Noise sensitivity in critical percolation and what else might we learn from it

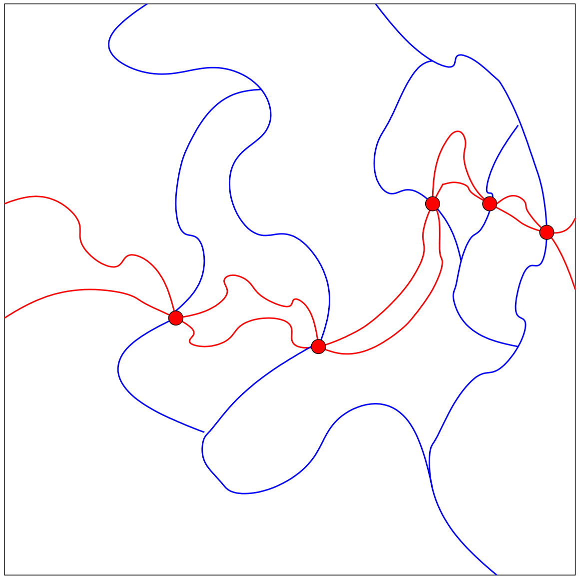
Mostly based on C. Garban, G. Pete & O. Schramm:
The Fourier spectrum of critical percolation, *Acta Math.* 2010.

Naive idea: how many pivotals are there?

A site (or bond) is pivotal in ω, if flipping it changes the existence of a left-right crossing.
$\mathbf{E}|\mathrm{Piv}_n| \asymp n^2\,\alpha_4(n) \quad (= n^{3/4+o(1)})$.

Furthermore, $\mathbf{E}[\,|\mathrm{Piv}_n|^2\,] \leqslant C\,(\mathbf{E}|\mathrm{Piv}_n|)^2$.
So, $\mathbf{P}\big[\,|\mathrm{Piv}_n| > \lambda\mathbf{E}|\mathrm{Piv}_n|\,\big] < C/\lambda^2$, any λ.

Tightness around mean also from below: $\mathbf{P}\big[\,0 < |\mathrm{Piv}_n| < \lambda\mathbf{E}|\mathrm{Piv}_n|\,\big] \asymp \lambda^{11/9+o(1)}$, as $\lambda \to 0$ (exponent only for Δ).



Cannot have many pivotals. $\Longrightarrow$ If $\epsilon_n\mathbf{E}[\,|\mathrm{Piv}_n|\,] \to 0$, then we don't hit any pivotals. $\Longrightarrow$ Asymptotically full correlation.

Cannot have few pivotals (if there is any). $\Longrightarrow$ If $\epsilon_n\mathbf{E}[\,|\mathrm{Piv}_n|\,] \to \infty$, then we do hit many pivotals. But this $\not\Longrightarrow$ asymptotic independence!

What is the Fourier spectrum and why is it useful?

$f_n : \{\pm 1\}^{V_n} \longrightarrow \{\pm 1\}$ indicator function of left-right crossing.

$(N_\epsilon f)(\omega) := \mathbf{E}[\, f(\omega^\epsilon) \mid \omega \,]$ is the noise operator, acting on the space $L^2(\Omega, \mu)$, where $\Omega = \{\pm 1\}^V$, μ uniform measure, inner product $\mathbf{E}[\, fg \,]$.

Correlation: $\mathbf{E}[\, f(\omega^\epsilon) f(\omega) \,] - \mathbf{E}[\, f(\omega) \,]\mathbf{E}[\, f(\omega^\epsilon) \,] = \mathbf{E}[\, f(\omega) N_\epsilon f(\omega) \,] - \mathbf{E}[\, f(\omega) \,]^2$. So, we would like to diagonalize the noise operator N_ϵ.

Let χ_i be the function $\chi_i(\omega) = \omega(i)$, $\omega \in \Omega$.

For $S \subset V$, let $\chi_S := \prod_{i \in S} \chi_i$, the parity inside S. Then

$$N_\epsilon \chi_i = (1-\epsilon)\, \chi_i \,; \qquad N_\epsilon \chi_S = (1-\epsilon)^{|S|}\, \chi_S .$$

Moreover, the family $\{\chi_S,\ S \subset V\}$ is an orthonormal basis of $L^2(\Omega, \mu)$.

Any function $f \in L^2(\Omega, \mu)$ in this basis (Fourier-Walsh series):

$$\hat{f}(S) := \mathbf{E}[\, f\chi_S \,]\,; \qquad f = \sum_{S\subset V} \hat{f}(S)\,\chi_S\,.$$

The correlation:

$$\mathbf{E}[\, f N_\epsilon f \,] - \mathbf{E}[\, f \,]^2 = \sum_S \sum_{S'} \hat{f}(S)\,\hat{f}(S')\,\mathbf{E}\big[\,\chi_S\, N_\epsilon \chi_{S'}\,\big] - \mathbf{E}[\, f\chi_\emptyset \,]^2$$

$$= \sum_{\emptyset\neq S\subset V} \hat{f}(S)^2\,(1-\epsilon)^{|S|} = \sum_{k=1}^{|V_n|} (1-\epsilon)^k \sum_{|S|=k} \hat{f}(S)^2.$$

By Parseval, $\sum_S \hat{f}(S)^2 = \mathbf{E}[\, f^2 \,] = 1$. So can define probability measure $\mathbf{P}\big[\,\mathscr{S}_f = S\,\big] := \hat{f}(S)^2/\mathbf{E}[\, f^2 \,]$, the spectral sample $\mathscr{S}_f \subset V$.

If, for some sequence k_n, we have $\nu\big[0 < |\mathscr{S}_n| < tk_n\big] \to 0$ as $t \to 0$, uniformly in n, then $(1-\epsilon)^k \sim \exp(-\epsilon k)$ implies that for $\epsilon_n \gg 1/k_n$ we have asymptotic independence. Maybe with $k_n = \mathbf{E}|\mathscr{S}_n|$?

Pivotals versus spectral sample

$\nabla_i f(\omega) := f(\sigma_i(\omega)) - f(\omega) \in \{-2, 0, +2\}$ gradient.

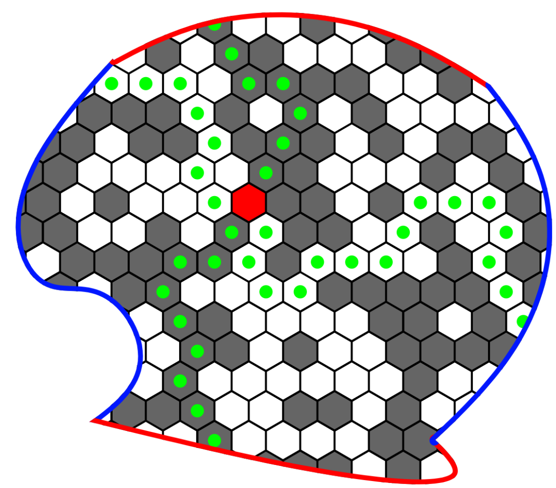
$\nabla_i f(\omega) = \sum_S \hat{f}(S)[\chi_S(\sigma_i(\omega)) - \chi_S(\omega)]$, hence $\widehat{\nabla_i f}(S) = 2\hat{f}(S)\mathbf{1}_{i\in S}$.

$\mathbf{P}[\, i \in \mathrm{Piv}_f \,] = \frac{1}{4}\|\nabla_i f\|_2^2 = \frac{1}{4}\sum_S \widehat{\nabla_i f}(S)^2 = \sum_{S\ni i} \hat{f}(S)^2 = \mathbf{P}[\, i \in \mathscr{S}_f \,]$.

Thus, $\mathbf{E}|\mathscr{S}_f| = \mathbf{E}|\mathrm{Piv}_f|$. So, the pivotal upper bound for noise sensitivity is sharp if there is tightness around $\mathbf{E}|\mathscr{S}|$.

Will see $\mathbf{P}[\, i,j \in \mathrm{Piv}_f \,] = \mathbf{P}[\, i,j \in \mathscr{S}_f \,]$, hence $\mathbf{E}|\mathscr{S}_f|^2 = \mathbf{E}|\mathrm{Piv}_f|^2$.

Not for more points and higher moments! Both random subsets measure the "influence" or "relevance" of bits, but in different ways.



For percolation, $\mathbf{E}[\,|\mathrm{Piv}_n|^2\,] \leqslant C\,(\mathbf{E}|\mathrm{Piv}_n|)^2$, hence $\exists c > 0$ s.t. $\mathbf{P}\big[\,|\mathscr{S}_n| > c\,\mathbf{E}|\mathscr{S}_n|\,\big] > c$. That's why one hopes for tightness around mean.

Three very simple examples

Dictator$_n(x_1, \ldots, x_n) := x_1$.
Here Cov$\big[$ Dic$_n(x)$, Dic$_n(x^\epsilon)$ $\big] = 1 - \epsilon$, so noise-stable.
And $\mathbf{P}\big[\, \mathscr{S}_n = \{x_1\} \,\big] = 1$.

Majority$_n(x_1, \ldots, x_n) := \mathrm{sgn}\,(x_1 + \cdots + x_n) \approx \frac{1}{\sqrt{n}}(x_1 + \cdots + x_n)$.
Here Cov$\big[$ Maj$_n(x)$, Maj$_n(x^\epsilon)$ $\big] = 1 - O(\epsilon)$, so noise-stable.
And $\mathbf{P}\big[\, \mathscr{S}_n = \{x_i\} \,\big] \asymp 1/n$, most of the weight is on singletons.
On the other hand, $\mathbf{E}|\mathscr{S}_n| = \mathbf{E}|Piv_n| \asymp \frac{1}{\sqrt{n}}\, n \asymp \sqrt{n}$.

Parity$_n(x_1, \ldots, x_n) := x_1 \cdots x_n$
Here Cov$\big[$ Par$_n(x)$, Par$_n(x^\epsilon)$ $\big] = (1 - \epsilon)^n$, the most sensitive to noise.
And $\mathbf{P}\big[\, \mathscr{S}_n = \{x_1, \ldots, x_n\} \,\big] = 1$.

Benjamini, Kalai & Schramm 1998

Theorem. A sequence f_n of monotone Boolean functions is noise sensitive, i.e., for any fixed $\epsilon > 0$,

$$\mathbf{E}\big[\, f_n(\omega)\, f_n(\omega^\epsilon)\,\big] - \mathbf{E}\big[\, f_n(\omega)\,\big]^2 \to 0$$

as $n \to \infty$, iff it is asymptotically uncorrelated with all weighted majorities $\mathrm{Maj}_w(x_1, \ldots, x_n) = \mathrm{sign} \sum_{i=1}^n x_i w_i$. Also, not very slow decorrelation with all subset-majorities is enough for sensitivity.

Theorem. The left-right percolation crossing in an $n \times n$ square is noise sensitive, even with $\epsilon = \epsilon_n > c/\log n$.

Steif & Schramm 2005

Theorem. If $f : \Omega \longrightarrow \mathbb{R}$ can be computed with a randomized algorithm with revealment δ, then

$$\sum_{S:|S|=k} \hat{f}(S)^2 \leqslant \delta\, k\, \|f\|_2^2 \,.$$

For left-right crossing in $n \times n$ box on the hexagonal lattice, exploration interface with random starting point gives revealment $n^{-1/4+o(1)}$ (it has length $n^{7/4+o(1)}$, given by 2-arm exponent), while $\sum_{k \leqslant m} k \asymp m^2$, thus:

Theorem. Left-right crossing on the triangular lattice is noise sensitive under $\epsilon_n > n^{-a}$, with any $a < 1/8$. Even on square lattice, can take some positive $a > 0$.

The revealment is at least $n^{-1/2+o(1)}$ for *any* algorithm computing the crossing, hence this method can give only $n^{-1/4+o(1)}$-sensitivity, far from the conjectured $\epsilon_n = n^{-3/4+o(1)}$.

The GPS approach, 2008

Although goal is to understand size, Gil Kalai suggested trying to understand entire distribution of $\mathscr{S}_f$. A strange random set of bits.

Effective sampling? If f is an effectively computable Boolean function, then there is an effective quantum algorithm for $\mathscr{S}_f$ [Bernstein-Vazirani 1993].

For $\mathscr{S}_{\mathcal{Q},n}$ (left-right crossing in a conformal rectangle $\mathcal{Q}$, mesh $1/n$), [Smirnov '01] + [Tsirelson '04] + [Schramm-Smirnov '11] implies that it has a conformally invariant scaling limit.

How to prove tightness for the size of strange random fractal-like sets?

Basic properties of the spectral sample

For $A \subseteq V$: $\mathbf{E}\big[\,\chi_S \mid \mathscr{F}_A\,\big] = \begin{cases} \chi_S & S \subseteq A\,, \\ 0 & \text{otherwise}\,. \end{cases}$

Therefore, $\mathbf{E}\big[\,f \mid \mathscr{F}_A\,\big] = \sum_{S \subseteq A} \widehat{f}(S)\,\chi_S$, a nice projection.

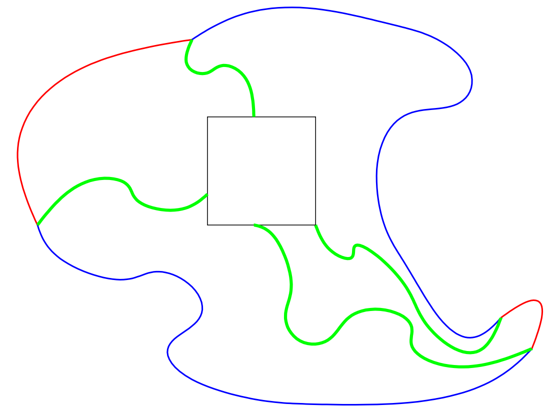
Also, for $S \subseteq A$: $\mathbf{E}\big[\,f\,\chi_S \mid \mathscr{F}_{A^c}\,\big] = \sum_{S' \subseteq A^c} \widehat{f}(S \cup S')\,\chi_{S'}$, hence

$$\mathbf{E}\Big[\mathbf{E}\big[\,f\,\chi_S \mid \mathscr{F}_{A^c}\,\big]^2\Big] = \sum_{S' \subseteq A^c} \widehat{f}(S \cup S')^2 = \mathbf{P}\big[\,\mathscr{S} \cap A = S\,\big]\,.$$

This is the Random Restriction Lemma of Linial-Mansour-Nisan '93. E.g.,

$$\begin{aligned} \mathbf{P}\big[\,i,j \in \mathscr{S}_f\,\big] &= \mathbf{E}\Big[\mathbf{P}\big[\,f\chi_{\{i,j\}} \mid \mathscr{F}_{\{i,j\}^c}\,\big]\Big] \\ &= \frac{1}{4}\mathbf{P}\big[\,\omega\big|_{\{i,j\}^c} \text{ is such that } i,j \text{ each may be pivotal}\,\big] \\ &= \mathbf{P}\big[\,i,j \in \mathrm{Piv}_f\,\big]\,. \end{aligned}$$

How does $\left[\mathscr{S}_n \cap B \,\middle|\, \mathscr{S}_n \cap B \neq \emptyset\right]$ look like?



B as set has to be pivotal.

Strong Separation Lemma. For $\mathrm{d}(B, \partial Q) > \mathrm{diam}(B)$, conditioned on the 4 interfaces to reach ∂B, with *arbitrary starting points*, with a uniformly positive conditional probability the interfaces are well-separated around ∂B. Very bad separation is very unlikely. [Simple proof by Damron-Sapozhnikov '09, following Kesten '87. Also explained in Appendix to GPS '10.]

Corollary 1. $\mathbf{P}\Big[\, \mathscr{S}_n \cap B_r \neq \emptyset \,\Big] \asymp \alpha_4(r, n)\,.$

Corollary 2. $\mathbf{E}\Big[\, |\mathscr{S}_n \cap B_r| \,\Big|\, \mathscr{S}_n \cap B_r \neq \emptyset \,\Big] \asymp r^2\, \alpha_4(1, r) \asymp \mathbf{E}|\mathscr{S}_r|\,.$

Self-similarity for left-right crossing of $n \times n$ square

$$\mathbf{E}|\mathscr{S}_n| = \mathbf{E}|\mathrm{Piv}_n| \asymp n^2\,\alpha_4(1,n) \overset{\Delta}{\asymp} n^{3/4+o(1)}\,,$$

$$\mathbf{E}|\mathscr{S}_n(r)| := \mathbf{E}\Big[\#\big\{r\text{-boxes } \mathscr{S}_n \cap B_r \neq \emptyset\big\}\Big] \asymp \frac{n^2}{r^2}\,\alpha_4(r,n) \asymp \mathbf{E}|\mathscr{S}_{n/r}|\,,$$

$$\mathbf{E}\Big[\,|\mathscr{S}_n \cap B_r| \;\Big|\; \mathscr{S}_n \cap B_r \neq \emptyset\Big] \asymp r^2\,\alpha_4(1,r) \asymp \mathbf{E}|\mathscr{S}_r|\,.$$

Of course, $r^2\,\alpha_4(1,r) \cdot \frac{n^2}{r^2}\,\alpha_4(r,n) \asymp n^2\,\alpha_4(1,n)$, by quasi-multiplicativity.

Self-similarity for left-right crossing of $n \times n$ square

$$\mathbf{E}|\mathscr{S}_n| = \mathbf{E}|\mathrm{Piv}_n| \asymp n^2\,\alpha_4(1,n) \overset{\Delta}{\asymp} n^{3/4+o(1)}\,,$$

$$\mathbf{E}|\mathscr{S}_n(r)| := \mathbf{E}\Big[\,\#\big\{r\text{-boxes } \mathscr{S}_n \cap B_r \neq \emptyset\big\}\,\Big] \asymp \frac{n^2}{r^2}\,\alpha_4(r,n) \asymp \mathbf{E}|\mathscr{S}_{n/r}|\,,$$

$$\mathbf{E}\Big[\,|\mathscr{S}_n \cap B_r|\;\Big|\;\mathscr{S}_n \cap B_r \neq \emptyset\Big] \asymp r^2\,\alpha_4(1,r) \asymp \mathbf{E}|\mathscr{S}_r|\,.$$

Of course, $r^2\,\alpha_4(1,r)\cdot\frac{n^2}{r^2}\,\alpha_4(r,n) \asymp n^2\,\alpha_4(1,n)$, by quasi-multiplicativity.

Similar to the zero-set of simple random walk: $\mathbf{E}|\mathscr{Z}_n| \asymp n\,n^{-1/2} = n^{1/2}$,

$$\mathbf{E}|\mathscr{Z}_n(r)| := \mathbf{E}\Big[\,\#\big\{r\text{-intervals } \mathscr{Z}_n \cap I_r \neq \emptyset\big\}\,\Big] \asymp \frac{n}{r}\,(n/r)^{-1/2} \asymp \mathbf{E}|\mathscr{Z}_{n/r}|\,,$$

$$\mathbf{E}\Big[\,|\mathscr{Z}_n \cap I_r|\;\Big|\;\mathscr{Z}_n \cap I_r \neq \emptyset\Big] \asymp r\,r^{-1/2} \asymp \mathbf{E}|\mathscr{Z}_r|\,.$$

These results are related to the existence of scaling limits.

What concentration can we expect?

$\mathscr{S}_n$ is very different from uniform set of similar density:
i.i.d. $\mathbf{P}[\, x \in \mathscr{U}_n \,] = n^{-5/4}$. Hence $\mathbf{E}|\mathscr{U}_n| = n^{3/4}$.

For large r ($\gg n^{5/8}$), this $\mathscr{U}_n$ intersects every r-box;
for small r, if it intersects one, there is just one point there.

Concentration of size: roughly within $\sqrt{\mathbf{E}|\mathscr{U}_n|} = n^{3/8}$.

A bit more similar: for $i = 1, \dots, (n/r)^2$, i.i.d. $\mathbf{P}\big[\, X_i = r^{3/4} \,\big] = (n/r)^{-5/4}$,
$X_i = 0$ otherwise. Then $S_{n,r} := \sum_i X_i$. Hence $\mathbf{E}|S_{n,r}| = n^{3/4}$.

For $r = n^\gamma$, size $|S_{n,r}|$ is concentrated within $n^{3/8(1+\gamma)}$, still $o(\mathbf{E}|S_{n,r}|)$.

For self-similar sets, we expect only tightness around the mean:
$\mathbf{P}\big[\, 0 < |\mathscr{S}_n| < \lambda\, \mathbf{E}|\mathscr{S}_n| \,\big] \to 0$ as $\lambda \to 0$, uniformly in n.

Proving tightness with a lot of independence

Assume we have the following ingredients, true for the zeroes:

(1) $$\mathbf{P}\Big[\,|\mathcal{Z}_n \cap I_r| > c\,\mathbf{E}|\mathcal{Z}_r| \;\Big|\; \mathcal{Z}_n \cap I_r \neq \emptyset,\ \mathscr{F}_{[n]\setminus I_r}\Big] \geqslant c > 0.$$

(2) $\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = k\,\big] \leqslant g(k)\,\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = 1\,\big]$, with sub-exponential $g(k)$:

when the r-intervals intersected are scattered, have to pay k times to get to and leave them, and this cost is not balanced by combinatorial entropy.

Proving tightness with a lot of independence

Assume we have the following ingredients, true for the zeroes:

(1) $\mathbf{P}\Big[\,|\mathcal{Z}_n \cap I_r| > c\,\mathbf{E}|\mathcal{Z}_r| \,\Big|\, \mathcal{Z}_n \cap I_r \neq \emptyset,\ \mathscr{F}_{[n]\setminus I_r}\Big] \geqslant c > 0.$

(2) $\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = k\,\big] \leqslant g(k)\,\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = 1\,\big]$, with sub-exponential $g(k)$:

when the r-intervals intersected are scattered, have to pay k times to get to and leave them, and this cost is not balanced by combinatorial entropy.

$$\mathbf{P}\big[\,0 < |\mathcal{Z}_n| < c\,\mathbf{E}|\mathcal{Z}_r|\,\big] = \sum_{k\geqslant 1}\mathbf{P}\Big[\,0 < |\mathcal{Z}_n| < c\,\mathbf{E}|\mathcal{Z}_r|\,,\ |\mathcal{Z}_n(r)| = k\,\Big]$$

$$\text{by (1):}\quad \leqslant \sum_{k\geqslant 1}(1-c)^k\,\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = k\,\big]$$

$$\text{by (2):}\quad \leqslant O(1)\,\mathbf{P}\big[\,|\mathcal{Z}_n(r)| = 1\,\big] \asymp (n/r)^{1-3/2},$$

which, using $\lambda = \frac{c\,\mathbf{E}|\mathcal{Z}_r|}{\mathbf{E}|\mathcal{Z}_n|} \asymp (r/n)^{1/2}$, reads as $\mathbf{P}\big[\,0 < |\mathcal{Z}_n| < \lambda\,\mathbf{E}|\mathcal{Z}_n|\,\big] \asymp \lambda$.

But we know much less independence for $\mathscr{S}_n$

$$(1') \quad \mathbf{P}\Big[\, |\mathscr{S}_n \cap B_r/3| > c\,\mathbf{E}|\mathscr{S}_r| \;\Big|\; \mathscr{S}_n \cap B_r \neq \emptyset = \mathscr{S}_n \cap W \,\Big] \geqslant c > 0,$$

for any W that is not too close to B_r.

Why only this negative conditioning? Inclusion formula:

$$\mathbf{P}[\,\mathscr{S}_f \subset U\,] = \sum_{S \subset U} \hat{f}(S)^2 = \mathbf{E}\Big[\Big(\sum_{S \subset U} \hat{f}(S)\,\chi_S\Big)^2\Big] = \mathbf{E}\Big[\mathbf{E}[\,f \mid \mathscr{F}_U\,]^2\Big].$$

From this, for disjoint subsets A and B,

$$\begin{aligned}
\mathbf{P}[\,\mathscr{S}_f \cap B \neq \emptyset = \mathscr{S}_f \cap A\,] &= \mathbf{P}[\,\mathscr{S}_f \subseteq A^c\,] - \mathbf{P}[\,\mathscr{S}_f \subseteq (A \cup B)^c\,] \\
&= \mathbf{E}\Big[\mathbf{E}[\,f \mid \mathscr{F}_{A^c}\,]^2 - \mathbf{E}[\,f \mid \mathscr{F}_{(A\cup B)^c}\,]^2\Big] \\
&= \mathbf{E}\Big[\big(\mathbf{E}[\,f \mid \mathscr{F}_{A^c}\,] - \mathbf{E}[\,f \mid \mathscr{F}_{(A\cup B)^c}\,]\big)^2\Big].
\end{aligned}$$

So, what are we going to do?

With quite a lot of work for both items,

(1') $\mathbf{P}\Big[\, |\mathscr{S}_n \cap B_r/3| > c\,\mathbf{E}|\mathscr{S}_r| \;\Big|\; \mathscr{S}_n \cap B_r \neq \emptyset = \mathscr{S}_n \cap W \,\Big] \geqslant c > 0.$

(2) $\mathbf{P}[\, |\mathscr{S}_n(r)| = k \,] \leqslant g(k)\, \mathbf{P}[\, |\mathscr{S}_n(r)| = 1 \,]$, with sub-exponential $g(k)$.

We could repeat (1') for many r-boxes only if "not enough points in one box" meant "we found nothing in that box".

So, take an independent random dilute sample: $\mathbf{P}[\, x \in \mathcal{R} \,] = 1/\mathbf{E}|\mathscr{S}_r|$ i.i.d.
Then, $|\mathscr{S}_n \cap B_r/3|$ is small $\implies$ $\mathcal{R} \cap \mathscr{S}_n \cap B_r/3 = \emptyset$ is likely,
and $|\mathscr{S}_n \cap B_r/3|$ is large $\implies$ $\mathcal{R} \cap \mathscr{S}_n \cap B_r/3 \neq \emptyset$ is likely.

So, what are we going to do?

With quite a lot of work for both items,

(1') $\mathbf{P}\Big[\,|\mathscr{S}_n \cap B_r/3| > c\,\mathbf{E}|\mathscr{S}_r| \;\Big|\; \mathscr{S}_n \cap B_r \neq \emptyset = \mathscr{S}_n \cap W\,\Big] \geqslant c > 0.$

(2) $\mathbf{P}\big[\,|\mathscr{S}_n(r)| = k\,\big] \leqslant g(k)\,\mathbf{P}\big[\,|\mathscr{S}_n(r)| = 1\,\big]$, with sub-exponential $g(k)$.

We could repeat (1') for many r-boxes only if "not enough points in one box" meant "we found nothing in that box".

So, take an independent random dilute sample: $\mathbf{P}[\,x \in \mathcal{R}\,] = 1/\mathbf{E}|\mathscr{S}_r|$ i.i.d.
Then, $|\mathscr{S}_n \cap B_r/3|$ is small $\implies \mathcal{R} \cap \mathscr{S}_n \cap B_r/3 = \emptyset$ is likely,
and $|\mathscr{S}_n \cap B_r/3|$ is large $\implies \mathcal{R} \cap \mathscr{S}_n \cap B_r/3 \neq \emptyset$ is likely.

But $\mathbf{P}\Big[\,\mathscr{S}_n \neq \emptyset = \mathcal{R} \cap \mathscr{S}_n \;\Big|\; |\mathscr{S}_n(r)| = k\,\Big]$ is still problematic conditioning.

A strange large deviations lemma solves the issue.

The strange large deviation lemma

Suppose $X_i, Y_i \in \{0,1\}$, $i = 1, \dots, n$, and that $\forall J \subset [n]$ and $\forall i \in [n] \setminus J$

$$\mathbf{P}\big[\, Y_i = 1 \,\big|\, \forall_{j \in J} Y_j = 0 \,\big] \geqslant c\, \mathbf{P}\big[\, X_i = 1 \,\big|\, \forall_{j \in J} Y_j = 0 \,\big].$$

Then

$$\mathbf{P}\Big[\forall_i Y_i = 0\Big] \leqslant c^{-1}\, \mathbf{E}\Big[\exp\Big(-(c/e) \sum_i X_i\Big)\Big].$$

We use this with $X_j := 1_{\{\mathscr{S} \cap B_j \neq \emptyset\}}$ and $Y_j := 1_{\{\mathscr{S} \cap B_j \cap \mathcal{R} \neq \emptyset\}}$.

Proof: Instead of sequential scan, average everything together.
Choose $J \subset [n]$ randomly, Bernoulli$(1-p)$. Get $\mathbf{E}\big[\, Y\, p^Y \,\big] \geqslant c\, \mathbf{E}\big[\, X\, p^{Y+1} \,\big]$.

So, $\mathbf{E}[\, Z \,] \geqslant 0$, where $Z := (Y - c\, p\, X)\, p^Y$. Choose $p := e^{-1}$. Maximize Z over Y, and get the bound $Z \leqslant \exp(-1 - c\, X/e)$. Altogether, $c\, e^{-1}\, \mathbf{P}\big[\, Y = 0 < X \,\big] \leqslant \mathbf{E}\big[\, 1_{X>0} \exp(-1 - c\, X/e) \,\big]$, and done.

Final result for the spectral sample

If $r \in [1, n]$, then $\{|\mathscr{S}_n| < \mathbf{E}|\mathscr{S}_r|\}$ is basically equivalent to being contained inside some $r \times r$ sub-square:

$$\mathbf{P}\big[\, 0 < |\mathscr{S}_n| < \mathbf{E}|\mathscr{S}_r| \,\big] \asymp \alpha_4(r,n)^2 \left(\frac{n}{r}\right)^2.$$

In particular, on the triangular lattice Δ,

$$\mathbf{P}\big[\, 0 < |\mathscr{S}_n| < \lambda\, \mathbf{E}|\mathscr{S}_n| \,\big] \asymp \lambda^{2/3}.$$

The *scaling limit* of $\mathscr{S}_n$ is a conformally invariant Cantor-set with Hausdorff-dimension $3/4$.

GPS (2010-11) proves that the scaling limit of dynamical percolation exists as a Markov process; for mesh $1/n$ the time-scale is $tn^{-3/4+o(1)}$. The above implies that this process is ergodic, with correlations decaying as $t^{-2/3}$.

Question 1: Can one build similar proofs for other Boolean functions?

Question 2: Self-similarity of Piv_n and $\mathscr{S}_n$ is a lot of restriction. The entropy of such random sets X_n should be at most $\mathbf{E}|X_n|$, i.e., there is no $\log$ factor as it would be in uniform:

Fractal percolation on a b-ary tree, by always choosing j random children, to depth h. This is uniform measure on

$$\binom{b}{j}\binom{b}{j}^{j}\cdots\binom{b}{j}^{j^{h-1}} = \binom{b}{j}^{\frac{j^h-1}{j-1}}$$

subsets, so entropy is $\mathrm{const}(b,j)\cdot j^h$, while size is j^h.

In particular, Influence-Entropy conjecture [Friedgut-Kalai 1996]: For some universal constant C, for any Boolean function f,

$$\mathrm{SpecEnt}(f) := \sum_{S\subset[n]} \widehat{f}(S)^2 \log\frac{1}{\widehat{f}(S)^2} \leqslant C \times$$

$$\times\ \mathrm{Influence}(f) := \mathbf{E}|\mathscr{S}_f| = \mathbf{E}|\mathrm{Piv}_f| = \sum_{S\subset[n]} \widehat{f}(S)^2 |S|\,.$$

I think I can do it for Piv_n, but not enough independence is known in $\mathscr{S}_n$.

Gil Kalai's motivating example: Recursive 3-Wise Majority:

Pivotal set is the leaves of a Galton-Watson tree with offspring distribution $\mathbf{P}\big[\,\pi = 0\,\big] = 1/4$ and $\mathbf{P}\big[\,\pi = 2\,\big] = 3/4$.

Spectral sample is the leaves of a Galton-Watson tree with offspring distribution $\mathbf{P}\big[\,\sigma = 1\,\big] = 3/4$ and $\mathbf{P}\big[\,\sigma = 3\,\big] = 1/4$.

Note that $\mathbf{E}[\,\pi\,] = \mathbf{E}[\,\sigma\,] = 3/2$ and $\mathbf{E}[\,\pi^2\,] = \mathbf{E}[\,\sigma^2\,] = 3$.

General Boolean functions?

Random Restriction Lemma + Strong Separation Lemma suggest that typical random restriction of a large Boolean function might look "generic". And then could continue recursively, to get tree-like structure (except that we don't have enough independence. . .). Is that naive?

Two results in similar directions:

Szemerédi Regularity Lemma '75: large dense graphs look random.

Chatterjee-Ledoux '09: If M is a large Hermitian matrix, and k is large, then the spectral measure of almost all principal submatrices of M of order k is almost the same (but depends on M, of course).