_{t_{m(\omega)}} = \omega_n \in A$ and there the procedure stops. The way the sequence $(C_1, \dots, C_{m(\omega)})$ is constructed shows that it is indeed a cycle-path. Moreover, by using the notion of initial cycle $C_A(\cdot)$ to define $C_1, \dots, C_{m(\omega)}$, they are automatically maximal cycles in $\mathcal{M}(\mathcal{X} \setminus A)$. Lastly, the fact that $\omega \in \Omega_{x,A}$ implies that $x \in C_1$ and that $\partial C_{m(\omega)} \cap A \neq \emptyset$, hence $\mathcal{C}_\omega \in \mathcal{P}_{x,A}$ and the mapping is well-defined. We remark that this mapping is not injective, since two different paths in $\Omega_{x,A}$ can be mapped into the same cycle-path in $\mathcal{P}_{x,A}$. In fact, a single cycle-path groups together all the paths that visit the same cycles (the same number of times and in the same order). Cycle-paths are the appropriate mesoscopic objects to investigate while studying the transition $x \to A$: Indeed one neglects in this way the microscopic dynamics of the process and focuses only on the relevant mesoscopic transitions from one maximal cycle to another. Furthermore, we note that for a given path $\omega \in \Omega_{x,A}$, the maximum energy barrier along ω is the maximum depth in its corresponding cycle-path $\mathcal{C}_\omega$, i.e.

$$\max_{z \in \omega} \Gamma(z, A) = \max_{C \in \mathcal{C}_\omega} \Gamma(C).$$

For every cycle $C \in \mathcal{C}(\mathcal{X})$ define

$$\mathcal{B}(C) := \begin{cases} \mathcal{F}(\partial C) & \text{if } C \text{ is a non-trivial cycle,} \\ \{z \in \partial C \mid H(z) \leq H(y)\} & \text{if } C = \{y\} \text{ is a trivial cycle.} \end{cases} \tag{4.36}$$

In other words, if C is a non-trivial cycle, then $\mathcal{B}(C)$ is its principal boundary, while when $C = \{y\}$ is a trivial cycle $\mathcal{B}(C)$ is the subset of states connected to y with energy lower than y. We will refer to $\mathcal{B}(C)$ as *principal boundary* of C also in the case where C is a trivial cycle.

We say that a cycle-path $(C_1, \dots, C_m)$ is *connected via typical jumps to* A or simply vtj-*connected* to A if

$$\mathcal{B}(C_i) \cap C_{i+1} \neq \emptyset, \quad \forall\, i = 1, \dots, m-1, \quad \text{and} \quad \mathcal{B}(C_m) \cap A \neq \emptyset. \tag{4.37}$$

The next lemma, presented in [38], guarantees that there always exists a cycle-path from the initial cycle $C_A(x)$ that is vtj-connected to A for any nonempty target subset $A \subset \mathcal{X}$ and $x \notin A$.

Lemma 4.2.11. [38, Proposition 3.22] *For any nonempty subset $A \subset \mathcal{X}$ and $x \notin A$, there exists a cycle-path $\mathcal{C}^* = (C_1, \dots, C_{m^*})$ vtj-connected to A with $x \in C_1$ and $C_1, \dots, C_m^* \subset \mathcal{X} \setminus A$.*

By inspecting the proof of [38, Proposition 3.22], one notices that the given cycle-path $\mathcal{C}^* = (C_1, \dots, C_{m^*})$ consists only of cycles in $\mathcal{X} \setminus A$ which are maximal by inclusion, so that $C_1, \dots, C_{m^*} \in \mathcal{M}(\mathcal{X} \setminus A)$, and, in particular, $C_1 = C_A(x)$. Hence $\mathcal{C}^* \in \mathcal{P}_{x,A}$ and therefore the collection $\mathcal{P}_{x,A}$ is not empty.

We define $\omega \in \Omega_{x,A}$ to be a *typical path from x to A* if its corresponding cycle-path $\mathcal{C}_\omega$ is vtj-connected to A, and we denote by $\Omega_{x,A}^{\text{vtj}}$ the collection of all typical paths from x to A, i.e.

$$\Omega_{x,A}^{\text{vtj}} := \{\omega \in \Omega_{x,A} \mid \omega \text{ is typical}\}.$$

The existence of a vtj-connected cycle-path $\mathcal{C}^* = (C_1, \dots, C_{m^*}) \in \mathcal{P}_{x,A}$ guarantees that $\Omega_{x,A}^{\text{vtj}} \neq \emptyset$. Indeed, take $y_0 = x$, $y_i \in \mathcal{B}(C_i) \cap C_{i+1}$, $i = 1, \dots, m^* - 1$ and $y_{m^*} \in \mathcal{B}(C_{m^*}) \cap A$ and consider a path ω^* that visits precisely the saddles $y_0, \dots, y_{m^*}$ in this order and stays in cycle C_i between the visit to y_{i-1} and y_i. Then ω^* is a typical path from x to A.

The following lemma gives an equivalent characterization of a typical path from x to A.

Lemma 4.2.12 (Equivalent characterization of a typical path)**.** *Take a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$. Then*

$$\omega \in \Omega_{x,A}^{\text{vtj}} \iff \omega \in \Omega_{x,A} \quad \text{and} \quad \Phi(\omega_{i+1}, A) \leq \Phi(\omega_i, A) \quad \forall\, i = 1, \dots, |\omega| - 1.$$

The proof of this result is presented in Section 4.3. Lemma 4.2.12 shows that every typical path from x to A is an optimal path from x to A, i.e.

$$\Omega_{x,A}^{\text{vtj}} \subseteq \Omega_{x,A}^{\text{opt}}, \tag{4.38}$$

since if $\omega \in \Omega_{x,A}^{\text{vtj}}$, then $\Phi(\omega_i, A) \leq \Phi(\omega_1, A) = \Phi(x, A)$ for every $i = 2, \dots, |\omega|$ and thus $\Phi_\omega = \Phi(x, A)$.

Let $\mathrm{T}_A(x)$ be the *tube of typical paths from x to A*, which is defined as

$$\mathrm{T}_A(x) := \{y \in \mathcal{X} \mid \exists\, \omega \in \Omega_{x,A}^{\text{vtj}} \,:\, y \in \omega\}. \tag{4.39}$$

In other words, $\mathrm{T}_A(x)$ is the subset of states $y \in \mathcal{X}$ that can be reached from x by means of a typical path which does not enter A before visiting y. The

endpoint of every path in $\Omega^{\text{vtj}}_{x,A}$ belongs to A, thus $\mathrm{T}_A(x) \cap A \neq \emptyset$. Since by (4.38) every typical path is an optimal path, it follows from definitions (4.30) and (4.39) that

$$\mathrm{T}_A(x) \subseteq R_A(x).$$

The tube of typical paths can be visualized as the *standard cascade* emerging from state x and reaching eventually A, in the sense that it is the part of the energy landscape that would be wet if a water source is placed at x and the water would "find its way" until the sink, that is subset A. This standard cascade possibly consists of basins/lakes (non-trivial cycles), saddle points (trivial cycles) and waterfalls (trivial cycles). From definition (4.39), it follows that if $z \in \mathrm{T}_A(x)$, then

$$\mathrm{T}_A(z) \subseteq \mathrm{T}_A(x). \tag{4.40}$$

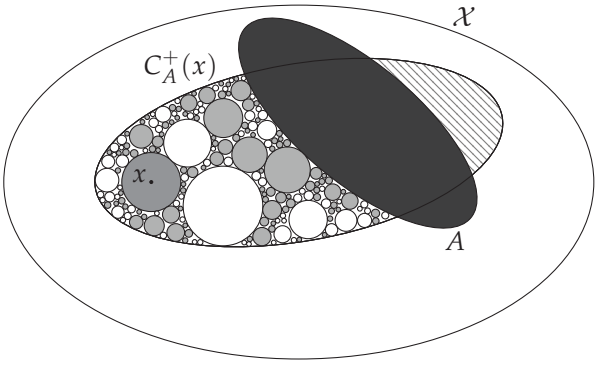


Figure 4.3: Example of an energy landscape with the tube of typical $\mathrm{T}_A(x)$ highlighted in gray

Denote by $\mathfrak{T}_A(x)$ the collection of cycles $C \in \mathcal{M}(\mathcal{X} \setminus A)$ for which there exists a vtj-connected cycle-path $C_1, \ldots, C_n \subset \mathcal{X} \setminus A$ with $C_1 = C_A(x)$ and $C_n = C$, i.e.

$$\mathfrak{T}_A(x) := \{C \in \mathcal{M}(\mathcal{X} \setminus A) \mid \exists\, C_1, \ldots, C_n \text{ vtj-connected cycle-path with } C_1 = C_A(x) \text{ and } C_n = C\}.$$

Note that the cycles in $\mathfrak{T}_A(x)$ form the partition into maximal cycles of $\mathrm{T}_A(x) \setminus A$, i.e.

$$\mathfrak{T}_A(x) = \mathcal{M}(\mathrm{T}_A(x) \setminus A),$$

and that, by construction, there exists $C \in \mathfrak{T}_A(x)$ such that $\mathcal{B}(C) \cap A \neq \emptyset$. The boundary of $\mathrm{T}_A(x)$ consists of states either in A or in the non-principal part of the boundary of a cycle in $C \in \mathfrak{T}_A(x)$:

$$\partial \mathrm{T}_A(x) \setminus A = \bigcup_{C \in \mathfrak{T}_A(x)} (\partial C \setminus \mathcal{B}(C)). \tag{4.41}$$

The typical paths in $\Omega^{\text{vtj}}_{x,A}$ are the only ones with non-vanishing probability of being visited by the Markov chain $\{X^\beta_t\}_{t\in\mathbb{N}}$ started in x before hitting A in the limit $\beta\to\infty$, as established by the next lemma which is proved in Section 4.3.

Lemma 4.2.13 (Exit from the typical tube $\mathrm{T}_A(x)$)**.** *Consider a nonempty subset $A\subset\mathcal{X}$ and $x\notin A$. Then there exists $\kappa>0$ such that for β sufficiently large*

$$\mathbb{P}\left(\tau^x_{\partial \mathrm{T}_A(x)}\le\tau^x_A\right)\le e^{-\kappa\beta}.$$

Given a nonempty subset $A\subset\mathcal{X}$ and $x\notin A$, define the following quantities:

$$\Theta_{\min}(x,A):=\min_{\omega\in\Omega^{\text{vtj}}_{x,A}}\max_{z\in\omega}\Gamma(z,A), \tag{4.42}$$

and

$$\Theta_{\max}(x,A):=\max_{\omega\in\Omega^{\text{vtj}}_{x,A}}\max_{z\in\omega}\Gamma(z,A). \tag{4.43}$$

In other words, definition (4.42) means that every typical path $\omega\in\Omega^{\text{vtj}}_{x,A}$ has to enter at some point a cycle of depth at least $\Theta_{\min}(x,A)$. On the other hand, definition (4.29) implies that all cycles visited by any typical path $\omega\in\Omega^{\text{vtj}}_{x,A}$ have depth less than or equal to $\Theta_{\max}(x,A)$. Hence, $\Theta_{\max}(x,A)$ can equivalently be characterized as the maximum depth (see definition (4.20)) of the tube $\mathrm{T}_A(x)$ of typical paths from x to A, as stated by the next lemma.

Lemma 4.2.14 (Equivalent characterization of $\Theta_{\max}(x,A)$)**.**

$$\Theta_{\max}(x,A)=\widetilde{\Gamma}(\mathrm{T}_A(x)\setminus A)=\max_{C\in\mathfrak{T}_A(x)}\Gamma(C). \tag{4.44}$$

Since by (4.38) every typical path from x to A is an optimal path from x to A, definitions (4.28), (4.29), (4.42) and (4.43) imply that

$$\Psi_{\min}(x,A)\le\Theta_{\min}(x,A)\le\Theta_{\max}(x,A)\le\Psi_{\max}(x,A). \tag{4.45}$$

We now have all the ingredients needed to formulate the first refined result for the hitting time τ^x_A, which is proved in Section 4.3. The main idea behind the next proposition is to look at the *shallowest-typical gorge* inside $\mathrm{T}_A(x)$ that the process has to overcome to reach A and at the *deepest-typical gorge* inside $\mathrm{T}_A(x)$ where the process has a non-vanishing probability to be trapped before hitting A.

Proposition 4.2.15 (Typical-cycles bounds)**.** *Consider a nonempty subset $A\subset\mathcal{X}$ and $x\notin A$. For any $\epsilon>0$ there exists $\kappa>0$ such that for β sufficiently large*

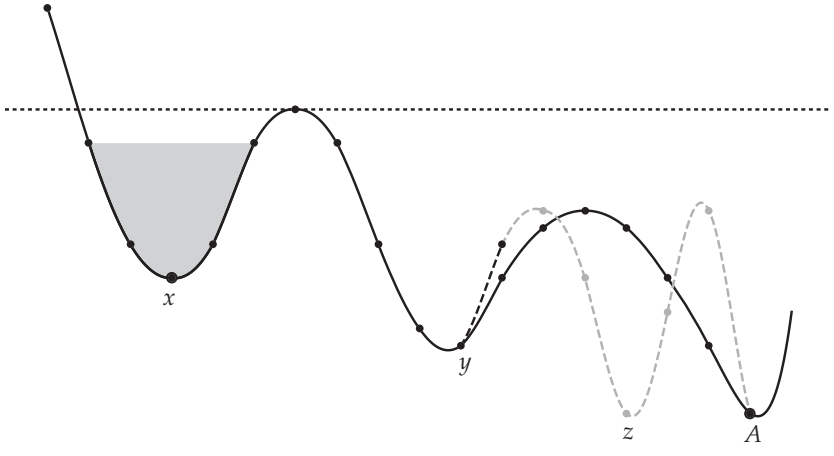
$$\mathbb{P}\left(\tau^x_A<e^{\beta(\Theta_{\min}(x,A)-\epsilon)}\right)<e^{-\kappa\beta}, \tag{4.46}$$

and

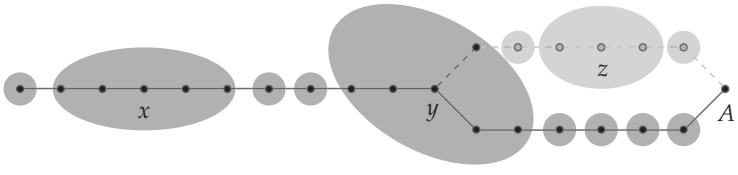
$$\mathbb{P}\left(\tau^x_A>e^{\beta(\Theta_{\max}(x,A)+\epsilon)}\right)<e^{-\kappa\beta}. \tag{4.47}$$

The proof, which is a refinement of that of Proposition 4.2.10, is presented in Section 4.3.

In general, the exponents $\Theta_{\min}(x, A)$ and $\Theta_{\max}(x, A)$ may not be equal, as illustrated by the energy landscape in Figure 4.4.



(a) Energy profile of the energy landscape with the initial cycle $C_A(x)$ (in grey) and the relevant cycle $C_A^+(x)$ (below the dashed black line)



(b) Partition into maximal cycles of $\mathcal{X} \setminus A$ for the same energy landscape

Figure 4.4: An example energy landscape for which $\Theta_{\min}(x, A) < \Theta_{\max}(x, A)$

Also in this example, there are two paths to go from x to A: The path ω which goes from x to y and then follows the solid path until A, and the path ω', which goes from x to y and then follows the dashed path through z and eventually hitting A. Both paths ω and ω' always move from a cycle to the next one visiting the principal boundary, hence they are both typical paths from x to A. By inspection, we get that $\Theta_{\max}(x, A) = \Gamma(z, A)$, since the typical path ω' visits the cycle $C_A(z)$. Using the path ω we deduce that $\Theta_{\min}(x, A) = \Gamma(y, A)$ and therefore $\Theta_{\min}(x, A) < \Theta_{\max}(x, A)$.

If the two exponents $\Theta_{\min}(x, A)$ and $\Theta_{\max}(x, A)$ coincide, then, in view of Proposition 4.2.15, we get sharp bounds in probability on a logarithmic scale for the hitting time τ_A^x, as stated in the next corollary.

Corollary 4.2.16. *Consider a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$. Assume that*

$$\Theta_{\min}(x, A) = \Theta(x, A) = \Theta_{\max}(x, A). \tag{4.48}$$

Then, for any $\epsilon > 0$

$$\lim_{\beta\to\infty} \mathbb{P}\Big(e^{\beta(\Theta(x,A)-\epsilon)} < \tau_A^x < e^{\beta(\Theta(x,A)+\epsilon)}\Big) = 1. \tag{4.49}$$

There are many examples of models and pairs (x, A) for which $\Theta_{\min}(x, A) = \Theta_{\max}(x, A)$. The most classical ones are the models that exhibit a *metastable behavior*: If one takes $x \in \mathcal{X}^m$ and $A = \mathcal{X}^s$, then it follows that $\Theta_{\min}(x, A) = \mathcal{V}_x = \Theta_{\max}(x, A)$ (by virtue of the definition (4.3) of stability level) and Corollary 4.2.16 holds, see also [106, Theorem 4.1].

4.2.5 *First moment convergence*

We now turn our attention to the asymptotic behavior of the mean hitting time $\mathbb{E}\tau_A^x$ as $\beta \to \infty$. In particular, we will show that it scales (almost) exponentially in β and we will identify the corresponding exponent. There may be some subexponential pre-factors, but, without further assumptions, one can only hope to get results on a logarithmic scale, due to the potential complexity of the energy landscape. We remark that a precise knowledge of the tube of typical paths is not always necessary to derive the asymptotic order of magnitude of the mean hitting time $\mathbb{E}\tau_A^x$, as established by Proposition 4.2.18.

To prove the convergence of the quantity $\frac{1}{\beta}\log \mathbb{E}\tau_A^x$, we need the following assumption.

Assumption A (Absence of deep typical cycles) *Given the energy landscape* (X, H, c)*, we assume*

(A1) $\Theta_{\min}(x, A) = \Theta(x, A) = \Theta_{\max}(x, A)$*, and*

(A2) $\Theta_{\max}(z, A) \le \Theta(x, A)$ *for every* $z \in \mathcal{X} \setminus A$.

Condition (A1) says that every path $\omega : x \to A$ visits one of the deepest typical cycles of the tube $\mathrm{T}_A(x)$. Condition (A2) guarantees that by starting in another state $z \neq x$, the deepest typical cycle the process can enter is not deeper than those in $\mathrm{T}_A(x)$. Checking the validity of Assumption A can be very difficult in general, but we give a sufficient condition in Proposition 4.2.18 which is satisfied in many models of interest, including the hard-core model on grid graphs and triangular grid graphs (as illustrated in Chapter 5 and Chapter 6, respectively) and the Widom-Rowlison model on grid graphs (see Chapter 7. We further remark that (A1) is precisely the assumption of Corollary 4.2.16. Therefore, in the scenarios where Assumption A holds, we also have the asymptotic result (4.49) in probability for the hitting time τ_A^x.

The next theorem says that if Assumption A is satisfied, then the asymptotic order of magnitude of the mean hitting time $\mathbb{E}\tau_A^x$ as $\beta \to \infty$ is $\Theta(x, A)$.

Theorem 4.2.17 (First moment convergence)**.** *If Assumption A is satisfied, then*

$$\lim_{\beta\to\infty} \frac{1}{\beta}\log \mathbb{E}\tau_A^x = \Theta(x, A).$$

In many models of interest, calculating $\widetilde{\Gamma}(\mathcal{X} \setminus A)$ is easier than calculating $\Theta_{\min}(x, A)$ or $\Theta_{\max}(x, A)$. Indeed, even if $\widetilde{\Gamma}(\mathcal{X} \setminus A)$ is a quantity that still requires a global analysis of the energy landscape, one needs to compute just the communication height $\Phi(z, A)$ between any state $z \in \mathcal{X} \setminus A$ and the target set A, without requiring a full understanding of the complex cycle structure of the energy landscape. Besides this fact, the main motivation to look at the quantity $\widetilde{\Gamma}(\mathcal{X} \setminus A)$ is that it allows to give a sufficient condition for Assumption A, as established in the following proposition.

Proposition 4.2.18 (Absence of deep cycles)**.** *If*

$$\Phi(x, A) - H(x) = \widetilde{\Gamma}(\mathcal{X} \setminus A), \tag{4.50}$$

then Assumption A holds.

Proof. From the inequality

$$\Phi(x, A) - H(x) \leq \Theta_{\min}(x, A) \leq \Theta_{\max}(x, A) \leq \widetilde{\Gamma}(\mathcal{X} \setminus A),$$

we deduce that $\Theta_{\min}(x, A) = \Theta_{\max}(x, A)$ and (A1) is proved. Moreover, by definition of $\widetilde{\Gamma}(\mathcal{X} \setminus A)$, we have $\Theta_{\max}(z, A) \leq \widetilde{\Gamma}(\mathcal{X} \setminus A)$ for every $z \in \mathcal{X} \setminus A$. This inequality, together with the fact that $\Theta_{\max}(x, A) = \widetilde{\Gamma}(\mathcal{X} \setminus A)$, proves that (A2) also holds and thus assumption A is satisfied. □

We now present two interesting scenarios for which (4.50) holds.

Example 1 (Metastability scenario)

Suppose that

$$x \in \mathcal{X}^m \quad \text{and} \quad A = \mathcal{X}^s.$$

In this first scenario, τ_A^x is the classical transition time between a metastable state and a stable state, a widely studied object in the statistical mechanics literature (see, e.g. [106]). Assumption A is satisfied in this case by applying Proposition 4.2.18, since condition (4.50) holds: The equality $\Phi(x, \mathcal{X}^s) - H(x) = \widetilde{\Gamma}(\mathcal{X} \setminus \mathcal{X}^s)$ follows from the assumption $x \in \mathcal{X}^m$, which means that there are no cycles in $\mathcal{X} \setminus \mathcal{X}^s$ that are deeper than $C_{\mathcal{X}^s}(x)$.

Example 2 (Tunneling scenario)

Suppose that $x \in \mathcal{X}^s$, $A = \mathcal{X}^s \setminus \{x\}$ and

$$\Phi(z, A) - H(z) \leq \Phi(x, A) - H(x) \qquad \forall z \in \mathcal{X} \setminus \{x\}. \tag{4.51}$$

In the second scenario, the hitting time τ_A^x is the *tunneling time* between any pair of stable states. Assumption (4.51) says that every cycle in the energy landscape which does not contain a stable state has depth strictly smaller than the cycle $C_A(x)$ and we generally refer to this property as *absence of deep*

cycles in the tunneling scenario. This condition immediately implies that (4.50) holds, i.e. $\widetilde{\Gamma}(\mathcal{X} \setminus A) = \Phi(x, A) - H(x)$, and hence in this scenario assumption A holds, thanks to Proposition 4.2.18.

We will show that the tunneling times of the hard-core model on grid graphs and on triangular grid graphs fall precisely in this second scenario, as we will show in Chapter 5 and Chapter 6, respectively.

4.2.6 *Asymptotic exponentiality*

We now present a sufficient condition for the scaled random variable $\tau_A^x / \mathbb{E}\tau_A^x$ to converge in distribution to an exponential unit-mean random variable as $\beta \to \infty$. Define

$$\Theta_*(x, A) := \lim_{\beta \to \infty} \frac{1}{\beta} \log \mathbb{E}\tau_A^x. \tag{4.52}$$

If Assumption A holds, then we know that $\Theta(x, A) = \Theta_*(x, A)$, but the result presented in this section does not require the exact knowledge of $\Theta_*(x, A)$. We prove asymptotic exponentiality of the scaled hitting time under the following assumption.

Assumption B ("Worst initial state") *Given an energy landscape (X, H, c), we assume that*

$$\Theta_*(x, A) > \widetilde{\Gamma}(\mathcal{X} \setminus (A \cup \{x\})). \tag{4.53}$$

This assumption guarantees that the following "recurrence" result holds: From any state $z \in \mathcal{X}$ the Markov chain reaches the set $A \cup \{x\}$ on a time scale strictly smaller than that at which the transition $x \to A$ occurs. Indeed, Proposition 4.2.7 gives that for any $\epsilon > 0$

$$\lim_{\beta \to \infty} \sup_{z \in \mathcal{X}} \mathbb{P}\left(\tau_{\{x\} \cup A}^z > e^{\beta(\widetilde{\Gamma}(\mathcal{X} \setminus (A \cup \{x\})) + \epsilon)}\right) = 0.$$

We can informally say that Assumption B requires x to be the "worst initial state" for the Markov chain when the target subset is A.

Proposition 4.2.20 gives a sufficient condition for Assumption B to hold, which is satisfied in many models of interest, in particular the hard-core model on grid graphs (see Chapter 5), the hard-core model on triangular grid graphs (see Chapter 6) and the Widom-Rowlison model on grid graphs (see Chapter 7.

Theorem 4.2.19 (Asymptotic exponentiality)**.** *If Assumption B is satisfied, then*

$$\frac{\tau_A^x}{\mathbb{E}\tau_A^x} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \beta \to \infty. \tag{4.54}$$

More precisely, there exist two functions $k_1(\beta)$ *and* $k_2(\beta)$ *with* $\lim_{\beta\to\infty} k_1(\beta) = 0$ *and* $\lim_{\beta\to\infty} k_2(\beta) = 0$ *such that for any* $s > 0$

$$\left|\mathbb{P}\Big(\frac{\tau_A^x}{\mathbb{E}\tau_A^x} > s\Big) - e^{-s}\right| \leq k_1(\beta)e^{-(1-k_2(\beta))s}.$$

The proof, presented in Section 4.3, readily follows from the consequences of Assumption B discussed above and by applying [49, Theorem 2.3],

We now present a condition which guarantees that Assumption B holds and show that it holds in two scenarios similar to those described in the previous subsection.

Proposition 4.2.20 ($C_A(x)$ is "the unique deepest cycle")**.** *If*

$$\Gamma(x, A) > \widetilde{\Gamma}(\mathcal{X} \setminus (A \cup \{x\})), \tag{4.55}$$

then Assumption B is satisfied.

The proof of this proposition is immediate from (4.34) and (4.45). We remark that if condition (4.55) holds, then the initial cycle $C_A(x)$ is the unique deepest cycle in $\mathcal{X} \setminus A$. Condition (4.55) is stronger than (4.53), but often easier to check, since one does not need to compute the exact value of $\Theta_*(x, A)$, but only the depth $\Gamma(x, A)$ of the initial cycle $C_A(x)$. We now present two scenarios of interest.

Example 3 (Unique metastable state scenario)

Suppose that

$$\mathcal{X}^m = \{z\}, \quad A = \mathcal{X}^s, \quad \text{and} \quad x \in C_A(z).$$

We remark that this scenario is a special case of the metastable scenario presented in Example 1 in Subsection 4.2.5. This scenario was already mentioned in [106], in the discussion following Theorem 4.15, but we briefly discuss here how to prove asymptotic exponentiality within our framework. Indeed, we have that

$$\Gamma(x, \mathcal{X}^s) = \Gamma(C_{\mathcal{X}^s}(z)) = \widetilde{\Gamma}(\mathcal{X} \setminus \mathcal{X}^s),$$

thanks to the fact that z is the configuration in $\mathcal{X} \setminus \mathcal{X}^s$ with the maximum stability level, which means that $C_{\mathcal{X}^s}(z)$ is the deepest cycle in $\mathcal{X} \setminus \mathcal{X}^s$. Moreover, the fact that z is the unique metastable state, implies that

$$\widetilde{\Gamma}(\mathcal{X} \setminus \mathcal{X}^s) > \widetilde{\Gamma}(\mathcal{X} \setminus (\mathcal{X}^s \cup \{z\})),$$

since every configuration in $\mathcal{X} \setminus (\mathcal{X}^s \cup \{z\})$ has stability level strictly smaller than $\mathcal{V}_z$.

Example 4 (Two stable states scenario)

Suppose that $\mathcal{X}^s = \{s_1, s_2\}$, $A = \{s_2\}$, $x \in C_A(s_1)$ and

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{s_1, s_2\}) < \Phi(s_1, s_2) - H(s_1).$$

This scenario is a special case of the tunneling scenario presented in Example 2 in Subsection 4.2.5. In this case condition (4.55) is obviously satisfied. In particular, it shows that the scaled tunneling time $\tau_{s_2}^{s_1}$ between two stable states in $\mathcal{X}$ is asymptotically exponential whenever $\mathcal{X}^s = \{s_1, s_2\}$ and the condition $\widetilde{\Gamma}(\mathcal{X} \setminus \{s_1, s_2\}) < \Phi(s_1, s_2) - H(s_1)$ is satisfied.

4.2.7 *An example of non-exponentiality*

Assumption B is a rather strong assumption. In fact, for many models and for most of choices of x and A, the scaled hitting time $\tau_A^x / \mathbb{E}\tau_A^x$ does not have an exponential distribution in the limit $\beta \to \infty$. Moreover, we do not claim that Assumption B is necessary to have asymptotically exponentiality of the scaled hitting time $\tau_A^x / \mathbb{E}\tau_A^x$. However, we will now show that for the hard-core model on complete K-partite graphs considered in Chapter 3 Assumption B does not always hold and that the model exhibits non-exponentially distributed scaled hitting times.

Let the conflict graph G be a complete K-partite graph as described in Section 3.1. Recall that this means that the sites in V can be partitioned into K disjoint subsets $V_1, \ldots, V_K$ called *components*, such that two sites are connected by an edge if and only if they belong to different components. An example of complete 5-partite graph is displayed in Figure 3.1 and the structure of the corresponding energy landscape $\mathcal{X}$ is illustrated by Figure 3.2a.

This choice for G results in a simple state space $\mathcal{X}$, for which a detailed analysis is possible. Moreover, for the same model the asymptotic behavior of the first hitting times in continuous time is already well understood, see Chapter 3. Before stating the results, we recall some definitions from Section 3.1 and introduce some new notation. Let L_k be the size of the k-th component V_k, for $k = 1, \ldots, K$. Clearly the total number of sites in V is $N = \sum_{k=1}^{K} L_k$. Define $L_{\max} := \max_{k=1,\ldots,K} L_k$. For $k = 1, \ldots, K$ define the configuration $\sigma_k \in \mathcal{X}$ as

$$\sigma_k(v) = \begin{cases} 1 & \text{if } v \in V_k, \\ 0 & \text{otherwise.} \end{cases}$$

The configurations $\{\sigma_1, \ldots, \sigma_K\}$ are precisely all the local minima of the energy function H on the state space $\mathcal{X}$. Moreover σ_k is a stable state if and only if $L_k = L_{\max}$. In addition, denote by $\mathbf{0}$ the configuration in $\mathcal{X}$ where all the sites are empty, i.e. the configuration such that $\mathbf{0}(v) = 0$ for every $v \in V$. Given $k_1, k_2 \in \{1, \ldots, K\}$, $k_1 \neq k_2$, we take σ_{k_1} and σ_{k_2} as starting and target configurations, respectively. Define $L_* = L_*(k_2) := \max_{k \neq k_2} L_k$ and let $K_* =$

$K_*(k_2) := \{k \neq k_2 \mid L_k = L_*\}$ be the set of indices of the components of size L_* different from k_2.

In Subsection 3.4.5 we presented the main results for hitting times on complete partite graphs in the case of homogeneous activation rates. Such results are for the continuous-time activity process, but they can be easily translated to discrete time as follows. Recall that given two real functions $f(\beta)$ and $g(\beta)$, we write $f \sim g$ as $\beta \to \infty$ when $\lim_{\beta\to\infty} f(\beta)/g(\beta) = 1$.

Proposition 4.2.21 (First moment convergence of the hitting time $\tau_{\sigma_{k_2}}^{\sigma_{k_1}}$)**.** *For any $k_1, k_2 \in \{1, \dots, K\}$ with $k_1 \neq k_2$, the first hitting time $\tau_{\sigma_{k_2}}^{\sigma_{k_1}}$ satisfies*

$$\mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}} \sim N\left(\frac{\mathbb{1}_{\{k_1 \in K_*\}}}{L_*} + \frac{|K_*|}{L_{k_2}}\right) e^{\beta L_*}, \quad \beta \to \infty.$$

In particular,

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}} = L_*.$$

Proposition 4.2.22 (Asymptotic distribution of the hitting time $\tau_{\sigma_{k_2}}^{\sigma_{k_1}}$)**.** *Take $k_1, k_2 \in \{1, \dots, K\}$ such that $k_1 \neq k_2$. If $k_1 \in K_*$, then*

$$\frac{\tau_{\sigma_{k_2}}^{\sigma_{k_1}}}{\mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}}} \xrightarrow{d} \mathrm{Exp}(1), \quad \beta \to \infty.$$

Instead, if $k_1 \notin K_*$, then

$$\frac{\tau_{\sigma_{k_2}}^{\sigma_{k_1}}}{\mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}}} \xrightarrow{d} Z, \quad \beta \to \infty,$$

where $Z \overset{d}{=} \sum_{i=1}^{M} Y_i$ and $(Y_i)_{i\geq 1}$ are i.i.d. exponential unit-mean random variables and M is an independent random variable with geometric distribution $\mathbb{P}(M = n) = (1-p)^n p$ for $n \in \mathbb{N} \cup \{0\}$ with success probability $p = L_{k_2}/(|K_*|L_* + L_{k_2})$.

As illustrated in Figure 3.2a, the energy landscape consists of K cycles, one for each component of G, and one trivial cycle $\{\mathbf{0}\}$ which links all the others. The depth of each of the cycles is equal to the size of the corresponding component of G. All the paths from σ_{k_1} to σ_{k_2} must at some point exit from the cycle corresponding to component k_1, at whose bottom lies σ_{k_1}. After hitting the configuration $\mathbf{0}$, they can go directly into the target cycle, i.e. the one at which bottom lies σ_{k_2}, or they may fall in one of the other $K-1$ cycles. Formalizing these simple considerations, we can prove the following proposition.

Proposition 4.2.23 (Structural properties of the energy landscape)**.** *For any $k_1, k_2 \in \{1, \dots, K\}$, $k_1 \neq k_2$,*

$$\Gamma(\sigma_{k_1}, \{\sigma_{k_2}\}) = L_{k_1} = \Psi_{\min}(\sigma_{k_1}, \{\sigma_{k_2}\}),$$

and

$$\Psi_{\max}(\sigma_{k_1}, \{\sigma_{k_2}\}) = L_* = \widetilde{\Gamma}(\mathcal{X} \setminus \{\sigma_{k_2}\}).$$

In particular, if $k_1 \notin K_*(k_2)$, Propositions 4.2.21 and 4.2.23 impy that

$$L_* = \lim_{\beta\to\infty} \frac{1}{\beta} \mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}} = \Theta(\sigma_{k_1}, \{\sigma_{k_2}\}) \not< \widetilde{\Gamma}(\mathcal{X} \setminus \{\sigma_{k_1}, \sigma_{k_2}\}) = L_*.$$

Assumption B is thus not satisfied for the the pair $(\sigma_{k_1}, \{\sigma_{k_2}\})$. Indeed there exists another configuration $\sigma_{k'}$, for some $k' \in K_*(k_2)$, $k' \neq k_1$, for which the recurrence probability

$$\mathbb{P}\left(\tau_{\{\sigma_{k_1}, \sigma_{k_2}\}}^{\sigma_{k'}} > e^{\beta(L_{k_1}+\epsilon)}\right)$$

does not vanish as $\beta \to \infty$, since component $V_{k'}$ has size $L_* > L_{k_1}$.

In fact, as established in Proposition 4.2.22, the scaled hitting time $\tau_{\sigma_{k_2}}^{\sigma_{k_1}} / \mathbb{E}\tau_{\sigma_{k_2}}^{\sigma_{k_1}}$ does not converge in distribution to an exponential random variable with unit mean as $\beta \to \infty$.

4.2.8 *Mixing time and spectral gap*

In this subsection we focus on the long-run behavior of the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ and in particular examine the rate of convergence to the stationary distribution. We measure the rate of convergence in terms of the total variation distance and the mixing time, which describes the time required for the distance to stationarity to become small. More precisely, for every $0 < \epsilon < 1$, we define the *mixing time* $t_\beta^{\text{mix}}(\epsilon)$ by

$$t_\beta^{\text{mix}}(\epsilon) := \min\{n \geq 0 \mid \max_{x\in\mathcal{X}} \|P_\beta^n(x, \cdot) - \mu_\beta(\cdot)\|_{\text{TV}} \leq \epsilon\}, \tag{4.56}$$

where $\|\nu - \nu'\|_{\text{TV}} := \frac{1}{2}\sum_{x\in\mathcal{X}} |\nu(x) - \nu'(x)|$ for any two probability distributions ν, ν' on $\mathcal{X}$. Another classical notion to investigate the speed of convergence of Markov chains is the *spectral gap*, which is defined as

$$\rho_\beta := 1 - a_\beta^{(2)}, \tag{4.57}$$

where $1 = a_\beta^{(1)} > a_\beta^{(2)} \geq \cdots \geq a_\beta^{(|\mathcal{X}|)} \geq -1$ are the eigenvalues of the matrix $(P_\beta(x, y))_{x,y\in\mathcal{X}}$. The spectral gap can be equivalently defined using the Dirichlet form associated with the pair (P_β, μ_β), see [99, Lemma 13.12]. The problem of studying the convergence rate towards stationarity for a Freidlin-Wentzell Markov chain has already been studied in [29, 65, 109, 126]. In particular, in [29] the authors characterize the order of magnitude of both its mixing time and spectral gap in terms of certain "critical depths" of the energy landscape associated with the Freidlin-Wentzell Markov chain. We summarize the results in the context of Metropolis Markov chains in the next proposition.

Proposition 4.2.24 (Mixing time and spectral gap for Metropolis Markov chains)**.** *For any* $0 < \epsilon < 1$ *and any* $s \in \mathcal{X}^s$,

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log t_\beta^{\text{mix}}(\epsilon) = \widetilde{\Gamma}(\mathcal{X} \setminus \{s\}) = \lim_{\beta\to\infty} -\frac{1}{\beta} \log \rho_\beta. \tag{4.58}$$

Furthermore, there exist two positive constants $0 < c_1 \leq c_2 < \infty$ *independent of* β *such that for every* $\beta \geq 0$

$$c_1 e^{-\beta\widetilde{\Gamma}(\mathcal{X}\setminus\{s\})} \leq \rho_\beta \leq c_2 e^{-\beta\widetilde{\Gamma}(\mathcal{X}\setminus\{s\})}. \tag{4.59}$$

4.3 PROOF OF RESULTS FOR GENERAL METROPOLIS MARKOV CHAIN

In this section we prove the results presented in Section 4.2 for a Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ with energy landscape $(\mathcal{X}, H, c)$ and inverse temperature β. For compactness, we will suppress the implicit dependence on the parameter β in the notation.

4.3.1 *Proof of Lemma 4.2.8*

If $\omega \in \Omega_{x,A}^{\text{opt}}$, then trivially $\omega \in \Omega_{x,A}$. Moreover, we claim that $\omega \in \Omega_{x,A}^{\text{opt}}$ implies $\omega \subseteq C_A^+(x)$. Indeed, by definition of an optimal path and inequality (4.25), it follows that an optimal path cannot exit from $C_A^+(x)$ since

$$\Phi_\omega = \Phi(x, A) < H(\mathcal{F}(\partial C_A^+(x))).$$

The reverse implication follows from the minimality of $C_A^+(x)$, which guarantees that $\Phi(x, A) = \max_{z\in C_A^+(x)} H(z)$. □

4.3.2 *Proof of Proposition 4.2.10*

In this proof we suppress the explicit dependence of $\Psi_{\min}(x, A)$ and $\Psi_{\max}(x, A)$ from x and A for compactness, and denote the same quantities by $\Psi_{\min}$ and $\Psi_{\max}$, respectively. We first prove the lower bound (4.32) and, in the second part of the proof, the upper bound (4.33). Consider the event $\{\tau_A^x < e^{\beta(\Psi_{\min}-\epsilon)}\}$ first. There are two possible scenarios: Either the process exits from the cycle $C_A^+(x)$ before hitting A or not. Hence,

$$\begin{aligned}
&\mathbb{P}\left(\tau_A^x < e^{\beta(\Psi_{\min}-\epsilon)}\right) = \\
&= \mathbb{P}\left(\tau_A^x < e^{\beta(\Psi_{\min}-\epsilon)}, \tau_A^x < \tau_{\partial C_A^+(x)}^x\right) + \mathbb{P}\left(\tau_{\partial C_A^+(x)}^x \leq \tau_A^x < e^{\beta(\Psi_{\min}-\epsilon)}\right) \\
&\leq \mathbb{P}\left(\tau_A^x < e^{\beta(\Psi_{\min}-\epsilon)}, \tau_A^x < \tau_{\partial C_A^+(x)}^x\right) + \mathbb{P}\left(\tau_{\partial C_A^+(x)}^x < e^{\beta(\Psi_{\min}-\epsilon)}\right).
\end{aligned} \tag{4.60}$$

The quantity $\mathbb{P}(\tau_{\partial C_A^+(x)}^x < e^{\beta(\Psi_{\min}-\epsilon)})$ is exponentially small in β for β sufficiently large, thanks to Theorem 4.2.2(i) and to the fact that $\Psi_{\min} < \Gamma(C_A^+(x))$.

In order to derive an upper bound for the first term in the right-hand side of (4.60), we introduce the following set

$$\mathcal{Z}_{\text{opt}} := \{z \in R_A(x) \setminus A \mid \Gamma(z, A) \geq \Psi_{\min}\}.$$

By definition (4.28) of $\Psi_{\min}$, every optimal path $\omega \in \Omega^{\text{opt}}_{x,A}$ must inevitably visit a cycle of depth not smaller than $\Psi_{\min}$ and therefore it has to enter the subset $\mathcal{Z}_{\text{opt}}$ before hitting A. Hence, for every $z \in \mathcal{Z}_{\text{opt}}$, conditioning on the event $\{\tau^x_A < \tau^x_{\partial C^+_A(x)}, X^\beta_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\}$, we can write

$$\tau^x_A \stackrel{d}{=} \tau^x_z + \tau^z_A,$$

and, in particular, $\tau^x_A \geq_{\text{st}} \tau^z_A$. Using this fact, we get that there exists some $k_2 > 0$ such that for β sufficiently large

$$\begin{aligned}
&\mathbb{P}\Big(\tau^x_A < e^{\beta(\Psi_{\min}-\epsilon)}, \tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) = \\
&= \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big)\mathbb{P}\Big(\tau^x_A < e^{\beta(\Psi_{\min}-\epsilon)} \mid \tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big)\cdot \\
&\quad \sum_{z\in\mathcal{Z}_{\text{opt}}} \mathbb{P}\Big(\tau^x_A < e^{\beta(\Psi_{\min}-\epsilon)} \mid \tau^x_A < \tau^x_{\partial C^+_A(x)}, X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big)\mathbb{P}\Big(X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) \sum_{z\in\mathcal{Z}_{\text{opt}}} \mathbb{P}\Big(\tau^z_A < e^{\beta(\Psi_{\min}-\epsilon)}\Big)\mathbb{P}\Big(X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) \sum_{z\in\mathcal{Z}_{\text{opt}}} \mathbb{P}\Big(\tau^z_A < e^{\beta(\Gamma(z,A)-\epsilon)}\Big)\mathbb{P}\Big(X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) \sum_{z\in\mathcal{Z}_{\text{opt}}} \mathbb{P}\Big(\tau^z_{\partial C_A(z)} < e^{\beta(\Gamma(z,A)-\epsilon)}\Big)\mathbb{P}\Big(X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) \sum_{z\in\mathcal{Z}_{\text{opt}}} e^{-k_2\beta}\cdot\mathbb{P}\Big(X_{\tau^x_{\mathcal{Z}_{\text{opt}}}} = z\Big) \\
&= \mathbb{P}\Big(\tau^x_A < \tau^x_{\partial C^+_A(x)}\Big)\cdot e^{-k_2\beta} \\
&\leq e^{-k_2\beta},
\end{aligned} \tag{4.61}$$

where we used Theorem 4.2.2(i) and the facts that $\tau^z_A \geq \tau^z_{\partial C_A(z)}$ and that $\Gamma(C_A(z)) = \Gamma(z, A) \geq \Psi_{\min}$ for every $z \in \mathcal{Z}_{\text{opt}}$.

For the upper bound, we can argue that

$$\begin{aligned}
&\mathbb{P}\Big(\tau^x_A > e^{\beta(\Psi_{\max}+\epsilon)}\Big) \\
&= \mathbb{P}\Big(\tau^x_A > e^{\beta(\Psi_{\max}+\epsilon)}, \tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) + \mathbb{P}\Big(\tau^x_A > e^{\beta(\Psi_{\max}+\epsilon)}, \tau^x_{\partial C^+_A(x)} \leq \tau^x_A\Big) \\
&\leq \mathbb{P}\Big(\tau^x_A > e^{\beta(\Psi_{\max}+\epsilon)}, \tau^x_A < \tau^x_{\partial C^+_A(x)}\Big) + \mathbb{P}\Big(\tau^x_{\partial C^+_A(x)} \leq \tau^x_A\Big).
\end{aligned}$$

The second term is exponentially small in β thanks to Theorem 4.2.2(iii) applied to the cycle $C_A^+(x)$, to which both x and at least one state of A belong.

We now turn our attention to the first term. If the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ hits the target set A before exiting from the cycle $C_A^+(x)$, then it has been following an optimal path and, in particular, before hitting A it can have visited only states in the set $R_A(x)\setminus A$. Consider a state $z \in R_A(x)\setminus A$. By definition of $R_A(x)$, z can be reached from x by means of an optimal path, i.e. there exists a path $\omega^* : z \to x$ such that $\Phi_{\omega^*} \le \Phi(x, A)$. This fact implies that $\Phi(z, A) \le \Phi(x, A)$ and thus for every path in $\omega \in \Omega_{z,A}^{\mathrm{opt}}$, we can obtain a path that belongs to $\Omega_{x,A}^{\mathrm{opt}}$ by concatenating ω^* and ω. Hence,

$$\Psi_{\max}(z, A) \le \Psi_{\max} = \Psi_{\max}(x, A). \tag{4.62}$$

Lemma 4.2.11 guarantees the existence of a cycle-path $C_1, \dots, C_n$ vtj-connected to A such that $z \in C_1$ and $C_1, \dots C_n \in \mathcal{M}(\mathcal{X}\setminus A)$. From the fact that this cycle-path is vtj-connected and Lemma 4.2.12, it follows that $H(\mathcal{B}(C_i)) \le \Phi(x, A)$. Definition (4.29), inclusion (4.38) and inequality (4.62) imply that

$$\Gamma(C_i) \le \Psi_{\max}, \quad i = 1, \dots, n.$$

For every $i = 2, \dots, n$ take a state $y_i \in \mathcal{B}(C_{i-1}) \cap C_i$. Furthermore, take $y_1 = z$ and $y_{n+1} \in \mathcal{B}(C_n) \cap A$. Consider the set of paths

$$\mathcal{E}_{\epsilon,z,A} := \mathcal{E}_{\epsilon,z,A}\,(y_1, C_1, y_2, C_2, \dots, y_n, C_n, y_{n+1})$$

consisting of the paths constructed by the concatenation of any n–tuple of paths $\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(n)}$ satisfying the following conditions:

(1) The path $\omega^{(i)}$ has length $|\omega^{(i)}| \le e^{\beta(\Psi_{\max}+\epsilon/4)}$, for any $i = 1, \dots, n$;

(2) The path $\omega^{(i)}$ joins y_i to y_{i+1}, i.e. $\omega^{(i)} \in \Omega_{y_i,y_{i+1}}$, for any $i = 1, \dots, n$;

(3) All the states $\omega_j^{(i)}$ belong to C_i for any $j = 1, \dots, |\omega^{(i)}| - 1$, for any $i = 1, \dots, n$.

We stress that the first condition restricts the set $\mathcal{E}_{\epsilon,z,A}$ to paths that spend less than $e^{\beta(\Psi_{\max}+\epsilon/4)}$ time in cycle C_i, for every $i = 1, \dots n$. Note that the length of any path $\omega \in \mathcal{E}_{\epsilon,z,A}$ satisfies the upper bound

$$|\omega| \le |\mathcal{X}| e^{\beta(\Psi_{\max}+\epsilon/4)}.$$

Moreover, since the state space $\mathcal{X}$ is finite, for β sufficiently large

$$|\omega| \le |\mathcal{X}| e^{\beta(\Psi_{\max}+\epsilon/4)} \le e^{\beta(\Psi_{\max}+\epsilon/2)} \quad \forall\, \omega \in \mathcal{E}_{\epsilon,z,A}.$$

Therefore, for every $z \in R_A(x)\setminus A$

$$\begin{aligned}\mathbb{P}\left(\tau_A^z \le e^{\beta(\Psi_{\max}+\epsilon/2)}\right) &\ge \mathbb{P}\left(\tau_A^z \le e^{\beta(\Psi_{\max}+\epsilon/2)}, (X_m)_{m=1}^{\tau_A^x} \in \mathcal{E}_{\epsilon,z,A}\right)\\ &= \mathbb{P}\left((X_m)_{m=1}^{\tau_A^z} \in \mathcal{E}_{\epsilon,z,A}\right).\end{aligned}$$

Using the Markov property, we obtain that for any $\epsilon' > 0$ and β sufficiently large

$$\begin{aligned}\mathbb{P}\Big((X_m)_{m=1}^{\tau_A^x} \in \mathcal{E}_{\epsilon,z,A}\Big) &= \prod_{i=1}^{n} \mathbb{P}\Big(\tau_{\partial C_i}^{y_i} \le e^{\beta(\Psi_{\max}+\epsilon/4)}, X_{\tau_{\partial C_i}^{y_i}}^{y_i} = y_{i+1}\Big) \\ &\ge e^{-\beta\epsilon' n} \ge e^{-\beta\epsilon'|\mathcal{X}|},\end{aligned}$$

where the second last inequality follows from Theorem 4.2.2(v). Since $e^{-\beta\epsilon'|\mathcal{X}|}$ does not depend on the initial state z,

$$\inf_{z\in R_A(x)\setminus A} \mathbb{P}\Big(\tau_A^z \le e^{\beta(\Psi_{\max}+\epsilon/2)}\Big) \ge e^{-\beta\epsilon'|\mathcal{X}|}.$$

Applying iteratively the Markov property at the times $ke^{\beta(\Psi_{\max}+\epsilon/2)}$, with $k = 1, \dots, e^{\beta\epsilon/2}$, we obtain that

$$\begin{aligned}&\mathbb{P}\Big(\tau_A^x > e^{\beta(\Psi_{\max}+\epsilon)}, \tau_A^x < \tau_{\partial C_A^+(x)}^x\Big) \\ &\quad\le \Big(\sup_{z\in R_A(x)\setminus A} \mathbb{P}\Big(\tau_A^z > e^{\beta(\Psi_{\max}+\epsilon/2)}\Big)\Big)^{e^{\beta\epsilon/2}} \\ &\quad\le \Big(1 - e^{-\beta\epsilon'|\mathcal{X}|}\Big)^{e^{\beta\epsilon/2}} \le e^{-e^{\beta(\epsilon/2-\epsilon')|\mathcal{X}|)}}.\end{aligned}$$

We remark that we can take the supremum over the states in $R_A(x) \setminus A$, since all the other states in $C_A^+(x) \setminus R_A(x)$ cannot be reached by means of an optimal path (i.e. without exiting from $C_A^+(x)$) before visiting the target subset A. Choosing $\epsilon' > 0$ small enough and β sufficiently large, we get that $e^{-e^{\beta(\epsilon/2-\epsilon')|\mathcal{X}|)}} \le e^{-k\beta}$ for any $k > 0$. □

4.3.3 *Proof of Lemma 4.2.12*

Take a path $\omega \in \Omega_{x,A}$ and the corresponding cycle-path $\mathcal{C}_\omega = (C_1, \dots, C_{m(\omega)})$.

We first show that $\omega \notin \Omega_{x,A}^{\text{vtj}}$ implies that $\Phi(\omega_{i+1}, A) > \Phi(\omega_i, A)$ for some $1 \le i \le |\omega|$. If $\omega \notin \Omega_{x,A}^{\text{vtj}}$, then the cycle-path $\mathcal{C}_\omega = (C_1, \dots, C_{m(\omega)})$ is not vtj-connected to A, which means that there exists an index $1 \le k \le m(\omega)$ such that $\partial C_k \cap C_{k+1} \neq \emptyset$, but $\mathcal{B}(C_k) \cap C_{k+1} = \emptyset$. Take the corresponding index i in the path ω such that $\omega_i \in C_k$ and $\omega_{i+1} \in \partial C_k \cap C_{k+1}$. From the fact that $\omega_{i+1} \notin \mathcal{B}(C_k)$, it follows that

$$\Phi(\omega_{i+1}, A) > \Phi(\omega_i, A).$$

Indeed, if C_k is a trivial cycle, i.e. $C_k = \{\omega_i\}$, then $\omega_{i+1} \notin \mathcal{B}(C_k)$ implies $H(\omega_{i+1}) > H(\omega_i)$ and thus

$$\Phi(\omega_{i+1}, A) \ge H(\omega_{i+1}) > H(\omega_i) = \Phi(\omega_i, A),$$

where the last equality holds since C_k is a trivial cycle in $\mathcal{M}(\mathcal{X} \setminus A)$. In the case where C_k is a non-trivial cycle, then

$$\Phi(\omega_{i+1}, A) \geq H(\omega_{i+1}) > H(\mathcal{F}(\partial C_k)) = \Phi(\omega_i, A),$$

where the last equality follows from the fact that $C_k = C_A(\omega_i)$.

We now focus on the converse implication. We want to prove that if $\omega \in \Omega^{\mathrm{vtj}}_{x,A}$ then $\Phi(\omega_{i+1}, A) \leq \Phi(\omega_i, A)$ for every $i = 1, \dots, |\omega|$. Consider the index k such that $\omega_i \in C_k$. If the states ω_i and ω_{i+1} both belong to C_k, then $C_A(\omega_i) = C_A(\omega_{i+1}) = C_k$ and $\Phi(\omega_{i+1}, A) = \Phi(\omega_i, A)$. If instead ω_i and ω_{i+1} belongs to different cycles, then $\omega_{i+1} \in \mathcal{B}(C_k) \cap C_{k+1}$ by definition of cycle-path. If $C_k = C_A(\omega_i)$ is a non-trivial cycle, then $H(\omega_{i+1}) = H(\mathcal{F}(\partial C_k)$ and thus

$$\Phi(\omega_{i+1}, A) \leq \max\{\Phi(\omega_i, A), H(\omega_{i+1})\} = H(\mathcal{F}(\partial C_k)) = \Phi(\omega_i, A).$$

Lastly, if C_k is instead a trivial cycle, then $H(\omega_{i+1}) \leq H(\omega_i) \leq \Phi(\omega_i, A)$ and thus

$$\Phi(\omega_{i+1}, A) \leq \max\{\Phi(\omega_i, A), H(\omega_{i+1})\} = \Phi(\omega_i, A).$$

4.3.4 *Proof of Lemma 4.2.13*

In (4.41) we have used the fact that the only way to exit from the tube $\mathrm{T}_A(x)$ without having hit the subset A first is to exit from the non-principal boundary of a cycle $C \in \mathfrak{T}_A(x)$. Therefore

$$\begin{aligned}
&\mathbb{P}\left(\tau^x_{\partial \mathrm{T}_A(x)} < \tau^x_A\right) = \\
&\quad = \sum_{C \in \mathfrak{T}_A(x)} \mathbb{P}\left(\tau^x_{\partial \mathrm{T}_A(x)} < \tau^x_A,\, X_{\tau^x_{\partial \mathrm{T}_A(x)} - 1} \in C,\, X_{\tau^x_{\partial \mathrm{T}_A(x)}} \notin \mathcal{B}(C)\right) \\
&\quad = \sum_{C \in \mathfrak{T}_A(x)} \sum_{z \in C} \mathbb{P}\left(\tau^x_{\partial \mathrm{T}_A(x)} < \tau^x_A,\, X_{\tau^x_{\partial \mathrm{T}_A(x)} - 1} = z,\, X_{\tau^x_{\partial \mathrm{T}_A(x)}} \notin \mathcal{B}(C)\right) \\
&\quad \leq \sum_{C \in \mathfrak{T}_A(x)} |C| \sup_{z \in C} \mathbb{P}\left(X_{\tau^z_{\partial C}} \notin \mathcal{B}(C)\right) \leq \sum_{C \in \mathfrak{T}_A(x)} |C| e^{-k_C \beta} < e^{-\kappa \beta},
\end{aligned}$$

for some $\kappa > 0$ and β sufficiently large. The second last inequality follows from Theorem 4.2.2(iv) when C is a non-trivial cycle and directly from definition (4.36) of $\mathcal{B}(C)$ and the transition probabilities (4.1) when C is a trivial cycle. Thanks to the definition (4.39) of the typical tube, $\mathbb{P}(\tau^x_{\partial \mathrm{T}_A(x)} = \tau^x_A) = 0$, since all the states of the target state A that can be hit starting from x by means of a typical path belong to $\mathrm{T}_A(x)$ and not to $\partial \mathrm{T}_A(x)$. □

4.3.5 *Proof of Proposition 4.2.15*

As mentioned in Subsection 4.2.4, this proposition is a refinement of Proposition 4.2.10, so instead of giving a full proof, we will just describe the necessary modifications. In this proof we suppress from our notation the explicit

dependence of $\Theta_{\min}(x, A)$ and $\Theta_{\max}(x, A)$ from x and A for compactness, and denote the same quantities by $\Theta_{\min}$ and $\Theta_{\max}$, respectively.

We first prove (4.46). Consider the event $\{\tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}\}$ first. There are two possible scenarios: Either the process exits the tube $\mathrm{T}_A(x)$ of typical paths before hitting A or it stays in $\mathrm{T}_A(x)$ until it hits A. Hence,

$$\begin{aligned}
&\mathbb{P}\Big(\tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}\Big) = \\
&= \mathbb{P}\Big(\tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}, \tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}\Big) + \mathbb{P}\Big(\tau^x_{\partial \mathrm{T}_A(x)} \le \tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}\Big) \\
&\le \mathbb{P}\Big(\tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}, \tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}\Big) + \mathbb{P}\Big(\tau^x_{\partial \mathrm{T}_A(x)} \le \tau_A^x\Big). \qquad (4.63)
\end{aligned}$$

Lemma 4.2.13 implies that the second term in the right-hand side of (4.63) is exponentially small in β. In order to derive an upper bound for the first term in (4.63), we introduce the set

$$\mathcal{Z}_{\mathrm{vtj}} := \{z \in \mathrm{T}_A(x) \setminus A \mid \Gamma(z, A) \ge \Theta_{\min}\}.$$

By definition (4.42) of $\Theta_{\min}$, every typical path $\omega \in \Omega^{\mathrm{vtj}}_{x,A}$ must inevitably visit a cycle of depth not smaller than $\Theta_{\min}$ and therefore has to enter the subset $\mathcal{Z}_{\mathrm{vtj}}$ before hitting A. Hence, for every $z \in \mathcal{Z}_{\mathrm{vtj}}$, conditioning on the event $\{\tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}, X_{\tau^x_{\mathcal{Z}_{\mathrm{vtj}}}} = z\}$, we can write

$$\tau_A^x \overset{d}{=} \tau_z^x + \tau_A^z,$$

and in particular we have that $\tau_A^x >_{\mathrm{st}} \tau_A^z$. Using this fact and arguing like in (4.61), we can prove that there exists $\kappa > 0$ such that β sufficiently large such that

$$\mathbb{P}\Big(\tau_A^x < e^{\beta(\Theta_{\min}-\epsilon)}, \tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}\Big) \le e^{-\kappa\beta}.$$

We now turn our attention to the proof of the upper bound (4.47). First note that

$$\begin{aligned}
&\mathbb{P}\Big(\tau_A^x > e^{\beta(\Theta_{\max}+\epsilon)}\Big) = \\
&= \mathbb{P}\Big(\tau_A^x > e^{\beta(\Theta_{\max}+\epsilon)}, \tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}\Big) + \mathbb{P}\Big(\tau_A^x > e^{\beta(\Theta_{\max}+\epsilon)}, \tau^x_{\partial \mathrm{T}_A(x)} \le \tau_A^x\Big) \\
&\le \mathbb{P}\Big(\tau_A^x > e^{\beta(\Theta_{\max}+\epsilon)}, \tau_A^x < \tau^x_{\partial \mathrm{T}_A(x)}\Big) + \mathbb{P}\Big(\tau^x_{\partial \mathrm{T}_A(x)} \le \tau_A^x\Big), \qquad (4.64)
\end{aligned}$$

where the the latter term is exponentially small in β for β sufficiently large, thanks to Lemma 4.2.13. For the first term in (4.64), we refine the argument given in the second part of the proof of Proposition 4.2.10. Consider a state $z \in \mathrm{T}_A(x) \setminus A$. Since $\mathrm{T}_A(z) \subseteq \mathrm{T}_A(x)$, it follows from (4.44) that

$$\Theta_{\max}(z, A) \le \Theta_{\max} = \Theta_{\max}(x, A). \qquad (4.65)$$

Thanks to Lemma 4.2.11, there exists a cycle-path of maximal cycles $C_1, \dots, C_n$ in $\mathcal{X} \setminus A$ that is vtj-connected to A and such that $z \in C_1$. The definition of vtj-connected cycle-path, Lemma 4.2.14 and inequality (4.65) imply that

$$\Gamma(C_i) \leq \Theta_{\max}, \quad \forall\, i = 1, \dots, n. \tag{4.66}$$

For each $i = 2, \dots, n$, take a state $y_i \in \mathcal{F}(\partial C_{i-1}) \cap C_i$. Furthermore, take $y_1 = z$ and $y_{n+1} \in \mathcal{F}(\partial C_n) \cap A$. We consider the collection of paths

$$\mathcal{E}^*_{\epsilon,z,A} := \mathcal{E}^*_{\epsilon,z,A}\left(y_1, C_1, y_2, C_2, \dots, y_n, C_n, y_{n+1}\right),$$

which consists of all paths obtained by concatenating any n–tuple of paths $\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(n)}$ satisfying the following conditions:

(1) The path $\omega^{(i)}$ has length $|\omega^{(i)}| \leq e^{\beta(\Theta_{\max}+\epsilon/4)}$, for any $i = 1, \dots, n$;

(2) The path $\omega^{(i)}$ joins y_i to y_{i+1}, i.e. $\omega^{(i)} \in \Omega_{y_i, y_{i+1}}$, for any $i = 1, \dots, n$;

(3) All the states $\omega^{(i)}_j$ belong to C_i for any $j = 1, \dots, |\omega^{(i)}| - 1$, for any $i = 1, \dots, n$.

This collection is similar to the collection $\mathcal{E}_{\epsilon,z,A}$ described in the proof of Proposition 4.2.10, but condition (1) here is stronger. Using (4.66) and arguing as in the proof of Proposition 4.2.10, we obtain that

$$\mathbb{P}\left(\tau^z_A \leq e^{\beta(\Theta_{\max}+\epsilon/2)}\right) \geq \mathbb{P}\left((X_m)_{m=1}^{\tau^x_A} \in \mathcal{E}^*_{\epsilon,z,A}\right) \geq e^{-\beta\epsilon'|\mathcal{X}|}.$$

Since $e^{-\beta\epsilon'|\mathcal{X}|}$ does not depend on the initial state z, we get for any $\epsilon' > 0$ and β sufficiently large

$$\inf_{z \in \mathrm{T}_A(x)} \mathbb{P}\left(\tau^z_A \leq e^{\beta(\Theta_{\max}+\epsilon/2)}\right) \geq e^{-\beta\epsilon'|\mathcal{X}|},$$

and thus

$$\begin{aligned}
&\mathbb{P}\left(\tau^x_A > e^{\beta(\Theta_{\max}+\epsilon)},\, \tau^x_A < \tau^x_{\partial \mathrm{T}_A(x)}\right) \\
&\quad \leq \left(\sup_{z \in \mathrm{T}_A(x) \setminus A} \mathbb{P}\left(\tau^z_A > e^{\beta(\Theta_{\max}+\epsilon/2)}\right)\right)^{e^{\beta\epsilon/2}} \\
&\quad \leq \left(1 - e^{-\beta\epsilon'|\mathcal{X}|}\right)^{e^{\beta\epsilon/2}} \leq e^{-e^{\beta(\epsilon/2-\epsilon'|\mathcal{X}|)}},
\end{aligned} \tag{4.67}$$

by applying iteratively the Markov property at the times $k e^{\beta(\Theta_{\max}+\epsilon/2)}$, with $k = 1, \dots, e^{\beta\epsilon/2}$. Choosing $\epsilon' > 0$ small enough and β sufficiently large, we get that the right-hand side of inequality (4.67) is super-exponentially small in β, which completes the proof of the upper bound (4.47). □

4.3.6 *Proof of Theorem 4.2.17*

Since Assumption (A1) holds, we set $\Theta(x,A) = \Theta_{\min}(x,A) = \Theta_{\max}(x,A)$. The starting point of the proof is the following technical lemma.

Lemma 4.3.1 (Uniform integrability)**.** *If Assumption* (A2) *holds, then for any* $\epsilon > 0$ *the variables* $Y_A^x(\beta) := \tau_A^x e^{-\beta(\Theta(x,A)+\epsilon)}$ *are uniformly integrable, i.e. there exists* $\beta_0 > 0$ *such that for any* $\delta > 0$ *there exists* $K \in (0,\infty)$ *such that for any* $\beta > \beta_0$

$$\mathbb{E}\left(Y_A^x(\beta)\mathbb{1}_{\{Y_A^x(\beta)>K\}}\right) < \delta.$$

Proof. The proof is similar to that of [106, Corollary 3.5]. It suffices to have exponential control of the tail of the random variable $Y_A^x(\beta)$ for β sufficiently large, i.e.

$$\mathbb{P}\left(Y_A^x(\beta) > n\right) = \mathbb{P}\left(\tau_A^x e^{-\beta(\Theta(x,A)+\epsilon)} > n\right) \le a^n,$$

with $a < 1$. Assumption (A2) implies that $\Theta_{\max}(z,A) \le \Theta(x,A)$ for every $z \in \mathcal{X} \setminus A$. Then, iteratively using the Markov property gives

$$\begin{aligned}\mathbb{P}\left(\tau_A^x > ne^{-\beta(\Theta(x,A)+\epsilon)}\right) &\le \left(\sup_{z\notin A}\mathbb{P}\left(\tau_A^z > e^{\beta(\Theta(x,A)+\epsilon)}\right)\right)^n \\ &\le \left(\sup_{z\notin A}\mathbb{P}\left(\tau_A^z > e^{\beta(\Theta_{\max}(z,A)+\epsilon)}\right)\right)^n,\end{aligned}$$

and the conclusion follows from Proposition 4.2.15. □

Proposition 4.2.15 implies that the random variable $Y_A^x(\beta) := \tau_A^x e^{-\beta(\Theta(x,A)+\epsilon)}$ converges to 0 in probability as $\beta \to \infty$. Lemma 4.3.1 guarantees that the sequence $(Y_A^x(\beta))_{\beta\ge\beta_0}$ is also uniformly integrable and thus $\lim_{\beta\to\infty}\mathbb{E}|Y_A^x(\beta)| = 0$. Therefore, for any $\epsilon > 0$ we have that for β sufficiently large

$$\mathbb{E}\tau_A^x < e^{\beta(\Theta(x,A)+\epsilon)}.$$

As far as the lower bound is concerned, for any $\epsilon > 0$ Proposition 4.2.15 and the identity $\Theta(x,A) = \Theta_{\min}(x,A)$ yield

$$\begin{aligned}\mathbb{E}\tau_A^x &> e^{\beta(\Theta(x,A)-\epsilon/2)}\mathbb{P}\left(\tau_A^x > e^{\beta(\Theta(x,A)-\epsilon/2)}\right) \\ &\ge e^{\beta(\Theta(x,A)-\epsilon/2)}(1-e^{-\kappa\beta}) \ge e^{\beta(\Theta(x,A)-\epsilon)}.\end{aligned}$$

Since ϵ is arbitrary, the conclusion follows.

4.3.7 *Proof of Theorem 4.2.19*

As mentioned before, the strategy is to show that the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ satisfies the assumptions of [49, Theorem 2.3], which for completeness we reproduce here. For $R > 0$ and $r \in (0,1)$, we say that the pair (x, A) with $A \subset \mathcal{X}$ satisfies $\mathrm{Rec}(R, r)$ if

$$\sup_{z\in\mathcal{X}} \mathbb{P}\Big(\tau^z_{\{x,A\}} > R\Big) \le r.$$

The quantities R and r are called *recurrence time* and *recurrence error*, respectively.

Theorem 4.3.2. [49, Theorem 2.3] *Consider a nonempty subset $A \subset \mathcal{X}$ and $x \notin A$ such that $\mathrm{Rec}(R(\beta), r(\beta))$ holds and*

(i) $\lim_{\beta\to\infty} R(\beta)/\mathbb{E}\tau^x_A(\beta) = 0$,

(ii) $\lim_{\beta\to\infty} r(\beta) = 0$.

Then there exist two functions $k_1(\beta)$ and $k_2(\beta)$ with $\lim_{\beta\to\infty} k_1(\beta) = 0$ and $\lim_{\beta\to\infty} k_2(\beta) = 0$ such that for any $s > 0$

$$\Big|\mathbb{P}\Big(\frac{\tau^x_A}{\mathbb{E}\tau^x_A} > s\Big) - e^{-s}\Big| \le k_1(\beta)e^{-(1-k_2(\beta))s}. \tag{4.68}$$

Since $\widetilde{\Gamma}(\mathcal{X}\setminus(A\cup\{x\})) < \Theta(x,A)$ by assumption, we can take $\epsilon > 0$ small enough such that $\widetilde{\Gamma}(\mathcal{X}\setminus(A\cup\{x\})) + \epsilon < \Theta(x,A)$. Proposition 4.2.7 implies that there exists $\kappa > 0$ such that for β sufficiently large the pair (x, A) satisfies $\mathrm{Rec}(e^{\beta\widetilde{\Gamma}(\mathcal{X}\setminus(A\cup\{x\}))+\epsilon)}, e^{-\kappa\beta})$, since

$$\sup_{z\in\mathcal{X}} \mathbb{P}\Big(\tau^z_{\{x,A\}} > e^{\beta(\widetilde{\Gamma}(\mathcal{X}\setminus(A\cup\{x\}))+\epsilon)}\Big) \le e^{-e^{\kappa\beta}}.$$

Clearly $r(\beta) = e^{-e^{\kappa\beta}} \to 0$ as $\beta\to\infty$ and thus assumption (ii) holds. Assumption (i) is also satisfied, since

$$\lim_{\beta\to\infty}\frac{1}{\beta}\log R(\beta) = \widetilde{\Gamma}(\mathcal{X}\setminus(A\cup\{x\})) + \epsilon < \Theta(x,A) = \lim_{\beta\to\infty}\frac{1}{\beta}\log\mathbb{E}\tau^x_A.$$

The conclusion of Theorem 4.2.19 then follows by applying Theorem 4.3.2. □

4.3.8 *Proof of Proposition 4.2.24*

The two limits in (4.58) are an almost immediate consequence of [29, Theorem 5.1] and [109, Proposition 2.1]. Indeed, we just need to show that the critical depths H_2 and H_3 (see below for their definitions) that appear in these two results are equal to $\widetilde{\Gamma}(\mathcal{X}\setminus\{s\})$, for any $s \in \mathcal{X}^s$. The critical depth H_2 is equal

to $\widetilde{\Gamma}(\mathcal{X} \setminus \{s\})$ by definition, see [29]. Note that this quantity is well defined, since its value is independent of the choice of s, as stated in [29, Theorem 5.1]. This critical depth is also known in the literature as *maximal internal resistance* of the state space $\mathcal{X}$, see [106, Remark 4.4].

The definition of the critical depth H_3 is more involved and we need some further notation. Consider the two-dimensional Markov chain $\{(X_t, Y_t)\}_{t\geq 0}$, where X_t and Y_t are two independent Metropolis Markov chains on the same energy landscape $(\mathcal{X}, H, c)$ and indexed by the same inverse temperature β. In other words, $\{(X_t, Y_t)\}_{t\geq 0}$ is the Markov chain on $\mathcal{X} \times \mathcal{X}$ with transition probabilities $P_\beta^{\otimes 2}$ given by

$$P_\beta^{\otimes 2}\Big((x,y),(w,z)\Big) = P_\beta(x,w)P_\beta(y,z) \quad \forall\,(x,y),(w,z) \in \mathcal{X}^2.$$

The critical depth H_3 is then defined as

$$H_3 := \widetilde{\Gamma}(\mathcal{X} \times \mathcal{X} \setminus D),$$

where $D := \{(x,x) \mid x \in \mathcal{X}\}$. Consider the *null-cost graph* on the set of stable states, i.e. the directed graph (V, E) with vertex set $V = \mathcal{X}^s$ and edge set

$$E = \Big\{(s,s') \in \mathcal{X}^s \times \mathcal{X}^s \;\Big|\; \lim_{\beta\to\infty} -\frac{1}{\beta}\log P_\beta(s,s') = 0\Big\}.$$

[29, Theorem 5.1] guarantees that $H_2 \leq H_3$ and states that if the null-cost graph has an aperiodic component, then $H_2 = H_3$. We claim that this condition is always satisfied by a Metropolis Markov chain with energy landscape $(\mathcal{X}, H, c)$ with a non-constant energy function H. It is enough to show that for any such a Markov chain there exists at least one stable state $s \in \mathcal{X}^s$ such that

$$\lim_{\beta\to\infty} -\frac{1}{\beta}\log P_\beta(s,s) = 0.$$

The subset $\mathcal{X} \setminus \mathcal{X}^s$ is a non-empty set, since H is non-constant. Since q is irreducible, there exists a state $s \in \mathcal{X}^s$ and $x \in \mathcal{X} \setminus \mathcal{X}^s$ such that $c(s,x) > 0$. Furthermore, we can choose $s \in \mathcal{X}^s$ and $x \in \mathcal{X} \setminus \mathcal{X}^s$ such that the difference $H(x) - H(s)$ is minimal. For this stable state s, the transition probability towards itself reads

$$\begin{aligned}
P_\beta(s,s) &= 1 - \sum_{y\neq s} c(s,y)e^{-\beta(H(y)-H(s))^+} \\
&= 1 - \sum_{s'\in\mathcal{X}^s, s'\neq s} c(s,s') - \sum_{y\in\mathcal{X}\setminus\mathcal{X}^s} c(s,y)e^{-\beta(H(y)-H(s))^+} \\
&\geq 1 - \sum_{s'\in\mathcal{X}^s, s'\neq s} c(s,s') - e^{-\beta(H(x)-H(s))^+} \sum_{y\in\mathcal{X}\setminus\mathcal{X}^s} c(s,y) \\
&\geq 1 - \sum_{s'\in\mathcal{X}^s, s'\neq s} c(s,s') - e^{-\beta(H(x)-H(s))^+}.
\end{aligned}$$

Since q is a stochastic matrix, it follows that $1 - \sum_{s' \in \mathcal{X}^s, s' \neq s} c(s, s') > 0$ independently of β and thus

$$\lim_{\beta \to \infty} -\frac{1}{\beta} \log P_\beta(s, s) = 0,$$

since for every $\epsilon > 0$ there exists $\beta_0 > 0$ such that for every $\beta > \beta_0$

$$P_\beta(s, s) \geq 1 - \sum_{s' \in \mathcal{X}^s, s' \neq s} c(s, s') - e^{-\beta(H(x) - H(s))^+} > e^{-\beta\epsilon}.$$

Finally, the bounds (4.59) follow immediately from [65, Theorem 2.1], since the quantity m which appears there is equal to $\widetilde{\Gamma}(\mathcal{X} \setminus \{s\})$ thanks to Lemma 4.2.6. □

5 HARD-CORE MODEL ON GRID GRAPHS

In this chapter we study the hard-core model on grid graphs with Metropolis dynamics, which is the discrete-time version of the saturated CSMA dynamics on grid networks described in 1.2.2. As observed in Chapter 1, these grid networks have two dominant states corresponding to the two chessboard patterns, which we denote by **e** and **o**. The goal of this chapter is to understand the asymptotic behavior of the transition times between these two dominant states by studying the tunneling time $\tau_{\mathbf{o}}^{\mathbf{e}}$ of the particle system with hard-core interaction.

Using a novel powerful combinatorial method, we identify the minimum energy barrier between **e** and **o** and prove absence of deep cycles for this model, which allows us to decouple the asymptotics for the hitting time $\tau_{\mathbf{o}}^{\mathbf{e}}$ and the study of the critical configurations. Applying the model-independent results developed in Chapter 4, we then obtain bounds in probability for $\tau_{\mathbf{o}}^{\mathbf{e}}$ and find the order of magnitude of $\mathbb{E}\tau_{\mathbf{o}}^{\mathbf{e}}$ on a logarithmic scale, which depends both on the grid dimensions and on the chosen boundary conditions. In addition, our analysis of the energy landscape shows that the scaled hitting time $\tau_{\mathbf{o}}^{\mathbf{e}}/\mathbb{E}\tau_{\mathbf{o}}^{\mathbf{e}}$ is exponentially distributed in the low-temperature regime and yields the order of magnitude of the mixing time of the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$.

After a detailed model description in Section 5.1, we present the main results of this chapter in Section 5.2, which are then proved in Section 5.3. We remark that from this chapter onwards we will use the classical notation and terminology used in the statistical physics literature. In particular, we will talk about particles rather than nodes and we will refer to dominant states as stable states; see Section 1.3 where this correspondence is described in more detail.

5.1 MODEL DESCRIPTION

The hard-core model on finite graphs has already been introduced in Subsection 1.3.1. Recall that the spatial structure of the volume where particles interact is described by means of a finite undirected graph $G = (V, E)$, where the N vertices represent the sites where particles reside and edges connect the pairs of sites where particles cannot reside simultaneously.

In Subsection 1.3.2 we described how the discrete-time evolution of this model fits in the framework of Metropolis Markov chains (see Chapter 4). Associating a variable $\sigma(v) \in \{0, 1\}$ with each site $v \in V$, indicating the absence (0) or the presence (1) of a particle in that site, the evolution of the particle system is described by the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ parametrized by

the inverse temperature β corresponding to the following energy landscape $(\mathcal{X}, H, c)$. The state space $\mathcal{X} \subset \{0,1\}^N$ is the set of *admissible configurations* on Λ, i.e. the configurations $\sigma \in \{0,1\}^N$ such that $\sigma(v)\sigma(w) = 0$ for every pair of neighboring sites v, w in G. The energy $H(\sigma)$ of an admissible configuration $\sigma \in \mathcal{X}$ is proportional to the total number of particles,

$$H(\sigma) := -\sum_{v \in V} \sigma(v). \tag{5.1}$$

The connectivity function $c : \mathcal{X} \times \mathcal{X} \to [0,1]$ allows only for single-site updates (possibly void): For any $\sigma, \sigma' \in \mathcal{X}$,

$$c(\sigma, \sigma') := \begin{cases} \frac{1}{N}, & \text{if } |\{v \in V \mid \sigma(v) \neq \sigma'(v)\}| = 1, \\ 0, & \text{if } |\{v \in V \mid \sigma(v) \neq \sigma'(v)\}| > 1, \\ 1 - \sum_{\eta \neq \sigma} c(\sigma, \eta), & \text{if } \sigma = \sigma'. \end{cases} \tag{5.2}$$

The transition probabilities of the Markov chain $\{X_t^\beta\}_{t \in \mathbb{N}}$ are uniquely determined by the inverse temperature β and the above energy landscape and they are given by (4.1). As argued in Section 4.1, the Markov chain $\{X_t^\beta\}_{t \in \mathbb{N}}$ is reversible with respect to the Gibbs measure (4.2). Thanks to the structure of the connectivity function c, the Markov chain $\{X_t^\beta\}_{t \in \mathbb{N}}$ is aperiodic and irreducible on $\mathcal{X}$, and thus ergodic on $\mathcal{X}$ with stationary distribution μ_β.

In this chapter we focus on the hard-core model on finite two-dimensional square lattices, to which we will simply refer to as *grid graphs*. More precisely, given two integers $K, L \geq 2$, we will take G to be a $K \times L$ grid graph $\Lambda = \Lambda_{K,L}$ with three possible boundary conditions: Toroidal (periodic), cylindrical (semiperiodic) and open. We denote them by $\Lambda_{K,L}^T$, $\Lambda_{K,L}^C$ and $\Lambda_{K,L}^O$, respectively. Figure 2.2 in Chapter 2 shows an example of the three possible types of boundary conditions. As illustrated in Section 2.2, each of the grid graphs described above has vertex set $V = \{0, \dots, L-1\} \times \{0, \dots, K-1\}$ and thus Λ has $N = KL$ sites in total. Every site $v \in V$ is described by its coordinates (v_1, v_2) and is called *even* (*odd*) if the sum of its two coordinates is even (odd, respectively). We denote by V_e and V_o the collection of even sites and that of odd sites of Λ, respectively.

The open grid $\Lambda_{K,L}^O$ is naturally a bipartite graph, since all neighbors of an even site are odd sites and vice versa. In contrast, the cylindrical and toric grids may not be bipartite, so that we further assume that K is an even integer for the cylindrical grid $\Lambda_{K,L}^C$ and that both K and L are even integers for the toric grid $\Lambda_{K,L}^T$. Since the bipartite structure is crucial for our methodology, we will tacitly work under these assumptions for the cylindrical and toric grids in the rest of the chapter. As a consequence, $\Lambda_{K,L}^T$ and $\Lambda_{K,L}^C$ are *balanced* bipartite graphs, i.e. $|V_e| = |V_o|$. The open grid $\Lambda_{K,L}^O$ has $|V_e| = \lceil KL/2 \rceil$ even sites and $|V_o| = \lfloor KL/2 \rfloor$ odd sites, hence it is a balanced bipartite graph if and only if the product KL is even.

We denote by $\mathbf{e}$ ($\mathbf{o}$) the configuration with a particle at each site in V_e (V_o, respectively). More precisely,

$$\mathbf{e}(v) := \begin{cases} 1 & \text{if } v \in V_e, \\ 0 & \text{if } v \in V_o, \end{cases} \quad \text{and} \quad \mathbf{o}(v) := \begin{cases} 0 & \text{if } v \in V_e, \\ 1 & \text{if } v \in V_o. \end{cases}$$

Note that $\mathbf{e}$ and $\mathbf{o}$ are admissible configurations for any of our three choices of boundary conditions, and that, in view of (5.1), $H(\mathbf{e}) = -|V_e| = -\lceil KL/2 \rceil$ and $H(\mathbf{o}) = -|V_o| = -\lfloor KL/2 \rfloor$. In the special case where $\Lambda = \Lambda_{K,L}^O$ with $KL \equiv 1 \pmod 2$, $H(\mathbf{e}) < H(\mathbf{o})$ and, as we will show in Section 5.3, $\mathcal{X}^s = \{\mathbf{e}\}$ and $\mathcal{X}^m = \{\mathbf{o}\}$. In all the other cases, we have $H(\mathbf{e}) = H(\mathbf{o})$ and $\mathcal{X}^s = \{\mathbf{e}, \mathbf{o}\}$; see Section 5.3 for details.

5.2 ASYMPTOTIC BEHAVIOR OF TUNNELING TIMES AND MIXING TIMES

In this section we present our main results, which describe the asymptotic behavior of the hard-core model on grid graphs in the low-temperature regime.

The first main result of this chapter describes the asymptotic behavior of the tunneling time $\tau_{\mathbf{o}}^{\mathbf{e}}$ for any grid graph Λ in the low-temperature regime $\beta \to \infty$. In particular, we prove the existence and find the value of an exponent $\Gamma(\Lambda) > 0$ that gives asymptotic control in probability of the hitting time $\tau_{\mathbf{o}}^{\mathbf{e}}$ on a logarithmic scale as $\beta \to \infty$ and characterizes the asymptotic order of magnitude of the mean tunneling time $\mathbb{E}\tau_{\mathbf{o}}^{\mathbf{e}}$. We further show that the tunneling time $\tau_{\mathbf{o}}^{\mathbf{e}}$ normalized by its mean converges in distribution to an exponential unit-mean random variable.

Theorem 5.2.1 (Asymptotic behavior of the tunneling time $\tau_{\mathbf{o}}^{\mathbf{e}}$)**.** *Consider the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ corresponding to hard-core dynamics on a $K \times L$ grid graph Λ as described in Section 5.1. There exists a constant $\Gamma(\Lambda) > 0$ such that*

(i) *For every $\epsilon > 0$, $\lim_{\beta\to\infty} \mathbb{P}\left(e^{\beta(\Gamma(\Lambda)-\epsilon)} < \tau_{\mathbf{o}}^{\mathbf{e}} < e^{\beta(\Gamma(\Lambda)+\epsilon)}\right) = 1$;*

(ii) $\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\mathbf{o}}^{\mathbf{e}} = \Gamma(\Lambda)$;

(iii) $\frac{\tau_{\mathbf{o}}^{\mathbf{e}}}{\mathbb{E}\tau_{\mathbf{o}}^{\mathbf{e}}} \xrightarrow{d} \mathrm{Exp}(1)$, as $\beta \to \infty$.

In the special case where $\Lambda = \Lambda_{K,L}^O$ with $KL \equiv 1 \pmod 2$, statements (i), (ii), *and* (iii) *hold also for the transition time $\tau_{\mathbf{e}}^{\mathbf{o}}$, but replacing $\Gamma(\Lambda)$ by $\Gamma(\Lambda) - 1$.*

Theorem 5.2.1 relies on the analysis of the energy landscape corresponding to the hard-core model on grid graphs, in combination with the results for hitting times presented in Chapter 4 in the general Metropolis Markov chains context. More precisely, in Section 5.3 we develop a powerful combinatorial approach that shows that Assumptions A (absence of deep cycles) and B are

satisfied by the hard-core model on a grid graph Λ. Similar combinatorial approaches will be used to analyze the hard-core model on triangular grid graphs in Chapter 6 and the Widom-Rowlison model on grid graphs in Chapter 7. Our analysis yields the value of the energy barrier $\Gamma(\Lambda)$ between $\mathbf{e}$ and $\mathbf{o}$, which turns out to depend both on the grid size and on the chosen boundary conditions, as established by the next theorem.

Theorem 5.2.2 (The exponent $\Gamma(\Lambda)$ for grid graphs)**.** *Let Λ be a $K \times L$ grid graph. Then the energy barrier $\Gamma(\Lambda)$ between $\mathbf{e}$ and $\mathbf{o}$ appearing in Theorem 5.2.1 takes the values*

$$\Gamma(\Lambda) = \begin{cases} \min\{K, L\} + 1 & \text{if } \Lambda = \Lambda^T_{K,L} \text{ and } K + L > 4, \\ \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} + 1 & \text{if } \Lambda = \Lambda^O_{K,L}, \\ \min\{K/2, L\} + 1 & \text{if } \Lambda = \Lambda^C_{K,L}. \end{cases}$$

The additional condition $K + L > 4$ leaves out the 2×2 toric grid graph, which is a special case for which our proof method does not work. However, Theorem 5.2.1 holds also in this case, since effectively $\Lambda^T_{2,2} = \Lambda^O_{2,2}$.

The proof of Theorem 5.2.2 is presented in Section 5.3. The crucial proof idea is that along the transition from $\mathbf{e}$ to $\mathbf{o}$, there must be a critical configuration where for the first time an entire row or an entire column coincides with the target configuration $\mathbf{o}$. In such a critical configuration particles reside both in even and odd sites and, due to the hard-core constraints, an interface of empty sites should separate particles with different parities. By quantifying the "inefficiency" of this critical configuration we get the minimum energy barrier that has to be overcome for the transition from $\mathbf{e}$ to $\mathbf{o}$ to occur. The proof is then concluded by exhibiting a path that achieves this minimum energy and by exploiting the absence of other deep cycles in the energy landscape. By proving that the energy landscape corresponding to the hard-core model on grid graphs exhibits absence of deep cycles, the study of the hitting time $\tau^{\mathbf{e}}_{\mathbf{o}}$ is decoupled from an exact control of the typical paths from $\mathbf{e}$ to $\mathbf{o}$. For this reason, the study of the critical configurations and of the minimal gates along the transition from $\mathbf{e}$ to $\mathbf{o}$ is beyond the scope of this chapter and will be the focus of future work.

Besides appearing in the asymptotic results for the tunneling times in Theorem 5.2.1, the exponent $\Gamma(\Lambda)$ also characterizes the asymptotic order of magnitude of the mixing time $t^{\mathrm{mix}}_\beta(\epsilon, \Lambda)$ and of the spectral gap $\rho_\beta(\Lambda)$ of the hard-core dynamics $\{X^\beta_t\}_{t \in \mathbb{N}}$ on Λ (see Subsection 4.2.8), as established by the next theorem.

Theorem 5.2.3 (Mixing time and spectral gap)**.** *For any grid graph Λ and for any $0 < \epsilon < 1$,*

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log t^{\mathrm{mix}}_\beta(\epsilon, \Lambda) = \Gamma(\Lambda) = \lim_{\beta \to \infty} -\frac{1}{\beta} \log \rho_\beta(\Lambda).$$

Furthermore, there exist two positive constants $0 < c_1 \leq c_2 < \infty$ *independent of* β *such that for every* $\beta \geq 0$

$$c_1 e^{-\beta\Gamma(\Lambda)} \leq \rho_\beta(\Lambda) \leq c_2 e^{-\beta\Gamma(\Lambda)}.$$

5.3 ENERGY LANDSCAPE ANALYSIS

This section is devoted to the analysis of the energy landscapes corresponding to the hard-core dynamics on the three different types of grid graphs presented in Section 5.1. Starting from geometrical and combinatorial properties of the admissible configurations, we prove some structural properties of the energy landscapes $\mathcal{X}_{\Lambda^T_{K,L}}$, $\mathcal{X}_{\Lambda^O_{K,L}}$ and $\mathcal{X}_{\Lambda^C_{K,L}}$. These results are precisely the model-dependent characteristics that are needed to exploit the general framework developed in Chapter 4 to obtain the main results presented in Section 5.2. These structural properties are stated in the next three theorems and the rest of this section is devoted to their proofs.

Theorem 5.3.1 (Structural properties of $\mathcal{X}_{\Lambda^T_{K,L}}$)**.** *Consider the energy landscape corresponding to the hard-core model on the* $K \times L$ *toric grid* $\Lambda^T_{K,L}$*. Then,*

(i) $\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \leq \min\{K, L\}$,

(ii) $\Gamma(\mathbf{e}, \{\mathbf{o}\}) = \min\{K, L\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\})$.

Theorem 5.3.1 implies that the conditions (4.50) and (4.55) (see Section 4.2) hold for the pair $(\mathbf{e}, \{\mathbf{o}\})$ in the energy landscape $(\mathcal{X}_{\Lambda^T_{K,L}}, H, c)$.

Hence Assumptions A and B are satisfied and the statements of Theorems 5.2.1 and 5.2.2 for a toric grid $\Lambda^T_{K,L}$ follow from Corollary 4.2.16 and Theorems 4.2.17 and 4.2.19, respectively.

Theorem 5.3.2 (Structural properties of $\mathcal{X}_{\Lambda^O_{K,L}}$)**.** *Consider the energy landscape corresponding to the hard-core model on the* $K \times L$ *open grid* $\Lambda^O_{K,L}$*. If* $KL \equiv 0 \pmod 2$*, then,*

(i) $\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \leq \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\}$,

(ii) $\Gamma(\mathbf{e}, \{\mathbf{o}\}) = \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\})$.

If instead $KL \equiv 1 \pmod 2$*, then,*

(iii) $\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) < \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\}$,

(iv) $\Gamma(\mathbf{e}, \{\mathbf{o}\}) = \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\})$ *and* $\Gamma(\mathbf{o}, \{\mathbf{e}\}) = \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}\})$.

We remark that in the case $KL \equiv 1 \pmod 2$ the inequality in (iii) is strict, while inequality in (i) is not, and this fact is crucial in order to conclude that $\mathbf{o}$ is the unique metastable state of the state space $\mathcal{X}_{\Lambda^O_{K,L}}$ when $KL \equiv 1 \pmod 2$.

Using Theorem 5.3.2, we can check that the pair $(\mathbf{e}, \{\mathbf{o}\})$ satisfies both Assumptions A and B (since both conditions (4.50) and (4.55) hold) and thus prove the asymptotic properties in Theorems 5.2.1 and 5.2.2 for the hitting times $\tau_{\mathbf{o}}^{\mathbf{e}}$ and $\tau_{\mathbf{e}}^{\mathbf{o}}$ in the case of an open grid $\Lambda_{K,L}^{O}$.

Theorem 5.3.3 (Structural properties of $\mathcal{X}_{\Lambda_{K,L}^{C}}$)**.** *Consider the energy landscape corresponding to the hard-core model on the $K \times L$ cylindrical grid $\Lambda_{K,L}^{C}$. Then,*

(i) $\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \leq \min\{K/2, L\}$,

(ii) $\Gamma(\mathbf{e}, \{\mathbf{o}\}) = \min\{K/2, L\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\})$.

Using Theorem 5.3.3, we can check that Assumptions A and B are satisfied by the pair $(\mathbf{e}, \{\mathbf{o}\})$, and then the statements of Theorems 5.2.1 and 5.2.2 for a cylindrical grid $\Lambda_{K,L}^{C}$ follow from Corollary 4.2.16 and Theorems 4.2.17 and 4.2.19.

The ideas behind the proofs of these three theorems are similar, but for clarity we present them separately in Subsections 5.3.2, 5.3.3 and 5.3.4.

Denote $\Gamma(\Lambda) := \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}\})$, where $(\mathcal{X}, H, c)$ is the energy landscape corresponding to the hard-core model on the grid Λ. In the case $\Lambda = \Lambda_{K,L}^{O}$ with $KL \equiv 1 \pmod 2$, Theorem 5.3.2 gives that $\Gamma(\Lambda) = \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\}$. In all the other cases by symmetry we have $\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}\}) = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\})$ and hence, from Theorems 5.3.1, 5.3.2 and 5.3.3 it then follows that

$$\Gamma(\Lambda) = \begin{cases} \min\{K, L\} + 1 & \text{if } \Lambda = \Lambda_{K,L}^{T}, \\ \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} + 1 & \text{if } \Lambda = \Lambda_{K,L}^{O} \text{ and } KL \equiv 0 \pmod 2, \\ \min\{\lceil K/2 \rceil, \lceil L/2 \rceil\} & \text{if } \Lambda = \Lambda_{K,L}^{O} \text{ and } KL \equiv 1 \pmod 2, \\ \min\{K/2, L\} + 1 & \text{if } \Lambda = \Lambda_{K,L}^{C}. \end{cases}$$

Lastly, the proof of Theorem 5.2.3 readily follows from the properties of the energy landscapes presented in Theorems 5.3.1, 5.3.2 and 5.3.3 and by applying Proposition 4.2.24.

5.3.1 *Preliminaries*

We next introduce some notation and definitions for grid graphs. Recall that Λ is a $K \times L$ grid graph with $K, L \geq 2$ which has $N = KL$ sites in total. We define the *energy wastage of a configuration* $\sigma \in \mathcal{X}$ on the grid graph Λ as the difference between its energy and the energy of the configuration $\mathbf{e}$, i.e.

$$U(\sigma) := H(\sigma) - H(\mathbf{e}). \tag{5.3}$$

Since $H(\mathbf{e}) = -\lceil N/2 \rceil$, we have that

$$U(\sigma) = H(\sigma) + \lceil N/2 \rceil = \lceil N/2 \rceil - \sum_{v \in V} \sigma(v).$$

Moreover, since $\mathbf{e}$ is a stable state, $U(\sigma) \geq 0$. The function $U : \mathcal{X} \to \mathbb{R}_+ \cup \{0\}$ is usually called *virtual energy* in the literature [29, 38] and satisfies the following identity

$$U(\sigma) = -\lim_{\beta \to \infty} \frac{1}{\beta} \log \mu_\beta(\sigma),$$

where μ_β is the Gibbs measure (4.2) of the Markov chain $\{\{X(t)\}_{t \geq 0}\}_{t \in \mathbb{N}}$.

We denote by c_j, $j = 0, \ldots, L-1$, the j-th column of Λ, i.e. the collection of sites whose horizontal coordinate is equal to j, and by r_i, $i = 0, \ldots, K-1$, the i-th row of Λ, i.e. the collection of sites whose vertical coordinate is equal to i, see Figure 5.1. In particular, a vertex is identified by the coordinates (j, i) if it lies at the intersection of row r_i and column c_j. In addition, define the *i-th horizontal stripe*, with $i = 1, \ldots, \lfloor K/2 \rfloor$, as

$$S_i := r_{2i-2} \cup r_{2i-1},$$

and the *j-th vertical stripe*, with $j = 1, \ldots, \lfloor L/2 \rfloor$ as

$$C_j := c_{2j-2} \cup c_{2j-1},$$

as illustrated in Figure 5.1.

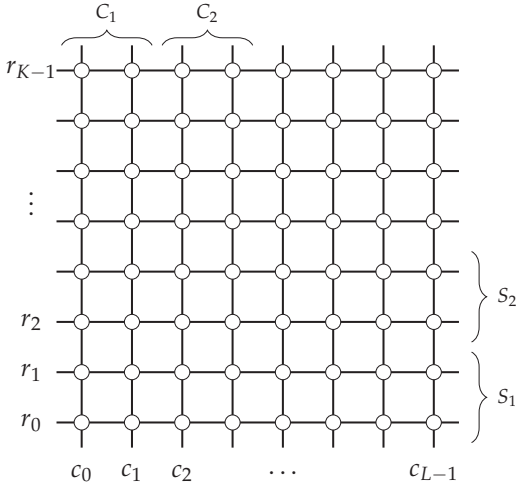


Figure 5.1: Illustration of row, column and stripe notation

An important feature of the energy wastage U for grid graphs, is that it can be seen as the sum of the energy wastages on each row (or on each horizontal stripe). More precisely, let $U_j(\sigma)$ be the energy wastage of a configuration $\sigma \in \mathcal{X}$ in the i-th row, i.e.

$$U_i(\sigma) := \lceil L/2 \rceil - \sum_{v \in r_i} \sigma(v). \tag{5.4}$$

Similarly, let $U_i^S(\sigma)$ be the energy wastage of a configuration $\sigma \in \mathcal{X}$ on the i-th horizontal stripe, i.e.

$$U_i^S(\sigma) := L - \sum_{v \in S_i} \sigma(v) = U_{2i-2}(\sigma) + U_{2i-1}(\sigma). \tag{5.5}$$

Then, we can rewrite the energy wastage of a configuration $\sigma \in \mathcal{X}$ as

$$U(\sigma) = \sum_{i=1}^{K} U_i(\sigma) = \sum_{i=1}^{\lceil K/2 \rceil} U_i^S(\sigma). \tag{5.6}$$

Given two configurations $\sigma, \sigma' \in \mathcal{X}$ and a subset of sites $W \subset V$, we write

$$\sigma_{|W} = \sigma'_{|W} \quad \Longleftrightarrow \quad \sigma(v) = \sigma'(v) \quad \forall\, v \in W.$$

We say that a configuration $\sigma \in \mathcal{X}$ has a *vertical odd (even) bridge* if there exists a column in which configuration σ perfectly agrees with $\mathbf{o}$ (respectively $\mathbf{e}$), i.e. if there exists an index $0 \le j \le L-1$ such that

$$\sigma_{|c_j} = \mathbf{o}_{|c_j} \quad (\text{respectively } \sigma_{|c_j} = \mathbf{e}_{|c_j}).$$

We define *horizontal odd and even bridges* in an analogous way and we say that a configuration $\sigma \in \mathcal{X}$ has an *odd (even) cross* if it has both vertical and horizontal odd (even) bridges.

We remark that the structure of the grid graph Λ and the hard-core constraints prohibit the existence of two perpendicular bridges of different parity, e.g. a vertical odd bridge and a horizontal even bridge. Bridges and crosses are the geometric feature of the configurations which will be crucial in the following subsections to prove Theorems 5.3.1, 5.3.2 and 5.3.3.

5.3.2 *Proofs for toric grids*

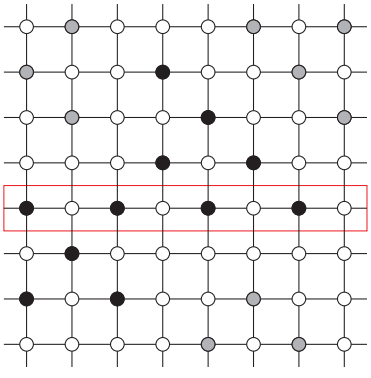
This subsection is devoted to the proof of Theorem 5.3.1 valid for the toric grid $\Lambda_{K,L}^T$. Without loss of generality, we assume henceforth that $K \le L$, and that $K + L > 4$, in view of the remark after Theorem 5.2.2. Recall that by construction of the toric grid Λ, both K and L are even integers. In the remainder of the section we will write $\mathcal{X}$ instead of $\mathcal{X}_{\Lambda_{K,L}^T}$ to keep the notation light.

We first introduce a *reduction algorithm*, which is used to construct a specific path in $\mathcal{X}$ from any given state in $\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$ to the subset $\{\mathbf{e}, \mathbf{o}\}$ and to show that

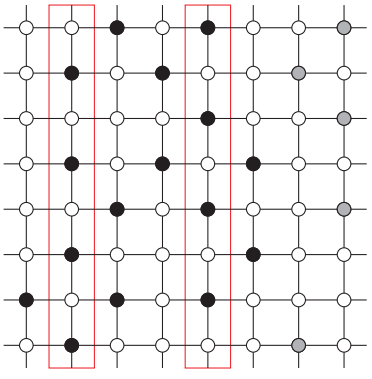
$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \le K, \tag{5.7}$$

which proves Theorem 5.3.1(i). Afterwards, we show in Proposition 5.3.5 that

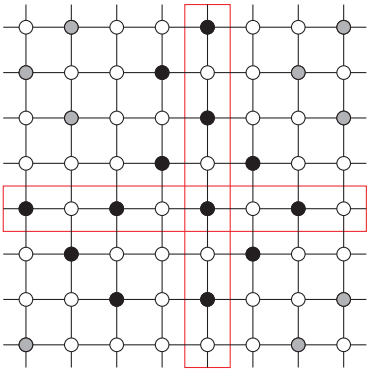
$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) \ge K + 1,$$



(a) Horizontal odd bridge



(b) Two vertical odd bridges



(c) Odd cross

Figure 5.2: Examples of configurations on the 8×8 toric grid displaying odd bridges or crosses

by giving lower bounds on the energy wastage along every path $\mathbf{e} \to \mathbf{o}$. The reduction algorithm is then used again in Proposition 5.3.6 to build a reference path $\omega^* : \mathbf{e} \to \mathbf{o}$ which shows that the lower bound is sharp and hence

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) = K + 1,$$

which, together with (5.7), proves Theorem 5.3.1(ii).

The starting point of the energy landscape analysis is a very simple observation: A configuration in $\mathcal{X}$ has zero energy wastage in a given row (column) if and only if it has an odd or even horizontal (vertical) bridge. The following lemma formalizes this property. We give the statement and the proof only for rows, since those for columns are analogous.

Lemma 5.3.4 (Equivalent characterization of bridges). *For any $\sigma \in \mathcal{X}$ and any $i = 0, \dots, K-1$,*

$$U_i(\sigma) = 0 \quad \Longleftrightarrow \quad \sigma_{|r_i} = \mathbf{e}_{|r_i} \quad \text{or} \quad \sigma_{|r_i} = \mathbf{o}_{|r_i}.$$

Proof. The i-th row of the toric grid graph Λ is a cycle graph with $L/2$ even sites and $L/2$ odd sites. If $\sigma_{|r_i} = \mathbf{e}_{|r_i}$ or $\sigma_{|r_i} = \mathbf{o}_{|r_i}$, then trivially there are $L/2$ occupied sites and hence $U_i(\sigma) = 0$. Noticing that the configurations $\mathbf{e}_{|r_i}$ and $\mathbf{o}_{|r_i}$ in row i correspond to the only two maximum independent sets of the cycle graph r_i proves the converse implication. □

Reduction algorithm for toric grids

We now describe an iterative procedure that builds a path ω in $\mathcal{X}$ from a suitable initial configuration σ (with specific properties, see below) to state $\mathbf{o}$. We call it a *reduction algorithm*, because along the path it creates the even clusters are gradually reduced and they eventually disappear, since the final configuration is $\mathbf{o}$.

The algorithm requires that the initial configuration σ is such that there are no particles in the even sites of the first vertical stripe C_1, i.e.

$$\sum_{v \in C_1 \cap V_e} \sigma(v) = 0. \tag{5.8}$$

This technical assumption is required because the algorithm needs "some room" to start working, as will become clear later. The path ω is the concatenation of L paths $\omega^{(1)}, \dots, \omega^{(L)}$. For every j along path $\omega^{(j)}$ the even sites of the $j+1$-th column are emptied, while the j-th column of σ is progressively changed to agree with $\mathbf{o}$. We will show that all configurations visited by ω are admissible and that the maximum energy achieved along the path ω is $H(\sigma) + 1$. More specifically, path $\omega^{(j)}$ goes from σ_j to σ_{j+1}, where we set $\sigma_1 = \sigma$ and recursively define for $j = 1, \dots, L$

$$\sigma_{j+1}(v) := \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_j \cup c_{j+1}), \\ \mathbf{o}(v) & \text{if } v \in c_j, \\ \sigma_j(v) & \text{if } v \in c_{j+1} \cap V_o, \\ 0 & \text{if } v \in c_{j+1} \cap V_e. \end{cases}$$

Clearly, due to the periodic boundary conditions, the column index should be taken modulo L. It can be checked that indeed $\sigma_{L+1} = \mathbf{o}$. We now describe in detail how to construct each of the paths $\omega^{(j)}$ for $j = 1, \dots, L$. We build a path $\omega^{(j)} = (\omega_0^{(j)}, \omega_1^{(j)}, \dots, \omega_K^{(j)})$ of length $K+1$ (but possibly with void moves), with $\omega_0^{(j)} = \sigma_j$ and $\omega_K^{(j)} = \sigma_{j+1}$. We start from configuration $\omega_0^{(j)} = \sigma_j$ and we repeat iteratively the following procedure for all $i = 0, \dots, K-1$:

- If $i \equiv 0 \pmod 2$, consider the even site $v = (j+1, i + (j+1 \pmod 2))$.
  - If $\omega_i^{(j)}(v) = 0$, we set $\omega_{i+1}^{(j)} = \omega_i^{(j)}$ and thus $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)})$.
  - If $\omega_i^{(j)}(v) = 1$, then we remove from configuration $\omega_i^{(j)}$ the particle in v increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(j)}$, which is such that $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) + 1$.

- If $i \equiv 1 \pmod 2$, consider the odd site $v = (j, i - 1 + (j + 1 \pmod 2))$.
  - If $\omega_i^{(j)}(v) = 1$, we set $\omega_{i+1}^{(j)} = \omega_i^{(j)}$ and thus $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)})$.
  - If $\omega_i^{(j)}(v) = 0$, then we add a particle in site v obtaining in this way a new configuration $\omega_{i+1}^{(j)}$ with energy $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) - 1$. This new configuration is admissible because all neighboring sites of v are unoccupied by construction: In particular, the particle at its right (i.e. that at the site $v + (1, 0)$) may have been removed exactly at the previous step.

Note that for the last path $\omega^{(L)}$ all the moves corresponding to even values of i are void (there are no particles in the even sites of c_0). The way the paths $\omega^{(1)}, \dots, \omega^{(L)}$ are constructed shows that for every $j = 1, \dots, L$,

$$H(\sigma_{j+1}) \le H(\sigma_j),$$

since the number of particles added in (the odd sites of) column c_j is greater than or equal to the number of particles removed in (the even sites of) column c_{j+1}. Moreover,

$$\Phi_{\omega^{(j)}} \le H(\sigma_j) + 1,$$

since along the path $\omega^{(j)}$ every particle removal (if any) is always followed by a particle addition. These two properties imply that the path $\omega : \sigma \to \mathbf{o}$ created by concatenating $\omega^{(1)}, \dots, \omega^{(L)}$ satisfies

$$\Phi_\omega \le H(\sigma) + 1.$$

Proof of Theorem 5.3.1(i). It is enough to show that for every $\sigma \in \mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$

$$\Phi(\sigma, \mathbf{o}) - H(\sigma) \le K,$$

since inequality (5.7) then follows the equivalent characterization of $\widetilde{\Gamma}$ given in Lemma 4.2.6. To prove such an inequality, we have to exhibit for every $\sigma \in \mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$ a path $\omega : \sigma \to \mathbf{o}$ in $\mathcal{X}$ such that $\Phi_\omega = \max_{\eta \in \omega} H(\eta) \le H(\sigma) + K$. We construct such a path ω as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \sigma \to \sigma'$, $\omega^{(2)} : \sigma' \to \mathbf{o}$ and σ' is a suitable configuration which depends on σ (see definition below).

Since $\sigma \neq \mathbf{e}$ by assumption, the configuration σ must have a vertical stripe with *strictly* less than K even occupied sites. Without loss of generality (modulo a cyclic rotation of column labels) we can assume that this vertical stripe is the first one, C_1, and we define

$$b := \sum_{v \in C_1 \cap V_e} \sigma(v) \le K - 1. \tag{5.9}$$

Define σ' as the configuration that differs from σ only in the even sites of the first vertical stripe, i.e.

$$\sigma'(v) := \begin{cases} \sigma(v) & \text{if } v \in V \setminus (C_1 \cap V_e), \\ 0 & \text{if } v \in C_1 \cap V_e. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{b+1}^{(1)})$, with $\omega_1^{(1)} = \sigma$ and $\omega_{b+1}^{(1)} = \sigma'$ can be constructed as follows. For $i = 1, \dots, b$, in step i we remove from configuration $\omega_i^{(1)}$ the first particle in $C_1 \cap V_e$ in lexicographic order obtaining in this way configuration $\omega_{i+1}^{(1)}$, increasing the energy by 1. Therefore, the configuration σ' is such that $H(\sigma') - H(\sigma) = b$ and

$$\Phi_{\omega^{(1)}} = \max_{\eta \in \omega^{(1)}} H(\eta) \leq H(\sigma) + b.$$

The path $\omega^{(2)} : \sigma' \to \mathbf{o}$ is then constructed by means of the reduction algorithm described earlier, choosing σ' as initial configuration and $\mathbf{o}$ as target configuration. The reduction algorithm guarantees that

$$\Phi_{\omega^{(2)}} = \max_{\eta \in \omega^{(2)}} H(\eta) \leq H(\sigma') + 1.$$

The concatenation of the two paths $\omega^{(1)}$ and $\omega^{(2)}$ gives a path $\omega : \sigma \to \mathbf{o}$ which satisfies the inequality $\Phi_\omega \leq H(\sigma) + b + 1$, which, using (5.9), implies that

$$\Phi(\sigma, \mathbf{o}) - H(\sigma) \leq b + 1 \leq K. \qquad \square$$

Proposition 5.3.5 (Lower bound for $\Phi(\mathbf{e}, \mathbf{o})$)**.** *Consider the $K \times L$ toric grid $\Lambda_{K,L}^T$ with $K \leq L$. The communication height between $\mathbf{e}$ and $\mathbf{o}$ in the corresponding energy landscape satisfies*

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) \geq K + 1.$$

Proof. We need to show that in every path $\omega : \mathbf{e} \to \mathbf{o}$, there is at least one configuration with energy wastage greater than or equal to $K + 1$. Take a path $\omega = (\omega_1, \dots, \omega_n)$ from $\mathbf{e}$ to $\mathbf{o}$. Without loss of generality, we may assume that there are no void moves in ω, i.e. at every step either a particle is added or a particle is removed, so that $H(\omega_{i+1}) = H(\omega_i) \pm 1$ for every $1 \leq i \leq n - 1$. Since $\mathbf{e}$ has no odd bridge and $\mathbf{o}$ does, at some point along the path ω there must be a configuration ω_{m^*} which is the first to display an odd bridge, horizontal or vertical, or both simultaneously. In symbols

$$m^* := \min\{m \leq n \mid \exists\, i \,:\, (\omega_m)_{|r_i} = \mathbf{o}_{|r_i} \quad \text{or} \quad \exists\, j \,:\, (\omega_m)_{|c_j} = \mathbf{o}_{|c_j}\}.$$

Clearly $m^* > 2$. We claim that $U(\omega_{m^*-1}) \geq K + 1$ or $U(\omega_{m^*-2}) \geq L + 1$. We distinguish the following three cases:

(a) ω_{m^*} displays an odd vertical bridge only;

(b) ω_{m^*} displays an odd horizontal bridge only;

(c) ω_{m^*} displays an odd cross.

These three cases cover all the possibilities, since the addition of a single particle cannot create more than one bridge in each direction.

For case (a), we claim that the energy wastage of configuration ω_{m^*} on every row is greater than or equal to one. Suppose by contradiction that there exists a row r such that $U^r(\sigma) = 0$. Then, by Lemma 5.3.4, there should be a bridge in row r; however, it cannot be an odd bridge, since otherwise we would be in case (c), and it cannot be an even bridge either, because it cannot coexist with the odd vertical bridge that configuration ω_{m^*} has. Therefore,

$$U(\omega_{m^*}) = \sum_{i=0}^{K-1} U^{r_i}(\omega_{m^*}) \geq K.$$

The previous configuration ω_{m^*-1} along the path ω differs from ω_{m^*} in a unique site v^*. By definition of m^*, v^* is an odd site and such that $\omega_{m^*-1}(v^*) = 0$ and $\omega_{m^*}(v^*) = 1$. Thus,

$$U(\omega_{m^*-1}) = U(\omega_{m^*-1}) + 1 \geq K + 1.$$

For case (b) we can argue as in case (a), but interchanging the role of rows and columns, and obtain that

$$U(\omega_{m^*-1}) \geq L + 1 \geq K + 1.$$

For case (c), the vertical and horizontal odd bridges that ω_{m^*} has, must necessarily meet in the odd site v^*. Having an odd cross, ω_{m^*} cannot have any horizontal or vertical even bridge. Consider the previous configuration ω_{m^*-1} along the path ω, which can be obtained from ω_{m^*} by removing the particle in v^*. From these considerations and from the definition of m^* it follows that ω_{m^*-1} has no vertical bridge (neither odd or even) and thus, by Lemma 5.3.4, it has energy wastage at least one in every column, which amounts to

$$U(\omega_{m^*-1}) \geq L.$$

If there is at least one column in which ω_{m^*-1} has energy wastage strictly greater than one, we get

$$U(\omega_{m^*-1}) \geq L + 1,$$

and the claim is proved. Consider now the other scenario, in which the configuration ω_{m^*-1} has energy wastage exactly one in every column, which means $U(\omega_{m^*-1}) = L$. Consider its predecessor in the path ω, namely the configuration ω_{m^*-2}. We claim that

$$U(\omega_{m^*-2}) = L + 1.$$

By construction, configuration ω_{m^*-2} must differ in exactly one site from ω_{m^*-1} and therefore

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1.$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = L - 1$. In this case the configuration ω_{m^*-2} must have a zero-energy-wastage column and by Lemma 5.3.4 it would be a vertical bridge. If it was an odd vertical bridge, the definition of m^* would be violated. If it was an even vertical bridge, it would be impossible to obtain the odd horizontal bridge (which ω_{m^*} has) in just two single-site updates, since three is the minimum number of single-site updates needed. Therefore

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) + 1 = L + 1. \qquad \square$$

The next proposition shows that the lower bound in Proposition 5.3.5 is sharp and concludes the proof of Theorem 5.3.1(ii), in view of (5.7).

Proposition 5.3.6 (Reference path)**.** *There exists a path $\omega^* : \mathbf{e} \to \mathbf{o}$ in $\mathcal{X}$ such that*

$$\Phi_{\omega^*} - H(\mathbf{e}) = K + 1.$$

Proof. We construct such a path ω^* as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{e} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{o}$, and prove that $\Phi_{\omega^{(1)}} = H(\sigma^*) = H(\sigma) + K$ and that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ are satisfied, so that $\Phi_{\omega^*} = \max_{\eta \in \omega^*} H(\eta) = H(\mathbf{e}) + K + 1$ as desired. The reason why ω is best described as the concatenation of two shorter paths is the following: The reduction algorithm cannot in general be started directly from $\mathbf{e}$ and the path $\omega^{(1)}$ indeed leads from $\mathbf{e}$ to σ^*, which is a suitable configuration to initialize the reduction algorithm. The configuration σ^* differs from $\mathbf{e}$ only in the even sites of the first vertical stripe:

$$\sigma^*(v) := \begin{cases} \mathbf{e}(v) & \text{if } v \in V \setminus C_1, \\ 0 & \text{if } v \in C_1. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{K+1}^{(1)})$, with $\omega_1^{(1)} = \mathbf{e}$ and $\omega_{K+1}^{(1)} = \sigma^*$ can be constructed as follows. For $i = 1, \dots, K$, at step i we remove from configuration $\omega_i^{(1)}$ the first particle in $C_1 \cap V_e$ in lexicographic order, increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore the configuration σ' is such that $H(\sigma^*) - H(\mathbf{e}) = K$ and $\Phi_{\omega^{(1)}} = H(\mathbf{e}) + K$. The second path $\omega^{(2)} : \sigma^* \to \mathbf{o}$ is then constructed by means of the reduction algorithm, which can be used since the configuration σ^* satisfies condition (5.8) and hence is a suitable initial configuration for the algorithm. The algorithm guarantees that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ and thus the conclusion follows. $\square$

5.3.3 *Proofs for open grids*

We now prove Theorem 5.3.2 for the open grid $\Lambda_{K,L}^O$. Also in this case, we assume without loss of generality that $K \le L$. Recall that K and L are positive

integers, not necessarily even as in the previous subsection. In the remainder of the section we will write $\mathcal{X}$ instead of $\mathcal{X}_{\Lambda^O_{K,L}}$.

We first introduce a modification of the previous reduction algorithm tailored for open grids. The scope of this reduction algorithm is twofold. It is used first to build a specific path in $\mathcal{X}$ from any given state in $\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$ to the subset $\{\mathbf{e}, \mathbf{o}\}$ and to prove that if $KL \equiv 0 \pmod 2$, then

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \leq \lceil K/2 \rceil, \tag{5.10}$$

which is Theorem 5.3.2(i). The same argument also shows that if $KL \equiv 1 \pmod 2$, then

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) < \lceil K/2 \rceil, \tag{5.11}$$

and also Theorem 5.3.2(iii) is proved. By giving a lower bound on the energy wastage along every path $\mathbf{e} \to \mathbf{o}$, we show in Proposition 5.3.8 that

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) \geq \lceil K/2 \rceil + 1.$$

Then, using again the reduction algorithm for open grids, we construct a reference path $\omega^* : \mathbf{e} \to \mathbf{o}$ which proves that the lower bound above is sharp and hence

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) = \lceil K/2 \rceil + 1. \tag{5.12}$$

In the special case $KL \equiv 1 \pmod 2$, since $\Phi(\mathbf{o}, \mathbf{e}) = \Phi(\mathbf{e}, \mathbf{o})$ and $H(\mathbf{o}) = H(\mathbf{e}) + 1$, we can easily derive from the last equality that

$$\Phi(\mathbf{o}, \mathbf{e}) - H(\mathbf{o}) = \lceil K/2 \rceil. \tag{5.13}$$

Lastly, we combine inequality (5.10) and equation (5.12) to obtain

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{o}\}) = \lceil K/2 \rceil + 1,$$

which concludes the proof of Theorem 5.3.2(ii). In the special case $KL \equiv 1 \pmod 2$, inequality (5.11) and equation (5.13) prove Theorem 5.3.2(iv), since they yield that

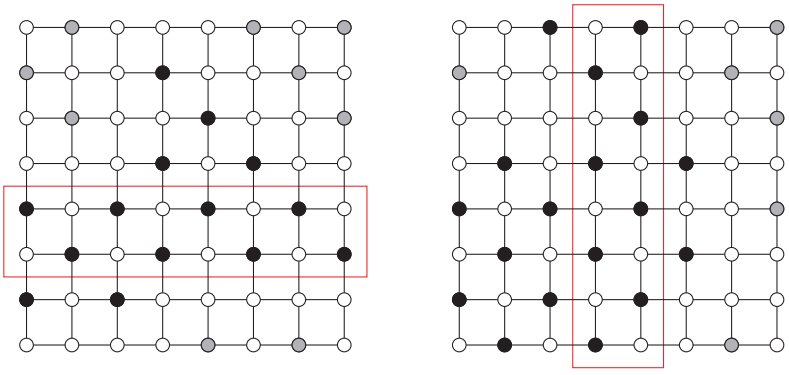
$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}\}) = \lceil K/2 \rceil.$$

We need one additional definition: Say that a configuration in $\mathcal{X}$ displays an *odd (even) vertical double bridge* if there exists at least one vertical stripe S_i in which configuration σ perfectly agrees with $\mathbf{o}$ (respectively $\mathbf{e}$), i.e. if there exists an index $1 \leq j \leq \lfloor L/2 \rfloor$ such that

$$\sigma_{|C_j} = \mathbf{o}_{|C_j} \quad (\text{respectively } \sigma_{|C_j} = \mathbf{e}_{|C_j}).$$

An *odd (even) horizontal double bridge* is defined analogously. The two types of double bridges are illustrated in Figure 5.3.

Observe that an admissible configuration on the open grid has zero energy wastage in a horizontal (vertical) stripe if and only if it has an odd or even horizontal (vertical) bridge in that stripe. The next lemma formalizes this property. We give the statement and the proof only for horizontal stripes, since those for vertical stripes are analogous. In the special case of an open grid where $KL \equiv 1 \pmod 2$, the topmost row and the leftmost column need special treatment, since they do not belong to any stripe. The second part of the following lemma shows that an admissible configuration has zero energy wastage in that row/column if and only if they agree perfectly with $\mathbf{e}$ therein. Again we will state and prove the result for the topmost row, the result for the leftmost column is analogous.



(a) Odd horizontal double bridge (b) Odd vertical double bridges

Figure 5.3: Examples of configurations on the 8×8 open grid displaying an odd double bridge

Lemma 5.3.7 (Equivalent characterization of double bridges)**.** *Consider a configuration $\sigma \in \mathcal{X}$.*

(i) *For any $i = 0, \dots, \lfloor K/2 \rfloor - 1$, the energy wastage $U_i^S(\sigma)$ in horizontal stripe S_i satisfies*

$$U_i^S(\sigma) = 0 \iff \sigma_{|S_i} = \mathbf{e}_{|S_i} \quad \text{or} \quad \sigma_{|S_i} = \mathbf{o}_{|S_i}.$$

(ii) *If additionally $KL \equiv 1 \pmod 2$, then the energy wastage in the topmost row $U_{K-1}(\sigma)$ satisfies*

$$U_{K-1}(\sigma) = 0 \iff \sigma_{|r_{K-1}} = \mathbf{e}_{|r_{K-1}}.$$

Proof. We prove statement (i) first. Consider the $2 \times L$ grid graph induced by the horizontal stripe S_i: It has L even sites and L odd sites. If $\sigma_{|S_i} = \mathbf{e}_{|S_i}$ or $\sigma_{|S_i} = \mathbf{o}_{|S_i}$, trivially $U_i^S(\sigma) = 0$. Let us prove the converse implication. Denote by e_t (e_b) the number of particles present in even sites in the top (bottom) row

of stripe S_i. Analogously, define o_t (o_b) as the number of particles present in odd sites in the top (bottom) row of stripe S_i. We will show that:

(1) $U_i^S(\sigma) = 0$ and $e_t + e_b = 0 \iff o_t + o_b = L$;

(2) $U_i^S(\sigma) = 0$ and $e_t + e_b > 0 \iff o_t + o_b = L$.

The proofs of statement (1) and of the implication (2 $\Leftarrow$) are immediate. Thus we focus on the implication (2 $\Rightarrow$).

Note that if $e_t + e_b \in [1, L-1]$ particles are present in even sites, then they block at least $e_t + e_b + 1$ odd sites, which must then be unoccupied. Indeed in the top row each of the e_t particles blocks the odd node at its right and in the bottom row each of the e_b particles blocks the odd node at its left. In one of the two rows, say the top one, there is at least one even unoccupied site and consider the even site at its right where a particle resides. This particle blocks also the odd site at its left. Hence $o_t + o_b \leq L - (e_t + e_b + 1)$, which gives $U_i^S(\sigma) = L - (e_t + e_b + o_t + o_b) > 0$.

We now turn to the proof of statement (ii). The topmost row has $\frac{L+1}{2}$ even sites and $\frac{L-1}{2}$ odd sites. Denote by e (respectively o) the number of particles present in even (respectively odd) sites in row r_{K-1}. The energy wastage of σ on this row can be computed as $U_{K-1}(\sigma) = \frac{L+1}{2} - e - o$. Trivially, if $\sigma_{|r_{K-1}} = \mathbf{e}_{|r_{K-1}}$, then $e = \frac{L+1}{2}$ and thus $U_{K-1}(\sigma) = 0$. Let us prove the opposite implication. Assume that $\sigma_{|r_{K-1}} \neq \mathbf{e}_{|r_{K-1}}$, i.e. $e < \frac{L+1}{2}$. If $e = 0$, then $U_{K-1}(\sigma) \geq 1$, since $o \leq \frac{L-1}{2}$. If instead $e \in [1, \frac{L+1}{2} - 1]$, then each particle residing in an even site blocks the odd site at its left, therefore $o \leq \frac{L-1}{2} - e$, which implies

$$U_{K-1}(\sigma) = \frac{L+1}{2} - e - o \geq \frac{L+1}{2} - e - \left(\frac{L-1}{2} - e\right) \geq 1. \qquad \square$$

Reduction algorithm for open grids

We now describe the *reduction algorithm for open grids*, which is a modification of the reduction algorithm for toric grids that builds a path ω in $\mathcal{X}$ from a given initial configuration σ to *either* $\mathbf{o}$ or $\mathbf{e}$. The reduction algorithm for open grids takes two inputs instead of one: The initial configuration σ and the target state which is either $\mathbf{o}$ or $\mathbf{e}$. This is the first crucial difference with the corresponding algorithm for toric grids, where the target configuration was always $\mathbf{o}$. In the following, we first assume that the target state is $\mathbf{o}$ and illustrate the procedure in this case. The necessary modifications when the target state is $\mathbf{e}$ are presented later.

The initial configuration σ for the reduction algorithm must be such that there are no particles in the even sites of the first column c_0, i.e.

$$\sum_{v \in c_0 \cap V_e} \sigma(v) = 0. \tag{5.14}$$

This condition ensures that the algorithm has enough "room" to work properly. Note that condition (5.14) is different from condition (5.8) for the reduction algorithm for toric grids, which requires instead that the even sites of *both* the first two columns c_0 and c_1 should be empty.

The path ω is the concatenation of L paths $\omega^{(1)}, \dots, \omega^{(L)}$. Path $\omega^{(j)}$ goes from σ_j to σ_{j+1}, where we set $\sigma_1 = \sigma$ and recursively define for $j = 1, \dots, L$

$$\sigma_{j+1}(v) = \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_j \cup c_{j+1}), \\ \mathbf{o}(v) & \text{if } v \in c_j, \\ \sigma_j(v) & \text{if } v \in c_{j+1} \cap V_o, \\ 0 & \text{if } v \in c_{j+1} \cap V_e. \end{cases}$$

This procedure guarantees that $\sigma_{L+1} = \mathbf{o}$. The path $\omega^{(j)}$ for $j = 1, \dots, L$ is constructed exactly as the path $\omega^{(j)}$ for the reduction algorithm for toric grids. Since their construction is identical, every path $\omega^{(j)}$ enjoys the same properties as those of the original reduction algorithm, namely

$$H(\sigma_{j+1}) \leq H(\sigma_j) \quad \text{and} \quad \Phi_{\omega^{(j)}} \leq H(\sigma_j) + 1.$$

This means that the path $\omega : \sigma \to \mathbf{o}$ created by their concatenation satisfies

$$\Phi_\omega \leq H(\sigma) + 1.$$

In the scenario where the target state is $\mathbf{e}$, three modifications are needed. First the initial state σ must be such that there are no particles in the *odd* sites of the first column c_0, i.e.

$$\sum_{v \in c_0 \cap V_e} \sigma(v) = 0.$$

Secondly, the sequence of intermediate configurations $\sigma_1, \dots, \sigma_{L+1}$ must be modified as follows: We set $\sigma_1 = \sigma$ and recursively define for $j = 1, \dots, L$

$$\sigma_{j+1}(v) = \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_j \cup c_{j+1}), \\ \mathbf{e}(v) & \text{if } v \in c_j, \\ \sigma_j(v) & \text{if } v \in c_{j+1} \cap V_e, \\ 0 & \text{if } v \in c_{j+1} \cap V_o. \end{cases}$$

Lastly, for step i of path $\omega^{(j)}$, we need a different offset to select the site v, namely $v = (j, i + (j \pmod 2))$ when $i \equiv 0 \pmod 2$ and $v = (j, i - 1 + (j \pmod 2))$ when $i \equiv 1 \pmod 2$. One can check that the resulting path $\omega : \sigma \to \mathbf{e}$ satisfies the inequality

$$\Phi_\omega \leq H(\sigma) + 1.$$

Proof of Theorem 5.3.2(i). It is enough to prove that for every $\sigma \in \mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$

$$\Phi(\sigma, \{\mathbf{e}, \mathbf{o}\}) - H(\sigma) \leq \lfloor K/2 \rfloor.$$

Indeed, this claim, together with the equivalent characterization of $\widetilde{\Gamma}$ given in Lemma 4.2.6, proves simultaneously inequality (5.10) when $KL \equiv 0 \pmod 2$ and the strict inequality (5.11) when $KL \equiv 1 \pmod 2$, since in this case $\lfloor K/2 \rfloor < \lceil K/2 \rceil$. To prove such an inequality, we have to exhibit for every $\sigma \in \mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}$ a path $\omega : \sigma \to \{\mathbf{e}, \mathbf{o}\}$ in $\mathcal{X}$ such that $\Phi_\omega = \max_{\eta \in \omega} H(\eta) \leq H(\sigma) + \lfloor K/2 \rfloor$.

Let b be the number of particles present in configuration σ in the odd sites of the leftmost column of Λ, i.e.

$$b := \sum_{v \in c_0 \cap V_o} \sigma(v).$$

Every column in Λ has $\lfloor K/2 \rfloor$ odd sites, and hence $0 \leq b \leq \lfloor K/2 \rfloor$. Differently from the proof of Theorem 5.3.1(i), here the value of b determines how the path ω will be constructed. We distinguish two cases: (a) $b = \lfloor K/2 \rfloor$ and (b) $b < \lfloor K/2 \rfloor$.

Consider case (a) first and assume that $b = \lfloor K/2 \rfloor$. In this case, we construct a path $\omega : \sigma \to \mathbf{o}$ by means of the reduction algorithm for open grids, choosing as initial configuration σ and as target configuration $\mathbf{o}$. The way this path is built guarantees that $\Phi_\omega \leq H(\sigma) + 1$, which implies that

$$\Phi(\sigma, \mathbf{o}) - H(\sigma) = 1 \leq \lfloor K/2 \rfloor.$$

Let us now focus on case (b) and assume that $b < \lfloor K/2 \rfloor$. In this case we create a path $\omega : \sigma \to \mathbf{e}$ as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \sigma \to \sigma'$, $\omega^{(2)} : \sigma' \to \mathbf{e}$ and σ' is a suitable configuration which depends on σ (see definition below). The reason why ω is best described as concatenation of two shorter paths is the following: Since $b < \lfloor K/2 \rfloor$, the reduction algorithm can not be started directly from σ and the path $\omega^{(1)}$ indeed leads from σ to σ', which is a suitable configuration to initialize the reduction algorithm for open grids. The configuration σ' differs from σ only in the odd sites of the first column, that is

$$\sigma'(v) := \begin{cases} \sigma(v) & \text{if } v \in V \setminus (c_0 \cap V_o), \\ 0 & \text{if } v \in c_0 \cap V_o. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \ldots, \omega_{b+1}^{(1)})$, with $\omega_1^{(1)} = \sigma$ and $\omega_{b+1}^{(1)} = \sigma'$, can be constructed as follows. For $i = 1, \ldots, b$, at step i we remove from configuration $\omega_i^{(1)}$ the topmost particle in $c_0 \cap V_o$ increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore the configuration σ' is such that $H(\sigma') - H(\sigma) = b$ and

$$\Phi_{\omega^{(1)}} = \max_{\eta \in \omega^{(1)}} H(\eta) \leq H(\sigma) + b.$$

The path $\omega^{(2)} : \sigma' \to \mathbf{e}$ is then constructed by means of the reduction algorithm for open grids described earlier, using σ' as initial configuration and $\mathbf{e}$ as target configuration. The reduction algorithm guarantees that

$$\Phi_{\omega^{(2)}} = \max_{\eta \in \omega^{(2)}} H(\eta) \leq H(\sigma') + 1.$$

The concatenation of the two paths $\omega^{(1)}$ and $\omega^{(2)}$ gives a path $\omega : \sigma \to \mathbf{e}$ which satisfies the inequality $\Phi_\omega \leq H(\sigma) + b + 1$ and therefore

$$\Phi(\sigma, \mathbf{e}) - H(\sigma) = b + 1 \leq \lfloor K/2 \rfloor. \qquad \square$$

Proposition 5.3.8 (Lower bound for $\Phi(\mathbf{e}, \mathbf{o})$)**.** *Consider the $K \times L$ open grid $\Lambda^O_{K,L}$ with $K \leq L$. The communication height between $\mathbf{e}$ and $\mathbf{o}$ in the corresponding energy landscape satisfies*

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) \geq \lceil K/2 \rceil + 1.$$

Proof. It is enough to show that in every path $\omega : \mathbf{e} \to \mathbf{o}$ there is at least one configuration with energy wastage greater than or equal to $\lceil K/2 \rceil + 1$. Take a path $\omega = (\omega_1, \ldots, \omega_n)$ from $\mathbf{e}$ to $\mathbf{o}$. Without loss of generality, we may assume that there are no void moves in ω, i.e. at every step either a particle is added or a particle is removed, so that $H(\omega_{i+1}) = H(\omega_i) \pm 1$ for every $1 \leq i \leq n-1$. Since $\mathbf{e}$ does not have an odd bridge while $\mathbf{o}$ does, at some point along the path ω there must be a configuration ω_{m^*} which is the first to display an odd bridge, horizontal or vertical, or both simultaneously. In symbols

$$m^* := \min\{m \leq n \mid \exists\, i \,:\, (\omega_m)_{|r_i} = \mathbf{o}_{|r_i} \quad \text{or} \quad \exists\, j \,:\, (\omega_m)_{|c_j} = \mathbf{o}_{|c_j}\}.$$

Clearly $m^* > 2$. We claim that $U(\omega_{m^*-1}) \geq \lceil K/2 \rceil + 1$ or $U(\omega_{m^*-2}) \geq \lceil L/2 \rceil + 1$. We distinguish the following three cases:

(a) ω_{m^*} displays an odd vertical bridge only;

(b) ω_{m^*} displays an odd horizontal bridge only;

(c) ω_{m^*} displays an odd cross.

These three cases cover all possibilities, since the addition of a single particle cannot create more than one bridge in each direction. Let $v^* \in V$ be the unique site where configuration ω_{m^*-1} and ω_{m^*} differ.

For case (a), assume first that v^* belong to the i^*-th horizontal stripe, i.e. $v^* \in S_{i^*}$ for some $0 \leq i^* \leq \lfloor K/2 \rfloor - 1$. By construction, v^* must be an odd site and $\omega_{m^*-1}(v^*) = 0$ and $\omega_{m^*}(v^*) = 1$ and thus $U^S_{i^*}(\omega_{m^*-1}) \geq 1$. We claim that in fact

$$U^S_{i^*}(\omega_{m^*-1}) \geq 2.$$

It is enough to show that $U^S_{i^*}(\omega_{m^*-1}) \neq 1$. Suppose by contradiction that $U^S_{i^*}(\omega_{m^*-1}) = 1$, then it must be the case that $U^S_{i^*}(\omega_{m^*}) = 0$, due the addition

of a particle in v^*, and by Lemma 5.3.7 the horizontal stripe S_{i^*} must agree fully with $\mathbf{o}$ ($\omega_{m^*} \neq \mathbf{e}$, since it has a particle residing in v^* which is an odd site). This fact would imply that ω_{m^*} has an odd horizontal bridge, which contradicts our assumption for case (a).

Assume instead that K is odd and that v^* does not belong to any horizontal stripe and belongs instead to the topmost row, i.e. $v^* \in r_{K-1}$. By construction, v^* must be an odd site and $\omega_{m^*-1}(v^*) = 0$ and $\omega_{m^*}(v^*) = 1$ and thus $U_{K-1}(\omega_{m^*-1}) \geq 1$. We claim that in fact

$$U_{K-1}(\omega_{m^*-1}) \geq 2.$$

It is enough to show that $U_{K-1}(\omega_{m^*-1}) \neq 1$. Suppose by contradiction that $U_{K-1}(\omega_{m^*-1}) = 1$, then it must be $U_{K-1}(\omega_{m^*}) = 0$, due to the addition of a particle in v^*. By Lemma 5.3.7 ω_{m^*} must agree fully with $\mathbf{e}$ on this topmost row, but this cannot be the case since ω_{m^*} has a particle residing in v^* which is an odd site.

Moreover, we claim that the energy wastage in every horizontal stripe that does not contain site v^* (and in the topmost row if $KL \equiv 1 \pmod 2$ and $v^* \notin r_{K-1}$) is also greater than or equal to 1. Indeed, configuration ω_{m^*-1} cannot display any horizontal odd bridge (by definition of i^*) and neither a horizontal even bridge, since $\omega_{m^*-1}(v^* + (1,0)) = 0$ and $\omega_{m^*-1}(v^* + (-1,0)) = 0$. Therefore for every $i = 1, \ldots, \lfloor K/2 \rfloor$ such that $v^* \notin S_j$ we have $(\omega_{m^*})_{|S_i} \neq \mathbf{o}_{|S_i}, \mathbf{e}_{|S_i}$ and hence, by Lemma 5.3.7

$$U_i^S(\omega_{m^*}) \geq 1.$$

If K is odd, then the topmost row r_{K-1} cannot be a horizontal odd bridge (our assumption would be violated) and neither a horizontal even bridge (it would be impossible to obtain the horizontal odd bridge which ω_{m^*} has in a single step, the minimum number of steps needed is two). Therefore, by Lemma 5.3.7,

$$U_{K-1}(\omega_{m^*-1}) \geq 1.$$

There are three possible scenarios:

- K even: There are $K/2 - 1$ horizontal stripes with positive energy wastage and $U_{i^*}^S(\omega_{m^*-1}) \geq 2$;
- K odd and $v^* \notin r_{K-1}$: There are $\lfloor K/2 \rfloor - 2$ horizontal stripes with positive energy wastage, $U_{K-1}(\omega_{m^*-1}) \geq 1$ and $U_{i^*}^S(\omega_{m^*-1}) \geq 2$;
- K odd and $v^* \in r_{K-1}$: There are $\lfloor K/2 \rfloor - 1$ horizontal stripes with positive energy wastage and $U_{K-1}(\omega_{m^*-1}) \geq 2$.

In all three scenarios, by summing the energy wastage of the horizontal stripes (and possibly that of the topmost row) we obtain

$$U(\omega_{m^*-1}) \geq \lceil K/2 \rceil + 1.$$

For case (b) we can argue in a similar way, but interchanging the roles of rows and columns, and obtain that

$$U(\omega_{m^*-1}) \geq \lceil L/2 \rceil + 1 \geq \lceil K/2 \rceil + 1.$$

For case (c), the vertical and horizontal odd bridges that ω_{m^*} has, must necessarily meet in the odd site v^*. Having an odd cross, ω_{m^*} cannot display any horizontal or vertical even bridge. Consider the previous configuration ω_{m^*-1} along the path ω, which can be obtained from ω_{m^*} by removing the particle in v^*. From these considerations and from the definition of m^* it follows that ω_{m^*-1} has no vertical bridge (neither odd or even) and thus, by Lemma 5.3.7, it has energy wastage at least one in each of the $\lfloor L/2 \rfloor$ vertical stripes and possibly in the leftmost column, if L is odd. In both cases, we have

$$U(\omega_{m^*-1}) \geq \lceil L/2 \rceil.$$

If there is at least one column in which ω_{m^*-1} has energy wastage strictly greater than one, then the proof is concluded, since

$$U(\omega_{m^*-1}) \geq \lceil L/2 \rceil + 1 \geq \lceil K/2 \rceil + 1.$$

Consider now the other scenario, in which the configuration ω_{m^*-1} has energy wastage exactly one in every vertical stripe (and possibly in the leftmost column, if L is odd), which means $U(\omega_{m^*-1}) = \lceil L/2 \rceil$. Consider its predecessor in the path ω, namely the configuration ω_{m^*-2}. We claim that

$$U(\omega_{m^*-2}) = \lceil L/2 \rceil + 1.$$

Indeed, by construction, configuration ω_{m^*-2} must differ in exactly one site from ω_{m^*-1} and therefore

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1.$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = \lceil L/2 \rceil - 1$. In this case the configuration ω_{m^*-2} must have a zero-energy-wastage vertical stripe and by Lemma 5.3.7 it would be a vertical double bridge. If it was a vertical odd double bridge, the definition of m^* would be violated. If it was an even vertical double bridge, it would be impossible to obtain the horizontal odd bridge (which ω_{m^*} has) in just two single-site updates, since three is the minimum number of single-site updates needed. Therefore

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) + 1 = \lceil L/2 \rceil + 1. \qquad \square$$

The lower bound for the communication height $\Phi(\mathbf{e}, \mathbf{o})$ we just proved is sharp, as established by the next proposition in which a reference path from $\mathbf{e}$ to $\mathbf{o}$ is constructed.

Proposition 5.3.9 (Reference path)**.** *There exists a path $\omega^* : \mathbf{e} \to \mathbf{o}$ in $\mathcal{X}$ such that*

$$\Phi_{\omega^*} - H(\mathbf{e}) = \lceil K/2 \rceil + 1.$$

Proof. We describe just briefly how the reference path ω^* is constructed, since it is very similar to the one given in the proof of Proposition 5.3.6. Also in this case, the path ω^* is the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{e} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{o}$, where σ^* is the configuration that differs from $\mathbf{e}$ only in the even sites of the leftmost column:

$$\sigma^*(v) := \begin{cases} \mathbf{e}(v) & \text{if } v \in V \setminus c_0, \\ 0 & \text{if } v \in c_0. \end{cases}$$

The path $\omega^{(1)}$ consists of $\lceil K/2 \rceil$ steps, at each of which we remove the first particle in $c_0 \cap V_e$ in lexicographic order from the previous configuration. The last configuration is precisely σ^*, which has energy $H(\sigma^*) = H(\mathbf{e}) + \lceil K/2 \rceil$, and, trivially, $\Phi_{\omega^{(1)}} = H(\mathbf{e}) + \lceil K/2 \rceil$. The second path $\omega^{(2)} : \sigma^* \to \mathbf{o}$ is then constructed by means of the reduction algorithm, which can be used since configuration σ^* is a suitable initial configuration for it, satisfying condition (5.14). The algorithm guarantees that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ and thus the concatenation of the two paths $\omega^{(1)}$ and $\omega^{(2)}$ yields a path ω^* with $\Phi_{\omega^*} = \max_{\eta \in \omega} H(\eta) = H(\mathbf{e}) + \lceil K/2 \rceil + 1$ as desired. □

The statements (ii) and (iv) of Theorem 5.3.2 can then be easily obtained from Propositions 5.3.8 and 5.3.9, as illustrated at the beginning of Subsection 5.3.3.

5.3.4 *Proofs for cylindrical grids*

In this subsection we briefly describe how to proceed to prove Theorem 5.3.3. The cylindrical grid $\Lambda^C_{K,L}$ is a hybrid between the toric grid and the open grid, since the columns of $\Lambda^C_{K,L}$ have the same structure as the columns of the toric grid $\Lambda^T_{K,L}$, while the horizontal stripes of $\Lambda^C_{K,L}$ enjoy the same structural properties of those of the open grid $\Lambda^O_{K,L}$. Along the lines of Lemmas 5.3.4 and 5.3.7 we can prove that the only columns with zero energy wastage are vertical bridges and the only horizontal stripes with zero energy wastage are horizontal double bridges.

In order to prove that

$$\Phi(\mathbf{e}, \mathbf{o}) - H(\mathbf{e}) \geq \min\{K/2, L\} + 1,$$

one can argue in a similar way as was done for the other two types of grids. Also for the cylindrical grid, in any path $\omega : \mathbf{e} \to \mathbf{o}$ there must be a configuration ω_{m^*} which is the first to display a horizontal odd bridge or a vertical odd bridge or both simultaneously, i.e.

$$m^* := \min\{m \leq n \mid \exists\, i \,:\, (\omega_m)_{|r_i} = \mathbf{o}_{|r_i} \quad \text{or} \quad \exists\, j \,:\, (\omega_m)_{|c_j} = \mathbf{o}_{|c_j}\}.$$

One can prove that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq \min\{K/2, L\} + 1.$$

We distinguish two cases, depending on whether $K/2 \geq L$ or $K/2 < L$. In these two cases, the proof can be obtained by studying the energy wastage either in the columns or in the horizontal stripes, in the same spirit as for the toric and open grids in Subsections 5.3.2 and 5.3.3, respectively. Moreover, depending on whether $K/2 \geq L$ or $K/2 < L$, we can take the reference path ω^* to be the same as in Subsections 5.3.2 and 5.3.3, respectively. Lastly, one can show that

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{e}, \mathbf{o}\}) \leq \min\{K/2, L\},$$

by exploiting the approach in Subsection 5.3.2, if $K/2 \geq L$, and the strategy adopted in Subsection 5.3.3, otherwise.

6 HARD-CORE MODEL ON TRIANGULAR GRID GRAPHS

The goal of this chapter is to understand the evolution of the activity process described in Subsection 1.2.2 on finite triangular grid networks. This class of conflict graphs has been introduced in Section 2.2 as an example of symmetric partite graphs with three dominant states. Adopting the approach used in Chapter 5, we study the hard-core model on triangular grid graphs with Metropolis transition probabilities. The main results of this chapter describe the asymptotic behavior of the tunneling times between any pair of the three stables states in the low-temperature regime and characterize the order of magnitude of the mixing time in the same regime. The proof method relies on the analysis of the energy landscape using geometrical and combinatorial properties of the admissible configurations, in combination with model-independent results from Chapter 4.

This chapter is structured as follows. Section 6.1 gives a model description of the hard-core model on triangular grid graphs. The main results of this chapter for tunneling times and mixing times are presented in Section 6.2 and proved later in Section 6.3.

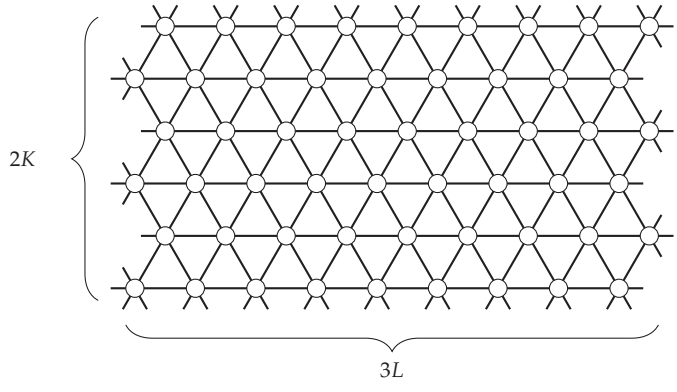
6.1 MODEL DESCRIPTION

We consider the hard-core model on finite rectangular regions of a triangular lattice with periodic boundary conditions. More precisely, given two integers $K \geq 2$ and $L \geq 1$, we consider as conflict graph the $2K \times 3L$ *triangular grid* $\mathcal{T}_{K,L}$ which consists of $2K$ rows and $6L$ columns. The graph $\mathcal{T}_{K,L} = (V, E)$ is the subgraph of the triangular lattice consisting of $N = |V| = 6KL$ sites placed on $2K$ rows of $3L$ sites each (or equivalently on $6L$ columns with K sites each), see Figure 6.1a. We impose periodic (wrap-around) boundary conditions on $\mathcal{T}_{K,L}$ to preserve symmetry.

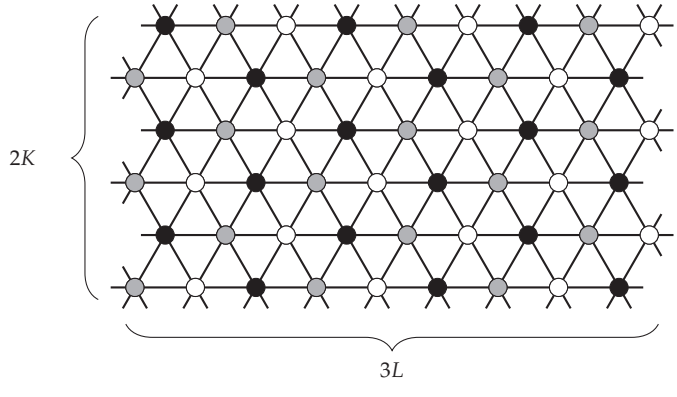
The graph $\mathcal{T}_{K,L}$ has a natural tri-partition $V = V_a \cup V_b \cup V_c$, which we illustrate in Figure 6.1b by coloring the three components in gray, black, and white, respectively. Thanks to the chosen dimensions of the triangular grid $\mathcal{T}_{K,L}$, these three components have the same number of sites

$$|V_a| = |V_b| = |V_c| = \frac{|V|}{3} = 2KL. \tag{6.1}$$

We associate a variable $\sigma(v) \in \{0, 1\}$ with each site $v \in V$, indicating the absence (0) or the presence (1) of a particle in that site. In view of the hard-core constraints, a configuration of particles on $\mathcal{T}_{K,L} = (V, E)$ is *admissible* if $\sigma(v)\sigma(w) = 0$ for every pair of neighboring sites v, w of $\mathcal{T}_{K,L}$. We denote by



(a) The triangular grid $\mathcal{T}_{K,L}$



(b) The three components of $\mathcal{T}_{K,L}$ highlighted using different colors

Figure 6.1: The triangular grid $\mathcal{T}_{K,L}$ corresponding to $K = 3$ and $L = 3$ and its three components

$\mathcal{X} \subset \{0,1\}^N$ the set of all admissible configurations on $\mathcal{T}_{K,L}$, which is in one-to-one correspondence with the collection of independent sets of the graph $\mathcal{T}_{K,L}$.

We are interested in studying the Metropolis dynamics for the hard-core model on triangular grid graphs, that is the family of Markov chains $\{X_t^\beta\}_{t\in\mathbb{N}}$ on $\mathcal{X}$ parametrized by $\beta > 0$ with transition probabilities

$$P_\beta(\sigma,\sigma') := \begin{cases} c(\sigma,\sigma')e^{-\beta[H(\sigma')-H(\sigma)]^+}, & \text{if } \sigma \neq \sigma', \\ 1 - \sum_{\eta\neq\sigma} P_\beta(\sigma,\eta), & \text{if } \sigma = \sigma', \end{cases}$$

where the energy function $H : \mathcal{X} \to \mathbb{R}$ and the connectivity function $c : \mathcal{X} \times \mathcal{X} \to [0,1]$ are the same as those defined in Chapter 5 for grid graphs,

see definitions (5.1) and (5.2), respectively. More precisely, the energy $H(\sigma)$ of a configuration $\sigma \in \mathcal{X}$ is proportional to the total number of particles in σ,

$$H(\sigma) := -\sum_{v \in V} \sigma(v), \tag{6.2}$$

while the connectivity function is given by

$$c(\sigma, \sigma') := \begin{cases} \frac{1}{N}, & \text{if } |\{v \in V \mid \sigma(v) \neq \sigma'(v)\}| = 1, \\ 0, & \text{if } |\{v \in V \mid \sigma(v) \neq \sigma'(v)\}| > 1, \\ 1 - \sum_{\eta \neq \sigma} c(\sigma, \eta), & \text{if } \sigma = \sigma'. \end{cases}$$

Recall that the parameter β is a positive number that represents the inverse temperature of the interacting particle system. We refer to Section 4.1 for more details and properties of Metropolis Markov chains.

Let $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ be the admissible configurations on $\mathcal{T}_{K,L}$ defined as

$$\mathbf{a}(v) := \mathbb{1}_{\{v \in V_a\}}(v), \quad \mathbf{b}(v) := \mathbb{1}_{\{v \in V_b\}}(v), \quad \text{and} \quad \mathbf{c}(v) := \mathbb{1}_{\{v \in V_c\}}(v).$$

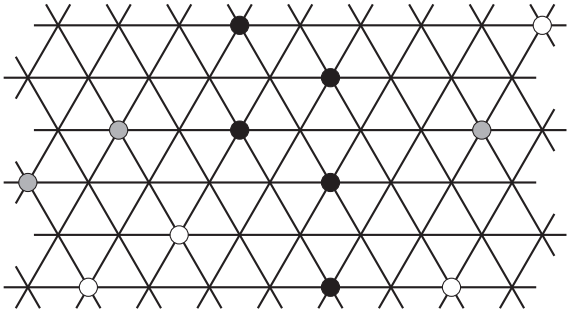
In view of (6.1) each of these three configurations has $2KL$ particles and hence

$$H(\mathbf{a}) = H(\mathbf{b}) = H(\mathbf{c}) = -2KL. \tag{6.3}$$

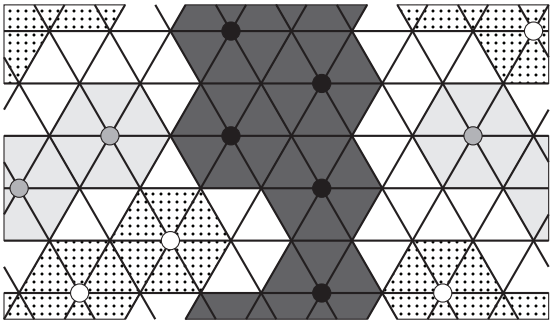
Later we will show that all other admissible configurations on $\mathcal{T}_{K,L}$ have fewer particles, proving that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are the only three stable states of the energy landscape $(\mathcal{X}, H, c)$.

In the rest of the chapter, we adopt the following coloring conventions for displaying a configuration $\sigma \in \mathcal{X}$: We put a node in site $v \in V$ if it is occupied, i.e. $\sigma(v) = 1$, and we color it gray, black or white depending on whether the site v belongs to V_a, V_b, V_c respectively; if instead a site $v \in V$ is vacant, i.e. $\sigma(v) = 0$, we do not display any node there. Figure 6.2a displays an admissible configuration using this color convention.

There is an equivalent way to represent admissible configurations. First notice that each triangle can have at most one occupied particle in its three vertices, being a clique. We adopt the following coloring convention for triangles: If a triangle has a particle in one of its three vertices, we color it gray, black or with a dashed pattern, depending on whether such particle belongs to V_a, V_b or V_c, respectively. Instead, we leave a triangle blank if none of its three vertices is occupied by a particle. Figure 6.2b displays the coloring corresponding to the admissible configuration of Figure 6.2a. Hence, placing particles with hard-core constraints on a triangular grid graph corresponds to placing hexagons on the same lattice, with the rule that hexagons of different colors cannot share an edge. This is the reason why the hard-core model on the triangular lattice is often called *hard-hexagon model* in the statistical physics literature.



(a) The occupied particles of σ colored according to the components to which they belong



(b) Colored areas visualize the blocked sites

Figure 6.2: An example of an admissible configuration σ on the triangular grid $\mathcal{T}_{3,3}$

In [61] the authors studied instead other local dynamics (Glauber dynamics) on triangular grid graphs, showing that the mixing time is slow for large values of the fugacity λ (which is equivalent to the low-temperature regime in view of the relation $\lambda = e^{\beta}$), since it scales exponentially with the triangular grid graph dimensions. Furthermore, in the statistical physics literature the hard-hexagon model has been studied on the infinite triangular lattice in [5, 9, 64], investigating mainly its phase transitions.

6.2 ASYMPTOTIC BEHAVIOR OF TUNNELING TIMES AND MIXING TIMES

In this section we present our main results, which describe the asymptotic behavior of the hard-core model on triangular grid graphs in the low-temperature regime.

The first main result describes the asymptotic behavior of the tunneling times $\tau_{\mathbf{b}}^{\mathbf{a}}$ and $\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}$ for the triangular grid graph $\mathcal{T}_{K,L}$ in the low-temperature regime $\beta \to \infty$. More specifically, we prove the existence and find the value of an exponent $\Gamma(\mathcal{T}_{K,L}) > 0$ that appears in the asymptotic upper and lower bounds in probability for these hitting times and characterizes the asymptotic

order-of-magnitude of their expected values as $\beta \to \infty$. Furthermore, we show that the tunneling times $\tau_{\mathbf{b}}^{\mathbf{a}}$ and $\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}$ normalized by their respective means converge in distribution to an exponential unit-mean random variable.

Theorem 6.2.1 (Asymptotic behavior of tunneling times for $\mathcal{T}_{K,L}$)**.** *Consider the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ corresponding to the hard-core dynamics on the triangular lattice $\mathcal{T}_{K,L}$ and define*

$$\Gamma(\mathcal{T}_{K,L}) := \min\{K, 2L\} + 1. \tag{6.4}$$

Then,

(i) $\lim_{\beta\to\infty} \mathbb{P}_\beta\left(e^{\beta(\Gamma(\mathcal{T}_{K,L})-\epsilon)} \le \tau_{\mathbf{b}}^{\mathbf{a}} \le \tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}} \le e^{\beta(\Gamma(\mathcal{T}_{K,L})+\epsilon)}\right) = 1;$

(ii) $\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}} = \Gamma(\mathcal{T}_{K,L}) = \lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}};$

(iii) $\frac{\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}}{\mathbb{E}\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \beta \to \infty;$

(iv) $\frac{\tau_{\mathbf{b}}^{\mathbf{a}}}{\mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}}} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \beta \to \infty.$

The proof of this theorem leverages the model-independent results presented in Chapter 4 in combination with the analysis of the energy landscape corresponding to the hard-core model on the triangular grid $\mathcal{T}_{K,L}$, to which Section 6.3 is devoted.

The same analysis also yields the following result for the mixing time, which turns out to be asymptotically of the same order of magnitude as the tunneling time between stable states, as established by the next theorem, which is our second main result. Recall that t_β^{mix} denotes the mixing time of the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ and ρ_β denotes its spectral gap, see definitions (4.56) and (4.57).

Theorem 6.2.2 (Mixing time and spectral gap)**.** *Consider the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ corresponding to the hard-core dynamics on the triangular lattice $\mathcal{T}_{K,L}$ and define $\Gamma(\mathcal{T}_{K,L})$ as in (6.4). Then, for any $0 < \epsilon < 1$,*

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log t_\beta^{\mathrm{mix}}(\epsilon) = \Gamma(\mathcal{T}_{K,L}) = \lim_{\beta\to\infty} -\frac{1}{\beta} \log \rho_\beta. \tag{6.5}$$

Furthermore, there exist two positive constants $0 < c_1 \le c_2 < \infty$ independent of β such that

$$c_1 e^{-\beta\Gamma(\mathcal{T}_{K,L})} \le \rho_\beta \le c_2 e^{-\beta\Gamma(\mathcal{T}_{K,L})} \qquad \forall\, \beta \ge 0. \tag{6.6}$$

The proofs of both theorems are presented in Subsection 6.3.3.

6.3 ENERGY LANDSCAPE ANALYSIS

This section is devoted to the analysis of the energy landscape associated with the hard-core dynamics on the triangular grid graph $\mathcal{T}_{K,L}$. This analysis, combined with the model-independent results presented in Chapter 4, yields Theorems 6.2.1 and 6.2.2, as proved in Subsection 6.3.3.

Using an approach similar to the one used in Chapter 5 for grid graphs, we study geometrical features of the admissible configurations on triangular grid graphs and leverage them to prove crucial properties of the energy landscape, as stated in Theorem 6.3.1. More specifically, we will prove that **a**, **b** and **c** are the only three stable states and find the value of the communication height between them, as a function of the dimensions K and L. Furthermore, we will show by means of two reduction algorithms that the energy landscape corresponding the hard-core dynamics on $\mathcal{T}_{K,L}$ exhibits absence of deep cycles (see condition (4.50) in Chapter 4).

In the rest of this chapter, we use the some notions and notation introduced in Chapter 4. In particular, recall that the communication height between a pair of configurations $\sigma, \sigma' \in \mathcal{X}$ has been defined as

$$\Phi(\sigma, \sigma') := \min_{\omega:\sigma\to\sigma'} \Phi_\omega = \min_{\omega:\sigma\to\sigma'} \max_{i=1,\dots,|\omega|} H(\omega_i), \tag{6.7}$$

while the communication height between any pair of disjoint subsets $A, B \subset \mathcal{X}$ is

$$\Phi(A, B) := \min_{\sigma\in A,\, \sigma'\in B} \Phi(\sigma, \sigma'). \tag{6.8}$$

Theorem 6.3.1 (Structural properties of the energy landscape)**.** *Let $(\mathcal{X}, H, q)$ be the energy landscape corresponding to the hard-core dynamics on the triangular grid graph $\mathcal{T}_{K,L}$. Then*

(i) $\mathcal{X}^s = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

(ii) $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) = \Phi(\mathbf{a}, \mathbf{c}) - H(\mathbf{a}) = \Phi(\mathbf{b}, \mathbf{c}) - H(\mathbf{b}) = \min\{K, 2L\} + 1$.

(iii) $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) - H(\sigma) \le \min\{K, 2L\} \quad \forall\, \sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

This latter theorem motivates the definition (6.4) of $\Gamma(\mathcal{T}_{K,L})$ in Theorem 6.2.1.

The rest of this section is devoted to the proof of Theorem 6.3.1 and we now briefly outline the proof strategy. In Subsection 6.3.1, after some preliminary definitions, we study the combinatorial properties of admissible configurations on horizontal and vertical stripes of the triangular grid $\mathcal{T}_{K,L}$, i.e. pairs of adjacent rows (triplets of adjacent columns, respectively). In particular, we find the maximum number of particles that an admissible configuration can have in a horizontal stripe and characterize in Lemma 6.3.3 how particles are arranged on such stripes. Theorem 6.3.1(i) is an almost immediate consequence of these combinatorial results. Afterwards, using geometrical proper-

ties of the admissible configurations, we prove Proposition 6.3.5, which gives the following lower bound for the communication height between **a** and **b**:

$$\Phi(\mathbf{a},\mathbf{b}) - H(\mathbf{a}) \geq \min\{K, 2L\} + 1. \tag{6.9}$$

We then introduce two *reduction algorithms* in Subsection 6.3.2, which are used in Proposition 6.3.6 to build a reference path $\omega^* : \mathbf{a} \to \mathbf{b}$. Such a path shows that the lower bound (6.9) is sharp, so that

$$\Phi(\mathbf{a},\mathbf{b}) - H(\mathbf{a}) = \min\{K, 2L\} + 1,$$

concluding the proof of Theorem 6.3.1(ii). The reduction algorithm is then used again to construct a path from every configuration $\sigma \neq \mathbf{a}, \mathbf{b}, \mathbf{c}$ to one of the three states $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, proving in this way Theorem 6.3.1(iii).

6.3.1 *Geometrical properties of admissible configurations*

In this subsection we first introduce some useful definitions and then prove some geometrical and combinatorial properties of admissible configurations.

We denote by c_j, $j = 0, \ldots, 6L-1$, the j-th column of $\mathcal{T}_{K,L}$, and by r_i, $i = 0, \ldots, 2K-1$, the i-th row of $\mathcal{T}_{K,L}$, see Figure 6.3.

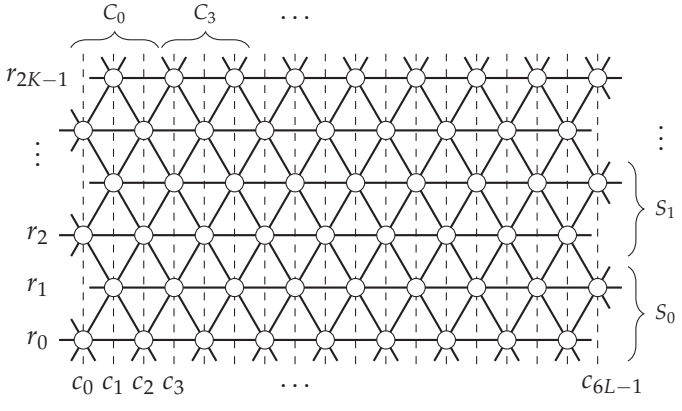


Figure 6.3: Illustration of row, column and stripe notation for the triangular grid

Note that every row has an equal number of sites from each component, since

$$|r_i \cap V_a| = |r_i \cap V_b| = |r_i \cap V_c| = L \quad \forall\, i = 0, \ldots, 2K-1, \tag{6.10}$$

while each column consists of sites from a single component, since

$$V_a = \bigcup_{j=0}^{L-1} c_{3j}, \quad V_b = \bigcup_{j=0}^{L-1} c_{3j+1}, \quad \text{and} \quad V_c = \bigcup_{j=0}^{L-1} c_{3j+2}. \tag{6.11}$$

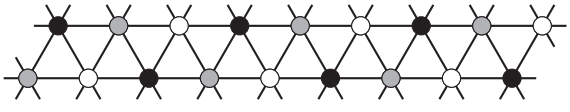
Each site v of $\mathcal{T}_{K,L}$ lies at the intersection of a row with a column and we associate to v the coordinates (i, j) if $v = r_i \cap c_j$. We call the collection of sites belonging to two adjacent rows a *horizontal stripe*. In particular, we denote by S_i, with $i = 0, \dots, K-1$, the horizontal stripe consisting of rows r_{2i} and r_{2i+1}, i.e. $S_i := r_{2i} \cup r_{2i+1}$, see Figure 6.3. When the index of a stripe is not relevant, we will simply denote it by S. We define a *vertical stripe* to be the collection of sites belonging to three adjacent columns, which we denote by C in general. In particular, for $j = 0, \dots, 3L-1$ we denote by C_j the vertical stripe consisting of columns c_j, c_{j+1} and c_{j+2} (note that the indices should be taken modulo $3L$), see Figure 6.3. For every horizontal stripe S note that $|S| = 6L$ and (6.10) implies that

$$|S \cap V_a| = |S \cap V_b| = |S \cap V_c| = 2L,$$

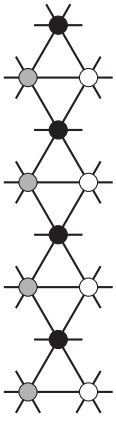
see also Figure 6.4a where we highlight the tripartition of a horizontal stripe. Similarly, for every vertical stripe C, we have $|C| = 3K$ and, in view of (6.11), we have

$$|C \cap V_a| = |C \cap V_b| = |C \cap V_c| = K.$$

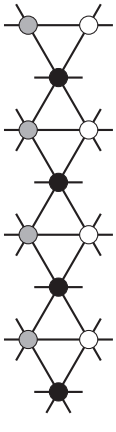
A special role will be played by the vertical stripes whose middle column belongs to V_b, which are those of the form C_{3j} for some $j = 0, \dots, 2L-1$, whose structure is displayed in Figures 6.4b and 6.4c.



(a) A horizontal stripe S_i of $\mathcal{T}_{K,3}$



(b) A vertical stripe C_{3j} of $\mathcal{T}_{4,L}$ for j even



(c) A vertical stripe C_{3j} of $\mathcal{T}_{4,L}$ for j odd

Figure 6.4: Illustration of horizontal and vertical stripes

Given a feasible configuration $\sigma \in \mathcal{X}$, we define its *energy wastage* $U(\sigma)$ as the energy difference between σ and configuration $\mathbf{a}$, i.e.

$$U(\sigma) := H(\sigma) - H(\mathbf{a}). \tag{6.12}$$

In view of the definition (6.2) of the energy function $H(\cdot)$ and identity (6.3), the energy wastage of configuration σ can be rewritten as

$$U(\sigma) = 2KL - \sum_{v \in V} \sigma(v).$$

Furthermore, we define the energy wastage of a configuration $\sigma \in \mathcal{X}$ on the horizontal stripe S by

$$U^S(\sigma) := 2L - \sum_{v \in S} \sigma(v), \tag{6.13}$$

and the energy wastage of a configuration $\sigma \in \mathcal{X}$ on the vertical stripe C by

$$U^C(\sigma) := K - \sum_{v \in C} \sigma(v).$$

The energy wastage (6.12) of a configuration σ can be written as the sum of the energy wastages of horizontal stripes (or of non-overlapping vertical stripes), i.e.

$$U(\sigma) = \sum_{i=0}^{K-1} U^{S_i}(\sigma) = \sum_{j=0}^{2L-1} U^{C_{3j}}(\sigma) = \sum_{j=0}^{2L-1} U^{C_{3j+1}}(\sigma) = \sum_{j=0}^{2L-1} U^{C_{3j+2}}(\sigma). \tag{6.14}$$

Given two configurations $\sigma, \sigma' \in \mathcal{X}$ and a subset of sites $W \subseteq V$, we write

$$\sigma_{|W} = \sigma'_{|W} \quad \Longleftrightarrow \quad \sigma(v) = \sigma'(v) \quad \forall\, v \in W.$$

We say that a configuration $\sigma \in \mathcal{X}$ has a *gray (black, white) horizontal bridge* in stripe S if σ perfectly agrees there with $\mathbf{a}$ (respectively $\mathbf{b}$, $\mathbf{c}$), i.e.

$$\sigma_{|S} = \mathbf{a}_{|S} \quad (\text{respectively } \sigma_{|S} = \mathbf{b}_{|S} \text{ or } \sigma_{|S} = \mathbf{c}_{|S}).$$

Similarly, we say that a configuration $\sigma \in \mathcal{X}$ has a *gray (black, white) vertical bridge* in stripe C if σ perfectly agrees there with $\mathbf{a}$ (respectively $\mathbf{b}$, $\mathbf{c}$), i.e.

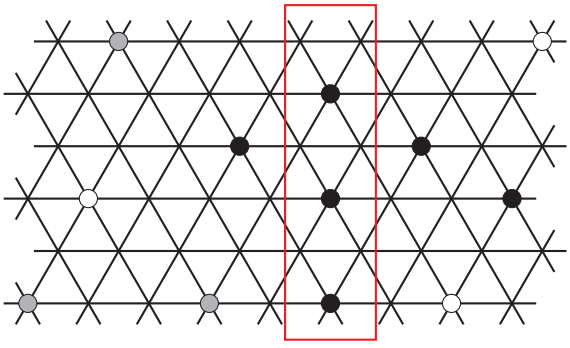
$$\sigma_{|C} = \mathbf{a}_{|C} \quad (\text{respectively } \sigma_{|C} = \mathbf{b}_{|C} \text{ or } \sigma_{|C} = \mathbf{c}_{|C}).$$

Two examples of bridges are shown in Figures 6.5a and 6.5b.

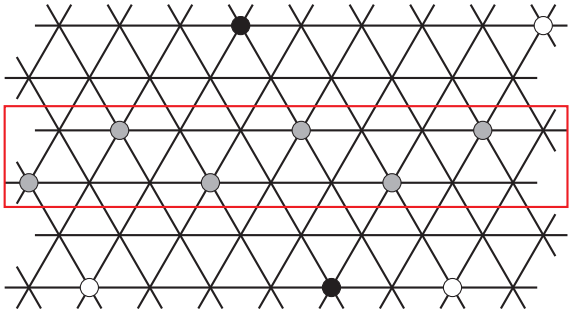
Lemma 6.3.2 (Geometric features of admissible configurations)**.** *An admissible configuration $\sigma \in \mathcal{X}$ cannot display simultaneously a vertical bridge and a horizontal bridge of different colors.*

Proof. We may assume that the vertical bridge is black, without loss of generality. Such a vertical bridge blocks two sites on every row, one belonging to V_a and one to V_c and thus no horizontal stripe can fully agree with $\mathbf{a}$ or $\mathbf{c}$. □

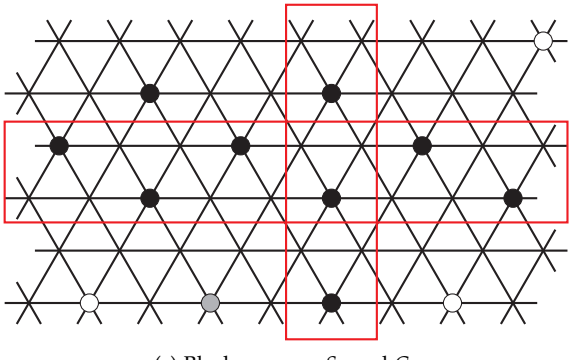
It is possible, however, that a vertical and a horizontal bridge coexist when they are of the same color and this fact motivates the next definition. We say that a configuration $\sigma \in \mathcal{X}$ has a *gray (black, white) cross* if it has simultaneously two gray (black, white) bridges, one vertical and one horizontal; see Figure 6.5c for an example of a black cross.



(a) Black vertical bridge in C_9



(b) Gray horizontal bridge in S_1



(c) Black cross on S_1 and C_9

Figure 6.5: Examples of admissible configurations on $\mathcal{T}_{3,3}$ displaying bridges or crosses

Lemma 6.3.3 (Efficient horizontal stripes structure)**.** *Consider an admissible configuration $\sigma \in \mathcal{X}$ on the triangular grid graph $\mathcal{T}_{K,L}$. For every horizontal stripe S, the energy wastage is non-negative, i.e. $U^S(\sigma) \geq 0$, and furthermore*

$$U^S(\sigma) = 0 \quad \Longleftrightarrow \quad \sigma \text{ has a horizontal bridge in stripe } S. \tag{6.15}$$

Proof. We first prove that the energy wastage $U^S(\sigma) \geq 0$ and later prove (6.15).

Suppose by contradiction that there exists an admissible configuration σ and a horizontal stripe $S = r \cup r'$ such that $U^S(\sigma) < 0$. Let p and p' the number of particles that σ has in rows r and r', respectively. By definition (6.13), this implies that σ has at least $2L+1$ particles in stripe S, which means that one of the two rows of S, say r, must have at least $L+1$ particles, i.e. $p \geq L+1$. Due to the structure of a horizontal stripe, each particle on r blocks two sites in row r' and no site in row r' is blocked simultaneously by two particles in row r, since to do so such particles should reside in adjacent sites. Hence, there are at least $2p$ blocked sites, so that in row r' there can be at most $3L - 2p$ particles. In view of the assumption $p \geq L+1$, this means that there are at most $p + (3L - 2p) = 3L - p \leq 2L - 1$ particles in stripe S, and thus $U^S(\sigma) > 0$, which is a contradiction.

Let us turn to the characterization (6.15) of the stripes with energy wastage equal to zero. If σ displays a bridge in a horizontal stripe S, then by definition $\sigma_{|S} = \mathbf{a}_{|S}$ or $\sigma_{|S} = \mathbf{b}_{|S}$ or $\sigma_{|S} = \mathbf{c}_{|S}$. In all three cases, we have $\sum_{v \in S} \sigma(v) = 2L$ and thus $U^S(\sigma) = 2L - \sum_{v \in S} \sigma(v) = 0$.

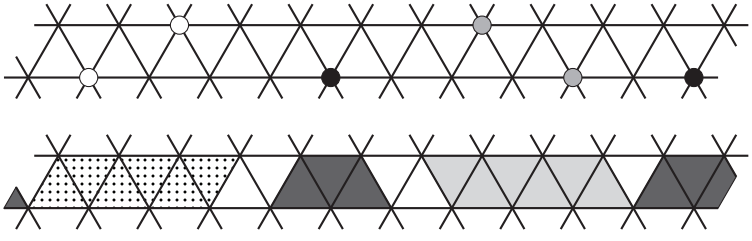


Figure 6.6: Equivalent representation of an admissible configuration in a horizontal stripe of $\mathcal{T}_{m,4}$

For the converse implication, we leverage the equivalent representation of an admissible configuration as hard hexagon introduced at the end of Section 6.1. Note that in a horizontal stripe, we see only half of any hexagon: This means that any particle blocks exactly 3 triangles on that stripe. We distinguish between blocked and free triangles and color the blocked triangles according to which particle blocks them. Denote by $b(S)$ the number of blocked triangles in stripe S and by $f(S)$ the number of free triangles on the same stripe. Since the total number of triangles of stripe S is $6L$, we have

$b(S) + f(S) = 6L$. Furthermore, as argued before, every particle blocks exactly 3 triangles on that stripe, i.e. $b(S) = 3\sum_{v\in S}\sigma(v)$. Therefore,

$$f(S) = 6L - b(S) = 3\left(2L - \sum_{v\in S}\sigma(v)\right) = 3\cdot U^S(\sigma). \tag{6.16}$$

Suppose now by contradiction that σ does not have a horizontal bridge in stripe S. If σ has only particles of the same color on S, then the number of particles must be strictly less than $2L$ and thus $U^S(\sigma) > 0$. If instead σ has at least a pair of particles belonging to different components, then there exist blocked triangles of different colors. The hard-hexagon representation implies that two blocked triangles of different color cannot be adjacent, so there must be at least one free triangle separating them, i.e. $f(S) > 0$, which implies that $U^S(\sigma) > 0$ in view of (6.16). □

Proof of Theorem 6.3.1(i). Lemma 6.3.3 guarantees that $U^S(\sigma) \geq 0$ on every horizontal stripe S. By definition (6.13), this means that on every horizontal stripe there can be at most $2L$ particles. Hence

$$\max_{\sigma\in\mathcal{X}}\sum_{v\in V}\sigma(v) \leq 2KL.$$

Configurations $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ have precisely $2KL$ particles each, in view of (6.1). Suppose by contradiction that there exists another configuration $\sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ that has $2KL$ particles. By definitions (6.12) and (6.13) such configuration should be such that $U^{S_i}(\sigma) = 0$ on every horizontal stripe S_i, $i = 0, \ldots, K-1$. By Lemma 6.3.3 configuration σ then has a horizontal bridge in each of these K stripes. Since two horizontal bridges of different colors cannot reside in adjacent stripes, it follows that all K bridges should be monochromatic and thus $\sigma \in \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, which is a contradiction. □

In order to prove Theorem 6.3.1(ii), we need the following lemma for vertical stripes, which is a complementary result of Lemma 6.3.3 characterizing the structure of vertical stripes with zero energy wastage.

Lemma 6.3.4 (Efficient vertical stripes structure)**.** *Consider an admissible configuration $\sigma \in \mathcal{X}$ on the triangular grid graph $\mathcal{T}_{K,L}$. For every vertical stripe C of the form $C = C_{3j}$, the energy wastage is non-negative, i.e. $U^C(\sigma) \geq 0$. Furthermore, if σ has at least one black particle on C, then*

$$U^C(\sigma) = 0 \quad \Longleftrightarrow \quad \sigma \text{ has a black vertical bridge in stripe } C. \tag{6.17}$$

Proof. Consider a vertical stripe C of the form C_{3j}, whose middle column is a subset of V_b. We first prove that the energy wastage $U^C(\sigma) \geq 0$ and later prove (6.17). Observe that, due to the structure of the vertical stripe C, an admissible configuration cannot have more than one particle in every row in C and, moreover, it cannot have particles residing in adjacent rows in C. From

these two observations, it follows that there are no admissible configurations with $K+1$ or more particles on the vertical stripe C and thus $U^C(\sigma) \geq 0$.

If σ displays a vertical black bridge in vertical stripe C, then by definition $\sigma_{|C} = \mathbf{b}_{|C}$ and thus trivially $U^C(\sigma) = 2L - \sum_{v \in C} \sigma(v) = 0$. Similarly to what we have done in the proof of Lemma 6.3.3, we prove the converse implication using the equivalent representation of an admissible configuration as hard hexagons mentioned in Section 6.1. Note that in a vertical stripe, we see only part of these hexagons; in particular, any particle blocks exactly 2 triangles on that stripe, as illustrated in Figure 6.7. We distinguish between blocked and free triangles and color the blocked triangles according to which particle blocks them. A key observation is that in such a vertical stripe, dashed and gray triangles can be adjacent to each other, as shown in Figure 6.7, but a black triangle cannot be adjacent to a triangle of different color. Denote by $b(C)$ the number of blocked triangles in stripe C and by $f(C)$ the number of free triangles on the same stripe. Since the total number of triangles in stripe C is $2K$, we have $b(C) + f(C) = 2K$. Furthermore, since every particle blocks exactly 2 triangles on that vertical stripe, we have $b(C) = 2\sum_{v \in C} \sigma(v)$. Therefore,

$$f(C) = 2K - b(C) = 2\left(K - \sum_{v \in C} \sigma(v)\right) = 2 \cdot U^C(\sigma). \tag{6.18}$$

Suppose by contradiction that σ does not have a vertical black bridge in stripe C. If σ has only black particles on C, then the number of black particles must

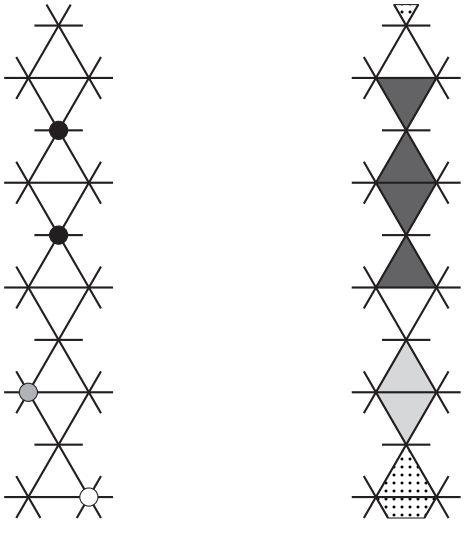


Figure 6.7: Equivalent representation of an admissible configuration in a vertical stripe of $\mathcal{T}_{5,L}$

be strictly less than $2K$ and thus $U^C(\sigma) > 0$. Instead, if σ has also at least one non-black particle in stripe C, then such particle blocks two triangles, which will then be non-black. Such triangles cannot be adjacent to the two black ones which exist by assumption, so there exists at least one free triangle separating them, i.e. $f(C) > 0$, which implies that $U^C(\sigma) > 0$ in view of (6.18). □

We are now ready to state and prove the lower bound on the communication height between any pair of stable states.

Proposition 6.3.5 (Lower bound on the communication height between **a** and **b**). *Consider the triangular grid graph $\mathcal{T}_{K,L}$. The communication height between* **a** *and* **b** *in the corresponding energy landscape satisfies*

$$\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) \geq \min\{K, 2L\} + 1. \tag{6.19}$$

Proof. We will show that in every path $\omega : \mathbf{a} \to \mathbf{b}$ there exists at least one configuration with energy wastage greater than or equal to $\min\{K, 2L\} + 1$. Consider a path $\omega = (\omega_1, \dots, \omega_n)$ from **a** to **b**. Without loss of generality, we may assume that there are no void moves in ω, i.e. at every step either a particle is added or a particle is removed, so that $H(\omega_{i+1}) = H(\omega_i) \pm 1$ for every $1 \leq i \leq n-1$. Since configuration **a** has no black bridges, while **b** does, at some point along the path ω there must be a configuration which is the first to display a black bridge, that is a column or a row occupied only by black particles. Denote by m^* the index corresponding to such configuration, i.e.

$$m^* := \{m \leq n \mid \exists\, i \,:\, (\omega_m)_{|r_i} = \mathbf{b}_{|r_i} \quad \text{or} \quad \exists\, j \,:\, (\omega_m)_{|c_j} = \mathbf{b}_{|c_j}\}.$$

Since a black bridge cannot be created in only two steps starting from **a**, we must have $m^* > 2$. We claim that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq \min\{K, 2L\} + 1.$$

Since the addition of a single black particle cannot create more than one bridge in each direction, it is enough to consider the following three cases:

(a) ω_{m^*} displays a black vertical bridge only;

(b) ω_{m^*} displays a black horizontal bridge only;

(c) ω_{m^*} displays a black cross.

For case (a), note that configuration ω_{m^*} does not have any horizontal bridge. Indeed, it cannot have a black horizontal bridge due to the definition of m^* and any gray or white horizontal bridge cannot coexist with the black vertical bridge, in view of Lemma 6.3.2. Hence, the energy wastage of every horizontal stripe is strictly positive, thanks to Lemma 6.3.3 and thus

$$U(\omega_{m^*}) = \sum_{i=0}^{K-1} U^{S_i}(\omega_{m^*}) \geq K.$$

Furthermore, configurations ω_{m^*-1} and ω_{m^*} differ in a unique site $v^* \in V_b$, which is such that $\omega_{m^*-1}(v^*) = 0$ and $\omega_{m^*}(v^*) = 1$. Hence, $U(\omega_{m^*-1}) = U(\omega_{m^*}) + 1$ and thus

$$U(\omega_{m^*-1}) \geq K + 1.$$

The argument for case (b) is similar to that of case (a). Firstly, configuration ω_{m^*} does not display any vertical bridge. Lemma 6.3.2 implies that there cannot be any gray or white vertical bridge due to the presence of a black horizontal bridge and a black vertical bridge cannot exist, since it would contradict the definition of m^*. Every vertical stripe has at least one black particle, due to the presence of a black horizontal bridge. Hence, $U^{C_j}(\omega_{m^*}) \geq 1$ for every $j = 0, \ldots, 2L-1$ in view of Lemma 6.3.4. Therefore,

$$U(\omega_{m^*}) = \sum_{j=0}^{2L-1} U^{C_j}(\omega_{m^*}) \geq 2L.$$

From this inequality it follows that $U(\omega_{m^*-1}) \geq 2L + 1$, because, as for case (a), the definition of m^* implies $U(\omega_{m^*-1}) = U(\omega_{m^*}) + 1$.

Consider now case (c), in which ω_{m^*} displays a black cross. The presence of both a vertical and a horizontal black bridge means that ω_{m^*} has a black particle in every vertical and horizontal stripe. This property is inherited by the configuration ω_{m^*-1}, since it differs from ω_{m^*} only by the removal of the black particle lying at the intersection of the vertical and horizontal bridge constituting the cross. Furthermore, by definition of m^*, configuration ω_{m^*-1} cannot have any black bridge, neither vertical nor horizontal. These two facts, in combination with Lemmas 6.3.3 and 6.3.4, imply that

$$U(\omega_{m^*-1}) \geq \min\{K, 2L\}.$$

If $U(\omega_{m^*-1}) \geq \min\{K, 2L\} + 1$, then the proof is completed. Otherwise, the energy wastage of configuration ω_{m*-1} is $U(\omega_{m^*-1}) = \min\{K, 2L\}$. The configuration preceding ω_{m^*-1} in the path ω satisfies

$$U(\omega_{m^*-2}) = \min\{K, 2L\} \pm 1, \tag{6.20}$$

since it differs from ω_{m^*-1} by a single site update. Suppose first that

$$U(\omega_{m^*-2}) = \min\{K, 2L\} - 1. \tag{6.21}$$

This means that ω_{m^*-2} differs from ω_{m^*-1} by the addition of a particle. Therefore, also configuration ω_{m^*-2} has at least one black particle in every horizontal stripe, i.e.

$$\sum_{v \in S_i \cap V_b} \omega_{m^*-2}(v) \geq 1 \quad \forall\, i = 0, \ldots, K, \tag{6.22}$$

and at least one black particle in every vertical stripe, i.e.

$$\sum_{v \in C_j \cap V_b} \omega_{m^*-2}(v) \geq 1 \quad \forall\, j = 0, \dots, 2L-1. \tag{6.23}$$

If $K \leq 2L$, (6.21) and the pigeonhole principle imply that there must be a horizontal stripe S such that $U^S(\omega_{m^*-2}) = 0$. In view of (6.22) and Lemma 6.3.3, ω_{m^*-2} must have a horizontal black bridge in S, which contradicts the definition of m^*. When instead $K > 2L$, it follows from (6.21) that there must be a vertical stripe C such that $U^C(\omega_{m^*-2}) = 0$. Also in this case, (6.23) and Lemma 6.3.4 imply that ω_{m^*-2} displays a vertical black bridge in C, in contradiction with the definition of m^*. We have in this way proved that assumption (6.21) always leads to a contradiction, so in view of (6.20) we have $U(\omega_{m^*-2}) = \min\{K, 2L\} + 1$ and the proof is concluded also for case (c). □

6.3.2 *Reference path and absence of deep cycles*

In this subsection we describe two iterative procedures that build a path from a suitable initial configuration σ to a target stable state. The first one modifies the original configuration row by row and we will refer to it as *reduction algorithm by rows*. The second one proceeds instead by modifying σ column by column and, for this reason, we will call it *reduction algorithm by columns*. These two procedures are similar in spirit, but need to be described separately, since for the triangular grid the structure of horizontal stripes is fundamentally different from that of vertical stripes. Such reduction algorithms will be used in Proposition 6.3.6 to build the reference path from **a** to **b** and in the proof of Theorem 6.3.1(iii).

Reduction algorithm by rows

We now describe in detail the reduction algorithm by rows with **b** as target state. In order for $\sigma \in \mathcal{X}$ to be a suitable initial configuration for the reduction algorithm by rows, we require that σ has no gray or white particles in the first horizontal stripe S_0, i.e.

$$\sum_{v \in S_0 \cap (V_a \cup V_c)} \sigma(v) = 0. \tag{6.24}$$

Figure 6.8 shows an admissible configuration that satisfies this initial condition.

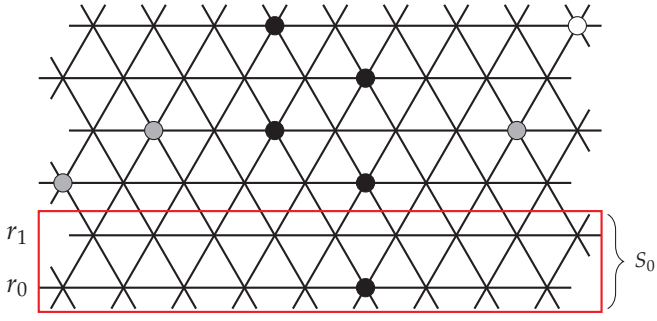


Figure 6.8: Example of an admissible configuration on $\mathcal{T}_{3,3}$ that satisfies (6.24)

The path ω is the concatenation of $2K$ paths $\omega^{(1)}, \dots, \omega^{(2K)}$. For every $i = 1, \dots, 2K$, path $\omega^{(i)}$ goes from σ_i to σ_{i+1}, where we set $\sigma_1 = \sigma$ and recursively define for $i = 1, \dots, 2K$

$$\sigma_{i+1}(v) := \begin{cases} \sigma_i(v) & \text{if } v \in V \setminus (r_i \cup r_{i+1}), \\ \mathbf{b}(v) & \text{if } v \in r_i, \\ \sigma_i(v) & \text{if } v \in r_{i+1} \cap V_b, \\ 0 & \text{if } v \in r_{i+1} \cap (V_a \cup V_c). \end{cases}$$

Clearly, due to the periodic boundary conditions, the row index should be taken modulo $2K$. It can be checked that indeed $\sigma_{2K+1} = \mathbf{b}$. We now describe in detail how to construct each of the paths $\omega^{(i)}$ for $i = 1, \dots, 2K$. We build a path $\omega^{(i)} = (\omega_0^{(i)}, \omega_1^{(i)}, \dots, \omega_{2L+1}^{(i)})$ of length $2L+1$ (but possibly comprising void moves), with $\omega_0^{(i)} = \sigma_i$ and $\omega_{2L+1}^{(i)} = \sigma_{i+1}$. We start from configuration $\omega_0^{(i)} = \sigma_i$ and we repeat iteratively the following procedure for all $j = 0, \dots, 2L-1$:

- If $j \equiv 0 \pmod 2$, consider the sites $v \in V_a$ and $v' \in V_c$ defined by

$$\begin{cases} v = (i+1, 3j), \ v' = (i+1, 3j+2) & \text{if } i \equiv 0 \pmod 2, \\ v = (i+1, 3j+1), \ v' = (i+1, 3j+3) & \text{if } i \equiv 1 \pmod 2. \end{cases}$$

  Note that the two sites v and v' are always neighbors, so that only one of the two can be occupied.
  - If $\omega_j^{(i)}(v) = 0 = \omega_j^{(i)}(v')$, we set $\omega_{j+1}^{(i)} = \omega_j^{(i)}$, so $H(\omega_{j+1}^{(i)}) = H(\omega_j^{(i)})$.
  - If $\omega_j^{(i)}(v) = 1$ or $\omega_j^{(i)}(v') = 1$, then we remove from configuration $\omega_j^{(i)}$ the particle in the unique occupied site between v and v', increasing the energy by 1 and obtaining in this way configuration $\omega_{j+1}^{(i)}$, which is such that $H(\omega_{j+1}^{(i)}) = H(\omega_j^{(i)}) + 1$.

- If $j \equiv 1 \pmod 2$, consider the site $v \in V_b$ defined as

$$v = \begin{cases} (i, 3j-1) & \text{if } i \equiv 0 \pmod 2, \\ (i, 3j-2) & \text{if } i \equiv 1 \pmod 2. \end{cases}$$

  - If $\omega_j^{(i)}(v) = 1$, we set $\omega_{j+1}^{(i)} = \omega_j^{(i)}$ and thus $H(\omega_{j+1}^{(i)}) = H(\omega_j^{(i)})$.
  - If $\omega_j^{(i)}(v) = 0$, then we add to configuration $\omega_j^{(i)}$ a particle in site v decreasing the energy by 1. We obtain in this way a configuration $\omega_{j+1}^{(i)}$, which is admissible because by construction all the first neighboring sites of v are unoccupied. In particular, the two particles residing in the two sites above v may have been removed exactly at the previous step. The new configuration has energy $H(\omega_{j+1}^{(i)}) = H(\omega_j^{(i)}) - 1$.

The way the path $\omega^{(i)}$ is constructed shows that for every $i = 1, \dots, 2K$,

$$H(\sigma_{i+1}) \leq H(\sigma_i),$$

since the number of particles added in row r_i is greater than or equal to the number of particles removed in row r_{i+1}. Moreover,

$$\Phi_{\omega^{(i)}} \leq H(\sigma_i) + 1$$

since along the path $\omega^{(i)}$ every particle removal (if any) is always followed by a particle addition. These two properties imply that the path $\omega : \sigma \to \mathbf{b}$ created by concatenating $\omega^{(1)}, \dots, \omega^{(2K)}$ satisfies

$$\Phi_\omega \leq H(\sigma) + 1.$$

Note that the reduction algorithm by rows can be tweaked in order to have **a** or **c** as target state. In particular, the condition (6.24) for the initial configuration σ should be adjusted accordingly, requiring that σ has no black and white (black and gray, respectively) particles in the first horizontal stripe S_0, depending on whether the target state is **a** or **c**, respectively.

Reduction algorithm by columns

We now illustrate how the reduction algorithm by columns works choosing **b** as target state. Note that the procedure we are about to describe can be tweaked to yield a path with target state **a** or **c**, but we omit the details. If the target state is **b**, we require that the initial configuration $\sigma \in \mathcal{X}$ for the reduction algorithm by columns satisfies the following condition:

$$\sum_{v \in c_2 \cup c_3} \sigma(v) = 0. \tag{6.25}$$

Since $c_2 = C_0 \cap V_c$ and $c_3 = C_1 \cap V_a$, condition (6.25) requires that in the two first vertical stripes C_0 and C_1 of configuration σ there are no white particles in

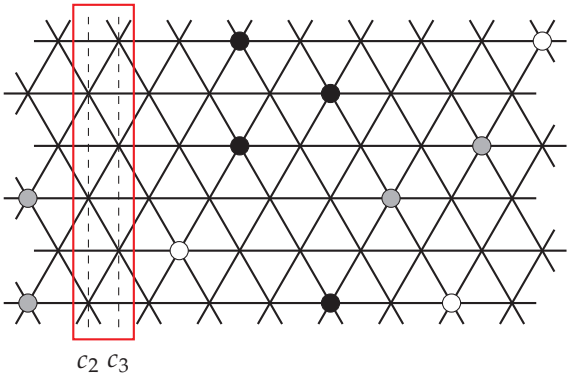


Figure 6.9: Example of admissible configuration on $\mathcal{T}_{3,3}$ that satisfies (6.25)

C_0 and no gray particles in C_1. Figure 6.9 shows an admissible configuration that satisfies this initial condition.

The output of this algorithm is a path ω from σ to $\mathbf{b}$, which we construct as concatenation of L paths $\omega^{(1)}, \dots, \omega^{(L)}$. For every $j = 1, \dots, L$, path $\omega^{(j)}$ goes from σ_j to σ_{j+1}, where we set $\sigma_1 = \sigma$ and recursively define for $j = 1, \dots, L$

$$\sigma_{j+1}(v) := \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_{3j-2} \cup c_{3j-1} \cup c_{3j}), \\ \mathbf{b}(v) & \text{if } v \in c_{3j-2}, \\ 0 & \text{if } v \in c_{3j-1} \cup c_{3j}. \end{cases}$$

The column index should of course be taken modulo $3L$, in view of the periodic boundary conditions. It can be easily checked that $\sigma_{L+1} = \mathbf{b}$. We now describe in detail how to construct each of the paths $\omega^{(j)}$ for $j = 1, \dots, L$. We remark that in the procedures we will now describe, also the row index has to be taken modulo $2K$.

We distinguish two cases, depending on whether (a) σ_j has a vertical bridge in column c_{3j-1} or (b) not.

Consider case (a) first. First notice that the presence of a vertical (white) bridge in column c_{3j-1} implies that all sites of the adjacent column c_{3j} must be empty in configuration σ_j.

We build a path $\omega^{(j)} = (\omega_1^{(j)}, \dots, \omega_{2K+1}^{(j)})$ of length $2K+1$ (but possibly comprising void moves), with $\omega_1^{(j)} = \sigma_j$ and $\omega_{2K+1}^{(j)} = \sigma_{j+1}$. Denote by $o(j) \in \{0,1\}$ the integer number such that $o(j) \equiv j - 1 \pmod 2$. We first remove the two white particles in column c_{3j-1} that lie in row $r_{o(j)}$ and $r_{o(j)+2}$ in two successive steps, obtaining in this way configuration $\omega_3^{(j)}$, which is such that $H(\omega_3^{(j)}) = H(\omega_1^{(j)}) + 2$. We then repeat iteratively the following procedure to obtain the configuration $\omega_{i+1}^{(j)}$ from $\omega_i^{(j)}$ for all $i = 3, \dots, 2K-1$:

- If $i \equiv 1 \pmod 2$, consider the site $v \in c_{3j-2} \subset V_b$ with coordinates $(3j-2, o(j)+i-2)$ and add a (black) particle there, obtaining in this

way configuration $\omega_{i+1}^{(j)}$. Such a particle can be added since all its six neighboring sites are empty. More specifically, the three left ones have been (possibly) emptied along the path $\omega^{(j-1)}$, while the one in c_{3j} is empty by assumption and the other two sites on its right have been emptied in the previous steps of $\omega^{(j)}$. Since we added one particle, $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) - 1$.

- If $i \equiv 0 \pmod 2$, consider the site $v \in c_{3j-1} \subset V_c$ with coordinates $(3j-1, o(j)+i-2)$ and remove the (white) particle lying there, obtaining in this way configuration $\omega_{i+1}^{(j)}$, which is such that $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) + 1$.

This procedure outputs configuration $\omega_{2K}^{(j)}$ which has no white particles in column c_{3j-1} and an empty site in column c_{3j-2}, the one with coordinates $(3j-2, 2K-1-o(j))$. All the neighboring sites of this site are empty by construction and, adding a black particle in this site, we obtain configuration $\omega_{2K+1}^{(j)} = \omega_{j+1}$, which is such that $H(\omega_j) = H(\omega_{2K}^{(j)}) - 1$. The way the path $\omega^{(j)}$ is constructed shows that

$$H(\sigma_{j+1}) = H(\sigma_j),$$

since we added K (black) particles in column c_{3j-2} and removed K (white) particles in columns c_{3j-1}. Moreover,

$$\Phi_{\omega^{(j)}} = \max_{\eta \in \omega^{(j)}} H(\eta) = H(\sigma_j) + 2 \tag{6.26}$$

since along the path $\omega^{(j)}$ every particle removal is followed by a particle addition, except at the beginning where we removed two particles consecutively.

Consider now case (b). We claim that, since there is no vertical (white) bridge in column c_{3j-1}, there exists a site v^* in column c_{3j-2} with at most one neighboring occupied site. First of all, all sites in column 3_{3j-3} and c_{3j-4} have been emptied along the path $\omega^{(j-1)}$, so all sites in c_{3j-2} have no left neighboring sites occupied. Let us look now at the right neighboring sites. Since there is no vertical white bridge in column c_{3j-1}, there exists an empty site in it, say w. Modulo relabeling the rows, we may assume that w has coordinates $(3j-1, o(j))$, where $o(j)$ is the integer in $\{0,1\}$ such that $o(j) \equiv j-1 \pmod 2$. The site $v^* = (3j-2, o(j)+1)$ has then the desired property, since at most one of its two remaining right neighboring sites (those with coordinates $(3j-1, o(j)+2)$ and $(3j, o(j)+1)$, respectively) can be occupied, since they are also neighbors of each other.

We build a path $\omega^{(j)} = (\omega_1^{(j)}, \dots, \omega_{2K+1}^{(j)})$ of length $2K+1$ (but possibly comprising void moves), with $\omega_1^{(j)} = \sigma_j$ and $\omega_{2K+1}^{(j)} = \sigma_{j+1}$. We then repeat iteratively the following procedure to obtain configuration $\omega_{i+1}^{(j)}$ from $\omega_i^{(j)}$ for all $i = 1, \dots, 2K$:

- If $i \equiv 1 \pmod 2$, consider the two sites $(3j-1, o(j)+i+1) \in V_c$ and $(3j, o(j)+i) \in V_a$. Since they are neighboring sites, at most one of them

is occupied. If they are both empty, we set $\omega_{i+1}^{(j)} = \omega_i^{(j)}$. If instead there is a particle in either of the two, we remove it, obtaining in this way configuration $\omega_{i+1}^{(j)}$, which is such that $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) + 1$.

- If $i \equiv 0 \pmod 2$, consider the site $v \in c_{3j-2} \subset V_b$ with coordinates $(3j-2, o(j)+i-1)$ and add a (black) particle there, obtaining in this way configuration $\omega_{i+1}^{(j)}$. Such a particle can be added since all its six neighboring sites are empty. More specifically, the three left ones have been (possibly) emptied along the path $\omega^{(j-1)}$, while the other two sites on its right have been emptied in the previous step of $\omega^{(j)}$. Since we added one particle, $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) - 1$.

The way the path $\omega^{(j)}$ is constructed shows that

$$H(\sigma_{j+1}) \leq H(\sigma_j),$$

since the number of (black) particles added in column c_{3j-2} is greater than or equal to the number of (white or gray) particles removed in columns c_{3j-1} and c_{3j}. Moreover, along the path $\omega^{(j)}$ every particle removal (if any) is always followed by a particle addition, and hence

$$\Phi_{\omega^{(j)}} = \max_{\eta \in \omega^{(j)}} H(\eta) \leq H(\sigma_j) + 1. \tag{6.27}$$

Consider now the path $\omega : \sigma \to \mathbf{b}$ created by concatenating $\omega^{(1)}, \dots, \omega^{(L)}$, which are constructed either using the procedure in case (a) or that in case (b). First notice that, regardless of which procedure has been used at step j, the following inequality always holds for every $j = 1, \dots, L$:

$$H(\sigma_{j+1}) \leq H(\sigma_j).$$

Using this fact in combination with (6.26) and (6.27) shows that the path ω always satisfies

$$\Phi_\omega \leq H(\sigma) + 2.$$

Furthermore, in the special case in which σ has no vertical white bridges, our procedure considers case (b) for every $j = 1, \dots, L$ and thus, by virtue of (6.27), the path ω satisfies

$$\Phi_\omega - H(\sigma) \leq 1.$$

We remark that the stable state **a** (or **c**) can also be the target state of the reduction algorithm by columns. In this scenario, one should adjust the condition (6.25) on the initial condition accordingly, requiring that σ has no particles in columns c_1 and c_2 (columns c_0 and c_1, respectively). The offset of rows and columns in the procedures described above should of course be tweaked appropriately.

We now use the reduction algorithms we just introduced to show that the lower bound for the communication height between $\mathbf{a}$ and $\mathbf{b}$ given in Proposition 6.3.5 is sharp, by explicitly giving a path that attains that value.

Proposition 6.3.6 (Reference path). *There exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ in $\mathcal{X}$ such that*

$$\Phi_{\omega^*} - H(\mathbf{a}) = \min\{K, 2L\} + 1. \tag{6.28}$$

Proof. We distinguish two cases, depending on whether (a) $K \leq 2L$ and (b) $K > 2L$.

For case (a), we construct such a path ω^* as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{a} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{b}$ for a suitable configuration $\sigma^* \in \mathcal{X}$, and prove that $\Phi_{\omega^{(1)}} = H(\sigma^*) = H(\sigma) + K$ and that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ are satisfied, so that $\Phi_{\omega^*} = H(\mathbf{a}) + K + 1$ as desired. The reason why ω is best described as the concatenation of two shorter paths is the following: The reduction algorithm by columns cannot in general be started directly from $\mathbf{a}$ and the path $\omega^{(1)}$ indeed leads from $\mathbf{a}$ to σ^*, which is a suitable configuration to initialize the reduction algorithm by columns. The configuration σ^* differs from $\mathbf{a}$ only in the sites of column c_3:

$$\sigma^*(v) := \begin{cases} \mathbf{a}(v) & \text{if } v \in V \setminus c_3, \\ 0 & \text{if } v \in c_3. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{K+1}^{(1)})$, with $\omega_1^{(1)} = \mathbf{a}$ and $\omega_{K+1}^{(1)} = \sigma^*$ can be constructed as follows. For $i = 1, \dots, K$, at step i we remove from configuration $\omega_i^{(1)}$ the particle in site $(3, 2i-1)$, increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore the configuration σ^* is such that $H(\sigma^*) - H(\mathbf{a}) = K$ and $\Phi_{\omega^{(1)}} = H(\mathbf{a}) + K$. The second path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ is then constructed by means of the reduction algorithm by columns, which can be used since the configuration σ^* satisfies condition (6.24) and hence is a suitable initial configuration for the algorithm. The algorithm guarantees that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$, since configuration σ^* has no vertical white bridges (see case (b) for the reduction algorithm by columns), and thus the conclusion follows. Figure 6.10 illustrates the reference path for the triangular grid $\mathcal{T}_{3,3}$.

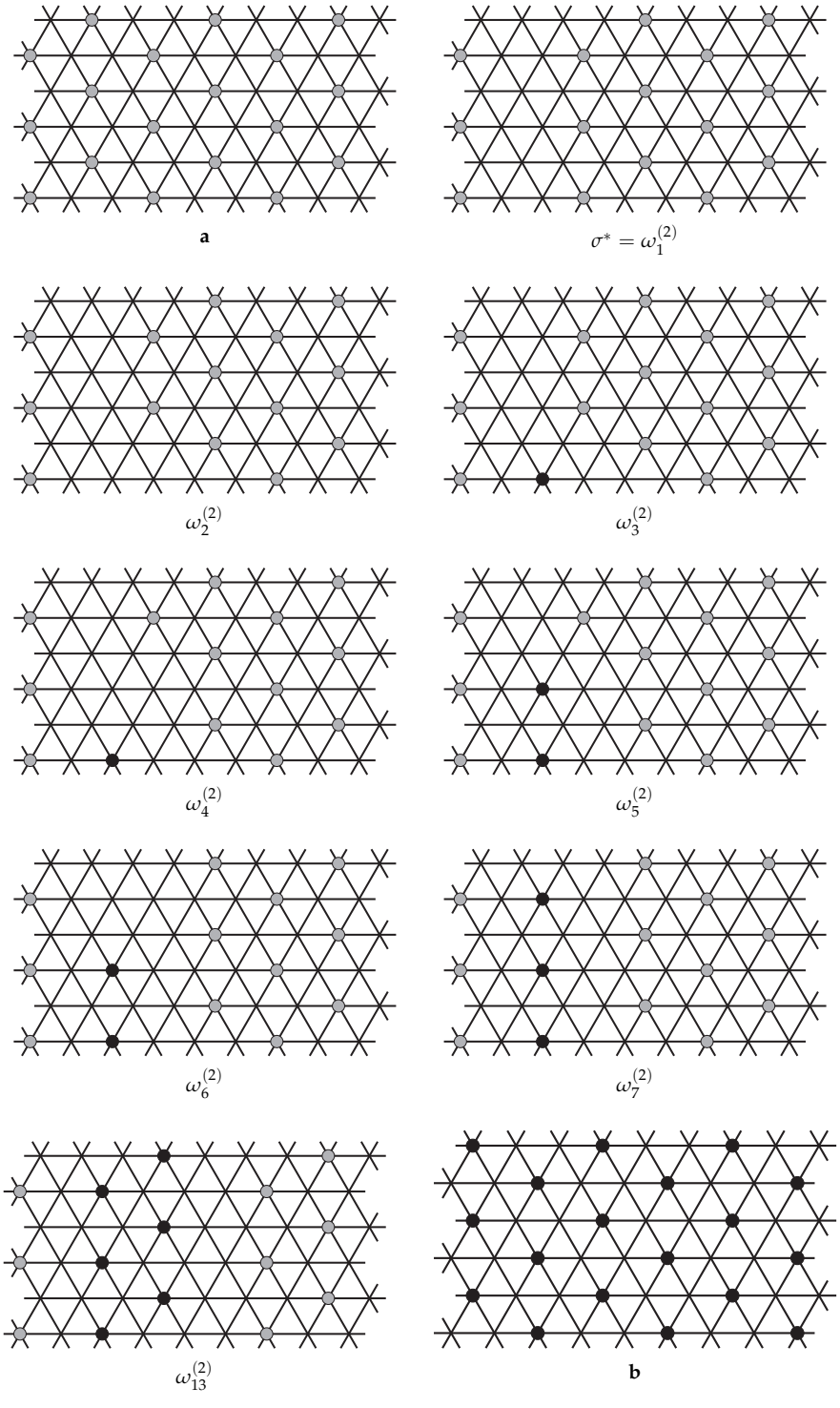


Figure 6.10: Illustration of the reference path $\omega^* : \mathbf{a} \to \mathbf{b}$ in the case $K \leq 2L$

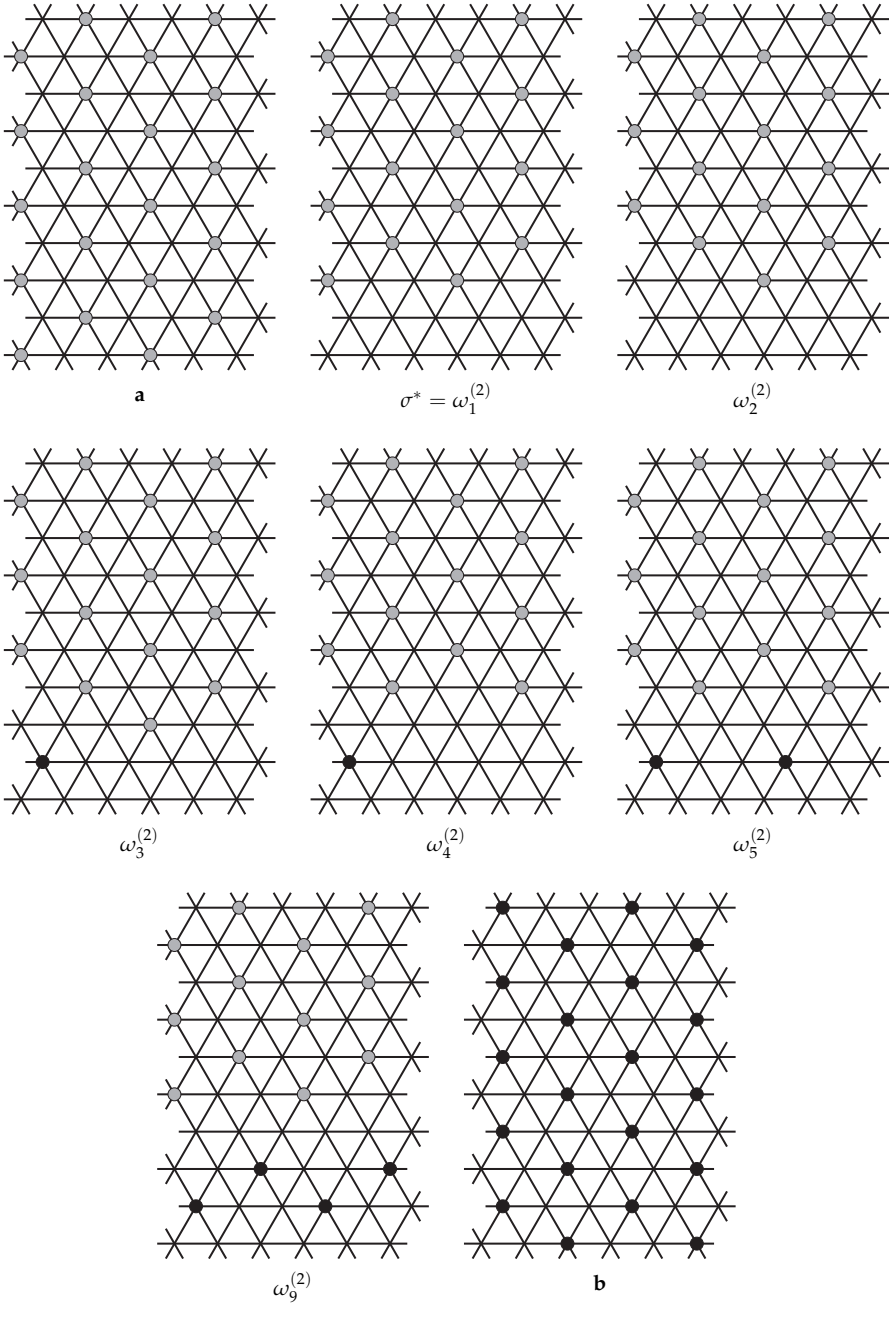


Figure 6.11: Illustration of the reference path $\omega^* : \mathbf{a} \to \mathbf{b}$ in the case $K > 2L$

For case (b), we construct such a path ω^* as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{a} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{b}$, and prove that $\Phi_{\omega^{(1)}} = H(\sigma^*) = H(\sigma) + 2L$ and that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ are satisfied, so that $\Phi_{\omega^*} = H(\mathbf{a}) + 2L + 1$ as desired. Some snapshots from the reference path from $\mathbf{a}$ to $\mathbf{b}$ on the triangular grid $\mathcal{T}_{5,2}$ are shown in Figure 6.11. Also in this case we construct ω as concatenation of two shorter paths because $\mathbf{a}$ is not a suitable initial configuration for a reduction algorithm. Consider the configuration σ^* that differs from $\mathbf{a}$ only in the sites of the first horizontal stripe S_0, namely

$$\sigma^*(v) := \begin{cases} \mathbf{a}(v) & \text{if } v \in V \setminus S_0, \\ 0 & \text{if } v \in S_0. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{2L+1}^{(1)})$, with $\omega_1^{(1)} = \mathbf{a}$ and $\omega_{2L+1}^{(1)} = \sigma^*$ can be constructed as follows. For $i = 1, \dots, 2L$, at step i we remove from configuration $\omega_i^{(1)}$ the first particle in lexicographic order in S_0, increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore the configuration σ^* is such that $H(\sigma^*) - H(\mathbf{a}) = 2L$ and $\Phi_{\omega^{(1)}} = H(\mathbf{a}) + 2L$. The second path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ is then constructed by means of the reduction algorithm by rows, which can be used since the configuration σ^* satisfies condition (6.24) and hence is a suitable initial configuration for the algorithm. The reduction algorithm by rows guarantees that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ and thus the conclusion follows. □

Proof of Theorem 6.3.1(ii). The proof of the identity involving $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a})$ readily follows by combining the lower bound in Proposition 6.3.5 and the statement of Proposition 6.3.6. The remaining identities can be derived invoking the symmetry of the triangular grid (or using the same arguments as in the proof of Propositions 6.3.5 and 6.3.6). □

Proof of Theorem 6.3.1(iii). We will show that for every admissible configuration σ on a triangular grid graph $\mathcal{T}_{K,L}$ with $\sigma \neq \mathbf{a}, \mathbf{b}, \mathbf{c}$, there exists a path ω from σ to one of the three stable states such that

$$\Phi_\omega - H(\omega) \leq \min\{K, 2L\}.$$

The idea is to construct such a path using the geometric features of the configuration σ and exploiting the reduction algorithms described earlier in this section. We start by distinguishing two cases: (a) $K \leq 2L$ and (b) $K > 2L$.

Consider case (a) first, where $K \leq 2L$. We distinguish two sub-cases, depending on whether σ has at least one vertical bridge or not.

If σ has a vertical bridge in a vertical stripe C, then σ is a suitable starting configuration for the reduction algorithm, which yields a path ω that goes from σ to the stable configuration in $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ on which σ agrees in stripe C. The path ω constructed in this way is such that

$$\Phi_\omega - H(\sigma) \leq 2$$

and thus $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) - H(\sigma) \leq 2 \leq K \leq \min\{K, 2L\}$, since by assumption K is an integer greater than 1.

Suppose now that there are no vertical bridges in σ. Since $\sigma \notin \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, which is the set of stable states in view of Theorem 6.3.1(i), configuration σ has a positive energy wastage $U(\sigma) > 0$. In view of (6.14), this means that there exists a vertical stripe C^* such that $U^{C^*}(\sigma) > 0$. Without loss of generality, we may assume (modulo relabeling) that C^* is the vertical stripe C_1, which consists of columns c_1, c_2 and c_3. By definition of energy wastage in a stripe, it follows that σ has at most $K - 1$ particles. Removing all the gray and white particles one by one, we construct a path $\omega^{(1)}$ from σ to a new configuration σ^* that satisfies

$$H(\sigma^*) - H(\sigma) \leq K - 1. \tag{6.29}$$

Since we remove all the gray particles from c_3 and all the white particles from c_2, σ^* is a suitable starting configuration for the reduction algorithm by columns with target state $\mathbf{b}$, in view of (6.25). We obtain in this way a second path $\omega^{(2)} : \sigma^* \to \mathbf{b}$, which is such that

$$\Phi_{\omega^{(2)}} - H(\sigma^*) \leq 1, \tag{6.30}$$

thanks to the absence of vertical bridges in σ (and thus in σ^*). The path $\omega : \sigma \to \mathbf{b}$ obtained by concatenating $\omega^{(1)}$ and $\omega^{(2)}$ is such that

$$\Phi_\omega - H(\sigma) \leq K,$$

since $\Phi_{\omega^{(1)}} - H(\sigma) \leq K - 1$ and in view of (6.29) and (6.30), and thus

$$\Phi(\sigma, \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) - H(\sigma) \leq K.$$

We remark that there is nothing special about $\mathbf{b}$ as target configuration of the path ω we just constructed. Indeed, by choosing the vertical stripe C^* with a different offset, we could have obtained a configuration σ^* which would have been a suitable initial configuration for the reduction algorithm by columns with target configuration $\mathbf{a}$ or $\mathbf{c}$.

We now turn to case (b), in which $K > 2L$. Thanks to Lemma 6.3.3, there must be a horizontal stripe S on which σ does not have a horizontal bridge, otherwise $\sigma \in \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. In particular, σ has at most $2L - 1$ particles on S, which without loss of generality we may assume to be S_0. We build a path $\omega^{(1)}$ from σ to a new configuration σ^* by removing all these particles one by one, so that $\Phi_{\omega^{(1)}} - H(\sigma) \leq 2L - 1$ and $H(\sigma^*) - H(\sigma) \leq 2L - 1$. Starting with configuration σ^* we can then use the reduction algorithm by rows to obtain a second path $\omega^{(2)}$ from σ^* to any of the three stable configurations. Since $\Phi_{\omega^{(2)}} - H(\sigma^*) \leq 1$, the path ω constructed by the concatenation of $\omega^{(1)}$ and $\omega^{(2)}$ satisfies

$$\Phi_\omega - H(\sigma) \leq 2L$$

and thus $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) - H(\sigma) \leq 2L$. □

6.3.3 *Proofs of Theorems 6.2.1 and 6.2.2*

In this subsection, we prove the two main theorems of the chapter exploiting the properties of the energy landscape proved earlier in this section. In these proofs we use some notation introduced in Chapter 4 to characterize the depths of the wells in the energy landscape, please refer to Subsection 4.2.2 and, in particular, to definitions (4.17) and (4.20).

Proof of Theorem 6.2.1. Exploiting the equivalent characterization of maximum depth given in Lemma 4.2.6, it immediately follows from Theorem 6.3.1(iii) that

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) \leq \min\{K, 2L\}. \tag{6.31}$$

Furthermore, we claim that the following identity holds:

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}, \mathbf{c}\}) = \min\{K, 2L\} + 1. \tag{6.32}$$

First notice that since $\mathbf{a} \in \mathcal{X} \setminus \{\mathbf{b}, \mathbf{c}\}$, we have

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}, \mathbf{c}\}) \geq \Phi(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\}) - H(\mathbf{a}) = \min\{K, 2L\} + 1.$$

In order to prove that this lower bound is sharp, we need to show that $\Phi(\sigma, \{\mathbf{b}, \mathbf{c}\}) - H(\sigma) \leq \min\{K, 2L\} + 1$ for every configuration $\sigma \neq \mathbf{a}, \mathbf{b}, \mathbf{c}$, but we will actually prove a stronger inequality, namely

$$\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq \min\{K, 2L\} + 1, \quad \forall\, \sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}. \tag{6.33}$$

Inspecting the proof of Theorem 6.3.1(iii), we notice that every configuration $\sigma \neq \mathbf{a}, \mathbf{b}, \mathbf{c}$ can be reduced either directly to $\mathbf{b}$, or otherwise to $\mathbf{a}$ or $\mathbf{c}$, depending on its geometrical features. If σ can be reduced directly to $\mathbf{b}$, then we proved therein that $\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq \min\{K, 2L\}$. Consider now the scenario where the configuration σ cannot be reduced to $\mathbf{b}$ directly, that is the case where $K \leq 2L$ and σ displays a vertical gray bridge or a vertical white bridge. In the proof of Theorem 6.3.1(iii) we built a path ω from σ to either $\mathbf{a}$ or $\mathbf{a}$ such that $\Phi_\omega \leq H(\sigma) + 2$, which, concatenated with the reference path from $\mathbf{a}$ to $\mathbf{b}$ given in Proposition 6.3.6 (or the analogous path from $\mathbf{c}$ to $\mathbf{b}$), shows that

$$\Phi(\sigma, \mathbf{b}) \leq \max\{H(\sigma) + 2, \Phi(\mathbf{a}, \mathbf{b})\}.$$

Thus,

$$\begin{aligned}\Phi(\sigma, \mathbf{b}) - H(\sigma) &\leq \max\{2, \Phi(\mathbf{a}, \mathbf{b}) - H(\sigma)\} \leq \max\{2, \Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a})\} \\ &= \max\{2, \min\{K, 2L\} + 1\} \leq \min\{K, 2L\} + 1,\end{aligned}$$

which implies that inequality (6.33) holds and proves in this way identity (6.32). The pair $(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\})$ then satisfies condition (4.50) (the so-called "absence of deep cycles"), since, thanks to identity (6.32),

$$\Phi(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\}) - H(\mathbf{a}) = \min\{K, 2L\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}, \mathbf{c}\}).$$

Proposition 4.2.18 then implies that Assumption A holds for the pair $(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\})$ and Corollary 4.2.16 and Theorem 4.2.17 yields statements (i) and (ii) of Theorem 6.2.1. Furthermore, combining Proposition 6.3.6 and inequality (6.31) gives

$$\begin{aligned}\Gamma(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\}) &= \Phi(\mathbf{a}, \{\mathbf{b}, \mathbf{c}\}) - H(\mathbf{a}) = \min\{K, 2L\} + 1 \\ &> \min\{K, 2L\} = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}),\end{aligned}$$

which implies that the initial cycle $C_{\{\mathbf{b},\mathbf{c}\}}(\mathbf{a})$ (see definition (4.16)) is the unique deepest cycle of the subset $\mathcal{X} \setminus \{\mathbf{b}, \mathbf{c}\}$ and proves that condition (4.55) holds. Proposition 4.2.20 and Theorem 4.2.19 then yield the asymptotic exponentiality of the scaled tunneling time $\tau^{\mathbf{a}}_{\{\mathbf{b},\mathbf{c}\}} / \mathbb{E}\tau^{\mathbf{a}}_{\{\mathbf{b},\mathbf{c}\}}$, i.e.

$$\frac{\tau^{\mathbf{a}}_{\{\mathbf{b},\mathbf{c}\}}}{\mathbb{E}\tau^{\mathbf{a}}_{\{\mathbf{b},\mathbf{c}\}}} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \beta \to \infty, \tag{6.34}$$

proving Theorem 6.2.1(iii).

Consider now the other tunneling time, namely $\tau^{\mathbf{a}}_{\mathbf{b}}$. From inequality (6.33) and the definition (4.20) of maximum depth it immediately follows that

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}\}) \leq \min\{K, 2L\} + 1,$$

which, in view of Proposition 6.3.6, implies that

$$\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) = \min\{K, 2L\} + 1 = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}\}).$$

Hence the pair $(\mathbf{a}, \{\mathbf{b}\})$ satisfies condition (4.50) and statements (i) and (ii) of Theorem 6.2.1 hold for the tunneling time $\tau^{\mathbf{a}}_{\mathbf{b}}$ using Proposition 4.2.18, Corollary 4.2.16 and Theorem 4.2.17.

The pair $(\mathbf{a}, \{\mathbf{b}\})$ does not satisfy condition (4.55) and neither Assumption B (4.53), due to the presence of a deep cycle (the one where the stable state $\mathbf{c}$ lies) different from the initial one (where $\mathbf{a}$ lies). Indeed, $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) \not< \Phi(\mathbf{c}, \mathbf{b}) - H(\mathbf{c})$, as shown in Theorem 6.3.1(ii). Hence, Theorem 6.2.1(iv) does not follow immediately from Theorem 4.2.19. The asymptotic exponentiality of the scaled hitting time $\tau^{\mathbf{a}}_{\mathbf{b}} / \mathbb{E}\tau^{\mathbf{a}}_{\mathbf{b}}$ in the limit $\beta \to \infty$ is proved differently, leveraging the symmetric structure of the energy landscape. Indeed the triangular grid $\mathcal{T}_{K,L}$ is a symmetric 3-partite graph, as shown in Section 2.2.

In Section 2.3 we derive (in the continuous-time setting) in Corollary 2.3.2 the asymptotic distribution of the scaled transition time between any two dominant states in the limit $\sigma \to \infty$. We will now argue that we can obtain a similar result in the current discrete-time setting.

Firstly, the low-temperature limit $\beta \to \infty$ in which we are interested is equivalent to the limit $\sigma \to \infty$ considered in Chapter 2, as illustrated in Section 1.3. Secondly, the set $\mathcal{D}$ of dominant states of the activity process correspond to the set of stable states of the hard-core model. Furthermore, Corollary 2.3.2 relies on three crucial properties proved in Proposition 2.3.1. The proofs of

these two results do not use in an essential way the fact that the activity process is a continuous-time process. In fact, it can be easily checked that the same statements can be proved for the uniformized Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ which we study in this chapter. In the notation of this chapter, the stochastic representation (2.11) for the the hitting time $\tau_{\mathbf{b}}^{\mathbf{a}}$ reads as

$$\tau_{\mathbf{b}}^{\mathbf{a}} \stackrel{d}{=} \sum_{i=1}^{N_{\mathcal{D}}} T^{(i)},$$

where $(T^{(i)})_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables with $T^{(i)} \stackrel{d}{=} \tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}$ and $N_{\mathcal{D}}$ is an independent geometric random variable with success probability equal to $\frac{1}{2}$, i.e.

$$\mathbb{P}(N_{\mathcal{D}} = m) = \left(\frac{1}{2}\right)^{m-1} \frac{1}{2}, \quad m \geq 1.$$

The random variable $N_{\mathcal{D}}$ counts the number of non-consecutive visits to stable states until the stable state $\mathbf{b}$ is hit. Non-consecutive visits to stable states here means that we count as actual visit to a stable state only the first one after the last visit to a different stable state. The random time between these non-consecutive visits does not depend on the last visited stable state, since

$$\tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}} \stackrel{d}{=} \tau_{\{\mathbf{a},\mathbf{b}\}}^{\mathbf{c}}.$$

Exploiting the symmetry of the energy landscape and arguing like in Section 2.3, we can show that if the Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ starts in state $\mathbf{a}$, the probability that $\mathbf{b}$ is the next visited stable state that is different from $\mathbf{a}$ is equal to $\frac{1}{2}$ (and similarly if the starting state is $\mathbf{c}$). This fact motivates the value of the success probability of the geometric random variable $N_{\mathcal{D}}$. Using the fact that $N_{\mathcal{D}}$ does not depend on β and arguing as in the proof of Corollary 2.3.2(iii), we can show that

$$\frac{\tau_{\mathbf{b}}^{\mathbf{a}}}{\mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}}} \xrightarrow{d} \frac{1}{2} \sum_{i=1}^{\mathrm{Geo}(1/2)} Y^{(i)}, \quad \text{as } \beta \to \infty,$$

where $\{Y^{(i)}\}_{i\in\mathbb{N}}$ are i.i.d. exponential random variables, thanks to (6.34) and the fact that $\tau_{\mathcal{D}\setminus\{\mathbf{a}\}}^{\mathbf{a}} = \tau_{\{\mathbf{b},\mathbf{c}\}}^{\mathbf{a}}$. Theorem 6.2.1(iv) then follows by noticing that a geometric sum of i.i.d. exponential random variables scaled by its mean is also exponentially distributed with unit mean. □

Proof of Theorem 6.2.2. The proof immediately follows from (6.33) in combination with Proposition 4.2.24. □

7 WIDOM-ROWLISON MODEL ON GRID GRAPHS

This chapter is devoted to the analysis of the Widom-Rowlison model on finite graphs with Metropolis dyanamics, which describes the evolution of a particle system where different types of particles interact subject to certain hard-core constraints. We focus in particular on the scenario in which there are exactly two types of particles and the spatial structure is modeled by grid graphs. We study the asymptotic behavior of this interacting particle system in the low-temperature regime, analyzing the tunneling time between its two stable states, which we denote by **a** and **b**, and the mixing time of the Markov chain. This chapter is organized as follows. In Section 7.1 we give a detailed description of the Widom-Rowlison model with two types of particles and show the connection between this model on a bipartite graph and the multi-channel CSMA network presented in Section 1.5 on the same graph with $C = 2$ available channels. In Section 7.2 we present our main results for the Widom-Rowlison model on grid graphs, concerning the asymptotic properties of tunneling time $\tau_{\mathbf{b}}^{\mathbf{a}}$ and the order of magnitude of the mixing time in the low-temperature regime. The rest of the chapter is then dedicated to the proofs of the main results. In Section 7.3 we introduce some notation and prove some preliminary results. Using a combinatorial method similar to that developed in Chapter 5, we analyze the energy landscapes corresponding to the Widom-Rowlison model on toric and open grid graphs in Sections 7.4 and 7.5, respectively.

7.1 MODEL DESCRIPTION

Consider a finite undirected graph $G = (V, E)$, which models the spatial structure of the finite volume where two types of particles dynamically interact subject to certain hard-core constraints. The N vertices of the graph G represent all possible sites where particles can reside. Particles can be of two types, A or B, and there is nearest-neighbor hard-core exclusion between unlike particles. These hard-core constraints are modeled by the set E of edges connecting the pairs of sites that cannot be occupied simultaneously by particles of different types. We associate a variable $\sigma(v) \in \{-1, 0, 1\}$ to each site $v \in V$, indicating the absence (0) or the presence of a particle of type A (1) or type B (-1) in that site. We say that a particle configuration $\sigma \in \{-1, 0, 1\}^N$ is *admissible* if it does not violate the hard-core constraints between unlike particles in neighboring sites and denote by $\mathcal{X}$ the collection of admissible configurations, namely

$$\mathcal{X} := \{\sigma \in \{-1, 0, 1\}^N \ : \ \sigma(v)\sigma(w) \neq -1 \quad \forall (v, w) \in E\}. \tag{7.1}$$

The evolution of this interacting particle system is described by the Metropolis dynamics introduced in Chapter 4. More specifically, we consider the single-site update Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ parametrized by a positive parameter β, representing the *inverse temperature*, with Metropolis transition probabilities

$$P_\beta(\sigma,\sigma') := \begin{cases} c(\sigma,\sigma')e^{-\beta[H(\sigma')-H(\sigma)]^+}, & \text{if } \sigma \neq \sigma', \\ 1-\sum_{\eta\neq\sigma} P_\beta(\sigma,\eta), & \text{if } \sigma = \sigma', \end{cases} \tag{7.2}$$

corresponding to the energy landscape $(\mathcal{X}, H, c)$ where the state space $\mathcal{X}$ is given in (7.1) and the energy and connectivity function are chosen as follows. The energy function $H : \mathcal{X} \to \mathbb{R}$ counts the number of particles, regardless of their type:

$$H(\sigma) := -\sum_{v\in V} \mathbb{1}_{\{\sigma(v)\neq 0\}} = -\sum_{v\in V} |\sigma(v)|, \quad x \in \mathcal{X}. \tag{7.3}$$

The connectivity function $c : \mathcal{X}\times\mathcal{X} \to [0,1]$ is given by

$$c(\sigma,\sigma') := \begin{cases} \frac{1}{2N}, & \text{if } d(\sigma,\sigma') = 1, \\ 0, & \text{if } d(\sigma,\sigma') > 1, \\ 1-\sum_{\eta\neq\sigma} c(\sigma,\eta), & \text{if } \sigma = \sigma', \end{cases} \tag{7.4}$$

where $d : \mathcal{X}\times\mathcal{X} \to \mathbb{N}$ is a distance function defined as

$$d(\sigma,\sigma') := \sum_{v\in V} \left(\mathbb{1}_{\{\sigma(v)\neq\sigma'(v)\}} + \mathbb{1}_{\{\sigma(v)\sigma'(v)\neq 0\}} \right). \tag{7.5}$$

In other words, the connectivity function allows only single-site updates and prescribes that a particle occupying a certain site cannot be replaced by a particle of the other type in a single step. The dynamics induced by this energy landscape can be described in words as follows. At every step a site $v \in V$ and a type (A or B) are selected uniformly at random; if the selected site is occupied, the particle therein is removed with probability $e^{-\beta}$; if instead the selected site v is empty, then a particle of the chosen type is created at v with probability 1 if and only if none of the neighboring sites is occupied by particles of the opposite type.

The Widom-Rowlinson model was originally introduced in the chemistry literature by [139] as a continuum model of particles living in $\mathbb{R}^d$ to study the vapor-liquid phase transition. The discrete-space variant we are interested in was first studied in [94], where the authors model a lattice gas where two types of particles interact. The Widom-Rowlinson model has in fact two equivalent standard formulations - one as a binary gas and the other as a single-species model of a dense (liquid) phase in contact with a rarefied (gas) phase. In the binary gas formulation, the only interaction is a hard-core exclusion between the two species of particles - call them A and B. If the B particles are

invisible, the resulting system of A particles yields a model for vapor-liquid phase transitions, which has been introduced and discussed by Widom and Rowlinson. The Widom-Rowlinson model has been proven to exhibit a phase transition both on $\mathbb{Z}^d$ [94] and on the continuum [95, 125]. On general graphs, however, its behavior is not always monotonic, as proved in [24]. Many variants of the Widom-Rowlison model have been studied in the literature: In [96] the authors study the Widom-Rowlison model with $M > 2$ types of particles, while the case where there is only strong repulsion (and no hard-core interaction) between unlike particles has been investigated in [94]. Further properties of the continuum Widom-Rowlison model have been studied in [33].

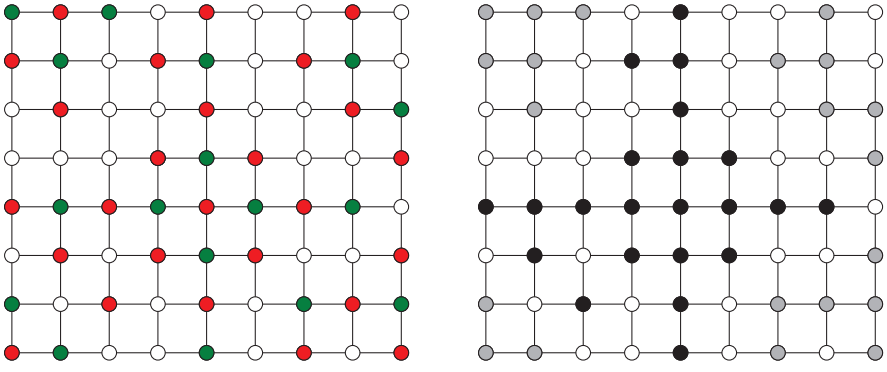
7.1.1 *Connection with multi-channel CSMA networks*

In this subsection we illustrate how the multi-channel CSMA dynamics on a conflict graph G described in Section 1.5 can be mapped into the Widom-Rowlison model on the same graph G when G is bipartite and there are $C = 2$ available channels.

Consider a bipartite connected conflict graph $G = (V, E)$ on N nodes, with partition $V = V_e \cup V_o$. Furthermore, we assume that the interference constraints on the two available channels are both described by the graph G. As a consequence, there is a one-to-one correspondence between 2-colorings of G and admissible activity states. In particular, there are then exactly two proper colorings of the graph G, which correspond to the only two possible ways of assigning two colors to the two components V_e and V_o. We will now introduce an equivalent representation of admissible activity states on the conflict graph G with $C = 2$ channels available. The main idea is that for each active node instead of keeping track of the channel on which it is active, it is enough to keep track which one of the two possible (global) colorings it agrees with. This intuition can be made rigorous by considering the function $\mathcal{W} : \{0,1,2\}^N \to \{-1,0,1\}^N$ that maps an admissible activity state $x \in \{0,1,2\}^N$ into the vector $\mathcal{W}(x) \in \{-1,0,1\}^N$ whose i-th component is defined as

$$\mathcal{W}(x)_i = \begin{cases} 1 & \text{if } x_i = 1 \text{ and } i \in V_e, \\ 1 & \text{if } x_i = 2 \text{ and } i \in V_o, \\ -1 & \text{if } x_i = 2 \text{ and } i \in V_e, \\ -1 & \text{if } x_i = 1 \text{ and } i \in V_o, \\ 0 & \text{if } x_i = 0. \end{cases}$$

The function $\mathcal{W}$ is illustrated in the case of the 8×9 open grid by Figure 7.1, where V_e and V_o are the collections of even and odd nodes, respectively. Since $x \in \{0,1,2\}^N$ is an admissible activity state on a graph G if $x_i x_j = 0$ or



(a) $x \in \mathcal{X} \subset \{0,1,2\}^N$, where the inactive nodes are white, nodes active on channel 1 are red and nodes active on channel 2 are green

(b) $\mathcal{W}(x) \in \{-1,0,1\}^N$, where particles of type A ($\mathcal{W}(x) = 1$) are gray and particles of type B ($\mathcal{W}(x) = -1$) are black

Figure 7.1: An activity state $x \in \mathcal{X}$ using two channels on the 8×9 grid graph and the corresponding Widom-Rowlison particle configuration $\mathcal{W}(x)$

$x_i \neq x_j$ for every pair of neighboring nodes $(i,j) \in E$, the corresponding state $\mathcal{W}(x) \in \{-1,0,1\}^N$ satisfies

$$\mathcal{W}(x)_i \mathcal{W}(x)_j = 0 \quad \text{or} \quad \mathcal{W}(x)_i = \mathcal{W}(x)_j \quad \forall\, (i,j) \in E. \tag{7.6}$$

Note that in this derivation, we used the key fact that, since G is bipartite, if i, j are neighbors in G, then they belong to different components. The constraints (7.6) imply that the image $\mathcal{W}(\mathcal{X})$ of the state space $\mathcal{X}$ under the mapping $\mathcal{W}$ is the collection of all states $\sigma \in \{-1,0,1\}^N$ such that

$$\sigma_i \sigma_j = 0 \quad \text{or} \quad \sigma_i = \sigma_j \quad \forall\, (i,j) \in E,$$

or, more compactly, $\sigma_i \sigma_j \neq -1$ for every pair of neighboring sites i and j, which we recognize to be the admissible configurations of the Widom-Rowlison model on the graph G, see (7.1).

7.2 ASYMPTOTIC BEHAVIOR OF TUNNELING TIMES AND MIXING TIME

In the rest of the chapter we focus on the study of the Widom-Rowlison model on grid graphs and, in particular, we want to understand the asymptotic behavior of this interacting particle system in terms of tunneling and mixing times in the low-temperature regime.

Given two integers $K, L \geq 2$, we take G to be the $K \times L$ grid graph $\Lambda = \Lambda_{K,L}$ described in Section 2.2 either with toroidal (periodic) boundary conditions, denoted by $\Lambda^T_{K,L}$, or with open (free) boundary conditions, denoted by $\Lambda^O_{K,L}$. Recall that the grid graph $\Lambda_{K,L} = (V, E)$ has vertex set $V = \{0, \dots, L-1\} \times \{0, \dots, K-1\}$ and $N = KL$ sites in total.

The equivalence between the Widom-Rowlison model and the multi-channel CSMA dynamics holds only when the underlying graph G is bipartite, as illustrated in Subsection 7.1.1. However, this assumption on G is not necessary for our results for the Widom-Rowlison model and for this reason we do not impose further conditions on the integers K and L, as we did instead in Chapters 2 and 5.

Let $\mathbf{a}$ and $\mathbf{b}$ be the two admissible configurations of the Widom-Rowlison model on the grid graph $\Lambda_{K,L}$ that correspond to the situation where all sites are occupied by particles of types A and B, respectively. Using the notation introduced in the previous subsection, we define

$$\mathbf{a}(v) := 1 \quad \forall\, v \in V, \qquad \text{and} \qquad \mathbf{b}(v) := -1 \quad \forall\, v \in V.$$

Clearly $H(\mathbf{a}) = -N = H(\mathbf{b})$. Furthermore, in Section 7.3 we will prove that $\mathbf{a}$ and $\mathbf{b}$ are the only two stable states of the energy landscape of the Widom-Rowlison model on any grid graph, regardless of the imposed boundary conditions.

We adopt the following coloring conventions for displaying a configuration $\sigma \in \mathcal{X}$: A site $v \in V$ is displayed as white if it is unoccupied (i.e. $\sigma(v) = 0$), while we color it in gray (black) if it is occupied by a particle of type A (by a particle of type B, respectively). In view of this coloring scheme, in the rest of the chapter we will refer to a particle of type A (B) as a *gray particle* (*black particle*, respectively). As a result, an admissible configuration consists of black and/or gray clusters separated by empty sites, since black and gray particles cannot reside in neighboring sites. Figure 7.2 illustrates three admissible configurations.

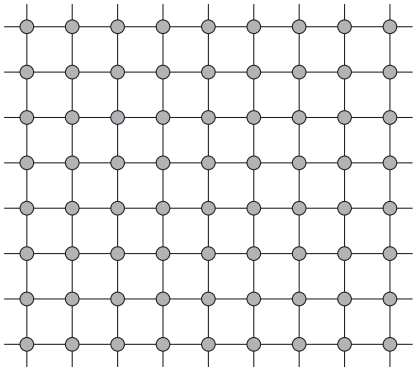
In order to capture how long the particle system takes to "switch" between the two stable states $\mathbf{a}$ and $\mathbf{b}$, we study the asymptotic behavior of the *tunneling time* $\tau_{\mathbf{b}}^{\mathbf{a}}$ of the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ in the low-temperature regime $\beta \to \infty$. The first main result proves the existence of an exponent $\Gamma(\Lambda) > 0$ for any grid graph Λ that gives an asymptotic control in probability of $\tau_{\mathbf{b}}^{\mathbf{a}}$ on a logarithmic scale as $\beta \to \infty$ and characterizes the asymptotic order of magnitude of the mean tunneling time $\mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}}$. We further show that the tunneling time $\tau_{\mathbf{b}}^{\mathbf{a}}$ scaled by its mean converges in distribution to an exponential unit-mean random variable.

Theorem 7.2.1 (Asymptotic behavior of the tunneling time $\tau_{\mathbf{b}}^{\mathbf{a}}$)**.** *Consider the Metropolis Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ corresponding to the Widom-Rowlison model on a $K \times L$ grid graph Λ. There exists a constant $\Gamma(\Lambda) > 0$ such that:*

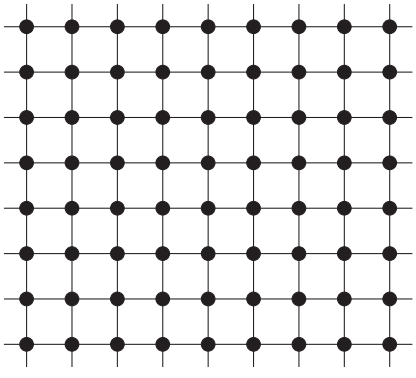
(i) *For every $\epsilon > 0$,* $\lim_{\beta\to\infty} \mathbb{P}\left(e^{\beta(\Gamma(\Lambda)-\epsilon)} < \tau_{\mathbf{b}}^{\mathbf{a}} < e^{\beta(\Gamma(\Lambda)+\epsilon)}\right) = 1;$

(ii) $\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}} = \Gamma(\Lambda);$

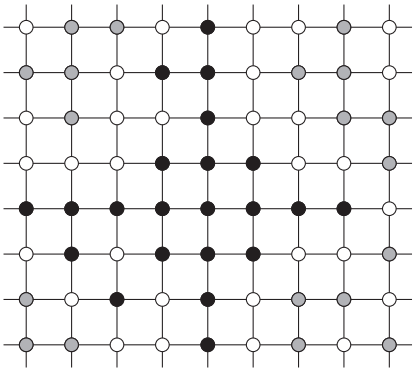
(iii) $\frac{\tau_{\mathbf{b}}^{\mathbf{a}}}{\mathbb{E}\tau_{\mathbf{b}}^{\mathbf{a}}} \xrightarrow{d} \mathrm{Exp}(1), \quad \text{as } \beta \to \infty.$



(a) Configuration **a**



(b) Configuration **b**



(c) An admissible configuration with particles of both types

Figure 7.2: Examples of admissible configurations on the 8×9 toric grid $\Lambda^T_{8,9}$

The proof of Theorem 7.2.1 relies on the general results for tunneling times for Metropolis Markov chains obtained in Chapter 4 and on the study of structural properties of the energy landscape $(\mathcal{X}, H, q)$ corresponding to the Widom-Rowlison model on Λ, which are presented in the next section.

The combinatorial approach developed in Sections 7.4 and 7.5 shows how the quantity $\Gamma(\Lambda)$ depends both on the grid size and on the chosen boundary conditions, as established by the next theorem.

Theorem 7.2.2 (The exponent $\Gamma(\Lambda)$ for grid graphs)**.** *Let Λ be a $K \times L$ grid graph. Then the energy barrier $\Gamma(\Lambda)$ between* **a** *and* **b** *appearing in Theorem 7.2.1 takes the values*

$$\Gamma(\Lambda) = \begin{cases} \min\{2K+1, 2L\} & \text{if } \Lambda = \Lambda^T_{K,L} \text{ with } K \leq L \text{ and } K+L \geq 6, \\ \min\{K, L\} + 1 & \text{if } \Lambda = \Lambda^O_{K,L}. \end{cases}$$

The additional assumption that $K + L \geq 6$ in the case of toric grids is necessary to leave out the two special cases $(K, L) = (2, 2)$ and $(K, L) = (2, 3)$, where the constant $\Gamma(\Lambda)$ takes the values 3 and 4, respectively.

Note that Theorem 7.2.2 states that $\Gamma(\Lambda) = 2K = 2L$ for a square toric grid (that is $\Lambda^T_{K,L}$ with equal dimensions $K = L$), while $\Gamma(\Lambda) = 2K + 1$ for a rectangular toric grid (namely $\Lambda^T_{K,L}$ with $K < L$).

We remark that from our analysis we could also immediately prove asymptotic results for the tunneling time $\tau^{\mathbf{a}}_{\mathbf{b}}$ analogous to those of Theorem 7.2.1 in the case of a cylindrical grid, that is the grid graph $\Lambda_{K,L}$ with semi-periodic boundary conditions (see Section 2.2 or 5.1). Indeed, arguing along similar lines as in Subsection 5.3.4, we can derive that for this class of grids $\Gamma(\Lambda) = \min\{K, 2L\} + 1$.

The value of the exponent $\Gamma(\Lambda)$ for the Widom-Rowlison model on a grid graph Λ is roughly twice as large as the exponent for the hard-core model on the same grid graph, see Theorem 5.2.2. This difference can be intuitively understood by noticing that the Widom-Rowlison model on a grid graph Λ is equivalent to the hard-core model on a 3-dimensional grid, which is the Cartesian product of Λ and the complete graph on two nodes, as illustrated in Figure 7.3. All the ingredients needed to make this equivalence rigorous have been introduced in Section 1.5 in the context of multi-channel CSMA networks.

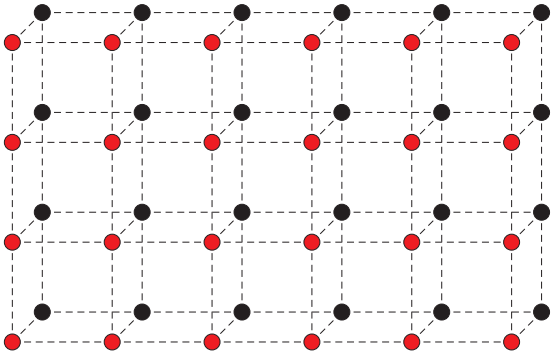
Besides appearing in the tunneling time asymptotics in Theorem 7.2.1, the exponent $\Gamma(\Lambda)$ also characterizes the asymptotic order of magnitude of the mixing time $t^{\mathrm{mix}}_\beta(\epsilon, \Lambda)$ and of the spectral gap $\rho_\beta(\Lambda)$ of the Metropolis Widom-Rowlison dynamics on Λ (see definitions in Subsection 4.2.8).

Theorem 7.2.3 (Mixing time of the Widom-Rowlison model on grid graphs)**.** *Let Λ be a $K \times L$ grid graph. For any $0 < \epsilon < 1$, the mixing time $t^{\mathrm{mix}}_\beta(\epsilon, \Lambda)$ of the Markov chain $\{X^\beta_t\}_{t \in \mathbb{N}}$ satisfies*

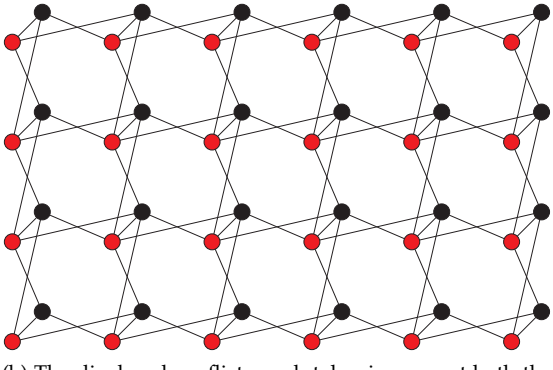
$$\lim_{\beta \to \infty} \frac{1}{\beta} \log t^{\mathrm{mix}}_\beta(\epsilon, \Lambda) = \Gamma(\Lambda) = \lim_{\beta \to \infty} -\frac{1}{\beta} \log \rho_\beta(\Lambda).$$

Furthermore, there exist two positive constants $0 < c_1 \leq c_2 < \infty$ independent of β such that

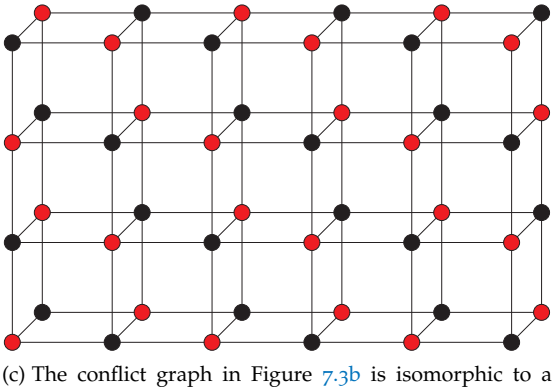
$$c_1 e^{-\beta \Gamma(\Lambda)} \leq \rho_\beta(\Lambda) \leq c_2 e^{-\beta \Gamma(\Lambda)}.$$



(a) Every site of the original grid graph Λ is doubled, so that we obtain two layers of sites. The occupied sites on the first layers are particle of type A, while those on the second layer are particles of type B



(b) The displayed conflict graph takes in account both the hard-core constraints between unlike particles in neighboring sites and the impossibility of having more than one particle in the each site



(c) The conflict graph in Figure 7.3b is isomorphic to a three-dimensional grid graph

Figure 7.3: The equivalence between the Widom-Rowlison model on an open grid and the hard-core model on a three-dimensional open grid

7.3 ENERGY LANDSCAPE ANALYSIS

In this section we present the structural properties of the energy landscapes of the Widom-Rowlison model on grid graphs and show show how they can be used to prove the main results presented in Section 7.2. In particular, we find the value of the energy barrier between the stable states **a** and **b** and show that this is the highest energy barrier of the whole energy landscape, proving the so-called "absence of deep cycles" property for the Widom-Rowlison model on grid graphs. The proof method is similar to that used in Chapter 5 for the hard-core model on grid graphs, and exploits combinatorial and geometrical features of admissible configurations. Our findings are summarized in the next theorem.

In the rest of the chapter we will repeatedly use the notions of path ω and its height Φ_ω, as well as that of communication height $\Phi(\cdot,\cdot)$ between two states or subset of states, which are all defined in Subsection 4.2.1.

Theorem 7.3.1 (Energy landscape properties for the Widom-Rowlison model on grid graphs)**.** *Let $(\mathcal{X}, H, q)$ be the energy landscape corresponding to the Metropolis dynamics of the Widom-Rowlison model on a $K \times L$ toric grid graph $\Lambda^T_{K,L}$ with $K \le L$ and $K + L \ge 6$. Then,*

(i) $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) = \min\{2K + 1, 2L\}$;

(ii) $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}\}) - H(\sigma) < \min\{2K + 1, 2L\} \quad \forall\, \sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}\}$.

If instead $(\mathcal{X}, H, q)$ is the energy landscape corresponding to the Metropolis dynamics of the Widom-Rowlison model on a $K \times L$ open grid graph $\Lambda^O_{K,L}$, then,

(iii) $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) = \min\{K, L\} + 1$;

(iv) $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}\}) - H(\sigma) \le \min\{K, L\} \quad \forall\, \sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}\}$.

Statements (i) and (ii) of Theorem 7.3.1 are proved in Section 7.4, while the proofs of statements (iii) and (iv) are presented in Section 7.5.

We remark that the two toric grid graphs $\Lambda^T_{2,2}$ and $\Lambda^T_{2,3}$ are not covered by the previous theorem, since the are special cases for which our proof method does not work. Nonetheless, it is easy to find the communication height between **a** and **b** and show that the corresponding energy landscapes exhibit absence of deep cycles as well, so that Theorem 7.2.1 holds also in these cases.

The rest of this section is organized as follows. In Subsection 7.3.1 we show how the three main results presented in Section 7.2 follows from Theorem 7.3.1. In Subsection 7.3.2 we introduce some useful notation and definitions and prove some results concerning geometrical and combinatorial properties of admissible Widom-Rowlison configurations on the grid graph $\Lambda_{K,L}$. These properties are then immediately used to show that **a** and **b** are the only two stable states of the Widom-Rowlison model on grid graphs, see Proposition 7.3.5. Several of these geometrical properties of admissible configurations will be heavily used in Sections 7.4 and 7.5.

7.3.1 *Proofs of Theorems 7.2.1, 7.2.2 and 7.2.3*

In this subsection we briefly illustrate how the three main theorems presented in Section 7.2 are proved. The proof method combines the model-independent results derived in Chapter 4 with the structural properties of the energy landscape corresponding to the Widom-Rowlison model on grid graphs.

We first show how the tunneling time $\tau_{\mathbf{b}}^{\mathbf{a}}$ falls precisely in the scenario described in Example 2 in Subsection 4.2.5 for both types of boundary conditions. We illustrate how to argue only in the case of the toric grid $\Lambda_{K,L}^{T}$, since in the case of an open grid the reasoning is almost identical. Theorem 7.3.1(ii) shows that for every configuration $\sigma \neq \mathbf{a}, \mathbf{b}$ there is a path from σ either directly to $\mathbf{b}$ and such that $\Phi(\sigma, \mathbf{b}) - H(\sigma) < \min\{2K+1, 2L\}$ or a path from σ to $\mathbf{a}$ such that $\Phi(\sigma, \mathbf{a}) - H(\sigma) = 0$. Using statement (i) of Theorem 7.3.1, we deduce that also in this latter case the inequality $\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq \min\{2K+1, 2L\}$ holds and thus $\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq \Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a})$ for every configuration $\sigma \neq \mathbf{a}$, proving that

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}\}) = \Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}). \tag{7.7}$$

In other words, we have shown that the Widom-Rowlison model on grid graphs exhibits absence of deep cycles. Therefore, Assumption A holds for the pair $(\mathbf{a}, \mathbf{b})$ in view of Proposition 4.2.18, and statements (i) and (ii) of Theorem 7.2.1 then follow from Corollary 4.2.16 and Theorem 4.2.17, respectively.

The value of the exponent $\Gamma(\Lambda)$ appearing in Theorem 7.2.1 for the two different types of grid graphs is immediately derived from the identities presented in Theorem 7.3.1(i) and (iii), since $\Gamma(\Lambda) = \Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) = \widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{b}\})$. This concludes the proof of Theorem 7.2.2.

Furthermore, Theorem 7.3.1 directly implies that the following inequality holds for the energy landscape of the Widom-Rowlison model on a grid graph Λ, regardless of the boundary conditions:

$$\widetilde{\Gamma}(\mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}\}) < \Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}).$$

In view of this latter inequality and of the fact that $\mathcal{X}^s = \{\mathbf{a}, \mathbf{b}\}$, we are precisely in the scenario illustrated in Example 4 in Subsection 4.2.6. In particular, Assumption B holds and, by applying Theorem 4.2.19, we obtain the conclusion of Theorem 7.2.1(iii).

Lastly, Proposition 4.2.24, in combination with (7.7), proves Theorem 7.2.3.

7.3.2 *Geometrical features of admissible configurations*

In this subsection we present some useful notation and results that are valid for every grid graph, regardless of the imposed boundary conditions.

Consider a $K \times L$ grid graph Λ as described in Section 7.2. We denote by c_j, $j = 0, \dots, L-1$, the j-th column of Λ, i.e. the collection of sites whose horizontal coordinate is equal to j, and by r_i, $i = 0, \dots, K-1$, the i-th row of Λ,

i.e. the collection of sites whose vertical coordinate is equal to i, see Figure 7.4. When we do not need to specify a precise row or column, we will denote a generic column by c and a generic row by r. We assign the coordinates (j, i) to the vertex v that lies at the intersection of column c_j and row r_i. In addition, define the *i-th horizontal stripe*, with $i = 1, \ldots, \lfloor K/2 \rfloor$, as

$$S_i := r_{2i-2} \cup r_{2i-1},$$

and the *j-th vertical stripe*, with $j = 1, \ldots, \lfloor L/2 \rfloor$ as

$$C_j := c_{2j-2} \cup c_{2j-1},$$

see Figure 7.4 for an illustration. Given an admissible configuration $\sigma \in \mathcal{X}$,

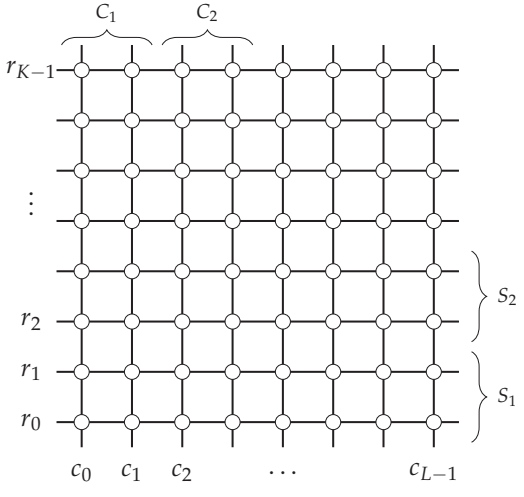


Figure 7.4: Illustration of row, column and stripe notation on a toric grid

we define its *energy wastage* $U(\sigma)$ the number of the empty sites that σ has as the difference between the energy of σ and that of $\mathbf{a}$, i.e.

$$U(\sigma) := N - \sum_{v \in V} \mathbb{1}_{\{\sigma(v) \neq 0\}}. \tag{7.8}$$

Note that $U(\sigma)$ can equivalently be seen as the energy difference between configuration σ and either $\mathbf{a}$ or $\mathbf{b}$, since $U(\sigma) = H(\sigma) - H(\mathbf{a})$. Furthermore, we define the energy wastage of a configuration $\sigma \in \mathcal{X}$ in row r by

$$U^r(\sigma) := L - \sum_{v \in r} \mathbb{1}_{\{\sigma(v) \neq 0\}} = L - \sum_{v \in r} |\sigma(v)|, \tag{7.9}$$

and the energy wastage of a configuration $\sigma \in \mathcal{X}$ in column c by

$$U^c(\sigma) := K - \sum_{v \in c} \mathbb{1}_{\{\sigma(v) \neq 0\}} = K - \sum_{v \in c} |\sigma(v)|. \tag{7.10}$$

Clearly the energy wastage (7.8) of a configuration σ can be written as the sum of the energy wastages in every row (or in every column). Given two configurations $\sigma, \sigma' \in \mathcal{X}$ and a subset of sites $W \subset \Lambda$, we write

$$\sigma_{|W} = \sigma'_{|W} \iff \sigma(v) = \sigma'(v) \quad \forall\, v \in W.$$

We say that a configuration $\sigma \in \mathcal{X}$ displays:

- A *vertical bridge* in column c if all sites of column c are occupied by particles of the same type, namely
$$\sigma_{|c} = \mathbf{a}_{|c} \quad \text{or} \quad \sigma_{|c} = \mathbf{b}_{|c};$$
- A *horizontal bridge* in row r if all sites of row r are occupied by particles of the same type, namely
$$\sigma_{|r} = \mathbf{a}_{|r} \quad \text{or} \quad \sigma_{|r} = \mathbf{b}_{|r};$$
- A *vertical quasi-bridge* in column c if all sites but one of column c are occupied by particles of the same type, i.e. if there exists $v \in c$ such that $\sigma(v) = 0$ and
$$\sigma_{|c\setminus\{v\}} = \mathbf{a}_{|c\setminus\{v\}} \quad \text{or} \quad \sigma_{|c\setminus\{v\}} = \mathbf{b}_{|c\setminus\{v\}};$$
- A *horizontal quasi-bridge* in row r if all sites but one of row r are occupied by particles of the same type, i.e. if there exists $v \in r$ such that $\sigma(v) = 0$ and
$$\sigma_{|r\setminus\{v\}} = \mathbf{a}_{|r\setminus\{v\}} \quad \text{or} \quad \sigma_{|r\setminus\{v\}} = \mathbf{b}_{|r\setminus\{v\}};$$
- A *cross* if it has both a vertical bridge and a horizontal bridge;
- A *quasi-cross* if it has both a vertical quasi-bridge and an horizontal quasi-bridge.

We will talk of *gray* bridges when they consist of particles of type A (i.e. they agree with $\mathbf{a}$), and of *black* bridges when they consist of particles of type B (i.e. they agree with $\mathbf{b}$). Analogously, we distinguish between *black* and *gray* quasi-bridges and crosses. Some examples are given in Figures 7.5 and 7.6.

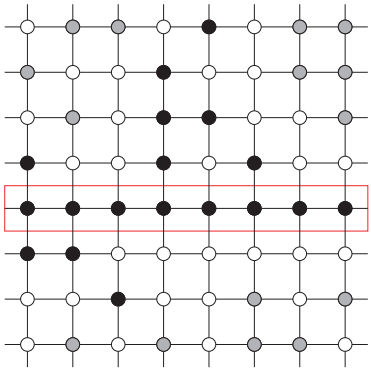
Lemma 7.3.2 (Geometric features of admissible configurations)**.** *In an admissible configuration $\sigma \in \mathcal{X}$ the following properties hold:*

(i) *Two bridges of different colors cannot be perpendicular to each other;*

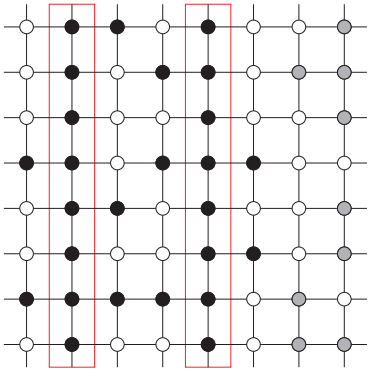
(ii) *A bridge and quasi-bridge of different colors cannot be perpendicular to each other;*

(iii) *Two quasi-bridges of different colors and perpendicular to each other must meet in their empty site;*

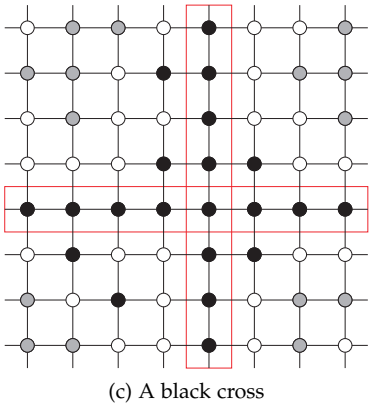
(iv) *A (quasi-)bridge can have another (quasi-)bridge in adjacent row/column only if they are of the same color.*



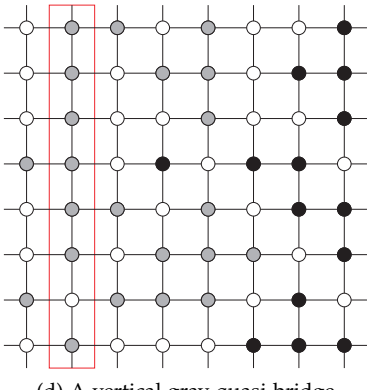
(a) A black horizontal bridge



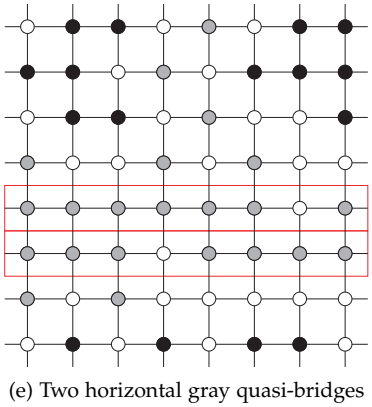
(b) Two vertical black bridges



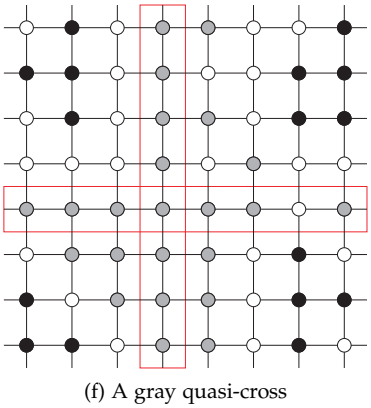
(c) A black cross



(d) A vertical gray quasi-bridge



(e) Two horizontal gray quasi-bridges

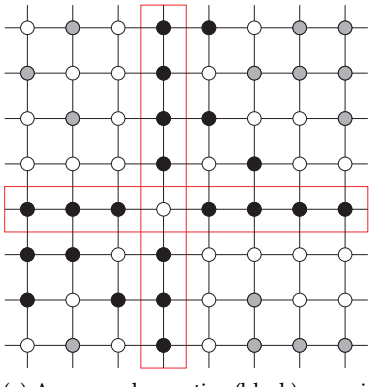


(f) A gray quasi-cross

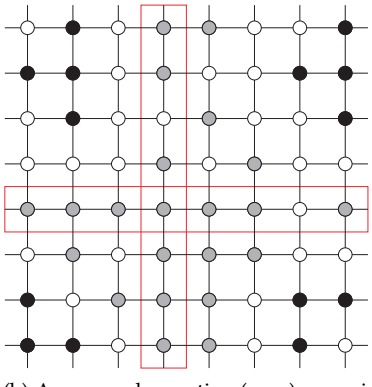
Figure 7.5: Examples of configurations on the 8×8 toric grid displaying bridges and quasi-bridges

The proof of Lemma 7.3.2 is immediate by looking at the hard-core constraints between unlike particles arising at the site in which the two bridges or quasi-bridges meet for statements (i)-(iii) and at the neighboring sites belonging to adjacent rows/columns for statement (iv), see for instance Figure 7.5.

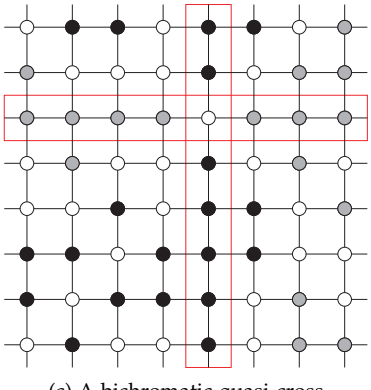
We call a quasi-cross *monochromatic* if the two quasi-bridges it consists of are of the same color, and *bichromatic* otherwise. Lemma 7.3.2(iii) implies that the two quasi-bridges of a bichromatic quasi-cross must intersect in their unique empty site. Furthermore, in view of Lemma 7.3.2(i), a cross can only be monochromatic and thus this distinction is not relevant for crosses.



(a) A monochromatic (black) quasi-cross



(b) A monochromatic (gray) quasi-cross



(c) A bichromatic quasi-cross

Figure 7.6: Examples of configurations on the 8×8 toric grid displaying quasi-crosses

The next two lemmas show that bridges are the unique particle displacements with zero energy wastage in a given row/column, regardless of the chosen boundary conditions. In addition, only in the case of toric grid graphs, we give an equivalent characterization of quasi-bridges. Lemmas 7.3.3 and 7.3.4 are stated and proved for rows and horizontal (quasi-)bridges, since those for columns and vertical (quasi-)bridges are analogous.

Lemma 7.3.3 (Bridges and quasi-bridges characterization on toric grid graphs)**.** *Let σ be an admissible Widom-Rowlison configuration on a toric grid graph $\Lambda^T_{K,L}$. Then,*

(i) $U^r(\sigma) = 0 \iff$ *σ has a horizontal bridge in row r;*

(ii) $U^r(\sigma) = 1 \iff$ *σ has a horizontal quasi-bridge in row r.*

In particular, if σ has no bridges nor quasi-bridges in row r, then $U^r(\sigma) \geq 2$.

Proof. The implications (i)$\Leftarrow$ and (ii)$\Leftarrow$ are immediate by definition of (quasi-)bridge and (7.9). For the converse implications, is enough to observe that, in order to have particles of different types in the same row r, there are at least two empty sites separating each gray cluster from the black particles, due to the toric boundary conditions, and, hence, one would have $U^r(\sigma) \geq 2$. Thus, all particles residing in row r are of the same type and their number is automatically determined by the value of the energy wastage $U^r(\sigma)$. □

Lemma 7.3.4 (Characterization of bridges in open grid graphs)**.** *Let σ be an admissible Widom-Rowlison configuration on an open grid graph $\Lambda^O_{K,L}$. Then,*

$$U^r(\sigma) = 0 \iff \text{σ has a horizontal bridge in row } r.$$

Proof. The implication $\Leftarrow$ is immediate in view of (7.9) and the definition of a bridge. Assume by contradiction that σ does not have a horizontal bridge in row r. If σ has only particles of one color in row r and does not display a bridge there, then the number of particle in row r must be strictly less than L, so by definition of energy wastage $U^r(\sigma) > 0$, contradiction. If instead σ has particles of both types in row r, there has to be at least one empty site separating the gray cluster(s) from the black cluster(s), and thus $U^r(\sigma) \geq 1$, which is again a contradiction. □

The fact that $\mathbf{a}$ and $\mathbf{b}$ are the only two stable states of the Widom-Rowlison model on grid graphs is an immediate consequence of the previous two lemmas and is formalized in the next proposition.

Proposition 7.3.5 (Stable states of the Widom-Rowlison model on grid graphs)**.** *For any $K \times L$ grid graph Λ, regardless of the boundary conditions,*

$$\min_{\sigma \in \mathcal{X}} H(\sigma) = -KL \quad \text{and} \quad \mathcal{X}^s = \{\mathbf{a}, \mathbf{b}\}.$$

Proof. At most one particle can reside in each site, so there cannot be more than N particles on Λ and trivially $\min_{\sigma \in \mathcal{X}} H(\sigma) \geq -N$. Moreover, by definition of $\mathbf{a}$ and $\mathbf{b}$ it follows that $H(\mathbf{a}) = -N = H(\mathbf{b})$. To conclude the proof, we then only need to show that $H(\sigma) > -N$ for every $\sigma \in \mathcal{X} \setminus \{\mathbf{a}, \mathbf{b}\}$. Since $\sigma \neq \mathbf{a}, \mathbf{b}$, there exists a row r such that $\sigma_{|r} \neq \mathbf{a}_{|r}$ and $\sigma_{|r} \neq \mathbf{b}_{|r}$. In view of Lemma 7.3.3 (if Λ is a toric grid) and Lemma 7.3.4 (if Λ is an open grid), the energy wastage in row r is strictly positive and thus $H(\sigma) = U(\sigma) - N > -N$. □

7.4 PROOFS FOR TORIC GRIDS

This section is devoted to the analysis of the energy landscape of the Widom-Rowlison model on the toric grid graph $\Lambda^T_{K,L}$, which leads to the proof of Theorem 7.3.1(i) and (ii).

The section is organized as follows. First, we present a lower bound for the communication height between the two stable states $\mathbf{a}$ and $\mathbf{b}$, see Proposition 7.4.1. We then introduce a *reduction algorithm*, which is then used in Proposition 7.4.2 to build a reference path $\omega^* : \mathbf{a} \to \mathbf{b}$, which shows that the lower bound given in Proposition 7.4.1 is sharp, concluding the proof of Theorem 7.3.1(i). Lastly, we use again the reduction algorithm to construct a specific path from every admissible configuration $\sigma \neq \mathbf{a}, \mathbf{b}$ to one of the two stable states and prove in this way Theorem 7.3.1(ii).

Proposition 7.4.1 (Lower bound for $\Phi(\mathbf{a}, \mathbf{b})$)**.** *Consider the Widom-Rowlison model on the $K \times L$ toric grid $\Lambda^T_{K,L}$ with $K \leq L$ and $(K, L) \neq (2, 2), (2, 3)$. The communication height between $\mathbf{a}$ and $\mathbf{b}$ in the corresponding energy landscape satisfies*

$$\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) \geq \begin{cases} 2K & \text{if } K = L, \\ 2K + 1 & \text{if } K < L. \end{cases} \tag{7.11}$$

Proof. We need to show that in every path $\omega : \mathbf{a} \to \mathbf{b}$ there is at least one configuration with energy wastage greater than or equal to $\min\{2K + 1, 2L\}$. Take a path $\omega = (\omega_1, \dots, \omega_n)$ from $\mathbf{a}$ to $\mathbf{b}$. Without loss of generality, we may assume that there are no void moves in ω, i.e. at every step either a particle is added or a particle is removed, so that $H(\omega_{i+1}) = H(\omega_i) \pm 1$ for every $1 \leq i \leq n - 1$. By virtue of definition (7.5), if two admissible configurations $\sigma, \sigma' \in \mathcal{X}$ with $\mathfrak{d}(\sigma, \sigma') = 1$ are such that σ does not display a black bridge in a certain row/column and σ' instead does, then σ must have a quasi-bridge in that row/column. Moreover in this case $H(\sigma') = H(\sigma) + 1$, since the bridge is created by adding a particle of type B in the only empty site of that row/column. Since $\mathbf{a}$ has no black bridges, while $\mathbf{b}$ does, at some point along the path ω there must be an index m^* such that configuration ω_{m^*} that is *the first* to display a black bridge (horizontal or vertical) or a black quasi-cross. Clearly $m^* > 2$. We claim that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq \min\{2K + 1, 2L\}.$$

We distinguish the following three cases:

(a) ω_{m^*} displays a black vertical bridge only;

(b) ω_{m^*} displays a black horizontal bridge only;

(c) ω_{m^*} displays a black quasi-cross.

Note that in case (c) we do not exclude the possibility that a black quasi-cross is created simultaneously with a black bridge.

For case (a), let c^* be the column where ω_{m^*} displays the black vertical bridge. The previous configuration ω_{m^*-1} along the path ω differs from ω_{m^*} in exactly one site, say $v' \in c^*$. By construction, $\omega_{m^*-1}(v') = 0$ and ω_{m^*-1} has a black vertical quasi-bridge in column c^*. In any row the configuration ω_{m^*-1} cannot display

- a black bridge, by definition of m^*;
- a black quasi-bridge, since otherwise ω_{m^*-1} would have a black quasi-cross together with the quasi-bridge in column c^*, violating the definition of m^*;
- a gray bridge, which cannot coexist with the black vertical quasi-bridge, in view of Lemma 7.3.2(ii);
- a gray quasi-bridge, since the black bridge in column c^* could not be created with a single-site update, as illustrated in Figure 7.7.

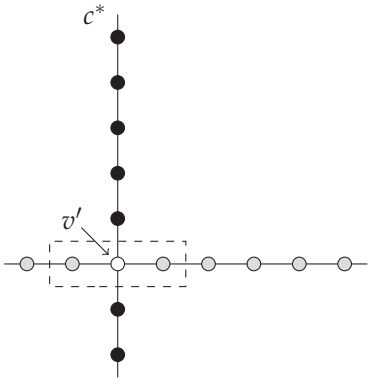


Figure 7.7: Three single-site updates are needed to create a black bridge in column c^*, since all the three sites in the dashed box must be updated

Therefore, by Lemma 7.3.3 $U^{r_i}(\omega_{m^*-1}) \geq 2$ for every $i = 0, \dots, K-1$ and hence

$$U(\omega_{m^*-1}) \geq \sum_{i=0}^{K-1} U^{r_i}(\omega_{m^*-1}) \geq 2K. \tag{7.12}$$

If $U(\omega_{m^*-1}) \geq 2K+1$, then the proof is complete. Suppose instead that $U(\omega_{m^*-1}) = 2K$ and consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, ω_{m^*-2} differs from ω_{m^*-1} by a single-site update, and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.13}$$

Suppose first that $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K - 1$, which means that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. The site where such a particle is added cannot be v', otherwise the definition of m^* would be violated. All the other sites in column c^* are already occupied, hence ω_{m^*-2} is identical to ω_{m^*-1} in column c^*. In particular, ω_{m^*-2} has a black quasi-bridge in column c^* as well. The configuration ω_{m^*-2} cannot have any horizontal bridge, since the existence of a black bridge would contradict the definition of m^* and that of a gray bridge is impossible by Lemma 7.3.2(ii). Since $U(\omega_{m^*-2}) = 2K - 1$, by the pigeonhole principle there exists then at least one row, say r', with $U^{r'}(\omega_{m^*-2}) = 1$, which means that ω_{m^*-2} has a horizontal quasi-bridge in row r'. Such a horizontal quasi-bridge cannot be black, otherwise ω_{m^*-2} would have a quasi-cross, violating the definition of m^*. By Lemma 7.3.2(iii) and the presence of a black quasi-bridge in column c^*, a gray quasi-bridge could exist only in the row containing site v'. However, in this case it would then be impossible to obtain a black vertical bridge in only two single-site updates, since the minimum number of steps required is three, as illustrated in Figure 7.7. Therefore, it is not possible that $U(\omega_{m^*-2}) = 2K - 1$ and, combining (7.12) and (7.13), we deduce that

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) + 1 = 2K + 1,$$

which concludes the proof of case (a).

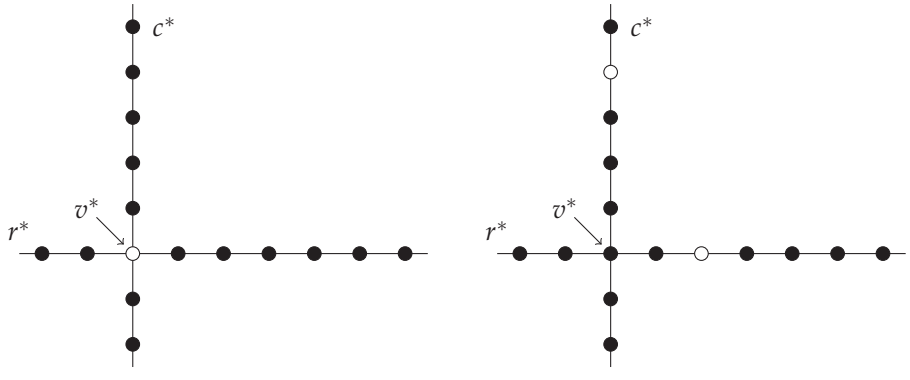
For case (b) we can argue as in case (a), but interchanging the role of rows and columns, and obtain that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq 2L + 1.$$

For case (c), let r^* and c^* be respectively the row and the column where the black quasi-cross lies in configuration ω_{m^*} and let v^* the site where they intersect. We distinguish two scenarios: (c1) the quasi-cross has exactly one empty site, which has to be v^*, and (c2) the quasi-cross has exactly two empty sites both different from v^*. Figure 7.8 illustrates these two possible scenarios.

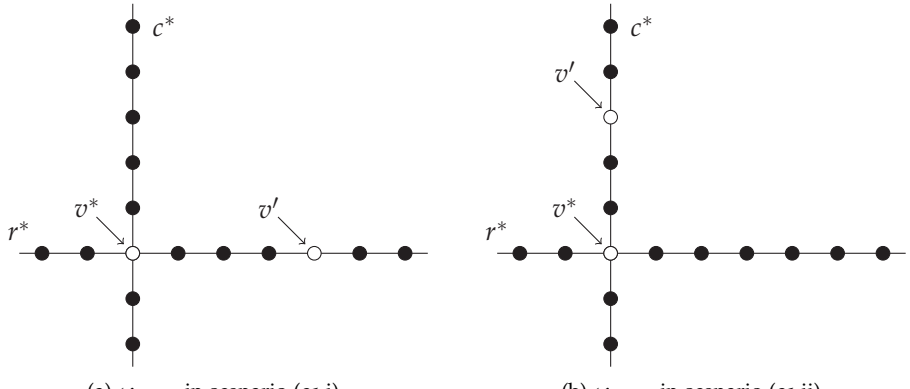
Consider scenario (c1) first. The previous configuration ω_{m^*-1} along the path ω differs from ω_{m^*} in exactly one site, say v'. By definition of m^*, configuration ω_{m^*-1} does not display a quasi-cross, so such a site v' has to lie either in row r^*, to which we refer as case (c1.i) (see Figure 7.9a), or in column c^*, to which we refer as case (c1.ii) (see Figure 7.9b). Furthermore, note that the quasi-cross present in ω_{m^*} could only have been created by the addition of a black particle, hence $\omega_{m^*-1}(v') = 0$ and $\omega_{m^*}(v') = \mathbf{b}(v')$.

In case (c1.i) ω_{m^*-1} is such that v' lies in row r^*, as in Figure 7.9a. Then $U^{r^*}(\omega_{m^*-1}) = 2$, since row r^* has exactly two empty sites, and $U^r(\omega_{m^*-1}) \geq 2$ for all rows $r \neq r^*$, since none of them can display a black bridge or quasi-



(a) ω_{m^*} in case (c1) (b) ω_{m^*} in case (c2)

Figure 7.8: Schematic representation of the two possible scenarios when the configuration ω_{m^*} displays a black quasi-cross



(a) ω_{m^*-1} in scenario (c1.i) (b) ω_{m^*-1} in scenario (c1.ii)

Figure 7.9: Schematic representation of configuration ω_{m^*-1} in case (c1)

bridge (by definition of m^*) and neither a gray bridge or quasi bridge (by Lemma 7.3.2). Hence,

$$U(\omega_{m^*-1}) = \sum_{i=0}^{K-1} U^{r_i}(\omega_{m^*-1}) \geq 2K.$$

If $U(\omega_{m^*-1}) \geq 2K+1$, then the proof of case (c1.a) is complete. Suppose instead that $U(\omega_{m^*-1}) = 2K$ and consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, the configuration ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.14}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K - 1$, which means that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. By virtue of the pigeonhole principle, the configuration ω_{m^*-2} must have at least one row, say r', such that $U^{r'}(\omega_{m^*-2}) \leq 1$. In view of Lemma 7.3.3, ω_{m^*-2} then has to display a bridge or a quasi-bridge in row r', which leads to a contradiction, since on this row ω_{m^*-2} cannot have

- a black horizontal bridge, by definition of m^*;
- a black horizontal quasi-bridge, since otherwise ω_{m^*-2} would have a quasi-cross in row r' and column c^*, violating the definition of m^*;
- a gray horizontal bridge or quasi-bridge, due to the presence of the black vertical quasi-bridge in column c^* in view of Lemma 7.3.2.

Hence, since $U(\omega_{m^*-1}) = 2K$, we deduce from (7.14) that

$$U(\omega_{m^*-2}) = 2K + 1,$$

which concludes the proof of case (c1.i).

In case (c1.ii) we can argue similarly to case (c1.i), by interchanging the role of rows and columns, and obtain that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq 2L + 1 \geq 2K + 1.$$

Consider now case (c2). We distinguish three scenarios, illustrated in Figure 7.10, depending where the last particle (that created the quasi-cross in configuration ω_{m^*}) has been added: (c2.i) in a site $v' \neq v^*$ in row r^* or (c2.ii) in a site $v' \neq v^*$ in column c^* or (c2.iii) in the site $v' = v^*$.

In case (c2.i), let c' be the column where site v' lies. We first notice that configuration ω_{m^*-1} cannot have a vertical gray bridge or quasi-bridge in column c', since otherwise it would not be possible to add a black particle in site v', see Figure 7.10a.

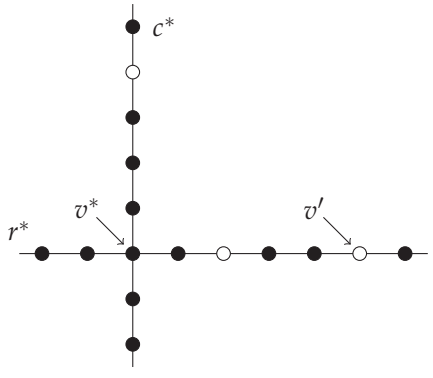
Moreover ω_{m^*-1} has no horizontal black quasi-bridges, since any of them would create, together with column c^*, a quasi-cross, violating the definition of m^*.

Suppose first that configuration ω_{m^*-1} has a vertical black quasi-bridge in column c', as in Figure 7.11, which means that

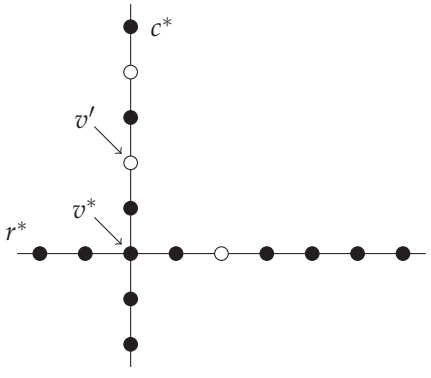
$$U^{c'}(\omega_{m^*-1}) = 1. \tag{7.15}$$

By virtue of Lemma 7.3.2, there cannot be any horizontal gray bridges or quasi-bridges. Furthermore, ω_{m^*-1} has no horizontal black quasi-bridges, which would create a quasi-cross together with column c', violating again the definition of m^*. In view of Lemma 7.3.3, $U^r(\omega_{m^*-1}) \geq 2$ for every row r and thus

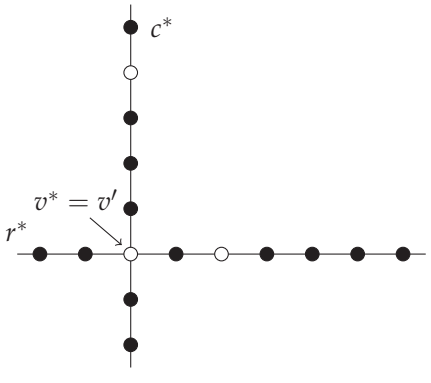
$$U(\omega_{m^*-1}) \geq 2K.$$



(a) ω_{m^*-1} in scenario (c2.i)



(b) ω_{m^*-1} in scenario (c2.ii)



(c) ω_{m^*-1} in scenario (c2.iii)

Figure 7.10: Schematic representation of the three possible scenarios in case (c2)

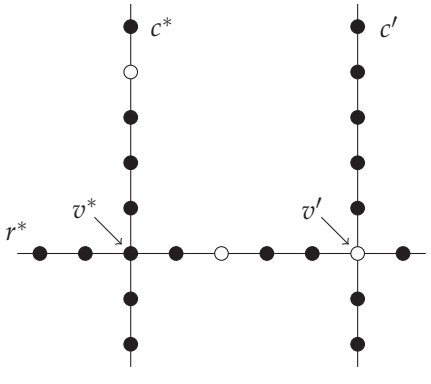


Figure 7.11: Schematic representation of configuration ω_{m^*-1} in scenario (c2.i) with a quasi-bridge in column c'

If $U(\omega_{m^*-1}) \geq 2K+1$, then the proof is complete. Suppose instead that $U(\omega_{m^*-1}) = 2K$ and consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.16}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K-1$, which means that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. By virtue of the the pigeonhole principle, the configuration ω_{m^*-2} must have at least one row, say r', such that $U^{r'}(\omega_{m^*-2}) \leq 1$. In view of Lemma 7.3.3, ω_{m^*-2} has then to display a bridge or a quasi-bridge in row r'. If $r' = r^*$, then ω_{m^*-2} would have a black quasi-cross or a black bridge, violating the definition of m^*. If $r' \neq r^*$, then we also obtain a contradiction, since in row r' ω_{m^*-2} cannot have

- a black bridge, by definition of m^*;
- a black quasi-bridge, since otherwise ω_{m^*-2} would have a quasi-cross in row r' and column c^*, violating the definition of m^*;
- a gray bridge or quasi-bridge, due to the presence of a black quasi-bridge in column c^* and Lemma 7.3.2.

Hence, $U(\omega_{m^*-2}) \neq 2K-1$, and from (7.16) it follows that

$$U(\omega_{m^*-2}) = 2K+1.$$

Suppose now that configuration ω_{m^*-1} does not have a vertical black quasi-bridge in column c'. By virtue of Lemma 7.3.3, $U^{c'}(\omega_{m^*-1}) \geq 2$. We first consider the case where

$$U^{c'}(\omega_{m^*-1}) = 2, \tag{7.17}$$

so that $U^{c'}(\omega_{m^*}) = 1$, which means that ω_{m^*} has an additional quasi-cross, namely the one lying in row r^* and column c', see Figure 7.12. In this case, we can conclude the proof by looking at this other quasi-cross and arguing as in sub-case (3) of scenario (c2.iii), which will be presented later.

Therefore, in view of (7.15) and (7.17), we can assume

$$U^{c'}(\omega_{m^*-1}) \geq 3. \tag{7.18}$$

We then distinguish three sub-cases, depending on whether ω_{m^*-1} has (1) no vertical quasi-bridges (see Figure 7.13a) or (2) at least one gray vertical quasi-bridge and no black vertical quasi-bridges (see Figure 7.13b) or (3) at least one black vertical quasi-bridge (see Figure 7.14).

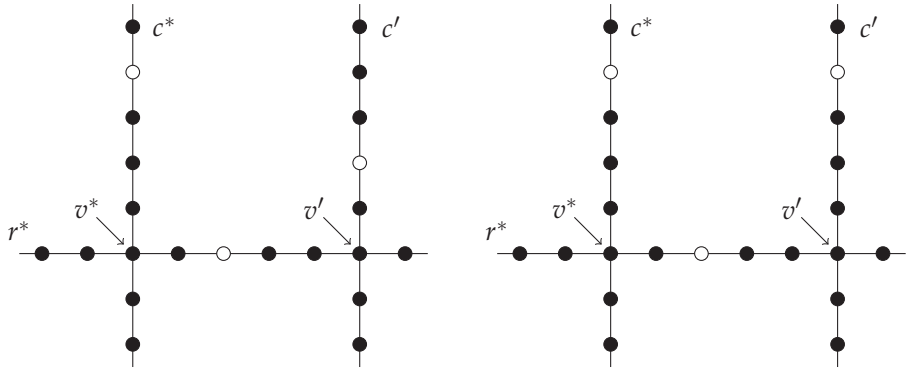
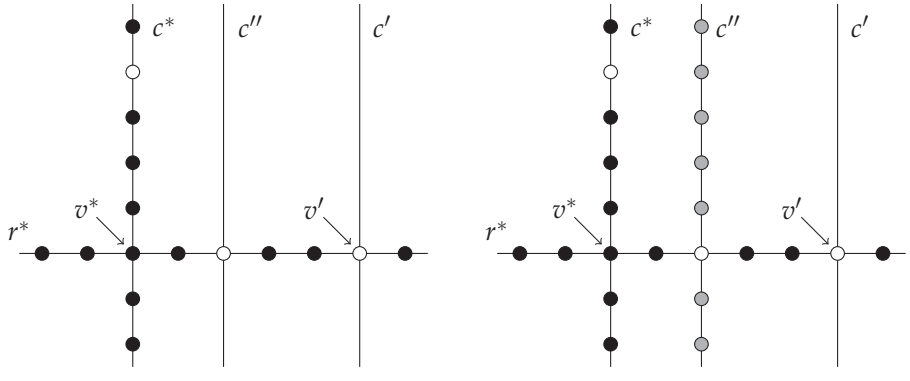


Figure 7.12: Schematic representation of configuration ω_{m^*} with two quasi-crosses



(a) ω_{m^*-1} in sub-case (1) (b) ω_{m^*-1} in sub-case (2)

Figure 7.13: Schematic representation of the sub-cases (1) and (2) of scenario (c2.i) when condition (7.18) is satisfied

In sub-case (1), ω_{m^*-1} has no vertical quasi-bridges except the one in column c^*, so by Lemma 7.3.3 $U^c(o_{m^*-1}) \geq 2$ for every $c \neq c^*$. This fact and (7.18) yield

$$U(\omega_{m^*-1}) = \sum_{j=0}^{L-1} U^{c_j}(\omega_{m^*-1}) \geq 2L,$$

and the proof of sub-case (1) is completed.

In sub-case (2), ω_{m^*-1} can have only one vertical gray quasi-bridge and it must lie in column c'' by Lemma 7.3.2, see Figure 7.13b. Lemma 7.3.3 gives

$U^{c''}(\omega_{m^*-1}) = 1$. All the columns $c \neq c^*, c'', c'''$ do not display a vertical quasi-bridge, so $U^c(\omega_{m^*-1}) \geq 2$ by Lemma 7.3.3. These facts and (7.18) yield

$$U(\omega_{m^*-1}) = \sum_{j=0}^{L-1} U^{c_j}(\omega_{m^*-1}) \geq 2L - 1.$$

If $U(\omega_{m^*-1}) \geq 2L$, the proof is complete. Suppose instead that $U(\omega_{m^*-1}) = 2L - 1$ and consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.19}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2L - 2$, which means that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. Such a particle cannot lie in site v', otherwise ω_{m^*-2} would have a quasi-cross, violating the definition of m^*. Furthermore, by virtue of the the pigeonhole principle, the configuration ω_{m^*-2} must have at least two horizontal quasi-bridges or one horizontal bridge, which cannot exists neither black or gray by Lemma 7.3.2 due to the presence of the black quasi-bridge in column c^* and the gray quasi-bridge in column c''. Hence, $U(\omega_{m^*-2}) \neq 2L - 2$, and from (7.19) it follows that

$$U(\omega_{m^*-2}) = 2L,$$

which completes the proof of sub-case (2).

Consider now sub-case (3), which is illustrated in Figure 7.14.

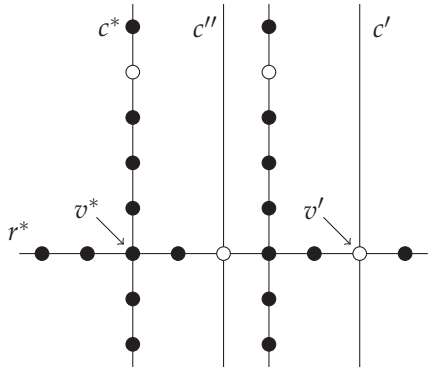
Since none of the rows can display a bridge or a quasi-bridge without contradicting Lemma 7.3.2 or violating the definition of m^*, it follows from Lemma 7.3.3 that

$$U(\omega_{m^*-1}) \geq 2K.$$

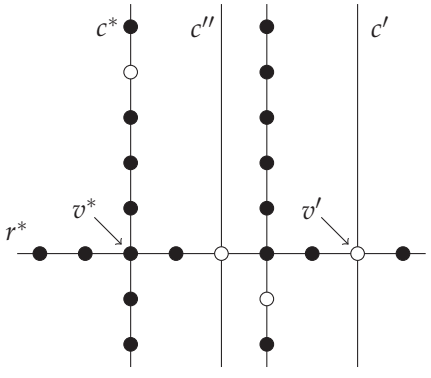
If $U(\omega_{m^*-1}) \geq 2K + 1$, the proof of the sub-case is complete. Consider now the remaining case, namely $U(\omega_{m^*-1}) = 2K$. Consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, the configuration ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.20}$$

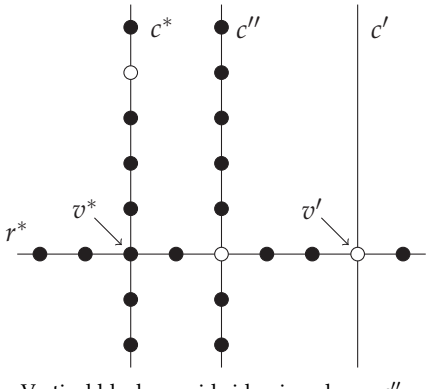
Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K - 1$, so that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1} (in a site which cannot be v'). Thanks to the pigeonhole principle, the configuration ω_{m^*-2} must have at least one row with energy wastage strictly less than 2. On such a row ω_{m^*-2} then has to display a bridge or a quasi-bridge, thanks to Lemma 7.3.3.



Vertical black quasi-bridge in a column different from c'' with empty site aligned with that of c^*



Vertical black quasi-bridge in a column different from c'' with empty site not aligned with that of c^*



Vertical black quasi-bridge in column c''

Figure 7.14: Schematic representation of ω_{m^*-1} in sub-case (3) of scenario (c2.i) when condition (7.18) is satisfied

Since ω_{m^*-2} has all the vertical black quasi-bridges that ω_{m^*-1} has, it is impossible for ω_{m^*-2} to display

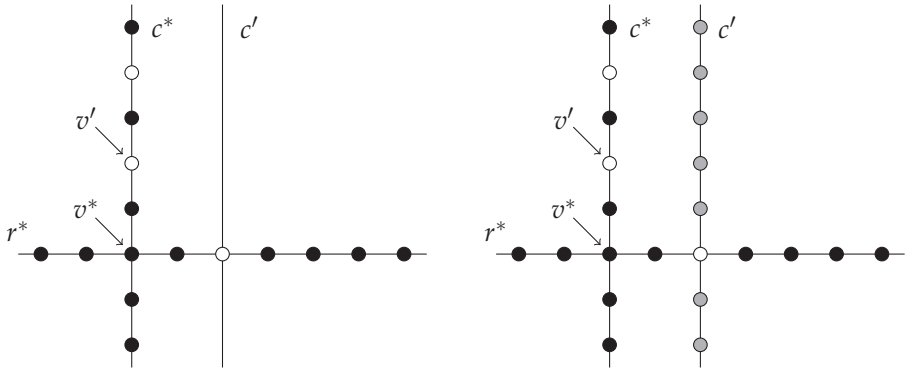
- a black horizontal bridge, by definition of m^*;
- a black horizontal quasi-bridge, since otherwise it would create a quasi-cross together with the vertical quasi-bridge in column c^*, violating the definition of m^*;
- a gray horizontal bridge by Lemma 7.3.2, due to the presence of the vertical black quasi-bridge in column c^*, see Figure 7.14;
- a gray horizontal quasi-bridge, since every row has either a black particle or two empty sites, see Figure 7.14.

Hence, $U(\omega_{m^*-2}) \neq 2K-1$, and from (7.20) it follows that

$$U(\omega_{m^*-2}) = 2K+1,$$

which completes the proof of sub-case (3).

In scenario (c2.ii), the configuration ω_{m^*-1} preceding ω_{m^*} along the path ω cannot have any vertical quasi-bridge, since it would create a quasi-cross together with row r^*. We distinguish two sub-cases, depending on whether ω_{m^*-1} has (1) no gray vertical quasi-bridges, see Figure 7.15a, or (2) at least one gray vertical quasi-bridge, see Figure 7.15b.



(a) ω_{m^*-1} in sub-case (1) (b) ω_{m^*-1} in sub-case (2)

Figure 7.15: Schematic representation of the two sub-cases in scenario (c2.ii)

In sub-case (1), by Lemma 7.3.3 $U^c(\omega_{m^*-1}) \geq 2$ in every column c and hence

$$U(\omega_{m^*-1}) \geq 2L.$$

In sub-case (2), it is clear that there can be only one gray vertical quasi-bridge, which intersects row r^* in the empty site different from v^*. By looking at the energy wastage in columns and arguing similarly to the final part of sub-case (2) of scenario (c2.i), we can conclude that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq 2L.$$

Consider now scenario (c2.iii). By definition of m^*, ω_{m^*-1} does not have any black bridge and by Lemma 7.3.3 it cannot have any gray bridge either. We distinguish three sub-cases, depending on whether ω_{m^*-1} has (1) no vertical quasi-bridges, see Figure 7.16a, (2) gray vertical quasi-bridges, but no black vertical quasi-bridges, see Figure 7.16b, or (3) black vertical quasi-bridges, see Figure 7.17.

In sub-case (1), by Lemma 7.3.3 $U^c(\omega_{m^*-1}) \geq 2$ in every column c and hence

$$U(\omega_{m^*-1}) \geq 2L.$$

In sub-case (2), ω_{m^*-1} can have only one vertical gray quasi-bridge and it must lie in column c' by Lemma 7.3.2, see Figure 7.16b. Lemma 7.3.3 gives $U^{c'}(o_{m^*-1}) = 1$. All the columns $c \neq c'$ do not display a vertical quasi-bridge, so $U^{c}(o_{m^*-1}) \geq 2$ by Lemma 7.3.3. These facts and (7.18) yield

$$U(\omega_{m^*-1}) = \sum_{j=0}^{L-1} U^{c_j}(o_{m^*-1}) \geq 2L - 1.$$

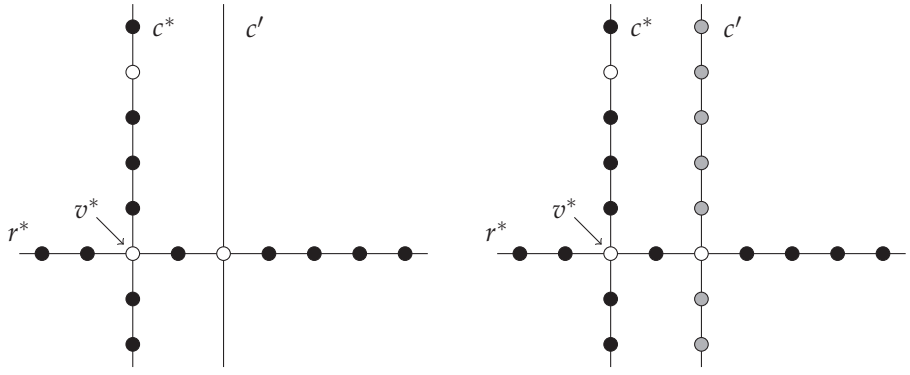
If $K < L$, then $U(\omega_{m^*-1}) \geq 2L - 1 \geq 2K + 1$ and the proof of sub-case (2) is complete. If $K = L$ and $U(\omega_{m^*-1}) \geq 2L$, then the proof is also complete. Suppose instead that $K = L$ and $U(\omega_{m^*-1}) = 2L - 1 = 2K - 1$ and consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.21}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K - 2$, which means that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. By virtue of the the pigeonhole principle, the configuration ω_{m^*-2} must have at least two horizontal quasi-bridges or one horizontal bridge, which cannot exists neither black or gray by Lemma 7.3.2 due to the presence of the black quasi-bridge in column c^* and the gray quasi-bridge in column c'. Hence, $U(\omega_{m^*-2}) \neq 2K - 1$, and from the fact that $U(\omega_{m^*-1}) = 2K - 1$ and (7.21) it follows that

$$U(\omega_{m^*-2}) = 2K,$$

which completes the proof of sub-case (2).

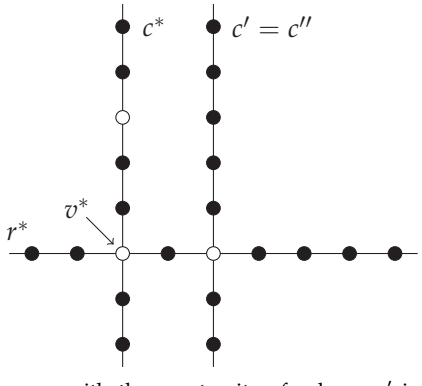


(a) ω_{m^*-1} in sub-case (1) (b) ω_{m^*-1} in sub-case (2)

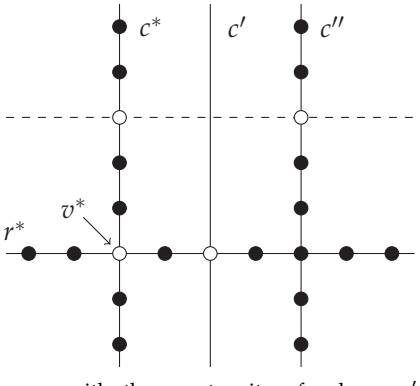
Figure 7.16: Schematic representation of sub-cases (1) and (2) in scenario (c2.iii)

In sub-case (3), illustrated in Figure 7.17 the presence of a vertical black quasi-bridge in a column $c'' \neq c^*$ means that there are no horizontal quasi-

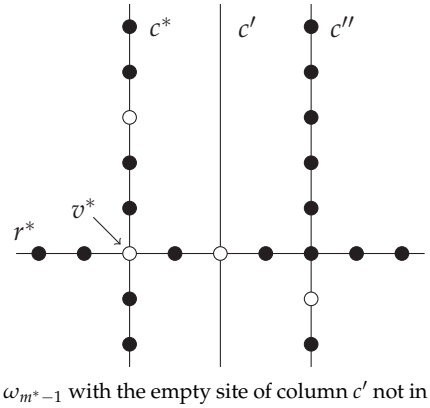
bridges in ω_{m^*-1}. Indeed the presence of a horizontal black quasi-bridge together with the quasi-bridge in column c'' would create a quasi-cross, violating the definition of m^*. Furthermore, it is impossible for ω_{m^*-1} to have a horizontal gray quasi-bridge, since every row has either a black particle or two empty sites, see Figure 7.17.



ω_{m^*-1} with the empty site of column c' in row r^*



ω_{m^*-1} with the empty site of column c' aligned with that of column c^*



ω_{m^*-1} with the empty site of column c' not in row r^* and not aligned with that of column c^*

Figure 7.17: Schematic representation of configuration ω_{m^*-1} in sub-case (3) of (c2.iii)

Since trivially ω_{m^*-1} displays no bridges either, by applying Lemma 7.3.3 to the rows, we get

$$U(\omega_{m^*-1}) \geq 2K.$$

If $U(\omega_{m^*-1}) \geq 2K+1$, then the proof of sub-case (3) is complete. Suppose instead that $U(\omega_{m^*-1}) = 2K$ and consider the configuration ω_{m^*-2} preceding

ω_{m^*-1} in the path ω. By construction, the configuration ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.22}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = 2K - 1$, so that configuration ω_{m^*-2} has exactly one more particle than configuration ω_{m^*-1}. Thanks to the pigeonhole principle, the configuration ω_{m^*-2} must have at least one row, say r' with energy wastage strictly less than 2. On such a row ω_{m^*-2} then has to display a bridge or a quasi-bridge, thanks to Lemma 7.3.3. This leads to a contradiction, since ω_{m^*-2} cannot have in row r'

- a black bridge, by definition of m^*;
- a black quasi-bridge, since otherwise ω_{m^*-2} would have a quasi-cross in row r' and column c'', violating the definition of m^*;
- a gray bridge or quasi-bridge, since every row has either a black particle or two empty sites, see Figure 7.17.

Hence, by (7.22) we obtain that

$$U(\omega_{m^*-2}) = 2K + 1. \qquad \square$$

Reduction algorithm for toric grids

We now describe an iterative procedure, to which we will refer as *reduction algorithm* that builds a path ω from a suitable initial configuration σ to the stable state $\mathbf{b}$. We require that the initial configuration σ has no gray particles in the first vertical stripe C_1 (or equivalently in the first two columns, since $C_1 = c_0 \cup c_1$), i.e.

$$\sum_{v \in C_1} \mathbb{1}_{\{\sigma(v) = \mathbf{a}(v)\}} = 0. \tag{7.23}$$

The desired path ω is built as the concatenation of L paths $\omega^{(1)}, \dots, \omega^{(L)}$. For every j along path $\omega^{(j)}$ the particles agreeing with $\mathbf{a}$ are removed from $j+1$-th column and simultaneously the j-th column is progressively changed to agree with $\mathbf{b}$. We will show that all configurations visited by ω are admissible and that the maximum energy achieved along the path ω is $H(\sigma) + 1$. Path $\omega^{(j)}$ goes from σ_j to σ_{j+1}, where we recursively define $\sigma_1 = \sigma$ and

$$\sigma_{j+1}(v) = \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_j \cup c_{j+1}), \\ \sigma_j(v) & \text{if } v \in c_{j+1} \text{ and } \sigma(v) = \mathbf{b}(v), \\ \mathbf{b}(v) & \text{if } v \in c_j, \\ 0 & \text{if } v \in c_{j+1} \text{ and } \sigma(v) = \mathbf{a}(v), \end{cases}$$

for $j = 1, \dots, L$. Clearly, due to the periodic boundary conditions, the column indices L and 0 should be identified. It can be checked that indeed $\sigma_{L+1} = \mathbf{b}$. We now describe in detail how to construct each of the paths $\omega^{(j)}$ for $j = 1, \dots, L$. We build a path $\omega^{(j)} = (\omega_1^{(j)}, \dots, \omega_{2K+1}^{(j)})$ of length $2K+1$ (but possibly with void moves), with $\omega_1^{(j)} = \sigma_j$ and $\omega_{2K+1}^{(j)} = \sigma_{j+1}$. We repeat iteratively the following procedure for every $i = 1, \dots, 2K$:

- If $i \equiv 1 \pmod 2$, consider site $v = (j+1, (i-1)/2)$.
  - If $\omega_i^{(j)}(v) = 0$ or $\omega_i^{(j)}(v) = \mathbf{b}(v)$, we set $\omega_{i+1}^{(j)}(v) = \omega_i^{(j)}(v)$.
  - If $\omega_i^{(j)}(v) = \mathbf{a}(v)$, then we remove the gray particle in site v from configuration $\omega_i^{(j)}$, obtaining in this way the configuration $\omega_{i+1}^{(j)}$ such that $\omega_{i+1}^{(j)}(v) = 0$ and thus $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) + 1$.
- If $i \equiv 0 \pmod 2$, consider site $v = (j, i/2 - 1)$.
  - If $\omega_i^{(j)}(v) = \mathbf{b}(v)$, we set $\omega_{i+1}^{(j)}(v) = \omega_i^{(j)}(v) = \mathbf{b}(v)$.
  - If $\omega_i^{(j)}(v) = 0$, then we add a black particle in site v, obtaining in this way a new configuration $\omega_{i+1}^{(j)}$ such that $\omega_{i+1}^{(j)}(v) = \mathbf{b}(v)$. This configuration is admissible because by construction there are no gray particles in any neighboring sites of v. In particular, the site at its right (that is site $v + (1,0)$) has possibly been emptied at the previous step, if there was a gray particle in it. Moreover, $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) - 1$.

Note that for the last path $\omega^{(L)}$ all steps corresponding to odd values of i are void (there are no gray particles in c_0 in configuration σ by virtue of (7.23)). For every $j = 1, \dots, L$, the configurations σ_j and σ_{j+1} and the path $\omega^{(j)}$ are constructed in such a way that

$$H(\sigma_{j+1}) \le H(\sigma_j),$$

since the number of black particles added in column c_j is greater than or equal to the number of gray particles removed in column c_{j+1}. Moreover,

$$\Phi_{\omega^{(j)}} \le H(\sigma_j) + 1,$$

since along the path $\omega^{(j)}$ every particle removal (if any) is always followed by a particle addition. These two properties imply that the path $\omega : \sigma \to \mathbf{b}$ created by concatenating $\omega^{(1)}, \dots, \omega^{(L)}$ satisfies $\Phi_\omega \le H(\sigma) + 1$, and thus $\Phi(\sigma, \mathbf{b}) - H(\sigma) \le 1$.

Note that a similar procedure can also be used to construct a path with target configuration the stable state $\mathbf{a}$, by inverting the role of gray and black

particles in the reduction algorithm we just described. It is possible to reduce a configuration σ to $\mathbf{a}$ using this procedure if

$$\sum_{v \in C_1} \mathbb{1}_{\{\sigma(v)=\mathbf{b}(v)\}} = 0, \tag{7.24}$$

that is when σ has no black particles in the first vertical stripe.

We now show how the reduction algorithm can be used to build a reference path between the two stable states. The maximum energy barrier along such a reference path matches the right-hand side of inequality (7.11), proving Theorem 7.3.1(i).

Proposition 7.4.2 (Reference path from $\mathbf{a}$ to $\mathbf{b}$). *Consider the $K \times L$ toric grid graph $\Lambda^T_{K,L}$. There exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ in $\mathcal{X}$ such that*

$$\Phi_{\omega^*} - H(\mathbf{a}) = \begin{cases} 2K & \text{if } K = L, \\ 2K+1 & \text{if } K < L. \end{cases}$$

Proof. We distinguish two cases, depending on whether (a) $K = L$ and (b) $K < L$.

Consider case (a) first. We will show that there exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ such that $\Phi_{\omega^*} - H(\mathbf{a}) = 2K = 2L$. If $K = 2$, ω^* is simply the path that gradually removes all the four gray particles that $\mathbf{a}$ has and then add four black particles one by one; one can immediately check that $\Phi_{\omega^*} - H(\mathbf{a}) = 4 = 2K$. We henceforth assume that $K \geq 3$. We construct the path ω^* as concatenation of three paths, which we denote by $\omega^{(1)}$, $\omega^{(2)}$ and $\omega^{(3)}$, respectively.

We first build a path $\omega^{(1)}$ of length $4(K-1)+3$ as follows. First remove in two steps the gray particles in sites $(0,0)$ and $(0,K-1)$, obtaining configuration $\omega^{(1)}_3$ with energy $H(\omega^{(1)}) = H(\mathbf{a}) + 2$, as illustrated in Figure 7.18. Then, iteratively define configuration $\omega^{(1)}_{4+i}$ from $\omega^{(1)}_{4+i-1}$ for $i = 0, \ldots, 4(K-2)+3$ as follows:

- If $i \equiv 0 \pmod 4$, then $\omega^{(1)}_{4+i-1}$ is the configuration obtained from $\omega^{(1)}_{4+i}$ by removing the gray particle in site $(L-1, \lfloor \frac{i}{4} \rfloor)$.
- If $i \equiv 1 \pmod 4$, then $\omega^{(1)}_{4+i-1}$ is the configuration obtained from $\omega^{(1)}_{4+i}$ by removing the gray particle in site $(1, \lfloor \frac{i}{4} \rfloor)$.
- If $i \equiv 2 \pmod 4$, then $\omega^{(1)}_{4+i-1}$ is the configuration obtained from $\omega^{(1)}_{4+i}$ by removing the gray particle in site $(0, \lfloor \frac{i}{4} \rfloor + 1)$.
- If $i \equiv 3 \pmod 4$, then $\omega^{(1)}_{4+i-1}$ is the configuration obtained from $\omega^{(1)}_{4+i}$ by adding a black particle in site $(0, \lfloor \frac{i}{4} \rfloor)$.

Note that the move corresponding to $i = 4(K-2)+2$ is void, since there is no particle in site $(0, K-1)$ to be removed, having been removed at the second

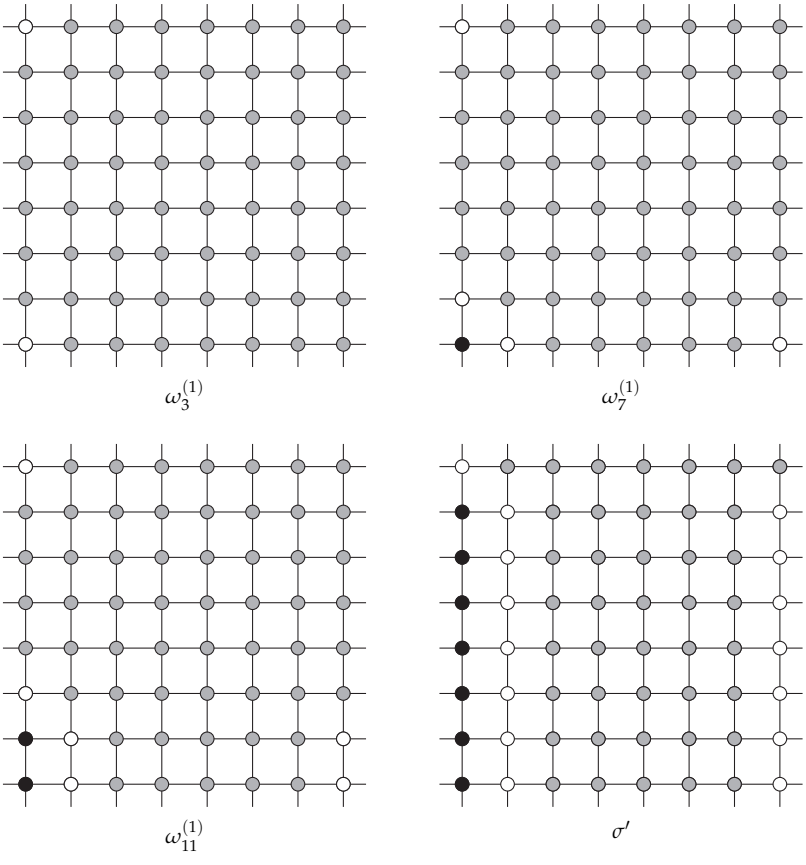


Figure 7.18: Some snapshots of the path $\omega^{(1)}$ for $\Lambda_{8,8}^T$

step of $\omega^{(1)}$. Denote by $\sigma' := \sigma^{(1)}_{4(K-1)+3}$ the configuration obtained by this procedure, which has energy wastage

$$U(\sigma') = H(\sigma') - H(\mathbf{a}) = 2K - 1.$$

The way the path $\omega^{(1)}$ is built guarantees that

$$\Phi_{\omega^{(1)}} - H(\mathbf{a}) = \max_{\eta \in \omega^{(1)}} H(\eta) - H(\mathbf{a}) = 2K. \tag{7.25}$$

If $K = 3$, the second path $\omega^{(2)}$ is not needed: Consider the configuration σ''' obtained from σ' by removing the gray particle in the site $(1, K-1)$ and thus $H(\sigma''') = H(\sigma') + 1$. The configuration σ''' satisfies the initial condition (7.23), so we can use the reduction algorithm to build the path $\omega^{(3)}$ from σ''' to $\mathbf{b}$. The concatenation of $\omega^{(1)}$ and $\omega^{(3)}$ yields a path ω^* from $\mathbf{a}$ to $\mathbf{b}$ such that

$$\Phi_{\omega^*} - H(\mathbf{a}) = 6 = 2K.$$

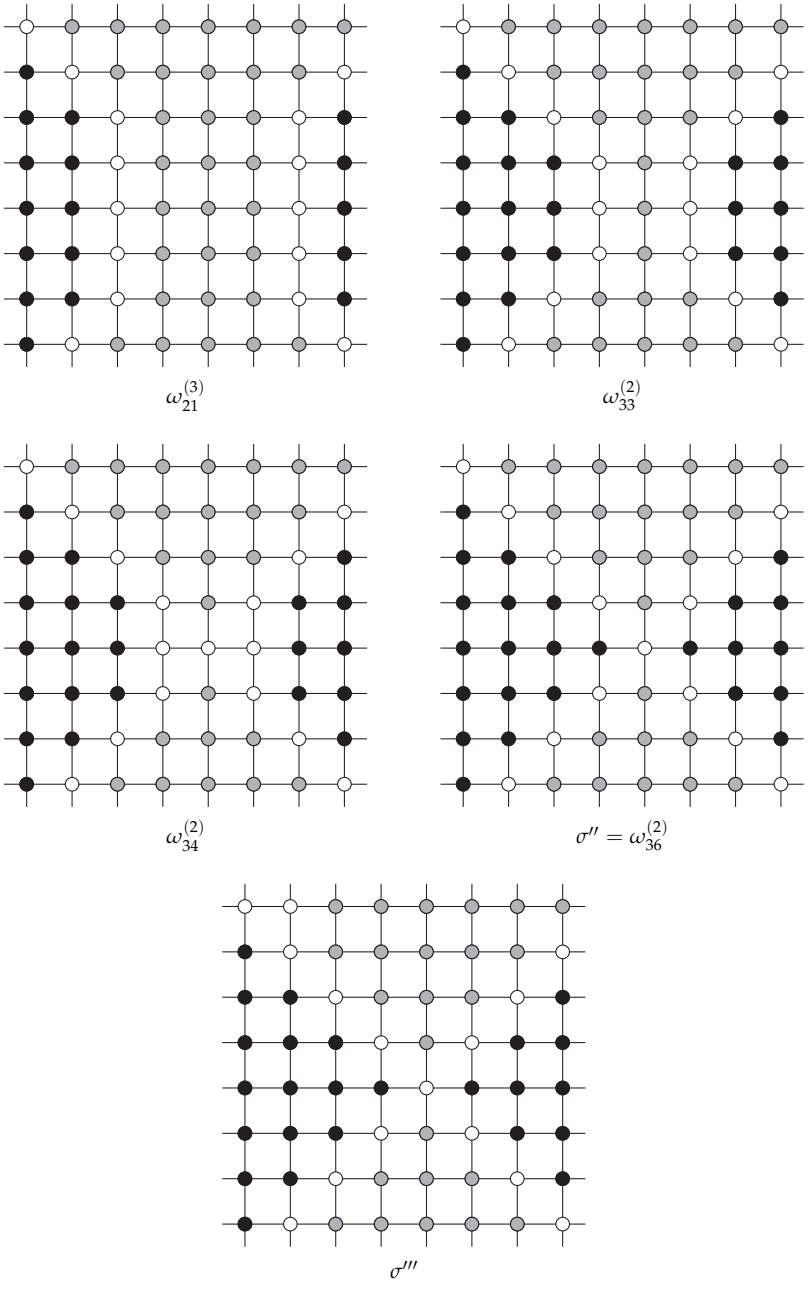


Figure 7.19: Some snapshots of the path $\omega^{(2)}$ for $\Lambda_{8,8}^{T}$

Assume now that $K \geq 4$. In this case, we build a path $\omega^{(2)}$ of length $K(K-4)+4$, which gradually enlarges the quasi-bridge (that configuration σ' has) to

obtain a configuration, say σ'', which has a "black rhombus" (see Figure 7.19) and energy wastage

$$U(\sigma'') = H(\sigma'') - H(\mathbf{a}) = 2K - 2.$$

It is possible to construct the path $\omega^{(2)}$ by alternatively removing a (gray) particle and adding immediately after a (black) particle, so that

$$\Phi_{\omega^{(2)}} - H(\sigma') = \max_{\eta \in \omega^{(2)}} H(\eta) - H(\sigma') = 1. \tag{7.26}$$

Consider now configuration σ''' obtained from σ'' by removing the gray particle in the site $(1, K-1)$ and thus $H(\sigma''') = H(\sigma'') + 1$. The configuration σ''' satisfies the initial condition (7.23), so we can use the reduction algorithm to build the path $\omega^{(3)}$ from σ''' to $\mathbf{b}$. As illustrated earlier, this procedure guarantees that

$$\Phi_{\omega^{(3)}} - H(\sigma''') = \max_{\eta \in \omega^{(3)}} H(\eta) - H(\sigma''') = 1. \tag{7.27}$$

In view of (7.25)-(7.27), the path ω^* obtained by concatenating $\omega^{(1)}$, $\omega^{(2)}$ and $\omega^{(3)}$ is such that

$$\Phi_{\omega^*} - H(\mathbf{a}) = \max_{\eta \in \omega^*} H(\eta) - H(\mathbf{a}) = 2K.$$

In case (b), where $K < L$, we will show that there exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ such that $\Phi_{\omega^*} - H(\mathbf{a}) = 2K + 1$. We construct such a path ω^* as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{a} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{b}$ for a suitable configuration $\sigma^* \in \mathcal{X}$. The configuration σ^* is the admissible configuration that differs from $\mathbf{a}$ only in the sites of the first vertical stripe (see also Figure 7.20):

$$\sigma^*(v) := \begin{cases} \mathbf{a}(v) & \text{if } v \in V \setminus C_1, \\ 0 & \text{if } v \in C_1. \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{2K+1}^{(1)})$, with $\omega_1^{(1)} = \mathbf{a}$ and $\omega_{2K+1}^{(1)} = \sigma^*$ can be constructed as follows. For $i = 1, \dots, 2K$, at step i we remove from configuration $\omega_i^{(1)}$ the particle in site $(\lfloor \frac{i}{K} \rfloor, i - K \cdot \lfloor \frac{i}{K} \rfloor)$, increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore the configuration σ^* is such that $H(\sigma^*) - H(\mathbf{a}) = 2K$ and

$$\Phi_{\omega^{(1)}} = H(\sigma^*) = H(\mathbf{a}) + 2K. \tag{7.28}$$

The second path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ is then constructed by means of the reduction algorithm presented before, which can be used since the configuration σ^* satisfies condition (7.23). The algorithm guarantees that

$$\Phi_{\omega^{(2)}} = H(\sigma^*) + 1. \tag{7.29}$$

The concatenation of $\omega^{(1)}$ and $\omega^{(2)}$ yields a path $\omega^* : \mathbf{a} \to \mathbf{b}$, that, in view of (7.28) and (7.29), satisfies $\Phi_{\omega^*} = H(\mathbf{a}) + 2K + 1$. □

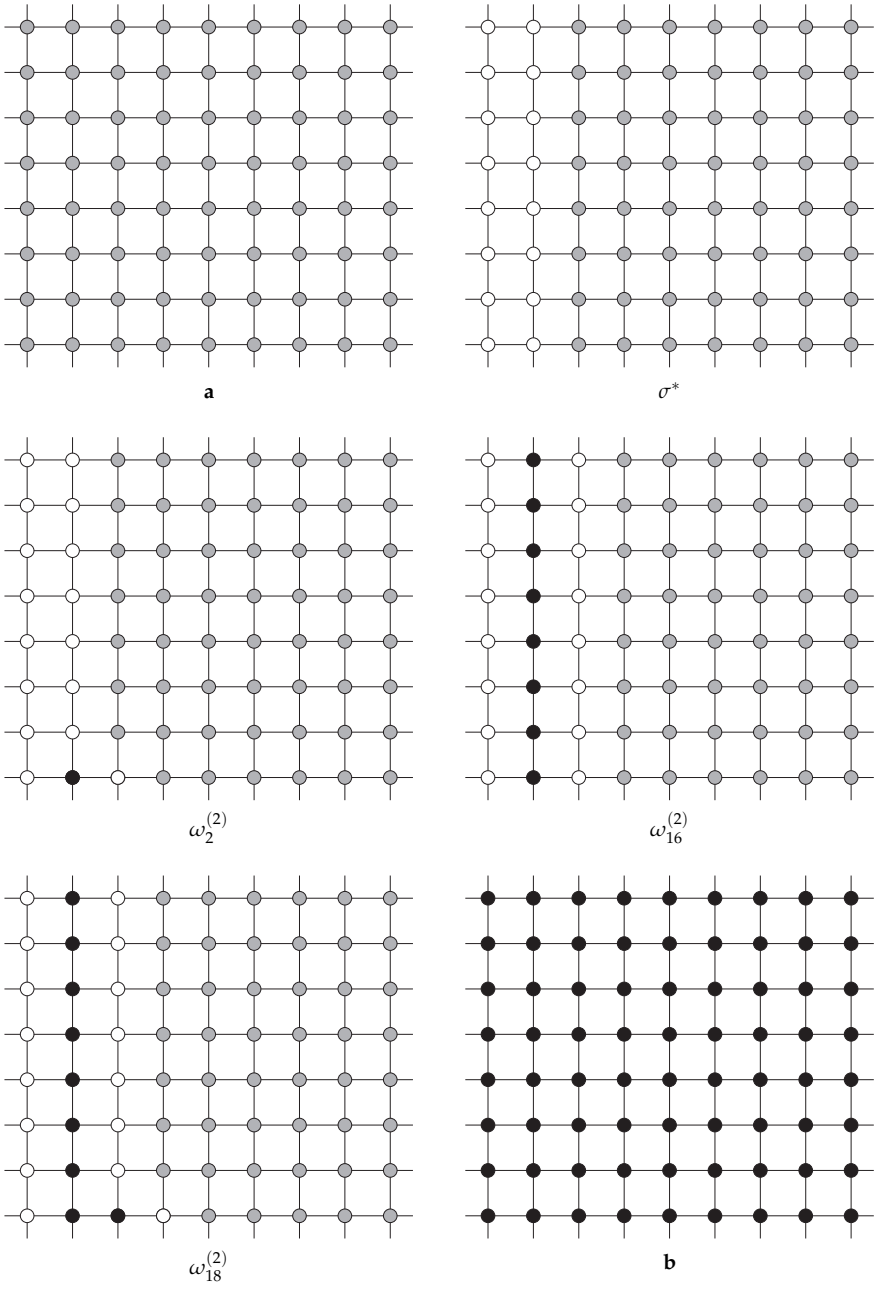


Figure 7.20: Some snapshots of the reference path ω^* for $\Lambda_{8,9}^T$

We will now use again the reduction algorithm for toric grids to prove Theorem 7.3.1(ii).

Proof of Theorem 7.3.1(ii). We want to show that for every admissible configuration $\sigma \neq \mathbf{a}, \mathbf{b}$, the following inequality holds

$$\Phi(\sigma, \{\mathbf{a}, \mathbf{b}\}) - H(\sigma) < 2K.$$

Let g be the number of gray particles that the configuration σ has in the first vertical stripe, that is

$$g := \sum_{v \in C_1} \mathbb{1}_{\{\sigma(v) = \mathbf{a}(v)\}}. \tag{7.30}$$

We distinguish two scenarios, depending on whether (a) $g \geq 2K - 1$ or (b) $g < 2K - 1$. In scenario (a) we will show that $\Phi(\sigma, \mathbf{a}) - H(\sigma) \leq 1$, while in scenario (b) we will prove that $\Phi(\sigma, \mathbf{b}) - H(\sigma) < 2K$.

Consider scenario (a) first, where σ has at least $2K - 1$ gray particles in the first vertical stripe. This means that σ has no black particles in the first vertical stripe and is then a suitable initial configuration for the reduction algorithm with target state $\mathbf{a}$, since condition (7.24) is satisfied. Using this procedure, we obtain a path ω such that $\Phi(\sigma, \mathbf{a}) - H(\sigma) \leq 1$, since at every step where a black particle is (possibly) removed, at the next step a gray particle is added.

As far as scenario (b) is concerned, let $\sigma^* \in \mathcal{X}$ be the configuration obtained from σ by removing all the g gray particles in the first vertical stripe, namely

$$\sigma^*(v) := \begin{cases} \sigma(v) & \text{if } v \in V \setminus C_1, \\ \sigma(v) & \text{if } v \in C_1 \text{ and } \sigma(v) = \mathbf{b}(v), \\ 0 & \text{if } v \in C_1 \text{ and } \sigma(v) = \mathbf{a}(v). \end{cases}$$

Clearly $H(\sigma^*) - H(\sigma) = g$. We construct a path $\omega^{(1)} = (\omega_1^{(1)}, \dots, \omega_{g+1}^{(1)})$ from $\omega_1^{(1)} = \sigma$ to $\omega_{g+1}^{(1)} = \sigma^*$ as follows. For $i = 1, \dots, g$, at step i we remove the first gray particle in lexicographic order from configuration $\omega_i^{(1)}$, obtaining in this way $\omega_{i+1}^{(1)}$, which is such that $H(\omega_{i+1}^{(1)}) = H(\omega_i^{(1)}) + 1$. Thus,

$$\Phi_{\omega^{(1)}} = H(\sigma^*) = H(\sigma) + g. \tag{7.31}$$

Note that the configuration σ^* satisfies condition (7.23), since it has no gray particles in the first vertical stripe C_1. Thus, σ^* is a suitable initial configuration for the reduction algorithm described before, which we can use to construct a path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ such that

$$\Phi_{\omega^{(2)}} \leq H(\sigma^*) + 1. \tag{7.32}$$

The concatenation of paths $\omega^{(1)}$ and $\omega^{(2)}$ yields a new path $\omega : \sigma \to \mathbf{b}$ that, in view of (7.31) and (7.32), satisfies inequality $\Phi_\omega \leq H(\sigma) + g + 1$. Therefore, using (7.30), we get

$$\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq g + 1 < 2K. \qquad \square$$

7.5 PROOFS FOR OPEN GRIDS

In this section we prove the structural properties for the energy landscape of the Widom-Rowlison model on the open grid graph $\Lambda^O_{K,L}$ given in Theorem 7.3.1(iii) and (iv).

We first prove a lower bound for the communication height between the two stable states $\mathbf{a}$ and $\mathbf{b}$, see Proposition 7.5.1 below. We then introduce a *reduction algorithm*, similar to the one described in the previous section. Using this algorithm, we exhibit in Proposition 7.5.2 a reference path from $\mathbf{a}$ to $\mathbf{b}$ that attains that lower bound, completing in this way the proof of Theorem 7.3.1(iii). Lastly, we use again the reduction algorithm to construct a specific path from every configuration $\sigma \neq \mathbf{a}, \mathbf{b}$ to the subset $\{\mathbf{a}, \mathbf{b}\}$ and prove Theorem 7.3.1(iv).

Proposition 7.5.1 (Lower bound for $\Phi(\mathbf{a}, \mathbf{b})$)**.** *Consider the Widom-Rowlison model on the $K \times L$ open grid $\Lambda^O_{K,L}$. The communication height between $\mathbf{a}$ and $\mathbf{b}$ in the corresponding energy landscape satisfies*

$$\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) \geq \min\{K, L\} + 1.$$

Proof. Without loss of generality, we may assume that $K \leq L$. We need to show that in every path $\omega : \mathbf{a} \to \mathbf{b}$ there is at least one configuration with energy wastage greater than or equal to $K + 1$. Take a path $\omega = (\omega_1, \ldots, \omega_n) \in \Omega_{\mathbf{a},\mathbf{b}}$. Without loss of generality, we may assume that there are no void moves in ω, i.e. at every step either a particle is added or a particle is removed, so that $H(\omega_{i+1}) = H(\omega_i) \pm 1$ for every $1 \leq i \leq n - 1$. Since $\mathbf{a}$ has no black bridges, while $\mathbf{b}$ does, at some point along the path ω there must be a configuration ω_{m^*} that is *the first* to display a black vertical bridge or a black horizontal bridge or both simultaneously (i.e. a black cross). We claim that

$$\max\{U(\omega_{m^*-1}), U(\omega_{m^*-2})\} \geq \min\{K, L\} + 1.$$

We distinguish the following three cases:

(a) ω_{m^*} displays a black vertical bridge only;

(b) ω_{m^*} displays a black horizontal bridge only;

(c) ω_{m^*} displays a black cross.

These three cases cover all the possibilities, since the addition of a single particle cannot create two parallel bridges simultaneously.

For case (a), we claim that the energy wastage in every row is greater than or equal to one. Suppose by contradiction that there exists a row r^* such that $U^{r^*}(\omega_{m^*}) = 0$. Then, by Lemma 7.3.4, ω_{m^*} would have a bridge in row r^*. Such a bridge cannot be black, since otherwise ω_{m^*} would have a black cross, and neither gray, which by Lemma 7.3.2 could not coexist with the black vertical

bridge that ω_{m^*} has, contradiction. Hence $U^r(\omega_{m^*}) \geq 1$ for every row r and thus

$$U(\omega_{m^*}) \geq \sum_{i=0}^{K-1} U^{r_i}(\omega_{m^*}) \geq K.$$

The previous configuration ω_{m^*-1} along the path ω differs from ω_{m^*} in exactly one site v^*, which is such that $\omega_{m^*-1}(v^*) = 0$ and $\omega_{m^*}(v^*) = \mathbf{b}(v^*)$. Hence $U(\omega_{m^*-1}) = U(\omega_{m^*}) + 1$ and therefore

$$U(\omega_{m^*-1}) \geq K + 1.$$

For case (b) we can argue as in case (a), but interchanging the role of rows and columns, and obtain that

$$U(\omega_{m^*-1}) \geq L + 1.$$

For case (c), let r^* and c^* be respectively the row and the column where the black cross lies in configuration ω_{m^*}. The previous configuration ω_{m^*-1} along the path ω differs from ω_{m^*} in exactly one site, denoted by v^*, which has to be the site where row r^* and column c^* intersect (otherwise configuration ω_{m^*-1} would have a black bridge, violating the definition of m^*). Furthermore, $U^{c^*}(\omega_{m^*-1}) \geq 1$, since $\omega_{m^*-1}(v^*) = 0$ by definition of m^*. We claim that the energy wastage of ω_{m^*-1} in all columns different from c^* is also greater than or equal to one, namely

$$U^c(\omega_{m^*-1}) \geq 1 \quad \forall\, c \neq c^*. \tag{7.33}$$

These inequalities follow from Lemma 7.3.4 after noticing that in each of these $L-1$ columns configuration ω_{m^*-1} cannot display

- a black vertical bridge, by definition of m^*;
- a gray vertical bridge, since every column $c \neq c^*$ has at least one black particle (at the intersection with row r^*).

Summing the energy wastage of all columns we get

$$U(\omega_{m^*-1}) \geq \sum_{j=0}^{L-1} U^{c_j}(\omega_{m^*-1}) \geq L. \tag{7.34}$$

If $U(\omega_{m^*-1}) \geq L + 1$, then the proof is complete. If instead $U(\omega_{m^*-1}) = L$, consider the configuration ω_{m^*-2} preceding ω_{m^*-1} in the path ω. By construction, the configuration ω_{m^*-2} differs from ω_{m^*-1} by a single-site update and thus

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) \pm 1. \tag{7.35}$$

Consider the case where $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = L - 1$. In this case, thanks to the pigeonhole principle, the configuration ω_{m^*-2} must have either

at least one column c with $U^c(\omega_{m^*-2}) = 0$. By Lemma 7.3.4, ω_{m^*-2} must have a bridge in this column, see Figure 7.21. However it cannot be a black bridge, due to the definition of m^*, and neither a gray bridge, since it would not be possible to obtain a black bridge in row r_* (which ω_{m^*} has) with only two single-site updates, since at least five are needed, as shown in Figure 7.21.

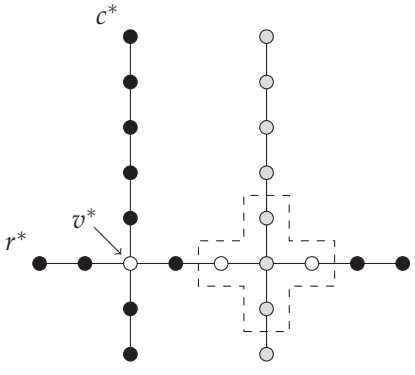


Figure 7.21: The dashed line encloses the five sites that should be updated if configuration ω_{m^*-2} had a gray bridge in column c

Therefore, it is not possible that $U(\omega_{m^*-2}) = U(\omega_{m^*-1}) - 1 = L - 1$, and (7.35) yields that

$$U(\omega_{m^*-2}) = U(\omega_{m^*-1}) + 1 = L + 1. \qquad \square$$

Reduction algorithm for open grids

We now present an iterative procedure similar to that for toric grids that constructs a path ω from an initial suitable configuration σ to the stable state $\mathbf{b}$. Thanks to the free boundary conditions of $\Lambda^O_{K,L}$, the condition on the initial configuration σ is weaker than (7.23). Indeed, we require that σ should not have gray particles in the first column c_0 only, i.e.

$$\sum_{v \in c_0} \mathbb{1}_{\{\sigma(v) = \mathbf{a}(v)\}} = 0. \tag{7.36}$$

The path $\omega : \sigma \to \mathbf{b}$ is built as the concatenation of L paths $\omega^{(1)}, \dots, \omega^{(L)}$. Path $\omega^{(j)}$ goes from σ_j to σ_{j+1}, where we recursively define $\sigma_1 = \sigma$ and

$$\sigma_{j+1}(v) = \begin{cases} \sigma_j(v) & \text{if } v \in V \setminus (c_j \cup c_{j+1}), \\ \sigma_j(v) & \text{if } v \in c_{j+1} \text{ and } \sigma(v) = \mathbf{b}(v), \\ \mathbf{a}(v) & \text{if } v \in c_j, \\ 0 & \text{if } v \in c_{j+1} \text{ and } \sigma(v) = \mathbf{a}(v), \end{cases}$$

for $j = 1, \dots, L$. It can be checked that indeed $\sigma_{L+1} = \mathbf{b}$. We now describe in detail how to construct each of the paths $\omega^{(j)}$ for $j = 1, \dots, L$. We build a path $\omega^{(j)} = (\omega_1^{(j)}, \dots, \omega_{2K+1}^{(j)})$ of length $2K+1$ (but possibly with void moves), with $\omega_1^{(j)} = \sigma_j$ and $\omega_{2K+1}^{(j)} = \sigma_{j+1}$. We repeat iteratively the following procedure for every $i = 1, \dots, 2K$:

- If $i \equiv 1 \pmod 2$, consider site $v = (j+1, (i-1)/2)$.
  - If $\omega_i^{(j)}(v) = 0$ or $\omega_i^{(j)}(v) = \mathbf{b}(v)$, we set $\omega_{i+1}^{(j)}(v) = \omega_i^{(j)}(v)$.
  - If $\omega_i^{(j)}(v) = \mathbf{a}(v)$, then we remove the gray particle in site v from configuration $\omega_i^{(j)}$, obtaining in this way configuration $\omega_{i+1}^{(j)}$ with $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) + 1$.
- If $i \equiv 0 \pmod 2$, consider site $v = (j, i/2 - 1)$.
  - If $\omega_i^{(j)}(v) = \mathbf{b}(v)$, we set $\omega_{i+1}^{(j)}(v) = \omega_i^{(j)}(v) = \mathbf{b}(v)$.
  - If $\omega_i^{(j)}(v) = 0$, then we add a black particle in site v, obtaining in this way a new configuration $\omega_{i+1}^{(j)}$ such that $\omega_{i+1}^{(j)}(v) = \mathbf{a}(v)$. This configuration is admissible because by construction there are no gray particles in any neighboring sites of v. In particular, the site at its right (that is $v + (1,0)$) has possibly been emptied at the previous step, if there was a gray particle in it. Moreover, $H(\omega_{i+1}^{(j)}) = H(\omega_i^{(j)}) - 1$

Note that for the last path $\omega^{(L)}$ all steps corresponding to odd values of i are void, since there is no column c_L.

In words, the reduction algorithm alternately removes a gray particle (if any) and adds a black particle, progressively column by column. For every $j = 1, \dots, L$, the configurations σ_j and σ_{j+1} and the path $\omega^{(j)}$ are constructed in such a way that

$$H(\sigma_{j+1}) \leq H(\sigma_j),$$

since the number of black particles added in column c_j is greater than or equal to the number of gray particles removed in column c_{j+1}. Moreover,

$$\Phi_{\omega^{(j)}} \leq H(\sigma_j) + 1,$$

since along the path $\omega^{(j)}$ every particle removal (if any) is always followed by a particle addition. These two properties imply that the path $\omega : \sigma \to \mathbf{b}$ created by concatenating $\omega^{(1)}, \dots, \omega^{(L)}$ satisfies

$$\Phi_\omega \leq H(\sigma) + 1,$$

which shows that $\Phi(\sigma, \mathbf{b}) - H(\sigma) \leq 1$.

We remark that, with a few minor tweaks, we can define a similar reduction algorithm that builds a path ω from a configuration σ to the stable state $\mathbf{a}$ such that $\Phi_\omega \leq H(\sigma) + 1$. In this scenario, however, the initial configuration σ should satisfy an initial condition analogous to (7.36), namely it should have no black particles in the first column.

The reduction algorithm will now be used in the next proposition to build a reference path from $\mathbf{a}$ to $\mathbf{b}$, which shows that $\Phi(\mathbf{a}, \mathbf{b}) - H(\mathbf{a}) \leq \min\{K, L\} + 1$.

Proposition 7.5.2 (Reference path)**.** *There exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ in $\mathcal{X}$ such that*

$$\Phi_{\omega^*} - H(\mathbf{a}) = \min\{K, L\} + 1.$$

Proof. Without loss of generality, we may assume $K \leq L$ and show that there exists a path $\omega^* : \mathbf{a} \to \mathbf{b}$ such that $\Phi_{\omega^*} - H(\mathbf{a}) = K + 1$. We describe just briefly how the reference path ω^* is constructed, since it is very similar to the one given in case (b) of the proof of Proposition 7.4.2, see Figure 7.22. Also in this case, the path ω^* is the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \mathbf{a} \to \sigma^*$ and $\omega^{(2)} : \sigma^* \to \mathbf{b}$, where σ^* is the configuration that differs from $\mathbf{a}$ only by having all sites of the leftmost column empty, i.e.

$$\sigma^*(v) := \begin{cases} \mathbf{a}(v) & \text{if } v \in V \setminus c_0, \\ 0 & \text{if } v \in c_0. \end{cases}$$

The path $\omega^{(1)}$ consists of K steps, at each of which we remove the first particle in c_0 in lexicographic order from the previous configuration. The last configuration is precisely σ^*, which has energy $H(\sigma^*) = H(\mathbf{a}) + K$, and, trivially, $\Phi_{\omega^{(1)}} = H(\mathbf{a}) + K$. The second path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ is then constructed by means of the reduction algorithm, which can be used since configuration σ^* is a suitable initial configuration for it, satisfying condition (7.36). The algorithm guarantees that $\Phi_{\omega^{(2)}} = H(\sigma^*) + 1$ and thus the concatenation of the two paths $\omega^{(1)}$ and $\omega^{(2)}$ yields a path ω^* with $\Phi_{\omega^*} = H(\mathbf{a}) + K + 1$ as desired. □

Proof of Theorem 7.3.1(iii). It is an immediate consequence of the lower bound for $\Phi(\mathbf{a}, \mathbf{b})$ in Proposition 7.5.1 and the matching upper bound that follows from Proposition 7.5.2. □

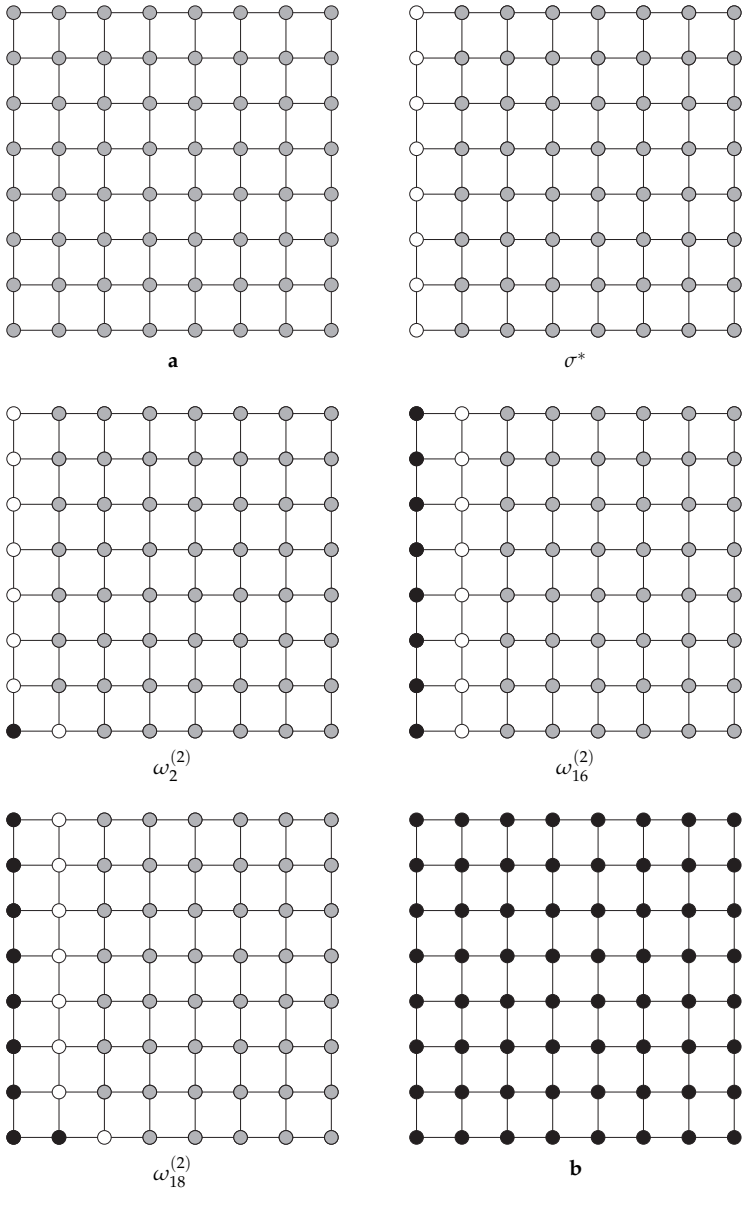


Figure 7.22: Some snapshots of the reference path from **a** to **b** for $\Lambda_{8,8}^O$

We now present the proof of Theorem 7.3.1(iv), which uses again the reduction algorithm for open grids described earlier in this section.

Proof of Theorem 7.3.1(iv). We want to show that for every configuration $\sigma \neq \mathbf{a}, \mathbf{b}$ we have $\Phi(\sigma, \{\mathbf{a}, \mathbf{b}\}) - H(\sigma) \leq \min\{K, L\}$. Since we may assume that $K \leq L$ without loss of generality, it is enough to prove that

$$\Phi(\sigma, \{\mathbf{a}, \mathbf{b}\}) - H(\sigma) \leq K, \quad \forall\, \sigma \neq \mathbf{a}, \mathbf{b}.$$

Let g be the number of gray particles present in configuration σ in the first column of Λ, i.e.

$$g := \sum_{v \in c_0} \mathbb{1}_{\{\sigma(v) = \mathbf{a}(v)\}}.$$

Every column in Λ has K sites, so $0 \leq g \leq K$. We distinguish two cases: (a) $g = K$ and (b) $g < K$.

In case (a), $g = K$ and thus σ has a gray bridge in column c_0. In particular, σ has no black particles in the first column and we can use the modified version of the reduction algorithm for open grids with $\mathbf{a}$ as target configuration. We build in this way a path $\omega : \sigma \to \mathbf{a}$ such that $\Phi_\omega \leq H(\sigma) + 1$, showing that in this case $\Phi(\sigma, \mathbf{a}) - H(\sigma) \leq 1$.

Consider now case (b), where $g < K$. In this case we create a path $\omega : \sigma \to \mathbf{b}$ as the concatenation of two shorter paths, $\omega^{(1)}$ and $\omega^{(2)}$, where $\omega^{(1)} : \sigma \to \sigma^*$, $\omega^{(2)} : \sigma^* \to \mathbf{b}$ and σ^* is a suitable configuration which depends on σ (see definition below). The reason why ω is best described as concatenation of two shorter paths is the following: Since $g > 0$, the reduction algorithm for open grids can not be started directly from σ and the path $\omega^{(1)}$ indeed leads from σ to σ^*, which is a suitable configuration to initialize the procedure. The configuration σ^* is obtained from σ by removing the g gray particles that reside in the leftmost column, i.e.

$$\sigma^*(v) := \begin{cases} \sigma(v) & \text{if } v \in V \setminus c_0, \\ \sigma(v) & \text{if } v \in c_0 \text{ and } \sigma(v) = \mathbf{b}(v), \\ 0 & \text{if } v \in c_0 \text{ and } \sigma(v) = \mathbf{a}(v). \end{cases}$$

The path $\omega^{(1)} = (\omega_1^{(1)}, \ldots, \omega_{g+1}^{(1)})$, with $\omega_1^{(1)} = \sigma$ and $\omega_{g+1}^{(1)} = \sigma^*$, can be constructed as follows. For $i = 1, \ldots, g$, at step i we remove from configuration $\omega_i^{(1)}$ the topmost gray particle in c_0 increasing the energy by 1 and obtaining in this way configuration $\omega_{i+1}^{(1)}$. Therefore, the configuration σ^* is such that $H(\sigma^*) - H(\sigma) = g$ and

$$\Phi_{\omega^{(1)}} \leq H(\sigma^*) = H(\sigma) + g.$$

The path $\omega^{(2)} : \sigma^* \to \mathbf{b}$ is then constructed by means of the reduction algorithm for open grids described earlier, using σ^* as initial configuration and $\mathbf{b}$ as target configuration. This procedure guarantees that

$$\Phi_{\omega^{(2)}} \leq H(\sigma^*) + 1.$$

The concatenation of the two paths $\omega^{(1)}$ and $\omega^{(2)}$ gives a path $\omega : \sigma \to \mathbf{b}$ that satisfies the inequality $\Phi_\omega \leq H(\sigma) + g + 1$ and, since $g < K$, we obtain

$$\Phi(\sigma, \mathbf{b}) - H(\sigma) = g + 1 \leq K. \qquad \square$$

BIBLIOGRAPHY

[1] N. Abramson. The ALOHA System – Another alternative for computer communication. In *AFIPS Fall Joint Computer Conference*, volume 37, pages 281–285, 1970.

[2] D.J. Aldous. Markov chains with almost exponential hitting times. *Stochastic Processes and their Applications*, 13(3):305–310, 1982.

[3] D.J. Aldous and M. Brown. Inequalities for rare events in time-reversible Markov chains I. In *Stochastic inequalities*, volume 22 of *Lecture Notes-Monograph Series*, pages 1–16. IMS, 1992.

[4] D.J. Aldous and M. Brown. Inequalities for rare events in time-reversible Markov chains II. *Stochastic Processes and their Applications*, 44(1):15–25, 1993.

[5] G.E. Andrews. The hard-hexagon model and Rogers-Ramanujan type identities. *Proceedings of the National Academy of Sciences of the United States of America*, 78(9):5290–5292, 1981.

[6] M. Andrews and L. Zhang. Utility optimization in heterogeneous networks via CSMA-based algorithms. In *2013 WiOpt Proceedings*, pages 372–379, 2013.

[7] N. Antunes, C. Fricker, P. Robert, and D. Tibi. Metastability of CDMA cellular systems. In *2006 MobiCom Proceedings*, pages 206–214, 2006.

[8] N. Antunes, C. Fricker, P. Robert, and D. Tibi. Stochastic networks with multiple stable points. *The Annals of Probability*, 36(1):255–278, 2008.

[9] R.J. Baxter. Hard hexagons: Exact solution. *Journal of Physics A: Mathematical and General*, 13(3):L61–L70, 1999.

[10] B. Bellalta, A. Checco, A. Zocca, and J. Barcelo. On the interactions between multiple overlapping WLANs using channel bonding. In press. Accepted in *IEEE Transactions on Vehicular Technology*, 2015.

[11] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *Journal of Statistical Physics*, 140(6):1065–1114, 2010.

[12] J. Beltrán and C. Landim. Metastability of reversible finite state Markov processes. *Stochastic Processes and their Applications*, 121(8):1633–1677, 2011.

[13] G. Ben Arous and R. Cerf. Metastability of the three dimensional Ising model on a torus at very low temperatures. *Electronic Journal of Probability*, 1:1–55, 1996.

[14] O. Benois, C. Landim, and M. Mourragui. Hitting times of rare events in Markov chains. *Journal of Statistical Physics*, 153(6):967–990, 2013.

[15] C. Berge. Minimax relations for the partial q-colorings of a graph. *Discrete Mathematics*, 74(1-2):3–14, 1989.

[16] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3): 535–547, 2000.

[17] R.R. Boorstyn and A. Kershenbaum. Throughput analysis of multihop packet radio networks. In *Proceedings of ICC '80*, pages 1361–1366, 1980.

[18] R.R. Boorstyn, A. Kershenbaum, B. Maglaris, and V. Sahin. Throughput analysis in multihop CSMA packet radio networks. *IEEE Transactions on Communications*, 35(3):267–274, 1987.

[19] C. Borgs, J.T. Chayes, A. Frieze, P. Tetali, E. Vigoda, H.V. Van. Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*, pages 218–229, 1999.

[20] N. Bouman, S.C. Borst, and J.S.H. van Leeuwaarden. Delay performance in random-access networks. *Queueing Systems*, 77(2):211–242, 2014.

[21] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low lying spectra in reversible Markov chains. *Communications in Mathematical Physics*, 228(2):219–255, 2002.

[22] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. *Journal of the European Mathematical Society*, 6(4):399–424, 2004.

[23] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes II: Precise asymptotics for small eigenvalues. *Journal of the European Mathematical Society*, 7(1):69–99, 2005.

[24] G.R. Brightwell, O. Häggström, and P. Winkler. Nonmonotonic behavior in hard-core and Widom-Rowlinson models. *Journal of Statistical Physics*, 94(3-4):415–435, 1999.

[25] K. Cameron. A min-max relation for the partial q-colourings of a graph. Part II: Box perfection. *Discrete Mathematics*, 74(1-2):15–27, 1989.

[26] M. Cassandro, A. Galves, E. Olivieri, and M.E. Vares. Metastable behavior of stochastic dynamics: A pathwise approach. *Journal of Statistical Physics*, 35(5-6):603–634, 1984.

[27] O. Catoni. Sharpe large deviations estimates for simulated annealing algorithms. *Ann. Inst. Henri Poincaré*, 27(3):291–383, 1991.

[28] O. Catoni. Rough large deviation estimates for simulated annealing: Application to exponential schedules. *The Annals of Probability*, 20(3): 1109–1146, 1992.

[29] O. Catoni. Simulated annealing algorithms and Markov chains with rare transitions. In *Séminaire de probabilités XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 69–119. Springer Berlin, 1999.

[30] O. Catoni and R. Cerf. The exit path of a Markov chain with rare transitions. *ESAIM: Probability and Statistics*, 1:95–144, 1997.

[31] O. Catoni and A. Trouvé. Parallel annealing by multiple trials: A mathematical study. In *Simulated Annealing: Parallelization techniques*, pages 129–144. John Wiley & Sons, Inc., 1992.

[32] F. Cecchi, S.C. Borst, and J.S.H. van Leeuwaarden. Throughput of CSMA networks with buffer dynamics. *Performance Evaluation*, 79:216–234, 2014.

[33] J.T. Chayes, L. Chayes, and R. Kotecký. The analysis of the Widom-Rowlinson model by stochastic geometric methods. *Communications in Mathematical Physics*, 172:551–569, 1995.

[34] E.N.M. Cirillo and J.L. Lebowitz. Metastability in the two-dimensional Ising model with free boundary conditions. *Journal of Statistical Physics*, 90(1/2):211–226, 1998.

[35] E.N.M. Cirillo and F.R. Nardi. Metastability for a stochastic dynamics with a parallel heat bath updating rule. *Journal of Statistical Physics*, 110 (1-2):183–217, 2003.

[36] E.N.M. Cirillo and F.R. Nardi. Relaxation height in energy landscapes: An application to multiple metastable states. *Journal of Statistical Physics*, 150(6):1080–1114, 2013.

[37] E.N.M. Cirillo, F.R. Nardi, and J. Sohier. A comparison between different cycle decompositions for Metropolis dynamics. *arXiv:1401.3522*, 2014.

[38] E.N.M. Cirillo, F.R. Nardi, and J. Sohier. Metastability for general dynamics with rare transitions: Escape time and critical configurations. *Journal of Statistical Physics*, 2015.

[39] E.N.M. Cirillo, F.R. Nardi, and C. Spitoni. Metastability for reversible probabilistic cellular automata with self-interaction. *Journal of Statistical Physics*, 132(3):431–471, 2008.

[40] E.N.M. Cirillo and E. Olivieri. Metastability and nucleation for the Blume-Capel model. Different mechanisms of transition. *Journal of Statistical Physics*, 83(3-4):473–554, 1996.

[41] F. den Hollander, F.R. Nardi, E. Olivieri, and E. Scoppola. Droplet growth for three-dimensional Kawasaki dynamics. *Probability Theory and Related Fields*, 125(2):153–194, 2003.

[42] F. den Hollander, E. Olivieri, and E. Scoppola. Metastability and nucleation for conservative dynamics. *Journal of Mathematical Physics*, 41(3): 1424, 2000.

[43] O. Dousse, P. Thiran, and M. Durvy. On the fairness of large CSMA networks. *IEEE Journal on Selected Areas in Communications*, 27(7):1093–1104, 2009.

[44] M. Durvy, O. Dousse, and P. Thiran. Modeling the 802.11 protocol under different capture and sensing capabilities. In *INFOCOM, 2007 Proceedings*, pages 2356–2360. IEEE, 2007.

[45] M. Durvy, O. Dousse, and P. Thiran. Border effects, fairness, and phase transition in large wireless networks. In *INFOCOM, 2008 Proceedings*, pages 601–609. IEEE, 2008.

[46] M. Durvy and P. Thiran. A packing approach to compare slotted and non-slotted medium access control. In *INFOCOM, 2006 Proceedings*, pages 1–12. IEEE, 2006.

[47] M. Dyer and C. Greenhill. On Markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

[48] A. Faridi, B. Bellalta, and A. Checco. Analysis of Dynamic Channel Bonding in Dense Networks of WLANs. *arXiv:1509.00290*, 2015.

[49] R. Fernandez, F. Manzo, F.R. Nardi, and E. Scoppola. Asymptotically exponential hitting times and metastability: A pathwise approach without reversibility. *arXiv:1406.2637*, 2014.

[50] M. Franceschetti, O. Dousse, D.N.C. Tse, and P. Thiran. Closing the gap in the capacity of wireless networks via percolation theory. *IEEE Transactions on Information Theory*, 53(3):1009–1018, 2007.

[51] D. Galvin. Sampling independent sets in the discrete torus. *Random Structures and Algorithms*, 33(3):356–376, 2008.

[52] D. Galvin, F. Martinelli, K. Ramanan, and P. Tetali. The multi-state hard core model on a regular tree. *SIAM Journal on Discrete Mathematics*, 25 (2):894–915, 2010.

[53] D. Galvin and P. Tetali. Slow mixing of Glauber dynamics for the hard-core model on the hypercube. *Proceedings of SODA '04*, pages 466–467, 2004.

[54] D. Galvin and P. Tetali. Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs. *Random Structures and Algorithms*, 28(4):427–443, 2006.

[55] M. Garetto, T. Salonidis, and E.W. Knightly. Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks. *IEEE/ACM Transactions on Networking*, 16(4):864–877, aug 2008.

[56] A. Gaudillière, E. Olivieri, and E. Scoppola. Nucleation pattern at low temperature for local Kawasaki dynamics in two dimensions. *Markov Processes and Related Fields*, 11(4):553–628, 2005.

[57] D.S. Gaunt and M.E. Fisher. Hard-sphere lattice gases. I. Plane-square lattice. *The Journal of Chemical Physics*, 43(8):2840–2863, 1965.

[58] J. Ghaderi and R. Srikant. On the design of efficient CSMA algorithms for wireless networks. In *49th IEEE Conference on Decision and Control (CDC)*, pages 954–959, 2010.

[59] J. Ghaderi and R. Srikant. The impact of access probabilities on the delay performance of Q-CSMA algorithms in wireless networks. *IEEE/ACM Transactions on Networking*, 21(4):1063–1075, 2013.

[60] R.J. Gibbens, P.J. Hunt, and F.P. Kelly. Bistability in communication networks. In *Disorder in Physical Systems*, pages 113–127. Oxford University Press, 1990.

[61] S. Greenberg and D. Randall. Slow mixing of Markov chains using fault lines and fat contours. *Algorithmica*, 58(4):911–927, 2010.

[62] R. Grübel and M. Reich. Rarity and exponentiality: An extension of Keilson's theorem, with applications. *Journal of Applied Probability*, 42(2): 393–406, 2005.

[63] M.M. Halldorsson and T. Tonoyan. How Well Can Graphs Represent Wireless Interference? In *Proceedings of the 47th Annual ACM on Symposium on Theory of Computing - STOC '15*, pages 635–644, 2015.

[64] O.J. Heilmann and E. Praestgaard. Phase transition of hard hexagons on a triangular lattice. *Journal of Statistical Physics*, 9(1):23–44, 1973.

[65] R.A. Holley and D.W. Stroock. Simulated annealing via Sobolev inequalities. *Communications in Mathematical Physics*, 115(4):553–569, 1988.

[66] P.-K. Huang and X. Lin. Improving the delay performance of CSMA algorithms: A Virtual Multi-Channel approach. In *INFOCOM, 2013 Proceedings*, pages 2598–2606. IEEE, 2013.

[67] K. Jain, J. Padhye, V.N. Padmanabhan, and L. Qiu. Impact of Interference on Multi-Hop Wireless Network Performance. *Wireless Networks*, 11(4):471–487, 2005.

[68] L. Jiang, M. Leconte, J. Ni, R. Srikant, and J. Walrand. Fast mixing of parallel Glauber dynamics and low-delay CSMA scheduling. In *INFOCOM, 2011 Proceedings*, pages 371–375. IEEE, 2011.

[69] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: Scheduling and congestion control. *IEEE Transactions on Information Theory*, 56(12):6182–6207, 2010.

[70] L Jiang and J Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In *2008 Allerton Conference Proceedings*, pages 1511–1519, 2008.

[71] L. Jiang and J. Walrand. A distributed CSMA algorithm for throughput and utility maximization in wireless networks. *IEEE/ACM Transactions on Networking*, 18(3):960–972, 2010.

[72] L. Jiang and J. Walrand. Stability and delay of distributed scheduling algorithms for networks of conflicting queues. *Queueing Systems*, 72(1-2): 161–187, 2012.

[73] O. Jouini and Y. Dallery. Moments of first passage times in general birth-death processes. *Mathematical Methods of Operations Research*, 68(1): 49–76, 2008.

[74] C.H. Kai and S.C. Liew. Temporal Starvation in CSMA Wireless Networks. In *2009 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2011.

[75] S. Karlin and J.L. McGregor. Coincidence properties of birth and death processes. *Pacific J. Math*, 9(4):1109–1140, 1959.

[76] J. Keilson. A review of transient behavior in regular diffusion and birth-death processes. Part II. *Journal of Applied Probability*, 2(2):405–428, 1965.

[77] J. Keilson. A limit theorem for passage times in ergodic regenerative processes. *The Annals of Mathematical Statistics*, 37(4):866–870, 1966.

[78] J. Keilson. Log-concavity and log-convexity in passage time densities of diffusion and birth-death processes. *Journal of Applied Probability*, 8(2): 391–398, 1971.

[79] J. Keilson. *Markov Chain Models - Rarity and Exponentiality*, volume 28 of *Applied Mathematical Sciences*. Springer, New York, NY, 1979.

[80] F.P. Kelly. *Reversibility and Stochastic Networks*. Wiley, New York, NY, 1979.

[81] F.P. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B*, 47(3):379–395, 1985.

[82] F.P. Kelly. One-Dimensional Circuit-Switched Networks. *The Annals of Probability*, 15(3):1166–1179, 1987.

[83] F.P. Kelly. Loss networks. *The Annals of Applied Probability*, 1(3):319–378, 1991.

[84] A. Kershenbaum, R.R. Boorstyn, and M.-S. Chen. An algorithm for evaluation of throughput in multihop packet radio networks with complex topologies. *IEEE Journal on Selected Areas in Communications*, 5(6):1003–1012, 1987.

[85] L. Kleinrock and J. Silvester. Spatial reuse in multihop packet radio networks. *Proceedings of the IEEE*, 75(1):156–167, 1987.

[86] L. Kleinrock and F.A. Tobagi. Packet switching in radio channels: Part I - Carrier Sense Multiple-Access modes and their throughput-delay characteristics. *IEEE Transactions on Communications*, 23(12):1400–1416, 1975.

[87] R. Kotecký and E. Olivieri. Droplet dynamics for asymmetric Ising model. *Journal of Statistical Physics*, 70(5-6):1121–1148, 1993.

[88] R. Kotecký and E. Olivieri. Shapes of growing droplets — A model of escape from a metastable phase. *Journal of Statistical Physics*, 75(3-4): 409–506, 1994.

[89] J. Kwak, C.-H. Lee, and D.Y. Eun. Exploiting the past to reduce delay in CSMA scheduling: A high-order Markov chain approach. *ACM SIGMETRICS Performance Evaluation Review*, 41(1):353, 2013.

[90] J. Kwak, C.-H. Lee, and D.Y. Eun. A high-order Markov chain based scheduling algorithm for low delay in CSMA networks. In *INFOCOM, 2014 Proceedings*, pages 1662–1670. IEEE, 2014.

[91] K.-K. Lam, C.-K. Chau, M. Chen, and S.C. Liew. Mixing time and temporal starvation of general CSMA networks with multiple frequency agility. In *2012 IEEE International Symposium on Information Theory (ISIT) Proceedings*, pages 2676–2680, 2012.

[92] R. Laufer and L. Kleinrock. On the capacity of wireless CSMA/CA multihop networks. In *INFOCOM, 2013 Proceedings*, pages 1312–1320. IEEE, 2013.

[93] G.F. Lawler and A.D. Sokal. Bounds on the L^2 spectrum for Markov chains and Markov processes: A generalization of Cheeger's inequality. *Transactions of the American Mathematical Society*, 309(2):557–557, 1988.

[94] J.L. Lebowitz and G. Gallavotti. Phase transitions in binary lattice gases. *Journal of Mathematical Physics*, 12(7):1129–1133, 1971.

[95] J.L. Lebowitz and E.H. Lieb. Phase transition in a continuum classical system with finite interactions. *Physics Letters A*, 39(2):98–100, 1972.

[96] J.L. Lebowitz, A.E. Mazel, P. Nielaba, and L. Amaj. Ordering and demixing transitions in multicomponent Widom-Rowlinson models. *Physical Review E*, 52(6):5985–5996, 1995.

[97] W. Ledermann and G.E.H. Reuter. Spectral Theory for the Differential Equations of Simple Birth and Death Processes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 246 (914):321–369, 1954.

[98] C.-H. Lee, D.Y. Eun, S.-Y. Yun, and Y. Yi. From Glauber dynamics to Metropolis algorithm: Smaller delay in optimal CSMA. In *2012 IEEE International Symposium on Information Theory (ISIT) Proceedings*, pages 2681–2685. IEEE, 2012.

[99] D.A. Levin, Y. Peres, and E.L. Wilmer. *Markov Chains and Mixing Times*. AMS, Providence, 2009.

[100] S.C. Liew, C.H. Kai, J. Leung, and B. Wong. Back-of-the-envelope computation of throughput distributions in CSMA wireless networks. In *2009 IEEE International Conference on Communications (ICC)*, pages 1–19, 2009.

[101] S.C. Liew, J. Zhang, C.-K. Chau, and M. Chen. Analysis of frequency-agile CSMA wireless networks. *arXiv:1007.5255*, 2010.

[102] J. Liu, Y. Yi, A. Proutière, M. Chiang, and H.V. Poor. Towards utility-optimal random access without message passing. *Wireless Communications and Mobile Computing*, 10(1):115–128, 2010.

[103] M. Lotfinezhad and P. Marbach. Throughput-optimal random access with order-optimal delay. In *INFOCOM, 2011 Proceedings*, pages 2867–2875. IEEE, 2011.

[104] M. Luby and E. Vigoda. Fast convergence of the Glauber dynamics for sampling independent sets. *Random Structures and Algorithms*, 1198(92): 229–241, 1999.

[105] B. Luen, K. Ramanan, and I. Ziedins. Nonmonotonicity of phase transitions in a loss network with controls. *The Annals of Applied Probability*, 16(3):1528–1562, 2006.

[106] F. Manzo, F.R. Nardi, E. Olivieri, and E. Scoppola. On the essential features of metastability: Tunnelling time and critical configurations. *Journal of Statistical Physics*, 115(1/2):591–642, 2004.

[107] V. Marbukh. Loss circuit switched communication network-performance analysis and dynamic routing. *Queueing Systems*, 13(1-3): 111–141, 1993.

[108] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures and Algorithms*, 31(2):134–172, 2007.

[109] L. Miclo. About relaxation time of finite generalized Metropolis algorithms. *The Annals of Applied Probability*, 12(4):1492–1515, 2002.

[110] F.R. Nardi and E. Olivieri. Low temperature stochastic dynamics for an Ising model with alternating field. *Markov Processes and Related Fields*, 2: 117–166, 1996.

[111] F.R. Nardi, E. Olivieri, and E. Scoppola. Anisotropy effects in nucleation for conservative dynamics. *Journal of Statistical Physics*, 119(3-4):539–595, 2005.

[112] F.R. Nardi, A. Zocca, and S.C. Borst. Hitting time asymptotics for hard-core interactions on grids. *Journal of Statistical Physics, under revision*, 2015.

[113] E.J. Neves and R.H. Schonmann. Critical droplets and metastability for a Glauber dynamics at very low temperatures. *Communications in Mathematical Physics*, 137(2):209–230, 1991.

[114] E.J. Neves and R.H. Schonmann. Behavior of droplets for a class of Glauber dynamics at very low temperature. *Probability Theory and Related Fields*, 91(3-4):331–354, 1992.

[115] J. Ni, B. Tan, and R. Srikant. Q-CSMA: Queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks. In *INFOCOM, 2010 Proceedings*, pages 1–5. IEEE, 2010.

[116] E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: First exit problem from a general domain. I. The reversible case. *Journal of Statistical Physics*, 79(3-4):613–647, 1995.

[117] E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: First exit problem from a general domain. II. The general case. *Journal of Statistical Physics*, 84(5-6):987–1041, 1996.

[118] E. Olivieri and M.E. Vares. *Large Deviations and Metastability*. CUP, Cambridge, 2005.

[119] E. Pinsky and Y. Yemini. The asymptotic analysis of some packet radio networks. *IEEE Journal on Selected Areas in Communications*, 4(6):938–945, 1986.

[120] A. Proutière, Y. Yi, T. Lan, and M. Chiang. Resource allocation over network dynamics without timescale separation. In *INFOCOM, 2010 Proceedings*, pages 1–5. IEEE, 2010.

[121] S. Rajagopalan, D. Shah, and J. Shin. Network adiabatic theorem: An efficient randomized protocol for contention resolution. *ACM SIGMETRICS Performance Evaluation Review*, 37(1):133–144, 2009.

[122] K. Ramanan, A. Sengupta, I. Ziedins, and P. Mitra. Markov random field models of multicasting in tree networks. *Advances in Applied Probability*, 34(1):58–84, 2002.

[123] D. Randall. Slow mixing of Glauber dynamics via topological obstructions. *Proceedings of SODA '06*, pages 870–879, 2006.

[124] L.G. Roberts. ALOHA packet system with and without slots and capture. *ACM SIGCOMM Computer Communication Review*, 5(2):28–42, 1975.

[125] D. Ruelle. Existence of a phase transition in a continuous classical system. *Physical Review Letters*, 27(16):1040–1041, 1971.

[126] E. Scoppola. Renormalization and graph methods for Markov chains. In *Advances in dynamical systems and quantum physics*, pages 260–281, Capri, 1993.

[127] D. Shah and J. Shin. Delay optimal queue-based CSMA. In *ACM SIGMETRICS*, pages 373–374, 2010.

[128] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *The Annals of Applied Probability*, 22(1):128–171, 2012.

[129] D. Shah, D.N.C. Tse, and J.N. Tsitsiklis. Hardness of low delay network scheduling. *IEEE Transactions on Information Theory*, 57(12):7810–7817, 2011.

[130] K. Sigman. Regenerative processes (lecture notes), 2009. URL `http://www.columbia.edu/{~}ks20/stochastic-I/stochastic-I-RP.pdf`.

[131] V.T. Stefanov and S. Wang. A note on integrals for birth-death processes. *Mathematical Biosciences*, 168(2):161–165, 2000.

[132] V.G. Subramanian and M. Alanyali. Delay performance of CSMA in networks with bounded degree conflict graphs. In *2011 IEEE International Symposium on Information Theory (ISIT) Proceedings*, pages 2373–2377, 2011.

[133] Y.M. Suhov and U.A. Rozikov. A hard-core model on a Cayley tree: An example of a loss network. *Queueing Systems*, 46(1/2):197–212, 2004.

[134] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12): 1936–1948, 1992.

[135] A. Trouvé. Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms. *Ann. Inst. H. Poincaré, Probab. Statist.*, 32:299–348, 1994.

[136] P.M. van de Ven, S.C. Borst, J.S.H. van Leeuwaarden, and A. Proutière. Insensitivity and stability of random-access networks. *Performance Evaluation*, 67(11):1230–1242, 2010.

[137] J. van den Berg and J.E. Steif. Percolation and the hard-core lattice gas model. *Stochastic Processes and their Applications*, 49(2):179–197, 1994.

[138] X. Wang and K. Kar. Throughput modelling and fairness issues in CSMA/CA based ad-hoc networks. In *INFOCOM, 2005 Proceedings*, volume 1, pages 23–34. IEEE, 2005.

[139] B. Widom and J.S. Rowlinson. New model for the study of liquid-vapor phase transitions. *The Journal of Chemical Physics*, 52(1970):1670, 1970.

[140] Y. Yemini. A statistical mechanics of distributed resource sharing mechanisms. In *INFOCOM, 1983 Proceedings*, pages 531–539. IEEE, 1983.

[141] S.-Y. Yun, Y. Yi, J. Shin, and D.Y. Eun. Optimal CSMA: A survey. In *2012 IEEE International Conference on Communication Systems (ICCS)*, pages 199–204. IEEE, 2012.

[142] S. Zachary and I. Ziedins. Loss networks and Markov random fields. *Journal of Applied Probability*, 36(2):403–414, 1999.

[143] X. Zhou, Z. Zhang, G. Wang, X. Yu, B.Y. Zhao, H. Zheng. Practical conflict graphs for dynamic spectrum distribution. *ACM SIGMETRICS Performance Evaluation Review*, 41(1):5–16, 2013.

[144] I. Ziedins and F.P. Kelly. Limit theorems for loss networks with diverse routing. *Advances in Applied Probability*, 21(4):804–830, 1989.

[145] A. Zocca. Hitting times asymptotics for the hard-core model on finite triangular lattice. *In preparation*, 2015.

[146] A. Zocca. Low-temperature behavior of the Widom-Rowlison model on grid graphs. *In preparation*, 2015.

[147] A. Zocca, S.C. Borst, and J.S.H. van Leeuwaarden. Mixing properties of CSMA networks on partite graphs. In *Proceedings of VALUETOOLS 2012*, pages 117–126. IEEE, 2012.

[148] A. Zocca, S.C. Borst, and J.S.H. van Leeuwaarden. Slow Transitions and Starvation in Dense Random-Access Networks. *Stochastic Models*, 31(3): 361–402, 2015.

[149] A. Zocca, S.C. Borst, J.S.H. van Leeuwaarden, and F.R. Nardi. Delay performance in random-access grid networks. *Performance Evaluation*, 70(10):900–915, 2013.

SUMMARY

In this thesis we study mathematical models that capture the collective behavior of devices sharing a wireless medium in a distributed fashion by viewing them as repelling particles. Our research is motivated by fundamental challenges in wireless networks, which typically are very large and lack centralized control. Instead these networks vitally rely on a distributed mechanism for regulating the access of the various devices to the shared medium.

Thanks to their low implementation complexity, randomized algorithms provide a popular mechanism for distributed medium-access control. The work in this thesis is particularly concerned with the so-called Carrier-Sense Multiple-Access (CSMA) protocol, various incarnations of which are implemented in IEEE 802.11 networks. Despite its simplicity on a local level, the macroscopic behavior of the CSMA protocol in large networks tends to be exceedingly complex, and critically depends on global spatial characteristics of the network in non-intuitive ways.

We consider stylized stochastic models to understand how the spatial deployment of the various transmitter-receiver pairs affects the global performance of the network. Specifically, we model the random-access network as an interacting particle system on a graph, which captures the interplay of conflicting transmissions due to interference. The global evolution of such a particle system is described by a continuous-time Markov process, which exhibits fascinating connections with the hard-core interaction between gas particles studied in chemistry and statistical physics. Specific attention is paid to scenarios in which nodes become more aggressive in trying to activate, which is relevant for networks in high-load regimes, where one cannot afford to leave network resources unutilized. This scenario corresponds to the low-temperature regime for the corresponding interacting particle system. The most likely activity states for the network in this regime are those with a maximum number of active nodes, to which we refer as dominant states. Even in scenarios where all nodes have an equal opportunity to be active in the long run or in symmetric scenarios where spatial fairness is automatically ensured, transient yet significant starvation effects can arise due to the fact that the dominant states become extremely rigid, by which we mean that the transitions between dominant states can be extremely slow, causing starvation for the nodes not in the currently active dominant state.

Chapter 1 introduces the stochastic models for CSMA networks considered in this thesis and illustrates their equivalent description as interacting particle systems. Furthermore, we briefly indicate how the usage of multiple frequencies in such networks leads to a non-trivial trade-off between the aggregate throughput and average packet delay.

We examine the impact of slow transitions between dominant states on the delay performance in random-access networks in Chapter 2. Focusing on highly symmetric networks, we show how delays in symmetric partite graphs are intimately connected with the transition times between dominant states in a high-load scenario. In particular, we prove an asymptotic lower bound for the average steady-state delay for such networks.

In Chapter 3 we then focus on complete partite graphs, which are graphs that are useful for modeling dense networks, and therefore provide a "worst-case" perspective. Leveraging the presence of a unique bottleneck configuration for these networks, we establish a geometric-sum representation for the transition times using regenerative arguments, which we then use to obtain the asymptotic order-of-magnitude and scaled distribution in the regime where the activity rates become large.

We then we turn our attention to regular meshes, such as square and triangular lattices, and analyze how the temporal starvation of the nodes depends on the structure of these networks in high-load scenarios. In order to do so, we generalize in Chapter 4 the existing Metropolis Markov chain framework, which is commonly used to study the meta-stability of particle systems in statistical mechanics. In particular, we obtain asymptotic results in probability, in expectation and in distribution for any first hitting times in the corresponding energy landscape by analyzing the most likely paths that the Markov process follows in the low-temperature regime.

Furthermore, in Chapters 5 and 6 we develop a novel combinatorial method to analyze the structure and the features of the typical transition paths between dominant states on regular meshes. As a result, we then quantify how the order-of-magnitude of the transition times between dominant states depends on the size and properties of these network topologies. In particular, Chapter 5 is dedicated to grid graphs, while Chapter 6 is devoted to triangular grid graphs.

Lastly, in Chapter 7 we study the Widom-Rowlison model, an interacting particle system that we show to be the discrete-time counterpart of a multi-channel CSMA network when there are two available channels and the conflict graph is bipartite. Focusing on grid graphs, we study the asymptotic behavior of the transition time between its two dominant states in the low-temperature regime.

ABOUT THE AUTHOR

Alessandro Zocca was born in Bussolengo, Italy, on November 6, 1988. He completed his secondary education in 2007 at "Liceo Scientifico G. Galilei" in Verona. In the same year Alessandro started his studies in Mathematics at the University of Padova, Italy, from which he graduated *cum laude* with a bachelor's degree in July 2010. Subsequently, he was admitted to the Master of Advanced Studies in Mathematics (Part III) at University of Cambridge, UK, where he obtained his Master's degree with merit in July 2011.

In September 2011 he started a PhD project at Eindhoven University of Technology (TU/e) in the Stochastics group under the supervision of Sem Borst, Johan van Leeuwaarden and Francesca Nardi. His PhD research falls in the area of stochastic networks and focuses on the performance analysis of spatial stochastic models, also using techniques from statistical physics. This research, whose results are presented in this dissertation, led to a number of publications in scientific journals and conference proceedings.

During his employment at TU/e, Alessandro attended many workshops at EURANDOM and several conferences, among which Stochastic Networks 2012 (Boston), ValueTools 2012 (Cargèse, France), Performance 2013 (Vienna), INFORMS 2013 (San José, Costa Rica), Stochastic Networks 2014 (Amsterdam), and INFORMS 2015 (Istanbul). He presented his work on various occasions, giving talks at some of the above-mentioned events, but also at EPFL in Lausanne and at the Hamilton Institute in Maynooth.

Alessandro defends his PhD thesis at TU/e on December 17, 2015. Starting January 2016, Alessandro will work as a post-doctoral researcher at Centrum Wiskunde & Informatica in Amsterdam.