The Annals of Applied Probability
2020, Vol. 30, No. 4, 1847–1883
https://doi.org/10.1214/19-AAP1545
© Institute of Mathematical Statistics, 2020

MIXING TIME AND CUTOFF FOR THE WEAKLY ASYMMETRIC SIMPLE EXCLUSION PROCESS

BY CYRIL LABBÉ[1] AND HUBERT LACOIN[2]

[1] *Université Paris-Dauphine, PSL Research University, labbe@ceremade.dauphine.fr*
[2] *IMPA, lacoin@impa.br*

We consider the simple exclusion process with k particles on a segment of length N performing random walks with transition $p > 1/2$ to the right and $q = 1 - p$ to the left. We focus on the case where the asymmetry in the jump rates $b = p - q > 0$ vanishes in the limit when N and k tend to infinity, and obtain sharp asymptotics for the mixing times of this sequence of Markov chains in the two cases where the asymmetry is either much larger or much smaller than $(\log k)/N$. We show that in the former case ($b \gg (\log k)/N$), the mixing time corresponds to the time needed to reach macroscopic equilibrium, like for the strongly asymmetric (i.e., constant b) case studied in (*Ann. Probab.* **47** (2019) 1541–1586), while the latter case ($b \ll (\log k)/N$) macroscopic equilibrium is not sufficient for mixing and one must wait till local fluctuations equilibrate, similarly to what happens in the symmetric case worked out in (*Ann. Probab.* **44** (2016) 1426–1487). In both cases, convergence to equilibrium is abrupt: we have a cutoff phenomenon for the total-variation distance. We present a conjecture for the remaining regime when the asymmetry is of order $(\log k)/N$.

1. Introduction. The simple exclusion process is a model of statistical mechanics that provides a simplified picture for a gas of interacting particles. Particles move on a lattice, each of them performing a nearest neighbor random walk independently of the others, and interact only via the exclusion rule that prevents any two particles from sharing the same site (when a particle tries to jump on a site which is already occupied, this jump is cancelled). In spite of its simplicity, this model displays a very rich behavior and has given rise to an abundant literature both in theoretical physics and mathematics; see, for instance [8, 16] and references therein.

In the present paper, we study relaxation to equilibrium for a particular instance of the simple exclusion process in which the lattice is a segment of length N and particles feel a bias towards the right that vanishes when N tends to infinity. This setup is often referred to as the Weakly Asymmetric Simple Exclusion Process (WASEP): it interpolates between the symmetric case (SSEP) and the asymmetric case (ASEP) where the bias is a positive constant.

While convergence to equilibrium for a particle system can be considered on a macroscopic scale via the evolution of the particle density or hydrodynamic profile (see, e.g., [8] and references therein), an alternative and complementary viewpoint (when the system is of finite size) consists in measuring the so-called ε-Total Variation Mixing Time [15]. It is defined as the first time at which the total variation distance to the stationary measure, starting from the "worst" initial condition, falls below a given threshold ε. Compared to the hydrodynamic profile, this provides a much more microscopic information on the particle system.

The problem of mixing time of the simple exclusion process on the segment has been extensively studied both in the symmetric [12, 13, 19] and the asymmetric setup [2, 11] and it has been proved in [13] and [11] respectively that in both cases, the total variation distance

Received April 2019; revised September 2019.
MSC2020 subject classifications. Primary 60J27; secondary 37A25, 82C22.
Key words and phrases. Exclusion process, WASEP, mixing time, cutoff.

drops abruptly from its maximal value 1 to 0, so that the mixing time does not depend at first order on the choice of the threshold ε. A phenomenon known as cutoff and conjectured to hold for a large class of Markov chains as soon as the mixing time is of a larger order than the relaxation time (which is defined as the inverse of the spectral gap of the generator). To be more precise, the mixing time being of a larger order than the relaxation time is a necessary condition, which is not sufficient (there are known counter examples, reversible and nonreversible), but it is believed that only a few additional assumptions are necessary. What these additional assumptions should be is one of the major open question in the field (see [15], Chapter 18).

However the patterns of convergence to equilibrium in the symmetric and asymmetric cases are very different. Let us for simplicity focus on the case with a density of particles $k = \alpha N$, $\alpha \in (0, 1)$. In the symmetric case, the time scale associated with the hydrodynamic profile is N^2 and the limit is given by the heat equation [9] (which takes an infinite time to relax to its equilibrium profile which is flat) and microscopic mixing occurs on a larger time scale $N^2 \log N$.

In the asymmetric setup the hydrodynamic limit is given by the inviscid Burgers' equation with a shorter time scale N [18] (see also [10, 11] for adaptations of this result to the segment). The equilibrium profile for this equation is reached after a finite time and in this case, the mixing time is of order N and corresponds exactly to the time at which macroscopic equilibrium is attained.

The aim of this paper is to understand better the role of the asymmetry in mixing and how one interpolates between the symmetric and asymmetric regimes. This leads us to consider the so-called Weakly Asymmetric Exclusion Process (WASEP) in which the asymmetry vanishes with the scale of observation. While hydrodynamic limit [4, 7, 8] and fluctuations scaling limits [3, 6, 10] for WASEP are now well understood, much less is known about how a weak asymmetry affects the mixing time of the system.

A first step in this direction was made in [14]. Therein the order of magnitude for the mixing time was identified for all possible intensities of vanishing bias, but with different constants for the upper and the lower bounds. Three regimes were distinguished (under the assumption that there is a density of particles):

(A) When $b_N \le N^{-1}$, the mixing time remains of the same order as that of the symmetric case $N^2 \log N$.
(B) When $N^{-1} \le b_N \le N^{-1} \log N$, the mixing time is of order $(b_N)^{-2} \log N$.
(C) When $N^{-1} \log N \le b_N \le 1$, the mixing time is of order $(b_N)^{-1} N$.

The transition occurring around $b_N \approx N^{-1}$ is the one observed for the hydrodynamic limit: It corresponds to a crossover regime where the limit is given by a viscous Burger's equation [4, 7, 8] which interpolates between the heat and the inviscid Burgers' equations. The one occurring for $b_N \approx N^{-1} \log N$ is however not observed in the macroscopic profile as it corresponds to the crossover regime for the position of the leftmost particle at equilibrium: for weaker asymmetries, the leftmost particle lies at a negligible distance (compared to N) from site 1 and for stronger asymmetries, it lies at a negligible distance (compared to N) from site $N - k + 1$ (which is the rightmost site at which this particle can lie).

In the present work, we identify the complete asymptotics of the mixing time (with the right constant) when the bias is either negligible compared to, or much larger than $N^{-1} \log N$ (or $N^{-1} \log k$ when the total number of particle is not of order N). This implies cutoff in these two regimes. Our result and its proof provide a better understanding of the effect of asymmetry on microscopic mixing: When $b_N \gg N^{-1} \log N$, the pattern of relaxation is identical to that of the fully asymmetric case and microscopic equilibrium is reached exactly when the macroscopic profile hits its equilibrium state. When $b_N \ll N^{-1} \log N$ the pattern of relaxation resembles that of the symmetric case, the mixing time corresponds to the time needed

to equilibrate local fluctuations, in particular in the case (B) described above (or more precisely when $N^{-1} \ll b_N \ll N^{-1}(\log N)$) this time does not correspond to the time needed to reach macroscopic equilibrium.

The crossover regime $b_N \approx N^{-1} \log N$ remains an open problem: In this case the time to reach macroscopic equilibrium and that to equilibrate local fluctuations are of the same order and the two phenomena are difficult to separate. In Section 2.3 we provide a conjecture for the mixing time in this regime in the case of vanishing density.

In the next section, we introduce precisely the model, state our main results and present the outline of the article.

2. Model and results.

2.1. *Mixing time for the WASEP.* Given $N \in \mathbb{N}$, $k \in [\![1, N-1]\!]$ (we use the notation $[\![a, b]\!] = [a, b] \cap \mathbb{Z}$) and $p \in (1/2, 1]$, the Asymmetric Simple Exclusion Process on $[\![1, N]\!]$ with k particles and parameter p is the random process on the state space

$$\Omega^0_{N,k} := \left\{ \xi \in \{0,1\}^N : \sum_{x=1}^{N} \xi(x) = k \right\},$$

associated with the generator

$$\mathcal{L}_{N,k} f(\xi) := \sum_{y=1}^{N-1} \left(q \mathbf{1}_{\{\xi(y) < \xi(y+1)\}} + p \mathbf{1}_{\{\xi(y) > \xi(y+1)\}}\right)\left(f(\xi^y) - f(\xi)\right), \tag{1}$$

where $q = 1 - p$ and

$$\xi^y(x) := \begin{cases} \xi(y+1) & \text{if } x = y, \\ \xi(y) & \text{if } x = y+1, \\ \xi(x) & \text{if } x \notin \{y, y+1\}. \end{cases} \tag{2}$$

In a more intuitive manner we can materialize the positions of 1 by particles, and say that the particles perform random walks with jump rates p to the right and $q = 1 - p$ to the left: These random walks are independent from one another except that any jump that would put a particle at a location already occupied by another particle is cancelled. Having in mind this particle representation, we let for $i \in [\![1, k]\!]$, ξ_i denote the position of the ith leftmost particle

$$\xi_i := \min\left\{ y \in [\![1, N]\!] : \sum_{x=1}^{y} \xi(x) = i \right\}. \tag{3}$$

We let $P_t^{N,k}$ denote the associated semi-group and $(\eta^\xi(t, \cdot))_{t \geq 0}$ denote the trajectory of the Markov chain starting from initial configuration $\xi \in \Omega^0_{N,k}$. This Markov chain is irreducible, and admits a unique invariant (and reversible) probability measure $\pi_{N,k}$ given by

$$\pi_{N,k}(\xi) := \frac{1}{Z_{N,k}} \lambda^{-A(\xi)}, \tag{4}$$

where $\lambda = p/q$, $Z_{N,k} := \sum_{\xi \in \Omega^0_{N,k}} \lambda^{-A(\xi)}$, and

$$A(\xi) := \sum_{i=1}^{k} (N - k + i - \xi_i) \geq 0 \tag{5}$$

denotes the minimal number of moves that are necessary to go from ξ to the configuration $\xi^{\min}$ where all the particles are on the right $\xi^{\min}(x) := \mathbf{1}_{[N-k+1,N]}(x)$ (this terminology is justified by the fact that $\xi^{\min}$ is minimal for the order introduced in Section 3).

Recall that the total-variation distance between two probability measures defined on the same space Ω is defined by

$$\|\alpha - \beta\|_{\mathrm{TV}} = \sup_{A \subset \Omega} \alpha(A) - \beta(A),$$

where the sup is taken over all measurable sets A.

The mixing time associated to the threshold $\varepsilon \in (0, 1)$ is defined by

$$T_{\mathrm{mix}}^{N,k}(\varepsilon) := \inf\{t \geq 0 : d^{N,k}(t) \leq \varepsilon\}, \tag{6}$$

where $d^{N,k}(t)$ denotes the total-variation distance to equilibrium at time t starting from the worst possible initial condition

$$d^{N,k}(t) := \max_{\xi \in \Omega_{N,k}^0} \|P_t^{N,k}(\xi, \cdot) - \pi_{N,k}\|_{\mathrm{TV}}. \tag{7}$$

We want to study the asymptotic behavior of the mixing time for this system when both the size of the system and the number of particles tend to infinity. A natural case to consider is when there is a nontrivial density of particles, that is $k/N \to \alpha \in (0, 1)$, but we decide to also treat the boundary cases of vanishing density ($\alpha = 0$) and full density ($\alpha = 1$). By symmetry we can restrict to the case when $k = k_N \leq N/2$: indeed, permuting the roles played by particles and empty sites boils down to reversing the direction of the asymmetry of the jump rates. Note that we will always impose $k \geq 1$ since when $k = 0$ the process is trivial (we do not assume however in general that k_N tends to infinity, this is only a requirement in the small bias regime, cf. (12)).

The asymptotic behavior of $T_{\mathrm{mix}}^{N,k}(\varepsilon)$ in the case of constant bias ($p > 1/2$ is fixed when N goes to infinity) has been obtained in a previous work.

THEOREM A (Theorem 2 in [11]). *We have for every $\varepsilon > 0$, every $\alpha \in [0, 1]$ and every sequence k_N such that $k_N/N \to \alpha$*

$$\lim_{N \to \infty} \frac{T_{\mathrm{mix}}^{N,k_N}(\varepsilon)}{N} = \frac{(\sqrt{\alpha} + \sqrt{1-\alpha})^2}{p - q}. \tag{8}$$

The result implies in particular that at first order, the mixing time does not depend on $\varepsilon \in (0, 1)$, meaning that on the appropriate time-scale, for large values of N the distance to equilibrium drops abruptly from 1 to 0. This phenomenon is referred to as *cutoff* and was first observed in the context of card shuffling [1, 5]. It is known to occur for a large variety of Markov chains; see for instance [15]. In the context of the exclusion process, it has been proved in [13] that cutoff holds for the Symmetric Simple Exclusion Process (SSEP) which is obtained by setting $p = 1/2$ in the generator (1).

THEOREM B (Theorem 2.4 in [13]). *When $p = 1/2$, for any sequence k_N that tends to infinity and satisfies $k_N \leq N/2$ for all N, we have*

$$\lim_{N \to \infty} \frac{T_{\mathrm{mix}}^{N,k_N}(\varepsilon)}{N^2 \log k_N} = \frac{1}{\pi^2}. \tag{9}$$

While cutoff occurs in the two cases, it appears to be triggered by different mechanisms. When $p > 1/2$, the mixing time is determined by the time needed for the particle density profile to reach its macroscopic equilibrium: After rescaling time and space by N, the evolution of the particle density has a nontrivial scaling limit (the inviscid Burgers' equation with zero-flux boundary conditions), which fixates at time $\frac{(\sqrt{\alpha}+\sqrt{1-\alpha})^2}{p-q}$. The first order asymptotic for the mixing time is thus determined by the time the density profile needs to reach equilibrium.

When $p = 1/2$, the right-time scale to observe a macroscopic motion for the particles is N^2, and it is worth mentioning that the scaling limit obtained for the particle density (the heat equation) does not fixate in finite time. To reach equilibrium, however, we must wait for a longer time, of order $N^2 \log k_N$, which is the time needed for local fluctuations in the particle density to come to equilibrium.

We are interested in studying the process when the drift tends to zero: this requires to understand the transition between these two patterns of relaxation to equilibrium. Hence we consider p to be a function of N which is such that the bias towards the right $b_N := p_N - q_N = 2p_N - 1$ vanishes, that is,

$$\lim_{N\to\infty} b_N = 0. \tag{10}$$

In this regime, the model is sometimes called WASEP for Weakly Asymetric Simple Exclusion Process. Its convergence to equilibrium has already been studied in [10, 14]. In [14] the authors identify the order of magnitude of the mixing time as a function of b_N in full generality. However the approach used in [14] does not allow to find the exact asymptotic for the mixing time nor to prove cutoff, and does not answer our question concerning the pattern of relaxation to equilibrium.

2.2. *Results*. We identify two main regimes for the pattern of relaxation to equilibrium. The *large bias* regime where

$$\lim_{N\to\infty} \frac{Nb_N}{(\log k_N) \vee 1} = \infty \tag{11}$$

and the *small bias* regime where

$$\begin{cases} \displaystyle\lim_{N\to\infty} \frac{Nb_N}{\log k_N} = 0, \\ \displaystyle\lim_{N\to\infty} k_N = \infty. \end{cases} \tag{12}$$

We identify the asymptotic expression for the mixing time in both regimes. In the large bias regime we show that the mixing time coincides with the time needed by the particle density to reach equilibrium like in the constant bias case.

THEOREM 1. *When* (11) *holds*, *and* $\lim_{N\to\infty} k_N/N = \alpha \in [0, 1]$, *we have for every* $\varepsilon \in (0, 1)$

$$\lim_{N\to\infty} \frac{b_N T_{\mathrm{mix}}^{N,k_N}(\varepsilon)}{N} = (\sqrt{\alpha} + \sqrt{1-\alpha})^2. \tag{13}$$

To state our result in the small bias regime, let us introduce the quantity

$$\mathrm{gap}_N := (\sqrt{p_N} - \sqrt{q_N})^2 + 4\sqrt{p_N q_N} \sin\left(\frac{\pi}{2N}\right)^2, \tag{14}$$

which corresponds to the spectral gap associated with the generator (1) (see, e.g., [11] for a proof). Notice that it does not depend on the number k of particles in the system. The pattern of relaxation is similar to the one observed in the symmetric case.

THEOREM 2. *When* (12) *holds*, *we have*

$$\lim_{N\to\infty} \frac{\mathrm{gap}_N T_{\mathrm{mix}}^{N,k_N}(\varepsilon)}{\log k_N} = \frac{1}{2}. \tag{15}$$

Using Taylor expansion for gap_N we have, whenever b_N tends to zero

$$\mathrm{gap}_N \overset{N\to\infty}{\sim} \frac{1}{2}\left(b_N^2 + \left(\frac{\pi}{N}\right)^2\right). \tag{16}$$

Thus in particular we have

$$T_{\mathrm{mix}}^{N,k_N}(\varepsilon) \sim \begin{cases} \dfrac{\log k_N}{b_N^2} & \text{if } 1/N \ll b_N \ll \log k_N/N, \\ \dfrac{1}{\pi^2} N^2 \log k_N & \text{if } 0 < b_N \ll 1/N, \\ \dfrac{1}{\pi^2+\beta^2} N^2 \log k_N & \text{if } b_N \sim \beta/N. \end{cases} \tag{17}$$

Note that our classification of regimes (11)–(12) does not cover all possible choices of b_N. Two cases have been excluded for very different reasons:

- When $b_N = O(N^{-1})$ and k_N is bounded, then we have a system of k diffusive interacting random walks. This system does not exhibit cutoff and has a mixing time of order N^2. (The upper bound can actually be deduced from an argument presented in Section 5.2.2 and the lower bound is achieved for example, by looking at the expectation and variance of the number of particles on the right half of the segment, like what is done in [17], Section 6.)
- When b_N is of order $\log k_N/N$ the time at which the density profile reaches its equilibrium and the time needed for local fluctuations to reach their equilibrium values are of the same order and we believe that there is an interaction between the two phenomena. We provide a more detailed conjecture in Section 2.3.

REMARK 1. We have not included here results concerning the biased card shuffling considered in [2, 11]. Let us mention that while our analysis should also yield optimal bounds for the mixing time of this process when (11) holds (i.e., $T_{\mathrm{mix}}^{k,N}(\varepsilon) \sim 2N/(p_N - q_N)$), it seems much more difficult to prove the equivalent of Theorem 2. The main reason is that the coupling presented in Section A cannot be extended to a coupling on the permutation process. Building on and adapting the techniques presented in [13], Section 5, it should a priori be possible to obtain a result concerning the mixing time starting from an extremal condition (the identity or its symmetric), but this is out of the scope of the present paper.

2.3. *Conjecture in the regime $b_N \asymp \log k_N/N$.* Let us here formulate, and heuristically support a conjecture concerning the mixing-time in the crossover regime where

$$\lim_{N\to\infty} \frac{b_N N}{\log k_N} = \beta \tag{18}$$

for some $\beta \in (0,\infty)$. For the ease of exposition, while it should be in principle possible to extend the heuristic to the case of positive density (see Remark 3 below) we restrict ourselves to the case $\lim_{N\to\infty} k_N/N = 0$. The justification we provide for the conjecture might be better understood after a first reading of the entire paper.

CONJECTURE 2. *When b_N and k_N display the asymptotic behavior given by* (18), *we have for every $\varepsilon > 0$*

$$\lim_{N\to\infty} \frac{T_{\mathrm{mix}}^{N,k_N}(\varepsilon)\log k_N}{N^2} = \begin{cases} \dfrac{2}{\beta} + \dfrac{1}{\beta^2} & \textit{if } \beta \le 1/2, \\ \left(\dfrac{\sqrt{2}+2\sqrt{\beta}}{2\beta}\right)^2 & \textit{if } \beta \ge 1/2. \end{cases} \tag{19}$$

To motivate this conjecture let us first describe the equilibrium measure and its dependence on β. As we are in the low-density regime, the equilibrium measure is quite close to the product measure one would obtain for the system without exclusion rules: the location of the k particles are approximately IID distributed, the distance from the right extremity being (also approximately) a geometric variable of parameter $q_N/p_N \approx e^{-\frac{2\beta \log k}{N}}$.

Hence the probability of having a particle at site $\lfloor zN \rfloor$ for $z \in [0, 1]$ is roughly of order $k^{[1-2\beta(1-z)]+o(1)}/N$. Thus, under the equilibrium measure, while particles are concentrated near the right extremity, the "logarithmic density" of particles exhibit a nontrivial profile in the sense that for any $z > 1 - (2\beta)^{-1}$ we have

$$\lim_{\varepsilon \to 0} \lim_{N \to \infty} \frac{\log(\sum_{i=(z-\varepsilon)N}^{(z+\varepsilon)N} \xi(i))}{\log k_N} = 1 - 2\beta(1-z), \tag{20}$$

where the convergence holds in distribution under the equilibrium measure $\pi_{N,k}$ when N tends to infinity and ε tends to zero in that order. The typical distance to zero of the left-most particle at equilibrium is also given by this profile in the sense that it is typically $o(N)$ when $\beta \le 1/2$ and of order $N(1 - \frac{1}{2\beta})$ when $\beta \ge 1/2$. While we only give heuristic justification for these statements concerning equilibrium, it is worth mentioning that they can be made rigorous by using the techniques exposed in Section 3.

To estimate the mixing time, we assume that the system gets close to equilibrium once the number of particles on any "mesoscopic" interval of the form $[(z-\varepsilon)N, (z+\varepsilon)N]$ is close to its equilibrium value. While the mean number of particle is of order $k^{1-2\beta(1-z)}$ (cf. (20)), the typical equilibrium fluctuation around this number should be given by the square root due to near-independence of different particles and thus be equal to $k^{\frac{1}{2}-\beta(1-z)}$.

To estimate the surplus of particles in this interval at time $tN^2(\log k)^{-1}$ for $t > 1/\beta$ (note that $t = 1/\beta$ is the time of macroscopic equilibrium where most particles are packed on the right), we consider the number of particles that end up there after keeping a constant drift of order $z(\log k)/(Nt)$, which is smaller than b_N. Neglecting interaction between particles and making a Brownian approximation for the random walk with drift, we obtain that the expected number of particles following this strategy is given by

$$k \exp\left(-\log k \frac{(\beta t - z)^2}{2t}\right) = k^{1-\frac{(\beta t - z)^2}{2t}}.$$

Hence equilibrium should be attained when this becomes negligible with respect to the typical fluctuation $k^{\frac{1}{2}-\beta(1-z)}$ for all values of z where we find particles at equilibrium. That is, when the inequality

$$1 - \frac{(\beta t - z)^2}{2t} < \frac{1}{2} - \beta(1-z), \tag{21}$$

is valid for all $z \in [0, 1]$ if $\beta \le 1/2$ or for all $z \ge 1 - \frac{1}{2\beta}$ if $\beta \ge 1/2$. A rapid computation show that one only needs to satisfy the condition for the smallest value of z (either 1 or $1 - \frac{1}{2\beta}$), which boils down to finding the roots of a degree two polynomial. This yields that we must have $t > t_0$ where

$$t_0 := \begin{cases} \dfrac{2}{\beta} + \dfrac{1}{\beta^2} & \text{if } \beta \le 1/2, \\ \left(\dfrac{\sqrt{2} + 2\sqrt{\beta}}{2\beta}\right)^2 & \text{if } \beta \ge 1/2. \end{cases} \tag{22}$$

REMARK 3. Describing the equilibrium "logarithmic profile" of particles when the system has positive density is also possible (note that on the right of $(1-\alpha)N$ it is the density of empty-sites that becomes the quantity of interest). It is thus reasonable to extend the heuristic to that case. However the best strategy to produce a surplus of particles in that case becomes more involved, as the zones with positive density of particles, which are described by the hydrodynamic evolution given in Proposition 6, play a role in the optimization procedure. For this reason we did not wish to bring the speculation one step further.

2.4. *Organization of the paper.* In the remainder of the paper we drop the subscript N in k_N in order to simplify the notation. The article is organised as follows. In Section 3, we introduce the representation through height functions and collect a few results on the invariant measure, the spectral gap and the hydrodynamic limit of the process. In Sections 4 and 5, we consider the large bias case and prove respectively the lower and upper bounds of Theorem 1: While the lower bound essentially follows from the hydrodynamic limit, the upper bound is more involved and is one of the main achievements of this paper. In Sections 6 and 7, we deal with the small bias case and prove respectively the lower and upper bounds of Theorem 2. Here again, the lower bound is relatively short and follows from similar argument as those presented by Wilson [19] in the symmetric case, while the upper bound relies on a careful analysis of the area between the processes starting from the maximal and minimal configurations and under some grand coupling.

3. Preliminaries and technical estimates.

3.1. *Height function ordering and grand coupling.* To any configuration of particles $\xi \in \Omega^0_{N,k}$, we can associate a so-called height function $h = h(\xi)$ defined by (see Figure 1) $h(\xi)(0) = 0$ and

$$h(\xi)(x) = \sum_{y=1}^{x} \big(2\xi(y) - 1\big), \quad x \in [\![1, N]\!].$$

For simplicity, we often abbreviate this in $h(x)$. The height function is a lattice path that increases by 1 from $\ell - 1$ to ℓ if there is a particle at site ℓ, and decreases by 1 otherwise. Its terminal value therefore only depends on k and N. The set of height functions obtained from $\Omega^0_{N,k}$ through the above map is denoted $\Omega_{N,k}$.

The particles dynamics can easily be rephrased in terms of height functions: every upward corner ($h(x) = h(x-1)+1 = h(x+1)+1$) flips into a downward corner ($h(x) = h(x-1) -$

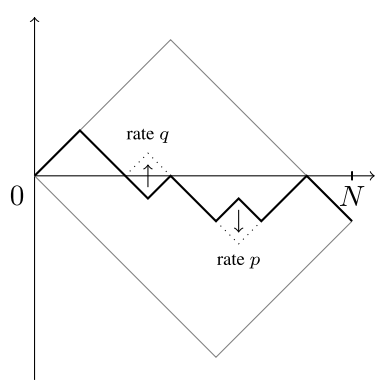

FIG. 1. *An example of height function with $k = 6$ particles over $N = 14$ sites. The interface lives within the grey rectangle.*

$1 = h(x+1) - 1)$ at rate p, while the opposite occurs at rate q. We denote by $(h^\zeta(t,\cdot), t \geq 0)$ the associated Markov process starting from some initial configuration $\zeta \in \Omega_{N,k}$.

It will be convenient to denote by $\wedge$ the maximal height function:

$$\wedge(x) = x \wedge (2k - x), \quad x \in [\![1, N]\!],$$

and by $\vee$ the minimal height function:

$$\vee(x) = (-x) \vee (x - 2N + 2k), \quad x \in [\![1, N]\!].$$

Though the dependence on k is implicit in the notations $\wedge$, $\vee$, this will never raise any confusion as the value k will be clear from the context.

It is possible to construct simultaneously on a same probability space and in a Markovian fashion, the height function processes $(h_t^\zeta, t \geq 0)$ starting from all initial conditions $\zeta \in \bigcup_k \Omega_{N,k}$ and such that the following monotonicity property is satisfied for all k and all $\zeta, \zeta' \in \Omega_{N,k}$:

$$\zeta \leq \zeta' \quad \Rightarrow \quad h_t^\zeta \leq h_t^{\zeta'}, \quad \forall t \geq 0. \tag{23}$$

Here, $\zeta \leq \zeta'$ simply means $\zeta(x) \leq \zeta'(x)$ for all $x \in [\![0, N]\!]$. We call such a construction a monotone Markovian grand coupling, and we denote by $\mathbb{P}$ the corresponding probability distribution. The existence of such a grand coupling is classical; see, for instance, [11], Proposition 4. In a portion of our proof, we require to use a specific grand coupling which is not the one displayed in [11] and for this reason we provide an explicit construction in Appendix A.

Once a coupling is specified, by enlarging our probability space, one can also define the process h_t^π which is started from an initial condition sampled from the equilibrium measure $\pi_{N,k}$, independently of h_t^ζ, $\zeta \in \Omega_{N,k}$.

Let us end up this subsection introducing the (less canonical) notation

$$\zeta < \zeta' \quad \Leftrightarrow \quad (\zeta \leq \zeta' \text{ and } \zeta \neq \zeta'). \tag{24}$$

We say that a function f on $\Omega_{N,k}$ is increasing (strictly) if $f(\zeta) \leq f(\zeta')$ ($f(\zeta) < f(\zeta')$) whenever $\zeta < \zeta'$. The minimal increment of an increasing function is defined by

$$\delta_{\min}(f) = \min_{\zeta, \zeta' \in \Omega_{N,k}, \zeta < \zeta'} f(\zeta') - f(\zeta). \tag{25}$$

3.2. *The equilibrium measure in the large-bias case.* For $\xi \in \Omega_{N,k}^0$ we set

$$\begin{aligned} \ell_N(\xi) &= \min\{x \in [\![1, N]\!] : \xi(x) = 1\}, \\ r_N(\xi) &= \max\{x \in [\![1, N]\!] : \xi(x) = 0\}. \end{aligned} \tag{26}$$

A useful observation on the invariant measure is the following. Given ξ, we define $\chi(\xi)$ as the sequence of particle spacings (recall (3)):

$$\chi_i := \xi_{i+1} - \xi_i, \quad \text{for } i \in [\![1, k-1]\!], \qquad \chi_k = N + 1 - \xi_k.$$

From (5), under $\pi_{N,k}$ the probability of a given configuration is proportional to

$$\lambda^{-\chi_1} \lambda^{-2\chi_2} \dots \lambda^{-k\chi_k}. \tag{27}$$

In other terms, under the invariant measure the particle spacings $(\chi_i)_{1 \leq i \leq k}$ are distributed like independent geometric variables, with respective parameters $1 - \lambda^{-i}$, conditioned to the event $\sum_{i=1}^k \chi_k \leq N$.

LEMMA 4. *When* (11) *holds we have for any* $\varepsilon > 0$

$$\begin{aligned} &\lim_{N\to\infty} \pi_{N,k}\big(\ell_N \le (N-k) - \varepsilon N\big) = 0, \\ &\lim_{N\to\infty} \pi_{N,k}\big(r_N \ge (N-k) + \varepsilon N\big) = 0. \end{aligned} \tag{28}$$

PROOF. By symmetry it is sufficient to prove the result for ℓ_N only, but for all $k \in [\![1, N-1]\!]$. Note that there is nothing to prove regarding ℓ_N if $\alpha := \lim_{N\to\infty} k/N = 1$, so we assume that $\alpha \in [0, 1)$.

Let $(X_i)_{1\le i\le k}$ be independent geometric variables, with respective parameters λ^{-i}. The sum of their means satisfies (recall that $\lambda - 1$ is of order b_N)

$$\sum_{i=1}^{k} \frac{1}{1-\lambda^{-i}} = k + \sum_{i=1}^{k} \frac{\lambda^{-i}}{1-\lambda^{-i}} \le k + Cb_N^{-1} \log \min(k, b_N^{-1}),$$

for some constant $C > 0$. The large bias assumption ensures that $b_N^{-1} \log \min(k_N, b_N^{-1}) = o(N)$. Hence using the Markov inequality, we obtain that if $(X_i)_{1\le i\le k}$ is a sequence of such geometric variables, and if (11) is satisfied, then for any $\varepsilon > 0$

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i \ge k + \varepsilon N\right) \le \frac{\mathbb{E}[|\sum_{i=1}^{k} X_i - k|]}{\varepsilon N} = \frac{\mathbb{E}[\sum_{i=1}^{k} X_i - k]}{\varepsilon N},$$

so that

$$\lim_{N\to\infty} \mathbb{P}\left(\sum_{i=1}^{k} X_i \ge k + \varepsilon N\right) = 0.$$

The above inequality for $\varepsilon < 1 - \alpha$ implies that $\mathbb{P}(\sum_{i=1}^{k} X_i \le N) \ge 1/2$ for all N large enough meaning that the conditioning only changes the probability by a factor at most 2. Then, we can conclude by noticing that $\ell_N = \xi_1 = N + 1 - \sum_{i=1}^{k} \chi_i$. $\square$

3.3. *The equilibrium measure in the small-bias case.* We aim at showing that with large probability the density of particles everywhere is of order $k^{1+o(1)}/N$. Given $\xi \in \Omega_{N,k}^0$ we let $Q_1(\xi)$, resp. $Q_2(\xi)$, denote the largest gap between two consecutive particles, resp. between two consecutive empty sites.

$$\begin{aligned} Q_1(\xi) &:= \max\{n \ge 1 : \exists i \in [\![0, N-n]\!], \forall x \in [\![i+1, i+n]\!], \xi(x) = 0\}, \\ Q_2(\xi) &:= \max\{n \ge 1 : \exists i \in [\![0, N-n]\!], \forall x \in [\![i+1, i+n]\!], \xi(x) = 1\}, \end{aligned} \tag{29}$$

and $Q(\xi) = \max(Q_1(\xi), Q_2(\xi))$.

PROPOSITION 5. *For all* $x \in [\![1, N]\!]$, *we have*

$$\frac{k}{N}\lambda^{x-N} \le \pi_{N,k}\big(\xi(x) = 1\big) \le \frac{k}{N}\lambda^{x-1}. \tag{30}$$

Furthermore, there exists a constant $c > 0$ *such that for all choices of* $N \ge 1$, $u > 1$ *and* $p_N \in (1/2, 1]$ *and all* $k \le N/2$

$$\pi_{N,k}\left(Q(\xi) \ge \frac{\lambda^N N u}{k}\right) \le 2ke^{-cu}. \tag{31}$$

PROOF. We set $A_x := \{\xi(x) = 1\}$, we first prove that for all $y \in [\![1, N-1]\!]$ we have

$$\pi_{N,k}(A_y) \le \pi_{N,k}(A_{y+1}) \le \lambda \pi_{N,k}(A_y). \tag{32}$$

We observe that the map $\xi \mapsto \xi^y$ defined in (2) induces a bijection from A_y to A_{y+1} and that for every $\xi \in A^y$,

$$\pi_{N,k}(\xi) \le \pi_{N,k}(\xi^y) \le \lambda \pi_{N,k}(\xi). \tag{33}$$

The reader can check indeed that (from the detailed balance condition) $\pi_{N,k}(\xi^y) = \lambda \pi_{N,k}(\xi)$ if $\xi \in A_y \setminus A_{y+1}$ and that $\xi^y = \xi$ if $\xi \in A_y \cap A_{y+1}$. The desired inequalities (32) are then obtained by summing over $\xi \in A_y$.

By iterating (32) we obtain

$$\lambda^{x-N} \pi_{N,k}(A_N) \le \pi_{N,k}(A_x) \le \lambda^{x-1} \pi_{N,k}(A_1). \tag{34}$$

By monotony of $\pi_{N,k}(A_y)$ in y and the fact that there are k particles

$$N \pi_{N,k}(A_1) \le \sum_{y=1}^{N} \pi_{N,k}(A_y) = k \le N \pi_{N,k}(A_N),$$

and thus (30) can be deduced.

We pass to the proof of (31). We can perform the same reasoning as above but limiting ourselves to configurations with no particles in some set $I \subset [\![1, N]\!]$. Setting $B_I := \{\forall y \in I, \xi(y) = 0\}$ we obtain similarly to (32) (exchanging directly the content of x and y instead of nearest neighbors) that for every $x, y \in [\![1, N]\!] \setminus I$ with $x < y$

$$\pi_{N,k}(A_x \cap B_I) \le \pi_{N,k}(A_y \cap B_I) \le \lambda^{y-x} \pi_{N,k}(A_x \cap B_I). \tag{35}$$

This allows us to deduce that

$$\pi_{N,k}\big(\xi(x) = 1 | \forall y \in I, \xi(y) = 0\big) \ge \frac{k}{N - |I|} \lambda^{x-N}. \tag{36}$$

We have hence

$$\pi_{N,k}\big(\forall y \in I \cup \{x\}, \xi(y) = 0\big) \le \left(1 - \lambda^{-N} \frac{k}{N}\right) \pi_{N,k}\big(\forall y \in I, \xi(y) = 0\big), \tag{37}$$

which yields by induction

$$\pi_{N,k}\big(\forall x \in I, \xi(x) = 0\big) \le \left(1 - \lambda^{-N} \frac{k}{N}\right)^{|I|} \le \exp\left(-|I| \lambda^{-N} \frac{k}{N}\right). \tag{38}$$

Then noticing that $\{Q_1(\xi) \ge 2m\}$ implies that an interval of the type $[\![mi + 1, m(i+1)]\!]$ is empty, a union bound yields that

$$\pi_{N,k}\big(Q_1(\xi) \ge 2m\big) \le \left\lfloor \frac{N}{m} \right\rfloor \exp\left(-m \lambda^{-N} \frac{k}{N}\right). \tag{39}$$

This remains true for $Q_2(\xi)$ upon replacing k by $N - k$, and this concludes the proof of (31) if one choses $m = \frac{\lambda^N N u}{2k}$. $\square$

3.4. *Eigenfunctions and contractions*. The exact expression of the principal eigenfunction/eigenvalue has been derived in previous works [11, 14]. It turns out that it can be obtained by applying a discrete Hopf–Cole transform to the generator of our Markov chain. Let us recall some identities in that direction as they will be needed later on; the details can be found in [11], Section 3.3. We set

$$\varrho_N := (\sqrt{p}_N - \sqrt{q}_N)^2 \sim \frac{b_N^2}{2}, \tag{40}$$

and we let $a_{N,k}$ be the unique solution of

$$\begin{cases} (\sqrt{p_N q_N}\Delta - \varrho_N)a(x) = 0 & x \in [\![1, N-1]\!], \\ a(0) = 1, \qquad a(N) = \lambda^{\frac{2k-N}{2}}, \end{cases}$$

where Δ denotes the discrete Laplace operator

$$\Delta(f)(x) = f(x+1) + f(x-1) - 2f(x), \quad x \in [\![1, N-1]\!]. \tag{41}$$

If $(h_t^\zeta, t \geq 0)$ denotes the height function process starting from some arbitrary initial condition $\zeta \in \Omega_{N,k}$, then the map

$$V(t,x) := \mathbb{E}\big[\lambda^{\frac{1}{2}h_t^\zeta(x)} - a_{N,k}(x)\big], \quad t \geq 0, \quad x \in [\![0, N]\!],$$

solves

$$\begin{cases} \partial_t V(t,x) = (\sqrt{p_N q_N}\Delta - \varrho_N)V(t,x) & x \in [\![1, N-1]\!]. \\ V(t,0) = V(t,N) = 0. \end{cases} \tag{42}$$

The spectral decomposition of the operator $\sqrt{p_N q_N}\Delta - \varrho_N$ thus allows us (see [11], Section 3.3) to identify $N-1$ eigenvalues and eigenfunctions of the generator $\mathcal{L}_{N,k}$ of the Markov chain: for every $j \in \{1, \dots, N-1\}$, the map

$$f_{N,k}^{(j)}(\zeta) = \sum_{x=1}^{N-1} \sin\left(\frac{jx\pi}{N}\right)\left(\frac{\lambda^{\frac{1}{2}\zeta(x)} - a_{N,k}(x)}{\lambda - 1}\right), \tag{43}$$

defines an eigenfunction with eigenvalue

$$-\gamma_j = -\varrho_N - 4\sqrt{p_N q_N}\sin\left(\frac{j\pi}{2N}\right)^2.$$

The eigenvalue $-\gamma_1$ corresponds to the spectral gap of the generator (this is related to the fact that the corresponding eigenfunction is monotone (see [11], Section 3.3) for more details), and for this reason we adopt the notation

$$\mathrm{gap}_N := \gamma_1 = \varrho + 4\sqrt{p_N q_N}\sin\left(\frac{\pi}{2N}\right)^2.$$

We also set $f_{N,k} := f_{N,k}^{(1)}(\zeta)$ for the corresponding eigenfunction. Notice that this is a *strictly* increasing function (recall (24)). An immediate useful consequence of the eigenvalue equation is that

$$\mathbb{E}\big[f_{N,k}(h_t^{\zeta'}) - f_{N,k}(h_t^\zeta)\big] = e^{-\mathrm{gap}_N t}\big(f_{N,k}(\zeta') - f_{N,k}(\zeta)\big). \tag{44}$$

To close this subsection, let us introduce another function which is not an eigenfunction, but is also strictly increasing and enjoys a similar contraction property

$$f_{N,k}^{(0)}(\zeta) := \sum_{x=1}^{N-1} \frac{\lambda^{\zeta(x)/2} - a_{N,k}(x)}{\lambda - 1}.$$

As a direct consequence of (42) at time zero, we have (using the notation introduced in (41))

$$
\begin{aligned}
(\mathcal{L}_{N,k} f_{N,k}^{(0)})(\zeta) &= -\varrho_N f_{N,k}^{(0)}(\zeta) + \sqrt{p_N q_N} \sum_{x=1}^{N-1} \frac{\Delta(\lambda^{\zeta/2} - a_{N,k})(x)}{\lambda - 1} \\
&= -\varrho f_{N,k}^{(0)}(\zeta) - \frac{\sqrt{p_N q_N}}{\lambda - 1}\big[\lambda^{\frac{\zeta(N-1)}{2}} + \lambda^{\frac{\zeta(1)}{2}} - a_{N,k}(N-1) - a_{N,k}(1)\big].
\end{aligned} \tag{45}
$$

In particular, we obtain for $\zeta \le \zeta'$

$$
\begin{aligned}
&(\mathcal{L}_{N,k} f_{N,k}^{(0)})(\zeta') - (\mathcal{L}_{N,k} f_{N,k}^{(0)})(\zeta) \\
&\quad = -\varrho_N\big(f_{N,k}^{(0)}(\zeta') - f_{N,k}^{(0)}(\zeta)\big) - \frac{\sqrt{p_N q_N}}{\lambda - 1}\big[\lambda^{\frac{\zeta'(N-1)}{2}} + \lambda^{\frac{\zeta'(1)}{2}} - \lambda^{\frac{\zeta(N-1)}{2}} - \lambda^{\frac{\zeta(1)}{2}}\big] \\
&\quad \le -\varrho_N\big(f_{N,k}^{(0)}(\zeta') - f_{N,k}^{(0)}(\zeta)\big).
\end{aligned}
$$

Considering a monotone coupling between $(h_t^{\zeta'})_{t\ge 0}$ and $(h_t^{\zeta})_{t\ge 0}$, we obtain that

$$
\begin{aligned}
\partial_t \mathbb{E}\big[f_{N,k}^{(0)}(h_t^{\zeta'}) - f_{N,k}^{(0)}(h_t^{\zeta})\big] &= \mathbb{E}\big[(\mathcal{L}_{N,k} f_{N,k}^{(0)})(h_t^{\zeta'}) - (\mathcal{L}_{N,k} f_{N,k}^{(0)})(h_t^{\zeta})\big] \\
&\le -\varrho_N \mathbb{E}\big[f_{N,k}^{(0)}(h_t^{\zeta'}) - f_{N,k}^{(0)}(h_t^{\zeta})\big],
\end{aligned} \tag{46}
$$

and thus

$$
\mathbb{E}\big[f_{N,k}^{(0)}(h_t^{\zeta'}) - f_{N,k}^{(0)}(h_t^{\zeta})\big] \le e^{-\varrho_N t}\big(f_{N,k}^{(0)}(\zeta') - f_{N,k}^{(0)}(\zeta)\big). \tag{47}
$$

3.5. *The hydrodynamic limit.* We are interested in the macroscopic evolution of the height function.

For $\alpha \in [0,1]$, we define $\vee_\alpha : [0,1] \to \mathbb{R}$, $\wedge_\alpha : [0,1] \to \mathbb{R}$ as

$$
\vee_\alpha(x) := \max\big(-x, x - 2(1-\alpha)\big), \qquad \wedge_\alpha(x) := \min(x, 2\alpha - x),
$$

and we let $g_\alpha : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ be defined as follows:

$$
\begin{aligned}
g_\alpha^0(t,x) &:= \begin{cases} \alpha - \dfrac{t}{2} - \dfrac{(x-\alpha)^2}{2t} & \text{if } |x-\alpha| \le t, \\ \wedge^\alpha(x) & \text{if } |x-\alpha| \ge t, \end{cases} \\
g_\alpha(t,x) &:= \max\big(\vee_\alpha(x), g_\alpha^0(t,x)\big).
\end{aligned}
$$

PROPOSITION 6. *Assume that $Nb_N = N(p_N - q_N) \to \infty$ and that $k_N/N \to \alpha \in (0,1)$. Then, after an appropriate space-time scaling, $h^\wedge(\cdot,\cdot)$ converges to g_α in probability as $N \to \infty$. More precisely we have for any $\varepsilon > 0$, $T > 0$,*

$$
\lim_{N\to\infty} \mathbb{P}\Big[\sup_{t\le T} \sup_{x\in[0,1]} \Big|\frac{1}{N} h^\wedge\Big(\frac{Nt}{b_N}, Nx\Big) - g_\alpha(t,x)\Big| \ge \varepsilon\Big] = 0. \tag{48}
$$

PROOF. This is essentially the content of [10], Th 1.3, where the hydrodynamic limit of the density of particles is shown to be given by the inviscid Burgers' equation with zero-flux boundary conditions: when starting from the maximal initial condition, this yields (after integrating the density in space) the explicit solution g_α.

Actually the setting of [10], Th 1.3, is more restrictive as the number of particles is taken to be $k = N/2$ and $p_N - q_N = 1/N^\alpha$ with $\alpha \in (0,1)$. However, a careful inspection of the proof shows that we only require $N^{1-\alpha}$ to go to infinity: this corresponds to the assumption $N(p_N - q_N) \to \infty$ which is in force in the statement of the proposition so that the proof carries through *mutatis mutandis*. $\square$

4. Lower bound on the mixing time for large biases. In the large bias case, the last observable that equilibrates is the position of the leftmost particle. Obtaining a lower bound on the mixing time is thus relatively simple: we have to show that for arbitrary $\delta > 0$ at time

$$s_\delta(N) := \big[(\sqrt{\alpha} + \sqrt{1-\alpha})^2 - \delta\big] N b_N^{-1},$$

the leftmost particle has not reached its equilibrium position given by Lemma 4. This is achieved by using the hydrodynamic limit for $\alpha > 0$, and a simple comparison argument for $\alpha = 0$.

PROPOSITION 7. *When* (11) *is satisfied, for every* $\delta > 0$ *we have*

$$\lim_{N\to\infty} \big\|\mathbb{P}\big(\ell_N(\eta^\wedge_{s_\delta(N)}) \in \cdot\big) - \pi_N(\ell_N \in \cdot)\big\|_{\mathrm{TV}} = 1. \tag{49}$$

As a consequence for all $\varepsilon > 0$ *and* N *sufficiently large*

$$T^{N,k}_{\mathrm{mix}}(1-\varepsilon) \geq s_\delta(N).$$

4.1. *The case* $\alpha = 0$. Given $\delta > 0$ we want to prove that the system is not mixed at time $s_\delta(N) := (1-\delta) N b_N^{-1}$. We know from Lemma 4, that when $\alpha = 0$ and (11) holds, at equilibrium we have

$$\lim_{N\to\infty} \pi_{N,k}\big(\ell_N \leq (1-\delta/2)N\big) = 0.$$

On the other hand, observing that the position of the first particle is dominated by a random walk on $\mathbb{N}$ with jump rates p_N to the right and q_N to the left, the reader can check that whenever $\lim_{N\to\infty} b_N N = \infty$ we have

$$\lim_{N\to\infty} \mathbb{P}\big(\ell_N(\eta^\wedge_{s_\delta(N)}) \leq (1-\delta/2)N\big) = 1. \tag{50}$$

If the stochastic comparison was with a random walk on $\mathbb{Z}$, it would be sufficient to observe that at time $s_\delta(N)$ the expected value of the walk is $b_N s_\delta(N) = (1-\delta)N$ while the variance equal to $s_\delta \ll N^2$ which via Chebychev's inequality would imply (50). The constraint to remain positive makes the analysis a bit more delicate but the argument remains quite standard.

4.2. *The case* $\alpha \in (0, 1/2]$. Setting $x_\delta := 1 - \alpha - c_\alpha \delta$, for some positive constant c_α sufficiently small, we observe that the hydrodynamic profile at the rescaled time corresponding to s_δ is above the minimum at x_δ

$$g_\alpha\big(x_\delta, (\sqrt{\alpha} + \sqrt{1-\alpha})^2 - \delta\big) > \vee_\alpha(x_\delta).$$

The reader can check that $c_\alpha = 1/3$ works for all $\alpha \in (0, 1/2)$. Thus whenever $\lim_{N\to\infty} b_N N = \infty$ Proposition 6 yields that

$$\lim_{N\to\infty} \mathbb{P}\big(\ell_N(\eta^\wedge_{s_\delta}) \leq (1-\alpha-c_\alpha\delta)N\big) = 1. \tag{51}$$

On the other hand, we know from Lemma 4 that when (11) holds, at equilibrium we have for any $\delta > 0$

$$\lim_{N\to\infty} \pi_{N,k}\big(\ell_N \leq (1-\alpha-c_\alpha\delta)N\big) = 0.$$

5. Upper bound on the mixing time for large biases. Let us recall how a grand coupling satisfying the order preservation property (23) is of help to establish an upper bound on the mixing time. Recall (7). By [15], Lemma 4.10, we have

$$d^{N,k}(t) \le \max_{\zeta,\zeta' \in \Omega_{N,k}} \| P_t^{N,k}(\zeta',\cdot) - P_t^{N,k}(\zeta,\cdot)\|_{\mathrm{TV}}. \tag{52}$$

On the other hand, if $\mathbb{P}$ is a monotone grand coupling, one observes that the extremal initial conditions are the last to couple so that one has

$$\| P_t^{N,k}(\zeta',\cdot) - P_t^{N,k}(\zeta,\cdot)\|_{\mathrm{TV}} \le \mathbb{P}[h_t^{\zeta} \ne h_t^{\zeta'}] \le \mathbb{P}[h_t^{\vee} \ne h_t^{\wedge}]. \tag{53}$$

Hence to establish an upper bound on the mixing time, it suffices to obtain a good control on the merging time

$$\tau := \inf\{t > 0 : h_t^{\vee} = h_t^{\wedge}\}. \tag{54}$$

Let us set for this section

$$t_\delta(N) := [(\sqrt{\alpha} + \sqrt{1-\alpha})^2 + \delta] N b_N^{-1}. \tag{55}$$

PROPOSITION 8. *When* (11) *is satisfied, for every* $\delta > 0$, *and any monotone grand coupling we have*

$$\lim_{N\to\infty} \mathbb{P}(\tau \le t_\delta(N)) = 1. \tag{56}$$

As a consequence for all $\varepsilon > 0$ *and* N *sufficiently large* $T_{\mathrm{mix}}^{N,k}(\varepsilon) \le t_\delta(N)$.

When $\alpha = 1/2$, a sharp estimate on τ can be obtained directly from spectral considerations (Section 5.1), but when $\alpha \in [0, 1/2)$ we need a refinement of the strategy used in [11]: The first step (Proposition 10) is to obtain a control on the position of the leftmost particle which matches the lower bound provided by the hydrodynamic limit. This requires a new proof since the argument used in [11] is not sharp enough to cover all biases. The second step is to use contractive functions once the system is at macroscopic equilibrium, this is sufficient to treat most cases. A third and new step is required to treat the case when the bias b_N of order $N^{-1}\log N$ or smaller: as we are working under the assumption (11) we only need to treat this case when $k = N^{o(1)}$. In this third step we use diffusive estimates to control the hitting time of zero for the function $f_{N,k}^{(0)}(h_t^{\wedge}) - f_{N,k}^{(0)}(h_t^{\vee})$ where $f^{(0)}$ was introduced in Section 3.4.

5.1. *The special case* $\alpha = 1/2$. In the special case $\alpha = 1/2$, a sharp upper-bound can be deduced in a rather direct fashion from spectral estimates repeating the computation performed in [19], Section 3.2, for the symmetric case. This fact is itself a bit surprising since this method does not yield the correct upper bound in the small bias case (including the zero bias symetric case) nor in the constant bias case.

Recall the definition of $f_{N,k}$ in equation (43) and below. It being a strictly monotone function and $\mathbb{P}$ being a monotone coupling, we obtain using Markov's inequality (recall (25))

$$\mathbb{P}(\tau > t) = \mathbb{P}[f_{N,k}(h_t^{\wedge}) > f_{N,k}(h_t^{\vee})] \le \frac{\mathbb{E}[f_{N,k}(h_t^{\wedge}) - f_{N,k}(h_t^{\vee})]}{\delta_{\min}(f_{N,k})}. \tag{57}$$

The expectation decays exponentially with rate gap_N (44) and it is not difficult to check that

$$\delta_{\min}(f_{N,k}) \ge \lambda^{(k-N)/2} \sin\left(\frac{\pi}{N}\right) \ge N^{-1}\lambda^{(k-N)/2}. \tag{58}$$

Hence equation (57) becomes

$$\mathbb{P}(\tau > t) \le N\lambda^{\frac{N-k}{2}} e^{-\mathrm{gap}_N t}\big(f_{N,k}(\wedge) - f_{N,k}(\vee)\big) \le \frac{N^2\lambda^{N/2}}{\lambda - 1} e^{-\mathrm{gap}_N t}, \tag{59}$$

where in the last inequality we used that

$$\sin\left(\frac{x\pi}{N}\right)\big(\lambda^{\wedge(x)/2} - \lambda^{\vee(x)/2}\big) \le \lambda^{\wedge(x)/2} \le \lambda^{k/2}.$$

Recall that we assume that (11) holds and b_N tends to zero. Recalling (16), we obtain the following asymptotic equivalent:

$$\mathrm{gap}_N \overset{N\to\infty}{\sim} b_N^2/2 \quad \text{and} \quad \log\lambda_N \overset{N\to\infty}{\sim} 2b_N. \tag{60}$$

Furthermore for N sufficiently large we have $(\lambda - 1)^{-1} \le N$. Hence recalling that $t_\delta = (2 + \delta)b_N^{-1}N$ and using (60) in (59) we obtain for all N sufficiently large

$$\mathbb{P}(\tau > t_\delta) \le \frac{N^2}{\lambda - 1}\exp\left(\frac{N}{2}\log\lambda - \mathrm{gap}_N t_\delta\right) \le N^3 e^{-\frac{\delta}{4}b_N N}, \tag{61}$$

and the left-hand side vanishes when N tends to infinity as a consequence of (11) (recall that as $\alpha = 1/2$, k is of order N). $\square$

REMARK 9. The reason why the computation above yields a sharp upper bound is because: (A) The difference of order between $\delta_{\min}(f_{N,k})$ and the typical fluctuation of $f_{N,k}$ at equilibrium is negligible in the computation. (B) Until shortly before the mixing time the quantity $\log[f_{N,k}(h_t^\wedge) - f_{N,k}(h_t^\vee)]$ has the same order of magnitude as $\log\mathbb{E}[f_{N,k}(h_t^\wedge) - f_{N,k}(h_t^\vee)]$. In the case of symmetric exclusion, (A) does not hold, while when the bias is constant, (B) fails to hold. In the weakly asymmetric case, when $\alpha \ne 1/2$, the reader can check by combining Propositions 6 and 10 that (B) holds until time 4α (in the macroscopic time-scale) after which $g_\alpha(t,\cdot)$ stops to display a local maximum in the interval $(l_\alpha(t), r_\alpha(t))$ (defined in (63) below) and $f_{N,k}(h_t^\wedge) - f_{N,k}(h_t^\vee)$ starts to decay much faster than its average.

5.2. *The case $\alpha \ne 1/2$: Scaling limit for the boundary processes.* In order to obtain a sharp upper-bound for $\alpha \ne 1/2$, we rely on a scaling limit result in order to control the value of $f_{N,k}(h_t^\wedge) - f_{N,k}(h_t^\vee)$ up to a time close to the mixing time, and then we use the contractive estimate (44) to couple $h_t^\wedge$ with $h_t^\vee$. Note that Proposition 6 is not sufficient to estimate $f_{N,k}(h_t^\wedge)$: we also need a control on the positions of the left-most particle and right-most empty site in our particle configuration.

In the case when $\alpha = 0$ and the bias is of order $\log N/N$ or smaller (this is possible when (11) is satisfied and k grows slower than any power of N), we need an additional step, based on diffusion estimates, to couple the two processes. In this last case also, the factor N^{-1} in (58) causes some difficulty. For that reason we use $f_{N,k}^{(0)}$ and (47) instead of $f_{N,k}$ and (44): observe that $\delta_{\min}(f_{N,k}^{(0)}) = \lambda^{\frac{k-N}{2}}$.

Let us define $[L_N(t), R_N(t)]$ to be the interval on which $h_t^\wedge$ and $\vee$ differ. More explicitly, we set

$$\begin{aligned} L_N(t) &:= \max\{x : h^\wedge(t,x) = -x\}, \\ R_N(t) &:= \min\{x : h^\wedge(t,x) = x - 2(N-k)\}, \end{aligned} \tag{62}$$

or equivalently $L_N(t) := \ell_N(\eta_t^\wedge) - 1$ and $R_N(t) := r_N(\eta_t^\wedge)$.

We let ℓ_α and r_α denote the most likely candidates for the scaling limits of L_N and R_N that can be inferred from the hydrodynamic behavior of the system (cf. Proposition 6):

$$\ell_\alpha(t) = \begin{cases} 0 & \text{if } t \le \alpha, \\ (\sqrt{t} - \sqrt{\alpha})^2 & \text{if } t \in \big(\alpha, (\sqrt{\alpha} + \sqrt{1-\alpha})^2\big), \\ 1-\alpha & \text{if } t \ge (\sqrt{\alpha} + \sqrt{1-\alpha})^2, \end{cases}$$
$$r_\alpha(t) = \begin{cases} 1 & \text{if } t \le 1-\alpha, \\ 1 - (\sqrt{t} - \sqrt{1-\alpha})^2 & \text{if } t \in \big(1-\alpha, (\sqrt{\alpha} + \sqrt{1-\alpha})^2\big), \\ 1-\alpha & \text{if } t \ge (\sqrt{\alpha} + \sqrt{1-\alpha})^2. \end{cases} \tag{63}$$

We prove that ℓ_α and r_α are indeed the scaling limits of L_N and R_N.

PROPOSITION 10. *If* (11) *holds and* $k_N/N \to \alpha$ *then for every* $t > 0$ *we have the following convergences in probability*:

$$\lim_{N\to\infty} \frac{1}{N} L_N\big(b_N^{-1} N t\big) = \ell_\alpha(t), \qquad \lim_{N\to\infty} \frac{1}{N} R_N\big(b_N^{-1} N t\big) = r_\alpha(t). \tag{64}$$

REMARK 11. The assumption (11) is optimal for the above result to hold. To see this, the reader can check that when (11) fails, at equilibrium ℓ_N and r_N are typically at a macroscopic distance from $(1-\alpha)N$. In other words, (28) does not hold.

The proof of Proposition 10 is presented in the next subsections. Let us now check that it yields the right bound on mixing time. First, notice that the inequalities (57) still hold with $f_{N,k}$ replaced by $f_{N,k}^{(0)}$ since the latter is also a strictly increasing function in the sense of (24). Next observe that Proposition 10 allows an acute control on the quantity

$$\frac{f_{N,k}^{(0)}(h_t^\wedge) - f_{N,k}^{(0)}(h_t^\vee)}{\delta_{\min}(f_{N,k}^{(0)})}.$$

We summarize the argument in a lemma.

LEMMA 12. *Set* $D_N(\zeta) := \max(|L_N(\zeta) - N + k|, |R_N(\zeta) - N + k|)$. *We have for* $\zeta' \ge \zeta$

$$\frac{f_{N,k}^{(0)}(\zeta') - f_{N,k}^{(0)}(\zeta)}{\delta_{\min}(f_{N,k}^{(0)})} \le N k \lambda^{D_N(\zeta')}. \tag{65}$$

PROOF. We assume that $\zeta' \ne \zeta$. Then,

$$\frac{\lambda^{\frac{1}{2}\zeta'(x)} - \lambda^{\frac{1}{2}\zeta(x)}}{\lambda - 1} = \sum_{n=0}^{\frac{\zeta'(x)-\zeta(x)}{2}-1} \lambda^{\frac{1}{2}\zeta(x)+n} \le \lambda^{\frac{\zeta'(x)}{2}} \frac{(\zeta'(x) - \zeta(x))}{2}. \tag{66}$$

Now for $x \le L_N(\zeta')$ or $x \ge R_N(\zeta')$ we necessarily have $\zeta(x) = \zeta'(x) = \vee(x)$. For $x \in [\![L_N(\zeta') + 1, R_N(\zeta') - 1]\!]$, the fact that ζ' is 1-Lipschitz yields

$$\zeta'(x) \le k - N + 2D_N(\zeta'). \tag{67}$$

Recall that $\delta_{\min}(f_{N,k}^{(0)}) = \lambda^{(k-N)/2}$. Hence one obtains from (66)

$$\begin{aligned} \frac{f_{N,k}^{(0)}(\zeta') - f_{N,k}^{(0)}(\zeta)}{\delta_{\min}(f_{N,k}^{(0)})} &\le \sum_{x=1}^{N} \lambda^{\frac{\zeta'(x)-(k-N)}{2}} \frac{(\zeta'(x) - \zeta(x))}{2} \\ &\le \lambda^{D_N(\zeta')} \sum_{x=1}^{N} \frac{(\zeta'(x) - \zeta(x))}{2} \le \lambda^{D_N(\zeta')} N k. \end{aligned}$$

In the last inequality we simply used that $\zeta'(x) - \zeta(x) \le 2k$ (there are at most k sites where the increment of ζ' is larger than that of ζ). $\square$

We can now apply Proposition 10 to obtain an estimate on the mixing time. For convenience we treat the case of smaller bias separately.

5.2.1. *Proof of Proposition 8 when $b_N \gg (\log N)/N$.* We assume that

$$\lim_{N\to\infty} (b_N N)/\log N = \infty. \tag{68}$$

We consider first the system at time $t_0(N) := (\sqrt{\alpha} + \sqrt{1-\alpha})^2 N b_N^{-1}$. From Proposition 10, we know that at time t_0, L_N and R_N are close to their equilibrium positions: we have for N sufficiently large and arbitrary $\delta, \varepsilon > 0$

$$\mathbb{P}\big[L_N(t_0) \ge N - k - (\delta/20)N;\, R_N(t_0) \le N - k + (\delta/20)N\big] =: \mathbb{P}[\mathcal{A}_N] \ge 1 - (\varepsilon/2). \tag{69}$$

We let $\mathcal{F}_t$ denote the canonical filtration associated with the process. For $t \ge t_0$, repeating (57) starting at time t_0 for $f_{N,k}^{(0)}$ and combining it with (47), we obtain that

$$\mathbb{P}[\tau > t | \mathcal{F}_{t_0}] \le e^{-\varrho(t-t_0)} \frac{f_{N,k}^{(0)}(h_{t_0}^{\wedge}) - f_{N,k}^{(0)}(h_{t_0}^{\vee})}{\delta_{\min}(f_{N,k}^{(0)})}. \tag{70}$$

Note that on the event $\mathcal{A}_N$, we have $D_N(h_{t_0}^{\wedge}) \le \delta N/20$. Thus using Lemma 12 to bound the r.h.s. we obtain

$$\mathbb{E}\left[\frac{f_{N,k}^{(0)}(h_{t_0}^{\wedge}) - f_{N,k}^{(0)}(h_{t_0}^{\vee})}{\delta_{\min}(f_{N,k}^{(0)})} \,\middle|\, \mathcal{A}_N\right] \le kN\lambda^{\frac{\delta N}{20}}. \tag{71}$$

Hence averaging (70) on the event $\mathcal{A}_N$ one obtains

$$\mathbb{P}(\tau > t) \le \varepsilon/2 + \mathbb{P}[\tau > t | \mathcal{A}_N] \le \varepsilon/2 + e^{-\varrho(t-t_0)} kN\lambda^{\frac{\delta N}{20}}. \tag{72}$$

For $t = t_\delta = t_0 + \delta b_N^{-1} N$, replacing ϱ and $\log\lambda$ by their equivalents given in (40) and (60), one can check that for N sufficiently large one has

$$Nk\lambda^{\frac{\delta N}{20}} e^{-\varrho(t_\delta - t_0)} \le Nke^{-\frac{\delta N b_N}{20}} \le \varepsilon/2, \tag{73}$$

where the last inequality is valid for N sufficiently large provided that (68) holds.

5.2.2. *Proof of Proposition 8: The general case.* If we no longer assume that (68) holds, then an additional step is needed in order to conclude: this step relies on diffusion estimates proved in Appendix B. From (71) and (47), for any $\varepsilon, \delta > 0$ we have for N sufficiently large (recall (55))

$$\mathbb{E}\left[\frac{f_{N,k}^{(0)}(h_{t_{\delta/2}}^{\wedge}) - f_{N,k}^{(0)}(h_{t_{\delta/2}}^{\vee})}{\delta_{\min}(f_{N,k}^{(0)})} \,\middle|\, \mathcal{A}_N\right] \le e^{-\varrho(t_{\delta/2}-t_0)} \lambda^{\frac{\delta N}{20}} kN \le e^{-\frac{\delta N b_N}{40}} kN \le e^{-\frac{\delta N b_N}{50}} N, \tag{74}$$

where the second inequality relies on the the asymptotic equivalence in (60) and the last one on (11).

Now we can conclude using Proposition 29(i) with $a := 4\varepsilon^{-1} e^{-\delta N b_N/50} N$ and

$$(M_s)_{s\ge 0} := \left(\frac{f_{N,k}^{(0)}(h_{t_{\delta/2}+s}^{\wedge}) - f_{N,k}^{(0)}(h_{t_{\delta/2}+s}^{\vee})}{\delta_{\min}(f_{N,k}^{(0)})}\right)_{s\ge 0}.$$

Indeed M_s is a nonnegative supermartingale whose jumps are of size at least 1 (recall that we have divided the weighted area by $\delta_{\min}(f_{N,k}^{(0)})$ in the definition of M). Furthermore, up to the merging time τ, the two interfaces $h^\wedge$ and $h^\vee$ differ on some interval: on this interval $h^\wedge$ makes an upward corner ($\Delta h^\wedge < 0$) and $h^\vee$ makes a downward corner ($\Delta h^\vee > 0$). Consequently, the jump rate of M is at least 1 up to its hitting time of 0. From Markov's inequality we have (recall (69))

$$\mathbb{P}[M_0 > a] \le \mathbb{P}[\mathcal{A}_N^{\complement}] + a^{-1}\mathbb{E}[M_0|\mathcal{A}_N] \le 3\varepsilon/4. \tag{75}$$

Setting $r_\delta := (\delta/2)Nb_N^{-1}$ and applying (138), we have for all N sufficiently large

$$\mathbb{P}[M_{r_\delta} > 0|M_0 \le a] \le 4a(r_\delta)^{-1/2} \le \frac{16\varepsilon^{-1}}{\sqrt{\delta/2}}\sqrt{Nb_N}e^{-\frac{\delta Nb_N}{50}} \le \varepsilon/4, \tag{76}$$

where the last inequality comes from the fact that Nb_N diverges. Hence we conclude by observing that for N sufficiently large

$$\mathbb{P}[\tau > t_\delta] = \mathbb{P}[M_{r_\delta} > 0] \le \mathbb{P}[M_0 > a] + \mathbb{P}[M_{r_\delta} > 0|M_0 \le a] \le \varepsilon. \tag{77}$$

5.3. *An auxiliary model to control the speed of the right-most particle.* Our strategy to prove Proposition 10 is to compare our particle system with another one on the infinite line, for which a stationary probability exists. We consider n particles performing the exclusion process on the infinite line with jump rate p and q (since the parameter N plays no role in our auxiliary model, it makes no sense to write p_N here, the presented results are valid for any choice of p), and we add a "slower" $n+1$th particle on the right to enforce existence of a stationary probability for the particle spacings. To make the system more tractable this extra particle is only allowed to jump to the right (so that it does not feel the influence of the n others). Note that in our application, the number of particles n does not necessarily coincide with k.

The techniques developed in this section present some similarities to those used for the constant bias case in [11], Section 6, but also present several improvements, the main conceptual change being the addition of a slow particle instead of modifying the biases in the process. This novelty presents two advantages: Firstly it considerably simplifies the computation since martingale concentration estimates are not needed any more. Second this allows us to obtain control for the whole large bias regime (11), something that cannot be achieved even by optimizing all the parameters involved in [11], Section 6.

More formally we consider a Markov process $(\hat{\eta}(t))_{t\ge 0}$ on the state space

$$\Theta_n := \{\xi \in \mathbb{Z}^{n+1} : \xi_1 < \xi_2 < \cdots < \xi_{n+1}\}.$$

The coordinate $\hat{\eta}_i(t)$ denotes the position of the ith leftmost particle at time t. The dynamics are defined as follows: the first n particles, $\hat{\eta}_i(t)$, $i \in [\![1, n]\!]$ perform an exclusion dynamics with jump rates p to the right and q to the left while the last one $\hat{\eta}_{n+1}(t)$ can only jump to the right and does so with rate $\beta b = \beta(p-q)$, for some $\beta < 1$.

We assume furthermore that initially we have $\hat{\eta}_{n+1}(0) = 0$. The initial position of the other particles is chosen to be random in the following manner. We define

$$\mu_i := \beta + \lambda^{-i}(1-\beta), \tag{78}$$

and we assume that the spacings $(\hat{\eta}_{i+1}(0) - \hat{\eta}_i(0))_{i=1}^n$ are independent with Geometric distribution

$$\mathbb{P}[\hat{\eta}_{i+1}(0) - \hat{\eta}_i(0) = m] = (1-\mu_i)\mu_i^{m-1}, \quad m \ge 1. \tag{79}$$

Our aim is to prove the following control on the position of the first particle in this system, uniformly in β and n. In Sections 5.4 and 5.5, we use this result in order to control the position of $L_N(t)$.

PROPOSITION 13. *We have*,

$$\lim_{A\to\infty}\sup_{t\geq 1, n\in\mathbb{N}, \beta\in(0,1)} \mathbb{P}\Big[\hat{\eta}_1(t) \leq t\beta b - A\Big(\sqrt{bt} + \frac{1}{1-\beta}[n + b^{-1}\log\min(n, b^{-1})]\Big)\Big] = 0. \tag{80}$$

The statement is not hard to prove, the key point is to observe that the distribution of particle spacings is stationary.

LEMMA 14. *For all $t \geq 0$, $(\hat{\eta}_{i+1}(t) - \hat{\eta}_i(t))_{i=1}^n$ are independent r.v. with distribution given by* (79).

PROOF. We use the notation $(m_i)_{i=1}^n \in \mathbb{N}^n$ to denote a generic element in the configuration space for the process $(\hat{\eta}_{i+1}(t) - \hat{\eta}_i(t))_{i=1}^n$. We need to show that the measure defined above is stationary.

A measure π is stationary if and only if we have

$$\begin{aligned} &p\pi(m_1+1,\dots,m_n) \\ &\quad + \sum_{i=1}^{n-1}[p\pi(\dots, m_i - 1, m_{i+1}+1, \dots) + q\pi(\dots, m_{i-1}+1, m_i - 1, \dots)]\mathbf{1}_{\{m_i\geq 2\}} \\ &\quad + q\pi(\dots, m_{n-1}+1, m_n - 1)\mathbf{1}_{\{m_n\geq 2\}} + \beta b\pi(m_1,\dots,m_n - 1)\mathbf{1}_{\{m_n\geq 2\}} \\ &= \pi(m_1,\dots,m_n)\Big(q + \sum_{i=1}^{n-1}(p+q)\mathbf{1}_{\{m_i\geq 2\}} + p\mathbf{1}_{\{m_n\geq 2\}} + \beta b\Big), \end{aligned}$$

where in the sums, the dots stand for coordinates that are not modified (and m_{i-1} simply has to be ignored when $i = 1$). If we assume that π is the product of geometric laws with respective parameters $1 - \mu_i$ (not yet fixed) then the equation above is equivalent to the system

$$\begin{cases} p\mu_1 = q + \beta(p-q), \\ q\dfrac{\mu_{i-1}}{\mu_i} + p\dfrac{\mu_{i+1}}{\mu_i} = p + q, \quad \forall i \in [\![1, n-1]\!], \\ q\dfrac{\mu_{n-1}}{\mu_n} + \dfrac{\beta(p-q)}{\mu_n} = p, \end{cases} \tag{81}$$

where we have taken the convention $\mu_0 = 1$. One can readily check that μ_i given by (78) satisfies this equation. □

REMARK 15. Note that the equations (81) can be obtained directly simply by using the fact that the expected drifts of the particles starting from the geometric distributions are given by $p\mu_i - q\mu_{i-1}$ for the i-th particle $i \in [\![1, n]\!]$ and $\beta(p-q)$ for the $n+1$-th particle, and that stationarity implies that the drifts are all equal. However, the proof is necessary to show that this condition is also a sufficient one.

PROOF OF PROPOSITION 13. Starting from stationarity allows us to control the distance between the first and last particle at all times. In particular we have

$$\begin{aligned} \mathbb{E}[\hat{\eta}_{n+1}(t) - \hat{\eta}_1(t)] = \mathbb{E}[\hat{\eta}_{n+1}(0) - \hat{\eta}_1(0)] &= \sum_{i=1}^n \frac{1}{1-\mu_i} = \frac{1}{1-\beta}\sum_{i=1}^n \frac{1}{1-\lambda^{-i}} \\ &\leq \frac{1}{1-\beta}\Big(n + \frac{C}{\lambda - 1}\log(\min(n, |\lambda - 1|^{-1}))\Big), \end{aligned} \tag{82}$$

for some universal constant C. By union bound, the probability in the l.h.s. of (80) is smaller than

$$\mathbb{P}\big[\hat{\eta}_{n+1}(t) \le t\beta b - A\sqrt{bt}\big] + \mathbb{P}\bigg[\hat{\eta}_{n+1}(t) - \hat{\eta}_1(t) \ge \frac{A}{1-\beta}\big[n + b^{-1}\log\min\big(n, b^{-1}\big)\big]\bigg]. \tag{83}$$

The first term is small because, the r.v. $\hat{\eta}_{n+1}(t)$ being Poisson of parameter $t\beta b$, its expectation and variance are equal to $t\beta b$. The second can be shown to be going to zero with A using (82) and Markov's inequality for $\hat{\eta}_{n+1}(t) - \hat{\eta}_1(t)$. $\square$

5.4. *Proof of Proposition 10 in the case $\alpha = 0$.* We restate and prove the result in this special case (observe that the result for R_N is trivial for $\alpha = 0$).

PROPOSITION 16. *Assume that $\alpha = 0$ and (11) holds. We have for any $\varepsilon > 0$*

$$\lim_{N\to\infty} \sup_{t\ge 0} \mathbb{P}\big[\big|L_N(t) - \min(b_N t, N)\big| \ge \varepsilon N\big] = 0. \tag{84}$$

PROOF. First let us remark that the convergence

$$\lim_{N\to\infty} \sup_{t\ge 0} \mathbb{P}\big[L_N(t) \ge \min(b_N t, N) + \varepsilon N\big] = 0 \tag{85}$$

is obvious when $t \ge N b_N^{-1}$ and follows from the fact that the first particle is stochastically dominated by a simple random walk with bias $b_N \gg N^{-1}$ on the segment, starting from position 1 (recall the observation below (50)). It remains to prove that

$$\lim_{N\to\infty} \sup_{t\ge 0} \mathbb{P}\big[L_N(t) \le \min(b_N t, N) - \varepsilon N\big] = 0. \tag{86}$$

We prove first

$$\lim_{N\to\infty} \sup_{t\in[0, b_N^{-1}N]} \mathbb{P}\big[L_N(t) \le b_N t - \varepsilon N\big] = 0, \tag{87}$$

and then we briefly explain how to extend the result to times larger than $b_N^{-1}N$. We couple $\eta^\wedge(t)$ with the system $\hat{\eta}(t)$ of the previous subsection, choosing $n = k$ and $\beta = 1 - (\varepsilon/2)$. The coupling is obtained by making the ith particle in both processes try to jump at the same time (for $i \in [\![1, k]\!]$) with rate p and q, and rejection of the moves occurs as consequences of the exclusion rule or boundary condition (for $\eta^\wedge$ only). Initially of course we have

$$\forall i \in [\![1, n]\!], \quad \eta_i^\wedge(0) \ge \hat{\eta}_i(0), \tag{88}$$

because of the choice of the initial condition for $\hat{\eta}$ (recall that by definition $\eta_i^\wedge(0) = i$). The boundary at zero, and the presence of one more particle on the right in $\hat{\eta}$ gives $\eta^\wedge$ only more pushes towards the right, so that the ordering is preserved at least until $\hat{\eta}_{n+1}$ reaches the right side of the segment and the effect of the other boundary condition starts to be felt:

$$\eta_i^\wedge(t) \ge \hat{\eta}_i(t), \quad \forall i \in [\![1, n]\!], \forall t \le \mathcal{T} \tag{89}$$

where $\mathcal{T} := \inf\{t \ge 0 : \hat{\eta}_{n+1}(t) = N + 1\}$. We have

$$\mathbb{E}\big[\hat{\eta}_{n+1}(t)\big] = \mathrm{Var}\big(\hat{\eta}_{n+1}(t)\big) = t\beta b_N. \tag{90}$$

Using the assumption (11), and the fact that $\beta < 1$, using the estimate on the variance and expectation we have

$$\lim_{N\to\infty} \mathbb{P}\big[\mathcal{T} \le b_N^{-1}N\big] = \lim_{N\to\infty} \mathbb{P}\big[\hat{\eta}_{n+1}\big(b_N^{-1}N\big) \ge N + 1\big] = 0,$$

and hence from (89)

$$\lim_{N\to\infty} \sup_{t\le b_N^{-1}N} \mathbb{P}\big[\eta_1^\wedge(t) \le \hat\eta_1(t)\big] = 0. \tag{91}$$

Therefore, it suffices to control the probability of $\hat\eta_1(t) \le b_N t - \varepsilon N$. Observe that the assumptions $(k_N/N) \to 0$ and (11) imply that for any given $A > 0$, for all N sufficiently large and for any $t \le b_N^{-1}N$ we have

$$\left(\sqrt{b_N t} + \frac{1}{1-\beta}\big[k_N + b_N^{-1}\log\min\big(k_N, b_N^{-1}\big)\big]\right) \le \varepsilon\frac{N}{2A}.$$

Furthermore $(1-\beta)b_N t \le \varepsilon N/2$ for all $t \in [0, b_N^{-1}N]$. Thus applying Proposition 13 we obtain that for N sufficiently large

$$\begin{aligned} &\sup_{t\in[0,b_N^{-1}N]} \mathbb{P}\big[\hat\eta_1(t) \le b_N t - \varepsilon N\big] \\ &\quad\le \sup_{t\in[0,b_N^{-1}N]} \mathbb{P}\left[\hat\eta_1(t) \le \beta b_N t - A\left(\sqrt{b_N t} + \frac{k_N + b_N^{-1}\log\min(k, b_N^{-1})}{1-\beta}\right)\right] \le \delta, \end{aligned} \tag{92}$$

where δ can be made arbitrarily small by choosing A large. This concludes the proof of (87). To check (86) for larger times it suffices to shift the particle system $\hat\eta(0)$ to the left by $\lfloor t - b_N^{-1}N\rfloor$ and to apply the same arguments. □

5.5. *Proof of Proposition 10 in the case* $\alpha \in (0,1)$. The roles of L_N and R_N being symmetric, we only need to prove the result for L_N (but we do not assume here that $\alpha \le 1/2$). A direct consequence of Proposition 6 is that for all $s \in \mathbb{R}$ and $\varepsilon > 0$ we have

$$\lim_{N\to\infty} \mathbb{P}\big[\eta_1^\wedge\big(b_N^{-1}Ns\big) \ge N\big(\ell_\alpha(s) + \varepsilon\big)\big] = 0. \tag{93}$$

Hence to conclude we want to prove that

$$\lim_{N\to\infty} \mathbb{P}\big[\eta_1^\wedge\big(b_N^{-1}Ns\big) \le N\big(\ell_\alpha(s) - \varepsilon\big)\big] = 0. \tag{94}$$

For the remainder of the proof s and ε are considered as fixed parameters. We set $\delta \in (0,\alpha)$, and $n = \lceil \delta N\rceil$. To prove (94), we are going to compare $(\eta_i^\wedge)_{i=1}^n$ to the particle system considered in Section 5.3.

First we observe that as a consequence of Proposition 6, we have, for any $T > 0$

$$\lim_{N\to\infty} \mathbb{P}\big[\exists t \in [0,T], \eta_{n+1}^\wedge\big(b_N^{-1}Nt\big) \le N\ell_\alpha(t)\big] = 0. \tag{95}$$

We define the process $\hat\eta$ as in Section 5.3 with $\beta = 1 - \varepsilon/(2s)$ but with a shifted initial condition. More precisely we set

$$\hat\eta_{n+1}(0) = N\big(\ell_\alpha(s) - s\big) \le 0,$$

and choose the initial particle spacings to be independent and with geometric distributions given by (79). As (88) is satisfied, we can couple the two processes in such a way that

$$\forall i \in [\![1,n]\!], \forall t \le \mathcal{T}', \quad \eta_i^\wedge(t) \ge \hat\eta_i(t), \tag{96}$$

where $\mathcal{T}' := \inf\{t : \hat\eta_{n+1}(t) = \eta_{n+1}^\wedge(t)\}$. It is a simple exercise to show that for every $T > 0$ the position of $\hat\eta_{n+1}$ satisfies the following law of large numbers:

$$\lim_{N\to\infty} \mathbb{P}\left[\sup_{t\in[0,T]}\left|\frac{\hat\eta_{n+1}(b_N^{-1}Nt)}{N} - \big(\ell_\alpha(s) - s\big) - \beta t\right| \ge \kappa\right] = 0, \quad \forall\kappa > 0, \tag{97}$$

which, combined with (95), yields

$$\lim_{N\to\infty} \mathbb{P}[\mathcal{T}' \geq b_N^{-1} Ns] = 1, \tag{98}$$

and thus we only need to prove (94) with $\eta_1^\wedge$ replaced by $\hat{\eta}_1$. More precisely we prove that given $\kappa > 0$, one can find δ sufficiently small such that

$$\mathbb{P}[\hat{\eta}_1(b_N^{-1} Ns) \leq N(\ell_\alpha(s) - \varepsilon)] \leq \kappa. \tag{99}$$

Using Proposition 13 for $t = b_N^{-1} Ns$ and $A = \delta^{-1/2}$ and taking into account the new initial condition, the event

$$\left\{\hat{\eta}_1(b_N^{-1} Ns) \leq N\left(\ell_\alpha(s) - \frac{\varepsilon}{2}\right) - \delta^{-1/2}\left(\sqrt{Ns} + \frac{2s}{\varepsilon}[\delta N + b_N^{-1}\log b_N^{-1}]\right)\right\}$$

has a probability which can be made arbitrarily small if δ is chosen sufficiently small. We can then conclude that (99) holds by observing that for δ sufficiently small and N sufficiently large

$$\delta^{-1/2}\left(\sqrt{Ns} + \frac{2s}{\varepsilon}[\delta N + b_N^{-1}\log b_N^{-1}]\right) \leq \varepsilon N/2.$$

6. Lower bound on the mixing time for small biases. Until the end of the section, we assume that the small bias assumption (12) holds.

Let us set $s_\delta(N) := (1-\delta)\log k/(2\mathrm{gap}_N)$. We show that at time s_δ, equilibrium is not reached if one starts from one of the extremal conditions (some moderate efforts allow to replace max by min in the statement of the proposition).

PROPOSITION 17. *When assumption* (12) *holds, we have*

$$\lim_{N\to\infty} \max_{\zeta\in\{\vee,\wedge\}} \|\mathbb{P}(h^\zeta_{s_\delta(N)} \in \cdot) - \pi_{N,k}\|_{\mathrm{TV}} = 1. \tag{100}$$

As a consequence for every $\varepsilon \in (0,1)$, $T^{N,k}_{\mathrm{mix}}(\varepsilon) \geq s_\delta(N)$ for N sufficiently large.

The method to obtain a lower bound on the mixing time for small biases is similar to the one used in the symmetric case (see [19], Section 3.3), and is based on the control of the two first moments of $f_{N,k}(h_t^\wedge) - f_{N,k}(\zeta)$ where ζ is independent of $h_t^\wedge$ and distributed according to $\pi_{N,k}$: if at time t the mean of $f_{N,k}(h_t^\wedge) - f_{N,k}(\zeta)$ is much larger than its standard deviation, then the system is not at equilibrium (cf [15], Proposition 7.12).

We present estimates for the first two moments that we prove at the end of the section. This first moment bound is elementary.

LEMMA 18. *We have*

$$f_{N,k}(\wedge) - f_{N,k}(\vee) \geq \frac{1}{8}\lambda^{(k-N)/2} Nk, \tag{101}$$

and as a consequence, for every $t \geq 0$

$$\max(f_{N,k}(\wedge), -f_{N,k}(\vee)) \geq \frac{1}{16}\lambda^{(k-N)/2} Nk. \tag{102}$$

The second moment estimates rely on the control of a martingale bracket.

LEMMA 19. *For all $t \geq 0$, $N \geq 1$ and all $k \in [\![1, N/2]\!]$ we have*

$$\mathrm{Var}(f_{N,k}(h_t^{\wedge}) \leq \frac{k\lambda^k}{2\mathrm{gap}_N}. \tag{103}$$

The same bound holds for $\mathrm{Var}(f_{N,k}(h_t^{\vee}))$ *and* $\mathrm{Var}_{\pi_{N,k}}(f_{N,k})$.

PROOF OF PROPOSITION 17. Let us assume for simplicity (recall (102)) that

$$f_{N,k}(\wedge) \geq \frac{1}{16}\lambda^{(k-N)/2}Nk, \tag{104}$$

(if not we apply the same proof to $f_{N,k}(\vee) \leq -\frac{1}{16}\lambda^{(k-N)/2}Nk$). Recalling Section 3.4, $f_{N,k}$ being an eigenfunction associated with the spectral gap we have

$$\mathbb{E}[f_{N,k}(h_t^{\wedge})] = e^{-\mathrm{gap}_N t} f_{N,k}(\wedge) \geq \frac{1}{16}e^{-\mathrm{gap}_N t}\lambda^{(k-N)/2}Nk. \tag{105}$$

Applying [15], Proposition 7.12, for the probability measures $P_t^{N,k}(\wedge, \cdot)$ and $\pi_{N,k}$ and the function $f_{N,k}$ (recall that $E_{\pi_{N,k}}[f_{N,k}] = 0$), we obtain that

$$\|P_t^{N,k}(\wedge, \cdot) - \pi_{N,k}\|_{\mathrm{TV}} \geq 1 - \frac{2(\mathrm{Var}(f_{N,k}(h_t^{\wedge})) + \mathrm{Var}_{\pi_{N,k}}(f_{N,k}))}{\mathbb{E}[f_{N,k}(h_t^{\wedge})]^2}. \tag{106}$$

Using Lemma 19 and (105), we obtain that

$$\frac{\mathrm{Var}(f_{N,k}(h_t^{\wedge})) + \mathrm{Var}_{\pi_{N,k}}(f_{N,k})}{\mathbb{E}[f_{N,k}(h_t^{\wedge})]^2} \leq \frac{16^2 e^{2\mathrm{gap}_N t}\lambda^N}{\mathrm{gap}_N N^2 k}. \tag{107}$$

Now if we apply this inequality at time $s_\delta = (1-\delta)\log k/(2\mathrm{gap}_N)$, then we obtain for any given $\varepsilon > 0$ and all N sufficiently large

$$d(t_1) \geq 1 - 2\frac{16^2\lambda^N}{k^\delta \mathrm{gap}_N N^2} \geq 1 - \varepsilon, \tag{108}$$

where we used the small bias assumption (12). This yields $T_{\mathrm{mix}}^{N,k}(\varepsilon) \geq s_\delta$. □

PROOF OF LEMMA 18. We have

$$\begin{aligned} f_{N,k}(\wedge) - f_{N,k}(\vee) &= \sum_{x=1}^{N-1} \sin\left(\frac{x\pi}{N}\right)\frac{\lambda^{\frac{1}{2}\wedge(x)} - \lambda^{\frac{1}{2}\vee(x)}}{\lambda - 1} \\ &\geq \sum_{x=1}^{N-1} \sin\left(\frac{x\pi}{N}\right)\lambda^{\frac{1}{2}\vee(x)}\frac{\wedge(x) - \vee(x)}{2}, \end{aligned}$$

where the last inequality is obtained similarly to (66). Since the terms inside the sum are nonnegative for all x, we obtain a lower bound on this quantity by restricting x to $\{N/4, \dots, 3N/4\}$. Since $\vee(x) \geq k - N$ for all x and $\wedge(x) - \vee(x) \geq k$ for all $x \in \{N/4, \dots, 3N/4\}$, we conclude that

$$\sum_{x=1}^{N-1} \sin\left(\frac{x\pi}{N}\right)\lambda^{\frac{1}{2}\vee(x)}\frac{\wedge(x) - \vee(x)}{2} \geq \frac{\sqrt{2}}{2}\lambda^{\frac{k-N}{2}}\frac{Nk}{4}.$$

□

PROOF OF LEMMA 19. Using (cf. Section 3.4) the fact that $f_{N,k}$ is an eigenfuntion of the generator of the process, we know that

$$M_t := f_{N,k}(h_t^{\wedge})e^{\mathrm{gap}_N t}$$

is a martingale. Its predictable bracket is given by

$$\langle M.\rangle_t = \int_0^t \sum_{x=1}^{N-1} \lambda^{h_s^\wedge(x)} \sin\left(\frac{\pi x}{N}\right)^2 e^{2\mathrm{gap}_N s} \times \left(p_N \mathbf{1}_{\{\Delta h_s^\wedge(x)<0\}}\lambda^{-2} + q_N \mathbf{1}_{\{\Delta h_s^\wedge(x)>0\}}\right) \mathrm{d}s,$$

and $M_t^2 - \langle M.\rangle_t$ is again a martingale. This yields the identity

$$\mathrm{Var}\big(f_{N,k}(h_t^\wedge)\big) = e^{-2\mathrm{gap}_N t}\mathbb{E}\big[\langle M.\rangle_t\big].$$

To bound the predictable bracket of M, let us observe that the number of possible particle transitions to the right and to the left (the number of sites x such that $\Delta h_s^\wedge(x) < 0$, resp. > 0) is bounded by k, and that for any x and $\zeta \in \Omega_{N,k}$ we have $\lambda^{\zeta(x)} \le \lambda^k$. Therefore, we obtain the bound

$$\mathbb{E}\big[\langle M.\rangle_t\big] \le \int_0^t e^{2\mathrm{gap}_N s}\lambda^k \sum_{x=1}^{N-1} \mathbb{P}\big(\Delta h_s^\wedge(x) \neq 0\big)\,\mathrm{d}s \le k\lambda^k \frac{e^{2\mathrm{gap}_N t}}{2\mathrm{gap}_N},$$

which yields the asserted bound. The case of $h_t^\vee$ is treated in the same manner by symmetry. Since the distribution of $h_t^\wedge$ converges to $\pi_{N,k}$ when t tends to infinity we deduce that

$$\mathrm{Var}_{\pi_{N,k}}(f_{N,k}) = \lim_{t\to\infty} \mathrm{Var}\big(f_{N,k}(h_t^\wedge)\big),$$

which allows us to conclude. □

7. Upper bound on the mixing time for small biases. Until the end of the section we assume that the small bias assumption (12) holds and that the different initial conditions are coupled using the monotone grand coupling $\mathbb{P}$ defined in Appendix A. We set for all $\delta > 0$

$$t_\delta(N) := (1+\delta)\frac{\log k}{2\mathrm{gap}_N}.$$

Recall the definition of the merging time τ from (54).

PROPOSITION 20. *Assume that* (12) *holds. We have*

$$\lim_{N\to\infty} \mathbb{P}\big[\tau < t_\delta(N)\big] = 1. \tag{109}$$

As a consequence, for every $\varepsilon > 0$ and all N sufficiently large, $T_{\mathrm{mix}}^{N,k}(\varepsilon) \le t_\delta(N)$.

Recall (see the paragraph after (23)) that h_t^π denotes the chain with stationary initial condition. For practical reasons, it is simpler to couple two processes when at least one of them is at equilibrium. We thus prove (109) by showing that $\lim_{N\to\infty}\mathbb{P}[\tau_i < t_\delta(N)] = 1$ for $i \in \{1,2\}$ where

$$\tau_1 := \inf\{t > 0 : h_t^\wedge = h_t^\pi\} \quad \text{and} \quad \tau_2 := \inf\{t > 0 : h_t^\vee = h_t^\pi\}. \tag{110}$$

The argument being completely symmetric, we focus only on τ_1. As in Sections 5.2.1 and 5.2.2, we interpret τ_1 as the time at which the weighted area A_t between the maximal and equilibrium interface vanishes

$$A_t := \frac{f_{N,k}^{(0)}(h_t^\wedge) - f_{N,k}^{(0)}(h_t^\pi)}{\delta_{\min}(f_{N,k}^{(0)})} = \lambda^{\frac{N-k}{2}} \sum_{x=1}^{N-1} \frac{\lambda^{\frac{1}{2}h_t^\wedge(x)} - \lambda^{\frac{1}{2}h_t^\pi(x)}}{\lambda - 1}. \tag{111}$$

A simple computation based on the identity (42) shows that A is a supermartingale.

Our strategy to control τ_1 is to control the rate at which A_t shrinks. In a first step (cf. Section 7.1) we use the contractive estimates given by the spectral gap to show that after a time of order $\frac{1}{2}(\mathrm{gap}_N)^{-1}\log k_N$, A_t has shrunk from A_0 (which is of order kN) to $\sqrt{k}N$. Then we need to show that in a time which is $o((\mathrm{gap}_N)^{-1}\log k_N)$ A_t shrinks from $\sqrt{k}N$ to 0. We prove this in Sections 7.2 and 7.3 by controlling the fluctuations of A_t, which implies controlling the predictive bracket of the associated martingale. These computations rely on diffusion bounds for continuous supermartingale (in Appendix B) which might have some other applications.

While in the large bias case (Section 5) the choice of the grand coupling does not matter, here it is crucial to use a coupling which maximizes in a certain sense the fluctuation of the weighted area A_t, so that this process reaches zero as quickly as possible. The coupling defined in Appendix A makes the transitions for the two processes $h^\wedge$ and h^π as independent as possible (some transitions must occur simultaneously for the two processes in order to preserve monotonicity).

We consider $\eta > 0$ small and introduce the successive stopping times $\mathcal{T}_i$ by setting

$$\mathcal{T}_0 := \inf\{t \geq t_{\delta/2} : A_t \leq k^{\frac{1}{2}-\frac{\delta}{5}}N\},$$

and

$$\mathcal{T}_i := \inf\{t \geq \mathcal{T}_{i-1} : A_t \leq k^{\frac{1}{2}-i\eta-\frac{\delta}{5}}N\}, \quad i \geq 1.$$

We also set for coherence $\mathcal{T}_\infty := \max(\tau_1, t_{\delta/2})$ the first time at which A_t reaches 0. Notice that some of these stopping times may be equal to $t_{\delta/2}$.

Set $T_N := \min(b_N^{-2}, N^2)$. To prove Proposition 20, we show first that A_t shrinks to $k^{\frac{1}{2}-\frac{\delta}{5}}N$ by time $t_{\delta/2}$ and then that it only needs an extra time $2T_N$ to reach 0. The second step is performed by controlling each increment $\Delta\mathcal{T}_i := \mathcal{T}_i - \mathcal{T}_{i-1}$ separately for each i smaller than some threshold $K := \lceil 1/(2\eta)\rceil$.

LEMMA 21. *Given δ, if η is chosen small enough and $K := \lceil 1/(2\eta)\rceil$, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\{\mathcal{T}_0 = t_{\delta/2}\} \cap \left(\bigcap_{i=1}^{K}\{\Delta\mathcal{T}_i \leq 2^{-i}T_N\}\right) \cap \{\mathcal{T}_\infty - \mathcal{T}_K \leq T_N\}\right) = 1.$$

Note that on the event defined in the lemma and for all N large enough, we have

$$\tau_1 \leq \mathcal{T}_\infty \leq t_{\delta/2} + 2T_N \leq t_\delta.$$

Hence Proposition 20 follows as a direct consequence.

The bound on $\mathcal{T}_0$ is proved in Section 7.1, while that of on $\mathcal{T}_\infty - \mathcal{T}_K$ follows from Lemma 23 in Section 7.2, the case of the other increments is more delicate and is detailed in Section 7.3.

7.1. *Contraction estimates.* The approach used in the first step bears some similarity with the one used in Section 5.2.1, the notable difference being that (47) is not sufficient here and we must work a bit more to show that $\mathbb{E}[A_t]$ decays with rate gap_N.

LEMMA 22. *Given $\delta > 0$ we have $\mathbb{P}(\mathcal{T}_0 > t_{\delta/2}) \to 0$ as $N \to \infty$.*

PROOF. Note that $a(t,x) := \mathbb{E}[\frac{\lambda^{\frac{1}{2}h_t^\wedge(x)} - \lambda^{\frac{1}{2}h_t^\pi(x)}}{\lambda-1}]$ is a solution of the equation

$$\partial_t a = (\sqrt{pq}\Delta - \varrho)a,$$

with $a(t,0) = a(t,N) = 0$. Diagonalising the operator on the right-hand side (see Section 3.4), we get the following bound on the ℓ^2-norm of the solution:

$$\sum_{x=1}^{N-1} a(t,x)^2 \le e^{-2\mathrm{gap}_N t} \sum_{x=1}^{N-1} a(0,x)^2,$$

and using Cauchy–Schwartz inequality we obtain

$$\lambda^{\frac{k-N}{2}} \mathbb{E}[A_t] = \sum_{x=1}^{N-1} a(t,x) \le \sqrt{N} e^{-\mathrm{gap}_N t} \sqrt{\sum_{x=1}^{N-1} a(0,x)^2} \le 2e^{-\mathrm{gap}_N t} N k \lambda^{k/2}. \tag{112}$$

Since $\lambda^{N/2}$ is, by the small bias assumption, asymptotically smaller than any power of k, Markov's inequality concludes the proof. □

7.2. *Diffusion estimate after time $t_{\delta/2}$.* Now this part is much more delicate than Section 5.2.2. The reason being that since T_N is extremely close to t_δ, we need very accurate control on the derivative of the predictable bracket of A_t. Our first task is to use Proposition 29 in order to control the increment of the bracket of A in between the $\mathcal{T}_i$'s. Let us set

$$\Delta_i \langle A \rangle := \langle A_\cdot \rangle_{\mathcal{T}_i} - \langle A_\cdot \rangle_{\mathcal{T}_{i-1}}, \qquad \Delta_\infty \langle A \rangle := \langle A_\cdot \rangle_{\mathcal{T}_\infty} - \langle A_\cdot \rangle_{\mathcal{T}_K}, \tag{113}$$

and consider the event

$$\mathcal{A}_N := \{\forall i \in [\![1,K]\!], \Delta_i \langle A \rangle \le k^{1-2(i-1)\eta - \frac{\delta}{4}} N^2\} \cap \{\Delta_\infty \langle A \rangle \le T_N\}.$$

LEMMA 23. *We have* $\lim_{N\to\infty} \mathbb{P}[\mathcal{A}_N^{\complement}] = 0$.

PROOF. We apply Proposition 29(ii) to $(A_{t+\mathcal{T}_{i-1}})_{t\ge 0}$, with $a = k^{\frac{1}{2}-(i-1)\eta-\frac{\delta}{5}} N$, $b = k^{\frac{1}{2}-i\eta-\frac{\delta}{5}} N$. We obtain that for all N sufficiently large and every $i \le K$

$$\mathbb{P}\big[\Delta_i \langle A \rangle \ge k^{1-2(i-1)\eta-\frac{\delta}{4}} N^2\big] \le k^{-\delta/100}.$$

Applying the same proposition to $(A_{t+\mathcal{T}_K})_{t\ge 0}$ with $a = k^{-\frac{\delta}{5}} N$ and $b = 0$, we obtain

$$\mathbb{P}\big[\Delta_\infty \langle A \rangle \ge T_N\big] \le 8Nk^{-\frac{\delta}{5}} (T_N)^{-1/2}, \tag{114}$$

and the r.h.s. tends to zero by assumption (12). □

The next step is to compare $\Delta_i \langle A \rangle$ with $\mathcal{T}_i - \mathcal{T}_{i-1}$. For the last increment this is easy: We have $\partial_t \langle A_\cdot \rangle \ge 1$ for any $t \le \mathcal{T}_\infty$ (from our construction A changes its value at rate at least 1, and its minimal increment in absolute value is 1). We have thus $\mathcal{T}_\infty - \mathcal{T}_K \le \Delta_\infty \langle A \rangle$, and thus when $\mathcal{A}_N$ holds we have

$$\mathcal{T}_\infty - \mathcal{T}_K \le T_N. \tag{115}$$

7.3. *Control of intermediate increments.* For all other increments we have to use a subtler control of the bracket. Let us set

$$\mathbf{H}(t) := \lambda^{\frac{N-k}{2}} \max_{x\in[\![0,N]\!]} \frac{\lambda^{\frac{1}{2}h_t^\wedge(x)} - \lambda^{\frac{1}{2}h_t^\vee(x)}}{\lambda - 1},$$

which corresponds roughly (up to a multiplicative factor λ^N) to the maximal height difference $\max h_t^\wedge(x) - h_t^\vee(x)$ and thus provides a bound for $\max_x h_t^\wedge(x) - h_t^\pi(x)$.

Recall $Q(\cdot)$ from Section 3.3, and set $Q(h_t^\pi) := Q(\eta_t^\pi)$ where η_t^π is the particle configuration associated with h_t^π.

LEMMA 24. *We have* $\partial_t \langle A. \rangle \geq \frac{A_t}{6\mathbf{H}(t)Q(h_t^\pi)}$.

PROOF. As mentioned above, all the jumps of A_t have amplitude larger than or equal to 1. Moreover, A_t performs a jump whenever h_t^π performs a transition while $h_t^\wedge$ does not, or when the opposite occurs. As any such transition occurs at rate larger than $q_N \geq 1/3$, only considering the transitions for h_t^π, we obtain the following lower bound for the drift of $\langle A. \rangle$ (recall (41)):

$$\partial_t \langle A. \rangle \geq \frac{1}{3}\#\{x \in \mathcal{C}_t : \Delta(h_t^\pi)(x) \neq 0\} =: \frac{1}{3}\#\mathcal{D}_t \tag{116}$$

where

$$\mathcal{C}_t := \{x \in [\![1, N-1]\!] : \exists y \in [\![x-1, x+1]\!], h_t^\wedge(y) > h_t^\pi(y)\}.$$

Now let $[\![a, b]\!]$ be a maximal connected component of $\mathcal{C}_t$, we claim that

$$\#(\mathcal{D}_t \cap [\![a, b]\!]) \geq \max\left(\left\lfloor \frac{b-a}{Q(h_t^\pi)} \right\rfloor, 1\right) \geq \frac{b-a}{2Q(h_t^\pi)}. \tag{117}$$

To check the first inequality, notice that $\#(\mathcal{D}_t \cap [\![a, b]\!]) \geq 1$ because h_t^π cannot be linear on the whole segment $[\![a, b]\!]$. On the other hand, considering the particle configuration associated to h_t^π and decomposing the segment $[\![a, b]\!]$ into maximal connected components containing either only particles or only holes, we see that any two consecutive components corresponds to a point in $\mathcal{D}_t$: since $Q(h_t^\pi)$ is an upper bound for the size of these components, we deduce that $\#(\mathcal{D}_t \cap [\![a, b]\!]) \geq \lfloor \frac{b-a}{Q(h_t^\pi)} \rfloor$.

Now we observe that

$$\lambda^{\frac{N-k}{2}} \sum_{x=a}^{b} \frac{\lambda^{\frac{h_t^\wedge(x)}{2}} - \lambda^{\frac{h_t^\pi(x)}{2}}}{\lambda - 1} \leq (b-a)\mathbf{H}(t). \tag{118}$$

Combining (117) and (118) and summing over all such intervals $[\![a, b]\!]$, we obtain

$$A_t \leq 2\#\mathcal{D}_t \mathbf{H}(t) Q(h_t^\pi), \tag{119}$$

and (116) allows us to conclude. □

The last ingredient needed is then a bound on $\mathbf{H}$: The proof of this proposition is postponed to Section 7.4. Recall that $t_0 = \log k/(2\mathrm{gap}_N)$.

PROPOSITION 25. *For any* $c > 0$ *we have*

$$\lim_{N\to\infty} \sup_{t \geq t_0} \mathbb{P}\left(\mathbf{H}(t) > k^{\frac{1}{2}+c}\right) = 0. \tag{120}$$

PROOF OF LEMMA 21. By Lemma 22, Lemma 23 and equation (115), we already know that

$$\lim_{N\to\infty} \mathbb{P}\left(\{\mathcal{T}_0 \leq t_{\delta/2}\} \cap \{\mathcal{T}_\infty - \mathcal{T}_K \leq T_N\}\right) = 1.$$

We define $\mathcal{H}_N$ to be the event on which particles are reasonably spread and $\mathbf{H}(t)$ is reasonably small for most of the times within the interval $[t_{\delta/2}, t_{\delta/2} + T_N]$,

$$\mathcal{H}_N := \left\{ \int_{t_{\delta/2}}^{t_{\delta/2}+T_N} \mathbf{1}_{\{\mathbf{H}(t) \leq k^{\frac{1}{2}+\frac{\delta}{80}}\} \cap \{Q(h_t^\pi) \leq N k^{\frac{\delta}{80}-1}\}} \, \mathrm{d}t \geq T_N\left(1 - 2^{-(K+1)}\right) \right\}. \tag{121}$$

We have

$$\begin{aligned}\mathbb{E}\Big[T_N^{-1}\int_{t_{\delta/2}}^{t_{\delta/2}+T_N}\mathbf{1}_{\{\mathbf{H}(t)>k^{\frac{1}{2}+\frac{\delta}{80}}\}\cup\{Q(h_t^\pi)>Nk^{\frac{\delta}{80}-1}\}}\Big]\\ \le\Big(\pi_{N,k}\big(Q(\xi)>Nk^{\frac{\delta}{80}-1}\big)+\sup_{t\ge t_0}\mathbb{P}\mathbf{H}(t)>k^{\frac{1}{2}+\frac{\delta}{80}}\Big).\end{aligned}\tag{122}$$

By Proposition 5 and Proposition 25, the r.h.s. tends to zero and hence the variable in the expecation tends to 0 in probability. As a consequence we have

$$\lim_{N\to\infty}\mathbb{P}(\mathcal{H}_N)=1.$$

We now work on the event $\mathcal{H}_N\cap\mathcal{A}_N\cap\{\mathcal{T}_0\le t_{\delta/2}\}$ whose probability tends to 1 according to Lemmas 22, 23. We prove by induction that $\Delta\mathcal{T}_j\le 2^{-j}T_N$ for all $j\in[\![1,K]\!]$. Let us reason by contradiction and let i be the smallest integer such that $\Delta\mathcal{T}_i>2^{-i}T_N$. We have

$$\Delta_i\langle A\rangle\ge\int_{\mathcal{T}_{i-1}}^{\mathcal{T}_{i-1}+2^{-i}T_N}\partial_t\langle A.\rangle\mathbf{1}_{\{\mathcal{H}(t)\le k^{\frac{1}{2}+\frac{\delta}{80}}\}\cap\{Q(h_t^\pi)\le Nk^{\frac{\delta}{80}-1}\}}\,\mathrm{d}t.\tag{123}$$

Now, Lemma 24 and the restriction with the indicator function provides a uniform lower bound on $\partial_t\langle A\cdot\rangle$. The assumption $\Delta\mathcal{T}_j\le 2^{-j}T_N$ for $j<i$ implies that $\mathcal{T}_{i-1}\le t_{\delta/2}+T_N(1-2^{-(i-1)})$, and thus the assumption that $\mathcal{H}_N$ holds implies that the indicator in (123) is equal to one on a set of measure at least $2^{-i}-2^{-(K+1)}\ge 2^{-(K+1)}$. All of this implies that

$$\Delta_i\langle A\rangle\ge\frac{1}{6}T_N2^{-(K+1)}k^{1-i\eta-\frac{\delta}{40}-\frac{\delta}{5}}.\tag{124}$$

On the other hand, since we work on $\mathcal{A}_N$ we have $\Delta_i\langle A\rangle\le k^{1-2(i-1)\eta-\frac{\delta}{4}}N^2$ so that we get a contradiction as soon as η is small enough compared to δ. $\square$

7.4. *Bounding the maximum.* Recall the function $a_{N,k}$ defined in Section 3.4. Set

$$H_1(t,x):=\lambda^{\frac{N-k}{2}}\frac{\lambda^{\frac{1}{2}h_t^\wedge(x)}-a_{N,k}(x)}{\lambda-1},\qquad H_2(t,x):=\lambda^{\frac{N-k}{2}}\frac{\lambda^{\frac{1}{2}h_t^\vee(x)}-a_{N,k}(x)}{\lambda-1},$$

so that

$$H_1(t,x)-H_2(t,x)=\lambda^{\frac{N-k}{2}}\frac{\lambda^{\frac{1}{2}h_t^\wedge(x)}-\lambda^{\frac{1}{2}h_t^\vee(x)}}{\lambda-1}.$$

For every $i=1,2$, we define

$$\mathbf{H}_i(t):=\max_{x\in[\![0,N]\!]}\big|H_i(t,x)\big|.$$

Notice that $\mathbf{H}(t)\le\mathbf{H}_1(t)+\mathbf{H}_2(t)$ so that Proposition 25 is a consequence of the following result.

PROPOSITION 26. *For any $c>0$, there exists $c'>0$ such that for all N large enough*

$$\sup_{t\ge t_0}\max_{i\in\{1,2\}}\mathbb{P}\big(\mathbf{H}_i(t)>k^{\frac{1}{2}+c}\big)\le e^{-k^{c'}}.$$

The proof of this bound is split into two lemmas. First, we show that $H_i(t,\cdot)$ can not decrease too much.

LEMMA 27. *We have for all N sufficiently large, all $x \in [\![1, N-1]\!]$, all $t \geq 0$, every $i \in \{1, 2\}$ and every $y \geq x$*

$$H_i(t, y) - H_i(t, x) \geq -\frac{k^2(y-x)}{4N}. \tag{125}$$

PROOF. It is of course sufficient to prove that

$$H_i(t, x) - H_i(t, x-1) \geq -\frac{k^2}{4N}.$$

We focus on $i = 1$, since similar arguments yield the inequality for $i = 2$. Since $a_{N,k}(x) = E_{\pi_{N,k}}[\lambda^{h(x)/2}]$ where $a_{N,k}$ was defined in Section 3.4, we have the identity

$$\begin{aligned} H_1(t, x) - H_1(t, x-1) &= \lambda^{\frac{N-k}{2}} \frac{\lambda^{\frac{1}{2}h_t^\wedge(x)} - \lambda^{\frac{1}{2}h_t^\wedge(x-1)}}{\lambda - 1} \\ &\quad - \lambda^{\frac{N-k}{2}} \frac{E_{\pi_{N,k}}[\lambda^{\frac{1}{2}h(x)} - \lambda^{\frac{1}{2}h(x-1)}]}{\lambda - 1}. \end{aligned} \tag{126}$$

We have for any $\eta \in \Omega^0_{N,k}$, setting $h = h(\eta)$,

$$\frac{\lambda^{\frac{1}{2}h(x)} - \lambda^{\frac{1}{2}h(x-1)}}{\lambda - 1} = \lambda^{\frac{1}{2}(h(x-1)-1)}\left(\eta(x) - \frac{1}{\sqrt{\lambda}+1}\right). \tag{127}$$

Hence applying (127) for $h_t^\wedge$ and for h at equilibrium, we obtain

$$\begin{aligned} &H_1(t, x) - H_1(t, x-1) \\ &\quad = \lambda^{\frac{N-k}{2}} \frac{E_{\pi_{N,k}}[\lambda^{\frac{h(x-1)-1}{2}}] - \lambda^{\frac{h_t^\wedge(x-1)-1}{2}}}{\sqrt{\lambda}+1} \\ &\qquad - \lambda^{\frac{N-k}{2}}\left(E_{\pi_{N,k}}\left[\lambda^{\frac{h(x-1)-1}{2}}\eta(x)\right] - \lambda^{\frac{h_t^\wedge(x-1)-1}{2}}\eta_t^\wedge(x)\right) \\ &\quad \geq \lambda^{\frac{N-k}{2}} \frac{E_{\pi_{N,k}}[\lambda^{\frac{h(x-1)-1}{2}}] - \lambda^{\frac{h_t^\wedge(x-1)-1}{2}}}{\sqrt{\lambda}+1} - \lambda^{\frac{N-k}{2}} E_{\pi_{N,k}}\left[\lambda^{\frac{h(x-1)-1}{2}}\eta(x)\right]. \end{aligned} \tag{128}$$

By Proposition 5 and the small bias assumption (12), we have for all N large enough

$$\lambda^{\frac{N-k}{2}} E_{\pi_{N,k}}\left[\lambda^{\frac{h(x-1)-1}{2}}\eta(x)\right] \leq \lambda^{\frac{N}{2}} E_{\pi_{N,k}}[\eta(x)] \leq \lambda^{2N}\frac{k}{N} \leq \frac{k^2}{8N}.$$

Regarding the first term on the r.h.s. of (128), we simply notice that for $\zeta, \zeta' \in \Omega_{N,k}$, we have $\zeta(x) - \zeta'(x) \leq 2k$ so that

$$\lambda^{\frac{N-k}{2}}\left|\lambda^{\frac{\zeta(x)}{2}} - \lambda^{\frac{\zeta'(x)}{2}}\right| \leq (\lambda - 1)\lambda^{\frac{N}{2}}k \leq \frac{k^2}{8N}. \tag{129}$$

This is sufficient to conclude. □

Let us introduce the average of $H_i(t, \cdot)$ over a box of size $\ell = \ell_{N,k} = \lceil \frac{N}{k^2} \rceil$

$$\bar{H}_i(t, y) := \frac{1}{\ell} \sum_{x=(\ell-1)y+1}^{y\ell} H_i(t, x).$$

As a consequence of Lemma 27 we have

$$\mathbf{H}_i(t) = \max_{x\in[\![0,N]\!]} \big|H_i(t,x)\big| \le \max_{y\in[\![1,N/\ell]\!]} \big|\bar{H}_i(t,y)\big| + 1. \tag{130}$$

The result is of course obvious when $\ell = 1$. For $\ell \ge 2$, let us briefly explain why $\max H_i(t,x) \le \max|\bar{H}_i(t,y)| + 1$ (the case for $-\min$ follows by symmetry). If $x_{\max}$ is the smallest x at which the max is attained, we must distinguish two cases

(A) $x_{\max} > \ell(\lfloor N/\ell\rfloor - 1) + 1 \ge N - 2\ell$, in which case (125) applied for $x_{\max}$ and N implies that $H_i(t, x_{\max}) \le 1$,

(B) $x_{\max} \le \ell(\lfloor N/\ell\rfloor - 1) + 1$ in which case one can compare $H_i(t, x_{\max})$ with $\bar{H}_i(t,y)$ for the smallest y such that $x_{\max} \le y(\ell-1)+1$ using (125) again.

Then, Proposition 26 is a direct consequence of the following bound on the averages of H_i.

LEMMA 28. *For any $a > 0$, there exists $a' > 0$ such that for all N large enough*

$$\sup_{t\ge t_0} \max_{i\in\{1,2\}} \mathbb{P}\Big(\max_{y\in[\![1,N/\ell]\!]} \big|\bar{H}_i(t,y)\big| > k^{\frac{1}{2}+a}\Big) \le e^{-k^{a'}}.$$

PROOF. We treat in details the bound of $\bar{H}_1$, since the bound of $\bar{H}_2$ follows from the same arguments. Using a decomposition of $\lambda^{\frac{k-N}{2}} H_1(t,\cdot)$, which is a solution of (42), on the basis of eigenfunction of the Laplacian formed by $\sin(i\pi\cdot)$, $i = 1,\dots,N-1$, we obtain the following expression for the mean:

$$\mathbb{E}\big[H_1(t,x)\big] = \lambda^{\frac{N-k}{2}} \sum_{i=1}^{N-1} \frac{2}{N} e^{-\gamma_i t} f_{N,k}^{(i)}(\wedge) \sin\Big(\frac{i\pi x}{N}\Big),$$

and the fluctuation around it

$$H_1(t,x) - \mathbb{E}\big[H_1(t,x)\big] = \lambda^{\frac{N-k}{2}} \sum_{i=1}^{N-1} \frac{2}{N}\big(f_{N,k}^{(i)}(h_t^{\wedge}) - e^{-\gamma_i t} f_{N,k}^{(i)}(\wedge)\big) \sin\Big(\frac{i\pi x}{N}\Big). \tag{131}$$

We bound separately the contributions to $\bar{H}_1$ coming from these two terms. We start with the mean. Since $\lambda^{\frac{1}{2}\wedge(y)} \ge a_{N,k}(y) \ge \lambda^{\frac{1}{2}\vee(y)}$ for every $y \in [\![0,N]\!]$, we have (recall (66))

$$\big|f_{N,k}^{(i)}(\wedge)\big| \le \sum_{y=1}^{N-1} \frac{\lambda^{\frac{1}{2}\wedge(y)} - a_{N,k}(y)}{\lambda - 1} \le \sum_{y=1}^{N-1} \lambda^{\frac{1}{2}\wedge(y)} \frac{\wedge(y) - \vee(y)}{2} \le \lambda^{\frac{k}{2}} kN.$$

Since, by the small bias assumption (12), $\lambda^{k/2}$ is negligible compared to any power of k, we deduce that for all $a > 0$ and all $t \ge t_0$ we have for all N large enough

$$\sup_{x\in[\![0,N]\!]} \Big|\sum_{i=1}^{N-1} \frac{2}{N} e^{-\gamma_i t} f_{N,k}^{(i)}(\wedge) \sin\Big(i\pi\frac{x}{N}\Big)\Big| \le k^{(1+a)/2} \sum_{i=1}^{N-1} e^{(\gamma_1-\gamma_i)t_0}.$$

Notice that there exists $c > 0$ such that for all $i \ge 2$ and all N large enough

$$\gamma_i - \gamma_1 \ge c\frac{i^2}{N^2}. \tag{132}$$

In addition, we have $N^2 \mathrm{gap}_N \ll (\log k)^2$ by the small bias assumption (12), so that we get for $i \ge 2$

$$e^{(\gamma_1-\gamma_i)t_0} \le e^{-c\frac{i^2}{N^2}\frac{\log k}{2\mathrm{gap}_N}} \le e^{-c'\frac{i^2}{\log k}} \le e^{-c'\frac{i}{\log k}},$$

so that for all N large enough we have

$$\sum_{i=2}^{N-1} e^{(\gamma_1-\gamma_i)t_0} \le \sum_{i=2}^{N-1} e^{-c'\frac{i}{\log k}} \le C\log k.$$

Recall that $\ell = \lceil N/k^2 \rceil$. Putting everything together and using assumption (12), we get that given $a > 0$ for N sufficiently large and all values of y we have

$$\mathbb{E}[\bar{H}_1(t,y)] = \frac{2\lambda^{\frac{N-k}{2}}}{N\ell}\left|\sum_{x=\ell(y-1)+1}^{\ell y}\sum_{i=1}^{N-1} e^{-\gamma_i t} f_{N,k}^{(i)}(\wedge)\sin\left(i\pi\frac{x}{N}\right)\right| \le \frac{1}{2}k^{\frac{1}{2}+a}. \tag{133}$$

We turn to the contribution coming from the second term (131). To that end, we rewrite it in the form

$$\bar{H}_1(s,y) - \mathbb{E}[\bar{H}_1(s,y)] = \lambda^{\frac{N-k}{2}}\sum_{i=1}^{N-1}\big(f_{N,k}^{(i)}(h_t^\wedge) - e^{-\gamma_i s} f_{N,k}^{(i)}(\wedge)\big)\Phi_{y,i}, \tag{134}$$

where (the second expression being obtained by summation by part)

$$\begin{aligned}\Phi_{y,i} &= \frac{2}{N\ell}\sum_{x=y(\ell-1)+1}^{y\ell}\sin\left(\frac{i\pi x}{N}\right)\\ &= \frac{1}{N\ell\sin(\frac{i\pi}{2N})}\left[\cos\left(\frac{[2y(\ell-1)+1]i\pi}{2N}\right) - \cos\left(\frac{[2y\ell+1]i\pi}{2N}\right)\right].\end{aligned} \tag{135}$$

Note that for all $N \ge 1$, y and i we have

$$|\Phi_{y,i}| \le 2\min\left(\frac{1}{N}, \frac{1}{i\ell}\right).$$

Now let us fix t and y and introduce the martingale

$$N_s^{(t,y)} = \lambda^{\frac{N-k}{2}}\sum_{i=1}^{N-1} e^{\gamma_i(s-t)}\big(f_{N,k}^{(i)}(h_s^\wedge) - e^{-\gamma_i s} f_{N,k}^{(i)}(\wedge)\big)\Phi_{y,i}, \quad s \in [0,t],$$

which satisfies

$$N_0^{(t,y)} = 0 \quad \text{and} \quad N_t^{(t,y)} = \bar{H}_1(t,y) - \mathbb{E}[\bar{H}_1(t,y)].$$

We wish to apply Lemma 31 to the martingale $N^{(t,y)}$: the maximal jump rate of this process is bounded by k and the maximal amplitude of the jump (cf. the notations introduced in Appendix C) satisfies

$$\forall s \in [0,t], \quad S(s) \le \lambda^N \sum_{i=1}^{N-1} e^{\gamma_i(s-t)}|\Phi_{y,i}|.$$

From (132) we have $\gamma_i \ge ci^2N^{-2}$ for all $i \ge 1$ for some positive constant c. We deduce that for some adequate choice of the constants $C, C' > 0$ we have

$$\begin{aligned}\int_0^t S(s)^2\,\mathrm{d}s &\le \lambda^{2N}\sum_{i,j=1}^{N-1}\frac{1}{\gamma_i+\gamma_j}|\Phi_{y,i}\Phi_{y,j}|\\ &\le C\lambda^{2N}\Bigg(\sum_{1\le i\le j\le \frac{N}{\ell}}\frac{1}{i^2+j^2} + \sum_{1\le i\le \frac{N}{\ell}<j}\frac{N}{j\ell}\frac{1}{i^2+j^2} + \sum_{\frac{N}{\ell}<i\le j}\frac{N^2}{ij\ell^2}\frac{1}{i^2+j^2}\Bigg)\\ &\le C'\lambda^{2N}\log k.\end{aligned}$$

Consequently, setting $\gamma = k^{-\frac{1}{2}-2a}$ and using the fact that from (12) we have $\gamma C'\lambda^{2N}\log k < 1$ for N sufficiently large, we apply (151) and obtain

$$\begin{aligned}\mathbb{P}\left(N_t^{(t,y)} > \frac{1}{2}k^{\frac{1}{2}+a}\right) &\le \mathbb{E}\big[e^{\gamma N_t^{(t,y)}}\big]e^{-\frac{1}{2}\gamma k^{\frac{1}{2}+a}} \\ &\le e^{C'e\gamma^2 k\lambda^{2N}\log k - \frac{1}{2}\gamma k^{\frac{1}{2}+a}} \le e^{-\frac{1}{4}k^{-a}}.\end{aligned} \tag{136}$$

A similar same computation for $N_t^{(t,y)} < -\frac{1}{2}k^{\frac{1}{2}+a}$ and a union bound yield

$$\begin{aligned}&\mathbb{P}\left(\sup_{y\in[\![1,N/\ell]\!]}\big|\bar{H}_i(t,y) - \mathbb{E}\big[H_i(t,y)\big]\big| > \frac{1}{2}k^{\frac{1}{2}+a}\right) \\ &\quad = \mathbb{P}\left(\sup_{y\in[\![1,N/\ell]\!]}\big|N_t^{(t,y)}\big| > \frac{1}{2}k^{\frac{1}{2}+a}\right) \le Ck^2 e^{-\frac{1}{4}k^{-a}},\end{aligned} \tag{137}$$

which combined with (133) allows us to conclude. □

APPENDIX A: A MONOTONE GRAND COUPLING

The construction below is similar to the one detailed in [13], Section 8.1, in the symmetric case. We consider a collection of independent Poisson clock processes $\mathcal{P}^{(i,\ell)}$ and $\mathcal{Q}^{(i,\ell)}$ with rate p and q respectively where $i \in [\![1, N]\!]$ and $\ell \in [\![-N, \dots, N]\!]$: For each (i, ℓ), $\mathcal{P}^{(i,\ell)}$ resp. $\mathcal{Q}^{(i,\ell)}$ is a random increasing sequence of positive real numbers (or equivalently a random locally finite subset of $(0, \infty)$) whose first term and increments are independent geometric variables of mean p^{-1} resp. q^{-1}.

For every k and every $\zeta \in \Omega_{N,k}$, we construct the process $(h_t^\zeta)_{t\ge 0}$ as follows: The process is càd-làg and may only jump at the times specified by the clock process $\mathcal{P}$ and $\mathcal{Q}$. We enumerate these Poisson times in increasing order and if $t \in \mathcal{P}^{(i,\ell)}$ and if h_{t-}^ζ displays a local maximum at i and height ℓ, that is if

$$h_{t-}^\zeta(i) = \ell = h_{t-}^\zeta(i-1) + 1 = h_{t-}^\zeta(i+1) + 1,$$

then we flip it downwards to a local minimum by setting, $h_t^\zeta(i) := h_{t-}^\zeta(i) - 2$, and $h_t^\zeta(j) = h_{t-}^\zeta(j)$ for $j \ne i$. A similar transition occurs if $Q(i, \ell)$ rings and if h_{t-}^ζ displays local minimum at i and height ℓ.

It is simple to check that under this construction, h^ζ indeed evolves according to the right dynamics, and that monotonicity is preserved.

APPENDIX B: DIFFUSION BOUNDS FOR CONTINUOUS-TIME SUPERMARTINGALES

In this section, we assume that $(M_t)_{t\ge 0}$ is a pure-jump supermartingale with bounded jump rate and jump amplitude. This implies in particular that, M_t is square integrable for all $t > 0$. With some abuse of notation, we use the notation $\langle M.\rangle_t$ for the predictable bracket associated with the martingale $\widetilde{M}_t = M_t - A_t$ where A is the compensator of M.

PROPOSITION 29. *Let $(M_t)_{t\ge 0}$ be as above*

(i) *Set $\tau = \inf\{t \ge 0 : M_t = 0\}$. Assume that M_t is nonnegative and that, until the absorption time τ, its jump amplitude and jump rate are bounded below by 1. Then we have for any $a \ge 1$ and all $u > 0$*

$$\mathbb{P}\big[\tau \ge a^2 u | M_0 \le a\big] \le 4u^{-1/2}. \tag{138}$$

(ii) *Given $a \in \mathbb{R}$ and $b \le a$, we set $\tau_b := \inf\{t \ge 0 : M_t \le b\}$. If the amplitude of the jumps of $(M_t)_{t\ge 0}$ is bounded above by $a-b$, we have for any $u \ge 0$*

$$\mathbb{P}\big[\langle M.\rangle_{\tau_b} \ge (a-b)^2 u | M_0 \le a\big] \le 8u^{-1/2}. \tag{139}$$

The important building block for the proof of the above proposition is the following result.

LEMMA 30. *Let $(M_t)_{t\ge 0}$ be as above*

(i) *If the amplitude of the jumps of $(M_t)_{t\ge 0}$ and the jump rate are bounded below by* 1 *then for all $\lambda \in (0,1)$,*

$$\big(e^{-\lambda M_t - \frac{\lambda^2 t}{4}}\big)_{t\ge 0}$$

is a submartingale,

(ii) *If the amplitude of the jumps of M_t is bounded above by a then for any $\lambda \in (0, a^{-1})$,*

$$\left(\exp\left(-\lambda M_t - \frac{\lambda^2}{4}\langle M.\rangle_t\right)\right)_{t\ge 0}$$

is a submartingale.

PROOF OF PROPOSITION 29. The result only needs to be proved for $u \ge 4$. Without loss of generality one can assume for the proof of both statements that $P[M_0 \le a] = 1$ and for the second one that $b = 0$. We set $\lambda = 2a^{-1}u^{-1/2}$.

For (i), we apply the Martingale Stopping Theorem to the submartingale given by Lemma 30(i). It yields:

$$\mathbb{E}\big[e^{-\frac{\lambda^2\tau}{4}}\big] \ge \mathbb{E}\big[e^{-\lambda M_0}\big] \ge e^{-\lambda a} = e^{-2u^{-1/2}}. \tag{140}$$

On the other hand, one has

$$\mathbb{E}\big[e^{-\frac{\lambda^2\tau}{4}}\big] \le 1 - \big(1 - e^{-\frac{\lambda^2}{4}a^2 u}\big)\mathbb{P}\big[\tau \ge a^2 u\big] \le 1 - \frac{1}{2}\mathbb{P}\big[\tau \ge a^2 u\big]. \tag{141}$$

The combination of the two yields

$$\mathbb{P}\big[\tau \ge a^2 u\big] \le 2\big(1 - e^{-2u^{-1/2}}\big) \le 4u^{-1/2}. \tag{142}$$

For (ii), the arguments of the previous case apply almost verbatim if one replaces τ by $T := \langle M.\rangle_\tau$. The only thing one has to take into account is that M_τ is not necessarily equal to 0, but the assumption on the amplitude of jumps yields $M_\tau \ge -a$. The Martingale Stopping Theorem gives us

$$\mathbb{E}\big[e^{-\frac{\lambda^2 T}{4}}\big] \ge e^{-\lambda a}\mathbb{E}\big[e^{-\lambda M_\tau - \frac{\lambda^2 T}{4}}\big] \ge e^{-\lambda a}\mathbb{E}\big[e^{-\lambda M_0}\big] \ge e^{-4u^{-1/2}}. \tag{143}$$

Repeating the rest of the computation yields

$$\mathbb{P}\big[T \ge a^2 u\big] \le 2\big(1 - e^{-4u^{-1/2}}\big) \le 8u^{-1/2}. \tag{144}$$

□

PROOF OF LEMMA 30. Until the end of the proof, we write $\mathbb{E}_t$ for the conditional expectation given $(M_s)_{s\le t}$.

Case (i). Take $\lambda \in (0,1)$. The submartingale identity we need to prove can be written as follows:

$$\forall s, t \ge 0, \quad \log \mathbb{E}_t\big[e^{-\lambda(M_{t+s} - M_t)}\big] \ge s\frac{\lambda^2}{4}.$$

Taking derivative, we deduce that it suffices to prove that for all $t, s \geq 0$ we have

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}_t\big[e^{-\lambda(M_{t+s+h}-M_t)} - e^{-\lambda(M_{t+s}-M_t)}\big] \geq \frac{\lambda^2}{4} \mathbb{E}_t\big[e^{-\lambda(M_{t+s}-M_t)}\big]. \tag{145}$$

Notice that for all $x \in \mathbb{R}$

$$e^{-x} + x - 1 \geq \frac{\min(1, x^2)}{4}. \tag{146}$$

Thus, using the supermartingale property of M we have for all $\lambda \in (0, 1)$

$$\begin{aligned}
&\mathbb{E}_t\big[e^{-\lambda(M_{t+s+h}-M_t)} - e^{-\lambda(M_{t+s}-M_t)}\big] \\
&\quad = \mathbb{E}_t\big[e^{-\lambda(M_{t+s}-M_t)} \mathbb{E}_{t+s}\big[e^{-\lambda(M_{t+s+h}-M_{t+s})} - 1\big]\big] \\
&\quad \geq \mathbb{E}_t\big[e^{-\lambda(M_{t+s}-M_t)} \mathbb{E}_{t+s}\big[e^{-\lambda(M_{t+s+h}-M_{t+s})} + \lambda(M_{t+s+h} - M_{t+s}) - 1\big]\big] \\
&\quad \geq \frac{\lambda^2}{4} \mathbb{E}_t\big[e^{-\lambda(M_{t+s}-M_t)} \mathbb{E}_{t+s}\big[\min\big(1, (M_{t+s+h} - M_{t+s})^2\big)\big]\big].
\end{aligned}$$

The assumption on the jump rates and the jump amplitudes yield

$$\liminf_{h \to 0} \frac{1}{h} \mathbb{E}_{t+s}\big[\min\big(1, (M_{t+s+h} - M_{t+s})^2\big)\big] \geq 1, \tag{147}$$

so that Fatou's Lemma concludes the proof.

Case (ii). We can assume without loss of generality that $a = 1$. Here again, taking the derivative of the submartingale identity that we want to establish, it suffices to prove that for all $t, s \geq 0$ we have

$$\liminf_{h \downarrow 0} \frac{1}{h} \mathbb{E}_t\big[e^{-\lambda M_{t+s+h} - \frac{\lambda^2}{4}\langle M.\rangle_{t+s+h}} - e^{-\lambda M_{t+s} - \frac{\lambda^2}{4}\langle M.\rangle_{t+s}}\big] \geq 0.$$

Taking the conditional expectation w.r.t. M_{t+s}, we see that it suffices to prove the existence of some deterministic constant $C > 0$ such that

$$\mathbb{E}_{t+s}\big[e^{-\lambda(M_{t+s+h}-M_{t+s}) - \frac{\lambda^2}{4}(\langle M.\rangle_{t+s+h} - \langle M.\rangle_{t+s})} - 1\big] \geq -Ch^2, \tag{148}$$

for all h small enough.

Without loss of generality, we can assume that $t + s = 0$ and $M_0 = 0$. Recall that $\widetilde{M} \geq M$ is the martingale which is obtained by subtracting the (negative) compensator. Thus

$$\begin{aligned}
e^{-\lambda M_h - \frac{\lambda^2}{4}\langle M.\rangle_h} - 1 &\geq e^{-\lambda \widetilde{M}_h - \frac{\lambda^2}{4}\langle M.\rangle_h} - 1 \\
&\geq \Big(1 - \lambda \widetilde{M}_h + \frac{1}{4}\min\big(1, \lambda^2 \widetilde{M}_h^2\big)\Big)\Big(1 - \frac{\lambda^2}{4}\langle M.\rangle_h\Big) - 1 \\
&\geq -\lambda \widetilde{M}_h + \frac{\lambda^2}{4}\big(\widetilde{M}_h^2 - \langle M.\rangle_h\big) - \frac{\lambda^2}{4}\big(\widetilde{M}_h^2 - \lambda^{-2}\big)_+ \\
&\quad + \frac{\lambda^2}{4}\langle M.\rangle_h\Big(\lambda \widetilde{M}_h - \frac{1}{4}\min\big(1, \lambda^2 \widetilde{M}_h^2\big)\Big),
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{E}\big[e^{-\lambda M_h - \frac{\lambda^2}{4}\langle M.\rangle_h} - 1\big] \geq \mathbb{E}\Big[&-\frac{\lambda^2}{4}\big(\widetilde{M}_h^2 - \lambda^{-2}\big)_+ \\
&+ \frac{\lambda^2}{4}\langle M.\rangle_h\Big(\lambda \widetilde{M}_h - \frac{1}{4}\min\big(1, \lambda^2 \widetilde{M}_h^2\big)\Big)\Big].
\end{aligned}$$

Take $\lambda \in (0,1)$. Our assumptions on the increments and jump rates imply that for some constant $C > 0$ we have

$$\begin{aligned} \mathbb{E}\big[(\widetilde{M}_h^2 - \lambda^{-2})_+\big] &\le Ch^2, \\ \langle M. \rangle_h &\le Ch, \\ \max\big(\mathbb{E}[|\widetilde{M}_h|], \mathbb{E}[\min(1, \lambda^2 \widetilde{M}_h^2)]\big) &\le Ch, \end{aligned} \tag{149}$$

(the compensator being of order h the estimates for $\widetilde{M}$ can be deduced from that for M), which allows us to conclude that (148) holds. □

APPENDIX C: EXPONENTIAL MOMENTS OF CONTINUOUS-TIME MARTINGALES

Let $(M_t)_{t\ge 0}$ be a martingale defined as a function of a continuous time Markov chain on a finite state space

$$M_t = f(t, X_t),$$

where f is differentiable in time. We let B denote the maximal jump rate for X and let $S(t)$ denote the maximal amplitude for a jump of M at time t:

$$S(t) := \max_{\xi \sim \xi'} \big| f(t,\xi) - f(t,\xi') \big|.$$

LEMMA 31. *For any $\lambda > 0$ we have*

$$\mathbb{E}\big[e^{\lambda M_t}\big] \le \exp\bigg(B \int_0^t \big[e^{\lambda S(s)} - \lambda S(s) - 1\big]\,\mathrm{d}s \bigg). \tag{150}$$

In particular if $\lambda S(t) \le 1$ for all $t \ge 0$ then we have

$$\mathbb{E}\big[e^{\lambda M_t}\big] \le \exp\bigg(Be\lambda^2 \int_0^t S^2(s)\,\mathrm{d}s \bigg). \tag{151}$$

PROOF. We are going to show that for all $t \ge 0$

$$\partial_t \log \mathbb{E}\big[e^{\lambda M_t}\big] = \frac{\partial_t \mathbb{E}[e^{\lambda M_t}]}{\mathbb{E}[e^{\lambda M_t}]} \le B\big[e^{\lambda S(t)} - \lambda S(t) - 1\big]. \tag{152}$$

To that end, it is sufficient to show that almost surely

$$\partial_s \mathbb{E}\big[e^{\lambda(M_{t+s} - M_t)} - 1 | \mathcal{F}_t\big]|_{s=0} \le B\big[e^{\lambda S(t)} - \lambda S(t) - 1\big]. \tag{153}$$

Indeed (152) is obtained by multiplying both sides of (153) by $e^{\lambda M_t}$ and by taking expectations. We let $\Delta_s M = M_{t+s} - M_t$ denote the martingale increment and as in the previous section write $\mathbb{E}_t$ for the conditional expectation w.r.t. M_t. By the martingale property, and the fact that $e^x - x \le e^{|x|} - |x|$ for every $x \in \mathbb{R}$, we have

$$\mathbb{E}_t\big[e^{\lambda \Delta_s M} - 1\big] = \mathbb{E}_t\big[e^{\lambda \Delta_s M} - \lambda \Delta_s M - 1\big] \le \mathbb{E}_t\big[e^{\lambda |\Delta_s M|} - \lambda |\Delta_s M| - 1\big].$$

Note that $|\Delta_s M|$ is stochastically dominated by

$$\Big[\max_{u \in [t,t+s]} S(u)\Big]\mathcal{W} + s \times \max_{u \in [t,t+s]} \big\| \partial_u f(u,\cdot) \big\|_\infty,$$

where $\mathcal{W}$ is a Poisson variable of parameter Bs. As S is Lipshitz we conclude that

$$\mathbb{E}_t\bigg[\frac{e^{\lambda|\Delta_s M|} - \lambda|\Delta_s M| - 1}{s}\bigg] \le B\big[e^{\lambda S(t)} - \lambda S(t) - 1\big] + cs. \tag{154}$$

□

Acknowledgments. The authors are grateful to Shangjie Yang for help on the manuscript. This work has been partially realized during a visit of H. L. to CEREMADE, he acknowledges kind hospitality and support. H. L. also acknowledges support from a productivity grant of CNPq and a JCNE grant from FAPERj. C. L. acknowledges support from the grant SINGULAR ANR-16-CE40-0020-01.

REFERENCES

[1] ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348. MR0841111 https://doi.org/10.2307/2323590

[2] 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

[3] 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

[4] DE MASI, A., PRESUTTI, E. and SCACCIATELLI, E. (1989). The weakly asymmetric simple exclusion process. *Ann. Inst. Henri Poincaré Probab. Stat.* **25** 1–38. MR0995290

[5] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR0626813 https://doi.org/10.1007/BF00535487

[6] DITTRICH, P. and GÄRTNER, J. (1991). A central limit theorem for the weakly asymmetric simple exclusion process. *Math. Nachr.* **151** 75–93. MR1121199 https://doi.org/10.1002/mana.19911510107

[7] GÄRTNER, J. (1988). Convergence towards Burgers' equation and propagation of chaos for weakly asymmetric exclusion processes. *Stochastic Process. Appl.* **27** 233–260. MR0931030 https://doi.org/10.1016/0304-4149(87)90040-8

[8] KIPNIS, C. and LANDIM, C. (1999). *Scaling Limits of Interacting Particle Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **320**. Springer, Berlin. MR1707314 https://doi.org/10.1007/978-3-662-03752-2

[9] KIPNIS, C., OLLA, S. and VARADHAN, S. R. S. (1989). Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.* **42** 115–137. MR0978701 https://doi.org/10.1002/cpa.3160420202

[10] LABBÉ, C. (2017). Weakly asymmetric bridges and the KPZ equation. *Comm. Math. Phys.* **353** 1261–1298. MR3652491 https://doi.org/10.1007/s00220-017-2875-0

[11] 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

[12] 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

[13] 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

[14] LEVIN, D. A. and PERES, Y. (2016). Mixing of the exclusion process with small bias. *J. Stat. Phys.* **165** 1036–1050. MR3575636 https://doi.org/10.1007/s10955-016-1664-z

[15] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. MR3726904

[16] LIGGETT, T. M. (2005). *Interacting Particle Systems*. *Classics in Mathematics*. Springer, Berlin. Reprint of the 1985 original. MR2108619 https://doi.org/10.1007/b138374

[17] MORRIS, B. (2006). The mixing time for simple exclusion. *Ann. Appl. Probab.* **16** 615–635. MR2244427 https://doi.org/10.1214/105051605000000728

[18] REZAKHANLOU, F. (1991). Hydrodynamic limit for attractive particle systems on $\mathbf{Z}^d$. *Comm. Math. Phys.* **140** 417–448. MR1130693

[19] 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