Inhomogeneous exclusion processes with extended objects: The effect of defect locations

J. J. Dong,* B. Schmittmann, and R. K. P. Zia

Center for Stochastic Processes in Science and Engineering, Department of Physics, Virginia Tech,
Blacksburg, Virginia 24061-0435, USA

We study the effects of local inhomogeneities, i.e., slow sites of hopping rate $q < 1$, in a totally asymmetric simple exclusion process for particles of size $\ell \geq 1$ (in units of the lattice spacing). We compare the simulation results of $\ell = 1$ and $\ell > 1$ and notice that the existence of local defects has qualitatively similar effects on the steady state. We focus on the stationary current as well as the density profiles. If there is only a single slow site in the system, we observe a significant dependence of the current on the *location* of the slow site for both $\ell = 1$ and $\ell > 1$ cases. When two slow sites are introduced, more intriguing phenomena emerge, e.g., dramatic decreases in the current when the two are close together. In addition, we study the asymptotic behavior when $q \rightarrow 0$. We also explore the associated density profiles and compare our findings to an earlier study using a simple mean-field theory. We then outline the biological significance of these effects.

PACS number(s): 05.70.Ln, 87.15.Aa, 05.40.−a

I. INTRODUCTION

A better understanding of nonequilibrium steady states in interacting complex systems forms a critical goal of much current research in statistical physics. In this pursuit, the totally asymmetric simple exclusion process (TASEP) [1–6] has played a paradigmatic role. It provides a nontrivial, yet exactly solvable, example of phase transitions far from equilibrium, taking place even in one-dimensional (1D) lattices. At the same time, it also serves as the starting point for the modeling of many physical (driven diffusive) processes, such as translation in protein synthesis [7–9], inhomogeneous growth processes (e.g., Kardar-Parisi-Zhang growth) [10,11], and vehicular traffic [12,13].

In its simplest version, the TASEP involves a single species of particles hopping to nearest-neighbor sites, in one direction only, along a homogeneous 1D lattice. Provided the destination site is empty, the rate for the particle hop is fixed at γ (typically chosen as unity without loss of generality). With periodic boundary conditions, the steady-state distribution is trivial [14] but the full dynamics is quite complex [15–17]. With open boundary conditions, particles are injected with rate α (in units of γ) at one end and drained with rate β at the other end. The competition of injection, transport, and drainage induces a nontrivial phase diagram in the α-β plane [1–6], reflecting a highly nontrivial steady state. Three phases are present: a maximum-current phase for $\alpha, \beta > 1/2$, and a low- (high-)density phase for $\alpha < \beta$, $\alpha < 1/2$ ($\beta < \alpha$, $\beta < 1/2$). Not surprisingly, there are also rich dynamical aspects [18,19].

To model protein synthesis, each site on the lattice represents a codon on the messenger RNA (mRNA), and the particles represent the ribosomes. Injection, hopping, and drainage are associated respectively with initiation, elongation, and termination in biological terms. The quantity of interest, namely, the (steady-state) protein production rate, is identical to the (stationary) particle current. Clearly, the simple

TASEP falls short of the biological system in several significant aspects. One is that an individual ribosome "covers" several codons [7,20,21], as opposed to a particle occupying only a single site. Another is that, in all naturally occurring mRNAs, the codons carry genetic information and therefore necessarily form an *inhomogeneous* sequence. Thus the elongation rate of a ribosome is unlikely to be uniform; instead, the hopping rate γ_i of a particle becomes a function of the site i. For example, it is well known that translation slows down at specific codons (see, e.g. [9,22–25]), with potentially significant consequences for protein production rates. Indeed, the steady-state current may depend sensitively on not only the *frequency* of each codon's occurrence, but also the *order* of their appearance in the sequence. Both of these issues—extended objects and inhomogeneous rates—have been addressed recently in separate contexts which we summarize briefly in the following.

The results associated with inhomogeneous (quenched random) rates fall into two broad categories, in the sense that the randomness can be associated with the particles [26–28] or with the sites. Randomness of the former type is more relevant for vehicular traffic where it accounts for a variety of driver preferences. In contrast, the disorder in the protein case is clearly *site dependent*, leading to *spatially* nonuniform hopping rates γ_i. Restricting ourselves to this class, we can consider the effect of having a whole *distribution*, or very *specific* configurations, of $\{\gamma_i\}$. Starting from given distributions, two groups [29,30] studied the resulting disorder in periodic systems. To mention just one significant effect, the current-density diagram develops a plateau: limited by the smallest rate in the system, the current becomes independent of density over a range of densities. Harris and Stinchcombe [30] also extended this work to open systems. While these studies may be of some interest to *mixtures* of many different mRNAs, our primary interest here is to understand how the production rate of a specific protein is associated with a specific genetic sequence. As a first step towards a solution, we adopt the approach of several other studies [9,31–33], by focusing on the effects of a few *localized* inhomogeneities, i.e., hopping rates which are uniform *except* at a handful of sites [34]. As a synthesis of these studies, we

*jjdong@vt.edu

will explore in some detail the consequences of having extended objects and locating one or two slow sites at a variety of positions on the lattice. In this manner, by introducing more and more sites with a range of rates, we hope to understand inhomogeneities in a systematic way, setting the stage for further investigations of the translation process.

A full comprehension of the effects of slow sites on the particle current may have potentially significant applications in biotechnology. While there are 64 distinct codons, proteins are chains composed of just 20 amino acids. So, many different mRNAs (codon sequences) can code for the same particular protein. Moreover, the amino acid is incorporated into the growing chain by an important intermediary, the so-called transfer RNA (tRNA), which carries the complementary anticodon. It turns out that the mapping between codons and tRNAs is also not precisely 1-1. For example, in *E. coli*, the genetic code actually involves 61 sense codons and about 46 tRNAs with associated anticodons [35]. Meanwhile, for a given mRNA sequence, the protein production rate is often modeled in terms of (generally accepted) charged-tRNA (aminoacyl-tRNA, or aa-tRNA) concentrations [24], so that different sequences can result in *different* production rates for the *same* protein. By elucidating how the spatial distribution of defects, especially of bottlenecks, affects translation rates, we can pinpoint those clusters of codons which are likely to have the most significant effect on the production rate of the associated protein. Exploiting the degeneracy in the mapping from mRNA sequence to protein, we can provide guidance as to how a few selected, local modifications of the mRNA can optimize the production rate of a given protein.

Our paper is organized as follows. In Sec. II, we define the model and provide a more detailed description of previous work, concerning exclusion processes with extended objects or spatially inhomogeneous rates. A brief discussion of a previous mean-field analysis is also included. In Sec. III, we present our Monte Carlo results. We focus especially on the implications of having extended objects by varying the particle size. We first consider the interaction between one slow site and the system boundary, and, motivated by the resulting findings, turn to the interactions between *two* slow sites. This provides interesting insights for genes containing clusters of slow codons, which occur frequently in, e.g., *E. coli, Drosophila*, yeast, and primates [9,36,37]. In Sec. IV, a complete investigation of systems with inhomogeneities is presented using a mean-field approach. Section V contains our conclusions and a summary of open questions.

II. MODEL SPECIFICATIONS AND KNOWN RESULTS

The TASEP is defined on a 1D lattice of N sites. We introduce an index $i=1,2,\ldots,N$ to label the sites. Each site (codon) is either occupied by a single particle (ribosome) of length ℓ (in units of sites) or empty. A microscopic configuration of the system can be uniquely characterized in terms of a set of occupation variables $\{n_i\}$, taking the value 1(0) if site i is occupied (empty). Of course, the extended nature of the particle induces strong correlations in $\{n_i\}$, in the sense that a single ribosome always covers ℓ *consecutive* sites. Yet,
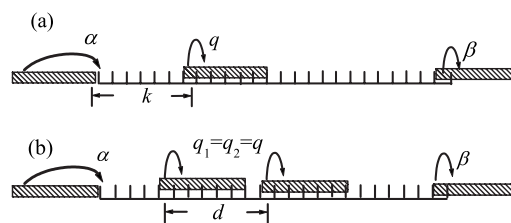

FIG. 1. Sketch of a TASEP for particle size $\ell=6$ with (a) a single slow site at position k, with rate q, and (b) two slow sites with rate q, separated by a distance d.

at any given time, only one of the covered codons is being "read" (i.e., the codon is "covered by the aminoacyl site," or A site, of the ribosome) and translated into an amino acid. Here, we refer to the associated location on the ribosome as the "reader" (of the genetic code). For our purposes, it is not essential which one of the ℓ sites is labeled as the reader, and so we follow the convention in [8] and choose the first (leftmost) site. Hence the statement "a ribosome (or particle) is located at site i" implies that the reader is located at site i and the subsequent

$$\bar{\ell} \equiv \ell - 1 \qquad (1)$$

sites are also occupied. Naturally, the position of the reader determines the elongation rate, i.e., γ_i, since the ribosome must wait for the arrival of the aa-tRNA with the i-specific anticodon before it can move to the next site. Clearly, the reader locations can also be used to label a microscopic configuration, i.e., we can define the reader occupation number at site i as r_i. The sets $\{n_i\}$ and $\{r_i\}$ are uniquely related to each other. Moreover, due to the extended size of a particle, strict *constraints* are built in (e.g., $r_i=1$ implies $r_{i+1}=\cdots=r_{i+\ell-1}=0$ and $n_i=\cdots=n_{i+\ell-1}=1$). As a consequence, neither set can be arbitrary and serious correlations arise as soon as $\ell>1$ [38].

In our simulations, we adopt a random sequential updating scheme and keep a list of locations of readers. In addition, the site $i=0$ is always occupied by a "virtual reader," which accounts for particles entering the system (initiation). At the beginning of each Monte Carlo step (MCS), we first find the number of particles in the system and label it M. Then, we randomly select an entry from this list of $M+1$ readers. If the chosen reader is virtual (i.e., $i=0$), a new particle enters the lattice with probability α, *provided* all the first ℓ sites are empty. If the chosen reader is real, say, at site $i>0$, the associated particle is then moved to site $i+1$ with probability γ_i, *provided* site $i+\ell$ is empty. With this notation, we can also write the initiation and termination probabilities (α and β) as γ_0 and γ_N, respectively. To be complete, the sites beyond the lattice are by definition empty, so that once a particle reaches $N-\ell+1$, it will not experience steric hindrance (see Fig. 1 for a sketch of this process). These processes have been termed "complete entry" and "incremental exit" [39]. Other entry and exit rules can be considered, but are believed to be inconsequential provided $\ell/N \ll 1$. Each MCS consists of $M+1$ such attempts, giving an even chance, on average, for each particle (ribosome) in the system to

elongate or terminate, as well as for an initiation event to occur.

Starting with an empty lattice, we typically discard 2×10^6 MCSs to ensure that the system has reached the steady state. Unless otherwise noted, good statistics result if we average over least 2×10^4 measurements, separated by 100 MCSs in order to avoid temporal correlations. Such steady state averages will be denoted by $\langle \cdots \rangle$. To reduce the number of parameters in the model, we study systems with $\alpha = \beta = \gamma_i = 1$, *except* at one or two sites. The system sizes (N) range from 200 to 1000, with most data taken from $N = 1000$.

To characterize the state of the system, we monitor several observables. The most obvious is $\rho_i^r \equiv \langle r_i \rangle$, a quantity we will refer to as the ribosome (or reader, or particle) density. Of course, $\Sigma_i \rho_i^r$ is just the average number of particles in the system (i.e., ribosomes on the mRNA). Thus the overall particle density $N^{-1} \Sigma_i \rho_i^r$ is bounded above by $1/\ell$. Another interesting variable $\rho_i \equiv \langle n_i \rangle$, labeled as the "coverage density," is the probability that site i is covered by a particle (regardless of the location of the reader). Needless to say, the profile for the vacancies is given by the local hole density, $\rho_i^h \equiv 1 - \rho_i$. The overall coverage density, $N^{-1} \Sigma_i \rho_i$, may reach unity and provides a good indication of how packed the system is. The two profiles are related by

$$\rho_i = \sum_{k=0}^{\ell-1} \rho_{i-k}^r \qquad (2)$$

with the understanding $\rho_i^r \equiv 0$ for $i \leq 0$.

A quantity of great importance to a biological system is the steady-state level of a given protein. If we assume that the degradation rates are (approximately) constant, i.e., independent of protein concentration, then these levels are directly related to the protein production rates. In our model, such a rate is just the average particle current J, defined as the average number of particles exiting the system per unit time. In the steady state, it is also the current measured across any section of the lattice. For simplicity and to ensure the best statistics, we count the total number of particles which enter the lattice over the entire measurement period (at least 2×10^6 MCSs in most cases).

In this study, we focus on two simple types of inhomogeneities: one or two "slow" sites (Fig. 1). Their locations specify the only inhomogeneities in the rates.

One slow site, at position k: We denote γ_k by q (<1). This corresponds to a bottleneck in the lattice. We are especially interested in the dependence of the current, denoted by $J_q(k)$, on the parameters q and k.

Two slow sites, at positions k_1 and k_2 with separation $d \equiv (k_2 - k_1)$ and rates $q_{1,2} \equiv \gamma_{k_{1,2}}$: We find that, when $q_1 \neq q_2$, the current is controlled mainly by the smaller of the two, with little dependence on d, in agreement with a simple mean-field theory to be discussed in Sec. IV. Therefore most of our attention will be devoted to the case with $q_1 = q_2 \equiv q < 1$. Moreover, we choose to limit our study to both sites being far from the boundaries. Then, the current is insensitive to their average position $(k_2 + k_1)/2$, and we can investigate $J_q(d)$. Note that these are precisely the systems studied
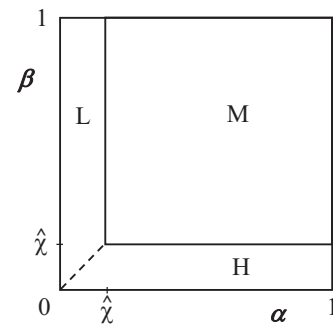

FIG. 2. Phase diagram for an ordinary TASEP. On the dashed line, the H and L phases coexist.

in [9], except that we consider particles with a *range* of sizes: $\ell = 1, 2, 4, 6$, and 12. While there are qualitative similarities, we will discuss the *quantitative* differences due to $\ell > 1$, as well as interesting phenomena associated with the density profiles.

Let us now provide the context of our work by briefly reviewing some related earlier studies. The homogeneous case ($\gamma_1 = \cdots = \gamma_{N-1} = 1$) with $\ell = 1$ is exactly soluble [2–6], and displays three phases in the α-β phase diagram. For $\ell > 1$, no exact solutions exist [40]. Analytic approximations using various mean-field approaches [7,8,20,39,41] predict the presence of the same phases, though the phase boundaries depend on ℓ (Fig. 2) through the combination [8]

$$\hat{\chi} \equiv \frac{1}{1 + \sqrt{\ell}}. \qquad (3)$$

Monte Carlo studies [8,39] largely confirm these conclusions.

The three phases carry different currents and display distinct density profiles [2–6,8,33,39]. Apart from "tails" near the boundaries, the (coverage) density profiles approach uniform bulk values in the thermodynamic limit, i.e., $\rho_i \rightarrow \rho_{bulk}$, for $1 \ll i \ll N$. For $\alpha < \hat{\chi}$ and $\alpha < \beta$, the system is in a low-density phase (L), characterized by $\rho_{bulk} = \ell \alpha / (1 + \alpha \bar{\ell})$ and $J = \alpha(1-\alpha)/(1+\alpha \bar{\ell})$. A high-density phase ($H$) prevails for $\beta < \hat{\chi}$ and $\beta < \alpha$, with bulk density $\rho_{bulk} = 1 - \beta$ and current $J = \beta(1-\beta)/(1+\beta \bar{\ell})$. For $\alpha, \beta > \hat{\chi}$, the system is in a maximum-current phase (M), where $\rho_{bulk} = 1 - \hat{\chi}$ and $J = \hat{\chi}^2$. On the $\alpha = \beta < \hat{\chi}$ line (dashed line in Fig. 2), the system consists of two macroscopic regions, characterized by a low (high) density region near the entry (exit) point. The two regions are joined by a shock front that performs a random walk. This is often referred to as the "shock phase" (S). Table I summarizes the J-ρ_{bulk} relation for TASEP with extended objects.

There is good agreement between simulations (with $\ell \leq 12$) and analytic results for these bulk quantities [8,39]. The details of the profile for $\ell > 1$, especially near the lattice boundaries, are less well understood. While periodic structures (of period ℓ) can be expected, mean-field theories [7,20,39] were successful in capturing only a limited part of the phenomena observed. We will return to these consider-

TABLE I. J-ρ_{bulk} relation for particles of size ℓ ($\bar{\ell} \equiv \ell - 1$).

Phase	Current J	Bulk density ρ_{bulk}
L	$\alpha(1-\alpha)/(1+\alpha\bar{\ell})$	$\ell\alpha/(1+\alpha\bar{\ell})$
H	$\beta(1-\beta)/(1+\beta\bar{\ell})$	$1-\beta$
M	$\hat{\chi}^2$	$1-\hat{\chi}$

TABLE II. Color coding scheme

Size ℓ	Online color
1	black
2	red
4	brown
6	green
12	blue

ations in Sec. IV. Beyond homogeneous systems, several studies introduced one or more "impurities" into TASEP with periodic boundary condition. A single "slow" site induces a shock in the density profile with some interesting statistics [42–45,47]. Subsequently, generalizations to systems with a finite fraction of slow sites, randomly located, were also investigated [29]. For the richer case of the open boundary TASEP [9,31–33,48], Kolomeisky focused on point particles ($\ell = 1$), with a *single* impurity at the *center* of the lattice [31], so as to mimic a defect situated deep in an infinitely long system. The consequences of the defect having both faster ($q > 1$) and slower ($q < 1$) rates were explored. By matching two ordinary TASEPs across the defect, the properties of such systems in the α-β plane can be well described [31,48]. While a fast site has no effect on the phase diagram, a slow site leads to a shift of the M-H and M-L phase boundaries to q-dependent, smaller values of α and β. The density profiles are quite sensitive to the existence of a defect site [9,31]. Kolomeisky's approach was generalized to the $\ell = 12$ case in [33], with similar levels of success. Below, we will provide further details of this work, on which we base much of the analysis of our problem. Ha and den Nijs also studied the $\ell = 1$ open boundary TASEP with a single defect at the center [32]. Focusing on the multicritical point $\alpha = \beta = 1/2$, they were mainly interested in the so-called "queueing transition" and its critical properties. Detailed results of density profiles, such as power law behavior and critical exponents, were obtained in the region $q \cong q_c$. Here, q_c denotes the critical value of q below which the bulk density in front of the slow site deviates from the density behind the blockage. By contrast, our focus here is essentially that of [9,48], namely, how does the *number* and the *locations* or *spacings* of the slow sites affect the current through the system? In the single slow site case [9], investigated mainly the overall current as a function of q, while [48] analyzed the current as a function of k (the distance between the defect and the entry point). For the case of two defects, both studies find that the spacing between them plays a significant role for the current. A finite-segment mean-field theory in [9] provides excellent agreement with data. In particular, clustered defects reduce the current much more effectively than well separated ones. In this sense, we can regard these investigations as exploring the "interactions" between the slow site(s) and/or the boundaries. Since both studies are restricted to point particles ($\ell = 1$), our intent is to explore the effects of having extended objects (with $\ell \leq 12$). Though we expect *qualitatively* similar behavior, as pointed out in [9], we also find noteworthy *quantitative* differences. Finally, we should mention that larger numbers of defects do not lead to significantly different effects [9], so that we limit ourselves to one or two slow sites here.

III. MONTE CARLO RESULTS

In this section, we present our Monte Carlo results. For convenience, we use a consistent color coding scheme (online only) for the various particle sizes, as specified in Table II.

The data here consist of the overall currents and the density profiles of both coverage and ribosomes. Our focus will be how these quantities depend on q, k (for the case with a single defect), and d (for the case with two defects). Although the profiles are difficult to extract experimentally, the reader profiles will be of interest in subsequent studies involving real gene sequences, since they provide information on how frequently a particular tRNA is bound to the mRNA. By contrast, the currents are easily measurable and our results here may generate more immediate interest.

A. One slow site

We begin by placing one slow site on the lattice as in Fig. 1(a). Figure 3 shows several coverage density profiles for a typical choice of parameters: $N = 1000$, $q = 0.2$, and $k = 82$ with $\ell = 1, 6, 12$. As expected, we observe pile-ups of particles due to the blockage—a high (low) density region before (after) the bottleneck in all three cases. However, due to the lack of ordinary particle-hole symmetry in the $\ell > 1$ cases, the average densities on either side of the slow site are not symmetric around 0.5. Instead, they are *roughly* related through the J-ρ_{bulk} relations in the H and L phases, summa-
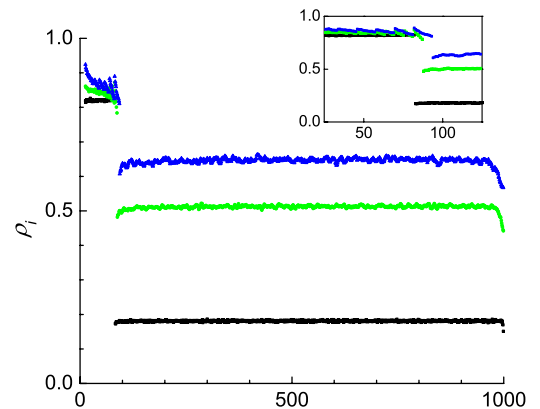

FIG. 3. (Color online) Coverage density profiles with one slow site of $q = 0.2$ at $k = 82$. $\ell = 1, 6, 12$ (from bottom to top in both subsections) and $N = 1000$. The inset is a magnified view of the $i \in [1, 150]$ interval, to expose the period ℓ structures.
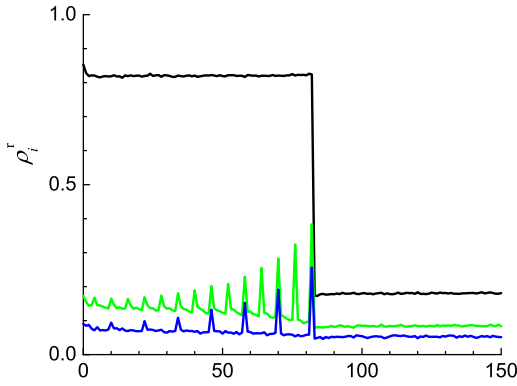
FIG. 4. (Color online) Ribosome density profiles with one slow site of $q=0.2$ at $k=82$. $\ell=1,6,12$ (from top to bottom in both subsections) and $N=1000$. Only the first 150 lattice sites are shown.

rized in Table I. In detail, the profiles are quite different: The "tails," i.e., the deviations from the bulk values, are quite noticeable in the vicinity of both the slow site and the edges of the system for the $\ell=6,12$ cases. The inset exposes more clearly that there are period ℓ structures in the profiles [7,20,39], especially just before the slow site.

A more dramatic difference between point particles and extended objects emerges when we plot the ribosome density ρ_i^r, in Fig. 4, corresponding to the inset in Fig. 3. Similar to profiles in [7,20,39], we find distinct period ℓ structures before the slow site. While the reader "waits" to pass the blockage, the readers of the following particles tend to catch up and pause at sites $k-n\ell$, where $n=1,2,\ldots$ The tails are even more marked than those in Fig. 3. To emphasize the difference between the reader and coverage profiles (ρ_i^r and ρ_i), we show a case with $q=0.05$, $k=948$, $\ell=12$ in Fig. 5. Though both profiles contain the same information, we see that ρ_i^r (lower plot) is far more sensitive than ρ_i (upper plot) in showing the very long tails (~ 1000 in this example) hidden in the collective behavior of the particles. At present, the crucial ingredients that control the characteristic decay length of the ρ_i^r envelopes have not yet been identified.
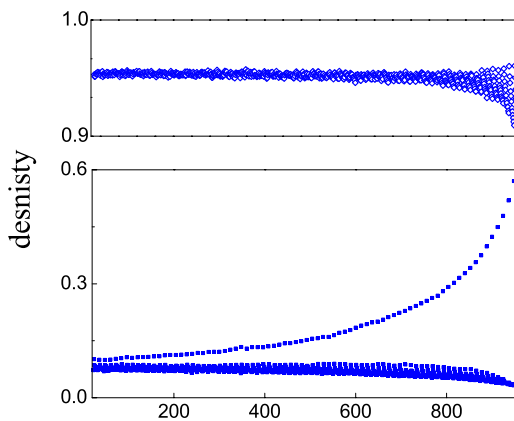

FIG. 5. (Color online) Coverage density profile (top) and ribosome density profile (bottom) with one slow site of $q=0.05$ at $k=948$. $\ell=12$ and $N=1000$.
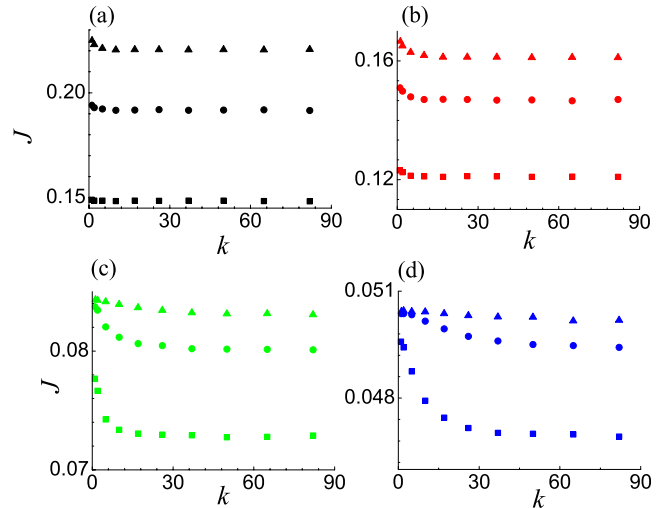

FIG. 6. (Color online) $J_q(k)$ as a function of the location k of the slow site for $q=0.2$ (lower set of squares), 0.3 (middle circles), and 0.4 (upper triangles). (a) $\ell=1$; (b) $\ell=2$; (c) $\ell=6$; (d) $\ell=12$. In all cases, $N=1000$.

Certainly, these very large length scales are completely absent from the $\ell=1$ systems deep within the H/L phases.

As for the current, Fig. 6 illustrates its dependence on q, k, and ℓ. Not surprisingly, the current is limited by the bottleneck and therefore varies monotonically with q. It is also reduced if the particle size increases, an effect that can be traced mainly to the particle density being effectively lower by the factor ℓ. For point particles, the current is not very sensitive to the location of the slow site. The enhancement as k approaches the boundary of the system—referred to as the "edge effect" [48]—is quite small. For larger ℓ, the enhancement is much more pronounced, especially for smaller q. Whatever the magnitude, in all cases the current increases monotonically as the slow site is located closer and closer to the entry point. For $\ell=1$, particle-hole symmetry is manifest in the microscopic dynamics, so that the symmetry of $J_q(k)$ under $k\to N+1-k$ inversion, is obvious [48]. For $\ell>1$, the density profiles confirm the *lack* of this particle-hole symmetry very clearly. Correspondingly, there is a systematic asymmetry in the current: $J_q(k)=J_q(N+1-k)$ is satisfied only for $k\lesssim\ell$. The origin of this behavior is not well understood.

The edge effect, and specifically its dependence on q and ℓ, can be quantified by the ratio

$$\Delta_1(q) = \frac{J_q(k=1)}{J_q(k\to\infty)}. \tag{4}$$

Figure 7 shows that $\Delta_1(q)$ depends on q in a nontrivial way. The maxima of $\Delta_1(q)$ occur at lower values of q as ℓ increases, reminiscent of the behavior of the phase boundary between M and L/H. With appropriate scaling, the curves of $\Delta_1(q)$ can be collapsed for large ℓ's. From the biological perspective, the edge effect is not easily observable since the current enhancement is less than 10% for the relevant ℓ.

Returning to Fig. 6, we note that significant deviations from the asymptotic value, $J_q(\infty)$, are found only for k's
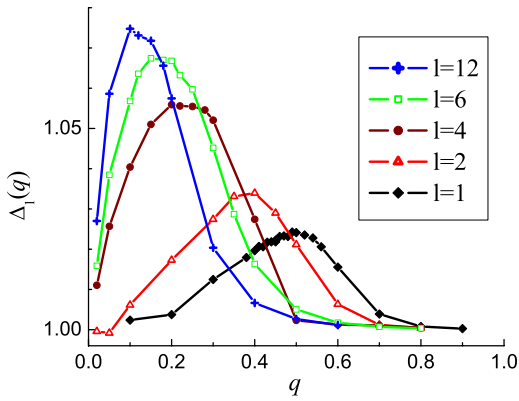
FIG. 7. (Color online) $\Delta_1(q)$ for $\ell = 1, 2, 4, 6$ and 12.

within a small distance from the boundaries. At a casual glance, this range appears to depend on both q and ℓ. On closer examination of, say, the most prominent case here, $(q, \ell) = (0.2, 12)$, we find that the decay of $J_q(k)$ into $J_q(\infty)$ fits an exponential quite well (Fig. 8), i.e., $J_q(k) - J_q(\infty) \propto \exp(-k/\delta)$, with $\delta \approx 10$. Assuming this behavior persists in the other cases, we can study the (q, ℓ) dependence of this characteristic length and denote it by $\delta(q, \ell)$. We believe that the origin of this length scale can be traced to the presence of tails in the density profiles near the lattice boundaries: As the slow site approaches the entry point, these tails may seriously affect the injection process, and thus the current as well. For the $\ell = 1$ case, we observe that $\delta(q, \ell) \sim \xi(\beta_{eff})$, where ξ is the characteristic length associated with the boundary layer of the *density profile* in the ordinary TASEP. Specifically, with entry rate $\alpha = 1$ and exit rate β, the system is in the H-phase and the profile decays exponentially into the bulk, as $\rho_x - \rho_{bulk} \propto \exp(-x/\xi)$ with [6]:

$$\xi(\beta) = -\frac{1}{\ln[4\beta(1 - \beta)]}. \tag{5}$$

Here, the left half of our system is such a TASEP, except that we have an effective β: $\beta_{eff} = q/(1+q)$ [33]. Using these ar-

guments on the three q's shown, we estimate decay lengths of about 5($q = 0.4$), 3($q = 0.3$), and 2($q = 0.2$) lattice constants. Though the data on the differences $J_q(k) - J_q(\infty)$ are small and noisy, simulation results are consistent with $\delta(q, 1) \sim \xi(\beta_{eff})$. However, for $\ell > 1$, there is no analytic result for the boundary layers of the density profiles. Moreover, the data suggest that they are quite complex (e.g., in Figs. 4 and 5). Thus it is unclear how to quantify the picture for point particles to the general case of $\delta(q, \ell)$. At present, a complete understanding of both "boundary layers"—in the density profiles and in $J_q(k)$—remains elusive.

If we consider the edge effect as an "interaction" between the slow site and the lattice boundaries, the natural next step is to explore the interactions between *two* slow sites. In order to avoid edge effects, we place the two slow sites sufficiently far away from the boundaries and vary their separation.

B. Two slow sites

As mentioned in the previous section, the currents in the $q_1 \neq q_2$ cases are essentially controlled by the slower of the two rates and so may be regarded as systems with a single slow site. These profiles can be interesting, but we choose to restrict our attention here to a study of the $q_1 = q_2 \equiv q$ case, in which the currents show a nontrivial dependence on d, the distance between the two slow sites. With two bottlenecks, the system consists of three sections: before the first blockage, in between the two, and after the second defect. Of course, for small q, the overall density in the first (last) section is expected to be high (low). In these cases, the effective entry and exit rates for the central section are also low, so that a wandering shock should be present. Hence the *average* profile should be linear for $\ell = 1$ (and essentially so for larger ℓ [8]) with a *positive* slope. This behavior is understandable, since the section between the two defects is comparable to an ordinary TASEP with small $\alpha = \beta$. These expectations are generally confirmed by simulations with $q \lesssim 0.5$ and various ℓ's up to 12. Figure 9 shows typical coverage profiles, for a relatively small rate of $q = 0.2$. The system appears to make a transition from this $H/S/L$ phase to an $M/M/M$ phase as q increases. The center profiles become essentially flat, as illustrated in the inset (where $q = 0.6$ and $\ell = 12$). Details of this transition are being explored.

More interesting are the finer features of the profiles in the small q cases. As in the single defect system, the profiles exhibit period ℓ structures near the slow sites. To resolve these more clearly, we plot the *reader* density profiles in Fig. 10. In all cases that involve extended particles ($\ell > 1$), the readers clearly pile up behind the slow sites. Apart from these "jams," another feature emerges, namely, a sequence of *depletion* zones, each of which precedes one of the period ℓ peaks. For $\ell = 2$, the differences between the upper and the lower envelope are especially dramatic. More remarkably, when the blockages are separated by small d's, two different "overlapping tails" are created, as illustrated in the inset of Fig. 10, where $d = 1$, $q = 0.2$, and $\ell = 12$. Indeed, there are further interesting structures for $d \lesssim \ell$, which will be presented elsewhere.
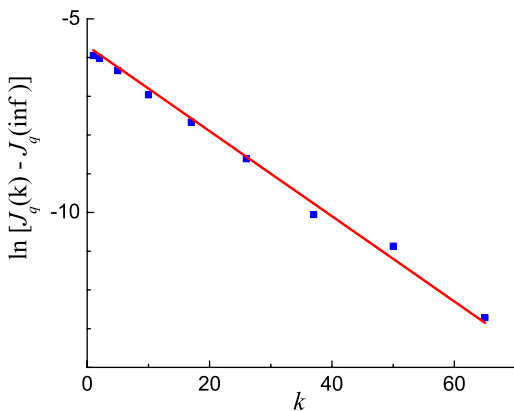

FIG. 8. (Color online) Dependence of $J_q(k)$ on k obtained from simulation is plotted in squares and the line is a linear fit with slope equal to -0.11. $q = 0.2$, $\ell = 12$, and $N = 1000$.
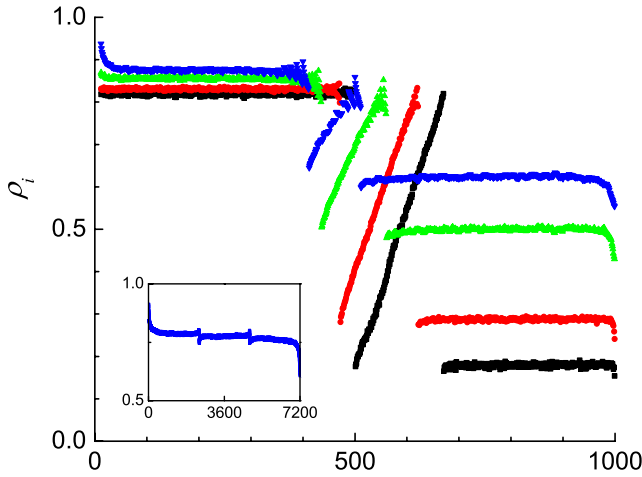
FIG. 9. (Color online) Coverage density profiles for two slow sites with $q=0.2$. $\ell=12$, $d=100$; $\ell=6$, $d=125$; $\ell=2$, $d=150$; and $\ell=1$, $d=170$ (curves are plotted from left to right in the midsubsection and top to bottom elsewhere). In all cases, $N=1000$. Inset: $q=0.6$, $\ell=12$, $d=2400$, and $N=7200$.

Compared to these remarkable characteristics in the profiles, the behavior of the currents seems lackluster. In Fig. 11, we plot four sets of currents [50], $J_q(d)$, associated with $\ell=1$, 2, 6, and 12. In all cases, we see that J is considerably suppressed when d is reduced. When the slow sites are very far apart, the current behaves as if there is only one slow site, consistent with expectations from mean-field theories. At the other extreme, when the two defect sites are nearest neighbors, the current reaches its minimum. Not surprisingly, period ℓ structures emerge as d is varied, illustrated in the inset of Fig. 11(d), but become less prominent for $d \gtrsim 50$. These plots also reveal that, unlike the dependence on k above, there are serious deviations from the $d \to \infty$ values when d is decreased. To quantify this deviation, we define

FIG. 10. (Color online) Ribosome density profiles with two slow sites of $q=0.2$. $\ell=12$, $d=100$; $\ell=6$, $d=125$; $\ell=2$, $d=150$; and $\ell=1$, $d=170$ (curves are plotted from left to right in the midsubsection and bottom to top elsewhere). Inset, $\ell=2$ and $d=1$. In all cases, $N=1000$.
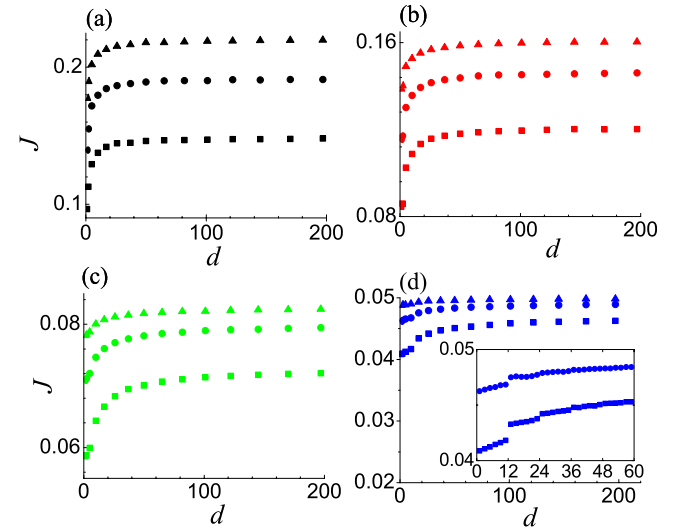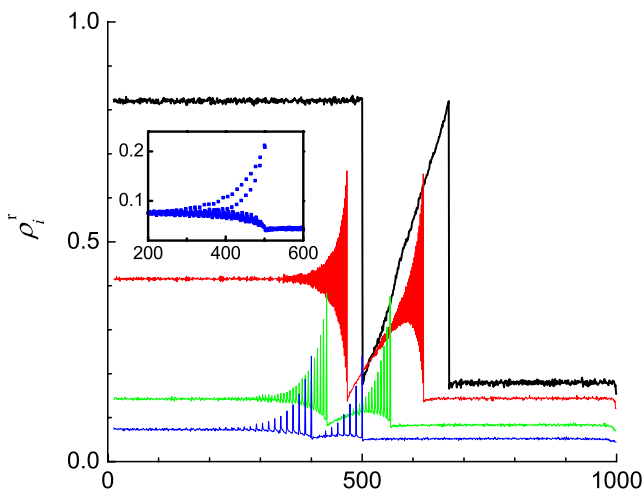

FIG. 11. (Color online) $J_q(d)$ as a function of the separation d between the two slow sites for $q=0.2$ (lower set of squares), 0.3 (middle circles), and 0.4 (upper triangles). (a) $\ell=1$; (b) $\ell=2$; (c) $\ell=6$; (d) $\ell=12$. The inset in (d) is a magnified view of the $d \in [1,60]$ interval, to expose the period ℓ structures. In all cases, $N=1000$.

$$\Delta_2(q) = \frac{J_q(d=1)}{J_q(d \to \infty)} \tag{6}$$

and plot this quantity vs q in Fig. 12. In contrast to $\Delta_1(q)$, we observe that $\Delta_2(q)$ exhibits a *sizable* dependence on q, especially for small values of q. In the limit of $q \to 0$ the current decreases by a factor of 2! In the following section, we will see that this factor can be understood via a mean-field approach.

To summarize our simulation results, two bottlenecks near each other have a dramatic effect on the current. We may regard this phenomenon as an "interaction" between the two slow sites, inducing far more "resistance" when they are close than when they are well separated.

Two additional comments are in order. First, we return to one of the predictions of the mean-field theory, namely that a second slow site, spaced far apart from its partner, should
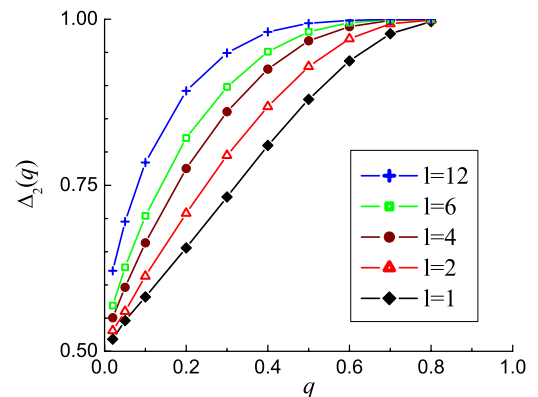

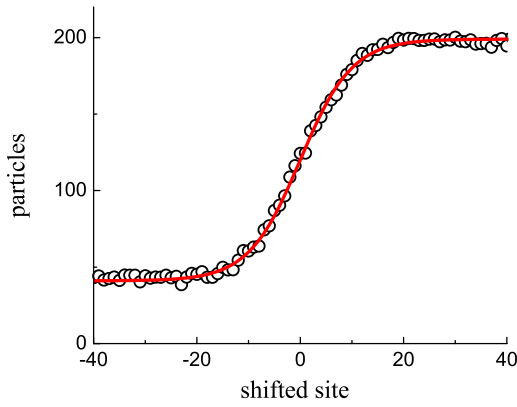FIG. 12. (Color online) $\Delta_2(q)$ for $\ell=1$, 2, 4, 6, and 12.

FIG. 13. (Color online) The open circles mark the average profile of a shock between slow sites 149 and 349 with $q=0.2$, compiled from of a very long run (3×10^8 MCSs) with an $N=1000$, $\ell =1$ system. Only the [210, 290] segment is shown and is scaled to center at 0. Details of how raw profiles are shifted (so that the shock is located at site $x=0$ shown here) will be published elsewhere [49]. A simple fit using $A+B\tanh(x/10)$ is also shown (solid line).

have no further effect on the current. Our data indicate that the current for two slow sites, spaced far apart, is systematically *lower* than the current for a single slow site, but only by a very small amount (less than 1%). Second, we can again attempt to identify a length scale which controls how $J_q(d)$ approaches $J_q(\infty)$, as d increases. Since the central section of the system displays a shock, it is natural to ask whether the intrinsic width of the shock sets this length scale. According to [42,45], this width covers only a few lattice spacings in the *periodic* TASEP with a single defect. Here, however, it appears that the shock is much broader. For example, the averaged profile of the shock for the case of $q=0.2$ with point particles is shown in Fig. 13, as well as a simple fit using a tanh function [46] with width of about 10. Intriguingly, this length appears to be comparable to the one appearing in Fig. 11(a). More work is needed to fully explore these issues.

IV. MEAN-FIELD THEORETIC APPROACHES

Mean-field theory is known to agree well with the exact results for a number of macroscopic quantities in the steady state of the $\ell=1$ TASEP (see, for example, [6]). For extended particles, no exact solution is available [40] so that mean-field (and more sophisticated cluster) approximations form the only route toward some understanding of the system's behavior. However, there are *many levels* of "mean-field" approximations [7,8,33,39], corresponding to neglecting different types of correlations. For certain quantities (e.g., currents in large systems), predictions from the simplest level are very close to the simulation results. For others (e.g., some reader profiles), only the most sophisticated level performs adequately. In all cases, no level of mean-field theory can give a good fit to both the current and the profile. A thorough discussion of all of these schemes is quite involved and will be provided elsewhere [49]. Here, we will restrict

ourselves to the simplest method and compare its performance to the simulations.

All approaches start with the exact expressions for the current,

$$J = \alpha\langle 1 - n_\ell \rangle \tag{7}$$

$$= \gamma_i \langle r_i(1 - n_{i+\ell}) \rangle; \quad i \in [1, N-\ell] \tag{8}$$

$$= \gamma_i \langle r_i \rangle; \quad i \in [N-\bar{\ell}, N-1] \tag{9}$$

$$= \beta \langle r_N \rangle. \tag{10}$$

In the absence of the steady-state distribution, the most naive approximation is to replace $\langle r_i n_j \rangle$ by $\langle r_i \rangle \langle n_j \rangle$. Unfortunately, the constraints due to particles with $\ell > 1$ are so severe that this approximation is entirely inadequate when $j=i+\ell$. Even for the simple case of TASEP on a periodic ring, it leads to an erroneous expression for J (except if $\ell=1$). Instead, the average (coverage) density at site $i+\ell$ is much larger than the (conditional) probability that it is actually covered *given* that the reader is at site i. MacDonald and Gibbs (MG) proposed [7] a much better approximation:

$$\langle r_i(1 - n_{i+\ell}) \rangle \simeq \frac{\rho_i^r(1 - \rho_{i+\ell})}{1 - \rho_{i+\ell} + \rho_{i+\ell}^r} = \frac{\rho_i^r \rho_{i+\ell}^h}{\rho_{i+\ell}^r + \rho_{i+\ell}^h}.$$

As discussed in Sec. II, the densities far from boundaries are uniform (e.g., $\rho_{i \to \infty}^r \to \rho_{bulk}/\ell$) and this fact provides a good description of the current-density relation

$$J(\rho_{bulk}) = \frac{\rho_{bulk}(1 - \rho_{bulk})}{\ell - \bar{\ell}\rho_{bulk}} \tag{11}$$

(for $\gamma=1$). Exploiting this relation and regarding our model as two or three TASEPs joined by slow sites, the simplest level of mean-field theories can be built. Ours is similar to, but simpler than, the approach in [33] for the single defect case. The main difference lies in the matching condition, i.e., what approximate expression for the current across the slow site to use. After comparing the two approaches, we proceed to build the case for TASEP with two defects.

A. One slow site

When a single slow site ($q<1$) is located at k, the system can be treated as two sublattices: $[1, k]$ and $[k+1, N]$, referred to as the left and right sublattices, respectively. Associated quantities will appear with subscripts L and R. The two sections are coupled through the slow site by having the same current in the steady state. Given this constraint, there are only two viable scenarios for the sublattices, out of the 3×3 logically possible ones: H/L and M/M.

First, let us consider the H/L case which was one studied extensively in [33]. The current for each sublattice can be written as

$$J_L = \frac{\beta_L(1 - \beta_L)}{1 + \beta_L \bar{\ell}}, \quad J_R = \frac{\alpha_R(1 - \alpha_R)}{1 + \alpha_R \bar{\ell}}, \quad (12)$$

where β_L and α_R are the effective exit and entry rates, to be determined later. By definition, the entire system reaches steady state when $J_L = J_R$, which yields $\beta_L = \alpha_R$. Of course, these are intimately related to the (bulk) densities through $\rho_L = 1 - \beta_L$ and $\rho_R = \ell \alpha_R/(1 + \alpha_R \bar{\ell})$, so that

$$\frac{(1 - \rho_L)}{1 + (1 - \rho_L)\bar{\ell}} = \frac{\rho_R}{\ell}.$$

Another way to regard this relation is that both densities lead to the same current, which we denote by J (a value to be determined, and equal to $J_L = J_R$). So, the high and low densities can be written as $\rho_+(J)$ and $\rho_-(J)$, respectively, being the two roots to Eq. (11). They will play a crucial role when we impose the matching condition, thereby fixing all quantities as a function of q.

The exact equation for "matching" is

$$J = q\langle r_k(1 - n_{k+\ell})\rangle, \quad (13)$$

in which r_k and $n_{k+\ell}$ lie in L and R, respectively. Now, the right can be expressed as, again exactly, $p(k|k+\ell)$, the probability for finding a ribosome at k, *conditioned* on the presence of a hole at $k+\ell$.

Since we have "broken" the system into two separate TASEPs, a naive approximation is to begin with

$$J_{NMF} = q\langle r_k\rangle\langle(1 - n_{k+\ell})\rangle,$$

where the subscript stands for "naive mean field." Regarding this as Eq. (7) for the R sublattice, we have $\alpha_R = q\rho_k^{\rm r}$. Now, $\rho_k^{\rm r}$ is in the L sublattice, and must be related to ρ_L in a mean-field approach. The most naive assumption is that $\rho_k^{\rm r}$ is the same as its average in the bulk, i.e., $\rho_{bulk}^{\rm r}$, which would be ρ_L/ℓ in this case. However, this turns out to underestimate $p(k|k+\ell)$ seriously. Indeed, the "pile-up" near a blockage (e.g., in Fig. 5) shows that $\rho_k^{\rm r}$ is significantly higher than its bulk value as well as the densities on the $\bar{\ell}$ sites before. Thus we propose that a better approximation would be to replace $\rho_k^{\rm r}$ by ρ_L, and we write

$$\alpha_R = q\rho_L. \quad (14)$$

Using $\rho_L = 1 - \beta_L$ and $\beta_L = \alpha_R$, so that $\alpha_R = \beta_L = q/(1+q)$ and

$$\rho_L = 1/(1 + q), \quad \rho_R = q\ell/(1 + q\ell),$$

we arrive at $J_{NMF} = q/[(1+q)(1+q\ell)]$. The premise behind this line of arguments is that the system is in H/L, so that both α_R and β_L should be less than $\hat{\chi}$. Therefore this expression for the current should be valid only if it is less than the maximal value ($\hat{\chi}^2$). In other words, the domain of its validity is limited to $q \lesssim 1/\sqrt{\ell}$. For higher q, this approach predicts that the system will be in an M/M phase, with maximal current. Note that such a phase cannot occur with a slow defect in the $\ell = 1$ case, where M/M can be accessed only with $q > 1$. In an earlier study [33], the parameters chosen

($q = 0.2$ and $\ell = 12$) also precluded the presence of this phase, although we believe (see below) that this phase cannot be present if the blockage is in the center ($k = N/2$) or deep in the bulk. We summarize this naive mean field by

$$J_{NMF} = \begin{cases} q/[(1 + q)(1 + q\ell)] & \text{for } q \lesssim 1/\sqrt{\ell} \\ \hat{\chi}^2 & \text{for } q \gtrsim 1/\sqrt{\ell} \end{cases}. \quad (15)$$

An alternative approximation for Eq. (13) was proposed earlier [33]:

$$J \cong q_{eff}\left(\frac{\rho_L}{\ell}\right)\left(\frac{1 - \rho_R}{1 - \rho_R \bar{\ell}/\ell}\right). \quad (16)$$

The last two factors can be recognized as $\langle r_k\rangle$ and the MG approximation for the effective hole density [7]. The first factor is a little more subtle [33]: Considering that the transit time for a single particle through the slow site (in the absence of steric hindrance) is $q^{-1} + \bar{\ell}$, q_{eff} is defined as the average rate to move just one step in this process:

$$q_{eff} \equiv \frac{q\ell}{1 + q\bar{\ell}}.$$

The end result for the current is the solution to the algebraic equation

$$J = q_{eff}\frac{\rho_+(J)[1 - \rho_-(J)]}{\ell - \rho_-(J)\bar{\ell}}.$$

Here, we give an explicit form (which displays the $\ell = 1$ limit well)

$$J_{SKL} = \frac{Q}{(1 - Q + \sqrt{1 - 2Q})\bar{\ell}}, \quad (17)$$

where

$$Q \equiv \frac{2q\bar{\ell}(1 + q\bar{\ell})}{(1 + q + 2q\bar{\ell})^2}.$$

Note that, for any $q < 1$, this approach predicts that the current is less than the maximal value of $\hat{\chi}^2$ and so the system is always in the H/L phase.

The results of both mean-field predictions for J as a function of q are shown in Fig. 14, along with two sets of data: $J_q(1)$ and $J_q(363)$. As expected, J_{SKL} was purpose-built for two infinite TASEP's connected by a slow site and provides a better fit to the data with the blockage deep in the bulk ($k = 363$ in $N = 1000$). On the other hand, it is understandable that, e.g., for $k = 1$, α_R must be very close to q. Thus we may expect that the system will have maximal current for $q \gtrsim 1/\sqrt{\ell}$. This behavior is confirmed by the data, as illustrated in the figure for $\ell = 12$. Since $J_{NMF}(q)$ has the property that it saturates at $\hat{\chi}^2$ for $q > \hat{\chi}$, it provides a better fit for $J_q(1)$. Of course, we recognize that, as mean-field theories, neither Eqs. (14) nor (16) are the first step in a systematic expansion, so that they may better be thought of as "semiphenomenological."
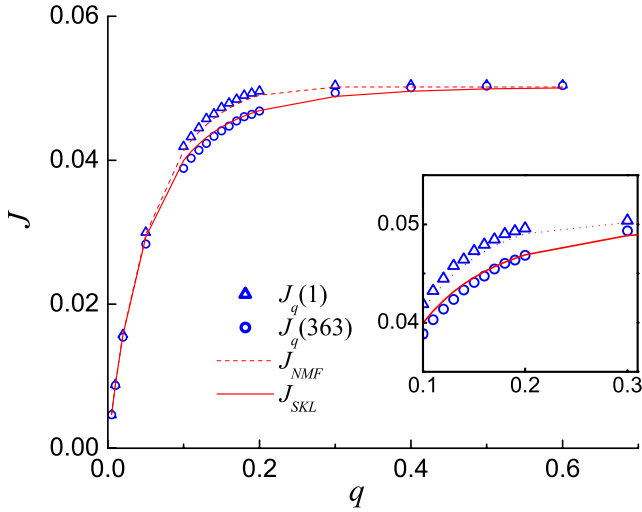
FIG. 14. (Color online) Comparisons of the current J as a function of q. The legend labels the two sets of simulation data (slow site at $k=1$ and 363) and predictions from two mean-field approximations.

More importantly, there is a more serious, inherent limitation in this level of mean-field theory. It cannot account for the full k dependence in $J_q(k)$, since it deals only with infinite systems. One possibility to incorporate finite-size effects (of the left sublattice) in this kind of theory is to exploit the MG expression [7]

$$J_{MG} = \frac{\rho_i^r \rho_{i+\ell}^h}{\rho_{i+\ell}^r + \rho_{i+\ell}^h}$$

for sites near the boundaries. Using it in a recursion relation for finding both the density profile and the current, we will arrive at a k-dependent expression, $J_{MG}(\alpha=1, \beta_L; k)$, which replaces J_L in the first expression of Eq. (12). However, the high density fixed point of this recursion relation is unstable and very careful numerical analysis will be necessary. Preliminary work on this approach is promising and will be reported elsewhere [49].

We end this subsection by noting that the effects of a single defect in TASEP have also been investigated in [32]. Unlike our focus here, the dependence of J on the *location* of the slow site, they are concerned with a "multicritical system," i.e., $\alpha = \beta = 1/2$ for the $\ell = 1$ case. Putting the defect at the center of the lattice, they explored density profiles in detail, finding power law tails on both sides of the defect with q-dependent exponents. By contrast, our choice of $\alpha = \beta = 1$ places us far from the multicritical point. We have no reason to expect similar power laws.

B. Two slow sites

The most general TASEP with just two slow sites can be quite involved, since the parameter space is four-dimensional: $\{q_1, q_2, k_1, k_2\}$. To carry out a manageable investigation, we let both sites be deep in the bulk, so that only the distance between them, $d \equiv (k_2 - k_1)$, plays a significant

role. Further, as pointed out above, the central section resembles an ordinary TASEP with α and β controlled by q_1 and q_2, respectively. Therefore it is the smaller (slower) of the two rates which limits that current, which in turn dictates the current through the whole system. Thus we will focus only on the $q_1 = q_2 = q$ case. Our parameter space will then *resemble* the single slow site case.

Following the single defect case, the simplest levels of mean-field theory treat our system as three subsections with obvious labels: L, C, and R. From our discussion, only two (out of the many logical possibilities) combinations of phases, $H/S/L$ and $M/M/M$, are expected to be viable. In addition to Eq. (12), we have

$$J_C = \frac{\alpha_C(1 - \alpha_C)}{1 + \alpha_C \bar{\ell}}. \tag{18}$$

Since the defect rates are identical, we fully expect that, for such a mean-field theory, $\beta_C = \alpha_C$. Now, matching the currents of the subsections, we immediately arrive at $J_L = J_C = J_R$ and so, $\beta_L = \alpha_C = \beta_C = \alpha_R$. From here, the "naive" mean-field approach for $H/S/L$ proceeds identically to the above. The argument relies on the presence of a shock in the central section, so that there is a low (high) density region near site $k_1 + 1 (k_2)$ and we can impose the same discontinuity in the densities across both defects, i.e., $\rho_+(J)$ before and $\rho_-(J)$ after. Thus we again arrive at $J_{NMF}(q)$, given explicitly in Eq. (15). The same argument can be applied to the next level of a mean-field approximation, which predicts $J_{SKL}(q)$, as in Eq. (17). The major difference between the two approaches, as in the single slow site case, is the absence of the $M/M/M$ phase in the latter. Meanwhile, their limitations are similar: The d dependence in $J_q(d)$ cannot be accommodated without serious modifications.

Nevertheless, the spirit of these approximations can be exploited to provide $J_q(1)$ in the $q \to 0$ limit. Since the central section consists of just one site, there can be no shock. Instead, the remnant of the shock is reflected in the average density there. It is more convenient to regard the system as two infinite TASEP's, with nontrivial matching across a "doubly slow site." Of course, we cannot expect to find any of the fascinating profile details (e.g., inset of Fig. 10), but we should be able to obtain the "coarser" information, such as the currents. The goal is to understand the behavior of $\Delta_2(q \to 0)$ (cf. Fig. 12) in, say, the first two nonvanishing orders in q.

Now, for $q \ll 1$, we are naturally in the H/L phase and the crudest approximation should suffice for the lowest order in the current. So, we let the bulk densities be at their extremes (i.e., 1 and 0) and simply consider the time it takes for a particle to move through the blockage, from the moment its predecessor is "released." The current is just the inverse of this quantity, i.e.,

$$\left[\frac{2}{q} + \ell - 2 \right]^{-1} \to \frac{q}{2} \left[1 - \frac{q}{2}(\ell - 2) + \cdots \right], \tag{19}$$

where we have included $O(q^2)$ terms for computing the next order. But, at this order, we should also take into account that, occasionally, the density before or after the blockage

deviates from unity or zero by virtue of the right hand side of Eq. (11) being nonzero. Thus these densities are

$$\rho_L \rightarrow 1 - J = 1 - q/2 + \cdots ,$$

$$\rho_R \rightarrow J\ell \approx q\ell/2 + \cdots$$

and further suppress the current at the next-to-lowest order through the factor

$$\rho_L(1 - \rho_R) \rightarrow 1 - \frac{q}{2}(\ell + 1) + \cdots . \qquad (20)$$

Combining these factors, we arrive at

$$J_{q\rightarrow 0}(d = 1) \rightarrow \frac{q}{2}\left[1 - q\left(\ell - \frac{1}{2}\right) + \cdots\right].$$

If we use exactly the same arguments for the $q \rightarrow 0$ limit current in the case of one slow site, we find, instead of Eq. (19),

$$\left[\frac{1}{q} + \ell - 1\right]^{-1} \rightarrow q[1 - q(\ell - 1) + \cdots], \qquad (21)$$

and, instead of Eq. (20),

$$\rho_L(1 - \rho_R) \rightarrow 1 - q(\ell + 1) + \cdots .$$

Finally, since $J_q(d \rightarrow \infty)$ is the same as the single-blockage current, we write

$$J_{q\rightarrow 0}(d \rightarrow \infty) \rightarrow q[1 - 2q\ell + \cdots] \qquad (22)$$

so that

$$\Delta_2(q \rightarrow 0) \rightarrow \frac{1}{2} + \frac{q}{2}\left(\ell + \frac{1}{2}\right) + \cdots .$$

It is remarkable how well this crude approximation agrees with the data in Fig. 12. There is no doubt that all curves extrapolate to the ℓ-independent value of $1/2$ at $q=0$. As for the slope at the origin, we can obtain a good estimate from the lowest q data points, using $[\Delta_2(q=0.02)-0.5]/0.02$. The values obtained from simulations for $\ell=1$, 2, 4, 6, and 12 are 0.92, 1.55, 2.53, 3.44, and 6.06, respectively.

We are aware that the expansion (22) differs from the small q limit of J_{NMF}. Unfortunately, it is difficult to implement the same scheme for J_{NMF} here, since we must start from the exact *pair* of equations:

$$J = q\langle r_k(1 - n_{k+\ell})\rangle = q\langle r_{k+1}(1 - n_{k+\ell+1})\rangle. \qquad (23)$$

Various attempts at approximating ρ_k^r or ρ_{k+1}^r led to poorer results. Alternatively, we could exploit the argument in Shaw *et al.* [33] and consider the average time to traverse both slow sites, $2/q+(\ell-2)$. This gives us a new effective q:

$$\tilde{q}_{eff} \equiv \frac{q\ell}{2 + q(\ell - 2)}$$

which can be inserted into Eq. (16). The result is $\Delta_2(q \rightarrow 0) \rightarrow \frac{1}{2} + \frac{q}{4}(\ell+2)+\cdots$, the $O(q)$ term of which differs from the data by about a factor of 2. Clearly, mean-field approaches are far from ideal for finding quantitative predic-

tions of $J_q(d)$. On the other hand, either $J_{NMF}(q)$ or $J_{SKL}(q)$ provide tolerable results when the blockages are from from each other or the boundaries. Such variations in the quality of mean-field theories point to the importance of correlations. Considerable efforts appear to be necessary for a comprehensive, yet relatively simple, theory.

Let us end this section with another method which could possibly improve the theoretical predictions [51], especially for the first few values of k. The idea is to find exact results, by solving the full master equation, for TASEPs (with extended objects) on very small lattices and then to match these to an infinite system (the R sublattice). One expectation is that, as in the $\ell=1$ case, the finite-size current is larger than its counterpart for the infinite TASEP at the same (α, β). This approach may eventually provide the essential argument to understand the increase in $J_q(k)$ as k becomes smaller. Similarly, this idea can be applied to the case with two slow sites when d is $O(1)$. Work is in progress to explore these avenues.

V. SUMMARY AND CONCLUSIONS

In this study, we consider an inhomogeneous TASEP with open boundaries and populated with particles of finite extent ℓ. The hopping rates are uniform (set at unity) *except* for one or two sites ("defect bonds"), where the rates q are different (faster or slower). We are interested in the effects of these local defects on the density profiles and the currents through the system. Simulations with various $\ell \leq 12$ show that fast sites have no effect on the current, but induce discontinuities in the density profiles. In contrast, slow sites generate a significant reduction of the current as well as nontrivial structures in the profiles, e.g., long tails behind the blockage, with period ℓ. These findings are entirely consistent with similar studies in the past, most of which were restricted to $\ell=1$ [31,32]. If the inhomogeneities are deep in the bulk and far from each other, the current depends only on q and can be understood through simple mean-field considerations. Through the current-density relationship (for an infinite homogeneous TASEP), the overall densities in each of the defect-free sections can be also predicted, so that the various "phases" of these subsystems can be understood. The distinguishing feature in our study is how the *location* of the defects affects the behavior of the system. For the case of one slow site, the current is slightly but measurably enhanced when the defect approaches the boundary. On the other hand, a drastic reduction of the current is observed when two slow sites are brought closer to each other. It is tempting to interpret these effects as "interactions" between the defects and to seek a formulation that can describe them quantitatively.

At present, neither the enhancement nor the suppression of the currents can be understood in terms of simple mean-field theories. The essential limitation is that they are based on matching homogeneous TASEPs of *infinite* length. More sophisticated versions, relying on recursion relations for the particle density at each site, may be exploited to deal with the finite subsections and provide some promise for a better understanding of such effects. For very small subsections, such as $k, d \leq 5$, it may be possible to find exact solutions

(even for $\ell > 1$) that can be used to match the mean-field descriptions for the macroscopic subsection(s). Work is in progress to investigate these approaches systematically.

Beyond one or two blockages, we should study systems with multiple slow sites, as our eventual goal is to understand the properties of fully inhomogeneous TASEPs. In particular, though restricted to just one or two slow sites, our findings—that the relative locations of blockages are important—will have implications for translation. For example, they are directly applicable to "designer genes," which consist of many repeats of the same codon, except at one or two locations. Using the abundance of associated aatRNAs as a control for the elongation rate across any particular codon, we can test our results directly on such genes. Thus it will be interesting to see the physical manifestation of, e.g., enhancement and suppression of production rates of such artificial proteins, depending on the placement of the slow codon(s). Similarly, reproducing the intriguing ribosome density profiles will be revealing. More important than "designer genes," we should consider the implications for real genes. Our results should provide, at the least, some simple qualitative insights. We can obviously maximize the production rate of a particular protein associated with a certain real gene by systematically replacing all slow codons

with synonymous, faster ones. However, for most genes this operation will require a large number of replacements. Instead, with our findings in mind, we can achieve considerable increases in the production rates by making only a few substitutions, namely, by replacing the slowest codons, or a cluster of nearby slow codons. The ratio of current enhancement to the number of codon replacements may be used to quantify how "optimal" a certain set of substitutions is. This idea can be applied to finding optimal means to *suppress* prodcution rates as well. Simulation work with TASEPs associated with real genes is in progress and we hope to demonstrate that these concepts are viable.

ACKNOWLEDGMENTS

[1] J. Krug, Phys. Rev. Lett. **67**, 1882 (1991).

[2] B. Derrida, E. Domany, and D. Mukamel, J. Stat. Phys. **69**, 667 (1992).

[3] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, J. Phys. A **26**, 1493 (1993).

[4] G. M. Schütz and E. Domany, J. Stat. Phys. **72**, 277 (1993).

[5] B. Derrida, Phys. Rep. **301**, 65 (1998).

[6] G. M. Schütz, in *Phase Transition and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic Press, San Diego, 2000), Vol. 19.

[7] C. MacDonald, J. Gibbs, and A. Pipkin, Biopolymers **6**, 1 (1968); C. MacDonald and J. Gibbs, *ibid.* **7**, 707 (1969).

[8] L. B. Shaw, R. K. P. Zia, and K. H. Lee, Phys. Rev. E **68**, 021910 (2003). This article also contains a review of earlier work.

[9] T. Chou and G. Lakatos, Phys. Rev. Lett. **93**, 198101 (2004).

[10] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).

[11] D. E. Wolf and L.-H. Tang, Phys. Rev. Lett. **65**, 1591 (1990).

[12] D. Chowdhury, L. Santen, and A. Schadschneider, Curr. Sci. **77**, 411 (1999).

[13] V. Popkov, L. Santen, A. Schadschneider, and G. M. Schütz, J. Phys. A **34**, L45 (2001).

[14] F. Spitzer, Adv. Math. **5**, 246 (1970).

[15] L. H. Gwa and H. Spohn, Phys. Rev. Lett. **68**, 725 1992; Phys. Rev. A **46**, 844 (1992).

[16] D. Kim, Phys. Rev. E **52**, 3512 (1995).

[17] S. Gupta, S. N. Majumdar, C. Godrèche, and M. Barma, Phys. Rev. E **76**, 021112 (2007).

[18] J. de Gier and F. H. L. Essler, Phys. Rev. Lett. **95**, 240601 (2005).

[19] D. A. Adams, R. K. P. Zia, and B. Schmittmann, Phys. Rev. Lett. **99**, 020601 (2007).

[20] R. Heinrich and T. Rapoport, J. Theor. Biol. **86**, 279 (1980).

[21] C. Kang and C. Cantor, J. Mol. Struct. **181**, 241 (1985).

[22] M. Robinson, R. Lilley, S. Little, J. S. Emtage, G. Yarranton, P. Stephens, A. Millican, M. Eaton, and G. Humphreys, Nucleic Acids Res. **12**, 6663 (1984).

[23] M. A. Sorensen, C. G. Kurland, and S. Pedersen, J. Mol. Biol. **207**, 365 (1989).

[24] J. Solomovici, T. Lesnik, and C. Reiss, J. Theor. Biol. **185**, 511 (1997).

[25] C. M. Stenström, H. Jin, L. L. Major, W. P. Tate, and L. A. Isaksson, Gene **263**, 273 (2001).

[26] Z. Csahók and T. Vicsek, J. Phys. A **27**, L591 (1994).

[27] J. Krug and P. A. Ferrari, J. Phys. A **29**, L465 (1996).

[28] J. Krug, Braz. J. Phys. **30**, 97 (2000).

[29] G. Tripathy and M. Barma, Phys. Rev. E **58**, 1911 (1998).

[30] R. J. Harris and R. B. Stinchcombe, Phys. Rev. E **70**, 016108 (2004).

[31] A. B. Kolomeisky, J. Phys. A **31**, 1153 (1998).

[32] M. Ha, J. Timonen, and M. den Nijs, Phys. Rev. E **68**, 056122 (2003). For more details, see also M. Ha, PhD thesis, University of Washington (Seattle, WA), 2003.

[33] L. B. Shaw, A. B. Kolomeisky, and K. H. Lee, J. Phys. A **37**, 2105 (2004).

[34] In much of the physics literature, the term "bond" is used instead of "site," since hopping is associated with a particle jump from site i to site $i+1$. However, in translation, a ribosome "at site i" can move to the next site only when the tRNA associated with site i arrives. Therefore it is natural to associate the jump rate with a site, and so we will use terms like

"slow site" and "slow bond" interchangeably. With protein synthesis in mind, we also use the phrase "slow codon."

[35] F. Neidhardt and H. Umbarger, in *Escherichia coli and Salmonella*, 2nd ed., edited by F. C. Neidhardt (ASM, Washington D.C., 1996).

[36] D. A. Phoenix and E. Korotkov, FEMS Microbiol. Lett. **155**, 63 (1997).

[37] S. Zhang, E. Goldman, and G. Zubay, J. Theor. Biol. **170**, 339 (1994).

[38] For TASEP on a ring, the total number of particles and holes are both conserved. If the hopping rates are *uniform*, it is possible to specify microscopic configurations in such a way that *no* such correlations are explicitly present, namely, the set of integers $\{h_k\}$, where h_k denotes the number of holes between the kth and $(k+1)$th particles. See, e.g., [29]. However, this mapping is quite impractical for open TASEPs, since the number of particles (or holes) is a fluctuating quantity.

[39] G. Lakatos and T. Chou, J. Phys. A **36**, 2027 (2003).

[40] Exact solutions for ASEP with extended objects have been obtained, but only for a *closed* system with *periodic* boundary conditions. See, e.g., F. C. Alcaraz and M. J. Lazo, Braz. J.

Phys. **33**, 533 (2003). Our main interest, motivated by translation in protein synthesis, is *open* systems, with nontrivial entry and exit rates (our α/β) and phases.

[41] G. Schonherr and G. M. Schutz, J. Phys. A **37**, 8215 (2004).

[42] S. A. Janowsky and J. L. Lebowitz, Phys. Rev. A **45**, 618 (1992).

[43] G. M. Schütz, J. Stat. Phys. **71**, 471 (1993).

[44] B. Derrida, S. A. Janowsky, J. L. Lebowitz, and E. R. Speer, J. Stat. Phys. **73**, 813 (1993).

[45] S. A. Janowsky and J. L. Lebowitz, J. Stat. Phys. **77**, 35 (1994).

[46] Since the shock diffuses throughout the region between the slow sites, a relatively nontrivial method must be used to compile averages. Details will be published elsewhere [49].

[47] K. Mallick, J. Phys. A **29**, 5375 (1996).

[48] J. J. Dong, B. Schmittmann, and R. K. P. Zia, J. Stat. Phys. **128**, 21 (2007).

[49] J. J. Dong, B. Schmittmann, and R. K. P. Zia (unpublished).

[50] Note that the argument in J_q now refers to the *distance* between the two slow sites.

[51] We thank T. Chou for suggesting the approach used in [9].