Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 68, 1–16.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2229

Exit time tails from pairwise decorrelation in hidden Markov chains, with applications to dynamical percolation

Alan Hammond* Elchanan Mossel† Gábor Pete‡

Dedicated to the memory of our friend and mentor Oded Schramm

Abstract

Consider a Markov process ω_t at stationarity and some event $\mathcal{C}$ (a subset of the state-space of the process). A natural measure of correlations in the process is the pairwise correlation $\mathbf{P}\big[\omega_0, \omega_t \in \mathcal{C}\big] - \mathbf{P}\big[\omega_0 \in \mathcal{C}\big]^2$. A second natural measure is the probability of the continual occurrence event $\big\{\omega_s \in \mathcal{C},\, \forall\, s \in [0,t]\big\}$. We show that for reversible Markov chains, and any event $\mathcal{C}$, pairwise decorrelation of the event $\mathcal{C}$ implies a decay of the probability of the continual occurrence event $\big\{\omega_s \in \mathcal{C}\, \forall\, s \in [0,t]\big\}$ as $t \to \infty$. We provide examples showing that our results are often sharp.

Our main applications are to dynamical critical percolation. Let $\mathcal{C}$ be the left-right crossing event of a large box, and let us scale time so that the expected number of changes to $\mathcal{C}$ is order 1 in unit time. We show that the continual connection event has superpolynomial decay. Furthermore, on the infinite lattice without any time scaling, the first exceptional time with an infinite cluster appears with an exponential tail.

Keywords: decorrelation, hidden Markov chains, hitting and exit times, spectral gap, dynamical percolation, exceptional times, scaling limits.
AMS MSC 2010: 60J25; 60K35; 82B43.
Submitted to EJP on December 26, 2011, final version accepted on May 31, 2012.

1 Introduction

We study the relationship between pairwise decorrelation of a specific event and the decay rate of the probability of continual occurrence of the event in **reversible Markov processes**. In particular, Theorem 1.1 below states that any decay of the pairwise correlations

$$\mathbf{P}[\omega_0, \omega_t \in A] - \mathbf{P}[\omega_0 \in A]^2$$

*Department of Statistics, University of Oxford, http://www.stats.ox.ac.uk/~hammond/
†Departments of Statistics and Computer Science, U.C. Berkeley, and Faculty of Mathematics and Computer Science, Weizmann Institute of Science, http://www.stat.berkeley.edu/~mossel/
‡Institute of Mathematics, Technical University of Budapest, http://www.math.bme.hu/~gabor/

for the process $(\omega_t)_{t\geq 0}$ in stationarity implies a comparable decay of the joint probability

$$\mathbf{P}[\omega_s \in A \text{ for all } 0 \leq s \leq t]\,.$$

Given a Markov process ω_t on S with stationary probability measure π, its time 1 Markov operator $(T_1 f)(\omega) := \mathbf{E}[f(\omega_1) \mid \omega_0 = \omega]$ on $f \in L^2(S,\pi)$ is a normal operator, hence it has a spectral decomposition, with $\mathsf{Spec}(T_1) \subseteq \{z : |z| \leq 1\} \subset \mathbb{C}$ and $z = 1$ being an obvious eigenvalue. Its **spectral gap** is defined as $g := \inf\big\{|1-\lambda| : \lambda \in \mathsf{Spec}(T_1)\setminus\{1\}\big\}$, while its **absolute spectral gap** is $g_* := 1 - \sup\big\{|\lambda| : \lambda \in \mathsf{Spec}(T_1)\setminus\{1\}\big\}$. It is well-known (and not hard to see) that $g_* > 0$ is equivalent to having an exponential decay of correlations for **any** function $f : S \longrightarrow \mathbb{R}$ with $\mathbf{E}[f] = 0$:

$$\mathbf{E}[f(\omega_0) f(\omega_t)] \leq (1-g_*)^t\, \mathbf{E}[f^2]\,,$$

for the process at stationarity. We write $\mathbf{P}$ both for both the law π on static configurations and for the measure of the process at stationarity. Similarly, $\mathbf{E}$ denotes expectation of a function with respect to one or other of these laws, depending on whether the function is defined on static or dynamic configurations It is quite classical for this case (in fact, $g > 0$ suffices), see [1], [2], [13, Theorem 3.6] and [3, Theorem 9.2.7], that for **any** set $A \subset S$ with stationary measure bounded away from 1, the exit-time tail $\mathbf{P}[\omega_s \in A \text{ for all } 0 \leq s \leq t]$ is exponentially small in t, with an exponent depending on the spectral gap and on $\pi(A)$. The strongest such bound is the one in [3], with a generalization in [16, Theorem 5.4].

Our Theorem 1.1 is a generalization of these results for the case when we have a pairwise correlation decay not for any function, but only for being in a given A — which might happen on a much faster time-scale than the mixing time of the entire chain. In other words, our generalization concerns the **hidden Markov chain** $\mathbb{1}_{\{\omega_t \in A\}}$. Furthermore, our Theorem 1.2 gives a generalization of [16, Theorem 5.4] in a different direction, by showing that, assuming a spectral gap, the exit-time tail from A is exponentially decaying provided that the probability that ω_t is in A at every moment of a fixed time interval is bounded away from one (which may be the case even if $\pi(A)$ is arbitrarily close to 1).

The exponential exit-time tail for Markov chains with spectral gap (such as random walks on expander graphs) has many applications in computer science including derandomization of algorithms [13, Section 3] and noise sensitivity [16], suggesting that our results may prove useful from such points of view. Nevertheless, our initial motivation comes from the study of dynamical percolation on planar lattices, which is the natural time evolution of critical percolation in the plane, a central model of statistical mechanics; see [11, 4, 18, 6, 7, 8, 12] for the original papers, and [20, 9] for surveys. The implications of our results to dynamical percolation will be explained in Section 4.

We now state our main results in detail.

1.1 Exiting an event with some pairwise decorrelation

We will consider continuous or discrete time Markov processes, $(\omega_t)_{t\in\mathbb{R}}$ or $(\omega_t)_{t\in\mathbb{Z}}$, on some state space S, with some (not necessarily unique) stationary probability measure π; we will always consider the process run in stationarity, i.e., with $\omega_0 \sim \pi$. For functions $f : S \longrightarrow \mathbb{R}$, consider the usual inner product $(f,g) := \mathbf{E}[fg]$, and the Markov operator $(T_t f)(\omega) := \mathbf{E}[f(\omega_t) \mid \omega_0 = \omega]$. Let $\mathcal{C}$ be a static event (i.e., measurable with respect to ω_0), suppose that $\pi(\mathcal{C}) = \mathbf{P}\big[\omega_0 \in \mathcal{C}\big] = p$, and let $f = \mathbb{1}_{\mathcal{C}}$. The decay of correlations of f in time is often quantified by the function $d : (0,\infty) \to [0,\infty)$ in one of the following two inequalities:

$$\mathbf{P}\big[\omega_0, \omega_t \in \mathcal{C}\big] - \mathbf{P}[\omega_0 \in \mathcal{C}]^2 = (f, T_t f) - (\mathbf{E} f)^2 \leq d(t)\,\mathrm{Var}[f] \tag{1.1}$$

and

$$\mathrm{Var}[T_t f] = (T_t f, T_t f) - (\mathbf{E} f)^2 \leq d(2t)\,\mathrm{Var}[f]\,, \tag{1.2}$$

for all $t \in [0,\infty)$. Of course, for reversible Markov processes, (1.1) is equivalent to (1.2). We will consider the cases where the decay of $d(t)$ as $t \to \infty$ is either polynomial or (stretched) exponential. Sometimes, one has a sequence of Markov processes $(\omega_t^n)_{t\in\mathbb{R}}$, $n \in \mathbb{N}$, on larger and larger finite state spaces, with the time parameter coming from the original time of the process rescaled by a function of n. In this case, the bounds are understood uniformly in n.

Theorem 1.1. *In the above setting, assuming (1.2), we have that*

$$\mathbf{P}\Big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\Big] \leq \min_{k\in\mathbb{N}^+} \left\{ \left(\frac{p+1}{2}\right)^k + \frac{16p}{(1-p)^2} d\left(\frac{2t}{k}\right) \right\}, \tag{1.3}$$

and therefore

$$\mathbf{P}\Big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\Big] \leq \begin{cases} t^{-\alpha+o(1)} & \text{if } d(t) = \Theta(t^{-\alpha}),\\ \exp\big(-t^{\frac{\alpha}{1+\alpha}+o(1)}\big) & \text{if } d(t) = \exp(-\Theta(t^\alpha))\,, \end{cases}$$

as $t \to \infty$, where the $o(1)$ terms depend only on p, α and the constant factors implicit in the $\Theta(\cdot)$ notation.

Examples and questions of sharpness and of non-sharpness in Theorem 1.1 appear in Section 3.

Remark. Note that we make no assumption of reversibility in Theorem 1.1. However, as we have noted, for a reversible Markov chain, we may replace the assumption (1.2) in the statement by (1.1), which is a more familiar form in which to express decorrelation in a Markov process.

Motivation. As we mentioned above, our main motivation is dynamical critical percolation on planar lattices: site percolation on the triangular lattice or bond percolation on $\mathbb{Z}^2$. Let $\mathcal{C}$ be the left-right crossing event of a large box, and let us scale time such that the expected number of changes to $\mathcal{C}$ is order 1 in unit time. Theorem 1.1 implies that the continual connection event has superpolynomial decay. See Corollary 4.1.

1.2 Exiting events defined on time intervals, assuming a spectral gap

We consider a continuous time Markov process semi-group $T_t = e^{tQ}$ on some state space S, reversible with respect to a probability measure π. Then the infinitesimal generator Q is self-adjoint (reversible) and negative semi-definite with respect to the usual inner product given by π, and its spectrum is contained in $(-\infty, 0]$. We will assume that Q has a spectral gap $\delta > 0$ around the obvious eigenvalue 0; then T_t has an absolute spectral gap $1 - e^{-\delta t}$, and the process is ergodic.

Let Ω be the space of all paths $\omega : \mathbb{R} \longrightarrow S$ of the Markov process under the probability measure $\mathbf{P}$, and let $L^2(\Omega, \mathbf{P})$ denote all L^2 integrable functions from Ω to $\mathbb{R}$. For a subset $I \subseteq R$ we denote by $\mathcal{F}_I$ the sigma algebra generated by $\{\omega(t) : t \in I\}$.

Theorem 1.2. *Suppose that the generator Q is reversible with respect to a probability measure π and has spectral gap $\delta > 0$. Let $k \in \mathbb{N}^+$, and let $a_i, b_i \in \mathbb{R}$, $0 \leq i \leq k$ satisfy $a_i < b_i$ for such i, as well as $b_i \leq a_{i+1}$ for $0 \leq i \leq k-1$; we also permit $a_0 = -\infty$ as well as $b_k = \infty$.*

Let $\mathcal{A}_0, \ldots, \mathcal{A}_k$ be subsets of Ω, each $\mathcal{A}_i$ measurable with respect to $\mathcal{F}_{[a_i,b_i]}$. Then

$$\mathbf{P}\Big[\bigcap_{i=0}^k \mathcal{A}_i\Big] \leq \sqrt{\mathbf{P}[\mathcal{A}_0]}\sqrt{\mathbf{P}[\mathcal{A}_k]} \prod_{i=0}^{k-1} \Big[\sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]} + e^{-\delta(a_{i+1}-b_i)}\Big(1 - \sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]}\Big)\Big].$$

In particular, suppose that $\mathcal{A}$ is measurable with respect to $\mathcal{F}_{[a,b]}$ with $\mathbf{P}[\mathcal{A}] = p$, and $[a_i, b_i] = [a + s_i, b + s_i]$. Setting $t_i = s_{i+1} - s_i - (b - a)$ for all $0 \le i < k$, then, provided that $t_i \ge 0$ for all such i,

$$\mathbf{P}\Big[\omega(t+s_i)_{t\in[a,b]} \in \mathcal{A} \text{ for } i = 0, 1, \ldots, k\Big] \le p \prod_{i=0}^{k-1} \big(p + e^{-\delta t_i}(1-p)\big).$$

Motivation. As we will show in Corollary 4.4, this theorem implies for dynamical percolation on the infinite lattice that the probability that there is no **exceptional time** in $[0, t]$ with the cluster of the origin being infinite is exponentially small. This corollary plays a significant role in [12].

Acknowledgments. We each benefitted greatly from conversations with Oded Schramm while working on this project and many others. We feel very fortunate to have known and to have worked with him.

We also thank Christophe Garban and Yuval Peres for useful conversations and remarks.

This work was started in the summer of 2007 at the Theory Group of Microsoft Research, Redmond. We are also grateful to the Fields Institute in Toronto, where some of the work by AH and GP was done in 2011. AH was supported principally by the EPSRC grant EP/I004378/1. EM was supported by NSF awards DMS 0548249 and DMS 1106999, DOD ONR grant N000141110140, ISF award 1300/08 and a Minerva award. GP was supported by an NSERC Discovery Grant at the University of Toronto, and an EU Marie Curie International Incoming Fellowship at the Technical University of Budapest.

2 Proofs

2.1 Exiting an event with some pairwise decorrelation

We now prove Theorem 1.1.

Proof. Let $p < \lambda < 1$. Consider the static event

$$A_s := \Big\{\omega \in S : \mathbf{P}[\omega_s \in \mathcal{C} \mid \omega_0 = \omega] < \lambda\Big\}.$$

Note that for $s = 0$ we have $\mathbf{P}\big[A_0^c\big] = p$, while for large s one expects $\mathbf{P}\big[A_s^c\big]$ to be small.

Let $\tau = t/k$, for some $k \in \mathbb{Z}_+$. We claim that, for $m \ge 0$,

$$\mathbf{P}\big[\omega_0 \in A_\tau^c \cap \mathcal{C};\ \omega_{j\tau} \in A_\tau \cap \mathcal{C} \text{ for } 1 \le j \le m\big] \le \lambda^{(m-1)\vee 0}\, \mathbf{P}\big[\omega_0 \in A_\tau^c\big]. \tag{2.1}$$

We may prove this by induction on m, the cases where $m \in \{0, 1\}$ being trivial. For $m \ge 2$, writing B_m for the event on the left-hand side of (2.1), we have that $\mathbf{P}\big[B_m\big] = \mathbf{P}\big[B_{m-1}\big] q_m$, where q_m is the conditional probability of $\omega_{m\tau} \in A_\tau \cap \mathcal{C}$ given B_{m-1}. Note that q_m is at most $\mathbf{P}\big[\omega_{m\tau} \in \mathcal{C} \bigm| B_{m-1}\big]$. The conditional distribution of $\omega_{(m-1)\tau}$ given B_{m-1} being supported on the event $\omega_{(m-1)\tau} \in A_\tau$, it follows from the Markov property that $\mathbf{P}\big[\omega_{m\tau} \in \mathcal{C} \bigm| B_{m-1}\big] \le \lambda$. Hence, the inductive hypothesis at $m - 1$ implies this statement at m, giving (2.1).

By the same argument, we see that, for each $m \ge 0$,

$$\mathbf{P}\big[\omega_{j\tau} \in A_\tau \cap \mathcal{C} \text{ for } 0 \le j \le m\big] \le \lambda^m. \tag{2.2}$$

We find then that

$$\begin{aligned}\mathbf{P}\Big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\Big] &\le \mathbf{P}\big[\omega_{j\tau} \in \mathcal{C} \text{ for } 0 \le j \le k\big] \\ &\le \mathbf{P}\big[\omega_{j\tau} \in \mathcal{C} \cap A_\tau \text{ for } 0 \le j \le k\big] + \\ &\quad \sum_{\ell=0}^{k} \mathbf{P}\big[\omega_{\ell\tau} \in A_\tau^c \cap \mathcal{C};\ \omega_{j\tau} \in A_\tau \cap \mathcal{C} \text{ for } \ell < j \le k\big] \\ &\le \lambda^k + \sum_{\ell=0}^{k} \lambda^{(k-\ell-1)\vee 0}\, \mathbf{P}\big[\omega_{\ell\tau} \in A_\tau^c\big] \\ &\le \lambda^k + \frac{2-\lambda}{1-\lambda}\, \mathbf{P}[A_\tau^c]\,. \end{aligned} \tag{2.3}$$

In the third inequality, (2.2) was used to bound the first term on its left-hand side, while the summand was bounded using (2.1) with $m = k - \ell$ and stationarity.

We need to find now an upper bound on $\mathbf{P}\big[A_s^c\big]$ for s large. By the definition of A_s,

$$\mathbf{E}\big[\mathbb{1}_{A_s^c}\, T_s f\big] = \mathbf{E}\big[f(\omega_s) \bigm| A_s^c\big]\, \mathbf{P}\big[A_s^c\big] \ge \lambda\, \mathbf{P}\big[A_s^c\big]\,,$$

where, as before, $f = \mathbb{1}_{\mathcal{C}}$. On the other hand,

$$\mathbf{E}\big[\mathbb{1}_{A_s^c}\, T_s f\big] = \mathbf{E}\big[\mathbb{1}_{A_s^c} p\big] + \mathbf{E}\Big[\mathbb{1}_{A_s^c}\, (T_s f - \mathbf{E} f)\Big]\,.$$

Putting these two things together,

$$\mathbf{E}\Big[\mathbb{1}_{A_s^c}\, (T_s f - \mathbf{E} f)\Big] \ge (\lambda - p)\, \mathbf{P}\big[A_s^c\big]\,.$$

Applying Cauchy-Schwarz to the left-hand side,

$$\|\mathbb{1}_{A_s^c}\|_2\, \|T_s f - \mathbf{E} f\|_2 \ge (\lambda - p)\, \mathbf{P}\big[A_s^c\big]\,,$$

so that we obtain

$$\mathrm{Var}[T_s f]^{1/2} = \|T_s f - \mathbf{E} f\|_2 \ge (\lambda - p)\, \mathbf{P}\big[A_s^c\big]^{1/2}\,.$$

Therefore, (1.2) implies that

$$\mathbf{P}\big[A_s^c\big] \le \frac{p - p^2}{(\lambda - p)^2}\, d(2s). \tag{2.4}$$

Now take $s = \tau$, where recall that $\tau = t/k$ for some $k \in \mathbb{Z}_+$ that we will shortly specify. Plugging (2.4) into (2.3), we find that

$$\mathbf{P}\Big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\Big] \le \lambda^k + \frac{p - p^2}{(\lambda - p)^2}\, \frac{2-\lambda}{1-\lambda}\, d(2t/k)\,.$$

Setting $\lambda = (p+1)/2$ yields (1.3).

For the case $d(t) = t^{-\alpha}$, setting $k = \lfloor K \log t \rfloor$ for a suitable constant $K = K(\lambda, \alpha)$ makes both terms $t^{-\alpha}$, as desired. For the case $d(t) = \exp(-t^\alpha)$, we set $k = \lfloor t^\beta \rfloor$, and optimize the upper bound by letting $\alpha(1-\beta) = \beta$, i.e., choosing $\beta = \alpha/(1+\alpha)$, and we are done. □

2.2 Exiting events defined on time intervals, assuming a spectral gap

In this subsection we prove Theorem 1.2. There are two main ideas. The first is that if a chain has a spectral gap, then the associated Markov operator will be a strict L^2-contraction not only on functions with zero mean, but also on any function whose support has a stationary measure bounded away from 1. The second idea is to use conditional expectation and the Markov property to extend the first idea to functions defined not on the state space, i.e., at individual times, but on time intervals. These two ideas are formalized in the following lemma.

Lemma 2.1. *Suppose that the infinitesimal generator Q is reversible with respect to a probability measure π and has spectral gap $\delta > 0$. Let $\mathcal{A}_1, \mathcal{A}_2$ be subsets of Ω that are measurable with respect to $\mathcal{F}_{\le a}$ and $\mathcal{F}_{\ge a+t}$, respectively. Let P_1 and P_2 be the corresponding projection operators on $L^2(\Omega, \mathbf{P})$, i.e., $P_i f(\omega) = f(\omega)\mathbb{1}_{\mathcal{A}_i}(\omega)$ for every function f on Ω. Then*

$$\|P_1P_2\| \le \sqrt{\mathbf{P}[\mathcal{A}_1]}\sqrt{\mathbf{P}[\mathcal{A}_2]} + e^{-\delta t}\left(1 - \sqrt{\mathbf{P}[\mathcal{A}_1]}\sqrt{\mathbf{P}[\mathcal{A}_2]}\right),$$

where the norm on the left is the operator norm for operators from $L^2(\Omega, \mathbf{P})$ into itself.

Proof. Since P_1 and P_2 are self-adjoint and commuting P_1P_2 is also self-adjoint. By the definition of norm, duality and the positivity of P_i,

$$\begin{aligned}\|P_1P_2\| &= \sup\left\{(P_1P_2f_2, f_1) : \|f_i\|_2 = 1, f_i \ge 0\right\}\\ &= \sup\left\{(P_2f_2, P_1f_1) : \|f_i\|_2 = 1, f_i \ge 0\right\}.\end{aligned}$$

For f_1, f_2 such that $P_1f_1 \ne 0, P_2f_2 \ne 0$, define

$$f_1' = \frac{P_1f_1}{\|P_1f\|_2}, \quad f_2' = \frac{P_2f_2}{\|P_2f_2\|_2},$$

so $\|f_i'\|_2 = 1$ and $\mathsf{supp}(f_i) \subseteq \mathcal{A}_i$, for $i = 1, 2$. Since $\|P_if_i\| \le 1$ and since P_1, P_2 are idempotent, we have

$$(P_2f_2, P_1f_1) \le (f_2', f_1') = (P_2f_2', P_1f_1').$$

Therefore,

$$\|P_1P_2\| = \sup\Big\{\mathbf{E}[f_1f_2] : \|f_i\|_2 = 1,\ f_i \ge 0,\ \mathsf{supp}(f_i) \subseteq \mathcal{A}_i \text{ for } i = 1, 2\Big\}.$$

Given f_1 and f_2 as in the last equation, let $g_i : S \longrightarrow \mathbb{R}$ be defined by $g_1(x) := \mathbf{E}[f_1(\omega) \mid \omega(a) = x]$ and $g_2(x) := \mathbf{E}[f_2(\omega) \mid \omega(a+t) = x]$. Clearly, $\mathbf{E}[g_i] = \mathbf{E}[f_i]$ and $\mathbf{E}[g_i^2] \le \mathbf{E}[f_i^2] = 1$. Since f_1 is $\mathcal{F}_{\le a}$ measurable and f_2 is $\mathcal{F}_{\ge a+t}$ measurable, by the Markov property it follows that

$$\mathbf{E}[f_1f_2] = \mathbf{E}\big[\mathbf{E}[f_1f_2 \mid \omega(a), \omega(a+t)]\big] = \mathbf{E}\big[g_1(\omega(a))\, g_2(\omega(a+t))\big] = \mathbf{E}[g_1T_tg_2].$$

Using the fact that the spectral gap of T_t is $1 - e^{-\delta t}$, it follows from Lemma 2.2 below that the last expression is bounded by

$$\mathbf{E}[g_1]\mathbf{E}[g_2] + e^{-\delta t}(1 - \mathbf{E}[g_1]\mathbf{E}[g_2]) = \mathbf{E}[f_1]\mathbf{E}[f_2] + e^{-\delta t}(1 - \mathbf{E}[f_1]\mathbf{E}[f_2]),$$

and using $\mathbf{E}[f_i] = \mathbf{E}[f_i\, \mathbb{1}_{\mathcal{A}_i}] \le \|f_i\|_2\|\mathbb{1}_{\mathcal{A}_i}\|_2 \le \sqrt{\mathbf{P}[\mathcal{A}_i]}$ we obtain that

$$\|P_1P_2\| \le \sqrt{\mathbf{P}[\mathcal{A}_1]}\sqrt{\mathbf{P}[\mathcal{A}_2]} + e^{-\delta t}\left(1 - \sqrt{\mathbf{P}[\mathcal{A}_1]}\sqrt{\mathbf{P}[\mathcal{A}_2]}\right),$$

as stated. □

The proof of the previous lemma used the following easy fact.

Lemma 2.2. *Let M be an ergodic transition matrix for a Markov chain on the set S which is reversible with respect to the probability measure π and which has spectral gap $\delta > 0$. Let $g_1, g_2 : S \longrightarrow \mathbb{R}$ be two functions with L^2 norm at most one. Then*

$$\mathbf{E}[g_1 M g_2] \leq \mathbf{E}[g_1]\mathbf{E}[g_2] + (1-\delta)\big(1 - \big|\mathbf{E}[g_1]\mathbf{E}[g_2]\big|\big).$$

Proof. Abbreviating $\mathbb{1} = \mathbb{1}_S$, we set $h_i = g_i - \mathbf{E}[g_i]\mathbb{1}$. Then h_i is orthogonal to the constant functions and $\mathbb{1}$ is a 1-eigenvector of M. Therefore,

$$\mathbf{E}[g_1 M g_2] = \mathbf{E}[g_1]\mathbf{E}[g_2] + \mathbf{E}[h_1 M h_2].$$

Using the spectral gap of M, we get:

$$\begin{aligned}\mathbf{E}[h_1 M h_2] \leq \|h_1\|_2 \|M h_2\|_2 \leq (1-\delta)\|h_1\|_2\|h_2\|_2 &\leq (1-\delta)\sqrt{(1-\mathbf{E}[g_1]^2)(1-\mathbf{E}[g_2]^2)} \\ &\leq (1-\delta)\big(1 - \big|\mathbf{E}[g_1]\mathbf{E}[g_2]\big|\big),\end{aligned}$$

where the last inequality follows from the inequality $(1-x^2)(1-y^2) \leq (1-xy)^2$, valid for all $x, y \in \mathbb{R}$, and we are done. □

Proof of Theorem 1.2. Let P_i denote the projection onto $\mathcal{A}_i$, as in Lemma 2.1. It is easy to see that

$$\mathbf{P}\Big[\bigcap_{i=0}^{k} \mathcal{A}_i\Big] = \mathbf{E}\Big[\mathbb{1}\,\big(\prod_{i=0}^{k} P_i\big)\,\mathbb{1}\Big] = \mathbf{E}\Big[\mathbb{1}\,\big(\prod_{i=0}^{k} P_i^2\big)\,\mathbb{1}\Big],$$

since the projection P_i satisfies $P_i^2 = P_i$, and the order in which the projections act does not matter. These operators are also self-adjoint, so that

$$\begin{aligned}\mathbf{E}\Big[\mathbb{1}\,\big(\prod_{i=0}^{k} P_i^2\big)\,\mathbb{1}\Big] = \mathbf{E}\Big[(P_k\mathbb{1})\,P_k\,\big(\prod_{i=1}^{k-1} P_i^2\big)\,P_0\,(P_0\mathbb{1})\Big] &\leq \|\mathbb{1}_{\mathcal{A}_0}\|_2\,\|\mathbb{1}_{\mathcal{A}_k}\|_2\,\Big\|P_k\,\big(\prod_{i=1}^{k-1} P_i^2\big)\,P_0\Big\| \\ &\leq \sqrt{\mathbf{P}[\mathcal{A}_0]}\sqrt{\mathbf{P}[\mathcal{A}_k]}\,\prod_{i=0}^{k-1}\|P_i P_{i+1}\|,\end{aligned}$$

where, in the rightmost expressions, the new notation denotes the operator norm from $L^2(\Omega, \mathbf{P})$ to itself. By Lemma 2.1, we have that, for each $i \in \{0, \ldots, k-1\}$,

$$\|P_i P_{i+1}\| \leq \sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]} + e^{-\delta(a_{i+1}-b_i)}\left(1 - \sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]}\right).$$

Hence,

$$\prod_{i=0}^{k-1}\|P_i P_{i+1}\| \leq \prod_{i=0}^{k-1}\left[\sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]} + e^{-\delta(a_{i+1}-b_i)}\left(1 - \sqrt{\mathbf{P}[\mathcal{A}_i]}\sqrt{\mathbf{P}[\mathcal{A}_{i+1}]}\right)\right],$$

and the proof is complete. □

3 Examples concerning Theorem 1.1

3.1 An example where Theorem 1.1 is sharp

We give our example (of a reversible Markov chain) in terms of conductances on edges; see [15, Chapter 2] for an exposition of the relevant theory. Let $\beta \in (1, 2)$. Consider the graph with vertex set $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ and edge set given by the nearest-neighbour edges among these vertices, plus the edge $(-1, 1)$, and a self-loop on each vertex. Equip the edges with conductances $c_{n,n+1} = c_{-(n+1),-n} := n^{-\beta}$ for each $n \geq 1$,

$c_{-1,1} = c_{1,1} = c_{-1,-1} := 1/2$, and $c_{n,n} := c_{n,n-1} + c_{n,n+1}$ for $|n| \geq 2$. Consider the discrete-time random walk Y on this graph equipped with this set of conductances. The sum of the conductances being finite, Y has a finite invariant measure π, whose value at a vertex is the sum of the conductances over incident edges; this is the unique invariant measure given its total mass.

Let $\mathbf{P}$ denote the law of Y run in stationarity. Set $\mathcal{C} = \mathbb{N}^+$, hence $\frac{\pi(\mathcal{C})}{\pi(\mathbb{Z}^*)} = 1/2$. We will show that

$$\mathbf{P}\big[Y_s \in \mathcal{C} \ \forall s \in [0,t]\big] = t^{\frac{1-\beta}{2}+o(1)} \tag{3.1}$$

and that

$$\mathbf{P}\big[Y_0, Y_t \in \mathcal{C}\big] - \mathbf{P}\big[Y_0 \in \mathcal{C}\big]^2 \leq \tfrac{1}{2}\mathbf{P}\big[Y_s \in \mathcal{C} \ \forall s \in [0,t]\big]\,; \tag{3.2}$$

moreover, for any $\epsilon > 0$, and for all $t > 0$ sufficiently high,

$$\mathbf{P}\big[Y_0, Y_t \in \mathcal{C}\big] - \mathbf{P}\big[Y_0 \in \mathcal{C}\big]^2 \geq \big(\tfrac{1}{2} - \epsilon\big)\mathbf{P}\big[Y_s \in \mathcal{C} \ \forall s \in [0,t]\big]\,. \tag{3.3}$$

These show that Theorem 1.1 is basically sharp in the regime of polynomial decay; in fact, (3.3) is a stronger bound than what follows from Theorem 1.1. As we will see from the proof of (3.2) and (3.3), this is really an example where not leaving the set $\mathcal{C}$ at all is "responsible" for almost all of the correlation between $\{Y_0 \in \mathcal{C}\}$ and $\{Y_t \in \mathcal{C}\}$.

We first prove (3.1), which might already be in the literature somewhere, but we could not locate a reference. We start with the lower bound. In essence, the bound holds because Y has probability of order $t^{(1-\beta)/2}$ to begin at a site of order at least $t^{1/2}$ from the vertex 1. From such a site on the positive half-line, the walk has positive probability to remain positive for t steps: indeed, at such sites the walk experiences an excess in leftward transition probability over rightward of order $t^{-1/2}$, so that, during t steps, this imbalance provides a drift towards the origin totalling an order of $t^{1/2}$ steps. This drift is thus comparable to the Gaussian fluctuation of the particle during this period, and the particle remains to the right of the origin with positive probability. Rather than make this heuristic rigorous, we prove the lower bound in (3.1) by invoking the Carne-Varopoulos bound (see [5, 21] or [15, Theorem 13.4]) which, in the case of a reversible Markov chain X having finite stationary measure (and hence spectral radius 1), asserts that

$$\mathbf{P}\big[X(s) = y \;\big|\; X(0) = x\big] \leq 2\sqrt{\frac{\pi(y)}{\pi(x)}}\exp\left\{-\frac{d(x,y)^2}{2s}\right\} \tag{3.4}$$

for all $x, y \in S$, $s > 0$; here, $d(\cdot,\cdot)$ denotes graphical distance on S, and π denotes the stationary measure. We apply this bound to the walk Y. Noting that $\frac{\pi(x)}{\pi(1)} \leq Cx^{-\beta}$ for $x \in \mathbb{N}$, we find that, if $s \in \{0,\ldots,t\}$, and $x \in \mathbb{N}$ satisfies $t^{1/2}\sqrt{2}(\log t)^{1/2} \leq x \leq t^{1/2}\sqrt{2}(\log t)^{1/2} + t^{1/2}$,

$$\mathbf{P}\big[Y(s) = 1 \;\big|\; Y(0) = x\big] \leq C\big(t^{1/2}\sqrt{2}(\log t)^{1/2} + t^{1/2}\big)^{\beta/2}\exp\big\{-2\log t\big\}.$$

Sum this bound over $s \in \{0,\ldots,t\}$ to arrive at

$$\mathbf{P}\big[\exists s \in \{0,\ldots,t\} : Y(s) = 1 \;\big|\; Y(0) = x\big] \leq Ct^{\beta/4-1}\big(\sqrt{2}(\log t)^{1/2} + 1\big)^{\beta/2}.$$

Recalling that $\beta < 2$, we find that, for t high enough, the conditional probability, given that $Y(0)$ assumes any one of the $t^{1/2}$ values of x described above, that Y reaches 1 before time t, is at most one-half. The probability that $Y(0)$ assumes some such value is at least $c\big(t^{1/2}\sqrt{2}(\log t)^{1/2} + t^{1/2}\big)^{-\beta}t^{1/2} = ct^{(1-\beta)/2}\big(\sqrt{2}(\log t)^{1/2} + 1\big)^{-\beta}$. Thus,

$$\mathbf{P}\big[Y(s) \geq 1 \,\forall s \in \{0,\ldots,t\}\big] \geq \tfrac{c}{2}t^{(1-\beta)/2}\big(\sqrt{2}(\log t)^{1/2} + 1\big)^{-\beta},$$

so that the lower bound in (3.1) is verified.

We turn to the upper bound in (3.1). Let Z denote the Markov chain on $\mathbb{Z}$ which shares its initial distribution with Y and which evolves as a simple random walk. The two processes may be coupled so that, should $Y(0)$ be positive, then $Y(m) \leq Z(m)$ for all m at most the hitting time of 1 by Y; for this reason, it suffices to establish the upper bound in (3.1) for the process Z.

Let $a, t \in \mathbb{N}$. Let $Z_a : \{0, \ldots, t\} \longrightarrow \mathbb{N}$ denote simple random walk with $Z_a(0) = a$. By [14, Theorem 2.17], $\mathbf{P}\big[\exists s \in \{0, \ldots, t\} : Z_a(s) = 0\big] \leq 12 a t^{-1/2}$. Thus,

$$\mathbf{P}\big[Z(s) \geq 1 \ \forall\ 0 \leq s \leq t \ \big|\ Z(0) = a\big] \leq 12 a t^{-1/2} . \tag{3.5}$$

Multiplying the inequality resulting from (3.5) by $\mathbf{P}\big[Z(0) = a\big]$, we sum over $a \in \mathbb{N}$ to obtain

$$\mathbf{P}\big[Z(s) \geq 1 \ \forall\ 0 \leq s \leq t\big] \leq 12 \sum_{a=1}^{t^{1/2}} \mathbf{P}\big[Z(0) = a\big] a t^{-1/2} \ + \ \mathbf{P}\big[Z(0) > t^{1/2}\big] . \tag{3.6}$$

Note that there exists $C > 0$ such that $\mathbf{P}\big[Z(0) = a\big] \leq C a^{-\beta}$ for $a \in \mathbb{N}$; thus, $\beta \in (1,2)$ implies that each of the two terms on the right-hand side of (3.6) is at most a constant multiple of $t^{(1-\beta)/2}$. That is, the upper bound in (3.1) holds for Z, as we sought to show. This completes the proof of (3.1).

We now show (3.2). Let U be the first time that Y makes the jump $(-1,1)$, and V be the first time that Y makes any of the three jumps $(-1,1)$, $(1,1)$, $(-1,-1)$. Obviously, $U \geq V$. The point of considering V is that $c_{-1,1} = c_{1,1} = c_{-1,-1}$ implies that $\mathbf{P}[Y_V \in \mathcal{C} \mid Y_{[0,V)}] = 1/2$, and then the symmetry of the entire chain and the strong Markov property implies that $\mathbf{P}[Y_t \in \mathcal{C} \mid V \leq t,\ Y_{[0,V)}] = 1/2$, as well. On the other hand, $\big\{Y_s \in \mathcal{C}\ \forall s \in [0,t]\big\} = \big\{Y_0 \in \mathcal{C},\ U > t\big\} \supseteq \big\{Y_0 \in \mathcal{C},\ V > t\big\}$, and hence $\mathbf{P}\big[Y_t \in \mathcal{C} \ \big|\ Y_0 \in \mathcal{C},\ V > t\big] = 1$. Therefore,

$$\begin{aligned}
\mathbf{P}\big[Y_0, Y_t \in \mathcal{C}\big] &= \mathbf{P}\big[Y_0 \in \mathcal{C},\ V \leq t\big] \cdot \mathbf{P}\big[Y_t \in \mathcal{C} \ \big|\ Y_0 \in \mathcal{C},\ V \leq t\big] \\
&\quad + \mathbf{P}\big[Y_0 \in \mathcal{C},\ V > t\big] \cdot \mathbf{P}\big[Y_t \in \mathcal{C} \ \big|\ Y_0 \in \mathcal{C},\ V > t\big] \\
&= \big(\mathbf{P}\big[Y_0 \in \mathcal{C}\big] - \mathbf{P}\big[Y_0 \in \mathcal{C},\ V > t\big]\big) \cdot 1/2 + \mathbf{P}\big[Y_0 \in \mathcal{C},\ V > t\big] \cdot 1 \\
&\leq \frac{1}{2} \mathbf{P}\big[Y_0 \in \mathcal{C}\big] + \frac{1}{2} \mathbf{P}\big[Y_0 \in \mathcal{C},\ U > t\big] .
\end{aligned} \tag{3.7}$$

Using that $\mathbf{P}[Y_t \in \mathcal{C}] = \pi(\mathcal{C}) = 1/2$ for all $t \geq 0$, we get (3.2).

By the second equality in (3.7), $\mathbf{P}\big[Y_0, Y_t \in \mathcal{C}\big] = \frac{1}{4} + \frac{1}{2}\mathbf{P}\big[Y_0 \in \mathcal{C},\ V > t\big]$. The bound (3.3) thus follows from the claim that for each $\epsilon > 0$, and for all $t > 0$ sufficiently high,

$$\mathbf{P}\big[Y_0 \in \mathcal{C},\ V > t\big] \geq \big(1 - \epsilon\big) \mathbf{P}\big[Y_0 \in \mathcal{C},\ U > t\big]. \tag{3.8}$$

The event on the right-hand side is simply $\{Y_s \geq 1\ \forall s \in [0,t]\}$; the event on the left-hand side contains the event $\{Y_s \geq 2\ \forall s \in [0,t]\}$. Hence, it is enough to argue that the conditional probability of the latter event given the former tends to one in a limit of high t. The event that $Y_s \geq 1$ for all $s \in [0,t]$ and $Y_s = 1$ for some such s entails either that Y remains positive for time $t/2$ after first reaching 1 after time 0, or that the same holds for the reversed chain, with time running backwards from t; either of these events has probability at most $C t^{-1/2}$ by (3.5) applied for $a = 1$. However, the event $Y_s \geq 1\ \forall s \in [0,t]$ has probability at least $c t^{(1-\beta)/2 + o(1)}$ by the lower bound in (3.1); this is much more probable under our hypothesis that $\beta < 2$, so that, given that Z is strictly positive on $\{0, \ldots, t\}$, the conditional probability that Z visits 1 during this interval tends to zero in high t. In this way, we obtain (3.8) and thus (3.3).

3.2 Some cases of non-sharpness

We first give an example where Theorem 1.1 is not at all sharp. Consider the process Y as above, and take $\mathcal{C}$ to be the set of even positive integers. The conductances on the self-loops are set in such a way that $\mathbf{P}[Y_{t+1} \in 2\mathbb{Z} \mid Y_t] = 1/2$, regardless of Y_t. Therefore, $\pi(2\mathbb{Z}) = 1/2$ and $\pi(\mathcal{C}) = 1/4$, and, using (3.2) and (3.3), then (3.1), we find that, for $t \geq 1$,

$$\begin{aligned}\mathbf{P}\big[Y_t \in \mathcal{C} \bigm| Y_0 \in \mathcal{C}\big] &= \frac{1}{2}\,\mathbf{P}\big[Y_{t-1} > 0 \bigm| Y_0 > 0\big] \\ &= \frac{1}{2}\Big(1/2 + \Theta(1)\,\mathbf{P}\big[Y_s > 0\ \forall\, 0 \leq s \leq t-1 \bigm| Y_0 > 0\big]\Big) \\ &= \frac{1}{4} + t^{\frac{1-\beta}{2}+o(1)}\,.\end{aligned}$$

Hence the correlation is polynomially large. On the other hand, $\mathbf{P}\big[Y_s \in \mathcal{C}\ \forall\, 0 \leq s \leq t\big]$ is clearly exponentially small.

A more complicated but more natural example is given by Corollary 4.1 below.

The previous subsection showed that Theorem 1.1 is sharp in the regime of polynomial decay. Examining the proof of the theorem, we have the feeling that this is not the case in the regime of superpolynomial decay. In particular, we have the following question.

Question 3.1. *Does the exponential pairwise decorrelation $d(t) = C\exp(-ct)$ in (1.2) for some event $\mathcal{C}$ imply an exit time exponential decay $\mathbf{P}\big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\big] < C'\exp(-c't)$, with c', C' depending only on c, C and $\mathbf{P}[\mathcal{C}] = p$?*

4 Applications to dynamical percolation

Critical planar percolation is a central object of probability theory and statistical mechanics; see [10, 22] for background. The best understood example is Bernoulli$(1/2)$ site percolation on the triangular lattice, where the existence of a conformally invariant scaling limit is known. Roughly, if we consider percolation on the lattice of mesh $1/n$, and any collection $\mathcal{Q}_1, \dots, \mathcal{Q}_k$ of conformal images of rectangles, then the joint distribution of the left-right crossing events inside these $\mathcal{Q}_i$'s has a limit that is conformally invariant. Moreover, one can define a continuum random limit object encoding all the macroscopic crossing events. See [17] and the explanations and references there. In **dynamical percolation**, every site is switching between being open and closed according to an independent exponential clock, in such a way that the stationary distribution on $\{0,1\}^{V_n}$ is critical percolation, where V_n is the set of sites, and 0 represents "closed" and 1 represents "open" . This model has been studied from three closely related points of view:

(1) *How long does it take to change macroscopic crossings? Or, how* **noise sensitive** *are the crossing events?* A reasonable guess is that this time-scale is given by the expected number of pivotal switches in the unit square (i.e., changes of the left-right crossing event) being of order one. Let Piv(n) be the expected number of sites in critical percolation in the unit square with mesh size $1/n$ that are pivotal for the left-right crossing; it is known for the triangular lattice that Piv$(n) = n^{3/4+o(1)}$ [19]. Then, in the stationary process, using Fubini's theorem and the linearity of expectation, the above time-scale is simply $n^2/\mathsf{Piv}(n) = n^{5/4+o(1)}$. This guess, based merely on the expectation, has been confirmed by [4, 18, 6]: if $t\,n^2/\mathsf{Piv}(n) = t\,n^{5/4+o(1)}$ sites are resampled in the unit square, then the correlation of crossing before and after the resampling is $t^{-2/3}$, up to constant factors, as

$t \to \infty$ [6, Eq. (8.7)]. Similar, though slightly weaker, results have been proved for general conformal rectangles $\mathcal{Q}$ with piecewise smooth boundary. Furthermore, even for critical (i.e., $\mathsf{Bernoulli}(1/2)$) bond percolation on $\mathbb{Z}^2$, where the existence of critical exponents such as that describing the growth of $\mathsf{Piv}(n)$ are not known, it follows from the proof of [6, Corollary 1.2], together with Eq. (2.6) there, that, after resampling $t\, n^2/\mathsf{Piv}(n)$ edges, the correlation is at most $O(t^{-\alpha})$ for some $\alpha > 0$.

(2) *On an infinite lattice, are there random times with exceptional behavior, e.g., with an infinite cluster? In other words, which events are* **dynamically sensitive**? It was proved in [18] that there are exceptional times with an infinite cluster, and in [6] that their Hausdorff dimension is almost surely 31/36, and that such exceptional times also exist for critical dynamical bond percolation on $\mathbb{Z}^2$. A natural law on the infinite cluster that appears at exceptional times will be introduced and studied in [12].

(3) *In the unit square (or in another conformal rectangle), with mesh $1/n$ and a well-chosen rate $r(n)$ for the exponential clocks, is there a* **scaling limit of the process***, giving a Markov process on continuum configurations?* If we choose $r(n) = 1/\mathsf{Piv}(n) = n^{-3/4+o(1)}$, then the expected number of pivotal switches in the unit square during a unit time will be exactly 1, independently of n. It is proved in [7, 8] that, with this scaling, such a scaling limit does indeed exist on the triangular lattice. Additionally, it follows from the results of [6], item (1) above, that the resulting Markov chain is ergodic; in particular, the correlation decay for the unit square, in rescaled large time t, is $t^{-2/3}$, up to constant factors.

We will have an application of Theorem 1.1 to the setup of items (1) and (3), and an application of Theorem 1.2 to the setup of item (2). Here is the first of these results:

Corollary 4.1. *In dynamical critical site percolation on the triangular lattice or bond percolation on $\mathbb{Z}^2$, with mesh $1/n$ and rate $1/\mathsf{Piv}(n)$ for the clocks, consider the left-right crossing event $\mathcal{C}$ in the unit square. There exist constants $\{C_K : K \in \mathbb{N}\}$ such that, for each $K \in \mathbb{N}$ and for all $t > 0$ and $n \in \mathbb{N}$,*

$$(1/4)^{\lceil 2t \rceil} \le \mathbf{P}\big[\omega_s \in \mathcal{C}\, \forall\, s \in [0,t]\big] \le C_K t^{-K}.$$

On the triangular lattice, it is known that $\mathsf{Piv}(n) = n^{3/4+o(1)}$, and the above bounds in t also hold for the scaling limit of dynamical percolation.

Before starting the proof, let us emphasize that this corollary concerns a natural question that has exactly the kind of setup for which Theorem 1.1 is designed. Namely, in the finite n version with the discrete-time chain (with sites being resampled one-by-one), the mixing time of the entire chain is $n^{2+o(1)}$ steps (it is just random walk on an n^2-dimensional hypercube), while the left-right crossing event $\mathcal{C}$ decorrelates on the scale $n^{5/4+o(1)}$, as mentioned in item (1). In the scaling limit of the chain, only the evolution of *macroscopic* crossing events is considered, so that $n^{5/4+o(1)}$ is the natural scaling factor needed to obtain this scaling limit. In particular, we are interested in the tail probability of exiting $\mathcal{C}$ on this time scale, a question for which analysis based on the spectral gap of the entire chain would clearly be too crude. Moreover, it turns out that the limit chain does not have a spectral gap (something which is clear from the polynomial decorrelation $t^{-2/3}$), hence the classical exponential exit time results [1] do not apply. One may nevertheless hope that at least there would be a "spectral gap restricted to $\mathcal{C}$", i.e., $\big(T_1(g\mathbb{1}_{\mathcal{C}}), g\mathbb{1}_{\mathcal{C}}\big) < (1-c)(g,g)$ for some $c > 0$, for all $g \in L^2(\{0,1\}^{V_n}, \mathbf{P}_{1/2})$, which, similarly to the proof of Theorem 1.2, would imply an exponentially small upper bound in Corollary 4.1. However, it is not hard to prove that $g = \mathbb{1}\{\text{density of open bits is} > 1/2 + n^{-3/4+\epsilon}\}$, with $\epsilon > 0$ fixed but small enough, is a counterexample. There are slightly more complicated counterexamples that make sense also in the scaling limit.

It is also interesting to note that Corollary 4.1, despite being a consequence of Theorem 1.1, provides a natural example in which the theorem by itself is not sharp: the correlation decay is polynomial, while the exit time tail is superpolynomial.

Proof of Corollary 4.1. We will work in the discrete lattice setting, i.e., with a fixed finite $n \in \mathbb{Z}^+$. All our results will hold uniformly in n, so that item (3) above implies that the results extend to the continuum scaling limit.

Firstly the upper bound. Let $L \in \mathbb{N}$, and decompose the unit square into L vertical slabs with dimensions $1/L \times 1$ (the induced subgraphs in the slabs will not be exactly isomorphic to each other, but this is not a problem). For $s \geq 0$ and $i \in \{1, \ldots L\}$, let $A_i(s)$ denote the event that the i^{th} such slab has an open left-right crossing at time s. We further write $\mathcal{A}_i(t)$ for the intersection of the events $A_i(s)$ over all $s \in [0,t]$. Clearly,

$$\{\omega_s \in \mathcal{C} \,\forall\, s \in [0,t]\} \subseteq \bigcap_{i=1}^{L} \mathcal{A}_i. \tag{4.1}$$

We may apply Theorem 1.1 to bound $\mathbf{P}\big[\mathcal{A}_i\big]$. To do so, we need to have a correlation decay between the events $A_i(0)$ and $A_i(t)$. Indeed,

$$\mathbf{P}\big[A_i(0), A_i(t)\big] - \mathbf{P}\big[A_i(0)\big]^2 \leq C'_L\, t^{-\alpha}$$

holds for all $t \geq 0$ and some constant C'_L, with $\alpha = 2/3$ in the case of the triangular lattice. This is simply the analogue for the slab of the decorrelation bound mentioned in item (1) above; if one does not want to optimize the constant C'_L, then the proof is identical to the one for the square case in [6, Corollary 1.2]. Noting that the variance of the left-right crossing event in any one of the slabs is an L-dependent constant, Theorem 1.1 may be applied and yields that

$$\mathbf{P}\big[\mathcal{A}_i(t)\big] \leq t^{-\alpha + o(1)}, \tag{4.2}$$

for each $i \in \{1, \ldots, L\}$, where the $o(1)$ term depends on L. We now use (4.1), (4.2) and the independence of dynamical percolation in disjoint regions to find that

$$\mathbf{P}\big[\omega_s \in \mathcal{C} \;\forall\, s \in [0,t]\big] \leq t^{-\alpha L + o(1)},$$

and the upper bound follows.

For the lower bound, we need a basic tool that we call the **dynamical FKG inequality**. The next lemma is not the strongest possible form of such a result, but it will suffice for our purposes. Firstly, we need some notation.

A realization of dynamical percolation on a finite graph G may be interpreted as a map $\omega : \mathbb{R} \longrightarrow \{0,1\}^{V(G)}$. Let $\mathcal{S}$ denote the space of such maps; we will write $\omega_s(x)$ (for $s \in \mathbb{R}$ and $x \in V(G)$) for the value at time s of ω in bit x. For static configurations, that is, elements $\eta \in \{0,1\}^{V(G)}$, we consider the natural component-wise partial order: $\eta \preceq \eta'$ iff $\eta(x) \leq \eta'(x)$ for all $x \in V(G)$. We extend this to a partial order on $\mathcal{S}$ by writing $\omega \preceq \omega'$ iff $\omega_s \preceq \omega'_s$ for all $s \in \mathbb{R}$. A function $f : \mathcal{S} \longrightarrow \mathbb{R}$ is called increasing if $f(\omega) \leq f(\omega')$ whenever $\omega \preceq \omega'$. An event $C \subseteq \mathcal{S}$ is called increasing if $\mathbb{1}_C$ is increasing.

The standard Harris-FKG inequality for percolation (see [10, Theorem 2.4]) says that increasing functions of static configurations are positively correlated. In particular, conditioning on an increasing static event makes the percolation configuration "larger". Our dynamical FKG inequality deals with conditioning on an increasing dynamical event:

Lemma 4.2. *Let $C \subseteq \mathcal{S}$ be an increasing event. Let $\mathbf{P}$ denote the law of dynamical percolation on $[0,\infty)$. Then there exists a coupling $\mathbf{Q}$ of the laws $\mathbf{P}$ and $\mathbf{P}\big[\cdot \bigm| C\big]$ such that, denoting the two marginals by ω and ω^C, we have that $\mathbf{Q}\big[\omega_0 \preceq \omega_0^C\big] = 1$.*

Assuming this lemma, if $\mathcal{A}(t)$ is the event that the square is crossed for all $s \in [0,t]$, then, for $k \in \mathbb{N}$,

$$\mathbf{P}\big[\mathcal{A}((k+1)/2)\big] \geq \mathbf{P}\big[\mathcal{A}(k/2)\big]\mathbf{P}\big[\mathcal{A}(1/2)\big]. \tag{4.3}$$

Indeed, Lemma 4.2 implies that, conditionally on $\mathcal{A}(k/2)$, the distribution of the marginal of dynamical percolation at time $k/2$ stochastically dominates critical percolation. The event $\mathcal{A}(k/2)$ being conditionally independent of the subsequent evolution of dynamical percolation given the configuration at time $k/2$, we see that, conditionally on $\mathcal{A}(k/2)$, the distribution of dynamical percolation on $[k/2,(k+1)/2]$ stochastically dominates its unconditioned counterpart, whence (4.3).

We claim now that $\mathbf{P}\big[\mathcal{A}(1/2)\big] \geq 1/4$. Indeed, the expected number of pivotal switches during a duration of one-half of scaled time is $1/2$, by Fubini's theorem, and, by symmetry, this remains the case conditionally on there being a left-right crossing at the start of this duration. So, by Markov's inequality, the conditional probability of having no pivotal switch during this time is at least $1/2$. The probability of a left-right crossing being $1/2$, we find that $\mathbf{P}\big[\mathcal{A}(1/2)\big] \geq 1/4$. Thus, iterating (4.3) gives the lower bound in Corollary 4.1. □

We still owe the proof of the dynamical FKG lemma that we used:

Proof of Lemma 4.2. Note that the law of ω may be constructed as follows. To each $x \in V(G)$, we associate a sequence $e_i(x)$, $i = 1,2,\ldots$ of independent exponential mean one random variables, and an independent sequence $b_i(x)$, $i = 1,2,\ldots$ of independent Bernoulli random variables. We set ω_0 to be a uniform element in $\{0,1\}^{V(G)}$. We further set $\omega_t(x) = b_i(x)$ where $i \in \mathbb{N}^+$ is minimal subject to $\sum_{j=1}^{i} e_j(x) \leq t$. If no such i exists, then $t < e_1(x)$; in this case, we set $\omega_t(x) = \omega_0(x)$.

Write Ω^+ denote the data $e_i(x)$ and $b_i(x)$ for $i \in \mathbb{N}^+$ and $x \in V(G)$. Note that Ω^+ and ω_0 comprise all of the data that specifies ω. As such, we may denote an instance ω of dynamical percolation on $[0,\infty)$ in the form $(\omega_0,\omega^+) \in \{0,1\}^{V(G)} \times \Omega^+$.

Suppose given an element $\omega^+ \in \Omega^+$. Note that, if $\omega_0, \omega_0' \in \{0,1\}^{V(G)}$ satisfy $\omega_0 \preceq \omega_0'$, then $(\omega_0,\omega^+) \in C$ implies that $(\omega_0',\omega^+) \in C$. Hence, by the Harris-FKG inequality for a static configuration, the distribution of ω_0 under $\mathbf{P}\big[\cdot \bigm| C,\omega^+\big]$ stochastically dominates its distribution under $\mathbf{P}\big[\cdot \bigm| \omega^+\big]$ (which is the uniform distribution).

Let $\mu_{C,+}$ denote the conditional distribution of ω^+ given C. Then

$$\mathbf{P}\big[\cdot \bigm| C\big] = \int \mathbf{P}\big[\cdot \bigm| C,\omega^+\big]\, d\mu_{C,+}(\omega^+). \tag{4.4}$$

We have shown that, for all choices of ω^+, the conditional distribution of ω_0 given C and ω^+ stochastically dominates the uniform distribution on $\{0,1\}^{V(G)}$. This statement remains true after the averaging in (4.4). Hence, we find that the law of ω_0 given C stochastically dominates its unconditioned law. □

Question 4.3. *What is the true decay of the probability for having a left-right crossing of the unit square during $[0,t]$ in the scaling limit of dynamical percolation? We expect it to be $\exp(-t^{\beta+o(1)})$, with $\beta \in (0,1)$.*

We now move to the application of Theorem 1.2 to the study of exceptional times, which is item (2) in the list at the start of Section 4. Consider critical dynamical site percolation $(\omega_t)_{t\geq 0}$ on the infinite triangular lattice or bond percolation on $\mathbb{Z}^2$ (no scaling of space or time), and let

$$\mathcal{E} := \{t \in [0,\infty) : \omega_t \text{ has } 0 \longleftrightarrow \infty\}$$

be the set of exceptional times when the cluster of the origin is infinite. For any fixed time t, we have $\mathbf{P}[t \in \mathcal{E}] = 0$; hence $\mathcal{E}$ has zero Lebesgue measure almost surely. However, as claimed in item (2), it is almost surely nonempty, with Hausdorff dimension 31/36 in the case of the triangular lattice. A natural question is how long one has to wait to see the first exceptional time. It is answered by the following corollary which will also be an important tool for [12] in studying the infinite clusters that appear in $\mathcal{E}$.

Corollary 4.4. *There exist $\infty > c_1 \geq c_2 > 0$ such that, for critical dynamical site percolation on the infinite triangular lattice or bond percolation on $\mathbb{Z}^2$,*

$$\exp(-c_1\, t) \leq \mathbf{P}\big[\mathcal{E} \cap [0,t] = \emptyset\big] \leq \exp(-c_2\, t)\,.$$

In other words, the first exceptional time $\mathsf{FET} := \min \mathcal{E}$ has an exponential tail. Note here that, by [11, Lemma 3.2], the set $\mathcal{E}$ is topologically closed, hence the minimum makes sense. Furthermore, $\mathscr{C}(\omega_0)$ is almost surely finite, hence it takes positive time until a bit in its boundary $\partial\mathscr{C}(\omega_0)$ first changes its status; thus $\mathsf{FET} > 0$ almost surely.

Naturally, the proof of the corollary will go through the finite approximations $\mathsf{FET}_R := \inf\big\{t \in [0,\infty) : \omega_t \text{ has } 0 \longleftrightarrow \partial B_R(0)\big\}$. But first of all we need to prove that these times are actually approximations to FET:

Lemma 4.5. *We have that $\mathsf{FET}_R \to \mathsf{FET}$ almost surely as $R \to \infty$.*

For the proof, we will need another result proved in [11, Lemma 3.2]:

Lemma 4.6. *Almost surely, the set $\mathcal{E}$ of exceptional times is disjoint from the set of times at which the status of a site is updated.*

Proof of Lemma 4.5. By item (2) above, $\mathcal{E} \cap (0,\infty) \neq \emptyset$, thus $\mathsf{FET} < \infty$, almost surely. The sequence FET_R is increasing and bounded above by FET; thus, there exists some random $\tau \in (0,\infty)$ such that $\mathsf{FET}_R \nearrow \tau$. Assume that $\tau < \mathsf{FET}$ with positive probability, which is to say that the cluster $\mathscr{C}_0(\omega_\tau)$ of the origin in ω_τ is finite with positive probability. Almost surely, the set of flip times for any given site is a locally finite subset of $\mathbb{R}$, and the same holds if we take the union of the flip times over the finite set $S := \mathscr{C}_0(\omega_\tau) \cup \partial\mathscr{C}_0(\omega_\tau)$. Therefore, on the event $|\mathscr{C}_0(\omega_\tau)| < \infty$, there exists some random $\epsilon > 0$ such that the interval $(\tau-\epsilon, \tau+\epsilon)$ either has no flip times for S (and hence the set $\mathscr{C}_0(\omega_t)$ remains unchanged), or it has a single flip time, τ. On the other hand, we know that $\limsup_{s\nearrow\tau} |\mathscr{C}_0(\omega_s)| = \infty$ almost surely, which is consistent with the above only if τ is a flip time for a site in $\partial\mathscr{C}_0(\omega_\tau)$, closing at exactly time τ, and if the reopening of this site creates a configuration in which $0 \longleftrightarrow \infty$. However, Lemma 4.6 shows that this circumstance has zero probability to occur at any time. We conclude that $\tau = \mathsf{FET}$ almost surely, which completes the proof. $\square$

Proof of Corollary 4.4. Lemma 4.5 shows that the events $\{\mathcal{E}_R \cap [0,t] = \emptyset\}$ increase to the event $\{\mathcal{E} \cap [0,t] = \emptyset\}$. This means that bounds for $\mathcal{E}_R$ that are uniform in R imply the same bounds for $\mathcal{E}$.

The lower bound is given simply by the probability that the bits (sites or bonds) neighbouring 0 are closed during $[0,t]$.

For the upper bound, in order to apply Theorem 1.2, we need the well-known fact that our Markov process has a spectral gap that is uniform in R (it is just continuous time random walk on the hypercube $\{0,1\}^{B_R(0)}$, with unit rates on the edges), and also need that $\mathbf{P}\big[\mathcal{E}_R \cap [0,1] \neq \emptyset\big] > c > 0$, uniformly in R. For the case of the triangular lattice, this is part of [18, Theorem 1.3], while, for the case of $\mathbb{Z}^2$, part of [6, Theorem 1.5]. □

To conclude, let us point out the following interesting phenomenon. Although the results of [6] behave well under the n vs. $1/\mathsf{Piv}(n) = n^{-3/4+o(1)}$ space-time scaling, and hence it is not hard to show that, in the scaling limit of dynamical percolation in the full plane (mentioned in item (3) above), exceptional times when the ball of radius 1 is connected to infinity do exist and have Hausdorff dimension 31/36 a.s., the tail for the first exceptional time is expected to behave differently in the scaling limit than in the discrete case. If we try to think of the scaling limit process as unit-order regions flipping between being well-connected and not-at-all-connected (analogues of being open and closed in the discrete process) at a roughly unit rate, then it seems reasonable that, similarly to the discrete case, the tail behaviour of not having the connection from radius 1 to infinity is comparable to the obvious lower bound, the tail for not having a connection across the annulus between radii 1 and 2. However, we expect this annulus-crossing tail to be subexponential (see Question 4.3), which would give a subexponential lower bound also here. The same issue from a different viewpoint is that we do not have any more the spectral gap that we needed in order to apply Theorem 1.2, hence there is no reason to hope for an exponential tail.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. Deterministic simulation in LOGSPACE. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pp. 132–140, 1987.

[2] N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized graph products. *Computational Complexity* **5** (1995), 60–75. `http://www.tau.ac.il/~nogaa/PDFS/derand5.pdf` MR-1319493

[3] N. Alon and J. Spencer. *The Probabilistic Method.* 2nd ed. Wiley, 2000. MR-1885388

[4] I. Benjamini, G. Kalai, and O. Schramm. Noise sensitivity of Boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.* **90** (1999), 5–43. arXiv:math.PR/9811157 MR-1813223

[5] T. K. Carne. A transmutation formula for Markov chains. *Bull. Sc. Math.* **109** (1985), 399–405. MR-0837740

[6] C. Garban, G. Pete, and O. Schramm. The Fourier spectrum of critical percolation. *Acta Math.* **205** (2010), no. 1, 19–104. arXiv:0803.3750 MR-2736153

[7] C. Garban, G. Pete, and O. Schramm. Pivotal, cluster and interface measures for critical planar percolation. *Preprint*, arXiv:1008.1378

[8] C. Garban, G. Pete, and O. Schramm. The scaling limits of dynamical and near-critical percolation. *In preparation.*

[9] C. Garban and J. E. Steif. Lectures on noise sensitivity and percolation. In: *Probability and statistical physics in two and more dimensions* (D. Ellwood, C. Newman, V. Sidoravicius and W. Werner, ed.). Proceedings of the Clay Mathematical Institute Summer School and XIV Brazilian School of Probability (Buzios, Brazil), Clay Mathematics Proceedings 15 (2012), 49–154. arXiv:1102.5761

[10] G. Grimmett. *Percolation.* 2nd ed. Grundlehren der mathematischen Wissenschaften 321. Springer-Verlag, Berlin, 1999. MR-1707339

[11] O. Häggström, Y. Peres, and J. E. Steif. Dynamical percolation. *Ann. Inst. H. Poincaré Probab. Statist.* **33** (1997), no. 4, 497–528. MR-1465800

[12] A. Hammond, G. Pete, and O. Schramm. Local time on the exceptional set of dynamical percolation, and the Incipient Infinite Cluster. *Preprint,* arXiv:1208.3826

[13] S. Hoory, N. Linial and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society* **43** (2006), No. 4, 439–561. `http://www.cs.huji.ac.il/~nati/PAPERS/expander_survey.pdf` MR-2247919

[14] D. Levin, Y. Peres and E. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. http://pages.uoregon.edu/dlevin/MARKOV/ MR-2466937

[15] R. Lyons, with Y. Peres. *Probability on trees and networks*. Book in preparation, present version is at http://mypage.iu.edu/~rdlyons.

[16] E. Mossel, R. O'Donnell, O. Regev, J. E. Steif, and B. Sudakov. Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality. *Israel J. Math.* **154** (2006), no. 1, 299–336. arXiv:math.PR/0410560 MR-2254545

[17] O. Schramm and S. Smirnov, with an appendix by C. Garban. On the scaling limits of planar percolation. *Ann. Probab.* **39** (2011), no. 5, 1768–1814. Memorial Issue for Oded Schramm. arXiv:1101.5820 MR-2884873

[18] O. Schramm and J. E. Steif. Quantitative noise sensitivity and exceptional times for percolation. *Ann. Math.* **171** (2010), no. 2., 619–672. arXiv:math.PR/0504586 MR-2630053

[19] S. Smirnov and W. Werner. Critical exponents for two-dimensional percolation. *Math. Res. Lett.* **8** (2001), no. 5-6, 729–744. arXiv:math.PR/0109120 MR-1879816

[20] J. E. Steif. A survey of dynamical percolation. In *Fractal geometry and stochastics IV*, Birkhäuser, pp. 145–174, 2009. arXiv:0901.4760. MR-2762676

[21] N. Th. Varopoulos. Long range estimates for Markov chains. *Bull. Sci. Math. (2)* **109** (1985), no. 3, 225–252. MR-0822826

[22] W. Werner. Lectures on two-dimensional critical percolation. In *Statistical Mechanics*, IAS/Park City Math. Ser., 16, pp. 297–360. Amer. Math. Soc., Providence, RI, 2009. arXiv:0710.0856 MR-2523462