

Power series solution of the inhomogeneous exclusion process

Juraj Szavits-Nossan*

SUPA, School of Physics and Astronomy, University of Edinburgh, Peter Guthrie Tait Road, Edinburgh EH9 3FD, United Kingdom

M. Carmen Romano

SUPA, Institute for Complex Systems and Mathematical Biology, Department of Physics, Aberdeen AB24 3UE, United Kingdom and Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB24 3FX, United Kingdom

Luca Ciandrini

DIMNP, Université de Montpellier, CNRS, Montpellier, France and L2C, Université de Montpellier, CNRS, Montpellier, France

(Received 2 March 2018; published 29 May 2018)

We develop a power series method for the nonequilibrium steady state of the inhomogeneous one-dimensional totally asymmetric simple exclusion process (TASEP) in contact with two particle reservoirs and with site-dependent hopping rates in the bulk. The power series is performed in the entrance or exit rates governing particle exchange with the reservoirs, and the corresponding particle current is computed analytically up to the cubic term in the entry or exit rate, respectively. We also show how to compute higher-order terms using combinatorial objects known as Young tableaux. Our results address the long outstanding problem of finding the exact nonequilibrium steady state of the inhomogeneous TASEP. The findings are particularly relevant to the modeling of mRNA translation in which the rate of translation initiation, corresponding to the entrance rate in the TASEP, is typically small.

DOI: [10.1103/PhysRevE.97.052139](https://doi.org/10.1103/PhysRevE.97.052139)**I. INTRODUCTION**

The exclusion process is a stochastic driven lattice gas of particles interacting by excluded volume. It appears as a model for diverse transport phenomena including mRNA translation [1], enzyme kinetics [2], molecular motors [3,4], and vehicle traffic [5]. Here we are interested in an exclusion process in which particles move unidirectionally along a discrete lattice in contact with two particle reservoirs. This model is called the totally asymmetric simple exclusion process (TASEP), whose distinctive feature is a nonequilibrium steady state that carries a net current of particles.

A fundamental difference between equilibrium and nonequilibrium steady states is that the former are independent of the dynamics (provided the detailed balance is respected) while the latter are not. As a result, nonequilibrium steady states are in general unknown, even in one-dimensional systems [6]. An important exception is the TASEP in which particles move along the lattice at constant rate, for which the steady state is known explicitly [2,7,8]. The exact solution reveals intriguing boundary-induced phase transitions [9] that have no equilibrium counterpart.

However, the steady state of the *inhomogeneous* TASEP, in which particles move at position-dependent rates (also called disordered TASEP or TASEP with sitewise disorder), is a long outstanding unsolved problem in nonequilibrium statistical physics. The majority of work on the inhomoge-

neous TASEP concerns the phenomenon of phase separation induced by isolated inhomogeneities [10–15] and full disorder [16–20], which is typically studied using numerical simulations and various types of mean-field approximations that neglect correlations between neighboring sites [1,10,13,20,21]. Importantly, there is a long-standing interest in the inhomogeneous TASEP as a model for mRNA translation, in which ribosomes progress along the mRNA at codon-dependent hopping rates [1,20,22–28].

Motivated by this lack of exact results, one of us recently developed a power series method for the steady state of the inhomogeneous TASEP with binary disorder in which site-dependent hopping rates are either $r < 1$ (“slow” sites) or 1 (“normal” sites) [29]. The power series was performed up to the quadratic order in the variable r for lattices with one and two slow sites. More recently, we extended this method to a TASEP-like model of mRNA translation that includes codon-dependent elongation rates, two-step hopping mechanism, and extended particles that cover 10 lattice sites [30]. The power series was performed up to the quadratic order in the rate of translation initiation α , which is typically rate limiting for translation under physiological conditions [27,31,32] and, therefore, it crucially reduces contact between particles [33].

In this paper we develop a power series method for the steady state of the inhomogeneous TASEP with open boundary conditions and site-dependent hopping rates in the bulk. We perform a power series in the entrance and exit rates α and β and find an expression for the current up to the cubic order. In addition, we show how to compute higher-order terms using combinatorial objects called Young tableaux of

*jszavits@staffmail.ed.ac.uk

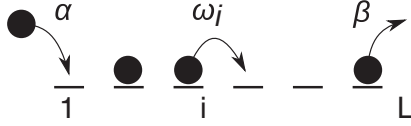


FIG. 1. Schematic of the TASEP with entrance rate α , site-dependent hopping rates ω_i for $i = 1, \dots, L-1$, and exit rate β .

shifted shape [34]. Interestingly, the connection between the inhomogeneous TASEP and Young tableaux that we establish here is different from the one in Ref. [35], which applies only to the homogeneous TASEP. Our method is robust and applicable to many other TASEP-like models, especially ones that describe mRNA translation in which the entrance rate is typically small.

The paper is organized as follows. The model is described in Sec. II A and its exact steady state is formally derived in Sec. II B. The main idea behind the power series method is summarized in Sec. III A. Sections III B and III C describe the power series in α and β , respectively. Main results are summarized in Sec. IV.

II. INHOMOGENEOUS EXCLUSION PROCESS

A. Model

We consider the totally asymmetric simple exclusion process on a one-dimensional lattice consisting of L sites, which is presented in Fig. 1. The lattice is in contact with two particle reservoirs that allow particles to enter at site 1 at rate α (provided the site 1 is empty) and exit from site L at rate β . Particles move unidirectionally along the lattice at site dependent rates ω_i , provided that the site $i+1$ in front is not occupied by another particle.

Each lattice site i is assigned an occupancy variable $\tau_i = 0$ if the site is empty and $\tau_i = 1$ if it is occupied by a particle. Throughout the text we consider continuous-time dynamics with entrance rate α , site-dependent hopping rates ω_i , $i = 1, \dots, L-1$, and exit rate β . The allowed transitions between configurations and their corresponding rates can be summarized as

$$0_1 \xrightarrow{\alpha} 1_1, \quad (1a)$$

$$1_i 0_{i+1} \xrightarrow{\omega_i} 0_i 1_{i+1}, \quad i = 1, \dots, L-1, \quad (1b)$$

$$1_L \xrightarrow{\beta} 0_L, \quad (1c)$$

where with 0_i (1_i) we represent the empty (occupied) site i .

The steady state of this process is determined by the following stationary master equation:

$$\alpha(1 - 2\tau_1)P(0_1\tau_2\dots) - \beta(1 - 2\tau_L)P(\dots\tau_{L-1}1_L) - \sum_{i=1}^{L-1} \omega_i(\tau_i - \tau_{i+1})P(\dots\tau_{i-1}1_i0_{i+1}\tau_{i+2}\dots) = 0, \quad (2)$$

where $P(C)$ is the probability to find the system in the configuration $C = \tau_1 \dots \tau_L$. The aim of this paper is to find $P(C)$ of an inhomogeneous exclusion process, which we then use to calculate the average particle current J , the local density

ρ_i , and the total density ρ , which are defined as

$$J = \alpha \sum_C [1 - \tau_1(C)]P(C), \quad (3)$$

$$\rho_i = \sum_C \tau_i(C)P(C), \quad \rho = \frac{1}{L} \sum_{i=1}^L \rho_i. \quad (4)$$

B. Exact solution

In this section we present an exact solution for any stationary ergodic master equation, which serves as the basis for our power series method presented in Sec. III. This solution is often overlooked in literature, because it is rarely practical for the reasons that we expose below. In that context, our power series method shows how to approximate $P(C)$ in a certain limit and extract information on the system described.

To this end, let us label configurations C_i , where i runs from 1 to $\mathcal{N} = 2^L$, which is the total number of configurations. This allows us to rewrite the master equation (2) in a more compact form,

$$M\mathbf{P} = \mathbf{0}, \quad (5)$$

where \mathbf{P} is a column vector of \mathcal{N} steady-state probabilities $P(C_i)$ and M is a $\mathcal{N} \times \mathcal{N}$ matrix whose elements M_{ij} are given by

$$M_{ij} = \begin{cases} W(C_j \rightarrow C_i), & i \neq j, \\ -\sum_{k \neq i} W(C_i \rightarrow C_k), & i = j. \end{cases} \quad (6)$$

Here $W(C_i \rightarrow C_j)$ is the transition rate from C_i to C_j that takes one of the values in Eqs. (1a)–(1c).

There are two general methods for solving Eq. (5) (and any other stationary ergodic master equation for that matter), one that uses determinants and the other that uses graphs; the latter is known as Schnakenberg's network theory [36] in physics or the matrix-tree theorem [37] in mathematics.

1. Solution using determinants

We recall that the sum of all elements in each column of M is zero, which means that the rows of M are linearly dependent and the determinant of M is zero. The Laplace expansion for the determinant $\det M$ gives

$$0 = \det M = \sum_{i=1}^{\mathcal{N}} M_{ij} A_{i,j}, \quad (7)$$

where $A_{i,j} = (-1)^{i+j} \det M^{(i,j)}$ is called i, j cofactor of M and $M^{(i,j)}$ is a matrix derived from M by removing the i th row and j th column. Equation (7) can be rewritten as

$$\begin{aligned} 0 = \det M &= \sum_{i \neq j} M_{ij} A_{i,j} + M_{jj} A_{j,j} \\ &= \sum_{i \neq j} (A_{i,j} - A_{j,j}) M_{ij}, \end{aligned} \quad (8)$$

which means that $A_{i,j} = A_{j,j}$. In the last passage of Eq. (8) we used the definition of $M_{jj} = -\sum_{k \neq j} M_{kj}$. Inserting this result back into Eq. (7) and comparing it to Eq. (5) shows that

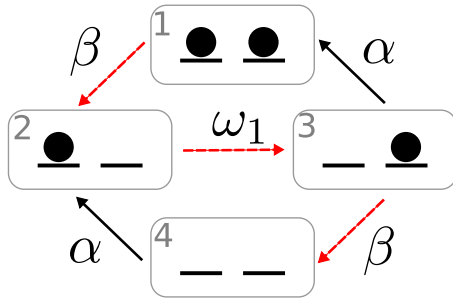


FIG. 2. Sketch of the graph connecting the configurations C_i in a lattice of length $L = 2$. The boxes represent all the possible configurations of the system, labeled by their index i and connected by weighted edges. In dashed red we highlight the only spanning tree T_4 rooted at C_4 (bottom configuration). By following Eq. (10), the probability of this configuration is proportional to $\beta^2\omega_1$.

$P(C_i)$ is given by

$$P(C_i) = \frac{\det M^{(i,i)}}{\sum_{j=1}^{\mathcal{N}} \det M^{(j,j)}}. \tag{9}$$

Alternatively, we can solve Eq. (5) using Cramer’s rule. Since the matrix M is singular ($\det M = 0$), we can replace one of the equations in Eq. (5) by the condition that $\sum_{i=1}^{\mathcal{N}} P(C_i) = 1$. This new system of equations has a nonsingular matrix, which can now be solved using Cramer’s rule.

2. Solution using Schnakenberg’s network theory

This method was first developed by Gustav Kirchhoff for electric circuits and was later refined by many others including the seminal Schnakenberg’s paper [36]. The reason we mention this method is because it will prove useful for our later analysis.

For a given Markov jump process, let us call V the set of all configurations C_i , $i = 1, \dots, \mathcal{N}$, and E the set of all pairs of configurations (C_i, C_j) for which $W(C_i \rightarrow C_j) \neq 0$. Schnakenberg’s network theory connects this stochastic process to the directed weighted graph $G = (V, E)$ consisting of vertices V and directed edges E weighted by the corresponding transition rates W .

Let us define a directed path in G as a sequence of vertices $C_{s(1)}, \dots, C_{s(n)}$ such that $(C_{s(i)}, C_{s(i+1)}) \in E$ for all $i = 1, \dots, n - 1$.¹ We define a spanning tree T_i rooted at C_i (also called spanning in tree with a sink at C_i) as a subgraph of G that contains all vertices in V and only a subset of edges in E such that there is exactly one directed path from every vertex $C_j \in V$ to C_i . An example of a spanning tree for a lattice of length $L = 2$ is shown in Fig. 2. Let us denote by $w(T_i)$ the weight of this spanning tree obtained by multiplying transition rates of all the edges in T_i . The following result, which is called the matrix-tree theorem, says that the minor $\det(-M^{(i,i)})^2$ is

equal to the sum of weights $w(T_i)$ of all spanning trees rooted at C_i

$$\det(-M^{(i,i)}) = (-1)^{\mathcal{N}-1} \det M^{(i,i)} = \sum_{T_i} w(T_i). \tag{10}$$

Inserting Eq. (10) into Eq. (9) gives

$$P(C_i) = \frac{\sum_{T_i} w(T_i)}{\sum_{j=1}^{\mathcal{N}} \sum_{T_j} w(T_j)}. \tag{11}$$

The main problem with Eqs. (9) and (11) is that the number of terms increases exponentially with L , which makes finding $P(C_i)$ intractable.

In the rest of this paper we develop a power series method that allows us to find $P(C_i)$ when either α or β are small, without solving the full system in Eq. (5).

III. POWER SERIES OF THE STEADY STATE

A. Main method

The main method exploits the fact that $P(C)$ is a quotient of two multivariate polynomials. We can choose any rate, which we denote by $\kappa \in \{\alpha, \omega_1, \dots, \omega_{L-1}, \beta\}$ and rewrite Eq. (9) as

$$P(C) = \frac{f(C)}{\sum_{C'} f(C')}, \quad f(C) = \sum_{n=0}^{K(C)} f_n(C) \kappa^n. \tag{12}$$

The unknown coefficients $f_n(C)$, $n = 0, \dots, K(C)$, depend on the configuration C and also on all the other rates, which we have omitted in order to simplify the notation.

According to the Schnakenberg network theory, $\sum_{n=0}^{K(C)} f_n(C) \kappa^n$ is equal to the sum of weights $w(T_C)$ of all spanning trees T_C rooted at C that have exactly n directed edges weighted by κ . The degree $K(C)$ is the maximum number of these edges over all spanning trees rooted at C . While in general we do not know the exact value of $K(C)$, we know that $K(C)$ is bounded from above by the maximum number of configurations (excluding C) that have an outward edge weighted by κ . For example, if we choose $\kappa = \alpha$, then $K(C) \leq 2^{L-1}$ if $\tau_1(C) = 1$ and $K(C) \leq 2^{L-1} - 1$ if $\tau_1(C) = 0$.

In order to understand what Eq. (12) implies for the coefficients $f_n(C)$, we write the master equation (2) in a generic form,

$$\sum_{C'} W(C \rightarrow C') P(C) = \sum_{C' \neq C} W(C' \rightarrow C) P(C'), \tag{13}$$

where we denote by $e(C)$ the exit rate from C , i.e.,

$$e(C) = \sum_{C'} W(C \rightarrow C'). \tag{14}$$

For any two configurations C and C' , we define $I_{C,C'}$ such that $I(C, C') = 1$ if there is a transition from C to C' at rate κ and $I_{C,C'} = 0$ otherwise,

$$I_{C,C'} = \begin{cases} 1, & W(C \rightarrow C') = \kappa, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

¹We used the index $s(i)$ for the i th configuration in the sequence S in order to remove any confusion with the indices $i = 1, \dots, \mathcal{N}$ that we introduced earlier.

²The minus sign is here because of our definition of the matrix M , which is natural for stochastic processes; in the graph theory one usually works with $-M$ instead.

We can now use $I_{C,C'}$ to decompose $W(C \rightarrow C')$ and $e(C)$ into

$$W(C \rightarrow C') = W(C \rightarrow C')(1 - I_{C,C'}) + \kappa I_{C,C'}, \quad (16)$$

$$e(C) = e_0(C) + \kappa \sum_{C'} I_{C,C'}, \quad (17)$$

$$e_0(C) = \sum_{C' \neq C} W(C \rightarrow C')(1 - I_{C,C'}). \quad (18)$$

Inserting Eqs. (12), (16), (17), and (18) into Eq. (13) and collecting all the terms of order κ^n gives the following recurrence relation for $f_n(C)$ in terms of $f_{n-1}(C')$:

$$\begin{aligned} e_0(C)f_n(C) + \sum_{C'} I_{C,C'} f_{n-1}(C) &= \sum_{C'} I_{C',C} f_{n-1}(C') \\ &+ \sum_{C'} W(C' \rightarrow C)(1 - I_{C',C}) f_n(C'). \end{aligned} \quad (19)$$

The equation above is true for $n > 0$. The terms containing $f_{n-1}(C)$ and $f_{n-1}(C')$ are absent for $n = 0$,

$$e_0(C)f_0(C) = \sum_{C'} W(C' \rightarrow C)(1 - I_{C',C})f_0(C'). \quad (20)$$

By construction, the configurations whose all outward edges are weighted by κ [so that $e_0(C) = 0$] are completely absent from Eq. (20). This equation thus has a trivial solution $f_0(C) = 0$, from which we conclude that $f_0(C) \neq 0$ only for configurations for which $e_0(C) = 0$. This conclusion is crucial for our later analysis.

Once we solve Eq. (19) recursively up to some order n we can Taylor expand $P(C)$ around $\kappa = 0$:

$$P(C) = \frac{\sum_{n=0}^{K(C)} f_n(C)\kappa^n}{\sum_{C'} \sum_{n=0}^{K(C')} f_n(C')\kappa^n} = \sum_{n=0}^{\infty} c_n(C)\kappa^n. \quad (21)$$

The coefficients $c_n(C)$ can be obtained by multiplying the denominator with the right-hand side (RHS) of Eq. (21) and collecting all terms of the same order in κ^n . The coefficient $c_0(C)$ is simply given by

$$c_0(C) = \frac{f_0(C)}{\sum_C f_0(C)}. \quad (22)$$

Next, we define b_n and K_0 as

$$b_n = \sum_C f_n(C), \quad n = 0, \dots, K_0, \quad (23)$$

$$K_0 = \max\{K(C)\}, \quad (24)$$

and we set the value of $f_n(C)$ to zero for $K(C) < n \leq K_0$. The coefficient $c_n(C)$ then reads

$$c_n(C) = \frac{f_n(C)}{b_0} - \sum_{m=0}^{n-1} \frac{b_{n-m}}{b_0} c_m(C) \quad (25)$$

for $0 < n \leq K(C)$ and

$$c_n(C) = - \sum_{m=1}^{\min\{K_0, n\}} \frac{b_m}{b_0} c_{n-m}(C) \quad (26)$$

for $n > K(C)$. Details of this calculation are presented in Appendix A.

Furthermore, we can use the fact that $\sum_C P(C) = 1$, which translates to the following condition on the coefficients $c_n(C)$ for any n :

$$\sum_C c_n(C) = \delta_{n,0}, \quad n = 0, \dots, K_0, \quad (27)$$

where $\delta_{n,0}$ is the Kronecker δ function.

In this paper we compute low-order coefficients for the Taylor expansion of $P(C)$ in $\kappa = \alpha \ll \beta, \omega_i$ (Sec. III B) and $\kappa = \beta \ll \alpha, \omega_i$ (Sec. III C). It is important to emphasize that these Taylor coefficients are exact. Thus the only approximation that we make is to replace the Taylor series with the corresponding Taylor polynomial, which is valid when the value of the expansion parameter is small compared to other transition rates.

B. Power series in α

1. Zero-order coefficients

As it can be seen from Eq. (18), the only configuration for which $e_0(C) = 0$ when $\kappa = \alpha$ is the empty lattice, which we denote by $C = \emptyset$. Following the argument given in Sec. III A, Eq. (20) applies for all configurations except the empty one and it has the trivial solution $f_0(C) = 0$ if $C \neq \emptyset$. We thus conclude that the value of $f_0(C)$ is nonzero only for the empty configuration. Therefore, we can write $f_0(C) = \delta_{C,\emptyset} f_0(\emptyset)$ and inserting it into Eq. (22) yields

$$c_0(C) = \frac{f_0(C)}{\sum_C f_0(C)} = \frac{\delta_{C,\emptyset} f_0(\emptyset)}{\sum_C \delta_{C,\emptyset} f_0(\emptyset)} = \delta_{C,\emptyset}, \quad (28)$$

which after insertion into (21) gives

$$P(C) = \delta_{C,\emptyset} + O(\alpha). \quad (29)$$

The zeroth-order solution is therefore equivalent to setting $\alpha = 0$ in the original master equation: when particles are not allowed to enter, the steady state is an empty lattice.

2. Proof that many coefficients $f_n(C)$ are equal to zero

We can use the Schnakenberg network theory to show that many coefficients $f_n(C)$ are zero too, not just for $n = 0$. Let us consider a configuration $C \neq \emptyset$ and let us define $N(C)$ as the number of particles in C :

$$N(C) = \sum_{i=1}^L \tau_i(C). \quad (30)$$

According to the Schnakenberg network theory,

$$P(C) \propto \sum_{n=0}^{K(C)} f_n(C)\alpha^n = \sum_{T_C} w(T_C), \quad (31)$$

where the last sum runs over all spanning trees T_C rooted at C and $w(T_C)$ is the product of the rates of all transitions (edges) contained in T_C . For a given T_C , there is a unique path from any $C' \neq C$ to C . Let us now consider C' to be the empty configuration, $C' = \emptyset$. A directed path from $C' = \emptyset$ to C must include *at least* $N(C)$ transitions at rate α , since a number $N(C)$ of particles must enter the lattice in order to reach configuration C (note that we say at least because some particles may as well leave the system). In order for the relation in Eq. (31) to hold,

we thus conclude that $f_n(C) = 0$ if n is smaller than the number of particles $N(C)$ present in configuration C , or equivalently,

$$f_n(C) \neq 0 \quad \text{if and only if } N(C) \leq n \leq K(C). \quad (32)$$

Relation (32) is crucial for our analysis, as it drastically reduces the number of unknowns. For $n = 0$, it shows that $f_0(C) \neq 0$ for only one configuration, the empty lattice. For $n = 1$, there are $L + 1$ configurations that have nonzero coefficients, L configurations with one particle, and the empty lattice, where L is the number of sites in the lattice. For $n = 2$, the number of nonzero coefficients is $L(L - 1)/2 + L + 1$, where $L(L - 1)/2$ is the number of configurations with two particles, and $L + 1$ is the number of configurations with 1 or zero particles. In general, the number of nonzero coefficients of order n grows polynomially in L and is given by the partial sum of binomial coefficients, $\sum_{j=0}^n \binom{L}{j}$.

3. First-order coefficients

According to relation (32), for $n = 1$, we need to consider configurations with at most one particle. Let us label a configuration with one particle at site i by 1_i . The equations for $f_1(1_i)$, $i = 1, \dots, L$, are given by

$$\omega_1 f_1(1_1) = f_0(\emptyset), \quad (33a)$$

$$\omega_i f_1(1_i) = \omega_{i-1} f_1(1_{i-1}), \quad i = 2, L - 1, \quad (33b)$$

$$\beta f_1(1_L) = \omega_{L-1} f_1(1_{L-1}), \quad (33c)$$

where we have set $f_1(C) = 0$ for all C that contain more than one particle, following relation (32). These equations can be easily solved, yielding

$$f_1(1_i) = \frac{f_0(\emptyset)}{\omega_i}, \quad (34)$$

where we have introduced the notation $\omega_L = \beta$. Inserting Eq. (34) into (23) and (25) gives

$$c_1(\emptyset) = - \sum_{i=1}^L \frac{1}{\omega_i}, \quad (35a)$$

$$c_1(1_i) = \frac{1}{\omega_i}, \quad i = 1, \dots, L, \quad (35b)$$

$$c_1(C) = 0, \quad 2 \leq \sum_{i=1}^L \tau_i(C) \leq L. \quad (35c)$$

The average particle current J is obtained by inserting Eqs. (28), (35a), and (35b) into Eq. (21) and then into (3), which up to the quadratic term yields

$$J = \alpha - \frac{1}{\omega_1} \alpha^2 + O(\alpha^3). \quad (36)$$

Similarly, the expressions for the local and total density ρ_i and ρ read, respectively,

$$\rho_i = \frac{\alpha}{\omega_i} + O(\alpha^2), \quad (37)$$

$$\rho = \frac{1}{L} \left(\sum_{i=1}^L \frac{1}{\omega_i} \right) \alpha + O(\alpha^2). \quad (38)$$

We can check that Eqs. (36) and (38) give the familiar result for the homogeneous case with $\omega_i = \omega$, $J = \alpha(1 - \alpha/\omega)$, and $\rho = \alpha/\omega$.

4. Second-order coefficients

According to relation (32), for $n = 2$ we need to consider configurations with at most two particles. As before, we label configurations by the positions of their particles, so that $C = 1_i 1_j$ for $i = 1, \dots, L - 1$ and $j = i + 1, \dots, L$ denote configurations with two particles and $C = 1_i$ for $i = 1, \dots, L$ denote configurations with one particle.

We first look at configurations with a particle at site 1. Essentially, we have to consider three main cases: $C = 1_1 1_2$, $C = 1_1 1_j$ with $j > 2$, and $C = 1_1$. Applying Eq. (19) to these three cases yields

$$\omega_2 f_2(1_1 1_2) = f_1(1_2), \quad (39a)$$

$$(\omega_1 + \omega_j) f_2(1_1 1_j) = f_1(1_j) + \omega_{j-1} f_2(1_1 1_{j-1}), \quad (39b)$$

$$\omega_1 f_2(1_1) = f_1(\emptyset) + \omega_L f_2(1_1 1_L), \quad (39c)$$

where $f_1(1_j) = f_0(\emptyset)/\omega_j$. Equation (39b) is a recurrence relation in j with Eq. (39a) as an initial condition. This is a nonhomogeneous recurrence relation with variable coefficients that can be turned into a nonhomogeneous recurrence relation with constant coefficients and solved explicitly, yielding

$$f_2(1_1 1_2) = \frac{f_0(\emptyset)}{\omega_2^2}, \quad (40a)$$

$$f_2(1_1 1_j) = \frac{f_0(\emptyset)}{\omega_1 \omega_j} \left[1 + \left(\frac{\omega_1}{\omega_2} - 1 \right) \prod_{q=3}^j \frac{\omega_q}{\omega_1 + \omega_q} \right]. \quad (40b)$$

We can now compute $f_2(1_1)$ by inserting Eq. (40b) for $j = L$ into Eq. (39c). Inserting Eqs. (39c), (40a), and (40b) into (25) gives the following expressions for $c_2(1_1 1_j)$ and $c_1(1_1)$, respectively:

$$c_2(1_1 1_2) = \frac{1}{\omega_2^2}, \quad (41a)$$

$$c_2(1_1 1_j) = \frac{1}{\omega_1 \omega_j} \left[1 + \left(\frac{\omega_1}{\omega_2} - 1 \right) \prod_{q=3}^j \frac{\omega_q}{\omega_1 + \omega_q} \right], \quad (41b)$$

$$c_2(1_1) = \frac{1}{\omega_1^2} \left[1 + \left(\frac{\omega_1}{\omega_2} - 1 \right) \prod_{q=3}^L \frac{\omega_q}{\omega_1 + \omega_q} \right] - \frac{1}{\omega_1} \sum_{i=1}^L \frac{1}{\omega_i}. \quad (41c)$$

It is important to note that these coefficients do not depend on the coefficients $f_0(\emptyset)$ and $f_1(\emptyset)$. Otherwise, our method would not be useful as we cannot say anything about these coefficients.³ We show in Sec. III B 7 that $f_n(\emptyset)$ are not needed until we want to compute $c_n(C)$ for $n > K(C)$, which is the equivalent of solving the full master equation (2).

We argue that these are the only second-order coefficients that we need in order to compute the cubic term in the power

³In Ref. [29] it was erroneously stated that all $f_n(\emptyset)$ are equal.

series of J , since

$$\begin{aligned}
 J &= \alpha \sum_C P(C)[1 - \tau_1(C)] = \sum_{n=0}^{\infty} \left(\sum_{\substack{C \\ \tau_1=0}} c_n(C) \right) \alpha^{n+1} \\
 &= \sum_{n=0}^{\infty} \left[\delta_{n,0} - \sum_{\substack{C \\ \tau_1=1}} c_n(C) \right] \alpha^{n+1} = \alpha - \frac{1}{\omega_1} \alpha^2 \\
 &\quad - \left[\sum_{j=2}^L c_2(1_1 1_j) + c_2(1_1) \right] \alpha^3 + O(\alpha^4), \quad (42)
 \end{aligned}$$

where we have used Eq. (27) in the second line. After inserting (41a)–(41c) into (42), the expression for J reads

$$\begin{aligned}
 J &= \alpha - \frac{1}{\omega_1} \alpha^2 + \left(\frac{1}{\omega_1} - \frac{1}{\omega_2} \right) \\
 &\quad \times \left[\frac{1}{\omega_2} + \sum_{j=3}^L \left(\frac{1}{\omega_j} + \delta_{j,L} \frac{1}{\omega_L} \right) \prod_{q=3}^j \frac{\omega_q}{\omega_1 + \omega_q} \right] \alpha^3 \\
 &\quad + O(\alpha^4). \quad (43)
 \end{aligned}$$

We note that the cubic term in Eq. (43) is equal to zero when $\omega_1 = \omega_2$. This includes the homogeneous case in which $\omega_1 = \dots = \omega_{L-1} = 1$, which was known before from the exact solution [8].

5. Power series solution in the mean-field approximation

Before we compare our predictions to exact (numerical) results from Monte Carlo simulations in Sec. III B 6, it is instructive to analyze how Eq. (43) compares to the mean-field (MF) solution used in other approaches [20,26].

The MF approximation amounts to replacing $\langle \tau_i \tau_j \rangle$ for $i \neq j$ with $\rho_i \rho_j$, where $\rho_i = \langle \tau_i \rangle$ is the local particle density. This approximation leads to the following mean-field equations for J and ρ_i :

$$\begin{aligned}
 J &= \alpha(1 - \rho_1) = \omega_1 \rho_1(1 - \rho_2) = \dots \\
 &= \omega_{L-1} \rho_{L-1}(1 - \rho_L) = \beta \rho_L. \quad (44)
 \end{aligned}$$

The exact (closed-form) solution of these equations is unknown. Instead, we can look for a perturbative solution for small α in the following form:

$$\rho_i = \sum_{n=0}^{\infty} \rho_i^{(n)} \alpha^n. \quad (45)$$

Inserting Eq. (45) into (44) and collecting terms of the same power of α for small n leads to

$$\begin{aligned}
 \rho_i^{(0)} &= 0, \quad \rho_i^{(1)} = \frac{1}{\omega_i}, \\
 \rho_i^{(2)} &= \begin{cases} \frac{1}{\omega_i} \left(\frac{1}{\omega_{i+1}} - \frac{1}{\omega_i} \right), & i = 1, \dots, L-1, \\ -\frac{1}{\omega_1 \omega_L}, & i = L. \end{cases} \quad (46)
 \end{aligned}$$

Inserting these coefficients back into Eq. (45) and then into (44), we get the following expression for current J_{MF} in the

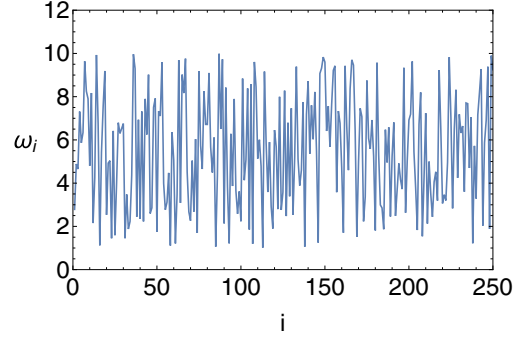


FIG. 3. Hopping rates ω_i for $i = 1, \dots, L = 250$ used to compute J in Fig. 4.

mean-field approximation:

$$J_{MF} = \alpha - \frac{1}{\omega_1} \alpha^2 + \left(\frac{1}{\omega_1} - \frac{1}{\omega_2} \right) \frac{1}{\omega_1} \alpha^3 + O(\alpha^4). \quad (47)$$

While the first two terms in Eqs. (43) and (47) are the same, the third term in Eq. (47) depends only on ω_1 and ω_2 , which is markedly different from the third term in Eq. (43), which depends on all ω_i . Hence the mean-field results clearly deviate from the main result of this paper, i.e., the power series solution derived in Eq. (43).

6. Monte Carlo simulations

We now compare the power series of the current $J(\alpha)$ to the exact current that we compute numerically using kinetic Monte Carlo simulations based on the Gillespie algorithm for $L = 250$ lattice sites. For that purpose we generated $\omega_1, \dots, \omega_L$ randomly from the uniform distribution on $[1, 10]$, which are plotted in Fig. 3.

In Fig. 4, the first three terms in the power series of J in α corresponding to $n = 0, 1$, and 2 were computed analytically, using Eqs. (36) and (43), respectively. The fourth term corresponding to $n = 3$ was computed numerically using a simple algorithm that we present in Appendix B. The

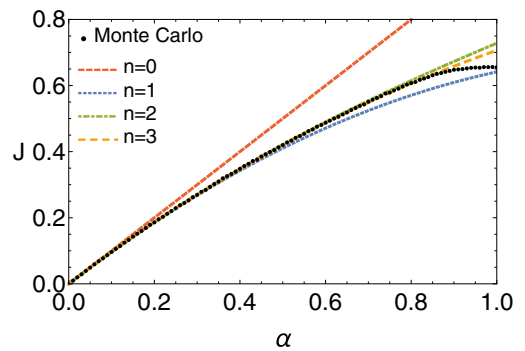


FIG. 4. Current J as a function of the entrance rate α . The black dots show the result of simulations obtained using the Gillespie algorithm (averaged over 500 independent runs). The dashed red ($n = 0$) and dotted blue ($n = 1$) lines are linear and quadratic approximations from Eq. (36), respectively. The dot-dashed green ($n = 2$) line is cubic approximation from Eq. (43). The dashed orange ($n = 3$) line is quartic approximation, which was computed numerically.

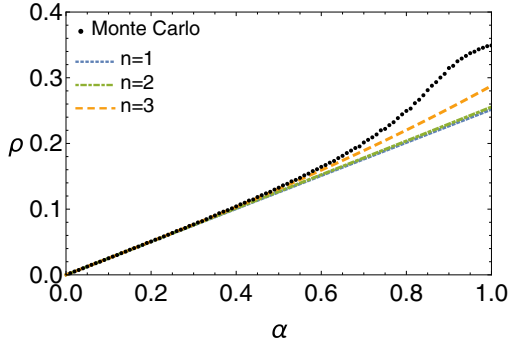


FIG. 5. Total density ρ as a function of the entrance rate α . The black dots show the result of simulations obtained using the Gillespie algorithm (averaged over 500 independent runs). The dashed blue line ($n = 1$) is linear approximation from Eq. (38). The dotted green ($n = 2$) and dashed orange ($n = 3$) lines are quadratic and cubic approximations, respectively, which were computed numerically.

agreement between the simulation results and the analytical expressions visibly increases as we include more terms in the power series.

Unlike J , the total density ρ depends on all second-order coefficients $c_2(C)$. The equation for $f_2(1_i 1_j)$ is a recurrence relation in two indices i and j , which is difficult to solve explicitly. In Fig. 5 we compare the power series of the total density $\rho(\alpha)$ for the same choice of ω_i as in Fig. 4 with the Monte Carlo simulations. The linear term corresponding to $n = 1$ was computed from Eq. (38), while the quadratic and cubic approximations corresponding to $n = 2$ and 3 were computed numerically.

As we can already see for the second-order coefficients, handling higher orders becomes progressively more difficult. In the next two sections we develop a general method for computing higher-order coefficients $c_n(C)$, revealing an interesting connection with combinatorial objects called Young tableaux.

7. Higher-order coefficients

Our starting point is the recurrence relation in Eq. (19), which after rearranging gives

$$f_n(C) = \sum_{C'} \frac{I_{C',C}}{e_0(C)} f_{n-1}(C') - \left(\sum_{C'} \frac{I_{C',C'}}{e_0(C)} \right) f_{n-1}(C) + \sum_{C'} \frac{(1 - I_{C',C})}{e_0(C)} W(C' \rightarrow C) f_n(C'), \quad C \neq \emptyset. \tag{48}$$

We cannot immediately find $f_n(C)$ by iterating over n , because the RHS of Eq. (48) contains coefficients of order n as well. The solution is to picture Eq. (48) as a recurrence relation not only in order n , but also in the configuration C . The iteration stops when all coefficients on the RHS of Eq. (48) belong to the empty configuration $C' = \emptyset$. That is because the recurrence relation in Eq. (48) does not hold for $C = \emptyset$. The final expression for $f_n(C)$ for $N(C) \leq n \leq K(C)$ must then be of the form

$$f_n(C) = \sum_{m=N(C)}^n \mu_m(C) f_{n-m}(\emptyset), \tag{49}$$

where $\mu_m(C)$ is an unknown coefficient. As we show later in more detail, the sum starts from $m = N(C)$ because of Eq. (32).

The expression for $f_n(C)$ in Eq. (49) can be used to calculate $c_n(C)$ in the power series of $P(C)$. According to Eqs. (23) and (25), the expression for $f_n(C)$ for $N(C) \leq n \leq K(C)$ is given by

$$f_n(C) = b_0 c_n(C) + \sum_{m=0}^{n-1} b_{n-m} c_m(C). \tag{50}$$

By inserting this expression for $C = \emptyset$ into Eq. (49) and comparing the term containing b_0 to the one in Eq. (50), we conclude that

$$c_n(C) = \sum_{m=N(C)}^n \mu_m(C) c_{n-m}(\emptyset). \tag{51}$$

The recurrence relation for $c_n(\emptyset)$ can be obtained by inserting Eq. (51) into Eq. (27), which gives

$$c_n(\emptyset) = - \sum_{C \neq \emptyset} \sum_{m=N(C)}^n \mu_m(C) c_{n-m}(\emptyset), \tag{52}$$

and the initial condition is $c_0(\emptyset) = 1$.

The expression for $c_n(C)$ in Eq. (51), together with the recurrence relation for $c_n(\emptyset)$ in Eq. (52) is one of the main results of this paper. It shows that the first $K(C)$ coefficients in the power series of $P(C)$ in α are determined by the coefficients $\mu_m(C)$ and not by the unknown coefficients $f_m(\emptyset)$ that our method cannot determine. The situation changes for $n > K(C)$, for which Eq. (50) is replaced by

$$c_n(C) = \sum_{m=1}^{\min\{K_0, n\}} \frac{b_m}{b_0} c_{n-m}(C), \quad n > K(C), \tag{53}$$

where we recall that $K_0 = \max_C \{K(C)\}$. In this case $c_n(C)$ depends on b_m , which in turn depends on the unknown coefficients $f_m(\emptyset)$. In other words, we cannot find $c_n(C)$ without finding all $f_m(\emptyset)$. That is not surprising: according to Eqs. (12) and (49), finding all $\mu_m(C)$ and $f_m(\emptyset)$ amounts to finding the full solution of the master equation. In that context, the result in Eq. (53) tells us that we cannot say anything about $c_n(C)$ for $n > K(C)$ without solving the original master equation fully. Since we expect $K(C)$ to grow exponentially with the system size L , these coefficients are inaccessible for all practical purposes.

In the next two sections we complete the power series of $P(C)$ in α by showing how to compute the coefficients $\mu_m(C)$.

8. Back-substitution method for finding $\mu_n(C)$

In this section we show how to find $f_n(C)$ and thus $\mu_n(C)$ by backward substitution until all terms on the RHS of Eq. (48) belong to the empty configuration \emptyset .

The first sum on the RHS of Eq. (48) is zero unless C contains a particle at site 1,

$$\sum_{C'} \frac{I_{C',C}}{e_0(C)} f_{n-1}(C') = \frac{\tau_1(C)}{e_0(C)} f_{n-1}(0, \tau_2, \dots). \tag{54}$$

This contribution to $f_n(C)$ is obtained from C by removing a particle at site 1, provided $\tau_1(C) = 1$. The resulting coefficient

TABLE I. Iteration steps (rules) for going from C' on the RHS to C on the LHS of Eq. (48).

Rule	Iteration ($C' \rightarrow C$)	Coefficient	Weight
1	$0_1 \rightarrow 1_1$	$f_{n-1}(C')$	$1/e_0(C)$
2	$0_1 \rightarrow 0_1$	$f_{n-1}(C)$	$(-1)/e_0(C)$
3	$1_i 0_{i+1} \rightarrow 0_i 1_{i+1}$	$f_n(C')$	$\omega_i/e_0(C)$
4	$1_L \rightarrow 0_L$	$f_n(C')$	$\beta/e_0(C)$

is $f_{n-1}(0, \tau_2(C), \dots, \tau_L(C))$, which is multiplied by the weight $1/e_0(C)$. We call this step *rule 1*.

The second sum is zero unless site 1 in C is empty:

$$\left(\sum_{C'} \frac{I_{C,C'}}{e_0(C)} \right) f_{n-1}(C) = \frac{1 - \tau_1(C)}{e_0(C)} f_{n-1}(C). \quad (55)$$

In this step we reduce the order of $f_n(C)$ to $f_{n-1}(C)$, provided $\tau_1(C) = 0$. The resulting coefficient $f_{n-1}(C)$ is multiplied by the weight $(-1)/e_0(C)$. We call this step *rule 2*.

The third sum runs over all configurations C' that lead to C by moving one particle forwards. Each of these moves contributes to $f_n(C)$ with $f_n(C')$, multiplied by the weight $W(C' \rightarrow C)/e_0(C)$. We call this step *rule 3* if a pair of variables $(\tau_i = 0, \tau_{i+1} = 1)$ in C is replaced by $(\tau_i = 1, \tau_{i+1} = 0)$ in C' and *rule 4* if $\tau_L = 0$ in C is replaced by $\tau_L = 1$ in C' .

In summary, rule 1 reduces order n to $n - 1$ and removes a particle from site 1. Rule 2 reduces order n to $n - 1$ but leaves the configuration unchanged. Rule 3 moves a particle at site $2 \leq i \leq L$ to $i - 1$, provided site $i - 1$ is empty. Rule 4 moves a particle from the right reservoir to site L , provided site L is empty. A summary of all possible transitions is presented in Table I.

The idea is to repeat these rules for all coefficients on the RHS of Eq. (48) until we get a coefficient $f_m(\emptyset)$ that belongs to the empty configuration. Since the recurrence relation in Eq. (48) does not apply to $C = \emptyset$, we leave this coefficient as it is. The other possibility is to get a coefficient $f_m(C')$ for which the number of particles $N(C') > m$. This can happen because of the rule 4 that increases the number of particles. However, this coefficient does not contribute to $f_n(C)$ because of Eq. (32). This is how we obtain the solution for $f_n(C)$ presented in Eq. (49).

Example. Let us say we want to compute $f_2(1_1 1_3)$ for a lattice of $L = 4$ sites. We can apply rule 1 to $C = 1_1 1_3$ leading to $C' = 1_3$ and rule 3 leading to $C' = 1_1 1_2$. The corresponding weights are $1/e_0(1_1 1_3)$ and $\omega_2/e_0(1_1 1_2)$, respectively. On this occasion we do not apply rule 4 leading to $C' = 1_1 1_3 1_4$ because we know from Eq. (32) that $f_2(1_1 1_3 1_4) = 0$. We can thus write $f_2(1_1 1_3)$ as

$$f_2(1_1 1_3) = \frac{1}{e_0(1_1 1_3)} f_1(1_3) + \frac{\omega_2}{e_0(1_1 1_3)} f_2(1_1 1_2). \quad (56)$$

Now we look at each of the configurations 1_3 and $1_1 1_2$ on the RHS separately. We apply rule 1 to $C = 1_3$ leading to $C' = 1_2$, whereby the corresponding weight is $\omega_2/e_0(1_3)$. Similarly, we apply rule 1 to $C = 1_1 1_2$ leading to $C' = 1_2$, whereby the corresponding weight is $1/e_0(1_2)$. Now $f_2(1_1 1_3)$

reads

$$f_2(1_1 1_3) = \frac{\omega_2}{e_0(1_1 1_3)e_0(1_3)} f_1(1_2) + \frac{\omega_2}{e_0(1_1 1_3)e_0(1_1 1_2)} f_1(1_2). \quad (57)$$

Next, we apply rule 3 to $C = 1_2$ leading to $C' = 1_1$, which is weighted by $\omega_1/e_0(1_2)$. In the final step we apply rule 1 to $C = 1_1$ leading to $C' = \emptyset$, which is weighted by $1/e_0(1_1)$. Altogether, the expression for $f_2(1_1 1_3)$ reads

$$f_2(1_1 1_3) = \frac{\omega_1 \omega_2}{e_0(1_1 1_3)e_0(1_3)e_0(1_2)e_0(1_1)} f_0(\emptyset) + \frac{\omega_1 \omega_2}{e_0(1_1 1_3)e_0(1_1 1_2)e_0(1_2)e_0(1_1)} f_0(\emptyset), \quad (58)$$

so that $\mu_2(1_1 1_3)$ is given by

$$\mu_2(1_1 1_3) = \frac{\omega_1 \omega_2}{e_0(1_1 1_3)e_0(1_3)e_0(1_2)e_0(1_1)} + \frac{\omega_1 \omega_2}{e_0(1_1 1_3)e_0(1_1 1_2)e_0(1_2)e_0(1_1)}. \quad (59)$$

In the next section we provide a formal expression for the coefficients $\mu_m(C)$. In particular, we show that the successive application of rules 1–4 can be graphically represented by combinatorial objects called Young tableaux of shifted shape.

9. Coefficients μ_m and Young tableaux

By definition, $\mu_m(C)$ is a sum of products of weights of all iteration steps in Table I that connect $f_n(C)$ to $f_{n-m}(\emptyset)$ in Eq. (49). By construction, we know immediately that

$$\mu_m(C) = 0, \quad m < N, \quad (60)$$

where N is the number of particles in C . We thus have to consider only $m \geq N$.

Let us first consider the case of $m = N$ for which $f_N(C) = \lambda_N(C) f_0(\emptyset)$. In order to compute $\mu_N(C)$, we have to find all directed paths from \emptyset to C that insert exactly N particles from the left reservoir. Let us label the N particles by integers $1, 2, \dots, N$ in the order in which they enter the lattice. Next, we represent a directed path from \emptyset to C by a sequence of moves, whereby each move in the sequence is labeled by the label of the particle that made that move.

For example, in order to get from \emptyset to $1_2 1_4$, we need to construct integer sequences consisting of four labels 1 and two labels 2. In the first move, we insert particle 1 from the left reservoir and the resulting sequence is 1. In the second move, we move particle 1 again and the resulting sequence is 11. In the third move, we can either move particle 1 leading to 111 or we can insert particle 2 into the lattice leading to 112. We repeat these steps until particle 1 reaches site 4 and particle 2 reaches site 2. The list of all final sequences representing directed paths from \emptyset to $1_2 1_4$ is 111122, 111212, 111221, 112112, and 112121, five in total.

A distinguished property of these integer sequences is that the number of 1's is strictly greater than the number of 2's. This is not only true for the final sequence, but also for all initial subsequences leading to the final sequence and is a direct consequence of exclusion that forbids two or more particles to share the same lattice site. For example, sequence 121112 in the example above is forbidden because its initial subsequence

is 12, which implies that particle 2 moves onto site 1 that is already occupied by particle 1. The same property extends to sequences and all their initial subsequences that contain more than two particles, so that the number of particles labeled by k is strictly larger than the number of particles labeled by $k + 1$, for all $k = 1, \dots, N - 1$.

Sequences of integers 1 and 2 in which the number of 1 remains greater than the number of integers 2 in all initial subsequences have a long history in mathematics (see Ref. [38], for example). The famous Bertrand’s Ballot Problem asks for the probability that candidate A receiving p votes in total stays ahead of candidate B receiving $q < p$ votes in total as the votes are counted [39]. When there are more than two candidates, we speak of generalized ballot sequences. Confusingly, the names *generalized ballot sequence*, *lattice permutation*, or *lattice word* that are nowadays used for such sequences assume that the number of integers k is *no less* than the number of integer $k + 1$, rather than being strictly greater. Terms “weak” in the former and “strict” in the latter case are sometimes used to distinguish these two cases.

There is a graphical way to represent generalized ballot sequences using combinatorial objects called Young tableaux [34]. A *Young diagram* of shape $\lambda = (\lambda_1, \dots, \lambda_k)$ is a left-justified shape of k rows of boxes of length $\lambda_1, \dots, \lambda_k$.⁴ A Young diagram whose boxes are filled with integers $1, 2, \dots, \lambda_1 + \dots + \lambda_k$ that are strictly increasing along rows and columns is called a *standard Young tableau*. A Young tableau of shifted shape $\lambda = (\lambda_1, \dots, \lambda_k)$ is a standard Young tableau whose i th row is indented by $i - 1$ boxes. In the following example,

$$\begin{array}{|c|c|c|} \hline & & \\ \hline & & \\ \hline & & \\ \hline \end{array}, \quad \begin{array}{|c|c|c|c|c|} \hline 1 & 2 & 5 & 6 \\ \hline 3 & 4 & & \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 5 \\ \hline & 4 & 6 & \\ \hline \end{array}, \quad (61)$$

the first is a Young diagram of shape (4,2), the second is a standard Young tableau of shape (4,2), and the third is a standard Young tableau of shifted shape (4,2). A standard Young tableau corresponds to a weak generalized ballot sequence and a standard Young tableau of shifted shape corresponds to a strict generalized ballot sequence.

For a given configuration $C = 1_{x(N)} \dots 1_{x(1)}$ with N particles at positions $x(1) > x(2) > \dots > x(N)$ (labeled in the order in which they enter the lattice), there is a one-to-one correspondence between a directed path from \emptyset to C and a Young tableau of shifted shape $x = [x(1), \dots, x(N)]$. We recall that a directed path $S(C)$ from \emptyset to C is a sequence of $x(1)$ integers 1, $x(2)$ integers 2, and so on, whereby the number of integers k is strictly larger than the number of integers $k + 1$ for all $k = 1, \dots, N - 1$ and for all subsequences of $S(C)$. Hence $x(i)$ designates both the position on the lattice of the particle i (used to write the configuration C), and the number of moves that the particle has experienced in going from \emptyset to C . The connection between $S(C)$ and the Young tableau of shifted shape x is made by filling each box with a number $\in \{1, \dots, x(1) + \dots + x(N)\}$ that corresponds to the position of that move in the sequence. For example, a Young tableau of

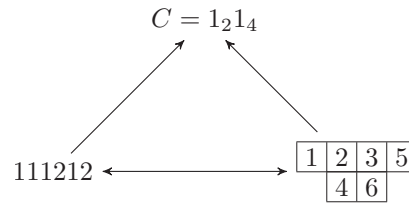


FIG. 6. Correspondence between a directed path from \emptyset to $C = 1_{x(N)} \dots 1_{x(1)}$ and a Young tableau of shifted shape $[x(1), \dots, x(N)]$ for $N = 2, x(1) = 4$, and $x(2) = 2$.

shifted shape (4,2) in Eq. (61) corresponds to a directed path from \emptyset to $C = 1_2 1_4$ that is represented by sequence 111212 (Fig. 6).

We are now ready to state the expression for $\mu_N(C)$, where N is the number of particles in $C = 1_{x(N)} \dots 1_{x(1)}$. Let us denote by $t(x)$ a standard Young tableau of shifted shape $x = [x(1), \dots, x(N)]$. This tableau corresponds to a sequence of $x(1)$ integers 1, $x(2)$ integers 2, and so on, which we denote by $S(C)$. We call C_k a configuration that corresponds to the initial subsequence of $S(C)$ of length k . The expression for $\mu_N(C)$ is then given by

$$\mu_N(1_{x(1)} \dots 1_{x(N)}) = \prod_{i=1}^N \prod_{j=2}^{x(i)} \omega_{j-1} \sum_{t(x)} \prod_{k=1}^l \frac{1}{e_0(C_k)}, \quad (62)$$

where the sum goes over all standard Young tableaux of shifted shape x . The product in front of the sum comes from the numerator of rule 3 in Table I and depends on the corresponding Young diagram but not on how the boxes are filled. The product after the sum goes over all iteration steps (l in total) that lead from \emptyset to C . For example, the expression for $\mu_2(1_2 1_4)$ is given by

$$\begin{aligned}
 \mu_2(1_2 1_4) = & \omega_1^2 \omega_2 \omega_3 \left[\frac{1}{\omega_1 \omega_2 \omega_3 \omega_4 (\omega_1 + \omega_4) (\omega_2 + \omega_4)} \right. \\
 & + \frac{1}{\omega_1 \omega_2 \omega_3 (\omega_1 + \omega_3) (\omega_1 + \omega_4) (\omega_2 + \omega_4)} \\
 & + \frac{1}{\omega_1 \omega_2 \omega_3 (\omega_1 + \omega_3) (\omega_2 + \omega_3) (\omega_2 + \omega_4)} \\
 & + \frac{1}{\omega_1 \omega_2^2 (\omega_1 + \omega_3) (\omega_1 + \omega_4) (\omega_2 + \omega_4)} \\
 & \left. + \frac{1}{\omega_1 \omega_2^2 (\omega_1 + \omega_3) \omega_3 (\omega_2 + \omega_4)} \right], \quad (63)
 \end{aligned}$$

in which the terms in square brackets correspond respectively to sequences 111122, 111212, 111221, 112112, and 112121.

So far we have discussed directed paths contributing to $\mu_m(C)$, where $m = N$ is the number of particles in C . The situation changes for $m > N$, because we may also apply rules 2 and 4 in addition to rules 1 and 3. Let us denote by m_1, m_2 , and m_4 the number of times that we applied rules 1, 2, and 4 in a directed path $S(C)$ from \emptyset to C . Obviously, $m_1 = N + m_4$ and $m = m_1 + m_2$.

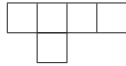
Let us consider the case $m_2 = 0$ first, so that the number of particles entering and leaving the system is $m_1 = m$ and $m_4 = N - m$, respectively. We set $x(1) = \dots = x(m - N) = L + 1$ for the particles that leave the system and the remaining

⁴We are using here the English notation for the appearance of Young diagrams; see Ref. [34] for other notations.

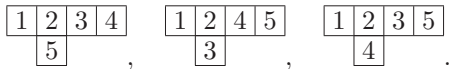
particles have positions $x(m - N + 1) > \dots > x(m)$. The only change for $m_4 > 0$ and $m_2 = 0$ compared to the previous case of $m_4 = m_2 = 0$ (i.e., $m = N$) is that a Young tableau $t(x)$ that corresponds to $S(C)$ now has a shifted shape $x = [x(1) = L + 1, \dots, x(m - N) = L + 1, x(m - N + 1), \dots, x(m)]$, i.e., the first m_4 rows are of length $L + 1$.

However, care must be taken to exclude directed paths from \emptyset to C that revisit the empty lattice configuration. Let us denote by $t_{k,j}(x)$ the value of the j th box in the k th row of $t(x)$ starting from top to bottom. The requirement that the empty configuration is not revisited translates to the condition that $t_{k,L+1}(x) \neq k(L + 1)$ for all $k = 1, \dots, m_4 = m_1 - N$. Another way of stating this condition is to require that none of the subtableaux of $t(x)$, which are obtained by retaining only the first $k = 1, \dots, m - 1$ rows of $t(x)$, are themselves *standard* Young tableaux of shifted shape $[x(1), \dots, x(k)]$.

Example. Let us consider the coefficient $\mu_2(1_1)$ for $L = 3$, so that $m = 2, N = 1, m_1 = 2, m_2 = 0$, and $m_4 = 1$. In order to get from \emptyset to $C = 1_1$, we need to insert two particles and remove one ending with $x(1) = L + 1 = 4$ and $x(2) = 1$. The corresponding Young diagram is



which we need to fill with integers $1, \dots, x(1) + x(2) = 5$ that increase in all rows and columns. There are three possible fillings, which are given by



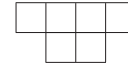
However, the first filling corresponds to the sequence 11112, which revisits the empty configuration after four jumps. Hence only the last two tableaux contribute to $\mu_2(1_1)$.

For $m_2 > 0$, we note that the rule 2 applied to any configuration $C' \in S(C)$ [provided $\tau_i(C') = 0$] changes only the weight of $S(C)$, but not the sequence itself. Let us denote by ϵ_k the number of times that we applied the rule 2 while being in configuration $C_k \in S(C)$, so that $\|\epsilon\| = \epsilon_1 + \dots + \epsilon_l = m_2$. Since rule 2 changes neither $S(C)$ nor the corresponding Young tableau, we call ϵ_k a *degeneracy* of C_k . The expression for $\mu_m(C)$ is then obtained by summing over all m_1 from N to m and over all combinations of degeneracies such that their sum $\|\epsilon\|$ is equal to $m_2 = m - m_1$,

$$\mu_m(C) = \sum_{m_1=N}^m (-1)^{m-m_1} \left(\prod_{i=1}^{m_1} \prod_{j=2}^{x(i)} \omega_{j-1} \right) \times \sum_{\|\epsilon\|=m-m_1} \sum_{t(x)} \prod_{k=1}^l \frac{1}{[e_0(C_k)]^{1+\epsilon_k}}, \quad (64)$$

where we recall that $l = x(1) + \dots + x(m_1)$, $\omega_L = \beta$, and $\|\epsilon\| = \epsilon_1 + \dots + \epsilon_l$. The primed sum means that the value of $(L + 1)$ th box in the k th row of $t(x)$ must not be equal to $k(L + 1)$ for all $k = 1, \dots, m_4 = m_1 - N$ in order to avoid visiting the empty lattice configuration. Together with Eqs. (21), (51), and (52), this result concludes our power series in α for the inhomogeneous TASEP.

Example. Let us compute $\mu_2(1_2)$ for the system of size $L = 3$, so that $m = 2$ and $N = 1$. Following Eq. (64), the first option is to have two particles entering ($m_1 = 2$), no rule 2 applied ($m_2 = 0$), and one particle leaving ($m_4 = 1$) so that $x(1) = L + 1 = 4$ and $x(2) = 2$ for which the corresponding Young diagram is



There are five fillings of this diagram with integers $1, \dots, l = x(1) + x(2) = 6$ corresponding to sequences 111122, 111212, 111221, 112112, and 112121, but the sequence 111122 is excluded because it revisits the empty configuration. The second option is to have one particle entering ($m_1 = 1$), rule 2 applied once ($m_2 = 1$), and no particles leaving ($m_4 = 0$), so that $x(1) = 2$ for which the Young diagram is



with only one filling corresponding to the sequence 11. Of the two configurations visited in the sequence 11, $C_1 = 1_1$ and $C_2 = 1_2$, only C_2 has $\tau_1 = 0$, so that $\epsilon_1 = 0$ and $\epsilon_2 = 1$. The final expression for $\mu_2(1_2)$ has five contributions in total; the first four are from the first option and the last one is from the second option,

$$\begin{aligned} \mu_2(1_2) = & \frac{\omega_1^2 \omega_2 \omega_3}{\omega_1 \omega_2 \omega_3 (\omega_1 + \omega_3) (\omega_1 + \omega_4) (\omega_2 + \omega_4)} \\ & + \frac{\omega_1^2 \omega_2 \omega_3}{\omega_1 \omega_2 \omega_3 (\omega_1 + \omega_3) (\omega_2 + \omega_3) (\omega_2 + \omega_4)} \\ & + \frac{\omega_1^2 \omega_2 \omega_3}{\omega_1 \omega_2^2 (\omega_1 + \omega_3) (\omega_1 + \omega_4) (\omega_2 + \omega_4)} \\ & + \frac{\omega_1^2 \omega_2 \omega_3}{\omega_1 \omega_2^2 (\omega_1 + \omega_3) \omega_3 (\omega_2 + \omega_4)} - \frac{\omega_1}{\omega_1 \omega_2^2}. \quad (65) \end{aligned}$$

C. Power series in β

Rather than repeating the calculations from the previous section, we can use the fact that the model is symmetric with respect to the following symmetry transformations:

$$1_i \longleftrightarrow 0_i, \quad (66a)$$

$$\alpha \longleftrightarrow \beta, \quad (66b)$$

$$i \longleftrightarrow L - i + 1. \quad (66c)$$

The first relation is the particle-hole symmetry which replaces particles with holes. Since holes move in the opposite direction, we also have to reverse the direction of the flow, which is done by the second and third relations. For example, applying these transformations to Eq. (43) gives the first three terms in the power series of J in the exit rate β :

$$\begin{aligned} J = & \beta - \frac{1}{\omega_L} \beta^2 + \left(\frac{1}{\omega_L} - \frac{1}{\omega_{L-1}} \right) \\ & \times \left[\frac{1}{\omega_{L-1}} + \sum_{j=1}^{L-2} \left(\frac{1}{\omega_j} + \delta_{j,1} \frac{1}{\omega_1} \right) \prod_{q=j}^{L-2} \frac{\omega_q}{\omega_L + \omega_q} \right] \beta^3 \\ & + O(\beta^4). \quad (67) \end{aligned}$$

D. Other cases

In Sec. III B we assumed that α is much smaller than any of the rates $\omega_1, \dots, \omega_L$ and β . The results of Sec. III B do not apply to the case in which one (or more) of these rates is equal to α . Instead, each of these cases has to be studied separately starting from the zeroth order. For example, there are $L + 1$ configurations for which $f_0(C) \neq 0$ in the power series in $\alpha = \beta$, which are $C = 1_i \dots 1_L$ for $i = 1, \dots, L$ and the empty lattice $C = \emptyset$.

Another case that we do not cover in this work is the power series in one of the rates ω_i . The simplest scenario is if $\omega_i \ll \alpha, \beta, \omega_j$ for all $j \neq i$ (the ‘‘slow’’ site problem). A special case in which $\omega_i \ll 1$ and $\omega_j = 1$ for all $j \neq i$ has been studied in [29]. Other scenarios in which one or more rates α , β , and ω_j for $j \neq i$ are equal to ω_i are in general more difficult to analyze and are beyond the scope of this paper.

IV. CONCLUSION

We have presented an analytic method for finding the steady state of the inhomogeneous TASEP as a power series in the entrance rate α , assuming that α is small. This is the case of mRNA translation for which the rate of ribosome recruitment α is typically one or two orders of magnitude smaller than the hopping rates ω_i [27,33].

A practical advantage of our method is that the steady state probability $P(C) \sim O(\alpha^N)$, where N is the number of particles in configuration C . Thus the computation of low-order terms is needed for only a small fraction of all configurations. In this paper we performed the expansion up to the second order, which allowed us to find an analytic expression for the density ρ up to the quadratic order and for the particle current J up to the cubic order in α . We also presented an algorithm for computing higher-order terms recursively for small n .

The exact steady state of the inhomogeneous TASEP is a long outstanding problem in nonequilibrium statistical physics. Additionally, the mean-field approximation can only be found numerically [20]. Our analytic method reduces this gap and reveals how density and particle current depend on the particular sequence of hopping rates. The outstanding challenge for future work is to get practical information from the higher-order terms.

Our framework, which is applied here to the standard TASEP, can be extended to other models based on the exclusion process. In Ref. [30] we applied this method to a two-state [40,41] TASEP with extended particles of size $\ell = 10$ [20], which allowed us to predict mRNA sequence determinants of the rate of translation in yeast. In the future, this approach may help to understand the role of codon usage bias, which remains one of the major unanswered questions in molecular biology, and could also be exploited in bioengineering.

ACKNOWLEDGMENTS

J.S.N. was supported by the Leverhulme Trust Early Career Fellowship under Grant No. ECF-2016-768. M.C.R. was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) Grant No. BB/N017161/1 and the Scottish Universities Life Sciences Alliance. L.C. would like

to thank the CNRS for having granted him a ‘‘demi-délégation’’ (2017–18).

APPENDIX A: DERIVATION OF EQ. (25)

Starting from Eq. (21), we multiply the denominator of $P(C_i)$ by the power series on the RHS, which gives

$$\sum_{n=0}^{K(C_i)} f_n(C_i) \kappa^n = \sum_{j=1}^{\mathcal{N}} \sum_{p=0}^{\infty} \sum_{m=0}^{K(C_j)} c_p f_m(C_j) \kappa^{m+p}. \quad (\text{A1})$$

Next, we introduce

$$K_0 = \max_j \{K(C_j)\} \quad (\text{A2})$$

and set $f_m(C_j) = 0$ for $K(C_j) < m \leq K_0$, so that

$$\sum_{n=0}^{K(C_i)} f_n(C_i) \kappa^n = \sum_{p=0}^{\infty} \sum_{m=0}^{K_0} c_p b_m \kappa^{m+p}, \quad (\text{A3})$$

where b_m is defined in Eq. (23). We now define $n = m + p$ so that

$$\sum_{n=0}^{K(C_i)} f_n(C_i) \kappa^n = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\min\{n, K_0\}} c_{n-m} b_m \right) \kappa^n. \quad (\text{A4})$$

Since by definition $K(C_i) \leq K_0$, for $0 \leq n \leq K(C_i)$ we write

$$\begin{aligned} f_n(C_i) &= \sum_{m=0}^n b_m c_{n-m} = \sum_{m=0}^n b_{n-m} c_m \\ &= b_0 c_n + \sum_{m=0}^{n-1} b_{n-m} c_m, \end{aligned} \quad (\text{A5})$$

which after rearranging gives Eq. (25). Similarly, we write

$$0 = \sum_{m=0}^{\min\{n, K_0\}} b_m c_{n-m} + b_0 c_n + \sum_{m=1}^{\min\{n, K_0\}} b_m c_{n-m}, \quad (\text{A6})$$

for $n > K(C_i)$, which after rearranging gives Eq. (26).

APPENDIX B: ALGORITHM FOR FINDING $c_n(C)$ RECURSIVELY FOR SMALL n

Our starting position for computing $c_n(C)$ is Eq. (25) for $n \leq K(C)$, which reads

$$c_n(C) = \frac{f_n(C) - \sum_{m=0}^{n-1} b_{n-m} c_m(C)}{b_0}. \quad (\text{B1})$$

In Eq. (49) we show that $f_n(C)$ can be written as a linear combination of unknown empty-lattice coefficients $f_0(\emptyset), \dots, f_{N(C)}(\emptyset)$,

$$f_n(C) = \sum_{m=N(C)}^n \mu_m(C) f_{n-m}(\emptyset), \quad (\text{B2})$$

where $N(C)$ is the number of particles in configuration C . Most importantly, we also show that $c_n(C)$ does not depend on any of these empty-lattice coefficients, which means that the numerator in Eq. (B1) must be proportional to $b_0 = f_0(\emptyset)$. In other words, none of the contributions to $f_n(C)$ and b_{n-m} from any of $f_m(\emptyset)$ other than $f_0(\emptyset)$ matter; they all cancel each other

out. For all practical purposes, that is numerically equivalent to setting the value of all $f_{n-m}(\emptyset)$ to zero except for $f_0(\emptyset)$, whose value is set to 1.

The algorithm for computing $c_n(C)$ first finds all $f_n(C)$ and then $c_n(C)$ according to Eq. (B1), assuming that we know $f_{n-1}(C)$ for all C and b_{n-m} for all $m = 0, \dots, n-1$. The first step is to compute $f_n(C)$ for configurations that have precisely n particles [we recall that $f_n(C) = 0$ if $N(C) > n$], starting from the configuration in which n particles occupy the first n sites, $C = 1_1 1_2 \dots 1_n$. In that case Eq. (48) reads

$$f_n(1_1 \dots 1_n) = \frac{1}{\omega_n} f_{n-1}(1_2 \dots 1_n). \quad (\text{B3})$$

The next step is to compute $f_n(1_1 1_{x(2)} \dots 1_{x(n)})$ for all $2 \leq x(2) < \dots < x(n) \leq L$, using the following recurrence equation:

$$\begin{aligned} f_n(1_1 1_{x(2)} \dots 1_{x(n)}) &= \frac{1}{e_0} f_{n-1}(1_{x(2)} \dots 1_{x(n)}) \\ &+ \sum_{k=2}^n \frac{\omega_{x(k)-1}}{e_0} f_n(1_1 \dots 1_{x(k)-1} \dots 1_{x(n)}), \end{aligned} \quad (\text{B4})$$

where $e_0(x)$, $x = [x(1), \dots, x(n)]$ is given by

$$e_0(x) = \sum_{k=1}^{n-1} (1 - \delta_{x(k), x(k+1)-1}) \omega_{x(k)} + \omega_{x(n)}. \quad (\text{B5})$$

In Eq. (B4) we have set the value of $f_n(1_{x(1)} \dots 1_{x(n)})$ to zero if any two neighboring particle coordinates $x(k), x(k+1)$ become equal,

$$f_n(\dots 1_{x(k)} 1_{x(k)} \dots) = 0, \quad (\text{B6})$$

which ensure that the exclusion principle is satisfied. We note that, since $f_n(C) = 0$, if the number of particles in C is larger than n , there is no term in Eq. (B4) that introduces a new particle from the right reservoir, i.e., $f_n(1_1 \dots 1_{x(n)} 1_{x(n+1)}) = 0$. That is the reason why we can solve Eq. (B4) recursively.

After we have computed all $f_n(1_1 1_{x(2)} \dots 1_{x(n)})$, we find all $f_n(1_2 1_{x(2)} \dots 1_{x(n)})$ using the following recurrence relation:

$$f_n(1_2 1_{x(2)} \dots 1_{x(n)}) = \sum_{k=1}^n \frac{\omega_{x(k)-1}}{e_0} f_n(\dots 1_{x(k)-1} \dots). \quad (\text{B7})$$

We repeat this procedure for all $x(1) = 2, \dots, L-n$, until we have computed all n -particle coefficients $f_n(1_{x(1)} \dots 1_{x(n)})$.

In the next step we consider configurations with $n-1$ particles, starting from the configurations in which particles occupy the first $n-1$ sites. The only difference compared to the previous step is that now we can also move a particle from the right reservoir,

$$\begin{aligned} f_n(1_1 1_{x(2)} \dots 1_{x(n-1)}) &= \frac{1}{e_0} f_{n-1}(1_{x(2)} \dots 1_{x(n-1)}) \\ &+ \sum_{k=2}^{n-1} \frac{\omega_{x(k)-1}}{e_0} f_n(1_1 \dots 1_{x(k)-1} \dots 1_{x(n-1)}) \\ &+ \frac{\omega_L}{e_0} f_n(1_1 \dots 1_{x(n-1)} 1_L). \end{aligned} \quad (\text{B8})$$

The last term contains n particles and has been computed in the previous step. A similar recurrence relation can be written for all other terms with $n-1$ particles. These steps are then repeated for $n-2$ particles, $n-3$ particles, and so on, until we reach configurations with only one particle. In particular, the equation for $f_n(C = 1_1)$ is given by

$$f_n(1_1) = \frac{1}{\omega_1} f_{n-1}(\emptyset) + \frac{\omega_L}{e_0} f_n(1_1 1_L), \quad (\text{B9})$$

in which we set the value of $f_{n-1}(\emptyset)$ to zero, as we discussed before.

Finally, after we have found all nonzero $f_n(C)$, we compute b_n from Eq. (23) and finally $c_n(C)$ from Eq. (B1).

-
- [1] C. T. MacDonald, J. H. Gibbs, and A. C. Pipkin, *Biopolymers* **6**, 1 (1968); C. T. MacDonald and J. H. Gibbs, *ibid.* **7**, 707 (1969).
[2] L. W. Shapiro and D. Zeilberger, *J. Math. Biol.* **15**, 351 (1982).
[3] R. Lipowsky, S. Klumpp, and T. M. Nieuwenhuizen, *Phys. Rev. Lett.* **87**, 108101 (2001).
[4] A. Parmeggiani, T. Franosch, and E. Frey, *Phys. Rev. Lett.* **90**, 086601 (2003).
[5] K. Nagel and M. Schreckenberg, *J. Phys. I France* **2**, 2221 (1992).
[6] *Nonequilibrium Statistical Mechanics in One Dimension*, edited by V. Privman (Cambridge University Press, Cambridge, UK, 1997).
[7] B. Derrida, E. Domany, and D. Mukamel, *J. Stat. Phys.* **69**, 667 (1992).
[8] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, *J. Phys. A: Math. Gen.* **26**, 1493 (1993); G. M. Schütz and E. Domany, *J. Stat. Phys.* **72**, 277 (1993).
[9] J. Krug, *Phys. Rev. Lett.* **67**, 1882 (1991).
[10] S. A. Janowsky and J. L. Lebowitz, *Phys. Rev. A* **45**, 618 (1992).
[11] G. M. Schütz, *J. Stat. Phys.* **71**, 471 (1993).
[12] S. A. Janowsky and J. L. Lebowitz, *J. Stat. Phys.* **77**, 35 (1994).
[13] A. B. Kolomeisky, *J. Phys. A: Math. Gen.* **31**, 1153 (1998).
[14] J. Schmidt, V. Popkov, and A. Schadschneider, *Europhys. Lett.* **110**, 20008 (2015).
[15] B. Scoppola, C. Lancia, and R. Mariani, *J. Stat. Phys.* **161**, 843 (2015).
[16] G. Tripathy and M. Barma, *Phys. Rev. Lett.* **78**, 3039 (1997).
[17] J. Krug, *Braz. J. Phys.* **30**, 97 (2000).
[18] R. J. Harris and R. B. Stinchcombe, *Phys. Rev. E* **70**, 016108 (2004).
[19] P. Greulich and A. Schadschneider, *J. Stat. Mech.* (2008) P04009.
[20] L. B. Shaw, J. P. Sethna, and K. H. Lee, *Phys. Rev. E* **70**, 021901 (2004).
[21] T. Chou and G. Lakatos, *Phys. Rev. Lett.* **93**, 198101 (2004).

- [22] R. Heinrich and T. A. Rapoport, *J. Theor. Biol.* **86**, 279 (1980).
- [23] N. Mitarai, K. Sneppen, and S. Pedersen, *J. Mol. Biol.* **382**, 236 (2008).
- [24] M. C. Romano, M. Thiel, I. Stansfield, and C. Grebogi, *Phys. Rev. Lett.* **102**, 198104 (2009).
- [25] R. K. P. Zia, J. J. Dong, and B. J. Schmittmann, *J. Stat. Phys.* **144**, 405 (2011).
- [26] S. Reuveni, I. Meilijson, M. Kupiec, E. Ruppin, and T. Tuller, *PLoS Comput. Biol.* **7**, e1002127 (2011).
- [27] L. Ciandrini, I. Stansfield, and M. C. Romano, *PLoS Comput. Biol.* **9**, e1002866 (2013).
- [28] H. Zur and T. Tuller, *Nucl. Acids Res.* **44**, 9031 (2016).
- [29] J. Szavits-Nossan, *J. Phys. A: Math. Theor.* **46**, 315001 (2013).
- [30] J. Szavits-Nossan, L. Ciandrini, and M. C. Romano, *Phys. Rev. Lett.* **120**, 128101 (2018).
- [31] G. Kudla, A. W. Murray, D. Tollervey, and J. B. Plotkin, *Science* **324**, 255 (2009).
- [32] H. M. Salis, E. A. Mirsky, and C. A. Voigt, *Nat. Biotechnol.* **27**, 946 (2009).
- [33] K. D. Duc, Z. H. Saleem, and Y. S. Song, *Phys. Rev. E* **97**, 012106 (2018).
- [34] R. Adin and Y. Roichman, *Handbook of Enumerative Combinatorics* (CRC Press, Boca Raton, 2015), p. 895.
- [35] S. Corteel and L. K. Williams, *Adv. Appl. Math.* **39**, 293 (2007).
- [36] J. Schnakenberg, *Rev. Mod. Phys.* **48**, 571 (1976).
- [37] S. Chaiken and D. J. Kleitman, *J. Comb. Theory, Ser. A* **24**, 377 (1978).
- [38] L. Addario-Berry and B. A. Reed, *Horizons of Combinatorics* (Springer, Berlin, 2008), Vol. 17, pp. 9–35.
- [39] W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley & Sons, New York, 1968), Vol. 1, p. 69.
- [40] S. Klumpp, Y. Chai, and R. Lipowsky, *Phys. Rev. E* **78**, 041909 (2008).
- [41] L. Ciandrini, I. Stansfield, and M. C. Romano, *Phys. Rev. E* **81**, 051904 (2010).