

Localization Schemes

Zhidan Li

1 Localization Schemes and Markov Chains

Now we introduce another framework to show the local-to-global theorem. This framework, named the *localization schemes*, is highly related to the recent breakthrough of the famous Kannan-Lovász-Simonovits Conjecture, and deeply studied in Chen and Eldan [CE22] to analyze the mixing time of the Markov chains.

We fix a state space Ω equipped with a σ -algebra Σ . Usually we assume that $\Sigma = 2^\Omega$ when Ω is finite and $\Sigma = \text{Borel}(\Omega)$ when Ω is a continuous space, and then we omit Σ . Let $\mathcal{M}(\Omega)$ be the space of all probability measures on Ω .

Definition 1.1 (Localization Process). A *localization process* $(\mu_t)_{t \geq 0}$ on the state space Ω is a stochastic process satisfying

- (P1) Almost surely μ_t is a probability measure on Ω for all $t \geq 0$.
- (P2) For every measurable $A \subseteq \Omega$, the process $(\mu_t(A))_{t \geq 0}$ is a martingale.
- (P3) For every measurable $A \subseteq \Omega$, the process $(\mu_t(A))_{t \geq 0}$ almost surely converges to either 0 or 1 as $t \rightarrow \infty$.

For convenience, we use Θ_t to denote the distribution of μ_t for every $t \geq 0$.

Definition 1.2 (Localization Scheme). A *localization scheme* \mathcal{L} on Ω is a mapping assigning to each probability measure $\mu \in \mathcal{M}(\Omega)$ a localization process $(\mu_t)_{t \geq 0}$ with $\mu_0 = \mu$. In this case, we say $(\mu_t)_t$ is the localization process associated with μ via the localization scheme \mathcal{L} .

1.1 Markov dynamics associated with the localization process

In this part we associate a localization process $(\mu_t)_{t \geq 0} = \mathcal{L}(\mu)$ with a Markov dynamics reversible with respect to the distribution $\mu \in \mathcal{M}(\Omega)$.

Definition 1.3 (Markov Chains Associated with Localization Processes). Let $(\mu_t)_{t \geq 0}$ be a localization process on Ω associated with μ via a localization scheme \mathcal{L} and $\tau > 0$ be a stopping time. The Markov dynamics $P = P^{(\mathcal{L}, \tau)}$ associated with $(\mu_t)_{t \geq 0}$ and τ is defined as

$$P(x, A) = \mathbf{E}_{\Theta_\tau} \left[\frac{\mu_\tau(x) \mu_\tau(A)}{\mu(x)} \right], \quad \forall x \in \Omega, A \in \Sigma.$$

Remark 1.4. An optional way to view Definition 1.3 is, let X, Y be two random variables taking values in $\Omega \times \Omega$ satisfying

$$\Pr [X \in A, Y \in B] = \mathbf{E} [\mu_\tau(A) \mu_\tau(B)], \quad \forall A, B \in \Sigma.$$

Then we define the kernel as

$$P(x, A) = \Pr [Y \in A \mid X = x].$$

Fact 1.5. Let $P = P^{(\mathcal{L}, \tau)}$ be the transition kernel defined as Definition 1.3. Then P is reversible with respect to μ .

Proof. For every $x \in \Omega$, it almost surely holds that

$$P(x, \Omega) = \mathbf{E}_{\Theta_t} \left[\frac{\mu_\tau(x) \mu_\tau(\Omega)}{\mu(x)} \right] = \mathbf{E}_{\Theta_t} \left[\frac{\mu_\tau(x)}{\mu(x)} \right] = 1.$$

Then we know $P(x, \cdot)$ is a probability measure on Ω almost surely. Also for every $A, B \in \Sigma$, it holds that

$$\begin{aligned} \int_{x \in A} P(x, B) \, d\mu(x) &= \int_{x \in A} \mathbf{E} \left[\frac{d\mu_\tau(x)}{d\mu(x)} \mu_\tau(B) \right] \, d\mu(x) \\ &= \mathbf{E} \left[\int_{x \in \Omega} \mu_\tau(B) \, d\mu_\tau(x) \right] \\ &= \mathbf{E} [\mu_\tau(A) \mu_\tau(B)] \\ &= \int_{y \in B} P(y, A) \, d\mu(y). \end{aligned}$$

Therefore we know P is reversible with respect to μ . □

1.2 Functional inequalities

Recall the Dirichlet form of a random walk P with stationary distribution μ : for two functions $f, g : \Omega \rightarrow \mathbb{R}$,

$$\mathcal{E}_P(f, g) := \int_{x \in \Omega} f(x)(I - P)g(x) \, d\mu(x)$$

and the spectral gap and modified log-Sobolev inequality constant of P :

$$\text{Gap}(P) := \inf_{f: \Omega \rightarrow \mathbb{R}} \frac{\mathcal{E}_P(f, f)}{\text{Var}_\mu[f]}, \quad \rho_{\text{LS}}(P) := \inf_{f: \Omega \rightarrow \mathbb{R}_{>0}} \frac{\mathcal{E}_P(f, \log f)}{\text{Ent}_\mu[f]}.$$

The following identity and the inequality illustrate the connection between the functional inequalities and the variance or entropy of the localization process.

Proposition 1.6. *Let $P = P^{(\mathcal{L}, \tau)}$ be a transition kernel associated with a localization process $(\mu_t)_{t \geq 0} = \mathcal{L}(\mu)$ and $\tau > 0$. Then it holds that*

$$\mathcal{E}_P(f, f) = \mathbf{E}_{\Theta_\tau} [\text{Var}_{\mu_\tau}[f]], \quad \mathcal{E}_P(f, \log f) \geq \mathbf{E}_{\Theta_\tau} [\text{Ent}_{\mu_\tau}[f]].$$

for every function f supported on Ω when the Dirichlet forms are well-defined.

Proof. We prove them one by one. By calculation,

$$\begin{aligned} \mathcal{E}_P(f, f) &= \int_{x \in \Omega} f(x)(I - P)f(x) \, d\mu(x) \\ &= \int_{x \in \Omega} (f(x)^2 - f(x)(Pf)(x)) \, d\mu(x) \\ &= \int_{x \in \Omega} f(x)^2 \, d\mu(x) - \int_{x \in \Omega} f(x) \left(\int_{y \in \Omega} f(y) \, dP(x, y) \right) \, d\mu(x) \\ &= \mathbf{E}_\mu[f^2] - \int_{x \in \Omega} \int_{y \in \Omega} f(x)f(y) \mathbf{E}_{\Theta_\tau} \left[\frac{d\mu_\tau(x)}{d\mu(x)} \, d\mu_\tau(y) \right] \, d\mu(x) \\ &= \mathbf{E}_\mu[f^2] - \mathbf{E}_{\Theta_\tau} \left[\int_{x \in \Omega} f(x) \left(\int_{y \in \Omega} f(y) \, d\mu_\tau(y) \right) \, d\mu(x) \right] \\ &= \mathbf{E}_{\Theta_\tau} [\mathbf{E}_{\mu_\tau}[f^2] - \mathbf{E}_{\mu_\tau}[f]^2] \\ &= \mathbf{E}_{\Theta_\tau} [\text{Var}_{\mu_\tau}[f]] \end{aligned}$$

where the identity $\mathbf{E}_\mu [f^2] = \mathbf{E}_{\Theta_\tau} [\mathbf{E}_{\mu_\tau} [f^2]]$ holds from the martingality of the process.

For the MLSI constant, by calculation, we know

$$\begin{aligned}
\mathcal{E}_P(f, \log f) &= \int_{x \in \Omega} f(x) ((I - P) \log f)(x) \, d\mu(x) \\
&= \int_{x \in \Omega} (f(x) \log f(x) - f(x)(P \log f)(x)) \, d\mu(x) \\
&= \int_{x \in \Omega} f(x) \log f(x) \, d\mu(x) - \int_{x \in \Omega} f(x) \left(\int_{y \in \Omega} \log f(y) \, dP(x, y) \right) \, d\mu(x) \\
&= \mathbf{E}_\mu [f \log f] - \int_{x \in \Omega} \int_{y \in \Omega} f(x) \log f(y) \mathbf{E}_{\Theta_\tau} \left[\frac{d\mu_\tau(x)}{d\mu(x)} \, d\mu_\tau(y) \right] \, d\mu(x) \\
&= \mathbf{E}_\mu [f \log f] - \mathbf{E}_{\Theta_\tau} \left[\int_{x \in \Omega} f(x) \left(\int_{y \in \Omega} \log f(y) \, d\mu_\tau(y) \right) \, d\mu(x) \right] \\
&= \mathbf{E}_{\Theta_\tau} [\mathbf{E}_{\mu_\tau} [f \log f] - \mathbf{E}_{\mu_\tau} [f] \mathbf{E}_{\mu_\tau} [\log f]] \\
&\geq \mathbf{E}_{\Theta_\tau} [\mathbf{E}_{\mu_\tau} [f \log f] - \mathbf{E}_{\mu_\tau} [f] \log \mathbf{E}_{\mu_\tau} [f]] \\
&= \mathbf{E}_{\Theta_\tau} [\mathbf{Ent}_{\mu_\tau} [f]]
\end{aligned}$$

where the inequality holds from the Jensen's inequality $\log \mathbf{E}_\pi [f] \geq \mathbf{E}_\pi [\log f]$ for every distribution π on Ω and every test function $f : \Omega \rightarrow \mathbb{R}_{>0}$. \square

2 Linear-Tilt Localization Processes

Now we introduce a family of localization processes which lies on the core of the analysis of the mixing time. For a distribution π on Ω , we use $\mathbf{b}(\pi)$ to denote the mass center of π , i.e.,

$$b(\pi) = \int_{x \in \Omega} x \, d\pi(x).$$

Definition 2.1 (Linear-Tilt Localization Processes). For a localization process $(\mu_t)_{t \geq 0}$, we say it is a *linear-tilt localization process* if:

- **(Discrete version)** For all $t \in \mathbb{N}$ and $x \in \Omega$,

$$\mu_{t+1}(x) = \mu_t(x) (1 + \langle x - \mathbf{b}(\mu_t), Z_t \rangle) \quad (1)$$

where Z_t is a random variable with $\mathbf{E} [Z_t \mid \mu_t] = 0$. Or,

- **(Continuous version)** For all $t \geq 0$ and $x \in \Omega$,

$$d\mu_t(x) = \mu_t(x) \langle x - \mathbf{b}(\mu_t), Z_t \rangle \quad (2)$$

where Z_t is a random variable with $\mathbf{E} [Z_t \mid \mu_t] = 0$.

For convenience, we say $(Z_t)_{t \geq 0}$ is the driving factor of $(\mu_t)_{t \geq 0}$.

We will focus on the following two kinds of localization schemes: (1) the coordinate-by-coordinate localization schemes; (2) the stochastic localization schemes driven by standard Brownian motion.

2.1 The coordinate-by-coordinate localization schemes

Given a distribution μ over $\Omega \subseteq \mathbb{R}^n$, we construct a discrete-time localization process $(\mu_t)_{t \geq 0}$ as follows:

- Firstly we pick a permutation k_1, \dots, k_n of $[n]$ uniformly at random.
- Let $X \sim \mu$. For $t \geq 0$, we set μ_t to be the law of X conditional on X_{k_1}, \dots, X_{k_i} where $i = \min\{n, \lfloor t \rfloor\}$.

Now we show the observation that the dynamics associated with the coordinate-by-coordinate localization process is the well-known *Glauber dynamics*.

Fact 2.2. *Given a coordinate-by-coordinate localization scheme \mathcal{L} over $\Omega \subseteq \mathbb{R}^n$ and an integer $\tau = n - 1$, the Markov chain $P = P^{(\mathcal{L}, \tau)}$ associated with $(\mu_t)_{t \geq 0} = \mathcal{L}(\mu)$ and τ is the single-site Glauber dynamics denoted by P^{GD} with stationary distribution μ .*

Proof. We verify the fact by definition. For every $x \in \Omega$ and $i \in [n]$, define $L_{x,i} := \{z \in \Omega \mid \forall j \in [n] \setminus \{i\}, z_j = x_j\}$. It's not hard to see that it suffices to show the case $\|x - y\|_0 = 1$.

Assume that x, y only differ at the coordinate $i \in [n]$, i.e., $x_i \neq y_i$ and $x_j = y_j$ for every $j \in [n] \setminus \{i\}$. Then by definition,

$$\begin{aligned}
 P(x, y) &= \mathbf{E}_{\Theta_{n-1}} \left[\frac{\mu_{n-1}(x) \mu_{n-1}(y)}{\mu(x)} \right] \\
 &= \sum_{j \in [n]} \frac{1}{n} \mathbf{E} \left[\frac{\mu_{n-1}(x) \mu_{n-1}(y)}{\mu(x)} \mid k_n = j \right] \\
 &= \frac{1}{n} \mathbf{E} \left[\frac{\mu_{n-1}(x) \mu_{n-1}(y)}{\mu(x)} \mid k_n = i \right] \\
 &= \frac{1}{n} \Pr[\text{supp}(\mu_{n-1}) = L_{x,i}] \mathbf{E} \left[\frac{\mu_{n-1}(x) \mu_{n-1}(y)}{\mu(x)} \mid k_n = i, \text{supp}(\mu_{n-1}) = L_{x,i} \right] \\
 &= \frac{1}{n} \frac{\mu(L_{x,i}) \mu(x) \mu(y)}{\mu(x) \mu(L_{x,i})^2} \\
 &= \frac{1}{n} \frac{\mu(y)}{\mu(L_{x,i})}.
 \end{aligned}$$

When $\|x - y\|_0 \geq 2$, it is easy to see $P(x, y) = P^{\text{GD}}(x, y) = 0$. Thus we conclude the statement. \square

Remark 2.3. When $\tau = n - \ell$, the corresponding Markov kernel associated with the coordinate-by-coordinate localization process and τ is the ℓ -uniform block dynamics $P^{\ell\text{-GD}}$.

2.1.1 The coordinate-by-coordinate localization process as a linear-tilt process

In this part we will show the coordinate-by-coordinate localization process $(\mu_t)_{t \geq 0}$ is a linear-tilt localization process. Fix a probability measure μ on $\Omega = \{-1, +1\}^n$. We pick a permutation k_1, \dots, k_n of $[n]$ uniformly at random. Let U_1, \dots, U_n be independent random variables uniformly distributed in $[-1, +1]$.

Let $\mu_0 = \mu$. For $i = 0, 1, \dots, n$, we define

$$\mu_{i+1}(x) = \mu_i(x) (1 + \langle x - \mathbf{b}(\mu_i), Z_i \rangle), \quad \forall x \in \Omega$$

where Z_i is a $\sigma(\mu_0, \dots, \mu_i)$ -measurable random variable defined as

$$Z_i := \mathbf{e}_{k_{i+1}} \times \begin{cases} \frac{1}{1 + \mathbf{b}(\mu_i)_{k_{i+1}}} & \mathbf{b}(\mu_i)_{k_{i+1}} \geq U_{i+1}, \\ \frac{-1}{1 - \mathbf{b}(\mu_i)_{k_{i+1}}} & \mathbf{b}(\mu_i)_{k_{i+1}} \leq U_{i+1}, \end{cases}$$

where $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is the standard basis of \mathbb{R}^n .

It is not hard to see $E[Z_i | \mu_i] = 0$, and

$$\begin{aligned}\mu_{i+1}(\Omega) &= \int_{x \in \Omega} d\mu_{i+1}(x) \\ &= \int_{x \in \Omega} (1 + \langle x - \mathbf{b}(\mu_i), Z_i \rangle) d\mu_i(x) \\ &= \mu_i(\Omega) + \left\langle \int_{x \in \Omega} (x - \mathbf{b}(\mu_i)) d\mu_i(x), Z_i \right\rangle \\ &= \mu_i(\Omega)\end{aligned}$$

meaning that $\mu_i(\Omega) = 1$ for each $i \in [n]$. To show μ_{i+1} is a pinning of μ_i , firstly note that the marginal distribution of the coordinate k_{i+1} is

$$\Pr_{X \sim \mu_t} [X_{k_{i+1}} = 1] = \frac{1 + \mathbf{b}(\mu_i)_{k_{i+1}}}{2}, \quad \Pr_{X \sim \mu_t} [X_{k_{i+1}} = 1] = \frac{1 - \mathbf{b}(\mu_i)_{k_{i+1}}}{2}.$$

By the definition of Z_i , when x is not identical to the pinned value, the inner product will be -1 and the probability will vanish.

2.2 Stochastic localization schemes driven by standard Brownian motion

Now we introduce a kind of linear-tilt localization schemes named the *stochastic localization scheme* firstly constructed by Eldan [Eld13]. Fix a probability measure μ on $\Omega \subseteq \mathbb{R}^n$. Let $(B_t)_{t \geq 0}$ be the standard Brownian motion in \mathbb{R}^n adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Let $(C_t)_{t \geq 0}$ be a stochastic process adapted to $(\mathcal{F}_t)_{t \geq 0}$ taking values in $n \times n$ positive semidefinite matrices. We define a measure-valued stochastic process $(\mu_t)_{t \geq 0}$ by $\frac{d\mu_t}{d\mu}(x) = F_t(x)$ as,

$$F_0(x) = 1, \quad dF_t(x) = F_t(x) \langle x - \mathbf{b}(\mu_t), C_t dB_t \rangle, \quad \forall x \in \Omega. \quad (3)$$

Proposition 2.4. *If $\int_{t=0}^{\infty} C_t^2 dt = \infty$, then $(\mu_t)_{t \geq 0}$ is a localization process. Moreover,*

$$\frac{d\mu_t}{d\mu}(x) = F_t(x) = \frac{1}{Z_t} \exp \left(-\frac{1}{2} \langle \Sigma_t x, x \rangle + \langle \mathbf{y}_t, x \rangle \right)$$

where Z_t is a normalizing factor to ensure that $\int_{x \in \Omega} F_t(x) d\mu(x) = 1$ and $(\Sigma_t)_{t \geq 0}, (\mathbf{y}_t)_{t \geq 0}$ are stochastic processes adapted to \mathcal{F}_t in the form of

$$d\mathbf{y}_t = C_t dB_t + C_t^2 \mathbf{b}(\mu_t) dt, \quad d\Sigma_t = C_t^2 dt.$$

Proof. We prove the proposition by solving (3). Consider the stochastic process $(\log F_t(x))_{t \geq 0}$. By Itô's formula,

$$\begin{aligned}d \log F_t(x) &= \frac{dF_t(x)}{F_t(x)} - \frac{d[F(x)]_t}{2F_t(x)^2} \\ &= \langle x - \mathbf{b}(\mu_t), C_t dB_t \rangle - \frac{1}{2} \|C_t(x - \mathbf{b}(\mu_t))\|_2^2 dt.\end{aligned}$$

This leads to the form

$$F_t(x) = \frac{1}{Z_t} \exp \left(-\frac{1}{2} \langle \Sigma_t x, x \rangle + \langle \mathbf{y}_t, x \rangle \right)$$

where $Z_t, \Sigma_t, \mathbf{y}_t$ are described as the proposition. Also we know $\mu_t(x) \geq 0$ for every $x \in \Omega$. By definition,

$$\begin{aligned}d\mu_t(\Omega) &= d \int_{x \in \Omega} d\mu_t(x) \\ &= \int_{x \in \Omega} F_t(x) \langle x - \mathbf{b}(\mu_t), C_t dB_t \rangle d\mu(x) \\ &= \left\langle \int_{x \in \Omega} (x - \mathbf{b}(\mu_t)) d\mu_t(x), C_t dB_t \right\rangle \\ &= 0.\end{aligned}$$

Then we know $\mu_t(\Omega) = 1$ for every $t \geq 0$ almost surely. Thus we know μ_t is almost surely a probability measure on Ω . The martingality comes directly from the definition, and to see the convergence of the process, note that when $\Sigma_t \rightarrow \infty$, by the form of F_t it will be a Dirac measure. \square

When $C_t \equiv Q^{-1/2}$, we know the law of y_t by El Alaoui and Montanari [EAM22].

Theorem 2.5 ([EAM22]). *Fix a probability measure μ on Ω and a positive semidefinite matrix $Q \in \mathbb{R}^{n \times n}$. Let $(\mu_t)_{t \geq 0}$ be a stochastic localization process starting from μ driven by $C_t \equiv Q^{-1/2}$. Define the stochastic process $(\Sigma_t)_{t \geq 0}$, $(y_t)_{t \geq 0}$ as above. Then*

$$\Sigma_t = tQ^{-1}, \quad y_t/t \sim \mu * \mathcal{N}(0, \Sigma_t), \quad \forall t \geq 0.$$

2.3 Variance contraction via linear-tilt localization processes

Now we show how to bound the spectral gap of the Glauber dynamics P^{GD} . The following property named the *variance conservation* is the key in our analysis.

Definition 2.6 (Variance Conservation - Discrete). Given a time-discrete localization process $(\mu_t)_{t \in \mathbb{N}}$ on Ω satisfying $(\kappa_1, \kappa_2, \dots)$ -variance conservation up to time $t \in \mathbb{N}$, if for every test function $f : \Omega \rightarrow \mathbb{R}$,

$$\mathbb{E} [\text{Var}_{\mu_i} [f] \mid \mu_{i-1}] \geq (1 - \kappa_i) \text{Var}_{\mu_{i-1}} [f], \quad \forall 1 \leq i \leq t.$$

Proposition 2.7. *Let $(\mu_t)_{t \in \mathbb{N}}$ be a time-discrete localization process on Ω satisfying $(\kappa_1, \kappa_2, \dots)$ -variance conservation up to time $t \in \mathbb{N}$. Let P be the random walk associated with $(\mu_t)_{t \in \mathbb{N}}$ and time t . Then its spectral gap $\text{Gap}(P)$ satisfies*

$$\text{Gap}(P) \geq \prod_{i=1}^t (1 - \kappa_i).$$

Proof. By Proposition 1.6, it suffices to show for every test function $f : \Omega \rightarrow \mathbb{R}^n$,

$$\frac{\mathbb{E}_{\Theta_t} [\text{Var}_{\mu_t} [f]]}{\text{Var}_{\mu} [f]} \geq \prod_{i=1}^t (1 - \kappa_i).$$

Note that $\mu_0 = \mu$. Then by direct calculation,

$$\begin{aligned} \frac{\mathbb{E}_{\Theta_t} [\text{Var}_{\mu_t} [f]]}{\text{Var}_{\mu} [f]} &= \mathbb{E}_{\Theta_t} \left[\frac{\text{Var}_{\mu_t} [f]}{\text{Var}_{\mu_0} [f]} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\dots \mathbb{E} \left[\frac{\text{Var}_{\mu_t} [f]}{\text{Var}_{\mu_{t-1}} [f]} \mid \mu_{t-1} \right] \dots \right] \frac{\text{Var}_{\mu_1} [f]}{\text{Var}_{\mu_0} [f]} \mid \mu_0 \right] \\ &\geq \prod_{i=1}^t (1 - \kappa_i) \end{aligned}$$

where the last inequality holds from Definition 2.6. \square

Now it's time for us to show the variance contraction for a linear-tilt localization process $(\mu_t)_{t \in \mathbb{N}}$. The first step is to show the form of the evolution of its variance.

Lemma 2.8. *Let $(\mu_t)_{t \in \mathbb{N}}$ be a time-discrete linear-tilt localization process and $(Z_t)_{t \in \mathbb{N}}$ be its driving factor. Then for every test function $f : \Omega \rightarrow \mathbb{R}$ and $t \in \mathbb{N}$,*

$$\mathbb{E} [\text{Var}_{\mu_{t+1}} [f] \mid \mu_t] = \text{Var}_{\mu_t} [f] - \langle V_t, C_t V_t \rangle$$

where

$$V_t := \int_{x \in \Omega} (x - \mathbf{b}(\mu_t)) f(x) d\mu_t(x), \quad C_t := \text{Cov}(Z_t \mid \mu_t).$$

Proof. Fix a test function $f : \Omega \rightarrow \mathbb{R}$. By direct calculation,

$$\begin{aligned}
\mathbf{E} [\mathbf{Var}_{\mu_{t+1}} [f] \mid \mu_t] &= \mathbf{E} \left[\int_{\Omega} f(x)^2 d\mu_{t+1}(x) - \left(\int_{\Omega} f(x) d\mu_{t+1}(x) \right)^2 \mid \mu_t \right] \\
&= \int_{\Omega} f(x)^2 d\mu_t(x) - \mathbf{E} \left[\left(\int_{\Omega} f(x) (1 + \langle x - \mathbf{b}(\mu_t), Z_t \rangle) d\mu_t(x) \right)^2 \mid \mu_t \right] \\
&= \int_{\Omega} f(x)^2 d\mu_t(x) - \left(\int_{\Omega} f(x) d\mu_t(x) \right)^2 - \mathbf{E} \left[\left(\int_{\Omega} f(x) \langle x - \mathbf{b}(\mu_t), Z_t \rangle d\mu_t(x) \right)^2 \mid \mu_t \right] \\
&= \mathbf{Var}_{\mu_t} [f] - \mathbf{E} [\langle V_t, Z_t \rangle^2 \mid \mu_t] \\
&= \mathbf{Var}_{\mu_t} [f] - V_t^\top \mathbf{E} [Z_t^\top Z_t \mid \mu_t] V_t \\
&= \mathbf{Var}_{\mu_t} [f] - V_t^\top C_t V_t.
\end{aligned}$$

□

Proposition 2.9. *Let $(\mu_t)_{t \in \mathbb{N}}$ be a time-discrete linear-tilt localization process and $(Z_t)_{t \in \mathbb{N}}$ be its driving factor. Then $(\mu_t)_{t \in \mathbb{N}}$ satisfies $(\kappa_1, \kappa_2, \dots)$ -variance conservation where*

$$\kappa_{t+1} = 1 - \left\| C_t^{1/2} \mathbf{Cov}(\mu_t) C_t^{1/2} \right\|_{\text{OP}}, \quad \forall t \in \mathbb{N}.$$

Proof. Firstly it is not hard to see that it suffices to show the case $\mathbf{E}_\mu [f] = \mathbf{E}_{\mu_t} [f] = 0$. By Lemma 2.8, we only need to bound the term $\langle V_t, C_t V_t \rangle$. By definition,

$$\begin{aligned}
\langle V_t, C_t V_t \rangle &= \left\| C_t^{1/2} V_t \right\|_2^2 \\
&= \sup_{\theta: \|\theta\|_2=1} \left\langle C_t^{1/2} V_t, \theta \right\rangle^2 \\
&= \sup_{\theta: \|\theta\|_2=1} \left(\int_{\Omega} \langle C_t(x - \mathbf{b}(\mu_t)), \theta \rangle f(x) d\mu_t(x) \right)^2 \\
&\leq \sup_{\theta: \|\theta\|_2=1} \mathbf{Var}_{\mu_t} [f] \int_{\Omega} \langle C_t(x - \mathbf{b}(\mu_t)), \theta \rangle^2 f(x) d\mu_t(x) \\
&= \left\| C_t^{1/2} \mathbf{Cov}(\mu_t) C_t^{1/2} \right\|_{\text{OP}} \mathbf{Var}_{\mu_t} [f]
\end{aligned}$$

where the inequality holds by the Cauchy-Schwarz inequality. □

2.3.1 Variance conservation via the coordinate-by-coordinate localization process

Now we show the main result of rapid mixing via the spectral independence by Anari, Liu and Oveis Shayan [ALOG20].

Lemma 2.10. *Fix a distribution μ on $\Omega \subseteq \{-1, +1\}^n$. Let $(\mu_t)_{t \in \mathbb{N}}$ be a coordinate-by-coordinate localization process starting from μ . Then $(\mu_t)_{t \in \mathbb{N}}$ satisfies $(\kappa_1, \kappa_2, \dots)$ -variance conservation up to n such that*

$$\kappa_{t+1} = 1 - \frac{\|\mathbf{Cor}(\mu_t)\|_{\text{OP}}}{n-t}, \quad \forall 0 \leq t < n$$

where $\mathbf{Cor}(\mu_t) = \text{diag}(\mathbf{Cov}(\mu_t))^{-1/2} \mathbf{Cov}(\mu_t) \text{diag}(\mathbf{Cov}(\mu_t))^{-1/2}$.

Proof. By Proposition 2.9, it suffices to show

$$C_t^{1/2} \mathbf{Cov}(\mu_t) C_t^{1/2} = \frac{\mathbf{Cor}(\mu_t)}{n-t}.$$

By direct calculation, for every unpinned $i \in [n]$,

$$\begin{aligned} C_t(i, i) &= \mathbf{Cov}(Z_t \mid \mu_t)_{i,i} \\ &= \frac{1}{n-t} \frac{1}{1 - \mathbf{b}(\mu_t)_i^2} \\ &= \frac{1}{(n-t) \mathbf{Cov}(\mu_t)_{i,i}}. \end{aligned}$$

Then the identity holds. □

Since we have already know $\|\Psi_{\mu_t}\|_{\text{OP}} = \|\mathbf{Cor}(\mu_t)\|_{\text{OP}}$, we can establish the result of [ALOG20].

Lemma 2.11 (A Reformulation of the Main Result in [ALOG20]). *Given an (η_0, \dots, η_n) -spectrally independent Gibbs distribution μ of some hardcore model over the state space $\Omega \subseteq \{-1, +1\}^n$, the spectral gap of the ℓ -uniform block dynamics is at least*

$$\text{Gap}(\mathbf{P}^{\ell\text{-GD}}) \geq \prod_{t=0}^{n-\ell-1} \left(1 - \frac{\eta_t}{n-t}\right).$$

References

- [ALOG20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS2020, Durham, NC, USA, November 16-19, 2020*, pages 1319–1330. IEEE, 2020. [7](#), [8](#)
- [CE22] Yuansi Chen and Ronen Eldan. Localization Schemes: A Framework for Proving Mixing Bounds for Markov Chains (extended abstract). In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 110–122, Los Alamitos, CA, USA, Nov. 2022. IEEE Computer Society. [1](#)
- [EAM22] Ahmed El Alaoui and Andrea Montanari. An Information-Theoretic View of Stochastic Localization. *IEEE Transactions on Information Theory*, 68(11):7423–7426, 2022. [6](#)
- [Eld13] Ronen Eldan. Thin Shell Implies Spectral Gap Up to Polylog via a Stochastic Localization Scheme. *Geometric and Functional Analysis*, 23(2):532–569, 2013. [5](#)