

# CUTOFF PROFILE OF THE METROPOLIS BIASED CARD SHUFFLING

BY LINGFU ZHANG<sup>a</sup>

Department of Statistics, University of California, <sup>a</sup>lfzhang@berkeley.edu

We consider the Metropolis biased card shuffling (also called the multi-species ASEP on a finite interval or the random Metropolis scan). Its convergence to stationarity was believed to exhibit a total-variation cutoff, and that was proved a few years ago by Labb   and Lacoin (*Ann. Probab.* **47** (2019) 1541–1586). In this paper, we prove that (for  $N$  cards) the cutoff window is in the order of  $N^{1/3}$ , and the cutoff profile is given by the Tracy–Widom GOE distribution function. This confirms a conjecture by Bufetov and Nejjar (*Probab. Theory Related Fields* **83** (2022) 229–253). Our approach is different from (*Ann. Probab.* **47** (2019) 1541–1586), by comparing the card shuffling with the multispecies ASEP on  $\mathbb{Z}$ , and using Hecke algebra and recent ASEP shift-invariance and convergence results. Our result can also be viewed as a generalization of the Oriented Swap Process finishing time convergence (*Ann. Appl. Probab.* **32** (2022) 753–763), which is the TASEP version (of our result).

**1. Introduction.** In this paper, we study the (*Metropolis*) *biased card shuffling* of size  $N$ , which can be viewed as a continuous time Markov chain with state space being the permutation group  $S_N$ . This Markov chain can be described as follows. Let  $[N]$  denote the set  $\{1, \dots, N\}$ , and let elements in  $S_N$  be represented by bijections from  $[N]$  to itself. Suppose that the current state is  $\lambda : [N] \rightarrow [N]$ , then for each  $i \in [N - 1]$  independently, if  $\lambda(i) < \lambda(i + 1)$ , with rate 1 we swap  $\lambda(i)$  and  $\lambda(i + 1)$ ; otherwise, with rate  $q$  we swap  $\lambda(i)$  and  $\lambda(i + 1)$ . Here,  $q \in [0, 1)$  is a constant. This Markov chain is known to have a stationary measure,  $\mathcal{M}_N$ , which is also called the *Mallows measure* of size  $N$ . We let  $\mathcal{W}_{N,t}^\lambda$  be the law of this chain at time  $t$ , starting from a deterministic state  $\lambda$ . Then  $\mathcal{W}_{N,t}^\lambda \rightarrow \mathcal{M}_N$  as  $t \rightarrow \infty$ . Our main result concerns the convergence of  $\mathcal{W}_{N,t}^\lambda$  to  $\mathcal{M}_N$  in the total-variation distance, which is defined as

$$\|\mathcal{P} - \mathcal{Q}\|_{\text{TV}} = \max_{A \subset S_N} |\mathcal{P}(A) - \mathcal{Q}(A)|$$

for any two probability measures  $\mathcal{P}$  and  $\mathcal{Q}$  of  $S_N$ .

**THEOREM 1.1.** *For any  $\tau \in \mathbb{R}$ , we have*

$$\lim_{N \rightarrow \infty} \max_{\lambda \in S_N} \|\mathcal{W}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^\lambda - \mathcal{M}_N\|_{\text{TV}} = 1 - F_{\text{GOE}}(2^{2/3}\tau),$$

where  $F_{\text{GOE}}$  is the distribution function of the Tracy–Widom GOE distribution.

The Tracy–Widom GOE distribution was introduced in [47], where it was shown to govern the fluctuation of the largest eigenvalue in the Gaussian Orthogonal Ensemble (GOE) random matrices. It can be defined as

$$(1.1) \quad F_{\text{GOE}}(s) = e^{-2^{-1} \int_s^\infty p(x) + (x-s)^2 p(x)^2 dx},$$

Received January 2023; revised September 2023.

**MSC2020 subject classifications.** Primary 60K35; secondary 60C05, 80C22.

**Key words and phrases.** Cutoff profile, biased card shuffling, Markov chain mixing, simple exclusion process, Tracy–Widom distribution, KPZ universality, Hecke algebra.

where  $p(x)$  is the Hastings–McLeod solution of the second Plainlevé equation  $p''(x) = 2p(x)^3 + xp(x)$ , with boundary condition  $p(x) \sim \text{Ai}(x)$  as  $x \rightarrow \infty$  for  $\text{Ai}$  being the Airy function. The appearance of this distribution here is due to that it also comes up in current fluctuations of 1D interacting particle systems (to be discussed shortly).

Our main result (Theorem 1.1) basically says that there is a cutoff phenomenon in the total-variation convergence of the biased card shuffling, with a cutoff window in the order of  $N^{1/3}$ , and the cutoff profile is Tracy–Widom GOE.

**1.1. Background and related works.** The word “cutoff” was first used by Aldous and Diaconis [4] to describe the phenomenon, where (usually in terms of the total-variation distance) a Markov chain stays away from its stationary distribution for some time, but then converges abruptly; see the review articles [18, 44] and the textbook [32]. To rigorously establish cutoff often requires a rather careful understanding of the particular Markov chains, as achieved in many instances in the past decades see, for example, [3, 7, 20, 21, 27, 34–36, 42] where cutoff was proved for various (families of) Markov chains using a variety of methods. Many of them were on different card shuffling processes.

The “cutoff window” refers to the small period when the total-variation distance decays from near 1 to near 0, and the “cutoff profile” describes the shape of the decay within the window. These are refined information on Markov chains and are not available even for most Markov chains where cutoffs have been established. Among those cases where cutoff profiles are known, many of them are given by the distribution function of Gaussian random variables; see, for example, [8, 30, 34] (and also [17] where the cutoff is in a slightly different setting). Some other types of cutoff profiles have also been discovered; see, for example, [15, 38, 45].

The biased card shuffling, the Markov chain we analyze in this paper, was first studied by Diaconis and Ram in [19] under the name “random Metropolis scan,” as it is an instance of the Metropolis algorithm that samples biased random permutations. The mixing of this Markov chain was later studied in [9], whose main result states that the mixing time is at most linear in the size  $N$ . A simpler proof of this result later appeared in [25]. Since then, whether a cutoff happens for the biased card shuffling remained an open problem. This was answered affirmatively by Labb   and Lacoin [27], who proved that the total-variation cutoff happens at time  $2(1 - q)^{-1}N$ . In the same paper, the spectral gap was also computed.

**Simple exclusion process.** A closely related Markov chain is the *Asymmetric Simple Exclusion Process* (ASEP), which has been used as a key tool in studying biased card shuffling. It is a classical interacting particle system and has been intensively studied in the literature (back to, e.g., [33]). It can be formulated as follows. For a (finite or infinite) sequence of sites (indexed by  $\mathbb{Z}$  or a finite discrete interval), each site may be occupied by a particle or be empty (later we shall represent this by a function from  $\mathbb{Z}$  or a finite discrete interval to  $\{0, 1\}$ , with 1 denoting a particle, and 0 denoting a hole). For each particle, if the site next to it in the right is empty, with rate 1 it jumps to the right; and if the site next to it in the left is empty, with rate  $q$  it jumps to the left. Such jumps happen independently for each particle in each direction. The ASEP is an important model in the so-called Kardar–Parisi–Zhang (KPZ) universality class and has exact-solvable structures. The hydrodynamics are well understood, and its diffusive scaling limit has been identified with the KPZ equation [10]. However, the scaling invariant and long time limit of the ASEP should be obtained under the KPZ scaling instead. That limit (when the space is  $\mathbb{Z}$ ) is termed the *KPZ fixed point*, and was first constructed in [37] as the KPZ scaling limit in the *totally asymmetric* case (i.e., when  $q = 0$ ), and later proved for the general  $q \in [0, 1]$  case in [41].

The ASEP on the finite discrete interval  $[N]$  can be easily connected to the biased card shuffling of size  $N$ . Namely, we can get the ASEP on  $[N]$  with  $k$  particles, if in the biased card shuffling we replace each number  $\leq k$  by a particle and each number  $> k$  by a hole (see

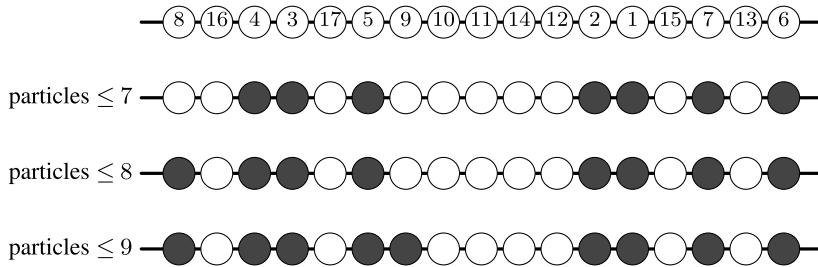


FIG. 1. *An illustration of projecting the biased card shuffling into single-species ASEP.*

Figure 1). In other words, we can project the biased card shuffling of size  $N$  into the ASEP, for each integer  $k$  between 0 and  $N$ . It is worth mentioning that given all such projections (for all  $k$ ) we can reconstruct the biased card shuffling (see Section 2.2 below for the details). For this reason, we also refer to the biased card shuffling of size  $N$  as “the multispecies ASEP” or “the colored ASEP” on  $[N]$ . As the ordinary single-species ASEP, the multispecies ASEP can also be defined on  $\mathbb{Z}$  or any finite discrete interval (not just  $[N]$ ). More precisely, we define the multispecies ASEP as an evolving bijection from  $\mathbb{Z}$  or a finite discrete interval to itself, such that for any pair of nearest neighbors, if the left number is smaller than the right one, they swap with rate 1; otherwise, they swap with rate  $q$  and such swaps happen independently for all pairs of nearest neighbors.

The paper [9] also studied the mixing of the single-species ASEP on a finite interval, and proved a “pre-cutoff.” Namely, it was shown that the total-variation distance decays from 1 to 0 in a certain time scale, but not necessarily in a small time window. The cutoff problem (of the single-species ASEP) was settled in [27] (together with the biased card shuffling cutoff), using hydrodynamics of the ASEP on  $\mathbb{Z}$ . In [27], Labb   and Lacoin expected that the cutoff window should be in the order of  $N^{1/3}$  with the cutoff profile related to the so-called (parabolic) Airy<sub>2</sub> process. These were later established by Bufetov and Nejjar in [15]. They verified the  $N^{1/3}$  order of cutoff window and showed that the cutoff profile is given by the Tracy–Widom GUE distribution (which is the one point distribution of the Airy<sub>2</sub> process and was introduced by Tracy and Widom in [46] as the scaling limit of the fluctuation of the largest eigenvalue of the Gaussian Unitary Ensemble (GUE) matrices). As noted in [15], Section 1.2, while proving the cutoff for the biased card shuffling and the single-species ASEP follow similar and largely related arguments (in [27]), it is “significantly more delicate” to study the cutoff profile of the biased card shuffling than the single-species ASEP.

In [15], the authors also conjectured the  $N^{1/3}$  cutoff window and the Tracy–Widom GOE cutoff profile for the multispecies ASEP, that is, the biased card shuffling. This conjecture is verified by our Theorem 1.1. The distinction in the cutoff profiles, between the multi and single-species ASEP, is due to that they are respectively connected to ASEP on  $\mathbb{Z}$  with step (i.e., having a particle at each site  $\leq 0$  and being empty at each site  $> 0$ ) or flat (i.e., having a particle at every other site) initial configurations. As will be explained in Section 1.2, under this connection, the mixing times (for multi and single-species ASEP on  $[N]$ ) are described by the current fluctuations (of ASEP on  $\mathbb{Z}$  with step and flat initial configurations). By the KPZ fixed-point limit of ASEP [41], these current fluctuations have scaling limits related to the parabolic Airy<sub>2</sub> process. More precisely, their scaling limits are respectively given by the one point and global maximum of the process, which are known to be Tracy–Widom GOE and GUE, respectively. We note that for the step limit configuration, the Tracy–Widom GUE limit has been established earlier in [48].

The behavior of these chains (single-species/multispecies simple exclusion process in a finite interval) in the slightly different symmetric setting, that is, where  $q = 1$  rather than

TABLE 1

A summary of some results on the mixing/cutoff of single-species/multispecies exclusion processes

	Mixing time and cutoff	(Single-species) exclusion process cutoff profile	Multispecies exclusion process cutoff profile
Totally Asymmetric ( $q = 0$ )	[6]	[6] Tracy–Widom GUE	[14] Tracy–Widom GOE
Partially Asymmetric ( $q \in (0, 1)$ )	[9, 27]	[15] Tracy–Widom GUE	Our Theorem 1.1 Tracy–Widom GOE
Symmetric ( $q = 1$ )	[31, 49]	[30] (on the cycle) Gaussian	unknown

$q \in [0, 1)$  is quite different. For both single-species and multispecies chains, the mixing time is in the order of  $N^2 \log(N)$  [49], and cutoff has been proven in [31]. For the single-species ASEP on a circle, the cutoff window is proven to be in the order of  $N^2$ , with Gaussian cutoff profile [30]. See also, e.g., [23, 24, 43], and the survey [29] and the references therein, for some recent developments in this direction.

*The totally asymmetric case.* For the biased card shuffling, the special case where  $q = 0$  has been studied in the literature, first in [6], under the name “Oriented Swap Process” (OSP) and as a type of random sorting algorithm. Note that in this case, there is a unique absorbing state  $\lambda \in S_N$  where  $\lambda(x) = N + 1 - x$  for each  $x \in [N]$ , and mixing of this Markov chain degenerates into the “absorbing time,” that is, the time of reaching this absorbing state. In [6], the authors studied various aspects of the OSP, including the trajectories of the numbers, and various aspects of the absorbing time. In particular, they showed that the absorbing time divided by  $N$  converges to 2 in probability. They also studied the finishing time of each number, that is, the time a number  $k \in [N]$  makes its last move. This can be viewed as the finishing time (in a segment) of the (single-species) *totally asymmetric simple exclusion process* (TASEP), which is the degeneration of the (single-species) ASEP to the  $q = 0$  case. In [6], the authors showed that the one number finishing time has fluctuation in the order of  $N^{1/3}$  with Tracy–Widom GUE scaling limit, using TASEP convergence results from [26].

A question asked in [6] is to figure out the order and distribution of the absorbing time (which can be viewed as the maximum of the finishing times over all  $k \in [N]$ ). This was settled in [14], which proved that the absorbing time fluctuation is in the order of  $N^{1/3}$  with Tracy–Widom GOE scaling limit, using recently proved symmetries of the six-vertex model [13]. (This is also the reason why the biased card shuffling cutoff profile is conjectured to be Tracy–Widom GOE in [15].) Another direction pointed out in [6] is to generalize the results of OSP to the partially asymmetric (i.e.,  $q > 0$ ) case. Our Theorem 1.1 can be viewed as such a generalization of [14]. We would also like to mention a recent interesting development [2], which likely leads to generalizing the results in [6] to the  $q > 0$  case.

Finally, we mention that similar exclusion processes in the continuous space setting have also been studied. More precisely, one considers real numbers  $0 \leq x_1 \leq \dots \leq x_{N-1} \leq N$ . For each  $x_i$ , with rate 1 it is resampled from a certain distribution in  $[x_{i-1}, x_{i+1}]$ . Such models are also referred to as the “adjacent walk on the  $N$ -simplex.” In the symmetric case where the resampling law is taken to be uniform or a symmetric log-concave Beta distribution, cutoff at  $\pi^{-2} N^2 \log(N)$  is proved in [16]. The settings of asymmetric Beta distribution resampling laws are studied more recently in [28], with order  $N$  mixing time and a pre-cutoff established.

Some of the above results are summarized in Table 1.

**1.2. Our contribution and main tools.** We next describe our strategy and the used tools, in comparison with related previous works (in particular, [15] and [14]). As discussed above,

[15] studies the general  $q \in [0, 1)$  problem, but at the level of a single-species ASEP projection; on the other hand, [14] studies the multispecies model, but in the special  $q = 0$  case.

Our proofs use ingredients from both papers. As [14], we can use a shift-invariance property of the multispecies ASEP (from [13, 22]), to reduce problems on the multispecies ASEP on  $\mathbb{Z}$  to problems on the single-species ASEP on  $\mathbb{Z}$  with flat initial configuration. In [14], classical convergence results on the TASEP were then used to get the Tracy–Widom GOE limit, while here we instead use the more recently proved ASEP convergence to the KPZ fixed point [41]. (Note that Tracy–Widom GOE limit in ASEP has been formally derived earlier in [39].) Then the main task is to connect the cutoff profile problem in a finite interval, with the distribution of the multispecies ASEP on  $\mathbb{Z}$ .

In the  $q = 0$  case, due to that swaps happen only in one direction, there is a natural coupling between the multispecies TASEP on a finite interval and on  $\mathbb{Z}$ , using some truncation operators, as first described in [6]. Such a relation is not available when  $q > 0$ . Besides, the total-variation distance (in the  $q > 0$  case) is also much less tractable than the absorbing time (in the  $q = 0$  case).

Our main contribution is to develop new ideas to overcome these difficulties, and build the connection in the above “main task.” We use the height function representation of the single-species ASEP, and Hecke algebra identities; and the core arguments are given in Section 2.3 and Section 3 below. In more detail, we couple the biased card shuffling (with a deterministic initial configuration) with the stationary one so that the swaps are synchronized. This is the *basic coupling* to be defined shortly in Section 2.2. Then the mixing time is reduced to the stopping time when they are equal for the first time. By considering the single-species ASEP projections and using ordering properties of the ASEP height function, this problem is eventually reduced to comparing two (single-species) ASEPs on  $\mathbb{Z}$  with different initial configurations.

For two (single-species) ASEPs under the basic coupling, such that the set of particle locations of one of them contains the set of particle locations of the other one, they are equivalent to the ASEP with *second-class particles*. A second-class particle is a particle that can swap with a hole next to it on the right, or a particle next to it on the left, with rate 1, respectively. It can also swap with a hole next to it on the left, or a particle next to it on the right, with rate  $q$ , respectively. It is straightforward to check that, by replacing each second-class particle with a hole or replacing each second-class particle with a particle, one can get two (single-species) ASEPs that are under the basic coupling. Thus to study the difference between these two (single-species) ASEPs, one just needs to track the locations of the second-class particles. For this, we use Hecke algebra identities, reducing it to events under the Mallows measure, which is well understood with explicit distribution functions (e.g., (2.1) and (2.3) below).

Some of these ingredients, such as the basic coupling and Hecke algebra, have also appeared in [15]. However, there are several key differences between the arguments in this paper and [15]. The most significant aspect is in our usage of Hecke algebra. In [15], Hecke algebra was used to translate events on certain single-species ASEPs to events on certain ASEPs with second-class particles, which were analyzed directly. In this paper, Hecke algebra is used in the opposite direction. Namely, we introduce second-class particles because we wish to compare two (single-species) ASEPs; and the Hecke algebra is used to reduce events on the ASEP with second-class particles to events on another (single-species) ASEP without second-class particles. Besides, while ordering of particle configurations was also used in [15], it is more extensively exploited in this paper, with the help of the more intuitive height function. Using these, we can directly compare ASEPs on a finite interval and  $\mathbb{Z}$ , and such comparison was not obtained in [15]. Finally, we mention that our proof can be easily adapted to the single-species ASEP setting, to get an alternative proof of the main result of [15] (see Remark 2.6 below).

*Notation and terminologies.* Some basic notation are used throughout this paper (and some of them have appeared already). For any  $x, y \in \mathbb{R} \cup \{-\infty, \infty\}$ , we let  $x \wedge y = \min\{x, y\}$  and  $x \vee y = \max\{x, y\}$ , and let  $\llbracket x, y \rrbracket$  be the discrete interval, containing all integers  $\geq x$  and  $\leq y$ . We also write  $[N] = \llbracket 1, N \rrbracket$  for any  $N \in \mathbb{N}$ . Below  $q$  always denotes the rate of reversed jumps or swaps, and we always assume that  $q \in [0, 1)$ .

*Organization of the remaining text.* The remaining text mainly focuses on the proofs. In Sections 2.1 and 2.2, we set up some notation and provide some preliminary lemmas; then in Section 2.3 we state the main steps of our arguments, as three propositions, and prove Theorem 1.1 assuming them. The most important one (of the three steps), where we compare the biased card shuffling and the multispecies ASEP on  $\mathbb{Z}$ , is implemented in Section 3. In the last two sections, we do the remaining steps, respectively. Specifically, in Section 4 we prove an estimate for certain “ground state hitting time” of the single-species ASEP, and in Section 5 we deduce convergence to the Tracy–Widom GOE distribution using known ASEP shift-invariance and convergence results.

## 2. Preliminaries and main steps.

**2.1. The stationary Mallows measure.** We start by giving the formal definition of the Mallows measures, which are stationary measures of the biased card shuffling or the multi-species ASEP on a finite interval. For the convenience of later arguments, we consider the interval  $\llbracket m, n \rrbracket$  for integers  $m \leq n$ . As in [15], we let  $S_{m,n}$  denote the set of all bijections between  $\llbracket m, n \rrbracket$  and itself, and let  $\mathcal{M}_{m,n}$  denote the Mallows measure on  $S_{m,n}$ , such that for any  $w \in S_{m,n}$ ,

$$(2.1) \quad \mathcal{M}_{m,n}(w) = q^{\kappa(w)} Z_{m,n},$$

where (and also for the rest of this paper) we denote  $\kappa(w) = \sum_{m \leq i < j \leq n} \mathbb{1}[w(i) < w(j)]$  as the “energy” for any  $w \in S_{m,n}$ , and

$$(2.2) \quad Z_{m,n} = \frac{1}{\sum_{w \in S_{m,n}} q^{\kappa(w)}} = \prod_{i=1}^{n-m+1} \frac{1-q}{1-q^i}.$$

Note that there is a unique  $w \in S_{m,n}$  with  $\kappa(w) = 0$ , that is,  $w : i \mapsto m + n - i$ ; and that is the “ground state.” It is straightforward to check that  $\mathcal{M}_{m,n}$  is the stationary measure of the biased card shuffling on  $\llbracket m, n \rrbracket$ .

For each  $k \in \llbracket 0, n - m + 1 \rrbracket$ , we let  $\mathcal{P}_{m,n}^k$  be the measure on  $\{0, 1\}^{\llbracket m, n \rrbracket}$ , obtained from  $\mathcal{M}_{m,n}$  under the projection where each  $w \in S_{m,n}$  is mapped to  $x \mapsto \mathbb{1}[w(x) \leq m + k - 1]$ . In words, under  $\mathcal{P}_{m,n}^k$  there are  $k$  particles and  $n - m + 1 - k$  holes if we view each 1 as a particle and each 0 as a hole. One can readily check that, for any  $\omega \in \{0, 1\}^{\llbracket m, n \rrbracket}$ ,

$$(2.3) \quad \mathcal{P}_{m,n}^k(\omega) = \mathbb{1} \left[ \sum_{i \in \llbracket m, n \rrbracket} \omega(i) = k \right] q^{\kappa(\omega)} Z_{m,n}^k,$$

where (as a slight misuse of notation, and also for the rest of this paper)  $\kappa(\omega) = \sum_{m \leq i < j \leq n} \mathbb{1}[\omega(i) = 1] \mathbb{1}[\omega(j) = 0]$  is the “energy” for any  $\omega \in \{0, 1\}^{\llbracket m, n \rrbracket}$ , and

$$Z_{m,n}^k = \frac{1}{\sum_{w \in \{0, 1\}^{\llbracket m, n \rrbracket}, \sum_{i \in \llbracket m, n \rrbracket} \omega(i) = k} q^{\kappa(w)}} = \frac{\prod_{i=1}^k (1-q^i) \prod_{i=1}^{n-m+1-k} (1-q^i)}{\prod_{i=1}^{n-m+1} (1-q^i)}.$$

Note that here the “ground state” is given by  $\omega : i \mapsto \mathbb{1}[i \geq n - k + 1]$ . From this definition of  $\mathcal{P}_{m,n}^k$  as a projection of  $\mathcal{M}_{m,n}$ , it is stationary under the (single-species) ASEP evolution on  $\llbracket m, n \rrbracket$ .

We will frequently use the following estimate of the “energy” under  $\mathcal{P}_{m,n}^k$ .

LEMMA 2.1. *There is a constant  $C > 0$  depending only on  $q$ , such that for any integers  $m \leq n$ ,  $k \in \llbracket 0, n - m + 1 \rrbracket$ , and any  $a \in \mathbb{N}$ ,*

$$(2.4) \quad \mathcal{P}_{m,n}^k(\{\omega \in \{0, 1\}^{\llbracket m, n \rrbracket} : \kappa(\omega) > a\}) < Cq^{a/2}.$$

PROOF. For each  $\omega \in \{0, 1\}^{\llbracket m, n \rrbracket}$  and  $i \in \llbracket m, n \rrbracket$ , denote  $r_i = \mathbb{1}[\omega(i) = 1] \times \sum_{i < j \leq n} \mathbb{1}[\omega(j) = 0]$ . Then  $\sum_{i=m}^n r_i = \kappa(\omega)$ , and  $r_i \geq r_j$  for any  $i < j$  with  $\omega(i) = \omega(j) = 1$ . Therefore, for each  $i \in \mathbb{Z}_{\geq 0}$ , we have

$$|\{\omega \in \{0, 1\}^{\llbracket m, n \rrbracket}, \kappa(\omega) = i\}| \leq \left| \left\{ \{a_j\}_{j=1}^{\infty} \in \mathbb{Z}_{\geq 0}^{\mathbb{N}} : a_j \geq a_{j+1}, \forall j \in \mathbb{N}, \sum_{j=1}^{\infty} a_j = i \right\} \right|.$$

We note that the right-hand side equals  $\alpha_i$ , where  $\sum_{i=0}^{\infty} \alpha_i z^i$  is the Taylor series for  $\prod_{i=1}^{\infty} (1 - z^i)^{-1}$ . By (2.3), and noting that  $Z_{m,n}^k \leq 1$ , the left-hand side of (2.4) is at most

$$\sum_{i=a+1}^{\infty} \alpha_i q^i < q^{(a+1)/2} \sum_{i=a+1}^{\infty} \alpha_i q^{i/2} < q^{(a+1)/2} \sum_{i=0}^{\infty} \alpha_i q^{i/2} = q^{(a+1)/2} \prod_{i=1}^{\infty} (1 - q^{i/2})^{-1}.$$

By taking  $C = q^{1/2} \prod_{i=1}^{\infty} (1 - q^{i/2})^{-1}$ , the conclusion follows.  $\square$

For  $N \in \mathbb{N}$  and  $k \in \llbracket 0, N \rrbracket$ , below we also write  $S_N = S_{1,N}$ ,  $\mathcal{M}_N = \mathcal{M}_{1,N}$  and  $\mathcal{P}_N^k = \mathcal{P}_{1,N}^k$ , for simplicity of notation.

2.2. *Setup for several processes.* In this subsection, we define some processes to be used in the proof of Theorem 1.1, and discuss some basic properties of them.

- Let  $\xi = (\xi_t)_{t \geq 0}$  be the multispecies ASEP on  $\mathbb{Z}$ , such that the configuration at any time  $t \geq 0$  is denoted as a bijection  $\xi_t : \mathbb{Z} \rightarrow \mathbb{Z}$ . Let the initial configuration  $\xi_0$  be the identity map of  $\mathbb{Z}$ .

In defining the following processes and for the rest of this paper, we take  $N \in \mathbb{N}$ :

- We let  $\xi = (\xi_t)_{t \geq 0}$  be the biased card shuffling of size  $N$ , with  $\xi_t : [N] \rightarrow [N]$  being the state at time  $t$ . Let the initial configuration  $\xi_0$  be the identity map of  $[N]$ .
- We let  $\lambda = (\lambda_t)_{t \geq 0}$  be the same as  $\xi = (\xi_t)_{t \geq 0}$ , except for that the initial configuration  $\lambda_0$  is some general deterministic element of  $S_N$ .
- Let  $\bar{\xi} = (\bar{\xi}_t)_{t \geq 0}$  be the stationary biased card shuffling of size  $N$ . Then for any  $t \geq 0$  the law of  $\bar{\xi}_t$  is  $\mathcal{M}_N$ .

*Basic coupling.* For the multispecies or single-species ASEP on  $\mathbb{Z}$  or a discrete interval, the evolution can be generated by two independent Poisson point processes on  $\mathbb{Z} \times \mathbb{R}_{\geq 0}$ , with rates 1 and  $q$ , respectively. In the setting of the multispecies ASEP on  $\mathbb{Z}$ , if there is a point at  $(x, t)$  in the rate 1 Poisson point process, at time  $t$  we swap the numbers at  $x$  and  $x + 1$ , if before the swap the number at  $x$  is smaller than the number at  $x + 1$ ; and if there is a point at  $(x, t)$  in the rate  $q$  Poisson point process, at time  $t$  we swap the numbers at  $x$  and  $x + 1$ , if before the swap the number at  $x$  is larger than the number at  $x + 1$ . In the discrete interval setting, one only considers such points  $(x, t)$  with both  $x$  and  $x + 1$  in the discrete interval. One can generate the single-species ASEP (from these Poisson point processes) similarly. The *basic coupling* between two or more multispecies or single-species ASEPs is the coupling under which all these processes are generated from the same pair of Poisson point processes.

For the defined processes  $\xi, \bar{\xi}, \lambda$  and  $\bar{\xi}$ , we couple all of them with the basic coupling.

*Projections.* For any  $k \in \llbracket 0, N \rrbracket$ , we define the processes  $\xi^k = (\xi_t^k)_{t \geq 0}$ ,  $\zeta^k = (\zeta_t^k)_{t \geq 0}$  and  $\bar{\xi}^k = (\bar{\xi}_t^k)_{t \geq 0}$ , as follows. For each  $t \geq 0$ , we let  $\xi_t^k, \lambda_t^k, \bar{\xi}_t^k : [N] \rightarrow \{0, 1\}$  be the projections, where

$$\xi_t^k(x) = \mathbb{1}[\xi_t(x) \leq k], \quad \lambda_t^k(x) = \mathbb{1}[\lambda_t(x) \leq k], \quad \bar{\xi}_t^k(x) = \mathbb{1}[\bar{\xi}_t(x) \leq k].$$

And for each  $k \in \mathbb{Z}$  and  $t \geq 0$ , we let  $\zeta_t^k : \mathbb{Z} \rightarrow \{0, 1\}$  be the projection where

$$\zeta_t^k(x) = \mathbb{1}[\zeta_t(x) \leq k].$$

Then these projections are single-species ASEP on  $\mathbb{Z}$  or  $[N]$ , and are also coupled together under the basic coupling (i.e., generated with the same Poisson point processes). We note that, for each  $t \geq 0$ , given  $\zeta_t^k$  for all  $k \in \mathbb{Z}$ , one can recover  $\zeta_t$  by letting

$$\zeta_t(x) = \min\{k : \zeta_t^k(x) = 1\}.$$

Similar recovery statements hold for  $\xi_t$ ,  $\lambda_t$ , and  $\bar{\xi}_t$ .

*Height function.* For any  $\omega : \mathbb{Z} \rightarrow \{0, 1\}$  such that  $\omega(x)$  stabilizes as  $x \rightarrow -\infty$ , we define the height function  $h\{\omega\} : \mathbb{Z} \rightarrow \mathbb{Z}$ , as follows:

1. For any  $x \in \mathbb{Z}$ , let  $h\{\omega\}(x) - h\{\omega\}(x-1) = 1 - 2\omega(x)$ .
2. If  $\omega(x) = 0$  for all  $x$  small enough, we let  $h\{\omega\}(x) = x$  for all  $x$  small enough.
3. If  $\omega(x) = 1$  for all  $x$  small enough, we let  $h\{\omega\}(x) = -x$  for all  $x$  small enough.

A quick and useful observation is that  $h\{\omega\}$  is 1-Lipschitz. For any integers  $m < n$  and  $\omega : \llbracket m, n \rrbracket \rightarrow \{0, 1\}$ , we also define  $h\{\omega\}$  by assuming  $\omega(x) = 0$  for any  $x \in \mathbb{Z}$ ,  $x < m$  and  $\omega(x) = 1$  for any  $x \in \llbracket n, \infty \rrbracket$  (see Figure 2). We note that for any  $k \in \llbracket 0, n-m+1 \rrbracket$ , among all the configurations in  $\{0, 1\}^{\llbracket m, n \rrbracket}$  that have positive probabilities under  $\mathcal{P}_{m,n}^k$ , the “ground state” (i.e.,  $\omega : i \mapsto \mathbb{1}[i \geq n-k+1]$ ) has the highest height function (i.e.,  $x \mapsto n-k-|x-n+k|$ ).

Such height functions are widely used in the study of interacting particle systems. For  $\omega$  evolving as the single-species ASEP, the height function  $h\{\omega\}$  evolves in the following way: for some  $x \in \mathbb{Z}$ , if  $h\{\omega\}(x-1) = h\{\omega\}(x+1) = h\{\omega\}(x) + 1$  for some  $x \in \mathbb{Z}$ , then  $h\{\omega\}(x)$  changes to  $h\{\omega\}(x) + 2$  with rate 1; if  $h\{\omega\}(x-1) = h\{\omega\}(x+1) = h\{\omega\}(x) - 1$ , then  $h\{\omega\}(x)$  changes to  $h\{\omega\}(x) - 2$  with rate  $q$ .

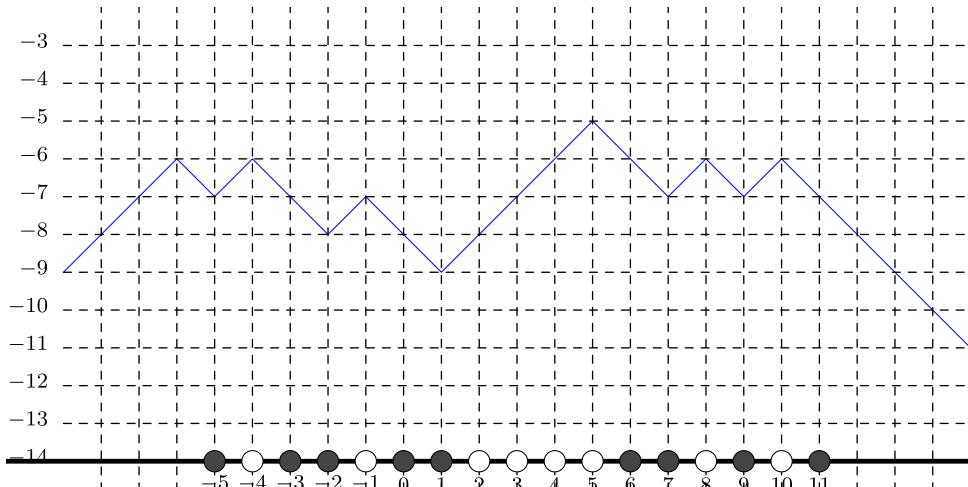


FIG. 2. An illustration of the height function for a single-species ASEP configuration in a finite interval.

We can define a partial ordering for these height functions. For any  $f, g : \mathbb{Z} \rightarrow \mathbb{R}$ , we write  $f \leq g$  if  $f(x) \leq g(x)$  for any  $x \in \mathbb{Z}$ . One reason for using these height functions is that, as one can check, this ordering is preserved under the ASEP evolution with the basic coupling. For example, for the processes defined above, the following orderings hold.

LEMMA 2.2. *We have  $h\{\xi_t^k\} \leq h\{\lambda_t^k\}$ ,  $h\{\xi_t^k\} \leq h\{\zeta_t^k\}$  and  $h\{\xi_t^k\} \leq h\{\bar{\xi}_t^k\}$ , for any  $k \in \llbracket 0, N \rrbracket$  and  $t \geq 0$ .*

PROOF. We first consider the case where  $t = 0$ . We have

$$h\{\xi_0^k\}(0) = h\{\lambda_0^k\}(0) = h\{\zeta_0^k\}(0) = h\{\bar{\xi}_0^k\}(0) = 0,$$

and

$$h\{\xi_0^k\}(N) = h\{\lambda_0^k\}(N) = h\{\zeta_0^k\}(N) = h\{\bar{\xi}_0^k\}(N) = N - 2k,$$

since

$$\sum_{x \in [N]} \xi_0^k(x) = \sum_{x \in [N]} \lambda_0^k(x) = \sum_{x \in [N]} \zeta_0^k(x) = \sum_{x \in [N]} \bar{\xi}_0^k(x) = k.$$

Then since  $h\{\xi_0^k\}$  has slope 1 in  $\llbracket -\infty, 0 \rrbracket \cup \llbracket k, N \rrbracket$  and slope  $-1$  in  $\llbracket 0, k \rrbracket \cup \llbracket N, \infty \rrbracket$ , and that  $h\{\lambda_0^k\}$ ,  $h\{\zeta_0^k\}$  and  $h\{\bar{\xi}_0^k\}$  are 1-Lipschitz, we have that these inequalities hold when  $t = 0$ .

For any  $t > 0$ , first note that the inequalities always hold in  $\llbracket -\infty, 0 \rrbracket$  and  $\llbracket N, \infty \rrbracket$ . Indeed, we have  $h\{\xi_t^k\}(x) = h\{\lambda_t^k\}(x) = h\{\bar{\xi}_0^k\}(x) = x$  for  $x \in \llbracket -\infty, 0 \rrbracket$ , and  $h\{\xi_t^k\}(x) = h\{\lambda_t^k\}(x) = h\{\bar{\xi}_0^k\}(x) = 2N - 2k - x$  for  $x \in \llbracket N, \infty \rrbracket$ ; and  $h\{\xi_t^k\}(x) \geq (-x) \vee (x - 2k)$  for all  $x \in \mathbb{Z}$  (since  $h\{\zeta_t^k\}(x) = -x$  for all small enough  $x$  and  $h\{\zeta_t^k\}(x) = x - 2k$  for all large enough  $x$ , and  $h\{\zeta_t^k\}$  is 1-Lipschitz). In the interval  $\llbracket 1, N - 1 \rrbracket$ , from the above description of how these height functions evolve in time, it is straightforward to check that, if the inequalities hold at any time, any nearest neighbor swap in the basic coupling would not break them. Then the inequalities hold for any  $t > 0$ .  $\square$

2.3. *Proof of the main result.* With the processes defined in the previous subsection, to understand the cutoff profile (thereby prove Theorem 1.1), we will (1) upper bound the first time  $T$  when  $\lambda_T = \bar{\xi}_T$ ; (2) lower bound the first time  $T$  when  $\bar{\xi}_T = \xi_T$ , and show that  $\bar{\xi}_t$  and  $\xi_t$  are “quite different” for  $t$  between  $T - o(N^{1/3})$  and  $T$ . We now give the main steps to accomplish these tasks.

We start by explaining how the Tracy–Widom GOE distribution is involved. This distribution appears in scaling limits of the single-species ASEP on  $\mathbb{Z}$ , as we will later see using the single-species ASEP convergence results from [41]. From these results and using an ASEP shift-invariance properties proved in [13, 22], we can deduce convergence results for the multispecies ASEP on  $\mathbb{Z}$ , as follows.

For  $b \in \mathbb{R}$  and  $t \geq 0$ , let  $\mathcal{D}_{N,t}^b$  denote the event where  $h\{\zeta_t^k\}(N - k) > N - k + b$  for any  $k \in \llbracket 0, N \rrbracket$ .

PROPOSITION 2.3. *For any  $\tau \in \mathbb{R}$ , and any real number sequence  $\{b_N\}_{N \in \mathbb{N}}$  such that  $\lim_{N \rightarrow \infty} N^{-1/3} b_N \rightarrow 0$ , we have*

$$\lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{D}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{b_N}] = F_{\text{GOE}}(2^{2/3}\tau).$$

The proof of this proposition will be given in Section 5.

In the TASEP setting (i.e., when  $q = 0$ ), the multispecies process on a finite interval can be obtained from the multispecies process on  $\mathbb{Z}$ , using truncation operators (see, e.g., [6]). Thus Proposition 2.3 directly implies that the absorbing time has Tracy–Widom GOE distribution asymptotically; and that is the proof in [14].

For the general  $q > 0$  setting, we need to study the more delicate mixing time (rather than absorbing time). A more important difference is that there is no such truncation operators and no exact coupling structure between the process on  $\mathbb{Z}$  and the process on a finite interval. Thus we instead show that, under the basic coupling, the processes on  $\mathbb{Z}$  and a finite interval are “close to each other” in terms of the height functions of their projections, in the following sense.

For  $a \in \mathbb{N}$ ,  $b \in \mathbb{R}$ ,  $k \in \llbracket 0, N \rrbracket$  and  $t \geq 0$ , let  $\mathcal{A}_{N,k,t}^{a,b}$  denote the event where  $h\{\xi_t^k\}(N-k) > N-k+b$ , and

$$\text{either } h\{\xi_t^k\}(N-k-a) < N-k-a \text{ or } h\{\xi_t^k\}(N-k+a) < N-k-a.$$

**PROPOSITION 2.4.** *There is a constant  $C > 0$  depending only on  $q$ , such that  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b}] < Cq^{a/2}$  when  $b > 4a$ .*

We will prove this proposition in Section 3, using Hecke algebra techniques.

In analyzing the first time  $T$  when  $\xi_T = \bar{\xi}_T$ , a key ingredient we need is that, if  $\xi_t$  and  $\bar{\xi}_t$  are “close to each other” at some time  $t$ , they will be the same in a short amount of time. We state this result as follows.

For  $a \in \mathbb{N}$ ,  $r > 0$ ,  $k \in \llbracket 0, N \rrbracket$  and  $t \geq 0$ , let  $\mathcal{B}_{N,k,t}^{a,r}$  denote the event where

$$h\{\xi_{t+i}^k\}(N-k-a) = h\{\xi_{t+i}^k\}(N-k+a) = N-k-a \quad \forall i \in \llbracket 0, r \rrbracket,$$

while  $h\{\xi_{t+s}^k\}(N-k) < N-k$  for each  $s \in [0, r]$  (in other words,  $\xi_{t+s}^k$  is not at the “ground state” for each  $s \in [0, r]$ , since  $h\{\xi_{t+s}^k\}(N-k) = N-k$  is equivalent to that  $\xi_{t+s}^k(x) = \mathbb{1}[x \geq N-k+1]$ ).

**PROPOSITION 2.5.** *There is a constant  $c > 0$  depending only on  $q$ , such that*

$$\mathbb{P}[\mathcal{B}_{N,k,t}^{a,\log(N)^4}] < e^{-c\log(N)^2},$$

when  $2 \leq a \leq \log(N)^2$ .

This proposition will be proved in Section 4, using an ASEP hitting time bound from [9].

We now give the proof of our main result (Theorem 1.1), which mainly consists of assembling these propositions.

**PROOF OF THEOREM 1.1: UPPER BOUND.** The key thing is to show that, for whatever initial configuration  $\lambda_0$ , we have

$$(2.5) \quad \liminf_{N \rightarrow \infty} \mathbb{P}[\lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4} = \bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}] \geq F_{\text{GOE}}(2^{2/3}\tau).$$

Let us first explain the strategy to prove it. In light of Proposition 2.3, the main idea is to show that the event in the left-hand side of (2.5) is roughly implied by  $\mathcal{D}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{8\log(N)^4}$ . We will show that, under  $\mathcal{D}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{8\log(N)^4}$ , for each  $k \in \llbracket 0, N \rrbracket$ ,  $\xi_{2(1-q)^{-1}(N+\tau N^{1/3})}^k$  is “close” to the ground state, that is, most locations in  $\llbracket N-k+1, N \rrbracket$  have particles, and most locations in  $\llbracket 1, N-k \rrbracket$  have holes. Then there exists some small  $s > 0$ , such that

$\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+s}^k$  is precisely at the ground state. These two steps are given by Propositions 2.4 and 2.5. Then  $h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+s}^k\}$  reaches the “largest possible state,” that is, it has slope 1 in  $[-\infty, N-k]$  and slope  $-1$  in  $[N-k, \infty]$ . By ordering of the height functions (Lemma 2.2), this implies the event in the left-hand side of (2.5).

We now give the details of proving (2.5). Consider the event

$$\mathcal{E}_0 = \mathcal{D}_{N, 2(1-q)^{-1}(N+\tau N^{1/3})}^{8\log(N)^4} \setminus \mathcal{E}_1 \cup \left( \bigcup_{\substack{k \in \llbracket 0, N \rrbracket, \\ i \in \llbracket 0, \log(N)^4 \rrbracket}} \mathcal{A}_{N, k, 2(1-q)^{-1}(N+\tau N^{1/3})+i}^{\lfloor \log(N)^2 \rfloor, 4\log(N)^4} \right) \cup \left( \bigcup_{k \in \llbracket 0, N \rrbracket} \mathcal{B}_{N, k, 2(1-q)^{-1}(N+\tau N^{1/3})}^{\lfloor \log(N)^2 \rfloor, \log(N)^4} \right).$$

Here,  $\mathcal{E}_1$  is the following event:

- For some  $k \in \llbracket 0, N \rrbracket$  and  $i \in \llbracket 0, \log(N)^4 \rrbracket$ , we have

$$(2.6) \quad h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+i}^k\}(N-k) > N-k + 8\log(N)^4,$$

and

$$(2.7) \quad h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+i}^k\}(N-k) \leq N-k + 4\log(N)^4.$$

Now let us assume that  $\mathcal{E}_0$  happens. By  $\mathcal{D}_{N, 2(1-q)^{-1}(N+\tau N^{1/3})}^{8\log(N)^4} \setminus \mathcal{E}_1$ , we have that for any  $k \in \llbracket 0, N \rrbracket$  and  $i \in \llbracket 0, \log(N)^4 \rrbracket$ , (2.7) does not hold. Then since  $\mathcal{A}_{N, k, 2(1-q)^{-1}(N+\tau N^{1/3})+i}^{\lfloor \log(N)^2 \rfloor, 4\log(N)^4}$  does not hold, we must have that

$$\begin{aligned} & h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+i}^k\}(N-k - \lfloor \log(N)^2 \rfloor) \\ &= h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+i}^k\}(N-k + \lfloor \log(N)^2 \rfloor) = N-k - \lfloor \log(N)^2 \rfloor. \end{aligned}$$

Using this for all  $k \in \llbracket 0, N \rrbracket$  and  $i \in \llbracket 0, \log(N)^4 \rrbracket$ , and that  $\mathcal{B}_{N, k, 2(1-q)^{-1}(N+\tau N^{1/3})}^{\lfloor \log(N)^2 \rfloor, \log(N)^4}$  does not hold, we have that  $\mathcal{E}_2$  holds, with  $\mathcal{E}_2$  being the event:

- For any  $k \in \llbracket 0, N \rrbracket$ , we have  $h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k\}(N-k) = N-k$  for some  $s_k \in [0, \log(N)^4]$ .

In summary, we have shown that  $\mathcal{E}_0$  implies  $\mathcal{E}_2$ .

We next show that under  $\mathcal{E}_2$ , we must have

$$(2.8) \quad \bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4} = \lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}.$$

Indeed, assuming  $\mathcal{E}_2$ , for any  $k \in \llbracket 0, N \rrbracket$  we have that  $h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k\}$  is the function with slope 1 in  $[-\infty, N-k]$ , and slope  $-1$  in  $[N-k, \infty]$ . Thus by Lemma 2.2, the same is true for  $h\{\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k\}$  and  $h\{\lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k\}$ . Then we have

$$\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k = \bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k = \lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+s_k}^k.$$

Since  $\bar{\xi}$  and  $\lambda$  are under the basic coupling and  $s_k \leq \log(N)^4$ , this implies that

$$\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}^k = \lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}^k.$$

As this holds for each  $k \in \llbracket 0, N \rrbracket$ , we have (2.8).

Now we have

$$\begin{aligned}
& \mathbb{P}[\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4} = \lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}] \\
& \geq \mathbb{P}[\mathcal{E}_2] \geq \mathbb{P}[\mathcal{D}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{8\log(N)^4}] - \mathbb{P}[\mathcal{E}_1] \\
(2.9) \quad & - \sum_{k \in \llbracket 0, N \rrbracket, i \in \llbracket 0, \log(N)^4 \rrbracket} \mathbb{P}[\mathcal{A}_{N,k,2(1-q)^{-1}(N+\tau N^{1/3})+i}^{\lfloor \log(N)^2 \rfloor, 4\log(N)^4}] \\
& - \sum_{k \in \llbracket 0, N \rrbracket} \mathbb{P}[\mathcal{B}_{N,k,2(1-q)^{-1}(N+\tau N^{1/3})}^{\lfloor \log(N)^2 \rfloor, \log(N)^4}].
\end{aligned}$$

To get (2.5), we next bound  $\mathbb{P}[\mathcal{E}_1]$ .

Fix  $k \in \llbracket 0, N \rrbracket$ . Let  $T_0 = 0$ . For each  $j \in \mathbb{N}$ , we let  $T_j$  be the smallest positive number with

$$h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+T_j}^k\}(N-k) = h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k) - 2j;$$

or  $T_j = \infty$  if no such positive number exists. Then given  $T_{j-1}$  and  $\xi_{2(1-q)^{-1}(N+\tau N^{1/3})+T_{j-1}}^k$ ,  $T_j - T_{j-1}$  stochastically dominates an  $\text{Exp}(q)$  random variable. Thus  $T_{\lfloor 2\log(N)^4 \rfloor}$  stochastically dominates the sum of  $\lfloor 2\log(N)^4 \rfloor$  independent  $\text{Exp}(q)$  random variables. On the other hand, if (2.6) and (2.7) happen for some  $i \in \llbracket 0, \log(N)^4 \rrbracket$ , we must have  $T_{\lfloor 2\log(N)^4 \rfloor} < \log(N)^4$ . Using a Chernoff bound, this happens with probability  $< Ce^{-c\log(N)^4}$ , for  $c, C > 0$  being universal constants. By a union bound over  $k$ , we have that  $\mathbb{P}[\mathcal{E}_1] < C(N+1)e^{-c\log(N)^4}$ .

Using this bound of  $\mathbb{P}[\mathcal{E}_1]$ , and (2.9), and Propositions 2.3, 2.4, 2.5 and sending  $N \rightarrow \infty$ , we have that (2.5) holds.

Finally, as the distributions of  $\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}$  and  $\lambda_{2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}$  are given by  $\mathcal{M}_N$  and  $\mathcal{W}_{N,2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}^{\lambda_0}$ , respectively, by (2.5) we have that

$$\limsup_{N \rightarrow \infty} \max_{\lambda \in S_N} \|\mathcal{W}_{N,2(1-q)^{-1}(N+\tau N^{1/3})+\log(N)^4}^{\lambda} - \mathcal{M}_N\|_{\text{TV}} \leq 1 - F_{\text{GOE}}(2^{2/3}\tau).$$

Since  $t \mapsto \|\mathcal{W}_{N,t}^{\lambda} - \mathcal{M}_N\|_{\text{TV}}$  is nonincreasing, and that  $F_{\text{GOE}}$  is a continuous function (which is evident from (1.1)), the upper bound of Theorem 1.1 follows.  $\square$

The lower bound proof is similar and simpler since for this we just need to show that  $\xi_{2(1-q)^{-1}(N+\tau N^{1/3})}^k$  and  $\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k$  are different with probability at least (roughly)  $1 - F_{\text{GOE}}(2^{2/3}\tau)$ . Only the ASEP on  $\mathbb{Z}$  convergence (Proposition 2.3) and ordering of the height functions (Lemma 2.2) will be used.

PROOF OF THEOREM 1.1: LOWER BOUND. We consider the events  $\mathcal{E}$ :

$$h\{\xi_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k) \leq N - k - \log(N)^2 \quad \exists k \in \llbracket 0, N \rrbracket.$$

and  $\bar{\mathcal{E}}$ :

$$h\{\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k) \leq N - k - \log(N)^2 \quad \exists k \in \llbracket 0, N \rrbracket.$$

Since the distribution of  $\xi_{2(1-q)^{-1}(N+\tau N^{1/3})}$  is  $\mathcal{W}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{\xi_0}$  and the distribution of  $\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}$  is  $\mathcal{M}_N$ , we have

$$\|\mathcal{W}_{N,2(1-q)^{-1}(N+\tau N^{1/3})}^{\xi_0} - \mathcal{M}_N\|_{\text{TV}} \geq \mathbb{P}[\mathcal{E}] - \mathbb{P}[\bar{\mathcal{E}}].$$

By Proposition 2.3 and Lemma 2.2, we have  $\liminf_{N \rightarrow \infty} \mathbb{P}[\mathcal{E}] \geq 1 - F_{\text{GOE}}(2^{2/3}\tau)$ . It now suffices to show that  $\mathbb{P}[\bar{\mathcal{E}}] \rightarrow 0$  as  $N \rightarrow \infty$ .

For each  $k \in \llbracket 0, N \rrbracket$ , the event  $h\{\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k) \leq N-k-\log(N)^2$  implies that

$$h\{\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k-\lfloor \log(N)^2/2 \rfloor + 1) \leq N-k-\log(N)^2 + \lfloor \log(N)^2/2 \rfloor - 1.$$

Then there exists an integer  $1 \leq x \leq N-k-\lfloor \log(N)^2/2 \rfloor + 1$  with  $\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k(x) = 1$ . This implies that

$$\kappa(\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k) \geq \lfloor \log(N)^2/2 \rfloor,$$

where we recall (from Section 2.1) that  $\kappa(\bar{\xi}_{2(1-q)^{-1}(N+\tau N^{1/3})}^k)$  is the ‘‘energy.’’ By Lemma 2.1, the probability of this event is bounded by  $Cq^{\log(N)^2/4}$ , for some  $C > 0$  depending only on  $q$ . Then by taking a union bound over  $k$  we have that  $\lim_{N \rightarrow \infty} \mathbb{P}[\bar{\mathcal{E}}] = 0$ , and we complete this proof.  $\square$

**REMARK 2.6.** Using Propositions 2.4 and 2.5 for a single  $k$ , and known one-point distribution convergence of ASEP to the Tracy–Widom GUE distribution (e.g., [48], Theorem 3) instead of Proposition 2.3, one can recover the main result of [15], via the same arguments as the proof of Theorem 1.1. Our Proposition 2.5 plays a similar role as the arguments in Section 3 of [15], and both use a hitting time bound from [9]. Our Proposition 2.4, on the other hand, plays a similar role as Section 4 of [15], but its proof (to be given in the next section) is quite different. As discussed above, while both our Proposition 2.4 and [15] use Hecke algebra as the main tool, it is used in quite different ways with different points of view, enabling us to analyze this more delicate multispecies setting.

The remaining three sections are devoted to proving the three propositions.

**3. ASEP on the line and an interval: Hecke algebra.** In this section, we prove Proposition 2.4. We can write  $\mathcal{A}_{N,k,t}^{a,b} = \mathcal{A}_{N,k,t}^{a,b,+} \cup \mathcal{A}_{N,k,t}^{a,b,-}$ , where

$$\begin{aligned} \mathcal{A}_{N,k,t}^{a,b,-} &: h\{\zeta_t^k\}(N-k) > N-k+b \text{ and } h\{\zeta_t^k\}(N-k-a) < N-k-a, \\ \mathcal{A}_{N,k,t}^{a,b,+} &: h\{\zeta_t^k\}(N-k) > N-k+b \text{ and } h\{\zeta_t^k\}(N-k+a) < N-k-a. \end{aligned}$$

By symmetry, below we just bound the probability  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b,+}]$ .

**3.1. Hecke algebra and basic properties.** We start by formally introducing Hecke algebra, the main tool of this section. Several notation below are from [15].

For any integers  $m < n$ , we can think of  $S_{m,n}$  as the permutation group of  $\llbracket m, n \rrbracket$ , such that for any  $w, v \in S_{m,n}$  we let  $wv = v \circ w$ . The Hecke algebra  $\mathcal{H}_{m,n}$  is the algebra with basis  $\{T_w\}_{w \in S_{m,n}}$ , and the rules

$$(3.1) \quad T_s T_w = \begin{cases} T_{sw} & \text{if } \kappa(sw) = \kappa(w) - 1, \\ (1-q)T_w + qT_{sw} & \text{if } \kappa(sw) = \kappa(w) + 1, \end{cases}$$

for any  $w \in S_{m,n}$  and  $s$  being any nearest neighbor transposition in  $S_{m,n}$  (i.e., there is some  $i \in \llbracket m, n-1 \rrbracket$  such that  $s(i) = i+1$ ,  $s(i+1) = i$  and  $s(j) = j$  for any  $j \in \llbracket m, n \rrbracket \setminus \{i, i+1\}$ ). Here, we recall (from Section 2.1) that  $\kappa(w) = \sum_{m \leq i < j \leq n} \mathbf{1}[w(i) < w(j)]$  is the ‘‘energy’’ for any  $w \in S_{m,n}$ .

We let  $\mathcal{H}_{m,n}^{\text{prob}} \subset \mathcal{H}_{m,n}$  be the *probability subalgebra*, containing all  $\sum_{w \in S_{m,n}} p_w T_w$  with  $\sum_{w \in S_{m,n}} p_w = 1$  and each  $p_w \geq 0$ . Then any element in  $\mathcal{H}_{m,n}^{\text{prob}}$  corresponds to a probability measure of  $S_{m,n}$ .

Note that for any  $m' \leq m \leq n \leq n'$ ,  $S_{m,n}$  is naturally embedded into  $S_{m',n'}$ . This gives a natural embedding of  $\mathcal{H}_{m,n}$  (resp.,  $\mathcal{H}_{m,n}^{\text{prob}}$ ) into  $\mathcal{H}_{m',n'}$  (resp.,  $\mathcal{H}_{m',n'}^{\text{prob}}$ ).

*Biased card shuffling.* From (3.1), it is not difficult to see that the evolution of the biased card shuffling can be written as multiplying a uniformly chosen random transposition, at rate  $n - m$ . Namely, we denote  $\hat{T}_{m,n} = (n - m)^{-1} \sum_{i=m}^{n-1} T_{s_i}$ , with  $s_i$  being the transposition between  $i$  and  $i + 1$ . Let

$$W_{m,n}(t) = e^{-(n-m)t} \sum_{i=0}^{\infty} \frac{((n-m)t)^i}{i!} \hat{T}_{m,n}^i,$$

where  $\hat{T}_{m,n}^i$  is the product of  $i$  copies of  $\hat{T}_{m,n}$  when  $i \geq 1$ , and  $\hat{T}_{m,n}^0 = T_{id}$  for  $id \in S_{m,n}$  being the identity element (i.e., the identity map of  $\llbracket m, n \rrbracket$ ). Then  $W_{m,n}(t)$  is in  $\mathcal{H}_{m,n}^{\text{prob}}$ , and its corresponding probability measure is just the time  $t$ -distribution of the ASEP on  $\llbracket m, n \rrbracket$ , starting from the identity element.

*Mallows element.* A particularly useful object is the *Mallows element*  $M_{m,n}$  of  $\mathcal{H}_{m,n}^{\text{prob}}$ , which is defined as

$$M_{m,n} = \sum_{w \in S_{m,n}} \mathcal{M}_{m,n}(w) T_w = \sum_{w \in S_{m,n}} q^{\kappa(w)} Z_{m,n} T_w,$$

where  $Z_{m,n}$  is defined in (2.2). Then the probability measure given by  $M_{m,n}$  is just the Mallows measure  $\mathcal{M}_{m,n}$ .

Note that since  $\mathcal{M}_{m,n}$  is the stationary measure of the biased card shuffling, we must have that  $W_{m,n}(t) \rightarrow M_{m,n}$  as  $t \rightarrow \infty$ .

*Involution.* Let  $i : \mathcal{H}_{m,n} \rightarrow \mathcal{H}_{m,n}$  be the linear map where  $i(T_w) = T_{w^{-1}}$  for any  $w \in S_{m,n}$ . It is straightforward to check that  $i$  is an involutive anti-homomorphism; namely, for any  $T_1, T_2 \in \mathcal{H}_{m,n}$  we have  $i(T_1 T_2) = i(T_2) i(T_1)$ , and  $i(i(T_1)) = T_1$  (see [15], Proposition 5.1).

We also have that  $i(W_{m,n}(t)) = W_{m,n}(t)$ , since  $s = s^{-1}$  for any nearest neighbor transposition  $s$  in  $S_{m,n}$ . We note that this equality can be interpreted as the ASEP ‘‘color-to-position symmetry,’’ as proved in [5, 6, 11]. By sending  $t \rightarrow \infty$ , we further have  $i(M_{m,n}) = M_{m,n}$ .

**3.2. Two additional processes.** We now explain our strategy of bounding  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b,+}]$  (thus proving Proposition 2.4).

In  $\mathcal{A}_{N,k,t}^{a,b,+}$ , we consider both  $\zeta^k$ , a process on  $\mathbb{Z}$ , and  $\xi^k$ , a process on  $[N]$ . To relate them, we shall extend  $\xi^k$ , by considering a process on  $\mathbb{Z}$ , which roughly speaking, is the same as  $\xi^k$  is  $[N]$ , while equals 0 is  $[-\infty, 0]$  and equals 1 on  $[N + 1, \infty]$ . This process is also coupled with  $\zeta^k$  under the basic coupling, and we shall compare it with  $\xi^k$ .

For the comparison, note that the initial configuration of this ‘‘extension’’ of  $\xi^k$  is also (roughly) the same as  $\zeta_0^k$  in  $[N]$ , while different in  $[-\infty, 0]$  and  $[N + 1, \infty]$ . We then consider an intermediate process, which is also coupled with  $\xi^k$  and  $\zeta^k$  under the basic coupling. Its initial configuration is the same as  $\zeta_0^k$  in  $[k + 1, \infty]$  (which is identically 0), and the same as the initial configuration of the ‘‘extension’’ of  $\xi^k$  in  $[-\infty, k]$ . This intermediate process together with  $\xi^k$  can be encoded as one ASEP, by placing a second-class particle at each location where they differ. Then comparing these two processes is reduced to tracking the locations of the second-class particles, and that can be achieved using the above involution of

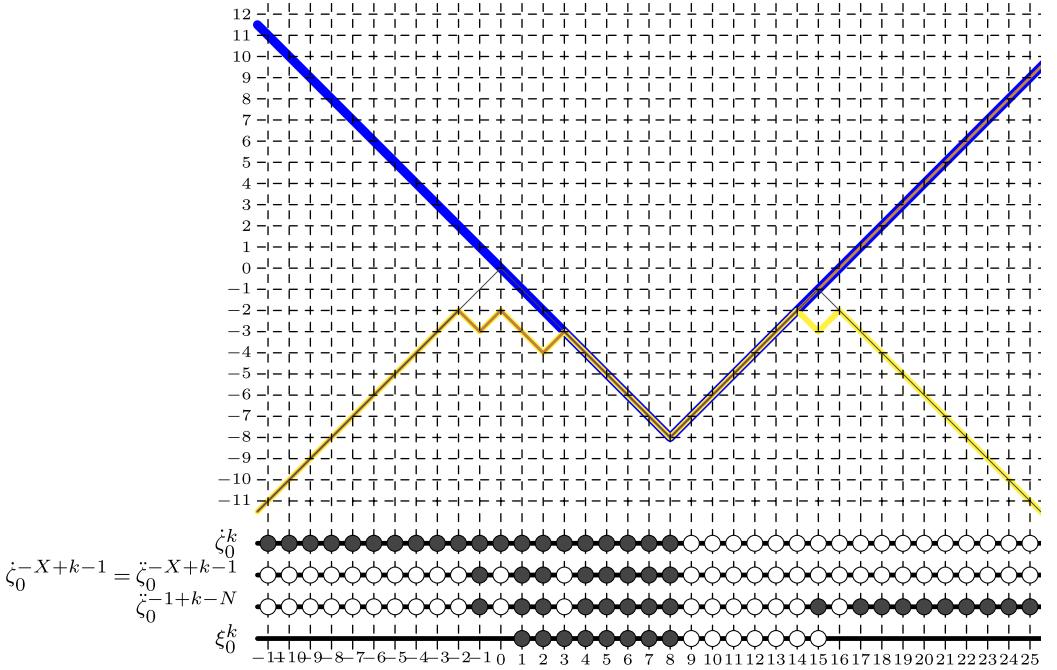


FIG. 3. An illustration of various processes and their height functions: the black, yellow, brown and blue functions are  $h\{\xi_0^k\}$ ,  $h\{\xi_0^{k-1+k-N}\}$ ,  $h\{\xi_0^{k-X+k-1}\} = h\{\xi_0^{k-X+k-1}\}$  and  $h'\{\xi_0^k\}$ , respectively. Note that  $h'\{\xi_0^k\}$  has slope  $-1$  outside  $[-X, X]$ , but that cannot be seen from this figure since we take  $X$  very large (at least  $X > 25$ ). The process  $\xi^{k-1+k-N}$  is the “extension” of  $\xi^k$ , and  $\xi^{k-X+k-1} = \xi^{k-X+k-1}$  is the “intermediate process.”

Hecke algebra. The comparison between the intermediate process and the “extension” of  $\xi^k$  is analyzed similarly. Putting these together, we get estimates on comparing  $\xi^k$  and  $\xi^k$ . See Figure 3 for an illustration of these processes, whose formal definitions are given below.

We start by defining the “extension” of  $\xi^k$  and the intermediate process, using the Hecke algebra Mallows elements. There are two key aspects to note. First, to save notation, we do not formally use the notion of second-class particles, and just write our proof using the multi-species ASEP setup. Second, to work with the Hecke algebra, we shall consider a large finite interval instead of  $\mathbb{Z}$ .

Take an integer  $X$  that is large enough depending on  $N, k, a, b$ . In particular, we require  $X > N$ . We let  $\dot{\xi} = (\dot{\xi}_t)_{t \geq 0}$  and  $\ddot{\xi} = (\ddot{\xi}_t)_{t \geq 0}$  be multispecies TASEPs on  $[-X, X]$ , such that the distribution of  $\dot{\xi}_0$  is given by  $M_{-X, k}$ , and the distribution of  $\ddot{\xi}_0$  is given by  $M_{-X, k} M_{-X+k, X}$  (both as elements of  $\mathcal{H}_{-X, X}^{\text{prob}}$ ). Thus  $\dot{\xi}_0$  and  $\ddot{\xi}_0$  can be described as follows:

$\dot{\xi}_0$ : In  $[-X, k]$ , roughly it is decreasing from  $k$  to  $-X$  (and the distribution is just  $M_{-X, k}$ ). In  $[k+1, X]$ , it is the identity map.

$\ddot{\xi}_0$ : It can be equivalently defined as follows: first, take the random configuration, which is the identity map in  $[-X, -X+k-1]$  and has distribution  $M_{-X+k, X}$  in  $[-X+k, X]$ ; then run the ASEP on  $[-X, k]$  to time infinity (this procedure is called “bringing into  $Q$ -equilibrium” in [15], Section 5.3). Thus, in  $[-X, 0]$ , roughly  $\ddot{\xi}_0$  is decreasing from  $X$  to 0; in  $[1, k]$ , roughly  $\ddot{\xi}_0$  is decreasing from  $-X+k-1$  to  $-X$ ; in  $[k+1, X]$ , roughly  $\ddot{\xi}_0$  is decreasing from  $-1$  to  $-X+k$ .

The process  $\dot{\xi}$  encodes the coupling between  $\xi^k$  and the intermediate process, and  $\ddot{\xi}$  encodes the coupling between the intermediate process and the “extension” of  $\xi^k$ . Namely, if we denote

$$\dot{\xi}_t^i(x) = \mathbb{1}[\dot{\xi}_t(x) \leq i], \quad \ddot{\xi}_t^i(x) = \mathbb{1}[\ddot{\xi}_t(x) \leq i],$$

for each  $i \in \llbracket -X, X \rrbracket$ , we make the following quick observations:

1.  $\dot{\zeta}_0^k(x) = \mathbb{1}[x \leq k]$  for  $x \in \llbracket -X, X \rrbracket$ . In particular, the limit of  $\dot{\zeta}_0^k$  as  $X \rightarrow \infty$  is  $\zeta_0^k$ .
2.  $\dot{\zeta}_0^{-X+k-1}$  and  $\ddot{\zeta}_0^{-X+k-1}$  are equal in distribution: it is straightforward to check that either of them in  $\llbracket -X, k \rrbracket$  has distribution  $\mathcal{P}_{-X,k}^k$ , and either of them is 0 at every location in  $\llbracket k+1, X \rrbracket$ . We couple  $\dot{\zeta}_0^{-X+k-1}$  and  $\ddot{\zeta}_0^{-X+k-1}$  so that they equal almost surely. This is the “intermediate process.”
3. We also consider  $\ddot{\zeta}_0^{-1+k-N}$ . Roughly, it equals 0 in  $\llbracket -X, 0 \rrbracket \cup \llbracket k+1, N \rrbracket$  (recall that  $X > N$ ), and equals 1 in  $\llbracket 1, k \rrbracket \cup \llbracket N+1, \infty \rrbracket$ . It is the “extension of  $\zeta_0^k$ .”

See Figure 3 for an illustration of these processes. Given the initial configurations, we couple the evolution of  $\dot{\zeta}$ ,  $\ddot{\zeta}$  with  $\zeta$ ,  $\xi$ , so that they are all under the same basic coupling (i.e., generated by the same Poisson point processes on  $\mathbb{Z} \times \mathbb{R}_{\geq 0}$ ).

For technical reasons, we need to slightly extend the definition of the height function. For any integers  $m < n$  and  $\omega : \llbracket m, n \rrbracket \rightarrow \{0, 1\}$ , we define  $h' \{\omega\} = h \{\omega'\}$ , where  $\omega' : \mathbb{Z} \rightarrow \{0, 1\}$  is the function satisfying  $\omega'(x) = \omega'(x)$  for  $x \in \llbracket m, n \rrbracket$ , and  $\omega'(x) = 1$  for  $x \in \mathbb{Z} \setminus \llbracket m, n \rrbracket$ . Note that compared with  $h \{\omega\}$ , the only difference is that we let the slope in  $\llbracket -\infty, m \rrbracket$  be  $-1$  rather than 1.

We then have the following ordering property (see Figure 3).

LEMMA 3.1. *We have  $h \{\ddot{\zeta}_t^{-1+k-N}\} \leq h \{\xi_t^k\}$  and  $h \{\ddot{\zeta}_t^{-1+k-N}\} \leq h \{\dot{\zeta}_t^{-X+k-1}\} = h \{\ddot{\zeta}_t^{-X+k-1}\} \leq h' \{\dot{\zeta}_t^k\}$  for any  $t \geq 0$ .*

PROOF. We first check that

$$(3.2) \quad h \{\ddot{\zeta}_0^{-1+k-N}\} \leq h \{\xi_0^k\}.$$

Note that for any  $x \in \mathbb{Z}$ , we always have that  $h \{\ddot{\zeta}_0^{-1+k-N}\}(x) \leq x \wedge (2N - 2k - x)$ , since (by definition) we have  $h \{\ddot{\zeta}_0^{-1+k-N}\}(-X) = -X$ , and  $\sum_{x \in \llbracket -X, X \rrbracket} h \{\ddot{\zeta}_0^{-1+k-N}\}(x) = X - N + k$ ; thus  $h \{\ddot{\zeta}_0^{-1+k-N}\}(X) = 2N - 2k - X$ , and that  $h \{\ddot{\zeta}_0^{-1+k-N}\}$  is 1-Lipschitz. From the definition of  $\ddot{\zeta}_0$ , for each  $x \in \llbracket -X, -X + k - 1 \rrbracket$ , we can find  $y \in \llbracket -X, k \rrbracket$  with  $\ddot{\zeta}_0(y) = x$ . This implies that

$$\sum_{x \in \llbracket -X, k \rrbracket} h \{\ddot{\zeta}_0^{-1+k-N}\}(x) \geq k,$$

so we have  $h \{\ddot{\zeta}_0^{-1+k-N}\}(k) \leq -k$ . Then for each  $x \in \llbracket 1, k \rrbracket$  there is  $h \{\ddot{\zeta}_0^{-1+k-N}\}(x) \leq -x$ , and for each  $x \in \llbracket k+1, N \rrbracket$  there is  $h \{\ddot{\zeta}_0^{-1+k-N}\}(x) \leq x - 2k$  (since  $h \{\ddot{\zeta}_0^{-1+k-N}\}$  is 1-Lipschitz). Putting all the bounds together, we get (3.2), since  $h \{\xi_0^k\}(x) = x \wedge (2N - 2k - x)$  for  $x \in \llbracket -\infty, 0 \rrbracket \cup \llbracket N, \infty \rrbracket$ ,  $h \{\xi_0^k\}(x) = -x$  for  $x \in \llbracket 0, k \rrbracket$  and  $h \{\xi_0^k\}(x) = x - 2k$  for  $x \in \llbracket k, N \rrbracket$ .

We next check that

$$(3.3) \quad h \{\ddot{\zeta}_0^{-1+k-N}\} \leq h \{\ddot{\zeta}_0^{-X+k-1}\} = h \{\dot{\zeta}_0^{-X+k-1}\} \leq h' \{\dot{\zeta}_0^k\}.$$

The first inequality holds since  $\sum_{y \in \llbracket -X, x \rrbracket} \ddot{\zeta}_0^{-1+k-N}(y) \geq \sum_{y \in \llbracket -X, x \rrbracket} \ddot{\zeta}_0^{-X+k-1}(y)$ , for any  $x \in \llbracket -X, X \rrbracket$ . The equality is by the second point above.

For the second inequality, by the second point above we have that

$$\sum_{x \in \llbracket -X, k \rrbracket} h \{\dot{\zeta}_0^{-X+k-1}\}(x) = k,$$

so we have  $h\{\dot{\zeta}_0^{-X+k-1}\}(k) = -k$ . Then for each  $x \in \llbracket 1, k \rrbracket$  there is  $h\{\dot{\zeta}_0^{-X+k-1}\}(x) \leq -x$ , and for each  $x \in \llbracket k+1, X \rrbracket$  there is  $h\{\dot{\zeta}_0^{-X+k-1}\}(x) \leq x - 2k$  (since  $h\{\dot{\zeta}_0^{-X+k-1}\}$  is 1-Lipschitz). We also have that  $h\{\dot{\zeta}_0^{-X+k-1}\}(x) \leq x \wedge (2X - 2k - x)$  for any  $x \in \mathbb{Z}$ , since  $\sum_{x \in \llbracket -X, X \rrbracket} h\{\dot{\zeta}_0^{-X+k-1}\}(x) = k$  (thus  $h\{\dot{\zeta}_0^{-X+k-1}\}(X) = X - 2k$ ), and that  $h\{\dot{\zeta}_0^{-X+k-1}\}$  is 1-Lipschitz. Putting all the bounds together, we get the second inequality in (3.2), since  $h'\{\dot{\zeta}_0^k\}(x) = -x$  for  $x \in \llbracket -\infty, k \rrbracket$ ,  $h'\{\dot{\zeta}_0^k\}(x) = x - 2k$  for  $x \in \llbracket k+1, X \rrbracket$  and  $h'\{\dot{\zeta}_0^k\}(x) = 2X - 2k - x$  for  $x \in \llbracket X, \infty \rrbracket$ .

Finally, using arguments similar to those in the proof of Lemma 2.2, the ordering of these height functions is preserved under the basic coupling, so the conclusion follows.  $\square$

We now define the events:

$$\dot{\mathcal{E}}: h'\{\dot{\zeta}_t^k\}(N-k) > N-k+b \text{ and } h\{\dot{\zeta}_t^{-X+k-1}\}(N-k+a) < N-k+a.$$

$$\ddot{\mathcal{E}}: h\{\ddot{\zeta}_t^{-X+k-1}\}(N-k+a) = N-k+a \text{ and } h\{\ddot{\zeta}_t^{-1+k-N}\}(N-k+a) < N-k-a.$$

By Lemma 3.1, and the fact that the  $X \rightarrow \infty$  limit of  $\dot{\zeta}_t^k$  is  $\dot{\zeta}_t^k$ , we have that the  $X \rightarrow \infty$  limit of  $\dot{\mathcal{E}} \cup \ddot{\mathcal{E}}$  contains  $\mathcal{A}_{N,k,t}^{a,b,+}$ .

Thus it suffices to bound the probabilities of the events  $\dot{\mathcal{E}}$  and  $\ddot{\mathcal{E}}$ . For this, note that using the above stated properties of the involution  $\mathbf{i}$ , we have

$$(3.4) \quad W_{-X,X}(t)M_{-X,k} = \mathbf{i}(W_{-X,X}(t))\mathbf{i}(M_{-X,k}) = \mathbf{i}(M_{-X,k}W_{-X,X}(t)),$$

and

$$(3.5) \quad \begin{aligned} W_{-X,X}(t)M_{-X,k}M_{-X+k,X} \\ = \mathbf{i}(W_{-X,X}(t))\mathbf{i}(M_{-X,k})\mathbf{i}(M_{-X+k,X}) = \mathbf{i}(M_{-X+k,X}M_{-X,k}W_{-X,X}(t)). \end{aligned}$$

These two identities can be interpreted as distributional identities, and are key inputs in bounding  $\mathbb{P}[\dot{\mathcal{E}}]$  and  $\mathbb{P}[\ddot{\mathcal{E}}]$ .

LEMMA 3.2. *We have  $\mathbb{P}[\dot{\mathcal{E}}] \leq Cq^{b/4-a/2}$  where  $C$  is a constant depending only on  $q$ .*

PROOF. Assuming  $\dot{\mathcal{E}}$ , we must have that  $h'\{\dot{\zeta}_t^k\}(N-k+a) > N-k+b-a$  (since  $h'\{\dot{\zeta}_t^k\}$  is 1-Lipschitz) and  $h\{\dot{\zeta}_t^{-X+k-1}\}(N-k+a) < N-k+a$ . In other words, we have

$$|\{x \in \mathbb{Z} : N-k+a+1 \leq x \leq X, \dot{\zeta}_t(x) \leq k\}| > k-a+b/2,$$

and

$$|\{x \in \mathbb{Z} : N-k+a+1 \leq x \leq X, \dot{\zeta}_t(x) \leq -X+k-1\}| < k.$$

Now we use the distributional identity implied by (3.4). Let  $\dot{\zeta}_t$  be the inverse of  $\dot{\zeta}_t$  (as a group element in  $S_{-X,X}$ ). Then the above two inequalities are equivalent to

$$(3.6) \quad |\{x \in \mathbb{Z} : \dot{\zeta}_t(x) \geq N-k+a+1, -X \leq x \leq k\}| > k-a+b/2,$$

and

$$(3.7) \quad |\{x \in \mathbb{Z} : \dot{\zeta}_t(x) \geq N-k+a+1, -X \leq x \leq -X+k-1\}| < k,$$

respectively. Note that these events only concern  $x \mapsto \mathbb{1}[\dot{\zeta}_t(x) \leq N-k+a]$ , which is a random configuration in  $\{0, 1\}^{\llbracket -X, X \rrbracket}$ . We next analyze its distribution.

By (3.4), the law of  $\dot{\zeta}_t$  is given by  $M_{-X,k}W_{-X,X}(t)$ , meaning that we can first sample  $w$  from the law given by  $W_{-X,X}(t)$ , then sample  $\dot{\zeta}_t$  from the law given by  $M_{-X,k}T_w$ . Equivalently,  $\dot{\zeta}_t$  is the same as  $w$  on  $\llbracket k+1, X \rrbracket$ ; and on  $\llbracket -X, k \rrbracket$ , the law of  $\dot{\zeta}_t$  (conditioned on  $w$ ) is the weak limit of running the (multispecies) ASEP from  $w$  to time  $\infty$ .

Let  $\phi : \llbracket -X, k \rrbracket \rightarrow \{w(x) : x \in \llbracket -X, k \rrbracket\}$  be the unique increasing bijection determined by  $w$  in  $\llbracket -X, k \rrbracket$ . Then  $\dot{\zeta}_t$  in  $\llbracket -X, k \rrbracket$  can be written as  $\phi \circ v$ , for  $v$  sampled from  $\mathcal{M}_{-X, k}$ . Now the event in (3.6) is equivalent to that

$$|\{x \in \mathbb{Z} : \phi(x) \geq N - k + a + 1, -X \leq x \leq k\}| > k - a + b/2.$$

Since  $\phi$  is increasing, this is further equivalent to that

$$\phi(-\lfloor -a + b/2 \rfloor) \geq N - k + a + 1.$$

The event in (3.7) is equivalent to that

$$\exists x \in \llbracket -X, -X + k - 1 \rrbracket, \quad \dot{\zeta}_t(x) = \phi(v(x)) \leq N - k + a.$$

So (since  $\phi$  is increasing) these two events together imply that  $v(x) < a - b/2$  for some  $x \in \llbracket -X, -X + k - 1 \rrbracket$ . Note that since  $v$  is sampled from  $\mathcal{M}_{-X, k}$ , there is always  $\sum_{-X \leq j \leq k} \mathbb{1}[v(j) \geq a - b/2] > k - (a - b/2)$  (when  $X$  is large enough depending on  $a, b$ ). If there is some  $x \in \llbracket -X, -X + k - 1 \rrbracket$  with  $v(x) < a - b/2$ , we would have  $\sum_{-X \leq j \leq -X + k - 1} \mathbb{1}[v(j) \geq a - b/2] \leq k - 1$ , so  $\sum_{-X + k \leq j \leq k} \mathbb{1}[v(j) \geq a - b/2] > b/2 - a + 1$ , which further implies that  $\sum_{-X \leq i < j \leq k} \mathbb{1}[v(i) < a - b/2, v(j) \geq a - b/2] > b/2 - a + 1$ . Note that  $x \mapsto \mathbb{1}[v(x) < a - b/2]$  has distribution  $\mathcal{P}_{-X, k}^{\lceil a-b/2 \rceil + X}$ , so by Lemma 2.1, the probability of this event is at most  $Cq^{b/4-a/2}$ , for some  $C > 0$  depending only on  $q$ . Thus the conclusion follows.  $\square$

The bound for  $\mathbb{P}[\ddot{\mathcal{E}}]$  is proved using a similar strategy.

LEMMA 3.3. *We have  $\mathbb{P}[\ddot{\mathcal{E}}] \leq Cq^{a/2}$  where  $C$  is a constant depending only on  $q$ .*

PROOF. Assuming  $\ddot{\mathcal{E}}$ , we must have

$$|\{x \in \mathbb{Z} : N - k + a + 1 \leq x \leq X, \ddot{\zeta}_t(x) \leq -X + k - 1\}| = k,$$

and

$$|\{x \in \mathbb{Z} : N - k + a + 1 \leq x \leq X, \ddot{\zeta}_t(x) \leq -1 + k - N\}| < X - N + k - a.$$

We let  $\ddot{\zeta}_t$  be the inverse of  $\dot{\zeta}_t$  (as a group element in  $S_{-X, X}$ ). Then the above two relations are equivalent to

$$(3.8) \quad |\{x \in \mathbb{Z} : \ddot{\zeta}_t(x) \geq N - k + a + 1, -X \leq x \leq -X + k - 1\}| = k,$$

and

$$(3.9) \quad |\{x \in \mathbb{Z} : \ddot{\zeta}_t(x) \geq N - k + a + 1, -X \leq x \leq -1 + k - N\}| < X - N + k - a,$$

respectively. By (3.5), the law of  $\ddot{\zeta}_t$  is given by  $M_{-X+k, X} M_{-X, k} W_{-X, X}(t)$ , meaning that we can first sample  $w$  from the law given by  $W_{-X, X}(t)$ , and then sample  $\ddot{\zeta}_t$  from the law given by  $M_{-X+k, X} M_{-X, k} T_w$ . Equivalently,  $\ddot{\zeta}_t$  can be obtained from  $w$ , by first running the (multispecies) ASEP on  $\llbracket -X, k \rrbracket$  to time  $\infty$  to get  $w'$ , then running the (multispecies) ASEP on  $\llbracket -X + k, X \rrbracket$  (from  $w'$ ) to time  $\infty$  to get  $\ddot{\zeta}_t$ .

From this construction of  $\ddot{\zeta}_t$ , we have that  $\ddot{\zeta}_t$  and  $w'$  are the same in  $\llbracket -X, -X + k - 1 \rrbracket$ . Then the event in (3.8) is equivalent to that

$$\ddot{\zeta}_t(x) = w'(x) \geq N - k + a + 1 \quad \forall x \in \llbracket -X, -X + k - 1 \rrbracket,$$

and this is further equivalent to that

$$(3.10) \quad |\{x \in \llbracket -X + k, X \rrbracket : w'(x) \leq N - k + a\}| = N - k + a + X + 1.$$

Let  $\phi : \llbracket -X + k, X \rrbracket \rightarrow \{w'(x) : x \in \llbracket -X + k, X \rrbracket\}$  be the unique increasing bijection determined by  $w'$  in  $\llbracket -X + k, X \rrbracket$ . Then  $\zeta_t$  in  $\llbracket -X + k, X \rrbracket$  can be written as  $\phi \circ v$ , for  $v$  sampled from  $\mathcal{M}_{-X+k, X}$ . Under the event in (3.10), we have that  $\phi(N + a) \leq N - k + a$  since  $\phi$  is increasing. The event in (3.9) is equivalent to that

$$\exists x \in \llbracket k - N, X \rrbracket, \quad \zeta_t(x) = \phi(v(x)) \geq N - k + a + 1.$$

Then under the events in (3.10) and (3.9), there exists  $x \in \llbracket k - N, X \rrbracket$  such that  $v(x) \geq N + a + 1$ . This further implies that

$$\sum_{-X+k \leq i < j \leq X} \mathbb{1}[v(i) \leq N + a, v(j) \geq N + a + 1] \geq \sum_{i \in \llbracket -X+k, k-N-1 \rrbracket} \mathbb{1}[v(i) \leq N + a] \geq a + 1.$$

Note that  $x \mapsto \mathbb{1}[v(x) \leq N + a]$  has distribution  $\mathcal{P}_{-X+k, X}^{N+a+X-k+1}$ , so by Lemma 2.1 the probability of this event is at most  $Cq^{a/2}$ , for some  $C > 0$  depending only on  $q$ . Thus the conclusion follows.  $\square$

Putting all these estimates together, we can now upper bound  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b}]$ .

**PROOF OF PROPOSITION 2.4.** From the definition of  $\dot{\mathcal{E}}$  and  $\ddot{\mathcal{E}}$ , we have that  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b,+}] \leq \limsup_{X \rightarrow \infty} \mathbb{P}[\dot{\mathcal{E}} \cup \ddot{\mathcal{E}}] \leq \limsup_{X \rightarrow \infty} \mathbb{P}[\dot{\mathcal{E}}] + \mathbb{P}[\ddot{\mathcal{E}}]$ . Then by Lemmas 3.2 and 3.3, we have  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b,+}] \leq Cq^{(b/4-a/2)\wedge a/2}$ , for  $C$  depending only on  $q$ . By symmetry, the same upper bound holds for  $\mathbb{P}[\mathcal{A}_{N,k,t}^{a,b,-}]$ . Thus the conclusion follows.  $\square$

**4. Hitting time bound.** In this section, we prove Proposition 2.5, for which we need the following known estimate on the “ground state hitting time” of the single-species ASEP.

For any  $n \in \mathbb{Z}$  and  $m \in \mathbb{N}$ , we let  $\eta^{n,m} = (\eta_t^{n,m})_{t \geq 0}$  be the (single-species) ASEP, with initial condition given by  $\eta_0^{n,m}(x) = 1$  for  $x \in \llbracket n - m + 1, n \rrbracket$  or  $x \in \llbracket n + m + 1, \infty \rrbracket$ , and  $\eta_0^{n,m}(x) = 0$  for  $x \in \llbracket -\infty, n - m \rrbracket$  or  $x \in \llbracket n + 1, n + m \rrbracket$ .

The following estimate is implied by [9], Theorem 1.9, which is also used in [15].

**LEMMA 4.1.** *There is a constant  $C_* > 0$  depending only on  $q$ , such that the following is true. Let  $H > 0$  be the smallest number with  $\eta_H^{n,m}(x) = \mathbb{1}[x \geq n + 1]$  for any  $x \in \mathbb{Z}$ . Then  $\mathbb{P}[H > C_*m] < 1/m$ .*

Below we also assume that  $C_* \in \mathbb{Z}$ , since otherwise, we can replace it with  $\lceil C_* \rceil$ .

Recall the setting of Proposition 2.5 and the parameters  $N, k, t, a$ . The idea of proving Proposition 2.5 is to split the time interval  $[0, \log(N)^4]$  into segments each with length  $C_*a$ , and use Lemma 4.1 to conclude that the probability of not reaching the “ground state” in each interval is at most  $1/a$ . Then we can multiply these probabilities for all the intervals, since the ASEP is Markovian.

For each  $i \in \mathbb{Z}_{\geq 0}$ , we let  $\mathcal{E}_i$  be the event where

$$h\{\xi_{t+iC_*a}^k\}(N - k - a) = h\{\xi_{t+iC_*a}^k\}(N - k + a) = N - k - a,$$

and  $\mathcal{E}'_i$  be the event where

$$h\{\xi_{t+iC_*a+s}^k\}(N - k) < N - k \quad \forall s \in [0, C_*a],$$

that is,  $\xi_{t+iC_*a+s}^k$  is not at the “ground state” for any  $s \in [0, C_*a]$ . Here,  $C_*$  is the constant in Lemma 4.1.

**LEMMA 4.2.** *We have  $\mathbb{P}[\mathcal{E}'_i | \xi_{t+iC_*a}^k] \leq 1/a$  for any  $\xi_{t+iC_*a}^k$  under the event  $\mathcal{E}_i$ .*

PROOF. Note that  $h\{\eta_0^{N-k,a}\}(N-k-a) = h\{\eta_0^{N-k,a}\}(N-k+a) = N-k-a$ , and  $h\{\eta_0^{N-k,a}\}$  has slope 1 in  $\llbracket -\infty, N-k-a \rrbracket \cup \llbracket N-k, N-k+a \rrbracket$ , and has slope  $-1$  in  $\llbracket N-k-a, N-k \rrbracket \cup \llbracket N-k+a, \infty \rrbracket$ . Thus assuming  $\mathcal{E}_i$ , we must have  $h\{\eta_0^{N-k,a}\} \leq h\{\xi_{t+iC_*a}^k\}$  since  $h\{\xi_{t+iC_*a}^k\}$  is 1-Lipschitz.

We couple  $(\xi_{t+iC_*a+s}^k)_{s \geq 0}$  and  $\eta^{N-k,a}$  under the basic coupling. Then  $h\{\eta_s^{N-k,a}\} \leq h\{\xi_{t+iC_*a+s}^k\}$  for any  $s \geq 0$ , since (by arguments similar to those in the proof of Lemma 2.2) ordering of these height functions is preserved under the basic coupling.

Now by Lemma 4.1, with probability  $\geq 1 - 1/a$  we have that  $\eta_s^{N-k,a}(x) = \mathbb{1}[x \geq N-k+1]$  for any  $x \in \mathbb{Z}$ , thus  $h\{\eta_s^{N-k,a}\}(N-k) = N-k$ , for some  $s \in [0, C_*a]$ . This implies that we must have  $h\{\xi_{t+iC_*a+s}^k\}(N-k) = N-k$ , so  $\mathcal{E}'_i$  does not hold. Thus the conclusion follows.  $\square$

PROOF OF PROPOSITION 2.5. Under the event  $\mathcal{B}_{N,k,t}^{a,\log(N)^4}$ , we must have that  $\mathcal{E}_i$  and  $\mathcal{E}'_i$  hold, for each  $i \in \llbracket 0, \log(N)^4/(C_*a) - 1 \rrbracket$ . By Lemma 4.2 and that the ASEP is Markovian, we have that

$$\mathbb{P}\left[\mathcal{E}'_i \mid \bigcap_{0 \leq j \leq i} \mathcal{E}_j \cap \bigcap_{0 \leq j < i} \mathcal{E}'_j\right] \leq 1/a.$$

Thus by taking the product of this over  $i \in \llbracket 0, \log(N)^4/(C_*a) - 1 \rrbracket$ , and using that  $2 \leq a \leq \log(N)^2$ , we have  $\mathbb{P}[\mathcal{B}_{N,k,t}^{a,\log(N)^4}] < a^{-\lfloor \log(N)^4/(C_*a) \rfloor} < e^{-\log(2) \lfloor \log(N)^2/C_* \rfloor}$ . So, the conclusion follows.  $\square$

**5. ASEP shift-invariance and convergence to the KPZ fixed point.** In this section, we deduce Proposition 2.3.

The first input we need is the shift-invariance property of the multispecies ASEP, which is deduced from the shift-invariance property of the colored six-vertex model from [13, 22]. Namely, we obtain the following result. Recall  $\zeta$  (the multispecies ASEP on  $\mathbb{Z}$ ) and its projections from Section 2.2.

**PROPOSITION 5.1.** *For any  $N \in \mathbb{N}$  and  $b \in \mathbb{R}$ ,  $t \geq 0$ , the event*

$$h\{\xi_t^k\}(N-k) > N-k+b \quad \forall k \in \llbracket 0, N \rrbracket,$$

*has the same probability as the event*

$$h\{\xi_t^0\}(N-2k) > N+b \quad \forall k \in \llbracket 0, N \rrbracket.$$

This is deduced directly from a special case of [13], Theorem 1.2, or [22], Theorem 1.6, using the limit transition from the colored six-vertex model to the multispecies ASEP. Such transition was first observed in [12], and then proved in [1] in the colorless setting; the colored version follows the same proof. See also [50], Section 3.1. We omit the proof of Proposition 5.1.

Now we just need to analyze the process  $\zeta^0 = (\zeta_t^0)_{\geq 0}$ , which is the single-species ASEP with step initial condition, and study how the minimum of its height function grows. By skew-time reversibility of ASEP (stated as, e.g., (1.4) of [41]), this is equivalent to understanding the one-point height function distribution of the single-species ASEP with flat (periodic) initial data (i.e., the process  $\eta^{N,*}$  to be introduced in the proof of Lemma 5.3). We then need that distribution (under scaling) to converge to the Tracy–Widom GOE distribution. Such convergence has been nonrigorously derived in [39], by formal asymptotic analysis of a Fredholm Pfaffian formula. It is then rigorously implied by the more recently proved

convergence of the single-species ASEP [41] to the KPZ fixed point, which is a random process  $\mathfrak{h} : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}$  and can be described as a continuous time Markov chain  $\mathfrak{h}(t, \cdot)$  with explicit transition probabilities. The state space of this Markov chain is the upper semi-continuous (UC) space, containing all upper semicontinuous functions  $f : \mathbb{R} \rightarrow [-\infty, \infty)$  with  $f(x) < C(|x| + 1)$  for some  $C < \infty$  and  $f(x) > -\infty$  for some  $x$ . We refer the readers to [37], Definition 3.12, for the precise definition of the KPZ fixed point.

We need the following convergence of the single-species ASEP to the KPZ fixed point.

**PROPOSITION 5.2** ([41], Theorem 2.2(2)). *For each  $\epsilon > 0$ , let  $\eta^\epsilon = (\eta_t^\epsilon)_{t \geq 0}$  be the (single-species) ASEP with deterministic initial configuration  $\eta_0^\epsilon$ , such that  $\eta_0^\epsilon(x)$  stabilizes as  $x \rightarrow -\infty$  (so that its height function is well defined) and let  $g_\epsilon \in \mathbb{R}$ . Let  $\mathfrak{h}_0 : \mathbb{R} \rightarrow \mathbb{R}$  be continuous with  $|\mathfrak{h}_0(x)| < C(1 + |x|^{1/2})$  for some  $C < \infty$ . Suppose that*

$$x \mapsto -\epsilon^{1/2} h\{\eta_0^\epsilon\}(2\epsilon^{-1}x) + g_\epsilon$$

*converges to  $\mathfrak{h}_0$ , uniformly on compact sets. Then for any  $t > 0$ ,*

$$x \mapsto -\epsilon^{1/2} h\{\eta_{2\epsilon^{-3/2}t}^\epsilon\}(2\epsilon^{-1}x) + g_\epsilon + \epsilon^{-1}(1-q)t$$

*converges to  $\mathfrak{h}((1-q)t, \cdot)$  in distribution, in the uniform on compact sets topology, with  $\mathfrak{h}$  being the KPZ fixed point with initial data  $\mathfrak{h}(0, \cdot) = \mathfrak{h}_0$ . Here, we regard  $h\{\eta_0^\epsilon\}$  and  $h\{\eta_{2\epsilon^{-3/2}t}^\epsilon\}$  as functions on  $\mathbb{R}$ , by linearly interpolating between integers.*

Using Propositions 5.1 and 5.2, we now prove the following statement, which directly implies Proposition 2.3.

**LEMMA 5.3.** *For any  $\tau \in \mathbb{R}$ , we have that*

$$-2^{2/3}N^{-1/3} \min\{h\{\zeta_{2(1-q)^{-1}(N+\tau N^{1/3})}^k\}(N-k) + k : k \in \llbracket 0, N \rrbracket\} + 2^{2/3}N^{2/3} + 2^{2/3}\tau$$

*converges in distribution to the Tracy–Widom GOE distribution, as  $N \rightarrow \infty$ .*

**PROOF.** Let  $\eta^{N,*} = (\eta_t^{N,*})_{t \geq 0}$  be the single-species ASEP with the following initial configuration:  $\eta_0^{N,*}(x) = 1$  for  $x \in \llbracket -\infty, -N \rrbracket$ ; and  $\eta_0^{N,*}(x) = 0$  for  $x \in \llbracket N+1, \infty \rrbracket$ ; and  $\eta_0^{N,*}(-N-1+2x) = 0$ ,  $\eta_0^{N,*}(-N+2x) = 1$ , for  $x \in \llbracket 1, N \rrbracket$ . Thus we have

$$h\{\eta_0^{N,*}\}(x) = \begin{cases} -x & \text{when } x \in \llbracket -\infty, -N \rrbracket, \\ N & \text{when } x \in 2\llbracket 0, N \rrbracket - N, \\ N+1 & \text{when } x \in 2\llbracket 0, N-1 \rrbracket - N+1, \\ x & \text{when } x \in \llbracket N, \infty \rrbracket. \end{cases}$$

By skew-time reversibility of ASEP (see, e.g., (1.4) of [41]), for any  $t \geq 0$  and  $b \in \mathbb{R}$ , we have

$$(5.1) \quad \mathbb{P}[h\{\zeta_t^0\}(N-2k) > N+b, \forall k \in \llbracket 0, N \rrbracket] = \mathbb{P}[h\{\eta_t^{N,*}\}(0) > 2N+b].$$

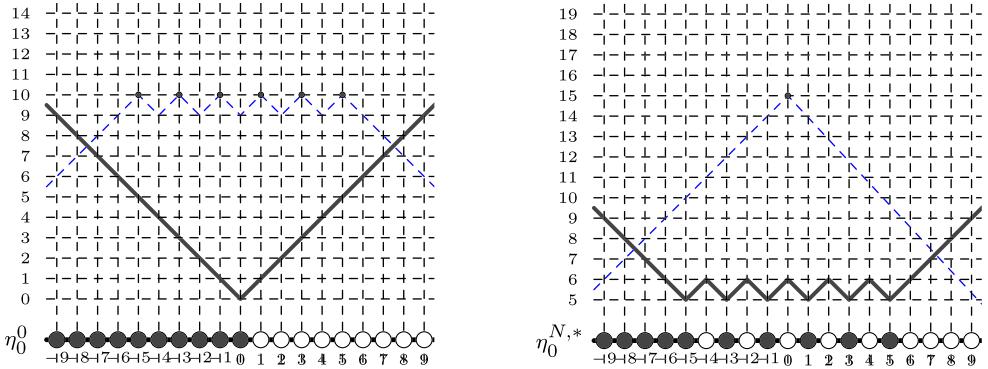
See Figure 4. Then by Proposition 5.1 we have

$$\mathbb{P}[h\{\zeta_t^k\}(N-k) > N-k+b, \forall k \in \llbracket 0, N \rrbracket] = \mathbb{P}[h\{\eta_t^{N,*}\}(0) > 2N+b].$$

Namely, we have

$$(5.2) \quad \min\{h\{\zeta_t^k\}(N-k) + k : k \in \llbracket 0, N \rrbracket\} - N \stackrel{d}{=} h\{\eta_t^{N,*}\}(0) - 2N,$$

where  $\stackrel{d}{=}$  denotes equal in distribution.



(a) An illustration of the configuration  $\zeta_0^0$  and its height function  $h\{\zeta_0^0\}$ . The above points indicate the lower bounds for  $h\{\zeta_t^0\}$  given in (5.1). If these bounds hold,  $h\{\zeta_t^0\}$  is also lower bounded by the dashed blue function since it is 1-Lipschitz.

(b) An illustration of the configuration  $\eta_0^{N,*}$  and its height function  $h\{\eta_0^{N,*}\}$ . The above point indicates the lower bound for  $h\{\eta_t^{N,*}\}$  given in (5.1). If these bounds hold,  $h\{\eta_t^{N,*}\}$  is also lower bounded by the dashed blue function, since it is 1-Lipschitz.

FIG. 4. An illustration of the skew-time reversibility, as given by, for example (1.4) of [41].

Note that for any  $\tau \in \mathbb{R}$ , we have that

$$x \mapsto -(N + \tau N^{1/3})^{-1/3} h\{\eta_0^{N,*}\} (2(N + \tau N^{1/3})^{2/3} x) + (N + \tau N^{1/3})^{-1/3} N$$

converges (uniformly in compact sets) to 0, as  $N \rightarrow \infty$ . Then by Proposition 5.2, we have that

$$-(N + \tau N^{1/3})^{-1/3} h\{\eta_{2(1-q)^{-1}(N+\tau N^{1/3})}^{N,*}\}(0) + (N + \tau N^{1/3})^{-1/3} N + (N + \tau N^{1/3})^{2/3}$$

converges in distribution to  $\mathfrak{h}(1, 0)$  as  $N \rightarrow \infty$ , for  $\mathfrak{h}$  being the KPZ fixed point with initial data  $\mathfrak{h}(0, \cdot) = 0$ . We note that with such initial data,  $x \mapsto 2^{-1/3} \mathfrak{h}(1, 2^{2/3} x)$  is the so-called Airy<sub>1</sub> process, and  $\mathfrak{h}(1, 0)$  has the same distribution as  $\sup_{x \in \mathbb{R}} \mathcal{A}(x) - x^2$ , for  $\mathcal{A}(x) - x^2$  being the parabolic Airy<sub>2</sub> process (see, e.g., (4.15) and Example 4.20 in [37]). Thus  $2^{2/3} \mathfrak{h}(1, 0)$  is Tracy–Widom GOE (see, e.g., (1.25) of [40]). By multiplying  $2^{2/3} N^{-1/3} (N + \tau N^{1/3})^{1/3}$ , we have that

$$-2^{2/3} N^{-1/3} h\{\eta_{2(1-q)^{-1}(N+\tau N^{1/3})}^{N,*}\}(0) + 2^{5/3} N^{2/3} + 2^{2/3} \tau$$

converges in distribution to the Tracy–Widom GOE distribution, as  $N \rightarrow \infty$ . Using (5.2), the conclusion follows.  $\square$

**Acknowledgments.** The author would like to thank Jimmy He and Amol Aggrawal for several valuable conversations, and Hubert Lacoin for his comments on an earlier version of this paper. The author would also like to thank the anonymous referees for carefully reading this manuscript and providing valuable feedback that helped improve the exposition. Part of this work was completed when the author was a Ph.D. student at Princeton University, Department of Mathematics.

**Funding.** The research of the author is supported by the Miller Institute for Basic Research in Science at the University of California, Berkeley, and NSF Grant DMS-2246664.

## REFERENCES

- [1] AGGARWAL, A. (2017). Convergence of the stochastic six-vertex model to the ASEP: Stochastic six-vertex model and ASEP. *Math. Phys. Anal. Geom.* **20** Paper No. 3. [MR3589552](#) <https://doi.org/10.1007/s11040-016-9235-8>
- [2] AGGARWAL, A., CORWIN, I. and GHOSAL, P. (2023). The ASEP speed process. *Adv. Math.* **422** Paper No. 109004, 57. [MR4576839](#) <https://doi.org/10.1016/j.aim.2023.109004>
- [3] ALDOUS, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability, XVII. Lecture Notes in Math.* **986** 243–297. Springer, Berlin. [MR0770418](#) <https://doi.org/10.1007/BFb0068322>
- [4] ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348. [MR0841111](#) <https://doi.org/10.2307/2323590>
- [5] AMIR, G., ANGEL, O. and VALKÓ, B. (2011). The TASEP speed process. *Ann. Probab.* **39** 1205–1242. [MR2857238](#) <https://doi.org/10.1214/10-AOP561>
- [6] ANGEL, O., HOLROYD, A. and ROMIK, D. (2009). The oriented swap process. *Ann. Probab.* **37** 1970–1998. [MR2561438](#) <https://doi.org/10.1214/09-AOP456>
- [7] BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313. [MR1161056](#)
- [8] BEN-HAMOU, A. and SALEZ, J. (2017). Cutoff for nonbacktracking random walks on sparse random graphs. *Ann. Probab.* **45** 1752–1770. [MR3650414](#) <https://doi.org/10.1214/16-AOP1100>
- [9] BENJAMINI, I., BERGER, N., HOFFMAN, C. and MOSSEL, E. (2005). Mixing times of the biased card shuffling and the asymmetric exclusion process. *Trans. Amer. Math. Soc.* **357** 3013–3029. [MR2135733](#) <https://doi.org/10.1090/S0002-9947-05-03610-X>
- [10] BERTINI, L. and GIACOMIN, G. (1997). Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** 571–607. [MR1462228](#) <https://doi.org/10.1007/s002200050044>
- [11] BORODIN, A. and BUFETOV, A. (2021). Color-position symmetry in interacting particle systems. *Ann. Probab.* **49** 1607–1632. [MR4260463](#) <https://doi.org/10.1214/20-aop1463>
- [12] BORODIN, A., CORWIN, I. and GORIN, V. (2016). Stochastic six-vertex model. *Duke Math. J.* **165** 563–624. [MR3466163](#) <https://doi.org/10.1215/00127094-3166843>
- [13] BORODIN, A., GORIN, V. and WHEELER, M. (2022). Shift-invariance for vertex models and polymers. *Proc. Lond. Math. Soc. (3)* **124** 182–299. [MR4400934](#) <https://doi.org/10.1112/plms.12427>
- [14] BUFETOV, A., GORIN, V. and ROMIK, D. (2022). Absorbing time asymptotics in the oriented swap process. *Ann. Appl. Probab.* **32** 753–763. [MR4414694](#) <https://doi.org/10.1214/21-aap1695>
- [15] BUFETOV, A. and NEJJAR, P. (2022). Cutoff profile of ASEP on a segment. *Probab. Theory Related Fields* **183** 229–253. [MR4421174](#) <https://doi.org/10.1007/s00440-021-01104-x>
- [16] CAPUTO, P., LABBÉ, C. and LACOIN, H. (2020). Mixing time of the adjacent walk on the simplex. *Ann. Probab.* **48** 2449–2493. [MR4152648](#) <https://doi.org/10.1214/20-AOP1428>
- [17] CHATTERJEE, S., DIACONIS, P., SLY, A. and ZHANG, L. (2022). A phase transition for repeated averages. *Ann. Probab.* **50** 1–17. [MR4385355](#) <https://doi.org/10.1214/21-AOP1526>
- [18] DIACONIS, P. (1996). The cutoff phenomenon in finite Markov chains. *Proc. Natl. Acad. Sci. USA* **93** 1659–1664. [MR1374011](#) <https://doi.org/10.1073/pnas.93.4.1659>
- [19] DIACONIS, P. and RAM, A. (2000). Analysis of systematic scan Metropolis algorithms using Iwahori–Hecke algebra techniques. *Michigan Math. J.* **48** 157–190. [MR1786485](#) <https://doi.org/10.1307/mmjj-1030132713>
- [20] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **57** 159–179. [MR0626813](#) <https://doi.org/10.1007/BF00535487>
- [21] DING, J., LUBETZKY, E. and PERES, Y. (2010). Total variation cutoff in birth-and-death chains. *Probab. Theory Related Fields* **146** 61–85. [MR2550359](#) <https://doi.org/10.1007/s00440-008-0185-3>
- [22] GALASHIN, P. (2021). Symmetries of stochastic colored vertex models. *Ann. Probab.* **49** 2175–2219. [MR4317703](#) <https://doi.org/10.1214/20-aop1502>
- [23] GANTERT, N., NESTORIDI, E. and SCHMID, D. (2023). Mixing times for the simple exclusion process with open boundaries. *Ann. Appl. Probab.* **33** 1172–1212. [MR4564424](#) <https://doi.org/10.1214/22-aap1839>
- [24] GONÇALVES, P., JARA, M., MARINHO, R. and MENEZES, O. (2021). Sharp convergence to equilibrium for the SSEP with reservoirs. Available at [arXiv:2110.06353](#).
- [25] GREENBERG, S., PASCOE, A. and RANDALL, D. (2009). Sampling biased lattice configurations using exponential metrics. In *Proceedings of the Twentieth Annual ACM–SIAM Symposium on Discrete Algorithms* 76–85. SIAM, Philadelphia, PA. [MR2809307](#)
- [26] JOHANSSON, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** 437–476. [MR1737991](#) <https://doi.org/10.1007/s002200050027>

- [27] LABBÉ, C. and LACOIN, H. (2019). Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *Ann. Probab.* **47** 1541–1586. MR3945753 <https://doi.org/10.1214/18-AOP1290>
- [28] LABBÉ, C. and PETIT, E. (2022). Hydrodynamic limit and cutoff for the biased adjacent walk on the simplex. Available at [arXiv:2203.04008](https://arxiv.org/abs/2203.04008).
- [29] LACOIN, H. (2021). Mixing time and cutoff for one dimensional particle systems. In *Proceedings of the ICM 2022*. To appear. Available at [arXiv:2111.06436](https://arxiv.org/abs/2111.06436).
- [30] LACOIN, H. (2016). The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.* **44** 3399–3430. MR3551201 <https://doi.org/10.1214/15-AOP1053>
- [31] LACOIN, H. (2016). Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.* **44** 1426–1487. MR3474475 <https://doi.org/10.1214/15-AOP1004>
- [32] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times: Second Edition*. Amer. Math. Soc., Providence, RI. MR3726904 <https://doi.org/10.1090/mkbk/107>
- [33] LIGGETT, T. M. (1985). *Interacting Particle Systems. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **276**. Springer, New York. MR0776231 <https://doi.org/10.1007/978-1-4613-8542-4>
- [34] LUBETZKY, E. and PERES, Y. (2016). Cutoff on all Ramanujan graphs. *Geom. Funct. Anal.* **26** 1190–1216. MR3558308 <https://doi.org/10.1007/s00039-016-0382-7>
- [35] LUBETZKY, E. and SLY, A. (2010). Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.* **153** 475–510. MR2667423 <https://doi.org/10.1215/00127094-2010-029>
- [36] LUBETZKY, E. and SLY, A. (2013). Cutoff for the Ising model on the lattice. *Invent. Math.* **191** 719–755. MR3020173 <https://doi.org/10.1007/s00222-012-0404-5>
- [37] MATETSKI, K., QUASTEL, J. and REMENIK, D. (2021). The KPZ fixed point. *Acta Math.* **227** 115–203. MR4346267 <https://doi.org/10.4310/acta.2021.v227.n1.a3>
- [38] NESTORIDI, E. and OLESKER-TAYLOR, S. (2022). Limit profiles for reversible Markov chains. *Probab. Theory Related Fields* **182** 157–188. MR4367947 <https://doi.org/10.1007/s00440-021-01061-5>
- [39] ORTMANN, J., QUASTEL, J. and REMENIK, D. (2017). A Pfaffian representation for flat ASEP. *Comm. Pure Appl. Math.* **70** 3–89. MR3581823 <https://doi.org/10.1002/cpa.21644>
- [40] QUASTEL, J. and REMENIK, D. (2014). Airy processes and variational problems. In *Topics in Percolative and Disordered Systems. Springer Proc. Math. Stat.* **69** 121–171. Springer, New York. MR3229288 [https://doi.org/10.1007/978-1-4939-0339-9\\_5](https://doi.org/10.1007/978-1-4939-0339-9_5)
- [41] QUASTEL, J. and SARKAR, S. (2023). Convergence of exclusion processes and the KPZ equation to the KPZ fixed point. *J. Amer. Math. Soc.* **36** 251–289. MR4495842 <https://doi.org/10.1090/jams/999>
- [42] SALEZ, J. (2023). Cutoff for non-negatively curved Markov chains. *J. Eur. Math. Soc.* Published online first. <https://doi.org/10.4171/JEMS/1348>
- [43] SALEZ, J. (2023). Universality of cutoff for exclusion with reservoirs. *Ann. Probab.* **51** 478–494. MR4546624 <https://doi.org/10.1214/22-aop1600>
- [44] SALOFF-COSTE, L. (2004). Random walks on finite groups. In *Probability on Discrete Structures. Encyclopaedia Math. Sci.* **110** 263–346. Springer, Berlin. MR2023654 [https://doi.org/10.1007/978-3-662-09444-0\\_5](https://doi.org/10.1007/978-3-662-09444-0_5)
- [45] TEYSSIER, L. (2020). Limit profile for random transpositions. *Ann. Probab.* **48** 2323–2343. MR4152644 <https://doi.org/10.1214/20-AOP1424>
- [46] TRACY, C. A. and WIDOM, H. (1994). Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159** 151–174. MR1257246
- [47] TRACY, C. A. and WIDOM, H. (1996). On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177** 727–754. MR1385083
- [48] TRACY, C. A. and WIDOM, H. (2009). Asymptotics in ASEP with step initial condition. *Comm. Math. Phys.* **290** 129–154. MR2520510 <https://doi.org/10.1007/s00220-009-0761-0>
- [49] WILSON, D. B. (2004). Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.* **14** 274–325. MR2023023 <https://doi.org/10.1214/aoap/1075828054>
- [50] ZHANG, L. (2023). Shift-invariance of the colored TASEP and finishing times of the oriented swap process. *Adv. Math.* **415** Paper No. 108884, 60. MR4541342 <https://doi.org/10.1016/j.aim.2023.108884>