FREDRICKSON-ANDERSEN ONE SPIN FACILITATED MODEL OUT OF EQUILIBRIUM

O. BLONDEL, N. CANCRINI, F. MARTINELLI, C. ROBERTO, AND C. TONINELLI

ABSTRACT. We consider the Fredrickson-Andersen one spin facilitated model (FA1f) on an infinite connected graph with polynomial growth. Each site with rate one refreshes its occupation variable to a filled or to an empty state with probability $p \in [0,1]$ or $q = 1-p$ respectively, provided that at least one of its nearest neighbours is empty. We study what happens when the evolution does not start from the equilibrium p-Bernoulli measure μ and prove convergence to equilibrium when the vacancy density q is above a proper threshold $\bar{q} < 1$. The convergence is exponential or stretched exponential, depending on the growth of the graph. In particular it is exponential on $\mathbb{Z}^d$ for $d = 1$ and stretched exponential for $d > 1$. The above result holds when the starting measure ν is such that the mean distance between two nearest empty sites is uniformly bounded. Our result can be generalized to other *non cooperative* models.

November 28, 2011

1. INTRODUCTION

Fredrickson-Andersen one spin facilitated model (FA1f) [6, 7] belongs to the class of interacting particle systems known as Kinetically Constrained Spin Models (KCSM), which have been introduced and very much studied in the physics literature to model liquid/glass transition and more generally glassy dynamics (see [16, 11] and references therein). A configuration for a KCSM is given by assigning to each vertex x of a (finite or infinite) connected graph $\mathcal{G}$ its occupation variable $\eta_x \in \{0,1\}$, which corresponds to an empty or filled site respectively. The evolution is given by Markovian stochastic dynamics of Glauber type. With rate one, each site refreshes its occupation variable to a filled or to an empty state with probability $p \in [0,1]$ or $q = 1-p$ respectively, provided that the current configuration satisfies an a priori specified local constraint. For FA1f the constraint at x requires at least one of its nearest neighbours to be empty. Note that (and this is a general feature of KCSM) the constraint which should be satisfied to allow creation/annihilation of a particle at x does not involve η_x, thus FA1f dynamics satisfies detailed balance w.r.t. the Bernoulli product measure at density p, which is therefore an invariant reversible measure for the process. Key features of FA1f model and more generally of KCSM are that a completely filled configuration is blocked (for generic KCSM other blocked configurations may occur) - namely all creation/destruction rates are identically equal to zero in this configuration -, and that due to the constraints the dynamics is not attractive, so that monotonicity arguments valid for e.g.

Work supported by the European Research Council through the "Advanced Grant" PTRELSS 228032 and the French Ministry of Education through ANR-2010-BLAN-0108.

ferromagnetic stochastic Ising models cannot be applied. Due to the above properties the basic issues concerning the large time behavior of the process are non-trivial.

In [2] it has been proved that the model on $\mathcal{G} = \mathbb{Z}^d$ is ergodic for any $q > 0$ with a positive spectral gap which shrinks to zero as $q \to 0$ corresponding to the occurrence of diverging mixing times. A key issue both from the mathematical and the physical point of view is what happens when the evolution does not start from the equilibrium measure μ. The analysis of this setting usually requires much more detailed information than just the positivity of the spectral gap, e.g. positivity of the logarithmic Sobolev constant or of the entropy constant uniformly in the system size. Since the latter requirement does certainly not hold (see Section 7.1 of [2]), even the basic question of whether convergence to μ occurs remains open in this non equilibrium setting. Of course, due to the existence of blocked configurations, convergence to μ cannot be true uniformly on the initial configuration and one could try to prove it a.e. or in mean w.r.t. a proper initial distribution $\nu \neq \mu$. From the point of view of physicists, a particularly relevant case (see e.g. [12]) is when ν is a product Bernoulli(p') measure with $p' \neq p$. In this case if $p' \neq 1$ the most natural guess is that convergence to equilibrium occurs for any local function f *i.e.*

$$\lim_{t\to\infty} \int d\nu(\eta)\mathbb{E}_\eta\big(f(\eta_t)\big) = \mu(f) \tag{1.1}$$

where η_t denotes the process started from η at time t and that the limit is attained exponentially fast. The only other case of KCSM where this result has been proved [4] is the East model, that is a one dimensional model in which the constraint at x requires that the neighbour to the right of x should be empty. The strategy used to prove convergence to equilibrium for East model in [4] relies however heavily on the oriented character of the East constraint and cannot be extended to FA1f model. We also recall that in [4] a perturbative result has been established proving exponential convergence for any one dimensional KCSM with finite range jump rates and positive spectral gap (thus including FA1f at any $q > 0$), provided the initial distribution ν is "not too far" from the reversible one (e.g. for ν Bernoulli at density p' with $p' - p$ chosen to be properly small).

Here we prove convergence to equilibrium for FA1f on a infinite connected graph $\mathcal{G}$ with polynomial growth (see the definition in sec. 2.1 below) when the equilibrium vacancy density q is above a proper threshold $\bar{q}$ (with $\bar{q} < 1$) and the starting measure ν is such that the mean distance between two nearest empty sites is uniformly bounded. This includes in particular Bernoulli measure at any $p' < 1$.The convergence is exponential or stretched exponential depending on the growth of the graph. In particular on $\mathbb{Z}^d$ it is exponential for $d = 1$ and stretched exponential for $d > 1$. When ν is the Dirac mass on some configuration this encodes the fact that the configuration has infinitely many empty sites and that, in addition to the case of the East model in [4], the distance between two nearest empty sites is uniformly bounded. Although our result can be generalized to other non cooperative models (see below for a the definition of this class), to let the paper be more readable we here consider the FA1f case. We recall that a KCSM model is *non cooperative* if there exists a finite set of the graph such that any configuration which is empty in all the sites of such set reaches the empty configuration under iteration of the naturally related bootstrap mapping [18, 2]. Otherwise the model will be called *cooperative*. The bootstrap mapping is deterministic and it is such that, given a configuration, at each step all the empty sites remain empty and all the sites whose constraint is satisfied are emptied. It follows immediately from this definition that for example East model is cooperative while FA1f model is non cooperative.

The road map follows. In section 2 we introduce the notations and give the main result. Our strategy is in section 3.1 and might be of independent interest. We first reduce the study of the evolution of the process from infinite volume to a finite ball of radius proportional to time t. Then to small sets on some ergodic component so that the log-Sobolev constant is much smaller than t. The first reduction is standard and known as the *finite speed of propagation*. The second one needs the estimate of the spectral gap of the process on the ergodic component and the study of the persistence of zeros out of equilibrium to control the behavior of the system out of the ergodic component. The persistence of zeros is studied in section 4 and gives the threshold on the density q of the reversible measure. In section 5 the proof of the main result is given. In section 6 we estimate the spectral gap of the FA1f process on the ergodic component on $\mathcal{G}$. This has been done in [3] but we present here an alternative proof, based on the technique introduced in [15], which has two qualities : it gives a more precise bound for very small q and can be generalized to other non cooperative models.

2. Notations and Result

2.1. **The graph.** Let $\mathcal{G} = (V, E)$ be an infinite, connected graph with vertex set V and edge set E and the associated graph distance $d(\cdot,\cdot)$. The set of neighbors of x, *i.e.* $y \in V$ such that $d(y, x) = 1$, will be denoted by $\mathcal{N}_x$. For all $\Lambda \subset V$ we call $\mathrm{diam}(\Lambda) = \sup_{x,y\in\Lambda} d(x, y)$ the diameter of Λ and $\partial\Lambda = \{x \in V \setminus \Lambda \colon d(x, \Lambda) = 1\}$ its (outer) boundary. Given a vertex x and an integer r, $B(x, r) = \{y \in V \colon d(x, y) \le r\}$ denotes the ball centered at x and of radius r. We introduce the growth function $F \colon \mathbb{N} \setminus \{0\} \to \mathbb{N} \cup \{\infty\}$ defined by

$$F(r) = \sup_{x\in V} |B(x, r)|$$

where $|\cdot|$ denotes the cardinality. Then we say that $\mathcal{G}$ has (k, D)-*polynomial growth* if $F(r) \le k\, r^D$ for all $r \ge 1$, with k and D two positive constants. An example of such a graph is given by the d-dimentional square lattice $\mathbb{Z}^d$ that has $(3^d, d)$-polynomial growth (with the constant 3^d certainly not optimal).

2.2. **The probability space.** The configuration space is $\Omega = \{0, 1\}^V$ equipped with the Bernoulli product measure μ of parameter p. Similarly we define Ω_Λ and μ_Λ for any subset $\Lambda \subset V$. Elements of Ω (Ω_Λ) will be denoted by Greek letters η, ω, σ (η_Λ, ω_Λ, σ_Λ) etc. Furthermore, we introduce the shorthand notation $\mu(f)$ to denote the expected value of f and $\mathrm{Var}(f)$ for its variance (when it exists).

2.3. **The Markov process.** The interacting particle model that will be studied here is a Glauber type Markov process in Ω, reversible w.r.t. the measure μ. It can be informally described as follows. Each vertex x waits an independent mean one exponential time and then, provided that the current configuration σ is such that one of the neighbors of x (i.e. one site $y \in \mathcal{N}_x$) is empty, the value $\sigma(x)$ is refreshed with a new value in $\{0,1\}$ sampled from a Bernoulli p measure and the whole procedure starts again.

The generator $\mathcal{L}$ of the process can be constructed in a standard way (see e.g. [13]). It acts on local functions as

$$\mathcal{L}f(\sigma) = \sum_{x\in V} c_x(\sigma)[q\sigma(x) + p(1-\sigma(x))][f(\sigma^x) - f(\sigma)] \tag{2.1}$$

where $c_x(\sigma) = 1$ if $\prod_{y\in\mathcal{N}_x}\sigma(y) = 0$ and $c_x(\sigma) = 0$ otherwise (namely the constraint requires at least one empty neighbor), σ^x is the configuration σ flipped at site x, $q \in [0,1]$ and $p = 1-q$. It is a non-positive self-adjoint operator on $\mathbb{L}^2(\Omega,\mu)$ with domain $Dom(\mathcal{L})$, core $\mathcal{D}(\mathcal{L}) = \{f\colon \Omega \to \mathbb{R} \text{ s.t. } \sum_{x\in V}\sup_{\sigma\in\Omega}|f(\sigma^x) - f(\sigma)| < \infty\}$ and Dirichlet form given by

$$\mathcal{D}(f) = \sum_{x\in V}\mu\left(c_x \operatorname{Var}_x(f)\right), \quad f \in Dom(\mathcal{L}).$$

Here $\operatorname{Var}_x(f) \equiv \int d\mu(\omega(x)) f^2(\omega) - \left(\int d\mu(\omega(x)) f(\omega)\right)^2$ denotes the local variance with respect to the variable $\omega(x)$ computed while the other variables are held fixed. To the generator $\mathcal{L}$ we can associate the Markov semigroup $P_t := e^{t\mathcal{L}}$ with reversible invariant measure μ. We denote by σ_t the process at time t starting from the configuration σ. Also, we denote by $\mathbb{E}_\eta(f(\eta_t))$ the expectation over the process generated by $\mathcal{L}$ at time t and started at configuration η at time zero and, with a slight abuse of notation, we let

$$\mathbb{E}_\nu(f(\sigma_t)) := \int d\nu(\eta)\mathbb{E}_\eta(f(\eta_t))$$

and let $\mathbb{P}_\nu$ be the distribution of the process started with distribution ν at time zero.

For any subset $\Lambda \subset V$ and any configuration $\eta \in \Omega$

$$\mathcal{L}^\eta_\Lambda f(\sigma) = \sum_{x\in\Lambda} c^\eta_{x,\Lambda}(\sigma)[q\sigma(x) + p(1-\sigma(x))][f(\sigma^x) - f(\sigma)] \tag{2.2}$$

where $c^\eta_{x,\Lambda}(\sigma) = c_x(\sigma_\Lambda\eta_{\Lambda^c})$ where $\sigma_\Lambda\eta_{\Lambda^c}$ is the configuration equal to σ on Λ and equal to η on Λ^c. When η is the empty configuration we write simply $c_{x,\Lambda}$ and $\mathcal{L}_\Lambda$.

2.4. **Main Result.** In order to state our main theorem, we need some notations. For any vertex $x \in V$, and any configuration $\sigma \in \Omega$, let

$$\xi^x(\sigma) = \min_{y\in V\colon \sigma(y)=0}\{d(x,y)\}$$

be the distance from x to the set of empty sites of σ.

Theorem 2.1. *Let $q > 1/2$. Assume that the graph $\mathcal{G}$ has k, D - polynomial growth and $f\colon \Omega \to \mathbb{R}$ is a local function with $\mu(f) = 0$. Let ν be a probability* *measure on Ω such that $\kappa := \sup_{x\in V} \mathbb{E}_\nu(\theta_o^{\xi x}) < \infty$ for some $\theta_o > 1$. Then, there exists a positive constant $c = c(q,k,D,\kappa,|\mathrm{supp(f)}|)$ such that*

$$|\mathbb{E}_\nu(f(\sigma_t))| \le c||f||_\infty \begin{cases} e^{-t/c} & \textit{if } D=1 \\ e^{-[t/(c\log t)]^{1/D}} & \textit{if } D>1. \end{cases} \qquad \forall t \ge 2.$$

Remark 2.2. *We expect that our results hold also for $0 < q \le \frac{1}{2}$. This needs a more precise control of the behavior of $\xi_t^x = \xi^x(\sigma_t)$ (the nearest zero to the site x). In dimension one we can obtain a better threshold for q just calculating further time derivatives of $u(t) = \mathbb{E}_\eta(\theta^{\xi_t})$, see Proposition 4.1 below.*

Remark 2.3. *Observe that if ν is a Dirac mass on some configuration η, the condition reads $\sup_{x\in V} \theta_o^{\xi^x(\eta)} < \infty$. This encodes the fact that η has infinitely many empty sites and that, in addition, the distance between two nearest empty sites is uniformly bounded. This condition is different from the case of the East model in* [4] *where the condition on the initial configuration was the presence of an infinite number of zeros.*

Remark 2.4. *If one considers the case ν product of Bernoulli-p' on $\mathcal{G}$, one has for all $\theta < 1/p'$ and all $x \in \mathcal{G}$,*

$$\mathbb{E}_\nu(\theta^{\xi^x}) = \sum_{k=0}^\infty \theta^k \mathbb{P}_\nu(\xi^x = k) \le \sum_{k=0}^\infty \theta^k (p')^{|B(x,k)|} \le \sum_{k=0}^\infty (\theta p')^k = \frac{1}{1-p'\theta}.$$

Hence, $\kappa \le \frac{1}{1-p'\theta_o}$ for $\theta_o \in (1, 1/p')$. In particular Theorem 2.1 applies to any initial probability measure, product of Bernoulli-p' on $\mathcal{G}$, with $p' \in [0,1)$.

Remark 2.5. *Note that graphs with polynomial growth are amenable. We stress anyway that there exist amenable graphs which do not enter in the framework of this paper. This is due to Proposition 3.1 below that gives a useless bound in the case of amenable graphs with intermediate growth (i.e. faster than any polynomial but slower than any exponential, see* [8]*). The same happens to any graph with exponential growth (such as for example any regular n-ary tree ($n \ge 2$)).*

3. From infinite to finite volume

This section provides a general result that will be the starting point of our analysis. The strategy developped here (and given in Section 3.1 below in a general setting) might be of indepedendent interest. The idea is first to reduce the study of the evolution of the process from infinite to a finite ball of radius proportional to t. Then to small sets on some ergodic component so that the log-Sobolev constant is much smaller than t.

The first reduction is standard and known as *the finite speed of propagation*. Namely, given a local function f with $\mathrm{supp(f)} \subset B(x,r)$ for some $x \in V$, and some integer r, we have (see e.g. [14]) for any initial measure ν on Ω

$$|\mathbb{E}_\nu(f(\sigma_t) - f(\sigma_t^\Lambda))| \le c\|f\|_\infty e^{-t}$$

where σ_t^Λ is the configuration at time t of the process starting from σ_Λ, on the finite volume $\Lambda = B(x, r+100t)$ with empty boundary condition and c is some positive constant depending on $|\mathrm{supp(f)}|$. Hence,

$$|\mathbb{E}_\nu(f(\sigma_t))| \le |\mathbb{E}_\nu(f(\sigma_t^\Lambda))| + c\|f\|_\infty e^{-t}. \tag{3.1}$$

Next we divide Λ into n connected subsets $\Lambda_1, \Lambda_2, \dots, \Lambda_n$ such that $\cup_i \Lambda_i = \Lambda$ and $\Lambda_i \cap \Lambda_j = \emptyset$ for all $i \neq j$. Such a decomposition will be called a partition of Λ.

Given such a partition of Λ, let $\mathcal{A}$ be the set of configurations containing at least two empty sites in each Λ_i. Namely,

$$\mathcal{A} = \bigcap_{i=1}^{n} \{\sigma \in \Omega_\Lambda \text{ s.t. } \sum_{x \in \Lambda_i} (1 - \sigma(x)) \geq 2\}. \tag{3.2}$$

With these notations we can now state our result.

Proposition 3.1. *Fix $\Lambda \subset V$ and f local, with* $\mathrm{supp(f)} \subset \Lambda$ *and* $\mu(f) = 0$. *Then, there exists a constant $c = c(q, |\mathrm{supp(f)}|)$ such that for any partition Λ_1, $\Lambda_2, \dots, \Lambda_n$ of Λ, for any initial probability measure ν on Ω, it holds,*

$$|\mathbb{E}_\nu(f(\sigma_t^\Lambda))| \leq c||f||_\infty \Bigg(ne^{-qm} + t \sup_{s \in [0,t]} \mathbb{P}_\nu(\eta_s^\Lambda \notin \mathcal{A}) + |\Lambda| e^{-t/3} + \exp\left\{ -\frac{t}{c} + c|\Lambda| e^{-t/(cM)} \right\} \Bigg)$$

where $m := \min\{|\Lambda_1|, \dots, |\Lambda_n|\}$ and $M := \max\{|\Lambda_1|, \dots, |\Lambda_n|\}$, provided that $ne^{-qm} < 1/2$.

In order to prove Proposition 3.1 we need first a general result on Markov processes.

3.1. **Preliminary results on Markov processes.** We here give a general result which links the behavior of a Markov process on a finite space to that of a restricted Markov process. We use it to reduce the evolution of the FA1f process to small sets on some ergodic component. In this section S is a finite space. Recall that a transition rate matrix $Q = (q(x,y))_{x,y \in S}$ is such that for any $x, y \in S$

$$q(x,y) \geq 0 \quad \text{for } x \neq y \quad \text{and } \sum_{y \in S} q(x,y) = 0.$$

and Q univocally defines a continuous time Markov chain $(X_t)_{t \geq 0}$ as follows [13]. If $X_t = x$, then the process stays at x for an exponential time with parameter $c(x) = -q(x,x)$. At the end of that time, it jumps to $y \neq x$ with probability $p(x,y) = q(x,y)/c(x)$, stays there for an exponential time with parameter $c(y)$, etc. Fix $\mathcal{A} \subset S$ and set $\hat{\mathcal{A}} = \mathcal{A} \cup \{y \notin \mathcal{A} \colon q(x,y) > 0 \text{ for some } x \in \mathcal{A}\}$. Let $(\hat{X}_t)_{t \geq 0}$ be a continuous time Markov chain with transition rate matrix $\hat{Q} = (\hat{q}(x,y))_{x,y \in \hat{\mathcal{A}}}$ such that $\forall\, x \in \mathcal{A}$ and $\forall y \in \hat{\mathcal{A}}$, $\hat{q}(x,y) = q(x,y)$. Assume that $(X_t)_{t \geq 0}$ and $(\hat{X}_t)_{t \geq 0}$ are reversible with respect to some probability measures π and $\hat{\pi}$ respectively. Then, we define the spectral gap $\hat{\gamma}$ of the hat chain as

$$\hat{\gamma} := \inf_{f \colon f \neq \text{const}} \frac{\sum_{x,y} \hat{\pi}(x) p(x,y) (f(y) - f(x))^2}{2 \operatorname{Var}_{\hat{\pi}}(f)}$$

and the log-Sobolev constant $\hat{\alpha}$ as

$$\hat{\alpha} := \sup_{f \colon f \neq \text{const}} \frac{2 \operatorname{Ent}_{\hat{\pi}}(f^2)}{\sum_{x,y} \hat{\pi}(x) p(x,y) (f(y) - f(x))^2}$$

where $\mathrm{Ent}_{\hat\pi}(f) = \hat\pi(f \log f) - \hat\pi(f)\log\hat\pi(f)$ denotes the entropy of f. See [1] for an introduction of these notions.

Proposition 3.2. *Let $(X_t)_{t\geq 0}$, $(\hat X_t)_{t\geq 0}$, π, $\hat\pi$, $\hat\gamma$ and $\hat\alpha$ as above. Then, for all initial probability measure ν on S and all $f\colon S \to \mathbb{R}$ with $\pi(f)=0$, it holds*

$$|\mathbb{E}_\nu(f(X_t))| \leq |\hat\pi(f)| + ||f||_\infty \left(4\mathbb{P}_\nu(\mathcal{A}_t^c) + \exp\left\{-\hat\gamma\frac{t}{2} + e^{-\frac{2t}{\hat\alpha}}\log\frac{1}{\hat\pi^*}\right\}\right) \tag{3.3}$$

where $\mathcal{A}_t = \{X_s \in \mathcal{A},\ \forall s \leq t\}$ and $\hat\pi^* := \min_{x\in S}\hat\pi(x)$.

Remark 3.3. *The usual argument (see* [10, 19]*) using the log-Sobolev constant would lead to*

$$|\mathbb{E}_\nu(f(X_t))| \leq ||f||_\infty \exp\left\{-\gamma\frac{t}{2} + \exp\{-\frac{2t}{\alpha}\}\log\frac{1}{\pi^*}\right\}.$$

In the setting of kinetically constrained spin models on the lattice $\mathbb{Z}^d$ (with $\pi=\mu$), it is proven that the spectral gap is constant and that the log-Sobolev constant in a finite volume Λ compares to the volume $|\Lambda|$ (see [2]*). In view of the finite speed property (Equation* (3.1)*), one has to consider a volume $\Lambda = B(x, r+100t)$ so that the log-Sobolev constant is of order $|B(x,r+100t)| \simeq t^d$. Hence the latter would lead to the following useless bound, since $\log\pi^*$ is of order $|B(x,r+100t)| \simeq t^d$,*

$$|\mathbb{E}_\nu(f(X_t))| \leq ||f||_\infty \exp\left\{-\gamma\frac{t}{2} + ct^d\right\}.$$

The main improvement here comes from the fact that we deal with a restricted (the hat) chain for which the log-Sobolev constant $\hat\alpha$ is much smaller than α, in particular much smaller than t so that the dominant term in $\exp\left\{-\hat\gamma\frac{t}{2}+\exp\{-\frac{2t}{\hat\alpha}\}\log\frac{1}{\hat\pi^*}\right\}$ is given by the gap $\hat\gamma$. The price to pay are the extra terms in (3.3) *that one has to analyze separately.*

Proof. Fix a probability measure ν and a function f with $\pi(f)=0$ and let $g = f - \hat\pi(f)$. Then

$$|\mathbb{E}_\nu(f(X_t))| \leq |\hat\pi(f)| + ||g||_\infty \mathbb{P}_\nu(\mathcal{A}_t^c) + |\mathbb{E}_\nu(g(X_t)\mathbb{1}_{\mathcal{A}_t})|. \tag{3.4}$$

We now concentrate on the last term in (3.4). By definition of the chains $(X_t)_{t\geq 0}$ and $(\hat X_t)_{t\geq 0}$ one has

$$\mathbb{E}_\nu(g(X_t)\mathbb{1}_{\mathcal{A}_t}) = \int_{\mathcal{A}} d\nu(x)\mathbb{E}_x(g(\hat X_t)\mathbb{1}_{\{\hat X_s\in\mathcal{A},\,\forall s\leq t\}}).$$

Hence, by Hölder inequality, we have

$$\begin{aligned}
|\mathbb{E}_\nu(g(X_t)\mathbb{1}_{\mathcal{A}_t})| &= |\int_{\mathcal{A}} d\nu(x)\mathbb{E}_x(g(\hat X_t)(1-\mathbb{1}_{\{\hat X_s\in\mathcal{A},\,\forall s\leq t\}^c})| \\
&\leq |\hat\pi(h\,\hat P_t g(\hat X_t))| + 2||f||_\infty\mathbb{P}_\nu(\mathcal{A}_t^c) \\
&\leq ||h||_{L^\beta(\hat\pi)}||g(\hat X_t)||_{L^{\beta'}(\hat\pi)} + 2||f||_\infty\mathbb{P}_\nu(\mathcal{A}_t^c)
\end{aligned}$$

where $h = d\nu/d\hat\pi$ and β, $\beta' \geq 1$, that will be chosen later, are such that $1/\beta + 1/\beta' = 1$. To bound the previous expression take $\beta' = 1 + e^{\frac{2t}{\hat\alpha}}$. Using the hypercontractivity property [9] (see e.g. [1, chapter 2]) and the spectral gap we obtain

$$||g(\hat{X}_t)||_{L^{\beta'}(\hat{\pi})} \leq ||g(\hat{X}_{\frac{t}{2}})||_{L^2(\hat{\pi})} \leq e^{-\hat{\gamma}\frac{t}{2}}||g||_{L^2(\hat{\pi})} \leq e^{-\hat{\gamma}\frac{t}{2}}||f||_\infty.$$

On the other hand

$$||h||_{L^\beta(\hat{\pi})} \leq \left(\int h d\hat{\pi}\right)^{\frac{1}{\beta}} ||h||_\infty^{\frac{\beta-1}{\beta}} = ||h||_\infty^{\frac{1}{\beta'}} \leq \exp\{e^{-\frac{2t}{\hat{\alpha}}}\log ||h||_\infty\}$$

and the proof is completed since $||h||_\infty \leq \log \frac{1}{\hat{\pi}^*}$. □

3.2. **Proof of Proposition 3.1.** This section is dedicated to the proof of Proposition 3.1.

Proof of Proposition 3.1. In all the proof c denotes a constant that depends on q and $|\text{supp}(\text{f})|$, and that may change from line to line.

Our aim is to apply Proposition 3.2. Let us define the setting. First $S = \Omega_\Lambda$. Define $\mathcal{A}$ as the set of configurations in Ω_Λ such that there exist at least two empty sites in each set Λ_i (see (3.2)), and $\mathcal{A}_t = \{\sigma_s^\Lambda \in \mathcal{A} \text{ for all } s \leq t\}$. Also, let $\hat{\mathcal{A}} = \{\sigma \in \Omega_\Lambda \colon \sum_{x\in\Lambda_i}(1-\sigma(x)) \geq 1, \text{ for all } i = 1,\dots,n\}$, i.e. the set of configurations that can be obtained from $\mathcal{A}$ by a legal flip for the process. Let ω be the entirely filled configuration (i.e. such that $\omega(x) = 1$ for all $x \in V$). The transition rates are, $\forall \sigma,\ \eta \in \Omega_\Lambda$,

$$q(\sigma,\eta) = \begin{cases} c_{x,\Lambda}(\sigma) & \text{if } \eta = \sigma^x \\ 0 & \text{otherwise} \end{cases}$$

and

$$\hat{q}(\sigma,\eta) = \begin{cases} c^\omega_{x,\Lambda}(\sigma) & \text{if } \sigma \in \mathcal{A} \text{ and } \eta = \sigma^x \in \mathcal{A} \text{ with } x \in \Lambda \\ 0 & \text{otherwise.} \end{cases}$$

The process $(X_t)_{t\geq 0}$ is $(\sigma_t^\Lambda)_{t\geq 0}$. The process $(\hat{X}_t)_{t\geq 0}$ is the process $(\sigma_t^\Lambda)_{t\geq 0}$ starting from $\sigma \in \mathcal{A}$ and killed on $\hat{\mathcal{A}}^c$. Then $\pi = \mu_\Lambda$ and $\hat{\pi}(\cdot) = \mu_\Lambda(\cdot \,|\, \hat{\mathcal{A}})$. Thus thanks to Proposition 3.2 we have

$$|\mathbb{E}_\nu(f(\sigma_t^\Lambda))| \leq |\hat{\pi}(f)| + ||f||_\infty \left(2\mathbb{P}_\nu(\mathcal{A}_t^c) + \exp\left\{-\hat{\gamma}\frac{t}{2} + e^{-\frac{2t}{\hat{\alpha}}}\log\frac{1}{\hat{\pi}^*}\right\}\right).$$

We now study each term of the last inequality separately.

Recalling that $\mu_\Lambda(f) = \mu(f) = 0$ and using a union bound, we have

$$|\hat{\pi}(f)| = \frac{|\mu_\Lambda(f(1-\mathbb{1}_{\hat{\mathcal{A}}^c}))|}{\mu_\Lambda(\hat{\mathcal{A}})} \leq ||f||_\infty \frac{\mu_\Lambda(\hat{\mathcal{A}}^c)}{\mu_\Lambda(\hat{\mathcal{A}})} \leq ||f||_\infty \frac{ne^{-qm}}{1-ne^{-qm}}. \tag{3.5}$$

Now we deal with the term $\mathbb{P}_\nu(\mathcal{A}_t^c)$. Let $\mathcal{I}_t$ be the event that there exists a site in Λ which has more than $2t$ rings of its Poisson clock in the time interval $[0,t]$. Then, by standard large deviation of Poisson variables and a union bound, there exists a universal positive constant d such that $\mathbb{P}_\nu(\mathcal{A}_t^c \cap \mathcal{I}_t) \leq d|\Lambda|e^{-t/3}$. Furthermore, using a union bound on all the rings on the event $\mathcal{I}_t^c$, we have

$$\mathbb{P}_\nu(\mathcal{A}_t^c \cap \mathcal{I}_t^c) \leq 2t \sup_{s\in[0,t]} \mathbb{P}_\nu(\sigma_s^\Lambda \notin \mathcal{A}).$$

We deduce that

$$\mathbb{P}_\nu(\mathcal{A}_t^c) \le c \left(t \sup_{s\in[0,t]} \mathbb{P}_\nu(\sigma_s^\Lambda \notin \mathcal{A}) + |\Lambda| e^{-t/3} \right).$$

Next we analyse the log-Sobolev constant $\hat{\alpha}$ and the spectral gap constant $\hat{\gamma}$. For that purpose, let us introduce a new process $\tilde{X}$ with transition rates :

$$\tilde{q}(\sigma,\eta) = \begin{cases} c^\omega_{x,\Lambda_i}(\sigma) & \text{if } \sigma \in \hat{\mathcal{A}} \text{ and } \eta = \sigma^x \in \hat{\mathcal{A}} \text{ with } x \in \Lambda_i \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, we have $\hat{\alpha} \le \tilde{\alpha}$ and $\hat{\gamma} \ge \tilde{\alpha}$, and $\hat{\pi}$ is a reversible measure for the tilde process. Observe that $\hat{\mathcal{A}} = \bigcap_{i=1}^n \hat{\Omega}_i$ where

$$\hat{\Omega}_i = \{\sigma \in \Omega_\Lambda \text{ s.t. } \exists x_i \in \Lambda_i \text{ with } \sigma(x_i) = 0\}.$$

If we denote by $\hat{\mu}_i(\cdot) = \mu_{\Lambda_i}(\cdot \,|\, \hat{\Omega}_i)$, the product structure of μ and $\mathcal{A}$ transfers to $\hat{\pi}$ so that $\hat{\pi} = \otimes_i \hat{\mu}_i$. Furthermore, the tilde process restricted to each $\hat{\Omega}_i$ is ergodic and reversible with respect to $\hat{\mu}_i$. Hence we can define the associated spectral gap $\tilde{\gamma}_i$ and log-Sobolev constant $\tilde{\alpha}_i$. By the well-known tensorisation property of the Poincaré and the log-Sobolev inequalities (see e.g. [1, Chapter 1]), we conclude that $\tilde{\gamma} = \min(\tilde{\gamma}_1, \dots, \tilde{\gamma}_n)$ and $\tilde{\alpha} = \max(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n)$. Then, Proposition 3.4 below shows that $\tilde{\gamma} \ge c$ and $\tilde{\alpha}_i \leqslant c|\Lambda_i|$. Hence,

$$\exp\left\{ -\hat{\gamma}\frac{t}{2} + \exp\{-\frac{2t}{\hat{\alpha}}\} \log \frac{1}{\hat{\pi}^*} \right\} \leqslant \exp\left\{ -\frac{t}{c} + c|\Lambda| e^{-t/(cM)} \right\}.$$

This ends the proof. □

Proposition 3.4 ([3]). *Let $A \subset V$ be connected, $\hat{\Omega}_A = \{\sigma \in \Omega\colon \sum_{x\in A}(1 - \sigma(x))) \ge 1\}$ and $\hat{\mu}_A(\cdot) = \mu_A(\cdot \,|\, \hat{\Omega}_A)$. Let ω be the entirely filled configuration (i.e. such that $\omega(x) = 1$ for all $x \in V$) and for $x \in A$, define $\hat{c}_x(\sigma) := c^\omega_{x,A}(\sigma) \mathbb{1}_{\sigma^x \in \hat{\Omega}_A}$ on $\hat{\Omega}_A$. Then, there exists a constant $c = c(q)$ such that*

$$\hat{\gamma}_A := \inf_{f: f \ne const.} \frac{\sum_{x\in A} \hat{\mu}_A(\hat{c}_x \operatorname{Var}_x(f))}{\operatorname{Var}_{\hat{\mu}_A}(f)} \ge c$$

and

$$\hat{\alpha}_A := \sup_{f: f \ne const.} \frac{\operatorname{Ent}_{\hat{\mu}_A}(f)}{\sum_{x\in A} \hat{\mu}_A(\hat{c}_x \operatorname{Var}_x(f))} \le c|A|.$$

Proof. The first part on the spectral gap is proven in [3, Theorem 6.4 page 336]. In section 6 we give an alternative proof which gives a better bound for small q and can be extended to non cooperative models different from FA1f.

The second part easily follows from the standard bound [5, 17]

$$\hat{\alpha}_A \le \hat{\gamma}_A^{-1} \log \frac{1}{\hat{\mu}_A^*}$$

where $\hat{\mu}_A^* := \min_{\sigma \in \hat{\Omega}_A} \hat{\mu}_A(\sigma) \ge \exp\{-c|A|\}$. □

4. Persistence of zeros out of equilibrium

In this section we study the behavior of the minimal distance from a fixed site at which one finds a vacancy. The result that we obtain will be used in Section 5 for the proof of our main Theorem 2.1. Indeed Proposition 4.1 will be a key ingredient in order to estimate the second term in the inequality of Proposition 3.1, namely the probability the process gets out of the component $\mathcal{A}$ of the configuration set which requires two empty sites on each small volume Λ_i. For any $\sigma \in \{0,1\}^V$ and any $x \in V$ define $\xi^x(\sigma)$ as the minimal distance at which one finds an empty site starting from x,

$$\xi^x(\sigma) = \min_{y\in V\colon \sigma(y)=0} \{d(x,y)\}$$

with the convention that $\min \emptyset = +\infty$, ($\xi^x(\sigma) = 0$ if $\sigma(x) = 0$).

Proposition 4.1. *Consider the FA1f process on a finite set $\Lambda \subset V$ with generator $\mathcal{L}_\Lambda$. Then, for all $x \in \Lambda$, all $\theta \geq 1$, all $q \in (\frac{\theta}{\theta+1}, 1]$ and all initial configuration η, it holds*

$$\mathbb{E}_\eta\left(\theta^{\xi^x(\sigma_t^\Lambda)}\right) \leq \theta^{\xi^x(\eta)}\, e^{-\lambda t} + \frac{q}{q(\theta+1)-\theta} \qquad \forall t \geq 0,$$

where $\lambda = \frac{\theta^2-1}{\theta}(q - \frac{\theta}{\theta+1})$.

Proof. Fix $\theta > 1$, $q > 0$ and $x \in \Lambda$. To simplify the notation we drop the superscript x from ξ^x and set $\xi_t = \xi(\sigma_t^\Lambda)$ in what follows. Recall that σ_t^Λ is defined with empty boundary condition so that $\xi_t \leq d(x, \Lambda^c)$. Let $u(t) = \mathbb{E}\_\eta(\theta^{\xi_t})$ and observe that

$$\frac{d}{dt}u(t) = \mathbb{E}_\eta(\mathcal{L}_\Lambda \theta^{\xi_t}).$$

To calculate the expected value above we distinguish two cases: **(i)** $\xi_t = 0$, **(ii)** $\xi_t \geq 1$.

Case **(i):** assume that $\xi_t = 0$. Then

$$(\mathcal{L}_\Lambda \theta^{\xi_t})\mathbb{1}_{\xi_t=0} = \theta^{\xi_t} c_x(\sigma_t^\Lambda) p(\theta-1)\mathbb{1}_{\xi_t=0}. \tag{4.1}$$

Case **(ii)**. Define $E(\sigma) = \{y \in V\colon d(x,y) = \xi(\sigma) \text{ and } \sigma(y) = 0\}$ and $F(\sigma) = \{y \in V\colon d(y,E) = 1 \text{ and } d(x,y) = \xi(\sigma) - 1\}$. Then one argues that ξ_t can increase by 1 only if there is exactly one empty site in the set E, and that it can always decrease by 1 by a flip (which is legal by construction) on each site of F (see Figure 4).

Hence

$$\begin{aligned}(\mathcal{L}_\Lambda \theta^{\xi_t})\mathbb{1}_{\xi_t\geq 1} &= \theta^{\xi_t}\left[p(\theta-1)\sum_{y\in E} c_y(\sigma_t^\Lambda)\mathbb{1}_{|E|=1} + q|F|(\frac{1}{\theta}-1)\right]\mathbb{1}_{\xi_t\geq 1} \\ &\leq \theta^{\xi_t}[p(\theta-1) - q\frac{\theta-1}{\theta}] + [q\frac{\theta-1}{\theta} - p(\theta-1)]\mathbb{1}_{\xi_t=0} \end{aligned} \tag{4.2}$$

Summing up (4.1) and (4.2) we end up with

$$\mathcal{L}_\Lambda \theta^{\xi_t} \leqslant \frac{\theta-1}{\theta}\left(\theta^{\xi_t}(p\theta - q) + q\right).$$

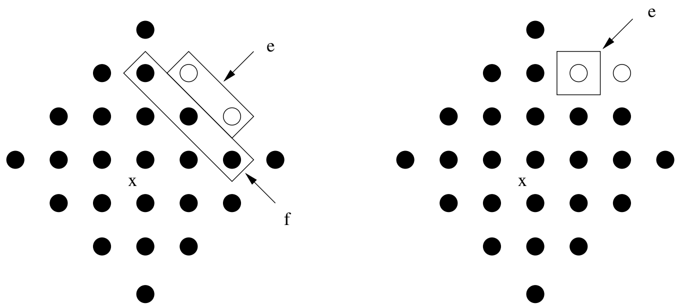

FIGURE 1. On the graph $\mathcal{G} = \mathbb{Z}^2$, two examples of configurations for which $\xi^x = 3$. On the left ξ^x cannot increase since $|E| \geq 2$, it can decrease by a flip (legal thanks to the empty sites in E) in any points of F. On the right ξ^x can either increase or decrease.

Therefore, since $p = 1 - q$,

$$u'(t) \leq \frac{\theta - 1}{\theta}\left((p\theta - q)u(t) + q\right) = -\lambda u(t) + q\frac{\theta - 1}{\theta}$$

and the expected result follows. □

5. Proof of the Main Theorem

In this section we prove Theorem 2.1

Proof of Theorem 2.1. In all the proof c denotes some positive constant depending on all the parameters of the system and that may change from line to line.

Fix $t \geq 2$ and a local function f. Thanks to (3.1) we deal with the process in finite volume $\Lambda = B(x, r + 100t)$ where $r \in \mathbb{N}$ and $x \in V$ are such that $\mathrm{supp}(f) \subset B(x, r)$.

Our aim is to apply Proposition 3.1. Observe first that for any positive integer $\ell \leqslant t$, there exists[1] a partition of (connected) sets $\Lambda_1, \dots, \Lambda_n$ of Λ, and vertices $x_1, \dots, x_n \in V$, such that for any i, $B(x_i, \ell) \subset \Lambda_i \subset B(x_i, 3\ell)$.

Then, take $\ell = \epsilon[t/\log t]^{1/D}$ if $D > 1$ and $\ell = \epsilon t$ if $D = 1$ for some $\epsilon > 0$ that will be chosen later and observe that, with this choice,

$$M = \max(|\Lambda_1|, \dots, |\Lambda_n|) \leq k3^D\ell^D$$

(since $\mathcal{G}$ has polynomial growth). Furhermore

$$m = \min(|\Lambda_1|, \dots, |\Lambda_n|) \geq \ell$$

[1]One can construct $\Lambda_1, \dots, \Lambda_n$, $x_1, \dots, x_n$ as follows. Fix a site $x_o \in \Lambda$ such that $B(x_o, \ell) \subset \Lambda$. Then order (arbitrarily) the sites $y_1, y_2, \dots, y_N$ of $\{x \in \Lambda \colon B(x, \ell) \subset \Lambda \text{ and } d(x, x_o) = 2i(\ell + 1) - 1 \text{ for some } i \geq 1\}$ and perform the following algorrithm: set $x_1 = x_o$, $i_0 = 0$, and for $k \geq 1$ set $x_{k+1} = y_{i_k}$ with $i_k := \inf\{j \geq i_{k-1} + 1 : B(y_j, \ell) \cap (\cup_{i=1}^k B(x_i, \ell)) = \emptyset\}$. Such a procedure gives the existence of n sites $x_1, \dots, x_n$ such that $B(x_i, \ell) \cap B(x_j, \ell) = \emptyset$, for all $i \neq j$, $B(x_i, \ell) \subset \Lambda$ for all i and any site $y_k \notin A := \cup_{i=1}^n B(x_i, \ell)$ is at distance at most $2\ell - 1$ from A. Now attach each connected component C of A^c to any (arbitrarily chosen) nearest ball $B(x_i, \ell)$, $i \in \{1, \dots, n\}$, with which C is connected, to obtain all the Λ_i with the desired properties.

Since $n \leq |\Lambda| \leq ct^D$, Equation (3.1) and Proposition 3.1 guarantee that

$$|\mathbb{E}_\nu(f(\sigma_t))| \leq c||f||_\infty t \sup_{s\in[0,t]} \mathbb{P}_\nu(\sigma_s^\Lambda \notin \mathcal{A}) + c||f||_\infty \begin{cases} e^{-t/c} & \text{if } D=1 \\ e^{-[t/(c\log t)]^{1/D}} & \text{if } D>1 \end{cases}$$

provided ϵ is small enough.

It remains to study the first term of the latter inequality. We partition each set Λ_i into two connected sets Λ_i^+ and Λ_i^- (i.e. $\Lambda_i = \Lambda_i^+ \cup \Lambda_i^-$ and $\Lambda_i^+ \cap \Lambda_i^- = \emptyset$) such that for some $x_i^+, x_i^- \in V$, $B(x_i^\pm, \ell/4) \subset \Lambda_i^\pm$ (the existence of such vertices are left to the reader). The event $\{\sigma_s^\Lambda \notin \mathcal{A}\}$ implies that there exists one index i such that at least one of the two halves Λ_i^+, Λ_i^- is completely filled. Assume that it is for example Λ_i^+, i.e. assume that for any $x \in \Lambda_i^+$, $\sigma_s^\Lambda(x) = 1$. This implies that $\xi^{x_i^+}(\sigma_s^\Lambda) \geq \ell/4$. Hence, thanks to a union bound, Markov's inequality, and Proposition 4.1, there exists $\theta > 1$ such that

$$\begin{aligned}
\mathbb{P}_\nu(\sigma_s^L \notin \mathcal{A}) &\leq 2n\mathbb{P}_\nu(\xi^{x_i^+}(\sigma_s^\Lambda) \geq \ell/4) \\
&\leq 2n\theta^{-\ell/4}\mathbb{E}_\nu(\theta^{\xi^{x_i^+}(\sigma_s^\Lambda)}) \\
&\leq cn\theta^{-\ell/4} \\
&\leq c \begin{cases} e^{-t/c} & \text{if } D=1 \\ e^{-[t/(c\log t)]^{1/D}} & \text{if } D>1 \end{cases}
\end{aligned}$$

where we used the definition of ℓ and that $n \leq |\Lambda| \leq ct^D$. This ends the proof.

□

6. Spectral gap on the ergodic component

In this section we estimate the spectral gap of the process FA1f on the ergodic component on $\mathcal{G} = (V, E)$. This has been done in [2, 3]. However, we present here an alternative proof, based on the ideas of [15], that, on the one hand, gives a somehow more precise bound for very small q and, on the other hand, can be generalized to *non cooperative* models different from FA1f on some ergodic component (not necessarily the largest one). An example of non cooperative model different from FA1f is the following. Each vertex x waits an independent mean one exponential time and then, provided that the current configuration σ is such that at least two of the sites at distance less or equal to 2 are empty ($\sum_{y\in\hat{\mathcal{N}}_x}(1-\sigma(y)) \geq 2$, where $\hat{\mathcal{N}}_x = \{y\colon d(x,y) \leq 2\}$), the value $\sigma(x)$ is refreshed with a new value in $\{0,1\}$ sampled from a Bernoulli p measure and the whole procedure starts again. For simplicity we deal with the FA-1f model.

For every $\Lambda \subset V$ finite, define

$$\hat{\Omega}_\Lambda = \{\sigma \in \Omega\colon \sum_{x\in\Lambda}(1-\sigma(x)) \geq 1\} \tag{6.1}$$

and $\hat{\mu}_\Lambda(\cdot) = \mu_\Lambda(\cdot \,|\, \hat{\Omega}_\Lambda)$. Let $\hat{c}_x(\sigma) = c^\omega_{x,\Lambda}(\sigma)\mathbb{1}_{\sigma^x\in\hat{\Omega}_\Lambda}$ for all $\sigma \in \hat{\Omega}_\Lambda$, where ω is the entirely filled configuration, i.e. $\omega(x) = 1$ for all $x \in V$. The spectral gap for the dynamics on $\hat{\Omega}_\Lambda$ is defined as

$$\hat{\text{gap}}(\Lambda) := \inf_{f\colon f \neq const.} \frac{\sum_{x\in\Lambda} \hat{\mu}_\Lambda(\hat{c}_x \operatorname{Var}_x(f))}{\operatorname{Var}_{\hat{\mu}_\Lambda}(f)} \tag{6.2}$$

where the infimum runs over all non constant functions $f\colon \hat{\Omega}_\Lambda \to \mathbb{R}$, and for simplicity we set $\operatorname{Var}_x(f) := \operatorname{Var}_{\mu_{\{x\}}}(f)$.

We now state the result on the spectral gap.

Theorem 6.1. *Let $\mathcal{G} = (V, E)$ be a graph with (k, D)-polynomial growth. Then there exists a positive constant $C = C(k, D)$ such that for any connected set $\Lambda \subset V$*

$$\hat{\text{gap}}(\Lambda) \geq C \, \frac{q^{D+4}}{\log(2/q)^{D+1}}$$

The proof of Theorem 6.1 is divided in two steps. At first we bound from below the spectral gap of the hat chain in Λ by the spectral gap of the FA1f model (not restricted to the ergodic component), on all subsets of V with minimal boundary condition. Then we study such a spectral gap following the strategy of [15].

We need some more notations. Given $A \subset V$, $z \in \partial A$ and $x \in A$ define $c^z_{x,A}(\sigma) = c^{\omega^{(z)}}_{x,A}(\sigma)$, $\sigma \in \Omega$, where $\omega^{(z)}$ is the entirely filled configuration, except at site z where it is 0: $\omega^{(z)}(x) = 1$ for all $x \neq z$ and $\omega^{(z)}(z) = 0$. The corresponding generator $\mathcal{L}^{\omega^{(z)}}_A$ will be simply denoted by $\mathcal{L}^z_A$. It corresponds to the FA1f process in A with minimal boundary condition.

The first step in the proof of Theorem 6.1 is the following result.

Proposition 6.2. *For every finite connected subsets Λ of V such that $8p^{\operatorname{diam}(\Lambda)/3} < \frac{1}{2}$ it holds*

$$\hat{\text{gap}}(\Lambda) \geq \frac{1}{48} \inf_{\substack{A \subset V, \text{connected} \\ z \in \partial A}} \operatorname{gap}(\mathcal{L}^z_A).$$

Observe that, combining [3, Theorem 6.1] and [2, Theorem 6.1] for any set A and any site z, we had $\operatorname{gap}(\mathcal{L}^z_A) \geq c q^{\log_2(1/q)}$ for some universal positive constant c. Hence, for the FA1f process, we had the lower bound

$$\hat{\text{gap}}(\Lambda) \geq c q^{\log_2(1/q)}.$$

We present below an alternative strategy (based on [15]) which can be applied to other non-cooperative models and gives a more accurate bound for the FA1f process when q is small.

Proof. Consider a non constant function $f\colon \hat{\Omega}_\Lambda \to \mathbb{R}$ and define $\tilde{f}\colon \Omega_\Lambda \to \mathbb{R}$ as

$$\tilde{f}(\sigma) = \begin{cases} f(\sigma) & \text{if } \sigma \in \hat{\Omega}_\Lambda \\ 0 & \text{otherwise} \end{cases}$$

We divide[2] Λ into two disjoint connected subsets A and B such that their diameter is larger then $|\Lambda|/3$.

[2]To construct A and B take two points x, y such that $d(x, y) = \ell := \operatorname{diam}(\Lambda)$ and define $A_0 = \{z \in \Lambda\colon d(x, z) \leq \ell/3\}$ and $B_0 = \{z \in \Lambda\colon d(y, z) \leq \ell/3\}$. Attach to A_0 all the connected components of $\Lambda \setminus (A_0 \cup B_0)$ connected to A_0 to obtain A, then attach all the remaining connected components of $\Lambda \setminus (A_0 \cup B_0)$ to B_0 to obtain B.

Thank to Lemma 6.5 below (our hypothesis implies that $\max(1-\mu(c_A), 1-\mu(c_B)) < 1/16$)

$$\mathrm{Var}_{\hat\mu_\Lambda}(f) \le 24\,\hat\mu_\Lambda[c_B \mathrm{Var}_{\mu_A}(\tilde f) + c_A \mathrm{Var}_{\mu_B}(\tilde f)]$$

where $c_A = \mathbb{1}_{\hat\Omega_A}$ and $c_B = \mathbb{1}_{\hat\Omega_B}$ and $\hat\Omega_A$ and $\hat\Omega_B$ are defined in (6.1).

Consider the first term. Define the random variable

$$\zeta := \sup_{x\in B}\{d(A,x)\colon \sigma(x)=0\}$$

where by convention the supremum of the empty set is ∞. The function c_B guarantees that $\zeta \in \{1, 2, \cdots, \mathrm{diam}(\Lambda)\}$. Following the strategy of [2] we have

$$\begin{aligned}
\hat\mu_\Lambda[c_B \mathrm{Var}_{\mu_A}(\tilde f)] &= \frac{1}{\mu_\Lambda(\hat\Omega_\Lambda)}\sum_{n\ge1}\mu_\Lambda[\mathbb{1}_\zeta \mathrm{Var}_{\mu_A}(\tilde f)]\\
&\le \frac{1}{\mu_\Lambda(\hat\Omega_\Lambda)}\sum_{n\ge1}\mu_\Lambda[\mathbb{1}_\zeta \mathrm{Var}_{\mu_{A_n}}(\tilde f)]
\end{aligned}$$

where $A_n = \{x\in\Lambda\colon d(A,x)\le n-1\}$ and we used the convexity of the variance (which is valid since the event $\{\zeta\}$ does not depend, by construction, on the value of the configuration σ_{A_n} inside A_n). The indicator function above $\mathbb{1}_\zeta$ guarantees the presence of a zero on the boundary ∂A_n of the set A_n. Order (arbitrarily) the points of ∂A_n and call Z the (random) position of the first empty site on ∂A_n. Then, for all $n\ge 1$,

$$\begin{aligned}
\mu_\Lambda[\mathbb{1}_\zeta \mathrm{Var}_{\mu_{A_n}}(\tilde f)] &= \sum_{z\in\partial A_n}\mu_\Lambda[\mathbb{1}_\zeta\mathbb{1}_{Z=z}\mathrm{Var}_{\mu_{A_n}}(\tilde f)]\\
&\le \sum_{z\in\partial A_n}\mathrm{gap}(\mathcal{L}^z_{A_n})^{-1}\sum_{y\in A_n}\mu_\Lambda[\mathbb{1}_\zeta\mathbb{1}_{Z=z}\mu_{A_n}(c^z_{y,A}\mathrm{Var}_y(\tilde f))]\\
&\le \gamma\sum_{z\in\partial A_n}\sum_{y\in A_n}\mu_\Lambda[\mathbb{1}_\zeta\mathbb{1}_{Z=z}c^z_{y,A}\mathrm{Var}_y(\tilde f)]
\end{aligned}$$

where we used the fact that the events $\{\zeta\}$ and $\{Z=z\}$ depend only on $\sigma^c_{A_n}$, and where $\gamma := \sup\mathrm{gap}(\mathcal{L}^z_A)^{-1}$, the supremum running over all connected subset A of V and all $z\in\partial A$. Now observe that $\mathbb{1}_\zeta\mathbb{1}_{Z=z}c^z_{y,A}\le\mathbb{1}_\zeta\mathbb{1}_{Z=z}\hat c_y$ for any $y\in A_n$. Hence,

$$\begin{aligned}
\hat\mu_\Lambda[c_B \mathrm{Var}_{\mu_A}(\tilde f)] &\le \frac{\gamma}{\mu_\Lambda(\hat\Omega_\Lambda)}\sum_{n\ge1}\sum_{z\in\partial A_n}\sum_{y\in A_n}\mu_\Lambda[\mathbb{1}_\zeta\mathbb{1}_{Z=z}\hat c_y\mathrm{Var}_y(\tilde f)]\\
&\le \frac{\gamma}{\mu_\Lambda(\hat\Omega_\Lambda)}\sum_{y\in\Lambda}\sum_{n\ge1}\sum_{z\in\partial A_n}\mu_\Lambda[\mathbb{1}_\zeta\mathbb{1}_{Z=z}\hat c_y\mathrm{Var}_y(\tilde f)]\\
&= \gamma\sum_{y\in\Lambda}\hat\mu_\Lambda[\hat c_y\mathrm{Var}_y(\tilde f)] = \gamma\sum_{y\in\Lambda}\hat\mu_\Lambda[\hat c_y\mathrm{Var}_y(f)].
\end{aligned}$$

The same holds for $\hat\mu_\Lambda[c_A \mathrm{Var}_{\mu_B}(\tilde f)]$, leading to the expected result. $\square$

The second step in the proof of Theorem 6.1 is a careful analysis of $\mathrm{gap}(\mathcal{L}^z_A)$ for any given connected set $A\subset V$ and $z\in\partial A$.

Proposition 6.3. *Let $\mathcal{G} = (V, E)$ be a graph with (k, D)-polynomial growth. Then, there exists a universal constant $C = C(k, D)$ such that for any connected set $A \subset V$, and any $z \in \partial A$, it holds*

$$\text{gap}(\mathcal{L}_A^z) \geq C \frac{q^{D+4}}{\log(2/q)^{D+1}}.$$

We postpone the proof of Proposition 6.3 to end the proof of Theorem 6.1.

Proof of Theorem 6.1. The result follows at once combining Proposition 6.2 and Proposition 6.3. □

In order to prove Proposition 6.3, we need a preliminary result on the spectral gap of some auxiliary chain, and to order the points of A in a proper way, depending on z. Let $N := \max_{x \in A} d(x, z)$, for any $i = 1, 2, \ldots, N$, we define

$$A_i := \{x \in A\colon d(x, z) = i\} = \{x_1^{(i)}, \ldots, x_{n_i}^{(i)}\}$$

where $x_1^{(i)}, \ldots, x_{n_i}^{(i)}$ is any chosen order. Then we say that for any $x, y \in A$, $x \leq y$ if either $d(x, z) > d(y, z)$ or $d(x, z) = d(y, z)$ and x comes before y in the above ordering. Then, we set $A_x = \{y \in A\colon y \geq x\}$ and $\tilde{A}_x = A_x \setminus \{x\}$.

Lemma 6.4. *Fix a connected set $A \subset V$, and $z \in \partial A$. For any $x \in A$ and $\sigma \in \Omega$, let $E_x \subset \Omega_{\tilde{A}_x}$, $\Delta_x = \text{supp}(E_x)$ and $\tilde{c}_x(\sigma) = \mathbb{1}_{E_x}(\sigma_{\tilde{A}_x})$. Assume that*

$$\sup_{x \in A} \mu(1 - \tilde{c}_x) \sup_{x \in A} |\{y \in A\colon \Delta_y \cup \{y\} \ni x\}| < \frac{1}{4}.$$

Then, for any $f : \Omega_A \to \mathbb{R}$ it holds

$$\text{Var}_{\mu_A}(f) \leq 4 \sum_{x \in A} \mu_A(\tilde{c}_x \operatorname{Var}_x(f)).$$

Proof. We follow [15]. In all the proof, to simplify the notations, we set $\text{Var}_B = \text{Var}_{\mu_B}$, for any B. First, we claim that

$$\text{Var}_A(f) = \sum_{x \in A} \mu_A(\text{Var}_{A_x}(\mu_{\tilde{A}_x}(f))). \tag{6.3}$$

Take $x = x_{n_N}^{(N)}$, by factorization of the variance, we have

$$\text{Var}_A(f) = \mu_A(\text{Var}_{\tilde{A}_x}(f)) + \text{Var}_A(\mu_{\tilde{A}_x}(f)).$$

The claim then follows by iterating this procedure, removing one site at a time, in the order defined above.

We analyze one term in the sum of (6.3) and assume, without loss of generality, that $\mu_{A_x}(f) = 0$. We write $\mu_{\tilde{A}_x}(f) = \mu_{\tilde{A}_x}(\tilde{c}_x f) + \mu_{\tilde{A}_x}((1 - \tilde{c}_x) f)$ so that

$$\mu_A[\text{Var}_{A_x}(\mu_{\tilde{A}_x}(f))] \leq 2\mu_A[\text{Var}_{A_x}(\mu_{\tilde{A}_x}(\tilde{c}_x f))] + 2\mu_A[\text{Var}_{A_x}(\mu_{\tilde{A}_x}((1 - \tilde{c}_x) f))]. \tag{6.4}$$

Observe that, by convexity of the variance and since $\tilde{c}_x$ does not depend on x, the first term of the latter can be bounded as

$$\mu_A[\text{Var}_{A_x}(\mu_{\tilde{A}_x}(\tilde{c}_x f))] = \mu_A[\text{Var}_x(\mu_{\tilde{A}_x}(\tilde{c}_x f))] \leq \mu_A[\tilde{c}_x \operatorname{Var}_x(f)].$$

Now we focus on the second term of (6.4). Note that $\mu_{\tilde{A}_x}[(1-\tilde{c}_x)f)] = \mu_{\tilde{A}_x}[(1-\tilde{c}_x)\mu_{\tilde{A}_x\setminus\Delta_x}(f))]$. Set $\delta := \sup_{x\in A}\mu(1-\tilde{c}_x)$. Hence, bounding the variance by the second moment and using Cauchy-Schwarz inequality, we get

$$\begin{aligned}\operatorname{Var}_{A_x}(\mu_{\tilde{A}_x}((1-\tilde{c}_x)f)) &\le \operatorname{Var}_{A_x}\left(\mu_{\tilde{A}_x}[(1-\tilde{c}_x)\mu_{\tilde{A}_x\setminus\Delta_x}(f)]\right)\\ &\le \mu_{A_x}\left(\mu_{\tilde{A}_x}[(1-\tilde{c}_x)\mu_{\tilde{A}_x\setminus\Delta_x}(f)]^2\right)\\ &\le \delta\left(\operatorname{Var}_{A_x}(\mu_{\tilde{A}_x\setminus\Delta_x}(f))\right)\end{aligned}$$

From all the previous computations (and using (6.3)) we deduce that

$$\operatorname{Var}_A(f) \le 2\sum_{x\in A}\mu_A(\tilde{c}_x\operatorname{Var}_x(f)) + 2\delta\sum_{x\in A}\mu_A\left(\operatorname{Var}_{A_x}(\mu_{\tilde{A}_x\setminus\Delta_x}(f))\right).$$

Hence if one proves that

$$\sum_{x\in A}\mu_A\left(\operatorname{Var}_{A_x}(\mu_{\tilde{A}_x\setminus\Delta_x}(f))\right) \le \sup_{y\in A}|\{x\in A\colon \Delta_x\cup\{x\}\ni y\}|\operatorname{Var}_A(f) \quad (6.5)$$

the result follows. We now prove (6.5). Using (6.3), we have

$$\operatorname{Var}_{A_x}(g) = \sum_{y\in A_x}\mu_{A_x}\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(g))\right) = \sum_{y\in\Delta_x\cup\{x\}}\mu_{A_x}\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(g))\right)$$

where $g = \mu_{\tilde{A}_x\setminus\Delta_x}(f)$ and we used that $\operatorname{supp}(g)\subset\Delta_x$. It follows that

$$\mu_A\left(\operatorname{Var}_{A_x}(g)\right) \sum_{y\in\Delta_x\cup\{x\}}\mu_A\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(g))\right) \le \sum_{y\in\Delta_x\cup\{x\}}\mu_A\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(f))\right)$$

since, by Cauchy-Schwarz,

$$\begin{aligned}\mu_A\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(g))\right) &= \mu_A\left(\left[\mu_{\tilde{A}_x\setminus\Delta_x}\left(\mu_{\tilde{A}_y}(f)-\mu_{A_y}(f)\right)\right]^2\right)\\ &\le \mu_A\left(\operatorname{Var}_{A_y}(\mu_{\tilde{A}_y}(f))\right).\end{aligned}$$

This ends the proof. □

Proof of Proposition 6.3. Our aim is to apply Lemma 6.4. Let us define the events E_x, for $x\in A$. Fix an integer ℓ that will be chosen later and set $n = \ell\wedge d(x,z)$. Let $(x_1,x_2,\dots,x_n)$ be an arbitrarily chosen ordered collection satisfying $d(x_i,x_{i+1}) = 1$, $d(x_i,x) = i$ and $d(x_i,z) = d(x,z)-i$ for $i = 0,\dots,n$, with the convention that $x_0 = x$, and set $E_x = \{\sigma\in\Omega\colon \sum_{i=1}^n(1-\sigma(x_i))\ge 1\}$, i.e. E_x is the event that at least one of the site of $\Delta_x = \{x_1,x_2,\dots,x_n\}$ is empty. Note that by construction $\Delta_x\subset A\cup\{z\}$ and is connected. Moreover for any x such that $d(x,z)\le\ell$, $E_x=\Omega$ so that $\tilde{c}_x\equiv 1$. Since $|\Delta_x|\leqslant k\ell^D$ for any $x\in A$, the assumption of Lemma 6.4 reads

$$p^\ell(1+k\ell^D) < 1/4$$

which is satisfied if one chooses $\ell = \frac{c}{q}\log\frac{2}{q}$ with $c = c(k,D)$ large enough. Hence for any $f:\Omega_A\to\mathbb{R}$ it holds

$$\operatorname{Var}_{\mu_A}(f) \le 4\sum_{x\in A}\mu_A(\tilde{c}_x\operatorname{Var}_x(f)).$$

and we are left with the analysis of each term $\mu_A(\tilde{c}_x \operatorname{Var}_x(f))$ for which we use a path argument. Fix $x \in A$ and the collection $(x_1, x_2, \dots, x_n)$ introduced above. Given a configuration σ such that $\tilde{c}_x(\sigma) = 1$, denote by ξ the (random) distance between x and the first empty site in the collection $(x_1, x_2, \dots, x_n)$: i.e. $\xi(\sigma) = \inf\{i \colon \sigma(x_i) = 0\}$. Then we write

$$\mu_A(\tilde{c}_x \operatorname{Var}_x(f)) = \sum_{i=1}^{n} \mu_A(\tilde{c}_x \mathbb{1}_{\xi=i} \operatorname{Var}_x(f))$$
$$pq \sum_{i=1}^{n} \sum_{\sigma:\xi(\sigma)=i} \mu_A(\sigma)(f(\sigma^x) - f(\sigma))^2$$

where the sum is understood to run over all σ such that $\tilde{c}_x(\sigma) = 1$ (and $\xi(\sigma) = i$).

Fix $i \in \{1, \dots, n\}$. For any $\sigma \in \Omega$ such that $\xi(\sigma) = i$, we construct a path of configurations $\gamma_x(\sigma) = (\sigma_0 = \sigma, \sigma_1, \sigma_2, \dots, \sigma_{4i-5} = \sigma^x)$ from σ to σ^x, of length $4i-5 \le 4\ell$. The idea behind the construction is to bring an empty site from x_i, step by step, toward x_1, make the flip in x and going back, keeping track of the initial configuration σ. For any j, σ_{j+1} can be obtained from σ_j by a legal flip for the FA1f process. Furthermore σ_j differs from σ on at most three sites (possibly counting x). More precisely, define $T_k(\sigma) := \sigma^{x_k}$ for any k and σ, and

$$\sigma_j = \begin{cases} T_{i-k-1}(\sigma) & \text{if } j = 2k+1, \text{ and } k = 0, 1, \dots, i-2 \\ T_{i-k} \circ T_{i-k-1}(\sigma) & \text{if } j = 2k, \text{ and } k = 1, \dots, i-2 \\ T_1(\sigma^x) & \text{if } j = 2i-2 \\ T_{k-i+2} \circ T_{k-i+3}(\sigma^x) & \text{if } j = 2k+1, \text{ and } k = i-1, \dots, 2i-4 \\ T_{k-i+2}(\sigma^x) & \text{if } j = 2k, \text{ and } k = i, \dots, 2i-3. \end{cases}$$

See Figure 6 for a graphical illustration of such a path.

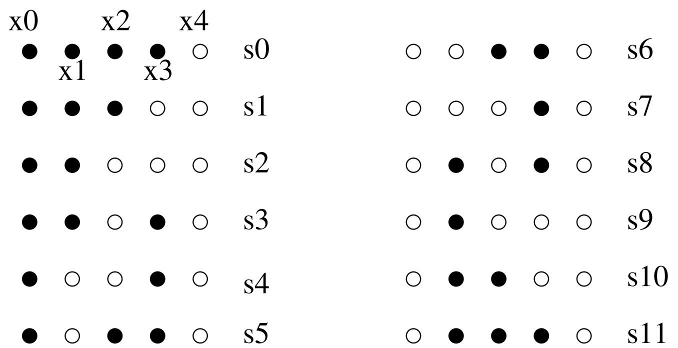

FIGURE 2. Illustration of the path from σ to σ^x for a configuration σ satisfying $\xi(\sigma) = x_4$. Here $i = 4$ and the lenght of the path is $4i - 5 = 11$.

Denote by $\Gamma_x(\sigma) = \{\sigma_0, \sigma_1, \dots, \sigma_{4i-6}\}$ (i.e. the configurations of the path $\gamma_x(\sigma)$ except the last one σ^x). For any $\eta = \sigma_j \in \Gamma_x(\sigma)$, $j \ge 1$, let

$y = y(x,\eta) \in \{x, x_1, x_2, \dots, x_\ell\}$ be such that $\eta = \sigma^y_{j-1}$. Then, by Cauchy-Schwarz inequality,

$$(f(\sigma^x) - f(\sigma))^2 = \left(\sum_{\eta \in \Gamma_x(\sigma)} (f(\eta^y) - f(\eta)) \right)^2 \le 4\ell \sum_{\eta \in \Gamma_x(\sigma)} (f(\eta^y) - f(\eta))^2$$
$$\le \frac{4\ell}{pq} \sum_{\eta \in \Gamma_x(\sigma)} c_y(\eta) \operatorname{Var}_y(f)(\eta).$$

Hence,

$$\mu_A(\tilde c_x \operatorname{Var}_x(f)) \le 4\ell K \sum_\eta \mu_A(\eta) c_y(\eta) \operatorname{Var}_y(f)$$

where

$$K = \sup_{\eta \in \Omega, x \in A} \left\{ \sum_\sigma \sum_{i=1}^{\ell} \frac{\mu_A(\sigma)}{\mu_A(\eta)} \mathbb{1}_{\xi(\sigma)=i} \mathbb{1}_{\Gamma_x(\sigma) \ni \eta} \right\} \le \frac{8}{q^3}.$$

Indeed $\mu_A(\sigma)/\mu_A(\eta) \le \frac{p^2}{q^2} \max(\frac{p}{q}, \frac{q}{p})$ since any $\eta \in \Gamma_x(\sigma)$ has at most two extra empty sites with respect to σ and differs from σ in at most three sites, and we used a computing argument.

Recall that $y = y(x,\eta)$. It follows fom the latter that

$$\operatorname{Var}_{\mu_A}(f) \le \frac{128\ell}{q^3} \sum_{x \in A} \sum_\eta \mu_A(\eta) c_y(\eta) \operatorname{Var}_y(f)$$
$$\le \frac{128\ell}{q^3} K' \sum_{u \in A} \sum_\eta \mu_A(\eta) c_u(\eta) \operatorname{Var}_u(f)$$

where

$$K' = \sup_\eta \sum_{x \in A} \mathbb{1}_{y(x,\eta)=u} \le \sup_{u \in A} |B(u,\ell)|.$$

The result follows since the graph has polynomial growth. □

In Proposition 6.2 we used the following lemma.

Lemma 6.5. *Take $\Lambda, A, B \subset V$ such that $\Lambda = A \cup B$ and $A \cap B = \emptyset$. Define $c_A = \mathbb{1}_{\hat\Omega_A}$ and $c_B = \mathbb{1}_{\hat\Omega_B}$ where $\hat\Omega_A$ and $\hat\Omega_B$ are defined in* (6.1). *Assume that $\max(1-\mu(c_A), 1-\mu(c_B)) < 1/16$. Then, for all $f \colon \hat\Omega_\Lambda \to \mathbb{R}$ with $\hat\mu_\Lambda(f) = 0$ it holds*

$$\operatorname{Var}_{\hat\mu_\Lambda}(f) \le 24 \hat\mu_\Lambda[c_B \operatorname{Var}_{\mu_A}(\tilde f) + c_A \operatorname{Var}_{\mu_B}(\tilde f)]$$

where $\tilde f \colon \Omega_\Lambda \to \mathbb{R}$ is defined as

$$\tilde f(\sigma) = \begin{cases} f(\sigma) & \text{if } \sigma \in \hat\Omega_\Lambda \\ 0 & \text{otherwise} \end{cases}$$

Proof. Recalling the definition of the variance we have

$$\begin{aligned}\mathrm{Var}_{\hat\mu_\Lambda}(f) &= \inf_{m\in\mathbb{R}} \hat\mu_\Lambda(|f-m|^2)\\ &\le \frac{1}{\mu_\Lambda(\hat\Omega_\Lambda)} \inf_{m\in\mathbb{R}} \mu_\Lambda((f\mathbb{1}_{\hat\Omega_\Lambda}-m)^2)\\ &= \frac{1}{\mu_\Lambda(\hat\Omega_\Lambda)} \mathrm{Var}_{\mu_\Lambda}(\tilde f).\end{aligned}$$

Observe now that, by construction, $\mu_\Lambda(\tilde f)=0$ and $(1-c_A)(1-c_B)\tilde f=0$ so that we can apply Lemma 6.6 below and obtain

$$\mathrm{Var}_{\hat\mu_\Lambda}(f) \le \frac{24}{\mu_\Lambda(\hat\Omega_\Lambda)} \mu_\Lambda[c_B \mathrm{Var}_{\mu_A}(\tilde f) + c_A \mathrm{Var}_{\mu_B}(\tilde f)]$$

and the result follows. □

The next Lemma might be heuristically seen as a result on the spectral gap of some constrained blocks dynamics (see [2]). Such a bound can be of independent interest.

Lemma 6.6. *Let $\Lambda = A\cup B$ with A, $B\subset V$ satisfying $A\cap B=\emptyset$. Define μ_A and μ_B two probability measures on $\{0,1\}^A$ and $\{0,1\}^B$ respectively, and $\mu=\mu_A\otimes\mu_B$. Take c_A, $c_B\colon\{0,1\}^\Lambda\to[0,1]$ with support in A and B respectively. For any function g on $\{0,1\}^\Lambda$ such that $(1-c_A)(1-c_B)g=0$ it holds*

$$\mathrm{Var}_\mu(g) \le 12\mu[c_B^2 \mathrm{Var}_{\mu_A}(g) + c_A^2 \mathrm{Var}_{\mu_B}(g)] + 8\max(1-\mu(c_A), 1-\mu(c_B)) \mathrm{Var}_\mu(g)$$

Proof. Fix g on $\{0,1\}^\Lambda$ such that $(1-c_A)(1-c_B)g=0$ and assume without loss of generality that $\mu(g)=0$. First we write

$$\begin{aligned} g &= c_B(g-\mu_A(g)) + (1-c_B)c_A(g-\mu_B(g)) + (1-c_B)c_A\mu_B(g)\\ &\quad - (1-c_B)c_A\mu_A(g) + (1-c_B)(1-c_A)(g-\mu_A(g)) + \mu_A(g)\\ &= c_B(g-\mu_A(g)) + (1-c_B)c_A(g-\mu_B(g)) + (1-c_B)c_A\mu_B(g) + c_B\mu_A(g)\end{aligned}$$

where we used the first hypothesis on g, $(1-c_A)(1-c_B)g=0$, and we arranged the terms. Therefore since we assumed $\mu(g)=0$ and c_A, $c_B\in[0,1]$

$$\begin{aligned}\mathrm{Var}_\mu(g) = \mu(g^2) &\le 4\mu(c_B^2(g-\mu_A(g))^2) + 4\mu(c_A^2(g-\mu_B(g))^2)\\ &\quad + 4\mu(\mu_B(g)^2) + 4\mu(\mu_A(g)^2)\\ &= 4\mu[c_B^2 \mathrm{Var}_{\mu_A}(g) + c_A^2 \mathrm{Var}_{\mu_B}(g)]\\ &\quad + 4\mu(\mu_B(g)^2) + 4\mu(\mu_A(g)^2).\end{aligned}$$

We now treat the fourth term in the latter inequality.

$$\begin{aligned}[\mu_A(g)]^2 &= [\mu_A(g)-\mu(g)]^2 = [\mu_A(g-\mu_B(g))]^2\\ &= [\mu_A(c_A[g-\mu_B(g)]) + \mu_A([1-c_A][g-\mu_B(g)])]^2\\ &\le 2\mu_A(c_A^2)\mu_A(c_A^2[g-\mu_B(g)]^2) + 2\mu_A((1-c_A)^2)\mu_A([g-\mu_B(g)]^2)\end{aligned}$$

If we average with respect to μ we have

$$\mu(\mu_A(c_A^2)\mu_A(c_A^2[g-\mu_B(g)]^2)) = \mu_A(c_A^2)\mu(c_A^2 \mathrm{Var}_{\mu_B}(g))$$

and, using Cauchy-Schwarz inequality and $x^2 \le x$ for $x \in [0,1]$,

$$\begin{aligned}\mu(\mu_A((1-c_A)^2)\mu_A([g-\mu_B(g)]^2)) &= \mu_A((1-c_A)^2)\mu([g-\mu_B(g)]^2) \\ &\le (1-\mu(c_A))\operatorname{Var}_\mu(g),\end{aligned}$$

so that

$$\mu(\mu_A(g)^2) \le 2\mu_A(c_A^2)\mu(c_A^2 \operatorname{Var}_{\mu_B}(g)) + 2(1-\mu(c_A))\operatorname{Var}_\mu(g).$$

An analogous calculation for $\mu(\mu_B(g)^2)$ allows to conclude the proof. □

REFERENCES

[1] C. Ané, S. Blachère D., Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Société Mathématique de France, Paris, 2000, With a preface by Dominique Bakry and Michel Ledoux.

[2] N. Cancrini, F. Martinelli, C. Roberto, and C. Toninelli, *Kinetically constrained spin models*, Probab. Theory Related Fields **140** (2008), no. 3-4, 459–504. MR 2365481 (2009e:60212)

[3] ———, *Facilitated spin models: recent and new results*, Methods of contemporary mathematical statistical physics, Lecture Notes in Math., vol. 1970, Springer, Berlin, 2009, pp. 307–340.

[4] N. Cancrini, F. Martinelli, R. Schonmann, and C. Toninelli, *Facilitated oriented spin models: some non equilibrium results*, J. Stat. Phys. **138** (2010), no. 6, 1109–1123.

[5] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750.

[6] G. H. Fredrickson and H. C. Andersen, *Kinetic ising model of the glass transition*, Phys. Rev. Lett. **53** (1984), 1244–1247.

[7] ———, *Facilitated kinetic ising models and the glass transition*, J.Chem.Phys. **83** (1985), 5822–5831.

[8] R. I. Grigorchuk, *On growth in group theory*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990) (Tokyo), Math. Soc. Japan, 1991, pp. 325–338. MR 1159221 (93e:20001)

[9] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083.

[10] Richard Holley and Daniel Stroock, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys. **46** (1987), no. 5-6, 1159–1194. MR 893137 (89e:82013)

[11] C.Toninelli J.P. Garrahan, P. Sollich, *Dynamical heterogeneities and kinetically constrained models*, preprint arXiv:1009.6113, 2011.

[12] Sébastien Léonard, Peter Mayer, Peter Sollich, Ludovic Berthier, and Juan P. Garrahan, *Non-equilibrium dynamics of spin facilitated glass models*, J. Stat. Mech. Theory Exp. (2007), no. 7, P07017, 69 pp. (electronic). MR 2335696 (2008e:82072)

[13] T. M. Liggett, *Interacting particle systems*, Springer-Verlag, New York, 1985.

[14] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, Lectures on probability theory and statistics (Saint-Flour, 1997), Springer, Berlin, 1999, pp. 93–191.

[15] F. Martinelli and C. Toninelli, Preprint (2011).

[16] F. Ritort and P. Sollich, *Glassy dynamics of kinetically constrained models*, Advances in Physics **52** (2003), no. 4, 219–342.

[17] L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Springer, Berlin, 1997, pp. 301–413.

[18] Roberto H. Schonmann, *On the behavior of some cellular automata related to bootstrap percolation*, Ann. Probab. **20** (1992), no. 1, 174–193. MR 1143417 (93b:60231)

[19] Daniel W. Stroock and Boguslaw Zegarliński, *The logarithmic Sobolev inequality for discrete spin systems on a lattice*, Comm. Math. Phys. **149** (1992), no. 1, 175–193. MR 1182416 (93j:82013)

E-mail address: `oriane.blondel@ens.fr`

UNIV. PARIS DIDEROT, SORBONNE PARIS CIT, LPMA, UMR 7599, F-75205 PARIS, FRANCE

E-mail address: `nicoletta.cancrini@roma1.infn.it`

DIP. MATEMATICA UNIV. L'AQUILA, I-67010 L'AQUILA, ITALY

E-mail address: `martin@mat.uniroma3.it`

DIP. MATEMATICA, UNIV. ROMA TRE, LARGO S.L.MURIALDO 00146, ROMA, ITALY

E-mail address: `cyril.roberto@univ-mlv.fr`

L.A.M.A., UNIV. MARNE-LA-VALLÉE, 5 BD DESCARTES 77454 MARNE-LA-VALLÉE FRANCE

E-mail address: `Cristina.Toninelli@lpt.ens.fr`

LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÈATOIRES CNRS-UMR 7599 UNIVERSITÉS PARIS VI-VII 4, PLACE JUSSIEU F-75252 PARIS CEDEX 05 FRANCE