SPECTRAL GAP AND CUTOFF PHENOMENON FOR THE GIBBS
SAMPLER OF $\nabla \varphi$ INTERFACES WITH CONVEX POTENTIAL

PIETRO CAPUTO, CYRIL LABBÉ, AND HUBERT LACOIN

ABSTRACT. We consider the Gibbs sampler, or heat bath dynamics associated
to log-concave measures on $\mathbb{R}^N$ describing $\nabla \varphi$ interfaces with convex poten-
tials. Under minimal assumptions on the potential, we find that the spectral
gap of the process is always given by $\text{gap}_N = 1 - \cos(\pi/N)$, and that for all
$\epsilon \in (0, 1)$, its $\epsilon$-mixing time satisfies $T_N(\epsilon) \sim \frac{\log N}{\text{gap}_N}$ as $N \to \infty$, thus estab-
lishing the cutoff phenomenon. The results reveal a universal behavior in that
they do not depend on the choice of the potential.

MSC 2010 subject classifications: Primary 60J25; Secondary 37A25, 82C22.
Keywords: Spectral gap; Mixing time; Cutoff.

CONTENTS

1. Introduction 1
2. Main tools 6
3. Lower bound on the mixing time 15
4. A first upper bound and the spectral gap 18
5. Upper bound on the mixing time 24
6. Proof of Theorem 5.1 32
References 37

1. INTRODUCTION

1.1. Model and result. We consider the 1+1 dimensional interface model defined
as follows. The state space of the interface is defined by

$$\Omega_N := \{(x_0, ..., x_N) \in \mathbb{R}^{N+1} : x_0 = 0, x_N = 0\}.$$ 

We fix a potential $V \in \mathcal{C}$, where $\mathcal{C}$ denotes the set of all functions $V : \mathbb{R} \to \mathbb{R}$
satisfying the following assumptions:

(i) $V$ is convex,
(ii) $V$ grows at most polynomially: there exist $C > 0$ and $K \geq 1$ such that for
all $x \in \mathbb{R},$

\begin{equation}
|V(x)| \leq C(1 + |x|)^K.
\end{equation}

(iii) $V$ is non-affine: namely we have $V'_+ > V'_-$ where

\begin{equation}
V'_+ := \lim_{x \to \infty} V(x)/x \quad \text{and} \quad V'_- := \lim_{x \to -\infty} V(x)/x
\end{equation}

Date: July 21, 2020.
The $\nabla \varphi$ interface with potential $V$ is the random element of $\Omega_N$ with distribution $\pi_N$, whose density with respect to Lebesgue measure is given by
\begin{equation}
\frac{d\pi_N}{dx_1 \ldots dx_{N-1}} = \frac{e^{-H}}{Z_N},
\end{equation}
where $Z_N$ is the normalization constant and $H = H_{N,V}$ is the Hamiltonian
\[ H(x) := \sum_{k=1}^{N} V(x_k - x_{k-1}). \]

The Gibbs sampler for the measure $\pi_N$ that we wish to consider is the heat-bath dynamics defined as follows. Let $Q_k$ be the operator that equilibrates the $k$-th coordinate of $x_k$ conditionally given the remaining coordinates. More precisely, letting $x^{(k,u)}_k$ denote the vector $(x_0, \ldots, x_{k-1}, u, x_{k+1}, \ldots, x_N)$, set
\begin{equation}
Q_k f(x) := \int f(x^{(k,u)}) e^{-H(x^{(k,u)})} du \int e^{-H(x^{(k,u)})} du
\end{equation}
\[ = \int f(x^{(k,u)}) \rho_{x_{k-1},x_{k+1}}(u) du, \]
where
\[ \rho_{b,c}(u) := \frac{e^{-V(u-b)-V(c-s)}}{\int e^{-V(u-b)-V(c-s)} ds}. \]

Define the Markov generator $L = L_{N,V}$ by
\begin{equation}
L f := \sum_{k=1}^{N-1} (Q_k f - f).
\end{equation}

Let $X^x = (X^x(t))_{t \geq 0}$ be the continuous time Markov chain on $\Omega_N$ with generator $L$ and initial condition $x$. Given $x \in \Omega_N$ and $\nu$ a probability measure on $\Omega_N$, let $P^x_t$ and $P^\nu_t$ denote the distribution at time $t$ of the Markov chain with initial condition $x$ and $\nu$ respectively. One can describe the evolution of the process as follows: each coordinate of $X^x(t)$ is updated with rate 1 independently. When an update is performed at time $t$ for coordinate $k$ the value of $X^x_k$ is resampled according to the conditional equilibrium measure, whose density is $\rho_{b,c}$ with $b = X^x_{k-1}(t)$, $c = X^x_{k+1}(t)$.

Since $L$ is a finite sum of orthogonal projectors, it is a bounded self-adjoint operator on $L^2 = L^2(\Omega_N, \pi_N)$, and therefore, the corresponding process is reversible with respect to $\pi_N$. The spectral gap of the Gibbs sampler is defined by
\begin{equation}
gap_N = \inf_{f \in L^2 \cap \pi_N(f)=0} \frac{\pi_N(f(-L f))}{\pi_N(f^2)},
\end{equation}
we use the notation $\pi_N(f) = \int f d\pi_N$. We do not know whether the operator $L$ has pure point spectrum in general, and therefore the spectral gap does not a priori coincide with (the opposite of) some eigenvalue of $L$. Our first result computes the value of $\gap_N$ and shows that it is indeed an eigenvalue.

**Theorem 1.1.** For any potential $V \in \mathcal{C}$, for all $N \geq 2$, the spectral gap of $L$ is given by
\[ \gap_N = 1 - \cos \left( \frac{\pi}{N} \right), \]
and the function

\[ f_N(x) = \sum_{k=1}^{N-1} \sin \left( \frac{k\pi}{N} \right) x_k, \]

is an eigenfunction of \( \mathcal{L} \) with eigenvalue \(-\text{gap}_N\).

We remark that the spectral gap of the dynamics is independent of the choice of the potential \( V \), as long as \( V \in \mathcal{C} \), and it coincides with the first Dirichlet eigenvalue of the discrete Laplace operator on the segment \( \{1, \ldots, N-1\} \).

Our next results concern the mixing time of the Gibbs sampler. Without restriction on the set of possible initial conditions, this mixing time is infinite. Consequently, we restrict ourselves to initial conditions with absolute height at most \( N \), and consider the distance to equilibrium at time \( t \) from a worst case initial condition:

\[ d_N(t) := \sup_{x \in \Omega_N : |x|_\infty \leq N} \|P_N^t x - \pi_N\|_{TV}, \]

where the total variation distance between two probability measures \( \mu, \nu \) on \( \Omega_N \) is defined as

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

the supremum ranging over all Borel subsets of \( \Omega_N \). Note that we do not condition the dynamics to keep the height of the interface within \([-N, N]\).

For any \( \epsilon \in (0, 1) \), the \( \epsilon \)-mixing time is then defined as

\[ T_N(\epsilon) := \inf \{t \geq 0 : d_N(t) < \epsilon\}. \]

**Theorem 1.2.** For any \( V \in \mathcal{C} \), for all \( \epsilon \in (0, 1) \):

\[ T_N(\epsilon) \sim \frac{\log N}{2 \text{gap}_N}. \]

We use the symbol “\( \sim \)” for asymptotic equivalence as \( N \to \infty \), so that in view of Theorem 1.1 (1.9) is equivalent to

\[ \lim_{N \to \infty} \frac{T_N(\epsilon)}{N^2 \log N} = \frac{1}{\pi^2}. \]

Theorem 1.2 shows that the \( \epsilon \)-mixing time is, to leading order, insensitive to the threshold parameter \( \epsilon \), that is, the Gibbs sampler satisfies the cutoff phenomenon. Note again the universal behavior, that is the fact that nothing depends on \( V \), as long as \( V \in \mathcal{C} \).

**Remark 1.3.** If the restriction on the absolute height \( |x|_\infty \leq N \) is replaced by \( |x|_\infty \leq a_N \) with \( a_N \gg \sqrt{N} \) then our proof carries over and yields

\[ T_N(\epsilon) \sim \frac{\log(a_N N^{-1/2})}{\text{gap}_N}. \]

For an interpretation of this result, observe that if the initial condition is \( x_i = a_N \) for every \( i \in \{1, \ldots, N-1\} \), then \( t = \frac{\log(a_N N^{-1/2})}{\text{gap}_N} \) is exactly the time it takes for

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

to drop from \( f_N(x) = \Theta(\sqrt{N a_N}) \) to \( \Theta(N^{3/2}) \), the latter being the order of fluctuations of \( f_N \) at equilibrium.
Remark 1.4. The Markov chain can be viewed as taking values in the larger space
\[(1.11) \quad \tilde{\Omega}_N := \{(x_0, \ldots, x_N) \in \mathbb{R}^{N+1} : x_0 = 0\}.
\]
In that case, the value at the endpoint \(X_N(t)\) remains equal to its initial value \(x_N\) for all \(t\). Moreover, we could have fixed the endpoint \(x_N = hN\) with \(h \neq 0\) and thus have considered the mixing property of the process within the set
\[\Omega_{N,h} := \{(x_0, \ldots, x_N) \in \mathbb{R}^{N+1} : x_0 = 0, x_N = hN\}.
\]
The results of Theorems 1.1 and 1.2 still hold in this more general setting. Indeed, using the transformation \((x_k) \mapsto (x_k - kh)\) which maps \(\Omega_{N,h}\) to \(\Omega_N\), and considering the modified potential \(V_h(\cdot) := V(h + \cdot)\), again an element of \(\mathcal{C}\), we are back to the original setting. In particular, it follows that the spectral gap is independent of \(h\). Concerning the mixing time, one can actually prove that the \(N \to \infty\) limit in Theorem 1.2 holds uniformly over \(h\) in compact sets.

Remark 1.5. Let us comment on our assumptions on the potential \(V \in \mathcal{C}\). The convexity hypothesis on \(V\) is the most important, and it is required at various points in the proof. In the language of interacting particle systems it makes the system attractive, in the sense that it entails the existence of a coupling that preserves monotonicity (see Lemma 2.2 below). The assumption \((1.1)\) about polynomial growth is merely technical. It helps us obtain certain estimates, and in practice it does not appear to be very restrictive. Finally the assumption \((1.2)\) is the easiest to justify: if \(V\) is an affine function then the measure \(\pi_N\) is not defined since \(e^{-H}\) would not be integrable in this case. Note that the definition \((1.3)\) remains unchanged if \(V(u)\) is replaced by \(V'(u) = V(u) + au + b\), since in that case \(H_N,V'(x) = H_N,V(x) + bN\) is only modified by a constant.

1.2. Related works. The relaxation to equilibrium of \(\nabla \varphi\) interfaces has been the object of many remarkable works in recent years, especially in conjunction with hydrodynamic limits, see e.g. [Gia02, Fun05] and the references therein. In particular, the validity of functional inequalities for the equilibrium measure \(\pi_N\) has been explored under various assumptions on the potential \(V\). The dynamics considered in these works is usually the conservative diffusion process, namely the Langevin dynamics associated to the Hamiltonian \(H\) in the state space \(\Omega_N\). For instance, when the potential is a bounded perturbation of a uniformly strictly convex function, then upper and lower bounds of order \(N^{-2}\) on the spectral gap of the Langevin diffusion have been obtained in [Cap03]. Moreover, the stronger logarithmic Sobolev inequality has been established by Menz and Otto [MO13]. These results were shown to hold uniformly in the tilt parameter \(h\) when the interface endpoint is fixed at \(x_N = hN\). The uniformity in \(h\) is a consequence of the assumption of uniform strict convexity and it cannot hold for the diffusion process if the potential is only assumed to be convex. In the non-uniformly convex case, spectral gap bounds with the correct dependence on the tilt \(h\) were obtained in [BW09, BM13] for certain potentials such as the solid-on-solid (SOS) potential \(V(u) = |u|\). While several of the techniques employed in these works carry over to the jump process we consider in this paper, as far as we know none of the previous works allows one to actually compute the spectral gap as we do here. As discussed in Remark 1.4, our results hold uniformly for \(h\) in a compact set. Comparison with the SOS case studied in [BM13] shows in particular that the spectral gap of the Gibbs sampler is much less sensitive than the spectral gap of the diffusion process.
regarding the choice of the tilt parameter $h$. Moreover, as already mentioned our results are largely insensitive to the choice of the potential $V$. On the other hand, we cannot handle perturbations of a convex potential, since we strongly rely on the FKG inequality and other monotonicity properties which in general do not hold if the convexity assumption is dropped.

Interface models of the form (1.3) are also commonly studied in the discrete setting, namely when the heights $x_i$ are restricted to take only integer values, in which case they form natural models for the interface separating two distinct phases in low temperature spin systems; see, e.g. [BIV00]. For the discrete SOS model, estimates that are tight up to multiplicative constants for the spectral gap and the mixing time of the Gibbs sampler were obtained in [MS12]. We believe that our main results Theorem 1.1 and Theorem 1.2 can be extended to include this case as well, with small modifications in the proof. Certainly more challenging would be the determination of the spectral gap and mixing time of the local dynamics for the discrete SOS interfaces considered e.g. in [Pos97, CMT12], where only $\pm 1$ increments of the height are allowed at each update.

The problem of determining whether a given Markov chain exhibits the cutoff phenomenon or not keeps attracting a lot of attention. While a general theory is still out of reach, more and more instances of the phenomenon are being understood. Most of the known results concern Markov chains with finite state space, see e.g. the monograph [LPW17]. Especially closely related to our analysis here are the results concerning the exclusion process [Lac16, LL19]. As in our recent work [CLL20], one of the motivations in the present paper is to investigate the phenomenon for Markov chains with continuous state space. Our previous paper [CLL20] establishes the cutoff phenomenon for a heat bath dynamics over the simplex, when the target distribution is uniform or some log-concave generalization thereof. While our assumptions on the potential $V$ here are general enough to handle target distributions $\pi_N$ from a very large family of log-concave measures, we note that they do not include the measures on the simplex considered in [CLL20] since the positivity constraint characterizing the simplex would require dropping the polynomial growth condition.

1.3. Overview. Section 2 introduces several tools and presents some estimates to be used in the sequel. In Section 3 we establish the lower bound on the mixing time of Theorem 1.2 by identifying an initial condition for which the process remains far from equilibrium until the putative mixing time: this initial condition is built in such a way that rather explicit computations can be performed on the law of the image through $f_N$ (from (1.7)) of the process. A first upper bound on the mixing time is obtained in Section 4: it catches the correct order but not the precise constant. This bound allows us to determine the spectral gap of the generator. In Sections 5 and 6 we refine the upper bound of the previous section by estimating, under some appropriate coupling, the merging time of two processes starting from a (random) ‘maximal’ initial condition and any arbitrary initial condition, and by proving that the process starting from this maximal initial condition reaches equilibrium by the putative mixing time.
2. Main tools

2.1. The gradient dynamics. The process $(\eta(t))_{t \geq 0}$ defined on $\mathbb{R}^N$ by the increments

$$\eta_k(t) = (X_k - X_{k-1})(t)$$

is also Markovian. We sometimes use the notation $\eta(t) = \nabla X(t)$. To describe its evolution, we introduce some notation.

Given $a \in \mathbb{R}$, we define the resampling potential $W_a$ as

$$W_a(u) := V(a + u) + V(a - u) - 2V(a),$$

and set

$$\theta_a(u) := \frac{e^{-W_a(u)}}{Z(a)} \quad \text{with} \quad Z(a) := \int_{\mathbb{R}} e^{-W_a(u)} \, du.$$  \hspace{1cm} (2.1)

The function $W_a(u)$ is symmetric (with respect to $u$, but not with respect to $a$ in general), convex and non-negative. It is minimized at 0 where it admits the value 0. Our assumption $V \in \mathcal{C}$ ensures that $Z(a)$ is finite. Note also that

$$\rho_{-a,a}(u) = \theta_a(u) \quad \text{and} \quad \rho_{b,c}(u) = \theta_{\frac{c-b}{2}}(u - \frac{c + b}{2}).$$  \hspace{1cm} (2.2)

The dynamics of the gradients is then described as follows. For each $k \in [1, N - 1]$ at rate one $(\eta_k, \eta_{k+1})$ jumps to $(\tilde{\eta}_k - U, \tilde{\eta}_k + U)$ where $\tilde{\eta}_k = \frac{\eta_{k+1} + \eta_k}{2}$ and $U$ is a r.v. with density $\theta_{\tilde{\eta}_k}$. The associated Markov generator is given by

$$\tilde{\mathcal{L}} f(\eta) = \sum_{k=1}^{N-1} \int \left( f(\eta^{(k,u)}) - f(\eta) \right) \theta_{\tilde{\eta}_k}(u) \, du,$$

where $f : \mathbb{R}^N \mapsto \mathbb{R}$ and $\eta^{(k,u)}$, $u \in \mathbb{R}$, denotes the vector $\eta$ with the pair $(\eta_{k-1}, \eta_k)$ replaced by $(\tilde{\eta}_k - u, \tilde{\eta}_k + u)$. Note that the invariant measure $\pi_N$, in terms of the gradient variables $\eta$, is nothing but the product probability measure with density proportional to $\otimes_{k=1}^{N} e^{-V}$, conditioned on $\sum_{k=1}^{N} \eta_k = 0$.

2.2. The action on linear functions. The generators $\mathcal{L}$ and $\tilde{\mathcal{L}}$ take a particularly simple form when applied to linear functions. If $g_k$ denotes the coordinate map $g_k : x \mapsto x_k$ then

$$\mathcal{L} g_k(x) = \frac{x_{k-1} + x_{k+1}}{2} - x_k = \frac{1}{2} \Delta x_k,$$

where $\Delta$ denotes the discrete Laplacian. Summation by parts and (2.3) then show that for every $j \in [1, N - 1]$ the map $f_N^{(j)} : \Omega_N \mapsto \mathbb{R}$ given by

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

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

$$\lambda_N^{(j)} := 1 - \cos \left( \frac{j \pi}{N} \right).$$  \hspace{1cm} (2.5)

Thus, linear functions form an invariant subspace, and the spectrum of $-\mathcal{L}$ restricted to that subspace consists of the $N$ eigenvalues

$$0 =: \lambda_N^{(0)} < \lambda_N^{(1)} < \lambda_N^{(2)} < \cdots < \lambda_N^{(N-1)}.$$
In the case \( j = 1 \), we simply write \( f_{N} \) for \( f_{N}^{(1)} \) and \( \lambda_{N} \) for \( \lambda_{N}^{(1)} \). In particular, it follows that \( \text{gap}_{N} \leq \lambda_{N} \). Theorem \ref{thm:gap} will establish that \( \lambda_{N} \) is actually equal to the spectral gap of \( L \).

2.3. **General spectral gap considerations.** Next, we give a rather general characterization of the spectral gap. Consider a reversible Markov process \((X_{t})_{t\geq 0}\) on a measurable space \( \Omega \) with generator \( L \) and stationary distribution \( \pi \). Assume that \( L \) is self-adjoint in \( L^{2} = L^{2}(\Omega,\pi) \), and define its spectral gap as

\[
\text{gap} = \inf_{f \in L^{2}} \frac{\langle f, -L f \rangle}{\langle f, f \rangle},
\]

where we write \( \langle \cdot, \cdot \rangle \) for the scalar product in \( L^{2} \).

Given a probability measure \( \nu \) on \( \Omega \), we let \( P_{\nu}^{t} \) denote the distribution of \( X_{t} \) starting with initial condition \( \nu \). Finally for a probability measure \( \nu \ll \pi \) we let \( \| \nu \|_{\infty} \) denote the \( L^{\infty} \) norm of density \( d\nu / d\pi \).

**Proposition 2.1.** The spectral gap satisfies

\[
\text{gap} = -\sup_{\| \nu \|_{\infty} < \infty} \limsup_{t \to \infty} \frac{1}{t} \log \| P_{\nu}^{t} - \pi \|_{TV}.
\]

If furthermore \( \Omega \) is a topological space exhausted by compact sets (and equipped with its Borel \( \sigma \)-algebra) we can restrict the supremum to \( \nu \) with compact support.

**Proof.** Suppose \( \nu \) is a probability measure on \( \Omega \) with \( \| \nu \|_{\infty} < \infty \). Let \( \rho \) and \( \rho_{t} \) denote respectively the density of \( \nu \) and \( P_{\nu}^{t} \) with respect to \( \pi \). Then \( \rho_{t} = e^{tL} \rho \), and the spectral theorem implies

\[
\| P_{\nu}^{t} - \pi \|_{TV} = \frac{1}{2} \| \rho_{t} - 1 \|_{L^{1}(\pi)}
\]

\[
\leq \frac{1}{2} \| \rho_{t} - 1 \|_{L^{2}(\pi)} \leq \frac{1}{2} \| \rho - 1 \|_{L^{2}(\pi)} e^{-\text{gap} t}.
\]

This proves that the spectral gap is at most the right hand side in (2.7). The other inequality requires a bit more work.

Let us first treat the simpler case where \( -\text{gap} \) is an eigenvalue of \( L \). Let \( f \) be a normalized eigenfunction such that \( L f = -\text{gap} f \). Assume without loss of generality that the positive part \( f_{+} \) satisfies \( \| f_{+} \|_{L^{2}(\pi)} \geq 1/2 \) (if not take the negative part). Given \( M > 0 \), consider the bounded density

\[
\rho_{M} = \frac{f_{+} \wedge M}{\| f_{+} \wedge M \|_{L^{1}(\pi)}}.
\]

By monotone convergence and using \( \| f_{+} \|_{L^{1}(\pi)} \leq \| f_{+} \|_{L^{2}(\pi)} \leq 1 \),

\[
\lim_{M \to \infty} \langle \rho_{M}, f \rangle = \frac{\| f_{+} \|_{L^{2}(\pi)}^{2}}{\| f_{+} \|_{L^{1}(\pi)}} \geq \frac{1}{2}.
\]

Let us thus fix \( M \) sufficiently large so that

\[
\langle \rho_{M}, f \rangle = \langle \rho_{M} - 1, f \rangle \geq \frac{1}{3}.
\]
Recall that \( \|\mu_1 - \mu_2\|_{TV} = \sup \{ \int h \, d(\mu_1 - \mu_2) : \|h\|_{\infty} \leq 1 \} \). If \( \nu \) has density \( \rho_M \), then
\[
\|P^\nu_t - \pi\|_{TV} \geq \int \frac{\rho_M}{\|\rho_M\|_{\infty}} \, d(P^\nu_t - \pi) = \int \frac{(\rho_M - 1)}{\|\rho_M\|_{\infty}} \, d(P^\nu_t - \pi)
\]
(\ref{2.8})
\[
= \frac{\langle e^{tL}(\rho_M - 1), (\rho_M - 1) \rangle}{\|\rho_M\|_{\infty}}.
\]
If \( f \) is an eigenfunction, then \( \rho_M - 1 = \langle \rho_M - 1, f \rangle f + g \), where \( g \) is orthogonal to \( f \) and \( e^{tL}g \) is orthogonal to \( f \). Therefore,
\[
\|P^\nu_t - \pi\|_{TV} \geq \frac{(\rho_M - 1, f)^2}{\|\rho_M\|_{\infty}} \langle e^{tL}f, f \rangle \geq \frac{1}{9\|\rho_M\|_{\infty}} e^{-t\text{gap}}.
\]
This implies the desired bound in the case where \(-\text{gap}\) is an eigenvalue of \( L \).

If \(-\text{gap}\) is not an eigenvalue we argue as follows. Let \( E_\delta \) denote the spectral projector of \(-L\) associated to the interval \([\text{gap}, \text{gap} + \delta]\), and let \( H_\delta \) denote the corresponding closed subspace of \( L^2(\pi) \). Suppose that \( f \) is normalized and \( f \in H_\delta \). Let \( \rho_M \) be defined as above and notice that (\ref{2.8}) continues to hold. Since \( \rho_M = E_\delta \rho_M + g \), where \( g \in H_\delta^\perp \) and \( e^{tL}g \in H_\delta^\perp \), one has
\[
\langle e^{tL}(\rho_M - 1), (\rho_M - 1) \rangle \geq e^{-(\text{gap} + \delta)}\|E_\delta \rho_M\|_{L^2(\pi)}^2.
\]
(\ref{2.9})
Since \( f \) is normalized and \( f \in H_\delta \), one has \( \|E_\delta \rho_M\|_{L^2(\pi)}^2 \geq \langle \rho_M, f \rangle^2 \). In conclusion, we have shown that if \( \nu \) has density \( \rho_M \) then
\[
\|P^\nu_t - \pi\|_{TV} \geq \frac{1}{9\|\rho_M\|_{\infty}} e^{-t(\text{gap} + \delta)}.
\]
By the arbitrariness of \( \delta \) this implies the desired inequality. If \( \Omega \) is exhausted by compact sets then we can modify the definition of \( \rho_M \) to make it compactly supported.

2.4. **Monotone grand coupling.** We will consider two partial orders on interface configurations:
\[
x \leq y \iff \forall k \in [0, N], \quad x_k \leq y_k.
x \preceq y \iff \forall k \in [1, N], \quad (x_k - x_{k-1}) \leq (y_k - y_{k-1}).
\]
(\ref{2.10})
Note that \( \leq \) is a natural partial order in both spaces \( \Omega_N \) and \( \tilde{\Omega}_N \) while \( \preceq \) is only relevant for the enlarged space \( \tilde{\Omega}_N \) (recall the definition in \( \text{(1.1)} \)).

We present a global coupling of the trajectories \( X^x \) (and therefore \( \eta \)) starting from all possible initial conditions \( x \) which preserves both types of monotonicity.

**Lemma 2.2.** There exists a coupling of \( \{X^x, x \in \tilde{\Omega}_N\} \) such that
- If \( x \leq y \), then \( X^x(t) \leq X^y(t) \) for all \( t \geq 0 \);
- If \( x \preceq y \), then \( X^x(t) \equiv X^y(t) \) for all \( t \geq 0 \).

**Proof.** The coupling is a version of the usual graphical construction (see e.g. \cite{Lig05}). To each \( k \in [1, N - 1] \) we associate a Poisson clock process \( (T^{(k)}_i)_{i \geq 1} \) whose increments are i.i.d. rate one exponentials, and a sequence \( (U^{(k)}_i)_{i \geq 1} \) of i.i.d. uniform r.v. Then, for every \( x \in \tilde{\Omega}_N \), \( (X^x(t))_{t \geq 0} \) is a càdlàg process that only evolves at
the update times \((T_t^{(k)})_{k \in \{1,N-1\},t \geq 1}\). More precisely at time \(t = T_t^{(k)}\), if \(U_t^{(k)} = u\) then the \(k\)-th coordinate is updated as follows

\[
X_k^x(t) = F_{X_{k-1}(t_-),X_{k+1}(t_-)}^{-1}(u) \quad \text{and} \quad X_j^y(t) := X_j^y(t_-) \quad \text{for} \quad j \neq k,
\]

where for \(b,c\) we define \(F_{b,c} : \mathbb{R} \to [0,1]\) as

\[
F_{b,c}(t) = \int_{-\infty}^t \rho_{b,c}(u) \, du.
\]

By construction, the law of \(X^x\) under this coupling is the desired one. To check that this coupling preserves the partial order “\(\preceq\)”, it is sufficient to check that for every \(t \in \mathbb{R}\)

\[
(2.11) \quad b \leq b' \text{ and } c \leq c' \implies F_{b,c}(t) \geq F_{b',c'}(t).
\]

For the partial order “\(\preceq\)” it suffices to show that if \(c - b \leq c' - b'\) then

\[
(2.12) \quad F_{b,c}(t + b) \geq F_{b',c'}(t + b') \quad \text{and} \quad F_{b,c}(t - b) \leq F_{b',c'}(c' - t).
\]

We start with (2.11). As \(\rho_{b,c}\) and \(\rho_{b',c'}\) are positive, continuous and integrate to the same value, there must exist \(u_0\) such that \(\rho_{b,c}(u_0) = \rho_{b',c'}(u_0)\). Let us show that \(u_0\) must satisfy

\[
(2.13) \quad \left\{ \begin{array}{l}
\forall u \leq u_0, \quad \rho_{b',c'}(u) \leq \rho_{b,c}(u), \\
\forall u \geq u_0, \quad \rho_{b',c'}(u) \geq \rho_{b,c}(u).
\end{array} \right.
\]

We note that the desired inequality (2.11) is a simple consequence of (2.13). To prove (2.13), we set \(W_{b,c} = \log \rho_{b,c}\), and show that \(W_{b',c'} - W_{b,c}\) is nondecreasing. Indeed, everywhere except on a countable set, \(W_{b',c'} - W_{b,c}\) is differentiable and we have by convexity

\[
(W_{b',c'} - W_{b,c})'(u) = V'(u - b) - V'(u - b') - V'(c - u) + V'(c - u) \geq 0.
\]

This proves (2.13). Now (2.12) only needs to be proved for \(b = b' = 0\) by translation invariance. With this in mind, the first inequality in (2.12) is a consequence of (2.11). Regarding the second inequality in (2.12), we observe that it is equivalent to

\[
\tilde{F}_{c',0}(-t + c') \geq \tilde{F}_{c,0}(-t + c),
\]

where \(\tilde{F}\) is the distribution function associated to the potential \(\tilde{V}(\cdot) = V(\cdot)\). The later inequality is then exactly of the same form as the first inequality in (2.12): since \(\tilde{V}\) satisfies the same assumptions as \(V\) we are done.

\[\square\]

2.5. The sticky coupling. In this section we construct a coupling of two trajectories \((X^x(t))_{t \geq 0}\) and \((X^y(t))_{t \geq 0}\) which is aimed at minimizing the merging time. This coupling is also monotone, that is if \(x \leq y\) then \(X^x(t) \leq X^y(t)\) at all times.

In contrast with that of the previous section, this construction cannot naturally be extending to a grand-coupling. It can (and will) also be used for two processes \(X^{(1)}\) and \(X^{(2)}\) with initial conditions \(X^{(1)}(0)\) and \(X^{(2)}(0)\) sampled according to some prescribed distributions on \(\Omega_N\).

As in the previous construction, to each \(k \in \{1,N-1\}\) we associate a Poisson clock process \((T_t^{(k)})_{t \geq 1}\) whose increments are i.i.d. rate one exponentials. Let us
now describe how the updates are performed. If $T_i^{(k)} = t$ for some $i$ we resample the values of $X^x_k$ and $X^y_k$. We use the short hand notation
\begin{equation}
\rho_x := \rho_{X^x_{k-1}(t_\cdot),X^x_{k+1}(t_\cdot)}, \quad \rho_y := \rho_{X^y_{k-1}(t_\cdot),X^y_{k+1}(t_\cdot)},
\end{equation}
and set
\begin{equation}
p(t,k) := \int_{\mathbb{R}} \rho_x(u) \wedge \rho_y(u) \, du.
\end{equation}

Finally we define three probability measures $\nu_1$, $\nu_2$ and $\nu_3$ with densities proportional to $(\rho_x - \rho_y)_+$, $(\rho_x \wedge \rho_y)$ and $(\rho_y - \rho_x)_+$ (in the case were $\rho_x = \rho_y$ we can set $\nu_1$ and $\nu_3$ to be the Dirac mass at 0, or any other arbitrary distribution). The update then goes as follows

- With probability $p = p(t,k)$, we set $X^x_k(t) = X^y_k(t)$, and we draw their common value according to $\nu_2$.
- With probability $q = 1 - p$, we draw $X^x_k(t)$ and $X^y_k(t)$ independently with respective distributions $\nu_1$ and $\nu_3$.

To see that this coupling preserves “≤” notice that if the configurations are ordered before the update (or more specifically if $X^x_{k\pm 1}(t_\cdot) \leq X^y_{k\pm 1}(t_\cdot)$) then there exists $u_0$ such that $\nu_1$ is supported on $(-\infty, u_0]$ and $\nu_3$ on $[u_0, \infty)$, the latter fact being a direct consequence of $(2.13)$.

**Remark 2.3.** More formally, we can define the coupling using, on top of the clock process, $4$ sequences of independent uniform random variables on $[0,1]$ for each coordinate, from which the updates are defined in a deterministic fashion: we couple if and only if the first uniform is smaller than $p$ and we use the three other uniforms to sample independent random variables with distribution $\nu_1$, $\nu_2$ and $\nu_3$ respectively.

2.6. **FKG inequalities.** Recall the partial order $\leq$ introduced in $(2.10)$. We say that $f : \Omega_N \to \mathbb{R}$ is increasing if
\[ x \leq y \Rightarrow f(x) \leq f(y). \]

For two probability measures $\mu, \nu$ on $\Omega_N$, we write $\mu \leq \nu$ and say that $\mu$ is stochastically dominated by $\nu$ if for all increasing $f : \Omega_N \to \mathbb{R}$ one has $\mu(f) \leq \nu(f)$. We also say that a set $A \subset \Omega_N$ is increasing if the map $1_A$ is increasing. Finally, for any two configurations $x, y \in \Omega_N$ we introduce the configurations $\min(x, y), \max(x, y)$ defined as
\[ \min(x, y)_i := \min(x_i, y_i), \quad \max(x, y)_i := \max(x_i, y_i). \]

**Proposition 2.4** (FKG inequalities). If $f, g$ are increasing then
\[ \pi_N(fg) \geq \pi_N(f) \pi_N(g). \]

Furthermore if $A, B \subset \Omega_N$ are increasing and satisfy
\[ \{x \in A \text{ and } y \in B\} \Rightarrow \min(x, y) \in B, \]
then
\begin{equation}
\pi_N(\cdot \mid A) \geq \pi_N(\cdot \mid B).
\end{equation}
Proof. By [Pre74, Thm 3], the first part of the statement is granted if we have for all \(x, y \in \Omega_N\)

\[(2.17)\quad H(\max(x, y)) + H(\min(x, y)) \leq H(x) + H(y).\]

The convexity of \(V\) is sufficient to ensure this inequality, see for instance [Gia02, Appendix B1]. We turn to the second part of the statement. Set \(\mu_A := \pi_N(\cdot | A)\) and define \(\mu_B\) similarly. The densities of these measures are proportional to \(e^{-H_A}, e^{-H_B}\) where

\[H_A(x) := \begin{cases} H(x) & \text{if } x \in A, \\ +\infty & \text{if } x \notin A. \end{cases}\]

By [Pre74, Prop 1], it suffices to check that for all \(x, y \in \Omega_N\)

\[H_A(\max(x, y)) + H_B(\min(x, y)) \leq H_A(x) + H_B(y).\]

This is granted by (2.17) and the assumption on \(A, B.\)

A useful example to keep in mind is as follows. Let \(\mathcal{X} \subset \{1, \ldots, N - 1\}\) be a set of labels and define, for some \(a \in \mathbb{R}\), the sets

\[(2.18)\quad A_i = \{x : x_i \geq a\}, \quad A = \bigcap_{i \in \mathcal{X}} A_i, \quad B = \bigcup_{i \in \mathcal{X}} A_i.\]

Then \(A, B \subset \Omega_N\) satisfy the requirement in Proposition 2.4 and the inequality (2.16) is crucially used in the proof of Proposition 6.5 below.

2.7. Absolute continuity. It will be useful to compare the conditional probability measure \(\pi_N\) to an unconditional measure under which the increments \(\eta_i\) are independent and have the same mean. We need a preliminary lemma.

**Lemma 2.5.** Let \(V \in \mathcal{C}\) and set \(I := (V'_-, V'_+),\) where \(V'_\pm\) are defined in (1.2).

The function \(\psi : I \mapsto \mathbb{R}\) defined by

\[(2.19) \quad \psi(\lambda) = \frac{\int u e^{-V(u) + \lambda u} \, du}{\int e^{-V(u) + \lambda u} \, du}\]

is bijective from \(I\) to \(\mathbb{R}\).

**Proof.** The function \(\psi\) is increasing since for any \(\lambda \in I:\)

\[(2.20) \quad \psi'(\lambda) = \frac{\int (u - \psi(\lambda)) e^{-V(u) + \lambda u} \, du}{\int e^{-V(u) + \lambda u} \, du} > 0.\]

To prove that \(\psi\) is surjective we show that \(\psi(\lambda) \uparrow \infty\) when \(\lambda \uparrow V'_+\) (a similar argument proves that \(\psi(\lambda) \downarrow -\infty\) when \(\lambda \downarrow V'_-\)). When \(V'_+ < \infty\) this follows from the fact that \(\psi(\lambda)\) is the derivative of \(\log \int e^{-V(u) + \lambda u} \, du\) which itself tends to infinity when \(\lambda \rightarrow V'_+\) (by convexity we have that \(V(u) \leq V'_+ u + V(0)\)). When \(V'_+ = \infty\) it is a standard task to check that \(\log \int e^{-V(u) + \lambda u} \, du\) grows superlinearly at infinity.

As a consequence there exists \(\lambda \in I\) such that for \(\widetilde{V}(x) : = V(x) - \lambda x\) we have \(\int x e^{-\widetilde{V}(x)} \, dx = 0\). Note that \(\widetilde{V} \in \mathcal{C}\). Let \(\nu_N\) be the probability measure under which the r.v. \(\eta_k, k \in [1, N]\) are i.i.d. with density proportional to \(e^{-\widetilde{V}}\). Under \(\nu_N\), the expectation of the r.v. \(x_N\) vanishes. The next lemma shows that the law of a fixed proportion (bounded away from 1) of the \(\eta_k\)'s under \(\pi_N\) is absolutely
continuous with respect to the law of the same r.v. under $\nu_N$, uniformly in $N$. The point here is that $\nu_N$ remains a product law and is therefore more tractable.

**Lemma 2.6.** Fix $a \in (0, 1)$ and write $N_a := \lfloor aN \rfloor$ for all $N \geq 1$. There exists a constant $C_a > 0$ such that for all $N \geq 1$ and all positive bounded measurable functions $f : \mathbb{R}^{N_a} \to \mathbb{R}_+$ we have

$$\pi_N[f(\eta_1, \ldots, \eta_{N_a})] \leq C_a \nu_N[f(\eta_1, \ldots, \eta_{N_a})].$$

**Remark 2.7.** Note that from exchangeability the above statement is also valid for the functional of an arbitrary subset of the increments of cardinality smaller than $aN$.

**Corollary 2.8.** There exist two constants $c, C > 0$ such that for all $u > 0$

$$\pi_N\left(\|x\|_{\infty} \geq u\sqrt{N}\right) \leq Ce^{-c[\sqrt{Nu}]}$$

and

$$\pi_N\left(\max_{i \in [1, N]} |\eta_i| \geq u\right) \leq Ne^{-cu}.$$

**Proof of Corollary 2.8.** For (2.22) we apply the lemma to $1_{|\eta_i| \geq u}$ and use the union bound. For (2.21) we only need to prove that

$$\pi_N\left(x_i \geq u\sqrt{N}\right) \leq \frac{C}{2}e^{-c[\sqrt{Nu}]}$$

for $i \leq N/2$ (the corresponding lower bound and the case $i \geq N/2$ can be dealt with by symmetry) and use union bound. Lemma 2.6 applied to $a = N/2$ allows to prove the bound for $\nu_N$ under which $x_i$ is a sum of IID exponentially integrable random variables. Reproducing the classic upper bound computation in the proof of Cramer's Theorem (see e.g. [DZ09, Chapter 1]) we have

$$\nu_N\left(x_i \geq u\sqrt{i}\right) \leq e^{-i\varphi(u^{-1/2})}$$

where $\varphi(x) := \max_{t \geq 0} \left( tx - \log \int e^{-\tilde{V}(u) + tu} du \right)$. Our assumptions on $\tilde{V}$ imply that $\varphi$ has quadratic behavior at zero. Since in addition $\varphi$ is convex, we have necessarily $\varphi(x) \geq cx^2$ and $x$ for all $x > 0$ (for some positive $c > 0$) yielding (2.23).

**Proof of Lemma 2.6.** Let $\sigma^2$ be the variance of $\eta$ under the measure with density proportional to $e^{-\tilde{V}}$. Let $q_k$ be the density of the random variable $\eta_1 + \ldots + \eta_k$ under $\nu_N$. The Local Limit Theorem [Pet75, Th. VII.2.7] gives

$$\lim_{k \to \infty} \sup_{y \in \mathbb{R}} |\varepsilon(k, y)| = 0,$$

where we define

$$\varepsilon(k, y) = \sigma \sqrt{k} q_k(y\sigma\sqrt{k}) - g(y),$$

and $g$ is the density of the standard Gaussian distribution. Since $g$ is maximized at 0, for $k$ sufficiently large we may estimate

$$\sup_{z \in \mathbb{R}} \sqrt{k} q_k(z) \leq 2g(0)$$
One can check that, for any \( f_0 \) which is a bounded measurable function of \( x_1, \ldots, x_{N-1} \), we have
\[
\pi_N[f_0] = \lim_{\delta \downarrow 0} \frac{\nu_N[f_0 \mathbf{1}_{x_N \in [-\delta, \delta]}]}{\nu_N(x_N \in [-\delta, \delta])}.
\]
Taking \( f \) as in the statement of the lemma, we thus get for all \( N \) sufficiently large,
\[
\pi_N[f(\eta_1, \ldots, \eta_N_a)] = \frac{\nu_N[f(\eta_1, \ldots, \eta_{N_a})q_{N-N_a}(-x_{N_a})]}{q_N(0)} \leq \frac{2g(0)}{\sqrt{N - N_a}q_{N}(0)} \nu_N[f(\eta_1, \ldots, \eta_{N_a})].
\]

The result of the lemma follows by adjusting the value of \( C_a \) in order to cover also the small values of \( N \).

### 2.8. Technical estimates for the resampling probability.

The goal of this subsection is to collect some useful estimates on the resampling distribution of our dynamics. All the constants are allowed to depend on the potential \( V \in \mathcal{C} \) and on nothing else. Let us mention before starting that, as a consequence of Assumptions (i) and (ii) on \( V \), we have
\[
|V'(u)| \leq C(1 + |u|)^{K-1},
\]
for all \( u \) where \( V \) is differentiable, and by continuity, also for the derivatives on the left and on the right when they differ. All issues concerning differentiability appearing in the proofs below can be resolved in this fashion, so we will not mention them.

Our first estimate guaranties that our distribution is sufficiently spread-out. Recall (2.1).

**Lemma 2.9.** There exists a constant \( C > 0 \) such that
\[
Z(a) \geq \frac{1}{C(1 \lor |a|^K)}.
\]

As a consequence, we have
\[
\|\rho_{b,c}\|_{\infty} \leq C(1 \lor |c - b|^K).
\]

Our second lemma ensures that the distribution \( \rho_{b,c} \) displays an exponential decay outside of the interval \([b, c] \).

**Lemma 2.10.** There exists positive constants \( \alpha \) and \( C \) such that for all \( s \geq 0 \) and all \( b, c \in \mathbb{R} \) we have
\[
\int_{(b \lor c) + s}^{\infty} \rho_{b,c}(u) du \leq Ce^{-\alpha s}.
\]

Symmetrically we have
\[
\int_{(b \land c) - s}^{-\infty} \rho_{b,c}(u) du \leq Ce^{-\alpha s}.
\]
In particular, the variance of the random variable with density \( \rho_{b,c} \) satisfies, for some possibly different choice of \( C \), for every \( b, c \in \mathbb{R} \):  

\begin{equation}
\text{Var}(\rho_{b,c}) := \int_{\mathbb{R}} \left( u - \frac{b + c}{2} \right)^2 \, du \leq C(|b - c| + 1)^2
\end{equation}

Finally the third lemma allows us to control the total variation distance between the distributions associated with \( \rho_{b,c} \) and \( \rho_{b',c'} \).

**Lemma 2.11.** There exists a constant \( C \) such that for any \( b, b', c, c' \)

\begin{equation}
q = \frac{1}{2} \int_{\mathbb{R}} |\rho_{b,c}(u) - \rho_{b',c'}(u)| \, du \leq C \Delta (1 \lor |c - c'|^K),
\end{equation}

where \( \Delta := (|b' - b| + |c' - c|)/2 \).

**Proof of Lemma 2.9** From (2.2), it suffices to prove (2.26) with \( a \) and \( \theta_a \) instead of \( (c - b) \) and \( \rho_{b,c} \). Since \( W_a(u) \geq W_a(0) = 0 \), we have \( \|\theta_a\|_\infty = Z(a)^{-1} \) and therefore we only need to prove (2.25). Let \( z_a \) be defined as the unique positive solution of \( e^{-W_a(z_a)} = 1 \). Existence and uniqueness of \( z_a \) follow from convexity of \( W_a \) and the fact that \( W_a \) is minimized at \( W_a(0) = 0 \). We have

\begin{equation}
Z(a) \geq \int_{|a| \leq z_a} e^{-W_a(z_a)} du \geq z_a.
\end{equation}

If \( z_a > 1 \), then (2.25) immediately follows. We now assume that \( z_a \leq 1 \). Writing

\( \log 2 = W_a(z_a) = \int_{0}^{z_a} (V'(a + u) - V'(a - u)) \, du \),

we deduce from (2.24) that

\begin{equation}
\log 2 \leq 2z_a \max_{|u-a| \leq 1} |V'(u)| \leq 2CZ(a)(|a| + 2)^K,
\end{equation}

thus concluding the proof.

**Proof of Lemma 2.10** Using translation invariance (2.2) we only need to prove an upper bound for the tail distribution associated with \( \theta_a \), that is, for \( \int_{|a|+s} \theta_a(u) \, du \).

Also, at the cost of changing the value of \( C \), we can assume that \( s \geq s_0 \) for some sufficiently large \( s_0 \geq 1 \) independent of \( a \). Recalling that \( \theta_a \) integrates to 1 and is decreasing on \( \mathbb{R}_+ \), we have

\begin{equation}
\int_{|a|+s}^\infty \theta_a(u) \, du \leq \int_{|a|+s_0}^\infty \theta_a(u) \, du \leq \frac{1}{|a| + s_0} \int_s^\infty \theta_a(|a| + s) \, du.
\end{equation}

We can then conclude if we show that for all \( u \geq s_0 \)

\begin{equation}
\frac{\theta_a(|a| + u)}{\theta_a(|a| + s_0)} \leq Ce^{-\alpha u}.
\end{equation}

From our assumptions (i) and (iii) on the potential \( V \), we have

\( \lim_{u \to \pm \infty} V'(u) - V'(-u) \in (0, \infty) \).

Therefore, there exist \( \alpha > 0 \) and \( s_0 > 1 \) such that for all \( u \geq s_0 \), we have \( V'(u) - V'(-u) \geq \alpha \). We then compute for all \( u \geq s_0 \)

\begin{align*}
\partial_u [\log \theta_a(|a| + u)] &= V'(a - |a| - u) - V'(|a| + a + u) \\
&\leq V'(-u) - V'(u) \leq -\alpha,
\end{align*}

\(\square\)
which readily yields
\begin{equation}
\frac{\theta_u(|a| + u)}{\theta_u(|a| + s_0)} \leq e^{-\alpha(u - s_0)} .
\end{equation}

\textbf{Proof of Lemma 2.11.} Note that we may assume \(|b' - b| + |c' - c| \leq 1\), otherwise the result is trivial. In particular, \(|c' - b'| \leq |c - b| + 1\). Using the triangle inequality
\[ |\rho_{b,c}(u) - \rho_{b',c'}(u)| \leq |\rho_{b,c}(u) - \rho_{b,c'}(u)| + |\rho_{b,c'}(u) - \rho_{b',c'}(u)| \]
it is sufficient to treat the case where either \(b = b'\) or \(c = c'\). By translation invariance we reduce to the case \(b = b' = 0\) (the case \(c = c'\) can be treated symmetrically). Interchanging the variables if necessary, we may further assume that \(Z(c/2) \geq Z(c'/2)\). Setting
\[ \Gamma_{c,c'}(u) = V(c' - u) - V(c - u) . \]
we observe that
\begin{align}
(2.36) & \quad q = \int_{\mathbb{R}} \rho_{0,c}(u) \left( 1 - \frac{Z(c)}{Z(c')} e^{-\Gamma_{c,c'}(u)} \right)_+ du \\
(2.37) & \quad \leq \int_{\mathbb{R}} \rho_{0,c}(u) \left( 1 - e^{-\Gamma_{c,c'}(u)} \right)_+ du \leq \int_{\mathbb{R}} \rho_{0,c}(u)(\Gamma_{c,c'}(u))_+ du.
\end{align}
Using (2.24) and \(2\Delta = |c - c'|\) we have
\begin{equation}
|\Gamma_{c,c'}(u)| \leq C\Delta(|u| + |c| + 1)^{K-1}.
\end{equation}
We can conclude using
\begin{equation}
\int_{\mathbb{R}} \rho_{0,c}(u)|u|^{K-1} du \leq C(1 \vee |c|)^K,
\end{equation}
which follows from Lemma 2.10. \qed

3. Lower bound on the mixing time

\textbf{Proposition 3.1.} There exists a constant \(c > 0\) such that, for every \(N\) and \(t \geq 0\),
\begin{equation}
d_N(t) \geq 1 - \frac{1}{1 + cN e^{-2\lambda_N t}},
\end{equation}
where \(\lambda_N = 1 - \cos(\pi/N)\). As a consequence, there exists another constant \(C\) such that, for all \(\varepsilon \in (0, 1)\),
\begin{equation}
T_N(\varepsilon) \geq \frac{1}{2\lambda_N} (\log N + \log(1 - \varepsilon) - C).
\end{equation}

To prove (3.1) we select a test function \(f\) and use the fact that if at time \(t\) the value \(f(X(t))\) is far from the equilibrium value \(\pi_N(f)\) with large probability then \(d_N(t)\) must be large. This is implemented by choosing a suitable initial condition and by estimating the first two moments of \(f(X(t))\). This is a variant of Wilson’s method [Wil04]. As for the exclusion process [Wil04] and for the Beta-sampler on the simplex [ CLI20], we take \(f = f_N\), the eigenfunction appearing in Theorem 1.1. For the remainder of this section we assume for notational simplicity that \(N\) is even and we write \(X\) for the process started from the random initial condition \(X(0)\) drawn according to the measure
\begin{equation}
\theta_N := \pi_N \left( \cdot \mid x_{N/2} = N/2, |x|_{\infty} \leq N \right).
\end{equation}
Proposition 3.2. There exists a constant $C$ such that for every $t \geq 0$
\[ \mathbb{E}[f_N(X(t))] \geq C^{-1}N^2 e^{-\lambda_N t}, \quad \text{Var}[f_N(X(t))] \leq CN^3. \]

Proof of Proposition 3.1 using Proposition 3.2

By definition,
\[ d_N(t) \geq \|P^N_t - \pi_N\|_{TV}. \]

From [LPW17, Proposition 7.12] one has
\[ \|P^N_t - \pi_N\|_{TV} \geq 1 - \left(1 + \frac{\mathbb{E}[f_N(X(t))] - \pi_N(f_N)}{2\text{Var}(f_N(X(t)) + 2\text{Var}_N(f_N))}\right)^{-1}, \]

where $\text{Var}_N(f_N)$ denotes the variance of $f_N$ with respect to $\pi_N$. Using $\pi_N(f_N) = 0$ and Fatou’s lemma for weak convergence to control $\text{Var}_N(f_N)$ through the variance $\text{Var}(f_N(X(t)))$ at $t = \infty$, Proposition 3.2 implies the estimate
\[ \|P^N_t - \pi_N\|_{TV} \geq 1 - \left(1 + \frac{Ne^{-\lambda_N t}}{4C^3}\right)^{-1}, \]

which proves (3.1) with $c = 1/4C^3$ if $C$ is the constant in Proposition 3.2. The lower bound (3.2) is a simple consequence of (3.1).

Proof of Proposition 3.2

As $f_N$ is an eigenfunction associated with the eigenvalue $-\lambda_N$, see Section 2.2, the process
\[ M_t := e^{\lambda_N t} f_N(X(t)) \]

is a martingale. In particular,
\[ \mathbb{E}[f_N(X(t))] = e^{-\lambda_N t} \mathbb{E}[f_N(X(0))] = e^{-\lambda_N t} \varphi_N(f_N). \]

Under $\pi_N = \pi_N(\cdot | x_{N/2} = N/2)$, the increments $\eta_i$, $i \in [1, N/2]$ are exchangeable and have all mean 1. The same can be said for $i \in [N/2 + 1, N]$ with mean $-1$. The distribution $\pi_N$ restricted to the variables in the first half of the segment is the distribution of $\{\eta_i + 1\}$ where the $\eta_i$ are distributed according to the measure $\pi_{N/2}$ for the shifted potential $V^+(u) = V(u + 1)$, see Remark 1.4. Similarly, for the second half of the segment with $V^-(u) = V(u - 1)$. Then, an application of Corollary 2.3 shows that for any $a_0 > 0$ there exists $c > 0$ such that for every $a \geq a_0$ and for all $N$ sufficiently large,

\[ \pi_N(\max |x_i - \pi_N(x_i)| \geq aN) \leq e^{-\alpha N}, \]

where $\pi_N(x_i) = i$ if $i \leq N/2$ and $\pi_N(x_i) = N - i$ if $i \geq N/2$. Moreover, using the Cauchy-Schwarz inequality and (2.21) one has

\[ \pi_N(f_N 1_{\|x\| > N}) = o(N^2). \]

It follows that $\varphi_N(f_N = \pi_N(f_N | \|x\| \leq N)$ satisfies
\[ \varphi_N(f_N) = \frac{2N^2}{\pi^2} (1 + o(1)). \]

Combined with (3.6) this proves the desired lower bound on $\mathbb{E}[f_N(X(t))]$.

To control the variance, we write
\[ \text{Var}[f_N(X(t))] = e^{-2\lambda_N t} \text{Var}[M_t] = e^{-2\lambda_N t} (\text{Var}[M_0] + \mathbb{E}[(M)_t]), \]

where $(M)_t$ is the increasing predictable process, or angle bracket, associated to the martingale $M_t$ defined above. The control of $\text{Var}[M_0] = \text{Var}_\varphi_N(f_N)$ can be
where $E$ with distribution $X$ bracket of the potential update at site $k$. The martingale bracket can be given an explicit expression. The contribution to the $\hat{\rho}$ of (3.12) $1$ than random variable $\mathbb{E}[(X - m)^2]$, one finds

\begin{equation}
\text{Var}_\theta(f_N) \leq \gamma_N \left( \left( f_N(x) - \sum_{i=1}^{N} \hat{\pi}_N(x_i) \sin(i\pi/N) \right)^2 \right) \leq 2CN^3.
\end{equation}

The martingale bracket can be given an explicit expression. The contribution to the bracket of the potential update at site $k$ at time $s$ is bounded by

\begin{equation}
e^{2\lambda_N s} \sin^2(k\pi/N) \left( (X_k(s) - X_k(s^-))^2 \right),
\end{equation}

where $\mathbb{E}[\cdot]$ is the expectation with respect to the resampling random variable $X_k(s)$ with distribution $p_{X_{k-1}(s^-),X_{k+1}(s^-)}$. Notice that

\begin{equation}
X_k(s) - X_k(s^-) = \frac{1}{2}(\eta_{k+1}(s^-) - \eta_k(s^-)) - U,
\end{equation}

where $U$ has distribution $\theta_{\eta_k(s^-)}$, see (2.2). Using Lemma 2.10 to estimate the variance of $U$, we see that (3.13) is bounded above by

\begin{equation}
Ce^{2\lambda_N s} \left[ 1 + \eta_k(s^-)^2 + \eta_{k+1}(s^-)^2 \right],
\end{equation}

for some constant $C > 0$. Hence,

\begin{equation}
\langle M \rangle_t \leq C \int_0^t e^{2\lambda_N s} \sum_{k=1}^{N-1} \left( 1 + \eta_k(s)^2 + \eta_{k+1}(s)^2 \right) \, ds.
\end{equation}

To conclude we prove that there exists $C > 0$ such that

\begin{equation}
\forall N \geq 1, \forall k \in [1, N], \forall s \geq 0, \quad \mathbb{E}[\eta_k(s)^2] \leq C.
\end{equation}

Indeed, (3.17) combined with (3.16) yields

\begin{equation}
e^{-2\lambda_N t} \mathbb{E}[(M)_t] \leq C \lambda_N^{-1} \leq C' N^3.
\end{equation}

By symmetry, it is sufficient to show (3.17) for $k \leq N/2$. Moreover, using (3.7) as above, we may consider the dynamics with initial distribution $\hat{\pi}_N$ instead of $\theta_N$. With slight abuse of notation we still use the notation $X$ for this process. We are going to prove a bound for $\mathbb{E}[\max(\eta_k(s), 0)^2]$, the analogous bound for the negative part being proved by a symmetric argument. Using Lemma 2.5 we fix $\lambda$ such that

\begin{equation}
\left( \int e^{-V(u)+\lambda u} \, du / (\int e^{-V(u)+\lambda u} \, du) \right) = 2.
\end{equation}

We consider the measure $\hat{\pi}_N$ under which the $\eta_i$ are IID with a distribution whose density with respect to Lebesgue is proportional to $e^{-V(u)+\lambda u}$, and note that $\hat{\pi}_N$ is an invariant measure for the generator $L$ in the enlarged state space $\tilde{\Omega}_N$. 

\[17\]
In the enlarged state space, we couple \( X \) with the process \( X' \) with initial condition distributed according to \( \tilde{\pi}_N(\cdot | x_{N/2} \geq N/2) \). Observe that the law of the increments \( (\eta_k)_{k \leq N/2} \) under \( \hat{\pi}_N \) coincides with the law of \( (\eta'_k)_{k \leq N/2} \) under \( \tilde{\pi}_N(\cdot | x_{N/2} = N/2) \). Therefore, by Lemma 2.2, \( X \) and \( X' \) can be coupled in such a way that \( \eta_k(s) \leq \eta'_k(s) \) for all \( s \geq 0 \) and \( k \leq N/2 \). Hence

\[
\mathbb{E} \left[ \max(\eta_k(s), 0)^2 \right] \leq \mathbb{E} \left[ (\eta'_k(s))^2 \right].
\]

Simple estimates for i.i.d. random variables show that \( \tilde{\pi}_N(x_{N/2} \geq N/2) \geq 1/2 \), and therefore, using the invariance of \( \tilde{\pi}_N \):

\[
\mathbb{E} \left[ \max(\eta_k(s), 0)^2 \right] \leq 2 \tilde{\pi}_N(\eta^2_k) = 2 \int u^2 e^{-V(u) + \lambda u} du / \int e^{-V(u) + \lambda u} du.
\]

\[
\square
\]

4. A FIRST UPPER BOUND AND THE SPECTRAL GAP

In this section, we establish an upper bound on the total-variation distance to equilibrium that holds for all \( N \geq 2 \). From this bound we will derive the value of the spectral gap of the generator. This upper bound is sharp enough to catch the order of the mixing time when \( N \to \infty \) but not the right prefactor: this will be sharpened in the next section. The main result of this section is formulated as follows. For a probability distribution \( \nu \) on \( \Omega_N \) we let \( B(\nu) \) denote the following quantity

\[
(4.1) \quad B(\nu) := \min_{x \sim \nu, x' \sim \pi_N} \sum_{k=1}^{N-1} \mathbb{E} \left[ |x_k - x'_k| \right]^2,
\]

where \( \mathbb{E} \) denotes the expectation with respect to a coupling of \( (x, x') \) with marginals \( \nu \) and \( \pi_N \), and the minimum is taken over all such couplings.

**Proposition 4.1.** There exists a constant \( C > 0 \) such that for any distribution \( \nu \) on \( \Omega_N \), all \( t \geq C \log N \) and all \( N \geq 2 \)

\[
(4.2) \quad \| P_t^\nu - \pi_N \|_{TV} \leq C \left( N^{1/2} B(\nu) t^C e^{-\lambda_N t} + N e^{-t} \right),
\]

where \( \lambda_N = 1 - \cos(\pi/N) \).

Before giving the proof of Proposition 4.1 we describe some of its consequences for the spectral gap and the mixing time.

4.1. **Proof of Theorem 1.1.** The upper bound in Proposition 4.1 is valid for all \( N \geq 2 \) and for all initial distributions \( \nu \), without restrictions on the maximal height. In particular, it allows us to identify the spectral gap of the generator and prove Theorem 1.1. We already saw that \( f_N \) is an eigenfunction of \( -\mathcal{L} \) associated with \( \lambda_N \). It remains to check that the latter is indeed the spectral gap of \( \mathcal{L} \). Using Proposition 2.1 it is sufficient to check that for any compactly supported distribution \( \nu \)

\[
\limsup_{t \to \infty} \frac{1}{t} \log \| P_t^\nu - \pi_N \|_{TV} \leq -\lambda_N.
\]

This follows from Proposition 4.1 since \( B(\nu) < \infty \) if \( \nu \) has compact support.
2.2 we obtain the following useful contraction bounds.

**Lemma 4.2.** For any \( x, y \in \Omega_N \), for all \( t \geq 0 \):

\[
(4.3) \quad \left( \sum_{k=1}^{N-1} \mathbb{E}[X_k^x(t) - X_k^y(t)] \right)^2 \leq N e^{-2 \lambda_N t} \sum_{k=1}^{N-1} (x_k - y_k)^2,
\]

where \( \mathbb{E} \) denotes the expectation with respect to an arbitrary coupling of \( X^x(t) \) and \( X^y(t) \). Moreover, for any distribution \( \nu \) on \( \Omega_N \) and \( t \geq 0 \), the quantity defined in (4.1) satisfies

\[
(4.4) \quad B(\nu^t) \leq B(\nu)e^{-\lambda_N t}.
\]

**Proof.** From (2.3) we have

\[
(4.5) \quad \partial_t a(t, k) = \frac{1}{2} \Delta a(t, k),
\]

where \( a(t, k) := \mathbb{E}[X_k^x(t) - X_k^y(t)] \). An orthonormal basis for \( \Delta \) on the segment \( \{1, \ldots, N-1\} \) with Dirichlet boundary condition at 0 and \( N \) is given by the eigenfunctions \( \varphi(j) \), \( j = 1, \ldots, N-1 \):

\[
(4.6) \quad \varphi_k^j = \sqrt{\frac{2}{N}} \sin \left( \frac{jk\pi}{N} \right), \quad (\Delta \varphi(j))_k = -2\lambda_k^j \varphi_k^j,
\]

where \( \lambda_k^j \) is given in (2.5). Expanding \( a(t, \cdot) \) along this basis one obtains

\[
\sum_{k=0}^{N} a(t, k)^2 \leq e^{-2 \lambda_N t} \sum_{k=0}^{N} a(0, k)^2,
\]

and the bound (4.3) follows from the Cauchy-Schwarz inequality. To prove (4.4) we argue as follows. By definition of \( B(\nu) \) we may choose a coupling \( P_0 \) of \( (\nu, \pi) \) such that

\[
(4.7) \quad \sum_{k=1}^{N-1} \mathbb{E}[|X_k^\nu(0) - X_k^\pi(0)|^2] = B(\nu)^2.
\]

Under this coupling we let \( Y \) and \( W \) denote the upper and lower envelope of \( \{X^\nu(0), X^\pi(0)\} \), setting \( Y_k = X_k^\nu(0) \vee X_k^\pi(0) \) and \( W_k = X_k^\nu(0) \wedge X_k^\pi(0) \). We have by definition

\[
\sum_{k=1}^{N-1} \mathbb{E}[|Y_k - W_k|^2] = B(\nu)^2.
\]

Now we couple four Markov chains \( X^\nu(t), X^\pi(t), X^Y(t), X^W(t) \) using the coupling \( P_0 \) to set the initial condition \( (X^Y(0) = Y \) and \( X^W(0) = W \) respectively) and using the monotone grand coupling from Section 2.4 for the dynamics. We let \( P \) denote the joint law. As the initial conditions are ordered we obtain from Lemma 2.2 that under \( P \) for any \( t \geq 0 \) we have

\[
X^W(t) \leq X^\nu(t) \leq X^Y(t) \quad \text{and} \quad X^W(t) \leq X^\pi(t) \leq X^Y(t).
\]
Therefore the argument used to prove (4.3) implies that
\[
\sum_{k=1}^{N-1} E[X_k^X(t) - X_k^W(t)]^2 \leq \sum_{k=1}^{N-1} E[Y_k - W_k]^2 \\
\leq e^{-2\lambda_N t} E[Y_k - W_k]^2 \\
= e^{-2\lambda_N t} B(\nu)^2.
\]
(4.8)

By stationarity of \( \pi \), under \( P \) the distribution of \( X_\nu(t) \) and \( X_\pi(k)(t) \) are respectively \( P_\nu(t) \) and \( \pi \), and (4.4) follows. \( \square \)

Next, we show that Proposition 4.1 provides an upper bound on the mixing time which is of order \( N^2 \log N \). This bound is off by a factor 4 with respect to Theorem 1.2. In the next section we will refine the proof in order to catch the right prefactor.

**Corollary 4.3.** For any \( \delta > 0 \), for all \( \varepsilon \in (0,1) \) and all \( N \geq N_0(\varepsilon, \delta) \) sufficiently large
\[
T_N(\varepsilon) \leq \frac{2 + \delta}{\lambda_N} \log N .
\]

**Remark 4.4.** An important observation here which is used in Section 6.1 is that not only the above estimate is also valid when the boundary condition \( x_N = 0 \) is replaced by \( x_N = hN \) (cf. Remark 1.4), but it is uniform when \( h \) takes value in a compact interval (say \([-C, C]\) for some constant \( C > 0 \)). Checking this uniformity is a tedious but rather straightforward procedure. We have chosen to omit it in the proof, but the reader can check that it boils down to making sure that all technical estimates in Section 2.8 are indeed uniform in this sense. A second observation (which can, this time, immediately be checked from the proof) is that if the bound on \( \|x\|_\infty \) is chosen to be \( N^{\alpha} \), with \( \alpha > 1/2 \) then the corresponding \( \varepsilon \)-mixing time is smaller than \( \frac{1}{\lambda_N} \log N \). Let us also remark that Corollary 4.3 is sufficient to establish the so-called pre-cutoff phenomenon, namely the fact that
\[
\limsup_{N \to \infty} \frac{T_N(\varepsilon)}{T_N(1-\varepsilon)}
\]
is uniformly bounded for \( \varepsilon \in (0,1/2) \).

**Proof.** Consider an initial condition \( x \in \Omega_N \) such that \( \|x\|_\infty \leq N \). We have \( B(\delta_x) \leq C N^{3/2} \) so that a direct application of Proposition 4.1 would yield \( T_N(\varepsilon) \leq C N^2 \log N \) for some constant \( C \) depending on \( C \) and \( N \) large enough. However one can sharpen this upper bound as follows.

By (4.4) we have for \( s \geq t \), \( B(P_{s-t}) \leq C N^{3/2} e^{-\lambda_N(s-t)} \). Now using Proposition 4.1 for \( \nu = P_{s-t}^x \) we obtain for some new constant \( C \):
\[
(4.9) \quad \|P_s^x - \pi_N\|_{TV} = \|P_{s-t}^{P_{s-t}} - \pi_N\|_{TV} \leq C \left( N^2 t^C e^{-\lambda_N s} + Ne^{-t} \right).
\]

Then choosing \( s = \frac{2 + \delta}{\lambda_N} \log N \) and \( t = (\log N)^2 \) we can conclude. \( \square \)
4.3. Proof of Proposition 4.1  

The rest of this subsection is devoted to the proof of Proposition 4.1. We are going to perform the proof for $N \geq 3$ (we require $\lambda_N < 1$ in (4.14)). For $N = 2$ since the system equilibrates after one update, we have

\begin{equation}
\|P_t^{\nu} - \pi_2\|_{TV} \leq e^{-t}. \tag{4.10}
\end{equation}

Moreover, since the total variation distance $\|P_t^{\nu} - \pi_N\|_{TV}$ is monotone as a function of $t$, we may assume without loss of generality that $t$ is an integer.

Fix $t \in \mathbb{N}$ and a distribution $\nu$ on $\Omega_N$. For notational simplicity we often write $\pi$ instead of $\pi_N$. We are going to construct a (non-Markovian) coupling $(X'(s), X''(s))_{s \in [0, t]}$ for the two processes starting with respective distributions $\nu$ and $\pi$. We let $P_t$ denote the law of this coupling. First we couple the initial conditions $X'(0), X''(0)$ in such a way that (4.7) holds. The second ingredient for our coupling is a set of independent, rate 1, Poisson clocks $(\tau_k)_{k=1}^{N-1}$ (which are independent of the initial conditions) indexed by coordinates from 1 to $N - 1$ (each $\tau_k$ is considered as a subset of $\mathbb{R}_+$). These clocks determine the update times for the coordinates of our processes. We then define the random time $T$ as the largest integer $\ell$ before $t$ such that all the Poisson clocks $\tau_k$ have rung at least once on $(\ell, t)$. More formally, we set (here $\sup \emptyset = 0$)

\begin{equation}
T := \sup \{ \ell \in [0, t] : \forall k, \tau_k \cap (\ell, t) \neq \emptyset \}. \tag{4.11}
\end{equation}

Note that we have

\begin{equation}
P_t(t - T = \ell) = \begin{cases}
(1 - e^{-1})^N & \text{if } \ell = 1, \\
(1 - e^{-\ell})^N - (1 - e^{-\ell+1})^N & \text{if } \ell \in [2, t - 1], \\
1 - (1 - e^{-t+1})^N & \text{if } \ell = t.
\end{cases} \tag{4.12}
\end{equation}

Observe that there exists $C > 0$ such that for all $N \geq 3$ and all $\ell \geq 0$

\begin{equation}
P_t(t - T \geq \ell) \leq CNe^{-\ell}. \tag{4.13}
\end{equation}

Thus, using the fact that $E[f(Z)] = f(0) + \sum_{k=1}^{\infty}[f(k) - f(k - 1)]P(Z \geq k)$ for non negative integer valued random variables $Z$ and any function $f$, provided that the sum in the r.h.s. converges, one has

\begin{equation}
E \left[ e^{\lambda_N(t - T)} \right] \leq 1 + 2CNe^{\ell(\lambda_N - 1)} \leq C', \tag{4.14}
\end{equation}

for some constant $C' > 0$.

Now we perform our coupling as follows

- For $s \leq T$, we use the monotone coupling of Subsection 2.4: At each update time we draw a uniform variable $U$ and the updated values of $X'_k, X''_k$ are constructed composing $U$ with the inverse of the conditional distribution function.

- For $s > T$, we use the sticky coupling of Subsection 2.5: At each update time we couple $X'_k$ and $X''_k$ with maximal probability.

To prove Proposition 4.1 we introduce the r.v.

\begin{equation}
A_s := \sum_{k=1}^{N-1} |X'_k(s) - X''_k(s)|, \quad s \in [0, t]. \tag{4.15}
\end{equation}
Lemma 4.5. There exist \( c', C' > 0 \) such that for all \( N \geq 2 \), all \( t \geq \log N \) and all \( \ell \in [1, t - 1] \) we have

\[
\mathbb{P}_t (X^\nu(t) \neq X^\pi(t) | \mathcal{T} = \ell) \leq C' \left( e^{-c't^2} + t^{2K+1} \mathbb{E}_t[A_{\ell} | \mathcal{T} = \ell] \right).
\]

Proof. For every \( k \in [1, N - 1] \), let us denote by \((t^{(i)}_k)_{i=1}^{n_k}\) the ordered set of update times occurring at site \( k \) on the time-interval \((\mathcal{T}, t)\). Let \( \tilde{\mathcal{F}} \) be the sigma-field generated by all the \((t^{(i)}_k)_{i=1}^{n_k}\), \( k \in [1, N - 1] \), and by the processes \( X^\nu, X^\pi \) up to time \( \mathcal{T} \). Denote by \( \tilde{\mathbb{P}}_t \) the associated conditional probability. We are going to show that, for some constant \( C > 0 \), on the event \( \{\mathcal{T} = \ell\} \) we have

\[
\tilde{\mathbb{P}}_t (X^\nu(t) \neq X^\pi(t)) \leq C \max_{k \in [1, N - 1]} n_k \left( e^{-c't^2} + t^{2K} A_{\ell} \right).
\]

and that

\[
\mathbb{E}_t \left[ \max_{k \in [1, N - 1]} n_k | \mathcal{T} = \ell \right] \leq Ct.
\]

Let us first show how we conclude from (4.17) and (4.18). Since \( \{\mathcal{T} = \ell\} \) is \( \tilde{\mathcal{F}} \)-measurable we have

\[
\mathbb{P}_t (X^\nu(t) \neq X^\pi(t) | \mathcal{T} = \ell) \leq C \mathbb{E}_t \left[ \max_{k \in [1, N - 1]} n_k \left( e^{-c't^2} + t^{2K} A_{\ell} \right) | \mathcal{T} = \ell \right]
\]

Observe that \( A_{\ell} \) and \( \max_{k \in [1, N - 1]} n_k \) are independent under \( \mathbb{P}_t (\cdot | \mathcal{T} = \ell) \). Therefore we get

\[
\mathbb{P}_t (X^\nu(t) \neq X^\pi(t) | \mathcal{T} = \ell) \leq C' \left( e^{-c't^2} + t^{2K+1} \mathbb{E}_t[A_{\ell} | \mathcal{T} = \ell] \right),
\]

as required.

Now let us prove (4.17). We introduce the event

\[
C_\ell := \{ \forall s \in [\ell, t] : \|\nabla X^\pi_s\|_\infty \leq t^2 \}.
\]

We note that \( C_\ell = \cap_{i,k} C_{\ell}^{i,k} \) where

\[
C_{\ell}^{i,k} := \{ \forall s \in [\ell, t^{(i)}_k] : \|\nabla X^\pi_s\|_\infty \leq t^2 \}.
\]

We say that the update at time \( t^{(i)}_k \) is successful if \( X^\nu(t^{(i)}_k) = X^\pi(t^{(i)}_k) \). We let \( \tau \) be the time of the first unsuccessful update among the update times \((t^{(i)}_k)_{i=1}^{n_k}, k \in [1, N - 1] \). If all the updates are successful, we set \( \tau := t \). We have

\[
\{ X^\nu(t) \neq X^\pi(t) \} \cap \{ \mathcal{T} = \ell \} \subset \{ \tau < t \} \cap \{ \mathcal{T} = \ell \},
\]

Indeed, on the event \( \{ \tau = t \} \cap \{ \mathcal{T} = \ell \} \), there is at least one update per coordinate on \((\ell, t)\) and all the updates are successful so that the two processes merge by time \( t \). Then we write

\[
\tilde{\mathbb{P}}_t (\tau < t) = \tilde{\mathbb{P}}_t (\cup_{i,k} \{ \tau = t^{(i)}_k \})
\]

\[
\leq \tilde{\mathbb{P}}_t (\cup_{i,k} C_{\ell}^{i,k}) + \tilde{\mathbb{P}}_t (\cup_{i,k} \{ \tau = t^{(i)}_k \} \cap C_{\ell}^{i,k})
\]

\[
\leq \tilde{\mathbb{P}}_t (C_{\ell}^C) + \sum_{i,k} \tilde{\mathbb{P}}_t (\{ \tau = t^{(i)}_k \} \cap C_{\ell}^{i,k}).
\]
Using Lemma 2.11 we have
\[
\widetilde{P}_t(\{\tau = t_k^{(i)}\} \cap C_i^{(k)})
= \widetilde{P}_t(\{X_k^{(i)}(t_k^{(i)}) \neq X_k^{(i)}(t_k^{(i)} - 1) \mid F_{t_k^{(i)} - 1}\} \cap \{\tau \geq t_k^{(i)}\} \cap C_i^{(k)})
\leq \widetilde{P}_t[C \max(1, \|\nabla X(t_k^{(i)} - 1)\|_\infty) K \Delta_k(t_k^{(i)} - 1) \mid F_{t_k^{(i)} - 1}\} \cap C_i^{(k)}],
\]
where
\[
2\Delta_k(s) := |X_k^{(i)}(s) - X_k^{(i)}(s - 1)| + |X_k^{(i)}(s) - X_k^{(i)}(s + 1)|.
\]
On the event \{\tau \geq t_k^{(i)}\}, all the updates are successful up to time \(t_k^{(i)}\) so that
\[
\Delta_k(t_k^{(i)} - 1) \leq \Delta_k(\ell).
\]
Consequently, we have
\[
\widetilde{P}_t(\{\tau = t_k^{(i)}\} \cap C_i^{(k)}) \leq C t^{2K} \Delta_k(\ell).
\]
Putting everything together, we find that on the event \{\mathcal{T} = \ell\} (which is \(\widetilde{F}\)-measurable):
\[
\widetilde{P}_t(X^{\nu}(t) \neq X^{\pi}(t)) \leq \widetilde{P}_t(\{\tau < t\})
\leq \widetilde{P}_t(C_\ell^0) + \left(\max_{k \in [1,N-1]} n_k\right) C' t^{2K} A_\ell.
\]
To bound the first term, we use stationarity and Corollary 2.8 to obtain
\[
\widetilde{P}_t(C_\ell^0) \leq \left(\max_{k \in [1,N-1]} n_k\right) \pi_N \left(\max_{i \in [1,N]} |\eta_i| > t^2/2\right)
\leq \left(\max_{k \in [1,N-1]} n_k\right) N e^{-c't^2}.
\]
Since \(t \geq \log N\), this yields (4.17).

Let us now estimate the conditional expectation of \(\max_{k \in [1,N-1]} n_k\). Let us first describe the conditional law of the \(n_k\)’s. Let \(G\) be the random number of Poisson clocks that have not rung on \((\ell + 1, t)\). On the event \{\mathcal{T} = \ell\} this number is positive. Given \{\mathcal{T} = \ell\} the \(n_k\)’s can be obtained as \(G\) i.i.d. Poisson r.v. of parameter 1 conditioned to be positive and \(N - 1 - G\) i.i.d. r.v. which are the independent sum of a Poisson r.v. of parameter 1 and a Poisson r.v. of parameter \(t - \ell - 1\) conditioned to be positive.

It is simple to check that the law of a Poisson r.v. of parameter \(q\) conditioned to be positive is stochastically increasing with \(q\). As a consequence of these observations, we deduce that \(\max_{k \in [1,N-1]} n_k\), conditionally given \{\mathcal{T} = \ell\}, is stochastically smaller than \(\max_{k \in [1,N-1]} Z_k\) where \(Z_k\) are i.i.d. r.v. obtained as the independent sum of a Poisson r.v. of parameter 1 and a Poisson r.v. of parameter \(t - 1\) conditioned to be positive. Recalling that a Poisson random variable \(W\) with parameter \(\lambda\) satisfies \(\mathbb{P}(W \geq k) \leq e^{-k(\log(k)/\lambda - 1)}\), and that \(t \geq \log N\), it is not difficult to check that
\[
\mathbb{E}\left[\max_{k \in [1,N-1]} Z_k\right] \leq C t,
\]
for some new constant \(C > 0\). This implies (4.18).
We now proceed to the proof of our proposition.

**Proof of Proposition 4.7** We start with an upper bound on the expectation of $A_\ell$ given $\{T = \ell\}$ for any $\ell \in [0, t - 1]$. Since up to time $T$ we use the monotone grand coupling, arguing as in the proof of Lemma 4.2, setting $Y_k = (X_k^\nu \vee X_k^\pi)(0)$ and $W_k = (X_k^\nu \wedge X_k^\pi)(0)$ one obtains

$$E_t[A_\ell | T = \ell]^2 \leq \left( \sum_{k=1}^{N-1} E_t[X_k^Y(\ell) - X_k^W(\ell) | T = \ell] \right)^2 \leq N \sum_{k=1}^{N-1} E_t[X_k^Y(\ell) - X_k^W(\ell) | T = \ell]^2 \leq Ne^{-2\lambda_N \ell} \sum_{k=1}^{N-1} E_t[Y_k - W_k]^2.$$ 

Therefore, by (4.7)

$$E_t[A_\ell | T = \ell] \leq \sqrt{NB(\nu)e^{-\lambda_N \ell}}.$$ 

By definition of the total-variation distance we have

$$\|P_t^\nu - \pi_N\|_{TV} \leq P_t(X^\nu(t) \neq X^\pi(t)) = \sum_{\ell=0}^{t-1} P_t(X^\nu(t) \neq X^\pi(t) | T = \ell) P_t(T = \ell).$$

We treat separately the case $\ell = 0$ (recall that $T = 0$ on the event where not all Poisson clocks have rung on $(0, t)$). Using (4.13) we have

$$P_t(X^\nu(t) \neq X^\pi(t) | T = 0) P_t(T = 0) \leq P_t(T = 0) \leq CN e^{-t}.$$

On the other hand, combining, (4.23) and Lemma 4.5 we find

$$\sum_{\ell=1}^{t-1} P_t(X^\nu(t) \neq X^\pi(t) | T = \ell) P_t(T = \ell) \leq C e^{-ct^2} + C'' N^{1/2} B(\nu) t^{2K+1} e^{-\lambda_N t} \mathbb{E}\left[e^{\lambda_N (t-T)}\right].$$

and we can conclude using (4.14). \(\square\)

5. Upper bound on the mixing time

5.1. **Proof strategy.** The overall strategy is similar to that in [CLL20]. First, we show that the ‘maximal’ evolution gets close to equilibrium by time $\log N / (2 \lambda_N)$. More precisely, let $\nu^\wedge$ denote the equilibrium measure $\pi$ conditioned to having $x_i \geq N$ for all $i = 1, \ldots, N - 1$. Let $X^\wedge$ denote the evolution with initial condition $\nu^\wedge$ and call $P_t^\wedge$ its law at time $t$. We have the following result.

**Theorem 5.1.** For any $\delta > 0$,

$$\lim_{N \to \infty} \|P_t^\wedge - \pi\|_{TV} = 0,$$

where $t_\delta := (1 + \delta) \log N / (2 \lambda_N)$.
Next, using Theorem 5.1 as an input, we compare the evolution $X^x$ for an arbitrary initial state $x$ with $\|x\|_\infty \leq N$ to $X^\wedge$ and show that they come close in total variation by time $t_\delta$.

**Theorem 5.2.** For any $\delta > 0$,

\[
\lim_{N \to \infty} \sup_{x \in \Omega_N: \|x\|_\infty \leq N} \|P^x_{t_\delta} - P^\wedge_{t_\delta}\|_{TV} = 0.
\]

The upper bound stated in Theorem 1.2 follows from the two results above and the triangle inequality. Although Theorem 1.2 can be deduced from Theorem 5.2 alone, the intermediate result provided by Theorem 5.1 is a crucial ingredient in our proof of Theorem 5.2.

Let us briefly explain the importance of Theorem 5.1 as an intermediary step. Our proof of Theorem 5.2 is based on a coupling argument that uses monotonicity. For this reason it is important to start with initial conditions that are ordered (for the order on $\Omega_N$). This is the case here since the random initial configuration $\nu^\wedge$ is by definition always above $x$ if $\|x\|_\infty \leq N$ (while using directly $\pi_N$ as an initial condition instead of $\nu^\wedge$ would not work).

On the other hand, our proof of Theorem 5.2 also requires to apply the equilibrium estimates of Lemma 2.6 to $X^\wedge$. It is the double requirement of having a trajectory which is already close to equilibrium and above $X^x(t)$ which makes Theorem 5.1 a necessity.

Observe that for all $t$ the density $dP^\wedge_{t_\delta}/d\pi_N$ is an increasing function. This allows for the use of various tools in order to control $\|P^\wedge_{t_\delta} - \pi\|_{TV}$, such as the FKG inequality as well as the censoring inequality. Our proof of Theorem 5.1 (which is postponed to Section 6) is entirely based on these tools and cannot be adapted to an arbitrary initial condition.

**Proof strategy for Theorem 5.2.** The remainder of this section is devoted to the proof of Theorem 5.2. From now on, the processes $X^\wedge(t)$ and $X^x(t)$ are coupled through the sticky coupling of Subsection 2.5 (we denote by $P$ the associated distribution). To prove Theorem 5.2 we want to estimate the time at which the trajectories $X^\wedge(t)$ and $X^x(t)$ merge using the auxiliary function

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

which corresponds to the area between the two configurations at time $t$. By monotonicity $A_t \geq 0$ and the merging time of the two trajectories is the hitting time of 0 by the random process $A_t$.

The control of the evolution of $A_t$ proceeds in several steps. First we use the heat equation for a time $t_{\delta/2}$ to bring the area $A_t$ between the ordered configurations $X_k^\wedge(t)$ and $X_k^x(t)$ below a first threshold equal to $N^{3/2-\eta}$ where $\eta > 0$ is a parameter that will be taken to be small depending on $\delta$. This step relies on Lemma 4.2.

In a second step, we show that within an additional time $T = O(N^2)$, with large probability, $A_t$ falls below a second threshold $N^{-\eta}$. This is a delicate step, which requires the application of diffusive estimates for super-martingales during a finite sequence of intermediate stages each running for a time $O(N^2)$. It relies
tremendously on the specificity of the sticky coupling, and also on the fact that one of the trajectories we are trying to couple is already at equilibrium (cf. Theorem 5.1).

The final step brings the area from $N^{-\eta}$ to zero, by using Proposition 5.1, the proof of which indicates that after the second threshold has been attained merging occurs with large probability as soon as every coordinate has been updated once, which by the standard coupon collector argument, takes a time of order $\log N$.

5.2. Proof of Theorem 5.2

We introduce the successive stopping times

$$T_i := \inf\{t \geq \frac{\delta}{2} : A_t \leq N^{3/2-i\eta}\}, \quad i \geq 0,$$

where $\eta > 0$ is a parameter that we are going to choose small enough depending on $\delta$.

**Step 1:** We want to show that by time $t = \frac{\delta}{2}$, $A_t$ is much smaller than $N^{3/2-4\eta}$ (here the factor 4 is present only for technical reason, and can be considered irrelevant since $\eta$ is arbitrary).

**Lemma 5.3.** Setting $A = A_N := \{ T_4 = \frac{\delta}{2} \}$, and fixing $\eta \leq \frac{\delta}{20}$ we have

$$\lim_{N \to \infty} P(A_N) = 1.$$

**Proof.** As in Lemma 4.2

$$\mathbb{E}[A_t] \leq \sqrt{N} \sum_{k=1}^{N-1} (\mathbb{E}[X_k^N(0) - X_k^N])^2 e^{-\lambda_N t} \leq 4N^2 e^{-\lambda_N t}.$$

In the last inequality we used the fact that $|x_k| \leq N$ (by definition) and the fact that $\mathbb{E}[X_k^N(0)] \leq 3N$ (cf. the proof of Proposition 6.6 for this estimate). Using this estimate for $t = \frac{\delta}{2}$ we obtain

$$\mathbb{E}[A_t] \leq 4N^{(3-\delta)/2}.$$

Since by monotonicity of the coupling, $A_t$ is positive, we can conclude using Markov’s inequality.

**Step 2:** The aim of the second step is to prove the following estimate

**Proposition 5.4.** Introduce $I := \min\{i \geq 1 : 3/2 - i\eta \leq -\eta\}$. We have

$$\lim_{N \to \infty} P(T_I \leq \frac{\delta}{2} + N^{2}/2) = 1.$$

To highlight better the main ideas of the proof, we postpone the proof of some of the technical lemmas (namely Lemma 5.5, Lemma 5.6 and Lemma 5.7) to the next subsection and focus on the main steps of the reasoning. By (4.5), we observe that $A_t$ is a super-martingale. More precisely, considering the natural filtration $(\mathcal{F}_s)_{s \geq 0}$ associated with the process $(X^N, X^x)$ the conditional version of (4.5) summed along the coordinates yield

$$(5.3) \quad \mathbb{E}[A_t | \mathcal{F}_s] = A_s - \int_s^t \mathbb{E}[X_k^N(u) - X_{k-1}^N(u) + X_{N-k}^N(u) - X_{N-k-1}^N(u)] \, du \leq A_s,$$

where again we have used the fact that our coupling preserves the ordering.
To prove Proposition 5.4 we would like to use diffusive estimates in the form of [CLL20 Proposition 21] but this requires a modification of \((A_t)\) in such a way that it becomes a super-martingale with bounded jumps. We thus define
\[
\mathcal{R}_i := \inf\{s \geq T_{i-1} : A_s \geq N^{3/2-(i-2)\eta}\}, \quad i \geq 1,
\]
\[
Q := \inf\{s \geq t_{3/2} : \|\nabla X^\wedge_s\|_\infty \geq (\log N)^2\}.
\]
and \(\mathcal{R} := \inf_{i\in[1, I]} \mathcal{R}_i \land Q\). We consider the super-martingale
\[
M_t := \begin{cases} 
A_t & \text{if } t < \mathcal{R} \\
A_\mathcal{R} \land N^{2/(i-2)\eta} & \text{if } t \geq \mathcal{R} \text{ and } \mathcal{R} = \mathcal{R}_i < \mathcal{R}_{i+1}, \\
A_{t} & \text{if } t \geq \mathcal{R} \text{ and } \mathcal{R} = Q < \inf_{i\in[1, I]} \mathcal{R}_i.
\end{cases}
\]

The construction of \(M_t\) is designed so that with large probability it coincides with \(A_t\). To show this we introduce a collection of events:
\[
\mathcal{B} = \mathcal{B}_N := \{\forall t \in [t_{3/2}, N^3] : \|\nabla X^\wedge_t\|_\infty < (\log N)^2\},
\]
\[
\mathcal{C} = \mathcal{C}_N := \{\forall i \in [4, I], \forall s \geq T_{i-1} : A_s \leq N^{3/2-(i-3)2\eta}\},
\]
\[
\mathcal{D} = \mathcal{D}_N := \{\forall t \in [t_{3/2}, N^3] : \max(\|X^\wedge_t\|_\infty, \|X^\vee_t\|_\infty) \leq \sqrt{N} (\log N)^2\},
\]

Note that on \(\mathcal{B} \cap \mathcal{C}\), we have \(\mathcal{R} \geq N^3\). We will show that \(\mathcal{B}, \mathcal{C}, \mathcal{D}\) are all very likely. This step of the proof requires Theorem 5.1 as an input.

**Lemma 5.5.** We have \(\lim_{N \to \infty} \mathbb{P}(\mathcal{B}_N \cap \mathcal{C}_N \cap \mathcal{D}_N) = 1\).

Then using the method developed in [CLL20] we control the increments of \(\langle M \rangle\), which denotes the angle bracket of the martingale part of \(M_t\), between each consecutive \(T_i\).

**Lemma 5.6.** The probability of the event
\[
\mathcal{E} = \mathcal{E}_N := \left\{\forall i \leq I, \langle M\rangle_{T_i} - \langle M\rangle_{T_{i-1}} \leq 4N^{3-2(i-2)\eta}\right\}.
\]
satisfies \(\lim_{N \to \infty} \mathbb{P}(\mathcal{E}_N) = 1\).

Then in order to compare \(T_i - T_{i-1}\) to \(\langle M\rangle_{T_i} - \langle M\rangle_{T_{i-1}}\), we prove the following estimates on the bracket derivative

**Lemma 5.7.** When \(\mathcal{B} \cap \mathcal{C} \cap \mathcal{D}\) holds, for all \(t \in [t_{3/2}, N^3 \cap T_i]\) we have
\[
\partial_t \langle M \rangle_t \geq \frac{1}{8(\log N)^C} \min \left(\frac{M_t}{\sqrt{N}}, \frac{M_t^2}{N}\right),
\]
for some constant \(C > 0\).

Then we can conclude by simply combining the control we have on the bracket increments, and that on the bracket derivative. The following Lemma, combined with the fact that \(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C} \cap \mathcal{D} \cap \mathcal{E}\) holds with large probability, implies Proposition 5.4.

**Lemma 5.8.** On the event \(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C} \cap \mathcal{D} \cap \mathcal{E}\) we have
\[
\forall i \in [5, I], T_i - T_{i-1} \leq 2^{-i} N^2.
\]
In particular we have \(T_I \leq t_{3/2} + N^2 / 2\).
Proof. We work on the event $A \cap B \cap C \cap D \cap E$. Let $j$ be the smallest $i \geq 5$ such that $T_i - T_{i-1} > 2^{-i} N^2$ and assume that $j \leq I$. Then, $T_{j-1} + 2^{-j} N^2 \leq N^3$ so that by Lemma 5.7
\[
\langle M \rangle_{T_{j-1} + 2^{-j} N^2} - \langle M \rangle_{T_{j-1}} \geq \left( \log N \right)^{-C'} 2^{-j} N^2 \left( N^{1-j \eta} \land N^{2(1-jn)} \right),
\]
where we use the fact that $A_t \geq N^{3/2-j\eta}$ if $t \leq T_j$ and $M_t = A_t$ on $B \cap C$. Moreover since we work on $E$ we have
\[
\langle M \rangle_{T_{j-1} + 2^{-j} N^2} - \langle M \rangle_{T_{j-1}} \leq 4 N^{3-2(j-2)\eta}.
\]
These two inequalities are incompatible for $N$ large enough and the lemma is proved.

Step 3: The last step consists in bringing the area to 0 within a short time after $t_{\delta/2} + N^2/2$. Introduce the event
\[
G := \{ A_{t_{\delta/2} + N^2/2} \leq N^{-\eta/2} \}.
\]
The following estimates can be proved as a variant of Lemma 4.5.

Lemma 5.9. There exists $C > 0$ such that for any $t \geq \log N$ we have
\[
P(X^\wedge(t_{\delta/2} + N^2/2 + t) \neq X^\wedge(t_{\delta/2} + N^2/2 + t) \mid G) \leq C(N e^{-t} + t^{2K + 1} N^{-\eta/2}) .
\]
Proof. This is an adaptation of the argument in Lemma 4.5. Denote by $(t_k^{(i)})_{i=1}^{n_k}$ the ordered set of updates times occurring at site $k$ on the time-interval $(t_{\delta/2} + N^2/2, t_{\delta/2} + N^2/2 + t)$. Let $\tilde{F}$ be the sigma-field generated by all the $(t_k^{(i)})$ and by $X^\wedge(t_{\delta/2} + N^2/2), X^\wedge(t_{\delta/2} + N^2/2)$, and let $\tilde{P}$ be the associated conditional probability. Define $\mathcal{H} := \{ \forall k \in [1, N-1] : n_k \geq 1 \}$. Then, the very same arguments as in the proof of (4.17) show that on the $\tilde{F}$-measurable event $G \cap \mathcal{H}$ we have
\[
\tilde{P}(X^\wedge(t_{\delta/2} + N^2/2 + t) \neq X^\wedge(t_{\delta/2} + N^2/2 + t)) \leq C \left( \max_{k \in [1, N-1]} n_k \right) (e^{-ct^2} + t^{2K} B),
\]
with
\[
B = \sum_{k=1}^{N-1} |X^\wedge_k(t_{\delta/2} + N^2/2) - X^\wedge_k(t_{\delta/2} + N^2/2)|
\]
\[
= \sum_{k=1}^{N-1} X^\wedge_k(t_{\delta/2} + N^2/2) - X^\wedge_k(t_{\delta/2} + N^2/2) .
\]
Furthermore, given $\mathcal{H}$, the $n_k$'s are i.i.d. Poisson r.v. of parameter $t$ conditioned to be positive. Therefore, reasoning as in (4.22), for all $t \geq \log N$
\[
\mathbb{E} \left[ \max_{k \in [1, N-1]} n_k \mid \mathcal{H} \right] \leq C t .
\]
Finally, we have
\[
P(\mathcal{H}^c) \leq N e^{-t} .
\]
Putting everything together we obtain the stated estimate. \hfill $\Box$

With the help of this final step, we can conclude the proof.
Proof of Theorem 5.2} By the Martingale Stopping Theorem, since \((A_t)_{t \geq 0}\) is a supermartingale then \((A_{T_t+t})_{t \geq 0}\) is also a càdlàg non-negative super-martingale (for the adequate filtration). A maximal inequality (sometimes referred to as Ville’s Maximal Inequality see [Dur19, Exercise 8.4.2] for the discrete time version and also [Vil39])

\[
\mathbb{P} \left( \sup_{t \geq 0} A_{T_{t+t}} > N^{-\eta/2} \right) \leq \mathbb{E} [A_{T_1}] N^{\eta/2}.
\]

Therefore,

\[
\lim_{N \to \infty} \mathbb{P} \left( \sup_{t \geq 0} A_{T_{t+t}} > N^{-\eta/2} \right) = 0.
\]

Combining this with Lemma 5.8 we deduce that the probability of the event \(G\) goes to 1. Applying Lemma 5.9 we thus deduce that for \(t = 2 \log N\) we have

\[
\lim_{N \to \infty} \mathbb{P}(X^\wedge(t_{\delta/2} + N^2/2 + t) \neq X^x(t_{\delta/2} + N^2/2 + t) \mid G) = 0.
\]

Since all our estimates hold uniformly over all \(x \in \Omega_N\) with \(\|x\|_\infty \leq N\), this suffices to deduce (5.1).

\[\square\]

5.3. Proof of the technical estimates of step 2.

Proof of Lemma 5.3} To prove that \(\mathcal{B}\) and \(\mathcal{D}\) have small probability, we are going to show that similar events have small probability for the stationary version of our Markov chain \((X^x(t))_{t \geq 0}\) and then use Theorem 5.1. By a simple coupling argument, for any \(\mathcal{A} \subset \Omega_N\) we have

\[
\mathbb{P}(\exists t \in [t_{\delta/2}, N^3] : X^\wedge_t \in \mathcal{A}) \leq \|P_{t_{\delta/2}} - \pi\|_{TV} + \mathbb{P}(\exists t \in [0, N^3 - t_{\delta/2}] : X^x_t \in \mathcal{A}),
\]

where, with slight abuse of notation we denote by \(\mathbb{P}\) the distribution of \(X^x\), the Markov chain starting from the equilibrium distribution.

By symmetry arguments (using the fact that \(\hat{V}(x) := V(-x)\) satisfies \(\hat{V} \in \mathcal{C}\)), (5.8) remains true upon replacing \(X^\wedge\) by \(X^\vee\) the dynamics with initial distribution \(\pi(\cdot \mid \forall i \in [1, N - 1], x_i \leq -N)\).

The first term in the r.h.s. of (5.8) goes to zero by Theorem 5.1. To bound the second term, we use a standard subdivision scheme and estimates on the invariant measure. More precisely, if one subdivides \([0, N^3]\) into intervals of length \(N^{-6}\) then with a probability \(1 - O(N^{-1})\), there are at most one resampling event per interval. Since the process is stationary, we can bound the second term in the r.h.s. of (5.8) by \(N^9 \pi_N(\mathcal{A}) + CN^{-1}\). To prove that \(\lim_{N \to \infty} \mathbb{P}(\mathcal{B}_N) = 0\) use (5.8) with \(\mathcal{A} = \{\|\nabla x\|_\infty > (\log N)^2\}\), and apply Corollary 2.8 which entails that \(N^9 \pi_N(\mathcal{A}) \leq N^{-1}\).

We turn now to \(\mathcal{D}\). Using (5.8) and the argument above with \(\mathcal{A} = \{\|x\|_\infty > \sqrt{N}(\log N)^2\}\) and Corollary 2.8 we deduce that

\[
\lim_{N \to \infty} \mathbb{P}(\exists t \in [t_{\delta/2}, N^3] : \|X^\wedge_t\|_\infty > \sqrt{N}(\log N)^2) = 0.
\]

and similarly for \(X^\vee\). To get a similar estimate for \(X^x\) it is sufficient to observe that from Lemma 2.2 \(X^x\) is stochastically dominated by \(X^\vee\) and stochastically
dominates $X^\gamma$, so that we can deduce from (5.9) the desired bound for $\max_k X^\gamma_k(t)$ and $\min_k X^\gamma_k(t)$ respectively, concluding the proof of $\lim_{N \to \infty} P(D_N) = 1$.

Finally let us focus on the event $C_N$. For every $i \geq 1$, by the Martingale Stopping Theorem and Ville’s Maximal Inequality (as in (5.7)) we have

$$
\mathbb{P} \left( \sup_{t \geq 0} A_{T_{i-1} + t} > N^{3/2-(i-3/2)\eta} \right) \leq \mathbb{E}[A_{T_{i-1}}] N^{-3/2+(i-3/2)\eta} \leq N^{-\eta/2}.
$$

Since $I$ is a fixed non-random integer, a union bound shows that $\lim_{N \to \infty} P(C_N) = 1$. □

**Proof of Lemma 5.6** The proof follows from a diffusivity bound developed in an earlier work [CLL20, Proposition 21], applied to the super-martingales

$$
M^{(i)}_s = M_{s + T_{i-1}},
$$

whose jump sizes are bounded above by $N^{3/2-(i-2)\eta}$. We refer to [CLL20] for more intuition about this inequality. □

To prove Lemma 5.7 we will require an intermediate technical result derived from the preliminary work of Section 2.8 which allows us to estimate the bracket derivative. Define

$$
\delta X_k(t) := X_k^\gamma(t) - X_k^\delta(t).
$$

**Lemma 5.10.** When $B \cap C$ holds, then for all $t \in [t^{\delta/2}, N^3 \wedge T_1]$ where $\partial_t \langle M \rangle_t$ is differentiable (all $t$ except a random countable set)

$$
\partial_t \langle M \rangle_t \geq \frac{1}{2} \sum_{k=1}^{N-1} \left[ (\delta X_k(t))^2 \wedge (\log N)^{-C_K} \right],
$$

for some constant $C_K > 0$.

**Proof of Lemma 5.7 assuming Lemma 5.10** Write $A_t = U_t + V_t$ where, for some $a > 0$:

$$
U_t = \sum_{k=1}^{N-1} \delta X_k(t) 1_{\{\delta X_k(t) < a\}}, \quad V_t = \sum_{k=1}^{N-1} \delta X_k(t) 1_{\{\delta X_k(t) \geq a\}}.
$$

The Cauchy-Schwarz inequality shows that

$$
U_t^2 \leq N \sum_{k=1}^{N-1} (\delta X_k(t))^2 1_{\{\delta X_k(t) \leq a\}}.
$$

Take $a = (\log N)^{-1/4}C_K$. If $U_t \geq A_t/2$, then Lemma 5.10 implies

$$
\partial_t \langle M \rangle_t \geq \frac{A_t^2}{8N}.
$$

If on the other hand $V_t \geq A_t/2$, then letting $n_t$ denote the number of indices $k$ such that $\delta X_k(t) \geq a$, Lemma 5.10 implies

$$
\partial_t \langle M \rangle_t \geq \frac{1}{2} (\log N)^{-C_K n_t}.
$$

Since $0 \leq \delta X_k(t) \leq 2 \max(\|X^\gamma_k\|_\infty, \|X^\delta_k\|_\infty)$, on the event $D$ we get

$$
n_t \geq \frac{1}{2\sqrt{N} (\log N)^2} \sum_{k=1}^{N-1} \delta X_k(t) 1_{\{\delta X_k(t) \geq a\}} \geq \frac{A_t}{4\sqrt{N} (\log N)^2}.
$$
Recall the sticky coupling of Subsection 2.5, in particular the laws $\nu_i$ defined therein. The derivative of the angle bracket $\partial_t(M)_t$ admits an explicit expression which can be derived from the sticky coupling description. For any $t \in [T_{i-1}, T_i \wedge R)$

\begin{equation}
\partial_t(M)_t = \sum_{k=1}^{N-1} \left( (1 - q_k)(\delta X_k(t_-))^2 + q_k \mathbb{E}[Y^2 \mid \mathcal{F}_t] \right),
\end{equation}

where

$$
Y = (Y^\wedge - Y^x - \delta X_k(t_-)) \wedge (R - M_{t_-}) , \quad R := N^{3/2-(i-2)\eta}
$$

and $(Y^\wedge, Y^x)$ are, conditionally given $\mathcal{F}_t$, independent r.v. with densities $\nu_3$ and $\nu_4$ respectively. The expression (5.15) simply comes from the fact that for each $k$, $M_t$ will jump by an amount $\delta X_k(t_-)$ with probability $1 - q_k$ and by an amount $Y$ with probability $q_k$. Note that the truncation with $R - M_{t_-}$ in the variable $Y$ comes from the definition of $M$ in terms of $A$.

We now work on the event $B \cap C$. From Lemma 2.9 we have

\begin{equation}
q_k \leq C(\delta X_k)(\log N)^{2K},
\end{equation}

for all $k$, where we use the notation

$$
\delta X_k = \frac{1}{2}(X_{k+1}^\wedge + X_{k-1}^\wedge - X_{k+1}^x - X_{k-1}^x).
$$

To prove Lemma 5.10 it is then sufficient to show that if $q_k \geq 1/2$ then

\begin{equation}
\mathbb{E}\left[ Y^2 \mid \mathcal{F}_{t_-} \right] \geq (\log N)^{-C_K},
\end{equation}

for some constant $C_K > 0$. Note that under the event $C$ we have $R - M_{t_-} \geq R/2$. Moreover, if $q_k \geq 1/2$, because of the event $B$ by Lemma 2.9 the density of the random variable $\tilde{Y} := Y^\wedge - Y^x - \delta X_k(t_-)$ is bounded above by $L := C'(\log N)^{2K}$. We next observe that we may assume $R \geq 2$. Indeed, if $R \leq 2$ and $q_k \geq 1/2$, then by (5.16) we also have $\delta X_k \geq (2C)^{-1}(\log N)^{-2K}$ and

$$
(2C)^{-1}(\log N)^{-2K} \leq M_{t_-} \leq N^{3/2-(i-3/2)\eta} = N^{-\eta/2} R \leq 2N^{-\eta/2},
$$

thus raising a contradiction. Hence assuming $R - M_{t_-} \geq R/2$ and $R \geq 2$ we may estimate

$$
\mathbb{E}\left[ Y^2 \mid \mathcal{F}_{t_-} \right] \geq \mathbb{E}\left[ \tilde{Y}^2 \wedge (R/2)^2 \mid \mathcal{F}_{t_-} \right] \geq \int_0^{1/2} 2v \mathbb{P}(|\tilde{Y}| > v \mid \mathcal{F}_{t_-}) dv.
$$

The bounded density property implies $\mathbb{P}(|\tilde{Y}| > v \mid \mathcal{F}_{t_-}) \geq 1 - 2Lv \geq 1/2$ for all $v \in [0, (4L)^{-1}]$. It follows that

$$
\mathbb{E}\left[ Y^2 \mid \mathcal{F}_{t_-} \right] \geq \int_0^{(4L)^{-1}} v dv = \frac{1}{32L^2}.
$$

This proves (5.17).

□
6. Proof of Theorem 5.1

The proof is based on ideas first introduced in [Lac16] for card shuffling by adjacent transpositions. An adaptation to the continuous setting was later developed in [CLL20], for the specific case of the adjacent walk on the simplex. Here we are going to follow the proof of [CLL20, Proposition 14], with some minor modifications due to the different setting. We start by recalling the Peres-Winkler censoring inequality.

6.1. Censoring. The censoring inequality of Peres and Winkler [PW13] compares the distance to equilibrium at time $t$ for two Markov processes, one of which is obtained as a censored version of the other by omitting some of the updates according to a given censoring scheme. The version of the result that we need here is formulated as Proposition 6.1 below. The proof is an adaptation to the present setting of the original argument for monotone finite spins systems in [PW13]. For completeness we give a brief self-contained account below.

A censoring scheme $C$ is defined as a càdlàg map

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

where $\mathcal{P}(A)$ denotes the set of all subsets of a set $A$. The subset $C(s)$, at any time $s \geq 0$, represents the set of labels whose update is to be suppressed at that time. More precisely, given a censoring scheme $C$, and an initial condition $x \in \Omega_N$, we write $P_{t,C}$ for the law of the random variable obtained by starting at $x$ and applying the standard graphical construction (see Section 2.4) with the proviso that if label $j$ rings at time $s$, then the update is performed if and only if $j \notin C(s)$.

In particular, the uncensored evolution $P_t$ corresponds to $P_{t,C}$ when $C(s) \equiv \emptyset$.

Let $\mathcal{S}_N$ denote the set of probability measures $\mu$ on $\Omega_N$ which are absolutely continuous with respect to $\pi_N$ and such that the density $d\mu/d\pi_N$ is an increasing function on $\Omega_N$. Recall the notation $\mu \leq \nu$ for stochastic domination.

**Proposition 6.1.** If $\mu \in \mathcal{S}_N$, and $C$ is a censoring scheme, then for all $t \geq 0$

$$\|\mu P_t - \pi_N\|_{TV} \leq \|\mu P_{t,C} - \pi_N\|_{TV}. \quad (6.1)$$

The proof is a consequence of the next two lemmas.

**Lemma 6.2.** If $\mu, \nu$ are two probability measures on $\Omega_N$ such that $\mu \in \mathcal{S}_N$ and $\mu \leq \nu$, then

$$\|\mu - \pi_N\|_{TV} \leq \|\nu - \pi_N\|_{TV}. \quad (6.2)$$

**Proof.** Setting $\varphi = d\mu/d\pi_N$, and $A = \{\varphi \geq 1\}$,

$$\|\mu - \pi_N\|_{TV} = \mu(A) - \pi_N(A). \quad (6.3)$$

Since $A$ is increasing, $\mu(A) \leq \nu(A)$, and therefore

$$\|\mu - \pi_N\|_{TV} \leq \nu(A) - \pi_N(A) \leq \|\nu - \pi_N\|_{TV}. \quad (6.4) \square$$
Let $Q_i : L^2(\Omega_N, \pi_N) \rightarrow L^2(\Omega_N, \pi_N)$, $i = 1, \ldots, N - 1$, denote the integral operator

\begin{equation}
Q_i f(x) = \int f(x^{(i,u)})\rho_{x_{i-1},x_{i+1}}(u)du,
\end{equation}

so that $Q_i f$ is the expected value of $f$ after the update of label $i$; see (1.4). If $\mu$ is a probability on $\Omega_N$, we write $\mu Q_i$ for the probability measure defined by

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

**Lemma 6.3.** If $\mu \in S_N$ then $\mu Q_i \in S_N$ and $\mu Q_i \leq \mu$, for all $i = 1, \ldots, N - 1$.

**Proof.** Set $\varphi = d\mu/d\pi_N$. Then $\mu Q_i$ has density $Q_i \varphi$ with respect to $\pi_N$. Since $\varphi$ is increasing, for any $x, y \in \Omega_N$ with $x \leq y$, from (6.5) and Lemma 2.2 (or more precisely (2.11)) it follows that

$$
Q_i \varphi(x) \leq Q_i \varphi(y).
$$

Therefore $\mu Q_i \in S_N$. To prove the stochastic domination $\mu Q_i \leq \mu$, we show that $\mu Q_i(g) \leq \mu(g)$ for any bounded measurable increasing function $g$. Notice that

$$
\mu Q_i(g) = \pi_N [\varphi Q_i g] = \pi_N [(Q_i \varphi)(Q_i g)].
$$

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

$$
\mu Q_i(g) \leq \pi_N [Q_i(\varphi g)] = \pi_N [\varphi g] = \mu(g).
$$

\hfill \Box

**Proof of Proposition 6.7.** By Lemma 6.2 it is sufficient to prove that $\mu P_t \in S_N$ and $\mu P_t \leq \mu P_t^C$ for all $t$. By conditioning on the realization $T_t$ of the Poisson clocks $\mathcal{T}^{(j)}$, $j \in [1, N - 1]$ up to time $t$ in the graphical construction, the uncensored evolution at time $t$ has a distribution of the form

\begin{equation}
\mu^z = \mu Q_{z_1} \cdots Q_{z_n},
\end{equation}

where $z := (z_1, \ldots, z_n) \in [1, N - 1]^n$ is a fixed sequence, while the censored evolution at time $t$ has distribution of the form $\mu^z'$, where $z'$ denotes a sequence obtained from $z$ by removing some of its entries. Taking the expectation over $T_t$ then shows that it is sufficient to prove that $\mu^z \in S_N$ and $\mu^z \leq \mu^z'$ for any pair of such sequences $z, z'$. Lemma 6.3 shows that $\mu^z \in S_N$ for any $\mu \in S_N$ and any sequence $z$. To prove $\mu^z \leq \mu^z'$ we may restrict to the case where $z$ and $z'$ differ by the removal of a single update, say $z_j$, so that

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

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 6.3 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, (cf. Equation (2.11)). \hfill \Box
6.2. Relaxation of skeletons. For any integer \( K \geq 2 \), consider the \( K - 1 \) labels \( u_i := \lfloor iN/K \rfloor, \ i = 1, \ldots, K - 1 \). We consider the evolution of the heights

\[
Y_i(t) = X_{u_i}(t), \quad i = 1, \ldots, K - 1,
\]

which will be referred to as the \( K \)-skeleton of the interface \( X(t) \).

**Proposition 6.4.** Fix an integer \( K \geq 2 \). Let \( \mu_t = P_t^\Lambda \) and let \( \bar{\mu}_t \) denote the marginal of \( \mu_t \) on the \( K \)-skeleton \( \{Y_i(t), i = 1, \ldots, K - 1\} \). If \( \bar{\pi}_N \) denotes the corresponding equilibrium distribution, then for any fixed \( \delta > 0 \), with \( t_\delta = (1 + \delta)\log \frac{N}{\text{gap}_N} \) one has

\[
\lim_{N \to \infty} \|\bar{\mu}_{t_\delta} - \bar{\pi}_N\|_{TV} = 0.
\]

Following [Lac16], the proof of Proposition 6.4 is based on a subtle use of the FKG inequality together with an explicit estimate on the expected value of the variables \( Y_i(t) \). Given a probability \( \mu \) on \( \Omega_N \), we write \( \bar{\mu} \) for the marginal of \( \mu \) on the \( K \)-skeleton \( y := (y_1, \ldots, y_{K-1}) \), where \( y_i = x_{u_i} \) for each \( i = 1, \ldots, K - 1 \).

We use the following notation for the area associated to \( K \)-skeleton variables \( y_i = x_{u_i} \):

\[
W = \sum_{i=1}^{K-1} y_i,
\]

and write \( \mu(W) = \bar{\mu}(W) \) for the expected value of \( W \) under \( \mu \).

**Proposition 6.5.** For any \( \varepsilon > 0 \), \( K \geq 2 \), there exists \( \eta = \eta(K, \varepsilon) > 0 \) such that for all \( N \geq 2 \), \( \mu \in \mathcal{S}_N \) one has:

\[
\mu(W) \leq \eta \sqrt{N} \implies \|\bar{\mu} - \bar{\pi}_N\|_{TV} \leq \varepsilon.
\]

The proof of Proposition 6.5 is omitted since it is identical to the proof of Proposition 36 in [CLL20]. Let us however point out that this proof uses in a crucial way the improved FKG inequality (2.16) in Proposition 2.4.

Next, we control the expected value of \( W \) at time \( t \). Let \( X^\Lambda(t) = \{X^{\Lambda}_k(t)\} \) denote the random variables with joint law \( P_t^\Lambda \).

**Proposition 6.6.** For any \( k = 1, \ldots, N - 1 \), any \( t \geq 0 \):

\[
\mathbb{E} \left[ X^\Lambda_k(t) \right] \leq 12Ne^{-\text{gap}_N t}.
\]

In particular, if \( \mu_t = P_t^\Lambda \), then for all \( t \geq 0 \):

\[
\mu_t(W) \leq 12KNe^{-\text{gap}_N t}.
\]

**Proof.** Set \( v(t) = (v_1(t), \ldots, v_{N-1}(t)) \), where \( v_k(t) = \mathbb{E}[X^\Lambda_k(t)] \). Expanding \( v_k(t) \) in the orthonormal basis (4.6), one finds \( v_k(t) = \sum_{j=1}^{N-1} a_j(t) \varphi_k^{(j)} \), where

\[
a_j(t) = \sum_{k=1}^{N-1} \varphi_k^{(j)} v_k(t).
\]

Since \( \frac{d}{dt} v_k(t) = \frac{1}{2} (\Delta v(t))_k \), it follows that

\[
a_j(t) = a_j(0)e^{-\lambda_j t}, \quad a_j(0) = \sum_{k=1}^{N-1} \varphi_k^{(j)} v_k(0).
\]

In particular, \( |a_j(0)| \leq \sqrt{2N} |v(0)|_\infty \), where \( |v(0)|_\infty = \max_k v_k(0) \). Therefore,

\[
v_k(t) \leq 2|v(0)|_\infty \sum_{j=1}^{N-1} e^{-\lambda_j t}.
\]

34
Let us show that $|v(0)|_\infty \leq 3N$ for all $N$ large enough. Raising the boundary condition from $(0,0)$ to $(2N,2N)$ and using monotonicity, we see that for all $k$ the random variable $X_k$ with distribution $\nu^\lambda$ is stochastically dominated by the random variable $X_k + 2N$ where $X_k$ has distribution $\pi(-\cdot \mid \min_i x_i \geq -N)$. The claimed monotonicity with respect to the boundary conditions can be checked using the FKG inequality for $\pi(-\cdot \mid \min_i x_i \geq n)$. Indeed the density of the measure with raised boundary with respect to the original one is equal (up to a renormalizing constant) to 

$$e^{V(2N+x_1) - V(x_1) + V(2N+x_{N-1}) - V(x_{N-1})}$$

which by convexity of $V$ is increasing for the order “≤” on $\Omega_N$. It follows that

(6.12)  
$$v_k(0) \leq 2N + \pi(x_k \mid \min_i x_i \geq -N).$$

From Corollary 2.8 and the union bound,

(6.13)  
$$\pi(\min_i x_i \geq -N) \geq 1 - Ne^{-cN},$$

for some constant $c > 0$ and all $N$ large enough. Moreover, Lemma 2.6 also shows that, uniformly in $k$,

(6.14)  
$$\pi(x_k \mid \min_i x_i \geq -N) \leq \pi(x_k^2)^{1/2} \leq C \sqrt{N},$$

for some constant $C > 0$ and all $N$ large enough. The estimates (6.12)-(6.14) imply $|v(0)|_\infty \leq 3N$ for $N$ large. From (6.11), using $\lambda_j \geq j \lambda_1$ it follows that

$$v_k(t) \leq \frac{6Ne^{-\lambda_1 t}}{1 - e^{-\lambda_1 T}}.$$ 

If $t$ is such that $e^{-\lambda_1 t} \leq 1/4$ then this implies $v_k(t) \leq 8Ne^{-\lambda_1 t}$. On the other hand if $e^{-\lambda_1 t} \geq 1/4$ then, using the monotonicity $P^\lambda_t \leq \nu^\lambda$ one has

$$v_k(t) \leq v_k(0) \leq 3N \leq 12Ne^{-\lambda_1 t}.$$

Since $\lambda_1 = \text{gap}_N$, this proves the desired upper bound. 

\textbf{Proof of Proposition 6.4} Proposition 6.6 shows that

(6.15)  
$$\lim_{N \to \infty} \frac{\bar{\mu}_s(W)}{\sqrt{N}} = 0,$$

and Proposition 6.5 shows that (6.15) is sufficient to achieve the desired convergence of $K$-skeletons. 

\textbf{6.3. Relaxation of the censored dynamics.} Consider the censored process obtained by suppressing all updates of the skeleton variables. That is, we use the censoring scheme $C$ such that $C(s) = \{u_1, \ldots, u_{K-1}\}$, $s \geq 0$.

\textbf{Proposition 6.7.} Let $P_{t,C}^x = \delta_t P_{t,C}$ and let $\pi_N(-\cdot \mid y)$ denote the equilibrium distribution given the skeleton heights $y_i = x_{u_i}, i = 1, \ldots, K - 1$. For any $\delta \in (0,1)$, define $K = \lfloor \delta^{-1} \rfloor$ and $s_\delta = \delta \frac{\log N}{2 \text{gap}_N}$, and let $B_{N,\delta}$ denote the event

(6.16)  
$$B_{N,\delta} = \left\{ x \in \Omega_N : \|x\|_\infty \leq 2N, \max_{i=1,\ldots,K} |x_{u_i}| \leq N/2K \right\}.$$ 

Then there exists $\delta_0 \in (0,1)$ such that for all fixed $\delta \in (0,\delta_0)$ and for all $N$ sufficiently large:

(6.17)  
$$\sup_{x \in B_{N,\delta}} \|P_{s_\delta,C}^x - \pi_N(-\cdot \mid y)\|_{TV} \leq \delta.$$
Proof. The censored process is a collection of $K$ independent processes each describing the evolution of an interface on a segment of length $n := \lfloor N/K \rfloor$, with fixed boundary heights $(y_{i-1}, y_i)$, where $y_i = x_{u_i}$. If $x \in B_{N, \delta}$ then the left and right boundary conditions of each interface satisfy

$$|y_{i-1} - y_i| \leq N/K \leq 2n.$$ 

Moreover, if $x \in B_{N, \delta}$ then the initial condition satisfies $\|x\|_\infty \leq 2N \leq n^2$, if $N$ is large enough. From the mixing time bound given in Corollary 4.4 (see Remark 4.4) it follows that for any given $\varepsilon \in (0, 1)$, when $N$ is sufficiently large, each individual process has $\varepsilon$-mixing time bounded above by

$$n^2 \log n \leq \frac{C}{K^2} N^2 \log(N) \leq s_\delta,$$

if $\delta > 0$ is small enough. Thus the entire censored process satisfies

$$\|P_{s_\delta, \mathcal{C}}^\varepsilon - \pi_N(\cdot | y)\|_{TV} \leq K \varepsilon. $$

The claimed inequality follows by taking $\varepsilon = K^{-1} \delta$. \hfill $\square$

6.4. Proof of Theorem 5.1 We want to prove that for any $\delta > 0$,

$$\lim_{N \to \infty} \|P_{t_\delta}^\lambda - \pi_N\|_{TV} = 0,$$

where $t_\delta = (1 + \delta) \frac{\log N}{\log \log N}$ and let $K = \lfloor \delta^{-1} \rfloor$ and let $\mathcal{C}'$ denote the censoring scheme defined by $\mathcal{C}'(s) = \emptyset$ for $s \in [0, t_\delta/2)$ and $\mathcal{C}'(s) = \{u_1, ..., u_{K-1}\}$ for $s \geq t_\delta/2$. Let also $P_{t_\delta, \mathcal{C}'}^\lambda = P_{t_\delta, \mathcal{C}}^\lambda$ denote the corresponding censored process. From Proposition 6.1, we have

$$\|P_{t_\delta}^\lambda - \pi_N\|_{TV} \leq \|P_{t_\delta, \mathcal{C}}^\lambda - \pi_N\|_{TV}. $$

We are going to construct a coupling of $P_{t_\delta, \mathcal{C}}^\lambda$ and $\pi_N$. We first couple the skeleton heights at time $t_\delta/2$. Set $\mu = P_{t_\delta/2, \mathcal{C}'}^\lambda$ and let $\mathbb{P}$ denote a coupling of $\mu$ and $\pi_N$. Let $(X, Z)$ denote the corresponding height variables, so that $X$ has distribution $\mu$ and $Z$ has distribution $\pi_N$. The coupling $\mathbb{P}$ can be chosen in such a way that the skeleton variables are optimally coupled, that is

$$\mathbb{P}(X_{u_i} = Z_{u_i}, i = 1, ..., K - 1) = 1 - \|\mu_{t_\delta/2} - \pi_N\|_{TV}. $$

Consider the event

$$E = \{x \in \Omega_N : \|x\|_\infty \leq 2N\}. $$

Monotonicity implies that $\pi_N \leq \mu \leq \nu^\lambda$ and therefore

$$\begin{align*}
\mu(E^c) & \leq N \max_i \mu(|x_i| > 2N) \\
& \leq N \max_i \pi_N(x_i < -2N) + N \max_i \nu^\lambda(x_i > 2N).
\end{align*}$$

Corollary 2.8 implies

$$\max_{i \in [1, N-1]} \pi_N(x_i < -N) \leq Ce^{-N/C},$$

for some constant $C > 0$. Raising the boundary condition from $(0, 0)$ to $(\frac{3}{2}N, \frac{3}{2}N)$ and using monotonicity, we see that for all $i$ the random variable $X_k$ with distribution $\nu^\lambda$ is stochastically dominated by the random variable $X_k + 3N/2$ where $X_k$ has distribution $\pi(\cdot | \min_i x_i \geq -N/2)$. Thus, reasoning as in (6.12) one finds

$$\max_{i \in [1, N-1]} \nu^\lambda(x_i > 2N) \leq Ce^{-N/C},$$

$$\begin{align*}
\int_{-N}^{-2N} \pi_N(x) \, dx & \leq \int_{-N}^{-2N} \mu(x) \, dx \\
& \leq \int_{-N}^{-2N} \nu^\lambda(x) \, dx \\
& \leq Ce^{-N/C}.
\end{align*}$$

36
for some constant $C > 0$. Define the event
\[ A = \{ X_{u_i} = Z_{u_i}, \ i = 1, \ldots, K - 1 \} \cap \{ X \in B_{N, \delta} \}, \]
where $B_{N, \delta}$ is given in Proposition 6.4. Let $F = \{ x \in \Omega_N : |x_{u_i}| \leq N/2K \}$ so that $\{ X \in B_{N, \delta} \} = \{ X \in E \cap F \}$. Then,
\[ (6.23) \quad A = \{ X \in E \} \cap \{ Z \in F \} \cap \{ X_{u_i} = Z_{u_i}, \ i = 1, \ldots, K - 1 \}. \]
Therefore,
\[ \mathbb{P}(A^c) \leq \| \mu_{3/2} - \pi_N \|_{TV} + \mu(X \notin E) + \pi_N(Z \notin F). \]
From (6.19)-(6.22) we have $\mu(X \notin E) \leq 2CNe^{-N/C}$. From Corollary 2.8 and the union bound one has that
\[ \pi_N(Z \notin F) \leq C_1 e^{-N/C_1}, \]
for some $C_1 = C_1(K) > 0$ independent of $N$.

If the event $A$ occurs, then we couple the interfaces at time $t_\delta = t_{3/2} + s_{3/2}$ with the optimal coupling attaining the total variation distance $\| P_{s_{3/2}, \delta} - \pi_N(\cdot) \|_{TV}$, where $C$ is as in Proposition 6.7. This shows that
\[ \| P_{t_\delta} - \pi_N \|_{TV} \leq \mathbb{P}(A^c) + \sup_{x \in B_{N, \delta}} \| P_{s_{3/2}, \delta} - \pi_N(\cdot) \|_{TV}. \]
From (6.23), Proposition 6.4 and Proposition 6.7,
\[ \limsup_{N \to \infty} \| P_{t_\delta} - \pi_N \|_{TV} \leq 2\delta. \]
The distance $\| P_{t_\delta} - \pi_N \|_{TV}$ is decreasing as a function of $\delta$, and therefore we may take $\delta \to 0$ in the right hand side above to conclude.

**Acknowledgements.** P.C. thanks University Paris-Dauphine for a funding of “Professeur Invité” and IMPA for the hospitality in the early stage of this work. C.L. acknowledges support from the grant SINGULAR ANR-16-CE40-0020-01. This work was realized in part during H.L. extended stay in Aix-Marseille University funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 837793.

**References**

[BIV00] T. Bodineau, D. Ioffe, and Y. Velenik. Rigorous probabilistic analysis of equilibrium crystal shapes. *Journal of Mathematical Physics* **41**, no. 3, (2000), 1033–1098.

[BM13] F. Barthe and E. Milman. Transference principles for log-sobolev and spectral-gap with applications to conservative spin systems. *Communications in Mathematical Physics* **323**, no. 2, (2013), 575–625.

[BW09] F. Barthe and P. Wolff. Remarks on non-interacting conservative spin systems: the case of gamma distributions. *Stochastic processes and their applications* **119**, no. 8, (2009), 2711–2723.

[Cap03] P. Caputo. Uniform poincaré inequalities for unbounded conservative spin systems: the non-interacting case. *Stochastic processes and their applications* **106**, no. 2, (2003), 223–244.

[CLL20] P. Caputo, C. Labbé, and H. Lacoin. Mixing time of the adjacent walk on the simplex. *Ann. Probab.* to appear.

[CMT12] P. Caputo, F. Martinelli, and F. L. Toninelli. Mixing times of monotone surfaces and sos interfaces: a mean curvature approach. *Communications in Mathematical Physics* **311**, no. 1, (2012), 157–189.

[Dur19] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 ed., 2019.
Cutoff phenomenon for $\nabla \phi$ interfaces

[A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.]

[T. Funaki. *Stochastic interface models. Lectures on Probability Theory and Statistics, Ecole d’Eté de Probabilites de Saint-Flour XXXIII-2003* (2005).]

[G. Giacomin. *Aspects of statistical mechanics of random surfaces*. Lecture Notes for course given at IHP. 2002.]

[H. Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.* 44, no. 2, (2016), 1426–1487. [doi:10.1214/15-AOP1004]]

[T. M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.]

[C. Labbé and H. Lacoin. Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *Annals of Probability* 47, no. 3, (2019), 1541–1586.]

[D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [ MR2466937], With a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.]

[G. Menz and F. Otto. Uniform logarithmic sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *The Annals of Probability* 41, no. 3B, (2013), 2182–2224.]

[F. Martinelli and A. Sinclair. Mixing time for the solid-on-solid model. *The Annals of Applied Probability* 22, no. 3, (2012), 1136–1166.]

[V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.]

[G. Posta. Spectral gap for an unrestricted kawasaki type dynamics. *ESAIM: Probability and Statistics* 1, (1997), 145–181.]

[C. J. Preston. A generalization of the FKG inequalities. *Comm. Math. Phys.* 36, (1974), 233–241.]

[Y. Peres and P. Winkler. Can extra updates delay mixing? *Communications in Mathematical Physics* 323, no. 3, (2013), 1007–1016.]

[J. Ville. Étude critique de la notion de collectif (1939).]

[D. B. Wilson. Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.* 14, no. 1, (2004), 274–325. [doi:10.1214/aoap/1075828054]]

DEPARTMENT OF MATHEMATICS AND PHYSICS, ROMA TRE UNIVERSITY, LARGO SAN MARI ALDO I, 00146 ROMA, ITALY.

E-mail address: caputo@mat.uniroma3.it

UNIVERSITÉ PARIS-DAUPHINE, PSL UNIVERSITY, CEREMADE, CNRS, 75775 PARIS CEDEX 16, FRANCE.

E-mail address: labbe@ceremade.dauphine.fr

IMPA, ESTRADA DONA CASTORINA 110, RIO DE JANEIRO, BRASIL.

E-mail address: lacoin@impa.br

38