Markov Processes Relat. Fields **19**, 383–406 (2013)

Markov Processes and Related Fields

Fredrickson – Andersen One Spin Facilitated Model out of Equilibrium*

O. Blondel[1], N. Cancrini[2], F. Martinelli[3], C. Roberto[4] and C. Toninelli[5]

[1] Univ. Paris Diderot, Sorbonne Paris Cité, LPMA, UMR 7599, F-75205 Paris, France. E-mail: oriane.blondel@ens.fr
[2] Dip. Matematica Univ. L'Aquila, I-67010 L'Aquila, Italy. E-mail: nicoletta.cancrini@roma1.infn.it
[3] Dip. Matematica, Univ. Roma Tre, Largo S.L.Murialdo 00146, Roma, Italy. E-mail: martin@mat.uniroma3.it
[4] MODAL'X, Université Paris Ouest Nanterre La Défense, 200 avenu de la République 92000 Nanterre, France. E-mail: croberto@math.cnrs.fr
[5] Lab. de Probabilités et Modèles Alèatoires CNRS-UMR 7599 Univ. Paris VI-VII, 4, Place Jussieu F-75252 Paris Cedex 05 France. E-mail: Cristina.Toninelli@lpt.ens.fr

Received January 18, 2013

Abstract. We consider the Fredrickson and 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 the non-equilibrium dynamics started from an initial distribution ν different from the stationary product p-Bernoulli measure μ. We assume that, under ν, the distance between two nearest empty sites has exponential moments. We then 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 $d > 1$. Our result can be generalized to other *non cooperative* models.

KEYWORDS: interacting particle systems, Fredrickson – Andersen model, convergence to equilibrium

AMS SUBJECT CLASSIFICATION: 60K35, 82C20

*Work supported by the European Research Council through the "Advanced Grant" PTRELSS 228032. The second author was also partially supported by PRIN 2009 protocollo n.2009TA2595.02 and the fourth author by the French Ministry of Education through ANR-2010-BLAN-0108

1. Introduction

Fredrickson – Andersen one spin facilitated model (FA1f) [7, 8] 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 [9, 17] 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 a single empty site is sufficient to ensure irreducibility of the chain. KCSM in which a fine subset of empty sites is able to move around and empty the whole space are called *non-cooperative* and are in general easier to analyze than cooperative ones. Note also 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. 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. boundedness of the logarithmic Sobolev constant or positivity of the entropy constant uniformly in the system size. The latter requirement certainly does not hold (see Section 7.1 of [2]) and even the basic question of whether convergence to μ occurs remains open in the infinite volume case. Of course, due to the existence of blocked configurations, convergence to μ cannot hold *uniformly* in 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. [13]) is when ν is a product Bernoulli (p') measure with $p' \neq p$ and $p' \neq 1$). In this case the most natural guess is that convergence to equilibrium occurs for any local (i.e. depending on finitely many occupation variables) 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 case of KCSM where this result has been proved [4] (see also [6]) is the East model, that is a one dimensional model in which the constraint at x requires the neighbour to the right of x to 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 $p' \sim p$).

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 distance between two nearest empty sites has exponential moments. That includes in particular any non-trivial Bernoulli product measure with $p' \neq p$ but also the case in which ν is the Dirac measure on a fixed configuration with infinitely many empty sites and such that the distance between two nearest empty sites is uniformly bounded. The derived convergence is either exponential or stretched exponential depending on the growth of the graph. In the particular case $\mathcal{G} = \mathbb{Z}^d$, we can prove exponential relaxation only for $d = 1$. If $d > 1$ we get a stretched exponential behavior. Although our results can be generalized to other *non cooperative* KCSM (see Section 6 for a specific example and [2] for the general definition of this class) we consider here only the FA1f case to let the paper be more readable.

We finish with a short road map of the paper. In Section 2 we introduce the notations and give the main result, Theorem 2.1, which is proven in Section 5. The strategy to derive this result can be summarized as follows. We first replace $\mathbb{E}_\nu(f(\eta_t))$ with a similar quantity but computed w.r.t. the FA1f finite volume process (actually a finite state, continuous time Markov chain evolving in a finite ball of radius proportional to time t around the support of f). This first reduction is standard and it follows easily from the so-called *finite speed of propagation*. Then we show that, with high probability, only the evolution of a restricted chain inside a suitable ergodic component matters. This reduction is performed via a general result on Markov processes which we derive in Section 3.

The ergodic component is chosen in such a way that the log-Sobolev constant for the restricted chain is much smaller than t. This second reduction is new and it is at this stage that the restriction on q appears and that all the difficulties of the non-equilibrium dynamics appear. Its implementation requires the estimate of the spectral gap of the process restricted to the ergodic component (see Section 6) and the study of the persistence of zeros out of equilibrium (see Section 4).

2. Notation and result

2.1. The graph

Let $\mathcal{G} = (V, E)$ be an infinite, connected graph with vertex set V, edge set E and graph distance $d(\cdot,\cdot)$. Given $x \in V$ the set of neighbors of x 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-dimensional 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. [14]). It acts on local functions as

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

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

$$\mathcal{D}(\mathcal{L}) = \Big\{ f\colon \Omega \to \mathbb{R} \text{ s.t. } \sum_{x\in V}\sup_{\sigma\in\Omega}|f(\sigma^x) - f(\sigma)| < \infty \Big\}$$

and Dirichlet form given by

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

Here $\operatorname{Var}_x(f) \equiv \int d\mu(\omega(x))f^2(\omega) - \big(\int d\mu(\omega(x))f(\omega)\big)^2$ 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 $\eta \in \Omega$

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

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 $\mathcal{L}_\Lambda$. We also let σ^Λ_t be the configuration at time t of the process starting from σ_Λ with empty boundary condition.

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

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

be the distance of x from 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 $f: \Omega \to \mathbb{R}$ 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 $c = c\big(q, k, D, \kappa, |\mathrm{supp}(f)|\big)$ such that*

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

Remark 2.1. We expect that our results hold also for $0 < q \le 1/2$. This needs a more precise control of the behavior of $\xi_t^x = \xi^x(\sigma_t)$. In dimension one we can obtain a better threshold by calculating further time derivatives of $u(t) = \mathbb{E}_\eta(\theta^{\xi_t})$, see Proposition 4.1 below.

Remark 2.2. 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 2.3. If one considers the case in which ν is the product of Bernoulli-p' on $\mathcal{G}$, one has that, 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, $\kappa \le 1/(1 - p'\theta_o)$ 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 2.4. Note that graphs with polynomial growth are amenable. We stress anyway that there exist amenable graphs which do not satisfy our assumption. This is due to Proposition 5.1 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 [10]). The same happens to any graph with exponential growth (such as for example any regular n-ary tree ($n \ge 2$)).

3. A preliminary result on Markov processes

We prove here a general result which relates the behavior of a Markov process on a finite space to that of a restricted Markov process. This result, which might be of independent interest, will be a key tool in our analysis. Indeed we will use it in the proof of Theorem 2.1 to reduce the evolution of the FA1f process on a large volume to the same process on smaller sets on a properly defined ergodic component.

We start by recalling some basic notions on continuous time Markov chains which will be used in the following. Let S be a finite space and $Q=(q(x,y))_{x,y\in S}$ be a transition rate matrix, namely a matrix such that for any $x,y\in S$ it holds

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

Recall that Q defines a continuous time Markov chain $(X_t)_{t\geq 0}$ on S as follows [14]. 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. Assume that $(X_t)_{t\geq 0}$ is reversible with respect to a probability measures π. Then, we define the spectral gap $\gamma(Q)$ and the log-Sobolev constant $\alpha(Q)$ of the chain as

$$\gamma(Q):=\inf_{f:f\neq\text{const}} \frac{\sum_{x,y}\pi(x)p(x,y)(f(y)-f(x))^2}{2\operatorname{Var}_\pi(f)}, \tag{3.1}$$

$$\alpha(Q):=\sup_{f:f\neq\text{const}} \frac{2\operatorname{Ent}_\pi(f^2)}{\sum_{x,y}\pi(x)p(x,y)(f(y)-f(x))^2}, \tag{3.2}$$

where $\operatorname{Ent}_\pi(f)=\pi(f\log f)-\pi(f)\log\pi(f)$ denotes the entropy of f. Let $(P_t)_{t\geq 0}$ be the semigroup of the Markov chain. Then

$$\operatorname{Var}_\pi(P_tf)\leq e^{-2tc}\operatorname{Var}_\pi(f) \quad \forall f \tag{3.3}$$

is equivalent to $\gamma\geq c$. On the other hand the positivity of the log-Sobolev constant is equivalent to the following hypercontractivity property [11]

$$\|P_tf\|_{L^p(\pi)}\leq\|f\|_{L^2(\pi)} \tag{3.4}$$

for all $t\geq 0$ and all $p\leq 1+\exp\{4t/\alpha\}$. We refer to [1] for an introduction of these notions.

We are now ready to introduce the restricted Markov chain. Fix $\mathcal{A}\subset S$ and set

$$\hat{\mathcal{A}}=\mathcal{A}\cup\big\{y\notin\mathcal{A}: q(x,y)>0 \ \text{ for some } \ x\in\mathcal{A}\big\}. \tag{3.5}$$

Let $(\hat{X}_t)_{t\geq 0}$ be a continuous time Markov chain (which we will call the *hat chain* on $\hat{\mathcal{A}}$ with transition rate matrix $\hat{Q}=(\hat{q}(x,y))_{x,y\in\hat{\mathcal{A}}}$ which satisfies

$$\hat{q}(x,y)=q(x,y) \quad \forall (x,y)\in\mathcal{A}\times\hat{\mathcal{A}} \tag{3.6}$$

and assume that the process is reversible with respect to a measure $\hat{\pi}$. We denote by $\hat{\gamma}$ and $\hat{\alpha}$ the spectral gap and log-Sobolev constant of the hat chain, namely $\hat{\gamma}:=\gamma(\hat{Q})$ and $\hat{\alpha}:=\alpha(\hat{Q})$.

Proposition 3.1. *Let* $(X_t)_{t\geq 0}$, $(\hat{X}_t)_{t\geq 0}$, π, $\hat{\pi}$, $\hat{\gamma}$ *and* $\hat{\alpha}$ *as above. Then, for any initial probability measure* ν *on* S *and all* $f\colon S\to\mathbb{R}$ *with* $\pi(f)=0$, *it holds for any* $t\geq 0$

$$|\mathbb{E}_\nu(f(X_t))| \leq |\hat{\pi}(f)| + 4\|f\|_\infty \mathbb{P}_\nu(\mathcal{A}_t^c) + \|f\|_\infty \exp\Big\{-\hat{\gamma}\frac{t}{2} + e^{-2t/\hat{\alpha}}\log\frac{1}{\hat{\pi}^*}\Big\} \tag{3.7}$$

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

Remark 3.1. The standard argument (see [12, 19]) using the log-Sobolev constant would lead to

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

with $\gamma=\gamma(Q)$ and $\alpha=\alpha(Q)$. The difference in Proposition 3.1 comes from the fact that we deal with the hat chain. This can be useful if the choice of $\hat{\mathcal{A}}$ and $\hat{Q}$ are done properly so that the log-Sobolev constant $\hat{\alpha}$ is smaller than α and/or the spectral gap $\hat{\gamma}$ is larger then γ. This will be the case for the application of the above result in the proof of Theorem 2.1 for which, fixed t, we will have to consider a state space which depends on t and the corresponding chain will have $\alpha\simeq t^d$ and $\log(1/\pi)\simeq t^d$. Hence the standard argument gives

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

and therefore does not prove decay in t. We will instead devise a hat chain for which $\hat{\gamma}\geq c>0$ and $\hat{\alpha}$ is much smaller than t so that the dominant term in $\exp\big\{-\hat{\gamma}t/2 + \exp\{-2t/\hat{\alpha}\}\log(1/\hat{\pi}^*)\big\}$ is given by the gap term $\hat{\gamma}t$. The price to pay are the first two extra terms in (3.7) that we will analyze separately.

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

$$|\mathbb{E}_\nu(f(X_t))| \leq |\hat{\pi}(f)| + \|g\|_\infty\mathbb{P}_\nu(\mathcal{A}_t^c) + \big|\mathbb{E}_\nu\big(g(X_t)\mathbb{1}\{\mathcal{A}_t\}\big)\big|. \tag{3.8}$$

We now concentrate on the last term in (3.8). By definition of the chains $(X_t)_{t\geq 0}$ and $(\hat{X}_t)_{t\geq 0}$ one has

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

Hence, by the Hölder inequality, we have

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

where for any $x \in \hat{\mathcal{A}}$ we let $h(x) = \nu(x)/\hat{\pi}(x)$ and β, $\beta' \geq 1$, that will be chosen later, are such that $1/\beta + 1/\beta' = 1$. To bound the previous expression take $\beta' = 1 + \exp\{2t/\hat{\alpha}\}$. Using (3.4) and (3.3) we obtain

$$\begin{aligned}\|\hat{P}_t g\|_{L^{\beta'}(\hat{\pi})} &= \|\hat{P}_{t/2}\hat{P}_{t/2} g\|_{L^{\beta'}(\hat{\pi})} \leq \|\hat{P}_{t/2} g\|_{L^2(\hat{\pi})} \\ &\leq e^{-\hat{\gamma}t/2}\|g\|_{L^2(\hat{\pi})} \leq e^{-\hat{\gamma}t/2}\|f\|_\infty.\end{aligned}$$

On the other hand

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

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

4. Persistence of zeros out of equilibrium

In this section we study the behavior of the minimal distance from a fixed site to the nearest site at which one finds a vacancy. The result that we obtain will be a key tool for the proof of our main Theorem 2.1.

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,

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

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 $q \in (\theta/(\theta+1), 1]$ and all initial configuration η, it holds*

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

where

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

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

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

Case (ii). Define $E(\sigma) = \big\{y \in V : d(x,y) = \xi(\sigma) \text{ and } \sigma(y) = 0\big\}$ and $F(\sigma) = \big\{y \in V : d(y,E) = 1 \text{ and } d(x,y) = \xi(\sigma) - 1\big\}$. 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 1).

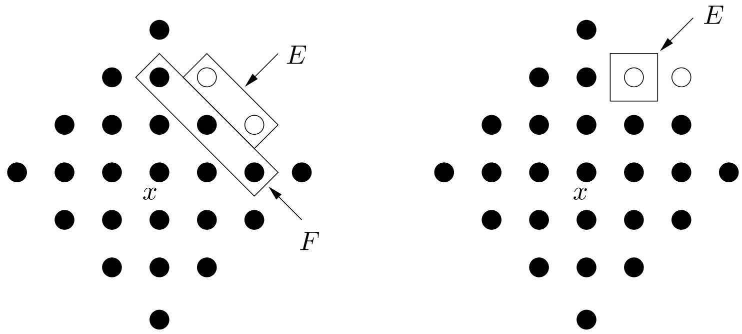

Figure 1. On the graph $\mathcal{G} = \mathbb{Z}^2$, two examples of configurations for which $\xi^x = 3$. On the left ξ^x cannot increase since $|E| \geq 2$, it can decrease by a flip (legal thanks to the empty sites in E) in any points of F. On the right ξ^x can either increase or decrease.

Hence

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

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

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

Therefore, since $p = 1 - q$,

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

and the expected result follows. □

5. Proof of Theorem 2.1

In this section we prove Theorem 2.1. We will first reduce the study of the evolution of the process from infinite volume to a finite ball of radius proportional to t thanks to finite speed of propagation. Then by using Proposition 3.1 we reduce to the study of a restricted process on smaller sets on some ergodic component so that the log-Sobolev constant of the restricted process is much smaller than t (recall Remark 3.1). In order to estimate the probability that the process gets out the ergodic component (namely to bound the second term in (3.7)) we will use Proposition 4.1 which allows to upper bound the probability of a region to be completely filled.

Proof of Theorem 2.1. Throughout the proof c denotes some positive constant $c = c(q, k, D, \kappa, |\mathrm{supp(f)}|)$ which may change from line to line.

Fix $t \geq 2$ and a local function f. Let $x \in V$ and r integer be such that $\mathrm{supp(f)} \subset B(x, r)$. Standard arguments using finite speed of propagation (see e.g. [15]) prove that for any initial measure ν on Ω it holds

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

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

$$\left|\mathbb{E}_\nu(f(\sigma_t))\right| \leq \left|\mathbb{E}_\nu\big(f(\sigma_t^\Lambda)\big)\right| + c\|f\|_\infty e^{-t}. \tag{5.1}$$

Let $\Lambda_1, \Lambda_2, \ldots, \Lambda_n \subset \Lambda$ be connected sets 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 *connected partition* of Λ. The following proposition holds true:

Proposition 5.1. *For any $\Lambda \subset V$ and any f local, with* $\mathrm{supp(f)} \subset \Lambda$ *and $\mu(f) = 0$ there exists a constant $c = c(q, |\mathrm{supp(f)}|)$ such that for any connected partition $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$ of Λ, for any initial probability measure ν on Ω, it holds that*

$$\begin{aligned}\left|\mathbb{E}_\nu\big(f(\sigma_t^\Lambda)\big)\right| \leq c\|f\|_\infty \Big(& n\exp\{-qm\} + t|\Lambda| \sup_{s\in[0,t]} \mathbb{P}_\nu\big(\sigma_s^\Lambda \notin \mathcal{A}\big) \\ &+ |\Lambda|\exp\{-t/3\} + \exp\Big\{-\frac{t}{c} + c|\Lambda|e^{-t/(cM)}\Big\}\Big)\end{aligned}$$

provided that $n\exp\{-qm\} < 1/2$ where $m := \min\{|\Lambda_1|, \ldots, |\Lambda_n|\}$, $M := \max\{|\Lambda_1|, \ldots, |\Lambda_n|\}$ and $\mathcal{A}$ is the set of configurations containing at least two empty sites in each Λ_i, namely

$$\mathcal{A} = \bigcap_{i=1}^{n} \Big\{\sigma \in \Omega_\Lambda \ s.t. \sum_{x\in\Lambda_i} (1 - \sigma(x)) \geq 2\Big\}. \tag{5.2}$$

We postpone the proof of this proposition to the end of this section.

Observe that for any positive integer $\ell \le t$, there exists[1] a connected partition $\Lambda_1, \dots, \Lambda_n$ of Λ, and vertices $x_1, \dots, x_n \in V$, such that for any i, $B(x_i, \ell) \subset \Lambda_i \subset B(x_i, 6\ell)$. Then, take $\ell = \varepsilon[t/\log t]^{1/D}$ if $D > 1$ and $\ell = \varepsilon t$ if $D = 1$ for some $\varepsilon > 0$ that will be chosen later and observe that, with this choice,

$$M = \max(|\Lambda_1|, \dots, |\Lambda_n|) \le k 6^D \ell^D$$

(since $\mathcal{G}$ has (k, D)-polynomial growth). Furthermore

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

Since $n \le |\Lambda| \le ct^D$, equation (5.1) and Proposition 5.1 guarantee that

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

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 $x_i^+, x_i^- \in V$, $B(x_i^\pm, \ell/4) \subset \Lambda_i^\pm$ (a proof of the existence of such vertices is 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 $\xi^{x_i^+}(\sigma_s^\Lambda) \ge \ell/4$. Hence, thanks to a union bound, Markov's inequality, and Proposition 4.1, there exists $\theta > 1$ such that

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

where we used the definition of ℓ, the assumption $\sup_{x\in V} \mathbb{E}_\nu(\theta_o^{\xi^x}) < \infty$ and the fact that $n \le |\Lambda| \le ct^D$. This ends the proof. $\square$

[1] One can construct $\Lambda_1, \dots, \Lambda_n, x_1, \dots, x_n$ as follows. Recall that $\Lambda = B(x, r+100t)$. Order (arbitrarily) the sites $y_1, y_2, \dots, y_N$ of $\{z \in \Lambda\colon B(z, \ell) \subset \Lambda$ and $d(z, x) = 2i(\ell+1)-1$ for some $i \ge 1\}$ and perform the following algorithm: set $x_1 = x$, $i_0 = 0$, and for $k \ge 1$ set $x_{k+1} = y_{i_k}$ with $i_k := \inf\{j \ge i_{k-1} + 1 : B(y_j, \ell) \cap (\cup_{i=1}^k B(x_i, \ell)) = \emptyset\}$. 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 $i \ne j$, $B(x_i, \ell) \subset \Lambda$ for all i and any site $y_k \notin A := \cup_{i=1}^n B(x_i, \ell)$ is at distance at most 5ℓ 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.

We are now left with proving Proposition 5.1.

Proof of Proposition 5.1. Fix $\Lambda \subset V$, f local, with $\mathrm{supp}(f) \subset \Lambda$ and $\mu(f) = 0$. Fix a connected partition $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$ of Λ and an initial probability measure ν on Ω.

Our aim is to apply Proposition 3.1. We let $S = \Omega_\Lambda$ and $(X_t)_{t\geq 0} = \big(\sigma_t^\Lambda\big)_{t\geq 0}$. The corresponding transition rates are, for all $\sigma, \eta \in \Omega_\Lambda$,

$$q(\sigma,\eta) = \begin{cases} c_{x,\Lambda}(\sigma)\big[q\sigma(x) + p(1-\sigma(x))\big], & \text{if } \eta = \sigma^x, \\ -\displaystyle\sum_{x\in\Lambda} q(\sigma,\sigma^x), & \text{if } \eta = \sigma, \\ 0 & \text{otherwise.} \end{cases}$$

We define $\mathcal{A}$ as in (5.2), namely the set of configurations in Ω_Λ such that there exist at least two empty sites in each set Λ_i, and $\hat{\mathcal{A}}$ as in (3.5). Next we define $(\hat{X}_t)_{t\geq 0}$ on $\hat{\mathcal{A}}$ via the rates $\hat{q}(\sigma,\eta) = q(\sigma,\eta)\ \forall \sigma,\eta \in \hat{\mathcal{A}}$. In words, $(\hat{X}_t)_{t\geq 0}$ corresponds to a modification of the FA1f process in which the moves that would lead the process to leave $\hat{\mathcal{A}}$ are suppressed. Let $\pi = \mu_\Lambda$ and $\hat{\pi}(\cdot) = \mu_\Lambda(\cdot \mid \hat{\mathcal{A}})$. It is immediate to verify that $(X_t)_{t\geq 0}$ and $(\hat{X}_t)_{t\geq 0}$ are reversible with respect to π and $\hat{\pi}$, respectively. By construction the above processes satisfy the property (3.5). Thus, thanks to Proposition 3.1, we have

$$\big|\mathbb{E}_\nu\big(f(\sigma_t^\Lambda)\big)\big| \leq |\hat{\pi}(f)| + \|f\|_\infty \Big(4\mathbb{P}_\nu\big(\mathcal{A}_t^c\big) + \exp\Big\{-\hat{\gamma}\frac{t}{2} + e^{-2t/\hat{\alpha}} \log\frac{1}{\hat{\pi}^*}\Big\}\Big). \quad (5.3)$$

We now study each term of the last inequality separately.

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

$$|\hat{\pi}(f)| = \frac{\big|\mu_\Lambda\big(f(1-\mathbb{1}\{\hat{\mathcal{A}}^c\})\big)\big|}{\mu_\Lambda(\hat{\mathcal{A}})} \leq \|f\|_\infty \frac{\mu_\Lambda(\hat{\mathcal{A}}^c)}{\mu_\Lambda(\hat{\mathcal{A}})} \leq \|f\|_\infty \frac{ne^{-qm}}{1-ne^{-qm}}. \quad (5.4)$$

We now 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 with more than $2t$ rings in the time interval $[0,t]$. Then, by standard large deviations of Poisson variables and a union bound, there exists a universal positive constant d such that $\mathbb{P}_\nu(\mathcal{A}_t^c \cap \mathcal{I}_t) \leq d|\Lambda|\exp\{-t/3\}$. Furthermore, using a union bound on all the rings on the event $\mathcal{I}_t^c$, we have

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

We deduce that

$$\mathbb{P}_\nu(\mathcal{A}_t^c) \leq c|\Lambda|\big(t \sup_{s\in[0,t]} \mathbb{P}_\nu\big(\sigma_s^\Lambda \notin \mathcal{A}\big) + e^{-t/3}\big). \quad (5.5)$$

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 $(\tilde{X}_t)_{t\geq 0}$ on $\hat{\mathcal{A}}$ via the rates, for $\sigma, \eta \in \hat{\mathcal{A}}$,

$$\tilde{q}(\sigma,\eta) = \begin{cases} c^{\omega}_{x,\Lambda_{i(x)}}(\sigma)\big[q\sigma(x) + p(1-\sigma(x))\big], & \text{if } \eta = \sigma^x, \\ -\sum_{x\in\Lambda} \tilde{q}(\sigma,\sigma^x), & \text{if } \eta = \sigma, \\ 0 & \text{otherwise} \end{cases}$$

where ω is the entirely filled configuration (i.e. such that $\omega(x) = 1$ for all $x \in V$) and $i(x)$ is such that $x \in \Lambda_{i(x)}$. In words, $(\tilde{X}_t)_{t\geq 0}$ corresponds to n independent FA1f processes inside the boxes Λ_i each evolving with filled boundary conditions on the ergodic component of the configurations with at least one zero, namely on $\hat{\Omega}_{\Lambda_i}$ where we set for any $A \subset \Lambda$

$$\hat{\Omega}_A = \big\{\sigma \in \Omega_\Lambda \text{ s.t. } \exists x_i \in A \text{ with } \sigma(x_i) = 0\big\}. \tag{5.6}$$

Note that $\hat{\mathcal{A}} = \bigcap_{i=1}^n \hat{\Omega}_{\Lambda_i}$ thus the FA1f constraint and the filled boundary condition on each box indeed guarantee that $(\tilde{X}_t)_{t\geq 0}$ does not exit $\hat{\mathcal{A}}$ and $\hat{\pi}$ is a reversible measure also for $(\tilde{X}_t)_{t\geq 0}$. Furthermore, since the occupied boundary conditions imply that for any $\sigma \in \Omega_\Lambda$ it holds $c^{\omega}_{x,\Lambda_{i(x)}}(\sigma) \leq c_{x,\Lambda}(\sigma)$ (a zero which is present in $\sigma_{\Lambda_{i(x)}}\omega_{\Lambda^c_{i(x)}}$ is also present in σ) , the following inequalities hold between the spectral gap and log-Sobolev constant of the hat and tilde process: $\hat{\alpha} \leq \tilde{\alpha}$ and $\hat{\gamma} \geq \tilde{\gamma}$, where $\hat{\alpha} := \alpha(\hat{Q})$, $\tilde{\alpha} := \alpha(\tilde{Q})$, $\hat{\gamma} := \gamma(\hat{Q})$ and $\tilde{\gamma} = \gamma(\tilde{Q})$ (see (3.1) and (3.2)). Observe now that $\tilde{X}_t$ restricted to each $\hat{\Omega}_{\Lambda_i}$ is ergodic and reversible with respect to $\hat{\mu}_i = \mu_{\Lambda_i}(\cdot \mid \hat{\Omega}_{\Lambda_i})$. Thus by the well-known tensorisation property of the Poincaré and the log-Sobolev inequalities (see e.g. [1, Chapter 1]), we conclude that $\tilde{\gamma} = \min(\tilde{\gamma}_1,\ldots,\tilde{\gamma}_n)$ and $\tilde{\alpha} = \max(\tilde{\alpha}_1,\ldots,\tilde{\alpha}_n)$ with $\tilde{\gamma}_i$ and $\tilde{\alpha}_i$ the spectral gap and log-Sobolev constant of an FA1f process on Λ_i with filled boundary condition on the ergodic component with at least one zero which, using (3.1) and (3.2), can be expressed as $\tilde{\gamma}_{\Lambda_i}$ (5.7) and $\tilde{\alpha}_{\Lambda_i}$ (5.8) respectively. Then, Proposition 5.2 below shows that $\tilde{\gamma} \geq c$ and $\tilde{\alpha} \leq c|\Lambda_i|$. Hence, for c as in Proposition 5.2 it holds

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

By collecting this inequality together with (5.3), (5.4) and (5.5) we end the proof. □

Proposition 5.2 ([3]). *Let $A \subset V$ be connected and $\hat{\mu}_A(\cdot) = \mu_A(\cdot \mid \hat{\Omega}_A)$. Let ω be the entirely filled configuration (i.e. such that $\omega(x) = 1$ for all $x \in V$). Then, there exists a constant $c = c(q)$ such that*

$$\tilde{\gamma}_A := \inf_{f:f\neq const.} \frac{\sum_{x\in A}\hat{\mu}_A(\,c^{\omega}_{x,A}\operatorname{Var}_x(f))}{\operatorname{Var}_{\hat{\mu}_A}(f)} \geq c \tag{5.7}$$

and

$$\tilde{\alpha}_A := \sup_{f: f \neq const.} \frac{\mathrm{Ent}_{\hat{\mu}_A}(f)}{\sum_{x\in A} \hat{\mu}_A(c^{\omega}_{x,A} \operatorname{Var}_x(f))} \le c|A|. \tag{5.8}$$

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, 18]

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

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

6. Spectral gap on the ergodic component

In this section we estimate the spectral gap of the process FA1f on a finite volume with occupied boundary conditions on the ergodic component of configurations with at least one zero. This result has been used in the proof of Theorem 2.1 as a key tool to prove Proposition 5.1. This has been done in [2,3]. We present here an alternative proof based on the ideas of [16] 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). The remaining of the proof of Theorem 2.1 for these models carries over along the same lines as for FA1f. 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 ($\sum_{y\in\hat{\mathcal{N}}_x}(1-\sigma(y)) \ge 2$, where $\hat{\mathcal{N}}_x = \{y\colon d(x,y) \le 2\}$), 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 FA1f model.

For every $\Lambda \subset V$ finite recall that $\hat{\Omega}_\Lambda$ is the set of configurations with at least one zero (5.6) and $\hat{\mu}_\Lambda(\cdot) = \mu_\Lambda(\cdot \mid \hat{\Omega}_\Lambda)$. By using (3.1) the spectral gap $\tilde{\gamma}_\Lambda$ for the dynamics on $\hat{\Omega}_\Lambda$ with filled boundary conditions can be expressed as

$$\tilde{\gamma}_\Lambda = \inf_{f: f\neq const.} \frac{\sum_{x\in\Lambda} \hat{\mu}_\Lambda(\hat{c}_x \operatorname{Var}_x(f))}{\operatorname{Var}_{\hat{\mu}_\Lambda}(f)} \tag{6.1}$$

where the infimum runs over all non constant functions $f : \hat{\Omega}_\Lambda \to \mathbb{R}$, $\operatorname{Var}_x(f) := \operatorname{Var}_{\mu_{\{x\}}}(f)$, and $\hat{c}_x(\sigma) := c^{\omega}_{x,\Lambda}(\sigma)$ with ω the entirely filled configuration, i.e. $\omega(x) = 1$ for all $x \in V$. We are now ready to 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 $\Lambda \subset V$*

$$\tilde{\gamma}_\Lambda \geq C \frac{q^{D+4}}{\log(2/q)^{D+1}}.$$

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 [16].

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 6.1. *For any finite connected $\Lambda \subset V$ with $8p^{\mathrm{diam}(\Lambda)/3} < 1/2$ it holds*

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

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

$$\tilde{\gamma}_\Lambda \geq c q^{\log_2(1/q)}.$$

We present below an alternative strategy (based on [16]) 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 $f \colon \hat{\Omega}_\Lambda \to \mathbb{R}$ with $\hat{\mu}(f) = 0$ 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 $d(x, y) = \ell := \mathrm{diam}(\Lambda)$ and define $A_0 = \{z \in \Lambda : d(x, z) \leq \ell/3\}$ and $B_0 = \{z \in \Lambda : d(y, z) \leq \ell/3\}$. 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.

Thanks to Lemma 6.2 below (our hypothesis implies that $\max(1-\mu(c_A), 1-\mu(c_B)) < 1/16$)

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

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 (5.6).

Consider the first term. Define the random variable

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

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

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

where $A_n = \{x \in \Lambda : d(A,x) \le n-1\}$ and we used the convexity of the variance (which is valid since the event $\{\zeta = n\}$ does not depend, by construction, on the value of the configuration σ_{A_n} inside A_n). The indicator function above $\mathbb{1}\{\zeta = n\}$ 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 $n \ge 1$,

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

where we used the fact that the events $\{\zeta = n\}$ 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

$$\mathbb{1}\{\zeta = n\}\,\mathbb{1}\{Z = z\}c^z_{y,A} \le \mathbb{1}\{\zeta = n\}\,\mathbb{1}\{Z = z\}\hat c_y$$

for any $y \in A_n$. Hence,

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

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

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

Proposition 6.2. *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*

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

We postpone the proof of Proposition 6.2 to end the proof of Theorem 6.1.

Proof of Theorem 6.1. The result follows at once combining Proposition 6.1 and Proposition 6.2. $\square$

In order to prove Proposition 6.2, 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 $x, y \in A$, $x \le y$ 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 $A_x = \{y \in A : y \ge x\}$ and $\tilde A_x = A_x \setminus \{x\}$.

Lemma 6.1. *Fix a connected set $A \subset V$, and $z \in \partial A$. For any $x \in A$ and $\sigma \in \Omega$, let $E_x \subset \Omega_{\tilde A_x}$, $\Delta_x = \operatorname{supp}\big(\mathbb{1}\{E_x\}\big)$ and $\tilde c_x(\sigma) = \mathbb{1}\{E_x\}(\sigma_{\tilde A_x})$. Assume that*

$$
\sup_{x\in A}\mu(1-\tilde c_x)\ \sup_{x\in A}\big|\{y\in A : \Delta_y\cup\{y\}\ni x\}\big| < \frac14.
$$

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

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

Proof. We follow [16]. During the proof, to simplify the notations, we set $\mathrm{Var}_B = \mathrm{Var}_{\mu_B}$, for any B. First, we claim that

$$\mathrm{Var}_A(f) \le \sum_{x\in A} \mu_A\big(\mathrm{Var}_{A_x}(\mu_{\tilde A_x}(f))\big). \tag{6.2}$$

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

$$\mathrm{Var}_A(f) = \mu_A(\mathrm{Var}_{\tilde A_x}(f)) + \mathrm{Var}_A(\mu_{\tilde A_x}(f)).$$

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 (6.2) 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

$$\begin{aligned}\mu_A&\left[\mathrm{Var}_{A_x}(\mu_{\tilde A_x}(f))\right] \\ &\le 2\mu_A\big[\mathrm{Var}_{A_x}\left(\mu_{\tilde A_x}(\tilde c_x f)\right)\big] + 2\mu_A\big[\mathrm{Var}_{A_x}\left(\mu_{\tilde A_x}((1-\tilde c_x)f)\right)\big].\end{aligned} \tag{6.3}$$

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

$$\mu_A\big[\mathrm{Var}_{A_x}\left(\mu_{\tilde A_x}(\tilde c_x f)\right)\big] = \mu_A\big[\mathrm{Var}_x\left(\mu_{\tilde A_x}(\tilde c_x f)\right)\big] \le \mu_A\big[\tilde c_x \,\mathrm{Var}_x(f)\big].$$

Now we focus on the second term of (6.3). Note that

$$\mu_{\tilde A_x}\big[(1-\tilde c_x)f\big] = \mu_{\tilde A_x}\big[(1-\tilde c_x)\mu_{\tilde A_x\setminus\Delta_x}(f)\big].$$

Set $\delta := \sup_{x\in A}\mu(1-\tilde c_x)$. Hence, bounding the variance by the second moment and using the Cauchy−Schwarz inequality, we get

$$\begin{aligned}\mathrm{Var}_{A_x}\left(\mu_{\tilde A_x}((1-\tilde c_x)f)\right) &\le \mathrm{Var}_{A_x}\left(\mu_{\tilde A_x}\big[(1-\tilde c_x)\mu_{\tilde A_x\setminus\Delta_x}(f)\big]\right) \\ &\le \mu_{A_x}\left(\mu_{\tilde A_x}\big[(1-\tilde c_x)\mu_{\tilde A_x\setminus\Delta_x}(f)\big]^2\right) \\ &\le \delta\big(\mathrm{Var}_{A_x}\left(\mu_{\tilde A_x\setminus\Delta_x}(f)\right)\big).\end{aligned}$$

From all the previous computations (and using (6.2)) we deduce that

$$\mathrm{Var}_A(f) \le 2\sum_{x\in A}\mu_A(\tilde c_x\,\mathrm{Var}_x(f)) + 2\delta\sum_{x\in A}\mu_A\big(\mathrm{Var}_{A_x}(\mu_{\tilde A_x\setminus\Delta_x}(f))\big).$$

Hence if one proves that

$$\sum_{x\in A}\mu_A\big(\mathrm{Var}_{A_x}(\mu_{\tilde A_x\setminus\Delta_x}(f))\big) \le \sup_{y\in A}\big|\{x\in A\colon \Delta_x\cup\{x\}\ni y\}\big|\,\mathrm{Var}_A(f) \tag{6.4}$$

the result follows. We now prove (6.4). Using (6.2), we have

$$\operatorname{Var}_{A_x}(g) \le \sum_{y\in A_x} \mu_{A_x}\big(\operatorname{Var}_{A_y}(\mu_{\tilde A_y}(g))\big) = \sum_{y\in\Delta_x\cup\{x\}} \mu_{A_x}\big(\operatorname{Var}_{A_y}(\mu_{\tilde A_y}(g))\big)$$

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

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

since, by Cauchy – Schwarz,

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

This ends the proof. □

Proof of Proposition 6.2. Our aim is to apply Lemma 6.1. 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 $E_x=\big\{\sigma\in\Omega : \sum_{i=1}^n(1-\sigma(x_i))\ge 1\big\}$, 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 $|\Delta_x|\le k\ell^D$ for any $x\in A$, the assumption of Lemma 6.1 reads $p^\ell(1+k\ell^D)<1/4$ which is satisfied if one chooses $\ell=(c/q)\log(2/q)$ 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. $\xi(\sigma)=\inf\{i:\sigma(x_i)=0\}$. Then we write

$$\begin{aligned}\mu_A(\tilde c_x\operatorname{Var}_x(f)) &= \sum_{i=1}^n \mu_A\big(\tilde c_x\mathbb{1}\{\xi=i\}\operatorname{Var}_x(f)\big)\\ &= pq\sum_{i=1}^n\sum_{\sigma:\xi(\sigma)=i}\mu_A(\sigma)(f(\sigma^x)-f(\sigma))^2\end{aligned}$$

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 2 for a graphical illustration of such a path.

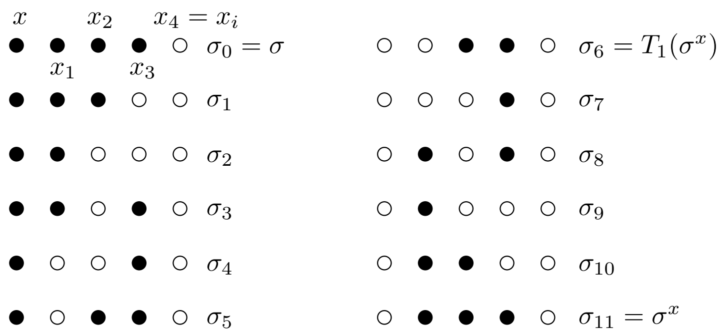

Figure 2. Illustration of the path from σ to σ^x for a configuration σ satisfying $\xi(\sigma) = x_4$. Here $i = 4$ and the length of the path is $4i-5 = 11$.

Denote by $\Gamma_x(\sigma) = \{\sigma_0, \sigma_1, \dots, \sigma_{4i-6}\}$ 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 $\eta = \sigma_{j-1}^y$. Then, by the Cauchy−Schwarz inequality,

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

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 $\mu_A(\sigma)/\mu_A(\eta) \le (p^2/q^2)\max(p/q, q/p)$, 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 from the latter that

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

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 6.1 we used the following lemma.

Lemma 6.2. *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* (5.6)*. Assume that* $\max(1-\mu(c_A), 1-\mu(c_B)) < 1/16$*. Then, for all $f : \hat\Omega_\Lambda \to \mathbb{R}$ with $\hat\mu_\Lambda(f) = 0$ it holds* $\operatorname{Var}_{\hat\mu_\Lambda}(f) \le 24\hat\mu_\Lambda\big[c_B \operatorname{Var}_{\mu_A}(\tilde f) + c_A \operatorname{Var}_{\mu_B}(\tilde f)\big]$ *where $\tilde f : \Omega_\Lambda \to \mathbb{R}$ is defined as*

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

Proof. Recalling the variational definition of the variance we have

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

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 6.3 below and obtain

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

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 6.3. *Let $\Lambda = A \cup B$ with $A, B \subset V$ satisfying $A \cap B = \emptyset$. Let μ_A and μ_B be two probability measures on $\{0,1\}^A$ and $\{0,1\}^B$ respectively,* *and* $\mu = \mu_A \otimes \mu_B$. *Take* c_A, $c_B : \{0,1\}^\Lambda \to [0,1]$ *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*

$$\begin{aligned}\operatorname{Var}_\mu(g) \le 12\mu\big[c_B^2 \operatorname{Var}_{\mu_A}(g) + c_A^2 \operatorname{Var}_{\mu_B}(g)\big]\\ + 8\max(1-\mu(c_A), 1-\mu(c_B)) \operatorname{Var}_\mu(g).\end{aligned}$$

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, $c_B \in [0,1]$,

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

We now treat the fourth term in the latter inequality.

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

If we average with respect to μ we have $\mu\big(\mu_A\big(c_A^2[g-\mu_B(g)]^2\big)\big) = \mu(c_A^2 \operatorname{Var}_{\mu_B}(g))$ and, using the Cauchy–Schwarz inequality and $x^2 \le x$ for $x \in [0,1]$,

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

so that $\mu\big(\mu_A(g)^2\big) \le 2\mu\big(c_A^2 \operatorname{Var}_{\mu_B}(g)\big) + 2(1-\mu(c_A)) \operatorname{Var}_\mu(g)$. 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 (2000) *Sur les Inégalités de Sobolev Logarithmiques.* Panoramas et Syntheses **10**, Soc. Math. de France, Paris.

[2] N. Cancrini, F. Martinelli, C. Roberto and C. Toninelli (2008) Kinetically constrained spin models. *Probab. Theory and Relat. Fields* **140** (3–4), 459–504.

[3] N. Cancrini, F. Martinelli, C. Roberto and C. Toninelli (2009) Facilitated spin models: recent and new results. In: *Methods of Contemporary Mathematical Statistical Physics*, Lect. Notes Math. **1970**, Springer, Berlin, 307–340.

[4] N. Cancrini, F. Martinelli, R. Schonmann and C. Toninelli (2010) Facilitated oriented spin models: some non equilibrium results. *J. Stat. Phys.* **138** (6), 1109–1123.

[5] P. Diaconis and L. Saloff-Coste (1996) Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** (3), 695–750.

[6] A. Faggionato, F. Martinelli, C. Roberto and C. Toninelli (2013) The East model: recent results and new progresses. *Markov Processes and Relat. Fields* **19** (3), 407–452.

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

[8] G.H. Fredrickson and H.C. Andersen (1985) Facilitated kinetic Ising models and the glass transition. *J. Chem. Phys.* **83**, 5822–5831.

[9] J.P. Garrahan, P. Sollich and C. Toninelli (2011) Dynamical heterogeneities and kinetically constrained models. In: *Dynamical Heterogeneities in Glasses, Colloids and Granular Media and Jamming Transitions*, L. Berthier, G. Biroli, J-P. Bouchaud, L. Cipelletti and W. van Saarloos (eds.), International Series of Monographs in Physics, Oxford press, 341–369.

[10] R.I. Grigorchuk (1991) On growth in group theory. In: *Proceedings of the International Congress of Mathematicians* (Kyoto, 1990), Math. Soc. Japan, Tokyo, 325–338.

[11] L. Gross (1975) Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** (4), 1061–1083.

[12] R. Holley and D. Stroock (1987) Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.* **46** (5–6), 1159–1194.

[13] S. Léonard, P. Mayer, P. Sollich, L. Berthier, and J.P. Garrahan (2007) Non-equilibrium dynamics of spin facilitated glass models. *J. Stat. Mech.: Theory and Exp.* **7**, P07017.

[14] T.M Liggett (1985) *Interacting Particle Systems.* Springer-Verlag, New York.

[15] F. Martinelli (1999) Lectures on Glauber dynamics for discrete spin models. In: *Lectures on Probability Theory and Statistics* (Saint-Flour, 1997), Springer, Berlin, 93–191.

[16] F. Martinelli and C. Toninelli (2013) Kinetically constrained models on trees. *Ann. Appl. Probab.* **23** (5), 1967–1987.

[17] F. Ritort and P. Sollich (2003) Glassy dynamics of kinetically constrained models. *Adv. Phys.* **52** (4), 219–342.

[18] L. Saloff-Coste (1997) Lectures on finite Markov chains. In: *Lectures on Probability Theory and Statistics* (Saint-Flour, 1996), J.A. Fill (ed.), Springer, Berlin, 301–413.

[19] D.W. Stroock and B. Zegarliński (1992) The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Commun. Math. Phys.* **149** (1), 175–193.