

MIXING TIME OF THE ADJACENT WALK ON THE SIMPLEX

BY PIETRO CAPUTO¹, CYRIL LABBÉ² AND HUBERT LACOIN³

¹*Department of Mathematics and Physics, Roma Tre University, caputo@mat.uniroma3.it*

²*Université Paris-Dauphine, PSL University, Ceremade, CNRS, labbe@ceremade.dauphine.fr*

³*IMPA, lacoin@impa.br*

By viewing the N -simplex as the set of positions of $N - 1$ ordered particles on the unit interval, the adjacent walk is the continuous-time Markov chain obtained by updating independently at rate 1 the position of each particle with a sample from the uniform distribution over the interval given by the two particles adjacent to it. We determine its spectral gap and prove that both the total variation distance and the separation distance to the uniform distribution exhibit a cutoff phenomenon, with mixing times that differ by a factor 2. The results are extended to the family of log-concave distributions obtained by replacing the uniform sampling by a symmetric log-concave Beta distribution.

1. Introduction. Randomized algorithms based on Markov chains are commonly used for sampling points uniformly at random in a convex body and to simulate other log-concave distributions in N -dimensional Euclidean space. The mixing time of the associated random walk in \mathbb{R}^N is often known to be polynomial in N ; see, for example, [6, 16]. In analogy with the setting of Markov chains with discrete state space, where the theory seems to be more advanced (see, e.g., the monographs [2, 14]), it is of interest to develop a finer analysis of the asymptotic growth of mixing times in high dimensions.

In this paper, we address that question for a specific model, namely the adjacent walk on the N -simplex. By viewing the N -simplex as the set of positions of $N - 1$ ordered particles on the unit interval, this is the process obtained by updating independently at rate 1 the position of each particle with a sample from the uniform distribution over the interval given by the two particles adjacent to it. The process defines a Gibbs sampler for the uniform distribution over the simplex.

The adjacent walk on the N -simplex has been previously analysed in [22], where the authors proved upper and lower bounds on the mixing time that are tight up to constant factors; see also [21, 23, 24] for similar estimates in related models. A version of this model with open boundary conditions was introduced in [11] to study the heat flow in a chain of one-dimensional oscillators.

Here we determine the mixing time to leading order and we establish the so-called cutoff phenomenon for the adjacent walk on the N -simplex. We also show that the same results hold if the uniform distribution over the sampling interval is replaced by a symmetric log-concave Beta law, in which case the chain converges to a log-concave distribution over the simplex.

1.1. Model and results. It is convenient to replace the unit interval by the interval $[0, N]$, that is to rescale the N -simplex to the set Ω_N of positions of $N - 1$ ordered particles on the interval $[0, N]$ defined by

$$(1) \quad \Omega_N := \{x = (x_1, \dots, x_{N-1}) \in \mathbb{R}^{N-1} : 0 \leq x_1 \leq \dots \leq x_{N-1} \leq N\}.$$

Received April 2019; revised January 2020.

MSC2020 subject classifications. Primary 60J25; secondary 37A25, 82C22.

Key words and phrases. Spectral gap, mixing time, cutoff, adjacent walk.

We also set $x_0 = 0, x_N = N$. An element x in Ω_N can be viewed as a nondecreasing interface $k \mapsto x_k$ connecting $(0, 0)$ to (N, N) : this viewpoint will be adopted throughout the article. Given an initial condition $x \in \Omega_N$, we consider the process

$$\mathbf{X}^x(t) = (X_1(t), \dots, X_{N-1}(t)),$$

where $\mathbf{X}(0) = x$, and independently for each i , at rate 1, X_i is resampled according to the uniform distribution on $[X_{i-1}, X_{i+1}]$, where we use the convention $X_0 := 0$ and $X_N := N$.

More formally, this is the continuous-time Markov chain with generator

$$(2) \quad (\mathcal{L}_N f)(x) = \sum_{i=1}^{N-1} \int_0^1 (f(x^{(i,u)}) - f(x)) du,$$

where, given $u \in [0, 1]$ and i , the transformation $x \mapsto x^{(i,u)}$ is defined by

$$x_j^{(i,u)} = \begin{cases} x_j & i \neq j, \\ ux_{i-1} + (1-u)x_{i+1} & i = j. \end{cases}$$

Let π_N denote the uniform distribution over Ω_N . Noting that $\int_0^1 f(x^{(i,u)}) du$ coincides with the conditional expectation $\pi_N[f \mid x_j, j \neq i]$ shows that the generator is a finite linear combination of orthogonal projections in $L^2(\Omega_N, \pi_N)$. In particular, \mathcal{L}_N is a bounded self adjoint operator, and the Markov chain is reversible with respect to π_N . Note that π_N can be obtained by conditioning N i.i.d. exponential random variables of parameter 1 to have a sum equal to N , and by taking the x_i 's to be their partial sums.

The generator is nonpositive definite and has the trivial eigenvalue zero associated to constant functions. Our first result is a characterization of the spectral gap, defined as the smallest nontrivial eigenvalue of $-\mathcal{L}_N$.

PROPOSITION 1. *For any $N \geq 2$, the spectral gap of the generator is given by*

$$(3) \quad \text{gap}_N = 1 - \cos\left(\frac{\pi}{N}\right),$$

and the corresponding eigenfunction is

$$(4) \quad f_N(x) = \sum_{k=1}^{N-1} \sin\left(\frac{\pi k}{N}\right)(x_k - k).$$

We are interested in the time needed for the total variation distance to equilibrium, starting from the “worst” initial condition, to pass below some given threshold $\varepsilon \in (0, 1)$. When this time is, to leading order in N , insensitive to the choice of ε , one speaks of a cutoff phenomenon. More precisely, we set

$$(5) \quad d_N(t) := \sup_{x \in \Omega_N} \|P_t^x - \pi_N\|_{\text{TV}},$$

where P_t^x is the law of $\mathbf{X}^x(t)$, and for μ, ν probability measures on Ω_N , the total variation distance is given by

$$\|\mu - \nu\|_{\text{TV}} = \sup_{B \in \mathcal{B}(\Omega_N)} (\mu(B) - \nu(B)),$$

the supremum ranging over all Borel subsets of Ω_N . For any $\varepsilon \in (0, 1)$, the mixing time is defined by

$$T_N(\varepsilon) := \inf\{t \geq 0 : d_N(t) < \varepsilon\}.$$

Our main result is the following.

THEOREM 1. For any $\varepsilon \in (0, 1)$,

$$\lim_{N \rightarrow \infty} \frac{T_N(\varepsilon)}{N^2 \log N} = \frac{1}{\pi^2}.$$

As a consequence, the sequence of Markov chains displays a cutoff phenomenon.

In view of Proposition 1, the above theorem can be restated as

$$T_N(\varepsilon) \stackrel{N \rightarrow \infty}{\sim} \frac{\log N}{2 \text{gap}_N}.$$

Remark that if f_N is as in (4) and we start with the extremal initial condition $x_i \equiv N$, then $t = \frac{\log N}{2 \text{gap}_N}$ is exactly the time it takes for the expected value

$$\mathbb{E}[f_N(\mathbf{X}^x(t))] = f_N(x)e^{-\text{gap}_N t},$$

to drop from the initial value $f_N(x) = \Theta(N^2)$ to a value $\Theta(N^{3/2})$, which is the typical size of fluctuations of f_N at equilibrium.

Strikingly, Proposition 1 and Theorem 1 take the exact same form in the case of the symmetric simple exclusion process on the N -segment; see [13, 25].

1.2. *Generalization to Beta-resampling.* Given $\alpha \in (0, \infty)$, the symmetric Beta distribution of parameter α is the probability measure on $[0, 1]$ with density

$$(6) \quad \rho_\alpha(u) := \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} [u(1-u)]^{\alpha-1}.$$

We define a generalization of the process described in (2) by resampling points according to a Beta distribution

$$(7) \quad (\mathcal{L}_{N,\alpha} f)(x) = \sum_{i=1}^{N-1} \int_0^1 (f(x^{(i,u)}) - f(x)) \rho_\alpha(u) du.$$

While $\rho_\alpha(u) du$ could be replaced by any probability on $[0, 1]$, the Beta laws are the only resampling laws that make the dynamics reversible with respect to probability measures with a product structure; see Remark 6 below. The associated equilibrium distribution is given by

$$(8) \quad \pi_{N,\alpha}(dx) := \frac{\Gamma(N\alpha)}{\Gamma(\alpha)^N N^{N\alpha-1}} \prod_{i=1}^N (x_i - x_{i-1})^{\alpha-1} dx,$$

where dx is Lebesgue’s measure on Ω_N . Since $\int_0^1 f(x^{(i,u)}) \rho_\alpha(u) du$ equals the conditional expectation $\pi_{N,\alpha}[f \mid x_j, j \neq i]$, it follows that $\mathcal{L}_{N,\alpha}$ is a bounded self adjoint operator in $L^2(\Omega_N, \pi_{N,\alpha})$. The following theorem extends the results of Proposition 1 and Theorem 1.

THEOREM 2. For any $\alpha \geq 1$, the spectral gap and the corresponding eigenfunction are given by

$$\text{gap}_N = 1 - \cos\left(\frac{\pi}{N}\right), \quad f_N(x) = \sum_{k=1}^{N-1} \sin\left(\frac{\pi k}{N}\right) (x_k - k).$$

Moreover, for any $\varepsilon \in (0, 1)$ the ε -mixing time associated with the Beta-resampling process satisfies

$$\lim_{N \rightarrow \infty} \frac{T_{N,\alpha}(\varepsilon)}{N^2 \log N} = \frac{1}{\pi^2}.$$

In particular, the sequence of Markov chains displays a cutoff phenomenon.

REMARK 2. We have chosen $x_N := N$ so that average inter-particle spacing at equilibrium is one. However, this convention has no influence on the result and we may take as well $x_N = 1$ or any other constant (the only effect of this change being a dilation or contraction of space). In the course of the proof it is sometimes convenient to consider also the case of a random x_N (since the process leaves x_N fixed this makes the dynamics nonirreducible in that case).

REMARK 3. The restriction $\alpha \geq 1$ is due to the fact that certain parts of our proof require log-concavity of the probability density ρ_α , which in turn implies the validity of the FKG property for the equilibrium measure $\pi_{N,\alpha}$, see Section 2.2. However, we believe the result to be valid for all $\alpha > 0$.

REMARK 4. The spectrum of $\mathcal{L}_{N,\alpha}$ restricted to the invariant subspace consisting of linear functions can be computed explicitly (see Section 2.3 below) and the fact that $\text{gap}_{N,\alpha} = 1 - \cos(\frac{\pi}{N})$ is equivalent to the statement that the spectral gap is attained within this subspace. As the following mean field example shows, this phenomenon is not always to be expected for the exchange dynamics obtained by replacing the segment with another graph. Consider for instance the exchange process on the complete graph with generator

$$(9) \quad (\mathcal{G}_{N,\alpha} f)(\eta) = \frac{1}{N} \sum_{1 \leq i < j \leq N} \int_0^1 (f(\eta^{(i,j,u)}) - f(\eta)) \rho_\alpha(u) du,$$

where

$$\eta_k^{(i,j,u)} = \begin{cases} \eta_k & k \neq i, j, \\ u(\eta_i + \eta_j) & k = i, \\ (1-u)(\eta_i + \eta_j) & k = j, \end{cases}$$

and $\eta = \{\eta_k\}$ denotes the increment variables $\eta_k = x_k - x_{k-1}$, $k = 1, \dots, N$. In other words, $\mathcal{G}_{N,\alpha}$ defines the mean field version of (7). Notice that $\mathcal{G}_{N,\alpha}$ is reversible w.r.t. $\pi_{N,\alpha}$. Using the arguments in [4] or [3] one can show that for each $N \geq 2$, $\alpha > 0$ the spectral gap of $\mathcal{G}_{N,\alpha}$ satisfies

$$(10) \quad \text{gap}(\mathcal{G}_{N,\alpha}) = \frac{\alpha N + 1}{(2\alpha + 1)N},$$

and that the corresponding eigenfunction is the symmetric quadratic function

$$g_N(\eta) = \text{const.} + \sum_{i=1}^N \eta_i^2.$$

1.3. *Mixing time for the separation distance.* Another distance for which mixing time can be considered is the separation distance (see [8]). Its use in the study of mixing times for Markovian systems was initially advocated by Aldous and Diaconis because of its appealing relation to strong-stationary times [1]. While, to our knowledge, separation distance has been considered so far only for countable state spaces, it can be generalized to a continuous setup via the definition

$$(11) \quad d_{N,\alpha}^{\text{sep}}(t) := 1 - \inf_{\substack{A \in \mathcal{B}(\Omega_N) \\ B \in \mathcal{B}(\Omega_N)}} \frac{P_t(A, B)}{\pi_{N,\alpha}(A)\pi_{N,\alpha}(B)},$$

where

$$P_t(A, B) := \int_A P_t^x(B) \pi_{N,\alpha}(dx),$$

and the infimum ranges over all Borel subsets of Ω_N of positive Lebesgue measure. We can then define the separation mixing time associated with our chain as

$$T_{N,\alpha}^{\text{sep}}(\varepsilon) := \inf\{t \geq 0 : d_{N,\alpha}^{\text{sep}}(t) < \varepsilon\}.$$

The separation mixing time is always larger than or equal to its total variation counterpart. It is known (cf. [14], Lemma 19.3, in the discrete setup, and Section 4 below for an adaptation of the argument to the continuous case) that in the case of reversible chains, the separation mixing time is at most twice as large as that in total variation. Because of this factor 2, cutoff in total variation does not necessarily imply that the same phenomenon holds for the separation distance (see [9] for counter-examples)

We prove nonetheless that our dynamics displays a cutoff phenomenon for the separation distance and that the associated mixing times are twice as large as those for the total variation distance.

THEOREM 3. *For all $\alpha \geq 1$ and every $\varepsilon \in (0, 1)$, the ε -mixing time for the separation distance satisfies*

$$\lim_{N \rightarrow \infty} \frac{T_{N,\alpha}^{\text{sep}}(\varepsilon)}{N^2 \log N} = \frac{2}{\pi^2}.$$

Here again, our result is analogous to the one obtained for the symmetric simple exclusion process [13, 25].

REMARK 5. Note that another natural definition of the separation distance in the continuous setup would be

$$(12) \quad d'_{N,\alpha}(t) := 1 - \inf_{\substack{x \in \Omega_N \\ B \in \mathcal{B}(\Omega_N)}} \frac{P_t(x, B)}{\pi_{N,\alpha}(B)}.$$

Using explicit bounds on the total-mass of the singular part of the law at time t , see the proof of Lemma 38, one can prove that the two distances have a cutoff phenomenon at the same time. Because of the regularity of our setting, we strongly believe that $d'_{N,\alpha}(t) = d_{N,\alpha}^{\text{sep}}(t)$, however this is certainly not true for every reversible Markov chain.

1.4. Comments on the proof and related work. The cutoff phenomenon is a widely studied topic in Markov chain theory; see [2, 5, 14] for an introduction. Thanks to recent remarkable efforts, many interesting examples are known of Markov chains exhibiting this particular type of phase transition. Unfortunately, these seem to be mostly confined to the setting of Markov chains with discrete state space; see, however, [10] for a recent exception. One of the reasons is possibly the fact that the analysis of total variation distance in a continuous state space is more demanding.

Let us briefly describe the main arguments used to establish the results of this article. Concerning the spectrum, the first observation is that when restricted to the invariant subspace of linear functions, the generator can be easily diagonalised. In particular, one finds that f_N is an eigenfunction of $\mathcal{L}_{N,\alpha}$ with eigenvalue $\cos(\pi/N) - 1$ and therefore, for all $\alpha > 0$, the spectral gap is at most $1 - \cos(\pi/N)$. Proving that this is actually the spectral gap is more delicate. We prove it only in the case $\alpha \geq 1$ since we rely on the FKG inequality. In this case, using the strict monotonicity of f_N , we show that any eigenfunction associated to an eigenvalue larger than $\cos(\pi/N) - 1$ needs to be nondecreasing as well, and the FKG inequality is then applied to conclude that the only such eigenfunction is the constant one; see Section 2.3 below. Concerning the case $\alpha \in (0, 1)$, we can only assert that the spectral gap is

not smaller than cN^{-2} for some constant $c = c(\alpha) > 0$. Indeed, this is a direct consequence of the mean field spectral gap (10) and a comparison argument (see, e.g., [3]).

To establish the lower bound on the mixing time of Theorem 2, we use Wilson’s method [25]. For $\alpha = 1$, this was already used in [22] to obtain the weaker lower bound

$$T_N(1/4) \geq cN^2 \log N,$$

for some constant $c > 0$. Here we sharpen this, and obtain for any $\alpha > 0$ and for all $\varepsilon > 0$

$$T_{N,\alpha}(\varepsilon) \geq \frac{1}{2 \operatorname{gap}_N} (\log N - C_{\alpha,\varepsilon}),$$

see Proposition 11 below: note that we prove this result also for $\alpha \in (0, 1)$, but in this case gap_N is not defined as the spectral gap but simply as being $1 - \cos(\pi/N)$. This suggests a cutoff window of order N^2 , but we do not have a corresponding upper bound. The proof of the above lower bound is based on Wilson’s method with a careful choice of the initial condition x and a comparison argument which allows us to control the variance of $f_N(\mathbf{X}^x(t))$; see Section 3.

The proof of the upper bound on the mixing time of Theorem 2 is the most involved part of the paper, and is worked out in Section 5 and Section 6. Our strategy can be roughly outlined as follows; see Section 5.1 below for a more detailed overview. In [22], and for $\alpha = 1$, it was shown that

$$(13) \quad T_{N,\alpha}(1/4) \leq CN^2 \log N,$$

for some constant $C > 0$. This upper bound was obtained by estimating, under some monotone grand coupling, the hitting time of 0 for the area comprised between the two extremal processes, that is, the processes starting from the highest ($x_1 = \dots = x_{N-1} = N$) and the lowest ($x_1 = \dots = x_{N-1} = 0$) initial conditions. Note that under such a coupling this hitting time bounds from above the coalescing time starting from any two initial conditions. The proof consisted of two steps that we now briefly recall. First, using the decay of the heat equation solved by the expectation of the area, one can show that after a time $C(\delta)N^2 \log N$, the area lies below $N^{-\delta}$ with large probability. Choosing δ large enough, the second step relies on a brute force argument that shrinks the area to 0 in a time of order $\log N$ with large probability. In the Appendix, we give a proof of (13) in the general case $\alpha > 0$: in contrast with the case $\alpha = 1$, there is no simple monotone grand coupling for a general α that achieves an efficient coalescing time of the two extremal processes. We are then lead to controlling the coalescing time of the stationary process and the process starting from some arbitrary initial condition.

The constant C of (13) is dictated by the smallest value $C(\delta)$ that one can take in the aforementioned strategy: since δ needs to be very large for the second step to apply, this value is much larger than desired. Our main contribution consists in introducing a sequence of intermediate steps that reduce the time necessary to bring the area to a small enough threshold from which a brute force argument can be applied. Namely, we use the first step above up to the target mixing time, which guarantees a shrinking of the expectation of the area down to $N^{3/2}$ (corresponding to $\delta = -3/2$), and then present a coupling under which, through a sequence of successive stages, we are able to bring the area from $N^{3/2}$ down to N^{-1} within a time of order N^2 with large probability, that is, a time negligible compared to the target mixing time. The last step is then a (slightly different) brute force argument that shrinks the area to 0 with large probability within a time of order $\log N$. This program is carried out in Section 5.

At a technical level, we use estimates on the derivative of the predictable bracket process associated to the evolution of the area together with diffusive estimates on supermartingales in order to upper bound the time needed for the area to descend through the intermediate

values mentioned above. A key ingredient of these hitting time estimates, is the control of the gradients of the particle positions for the extremal processes, that is the process with maximal and minimal initial particle positions. The control of the gradients, in turn, is obtained by coupling the extremal processes with the stationary distribution. To this end we need to establish, with an independent argument, a sharp upper bound on the time needed for convergence to equilibrium of the two extremal processes.

The sharp upper bound for the extremal processes (by symmetry, we may consider only the highest process) is performed in Section 6. This part of the proof is based on a strategy developed in the discrete setting by [13] for the adjacent transposition process. It relies on the FKG inequality and a version of the censoring inequality from [18]. Censoring inequalities are known to hold provided one has monotonicity of the density of the initial condition with respect to the invariant measure. While in the discrete setting this adapts well to the extremal process, in our context a nontrivial adaptation is required since the initial condition is singular with respect to the invariant measure.

2. Preliminaries. Let us consider a larger state-space where there is no constraint on x_N

$$(14) \quad \Omega_N^+ := \{x = (x_1, \dots, x_N) \in \mathbb{R}^N : 0 \leq x_1 \leq x_2 \leq \dots \leq x_N\}.$$

The generator $\mathcal{L}_{N,\alpha}$ defines a dynamics also on this enlarged space. If we let

$$(\eta_i(x))_{i=1}^N := (x_i - x_{i-1})_{i=1}^N$$

denote the increments of x , one can easily check that the distribution under which the increments η_i are i.i.d. random variables with distribution $\Gamma(\alpha, \lambda)$, $\lambda > 0$ being an arbitrary parameter, is a reversible measure for the dynamics. Here $\Gamma(\alpha, \lambda)$ stands for the Gamma distribution with mean α/λ and variance α/λ^2 .

REMARK 6. We remark that the Beta resampling rule (6) is a natural choice in our context. Indeed, suppose we define a generator as in (7) with $\rho_\alpha(u) du$ replaced by another probability $\nu(du)$ on the unit interval. Then one can check that if μ is a probability measure on Ω_N^+ under which the increments η_i are independent, the generator is reversible with respect to μ if and only if μ is the product of $\Gamma(\alpha, \lambda)$, and $\nu(du) = \rho_\alpha(u) du$, for some $\alpha, \lambda > 0$. This rests on the well-known fact that if X, Y are two independent random variables such that $X + Y$ and $X/(X + Y)$ are independent, then X and Y must have the gamma distribution, see [17].

2.1. Monotone grand coupling. There are two natural orders on Ω_N^+ associated with the dynamics. The coordinate order

$$(15) \quad x \geq x' \Leftrightarrow \{\forall i \in \llbracket 1, N \rrbracket, x_i \geq x'_i\},$$

and the gradient order

$$(16) \quad x \succcurlyeq x' \Leftrightarrow \{\forall i \in \llbracket 1, N \rrbracket, \eta_i(x) \geq \eta'_i(x')\}.$$

Here and below we use the notation $\llbracket i, j \rrbracket$ for the integer interval $[i, j] \cap \mathbb{Z}$. Note that the coordinate order is natural in both Ω_N and Ω_N^+ , while the gradient order is only relevant for the unconstrained space Ω_N^+ . An important observation is that the Beta resampling dynamics preserves both orders in the following sense.

PROPOSITION 7 (Existence of a grand coupling). *For any $\alpha > 0$, and $N \geq 2$, we can construct the trajectories $(\mathbf{X}^x(t))_{t \geq 0}$ of the Markov chain on Ω_N^+ with generator $\mathcal{L}_{N,\alpha}$ on the same probability space in such a way that \mathbb{P} -a.s., for all $x, x' \in \Omega_N^+$*

$$(17) \quad \begin{aligned} x \geq x' &\Leftrightarrow \{\forall t \geq 0, \mathbf{X}^x(t) \geq \mathbf{X}^{x'}(t)\}, \\ x \succcurlyeq x' &\Leftrightarrow \{\forall t \geq 0, \mathbf{X}^x(t) \succcurlyeq \mathbf{X}^{x'}(t)\}. \end{aligned}$$

PROOF. The coupling invoked above is the usual graphical construction (see, e.g., [15]). To each $k \in \llbracket 1, N - 1 \rrbracket$ we associate a Poisson clock process $(\mathcal{T}_i^{(k)})_{i \geq 1}$ whose increments are i.i.d. rate one exponentials, and a sequence $(U_i^{(k)})_{i \geq 1}$ of i.i.d. symmetric Beta variables of parameter α . Then, for all $x \in \Omega_N^+$, $(\mathbf{X}^x(t))_{t \geq 0}$ is chosen to be càd-làg and constant outside of the update times $(\mathcal{T}_i^{(k)})_{k \in \llbracket 1, N - 1 \rrbracket, i \geq 1}$. At time $t = \mathcal{T}_i^{(k)}$, if $U_i^{(k)} = u$ the k th coordinate is updated as follows

$$X_k^x(t) = uX_{k-1}^x(t_-) + (1 - u)X_{k+1}^x(t_-) \quad \text{and} \quad X_j^x(t) := X_j^x(t_-) \quad \text{for } j \neq k.$$

One then checks by inspection that the process generated in this manner is the desired Markov chain. Moreover, the above construction implies that it preserves the two above mentioned orders in the sense of (17). \square

Let us finally introduce the maximal \wedge and minimal \vee configurations for the coordinate order \geq that we defined above:

$$(18) \quad \wedge_k = N, \quad \vee_k = 0 \quad \forall k \in \llbracket 1, N - 1 \rrbracket.$$

2.2. *The FKG inequality.* When $\alpha \geq 1$, the equilibrium measure $\pi_{N,\alpha}$ can be interpreted as a one dimensional gradient field associated with a convex potential. More precisely, we can write

$$\pi_{N,\alpha}(dx) := C(N, \alpha) \exp\left(-\sum_{k=1}^N V(x_k - x_{k-1})\right) dx,$$

where V is the convex potential

$$(19) \quad V(u) = \begin{cases} +\infty & \text{if } u \leq 0, \\ -(\alpha - 1) \log u & \text{if } u > 0. \end{cases}$$

Below we use a standard application of Holley’s criterion (see [20]) to show that $\pi_{N,\alpha}$ satisfies the so-called FKG inequality which entails positive correlation between increasing functions.

We say that f defined on Ω_N is increasing if

$$(20) \quad \forall x, x' \in \Omega_N, \quad x \geq x' \quad \Rightarrow \quad f(x) \geq f(x').$$

A subset $A \subset \Omega_N$ is said to be increasing if $\mathbf{1}_A$ is an increasing function. We let \wedge and \vee denote the following operations in Ω_N :

$$(21) \quad (x \vee x') := (\max(x_i, x'_i))_{i=1}^{N-1} \quad \text{and} \quad (x \wedge x') := (\min(x_i, x'_i))_{i=1}^{N-1}.$$

(Note that there is a little clash of notation with the maximal and minimal configurations introduced in (18): however, we believe that this will never raise any confusion in the sequel.)

PROPOSITION 8 (FKG inequality). *For any $\alpha \geq 1$ and $N \geq 2$, if f and g are increasing then*

$$\pi_{N,\alpha}(fg) \geq \pi_{N,\alpha}(f)\pi_{N,\alpha}(g).$$

Moreover, the inequality remains valid if $\pi_{N,\alpha}$ is replaced by its restriction $\pi_{N,\alpha}(\cdot | A)$ to any set A which is stable under the operations \vee and \wedge .

PROOF. Setting $H(x) := \sum_{k=1}^N V(x_k - x_{k-1})$, the density of $\pi_{N,\alpha}$ with respect to Lebesgue is given by $e^{-H(x)}$ and Holley’s criterion shows that the inequality holds provided that

$$(22) \quad H(x \vee x') + H(x \wedge x') \leq H(x) + H(x').$$

The inequality (22) can be deduced as a consequence of the convexity of V see [7], Appendix B1. For the measure restricted to a set A it is sufficient to check that (22) is valid when H is replaced by

$$(23) \quad H_A(x) := \begin{cases} \sum_{k=1}^N V(x_i - x_{i-1}) & \text{if } x \in A, \\ +\infty & \text{if } x \in A^c, \end{cases}$$

which is an immediate consequence of (22) under our assumption on A . \square

Let us point out that the first inequality of Proposition 8 could be obtained directly from the monotonicity stated in Proposition 7, see, for instance, [14], Th 22.16, and therefore does not require the convexity of the potential (that is, it holds also for $\alpha \in (0, 1)$).

For two probability measures μ, ν on Ω_N , we write $\mu \leq \nu$ and say that μ is stochastically dominated by ν if for all increasing f (in the sense of (20)) one has $\mu(f) \leq \nu(f)$. Let us mention another application of Holley’s criterion which we are going to use in the proof.

LEMMA 9. *Let A and B be two increasing subsets of Ω_N such that*

$$\{x \in A \text{ and } x' \in B\} \Rightarrow x \wedge x' \in B.$$

Then

$$\pi_{N,\alpha}(\cdot | A) \geq \pi_{N,\alpha}(\cdot | B).$$

PROOF. Let $\mu_A := \pi_{N,\alpha}(\cdot | A)$ and $\mu_B := \pi_{N,\alpha}(\cdot | B)$. These probability measures have density proportional to $\exp(-H_A(x))$ and $\exp(-H_B(x))$ respectively, where the potentials H_A, H_B are defined as in (23). The result then directly follows from [20], Proposition 1, if one can show that for every $x, x' \in \Omega_N$

$$(24) \quad H_A(x \vee x') + H_B(x \wedge x') \leq H_A(x) + H_B(x').$$

This in turn follows from the inequality (22) for H and our assumption on A, B . \square

2.3. Identification of the spectral gap. Here we prove the first statement of Theorem 2. Fix $\alpha \geq 1$ and $N \geq 2$, and write \mathcal{L} for the generator $\mathcal{L}_{N,\alpha}$. Using the expression (7), the action of the generator on the coordinate map $h_k(x) := x_k, k = 1, \dots, N - 1$, is given by

$$(25) \quad (\mathcal{L}h_k)(x) = \frac{1}{2} \Delta h_k(x),$$

where Δ denotes the discrete Laplace operator

$$\Delta x_k := x_{k+1} - 2x_k + x_{k-1}.$$

Summation by parts and (25) then shows that for every $j \in \llbracket 1, N - 1 \rrbracket$ the map

$$(26) \quad f_N^{(j)}(x) := \sum_{k=1}^{N-1} \sin\left(\frac{j\pi k}{N}\right)(x_k - k),$$

is an eigenfunction of \mathcal{L} with the eigenvalue $-\lambda_N^{(j)}$ where

$$\lambda_N^{(j)} := 1 - \cos\left(\frac{j\pi}{N}\right).$$

In the case $j = 1$, we simply write f_N for $f_N^{(1)}$ and λ_N for $\lambda_N^{(1)}$. We now turn to show that λ_N is actually the spectral gap.

It is not hard to check that for any $n \geq 1$, the set of all polynomials of degree at most n in the variables x_1, \dots, x_{N-1} is stable under the action of the generator. When restricted to any such set, the generator admits a finite complete decomposition into eigenvalues / eigenfunctions. By density of polynomials in $L^2(\Omega_N, \pi_{N,\alpha})$, there exists an orthonormal basis of polynomial eigenfunctions and therefore the generator has pure point spectrum in $L^2(\Omega_N, \pi_{N,\alpha})$.

\mathcal{L} maps to zero all constant functions and any nontrivial eigenvalue of \mathcal{L} must be associated with an eigenfunction with mean zero. Therefore, it is sufficient to show that if g is a normalized polynomial eigenfunction such that $\pi_{N,\alpha}(g) = 0$ and $\mathcal{L}g = -\mu g$, with $\mu < \lambda_N$, then $g = 0$. Since g is polynomial, and since f_N is strictly increasing in all its variables, there exists $\epsilon > 0$ such that $f_N + \epsilon g$ remains increasing in all its variables. Next, we define the following normalized function:

$$v_t := \frac{P_t(f_N + \epsilon g)}{\|P_t(f_N + \epsilon g)\|_2},$$

where $P_t = e^{t\mathcal{L}}$ denotes the semigroup generated by \mathcal{L} , and $\|\cdot\|_2$ is the $L^2(\Omega_N, \pi_{N,\alpha})$ -norm. From our assumptions, one has $\pi_{N,\alpha}(f_N g) = 0$ and

$$P_t(f_N + \epsilon g) = e^{-\lambda_N t} f_N + \epsilon e^{-\mu t} g,$$

with $\mu < \lambda_N$. It follows that $v_t \rightarrow g$ as $t \rightarrow \infty$ in $L^2(\Omega_N, \pi_{N,\alpha})$. On the other hand, the semigroup preserves monotonicity by Proposition 7, so that v_t is a nondecreasing function at any time t . Thus, g must be also nondecreasing. Notice that so far the argument is valid for any $\alpha > 0$. We shall now use the assumption $\alpha \geq 1$.

The FKG inequality and the orthogonality of f_N and g imply that the centered coordinate maps $\bar{h}_k(x) := x_k - k$ satisfy

$$\pi_{N,\alpha}(\bar{h}_k g) = 0, \quad k \in \llbracket 1, N - 1 \rrbracket.$$

Indeed, by Proposition 8 one has $\pi_{N,\alpha}(\bar{h}_k g) \geq 0$ and if this is positive for some k , then $\pi_{N,\alpha}(f_N g) > 0$.

Let A_ϵ be the event $h_1 \geq N - \epsilon$, with $\epsilon \in (0, 1)$. Since both A_ϵ and A_ϵ^c are stable for the operations \wedge and \vee introduced in (21), by Proposition 8 the restrictions $\pi_{N,\alpha}(\cdot | A_\epsilon)$ and $\pi_{N,\alpha}(\cdot | A_\epsilon^c)$ satisfy the FKG inequality. Therefore,

$$(27) \quad \begin{aligned} 0 &= \pi_{N,\alpha}(\bar{h}_1 g) = \pi_{N,\alpha}(A_\epsilon) \pi_{N,\alpha}(\bar{h}_1 g | A_\epsilon) + \pi_{N,\alpha}(A_\epsilon^c) \pi_{N,\alpha}(\bar{h}_1 g | A_\epsilon^c) \\ &\geq \pi_{N,\alpha}(\bar{h}_1 \mathbf{1}_{A_\epsilon}) \pi_{N,\alpha}(g | A_\epsilon) + \pi_{N,\alpha}(\bar{h}_1 \mathbf{1}_{A_\epsilon^c}) \pi_{N,\alpha}(g | A_\epsilon^c). \end{aligned}$$

The FKG inequality implies that both terms in the last expression are nonnegative: indeed, A_ε is increasing and therefore

$$\begin{aligned} \pi_{N,\alpha}(\bar{h}_1 \mathbf{1}_{A_\varepsilon}) &\geq \pi_{N,\alpha}(\bar{h}_1) \pi_{N,\alpha}(A_\varepsilon) = 0, \\ \pi_{N,\alpha}(g \mathbf{1}_{A_\varepsilon}) &\geq \pi_{N,\alpha}(g) \pi_{N,\alpha}(A_\varepsilon) = 0, \\ \pi_{N,\alpha}(\bar{h}_1 \mathbf{1}_{A_\varepsilon^c}) &\leq \pi_{N,\alpha}(\bar{h}_1) \pi_{N,\alpha}(A_\varepsilon^c) = 0, \\ \pi_{N,\alpha}(g \mathbf{1}_{A_\varepsilon^c}) &\leq \pi_{N,\alpha}(g) \pi_{N,\alpha}(A_\varepsilon^c) = 0. \end{aligned}$$

Hence the first term in the right-hand side of (27) must be zero. Since $\bar{h}_1 \mathbf{1}_{A_\varepsilon} \geq N - 1 - \varepsilon > 0$, this is possible only if $\pi_{N,\alpha}(g \mid A_\varepsilon) = 0$. However, that and the continuity of g imply that the extremal configuration $x_i \equiv N$ satisfies

$$g(N, \dots, N) = \lim_{\varepsilon \rightarrow 0^+} \inf_{x \in A_\varepsilon} g(x) = 0.$$

Hence, $g(x) \leq 0$ for all $x \in \Omega_N$ and therefore $g \equiv 0$.

2.4. *Absolute continuity.* Recall that under $\pi_{N,\alpha}$, the r.v. $(\eta_k)_{k=1}^N$ sum up to N . At several places in the proofs, it will be convenient to deal with independent r.v. instead. To that end, we use the following informal fact: for any $p \in (0, 1)$, the law of $(\eta_1, \eta_2, \dots, \eta_{\lfloor pN \rfloor})$ under $\pi_{N,\alpha}$ is uniformly over N absolutely continuous w.r.t. the law of $\lfloor pN \rfloor$ independent $\Gamma(\alpha, \alpha)$ r.v. The formal version is stated below.

LEMMA 10. Fix $p \in (0, 1)$. There exists a constant $C_p > 0$ such that for all $N \geq 1$, writing $n = \lfloor pN \rfloor$, for any bounded measurable function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ we have

$$\pi_{N,\alpha}[f(\eta_1, \dots, \eta_n)] \leq C_p \nu_N[f(\eta_1, \dots, \eta_n)],$$

where ν_N is the law on Ω_N^+ under which the η_k 's are IID $\Gamma(\alpha, \alpha)$ r.v.

PROOF. Let $g(\cdot)$ be the density function of a centered Gaussian distribution of unit variance. Let q_k be the density function of a centered $\Gamma(k\alpha, \alpha)$ r.v. By the local limit theorem ([19], Th.VII.2.7), we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \sqrt{\frac{N}{\alpha}} q_N(0) - g(0) &= 0, \\ \lim_{N \rightarrow \infty} \sup_{y \in \mathbb{R}} \left| \sqrt{\frac{N - n_N}{\alpha}} q_{N - n_N} \left(y \sqrt{\frac{N - n_N}{\alpha}} \right) - g(y) \right| &= 0. \end{aligned}$$

Using the above together with the fact that g is maximized at 0, we have for all N sufficiently large

$$\pi_{N,\alpha}[f(\eta_1, \dots, \eta_n)] = \frac{\nu_N[f(\eta_1, \dots, \eta_n) q_{N-n}(n - x_n)]}{q_N(0)} \leq \frac{2\nu_N[f(\eta_1, \dots, \eta_n)]}{\sqrt{1 - p}}.$$

The result follows by tuning the value of C_p to also include the first few values of N . \square

3. Lower bound on total variation distance. In this section, we prove the lower bound on the mixing time displayed in Theorems 1 and 2. We obtain in fact a more quantitative lower bound which is valid for all values $\alpha > 0$.

PROPOSITION 11. *For any $\alpha > 0$, for any $\varepsilon \in (0, 1)$ there exists $C_{\alpha,\varepsilon} > 0$ such that for all $N \geq 2$*

$$(28) \quad T_{N,\alpha}(1 - \varepsilon) \geq \frac{N^2}{\pi^2}(\log N - C_{\alpha,\varepsilon}).$$

We are going to follow a variant of the method introduced by Wilson in [25]: to obtain the lower bound (28), we select a specific test function f and show that by time $t = \frac{N^2}{\pi^2}(\log N - C_{\alpha,\varepsilon})$, the value $f(\mathbf{X}(t))$ is far from the equilibrium value $\pi_{N,\alpha}(f)$ with large probability. This is achieved by picking a suitable initial condition and by evaluating the first two moments of $f(\mathbf{X}(t))$. As in the case of the exclusion process, we select $f = f_N$, the eigenfunction appearing in Proposition 1. As noted in the proof of that proposition, f_N is an eigenfunction with eigenvalue gap_N for all $\alpha > 0$.

The initial condition $\mathbf{X}(0)$ is defined as follows. When $\alpha = 1$ we let $\eta_k, k = 1, \dots, N/2$ be¹ i.i.d. exponential random variables with mean 2 conditioned on $\eta_1 + \dots + \eta_{N/2} = N$, and set $\eta_k = 0$ for all $k \in \llbracket N/2 + 1, N \rrbracket$. More generally, for arbitrary $\alpha > 0$ we choose the distribution of η_k for $k \leq N/2$ to be i.i.d. $\Gamma(\alpha, \alpha/2)$, with the same conditioning on the sum, and set again $\eta_k = 0$ for all $k \in \llbracket N/2 + 1, N \rrbracket$. For the rest of this section, we let $\mathbf{X}(t)$ be the process starting from the random initial condition with increments η_k ,

$$(29) \quad X_k(0) = \sum_{i=1}^k \eta_i, \quad k \in \llbracket 1, N \rrbracket.$$

We shall use \mathbb{P}, \mathbb{E} for the corresponding probability measure and expectation, and Var for the associated variance. For later use, we also prove a result concerning the variance of other Fourier coefficients of $\mathbf{X}(t)$, namely the functions $f_N^{(j)}$ from (26).

LEMMA 12. *For the process described above with initial condition (29), there exists $C > 0$ such that for all $t \geq 0$ and all N large enough*

$$\mathbb{E}[f_N(\mathbf{X}(t))] \geq C^{-1}N^2e^{-\text{gap}_N t}, \quad \text{and} \quad \text{Var}[f_N(\mathbf{X}(t))] \leq CN^3,$$

and for every $j \in \llbracket 2, N - 1 \rrbracket$

$$\text{Var}[f_N^{(j)}(\mathbf{X}(t))] \leq Cj^{-2}N^3,$$

With this lemma at hand, the proof of the lower bound is simple.

PROOF OF PROPOSITION 11. Define

$$B_t = \left\{ x \in \Omega_N : f_N(x) \geq \frac{1}{2}\mathbb{E}[f_N(\mathbf{X}(t))] \right\},$$

and note that for all $t \geq 0$:

$$d^N(t) \geq \|\mathbb{P}(\mathbf{X}(t) \in \cdot) - \pi_{N,\alpha}\|_{\text{TV}} \geq \mathbb{P}(\mathbf{X}(t) \in B_t) - \pi_{N,\alpha}(B_t).$$

Chebyshev’s inequality and Lemma 12 imply that

$$\mathbb{P}(\mathbf{X}(t) \in B_t) \geq 1 - 4C^3N^{-1}e^{2\text{gap}_N t}.$$

Similarly, noting that $\text{Var}_{\pi_{N,\alpha}}(f_N) = \pi_{N,\alpha}(f_N^2) = \lim_{t \rightarrow \infty} \text{Var}[f_N(\mathbf{X}(t))]$, one has

$$\pi_{N,\alpha}(B_t) \leq 4C^3N^{-1}e^{2\text{gap}_N t}.$$

Recalling that $\text{gap}_N = 1 - \cos(\pi/N)$ concludes the proof. \square

¹In this section, we write $N/2$ for $\lfloor N/2 \rfloor$ in order to alleviate notations.

We close this section with the proof of the lemma.

PROOF OF LEMMA 12. First of all, since f_N is an eigenfunction of $\mathcal{L}_{N,\alpha}$ associated with the eigenvalue gap_N , we have

$$\mathbb{E}[f_N(\mathbf{X}(t))] = \mathbb{E}[f_N(\mathbf{X}(0))]e^{-\text{gap}_N t} \geq \frac{N^2}{20}e^{-\text{gap}_N t},$$

where the last bound follows from the fact that the initial condition satisfies

$$\mathbb{E}[X_k(0)] \geq \min\{2k, N\}.$$

To compute the variance at a fixed time $t > 0$ (we focus on the case $j = 1$ to keep the notation light and explain the general case at the end of the proof), we introduce the process

$$M_s^t := e^{\text{gap}_N(s-t)} f_N(\mathbf{X}(s)), \quad s \in [0, t].$$

Since $\mathbb{E}[f_N(\mathbf{X}(s)) \mid \mathbf{X}(u)] = e^{-\text{gap}_N(s-u)} f_N(\mathbf{X}(u))$, $u \in [0, s]$, it follows that $(M_s^t, s \in [0, t])$ is a martingale. The associated increasing predictable process, or angle bracket, is denoted $\langle M^t \rangle_s, s \in [0, t]$. In particular, we look for an upper bound on $\text{Var}[f_N(\mathbf{X}(t))] = \mathbb{E}[\langle M^t \rangle_t]$.

Note that independently of the value of α , when an update is performed at coordinate k , the value of $f_N(\mathbf{X}(t))$ varies at most by $\eta_k(s) + \eta_{k+1}(s)$, where $\eta_k(s) = (X_k - X_{k-1})(s)$. Thus,

$$\begin{aligned} \partial_s \langle M^t \rangle_s &\leq e^{2\text{gap}_N(s-t)} \sum_{k=1}^{N-1} (\eta_k(s) + \eta_{k+1}(s))^2 \\ &\leq 4e^{2\text{gap}_N(s-t)} \sum_{k=1}^N (\eta_k(s))^2. \end{aligned}$$

Therefore,

$$(30) \quad \text{Var}[f_N(\mathbf{X}(t))] = \mathbb{E}[\langle M^t \rangle_t] \leq 4 \int_0^t e^{2\text{gap}_N(s-t)} \sum_{k=1}^N \mathbb{E}[\eta_k(s)^2] ds.$$

Now using monotonicity of the dynamics for the gradients (Proposition 7), we can replace $\eta_k(s)$ by the gradients corresponding to a dynamics on Ω_N^+ starting from a larger initial condition for the order \succ defined in (16). A natural choice is to pick an initial condition which is stationary for the dynamics so that the dependence in s vanishes in the integral.

Let $(\eta'_k)_{k=1}^N$ be i.i.d. variables with distribution $\Gamma(\alpha, \alpha/2)$, conditioned to

$$\sum_{j=1}^{N/2} \eta'_j \geq N,$$

and consider $\mathbf{X}'(t)$ the dynamics on Ω_N^+ with initial condition $X'_k(0) := \sum_{i=1}^k \eta'_i$. Note that one can construct η and η' on the same probability space in such a way that $\eta_k \leq \eta'_k$ for all k by setting

$$(31) \quad \eta_k := \frac{N\eta'_k}{\sum_{j \leq N/2} \eta'_j} \quad \text{for } k \leq N/2,$$

and $\eta_k = 0$, for $k \in \llbracket N/2 + 1, N \rrbracket$. Indeed, by a standard property of the gamma distribution, the variables $\eta'_k / \sum_{j \leq N/2} \eta'_j$ and $\sum_{j \leq N/2} \eta'_j$ are independent and therefore the η_k defined in (31) has the correct distribution. Hence, using Proposition 7 we can assume $\mathbf{X}(t) \preceq \mathbf{X}'(t)$ for all $t \geq 0$ and thus in particular $\mathbb{E}[\eta_k(s)^2] \leq \mathbb{E}[\eta'_k(s)^2]$.

Finally let $(\eta''_k)_{k=1}^N$ be (unconditioned) i.i.d. with distribution $\Gamma(\alpha, \alpha/2)$. Defining the dynamics $\eta''_k(s)$ with this initial condition one has

$$\mathbb{E}[\eta'_k(s)^2] = \mathbb{E}\left[\eta''_k(s)^2 \mid \sum_{j \leq N/2} \eta''_j(0) \geq N\right] \leq \frac{\mathbb{E}[\eta''_k(s)^2]}{\mathbb{P}[\sum_{j \leq N/2} \eta''_j(0) \geq N]}.$$

By stationarity,

$$\mathbb{E}[\eta''_k(s)^2] = \mathbb{E}[\eta''_k(0)^2] = 4 + \frac{4}{\alpha}.$$

Since the expected value of each $\eta''_k(0)$ is 2, the central limit theorem shows that the event $\sum_{j \leq N/2} \eta''_j(0) \geq N$ has probability at least 1/3 if N is sufficiently large. In conclusion,

$$\mathbb{E}[\eta_k(s)^2] \leq \mathbb{E}[\eta'_k(s)^2] \leq 12(1 + \alpha^{-1}).$$

Using this bound in (30) shows that $\text{Var}[f_N(\mathbf{X}(t))] \leq CN / \text{gap}_N \leq C'N^3$ uniformly in $t \geq 0$.

For $j \in \llbracket 2, N - 1 \rrbracket$, we repeat the above procedure for the martingale

$$M_s^{t,j} := e^{-\lambda_N^{(j)}(s-t)} f_N^{(j)}(\mathbf{X}(t))$$

and obtain $\text{Var}[f_N^{(j)}(\mathbf{X}(t))] \leq CN / \lambda_N^{(j)} \leq CN^3 j^{-2}$. This concludes the proof of Lemma 12. □

4. Mixing time for the separation distance. Here we prove Theorem 3. The main result of this section is the following lower bound.

PROPOSITION 13. *For any $\alpha > 0$, there exists $C_{\alpha,\varepsilon} > 0$ such that for any $\varepsilon \in (0, 1)$ and for all $N \geq N_\varepsilon$ sufficiently large we have*

$$(32) \quad T_{N,\alpha}^{\text{sep}}(\varepsilon) \geq N^2 \left(\frac{2}{\pi^2} \log N - C_{\alpha,\varepsilon} \log \log N \right).$$

With this result at hand, and assuming the validity of Theorem 2, the derivation of the asymptotic of the separation mixing times is somewhat standard.

PROOF OF THEOREM 3. Proposition 13 gives the desired lower bound on the mixing times. Regarding the upper bound, we adapt the argument used in the discrete setup (see, e.g., [14], Lemma 19.3) to show that

$$(33) \quad d_{N,\alpha}^{\text{sep}}(2t) \leq 1 - (1 - 2d_{N,\alpha}(t))_+^2,$$

where $(\cdot)_+$ denotes the positive part. Theorem 2 and (33) clearly imply the desired upper bound in Theorem 3.

Recalling the notation $P_t(A, B)$ defined below (11) we notice that reversibility implies that $P_t(A, B) = P_t(B, A)$ and as a consequence

$$(34) \quad P_t(B, dz) = P_t^z(B) \pi_{N,\alpha}(dz).$$

Thus for any A, B with positive Lebesgue measure, using the semigroup property at the first line and reversibility at the second line

$$\begin{aligned} P_{2t}(A, B) &= \int_{\Omega_N} P_t(A, dz) P_t^z(B) \\ &= \int_{\Omega_N} P_t^z(A) P_t^z(B) \pi_{N,\alpha}(dz). \end{aligned}$$

By Schwarz' inequality, we have

$$\begin{aligned} \sqrt{\frac{P_{2t}(A, B)}{\pi_{N,\alpha}(A)\pi_{N,\alpha}(B)}} &\geq \int_{\Omega_N} \sqrt{\frac{P_t^z(A)P_t^z(B)}{\pi_{N,\alpha}(A)\pi_{N,\alpha}(B)}} \pi_{N,\alpha}(dz) \\ &\geq \int_{\Omega_N} \frac{P_t^z(A)}{\pi_{N,\alpha}(A)} \wedge \frac{P_t^z(B)}{\pi_{N,\alpha}(B)} \pi_{N,\alpha}(dz). \end{aligned}$$

Let $\bar{P}_t(A, dz) = (\pi_{N,\alpha}(A))^{-1} P_t(A, dz)$ denote the normalized version of $P_t(A, dz)$. From (34) it follows that $\frac{P_t^z(A)}{\pi_{N,\alpha}(A)}$ and $\frac{P_t^z(B)}{\pi_{N,\alpha}(B)}$ are the respective densities of $\bar{P}_t(A, \cdot)$ and $\bar{P}_t(B, \cdot)$ w.r.t. $\pi_{N,\alpha}$. Therefore, using the triangular inequality one has

$$\sqrt{\frac{P_{2t}(A, B)}{\pi_{N,\alpha}(A)\pi_{N,\alpha}(B)}} \geq 1 - \|\bar{P}_t(A, \cdot) - \bar{P}_t(B, \cdot)\|_{TV} \geq (1 - 2d_{N,\alpha}(t))_+.$$

Taking the infimum over A and B yields (33). \square

PROOF OF PROPOSITION 13. We are going to show that for all $\alpha, \varepsilon > 0$, there exists $C > 0$ such that if

$$t_0 := N^2 \left(\frac{2}{\pi^2} \log N - C \log \log N \right),$$

then one can find measurable sets $A, B \subset \Omega_N$ such that

$$P_{t_0}(A, B) \leq \varepsilon \pi_{N,\alpha}(A) \pi_{N,\alpha}(B).$$

A natural choice to minimize $P_{t_0}(A, B)$ is to choose A, B as tiny neighbourhoods of the opposite extremal configurations. We define the neighbourhoods of our extremal configurations $\bar{\vee}$ and $\bar{\wedge}$ as follows:

$$(35) \quad \begin{aligned} \bar{\wedge} &:= \{x \in \Omega_N : \forall i \in \llbracket 1, N-1 \rrbracket, x_i \geq N-1/N\}, \\ \bar{\vee} &:= \{x \in \Omega_N : \forall i \in \llbracket 1, N-1 \rrbracket, x_i \leq 1/N\} \end{aligned}$$

and we are going to prove

$$(36) \quad \int_{\bar{\wedge}} P_{t_0}(x, \bar{\vee}) \pi_{N,\alpha}(dx) \leq \varepsilon \pi_{N,\alpha}(\bar{\wedge}) \pi_{N,\alpha}(\bar{\vee}).$$

We assume for simplicity that N is even just for the sake of notation. The first step of the proof is to reduce (36) to a simple statement about the process at time $t_0/2$,

$$(37) \quad \frac{1}{\pi_{N,\alpha}(\bar{\wedge})} \int_{\bar{\wedge}} P_{t_0/2}^x[x_{N/2} < N/2] \pi_{N,\alpha}(dx) \leq \varepsilon/4,$$

and the second step is to prove (37).

Let us now show that (36) follows from (37) as a consequence of reversibility and the FKG inequality. The proof is in fact very similar to the one developed in [13], Section 7.1, in a discrete setup. We have

$$(38) \quad \begin{aligned} &\int_{\bar{\wedge}} P_{t_0}(x, \bar{\vee}) \pi_{N,\alpha}(dx) \\ &= \int_{x \in \bar{\wedge}} \int_{y \in \Omega_N} \int_{z \in \bar{\vee}} \pi_{N,\alpha}(dx) P_{t_0/2}(x, dy) P_{t_0/2}(y, dz). \end{aligned}$$

Now we can split the integral on y into two contributions

$$U_+ = \{x \in \Omega_N : x_{N/2} \geq N/2\} \quad \text{and} \quad U_- = \{x \in \Omega_N : x_{N/2} \leq N/2\},$$

with $U_+ \cup U_- = \Omega_N$ and $\pi_{N,\alpha}(U_+ \cap U_-) = 0$. Using reversibility and symmetry we have

$$\begin{aligned}
 (39) \quad & \int_{x \in \bar{\lambda}} \int_{y \in U_+} \int_{z \in \underline{\vee}} \pi_{N,\alpha}(dx) P_{t_0/2}(x, dy) P_{t_0/2}(y, dz) \\
 &= \int_{x \in \bar{\lambda}} \int_{y \in U_+} \int_{z \in \underline{\vee}} \pi_{N,\alpha}(dz) P_{t_0/2}(y, dx) P_{t_0/2}(z, dy) \\
 &= \int_{x \in \underline{\vee}} \int_{y \in U_-} \int_{z \in \bar{\lambda}} \pi_{N,\alpha}(dz) P_{t_0/2}(y, dx) P_{t_0/2}(z, dy),
 \end{aligned}$$

so that the right-hand side of (38) is equal to

$$(40) \quad 2 \int_{x \in \bar{\lambda}} \int_{y \in U_-} \pi_{N,\alpha}(dx) P_{t_0/2}(x, dy) P_{t_0/2}(y, \underline{\vee}).$$

Now according to the observations made in Section 6 (see Lemma 32), the measure

$$\int_{x \in \bar{\lambda}} \pi_{N,\alpha}(dx) P_{t_0/2}(x, dy)$$

has an increasing density with respect to $\pi_{N,\alpha}$ (call it ρ_+). Also we observe that by monotone coupling $P_{t_0/2}(y, \underline{\vee})$ is decreasing in y . Hence, using the FKG inequality (Proposition 8 for the restriction to the stable set U_-) we obtain that the quantity (40) satisfies

$$\begin{aligned}
 (41) \quad & 2 \int_{U_-} \rho_+(y) P_{t_0/2}(y, \underline{\vee}) \pi_{N,\alpha}(dy) \\
 & \leq \frac{2}{\pi_{N,\alpha}(U_-)} \int_{U_-} \rho_+(y) \pi_{N,\alpha}(dy) \int_{U_-} P_{t_0/2}(y, \underline{\vee}) \pi_{N,\alpha}(dy).
 \end{aligned}$$

Using stationarity we see that the second integral is smaller than

$$\int_{\Omega_N} P_{t_0/2}(y, \underline{\vee}) \pi_{N,\alpha}(dy) = \pi_{N,\alpha}(\underline{\vee}).$$

Hence, using the fact that $\pi_{N,\alpha}(U_-) = 1/2$ and replacing $\rho_+(x) \pi_{N,\alpha}(dx)$ by its definition we can conclude that

$$(42) \quad \int_{\bar{\lambda}} P_{t_0}(x, \underline{\vee}) \pi_{N,\alpha}(dx) \leq 4 \pi_{N,\alpha}(\underline{\vee}) \int_{\bar{\lambda}} P_{t_0/2}(x, U_-) \pi_{N,\alpha}(dx).$$

This proves that (36) follows from (37). It remains to prove (37), or in other words that starting from a random initial condition in $\bar{\lambda}$ the probability that $\mathbf{X}_{t_0/2} \in U_-$ is small.

Note that (37) follows from a slightly stronger result for the maximal initial condition

$$(43) \quad \mathbb{P}[X_{N/2}^\wedge(t_0/2) \geq N/2 + 1] \geq 1 - \varepsilon/4.$$

Indeed, with the grand coupling described in the proof of Proposition 7, for any $x \in \bar{\lambda}$, $i \in \llbracket 1, N - 1 \rrbracket$ and any $t > 0$ we have $X_i^x(t) \geq X_i^\wedge - 1$.

Now by monotonicity we can replace $\mathbf{X}^\wedge(t)$ by the dynamics with the random initial condition considered in Lemma 12.

The function $i \mapsto X_i(t) - i$ can be decomposed on the orthonormal basis of the discrete Laplacian (endowed with Dirichlet b.c.). Using (26), we thus obtain

$$(44) \quad X_i(t) - i = \frac{2}{N} \sum_{j=1}^{N-1} \sin\left(\frac{ij\pi}{N}\right) f_N^{(j)}(\mathbf{X}(t)).$$

Using the fact that $\mathbb{E}[f_N^{(j)}(\mathbf{X}(t))] = e^{-\lambda_N^{(j)} t} \mathbb{E}[f_N^{(j)}(\mathbf{X}(0))]$, and observing that all the terms with $j \geq 2$ become negligible, it follows that if the constant C is large enough, then

$$(45) \quad \mathbb{E}[X_{N/2}(t_0/2) - N/2] \geq N^{1/2}(\log N)^2.$$

Now the important part is to control the variance of $X_{N/2}(t)$. Our control is uniform in time.

From the very rough bound $\text{Var}(\sum_{i \in I} Z_i) \leq (\sum_{i \in I} \sqrt{\text{Var} Z_i})^2$ valid for any sequence of random variables $(Z_i)_{i \in I}$, and using Lemma 12 one obtains

$$(46) \quad \text{Var}(X_{N/2}(t)) \leq 4N^{-2} \left(\sum_{j=1}^{N-1} \sqrt{\text{Var} f_N^{(j)}(\mathbf{X}(t))} \right)^2 \leq CN(\log N)^2.$$

Then (43) can easily be deduced by combining the inequalities for the two first moments (45) and (46). \square

5. Upper bound on total variation distance.

5.1. *Decomposition of the proof.* In this section, we prove the upper bound on the total variation mixing time displayed in Theorems 1 and 2. Given $\delta > 0$ we set

$$t_{N,\delta} = t_\delta := (1 + \delta) \frac{\log N}{2 \text{gap}_N},$$

and we are going to prove that

$$(47) \quad \lim_{N \rightarrow \infty} \sup_{x \in \Omega_N} \|P_{t_{N,\delta}}^x - \pi_{N,\alpha}\|_{\text{TV}} = 0.$$

Since $\delta > 0$ is arbitrary and $\text{gap}_N \sim \frac{1}{2}\pi^2 N^{-2}$ this yields the desired upper bound. We establish this statement via two intermediate propositions. Firstly, we show that by time t_δ we can bring to equilibrium the extremal process started in the maximal configuration \wedge .

PROPOSITION 14. *For any $\delta > 0$, we have*

$$\lim_{N \rightarrow \infty} \|P_{t_{N,\delta}}^\wedge - \pi_{N,\alpha}\|_{\text{TV}} = 0.$$

Secondly, we show that with large probability and for any given initial condition x , we can couple the process starting from x with the maximal process.

PROPOSITION 15. *For any $\delta > 0$ we have*

$$\lim_{N \rightarrow \infty} \sup_{x \in \Omega_N} \|P_{t_{N,\delta}}^x - P_{t_{N,\delta}}^\wedge\|_{\text{TV}} = 0.$$

By the triangle inequality, (47) is an immediate consequence of the two propositions above.

Even though Proposition 14 is a consequence of Proposition 15, our proof of the latter will actually rely on the former result; see also Remark 42 below. Consequently, we will need an independent proof of Proposition 14 and this will be carried out in Section 6. The remainder of this section is concerned with the proof of Proposition 15.

From now on, $x \in \Omega_N$ is fixed; however, all the constants that will appear below will be *independent* of this chosen x . In the forthcoming Section 5.2, we construct the processes \mathbf{X}^\wedge and \mathbf{X}^x on the same probability space in such a way that (recall (15))

$$(48) \quad \forall t \geq 0, \quad \mathbf{X}^\wedge(t) \geq \mathbf{X}^x(t).$$

In the remainder of this proof \mathbb{P} denotes the probability associated with the probability space on which this coupling is constructed. We let A_t denote the area comprised between the graphs of \mathbf{X}^x and \mathbf{X}^\wedge :

$$A_t := \sum_{k=1}^{N-1} (X_k^\wedge(t) - X_k^x(t)),$$

and we aim at bounding its hitting time of 0

$$\tau := \inf\{t \geq 0 : A_t = 0\}.$$

As the ordering given by (48) implies that $\mathbf{X}^\wedge(t) = \mathbf{X}^x(t)$ for $t \geq \tau$, we have

$$(49) \quad \|P_t^x - P_t^\wedge\|_{TV} \leq \mathbb{P}[\tau > t].$$

Of course the distribution of τ depends tremendously on the coupling. For instance, the reader can check that for the coupling provided by Proposition 7 which satisfies (48), we have $\tau = \infty$ with probability one. The coupling presented below is constructed with the aim of minimizing the merging time τ . The proof is split into three main steps:

- (1) The area passes below $N^{3/2}$ by time $t_{\delta/2}$ with large probability.
- (2) Within an additional time of order N^2 , the area is very likely to decline from $N^{3/2}$ to N^{-1} .
- (3) The area goes from N^{-1} down to 0 within a time of order $\log N$ with large probability.

The above steps clearly ensure that the event $\tau > t_\delta$ has a small probability, which will conclude the proof of Proposition 15.

The first step is a rather simple consequence of the fact that $\mathbb{E}[X_k^\wedge(t) - X_k^x(t)]$ is a solution of the discrete heat equation.

The third step is a brute force argument inspired by Randall and Winkler [22]. More precisely, for this last step we build on the main idea of [22], but we improve it in a quantitative manner using estimates on the minimum gradient of \mathbf{X}^\wedge .

The second step above is by far the most delicate one. The strategy is to introduce a sequence of intermediate thresholds between the values $N^{3/2}$ and N^{-1} , and then analyse the associated hitting times for the area process. Our control of these hitting times relies on diffusive estimates for the supermartingale A_t and on estimates on the corresponding angle bracket process. One of the ingredients of the latter estimates is a fine control of the gradients of \mathbf{X}^\wedge , which is in turn derived from a combination of equilibrium estimates and Proposition 14.

5.2. *The coupling and preliminary lemmas.* We start by defining the coupling \mathbb{P} . For any $k \in \llbracket 1, N - 1 \rrbracket$ and any configuration $x \in \Omega_N$, we define the “interval of resampling of the k th coordinate” as $I(x, k) := [x_{k-1}, x_{k+1}]$ and write $|I(x, k)| = \nabla x_k$ for its length, where the “gradient” ∇x_k is defined by

$$(50) \quad \nabla x_k := x_{k+1} - x_{k-1}.$$

For simplicity we sometimes use the short-hand notation $I^\wedge = I(\mathbf{X}^\wedge(t_-), k)$ and $I^x = I(\mathbf{X}^x(t_-), k)$. In our construction, we are going to try, at each resampling event, to couple \mathbf{X}_k^x and \mathbf{X}_k^\wedge with the maximal probability. Letting $\text{Beta}_\alpha(I)$ denoting the distribution with density given by (when $I = [a, b]$)

$$B_\alpha[a, b](x) := \frac{\Gamma(2\alpha)(x - a)^{\alpha-1}(b - x)^{\alpha-1}}{\Gamma(\alpha)^2(b - a)^{2\alpha-1}} \mathbf{1}_{[a, b]}(x),$$

we set

$$(51) \quad \begin{aligned} p &= p(t, k) := 1 - \|\text{Beta}_\alpha(I^x) - \text{Beta}_\alpha(I^\wedge)\|_{TV} \\ &= \int_{I^\wedge \cap I^x} \min(B_\alpha(I^x), B_\alpha(I^\wedge))(x) \, dx, \end{aligned}$$

and $q := 1 - p$. In the case $\alpha = 1$, we have

$$p = |I^\wedge \cap I^x| / \max(\nabla X_k^\wedge(t_-), \nabla X_k^x(t_-))$$

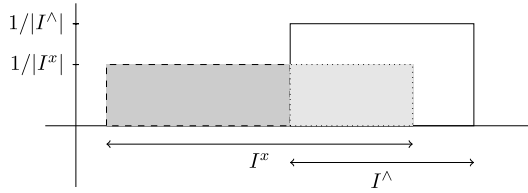


FIG. 1. A plot of the densities $B_\alpha(I^x)$ and $B_\alpha(I^\wedge)$ when $\alpha = 1$. The light, respectively dark, shaded region has area p , respectively q .

but there is no such simple expression for p when $\alpha > 1$; however, we will be able to provide good estimates for it (cf. Lemma 16).

Then, we define ν_1, ν_2, ν_3 to be the probability measures with respective densities $\rho_1, \rho_2,$ and ρ_3 given by

$$\begin{aligned}
 \rho_1 &= q^{-1}[B_\alpha(I^x) - B_\alpha(I^\wedge)]_+, \\
 \rho_2 &= p^{-1} \min(B_\alpha(I^x), B_\alpha(I^\wedge)), \\
 \rho_3 &= q^{-1}[B_\alpha(I^\wedge) - B_\alpha(I^x)]_+.
 \end{aligned}
 \tag{52}$$

As a consequence of our assumption $\alpha \geq 1$ (which makes the functions $B_\alpha(I)$ unimodal), and the fact that by monotonicity both extremities of I^\wedge are larger than their counterparts in I^x , the supports of ρ_1 and ρ_3 are intervals I_1 and I_3 , the lower extremity of I_3 being larger or equal than the upper extremity of I_1 . To see this it is sufficient to check that there exists at most one value u such that the equation $B_\alpha(I^x)(u) = B_\alpha(I^\wedge)(u) > 0$ has at most one solution. We refer to Figure 1 for the case $\alpha = 1$ and to Figure 2 for the general case of a unimodal density.

Our coupled dynamics is defined as follows. Each pair of coordinates $(X_k^x, X_k^\wedge), k = 1, \dots, N - 1,$ is updated with rate one independently, and when an update occurs at time t then:

- with probability p , the two new coordinates are set to the same value $X_k^x(t) = X_k^\wedge(t)$ drawn from the distribution ν_2 ,
- with probability q , the two new coordinates are sampled independently with respective distributions ν_1 and ν_3 , and therefore satisfy $X_k^x(t) \leq X_k^\wedge(t)$.

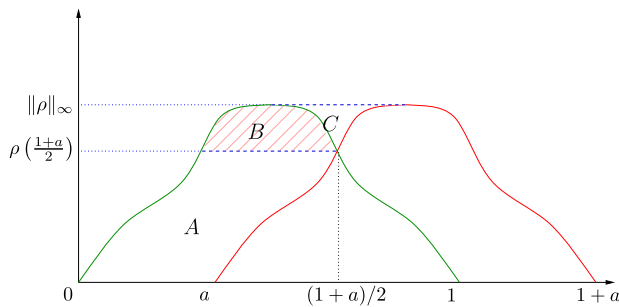


FIG. 2. We present a graphic proof of the inequalities (59) and (60). Using symmetry and unimodality of ρ we see that the vertical stripe of constant width delimited by the two increasing portions of the graphs ρ and $\rho_{[a, 1+a]}$ can be divided into 3 regions A, B and C, the separation being given by the horizontal line $y = \rho(\frac{1+a}{2})$ and decreasing part of the graph of ρ . The integral we wish to estimate $\int_{[0, 1]} [\rho(u) - \rho_{[a, 1+a]}(u)]_+ du$ corresponds to $|A| + |B|$ ($|\cdot|$ standing here for Lebesgue measure) while the upper and lower bounds $\|\rho\|_\infty a$ and $\rho(\frac{1+a}{2})a$ correspond respectively, to $|A| + |B| + |C|$ and $|A|$.

For convenience, we set

$$\delta X_k(t) := X_k^\wedge(t) - X_k^x(t)$$

(in the remainder of the proof δ as a positive parameter is not used anymore so that this should not yield confusion). We also introduce the “mean” interfaces \bar{X}^\wedge and \bar{X}^x by setting

$$\bar{X}_k^\wedge(t) := \frac{X_{k-1}^\wedge(t) + X_{k+1}^\wedge(t)}{2},$$

and similarly for \bar{X}^x . Note that $\bar{X}_k^\wedge(t)$ is the midpoint of the interval of resampling of the k th coordinate. We finally set

$$\delta \bar{X}_k(t) := \bar{X}_k^\wedge(t) - \bar{X}_k^x(t).$$

We now collect a few facts on our coupling. We start by showing that the probability $q(t, k)$ can be fairly approximated by

$$(53) \quad Q(t, k) := \min\left(\frac{\delta \bar{X}_k(t_-)}{\max(\nabla X_k^\wedge(t_-), \nabla X_k^x(t_-))}, 1\right),$$

and that, as soon as q is close enough to 1, the overlap between the resampling intervals represents a small fraction of the largest resampling interval. This second bound will be convenient in order to bound from below the derivative of the bracket of the area, see the proof of Proposition 18.

LEMMA 16. *For any $\alpha \geq 1$, there is a constant C (that depends on α) such that for all $t \geq 0$ and $k \in \llbracket 1, N - 1 \rrbracket$*

$$(54) \quad \frac{1}{C} Q(t, k) \leq q(t, k) \leq C Q(t, k).$$

Furthermore, there exists a constant $c_1 > 0$ such that

$$(55) \quad q(t, k) \geq 1 - c_1 \implies \max(|I^x|, |I^\wedge|) \geq 2|I^x \cap I^\wedge|$$

(where the notation used is the one introduced below (50)).

PROOF. Note that (54) is simply a result concerning the total variation between two $Beta_\alpha$ variables defined on two intervals $I_1 := [l_1, r_1]$ and $I_2 := [l_2, r_2]$ such that $l_1 \leq l_2$ and $r_1 \leq r_2$. Recalling the definition (51),

$$(56) \quad q = \int_0^1 [B_\alpha[l_1, r_1](u) - B_\alpha[l_2, r_2](u)]_+ du,$$

and (54) holds if we can prove (for a different constant C) that

$$(57) \quad \frac{1}{C} \min\left(\frac{\max(|l_2 - l_1|, |r_2 - r_1|)}{\max(r_1 - l_1, r_2 - l_2)}, 1\right) \leq q \leq C \frac{\max(|l_2 - l_1|, |r_2 - r_1|)}{\max(r_1 - l_1, r_2 - l_2)}.$$

By symmetry, we can assume that I_1 is the largest of the two intervals and by scaling invariance we can assume without loss of generality that $I_1 = [0, 1]$ and $I_2 =: [a, a + b]$, $a > 0$, $b \in (1 - a, 1]$, in which case

$$\max(|l_2 - l_1|, |r_2 - r_1|) / \max(r_1 - l_1, r_2 - l_2) = a.$$

We can further assume that $a < 1$ as the result is trivially valid when $a \geq 1$.

For better readability of the proof, we replace $Beta_\alpha$ by a generic unimodal function ρ which is positive on $(0, 1)$, integrates to 1 and is symmetric around $1/2$. We replace

$\text{Beta}_\alpha[a, a + b]$ by $\rho_{[a, a+b]} = b^{-1}\rho((\cdot - a)/b)$ with the convention that $\rho(u) = 0$ for $u \notin [0, 1]$.

We first let the reader check that $\int_{[0,1]}[\rho(u) - \rho_{[a, a+b]}(u)]_+ du$ is increasing in a and b simply because the integrand displays the same monotonicities.

For (55), we simply observe, using these monotonicities in a and b , that if $a < 1/2$ then

$$\int_{[0,1]}[\rho(u) - \rho_{[a, a+b]}(u)]_+ du < \int_{[0,1]}[\rho(u) - \rho_{[1/2, 3/2]}(u)]_+ du.$$

As a consequence, setting $c_1 = 1 - \int_{[0,1]}[\rho(u) - \rho_{[1/2, 3/2]}(u)]_+ du$, we deduce that if $q \geq 1 - c_1$, we have $a \geq 1/2$ and $|I_1 \cap I_2| \leq 1/2$ thus yielding (55).

To prove (54), we show that for every $a \in [0, 1]$

$$(58) \quad \frac{\rho(1/4)a}{4} \leq \int_{[0,1]}[\rho(u) - \rho_{[a, a+b]}(u)]_+ du \leq \|\rho\|_\infty a,$$

which allows to prove (57) with $C = \max(\|\rho\|_\infty, 4/\rho(1/4))$. By monotonicity it is sufficient to check the upper bound for $b = 1$. Figure 2 provides a graphical proof of the following inequality:

$$(59) \quad \int_{[0,1]}[\rho(u) - \rho_{[a, 1+a]}(u)]_+ du \leq \|\rho\|_\infty a.$$

Now concerning the lower bound, using again the graphical proof of Figure 2 (by symmetry of ρ the two curves intersect at $(1 + a)/2$), we have for $a \leq 1/2$

$$(60) \quad \int_{[0,1]}[\rho(u) - \rho_{[a, 1+a]}(u)]_+ du \geq a\rho\left(\frac{1+a}{2}\right) \geq \rho(1/4)a,$$

and thus the l.h.s. being increasing in a we conclude that for all $a \in (0, 1)$ (the factor $1/2$ is present so that the inequality is also valid for $a \in (1/2, 1]$)

$$\int_{[0,1]}[\rho(u) - \rho_{[a, 1+a]}(u)]_+ du \geq \frac{\rho(1/4)a}{2}.$$

Using symmetry and invariance by translation at the first line and the triangle inequality for the total variation distance at the second line, we have

$$\begin{aligned} & \int_{[0,1]}[\rho(u) - \rho_{[a, 1]}(u)]_+ du \\ &= \frac{1}{2} \left(\int_{[0,1]}[\rho(u) - \rho_{[a, 1]}(u)]_+ du + \int_{[0,1]}[\rho_{[a, 1]}(u) - \rho_{[a, 1+a]}(u)]_+ du \right) \\ &\geq \frac{1}{2} \int_{[0,1]}[\rho(u) - \rho_{[a, 1+a]}(u)]_+ du \geq \frac{\rho(1/4)a}{4}. \end{aligned}$$

This lower bound is also valid for $b \in (1 - a, 1)$ by monotonicity in b , thus completing the proof of (58). \square

REMARK 17. Let us observe for latter use that the upper bound in (57) is also valid without the assumption $l_1 \leq l_2, r_1 \leq r_2$ (at the cost of taking C twice as large). This can be deduced from the other case. Indeed, assume without loss of generality that $[l_1, r_1] \subset [l_2, r_2]$ and set $I_3 = [l_1, r_2]$. Then, using the previous bounds for the pairs (I_1, I_3) and (I_2, I_3) ,

$$\begin{aligned} & \|\text{Beta}_\alpha(I_1) - \text{Beta}_\alpha(I_2)\|_{\text{TV}} \\ &\leq \|\text{Beta}_\alpha(I_1) - \text{Beta}_\alpha(I_3)\|_{\text{TV}} + \|\text{Beta}_\alpha(I_3) - \text{Beta}_\alpha(I_2)\|_{\text{TV}} \\ &\leq C \frac{|r_2 - r_1|}{r_2 - l_1} + C \frac{|l_2 - l_1|}{r_2 - l_2} \leq C \frac{\max(|l_2 - l_1|, |r_2 - r_1|)}{\max(r_1 - l_1, r_2 - l_2)}. \end{aligned}$$

Using the expression for the generator (25) and the fact that our coupling preserves monotonicity the reader can check that for every $t \geq 0$

$$\partial_s \mathbb{E}[A_{t+s} \mid \mathcal{F}_t]_{s=0} = \frac{1}{2}(X_1^x(t) - X_1^\wedge(t)) + \frac{1}{2}(X_{N-1}^x(t) - X_{N-1}^\wedge(t)) \leq 0,$$

and hence that A_t is a supermartingale for the filtration $(\mathcal{F}_t)_{t \geq 0}$ defined by $\mathcal{F}_t := \sigma(\mathbf{X}_s^x, \mathbf{X}_s^\wedge, s \leq t)$. We write $\{\langle A \cdot \rangle_t, t \geq 0\}$, for the associated angle bracket process, namely the increasing predictable process that compensates the square of the martingale part of A_t .

PROPOSITION 18. *There exists a constant $c > 0$ such that for all N large enough and for all $t \geq 0$, we have*

$$\partial_t \langle A \cdot \rangle_t \geq c \sum_{k=1}^{N-1} \min(\delta \bar{X}_k(t-) \nabla X_k^\wedge(t-), (\nabla X_k^\wedge(t-))^2).$$

REMARK 19. The proof will actually establish the same bound but with $\nabla X_k^\wedge(t)$ replaced by the maximum of the latter and $\nabla X_k^x(t)$.

PROOF. First of all, recalling the coupling defined in Section 5.2, we have

$$\partial_t \langle A \cdot \rangle_t = \sum_{k=1}^{N-1} p(t, k) (\delta X_k(t-))^2 + q(t, k) J(t, k),$$

where $J(t, k)$ is the mean square displacement corresponding to an instantaneous uncoupled jump of $X_k^\wedge - X_k^x$ at time t . We are going to prove a lower bound for each term in the sum, and as before we omit from now on the dependence in k and t . Without loss of generality we assume that $\nabla X^\wedge \geq \nabla X^x$.

We let Z^x and Z^\wedge denote the two independent variables with respective distribution ν_1 and ν_3 (whose densities are described in equation (52)) that are used in the coupling. We are going to prove first that

$$(61) \quad p(\delta X)^2 + qJ \geq q(\text{Var}(Z^x) + \text{Var}(Z^\wedge)) + \frac{p}{q}(\delta \bar{X})^2.$$

This is achieved by computing explicitly J (here E denotes expectation for the pair of variables (Z^x, Z^\wedge))

$$(62) \quad J = E[(Z^\wedge - Z^x - \delta X)^2] = \text{Var}(Z^x) + \text{Var}(Z^\wedge) + (E[Z^\wedge - Z^x] - \delta X)^2.$$

Replacing J by its value, observing that $\delta \bar{X} = qE[Z^\wedge - Z^x]$, and taking the minimum over all possible values for δX we obtain that the l.h.s. of (61) is larger than

$$q(\text{Var}(Z^x) + \text{Var}(Z^\wedge)) + \min_{u \in \mathbb{R}} [q(q^{-1}\delta \bar{X} - u)^2 + pu^2],$$

which is the desired result.

To conclude from (61), we consider c_1 from (55) in Lemma 16. If $p \geq c_1$ then the r.h.s. of (61) is larger than $c_1(\delta \bar{X})^2/q$ and we can conclude using the upper bound in (58).

When $p \leq c_1$ then we use (55) which ensures that

$$|I^\wedge \setminus I^x| \geq |I^\wedge|/2,$$

so that with probability at least $1/2$, Z^\wedge coincides with a $\text{Beta}_\alpha(I^\wedge)$ r.v. conditioned on being larger than its median \bar{X}^\wedge . Since the variance of the latter conditional law is of order $|I^\wedge|^2 = (\nabla X^\wedge)^2$, so that for some adequate choice of $c > 0$ we have we have

$$(63) \quad q \text{Var}(Z^\wedge) \geq cq(\nabla X^\wedge)^2 \geq c(1 - c_1)(\nabla X^\wedge)^2,$$

which concludes the proof. \square

5.3. *Successive hitting times.* To prove Proposition 15, our argument is to show that by time $t_{\delta/2} + N^2$, the area has become very small (smaller than N^{-1}) and then to use some brute force argument to show that τ cannot be much larger.

Our strategy to control the decay of the area requires several steps and we introduce the successive hitting times by the area of a sequence of well-chosen thresholds. For $\eta > 0$ small, we define

$$\mathcal{T}_i := \inf\{t \geq t_{\delta/2} : A_t \leq N^{\frac{3}{2}-i\eta}\}, \quad i \geq 1.$$

We first show that with large probability \mathcal{T}_2 is equal to $t_{\delta/2}$. Then setting $L := \min\{i \geq 1 : \frac{3}{2} - i\eta < -1\}$, we show that the increments $\mathcal{T}_i - \mathcal{T}_{i-1}$ are small for all $i \leq L$. The argument to control the increment $\mathcal{T}_i - \mathcal{T}_{i-1}$ differs for different ranges of i so that it is practical for us to introduce the following intermediate thresholds:

$$I := \max\left\{i \geq 1 : \frac{3}{2} - i\eta > 1\right\}, \quad K := \max\left\{i \geq 1 : \frac{3}{2} - i\eta > 1/4\right\}.$$

We now introduce the following events which allow us to split our proof in four parts:

$$\begin{aligned} \mathcal{B}_{(1)}^N &:= \{\mathcal{T}_2 = t_{\delta/2}\}, \\ \mathcal{B}_{(2)}^N &:= \{\forall i \in \llbracket 3, I \rrbracket, \mathcal{T}_i - \mathcal{T}_{i-1} \leq 2^{-i} N^2\}, \\ \mathcal{B}_{(3)}^N &:= \{\forall i \in \llbracket I + 1, K \rrbracket, \mathcal{T}_i - \mathcal{T}_{i-1} \leq 2^{-i} N^2\}, \\ \mathcal{B}_{(4)}^N &:= \{\forall i \in \llbracket K + 1, L \rrbracket, \mathcal{T}_i - \mathcal{T}_{i-1} \leq 2^{-i} N^2\}. \end{aligned} \tag{64}$$

We prove in the forthcoming Sections 5.4, 5.7, 5.8 and 5.9, respectively, that each of these four events holds with probability tending to 1. This implies in particular that

$$\lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{T}_L \leq t_{\delta/2} + N^2] = 1.$$

In Section 5.10, we use this last statement to conclude the proof of Proposition 15. The remaining Sections 5.5–5.6 are dedicated to the introduction of technical material which is used throughout Sections 5.7–5.10.

5.4. *Initial contraction of A_t and control of \mathcal{T}_2 .* The probability of $\mathcal{B}_{(1)}^N$ can be controlled using Markov’s inequality for the nonnegative random variable A_t . Indeed $\mathbb{E}[A_t]$ has an explicit expression in terms of the discrete heat equation. This argument does not exploit our specific coupling \mathbb{P} .

LEMMA 20. *For any $\eta > 0$ we have as $N \rightarrow \infty$*

$$\lim_{N \rightarrow \infty} \mathbb{P}(A_{t_{5\eta}} > N^{\frac{3}{2}-2\eta}) = 0.$$

In particular, if $\eta \leq \delta/10$ then $\lim_{N \rightarrow \infty} \mathbb{P}(\mathcal{B}_{(1)}^N) = 1$.

PROOF. Let us set $a(t, k) := \mathbb{E}[X_k^\wedge(t) - X_k^x(t)]$. From the expression (25), we deduce that

$$\partial_t a(t, k) = \frac{1}{2} \Delta a(t, k), \quad k \in \llbracket 1, N - 1 \rrbracket.$$

In addition we have $a(t, 0) = a(t, N) = 0$. Expanding $k \mapsto a(t, k)$ on the orthonormal basis of the discrete Laplacian as in (44) and estimating all the corresponding eigenvalues by the main eigenvalue, it follows that

$$\sum_{k=0}^N a(t, k)^2 \leq e^{-2 \text{gap}_N t} \sum_{k=0}^N a(0, k)^2.$$

By Cauchy–Schwarz’s inequality,

$$(65) \quad \mathbb{E}[A_t]^2 \leq N \sum_{k=0}^N a(t, k)^2 \leq N e^{-2 \text{gap}_N t} \sum_{k=0}^N a(0, k)^2 \leq N^4 e^{-2 \text{gap}_N t}.$$

Markov’s inequality then yields the asserted result. \square

5.5. *Technical preliminaries to control the probability of $\mathcal{B}_{(2)}^N, \mathcal{B}_{(3)}^N, \mathcal{B}_{(4)}^N$.* Before going into the specifics of each case let us introduce the common framework which allows us to control $\mathcal{T}_i - \mathcal{T}_{i-1}$ for $i \in \llbracket 3, L \rrbracket$. Our main idea is to exploit the fact that $(A_t)_{t \geq 0}$ is a supermartingale for which we have a reasonable control on the jumps. For such processes the hitting time can be estimated if one can control the angle bracket of the martingale, as shown in the following result from [12].

PROPOSITION 21. *Let $(M_t)_{t \geq 0}$ be a pure-jump supermartingale with bounded jump rate and jump amplitude. Given $a \in \mathbb{R}$ and $b \leq a$, set*

$$\tau_b := \inf\{t \geq 0 : M_t \leq b\}.$$

If the amplitude of the jumps of $(M_t)_{t \geq 0}$ is bounded above by d , then we have for any $v \geq 4d^2$

$$(66) \quad \mathbb{P}[\langle M \cdot \rangle_{\tau_b} \geq v \mid M_0 \leq a] \leq 8(a - b)v^{-1/2},$$

where $\langle M \cdot \rangle$ denotes the bracket of $(M_t)_{t \geq 0}$ (the predictable processes which compensates the square of the martingale part of M).

PROOF. The only difference with the statement of [12], Proposition 29, is that d is not necessarily below $a - b$. However, a careful inspection of the proof therein shows that the present statement holds (note that the proof therein relies on [12], Lemma 30, and this result does not need any modification to cover our setting). \square

We apply the previous proposition to the supermartingale $M_t = A_t$. Our idea is to combine the above result with estimates on the increments of the bracket of A , $\Delta_i \langle A \rangle := \langle A \cdot \rangle_{\mathcal{T}_i} - \langle A \cdot \rangle_{\mathcal{T}_{i-1}}$ in order to obtain upper-tail probability for the increments of the hitting times $\mathcal{T}_i - \mathcal{T}_{i-1}$. The control of $\Delta_i \langle A \rangle$ as a function of $\mathcal{T}_i - \mathcal{T}_{i-1}$ is technically involved. Our general strategy is to restrict ourselves to an event of large probability for which we have (for some adequate constant $C(i, N)$)

$$\Delta_i \langle A \rangle \geq C(i, N)(\mathcal{T}_i - \mathcal{T}_{i-1}).$$

The specific events that we have to consider is introduced in the following section.

5.6. *Restriction to the right set of events.* Our convenient event is the intersection of two events $\mathcal{A}_1^{(N)}$ and $\mathcal{A}_2^{(N)}$ that we now introduce (for these events and others we make the dependence in N appear only when necessary). Regarding \mathcal{A}_1 , we only impose that the increments of the higher interface are not too large “at all times”:

$$(67) \quad \mathcal{A}_1 := \left\{ \forall t \in [t_{\delta/2}, t_{\delta/2} + N^2] : \max_{k \in \llbracket 1, N-1 \rrbracket} \eta_k^\wedge(t) \leq 10 \log N \right\}.$$

That the probability of \mathcal{A}_1 goes to 1 will follow from the fact that the higher interface is close to equilibrium by Proposition 14 and from simple estimates under the invariant measure. To define \mathcal{A}_2 , we introduce the following events that impose some restrictions on the interfaces at

a given time t . The events \mathcal{C}_1 and \mathcal{C}_2 require respectively the gradients of the higher interface to be not too small, and the distance between the two interfaces to be not too large:

$$(68) \quad \begin{aligned} \mathcal{C}_1(t) &:= \left\{ \min_{k \in \llbracket 1, N-1 \rrbracket} \nabla X_k^\wedge(t) \geq (N \log N)^{-1/2} \right\}, \\ \mathcal{C}_2(t) &:= \left\{ \max_{k \in \llbracket 1, N-1 \rrbracket} |X_k^\wedge - X_k^x| \leq \sqrt{N} \log N \right\}. \end{aligned}$$

Given $x \in \Omega_N$, we let $(a_i(x))_{i=1}^{N-1}$ be the increasingly ordered sequence of the values $(\nabla x_k)_{k=1}^{N-1}$. Then we set

$$(69) \quad \mathcal{C}_3(t) := \left\{ \forall i \in \llbracket 1, N \rrbracket : (a_i(X^\wedge(t)))^2 \geq \frac{i}{N \log N} \right\}.$$

Finally, we define

$$(70) \quad \mathcal{C}_4(t) := \left\{ \forall i \in \llbracket 0, N - (\log N)^2 \rrbracket : \sum_{k=i+1}^{i+(\log N)^2} (\nabla X_k^\wedge(t))^2 \geq \frac{(\log N)^2}{100} \right\}.$$

The event \mathcal{A}_2 then requires that for a large proportion of the interval of time $[t_{\delta/2}, t_{\delta/2} + N^2]$, the four events $\mathcal{C}_i(t)$ are satisfied:

$$(71) \quad \mathcal{A}_2 := \left\{ \int_{t_{\delta/2}}^{t_{\delta/2} + N^2} \mathbf{1}_{\mathcal{C}_1(s) \cap \mathcal{C}_2(s) \cap \mathcal{C}_3(s) \cap \mathcal{C}_4(s)} \, ds \geq (1 - 2^{-(L+1)})N^2 \right\}.$$

Note that the probability of \mathcal{A}_2 still goes to 1 if $1 - 2^{-(L+1)}$ is replaced by any factor $1 - c \in [0, 1)$. However, for latter use we need this factor to be larger than $1 - 2^{-L}$, and this explains our particular choice $1 - 2^{-(L+1)}$.

PROPOSITION 22. *We have $\lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{A}_1^{(N)} \cap \mathcal{A}_2^{(N)}] = 1$.*

PROOF. Recall that $\pi_{N,\alpha}$ is the invariant measure of our dynamics, and let ν_N be the law on Ω_N^+ under which the η_k 's are independent $\Gamma(\alpha, \alpha)$ r.v. Without further mention, we will apply Lemma 10 that allows us to bound some functionals under $\pi_{N,\alpha}$ by the same functionals under ν_N .

We start with the event \mathcal{A}_1 . We have (recall that $\alpha \geq 1$)

$$(72) \quad \pi_{N,\alpha}(\exists k : \eta_k \geq 10 \log N) \leq C \sum_{k=1}^N \nu_N(\eta_k \geq 10 \log N) \leq N^{-8}.$$

By Proposition 14, it follows that

$$\lim_{N \rightarrow \infty} \sup_{t \geq t_{N,\delta/2}} \|P_t^\wedge - \pi_{N,\alpha}\|_{\text{TV}} = 0.$$

Therefore, it suffices to work with the process starting from the stationary measure $\pi_{N,\alpha}$: we denote by \mathbf{P} the law of such a process. Let us subdivide the interval $[t_{\delta/2}, t_{\delta/2} + N^2]$ into disjoint intervals $[t_i, t_{i+1}]$ of size N^{-5} : note that there are of order N^7 such intervals. Then, a standard argument on independent Poisson clocks ensures that the probability that on each interval $[t_i, t_{i+1}]$ there is no more than 1 resampling event is larger than $1 - CN^{-1}$ for some constant $C > 0$, hence on that event, at any time $s \in [t_{\delta/2}, t_{\delta/2} + N^2]$, $\eta_k(s)$ is equal to some $\eta_k(t_i)$ for some i . Moreover, a simple union bound combined with (72) shows that the probability that for all t_i and all k , $\eta_k(t_i) < 10 \log N$ goes to 1. Therefore,

$$\lim_{N \rightarrow \infty} \mathbf{P}(\exists t \in [t_{\delta/2}, t_{\delta/2} + N^2], \exists k \in \llbracket 1, N \rrbracket : \eta_k(t) \geq 10 \log N) = 0.$$

We turn to the event \mathcal{A}_2 . By Markov’s inequality, it suffices to show that for every $i \in \llbracket 1, 4 \rrbracket$

$$\lim_{N \rightarrow \infty} \sup_{t \geq t_{\delta, N}} \mathbb{P}(\mathcal{C}_i^{(N)}(t)^{\complement}) = 0.$$

To handle the events $\mathcal{C}_1, \mathcal{C}_3$ and \mathcal{C}_4 , Proposition 14 ensures that one can work under the stationary measure $\pi_{N, \alpha}$. We are going to make extensive use of the following tail estimates for the $\Gamma(2\alpha, \alpha)$ distribution:

$$(73) \quad \nu_N(\nabla x_k \leq t) = \nu_N(\eta_1 + \eta_2 \leq t) \leq C' t^{2\alpha} \leq C t^2 \quad \forall t \in [0, 1].$$

By Lemma 10, the bound is also valid under $\pi_{N, \alpha}$ (with a different constant). To control the probability of $\mathcal{C}_1^{\complement}$ it is sufficient to observe that by union bound and exchangeability of the increments

$$(74) \quad \pi_{N, \alpha}(\mathcal{C}_1^{\complement}) \leq N \pi_{N, \alpha} \left(\nabla x_1 < \frac{1}{\sqrt{N \log N}} \right) \leq C (\log N)^{-1}.$$

We turn to \mathcal{C}_4 . By union bound and exchangeability again we have

$$\begin{aligned} \pi_{N, \alpha}(\mathcal{C}_4^{\complement}) &\leq N \pi_{N, \alpha} \left(\sum_{k=1}^{(\log N)^2} (\nabla x_k)^2 \leq \frac{(\log N)^2}{100} \right) \\ &\leq C N \nu_N \left(\sum_{k=1}^{\frac{1}{2}(\log N)^2} (\nabla x_{2k})^2 \leq \frac{(\log N)^2}{100} \right) \\ &\leq C N (\nu_N[e^{-\lambda(\nabla x_2)^2}]) e^{\frac{\lambda}{50} \frac{(\log N)^2}{2}}, \end{aligned}$$

for all $\lambda \geq 0$. Since $\nu_N[(\nabla x_{2k})^2] = (4 + 2\alpha^{-1}) > 1$, it is simple to check that for small enough λ , we have $\mathbb{E}[e^{-\lambda(\nabla x_2)^2}] \leq e^{-\lambda}$ which suffices to conclude.

We now consider \mathcal{C}_3 . We let $b_i(x)$ denote the set of increasingly ordered values of $(\nabla x_{2k})_{k=1}^{N/2}$. We are going to show that $\pi_{N, \alpha}(\mathcal{C}_3^{\complement})$ tends to zero where

$$\mathcal{C}'_3 := \left\{ \forall i \leq N/4, b_i(x)^2 \geq \frac{4i}{N \log N} \right\}.$$

The same bound concerning the odd gradients $(\nabla x_{2k-1})_{k=1}^{N/2}$ is then sufficient to conclude. By union bound and exchangeability of the variables $(\nabla x_{2k})_{k=1}^{N/2}$, we have

$$\pi_{N, \alpha}((\mathcal{C}'_3)^{\complement}) \leq \sum_{i=1}^{N/4} \binom{N}{i} \pi_{N, \alpha} \left(\forall j \in \llbracket 1, i \rrbracket, (\nabla x_{2j})^2 \leq \frac{4i}{N \log N} \right).$$

Using Lemma 10 and (73), we obtain for some positive constant C

$$(75) \quad \pi_{N, \alpha}((\mathcal{C}'_3)^{\complement}) \leq \sum_{i=1}^{N/4} \frac{N^i}{i!} \left(\frac{C i}{N \log N} \right)^i \leq \frac{C'}{\log N}.$$

Regarding \mathcal{C}_2 , we first show that for N sufficiently large

$$(76) \quad \pi_{N, \alpha} \left(\max_{k \in \llbracket 1, N-1 \rrbracket} |x_k - k| > \sqrt{N} \log N / 2 \right) \leq N^{-1}.$$

By symmetry, we can restrict to $k \in \llbracket 0, N/2 \rrbracket$ and by Lemma 10, it suffices to prove this bound under ν_N . For every $k \in \llbracket 0, N/2 \rrbracket$ the law of x_k under ν_N is $\Gamma(k\alpha, \alpha)$. Using a Chernoff bound for all N large enough, we obtain

$$\nu_N(|x_k - k| > \sqrt{N} \log N / 2) \leq e^{-10 \log N}.$$

(Note that a finer upper bound would be $e^{-c(\log N)^2}$ for some small enough $c > 0$.) Consequently, (76) follows. By Proposition 14, we deduce that

$$(77) \quad \lim_{N \rightarrow \infty} \sup_{t \geq t_{\delta/2}} \mathbb{P} \left(\max_{k \in \llbracket 1, N-1 \rrbracket} |X_k^\wedge(t) - k| > \sqrt{N} \log N/2 \right) = 0.$$

By symmetry the same is valid for $X_k^\vee(t)$. By Proposition 7 $X_k^x(t)$ is stochastically dominated by $X_k^\wedge(t)$ and stochastically dominates $X_k^\vee(t)$. This implies that

$$(78) \quad \lim_{N \rightarrow \infty} \sup_{x \in \Omega_N} \sup_{t \geq t_{\delta/2}} \mathbb{P} \left(\max_{k \in \llbracket 1, N-1 \rrbracket} |X_k^x(t) - k| > \sqrt{N} \log N/2 \right) = 0$$

and the result is obtained by observing that $\mathcal{C}_2(t)^\complement$ is contained in the union of the two events in (77) and (78). \square

5.7. *Controlling the \mathcal{T}_i increments for $i \in \llbracket 3, I \rrbracket$.* We are now ready to prove the following lemma.

LEMMA 23. *The event $\mathcal{B}_{(2)}^N$ satisfies*

$$(79) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{B}_{(2)}^N] = 1.$$

Recall that for any $t \in [t_{\delta/2}, \mathcal{T}_I)$, the area satisfies $A_t > N^{1+\kappa}$ for some $\kappa > 0$. Our proof relies then on the following observation.

LEMMA 24. *For all $t \in [t_{\delta/2}, t_{\delta/2} + N^2]$, on the event $\mathcal{A}_1 \cap \mathcal{C}_2(t) \cap \mathcal{C}_4(t) \cap \{A_t \geq N(\log N)^4\}$ we have*

$$(80) \quad \partial \langle A. \rangle_t \geq \frac{A_t}{\sqrt{N} \log N}.$$

PROOF OF LEMMA 23. Proposition 21 applied to the supermartingale $(A_{\mathcal{T}_{i-1}+s})_{s \geq 0}$ (whose maximal jump size is N) with $v = N^{3-2(i-1)\eta} \log N$, yields that $\lim_{N \rightarrow \infty} \mathbb{P}(\mathcal{A}_3^{(N)}) = 1$ where

$$\mathcal{A}_3 := \bigcap_{i=3}^I \{ \Delta_i \langle A \rangle < N^{3-2(i-1)\eta} \log N \}.$$

To conclude, we only need to show that

$$\mathcal{B}_{(2)}^N \supset \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3 \cap \mathcal{B}_{(1)}^N.$$

We proceed by contradiction. On the event $\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3$ consider the smallest integer j in $\llbracket 3, I \rrbracket$ such that $\mathcal{T}_j - \mathcal{T}_{j-1} > 2^{-j} N^2$. Applying Lemma 24 on the interval $[\mathcal{T}_{j-1}, \mathcal{T}_{j-1} + 2^{-j} N^2] \subset [t_{\delta/2}, t_{\delta/2} + N^2]$, on which $A_t \geq N^{3/2-j\eta}$ we obtain

$$(81) \quad \begin{aligned} \Delta_j \langle A \rangle &\geq \langle A \rangle_{\mathcal{T}_{j-1}+2^{-j}N^2} - \langle A \rangle_{\mathcal{T}_{j-1}} \\ &\geq \frac{N^{3/2-j\eta}}{\sqrt{N} \log N} \int_{\mathcal{T}_{j-1}}^{\mathcal{T}_{j-1}+2^{-j}N^2} \mathbf{1}_{\mathcal{C}_2(s) \cap \mathcal{C}_4(s)} \, ds \geq \frac{2^{-(j+1)} N^2 N^{3/2-j\eta}}{\sqrt{N} \log N}, \end{aligned}$$

where the last inequality uses the definition of \mathcal{A}_2 to assert that

$$\int_{\mathcal{T}_{j-1}}^{\mathcal{T}_{j-1}+2^{-j}N^2} \mathbf{1}_{\mathcal{C}_4(s) \cap \mathcal{C}_2(s)} \, ds \geq (2^{-j} - 2^{-(L+1)}) N^2. \quad \square$$

PROOF OF LEMMA 24. Since $A_t \geq N(\log N)^4$, we have for N large enough

$$\sum_{k=1}^{N-1} \delta X_k(t) \mathbf{1}_{\{\delta X_k(t) \leq 20(\log N)^3\}} \leq 20N(\log N)^3 \leq \frac{1}{2} A_t.$$

Since we work on $\mathcal{C}_2(t)$, we get

$$(82) \quad \frac{1}{2} A_t \leq \left(\max_{k \in \llbracket 1, N-1 \rrbracket} \delta X_k(t) \right) \#\{k : \delta X_k(t) > 20(\log N)^3\}$$

$$(83) \quad \leq \sqrt{N} \log N \#\{k : \delta X_k(t) > 20(\log N)^3\}.$$

The bound on the gradients given by \mathcal{A}_1 ensures that

$$\delta X_{k-1}(t) > \delta X_k(t) - 10 \log N$$

so that for N large enough we get

$$\{\delta X_k(t) > (\log N)^3\} \Rightarrow \left\{ \delta \bar{X}_k(t) > \frac{(\log N)^3}{3} \right\},$$

and

$$\{\delta X_\ell(t) > 20(\log N)^3\} \Rightarrow \{\forall k \in \llbracket \ell - (\log N)^2, \ell \rrbracket, \delta X_k(t) > (\log N)^3\}.$$

Consequently by Proposition 18 on the event \mathcal{A}_1

$$\begin{aligned} \partial_t \langle A. \rangle_t &\geq c \sum_{k=1}^{N-1} (\nabla X_k^\wedge(t))^2 \mathbf{1}_{\{\delta X_k(t) > (\log N)^3\}} \\ &\geq c \sum_{k=1}^{N-1} (\nabla X_k^\wedge(t))^2 \mathbf{1}_{\{\exists \ell \in \llbracket k, k + (\log N)^2 \rrbracket : \delta X_\ell(t) > 20(\log N)^3\}}. \end{aligned}$$

The set over which the sum is taken on the r.h.s. can be decomposed into connected components of size at least $(\log N)^2$. On $\mathcal{C}_4(t)$, we thus deduce that there exists $c' > 0$ such that

$$\begin{aligned} \partial_t \langle A. \rangle_t &\geq c' \#\{k : \exists \ell \in \llbracket k, k + (\log N)^2 \rrbracket, \delta X_\ell(t) > 20(\log N)^3\} \\ &\geq c' \#\{\ell : \delta X_\ell(t) > 20(\log N)^3\} \\ &\geq \frac{c'}{2} \frac{A_t}{\sqrt{N} \log N}, \end{aligned}$$

where in the last line we simply used (82). \square

5.8. *Controlling the \mathcal{T}_i increments for $i \in \llbracket I + 1, K \rrbracket$.* We can prove now

LEMMA 25. *We have*

$$(84) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{B}_{(3)}^N] = 1.$$

We follow essentially the same line of proof as in the previous section, but using a different inequality for the bracket derivative. We also need an extra trick to control the maximal amplitude of jumps. Recall that for every $t \in [\mathcal{T}_I, \mathcal{T}_K)$, we have $A_t > 2N^{1/4}$.

LEMMA 26. *Fix $\varepsilon > 0$. On the event $\mathcal{A}_1 \cap \mathcal{C}_3(t) \cap \mathcal{C}_4(t) \cap \{A_t > 2N^\varepsilon\}$, we have for all $t \in [t_{\delta/2}, t_{\delta/2} + N^2]$*

$$\partial_t \langle A. \rangle_t \geq N^{\frac{\varepsilon}{2}} \wedge \frac{A_t^2}{N^{1+3\varepsilon}}.$$

PROOF OF LEMMA 25. In order to diminish the restriction on v imposed in Proposition 21, we introduce the following stopping times with the objective of considering a process with smaller jump amplitude:

$$\mathcal{R}_i := \inf \left\{ t \geq \mathcal{T}_{i-1} : \max_{k \in \llbracket 1, N-1 \rrbracket} \nabla X_k^\wedge(t) > 10 \log N \text{ or } A_t > N^{\frac{3}{2}-(i-2)\eta} \right\}.$$

We consider the supermartingale $(A_{(\mathcal{T}_{i-1}+t) \wedge \mathcal{R}_i})_{t \geq 0}$. For every $i \leq L$, its maximal jump size up to time $\mathcal{T}_i - \mathcal{T}_{i-1}$ is bounded by

$$10 \log N + N^{\frac{3}{2}-(i-2)\eta}.$$

Indeed, the maximal variation of the area due to an update of site k at time t is bounded above by

$$X_{k+1}^\wedge(t_-) - X_{k-1}^x(t_-) \leq \delta X_{k-1}(t_-) + \nabla X_k^\wedge(t_-) \leq A_{t_-} + 10 \log N.$$

We consider now the event

$$\begin{aligned} \mathcal{A}_{4,1}^{(N)} &:= \{ \forall i \in \llbracket I+1, L \rrbracket, \forall t \geq \mathcal{T}_{i-1}, A_t \leq N^{\frac{3}{2}-(i-2)\eta} \}, \\ (85) \quad \mathcal{A}_{4,2}^{(N)} &:= \{ \forall i \in \llbracket I+1, L \rrbracket, \langle A \rangle_{\mathcal{T}_i \wedge \mathcal{R}_i} - \langle A \rangle_{\mathcal{T}_{i-1} \wedge \mathcal{R}_i} \leq v_i \}, \\ \mathcal{A}_4^{(N)} &:= \mathcal{A}_{4,1}^{(N)} \cap \mathcal{A}_{4,2}^{(N)} \end{aligned}$$

with $v_i := 4(10 \log N + N^{\frac{3}{2}-(i-2)\eta})^2$. The standard union bound and a supermartingale version of Doob’s maximal inequality (sometimes referred to as Ville’s inequality) show that $\mathbb{P}[(\mathcal{A}_{4,1}^{(N)})^c] \leq 3(L - I)N^{-\eta}$, and hence converges to 0. Now, applying Proposition 21 to $(A_{(\mathcal{T}_{i-1}+t) \wedge \mathcal{R}_i})_{t \geq 0}$ with $v = v_i$ yields $\lim_{N \rightarrow \infty} \mathbb{P}(\mathcal{A}_{4,2}^{(N)}) = 1$.

Our final step is to prove

$$\mathcal{B}_{(3)} \supset \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_4 \cap \mathcal{B}_{(1)} \cap \mathcal{B}_{(2)}.$$

We place ourselves on the event $\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_4 \cap \mathcal{B}_{(2)}$ and we proceed by contradiction. Consider the smallest element j of $\llbracket 1, K \rrbracket$ such that $\mathcal{T}_j - \mathcal{T}_{j-1} \geq 2^{-j}N^2$ (note that as we are on $\mathcal{B}_{(2)}^N$ we must have $j \geq I + 1$). Because we are on the event $\mathcal{A}_1 \cap \mathcal{A}_{4,1}$, the reader can check that we must have

$$\mathcal{R}_j \geq t_{\delta/2} + N^2 \geq \mathcal{T}_{j-1} + 2^{-j}N^2.$$

Hence by Lemma 26, $\mathcal{A}_{4,2}$ and \mathcal{A}_2 imply that

$$\begin{aligned} (86) \quad v_j &\geq \langle A \rangle_{\mathcal{T}_{j-1}+2^{-j}N^2} - \langle A \rangle_{\mathcal{T}_{j-1}} \\ &\geq (N^{\frac{1}{8}} \wedge N^{\frac{5}{4}-2j\eta}) \int_{\mathcal{T}_{j-1}}^{\mathcal{T}_{j-1}+2^{-j}N^2} \mathbf{1}_{\mathcal{C}_3(s) \cap \mathcal{C}_4(s)} \, ds \geq 2^{-j-1} (N^{\frac{17}{8}} \wedge N^{\frac{13}{4}-2j\eta}), \end{aligned}$$

where for the inequality we proceed as in the proof of Lemma 23 to show that the integral is larger than $2^{-j-1}N^2$. This yields the desired contradiction. \square

We are left with proving Lemma 26. The first step is the following technical estimate.

LEMMA 27. Given $B > 0$, let $(a_i)_{i=1}^n$ be an increasing sequence of positive real numbers with $\max a_i \leq B$, and $(b_i)_{i=1}^n$ an arbitrary sequence in \mathbb{R}_+ with $\max b_i \leq B$, and let us set $\sigma := \sum_{i=1}^n b_i$, and $K := \lfloor \sigma/B \rfloor$. Then we have

$$(87) \quad \sum_{i=1}^n a_i (a_i \wedge b_i) \geq \sum_{i=1}^K a_i^2.$$

If $K = 0$, we have

$$(88) \quad \sum_{i=1}^n a_i(a_i \wedge b_i) \geq a_1(a_1 \wedge \sigma).$$

PROOF. Let us start with the case when $\sigma \geq B$. Using that a_i and b_i are bounded by B we have

$$\sum_{i=1}^n a_i(a_i \wedge b_i) \geq \sum_{i=1}^n \frac{a_i^2 b_i}{B}.$$

Now the reader can check that if σ is fixed, the r.h.s. is minimized when $b_i = B$ for $i \leq K$, $b_{K+1} = \sigma - BK$, and $b_i = 0$ for $i > K + 1$. When $\sigma < B$, it is sufficient to check (88) when $\sigma \leq a_1$ by monotonicity, and in that case we have

$$\sum_{i=1}^n a_i(a_i \wedge b_i) = \sum_{i=1}^n a_i b_i \geq a_1 \sum_{i=1}^n b_i. \quad \square$$

PROOF OF LEMMA 26. First assume that $\max_k |X_k^\wedge(t) - X_k^x(t)| > N^\varepsilon$. As in the proof of Lemma 24, the bound on the gradients given by \mathcal{A}_1 ensures that

$$\delta X_\ell(t) > N^\varepsilon \quad \Rightarrow \quad \left(\forall k \in \llbracket \ell - N^{\frac{2\varepsilon}{3}}, \ell \rrbracket, \delta X_k(t) > \frac{N^\varepsilon}{2} \right).$$

Consequently, if we let ℓ be an integer for which $\delta X_\ell(t) > N^\varepsilon$ we get the bound

$$\partial_t \langle A. \rangle_t \geq c \sum_{k \in \llbracket \ell - N^{\frac{2\varepsilon}{3}}, \ell \rrbracket} (\nabla X_k^\wedge(t))^2.$$

By $\mathcal{C}_4(t)$, we thus deduce that

$$\partial_t \langle A. \rangle_t \geq N^{\frac{\varepsilon}{2}}.$$

Next, assume that $\max_k |X_k^\wedge(t) - X_k^x(t)| \leq N^\varepsilon$. Set $B = N^\varepsilon$. We apply Lemma 27 with $(a_i)_{i=1}^N$ being the ordered sequence of the $(\nabla X_k^\wedge(t))_{k=1}^N$ and $(b_i)_{i=1}^N$ being the corresponding $\delta \bar{X}_k(t)$'s. Since $\sum_k \delta \bar{X}_k(t) \geq A_t/2$ and since $A_t \geq 2N^\varepsilon$, the lemma combined with $\mathcal{C}_3(t)$ and Proposition 18 gives the following lower bound:

$$\partial_t \langle A. \rangle \geq c \sum_{i=1}^{\lfloor (A_t N^{-\varepsilon})/2 \rfloor} a_i^2 \geq \frac{c'(A_t N^{-\varepsilon})^2}{N \log N} \geq \frac{A_t^2}{N^{1+3\varepsilon}}. \quad \square$$

5.9. *Controlling the \mathcal{T}_i increments for $i \in \llbracket K + 1, L \rrbracket$.* Following our plan, we now prove the following lemma.

LEMMA 28. *We have*

$$(89) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{B}_{(4)}^N] = 1.$$

This time we use the following control for the martingale bracket, which is almost immediate to prove the following lemma.

LEMMA 29. *On the event $\mathcal{A}_1 \cap \mathcal{C}_1(t)$ we have*

$$\partial_t \langle A. \rangle_t \geq \frac{1 \wedge (\sqrt{N} A_t)}{N (\log N)^2}.$$

PROOF. Combining Proposition 18 and Lemma 27, we obtain

$$\partial_t \langle A. \rangle_t \geq ca_1(a_1 \wedge A_t) \geq \frac{c}{\sqrt{N \log N}} \left(\frac{1}{\sqrt{N \log N}} \wedge A_t \right). \quad \square$$

PROOF OF LEMMA 28. We only need to show that

$$\mathcal{B}_{(4)}^N \supset \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_4 \cap \mathcal{B}_{(1)}^N \cap \mathcal{B}_{(2)}^N \cap \mathcal{B}_{(3)}^N.$$

A contradiction can be obtained exactly as in the proof of Lemma 28. More precisely assuming that $\mathcal{B}_{(4)}^N$ does not hold we obtain that for some $j \in \llbracket K + 1, L \rrbracket$

$$(90) \quad v_j \geq 2^{-j-1} N (\log N)^{-2} (1 \wedge N^{2-j\eta}). \quad \square$$

5.10. Proof of Proposition 15. We finally conclude the proof by proving

$$(91) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\tau \leq t_{\delta/2} + N^2 + 2 \log N] = 1.$$

We first need to restrict ourself to an adequate event of large probability. Using the fact that $(A_{\mathcal{T}_L+t})_{t \geq 0}$ is a supermartingale, combining the Martingale stopping theorem and Doob's inequality we obtain that

$$(92) \quad \mathbb{P}[\exists t \geq 0, A_{\mathcal{T}_L+t} \geq N^{-1}] \leq 3N^{5/2-L\eta}.$$

Since Lemmas 20, 23, 25 and 28 assert that $\mathcal{T}_L \leq t_{\delta/2} + N^2$ with large probability, we have $\lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{A}_{5,1}^{(N)}] = 1$ where

$$\mathcal{A}_{5,1}^{(N)} := \{A_{t_{\delta/2}+N^2} < N^{-1}\}.$$

We also need to make sure that the gradients in our dynamics are not too small.

LEMMA 30. Setting

$$\mathcal{A}_{5,2}^{(N)} := \left\{ \min_{k \in \llbracket 1, N-1 \rrbracket, s \in [0, 2 \log N]} \nabla X_k^\wedge(t_{\delta/2} + N^2 + s) > \frac{1}{\sqrt{N \log N}} \right\},$$

we have

$$(93) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\mathcal{A}_{5,2}^{(N)}] = 1.$$

PROOF. First of all, by Proposition 14 it suffices to control the probability of this event for the stationary process in the time interval $[0, 2 \log N]$, which we denote by \mathbf{X} for the rest of this proof. Recalling (73) we have

$$(94) \quad \pi_{N,\alpha} \left(\nabla x_k < \frac{1}{\sqrt{N \log N}} \right) \leq CN^{-1} (\log N)^{-2}.$$

Now for each $k \in \llbracket 1, N - 1 \rrbracket$, consider the ordered set of update times $(t_i^{(k)})_{i \geq 1}$ of the coordinate k for our dynamics. We let n_k be the number of updates of the site k occurring in the time interval $[0, 2 \log N]$.

As the dynamics is of heat-bath type, for every i and k , $\mathbf{X}(t_i^{(k)})$ is distributed like $\pi_{N,\alpha}$. This is also valid conditionally on the realization $(t_i^{(j)})_{i,j}$ of the update times. Hence, considering that ∇X_k can only be altered at times $t_i^{(k \pm 1)}$, from the union bound and (94),

$$\begin{aligned} & \mathbb{P}[\exists t \in [0, 2 \log N], \exists k \in \llbracket 1, N - 1 \rrbracket : \nabla X_k^\wedge(t) \leq 1/(\sqrt{N \log N}) \mid (t_i^{(j)})_{i,j}] \\ & \leq 2CN^{-1} (\log N)^{-2} \sum_{k=1}^{N-1} (n_k + 1). \end{aligned}$$

Taking the expectation of the above we obtain the desired result. \square

For every site $k \in \llbracket 1, N - 1 \rrbracket$, let $(t_i^{(k)})_{i=1}^{n_k}$ denote the random set of update times occurring in the interval $[t_{\delta/2} + N^2, t_{\delta/2} + N^2 + 2 \log N]$. We set

$$\mathcal{A}_{5,3}^{(N)} := \bigcap_{k=1}^{N-1} \{n_k \in [1, 10 \log N]\}.$$

The probability of $\mathcal{A}_{5,3}^{(N)}$ goes to one (by a standard coupon collector argument for the lower bound while the upper bound is a direct union bound) and one can thus safely restrict oneself to the event $\mathcal{A}_5 := \bigcap_{i=1}^3 \mathcal{A}_{5,i}$.

To conclude the proof, we are going to show that

$$(95) \quad \lim_{N \rightarrow \infty} \mathbb{P}[\exists k \in \llbracket 1, N - 1 \rrbracket, \exists i \leq n_k : X_k^\wedge(t_i^{(k)}) \neq X_k^x(t_i^{(k)})] = 0.$$

We are in fact going to prove an upper bound for this probability conditioned on both $(t_i^{(k)})_{i=1}^{n_k}$ and the state of the system at the initial time $t_{\delta/2} + N^2$. We use the short-hand notation $\tilde{\mathbb{P}}, \tilde{\mathbb{E}}$ for this conditional probability and the corresponding expectation.

Say that an update at time $t_i^{(k)}$ is *successful* if $X_k^\wedge(t_i^{(k)}) = X_k^x(t_i^{(k)})$. The strategy is to work recursively on the successive update times, and to use the following fact. If all the previous updates have been successful, then using Lemma 16 the probability that the next update (occurring at site k say) is not successful is bounded from above by (a constant times) $\delta \bar{X}_k$ divided by the largest gradient, this ratio being bounded by $\delta \bar{X}_k \sqrt{N} \log N$ on the event $\mathcal{A}_{5,2}$. Since there are at most $10 \log N$ updates per site, since $\sum_k \delta \bar{X}_k$ is bounded by twice the area and since the area is nonincreasing as long as the updates are all successful, we deduce that the probability (95) that there exists an unsuccessful update is bounded by $A_{t_{\delta/2} + N^2} \sqrt{N} \log N < N^{-1/2} \log N$ on the event $\mathcal{A}_{5,1}$. Thus, one can conclude. To put these heuristic observations on a firm ground, we need to introduce some notations.

Let $\tilde{\tau}$ be, among the set of update times $(t_i^{(k)})_{i=1}^{n_k}$, the time of the first unsuccessful update: on the event that there are no unsuccessful updates, we set arbitrarily $\tilde{\tau} := t_{\delta/2} + N^2 + 2 \log N$. Note that the event $\mathcal{A}_{5,1} \cap \mathcal{A}_{5,3}$ is measurable w.r.t. the sigma-field generated by $(t_i^{(k)})_{i=1}^{n_k}$ and the state of the system at the initial time $t_{\delta/2} + N^2$. Since the probability of this event goes to 1, it suffices to show that

$$(96) \quad \lim_{N \rightarrow \infty} \mathbb{P}(\mathbf{1}_{\{\mathcal{A}_{5,1} \cap \mathcal{A}_{5,3}\}} \tilde{\mathbb{P}}(\tilde{\tau} < t_{\delta/2} + N^2 + 2 \log N)) = 0,$$

in order to deduce that the merging time τ satisfies

$$\lim_{N \rightarrow \infty} \mathbb{P}(\tau \geq t_{\delta/2} + N^2 + 2 \log N) = 0.$$

To prove (96), we set

$$\mathcal{A}_{5,2}^{(i,k)} := \left\{ \forall s \in [t_{\delta/2} + N^2, t_i^{(k)}), \nabla X_k^\wedge(s) > \frac{1}{\sqrt{N} \log N} \right\}.$$

Then, almost surely

$$\begin{aligned} \tilde{\mathbb{P}}(\tilde{\tau} < t_{\delta/2} + N^2 + 2 \log N) &= \tilde{\mathbb{P}}\left[\bigcup_{i,k} \{\tilde{\tau} = t_i^{(k)}\}\right] \\ &\leq \tilde{\mathbb{P}}\left[\bigcup_{i,k} (\mathcal{A}_{5,2}^{(i,k)})^c\right] + \tilde{\mathbb{P}}\left[\bigcup_{i,k} \{\tilde{\tau} = t_i^{(k)}\} \cap \mathcal{A}_{5,2}^{(i,k)}\right] \\ &\leq \tilde{\mathbb{P}}[(\mathcal{A}_{5,2})^c] + \sum_{i,k} \tilde{\mathbb{P}}[\{\tilde{\tau} = t_i^{(k)}\} \cap \mathcal{A}_{5,2}^{(i,k)}]. \end{aligned}$$

The first term on the r.h.s. goes to 0 by Lemma 30. Regarding the second term, we argue as follows. Recall that \mathcal{F}_t is the sigma-field generated by the system up to time t , and let $\mathcal{F}_{t-} = \sigma(\bigcup_{s < t} \mathcal{F}_s)$. Using Lemma 16 at the second line, for all i and k we have

$$\begin{aligned} \tilde{\mathbb{P}}[\{\tilde{\tau} = t_i^{(k)}\} \cap \mathcal{A}_{5,2}^{(i,k)}] &= \tilde{\mathbb{E}}[\tilde{\mathbb{E}}[\mathbf{1}(X_k^\wedge(t_i^{(k)}) \neq X_k^x(t_i^{(k)})) \mid \mathcal{F}_{(t_i^{(k)})-}] \mathbf{1}_{\mathcal{A}_{5,2}^{(i,k)} \cap \{\tilde{\tau} \geq t_i^{(k)}\}}] \\ &\leq \tilde{\mathbb{E}}\left[C \frac{\delta \bar{X}_k((t_i^{(k)})_-)}{\nabla X_k^\wedge((t_i^{(k)})_-)} \mathbf{1}_{\mathcal{A}_{5,2}^{(i,k)} \cap \{\tilde{\tau} \geq t_i^{(k)}\}}\right] \\ &\leq C\sqrt{N} \log N \delta \bar{X}_k(t_{\delta/2} + N^2). \end{aligned}$$

Consequently,

$$\sum_{i,k} \tilde{\mathbb{P}}[\{\tilde{\tau} = t_{i,k}\} \cap \mathcal{A}_{5,2}^{(i,k)}] \leq \sum_{i,k} C\sqrt{N} \log N \delta \bar{X}_k(t_{\delta/2} + N^2),$$

and on the event $\mathcal{A}_{5,1} \cap \mathcal{A}_{5,3}$ this last expression is bounded by

$$C'\sqrt{N}(\log N)^2 A_{t_{\delta/2}+N^2} \leq N^{-1/3},$$

for some new constant $C' > 0$. This concludes the proof of (96).

6. From the top down to equilibrium. The goal of this section is to prove Proposition 14, that is: when started from the maximal configuration $\wedge = (N, \dots, N)$, setting $t_\delta = (1 + \delta) \frac{\log N}{2 \text{gap}_N}$, one has

$$(97) \quad \lim_{N \rightarrow \infty} \|P_{t_\delta}^\wedge - \pi_{N,\alpha}\|_{\text{TV}} = 0,$$

for any fixed $\delta > 0$.

Inspired by a strategy that was introduced in [13] in the context of the adjacent interchange process, we shall base our proof on a two-scale argument, that can be roughly described as follows. For any integer $K \geq 2$, consider the $K - 1$ particles with labels $u_i := \lfloor iN/K \rfloor$, $i = 1, \dots, K - 1$. These will be called the *special particles*. The proof of (97) consists of three main steps.

Step 1. Starting from \wedge , after a time $t = t_{\delta/2}$, if K is fixed and N tends to ∞ , then the joint distribution of the positions of the special particles

$$(98) \quad Y_i(t) = X_{u_i}(t), \quad i = 1, \dots, K - 1$$

is arbitrarily close to the corresponding equilibrium distribution, see Proposition 39 below. This step is based on a subtle use of the FKG inequality together with the control of the expected value of the variables $Y_i(t)$.

Step 2. Consider the censored dynamics obtained by freezing the positions of the special particles and letting the rest of the particles evolve as usual. We will show that if K is taken proportional to δ^{-1} , uniformly in the initial condition, the censored dynamics at time $s_\delta := \delta \frac{\log N}{2 \text{gap}_N}$ has essentially reached the conditional equilibrium given by $\pi_{N,\alpha}$ conditioned on the positions of the special particles. For this step, it will be sufficient to exploit an upper bound on the mixing time that is tight up to a constant factor as, for example, the one obtained in [22] in the case $\alpha = 1$, see (13) above.

Step 3. We combine the results in the previous two steps to obtain the desired conclusion. The key point is that the distance to equilibrium at time $t_\delta = t_{\delta/2} + s_{\delta/2}$ appearing in (97) satisfies

$$(99) \quad \|P_{t_\delta}^\wedge - \pi_{N,\alpha}\|_{\text{TV}} \leq \|P_{t_\delta,*}^\wedge - \pi_{N,\alpha}\|_{\text{TV}},$$

where the distribution $P_{t_{\delta},*}^{\wedge}$ is obtained by running the standard dynamics, starting from \wedge , for a time $t_{\delta/2}$ and then by running from there the censored dynamics for a time $s_{\delta/2}$. This step requires an adaptation to our continuous setting of the so-called censoring inequality of Peres and Winkler [18].

We start developing the above program with a discussion of the censoring inequality. We then move to the proof of the mentioned steps in the given order.

6.1. Censoring lemma. The censoring inequality established by Peres and Winkler [18] allows one to compare the distance to equilibrium at time t for the process under consideration with the distance to equilibrium at time t for a censored process in which some of the updates have been omitted according to a given censoring scheme. In the context of Glauber dynamics for monotone, finite state spin systems, their argument rests on the following two key properties of a monotone dynamics: If the initial state has a distribution ν whose density w.r.t. the equilibrium measure π is increasing, then for any $t \geq 0$, the distribution νP_t of the state at time t satisfies:

- (1) νP_t has an increasing density w.r.t. π ,
- (2) νP_t is stochastically lower than the distribution of the state of the censored dynamics, say $\nu P_{t,*}$, for any valid censoring scheme.

Properties 1 and 2 then allow one to prove the censoring inequality

$$(100) \quad \|\nu P_t - \pi\|_{\text{TV}} \leq \|\nu P_{t,*} - \pi\|_{\text{TV}}.$$

We shall follow the same line of reasoning here. However, a technical problem arises with respect to the usual discrete spin setting: to prove the censoring inequality (99) we need to start with the distribution $\nu = \delta_{\wedge}$ which has no density w.r.t. equilibrium, while the strategy outlined above is crucially based on the existence of such a density. Notice also that $P_t^{\wedge} = \delta_{\wedge} P_t$ has no density w.r.t. equilibrium at any time $t \geq 0$ and so one cannot get around this problem by regularising the measure with a burn-in time. We shall need a more general version of the above properties which extends to a certain family of measures with a singular part.

We start by defining the latter. For $k \in \llbracket 1, N \rrbracket$, define the nested sets

$$(101) \quad \Omega_{k,N} = \{x \in \Omega_N : x_k = N\}.$$

Thus, $\Omega_{1,N} = \{\wedge\}$ is the maximal configuration, while $\Omega_{N,N} = \Omega_N$ is the whole set of particle positions. Let $\pi_{k,\alpha}$ denote the probability measure supported on $\Omega_{k,N}$ defined as the law of the random vector (x_1, \dots, x_{N-1}) of the partial sums $x_j = \sum_{i=1}^j \eta_i$ where η_1, \dots, η_k are i.i.d. with distribution $\Gamma(\alpha, \lambda)$, for some arbitrary $\lambda > 0$, conditioned to $\sum_{i=1}^k \eta_i = N$, while $\eta_{k+1}, \dots, \eta_N \equiv 0$. Notice that this notation is consistent with our notation $\pi_{N,\alpha}$ for the equilibrium measure on Ω_N . If a probability measure μ supported on $\Omega_{k,N}$ is absolutely continuous w.r.t. $\pi_{k,\alpha}$, we write $\frac{d\mu}{d\pi_{k,\alpha}}$ for the corresponding density, and say that μ belongs to the family \mathcal{S}_k if $\frac{d\mu}{d\pi_{k,\alpha}}$ is an increasing function. In particular, $\mathcal{S}_1 = \{\delta_{\wedge}\}$, while \mathcal{S}_N coincides with the set of distributions on Ω_N with an increasing density with respect to equilibrium. Finally, we define the family of measures \mathcal{S} consisting of all probability measures μ on Ω_N such that

$$(102) \quad \mu = \sum_{k=1}^N \gamma_k \mu_k, \quad \mu_k \in \mathcal{S}_k,$$

for some $\gamma_k \geq 0$ with $\sum_{k=1}^N \gamma_k = 1$. Notice that (102) is a decomposition into mutually singular measures, since if $1 \leq j < k \leq N$, then $\mu_j(\Omega_{j,N}) = 1$ while $\mu_k(\Omega_{j,N}) = 0$. The following lemma is a key fact about the set \mathcal{S} .

LEMMA 31. *If μ, ν are two probability measures on Ω_N such that $\mu \in \mathcal{S}$ and $\mu \leq \nu$, then*

$$(103) \quad \|\mu - \pi_{N,\alpha}\|_{TV} \leq \|\nu - \pi_{N,\alpha}\|_{TV}.$$

PROOF. Write

$$\mu = \sum_{j=1}^N \gamma_j \mu_j = (1 - \gamma_N) \mu' + \gamma_N \mu_N,$$

where $\mu' = (1 - \gamma_N)^{-1} \sum_{j=1}^{N-1} \gamma_j \mu_j$ is singular w.r.t. $\pi_{N,\alpha}$ while μ_N is absolutely continuous w.r.t. $\pi_{N,\alpha}$, with an increasing density φ_N . We have

$$\begin{aligned} \|\mu - \pi_{N,\alpha}\|_{TV} &= \sup_{C \in \mathcal{B}(\Omega_{N-1,N}), A \in \mathcal{B}(\Omega_{N-1,N}^c)} \mu(A \cup C) - \pi_{N,\alpha}(A \cup C) \\ &= \sup_{C \in \mathcal{B}(\Omega_{N-1,N}), A \in \mathcal{B}(\Omega_{N-1,N}^c)} (1 - \gamma_N) \mu'(C) + \gamma_N \mu_N(A) - \pi_{N,\alpha}(A) \\ &= 1 - \gamma_N + \sup_{A \in \mathcal{B}(\Omega_{N-1,N}^c)} \gamma_N \mu_N(A) - \pi_{N,\alpha}(A). \end{aligned}$$

We then define $A = \{x \in \Omega_N : \varphi_N(x) \geq 1/\gamma_N\}$. It is easy to check that this event maximizes the second term on the r.h.s. of the last equation. Therefore, setting $B = A \cup \Omega_{N-1,N}$, and observing that one has $\mu'(B) = \mu'(\Omega_{N-1,N}) = 1$ and $\mu_N(B) = \mu_N(A)$, $\pi_{N,\alpha}(B) = \pi_{N,\alpha}(A)$, we deduce that

$$\|\mu - \pi_{N,\alpha}\|_{TV} = \mu(B) - \pi_{N,\alpha}(B).$$

Since A and $\Omega_{N-1,N}$ are increasing, the set B is also increasing. Therefore, $\mu(B) \leq \nu(B)$ and

$$\|\mu - \pi_{N,\alpha}\|_{TV} \leq \nu(B) - \pi_{N,\alpha}(B) \leq \|\nu - \pi_{N,\alpha}\|_{TV}. \quad \square$$

Let $Q_i : L^2(\Omega_N, \pi_{N,\alpha}) \mapsto L^2(\Omega_N, \pi_{N,\alpha})$, $i = 1, \dots, N - 1$, denote the orthogonal projection onto functions that do not depend on the position of the i th particle:

$$Q_i f(x) = \pi_{N,\alpha}[f \mid x_j, j \neq i].$$

If μ is a probability on Ω_N , we write μQ_i for the probability measure defined by

$$\mu Q_i(f) = \int \mu(dx) Q_i f(x).$$

This is the distribution obtained from μ after one update at i .

LEMMA 32. *If $\mu \in \mathcal{S}$ then for any $i = 1, \dots, N - 1$:*

- (1) $\mu Q_i \in \mathcal{S}$;
- (2) μQ_i is stochastically lower than μ .

PROOF. Pick $\mu_k \in \mathcal{S}_k$. If $i > k$, then $\mu_k Q_i = \mu_k$, since $x_k = N$ forces $x_j = N$ for all $j > k$. If $i < k$, writing $\varphi_k = \frac{d\mu_k}{d\pi_{k,\alpha}}$, for any bounded measurable f one has

$$\mu_k Q_i(f) = \pi_{k,\alpha}[\varphi_k Q_i f] = \pi_{k,\alpha}[(Q_i \varphi_k) f],$$

where we note that if $x_k = N$ and $i < k$ then $Q_i f(x) = \pi_{k,\alpha}[f \mid x_j, j \neq i]$, and therefore Q_i is self-adjoint in $L^2(\Omega_{k,N}, \pi_{k,\alpha})$. Thus, $\mu_k Q_i$ has density $Q_i \varphi_k$ w.r.t. $\pi_{k,\alpha}$. Recall the

notation $x^{(i,u)}$ for the configuration x updated at the i th particle. For any $x, y \in \Omega_{k,N}$ with $x \leq y$, using $x^{(i,u)} \leq y^{(i,u)}$ for all $u \in [0, 1]$, it follows that

$$\begin{aligned} Q_i \varphi_k(x) &= \int_0^1 \varphi_k(x^{(i,u)}) \rho_\alpha(u) \, du \\ &\leq \int_0^1 \varphi_k(y^{(i,u)}) \rho_\alpha(u) \, du = Q_i \varphi_k(y). \end{aligned}$$

In other words $Q_i \varphi_k$ is increasing, and $\mu_k Q_i \in \mathcal{S}_k$ if $i < k$. When $i = k$, observe that if $x_{k+1} = N$, then $Q_k(f) = \pi_{k+1,\alpha}[f \mid x_j, j < k]$. Therefore, if $\psi_{k,\alpha}$ denotes the density of the marginal of $\pi_{k,\alpha}$ on (x_1, \dots, x_{k-1}) w.r.t. the marginal of $\pi_{k+1,\alpha}$ on the same variables,

$$\begin{aligned} \mu_k Q_k(f) &= \pi_{k,\alpha}[\varphi_k \pi_{k+1,\alpha}[f \mid x_j, j < k]] \\ &= \pi_{k+1,\alpha}[\psi_{k,\alpha} \varphi_k \pi_{k+1,\alpha}[f \mid x_j, j < k]] \\ &= \pi_{k+1,\alpha}[\pi_{k+1,\alpha}[\psi_{k,\alpha} \varphi_k \mid x_j, j < k] f] \\ &= \pi_{k+1,\alpha}[\psi_{k,\alpha} \varphi_k f]. \end{aligned}$$

This shows that $\mu_k Q_k$ is supported on $\Omega_{k+1,N}$ and has density $\psi_{k,\alpha} \varphi_k$ w.r.t. $\pi_{k+1,\alpha}$. A direct computation shows that

$$\psi_{k,\alpha} = C_{\alpha,k,N} (N - x_{k-1})^{-\alpha},$$

for some positive constant $C_{\alpha,k,N}$. In particular, $\psi_{k,\alpha}$ is increasing for any $\alpha > 0$. It follows that if $\mu_k \in \mathcal{S}_k$, then $\mu_k Q_k \in \mathcal{S}_{k+1}$, for all $k = 1, \dots, N - 1$. Taking a generic $\mu \in \mathcal{S}$, by linearity the above implies that $\mu Q_i \in \mathcal{S}$ for any $i = 1, \dots, N - 1$.

To prove the stochastic domination $\mu Q_i \leq \mu$, for $\mu \in \mathcal{S}$, it is sufficient to show that $\mu_k Q_i \leq \mu_k$, for $\mu_k \in \mathcal{S}_k$, for all k, i . Pick $\mu_k \in \mathcal{S}_k$ for some $k = 1, \dots, N$ and an increasing function g on Ω_N . We are going to show that $\mu_k Q_i(g) \leq \mu_k(g)$ for any $i = 1, \dots, N - 1$. If $i > k$ then $\mu_k Q_i = \mu_k$ and there is nothing to prove. If $i < k$, as above we may write

$$\mu_k Q_i(g) = \pi_{k,\alpha}[(Q_i \varphi_k) g] = \pi_{k,\alpha}[(Q_i \varphi_k)(Q_i g)].$$

Since φ_k is also increasing, the FKG inequality on \mathbb{R} , which is valid for any probability measure, implies that $(Q_i \varphi_k)(Q_i g) \leq Q_i(\varphi_k g)$ pointwise. Therefore,

$$\mu_k Q_i(g) \leq \pi_{k,\alpha}[Q_i(\varphi_k g)] = \pi_{k,\alpha}[\varphi_k g] = \mu_k(g).$$

Finally, if $i = k$, then as before we have

$$\mu_k Q_k(g) = \pi_{k+1,\alpha}[\psi_{k,\alpha} \varphi_k g].$$

On the other hand, defining the function $\bar{g}(x) := g(x_1, \dots, x_{k-1}, N, \dots, N)$, one has

$$\mu_k(g) = \pi_{k+1,\alpha}[\psi_{k,\alpha} \varphi_k \bar{g}].$$

Thus, the conclusion $\mu_k Q_k(g) \leq \mu_k(g)$ follows from the fact that $g \leq \bar{g}$. \square

In the next lemma we consider the effect of a sequence of updates on a measure $\mu \in \mathcal{S}$ and compare it with the effect of another sequence obtained from the first by removing some of the updates.

LEMMA 33. *Pick $n \in \mathbb{N}$ and fix a sequence $z := (z_1, \dots, z_n) \in \llbracket 1, N - 1 \rrbracket^n$. For any $\mu \in \mathcal{S}$, if μ^z denotes the new measure*

$$(104) \quad \mu^z = \mu Q_{z_1} \cdots Q_{z_n},$$

then $\mu^z \in \mathcal{S}$. Moreover, if z' denotes a sequence obtained from z by removing some of the entries, then $\mu^z \leq \mu^{z'}$ and

$$(105) \quad \|\mu^z - \pi_{N,\alpha}\|_{\text{TV}} \leq \|\mu^{z'} - \pi_{N,\alpha}\|_{\text{TV}}.$$

PROOF. Lemma 32 shows that $\mu^z \in \mathcal{S}$ for any $\mu \in \mathcal{S}$ and any sequence z . For the second part of the lemma, by a telescoping argument it is sufficient to consider the case where z and z' differ by the removal of a single update, say z_j , so that

$$z = (z_1, \dots, z_{j-1}, z_j, z_{j+1}, \dots, z_n), \quad z' = (z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_n).$$

Let $\mu_1 = \mu Q_{z_1} \cdots Q_{z_j}$, and $\mu_2 = \mu Q_{z_1} \cdots Q_{z_{j-1}}$. Then $\mu_1 = \mu_2 Q_{z_j}$ and thus, by Lemma 32 one has $\mu_1 \leq \mu_2$. Moreover,

$$\mu^z = \mu_1 Q_{z_{j+1}} \cdots Q_{z_n} \leq \mu_2 Q_{z_{j+1}} \cdots Q_{z_n} = \mu^{z'},$$

where the inequality follows from the fact that each update preserves the monotonicity, see Proposition 7. The conclusion follows from Lemma 31. \square

We can now state and prove the censoring inequality in our setup. A censoring scheme \mathcal{C} is defined as a càdlàg map

$$\mathcal{C} : [0, \infty) \mapsto \mathcal{P}(\{1, \dots, N - 1\}),$$

where $\mathcal{P}(A)$ denotes the set of all subset of a set A . The subset $\mathcal{C}(s)$, at any time $s \geq 0$, represents the set of particles whose update is to be suppressed at that time. More precisely, given a censoring scheme \mathcal{C} , and an initial condition $x \in \Omega_N$, we write $P_{t,\mathcal{C}}^x$ for the law of the random variable obtained by starting at x and applying the standard graphical construction (see Proposition 7) with the proviso that if the particle with label j rings at time s , then the update is performed if and only if $j \notin \mathcal{C}(s)$. In particular, the uncensored evolution P_t^x corresponds to $P_{t,\mathcal{C}}^x$ when $\mathcal{C}(s) \equiv \emptyset$. Given a distribution μ on Ω_N , we write

$$\mu P_{t,\mathcal{C}} = \int P_{t,\mathcal{C}}^x \mu(dx).$$

LEMMA 34. For any $\alpha > 0$, if $\mu \in \mathcal{S}$, and \mathcal{C} is a censoring scheme, then for all $t \geq 0$:

- (1) $\mu P_t \in \mathcal{S}$ and $\mu P_{t,\mathcal{C}} \in \mathcal{S}$,
- (2) $\mu P_{t,\mathcal{C}}$ is stochastically higher than μP_t .

Moreover, for all $t \geq 0$,

$$(106) \quad \|\mu P_t - \pi_{N,\alpha}\|_{\text{TV}} \leq \|\mu P_{t,\mathcal{C}} - \pi_{N,\alpha}\|_{\text{TV}}.$$

PROOF. It is sufficient to prove (1) and (2) above, since the conclusion (106) is then a consequence of Lemma 31. To prove (1) and (2) note that by conditioning on the realization \mathcal{T}_t of the Poisson clocks $\mathcal{T}^{(j)}$, $j \in \llbracket 1, N - 1 \rrbracket$ up to time t in the graphical construction, one has that the uncensored and the censored evolution are measures of the form μ^z and $\mu^{z'}$ respectively; see (104). By Lemma 33 one has $\mu^z, \mu^{z'} \in \mathcal{S}$ and $\mu^z \leq \mu^{z'}$. Taking the expectation over \mathcal{T}_t shows that $\mu P_t \in \mathcal{S}$ and $\mu P_{t,\mathcal{C}} \in \mathcal{S}$, and that $\mu P_t \leq \mu P_{t,\mathcal{C}}$. \square

6.2. Relaxation of the special particles. Here we show that special particles have reached equilibrium by time $t_{\delta/2}$; see Proposition 39. The key to this result will be Proposition 36 below. Recall the notation \mathcal{S}_N introduced in Section 6.1 for the set of probability measures on Ω_N with an increasing density w.r.t. $\pi_{N,\alpha}$. Given a probability μ on Ω_N , we write $\bar{\mu}$ for the marginal of μ on the special particle positions $y := (y_1, \dots, y_{K-1})$.

LEMMA 35. If $\mu \in \mathcal{S}_N$, then $\bar{\mu}$ is absolutely continuous w.r.t. $\bar{\pi}_{N,\alpha}$ and the corresponding density is increasing on \mathbb{R}^{K-1} .

PROOF. Let $\varphi = d\mu/d\pi_{N,\alpha}$. The density of $\bar{\mu}$ w.r.t. $\bar{\pi}_{N,\alpha}$ is given by the conditional expectation

$$\bar{\varphi}(y) = \pi_{N,\alpha}(\varphi \mid y).$$

To prove that it is increasing, we have to show that $\pi_{N,\alpha}(\varphi \mid y) \geq \pi_{N,\alpha}(\varphi \mid y')$ whenever $y \geq y'$. The latter domination can be seen as follows. Let x be the highest configuration of particle positions such that $y_i = x_{u_i}$, $i = 1, \dots, K - 1$, and let x' be the lowest configuration of particle positions such that $y'_i = x'_{u_i}$, $i = 1, \dots, K - 1$. Clearly, $x \geq x'$. Then use (x, x') as initial conditions in the graphical construction (Proposition 7) for the censored dynamics where all updates of the special particles are suppressed. As time goes to infinity the two distributions converge weakly to $\pi_{N,\alpha}(\cdot \mid y)$, $\pi_{N,\alpha}(\cdot \mid y')$, respectively. Since $x \geq x'$ and the graphical construction preserves the order, this shows that $\pi_{N,\alpha}(\cdot \mid y) \geq \pi_{N,\alpha}(\cdot \mid y')$. \square

We use the following notation for the centered height of the special particles:

$$w_i = x_{u_i} - u_i, \quad W = \sum_{i=1}^{K-1} w_i,$$

and write $\mu(W) = \bar{\mu}(W)$ for the expected value of W . Note that at equilibrium, for N large, the vector (w_1, \dots, w_{K-1}) behaves roughly as a normal vector, and the fluctuations of W are of order \sqrt{N} for each fixed K . The results below are valid for all $\alpha \geq 1$.

PROPOSITION 36. *For any $\varepsilon > 0$, $K \in \mathbb{N}$, there exists $\eta = \eta(K, \varepsilon) > 0$ such that for all $N \geq 2$, $\mu \in \mathcal{S}_N$ one has*

$$(107) \quad \mu(W) \leq \eta\sqrt{N} \quad \Rightarrow \quad \|\bar{\mu} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} \leq \varepsilon.$$

PROOF. We follow [13], Section 5, where a similar statement was proved in the context of random permutations. Given a constant $\lambda > 0$, define the events

$$A_i = \{x : w_i \geq \lambda\sqrt{N}\}, \quad A = \bigcap_{i=1}^{K-1} A_i, \quad B = \bigcup_{i=1}^{K-1} A_i.$$

Let us first show that for any $\mu \in \mathcal{S}_N$, $a > 0$

$$(108) \quad \mu(A) \geq (1 + a)\pi_{N,\alpha}(A) \quad \Rightarrow \quad \mu(W) \geq a\lambda\sqrt{N}\pi_{N,\alpha}(A).$$

Let φ denote the density $d\mu/d\pi_{N,\alpha}$. The sets A_i and A_i^c are stable under the operations \wedge and \vee introduced in (21), and the FKG inequality applied to $\pi_{N,\alpha}(\cdot \mid A_i)$ shows that

$$\begin{aligned} \mu(w_i; A_i) &:= \mu(w_i \mathbf{1}_{A_i}) \\ &= \pi_{N,\alpha}(A_i)\pi_{N,\alpha}(\varphi w_i \mid A_i) \\ &\geq \pi_{N,\alpha}(\varphi; A_i)\pi_{N,\alpha}(w_i \mid A_i) = \mu(A_i)\pi_{N,\alpha}(w_i \mid A_i). \end{aligned}$$

Similarly, the FKG inequality for $\pi_{N,\alpha}(\cdot \mid A_i^c)$ shows that

$$\mu(w_i; A_i^c) \geq \mu(A_i^c)\pi_{N,\alpha}(w_i \mid A_i^c).$$

Therefore, using $\pi_{N,\alpha}(w_i) = 0$,

$$\mu(w_i) \geq (\mu(A_i) - \pi_{N,\alpha}(A_i))\pi_{N,\alpha}(w_i \mid A_i) + (\mu(A_i^c) - \pi_{N,\alpha}(A_i^c))\pi_{N,\alpha}(w_i \mid A_i^c).$$

Since A_i and w_i are increasing and since $\pi_{N,\alpha}(w_i) = 0$, one has $(\mu(A_i^{\mathbb{C}}) - \pi_{N,\alpha}(A_i^{\mathbb{C}})) \times \pi_{N,\alpha}(w_i | A_i^{\mathbb{C}}) \geq 0$ and therefore

$$(109) \quad \mu(w_i) \geq (\mu(A_i) - \pi_{N,\alpha}(A_i))\pi_{N,\alpha}(w_i | A_i) \geq (\mu(A_i) - \pi_{N,\alpha}(A_i))\lambda\sqrt{N}.$$

Consider the function

$$\psi = \left(\sum_{i=1}^{K-1} \mathbf{1}_{A_i} \right) - \mathbf{1}_A.$$

Then ψ is increasing, and applying FKG we obtain

$$\begin{aligned} \sum_{i=1}^{K-1} (\mu(A_i) - \pi_{N,\alpha}(A_i)) &= \mu(A) - \pi_{N,\alpha}(A) + \pi_{N,\alpha}((\varphi - 1)\psi) \\ &\geq \mu(A) - \pi_{N,\alpha}(A). \end{aligned}$$

Summing in (109),

$$\mu(W) \geq (\mu(A) - \pi_{N,\alpha}(A))\lambda\sqrt{N}.$$

This proves (108). Next, let us show that for any $\mu \in \mathcal{S}_N$, $a > 0$

$$(110) \quad \mu(A) \leq (1 + a)\pi_{N,\alpha}(A) \implies \|\bar{\mu} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} \leq a + \pi_{N,\alpha}(B).$$

The sets A, B satisfy the assumptions of Lemma 9. Therefore,

$$\frac{\mu(A)}{\pi_{N,\alpha}(A)} = \pi_{N,\alpha}(\varphi | A) \geq \pi_{N,\alpha}(\varphi | B) = \frac{\mu(B)}{\pi_{N,\alpha}(B)}.$$

If $\mu(A) \leq (1 + a)\pi_{N,\alpha}(A)$, then

$$(111) \quad \mu(B) \leq (1 + a)\pi_{N,\alpha}(B).$$

From Lemma 35 we know that $\bar{\varphi} = d\bar{\mu}/d\bar{\pi}_{N,\alpha}$ is increasing. If x, x' are two particle configurations such that $x' \in B^{\mathbb{C}}$ and $x \in A$, then $y_i = x_{u_i} \geq x'_{u_i} = y'_i$, for all $i = 1, \dots, K - 1$. Since A, B are measurable w.r.t. the y variables, we write \bar{A}, \bar{B} for the corresponding subsets of \mathbb{R}^{K-1} , so that $x \in A \iff y \in \bar{A}$ and $x \in B \iff y \in \bar{B}$. Thus, we have

$$\bar{\varphi}(y') \leq \bar{\varphi}(y), \quad y' \in \bar{B}^{\mathbb{C}}, y \in \bar{A}.$$

Integrating over y w.r.t. $\bar{\pi}_{N,\alpha}(\cdot | \bar{A})$ in the above inequality one finds

$$\bar{\varphi}(y') \leq \frac{\bar{\mu}(\bar{A})}{\bar{\pi}_{N,\alpha}(\bar{A})} = \frac{\mu(A)}{\pi_{N,\alpha}(A)} \leq 1 + a, \quad y' \in \bar{B}^{\mathbb{C}},$$

where we have used the assumption $\mu(A) \leq (1 + a)\pi_{N,\alpha}(A)$. In conclusion,

$$\begin{aligned} \|\bar{\mu} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} &= \int (\bar{\varphi}(y) - 1)_+ \bar{\pi}_{N,\alpha}(dy) \\ &\leq \int_{\bar{B}} \bar{\varphi}(y) \bar{\pi}_{N,\alpha}(dy) + \int_{\bar{B}^{\mathbb{C}}} (\bar{\varphi}(y') - 1)_+ \bar{\pi}_{N,\alpha}(dy') \\ &\leq \bar{\mu}(\bar{B}) + a\bar{\pi}_{N,\alpha}(\bar{B}^{\mathbb{C}}) = \mu(B) + a\pi_{N,\alpha}(B^{\mathbb{C}}). \end{aligned}$$

Using (111) we obtain (110).

Finally, we need a lower bound on the probability $\pi_{N,\alpha}(A)$ and an upper bound on the probability $\pi_{N,\alpha}(B)$. At equilibrium w_i is approximately normal with mean zero and variance iN/K , and therefore

$$b_1(\lambda, K) \leq \pi_{N,\alpha}(A_i) \leq b_2(\lambda, K)$$

for all $i = 1, \dots, K - 1, N \geq 2$, for some constants $b_2(\lambda, K) \rightarrow 0$ as $\lambda \rightarrow \infty$ and $b_1(\lambda, K) > 0$ for all $\lambda > 0$. As for the lower bound on $\pi_{N,\alpha}(A)$ observe that by the FKG inequality one has

$$\pi_{N,\alpha}(A) \geq \prod_{i=1}^{K-1} \pi_{N,\alpha}(A_i) \geq b_1(\lambda, K)^{K-1}.$$

On the other hand,

$$\pi_{N,\alpha}(B) \leq \sum_{i=1}^{K-1} \pi_{N,\alpha}(A_i) \leq (K - 1)b_2(\lambda, K).$$

Then, for any $\varepsilon > 0$, any fixed K , taking λ large enough, we find that there exists a constant $\delta_1 = \delta_1(\varepsilon, K) > 0$ such that

$$(112) \quad \delta_1 \leq \pi_{N,\alpha}(A) \leq \pi_{N,\alpha}(B) \leq \frac{\varepsilon}{2}.$$

Once we have (108), (110) and (112) we can conclude as follows. Suppose that $\mu(W) \leq \eta\sqrt{N}$. If $a = 2\eta/\lambda\delta_1$, then by (108) and (112) one must have $\mu(A) \leq (1 + a)\pi_{N,\alpha}(A)$. Therefore by (110) and (112)

$$\|\bar{\mu} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} \leq \frac{2\eta}{\lambda\delta_1} + \frac{\varepsilon}{2}.$$

The desired conclusion follows by taking $\eta = \varepsilon\lambda\delta_1/4$. \square

Next, we address the problem of controlling the expected value of W at time t when started from the maximal configuration.

PROPOSITION 37. *For any $k = 1, \dots, N - 1$, any $t \geq 0$:*

$$\mathbb{E}[X_k^\wedge(t)] \leq k + 2Ne^{-\text{gap}_N t}.$$

In particular, if $\mu_t = \delta_\wedge P_t$, then for all $t \geq 0$:

$$(113) \quad \mu_t(W) \leq 2KN e^{-\text{gap}_N t}.$$

PROOF. Defining

$$\varphi_j(k) = \sqrt{\frac{2}{N}} \sin\left(\frac{jk\pi}{N}\right),$$

one has for all $j, k = 1, \dots, N - 1$:

$$(114) \quad \Delta\varphi_j(k) = \varphi_j(k + 1) + \varphi_j(k - 1) - 2\varphi_j(k) = -2\lambda_j\varphi_j(k),$$

where

$$(115) \quad \lambda_j = 1 - \cos\left(\frac{j\pi}{N}\right).$$

Set $v(t) = (v_1(t), \dots, v_{N-1}(t))$, where

$$v_k(t) = \mathbb{E}[X_k^\wedge(t)] - k.$$

Expanding the vector $v(t)$ in the orthonormal basis $\varphi_j, j = 1, \dots, N - 1$, one finds $v_k(t) = \sum_{j=1}^{N-1} a_j(t)\varphi_j(k)$, where $a_j(t) = \sum_{k=1}^{N-1} \varphi_j(k)v_k(t)$. Using (25) one finds

$$a_j(t) = a_j(0)e^{-\lambda_j t}, \quad a_j(0) = \sum_{k=1}^{N-1} \varphi_j(k)(N - k).$$

In particular, $|a_j(0)| \leq N^{3/2}/\sqrt{2}$, and

$$v_k(t) \leq N \sum_{j=1}^{N-1} e^{-\lambda_j t}.$$

Using $\lambda_j \geq j\lambda_1$ it follows that

$$v_k(t) \leq \frac{Ne^{-\lambda_1 t}}{1 - e^{-\lambda_1 t}}.$$

If t is such that $e^{-\lambda_1 t} \leq 1/2$ then this implies $v_k(t) \leq 2Ne^{-\lambda_1 t}$. On the other hand if $e^{-\lambda_1 t} \geq 1/2$ then clearly $v_k(t) \leq N \leq 2Ne^{-\lambda_1 t}$. Since $\lambda_1 = \text{gap}_N$, this proves the desired upper bound. \square

Next, we want to use the bound in Proposition 37 to obtain, via Proposition 36, the desired control on the convergence of special particles. However, again a technical problem arises due to the fact that $\mu_t = \delta_{\wedge} P_t$ has no density w.r.t. equilibrium. We overcome this by showing that the singular part of μ_t has very small mass if t is large.

LEMMA 38. *For any $t \geq 0$, the measure $\mu_t = \delta_{\wedge} P_t$ satisfies*

$$(116) \quad \mu_t = (1 - \gamma_{t,N})\mu'_t + \gamma_{t,N}\mu_{t,N},$$

where $\mu_{t,N} \in \mathcal{S}_N$, and $\gamma_{t,N} \geq 1 - Ne^{-t/N}$, for some probability measure μ'_t .

PROOF. We know from Lemma 34 that $\mu_t \in \mathcal{S}$ for all $t \geq 0$. Thus,

$$\mu_t = (1 - \gamma_{t,N})\mu'_t + \gamma_{t,N}\mu_{t,N},$$

with $\mu_{t,N} \in \mathcal{S}_N$ for some coefficients $\gamma_{t,N}$, and some singular measure μ'_t . It remains to show that $\gamma_{t,N} \geq 1 - Ne^{-t/N}$. By conditioning on the realization \mathcal{T}_t of the Poisson clocks $\mathcal{T}^{(j)}$, $j \in \llbracket 1, N - 1 \rrbracket$ up to time t in the graphical construction, one has that the distribution of the particles at time t is of the form μ^z , for some sequence of updates z ; see (104). In the proof of Lemma 32 we have seen that if $\mu \in \mathcal{S}_k$, then $\mu Q_k \in \mathcal{S}_{k+1}$, and therefore if the sequence z contains the full sweep $(1, 2, \dots, N - 1)$ as a subsequence, then $\mu^z \in \mathcal{S}_N$. Let E_t denote the event that z contains $(1, 2, \dots, N - 1)$ as a subsequence. Since \mathcal{S}_N is stable under convex combinations, taking the expectation over \mathcal{T}_t one finds that $\gamma_{t,N} \geq \mathbb{P}(E_t)$. A rough lower bound on the latter can be obtained by dividing the interval $[0, t]$ in $N - 1$ intervals and by requiring that for each i the clock of particle with label i rings during the i th time interval. This shows that $\mathbb{P}(E_t) \geq 1 - Ne^{-t/N}$. \square

We are ready to accomplish the first and most delicate step in the program outlined at the beginning of this section.

PROPOSITION 39. *Fix $\alpha \geq 1$ and $K \in \mathbb{N}$. Let $\mu_t = \delta_{\wedge} P_t$ and let $\bar{\mu}_t$ denote the marginal of μ_t on the special particle positions $Y_i(t)$, $i = 1, \dots, K - 1$. If $\bar{\pi}_{N,\alpha}$ denotes the corresponding equilibrium distribution, for any fixed $\delta > 0$, with $t_\delta = (1 + \delta) \frac{\log N}{2 \text{gap}_N}$ one has*

$$(117) \quad \lim_{N \rightarrow \infty} \|\bar{\mu}_{t_\delta} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} = 0.$$

PROOF. From Lemma 38 it follows that

$$\|\bar{\mu}_{t_\delta} - \bar{\pi}_{N,\alpha}\|_{\text{TV}} \leq Ne^{-t_\delta/N} + \|\bar{\mu}_{t_\delta,N} - \bar{\pi}_{N,\alpha}\|_{\text{TV}},$$

where $\mu_{t_\delta, N} \in \mathcal{S}_N$. Since $Ne^{-t_\delta/N} \rightarrow 0, N \rightarrow \infty$, from Proposition 36, (117) follows if we show that

$$(118) \quad \lim_{N \rightarrow \infty} N^{-\frac{1}{2}} \mu_{t_\delta, N}(W) = 0.$$

By Proposition 37, we know that

$$0 \leq \mu_{t_\delta}(W) \leq 2KN^{\frac{1}{2}(1-\delta)}.$$

On the other hand, Lemma 38 shows that

$$\begin{aligned} 0 \leq \mu_{t_\delta, N}(W) &= -\gamma_{t_\delta, N}^{-1}(1 - \gamma_{t_\delta, N})\mu'_{t_\delta}(W) + \gamma_{t_\delta, N}^{-1}\mu_{t_\delta}(W) \\ &\leq 2\mu_{t_\delta}(W) = O(N^{\frac{1}{2}(1-\delta)}). \end{aligned} \quad \square$$

6.3. *Relaxation of the censored dynamics.* Here we establish the second step in the proof of Proposition 14. Consider the censored process obtained by suppressing all updates of the special particles. In other words, we use the censoring scheme \mathcal{C} such that $\mathcal{C}(s) = \{u_1, \dots, u_{K-1}\}, s \geq 0$.

PROPOSITION 40. *Fix $\alpha > 0$. Let $P_{t, \mathcal{C}}^x = \delta_x P_{t, \mathcal{C}}$ and let $\pi_{N, \alpha}(\cdot | y)$ denote the equilibrium distribution given the special particle positions $y_i = x_{u_i}, i = 1, \dots, K - 1$. For any $\delta > 0$ small enough, setting $K = \lfloor \delta^{-1} \rfloor$, and $s_\delta = \delta \frac{\log N}{2 \text{gap}_N}$, one has, for N sufficiently large*

$$(119) \quad \sup_{x \in \Omega_N} \|P_{s_\delta, \mathcal{C}}^x - \pi_{N, \alpha}(\cdot | y)\|_{\text{TV}} \leq \delta.$$

PROOF. By construction, the censored process corresponds to the product of K independent adjacent walks each on the n -simplex, with $n := \lfloor N/K \rfloor$. The bound (120) implies that when N is sufficiently large, the mixing time of a system of size n satisfies $T_n(\varepsilon) \leq Cn^2 \log n$ for any given $\varepsilon \in (0, 1)$ if n is large enough. Therefore,

$$T_n(K^{-1}\delta) \leq C \frac{N^2}{K^2} \log(N/K) \leq s_\delta,$$

for $\delta > 0$ small enough. Thus if $d_N(t)$ is the distance defined in (5), then

$$\sup_{x \in \Omega_N} \|P_{s_\delta, \mathcal{C}}^x - \pi_{N, \alpha}(\cdot | y)\|_{\text{TV}} \leq K d_n(s_\delta) \leq \delta,$$

as required. \square

6.4. *Proof of Proposition 14.* With the previous results at hand it is relatively simple to conclude the proof of the desired estimate. We formulate the result as follows. Recall the notation $t_\delta = (1 + \delta) \frac{\log N}{2 \text{gap}_N}$.

PROPOSITION 41. *Fix $\alpha \geq 1$. For any $\delta > 0$,*

$$\lim_{N \rightarrow \infty} \|P_{t_\delta}^\wedge - \pi_{N, \alpha}\|_{\text{TV}} = 0.$$

PROOF. Set $K = \lfloor \delta^{-1} \rfloor$ and let \mathcal{C}' denote the censoring scheme defined by $\mathcal{C}'(s) = \emptyset$ for $s \in [0, t_\delta/2)$ and $\mathcal{C}'(s) = \{u_1, \dots, u_{K-1}\}$ for $s \geq t_\delta/2$. Let also $P_{t, * }^\wedge = \delta_\wedge P_{t, \mathcal{C}'}$ denote the corresponding censored process. From Lemma 34, we have

$$\|P_{t_\delta}^\wedge - \pi_{N, \alpha}\|_{\text{TV}} \leq \|P_{t_\delta, * }^\wedge - \pi_{N, \alpha}\|_{\text{TV}}.$$

A coupling of $P_{t_\delta, *}$ and $\pi_{N, \alpha}$ can be achieved as follows. Call $\mu_t = \delta \wedge P_t$. Use the optimal coupling attaining the total variation distance $\|\bar{\mu}_{t_{\delta/2}} - \bar{\pi}_{N, \alpha}\|_{TV}$ to couple the special particles at time $t_{\delta/2}$. If the special particles are coupled at time $t_{\delta/2}$ and if x is the configuration at that time, then couple the remaining particles at time $t_\delta = t_{\delta/2} + s_{\delta/2}$ with the optimal coupling attaining the total variation distance $\|P_{s_{\delta/2}, \mathcal{C}}^x - \pi_{N, \alpha}(\cdot | y)\|_{TV}$, where \mathcal{C} is as in Proposition 40. This shows that

$$\|P_{t_\delta}^\wedge - \pi_{N, \alpha}\|_{TV} \leq \|\bar{\mu}_{t_{\delta/2}} - \bar{\pi}_{N, \alpha}\|_{TV} + \sup_{x \in \Omega_N} \|P_{s_{\delta/2}, \mathcal{C}}^x - \pi_{N, \alpha}(\cdot | y)\|_{TV}.$$

From Proposition 39 and Proposition 40,

$$\limsup_{N \rightarrow \infty} \|P_{t_\delta}^\wedge - \pi_{N, \alpha}\|_{TV} \leq 2\delta.$$

The distance $\|P_{t_\delta}^\wedge - \pi_{N, \alpha}\|_{TV}$ is decreasing as a function of δ , and therefore we may take $\delta \rightarrow 0$ in the right-hand side above to conclude. \square

REMARK 42. Proposition 41 can be strengthened to obtain that for any $\delta > 0$,

$$\lim_{N \rightarrow \infty} \|\mu P_{t_\delta} - \pi_{N, \alpha}\|_{TV} = 0$$

for an arbitrary $\mu \in \mathcal{S}$. Indeed, for any $\mu \in \mathcal{S}$ one has $\mu P_t \in \mathcal{S}$ and $\mu P_t \leq P_t^\wedge$ and therefore the claim follows from Lemma 31. However, at this point we cannot infer that the same holds for arbitrary initial distribution without the requirement that it belongs to \mathcal{S} .

APPENDIX: AN EXTENSION OF THE RANDALL–WINKLER UPPER BOUND

As already discussed in the previous sections, one of the ingredients of our main results is the upper bound (13) on the mixing time that captures the right order of magnitude up to a multiplicative constant, which in the case $\alpha = 1$ was obtained by Randall and Winkler [21]. In this section, we explain how one can extend that bound to the Beta-resampling case with $\alpha \geq 1$. We are going to show more precisely that for any $\varepsilon > 0$, if N is sufficiently large, then

$$(120) \quad T_{N, \alpha}(\varepsilon) \leq 5(\log N) \text{gap}_N^{-1}.$$

Let us fix $x \in \Omega_N$ and construct a coupling of the two processes \mathbf{X}^x and \mathbf{X}^{eq} , where \mathbf{X}^x has initial state x and \mathbf{X}^{eq} has initial distribution $\pi_{N, \alpha}$. Clearly,

$$(121) \quad \|P_t^x - \pi_{N, \alpha}\|_{TV} \leq \mathbb{P}[\mathbf{X}^x(t) \neq \mathbf{X}^{\text{eq}}(t)].$$

Our manner of coupling the two processes is time dependent and is as follows:

(A) For $t \leq t_1 := 4(\log N) \text{gap}_N^{-1}$, we couple the two trajectories using the construction presented in the proof of Proposition 7.

(B) For $t > t_1$, we couple the two trajectories according to the construction described in Section 5.2, namely we use a coupling that maximizes the probability of sticking at each update.

We first observe that at time t_1 the two processes have come close to one another:

$$(122) \quad \mathbb{E} \left[\sum_{k=1}^{N-1} |X_k^x(t_1) - X_k^{\text{eq}}(t_1)| \right] \leq N^{-2}.$$

To see this, referring to the proof of Proposition 7, we consider the two processes $(\mathbf{X}^\wedge(t))_{t \in [0, t_1]}$, $(\mathbf{X}^\vee(t))_{t \in [0, t_1]}$ obtained by using the same clock processes \mathcal{T} and update

variables U as in the construction of $\mathbf{X}^x(t)$ and $\mathbf{X}^{\text{eq}}(t)$. Since this construction is order preserving, we necessarily have

$$|X_k^x(t_1) - X_k^{\text{eq}}(t_1)| \leq X_k^\wedge(t_1) - X_k^\vee(t_1),$$

and (122) can be proved for $X_k^\wedge(t_1) - X_k^\vee(t_1)$ repeating the computation leading to (65).

Now for $t > t_1$, we remark that in contrast with Randall–Winkler’s case $\alpha = 1$, the step B described above does not descend from a monotone grand coupling, and thus it is not sufficient to couple $X_k^\wedge(t)$ and $X_k^\vee(t)$ by time $5 \log N \text{gap}_N^{-1}$ to prove the desired mixing time bound.

Instead, we need to couple $\mathbf{X}^x(t)$ and $\mathbf{X}^{\text{eq}}(t)$ for arbitrary x . To this end we are going to repeat the recursive argument of Section 5.10 with $X_k^\wedge(t)$ replaced by $X_k^{\text{eq}}(t)$. In particular, when we perform the update at a site k , we have to replace the interval I^\wedge (see Section 5.2) by I^{eq} , which we define as the resampling interval of X_k^{eq} . As a consequence of this modification, the supports of the densities ρ_1 and ρ_3 presented in (52) are not necessarily intervals. Define the events

$$\begin{aligned} \bar{\mathcal{A}}_1^{(N)} &:= \left\{ \sum_{k=1}^{N-1} |X_k^x(t_1) - X_k^{\text{eq}}(t_1)| < N^{-1} \right\}, \\ \bar{\mathcal{A}}_2^{(N)} &:= \left\{ \min_{k \in \llbracket 1, N-1 \rrbracket, s \in [0, 2 \log N]} \nabla X_k^{\text{eq}}(t_1 + s) > \frac{1}{\sqrt{N \log N}} \right\}. \end{aligned}$$

From (122) and Markov’s inequality it follows that $\mathbb{P}(\bar{\mathcal{A}}_1^{(N)}) \rightarrow 1$, and from the proof of Lemma 30 it follows that $\mathbb{P}(\bar{\mathcal{A}}_2^{(N)}) \rightarrow 1$.

As in Section 5.10 we consider the probability of an unsuccessful update in the time interval $[t_1, t_1 + 2 \log N]$. Let $\bar{\mathcal{A}}_3^{(N)}$ denote the event that each site is updated at least once and no more than $10 \log N$ times in this time interval. The probability of $\bar{\mathcal{A}}_3^{(N)}$ tends to 1 (cf. Section 5.10).

Using the upper bound (57) and recalling Remark 17, we see that conditioned on the realization of the update times $(t_i^{(k)})_{i,k}$ in the time interval $[t_1, t_1 + 2 \log N]$, on the event $\bar{\mathcal{A}}_1^{(N)} \cap \bar{\mathcal{A}}_3^{(N)}$, the probability that there exists at least one unsuccessful update within $[t_1, t_1 + 2 \log N]$ is bounded above by $C N^{-1/2} (\log N)^2 + o(1)$. This shows that the probability that $\mathbf{X}^x(t_1 + 2 \log N) \neq \mathbf{X}^{\text{eq}}(t_1 + 2 \log N)$ vanishes as $N \rightarrow \infty$, which implies (120). Note that the constant 5 in this bound is not optimal.

Finally, we remark that the proof given above can be extended to the case $\alpha \in (0, 1)$, with 5 replaced by a constant which depends on α . However, this requires to prove an inequality that replaces the upper bound of (57), namely, that when $\alpha \in (0, 1)$ there exists $C_\alpha > 0$ such that for any pair of intervals

$$(123) \quad \|\text{Beta}_\alpha[l_1, r_1] - \text{Beta}_\alpha[l_2, r_2]\|_{\text{TV}} \leq C_\alpha \left(\frac{\max(|l_2 - l_1|, |r_2 - r_1|)}{\max(r_1 - l_1, r_2 - l_2)} \right)^{1/\alpha}.$$

We leave the details of the adaptation as well as the proof of (123) to the interested reader.

Acknowledgements. Hubert Lacoïn is grateful to Werner Krauth for bringing to his knowledge the work of Randall and Winkler and for a stimulating discussion on the subject. This work was initiated during a stay of Pietro Caputo and Cyril Labbé at IMPA, and continued during a stay of Pietro Caputo and Hubert Lacoïn at Dauphine: we are grateful for hospitality and support provided by these two institutions. Hubert Lacoïn also acknowledges support from a productivity grant of CNPq and a JCNE grant from FAPERj. Cyril Labbé acknowledges support from the grant SINGULAR ANR-16-CE40-0020-01.

REFERENCES

- [1] ALDOUS, D. and DIACONIS, P. (1987). Strong uniform times and finite random walks. *Adv. in Appl. Math.* **8** 69–97. MR0876954 [https://doi.org/10.1016/0196-8858\(87\)90006-6](https://doi.org/10.1016/0196-8858(87)90006-6)
- [2] ALDOUS, D. and FILL, J. (2002). *Reversible Markov Chains and Random Walks on Graphs*. Berkeley, CA.
- [3] CAPUTO, P. (2008). On the spectral gap of the Kac walk and other binary collision processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 205–222. MR2429910
- [4] CARLEN, E. A., CARVALHO, M. C. and LOSS, M. (2003). Determination of the spectral gap for Kac’s master equation and related stochastic evolution. *Acta Math.* **191** 1–54. MR2020418 <https://doi.org/10.1007/BF02392695>
- [5] DIACONIS, P. (1996). The cutoff phenomenon in finite Markov chains. *Proc. Natl. Acad. Sci. USA* **93** 1659–1664. MR1374011 <https://doi.org/10.1073/pnas.93.4.1659>
- [6] DYER, M., FRIEZE, A. and KANNAN, R. (1991). A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.* **38** 1–17. MR1095916 <https://doi.org/10.1145/102782.102783>
- [7] GIACOMIN, G. (2002). *Aspects of Statistical Mechanics of Random Surfaces. Lecture Notes for Course Given at IHP*.
- [8] GIBBS, A. L. and SU, F. E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* 419–435.
- [9] HERMON, J., LACONIN, H. and PERES, Y. (2016). Total variation and separation cutoffs are not equivalent and neither one implies the other. *Electron. J. Probab.* **21** Paper No. 44, 36. MR3530321 <https://doi.org/10.1214/16-EJP4687>
- [10] HOUGH, B. and JIANG, Y. (2017). Cut-off phenomenon in the uniform plane Kac walk. *Ann. Probab.* **45** 2248–2308. MR3693962 <https://doi.org/10.1214/16-AOP1111>
- [11] KIPNIS, C., MARCHIORO, C. and PRESUTTI, E. (1982). Heat flow in an exactly solvable model. *J. Stat. Phys.* **27** 65–74. MR0656869 <https://doi.org/10.1007/BF01011740>
- [12] LABBÉ, C. and LACONIN, H. (2018). Mixing time and cutoff for the weakly asymmetric simple exclusion process. ArXiv e-prints.
- [13] LACONIN, 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. (2017). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. Second edition of [MR2466937]. MR3726904
- [15] LIGGETT, T. M. (2005). *Interacting Particle Systems. Classics in Mathematics*. Springer, Berlin. Reprint of the 1985 original. MR2108619 <https://doi.org/10.1007/b138374>
- [16] LOVÁSZ, L. and VEMPALA, S. (2003). Hit-and-run is fast and fun. Preprint, Microsoft Research.
- [17] LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Stat.* **26** 319–324. MR0069408 <https://doi.org/10.1214/aoms/1177728549>
- [18] PERES, Y. and WINKLER, P. (2013). Can extra updates delay mixing? *Comm. Math. Phys.* **323** 1007–1016. MR3106501 <https://doi.org/10.1007/s00220-013-1776-0>
- [19] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York. Translated from the Russian by A. A. Brown, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 82. MR0388499
- [20] PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Comm. Math. Phys.* **36** 233–241. MR0341553
- [21] RANDALL, D. and WINKLER, P. (2005). Mixing points on a circle. In *Approximation, Randomization and Combinatorial Optimization. Lecture Notes in Computer Science* **3624** 426–435. Springer, Berlin. MR2193706 https://doi.org/10.1007/11538462_36
- [22] RANDALL, D. and WINKLER, P. (2005). Mixing points on an interval. In *Proceedings of the Second Workshop on Analytic Algorithms and Combinatorics, Vancouver, 2005* 216–221.
- [23] SMITH, A. (2013). Analysis of convergence rates of some Gibbs samplers on continuous state spaces. *Stochastic Process. Appl.* **123** 3861–3876. MR3084162 <https://doi.org/10.1016/j.spa.2013.05.003>
- [24] SMITH, A. (2014). A Gibbs sampler on the n -simplex. *Ann. Appl. Probab.* **24** 114–130. MR3161643 <https://doi.org/10.1214/12-AAP916>
- [25] 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>