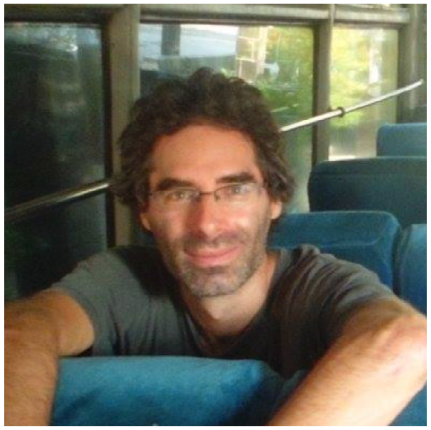
Sparse reconstruction of Boolean functions in spin systems

Gábor Pete
Alfréd Rényi Institute of Mathematics & Technical University, Budapest
http://www.math.bme.hu/∼gabor

Joint work with Pál Galicza
Rényi Institute



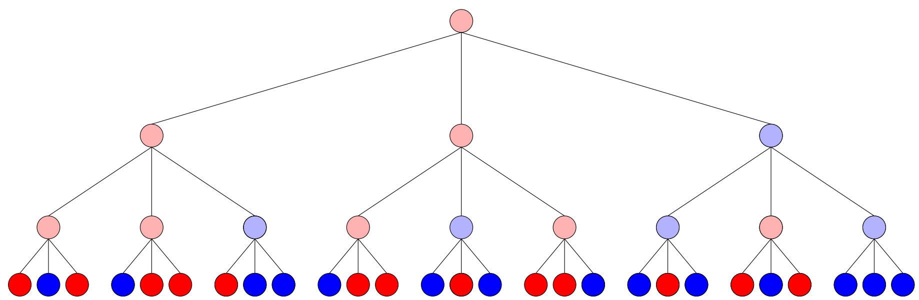
[arXiv:2010.10483], and
in preparation

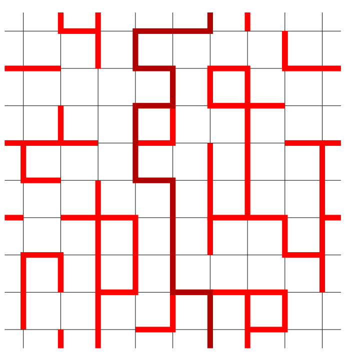
Guessing the output from partial input

A transitive Boolean function $f : \{-1,1\}^n \longrightarrow \{-1,1\}$ is given (so that every bit has the same role), with iid fair random input bits.

Example 1: $\mathbf{Maj_n}(\omega_1, \ldots, \omega_n) := \mathrm{sign} \sum_{i=1}^n \omega_i$, with an odd n.

Example 2:
Iterated 3-majority
on $n = 3^k$ bits.





Example 3: In **critical percolation** on the torus $\mathbb{Z}_k^2$, is there a non-contractible cycle?

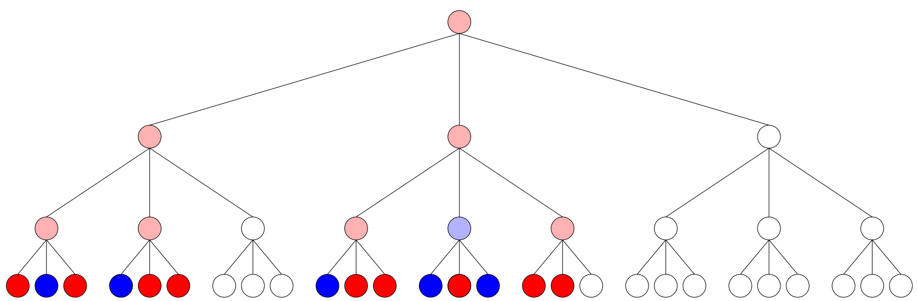
Q: Is there a *small* subset $U \subset [n]$ s.t. from ω_U we can guess the output?

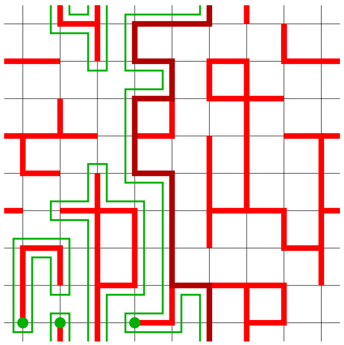
Guessing the output from partial input

A transitive Boolean function $f : \{-1,1\}^n \longrightarrow \{-1,1\}$ is given (so that every bit has the same role), with iid fair random input bits.

Example 1: $\mathbf{Maj_n}(\omega_1, \ldots, \omega_n) := \operatorname{sign} \sum_{i=1}^n \omega_i$, with an odd n. NO!

Example 2:
Iterated 3-majority
on $n = 3^k$ bits.
YES! $\approx (5/2)^k$ bits.





Example 3: In **critical percolation** on the torus $\mathbb{Z}_k^2$, is there a non-contractible cycle?
YES! Exploration interfaces have length $k^{2-\delta}$.

Q: Is there a *small* subset $U \subset [n]$ s.t. from ω_U we can guess the output?
One answer: for some functions *yes*, if we can choose U *adaptively*.

Guessing the output from partial input

Adaptive algorithms computing the output by asking few bits, possibly using extra randomness (also called *randomized decision trees*), have been used by P. Hajnal (1991), O'Donnell, Saks, Schramm & Servedio (2005), Schramm & Steif (2010), Duminil-Copin, Raoufi & Tassion (2019), . . .

Itai Benjamini: what if U has to be given *in advance*? Are there transitive functions whose value can be reconstructed from a vanishingly small subset?

$$\frac{|U_n|}{n} \to 0, \text{ but } \operatorname{Corr}\Big[f_n(\omega),\ \mathbf{E}\big[\, f_n(\omega) \bigm| \omega_{U_n} \big] \Big] \not\to 0\,, \text{ or even } \to 1\,?$$

Version not requiring transitivity: are there *any* functions f_n for which exist *random* subsets $\mathcal{U}_n \subseteq [n]$ with *small revealment* $\delta_{\mathcal{U}} := \sup_{j\in[n]} \mathbf{P}\big[\, j \in \mathcal{U}_n \,\big] \to 0$, but high expected correlation?

If a transitive function f_n has a small U_n, then it also has a low revealment random $\mathcal{U}_n$: just take a uniform random translate of U_n.

No sparse reconstruction for iid bits

Theorem (Galicza & P). No sparse reconstruction for any transitive f. Also, no random sparse reconstruction for any f.

Proof. Fourier spectrum! $\widehat{f}(S)^2 := \mathbf{E}\big[\, f(\omega)\, \chi_S(\omega)\,\big]$, $\chi_S(\omega) := \prod_{i \in S} \omega_i$.

Spectral sample: $\mathbf{P}\big[\,\mathscr{S}_f = S\,\big] := \widehat{f}(S)^2/\|f\|^2$, used by Garban, Pete & Schramm (2010) for noise sensitivity of critical planar percolation.

Proof for transitive f:

$$\text{clue}(f \mid U) := \frac{\mathrm{Var}(\mathbf{E}[\, f \mid \omega_U\,])}{\mathrm{Var}(f)} = \frac{\sum_{\emptyset \neq S \subseteq U} \widehat{f}(S)^2}{\sum_{\emptyset \neq S \subseteq [n]} \widehat{f}(S)^2}$$
$$= \mathbf{P}\big[\,\mathscr{S}_f \subseteq U \;\big|\; \mathscr{S}_f \neq \emptyset\,\big] \leqslant \tilde{\mathbf{P}}[X_f \in U]\,,$$

where X_f is a uniform random element of $\mathscr{S}_f$ conditioned to be non-empty.

$$\tilde{\mathbf{P}}[X_f \in U] = \sum_{j \in U} \tilde{\mathbf{P}}[X_f = j] = \frac{|U|}{n}\,. \qquad \square$$

Entropy proof of small clue

Entropy: $H(X) := -\sum_x \mathbf{P}[X = x] \log \mathbf{P}[X = x]$. Mutual information:

$$I(X,Y) := H(X) + H(Y) - H(X,Y) = H(X) - H(X \mid Y)\,.$$

Information-theoretic clue:

$$\mathsf{clue}^I(f \mid U) := \frac{I(f(\omega), \omega_U)}{H(f(\omega))}\,.$$

For non-degenerate Boolean f, this is small exactly when clue$(f \mid U)$ is.

Entropy proof of small clue

Entropy: $H(X) := -\sum_x \mathbf{P}[X = x] \log \mathbf{P}[X = x]$. Mutual information:

$$I(X, Y) := H(X) + H(Y) - H(X, Y) = H(X) - H(X \mid Y) .$$

Theorem (Galicza & P). For any transitive function f,

$$\text{clue}^I(f \mid U) := \frac{I(f(\omega), \omega_U)}{H(f)} \leqslant \frac{|U|}{n} .$$

Proof. Shearer's inequality: if $X_1, \dots, X_n$ are random variables with any joint distribution, and $\{U_j\}$ is a k-cover of $[n]$, then

$$\sum_j H(X_{U_j}) \geqslant k\, H(X_{[n]}) .$$

Cultural remarks:

Follows from submodularity: $H(X_{S \cup T}) + H(X_{S \cap T}) \leqslant H(X_S) + H(X_T)$, implies Loomis-Whitney $\mathbb{Z}^d$ isoperimetric inequality: $|A|^{d-1} \leqslant \prod_{j=1}^d |\pi_i(A)|$

Entropy proof of small clue

Theorem (Galicza & P). For any transitive function f,

$$\text{clue}^I(f \mid U) := \frac{I(f(\omega), \omega_U)}{H(f)} \leqslant \frac{|U|}{n}.$$

Proof. From Shearer's inequality, for n translates of U, forming a $|U|$-cover,

$$-\sum_j H(\omega_{U_j} \mid f(\omega)) \leqslant -|U|\, H(\omega_{[n]} \mid f(\omega)).$$

On the other hand, for independent variables:

$$\sum_j H(\omega_{U_j}) = \sum_j \sum_{i \in U_j} H(\omega_i) = |U|\, H(\omega_{[n]}),$$

Altogether:

$$n\, I(f(\omega), \omega_U) = \sum_j I(f(\omega), \omega_{U_j}) \leqslant |U|\, I(f(\omega), \omega_{[n]}) = |U|\, H(f).$$

$\square$

Clue and cooperative game theory

Why do we have the same bound $|U|/n$ for two different notions of clue?

Theorem (Galicza & P). For any notion of $\text{clue}(f \mid U)$ that is supermodular (e.g., the L^2-clue and clue^I), and $\text{clue}(f \mid [n]) = 1$ and $\text{clue}(f \mid \emptyset) = 0$, the bound $\text{clue}(f \mid U) \leqslant |U|/n$ holds for any transitive f.

Proof. Consider X_f distributed according to the Shapley value of the cooperative game with payoff $\text{clue}(f \mid U)$. $\square$

What happens for non-iid spins?

G_n finite transitive graphs, often $G_n \to G$ locally to an infinite graph.
$\sigma \in \{-1,+1\}^{V(G_n)}$ translation invariant Markov random field.

E.g., the Ising model at inverse temperature $\beta \in (0,\infty)$:

$$\mu_\beta^{G_n}(\sigma) := \frac{1}{Z_\beta^{G_n}} \exp\Big(-\beta \sum_{(x,y)\in E(G_n)} \mathbf{1}_{\sigma(x)\neq\sigma(y)} \Big) .$$

Subcritical phase, $\beta < (1-\epsilon)\beta_c(G_n)$: correlations decay fast with distance.
Total magnetization $M_n(\sigma) := \sum_{x\in V(G_n)} \sigma(x)$ has $\mathbf{SD}[M_n] \asymp \sqrt{|V(G_n)|}$.
If $G_n \to G$, often $\beta_c(G_n) \to \beta_c(G)$, and unique Gibbs measure on G.

Supercritical phase, $\beta > (1+\epsilon)\beta_c(G_n)$: correlations do not decay.
Long range order: $M_n(\sigma) \asymp \pm|V(G_n)|$ typically.
More than one Gibbs measure on limiting infinite graph G.

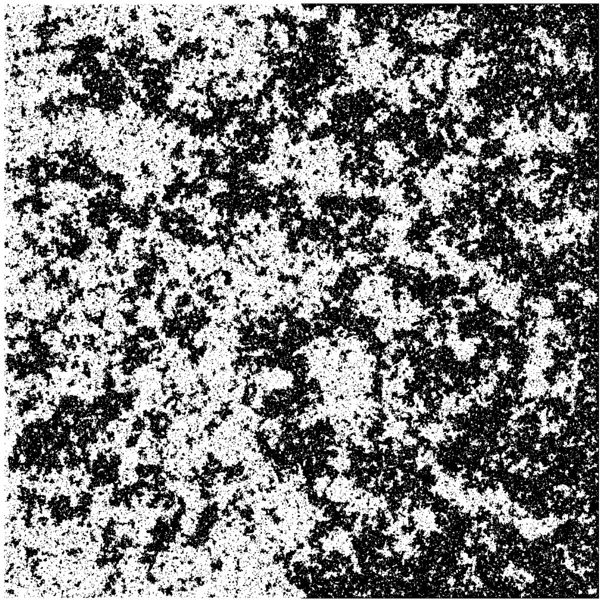
Critical phase, $\beta \sim \beta_c(G_n)$, **typically**: correlations decay, but not fast.
$\sqrt{|V(G_n)|} \ll \mathbf{SD}[M_n] \ll |V(G_n)|$.
Unique Gibbs measure on limiting infinite graph G.

What happens for non-iid spins?

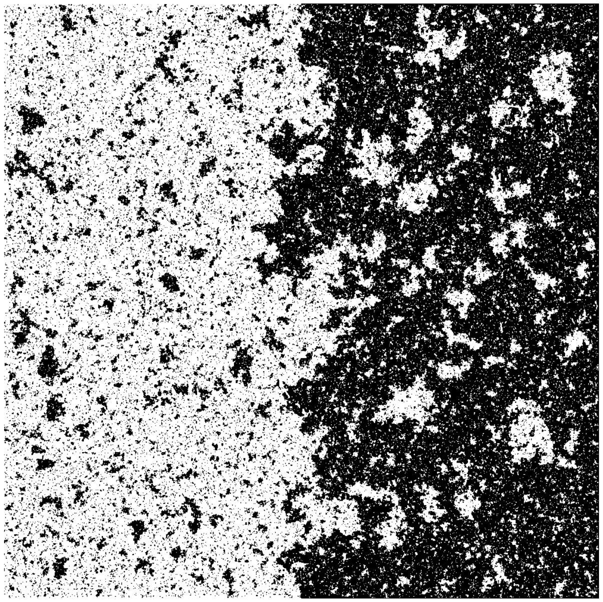
G_n finite transitive graphs, often $G_n \to G$ locally to an infinite graph.
$\sigma \in \{-1,+1\}^{V(G_n)}$ translation invariant Markov random field.

E.g., the Ising model at inverse temperature $\beta \in (0,\infty)$:

$$\mu_\beta^{G_n}(\sigma) := \frac{1}{Z_\beta^{G_n}} \exp\Big(-\beta \sum_{(x,y)\in E(G_n)} \mathbf{1}_{\sigma(x)\neq\sigma(y)} \Big).$$



$\beta = 0.881374$



$\beta = 0.9$

Sparse reconstruction in the supercritical phase

Low temperature Ising, $\beta > \beta_c(\mathbb{Z}^d)$. Then $\mu_\beta^{\mathbb{Z}_n^d}$ converges weakly to $(\mu_\beta^+ + \mu_\beta^-)/2$, a non-ergodic measure.

Moreover, with probability $1/2$, the finite system looks locally like μ_β^+, and with probability $1/2$, looks like μ_β^-.

So, sparse reconstruction is easy: if $|U_n| \to \infty$, then $\operatorname{sign} \sum_{x \in U_n} \sigma(x)$ tells us with large probability if we are in μ_β^+ or μ_β^-, hence has clue close to 1 about $\mathrm{Maj}(\sigma) := \operatorname{sign} \sum_{x \in \mathbb{Z}_n^2} \sigma(x)$.

Similar argument for measures on expander graphs G_n with a non-ergodic limit. However, if the limit is ergodic but non-extremal?

Lemma (Lanford & Ruelle '69). For Markov fields, non-extremal $\Leftrightarrow$ not tail-trivial $\Leftrightarrow$ spin reconstruction from a large distance.

E.g., the unique automorphism-invariant random perfect matching on $\mathbb{T}_3$ is ergodic, but non-extremal. Does the approximating random matching on the 3-regular random graph have (random) sparse reconstruction?

No sparse reconstruction for subcritical Curie-Weiss

Ising model on the complete graph K_n. Scale β with n:

$$\mu_\beta^{K_n}(\sigma) := \frac{1}{Z_\beta^{K_n}} \exp\Big(- \frac{\beta}{n} \sum_{(x,y)\in E(K_n)} \sigma(x)\sigma(y) \Big)\,.$$

Quite analogously to the Erdős-Rényi random graph (via the FK random cluster representation), phase transition at $\beta_c = 1$.

For $\beta < \beta_c = 1$, one has $\frac{M_n}{\sqrt{n}} \xrightarrow{d} N\left(0, \frac{1}{1-\beta}\right)$, even a Local CLT.

This can be used to prove that $H(\sigma_{[n]}^\beta) \geqslant n - C_\beta$, with $\log_2$-entropy. Then, in the proof with Shearer's inequality,

$$\sum_j H(\sigma_{U_j}) \leqslant \sum_j \sum_{i\in U_j} H(\sigma_i) = k\,n \leqslant k\big(H(\sigma_{[n]}) + C_\beta\big)\,,$$

and we get $\mathsf{clue}^I(f \,|\, U) \leqslant \frac{|U|}{n}\left(1 + \frac{C_\beta}{H(f)}\right) \to 0.$ $\square$

Spectral sample for non-iid spins?

Can we define a random set $\mathscr{S} = \mathscr{S}_f$, based on clue?

$$\mathbf{P}[\, \mathscr{S} \subseteq U \,] := \left\| \mathbf{E}[\, f \mid \sigma_U \,] \right\|^2 ,$$

and then inclusion-exclusion formula:

$$\mathbf{P}[\, \mathscr{S} = S \,] := \sum_{T \subseteq S} (-1)^{|S|-|T|} \, \mathbf{P}[\, \mathscr{S} \subseteq T \,] \,.$$

Eigenfunctions of Glauber dynamics are typically not indexed by subsets of bits, hence this would be a different generalization of Fourier transform.

Issue: why would this be non-negative for all S?

Efron-Stein decomposition '81: works for arbitrary product measures!

Hence the one-line Small Clue Theorem works.
And this can be used for non-iid!

Ising as a factor of iid

A spin system σ on $\{-1,+1\}^{\mathbb{Z}^d}$ is a **factor of iid** if there is a measurable map $\psi : [0,1]^{\mathbb{Z}^d} \longrightarrow \{-1,+1\}$ such that for $\omega \sim \mathrm{Unif}[0,1]^{\mathbb{Z}^d}$,

$$\sigma(x) = \psi\big(\omega(x+\cdot)\big)\,, \quad x \in \mathbb{Z}^d\,.$$

This factor map is **finitary** if there is a random coding radius $R(\omega) < \infty$ such that $R(\omega)$ and $\psi(\omega)$ are determined by $\big\{\omega(x) : x \in [-R,R]^d\big\}$.

Theorem (vdBerg & Steif '99). For $\beta < \beta_c$, the unique **Ising** measure on $\mathbb{Z}^d$ is a finitary factor of $\mathrm{Unif}[0,1]^{\mathbb{Z}^d}$, coding radius $\mathbf{P}[\,R > t\,] < \exp(-ct)$.

(Uses exponential convergence of Glauber dynamics, Martinelli & Olivieri '94, and "Coupling From The Past" perfect sampling Propp & Wilson '96).

At β_c: finitary factor, but only with $\mathbf{E}[\,R^d\,] = \infty$ (joint with Peres).

For $\beta > \beta_c$: $+$ measure is fiid, but not finitary (uses Marton & Shields '94).

Small clue for FFIID with exponential decay

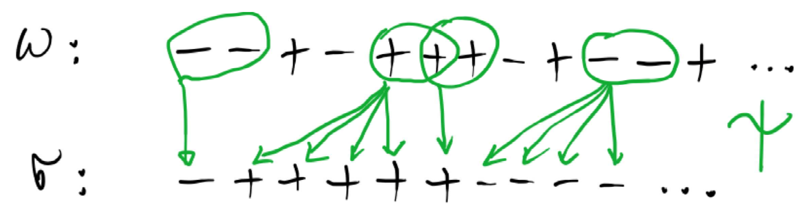
Theorem (Galicza & P). If σ is a finitary factor of iid on $\mathbb{Z}^d$ with $\mathbf{P}[\,R > t\,] < \exp(-ct)$, and σ_n is any version on the torus $\mathbb{Z}_n^d$, then, for any function f_n of the spins, and any random subset with revealment $\delta_{\mathcal{U}_n} = o(1/\log^d n)$, independent of σ_n, we have

$$\mathbf{E}\big[\,\mathsf{clue}(f_n \,|\, \mathcal{U}_n)\,\big] := \mathbf{E}\left[\frac{\mathrm{Var}(\mathbf{E}[\,f \mid \sigma_{\mathcal{U}_n}\,])}{\mathrm{Var}(f_n)}\right] \to 0\,.$$

Proof sketch. Take $\mathcal{W}_n := \bigcup_{u \in \mathcal{U}_n} B_{C \log n}(u)$, with C large enough.
Then $\omega_{\mathcal{W}_n}$ determines $\sigma_{\mathcal{U}_n}$ with high probability.
But the revealment $\delta_{\mathcal{W}_n}$ on ω is still small, hence the clue is small. $\square$

Seems wasteful, because $B_{C \log n}(u)$ is the worst case for each u. However:

Example of sharpness.
2^n iid copies of ψ on $\mathbb{Z}_n \times \mathbb{Z}_{2^n}$.
$f_n = \mathbf{1}\{\exists \text{ alternating} \pm \text{copy}\}$.
$\mathcal{U}_n$: 3 consecutive spins in each copy.



Sparse reconstruction in critical Ising

Critical Ising on $\mathbb{Z}^d$ is not a ffiid with finite expected coding volume, because the **susceptibility** $S_\beta := \sum_{x\in\mathbb{Z}^d} \mathrm{Cov}_\beta[\sigma_0, \sigma_x]$ is infinite at $\beta = \beta_c$. (Aizenman & Fernandez '86, Aizenman, Duminil-Copin & Sidoravicius '13)

Theorem (Galicza & P). On the tori $\mathbb{Z}_n^d$, $d \geqslant 2$, at β_c, the total magnetization $M_n(\sigma) := \sum_{x\in\mathbb{Z}_n^d} \sigma_x$ can be reconstructed from some low revealment subset $\mathcal{U}_n$. Also true for $\mathsf{Maj}_n(\sigma) := \mathrm{sign}\, M_n(\sigma)$.

Proof sketch. $\mathrm{Var}_{\beta_c}[M_n(\sigma)] = n^d S_{\beta_c,n}$, and $S_{\beta_c,n} \to S_{\beta_c} = \infty$, hence $1 \gg \delta_n \gg 1/S_{\beta_c,n}$ will make $\mathrm{Corr}_{\beta_c}\left[M_n(\sigma), \sum_{u\in\mathcal{B}(\delta_n)} \sigma_u\right] \to 1$. $\square$

In particular, on $\mathbb{Z}_n^2$, revealment $\delta_{\mathcal{U}_n} \gg n^{-7/4}$ is enough for magnetization.

On the other hand, $\delta_{\mathcal{U}_n} \ll n^{-15/8}$ is *not* enough for *any odd* function. Simple but inspiring proof for magnetization by Christophe Garban:

Couple σ_n and $\tilde\sigma_n$ by sampling an FK-representation ω_n, then same $\pm$ spins on ω_n-clusters intersecting $\mathcal{U}_n$, while independent spins on the other clusters. Thus $\sigma_{\mathcal{U}_n} = \tilde\sigma_{\mathcal{U}_n}$, but $\mathrm{Cov}[M_n(\sigma_n), M_n(\tilde\sigma_n)] \leqslant |\mathcal{U}_n|\, \mathbf{E}\big[\, |\mathsf{Cluster}_o|^2 \,\big]$.

From strong spatial mixing to sparse reconstruction

A Markov random field $\{-1,1\}^{\mathbb{Z}^d}$ has **strong spatial mixing** if for any finite box V, given two boundary configurations $\sigma_{\partial V}$ and $\tilde{\sigma}_{\partial V}$ that differ only at a single vertex $v \in \partial V$, for any radius R, the conditional distributions inside V satisfy

$$d_{\mathrm{TV}}\left(\sigma_{V \setminus B_R(v)},\ \tilde{\sigma}_{V \setminus B_R(v)}\right) \leqslant \exp(-cR).$$

Ising on $\mathbb{Z}^d$, $d \leqslant 2$, all $\beta < \beta_c$, Martinelli, Olivieri & Schonmann '94 and Alexander '98 together imply SSM. For $d \geqslant 3$, small enough β, follows from Stroock & Zegarlinski '92 or Marton '19.

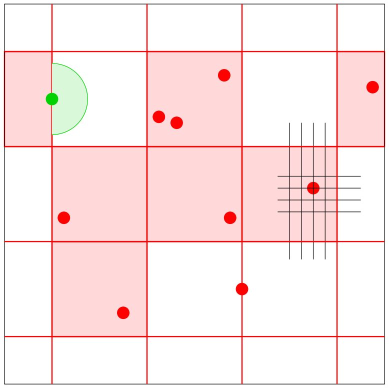
Blanca, Caputo, Sinclair & Vigoda '19 proved that SSM implies a uniformly positive **spectral gap** for certain block dynamics (e.g., for Swendsen-Wang). Inspired by this, and the previous proof of Christophe:

Theorem (Galicza & P). For any SSM Markov field on $\{-1,1\}^{\mathbb{Z}^d}$, for any function f_n and any random subset with revealment $\delta_{\mathcal{U}_n} \to 0$, independent of σ_n, we have $\mathbf{E}\big[\,\mathrm{clue}(f_n \,|\, \mathcal{U}_n)\,\big] \to 0$.

From strong spatial mixing to sparse reconstruction

Proof sketch. Block Glauber dynamics $\sigma \mapsto \sigma^{\mathcal{U}}$, where $\mathcal{U}$ is sampled, then $\sigma_{\mathcal{U}}$ gets fixed and $\sigma_{\mathbb{Z}_n^d \setminus \mathcal{U}}$ gets resampled from the conditional distribution. Enough to prove that spectral gap of this chain is close to 1 if $\delta_{\mathcal{U}}$ is small.

Take $1 \ll L := (1/\delta)^{1/(d+1)} \log^{d/(d+1)}(1/\delta) \ll n$, let $\mathcal{L}$ be a randomly shifted hyperplane sublattice of mesh size L, and let $\mathcal{H}$ be $\mathcal{L}$ together with all the inner boxes that intersect $\mathcal{U}$.



Enough: spectral gap of $\sigma \mapsto \sigma^{\mathcal{H}}$ is close to 1.

Path coupling method of Bubley & Dyer '97: whenever σ and $\tilde{\sigma}$ differ only at a single vertex $v \in \mathbb{Z}_n^d$, if $\mathbf{E}\left[d_{\text{Hamming}}\left(\sigma^{\mathcal{H}}, \tilde{\sigma}^{\mathcal{H}}\right) \right] < \epsilon$, then OK.

If $v \in \mathcal{H}^\circ$, with prob $O(\delta L^d)$, error remains 1.
If $v \in \partial\mathcal{H}$, with prob $O(1/L)$, error propagates to $O(R^d + e^{-cR} L^d)$ spins.
If $v \in \mathbb{Z}_n^d \setminus \mathcal{H}$, error becomes 0.
Choose R with $\log L \ll R \ll L^{1/d}$. $\square$

Open problems on sparse reconstruction

1. Subcritical Ising on $\mathbb{Z}_n^d$, $d \geqslant 3$, all $\beta < \beta_c$: shave off the $\log^d n$.

2. Critical Ising $\mathbb{Z}_n^2$: what is the exact sparse reconstruction threshold?

3. In reasonable spin systems, if total magnetization cannot be sparse reconstructed (finite susceptibility), then nothing can?

For instance, if σ is a finitary fiid system with finite expected coding volume, then susceptibility is finite. If σ_n is a sequence of finite systems such that, with probability tending to 1, at every vertex the finitary factor works, then there is never sparse reconstruction in σ_n?

4. Possible example for sparse reconstruction but not for magnetization: $+$ phase of supercritical Ising on $\mathbb{Z}_n^2$.

5. Balázs Szegedy: does every fiid system have a trivial sparse tail? Would imply no sparse reconstruction for local functions. (E.g., in the almost perfect matching example earlier.)

True for amenable transitive graphs, by using entropy.