UNIVERSITÀ DEGLI STUDI DELL'AQUILA

Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica

Dottorato di Ricerca in Matematica e Modelli

XXXVIII CICLO

Titolo della tesi

The Matrix Product Ansatz from a probabilistic viewpoint

MATH-04/A

Dottoranda

Federica Iacovissi

Coordinatrice del corso

Prof.ssa Debora Amadori

Supervisor

Prof. Davide Gabrielli

Cosupervisor

Prof.ssa Anna De Masi

Anno Accademico 2024/2025

Acknowledgements

First of all, I would like to thank my supervisor, Davide Gabrielli, for all the work we have done together over these years. He has been a mentor and a constant reference point. Over these three years, I have learned so much. From his deep knowledge and his constant interest and intellectual curiosity, I have learned a great deal scientifically; from his availability, support, and enthusiasm, I have learned just as much on a personal level. Learning so much has deeply enriched me, both professionally and as a person. It is a privilege and an honor to work with him.

I would like to thank Giada Basile, my Master's thesis advisor, an encounter that was decisive in some of the choices I made, including my decision to pursue a PhD. Her help and advice have been fundamental in guiding me toward the direction I chose.

I would also like to thank my high school philosophy teacher, Daniela Papitto, who one September almost fourteen years ago entered a classroom of sixteen-year-olds - among whom I was one - reading a passage from Aristotle's *Metaphysics* on the sense of wonder with which human beings look at the world and begin to ask questions and seek answers. Many of the things that have led me here today have their roots there.

My thanks always go to my family for their quiet and ever-present support.

I am grateful to all the people I have met during these PhD years; each of them has contributed to this journey. Many of them have become close friends. A particular thanks to Antonio and Giulia, who over these years have been constant points of reference.

Contents

Abstract

In this thesis, we provide a probabilistic characterization of the class of probability measures that can be represented by the Matrix Product Ansatz (MPA). We introduce a constructive procedure, based on a suitable enlargement of the state space, showing that a probability measure admits a representation in terms of non-negative matrices via the MPA if and only if it can be expressed as a mixture of inhomogeneous product measures, where the mixing law is given by a Markov bridge. We illustrate this construction by applying it to several examples of interacting particle systems. Finally, we exploit the resulting probabilistic structure to derive large deviation principles for this class of measures.

Introduction

In equilibrium statistical mechanics, the Gibbs distribution provides a complete probabilistic description of a system. This measure, determined by the Hamiltonian rather than by the underlying dynamics, describes the probability of each configuration and allows one to derive macroscopic quantities from the microscopic behavior.

By contrast, in non-equilibrium statistical mechanics no universal principle exists; when a system is out of equilibrium very little is known about the statistics of its microscopic states, even in the stationary case. The stationary measure of a driven system is not known a priori and, in general, cannot be expressed in closed form: it depends on the specific microscopic dynamics.

To illustrate this setting, consider a finite volume containing interacting particles in contact with different reservoirs at the boundaries or under the action of external fields. In this situation currents (heat, electrical currents,...) are generated throughout the system. In general, when heat, particle or volume exchanges are reversible, the system is at equilibrium with its environment; when the exchanges are irreversible, the system is driven out of equilibrium.

The simplest non-equilibrium states are the stationary states of systems that do not conserve the total number of particles due to the interaction with different external reservoirs. This provides a way to model the external world: reservoirs can be viewed as mechanisms that inject and remove particles from the system.

In this context, stochastic lattice gases have provided useful tools for the study of properties of non-equilibrium states. Among them, there exists a class of models that admit a representation of the stationary measure in terms of a product of matrices, a construction known as the *Matrix Product Ansatz (MPA)* which provides an exact solution for the steady state.

The MPA is a remarkable algebraic construction which enables one to get a combinatorial representation of the invariant measure of a variety of non equilibrium stochastic particle systems. The power of this method is manifold: it has not only allowed to study several macroscopic quantities, such as density profiles, correlation functions, large deviations functionals, but it has also revealed the deep connection with some combinatorial objects, thus enlightening the mathematical structure behind the stationary states.

The aim of this thesis is to explore the probabilistic structure underlying the matrix product ansatz. Specifically, we provide a probabilistic characterization of the set of measures that can be represented in terms of product of matrices. By suitably enlarging the state space, we show that a probability measure can be described in terms of non negative matrices by the MPA, if and only if it can be written as a mixture of inhomogeneous product measure where the mixing law is a Markov bridge. Moreover, as we shall see, the measure on the enlarged state space is itself a Markov bridge.

We give a constructive procedure to identify such probabilistic features, based on the enlargement of the state space by a special coupling and on a spectral conjugation that produces stochastic matrices. We illustrate the result by examples and show that existing probabilistic representations of the invariant measures of non equilibrium interacting particle systems can be obtained from the matrix product ansatz by this general procedure.

Finally, we shall see that this construction, highlighting a hidden Markovian structure underlying the MPA, provides a suitable setting for deriving large deviations principles.

This thesis is organized as follows.

The first two chapters provide the necessary background and review known results that will be used in the sequel.

In Chapter 1, we introduce several examples of stochastic lattice gases that will be used throughout the thesis. Specifically, we deal with the exclusion processes (TASEP, ASEP and the 2-TASEP) and the harmonic process.

Chapter 2 is dedicated to the Matrix Product Ansatz. We first recall its application to the TASEP and explain how the stationary measure can be obtained explicitly. We then discuss its generalization to the ASEP and finally describe the case with second-class particles.

In Chapter 3 we investigate the probabilistic structure underlying the MPA, [43]. We state and prove a characterization theorem for the family of measures which can be represented in terms of a product of matrices. Using the constructive procedure on which the proof is based, we illustrate the result by applying the construction to some examples: the TASEP, the ASEP with two and four boundary parameters, the 2-TASEP on a ring and the harmonic process. Moreover, we show that, starting from the corresponding MPA representation and applying our transformation, one can recover some remarkable probabilistic representation of the invariant measure of the open TASEP and of the 2-TASEP, namely the *two lines ensemble* and the *queue representation*, respectively.

In Chapter 4, building on the construction introduced in the previous chapter, we derive some large deviation principles for MPA measures. We present several results that appear in [44]. Specifically, exploiting the Markov bridge representation on the enlarged state space we study the large deviations for this class of measure. We focus on the major distinction when the matrices are finite or infinite. In the finite case, we give a variational formula both for the algebraic and the spatial empirical measures that can be solved in special cases. For the infinite case we illustrate the method by an example, namely the invariant measure of the boundary driven TASEP model in a special regime.

Chapter 1

Models and Preliminaries

The aim of this chapter is to present the definitions and basic properties of the stochastic models that will be central to this work. We begin with the exclusion processes (TASEP and ASEP) and the notion of second–class particles, and then briefly recall the harmonic models that will also play a role in later chapters.

Informally, stochastic lattice gases are models of interacting particles evolving on a discrete lattice, whose evolution is governed by probabilistic local dynamics.

Time is continuous and each site or edge is equipped with an independent exponential clock. When the clock at site i rings, the configuration is updated: particles at that site may attempt to move to neighboring locations, or interact with the surrounding environment, depending on the dynamical rules of the process.

The models we will introduce are not only a common framework for the matrix product approach, but also the systems on which new results will be developed in the subsequent chapters.

1.1 The totally asymmetric simple exclusion process

The totally asymmetric exclusion process (TASEP) is a stochastic system of hopping particles which evolve according to a continuous–time Markov process. It is a fundamental and widely studied model of interacting particle systems and represents one of the simplest examples of a non-equilibrium system. For this reason, it has become a paradigm for the mathematical analysis of non–equilibrium phenomena. The TASEP belongs to the class of exclusion processes in which each site of the lattice can be occupied by at most one particle, in accordance with the exclusion principle that forbids multiple occupancy.

The particles are indistinguishable and the configuration space is $\{0,1\}^\Lambda$ with Λ a countable set. A particle configuration is defined by $\eta = (\eta_i \ : \ i \in \Lambda)$ where $\eta_i = 1$ denotes the presence of a particle at site i and $\eta_i = 0$ denotes an empty space. The lattice can, in principle, be any countable set: typical examples are regular lattices $\mathbb{Z}^d$, subsets of $\mathbb{Z}^d$, or the vertex set of a graph; in the one-dimensional TASEP, the latter choice is the most common. Depending on the choice of Λ one can consider different boundary conditions. In one dimension, for instance, if $\Lambda = \{1, ..., N\}$ the process is defined on a finite 1-dimensional open system with reservoirs attached to the boundaries; if instead $\Lambda = \mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ the dynamics are imposed on a ring, corresponding to periodic conditions.

In the present work, we will focus on 1-dimensional systems; for the TASEP and ASEP we consider the open system on the finite lattice $\{1, ..., N\}$ where particles can enter and exit through the boundaries, while for the 2-TASEP we will work on a 1-dimensional torus.

The TASEP is a continuous time Markov chain on $\{0,1\}^N$ and it is totally asymmetric, which means that the particles can move just in one direction. They jump with rate one to the nearest neighbor site to the right (i.e. from site i to site $i+1$) and the jump is suppressed

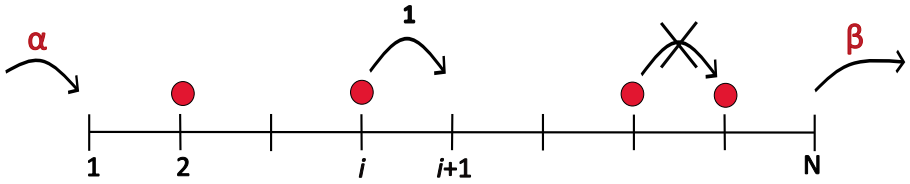

Figure 1.1: Dynamics of the TASEP on the finite lattice $\{1, ..., N\}$

if this state is already occupied. At the boundaries, particles are injected with rate α on the leftmost site 1, when empty, and particles in the rightmost site N are destroyed with rate β, see Figure 1.1. Formally, the bulk rates of jump of one particle are:

$$c_{i,i+1}(\eta) = \eta_i(1-\eta_{i+1}) \qquad c_{i+1,i} = 0 \qquad i = 1, ..., N-1$$

and the boundary rates are:

$$c_{0,1}(\eta) = (1-\eta_1)\alpha, \qquad c_{N,N+1}(\eta) = \eta_N\beta, \qquad c_{1,0}(\eta) = c_{N+1,N}(\eta) = 0.$$

where $0 \leq \alpha, \beta \leq 1$ are parameters. To clarify the notation we may think the reservoirs as site 0 with density α and site N+1 with density $1-\beta$. The generator is given by

$$\mathcal{L}_N^{TASEP} f(\eta) = \sum_{i=1}^{N-1} \eta_i(1-\eta_{i+1}) \left[f(\eta^{i,i+1}) - f(\eta) \right] \tag{1.1.1}$$

$$+ (1-\eta_1)\alpha \left[f(\eta^{1,+}) - f(\eta) \right] + \eta_N\beta \left[f(\eta^{N,-}) - f(\eta) \right], \tag{1.1.2}$$

with:

$$\eta_k^{i,j} = \begin{cases} \eta_k & \text{if } k \neq i,j \\ \eta_i & \text{if } k = j \\ \eta_j & \text{if } k = i \end{cases} \tag{1.1.3}$$

and

$$\eta_k^{i,+} = \begin{cases} \eta_k & \text{if } k \neq i \\ 1 & \text{if } k = i \end{cases} \qquad \eta_k^{i,-} = \begin{cases} \eta_k & \text{if } k \neq i \\ 0 & \text{if } k = i\,. \end{cases} \tag{1.1.4}$$

The process is irreducible except in the two degenerate cases $\alpha = 0$ or $\beta = 0$ for which the system admits absorbing states (e.g. $00\cdots011\cdots11$). The measures concentrated on these states are invariant. While, if $\alpha = 1-\beta$ the invariant measure of the process is $\prod_{i=1}^{N} \mathcal{B}_\alpha(\eta_i)$, where $\mathcal{B}_\alpha$ denotes the Bernoulli measure of parameter α.

In general, for any values of $\alpha, \beta \neq 0$, the invariant measure of the system admits a matrix product representation, as we shall see in more detail in the next chapter. Moreover, we will briefly explain that Catalan numbers, surprisingly, appear within the MPA formulation. This has led to the study and development of many combinatorial structures underlying the TASEP, thus providing several different expressions for the invariant measure.

Many works have focused on this topic, involving Catalan combinatorics of weighted paths and Catalan tableaux, in relation to different solutions of the matrix ansatz; we mention only

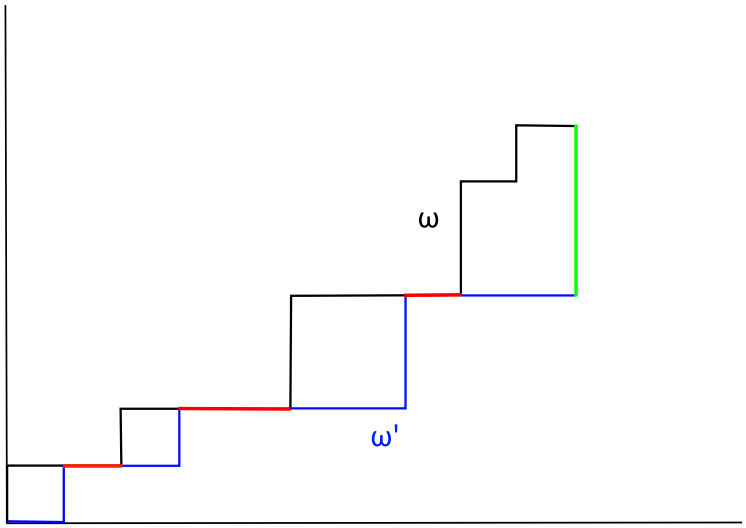

Figure 1.2: Example of two compatible paths ω (black) and ω' (blue), with ω' lying below ω. The red segments correspond to the common horizontal steps counted by $f(\omega, \omega')$, while the green segment indicates the vertical steps contributing to $g(\omega, \omega')$.

a few of them.

In [15, 16], Corteel and Williams proved a close connection between the *permutation tableaux* and the Matrix Product Ansatz, providing an interpretation of the invariant measure in terms of *Catalan tableaux*. Subsequently, Viennot [62], building on this work, interpreted the measure in terms of binary trees and consequently, in terms of paths.

In particular in [62] he provided an alternative characterization of the stationary distribution by establishing a correspondence between particle configurations and certain families of lattice paths $(\omega_i)_{i=0}^{N} \in (\mathbb{Z}^+)^2$. The path ω starts at $\omega_0 = (0, 0)$ and is composed of horizontal steps denoted by $e^{(1)} = (1, 0)$ and vertical steps denoted by $e^{(2)} = (0, 1)$. In this representation, each upward step encodes the presence of a particle, while each horizontal step encodes an empty site. More precisely, $\omega_x - \omega_{x-1} = e^{(2)}$ if and only if $\eta_x = 1$, and $\omega_x - \omega_{x-1} = e^{(1)}$ if and only if $\eta_x = 0$.

Therefore, a trajectory ω uniquely represents a configuration η and vice versa. Two trajectories are said to be *compatible* if they have the same endpoints and if one remains below the other. For a compatible pair (ω, ω') one defines $f(\omega, \omega')$ the number of common horizontal steps and $g(\omega, \omega')$ the number of final vertical steps in ω' after its last horizontal step (see Figure 1.2). This construction allows one to express the invariant measure in purely combinatorial terms. In particular, a probability measure depending on two parameters can be defined on pairs of compatible paths by:

$$\hat{\nu}_{\alpha,\beta}(\omega, \omega') := \frac{\alpha^{-f(\omega,\omega')}\beta^{-g(\omega,\omega')}}{Z_{\alpha,\beta}} . \tag{1.1.5}$$

Denote by $\nu_{\alpha,\beta}(\omega) = \sum_{\omega'} \hat{\nu}_{\alpha,\beta}(\omega, \omega')$, its marginal distribution. Since particle configurations and lattice paths are in bijection, this marginal directly induces a measure on configurations: the stationary weight of η is given by evaluating $\nu_{\alpha,\beta}$ at the corresponding path $\omega[\eta]$.

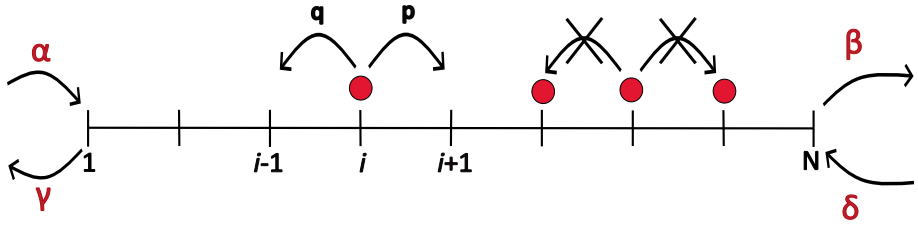

Figure 1.3: Dynamics of the ASEP on $\{1, ..., N\}$

In a different approach, Duchi and Schaeffer [30] avoided the matrix ansatz and gave a direct combinatorial interpretation by introducing a space of closed two-rows systems, called *complete configuration*, and defining a Markov process on this larger space that reproduces TASEP dynamics on the top row.

Recently, another representation of the invariant measure of the boundary driven TASEP has been obtained in [6] in terms of two lines ensemble, based on the approach of [2]. Further details will be discussed in Chapter 3, where we show its connection with our results.

1.2 The asymmetric simple exclusion process

The asymmetric simple exclusion process is the generalization of the totally asymmetric process. In this process, differently to the TASEP, particles are allowed to move in both directions with possibly different rates. As in the TASEP, the exclusion principle holds: each site of the lattice can be occupied by at most one particle, and attempted jumps are suppressed whenever the site is already occupied.

Consider a finite lattice $\{1, ..., N\}$, a particle at site i can jump to site $i+1$ with rate p or to site $i-1$ with rate q, when the destination is empty. At the boundaries, particles are injected at the leftmost site 1 with rate α and removed from it with rate γ; analogously, they are injected at the rightmost site with rate δ and removed with rate β, see Figure 1.3.

The bulk rates of jump of the process are defined by:

$$c_{i,i+1}(\eta) = p\eta_i(1-\eta_{i+1}), \qquad c_{i+1,i}(\eta) = q\eta_{i+1}(1-\eta_i), \qquad i = 1, ..., N-1$$

and the boundary rates:

$$c_{0,1}(\eta) = (1-\eta_1)\alpha, \quad c_{1,0}(\eta) = \eta_1\gamma, \quad c_{N+1,N}(\eta) = (1-\eta_N)\delta, \quad c_{N,N+1}(\eta) = \eta_N\beta,$$

with $p, q, \alpha, \beta, \gamma, \delta \geq 0$. As in the previous process, sites 0 and $N+1$ are interpreted as reservoirs with effective densities determined by the parameters. Then, the generator of the process is given by:

$$\begin{aligned}\mathcal{L}_N^{ASEP} f(\eta) = &\sum_{i=1}^{N-1} \Big[p\eta_i(1-\eta_{i+1})(f(\eta^{i,i+1}) - f(\eta)) + q\eta_{i+1}(1-\eta_i)(f(\eta^{i+1,i}) - f(\eta))\Big] \\ &+ (1-\eta_1)\alpha\left[f(\eta^{1,+}) - f(\eta)\right] + \eta_1\gamma\left[f(\eta^{1,-}) - f(\eta)\right] \\ &+ (1-\eta_N)\delta\left[f(\eta^{N,+}) - f(\eta)\right] + \eta_N\beta\left[f(\eta^{N,-}) - f(\eta)\right].\end{aligned}$$

The process is a continuous–time Markov chain on $\{0,1\}^N$. Depending on the values of the parameters, one recovers classical exclusion processes as special cases:

- The TASEP when $p = 1,\ q = 0,\ \gamma = \delta = 0$
- The symmetric exclusion process (SSEP) when $p = q$.

As in the TASEP, the characterization of the invariant measures for the open ASEP depends on the boundary parameters. This process also admits a description of the invariant measures through the Matrix Product Ansatz, we shall explain it in the next chapter, but there are some special situations in which the invariant measure takes a simpler form.

Consider first the symmetric case $p = q$, and denote with ρ_L and ρ_R the density of the left and right reservoirs:

$$\rho_L = \frac{\alpha}{\alpha+\gamma}, \qquad \rho_R = \frac{\delta}{\beta+\delta}.$$

If the density of the two reservoirs is equal, i.e. $\rho = \rho_L = \rho_R$, then the invariant measure of the system is the product of Bernoulli measures $\prod_i \mathcal{B}_\rho(\eta_i)$. From the physical point of view, this occurs because in the symmetric case the bulk cannot sustain any net current since rightward and leftward jumps occur with the same rate. Consequently, when the two reservoirs impose the same density the particles injection and exit at the boundaries balance exactly, and therefore the invariant measure is the Bernoulli product.

Moreover, this measure is not only invariant but also reversible for the process since the dynamics satisfies the detailed balance condition with respect to it.

In the asymmetric case, if the boundaries parameters satisfy the condition:

$$(p-q)(\alpha+\delta)(\beta+\gamma) = (\alpha+\delta+\beta+\gamma)(\alpha\beta-\gamma\delta) \tag{1.2.1}$$

the process admits as invariant distribution the product of Bernoulli measures $\prod_i \mathcal{B}_\rho(\eta_i)$ with parameter:

$$\rho = \frac{\alpha+\delta}{\alpha+\beta+\gamma+\delta}.$$

Since the bulk current is non-zero, the Bernoulli product measure is stationary only when this current is exactly compensated by the flux at the boundaries, which is precisely the content of the above condition.

Finally, degenerate choices of the boundary parameters may lead to absorbing states: if both entry rates vanish ($\alpha = \delta = 0$) the empty configuration is absorbing while if both exit rates are equal to zero ($\gamma = \beta = 0$) the fully occupied configuration is absorbing. In these situations the process is not irreducible because certain configurations cannot be reached.

1.3 Two species TASEP

The asymmetric exclusion process can be naturally extended to settings where particles are no more indistinguishable but belong to different classes. In these multi-species processes, particles, each represented by a distinct label according to the species, are indistinguishable within the same class but distinguishable across classes. Each site of the lattice may be

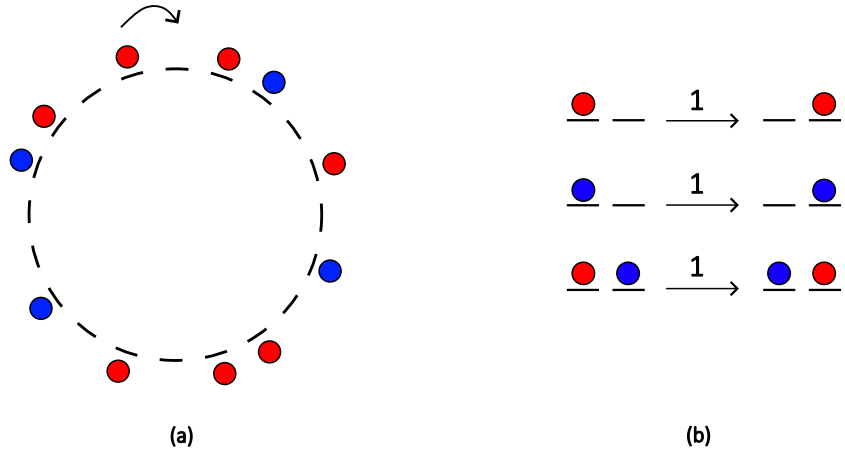

Figure 1.4: **(a)** Schematic representation of the two-species TASEP on a ring. Red balls denote first-class particles, blue balls denote second-class particles **(b)** Allowed transitions of the dynamics: a first-class particle moves into a vacancy, a second-class particle moves into a vacancy, and a first-class particle exchanges position with a second-class particle. Every jumps occur with rate 1.

occupied by particles of various type and the dynamics is governed by priority rules: particles of higher class have precedence over those of lower class, which implies that different types of particles interact differently when attempting to exchange positions.

This generalization enriches the model substantially; the study of two-species and, more generally, multi-species exclusion processes provides a natural framework for investigating systems governed by hierarchical constraints and competing priorities.
For giving a wider background, in the 2-ASEP on $\mathbb{Z}$, from a probabilistic point of view, the presence of a second-class particle offers a microscopic marker to track both shocks and characteristics in the hydrodynamic limit: when placed at a density discontinuity, its trajectory coincides with the position of the shock, while if the density is homogeneous, it follows the characteristic speed. This idea, developed by Ferrari (1992) [36], has become a fundamental result in the analysis of shocks.

On the other hand, in statistical physics the two-species TASEP has also been studied on the finite ring by Derrida *et al.* (1993) [22, 23], where the invariant measure can be obtained through the Matrix Product Ansatz; taking the infinite-volume limit then allows one to construct shock profiles.

Beyond shock analysis, multi-species exclusion processes have been explored in different contests. For instance, Ferrari and Martin [38] introduced multiline queues to describe the stationary distribution of N-TASEP, while Corteel and Williams [17] established connections with tableaux combinatorics. In general, these models are of independent interest; they not only provide examples of non-equilibrium phenomena but also show a natural representation of transport systems with multiple priority classes.

In what follows, we shall focus on the case of the two-species TASEP on the finite ring, since this is the setting to which we will return later in this thesis.
We consider the TASEP on $\mathbb{Z}_N$ the ring with N sites where particles can be of two different type, first and second class particles. The space configuration is $\{0,1,2\}^N$ and the value of η_i represents the class of the particle at i; this means that if $\eta_i = 0$ then the site $i \in \mathbb{Z}_N$ is

empty, if $\eta_i = 1$ there is a first class particle at i, while when $\eta_i = 2$ there is a second class particle at i. Since we are on a ring we use an equivalence modulo N for the indices on the lattice. The stochastic dynamics is defined by the generator

$$\mathcal{L}_N f(\eta) := \sum_{i=1}^{N} \mathbb{1}(r(\eta_i) > r(\eta_{i+1})) \left[f(\eta^{i,i+1}) - f(\eta) \right] , \tag{1.3.1}$$

where the symbol $\eta^{i,i+1}$ is defined in (1.1.3) and $r : \{0,1,2\} \to \{0,1,2\}$ is the priority function defined by: $r(0) = 0\,, r(1) = 2\,, r(2) = 1$; particles of higher class have lower priority. The informal description of the dynamics is as follows: at rate one the occupation variables at the extreme of a bond $(i, i+1)$ exchange their value when the priority of the variable at vertex i is higher than that at vertex $i+1$. The dynamics has two conserved quantities that are the number of first class particles (i.e. the number of particles with priority 2) $\mathcal{N}_1(\eta) := \sum_{i=1}^{N} \mathbb{1}(\eta_i = 1)$ and the number of second class particles (i.e. the number of particles with priority 1) $\mathcal{N}_2(\eta) := \sum_{i=1}^{N} \mathbb{1}(\eta_i = 2)$. This means that there is a two parameter family of invariant measures depending on the pair of integer numbers (n_1, n_2) that represents respectively the number of first and second class particles and are such that $n_1 + n_2 \leq N$.

As mentioned previously, we note that the model admits a natural extension to an arbitrary number of classes, the N-TASEP, with corresponding priority rules and conserved quantities; however, in this thesis, we restrict to the two-species case, which is the only one that will be used in the sequel.

1.4 Harmonic models

The harmonic models are a family of integrable interacting particle systems with both discrete and continuous variants. These models have been introduced in [42] and a combinatorial representation of the invariant measure, on a N lattice with open boundary, has been obtained in [41] by combining probabilistic arguments with techniques inspired by integrable systems. In the discrete version, each site of a lattice can be occupied by an arbitrary number of particles and if a site contains η_i particles, any block of $k \leq \eta_i$ of them may jump together to a neighboring site with a certain rate. A continuous analogue, which can be interpreted as a stochastic model of heat transport, arises as a scaling limit of the discrete process. We will not discuss the continuous version in detail here, and we only mention it for completeness.
Both models form a one-parameter family indexed by a positive spin value $s > 0$. This parameter originates from the representation theory of non-compact quantum spin chains and determines the precise form of the dynamics and invariant measures. The special case $s = \frac{1}{2}$ corresponds to the version where blocks of k particles move together with rate $1/k$.
Within this algebraic framework, the discrete and continuous variants are further connected by a moment duality, which establishes a direct correspondence between their observables [39]. For their structure, the harmonic models belong to the broad family of zero-range processes, in the sense that jump rates depend only on the number of particles at the departure site. However, they differ from the classical zero-range dynamics: rather than moving single particles with a rate function $g(\eta_i)$, they allow jumps of multiple particles. Thus, in the presence of boundary reservoirs, the stationary non-equilibrium states are no longer of product type and exhibit non-trivial correlations across sites.

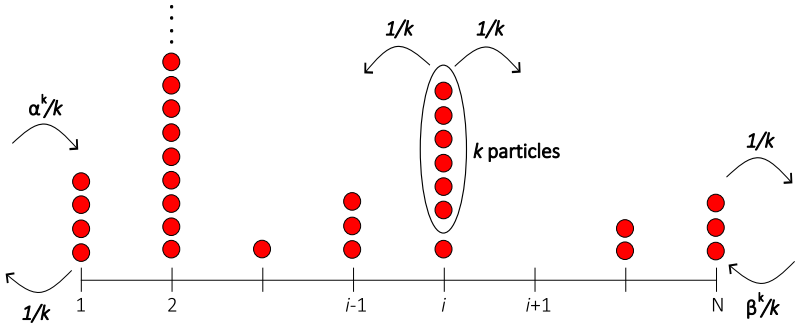

Figure 1.5: Boundary-driven harmonic model on a one-dimensional lattice.

We consider only the simplest case, which will be useful for future discussion. We have one-dimensional lattice $\{1, ..., N\}$ with two external sources at the left and right boundaries having two parameters $0 < \alpha \leq \beta < 1$, respectively (the case $\beta < \alpha$ can be considered just exchanging left with right). On each lattice site we have an arbitrary number of particles and we denote by $\eta_i \in \mathbb{N}_0$ the number (possibly zero) of particles at site i. From site i a block of $k \leq \eta_i$ particles may jump to site $i-1$ or $i+1$ with rate $1/k$. The boundary dynamics is defined as follows: at site 1, a block of k particles exits with rate $1/k$ and enters with rate α^k/k. At site N, the same mechanism holds with the density of the reservoir given by β, see Figure 1.5. Then, the stochastic dynamics has a bulk and a boundary part which are described in terms of the generator $\mathcal{L}_N^H$ defined below. Denoting by δ^i the configuration defined by $\delta_j^i = 0$ when $j \neq i$ and $\delta_i^i = 1$ for any $i \in \{1, \ldots, N\}$ we have:

$$\mathcal{L}_N^H := \mathcal{L}_N^{\text{bulk}} + \mathcal{L}_N^{\text{bound}}. \tag{1.4.1}$$

The bulk generator is given by:

$$\mathcal{L}_N^{\text{bulk}} f(\eta) = \sum_{\substack{i,j \\ |i-j|=1}} \sum_{k=1}^{\eta_i} \frac{1}{k} \left[f(\eta - k\delta^i + k\delta^j) - f(\eta) \right]. \tag{1.4.2}$$

The boundary part which encodes the interaction with the reservoirs is:

$$\begin{aligned} \mathcal{L}_N^{\text{bound}} f(\eta) = &\sum_{k=1}^{\eta_1} \frac{1}{k} \left[f(\eta - k\delta^1) - f(\eta) \right] + \sum_{k=1}^{\infty} \frac{\alpha^k}{k} \left[f(\eta + k\delta^1) - f(\eta) \right] \\ &+ \sum_{k=1}^{\eta_N} \frac{1}{k} \left[f(\eta - k\delta^N) - f(\eta) \right] + \sum_{k=1}^{\infty} \frac{\beta^k}{k} \left[f(\eta + k\delta^N) - f(\eta) \right]. \end{aligned} \tag{1.4.3}$$

The process is a continuous-time Markov chain $\{\eta(t), t \geq 0\}$ whose state space is the set $\mathbb{N}_0^N$ of configurations $(\eta_1, ..., \eta_N)$.
For equal boundary densities $\alpha = \beta$ the process is reversible and the stationary distribution is a product of geometric marginals. Conversely, if $\alpha < \beta$ the process is irreversible; in [11] it has been proved that for unequal reservoirs densities the invariant measure becomes a convex mixture of inhomogeneous products of geometric distributions, with the mixing law described

in terms of the order statistics of i.i.d. uniform random variables in the interval $[\alpha, \beta]$. This result, first obtained for the harmonic model with $s = \frac{1}{2}$, has been generalized to arbitrary spin value s in [10].

This representation provides a probabilistic interpretation of the long-range correlations that arise in the stationary regime. We will return to this explicit characterization of the invariant measure in Chapter 3 where it will be discussed in more detail in connection with the novel method that we introduce in this work.

Chapter 2

Matrix product ansatz

2.1 Background and motivation

The Matrix Product Ansatz (MPA) is an algebraic technique introduced to explicitly describe the invariant measures of a variety of non-equilibrium stochastic particle systems. It has a long history, starting from the early 1990s when it was first introduced by statistical physicists as a matrix product solution of the ASEP, and it has since evolved through a large number of applications to many different models (see [4, 63] and [20] for a review of some applications). The family of asymmetric exclusion processes in one dimension is one of the simplest and most studied classes of driven diffusive systems. They are motivated by transport phenomena in physics, and their analytical tractability has made them key models for understanding how non-equilibrium behavior differs from equilibrium. In some cases, these models exhibit phase transitions in one dimension while the equilibrium counterpart does not show this phenomenon. In addition, the stationary distribution in the non-equilibrium regime exhibits a rich structure of considerable mathematical interest.

The first solution [21] relied on recursion relations between the statistical weights of configurations on lattices of different lengths. In particular, the steady state for a system of size N can be derived exactly as a simple recursion relation on the size of the system. We point out that an exact solution for the stationary weights based on recursion relations was independently found in [55], building on a previous work [21] where the reservoir rates were taken to be $\alpha = \beta = 1$.

The recursion method made it possible to compute, in the steady state, exact expressions for the density profile as well as for correlation functions of order higher than one. However, the step from the recursion to the explicit expressions of these quantities relied on a very complicated method involving generating functions.

The main advance came with the formulation of the MPA [23], where these recursion relations were replaced by simple algebraic rules between fixed matrices and vectors. This allowed for a compact and explicit representation of the stationary measure, and at the same time simplified the derivation of previous results and facilitated their generalization.

A matrix product state is built by associating a matrix to each site according to its occupation state and taking the ordered product over the entire lattice. As we shall see, this product is then reduced to a scalar probability by appropriately contracting: through two boundary vectors in the case of open systems with boundary reservoirs, or involving the trace operator in the case of closed systems, such as on a ring.

The MPA has had a significant impact on non-equilibrium statistical mechanics, as it provides a unified algebraic framework, reveals deep combinatorial connections, and applies broadly to models with different boundary conditions and to infinite-volume settings. The exact form of the invariant measure of the ASEP reveals, as the system size increases, number

sequences that are recurrent in enumerative combinatorics, such as Catalan numbers, ballot numbers and other classical families. These sequences emerge from one-to-one correspondences between particle configurations in the ASEP and larger classes of combinatorial objects, often lattice paths (e.g. Motzkin paths, dominated paths etc.). The recurrent appearance of these combinatorial structures has highlighted that the MPA is not just a "physical trick" or an "algebraic device" but rather reveals the deep mathematical structure underlying the stationary states.

Moreover, in several cases, the Matrix Product Ansatz provides an explicit and tractable structure for the invariant measure, which makes it possible to analyze macroscopic properties such as current, density profile, correlation functions and large deviation principles. This applies in particular to algebraic and spatial empirical measures, although the nature of the rate function may depend strongly on the finiteness or infiniteness of the underlying matrices.

We point out that measures of MPA type are also studied in informatics and pattern statistics under the name of *rational models* (see, for example, [48]). This underlines the unifying character of the Ansatz and its relevance across different domains.

In this chapter, after presenting the abstract formulation of the Matrix Product Ansatz, we illustrate its application to the TASEP, discuss algebraic properties and explicit representations, and extend the framework to the ASEP and to multi-species systems. The MPA developed for these models, as mentioned in the Introduction, will provide useful tools in the following chapters.

The chapter includes an analysis of partition functions and correlations, which illustrate the efficiency of the MPA in producing exact results for physically relevant observables, and briefly show its natural connection with combinatorics.

2.2 Abstract formulation of the Matrix Product Ansatz

We consider two finite or countable alphabets A and B. The set A is the state space in which the variables we study assume values, while B is an auxiliary state space. We discuss notation using A but a similar notation holds for B too and any other set.

Given $\eta = (\eta_i)_{i=1}^N, \in A^N$, we denote by $\eta_i^j = (\eta_i, \dots, \eta_j) \in A^{j-i+1}$, the finite portion of the word η contained between the indices $i < j$. Given $\eta, \gamma \in A^N$ we denote by $\eta_i^j \gamma_l^m = (\eta_i, \dots, \eta_j, \gamma_l \dots, \gamma_m) \in A^{j+m-i-l+2}$ the concatenation.

We remind that, given a set S, we denote by $\mathcal{M}^1(S)$ the probability measures on S. Given $\mu_N \in \mathcal{M}^1(A^N)$ and $\eta \in A^N$ we denote by $\mu_N(\eta_i^j) := \sum_{\{\gamma \in A^N : \gamma_i^j = \eta_i^j\}} \mu_N(\gamma)$ the probability of the cylinder set associated to η_i^j. The probability of the cylinder set determined by the single value η_i is denoted by $\mu_N(\eta_i)$.

We introduce a family of probability measures, which we call of *Matrix Product Ansatz* (MPA) type, that forms a special subset of $\mathcal{M}^1(A^N)$. For any element $a \in A$ we have a $B \times B$ matrix M^a having non-negative entries, i.e. for any $b, b' \in B$ we have $M^a_{b,b'} \geq 0$. We also consider $\vec{x} := (x_b)_{b \in B}$, $\vec{y} = (y_b)_{b \in B} \in \mathbb{R}_+^B$ two column vectors having non negative entries, and denote by $\vec{x}^T, \vec{y}^T$ the corresponding row vectors obtained by transposition.

Throughout this work, in order to avoid any ambiguity in the notation, we use the symbol $M^{(n)}$ to denote the n-th power of a matrix and the symbol M^a to denote the matrix associated with the element $a \in A$.

The Matrix Product Ansatz is a way of constructing probability measures on words of

length N by assigning to each symbol a non-negative matrix and evaluating matrix product between boundary vectors. The probability of a configuration is proportional to this product and the normalization factor ensures that the total probability sums to one.

Definition 2.1. *An element $\mu_N \in \mathcal{M}^1(A^N)$ is of Matrix Product Anstatz (MPA) type if it can be written as*

$$\mu_N(\eta) := \frac{\vec{y}^T \left(\prod_{i=1}^N M^{\eta_i}\right) \vec{x}}{Z_N}, \qquad \eta \in A^N, \tag{2.2.1}$$

where $\vec{x}, \vec{y}$ and $(M^a)_{a\in A}$ are non negative vectors and matrices and Z_N is a normalization factor. The product of the matrices and vectors is the usual row by column product.

The probability measure μ_N depends on the family of matrices and on the vectors but we will make this dependence explicit only when necessary. When $|A|, |B| < +\infty$ the measure (2.2.1) is always well defined while in the countable infinite cases the measure exists only when Z_N is finite for every N, i.e. only if the sum of weights associated with all possible configurations is convergent.

We call $M := \sum_{a\in A} M^a$ and the normalization term is given by

$$Z_N = \sum_{\eta\in A^N} \vec{y}^T \left(\prod_{i=1}^N M^{\eta_i}\right) \vec{x} = \vec{y}^T M^{(N)} \vec{x}. \tag{2.2.2}$$

with $M^{(N)}$ the N-th power of the matrix M.

In the next section we introduce and discuss the MPA for the TASEP, which corresponds to the case where the state space is $A = \{0, 1\}$ and the auxiliary space, in general, is $B = \mathbb{N} \cup \{0\} := \mathbb{N}_0$, apart from some special cases that will be mentioned below.

2.3 The matrix approach of the TASEP

For the sake of simplicity, and since all other cases can be adapted along similar lines, we shall first focus on proving the matrix product form of the invariant measure for the boundary-driven TASEP, following the original formulation of the MPA introduced in [23].

Let us first recall briefly the dynamics of the model with open boundary conditions we introduced in the previous chapter. A particle jumps with rate one to the nearest neighbor to the right if the site is empty otherwise if it is occupied, the particle remains in its position. Particles are injected at site 1 with rate α and exit from site N with rate β.

As time evolves, the chain converges to a steady state in which the probabilities $\mu_N(\eta_1, \eta_2, ..., \eta_N)$ of finding the system in configuration $\eta = (\eta_1, \eta_2, ...\eta_N)$ are stationary:

$$\frac{d}{dt}\mu_N(\eta_1, \eta_2, ..., \eta_N) = 0. \tag{2.3.1}$$

Given $\eta = (\eta_1, \eta_2, ..., \eta_N) \in \{0, 1\}^N$ a TASEP configuration of size N, we consider the measure defined as an ordered product of matrices:

$$\mu_N(\eta) = \frac{f_N(\eta)}{Z_N} = \frac{\vec{y}^T M^{\eta_1} M^{\eta_2} \cdots M^{\eta_N} \vec{x}}{Z_N} \tag{2.3.2}$$

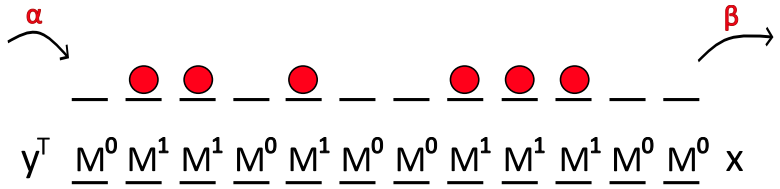

Figure 2.1: Matrix product representation obtained by encoding a particle configuration

where Z_N is the normalization factor, the two column vectors $\vec{x}, \vec{y}$ reduce the product to a scalar. The matrix M^{η_i} is defined to be M^1 if site i is occupied by a particle, i.e. $\eta_i = 1$, and M^0 if site i is empty, i.e. $\eta_i = 0$. The matrix M^{η_i} does not depend on the site label i but on the state of that site. Moreover, we note that this is a sort of procedure in which a particles configuration is encoded as a string of matrices and the initial and final vectors describe the presence of the reservoirs, see Figure 2.1.

The matrix product form (2.3.2) is a solution of the steady-state master equation (2.3.1) if the matrices M^0, M^1 and the vectors $\vec{x}, \vec{y}$ satisfy the following algebraic relations:

$$\begin{aligned} M^0 + M^1 &= M^1 M^0 \\ \vec{y}^T M^0 &= \frac{1}{\alpha}\vec{y}^T \\ M^1 \vec{x} &= \frac{1}{\beta}\vec{x}. \end{aligned} \tag{2.3.3}$$

Let us start by writing down the master equation:

$$\frac{d}{dt}\mu_N = L^T \mu_N \tag{2.3.4}$$

where L is the generator associated with the process (1.1.1) and describes the transition rates $c(\eta', \eta)$ from a particle configuration η' to another η. More precisely, the operator L^T is defined as:

$$L^T(\eta, \eta') = \begin{cases} \tilde{c}(\eta, \eta') & \text{if } \eta \neq \eta' \\ -\sum_{\eta \neq \eta'} \tilde{c}(\eta, \eta') & \text{otherwise} \end{cases}$$

where $\tilde{c}(\eta, \eta') = c(\eta', \eta)$. The non-diagonal terms $L^T(\eta, \eta') = \tilde{c}(\eta, \eta')$ represent the effective transition rates from a configuration η' to η, while the diagonal entries do not correspond to physical transition rates, but to the total exit rate from state η with a negative sign. Moreover, we observe that the master equation can be rewritten in the same way in terms of f_N.

The crucial idea underlying the proof is to rewrite the right-hand of the equation (2.3.4) in a suitable expanded form. We now analyze all possible transitions at the boundaries and in the bulk. To make explicit the structure of the global operator L^T, it is convenient to decompose it into local operators acting on small subsets of sites:

$$L^T = L_1^T + \sum_{i=1}^{N-1} L_{i,i+1}^T + L_N^T$$

where L_1^T denotes the local operator that acts on the first site and encodes the interaction with the left boundary reservoir, L_N^T the local operator that acts on the last site and encodes the interaction with the right boundary reservoirs and, finally, $L_{i,i+1}^T$ that one which acts on two neighboring sites $(i, i+1)$ in the bulk, describing the internal jumps. Since, in the exclusion process, each site can be either occupied or empty, the single site operators, such as L_1^T or L_N^T, act on the two possible local configurations of the single site, that is $(0), (1)$ and are therefore represented by 2×2 matrices. Differently, the operators acting on two sites, such as $L_{i,i+1}^T$, act on the four possible pairs of local states $(0,0), (0,1), (1,0), (1,1)$ and are represented by $2^2 \times 2^2$ matrices. More generally, if each site could take m possible values, the corresponding local operators would act on spaces of dimensions m and m^2, respectively. Each local operator acts in a non-trivial way only on the sites it corresponds to, and as the identity on the other sites. In this sense, the matrix that describes the global operator L^T is obtained by extending these local matrices into the full system through tensor products with identity operators acting on the remaining sites.

At site 1, when it is empty, the only possible transition is due to the particle injection:

$$(0, \eta_2, \eta_3, ..., \eta_N) \longrightarrow (1, \eta_2, \eta_3, ..., \eta_N)$$

which occurs with rate $c_1(0,1) = \alpha$. Consequently, the total exit rate from state $(0, \eta_2, ..., \eta_N)$ is α, so that the corresponding diagonal entry in L_1^T is $c_1(0,0) = -\alpha$. We represent the local operator L_1^T as a matrix by indexing its rows and columns according to the local states of the first site, ordered as $\{0; 1\}$, where 0 denotes an empty site and 1 an occupied one. With this convention, the matrix associated with the left-boundary transitions is given by:

$$L_1^T = \begin{bmatrix} -\alpha & 0 \\ \alpha & 0 \end{bmatrix}.$$

Similarly, at site N, when it is occupied by a particle, the only possible transition is due to particle exit:

$$(\eta_1, \eta_2, ..., \eta_{N-1}, 1) \longrightarrow (\eta_1, \eta_2, ..., \eta_{N-1}, 0)$$

which occurs with rate $c_N(1,0) = \beta$. The total exit rate from state $(\eta_1, ..., \eta_{N-1}, 1)$ is β and the corresponding diagonal entry in L_N^T is $c_N(1,1) = -\beta$ and then the matrix associated with the right-boundary transition is of the form:

$$L_N^T = \begin{bmatrix} 0 & \beta \\ 0 & -\beta \end{bmatrix}.$$

Regarding the bulk, the total asymmetry of the model implies that the only possible transition is a particle jump from site i to site $i+1$:

$$(\eta_1, ..., \eta_{i-1}, 1, 0, \eta_{i+2}, ..., \eta_N) \longrightarrow (\eta_1, ..., \eta_{i-1}, 0, 1, \eta_{i+2}, ..., \eta_N)$$

whose rate is $c((1,0),(0,1)) = 1$ and the total exit rate from this state is 1, then $c((1,0),(1,0)) = -1$. We represent the operator $L_{i,i+1}^T$ as a matrix by indexing its rows and columns according to the four local configurations $\{(0,0); (0,1); (1,0); (1,1)\}$, taken in this order. With this

convention, the matrix describing the bulk transition is a 4×4 matrix given by:

$$L^T_{i,i+1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We explicitly write the master equation (2.3.4) in the following way:

$$\begin{aligned} \frac{d}{dt}\mu_N(\eta_1, \eta_2, ..., \eta_N) = &\sum_{\tau_1 \in \{0,1\}} c_1(\tau_1, \eta_1)\mu_N(\tau_1, \eta_2, ..., \eta_N) + \sum_{\tau_N \in \{0,1\}} c_N(\tau_N, \eta_N)\mu_N(\eta_1, ..., \eta_{N-1}, \tau_N) \\ &+ \sum_{i=1}^{N-1} \sum_{\tau_i, \tau_{i+1} \in \{0,1\}} c((\tau_i, \tau_{i+1}), (\eta_i, \eta_{i+1}))\mu_N(\eta_1, ..., \tau_i, \tau_{i+1}, ..., \eta_N). \end{aligned} \tag{2.3.5}$$

We note that if there exist two coefficients x_0, x_1 such that the measure μ_N satisfies, for any choice of η_i, the conditions:

$$\sum_{\tau_1 \in \{0,1\}} c_1(\tau_1, \eta_1)\mu_N(\tau_1, \eta_2, ..., \eta_N) = x_{\eta_1}\mu_{N-1}(\eta_2, ..., \eta_N) \tag{2.3.6}$$

$$\sum_{\tau_N \in \{0,1\}} c_N(\tau_N, \eta_N)\mu_N(\eta_1, ...\eta_{N-1}, \tau_N) = -x_{\eta_N}\mu_{N-1}(\eta_1, ..., \eta_{N-1}) \tag{2.3.7}$$

$$\begin{aligned} \sum_{\tau_i, \tau_{i+1} \in \{0,1\}} &c((\tau_i, \tau_{i+1}),(\eta_i, \eta_{i+1}))\mu_N(\eta_1, ..., \tau_i, \tau_{i+1}, ..., \eta_N) \\ &= - x_{\eta_i}\mu_{N-1}(\eta_1, ..., \eta_{i-1}, \eta_{i+1}, ..., \eta_N) \\ &\quad + x_{\eta_{i+1}}\mu_{N-1}(\eta_1, ..., \eta_i, \eta_{i+2}, ..., \eta_N) \end{aligned} \tag{2.3.8}$$

then in (2.3.5) summing over all internal sites, the bulk contributions yield a telescopic sum which cancels with the boundary terms. Therefore, the right-hand site of (2.3.5) is equal to zero, and the measure μ_N is an invariant measure for the process generated by L.
The problem is reduced to determining two coefficients x_0 and x_1 such that the matrices M^1, M^0 and the vectors $\vec{x}$, $\vec{y}$ satisfy:

$$\begin{aligned} &\alpha \vec{y}^T M^0 = x_1 \vec{y}^T = -x_0 \vec{y}^T \\ &\beta M^1 \vec{x} = x_1 \vec{x} = -x_0 \vec{x} \\ &M^1 M^0 = -x_0 M^1 + x_1 M^0. \end{aligned}$$

From these relations it follows that $x_1 = -x_0$, and by choosing $x_1 = 1$ we obtain exactly the algebraic conditions in (2.3.3). Note that fixing $x_1 = 1$ is without loss of generality, since choosing any other constant would only rescale the matrices. For a more detailed explanation, see [23].

2.3.1 Algebraic properties and explicit matrix representations

In the framework of the Matrix Product Ansatz, the matrices M^0 and M^1 are introduced as abstract generators of a non-commutative algebra specified by defining bulk relations and boundary conditions. This algebraic formulation already suffices to characterize the stationary measure and to derive many of its properties. Nevertheless, explicit matrix representations of the algebra play an important role, as they provide concrete realizations of the framework and can be exploited for further analytical and combinatorial computations. As we shall see in the next chapter, these representations will also serve as the basis for the novel results we obtained.
Before introducing some explicit matrix forms, we first analyze their properties:

- If M^0 and M^1 commute, using the algebraic relations (2.3.3), one obtains:

$$\frac{1}{\alpha\beta}\vec{y}\cdot\vec{x} = \vec{y}(M^0M^1)\vec{x} = \vec{y}(M^0+M^1)\vec{x} = \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\vec{y}\cdot\vec{x}.$$

  Since $\vec{y}\cdot\vec{x}\neq 0$, it follows that $\alpha+\beta=1$. In this special case the matrices reduce to scalars:

$$M^1 = \frac{1}{1-\alpha}, \qquad M^0 = \frac{1}{\alpha}$$

  and $\vec{y}=\vec{x}=1$. The unique stationary distribution is the Bernoulli product measure $\prod_i \mathcal{B}_\alpha(\eta_i)$.

- If M^0 and M^1 do not commute, the representation is necessarily infinite dimensional.

The previous relations (2.3.2) do not uniquely identify the matrices M^0, M^1 indeed there are several different representations for different values of the parameters. By convention, we fix the normalization to $\vec{y}^T\vec{x}=1$. Other choices would only differ by a constant rescaling, without modifying the stationary measure.
For $\alpha+\beta>1$, one possible choice of matrices and vectors that satisfy the algebraic relations of the MPA for the TASEP is given by:

$$M^0 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots \\ 1 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad M^1 = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 1 & \ddots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.3.9}$$

$$\vec{y}^T = \hat{k}\left(1, \frac{1-\alpha}{\alpha}, \left(\frac{1-\alpha}{\alpha}\right)^2, \cdots\right) \qquad \vec{x} = \hat{k}\begin{pmatrix} 1 \\ \frac{1-\beta}{\beta} \\ \left(\frac{1-\beta}{\beta}\right)^2 \\ \vdots \end{pmatrix}$$

where $\hat{k} = \sqrt{(\alpha+\beta-1)/\alpha\beta}$ is the normalization factor that we fix in such a way that $\vec{y}^T\vec{x} = 1$.
Another possible choice, which instead holds for every value of $\alpha, \beta \in (0,1)$ is:

$$M^0 = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & \cdots \\ 1 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad M^1 = \begin{bmatrix} \frac{1}{\beta} & \frac{1}{\beta} & \frac{1}{\beta} & \cdots & \frac{1}{\beta} & \cdots \\ 0 & 1 & 1 & 1 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & 1 & \ddots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.3.10}$$

with the vectors:

$$\vec{y}^T = \left(1, \frac{1}{\alpha}, \left(\frac{1}{\alpha}\right)^2, \cdots\right), \qquad \vec{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

The first representation has the advantage of making the particle-hole symmetry of the problem clear, since the matrices M^0 and M^1 have very similar forms and the boundary parameters α, β only appear in the vectors. However, as already mentioned, this choice is not valid in the whole parameter range, the product $\vec{y}^T\vec{x}$ leads to divergent geometric series when $\alpha + \beta < 1$. The second representation, although more complex, is well defined for all $\alpha, \beta \in (0,1)$.
Both representations of the matrices can be found in the original work of Derrida *et al.* [23]. We have presented these explicit forms because they will be used in the following chapters, but it is important to stress that the algebra admits infinitely many possible representations.

2.4 General formulation of the ASEP

One of the main reasons for the impact of the MPA is the possibility to extend it beyond the totally asymmetric case to more general situations. Having established the matrix product formulation for the TASEP, we now turn to its generalization to the ASEP. In this case the state space remains $A = \{0, 1\}$ and the auxiliary space is, in general, again $\mathbb{N}_0$. In this more general setting, the algebraic relations acquire a q-deformed structure that reflects the presence of both rightward and leftward jumps.
Let us recall that particles jump one step to the right with rate p and one step to the left with rate q. To simplify the notation, we first set $p = 1$ so that the bulk relation reduces to the q-deformed algebra:

$$M^1M^0 - qM^0M^1 = M^1 + M^0.$$

At the boundary we consider the general parametrization with four independent rates $(\alpha, \beta, \gamma, \delta)$: at the left boundary particles enter with rate α and exit with rate γ and at the right boundary particles leave with rate β and enter with rate δ. The general case with arbitrary p will be discussed in the remark below.

For this system, the invariant measure can be written in terms of matrices:

$$\mu_N(\eta) = \frac{\vec{y}^T \left[\prod_{i=1}^N M^{\eta_i}\right] \vec{x}}{Z_N}$$

when $M^0, M^1, \vec{y}^T$ and $\vec{x}$ satisfy:

$$\begin{cases} M^1 M^0 - q M^0 M^1 = M^1 + M^0 \\ \vec{y}^T[\alpha M^0 - \gamma M^1] = \vec{y}^T \\ [\beta M^1 - \delta M^0]\vec{x} = \vec{x}. \end{cases} \tag{2.4.1}$$

The proof that the measure μ_N, expressed as a product of matrices under conditions (3.4.1) is invariant for the ASEP follows the same line of argument as for the TASEP, with the transition rates appropriately replaced. Note that by taking $q = 1$ the model reduces to the symmetric exclusion process (SEP) with the corresponding algebraic relations for the matrices, whereas if we take $q = 0$ and $\gamma = \delta = 0$ we recover the totally asymmetric exclusion process discussed previously with the algebraic relation (2.3.3).
Similarly to the TASEP, the MPA for the open ASEP generally involves infinite-dimensional matrices. For special values of the parameters, discussed in the previous chapter (see formula (1.2.1)), the matrices reduce to scalars, and the corresponding invariant measure becomes the Bernoulli product.

Remark 2.2 (Bulk relation with generic p, q)**.** *In the general ASEP with arbitrary hopping rate p and q, the relation of the bulk becomes:*

$$pM^1M^0 - qM^0M^1 = M^1 + M^0$$

with the same boundary conditions for the matrices M^0 and M^1. This choice does not involve any loss of generality, since all results can be recovered after a rescaling of time by p. Explicit matrix representations associated with this general process can be found in [24].

Remark 2.3 (Finite-dimensional representations)**.** *For completeness, we point out that it was shown in [33] that finite-dimensional representation can occur under specific conditions. In this work, the authors introduce the function:*

$$\kappa_+(\alpha,\gamma) = \frac{p - q + \alpha - \gamma + \sqrt{(p - q + \alpha - \gamma)^2 + 4\alpha\gamma}}{2\alpha},$$

analogously $\kappa_+(\beta, \delta)$. They proved that finite-dimensional representations of dimension at least two exist in the region of the phase diagram characterized by the condition:

$$\kappa_+(\beta,\delta) > \frac{1}{\kappa_+(\alpha,\gamma)}.$$

The region identified by this inequality lies within the low-density (LD) and high-density (HD) phases of the ASEP.
When the inequality is replaced by an equality, the algebra admits a one-dimensional representation. In the totally asymmetric case $p = 1$, $q = 0$ and $\gamma = \delta = 0$, one recovers the result, we

described previously noted by Derrida et al.: along the line $\alpha = 1 - \beta$ the stationary state is the Bernoulli product measure which corresponds to a one-dimensional representation algebra. In [33] one and two dimensional representations were constructed. Later, in [53] a complete classification of all finite-dimensional irreducible representations of this algebra was provided, with an explicit construction of such representations for every possible finite dimension.

2.5 Partition functions and correlations

In the present section, where powers of matrices will appear, we recall that parentheses in the exponent are reserved for matrix powers: in particular, $M^{(n)}$ denotes the n-th power of M, whereas M^a continues to denote the matrix associated with the state.

In general, for any measure $\mu_N \in \mathcal{M}^1(A^N)$ of MPA type, introduced in the definition 2.1, the partition function for a system of size N can be written in the compact form:

$$Z_N = \vec{y}^T M^{(N)} \vec{x}, \qquad M = \sum_{a \in A} M^a.$$

The expansion of $M^{(N)}$ in the non-commutative setting produces all ordered products of length N of the matrices $\{M^a\}_{a \in A}$, thus accounting for all possible configurations of the system. Given $\eta \in A^N$ and an element $a_1 \in A$, it follows immediately that the first-order correlation function can be calculated by:

$$\mu_N(\eta : \eta_i = a_1) = \frac{\vec{y}^T M^{(i-1)} M^{a_1} M^{(N-i)} \vec{x}}{Z_N}$$

Similarly, higher-order correlation functions can be computed in the same way. For instance, given $i_1 < i_2$ and $a_1, a_2 \in A$ the second-order correlation function is:

$$\mu_N(\eta : \eta_{i_1} = a_1, \eta_{i_2} = a_2) = \frac{\vec{y}^T M^{(i_1-1)} M^{a_1} M^{(i_2-i_1-1)} M^{a_2} M^{(N-i_2)} \vec{x}}{Z_N}.$$

When each site can only take two possible local states, namely $A = \{0, 1\}$, i.e. for the TASEP, the partition function reduces to

$$Z_N = \vec{y}^T (M^0 + M^1)^N \vec{x}$$

where the N-th power of the sum of the two matrices generates all possible strings of 0 and 1 of length N. The same principle can be applied to correlation functions, which can be obtained by inserting the corresponding matrices M^0 or M^1 at the specified sites within the product. Note that for the binary case, the first-order correlation function reduces to the average occupation number:

$$\langle \eta_i \rangle_N = \frac{\vec{y} M^{(i-1)} M^1 M^{(N-1)} \vec{x}}{Z_N}.$$

Moreover, in the TASEP case the relation between the matrices $M^1 M^0 = M^1 + M^0$ allow to show the independence of the stationary current from each site i and to compute it in terms

of the partition function:

$$J_N := \mu_N(\eta : \eta_i = 1, \eta_{i+1} = 0) = \frac{\vec{y}^T M^{(i-1)} M^1 M^0 M^{(N-i-1)} \vec{x}}{\vec{y}^T M^{(N)} \vec{x}} = \frac{\vec{y}^T M^{(N-1)} \vec{x}}{\vec{y}^T M^{(N)} \vec{x}} = \frac{Z_{N-1}}{Z_N}.$$

If the explicit form of M^0, M^1, $\vec{y}$ and $\vec{x}$ is suitable, the entries of any power of M admit a simple expression, then the density profile and higher-order correlation functions can be computed directly.
Moreover, in the specific case of the TASEP (and more generally, of the ASEP and SEP), the algebraic relations (2.3.3) satisfied by the matrices allow one to derive explicit expressions for the partition function, without relying on an explicit matrix representation [23, 51]. By the same method, making use of the algebraic relations, the density profile as well as higher-order correlation functions can also be obtained explicitly [23]. Alternatively, they can be computed using the generating function method [4].
For the open TASEP, an explicit closed form for the partition function can be obtained, which reveals the combinatorial nature of the model. With the normalization $\vec{y}^T \vec{x} = 1$ fixed above, the partition function can be written as:

$$Z_N = \sum_{p=1}^{N} \frac{p(2N-p-1)}{N!(N-p)!} \sum_{q=0}^{p} \left(\frac{1}{\alpha}\right)^q \left(\frac{1}{\beta}\right)^{p-q} \tag{2.5.1}$$

which in the special case $\alpha = \beta = 1$ reduces to

$$Z_N = \frac{1}{N+2} \binom{2N+2}{N+1} = C_{N+1} \tag{2.5.2}$$

where C_{N+1} is the $(N+1)$-st Catalan number.
Similarly to the TASEP, also for the SEP the partition function admits a compact closed form. Consider the SEP in the simplified setting with open boundary conditions where particles enter at site 1 with rate α and exit from site N with rate β. It can be shown [63] that the partition function Z_N can be expressed in terms of the Gamma function Γ as follows:

$$Z_N = \frac{\Gamma\left(\frac{1}{\alpha} + \frac{1}{\beta} + N\right)}{\Gamma\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)}.$$

In the case $\alpha = \beta = 1$ it becomes $Z_N = (N+1)!$.
The computation of the partition function for the ASEP with general q is more involved; we will not provide details here, but we stress that, using the matrix approach, it can be obtained explicitly in a compact form. The exact expressions for the partition function and the currents, for general q, α and β, can be found in the work [5].

Remark 2.4 (Combinatorial connections)**.** *The explicit expression of the partition function*

generates two classical sequences in enumerative combinatorics which emerge naturally in this non-equilibrium setting. Through the matrix product representation (2.3.9), in the case $\alpha = \beta = 1$, Z_N can be interpreted in terms of bicoloured Motzkin paths of length N, in bijection with Dyck paths of length $2(N+1)$. It follows that the partition function enumerates all such paths and coincides with the Catalan number C_{N+1}, consistent with the result given in (2.5.2). In the general case, (2.5.1) contains as a factor the ballot number:

$$B_{N,p} = \frac{p(2N-p-1)!}{N!(N-p)!}$$

which counts Dyck paths of length $2N$ with exactly p returns to the origin. Thus the partition function admits a natural interpretation as a weighted enumeration of Dyck paths, where returns from above and below carry weights $1/\alpha$ and $1/\beta$, respectively. These examples show how the matrix product representation naturally allows for a combinatorial interpretation, linking stationary configurations of exclusion processes to classical objects such as lattice paths. A broader overview of these connections is presented in [63].

In the specific case of the TASEP the partition function (2.5.1) can be rewritten in terms of the polynomial R_N defined by:

$$R_N(x) := \sum_{k=0}^{N} \frac{k}{2N-k} \binom{2N-k}{N} x^{k+1}$$

in the following form:

$$Z_N = \vec{y}^T M^N \vec{x} = \frac{R_N(\beta^{-1}) - R_N(\alpha^{-1})}{\beta^{-1} - \alpha^{-1}}.$$

We observe that the asymptotic analysis of the partition function as $N \to +\infty$, through the asymptotics of R_N, allows the investigation of the large-N behavior of the stationary current J_N and even for the limiting density profile.

2.5.1 Phase diagram of the TASEP

The TASEP exhibits three phases depending on the values of the boundary parameter α and β (Figure 2.2):

1. *Maximal current phase (MC)* $\alpha, \beta > \frac{1}{2}$. For these values of the boundary parameters, the stationary current and the density profile for large N are:

$$J_N \simeq \frac{1}{4}, \qquad \langle \eta_i \rangle_N \simeq \frac{1}{2}$$

where $i = Nx$ with $0 < x < 1$. In this phase, the injection and exit processes are sufficiently fast so that the boundaries cease to impose any constraint on the dynamics. The system established a stationary state with an average density of $\frac{1}{2}$, which ensures the most efficient transport through the bulk. Consequently, the stationary current reaches the maximal value and becomes independent of the boundary parameters.

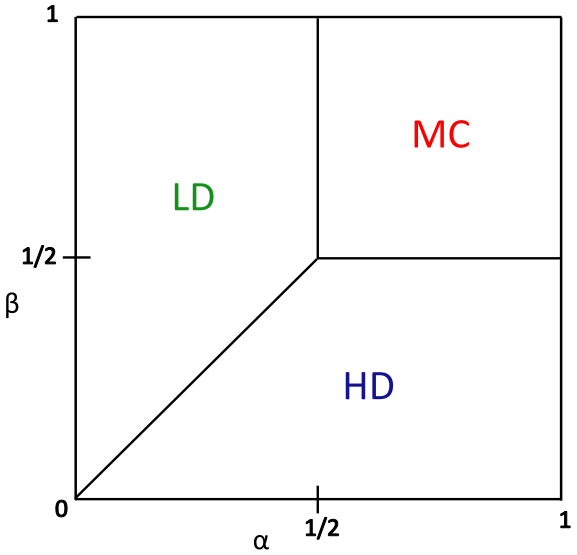

Figure 2.2: Phase diagram of the TASEP.

2. *High density phase (HD)* $\alpha > \beta, \beta < \frac{1}{2}$. In this parameters region:

$$J_N \simeq \beta(1-\beta), \qquad \langle \eta_i \rangle_N \simeq (1-\beta)$$

$i = Nx$, $0 < x < 1$. In this regime the exit rate at the right boundary becomes the limiting factor of the dynamics. Particles are injected faster than they can be removed leading to an accumulation within the system. This congestion propagates backward into the bulk, so that the entire system is characterized by a high average density. In this phase the stationary current is determined by the boundary parameter β indicating that the global transport capacity is restricted by the efficiency of the right boundary.

3. *Low density (LD)*, $\alpha < \beta$, $\alpha < \frac{1}{2}$. In this regime:

$$J_N \simeq \alpha(1-\alpha) \qquad \langle \eta_i \rangle_N \simeq \alpha$$

$i = Nx$, $0 < x < 1$. In the low-density phase, in contrast, is the injection rate the limiting factor of the dynamics. Particles are introduced more slowly than they can be transported through the bulk resulting in a stationary state with a low average density. In this regime the stationary current is constrained by the injection rate.

A more detailed analysis of the phase-diagram which includes the behavior along the boundaries between the three density regions, can be found in [4].

2.6 MPA representation of the two-species TASEP

One of the earliest extensions of the matrix product approach to the TASEP concerns the case with two particle classes. We recall that in this version of the totally asymmetric exclusion process both species attempt to jump to the right with rate 1. If a first–class particle finds a second–class particle to its right, they exchange positions with rate 1. Conversely, if a

second–class particle has a first–class particle on its right, it cannot move. Thus, from the perspective of first–class particles, second–class ones play the role of holes, while from the perspective of the holes, second–class particles behave as particles.
The first result in this framework was obtained in [22] and concerns the matrix product formulation for the stationary state of a system containing two species particle on a ring $\mathbb{Z}_N$. The state space is $A = \{0, 1, 2\}$ and, as for the TASEP, the matrices involved are infinite-dimensional, hence auxiliary space is $\mathbb{N}_0$. The variables $\eta_i = 0, 1, 2$ define, correspondingly, if the site is empty, if contains a first-class particle or a second-class one. The matrix approach proved that the stationary measure of the system on a periodic N-lattice, assuming that there is at least one second class particle ($n_2 > 0$), can be written in terms of matrices as:

$$\mu_N^{n_1,n_2}(\eta) = \frac{\mathrm{Tr}\left[\prod_{i=1}^N M^{\eta_i}\right]}{Z_N(n_1, n_2)} \mathbb{I}\Big(\mathcal{N}_1(\eta) = n_1\,, \mathcal{N}_2(\eta) = n_2\Big)\,, \tag{2.6.1}$$

where the normalization factor $Z_N(n_1, n_2)$ depends on the numbers n_1, n_2 of the first and second class particles. Note that the trace operation on matrix products encodes the lattice periodicity via its cyclic symmetry. In order for the measure $\mu_N^{n_1,n_2}$ to be stationary for the process, the matrices M^0 M^1 M^2 must satisfy the following conditions:

$$M^1 M^0 = M^1 + M^0, \qquad M^1 M^2 = M^2, \qquad M^2 M^0 = M^2.$$

The first equation coincides with that of the TASEP, whereas the additional relations involve the matrix M^2. In addition, it is required that the trace of the matrix product remain finite. In particular, if M^2 has only finitely nonzero entries on its diagonal - regardless of the specific forms of M^0 and M^1 - then the assumption of at least one second-class particle guarantees the finiteness of the trace operator. Indeed if M^2 has this specific form, any matrix product involving at least one occurrence of M^2 contributes only a finite number of nonzero terms to the diagonal of the product, ensuring that the trace is well defined.
A representation of matrices which satisfies the previous relations and the finiteness of the trace is (see [22]):

$$M^0 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots \\ 1 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad M^1 = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 1 & \ddots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.6.2}$$

$$M^2 = \vec{0}\,\vec{0}^T = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{2.6.3}$$

The matrix M^2 takes the form of the rank-one operator $M^2 = \vec{0}\ \vec{0}^T$, with $\vec{0}^T = (1, 0, 0, ...)$. This projector obeys $M^2 M^2 = M^2$ and due to its form, this representation shows easily an important property of the stationary state, (2.6.1) reduces to:

$$\mu_N^{n_1,n_2}(\eta) = \frac{1}{Z_N(n_1, n_2)} \prod_{j=1}^{n_2} \omega(B_j) \tag{2.6.4}$$

where B_j is a configuration of 1s and 0s between two 2-class particles:

$$\omega(B_j) = \vec{0}^T \left(\prod_{i=1}^{l} M^{\eta_i} \right) \vec{0}$$

where l is the length of the binary string B_j and M^{η_i} is M^0 or M^1 according to the occupation of the i-th element of the string. Hence, the stationary state factorizes with respect to the positions of the second–class particles. This factorization highlights the role of second-class particles as natural separators of independent blocks, with each block corresponding to a binary configuration of first-class particles and holes.

The representation of the stationary measure of the 2-TASEP [22], inspired Ferrari, Fontes and Kohayakawa in [37] to introduce a probabilistic construction of the measure, later refined by Angel [1] through a combinatorial characterization. Based on the Angel's work, Ferrari and Martin in [38] reinterpreted his approach within a queuing framework and generalized it to the case of N particles species. Further progress in this direction was achieved by Evans, Ferrari and Mallick in [34]. They showed how the trajectories of the queueing process are counted within the matrix product formulation. Moreover, they provided an explicit matrix formulation for the multispecies model.

Chapter 3

Mixtures, Markov Bridges and MPA

In the previous chapters we introduced the main features of stochastic particle systems that will serve as the main examples throughout this work, and discussed the Matrix Product Ansatz (MPA), focusing on its algebraic properties and providing a brief account of its combinatorial structure.

These elements offer the background for the developments of this chapter, where we shift our attention to the probabilistic structure of the MPA. As we have seen, the Matrix Product Ansatz provides an algebraic description of stationary measures for a large class of interacting particle systems; here we explore the probabilistic structures that such measures may exhibit, focusing on their characterization in terms of mixtures of non-homogeneous product measures. The following sections present results from [43].

Our main result is that a probability measure can be written by the MPA using matrices with non negative entries if and only if the measure has a special canonical probabilistic structure. This special structure is a *mixture of non homogeneous product measures* where the parameters that determine the marginals of the product are distributed according to a *Markov bridge*.

This probabilistic characterization provides a unified viewpoint and allows one to interpret matrix-product measures as mixtures governed by hidden Markovian variables.

As we shall see, this result is obtained by enlarging the state space using a special coupling and then applying a generalized Doob transform that associates a stochastic matrix to a non-negative one. This transformation is based on a spectral construction. We prove our general statement in the case of finite alphabets, when the spectral structure is simpler and we do not need to discuss special and detailed assumptions. However, the basic mechanism holds also in the infinite setting; indeed, we illustrate the construction by several examples of particle systems in which the matrices involved in the MPA are infinite dimensional. In particular we show that some remarkable probabilistic representations of the invariant measures of non reversible particle systems, such as the two lines ensemble for the boundary driven TASEP [2, 6], or the queue representation of a multiclass TASEP [1, 38, 34], can be obtained from the corresponding MPA by using our transformation.

Recently, representations in terms of mixtures of inhomogeneous product measures of the invariant measures of boundary driven generalized KMP and zero range models have been obtained in [11, 18] and in [10, 46, 54]. For the zero range case, considering just the simplest model, we show (this is one of the examples we discuss) that this representation is a special case of our construction corresponding to a Markov process and marginal laws of the product measures having special features. The general case is more involved; more comments on it and on KMP process are provided in the remark 3.12.

Our enlargement of the state space and the consequent Markov characterization are similar to the construction, in a stationary setting, of the algebraic and manifestly positive algebraic

measures in [35]. In [35] a characterization of such measures is obtained as a function of Markov processes for a sufficiently enlarged state space; see remark 3.9 for more details. Our mixture representation does not require an enlargement of the state space and follows from a further step using the fact that the Markov processes involved are of a special type.

The mixture representation introduced here provides a suitable framework for studying the corresponding large deviation asymptotics, an aspect that will be discussed in the next chapter.

Finally, we note that it would be also interesting to compare our construction with recent results for quantum models [57].

The chapter is organized as follows.

In the first section, we introduce the probabilistic framework and the main objects of interest - mixtures of product measures and Markov bridges - providing the definitions and notation that will be used.

In the second section, we prove the equivalence between the Matrix Product Ansatz and mixtures of inhomogeneous product measures whose parameters are distributed as a Markov bridge; in particular, we state and prove the main characterization theorem.

The following sections focus on applications of the construction to specific models: we discuss the boundary-driven TASEP and ASEP and the two-class particles system on a ring.

Finally, we move our attention to the harmonic model, where we start directly from a known mixture representation and show how a generalized *matrix*-product structure can be defined.

3.1 Probabilistic framework

In the previous chapter we introduced the Matrix Product Ansatz in its abstract formulation, as a family of probability measures on words of length N constructed through products of non-negative matrices and boundary vectors.

Throughout this chapter we will use the same notation introduced in Section 2.2; we recall that $M^{(n)}$ denote the n-th power of a matrix.

As previously stated, we consider two finite or countable sets A and B: the set A is the state space where the variables take values while B is an auxiliary space. For any $a \in A$ we have a $B \times B$ matrix M^a having non-negative entries, and we also consider two column vectors $\vec{x} := (x_b)_{b\in B}, \vec{y} := (y_b)_{b\in B} \in \mathbb{R}_+^B$ having non-negative coordinates.

For convenience, we recall that such measures can be written in terms of product of matrices as:

$$\mu_N(\eta) = \frac{\vec{y}^T \left(\prod_{i=1}^N M^{\eta_i} \right) \vec{x}}{Z_N}, \qquad \eta \in A^N, \tag{3.1.1}$$

where Z_N is the normalization factor.

This algebraic structure will serve as the starting point for the probabilistic representation developed below; we first introduce the notion of a Markov bridge and of mixtures of product measures.

Definition 3.1 (Markov bridge)**.** *Consider a $B \times B$ stochastic matrix P and two column vectors $f, g \in \mathbb{R}_+^B$. For notational convenience, given a $B \times B$ stochastic matrix P and $\zeta \in$*

B^{N+1}, *we denote by*

$$\mathbb{P}_P(\zeta) := \prod_{i=1}^{N} P_{\zeta_i,\zeta_{i+1}}\,,$$

i.e. the probability of the path ζ *with fixed initial condition* ζ_1 *according to the Markov law induced by* P *(a similar notation is also used for other state spaces). We define a Markov bridge to be the family of elements* $\rho_{N+1} \in \mathcal{M}^1(B^{N+1})$ *by*

$$\rho_{N+1}(\zeta) := \frac{g_{\zeta_1}\mathbb{P}_P(\zeta) f_{\zeta_{N+1}}}{g^T P^{(N)} f}\,, \qquad \zeta \in B^{N+1}\,. \tag{3.1.2}$$

The notation is the same as in the formula (3.1.1) and again we do not write explicitly the dependence of the measure on the parameters f, g, P. The motivation for the presence of the $N+1$ instead of N will be clarified later. The term in the denominator is the normalization factor. Again, in the finite case the measure (3.1.2) is always well-defined, while a summability condition is necessary in the infinite case.

The measure (3.1.2) is a non-homogeneous Markov measure, since, by a direct computation, we have

$$\rho_{N+1}(\zeta_{k+1}|\zeta_1^k) = \frac{P_{\zeta_k,\zeta_{k+1}} \sum_{\xi_{N+1}} P^{(N-k)}_{\zeta_{k+1},\xi_{N+1}} f_{\xi_{N+1}}}{\sum_{\xi_{N+1}} P^{(N+1-k)}_{\zeta_k,\xi_{N+1}} f_{\xi_{N+1}}}\,, \qquad \zeta \in B^{N+1}\,. \tag{3.1.3}$$

The Markovianity follows from the fact that the right hand side in the above formula does not depend on ζ_1^{k-1}, i.e. $\rho_{N+1}(\zeta_{k+1}|\zeta_1^k) = \rho_{N+1}(\zeta_{k+1}|\zeta_k)$.

In the case that $f_\cdot = \delta_{\cdot,b'}$ and $g_\cdot = \delta_{\cdot,b}$, the Markov bridge (3.1.2) is pinned at b at time 1 and at b' at time $N+1$. The resulting process coincides with the law of the Markov chain with transition probability P, conditioned to start at b and to be in b' at time $N+1$. We call $\rho^{b,b'}_{N+1}$ this measure that is given by

$$\rho^{b,b'}_{N+1}(\zeta) = \begin{cases} \frac{\mathbb{P}_p(\zeta)}{P^{(N)}_{b,b'}}\,, & \text{if } \zeta_1 = b\,, \zeta_{N+1} = b'\,, \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3.1.4}$$

The above measure also depends, of course, on the stochastic matrix P but we do not write this dependence explicitly.

In the general case, the joint law of the pair (ζ_1, ζ_{N+1}) is given by $m \in \mathcal{M}^1(B^2)$ which is defined by

$$m(\zeta_1, \zeta_{N+1}) = \frac{g_{\zeta_1} P^{(N)}_{\zeta_1,\zeta_{N+1}} f_{\zeta_{N+1}}}{g^T P^{(N)} f} \tag{3.1.5}$$

and we then have the convex decomposition

$$\rho_{N+1}(\zeta) = \sum_{b,b'} m(b,b')\rho^{b,b'}_{N+1}(\zeta)\,.$$

This means that a sample of the measure (3.1.2) can be generated by choosing initial and final values according to m in (3.1.5), and then generating a pinned Markov bridge with those

initial and final values.

We now consider a map $\hat{p} : \{1, \dots, N\} \times B^{N+1} \to \mathcal{M}^1(A)$, which associates to the pair $i \in \{1, \dots, N+1\}$ and $\zeta \in B^{N+1}$ the probability measure $\hat{p}^i_\zeta(\cdot) \in \mathcal{M}^1(A)$.

Definition 3.2 (Mixtures)**.** *We say that a probability measure $\nu_N \in \mathcal{M}^1(A^N)$ is a mixture of product measures with marginal measures determined by $\hat{p}$ and parameter distribution $\rho_{N+1} \in \mathcal{M}^1(B^{N+1})$ if*

$$\nu_N(\eta) = \sum_{\zeta \in B^{N+1}} \rho_{N+1}(\zeta) \left[\prod_{i=1}^{N} \hat{p}^i_\zeta(\eta_i) \right], \qquad \eta \in A^N . \tag{3.1.6}$$

We denote ρ_{N+1} the parameter distribution, using the same symbol as for a Markov bridge for simplicity of notation. In Definition 3.2 the measure $\rho_{N+1} \in \mathcal{M}^1(B^{N+1})$ is arbitrary, but in the following we will consider the special cases in which it is exactly a Markov bridge.

In principle, we could consider parameter distributions on longer sequences over the alphabet B, but the length $N+1$ is natural since we will study the special cases in which

$$\hat{p}^i_\zeta(\cdot) = p_{\zeta_i, \zeta_{i+1}}(\cdot) , \tag{3.1.7}$$

for a suitable map $p : B^2 \to \mathcal{M}^1(A)$ that associates to the pair $(b, b') \in B^2$ the probability measure $p_{b,b'}(\cdot) \in \mathcal{M}^1(A)$.

Comments on the definition 3.2 and the related constructions are postponed to the following sections, and we move directly to the main result.

3.2 The equivalence: Mixtures and MPA

In this section, we state and prove our main result in the cases of finite alphabets where we can avoid any technical assumption; in the following sections we illustrate the result with examples that are not finite, since the basic construction works also in that case modulo some technical assumptions, that are satisfied in the examples we consider.

Before stating the theorem establishing the equivalence between MPA measures and mixtures, we first briefly recall the classical Perron-Frobenius theorem which will play a central role in the subsequent construction. We refer to [50] for a complete discussion, also in the infinite case.

Theorem 3.3 (Perron-Frobenius)**.** *Consider a finite $|B| \times |B|$ irreducible and aperiodic non-negative matrix M, then there exists a positive maximal eigenvalue $\lambda > 0$ such that all the remaining eigenvalues λ_i are such that $|\lambda_i| < \lambda$. The unique eigenvector $\vec{e} = (e(b))_{b \in B}$ corresponding to the maximal eigenvalue can be fixed in such a way that $e(b) > 0$ for any $b \in B$.*

Remark 3.4. *When the matrix M is irreducible but not aperiodic, theorem 3.3 has to be modified and there exists still a maximal real eigenvalue λ and a corresponding strictly positive eigenvector $\vec{e}$, we have, however, for the other eigenvalues not a strict inequality but instead $|\lambda_i| \leq \lambda$. If the period of the irreducible chain is p then there are exactly p eigenvalues of maximal modulus λ, see [50] for more details.*

As a consequence of the Perron-Frobenius theorem, given a finite irreducible matrix M, by a simple transformation, sometimes called a generalized Doob transform or ground state

transform, we can associate to M a stochastic matrix given by

$$P = \lambda^{-1}\mathcal{E}^{-1}M\mathcal{E}\,, \tag{3.2.1}$$

where $\mathcal{E}$ is the diagonal matrix having elements $\mathcal{E}_{i,i} = e_i$. The entries of the matrix P are then given by $P_{b,b'} = \frac{M_{b,b'}e_{b'}}{\lambda e_b}$. More precisely, we have the following lemma:

Lemma 3.5. *Let M be a non-negative irreducible and aperiodic matrix and let λ, $\vec{e}$ be the corresponding maximal eigenvalue and positive right eigenvector; then the matrix*

$$P_{b,b'} = \frac{1}{\lambda}e(b)^{-1}M_{b,b'}e(b') \qquad b,b' \in B\,, \tag{3.2.2}$$

is a stochastic matrix and we have that

$$P^{(k)}_{b,b'} = \frac{1}{\lambda^k}e(b)^{-1}\left(M^{(k)}\right)_{b,b'}e(b') \qquad k \in \mathbb{N},. \tag{3.2.3}$$

If M is irreducible and aperiodic, the matrix S is also irreducible and aperiodic.

Proof. The proof of (3.2.2) follows by a direct computation. We need just to check that the sum along each row is one

$$\sum_{b'\in B} P_{b,b'} = \frac{1}{\lambda}e(b)^{-1}\sum_{b'\in B} M_{b,b'}e(b') = \frac{1}{\lambda}e(b)^{-1}\lambda e(b) = 1\,.$$

Formula (3.2.3) follows directly by the representation (3.2.1). Irreducibility and aperiodicity follow directly. □

This construction is also related to Hammersley–Clifford theorem (see, e.g., [14]).

Building on the previous construction, we are now ready to state the main theorem which provides a complete characterization of the correspondence between MPA measures and mixtures.

Theorem 3.6. *Consider $|A|, |B| < +\infty$; a probability measure $\mu_N \in \mathcal{M}^1(A^N)$ is of MPA type (3.1.1) with $M := \sum_{a\in A} M^a$ an irreducible non negative matrix, if and only if μ_N is a mixture of product measures 3.2 having the distribution of the parameters ρ_{N+1} given by a Markov bridge (3.1.2) and marginal measures of the form (3.1.7). Moreover, the Markov bridge has transition probability given by the stochastic matrix P defined in (3.2.1); the vectors f, g are related to x, y (those of the MPA) by the relations $f = \mathcal{E}^{-1}x$, $g = \mathcal{E}y$ and the marginal distributions are of the form (3.1.7) with*

$$p_{b,b'}(a) = \frac{M^a_{b,b'}}{M_{b,b'}}\,, \qquad b,b' \in B\,,\, a \in A\,. \tag{3.2.4}$$

Finally, we have the relation $g^T P^N f = \frac{Z_N}{\lambda^N}$.

The proof is based on a special coupling construction and on the transform (3.2.1). In the proof we will introduce a coupling measure and discuss some of its properties.

Proof. First, we start with a measure of MPA type (3.1.1) and show that is a mixture. Consider a measure μ_N of the form (3.1.1). We enlarge the state space and construct a probability measure $C_N \in \mathcal{M}^1(A^N \times B^{N+1})$ defined by

$$C_N(\eta,\zeta) := \frac{y_{\zeta_1}\left(\prod_{i=1}^N M^{\eta_i}_{\zeta_i,\zeta_{i+1}}\right) x_{\zeta_{N+1}}}{Z_N}\,, \qquad \eta \in A^N\,,\ \zeta \in B^{N+1}\,. \tag{3.2.5}$$

The coupling measure C_N has the following remarkable properties:

1. The η marginal of C_N is μ_N in (3.1.1), i.e. $\mu_N(\eta) = \sum_{\zeta\in B^{N+1}} C_N(\eta,\zeta)$.
2. The measure C_N is a Markov bridge with state space $A \times B$.
3. The ζ marginal of C_N, $\rho_{N+1}(\zeta) = \sum_{\eta\in A^N} C_N(\zeta,\eta)$, is a Markov bridge with the features described in the statement of the theorem.
4. Conditioned on the ζ variables, the η variables are independent, and moreover

$$C_N(\eta|\zeta) = \prod_{i=1}^N p_{\zeta_i,\zeta_{i+1}}(\eta_i)\,, \tag{3.2.6}$$

where the $p_{b,b'}(\cdot)$ are defined in (3.2.4).

Item (1) follows directly from the definition, since

$$\sum_{\zeta\in B^{N+1}} y_{\zeta_1}\left(\prod_{i=1}^N M^{\eta_i}_{\zeta_i,\zeta_{i+1}}\right) x_{\zeta_{N+1}} = y\prod_{i=1}^N M^{\eta_i} x\,.$$

Item (2) will not be used in this chapter but it will be central in the next one. We skip some details here, related to the irreducibility of the positive matrix we are going to introduce; these details will be explained in the next chapter (see Section 4.4).

We introduce the $(A \times B) \times (A \times B)$ matrix $\mathfrak{M}_{(a,b),(a',b')} := M^a_{b,b'}$ which has non negative entries. As will be shown in Sec. 4.4 if M is irreducible then T is also irreducible. By (3.2.1) we can construct the $(A \times B) \times (A \times B)$ stochastic matrix $\mathfrak{S} := \Lambda^{-1}\hat{\mathcal{E}}^{-1}\mathfrak{M}\hat{\mathcal{E}}$ where Λ is the Perron eigenvalue of $\mathfrak{M}$ and $\hat{\mathcal{E}}$ is the diagonal matrix having elements $\hat{\mathcal{E}}_{(a,b),(a,b)} = \varepsilon_{a,b}$, where ε is the positive eigenvector associated with the eigenvalue Λ. We introduce the vectors $\hat{x}, \hat{y} \in \mathbb{R}_+^{A\times B}$ defined by $\hat{x}_{a,b} := x_b$ and $\hat{y}_{a,b\cdot} = y_b$. We can then write

$$C_N(\eta,\zeta) = \frac{\hat{y}_{\eta_1,\zeta_1}\prod_{i=1}^N \mathfrak{M}_{(\eta_i,\zeta_i),(\eta_{i+1},\zeta_{i+1})}\hat{x}_{\eta_{N+1},\zeta_{N+1}}}{Z_N}\,,$$

where we observe that, differently from (2.2.1), only one matrix appears in the formula, namely $\mathfrak{M}$. We also stress that, although the variable η_{N+1} appears in the right-hand side of the formula, it is not defined and does not appear on the left-hand side; however, due to the special form of $\mathfrak{M}$ and $\hat{x}$, the right-hand side is also independent of η_{N+1}. By a telescoping

argument we can write the above formula as

$$C_N(\eta,\zeta) = \frac{\Lambda^N \hat{y}_{\eta_1,\zeta_1} \hat{\mathcal{E}}_{(\eta_1,\zeta_1),(\eta_1,\zeta_1)} \mathbb{P}_{\mathfrak{S}}(\eta,\zeta) \hat{\mathcal{E}}^{-1}_{(\eta_{N+1},\zeta_{N+1}),(\eta_{N+1},\zeta_{N+1})} \hat{x}_{\eta_{N+1},\zeta_{N+1}}}{Z_N},$$

that has exactly the form (3.1.2) but on the state space $A \times B$ with the transition matrix P given by $\mathfrak{S}$ and the vectors g, f given by $\hat{\mathcal{E}}\hat{y}$, $\hat{\mathcal{E}}^{-1}\hat{x}$.

Item (3) is obtained by

$$\sum_{\eta} y_{\zeta_1} \left(\prod_{i=1}^{N} M^{\eta_i}_{\zeta_i,\zeta_{i+1}} \right) x_{\zeta_{N+1}} = y_{\zeta_1} \left(\prod_{i=1}^{N} M_{\zeta_i,\zeta_{i+1}} \right) x_{\zeta_{N+1}} . \tag{3.2.7}$$

Using now formula (3.2.1), and by the same telescoping argument as before, we obtain

$$\rho_{N+1}(\zeta) = \sum_{\eta} C_N(\eta,\zeta) \tag{3.2.8}$$

$$= \frac{y_{\zeta_1} \left(\prod_{i=1}^{N} M_{\zeta_i,\zeta_{i+1}} \right) x_{\zeta_{N+1}}}{Z_N} \tag{3.2.9}$$

$$= \frac{\lambda^N \left(\mathcal{E} y\right)_{\zeta_1} \mathbb{P}_P(\zeta) \left(\mathcal{E}^{-1} x\right)_{\zeta_{N+1}}}{Z_N} , \tag{3.2.10}$$

which is of the form (3.1.2), and we deduce the validity of the statement of item (3).

Item (4) is obtained by a direct computation using (3.2.5) and (3.2.7)

$$C_N(\eta|\zeta) = \frac{C_N(\eta,\zeta)}{\rho_{N+1}(\zeta)} = \prod_{i=1}^{N} \frac{M^{\eta_i}_{\zeta_i,\zeta_{i+1}}}{M_{\zeta_i,\zeta_{i+1}}} = \prod_{i=1}^{N} p_{\zeta_i,\zeta_{i+1}}(\eta_i) .$$

The relation $g^T P^{(N)} f = \frac{Z_N}{\lambda^N}$ follows directly from (2.2.2) and (3.2.1).

Conversely, consider now a measure written in terms of mixtures as in Definition 3.2 with the parameter distribution ρ_{N+1} given a Markov bridge (3.1.2) and the marginal distributions of the form (3.1.7). We introduce the vectors $x = f$ and $y = g$ and the matrices M^a whose elements are defined by $M^a_{b,b'} = P_{b,b'} p_{b,b'}(a)$. We obtain in this way a representation of the mixture by the MPA; with this choice, the matrix M is stochastic and coincides with P.

□

We now conclude this section with few remarks that highlight the generality, implications, and some technical aspects of Theorem 3.6.

Remark 3.7 (Non-uniqueness)**.** *The mixture representation* (3.1.6)*, without imposing any restrictions on the measures involved, is a very general statement and it can be shown under general assumptions that such a representation of a probability measure μ_N exists and is moreover not unique. Consider, for example $A = B = \{0,1\}$ and the marginal distributions given by $\hat{p}^i_\zeta(\cdot) = \mathcal{B}_{\zeta_i}(\cdot)$ where $\mathcal{B}_p(\cdot)$ denotes the Bernoulli distribution of parameter p. We have that*

any $\mu_N \in \mathcal{M}^1\left(\{0,1\}^N\right)$ *always admits a representation of the form* (3.1.6) *as*

$$\mu_N(\eta) = \sum_{\zeta \in \{0,1\}^{N+1}} \rho_{N+1}(\zeta) \left[\prod_{i=1}^{N} \mathcal{B}_{\zeta_i}(\eta_i)\right] ,$$

provided that $\rho_{N+1}(\zeta_1^N) = \mu_N(\zeta_1^N)$.

To illustrate the absence of uniqueness, let us consider the following simple case. Consider the product measure $\mu_N \in \mathcal{M}^1\left(\{0,1\}^N\right)$ *given by* $\mu_N = \prod_{i=1}^N \mathcal{B}_{q_i}$ *for an arbitrary collection of parameters* $q_i \in [0,1]$. *Let* $\rho_{N+1} \in \mathcal{M}^1\left(B^{N+1}\right)$ *be also a product measure with* $i-$*marginal given by* $\rho^{[i]} \in \mathcal{M}^1(B)$. *Let* $m_i : B \to [0,1]$ *be functions with mean values* $\sum_{b\in B} \rho^{[i]}(b) m_i(b) = q_i$ *and consider the marginal distributions* $\hat{p}^i_\zeta(\cdot) = \mathcal{B}_{m_i(\zeta_i)}(\cdot)$. *We the have, under this single assumption, that*

$$\prod_{i=1}^{N} \mathcal{B}_{q_i}(\eta_i) = \sum_{\zeta \in B^{N+1}} \rho_{N+1}(\zeta) \left[\prod_{i=1}^{N} \mathcal{B}_{m_i(\zeta_i)}(\eta_i)\right] .$$

Since the condition is only on the mean values of the marginals of ρ_{N+1} *it is easy to see that uniqueness does not hold.*

Remark 3.8 (Consequences of the mixture representation)**.** *The importance and power of the representation as a mixture* (3.1.6) *for measures of MPA type lies in the fact that the law* ρ_{N+1} *of the parameters is Markovian and the marginal distributions* $\hat{p}^i_\zeta(\cdot)$ *are local (i.e. depends just on* ζ_i, ζ_{i+1}*) and translationally covariant. In particular, mean values, covariances and other properties of the measure* μ_N *are strictly related to those of the underlying Markov bridge. Consider for example correlations. Let us define* $V_{\zeta_i,\zeta_{i+1}} := \sum_{a\in A} p_{\zeta_i,\zeta_{i+1}}(a)a$. *By a direct computation we have*

$$\mathbb{E}_{\mu_N}\left(\eta_i \eta_j\right) = \mathbb{E}_{\rho_{N+1}}\left(V_{\zeta_i,\zeta_{i+1}} V_{\zeta_j,\zeta_{j+1}}\right) ,$$

i.e. the correlations for the measure μ_N *coincide with the correlations of the local function* V *for the Markov measure* ρ_{N+1}. *Other features of the measure* μ_N *can be deduced from those of* ρ_{N+1} *as for example large deviations [45] and FKG type inequalities [11].*

Remark 3.9. *Item* (2) *in the proof of Theorem 3.6 is a statement similar to the result in [35]. In an infinite stationary framework, the authors of [35] introduce the class of algebraic measures and the class of manifestly positive algebraic measures, for which probabilities of cylinder sets can be computed, like in* (2.2.1). *They prove an equivalence between the family of manifestly positive algebraic measures and the family of functions of Markov processes. A process* $(Y_n)_{n\in\mathbb{N}}$ *is called a function of a Markov process if there exists a Markov process* $(X_n)_{n\in\mathbb{N}}$ *and a function* Φ *such that* $Y_n = \Phi(X_n)$. *For simplicity, the authors of [35] also consider finite dimensional cases and obtain bounds on the size of the configuration space of the Markov process. In the case that the manifestly positive algebraic measure has state space* A *and the probabilities are computed using matrices of size* $|B|$ *then the process can be obtained as a function of a Markov process whose configuration space is bounded by* $\left(\max\{|A|,|B|\}\right)^4$.

By item (2) *of Theorem 3.6 we can deduce that the process* $(\eta_i)_{i=1}^N$ *is a function of the*

inhomogeneous Markov process $\left(\eta_i, \zeta_i\right)_{i=1}^{N}$ *via the function* $\Phi(a,b) = a$. *In this case the size of the configuration space of the Markov process is given by* $|A||B|$; *however, by using the remaining statements of Theorem 3.6 we can instead obtain the mixture representation in which several features of the process* $(\eta_i)_{i=1}^{N}$ *can be deduced from those of the Markov process with a state space whose cardinality is simply* $|B|$. *This fact greatly simplifies the probabilistic description.*

Remark 3.10 (On the uniqueness of Doob transform)**.** *In the case of finite-dimensional matrices, i.e.* $|B| < +\infty$, *the Perron eigenvalue coincides with the spectral radius of the matrix, and the matrix admits a unique positive eigenvector. As a consequence, the transition matrix associated with a non-negative matrix by* (3.2.1) *is uniquely determined. In the case of infinite-dimensional matrices the Perron value does not, in general, coincide with the spectral radius and has a different definition; see, for example, [50]. In this setting, there may exist several positive eigenvalues larger than the Perron value, each admitting a positive eigenvector. Moreover, the operator may admit no eigenvectors. There is uniqueness when the stochastic matrix constructed through* (3.2.1), *using the spectral radius as eigenvalue, is recurrent. We do not go into details, and refer, for example, to [50] for some statements. In the infinite-dimensional case, one may therefore obtain more than one stochastic matrix associated with* M *via* (3.2.1). *This can be related to Doob h-transforms and to the existence of positive non-constant harmonic functions. Suppose, for example, that* P *is constructed from* M *through* (3.2.1) *and that* h *is a non-constant harmonic function for* P *(i.e. satisfies* $Ph = h$*). Then the matrix* P^h, *defined by* $P^h_{b,b'} = h_b^{-1} P_{b,b'} h_{b'}$, *is related to* M *by a generalized Doob transform of the form* (3.2.1). *By the same telescopic arguments used in the proof of Theorem 3.6, the families of Markov bridges constructed using* P *and* P^h *coincide.*

In the next sections, we illustrate the construction by applying it to several models of interacting particle systems.
As we mentioned previously, the same constructions can be applied to examples with countable alphabets; they are the same of the finite case, one only needs to verify summability conditions, which are indeed satisfied in the examples considered.

Before turning to specific models, we discuss a technical result concerning the spectral properties of an infinite tridiagonal matrix. This computation will be useful in the analysis of the model discussed below.

Eigenvalues and eigenvectors for a class of infinite matrices

We consider the eigenvalue problem for the infinite-dimensional matrices, having rows and columns labeled by $\mathbb{N}_0$, of the form:

$$A = \begin{bmatrix} \alpha & \beta_1 & 0 & \cdots & 0 & \cdots & \cdots \\ \beta_2 & \alpha & \beta_1 & 0 & \cdots & 0 & \cdots \\ 0 & \beta_2 & \alpha & \beta_1 & 0 & \cdots & \\ 0 & 0 & \beta_2 & \alpha & \beta_1 & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \tag{3.2.11}$$

that is

$$A_{ij} = \begin{cases} \beta_1 \delta_{i,j+1} \\ \alpha \delta_{i,j} \\ \beta_2 \delta_{i-1,j} \end{cases} \qquad i,j = 0,1,2,...$$

with $\alpha, \beta_1, \beta_2 > 0$. The spectral theory in the case of infinite matrices is more complex and we illustrate just a few computations. We refer, for example, to [50] for more details.

The Perron value for the matrix (3.2.11) is defined as

$$\lim_{n\to+\infty} \sqrt[n]{(A^n)_{i,j}}\,, \tag{3.2.12}$$

and the limit does not depend on i, j. These matrices have a continuum of positive eigenvalues larger than the Perron value, for which the corresponding eigenvectors are positive. We have

Lemma 3.11. *The Perron value of the matrix (3.2.11) is given by* $\alpha + 2\sqrt{\beta_1\beta_2}$ *and the corresponding column eigenvector is given by* $\vec{e} = \left((n+1) \left(\frac{\beta_2}{\beta_1}\right)^{\frac{n+1}{2}} \right)_{n\in\mathbb{N}_0}$.

Proof. We give only an outline of the argument, since the computation is equivalent to the classic combinatorial computation of the partition function based on bicoloured Motzkin and Dyck paths [63].

We consider a discrete time random walk on $\mathbb{N}_0$, that starts at zero and performs jumps of 0 or ± 1. We assign the weight β_1 to jumps of $+1$, the weight β_2 to jumps of -1 and the weight α to jumps of 0. The weights of paths of even length that start and end at 0 is given by

$$(A^{2n})_{0,0} = \sum_{k=0}^{n} \mathcal{C}_{n-k} \binom{2n}{2k} \alpha^{2k} \beta_1^{n-k} \beta_2^{n-k}$$

where $\mathcal{C}_n$ is the n-th Catalan number. By Stirling formula and classic asymptotic estimates, we deduce the Perron value. The eigenvector can be verified by a direct computation. □

3.3 Boundary-driven TASEP

We start by considering the boundary-driven totally asymmetric exclusion process (TASEP); the model and its matrix product formulation have already been introduced in Sections 1.1 and 2.3, respectively; here we only recall the features relevant to the present discussion.
As shown in Section 2.3, the stationary measure of the open TASEP with boundary parameters α and β admits a matrix-product representation of the form (3.1.1) with $A = \{0,1\}$, when the matrices M^0,M^1, and boundary vectors $\vec{x}$, $\vec{y}$ which satisfy:

$$\begin{cases} M^1 M^0 = M^1 + M^0\,, \\ M^1 x = \frac{1}{\beta}\vec{x}\,, \\ y^T M^0 = \frac{1}{\alpha} y^T\,. \end{cases} \tag{3.3.1}$$

As explained in the previous chapter, equations (3.3.1) do not identify uniquely the matrices $(M^a)_{a=0,1}$ and there are several different solutions of (3.3.1) for different values of the pa-

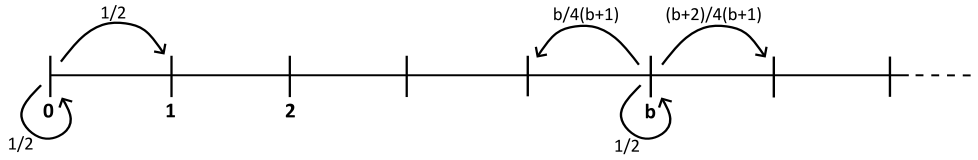

Figure 3.1: The transition graph associated to the matrix P in (3.3.4).

rameters. In the following, we discuss two explicit representations and interpret them in the probabilistic framework introduced above, by identifying the stochastic matrix, the Markov bridge, and the corresponding mixture representation.

3.3.1 The case $\alpha + \beta > 1$

In this regime, the algebraic relations (3.3.1) can be satisfied by taking $B = \mathbb{N} \cup \{0\}$ and using the matrices and the vectors described in (2.3.9).
The corresponding matrix $M = M^0 + M^1$ has entries:

$$M_{b,b'} = \begin{cases} 2\delta_{b,b'} + \delta_{b,b'+1} + \delta_{b,b'-1}\,, & b \geq 1\,, \\ 2\delta_{b,b'} + \delta_{b,b'+1}\,, & b = 0\,, \end{cases} \tag{3.3.2}$$

and the normalization factor (2.2.2) can be easily computed and is finite so that the measure (3.1.1) is well defined.

The eigenvalue problem for an eigenvalue λ with corresponding eigenvector $(e_b)_{b\in\mathbb{N}_0}$ is given by

$$\begin{cases} 2e_0 + e_1 = \lambda e_0\,, \\ e_{i-1} + 2e_i + e_{i+1} = \lambda e_i \quad i \geq 1\,, \end{cases} \tag{3.3.3}$$

This is a special case discussed in Lemma 3.11 with $\alpha = 2$, $\beta_1 = \beta_2 = 1$. It follows that the spectral radius of the matrix M is $\lambda = 4$ and the eigenvector associated is $\vec{e} = (b+1)_{b\in\mathbb{N}_0}$. The stochastic matrix (3.2.1) can therefore be explicitly computed and it is given by

$$P_{b,b'} = \begin{cases} \frac{1}{2} & b' = b \\ \frac{(b+2)}{4(b+1)} & b' = b+1 \\ \frac{b}{4(b+1)} & b' = b-1\,, \end{cases} \tag{3.3.4}$$

and its transition graph is represented in Figure 3.1 We have also

$$\begin{cases} g_b = y_b(b+1) = \frac{(1-\alpha)^b(b+1)}{\alpha^b} & b \in \mathbb{N}_0\,, \\ f_b = \frac{x_b}{b+1} = \frac{(1-\beta)^b}{\beta^b(b+1)} & b \in \mathbb{N}_0\,. \end{cases} \tag{3.3.5}$$

Finally, since $A = \{0,1\}$ and recalling that $\mathcal{B}_q \in \mathcal{M}^1(\{0,1\})$ is the Bernoulli measure of

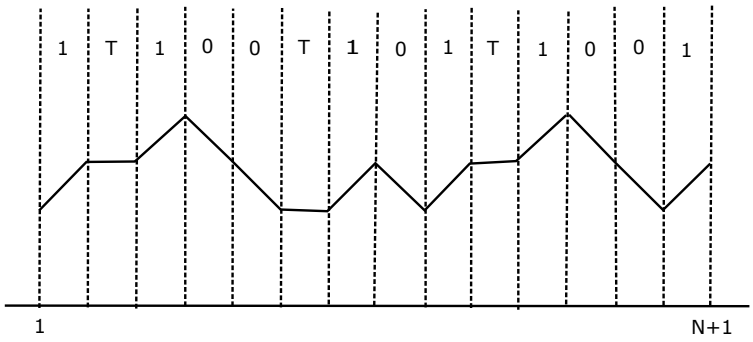

Figure 3.2: A trajectory ζ of a Markov bridge for a random walk on $\mathbb{N}_0$ with jumps $0, \pm 1$. Given the trajectory, the configuration η is completely determined on the intervals where the slope is ± 1 while we need to toss independent fair coins (T=Toss) in correspondence of horizontal steps. The trajectory of ζ may also assume the value zero.

parameter q, using (3.2.4) we have

$$p_{b,b'} = \begin{cases} \mathcal{B}_1 & \text{if } b' = b+1\,, \\ \mathcal{B}_0 & \text{if } b' = b-1\,, \\ \mathcal{B}_{\frac{1}{2}} & \text{if } b' = b\,. \end{cases} \tag{3.3.6}$$

This completes the probabilistic characterization of the invariant measure of the boundary driven TASEP in the regime considered in this section. The Markov bridge ρ_{N+1} for the variables ζ is associated with the transition matrix (3.3.4) and the vectors (3.3.5), while the marginal laws of the mixtures are given by (3.3.6). The construction of a sample of η distributed according to the invariant measure, given a sample of ζ is illustrated in Figure 3.2.

As discussed in the Section 1.1, several combinatorial representations of the stationary measure are known in the literature [62, 9, 30], including a recent representation obtained in terms of *two-lines ensemble* [6, 2]. We now show that the mixture representation obtained above is related and equivalent to this representation of two lines.

Comparison with the two lines ensemble. The *two-lines ensemble* representation derived in [6] on the approach of [2] is obtained using a pair of right-up independent random walks $s^1 = (s^1_i)_{i=1}^{N+1}$, $s^2 = (s^2_i)_{i=1}^{N+1}$. Both representations describe the same invariant measure and must therefore be equivalent in this sense. However, the two are based on an enlargement of the state space with the introduction of a coupling with some hidden variables; what we show here is that the coupled measures, which could in principle be different, are instead the same modulo a simple mapping.

The two lines ensemble is as follows. We have that $s^1_1 = s^2_1 = 0$, and for $2 \le i \le N+1$ we have $s^1_i = \sum_{j=1}^{i-1} \eta_j$, $s^2_i = \sum_{j=1}^{i-1} \gamma_j$, where $\eta, \gamma \in \{0,1\}^N$. We call η the increments of the walk s^1 since they will be finally distributed according to the invariant measure of TASEP. The two-lines ensembles representation in [6, 2] is written in terms of the two functions s^1, s^2

defining the coupling measure

$$C_N^{\text{2-lines}}\left(s^1,s^2\right) := \frac{\left(\frac{1-\beta}{\beta}\right)^{\left\{s^1_{N+1}-s^2_{N+1}\right\}}}{\mathcal{Z}_N\left(\frac{(1-\beta)(1-\alpha)}{\beta\alpha}\right)^{\min_{1\le i\le N+1}\left\{s^1_i-s^2_i\right\}}}\left(\frac{1}{4}\right)^N . \tag{3.3.7}$$

Here $\mathcal{Z}_N$ is a normalization factor and $\left(1/4\right)^N = \mathbb{P}_{\text{Unif.}}(s^1,s^2)$ denotes the uniform measure on the pair of paths which we do not include in the normalization factor just for convenience. In [6] it is proved that $\mu_N^{\text{2-lines}}(s^1) = \sum_{s^2} C_N^{\text{2-lines}}\left(s^1,s^2\right)$, coincides with $\mu_N(\eta)$, by using the bijection between s^1 and η.

Let us introduce the $\mathbb{N}_0\times\mathbb{N}_0$ stochastic matrix $\hat{P}$ defined by

$$\hat{P}_{b,b'} = \begin{cases} \frac{1}{2} & b'=b \\ \frac{1}{4} & b'=b+1 \\ \frac{1}{4} & b'=b-1\,, b>0\,, \\ \frac{1}{2} & b=0, b'=1\,. \end{cases} \tag{3.3.8}$$

This stochastic matrix is related to (3.3.4) through the following relations

$$\hat{P}_{b,b'} = \begin{cases} \frac{P_{b,b'}e_b}{e_{b'}}\,, & b>0\,, \\ P_{b,b'}\,, & b=0\,. \end{cases}$$

This means that the stochastic matrix $\hat{P}$ is almost related by a Doob transformation to the matrix P; the only transition for which this relation fails is the transition $0\to 1$. Consider $\zeta\in\mathbb{N}_0^{N+1}$ and call $\mathcal{N}_{0,1}(\zeta) := \sum_{i=1}^N \mathbb{1}\Big(\zeta_i^{i+1}=(0,1)\Big)$. By the above relations, we obtain

$$\mathbb{P}_P(\zeta) = \mathbb{P}_{\hat{P}}(\zeta)\frac{\zeta_{N+1}+1}{\zeta_1+1}\left(\frac{1}{2}\right)^{\mathcal{N}_{0,1}(\zeta)}, \qquad \zeta\in\mathbb{N}_0^{N+1}. \tag{3.3.9}$$

This allows to write the Markov bridge for the stochastic matrix P in terms of the stochastic matrix $\hat{P}$ as

$$\rho_{N+1}(\zeta) = \frac{\left(\frac{1-\alpha}{\alpha}\right)^{\zeta_1}\mathbb{P}_{\hat{P}}(\zeta)\left(\frac{1-\beta}{\beta}\right)^{\zeta_{N+1}}\left(\frac{1}{2}\right)^{\mathcal{N}_{0,1}(\zeta)}}{g^T P^N f}, \tag{3.3.10}$$

where the denominator is simply the normalization factor.

Consider a random walk $\xi_i\in\mathbb{Z}$ that jumps to $\xi_i\pm 1$ with probability $1/4$ and stays at ξ_i with probability $1/2$. Its $\mathbb{Z}\times\mathbb{Z}$ transition probability $\tilde{P}$ is given by

$$\tilde{P}_{b,b'} = \begin{cases} \frac{1}{2} & b,b'\in\mathbb{Z}\,,\ b'=b \\ \frac{1}{4} & b,b'\in\mathbb{Z}\,,\ b'=b\pm 1\,. \end{cases} \tag{3.3.11}$$

The stochastic matrix $\hat{P}$ is the transition probability of the Markov process $\zeta_i = |\xi_i|$. We

have, in addition, that for any $\zeta \in \mathbb{N}_0^{N+1}$

$$\mathbb{P}_{\hat{P}}(\zeta)\left(\frac{1}{2}\right)^{\mathcal{N}_{0,1}(\zeta)} = \mathbb{P}_{\tilde{P}}(b+\zeta) \qquad \forall b \in \mathbb{Z}\,, \tag{3.3.12}$$

where by $b+\zeta$ we denote the shifted path $\left(b+\zeta_i\right)_{i=1}^{N+1}$ (this means that $\mathbb{P}_{\tilde{P}}(b+\zeta) = \left[\prod_{i=1}^{N} \tilde{P}_{b+\zeta_i, b+\zeta_{i+1}}\right]$). The arbitrary shift factor follows from the fact that $\tilde{P}$ is associated with a space homogeneous random walk on $\mathbb{Z}$.

Recall the paths $(s_i^1)_{i=1}^{N+1}$ and $(s_i^2)_{i=1}^{N+1}$ introduced for defining the two lines ensemble and call $\xi_i = s_i^1 - s_i^2$. We can describe the independent uniform measure on the two paths either in terms of the pair (s^1, s^2) or in terms the pairs (η, ξ); we recall that η are the increments of s^1 and ξ has just been defined. With a slight abuse of notation we denote by $\mathbb{P}_{\text{Unif.}}(\eta, \xi)$ the measure $\mathbb{P}_{\text{Unif.}}(s^1, s^2)$ when expressed in terms of the variables (η, ξ). By a direct computation, we have

$$\mathbb{P}_{\text{Unif.}}(\eta, \xi) = \mathbb{P}_{\tilde{P}}(\xi)\mathbb{P}_{\text{Unif.}}(\eta|\xi) = \mathbb{P}_{\tilde{P}}(\xi)\prod_{i=1}^{N} p_{\xi_i, \xi_{i+1}}(\eta_i)\,, \tag{3.3.13}$$

where the probability measures $p_{b,b'}(\cdot)$ are those defined in (3.3.6). An important feature of such measures is that they are invariant by a joint shift of the indices, that is, $p_{b,b'}(\cdot) = p_{b+c,b'+c}(\cdot)$ for any $c \in \mathbb{Z}$. This means that the Markov bridge (3.3.10) or any translation of this measure are good parameter laws for the mixture in order to obtain the invariant measure of boundary driven TASEP.

Fix $\xi = (\xi_i)_{i=1}^{N+1} \in \mathbb{Z}^{N+1}$ such that $\xi_1 = 0$ and $|\xi_i - \xi_{i+1}| \leq 1$ and compute the weight assigned by the probability measure (3.3.10) to all the translated paths of ξ. We therefore have to sum the probability given by (3.3.10) over all paths of the form $\zeta = b + \xi$, where $b + \xi$, as before, is the path shifted by $b \in \mathbb{Z}$. Since $\zeta \in \mathbb{N}_0^{N+1}$, we have the constraint $b \geq -\min_{1\leq i\leq N+1} \xi_i := -m_{N+1}(\xi)$. In this way, we obtain the probability measure $\tilde{\rho}_{N+1}$ on paths in $\mathbb{Z}$ starting from the origin, defined by

$$\tilde{\rho}_{N+1}(\xi) := \sum_{\{b \geq -m_{N+1}(\xi)\}} \rho_{N+1}(b+\xi) \tag{3.3.14}$$

$$= \frac{\mathbb{P}_{\tilde{P}}(\xi)\sum_{\{b\geq -m_{N+1}(\xi)\}}\left(\frac{1-\alpha}{\alpha}\right)^b\left(\frac{1-\beta}{\beta}\right)^{b+\xi_{N+1}}}{g^T P^N f} \tag{3.3.15}$$

$$= \frac{\mathbb{P}_{\tilde{P}}(\xi)\left(\frac{1-\beta}{\beta}\right)^{\xi_{N+1}}}{\left(\frac{(1-\alpha)(1-\beta)}{\alpha\beta}\right)^{m_{N+1}(\xi)}(g^T P^N f)\left(1-\frac{(1-\alpha)(1-\beta)}{\alpha\beta}\right)}\,. \tag{3.3.16}$$

where we used (3.3.12) in the second equality, while the last equality follows from a direct computation of the geometric sum. By the translational invariance of (3.3.6) we have that the coupled measure

$$C_N(\eta, \xi) = \tilde{\rho}_{N+1}(\xi)\prod_{i=1}^{N} p_{\xi_i, \xi_{i+1}}(\eta_i) \tag{3.3.17}$$

is such that $\mu_N(\eta) = \sum_\xi C_N(\eta, \xi)$ is the invariant measure of boundary driven TASEP. Using (3.3.13) and writing the measure in terms of the paths (s^1, s^2), we obtain the two lines ensemble

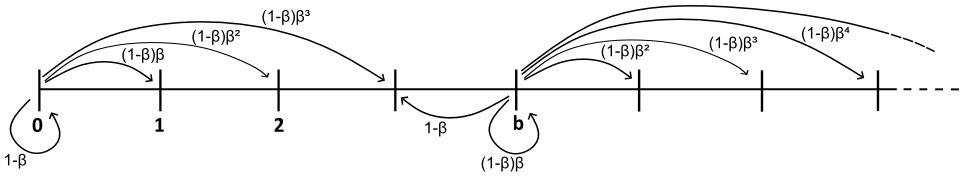

Figure 3.3: The transition graph on $\mathbb{N}_0$ for the transition matrix (3.3.20); the Markov dynamics may have arbitrary long jumps.

(3.3.7) with $\mathcal{Z}_N = \left(g^T P^N f\right)\left(\frac{\alpha+\beta-1}{\alpha\beta}\right)$.

3.3.2 The general case

A different solution of the algebraic relations (3.3.1), valid for any value of the boundary parameters α and β, can be obtained again with $B = \mathbb{N}_0$, using the matrices and vectors introduced in (2.3.10).

We recall that this representation satisfies the same algebraic relations as before and even if it is more complex, it extends to the full range of parameters, without requiring any restriction. The matrix $M = M^0 + M^1$ has the entries defined by:

$$M_{b,b'} = \begin{cases} \frac{1}{\beta} & b = 0\,, \\ 1 & \forall\ b' \geq b-1\,. \end{cases} \tag{3.3.18}$$

The eigenvalue problem for an eigenvalue λ and a corresponding eigenvector $e = (e_b)_{b\in\mathbb{N}_0}$ is obtained by solving the equations

$$\begin{cases} \lambda e_0 = \frac{1}{\beta}\sum_{j=0}^{+\infty} e_j\,, \\ \lambda e_i = \sum_{j=i-1}^{+\infty} e_j\,, \quad i \geq 1\,. \end{cases} \tag{3.3.19}$$

One can check that a positive eigenvalue is given by $\lambda = \frac{1}{\beta(1-\beta)}$ with corresponding positive eigenvector $e = (\beta^b)_{b\in\mathbb{N}_0}$. We can then apply the transformation (3.2.1), obtaining explicitly the corresponding stochastic matrix:

$$P_{b,b+k} = \begin{cases} (1-\beta)\beta^k & b = 0,\ k \geq 0\,, \\ (1-\beta)\beta^{k+1} & \forall\ b \geq 1,\ k \geq -1\,, \end{cases} \tag{3.3.20}$$

and all the remaining entries are zero. The transition graph of the stochastic matrix P on $\mathbb{N}_0$ is drawn in Figure 3.3. The vectors defining the Markov bridge are $g = \left(\left(\frac{\beta}{\alpha}\right)^b\right)_{b\in\mathbb{N}_0}$ and $f = \left(\delta_{0,b}\right)_{b\in\mathbb{N}_0}$. The marginal distributions (3.2.4) can also be easily computed and are given by:

$$p_{b,b'} = \begin{cases} \mathcal{B}_0 & b' = b-1\,, \\ \mathcal{B}_1 & \forall\ b' \geq b\,. \end{cases} \tag{3.3.21}$$

Recall that $\mathcal{B}_q$ is a Bernoulli measure of parameter q. Since the parameters take only the values 1 and 0, it follows the particle configuration η is completely determined by the Markov

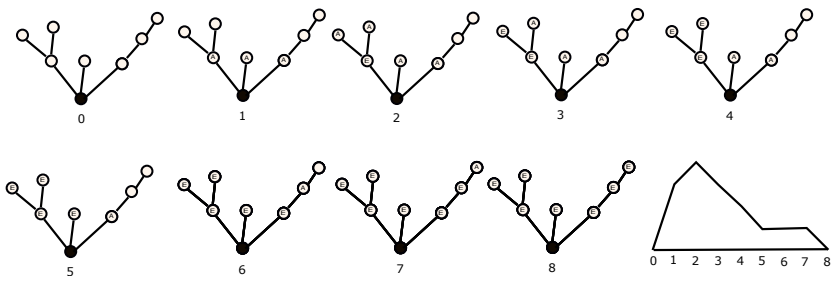

Figure 3.4: The exploration sequence of a rooted tree; the root is the black node. The symbol A means Activated, while the symbol E means Explored; at the end of the figure, the path of the corresponding walk ζ, associated to the exploration procedure is shown.

bridge ζ.

An exploration process. We now discuss a natural and simple interpretation of the Markov process obtained above. There is a well-known exploration algorithm for branching processes (see, for example, Section 3.3 in [61]) that connects branching processes to random walks on $\mathbb{N}_0$ which may have arbitrarily large positive jumps and negative jumps only of size -1. The Markov bridge obtained above is constructed starting from Markov chains of this type. We describe a small variation of the classical construction; in particular, we consider an exploration of infinite independent branching processes, each of them attached to its own root vertex. To each root, we attach an independent copy of a branching process with branching law $q \in \mathcal{M}^1(\mathbb{N}_0)$. Consider now the random walk $(\zeta_n)_{n\in\mathbb{N}}$ associated with the following exploration procedure. We start with $\zeta_0 = 0$ and select one root. The value ζ_1 is the number of nodes attached to the root, and we set all of them in the *activated* status. Select one node among the activated ones, according, for example, to the depth-first search and set it in the *explored* status. This node no longer belongs to the set of activated nodes (it has been explored), and we add to the set of activated nodes all the vertices in the progeny of the node just explored. The variable ζ_2 is the number of activated nodes after this construction. We can now iterate the procedure and denote by ζ_n the number of activated nodes after n iterations. If after n steps, all nodes of the tree associated with the branching process have been explored, then $\zeta_n = 0$, since there are no more activated nodes. One then proceeds, in the next step, to the exploration of the branching tree attached to another root, and the variable ζ_{n+1} is the number of nodes attached to the new root, and so on. The random walk described above has the transition probability equal to P in (3.3.20) when the branching law q is geometric of parameter $(1-\beta)$, that is, $q_k = (1-\beta)\beta^k$, $k = 0, 1, 2, ...$

In conclusion, these examples suggest that distinct matrix representations which satisfy the algebraic relations (3.3.1) may correspond to several hidden stochastic mechanisms that nevertheless produce the same invariant measure.

In this sense, different matrix representations can be viewed as different probabilistic constructions of the same object. Here we have explicitly examined two of these for the TASEP, but, as mentioned in the previous chapter, several others are known in the literature even in

the finite case. We now conclude the discussion of the TASEP and turn to the asymmetric case (ASEP).

3.4 Boundary-driven ASEP

We consider the boundary-driven asymmetric exclusion process (ASEP), where particles jump from site x to $x+1$ with rate q and from site $x+1$ to site x with rate 1. In the general case with four boundary parameters, as explained in the Section 2.4, its invariant measure can be expressed in MPA form (3.1.1) with the state space $A = \{0, 1\}$, when the matrices M^0, M^1 and the vectors $\vec{y}$ and $\vec{x}$ satisfy the algebraic conditions:

$$\begin{aligned} &M^1 M^0 - qM^0 M^1 = M^1 + M^0 \\ &\vec{y}^T[\alpha M^0 - \gamma M^1] = \vec{y}^T \\ &[\beta M^1 - \delta M^0]\vec{x} = \vec{x}. \end{aligned} \tag{3.4.1}$$

In what follows, we first apply Theorem 3.6 to the case with two boundary parameters (α,β) and subsequently to the general case with four parameters $(\alpha, \beta, \gamma, \delta)$.

3.4.1 The two-parameter case ($\delta = \gamma = 0$)

We start by considering the simplified version with open boundary conditions where particles enter at site 1 with rate α and exit from site N with rate β. In this setting, we do not consider the exit of particles at site 1 and the entrance of particles at site N, that is, $\gamma = \delta = 0$.
Under these assumptions, the boundary relations reduce to the condition that the vector $\vec{y}$ is a left-eigenvector of M^0 with eigenvalue $\frac{1}{\alpha}$ and the vector $\vec{x}$ is a the right-eigenvector of M^1 with eigenvalue $\frac{1}{\beta}$.
One possible representation, satisfying these relations is obtained by taking the auxiliary space $B = \mathbb{N}_0$ and defining the matrices:

$$M^0 = \frac{1}{1-q}\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots \\ 1-q & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1-q^2 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1-q^3 & 1 & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \; M^1 = \frac{1}{1-q}\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 1 & \ddots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{3.4.2}$$

The corresponding boundary vectors take the form:

$$\vec{y} = \left(\frac{1}{(q;q)_b}\left(\frac{1-q-\alpha}{\alpha}\right)^b\right)_{b\in\mathbb{N}_0}, \qquad \vec{x} = \left(\left(\frac{1-q-\beta}{\beta}\right)^b\right)_{b\in\mathbb{N}_0}$$

where the symbol $(x;q)_n$ denotes the q-shifted factorial, defined by $(x;q)_0 = 1$ and, for $n > 0$:

$$(x;q)_n = \prod_{k=0}^{n-1}(1 - xq^k).$$

In order to ensure that the boundary vectors are positive, we restrict to the parameter region with $\alpha, \beta < 1 - q$.
The matrix $M = M^0 + M^1$ can be written as $M = \frac{1}{1-q}\tilde{M}$ where $\tilde{M}$ has the entries defined as:

$$\tilde{M}_{b,b'} = \begin{cases} 2\delta_{b,b'} + \delta_{b,b'+1} + (1-q^b)\delta_{b,b'-1}\,, & b \geq 1\,, \\ 2\delta_{b,b'} + \delta_{b,b'+1}\,, & b = 0\,. \end{cases} \tag{3.4.3}$$

Then, the eigenvalue problem $M\vec{e} = \lambda_q \vec{e}$ reduces to solve the following recurrence system:

$$\begin{cases} 2e_0 + e_1 = \lambda_q(1-q)e_0\,, \\ (1-q^n)e_{n-1} + 2e_n + e_{n+1} = \lambda_q(1-q)e_n \quad n \geq 1\,. \end{cases} \tag{3.4.4}$$

Note that for the specific choice of parameters $\lambda_q = \frac{4}{1-q}$ and $e_0 = 1$ the recurrence system (3.4.4) simplifies considerably and takes a well-known form.
In this case, the coefficients appearing in the relation coincide exactly with those characterizing the recurrence that defines the family of continuous q-Hermite polynomials. This correspondence is exact: with this choice of the eigenvalue λ_q and of the initial condition $e_0 = 1$, both the recurrence relation and the starting values coincide with those of the continuous q-Hermite polynomials evaluated at $x = 1$.
Recall that the continuous q-Hermite polynomials $H_n(x|q)$ are defined recursively by:

$$2xH_n(x|q) = H_{n+1}(x|q) + (1-q^n)H_{n-1}(x|q), \qquad H_0(x|q) = 1 \tag{3.4.5}$$

which determines the whole sequence $(H_n)_{n\geq 0}$, (for a review on q-Hermite polynomials see [58] or [13]).
Comparing the recurrence (3.4.4) with the one above, it is evident that the two problems have the same algebraic structure. More precisely, the eigenvalue problem for the matrix M can be interpreted as a particular realization of the recurrence relation that defines the continuous q-Hermite polynomials at the specific point $x = 1$.
It follows that the eigenvector of the matrix M associated with the eigenvalue $\lambda = 4/(1-q)$ has, as components, the succession of the continuous q-Hermite polynomials centered in $x = 1$ which can be written as:

$$e_n = H_n(1|q) = \sum_{k=0}^{n} \binom{n}{k}_q \tag{3.4.6}$$

where $\binom{n}{k}_q = \frac{(q;q)_n}{(q,q)_k (q;q)_{n-k}}$ denotes the q-binomial coefficient. Now, the stochastic matrix (3.2.1) can be explicitly computed. For convenience in notation, we introduce the quantity:

$$K_{n,q} = \frac{\sum_{k=0}^{n} q^k \binom{n}{k}_q}{\sum_{k=0}^{n} \binom{n}{k}_q} \tag{3.4.7}$$

which can be viewed as the expected value of q^k with respect to the normalized q-binomial distribution. Using this definition, the ratio between consecutive continuous q-Hermite poly-

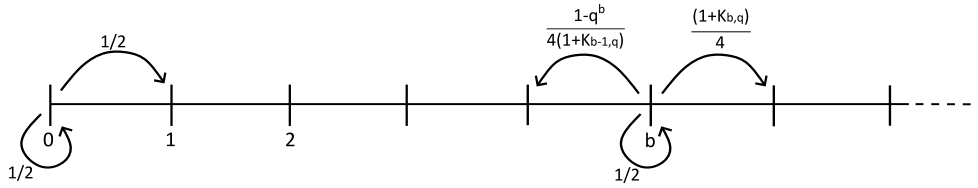

Figure 3.5: The transition graph associated with the stochastic matrix P in (3.4.8) with $K_{b,q}$ defined in (3.4.7)

nomials is given by:

$$\frac{H_{n+1}(1|q)}{H_n(1|q)} = 1 + K_{n,q}$$

and the stochastic matrix can be written in the compact form:

$$P_{b,b'} = \begin{cases} \frac{1}{2} & b' = b \\ \frac{1}{4}(1 + K_{b,q}) & b' = b+1 \\ \frac{(1-q^b)}{4(1+K_{b-1,q})} & b' = b-1\,, \end{cases} \tag{3.4.8}$$

The transition graph of the stochastic matrix P on $\mathbb{N} \cup \{0\}$ is represented in Figure 3.5. The vectors which define the Markov bridge become:

$$\begin{cases} f_b = \frac{x_b}{H_b(1|q)} = \frac{(1-q-\beta)^b}{\beta^b H_b(1|q)} & b \in \mathbb{N}_0\,, \\ g_b = y_b H_b(1|q) = \frac{(1-q-\alpha)^b}{\alpha^b (q;q)_b} H_b(1|q) & b \in \mathbb{N}_0\,. \end{cases} \tag{3.4.9}$$

and the marginal distributions of the product measure are:

$$p_{b,b'} = \begin{cases} \mathcal{B}_1 & \text{if } b' = b+1\,, \\ \mathcal{B}_0 & \text{if } b' = b-1\,, \\ \mathcal{B}_{\frac{1}{2}} & \text{if } b' = b\,. \end{cases} \tag{3.4.10}$$

Finally, in the limiting case $q = 0$, the present construction reproduces exactly the totally asymmetric representation. Indeed, setting $q = 0$ in the matrices M^1, M^0 and in the vectors $\vec{y}, \vec{x}$, the eigenvalue problem and its eigenvectors coincide with those computed for the TASEP representation in Section 3.3.1, yielding the same stochastic matrix P with identical boundary vectors defining the Markov bridge and the same behavior for the marginal distributions.
Hence, the probabilistic representation developed here provides a unified description that continuously interpolates between the totally asymmetric and the partially asymmetric regimes, preserving the structure of the stationary measure across different values of $q \in [0, 1)$.

3.4.2 The general boundary case $(\alpha, \beta, \gamma, \delta)$

We now extend the analysis to the general boundary-driven ASEP characterized by four parameters $(\alpha, \beta, \gamma, \delta)$. In this setting, we recall that particles can both enter and exit the system at each boundary, at site 1 with rate α and γ and at site N with δ and β, respectively (Figure 1.3).
The bulk dynamics and the algebraic relation between the matrices remain unchanged while

the effect of the boundary terms is entirely encoded in the conditions involving the vectors $\vec{y}, \vec{x}$, (see formula (3.4.1)). Therefore, the q-Hermite structure in the bulk is preserved, while the additional parameters γ and δ deform the boundary conditions of the probabilistic representation.
Consequently, the stochastic matrix P associated with the process has the same form as in the two-parameter case (3.4.8). The problem reduces to finding the vectors $\vec{x}$ and $\vec{y}$ which satisfy the conditions:

$$\begin{aligned} \vec{y}^T[\alpha M^0 - \gamma M^1] &= \vec{y}^T \\ [\beta M^1 - \delta M^0]\vec{x} &= \vec{x}. \end{aligned} \tag{3.4.11}$$

Once these boundary vectors are determined, the whole probabilistic construction - including the Doob-transformed dynamics and the Markov bridge representation - follows exactly as in the simpler case discussed above.
To determine the vector $\vec{y}$ satisfying the first relation in (3.4.11) we expand it in components and use the explicit form of the matrices M^0 and M^1:

$$\begin{cases} (\alpha - \gamma)y_0 + \alpha(1-q)y_1 = (1-q)y_0, \\ -\gamma y_{n-1} + (\alpha - \gamma)y_n + \alpha(1-q^{n+1})y_{n+1} = (1-q)y_n \quad n \geq 1 \end{cases} \tag{3.4.12}$$

By fixing $y_0 = 1$ and after a suitable rescaling, one obtains a solution of the form:

$$y_n = \frac{(\sqrt{\gamma/\alpha})^n}{(q;q)_n} i^n H_n(iA/2|q) \qquad \text{with} \quad A = \frac{(\alpha - \gamma - 1 + q)}{\sqrt{\gamma\alpha}}$$

where the family $(H_n)_{n\geq 0}$ satisfies the recurrence relation defining the continuous q-Hermite polynomials, see (3.4.5).
Each of the three prefactors plays a specific role in this representation. The factor $(\sqrt{\gamma/\alpha})^n$ symmetrizes the boundary recurrence, compensating the asymmetry between the entrance and exit rates α and γ, and allowing the equation to be rewritten in a form directly comparable to the recurrence of the continuous $q-$Hermite polynomials. The term $(q;q)_n$ is introduced to absorb the factor $(1-q^{n+1})$ in order to transfer its dependence from the forward term to the backward one. Finally, the factor i^n is introduced as a formal rescaling that adjusts the global sign of the recurrence, aligning it with the canonical convention used for the continuous q-Hermite polynomials. Although the expression involves the imaginary unit i, the sequence $(y_n)_{n\geq 0}$ is real. Indeed, each polynomial $H_n(x|q)$ can be expanded as:

$$H_n(x|q) = \sum_{k=0}^{\lfloor n/2 \rfloor} a_{n,k}(q)\ x^{n-2k}, \qquad a_{n,k}(q) \in \mathbb{R}$$

which shows that $H_n(x|q)$ contains only power of x with the same parity as n. As a consequence, substituting $x = i\xi$ with $\xi \in \mathbb{R}$ we obtain:

$$H_n(i\xi|q) = i^n \sum_{k=0}^{\lfloor n/2 \rfloor} a_{n,k}(q)(-\xi^2)^k$$

hence the element $i^n H_n(i\xi|q)$ is real for every n and for all real ξ. In particular, this implies that each component y_n is real.
The positivity of the sequence $(y_n)_{n\geq 0}$ can be determined by induction on the recurrence, which requires $A < 0$, that is $\alpha - \gamma < 1 - q$. This condition ensures that all components y_n are positive and the boundary vector $\vec{y}$ is well defined.
Using analogous arguments, one can prove that the vector $\vec{x}$ which solves the second relation in (3.4.11) is given by:

$$x_n = \left(\sqrt{\frac{\delta}{\beta}}\right)^n i^n H_n(iB/2|q), \qquad \text{with} \quad B = \frac{\beta - \delta - 1 + q}{\sqrt{\delta\beta}}$$

which has positive components when $B < 0$, that is $\beta - \delta < 1 - q$.
The vectors defining the Markov bridge are:

$$\begin{cases} f_b = \frac{x_b}{H_b(1|q)} & b \in \mathbb{N}_0\,, \\ g_b = y_b H_b(1|q) & b \in \mathbb{N}_0\,. \end{cases} \tag{3.4.13}$$

where $H_b(1|q)$ is defined in (3.4.6) and $\vec{x}, \vec{y}$ are the vectors derived above. These expressions complete the probabilistic representation for the four-parameter boundary ASEP.

As we have seen in the previous chapter, the multispecies models also admit an invariant measure written in terms of matrices; in the following section we move the attention to the 2-TASEP by computing the stochastic matrix associated with a specific form of matrices, the Markov bridge and the corresponding mixture representation. Finally, we show how to recover the queue representation of the invariant measure [1, 34, 38] from our construction.

3.5 TASEP with second class particles on a ring

In this section we study the TASEP on $\mathbb{Z}_N$, the ring with N sites where particles can be of two different type, first and second class particles. We recall that there exists a priority scales between the particles, those of higher class have lower priority (for more details on the model refer to Section 1.3 and 2.6).
We also remind that the matrix product formulation on a periodic lattice is slightly different from (3.3.1): we have that the stationary measure is given by [22]:

$$\mu_N^{n_1,n_2}(\eta) = \frac{\mathrm{Tr}\left[\prod_{i=1}^N M^{\eta_i}\right]}{Z_N(n_1,n_2)} \mathbb{1}\Big(\mathcal{N}_1(\eta) = n_1\,, \mathcal{N}_2(\eta) = n_2\Big)\,, \tag{3.5.1}$$

where the normalization factor $Z_N(n_1, n_2)$ depends on the numbers n_1, n_2 of the first and second class particles. The matrices M^i, $i = 0, 1, 2$, have to satisfy the following conditions:

$$M^1 M^0 = M^1 + M^0, \qquad M^1 M^2 = M^2, \qquad M^2 M^0 = M^2.$$

A solution to the above algebraic relations is obtained with $B = \mathbb{N}_0$ and the matrices given by (2.6.2).

We define the matrix $M = M^0 + M^1 + M^2$ that has the form

$$M = \begin{bmatrix} 3 & 1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots \\ 1 & 2 & 1 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 1 & 2 & 1 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & 2 & 1 & 0 & \cdots & \cdots \\ \vdots & \vdots & 0 & 1 & 2 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

The sum of each row of M is constant and equal to 4 so that $\lambda = 4$ is a positive eigenvalue with corresponding constant eigenvector $e = (1)_{b\in\mathbb{N}_0}$. By the transformation (3.2.1) we obtain the stochastic matrix P defined by

$$P_{b,b'} = \begin{cases} 1/4 & b > 0\,,\, b' = b \pm 1\,, \\ 1/2 & b > 0\,,\, b' = b\,, \\ 1/4 & b = 0\,,\, b' = 1\,, \\ 3/4 & b = 0\,,\, b' = 0\,, \end{cases} \tag{3.5.2}$$

whose transition graph is represented in Figure 3.6.

The set of invariant probability measures is a convex set whose extremal elements are the measures in (3.5.1). In [22] the authors introduce a grand canonical invariant probability measure

$$\mu_N(\eta) = \frac{\mathrm{Tr}\left[\prod_{i=1}^N M^{\eta_i}\right]}{Z_N} \mathbb{1}\Big(\mathcal{N}_2(\eta) > 0\Big)\,, \tag{3.5.3}$$

that is a non trivial convex combination of the extremal invariant measures. We indeed have that $\mu_N = \sum_{n_1,n_2} \frac{Z_N(n_1,n_2)}{Z_N} \mu_N^{n_1,n_2}$, where the sum is over the pairs (n_1, n_2) such that $n_1 + n_2 \leq N$ and $n_2 > 0$. The extremal measures can be obtained as the canonical measures of the grand canonical one

$$\mu_N\left(\cdot|\mathcal{N}_1 = n_1, \mathcal{N}_2 = n_2\right) = \mu_N^{n_1,n_2}(\cdot)\,.$$

To give a mixture representation of the invariant measures we need to introduce another probability measure, since the construction of Theorem 3.6 using the measure (3.5.3) does not work properly. This is due to the fact that, if we start from the measure (3.5.3) and enlarge it to the coupling measure C_N, as we did before, our general strategy fails. In particular it is not possible to implement the marginal computation in the item 3) of the proof of Theorem (3.6), since the coupled measure contains constraints written in terms of the variables η. This constraint is contained in the characteristic function and, as a result, the marginal ζ law of the coupling is not a simple Markov process. Therefore, we construct therefore another grand canonical measure whose canonical measures are again the measures $\mu_N^{n_1,n_2}$.

Fix j a node of the ring and let us define

$$\mu_{N,j} := \frac{\mathrm{Tr}\left[\prod_{i=1}^N M^{\eta_i}\right]}{Z_N^*} \mathbb{1}\Big(\eta_j = 2\Big)\,, \tag{3.5.4}$$

where by symmetry the normalization factor Z_N^* does not depend on j. We also define the translational invariant measure $\hat{\mu}_N := \frac{1}{N}\sum_{j=1}^N \mu_{N,j}$ for which we have

$$\hat{\mu}_N(\eta) = \frac{\mathrm{Tr}\left[\prod_{i=1}^N M^{\eta_i}\right]\sum_{j=1}^N \mathbb{1}(\eta_j = 2)}{NZ_N^*} = \frac{Z_N \mu_N(\eta)\mathcal{N}_2(\eta)}{NZ_N^*}\,. \tag{3.5.5}$$

By a direct computation it is easy to see that the measure (3.5.5) can be written as $\hat{\mu}_N = \sum_{n_1,n_2} \frac{n_1 Z_N(n_1,n_2)}{NZ_N^*}\mu_N^{n_1,n_2}$ (we recall once again that the sum is over (n_1,n_2) such that $n_1+n_2 \le N$ and $n_2 > 0$). In particular, the measure (3.5.1) are again the conditional measures of $\hat{\mu}_N$.

The coupling construction of Theorem 3.6 is particularly simple for the measures $\mu_{N,j}$ and consequently for $\hat{\mu}_N$. The motivation is that, in this case, the constraint on the η variables can be transformed into a constraint on the ζ variables. By symmetry the construction for different values of j is the same up to a rotation; for simplicity of notation we discuss the case $\mu_{N,N}$. For each j we can construct the coupling $C_{N,j} \in \mathcal{M}^1(A^N \times B^N)$ which in the case $j = N$ is

$$\begin{aligned} C_{N,N}(\eta,\zeta) &:= \frac{\left[\prod_{i=1}^N M^{\eta_i}_{\zeta_i,\zeta_{i+1}}\right]\mathbb{1}\Big(\eta_N = 2\Big)}{Z_N^*} \\ &= \frac{3\left[\prod_{i=1}^{N-1} M^{\eta_i}_{\zeta_i,\zeta_{i+1}}\right]\mathbb{1}\Big(\eta_N = 2, \zeta_1 = 0, \zeta_N = 0\Big)}{Z_N^*}\,. \end{aligned}$$

The ζ marginal of the coupled measure $C_{N,N}$ can be easily computed since the sum over the η_i with $i \neq N$ has no constraints while there is no sum over η_N that is fixed equal to 2; we thus obtain

$$\rho_{N,N}(\zeta) = \sum_\eta C_{N,N}(\eta,\zeta) = \frac{3\prod_{i=1}^{N-1} P_{\zeta_i,\zeta_{i+1}}}{4^{N-1}Z_N^*}\mathbb{1}\Big(\zeta_1 = 0, \zeta_N = 0\Big)\,. \tag{3.5.6}$$

In the case of periodic boundary conditions, since the system is defined on a ring, the number of ζ variables coincides with the number of the η variables. We therefore obtain a Markov bridge of length N rather than $N+1$ as in the open boundary case. According to the notation (3.1.4), the measure (3.5.6) is the Markov Bridge $\rho_N^{0,0}$ of length N with transition matrix P and pinned to start and finish at $\zeta_1 = \zeta_N = 0$. It is geometrically natural to associate the variables ζ with the dual lattice of the one associated with the variables η, in such a way that ζ_i and ζ_{i+1} are associated with the points of the dual lattice corresponding to the edges exiting from the vertex of the original lattice associated with the variable η_i. We will use this shifted association in the discussion of the comparison with other representations. A similar construction can be done in the open boundary case. By rotational symmetry, the measures $\rho_{N,j}$ have the same structure after a rotation.

Also in this case, the marginals of the product measures can be directly computed. For any $b, b' \in B$ we have $p_{b,b'}(\cdot) \in \mathcal{M}^1(\{0,1,2\})$ and we use the notation $p_{b,b'} = (p_{b,b'}(0), p_{b,b'}(1), p_{b,b'}(2))$. By computing (3.2.4) we obtain four different probability measures depending on the values

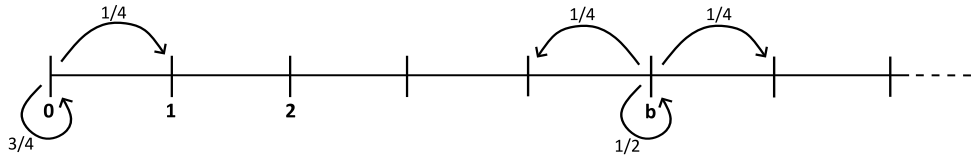

Figure 3.6: The transition graph associated to the Markov matrix P (3.5.2)

of b, b'

$$\begin{cases} p_{0,0} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) , \\ p_{b,b+1} = (0,1,0) , \\ p_{b,b-1} = (1,0,0) , \\ p_{b,b} = \left(\frac{1}{2}, \frac{1}{2}, 0\right) , \quad b \geq 1 . \end{cases} \tag{3.5.7}$$

To construct a configuration η distributed according to the grand canonical invariant measure (3.5.5), we place a second class particle at N, i.e. we fix $\eta_N = 2$. We then construct a sample path ζ according to the Markov Bridge $\rho_{N,N} = \rho_N^{0,0}$. Given ζ we generate the η_i, $i \neq N$, independently using (3.5.7). Once the configuration η has been obtained in this way, we insert it into the ring by simply shifting it by a uniform rotation.

Comparison with the queue representation. We start by recalling the collapsing representation of the invariant measure [1] and its queue interpretation and generalization [38]. An interesting issue would be to consider the mixture representation in the multiclass case of [38] whose representation in MPA form is obtained in [34].

We use the language of [38] with a left-right symmetry exchange of the arguments due to the different orientation of the dynamics considered there. Consider two subsets $\mathcal{A}, \mathcal{S} \subseteq \mathbb{Z}_N$ such that $|\mathcal{A}| < |\mathcal{S}|$; they may have elements in common. The set $\mathcal{A}$ represents the set of arrivals of clients while the subset $\mathcal{S}$ represents the set of services. The time axis for the queue interpretation runs from right to left, according to the same order of going from $i+1$ to i. We introduce the variables $(A_i)_{i\in\mathbb{Z}_N}$ and $(S_i)_{i\in\mathbb{Z}_N}$, defined by $A_i = 1$ if $i \in \mathcal{A}$ and 0 otherwise, while $S_i = 1$ if $i \in \mathcal{S}$ and 0 otherwise. The construction in [1, 38] is as follows. We construct iteratively a subset $\mathcal{S}^+ \subseteq \mathcal{S}$, which is the set of *used* service times. By construction, we will have $|\mathcal{S}^+| = |\mathcal{A}|$ and we denote by $\mathcal{S}^- := \mathcal{S} \setminus \mathcal{S}^+$ the set of *unused* services. Initially, $\mathcal{S}^+$ is empty. Select an arbitrary client (that is, an $i \in \mathcal{A}$) and serve it in the nearest unused service time to its left (a $j \in \mathcal{S} \setminus \mathcal{S}^+$). The selected service time is then added to the set $\mathcal{S}^+ \subseteq \mathcal{S}$ of used services. We iterate the procedure until a service time has been assigned to each client. Given $\mathcal{S}^\pm$, we assign a particle configuration by placing a first class particle at each vertex in $\mathcal{S}^+$, a second class particle at each vertex in $\mathcal{S}^-$ and leaving the remaining vertices empty. When $|\mathcal{S}| = n_1 + n_2$, $|\mathcal{A}| = n_1$ and they are uniformly distributed among all sets satisfying these constraints, then the resulting particle configuration is distributed according to $\mu_N^{n_1,n_2}$ in (3.5.1).

Since our mixture representation is associated with the measures (3.5.4), we need to slightly modify this construction in order to make a direct comparison. We consider the case $j = N$ and we denote by $\mathbb{Z}_N^*$ the dual lattice. Consider two subsets $\mathcal{S}, \mathcal{A}$ of $\mathbb{Z}_N$ such that $|\mathcal{A}| < |\mathcal{S}|$. We represent the elements of $\mathbb{Z}_N$ periodically on $\mathbb{Z}$ and fix an arbitrary reference starting

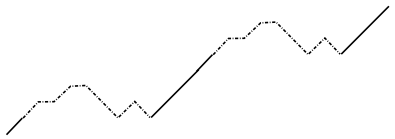

Figure 3.7: An example of a function $\tilde{W}$ with its subdivision of the graph into excursions (dotted part) and records (continuous part).

point $j \in \mathbb{Z}^*$ on the dual lattice. Define on $\mathbb{Z}^*$ the function $\tilde{W}_i$, $i \in \mathbb{Z}^*$, defined by

$$\tilde{W}_i := \begin{cases} \sum_{\ell \in \mathbb{Z} \cap [i,j]} (A_\ell - S_\ell) \,, & i \le j \,, \\ -\sum_{\ell \in \mathbb{Z} \cap [j,i]} (A_\ell - S_\ell) \,, & i > j \,. \end{cases}$$

By periodicity, we have $\tilde{W}_{i+N} = \tilde{W}_i + |\mathcal{S}| - |\mathcal{A}|$ and therefore $\lim_{i \to \pm\infty} \tilde{W}_i = \pm\infty$, since $|\mathcal{S}| > |\mathcal{A}|$. We extend the function $\tilde{W}$ to a function defined on the whole real axis by linear interpolation. The graph of $\tilde{W}$ is divided into records and excursions; a segment of the graph between the coordinates $i \in \mathbb{Z}^*$ and $i+1 \in \mathbb{Z}^*$ is a record if $\tilde{W}_i < \tilde{W}_k$ for any $k > i$ in $\mathbb{Z}^*$, otherwise it belongs to an excursion; see Figure 3.7 for an illustrative example. By periodicity, this classification is invariant under shifts of N, and we obtain a division of the edges of $\mathbb{Z}^*_N$, which are in bijection with the nodes of $\mathbb{Z}_N$, into edges belonging to excursions and to records. The values of (A_i, S_i) with $i \in \mathbb{Z}_N$ corresponding to an edge of the dual graph that is a record can be only $(0,1)$. Since $|\mathcal{A}| < |\mathcal{S}|$ and, for each excursion, we have the same number of values $(0,1)$ (corresponding to decreasing steps when moving from right to left in $\tilde{W}$) and $(1,0)$ (corresponding to increasing steps when moving from right to left in $\tilde{W}$), we obtain that there must exist at least one $j \in \mathbb{Z}_N$ such that $(A_j, S_j) = (0,1)$ and, moreover, that j corresponds to a record. We fix the coordinates on the torus in such a way that $j = N$. We consider the function $\tilde{W}$ constructed by taking as a starting reference point on the dual lattice the site $j - \frac{1}{2}$ which, after changing the reference coordinates, becomes $N - \frac{1}{2}$. We consider this function $\tilde{W}$ in the interval $\left[\frac{1}{2}, N - \frac{1}{2}\right]$ and define $W := \left[\tilde{W}\right]_+$, where $[\cdot]_+$ denotes the positive part. By the special choice of the starting point, the function W is simply obtained from $\tilde{W}$ by preserving the shape of excursions and transforming the records into constant pieces with value zero; in particular, $W_{\frac{1}{2}} = W_{N-\frac{1}{2}} = 0$. The function W represents the length of the queue associated with the arrivals and services determined, respectively, by $\mathcal{A}, \mathcal{S}$. starting the counting from the initial reference point. Apart from the $1/2$ shift, which is introduced for a simpler and clearer description, the graph of W has the same geometric features as the sample paths of ζ distributed according to $\rho_N^{0,0}$, and indeed we are going to show that they can be identified.

The identification of the sets $\mathcal{S}^\pm$ in the special reference frame we selected is particularly simple. The length of the queue starting from the reference point is given by W. If at site i the queue is non-zero, i.e. $W_i > 0$, and there is a service time available, i.e. $S_i = 1$, then one client in the queue is served at this time; this means that i is a used service time and

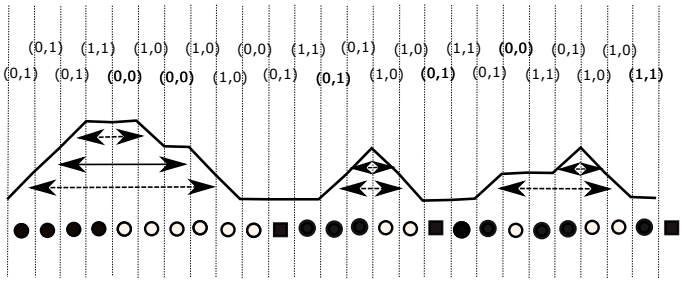

Figure 3.8: An example of the identification of used and unused services generating a configuration of particles of first and second class. In the top part of the picture, there are the values of the variables (A_i, S_i), while below them there is the graph of W associated with the waiting queue for the service. The horizontal arrows indicate the matching between clients and services according to the last arrived first served procedure. At the bottom, we draw the resulting particle configuration: round black dots represent first class particles, black squares represent second class particles and round white dots represent empty sites. Note that there is an extra second class particle at the far right, corresponding to the coordinate N.

$i \in \mathcal{S}^+$. Moreover, if the queue at i is zero but $(A_i, S_i) = (1, 1)$, then the service at i is also used and $i \in \mathcal{S}^+$ (the client arrives and is immediately served). Note that any i such that $(A_i, S_i) = (1, 1)$ belongs to an excursion. All the remaining $i \in \mathcal{S}$ are instead unused services and belong to $\mathcal{S}^-$. This construction corresponds to identifying the vertices in $\mathcal{S}^+$ with the elements of $\mathcal{S}$ that belong to excursions, while the vertices in $\mathcal{S}^-$ with the elements of $\mathcal{S}$ that belong to records. This is obtained by considering the unfair last arrived first served procedure, which yields the pairing inside each excursion illustrated in Figure 3.8.

Starting from the sets $\mathcal{A}$ and $\mathcal{S}$, we now construct the variables ζ. We first construct the unconditioned Markov chain with transition matrix (3.5.2), and then the variables η, which, conditioned on ζ, have a product distribution with marginals given by (3.5.7). To allow for a direct comparison with the construction in the general part, we identify, as before, the element of the dual lattice between the vertices i and $i + 1$ of the original lattice with i. We consider A_i and S_i that are independent i.i.d and uniformly distributed, taking values in $\{0, 1\}$ with the same probability.

We consider the walks $(s_i^1)_{i=1}^N$ defined by $s_1^1 = 0$ and $s_i^1 = \sum_{\ell=1}^{i-1} S_i$, $i > 1$, and $(s_i^2)_{i=1}^N$ defined by $s_1^2 = 0$ and $s_i^2 = \sum_{\ell=1}^{i-1} A_i$, $i > 1$. We then define $\xi_i = s_i^1 - s_i^2$ and finally $\zeta_i := |\xi_i|$. It follows that the variables ζ constructed in this way form a Markov chain with initial condition $\zeta_1 = 0$ and transition matrix (3.5.2). The variables η can now be defined as follows, in such a way that, conditioned on the ζ they have law (3.5.7)

$$\eta_i = S_{i+1}\mathbb{1}(\zeta_i > 0) + \mathbb{1}(\zeta_i = 0)\Big[2S_{i+1} - S_{i+1}A_{i+1}\Big], \qquad i = 1, \dots, N-1\,. \tag{3.5.8}$$

Considering subsets $\mathcal{S}, \mathcal{A}$ such that $\zeta_N = 0$, we see that formula (3.5.8) and the queue construction obtained from the same subsets identify exactly the same splitting $\mathcal{S} = \mathcal{S}^+ \cup \mathcal{S}^-$. This

implies that, when $\mathcal{A}, \mathcal{S}$ have the correct conditional distribution, the mixture construction and the queue construction identify the same configuration of particles. We do not discuss further details, the equivalence is even simpler and more direct in the infinite setting of $\mathbb{Z}$.

The last model we study is the harmonic model. As we shall see, the construction in this section will be different from the previous ones. We start by considering the invariant measure of the process represented as a mixture of product of inhomogeneous geometric distribution [11] and we show how it can be interpreted as a probability measure of MPA type, where the matrices are replaced by operators determined by continuous kernels.

3.6 Harmonic models

As introduced in Chapter 1 the harmonic model belongs to the broad class of processes in which particles may jump together with a rate depending on the number of particles. We study the simplest case: we consider a one-dimensional lattice of N sites, a block of k particles can jump from site x to site $x+1$ or $x-1$ with rate $1/k$. A block of k particles is injected at site 1, and site N, with rate α^k/k, and β^k/k, respectively, and is removed from site 1, and site N, with rate $1/k$ (for more details, see Section 1.4).

Let $\mathcal{G}_m(k) = \frac{1}{1+m}\left(\frac{m}{1+m}\right)^k$, $k = 0, 1, \dots$, be a geometric distribution of mean m. Given $\underline{m} = (m_1, \dots, m_N)$ and $\underline{k} = (k_1, \dots, k_N)$ we denote by $\mathcal{G}_{\underline{m}}(\underline{k}) := \prod_{i=1}^N \mathcal{G}_{m_i}(k_i)$. Given $0 < \alpha < \beta < 1$ we call $m_L := \frac{\alpha}{1-\alpha} < \frac{\beta}{1-\beta} := m_R$ and introduce $O_N^{m_L,m_R} \subseteq [m_L, m_R]^N$ as the set defined by

$$O_N^{m_L,m_R} := \{\underline{m} \,:\, m_L \le m_1 \le \cdots \le m_N \le m_R\}\,,$$

whose Lebesgue volume is given by $|O_N^{m_L,m_R}| = \frac{(m_L-m_R)^N}{N!}$. Inspired by [7, 12], in [11] it has been proved that the invariant measure of this model can be represented as

$$\mu_N^{m_L,m_R}(\eta) = \frac{1}{|O_N^{m_L,m_R}|}\int_{O_N^{m_L,m_R}} d\underline{m}\ \mathcal{G}_{\underline{m}}(\eta)\,. \tag{3.6.1}$$

For convenience we make explicit the dependence of the invariant measure on the parameters m_L, m_R, N.

Formula (3.6.1) represents the invariant measures as a mixture of product of inhomogeneous geometric distributions. We now show that formula (3.6.1) can be interpreted as a probability measure of the MPA type (2.2.1) but with operators $(M^k)_{k\in\mathbb{N}_0}$ determined by continuous kernels $\left(M^k_{m,m'}\right)^{k\in\mathbb{N}_0}_{m,m'\in\mathbb{R}^+}$. Using our equivalence, we then show that (3.6.1) is a particular case of (3.1.6) for a special Markov bridge and special marginal distributions (3.1.7). We note that. recently, a representation of (3.6.1) as a measure of MPA type with matrices has been obtained in [40]; it would be interesting to explore the corresponding mixture representation that should be different from (3.6.1) (recall Remark 3.7 on non uniqueness).

Let us define the kernels

$$M^k_{m,m'} := \mathcal{G}_m(k)\mathbb{1}(m \le m')\,, \qquad k \in \mathbb{N}_0\,,\ m, m' \in \mathbb{R}_+\,, \tag{3.6.2}$$

which define operators in a functional space that, for simplicity, we do not describe formally. We use the bra-ket formalism of quantum mechanics to denote elements of the space on which

the operators act. For a state $|f\rangle$ associated with a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ we have

$$\begin{cases} M^k |f\rangle = \int_0^{+\infty} dm' \, M^k_{m,m'} f(m') = \mathcal{G}_m(k) \int_m^{+\infty} f(m') d\, m' \,, \\ \langle f| \, M^k = \int_0^{+\infty} dm \, M^k_{m,m'} f(m) = \int_0^{m'} \mathcal{G}_m(k) f(m) d\, m \,. \end{cases} \tag{3.6.3}$$

We also introduce delta states $|z\rangle$ associated with elements $z \in \mathbb{R}_+$ and the action of the operators is defined in this case as

$$\begin{cases} M^k |z\rangle = \int_0^{+\infty} dm' \, M^k_{m,m'} \delta(m' - z) = \mathcal{G}_m(k) \mathbb{1}(m \le z) \,, \\ \langle z| \, M^k = \int_0^{+\infty} dm \, M^k_{m,m'} \delta(m - z) = \mathcal{G}_z(k) \mathbb{1}(z \le m') \,. \end{cases} \tag{3.6.4}$$

Using these operators we can introduce the family of probability measures

$$\hat{\mu}^{m_L, m_R}_{N+1}(\eta) := \frac{\langle m_L| \prod_{i=0}^N M^{\eta_i} |m_R\rangle}{Z_N} \,, \tag{3.6.5}$$

where the states $\langle m|_L$ and $|m\rangle_R$ are delta states associated with the values $m_L, m_R \in \mathbb{R}_+$. The index $N+1$ is due to the fact that in (3.6.5) we have $\eta = (\eta_0, \eta_1, \dots, \eta_N)$ which contains the extra variable η_0 that does not belong to our particle system; however, we have the relation

$$\mu^{m_L, m_R}_N(\eta_1, \dots, \eta_N) = \sum_{\eta_0} \hat{\mu}^{m_L, m_R}_{N+1}(\eta_0, \dots, \eta_N) \,.$$

The operator $M = \sum_{k \in \mathbb{N}_0} M^k$ has a continuous family of positive eigenvectors; consider the family of functions $\Big(e_\lambda(m) = \lambda e^{-\lambda m} \,; m \in \mathbb{R}_+\Big)_{\lambda > 0}$, then we have

$$M e_\lambda = \int_m^{+\infty} \lambda e^{-\lambda m'} \, dm' = e^{-\lambda m} = \frac{e_\lambda}{\lambda} \,.$$

For any $\lambda > 0$ we can therefore apply the transformation (3.2.1), obtaining the transition kernel

$$P^\lambda_{m,m'} = \lambda e^{-\lambda(m'-m)} \mathbb{1}(m' - m \ge 0) \, d\, m' \,. \tag{3.6.6}$$

Since the states in (3.6.5) are delta states, the process $(\zeta_0, \dots, \zeta_{N+1})$ is a Markov bridge with transition probabilities given in (3.6.6) and pinned at initial and final time at $\zeta_0 = m_L$ and $\zeta_{N+1} = m_R$. The Markov bridge has length $N+2$ since in (3.6.5) there is an extra variable η_0. Therefore, the random variables $(\zeta_0, \dots, \zeta_{N+1})$ are distributed as the random walk $\zeta_n = \zeta_0 + \sum_{i=1}^n \gamma_i$, where the $(\gamma_i)_{i \in \mathbb{N}}$ are i.i.d. exponential random variables of parameter $\lambda > 0$. For different values of λ we obtain different Markov processes, all having the same conditioned Markov bridge, which indeed coincides with the order statistics law of uniform random variable on the interval $[m_L, m_R]$. A special feature of these Markov processes is that they are increasing, so that the Markov property is equivalent to the spatial Markov property of the point process $\{\zeta_1, \dots, \zeta_N\} \subseteq [m_L, m_R]$ discussed in [10, 54]; this is not true in the general case.

By a simple direct computation we obtain that the marginals (3.2.4) of the mixture are given by $p_{m,m'}(k) = \mathcal{G}_m(k)$; also in this case this model has a special feature that simplifies

both the computation and the representation, namely the fact that the marginal of the mixture depends only on m and not on m'.

A similar generalized construction to the one described above can be carried out for the whole class of harmonic models, whose representation of the invariant measures as a mixture has been obtained in [10].

Remark 3.12 (The general case and KMP processes)**.** *We mentioned in the Introduction that recently in [11, 18] and in [10, 46, 54] have been found representations of the invariant measure of boundary driven generalized KMP and zero range models, respectively, in terms of mixture of inhomogeneous product measures. The general case is more involved than the one described in this section: the marginals of the inhomogeneous product measure depend on two values of the Markovian parameter hidden variables while for the simplest model the dependence reduces to just one. Moreover, in this case the Markov bridges are associated to increasing processes so that the Markov property implies the spatial Markov property of [54]. In the case of the KMP and generalized KMP processes it is not clear if the hidden variables that determine the parameters of the inhomogeneous products are Markovian, so that is not clear if the invariant measures are of MPA type.*

3.7 Miscellany

In the finite dimensional case, under irreducibility conditions, there is a unique stochastic matrix P associated with M by (3.2.1). This means that, disregarding for the moment the vectors x, y, the whole class of probability measures of MPA type (2.2.1) can be parametrized by a stochastic matrix P and by $|B|^2$ elements of $\mathcal{M}^1(A)$, which determine the marginals $p_{b,b'}(\cdot)$, $b, b' \in B$, of the mixture. This is a smaller family with respect to the one obtained from the family of $|A|$ non-negative $B \times B$ matrices M^a. The mixture representation (3.1.6) can therefore be interpreted as a canonical form of MPA-type measures; given P and $\big(p_{b,b'}\big)_{b,b' \in B}$, there is a whole family of matrices $(M^a)_{a \in A}$ associated with them, consisting of all matrices of the form

$$M^a_{b,b'} = p_{b.b'} \lambda \, e_b P_{b,b'} e_{b'}^{-1} \,, \;\; \lambda \in \mathbb{R}_+ \,, \;\; e \in \mathbb{R}_+^B \,, \qquad a \in A \,, \, b, b' \in B \,, \tag{3.7.1}$$

where $\lambda \in \mathbb{R}_+$ and $e \in \mathbb{R}_+^B$ (which can be fixed so that $\sum_b e_b = 1$) are arbitrary parameters.

Relation (3.7.1) identifies equivalence classes of families of matrices $(M^a)_{a \in A}$ all corresponding to the same mixture. Note that this is not equivalent to the problem of different representations of matrices satisfying, for example, the algebraic relations (**??**). In that case, one deals with infinite dimensional representations and moreover the matrices solving (**??**) are not invariant by an arbitrary scaling factor, as is instead the case of (3.7.1).

The Markov bridge measure is also determined by the measure m in (3.1.5), which depends on the vectors f, g themselves related to the vectors x, y through the matrices $\mathcal{E}$, $\mathcal{E}^{-1}$. It can be shown that, for a given P, not all the probability measures can be written in the form (3.1.5), but any pair of one-dimensional marginals can indeed be obtained, for large enough N and $|B| < +\infty$. The vectors f, g can be fixed by imposing the normalization $\sum_b f_b = \sum_b g_b = 1$. We do not discuss this topic further.

We say that μ_N of MPA type is stationary when $\mu_N(\eta_i) = \mu_N(\eta_j)$ for any i, j. In this case, the measure can be extended to a measure on infinite sequences. Using the mixture

representation, this occurs when the Markov bridge is stationary, which is the case when $g_b = \pi_b$ and $f_b = 1$, $b \in B$, where π is the invariant measure of P. For example, when $|A| = |B| = 2$, the ζ process is a stationary $\{0,1\}$-valued Markov chain, and the η process is constructed, using four numbers $0 < p_{b,b'} < 1$, starting from a sample path of ζ and tossing independent coins of parameter $p_{b,b'}$ at each occurrence of the pair b, b' in ζ. Among these processes, there are the special cases described in Example 5 of Section 3.1 of [35]. Differently from [35], where one needs to suitably enlarge the alphabets in order to describe the processes as functions of Markov processes, the description in terms of mixture keeps the original sizes of A, B.

Still in the stationary case, we consider the *one dependent* measures of Example 4 in Section 3.1 of [35]. This corresponds to measures for which η_i^{j-1} and η_{j+1}^k are independent for any $i < j < k$, and hence to having $\sum_{\eta_j} \mu_N(\eta) = \mu_N(\eta_1^{j-1})\mu_N(\eta_{j+1}^N)$.

As explained in [35], a family of matrices $(M^a)_{a\in A}$ and vectors x, y, provides a stationary one dependent measure when $M = cxy^T$ is a rank-one matrix for an arbitrary constant $c > 0$. In this case, the Perron eigenvector is x with corresponding eigenvalue $cy^T x$, and the stochastic matrix (3.2.1) is given by $P_{b,b'} = \frac{y_{b'} x_{b'}}{y^T x}$, which corresponds to an i.i.d. process. We then obtain a simple characterization of one dependent measures that are also manifestly positive algebraic measures, in the terminology of [35]. These are measures obtained by a double independent construction, parametrized by a $q \in \mathcal{M}^1(B)$ and a family $p_{b,b'} \in \mathcal{M}^1(A)$, $b, b' \in B$. First, one constructs a sample path of ζ using *i.i.d.* variables distributed according to q; conditioned on this sample path, one then generates independent variables η such that η_i is distributed as $p_{\zeta_i,\zeta_{i+1}}$. The result is a one dependent stationary measure.

Chapter 4

Large Deviations for MPA measures

In the previous sections we have explored the probabilistic structure of the Matrix Product Ansatz, providing a characterization of MPA measures in terms of mixtures of inhomogeneous product measure where the mixing law is a Markov bridge. In particular, in the proof of this result, (see Theorem 3.6, item (2)) we have shown that such measures can also be represented as a Markov bridge on the enlarged state space.

In this chapter, we focus on the Markov bridge representation on the enlarged space to study large deviations for MPA measures. We remark that measures of this type also appear in different context under the name of rational models, introduced in the area of theoretical informatics in connection with the analysis of pattern statistics (see for example [3, 47, 48] and references therein).

A rational model is a probability measure on symbolic sequences defined by using a collection of matrices with non-negative elements and two vectors with non-negative coordinates; this corresponds exactly to the abstract formulation given in the previous chapters of matrix product ansatz type measures.

In the following sections, we prove some large deviations principles for MPA measures. In particular we will introduce several notions of *Empirical Measures* both algebraically defined in terms of frequencies, and spatially defined considering the sequence rescaled and embedded on a unit interval. This second definition is natural considering the symbolic sequence as a configuration of particles and having in mind a scaling limit. The phenomenology of this class of measures can be very different depending on several features of the matrices and of the vectors. We will concentrate on the major distinction when the matrices are finite or infinite. For the finite case we will just consider the simpler case of irreducible and aperiodic matrices, while for the infinite case we will consider a remarkable example, the totally asymmetric exclusion process.

We will see that, while in the finite case (at least under the assumptions on the matrices) we have a general compact variational representation for the rate functional of both algebraic and spatial empirical measures, in the infinite case we may have different behaviors. This is the counterpart of the fact that irreducible finite Markov chains are recurrent while in the infinite different behaviors may occur. The specific infinite example that we discuss has not a large deviations principle for the algebraic empirical measures while that for the spatial empirical measure has a non local structure.

As anticipated, the strategy of the proof is based on the previous general construction that any MPA measure can be represented as an *Hidden Markov measure* on a suitable enlarged state space. Consequently, we can deduce the large deviations results by contraction from the corresponding ones for the enlarged Markov chain.

We point out that the rate functionals obtained by this procedure will always have a variational representation as an infimum over rate functionals for enlarged Markov measures.

Finally, we stress that the additional structure of these models - mixture of inhomogeneous product measure enlightened in the previous chapter - gives an additional structure to the rate functional. In particular, it can be represented by variational problems over a double entropy functional. This approach is discussed in the last section of the chapter in the simplest setting; a more general analysis will be given in [45].

The results presented in the following sections are contained in [44].

The chapter is organized as follows.

Section 1 is dedicated to some preliminaries on large deviations theory which will be useful in what follows.

In Section 2 we introduce notation and define the empirical measures. For completeness, we also include a short subsection on rational models, where we briefly recall their definition and explain their relation to MPA measures.

In Section 3 we discuss some details we postponed in the previous section, precisely in Theorem 3.6, item (2), and we prove a general large deviations Theorem for the finite case with a variational representation of the rate functional both for the algebraic and the spatial empirical measures. We also illustrate some special cases when the variational formula can be expressed explicitly.

In Section 4 we apply the ideas developed in the finite setting to the boundary driven TASEP recovering the results in [26, 28] and in particular getting a representation of the rate functional as an infimum similar to that in [6].

Finally, in Section 5 we briefly discuss the simplified representation of the rate functional as a double entropy variational problem, obtained by exploiting the mixture representation of MPA measures in the finite and irreducible setting.

4.1 Preliminaries on Large Deviations Theory

In this section we introduce and develop the theoretical and mathematical framework of large deviation theory, which describes how the probability of rare fluctuations behaves in systems with many interacting or independent components. Large deviation theory can be seen as a bridge between the microscopic fluctuations and macroscopic laws, providing a probabilistic point of view to the study of collective behavior.

The general background presented here will be used as a basis for the analysis carried out in the following sections. To make the exposition self-contained, we first recall the physical motivation and historical background that connect large deviations to statistical mechanics, before turning to the formal definitions and main theorems that will be applied later to the study of large deviations for the MPA measures.

4.1.1 Physical background and motivation

The mathematical theory of large deviations was first developed in the 1930s by Cramér in the context of probability theory, later extended by Donsker and Varadhan to Markov processes and generalized by Freidlin and Wentzell to stochastic differential equations in the 1970s.

Originally conceived in a mathematical context, it developed independently of the questions that motivated statistical physics. In reality, its fundamental ideas had long been applied in physics, often implicitly, through the use of entropy, free energy, and variational principles to describe the collective behavior of many-particle systems. Therefore, over time, large de-

viation theory has been considered the natural mathematical language of statistical mechanics, providing a rigorous probabilistic framework that connects microscopic fluctuations with macroscopic thermodynamic behavior, both at and out of equilibrium (for a general review, see [59]; for a discussion focused on non-equilibrium systems, see [60]).

The roots of the large deviation theory in physics can be traced back to Boltzmann and Einstein. In equilibrium statistical mechanics, a well-defined relationship established by Boltzmann connects the probability of a state with its entropy, establishing a probabilistic interpretation of thermodynamic quantities. Einstein later extended this idea to describe the probability of observing fluctuations of macroscopic observables away from their equilibrium values. He proposed that the probability of observing a macroscopic deviation from equilibrium decreases exponentially with the "entropic cost" of that fluctuation. This idea anticipates the large deviation principle: given a macroscopic observable $M_N := M_N(\eta)$, depending on microscopic configurations η, probabilities of atypical events decay exponentially with the system size N, according to:

$$P(M_N \approx m) \sim e^{-NI(m)}$$

where $I(m)$ is a positive convex function, called *rate function*, that measures the "cost" of the fluctuation from the equilibrium.

At equilibrium, the Gibbs distribution provides a complete statistical description of the system, thus giving an explicit connection between probability and thermodynamics.

Out of equilibrium, conversely, the situation changes deeply and becomes more complex: the stationary states of (irreversible) driven systems, such as those in contact with two reservoirs, cannot be described by a Gibbs measure. Their invariant measure is not known in a closed form a priori, but it is determined by the stochastic microscopic dynamics.
It follows that there is no universal relation linking probability and entropy. Even if local equilibrium allows one to define local thermodynamic variables, such as density or temperature, global thermodynamic potentials are not uniquely defined. Adopting the Boltzmann-Einstein viewpoint, a natural way to generalize thermodynamic potentials to the non-equilibrium setting is to define them through fluctuation theory. Large deviation principles provide precisely this probabilistic structure, quantifying both typical behavior and rare events.

This perspective is formalized by the Macroscopic Fluctuation Theory (MFT) ([7, 8]), which formulates non-equilibrium statistical mechanics in terms of large deviation functionals for density and current fields. The theory is inspired by stochastic lattice gases and is based on the concept of the hydrodynamic limit, which describes how deterministic macroscopic laws emerge from stochastic microscopic dynamics. Complementary to this macroscopic perspective, large deviations can, in some cases, be derived explicitly from the microscopic dynamics, in particular for exactly solvable models. This approach has been successfully applied to exclusion processes, where the stationary measure admits a matrix-product representation and the rate function for the density profile can be determined explicitly. Starting from the diffusive symmetric case [27], later extended to the driven asymmetric case [28], these results provide microscopic realizations of non-equilibrium large deviation principles. In the next sections we follow the microscopic perspective; before presenting our results, we recall some fundamental mathematical notions of large deviation theory that will be used in the following.

4.1.2 Elements of large deviation theory

In this section, we recall the general formalism and the main results of large deviation theory; unless otherwise specified, the exposition follows [19].

We begin with the first fundamental result in the field, which, as mentioned in the background, goes back to Cramér [9].

As discussed in the previous section, the central idea of large deviation theory is that probabilities of rare fluctuations decay exponentially with the system size, with a rate quantified by a suitable function I. We now give a more precise formulation of this principle.

Let $\{A_n\}$ be a sequence of random variables taking values in a measurable space E and define $\mu_n(B) := \mathbb{P}(A_n \in B)$ with $B \subset E$ measurable. We say that the sequence $\{A_n\}$ satisfies a *Large Deviation Principle (LDP)* with speed n and rate function I if

- for every closed set $F \subset E$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x\in F} I(x), \tag{4.1.1}$$

- for every open set $G \subset E$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x\in G} I(x). \tag{4.1.2}$$

The definition above admits an equivalent and more general formulation directly in terms of sequences of probability measures on E. In what follows, we will adopt either the language of random variables or that of probability measures, depending on the context.

The function $I(x) \geq 0$ determines the exponential rate at which the probability of atypical values decays as n increases. Typical values correspond to the minima of $I(x)$, while large values of $I(x)$ describe exponentially rare events.

In his work, Cramér established a large deviation principle for the *empirical average*:

$$\frac{1}{n} S_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

where $X_1, ..., X_n$ are independent and identically distributed random variables. This theorem provides the starting point for more general formulations developed later in the theory.

Theorem (Cramèr). Let $\{X_i\}_{i=1}^n$ be i.i.d. real random variables such that the moment generating function satisfies:

$$\varphi(\lambda) = \mathbb{E}[e^{\lambda X_1}] < +\infty \qquad \forall\ \lambda \in \mathbb{R} \tag{4.1.3}$$

Then the sequence $\{S_n/n\}_{n\geq 1}$ satisfies a large deviation principle with speed n and rate function:

$$I(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \log \varphi(\lambda)\}.$$

The rate function $I(x)$ is convex, lower semicontinuous, and satisfies $I(x) = 0$ if and only if $x = \mathbb{E}[X_1]$.

It follows that a large deviation principle for the empirical mean S_n/n can be derived directly by computing the cumulant generating function of a single random variable and then applying a Legendre–Fenchel transform. This is a mechanism that provides the basic example for many results in large deviation theory, in which the rate function admits a variational Legendre-Fenchel representation.

In the i.i.d. setting, a natural extension of Cramér's theorem concerns the empirical distribution of the variables themselves, rather than their average. This was established by Sanov, who derived a large deviation principle for *empirical measures* of i.i.d. sequences.

Theorem (Sanov). Let $\{X_i\}_{i\geq 1}$ be random variables taking values in a finite set $\Gamma = \{1, ..., r\} \subset \mathbb{N}$, independent and identically distributed with marginal law $\mu = (\mu(s))_{s\in\Gamma}$ such that $\mu(s) > 0 \ \ \forall s \in \Gamma$. We define the empirical measure

$$L_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i} \ \in \ \mathcal{M}^1(\Gamma)$$

where δ_x is the point-mass at $x \in \mathbb{R}$. Then the sequence $\{L_n\}_{n\geq 1}$ satisfies a large deviation principle with speed n and rate function

$$I_\rho(\nu) = \sum_{s\in\Gamma} \nu(s) \log\left(\frac{\nu(s)}{\mu(s)}\right).$$

This result concerns the empirical measure L_n that is the random probability measure that assigns to each possible outcome $x \in \Gamma$ the relative frequency with which x appears in the sample $(X_1, ..., X_n)$. It states that the probability that L_n deviates significantly from the distribution $\mu = (\mu(1), ..., \mu(r))$ decays exponentially with the system size, with a rate quantified by the *relative entropy* between ν and μ.

Sanov's theorem thus extends Cramér's result from empirical means to empirical distributions, describing the fluctuation of the entire empirical measure and providing a complete statistical description of i.i.d. sequences.

The theorem given above refers to the discrete case where Γ is finite and $\mu(s) > 0$ for all $s \in \Gamma$, but it can be generalized much further. It can be extended to $\mathcal{M}^1(\mathbb{R})$, or more generally to any space of probability measure defined on a Polish space $\mathcal{X}$. In this general setting Sanov's theorem still holds with the same structure, and the rate function is expressed in terms of the relative entropy

$$H(\nu|\mu) = \int_{\mathcal{X}} \log \frac{d\nu}{d\mu} d\nu$$

if ν is absolutely continuous with respect to μ and $H(\nu|\mu) = +\infty$ otherwise.

Having established large deviation principles for independent random variables, one may ask how such principles behave under transformations of these random quantities. This connection is formalized by the *contraction principle*, which shows that if an LDP holds for a sequence of random variables, then it also holds for any of their continuous images, with

a suitably modified rate function. In particular, Cramér's theorem can be recovered from Sanov's by an application of this principle.

Before going on, we recall that the definition of a large deviation principle given in (4.1.1)-(4.1.2) can equivalently be formulated in terms of sequences of probability measures.

Contraction principle. Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces and $f : \mathcal{X} \to \mathcal{Y}$ a continuous function. Assume that a LDP holds for a family of probability measures $\{\mu_n\}$ defined on $\mathcal{X}$ with speed r_n and rate function I. Let $\nu_n = \mu_n \circ f^{-1}$ be the *pushforward* measure of μ_n induced by f. For each $y \in \mathcal{Y}$, define

$$I'(y) := \inf\{I(x) : x \in \mathcal{X},\ y = f(x)\}.$$

Then the family of probability measures $\{\nu_n\}$ satisfies an LDP on $\mathcal{Y}$ with speed r_n and rate function I'. (For more details see [31], Sec. 4.2).

This result expresses that the large deviation principle is preserved under continuous mappings. The contraction principle is widely used in statistical mechanics, in particular if the microscopic variables of a system satisfy an LDP, then any macroscopic observable obtained from them through a continuous transformation, also satisfies an LDP. Moreover, it is a very powerful calculation device, indeed it provides a practical way to compute the rate function of a transformed sequence from a known one, through a variational minimization of the original rate function over the pre-image of the transformation.

This principle will play a central role later in the derivation of large deviation results for matrix-product measures.

Another useful tool in large deviation theory is the *Varadhan's lemma* which allows one to compute the asymptotic behavior of exponential integrals with respect to a sequence of measures which satisfy a large deviation principle.

Varadhan's lemma. Let $\{\mu_n\}$ be a sequence of probability measures on a Polish space $\mathcal{X}$ satisfying an LDP with speed n and rate function I. Let $F : \mathcal{X} \to \mathbb{R}$ be a continuous function that is bounded from above. Then

$$\lim_{n\to+\infty} \frac{1}{n} \log \int_{\mathcal{X}} e^{nF(x)} \mu_n(dx) = \sup_{x\in\mathcal{X}} [F(x) - I(x)].$$

This result will be used later, in Section 4.5, to justify the passage between large deviation functionals associated with different, but related, probability measures.

A more general formulation of Cramér's theorem, valid not only in the i.i.d. setting, is provided by the Gärtner-Ellis theorem. It defines a large deviation principle for more general sequence of random variables taking values in $\mathbb{R}^d$, under suitable regularity assumptions on their cumulant generating function. The rate function is again obtained through a Legendre-Fenchel transform, as in Cramér's theorem. We mention it here for completeness, but we do not discuss it since it will not be needed in what follows.

While the previous results concern large deviations of empirical means or measures, in

many problems the quantity of interest is not a single observable but the evolution of an entire trajectory. In these terms, a further extension of Cramér's result, from empirical means to entire trajectories of the partial-sum process, is provided by Mogulskii's theorem (see [31], Thm. 5.1.2). It defines a functional large deviation principle on the space of paths.

For our purposes, and for simplicity, we state the theorem in $\mathbb{R}$; its general formulation holds in $\mathbb{R}^d$.

Theorem (Mogulskii) Let $\{X_i\}_{i\geq 0}$ be real i.i.d. random variables satisfying the Cramér's condition (4.1.3). Define:

$$S_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} X_i, \qquad t \in [0,1]$$

and let μ_n be the distribution of $S_n(\cdot)$ in $L^\infty([0,1])$. Then, the measure μ_n satisfies the large deviation principle with speed n and rate function:

$$I(\varphi) = \begin{cases} \displaystyle\int_0^1 I_C(\dot{\varphi}(t))\,dt, & \text{if } \varphi \in \mathcal{AC} \text{ and } \varphi(0) = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

where I_C is the Cramér's rate function for the variables X_i.

This result provides a functional large deviation principle, showing that the probability of observing an entire trajectory φ decays exponentially with a rate determined by the function I_C computed along its derivative.

Although Mogulskii's theorem is often presented as a pathwise extension of Cramér's theorem, it can equivalently be derived from Sanov's theorem combined with the contraction principle, viewing the path as a functional of the empirical measure of the increments.

We conclude this section with the formulation of the Donsker–Varadhan theorem for discrete-time Markov chains. In this setting, the most natural object to study is the *pair empirical measure*, which records the frequency of transitions observed along the chain.

Theorem (LDP for discrete-time Markov chain). Let $X_1, ..., X_n$ be a Markov chain taking value in $\Gamma = \{1, ..., r\} \in \mathbb{N}$ and described by the transition matrix $P = (P_{st})_{s,t\in\Gamma}$ with $P_{st} > 0$ for all $s, t \in \Gamma$. We define the *pair empirical measure*:

$$L_n^2 = \frac{1}{n}\sum_{i=1}^{n} \delta_{(X_i, X_{i+1})}$$

with periodic boundary conditions $X_{n+1} = X_1$. Then, the sequence (L_n^2) satisfies a large deviation principle on $\mathcal{M}^1(\Gamma^2)$ with speed n and rate function:

$$I^2(\nu) = \sum_{s,t} \nu_{s,t} \log\left(\frac{\nu_{s,t}}{\bar{\nu}_s P_{s,t}}\right)$$

where $\bar{\nu} = \sum_t \nu_{s,t}$.

We note that there exist extended versions of this theorem: the large deviation principle holds under weaker assumptions, for example under the irreducibility of P (see [31], Sec. 6.5.2).

The preliminaries recalled in this section provide the probabilistic and variational background needed for the study of large deviations in the remainder of the chapter. They offer the theoretical framework upon which the following results are built.

4.2 Rational models and definitions

In this and the following sections, we will use the same notation as in the previous chapters, introduced in Section 2.2. We first recall and introduce a few notions that will be needed in what follows. Then we provide a more general definition of empirical measures, discussed only briefly in the previous section. Finally, for completeness, we devote a short subsection to the definition of rational models.

Given A, B finite or countable sets, we denote by $A^* = \cup_{N=1}^{+\infty} A^N$ the set of non empty finite words over the alphabet A. If $\eta \in A^*$ is a word with symbols written in the alphabet A, with $\eta \in A^N$, we write $|\eta| = N$. Moreover, as explained in Section 2.2, we denote by $\eta_n^m := (\eta_n, \eta_{n+1}, \dots \eta_m)$ the finite portion of the string η between the indices n and m when $n < m \leq |\eta|$. In some constructions, terms of the form η_i with $i > |\eta| = N$ may appear; in this case we mean that the index of the position of the letter in the string is taken modulus $|\eta|$, i.e. $\eta_{i+N} = \eta_i$ for any $i \in \mathbb{Z}$.

Consider M a $B \times B$ square matrix. We say that M is *non-negative* if $M_{b,b'} \geq 0$ for any $b, b' \in B$. We associate a directed graph (B, E) to the matrix M considering $(b, b') \in E$ if and only if $M_{b,b'} > 0$. The matrix M is called *irreducible* if (B, E) is strongly connected. Given a state $b \in B$ we call the period of b as the greatest common divisor of the set $\{k \in \mathbb{N} : M_{b,b}^k > 0\}$. For irreducible matrices the period of each state is the same. We call *aperiodic* a matrix having the period of each state equal to 1.

We call respectively $\mathcal{M}^1(S)$ and $\mathcal{M}^+(S)$ the set of probability measures and positive measures on a given set S endowed with the topology induced by the weak convergence, i.e. $\mu_n \to \mu$ when $\int f d\mu_n \to \int f d\mu$ for any continuous and bounded $f : S \to \mathbb{R}$.

Given $\nu \in \mathcal{M}^1(A^k)$ we say that the measure satisfies a *finite stationary condition* if

$$\sum_{a \in A} \nu(\eta a) = \sum_{a \in A} \nu(a\eta)\,, \qquad \forall \eta \in A^{k-1}\,, \tag{4.2.1}$$

and we call $\mathcal{M}_{st}^1(A^k)$ the set of such probability measures.

Consider the real interval $[0, 1] \subseteq \mathbb{R}$. We call a function $\pi : [0, 1] \to \mathbb{R}$ *absolutely continuous* if there exists a function $\dot{\pi} \in L^1([0, 1])$ such that for almost all $x \in [0, 1]$ we have $\pi(x) = \int_0^x \dot{\pi}(y) dy$. We call $\mathcal{AC}([0, 1])$ the set of absolutely continuous functions.

4.2.1 Empirical measures: algebraic and spatial

We now introduce two complementary notions of empirical measures. The first one is algebraic, defined in terms of frequencies, while the second one is spatial, obtained by rescaling the sequence on a unit interval.

The set A will always be finite while B may be either finite or countably infinite. We assume, for simplicity, that in the finite case $A := \{0, 1, \dots\ |A| - 1\}$ and $B := \{0, 1, \dots, |B| - 1\}$ while instead in the infinite countable case we have $B = \mathbb{N} \cup \{0\} =: \mathbb{N}_0$.

Given $\eta \in A^*$ we define its *algebraic empirical measure* of order one as

$$\hat{\nu}^1 = \hat{\nu}^1[\eta] := \frac{1}{|\eta|} \sum_{i=1}^{|\eta|} \delta_{\eta_i} \in \mathcal{M}^1(A)\,, \tag{4.2.2}$$

where δ_a is the delta measure at $a \in A$. When $\eta \in A^N$ so that $|\eta| = N$ we call $\hat{\nu}^1_N$ the empirical measure (4.2.2). The empirical measures are constructed starting from a sample, that is η in this case, and later on it will be a pair (η, ζ); we denote such dependence on the sample using squared parenthesis.

For any positive integer k we define the *algebraic empirical measure of order k* as

$$\hat{\nu}^k = \hat{\nu}^k[\eta] := \frac{1}{|\eta|} \sum_{i=1}^{|\eta|} \delta_{\eta_i^{i+k-1}} \in \mathcal{M}^1_{st}(A^k)\,, \tag{4.2.3}$$

that, by definition, satisfies the finite version of the stationary condition (4.2.1); again when $|\eta| = N$ we use the notation $\hat{\nu}^k_N$ for (4.2.3).

For any fixed η, we have that $\left(\hat{\nu}^k\right)_{k\in\mathbb{N}}$ is a sequence of compatible probability measures, i.e.

$$\sum_{a\in A} \hat{\nu}^{k+1}(\xi a) = \hat{\nu}^k(\xi)\,, \qquad \forall \xi \in A^k\,,\ \forall k \geq 1\,, \tag{4.2.4}$$

and, by Kolmogorov theorem, are therefore the marginals of a shift invariant measure on $\mathcal{M}^1(A^{\mathbb{N}})$ called the stationary process (see for example [19] for details and notice that you do not need to have $k < |\eta|$).

We now introduce some natural empirical measures having a spatial structure. Consider the interval $[0, 1]$ as reference geometric space. Given $\eta \in A^*$, we define the corresponding *spatial empirical measure* as

$$\hat{\pi} = \hat{\pi}[\eta] := \frac{1}{|\eta|} \sum_{i=1}^{|\eta|} \eta_i \delta_{i/|\eta|}\,. \tag{4.2.5}$$

We have that $\hat{\pi} \in \mathcal{M}^+([0, 1])$ is a positive measure on the interval $[0, 1]$. When $|\eta| = N$ we call $\hat{\pi}_N$ the spatial empirical measure (4.2.5). Equivalently we can define the empirical measure by its action on continuous functions $f : [0, 1] \to \mathbb{R}$

$$\int_{[0,1]} f\, d\hat{\pi}_N = \frac{1}{N} \sum_{i=1}^{N} \eta_i f(i/N)\,. \tag{4.2.6}$$

We introduce also a *generalized spatial empirical measure* defined by

$$\hat{\Pi}^k = \hat{\Pi}^k[\eta] := \frac{1}{|\eta|} \sum_{i=1}^{|\eta|} \delta_{\eta_i^{i+k-1}} \delta_{i/|\eta|}(dx)\,, \qquad k = 1, 2, \dots \tag{4.2.7}$$

where we added the symbol dx to distinguish $\delta_y(dx)$, the spatial delta measure at $y \in [0,1]$, from δ_η, the element of $\mathcal{M}^1(A^k)$ concentrated on η such that $|\eta| = k$. We have that $\hat{\Pi}^k \in \mathcal{M}([0,1];\mathcal{M}^+(A^k))$, i.e. it is a measure valued measure on the integral $[0,1]$, that can be equivalently characterized by its action on continuous vector fields. Consider $\mathfrak{f} = (\mathfrak{f}_\zeta)_{\zeta \in A^k}$ where for any ζ, we have that $\mathfrak{f}_\zeta : [0,1] \to \mathbb{R}$ is a continuous function. We have then

$$\hat{\Pi}^k[\eta](\mathfrak{f}) := \frac{1}{|\eta|} \sum_{i=1}^{|\eta|} \mathfrak{f}_{\eta_i^{i+k-1}}(i/|\eta|)\,.$$

Given a measurable $S \subset [0,1]$ and $\eta \in A^k$ we have that

$$\hat{\Pi}^k(S) = \frac{1}{|\eta|} \sum_{\{i : i/|\eta| \in S\}} \delta_{\eta_i^{i+k-1}} \in \mathcal{M}^+(A^k)$$

and $\hat{\Pi}^k(S;\zeta) \in \mathbb{R}_+$ is the weight associated with $\zeta \in A^k$ by $\hat{\Pi}^k(S) \in \mathcal{M}^+(A^k)$. Some simple relations are $\hat{\Pi}^k([0,1]) = \hat{\nu}^k$ and $\hat{\pi}(S) = \sum_{a \in A} \hat{\Pi}^1(S;a)a$.

The topology considered for the spatial empirical measures is always that induced by the weak convergence. Consider a measure $\Pi^k \in \mathcal{M}([0,1];\mathcal{M}^+(A^k))$; when Π^k is absolutely continuous with respect to Lebesgue measure we use the same symbol for the measure and its density, i.e. $\Pi^k = \Pi^k(x)dx$ and we have $\Pi^k(S) = \int_S \Pi^k(x)dx$. For almost any $x \in [0,1]$ we have that the density $\Pi^k(x) \in \mathcal{M}^+(A^k)$.

Having introduced spatial empirical measures, for completeness, we now recall the classical large deviation principles for spatial empirical measures of order one and two.

In the case that η_i are i.i.d. with distribution ν we have that $\hat{\Pi}^1_N$ satisfy a LDP when $N \to +\infty$ with rate functional

$$J^1(\rho) = \begin{cases} \int_0^1 H(\rho(x)|\nu)dx & \text{if } \rho = \rho(x)dx\,, \\ +\infty & \text{otherwise}\,, \end{cases} \tag{4.2.8}$$

where $\rho(x) = (\rho_i(x))_{i \in A}$ with $\sum_i \rho_i(x) = 1$ for *a.e.* $x \in [0,1]$.

In the case η_i are Markov with transition probability P we have that $\hat{\Pi}^2_N$ satisfy a LDP when $N \to +\infty$ with rate functional

$$J^2(\rho) = \begin{cases} \int_0^1 \sum_{i,j} \rho_{i,j}(x) \log \frac{\rho_{i,j}(x)}{\rho_i(x)P_{i,j}} dx & \text{if } \rho = \rho(x)dx\,, \\ +\infty & \text{otherwise}\,, \end{cases} \tag{4.2.9}$$

where $\rho(x) = (\rho_{i,j}(x))_{i,j \in A}$ and for *a.e.* $x \in [0,1]$ we have $\rho(x) \in \mathcal{M}^1_{st}(A^2)$.

4.2.2 Rational models

Rational models are a class of stochastic models for the random generation of words over a given alphabet. An informal definition of the measure is the following. Given a finite alphabet we associate a matrix with each of its elements. To any symbolic sequence we associate the matrix obtained as the ordered product of the matrices corresponding to its symbols. Finally, we compute the bilinear form of this matrix between the two fixed vectors. The resulting positive number is the weight associated with the symbolic sequence. Normalizing to one we

obtain a probability measure that we call a rational model.

More formally, we consider two sets A and B, we have for each $a \in A$ a non-negative $B \times B$ matrix M^a whose rows and columns are labeled by the elements of B. We alsorecall that the term non-negative here refers to the fact that all the elements $M^a_{b,b'}$ are non negative. We call $M := \sum_{a\in A} M^a$.

We consider two vectors with non-negative coordinates that we denote by $\vec{x} = \big(x(b)\big)_{b\in B}$, $\vec{y} = \big(y(b)\big)_{b\in B}$. In the cases of infinite alphabets, we will consider suitable summability conditions so that the model is well defined; we will discuss such cases afterwards.

We define a probability measure μ_N on A^N by

$$\mu_N(\eta) := \frac{\vec{y}^T \prod_{i=1}^N M^{\eta_i} \vec{x}}{Z_N}\,, \qquad \eta \in A^N\,, \tag{4.2.10}$$

where Z_N is the normalization factor

$$Z_N := \vec{y}^T M^{(N)} \vec{x}\,. \tag{4.2.11}$$

Note that the measure and the normalization factor depend on $\vec{x}, \vec{y}$ and $(M^a)_{a\in A}$. In cases where we need to underline the dependence on such factors, we will make the dependence explicit, for example by writing $\mu_N = \mu_N^{x,y}$.

We recall that throughout this work we have considered Matrix Product Ansatz (MPA) measures associated with matrices having non-negative entries, as this assumption ensures a probabilistic interpretation. For completeness, we point out that in the more general MPA formulation this condition is not necessary: the matrices may have complex entries. It follows that a probability measure represented by the MPA often corresponds to a rational model with a countable infinite alphabet B. However, there is not a complete equivalence: unlike rational models, in the MPA setting the matrices may be complex, although in many cases they are real and non-negative, thus corresponding to a rational model.

4.3 Large deviations: the strategy

In order to understand the statistics and the large deviations of the measure μ_N, we recall the construction introduced in the previous chapter.

We define a coupling measure C_N on $A^N \times B^{N+1}$ (i.e. $C_N \in \mathcal{M}^1(A^N \times B^{N+1})$) defined as follows. Consider $\eta = (\eta_1, \dots, \eta_N) \in A^N$ and $\zeta = (\zeta_1, \dots, \zeta_{N+1}) \in B^{N+1}$. We define

$$C_N(\eta, \zeta) := \frac{y(\zeta_1) \left(\prod_{i=1}^N M^{\eta_i}_{\zeta_i,\zeta_{i+1}}\right) x(\zeta_{N+1})}{Z_N}\,. \tag{4.3.1}$$

Multiplying by $|A|$ the normalization constant, the above measure can also be thought of as a probability measure on $A^{N+1} \times B^{N+1}$ whose value does not depend on η_{N+1}. As we have shown in Theorem 3.6, by construction we have

$$\sum_{\zeta\in B^{N+1}} C_N(\eta, \zeta) = \mu_N(\eta)\,. \tag{4.3.2}$$

It turns out that a very natural approach to compute large deviations rate functionals for the

measures $(\mu_N)_{N\in\mathbb{N}}$ is to prove large deviations principles for the measures $(C_N)_{N\in\mathbb{N}}$ and then apply the contraction principle.

We recall that the measure C_N is a Markov bridge, see item (2) in Theorem 3.6; we have not discussed some details regarding this proof and, as mentioned previously, we will focus on them in the next section.

4.4 The finite case

In this section we consider the case when the matrices $(M^a)_{a\in A}$ are finite. Since our main aim is to illustrate the general method, we consider the simplest situation where all the matrices $(M^a)_{a\in A}$ are irreducible and aperiodic and all the components of the vectors $\vec{x}, \vec{y}$ are strictly positive. We will discuss more general cases in a forthcoming paper [45] where also the mixture structure discussed in the previous chapter will be used to get a different representation. We refer to [47, 48] for results in the general case, obtained by a different approach in the context of rational models. Our approach can be adapted also to the case of infinite matrices as we will see in the next section. For simplicity, we compute the large deviations for some lower order algebraic and spatial empirical measures; with a more detailed analysis the result can be extended to higher order.

Given the collection of matrices $(M^a)_{a\in A}$ we define a $\big(|A||B|\big)\times\big(|A||B|\big)$ matrix $\mathfrak{M}$ having rows and columns labeled by elements of $A\times B$ and defined as

$$\mathfrak{M}_{(a,b),(a',b')} := M^a_{b,b'}\,, \qquad a,a'\in A\,;\ b,b'\in B\,. \tag{4.4.1}$$

Lemma 4.1. *If the $B\times B$ matrices $(M^a)_{a\in A}$ are irreducible and aperiodic, then the $(A\times B)\times(A\times B)$ matrix $\mathfrak{M}$ defined in (4.4.1) is irreducible and aperiodic.*

Proof. Since all the matrices $(M^a)_{a\in A}$ are irreducible then also $M=\sum_{a\in A}M^a$ is irreducible. Given any pair $b_I, b_F\in B$ then there exists a path $(b_1^*,\dots b_\ell^*)$ such that $b_1^*=b_I$, $b_\ell^*=b_F$ and $M_{b_i^*,b_{i+1}^*}>0$. Let us call $a_i^*\in A$, $i=1,\dots,\ell-1$ an element such that $M^{a_i^*}_{b_i^*,b_{i+1}^*}>0$. We need to show that for any $(a,b),(a'.b')\in A\times B$ there exists a path $(a_i,b_i)_{i=1}^k$ that satisfies $(a_1,b_1)=(a,b)$, $(a_k,b_k)=(a',b')$ and $\mathfrak{M}_{(a_i,b_i),(a_{i+1},b_{i+1})}>0$. Since by irreducibility there exists b_I such that M^a_{b,b_I} is strictly positive, we define $(a_1,b_1)=(a,b)$, $(a_2,b_2)=(a_1^*,b_I)$ where a_i^* and b_i^* are the sequences associated, as before, with a path from b_I to $b_F=b'$ and then $(a_i,b_i)=(a_{i-1}^*,b_{i-1}^*)$, $i=2,\dots k$ (observe that a_ℓ^* can be fixed equal to a'), and we have $k=\ell+1$. Aperiodicity follows easily by aperiodicity of all the $(M^a)_{a\in A}$. □

Recall the Perron-Frobenius theorem, see theorem 3.3. We have the following relations among maximal eigenvalues and eigenvectors of the matrices M and $\mathfrak{M}$.

Lemma 4.2. *If all the $(M^a)_{a\in A}$ are irreducible and aperiodic, we have that the maximal eigenvalue Λ of the matrix $\mathfrak{M}$ is equal to the maximal eigenvalue λ of the matrix M; moreover, if $\vec{e}$ is the positive maximal eigenvector oof M then the positive maximal eigenvector $\vec{\varepsilon}$ of $\mathfrak{M}$ is given by*

$$\varepsilon(a,b) := \frac{1}{\lambda}\sum_{b'\in B} M^a_{b,b'}e(b')\,. \tag{4.4.2}$$

Proof. By the previous Lemma, $\mathfrak{M}$ is irreducible and aperiodic and therefore it has a unique maximal eigenvalue with a positive eigenvector by Perron Frobenius theorem. By irreducibility and the positivity of $\vec{e}$ we have that the vector $\vec{\varepsilon}$, as defined by formula (4.4.2), has all the elements positive.

Taking the sum over $a \in A$ on both sides of (4.4.2) we observe that

$$\sum_{a\in A} \varepsilon(a,b) = \frac{1}{\lambda} \sum_{b'\in B} M_{b,b'} \underline{e}(b') = e(b)\,. \tag{4.4.3}$$

To finish the proof it is enough to show that $\vec{\varepsilon}$ is an eigenvector of $\mathfrak{M}$ with eigenvalue $\lambda > 0$ and this follows from

$$\sum_{a',b'} \mathfrak{M}_{(a,b),(a',b')} \varepsilon(a',b') = \sum_{a',b'} M^a_{b,b'} \varepsilon(a',b') \tag{4.4.4}$$

$$= \sum_{b'} M^a_{b,b'} e(b') = \lambda \varepsilon(a,b)\,, \tag{4.4.5}$$

where we used (4.4.3) and the definition (4.4.2) of ε. □

We recall the stochastic matrices P and $\mathfrak{S}$ associated with M and $\mathfrak{M}$, respectively, introduced in the previous chapter:

$$\begin{cases} P_{b,b'} = \frac{1}{\lambda} e(b)^{-1} M_{b,b'} e(b')\,, \\ \mathfrak{S}_{(a,b),(a',b')} = \frac{1}{\lambda} \varepsilon^{-1}(a,b) \mathfrak{M}_{(a,b),(a',b')} \varepsilon(a',b'). \end{cases} \tag{4.4.6}$$

By lemmas 4.1, 4.2 and 3.5, both matrices are stochastic, irreducible and aperiodic. For completeness, let us also recall that the previous relations (4.4.6) can be compactly written as:

$$\begin{cases} P = \frac{1}{\lambda} E^{-1} M E\,, \\ \mathfrak{S} = \frac{1}{\lambda} \mathcal{E}^{-1} \mathfrak{M} \mathcal{E}. \end{cases} \tag{4.4.7}$$

where E is the $B \times B$ diagonal matrix diagonal elements $E_{b,b} = e(b)$ and $\mathcal{E}$ is the $(A \times B) \times (A \times B)$ diagonal matrix having diagonal elements $\mathcal{E}_{(a,b),(a,b)} = \varepsilon(a,b)$.

Before going on, we state and prove the following lemma on unique invariant measures for completeness, although it will not be used directly.

Lemma 4.3. *If $\vartheta = (\vartheta_b)_{b\in B}$ is the invariant measure of the stochastic matrix P, then*

$$\Theta(a,b) = \vartheta(b) \frac{\varepsilon(a,b)}{e(b)}\,, \qquad a \in A, b \in B\,, \tag{4.4.8}$$

is the invariant measure of $\mathfrak{S}$.

Proof. Let us introduce the matrices $(P^a)_{a\in A}$ defined by $P^a_{b,b'} := \frac{1}{\lambda} e(b)^{-1} M^a_{b,b'} e(b')$. A direct relation between $\mathfrak{S}$ and the matrices $(P^a)_{a\in A}$ is

$$\mathfrak{S}_{(a,b),(a',b')} = \frac{1}{\lambda} \varepsilon^{-1}(a,b) M^a_{b,b'} \varepsilon(a',b')$$

$$= \left(\frac{\varepsilon(a,b)}{e(b)}\right)^{-1} P^a_{b,b'} \left(\frac{\varepsilon(a',b')}{e(b')}\right) .$$

Using the latter relation we can show that (4.4.8) is the invariant measure of $\mathfrak{S}$:

$$\begin{aligned}
&\sum_{a,b} \Theta(a,b)\mathfrak{S}_{(a,b),(a',b')} = \sum_{a,b} \Theta(a,b) \left(\frac{\varepsilon(a,b)}{e(b)}\right)^{-1} P^a_{b,b'} \left(\frac{\varepsilon(a',b')}{e(b')}\right) \\
&= \sum_{a,b} \vartheta(b) P^a_{b,b'} \left(\frac{\varepsilon(a',b')}{e(b')}\right) = \sum_{b\in B} \vartheta(b) P_{b,b'} \left(\frac{\varepsilon(a',b')}{e(b')}\right) \\
&= \vartheta(b') \left(\frac{\varepsilon(a',b')}{e(b')}\right) = \Theta(a',b') .
\end{aligned}$$

□

Using the previous lemmas, more precisely Lemmas 4.1, 4.2, we have proved that the aperiodicity and irreducibility of M imply the aperiodicity and irreducibility for the matrix $\mathfrak{M}$, as well as the relation between the maximal eigenvalue and its corresponding eigenvectors. With these additional details we complete the discussion in Theorem 3.6, item (2), where we showed that the enlarged measure C_N corresponds exactly to a Markov bridge associated with the $(A \times B) \times (A \times B)$ stochastic matrix $\mathfrak{S}$ and to the functions $g, f; A \times B \to \mathbb{R}$ defined by

$$\begin{cases} g(a,b) = y(b)\varepsilon(a,b) \\ f(a,b) = \varepsilon^{-1}(a,b)x(b) . \end{cases}$$

Note that in the previous chapter, when we referred to a Markov bridge, we did not make explicit in the notation the dependence on the stochastic matrix P and on the boundary functions g and f, since it was not needed there; we simply used the symbol ρ_N. From now on, however, this dependence will play a role and for this reason we will denote the Markov bridge by $\mathbb{P}^{P,g,f}$.

Before presenting our main result in the finite case we state a simple but powerful result that will be useful in the following.

Lemma 4.4. *Suppose that we have two sequences of probability measures μ_N and μ'_N such that there exist two sequences of positive constants k_N, K_N such that*

$$0 \le \liminf_{N\to+\infty} \frac{1}{N} \log k_N \le \limsup_{N\to+\infty} \frac{1}{N} \log K_N \le 0 , \tag{4.4.9}$$

with $k_N\mu'_N \le \mu_N \le K_N\mu'_N$. If μ'_N satisfies a large deviations principle then also μ_N satisfies a large deviations principle with the same rate.

Proof. The proof follows from a direct verification of the validity of the lower and upper bounds. □

We can now formulate the general statement in the finite case. We denote by $\mathcal{C}^2$ a generic

element of $\mathcal{M}^2_{stat}\left((A\times B)^2\right)$ and by $\mathcal{C}^1$ its one marginal, i.e.

$$\mathcal{C}^1(a,b) := \sum_{(a',b')\in A\times B} \mathcal{C}^2\left[(a,b),(a',b')\right] = \sum_{(a',b')\in A\times B} \mathcal{C}^2\left[(a',b'),(a,b)\right], \tag{4.4.10}$$

where the second equality follows from the finite stationarity of $\mathcal{C}^2$. We instead denote by ν^2 a generic element of $\mathcal{M}^2_{stat}(A^2)$.

Theorem 4.5. *Assume that $|A|<+\infty$, every $(M^a)_{a\in A}$ is irreducible and aperiodic and $|x\rangle, |y\rangle$ have strictly positive components. When η is distributed according to μ_N in (4.2.10), we have that $\hat{\nu}^2_N$, defined in (4.2.3), satisfies a large deviation principle on $\mathcal{M}^1_{st}(A^2)$ with speed N and rate functional*

$$I^2\left(\nu^2\right) = \inf\left\{\sum_{a,a'\in A;b,b'\in B} \mathcal{C}^2\left[(a,b),(a',b')\right]\log\frac{\mathcal{C}^2\left[(a,b),(a',b')\right]}{\mathcal{C}^1\left[(a,b)\right]M^{(a)}_{b,b'}}+\log\lambda\right\}, \tag{4.4.11}$$

where the infimum is over

$$\left\{\mathcal{C}^2\in\mathcal{M}^2_{stat}\left((A\times B)^2\right) : \sum_{b,b'\in B}\mathcal{C}^2\left[(a,b),(a',b')\right]=\nu^2(a,a'),\forall a,a'\in A\right\} \tag{4.4.12}$$

and λ is the maximal eigenvalue of M. We have also that $\hat{\Pi}^2_N$ satisfies a large deviations principle on $\mathcal{M}([0,1];\mathcal{M}^+(A^2))$ with rate functional

$$J^2\left(\Pi^2\right) = \begin{cases} \int_0^1 I^2\left(\Pi^2(x)\right)dx, & \text{if } \Pi^2=\Pi^2(x)dx, \\ +\infty & \text{otherwise.}\end{cases} \tag{4.4.13}$$

where $\Pi^2(x)\in\mathcal{M}^1_{st}(A^2)$ for almost all x.

Proof. By lemmas (4.1), (4.2) we have that the matrix

$$\mathfrak{S}_{(a,b),(a',b')} = \frac{1}{\lambda}\varepsilon(a,b)^{-1}\mathfrak{M}_{(a,b),(a',b')}\varepsilon(a',b') \tag{4.4.14}$$

is stochastic. By the positivity of the vectors $\vec{x},\vec{y},\vec{\varepsilon}$ we have that there exist positive constants k,K such that

$$k\mathbb{P}^{\mathfrak{S},\mathrm{st}}_N \le \mathbb{P}^{\mathfrak{S},g,f}_N \le K\mathbb{P}^{\mathfrak{S},\mathrm{st}}_N$$

where we called $\mathbb{P}^{\mathfrak{S},\mathrm{st}}_N$ the stationary Markov measure with transition probability $\mathfrak{S}$, i.e.

$$\mathbb{P}^{\mathfrak{S},\mathrm{st}}_N(\eta,\zeta) = \Theta(\eta_1,\zeta_1)\prod_{i=1}^N \mathfrak{S}_{(\eta_i,\zeta_i),(\eta_{i+1},\zeta_{i+1})}, \tag{4.4.15}$$

where Θ is the unique stationary measure of the transition matrix $\mathfrak{S}$ and the value of η_{N+1} is irrelevant by the form of $\mathfrak{S}$. We can therefore apply lemma (4.4) deducing the large deviations for the empirical measure for $\mathbb{P}^{\mathfrak{S},g,f}_N$ from the large deviations of the Markov measure $\mathbb{P}^{\mathfrak{S},\mathrm{st}}_N$, that are well established since it is a finite state irreducible stationary Markov measure.

Let us call

$$\hat{\mathcal{C}}^2_{N+1} := \frac{1}{N+1}\sum_{i=1}^{N+1} \delta_{((\eta_i,\zeta_i),(\eta_{i+1},\zeta_{i+1}))} \in \mathcal{M}^2_{st}\left((A\times B)^2\right), \tag{4.4.16}$$

where as usual we have that the sums are taken modulo $N+1$. By the general theory of large deviations for Markov chains, which we briefly discussed in the preliminaries section, we have that (4.4.16) satisfies a LDP with rate functional $\mathcal{I}^2$ on $\mathcal{M}^2_{st}\left((A\times B)^2\right)$ that is defined as follows. Let $\mathcal{C}^2 \in \mathcal{M}^2_{st}\left((A\times B)^2\right)$ and recall that $C^1 \in \mathcal{M}^1(A\times B)$ is its one marginal (4.4.10); it follows

$$\mathcal{I}^2(\mathcal{C}^2) = \sum_{a,a',b,b'} \mathcal{C}^2\left[(a,b),(a',b')\right] \log \frac{\mathcal{C}^2\left[(a,b),(a',b')\right]}{\mathcal{C}^1\left[(a,b)\right]\mathfrak{S}_{(a,b),(a',b')}}, \tag{4.4.17}$$

$$= \sum_{a,a',b,b'} \mathcal{C}^2\left[(a,b),(a',b')\right] \log \frac{\mathcal{C}^2\left[(a,b),(a',b')\right]}{\mathcal{C}^1\left[(a,b)\right]M^{(a)}_{b,b'}} + \log\lambda, \tag{4.4.18}$$

where the second inequality follows since $\mathcal{C}^2 \in \mathcal{M}^2_{st}\left((A\times B)^2\right)$. We can now apply the contraction principle obtaining for $\nu^2 \in \mathcal{M}^2_{stat}(A^2)$

$$I^2(\nu^2) = \inf_{\{\sum_{b,b'}\mathcal{C}^2\left[(a,b),(a',b')\right]=\nu^2(a,a'),\forall a,a'\}} \mathcal{I}^2(\mathcal{C}^2), \tag{4.4.19}$$

and this finishes the proof for the algebraic case. The proof of (4.4.13) can be obtained also following classic arguments for Markov measures. □

Large deviations principles for lower order empirical measures can be obtained by contraction, LDP for higher order empirical measures can be obtained following the same strategies used above and using the classic results for Markov measures (see, for example, [19]) on the enlarged state space and then applying the contraction principle.

In the general case it is difficult to get a more explicit form of the rate functional with respect to (4.4.11). We discuss, however, some special cases.

A general form. The critical condition in the minimization (4.4.11) is obtained considering several constraints that the measure C^2 has to satisfy. The first collection of constraints is given by (4.4.12), while the other collection of constraints is obtained imposing $C^2 \in \mathcal{M}_{st}\left((A\times B)^2\right)$. To each constraint there corresponds a Lagrange multiplier. By a direct computation the critical conditions are equivalent to

$$\frac{\mathcal{C}^2\left[(a,b),(a',b')\right]}{\mathcal{C}^1\left[(a,b)\right]} = k^{-1}M^a_{b,b'}p(a,a')\frac{\gamma_{a',b'}}{\gamma_{a,b}}, \tag{4.4.20}$$

where the positive matrix p encodes the set of Lagrange multipliers for constraints (4.4.12), while the numbers γ and the constant k are the Lagrange multipliers for the second group of constraints. Since on the left-hand side we have a stochastic matrix, we obtain that the minimizer is achieved when $\left(\gamma_{a,b}\right)_{a\in A,b\in B}$ is the maximal eigenvector of the positive $(A\times B)\times$

$(A \times B)$ matrix $\left(M^a_{b,b'} p(a,a')\right)^{b,b' \in B}_{a,a' \in A}$ and k is the corresponding maximal eigenvalue.

Inserting (4.4.20) as the minimizer in (4.4.11) we obtain, using that $\mathcal{C}^2$ is stationary and has ν^2 as $A \times A$ marginal (that are the constraints (4.4.12)),

$$I(\nu^2) = \sum_{a,a'} \nu^2(a,a') \log p(a,a') + \log(\lambda/k)\,. \tag{4.4.21}$$

It can be shown that in (4.4.20) and (4.4.21) we can restrict to matrices $p(a,a')$ that are stochastic. Similar considerations can also be applied for the spatial empirical measure.

Parallel eigenvectors. We consider the special case when

$$M^a_{b,b'} = m(a) P^a_{b,b'} \frac{\varphi(b)}{\varphi(b')} \tag{4.4.22}$$

where $(P^a)_{a \in A}$ is a collection of $B \times B$ stochastic matrices, $(m(a))_{a \in A}$ is a collection of positive numbers and $(\varphi(b))_{b \in B}$ is a positive vector. This corresponds to a family of positive matrices M^a, each having the same maximal eigenvector $\vec{\varphi}$ with corresponding maximal eigenvalue $m(a)$. Indeed lemma 3.5 implies that all the positive matrices having $\vec{\varphi}$ as a maximal eigenvector are of the form (4.4.22) for some stochastic matrix P^a. We have the following.

Lemma 4.6. *For matrices $(M^a)_{a \in A}$ of the form (4.4.22) we have that the maximal eigenvalue and eigenvector k, $\vec{\gamma}$ of the $(A \times B) \times (A \times B)$ matrix $\left(M^a_{b,b'} p(a,a')\right)^{b,b' \in B}_{a,a' \in A}$ are related to μ, $\vec{\psi}$ those of the $A \times A$ matrix $\left(p(a,a')m(a')\right)_{a,a' \in A}$ by the relations*

$$\begin{cases} k = \mu\,, \\ \gamma(a,b) = \psi(a) m(a) \varphi(b)\,. \end{cases} \tag{4.4.23}$$

Moreover, we have that the maximal eigenvalue λ of the $B \times B$ matrix M is $\sum_a m(a)$ with corresponding eigenvector $\vec{\varphi}$.

Proof. Since a positive eigenvector is necessarily the one associated with the maximal eigenvalue, by a direct computation we have

$$\sum_{a',b'} M^a_{b,b'} p(a,a') \gamma(a',b') \tag{4.4.24}$$

$$= \sum_{a',b'} m(a) S^a_{b,b'} \frac{\varphi(b)}{\varphi(b')} p(a,a') m(a') \psi(a') \varphi(b') \tag{4.4.25}$$

$$= \varphi(b) m(a) \sum_{a'} p(a,a') m(a') \psi(a') = \mu \gamma(a,b)\,. \tag{4.4.26}$$

The last statement also follows from the direct computation

$$\sum_{b'} M_{b,b'} \varphi(b') = \varphi(b) \sum_a m(a) \sum_{b'} S^a_{b,b'} = \left(\sum_a m(a) \right) \varphi(b)\,.$$

□

We can now deduce the following.

Proposition 4.7. *When the matrices $(M^a)_{a\in A}$ have all the same Perron eigenvector, then the large deviations rate functional I^2 in (4.4.11) coincides with the large deviations rate functional for the pair empirical measure of a sequence of i.i.d. random variables taking values on A and having probability distribution $\mathbb{P}(a) = \frac{m(a)}{\sum_{a'\in A} m(a')}$, $a \in A$; i.e. we have*

$$I(\nu^2) = \sum_{a,a'} \nu^2(a,a') \log \frac{\nu^2(a,a')}{\nu^1(a)\mathbb{P}(a')} \,. \tag{4.4.27}$$

Proof. Inserting the eigenvalues obtained in the previous lemma into the relation (4.4.20) for the minimizer we get

$$\mathcal{C}^2((a,b),(a',b')) = \mathcal{C}^1(a,b)k^{-1}P^a_{b,b'}p(a,a')\frac{\psi(a')m(a')}{\psi(a)} \,,$$

and we can determine p satisfying (4.4.12) just summing over $(b,b') \in B^2$ the above equality getting

$$p(a,a') = \frac{k\nu^{(2)}(a,a')\psi(a)}{\nu^{(1)}(a)\psi(a')m(a')} \,. \tag{4.4.28}$$

Inserting this value in (4.4.21), recalling that $\nu^2 \in \mathcal{M}_{st}\left(A^2\right)$ and the value of λ obtained in Lemma 4.6, by a direct computation and using Lemma 4.6 we obtain (4.4.27), which is the large deviations rate functional for the pair empirical measure for a product measure on A, assigning probability $\mathbb{P}(a)$ to each element $a \in A$ (see for example [19]). □

As in the general case, also in this case the rate functional for the empirical measures of order different from 2 and for the spatial empirical measures can be obtained by the same approach.

Stochastic matrices. A special case of the situation of the previous subsection is when all the matrices $(M^a)_{a\in A}$ are stochastic. In this case all the Perron eigenvalues are equal to one and the distribution $\mathbb{P}(a) = \frac{1}{|A|}$ is the uniform one. In this case, the results can also be obtained directly as follows. The first observation is that, when all the matrices $(M^a)_{a\in A}$ are stochastic, the measure $\mu_N^{1,1}$ associated with the vectors $\vec{x}, \vec{y}$ having all the coordinates identically equal to one, is the uniform measure on A^N. This follows by a direct check.

Let us call $|\vec{x}|_M = \sup_a x(a)$ and $|\vec{x}|_m = \inf_a x(a)$ that are both finite and strictly positive, we use a similar notation for $\vec{y}$ too. We have

$$|x|_m|y|_m\mu_N^{1,1} \le \mu_N^{x,y} \le |x|_M|y|_M\mu_N^{1,1} \,,$$

and from lemma 4.4 we deduce

$$I^k(\nu^k) = \begin{cases} \sum_{a_1^k\in A^k} \nu^k(a_1^k)\left(\log \nu^k(a_1^k) - k\log|A|\right) & \text{if } \nu^k \in \mathcal{M}^1_{st}(A^k) \\ +\infty & \text{otherwise} \end{cases} \tag{4.4.29}$$

since the rate functional (4.4.29) is exactly the rate functionals associated with a product of uniform measures. Similar considerations can be done also for the spatial empirical measures.

4.5 The infinite case: boundary driven TASEP

As we proved in the previous section, for Markov measures in the finite case an irreducibility condition is sufficient to have a LDP both for the algebraic and spatial empirical measures. This is not the case for a countable infinite alphabet where suitable conditions have to be satisfied; see [29]. In the example we consider, only a LDP for the spatial empirical measures can be obtained directly, and we therefore restrict to that case. A proof of the general statements in the infinite case is more subtle, but for our specific example we proceed by direct computation.

We consider the boundary driven TASEP and use the representation of the matrices valid for $\alpha + \beta > 1$, discussed in Section 3.3.1. We recall that, with this choice of representation, the matrix $M = M^0 + M^1$ is the infinite matrix with all the elements on the main diagonal equal to 2, the ones on the lower and on the upper diagonal equal to 1 and all the other entries equal to 0 (see (2.3.9) and (3.3.2)). Its Perron eigenvalue is $\lambda = 4$ with the corresponding eigenvector equal to $\vec{e} = (b+1)_{b\in\mathbb{N}_0}$.

We define the matrix $\mathfrak{M}_{(a,b),(a',b')} = M^a_{b,b'}$, with $(a,b) \in \{0,1\} \times \mathbb{N}_0$, which has the same Perron eigenvalue $\Lambda = \lambda = 4$ as the matrix M and the corresponding eigenvector is defined according to (4.4.2) by the formula:

$$\varepsilon(a,b) = \frac{1}{4} \sum_{b' \in \mathbb{N}_0} M^a_{b,b'}(b'+1)\,.$$

By a direct computation, we obtain:

$$\varepsilon(a,b) = \frac{1}{4}(2b + 2a + 1)\,, \qquad a \in \{0,1\}\,,\ b \in \mathbb{N}_0\,.$$

We can now write the stochastic $\Big(\{0,1\} \times \mathbb{N}_0\Big) \times \Big(\{0,1\} \times \mathbb{N}_0\Big)$ matrix $\mathfrak{S}$ explicitly:

$$\mathfrak{S}_{(a,b),(a',b')} = \frac{1}{4}\frac{2(b'+a')+1}{2(b+a)+1} M^a_{b,b'}\,, \qquad a \in \{0,1\}\,,\, b \in \mathbb{N}_0\,. \tag{4.5.1}$$

Recalling the form of the matrix M in (3.3.2), we have that all the elements of the matrix $\mathfrak{S}$ are equal to zero except for those of the form specified below. The non zero elements associated with transitions from the state $(0,0)$ are

$$\mathfrak{S}_{(0,0),(0,0)} = \frac{1}{4}\,, \qquad \mathfrak{S}_{(0,0),(1,0)} = \frac{3}{4}\,;$$

the non zero elements associated with transitions from the state $(1,0)$ are

$$\mathfrak{S}_{(1,0),(1,0)} = S_{(1,0)(0,1)} = \frac{1}{4}\,, \qquad \mathfrak{S}_{(1,0),(0,0)} = \frac{1}{12}\,, \qquad \mathfrak{S}_{(1,0),(1,1)} = \frac{5}{12}\,;$$

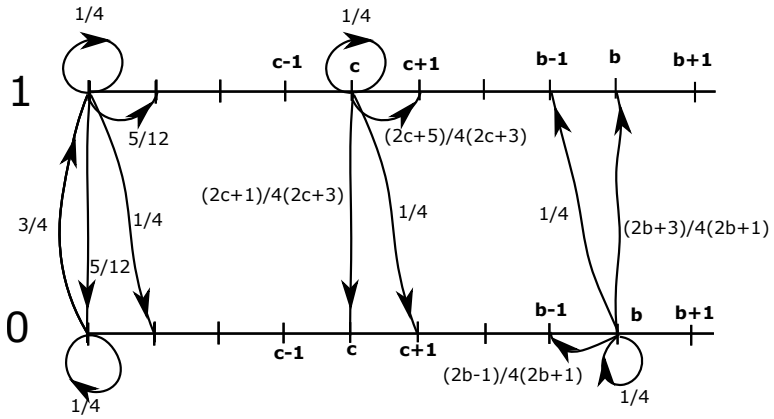

Figure 4.1: A graphical representation of the transition probabilities of the Markov chain $\mathfrak{S}$ on $\{0,1\}\times\mathbb{N}_0$. The coordinates $\mathbf{c},\mathbf{b}$ on the axis $\mathbb{N}_0$ are written in boldface to distinguish them from the values of the transition probabilities.

the non zero elements associated with transitions from the state $(0,b)$ with $b\geq 1$ are

$$\mathfrak{S}_{(0,b),(0,b)}=\mathfrak{S}_{(0,b),(1,b-1)}=\frac{1}{4}\,,\ \mathfrak{S}_{(0,b)(1,b)}=\frac{2b+3}{4(2b+1)}\,,\ \mathfrak{S}_{(0,b),(0,b-1)}=\frac{2b-1}{4(2b+1)}\,;$$

the non zero elements associated with transitions from the state $(1,b)$ with $b\geq 1$ are

$$\mathfrak{S}_{(1,b),(1,b)}=\mathfrak{S}_{(1,b),(0,b+1)}=\frac{1}{4}\,,\ \mathfrak{S}_{(1,b),(0,b)}=\frac{2b+1}{4(2b+3)}\,,\ \mathfrak{S}_{(1,b),(1,b+1)}=\frac{2b+5}{4(2b+3)}\,.$$

The transition graph associated with the matrix $\mathfrak{S}$ is drawn in Figure 4.1. The corresponding Markov chain is a special random walk on two infinite lines with a left boundary at the origin of the two lines. Transitions are possible between the two lines. The transition probabilities of the random walk are spatially non homogeneous. According to item (2) in the proof of Theorem 3.6, in Section 3.2, the enlarged measure of the boundary driven TASEP is a Markov bridge of length $N+1$ associated with above random walk.

It will be useful in order to simplify the proofs to introduce an effective random walk on $\{0,1\}\times\mathbb{N}_0$ with a simpler transition matrix $\mathcal{S}$. This is again a random walk on the two infinite lines but it is spatially homogeneous with a different behavior just at the boundary sites. The transition matrix is given by

$$\begin{cases}\mathcal{S}_{(0,b)(a',b+s)}=\frac{1}{4}\,, & b\neq 0;\ a'=0,1;\ s=0,-1;\\ \mathcal{S}_{(1,b)(a',b+s)}=\frac{1}{4}\,, & a'=0,1;\ s=0,+1;\\ \mathcal{S}_{(0,0)(a',0)}=\frac{1}{2}\,, & a'=0,1\,;\end{cases}\tag{4.5.2}$$

we give a graphical representation of the transition probabilities in Figure 4.2.

The stochastic matrix $\mathcal{S}$ is almost related by a generalized Doob transform to the stochastic matrix $\mathfrak{S}$; the relation is violated only for transitions exiting from the state $(0,0)$. By a direct

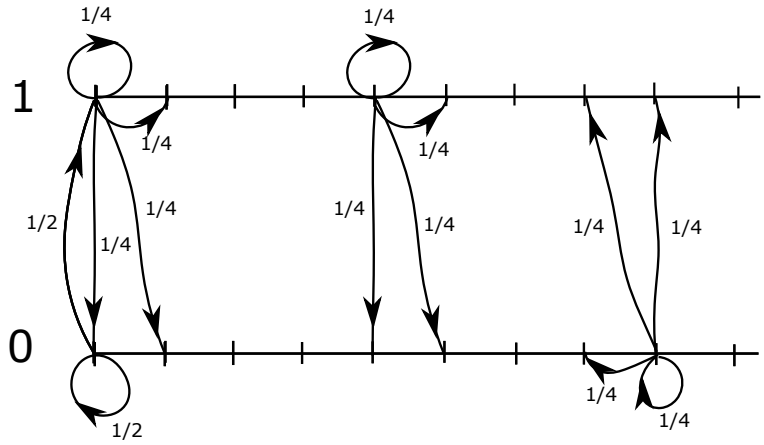

Figure 4.2: A graphical representation of the transition probabilities of the effective Markov chain $\mathcal{S}$ on $\{0,1\}\times\mathbb{N}_0$.

computation, we obtain the following relation

$$\mathcal{S}_{(a,b)(a',b')} = \mathfrak{S}_{(a,b)(a',b')}\frac{2(a+b)+1}{2(a'+b')+1}2^{\mathbb{1}\left((a,b)=(0,0)\right)} = \frac{1}{4}M^a_{b,b'}2^{\mathbb{1}\left((a,b)=(0,0)\right)}\,. \tag{4.5.3}$$

The possibility to work with this simple homogeneous random walk is due to the special form of the matrices M^0, M^1.

Let us recall that $\mathbb{P}_N^{\mathfrak{S},g,f}$ and $\mathbb{P}_N^{\mathcal{S},g',f'}$ denote Markov bridges with transition matrix $\mathfrak{S}$ and $\mathcal{S}$, respectively. We call $\mathfrak{Z}_N(g,f)$ and $\mathcal{Z}_N(g',f')$ the corresponding normalization factors appearing in the definition (3.1.2). We denote by $(\eta,\zeta)=(\eta_k,\zeta_k)_{k=1}^{N+1}$ a trajectory on $\{0,1\}\times\mathbb{N}_0$ of length $N+1$ and having positive probability with respect to the Markov measures. We call $\mathbb{P}_N^{\mathfrak{S}}$ and $\mathbb{P}_N^{\mathcal{S}}$ the corresponding Markov measures with fixed initial condition; we have, for example, $\mathbb{P}_N^{\mathfrak{S}}(\eta,\zeta) := \prod_{i=1}^N \mathfrak{S}_{(\eta_i,\zeta_i),(\eta_{i+1},\zeta_{i+1})}$, and likewise for the other matrix. Finally, considering (4.2.2) for the product space $A\times B$, we call $\hat{\nu}_N^1[\eta,\zeta] = \frac{1}{N+1}\sum_{i=1}^{N+1}\delta_{(\eta_i,\zeta_i)} \in \mathcal{M}^1(A\times B)$. The introduction of the simplified homogeneous Markov measure with transition probability $\mathcal{S}$ is due to the fact that we can deduce LDP's for Markov measures with transition probability $\mathfrak{S}$ from the corresponding ones for the matrix $\mathcal{S}$. This follows from the next lemma. Let us introduce the function $h:\{0,1\}\times\mathbb{N}_0\to\mathbb{R}$ defined by $h(a,b):=2(a+b)+1$.

Lemma 4.8. *When $\mathfrak{Z}_N,\mathcal{Z}_N<+\infty$ we have*

$$\mathbb{P}_N^{\mathfrak{S},g,f}(\eta,\zeta) = \mathbb{P}_N^{\mathcal{S},g/h,fh}(\eta,\zeta)2^{-N\hat{\nu}_N^1(0,0)}\frac{\mathcal{Z}_N(g/h,fh)}{\mathfrak{Z}_N(g,f)}\,. \tag{4.5.4}$$

Proof. By a direct computation we have for any fixed (η,ζ)

$$\mathbb{P}_N^{\mathfrak{S}}(\eta,\zeta) = \mathbb{P}_N^{\mathcal{S}}(\eta,\zeta)\frac{h(\eta_{N+1},\zeta_{N+1})}{h(\eta_1,\zeta_1)}2^{-N\hat{\nu}_N^1(0,0)}\,.$$

Multiplying both sides by $g(\eta_1,\zeta_1)$ and $f(\eta_{N+1},\zeta_{N+1})$ and inserting the constants we get (4.5.4). □

Now we illustrate the strategy to prove a spatial LDP for the Markov measure (4.3.1) for the matrices and the vectors defined in (2.3.9). We prove first the corresponding result for the simple homogeneous Markov measure with transition $\mathcal{S}$, and then use Lemma 4.8 to deduce the result for the original matrix $\mathfrak{S}$. We will then apply the contraction principle to get the LD rate functional for the boundary driven TASEP.

4.5.1 The probabilistic structure

The simple homogeneous random walk with transition matrix $\mathcal{S}$ has some special features. Since $\sum_{b'\in\mathbb{N}_0}\mathcal{S}_{(a,b),(a',b')}=\frac{1}{2}$ for any $a,a'\in\{0,1\}$ and $b\in\mathbb{N}_0$, we deduce that, under the measure $\mathbb{P}_N^{\mathcal{S}}$, the variables η_2^{N+1} are i.i.d. all having distribution $\mathcal{B}_{1/2}$. Moreover, we can construct samples of the Markov measure $\mathbb{P}_N^{\mathcal{S}}$ using sequences of i.i.d. random vectors. More precisely, we introduce $(X_i^I,Y_i^I)_{i\in\mathbb{N}}$ a sequence of i.i.d. random vectors such that

$$\mathbb{P}\Big[(X_i^I,Y_i^I)=(a,b)\Big]=\begin{cases}\frac{1}{4} & \text{if } (a,b)=(0,0),(0,-1),(1,0),(1,1)\,,\\ 0 & \text{otherwise}\,.\end{cases}$$

We also introduce $(X_i^B,Y_i^B)_{i\in\mathbb{N}}$ a sequence of i.i.d. random vectors such that

$$\mathbb{P}\Big[(X_i^B,Y_i^B)=(a,b)\Big]=\begin{cases}\frac{1}{4} & \text{if } (a,b)=(1,0),(1,1)\,,\\ \frac{1}{2} & \text{if } (a,b)=(0,0)\,,\\ 0 & \text{otherwise}\,.\end{cases}$$

The two collection of i.i.d. random vectors are also independent of each other. The random variables $(X_i^I,Y_i^I)_{i\in\mathbb{N}}$ have to be used to construct the path of the Markov measure $\mathbb{P}_N^{\mathcal{S}}$ in the bulk, while the random variables $(X_i^B,Y_i^B)_{i\in\mathbb{N}}$ have to be used to construct the path in the boundary.

We call $\mu^I,\mu^B\in\mathcal{M}^1\Big(\{0,1\}\times\{-1,0,1\}\Big)$ the distribution of the pair (X_i^I,Y_i^I) and (X_i^B,Y_i^B), respectively; namely

$$\begin{cases}\mu^I_{0,0}=\mu^I_{0,-1}=\mu^I_{1,0}=\mu^I_{1,1}=\frac{1}{4}\,,\\ \mu^I_{0,1}=\mu^I_{1,-1}=0\,,\end{cases}\qquad\qquad\begin{cases}\mu^B_{1,0}=\mu^B_{1,1}=\frac{1}{4}\,,\\ \mu^B_{0,0}=\frac{1}{2}\,,\\ \mu^B_{1,-1}=\mu^B_{0,1}=0\,.\end{cases}$$

We iteratively define the sequence of random variables (η,ζ) distributed according to $\mathbb{P}_N^{\mathcal{S}}$. The variables (η_1,ζ_1) are fixed and we generate ζ_2 as follows. If $(\eta_1,\zeta_1)=(0,0)$ then $\zeta_2=0$; otherwise $\zeta_2=\zeta_1$ or $\zeta_2=\zeta_1+2\eta_1-1$ with equal probability and independently of the sequences of i.i.d. vectors. Once that ζ_2 is determined, we generate the whole sequence as follows. For any $i>1$ we have that, when $\zeta_i>0$ then $(\eta_i,\zeta_{i+1}-\zeta_i)=(X_i^I,Y_i^I)$, when instead $\zeta_i=0$ then $(\eta_i,\zeta_{i+1}-\zeta_i)=(X_i^B,Y_i^B)$. By a direct inspection, it is possible to see that the sequence $(\eta_i,\zeta_i)_{i=1}^{N+1}$, constructed in this way, has law $\mathbb{P}_N^{\mathcal{S}}$. In the case of an initial condition with the variables (η_1,ζ_1) distributed as $\mathcal{B}_{1/2}(\eta_1)\delta_{\zeta_1}$, the whole collection of random variables can be constructed by the same iteration. In this case, as input to start the iteration, we only

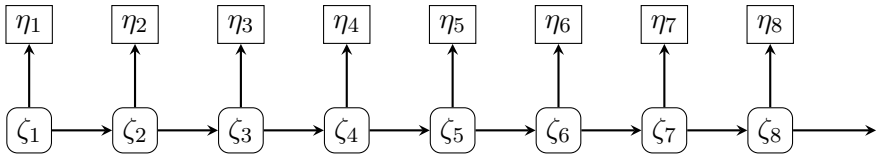

Figure 4.3: Iterative construction of the sequence (η, ζ) distributed according to $\mathbb{P}_N^{\mathrm{S}}$, starting from the value ζ_1.

need assign the value ζ_1. In the Figure 4.3 the diagram shows how to construct iteratively the variables starting form ζ_1, and the arrows exiting from each ζ_i show which are the variables that can be constructed directly using the i.i.d. sequences. As follows from the diagram in the Figure 4.3, the sequence ζ alone is Markovian; this property is related to the representation in terms of mixtures of inhomogeneous product measures discussed in the previous chapter.

4.5.2 The empirical observables

It is convenient to study the large deviations asymptotic in terms of some empirical observables that we are going to introduce. First of all, we consider two empirical measures defined by

$$\begin{cases} \hat{\Pi}_N^B[\eta,\zeta] := \frac{1}{N}\sum_{i=1}^N \mathbb{1}(\zeta_i = 0)\delta_{(\eta_i,\tilde{\zeta}_i)}\delta_{i/N} \\ \hat{\Pi}_N^I[\eta,\zeta] := \frac{1}{N}\sum_{i=1}^N \mathbb{1}(\zeta_i > 0)\delta_{(\eta_i,\tilde{\zeta}_i)}\delta_{i/N}\,, \end{cases} \tag{4.5.5}$$

where we use the random variables $\tilde{\zeta}_i := \zeta_{i+1} - \zeta_i$. The two empirical measures in (4.5.5) are random elements of $\mathcal{M}\Big([0,1];\mathcal{M}^+\Big(A\times\{-1,0,+1\}\Big)\Big)$. Some other natural empirical measures that can be constructed starting from (4.5.5) are

$$\hat{\Pi}_N[\eta,\zeta] := \frac{1}{N}\sum_{i=1}^N \delta_{(\eta_i,\tilde{\zeta}_i)}\delta_{i/N} \in \mathcal{M}\Big([0,1];\mathcal{M}^+\Big(A\times\{-1,0,+1\}\Big)\Big)\,, \tag{4.5.6}$$

and

$$\hat{\pi}_N^B[\zeta] := \frac{1}{N}\sum_{i=1}^N \mathbb{1}(\zeta_i = 0)\delta_{i/N} \in \mathcal{M}^+([0,1])\,. \tag{4.5.7}$$

The empirical measure (4.5.6) can be obtained by $\hat{\Pi}_N = \hat{\Pi}_N^B + \hat{\Pi}_N^I$, while the empirical measure (4.5.7) is obtained by $\hat{\pi}_N^B = \sum_a \sum_{\tilde{b}} \hat{\Pi}_N^B(a,b)$, where, to better explain again the notation, $\hat{\Pi}_N^B \in \mathcal{M}\Big([0,1];\mathcal{M}^+\Big(A\times\{-1,0,+1\}\Big)\Big)$ is given by $\Big(\hat{\Pi}_N^B(a,b) \in \mathcal{M}\Big([0,1];\mathbb{R}\Big)\Big)_{a\in A}^{b=-1,0,+1}$. We use similar notation for other measures.

The above empirical measures depend on the variable η, ζ and we underline such a dependence inside squared parenthesis $[\cdot]$.

The empirical measure (4.2.5) associated with the configuration of the TASEP model is obtained starting from (4.5.6) by

$$\hat{\pi}_N[\eta] = \sum_a \sum_b \hat{\Pi}_N[\eta,\zeta](a,b)\,a\,; \tag{4.5.8}$$

our goal is to deduce a LDP for (4.5.8) by contraction from a LD on a larger state space.

It will be useful to consider the following relation for the empirical measure appearing in Lemma 4.8

$$\hat{\nu}_N(0,0) = \int_{[0,1]} \sum_a \sum_b (1-a)\hat{\Pi}_N^B(a,b;dx)\,. \tag{4.5.9}$$

We introduce a last empirical object and then proceed to describe their continuous counterpart, scaling limits and large deviations. This is a piecewise linear path $\hat{z}_N = \big(\hat{z}_N(x)\big)_{x\in[0,1]}$ that is defined setting $\hat{z}_N\left(\frac{i}{N}\right) = \frac{\zeta_{i+1}}{N}$, $i = 0,\dots,N$ and then defining it at the remaining points by linear interpolation. We have the following relation

$$\hat{z}_N\left(\frac{j}{N}\right) - \hat{z}_N\left(\frac{i}{N}\right) = \int_{(i/N,j/N]} \sum_{a=0,1} \sum_{b=-1,0,1} \hat{\Pi}_N(a,b;dx)b\,. \tag{4.5.10}$$

Note that if $i = 0$ and $j = N$, we have on the left hand side above $\hat{z}_N(1) - \hat{z}_N(0)$ and on the right hand side the integral with respect to the whole interval giving a continuous correspondence.

We are interested in studying the LD for the joint empirical variables

$$\left(\hat{z}_N, \hat{\Pi}_N^B, \hat{\Pi}_N^I\right) \in C([0,1]) \times \mathcal{M}\Big([0,1]; \mathcal{M}^+\Big(A \times \{-1,0,+1\}\Big)\Big)^2\,, \tag{4.5.11}$$

when (η,ξ) are distributed according to the measure (4.3.1) with matrices and vectors defined in (2.3.9), and where we consider the space of continuous functions with the topology induced by the supremum norm and the product space with the product topology.

Let $\mathbb{P}_N^{\mathcal{S},*}$ denote the Markov measure with initial distribution $\mathcal{B}_{1/2}(\eta_1)\delta_{\zeta_1}$. Since the probability of each sequence (η,ζ) depends only on the empirical measures in (4.5.5), the measure $\mathbb{P}_N^{\mathcal{S},*}$ can be compactly written as

$$\mathbb{P}_N^{\mathcal{S},*} = \prod_{a,b} \left[\mu^B(a,b)^{\left(N\int_{[0,1]},\hat{\Pi}_N^B(a,b)\right)} \mu^I(a,b)^{\left(N\int_{[0,1]} d\hat{\Pi}_N^I(a,b)\right)}\right]\,, \tag{4.5.12}$$

where we recall that in the product $a \in \{0,1\}$ and $b \in \{-1,0,1\}$.

According to the definition (3.1.2), as discussed in the proof of Theorem 3.6 (see item (2)), the measure (4.3.1) corresponds to the Markov bridge $\mathbb{P}^{\mathfrak{S},g,f}$ associated with the transition matrix $\mathfrak{S}$ and the two functions g,f determined by the vectors $\vec{y}$ and $\vec{x}$ and defined by

$$\begin{cases} g(a,b) := \frac{(1-\alpha)^b(2a+2b+1)}{4\alpha^b}\,, \\ f(a,b) := \frac{4(1-\beta)^b}{\beta^b(2a+2b+1)}\,. \end{cases} \tag{4.5.13}$$

We will then deduce the LD for the empirical measure (4.2.5) when η is distributed according to the invariant measure of the TASEP by contraction.

We will use Lemma 4.8 and deduce LD for $\mathbb{P}^{\mathfrak{S},g,f}$ from LD for $\mathbb{P}_N^{\mathcal{S},g/h,fh}$. By the expression

(4.5.13), it follows that the functions $g/h, fh$ do not depend on the variable a then we have

$$\mathbb{P}_N^{\mathcal{S},g/h,fh}(\eta,\zeta) = \left(\frac{1-\alpha}{\alpha}\right)^{\zeta_1} \mathbb{P}_N^{\mathcal{S},*}(\eta,\zeta) \left(\frac{1-\beta}{\beta}\right)^{\zeta_{N+1}} (\mathcal{Z}_N)^{-1} , \tag{4.5.14}$$

where $\mathcal{Z}_N$ is the same normalization constant appearing in 4.5.4.

4.5.3 The rate functionals

We now describe some joint LD rate functionals associated with the empirical observables we described before when the variables (η,ζ) are distributed according to different measures. The LD rate functionals are defined on the space where the empirical triple of observables lives, namely the one introduced in (4.5.11). The LD rate functionals are concentrated on absolutely continuous paths and measures; this follows from the fact that for any configuration (η,ζ), any $(a,b) \in \{0,1\} \times \{-1,0,1\}$ and any continuous function G we have

$$\left| \int_{[0,1]} \Pi_N(a,b;dx) G(x) \right| \leq \frac{1}{N} \sum_{i=1}^{N} \left| G(i/N) \right| ,$$

so that any limit measure Π satisfies $\left| \int_{[0,1]} \Pi(a,b;dx) G(x) \right| \leq \int_{[0,1]} \left| G(x) \right| dx$ and moreover

$$\mathbb{P}_N^{\mathcal{S},*} \left(\left| \int_{[0,1]} \Pi_N(a,b;dx) G(x) \right| \geq \int_{[0,1]} \left| G(x) \right| + \varepsilon \right) = 0 ,$$

for N large enough. This simple estimate, and similar ones for any empirical measure, allows to show that the LD rate functional is identically $+\infty$ on measures and paths that are not absolutely continuous with respect to the Lebesgue measure, or whose density is not bounded by one. We therefore only need to show the form of the rate functional on the closed subset of absolutely continuous measures with densities bounded by one. Moreover, by relation (4.5.10) we have

$$\mathbb{P}_N^{\mathcal{S},*} \left(\sup_{x,y} \left| \hat{z}_N(x) - \hat{z}_N(y) - \int_{[x,y]} \sum_{a=0,1} \sum_{b=-1,0,1} \hat{\Pi}_N(a,b;dz) b \right| > \varepsilon \right) = 0 , \tag{4.5.15}$$

for any ε and N big enough. We deduce, therefore, that the rate functional has a domain that is concentrated on the closed set of absolutely continuous measures and paths where the following relations are satisfied.

Consider an absolutely continuous triple

$$\left(z, \Pi^B, \Pi^I\right) \in C([0,1]) \times \mathcal{M}\Big([0,1]; \mathcal{M}^+\Big(A \times \{-1,0,+1\}\Big)\Big)^2 ,$$

we will use the same notation for an absolutely continuous measure and its density. We write $\Pi^B = m(x)\nu^B(x)dx$, and $\Pi^I = (1-m(x))\nu^I(x)dx$ where $\nu^B(x), \nu^I(x)$ are space dependent probability measures on $\{0,1\} \times \{-1,0,+1\}$ and $0 \leq m(x) \leq 1$.

This means that $\nu^B(x) = \left(\nu^B_{a,b}\right)_{a=0,1}^{b=-1,0,+1}$ and $\sum_{a,b}\nu^B_{a,b}(x) = 1$ and $\nu^I(x) = \left(\nu^I_{a,b}\right)_{a=0,1}^{b=-1,0,+1}$ and $\sum_{a,b}\nu^I_{a,b}(x) = 1$ for a.e. $x \in [0,1]$. As in (4.5.6) we have

$$\Pi(x)dx = \Pi^B(x)dx + \Pi^I(x)dx = \Big[m(x)\nu^B(x) + (1-m(x))\nu^I(x)\Big]dx\,.$$

It is possible to write in this form Π^B and Π^I since the density of $\sum_{a,b}\Pi(a,b)$ is identically one. For any x, the support of $\nu^B(x)$ is contained in that of μ^B and the support of $\nu^I(x)$ is contained in that of μ^I.

The continuous counterpart of (4.5.7) is given by

$$m(x)dx = \sum_{a,b}\Pi^B_{a,b}(x)dx\,.$$

By (4.5.15), we deduce that the rate functional is finite only when the following relation is satisfied a.e.

$$\begin{aligned}\dot z(x) &= \sum_a\sum_b \Pi_{a,b}(x)b dx = \Pi_{1,1}(x) - \Pi_{0,-1}(x)\\ &= m(x)\nu^B_{1,1}(x) + (1-m(x))\left(\nu^I_{1,1}(x) - \nu^I_{0,-1}(x)\right)\,.\end{aligned} \tag{4.5.16}$$

Finally, we call $\rho(x)dx$ the continuous counterpart of the empirical measure (4.5.8) and we have $\rho(x) = \Pi_{1,1}(x) + \Pi_{1,0}(x)$.

We summarize the basic constraints; when these are violated, the LD rate functionals are identically $+\infty$.

A) The triple $\left(z, \Pi^B, \Pi^I\right)$ is absolutely continuous and $\Pi^B = m(x)\nu^B(x)dx$, and $\Pi^I = (1-m(x))\nu^I(x)dx$ for x dependent probability measures ν^B, ν^I and some $0 \le m(x) \le 1$.

B) The relation (4.5.16) is satisfied a.e.

C) The constraint $z(x) \ge 0$ is satisfied.

D) It holds $m(x) = 0$ when $z(x) > 0$.

Given ν, μ two probability measures on $\{0,1\}\times\{-1,0,+1\}$ we call

$$h(\nu|\mu) = \begin{cases} \sum_{a,b}\nu_{a,b}\log\frac{\nu_{a,b}}{\mu_{a,b}} & \nu \ll \mu\,,\\ +\infty & \text{otherwise}\,,\end{cases}$$

the relative entropy.

Proposition 4.9. *Under the initial condition $B_{1/2}(\eta_1)\delta_{Nz_0}(\zeta_1)$, when (η,ζ) are distributed according to $\mathbb{P}_N^{\mathsf{S},*}$ in (4.5.12), the triple $\left(\hat z_N, \hat\Pi^B_N, \hat\Pi^I_N\right)$ satifies a LD principle with rate functional $\mathfrak{I}^*$, that is $+\infty$ if at least one among the conditions A, B, C, D and $z(0) \ne z_0$ is violated,*

and it is otherwise equal to

$$\mathcal{I}^*\left(z,\Pi^B,\Pi^I\right)=\int_0^1\left[m(x)h\left(\nu^B(x)|\mu^B\right)+(1-m(x))h\left(\nu^I(x)|\mu^I\right)\right]dx\,. \tag{4.5.17}$$

We remark that the dependence on the path z in the rate functional appears through the constraint on the initial condition and the conditions B, C and D. Starting from Proposition 4.9 we can deduce the following LD rate functional whose proof is obtained by applying the Varadhan lemma.

Lemma 4.10. *When the variables (η,ζ) are distributed according to the Markov bridge (4.5.14), then the triple $\left(\hat{z}_N,\hat{\Pi}^B_N,\hat{\Pi}^I_N\right)$ satisfies a LD principle with a rate functional that is $+\infty$ if at least one among the constraints A, B, C and D is violated, and it is otherwise given by*

$$\mathcal{I}^{\mathcal{S}}\left(z,\Pi^B,\Pi^I\right)=z(0)\log\frac{\alpha}{1-\alpha}+z(1)\log\frac{\beta}{1-\beta}+\mathcal{I}^*\left(z,\Pi^B,\Pi^I\right)+c\,, \tag{4.5.18}$$

where c is a suitable constant and $\mathcal{I}^*$ is the rate functional in (4.5.17).

By Lemma 4.8 we can now deduce the LD for the Markov bridge $\mathbb{P}_N^{\mathfrak{S},g,f}$, again using the Varadhan lemma with respect to the continuous map (4.5.9). By the structure of support of the measure μ^B the map on the right-hand side of (4.5.9) for the absolutely continuous triple is given by $\int_0^1 m(x)\nu^B_{0,0}(x)dx$. We therefore have the following.

Lemma 4.11. *When the variables (η,ζ) are distributed according to the Markov bridge $\mathbb{P}_N^{\mathfrak{S},g,f}$, then the triple $\left(\hat{z}_N,\hat{\Pi}^B_N,\hat{\Pi}^I_N\right)$ satisfies a LD principle with rate functional that is $+\infty$ if at least one of the constraints A, B, C and D are violated and it is otherwise given by*

$$\mathcal{I}^{\mathfrak{S}}\left(z,\Pi^B,\Pi^I\right)=\mathcal{I}^{\mathcal{S}}\left(z,\Pi^B,\Pi^I\right)+\left[\int_0^1 m(x)\nu^B_{0,0}(x)dx\right]\log 2+C\,, \tag{4.5.19}$$

where C is a suitable constant.

Strategy of the proof. We give an hint of the strategy of the proof of Proposition 4.9; the following Lemmas 4.10 and 4.11 follow by an application of the Varadhan lemma. The basic property that we use is the fact the measure (4.5.12) depends just on the empirical measures $\hat{\Pi}^I_N$ and $\hat{\Pi}^B_N$.

Let us introduce a family of perturbations of the probability measure (4.5.12): consider two space dependent families of probability measures $\gamma^B(x)=(\gamma^B_{a,b}(x))_{a=0,1}^{b=-1,0,1}$ and $\gamma^I(x)=(\gamma^I_{a,b}(x))_{a=0,1}^{b=-1,0,1}$, and consider the corresponding absolutely continuous measures $\gamma^B(x)dx$ and $\gamma^I(x)dx$ that are elements in $\mathcal{M}\left([0,1],\mathcal{M}^1\left(A\times\{-1,0,+1\}\right)\right)$. For any x, $\gamma^B(x)$ is strictly positive on the support of μ^B and $\gamma^I(x)$ is strictly positive on the support of μ^I and, moreover, the x dependent components are Lipschitz continuous. For any pair γ^B,γ^I we introduce a corresponding probability measure on the sequences (η,ζ) defined by

$$\mathbb{P}_N^{\gamma^I,\gamma^B} := \exp\left\{ N \sum_{a,b} \int_{[0,1]} \left[\hat{\Pi}_N^I(a,b;dx) \log \gamma_{a,b}^I(x) + \hat{\Pi}_N^B(a,b;dx) \log \gamma_{a,b}^B(x) \right] \right\}, \quad (4.5.20)$$

and notice that (4.5.12) corresponds to the cases of γ^B and γ^I constant and constantly equal to μ^B and μ^I respectively.

The probability measure (4.5.20) is absolutely continuous with respect to (4.5.12) and we can compute the Radon-Nikodym derivative that is given by

$$\frac{1}{N} \log \frac{d\mathbb{P}_N^{\gamma^I,\gamma^B}}{d\mathbb{P}_N^{S,*}} = \sum_{a,b} \int_{[0,1]} \left[\hat{\Pi}_N^I(a,b;dx) \log \frac{\gamma_{a,b}^I(x)}{\mu_{a,b}^I} + \hat{\Pi}_N^B(a,b;dx) \log \frac{\gamma_{a,b}^B(x)}{\mu_{a,b}^B} \right]. \quad (4.5.21)$$

The proof of the large deviations principle is based on the law of large numbers for the perturbed measures according to the following fact.

Proposition 4.12. *When (η,ζ) are distributed according to $\mathbb{P}_N^{\gamma^I,\gamma^B}$ with initial condition on the ζ variables given by $\zeta_1 = Nz_0$, then the triple $\left(\hat{z}_N, \hat{\Pi}_N^B, \hat{\Pi}_N^I\right)$ converges in probability to the absolutely continuous triple (z,Π^B,Π^I) such that*

$$\begin{cases} \Pi^B(x) = m(x)\gamma^B(x)\,, \\ \Pi^I(x) = (1-m(x))\gamma^I(x)\,, \end{cases} \quad (4.5.22)$$

and $z(x)$ is the solution of

$$\begin{cases} z(0) = z_0 \\ \dot{z}(x) = m(x)\gamma_{1,1}^B(x) + (1-m(x))(\gamma_{1,1}^I(x) - \gamma_{0,-1}^I(x))\,. \end{cases} \quad (4.5.23)$$

In the above formulas $m(x) = 0$ when $z(x) > 0$ or when $z(x) = 0$ and $\gamma_{1,1}^I(x) - \gamma_{0,-1}^I(x) > 0$, otherwise when $z(x) = 0$ and $\gamma_{1,1}^I(x) - \gamma_{0,-1}^I(x) \le 0$, then $0 \le m(x) \le 1$ is the unique value such that

$$m(x)\gamma_{1,1}^B(x) + (1-m(x))(\gamma_{1,1}^I(x) - \gamma_{0,-1}^I(x)) = 0\,; \quad (4.5.24)$$

this means

$$m(x) = \frac{\left[\gamma_{0,-1}^I(x) - \gamma_{1,1}^I(x)\right]_+}{\gamma_{011}^B(x) + \gamma_{0,-1}^I(x) - \gamma_{1,1}^I(x)} \mathbb{1}(z(x) = 0)\,, \quad (4.5.25)$$

where $[\cdot]_+$ denotes the positive part.

Proof. We give an outline of the proof; see [32, 56] for results of this type. Due to the Lipschitz continuous assumption on γ^B, γ^I we have existence and uniqueness of the solution of (4.5.23). We already showed that any limit triple need to be absolutely continuous and, moreover, by independent sampling any limit measure has to be of the form (4.5.22) for a suitable $m(x)$. The proof of the lemma for the intervals on which $z(x) > 0$ follows directly since it is a special case of the classic fluid limit [50, 56] with $m = 0$. On the intervals on which $z(x) = 0$ and $\gamma_{1,1}^I(x) - \gamma_{0,-1}^I(x) \le 0$, the result follows since any scaling limit has to be absolutely continuous,

by (4.5.15) it has to satisfy (4.5.16), and there is an unique $m(x)$ satisfying (4.5.24). □

Exponential tightness To show exponential tightness we need to construct a family of compact sets $(\mathcal{K}_\alpha)_{\alpha\in\mathbb{R}^+}$ such that

$$\limsup_{N\to+\infty}\frac{1}{N}\log P_N^{\mathcal{S},*}\left((\hat z_N,\hat\Pi^B,\hat\Pi^I)\in\mathcal{K}_\alpha\right)<-\alpha\,.$$

This statement allows to extend the following upper bound from compact set to closed ones. This is a technical issue that we do not discuss in detail and that can be treated considering $\mathcal{K}_\alpha=\mathcal{K}_\alpha^z\times\mathcal{K}_\alpha^B\times\mathcal{K}_\alpha^I$ and following the arguments in [31] Lemma 5.1.7 and [49] Section 4 of Chapter 10.

Lower bound Consider a triple (z,Π^B,Π^I) that satisfies the conditions A, B, C and D and such that $\Pi^B(x)=m(x)\gamma^B(x)$ and $\Pi^I(x)=(1-m(x))\gamma^I(x)$ with γ^B,γ^I that are strictly positive and Lipschitz continuous. Then, by Proposition 4.12, it follows that, with respect to the measures $\mathbb{P}_N^{\gamma^I,\gamma^B}$, the triple $\left(\hat z_N,\hat\Pi_N^B,\hat\Pi_N^I\right)$ converges in probability to the triple (z,Π^B,Π^I) and, moreover, by the formula (4.5.21) we have

$$\lim_{N\to+\infty}\frac{1}{N}h\left(\mathbb{P}_N^{\gamma^I,\gamma^B}\middle|\mathbb{P}_N^{\mathcal{S},*}\right)=\lim_{N\to+\infty}\frac{1}{N}\mathbb{E}_{\mathbb{P}_N^{\gamma^I,\gamma^B}}\left[\log\frac{d\mathbb{P}_N^{\gamma^I,\gamma^B}}{d\mathbb{P}_N^{\mathcal{S},*}}\right]=\mathcal{I}^*\left(z,\Pi^B,\Pi^I\right).\tag{4.5.26}$$

Let $\tilde{\mathcal{I}}^*$ be the functional equal to $\mathcal{I}^*$ for measures $\Pi^B=m\nu^B$ and $\Pi^I=(1-m)\nu^I$ with ν^I,ν^B that are strictly positive and Lipschitz continuous and equal to $+\infty$ otherwise. By Theorem 3.4 in [52] we obtain a LD lower bound with rate functional that is $\tilde{\mathcal{I}}^*_{\text{lse}}$, the lower semicontinuous envelope of $\tilde{\mathcal{I}}^*$. The lower bound is obtained by showing that $\tilde{\mathcal{I}}^*_{\text{lse}}=\mathcal{I}^*$.

Upper bound The upper bound is obtained by the classic argument; by the exponential tightness we can consider compact subsets $\mathcal{K}$. We have the following estimates. Let γ^I,γ^B be positive x dependent probability measures, then we have the following classic steps

$$\begin{aligned}&\mathbb{P}_N^{\mathcal{S},*}\left((\hat z_N,\hat\Pi_N^B,\hat\Pi_N^I)\in\mathcal{K}\right)=\mathbb{E}_{\mathbb{P}_N^{\gamma^I,\gamma^B}}\left[\frac{d\mathbb{P}_N^{\mathcal{S},*}}{d\mathbb{P}_N^{\gamma^I,\gamma^B}}\mathbb{1}\left((\hat z_N,\hat\Pi_N^B,\hat\Pi_N^I)\in\mathcal{K}\right)\right]\\ \le\;&\exp\left\{-\inf_{\{(\pi,\Pi^B,\Pi^I)\in\mathcal{K}\}}\sum_{a,b}\int_{[0,1]}\left[\Pi^I(a,b;dx)\log\frac{\gamma^I_{a,b}(x)}{\mu^I_{a,b}}+\Pi^B(a,b;dx)\log\frac{\gamma^B_{a,b}(x)}{\mu^B_{a,b}}\right]\right\}.\end{aligned}$$

Since the above estimate is true for any positive γ^B,γ^I we can optimize getting

$$\begin{aligned}&\limsup_{N\to+\infty}\frac{1}{N}\log\mathbb{P}_N^{\mathcal{S},*}\left((\hat z_N,\hat\Pi_N^B,\hat\Pi_N^I)\in\mathcal{K}\right)\le\\ -\;&\sup_{\{\gamma^B,\gamma^I\}}\inf_{\{(\pi,\Pi^B,\Pi^I)\in\mathcal{K}\}}\sum_{a,b}\int_{[0,1]}\left[\Pi^I(a,b;dx)\log\frac{\gamma^I_{a,b}(x)}{\mu^I_{a,b}}+\Pi^B(a,b;dx)\log\frac{\gamma^B_{a,b}(x)}{\mu^B_{a,b}}\right]\\ =\;&-\inf_{\{(\pi,\Pi^B,\Pi^I)\in\mathcal{K}\}}\sup_{\{\gamma^B,\gamma^I\}}\sum_{a,b}\int_{[0,1]}\left[\Pi^I(a,b;dx)\log\frac{\gamma^I_{a,b}(x)}{\mu^I_{a,b}}+\Pi^B(a,b;dx)\log\frac{\gamma^B_{a,b}(x)}{\mu^B_{a,b}}\right]\end{aligned}$$

$$= \quad - \inf_{\{(\pi,\Pi^B,\Pi^I)\in\mathcal{K}\}} \mathcal{I}^*\Big(z,\Pi^B,\Pi^I\Big)\,,$$

and this is the LD upper bound. In the first inequality we optimized over $\gamma^B\gamma^I$; in the second equality we used a minmax lemma (see for example lemmas 3.2 and 3.3 in Appendix 2 of [49]) to exchange infimum and supremum, and the last equality follows from a direct computation by using Lagrange multipliers.

4.5.4 The contraction

By contraction we can deduce the LD rate functional for the empirical measure of the TASEP configuration (4.5.8). We proceed by successive minimizations. We first contract the rate functional $\mathcal{I}^{\mathfrak{S}}(z,\Pi^B,\Pi^I)$ to the rate functional $\mathfrak{I}^{\mathfrak{S}}(z,\Pi)$ for the pair $\Big(\hat{z}_N,\hat{\Pi}_N\Big)$; namely

$$\mathfrak{I}^{\mathfrak{S}}(z,\Pi) = \inf_{\{\Pi^B+\Pi^I=\Pi\}} \mathcal{I}^{\mathfrak{S}}(z,\Pi^B,\Pi^I)\,. \tag{4.5.27}$$

We have the following result.

Proposition 4.13. *When the variables (η,ζ) are distributed according to the Markov bridge (4.5.14), then the pair $\Big(\hat{z}_N,\hat{\Pi}_N\Big)$ satisfies a LD principle with a rate functional that is $+\infty$ if at least one among the constraints A, B, C and D is violated and it is otherwise given by*

$$\mathfrak{I}^{\mathfrak{S}}\Big(z,\Pi\Big) = z(0)\log\frac{\alpha}{1-\alpha} + z(1)\log\frac{\beta}{1-\beta} + \int_0^1 h\left(\Pi(x)|\mu^I\right)\,dx \;\; + C\,, \tag{4.5.28}$$

where C is the same constant in Lemma 4.11.

Proof. Since z is fixed, to perform the minimization in (4.5.27), we need to minimize, for each x such that $z(x)=0$, the following expression

$$\left[m(x)h\left(\nu^B(x)|\mu^B\right) + (1-m(x))h\left(\nu^I(x)|\mu^I\right)\right] + m(x)\nu^B_{0,0}(x)\log 2\,, \tag{4.5.29}$$

under the condition

$$\Big\{m(x),\nu^B(x),\nu^I(x)\;:\; m(x)\nu^B(x)+(1-m(x))\nu^I(x)=\Pi(x)\Big\} \tag{4.5.30}$$

where we recall that $\sum_{a,b}\nu^B_{a,b}(x)=1$ and $\sum_{a,b}\nu^I_{a,b}(x)=1$. When $z(x)>0$ we do not need to do a minimization since $m(x)$ is constrained to be zero.

By writing the variational conditions with the Lagrange multipliers for the variational problem (4.5.29), namely (4.5.30) with also the normalization of ν^B and ν^I, one finds that the resulting system of equations has no critical points for any $m(x)\in(0,1)$. The minimizer must be therefore on the boundary, i.e. corresponding to $m(x)=0$ or $m(x)=1$. Note that if $m(x)=0$ it follows that $\Pi(x)=\nu^I(x)$ while if $m(x)=1$ then $\Pi(x)=\nu^B(x)$.

If the probability measure $\Pi(x)$ satisfies $\Pi_{0,-1}(x)>0$ then the minimum is necessarily attained at $m(x)=0$. Indeed, when $m(x)=1$ we have $\Pi(x)=\nu^B(x)$ and we have $\nu^B_{0,-1}(x)=0$ for any ν^B ; hence the value $m(x)=1$ is not admissible when $\Pi_{0,-1}(x)>0$.

When $\Pi_{0,-1}(x) = 0$ we obtain the same value if we compute (4.5.29) for $m(x) = 0$ and $m(x) = 1$; therefore, also in this case, we can consider $m(x) = 0$. We conclude that the large deviations rate functional $\mathfrak{I}^{\mathfrak{S}}$ is obtained having $\Pi^B = 0$ and $\Pi^I = \Pi$ as the minimizer in (4.5.27), thus obtaining (4.5.28) as a result. □

In Proposition 4.13 the constant C that appears in the large deviation rate functional is fixed by the normalization factor of the Markov bridge $\mathbb{P}_N^{\mathfrak{S},g,f}$, i.e., by $\mathfrak{Z}_N$; we recall that this constant C coincides with the one in Lemma 4.11. By construction, $\mathfrak{Z}_N$ is given by $\mathfrak{Z}_N = \frac{Z_N}{4^N}$ with $Z_N = \vec{y}(M^0 + M^1)^N\vec{x}$. More precisely, C is determined by the asymptotic behaviour of Z_N (see [23, 51]), up to the additive constant $-\log 4$. In particular $C := C(\alpha, \beta) = -\log 4 - \log(\bar\rho(1-\bar\rho))$, where $\bar\rho$ is the limiting particle density:

$$\bar\rho = \begin{cases} \frac{1}{2} & \alpha \geq 1, \quad \beta \geq 1 \\ \alpha & \alpha > \frac{1}{2}, \quad \beta > \alpha \\ 1-\beta & \beta > \frac{1}{2}, \quad \alpha > \beta. \end{cases}$$

Our next step is to deduce by contraction the LD rate functional $I^{\mathfrak{S}}(z,\rho)$ for the pair $(\hat z_N, \hat\pi_N)$ from (4.5.28). Let us introduce the function $H(x) := x\log x + (1-x)\log(1-x)$.

Proposition 4.14. *When the variables (η, ζ) are distributed according to the Markov bridge (4.5.14), then the pair $(\hat z_N, \hat\pi_N)$ satisfies a LD principle with rate functional $I^{\mathfrak{S}}(z,\rho)$ that is $+\infty$ if ρ and z are not absolutely continuous or if condition C is violated. Moreover, the rate functional is also $+\infty$ if, for almost all x, the densities $(\dot z(x), \rho(x))$ do not belong to the region $R \subset \mathbb{R}^2$ defined by*

$$R = \left\{(\dot z, \rho) \,:\, 0 \leq \rho \leq 1\,, 0 \leq \rho - \dot z \leq 1\right\}. \tag{4.5.31}$$

In the remaining cases the rate functional is finite and it is given by

$$I^{\mathfrak{S}}(z,\rho) = z(0)\log\frac{\alpha}{1-\alpha} + z(1)\log\frac{\beta}{1-\beta} + \int_0^1 \left[H\Big(\rho(x)\Big) + H\Big(\rho(x) - \dot z(x)\Big)\right] dx + C' \tag{4.5.32}$$

where C' is a normalization constant.

Proof. By the contraction principle, the proof follows by a direct minimization of (4.5.28) over Π with the constraints

$$\begin{cases} \rho(x) = \Pi_{1,1}(x) + \Pi_{1,0}(x) \\ \dot z(x) = \Pi_{1,1}(x) - \Pi_{0,-1}(x)\,. \end{cases} \tag{4.5.33}$$

The geometric constraints (4.5.31) follow by the above relation and the fact that Π is a probability measure. □

A direct computation shows that the constant in (4.5.32) is given by $C' = C + \log 4$ where C is the constant in (4.5.28).

We can now obtain the LD rate function $I^{\mathrm{TASEP}}(\rho)$ for the empirical measure (4.5.8), when η is distributed according to the invariant measure of boundary driven TASEP, by contraction

from (4.5.32), i.e. we have $I^{\text{TASEP}}(\rho) = \inf_z I^{\mathfrak{S}}(z,\rho)$. We define:

$$a = \frac{1-\alpha}{\alpha}, \quad b = \frac{1-\beta}{\beta}$$

with this notation we are considering the regime of TASEP corresponding to $ab < 1$. This last contraction gives as a result the rate functional in [25, 28]. Our representation is obtained as an infimum and we write the final expression in the same form of [6] where its equivalence with the representation in [25, 28] is shown.

Theorem 4.15. *Let $F(x) = \int_0^x \rho(y)dy$ and $\mathcal{G}$ the set of absolutely continuous functions G such that $G(0) = 0$ and $0 \leq \dot{G}(x) \leq 1$. Let η be distributed according to the invariant measure of the boundary driven TASEP when the boundary parameters are such that $ab < 1$. We have that the large deviations rate functional for $\hat{\pi}_N$ in (4.5.8) is given by*

$$I^{\text{TASEP}}(\rho) = \inf_{G\in\mathcal{G}} \Bigg\{ \int_0^1 \Big[H\left(\dot{F}(y)\right) + H\left(\dot{G}(y)\right) \Big] dy + \ln(ab) \min_{x\in[0,1]} \Big[F(x) - G(x) \Big] - \ln(b) \Big(F(1) - G(1) \Big) \Bigg\}.$$

Proof. By the contraction principle we have to minimize the functional (4.5.32) over the functions $z(x) \geq 0$ such that $0 \leq \rho(x) - \dot{z}(x) \leq 1$ for almost any $x \in [0,1]$.

We define:

$$F(x) = \int_0^x \rho(y)\, dy$$
$$G(x) = F(x) - \Big(z(x) - z(0) \Big),$$

and note that $G(0) = 0$ and $\dot{G}(x) = \rho(x) - \dot{z}(x)$ so that $G \in \mathcal{G}$; moreover $F(x) + z(0) - G(x) = z(x) \geq 0$. The functional (4.5.32) can be written in terms of $F, G, z(0)$ as

$$\int_0^1 \Big[H\left(\dot{F}(y)\right) + H\left(\dot{G}(y)\right) \Big] dy - z(0)\ln(ab) - \Big[F(1) - G(1) \Big] \ln(b) := \mathcal{L}(F,G,z(0))\,.$$

By the contraction principle we have

$$I^{\text{TASEP}}(\rho) = \inf_{\{z(0)\geq 0\}} \inf_{\{G\in\mathcal{G}: G(x)\leq F(x)+z(0)\}} \mathcal{L}(F,G,z(0))\,.$$

We exchange the order of the infima obtaining that the above variational problem is equivalent to

$$I^{\text{TASEP}}(\rho) = \inf_{\{G\in\mathcal{G}\}} \inf_{\{z(0)\geq \max_{x\in[0,1]}[G(x)-F(x)]\}} \mathcal{L}(F,G,z(0))\,.$$

Since $ab < 1$ the value of $z(0)$ that minimizes the first infimum is exactly $z(0) = \max_{x\in[0,1]}[G(x) - F(x)]$ and this finishes the proof. □

4.6 Rate functional as a double entropy

In the previous section, we proved a large deviations principle using the joint Markov property. Here, we instead exploit the mixture representation, which leads to a different and simpler characterization of the rate functional.

We restrict our analysis to the simplest case, namely finite alphabets and under the assumption of irreducibility; a more general analysis will be discussed in a forthcoming publication [45]. In the finite-dimensional irreducible case, the resulting LDP does not depend on the boundary vectors $\vec{x}$ and $\vec{y}$. For this reason, we consider the simplest situation that corresponds to periodic boundary conditions.

Consider a double sequence $(\eta;\zeta) = \Big(\eta_1,\dots\eta_{N+1};\zeta_1,\dots,\zeta_{N+1}\Big)$ (we use $N+1$ instead of N in order to remain consistent with the previous cases). We introduce a shift and a periodization of these sequences, so that the variables ζ_i are associated with the vertices of a discrete ring with $N+1$ vertices, while the variables η_i are associated with the edges, with η_i corresponding to the edge connecting vertices i and $i+1$.

Consider $\mathcal{M}^1(B^2\times A)$ as the set of probability measures on local configurations of the double sequence consisting of two adjacent vertices (conventionally referred to as the left and right ones) together with the edge connecting them.

For $G\in\mathcal{M}^1(B^2\times A)$ we denote by $G(b,b';a)$ the joint probability to observe the pair (b,b') at the left and right vertices, respectively, and the value a on the corresponding edge. Given a double sequence (η,ζ), we associate with it the empirical measure $\hat{G}_N = \hat{G}_N[\eta,\zeta]$ defined by

$$\hat{G}_N(b,b';a) := \frac{1}{N+1}\sum_{i=1}^{N+1} \mathbb{1}(\zeta_i = b, \zeta_{i+1} = b', \eta_i = a)\,, \tag{4.6.1}$$

where the indices are taken modulus $N+1$.

Since the variables η are distributed according to a measure of MPA type and the variables ζ play the role of hidden variables, we can canonically associate to the collection of $B\times B$ matrices $(M^a)_{a\in A}$ a single stochastic $B\times B$ matrix S and a family $\Big(p_{b,b'}(\cdot)\in\mathcal{M}^1(A)\Big)_{b,b'\in B}$. A key feature of the MPA measure is that the probability of any double sequence (η,ζ) depends only on the empirical measure defined in (4.6.1). More precisely, one has

$$\mathbb{P}(\eta,\zeta) = K_N \prod_{b,b'\in B}\prod_{a\in A}\big(S_{b,b'}p_{b,b'}(a)\big)^{(N+1)\hat{G}_N(b,b';a)}\,, \tag{4.6.2}$$

where K_N is a suitable normalization constant. The above formula is exact when MPA measure is considered on a ring.

Formula (4.6.2) can be interpreted as a generalization of both the independence and Markov conditions. Indeed, for independent variables the probability of any sequence depends only on the frequencies of symbols, while for Markov measures it depends only on the frequencies of pairs. Moreover, these properties characterize the corresponding classes of models in terms of De Finetti's Theorem and its generalizations.

The MPA variables together with the associated hidden variables have a joint probability that depends only on the special frequencies defined in (4.6.1).

From formula (4.6.2), one can obtain the large deviations rate functional for the empirical

measure $\hat{G}_N$ by a direct combinatorial computation. This empirical measure is the natural object to consider, since the probability of each sequence depends only on $\hat{G}_N$. This is exactly the the same reason why, in the study of large deviations for Markov chains, the natural object to consider is the pair empirical measure.

The approach used to deduce large deviations principle is analogous to the one used in the direct proof of LD for the pair empirical measure of Markov chains (see [19]). We do not present the complete argument, which can be obtained by a direct extension of the proof for the Markov case given in [19]. We instead focus on the main combinatorial estimates from which, by applying Stirling's approximation and suitable discretization arguments, a full large deviations principle can be derived.

Let us call $\hat{g}_N(b,b') := \sum_a \hat{G}_N(b,b';a)$ the marginal of the empirical measure $\hat{G}_N$ on the B-variables. The basic idea is the following: given a probability measure G on $B^2 \times A$, we can factorize the probability as follows

$$\mathbb{P}\left(\left\{(\eta,\zeta) : \hat{G}_N \sim G\right\}\right) \simeq \left|\left\{(\eta,\zeta) : \hat{G}_N \sim G\right\}\right| K_N \prod_{b,b'\in B}\prod_{a\in A}\left(S_{b,b'}p_{b,b'}(a)\right)^{(N+1)G(b,b';a)} . \tag{4.6.3}$$

In the above formula, the symbol $\sim$ denotes closeness with respect to a suitable norm, while $\simeq$ represents the same asymptotic behavior at exponential scale. The factorization follows from (4.6.2): all the sequences (η,ζ) whose empirical measure satisfies $\hat{G}_N \sim G$ have the same probability, given by the second factor on the right-hand side of (4.6.3). To obtain the asymptotic behavior, it therefore remains to estimate the first combinatorial factor on the right-hand side of (4.6.3). We stress that formula (4.6.3) is written in an informal way but there are no real difficulties in translating into a rigorous argument in this finite-dimensional setting, as done in [19].

We obtain this estimate by first estimating the number of sequences ζ whose empirical measure $\hat{g}_N$ is equal to a fixed measure g. We note that, by construction, both $\hat{g}_N$ and g must belong to $\mathcal{M}^1_{stat}(B^2)$. This is a classical combinatorial problem, discussed for example in the BEST theorem in [19]. It follows that

$$\left|\{\zeta : \hat{g}_N \sim g\}\right| \simeq \frac{\prod_b \left[(N+1)g_1(b)\right]!}{\prod_{b,b'}\left[(N+1)g(b,b')\right]!}, \tag{4.6.4}$$

where we used the notation $g_1(b) = \sum_{b'} g(b,b') = \sum_{b'} g(b',b)$ for the one marginal of g. By Stirling formula we obtain

$$\left|\{\zeta : \hat{g}_N \sim g\}\right| \simeq \exp\left\{-(N+1)\left[\sum_{b,b'} g(b,b')\log\frac{g(b,b')}{g_1(b)}\right]\right\} . \tag{4.6.5}$$

Consider now a fixed configuration ζ such that the corresponding empirical measure $\hat{g}_N$ is equal to a given measure g.Let G be a measure having g as its B-marginal, and denote by $\mathcal{N}_\zeta(G)$ the following set

$$\mathcal{N}_\zeta(G) := \left\{\eta \,:\, \hat{G}_N \sim G\,;\ \zeta \text{ fixed}\right\} . \tag{4.6.6}$$

Since in the fixed configuration ζ there are exactly $(N+1)g(b,b')$ pairs $(\zeta_i,\zeta_{i+1})=(b,b')$, the estimation of the cardinality of (4.6.6) follows from a multinomial computation. More precisely, we have

$$\big|\mathcal{N}_\zeta(p)\big| \simeq \prod_{b,b'} \frac{(N+1)g(b,b')}{\prod_a (N+1)G(b,b';a)}\,. \tag{4.6.7}$$

Using Stirling's formula again, we obtain

$$\begin{aligned}
\Big|\Big\{(\eta,\zeta):\hat G_N\sim G\Big\}\Big| &= \big|\{\zeta:\hat g_N\sim g\}\big|\times\big|\mathcal{N}_\zeta(p)\big| \\
&\simeq \exp\left\{-(N+1)\left[\sum_{b,b'} g(b,b')\log\frac{g(b,b')}{g_1(b)} + \sum_{b,b',a} G(b,b';a)\log\frac{G(b,b';a)}{g(b,b')}\right]\right\} \\
&= \exp\left\{-(N+1)\left[\sum_{b,b',a} G(b,b';a)\log\frac{G(b,b';a)}{g_1(b)}\right]\right\}.
\end{aligned}$$

Combining this estimate with (4.6.3) we obtain the following estimate

$$\mathbb{P}\left(\Big\{(\eta,\zeta):\hat G_N\sim G\Big\}\right) \simeq K_N \exp\left\{-(N+1)\left[\sum_{b,b',a} G(b,b';a)\log\frac{G(b,b';a)}{g_1(b)S_{b,b'}p_{b,b'}(a)}\right]\right\}; \tag{4.6.8}$$

This implies that the large deviations rate functional for $\hat G_N$ is given by

$$\sum_{b,b',a} G(b,b';a)\log\frac{G(b,b';a)}{g_1(b)S_{b,b'}p_{b,b'}(a)}\,, \tag{4.6.9}$$

where the normalization factor K_N does not contribute, since the above functional is non-negative and attains its minimum value zero. Calling $p^G_{b,b'}(a) := \frac{G(b,b';a)}{g(b,b')}$ we can rewrite (4.6.9) as a double entropy

$$\sum_{b,b'} g(b,b')\log\frac{g(b,b')}{g_1(b)S_{b,b'}} + \sum_{b,b'} g(b,b')\sum_a p^G_{b,b'}(a)\log\frac{p^G_{b,b'}(a)}{p_{b,b'}(a)}\,. \tag{4.6.10}$$

In (4.6.10) the first term is the rate functional of the pair empirical measure associated with the Markov process ζ, while the second term is the average of the relative entropy between the conditional law p^G of the A variable and the conditional law p of the MPA distribution.

A special case occurs when all the matrices $(M^a)_{a\in A}$ have the same eigenvector, with associated eigenvalues λ_a. In this case, we have $S=\sum_a q_a S^a$ where S^a are the stochastic matrices associated with the matrices M^a and $q_a=\frac{\lambda_a}{\lambda}$. By direct computation, using that $p_{b,b'}(a)=q_a\frac{S^a_{b,b'}}{S_{b,b'}}$, we obtain

$$\mathbb{P}(\eta,\zeta) = K_N \prod_{i=1}^{N+1} q_{\eta_i} S^{\eta_i}_{\zeta_i,\zeta_{i+1}}\,. \tag{4.6.11}$$

Note that this form of the distribution is rather special. For example, if one considers the case where vectors x, y are identically equal to one, the marginal distribution of the variables η is a product measure of single-site measures equal to q. Let us now show, as we already did using the different approach, that even if the distribution of η is not exactly a product measure in the periodic case, the associated large deviations rate functional coincides with that of the product measure. We have the following

Proposition 4.16. *The infimum of the rate functional* (4.6.9) *over*

$$\mathcal{G}_\nu := \left\{ G \,:\, \sum_{b,b'} G(b,b';a) = \nu(a) \right\},$$

for models for which we have (4.6.11)*, is given by*

$$\sum_a \nu(a) \log \frac{\nu(a)}{\lambda_a}\,. \tag{4.6.12}$$

Proof. All the measures $G \in \mathcal{G}_\nu$ can be written in the form $G(b,b';a) = \nu(a)Q_a(b,b')$ where $\big(Q_a(\cdot,\cdot) \in \mathcal{M}^1(B^2)\big)_{a\in A}$ is an arbitrary family. When (4.6.11) holds, then (4.6.9) can be written as

$$\sum_{b,b',a} G(b,b';a) \log \frac{G(b,b';a)}{g_1(b) S^a_{b,b'} q_a}\,. \tag{4.6.13}$$

Using $G \in \mathcal{G}_\nu$ this becomes

$$\sum_a \nu(a) \sum_{b,b'} Q_a(b,b') \log \frac{Q_a(b,b')}{g_1(b) S^a_{b,b'}} + \sum_a \nu(a) \log \frac{\nu(a)}{\lambda_a}\,. \tag{4.6.14}$$

The first term is greater than or equal to zero and vanishes if and only if $Q_a(b,b') = g_1(b) S^a_{b,b'}$. In this case (4.6.12) holds, and the result is proved. □

Bibliography

[1] Angel O. *The stationary measure of a 2-type totally asymmetric exclusion process* Journal of Combinatorial Theory, Series A, Vol. **113**, Issue 4, 625-635, (2006).

[2] Barraquand G., Corwin I., Yang, Z. *Stationary measures for integrable polymers on a strip.* Invent. math. **237**, 1567–1641 (2024).

[3] Bertoni A., Choffrut C., Goldwurm M., Lonati V. *On the number of occurrences of a symbol in words of regular languages.* Theoret. Comput. Sci., **302**, 431–456, 2003

[4] Blythe R A, Evans M. R. *Nonequilibrium steady states of matrix-product form: a solver's guide* 2007 J. Phys. A: Math. Theor. **40** R333.

[5] Blythe R A, Evans M. R., Colaiori F., Essler F.H.L. *Exact solution of a partially asymmetric exclusion model using a deformed oscillator algebra* 2000 J. Phys. A: Math. and Gen. **33** (12) 2313-2332

[6] Bryc W., Zatitskii P. *A two-line representation of stationary measure for open TASEP* Electron. J. Probab. **29**: 1-24 (2024).

[7] Bertini L., De Sole A., Gabrielli D., Jona-Lasinio G., Landim C. *Stochastic interacting particle systems out of equilibrium.* J. Stat. Mech. (2007) P07014.

[8] Bertini L., De Sole A., Gabrielli D., Jona-Lasinio G., Landim C. *Macroscopic fluctuation theory for stationary non-equilibrium states.* Rev. Mod. Phys. **87**, 593–636 (2015).

[9] Cramér H., *Sur un nouveau théorème limite dans la théorie des probabilités* Colloque consacré à la théorie des probabilités, Hermann, 1938, **3**, 2-23.

[10] Carinci G., Franceschini C., Frassek R., Giardinà C., Redig F. *Large Deviations and Additivity Principle for the Open Harmonic Process.* Commun. Math. Phys. **406**, 103 (2025)

[11] Carinci G., Franceschini C., Gabrielli D., Giardinà C. , Tsagkarogiannis D. *Solvable Stationary Non Equilibrium States.* J Stat Phys **191**, 10 (2024)

[12] Capanna M., Gabrielli D., Tsagkarogiannis D. *On a Class of Solvable Stationary Non Equilibrium States for Mass Exchange Models.* J Stat Phys **191**, 25 (2024).

[13] Cigler J. *Continuous q-Hermite polynomials: An elementary approach.* arXiv preprint, arXiv:1307.0357 (2013).

[14] Clifford P. *Markov random fields in statistics.* In: *Disorder in Physical Systems*, Oxford University Press (1990).

[15] Corteel S., Williams L. K. *Tableaux combinatorics for the asymmetric exclusion process* Advances in Applied Mathematics **39** (2007), 293-310

[16] Corteel S., Williams L. K. *Tableaux combinatorics for the asymmetric exclusion process and Askey-Wilson polynomials* Duke Mathematical Journal, 2011 **159** (3), 385-427

[17] Corteel S., Mandelshtam O., Williams L. K. *Combinatorics of the two-species ASEP and Koornwinder moments* Advances in Mathematics, 2017 **321**, 160-204

[18] De Masi A., Ferrari P.A., Gabrielli D. *Hidden Temperature in the KMP Model.* J Stat Phys **191**, 150 (2024).

[19] den Hollander F. *Large Deviations.* Fields Institute Monographs, vol. 14, American Mathematical Society, Providence, RI (2000).

[20] Derrida B. *Non-equilibrium steady states: fluctuations and large deviations of the density and of the current.* J. Stat. Mech. (2007) P07023.

[21] Derrida B., Domany E., Mukamel D. *An Exact Solution of a One-Dimensional Asymmetric Exclusion Model with Open Boundaries.* Journal of Statistical Physics, 1992, **69** (3-4), pp.667-687.

[22] Derrida B., Janowsky S. A., Lebowitz J. L., Speer E. R. *Exact solution of the totally asymmetric simple exclusion process: Shock profiles.* Journal of Statistical Physics, 1993, **73** (5-6), pp.813-842.

[23] Derrida B., Evans M. R., Hakim V., Pasquier V. *Exact solution of a 1D asymmetric exclusion model using a matrix formulation.* Journal of Physics A: Mathematical and General, 1993, **26** (7), pp.1493-1517.

[24] Derrida B., Enaud C. *Large Deviation Functional of the Weakly Asymmetric Exclusion Process.* Journal of Statistical Physics 2004, **114** (3), pp.537-562.

[25] Derrida B., Lebowitz J.L., Speer E.R. *Free Energy Functional for Nonequilibrium Systems : An Exactly Soluble Case Phys.* Rev. Lett. **87**, 150601 (2001).

[26] Derrida B., Lebowitz J.L., Speer E.R. *Exact free energy functional for a driven diffusive open stationary nonequilibrium system* Phys. Rev. Lett. **89**, 030601 (2002)

[27] Derrida B., Lebowitz J.L., Speer E.R. *Large deviation of the density profile in the steady state of the symmetric simple exclusion process.* J. Stat. Phys. **107**, 599–634 (2002).

[28] Derrida B., Lebowitz J.L., Speer E.R. *Exact large deviation functional of a stationary open driven diffusive system: the asymmetric exclusion process.* J. Stat. Phys. **110**, 775–810 (2003).

[29] Donsker M.D., Varadhan S.R.S. *Asymptotic evaluation of certain Markov process expectations for large time.* Comm. Pure Appl. Math. *(I)*, **28**, 1-47 (1975), *(II)*, **28**, 279-301 (1975), *(III)*, **29**, 389-461 (1976), *(IV)*, **36**, 183-212 (1983).

[30] Duchi E., Schaeffer G. *A combinatorial approach to jumping particles*, J Combin Theory Series A **110** (2004), 1–29

[31] Dembo A., Zeitouni O. *Large Deviations Techniques and Applications.* 2nd ed., Springer, New York (1998).

[32] Ethier S. N., Kurtz T. *Markov processes. Characterization and convergence* Wiley Ser. Probab. Math. Statist. Probab. Math. Statist. John Wiley and Sons, Inc., New York, 1986.

[33] Essler F.H.L., Rittenberg V. *Representations of the Quadratic Algebra and Partially Asymmetric Diffusion with Open Boundaries* J of Physics A: Mathematical and General **29**, (1996) 3375-3407

[34] Evans M.R., Ferrari P.A., Mallick K. *Matrix representation of the stationary measure for the multispecies TASEP* J Stat Phys **135**, 217–239 (2009)

[35] Fannes M., Nachtergaele B., Slegers L. *Functions of Markov processes and algebraic measures.* Rev. Math. Phys **4** (1), 39-64 (1992).

[36] Ferrari P.A. *Shock fluctuations in asymmetric simple exclusion* J Probability Theory and Related Fields, 1992, **91**, 81-101

[37] Ferrari P.A., Fontes L.R.G., Kohayakawa Y. *Invariant measures for a two-species asymmetric process.* J Stat Phys **76**, 1153–1177 (1994)

[38] Ferrari P.A., Martin J.B. *Stationary distributions of multi-type totally asymmetric exclusion processes.* Ann. of Prob., Vol. **35**, No. 3, 807–832 (2007)

[39] Franceschini C., Frassek R., Giardinà C. *Integrable heat conduction model.* J. Math. Phys **64**, 6, (2023): 063301.

[40] Frassek R. *Steady state representations for the harmonic process* arXiv:2507.20436

[41] Frassek R., Giardinà C. *Exact solution of an integrable non-equilibrium particle system.* J. Math. Phys **63.10** (2022): 103301.

[42] Frassek R., Giardinà C., Kurchan J. *Non-compact quantum spin chains as integrable stochastic particle processes.* J. Stat. Phys., **180**, 1-6, 135–171 (2020).

[43] Gabrielli D., Iacovissi F. *Mixtures, Markov bridges and the matrix product ansatz*, preprint arXiv:2509.16455

[44] Gabrielli D., Iacovissi F. *Large deviations and the matrix product ansatz*, preprint arXiv:2601.11820.

[45] Gabrielli D., Goldwurm M., Iacovissi F., Mignosi F. in preparation.

[46] Giardinà C., Redig F., Van Tol, B. *Intertwining and Propagation of Mixtures for Generalized KMP Models and Harmonic Models.* J Stat Phys **192**, 21 (2025).

[47] Goldwurm M., Radicioni R. *Probabilistic models for pattern statistics*, RAIRO-Theor. Inf. Appl. **40**, Number 2, 207-225, (2006).

[48] Goldwurm M., Vignati M. *Large deviation properties for pattern statistics in primitive rational models.* Theoret. Comput. Sci., vol. **1030**, article 115055, (2025).

[49] Kipnis C., Landim C. *Scaling Limits of Interacting Particle Systems* Springer–Verlag, Berlin, Heidelberg (1999).

[50] Kitchens B.P. *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts.* Universitext Springer-Verlag, Berlin, 1998.

[51] Ligget T. M. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* Springer, Grundlehren Math. Wiss. **324**, 1999

[52] Mariani M. *A Γ-convergence approach to large deviations.* Annali Scuola Normale Superiore - Classe di Scienze, **18** (3), 951-976, (2018).

[53] Mallick K., Sandow, S. *Finite-dimensional representations of the quadratic algebra: Applications to the exclusion process*Journal of Physics A: Mathematical and General, **30**, 13 (1997), 4513–4526.

[54] Redig F., Van Tol B. *Non-equilibrium Steady States with a Spatial Markov Structure.* J Stat Phys, **192**, 89 (2025).

[55] Schütz G. M., Domany E. *Phase Transitions in an Exactly Soluble One-Dimensional Exclusion Process.* J. Stat. Phys., **72**, (1-2), 277-296 (1993).

[56] Shwartz A., Weiss A. *Large deviations for performance analysis. Queues, communications, and computing. With an appendix by Robert J. Vanderbei* Stochastic Model. Ser. Chapman and Hall, London, 1995.

[57] Souissi A. *Matrix Product States as Observations of Entangled Hidden Markov Models.* J Stat Phys **192**, 88 (2025).

[58] Szabłowski P. J. *On the q-Hermite polynomials and their relationship with some other families of orthogonal polynomials.* Demonstr. Math. **44** (2011), no. 4, 679–708.

[59] Touchette H. *The large deviation approach to statistical mechanics.* Phys. Rep. **478**, 1–69 (2009).

[60] Touchette H., Harris R.J. *Large deviation approach to nonequilibrium systems.* In: Lecture Notes of the Les Houches Summer School, vol. 101, Oxford University Press (2012).

[61] Van Der Hofstad R. *Random graphs and complex networks. Vol. 1.* Camb. Ser. Stat. Probab. Math., **43** Cambridge University Press, Cambridge, 2017.

[62] Viennot X. G. *Catalan tableaux and the asymmetric exclusion process.* Paper presented at the FPSAC'07 - 19th International Conference on Formal Power Series and Algebraic Combinatorics, (2007).

[63] Wood A.J., Blythe R.A., Evans M.R. *Combinatorial mappings of exclusion processes* J. Phys. A **53** (2020), no. 12, 123001.