Concentration Inequalities for Gibbs Random Fields

Jean-René Chazottes
CNRS & Ecole Polytechnique, Paris

May 17, 2018
**Introduction**

Gibbs measures are (non-product) measures on the configuration space $\mathcal{S}^{\mathbb{Z}^d}$, $d \geq 2$.

In this talk: $\mathcal{S} = \{-1, +1\}$ (spins) for simplicity but any finite set $\mathcal{S}$ is ok.

**Abstract:**

- At sufficiently high temperature, we have a Gaussian concentration bound. 
  In fact, such a bound holds in Dobrushin’s uniqueness regime.
- For some Gibbs measures at sufficiently low temperature, we have a ‘stretched exponential’ concentration bound.
- These bounds have many consequences.
Boltzmann-Gibbs kernel

$$
\gamma^{(\beta)}_\Lambda(\omega|\eta) = \frac{\exp\left(-\beta \mathcal{H}_\Lambda(\omega|\eta)\right)}{Z^{(\beta)}(\eta)}, \quad \Lambda \subseteq \mathbb{Z}^d, \omega, \eta \in \mathcal{S}^{\mathbb{Z}^d}.
$$

\(\sim\) (DLR equation) Gibbs measures on \(\mathcal{S}^{\mathbb{Z}^d}\) depending on \(\eta\) in general.

Parameter \(\beta \geq 0\): inverse temperature.

**Special case:** \(\beta = 0\) (infinite temperature)

\(\sim\) uniform **product measure** \(\sim \) Gaussian concentration bound).
The ferromagnetic Ising model (Markov random field)

\[ \mathcal{H}_\Lambda(\omega|\eta) = - \sum_{i,j \in \Lambda} \omega_i \omega_j - \sum_{i \in \partial \Lambda, j \notin \Lambda} \omega_i \eta_j \]

\[ \eta_j = +1, \forall j \in \mathbb{Z}^d \text{ ("+-boundary condition"), gives rise to } \mu^+. \]

**Fact** \((d \geq 2)\): there exists a unique Gibbs measure \(\mu\) for all \(\beta < \beta_c\), whereas there are several ones for all \(\beta > \beta_c\), depending on \(\eta\), in fact, two extremal ones: \(\mu^+\) and \(\mu^-\) (i.e., ergodic under the shift action).
Phase transition for $d = 2$

$\beta$ increases from left to right
‘+’ $\leftrightarrow$ black, ‘−’ $\leftrightarrow$ white

$\beta_c = (1/2) \sinh^{-1}(1) \approx 0.4407$
The magnetization

\[ M_n(\omega) := \sum_{i \in C_n} s_0(T_i \omega), \]
where \( s_0(\omega) = \omega_0 \), be the total magnetization in \( C_n \), and where \( (T_i \omega)_j = \omega_{j-i} \) (shift operator). Then

\[ \frac{M_n(\omega)}{(2n + 1)^d} \]

is the magnetization per spin in \( C_n \). For any shift-invariant probability measure \( \nu \) on \( \mathcal{S}^{\mathbb{Z}^d} \),

\[ \mathbb{E}_\nu \left[ \frac{M_n(\omega)}{(2n + 1)^d} \right] = \mathbb{E}_\nu [s_0] \]

is the mean magnetization per site (magnetization, for short) wrt \( \nu \).

The following is well-known for the Ising model \((d \geq 2)\):

- for \( \beta < \beta_c \), \( \mathbb{E}_\mu [s_0] = 0 \);
- for \( \beta > \beta_c \), \( \mathbb{E}_\mu^+ [s_0] \neq 0 \).
CONCENTRATION FOR THE ISING MODEL

Gaussian concentration

Stretched exponential concentration

Uniqueness

Non-uniqueness
Let $F : \mathcal{S}^\mathbb{Z}_d \to \mathbb{R}$ and

$$\ell_i(F) = \sup_{\omega \in \mathcal{S}^\mathbb{Z}_d} |F(\omega(i)) - F(\omega)|, \ i \in \mathbb{Z}_d,$$

where $\omega(i)$ is obtained from $\omega$ by flipping the spin at $i$.

**Theorem:** Gaussian concentration bound ($\beta < \underline{\beta}$)

Let $\mu$ be the (unique) Gibbs measure for the Ising model. There exists a constant $D > 0$ such that, for all functions $F$ with

$$\sum_{i \in \mathbb{Z}_d} \ell_i(F)^2 < +\infty,$$

one has

$$\mathbb{E}_\mu\left[ \exp(F - \mathbb{E}_\mu(F)) \right] \leq \exp \left( D \sum_{i \in \mathbb{Z}_d} \ell_i(F)^2 \right).$$

**Remark.** As shown by C. Külske, the Gaussian concentration bound holds in the Dobrushin uniqueness regime with

$$D = 2(1 - c(\gamma))^{-2},$$

where $c(\gamma)$ is Dobrushin’s contraction coefficient.
Recall that the Gaussian concentration implies that for all $u \geq 0$ one has

$$
\mu \left( \omega \in \mathcal{L}^{\mathbb{Z}^d} : |F(\omega) - \mathbb{E}_\mu[F]| \geq u \right) \leq 2 \exp \left( - \frac{u^2}{4D \sum_{i \in \mathbb{Z}^d} \ell_i(F)^2} \right).
$$

**Remark.** All local functions satisfy $\sum_{i \in \mathbb{Z}^d} \ell_i(F)^2 < +\infty$. 
At sufficiently low temperature, we can gather all moment bounds to obtain the following. We denote by $\mu^+$ the Gibbs measure for the +--phase of the Ising model.

**Theorem: Stretched-exponential concentration bound ($\beta > \bar{\beta}$)**

There exists $\varrho = \varrho(\beta) \in (0, 1)$ and $c_\varrho > 0$ such that for all functions $F$ with $\sum_{i \in \mathbb{Z}^d} \ell_i(F)^2 < +\infty$, for all $u \geq 0$, one has

$$\mu^+\left(\omega \in \mathcal{L}^{\mathbb{Z}^d} : |F(\omega) - \mathbb{E}_{\mu^+}[F]| \geq u\right) \leq 4 \exp\left(\frac{-c_\varrho u^\varrho}{\left(\sum_{i \in \mathbb{Z}^d} \ell_i(F)^2\right)^{\frac{\varrho}{2}}} \right).$$
The basic ingredients in proofs

Enumeration of $\mathbb{Z}^d$:

$$e : \mathbb{Z}^d \to \mathbb{N}$$

$$(\leq i) := \{j \in \mathbb{Z}^d : e(j) \leq e(i)\}$$

$\mathcal{F}_{\leq i} : \sigma -$-field generated by $\omega_j, j \leq i$

We have $F - \mathbb{E}[F] = \sum_{i \in \mathbb{Z}^2} \Delta_i, \quad \Delta_i := \mathbb{E}[F | \mathcal{F}_{\leq i}] - \mathbb{E}[F | \mathcal{F}_{< i}]$

and

$$\Delta_i \leq (D_{i,j}^{\omega \leq i} \ell(F))_i$$

where $D_{i,j}^{\omega \leq i} := \hat{P}_{i,+,-}(\omega_j^{(1)} \neq \omega_j^{(2)})$

where we maximally couple $P(\cdot | \omega_{< i,+})$ and $P(\cdot | \omega_{< i,-})$. 
Other models besides the standard Ising model: Potts, long-range Ising, etc.

- Ergodic sums in *arbitrarily shaped* volumes;
- Fluctuations in the Shannon-McMillan-Breiman theorem;
- First occurrence of a pattern of a configuration in another configuration;
- Bounding $d$-distance by relative entropy;
- Fattening patterns;
- Speed of convergence of the empirical measure;
- Almost-sure central limit theorems.
Take $\Lambda \subseteq \mathbb{Z}^d$ and $\omega \in \mathcal{S}^{\mathbb{Z}^d}$ and let

$$\mathcal{E}_\Lambda(\omega) = \frac{1}{|\Lambda|} \sum_{i \in \Lambda} \delta_{T_i \omega}$$

where $(T_i \omega)_j = \omega_{j-i}$ (shift operator).

Let $\mu$ be an ergodic measure on $\mathcal{S}^{\mathbb{Z}^d}$. If $(\Lambda_n)_n$ is a sequence of cube $\uparrow \mathbb{Z}^d$ (more generally, a van Hove sequence), then

$$\mathcal{E}_{\Lambda_n}(\omega) \xrightarrow{n \to \infty} \text{weakly } \mu.$$

**Question:** If $\mu$ is a Gibbs measure, what is the “speed” of this convergence?
KANTOROVICH DISTANCE on the set of probability measures on $\mathcal{S} \mathbb{Z}^d$:

$$d_{Kanto}(\mu_1, \mu_2) = \sup_{G: \mathcal{S} \mathbb{Z}^d \to \mathbb{R}, \quad G \text{ 1-Lipshitz}} (\mathbb{E}_{\mu_1}(G) - \mathbb{E}_{\mu_2}(G))$$

where $|G(\omega) - G(\omega')| \leq d(\omega, \omega') = 2^{-k}$, where $k$ is the sidelength of the largest cube in which $\omega$ and $\omega'$ coincide.

**Lemma.** Let $\mu$ be a probability measure and

$$F(\omega) = \sup_{G: \mathcal{S} \mathbb{Z}^d \to \mathbb{R}, \quad G \text{ 1-Lipshitz}} \left( \frac{1}{|\Lambda|} \sum_{i \in \Lambda} G(T_i \omega) - \mathbb{E}_{\mu}(G) \right).$$

Then

$$\sum_{i \in \mathbb{Z}^d} \ell_i(F)^2 \leq \frac{c_d}{|\Lambda|}$$

where $c_d > 0$ depends only on $d$. 

Proof
Gaussian concentration for the empirical measure ($\beta < \hat{\beta}$)

Let $\mu$ be the (unique) Gibbs measure of the Ising model. There exists a constant $C > 0$ such that, for all $\Lambda \subseteq \mathbb{Z}^d$ and for all $u \geq 0$, one has

$$
\mu \left\{ \omega \in \mathcal{P}^{\mathbb{Z}^d} : \left| d_{Kanto}(\mathcal{E}_\Lambda(\omega), \mu) - \mathbb{E}_\mu \left[ d_{Kanto}(\mathcal{E}_\Lambda(\cdot), \mu) \right] \right| \geq u \right\} 
\leq 2 \exp \left( - C |\Lambda| u^2 \right).
$$
We denote by $\mu^+$ the Gibbs measure for the $+$-phase of the Ising model.

**Stretched-exponential concentration for the empirical measure ($\beta > \overline{\beta}$)**

There exist $\varrho = \varrho(\beta) \in (0, 1)$ and a constant $c_\varrho > 0$ such that, for all $\Lambda \subseteq \mathbb{Z}^d$ and for all $u \geq 0$, one has

$$
\mu^+ \left\{ \omega \in \mathcal{S}^{\mathbb{Z}^d} : \left| d_{Kanto}(\mathcal{E}_\Lambda(\omega), \mu^+) - \mathbb{E}_{\mu^+}[d_{Kanto}(\mathcal{E}_\Lambda(\cdot), \mu^+)] \right| \geq u \right\} 
\leq 4 \exp \left( -c_\varrho |\Lambda|^{\frac{\varrho}{2}} u^{\frac{\varrho}{\varrho}} \right).
$$
Can we estimate $\mathbb{E}_\mu \left[ d_{Kanto}(\mathcal{E}_\Lambda(\cdot), \mu) \right]$?

Let

$$\mathcal{L} = \{ G : \mathcal{S}^Z \rightarrow \mathbb{R} : G \text{ 1-Lipschitz} \}$$

and

$$Z_G^\Lambda := \frac{1}{|\Lambda|} \sum_{i \in \Lambda} (G \circ T_i - \mathbb{E}_\mu(G)), \Lambda \subseteq \mathbb{Z}^d.$$  

Then

$$\mathbb{E}_\mu \left[ d_{Kanto}(\mathcal{E}_\Lambda(\cdot), \mu) \right] = \mathbb{E}_\mu \left( \sup_{G \in \mathcal{L}} Z_G^\Lambda \right).$$

Notice that we have functions defined on a Cantor space, which is really different from the case of, say, $[0, 1]^k \subset \mathbb{R}^k$. 
Let $\mu$ be a probability measure on $\mathcal{S} \mathbb{Z}^d$ satisfying the Gaussian concentration bound. Then

$$\mathbb{E}_\mu \left[ d_{Kanto} \left( \mathcal{E}_\Lambda (\cdot), \mu \right) \right] \leq \begin{cases} |\Lambda|^{-\frac{1}{2}} (1 + \log |\mathcal{S}|)^{-1} & \text{if } d = 1 \\ \exp \left( -\frac{1}{2} \left( \frac{\log |\Lambda|}{\log |\mathcal{S}|} \right)^{1/d} \right) & \text{if } d \geq 2. \end{cases}$$

For $(a_\Lambda)$ and $(b_\Lambda)$ indexed by finite subsets of $\mathbb{Z}^d$ we denote $a_\Lambda \preceq b_\Lambda$ if, for every sequence $(\Lambda_n)$ such that $|\Lambda_n| \to +\infty$ as $n \to +\infty$, we have $\limsup_n n \log a_{\Lambda_n} / \log b_{\Lambda_n} \leq 1$.

It is possible to get bounds but they are really messy.
Application 2: Almost-sure Central Limit Theorems (only part of the story)

This application shows that one can also get limit theorems out of concentration inequalities.

Informal statement:
If you know that the central limit theorem holds for some function \( f : \mathcal{P}^d \rightarrow \mathbb{R} \) wrt to a shift-invariant probability measure, and if you can prove that this measure satisfies a moment concentration bound of order 2, then the almost-sure central limit theorem holds in the sense of Kantorovich distance.
Given $f : \mathcal{L}^d \rightarrow \mathbb{R}$ and $\nu$ a shift-invariant probability measure on $\mathcal{L}^d$, the usual form of the CLT is: for all $u \in \mathbb{R}$

$$\lim_{n \to \infty} \nu \left\{ \omega \in \mathcal{L}^d : \frac{\sum_{i \in C_n} f(T_i \omega)}{(2n + 1)^{d/2}} \leq u \right\} = G_{0,\sigma^2 f}((\infty, u])$$

where

$$\sigma^2 f = \sum_{i \in \mathbb{Z}^d} \int f \cdot f \circ T_i \, d\nu \in (0, +\infty)$$

and where $G_{0,\sigma^2 f}$ is the Gaussian measure with mean 0 and variance $\sigma_f$. 
The CLT can be re-written as
\[
\lim_{n \to \infty} \mathbb{E}_\nu \left[ 1 \left\{ \sum_{i \in C_n} f(T_i \cdot)/(2n+1)^{d/2} \leq u \right\} \right] = G_0,\sigma_f\left( (-\infty, u] \right).
\]

The ASCLT consists in replacing \( \mathbb{E}_\nu \) by a point-wise logarithmic average and get an almost-sure version of the CLT: for all \( u \in \mathbb{R} \)
\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left[ \sum_{i \in C_n} f(T_i \omega)/(2n+1)^{d/2} \leq u \right] = G_0,\sigma_f\left( (-\infty, u] \right)
\]
for \( \nu \)-a.e. \( \omega \).
We will only formulate two results for $f = s_0$ (magnetization).

To state the theorems, define

$$d_{Kanto}(\nu_1, \nu_2) = \sup (\mathbb{E}_{\nu_1}(g) - \mathbb{E}_{\nu_2}(g))$$

where the sup is taken over all functions $g : \mathbb{R} \to \mathbb{R}$ that are 1-Lipschitz.

Metrizes the weak topology on the set of probability measures on $\mathbb{R}$ with a first moment.
Let $\beta < \underline{\beta}$. Then, for $\mu$-a.e. $\omega \in \mathcal{S}^{\mathbb{Z}^d}$, we have

$$
\lim_{N \to \infty} d_{Kanto} \left( \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \delta_{M_n(\omega)/(2n+1)^{\frac{d}{2}}} , G_0, \sigma^2 \right) = 0
$$

where

$$
\sigma^2 = \sum_{\mathbf{i} \in \mathbb{Z}^d} \int s_0 \cdot s_0 \circ T_i \, d\mu \in (0, \infty).
$$
Let $\beta > \bar{\beta}$. Then, for $\mu^+$-a.e. $\omega \in \mathcal{S}^\mathbb{Z}^d$, we have

$$
\lim_{N \to \infty} d_{\text{Kanto}} \left( \frac{1}{\ln N} \sum_{n=1}^{N} \frac{1}{n} \delta_{(M_n(\omega) - \mathbb{E}_{\mu^+}[s_0])/(2n+1)^{d/2}} , G_0, \sigma^2 \right) = 0
$$

where

$$
\sigma^2 = \sum_{i \in \mathbb{Z}^d} \int s_0 \cdot s_0 \circ T_i \, d\mu^+ \in (0, \infty).
$$
Some open questions

1. ‘Close the gap’ between $\beta$ and $\bar{\beta}$.
2. Write the proof in the low temperature regime in the setting of Pirogov-Sinai theory.
3. Get the optimal $\varrho$ in

$$\exp\left(\frac{-c_\varrho u^\varrho}{\left(\sum_{i \in \mathbb{Z}^d} \ell_i (F)^2\right)^{\frac{\varrho}{2}}}\right).$$


http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik/lehrstuhl-ii/kuelske-pub/autocorrdobrushin-fulltext.pdf
\( \mu \) is a Gibbs measure for a given potential \( \Phi \) if, for all \( \Lambda \subseteq \mathbb{Z}^d \) and for all \( A \in \mathcal{B}(\mathcal{S}\mathbb{Z}^d) \)

\[
\mu(A) = \int d\mu(\eta) \sum_{\omega' \in \Lambda} \gamma_\Lambda(\omega'|\eta) 1_A(\omega'_\Lambda \eta_{\Lambda^c})
\]

where \( \Phi \) is a real-valued function having two arguments: a finite subset of \( \mathbb{Z}^d \) and a configuration \( \omega \in \mathcal{S}\mathbb{Z}^d \), and where

\[
\mathcal{H}_\Lambda(\omega|\eta) = \sum_{\Lambda' \cap \Lambda \neq \emptyset} \Phi(\Lambda', \omega_\Lambda \eta_{\mathbb{Z}^d \backslash \Lambda})
\]

where \( \Lambda' \) runs through the set of finite subsets of \( \mathbb{Z}^d \).
Let

\[ C_{i,j}(\gamma) = \sup_{\omega, \omega' \in \mathcal{Z}^d} \| \gamma\{i\}(\cdot|\omega) - \gamma\{i\}(\cdot|\omega') \|_\infty. \]

Then in our context \( C_{i,j} \) only depends on \( i - j \) and we define

\[ c(\gamma) = \sum_{i \in \mathbb{Z}^d} C_{0,i}(\gamma). \]

Dobrushin’s uniqueness regime: \( c(\gamma) < 1 \).
van Hove sequence

A sequence $(\Lambda_n)_n$ of nonempty finite subsets of $\mathbb{Z}^d$ is said to tend to infinity in the sense of van Hove if, for each $i \in \mathbb{Z}^d$, one has

$$\lim_{n \to +\infty} |\Lambda_n| = +\infty \quad \text{and} \quad \lim_{n \to +\infty} \frac{|(\Lambda_n + i) \setminus \Lambda_n|}{|\Lambda_n|} = 0.$$
Proof of the Lemma

Let $\omega, \omega' \in \mathcal{S}^{\mathbb{Z}^d}$ and $G : \mathcal{S}^{\mathbb{Z}^d} \to \mathbb{R}$ be a 1-Lipschitz function. Without loss of generality, we can assume that $\mathbb{E}_{\mu}(G) = 0$. We have

$$
\sum_{i \in \Lambda} G(T_i \omega) \leq \sum_{i \in \Lambda} G(T_i \omega') + \sum_{i \in \Lambda} d(T_i \omega, T_i \omega').
$$

Taking the supremum over 1-Lipschitz functions thus gives

$$
F(\omega) - F(\omega') \leq \sum_{i \in \Lambda} d(T_i \omega, T_i \omega').
$$

We can interchange $\omega$ and $\omega'$ in this inequality, whence

$$
|F(\omega) - F(\omega')| \leq \sum_{i \in \Lambda} d(T_i \omega, T_i \omega').
$$
Now we assume that there exists $k \in \mathbb{Z}^d$ such that $\omega_j = \omega'_j$ for all $j \neq k$. This means that $d(T_i \omega, T_i \omega') \leq 2^{-\|k-i\|_\infty}$ for all $i \in \mathbb{Z}^d$, whence

$$\ell_k(F) \leq \sum_{i \in \Lambda} 2^{-\|k-i\|_\infty}.$$ 

Therefore, using Young’s inequality,

$$\sum_{i \in \mathbb{Z}^d} \ell_i(F)^2 \leq \sum_{k \in \mathbb{Z}^d} \left( \sum_{i \in \mathbb{Z}^d} 1_{\Lambda}(i) 2^{-\|k-i\|_\infty} \right)^2 \leq \sum_{i \in \mathbb{Z}^d} 1_{\Lambda}(i) \times \left( \sum_{k \in \mathbb{Z}^d} 2^{-\|k\|_\infty} \right)^2.$$

We thus obtain the desired estimate with

$$c_d = \left( \sum_{k \in \mathbb{Z}^d} 2^{-\|k\|_\infty} \right)^2.$$  

$\Box$ Kantorovich distance