An exactly solvable model for RNA polymerase during the elongation stage

IOP Publishing Phys. Biol. 22 (2025) 016001 https://doi.org/10.1088/1478-3975/ad899e

Physical Biology

OPEN ACCESS

RECEIVED
4 July 2024

REVISED
21 September 2024

ACCEPTED FOR PUBLICATION
21 October 2024

PUBLISHED
6 November 2024

PAPER

An exactly solvable model for RNA polymerase during the elongation stage

Ngo P N Ngoc[1,2], **Vladimir Belitsky**[3] and **Gunter M Schütz**[4,*]

[1] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[2] Faculty of Natural Sciences, Duy Tan University, Da Nang 550000, Vietnam
[3] Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, CEP 05508-090 São Paulo-SP, Brazil
[4] IAS 2, Forschungszentrum Jülich, 52425 Jülich, Germany
[*] Author to whom any correspondence should be addressed.

E-mail: g.schuetz@fz-juelich.de, ngopnguyenngoc@duytan.edu.vn and belitsky@ime.usp.br

Keywords: RNA polymerase, transcription elongation, cooperativity, Markov models, exclusion process, Ising measure

Abstract

We consider a Markovian model for the kinetics of RNA Polymerase (RNAP) which provides a physical explanation for the phenomenon of cooperative pushing during transcription elongation observed in biochemical experiments on *Escherichia coli* and yeast RNAP. To study how backtracking of RNAP affects cooperative pushing we incorporate into this model backward (upstream) RNAP moves. With a rigorous mathematical treatment of the model we derive conditions on the mutual static and kinetic interactions between RNAP under which backtracking preserves cooperative pushing. This is achieved by exact computation of several key properties in the steady state of this model, including the distribution of headway between two RNAP along the DNA template and the average RNAP velocity and flux.

1. Introduction

RNA polymerase (RNAP) is an enzyme that functions as a molecular motor responsible for transcribing the genetic information encoded in the DNA base pair sequence into RNA [1]. The transcription process unfolds in three distinct phases: initiation, elongation, and termination. At initiation RNAP binds to a specific region on the DNA called promoter sequence and starts separating the two strands of the DNA, thus creating the conditions for the onset of transcription elongation: After forming the transcription elongation complex (TEC), the RNAP proceeds along the DNA base pairs. Within this complex, the enzyme polymerizes the monomeric subunits of an RNA by adding nucleotides in accordance with the corresponding sequence on the DNA template. Each elongation step involves a catalytic mechanism encompassing several key stages, including as elaborated in [2–4], (1) binding of nucleoside triphosphate (NTP), (2) hydrolysis of NTP, (3) release of pyrophosphate (PP_i) as a product of hydrolysis, and (4) concurrent forward movement of RNAP along the DNA template by one base pair, called translocation. Termination marks the conclusion of the transcription elongation process, occurring when the TEC encounters a specific termination sequence and the RNAP detaches from the template DNA.

A model for the kinetics of translocation must take into account that thermal noise and other factors such as sufficient supply of NTP introduce randomness into the amount of time that is necessary to complete the mechanochemical cycle involved in a translocation step. Also, many RNAP move simultaneously on the same promoter sequence, so that one cannot ignore their mutual interactions. In particular, due to steric hindrance they cannot occupy the same region on the DNA template which is incorporated in most modeling approaches to molecular motor traffic as a hard core repulsion [5–11]. It is remarkable that with this steric excluded volume interaction alone, one can successfully capture the appearance of RNAP 'traffic jams' which is a collective phenomenon that occurs when a pausing RNAP prevents a trailing RNAP from moving forward and thus leads to a reduction of the average flux of RNAP along the DNA and consequently decreases the rate of elongation. Using the asymmetric simple exclusion process (ASEP) [12–14] as a prototypical model for molecular motors [5] this has been demonstrated in the context of protein

synthesis by ribosomes in [15, 16] and more rigorously and in considerable detail from a mathematical perspective in [17–22].

However, as empirically demonstrated both *in vitro* and *in vivo* already some time ago for *E. coli* [23–25] and for yeast [26], interactions between RNAP may also be cooperative and lead to an enhancement of the rate of elongation. This has been argued to originate in a process where trailing RNAP 'pushes' the leading RNAP out of pause sites [27–29]. To account for this mechanism in stochastic models of transcription elongation, RNAP interactions more complex than just excluded volume due to steric hindrance need to be considered.

Belitsky and Schütz in [30, 31] introduced such a model of RNAP interactions which predicts conditions under which either jamming or cooperative pushing arise. This model is a generalization of the ASEP [12–14], originally introduced in the seminal work [15] on the kinetics of protein synthesis by ribosomes and since then widely used as a starting point for modelling many different kinds of molecular motors [5]. The generalized ASEP first introduced in [30] augments the original ASEP with a next-nearest interaction and with an internal degree of freedom to account in the spirit of [32] for the mechanochemical cycle that RNAP undergoes during transcription elongation. The transition rates for translocation in this model depend on the presence of nearby RNAP to account for both blocking and pushing.

It should be stressed that on the microscopic level of interactions between individual RNAP these configuration-dependent rates describe the mutual interactions of blocking and pushing between neighboring RNAPs in an explicit way detailed below, while on the macroscopic empirical level studied in the biochemical experiments [23–26], those rates lead in an intricate way to the competing *collective* phenomena of jamming and cooperative pushing. Shedding light on the emergence of these phenomena requires a detailed mathematical analysis of the macroscopic properties of the microscopic model that is difficult to achieve by commonly used numerical simulations or analytic approximation schemes such as mean field theories.

Such a mathematical analysis is possible for the model of [30] which is an exclusion process with short-range interactions in addition to pure excluded volume interaction. Studying the stationary distribution has revealed that an enhancement of the rate of elongation cannot be explained by the mere existence of microscopic pushing between RNAP. For cooperative pushing in macroscopic to emerge, sufficiently strong repulsive interactions in addition to excluded volume interaction which are reflected in pushing above a certain critical strength are necessary. However, given the complexity of the mechanochemical translocation step that is only very partially taken into account in [30] this deeper insight into cooperative pushing is not yet fully satisfactory. The robustness of the argument for a minimal critical pushing strength and how pushing competes with potentially opposing forces still need to be probed.

Indeed, a drawback of the model of [30] is the feature of purely unidirectional motion of RNAP along the DNA template. This simplification does not capture backtracking, i.e. a backward jump of the RNAP during transcription [33, 34]. This may happen when the polymerase tries to incorporate a noncognate NTP and plays a role in error correction to enhance transcription fidelity [35]. Despite a generically small error rate [36, 37] error correction is an important process since a single mutated RNA transcript can have a large effect. Since backtracking is an upstream movement of RNAP against the mean forward (downstream) flow due to translocation, one might wonder whether backtracking would not only reduce the rate of RNA polymerization but also overcome the effect of pushing between individual RNAP and thus prevent the emergence of cooperative pushing, i.e. the boosting of the efficiency of transcription by pushing of stalled RNAP.

Backtracking has been investigated not only as error correction mechanism but in some detail in [27] on a molecular level, showing that translocation of RNAP might occur through a power stroke. A recent study based on a different exclusion process with configuration dependent rates investigated the interplay of pushing and backtracking [38]. The microscopic dynamics of that model, which was studied analytically using a mean-field approximation, bias the system to make the incorporation of noncognate nucleotides more likely. Here we address the role of backtracking in cooperative pushing by extending the model of [30, 31] to include backward translocation in a way that maintains the rigorous mathematical tractability. As an advantage of this approach we note that one can still calculate *exact* stationary bulk properties of the kinetics of transcription elongation of this more sophisticated exclusion model, without any uncontrolled approximation like mean field theory and thus explore *quantitatively* how macroscopic stationary properties arise from kinetic interactions between single RNAP that are encoded in the microscopic transition rates.

In particular, one can compute how in the presence of backtracking the flux of RNAP along the DNA template depends on the density of RNAP and the strength of interactions between them, which is the purpose of the present work. The focus is on the important role of collisions between RNAP during transcription elongation, reviewed recently in [29]. It is not intended to provide a comprehensive analysis of all microscopic mechanisms of backtracking and translocation that occur during transcription elongation but to highlight the role of the backtracking that arises from the reverse mechanochemical process

associated with the main process of translocation. The regulation of the RNAP density by the kinetics of initiation, termination [15, 39, 40], and bulk factors such as bulk attachment and detachment of RNAP [41–44], or defects [45–49] is not considered. We also neglect interactions between RNAP that may arise from DNA supercoiling [50]. Such interactions would lead to long-range interactions along the template which is out of the scope of the present framework. We only mention that exclusion processes with nontrivial exactly solvable stationary distributions and long-range interactions may be constructed along the lines described in [51]. Notice that although we treat a particular backtracking mechanism separately, we aim, in future work, at a unified treatment of such models, via the exact mathematical consistency approach used in [30, 31] and in the present paper.

To facilitate the distinction between microscopic processes involving interactions between individual RNAP and the collective macroscopic result of these interactions that can be observed in biochemical experiments e.g. in terms of the rate of elongation, and understand how the interplay of microscopic processes generate collective behaviour we shall from now on refer to *blocking* and *pushing* when refering to microscopic forces acting on RNAP and to *jamming* and *boosting* when discussing the resulting collective phenomena of a reduction or enhancement respectively of the flux of RNAP along the DNA template and thus to a corresponding change in the rate of elongation.

The paper is organized as follows. In the following section, we present the mathematical setting of our model. We begin by defining the state space of RNAP configurations allowed in the framework of the model and an exposition of the microscopic model dynamics (section 2.1). In section 2.2 we introduce the stationary distribution of the model and present the central mathematical result of this work. In section 3, we explore various stationary properties of the model, including the RNAP headway distribution, average excess, and the average elongation rate in terms of the stationary RNAP flux and discuss the impact of backtracking on these quantities. In the concluding section, we provide a concise summary of our findings and present some open problems that are triggered by the present results.

2. Methods

The basic idea of our consistency approach is to introduce a generalized Ising measure in a parametric form and determine transition rates such that this measure is stationary. The transitions that occur with these rates are chosen to mimick the translocation process of RNAP. This approach involves a series of steps. (i) We envisage the DNA template as a one-dimensional lattice with a length of L, where individual lattice sites are numbered from 1 to L. RNAPs are depicted as rods covering l_{rod} consecutive sites, reflecting the physical reality where each RNAP covers l_{rod} nucleotides [15, 52–56]. (ii) We propose a stationary distribution with *static interactions* as in [30] that take into account *static* interactions between these rods, viz., excluded volume interaction like in the ASEP and a short-range interaction with the nearest RNAP on the lattice, leading to phenomena beyond what the simple exclusion process can demonstrate. Weak logarithmic long-range interactions of entropic origin [57–60] are neglected. (iii) Following [32] we define the chemical cycle that an RNAP undergoes in each translocation step in a reduced fashion in terms of transition rates between two states in which each RNAP may exist. (iv) We postulate the transition rates governing the motion of the RNAP along the DNA template which describe *kinetic* interactions that reflect the static excluded volume and nearest neighbor interactions. For these we derive a consistency condition that ensures that the envisioned distribution is indeed stationary for the dynamics specified by those rates.

2.1. Mathematical modelling of the process

As mentioned above, we represent the DNA template as a one-dimensional lattice with a length L in units of a step length of $\delta \approx 0.34$ nm determined by the size of single base pair. In our approach, we do not differentiate between RNAP and TEC, even in the presence of the intricate TEC structure. Instead, we simplify the TEC by modeling it as a hard rod with a defined length denoted as l_{rod}. This parameter represents the extent of nucleotides covered by an RNAP. In a scenario with N RNAPs on the lattice, they are consecutively labeled by integers, with i ranging from 1 to N. Specifying the position k_i of an RNAP on the lattice then only requires knowledge of the position of the leftmost nucleotide it covers, which we refer to the position of the RNAP. Due to excluded volume interaction, no lattice site can be simultaneously covered by more than one RNAP. Furthermore, to account for the mechanochemical cycle we allow for RNAP to occur in two distinct polymerization states: one without PP_i bound (state 1) and the other with PP_i bound (state 2).

Once RNAP has released PP_i, it can advance along the DNA template by a single base pair, equivalent to a step length 1 on the lattice. In terms of our lattice model, this translocation thus implies that an RNAP positioned at location k_i in state 1 can progress one site forward, shifting from k_i to $k_i + 1$, provided that the site $k_i + l_{\mathrm{rod}}$ is unoccupied. Conversely, the RNAP can move in reverse, leading to the depolymerization of RNA from its position at k_i to $k_i - 1$, provided that the site $k_i - 1$ is vacant. This backtracking occurs only when the RNAP is in state 2, indicating that PP_i is bound to it. Therefore, the presence and status of RNAPs along the same DNA segment can be described at any moment of time by their positions and states. Thermal noise, availability of NTP and

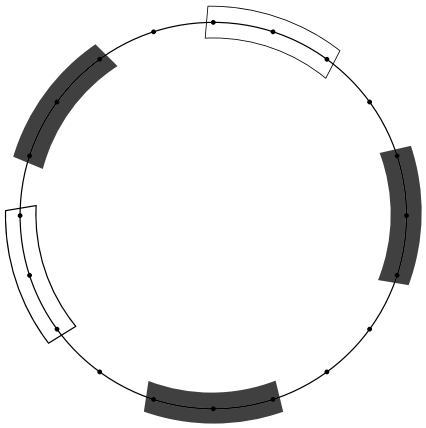


Figure 1. An allowed configuration on a ring with $L = 20, N = 5, l_{\text{rod}} = 3$. Black rods are in state 1 and blank rods are in state 2.

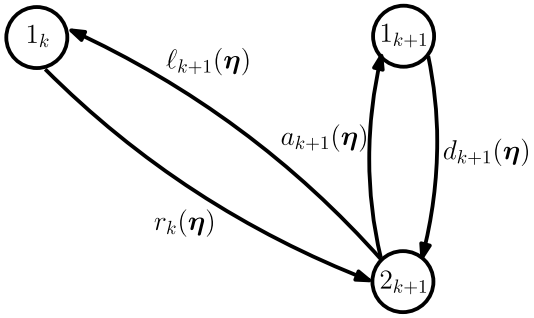


Figure 2. Minimal scheme of the mechano-chemical cycle of an RNAP. The RNAP without PP_i bounds to it is in state 1 and with PP_i is in state 2. The integer subscript k labels the position of the RNAP on the DNA template. Within this scheme, the ith RNAP, situated in state 1 at position k and denoted as 1_k, possesses the ability to move from base pair k to $k+1$. This translocation is contingent on the current system configuration and is quantified by the configuration-dependent rate $r_i(\boldsymbol{\eta})$. However, the subsequent translocation step for the RNAP can only occur following the release of PP_i, a process governed by a rate that is denoted as $a_i(\boldsymbol{\eta})$. This transition leads the RNAP from state 2_{k+1} to state 1_{k+1}. In the event that the RNAP is positioned at base pair $k+1$ and in state 2_{k+1}, it can move back to base pair k through the depolymerization of RNA. This backward movement is associated with a rate represented as $\ell_i(\boldsymbol{\eta})$, resulting in a transition from state 2_{k+1} to state 1_k. Finally, the association of PP_i is accompanied by a rate $d_i(\boldsymbol{\eta})$, enabling the transition from state 1_{k+1} to state 2_{k+1}.

other molecules required for these processes to happen leads to a translocation dynamics of individual RNAP that is subject to randomness. To define the stochastic mathematical model for this random process we refer to rods rather than RNAP and to the lattice rather than DNA template.

2.1.1. Rod configurations

We define a complete configuration of rods on the lattice, denoted as $\boldsymbol{\eta}$, by using the set of positions $k_i \in \{1,2,\ldots,L\}$ and corresponding states $\alpha_i \in \{1,2\}$ of the rods. It is important to note that if k_i represents the position of a rod, then $k_i + l_{\text{rod}} - 1$ represents the lattice position of the 'front' edge of the rod which below we refer to as the right edge of the rod, as opposed to the left edge at site k_i, which corresponds to the back edge of the rod. In an allowed configuration the ordering condition $k_{i+1} \geqslant k_i + l_{\text{rod}}$ must be satisfied due to the excluded volume rule. This constraint is expressed as $k_{i+1} \geqslant k_i + l_{\text{rod}}$ and we say that two rods i and $i+1$ are neighbors when the right end of rod i and the left edge of rod $i+1$ occupy neighboring lattice sites, i.e. when $k_{i+1} = k_i + l_{\text{rod}}$. Since we are interested only in the elongation stage of transcription, we take a lattice of L sites with periodic boundary conditions, see figure 1.

2.1.2. Transition rates for the mechanochemical cycle

The rate at which the forward step, i.e. translocation, of rod i occurs is denoted as $r_i(\boldsymbol{\eta})$. The rate of the backward movement. i.e. backtracking, is represented as $\ell_i(\boldsymbol{\eta})$. Additionally, we denote the rate of PP_i release as $a_i(\boldsymbol{\eta})$ and the rate of PP_i binding as $d_i(\boldsymbol{\eta})$. This minimal reaction scheme aligns with the description found in [3, 30, 32] for a single RNAP and is illustrated in figure 2.

Let $\boldsymbol{\eta}$ be an allowed configuration with the coordinate vector $\mathbf{k} = (k_1,\ldots,k_N)$ and state vector $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_N)$. The above-mentioned rates are of the form

$$\begin{aligned} r_i(\boldsymbol{\eta}) = {} & r\,\delta_{\alpha_i,1}\Big(1 + r^{\star\vec{\bullet}}\,\delta_{k_{i-1}+l_{\text{rod}},k_i} \\ & + r^{\vec{\bullet}\circ\star}\,\delta_{k_i+l_{\text{rod}}+1,k_{i+1}}\Big)\left(1-\delta_{k_i+l_{\text{rod}},k_{i+1}}\right), \end{aligned} \tag{1}$$

$$\begin{aligned} \ell_i(\boldsymbol{\eta}) = {} & \ell\,\delta_{\alpha_i,2}\Big(1 + \ell^{\star\circ\overleftarrow{\bullet}}\,\delta_{k_{i-1}+l_{\text{rod}}+1,k_i} \\ & + \ell^{\overleftarrow{\bullet}\star}\,\delta_{k_i+l_{\text{rod}},k_{i+1}}\Big)\left(1-\delta_{k_{i-1}+l_{\text{rod}},k_i}\right), \end{aligned} \tag{2}$$

$$\begin{aligned} a_i(\boldsymbol{\eta}) = {} & a\,\delta_{\alpha_i,2}\big[1 + a^{\star\bullet}\delta_{k_{i-1}+l_{\text{rod}},k_i} + a^{\bullet\star}\delta_{k_i+l_{\text{rod}},k_{i+1}} \\ & + a^{\star\bullet\star}\delta_{k_{i-1}+l_{\text{rod}},k_i}\delta_{k_i+l_{\text{rod}},k_{i+1}} \\ & + a^{\star\circ\bullet}\left(1-\delta_{k_{i-1}+l_{\text{rod}},k_i}\right)\delta_{k_{i-1}+l_{\text{rod}}+1,k_i} \\ & + a^{\bullet\circ\star}\left(1-\delta_{k_i+l_{\text{rod}},k_{i+1}}\right)\delta_{k_i+l_{\text{rod}}+1,k_{i+1}}\big], \end{aligned} \tag{3}$$

$$\begin{aligned} d_i(\boldsymbol{\eta}) = {} & d\,\delta_{\alpha_i,1}\big[1 + d^{\star\bullet}\delta_{k_{i-1}+l_{\text{rod}},k_i} + d^{\bullet\star}\delta_{k_i+l_{\text{rod}},k_{i+1}} \\ & + d^{\star\bullet\star}\delta_{k_{i-1}+l_{\text{rod}},k_i}\delta_{k_i+l_{\text{rod}},k_{i+1}} \\ & + d^{\star\circ\bullet}\left(1-\delta_{k_{i-1}+l_{\text{rod}},k_i}\right)\delta_{k_{i-1}+l_{\text{rod}}+1,k_i} \\ & + d^{\bullet\circ\star}\left(1-\delta_{k_i+l_{\text{rod}},k_{i+1}}\right)\delta_{k_i+l_{\text{rod}}+1,k_{i+1}}\big]. \end{aligned} \tag{4}$$

In this setting, the transitions are contingent on the configuration as given by the Kronecker-δ factors, and their rates depend on 16 parameters all of which describe the kinetic interactions between neighboring RNAPs. The notation for the rates and kinetic interaction parameters is chosen as follows.

- The subscript i on the rates refers to the rod with label i at position k_i in state α_i in the configuration $\boldsymbol{\eta}$.

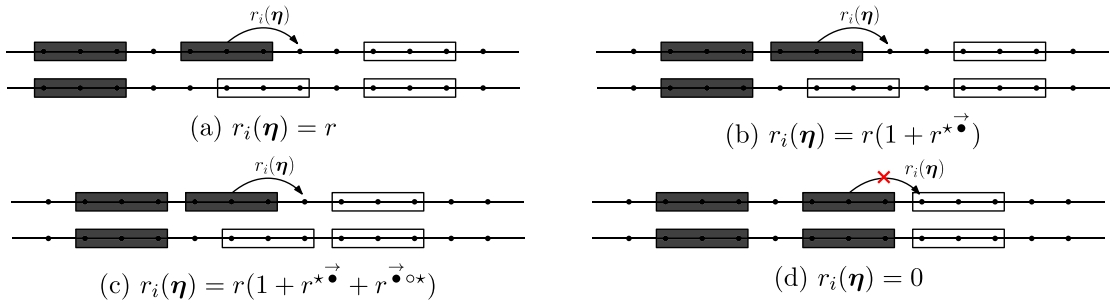


Figure 3. Some translocation rates $r_i(\boldsymbol{\eta})$ for a rod in state 1. In these visual representations, black rods are depicted in state 1, while blank rods are shown in state 2.

- The parameters r, ℓ, a, d are rates in units of seconds. They would be the transition rates of the rods if only excluded volume interaction was taken into account. Hence we call them *bare rates*.
- The parameters with superscripts are dimensionless numbers that describe the kinetic interactions by multiplying the bare rates in a way that depends on the location $k_{i\pm1}$ of the neighboring rods $i \pm 1$ as determined by the Kronecker-δ factors. We call these quantities *kinetic interaction parameters*.
- The quantities denoted by $r, r^{\star\vec{\bullet}}, r^{\vec{\bullet}\circ\star}$ determine the jump rates to the right (translocation).
- The quantities denoted by $\ell, \ell^{\star\circ\overleftarrow{\bullet}}, \ell^{\overleftarrow{\bullet}\star}$ determine the jump rates to the left (backtracking).
- The quantities denoted by $a, a^{\star\bullet}, a^{\bullet\star}, a^{\star\bullet\star}, a^{\star\circ\bullet}$, $a^{\bullet\circ\star}$ determine the release rates of PP_i.
- The quantities denoted by $d, d^{\star\bullet}, d^{\bullet\star}, d^{\star\bullet\star}, d^{\star\circ\bullet}$, $d^{\bullet\circ\star}$ determine the binding rates of PP_i.
- The superscript $\bullet$ refers to rod i and an arrow above it indicates the jump direction.
- The superscript $\star$ refers to a neighboring rod $i \pm 1$ and is placed to the right of $\bullet$ for kinetic interactions with the right neighboring rod $i+1$ and to the left of $\bullet$ for kinetic interactions with the left neighboring rod $i-1$ or on both sides for kinetic interactions influenced by both neighbors.
- The superscript $\circ$ refers to one empty site next to rod i and is placed to the right (left) of $\bullet$ for one empty site to the right (left) of rod i.
- In the absence of the superscript $\circ$ the next rod $i \pm 1$ indicated by $\star$ in the superscript to the right or left of $\bullet$ is on a nearest neighbor site (without an empty site in between) while the presence the superscript $\circ$ there is one empty site between rod i and rod $i \pm 1$, and we say that rod $i \pm 1$ is on a next-nearest neighbor site.

In figures 3 and 4 some of these transitions are illustrated, with figures 3(a)–(d) showing how the rate of translocation depends on the presence rods on neighboring sites , and figures 4(a)–(d) displaying various transitions between states 1 and 2.

The excluded volume interaction is taken into account by the overall factors $(1 - \delta_{k_i + l_{\mathrm{rod}}, k_{i+1}})$ and $(1 - \delta_{k_{i-1} + l_{\mathrm{rod}}, k_i})$ in the rates for translocation and backtracking which forbid jumps onto an occupied site, corresponding to blocking. Below we indicate this by defining hypothetical interaction parameters $r^{\vec{\bullet}\star} = \ell^{\star\overleftarrow{\bullet}} = -1$ for jumps onto occupied sites that would violate the exclusion rule. The factors $\delta_{k_{i-1}+l_{\mathrm{rod}}+1, k_i}$ capture the kinetic next-nearest neighbor interaction. The overall factors $\delta_{\alpha_i, \beta}$ ensure that the transitions between the chemical states 1 and 2 occur as described by the simplified mechanochemical cycle we consider in this work.

2.1.3. Choice of rates and kinetic interaction range

2.1.3.1. Bare rates

In the setting of Wang *et al* [3] the bare rates a, r, ℓ, and d take the values

$$r = [\mathrm{NTP}]\,(\mu\mathrm{M})^{-1}\,\mathrm{s}^{-1}, \quad \ell = 0.21\,\mathrm{s}^{-1},$$
$$a = 31.4\,\mathrm{s}^{-1}, \quad d = [\mathrm{PP}_i]\,(\mu\mathrm{M})^{-1}\,\mathrm{s}^{-1}. \qquad (5)$$

Here $[\mathrm{NTP}]$ and $[\mathrm{PP}_i]$ are the NTP and PP_i concentrations which following [32] are chosen as $[\mathrm{NTP}] = 10^{-3}$, $[\mathrm{PP}_i] = 10^{-5}$. We stress that is not the purpose of this study to predict elongation rates for any concrete biological process but to study how interactions between RNAP affect the elongation rate qualitatively. Hence we adopt these specific empirical parameters as reference constants throughout this work.

2.1.3.2. Kinetic interaction parameters

We have no empirical data on the kinetic interaction parameters at our disposal. Hence they are taken as variables and the main characteristics of the model are computed for different values of these variables to explore how the main quantitites depend on these unknown quantities which, in principle, are measurable in experiments. To ensure the positivity of the rates and ergodicity of the process for all allowed configurations, all interaction rates must be individually larger than or equal to -1 and in combination with others satisfy the inequalities $r^{\star\vec{\bullet}} + r^{\vec{\bullet}\circ\star} \geqslant -1, \ell^{\star\circ\overleftarrow{\bullet}} + \ell^{\overleftarrow{\bullet}\star} \geqslant -1, a^{\star\bullet} + a^{\bullet\star} \geqslant -1, a^{\star\bullet} + a^{\bullet\circ\star} \geqslant -1, a^{\star\circ\bullet} + a^{\bullet\star} \geqslant -1, a^{\star\circ\bullet} + a^{\bullet\circ\star} \geqslant -1, a^{\star\bullet\star} \geqslant -1$.

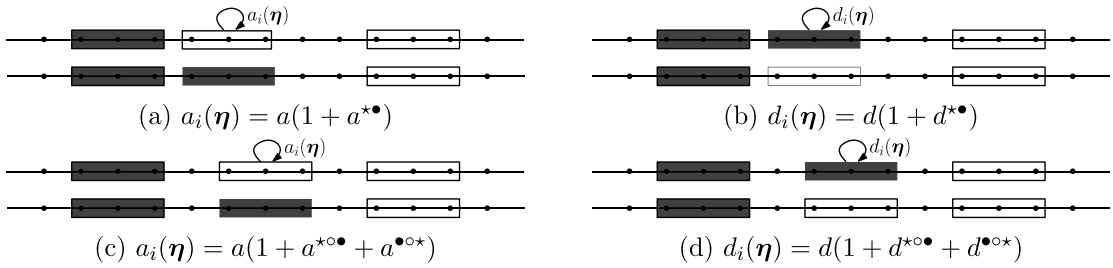


Figure 4. Some binding and release rates $a_i(\boldsymbol{\eta})$ and $d_i(\boldsymbol{\eta})$. Black rods represent RNAP in state 1, while the blank rods signify RNAP in state 2 (with PP_i bound to it).

Depending on the sign of the interaction parameters for translocation they describe kinetic repulsion or kinetic attraction as follows.

When $r^{\star\vec{\bullet}} > 0$ the bare translocation rate r is increased by the presence of a trailing RNAP which means RNAP pushing to the right. Similarly, $\ell^{\vec{\bullet}\star} > 0$ increases the bare backtracking rate ℓ in the presence of a neighboring RNAP upstream, which means RNAP pushing to the left. As mentioned in the introduction, we refer these processes, which correspond to a kinetic repulsion, as pushing, as opposed to the boosting (i.e. cooperative pushing) reported for the biochemical experiments [23–26]. We stress once more that as shown in [30], RNAP pushing on the level of individual RNAP does *not* automatically imply boosting.

When $r^{\vec{\bullet}\circ\star} < 0$ the bare translocation rate r is reduced by the presence of a next-nearest neighbor upstream RNAP. We refer to this effect as blocking enhancement as it corresponds to a kinetic repulsion that is of longer interaction range, but less strong than the full suppression of translocation in the presence nearest neighbor upstream RNAP due to excluded volume interaction. Similarly, $\ell^{\star\circ\vec{\bullet}} < 0$ corresponds to repulsive blocking enhancement for backtracking. We recall that blocking does not necessarily imply jamming.

On the contrary, when $r^{\star\vec{\bullet}} < 0$ then the bare translocation rate r is reduced by the presence of a trailing nearest neighbor RNAP, and similarly when $\ell^{\vec{\bullet}\star} < 0$ then the bare backtracking rate ℓ decreases due to the presence of a nearest neighbor upstream RNAP. These effects may be described as 'clinging', corresponding a kinetic attraction. Also $r^{\vec{\bullet}\circ\star} > 0$ and $\ell^{\star\circ\vec{\bullet}} > 0$ describe a form of kinetic attraction due to a 'pulling' by a next-nearest neighbor RNAP upstream in case of translocation or downstream in case of backtracking.

2.1.3.3. Interaction range

Notice that for $r^{\vec{\bullet}\circ\star} = \ell^{\star\circ\vec{\bullet}} = 0$ the transition rates depend only on whether the *nearest* neighbor site if rod i is occupied by another rod via excluded volume interaction and the through the interaction parameters $r^{\star\vec{\bullet}}, \ell^{\vec{\bullet}\star}$, as opposed to $r^{\vec{\bullet}\circ\star} \neq 0$ or $\ell^{\star\circ\vec{\bullet}} \neq 0$ when the transition rates depend also on occupation of the next-nearest neighbor site. We call the simplified scenario $r^{\vec{\bullet}\circ\star} = \ell^{\star\circ\vec{\bullet}} = 0$ *minimal interaction range* while otherwise speak of *extended interaction range*.

We summarize the role of the interaction terms in itemized form

- Kinetic repulsion:

$$\begin{cases} r^{\vec{\bullet}\star} = \ell^{\star\vec{\bullet}} = -1 & \text{(Blocking)} \\ r^{\vec{\bullet}\circ\star} < 0, \ell^{\star\circ\vec{\bullet}} < 0 & \text{(Blocking enhancement)} \\ r^{\star\vec{\bullet}} > 0, \ell^{\vec{\bullet}\star} > 0 & \text{(Pushing)} \end{cases} \tag{6}$$

- Kinetic attraction:

$$\begin{cases} r^{\star\vec{\bullet}} < 0, \ell^{\vec{\bullet}\star} < 0 & \text{(Clinging)} \\ r^{\vec{\bullet}\circ\star} > 0, \ell^{\star\circ\vec{\bullet}} > 0 & \text{(Pulling).} \end{cases} \tag{7}$$

Kinetic repulsion allows for repulsive forces that reach further than the on-site steric excluded volume interaction implemented by taking $r^{\vec{\bullet}\star} = \ell^{\star\vec{\bullet}} = -1$. Kinetic attraction represents a stylized form of Lennard–Jones forces which are repulsive at very close distance (blocking due to the steric excluded volume interaction), attractive at small distance (clinging and pulling), and eventually absent at larger distances. By introducing the notions of clinging and pulling we do not presume that these mechanisms exist in any specific process of transcription elongation. They are features that arise naturally in the RNAP model studied here and they may or may not have counterparts in biological systems. When all interaction parameters are taken to zero then only excluded volume interaction is taken into account.

2.1.4. Master equation

In a nutshell, the Markovian microscopic dynamics unfold as follows. Each rod is associated with four random Poissonian clocks, labeled as 1, 2, 3, and 4, each operating with configuration-dependent rates denoted as $r_i(\boldsymbol{\eta}), d_i(\boldsymbol{\eta}), \ell_i(\boldsymbol{\eta})$ and $a_i(\boldsymbol{\eta})$, respectively. When one of these four clocks for rod i activates, the following scenarios can occur:

- For a rod i in state 1:
  - If the clock is 3 or 4, no action takes place.
  - If the clock is 1, the rod i advances one site, provided that the target site $k_i + l_{\rm rod}$ is unoccupied. Consequently, the coordinate of the ith rod changes to $k_i + 1$, and its state instantly switches to 2.
  - If the clock is 2, the position of the rod remains unchanged, but its state transitions to 2.
- For a rod i in state 2:
  - If the clock is 1 or 2, there is no effect.
  - If the clock is 3, the rod i moves backward by one site, contingent on the target site $k_i - 1$ being unoccupied. Consequently, the coordinate of the ith rod becomes $k_i - 1$, and its state promptly shifts to 1.
  - If the clock is 4, the position of the rod remains unchanged, but its state changes to 1.

With this definition the master equation for the probability $\mathbb{P}_t(\boldsymbol{\eta})$ of finding the rods at time t in the configuration $\boldsymbol{\eta}$ is as follows

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(\boldsymbol{\eta},t) = \sum_{i=1}^{N}\Big[&r_i\left(\boldsymbol{\eta}^i_{\rm tlf}\right)\mathbb{P}\left(\boldsymbol{\eta}^i_{\rm tlf},t\right) + \ell_i\left(\boldsymbol{\eta}^i_{\rm tlb}\right)\mathbb{P}\left(\boldsymbol{\eta}^i_{\rm tlb},t\right)\\ &+ a_i\left(\boldsymbol{\eta}^i_{\rm rel}\right)\mathbb{P}\left(\boldsymbol{\eta}^i_{\rm rel},t\right) + d_i\left(\boldsymbol{\eta}^i_{\rm bin}\right)\mathbb{P}\left(\boldsymbol{\eta}^i_{\rm bin},t\right)\\ &- \left(r_i(\boldsymbol{\eta}) + \ell_i(\boldsymbol{\eta}) + a_i(\boldsymbol{\eta}) + d_i(\boldsymbol{\eta})\right)\mathbb{P}(\boldsymbol{\eta},t)\Big]\end{aligned} \tag{8}$$

where $\boldsymbol{\eta}^i_{\rm tlf}$ is the configuration that leads to $\boldsymbol{\eta}$ before a forward translocation of RNAP i (i.e. with coordinate $k_i^{\rm tlf} = k_i - 1$ and state $\alpha_i^{\rm tlf} = 3 - \alpha_i$), $\boldsymbol{\eta}^i_{\rm tlb}$ is the configuration that leads to $\boldsymbol{\eta}$ before a backward translocation of RNAP i (i.e. $k_i^{\rm tlb} = k_i + 1, \alpha_i^{\rm tlb} = 3 - \alpha_i$), $\boldsymbol{\eta}^i_{\rm rel}$ is the configuration $\boldsymbol{\eta}$ before PP$_i$ release at RNAP i (i.e. $k_i^{\rm rel} = k_i$ and $\alpha_i^{\rm rel} = 3 - \alpha_i$), and $\boldsymbol{\eta}^i_{\rm bin}$ is the configuration leads to $\boldsymbol{\eta}$ before PP$_i$ binding at RNAP i (i.e. $k_i^{\rm bin} = k_i, \alpha_i^{\rm bin} = 3 - \alpha_i$). Notice here that due to periodicity, the positions k_i of the rods are counted modulo L and labels i are counted modulo N. The stationary master equation, denoted below by $\hat{\pi}(\boldsymbol{\eta})$, satisfies (8) with the left hand side taken to be zero.

2.2. Stationary distribution

Following [30, 31] the stationary probability for the presence of rods at positions $\mathbf{k} = (k_1, \ldots, k_N)$ with states $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ within the configuration $\boldsymbol{\eta}$ is expressed as follows:

$$\hat{\pi}(\boldsymbol{\eta}) = \frac{1}{Z}\pi(\boldsymbol{\eta}) \tag{9}$$

where $\pi(\boldsymbol{\eta})$ is the Boltzmann weight which is of the form

$$\pi(\boldsymbol{\eta}) = \exp\left[-\frac{1}{k_B T}\left(U(\mathbf{k}) + \lambda \mathrm{B}(\boldsymbol{\alpha})\right)\right]. \tag{10}$$

The quantity T is an effective temperature that is considered as a constant. The quantity $U(\mathbf{k})$ is the static short-range interaction energy described by

$$U(\mathbf{k}) = J\sum_{i=1}^{N}\delta^L_{k_{i+1},k_i+l_{\rm rod}}. \tag{11}$$

A positive value of J corresponds to repulsive static interaction between neighboring rods. Here, δ^L represents the Kronecker symbol, computed modulo L due to the presence of periodic boundary conditions.

The quantity

$$B(\boldsymbol{\alpha}) := \sum_{i=1}^{N}(3 - 2\alpha_i) = N^1(\boldsymbol{\eta}) - N^2(\boldsymbol{\eta}) \tag{12}$$

signifies the excess in the number $N^\alpha(\boldsymbol{\eta})$ of rods in state $\alpha \in \{1,2\}$ in a configuration $\boldsymbol{\eta}$. The chemical potential λ acts as a Lagrange multiplier, which parametrizes the mean excess and describes the fluctuations of the excess that arises from the interplay of NTP hydrolysis and PP$_i$ release. The partition function

$$Z = \sum_{\boldsymbol{\eta}}\pi(\boldsymbol{\eta}) \tag{13}$$

is not needed in explicit form in the computations below. For the convenience of computation, one introduces

$$x = e^{\frac{2\lambda}{k_B T}}, \quad y = e^{\frac{J}{k_B T}}, \tag{14}$$

so that $x > 1$ corresponds to an excess of RNAP in state 1 and repulsive static interaction corresponds to $y > 1$.

To ensure that the process governed by the dynamics (1)–(4) admits a measure of the form (9) to be its invariant distribution, a price to pay is that the parameters of the model must satisfy the three consistency conditions

$$x = \frac{r + d}{\ell + a} \tag{15}$$

$$y = \frac{1 + r^{\star\vec{\bullet}}}{1 + r^{\vec{\bullet}\circ\star}} = \frac{1 + \ell^{\overleftarrow{\bullet}\star}}{1 + \ell^{\star\circ\overleftarrow{\bullet}}} \tag{16}$$

relating the parameters of the stationary distribution to the four bare rates and the four interaction parameters for translocation and backtracking and the five consistency conditions

$$\begin{aligned}xaa^{\star\bullet} - dd^{\star\bullet} = &\frac{1}{1+x}(-r + x\ell)\\ &- \frac{x}{1+x}\left(-rr^{\star\vec{\bullet}} + x\ell\ell^{\overleftarrow{\bullet}\star}\right)\end{aligned} \tag{17}$$

$$\begin{aligned}xaa^{\bullet\star} - dd^{\bullet\star} = &\frac{x}{1+x}(-r + x\ell)\\ &- \frac{1}{1+x}\left(-rr^{\star\vec{\bullet}} + x\ell\ell^{\overleftarrow{\bullet}\star}\right)\end{aligned} \tag{18}$$

$$xaa^{\star\bullet\star} - dd^{\star\bullet\star} = -rr^{\star\vec{\bullet}} + x\ell\ell^{\overleftarrow{\bullet}\star} \tag{19}$$

$$xaa^{\star\circ\bullet} - dd^{\star\circ\bullet} = \frac{1}{1+x}\left(rr^{\vec{\bullet}\circ\star} - x\ell\ell^{\star\circ\overleftarrow{\bullet}}\right) \quad (20)$$

$$xaa^{\bullet\circ\star} - dd^{\bullet\circ\star} = \frac{x}{1+x}\left(rr^{\vec{\bullet}\circ\star} - x\ell\ell^{\star\circ\overleftarrow{\bullet}}\right). \quad (21)$$

Involving the nine interaction parameters for binding and release of PP_i.

This is proved rigorously in appendix and allows us to present the main mathematical result of the present work as a formal theorem.

Theorem 2.1. *If the parameters appearing in the rates (1)–(4) satisfy the consistency conditions (15)–(21), then the invariant measure of the rod process defined by the master equation (8) is given by*

$$\hat{\pi}(\boldsymbol{\eta}) = \frac{1}{Z}\left(\frac{r+d}{\ell+a}\right)^{\sum_{i=1}^{N} -3/2+\alpha_i} \times \left(\frac{1+r^{\star\vec{\bullet}}}{1+r^{\vec{\bullet}\circ\star}}\right)^{-\sum_{i=1}^{N}\delta^{L}_{k_{i+1},k_i+l_{\mathrm{rod}}}} \quad (22)$$

where Z is the partition function.

Notice that the excess part of the stationary distribution involving the Langrange multiplyer λ depends only on the bare transition rates while the static interaction part of the stationary distribution involving the Kronecker-δ terms depends only on the kinetic interaction parameters which satisfy the symmetry (16). Remarkably, comparing this consistency condition with the role of the interaction parameters shows that repulsive *kinetic* interactions are consistent only with repulsive *static* interaction and similarly, attractive *kinetic* interactions are consistent only with attractive *static* interaction. While this is what one may expect on physical grounds the consistency conditions (15)–(21) between static and kinetic interaction parameters are *not* a feature built into the definition of the model but a purely mathematical result that comes out in the proof of the theorem. The significance of the relations between the parameters of the stationary distribution and the transition rates are discussed in the following section.

3. Results and discussion

Given an average density $\rho = N/L$ of rods of the lattice, a central quantity of interest are the statistical properties of the distance between rods, expressed in terms of the headway m_i which is the number of empty sites between neighboring rods ith and $(i+1)$th. This quantity, apart from its intrinsic interest, also determines further important properties of the stationary translocation kinetics, in particular the stationary flux related to the rate of elongation and the average excess of bound and unbound RNAP. These quantities are computed below. It is not surprising that some formulas in the present work turn out to be resemble corresponding expressions in [30, 31] since the Boltzmann factor (9) is of similar form as in those papers. However, the parameter x in our setting depends not only on the rates r, a (as in [30]) but also on the rates ℓ, d. Moreover, the value y in this work is also different from the one in [30] since it depends on the parameters $r^{\star\vec{\bullet}}, r^{\vec{\bullet}\circ\star}$ and $\ell^{\overleftarrow{\bullet}\star}, \ell^{\star\circ\overleftarrow{\bullet}}$ characterizing translocation and backtracking, respectively, while the same value in [30, 31] depends only the forward translocation.

3.1. Average excess

The simplest measure that characterizes the distribution of RNAP is the average excess density with no PP_i bound over the PP_i bound state of RNAP given by

$$\sigma = \frac{\langle N^1\rangle - \langle N^2\rangle}{L} \quad (23)$$

where N^α is the number of rods in state α. For a configuration $\boldsymbol{\eta}$ with N rods one has by definition $N^1(\boldsymbol{\eta}) + N^2(\boldsymbol{\eta}) = N$. With the second equality in the definition (12) the factorization of the Boltzmann weight (10) in the invariant measure (9) into an interaction part and the excess part with the Lagrange multiplier λ thus yields

$$\sigma = \frac{1-x}{1+x}\rho. \quad (24)$$

We denote by

$$\rho_\alpha := \langle \delta_{\alpha_i,\alpha}\rangle = \frac{1}{L}\langle N^\alpha\rangle, \quad \alpha \in \{1,2\}, \quad (25)$$

the average densities of rods in states $1, 2$. Since $\rho_1 + \rho_2 = \rho$, one gets from (24)

$$\rho_1 = \frac{1}{1+x}\rho, \quad \rho_2 = \frac{x}{1+x}\rho. \quad (26)$$

The prefactors

$$\tau_1 := \frac{1}{1+x}, \quad \tau_2 := \frac{x}{1+x} \quad (27)$$

appearing in the consistency relations for the interaction parameters (21) thus play the role of the fraction of RNAP in states 1 and 2 respectively. Correspondingly,

$$x = \frac{\rho_2}{\rho_1} \quad (28)$$

is the stationary ratio of RNAP in states 1 and 2. According to the consistency relation (15). This quantity depends only on the bare rates r, ℓ, a, d not on the interaction parameters. With the empirical values of r, ℓ, a, d as in (5) one finds $x = 31.95$.

3.2. Absence of headway correlations

As a first result we note that the headways between rods are uncorrelated. To prove this we note that due to translation invariance, an allowed configuration of RNAPs can be specified by the headway vector $\mathbf{m} := (m_1, \ldots, m_N)$ and the state vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$. Thus, one has $m_i = k_{i+1} - (k_i + l_{\rm rod}) \mod L$ and the total number of vacant sites is $M = L - l_{\rm rod} N$. We denote by

$$\theta_i^p := \delta_{m_i,p} = \delta_{k_{i+1},k_i+l_{\rm rod}+p}. \tag{29}$$

the indicator functions on a headway of length p (in units of base pair) with the index i taken modulo N, i.e. $\theta_0^p \equiv \theta_N^p$. In terms of the parameters (14) and the new distance variables (29) one rewrites the stationary distribution (9) as follows

$$\tilde{\pi}(\boldsymbol{\zeta}) = \frac{1}{Z} \prod_{i=i}^{N} \left(x^{-3/2+\alpha_i} y^{-\theta_i^0} \right) \tag{30}$$

where $\boldsymbol{\zeta}$ is an allowed configuration defined by state vector $\boldsymbol{\alpha}$ and headway vector $\mathbf{m}$. Notice that the measure (30) is of factorized form which indicates the absence of headway correlations.

As in [30, 31], we work in the grand-canonical ensemble defined by

$$\tilde{\pi}_{\rm gc}(\boldsymbol{\zeta}) = \frac{1}{Z_{\rm gc}} \prod_{i=1}^{N} \left(x^{-3/2+\alpha_i} y^{-\theta_i^0} z^{m_i} \right), \tag{31}$$

where $Z_{\rm gc} = (Z_1 Z_2)^N$ with

$$Z_1 = \frac{1 + (y-1)z}{y(1-z)}, \quad Z_2 = x^{1/2} + x^{-1/2}, \tag{32}$$

and the solution of the quadratic equation

$$(y-1)z^2 + z\left(y \frac{1 - (l_{\rm rod}-1)\rho}{1 - l_{\rm rod}\rho} - 2(y-1) \right) - 1 = 0 \tag{33}$$

given by

$$z := z(\rho, y) = 1 - \frac{1 - (l_{\rm rod}-1)\rho - \sqrt{(1-(l_{\rm rod}-1)\rho)^2 - 4\rho(1-l_{\rm rod}\rho)(1-y^{-1})}}{2(1-l_{\rm rod}\rho)(1-y^{-1})}. \tag{34}$$

Which parametrizes the density of rods. In the absence of nearest-neighbor static interaction, i.e. when RNAP only experience excluded volume interaction, this relation reduces to

$$z_0 := z(\rho, 1) = \frac{1 - l_{\rm rod}\rho}{1 - (l_{\rm rod}-1)\rho}. \tag{35}$$

By definition, for any static interaction strength the mathematically maximal density of rods is $\rho_{\max} = l_{\rm rod}^{-1}$ which expresses full coverage of the lattice by rods. For all y one has $z(0,y) = 1 \geqslant z(\rho,y) \geqslant 0 = z(\rho_{\max}, y)$. Hence for rod densities of interest, i.e. $\rho \neq 0, \rho_{\max}$ one has $0 < z < 1$ and z is strictly monotonically decreasing in ρ.

3.3. Headway distribution

Since the invariant measure is of the same form as in [30, 31], mean headway and headway distribution are the same form as in those papers as functions of the parameters x, y, z, the difference being the dependence of these parameters on the microscopic transition rates (15), (16), and the density parameter (34).

Denote by $P_h(r)$ the distribution of the headway between the right edge of a trailing rod i and the left edge of a leading rod $i+1$ which means $P_h(r) = \frac{1}{\rho} \left\langle \delta_{k_{i+1}-k_i-l_{\rm rod},r} \right\rangle = \langle \theta_i^r \rangle$, $r \in \mathbb{N}$ where $\mathbb{N} = \{0, 1, 2 \ldots\}$ are the natural numbers. This distribution depends on the rod density ρ and the interaction parameter y. However, to keep notation light we omit this dependence. From (31) one finds

$$P_h(r) = \begin{cases} \dfrac{1-z}{1+(y-1)z} & \text{for } r = 0, \\ y P_h(0) z^r & \text{for } r \geqslant 1 \end{cases} \tag{36}$$

with the mean headway

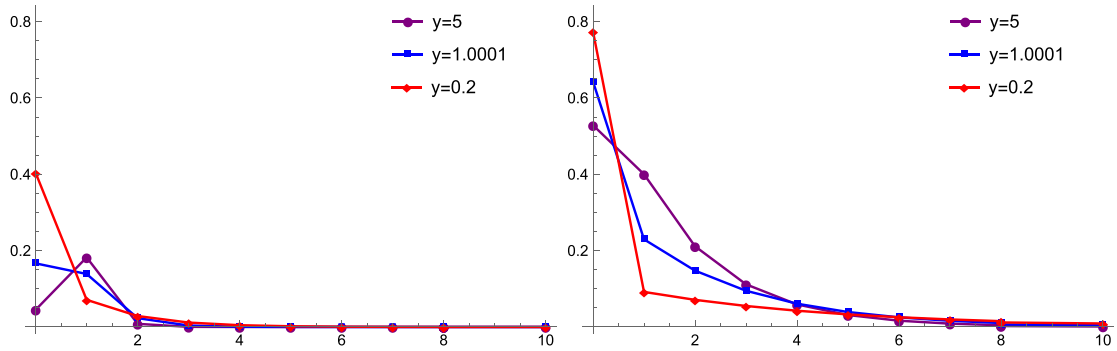
$$\overline{\lambda}(\rho) := \langle m_i \rangle = \frac{yz}{(1-z)(1+(y-1)z)} = \frac{1 - l_{\rm rod}\rho}{\rho} \tag{37}$$

in units of the lattice constant δ given by the size of a DNA base pair. The mean headway does not depend on the static interaction strength.

The nearest-neighbor probability

$$p_0 := P_h(0) = \frac{1-z}{1+(y-1)z} \tag{38}$$

of having headway 0 plays a special role. This is the probability of finding two rods as nearest neighbors which determines the mean static interaction energy density



(a) RNAP density $\rho = 0.1$. (b) RNAP density $\rho = 0.18$.

Figure 5. RNAP headway distribution $P_h(r)$ for different static interaction strengths y as a function of the integer lattice distance at average RNAP densities $\rho = 0.1$ and $\rho = 0.18$. The connecting lines are guides to the eye.

$$\overline{u}(\rho, y) := \frac{1}{LJ}\sum_{\mathbf{k}} \langle U(\mathbf{k})\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle \theta_i^0\rangle = p_0 \quad (39)$$

in units of the interaction constant J.

To examine the impact of the static interaction strength parametrized by y on the full distribution we first note that $P_h(r)$ is strictly decreasing for headways $r \geqslant 1$ independently of interaction strength and rod density. This is an entropic effect which indicates that the number of allowed configurations decreases with the headway between them. However, the probability p_0 of finding two neighboring rods depends non-trivially on the interplay of interaction strength and rod density. Since $p_0 = P_h(1)/(yz)$ and since $z < 1$, any attractive interaction (which corresponds to $y < 1$) yields $p_0 > P_h(1)$ which is indeed expected for attraction. However, somewhat contrary to intuition, the probability of finding two rods as immediate neighbors is smaller than the probabilty of finding them with an empty site between them even for repulsive interaction as long as it is not too strong. Only above a critical repulsive interaction strength that depends on the density through the relation $y > 1/z$ the next-nearest neighbor headway probability $P_h(1)$ exceeds the nearest neighbor headway probability p_0. This effect is demonstrated in figure 5 for rods of length $l_{\mathrm{rod}} = 5$ for two different rod densities and three different interaction parameters $y = 0.2$ (attractive static interaction), $y = 1.0001$ (very weak repulsive interaction), and $y = 5$ (strong repulsive static interaction). Since backtracking is reduced when there are many neighboring rods we conclude that backtracking due to neighbor depletion sets in for strong repulsive static interactions above the critical value $y_c = 1/z$.

3.4. Average elongation rate

The main quantity of interest is the average elongation rate which is related to the flux of rods along the chain. To elucidate the effect of backtracking we first compute for the mean velocity of a single RNAP which experiences no interaction with another RNAP and how it changes qualitatively if we assume a rate of backtracking ℓ different from the empirical value reported in [3].

3.4.1. Mean velocity of a single RNAP

For a single rod the process reduces to a biased random walk of a particle with an internal degree of freedom that is given by the two chemical states ini which the RNAP can be. Following the approach of Wang *et al* [3], one finds by straightforward computation

$$v_0 = \frac{ra - \ell d}{r + \ell + a + d} = r\tau_1 - \ell\tau_2 = \frac{r - \ell x}{1 + x}. \quad (40)$$

If an RNAP would perform a simple random walk, then its velocity would be $v_0 = r - \ell$ which differs from (40) by the prefactors $\tau_1 = \dfrac{\rho_1}{\rho}, \tau_2 = \dfrac{\rho_2}{\rho}$ which are the number fraction of the chemical states 1 and 2, respectively. This difference quantifies the effect of the mechanochemical cycle on the average velocity v_0 of an RNAP. At low density ρ of RNAP, i.e. in a scenario when RNAP would almost never become neighbours on the DNA strand, the RNAP flux is then given by

$$j_0 = \rho v_0. \quad (41)$$

We can also read off the effect of backtracking on the velocity of a single RNAP. To quantify this effect we denote by v_0^{ref} the hypothetical velocity in the absence of backtracking ($\ell = 0$) and by $v_0(10)$ the velocity for strong backtracking for which we take a tenfold backtracking rate 10ℓ compared to the empirical rates reported in [3]. This yields the ratios

$$\frac{v_0}{v_0^{ref}} = 0.9997, \quad \frac{v_0(10)}{v_0^{\mathrm{ref}}} = 0.974. \quad (42)$$

Hence assuming a complete absence of backtracking yields no perceptible change in the avarage velocity of a single RNAP. The reduction of the velocity for strong backtracking compared to v_0^{ref} is small (about 2.6%) even though the backtracking rate has been

taken as 10 times the empirical rate ℓ. This observation leads us to conclude that the effect of backtracking on the velocity of single RNAP is, at most, small.

3.4.2. Average velocity

To elucidate how the interplay of backtracking with the static and kinetic interactions affects the average elongation rate we investigate the average velocity v of an RNAP in an ensemble of interacting RNAP, as opposed to the single RNAP discussed above. The stationary average flux

$$j = \rho v. \tag{43}$$

of RNAP along the DNA template is the average number of RNAP crossing a lattice bond per unit time (second) [32] is thus a measure of the average elongation rate.

In the framework of our model, j is given by the expectation of the right and left rod jump rates $r_i(\boldsymbol{\eta})$, $\ell_i(\boldsymbol{\eta})$ with respect to the stationary distribution (9) through the difference $j = \langle r_i - \ell_i \rangle$. From the definitions (1), (2) one has

$$\begin{aligned} j = {} & \frac{r}{L}\left\langle N_1\left(1 + r^{\star\vec{\bullet}}\theta^0_{i-1}\right)\left(1-\theta^0_i\right) + r^{\vec{\bullet}\circ\star}\theta^1_i\right\rangle \\ & - \frac{\ell}{L}\left\langle N_2\left(1 + \ell^{\overleftarrow{\bullet}\star}\theta^0_i\right)\left(1-\theta^0_{i-1}\right) + \ell^{\star\circ\overleftarrow{\bullet}}\theta^1_{i-1}\right\rangle \end{aligned} \tag{44}$$

where we have used that $\delta_{k_i+l_{\rm rod},k_{i+1}} \cdot \delta_{k_i+l_{\rm rod}+1,k_{i+1}} = \delta_{k_{i-1}+l_{\rm rod},k_i} \cdot \delta_{k_{i-1}+l_{\rm rod}+1,k_i} = 0$. The expectation (44) does not depend on the rod i because of stationarity and the conservation of the total number of rods during translocation.

The factorization property of the stationary distribution (30) allows for expressing the expectation of the products appearing in this formula by the product of expectations involving the stationary headway probabilities $\langle\theta^0_i\rangle = P_h(0) = p_0$, $\langle\theta^1_i\rangle = P_h(1) = yzp_0$ given by (36). The headway distribution (36) and the consistency relation (16) then yields

$$j = j_0 z\left[1 + \left(r\rho_1 r^{\star\vec{\bullet}} - \ell\rho_2 \ell^{\overleftarrow{\bullet}\star}\right)(1-z)\frac{y+1+(y-1)z}{\left[1+(y-1)z\right]^2}\right] \tag{45}$$

$$v = v_0 z\left[1 + \left(r\tau_1 r^{\star\vec{\bullet}} - \ell\tau_2 \ell^{\overleftarrow{\bullet}\star}\right)(1-z)\frac{y+1+(y-1)z}{\left[1+(y-1)z\right]^2}\right] \tag{46}$$

in terms the kinetic interaction parameters for pushing.

To discuss the effect of the microscopic interactions on the collective behaviour of a stationary ensemble of RNAP in terms of the average RNAP velocity (46) we define the *interaction factor*

$$\begin{aligned} q := {} & \frac{v}{v_0} = \frac{j}{j_0} \\ = {} & z\left[1 + \frac{r\tau_1 r^{\star\vec{\bullet}} - \ell\tau_2 \ell^{\overleftarrow{\bullet}\star}}{r\tau_1 - \ell\tau_2}(1-z)\frac{y+1+(y-1)z}{\left[1+(y-1)z\right]^2}\right] \end{aligned} \tag{47}$$

which quantifies how much the velocity (or flux) is affected by the presence of RNAP interactions. Notice that q depends both on the average RNAP density ρ an the various rates and interaction parameters that define the microscopic interactions between individual RNAP. Thus q characterizes whether the average velocity of an interacting system of RNAP is enhanced or reduced compared to a hypothetical system of noninteracting RNAP that is effectively described by translocation at very low RNAP density. We speak of boosting when $q > 1$ for a range of RNAP densities and system parameters and of jamming when $q < 1$.

3.5. Role of backtracking for boosting

To examine the relationship between microscopic backtracking and the collective behaviour that leads to boosting we work with the length of RNAP $l_{\rm rod} = 5$ and use the values of r, ℓ, a, d given in (5) and below. Kinetic interaction strengths are varied in different ways and the effect on boosting is discussed for the full range of rod densities, ranging from 0 to the maximal rod density $1/\ell = 0.2$

3.5.1. Minimal kinetic interaction range

It is interesting that if $r^{\vec{\bullet}\circ\star} = \ell^{\star\circ\overleftarrow{\bullet}} = 0$ (minimal kinetic interaction range), from identity (16) one has $r^{\star\vec{\bullet}} = \ell^{\overleftarrow{\bullet}\star} = y - 1$. Then the average velocity takes the simple form

$$v = v_0 z\left(\frac{y}{1+(y-1)z}\right)^2 \tag{48}$$

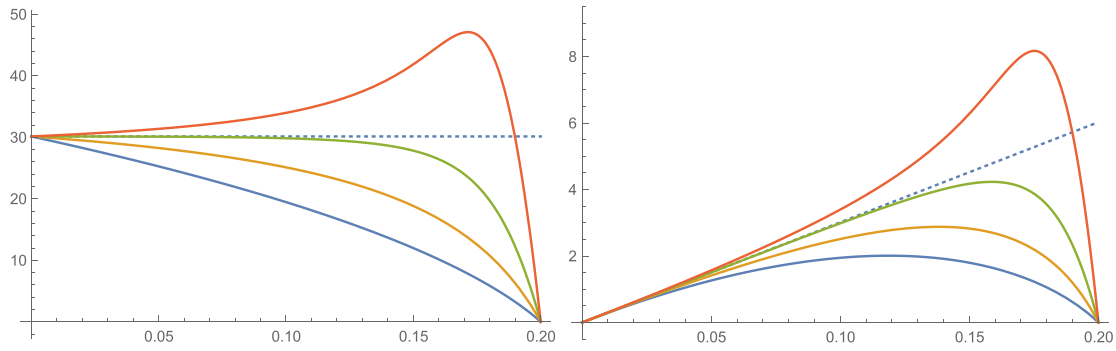
where z given by (34) is a function of the RNAP density and static interaction strength. Thus the interaction factor (47)

$$q = z\left(\frac{y}{1+(y-1)z}\right)^2 \tag{49}$$

is a function only of the RNAP density and the interaction terms of the model and can be expressed in terms of the density and the static interaction alone.

The result (48) for the average velocity shows that backtracking manifests itself through the bare backtracking rate ℓ in the overall amplitude v_0 given by the velocity of an isolated RNAP. However, as discussed above, the effect is so small that curves for different values of ℓ would collapse onto the same curves shown in figure 6 for ℓ given by (5). Moreover, this particular effect of backtracking has no impact on boosting.

To examine backtracking affects on boosting through the RNAP interactions we first note that by definition, $q = 1$ at density $\rho = 0$. As a function of the density, the interaction factor has a maximum at a density ρ^* given by $z = 1/(y-1)$ where the derivative of w.r.t. the density vanishes. Since $0 \leqslant z < 1$, this can happen only if $y > 2$ which implies that the velocity increases and reaches a maximum $v^* > v_0$ only



(a) Average velocity v (b) Average flux j

Figure 6. RNAP velocity v and RNAP flux j as functions of RNAP density with $r^{\vec{\bullet}\circ\star} = \ell^{\star\circ\overleftarrow{\bullet}} = 0$ and different static interaction strengths y. Curves from top to bottom with value of y : $5, 2, 1.001, 0.5$. The dotted lines correspond to non-interacting RNAP.

for a sufficiently strong static repulsion given by the critical value

$$y_c = 2. \tag{50}$$

Above ρ^* the interaction factor decreases and boosting disappears at a critical density ρ_c given by the nonzero solution of the equation $q(\rho_c, y) = 1$. Thus for $y > y_c$ boosting occurs in the density range $0 < \rho < \rho_c$ as shown in figure 6. In the high density range $\rho_c < \rho \leqslant 1/l_{\rm rod}$ jamming takes over.

The same observation was made in [30] in the absence of backtracking and shows that the emergence of boosting arises by the same interplay between static interaction and kinetic interactions even when the interaction parameter $\ell^{\overleftarrow{\bullet}\star}$ for backtracking is as strong as the interaction parameter $r^{\star\vec{\bullet}}$ for translocation. Thus, in the minimal kinetic interaction range, only for sufficiently strong static repulsion $y > 2$ boosting appears and reaches global maximum at a density ρ^*, above which the velocity drops from the maximum to zero at the maximal rod density 0.2. Above the critical static repulsion strength boosting thus occurs for densities between 0 and a critical density that is close to the maximally possible density $1/l_{\rm rod}$. When y is not strong enough ($y \leqslant 2$), the velocity is less than the velocity of a single RNAP (dotted line). See figure 6(a) for these features and figure 6(b) that shows how the corresponding the average flux j varies with the density. The interaction factor q is given by the same curves as in figure 6(a), with rescaled y-axis where the dotted reference line for non-interaction RNAP at $y = 1$.

The interaction parameter $\ell^{\overleftarrow{\bullet}\star} = r^{\star\vec{\bullet}}$, related to both backtracking and translocation, however, has a significant impact on boosting as it is linked with the static interaction strength via the consistency relation $\ell^{\overleftarrow{\bullet}\star} = y - 1$ which shows that boosting arises if and only if $\ell^{\overleftarrow{\bullet}\star} = r^{\star\vec{\bullet}} > 1$. This means that for boosting to emerge it is not enough that the RNAP pushing on the level of individual RNAP, which corresponds to $\ell_c^{\overleftarrow{\bullet}\star} = r_c^{\star\vec{\bullet}} > 0$, exists. It has to be sufficiently strong and exceed the critical value determined by the interaction parameter $\ell_c^{\overleftarrow{\bullet}\star} = r_c^{\star\vec{\bullet}} = 1$. Moreover, when sufficiently strong it is the RNAP pushing itself that is important, not its direction. Hence backtracking has no significant impact on the rate of elongation in the scenario of minimal kinetic interaction range.

To go beyond the basic minimal interaction scheme we account in what follows for next-nearest neighbor interaction (extended kinetic interactions).

3.5.2. Extended kinetic interaction range

The picture changes somewhat for extended kinetic interaction range. Using the exact expressions (46) for the average velocity in an ensemble of interacting RNAP and (40) for a single rod as well as the consistency conditions (16) one finds that the interaction parameter is larger than 1 in a density range is given by the inequality

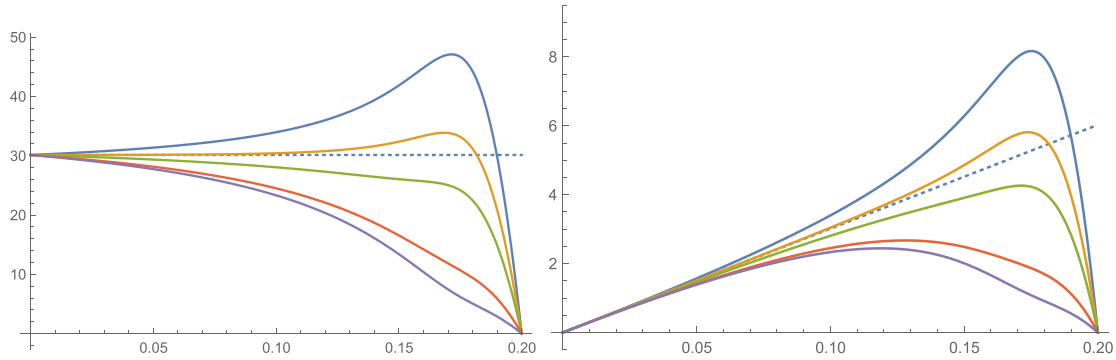
$$z\frac{1+(y-1)z+y}{(1+(y-1)z)^2} < \frac{r\tau_1 r^{\star\vec{\bullet}} - \ell\tau_2\ell^{\overleftarrow{\bullet}\star}}{r\tau_1 - \ell\tau_2} \tag{51}$$

relating density (parametrized by z) and the static interaction parameter y to the bare and nearest-neighbor backtracking and translocation rates. We illustrate this inequality for some scenarios.

3.5.2.1. Strong static repulsion

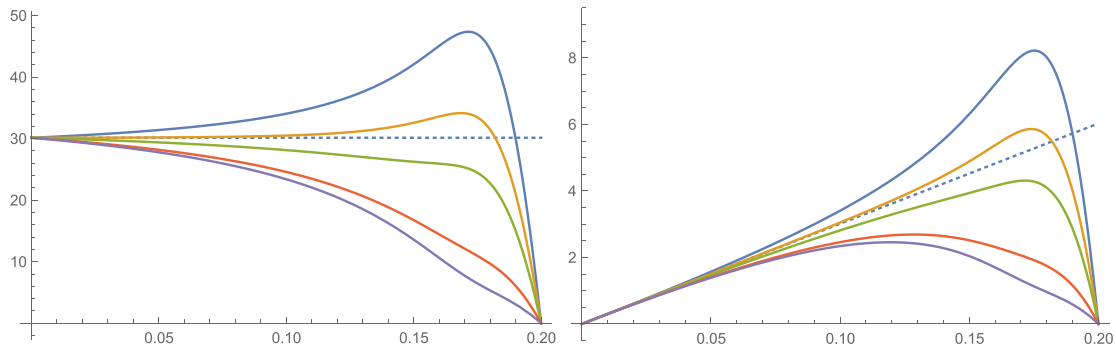
First we consider strong repulsion with $y = 5$ well above the critical value for which boosting occurs and explore two scenarios for backtracking.

(i) We take $\ell^{\star\circ\overleftarrow{\bullet}} = 0$ which means that the next-nearest upstream neighbor has no effect on the rate of RNAP backtracking (no blocking enhancement for backtracking RNAP). Static repulsion is then realized by $\ell^{\overleftarrow{\bullet}\star} = 4$, i.e. pushing to the left. In figures 7(a) and (b) it is shown how boosting changes as blocking enhancement for translocation is increased. When the blocking enhancement for translocation is too strong



(a) Average velocity v (b) Average flux j

Figure 7. RNAP velocity v and RNAP flux j as functions of RNAP density with static interaction strength $y = 5$ and kinetic interaction parameter $\ell^{\star\circ\overleftarrow{\bullet}} = 0$ for different values $r^{\vec{\bullet}\circ\star}$. Curves from top to bottom with value of $r^{\vec{\bullet}\circ\star}$: 0, −0.3, −0.5, −0.8, −0.9. The dotted lines correspond to non-interacting RNAP.



(a) Average velocity v (b) Average flux j

Figure 8. RNAP velocity v and RNAP flux j as functions of RNAP density with static interaction strength $y = 5$ and kinetic interaction parameter $\ell^{\star\circ\overleftarrow{\bullet}} = -0.9$ for different values $r^{\vec{\bullet}\circ\star}$. Curves from top to bottom with value of $r^{\vec{\bullet}\circ\star}$: 0, −0.3, −0.5, −0.8, −0.9. The dotted lines correspond to non-interacting RNAP.

($r^{\vec{\bullet}\circ\star}$ is close to -1), then even strong pushing (y arbitrarily large) does not lead to boosting.

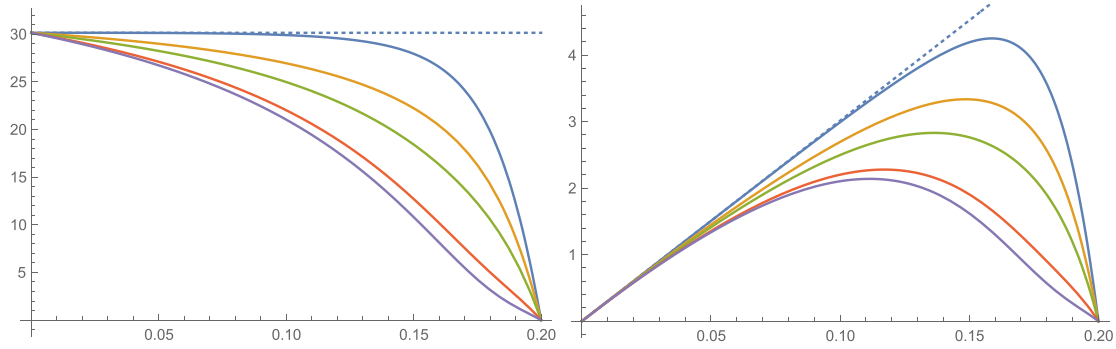
(ii) For $\ell^{\star\circ\overleftarrow{\bullet}} = -0.9$ the presence of a next-nearest upstream neighbor strongly suppresses backtracking. For the same choice interaction parameters for translocation and the same static repulsion strength. The curves shown in figures 8(a) and (b) for $\ell^{\star\circ\overleftarrow{\bullet}} = -0.9$ are nearly indistinguishable from the curves in figures 7(a) and 7(b) for $\ell^{\star\circ\overleftarrow{\bullet}} = 0$. Hence the occurrence of boosting phenomenon is insensitive to the choice of the phenomenological static parameter $\ell^{\star\circ\overleftarrow{\bullet}}$, indicating that suppression of boosting by sufficiently strong blocking enhancement and persistence of of boosting for low blocking enhancement is robust, also in the presence of backtracking.

3.5.2.2. Critical and weak static repulsion

We consider blocking enhancement both for backtracking and translocation with the same parameters as above but for critical static repulsion strength and extremely weak static repulsion, which can be realized by clinging as mechanism that reduces backtracking and translocation. As expected from the general discussion above there is no boosting. RNAP blocking enhancement leads to jamming for all densities (figures 9(a) and (b)). This jamming is stronger as the repulsion gets weaker, as demonstrated by the plots in figures 10(a) and (b).

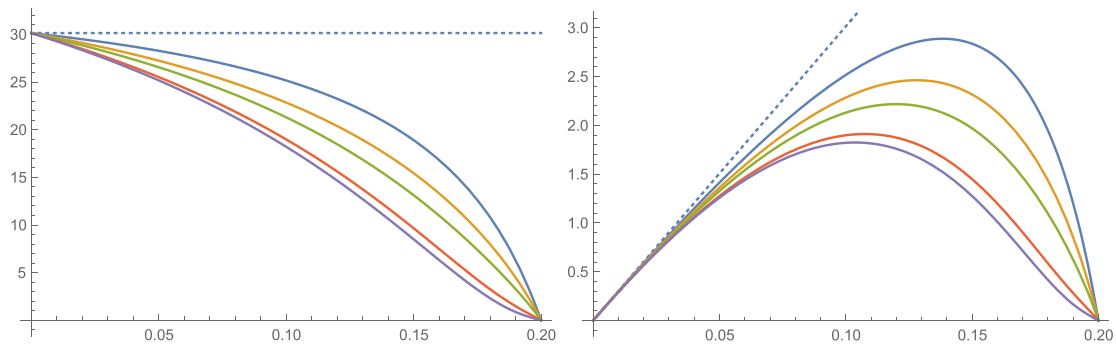
4. Summary and conclusions

In this work we have studied the effect of backtracking of RNAP on the average flux and velocity of RNAP along the DNA template during transcription elongation by considering the role of reverse reactions in the mechanochemical cycle that drives translocation. As starting point we have used the mathematically tractable model of [30] which has proven to be successful in understanding the role of interactions between RNAP in the emergence of cooperative pushing [27], called boosting in the present work. Boosting is a macroscopic phenomenon that is observed in biochemical experiments [23–26] and



(a) Average velocity v (b) Average flux j

Figure 9. RNAP velocity v and RNAP flux j as functions of RNAP density with static interaction strength $y = 2$ and kinetic interaction parameter $\ell^{\star\circ\overleftarrow{\bullet}} = -0.9$ for different values $r^{\overrightarrow{\bullet}\circ\star}$. Curves from top to bottom with value of $r^{\overrightarrow{\bullet}\circ\star}$: 0, −0.3, −0.5, −0.8, −0.9. The dotted lines correspond to non-interacting RNAP.



(a) Average velocity v (b) Average flux j

Figure 10. RNAP velocity v and RNAP flux j as functions of RNAP density with static interaction strength $y = 1.0001$ and kinetic interaction parameter $\ell^{\star\circ\overleftarrow{\bullet}} = -0.9$ for different values $r^{\overrightarrow{\bullet}\circ\star}$. Curves from top to bottom with value of $r^{\overrightarrow{\bullet}\circ\star}$: 0, −0.3, −0.5, −0.8, −0.9. The dotted lines correspond to non-interacting RNAP.

signifies an enhancement of the overall rate of transcription elongation through an increase of the average RNAP velocity that has its origin in pushes of stalled RNAP by trailing RNAP. It thus overcompensates jamming which arises from blocking the translocation of active RNAP by stalled RNAP and thus leads to a 'traffic jam' [5] that reduces the RNAP flux and thus the average RNAP velocity.

Significantly, as already noticed in [30], while simple steric excluded volume interaction between RNAP is enough to explain the emergence of jamming, the mere of existence of individual RNAP pushing, is *not* sufficient to explain boosting. Likewise, it was found in the present work that the presence of backtracking alone (which enhances the role of blocking) does not predict whether jamming takes over or whether boosting persists. Several key concepts are found to be crucial to understand how the interplay of microscopic forces that arise from interactions between individual RNAP moving along on the DNA template leads to the macroscopic collective phenomena of jamming and boosting.

The most important (and perhaps obvious) concept is the distinction between microscopic interactions between individual RNAP and the emergent collective phenomena that appear on experimental macroscopic scale. This is reflected in terminology adopted in the present paper: on the microscopic level we speak of blocking and pushing of RNAP, while the collective macroscopic counter parts are called jamming and boosting. With the latter term we deviate from the more standard notion of cooperative pushing which is what we mean by boosting, but which somewhat obscures the fundamental distinction between microscopic interaction between individual objects and collective outcome of this interaction.

The second most important (and perhaps less obvious) notion is the distinction on microscopic level between two kinds of interactions between individual RNAP, viz., (i) *static interactions* that determine the stationary distribution of the microscopically stochastic dynamics of translocation during transcription elongation, and (ii) *kinetic interactions* that

determine the rates with which the various microscopic processes in the mechanochemical cycle of RNAP translocation occur. These two kinds of interactions are conceptually different, but both physically and mathematically linked: A probability distribution for the microscopic dynamics determined by a given set of static interaction energies cannot be stationary for just any set of transition rates that encode the kinetic interactions. They have to be both physically and mathematically *consistent*. For our model this consistency is proved mathematically and allows us to make exact predictions within the framework of this model. In particular, it shows that (and how precisely) static repulsion or attraction correlate with kinetic blocking and pushing.

Finally, the third fundamental concept that (perhaps not surprisingly) plays a determining role are general characteristics of these interactions. The point in the present context is that the process of translocation of RNAP is permanently out of thermal equilibrium. Hence the interaction energy appearing in the stationary distribution has to be understood as a phenomenological *effective* energy. Therefore it cannot be derived from fundamental principles of Newtonian classical mechanics but needs to be postulated. When developing models it makes therefore makes sense to be guided on the one hand by empirical data (which are usually in short supply for the processes we have in mind) and by general theoretical notions such as interaction range (short-ranged or long-ranged), sign (attractive or repulsive), and strength. As empirical data are not readily available for quantitative predictions by the present simplified model we consider most parameters as variables and study how the quantities that we have computed change as these parameters are changed.

Taking these general insights as guide line, the main insight of the present work is that also in the presence of backtracking the strength of boosting, i.e. the phenomenon of cooperative pushing, is primarily determined by the strength of the effective static interaction between RNAP. As in the absence of backtracking, this static interaction needs to be repulsive and sufficiently strong, i.e. above a critical value that is determined by the interplay of the microscopic forces between two RNAP located at nearest neighbor or next-nearest sites. If pushing is strong enough then boosting occurs in a range of RNAP densities which is determined by the strength of blocking due to steric excluded volume interaction and blocking enhancement.

This conclusion is deduced from two observations. The first point to note is that backtracking arising from the reversed mechanochemical cycle appears in the rate of elongation in a direct reduction of the average speed of individual RNAP that is determined by the rate of backward pushing and which arises already in the absence of interaction. It is a straightforward consequence of the fact that occasional backtracking reduces the average speed of a single RNAP that mostly moves by forward translocation. Also in the interacting case this direct reduction of the average of an RNAP remains very small as it is shown to be proportional to the small bare rate of backtracking. Hence this effect has no significant bearing on whether or not boosting occurs.

The second and more subtle manifestation of backtracking in the rate of elongation is in the static interaction strength itself. The consistency relations show that this effect is linked to the rate of pushing and blocking enhancement in forward translocation and hence indepedent of the overall bare rate of backtracking. When pushing and blocking enhancement in forward translocation are sufficiently strong to cause boosting, then by consistency also blocking enhancement of backtracking RNAP is necessarily large and even consistent with a very short-ranged attraction that may cause clinging and pulling. Thus the strength of boosting is indepedent of whether backtracking takes place at all.

The present model is highly stylized and in order to examine which microscopic mechanims of backtracking affect boosting we have focussed on a specific one. An open question to be addressed in future theoretical work is the role of other modes of backtracking, in particular, backtracking directly from state 1 without PP_i bound which was considered in [32] but without taking into account pushing. A second open problem is the range of interaction which in the present work was taken to be at most next-nearest neighbor to allow for exact computations in closed form. Current work on exclusion processes without internal degree of freedom shows that the exact stationary distribution can be constructed also for interactions with longer range [60]. It is interesting to extend this approach to allow for mechanochemical cycles with 2 or more states. This will open up the possibility to adjust interaction parameters to less stylized interaction forces and to experimental data that may be expected from technological advances in the observation of motion of single RNAP.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

This work is financially supported by CAPES (Brazil) Finance Code 001, by CNPq (Brazil), Grant Number 140797/2018-1, by FAPESP (Brazil), Grant 2017/10555-0, and by FCT (Portugal) through CAMGSD, IST-ID, Projects UIDB/04459/2020, and UIDP/04459/2020, and by the FCT Grants 2020.03953.CEECIND and 2022.09232.PTDC.

N Ngoc gratefully acknowledges the financial support of CAPES and CNPq during his studies at the Doctorate Program in Statistics at the University of São Paulo.

Appendix. Consistency relations and proof of the theorem

We provide a way how to find conditions on parameters appearing in the rates (1)–(4) such that the invariant measure of the process is of the form (9). In order to do that we employ the method used in [30]. Namely, at equilibrium one can rewrite the master equation of the process in a local divergence by using a specific discrete form of Noether's theorem (57).

A.1. Stationary condition

Dividing (8) by the stationary distribution (9), the stationary condition becomes

$$\sum_{i=1}^{N}\Bigg[r_i\left(\boldsymbol{\eta}^i_{\mathrm{tlf}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{tlf}}\right)}{\pi\left(\boldsymbol{\eta}\right)}+\ell_i\left(\boldsymbol{\eta}^i_{\mathrm{tlb}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{tlb}}\right)}{\pi\left(\boldsymbol{\eta}\right)} + a_i\left(\boldsymbol{\eta}^i_{\mathrm{rel}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{rel}}\right)}{\pi\left(\boldsymbol{\eta}\right)}+d_i\left(\boldsymbol{\eta}^i_{\mathrm{bin}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{bin}}\right)}{\pi\left(\boldsymbol{\eta}\right)} - \left(r_i\left(\boldsymbol{\eta}\right)+\ell_i\left(\boldsymbol{\eta}\right)+a_i\left(\boldsymbol{\eta}\right)+d_i\left(\boldsymbol{\eta}\right)\right)\Bigg]=0. \quad (52)$$

Now we introduce the quantities

$$R_i\left(\boldsymbol{\eta}\right)=r_i\left(\boldsymbol{\eta}^i_{\mathrm{tlf}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{tlf}}\right)}{\pi\left(\boldsymbol{\eta}\right)}-r_i\left(\boldsymbol{\eta}\right), \quad (53)$$

$$L_i\left(\boldsymbol{\eta}\right)=\ell_i\left(\boldsymbol{\eta}^i_{\mathrm{tlb}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{tlb}}\right)}{\pi\left(\boldsymbol{\eta}\right)}-\ell_i\left(\boldsymbol{\eta}\right), \quad (54)$$

$$A_i\left(\boldsymbol{\eta}\right)=a_i\left(\boldsymbol{\eta}^i_{\mathrm{rel}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{rel}}\right)}{\pi\left(\boldsymbol{\eta}\right)}-a_i\left(\boldsymbol{\eta}\right), \quad (55)$$

$$D_i\left(\boldsymbol{\eta}\right)=d_i\left(\boldsymbol{\eta}^i_{\mathrm{bin}}\right)\frac{\pi\left(\boldsymbol{\eta}^i_{\mathrm{bin}}\right)}{\pi\left(\boldsymbol{\eta}\right)}-d_i\left(\boldsymbol{\eta}\right). \quad (56)$$

Taking into account periodicity, the stationarity condition (52) is satisfied if the lattice divergence condition

$$R_i\left(\boldsymbol{\eta}\right)+L_i\left(\boldsymbol{\eta}\right)+A_i\left(\boldsymbol{\eta}\right)+D_i\left(\boldsymbol{\eta}\right)=\Phi_i\left(\boldsymbol{\eta}\right)-\Phi_{i+1}\left(\boldsymbol{\eta}\right) \quad (57)$$

holds for all allowed configurations with a family of functions $\Phi_i(\boldsymbol{\eta})$ satisfying $\Phi_{N+1}(\boldsymbol{\eta})=\Phi_1(\boldsymbol{\eta})$. The lattice divergence condition can be understood as a specific discrete form of Noether's theorem.

A.2. Mapping to the headway process

Due to steric hard core repulsion, a translocation of the ith rod from k_i to k_i+1, corresponding to the transition $(m_{i-1},m_i)\to(m_{i-1}+1,m_i-1)$, takes place if $m_i>0$. Similarly, only if $m_{i-1}>0$ the backtracking corresponding to the transition $(m_{i-1},m_i)\to(m_{i-1}-1,m_i+1)$ can occur.

In terms of the new stochastic variables $\boldsymbol{\zeta}=(\mathbf{m},\boldsymbol{\alpha})$ given by the distance vector $\mathbf{m}$ and the state vector $\boldsymbol{\alpha}$ the transition rates (1)–(4) become

$$\tilde{r}_i\left(\boldsymbol{\zeta}\right)=r\delta_{\alpha_i,1}\left(1+r^{\star\vec{\bullet}}\theta^0_{i-1}+r^{\vec{\bullet}\circ\star}\theta^1_i\right)\left(1-\theta^0_i\right); \quad (58)$$

$$\tilde{\ell}_i\left(\boldsymbol{\zeta}\right)=\ell\delta_{\alpha_i,2}\left(1+\ell^{\star\circ\overleftarrow{\bullet}}\theta^1_{i-1}+\ell^{\overleftarrow{\bullet}\star}\theta^0_i\right)\left(1-\theta^0_{i-1}\right); \quad (59)$$

$$\begin{aligned}\tilde{a}_i\left(\boldsymbol{\zeta}\right)=a\delta_{\alpha_i,2}\big(&1+a^{\star\bullet}\theta^0_{i-1}+a^{\bullet\star}\theta^0_i+a^{\star\bullet\star}\theta^0_{i-1}\theta^0_i\\&+a^{\star\circ\bullet}\theta^1_{i-1}+a^{\bullet\circ\star}\theta^1_i\big);\end{aligned} \quad (60)$$

$$\begin{aligned}\tilde{d}_i\left(\boldsymbol{\zeta}\right)=d\delta_{\alpha_i,1}\big(&1+d^{\star\bullet}\theta^0_{i-1}+d^{\bullet\star}\theta^0_i+d^{\star\bullet\star}\theta^0_{i-1}\theta^0_i\\&+d^{\star\circ\bullet}\theta^1_{i-1}+d^{\bullet\circ\star}\theta^1_i\big).\end{aligned} \quad (61)$$

Before writing the master equation for the headway process, we introduce notation for the configuration that leads to a given configuration $\boldsymbol{\zeta}$. Namely, $\boldsymbol{\zeta}^{i-1,i},\boldsymbol{\zeta}^{i,i-1}$ correspond to translocation and backtracking respectively, and $\boldsymbol{\zeta}^{i,\mathrm{rel}},\boldsymbol{\zeta}^{i,\mathrm{bin}}$ correspond to PP_i release and binding respectively. Before introducing these configurations, we denote by (k,l) the pair $(i-1,i)$ or $(i,i-1)$ and by $\sharp$ the superscript rel or bin. Thus, the configurations $\boldsymbol{\zeta}^{i-1,i},\boldsymbol{\zeta}^{i,i-1},\boldsymbol{\zeta}^{i,\mathrm{rel}}$, and $\boldsymbol{\zeta}^{i,\mathrm{bin}}$ are defined by

$$m^{k,l}_j:=m_j+\delta_{j,l}-\delta_{j,k}\ \text{and}\ s^{k,l}_j:=\alpha_j+\left(3-2\alpha_j\right)\delta_{j,i}, \quad (62)$$

$$m^{i,\sharp}_j:=m_j\ \text{and}\ \alpha^{i,\sharp}_j:=\alpha_j+\left(3-2\alpha_j\right)\delta_{j,i}. \quad (63)$$

This yields the master equation

$$\frac{\mathrm{d}\mathbb{P}\left(\boldsymbol{\zeta},t\right)}{\mathrm{d}t}=\sum_{i=1}^{N}Q_i\left(\boldsymbol{\zeta},t\right) \quad (64)$$

with

$$\begin{aligned}Q_i\left(\boldsymbol{\zeta},t\right)=&\,\tilde{r}_i\left(\boldsymbol{\zeta}^{i-1,i}\right)\mathbb{P}\left(\boldsymbol{\zeta}^{i-1,i},t\right)-\tilde{r}_i\left(\boldsymbol{\zeta}\right)\mathbb{P}\left(\boldsymbol{\zeta},t\right)\\&+\tilde{\ell}_i\left(\boldsymbol{\zeta}^{i,i-1}\right)\mathbb{P}\left(\boldsymbol{\zeta}^{i,i-1},t\right)-\tilde{\ell}_i\left(\boldsymbol{\zeta}\right)\mathbb{P}\left(\boldsymbol{\zeta},t\right)\\&+\tilde{a}_i\left(\boldsymbol{\zeta}^{i,\mathrm{rel}}\right)\mathbb{P}\left(\boldsymbol{\zeta}^{i,\mathrm{rel}},t\right)-\tilde{a}_i\left(\boldsymbol{\zeta}\right)\mathbb{P}\left(\boldsymbol{\zeta},t\right)\\&+\tilde{d}_i\left(\boldsymbol{\zeta}^{i,\mathrm{bin}}\right)\mathbb{P}\left(\boldsymbol{\zeta}^{i,\mathrm{bin}},t\right)-\tilde{d}_i\left(\boldsymbol{\zeta}\right)\mathbb{P}\left(\boldsymbol{\zeta},t\right)\end{aligned} \quad (65)$$

where

$$\tilde{r}_i\left(\boldsymbol{\zeta}^{i-1,i}\right)=r\delta_{\alpha_i,2}\left(1+r^{\star\vec{\bullet}}\theta^1_{i-1}+r^{\vec{\bullet}\circ\star}\theta^0_i\right)\left(1-\theta^0_{i-1}\right), \quad (66)$$

$$\tilde{\ell}_i\left(\boldsymbol{\zeta}^{i,i-1}\right)=\ell\delta_{\alpha_i,1}\left(1+\ell^{\star\circ\overleftarrow{\bullet}}\theta^0_{i-1}+\ell^{\overleftarrow{\bullet}\star}\theta^1_i\right)\left(1-\theta^0_i\right), \quad (67)$$

$$\tilde{a}_i\left(\zeta^{i,\mathrm{rel}}\right) = a\delta_{\alpha_i,1}\left(1 + a^{\star\bullet}\theta^0_{i-1} + a^{\bullet\star}\theta^0_i + a^{\star\bullet\star}\theta^0_{i-1}\theta^0_i + a^{\star\circ\bullet}\theta^1_{i-1} + a^{\bullet\circ\star}\theta^1_i\right), \tag{68}$$

$$\tilde{d}_i\left(\zeta^{i,\mathrm{bin}}\right) = d\delta_{\alpha_i,2}\left(1 + d^{\star\bullet}\theta^0_{i-1} + d^{\bullet\star}\theta^0_i + d^{\star\bullet\star}\theta^0_{i-1}\theta^0_i + d^{\star\circ\bullet}\theta^1_{i-1} + d^{\bullet\circ\star}\theta^1_i\right). \tag{69}$$

Thanks to the discrete version of Noether theorem, one can rephrase the stationarity condition for the headway process in a local divergence form which is equivalent to (57). Let us first introduce the following notations

$$\tilde{R}_i(\zeta) = \tilde{r}_i\left(\zeta^{i-1,i}\right)\frac{\tilde{\pi}\left(\zeta^{i-1,i}\right)}{\tilde{\pi}(\zeta)} - \tilde{r}_i(\zeta), \tag{70}$$

$$\tilde{L}_i(\zeta) = \tilde{\ell}_i\left(\zeta^{i,i-1}\right)\frac{\tilde{\pi}\left(\zeta^{i,i-1}\right)}{\tilde{\pi}(\zeta)} - \tilde{\ell}_i(\zeta), \tag{71}$$

$$\tilde{A}_i(\zeta) = \tilde{a}_i\left(\zeta^{i,\mathrm{rel}}\right)\frac{\tilde{\pi}\left(\zeta^{i,\mathrm{rel}}\right)}{\tilde{\pi}(\zeta)} - \tilde{a}_i(\zeta), \tag{72}$$

$$\tilde{D}_i(\zeta) = \tilde{d}_i\left(\zeta^{i,\mathrm{bin}}\right)\frac{\tilde{\pi}\left(\zeta^{i,\mathrm{bin}}\right)}{\tilde{\pi}(\zeta)} - \tilde{d}_i(\zeta). \tag{73}$$

Again, we make use of (k,l) for $(i-1,i)$ or $(i,i-1)$ and $\sharp$ for the superscript rel or bin. Notice that

$$\theta^p_j\left(\zeta^{k,l}\right) = \delta_{m_j+\delta_{j,l}-\delta_{j,k},p} = \theta^{p-\delta_{j,l}+\delta_{j,k}}_j(\zeta) \text{ and } \delta_{\alpha_i^{k,l},\alpha} = \delta_{\alpha_i,3-\alpha}; \tag{74}$$

$$\theta^p_j\left(\zeta^{i,\sharp}\right) = \theta^p_j(\zeta) \text{ and } \delta_{\alpha_i^{i,\sharp},\alpha} = \delta_{\alpha_i,3-\alpha}, \tag{75}$$

so that one gets

$$\frac{\tilde{\pi}\left(\zeta^{i,\sharp}\right)}{\tilde{\pi}(\zeta)} = x^{3-2\alpha_i}, \tag{76}$$

$$\frac{\tilde{\pi}\left(\zeta^{k,l}\right)}{\tilde{\pi}(\zeta)} = x^{-3+2\alpha_i}y^{\theta^0_k-\theta^1_k+\theta^0_l}. \tag{77}$$

Hence,

$$\tilde{R}_i(\zeta) = x^{-1}y^{\theta^0_{i-1}+\theta^0_i-\theta^1_{i-1}}r\delta_{\alpha_i,2}\left(1 + r^{\star\vec{\bullet}}\theta^1_{i-1} + r^{\vec{\bullet}\circ\star}\theta^0_i\right)\left(1-\theta^0_{i-1}\right) - r\delta_{\alpha_i,1}\left(1 + r^{\star\vec{\bullet}}\theta^0_{i-1} + r^{\vec{\bullet}\circ\star}\theta^1_i\right)\left(1-\theta^0_i\right), \tag{78}$$

$$\tilde{L}_i(\zeta) = xy^{\theta^0_i-\theta^1_i+\theta^0_{i-1}}\ell\delta_{\alpha_i,1}\left(1 + \ell^{\star\circ\overleftarrow{\bullet}}\theta^0_{i-1} + \ell^{\overleftarrow{\bullet}\star}\theta^1_i\right)\left(1-\theta^0_i\right) - \ell\delta_{\alpha_i,2}\left(1 + \ell^{\star\circ\overleftarrow{\bullet}}\theta^1_{i-1} + \ell^{\overleftarrow{\bullet}\star}\theta^0_i\right)\left(1-\theta^0_{i-1}\right), \tag{79}$$

$$\tilde{A}_i(\zeta) = \left(x\delta_{\alpha_i,1} - \delta_{\alpha_i,2}\right)a\left(1 + a^{\star\bullet}\theta^0_{i-1} + a^{\bullet\star}\theta^0_i + a^{\star\bullet\star}\theta^0_{i-1}\theta^0_i + a^{\star\circ\bullet}\theta^1_{i-1} + a^{\bullet\circ\star}\theta^1_i\right), \tag{80}$$

$$\tilde{D}_i(\zeta) = \left(x^{-1}\delta_{\alpha_i,2} - \delta_{\alpha_i,1}\right)d\left(1 + d^{\star\bullet}\theta^0_{i-1} + d^{\bullet\star}\theta^0_i + d^{\star\bullet\star}\theta^0_{i-1}\theta^0_i + d^{\star\circ\bullet}\theta^1_{i-1} + d^{\bullet\circ\star}\theta^1_i\right). \tag{81}$$

One requires

$$\tilde{R}_i + \tilde{L}_i + \tilde{A}_i + \tilde{D}_i = \tilde{\Phi}_{i-1} - \tilde{\Phi}_i, \tag{82}$$

where $\tilde{\Phi}_i$ is of the form $\tilde{\Phi}_i = (e + f\theta^0_i + h\theta^1_i)(\delta_{\alpha_i,1} + \delta_{\alpha_i,2}) = e + f\theta^0_i + h\theta^1_i$. Notice that $\tilde{\Phi}_i$ must be of that form since $\tilde{R}_i, \tilde{L}_i, \tilde{A}_i, \tilde{D}_i$ depend on the state of rod i and variables $\theta^0_{i-1}, \theta^1_{i-1}, \theta^0_i, \theta^1_i$ belonging to $\{0,1\}$.

By considering all possible cases of (82), one first gets

$$f = \frac{-r\left(1 + r^{\star\vec{\bullet}}\right) + x\ell\left(1 + \ell^{\vec{\bullet}\star}\right)}{1+x} \tag{83}$$

$$h = \frac{rr^{\vec{\bullet}\circ\star} - x\ell\ell^{\star\circ\overleftarrow{\bullet}}}{1+x} \tag{84}$$

and then one gets the stationary conditions (15)–(21). However, we shall give a short proof of the above claim in the next subsection after knowing the results.

A.3. Proof of theorem 2.1

As previously noted, it is necessary to account for all instances of (82) in order to identify the constraints. Yet, this task may seem tedious in ensuring the accuracy of the results. In this context, we aim to present a concise proof of theorem 2.1. Consequently, we will demonstrate that given conditions (15)–(21), the process's invariant measure adheres to the structure described in (22). Our objective is achieved by confirming that (82) holds for all configurations.

- If $m_{i-1}, m_i > 1$, one has $\tilde{R}_i = x^{-1}r\delta_{\alpha_i,2} - r\delta_{\alpha_i,1}, \tilde{L}_i = x\ell\delta_{\alpha_i,1} - \ell\delta_{\alpha_i,2}, \tilde{A}_i = (x\delta_{\alpha_i,1} - \delta_{\alpha_i,2})a$, $\tilde{D}_i = (x^{-1}\delta_{\alpha_i,2} - \delta_{\alpha_i,1})d$. Notice that in this case, the left-hand side of (82) is 0. If the state of rod is 1 meaning that $\alpha_i = 1$, one has $\tilde{R}_i + \tilde{L}_i + \tilde{A}_i + \tilde{D}_i = -r + x\ell + xa - d$ which is 0 due to (15). Similarly, for the case $\alpha_i = 2$, (82) holds.
- If $m_{i-1} > 1, m_i = 1$, one has $\tilde{R}_i = x^{-1}r\delta_{\alpha_i,2} - r\delta_{\alpha_i,1}(1 + r^{\vec{\bullet}\circ\star}), \tilde{L}_i = xy^{-1}\ell\delta_{\alpha_i,1}(1 + \ell^{\overleftarrow{\bullet}\star}) - \ell\delta_{\alpha_i,2}$, $\tilde{A}_i = (x\delta_{\alpha_i,1} - \delta_{\alpha_i,2})a(1 + a^{\bullet\circ\star})$, $\tilde{D}_i =$

$(x^{-1}\delta_{\alpha_i,2} - \delta_{\alpha_i,1})d(1+d^{\bullet\circ\star})$. Notice that in this case the left-hand side of (82) is $-h$. If $\alpha_i = 1$, one has $\tilde{R}_i + \tilde{L}_i + \tilde{A}_i + \tilde{D}_i = -r(1+\vec{r^{\bullet}}^{\circ\star}) + xy^{-1}\ell(1+\overleftarrow{\ell^{\bullet}}^{\star}) + xa(1+a^{\bullet\circ\star}) - d(1+d^{\bullet\circ\star})$. It is easy to check from (15), (16), and (21) that $\tilde{R}_i + \tilde{L}_i + \tilde{A}_i + \tilde{D}_i = -h$ where h is defined in (84). Thus, (82) holds for this case. Similarly, for the case $\alpha_i = 2$, (82) holds as well.

- For the rest of the cases: $m_{i-1} > 1, m_i = 0$; $m_{i-1} = 1, m_i > 1$; $m_{i-1} = 0, m_i > 1$; $m_{i-1} = 1, m_i = 1$; $m_{i-1} = 1, m_i = 0$; $m_{i-1} = 0, m_i = 1$; $m_{i-1} = 0, m_i = 0$, one considers similarly to show that (82) holds.

The proof is complete.

ORCID iDs

Ngo P N Ngoc https://orcid.org/0000-0002-5261-726X
Vladimir Belitsky https://orcid.org/0000-0003-2231-6877
Gunter M Schütz https://orcid.org/0000-0002-1050-6617

References

[1] Alberts B, Bray D, Hopkin K, Johnson A, Lewis J, Roberts M K and Walter P 2013 *Essential Cell Biology* 4th edn (Garland Science)
[2] Bai L, Santangelo T J and Wang M D 2006 Single-molecule analysis of RNA polymerase transcription *Annu. Rev. Biophys. Biomol. Struct.* **35** 343
[3] Wang H Y, Elston T, Mogilner A and Oster G 1998 Force generation in RNA polymerase *Biophys. J.* **74** 1186–202
[4] Chowdhury D 2013 Stochastic mechano-chemical kinetics of molecular motors: a multidisciplinary enterprise from a physicist's perspective *Phys. Rep.* **529** 1–197
[5] Schadschneider A, Chowdhury D and Nishinari K 2010 *Stochastic Transport in Complex Systems* (Elsevier)
[6] Lipowsky R, Beeg J, Dimova R, Klumpp S and Müller M J I 2010 Cooperative behavior of molecular motors: cargo transport and traffic phenomena *Physica* E **42** 649–61
[7] Chou T, Mallick K and Zia R K P 2011 Non-equilibrium statistical mechanics: from a paradigmatic model to biological transport *Rep. Prog. Phys.* **74** 116601
[8] Kolomeisky A B 2015 *Motor Proteins and Molecular Motors* (CRC Press)
[9] Graf I R and Frey E 2017 Generic transport mechanisms for molecular traffic in cellular protrusions *Phys. Rev. Lett.* **118** 128101
[10] Fang X, Kruse K, Lu T and Wang J 2019 Nonequilibrium physics in biology *Rev. Mod. Phys.* **91** 045004
[11] Cavallaro M, Wang Y, Hebenstreit D and Dutta R 2023 Bayesian inference of polymerase dynamics over the exclusion process *R. Soc. Open Sci.* **10** 221469
[12] Derrida B 1998 An exactly soluble non-equilibrium system: the asymmetric simple exclusion process *Phys. Rep.* **301** 65–83
[13] Schütz G M 2001 Exactly solvable models for many-body systems far from equilibrium *Phase Transitions and Critical Phenomena* vol 19 (Academic)
[14] Liggett T M 2010 *Continuous Time Markov Processes: An Introduction* (American Mathematical Society)
[15] MacDonald J T, Gibbs J H and Pipkin A C 1968 Kinetics of biopolymerization on nucleic acid templates *Biopolymers* **6** 1–25
[16] Schütz G M 1997 The Heisenberg chain as a dynamical model for protein synthesis—some theoretical and experimental results *Int. J. Mod. Phys.* B **11** 197–202
[17] Ferrari P A, Kipnis C and Saada E 1991 Microscopic structure of travelling waves in the asymmetric simple exclusion process *Ann. Probab.* **19** 226–44
[18] Derrida B, Janowsky S A, Lebowitz J L and Speer E R 1993 Exact solution of the totally asymmetric simple exclusion process: Shock profiles *J. Stat. Phys.* **73** 813–42
[19] Ferrari P A and Fontes L R G 1994 Shock fluctuations in the asymmetric simple exclusion process *Probab. Theory Relat. Fields* **99** 305–19
[20] Dudziński M and Schütz G M 2000 Relaxation spectrum of the asymmetric exclusion process with open boundaries *J. Phys. A: Math. Gen.* **33** 8351–64
[21] Belitsky V and Schütz G M 2002 Diffusion and scattering of shocks in the partially asymmetric simple exclusion process *Electron. J. Probab.* **7** 1–21
[22] de Gier J and Essler F H L 2006 Exact spectral gaps of the asymmetric exclusion process with open boundaries *J. Stat. Mech.* P12011
[23] Epshtein V and Nudler E 2003 Cooperation between RNA polymerase molecules in transcription elongation *Science* **300** 801–5
[24] Epshtein V, Toulme F, Rahmouni A R, Borukhov S and Nudler E 2003 Transcription through the roadblocks: the role of RNA polymerase cooperation *SEMBO J.* **22** 4719
[25] Jin J, Bai L, Johnson D S, Fulbright R M, Kireeva M L, Kashlev M and Wang M D 2010 Synergistic action of RNA polymerases in overcoming the nucleosomal barrier *Nat. Struct. Mol. Biol.* **17** 745–52
[26] Saeki H and Svejstrup J Q 2009 Stability, flexibility and dynamic interactions of colliding RNA polymerase II elongation complexes *Mol. Cell* **35** 191–205
[27] Galburt E A, Parrondo J M R and Grill S W 2011 RNA polymerase pushing *Biophys. Chem.* **157** 43–47
[28] Stevenson-Jones F, Woodgate J, Castro-Roa D and Zenkin N 2020 Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest *Proc. Natl Acad. Sci. USA* **117** 8462–7
[29] Wang L 2024 RNA polymerase collisions and their role in transcription *Transcription* **15** 38–47
[30] Belitsky V and Schütz G M 2019 RNA Polymerase interactions and elongation rate *J. Theor. Biol.* **462** 370–80
[31] Belitsky V and Schütz G M 2019 Stationary RNA polymerase fluctuations during transcription elongation *Phys. Rev.* E **99** 012405
[32] Tripathi T and Chowdhury D 2008 Interacting RNA polymerase motors on DNA track: effects of traffic congestion and intrinsic noise on RNA synthesis *Phys. Rev.* E **77** 011921
[33] Shaevitz J W, Abbondanzieri E A, Landick R and Block S M 2003 Backtracking by single RNA polymerase molecules observed at near-base-pair resolution *Nature* **426** 684–7
[34] Galburt E, Grill S W, Wiedmann A, Lubhowska L, Choy J, Nogales E, Kashlev M and Bustamante C 2007 Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner *Nature* **446** 820–3
[35] Sahoo M and Klumpp S 2013 Backtracking dynamics of RNA polymerase: pausing and error correction *J. Phys.: Condens. Matter* **25** 374104
[36] Gout J F, Thomas W K, Smith Z, Okamoto K and Lynch M 2013 Large-scale detection of *in vivo* transcription errors *Proc. Natl Acad. Sci. USA* **110** 18584–9
[37] Gout J F *et al* 2017 The landscape of transcription errors in eukaryotic cells *Sci. Adv.* **3** e1701484
[38] Zuo X and Chou T 2022 Density- and elongation velocity-dependent error correction in RNA polymerization *Phys. Biol.* **19** 026001

[39] Schütz G and Domany E 1993 Phase transitions in an exactly soluble one- dimensional asymmetric exclusion model *J. Stat. Phys.* **72** 277–96
[40] Derrida B, Evans M R, Hakim V and Pasquier V 1993 Exact solution of a 1D asymmetric exclusion model using a matrix formulation *J. Phys. A: Math. Gen.* **26** 1493–517
[41] Schütz G M 1995 Diffusion-annihilation in the presence of a driving field *J. Phys. A: Math. Gen.* **28** 3405–15
[42] Parmeggiani A, Franosch T and Frey E 2003 Phase coexistence in driven one-dimensional transport *Phys. Rev. Lett.* **90** 086601
[43] Popkov V, Rákos A, Willmann R D, Kolomeisky A B and Schütz G M 2003 Localization of shocks in driven diffusive systems without particle number conservation *Phys. Rev.* E **67** 066117
[44] Gomes L, Midha T, Kumar Gupta A and Kolomeisky A B 2019 The effect of local dissociation on dynamics of interacting molecular motors *J. Phys. A: Math. Theor.* **52** 365001
[45] Chai Y, Lipowsky R and Klumpp S 2009 Transport by molecular motors in the presence of static defects *J. Stat. Phys.* **135** 241–60
[46] Tholstrup J, Oddershede L B and Sørensen M A 2012 mRNA pseudoknot structures can act as ribosomal roadblocks *Nucl. Acids Res.* **40** 303–13
[47] Appert-Rolland C, Ebbinghaus M and Santen L 2015 Intracellular transport driven by cytoskeletal motors: general mechanisms and defects *Phys. Rep.* **593** 1–59
[48] Erdmann-Pham D D, Duc K D and Song Y S 2020 The key parameters that govern translation efficiency *Cell Syst.* **10** 183–92
[49] Schütz G M 2021 Defect-induced anticorrelations in molecular motor traffic *J. Phys. A: Math. Theor.* **54** 255601
[50] Chatterjee P, Goldenfeld N and Kim S 2021 DNA supercoiling drives a transition between collective modes of gene synthesis *Phys. Rev. Lett.* **127** 218101
[51] Karevski D and Schütz G M 2017 Conformal invariance in driven diffusive systems at high currents *Phys. Rev. Lett.* **118** 030601
[52] Lakatos G and Chou T 2003 Totally asymmetric exclusion process with extended objects: a model for protein synthesis *J. Phys. A: Math. Gen.* **36** 2027–41
[53] Shaw L B, Zia R K P and Lee K H 2003 Totally asymmetric exclusion process with extended objects: a model for protein synthesis *Phys. Rev.* E **68** 021910
[54] Schönherr G and Schütz G M 2004 Exclusion process for particles of arbitrary extension: hydrodynamic limit and algebraic properties *J. Phys. A: Math. Gen.* **37** 8215–31
[55] Gupta S, Barma M, Basu U and Mohanty P K 2011 Driven *k*-mers: correlations in space and time *Phys. Rev.* E **84** 041102
[56] Dao Duc K, Saleem Z H and Song Y S 2018 Theoretical analysis of the distribution of isolated particles in totally asymmetric exclusion processes: application to mRNA translation rate estimation *Phys. Rev.* E **97** 012106
[57] Carlon E, Orlandini E and Stella A L 2002 Roles of stiffness and excluded volume in DNA denaturation *Phys. Rev. Lett.* **88** 198101
[58] Bar A, Kafri Y and Mukamel D 2009 Dynamics of DNA melting *J. Phys.: Condens. Matter* **21** 034110
[59] Hirschberg O, Mukamel D and Schütz G M 2011 Approach to equilibrium of diffusion in a logarithmic potential *Phys. Rev.* E **84** 041111
[60] Belitsky V, Ngoc N P N and Schütz G M 2024 Asymmetric exclusion process with long-range interactions (arXiv:2409.05017)