Convex decay of Entropy for interacting systems

Paolo Dai Pra

Università degli Studi di Padova

Cambridge ∼ March 30, 2011

Based on:

- BOUDOU A.S., CAPUTO P., D.P., POSTA G. (2006). Spectral gap estimates for interacting particle systems via a Bochner-type identity. JOURNAL OF FUNCTIONAL ANALYSIS, vol. 232; p. 222-25
- CAPUTO P, D.P., POSTA G (2009). Convex Entropy Decay via the Bochner-Bakry-Emery approach. ANNALES DE L'INSTITUT HENRI POINCARE-PROBABILITES ET STATISTIQUES, vol. 45; p. 734- 753
- Some more recent progresses, with G. POSTA.

Convergence to equilibrium

Let π be the stationary distribution of an ergodic, continuous-time, Markov chain on a finite or countable set X, with semigroup $S_t = e^{tL}$. Denote by μS_t the law at time t of the chain started from the law μ.

Convergence to equilibrium

Let π be the stationary distribution of an ergodic, continuous-time, Markov chain on a finite or countable set X, with semigroup $S_t = e^{tL}$. Denote by μS_t the law at time t of the chain started from the law μ.

Denoting by

$$\|\mu - \pi\|_{TV} := \frac{1}{2}\sum_{x \in S} |\mu(x) - \pi(x)|$$

the total variation distance, one of the aims of quantitative ergodic theory is to estimate the rate of convergence of

$$\|\mu S_t - \pi\|_{TV} \to 0$$

as $t \to +\infty$.

It is usually simpler to obtain rates for other modes of convergence.

It is usually simpler to obtain rates for other modes of convergence.

For $f : X \to \mathbb{R}$ set $\pi[f] := \sum_x \pi(x) f(x)$.

It is usually simpler to obtain rates for other modes of convergence.

For $f : X \to \mathbb{R}$ set $\pi[f] := \sum_x \pi(x) f(x)$.

- L^2 convergence:

$$\|S_t - \pi\|_{2\to 2} := \sup\{\|S_t f - \pi[f]\|_2 : \|f\|_2 = 1\}.$$

where $\|f\|_2 := \sqrt{\pi\left[f^2\right]}$.

It is usually simpler to obtain rates for other modes of convergence.

For $f : X \to \mathbb{R}$ set $\pi[f] := \sum_x \pi(x) f(x)$.

- L^2 convergence:

$$\|S_t - \pi\|_{2\to 2} := \sup\{\|S_t f - \pi[f]\|_2 : \|f\|_2 = 1\}.$$

where $\|f\|_2 := \sqrt{\pi\left[f^2\right]}$.

The following bound holds: letting $\pi^* := \min_x \pi(x)$,

$$\sup_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \frac{1}{\pi^*} \|S_t - \pi\|_{2\to 2}.$$

- Convergence in relative entropy: letting

$$h(\mu|\pi) := \sum_x \pi(x) \left(\frac{\mu(x)}{\pi(x)} \log \frac{\mu(x)}{\pi(x)} \right) = \pi \left[\frac{\mu}{\pi} \log \frac{\mu}{\pi} \right].$$

one obtains bounds for

$$\sup_{x \in S} h(\delta_x S_t | \pi).$$

- Convergence in relative entropy: letting

$$h(\mu|\pi) := \sum_x \pi(x) \left(\frac{\mu(x)}{\pi(x)} \log \frac{\mu(x)}{\pi(x)} \right) = \pi \left[\frac{\mu}{\pi} \log \frac{\mu}{\pi} \right].$$

one obtains bounds for

$$\sup_{x \in S} h(\delta_x S_t|\pi).$$

These bounds are then transferred to total variation by Czisar's inequality:

$$\sup_{x \in S} h(\delta_x S_t|\pi) \geq \sup_{x \in S} \|\delta_x S_t - \pi\|_{TV}^2$$

Functional inequalities

We are therefore interested in the rate of convergence to zero of quantities as

$$\|S_t - \pi\|_{2\to 2} \quad \text{and} \quad h(\delta_x S_t|\pi).$$

Functional inequalities

We are therefore interested in the rate of convergence to zero of quantities as

$$\|S_t - \pi\|_{2\to 2} \quad \text{and} \quad h(\delta_x S_t|\pi).$$

A key notion is that of *Dirichlet form*:

$$\mathcal{E}(f,g) := -\pi\left[fLg\right].$$

Functional inequalities

We are therefore interested in the rate of convergence to zero of quantities as

$$\|S_t - \pi\|_{2\to 2} \quad \text{and} \quad h(\delta_x S_t|\pi).$$

A key notion is that of *Dirichlet form*:

$$\mathcal{E}(f,g) := -\pi\,[fLg]\,.$$

If the chain is reversible (i.e. $\pi(x)L_{xy} = \pi(y)L_{yx}$)

$$\mathcal{E}(f,g) = \frac{1}{2}\sum_{y\in S}\pi\,[L_{xy}(f(y)-f(x))(g(y)-g(x))]$$

(true for $g = f$ without reversibility)

Functional inequalities

We are therefore interested in the rate of convergence to zero of quantities as

$$\|S_t - \pi\|_{2\to 2} \quad \text{and} \quad h(\delta_x S_t|\pi).$$

A key notion is that of *Dirichlet form*:

$$\mathcal{E}(f,g) := -\pi\left[fLg\right].$$

If the chain is reversible (i.e. $\pi(x)L_{xy} = \pi(y)L_{yx}$)

$$\mathcal{E}(f,g) = \frac{1}{2}\sum_{y\in S}\pi\left[L_{xy}(f(y)-f(x))(g(y)-g(x))\right]$$

(true for $g = f$ without reversibility)
Also let, for $f > 0$

$$Ent_\pi(f) := \pi[f\log f] - \pi[f]\log\pi[f],$$

so that $h(\mu|\pi) = Ent_\pi\left(d\mu/d\pi\right)$.

Theorem

$\|S_t - \pi\|_{2\to 2} \le e^{-\gamma t}$ *is equivalent to*

$$\textbf{(PI)} \qquad Var_\pi[f] := \pi\left[f^2\right] - \pi^2[f] \le \frac{1}{\gamma}\mathcal{E}(f,f)$$

for every f, *which is called the* Poincaré inequality. *The best constant in* (**PI**) *is the spectral gap of* $\frac{L+L^*}{2}$.

Theorem

$\|S_t - \pi\|_{2\to 2} \le e^{-\gamma t}$ *is equivalent to*

$$\textbf{(PI)} \qquad Var_\pi[f] := \pi\left[f^2\right] - \pi^2[f] \le \frac{1}{\gamma}\mathcal{E}(f,f)$$

for every f, *which is called the* Poincaré inequality. *The best constant in* (**PI**) *is the spectral gap of* $\frac{L+L^*}{2}$.

Suppose the chain is reversible. Then $h(\mu S_t|\pi) \le h(\mu|\pi)e^{-\alpha t}$ *for every initial distribution* μ, *is equivalent to*

$$\textbf{(MLSI)} \qquad Ent_\pi(f) \le \frac{1}{\alpha}\mathcal{E}(f, \log f),$$

for every $f > 0$. *The inequality above is called the* modified logarithmic Sobolev inequality.

In particular (**PI**) implies

$$\max_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \frac{1}{\pi^*} e^{-\gamma t}.$$

In particular (**PI**) implies

$$\max_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \frac{1}{\pi^*} e^{-\gamma t}.$$

while (**MLSI**) and Czisar's inequality yield

$$\max_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \sqrt{\log\left(\frac{1}{\pi^*}\right)} e^{-\alpha t/2}$$

In particular (**PI**) implies

$$\max_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \frac{1}{\pi^*} e^{-\gamma t}.$$

while (**MLSI**) and Czisar's inequality yield

$$\max_{x \in S} \|\delta_x S_t - \pi\|_{TV} \leq \sqrt{\log\left(\frac{1}{\pi^*}\right)} e^{-\alpha t/2}$$

It turns out that $\alpha \leq 2\gamma$. Thus the estimate obtained with the (**MLSI**) could be worse in the exponential rate, but for moderate times could be much better since, for large S, $\sqrt{\log\left(\frac{1}{\pi^*}\right)} \ll \frac{1}{\pi^*}$.

It is customary to deal with a third functional inequality, the *logarithmic Sobolev inequality*:

$$\textbf{(LSI)} \qquad Ent_{\pi}(f) \leq \frac{1}{s}\mathcal{E}(\sqrt{f}, \sqrt{f})$$

for all $f \geq 0$.

It is customary to deal with a third functional inequality, the *logarithmic Sobolev inequality*:

$$(\textbf{LSI}) \qquad Ent_\pi(f) \leq \frac{1}{s}\mathcal{E}(\sqrt{f}, \sqrt{f})$$

for all $f \geq 0$.
The deepest meaning of this inequality will not be dealt with here. We mention the fact that in the case of *diffusion processes*, i.e. when L is a second order differential operator, (**MLSI**) and (**LSI**) coincide.

Theorem

Consider an irreducible, reversible, finite state Markov chain, and let γ, α, s denote the largest constants in (**PI**), (**MLSI**) *and* (**LSI**) *respectively. Then*

$$2\gamma \geq \alpha \geq 4s > 0$$

Theorem

*Consider an irreducible, reversible, finite state Markov chain, and let γ, α, s denote the largest constants in (**PI**), (**MLSI**) and (**LSI**) respectively. Then*

$$2\gamma \geq \alpha \geq 4s > 0$$

For Markov chains with infinite state space the inequalities between the best constants still hold, but the constants may be zero.

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.
  - Sometimes hard to get the right scaling of the constants in terms of the parameters of the model.

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.
  - Sometimes hard to get the right scaling of the constants in terms of the parameters of the model.
- For model with a *spatial* structure, one relates functional inequalities to *spatial mixing properties* (Stroock, Zegarlinski, Yau, Martinelli, Olivieri,.....)

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.
  - Sometimes hard to get the right scaling of the constants in terms of the parameters of the model.
- For model with a *spatial* structure, one relates functional inequalities to *spatial mixing properties* (Stroock, Zegarlinski, Yau, Martinelli, Olivieri,.....)
  - Quite powerful but spatial mixing can be hard to prove.

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.
  - Sometimes hard to get the right scaling of the constants in terms of the parameters of the model.
- For model with a *spatial* structure, one relates functional inequalities to *spatial mixing properties* (Stroock, Zegarlinski, Yau, Martinelli, Olivieri,.....)
  - Quite powerful but spatial mixing can be hard to prove.
  - Applications to inhomogeneous models may be problematic.

Main approaches to functional inequalities for Markov Chains

- Path (geometric) methods (Saloff-Coste, Roberto, ...)
  - Easy to use to get rough bounds.
  - Sometimes hard to get the right scaling of the constants in terms of the parameters of the model.
- For model with a *spatial* structure, one relates functional inequalities to *spatial mixing properties* (Stroock, Zegarlinski, Yau, Martinelli, Olivieri,.....)
  - Quite powerful but spatial mixing can be hard to prove.
  - Applications to inhomogeneous models may be problematic.
- Lyapunov function methods (Cattiaux, Guillin).
  - Natural approach, but effectiveness on models with spatial structure is still unclear (to me)

The Bochner-Bakry-Emery approach to PI and MLSI

In this section we assume reversibility of L. We recall that the proof that the rate of L^2 convergence to equilibrium is the best constant in (**PI**) is based on

$$\frac{d}{dt} Var_{\pi}(S_t f) = -2\mathcal{E}(S_t f, S_t f).$$

The Bochner-Bakry-Emery approach to PI and MLSI

In this section we assume reversibility of L. We recall that the proof that the rate of L^2 convergence to equilibrium is the best constant in (**PI**) is based on

$$\frac{d}{dt} Var_\pi(S_t f) = -2\mathcal{E}(S_t f, S_t f).$$

Taking one more derivative we get

$$\frac{d^2}{dt^2} Var_\pi(S_t f) = -2\frac{d}{dt}\mathcal{E}(S_t f, S_t f) = 4\pi\left[(LS_t f)^2\right].$$

The Bochner-Bakry-Emery approach to PI and MLSI

In this section we assume reversibility of L. We recall that the proof that the rate of L^2 convergence to equilibrium is the best constant in (**PI**) is based on

$$\frac{d}{dt} Var_\pi(S_t f) = -2\mathcal{E}(S_t f, S_t f).$$

Taking one more derivative we get

$$\frac{d^2}{dt^2} Var_\pi(S_t f) = -2\frac{d}{dt}\mathcal{E}(S_t f, S_t f) = 4\pi\left[(LS_t f)^2\right].$$

Suppose now $\exists k > 0$ such that for every f

$$\mathcal{E}(f, f) \leq \frac{1}{k}\pi\left[(Lf)^2\right].$$

The Bochner-Bakry-Emery approach to PI and MLSI

In this section we assume reversibility of L. We recall that the proof that the rate of L^2 convergence to equilibrium is the best constant in (**PI**) is based on

$$\frac{d}{dt} Var_\pi(S_t f) = -2\mathcal{E}(S_t f, S_t f).$$

Taking one more derivative we get

$$\frac{d^2}{dt^2} Var_\pi(S_t f) = -2\frac{d}{dt}\mathcal{E}(S_t f, S_t f) = 4\pi\left[(LS_t f)^2\right].$$

Suppose now $\exists k > 0$ such that for every f

$$\mathcal{E}(f, f) \leq \frac{1}{k}\pi\left[(Lf)^2\right].$$

We get

$$\frac{d}{dt}\mathcal{E}(S_t f, S_t f) \leq -2k\mathcal{E}(S_t f, S_t f)$$

In particular $\mathcal{E}(S_t f, S_t f) \to 0$ as $t \to +\infty$. Rewriting the last inequality as

$$\frac{d}{dt}\mathcal{E}(S_t f, S_t f) \leq k \frac{d}{dt} Var_\pi(S_t f)$$

and integrating from t to $+\infty$ we get

$$\mathcal{E}(f, f) \geq k Var_\pi(f) \;\Rightarrow\; k \leq \gamma \;!$$

In particular $\mathcal{E}(S_t f, S_t f) \to 0$ as $t \to +\infty$. Rewriting the last inequality as

$$\frac{d}{dt}\mathcal{E}(S_t f, S_t f) \le k \frac{d}{dt} Var_\pi(S_t f)$$

and integrating from t to $+\infty$ we get

$$\mathcal{E}(f,f) \ge k Var_\pi(f) \;\Rightarrow\; k \le \gamma\ !$$

By a bit of spectral Theory it can be shown that the best constant k in

$$(\mathbf{PI'}) \qquad \mathcal{E}(f,f) \le \frac{1}{k}\pi\left[(Lf)^2\right]$$

is *equal* to the spectral gap γ.

The same argument can be implemented with he entropy replacing the variance. We obtain, for $f > 0$

$$\frac{d^2}{dt^2} Ent_\pi(S_t f) = -\frac{d}{dt}\mathcal{E}(S_t f, \log S_t f)$$
$$= \pi\left[LS_t f L \log S_t f\right] + \pi\left[\frac{(LS_t f)^2}{S_t f}\right]$$

The same argument can be implemented with he entropy replacing the variance. We obtain, for $f > 0$

$$\frac{d^2}{dt^2} Ent_\pi(S_t f) = -\frac{d}{dt}\mathcal{E}(S_t f, \log S_t f)$$
$$= \pi\left[LS_t f L \log S_t f\right] + \pi\left[\frac{(LS_t f)^2}{S_t f}\right]$$

We therefore have that the inequality

$$(\mathbf{MLSI}') \qquad k\mathcal{E}(f, \log f) \leq \pi\left[LfL \log f\right] + \pi\left[\frac{(Lf)^2}{f}\right]$$

for every $f > 0$, implies the (**MLSI**) $kEnt_\pi(f) \leq \mathcal{E}(f, \log f)$.

The same argument can be implemented with he entropy replacing the variance. We obtain, for $f > 0$

$$\frac{d^2}{dt^2} Ent_\pi(S_t f) = -\frac{d}{dt}\mathcal{E}(S_t f, \log S_t f)$$
$$= \pi\left[LS_t f L \log S_t f\right] + \pi\left[\frac{(LS_t f)^2}{S_t f}\right]$$

We therefore have that the inequality

$$(\mathbf{MLSI'}) \qquad k\mathcal{E}(f, \log f) \leq \pi\left[Lf L \log f\right] + \pi\left[\frac{(Lf)^2}{f}\right]$$

for every $f > 0$, implies the (**MLSI**) $kEnt_\pi(f) \leq \mathcal{E}(f, \log f)$. This time the converse is not necessarily true: the entropy may decay exponentially fast, but not necessarily in a convex way.

In order to understand how useful the above inequalities are, we write generators of Markov chains in the following form:

$$Lf(x) = \sum_{\gamma \in G} c(x,\gamma)[f(\gamma(x)) - f(x)] =: \sum_{\gamma \in G} c(x,\gamma) \nabla_\gamma f(x)$$

where G is some set of functions from X to X (*allowed movements*). It is clear that every countable Markov chain can be written in this way.

In order to understand how useful the above inequalities are, we write generators of Markov chains in the following form:

$$Lf(x) = \sum_{\gamma \in G} c(x, \gamma)[f(\gamma(x)) - f(x)] =: \sum_{\gamma \in G} c(x, \gamma) \nabla_\gamma f(x)$$

where G is some set of functions from X to X (*allowed movements*). It is clear that every countable Markov chain can be written in this way.

For L to be reversible w.r.t. a probability π we need

(Rev) For every $\gamma \in G$ there exists $\gamma^{-1} \in G$ such that $\gamma^{-1}\gamma(x) = x$ for every $x \in X$ such that $c(x, \gamma) > 0$. Moreover

$$\pi(x)c(x, \gamma) = \pi(\gamma(x))c(\gamma(x), \gamma^{-1})$$

The inequalities (**PI'**) and (**MLSI'**) become, respectively

$$\frac{1}{2}\pi\left[\sum_{\gamma\in G} c(x,\gamma)\nabla_\gamma f(x)\nabla_\gamma f(x)\right] \leq \frac{1}{k}\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(x,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

and

$$\frac{1}{2}\pi\left[\sum_{\gamma\in G} c(x,\gamma)\nabla_\gamma f(x)\nabla_\gamma \log f(x)\right]$$
$$\leq \frac{1}{k}\pi\left[\sum_{\gamma,\delta} c(x,\gamma)c(x,\delta)\left(\nabla_\gamma f(x)\nabla_\delta \log f(x) + \frac{\nabla_\gamma f(x)\nabla_\delta f(x)}{f(x)}\right)\right]$$

The "quadratic form"

$$\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(x,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

can be modified as follows.

The "quadratic form"

$$\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(x,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

can be modified as follows.

To fix ideas, set $\pi(x) = \frac{1}{Z}e^{-H(x)}$, $c(x,\gamma) = \exp[-\nabla_\gamma H(x)/2]$, and assume $\gamma\circ\delta = \delta\circ\gamma$ for every γ,δ. Then one shows that, for every f

$$\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(\gamma(x),\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$
$$= \frac{1}{4}\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(\gamma(x),\delta)\left(\nabla_\gamma\nabla_\delta f(x)\right)^2\right] \geq 0.$$

Thus

$$\frac{1}{2}\pi\left[\sum_{\gamma\in G} c(x,\gamma)\nabla_\gamma f(x)\nabla_\gamma f(x)\right] \leq$$

$$\frac{1}{k}\pi\left[\sum_{\gamma,\delta\in G}\left(c(x,\gamma)c(x,\delta)-c(x,\gamma)c(\gamma(x),\delta)\right)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

is stronger than (**PI'**).

Thus

$$\frac{1}{2}\pi\left[\sum_{\gamma\in G}c(x,\gamma)\nabla_\gamma f(x)\nabla_\gamma f(x)\right]\leq$$

$$\frac{1}{k}\pi\left[\sum_{\gamma,\delta\in G}\left(c(x,\gamma)c(x,\delta)-c(x,\gamma)c(\gamma(x),\delta)\right)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

is stronger than (**PI'**).

This may be useful if

$$c(x,\gamma)c(x,\delta)\simeq c(x,\gamma)c(\gamma(x),\delta)$$

for "many" pairs (γ,δ).

Theorem

Let $R : S \times G \times G \to [0, +\infty)$ *be such that for each* x, γ, δ *with* $R(x,\gamma,\delta) > 0$ *we have*

$$\begin{aligned}
\mathbf{P1}: &\quad R(x,\gamma,\delta) = R(x,\delta,\gamma) \\
\mathbf{P2}: &\quad \pi(x)R(x,\gamma,\delta) = \pi(\gamma(x))R(\gamma(x),\gamma^{-1},\delta) \\
\mathbf{P3}: &\quad \gamma\delta(x) = \delta\gamma(x)
\end{aligned}$$

Set $\Gamma_R(x,\gamma,\delta) := c(x,\gamma)c(x,\delta) - R(x,\gamma,\delta)$. *Then we have*

$$\pi\left[(Lf)^2\right] = \pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(x,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

$$\geq \pi\left[\sum_{\gamma,\delta\in G} \Gamma_R(x,\gamma,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$

The result above follows from the identity:

$$\pi\left[\sum_{\gamma,\delta\in G} c(x,\gamma)c(x,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$
$$= \pi\left[\sum_{\gamma,\delta\in G} \Gamma_R(x,\gamma,\delta)\nabla_\gamma f(x)\nabla_\delta f(x)\right]$$
$$+ \pi\left[\sum_{\gamma,\delta\in G} \Gamma_R(x,\gamma,\delta)\left(\nabla_\gamma\nabla_\delta f(x)\right)^2\right]$$

which is reminiscent of Bochner's identities.

Similarly

Theorem

$$\pi\left[LfL\log f\right] + \pi\left[\frac{(Lf)^2}{f}\right]$$

$$= \pi\left[\sum_{\gamma,\delta} c(x,\gamma)c(x,\delta)\left(\nabla_\gamma f(x)\nabla_\delta \log f(x) + \frac{\nabla_\gamma f(x)\nabla_\delta f(x)}{f(x)}\right)\right]$$

$$\geq \pi\left[\sum_{\gamma,\delta} \Gamma_R(x,\gamma,\delta)\left(\nabla_\gamma f(x)\nabla_\delta \log f(x) + \frac{\nabla_\gamma f(x)\nabla_\delta f(x)}{f(x)}\right)\right]$$

Summing up:

$$\pi \left[\sum_{\gamma,\delta \in G} \Gamma_R(x,\gamma,\delta) \nabla_\gamma f(x) \nabla_\delta f(x) \right] \geq k\mathcal{E}(f,f)$$

Summing up:

$$\pi \left[\sum_{\gamma,\delta \in G} \Gamma_R(x,\gamma,\delta) \nabla_\gamma f(x) \nabla_\delta f(x) \right] \geq k\mathcal{E}(f,f)$$

and

$$\pi \left[\sum_{\gamma,\delta} \Gamma_R(x,\gamma,\delta) \left(\nabla_\gamma f(x) \nabla_\delta \log f(x) + \frac{\nabla_\gamma f(x) \nabla_\delta f(x)}{f(x)} \right) \right] \geq k\mathcal{E}(f, log f)$$

for every $f > 0$, imply (**PI**) and (**MLSI'**), respectively.

Glauber dynamics for unbounded particles

Let $\Lambda \subseteq \mathbb{Z}^d$, $K, \lambda : \mathbb{Z}^d \to [0, +\infty)$ with $K(0) = 0$,

$$\sum_{x \in \mathbb{Z}^d} K(x) < +\infty \qquad \overline{\lambda} := \sup_{x \in \mathbb{Z}^d} \lambda(x) < +\infty.$$

Glauber dynamics for unbounded particles

Let $\Lambda \subseteq \mathbb{Z}^d$, $K, \lambda : \mathbb{Z}^d \to [0, +\infty)$ with $K(0) = 0$,

$$\sum_{x \in \mathbb{Z}^d} K(x) < +\infty \qquad \overline{\lambda} := \sup_{x \in \mathbb{Z}^d} \lambda(x) < +\infty.$$

$X := \mathbb{N}^\Lambda$. So, for $\eta \in X$, $\eta_x =$ number of particles at $x \in \Lambda$.

Glauber dynamics for unbounded particles

Let $\Lambda \subseteq \mathbb{Z}^d$, $K, \lambda : \mathbb{Z}^d \to [0, +\infty)$ with $K(0) = 0$,

$$\sum_{x \in \mathbb{Z}^d} K(x) < +\infty \qquad \overline{\lambda} := \sup_{x \in \mathbb{Z}^d} \lambda(x) < +\infty.$$

$X := \mathbb{N}^\Lambda$. So, for $\eta \in X$, $\eta_x =$ number of particles at $x \in \Lambda$.

$$\pi(\eta) := \frac{1}{Z} \prod_{x \in \Lambda} \frac{\lambda(x)^{\eta_x}}{\eta_x!} \exp\left[-\beta \sum_{\{x,y\} \subseteq \Lambda} K(x-y)\eta_x \eta_y\right],$$

Glauber dynamics for unbounded particles

Let $\Lambda \subseteq \mathbb{Z}^d$, $K, \lambda : \mathbb{Z}^d \to [0, +\infty)$ with $K(0) = 0$,

$$\sum_{x \in \mathbb{Z}^d} K(x) < +\infty \qquad \overline{\lambda} := \sup_{x \in \mathbb{Z}^d} \lambda(x) < +\infty.$$

$X := \mathbb{N}^\Lambda$. So, for $\eta \in X$, $\eta_x =$ number of particles at $x \in \Lambda$.

$$\pi(\eta) := \frac{1}{Z} \prod_{x \in \Lambda} \frac{\lambda(x)^{\eta_x}}{\eta_x!} \exp\left[-\beta \sum_{\{x,y\} \subseteq \Lambda} K(x-y)\eta_x \eta_y\right],$$

$\gamma_x^+ =$ creation of a particle at x.
$\gamma_x^- =$ deletion of a particle at x.

Glauber dynamics for unbounded particles

Let $\Lambda \subseteq \mathbb{Z}^d$, $K, \lambda : \mathbb{Z}^d \to [0, +\infty)$ with $K(0) = 0$,

$$\sum_{x \in \mathbb{Z}^d} K(x) < +\infty \qquad \overline{\lambda} := \sup_{x \in \mathbb{Z}^d} \lambda(x) < +\infty.$$

$X := \mathbb{N}^\Lambda$. So, for $\eta \in X$, η_x = number of particles at $x \in \Lambda$.

$$\pi(\eta) := \frac{1}{Z} \prod_{x \in \Lambda} \frac{\lambda(x)^{\eta_x}}{\eta_x!} \exp\left[-\beta \sum_{\{x,y\} \subseteq \Lambda} K(x-y)\eta_x \eta_y\right],$$

γ_x^+ = creation of a particle at x.
γ_x^- = deletion of a particle at x.

$$c(\eta, \gamma_x^-) := \eta_x \qquad c(\eta, \gamma_x^+) := \lambda(x) \exp\left[-\beta \sum_{y \in \Lambda} K(x-y)\eta_y\right]$$

Choose

$$
\begin{aligned}
R(\eta, \gamma_x^+, \gamma_y^+) &= c(\eta, \gamma_y^+) c(\gamma_x^+(\eta), \gamma_y^+) \\
R(\eta, \gamma_x^-, \gamma_y^-) &= c(\eta, \gamma_y^-) c(\gamma_x^-(\eta), \gamma_y^-) \\
R(\eta, \gamma_x^-, \gamma_y^+) &= c(\eta, \gamma_x^-) c(\eta, \gamma_y^+)
\end{aligned}
$$

and write $\nabla_x^\pm$ for $\nabla_{\gamma_x^\pm}$.

We get

$$\pi\left[(Lf)^2\right] \geq \pi\left[\sum_{\gamma,\delta}\Gamma_R(\eta,\gamma,\delta)\nabla_\gamma f(\eta)\nabla_\delta f(\eta)\right] = \sum_{x\in\Lambda}\pi\left[\eta_x\left(\nabla_x^- f(\eta)\right)^2\right]$$
$$+\sum_{x,y\in\Lambda}\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-\exp[-\beta K(x-y)]\right)\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\right]$$
$$=\mathcal{E}(f,f)+\dots$$

We get

$$\pi\left[(Lf)^2\right] \geq \pi\left[\sum_{\gamma,\delta}\Gamma_R(\eta,\gamma,\delta)\nabla_\gamma f(\eta)\nabla_\delta f(\eta)\right] = \sum_{x\in\Lambda}\pi\left[\eta_x\left(\nabla_x^- f(\eta)\right)^2\right]$$

$$+\sum_{x,y\in\Lambda}\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-\exp[-\beta K(x-y)]\right)\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\right]$$

$$=\mathcal{E}(f,f)+\dots$$

Using:

- $c(\eta,\gamma_y^+)\leq\overline{\lambda}$,
- $2\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\geq-\left[\nabla_x^+ f(\eta)\right]^2-\left[\nabla_y^+ f(\eta)\right]^2$

we obtain

We get

$$\pi\left[(Lf)^2\right] \geq \pi\left[\sum_{\gamma,\delta}\Gamma_R(\eta,\gamma,\delta)\nabla_\gamma f(\eta)\nabla_\delta f(\eta)\right] = \sum_{x\in\Lambda}\pi\left[\eta_x\left(\nabla_x^- f(\eta)\right)^2\right]$$

$$+\sum_{x,y\in\Lambda}\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-\exp[-\beta K(x-y)]\right)\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\right]$$

$$=\mathcal{E}(f,f)+\dots$$

Using:

- $c(\eta,\gamma_y^+)\leq\overline{\lambda}$,
- $2\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\geq -\left[\nabla_x^+ f(\eta)\right]^2-\left[\nabla_y^+ f(\eta)\right]^2$

we obtain

$$\pi\left[(Lf)^2\right]\geq[1-\overline{\lambda}\varepsilon(\beta)]\mathcal{E}(f,f)$$

with

$$\varepsilon(\beta):=\sum_{x\in\mathbb{Z}^d}\left(1-\exp[-\beta K(x)]\right).$$

This proves **(PI)** with

$$\gamma = 1 - \overline{\lambda}\varepsilon(\beta),$$

reproving results by Bertini, Cancrini, Cesi (2002), Kondratiev & Lytvynov (2005).

This proves **(PI)** with

$$\gamma = 1 - \overline{\lambda}\varepsilon(\beta),$$

reproving results by Bertini, Cancrini, Cesi (2002), Kondratiev & Lytvynov (2005).

STRONGER INEQUALITIES?

This proves **(PI)** with

$$\gamma = 1 - \overline{\lambda}\varepsilon(\beta),$$

reproving results by Bertini, Cancrini, Cesi (2002), Kondratiev & Lytvynov (2005).

STRONGER INEQUALITIES?

(LSI) is known to fail

One key step was to bound the "non-diagonal terms"

$$\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-\exp[-\beta K(x-y)]\right)\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\right]$$

using $2\nabla_x^+ f(\eta)\nabla_y^+ f(\eta) \geq -\left[\nabla_x^+ f(\eta)\right]^2 - \left[\nabla_y^+ f(\eta)\right]^2$.

One key step was to bound the "non-diagonal terms"

$$\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-\exp[-\beta K(x-y)]\right)\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)\right]$$

using $2\nabla_x^+ f(\eta)\nabla_y^+ f(\eta) \geq -\left[\nabla_x^+ f(\eta)\right]^2 - \left[\nabla_y^+ f(\eta)\right]^2$.

This is an instance of the following bound: for a symmetric matrix $A = (A_{ij})$,

$$A \geq D$$

with D diagonal and

$$d_{ii} = a_{ii} - \sum_{j \neq i} |a_{ij}|.$$

For the **(MLSI')** we get

$$\pi\left[LfL\log f\right]+\pi\left[\frac{(Lf)^2}{f}\right]$$

$$\geq \pi\left[\sum_{\gamma,\delta}\Gamma_R(x,\gamma,\delta)\left(\nabla_\gamma f(x)\nabla_\delta \log f(x)+\frac{\nabla_\gamma f(x)\nabla_\delta f(x)}{f(x)}\right)\right]$$

$$=\mathcal{E}(f,\log f)+\sum_{x\in\Lambda}\pi\left[\eta(x)\frac{(\nabla_x^- f(\eta))^2}{f(\eta)}\right]$$

$$+\sum_{x,y\in\Lambda}\pi\left[c(\eta,\gamma_x^+)c(\eta,\gamma_y^+)\left(1-e^{-\beta K(x-y)}\right)\left(\nabla_x^+ f(\eta)\nabla_y^+\log f(\eta)+\frac{\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)}{f(\eta)}\right)\right]$$

The role of $2ab \geq -a^2 - b^2$ is now played by the following inequality:

The role of $2ab \geq -a^2 - b^2$ is now played by the following inequality:

$$
\begin{aligned}
&(a-1)\log b + (b-1)\log a + 2(a-1)(b-1) \\
&\qquad \geq -\frac{(a-1)^2}{a} - \frac{(b-1)^2}{b} - [(a-1)\log a + (b-1)\log b]
\end{aligned}
$$

The role of $2ab \geq -a^2 - b^2$ is now played by the following inequality:

$$(a-1)\log b + (b-1)\log a + 2(a-1)(b-1)$$
$$\geq -\frac{(a-1)^2}{a} - \frac{(b-1)^2}{b} - [(a-1)\log a + (b-1)\log b]$$

which yields

$$\nabla_x^+ f(\eta)\nabla_y^+ \log f(\eta) + \nabla_y^+ f(\eta)\nabla_x^+ \log f(\eta) + 2\frac{\nabla_x^+ f(\eta)\nabla_y^+ f(\eta)}{f(\eta)}$$
$$\geq -\frac{\left(\nabla_x^+ f(\eta)\right)^2}{f(\eta+\delta_x)} - \frac{\left(\nabla_y^+ f(\eta)\right)^2}{f(\eta+\delta_y)} - \nabla_x^+ f(\eta)\nabla_x^+ \log f(\eta) - \nabla_y^+ f(\eta)\nabla_y^+ \log f(\eta)$$

from which we obtain

$$\pi\left[LfL\log f\right] + \pi\left[\frac{(Lf)^2}{f}\right] \geq [1-\varepsilon(\beta)]\left[\mathcal{E}(f,\log f) + \sum_{x\in\Lambda}\pi\left[\eta(x)\frac{(\nabla_x^- f(\eta))^2}{f(\eta)}\right]\right]$$

which means

from which we obtain

$$\pi\left[LfL\log f\right] + \pi\left[\frac{(Lf)^2}{f}\right] \geq [1-\varepsilon(\beta)]\left[\mathcal{E}(f,\log f) + \sum_{x\in\Lambda}\pi\left[\eta(x)\frac{(\nabla_x^- f(\eta))^2}{f(\eta)}\right]\right]$$

which means

Theorem

Assuming $\overline{\lambda}\varepsilon(\beta) < 1$, **(MLSI')** *holds with constant* $[1-\overline{\lambda}\varepsilon(\beta)]$.

Some concluding remarks

For the spectral gap, the methods used for bounding from below a quadratic form is very naive. Better estimates could be obtained by estimating the spectral radius of the form.

Some concluding remarks

For the spectral gap, the methods used for bounding from below a quadratic form is very naive. Better estimates could be obtained by estimating the spectral radius of the form.

We worked out some examples where the “naive” estimates are not good enough.

Some concluding remarks

For the spectral gap, the methods used for bounding from below a quadratic form is very naive. Better estimates could be obtained by estimating the spectral radius of the form.

We worked out some examples where the "naive" estimates are not good enough.

E.g. the Loss Networks representation of Peierls contours (Fernandez, Ferrari, Garcia, 2001). This is a birth & death dynamics on Peierls contours, having the Ising model as invariant measure.

We can give a uniform lower bound of the gap under the same low temperature condition as in [FFG].

Some concluding remarks

However, *global, non-diagonal bounds* are not yet available for **(MLSI')**.

THANKS!