Exponential type inequalities for Markov Chains

Witold Bednorz[1]

[1]Department of Mathematics
University of Warsaw

MCMC, Warwick 16-20 March 2009
Joint work with R.Adamczak from Warsaw University

Outline

Independent Random Variables

- Let $X_1, X_2, ..., X_n$ be family of independent random variables, such that $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$.

Independent Random Variables

- Let $X_1, X_2, ..., X_n$ be family of independent random variables, such that $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$.
- To prove the concentration inequality we use Markov inequality

$$\mathbf{P}(\sum_{i=1}^{n} X_i > t) \leqslant e^{-\lambda t} \prod_{i=1}^{n} \mathbf{E} \exp(\lambda X_i).$$

Independent Random Variables

- Let $X_1, X_2, ..., X_n$ be family of independent random variables, such that $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$.
- To prove the concentration inequality we use Markov inequality

$$\mathbf{P}(\sum_{i=1}^{n} X_i > t) \leqslant e^{-\lambda t} \prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i).$$

- The simplest bound on the Laplace transform is

$$\mathbf{E}\exp(\lambda X_i) \leqslant 1+\sum_{k=2}^{\infty} \frac{\mathbf{E}\lambda^k |X_i|^k}{k!} \leqslant 1+\sum_{k=2}^{\infty} \lambda^k c^k \leqslant \exp(\frac{\lambda^2 c^2}{1-\lambda c}).$$

Independent Random Variables

- Let $X_1, X_2, ..., X_n$ be family of independent random variables, such that $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$.
- To prove the concentration inequality we use Markov inequality

$$\mathbf{P}(\sum_{i=1}^{n} X_i > t) \leqslant e^{-\lambda t} \prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i).$$

- The simplest bound on the Laplace transform is

$$\mathbf{E}\exp(\lambda X_i) \leqslant 1 + \sum_{k=2}^{\infty} \frac{\mathbf{E}\lambda^k |X_i|^k}{k!} \leqslant 1 + \sum_{k=2}^{\infty} \lambda^k c^k \leqslant \exp(\frac{\lambda^2 c^2}{1-\lambda c}).$$

- Therefore

$$\mathbf{P}(\sum_{i=1}^{n} X_i > t) \leqslant e^{-\lambda t} \exp(\frac{\lambda^2 c^2 n}{1-\lambda c}).$$

Variance Improvement

- We keep the assumption that $X_1, ..., X_n$ are independent, $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$. Let $\sigma_i^2 = \mathbf{E}X_i^2$.

Variance Improvement

- We keep the assumption that $X_1, ..., X_n$ are independent, $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$. Let $\sigma_i^2 = \mathbf{E}X_i^2$.
- If $|X_i| \leqslant M$, then the previous argument shows that

$$\prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i) \leqslant \exp(\frac{\lambda^2 \sum_{i=1}^{n} \sigma_i^2}{2(1-(1/3)\lambda M)}).$$

Variance Improvement

- We keep the assumption that $X_1, ..., X_n$ are independent, $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$. Let $\sigma_i^2 = \mathbf{E}X_i^2$.
- If $|X_i| \leqslant M$, then the previous argument shows that

$$\prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i) \leqslant \exp\left(\frac{\lambda^2 \sum_{i=1}^{n} \sigma_i^2}{2(1-(1/3)\lambda M)}\right).$$

- When there is no bound $|X_i| \leqslant M$, we can artificially generate it using decomposition

$$X_i = (X_i 1_{|X_i| \leqslant M} - \mathbf{E}X_i 1_{|X_i| \leqslant M}) + (X_i 1_{|X_i| > M} - \mathbf{E}X_i 1_{|X_i| > M}).$$

Variance Improvement

- We keep the assumption that $X_1, ..., X_n$ are independent, $\mathbf{E}X_i = 0$ and $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$. Let $\sigma_i^2 = \mathbf{E}X_i^2$.
- If $|X_i| \leqslant M$, then the previous argument shows that

$$\prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i) \leqslant \exp\left(\frac{\lambda^2 \sum_{i=1}^{n} \sigma_i^2}{2(1-(1/3)\lambda M)}\right).$$

- When there is no bound $|X_i| \leqslant M$, we can artificially generate it using decomposition

$$X_i = (X_i 1_{|X_i| \leqslant M} - \mathbf{E}X_i 1_{|X_i| \leqslant M}) + (X_i 1_{|X_i| > M} - \mathbf{E}X_i 1_{|X_i| > M}).$$

- Choosing $M = c \log n$ one can show that

$$\prod_{i=1}^{n} \mathbf{E}\exp(\lambda X_i) \leqslant L \exp\left(\frac{\lambda^2 \sum_{i=1}^{n} \sigma_i^2}{1-(2/3)\lambda(c\log n)}\right).$$

Basic Inequalities

- (Bennet's Inequality) Suppose that $X_1, X_2, ..., X_n$ are real independent random variables such that $\mathbf{E}X_i = 0$, $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$, then

$$\mathbf{P}(|\sum_{i=1}^{n} X_i| \geqslant t) \leqslant 2\exp(-\frac{t^2}{nc^2 + ct}).$$

Basic Inequalities

- (Bennet's Inequality) Suppose that $X_1, X_2, ..., X_n$ are real independent random variables such that $\mathbf{E}X_i = 0$, $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$, then

$$\mathbf{P}(|\sum_{i=1}^{n} X_i| \geqslant t) \leqslant 2\exp(-\frac{t^2}{nc^2 + ct}).$$

- (Bernstein's Inequality) Suppose that $X_1, X_2, ..., X_n$ are real independent random variables such that $\mathbf{E}X_i = 0$, $|X_i| \leqslant M$, $\sum_{i=1}^{n} \mathbf{E}X_i^2 = n\sigma^2$, then

$$\mathbf{P}(|\sum_{i=1}^{n} X_i| \geqslant t) \leqslant 2\exp(-\frac{t^2}{2(n\sigma^2 + (1/3)Mt)}).$$

CLT Inequality

- Suppose that $X_1, X_2, ..., X_n$ are real independent random variables such that $\mathbf{E}X_i = 0$, $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$, $\sum_{i=1}^{n} \mathbf{E}X_i^2 = n\sigma^2$, then

$$\mathbf{P}(|\sum_{i=1}^{n} X_i| \geqslant t) \leqslant K \exp(-\frac{t^2}{(n\sigma^2 + (2/3)(c \log n)t)}),$$

for some universal constant K.

CLT Inequality

- Suppose that $X_1, X_2, ..., X_n$ are real independent random variables such that $\mathbf{E}X_i = 0$, $\mathbf{E}(\exp(|X_i|/c) - 1) \leqslant 1$, $\sum_{i=1}^n \mathbf{E}X_i^2 = n\sigma^2$, then

$$\mathbf{P}(|\sum_{i=1}^n X_i| \geqslant t) \leqslant K \exp(-\frac{t^2}{(n\sigma^2 + (2/3)(c \log n)t)}),$$

  for some universal constant K.
- The meaning of the result is that for exponentially fast decaying random variables the CLT (Gaussian concentration) can be observed for $t \leqslant (\sigma^2/c)(\sqrt{n}/\log n)$

Markov Chains

- Let $(X_i)_{i \geqslant 0}$ be a Markov chain on a general state space $(\mathcal{X}, \mathcal{B})$.

Markov Chains

- Let $(X_i)_{i \geqslant 0}$ be a Markov chain on a general state space $(\mathcal{X}, \mathcal{B})$.
- We assume that there exists a stationary distribution π on $(\mathcal{X}, \mathcal{B})$.

Markov Chains

- Let $(X_i)_{i\geqslant 0}$ be a Markov chain on a general state space $(\mathcal{X},\mathcal{B})$.
- We assume that there exists a stationary distribution π on $(\mathcal{X},\mathcal{B})$.
- We assume that $(X_n)_{n\geqslant 0}$ verifies geometric drift condition: there exist a small set C, constants $b<\infty$, $\beta>0$ and a function $V\geqslant 1$ satisfying

$$PV(x)-V(x)\leqslant -\beta V(x)+b1_C(x),\quad x\in\mathcal{X}.$$

Markov Chains

- Let $(X_i)_{i \geqslant 0}$ be a Markov chain on a general state space $(\mathcal{X}, \mathcal{B})$.
- We assume that there exists a stationary distribution π on $(\mathcal{X}, \mathcal{B})$.
- We assume that $(X_n)_{n \geqslant 0}$ verifies geometric drift condition: there exist a small set C, constants $b < \infty$, $\beta > 0$ and a function $V \geqslant 1$ satisfying

$$PV(x) - V(x) \leqslant -\beta V(x) + b 1_C(x), \quad x \in \mathcal{X}.$$

- The geometric drift condition is equivalent to the existence of a small set $C \in \mathcal{B}(\mathcal{X})$ and a constant $c > 0$ such that

$$\mathbf{E}_x(\exp(\frac{\tau_C}{c}) - 1) \leqslant 1 \quad \text{for} \quad x \in C.$$

Simplified Setting

- Let $f : \mathcal{X} \to \mathbb{R}$ be a π-integrable function.

Simplified Setting

- Let $f : \mathcal{X} \to \mathbb{R}$ be a π-integrable function.
- The goal is to show concentration inequalities of the type

$$\mathbf{P}_{\mu}(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant \exp(-\Phi(\mu, t, n)),$$

for suitably chosen Φ, where μ is the initial distribution.

Simplified Setting

- Let $f : \mathcal{X} \to \mathbb{R}$ be a π-integrable function.
- The goal is to show concentration inequalities of the type

$$\mathbf{P}_{\mu}(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant \exp(-\Phi(\mu, t, n)),$$

  for suitably chosen Φ, where μ is the initial distribution.
- We first consider the concentration inequalities for a bounded f, i.e. we assume that $|f| \leqslant a$.

Simplified Setting

- Let $f : \mathcal{X} \to \mathbb{R}$ be a π-integrable function.
- The goal is to show concentration inequalities of the type

$$\mathbf{P}_\mu(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant \exp(-\Phi(\mu, t, n)),$$

  for suitably chosen Φ, where μ is the initial distribution.
- We first consider the concentration inequalities for a bounded f, i.e. we assume that $|f| \leqslant a$.
- Whenever geometric drift condition holds, then there exists a set S of full π-measure such that

$$\mathbf{E}_x(\exp((2ac_x)^{-1} \sum_{i=0}^{\tau_C - 1} \bar{f}(X_i)) - 1) \leqslant 1, \quad \text{for } x \in S.$$

Renewal Decomposition

- We can start the Markov chain from any regular point $x \in S$.

Renewal Decomposition

- We can start the Markov chain from any regular point $x \in S$.
- Using the split chain construction the problem can be reduced to the case where a true atom $C = \alpha$-exists.

Renewal Decomposition

- We can start the Markov chain from any regular point $x \in S$.
- Using the split chain construction the problem can be reduced to the case where a true atom $C = \alpha$-exists.
- Let $\tau_\alpha(1) = \tau_\alpha$, $\tau_\alpha(k+1) = \inf\{n > \tau_\alpha(k) : X_n \in \alpha\}$ we define

$$Y_1 = \sum_{i=1}^{\tau_\alpha} \bar{f}(X_i), \quad Y_2 = \sum_{i=\tau_\alpha(N)+1}^{n} \bar{f}(X_i), \quad N = \sum_{k=1}^{n} 1_{X_k \in \alpha}$$

Renewal Decomposition

- We can start the Markov chain from any regular point $x \in S$.
- Using the split chain construction the problem can be reduced to the case where a true atom $C = \alpha$-exists.
- Let $\tau_\alpha(1) = \tau_\alpha$, $\tau_\alpha(k+1) = \inf\{n > \tau_\alpha(k) : X_n \in \alpha\}$ we define

$$Y_1 = \sum_{i=1}^{\tau_\alpha} \bar{f}(X_i), \quad Y_2 = \sum_{i=\tau_\alpha(N)+1}^{n} \bar{f}(X_i), \quad N = \sum_{k=1}^{n} 1_{X_k \in \alpha}$$

- Then

$$\sum_{i=1}^{n} \bar{f}(X_i) = Y_1 + \sum_{i=1}^{N-1} Z_i + Y_2,$$

where $Z_i = \sum_{\tau_\alpha(i)+1}^{\tau_\alpha(i+1)} \bar{f}(X_i)$.

Estimates for the Entrance and Exit Part.

- We use assumption $\mathbf{E}_x(\exp(\frac{\tau_\alpha}{c_x}) - 1) \leqslant 1$, to get

$$\mathbf{P}_x(|Y_1| > t) \leqslant 2\exp(-\frac{t}{2ac_x}).$$

Estimates for the Entrance and Exit Part.

- We use assumption $\mathbf{E}_x(\exp(\frac{\tau_\alpha}{c_x}) - 1) \leqslant 1$, to get

$$\mathbf{P}_x(|Y_1| > t) \leqslant 2\exp(-\frac{t}{2ac_x}).$$

- We use the assumption $\mathbf{E}_\alpha(\exp(\frac{\tau_\alpha}{c}) - 1) \leqslant 1$, to get

$$\mathbf{P}_x(|Y_2| > t) \leqslant 2\exp(-\frac{t}{2\pi(\alpha)ac^2}).$$

Estimates for the Entrance and Exit Part.

- We use assumption $\mathbf{E}_x(\exp(\frac{\tau_\alpha}{c_x}) - 1) \leqslant 1$, to get

$$\mathbf{P}_x(|Y_1| > t) \leqslant 2\exp(-\frac{t}{2ac_x}).$$

- We use the assumption $\mathbf{E}_\alpha(\exp(\frac{\tau_\alpha}{c}) - 1) \leqslant 1$, to get

$$\mathbf{P}_x(|Y_2| > t) \leqslant 2\exp(-\frac{t}{2\pi(\alpha)ac^2}).$$

- In the similar way way we show that

$$\mathbf{P}_x(|Z_N| > t) \leqslant 2\exp(-\frac{t}{2\pi(\alpha)ac^2}).$$

Estimate for the Main Body

- To bound $\sum_{i=1}^{N} Z_i$ we use the martingale method, which leads to

$$\mathbf{E}_x(\exp(\lambda \sum_{i=1}^{N} Z_i)) \leqslant (\mathbf{E}_x(\mathbf{E}_\alpha \exp(2\lambda Z_1))^N)^{1/2}.$$

Estimate for the Main Body

- To bound $\sum_{i=1}^{N} Z_i$ we use the martingale method, which leads to

$$\mathbf{E}_x(\exp(\lambda \sum_{i=1}^{N} Z_i)) \leqslant (\mathbf{E}_x(\mathbf{E}_\alpha \exp(2\lambda Z_1))^N)^{1/2}.$$

- Then one can apply bounds on the Laplace transform of Z_1, e.g.

$$\mathbf{E}_\alpha \exp(2\lambda Z_1) \leqslant \exp(\frac{4\lambda^2 c^2 a^2}{1 - 2\lambda ca}),$$

which gives

$$\mathbf{E}_x(\exp(\lambda \sum_{i=1}^{N} Z_i)) \leqslant (\mathbf{E}_x \exp(\frac{4N\lambda^2 c^2}{1 - 2\lambda c}))^{1/2}.$$

Bounded Case Result

- The construction of N, i.e. $\{N \leqslant k\} = \{\tau_\alpha(k+1) > n\}$, provides that

$$\mathbf{P}_x(N > k) \leqslant \exp(-\frac{k}{4\pi(\alpha)^2 c^2}), \quad \text{for} \quad k \geqslant 2n\pi(\alpha).$$

Bounded Case Result

- The construction of N, i.e. $\{N \leqslant k\} = \{\tau_\alpha(k+1) > n\}$, provides that

$$\mathbf{P}_x(N > k) \leqslant \exp(-\frac{k}{4\pi(\alpha)^2 c^2}), \quad \text{for} \quad k \geqslant 2n\pi(\alpha).$$

Theorem (R.Adamczak,W.B, 2009)

Let $(X_n)_{n\geqslant 0}$ be Markov chain with values in $(\mathcal{X}, \mathcal{B})$ satisfying the geometric drift condition. Then for any regular initial point $x \in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant K \exp(-\frac{1}{K} \min(\frac{t^2}{n\pi(\alpha)c^2 a^2}, \frac{t}{\pi(\alpha)c^2 a}, \frac{t}{c_x a})).$$

Remarks

- The result shows that gaussian type concentration works on $|t| \leqslant a\sqrt{n}$.

Remarks

- The result shows that gaussian type concentration works on $|t| \leqslant a\sqrt{n}$.
- If we apply $N \leqslant n$ and the variance improvement to bound $\mathbf{E}_\alpha \exp(\lambda Z_1)$. Consequently CLT works well for $t \leqslant (\sigma^2/(ac^2))(\sqrt{n}/\log n)$. then the following can be obtained

Remarks

- The result shows that gaussian type concentration works on $|t| \leqslant a\sqrt{n}$.
- If we apply $N \leqslant n$ and the variance improvement to bound $\mathbf{E}_\alpha \exp(\lambda Z_1)$. Consequently CLT works well for $t \leqslant (\sigma^2/(ac^2))(\sqrt{n}/\log n)$. then the following can be obtained

Theorem (R.Adamczak,W.B 2009)

Let $(X_n)_{n \geqslant 0}$ be Markov chain with values in $(\mathcal{X}, \mathcal{B})$ satisfying the geometric drift condition. Then for any regular initial point $x \in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant K \exp(-\frac{1}{K} \min(\frac{t^2}{n\pi(\alpha)\sigma^2}, \frac{t(\log n)^{-1}}{\pi(\alpha)c^2 a}, \frac{t}{c_x a})),$$

where $\sigma^2 = \mathbf{E}_\alpha Z_1^2$

General Setting

- In the general case we assume the multiplicative drift condition: there exists a small set C and a function $V : \mathcal{X} \to \mathbb{R}_+$ and constants $b < \infty, \beta > 0$ such that

$$\exp(-V(x))P(\exp(V))(x) \leqslant \exp(-\beta|f(x)| + b1_C(x)).$$

General Setting

- In the general case we assume the multiplicative drift condition: there exists a small set C and a function $V : \mathcal{X} \to \mathbb{R}_+$ and constants $b < \infty, \beta > 0$ such that

$$\exp(-V(x))P(\exp(V))(x) \leqslant \exp(-\beta|f(x)| + b1_C(x)).$$

- The multiplicative drift condition is a natural setting to show multiplicative ergodicity. It guarantees (I.Kontoyiannis, S.P.Meyn 2005) that there exists a small set C, $d < \infty$ s.t.

$$\mathbf{E}_x(\exp(d^{-1} \sum_{k=0}^{\tau_C - 1} |f(X_k)|) - 1) \leqslant 1, \quad \text{for} \quad x \in C.$$

General Setting

- In the general case we assume the multiplicative drift condition: there exists a small set C and a function $V : \mathcal{X} \to \mathbb{R}_+$ and constants $b < \infty, \beta > 0$ such that

$$\exp(-V(x))P(\exp(V))(x) \leqslant \exp(-\beta|f(x)| + b1_C(x)).$$

- The multiplicative drift condition is a natural setting to show multiplicative ergodicity. It guarantees (I.Kontoyiannis, S.P.Meyn 2005) that there exists a small set C, $d < \infty$ s.t.

$$\mathbf{E}_x(\exp(d^{-1}\sum_{k=0}^{\tau_C-1}|f(X_k)|) - 1) \leqslant 1, \quad \text{for} \quad x \in C.$$

- There exists a set S of full π-measure, such that

$$\mathbf{E}_x(\exp((2d_x)^{-1}\sum_{k=0}^{\tau_C-1}|f(X_k)|) - 1) \leqslant 1, \quad \text{for } x \in \mathcal{X}.$$

Main Result

- Following the idea with the split chain we reduce the problem to the case where a true atom $C = \alpha$ exists.

Main Result

- Following the idea with the split chain we reduce the problem to the case where a true atom $C = \alpha$ exists.
- In the atomic case our assumptions can be rewritten as

$$\mathbf{E}_\alpha(\exp(d^{-1}\sum_{k=0}^{\tau_\alpha - 1}|f(X_k)|) - 1) \leqslant 1, \quad \mathbf{E}_\alpha(\exp(c^{-1}\tau_\alpha) - 1) \leqslant 1$$

Main Result

- Following the idea with the split chain we reduce the problem to the case where a true atom $C = \alpha$ exists.
- In the atomic case our assumptions can be rewritten as

$$\mathbf{E}_\alpha(\exp(d^{-1}\sum_{k=0}^{\tau_\alpha-1}|f(X_k)|)-1)\leqslant 1,\quad \mathbf{E}_\alpha(\exp(c^{-1}\tau_\alpha)-1)\leqslant 1$$

Theorem (R.Adamczak, W.B. 2009)

Let $(X_n)_{n\geqslant 0}$ be Markov chain with values in $(\mathcal{X},\mathcal{B})$ satisfying the multiplicative drift condition. Then for any regular initial point $x\in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n}\bar{f}(X_i)|>t)\leqslant K\exp(-\frac{1}{K}\min(\frac{t^2}{n\pi(\alpha)d^2},\frac{t}{\pi(\alpha)cd},\frac{t}{d_x})).$$

Main Result

Theorem (R.Adamczak, W.B. 2009)

Let $(X_n)_{n\geqslant 0}$ be Markov chain with values in $(\mathcal{X},\mathcal{B})$ satisfying the multiplicative drift condition. Then for any regular initial point $x\in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n}\bar{f}(X_i)|>t)\leqslant K\exp(-\frac{1}{K}\min(\frac{t^2}{n\pi(\alpha)\sigma^2},\frac{t(\log n)^{-1}}{\pi(\alpha)cd},\frac{t}{d_x})),$$

where $\sigma^2=\mathbf{E}_\alpha Z_1^2$

Main Result

Theorem (R.Adamczak, W.B. 2009)

Let $(X_n)_{n\geqslant 0}$ be Markov chain with values in $(\mathcal{X},\mathcal{B})$ satisfying the multiplicative drift condition. Then for any regular initial point $x \in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n}\bar{f}(X_i)| > t) \leqslant K\exp(-\frac{1}{K}\min(\frac{t^2}{n\pi(\alpha)\sigma^2}, \frac{t(\log n)^{-1}}{\pi(\alpha)cd}, \frac{t}{d_x})),$$

where $\sigma^2 = \mathbf{E}_\alpha Z_1^2$

- It shows that the CLT-type concentration can be seen for $t \leqslant (\sigma^2/cd)(\sqrt{n}/\log(n))$.

Main Result

Theorem (R.Adamczak, W.B. 2009)

Let $(X_n)_{n\geqslant 0}$ be Markov chain with values in $(\mathcal{X},\mathcal{B})$ satisfying the multiplicative drift condition. Then for any regular initial point $x \in S$

$$\mathbf{P}_x(|\sum_{i=1}^{n} \bar{f}(X_i)| > t) \leqslant K \exp(-\frac{1}{K} \min(\frac{t^2}{n\pi(\alpha)\sigma^2}, \frac{t(\log n)^{-1}}{\pi(\alpha)cd}, \frac{t}{d_x})),$$

where $\sigma^2 = \mathbf{E}_\alpha Z_1^2$

- It shows that the CLT-type concentration can be seen for $t \leqslant (\sigma^2/cd)(\sqrt{n}/\log(n))$.
- The result allows to compute the independent-like concentration inequality in terms of parameters given in drift conditions.