How long do we have to wait for the exceptional, and what will it be like?

Gábor Pete
http://www.math.bme.hu/~gabor

Joint works with Alan Hammond (University of Oxford)
and Elchanan Mossel (UC Berkeley)
and Oded Schramm (Microsoft Research, 1961-2008)

AH & EM & GP [arXiv:1111.6618] and AH & GP & OS in preparation

A basic question

Pairwise independence is weaker than full independence. E.g., can have $X_1, X_2, \ldots, X_n$ on $(\Omega, \mathbf{P})$, with values in some V, and a subset $A \subset V$, s.t.

- the events $\{X_i \in A\}$ are pairwise independent, $\mathbf{P}[\, X_i \in A \,] = 1/2$,
- but the joint probability is large: $\mathbf{P}[\, X_1, \ldots, X_n \in A \,] = 1/n$.

Namely, let σ_i, $i = 1, \ldots, k$ be independent uniform ± 1 bits, $n = 2^k$, and $x_S := \prod_{i \in S} \sigma_i \in \{-1, 1\}$, for all $S \subseteq [k]$. Then x_S and x_T are independent for $S \neq T$, but $\mathbf{P}\big[\, x_S = 1 \;\forall\, S \subseteq [k] \,\big] = \mathbf{P}\big[\, \sigma_i = 1 \;\forall\, i \in [k] \,\big] = 2^{-k} = 1/n$.

Can this happen for stationary reversible Markov processes? I.e.,

- a fast pairwise decorrelation $\mathbf{P}[\, X_0, X_t \in A \,] - \mathbf{P}[\, X_0 \in A \,]^2$,
- but a fat exit tail $\mathbf{P}[\, X_s \in A \text{ for all } 0 \leqslant s \leqslant t \,]$?

Say, can the first one be exponential but the second one only polynomial?

A famous example: random walk on expanders

Degree $\leqslant d$ finite graph $G(V,E)$ is (n,d,c)-expander if $|\partial S|/|S| \geqslant c$ for all $|S| \leqslant \frac{|V|}{2} = \frac{n}{2}$. **With random walks**: $\mathbf{P}[\, X_1 \notin S \mid X_0 \sim \mathrm{Unif}(S)\,] > c$.

With functional analysis: Markov operator $Pf(x) := \mathbf{E}[\, f(X_1) \mid X_0 = x\,]$, self-adjoint on $L^2(V,\pi)$, where π is the stationary measure. Then $(P\mathbf{1}_S, \mathbf{1}_S) \leqslant (1-c)(\mathbf{1}_S, \mathbf{1}_S)$. More generally, if $\pi(\mathrm{supp} f) \leqslant 1-\epsilon$, then

$$(Pf,f) \leqslant (1-\delta_1)(f,f) \quad \text{and} \quad (Pf,Pf) \leqslant (1-\delta_2)(f,f)\,.$$

Equivalently, spectral gap $g := 1 - \lambda_2 > 0$, where

$$\lambda_2 := \sup_{f \perp \mathbf{1}} \frac{(Pf,f)}{(f,f)} = \sup_{f \perp \mathbf{1}} \frac{(Pf,Pf)^{1/2}}{(f,f)^{1/2}}\,.$$

Spectral gap is not enough for exponential decay of correlations, but absolute spectral gap $g_* := 1 - \sup\big\{|\lambda| : \lambda \in \mathrm{Spec}(P) \setminus \{1\}\big\}$ (non-bipartiteness) is: For any $\mathbf{E}_\pi[f] = 0$, we have $\mathbf{E}[\, f(X_0) f(X_t)\,] \leqslant (1-g_*)^t\, \mathbf{E}_\pi[f^2]$.

Theorem (Ajtai, Komlós & Szemerédi 1987). Let $(X_i)_{i=0}^{\infty}$ be a stationary reversible chain with P and π and $\lambda_2 < 1$, and let $\pi(A) \leqslant \beta < 1$. Then there exists $\gamma(\lambda_2, \beta) > 0$ with

$$\mathbf{P}\big[\, X_i \in A \text{ for all } i = 0, 1, \ldots, t \,\big] \leqslant C(1-\gamma)^t \,.$$

Proof. Consider the projection $Q : f \mapsto f\mathbf{1}_A$. Then,

$$\begin{aligned}
\mathbf{P}\big[\, X_i \in A \text{ for } i = 0, 1, \ldots, 2t+1 \,\big] &= \big(Q(PQ)^{2t+1}\mathbf{1}, \mathbf{1}\big) \\
&= \big(P(QP)^t Q\mathbf{1}, (QP)^t Q\mathbf{1}\big)\,, \text{ by self-adjointness of } P \text{ and } Q \\
&\leqslant (1-\delta_1)\,\big((QP)^t Q\mathbf{1}, (QP)^t Q\mathbf{1}\big)\,, \text{ by } \pi(\mathrm{supp}(Qg)) \leqslant \beta \\
&\leqslant (1-\delta_1)\,\big(P(QP)^{t-1} Q\mathbf{1}, P(QP)^{t-1} Q\mathbf{1}\big)\,, \text{ by } Q \text{ being a projection} \\
&\leqslant (1-\delta_1)\,(1-\delta_2)\,\big((QP)^{t-1} Q\mathbf{1}, (QP)^{t-1} Q\mathbf{1}\big)\,, \text{ by } \pi(\mathrm{supp}(Qg)) \leqslant \beta \\
&\leqslant (1-\delta_1)\,(1-\delta_2)^t\,\big(Q\mathbf{1}, Q\mathbf{1}\big)\,, \text{ by iterating previous step} \\
&\leqslant (1-\delta_1)\,(1-\delta_2)^t\,\beta\,,
\end{aligned}$$

and done for odd times. For even times, use monotonicity in t. ∎

A general result

Stationary Markov process ω_t, operator T_t. Let $\pi(\mathcal{C}) = \mathbf{P}\big[\,\omega_0 \in \mathcal{C}\,\big] = p$, and let $f = \mathbf{1}_{\mathcal{C}}$. The decay of correlations of f can be quantified by

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

or $$\mathrm{Var}[T_t f] = (T_t f, T_t f) - (\mathbf{E} f)^2 \leqslant d(2t)\,\mathrm{Var}[f]\,.$$

These two are the same for reversible Markov processes.

Theorem (Hammond, Mossel & P 2011). Under the second condition,

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

Sharp in the regime of polynomial decay. Open in the exponential case.

Proof of Correlation decay $\implies$ exit time tail

Let $p < \lambda < 1$. Consider $H_s := \Big\{\omega \in S : \mathbf{P}[\,\omega_s \in \mathcal{C} \mid \omega_0 = \omega\,] > \lambda\Big\}$, the set of very good hiding places.

Fix large k, let $\tau = t/k$. Check $\omega_s \in \mathcal{C}$ at $s = j\tau$, for $j = 0, \dots, k$.

$$
\begin{aligned}
\mathbf{P}\Big[\omega_s \in \mathcal{C}\ \forall s \in [0,t]\Big] &\leqslant \mathbf{P}\big[\,\forall j : \omega_{j\tau} \in H_\tau^c \cap \mathcal{C}\,\big] + \mathbf{P}\big[\,\exists j : \omega_{j\tau} \in H_\tau \cap \mathcal{C}\,\big] \\
&\leqslant \lambda^k + \sum_{\ell=0}^{k} \lambda^{(k-\ell-1)\vee 0}\,\mathbf{P}\big[\,\omega_{\ell\tau} \in H_\tau\,\big] \\
&\leqslant \lambda^k + \frac{2-\lambda}{1-\lambda}\,\mathbf{P}[\,H_\tau\,]\,.
\end{aligned}
$$

On the other hand, if s is large, then the pairwise decorrelation suggests that $\mathbf{P}\big[\,H_s\,\big]$ has to be small. Indeed,

$$\lambda\,\mathbf{P}\big[\,H_s\,\big] \;\leqslant\; \mathbf{E}\big[\,f(\omega_s)\;\big|\;H_s\,\big]\,\mathbf{P}\big[\,H_s\,\big] = \mathbf{E}\big[\,\mathbf{1}_{H_s}\,T_s f\,\big]$$

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

Rearranging and using Cauchy-Schwarz,

$$(\lambda - p)\,\mathbf{P}\big[\,H_s\,\big] \;\leqslant\; \mathbf{E}\Big[\,\mathbf{1}_{H_s}\,(T_s f - \mathbf{E} f)\,\Big] \leqslant \|\mathbf{1}_{H_s}\|_2\,\|T_s f - \mathbf{E} f\|_2\,,$$

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

Thus

$$\mathbf{P}\big[\,H_\tau\,\big] \leqslant \frac{p - p^2}{(\lambda - p)^2}\,d(2t/k)\,,$$

and can optimize the sum of two terms over k. ∎

Exceptional times in dynamical percolation

Lower bound on Hausdorff dimension needs decay of correlations:

1. $\mathbf{E}\big[\, f_{\mathcal{Q},\eta}(\omega_0)\, f_{\mathcal{Q},\eta}(\omega_{t\eta^{3/4+o(1)}})\,\big] - \mathbf{E}\big[\, f_{\mathcal{Q},\eta}\,\big]^2 \asymp_{\mathcal{Q}} t^{-2/3}$ as $t \to \infty$, uniformly in mesh η, for the indicator of left-right crossing in the quad $\mathcal{Q}$.

2. $\mathbf{E}\big[\, f_R(\omega_0)\, f_R(\omega_t)\,\big] / \mathbf{E}\big[\, f_R(\omega)\,\big]^2 \asymp t^{-(4/3)\,\xi_1 + o(1)}$, as $t \to 0$, for the indicator of the one-arm event to radius R.

Now, Mass Distribution Principle for the measure $\overline{\mu}_R[a,b] = \int_a^b \frac{\mathbf{1}\{0 \overset{\omega_t}{\longleftrightarrow} R\}}{\mathbf{P}[\,0 \longleftrightarrow R\,]}\, dt$ on $\mathscr{E}_R = \{t : 0 \overset{\omega_t}{\longleftrightarrow} R\}$ and some compactness: if

$$\sup_R \int_0^1 \int_0^1 \frac{\mathbf{E}\big[\, f_R(\omega_t)\, f_R(\omega_s)\,\big]}{\mathbf{E}\big[\, f_R(\omega)\,\big]^2\, |t-s|^\gamma}\, dt\, ds < \infty\,,$$

then $\dim(\mathscr{E}) \geqslant \gamma$ a.s. Hence $\dim(\mathscr{E}) \geqslant 1 - \frac{4}{3}\xi_1$.

Two natural questions on the exceptional set

How do exceptional infinite clusters look like? The first one? A typical one?

There is an "infinite critical cluster" in the static world, Kesten's Incipient Infinite Cluster measure (1986): for $H \subset \Delta$ and ω^H configuration in H, the limit $\mathrm{IIC}(\omega^H) = \lim_{R\to\infty} \mathbf{P}[\,\omega^H \,|\, 0 \leftrightarrow R\,]$ exists.

All other natural definitions give the same measure (Járai 2003).

What is the hitting time tail $\mathbf{P}\big[\,\mathscr{E} \cap [0,t] = \emptyset\,\big]$?

To answer the first question, we needed to answer the second one:

Theorem (Hammond, Mossel & P. 2011). The hitting time tail is exponentially small.

Theorem (Hammond, P. & Schramm 2012). The configuration at a "typical" exceptional time has the law of IIC, but the First Exceptional Time Infinite Cluster (FETIC) is thinner.

Proof of exponential tail for FET

Dynamical percolation in B_R is just continuous time random walk on the hypercube $\{0,1\}^{B_R}$, with rate 1 clocks on the edges. On $\{0,1\}^n$, discrete time random walk has spectral gap $1/n$, but in continuous time, the gap is uniformly positive, so could try to use [AKSz'87].

Of course, $\mathbf{P}[\,0 \longleftrightarrow R\,]$ is tiny, so we don't want to hit that set. But $\mathbf{P}[\,\mathscr{E}_R \cap [0,1] \neq \emptyset\,]$ is uniformly positive!

So, first idea: Markov chain $\{\omega[2t, 2t+1] : t = 0, 1, 2\}$ on a huge state space. This again has a uniform spectral gap. However, it's not reversible!

So, another trick: $L^2(\Omega, \mathbf{P})$ is the space of trajectories $\{\omega_t : t \in \mathbb{R}\}$, on it the event $A_t := \{\mathscr{E}_R \cap [t, t+1] = \emptyset\}$ for any $t \in \mathbb{R}$, then the projection $Q_t f := f \mathbf{1}_{A_t}$ is still self-adjoint and $\mathbf{P}[\,\mathrm{supp}(Q_t g)\,] \leqslant \beta < 1$ for any g. On the other hand, for $g_i(\omega) := \mathbf{E}[\,f_i(\omega[0,1]) \mid \omega_0 = \omega\,]$, we have $\mathbf{E}\big[\,f_1(\omega[0,1])\,f_2(\omega[t,t+1])\,\big] = \mathbf{E}_\pi[g_1\,T_t g_2]$, hence the spectral gap of T_t can be used.

Local time measure for exceptional times

$$\overline{M}_r(\omega_s) := \frac{\mathbf{1}\{0 \leftrightarrow r\}}{\mathbf{P}[\,0 \leftrightarrow r\,]}, \quad \overline{\mu}_r[a,b] := \int_a^b \overline{M}_r(\omega_s)\,ds, \quad \overline{\mu}[a,b] := \lim_{r\to\infty} \overline{\mu}_r[a,b].$$

This $\overline{M}_r(\omega)$ is a martingale w.r.t. the filtration $\overline{\mathscr{F}}_r$ of the percolation space generated by the variables $\mathbf{1}\{0 \leftrightarrow r\}$. Moreover, $\mathbf{E}\overline{\mu}_r[a,b] = b-a$, and, by the correlation decay, $\sup_r \mathbf{E}\big[\,\overline{\mu}_r[a,b]^2\,\big] < C_1$. So $\lim_r$ exists.

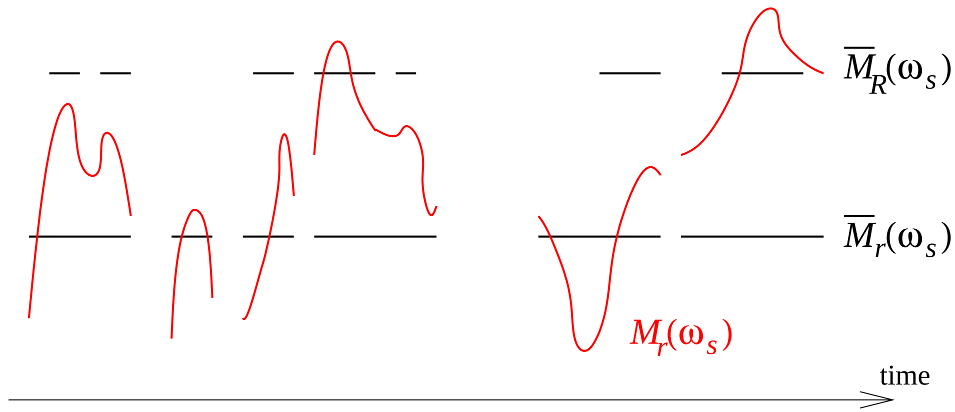
$$M_H(\omega) := \lim_{R\to\infty} \frac{\mathbf{P}[\,0 \leftrightarrow R \,|\, \omega^H\,]}{\mathbf{P}[\,0 \leftrightarrow R\,]} = \lim_{R\to\infty} \frac{\mathbf{P}[\,\omega^H \,|\, 0 \leftrightarrow R\,]}{\mathbf{P}[\,\omega^H\,]} = \frac{\mathsf{IIC}(\omega^H)}{\mathbf{P}[\,\omega^H\,]}.$$

$$M_r(\omega_s) := M_{B_r}(\omega_s), \quad \mu_r[a,b] := \int_a^b M_r(\omega_s)\,ds, \quad \mu[a,b] := \lim_{r\to\infty} \mu_r[a,b].$$

Now $M_r(\omega)$ is a MG w.r.t. the full filtration $\mathscr{F}_r$ generated by $\omega(B_r)$, again $\mathbf{E}\mu_r[a,b] = b-a$, and $M_r(\omega) \leqslant C_2\,\overline{M}_r(\omega)$ because of quasi-multiplicativity:

$$\frac{\mathbf{P}[\,0 \leftrightarrow R \,|\, \omega^{B_r}\,]}{\mathbf{P}[\,0 \leftrightarrow R\,]} \asymp \frac{\mathbf{P}[\,0 \leftrightarrow R \,|\, \omega^{B_r}\,]}{\mathbf{P}[\,0 \leftrightarrow r\,]\mathbf{P}[\,r \leftrightarrow R\,]}$$
$$\leqslant \frac{\mathbf{P}[\,r \leftrightarrow R \,|\, \omega^{B_r}\,]\mathbf{1}\{0 \leftrightarrow r\}}{\mathbf{P}[\,0 \leftrightarrow r\,]\mathbf{P}[\,r \leftrightarrow R\,]} = \frac{\mathbf{1}\{0 \leftrightarrow r\}}{\mathbf{P}[\,0 \leftrightarrow r\,]}.$$

Hence, both local time measures exist, and are clearly supported inside $\mathscr{E}$.



$$\mathbf{E}\big[\,\overline{M}_R \,\big|\, \mathscr{F}_r\,\big] = \frac{\mathbf{P}[\,0 \leftrightarrow R \,|\, \mathscr{F}_r\,]}{\mathbf{P}[\,0 \leftrightarrow R\,]} \xrightarrow[L^\infty]{\text{a.s.}} M_r\,, \quad \text{for fixed } r \text{ and } R \to \infty\,.$$

Theorem (Hammond, P & Schramm 2012). $\overline{\mu} = \mu$ a.s. At a μ-typical time, the configuration has the distribution of IIC.

Question: is it true that $\operatorname{supp}(\mu) = \mathscr{E}$?

FETIC versus IIC

Mutual singularity should hold, but let's just show that there is some ω^{B_r} such that $\lim_{R\to\infty} \mathsf{FETIC}_R(\omega^{B_r}) \neq \lim_{R\to\infty} \mathsf{IIC}_R(\omega^{B_r})$.

The configuration at a typical switch time for $\{0 \longleftrightarrow R\}$ is size-biased by the number of pivotals. Because of the many pivotals far from the origin, inside B_r this bias becomes negligible as $R \to \infty$, so we still have IIC.

The configuration at FET_R is further size-biased by the length of the non-connection interval ending at the switch time.

For any $\omega = \omega^{B_R}$ satisfying $\{0 \longleftrightarrow R\}$, get $\mathsf{THIN}_r(\omega)$ by thinning inside B_r.

Want to show that the reconnection time $V = V_{r,R}$ started from $\mathsf{THIN}_r(\omega^{B_R})$ is larger in expectation than $N = N_{r,R}$, the one started from the normal ω^{B_R}, uniformly as $R \to \infty$. (While both are very small.)

Because of the thinning, there is some $\epsilon(r) \to 0$ and $g(r) \to \infty$ with

$$\mathbf{P}\big[\, V > g(r) \;\big|\; V > \epsilon(r) \,\big] \;>\; c_1 \,. \tag{1}$$

Also, from stochastic domination, $\mathbf{P}[\, V > \epsilon(r) \,] \;\geqslant\; \mathbf{P}[\, N > \epsilon(r) \,]\,.$ (2)

(1) would be hard, so our thinning is different, and (2) doesn't quite hold.

Write $X^\epsilon = X\,\mathbf{1}_{\{X > \epsilon(r)\}}$. Note that size-biased $\widehat{N}$ times Unif$[0,1]$ is FET.

A size-biasing lemma: $\mathbf{P}\big[\,\widehat{N} > \epsilon(r)\,\big] = \frac{\mathbf{E}[\,N^\epsilon\,]}{\mathbf{E}[\,N\,]} > c_2$ and $\mathbf{E}\big[\,\widehat{N}\,\big] < C_1$ imply

$$\mathbf{E}\big[\, N \;\big|\; N > \epsilon(r) \,\big] \;<\; C_2 \,. \tag{3}$$

From these three,

$$\mathbf{E}[\,V^\epsilon\,] \;\geqslant\; c_1\, g(r)\, \mathbf{P}[\, N > \epsilon(r) \,] \;\geqslant\; c_1\, g(r)\, \frac{\mathbf{E}[\,N^\epsilon\,]}{C_2}\,,$$

hence

$$\mathbf{E}V \;\geqslant\; \mathbf{E}[\,V^\epsilon\,] \;\gg_r\; \mathbf{E}[\,N^\epsilon\,] \;\geqslant\; c_2\,\mathbf{E}N\,.$$

What about the tail of left-right connection?

As mentioned before, $\mathbf{E}\big[\, f_{\mathcal{Q},\eta}(\omega_0)\, f_{\mathcal{Q},\eta}(\omega_{t\eta^{3/4+o(1)}})\,\big] - \mathbf{E}[\, f_{\mathcal{Q},\eta}\,]^2 \asymp_{\mathcal{Q}} t^{-2/3}$ as $t \to \infty$, uniformly in mesh η, hence natural to rescale time like this.

In fact, there exists a scaling limit of dynamical percolation [Garban, P. & Schramm 2012], so one can either talk about the rescaled finite chains, "uniformly in η", or about the scaling limit process.

Earlier theorem [HMP'11] gives $\mathbf{P}\big[\, f_{\mathcal{Q}}(\omega_s) = 1\ \forall\, s \in [0,t]\,\big] \leqslant t^{-2/3+o(1)}$.

In fact, by cutting $\mathcal{Q}$ vertically into L slabs: $\leqslant t^{-2L/3+o(1)}$ for any L, superpolynomial decay.

Exponential lower bound is easy from dynamical FKG inequality.

Conjecture. $\mathbf{P}\big[\, f_{\mathcal{Q}}(\omega_s) = 1$ for all $s \in [0,t]\,\big] = \exp(-t^{2/3+o(1)})$.

Supported by a *very* non-rigorous renormalization argument.