POLITECNICO DI TORINO

Corso di Laurea
in Matematica per l'Ingegneria

Master's Thesis

Non-equilibrium unidimensional traffic model with pausing particles: a biological-motivated extension of the exclusion process

Relatore
prof. Luigi Preziosi
firma del relatore

. . . . . .

Candidato
Lorenzo Vito Dal Zovo
firma del candidato

. . . . . .

Collaboratore esterno
Luca Ciandrini
firma del collaboratore esterno

. . . . . .

Anno Accademico 2024-2025

Summary

This thesis deals with the study of nonequilibrium, one-dimensional traffic phenomena which are modelled as lattices gasses, i.e. systems of particles that inhabit a discrete, one-dimensional lattice and evolve by choosing an action out of a finite set of possible ones, according to a continuous time Poisson process. In particular we study the well-known Totally Asymmetric Simple Exclusion Process (TASEP), in which particles hop from site to site in one direction only (totally asymmetric) on the condition that the landing site is empty (simple exclusion), and a TASEP-derived model called pausing TASEP (pTASEP) in which, on top of the previous dynamics, particles can randomly enter and exit a paused state during which they are unable to advance.

The first aim of this work is to explore the feasibility of using the pTASEP as a mean to model the phenomenon of antibiotics inhibited protein translation, that is the process in which mRNA templates (the lattice) are read site by site by protein-synthetizing ribosomes (the particles), while subject to stochastic pauses due to antibiotics. Unlike for the standard TASEP, for the pTASEP there are no known exact, closed-form expressions for the physical observables of interest. Preexisting literature contains only mean field theories in very restrictive geometries, such as the closed ring (periodic boundary conditions) rather than the more useful open segment (open boundaries conditions), and no inquiry into the phenomena taking place in the biologically relevant parameter space. In this work we provide a numerical analysis, through simulations based on Gillespie's algorithm, of said region of the parameter space, proving that finite size effects are predominant in a non-negligible portion.
We extended the previous mean field theories to the open boundaries case, showing a similar accuracy compared to the periodic case in the absence of finite size effects.

Finally, we provide minimal, approximate models to predict such finite size phenomena in both the periodic and open boundaries cases, finding good agreement with numerical results and thus providing greater insight into the effects that the system size has on the emitted particle current. Both approximate models were obtained by considering the clear time scale separation between the ribosome dynamics and the antibiotics dynamics, as the latter are several orders of magnitude slower than the former. In the open case this implies that, for biologically relevant lattice sizes and system parameters, pauses are rare, in the sense that the average time spent on the lattice is much smaller than the antibiotics kinetics timescales. These considerations suggest seeing the system as an intermittent one, oscillating between a behaviour described by the standard TASEP and one dominated by the

antibiotics-induced pauses. These simplifications allowed to obtain a simple relation between the particle current emitted by the lattice and the antibiotics binding rate, which is the parameter of greatest interest as it is directly linked with antibiotics concentration. The results show a qualitative agreement across several orders of magnitude of variations of the binding parameters, which could not be obtained by the usual mean field theory. The same considerations were used for the model equipped with periodic boundary conditionss, with the difference that we had to describe the transient between the two behaviours explicitly to obtain agreements with the simulations. To do so I made use of a continuum limit approximation, in which the system is described by a quasi-linear, Burgers-like partial differential equation. The resulting model provides better agreement with numerical results than the simple mean field model, with the difference becoming more dramatic as the antibiotics dynamics become slower and slower, compared to the ribosome dynamics.

The second part of the present work presents a novel link between the algebraic properties of the matrix appearing in the master equation governing the evolution of the system (the transition rate matrix) and the closed form expressions of physical observables in the standard TASEP. We first showed how all physical observables of the system can be written as functions of the trace of powers of the transition rate matrix, making use of the Faddeev-Leverrier algorithm. We then decomposed the transition rate matrix as a sum of terms, in what is known as the quantum Hamiltonian formalism, to show that the values of said traces depend on how the terms of the quantum Hamiltonian commute with each other. This part of the work constitutes a step in the attempt to recover already known analytical results for the standard TASEP, but without making use of the ad hoc Ansatz currently needed to obtain them. The appeal of this line of research is to develop a more general solution method that could be applied to models for which the Ansatz used for the standard TASEP is not useful (such as the aforementioned pTASEP), which to this day still lack an exact treatment.

Acknowledgements

Giunto alla fine di questo percorso vorrei ringraziare le persone che fino a questo momento mi hanno sostenuto emotivamente e materialemnte.

Ringrazio mio padre e mia madre per il loro sostegno, per avermi permesso di trasferirmi in questa città e avermi lasciato i mezzi materiali per potermi concentrare sullo studio. Ringrazio i miei fratelli, Thomas e Checco, per essersi sempre fatti senitre (più di me) e per non essersi dimenticati di me durante questi anni di assenza. Ringrazio mia zia Mimi e mia nonna, che avrebbe voluto assistere a questo momento.

Ringrazio tutti i miei amici per il loro illimitato e continuo supporto, emotivo e non. Ringrazio Isa per avermi ospitato, ascoltato, sostenuto e permesso di migliorare. Fabri per avermi ospitato per mezza Europa e non aver mai mancato una chiamata in anni di amicizia a distanza. Ringrazio Leti per avermi sopportato ed essere rimasta nonostante tutto, Alessio per avermi tollerato come coinquilino per anni e sostenuto in infinite ore di aula studio. Vorrei ringraziare Paola per avermi ospitato a casa sua, per esserci stata anche quando non c'ero io, per essermi sempre venuta incontro nel momento del bisogno. Ringrazio ancora Gian per essermi stato vicino e per aver salvato tutte le bici che ho distrutto, Frank per aver rallegrato le sessioni più disperate al Comala, Enrico per avere condiviso con me tragedie e successi lungo questo percorso.

Vorrei ringraziare Letizia, con cui ho condiviso questi ultimi mesi e che mi ha permesso di passarli felicemente, nonostante le difficoltà.

Ringrazio il professor Luca Ciandrini, che mi ha dato la possibilità di lavorare a questo progetto, ma sopratutto che mi ha accolto nel suo gruppo di ricerca come un amico prima che come uno stagista. Ringrazio Marco Morandotti per aver creduto in me, nel mio progetto e per la disponibilità che ha sempre dimostrato dal primo giorno in cui l'ho conosciuto, nel primo anno di triennale. Ringrazio il professor Preziosi per la sua disponibilità la tesi e il tirocinio.

I would like to thank Johannes Keisers and Philppe Fuchs, for sharing with me this project, too many Aperós and many climbing sessions. I don't think many interns find such a lovely environment at their arrival, and you are the ones to thank for it.

Contents

List of Figures

Chapter 1

Nonequilibrium lattice models

In this chapter we will introduce the field of lattice based non-equilibrium models and showcase their usefulness both in providing predictions for physical phenomena and in being a paradigmatic example for non-equilibrium physics as whole. In the second section we will introduce the totally asymmetric simple exclusion process (i.e. the TASEP) and discuss its main features.

1.1 Background and motivation

Throughout this work we will discuss models describing out of equilibrium systems, that are systems exhibiting an irreversible exchange of some physical quantities, such as mass or heat. These systems, despite said net exchange, can still reach states in which all physical observables become time invariant. Such states are referred to as *nonequilibrium steady-states* and they represent the main object of inquiry of this thesis. When a system is in equilibrium with its environment and reaches a steady state it is well known that its micro-states follow a Boltzmann distribution, but there is no such equivalent in nonequilibrium physics and, as such, each model has to be treated individually. Compared to more complete physical theories, the field of nonequilibrium statistical physics appears as a collection of results and techniques lacking a unifying theory [18].
We will focus on a subclass of models that could be categorized as *driven, one dimensional, lattice gases*. These systems are essentially stochastic processes taking place in continuous time and on a discrete lattice composed of a finite number L of sites. Each site can exist in a finite number m of states and interact with other sites of the lattice, generally its nearest neighbors. As a consequence the configuration space χ will be finite ($|\chi| = n = L^m$) and full knowledge of the

system at a certain time t can be encoded in a vector $x(t) \in \{1, ..., m\}^L$ called the *configuration vector*, where numbers $1, .., m$ identify the state of a given site. A transition between two configurations, say i and j, is a Poisson process governed by a transition rate λ_{ij}, meaning that after an infinitesimal time period dt the system will have transitioned from j to i with probability $\lambda_{ij}dt$. If we define as $P_j(t)$ the probability of being in configuration j at time t, the above behaviour can be properly described by a master equation [21]:

$$\frac{d}{dt}P_j(t) = \sum_{\substack{i\in\chi \\ i\neq j}} \left[\lambda_{ij}P_i(t) - \lambda_{ji}P_j(t)\right]. \tag{1.1}$$

It is convenient to define a probability vector, assigning to each configuration its probability at a given time, whose entries obey Eq.(1.1). We obtain

$$\frac{d}{dt}P(t) = -HP(t), \tag{1.2}$$

where $H \in \mathbb{R}^{n,n}$ is the matrix defined as $H_{i,j} = -\lambda_{ji}$ if $i \neq j$, while the diagonal elements are chosen so that all columns sum up to zero. We will refer to the solution of the linear problem formed by Eq.(1.2) together with an initial condition $P_0 = P(0)$ as $P(t|P_0)$. We will say that the stochastic system is *ergodic* [2] if

$$\lim_{t\to+\infty} P(t|P_0) = \overline{P} \quad \forall P_0 \in [0,1]^n$$

that is if there exists a unique steady-state probability distribution and if said distribution can be reached by all possible initial conditions. All systems in the following work will be ergodic, so the question of finding the unique steady-state probability distribution is well posed. We say that a given system has been solved exactly when there are known, closed-form relationships for the physical observables of interest that depend explicitly on the system size L and the transition rates of the system and, as a consequence, it is possible to compute their infinite size limits. The existence of said limits means that, while for small sizes a system can experience finite size effects, there exists a size after which the dynamics can be considered size independent. This asymptotic study is of great interest both applicative, because often real life phenomena happen on lattices that are well beyond the scale of size-dependency, and theoretical, for the insurgence of limit phenomena such as phase transitions.

1.2 The Totally Asymmetric Exclusion Process (TASEP)

The Totally Asymmetric Simple Exclusion Process (TASEP) was proposed for the first time by MacDonald et al. [15] to describe the kinetics of bio-polymerization on

nucleic acid templates, i.e. protein synthesis. It remains the most paradigmatic model in nonequilibrium statistical physics as it is simple enough to be exactly solvable while still exhibiting non trivial phenomena exclusive of non equilibrium physics, such as boundary driven phase transitions. For this reasons the TASEP holds the same role that the Ising model has for equilibrium physics [18]. The model consists of a finite dimensional lattice of size L, where every site i ($1 \leq i \leq L$) can either be occupied ($\tau_i = 1$) or empty ($\tau_i = 0$). After each infinitesimal time step dt all particles can perform a hop to the right with probability γdt, where γ is known as the hopping rate. Thus, if at time t the system exists in the configuration $x(t) = (\tau_1, ..., \tau_n)$ we will have:

$$\tau_i(t+dt) = \begin{cases} 1 & \text{with probability } p_i = \tau_i + \gamma[\tau_{i-1}(1-\tau_i) - \tau_i(1-\tau_{i+1})]dt \\ 0 & \text{with probability } 1 - p_i \end{cases} \tag{1.3}$$

for $i = 1, ..., L-1$.
To complete the model we have to add to (1.3) other rules concerning the boundaries to describe how particles enter site $i = 1$ and exist $i = L$, known as boundary conditions. There exists two options of interest: open and periodic. The periodic system is essentially a ring and, even tough it is not a realistic model of biosynthesis, it is of theoretical interest as it is more easily treated than the open one. For the periodic boundaries we have:

$$\begin{aligned} \tau_1(t+dt) &= \begin{cases} 1 & \text{with probability } p_1 = \tau_1 + \gamma[\tau_L(1-\tau_1) - \tau_1(1-\tau_2)]dt \\ 0 & \text{with probability } 1 - p_1 \end{cases}; \\ \tau_L(t+dt) &= \begin{cases} 1 & \text{with probability } p_L = \tau_L + \gamma[\tau_{L-1}(1-\tau_L) - \tau_L(1-\tau_1)]dt \\ 0 & \text{with probability } 1 - p_L \end{cases}. \end{aligned} \tag{1.4}$$

In the open system, instead, the lattice is a segment onto which particles are injected at site $i = 1$ to then be ejected at site $i = L$. Both injection and ejection are Poisson processes, with rates α and β respectively. Thus we have the following update rules:

$$\begin{aligned} \tau_1(t+dt) &= \begin{cases} 1 & \text{with probability } p_1 = [\alpha(1-\tau_1) - \gamma\tau_1(1-\tau_2)]dt \\ 0 & \text{with probability } 1 - p_1 \end{cases}; \\ \tau_L(t+dt) &= \begin{cases} 1 & \text{with probability } p_L = [\gamma\tau_{L-1}(1-\tau_L) - \beta\tau_L]dt \\ 0 & \text{with probability } 1 - p_L \end{cases}. \end{aligned} \tag{1.5}$$

It is worth noting that in both cases there exists a complete symmetry between particles and holes. If we inverted the labels between particles and holes we would

end up again with a TASEP but with opposite hopping direction with respect to the original. Throughout this work our main observable of interest will be the steady state particle current J, that is the mean number of particles moving per time unit out of a site. In particular in the case a of periodic boundaries we are interested in current between any two sites i and $i+1$, while in the open system we are interested in the exit current at site L.

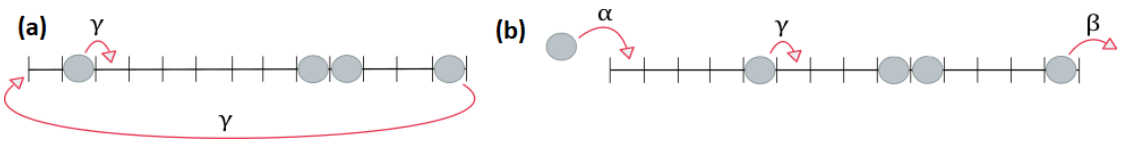

Figure 1.1: Panel (a) shows the TASEP with periodic boundary conditions, while panel (b) with open boundary conditions.

1.2.1 Periodic boundary conditions

The periodic system represents the simpler scenario as the number of particles N on the lattice remains constant and is a fixed parameter. Being that the system exhibits no bias towards any particular site, all possible lattice configurations are equally likely. This in turn allows us to compute the aforementioned particle current. The particle current between two consecutive sites, say i and $i+1$, can be seen as an effective hopping rate describing the number of successful hopping events between two sites. The steady state current J is identical for every site as a consequence of all configurations being equally likely and it will be given by the hopping rate times the probability that the event is allowed. Knowing that in total there are $\binom{L}{N}$ possible configurations we can write:

$$J = \gamma \frac{\binom{L-2}{N-1}}{\binom{L}{N}} = \gamma \frac{N}{L} \frac{L-N}{L-1}$$

if we now take the limit for $N, L \longrightarrow \infty$, $\rho = \frac{N}{L} = constant$, which we will refer to as the *thermodynamic limit*, we get

$$J = \gamma\rho(1-\rho), \tag{1.6}$$

which is the celebrated current-density relation, see Fig.(1.2). TASEP-like systems can indeed exhibit size dependent behaviour, and such finite size effects will be the key focus of the coming chapters. For the standard TASEP here described they become negligible already for moderate system sizes, making the study of thermodynamic limits very appealing.

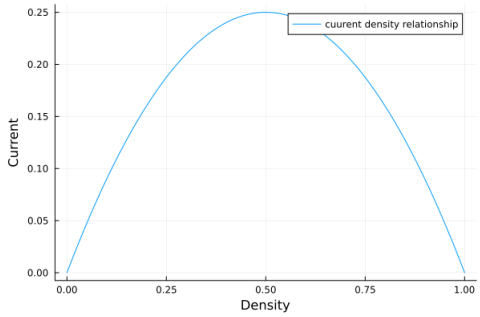

Figure 1.2: Current density relation for the periodic TASEP where $\gamma = 1$.

1.2.2 Open boundary conditions

The open system is of greater applicative interest as it describes a transport process between two reservoirs maintained at fixed densities. Said density reservoirs can be imagined as two virtual sites labelled as $i = 0$ and $i = L + 1$, where a site density has to be interpreted as the mean site occupation . From now onward when speaking of the standard open TASEP we will consider all rates to have been renomarlized over γ, so that, without loss of generality, we can write $\gamma = 1$. In this context the virtual site $i = 0$ will have fixed density $\rho_0 = min(\alpha,1)$, while virtual site $i = L + 1$ will be associated with a density $\rho_{L+1} = max(1 - \beta,0)$. It can be observed that during the system steady state there exist a portion of the lattice where the particle density is roughly constant and its determined by the choice of the ejection and injection rates. We will refer to this region as the *bulk of the system* and to the bulk density as ρ_{bulk}. The fact that the boundaries of the system determine the behaviour of its bulk is a unique feature of nonequilibrium systems as, unlike in their equilibrium counter parts, the boundaries effects are transported deep within the lattice, rather then vanishing as the system is probed deeper and deeper within the bulk. Interestingly, the bulk density experiences a first order phase transition as the parameters are varied, leading to the following phase plane:

$$\rho_{bulk} = \begin{cases} \alpha & \text{if } \alpha < \frac{1}{2},\ \beta > \alpha \\ \beta & \text{if } \beta < \frac{1}{2},\ \alpha > \beta \\ \frac{1}{2} & \text{if } \alpha, \beta > \frac{1}{2} \end{cases}.$$

This behaviour can be rationalized trough the use of a mean field approach, where one ignores site correlations and considers Eq.(1.6) as always locally valid. Interestingly this simplified approach correctly predicts the real current behaviour, as the the bulk density determines the overall current and the current density relation

still holds:

$$J = \begin{cases} \alpha(1-\alpha) & \text{if } \alpha < \frac{1}{2},\ \beta > \alpha \\ \beta(1-\beta) & \text{if } \beta < \frac{1}{2},\ \alpha > \beta \\ \frac{1}{4} & \text{if } \alpha, \beta > \frac{1}{2} \end{cases}.$$

We will refer to the region where $J = \frac{1}{4}$ as the *maximal current phase*, to the region where the current is constrained by the injection rate as the *low density phase* and to the region where the ejection rate is limiting as the *high density phase*. The complete phase diagram can be seen in Fig.(1.3). We will not provide

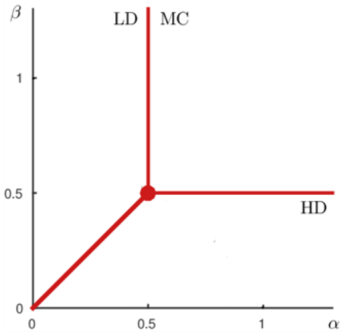

Figure 1.3: Boundaries driven phase transitions in the TASEP

the details of the mean-field theory of the TASEP, which can be found, for example, in [18]. Although it succeeds in predicting the current phase plane, we will move directly to the exact theory developed by Derrida et al. [5]. This theory not only offers a mathematically rigorous explanation for the phase diagram but is also able to deliver results beyond the reach of mean-field theories, such as exact, size-dependent expressions for the physical observables of interest, i.e. particle current and density.

Chapter 2

Mathematical context and exact results

In the first two sections of this chapter we will present the solution method known as the *matrix factorization method*, first utilized by Derrida et al. in 1992 [5] for solving the open TASEP. We will then review their derivation of the main features of the open TASEP, with a special focus on phase transitions. In the last section we will instead present a result, due to Krebs et al. [13], providing a greater understanding of the applicability and limits of the matrix factorization method. The main takeaway of this part is that, even when the method is applicable, it has no guarantee of leading to closed form expressions for the physical observables of the system, thus leaving mean field and approximate theories as the only tools for more general models.

2.1 Mathematical formalization of the TASEP

Throughout the coming two sections we will assign to each site of the system a boolean variable τ_i, $1 \leq i \leq L$, where $\tau_i = 0$ will mean that site i is empty and vice versa. We will refer to the probability of finding a system of size L in configuration $\tau_1, ..., \tau_L$ at time t as $P_L(\tau_1, ...\tau_L; t)$, while the absence of the variable t will mean that we are referring to the unique steady state probability of the configuration. Our goal is to write Eq.(1.1) explicitly for the open TASEP update rules (1.3),(1.5). To do so we will introduce matrices representing transition out of (diagonal elements) and into (off diagonal elements) a given configuration. We will call h_l and h_r the matrices describing, respectively, configuration changes due

to a particle entering and exiting the lattice. By choosing as base $\{0;1\}$ we have

$$h_l = \begin{pmatrix} -\alpha & 0 \\ \alpha & 0 \end{pmatrix}, \; h_r = \begin{pmatrix} 0 & -\beta \\ 0 & \beta \end{pmatrix}.$$

Transition between configurations due to a particle hopping forward require a 4 by 4 matrix, as they describe a two site interaction. In particular rules (1.5) describe the following transition:

$$1\ 0 \xrightarrow{\gamma} 0\ 1,$$

by choosing as a base $B = \{(0,0);(0,1);(1,0);(1,1)\}$ we can write:

$$h = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where again we assume, without loss of generality, that the transition rates have been normalized as to have $\gamma = 1$. We can now write Eq.(1.1) for the open TASEP:

$$\begin{aligned} \frac{\partial}{\partial t} P_L(\tau_1, ..., \tau_L; t) = & \sum_{\sigma_1 \in \{0,1\}} (h_l)_{\tau_1,\sigma_1} P_L(\sigma_1, \tau_2, ..., \tau_L; t) + \\ & + \sum_{i=1}^{L-1} \sum_{(\sigma_i,\sigma_{i+1}) \in B} (h)_{(\tau_i,\tau_{i+1})),(\sigma_i,\sigma_{i+1})} P_L(\tau_1, ..., \sigma_i, \sigma_{i+1}, ..., \tau_L; t) + \\ & + \sum_{\sigma_L \in \{0,1\}} (h_r)_{\tau_L,\sigma_L} P_L(\tau_1, \tau_2, ..., \sigma_L; t)\ . \end{aligned} \tag{2.1}$$

Let $i \in I$ be an indexing of the set of all possible configurations, so that $\forall(\tau_1, ..., \tau_L)$ it's possible to univocally write $P_L(\tau_1, ..., \tau_L; t) = (P_L(t))_i$, where $P_L(t) \in [0,1]^L$ is the vector having for the i-th entry the probability of finding the system at time t in the configuration indexed by $i \in I$. We are interested in computing the steady state probability vector $\overline{P_L}$, but it will be more convenient to first compute a vector parallel to it whose entries do not in general sum up to one. We will refer to said vectors $f_L \in \mathbb{R}^L$ for which $\exists k \in \mathbb{R}$ s.t. $f_L = k\overline{P_L}$ as vectors of *statistical weights*.

2.2 Exact results for the TASEP

2.2.1 The matrix factorization method

Our goal is to find the probability vector $\overline{P}_L$, or a statistical weights vector f_L, for which Eq.(2.1) is equal to 0. Let us assume that there exists two scalar quantities

$x_0,\ x_1 \in \mathbb{R}$ such that the following expressions are true:

$$\sum_{\sigma_1 \in \{0,1\}} (h_l)_{\tau_1,\sigma_1} P_L(\sigma_1, \tau_2, ..., \tau_L; t) = x_{\tau_1} P_{L-1}(\tau_2, ..., \tau_L; t) \tag{2.2}$$

$$\sum_{i=1}^{L-1} \sum_{(\sigma_i,\sigma_{i+1}) \in B} (h)_{(\tau_i,\tau_{i+1})),(\sigma_i,\sigma_{i+1})} P_L(\tau_1, ..., \sigma_i, \sigma_{i+1}, ..., \tau_L; t) = \tag{2.3}$$
$$= -x_{\tau_i} P_{L-1}(\tau_1, ..., \tau_{i-1}, \tau_{i+1}, ..., \tau_L; t) + x_{\tau_{i+1}} P_{L-1}(\tau_1, ..., \tau_i, \tau i + 2, ..., \tau_L; t)$$

$$\sum_{\sigma_L \in \{0,1\}} (h_r)_{\tau_L,\sigma_L} P_L(\tau_1, \tau_2, ..., \sigma_L; t) = -x_{\tau_L} P_{L-1}(\tau_1, \tau_2, ..., \tau_{L-1}; t). \tag{2.4}$$

If conditions (2.2) to (2.4) are verified it follows immediately that the master equation Eq.(2.1) is identically zero. If we now find an expression for the steady state probability of a given configuration satisfying the aforementioned conditions then the problem is solved. To achieve this we introduce the *matrix product Ansatz*, which states that for every configuration $\tau_1, ..., \tau_L$ one can have a statistical weight of the form:

$$f_L(\tau_1, ..., \tau_L) = w^\intercal \left[\prod_{i=1}^{L} (\tau_i D + (1 - \tau_i) E) \right] v \tag{2.5}$$

where D, E are matrices associated respectively with occupied and empty sites, while w, v are appropriately sized vectors so that the product gives a scalar. If we now substitute Eq.(2.5) into conditions (2.2)-(2.4) we get:

$$\begin{aligned} &\alpha w^\intercal E = x_1 w^\intercal = -x_0 w^\intercal, \\ &DE = -x_0 D + x_1 E, \\ &\beta D v = x_1 v = -x_0 v. \end{aligned} \tag{2.6}$$

Finally, choosing $x_0 = -x_1 = 1$ and substituting in conditions (2.6) we obtain

$$\begin{aligned} &w^\intercal E = \frac{1}{\alpha} w^\intercal, \\ &DE = D + E, \\ &Dv = \frac{1}{\beta} v. \end{aligned} \tag{2.7}$$

It is possible to explicitly write a choice of $v,\ w,\ E,\ D$ fulfilling the conditions (2.7) and actually there are infinite many of them, but it will be shown in the following that the algebraic conditions themselves are sufficient to compute all observables of interest. As such we will not provide examples, but some can be found in the original work by Derrida et al. [5].

2.2.2 Physical observables as a function of the factorization matrices

To compute any physical observable we first need to calculate the *normalization function* Z_L, that is the sum of all statistical weights:

$$Z_L = \sum_{(\tau_1,...,\tau_L)} f_L(\tau_1, ..., \tau_L),$$

where the sum has to be interpreted as over all possible configurations. Substituting Eq.(2.5) and using conditions (2.7) we get

$$Z_L = w^\intercal (D + E)^L v := w^\intercal C^L v, \tag{2.8}$$

where we have defined the matrix $C = D + E$. Let $\rho_{i,L}$ be the average occupation of a site i, which can be seen as a particle density. We have:

$$\rho_{i,L} = \frac{\sum_{(\tau_1,....,\tau_L)} \tau_i f_L(\tau_1, ..., \tau_L)}{Z_L} = \frac{w^\intercal C^{i-1} D C^{L-i} v}{w^\intercal C^L v}, \tag{2.9}$$

where one can compute the second equality by using the matrix product form of f_L and the algebraic rules just as before. The current emitted by the system is given by:

$$J_L = \beta \rho_{L,L},$$

and by using Eq.(2.9) one obtains:

$$J_L = \beta \frac{w^\intercal C^{L-1} D v}{Z_L} = \beta \frac{1}{\beta} \frac{Z_{L-1}}{Z_L} = \frac{Z_{L-1}}{Z_L}. \tag{2.10}$$

Thus, we only need to find an explicit expression for Z_L to compute the particle current.

2.2.3 Normalization generating function

There are several ways to obtain exact combinatorial expressions for the normalization function Z_L and the particle current J_L from the above relations. Here we will present a *generating function approach* as it is very elegant and arguably the fastest. The derivation presented from now onwards is found in [18]. By definition the generating function of the quantity Z_L is

$$\mathcal{Z}(z) = \sum_{L=0}^{+\infty} z^L Z_L, \tag{2.11}$$

so that $\forall L > 0$ we have:

$$Z_L = \frac{1}{L!}\left(\frac{d^L}{dz^L}\mathcal{Z}(z)\right)\Big|_{z=0}.$$

Consider the formal matrix series:

$$\frac{1}{1-zC} = \sum_{i=0}^{\infty} z^i C^i,$$

so that Eq.(2.11) can be written as:

$$\mathcal{Z}(z) = w^\intercal \frac{1}{1-zC} v. \tag{2.12}$$

By introducing the auxiliary variable $\eta(z)$ defined as $z = \eta(1-\eta)$ and using the identities (2.7) is possible to prove that:

$$\frac{1}{1-zC} = \frac{1}{1-\eta E}\frac{1}{1-\eta D},$$

which substituted in Eq.(2.12) along with conditions (2.7) gives us:

$$\mathcal{Z}(z) = \left(1 - \frac{\eta(z)}{\alpha}\right)^{-1}\left(1 - \frac{\eta(z)}{\beta}\right)^{-1}. \tag{2.13}$$

The conventional condition $\mathcal{Z}(0) = 1$ implies that we have to choose the root $\eta(z) = \frac{1}{2}(1-\sqrt{1-4z})$, completing Eq.(2.13). We are interested now in providing an explicit expression for the quantities Z_L, which are the coefficients of the terms in the power series expansion of the generating function. Given a formal series $\mathcal{G}(x)$ we will refer to the coefficient of x^n as $\{x^n\}\mathcal{G}(x)$. It's possible to recover the aforementioned coefficient trough an application of the Lagrange inversion formula for generating functions ([18], [27]), obtaining:

$$Z_L = \{z^L\}\mathcal{Z}(z) = \sum_{p=0}^{L} \frac{p(2L-p-1)!}{L!(L-p)!} \sum_{q=0}^{p} \left(\frac{1}{\alpha}\right)^p \left(\frac{1}{\alpha}\right)^{p-q}. \tag{2.14}$$

Eq.(2.14) is already sufficient to derive a phase diagram by studying its asymptotic form, but we will first focus on obtaining similar exact combinatorial identities for the particle densities $\rho_{i,L}$ at a given site site i.

2.2.4 Exact density profiles

Lets recall Eq.(2.9):

$$\rho_{i,L} = \frac{w^\intercal C^{i-1} D C^{L-i} v}{Z_L},$$

which we can rewrite by investigating the matrix product DC^j, for a generic j. By computing manually the products for a few manageable values of j one can deduce (and subsequently prove by induction) the following relation:

$$DC^j = \sum_{k=0}^{j-1} B_{k+1,1} C^{j-k} + \sum_{k=2}^{j-1} B_{j,k-1} D^k, \tag{2.15}$$

where we have introduced the combinatorial quantity $B_{N,p}$, defined as:

$$B_{N,p} = \begin{cases} \frac{p(2N-p-1)!}{N!(N-p)!} & \text{for } 0 < p \leq N \\ 0 & \text{otherwise} \end{cases}.$$

By substituting Eq.(2.15) into Eq.(2.9) one gets:

$$\rho_{i,L} = \begin{cases} \sum_{n=1}^{L-i} B_{n,i} \frac{Z_{L-n}}{Z_L} + \frac{Z_{i-1}}{Z_L} \sum_{p=1}^{L-i} B_{L-i,p} \frac{1}{\beta^{p+1}} & \text{for } 0 < i < L \\ \frac{Z_{L-1}}{\beta Z_L} & \text{for } i = L \end{cases}. \tag{2.16}$$

Now we can write both the particle current and the density profile for finite sizes as functions of the normalization, which is itself known for all systems sizes L. Our next goal is to describe phase transitions by looking of the asymptotics of the system, which we can do reasonably fast thanks to knowing the generating function of the sequence of normalization constants $\{Z_L\}_{L=1}^{\infty}$, given by Eq.(2.13).

2.2.5 Asymptotics and the exact phase plane

The asymptotic behaviour of an infinite sequence, such as $\{Z_L\}_{L=1}^{\infty}$, is determined by the singularity of least module of the sequence's generating function [27]. The generating function Eq.(2.13) always exhibits, no matter the values of of α and β, a square root singularity in $z = \frac{1}{4}$. Function $\mathcal{Z}(z)$ can also exhibit poles depending on the values of the parameters α and β: the first factor contributes a pole whenever $a = \eta(z)$, which is possible only for $\alpha < \frac{1}{2}$, same goes for the second factor and β. The two nodes, when they exist, have always smaller module then the square root singularity, while the singularity of least module between the two nodes is determined by which parameter is smaller. Regions of the parameter space in which different singularities dominate correspond to regions in which the sequence $\{Z_L\}_{L=1}^{\infty}$ has different asymptotic behaviour, thus providing the phase plane. From the analysis above follows that there exist three such regions:

- region $\{\alpha, \beta \in \mathbb{R}^+ |\ \alpha, \beta > \frac{1}{2}\}$, where the square root singularity dominates;
- region $\{\alpha, \beta \in \mathbb{R}^+ |\ \alpha < \frac{1}{2},\ \beta > \alpha\}$, where the pole due to the first factor of Eq.(2.13) dominates;

- region $\{\alpha, \beta \in \mathbb{R}^+ |\ \beta < \frac{1}{2},\ \alpha > \beta\}$, where the pole due to the second factor of Eq.(2.13) dominates.

The above phase diagram corresponds exactly to the well known one shown in Fig.(1.3). To compute the infinite size limit of the current we first need to remind some properties of generating functions [27]. Consider a generating function $f(x)$ and its associated series $[x^n]f(x)$. Let x_0 be the singularity of least module of $f(x)$ so that one can write:

$$f(x) = f_r(x)\left(1 - \frac{x}{x_0}\right)^{-\eta}.$$

Then, for integer η one has:

$$[x^n]f(x) \sim \binom{\eta + n - 1}{n} f_r(x_0) x_0^{-n}, \tag{2.17}$$

while for non-integer η:

$$[x^n]f(x) \sim \frac{f_r(x_0)}{\Gamma(\eta) n^{\eta-1}} f_r(x_0) x_0^{-n}, \tag{2.18}$$

where the factor f_r is given by:

$$f_r(x_0) = \lim_{x \to 0} \left(1 - \frac{x}{x_0}\right)^{\eta} f(x). \tag{2.19}$$

Equations (2.17),(2.18) and (2.19) allow to compute the asymptotic form of $\{Z_L\}_{L=1}^{\infty}$, which in turn, together with the current equation (2.10), allows to compute the infinite size limit of the particle current $J(\alpha, \beta)$, trough straightforward, albeit lengthy, computations. One eventually obtains:

$$Z_L \sim \begin{cases} \frac{\alpha\beta(1-2\alpha)}{(\beta-\alpha)}[\alpha(1-\alpha)]^{-L-1} & \text{if } \alpha < \frac{1}{2},\ \beta > \alpha \\ \frac{\alpha\beta(1-2\beta)}{(\alpha-\beta)}[\beta(1-\beta)]^{-L-1} & \text{if } \beta < \frac{1}{2},\ \alpha > \beta \\ \frac{4\alpha\beta(\alpha+\beta-1)}{\sqrt{\pi}(2\alpha-1)^2(2\beta-1)^2}\frac{4^L}{L^{3/2}} & \text{if } \alpha, \beta > \frac{1}{2} \end{cases},$$

to then obtain trough Eq.(2.10):

$$J = \begin{cases} \alpha(1-\alpha) & \text{if } \alpha < \frac{1}{2},\ \beta > \alpha \\ \beta(1-\beta) & \text{if } \beta < \frac{1}{2}\ \alpha > \beta \\ \frac{1}{4} & \text{if } \alpha, \beta > \frac{1}{2} \end{cases},$$

which corresponds exactly to the mean field prediction.

2.3 Closing remarks on the matrix fact factorization method

Krebs et al. demonstrated [13] that the unique stationary solution $\overline{P}$ of any master equation of the form (1.2) describing a one-dimensional system with nearest-neighbor interactions in the bulk and one-site interactions at the boundaries can be expressed as a matrix product. That means that it is always possible to choose an auxiliary vectors space V, two vectors $v,\ w \in V$ and $2m$ operators $A_1, \ldots, A_m, B_1, \ldots, B_m$ such that:

$$h\left[\begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix} \otimes \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix}\right] = \begin{pmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_m \end{pmatrix} \otimes \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix} - \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix} \otimes \begin{pmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_m \end{pmatrix},$$

$$w^\intercal h^{(l)} \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix} = -w^\intercal \begin{pmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_m \end{pmatrix} \quad \text{and} \quad h^{(r)} \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix} v = \begin{pmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_m \end{pmatrix} v,$$

and

$$\overline{P}(\tau_1, \ldots, \tau_n) = w^\intercal A_{\tau_1} A_{\tau_2} \ldots A_{\tau_n} v,$$

where m is the number of sates a single site can be in, n is the numbers of configurations of the whole lattice, h, $h^{(r)}$, $h^{(l)}$ are respectively the matrices describing the bulk and boundaries interactions and $\tau_i \in \{1, \ldots, m\}$ is the state of site i. In the case discussed earlier the standard TASEP one has $m = 2$, $(A_1, A_2) = (E, D)$, $B_1 = -B_2 = 1$ and the matrices h, $h^{(r)}$, $h^{(l)}$ are defined at the beginning of Section 2.1. The reason closed-form relations like Eq.(2.8) are possible in the standard TASEP is that the operators B_i happen to actually be numbers, which is not generally the case. The matrix factorization method is a reformulation of the standard eigenvector problem associated with the transition rate matrix H and does not provide additional insight beyond that. Therefore, it is not strictly an Ansatz, although it is often referred to as such, but rather a resolution method that may or may not simplify the original eigenvalue problem.

Chapter 3

Biopolymerization processes as lattice gases

Protein synthesis is a fundamental biological process involving two distinct biopolymerization events: transcription and translation. During transcription, DNA is transcribed into messenger RNA (mRNA) by the enzyme RNA polymerase (RNAP). This mRNA then serves as a template for translation, during which ribosomes assemble proteins based on the mRNA sequence. In both stages, molecular machines—RNAP in transcription and ribosomes in translation—move unidirectionally along their respective templates, which are linear sequences of nucleotides (DNA or RNA).

The transcription process begins when RNAP binds to a specific region of the DNA, initiating the production of an mRNA strand complementary to the DNA template. This mRNA carries the genetic information from the nucleus to the cytoplasm, where translation occurs. During translation, ribosomes decode the mRNA sequence into a chain of amino acids, producing functional proteins. The information encoded in these templates is represented by a finite sequence of nucleotides—four bases (A, U, G, C) in the case of mRNA—organized in codons, each specifying a particular amino acid. This finite sequence behaves like a discrete, one-dimensional lattice along which the molecular machinery moves step by step, synthesizing proteins according to the encoded instructions.

A key feature of this process, confirmed by experimental evidence [16], is that multiple instances of a given protein can be synthesized simultaneously on the same mRNA template. This results in the formation of polyribosomes, or polysomes, where several ribosomes are translating a single mRNA at once. Similarly, in transcription, multiple RNA polymerases can transcribe the same gene simultaneously. However, due to the linear nature of the template and the need to preserve its integrity, meaning no section of the template can be skipped, there is only one active

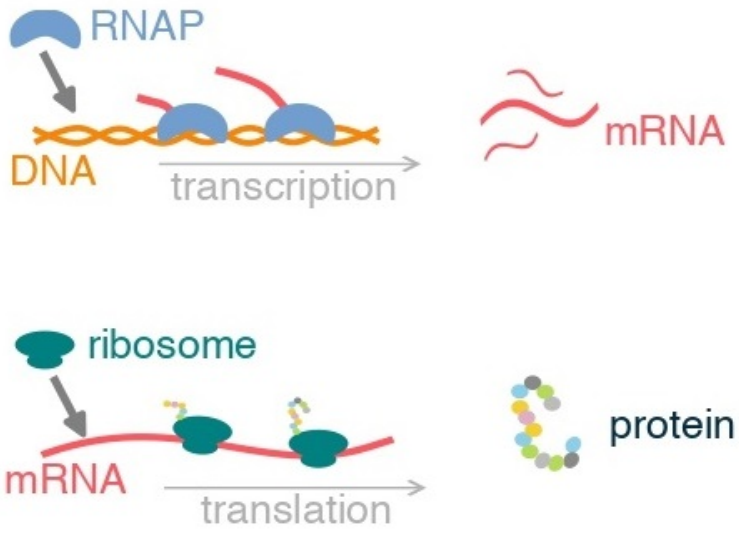

Figure 3.1: Schematic representation of the processes of transcription and translation.

site at each point along the sequence, and molecular machines cannot overtake one another. This restriction leads to traffic-like behavior, where queuing phenomena naturally arise.

These observations suggest that queuing effects must be accounted for when modeling the dynamics of transcription and translation. Recognizing this, MacDonald et al. introduced the TASEP in 1968 [15] as a mathematical model to describe such traffic behavior. While the standard TASEP simplifies some aspects of polymerization dynamics, such as neglecting the fact that particles (like ribosomes or RNAPs) occupy multiple sites on the template due to their physical size, it remains the only exactly solvable model of driven lattice gases. As such, TASEP provides an invaluable tool for quantitatively understanding the effects of traffic during transcription and translation.

Even though more complex models do not have an exact analytical solution, they can still be effectively simulated using methods such as Gillespie's algorithm. Numerical analysis can provide biologically relevant insights. In the following, we will introduce one of the many possible extensions of the TASEP: a model where particles can stochastically enter, and subsequently exit, a paused state, known as the pTASEP.

This pause mechanism was originally introduced in [12] to model the stalling of RNAPs during transcription, however it will be shown to have much broader

applicability in various polymerization phenomena.

3.1 The pausing TASEP (pTASEP)

The pTASEP introduces a new state configuration, which is the paused state. In the same vein as Section 2, one can equip every site i of a lattice of length L with a function $\tau_i(t)$, which takes values in $\{0,1,2\}$ depending on whether site i is empty, hosts a "free" particle or a paused one, respectively.

This allows us to write the update rules for the bulk of the system ($1 < i < L$):

$$\tau_i(t+dt) = \begin{cases} 1 & \text{with probability } p_1 = \delta_{\tau_i,1} + \gamma\left(\delta_{\tau_{i-1},1}\delta_{\tau_i,0} - \delta_{\tau_{i+1},0}\delta_{\tau_i,1}\right) dt + \\ & \qquad + \left(k_-\delta_{\tau_i,2} - k_+\delta_{\tau_i,1}\right) dt \\ 2 & \text{with probability } p_2 = \left(k_+\delta_{\tau_i,1} - k_-\delta_{\tau_i,2}\right) dt \\ 0 & \text{with probability } 1 - p_1 - p_2 \end{cases},$$

while the boundaries update rules, either periodic or open, can be written analogously to Eq.(1.4).

As stated earlier, the subsequent expansions discussed in this thesis are motivated by the existence of a class of antibiotics that disrupts protein synthesis by binding temporarily to ribosomes, blocking them for a stochastically determined period of time [23, 28, 20]. The TASEP with dynamical defects (ddTASEP) represents another, seemingly distinct, extension of the TASEP, originally introduced in [24] for a single defect and later extended in [19, 25]. In the ddTASEP, defects can bind and unbind to empty sites, obstructing particle movement. These defects can, for example, model the effects of regulatory proteins in transcription, or secondary structures in translation.

Although the dynamics of pTASEP and ddTASEP differ, the two models are equivalent through a mapping between particles (occupied sites) in one model and holes (unoccupied sites) in the other. In fact, just as the pTASEP features advancing particles that pause/unpause, the ddTASEP can be interpreted as receding holes that become blocked/unblocked. Thus, the two models are equivalent upon exchanging ρ with $1-\rho$ ([11])[1].

A mean-field theory, ignoring correlation between particles, was presented in [26] for the pTASEP. The following subsections will present said theory and a

[1] Whenever the paper [11] is cited in this chapter, it refers to results obtained by PhD student Johannes Keisers, under the supervision of Luca Ciandrini and Norbert Kern. In contrast, the results within the same paper cited in Chapter 5 were developed by myself, with the same supervision. The numerical data required for the figures were produced by Johannes Keisers, while all analytical results used in the images were derived by myself.

recent expansion to open boundary conditions [11], along with a discussion about its performance in predicting actual numerical results.

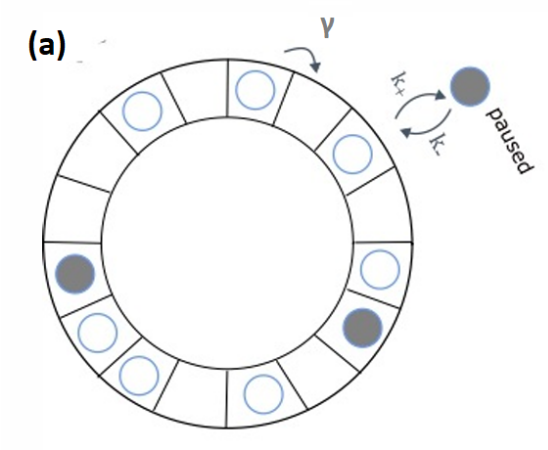

Figure 3.2: The pTASEP with open and periodic boundary conditions

3.1.1 Mean-field treatment of the pTASEP

The pTASEP involves two interconnected dynamical processes. The first is the standard exclusion process, discussed in the previous section, which operates far from equilibrium. The second process, however, involves particles transitioning between paused and unpaused states, and this occurs at equilibrium. As a result, in the steady state, the proportion of time a particle remains in either the paused or unpaused state is determined directly by equilibrium conditions:

$$f_p = \frac{k_+}{k_+ + k_-}\,, \qquad f_a = \frac{k_-}{k_+ + k_-}\,.$$

Thus, f_p denotes the fraction of paused particles (i.e. paricles that cannot move due to their internal state), and f_a the fraction of active (non-paused) particles, $f_p + f_a = 1$. Wang *et. al* [26] developed a mean-field approximation for the pTASEP by assuming that only particles that are both active and not blocked by a paused particle ahead, contribute to the current J. The mean-field approximation of the current obtained in [26] reads:

$$J_{\mathrm{MF}} = \gamma\rho f_{\mathrm{J}}(1-\rho), \tag{3.1}$$

where

$$f_{\mathrm{J}} := \frac{k_-}{k_- + \hat{k}_+}\,, \qquad \hat{k}_+ := k_+ + \gamma\rho f_p\,.$$

This reflects the fact that a particle contributing to the current will now longer do so after becoming paused (rate k_+) or after moving (rate γ) to encounter a paused particle (probability ρf_p). This should take into account both the internal pausing dynamics and the new queuing phenomena

These considerations constitutes the mean-field theory introduced in Wang *et. al* [26], resulting in Eq. (3.1).

In this context, f_J can be understood as the effective probability that a particle contributes to the current. This is derived from the typical time a particle remains in the paused state ($1/k_-$) and the corresponding effective time ($1/\hat{k}_+$) spent in the state where it contributes to the current.

Equation (3.1) effectively captures the general behavior observed in stochastic simulations. However, while the analytical solution aligns with numerical results in many cases, it becomes less precise when k_- and k_+ are reduced. In the next chapter, we will analyze the accuracy of the mean-field predictions in the parameter space of our interest.

Given that the fundamental relation $J(\rho)$ for the pTASEP is known, as shown in Eq. (3.1), it is possible to extend the mean-field solution to systems with open boundary conditions, which allows us to explore how the rates α and β control the different phases.

To achieve this, we make use of the maximum current principle, developed by Krug [14], to describe boundary-induced phase transitions. This principle asserts that

$$J = \max_{\rho \in [min(\rho_{L+1},\rho_0),max(\rho_{L+1},\rho_0)]} J(\rho) \ , \tag{3.2}$$

where 0 and $L+1$ refer to two fictitious reservoir sites located at the ends of the lattice, extending the system. These reservoirs have densities selected to simulate the desired entry and exit rates, under the assumption that the dynamics between the reservoirs and the lattice follow the same microscopic rules as in the bulk.

The current from the reservoir with density ρ_0 to the first site is given by $\gamma\,\rho_0\,f_\mathrm{J}(1-\rho_1)$, which corresponds to the current dictated by the entry rate, $\alpha(1-\rho_1)$. Therefore, we set $\alpha = \gamma\,\rho_0\,f_\mathrm{J}$. Similarly, the exit current is determined by the fraction of active particles leaving the last site at rate β, i.e., $\beta\rho_L f_a$. Importantly, all active particles on the last site contribute to the current since they cannot be blocked by traffic or paused particles downstream. The exit current is then replicated by the density ρ_{L+1} in the right reservoir. Since the current in this case is $\gamma\,\rho_L\,f_\mathrm{J}\,(1-\rho_{L+1})$, we require $\beta = \gamma(1-\rho_{L+1})f_\mathrm{J}/f_a$.

Thus, the reservoir densities ρ_0 and ρ_{L+1} are directly determined by the entry and exit rates, respectively.

The phase diagram is derived from the maximum current principle, which is

applied by setting the derivative of Eq. (3.1) to zero:

$$\begin{aligned} J &= J_{LD} \quad \text{when} \quad \rho_0 < 1 - \rho_{L+1},\ \rho_0 < \rho_{max} \\ J &= J_{HD} \quad \text{when} \quad \rho_0 > 1 - \rho_{L+1},\ \rho_{L+1} > \rho_{max} \\ J &= J_{MC} \quad \text{when} \quad \rho_0 \geq \rho_{max},\ \rho_{L+1} \leq \rho_{max} \end{aligned}$$

with

$$\rho_{max} := -\chi + \sqrt{\chi^2 + \chi} \quad \text{and} \quad \chi := \frac{k_+ + k_-}{\gamma f_p} \ .$$

We are thus able to recover a phase diagram containing the same three phases characterising the standard TASEP, even though the boundaries between them are shifted.

3.2 Comparison with numerical results

The aim of this section is to explore the performance of the mean -field theory presented above, comparing its predictions to numerical simulations, with a focus in the biologically-relevant section of the parameter space.

Regarding transcription, single-molecule studies of in vitro transcription have shown that RNA polymerases (RNAPs) transition into and out of a paused state at rates between 0.07 - 0.15 s^{-1} and roughly 1 s^{-1}, respectively [12]. The elongation speed of RNAPs varies between 20 and 80 nucleotides per second, leading to a timescale difference between the pausing and resuming processes of one to three orders of magnitude. Additionally, the ddTASEP model has been employed to describe how DNA blockages influence RNAP traffic. These blockages can be caused by regulatory proteins that bind and unbind from the DNA template, such as histones or methylated-DNA binding proteins. For example, the MeCP2 binding protein has a rate of $k_\pm$ $0.04/s$, again showing a similar time-scale separation.

In the case of antibiotics-inhibited translation, the parameters depend on the choiche of antibiotic. For example chloramphenicol, erythromycin and tetracycline all disrupt the process of elongation during translation, which is what we are modelling. The binding and unbinding rates of these translation inhibiting antibiotics are again several orders of magnitude lower than the elongation rate of the ribosome [7, 4, 23, 28].

For instance, chloramphenicol has a binding rate constant of ($k_{on} = 5.6 \times 10^{-4} \mu M^{-1} s^{-1}$) and an unbinding rate of ($k_- = 1.23 \times 10^{-3} s^{-1}$) [8]. As typical sublethal doses of antibiotics range from 2 μM to 12 μM [22], the timescales are up to five orders of magnitude apart.

Typical lattice lengths ranges from the low hundreds to the few thousands of sites, with the vast majority of known proteins being less than a thousand sites long.

With these number in mind we investigate pausing parameters up to six order of magnitude smaller then the elongation parameter. In the following all parameters will be adimensionalised taking the elongation parameter as a reference, so that $\gamma = 1$.

3.2.1 Periodic boundary conditions

The parameter space was explored by choosing three lattice lengths (200, 500 and 750 sites) and for each constructing a current-density plot for around 50 couples of the binding and binding parameters. To achieve a current-density plot the TASEP dynamics was simulated for 50 different, equally spaced density values spanning from $\rho \approx 0.01$ to $\rho \approx 0.99$, measuring each time the steady-state current. The error was maesured by taking the relative error between the mean-field prediction for the current maximum and the simulated current at the corresponding density value. This was done as often the main objective behind traffic modelling is to predict the optimal density choice to maximise the current. The results show a good agreement, at least qualitatively, the closer the binding and unbinding rates are to the elongation rate. Moreover, the deterioration of the reliability of the prediction happens sooner the shorter the lattice length, as shown in Fig(3.3). We conclude that a major part of the error is due to the emergence of finite size phenomena, which will be the subject of the next chapter.

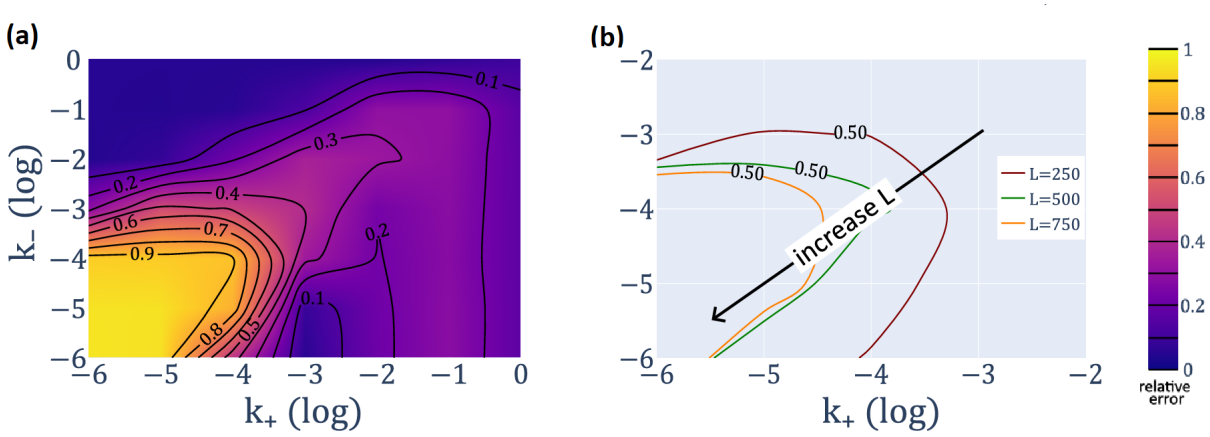

Figure 3.3: Panel (a) shows the behaviour of the mean-field theory's error in the parameter space for lattice size 750, while panel (b) shows the regions where the error rises above 50% as the lattice size is varied

When the binding and unbinding parameters are close to the elongation one it can be seen that, although the current is systematically over-estimated, the point of maximum current is predicted with good precision, as seen in Fig(3.4).

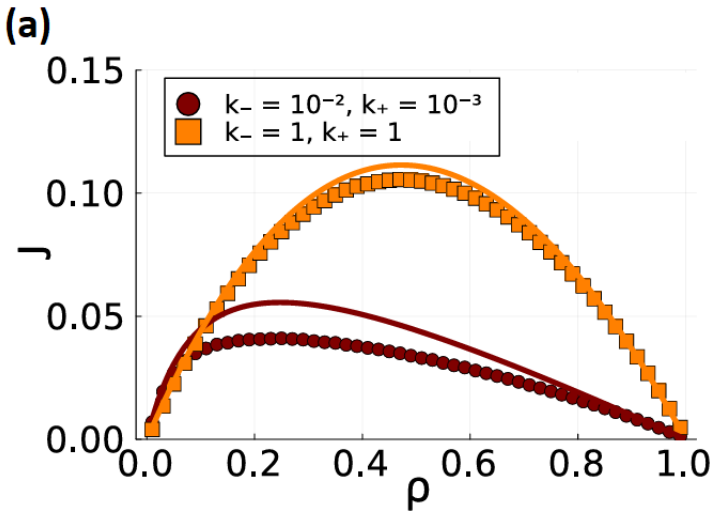

Figure 3.4: Current-density plot for two choices of binding and unbind rates, for a lattices of 1000 sites.

3.2.2 Open boundary conditions

Again for moderate time scale separation there exists good agreement between the mean-field prediction and the simulations results. In Fig(3.5) a current-density plot is reobtained by varying the parameters α and β. The low density branch, where the initiation rate α is the limiting factor for the current, is obtained by varying α at a fixed β. Both current and density are measured empirically and plotted against the respective prediction. Similarly the high density branch is computed by varying β at a fixed α. The two branches meet at the point of maximum current, where the limiting factor for the current is again the elongation rate. Fig(3.5) thus shows both good agreement in the current prediction and that the overall behaviour follows the predicted phase diagram. It's worth noting how this case too shows a decay in the accuracy of the prediction as the time separations increases, leading to similar conclusions as before.

Said deterioration in the quality of the approximation can be seen in Fig(3.6), showing the behaviour of the current as the initiation rate (panel a) and the termination rate (panel b) are varied, all other parameters being equal. All plots reach a plateau as when the parameter of interest is increased to the point of not being limiting anymore, which corresponds to the value at which the system experiences a phase transition.

The proposed phase diagram is explored in the case of $k_+ = k_- = \gamma$, as to exclude the aforementioned finite size effects. Fig(3.7) shows that the mean-field prediction is indeed descriptive of the phenomenology of the system, showing its

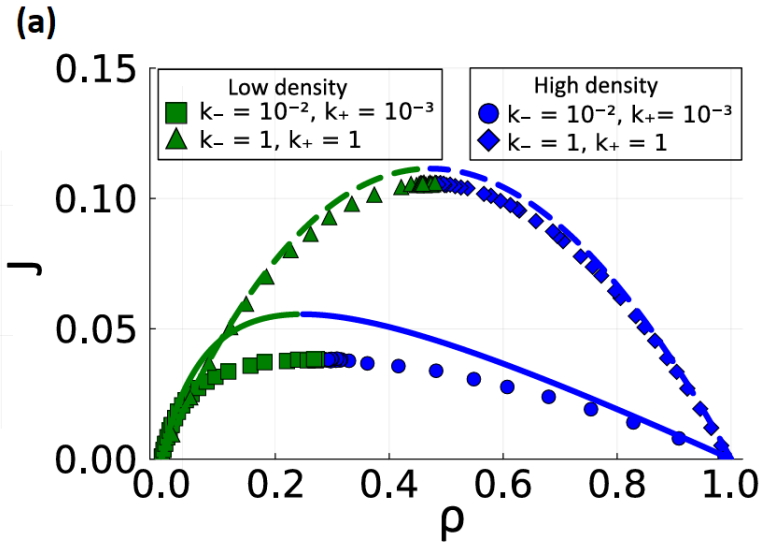

Figure 3.5: Current-density relation in the open boundaries case. The lattice is 1000 sites long.

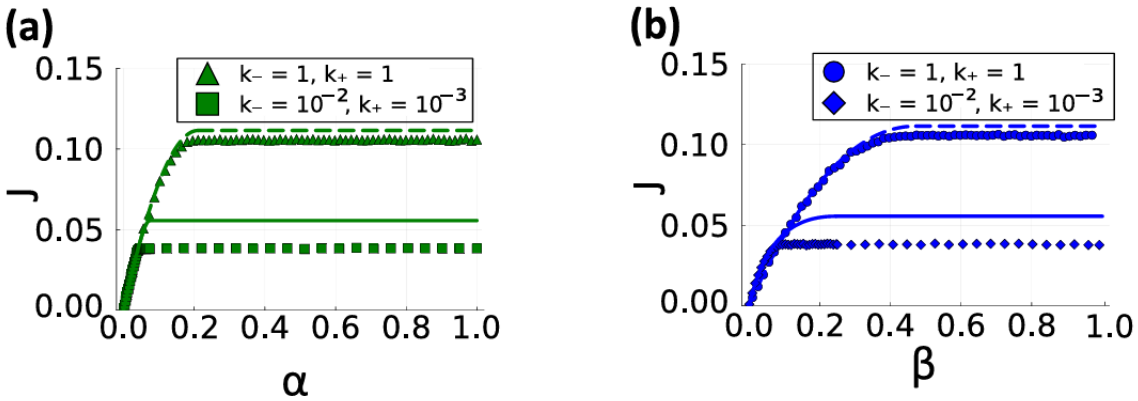

Figure 3.6: Plot of the current against (a) the initiation rate and (b) the termination rate for a lattice 1000 sites long.

usefulness in the absence of finite size effects, that is, unfortunately, in a very limited portion of the biologically relevant portion of the parameter space.

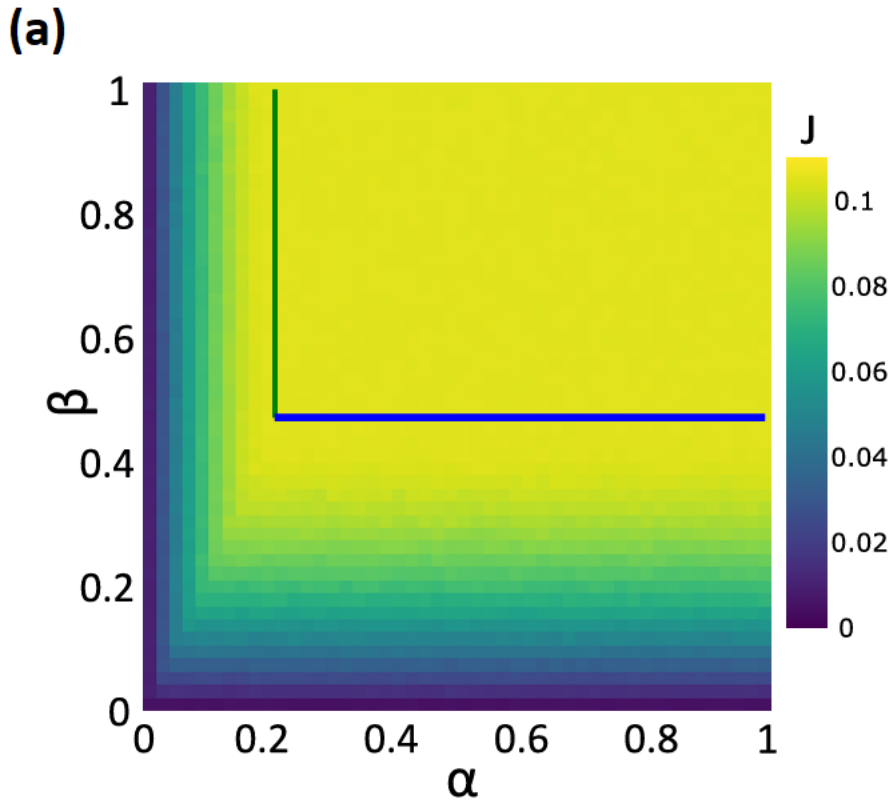

Figure 3.7: The entire phase diagram for $k_- = k_+ = \gamma$. Straight lines correspond to the predictions of the phase transitions while the heatmap represents the outcome of the simulated current

Chapter 4

Modelling finite-size effects

The following chapter will be devoted to deriving minimal models describing the finite-size effects discussed earlier. The motivation behind is twofold: firstly corroborating the hypothesis that the main source of error are indeed the finite-size effects and, secondly, to show that even in their presence it is possible to explain the system behaviour simply, without having to rely on simulation exclusively.

The approach for both periodic and open boundary conditions will rely on assuming the system to oscillate intermittently between periods of time where no paused particle is present on the lattice and "paused" periods, where at least one is present. This is intrinsically a finite-size consideration as it is clear that, as the number of particles on the lattice grows to infinity, the chances of observing no paused particles go to zero.

4.1 Periodic boundary conditions

When the particle movement is much faster than the pausing dynamics (i.e. in the biologically relevant regime $k_-, k_+ \ll \gamma$) it is reasonable to first assume that all particles cluster together behind a particle in the paused state. In this state, unpausing events of the leading particle determine an intermittent current J_c established by the detaching fronts. When finite-size effects are relevant, there is a non-negligible probability of dissolving the complete cluster, which effectively goes to zero for high density. In this case, the system will relax from the 'cluster state' to the standard TASEP state (with no particles paused), with a particle current J_0.

We address this situation by developing a two-term equation, with one term describing the current produced in the absence of pauses, weighted by the probability P_0 of having no paused particles on the lattice, and the other describing the current produced by the system in the complementary case, weighted by $1 - P_0$.

Since the particle number ρL is constant, as well as the average ratio of active particles f_a (at steady state), P_0 is given by

$$P_0 = (f_a)^{\rho L} = \left(\frac{k_-}{k_- + k_+}\right)^{\rho L} \tag{4.1}$$

and the current equation reads

$$J = P_0 J_0 + (1 - P_0) J_c \,, \tag{4.2}$$

where the current is J_0 when no particles are in the paused state, and J_c when at least one particle is paused. This hypothesis, which might seem too simplistic at first, is corroborated by numerical measurements: Fig(4.1) shows the empirical probability of observing a certain number of clusters on a lattice, as measured at the system's steady state; one can observe a bimodal distribution made of a Gaussian bell, typical with the standard TASEP, and a sharp peak on the one cluster occurrence. Thus we have some evidence of the duality between a completely clustered condition and one relaxing towards an homogeneous density. For this reason we will refer this approximation as the *single-cluster approximation*.

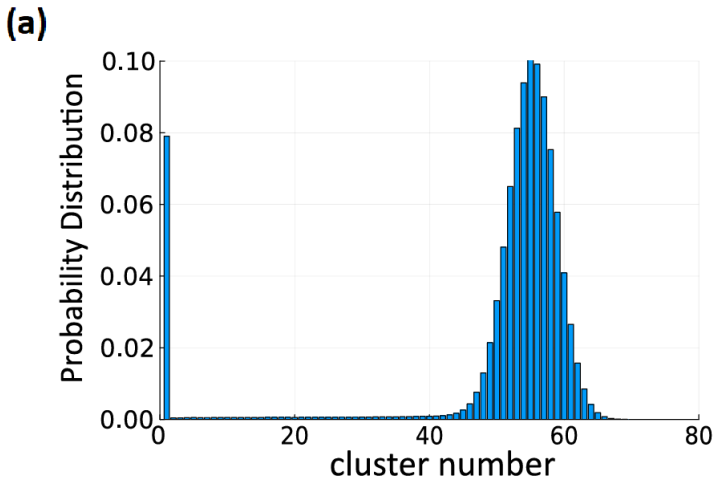

Figure 4.1: Probability distribution of the number of clusters for $L = 250$, $k^- = 10^{-3}$, $k^+ = 10^{-6}$, $\rho = 0.33$.

4.1.1 Cluster and relaxation current

When the leading particle becomes unpaused, $d + 1$ particles detach from the cluster, representing an effective density $(d + 1)/L$. Due to the aforementioned

timescale separation between the two dynamics and the modest lattice lengths of the systems of our interest, we make the hypothesis that these particles travel the $L(1-\rho)$ empty sites without becoming paused and rejoin the cluster from the other end. This sub-cluster detachment occurs intermittently, with a typical time between events given by $1/k_-$, the inverse of the unpausing rate of the particle at the front of the cluster. Putting all this together, in the clustered regime the mean particle current reads

$$J_c = k_-(1-\rho)(d+1)\,. \tag{4.3}$$

The number of particles d detaching together with the leading particle can be computed explicitly as the expected value of the number of negative results (finding a free particle) of a Bernoulli experiment before a successful one (finding a paused particle). If x is said random variable then we know that it is distributed as a negative binomial distribution and its expected value is:

$$\langle x \rangle = \sum_{i=0}^{+\infty} (f_a)^i (1-f_a)^i \, i\,.$$

However, as the cluster is finite, we constrain $d+1 \leq N$, i.e. $d \leq \rho L - 1$. We thus obtain

$$d = \sum_{i=0}^{\rho L-2} (f_a)^i (1-f_a) i + \sum_{i=\rho L-1}^{+\infty} (f_a)^i (1-f_a)(\rho L - 1)\,,$$

where the last sum bounds x with $(\rho L - 1)$ in the Bernoulli experiment. After using the geometric series (and semi-series) identities we obtain:

$$d = \frac{k_-}{k_+}\left(1 - \left(\frac{k_-}{k_- + k_+}\right)^{\rho L}\right)\,, \tag{4.4}$$

which reduces to k_-/k_+ for clusters of infinite length. The value of d can then be plugged into Eq. (4.3) to evaluate J_c.

To obtain the relaxation current J_0 we model the dissolution of the cluster and study how the density profile relaxes to the homogeneous profile of the standard TASEP. Put differently, we compute the time-dependent particle current $J_0(t)$ which relaxes to the steady state value given by Eq. (4.2). The evolution of density profiles of the exclusion process can be studied in the hydrodynamic limit [18], considering $\rho(x,t)$ as a continuous variable in both space and time. The derivation of the time dependent density profile $\rho(x,t)$ will be the subject of the next section. Assuming it to be known for the moment we can compute the instantaneous current:

$$j(t) = \frac{\gamma}{L} \int_0^L \rho(x,t)\,(1-\rho(x,t))\; dx\,, \tag{4.5}$$

which is the space-averaged current at time t.

We are interested in the current from time 0, that is when the fully packed cluster (density 1) starts to dissolve due to the absence of paused particles, to the expected time required for a new pause, which on average will occur after a time $\tau = \frac{1}{\rho L k_+}$ that corresponds to the inverse of the rate of pausing of any of the ρL particles. Making use of the aforementioned density profile one obtains:

$$j(t) = \begin{cases} \frac{\gamma^2}{6L} t & \text{if } t \in [0, \frac{\rho L}{\gamma}) \\ \gamma\rho\left(1 - \frac{2}{3}\sqrt{\frac{\rho L}{\gamma t}}\right) & \text{if } t \in \left[\frac{\rho L}{\gamma}, \frac{L}{4\rho\gamma}\right) \\ \gamma\left[\rho(1-\rho) - \frac{L}{48\gamma^2 t^2}\right] & \text{if } t \in \left[\frac{L}{4\rho\gamma}, +\infty\right) \end{cases}. \tag{4.6}$$

Finally, J_0 in Eq. (4.2) is given by the function

$$J_0(\tau) = \frac{1}{\tau} \int_0^\tau j(t)\, dt \tag{4.7}$$

evaluated at $\tau = \frac{1}{\rho L k_+}$.

We now have all components of Eq. (4.2), which can thus be tested against numerical results. Fig. 4.4 shows how the current J_0 compares to a numerical ensemble average of the current, where $t = 0$ corresponds to the instant in which the last paused particle is freed.

4.1.2 Density profile in the continuum limit

In this section we derive the aforementioned time-dependent density profile required to compute the integral in Eq.(4.5). It has been proven [1] that the evolution of the density profile, seen as a continuous variable in time and space, in the TASEP is given by the unique entropic weak solution of a generalized Burgers' equation with appropriate boundary and initial conditions. From now onward, to avoid ambiguity, we will use $\overline{\rho}$ to refer to the number of particles divided by the lattice length, while for the density profile in the continuum limit we will use $\rho(x,t)$. Thus, in the case of a clustered initial state and periodic boundaries on has:

$$\begin{cases} \frac{\partial\rho(x,t)}{\partial t} - \gamma\,(1-2\rho(x,t))\,\frac{\partial\rho(x,t)}{\partial x} = 0 & \forall t \geq 0,\ \forall x \in [0,L] \\ \rho(x,0) = 1 & \forall x \in [0, \overline{\rho}L] \\ \rho(x,0) = 0 & \forall x \in (\overline{\rho}L, L] \\ \rho(0,t) = \rho(L,t) & \forall t \geq 0 \end{cases},$$

where we have placed the beginning of the cluster on the first site for convenience.

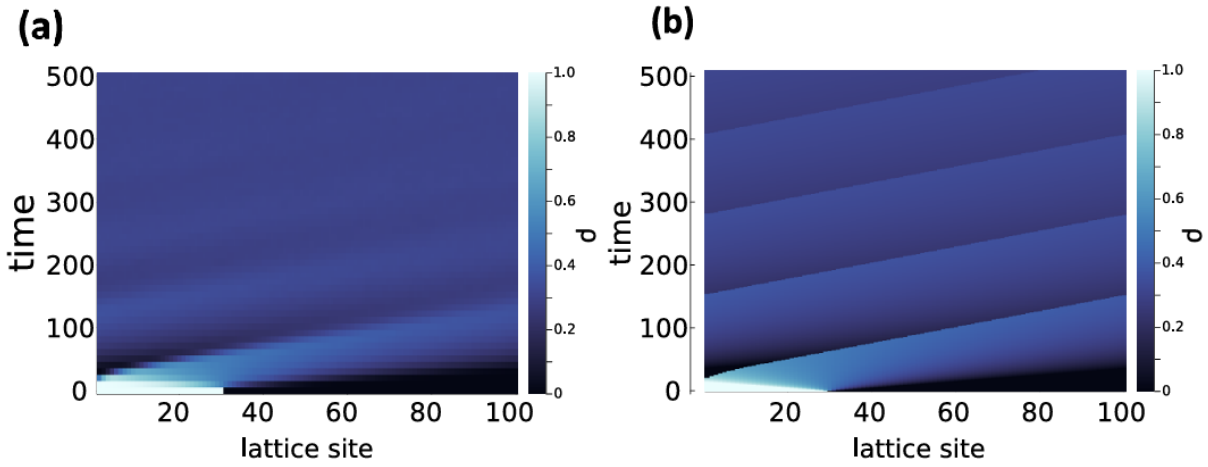

Figure 4.2: Cluster dissolution of the single-cluster configuration ($L = 100$, $\rho = 0.3$). Panel (a) shows the numerical time evolution of the cluster and Panel (b) shows the time evolution of the single cluster according to the continuum limit approach.

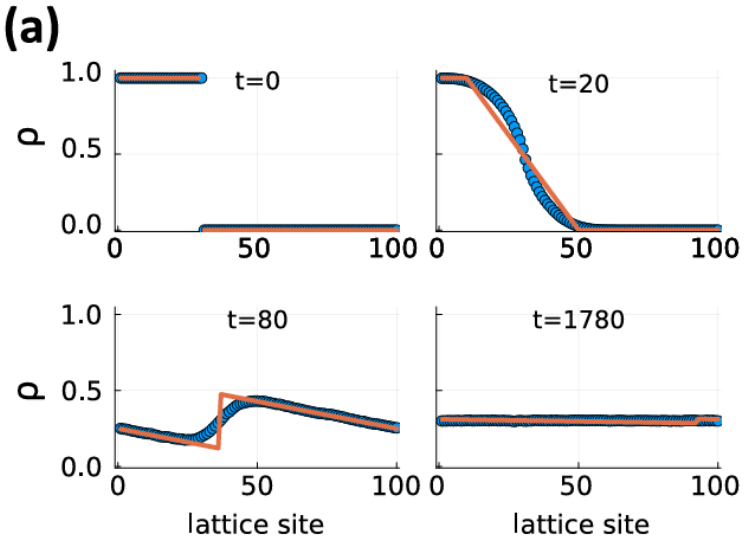

Figure 4.3: Cluster dissolution of the single-cluster configuration at four different time instants, the solid line represents the continuum limit solutions while dots are the numerical ensemble average.($L = 100$, $\rho = 0.3$).

To lighten the notation it is convenient to adimensionalize the time variable trough the transformation $t' = t\gamma$, yielding:

$$\begin{cases} \frac{\partial \rho(x,t)}{\partial t} - (1 - 2\rho(x,t)) \frac{\partial \rho(x,t)}{\partial x} = 0 & \forall t \geq 0,\ \forall x \in [0, L] \\ \rho(x,0) = 1 & \forall x \in [0, \overline{\rho} L] \\ \rho(x,0) = 0 & \forall x \in (\overline{\rho} L, L] \\ \rho(0,t) = \rho(L,t) & \forall t \geq 0 \end{cases} \tag{4.8}$$

where we have relabeled t' as t. Owing to the quasi linear nature of the Burgers' equation, the initial value contained at any given point (x_0,0) remains constant along the characteristic lines:

$$x(t) = x_0 + (1 - 2\rho(x_0,0))\, t$$

up until the point in the space-time cylinder where two distinct characteristics meet. At that point, to obtain a weak solution of Eq. (4.8), we consider a shock wave originating from that point and obeying the ODE:

$$\dot{x}(t) = 1 - \rho_l - \rho_r \tag{4.9}$$

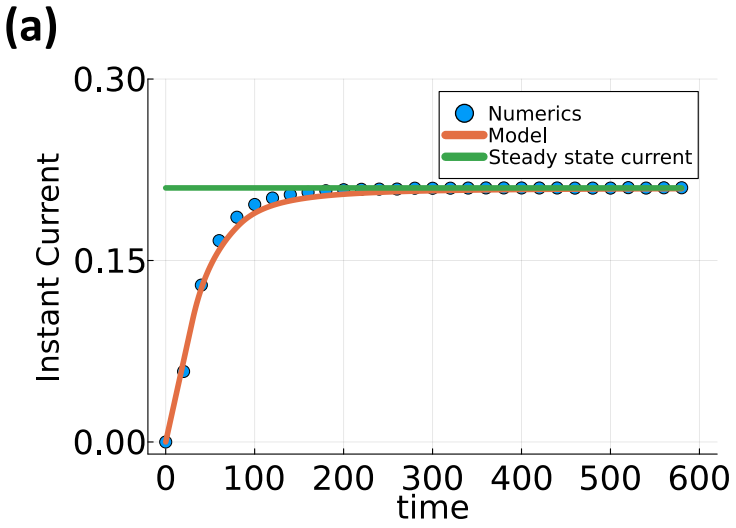

Figure 4.4: Comparison between the simulated and predicted current as a function of time, starting from a configuration with all particles clustered on the first sites. The horizontal line correspond to the expected steady-state value.

that is, the time derivative of the shock profile is the mean of the time derivatives of the characteristics to its left and to its right [1]. When two characteristics diverge one from the other, and as such there exist points on the space-time cylinder which

do not belong to the support of any characteristic line, there can be multiple weak solutions satisfying (4.8). The entropic one, which is the one actually describing the system, is obtained by connecting linearly (in space) the solutions existing at the boundaries of the divergence region, giving rise to what is often called a rarefaction fan. From $t = 0$ up until $t = \overline{\rho}L$ there will be a vertical shock separating a region with particle density $\rho = 1$ from a region with $\rho = 0$ at position $x = 0$, while a rarefaction fan will form at the point $x = \overline{\rho}L$. Thus we have

$$\rho(x,t) = \begin{cases} 1 & 0 \leq x < \overline{\rho}L - t \\ \frac{\overline{\rho}L+t-x}{2t} & \overline{\rho}L - t \leq x < \overline{\rho}L + t \ , \ t \in [0, \overline{\rho}L) \\ 0 & \overline{\rho}L + t \leq x < L \, . \end{cases}$$

Starting from $t = \overline{\rho}L$ the $0-$density characteristics will intersect with the rarefaction fan's characteristics producing a shock, which will itself exist up until the last 0-characteristic will have intersected it. Depending on the starting cluster length the shock may or may not cross the system boundary, in which case its periodicity has to be accounted for. The shock curve $s(t)$ can be computed by solving the ODE (4.9) along with the condition $x(\overline{\rho}L) = 0$, leading to:

$$s(t) = t - 2\sqrt{\overline{\rho}Lt} + \overline{\rho}L \, .$$

To understand where this equation is defined we have to compute the intersection point x^* with $x = t - (1-\rho)L$ (the last $0-$characteristic) in non-periodic space and then check how many lattice lengths are spanned by the two curves before intersecting. We obtain $x^* = \left(\frac{1}{4\overline{\rho}} - 1 + \overline{\rho}\right) L$, as such the shock will cross the boundary $n = \left[\frac{x^*}{L}\right] = \left[\frac{1}{4\overline{\rho}} - 1 + \overline{\rho}\right]$ times, where the square brackets mean that we are considering the integer part of their argument. Thus we obtain:

$$\rho(x,t) = \begin{cases} 0 & 0 \leq x < s(t) \\ \frac{\overline{\rho}L+t-x}{2t} & s(t) \leq x < t + \overline{\rho}L \\ 0 & t + \overline{\rho}L \leq x < L \end{cases}$$

for $\forall t \in [\overline{\rho}L, (1-\overline{\rho})L)$, while for $i = 1, .., n$ we have

$$\rho(x,t) = \begin{cases} 0 & 0 \leq x < s(t) - iL \\ \frac{(\overline{\rho}-i)L+t-x}{2t} & s(t) - iL \leq x < t + (\overline{\rho}-i)L \\ 0 & t + (\overline{\rho}-i)L \leq x < s(t) - (i-1)L \\ \frac{(\overline{\rho}-i+1)L+t-x}{2t} & s(t) - (i-1)L \leq x < L \end{cases}$$

for $t \in [(1-\overline{\rho})L + (i-1)L, (1-\overline{\rho})L + iL)$. Finally, we get to the time instant

$t^* = \frac{L}{4\overline{\rho}}$ where the last 0−characteristic line meets the shock:

$$\rho(x,t) = \begin{cases} \frac{(\overline{\rho}-n-1)L+t-x}{2t} & 0 \leq x < t - nL - (1-\overline{\rho})L \\ 0 & t - nL - (1-\overline{\rho})L \leq x < s(t-nL) \\ \frac{(\overline{\rho}-n)L+t-x}{2t} & s(t-nL) \leq x < L \end{cases}$$

for $t \in \left[(1-\overline{\rho})L + nL, \frac{L}{4\overline{\rho}}\right)$.

Some of the characteristics will cross the lattice to intercept other characteristics from the same rarefaction fan, leading to a new shock $z(t)$, which is going to propagate to infinity. Solving (4.9) we obtain:

$$z(t) = \left(\overline{\rho} - n - \frac{1}{2}\right) L + ct$$

and imposing the condition $z(\frac{L}{4\rho}) = x^* - nL =$ we get:

$$c = 1 - 2\rho.$$

The shock will meet the boundary at time:

$$t^{**} = \frac{L}{4\overline{\rho}} + \frac{L - (x^* - nL)}{c}$$

so we have

$$\rho(x,t) = \begin{cases} \frac{(\overline{\rho}-n-1)L+t-x}{2t} & 0 \leq x < z(t) \\ \frac{(\overline{\rho}-n)L+t-x}{2t} & z(t) \leq x < L \end{cases}$$

for $t \in \left[\frac{L}{4\overline{\rho}}, t^{**}\right)$. From t^{**} onwards the system will repeat periodically up to infinity, with a time period τ given by:

$$\tau = \frac{L}{c} = \frac{L}{1-2\rho}\,.$$

Thus $\forall t \geq t^{**}$ one has:

$$\rho(x,t) = \begin{cases} \frac{(\overline{\rho}-n-m-2)L+t-x}{2t} & 0 \leq x < c(t-t^{**}) - mL \\ \frac{(\overline{\rho}-n-m)L+t-x}{2t} & c(t-t^{**}) - mL \leq x < L \end{cases}$$

where $m = \left[\frac{t-t^{**}}{\tau}\right]$. Now all one needs to obtain Eq.(4.7) is to transform back the time unit and compute the integral.

In Fig. (4.2) we compare the predictions for the evolution of the density profile compared to simulations at four different time instants. Fig. (4.3) show the continuous evolution of the density profile in time and space for the simulated system (a) and for the analytical approach used in this appendix (b).

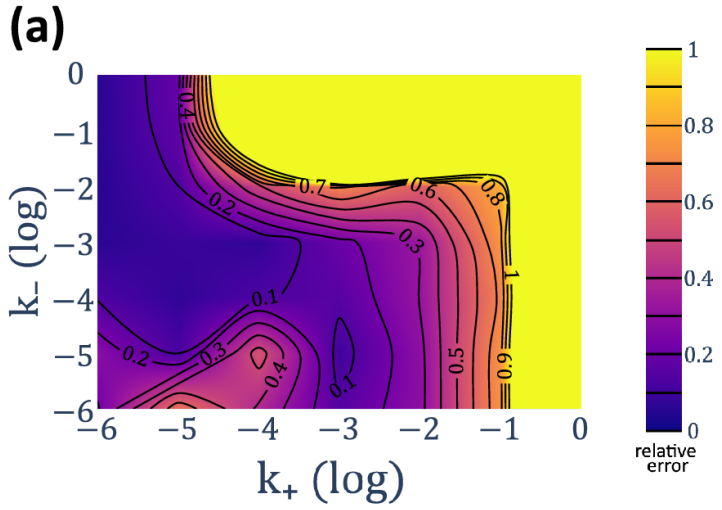

Figure 4.5: Heatmap showing the behaviour of the error made by the single cluster approximation. The lattice is 250 sites long.

4.1.3 Results

We now proceed to analyze and compare the previously developed mean-field model for the pTASEP from [26], alongside the single-cluster approximation introduced in this thesis. The single-cluster model effectively predicts the system's behavior in the presence of strong finite-size effects, as demonstrated in Fig.(4.6). To systematically evaluate the accuracy of this approximation within the k_-, k_+ parameter space, the predicted particle current is compared to simulation results across the parameter space in Fig.(4.5).

Comparing it to the previously presented heatmap in Fig.(3.3) reveals how the previous models fail to capture the finite-size effects present in the lower-left region of the parameter space, where the single-cluster approximation performs better. However, the increasing error for large values of k_- and k_+ in the single-cluster model is a consequence of the breakdown of the single-cluster assumption in regions where the separation of timescales between these rates and γ becomes unclear. In this regime, the single-cluster approximation is expected to fail, making it a complementary approach to the earlier mean-field models of pTASEP, which similarly struggle to account for finite-size effects.

The finite-size dominated regime differs not only in a quantitative sense but also qualitatively. Fig. (4.6) illustrates how, in the presence of finite-size effects, the typical bell-shaped current-density relationship of TASEP now exhibits a "bump" at low densities. This phenomenon can be understood by noting that the current during the relaxation phase is significantly higher than that generated by the

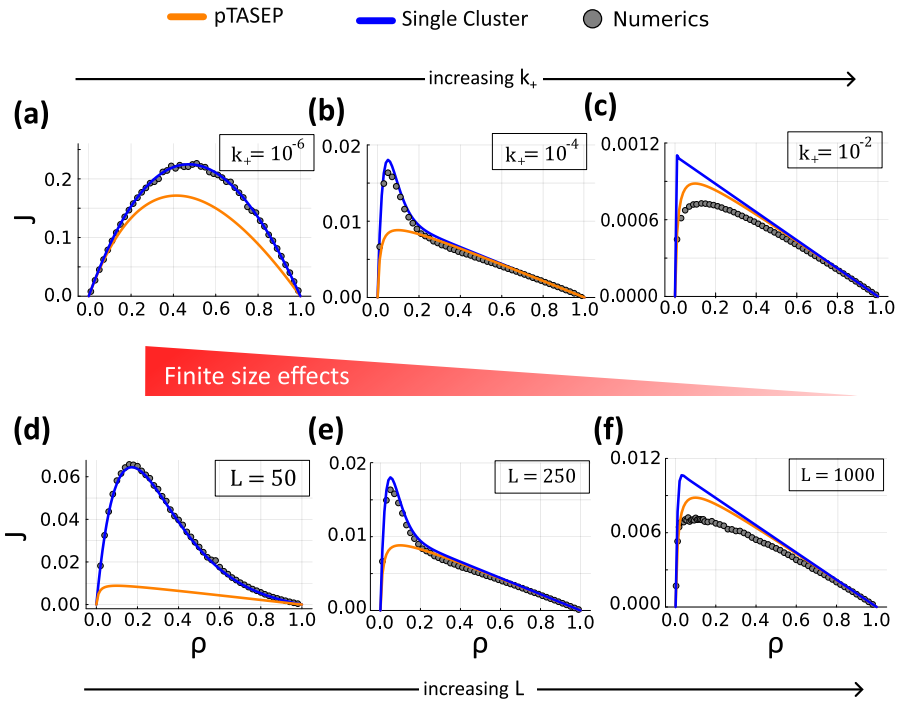

Figure 4.6: Impact of finite-size effects on the $J(\rho)$ relationship. The pTASEP solution is represented in blue, the single-cluster approximation in orange, and the numerical simulation in grey. Panels (a) to (c) illustrate the emergence of finite-size effects as k_+ is varied from 10^{-6} to 10^{-4} and 10^{-2}, with $L = 250$ and $k_- = 10^{-3}$ held constant. Panels (d) to (f) show a similar emergence of finite-size effects as the system size L is varied, while keeping $k_- = 10^{-3}$ and $k_+ = 10^{-4}$ fixed.

breaking fronts in the clustered phase. Consequently, in addition to the usual competing effects, where increasing particle density enhances movement but also introduces traffic, this regime introduces an increased likelihood of long clusters forming due to pauses.

4.2 Open boundary conditions

Handling open boundaries is significantly more complicated than the periodic case, as the number of particles on the lattice is no longer conserved over time. This makes it impossible to use a fixed particle count to calculate probabilities or to assume that the pausing dynamics have reached a steady state.

However, within the parameter range of our interest, we expect that the single-blockage behavior observed for periodic boundary conditions can also apply to the

open boundary case. In this scenario, the system can exist in one of two states: a 'clogged' state (similar to the clustered phase seen in periodic boundaries), where the lattice is filled up to the first paused particle, and a pause-free state, where the system behaves like the standard TASEP.

Due to the separation of timescales between k_-, k_+, and γ, we can assume that particles almost instantaneously fill the lattice up to the position of the first paused particle. Similarly, particles downstream of the rightmost paused particle quickly exit the system. Following the same reasoning as used to derive Eq. (4.2), we introduce a two-term model for the current:

$$J = \frac{\tau_0}{\tau_c + \tau_0} J_0 + \frac{\tau_c}{\tau_c + \tau_0} J_c \,, \tag{4.10}$$

where τ_0 and τ_c represent the average durations of the pause-free state and the clogged state, respectively. The pause-free phase behaves like the standard TASEP, so that J_0 is given by the standard TASEP's steady-state current.

To evaluate the current in the clogged phase, let x denote the position where a particle first transitions into the paused state, starting from the pause-free regime. The position x (measured in lattice sites) also corresponds to the number of particles that form the initial cluster. In the standard TASEP, particle density is uniform, and since each particle has an equal probability of pausing, the average value of x is $\langle x \rangle = L/2$. The current in the clogged state, J_c, can thus be approximated as the number of particles in the cluster divided by the typical duration of the clogged state, $J_c = L/(2\tau_c)$. Substituting this into Eq. (4.10) gives:

$$J = \frac{\tau_0}{\tau_0 + \tau_c} J_0 + \frac{1}{\tau_0 + \tau_c} \frac{L}{2}. \tag{4.11}$$

The timescale τ_0 is determined by the rate at which particles transition from the pause-free state, which is proportional to the total number of particles times k_+. This gives:

$$\tau_0 = \frac{1}{\rho L k_+} = \frac{\gamma}{\alpha L k_+} \,, \tag{4.12}$$

where, in the last expression, we assume the lattice is in the low-density (LD) phase, with $\rho = \alpha/\gamma$.

The average duration of the clogged state, τ_c, is more difficult to express in a closed form and must be computed numerically. This is because τ_c can be shorter or comparable to the time required for the paused particle fraction f_p to reach equilibrium. After the first pausing event forms the extensive cluster, other particles may also pause, further slowing down the return to the pause-free state. The clogged state ends when all paused particles become active again and leave the lattice, along with the particles that were blocked behind them. Implicit in

this reasoning is the assumption that active particles not trapped in the cluster will exit the system without pausing again.

To compute τ_c, we first calculate the time evolution of f_p, starting from a state where all particles are active:

$$f_p(t) = \frac{k_+}{k_+ + k_-}\left(1 - e^{-(k_+ + k_-)t}\right), \tag{4.13}$$

which asymptotically approaches the steady-state value of $f_p = \frac{k_+}{k_+ + k_-}$. At any given time t, the mean number of particles d_t trapped between the first and second paused particles can be estimated using the negative binomial distribution, which gives the average number of failures before another paused particle appears. This is given by:

$$d_t = \frac{1 - f_p(t)}{f_p(t)}. \tag{4.14}$$

The average lifetime $\tau(x)$ of a paused particle located at position x is $t_- = 1/k_-$, and the expected number of particles between two paused particles is given by d_t evaluated at $t = t_-$. With this, we can numerically estimate the time required to free all particles trapped in the clogged state. Specifically, after the leading particle becomes active, the number of freed particles is $d_1 + 1$, where d_1 represents the number of particles between the first and second paused particles at time $t_1 = t_-$. If $d_1 + 1$ exceeds x, we stop and set $\tau(x) = t_-$. Otherwise, we compute subsequent distances d_i, evaluating Eq. (4.14) at times $t_i = i \times t_-$, and continue until the total number of freed particles exceeds x. The duration $\tau(x)$ is then approximated as $j \times t_-$, where j is the number of intervals needed.

Finally, averaging $\tau(x)$ over all possible cluster sizes gives the value of τ_c to be used in Eq. (4.10). Although this method relies on numerical calculations, it is still significantly faster than directly simulating the entire system, and its success underscores both the importance of finite-size effects and the validity of our approximation.

4.2.1 Results

Our theory is in good agreement with the simulations, particularly for small lattices. In Fig. (4.7), we demonstrate that the two-state approximation in Eq. (4.10) accurately predicts the particle current J, even for small systems ($L = 50$), and when the binding rate k_+ is varied over several orders of magnitude.

The primary interest in this work lies in establishing a clear relationship between the particle current and the binding rate k_+. This is motivated by findings in the literature suggesting that the concentration of antibiotics in the cytoplasm directly affects the binding rate. As the antibiotic concentration increases, the likelihood

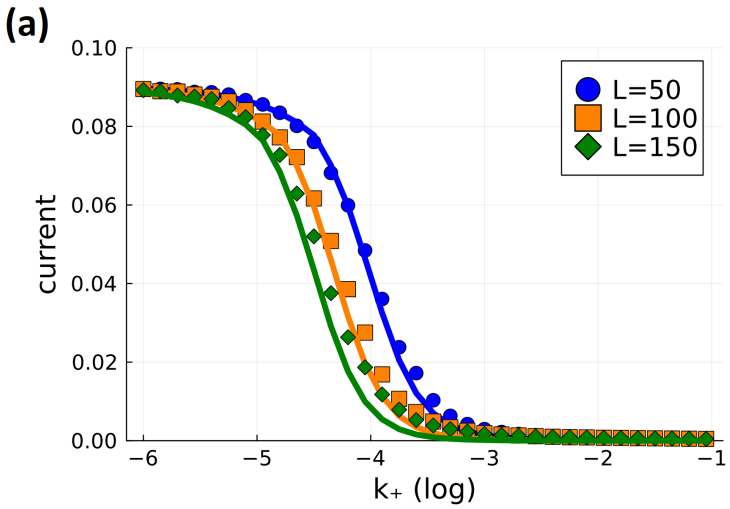

Figure 4.7: The current as a function of k_+ is compared for various system sizes, with parameters $k_- = 10^{-3}$, $\alpha = 0.1$, and $\beta = 1$. The solution provided by Eq. **??** shows good agreement for small system sizes but its accuracy diminishes as the lattice size increases

of particles entering the paused state rises, leading to a greater binding rate k_+. In contrast, the unbinding rate k_- appears to be independent of concentration and is instead determined by the specific choice of antibiotic. This makes it critical to understand how the system's behavior, particularly the current J, responds to variations in the binding rate, as it may offer insights into how different antibiotics influence intracellular processes.

For small lattices, where finite-size effects dominate, our approximation shows strong predictive power. However, for larger lattices, the accuracy of the two-state model diminishes, as the approximation fails to capture the full complexity of the system in this regime. Despite this limitation, the results presented here demonstrate that the model provides a useful framework for understanding the impact of binding and unbinding dynamics on particle transport, particularly in scenarios where the binding rate is subject to external modulation, such as in the presence of varying antibiotic concentrations.

Chapter 5

Algebraic treatment and future expansions

In this chapter we provide novel approach to achieve the normalization function for a TASEP- like system, giving it a completely combinatorial meaning. We apply this line of reasoning to the standard TASEP and we show the it provides a promising way to attempt to recover known results without making use of the *Matrix Factorization Ansatz*. The interest to do so lies in the hope to one day develop a resolution method applicable to still unsolved problems where the aforementioned Ansatz is not valid.

5.1 Transition matrix co-factors and the adjoint matrix

Consider a continuous time, ergodic stochastic process, taking place in a finite state space χ. Call $P(t) \in [0,1]^n$, where $n = |\chi|$, the probability distribution on the states of the system at time t. The probability distribution is governed by the master equation:

$$\frac{dP(t)}{dt} = AP(t), \tag{5.1}$$

where $A \in \mathbb{R}^n$ is called the transition rate matrix and $A_{i,j} = \lambda_{j\to i}$ is the rate of transition from configuration j to configuration i. Being that the process is assumed to be ergodic [2] there exists a unique steady state probability vector $\overline{P}$, such that, no matter the initial condition, we have $\lim_{t\to\infty} P(t) = \overline{P}$ and

$$A\overline{P} = 0. \tag{5.2}$$

Equation (5.2) can, in theory, always be solved exactly, giving:

$$\overline{P}_i = \frac{\det(A^{(i,i)})}{\sum_{j=1}^{n} \det(A^{(j,j)})} = \frac{f_i}{\sum_{j=1}^{n} f_j}, \tag{5.3}$$

where $A^{(i,j)}$ refers to the square matrix obtained by removing the i-th row and j-th row from A and $f_i := det(A^{(i,i)})$. The value f_i is called the statistical weight of configuration i and $Z = \sum_{j=1}^{n} f_j$ is the associated normalization function. Its worth noting that [2] the normalization function Z corresponds, up to a sign, to the coefficient c_{n-1} of the linear term of the characteristic polynomial of A, that is:

$$p_A(\lambda) = \lambda^n + \sum_{i=1}^{n-1} c_i \lambda^{n-i},$$

where in our case $c_n = 0$ as it is the determinant of A. As it was shown in the first chapter the normalization function is extremely informative in and on itself, providing for example the locations in the phase space of phase transitions. As such the focus of the chapter will be on studying the normalization function Z rather then specific physical observables.

5.2 Faddeev-LeVerrier algorithm and Bell's polynomials

The Faddeev-Leverrier algorithm is a method to construct recursively the coefficients of the characteristic polynomial of a matrix , including the one we are after. It is of particular interest to our case as it links said coefficients to the traces of consecutive powers of the original matrix. In the next section I will give a combinatorial meaning to this method in the context of the TASEP's transition rate matrix. The algorithm states [10] that the coefficients can be found recursively trough the formulas:

$$\begin{array}{ll} N_1 = I, & c_1 = -\operatorname{tr} AN_1, \\ N_2 = AN_1 + c_1 I, & c_2 = -\frac{1}{2}\operatorname{tr} AN_2, \\ & \vdots \\ N_n = AN_{n-1} + c_{n-1} I, & c_n = -\frac{1}{n}\operatorname{tr} AN_n, \end{array} \quad . \tag{5.4}$$

where I is the identity matrix.

Interestingly the coefficient values can be given [17] a closed form combinatorial expression trough the use of the complete exponential Bell polynomials. Given a sequence $\{x_n\}$ of variables, the complete exponential Bell polynomials are defined as the coefficients of the generating function [3]:

$$\Psi(t) = e^{\sum_{m>1} x_m \frac{t^m}{m!}},$$

that is:

$$e^{\sum_{m>1} x_m \frac{t^m}{m!}} = 1 + \sum_{n>1} B_n \frac{t^n}{n!},$$

where $B_n = B_n(x_1, \ldots, x_n)$ is the n-th Bell polynomial in the variables $x_1, \ldots, x_n$. Trough direct calculation one can then find the closed form expression for the $n-$th Bell polynomial [3]:

$$B_n = n! \sum_{k=1}^{n} \sum_{\{a_1,\ldots,a_{n-k+1}\}} \prod_{i=1}^{n-k+1} \frac{x_i^{a_i}}{(i!)^{a_i} a_i!}$$

where the second sum is to be taken over all sets of integers $\{a_1, ..., a_{n-k+1}\}$ such that:

$$\begin{array}{ll} \sum_{i=1}^{n-k-1} & a_i = k, \\ \sum_{i=1}^{n-k-1} & i a_i = n. \end{array}$$

Both the coefficients in (5.4) and the Bell polynomials can be written in matrix form [17], leading to:

$$c_m = \frac{1}{m!} B_m \left(-0! tr(A), -1! tr(A^2), \ldots, -(m-1)! tr(A^m) \right).$$

Now, if A is the transition rate matrix of a stochastic system of the type described earlier, we can write the partition function as:

$$Z = c_{n-1} = \frac{1}{(n-1)!} B_{n-1} \left(-0! tr(A), -1! tr(A^2), \ldots, -(n-2)! tr(A^{n-1}) \right), \quad (5.5)$$

where n is the number of states. Out of all the ways in which the partition function can be written, Eq.(5.5) is of interest because, in the case of lattice systems such as the TASEP, the traces of the powers of the transition rate matrix can be given a combinatorial meaning, which is the subject of the next section.

5.3 Application to the TASEP

Consider a one-dimensional lattice of L sites equipped with the TASEP dynamics described in the first chapter. Let the transition rate be normalized so that the hopping rate can be taken equal to 1, while we will refer to the initiation rate as α and to the termination rate as β. To write explicitly the opposite of the transition rate matrix $H = -A$ of the system one can start by writing three matrices describing the transition between two configuration due, respectively to a particle entering, hopping and exiting the lattice. Doing so leads to:

$$h^{(l)} = \begin{pmatrix} \alpha & 0 \\ -\alpha & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad h^{(r)} = \begin{pmatrix} 0 & -\beta \\ 0 & \beta \end{pmatrix}. \tag{5.6}$$

Following [13] we can write H as a sum of terms, one for each stochastic event that can take place on the lattice. Given that the first site can experience both initiation and hopping event and that the last site can only experience termination events, there will be a total of L + 1 such terms. Thus, the matrix H can be written as:

$$H = h^{(l)} \otimes I_m^{\otimes L} + \sum_{j=1}^{L-1} I_m^{\otimes j-1} \otimes h \otimes I_m^{\otimes L-j-1} + I_m^{\otimes L-1} \otimes h^{(r)}, \tag{5.7}$$

where $\otimes$ is the *Kronecker product*. The Kronecker product of two matrices of arbitrary dimensions $A \in \mathbb{R}^{m,n}$ and $B \in \mathbb{R}^{p,q}$ is defined as the matrix:

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} \in \mathbb{R}^{mp,nq}.$$

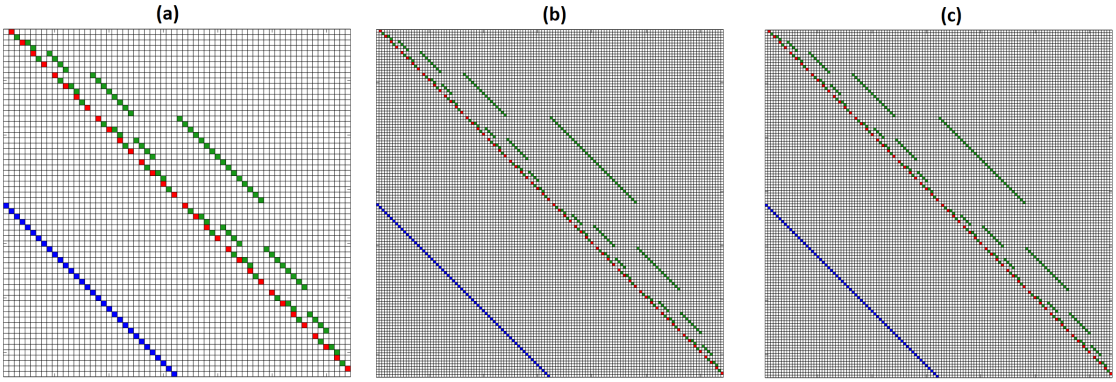

Figure 5.1: Matrices referring to systems of size 5 (a), 6 (b) and 7 (c), where the values -1, $-\alpha$, $-\beta$ are represented by the colors green, blue and red respectively. The diagonal has been set to zero for ease of representation, the diagonal elements are chosen so that all elements of a column sum up to zero.

From the definition it immediately follows that $tr(A \otimes B) = tr(A)tr(B)$. Here is as summary of properties of the Kronecker product that will be used later on [9]:

- $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$ for any scalar α and matrices A, B.
- $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ for any three matrices A, B, C.
- $(A + B) \otimes C = A \otimes C + B \otimes C$ for any three matrices A, B, C.
- If I_ℓ is the $\ell-$dimensional square identity matrix then $I_\ell^{\otimes m} = I_{\ell^m}$ and $I_\ell \otimes I_m = I_{\ell m}$.
- $(A_1 \otimes A_2 \otimes \cdots \otimes A_k)(B_1 \otimes B_2 \otimes \cdots \otimes B_k) = A_1B_1 \otimes A_2B_2 \otimes \cdots \otimes A_kB_k$ for appropriate dimensions.

Going back to the results of the previous section, it is of interest to be able to compute $tr(H^n)$ for an arbitrary integer exponent n and arbitrary system's size L as it would lead to a combinatorial expression for the normalization function trough Eq.(5.5). First, eq.(5.7) can be rewritten as

$$H = H_0 + \sum_{j=1}^{L-1} H_i + H_L,$$

now we have:

$$tr\left(H^n\right) = tr\left(H_0 + \sum_{j=1}^{L-1} H_i + H_L\right)^n = \sum_{s \in \Sigma^n} tr(s), \tag{5.8}$$

where, borrowing from jargon of trace theory [6], we call Σ^n the set of all strings of length n over the alphabet $\Sigma = \{H_i\}_{i=0}^{L}$. The next subsection will characterize a subset of Σ^n and provide its contribution to Eq.(5.8). Examples of the matrices H for some values of L can be seen in Fig.(5.1). From now onwards the set Σ^n will be referred to as a dictionary, Σ as an alphabet and the elements in Σ^n as either strings or words.

5.3.1 Algebraic properties of Σ

Given the alphabet $\Sigma = \{H_i\}_{i=0}^{L}$ defined above, a few important properties can be observed immediately. That is:

$$H_iH_{i\pm1}H_i = 0 \qquad \forall i \in \{1, \ldots, L-1\}, \tag{5.9}$$
$$H_i^m = H_i \qquad \forall i \in \{1, \ldots, L-1\}\ \forall m \in \mathbb{N}, \tag{5.10}$$
$$H_0^m = \alpha^{m-1}H_0 \qquad \forall m \in \mathbb{N}, \tag{5.11}$$
$$H_L^m = \beta^{m-1}H_L \qquad \forall m \in \mathbb{N}, \tag{5.12}$$
$$[H_i, H_j] = 0 \qquad |i - j| > 2. \tag{5.13}$$

From now onwards, a subset of Σ such that all of its elements commute one with the will be referred other to as an *independent set*. A string composed of letters from an independent set will itself be called *independent*, while one where every letter does not commute with at least one other letter will be called *dependent*. We will call a string t a rewriting of s if they are equal or if they differ by a scalar factor of the form $\alpha^a\beta^b$, for some a, b. To to be able to characterize Σ^n is important to observe:

Observation 1. *Consider $s \in \Sigma^n$. Then, either s can be rewritten such that each factor H_i, $i = 1, \dots, L-1$, appears at most twice, or $s = 0$. This rewriting allows a factorization of s into sub-strings $s = \prod s_i$, where no pair (s_i, s_j) shares any letter except H_0 and H_L, and each s_i either consists of non-repeated letters or has the form:*

$$s_i = b_{m,n}\left[H_{m-1}H_{n+1}\right]b_{m,n},$$

where $b_{m,n}$ is a permutation of the string $H_mH_{m+1}\dots H_n$ with $0 < m, n < L$.

The following proposition provides a sufficient and necessary condition for a generic $s \in \Sigma^n$ to equal the null matrix, which will be useful in the following calculations.

Proposition 1. *Given $s \in \Sigma^n$, one has $s = 0$ if and only if it contains a sub-string of the form $H_imH_{i\pm1}nH_i$, where m, n are strings of elements that commute with H_i.*

Proof. The proof will proceed by showing that the matrices in the factors described in Obs.(1) are always non-zero, to then show that so are their products. Products of independent letters are non-zero as an immediate consequence of mixed product property, as such only strings of dependent letters will be considered.

Part 1: Sub-strings with non repeated letters

To expand the products in Σ^n its convenient to rewrite the matrices in (5.6) as:

$$h^{(l)} = \alpha(a_2 + a_4),\ h^{(r)} = \beta(a_1 + a_3),\ h = a_1 \otimes a_2 - a_3 \otimes a_4, \tag{5.14}$$

where $a_1,\ a_2,\ a_3,\ a_4 \in \mathbb{R}^{2,2}$. The following multiplicative table holds:

$*$	a_1	a_2	a_3	a_4
a_1	a_1	0	a_3	0
a_2	0	a_2	0	a_4
a_3	0	a_3	0	a_1
a_4	a_4	0	a_2	0

.

Consider a string s_i in which all letters in $\{H_1, \ldots, H_{L-1}\}$ once, where the requirement for all the letters to be present is asked only to ease the notation and the results can trivially be extended to any string of independent, non repeated letters. One can use (5.14) and the multiplicative table to expand s_i in a sum where each term is a series of Kronecker products of the matrices I, a_i for $i = 1, \ldots, 4$. Lets call t the term formed by multiplying matrices of the form:

$$-\left(I \otimes \cdots \otimes I \otimes a_3 \otimes a_4 \otimes I \otimes \cdots \otimes I\right),$$

direct computation gives:

$$t = (-1)^{L-1} a_3 \otimes \left[\bigotimes_{i=2}^{L-2} b_i\right] \otimes a_4, \tag{5.15}$$

where $b_i = a_2$ if H_{i-1} comes first in t than H_i, in the opposite case $b_i = a_1$. One can see that $t \neq 0$ and it can't cancel out with other terms as no other shares its Kronecker-factorization. Direct computation shows that $H_0 t,\ tH_0,\ H_0 t H_0 \neq 0$ and
$H_L t,\ tH_L,\ H_L t H_L \neq 0$, proving that strings of non repeated letters can never equal the null matrix.

Part 2: Sub-strings with repeated letters

Consider a string s of repeated terms of the form described in Obs.(1). Again, without real loss of generality, consider the case where s contain all the letters in $\Sigma \backslash \{H_0, H_L\}$. As before on can focus on one of the terms in the expansion of s obtaining:

$$t = (-1)^{L-1} a_3 \otimes a_3 \otimes \left[\bigotimes_{i=3}^{L--3} b_i\right] \otimes a_4 \otimes a_4$$

where b_i are defined as before. By the same argument this proves that $s \neq 0$. From this formula is also clear that arbitrary products of the type of sub-strings described above will always have at least one non-zero term, thus proving the theorem. □

To proceed in computing (5.8) one needs to more information about the traces of the non-null elements of Σ^n. On that regard one has:

Proposition 2. *Let $s \in \Sigma^n$ such that $s \neq 0$. Then:*

1. *if the factor of s all commute with each other then:*

$$tr(s) = \alpha^{n_\alpha}\ \beta^{n_\beta}\ 2^{L - min(n_\alpha, 1) - min(n_\beta, 1) - 2M},$$

where M is the number of distinct symbols appearing in s other then H_0 and H_n, while n_α and n_β are the number of occurrences of letter H_0 and H_L, respectively;

2. *if s contains dependent terms and $tr(s) \neq 0$ then all letters in Σ must appear at least once.*

Proof. Consider again the expansion of s in terms of the matrices a_1, a_2, a_3 and a_4. Being that $tr(a_1) = tr(a_2) = 1$ and $tr(a_3) = tr(a_4) = 0$, a term in the expansion has to be composed only of a_2 and a_1 Kronecker-factors for its trace to be non zero. If s is a product of independent terms then there is only one such term in its expansion. Each $H_i \in \Sigma \backslash \{H_0, H_L\}$ contributes an $a_1 \otimes a_2$ factor, while $\{H_0, H_L\}$ contribute an a_2 and a_1 factor each, proving statement (1). Now consider products of independent, non repeated letters in $\Sigma \backslash \{H_0, H_L\}$, let t be one such product. Repeating letters only add to the conditions to be satisfied for $tr(s) \neq 0$, so asking for s to not contain repetition does not lead to a loss of generality. Then among the terms in its expansion the only candidate to have non-zero trace is the one formed by multiplying the factors of the form:

$$-\left(I \otimes \cdots \otimes I \otimes a_3 \otimes a_4 \otimes I \otimes \cdots \otimes I\right).$$

Calling it t_1 we have the following result:

$$t_1 = (-1)^{m-1} I \otimes \cdots \otimes I \otimes a_3 \otimes \left[\bigotimes_{i=2}^{m-2} b_i\right] \otimes a_4 \otimes I \cdots \otimes I,$$

where m is the number of letters in t_1 and b_i as stated earlier is either a_1 or a_2. It follows that $tr(t) = 0$, but if t contained all the letters in $\Sigma \backslash \{H_0, H_L\}$ then its product with H_0 and H_L (no matter their position and repetitions) would have trace equal to $(-1)^L$, multiplied by some monomial in α and β given by the number of times that H_0 and H_L are repeated. □

Point (1) of the above proposition provides a complete characterization of the independent strings in Σ^n, while point (2) implies that dependent strings contribute to Eq.(5.8) only monomial of degree between 2 and $n - L$. To proceed in the computation of Eq.(5.8), it is convenient to consider it as a sum between the contribution of independent strings and dependent ones

$$tr(H^n) = I_L^n + D_L^n, \tag{5.16}$$

where both terms in the r.h.s. are polynomials in α and β. Surprisingly, it was observed that I_L^n seems to follow a relatively simple closed form expression:

$$\begin{aligned} I_L^n(\alpha, \beta) = &\sum_{\substack{0 \leq i < L \\ i \text{ even}}} \binom{L-1}{i} \left[\left(\alpha + \frac{i}{2}\right)^n + \left(\beta + \frac{i}{2}\right)^n\right] + \\ &+ \sum_{\substack{0 \leq i < L \\ i \text{ odd}}} \binom{L-1}{i} \left[\left(\alpha + \beta + \frac{i-1}{2}\right)^n + \left(\frac{i+1}{2}\right)^n\right]. \end{aligned} \tag{5.17}$$

The above expression is, as of now, without proof and as such it should be seen as a conjecture. It was checked extensively for different choices of (L, n) trough the use of a *Julia* without revealing counterexamples.

While obtaining a closed form for D_L^n still proves elusive, the above statement gets us closer to the stated goal of solving the TASEP exactly without making use of the *Matrix Factorization Ansatz*. The above discussion about independent terms can also be extended to different, unsolved problems without apparent impediments, as it was hoped.

5.4 The case of the two sites TASEP

In this subsection, we apply the methodology proposed earlier to the case of standard TASEP dynamics on a two-site lattice. The choice of a two-site lattice is motivated by the fact that it represents the minimal system size that allows for both hopping between sites and boundary phenomena. The goal of this subsection is to clarify the proposed methodology and to compare the preliminary results with exact results, which are feasible to obtain due to the modest system size.

Applying Eq.(5.7) to this case gives:

$$H = h^{(l)} \otimes I_2 + h + I_2 \otimes h^{(r)} = H_0 + H_1 + H_2 = \begin{pmatrix} \beta & -\alpha & 0 & 0 \\ 0 & \alpha+\beta & -1 & 0 \\ -\beta & 0 & 1 & -\alpha \\ 0 & -\beta & 0 & \alpha \end{pmatrix}.$$

To be able to compute the normalization function Z_2 using Eq.(5.5) we first have to compute the quantities $tr(H^i)$ for $i = 1,\ 2,\ 3$. We take the chance given by the minute size of the system to show case the combinaotrial approach discussed earlier, that is taking advantage of Eq.(5.8). We have:

$$tr(H^2) = \sum_{s \in \Sigma^2} tr(s),$$

where $\Sigma = \{H_0, H_1, H_2\}$ and Σ^2 is the set of all words of two letters, where each letters comes from the alphabet Σ. Is easy to see that Σ^2 contains $3^2 = 9$ elements, of which 5 being independent, which are: $H_0^2,\ H_0H_2,\ H_2H_0,\ H_1^2, H_2^2$. For the reasons stated earlier we expect $tr(H^2)$ to coincide exactly with the contribution of the independent terms, that is $tr(H^2) = I_2^2(\alpha, \beta)$. We can immediately check the validity of Obs(5.17) against direct computation, observing that it is indeed true and obtaining:

$$I_2^2 = 1 + \alpha^2 + \beta^2 + (\alpha+\beta)^2 = tr(H^2),$$

where the final equality was verified trough direct computation of the matrix power.

Moving on to $tr(H^3)$, this time we expect to differ from $I_2^3(\alpha,\beta)$. In particular, according to Prop.(2), out of all the dependent words in Σ^3 only $H_0H_1H_2$, $H_0H_1H_2$, $H_1H_0H_2$, $H_1H_2H_0$, $H_2H_1H_0$ and $H_2H_0H_1$ can contribute to $tr(H^3)$, as they contain all elements of the alphabet Σ. It is thus straightforward to compute the polynomial D_2^3, as defined in Eq.(5.16), by summing the tarce of the terms above. One obtains:

$$D_2^3(\alpha,\beta) = -6\alpha\beta,$$

while from Obs.(5.17) one has:

$$I_2^3(\alpha,\beta) = \alpha^3 + \beta^3 + (\alpha+\beta)^3 + 1.$$

By direct computation the trace ends up being:

$$tr(H^3) = \alpha^3 + \beta^3 + (\alpha+\beta)^3 + 1 - 6\alpha\beta,$$

which does indeed coincide with the polynomial $I_2^3(\alpha,\beta) + D_2^3(\alpha,\beta)$, in accordance with the results of this chapter. Being that $tr(H) = 2\alpha + 2\beta + 1$, we can now proceed to apply Eq.(5.5):

$$\begin{aligned} Z_2 &= -\frac{1}{6}B_3\left(-tr(H), -1!tr(H^2), -2!tr(H^3)\right) = \\ &= \alpha^2\beta + \alpha^2 + \alpha\beta^2 + \alpha\beta + \beta^2. \end{aligned} \tag{5.18}$$

The validity of the above equation can be checked by explicitly computing the linear coefficient of the characteristic polynomial of matrix $-H$. This can be done easily trough any symbolic linear algebra routine, for example trough the use of the *polychar()* function in MATLAB. Doing so reveals that Eq.(5.18) is indeed correct.

Chapter 6

Conclusions and further developments

The first part of this work, from Chapter 1 to Chapter 4, aims to explore the feasibility of modeling the process of antibiotics-inhibited translation through the pTASEP. In doing so, we propose that the observed inadequacy of the existing mean-field theory in the parameter space of interest is due to the emergence of severe finite-size effects. This explanation is corroborated by an extensive numerical exploration of the parameter space and by the success in recovering predictive power trough the models proposed here.

In fact, we succeed in producing minimal but effective models, for both boundary conditions, describing the current of the system in the presence of extreme finite-size effects. Still, to be able to describe the actual behavior of ribosomes, the model needs further expansion, both in its definition and in the effectiveness of its approximations.

On the first front, the pTASEP as described here fails to take into account that the mRNA template itself (i.e., the lattice) has a limited lifespan, which is itself a random variable, after which it is recycled. On the second front, we lack a convincing approximate theory able to describe both the presence and the absence of finite-size effects and, most importantly, the transition between the two regimes. In addition, to truly eliminate the need for numerical simulations, an exact theory, similar to the one developed for the standard TASEP and briefly presented in this thesis, is needed.

To this end, in the final chapter, an effort was made to provide a new framework to obtain exact results in lattice systems. The primary goal, for now, is to reproduce the known standard TASEP results, but without relying on the Ansatz currently required to derive them. We propose a novel approach to obtain the normalization function for the system, which allows us to assign it a combinatorial

meaning. Although this approach is still in its early stages, we at least have a clear path for further development. In particular, the most immediate deficiencies in the results presented lie in Prop. (1) and Eq. (5.17). The first provides only a partial characterization of the elements in Σ^n, which, to be complete, requires a necessary and sufficient condition for dependent elements to have a non-zero trace, as well as a formula for said trace. The second, as of now, still lacks a rigorous proof and is corroborated only by numerical evidence. Developing a proof, other then corroborating the result, may also help in deducing an expression for the still elusive polynomial D_L^n. The lack of results for D_L^n led to the inability to recover Eq. (5.18) without making use of direct computations, as shown in Section (5.4).

In conclusion, the research conducted in this thesis provides valuable insights into the limitations of existing models for antibiotics-inhibited translation, particularly in the presence of finite-size effects. While we have succeeded in developing minimal models that capture these effects, further refinements and expansions are needed to describe biological systems more accurately. On a broader theoretical level, the novel approach to exact results in lattice systems opens a promising avenue of exploration, although significant challenges remain. Addressing the open problems identified in this work constitutes a clear direction for future research, with the potential to shed some light on still unresolved lattice systems.

Bibliography

[1] A. Benassi and JP. Foque. Hydrodynamical limit for the asymmetric simple exclusion process. *The Annals of Probability*, 1987.

[2] R. A. Blythe. Nonequilibrium phase transitions and dynamical scaling regimes. *KB thesis scanning project 2015*, 2001.

[3] Louis Comtet and J. W. Nienhuys. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. 1974.

[4] Xiongfeng Dai, Manlu Zhu, Mya Warren, Rohan Balakrishnan, Vadim Patsalo, Hiroyuki Okano, James R. Williamson, Kurt Fredrick, Yi-Ping Wang, and Terence Hwa. Reduction of translating ribosomes enables escherichia coli to maintain elongation rates during slow growth. *Nature Microbiology*, 2:16231, 2016.

[5] Bernard Derrida, Martin R. Evans, Vincent Hakim, and Vincent Pasquier. Exact solution of a 1d asymmetric exclusion model using a matrix formulation. *Journal of Physics A*, 26:1493–1517, 1993.

[6] V. Diekert and G. Rozenberg. *The Book of Traces*. G - Reference,Information and Interdisciplinary Subjects Series. World Scientific, 1995.

[7] Philip Greulich, Matthew Scott, Martin R. Evans, and Rosalind J. Allen. Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics. *Molecular Systems Biology*, 11(3):796, 2015.

[8] R J Harvey and A L Koch. How partially inhibitory concentrations of chloramphenicol affect the growth of escherichia coli. *Antimicrobial Agents and Chemotherapy*, 18(2):323–337, 1980.

[9] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[10] Shui-Hung Hou. Classroom Note:A Simple Proof of the Leverrier–Faddeev Characteristic Polynomial Algorithm. *SIAM Review*, 40(3):706–709, January 1998.

[11] Johannes Keisers, Lorenzo Vito Dal Zovo, Norbert Kern, and Luca Ciandrini. Biologically relevant finite-size effects in a driven lattice gas with particle pausing and dynamical defects. 2024.

[12] Stefan Klumpp and Terence Hwa. Stochasticity and traffic jams in the transcription of ribosomal rna: Intriguing role of termination and antitermination. *Proceedings of the National Academy of Sciences*, 105(47):18159–18164, November 2008.

[13] Klaus Krebs and Sven Sandow. Matrix product eigenstates for one-dimensional stochastic models and quantum spin chains. *Journal of Physics A: Mathematical and General*, 30(9):3165–3173, May 1997.

[14] Joachim Krug. Boundary-induced phase transitions in driven diffusive systems. *Phys. Rev. Lett.*, 67(14):1882–1885, sep 1991.

[15] Carolyn T. MacDonald, Julian H. Gibbs, and Allen C Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, 6:1–25, 1968.

[16] Carolyn T. MacDonald, Julian H. Gibbs, and Allen C. Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, 6(1):1–25, 1968.

[17] M Pathan, Hemant Kumar, Taekyun Kim, and José Bonilla. On the characteristic polynomial of chebyshev matrices. 53:1–6, 01 2023.

[18] M. R. Evans R. A. Blythe. Nonequilibrium steady states of matrix product form: A solver's guide. *J. Phys. A Math. Theor*, 2008.

[19] Mamata Sahoo and Stefan Klumpp. Asymmetric exclusion process with a dynamic roadblock and open boundaries. *J. Phys. A: Math. Theor*, 49(31):315001, 2016.

[20] Andreas Savelsbergh, Marina V. Rodnina, and Wolfgang Wintermeyer. Distinct functions of elongation factor g in ribosome recycling and translocation. *RNA*, 15(5):772–780, 2009.

[21] Gunter Schuetz. Exactly solvable models for many-body systems far from equilibrium. *Phase Transitions and Critical Phenomena*, 19:1–251, 12 2001.

[22] Matthew Scott, Eduard M Mateescu, Zhongge Zhang, and Terence Hwa. Interdependence of cell growth origins and consequences. *Science*, 330:1099–1102, 2010.

[23] Thomas R. Tritton. Ribosome-tetracycline interactions. *Biochemistry*, 16(18):4133–4138, 1977.

[24] Francesco Turci, Andrea Parmeggiani, Estelle Pitard, M. Carmen Romano, and Luca Ciandrini. Transport on a lattice with dynamical defects. *Phys Rev E*, 87(1):012705, 2013.

[25] Bartlomiej Waclaw, Justyna Cholewa-Waclaw, and Philip Greulich. Totally asymmetric exclusion process with site-wise dynamic disorder. *J. Phys. A: Math. Theor.*, 52(6):065002, 2019.

[26] Jingkui Wang, Benjamin Pfeuty, Quentin Thommen, M. Carmen Romano, and Marc Lefranc. Minimal model of transcriptional elongation processes with pauses. *Phys. Rev. E*, 90(5):050701, 2014.

[27] Herbert S. Wilf. *generatingfunctionology: Third Edition*. 1990.

[28] Daniel N. Wilson. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology*, 12(1):35–48, 2014.