

## Dependence of the transportation time on the sequence in which particles with different hopping probabilities enter a lattice

Hiroki Yamamoto,<sup>1,\*</sup> Daichi Yanagisawa,<sup>2,3</sup> and Katsuhiro Nishinari<sup>2,3</sup>

<sup>1</sup>*School of Medicine, Hirosaki University, 5 Zaifu-cho Hirosaki city, Aomori, 036-8562, Japan*

<sup>2</sup>*Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan*

<sup>3</sup>*Department of Aeronautics and Astronautics, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*



(Received 1 March 2019; revised manuscript received 10 July 2019; published 7 October 2019)

Smooth transportation has drawn the attention of many researchers and practitioners in several fields. In the present study, we propose a modified model of a totally asymmetric simple exclusion process (TASEP), which includes multispecies of particles and takes into account the sequence in which the particles enter a lattice. We investigate the dependence of the transportation time on this “entering sequence” and show that, for a given collection of particles, group sequence in some cases minimizes the transportation time better than a random sequence. We also introduce the “sorting cost” necessary to transform a random sequence into a group sequence and show that when this is included a random sequence can become advantageous in some conditions. We obtain these results not only from numerical simulations but also by theoretical analyses that generalize the simulation results for some special cases.

DOI: [10.1103/PhysRevE.100.042106](https://doi.org/10.1103/PhysRevE.100.042106)

### I. INTRODUCTION

Transportation systems are key topics in social or biological systems [1]. In social systems, researchers have sought to obtain smooth transportation in various situations, such as production flow [2,3], vehicular traffic [4–7], and pedestrian evacuation [8–11]. On the other hand, for biological systems, intracellular transportation along microtubules has been vigorously investigated [12–14].

Among various transportation models, the asymmetric simple exclusion process (ASEP), pioneered by MacDonald and Gibbs [15,16], has attracted much attention. It is a stochastic process on a one-dimensional lattice in which particles move asymmetrically. A derivative of ASEP, in which particles are allowed to hop unidirectionally (left to right in the present study) is called a totally asymmetric simple exclusion process (TASEP). In the field of nonequilibrium statistical mechanics, researchers have applied TASEP to various transportation problems, such as molecular-motor traffic [17–20], vehicular traffic [5–7,21–23], and the exclusive-queuing process [24–26], especially since the TASEP with open boundary conditions has been solved exactly [27–29].

In practice, researchers struggle to achieve smooth operation for various tasks, smooth logistics for various products, and an effective evacuation method for pedestrians in various situations, such as exit plans from sports stadiums and concert venues. To attain smooth flow in such situations, we often consider the sequence in which we perform tasks and pedestrians move because this sequence may affect the total performance of the systems. For example, slow pedestrians may block fast ones at the back of a narrow street, which worsens pedestrian

flow. To investigate how the above-mentioned sequences affect pedestrian flow in various systems, herein, we propose a modified TASEP comprising a finite number of multispecies particles, in which the entering sequences of the particles are considered.

In the proposed model, we consider the number of particles to be finite and study the transportation times of those particles. Note that we do not consider the steady state of the system itself. Minemura *et al.* [30] investigated the transportation time for a hopping probability that depends on the lot size, using the single-species TASEP with a finite number of particles. Other related works [31–34] also adopted a finite number of particles. In those models, however, particles circulated through a system comprising a lattice and a particle pool while the input or output rate was varied.

Additionally, the concept of multiple particles has already been extensively studied [35–58]. For example, second-class particles were introduced in Refs. [35–45] and more than two-species particles were introduced in Refs. [46–55]. However, most of these studies focused on mathematically exact solutions to the systems under consideration by using matrix product Ansatz and did not much consider the application of the studied model to real-world situations. Furthermore, owing to their simplicity, periodic-boundary conditions have been adopted in many studies [35,39,47,49–51,53,54,56–58]. Studies on multispecies ASEP with open boundaries and random updating were undertaken only recently [37,40,41,44,45,52,55]; for example, Ref. [55] obtained the exact phase diagram for a multispecies (more than two-species) ASEP. The present investigation primarily focuses on the problem of minimizing the transportation time, adopting open-boundary conditions and parallel updating. With the same boundary conditions and updating rules as the present study, Ref. [59] adopted particles with disorder,

\*h18m1140@hirosaki-u.ac.jp

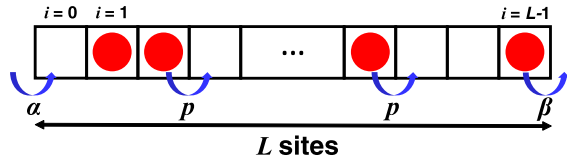


FIG. 1. Schematic illustration of the original TASEP with open-boundary conditions.

whereas jumping particles were introduced in Ref. [60]. Note that majority related works considering multispecies particles assume that swapping between different types of particles, i.e., bidirectional particle hopping, can occur, whereas our model prohibits swapping [61].

In contrast, to the best of our knowledge, no TASEP investigations that focus on the entering sequence of the particles (the key highlight of our model) have been reported thus far. Herein we have considered two special types of sequences in particular: “random sequences” and “group sequences,” and we have compared the transportation times for these two types of sequences. In association with the entering sequence, we have introduced the sorting cost in our model. Without sorting, particles are typically transported at random, i.e., in a random sequence. Therefore, considering the cost of sorting particles from a random sequence into a group sequence is useful. In the present study, we define this sorting cost and compare the results obtained with and without sorting.

We have determined the dependence of the transportation time on the entering sequence of the particles from numerical simulations based on our model. Moreover, we find that the optimal sequence can vary, depending on choice of parameter set, when the sorting cost is considered. In addition, we have succeeded in obtaining mathematical proofs of the simulation results for some special cases.

The remainder of the present study is organized as follows. Section II describes the details of our proposed model and some important parameters, modifying the original TASEP. In Sec. III, we present and discuss the results of numerical simulations using the modified TASEP. Section IV presents theoretical analyses of the simulation results for some special cases. The paper concludes in Sec. V.

## II. MODEL DESCRIPTION

### A. Original (single-species) TASEP with open-boundary conditions

The original TASEP with open-boundary conditions is defined as a one-dimensional lattice of  $L$  sites, labeled from left to right  $i = 0, 1, \dots, L - 1$  (see Fig. 1). Each site can be either empty or occupied by a single particle. In the present study, we adopt discrete time steps and parallel updating. In parallel updating, the states of all the particles on the lattice are determined simultaneously in the next time step. Notably, we can use random updating, which is usually adopted in the ASEP; however, we intentionally adopt parallel updating in the present study (see the specific reasons in Ref. [62]). Particles enter the lattice from the left boundary with probability  $\alpha$  and leave the lattice from the right boundary with probability  $\beta$ . In the bulk of the lattice, if the right-neighboring site is empty, a particle hops to that site with probability  $p$ ;

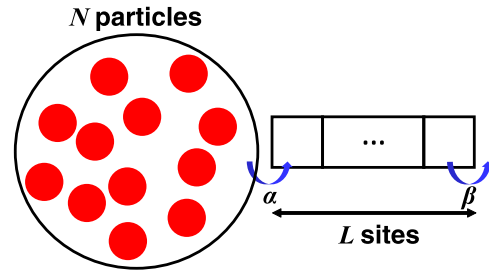


FIG. 2. Schematic illustration of TASEP with a finite number of particles. This figure shows the case  $N = 12$ .

otherwise, it remains at its present site. Our modified TASEP differs from this original one in the following four ways.

### B. Difference 1: Finite number of particles

First, the number of particles  $N$  is finite, as illustrated in Fig. 2. The system evolves until the  $N$ th particle leaves the lattice. We define the transportation time  $T$  as the time gap between the start of the simulation and the time when the  $N$ th particle leaves the lattice.

### C. Difference 2: Multispecies particles

Second, our model adopts multispecies particles, i.e., particles with different hopping probabilities, as illustrated in Fig. 3. Specifically, each of the  $N$  particles is allocated to one of  $S$  species, where  $1 \leq S \leq N$ . Particles that belong to each species  $s$  ( $s = 1, 2, \dots, S$ ) all have the same hopping probability  $p = p_s$  ( $0 < p_s \leq 1$ ). Note that with  $S = 1$  our model reduces to the single-species TASEP, whereas with  $S = N$  all particles have different hopping probabilities. The fraction of all the  $N$  particles allocated to each species  $s$  is defined as  $r_s$ , obviously satisfying  $\sum_{s=1}^S r_s = 1$ .

### D. Difference 3: Consideration of entering sequence of particles

Third, we consider the sequence in which the particles enter the lattice (i.e., the “entering sequence”), which is the most important feature in our model. Specifically, particles form a queue before the left boundary and enter the lattice according to the sequence, as illustrated in Fig. 4. In the present study, we investigate two types of sequences: “random sequences” and “group sequences,” as illustrated in Fig. 5.

In a random sequence, particles line up randomly regardless of their hopping probabilities. A random sequence thus has  $N! / \prod_{s=1}^S (r_s N)!$  patterns. Note that in real situations

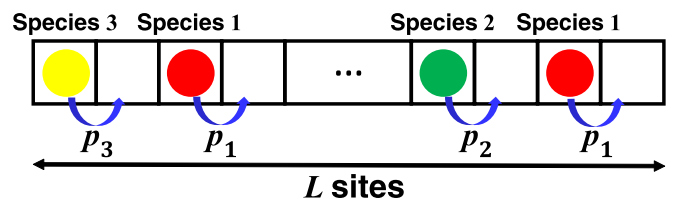


FIG. 3. Schematic illustration of TASEP with multispecies particles. In this figure, we show a case with  $S = 3$ , where the red particles belong to species 1, green ones to species 2, and the yellow ones to species 3.

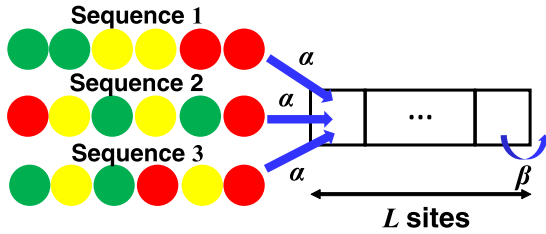


FIG. 4. Schematic illustration of the entering sequences of particles. In this figure, we show three examples among all 90  $\{= 6!/(2!2!2!)\}$  possible sequences for the case  $N = 6$ ,  $S = 3$ , and  $r_1 = r_2 = r_3 = 1/3$ . Note that Sequence 1 is one example of a group sequence, whereas the others are examples of random sequences.

without any controls, random sequences can be assumed to occur spontaneously.

On the other hand, in a group sequence, particles form groups of the same species and line up group by group. There are  $S!$  possible patterns of group sequences, which are clearly among the random sequences.

For the case  $S = N$ , where all hopping probabilities are different, we bunch the particles with similar hopping probabilities close together with each other as much as possible, imaginarily considering them as “continuous groups.” Consistent with this idea, we define a group sequence with  $S = N$  as either an ascending or a descending sequence. Note that we define such a sequence by considering the rightmost particle to be the first particle in the sequence.

We define the transportation times for the random and group sequences to be  $T_R$  and  $T_G$ , respectively.

**E. Difference 4: Introduction of the sorting cost**

Finally, we introduce the cost of sorting the particles and investigate the effect of the sorting cost on the transportation

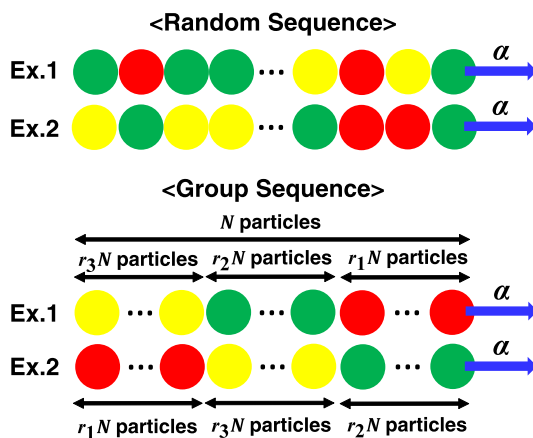


FIG. 5. Schematic illustration of random (upper panel) and group (lower panel) sequences for  $S = 3$ , where the red particles belong to species 1, the green ones to species 2, and the yellow ones to species 3. In the upper panel, we show two possible examples out of all  $N!/\prod_{s=1}^3(r_s N)!$  possible random sequences, whereas in the lower panel, we display two examples of all 6  $(=3!)$  possible group sequences. Note that in each case  $r_1 + r_2 + r_3 = 1$ .

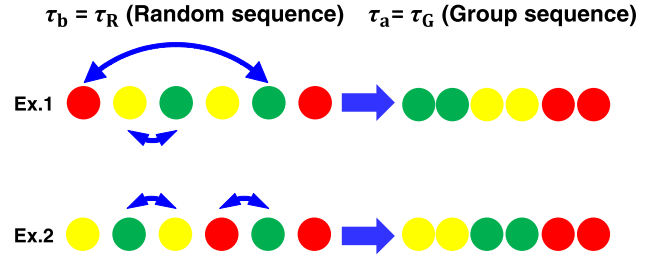


FIG. 6. Two examples with  $K(\tau_a, \tau_b) = 2$  for the case  $N = 6$  and  $S = 3$ . For each sequence  $\tau_R$ , we chose the one of all 6  $(=3!)$  possible sequences  $\tau_G$  so that  $K(\tau_a, \tau_b)$  is minimized.

time. Here we define the sorting cost as the minimal number of exchanges  $K(\tau_a, \tau_b)$  necessary to sort the particles from sequence  $\tau_b$  to sequence  $\tau_a$ , where  $\tau_a$  and  $\tau_b$  represent the sequence after sorting and before sorting, respectively. Note that the arguments of  $K(\tau_a, \tau_b)$  will be abbreviated in obvious cases.

In the present study,  $\tau_a$  ( $\tau_b$ ) correspond to  $\tau_G$  ( $\tau_R$ ), where  $\tau_R$  and  $\tau_G$  represent a random sequence and a group sequence, respectively. The sequence  $\tau_G$  can differ depending on  $\tau_R$ ; that is,  $\tau_G$  is determined so that the number of exchanges is minimized for each  $\tau_R$ . Figure 6 shows two examples for which  $K(\tau_a, \tau_b) = 2$  when  $N = 6$  and  $S = 3$ . Note that we do not consider the distance between the exchanged particles.

We define the number of time steps necessary to sort the particles to be  $\lambda K$ , where the parameter  $\lambda$  is the ratio of the sorting cost to number of TASEP time steps.

**III. SIMULATION RESULTS**

In this section, we use numerical simulations to investigate the dependence of the transportation time on the entering sequence of the particles.

In all the simulations below, we set  $L = 200$  and  $N = 10\,000$ ; we validate this selection of  $L$  and  $N$  in Appendix A. We determine the value of  $T$  for each parameter and average  $T$  over 100 trials for Fig. 7 and over 10 trials for Figs. 8–10.

**A. Without sorting cost ( $\lambda = 0$ )**

In this subsection, we set  $\lambda = 0$ , i.e., we do not include the sorting cost.

In Fig. 7 we plot the simulation values of the number of particles that have not yet exited the lattice at time  $t$  for  $S \in \{2, 3, N\}$ . We fix  $(\alpha, \beta) = (1, 1)$  for Figs. 7(a)–7(c) and  $(\alpha, \beta) = (0.1, 0.2)$  for Figs. 7(d)–7(f). In the figures, we refer to the number of particles that have not yet exited the lattice at time  $t$  simply as the “remaining particles.” The simulation starts at  $t = 0$ , and the number of particles becomes 0, i.e., the  $N$ th particle exits the lattice, at  $t = T$ .

We note two important phenomena in Figs. 7(a)–7(c). First, surprisingly,  $T_G$  is smaller than  $T_R$  for all three values of  $S$  when  $\alpha = \beta = 1$ . This result implies that the group sequences yield smoother transportation than the random ones for the cases  $(\alpha, \beta) = (1, 1)$ . Second,  $T_G$  seems not to depend on the order of each group in the group sequence, which can take  $S!$  possible patterns.

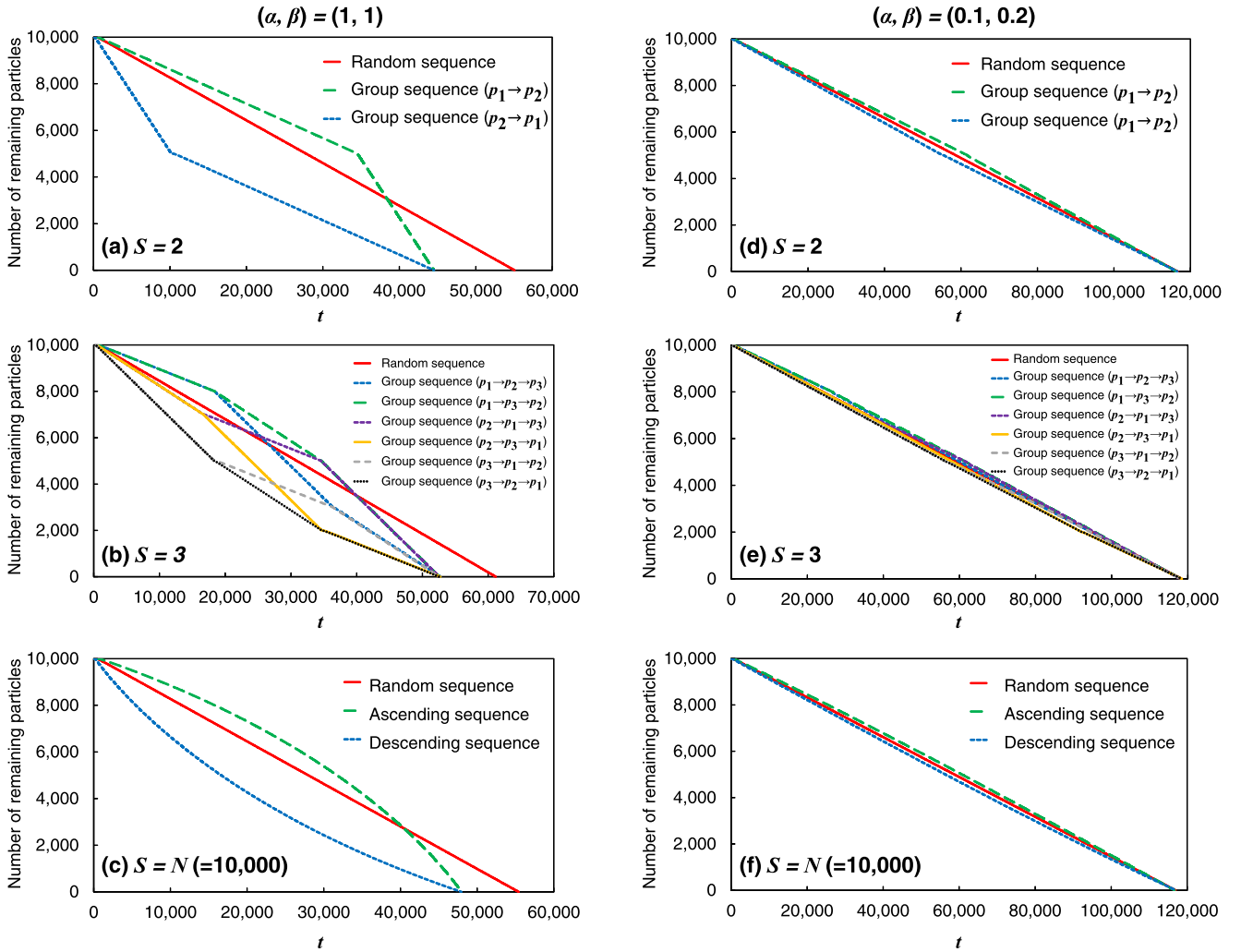


FIG. 7. Simulation values of the number of particles remaining at time  $t$  with  $\lambda = 0$  for (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  with  $(\alpha, \beta) = (1, 1)$  and for (d)  $S = 2$ , (e)  $S = 3$ , and (f)  $S = N$  with  $(\alpha, \beta) = (0.1, 0.2)$ . For  $S = 2$ ,  $S = 3$ , and  $S = N$ , we set  $(p_1, p_2; r) = (0.5, 1; 0.5)$ ,  $(p_1, p_2, p_3; r_1, r_2, r_3) = (0.4, 0.6, 0.8; 0.2, 0.3, 0.5)$ , and  $p_s = 1 - 0.6(N - s)/(N - 1)$  ( $s = 1, 2, \dots, N$ ), respectively, fixing  $\lambda = 0$ . The notation “ $p_s \rightarrow p_t$ ” means that a group of species  $s$  is followed by a group of species  $t$ .

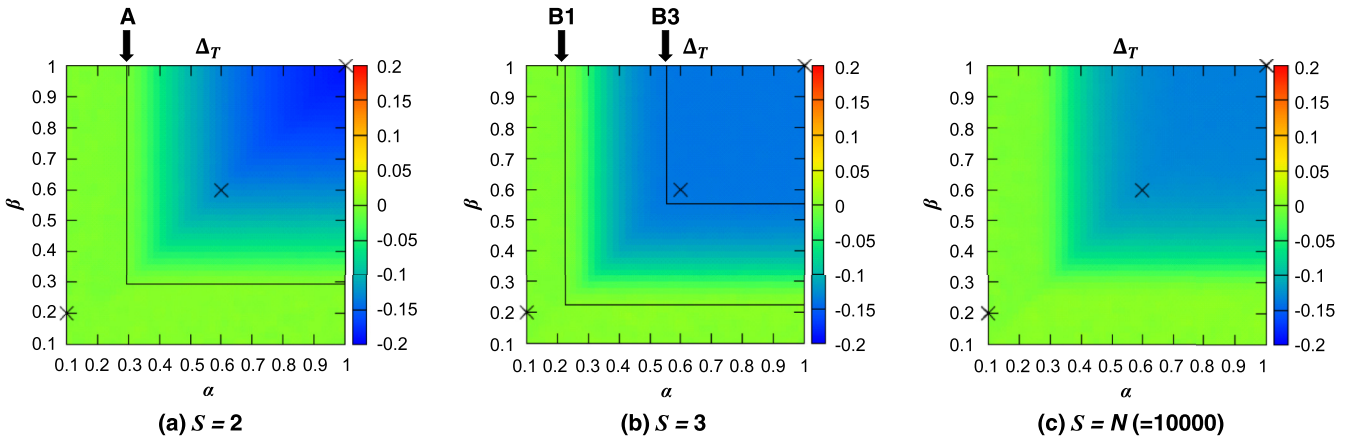


FIG. 8. Simulation values of  $\Delta_T$  for various  $(\alpha, \beta)$  with (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$ . The parameters other than  $(\alpha, \beta)$  are the same as in Fig. 7. Note that three black crosses in each panel represent  $(\alpha, \beta) = (0.1, 0.2)$ ,  $(0.6, 0.6)$ , and  $(1, 1)$ , respectively. The color scale at the right of each panel represents the value of  $\Delta_T$ .



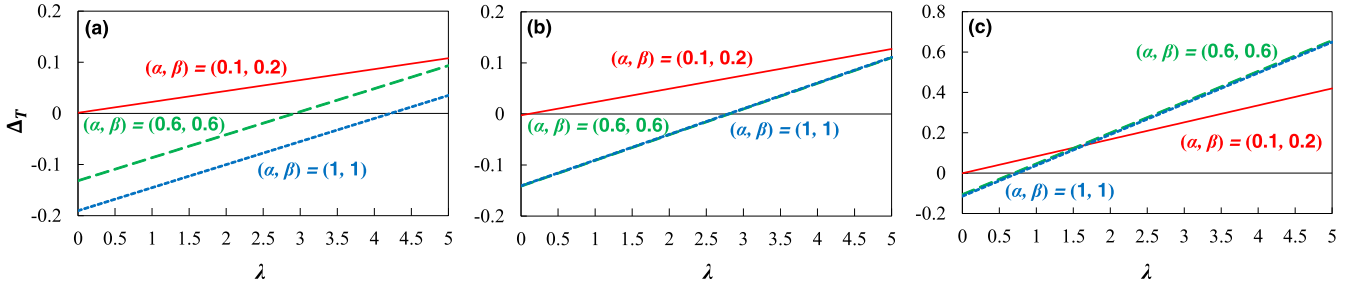


FIG. 9. Simulation values of  $\Delta_T$  with (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  as functions of  $\lambda$  for various  $(\alpha, \beta) \in \{(0.1, 0.2), (0.6, 0.6), (1, 1)\}$ . The parameters other than  $(\alpha, \beta)$  are the same as in Fig. 7.

On the other hand, in Figs. 7(d)–7(f), unlike the cases in Figs. 7(a)–7(c), the difference between  $T_R$  and  $T_G$  seems to vanish.

In order to compare the difference between  $T_R$  and  $T_G$  for various  $(\alpha, \beta)$ , we define  $\Delta_T$  as the ratio of the change from  $T_R$  to  $T_G$ , that is,

$$\Delta_T = \frac{T_G - T_R}{T_R}. \quad (1)$$

From this definition of  $\Delta_T$ ,  $\Delta_T < 0$  ( $\Delta_T > 0$ ) indicates that group (random) sequences are preferable for smooth transportation. Note that in the following, to calculate  $\Delta_T$  we assume that each group in a group sequence is arranged in ascending order in terms of species number  $s$ .

The simulation values of  $\Delta_T$  for various  $(\alpha, \beta)$  with (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  are plotted in Fig. 8. Note that the black lines represent the boundaries between the low-density–high-density (LD-HD) and the maximal current (MC) phases of the single-species TASEP with hopping probability  $p_1$  (boundary A) in Fig. 8(a) and  $p_1$  (boundary B1) and  $p_3$  (boundary B3) in Fig. 8(b), respectively.

Figure 8 shows that for all three values of  $S$ ,  $\Delta_T$  is small in the region where  $\min(\alpha, \beta)$  is relatively large. [In Fig. 8(b),  $\Delta_T$  finally yields to a constant value in the upper-right region beyond boundary B3.] On the other hand,  $\Delta_T$  is small in the region where  $\min(\alpha, \beta)$  is relatively small. [In Figs. 8(a) and 8(b),  $\Delta_T$  is almost 0, especially in the lower-left region

beyond the boundary A or B1.] Here, we term the region with  $\Delta_T < 0$  as the “group-advantageous region” ( $T_R > T_G$ ), whereas we designate the region with  $\Delta_T \approx 0$  as a “neutral region” ( $T_R \approx T_G$ ), if it exists.

These results indicate that group sequences can make transportation smoother than random sequences when the system is mainly governed by the bulk region of the lattice, but the dependence on the type of sequences vanishes (or decreases) when the system is mainly governed by the boundaries.

### B. With sorting cost ( $\lambda > 0$ )

In this subsection, we consider the sorting cost by varying  $\lambda$  for the same parameter sets in the previous subsection. Appendix B presents specific schemes for obtaining the minimal number of exchanges necessary to sort the particles in the simulations.

Figure 9 plots  $\Delta_T$  for (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  as functions of  $\lambda$  for various  $(\alpha, \beta) \in \{(0.1, 0.2), (0.6, 0.6), (1, 1)\}$ , which are plotted as black crosses in Fig. 8. We emphasize again that in the region with  $\Delta_T < 0$  group sequences are preferred, even if when the sorting cost is considered, whereas in the region with  $\Delta_T > 0$  random sequences are preferred. Note that the cases with  $\lambda = 0$  correspond to those obtained without considering the sorting cost.

As discussed in the previous subsection, we note that  $\Delta_T \leq 0$  for almost all  $(\alpha, \beta)$  when  $\lambda = 0$ , indicating that

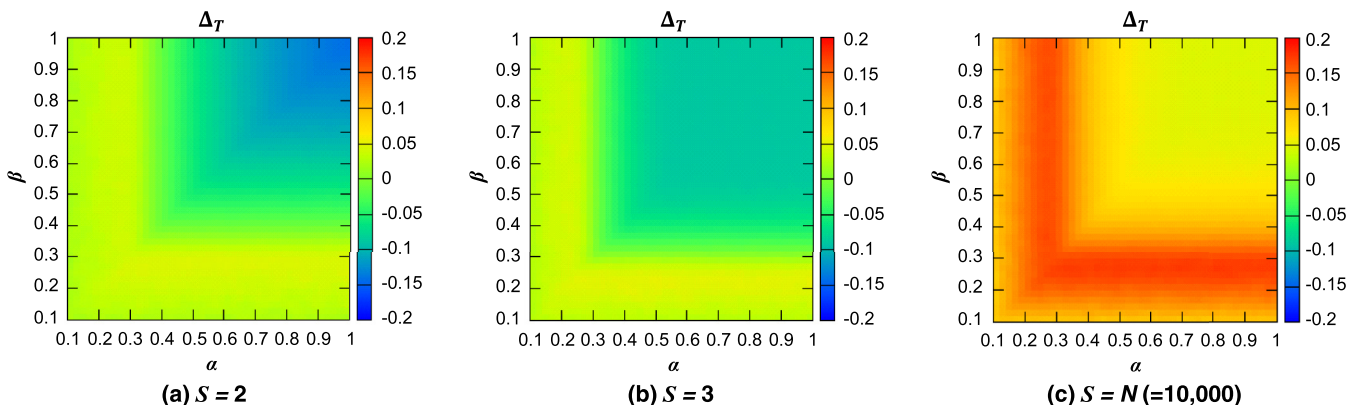


FIG. 10. Simulation values of  $\Delta_T$  for various  $(\alpha, \beta)$  with (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  for  $\lambda = 1$ . The other parameters are the same as in Fig. 7.

sorting is almost always beneficial for smooth transportation. However, once the sorting cost is considered, the sign of  $\Delta_T$  can become positive, especially in the region where  $\min(\alpha, \beta)$  is relatively small, indicating that sorting is not always beneficial. Note that the curves of  $(\alpha, \beta) = (0.6, 0.6)$  and  $(1, 1)$  are observed to overlap each other in Figs. 9(b) and 9(c), unlike Fig. 9(a). This happens because there is no difference in  $\Delta_T$  at these two points when  $\lambda = 0$ , as we can see in Figs. 8(b) and 8(c).

Figure 10 plots  $\Delta_T$  for (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N$  for various  $(\alpha, \beta)$  with  $\lambda = 1$ . In this figure, we note the existence of a new region in which  $\Delta_T > 0$ , which we term a “random-advantageous region” ( $T_R < T_G$ ). This new region widens as  $\lambda$  increases, finally resulting in the complete disappearance of the group-advantageous region for large-enough  $\lambda$ . Note that Fig. 10(c) exhibits only a random-advantageous region.

$$\begin{pmatrix} P_{00} \\ P_{0*} \\ P_{10} \\ P_{1*} \\ P_{20} \\ P_{2*} \end{pmatrix} = \begin{pmatrix} 1 - \alpha & (1 - \alpha)\beta \\ 0 & (1 - \alpha)(1 - \beta) \\ r\alpha & r\alpha\beta \\ 0 & r\alpha(1 - \beta) \\ (1 - r)\alpha & (1 - r)\alpha\beta \\ 0 & (1 - r)\alpha(1 - \beta) \end{pmatrix} \begin{pmatrix} 0 \\ p_1 \\ 1 - p_1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & 1 - p_2 & \beta \\ 0 & 0 & 0 & 1 - \beta \end{pmatrix} \begin{pmatrix} P_{00} \\ P_{0*} \\ P_{10} \\ P_{1*} \\ P_{20} \\ P_{2*} \end{pmatrix} \quad (2)$$

$$L = 2$$

This subsection presents the derivation of an approximate  $Q_S$  based on a Markov chain model. Due to the difficulty of considering general values of  $L$  (the length of the lattice) and  $S$  (the number of particle species), we consider the simplest case—with  $L = 2$  and  $S = 2$ . As two species of particles exist—that is, particles with hopping probability  $p_1$  and particles with  $p_2$ —each site may have three states: “unoccupied (state 0),” “occupied by a particle 1 (state 1),” and “occupied by a particle 2 (state 2).” This results in nine possible states; however, noting that it is not necessary to distinguish the particle at site 1 because it always leaves the lattice with probability  $\beta$ , the number of possible states can be reduced to 6. Here, we define the probability distribution  $P_{ij}$  ( $i = 0, 1, 2, j = 0, *$ ), where  $i$  and  $j$  represent the state number of site 0 and 1, respectively. Note that state  $*$  indicates either of state 1 or 2.

The master equations for the steady state are summarized in Eq. (2), using the relation  $r_1 + r_2 = 1$ . Note that  $r_1$  and  $r_2$  are replaced with  $r$  and  $1 - r$ , respectively, for the case  $S = 2$ . In addition,  $P_{ij}$  must satisfy the normalization condition

$$\sum_{i=0}^2 P_{i0} + \sum_{i=0}^2 P_{i*} = 1. \quad (3)$$

From Eqs. (2) and (3), the flow of the system can be written as a function of  $p_1, p_2$ , and  $r$ ; that is,  $Q_2(p_1, p_2; r)$ , is given by the following expression:

$$Q_2(p_1, p_2; r) = p_1 P_{10} + p_2 P_{20} = \frac{p_1 p_2 A}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B}, \quad (4)$$

## IV. THEORETICAL ANALYSES

In this section, we show that the simulation results can be theoretically reproduced in some special cases. Specifically, we have succeeded in obtaining a mathematical proof of the appearance of the group-advantageous region for any group number  $S(>1)$  when  $\lambda = 0$ .

### A. Approximate flow of a multispecies TASEP

In this subsection, before calculating  $T$ , we briefly discuss the steady-state flow  $Q_S$  of the multispecies TASEP that corresponds to a random sequence. We write  $Q_S = Q_S(p_1, \dots, p_S; r_1, \dots, r_S)$  with the arguments abbreviated in obvious cases. When the flow  $Q$  is simulated for each parameter set, we first evolve the system for  $10^5$  time steps and then average over the next  $10^6$  time steps.

where

$$A = \alpha\beta(\alpha + \beta - \alpha\beta) \quad (5)$$

and

$$B = \alpha^2 + \beta^2 - \alpha^2\beta - \alpha\beta^2 + \alpha\beta. \quad (6)$$

The specific forms of the probability distributions are summarized in Appendix C.

For  $r = 1$  and  $p_1 = p$ , the system reduces to the single-species TASEP with the flow  $Q_1(p)$ , where

$$Q_1(p) = \frac{pA}{pB + A}. \quad (7)$$

Note that the flow of the single-species TASEP for general  $L$  is exactly solved in Ref. [65]. Therefore, assuming that the value  $p = p_h$  satisfies the condition  $Q_2(p_1, p_2; r) = Q_1(p)$ , we can derive

$$p_h = \frac{p_1 p_2}{(1 - r)p_1 + r p_2}. \quad (8)$$

The quantity  $p_h$  is termed the harmonic mean of  $p_1$  and  $p_2$ . Accordingly, for  $L = 2$ ,  $Q_2$  is equivalent to  $Q_1(p = p_h)$ . This relation holds for any species number  $S(>2)$ , as we show in Appendix D.

Figure 11 compares the simulation and theoretical curves for various  $\beta \in \{0.2, 0.6, 1\}$  with (a)  $S = 2$ , (b)  $S = 3$ , and (c)  $S = N (= 10\,000)$ . In all the figures, the simulations show very good agreement with our exact analyses.

### 2. General $L(>2)$

For general  $L$  and  $S$ , it is complicated to solve the master equations. Therefore, in this subsection, we instead assume an inequality, based on the results in the previous subsection

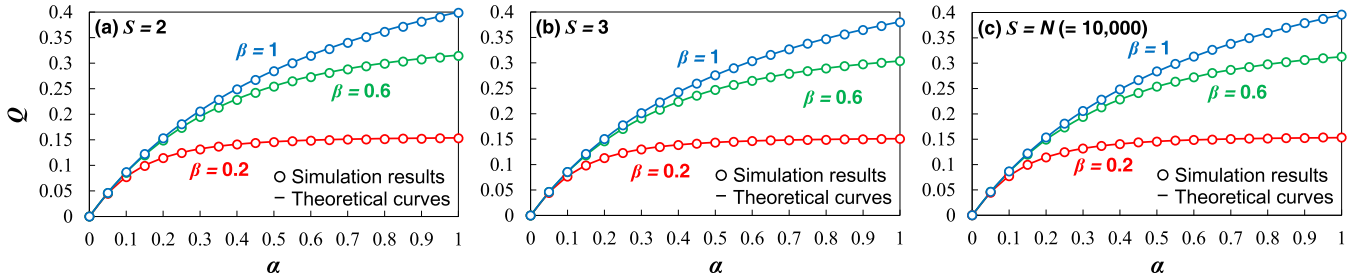


FIG. 11. Simulation (circles) and theoretical (curves) values of (a)  $Q_2$ , (b)  $Q_3$ , and (c)  $Q_N$  for  $L = 2$  as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$ . The other parameters are fixed at (a)  $(p_1, p_2; r) = (0.5, 1; 0.5)$ , (b)  $(p_1, p_2, p_3; r_1, r_2, r_3) = (0.4, 0.6, 0.8; 0.2, 0.3, 0.5)$ , (c)  $p_s = 1 - 0.6(S - s)/(S - 1)$  ( $s = 1, 2, \dots, N$ ) and  $r_s = 1/N (= 1/10\,000)$ .

and the qualitative discussions, and confirm the validity of the inequality by the simulations.

First, for general  $L$  and  $S$ ,  $Q_S$  is clearly larger than  $Q_1(p = p_{\min})$ , where  $p_{\min} = \min\{p_1, p_2, \dots, p_S\}$ .

In addition, for  $L > 2$ , a platoon can be observed in the bulk of the lattice, in which a slower particle behaves as a bottleneck, and faster particles behind it cannot hop with a probability larger than that of the smaller one, i.e., less than  $p_h$ , as shown in Fig. 12. This phenomenon suppresses the flow, implying that  $Q_S$  is smaller than  $Q_1(p = p_h)$ .

Consequently,  $Q_S$  satisfies the following inequality:

$$Q_1(p = p_{\min}) < Q_S < Q_1(p = p_h). \quad (9)$$

In this subsection, we hereafter consider the case  $S = 2$  excluding Fig. 16 and its explanation.

Figure 13 shows the phase diagrams obtained by plotting the simulation values for (a) the single-species TASEP with  $p = p_{\min}$ , (b) the two-species TASEP, and (c) the single-species TASEP with  $p = p_h$ , respectively. Note that  $Q_2$  ( $Q_1$ ) are the simulation (theoretical) values (and similarly hereafter).

Comparing these three figures shows that Eq. (9) obviously holds. In addition, as in Figs. 13(a) and 13(c), we find that three different phases—HD, LD, and MC—also exist in Fig. 13(b). Due to Eq. (9), the boundaries between the LD (HD) and MC phases of Fig. 13(b) lie between those of Figs. 13(a) and 13(c). Note that the black lines in Figs. 13(a) and 13(c) are theoretical boundaries, based on the fact that the boundary between the LD and MC phases of the single-species TASEP with hopping probability  $p$  [27] can be written as

$$\begin{aligned} \alpha &= 1 - \sqrt{1 - p} \wedge 1 - \sqrt{1 - p} < \beta < 1 \\ \beta &= 1 - \sqrt{1 - p} \wedge 1 - \sqrt{1 - p} < \alpha < 1. \end{aligned} \quad (10)$$

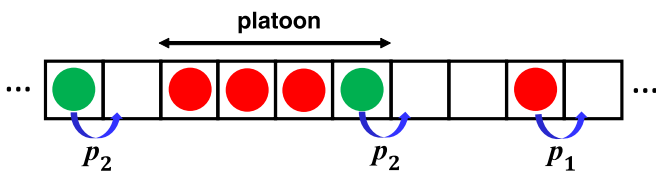


FIG. 12. Schematic illustration of a platoon. In this figure, we set  $S = 2$ , with the red particles belonging to species 1 (faster) and the green ones to species 2 (slower). A green particle blocks the red particles behind it, so that the trailing red particles cannot hop with probability  $p_1$  but only with probability  $p_2$ , which is less than  $p_h$ .

Here, as with  $\Delta_T$ , we define  $\Delta_Q$  as the ratio of the change from  $Q_1(p = p_h)$  to  $Q_2$ ; that is,

$$\Delta_Q = \frac{Q_2 - Q_1(p = p_h)}{Q_1(p = p_h)}, \quad (11)$$

and we note that  $\Delta_Q = 0$  when  $Q_1(p = p_h) = 0$ .

Figure 14 shows  $\Delta_Q$  for various  $(\alpha, \beta)$ , for the fixed parameter set  $(p_1, p_2; r) = (0.5, 1; 0.5)$ . The black lines represent the boundaries between the LD-HD and MC phases of the single-species TASEP with hopping probability  $p_1$  (boundary 1) and the single-species TASEP with  $p = p_h$  (boundary 2). Therefore, the lower-left (upper-right) region beyond boundary 1 (boundary 2) corresponds to the LD-HD (MC) phases both for the two-species TASEP and for the single-species TASEP with  $p = p_h$ . This figure confirms that  $\Delta_Q$  starts from 0 in the LD-HD phase, decreases, and finally yields to a constant value in the MC phase as  $(\alpha, \beta)$  approaches the upper right.

Figure 15 plots  $\Delta_Q$  as a function of  $(p_1, p_2)$  for various  $(\alpha, \beta) \in \{(0.1, 0.2), (0.2, 0.1), (1, 1)\}$ , fixing  $r = 0.5$ . Note that both the single-species TASEP with hopping probability  $p_h$  and the two-species TASEP exhibit the LD, HD, and MC phases with  $(\alpha, \beta) = (0.1, 0.2)$ ,  $(0.2, 0.1)$ , and  $(1, 1)$ , respectively. This is because  $\forall (p_1, p_2) (0.2 \leq p_1, p_2 \leq 1) 0.1 < 1 - \sqrt{1 - p_{\min}} \leq 1 - \sqrt{1 - p_h}$ . For example, Fig. 13 confirms that those three points exist within each corresponding phase for  $(p_1, p_2) = (0.5, 1)$ .

In Figs. 15(a) and 15(b), we find that  $\Delta_Q$  is approximately 0, whereas  $\Delta_Q$  deviates from 0 in Fig. 15(c), as is also observed in Fig. 14. These phenomena can be explained as follows.

First, in the LD (HD) phase,  $Q_2$  is mainly governed by the input (output) probability, leading to  $\Delta_Q \rightarrow 0$ , i.e.,  $Q_2$  approaches  $Q_1(p = p_h)$ . This is because  $Q_2$  deviates from  $Q_1(p = p_h)$  mainly due to the existence of platoons, which do not influence the flow much in this phase. Note that  $\Delta_Q$  decreases as  $\alpha$  or  $\beta$  approaches 0.2, because the influence of platoons increases, approaching the MC phase of the two-species TASEP.

On the other hand, in the MC phase,  $Q_2$  is mainly governed by the bulk region of the lattice. Therefore, the existence of platoons has a more critical influence on  $Q_2$ , causing  $\Delta_Q$  to deviate from 0; i.e.,  $Q_2 < Q_1(p = p_h)$ . Especially as  $|p_1 - p_2|$  increases, the extent of the deviation also increases. This is because the effect of platoons increases when there is a large gap between  $p_1$  and  $p_2$ .

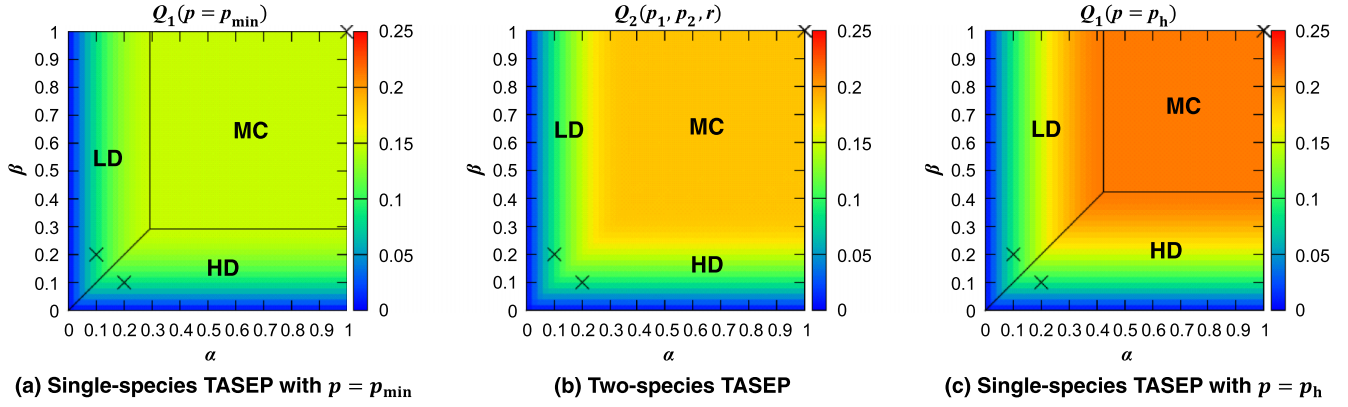


FIG. 13. Phase diagrams. The color bars indicate the simulation values of (a)  $Q_1(p = p_{\min})$ , (b)  $Q_2(p_1, p_2; r)$ , and (c)  $Q_1(p = p_h)$ , respectively, obtained by fixing  $(p_1, p_2; r) = (0.5, 1; 0.5)$ ,  $p_{\min} = 0.5$ , and  $p_h = 1/(0.5/0.5 + 0.5/1) = 2/3$ . Note that the three black crosses in each figure represent  $(\alpha, \beta) = (0.1, 0.2)$ ,  $(0.2, 0.1)$ , and  $(1, 1)$ , respectively.

In the end of this subsection, we briefly discuss the case of general  $S$ . Figure 16 plots (a)  $Q_2/Q_1(p = p_h)$ , (b)  $Q_3/Q_1(p = p_h)$ , and (c)  $Q_N/Q_1(p = p_h)$  for various  $L \in \{2, 3, 4, 10, 100, 200\}$ . Both of  $Q_S$  ( $S = 2, 3, 10\,000$ ) and  $Q_1(p = p_h)$  are obtained by the simulations. In all the figures, we observe that for  $L > 2$ ,  $Q_S/Q_1(p = p_h)$  generally becomes less than 1, i.e.,  $Q_S < Q_1(p = p_h)$ , especially when  $\alpha$  increases, i.e., the system approaches and exhibits the MC phase. Note that for  $S = 10\,000$ , the difference among  $L = 2, 3, 4$  is unclear; however, it becomes obvious when  $L \geq 10$ . Those results imply that Eq. (9) and its qualitative discussions can be applicable for general  $L$  and  $S$ .

### B. Relation between $T_R$ and $T_G$ without the sorting cost

Hereafter, we assume  $p_1 < p_2 < \dots < p_S$ ,  $S > 1$ ,  $\forall s r_s > 0$ , and  $\alpha > 0$ .

In this subsection, we fix  $\lambda = 0$ , i.e., we do not consider the sorting cost. If for any number of particle species  $s$ ,  $r_s N$  is large enough for  $T_G$  and  $T_R$  to be determined by the steady-

state flow (see Appendix E), we obtain

$$T_R \approx \frac{N}{Q_S} \quad (12)$$

and

$$T_G \approx \sum_{s=1}^S \frac{r_s N}{Q_1(p = p_s)}. \quad (13)$$

Note that this approximation immediately implies the independence of  $T_G$  from the order of the group sequence. Strictly speaking,  $T_G$  can differ depending on the order of each group in the group sequence. However, that difference can be ignored for large  $N$  (see Appendix E).

In addition, we define the transportation times of the particles with the same hopping probabilities  $p_h$  and  $p_{\min}$  as

$$T_H \approx \frac{N}{Q_1(p = p_h)} \quad (14)$$

and

$$T_M \approx \frac{N}{Q_1(p = p_{\min})}, \quad (15)$$

respectively. From Eqs. (9), (12), (14), and (15), we immediately obtain the inequality

$$T_H < T_R < T_M. \quad (16)$$

In the following, we show that a general relation between  $T_R$  and  $T_G$  can be obtained mathematically for general  $S (> 1)$ . We emphasize that this relation can be proven by comparing  $T_H$  and  $T_G$ , and not by comparing  $T_R$  and  $T_G$  directly. Here, we introduce the new function  $f(\alpha, \beta; p_1, \dots, p_S; r_1, \dots, r_S)$ , which is defined as follows:

$$f(\alpha, \beta; p_1, \dots, p_S; r_1, \dots, r_S) = T_H - T_G. \quad (17)$$

Because we can assume  $\alpha < \beta$  without loss of generality, we adopt this assumption in the following discussion, writing in abbreviated form  $f(\alpha, \beta; p_1, \dots, p_S; r_1, \dots, r_S) = f(\alpha)$ . Note that for cases with  $\alpha \geq \beta$ , the theoretical results can be obtained simply by replacing  $\alpha$  (LD) with  $\beta$  (HD).

A contour map of  $f(\alpha)$  in the  $(\alpha, \beta)$  plane exhibits four large regions, which are summarized in Table I.

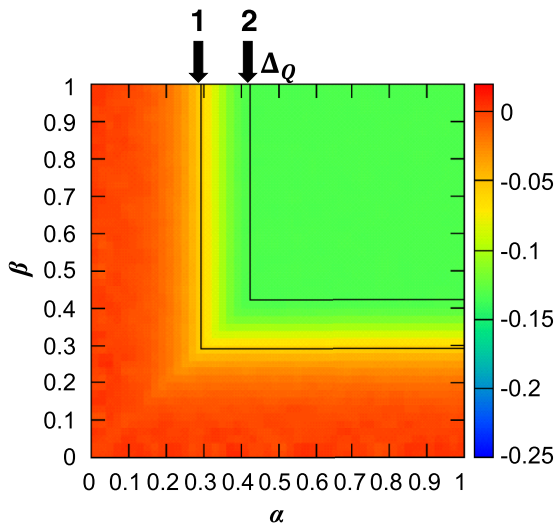


FIG. 14. Simulation values of  $\Delta_Q$  for various  $(\alpha, \beta)$ , fixing  $(p_1, p_2; r) = (0.5, 1; 0.5)$ .



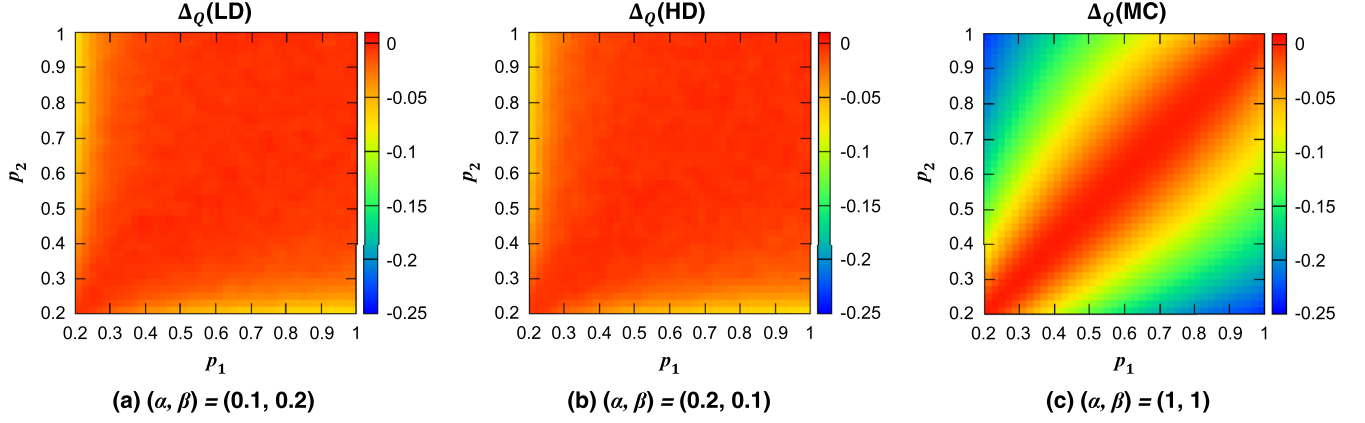


FIG. 15. Simulation values of  $\Delta_Q$  for various  $(p_1, p_2)$ ,  $S = 2$ , and  $r = 0.5$ , (a)  $(\alpha, \beta) = (0.1, 0.2)$  (LD), (b)  $(\alpha, \beta) = (0.2, 0.1)$  (HD), and (c)  $(\alpha, \beta) = (1, 1)$  (MC).

In the following subsections, we examine the behavior of  $f(\alpha)$  according to this classification.

### 1. Region 1: $\alpha < 1 - \sqrt{1 - p_1}$

In this region, all the steady-state phases of the single-species TASEP for any  $p_s$  exhibit the LD phase. Here the steady-state flow for the single-species TASEP with parallel updating [27] is given by

$$Q_1(p) = \begin{cases} \alpha \frac{p - \alpha}{p - \alpha^2} & \text{for LD phase} \\ \frac{1}{2}(1 - \sqrt{1 - p}) & \text{for MC phase} \end{cases} \quad (18)$$

Therefore, we obtain

$$T_G \approx \sum_{s=1}^S \frac{r_s N (p_s - \alpha^2)}{\alpha (p_s - \alpha)} \quad (19)$$

and

$$T_H \approx \frac{N(p_h - \alpha^2)}{\alpha(p_h - \alpha)}. \quad (20)$$

From Eqs. (19) and (20), we obtain  $f(\alpha)$  as

$$f(\alpha) \approx \frac{N(p_h - \alpha^2)}{\alpha(p_h - \alpha)} - \sum_{s=1}^S \frac{r_s N (p_s - \alpha^2)}{\alpha (p_s - \alpha)}. \quad (21)$$

After some calculations, we obtain

$$f(\alpha) < 0; \quad (22)$$

the detailed derivation is given in Appendix F.

### 2. Region 2: $1 - \sqrt{1 - p_1} \leq \alpha < 1 - \sqrt{1 - p_h}$

This region is further divided into  $(u - 1)$  subregions, as summarized in Table II.

In Subregion  $2-v$  ( $v = 1, 2, \dots, u - 1$ ), the single-species TASEP with  $p = p_1, \dots, p_v, p_h$  exhibits the MC phase, whereas that with  $p = p_{v+1}, \dots, p_S$  displays the LD phase. Therefore, using Eq. (18), we obtain

$$Q_1(p) = \begin{cases} \frac{1}{2}(1 - \sqrt{1 - p}) & \text{for } p = p_1, \dots, p_v, p_h \\ \alpha \frac{p - \alpha}{p - \alpha^2} & \text{for } p = p_{v+1}, \dots, p_S \end{cases} \quad (23)$$

From Eqs. (13), (14), and (23),  $T_G$  and  $T_H$  can be written as follows;

$$T_G \approx \sum_{s=1}^v \frac{2r_s N}{1 - \sqrt{1 - p_s}} + \sum_{s=v+1}^S \frac{r_s N (p_s - \alpha^2)}{\alpha (p_s - \alpha)} \quad (24)$$

and

$$T_H \approx \frac{N(p_h - \alpha^2)}{\alpha(p_h - \alpha)}. \quad (25)$$

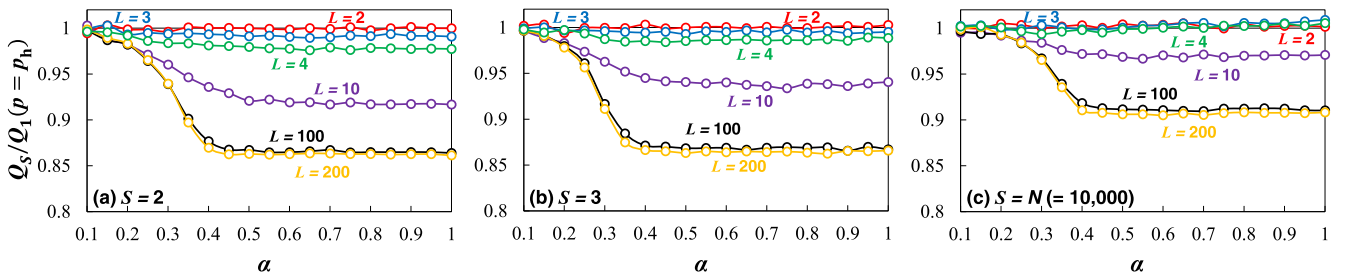


FIG. 16. Simulation values of the ratio (a)  $Q_2/Q_1(p = p_h)$ , (b)  $Q_3/Q_1(p = p_h)$ , and (c)  $Q_N/Q_1(p = p_h)$  as functions of  $\alpha$  for various  $L \in \{2, 3, 4, 10, 100, 200\}$ , fixing  $\beta = 0.6$ . The other parameters are fixed at (a)  $(p_1, p_2; r) = (0.5, 1; 0.5)$ , (b)  $(p_1, p_2, p_3; r_1, r_2, r_3) = (0.4, 0.6, 0.8; 0.2, 0.3, 0.5)$ , (c)  $p_s = 1 - 0.6(S - s)/(S - 1)$  ( $s = 1, 2, \dots, N$ ) and  $r_s = 1/N (= 1/10000)$ .

TABLE I. Classification of Regions.

Region no.	Range
1	$\alpha < 1 - \sqrt{1 - p_1}$
2	$1 - \sqrt{1 - p_1} \leq \alpha < 1 - \sqrt{1 - p_h}$
3	$1 - \sqrt{1 - p_h} \leq \alpha < 1 - \sqrt{1 - p_S}$
4	$1 - \sqrt{1 - p_S} \leq \alpha$

From Eqs. (24) and (25), we thus obtain  $f(\alpha)$  in the form

$$f(\alpha) \approx \frac{N(p_h - \alpha^2)}{\alpha(p_h - \alpha)} - \sum_{s=1}^v \frac{2r_s N}{1 - \sqrt{1 - p_s}} - \sum_{s=v+1}^S \frac{r_s N(p_s - \alpha^2)}{\alpha(p_s - \alpha)}. \quad (26)$$

For  $1 - \sqrt{1 - p_v} \leq \alpha < 1 - \sqrt{1 - p_{v+1}}$ , due to  $\alpha < p_{v+1} < \dots < p_S$  and  $\alpha < p_h$ , we obtain  $p_s - \alpha > 0$  for  $s = v + 1, \dots, S$  and  $p_h - \alpha > 0$ . Therefore, the function  $f(\alpha)$  is continuous and differentiable with respect to  $\alpha$ , including at each boundary (see Appendix G). However, the signs of  $f(\alpha)$  and  $df(\alpha)/d\alpha$  are not specified. Note that the following condition

$$\lim_{\alpha \rightarrow q_h - 0} \frac{df(\alpha)}{d\alpha} > 0, \quad (27)$$

where  $q_h = 1 - \sqrt{1 - p_h}$  indicates that  $f(\alpha)$  increases monotonically at least near the boundary between Subregion 2- $(u - 1)$  and Region 3. This is discussed in Appendix H.

**3. Region 3:  $1 - \sqrt{1 - p_h} \leq \alpha < 1 - \sqrt{1 - p_S}$**

Similarly to Region 2, this region is further divided into  $(S - u + 1)$  subregions, as summarized in Table III. Note that Subregion 3- $u$  vanishes in the case  $p_h = p_u$ , resulting in the  $(S - u)$  subregions.

In Region 3- $v$  ( $v = u, u + 1, \dots, S$ ), the single-species TASEP with  $p = p_1, \dots, p_{v-1}, p_h$  exhibits the MC phase, whereas that with  $p = p_v, \dots, p_S$  displays the LD phase. Therefore, using Eq. (18), we obtain  $Q_1(p)$

$$Q_1(p) = \begin{cases} \frac{1}{2}(1 - \sqrt{1 - p}) & \text{for } p = p_1, \dots, p_{v-1}, p_h. \\ \alpha \frac{p - \alpha}{p - \alpha^2} & \text{for } p = p_v, \dots, p_S. \end{cases} \quad (28)$$

TABLE II. Classification of subregions in Region 2.

Subregion no.	Range
2-1	$1 - \sqrt{1 - p_1} \leq \alpha < 1 - \sqrt{1 - p_2}$
2-2	$1 - \sqrt{1 - p_u} \leq \alpha < 1 - \sqrt{1 - p_{u+1}}$
...	...
2- $v$	$1 - \sqrt{1 - p_v} \leq \alpha < 1 - \sqrt{1 - p_{v+1}}$
...	...
2- $(u - 1)$	$1 - \sqrt{1 - p_{u-1}} \leq \alpha < 1 - \sqrt{1 - p_h}$

TABLE III. Classification of subregions in Region 3.

Subregion no.	Range
3- $u$	$1 - \sqrt{1 - p_h} \leq \alpha < 1 - \sqrt{1 - p_u}$
3- $(u + 1)$	$1 - \sqrt{1 - p_u} \leq \alpha < 1 - \sqrt{1 - p_{u+1}}$
...	...
3- $v$	$1 - \sqrt{1 - p_{v-1}} \leq \alpha < 1 - \sqrt{1 - p_v}$
...	...
3- $S$	$1 - \sqrt{1 - p_{S-1}} \leq \alpha < 1 - \sqrt{1 - p_S}$

From Eqs. (13), (14), and (28),  $T_G$  and  $T_H$  can be written as follows;

$$T_G \approx \sum_{s=1}^{v-1} \frac{2r_s N}{1 - \sqrt{1 - p_s}} + \sum_{s=v}^S \frac{r_s N(p_s - \alpha^2)}{\alpha(p_s - \alpha)} \quad (29)$$

and

$$T_H \approx \frac{2N}{1 - \sqrt{1 - p_h}}. \quad (30)$$

From Eqs. (29) and (30),  $f(\alpha)$  becomes

$$f(\alpha) \approx \frac{2N}{1 - \sqrt{1 - p_h}} - \sum_{s=1}^{v-1} \frac{2r_s N}{1 - \sqrt{1 - p_s}} - \sum_{s=v}^S \frac{r_s N(p_s - \alpha^2)}{\alpha(p_s - \alpha)}. \quad (31)$$

For  $1 - \sqrt{1 - p_{v-1}} \leq \alpha < 1 - \sqrt{1 - p_v}$ , due to  $\alpha < p_v < \dots < p_S$ , we obtain  $p_s - \alpha > 0$  for  $s = v, \dots, S$ . Similarly to Region 2,  $f(\alpha)$  is continuous and differentiable with respect to  $\alpha$  including at each boundary.

After some calculations, we find that  $df(\alpha)/d\alpha$  satisfies

$$\frac{df(\alpha)}{d\alpha} > 0 \quad (32)$$

in each subregion, as discussed in detail in Appendix I. Equation (32) indicates that  $f(\alpha)$  is a monotonically increasing function of  $\alpha$  throughout Region 3.

**4. Region 4:  $1 - \sqrt{1 - p_S} \leq \alpha$**

In this region, all the steady-state phases of the single-species TASEP for any hopping probabilities  $p_s$  and  $p_h$  exhibit the MC region. Note that this region vanishes in the case  $p_S = 1$  because  $1 - \sqrt{1 - p_S} = 1$ .

Therefore, we obtain

$$T_G \approx \sum_{s=1}^S \frac{2r_s N}{1 - \sqrt{1 - p_s}} \quad (33)$$

and

$$T_H \approx \frac{2N}{1 - \sqrt{1 - p_h}}, \quad (34)$$

both of which are independent of  $\alpha$ . From Eqs. (33) and (34), we thus obtain  $f(\alpha)$  as

$$f(\alpha) \approx \frac{2N}{1 - \sqrt{1 - p_h}} - \sum_{s=1}^S \frac{2r_s N}{1 - \sqrt{1 - p_s}}. \quad (35)$$

TABLE IV. Sign of  $f(\alpha)$  and  $df(\alpha)/d\alpha$ .

Region no.	$f(\alpha)$	$df(\alpha)/d\alpha$
1	–	U
2	U	U
3	U	+
4	+	0

After some calculations, we obtain

$$f(\alpha) > 0, \quad (36)$$

the detailed derivation of which is given in Appendix J.

### 5. Relation between $T_R$ and $T_G$

With the results of Secs. IV B 1–IV B 4, we can obtain a general relation between  $T_R$  and  $T_G$  for some special cases. Table IV summarizes the signs of  $f(\alpha)$  and  $df(\alpha)/d\alpha$  in each region. Note that “U” indicates that the sign is unclear.

Considering Table IV and the continuity of  $f(\alpha)$  including at each boundary (see Appendix G), we find from the intermediate value theorem that  $\exists \alpha_{cr}$  such that  $f$  satisfies

$$f(\alpha = \alpha_{cr}) = 0 \Leftrightarrow T_H = T_G \quad (37)$$

in Region 2 or 3. The specific conditions that  $\alpha_{cr}$  must satisfy are given in Appendix K.

Defining  $\alpha_{cr,max}$  as the largest value among the quantities  $\alpha_{cr}$ , we obtain  $f(\alpha) > 0$ —i.e.,  $T_H > T_G$ —in the region where  $\alpha > \alpha_{cr,max}$ . This is because  $f(\alpha)$  is continuous and increases monotonically from a point in Region 2 (and through Region 3), to yield  $f(\alpha) > 0$  in Region 4.

Considering Eq. (16), we finally obtain

$$T_G < T_H < T_R \quad (38)$$

in the region  $\alpha > \alpha_{cr,max}$ . Equation (38) means that  $\Delta_T < 0$ , reproducing the simulation results in the region where  $\min(\alpha, \beta)$  is relatively large. This result indicates that the group-advantageous region must appear even in a case with  $p_S = 1$ , for which Region 4 vanishes.

In analogy with the discussion above, we can also predict that a region with  $\Delta_T < 0$  must appear in the case  $S = N$ .

### C. Relation between $T_R$ and $T_G$ with sorting cost

In this subsection, we discuss the change in the relation between  $T_R$  and  $T_G$  when  $\lambda > 0$ , i.e., when the sorting cost is included. In the following, we first obtain a general formula for the sorting cost and then evaluate upper and lower limits to  $\lambda$ .

#### 1. General formula for the sorting cost

First, we calculate mathematically the averaged minimal number of exchanges necessary to sorting the particles from random to group sequences.

We here define  $\bar{K}$  as the averaged value of  $K$ , using the fact that  $\tau_R$  can take  $N!/\prod_{s=1}^S (r_s N)!$  patterns with equal probability. We thus have

$$\bar{K} = \frac{\prod_{s=1}^S (r_s N)!}{N!} \sum_{\forall \tau_R} K(\tau_G, \tau_R). \quad (39)$$

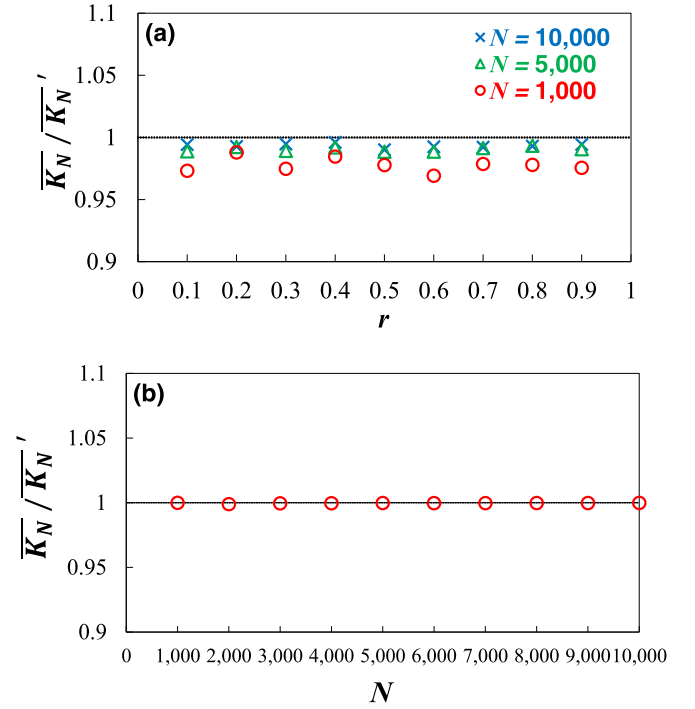


FIG. 17. (a) Simulation values of the ratio  $\bar{K}/\bar{K}'$  as functions of  $r$  for various  $N \in \{1000 \text{ (red)}, 5000 \text{ (green)}, 10000 \text{ (blue)}\}$  for  $S = 2$ . (b) Simulation values of the ratio  $\bar{K}/\bar{K}'$  as functions of  $N$  for  $S = N$ . Note that  $\bar{K}$  and  $\bar{K}'$  are each simulation value obtained by respectively averaging over 100 trials.

If  $K'(\tau_G, \tau_R)$  is the minimal number of exchanges necessary to sort the particles from a random sequence  $\tau_R$  to a given fixed group sequence  $\tau_G$ , then  $K'(\tau_G, \tau_R)$  satisfies

$$K(\tau_G, \tau_R) = \min\{K'(\tau_G, \tau_R)\}. \quad (40)$$

Note that the number of elements of  $\{K'(\tau_G, \tau_R)\}$  is equal to that of  $\{\tau_G\}$  from the definition. Equations (39) and (40) indicate that the best group sequence  $\tau_G$  can vary depending on the particular random sequence  $\tau_R$ .

Due to the difficulty of a general calculation of  $\bar{K}$ , we instead calculate  $\bar{K}'$ , which is defined as follows:

$$\bar{K}' = \frac{\prod_{s=1}^S (r_s N)!}{N!} \sum_{\forall \tau_R} K'(\tilde{\tau}_G, \tau_R), \quad (41)$$

where  $\tilde{\tau}_G$  is a fixed sequence out of the set  $\{\tau_G\}$  for all possible  $\tau_R$ .

For  $S = 2$  and  $S = N$ ,  $\bar{K}'$  can be generally calculated as

$$\bar{K}' = \begin{cases} r(1-r)N & \text{for } S = 2 \\ N - \sum_{k=1}^N \frac{1}{k} & \text{for } S = N \end{cases} \quad (42)$$

the detailed derivations of which are discussed in Appendix L.

Figure 17 shows the ratio  $\bar{K}/\bar{K}'$  for (a)  $S = 2$  and (b)  $S = N$ . Both figures show that  $\bar{K}/\bar{K}' \approx 1$ , i.e.,  $\bar{K} \approx \bar{K}'$ , indicating that there is no problem in substituting  $\bar{K}'$  for  $\bar{K}$  for large-enough  $N$ .

In the following calculations, we therefore use  $\bar{K}'$  instead of  $\bar{K}$  because  $\bar{K}'$  can be represented by a general formula, whereas  $\bar{K}$  cannot.

TABLE V. Upper and lower limits to  $\lambda_{\text{cr}}$ .

Region no.	Upper and lower limits to $\lambda_{\text{cr}}$
1	$0 \leq \lambda_{\text{cr}} < \frac{1}{r\alpha} \left( \frac{p_1 - \alpha^2}{p_1 - \alpha} - \frac{p_2 - \alpha^2}{p_2 - \alpha} \right)$
2	$\max\left\{0, \frac{1}{r(1-r)} \left[ \frac{p_h - \alpha^2}{\alpha(p_h - \alpha)} - \frac{2r}{1 - \sqrt{1-p_1}} - \frac{(1-r)(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)} \right]\right\} \leq \lambda_{\text{cr}} < \frac{1}{r} \left[ \frac{2}{1 - \sqrt{1-p_1}} - \frac{p_2 - \alpha^2}{\alpha(p_2 - \alpha)} \right]$
3	$\max\left\{0, \frac{1}{r(1-r)} \left[ \frac{2}{1 - \sqrt{1-p_h}} - \frac{2r}{1 - \sqrt{1-p_1}} - \frac{(1-r)(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)} \right]\right\} \leq \lambda_{\text{cr}} < \frac{1}{r} \left[ \frac{2}{1 - \sqrt{1-p_1}} - \frac{p_2 - \alpha^2}{\alpha(p_2 - \alpha)} \right]$
4	$\frac{2}{r(1-r)} \left( \frac{1}{1 - \sqrt{1-p_h}} - \frac{r}{1 - \sqrt{1-p_1}} - \frac{1-r}{1 - \sqrt{1-p_2}} \right) < \lambda_{\text{cr}} < \frac{2}{r} \left( \frac{1}{1 - \sqrt{1-p_1}} - \frac{1}{1 - \sqrt{1-p_2}} \right)$

## 2. Upper and lower limits to $\lambda_{\text{cr}}$

We first define  $\lambda = \lambda_{\text{cr}} \geq 0$ , as the value for which  $T_{\text{R}} = T_{\text{G}}$ . Note that  $\lambda_{\text{cr}}$  is defined to be equal to 0 if  $T_{\text{R}} \leq T_{\text{G}}$  when  $\lambda = 0$ . From the definition of  $\lambda_{\text{cr}}$ , the random-advantageous region appears when  $\lambda > \lambda_{\text{cr}}$ . Based on the discussions in Secs. IV B and IV C 1, we here evaluate  $\lambda_{\text{cr}}$  for  $S = 2$  and  $\alpha < \beta$ .

To take into account the sorting cost, we add the term  $\lambda K$  ( $\lambda > 0$ ) to  $T_{\text{G}}$ ; that is,

$$T_{\text{G}} \approx \lambda \bar{K}' + \sum_{s=1}^S \frac{r_s N}{Q_1(p = p_s)}. \quad (43)$$

Conversely, we do not add that term to  $T_{\text{R}}$  because a random sequence means a sequence without sorting.

We also define  $\lambda_{\text{H}}$  and  $\lambda_{\text{M}}$  as the values of  $\lambda$  for which  $T_{\text{H}} = T_{\text{G}}$  and  $T_{\text{M}} = T_{\text{G}}$ , respectively. Note that  $\lambda_{\text{H}}$  can have negative values, because  $T_{\text{H}}$  can be less than  $T_{\text{G}}$ .

From Eq. (16) and Sec. IV B, when  $\lambda = 0$  the relations among  $T_{\text{R}}$ ,  $T_{\text{G}}$ ,  $T_{\text{H}}$ , and  $T_{\text{M}}$  must satisfy one of the following three inequalities:

$$T_{\text{H}} < T_{\text{R}} \leq T_{\text{G}} < T_{\text{M}}, \quad (44)$$

$$T_{\text{H}} < T_{\text{G}} < T_{\text{R}} < T_{\text{M}}, \quad (45)$$

or

$$T_{\text{G}} < T_{\text{H}} < T_{\text{R}} < T_{\text{M}}. \quad (46)$$

Therefore, the relations of  $\lambda_{\text{cr}}$ ,  $\lambda_{\text{H}}$ , and  $\lambda_{\text{M}}$  can be written as follows:

$$\max(0, \lambda_{\text{H}}) \leq \lambda_{\text{cr}} < \lambda_{\text{M}}, \quad (47)$$

where we note that by definition  $\lambda_{\text{cr}} \geq 0$ , whereas  $\lambda_{\text{H}}$  can be either negative or positive, while  $\lambda_{\text{M}}$  must be positive.

In Region 1, i.e.,  $\alpha < 1 - \sqrt{1-p_1}$ , where  $T_{\text{H}} < T_{\text{G}}$  with  $\lambda = 0$ ,  $\lambda_{\text{H}}$  must be negative, while  $\lambda_{\text{M}}$  must be positive, and satisfy

$$\begin{aligned} \frac{N(p_1 - \alpha^2)}{\alpha(p_1 - \alpha)} &\approx \lambda_{\text{M}} r(1-r)N + \frac{rN(p_1 - \alpha^2)}{\alpha(p_1 - \alpha)} \\ &+ \frac{(1-r)N(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)}. \end{aligned} \quad (48)$$

In Region 2, i.e.,  $1 - \sqrt{1-p_1} \leq \alpha < 1 - \sqrt{1-p_h}$ ,  $\lambda_{\text{H}}$  can be either negative or positive, whereas  $\lambda_{\text{M}}$  must be

positive. The quantities  $\lambda_{\text{H}}$  and  $\lambda_{\text{M}}$  satisfy

$$\begin{aligned} \frac{rN(p_h - \alpha^2)}{\alpha(p_h - \alpha)} &\approx \lambda_{\text{H}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{(1-r)N(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)} \end{aligned} \quad (49)$$

and

$$\begin{aligned} \frac{2N}{1 - \sqrt{1-p_1}} &\approx \lambda_{\text{M}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{(1-r)N(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)}, \end{aligned} \quad (50)$$

respectively.

In Region 3, i.e.,  $1 - \sqrt{1-p_h} \leq \alpha < 1 - \sqrt{1-p_2}$ ,  $\lambda_{\text{H}}$  can be either negative or positive, and  $\lambda_{\text{M}}$  must be positive. Thus,  $\lambda_{\text{H}}$  and  $\lambda_{\text{M}}$  satisfy

$$\begin{aligned} \frac{2N}{1 - \sqrt{1-p_h}} &\approx \lambda_{\text{H}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{(1-r)N(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)} \end{aligned} \quad (51)$$

and

$$\begin{aligned} \frac{2N}{1 - \sqrt{1-p_1}} &\approx \lambda_{\text{M}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{(1-r)N(p_2 - \alpha^2)}{\alpha(p_2 - \alpha)}, \end{aligned} \quad (52)$$

respectively.

In Region 4, i.e.,  $1 - \sqrt{1-p_2} \leq \alpha$ , due to  $T_{\text{G}} < T_{\text{H}} < T_{\text{M}}$  when  $\lambda = 0$ ,  $\lambda_{\text{H}}$  and  $\lambda_{\text{M}}$  must both be positive. The quantities  $\lambda_{\text{H}}$  and  $\lambda_{\text{M}}$  therefore satisfy

$$\begin{aligned} \frac{2N}{1 - \sqrt{1-p_h}} &\approx \lambda_{\text{H}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{2(1-r)N}{1 - \sqrt{1-p_2}} \end{aligned} \quad (53)$$

and

$$\begin{aligned} \frac{2N}{1 - \sqrt{1-p_1}} &\approx \lambda_{\text{M}} r(1-r)N + \frac{2rN}{1 - \sqrt{1-p_1}} \\ &+ \frac{2(1-r)N}{1 - \sqrt{1-p_2}}, \end{aligned} \quad (54)$$

respectively.

Table V summarizes the upper and lower limits to  $\lambda_{\text{cr}}$  in each region. Furthermore, Fig. 18 shows the simulation



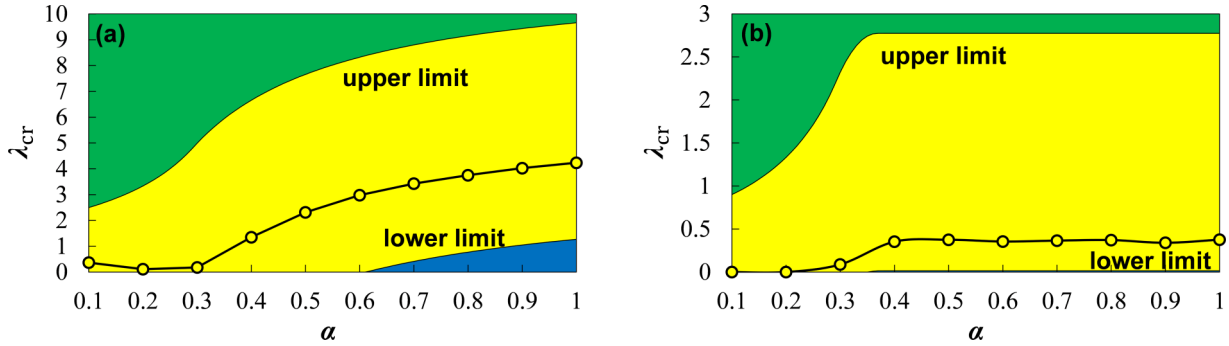


FIG. 18. Simulation values (black circles) and the theoretical existence range of  $\lambda_{\text{cr}}$  (yellow region) as functions of  $\alpha$ . The other parameters are fixed at  $(\beta; p_1, p_2; r) =$  (a)  $(1; 0.5, 1; 0.5)$  and (b)  $(1; 0.5, 0.6; 0.5)$ .

values (black circles) and the theoretical existence range of  $\lambda_{\text{cr}}$  (yellow region) as functions of  $\alpha$ . Note that we calculated the simulation values using with 10-trial-averaged values of  $T_R$ ,  $T_G$ , and  $K$ .

We can interpret Fig. 18 as demonstrating that a group (random) sequence is preferable in the region below (above) the black line. In the blue (green) region, a group (random) sequence is in fact theoretically verified to be preferable. Comparing Figs. 18(a) and 18(b), the simulation values approach the lower limit—i.e., the accuracy of approximating  $T_R$  by  $T_H$  increases—as  $|p_1 - p_2|$  decreases. We admit that the yellow region is extensive, especially when  $|p_1 - p_2|$  is relatively large; however, we emphasize that the simulation values always exist within the expected region and that the region can be limited easily without numeric calculations, which is convenient for applications to actual situations.

## V. CONCLUSION

In the present study, we have used a modified TASEP to analyze the dependence of the transportation time on the entering sequences of particles, using both the numerical simulations and theoretical analyses.

Here we summarize a number of important results. In Sec. III, we discovered that there exists an important “group-advantageous region” where  $T_R > T_G$  when  $\min(\alpha, \beta)$  is relatively large and the sorting costs are neglected. When sorting costs are introduced, a new region called a “random-advantageous region” appears with  $T_R < T_G$ . In addition, the group-advantageous region shrinks and finally disappears as  $\lambda$  increases. We explored these phenomena for various  $S \in \{2, 3, N\}$ .

In Sec. IV, we analyzed the simulation results by employing mathematical approaches for certain special cases. Using some approximations, we have shown theoretically that without the sorting cost the group-advantageous region must appear for any parameter sets  $(S, p_s, r_s)$ . Moreover, we have succeeded in deriving the upper and lower limits to the value of  $\lambda_{\text{cr}}$  where  $T_R = T_G$  by obtaining a general formula for the sorting cost.

Our findings can be applied to real-world situations, such as providing efficient operation for various tasks and smooth logistics for various products and yielding an effective evacuation method for pedestrians. Specifically, for smooth op-

eration, we can determine whether we should begin tasks without considering the operation sequence or otherwise. Similarly, for smooth logistics, we can select whether the products should be bunched with nearly equal sizes. Furthermore, for ensuring effective evacuation of pedestrians, we can determine whether the bunching of pedestrians having nearly equal velocities should be conducted before transportation. The criteria for these judgments depend on the magnitude of the consideration or bunching cost ( $\lambda$ ). Note that these magnitudes significantly differ from each other, i.e., considering only the sequence of tasks is typically deemed cheaper (have a smaller  $\lambda$ ) than sorting various pedestrians and products.

## ACKNOWLEDGMENTS

This work was partially supported by JST-Mirai Program Grant No. JPMJMI17D4, Japan, JSPS KAKENHI Grant No. JP15K17583, and MEXT as “Post-K Computer Exploratory Challenges” (Exploratory Challenge 2: Construction of Models for Interaction Among Multiple Socioeconomic Phenomena, Model Development and its Applications for Enabling Robust and Optimized Social Transportation Systems) (Project ID: hp180188).

## APPENDIX A: VALIDITY OF OUR SELECTION OF $L$ AND $N$

In this Appendix, we briefly discuss the validity of selecting  $L = 200$  and  $N = 10\,000$ .

As finite-size effects may occur for small  $L$ , we compare the simulation values of  $Q_2$  for  $L = 200$  and  $L = 1000$ . Figure 19 shows the ratio  $Q_2/Q'_2$ , where  $Q_2$  and  $Q'_2$  represent the flow of the multispecies TASEP with  $L = 200$  and  $L = 1000$ , respectively, as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$ . The result that  $Q_2/Q'_2 \approx 1$  indicates that the effect can be ignored for  $L = 200$ . Thus, we choose  $L = 200$  to decrease the simulation time.

On the other hand, the assumption that  $T$  is determined by a steady-state flow may be inappropriate for small  $N$ . Therefore, we have compared the results for  $N = 10\,000$  and  $N = 20\,000$ , in both cases for  $L = 200$ . Figure 20 shows the ratio  $T/T'$ , where  $T$  and  $T'$  represent the transportation times for  $N = 10\,000$  and  $N = 20\,000$ , respectively, as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$ . The result that  $T/T' \approx 0.5$ , i.e., that  $T$  is proportional to  $N$ , indicates that the assumption

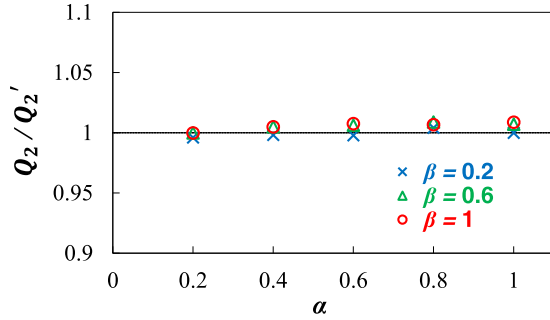


FIG. 19. Simulation values of the ratio  $Q_2/Q_2'$  as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$ . The other parameters are fixed at  $(p_1, p_2; r) = (0.5, 1; 0.5)$ .

can be regarded as valid for  $N = 10\,000$ . Thus, we choose  $N = 10\,000$  similarly to decrease the simulation time.

#### APPENDIX B: SIMULATION SCHEMES FOR OBTAINING THE MINIMAL NUMBER OF NECESSARY EXCHANGES

In this Appendix, we briefly describe the specific simulation schemes we used to obtain  $K$ . We emphasize that the cost of counting or comparing particles and the distances between exchanged particles are both ignored in the following.

First, for  $S = 2$ ,  $\tau_a = \tau_G$  can have only one of two patterns. Once  $\tau_G$  is fixed to be either of these two sequences, we can immediately obtain the number of particles placed at the wrong areas in sequence  $\tau_b = \tau_R$ , which is twice as large as the number of necessary exchanges (see also Appendix L). Consequently, comparing the results for the two  $\tau_G$  gives the smaller number as  $K$ .

Second, for  $S = 3$ ,  $\tau_a = \tau_G$  can have six patterns. Once  $\tau_G$  is fixed at one of these six sequences, we can immediately obtain the number of particles placed at the wrong areas in sequence  $\tau_b = \tau_R$ . After selecting one species, which we first replace at the correct location, we exchange all particles of that species that are placed in the wrong areas in sequence  $\tau_b = \tau_R$ . The subsequent procedure is similar to the case for  $S = 2$ . Consequently, comparing the six results for each  $\tau_G$  again gives the smallest number as  $K$ . Note that we can similarly calculate the numbers for general  $S > 4$ .

Finally, for  $S = N$ ,  $\tau_a = \tau_G$  can have one of two patterns: either an ascending or a descending sequence. One exchange

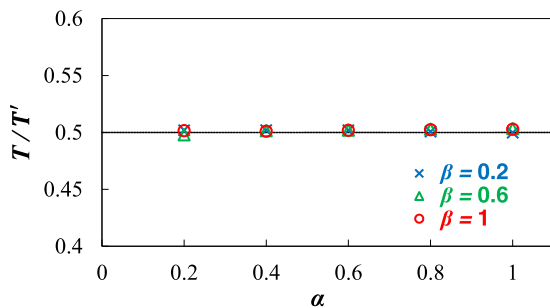


FIG. 20. Simulation values of the ratio  $T/T'$  as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$ . The other parameters are fixed at  $L = 200$  and  $(p_1, p_2; r) = (0.5, 1; 0.5)$ .

is needed for each particle in  $\tau_b = \tau_R$  for which there exists a particle with a smaller (larger) hopping probability than the noted particle. This is termed a “selection sort.” This procedure starts from the leading particle. Consequently, by comparing the results for the two  $\tau_G$ , the smaller number is again selected as the minimal number of necessary exchanges.

#### APPENDIX C: PROBABILITY DISTRIBUTION WITH $L = 2$ AND $S = 2$

Here, we summarize the probability distributions with  $L = 2$  and  $S = 2$ , which can be obtained from Eqs. (2) and (3). The specific forms are described as follows:

$$\begin{aligned}
 P_{00} &= \frac{p_1 p_2 (1 - \alpha) \beta^2}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{p_h (1 - \alpha) \beta^2}{A + p_h B} \\
 P_{0*} &= \frac{p_1 p_2 \alpha \beta}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{p_h \alpha \beta}{A + p_h B} \\
 P_{10} &= \frac{r p_2 A}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{r p_h A}{p_1 (A + p_h B)} \\
 P_{1*} &= \frac{r p_1 p_2 \alpha^2 (1 - \beta)}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{r p_h \alpha^2 (1 - \beta)}{A + p_h B} \\
 P_{20} &= \frac{(1 - r) p_1 A}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{(1 - r) p_h A}{p_2 (A + p_h B)} \\
 P_{2*} &= \frac{(1 - r) p_1 p_2 \alpha^2 (1 - \beta)}{\{(1 - r)p_1 + r p_2\}A + p_1 p_2 B} = \frac{(1 - r) p_h \alpha^2 (1 - \beta)}{A + p_h B},
 \end{aligned} \tag{C1}$$

where

$$A = \alpha \beta (\alpha + \beta - \alpha \beta), \tag{C2}$$

$$B = \alpha^2 + \beta^2 - \alpha^2 \beta - \alpha \beta^2 + \alpha \beta, \tag{C3}$$

and

$$p_h = \frac{p_1 p_2}{(1 - r)p_1 + r p_2}. \tag{C4}$$

#### APPENDIX D: $Q_S$ FOR GENERAL $S$ WITH $L = 2$

In this Appendix, we prove that for general  $S$  with  $L = 2$ ,  $Q_S$  is equal to  $Q_1(p = p_h)$ .

From the results with  $L = 2$  and  $S = 2$  (see Appendix C), we can conjecture the probability distributions for general  $S$  with  $L = 2$  as

$$\begin{aligned}
 P_{00} &= \frac{p_h (1 - \alpha) \beta^2}{A + p_h B} \\
 P_{0*} &= \frac{p_h \alpha \beta}{A + p_h B} \\
 P_{s0} &= \frac{r_s p_h A}{p_s (A + p_h B)} \\
 P_{s*} &= \frac{r_s p_h \alpha^2 (1 - \beta)}{A + p_h B},
 \end{aligned} \tag{D1}$$

where  $s = 1, 2, \dots, S$ .

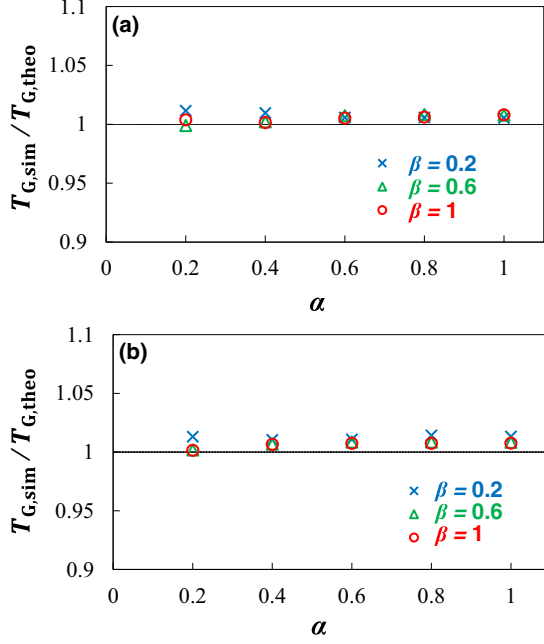


FIG. 21. Simulation values of the ratio  $T_{G,\text{sim}}/T_{G,\text{theo}}$  as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$  with (a)  $S = 2$  and (b)  $S = 3$ . The other parameters are fixed at (a)  $(p_1, p_2; r) = (0.5, 1; 0.5)$  and (b)  $(p_1, p_2, p_3; r_1, r_2, r_3) = (0.4, 0.6, 0.8; 0.2, 0.3, 0.5)$ .

On the other hand, the master equations of the steady state are summarized as  $2(S + 1)$  equations:

$$\begin{aligned} P_{00} &= (1 - \alpha)P_{00} + (1 - \alpha)\beta P_{0*} \\ P_{0*} &= (1 - \alpha)(1 - \beta)P_{0*} + \sum_{k=1}^S p_k P_{k0} \\ P_{s0} &= r_s \alpha P_{00} + r_s \alpha \beta P_{0*} + (1 - p_s)P_{s0} + \beta P_{s*} \\ P_{s*} &= (1 - \beta)P_{s*} + r_s \alpha (1 - \beta)P_{0*}, \end{aligned} \quad (\text{D2})$$

where  $s = 1, 2, \dots, S$ . In addition,  $P_{ij}$  must satisfy the normalization condition

$$\sum_{i=0}^S P_{i0} + \sum_{i=0}^S P_{i*} = 1. \quad (\text{D3})$$

We can confirm that Eqs. (D1) satisfy Eqs. (D2) and (D3). With Penron-Frobenius theorem regarding stochastic matrix, this indicates that Eqs. (D1) are unique solutions for Eqs. (D2) and (D3).

From Eqs. (D1), the flow of the system is given by the following expression:

$$Q_S = \sum_{s=1}^S p_s P_{s0} = \sum_{s=1}^S p_s \frac{r_s p_h A}{p_s (A + p_h B)} = Q_1(p = p_h). \quad (\text{D4})$$

#### APPENDIX E: VALIDITY OF THE APPROXIMATION FOR $T$

In this Appendix, we briefly demonstrate the validity of Eq. (13).

Figure 21(a) shows the ratio  $T_{G,\text{sim}}/T_{G,\text{theo}}$  as functions of  $\alpha$  for various  $\beta \in \{0.2, 0.6, 1\}$  with  $S = 2$  and Fig. 21(b) shows

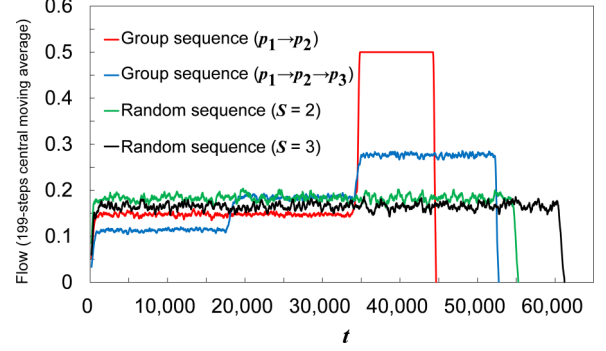


FIG. 22. Simulation values of 199-steps central moving average of flow at time  $t$  with  $(N, \alpha, \beta) = (10\,000, 1, 1)$ . For  $S = 2$  and  $S = 3$ , respectively, we set  $(p_1, p_2; r) = (0.5, 1; 0.5)$ ,  $(p_1, p_2, p_3; r_1, r_2, r_3) = (0.4, 0.6, 0.8; 0.2, 0.3, 0.5)$ . Note that each plot ends just after it reaches 0.

the same ratio for  $S = 3$ . Note that  $T_{G,\text{sim}}$  and  $T_{G,\text{theo}}$  represent the values of  $T_G$  from the simulations and that given by Eq. (13), respectively. Both figures show that  $T_{G,\text{sim}}/T_{G,\text{theo}} \approx 1$ , indicating that Eq. (13) provides a good approximation for  $T_G$ .

Strictly speaking,  $T_{G,\text{sim}}/T_{G,\text{theo}}$  must be larger than 1 on average. This is mainly due to the fact that  $T_{G,\text{sim}}$  includes  $T_1$ , which is the time required for the first particle to reach the right-hand boundary, whereas  $T_{G,\text{theo}}$  ignores that time. This also indicates that  $T_G$  can differ depending on the order of each group in the group sequence (i.e., the hopping probability of the leading group). However, this difference has little influence on the theoretical results, as explained below.

First,  $T_1$  can be estimated as

$$T_1 \approx \frac{L}{p_s}, \quad (\text{E1})$$

where  $s = 1, 2, \dots, S$  and the time steps before the first particle enters the lattice are assumed to be small enough to be ignored. The quantities  $T_1$  and  $T_G$  without  $T_1$  satisfy

$$T_1 \approx \frac{L}{p_s} < \frac{L}{p_s} < \frac{L}{1 - \sqrt{1 - p_s}} \quad (\text{E2})$$

and

$$T_G \approx \sum_{s=1}^S \frac{r_s N}{Q_1(p = p_s)} > \frac{2N}{1 - \sqrt{1 - p_s}}, \quad (\text{E3})$$

respectively. Therefore,  $T_1/T_G$  reduces to

$$\frac{T_1}{T_G} < \frac{L}{2N}. \quad (\text{E4})$$

Under the proposition that  $N$  is large enough, we can assume  $L/2N \ll 1$  ( $L/2N = 0.01$  in the present study). In fact, observing the time series of the flows (199-steps central moving average) in Fig. 22, we find that nearly the entire duration during transportation can be regarded to be in the steady state for large-enough  $r_s N$ . Note that we calculate the flows at time  $t$  by averaging number of moving particles per bond between  $(t - 1)$  and  $t$ . Moreover, all the transportation times  $T$  ( $T_R$ ,  $T_H$ , and  $T_M$ ) originally include  $T_1$ , so that this term disappears when they are subtracted from each other.

Consequently,  $T_1$  (and therefore, the dependence of  $T_G$  on the order of each group in the group sequence) can be assumed to be ignorable.

#### APPENDIX F: DISCUSSION OF THE SIGN OF $f(\alpha)$ IN REGION 1

In this Appendix, we give a detailed derivation of Eq. (22) for Region 1, where  $\alpha < 1 - \sqrt{1 - p_1}$ .

Equation (21) gives  $f(\alpha)$

$$\begin{aligned} f(\alpha) &= \frac{N(p_h - \alpha^2)}{\alpha(p_h - \alpha)} - \sum_{s=1}^S \frac{r_s N(p_s - \alpha^2)}{\alpha(p_s - \alpha)} \\ &= \frac{N}{\alpha} \sum_{s=1}^S \left[ \frac{r_s(p_h - \alpha^2)}{p_h - \alpha} - \frac{r_s(p_s - \alpha^2)}{p_s - \alpha} \right] \\ &= \frac{N(\alpha - 1)}{(p_h - \alpha) \prod_{s=1}^S (p_s - \alpha)} C, \end{aligned} \quad (\text{F1})$$

where

$$C = \sum_{s=1}^S \left[ r_s(p_h - p_s) \prod_{k \neq s} (p_k - \alpha) \right]. \quad (\text{F2})$$

The quantity  $C$  is calculated as follows:

$$\begin{aligned} C &= \sum_{s=1}^S \left[ r_s \left( \frac{1}{\sum_{t=1}^S r_t/p_t} - p_s \right) \prod_{k \neq s} (p_k - \alpha) \right] \\ &= \frac{1}{D} \sum_{s=1}^S \left\{ r_s \left[ \prod_{k=1}^S p_k - p_s \sum_{t=1}^S \left( r_t \prod_{k \neq t} p_k \right) \right] \prod_{k \neq s} (p_k - \alpha) \right\} \\ &= \frac{1}{D} \sum_{s=1}^S \sum_{t=1}^S \left[ r_s r_t \left( \prod_{k=1}^S p_k - p_s \prod_{k \neq t} p_k \right) \prod_{k \neq s} (p_k - \alpha) \right] \\ &= \frac{1}{D} \sum_{s=1}^S \sum_{t \neq s} \left[ r_s r_t (p_t - p_s) \prod_{k \neq t} p_k \prod_{k \neq s} (p_k - \alpha) \right], \end{aligned} \quad (\text{F3})$$

where

$$D = \sum_{s=1}^S \left( r_s \prod_{k \neq s} p_k \right). \quad (\text{F4})$$

By regarding the sum of the term with  $(s, t) = (x, y)$  and that with  $(s, t) = (y, x)$  as a new term for  $\exists(x, y)$  ( $x, y = 1, 2, \dots, S, x < y$ ), we can rewrite Eq. (F3) as follows:

$$\begin{aligned} C &= \frac{1}{D} \sum_{s=1}^S \sum_{t < s} \left[ r_s r_t (p_t - p_s) \prod_{k \neq t} p_k \prod_{k \neq s} (p_k - \alpha) \right. \\ &\quad \left. + r_t r_s (p_s - p_t) \prod_{k \neq s} p_k \prod_{k \neq t} (p_k - \alpha) \right] \\ &= \frac{1}{D} \sum_{s=1}^S \sum_{t < s} \left[ \alpha r_s r_t (p_t - p_s)^2 \prod_{k \neq s, t} p_k \prod_{k \neq s, t} (p_k - \alpha) \right]. \end{aligned} \quad (\text{F5})$$

Because  $D > 0$ ,  $p_s - \alpha > 0$  ( $\forall s$ ), and  $p_h - \alpha > 0$ , we obtain  $C > 0$ .

Considering  $\alpha - 1 < 0$  and  $C > 0$ , we finally obtain

$$f(\alpha) < 0. \quad (\text{F6})$$

#### APPENDIX G: CONTINUITY AND DIFFERENTIABILITY OF $f(\alpha)$ AT EACH BOUNDARY

In this Appendix, we briefly discuss the continuity and differentiability of  $f(\alpha)$  at each boundary.

Defining  $g(x)$  for  $0 < x \leq 1$  as

$$g(x) = \begin{cases} \frac{p-x^2}{x(p-x)} & \text{for } 0 < x \leq 1 - \sqrt{1-p} \\ \frac{2}{1-\sqrt{1-p}} & \text{for } 1 - \sqrt{1-p} < x \leq 1 \end{cases}, \quad (\text{G1})$$

where  $0 < p \leq 1$ , the following equations hold:

$$\lim_{x \rightarrow q-0} g(x) = \lim_{x \rightarrow q+0} g(x) = \frac{2}{1 - \sqrt{1-p}} \quad (\text{G2})$$

and

$$\lim_{\delta \rightarrow -0} \frac{g(x+\delta) - g(x)}{\delta} = \lim_{\delta \rightarrow +0} \frac{g(x+\delta) - g(x)}{\delta} = 0, \quad (\text{G3})$$

where  $q = 1 - \sqrt{1-p}$ . Therefore,  $g(x)$  is continuous and differentiable at  $x = q = 1 - \sqrt{1-p}$ , resulting in the continuity and differentiability of  $g(x)$  for  $0 < x \leq 1$ .

As a result, because  $f(\alpha)$  is represented as a linear sum of terms  $g(\alpha)$ , where  $p$  is substituted for  $p_s$  or  $p_h$  ( $0 < p_s, p_h \leq 1$ ),  $f(\alpha)$  is clearly continuous and differentiable at each boundary.

#### APPENDIX H: DISCUSSION OF THE SIGN OF $df(\alpha)/d\alpha$ IN SUBREGION 2-v

In this Appendix, we discuss the sign of  $df(\alpha)/d\alpha$  in Subregion 2-v, i.e.,  $1 - \sqrt{1-p_v} \leq \alpha < 1 - \sqrt{1-p_{v+1}}$

From Eq. (26),  $df(\alpha)/d\alpha$  can be calculated as follows:

$$\begin{aligned} \frac{df(\alpha)}{d\alpha} &\approx \frac{N(2\alpha - \alpha^2 - p_h)}{\alpha^2(p_h - \alpha)^2} - \sum_{s=u}^S \frac{r_s N(2\alpha - \alpha^2 - p_s)}{\alpha^2(p_s - \alpha)^2} \\ &= \sum_{s=1}^v \frac{r_s N(2\alpha - \alpha^2 - p_h)}{\alpha^2(p_h - \alpha)^2} \\ &\quad + \sum_{s=v+1}^S \left[ \frac{r_s N(2\alpha - \alpha^2 - p_h)}{\alpha^2(p_h - \alpha)^2} - \frac{r_s N(2\alpha - \alpha^2 - p_s)}{\alpha^2(p_s - \alpha)^2} \right] \\ &= \sum_{s=1}^v \frac{r_s N(2\alpha - \alpha^2 - p_h)}{\alpha^2(p_h - \alpha)^2} \\ &\quad + \sum_{s=v+1}^S \frac{r_s N(p_s - p_h)[(\alpha - 1)^2 + p_s + p_h - p_s p_h - 1]}{p_h p_s (p_h - \alpha)^2 (p_s - \alpha)^2}. \end{aligned} \quad (\text{H1})$$



For  $1 - \sqrt{1 - p_v} \leq \alpha < 1 - \sqrt{1 - p_{v+1}}$ , the following two inequalities hold:

$$2\alpha - \alpha^2 - p_h < 2(1 - \sqrt{1 - p_{v+1}}) - (1 - \sqrt{1 - p_{v+1}})^2 - p_h \\ = p_{v+1} - p_h < 0 \quad (\text{H2})$$

and

$$(\alpha - 1)^2 + p_s + p_h - p_s p_h - 1 \\ > (1 - \sqrt{1 - p_{v+1}} - 1)^2 + p_s + p_h - p_s p_h - 1 \\ > (1 - \sqrt{1 - p_h} - 1)^2 + p_s + p_h - p_s p_h - 1 \\ = p_s(1 - p_h) > 0. \quad (\text{H3})$$

We cannot specify the sign of  $df(\alpha)/d\alpha$  in this subregion from Eqs. (H1)–(H3). However, near the boundary between Subregions 2–( $u - 1$ ) and 3– $u$ , we obtain the following conditions:

$$\lim_{\alpha \rightarrow q_h - 0} (2\alpha - \alpha^2 - p_h) \\ = 2(1 - \sqrt{1 - p_h}) - (1 - \sqrt{1 - p_h})^2 - p_h = 0 \quad (\text{H4})$$

and

$$\lim_{\alpha \rightarrow q_h - 0} \{(\alpha - 1)^2 + p_s + p_h - p_s p_h - 1\} \\ = (1 - \sqrt{1 - p_h} - 1)^2 + p_{u-1} + p_h - p_{u-1} p_h - 1 \\ = p_{u-1}(1 - p_h) > 0, \quad (\text{H5})$$

where  $q_h = 1 - \sqrt{1 - p_h}$ . Therefore, noting the obvious continuity of  $df(\alpha)/d\alpha$  for  $1 - \sqrt{1 - p_v} \leq \alpha < 1 - \sqrt{1 - p_{v+1}}$ ,

the region of  $df(\alpha)/d\alpha > 0$  must exist at least in Subregion 2–( $u - 1$ ).

#### APPENDIX I: DISCUSSION OF THE SIGN OF $df(\alpha)/d\alpha$ IN SUBREGION 3– $v$

In this Appendix, we give a proof on Eq. (32) in Subregion 3– $v$ , i.e.,  $1 - \sqrt{1 - p_{v-1}} \leq \alpha < 1 - \sqrt{1 - p_v}$ .

From Eq. (31),  $df(\alpha)/d\alpha$  can be calculated as follows:

$$\frac{df(\alpha)}{d\alpha} \approx \sum_{s=v}^S \frac{r_s N (p_s - 2\alpha + \alpha^2)}{\alpha^2 (p_s - \alpha)^2}. \quad (\text{I1})$$

For  $s = v, \dots, S$ , the quantity  $p_s - 2\alpha + \alpha^2$  satisfies

$$p_s - 2\alpha + \alpha^2 > p_s - 2(1 - \sqrt{1 - p_v}) + (1 - \sqrt{1 - p_v})^2 \\ = p_s - p_v > 0. \quad (\text{I2})$$

From Eqs. (I1) and (I2), we finally obtain

$$\frac{df(\alpha)}{d\alpha} > 0. \quad (\text{I3})$$

#### APPENDIX J: DISCUSSION OF THE SIGN OF $f(\alpha)$ IN REGION 4

In this Appendix, we give a detailed derivation of Eq. (36), where  $1 - \sqrt{1 - p_S} \leq \alpha$ .

From Eqs. (33) and (34),  $f(\alpha)$  can be represented as follows:

$$T_G \approx 2N \sum_{s=1}^S \frac{r_s(1 + \sqrt{1 - p_s})}{p_s} = \frac{2N}{\prod_{s=1}^S p_s} \left[ \sum_{s=1}^S (r_s + r_s \sqrt{1 - p_s}) \prod_{t \neq s} p_t \right] \\ = \frac{2N}{\prod_{s=1}^S p_s} \left[ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) + \sum_{s=1}^S \left( r_s \sqrt{1 - p_s} \prod_{t \neq s} p_t \right) \right] \quad (\text{J1})$$

and

$$T_H \approx 2N \frac{1 + \sqrt{1 - p_h}}{p_h} = 2N \frac{1 + \sqrt{1 - 1/\sum_{s=1}^S (r_s/p_s)}}{1/\sum_{s=1}^S (r_s/p_s)} \\ = \frac{2N}{\prod_{s=1}^S p_s} \left\{ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) + \sqrt{\left[ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) \right]^2 - \prod_{s=1}^S p_s \times \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right)} \right\}. \quad (\text{J2})$$

From Eqs. (35), (J1), and (J2),  $f(\alpha)$  is given by

$$f(\alpha) \approx \frac{2N}{\prod_{s=1}^S p_s} \left\{ \sqrt{\left[ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) \right]^2 - \prod_{s=1}^S p_s \times \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right)} - \sum_{s=1}^S \left( r_s \sqrt{1 - p_s} \prod_{t \neq s} p_t \right) \right\} \\ = \frac{2N}{\prod_{s=1}^S p_s} (E - F), \quad (\text{J3})$$

where

$$E = \sqrt{\left[ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) \right]^2 - \prod_{s=1}^S p_s \times \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right)} \quad (\text{J4})$$

and

$$F = \sum_{s=1}^S \left( r_s \sqrt{1 - p_s} \prod_{t \neq s} p_t \right). \quad (\text{J5})$$

From Eqs. (J4) and (J5),  $E^2 - F^2$  becomes

$$\begin{aligned} E^2 - F^2 &= \left[ \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) \right]^2 - \prod_{s=1}^S p_s \times \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) - \left[ \sum_{s=1}^S \left( r_s \sqrt{1 - p_s} \prod_{t \neq s} p_t \right) \right]^2 \\ &= \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right)^2 - \sum_{s=1}^S \sum_{t \neq s} \left( r_s r_t \prod_{k \neq s} p_k \prod_{l \neq t} p_l \right) - \prod_{s=1}^S p_s \times \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right) \\ &\quad - \sum_{s=1}^S \left( r_s \prod_{t \neq s} p_t \right)^2 + \sum_{s=1}^S \left[ r_s^2 p_s \left( \prod_{t \neq s} p_t \right)^2 \right] - \sum_{s=1}^S \sum_{t \neq s} \left( r_s r_t \sqrt{1 - p_s} \sqrt{1 - p_t} \prod_{k \neq s} p_k \prod_{l \neq t} p_l \right) \\ &= \sum_{s=1}^S \left[ \sum_{t \neq s} r_s r_t \prod_{k \neq s} p_k \prod_{l \neq t} p_l - r_s (1 - r_s) p_t \prod_{k \neq s} p_k \prod_{l \neq t} p_l - \sum_{t \neq s} \left( r_s r_t \sqrt{1 - p_s} \sqrt{1 - p_t} \prod_{k \neq s} p_k \prod_{l \neq t} p_l \right) \right] \\ &= \sum_{s=1}^S \sum_{t \neq s} \left\{ r_s r_t \prod_{k \neq s} p_k \prod_{l \neq t} p_l [1 - p_t - \sqrt{(1 - p_s)(1 - p_t)}] \right\}. \quad (\text{J6}) \end{aligned}$$

Here, regarding the sum of the term with  $(s, t) = (x, y)$  and that with  $(s, t) = (y, x)$  as a new term for  $\Xi(x, y)$  ( $x, y = 1, 2, \dots, S, x < y$ ), Eq. (J6) can be rewritten as

$$\begin{aligned} E^2 - F^2 &= \sum_{s=1}^S \sum_{t < s} \left\{ r_s r_t \prod_{k \neq s} p_k \prod_{l \neq t} p_l [2 - p_s - p_t - \sqrt{(1 - p_s)(1 - p_t)}] \right\} \\ &= \sum_{s=1}^S \sum_{t < s} \left[ r_s r_t (\sqrt{1 - p_s} - \sqrt{1 - p_t})^2 \prod_{k \neq s} p_k \prod_{l \neq t} p_l \right]. \quad (\text{J7}) \end{aligned}$$

Due to Eq. (J7) and the non-negativity of both  $E$  and  $F$ , we have  $E > F$ , thereby resulting in

$$f(\alpha) > 0. \quad (\text{J8})$$

#### APPENDIX K: SPECIFIC CONDITIONS ON $\alpha_{\text{cr,max}}$

Here, we discuss the specific conditions on  $\alpha_{\text{cr,max}}$ .

Table VI summarizes explicit expressions for  $f(\alpha = \alpha_{\text{cr}}) = 0$ , where the upper (lower) expression holds in Region 2 (Region 3). For  $S = 2$ , the lower expression becomes a quadratic equation in  $\alpha_{\text{cr}}$ . However, the upper expression becomes a quartic equation that is too difficult to solve analytically those conditions. Note that for  $S > 2$ , both equations become more than quartic.

From its definition of  $\alpha_{\text{cr,max}}$ ,  $\alpha_{\text{cr,max}}$  can be written as

$$\alpha_{\text{cr,max}} = \max\{\alpha_{\text{cr}}\}, \quad (\text{K1})$$

where  $\{\alpha_{\text{cr}}\}$  represents the set of  $\alpha_{\text{cr}}$ .

#### APPENDIX L: DERIVATION OF $\bar{K}'$

In this Appendix, we derive the approximate averaged minimal number of exchanges  $\bar{K}'$  necessary to sort the particles for two special cases:  $S = 2$  and  $S = N$ .

##### 1. $S = 2$

First, for a general calculation of  $\bar{K}'$ ,  $\tau_G$  has to be fixed to be either of the two possible patterns. Once  $\tau_G$  is fixed to  $\tilde{\tau}_G$ ,  $K'(\tau_G, \tau_R)$  can be determined uniquely for all possible  $\tau_R$ . Without loss of generality, we can assume  $rN \leq (1 - r)N$  and  $\tau_G$  can be fixed as illustrated in the lower panel of Fig. 23.

TABLE VI. Explicit expressions for  $f(\alpha = \alpha_{\text{cr}}) = 0$ .

Region No.	Explicit expressions for $f(\alpha = \alpha_{\text{cr}}) = 0$
2	$\frac{N(p_h - \alpha_{\text{cr}}^2)}{\alpha_{\text{cr}}(p_h - \alpha_{\text{cr}})} - \sum_{s=1}^v \frac{2r_s N}{1 - \sqrt{1 - p_s}} - \sum_{s=v+1}^S \frac{r_s N(p_s - \alpha_{\text{cr}}^2)}{\alpha_{\text{cr}}(p_s - \alpha_{\text{cr}})} = 0$ $\wedge \begin{cases} 1 - \sqrt{1 - p_v} \leq \alpha_{\text{cr}} < 1 - \sqrt{1 - p_{v+1}} & \text{for } 1 \leq v \leq u - 2 \\ 1 - \sqrt{1 - p_v} \leq \alpha_{\text{cr}} < 1 - \sqrt{1 - p_h} & \text{for } v = u - 1 \end{cases}$
3	$\frac{2N}{1 - \sqrt{1 - p_h}} - \sum_{s=1}^{v-1} \frac{2r_s N}{1 - \sqrt{1 - p_s}} - \sum_{s=v}^S \frac{r_s N(p_s - \alpha_{\text{cr}}^2)}{\alpha_{\text{cr}}(p_s - \alpha_{\text{cr}})} = 0$ $\wedge \begin{cases} 1 - \sqrt{1 - p_h} \leq \alpha_{\text{cr}} < 1 - \sqrt{1 - p_v} & \text{for } v = u \\ 1 - \sqrt{1 - p_{v-1}} \leq \alpha_{\text{cr}} < 1 - \sqrt{1 - p_v} & \text{for } u + 1 \leq v \leq S \end{cases}$

Suppose that for  $\tau_R$ ,  $k$  ( $0 \leq k \leq rN$ ) particles of species 1 are located in the Area 2 ( $k$  particles of species 2 are located in the Area 1, conversely) as described in the lower of Fig. 23. Under this supposition,  $k$ -time exchanges are necessary for sorting particles from  $\tau_R$  to  $\tau_G$ . Considering that  $\tau_R$  satisfying this supposition possibly has  $\binom{rN}{k} \times \binom{(1-r)N}{k}$  sequences,  $a_N = \sum_{\forall \tau_R} K'(\tau_G, \tau_R)$  can be written as follows;

$$\begin{aligned}
 a_N &= \sum_{\forall \tau_R} K'(\tau_G, \tau_R) \\
 &= \sum_{k=1}^{rN} k \binom{rN}{k} \binom{(1-r)N}{k} \\
 &= \sum_{k=1}^{rN} rN \binom{rN-1}{k-1} \binom{(1-r)N}{k} \\
 &= rN \sum_{k=1}^{rN} \left[ \binom{rN}{k} \binom{(1-r)N}{k} - \binom{rN-1}{k} \binom{(1-r)N}{k} \right].
 \end{aligned} \tag{L1}$$

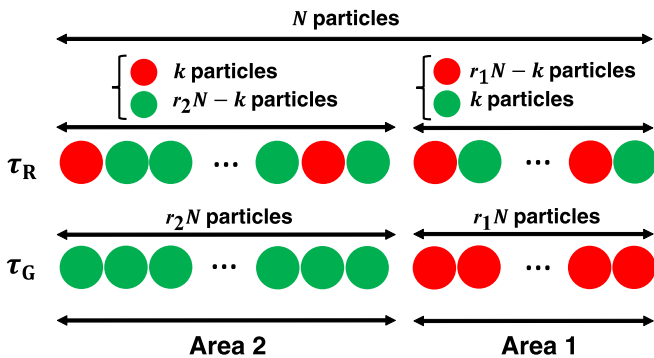


FIG. 23. Schematic illustration of  $\tau_R$  (upper panel) and  $\tau_G$  (lower panel), where the red particles belong to species 1 and the green ones to species 2. In the upper panel, we show one example from among all  $\binom{rN}{k} \times \binom{(1-r)N}{k}$  possible random sequences, whereas in the lower panel we show one of the two possible group sequences. Note that  $r_1 = r$ ,  $r_2 = 1 - r$ .

Using the Vandermonde convolution formula, Eq. (L1) can be rewritten as follows:

$$a_N = rN \left[ \binom{N}{rN} - \binom{N-1}{rN} \right]. \tag{L2}$$

Because the sequence  $\tau_R$  can take any of  $N!/\{(rN)!((1-r)N)!\}$  possible patterns with equal probability, we can finally reduce  $\bar{K}'$  to

$$\bar{K}' = \frac{(rN)![(1-r)N]!}{N!} a_N = r(1-r)N. \tag{L3}$$

Figure 24 compares the simulation (circles) and theoretical (curves) values for various  $N \in \{1000$  (red),  $5000$  (green),  $10\,000$  (blue) $\}$  for  $S = 2$ . The simulations show very good agreement with our exact analysis.

## 2. $S = N$

When  $S = N$ ,  $\tau_G$  also has to be fixed as either of the two possible patterns—an ascending or a descending sequence—for a general calculation of  $\bar{K}'$ , as illustrated in the upper panel of Fig. 25. Once  $\tau_G$  is fixed to  $\bar{\tau}_G$ ,  $K'(\tau_G, \tau_R)$  can be determined uniquely for all possible  $\tau_R$ .

If we regard the entire sequence as consisting of two parts—the first (blue) particle and other  $(N-1)$  particles, as described in the lower panel of Fig. 25—the sorting procedure can also be divided into two parts: sorting  $(N-1)$  particles plus the last exchange for the first particle. If the first particle

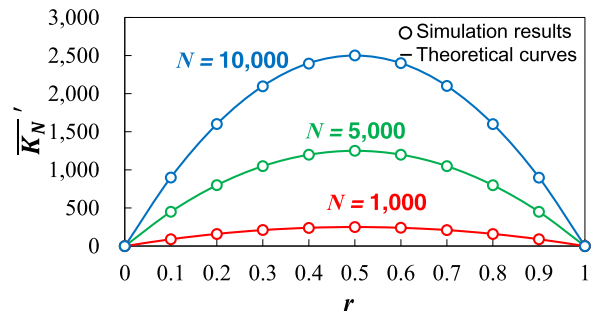


FIG. 24. Simulation (circles) and theoretical (curve) values of  $\bar{K}'$  as functions of  $r$  for various  $N \in \{1000$  (red),  $5000$  (green),  $10\,000$  (blue) $\}$  with  $S = 2$ . We obtained each of the simulation values by averaging over 100 trials.

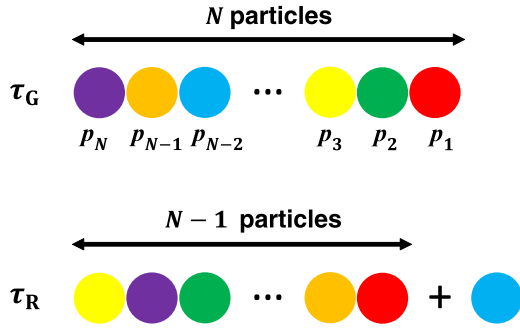


FIG. 25. Schematic illustration of  $\tau_G$  (upper panel) and  $\tau_R$  (lower panel) for the case  $S = N$ . In the lower panel, we show one example of all  $N \times (N - 1)!$  possible sequences. Note that  $p_1 < p_2 < \dots < p_N$  for the ascending sequence, whereas  $p_1 > p_2 > \dots > p_N$  for the descending one.

corresponds to the particle with hopping probability  $p_l$  ( $l = 1, 2, \dots, N$ ), and noting that the sequence for the remaining  $(N - 1)$  particles has  $(N - 1)!$  possible patterns, we can calculate the quantity  $b_{N,l} = \sum_{\forall \tau'_{R,l}} K'(\tilde{\tau}_G, \tau'_{R,l})$  as follows:

$$b_{N,l} = \sum_{\forall \tau'_{R,l}} K'(\tilde{\tau}_G, \tau'_{R,l}) = \begin{cases} a_{N-1} & \text{for } l = 1, \\ a_{N-1} + (N - 1)! & \text{for } l = 2, 3, \dots, N, \end{cases} \quad (\text{L4})$$

where  $N > 1$  and  $\tau'_{R,l}$  represents the sequence for which the first particle is the particle with hopping probability  $p_l$ . Note that the last sort is not necessary in the case where  $l = 1$ .

Therefore, for  $N > 1$ , we can write  $a_N = \sum_{\forall \tau_R} K'(\tilde{\tau}_G, \tau_R)$ :

$$a_N = \sum_{\forall \tau_R} K'(\tilde{\tau}_G, \tau_R) = \sum_{l=1}^N \sum_{\forall \tau'_{R,l}} K'(\tilde{\tau}_G, \tau'_{R,l})$$

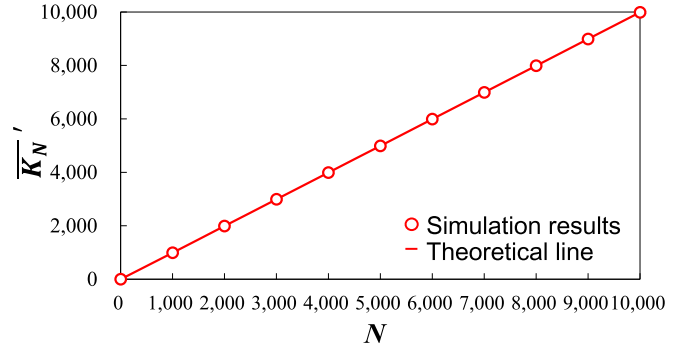


FIG. 26. Simulation (circles) and theoretical (line) values of  $\overline{K'_N}$  as a function of  $N$ . We obtained each of the simulation values by averaging over 100 trials.

$$= \sum_{l=1}^N b_{N,l} = (N - 1) \times (N - 1)! + Na_{N-1}. \quad (\text{L5})$$

Dividing both sides of Eq. (L5) by  $N!$ , we obtain

$$c_N = c_{N-1} + \frac{N - 1}{N} = c_1 + \sum_{k=1}^N \frac{k - 1}{k}, \quad (\text{L6})$$

where  $c_N = a_N/N!$  and  $N > 1$ . With the initial condition  $c_1 = a_1 = 0$ ,  $a_N$  is finally reduced to

$$a_N = N! \left( N - \sum_{k=1}^N \frac{1}{k} \right), \quad (\text{L7})$$

which we note holds for the case  $N = 1$ .

The sequence  $\tau_R$  can take  $N!$  patterns with equal probability, and therefore,  $\overline{K'}$  is finally reduced to

$$\overline{K'} = \frac{a_N}{N!} = N - \sum_{k=1}^N \frac{1}{k}. \quad (\text{L8})$$

Figure 26 compares the simulation (circles) and theoretical (line) values for  $S = N$ . The simulations again show a very good agreement with our exact analysis.

[1] C. Appert-Rolland, S. Klein, M. Ebbinghaus, and L. Santen, in *Proceedings of the Asia-Pacific Econophysics Conference 2016 Big Data Analysis and Modeling toward Super Smart Society (APEC-SSS2016)* (Physical Society of Japan, Tokyo, Japan, 2017), p. 011001.

[2] S. Croom, P. Romano, and M. Giannakis, *Eur. J. Purch. Supply Manag.* **6**, 67 (2000).

[3] T. Ezaki, D. Yanagisawa, and K. Nishinari, *Physica A* **427**, 62 (2015).

[4] Y. Taniguchi, R. Nishi, T. Ezaki, and K. Nishinari, *Physica A* **433**, 304 (2015).

[5] M. Woelki, *Phys. Rev. E* **87**, 062818 (2013).

[6] T. Imai and K. Nishinari, *Phys. Rev. E* **91**, 062818 (2015).

[7] H. Yamamoto, D. Yanagisawa, and K. Nishinari, *J. Stat. Mech.* (2017) 043204.

[8] K. Christensen and Y. Sasaki, *J. Artif. Soc. Soc. Simul.* **11**, 9 (2008).

[9] N. Pelechano and A. Malkawi, *Automat. Constr.* **17**, 377 (2008).

[10] D. Yanagisawa, A. Kimura, A. Tomoeda, R. Nishi, Y. Suma, K. Ohtsuka, and K. Nishinari, *Phys. Rev. E* **80**, 036110 (2009).

[11] T. Ezaki, D. Yanagisawa, K. Ohtsuka, and K. Nishinari, *Physica A* **391**, 291 (2012).

[12] J. E. Rothman, *Nature* **372**, 55 (1994).

[13] M. J. Müller, S. Klumpp, and R. Lipowsky, *Proc. Natl. Acad. Sci. USA* **105**, 4609 (2008).

[14] P. C. Bressloff and J. M. Newby, *Rev. Mod. Phys.* **85**, 135 (2013).



- [15] C. T. MacDonald, J. H. Gibbs, and A. C. Pipkin, *Biopolymers* **6**, 1 (1968).
- [16] C. T. MacDonald and J. H. Gibbs, *Biopolymers* **7**, 707 (1969).
- [17] A. Parmeggiani, T. Franosch, and E. Frey, *Phys. Rev. Lett.* **90**, 086601 (2003).
- [18] D. Chowdhury, A. Schadschneider, and K. Nishinari, *Phys. Life Rev.* **2**, 318 (2005).
- [19] T. Chou, K. Mallick, and R. Zia, *Rep. Prog. Phys.* **74**, 116601 (2011).
- [20] C. Appert-Rolland, M. Ebbinghaus, and L. Santen, *Phys. Rep.* **593**, 1 (2015).
- [21] D. Helbing, *Rev. Mod. Phys.* **73**, 1067 (2001).
- [22] H. Ito and K. Nishinari, *Phys. Rev. E* **89**, 042813 (2014).
- [23] S. Tsuzuki, D. Yanagisawa, and K. Nishinari, *Phys. Rev. E* **97**, 042117 (2018).
- [24] D. Yanagisawa, A. Tomoeda, R. Jiang, and K. Nishinari, *JSIAM Lett.* **2**, 61 (2010).
- [25] C. Arita and D. Yanagisawa, *J. Stat. Phys.* **141**, 829 (2010).
- [26] C. Arita and A. Schadschneider, *Math. Mod. Meth.* **25**, 401 (2015).
- [27] J. de Gier and B. Nienhuis, *Phys. Rev. E* **59**, 4899 (1999).
- [28] R. A. Blythe and M. R. Evans, *J. Phys. A* **40**, R333 (2007).
- [29] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, *J. Phys. A* **26**, 1493 (1993).
- [30] T. Minemura, K. Nishinari, and A. Schadschneider, in *Proceedings of the International Conference on Cellular Automata* (Springer, Berlin, 2010), pp. 593–599.
- [31] D. A. Adams, B. Schmittmann, and R. K. P. Zia, *J. Stat. Mech.* (2008) P06009.
- [32] L. J. Cook, R. K. P. Zia, and B. Schmittmann, *Phys. Rev. E* **80**, 031142 (2009).
- [33] L. J. Cook and R. K. P. Zia, *J. Stat. Mech.* (2009) P02012.
- [34] L. J. Cook and R. K. P. Zia, *J. Stat. Mech.* (2012) P05008.
- [35] O. Angel, *J. Comb. Theory A* **113**, 625 (2006).
- [36] P. A. Ferrari, L. R. Fontes, and Y. Kohayakawa, *J. Stat. Phys.* **76**, 1153 (1994).
- [37] C. Arita, *J. Stat. Mech.* (2006) P12008.
- [38] K. Mallick, *J. Phys. A* **29**, 5375 (1996).
- [39] B. Derrida and M. Evans, *J. Phys. A* **32**, 4833 (1999).
- [40] A. Ayyer, J. L. Lebowitz, and E. R. Speer, *J. Stat. Phys.* **135**, 1009 (2009).
- [41] V. Privman, *Nonequilibrium Statistical Mechanics in One Dimension* (Cambridge University Press, Cambridge, 2005).
- [42] P. A. Ferrari, J. B. Martin *et al.*, *Ann. Probab.* **35**, 807 (2007).
- [43] L. Cantini, *J. Phys. A* **41**, 095001 (2008).
- [44] M. Uchiyama, *Chaos Soliton Fract.* **35**, 398 (2008).
- [45] N. Crampe, K. Mallick, E. Ragoucy, and M. Vanicat, *J. Phys. A* **48**, 175002 (2015).
- [46] B. Aneva, *Eur. Phys. J. C* **31**, 403 (2003).
- [47] C. Arita, A. Kuniba, K. Sakai, and T. Sawabe, *J. Phys. A* **42**, 345002 (2009).
- [48] E. Duchi and G. Schaeffer, *J. Comb. Theory A* **110**, 1 (2005).
- [49] M. Evans, *J. Phys. A* **30**, 5669 (1997).
- [50] J. Krug and P. A. Ferrari, *J. Phys. A* **29**, L465 (1996).
- [51] K. Mallick, S. Mallick, and N. Rajewsky, *J. Phys. A* **32**, 8399 (1999).
- [52] N. Crampe, C. Finn, E. Ragoucy, and M. Vanicat, *J. Phys. A* **49**, 375201 (2016).
- [53] C. Arita and K. Mallick, *J. Phys. A* **46**, 085002 (2013).
- [54] S. Prolhac, M. R. Evans, and K. Mallick, *J. Phys. A* **42**, 165004 (2009).
- [55] A. Ayyer and D. Roy, *Sci. Rep.* **7**, 13555 (2017).
- [56] P. F. Arndt, T. Heinzel, and V. Rittenberg, *J. Phys. A* **31**, L45 (1998).
- [57] M. Evans, *Europhys. Lett.* **36**, 13 (1996).
- [58] T. Sasamoto, *Phys. Rev. E* **61**, 4980 (2000).
- [59] M. Bengrine, A. Benyoussef, H. Ez-Zahraouy, J. Krug, M. Loulidi, and F. Mhirech, *J. Phys. A* **32**, 2527 (1999).
- [60] E. Duchi and G. Schaeffer, *Random Struct. Algor.* **33**, 434 (2008).
- [61] To be precise, if we regard vacant sites as particle 0, swapping between particle 0 and other kinds of particles (1, 2, 3, ...) occurs. However, swapping between particles 1, 2, 3, ... does not occur.
- [62] This system can be simulated also with random updating; however, we adopt parallel updating for the following three reasons. First, we consider that the results will be mainly applied to real systems, where particles and agents move simultaneously. It is generally natural to adopt parallel updating in the contexts of pedestrian dynamics or traffic flows [63,64]. Second, it is more natural to consider that time for swapping two particles is compatible with a one-time step with parallel updating than with random-sequential updating. Finally, the simulation times can be suppressed drastically with parallel updating compared to with random updating.
- [63] K. Nagel and M. Schreckenberg, *J. Phys. I (France)* **2**, 2221 (1992).
- [64] C. Burstedde, K. Klauk, A. Schadschneider, and J. Zittartz, *Physica A* **295**, 507 (2001).
- [65] M. R. Evans, N. Rajewsky, and E. R. Speer, *J. Stat. Phys.* **95**, 45 (1999).