

Singular-Value Gap of Nonreversible Markov Processes

Ruochuan Xu

May 2026

Stanford University

Outline

Motivation

Singular-value gap

Main results

Rotating OU process

Relations to other mixing quantities

Two applications

Past and future directions

Motivation

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Two notions of convergence

- Let $(X_t)_{t \geq 0}$ be a Markov process with an invariant probability measure μ .
- The semigroup $(P_t)_{t \geq 0}$ acts on $L^2(\mu)$ functions as $P_t f(x) = \mathbb{E}_x[f(X_t)]$
- The infinitesimal generator L is given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.

Two natural questions:

1. **Time- t law:** how fast does the law of X_t approach μ ?
2. **Empirical averages:** for centered $g \in L_0^2(\mu)$, how fast does

$$\mu_t g := \frac{1}{t} \int_0^t g(X_s) ds$$

approach 0?

Under stationarity $X_0 \sim \mu$, we measure convergence of empirical averages by the empirical variance

$$\mathbb{E}_\mu |\mu_t g|^2.$$

MCMC interpretation

These convergence questions address the burn-in time to reach (approximate) stationarity and time required for accurate sampling in stationarity.

Reversible spectral gap and convergence

Assume $L = L^*$ on $L^2(\mu)$ and $-L \geq \lambda I$ on $L_0^2(\mu)$. Let $\|\cdot\|$ denote the $L^2(\mu)$ norm.

Semigroup relaxation

For centered g ,

$$\|P_t g\| \leq e^{-\lambda t} \|g\|.$$

Stationary empirical averages

Using the covariance identity and reversibility,

$$\begin{aligned}\mathbb{E}_\mu |\mu_t g|^2 &= \frac{2}{t^2} \int_0^t (t-s) \langle g, P_s g \rangle \, ds \\ &= \frac{2}{t^2} \int_0^t (t-s) \|P_{s/2} g\|^2 \, ds \\ &\leq \frac{2}{\lambda t} \|g\|^2.\end{aligned}$$

Finite-state analogue: Levin-Peres-Wilmer, Lemma 12.22.

Reversible spectral gap and convergence

Assume $L = L^*$ on $L^2(\mu)$ and $-L \geq \lambda I$ on $L_0^2(\mu)$. Let $\|\cdot\|$ denote the $L^2(\mu)$ norm.

Semigroup relaxation

For centered g ,

$$\|P_t g\| \leq e^{-\lambda t} \|g\|.$$

Stationary empirical averages

Using the covariance identity and reversibility,

$$\begin{aligned}\mathbb{E}_\mu |\mu_t g|^2 &= \frac{2}{t^2} \int_0^t (t-s) \langle g, P_s g \rangle \, ds \\ &= \frac{2}{t^2} \int_0^t (t-s) \|P_{s/2} g\|^2 \, ds \\ &\leq \frac{2}{\lambda t} \|g\|^2.\end{aligned}$$

Finite-state analogue: Levin-Peres-Wilmer, Lemma 12.22.

Reversible spectral gap and convergence

Assume $L = L^*$ on $L^2(\mu)$ and $-L \geq \lambda I$ on $L_0^2(\mu)$. Let $\|\cdot\|$ denote the $L^2(\mu)$ norm.

Semigroup relaxation

For centered g ,

$$\|P_t g\| \leq e^{-\lambda t} \|g\|.$$

Stationary empirical averages

Using the covariance identity and reversibility,

$$\begin{aligned} \mathbb{E}_\mu |\mu_t g|^2 &= \frac{2}{t^2} \int_0^t (t-s) \langle g, P_s g \rangle \, ds \\ &= \frac{2}{t^2} \int_0^t (t-s) \|P_{s/2} g\|^2 \, ds \\ &\leq \frac{2}{\lambda t} \|g\|^2. \end{aligned}$$

Finite-state analogue: Levin-Peres-Wilmer, Lemma 12.22.

Reversible spectral gap and convergence

Assume $L = L^*$ on $L^2(\mu)$ and $-L \geq \lambda I$ on $L_0^2(\mu)$. Let $\|\cdot\|$ denote the $L^2(\mu)$ norm.

Semigroup relaxation

For centered g ,

$$\|P_t g\| \leq e^{-\lambda t} \|g\|.$$

Stationary empirical averages

Using the covariance identity and reversibility,

$$\begin{aligned} \mathbb{E}_\mu |\mu_t g|^2 &= \frac{2}{t^2} \int_0^t (t-s) \langle g, P_s g \rangle \, ds \\ &= \frac{2}{t^2} \int_0^t (t-s) \|P_{s/2} g\|^2 \, ds \\ &\leq \frac{2}{\lambda t} \|g\|^2. \end{aligned}$$

Finite-state analogue: Levin-Peres-Wilmer, Lemma 12.22.

Nonreversibility separates the stories

For nonreversible L , several spectral quantities no longer agree:

real-part spectral gap
semigroup contraction rate
symmetrized spectral gap
pseudo-spectral gap
singular-value gap

We study the singular-value gap as the quantity controlling empirical averages.

Key observation

singular-value gap $\gamma > 0 \iff -L^{-1}$ exists on $L_0^2(\mu)$ with norm bounded by γ^{-1} .

- Solving the Poisson equation allows us to control g through f that satisfies

$$-Lf = g$$

- Control of empirical averages by the singular-value gap is therefore not surprising, because we are in an “easy” CLT regime (Bhattacharya’82).

Nonreversibility separates the stories

For nonreversible L , several spectral quantities no longer agree:

real-part spectral gap
semigroup contraction rate
symmetrized spectral gap
pseudo-spectral gap
singular-value gap

We study the singular-value gap as the quantity controlling empirical averages.

Key observation

singular-value gap $\gamma > 0 \iff -L^{-1}$ exists on $L_0^2(\mu)$ with norm bounded by γ^{-1} .

- Solving the Poisson equation allows us to control g through f that satisfies

$$-Lf = g$$

- Control of empirical averages by the singular-value gap is therefore not surprising, because we are in an “easy” CLT regime (Bhattacharya’82).

Nonreversibility separates the stories

For nonreversible L , several spectral quantities no longer agree:

real-part spectral gap
semigroup contraction rate
symmetrized spectral gap
pseudo-spectral gap
singular-value gap

We study the singular-value gap as the quantity controlling empirical averages.

Key observation

singular-value gap $\gamma > 0 \iff -L^{-1}$ exists on $L_0^2(\mu)$ with norm bounded by γ^{-1} .

- Solving the Poisson equation allows us to control g through f that satisfies

$$-Lf = g$$

- Control of empirical averages by the singular-value gap is therefore not surprising, because we are in an “easy” CLT regime (Bhattacharya’82).

Nonreversibility separates the stories

For nonreversible L , several spectral quantities no longer agree:

real-part spectral gap
semigroup contraction rate
symmetrized spectral gap
pseudo-spectral gap
singular-value gap

We study the singular-value gap as the quantity controlling empirical averages.

Key observation

singular-value gap $\gamma > 0 \iff -L^{-1}$ exists on $L_0^2(\mu)$ with norm bounded by γ^{-1} .

- Solving the Poisson equation allows us to control g through f that satisfies

$$-Lf = g$$

- Control of empirical averages by the singular-value gap is therefore not surprising, because we are in an “easy” CLT regime (Bhattacharya’82).

Nonreversibility separates the stories

For nonreversible L , several spectral quantities no longer agree:

real-part spectral gap
semigroup contraction rate
symmetrized spectral gap
pseudo-spectral gap
singular-value gap

We study the singular-value gap as the quantity controlling empirical averages.

Key observation

singular-value gap $\gamma > 0 \iff -L^{-1}$ exists on $L_0^2(\mu)$ with norm bounded by γ^{-1} .

- Solving the Poisson equation allows us to control g through f that satisfies

$$-Lf = g$$

- Control of empirical averages by the singular-value gap is therefore not surprising, because we are in an “easy” CLT regime (Bhattacharya’82).

Singular-value gap

Definition

- For a Markov chain with transition operator P , the generator is $L = P - I$.
- For a strongly continuous Markov semigroup $(P_t)_{t \geq 0}$, the generator L is a possibly unbounded operator on $L^2(\mu)$ given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.
- Conservation of probability implies $L1 = 0$.

Singular-value gap

$$\gamma := \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|}.$$

Definition

- For a Markov chain with transition operator P , the generator is $L = P - I$.
- For a strongly continuous Markov semigroup $(P_t)_{t \geq 0}$, the generator L is a possibly unbounded operator on $L^2(\mu)$ given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.
- Conservation of probability implies $L1 = 0$.

Singular-value gap

$$\gamma := \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|}.$$

Definition

- For a Markov chain with transition operator P , the generator is $L = P - I$.
- For a strongly continuous Markov semigroup $(P_t)_{t \geq 0}$, the generator L is a possibly unbounded operator on $L^2(\mu)$ given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.
- Conservation of probability implies $L\mathbf{1} = 0$.

Singular-value gap

$$\gamma := \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|}.$$

Definition

- For a Markov chain with transition operator P , the generator is $L = P - I$.
- For a strongly continuous Markov semigroup $(P_t)_{t \geq 0}$, the generator L is a possibly unbounded operator on $L^2(\mu)$ given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.
- Conservation of probability implies $L\mathbf{1} = 0$.

Singular-value gap

$$\gamma := \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|}.$$

Definition

- For a Markov chain with transition operator P , the generator is $L = P - I$.
- For a strongly continuous Markov semigroup $(P_t)_{t \geq 0}$, the generator L is a possibly unbounded operator on $L^2(\mu)$ given by $Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$ on $\text{Dom}(L)$.
- Conservation of probability implies $L\mathbf{1} = 0$.

Singular-value gap

$$\gamma := \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|}.$$

Alternative characterizations of γ

Singular-value gap

$$\gamma := \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \|Lf\| / \|f\|.$$

The absolute value of L is

$$|L| := (L^*L)^{1/2}.$$

Then

$$\|Lf\| = \||L|f\|.$$

Since $|L|$ is nonnegative and self-adjoint,

$$\gamma = \inf \text{spec}(|L|_{L_0^2(\mu)}) \quad \text{equivalently} \quad \gamma^2 = \inf \text{spec}(L^*L|_{L_0^2(\mu)})$$

Finite state space

γ is exactly the second-smallest singular value of the matrix $-L = I - P$.

If L is normal

The spectral theorem gives

$$\gamma = \text{dist}(0, \text{spec}(L)).$$

Alternative characterizations of γ

Singular-value gap

$$\gamma := \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \|Lf\| / \|f\|.$$

The absolute value of L is

$$|L| := (L^*L)^{1/2}.$$

Then

$$\|Lf\| = \||L|f\|.$$

Since $|L|$ is nonnegative and self-adjoint,

$$\gamma = \inf \text{spec}(|L|_{L_0^2(\mu)}) \quad \text{equivalently} \quad \gamma^2 = \inf \text{spec}(L^*L|_{L_0^2(\mu)})$$

Finite state space

γ is exactly the second-smallest singular value of the matrix $-L = I - P$.

If L is normal

The spectral theorem gives

$$\gamma = \text{dist}(0, \text{spec}(L)).$$

Alternative characterizations of γ

Singular-value gap

$$\gamma := \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \|Lf\| / \|f\|.$$

The absolute value of L is

$$|L| := (L^*L)^{1/2}.$$

Then

$$\|Lf\| = \||L|f\|.$$

Since $|L|$ is nonnegative and self-adjoint,

$$\gamma = \inf \text{spec}(|L|_{L_0^2(\mu)}) \quad \text{equivalently} \quad \gamma^2 = \inf \text{spec}(L^*L|_{L_0^2(\mu)})$$

Finite state space

γ is exactly the second-smallest singular value of the matrix $-L = I - P$.

If L is normal

The spectral theorem gives

$$\gamma = \text{dist}(0, \text{spec}(L)).$$

Alternative characterizations of γ

Singular-value gap

$$\gamma := \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \|Lf\| / \|f\|.$$

The absolute value of L is

$$|L| := (L^*L)^{1/2}.$$

Then

$$\|Lf\| = \||L|f\|.$$

Since $|L|$ is nonnegative and self-adjoint,

$$\gamma = \inf \text{spec}(|L|_{L_0^2(\mu)}) \quad \text{equivalently} \quad \gamma^2 = \inf \text{spec}(L^*L|_{L_0^2(\mu)})$$

Finite state space

γ is exactly the second-smallest singular value of the matrix $-L = I - P$.

If L is normal

The spectral theorem gives

$$\gamma = \text{dist}(0, \text{spec}(L)).$$

Alternative characterizations of γ

Singular-value gap

$$\gamma := \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \|Lf\| / \|f\|.$$

The absolute value of L is

$$|L| := (L^*L)^{1/2}.$$

Then

$$\|Lf\| = \||L|f\|.$$

Since $|L|$ is nonnegative and self-adjoint,

$$\gamma = \inf \text{spec}(|L|_{L_0^2(\mu)}) \quad \text{equivalently} \quad \gamma^2 = \inf \text{spec}(L^*L|_{L_0^2(\mu)})$$

Finite state space

γ is exactly the second-smallest singular value of the matrix $-L = I - P$.

If L is normal

The spectral theorem gives

$$\gamma = \text{dist}(0, \text{spec}(L)).$$

Singular-value gap $\neq$ spectral gap of L

For nonnormal operators, eigenvalues do not determine singular values.

Let S_ℓ be the nilpotent shift on ℓ states

$$e_1 \mapsto e_2 \mapsto \cdots \mapsto e_\ell \mapsto 0.$$

The generator is

$$L_\ell = S_\ell - I.$$

Then

$$\text{spec}(L_\ell) = \{-1\}, \quad \text{dist}(0, \text{spec}(L_\ell)) = 1.$$

But

$$\gamma(L_\ell) = s_2(-L_\ell) = 2 \sin\left(\frac{\pi}{4\ell + 2}\right) \asymp \frac{1}{\ell}.$$

Singular-value gap $\neq$ spectral gap of L

For nonnormal operators, eigenvalues do not determine singular values.

Let S_ℓ be the nilpotent shift on ℓ states

$$e_1 \mapsto e_2 \mapsto \cdots \mapsto e_\ell \mapsto 0.$$

The generator is

$$L_\ell = S_\ell - I.$$

Then

$$\text{spec}(L_\ell) = \{-1\}, \quad \text{dist}(0, \text{spec}(L_\ell)) = 1.$$

But

$$\gamma(L_\ell) = s_2(-L_\ell) = 2 \sin\left(\frac{\pi}{4\ell + 2}\right) \asymp \frac{1}{\ell}.$$

Singular-value gap $\neq$ spectral gap of L

For nonnormal operators, eigenvalues do not determine singular values.

Let S_ℓ be the nilpotent shift on ℓ states

$$e_1 \mapsto e_2 \mapsto \cdots \mapsto e_\ell \mapsto 0.$$

The generator is

$$L_\ell = S_\ell - I.$$

Then

$$\text{spec}(L_\ell) = \{-1\}, \quad \text{dist}(0, \text{spec}(L_\ell)) = 1.$$

But

$$\gamma(L_\ell) = s_2(-L_\ell) = 2 \sin\left(\frac{\pi}{4\ell + 2}\right) \asymp \frac{1}{\ell}.$$

Singular-value gap $\neq$ symmetrized spectral gap

Let γ_s be the spectral gap of the self-adjoint operator $\frac{L+L^*}{2}$. Consider

$$dX_t = \alpha dt + \sqrt{2} dB_t \quad \text{on } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}, \quad L_\alpha = \partial_\theta^2 + \alpha \partial_\theta.$$

Fourier modes $e_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$ satisfy

$$L_\alpha e^{ik\theta} = (-k^2 + i\alpha k) e^{ik\theta}.$$

Singular-value gap

$$\gamma = \min_{k \neq 0} \sqrt{k^4 + \alpha^2 k^2} = \sqrt{1 + \alpha^2}$$

$$\mathbb{E}_\mu |\mu_t e_k|^2 \asymp \frac{2}{(k^2 + \alpha^2)t}$$

Symmetrized spectral gap

$$\gamma_s = \min_{k \neq 0} k^2 = 1$$

$$\|P_t\|_{L_0^2(\mu) \rightarrow L_0^2(\mu)} = e^{-t}$$

Empirical averages converge faster with larger $|\alpha|$, but the time- t distribution does not. The separation of two notions of convergence is captured, respectively, by γ_s and γ .

Singular-value gap $\neq$ symmetrized spectral gap

Let γ_s be the spectral gap of the self-adjoint operator $\frac{L+L^*}{2}$. Consider

$$dX_t = \alpha dt + \sqrt{2} dB_t \quad \text{on } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}, \quad L_\alpha = \partial_\theta^2 + \alpha \partial_\theta.$$

Fourier modes $e_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$ satisfy

$$L_\alpha e^{ik\theta} = (-k^2 + i\alpha k) e^{ik\theta}.$$

Singular-value gap

$$\gamma = \min_{k \neq 0} \sqrt{k^4 + \alpha^2 k^2} = \sqrt{1 + \alpha^2}$$

$$\mathbb{E}_\mu |\mu_t e_k|^2 \asymp \frac{2}{(k^2 + \alpha^2)t}$$

Symmetrized spectral gap

$$\gamma_s = \min_{k \neq 0} k^2 = 1$$

$$\|P_t\|_{L_0^2(\mu) \rightarrow L_0^2(\mu)} = e^{-t}$$

Empirical averages converge faster with larger $|\alpha|$, but the time- t distribution does not.
The separation of two notions of convergence is captured, respectively, by γ_s and γ .

Singular-value gap $\neq$ symmetrized spectral gap

Let γ_s be the spectral gap of the self-adjoint operator $\frac{L+L^*}{2}$. Consider

$$dX_t = \alpha dt + \sqrt{2} dB_t \quad \text{on } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}, \quad L_\alpha = \partial_\theta^2 + \alpha \partial_\theta.$$

Fourier modes $e_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$ satisfy

$$L_\alpha e^{ik\theta} = (-k^2 + i\alpha k) e^{ik\theta}.$$

Singular-value gap

$$\gamma = \min_{k \neq 0} \sqrt{k^4 + \alpha^2 k^2} = \sqrt{1 + \alpha^2}$$

$$\mathbb{E}_\mu |\mu_t e_k|^2 \asymp \frac{2}{(k^2 + \alpha^2)t}$$

Symmetrized spectral gap

$$\gamma_s = \min_{k \neq 0} k^2 = 1$$

$$\|P_t\|_{L_0^2(\mu) \rightarrow L_0^2(\mu)} = e^{-t}$$

Empirical averages converge faster with larger $|\alpha|$, but the time- t distribution does not.
The separation of two notions of convergence is captured, respectively, by γ_s and γ .

Singular-value gap $\neq$ symmetrized spectral gap

Let γ_s be the spectral gap of the self-adjoint operator $\frac{L+L^*}{2}$. Consider

$$dX_t = \alpha dt + \sqrt{2} dB_t \quad \text{on } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}, \quad L_\alpha = \partial_\theta^2 + \alpha \partial_\theta.$$

Fourier modes $e_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$ satisfy

$$L_\alpha e^{ik\theta} = (-k^2 + i\alpha k) e^{ik\theta}.$$

Singular-value gap

$$\gamma = \min_{k \neq 0} \sqrt{k^4 + \alpha^2 k^2} = \sqrt{1 + \alpha^2}$$

$$\mathbb{E}_\mu |\mu_t e_k|^2 \asymp \frac{2}{(k^2 + \alpha^2)t}$$

Symmetrized spectral gap

$$\gamma_s = \min_{k \neq 0} k^2 = 1$$

$$\|P_t\|_{L_0^2(\mu) \rightarrow L_0^2(\mu)} = e^{-t}$$

Empirical averages converge faster with larger $|\alpha|$, but the time- t distribution does not.
The separation of two notions of convergence is captured, respectively, by γ_s and γ .

Singular-value gap $\neq$ symmetrized spectral gap

Let γ_s be the spectral gap of the self-adjoint operator $\frac{L+L^*}{2}$. Consider

$$dX_t = \alpha dt + \sqrt{2} dB_t \quad \text{on } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}, \quad L_\alpha = \partial_\theta^2 + \alpha \partial_\theta.$$

Fourier modes $e_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$ satisfy

$$L_\alpha e^{ik\theta} = (-k^2 + i\alpha k) e^{ik\theta}.$$

Singular-value gap

$$\gamma = \min_{k \neq 0} \sqrt{k^4 + \alpha^2 k^2} = \sqrt{1 + \alpha^2}$$

$$\mathbb{E}_\mu |\mu_t e_k|^2 \asymp \frac{2}{(k^2 + \alpha^2)t}$$

Symmetrized spectral gap

$$\gamma_s = \min_{k \neq 0} k^2 = 1$$

$$\|P_t\|_{L_0^2(\mu) \rightarrow L_0^2(\mu)} = e^{-t}$$

Empirical averages converge faster with larger $|\alpha|$, but the time- t distribution does not. The separation of two notions of convergence is captured, respectively, by γ_s and γ .

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Connection with Chatterjee's gap

Chatterjee's finite-state discrete-time gap is defined as the second-smallest singular value of $I - P$:

$$\gamma_{\text{Ch}} = s_2(I - P),$$

Our definition of γ is a consistent extension. Chatterjee's main theorem also characterizes the convergence of empirical averages with the singular-value gap.

New ingredients

- continuous time Markov processes
- general measurable state spaces
- subspaces of observables
- Poisson solver and asymptotic-variance

Same philosophy:

singular-value gap of the generator $\implies$ uniform convergence of empirical-averages

Main results

Local observable classes

We formulate our results locally, because under an irreversible perturbation, empirical averages may converge faster only for a subspace of observables.

Closed invariant subspace

We fix a closed subspace $K \subset H$ such that

$$P_t K \subset K, \quad t \geq 0.$$

Let L_K be the generator of the restricted semigroup:

$$\text{Dom}(L_K) = \text{Dom}(L) \cap K.$$

Local singular-value gap

$$\gamma_K := \inf_{\substack{f \in \text{Dom}(L_K) \\ f \neq 0}} \frac{\|L_K f\|}{\|f\|}.$$

Local observable classes

We formulate our results locally, because under an irreversible perturbation, empirical averages may converge faster only for a subspace of observables.

Closed invariant subspace

We fix a closed subspace $K \subset H$ such that

$$P_t K \subset K, \quad t \geq 0.$$

Let L_K be the generator of the restricted semigroup:

$$\text{Dom}(L_K) = \text{Dom}(L) \cap K.$$

Local singular-value gap

$$\gamma_K := \inf_{\substack{f \in \text{Dom}(L_K) \\ f \neq 0}} \frac{\|L_K f\|}{\|f\|}.$$

Local observable classes

We formulate our results locally, because under an irreversible perturbation, empirical averages may converge faster only for a subspace of observables.

Closed invariant subspace

We fix a closed subspace $K \subset H$ such that

$$P_t K \subset K, \quad t \geq 0.$$

Let L_K be the generator of the restricted semigroup:

$$\text{Dom}(L_K) = \text{Dom}(L) \cap K.$$

Local singular-value gap

$$\gamma_K := \inf_{\substack{f \in \text{Dom}(L_K) \\ f \neq 0}} \frac{\|L_K f\|}{\|f\|}.$$

Main theorem: finite-time empirical variance

For a subspace $K \subset L_0^2(\mu)$, define the worst-case standard deviation in K :

$$\Delta_t^K := \sup_{\substack{g \in K \\ \|g\|=1}} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

When $K = L_0^2(\mu)$, we use the notation Δ_t .

Theorem 1: Upper and lower bounds for empirical variance

Assume $\gamma_K > 0$. Then

$$\Delta_t^K \leq \frac{2}{\sqrt{\gamma_K t}}.$$

Moreover, we have the worst-case over bound over a time interval:

$$\sup_{s \in [t, 2t]} \Delta_s^K \geq \frac{1}{\frac{3}{2}\gamma_K t + 3}.$$

- Pointwise lower bound is not applicable due to deterministic drift.
- Taking $K = L_0^2(\mu)$ gives the global version when $\gamma > 0$.
- We can also deduce $\Delta_t \geq 1/36$ when $t \leq 1/(3\gamma)$.

Main theorem: finite-time empirical variance

For a subspace $K \subset L_0^2(\mu)$, define the worst-case standard deviation in K :

$$\Delta_t^K := \sup_{\substack{g \in K \\ \|g\|=1}} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

When $K = L_0^2(\mu)$, we use the notation Δ_t .

Theorem 1: Upper and lower bounds for empirical variance

Assume $\gamma_K > 0$. Then

$$\Delta_t^K \leq \frac{2}{\sqrt{\gamma_K t}}.$$

Moreover, we have the worst-case over bound over a time interval:

$$\sup_{s \in [t, 2t]} \Delta_s^K \geq \frac{1}{\frac{3}{2}\gamma_K t + 3}.$$

- Pointwise lower bound is not applicable due to deterministic drift.
- Taking $K = L_0^2(\mu)$ gives the global version when $\gamma > 0$.
- We can also deduce $\Delta_t \geq 1/36$ when $t \leq 1/(3\gamma)$.

Main theorem: finite-time empirical variance

For a subspace $K \subset L_0^2(\mu)$, define the worst-case standard deviation in K :

$$\Delta_t^K := \sup_{\substack{g \in K \\ \|g\|=1}} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

When $K = L_0^2(\mu)$, we use the notation Δ_t .

Theorem 1: Upper and lower bounds for empirical variance

Assume $\gamma_K > 0$. Then

$$\Delta_t^K \leq \frac{2}{\sqrt{\gamma_K t}}.$$

Moreover, we have the worst-case over bound over a time interval:

$$\sup_{s \in [t, 2t]} \Delta_s^K \geq \frac{1}{\frac{3}{2}\gamma_K t + 3}.$$

- Pointwise lower bound is not applicable due to deterministic drift.
- Taking $K = L_0^2(\mu)$ gives the global version when $\gamma > 0$.
- We can also deduce $\Delta_t \geq 1/36$ when $t \leq 1/(3\gamma)$.

Main theorem: finite-time empirical variance

For a subspace $K \subset L_0^2(\mu)$, define the worst-case standard deviation in K :

$$\Delta_t^K := \sup_{\substack{g \in K \\ \|g\|=1}} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

When $K = L_0^2(\mu)$, we use the notation Δ_t .

Theorem 1: Upper and lower bounds for empirical variance

Assume $\gamma_K > 0$. Then

$$\Delta_t^K \leq \frac{2}{\sqrt{\gamma_K t}}.$$

Moreover, we have the worst-case over bound over a time interval:

$$\sup_{s \in [t, 2t]} \Delta_s^K \geq \frac{1}{\frac{3}{2}\gamma_K t + 3}.$$

- Pointwise lower bound is not applicable due to deterministic drift.
- Taking $K = L_0^2(\mu)$ gives the global version when $\gamma > 0$.
- We can also deduce $\Delta_t \geq 1/36$ when $t \leq 1/(3\gamma)$.

Main theorem: finite-time empirical variance

For a subspace $K \subset L_0^2(\mu)$, define the worst-case standard deviation in K :

$$\Delta_t^K := \sup_{\substack{g \in K \\ \|g\|=1}} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

When $K = L_0^2(\mu)$, we use the notation Δ_t .

Theorem 1: Upper and lower bounds for empirical variance

Assume $\gamma_K > 0$. Then

$$\Delta_t^K \leq \frac{2}{\sqrt{\gamma_K t}}.$$

Moreover, we have the worst-case over bound over a time interval:

$$\sup_{s \in [t, 2t]} \Delta_s^K \geq \frac{1}{\frac{3}{2}\gamma_K t + 3}.$$

- Pointwise lower bound is not applicable due to deterministic drift.
- Taking $K = L_0^2(\mu)$ gives the global version when $\gamma > 0$.
- We can also deduce $\Delta_t \geq 1/36$ when $t \leq 1/(3\gamma)$.

Sharpness examples

Upper bound is sharp: stationary refreshment

At rate q , replace the current state by an independent sample from μ . On $L_0^2(\mu)$,

$$L = -qI, \quad \gamma = q.$$

Then

$$(\mathbb{E}_\mu |\mu_t g|^2)^{1/2} \sim \sqrt{\frac{2}{\gamma t}} \|g\|.$$

This is just the continuous time analog of i.i.d. sampling and classical CLT fluctuations.

Lower bound is sharp: pure rotation

On $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$,

$$L = \alpha \partial_\theta, \quad \gamma = |\alpha|.$$

For a Fourier mode $e_k(\theta) = e^{ik\theta}$,

$$(\mathbb{E}_\mu |\mu_t e_k|^2)^{1/2} = \left| \frac{e^{i\alpha k t} - 1}{i\alpha k t} \right| \leq \frac{2}{|\alpha k| t}.$$

The dyadic lower bound has the correct $1/(\gamma t)$ order.

Sharpness examples

Upper bound is sharp: stationary refreshment

At rate q , replace the current state by an independent sample from μ . On $L_0^2(\mu)$,

$$L = -qI, \quad \gamma = q.$$

Then

$$(\mathbb{E}_\mu |\mu_t g|^2)^{1/2} \sim \sqrt{\frac{2}{\gamma t}} \|g\|.$$

This is just the continuous time analog of i.i.d. sampling and classical CLT fluctuations.

Sharpness examples

Upper bound is sharp: stationary refreshment

At rate q , replace the current state by an independent sample from μ . On $L^2_0(\mu)$,

$$L = -qI, \quad \gamma = q.$$

Then

$$(\mathbb{E}_\mu |\mu_t g|^2)^{1/2} \sim \sqrt{\frac{2}{\gamma t}} \|g\|.$$

This is just the continuous time analog of i.i.d. sampling and classical CLT fluctuations.

Lower bound is sharp: pure rotation

On $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$,

$$L = \alpha \partial_\theta, \quad \gamma = |\alpha|.$$

For a Fourier mode $e_k(\theta) = e^{ik\theta}$,

$$(\mathbb{E}_\mu |\mu_t e_k|^2)^{1/2} = \left| \frac{e^{i\alpha k t} - 1}{i\alpha k t} \right| \leq \frac{2}{|\alpha k| t}.$$

The dyadic lower bound has the correct $1/(\gamma t)$ order.

Positive gap implies invertible generator

Now we sketch the proof for the main theorem.

Key general state-space input: Let L be the generator, in discrete or continuous time.

Positive gap implies invertible generator

Now we sketch the proof for the main theorem.

Key general state-space input: Let L be the generator, in discrete or continuous time.

Invertibility lemma

If

$$\gamma = \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \frac{\|Lf\|}{\|f\|} > 0,$$

then

$$-L : \text{Dom}(L) \cap L_0^2(\mu) \rightarrow L_0^2(\mu)$$

is bijective and

$$\|-L^{-1}\| \leq \gamma^{-1}.$$

Idea: consider the resolvent on $L_0^2(\mu)$

$$(\lambda I - L)^{-1};$$

send $\lambda \downarrow 0$ to obtain $-L^{-1}$.

This is the analog of the finite-dimensional fact: a matrix is invertible if it has positive singular values.

Positive gap implies invertible generator

Now we sketch the proof for the main theorem.

Key general state-space input: Let L be the generator, in discrete or continuous time.

Invertibility lemma

If

$$\gamma = \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \frac{\|Lf\|}{\|f\|} > 0,$$

then

$$-L : \text{Dom}(L) \cap L_0^2(\mu) \rightarrow L_0^2(\mu)$$

is bijective and

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

Idea: consider the resolvent on $L_0^2(\mu)$

$$(\lambda I - L)^{-1};$$

send $\lambda \downarrow 0$ to obtain $-L^{-1}$.

This is the analog of the finite-dimensional fact: a matrix is invertible if it has positive singular values.

Positive gap implies invertible generator

Now we sketch the proof for the main theorem.

Key general state-space input: Let L be the generator, in discrete or continuous time.

Invertibility lemma

If

$$\gamma = \inf_{f \in \text{Dom}(L) \cap L_0^2(\mu) \setminus \{0\}} \frac{\|Lf\|}{\|f\|} > 0,$$

then

$$-L : \text{Dom}(L) \cap L_0^2(\mu) \rightarrow L_0^2(\mu)$$

is bijective and

$$\|-L^{-1}\| \leq \gamma^{-1}.$$

Idea: consider the resolvent on $L_0^2(\mu)$

$$(\lambda I - L)^{-1};$$

send $\lambda \downarrow 0$ to obtain $-L^{-1}$.

This is the analog of the finite-dimensional fact: a matrix is invertible if it has positive singular values.

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Take $g \in L_0^2(\mu)$ and choose $f \in L_0^2(\mu) \cap \text{Dom}(L)$ such that

$$g = -Lf.$$

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Take $g \in L_0^2(\mu)$ and choose $f \in L_0^2(\mu) \cap \text{Dom}(L)$ such that

$$g = -Lf.$$

Then the *Cesàro average* of g is

$$A_t g := \frac{1}{t} \int_0^t P_s g \, ds = \frac{1}{t} \int_0^t P_s (-Lf) \, ds = \frac{f - P_t f}{t}.$$

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Take $g \in L_0^2(\mu)$ and choose $f \in L_0^2(\mu) \cap \text{Dom}(L)$ such that

$$g = -Lf.$$

Then the *Cesàro average* of g is

$$A_t g := \frac{1}{t} \int_0^t P_s g \, ds = \frac{1}{t} \int_0^t P_s (-L f) \, ds = \frac{f - P_t f}{t}.$$

Since P_t is contractive,

$$\|A_t g\| \leq \frac{2}{t} \|f\| \leq \frac{2}{\gamma t} \|g\|.$$

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Take $g \in L_0^2(\mu)$ and choose $f \in L_0^2(\mu) \cap \text{Dom}(L)$ such that

$$g = -Lf.$$

Then the *Cesàro average* of g is

$$A_t g := \frac{1}{t} \int_0^t P_s g \, ds = \frac{1}{t} \int_0^t P_s (-L f) \, ds = \frac{f - P_t f}{t}.$$

Since P_t is contractive,

$$\|A_t g\| \leq \frac{2}{t} \|f\| \leq \frac{2}{\gamma t} \|g\|.$$

Stationary covariance identity:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{2}{t^2} \operatorname{Re} \int_0^t u \langle g, A_u g \rangle du.$$

Proof sketch: upper bound

$$\Delta_t \lesssim 1/\sqrt{\gamma t}$$

Take $g \in L_0^2(\mu)$ and choose $f \in L_0^2(\mu) \cap \text{Dom}(L)$ such that

$$g = -Lf.$$

Then the *Cesàro average* of g is

$$A_t g := \frac{1}{t} \int_0^t P_s g \, ds = \frac{1}{t} \int_0^t P_s (-Lf) \, ds = \frac{f - P_t f}{t}.$$

Since P_t is contractive,

$$\|A_t g\| \leq \frac{2}{t} \|f\| \leq \frac{2}{\gamma t} \|g\|.$$

Stationary covariance identity:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{2}{t^2} \operatorname{Re} \int_0^t u \langle g, A_u g \rangle \, du.$$

Combine with Cauchy-Schwarz to get

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

Proof sketch: lower bound

$$\sup_{s \in [t, 2t]} \Delta_s \gtrsim 1/(1 + \gamma t)$$

Choose a near-minimizer

$$f \in \text{Dom}(L) \cap L_0^2(\mu), \quad \|f\| = 1, \quad \|Lf\| \approx \gamma.$$

Set

$$g = -Lf.$$

Decompose f using Cesaro averages over $[t, 2t]$:

$$f = \left(f - \frac{1}{t} \int_t^{2t} P_s f \, ds \right) + \frac{1}{t} \int_t^{2t} P_s f \, ds =: u_t + v_t.$$

The first term is controlled by averages of g ; the second by averages of f .

$1 = \|f\| \leq \|u_t\| + \|v_t\|$ forces large worst-case empirical variance on $[t, 2t]$.

Proof sketch: lower bound

$$\sup_{s \in [t, 2t]} \Delta_s \gtrsim 1/(1 + \gamma t)$$

Choose a near-minimizer

$$f \in \text{Dom}(L) \cap L_0^2(\mu), \quad \|f\| = 1, \quad \|Lf\| \approx \gamma.$$

Set

$$g = -Lf.$$

Decompose f using Cesaro averages over $[t, 2t]$:

$$f = \left(f - \frac{1}{t} \int_t^{2t} P_s f \, ds \right) + \frac{1}{t} \int_t^{2t} P_s f \, ds =: u_t + v_t.$$

The first term is controlled by averages of g ; the second by averages of f .

$1 = \|f\| \leq \|u_t\| + \|v_t\|$ forces large worst-case empirical variance on $[t, 2t]$.

Proof sketch: lower bound

$$\sup_{s \in [t, 2t]} \Delta_s \gtrsim 1/(1 + \gamma t)$$

Choose a near-minimizer

$$f \in \text{Dom}(L) \cap L_0^2(\mu), \quad \|f\| = 1, \quad \|Lf\| \approx \gamma.$$

Set

$$g = -Lf.$$

Decompose f using Cesaro averages over $[t, 2t]$:

$$f = \left(f - \frac{1}{t} \int_t^{2t} P_s f \, ds \right) + \frac{1}{t} \int_t^{2t} P_s f \, ds =: u_t + v_t.$$

The first term is controlled by averages of g ; the second by averages of f .

$1 = \|f\| \leq \|u_t\| + \|v_t\|$ forces large worst-case empirical variance on $[t, 2t]$.

Proof sketch: lower bound

$$\sup_{s \in [t, 2t]} \Delta_s \gtrsim 1/(1 + \gamma t)$$

Choose a near-minimizer

$$f \in \text{Dom}(L) \cap L_0^2(\mu), \quad \|f\| = 1, \quad \|Lf\| \approx \gamma.$$

Set

$$g = -Lf.$$

Decompose f using Cesaro averages over $[t, 2t]$:

$$f = \left(f - \frac{1}{t} \int_t^{2t} P_s f ds \right) + \frac{1}{t} \int_t^{2t} P_s f ds =: u_t + v_t.$$

The first term is controlled by averages of g ; the second by averages of f .

Proof sketch: lower bound

$$\sup_{s \in [t, 2t]} \Delta_s \gtrsim 1/(1 + \gamma t)$$

Choose a near-minimizer

$$f \in \text{Dom}(L) \cap L_0^2(\mu), \quad \|f\| = 1, \quad \|Lf\| \approx \gamma.$$

Set

$$g = -Lf.$$

Decompose f using Cesaro averages over $[t, 2t]$:

$$f = \left(f - \frac{1}{t} \int_t^{2t} P_s f \, ds \right) + \frac{1}{t} \int_t^{2t} P_s f \, ds =: u_t + v_t.$$

The first term is controlled by averages of g ; the second by averages of f .

$1 = \|f\| \leq \|u_t\| + \|v_t\|$ forces large worst-case empirical variance on $[t, 2t]$.

Discrete-time general-state version

Let P be a Markov kernel with invariant μ , and let

$$\gamma = \inf_{f \in L_0^2(\mu), f \neq 0} \frac{\|(P - I)f\|}{\|f\|}.$$

Discrete-time general-state version

Let P be a Markov kernel with invariant μ , and let

$$\gamma = \inf_{f \in L_0^2(\mu), f \neq 0} \frac{\|(P - I)f\|}{\|f\|}.$$

For

$$\Delta_n = \sup_{\|g\|=1} \left(\mathbb{E}_\mu \left| \frac{1}{n} \sum_{k=0}^{n-1} g(X_k) \right|^2 \right)^{1/2},$$

we obtain the exact analog of Chatterjee's main theorem:

$$\Delta_n \leq \frac{2}{\sqrt{n\gamma}}, \quad \sup_{k \in [n, 2n]} \Delta_k \geq \frac{1}{C(1+n\gamma)},$$

Discrete-time general-state version

Let P be a Markov kernel with invariant μ , and let

$$\gamma = \inf_{f \in L_0^2(\mu), f \neq 0} \frac{\|(P - I)f\|}{\|f\|}.$$

For

$$\Delta_n = \sup_{\|g\|=1} \left(\mathbb{E}_\mu \left| \frac{1}{n} \sum_{k=0}^{n-1} g(X_k) \right|^2 \right)^{1/2},$$

we obtain the exact analog of Chatterjee's main theorem:

$$\Delta_n \leq \frac{2}{\sqrt{n\gamma}}, \quad \sup_{k \in [n, 2n]} \Delta_k \geq \frac{1}{C(1+n\gamma)},$$

thanks to being able to solve the Poisson equation $-Lf = g$.

Discrete-time general-state version

Let P be a Markov kernel with invariant μ , and let

$$\gamma = \inf_{f \in L_0^2(\mu), f \neq 0} \frac{\|(P - I)f\|}{\|f\|}.$$

For

$$\Delta_n = \sup_{\|g\|=1} \left(\mathbb{E}_\mu \left| \frac{1}{n} \sum_{k=0}^{n-1} g(X_k) \right|^2 \right)^{1/2},$$

we obtain the exact analog of Chatterjee's main theorem:

$$\Delta_n \leq \frac{2}{\sqrt{n\gamma}}, \quad \sup_{k \in [n, 2n]} \Delta_k \geq \frac{1}{C(1 + n\gamma)},$$

thanks to being able to solve the Poisson equation $-Lf = g$.

Note

No topology is needed: this is an $L^2(\mu)$ operator argument on a measurable space.

Reminder: asymptotic variance and CLT

For a fixed centered observable g , the asymptotic variance is

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

If

$$-Lf = g$$

has a suitable solution f , then Dynkin's formula gives

$$\int_0^t g(X_s) ds = M_t - (f(X_t) - f(X_0)).$$

The boundary term is negligible at CLT scale.

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

If

$$-Lf = g$$

has a suitable solution f , then Dynkin's formula gives

$$\int_0^t g(X_s) ds = M_t - (f(X_t) - f(X_0)).$$

The boundary term is negligible at CLT scale.

Even without convergence in distribution, we have the second-moment formula

$$\sigma_q^2 = 2 \operatorname{Re} \langle g, -L^{-1}g \rangle.$$

Reminder: asymptotic variance and CLT

For a fixed centered observable g , the asymptotic variance is

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

Poisson equation perspective

If

$$-Lf = g$$

has a suitable solution f , then Dynkin's formula gives

$$\int_0^t g(X_s) ds = M_t - (f(X_t) - f(X_0)).$$

The boundary term is negligible at CLT scale.

Even without convergence in distribution, we have the second-moment formula

$$\sigma_g^2 = 2 \operatorname{Re} \langle g, -L^{-1}g \rangle.$$

- Bhattacharya'82: CLT for g in the range of L .
- Kipnis-Varadhan'86: CLT for reversible semigroup and $g \in H_{-1} \cap L_0^2(\mu)$
- Our setting: a positive singular-value gap gives a bounded Poisson solver, so CLT holds for all functions in $L_0^2(\mu)$.

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

If

$$-Lf = g$$

has a suitable solution f , then Dynkin's formula gives

$$\int_0^t g(X_s) ds = M_t - (f(X_t) - f(X_0)).$$

The boundary term is negligible at CLT scale.

Even without convergence in distribution, we have the second-moment formula

$$\sigma_q^2 = 2 \operatorname{Re} \langle g, -L^{-1}g \rangle.$$

- Bhattacharya'82: CLT for g in the range of L .
- Kipnis-Varadhan'86: CLT for reversible semigroup and $g \in H_{-1} \cap L_0^2(\mu)$
- Our setting: a positive singular-value gap gives a bounded Poisson solver, so CLT holds for all functions in $L_0^2(\mu)$.

$$\sigma_g^2 := \lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t g|^2$$

If

$$-Lf = g$$

has a suitable solution f , then Dynkin's formula gives

$$\int_0^t g(X_s) ds = M_t - (f(X_t) - f(X_0)).$$

The boundary term is negligible at CLT scale.

Even without convergence in distribution, we have the second-moment formula

$$\sigma_g^2 = 2 \operatorname{Re} \langle g, -L^{-1}g \rangle.$$

- Bhattacharya'82: CLT for g in the range of L .
- Kipnis-Varadhan'86: CLT for reversible semigroup and $g \in H_{-1} \cap L_0^2(\mu)$
- Our setting: a positive singular-value gap gives a bounded Poisson solver, so CLT holds for all functions in $L_0^2(\mu)$.

Refinement of finite-time bounds by asymptotic variance

Assume $K \subset L_0^2(\mu)$ is closed invariant and $\gamma_K > 0$. For $g \in K$,

$$\sigma_g^2 = 2 \operatorname{Re} \left\langle g, -L_K^{-1} g \right\rangle \leq \frac{2}{\gamma_K t}.$$

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

Define the worst-case quantities

$$\sigma_K^2 := \sup_{g \in K, \|g\|=1} \sigma_g^2 \qquad \Delta_t^K := \sup_{g \in K, \|g\|=1} \left(\mathbb{E}_\mu |\mu_t g|^2 \right)^{1/2}.$$

Then

$$\frac{\sigma_K^2}{t} - \frac{4}{\gamma_K^2 t^2} \leq (\Delta_t^K)^2 \leq \frac{\sigma_K^2}{t} + \frac{4}{\gamma_K^2 t^2}.$$

- While γ^{-1} is the time scale for approximate convergence of empirical averages, the long-time coefficient for finite-time variance is still the asymptotic variance, when it is finite and positive.
- Improvement in the upper bound is constant; the lower bound is more interesting.

Refinement of finite-time bounds by asymptotic variance

Assume $K \subset L_0^2(\mu)$ is closed invariant and $\gamma_K > 0$. For $g \in K$,

$$\sigma_g^2 = 2 \operatorname{Re} \left\langle g, -L_K^{-1} g \right\rangle \leq \frac{2}{\gamma_K t}.$$

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

Define the worst-case quantities

$$\sigma_K^2 := \sup_{g \in K, \|g\|=1} \sigma_g^2 \qquad \Delta_t^K := \sup_{g \in K, \|g\|=1} (\mathbb{E}_\mu |\mu_t g|^2)^{1/2}.$$

Then

$$\frac{\sigma_K^2}{t} - \frac{4}{\gamma_K^2 t^2} \leq (\Delta_t^K)^2 \leq \frac{\sigma_K^2}{t} + \frac{4}{\gamma_K^2 t^2}.$$

- While γ^{-1} is the time scale for approximate convergence of empirical averages, the long-time coefficient for finite-time variance is still the asymptotic variance, when it is finite and positive.
- Improvement in the upper bound is constant; the lower bound is more interesting.

Refinement of finite-time bounds by asymptotic variance

Assume $K \subset L_0^2(\mu)$ is closed invariant and $\gamma_K > 0$. For $g \in K$,

$$\sigma_g^2 = 2 \operatorname{Re} \left\langle g, -L_K^{-1} g \right\rangle \leq \frac{2}{\gamma_K t}.$$

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

Define the worst-case quantities

$$\sigma_K^2 := \sup_{g \in K, \|g\|=1} \sigma_g^2 \qquad \Delta_t^K := \sup_{g \in K, \|g\|=1} \left(\mathbb{E}_\mu |\mu_t g|^2 \right)^{1/2}.$$

Then

$$\frac{\sigma_K^2}{t} - \frac{4}{\gamma_K^2 t^2} \leq (\Delta_t^K)^2 \leq \frac{\sigma_K^2}{t} + \frac{4}{\gamma_K^2 t^2}.$$

- While γ^{-1} is the time scale for approximate convergence of empirical averages, the long-time coefficient for finite-time variance is still the asymptotic variance, when it is finite and positive.
- Improvement in the upper bound is constant; the lower bound is more interesting.

Refinement of finite-time bounds by asymptotic variance

Assume $K \subset L_0^2(\mu)$ is closed invariant and $\gamma_K > 0$. For $g \in K$,

$$\sigma_g^2 = 2 \operatorname{Re} \left\langle g, -L_K^{-1} g \right\rangle \leq \frac{2}{\gamma_K t}.$$

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

Define the worst-case quantities

$$\sigma_K^2 := \sup_{g \in K, \|g\|=1} \sigma_g^2 \qquad \Delta_t^K := \sup_{g \in K, \|g\|=1} \left(\mathbb{E}_\mu |\mu_t g|^2 \right)^{1/2}.$$

Then

$$\frac{\sigma_K^2}{t} - \frac{4}{\gamma_K^2 t^2} \leq (\Delta_t^K)^2 \leq \frac{\sigma_K^2}{t} + \frac{4}{\gamma_K^2 t^2}.$$

- While γ^{-1} is the time scale for approximate convergence of empirical averages, the long-time coefficient for finite-time variance is still the asymptotic variance, when it is finite and positive.
- Improvement in the upper bound is constant; the lower bound is more interesting.

Refinement of finite-time bounds by asymptotic variance

Assume $K \subset L_0^2(\mu)$ is closed invariant and $\gamma_K > 0$. For $g \in K$,

$$\sigma_g^2 = 2 \operatorname{Re} \left\langle g, -L_K^{-1} g \right\rangle \leq \frac{2}{\gamma_K t}.$$

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

Define the worst-case quantities

$$\sigma_K^2 := \sup_{g \in K, \|g\|=1} \sigma_g^2 \qquad \Delta_t^K := \sup_{g \in K, \|g\|=1} \left(\mathbb{E}_\mu |\mu_t g|^2 \right)^{1/2}.$$

Then

$$\frac{\sigma_K^2}{t} - \frac{4}{\gamma_K^2 t^2} \leq (\Delta_t^K)^2 \leq \frac{\sigma_K^2}{t} + \frac{4}{\gamma_K^2 t^2}.$$

- While γ^{-1} is the time scale for approximate convergence of empirical averages, the long-time coefficient for finite-time variance is still the asymptotic variance, when it is finite and positive.
- Improvement in the upper bound is constant; the lower bound is more interesting.

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale

$$t^{-1/2}.$$

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as

$$t^{-1}.$$

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale

$$t^{-1/2}.$$

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as

$$t^{-1}.$$

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale

$$t^{-1/2}.$$

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as

$$t^{-1}.$$

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale
 $t^{-1/2}.$

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as
 $t^{-1}.$

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale
 $t^{-1/2}$.

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as
 t^{-1} .

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Two regimes: diffusive vs. oscillatory

Refined finite-time variance bounds

$$\frac{\sigma_g^2}{t} - \frac{4}{\gamma_K^2 t^2} \|g\|^2 \leq \mathbb{E}_\mu |\mu_t g|^2 \leq \frac{\sigma_g^2}{t} + \frac{4}{\gamma_K^2 t^2} \|g\|^2.$$

The refined lower bound is vacuous unless $\sigma_g^2 > 0$

Positive asymptotic variance

If $\sigma_g^2 > 0$,

$$\mathbb{E}_\mu |\mu_t g|^2 \sim \frac{\sigma_g^2}{t}.$$

Empirical standard deviation has the CLT scale
 $t^{-1/2}$.

Zero asymptotic variance

If deterministic dynamics dominate s.t.

$$\sigma_g^2 = 0.$$

Then empirical standard deviation can decay as
 t^{-1} .

- Example: pure rotation on the circle has $\gamma > 0$ but $\sigma_g^2 = 0$ for every $g \in L_0^2(\mu)$.
- The refined upper bound can also fail if $\gamma_K = 0$ and we consider the worst case over an infinite-dimensional observable class K .

Rotating OU process

Rotating Ornstein-Uhlenbeck process

Let $\mu = N(0, I_2)$ and

$$dX_t = (-X_t + \alpha J X_t) dt + \sqrt{2} dW_t, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

In polar coordinates,

$$L_\alpha = \Delta - x \cdot \nabla + \alpha \partial_\theta.$$

The symmetric part is the usual Ornstein-Uhlenbeck generator; the skew-symmetric part rotates around the origin. L_α is normal with discrete spectrum:

Eigenbasis

$$\psi_{k,m}(r, \theta) = c_{k,m} r^{|m|} L_k^{(|m|)}(r^2/2) e^{im\theta}, \quad k \geq 0, m \in \mathbb{Z}.$$

$$L_\alpha \psi_{k,m} = (-(2k + |m|) + i\alpha m) \psi_{k,m}.$$

Bounds by singular-value gap(s) do not give more information because empirical variance is explicitly computable in the eigenbasis, but they offer a new perspective.

Rotating Ornstein-Uhlenbeck process

Let $\mu = N(0, I_2)$ and

$$dX_t = (-X_t + \alpha J X_t) dt + \sqrt{2} dW_t, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

In polar coordinates,

$$L_\alpha = \Delta - x \cdot \nabla + \alpha \partial_\theta.$$

The symmetric part is the usual Ornstein-Uhlenbeck generator; the skew-symmetric part rotates around the origin. L_α is normal with discrete spectrum:

Eigenbasis

$$\psi_{k,m}(r, \theta) = c_{k,m} r^{|m|} L_k^{(|m|)}(r^2/2) e^{im\theta}, \quad k \geq 0, \quad m \in \mathbb{Z}.$$

$$L_\alpha \psi_{k,m} = (-(2k + |m|) + i\alpha m) \psi_{k,m}.$$

Bounds by singular-value gap(s) do not give more information because empirical variance is explicitly computable in the eigenbasis, but they offer a new perspective.

Rotating Ornstein-Uhlenbeck process

Let $\mu = N(0, I_2)$ and

$$dX_t = (-X_t + \alpha J X_t) dt + \sqrt{2} dW_t, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

In polar coordinates,

$$L_\alpha = \Delta - x \cdot \nabla + \alpha \partial_\theta.$$

The symmetric part is the usual Ornstein-Uhlenbeck generator; the skew-symmetric part rotates around the origin. L_α is normal with discrete spectrum:

Eigenbasis

$$\psi_{k,m}(r, \theta) = c_{k,m} r^{|m|} L_k^{(|m|)}(r^2/2) e^{im\theta}, \quad k \geq 0, m \in \mathbb{Z}.$$

$$L_\alpha \psi_{k,m} = (-(2k + |m|) + i\alpha m) \psi_{k,m}.$$

Bounds by singular-value gap(s) do not give more information because empirical variance is explicitly computable in the eigenbasis, but they offer a new perspective.

Rotating Ornstein-Uhlenbeck process

Let $\mu = N(0, I_2)$ and

$$dX_t = (-X_t + \alpha J X_t) dt + \sqrt{2} dW_t, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

In polar coordinates,

$$L_\alpha = \Delta - x \cdot \nabla + \alpha \partial_\theta.$$

The symmetric part is the usual Ornstein-Uhlenbeck generator; the skew-symmetric part rotates around the origin. L_α is normal with discrete spectrum:

Eigenbasis

$$\psi_{k,m}(r, \theta) = c_{k,m} r^{|m|} L_k^{(|m|)}(r^2/2) e^{im\theta}, \quad k \geq 0, \quad m \in \mathbb{Z}.$$

$$L_\alpha \psi_{k,m} = (-(2k + |m|) + i\alpha m) \psi_{k,m}.$$

Bounds by singular-value gap(s) do not give more information because empirical variance is explicitly computable in the eigenbasis, but they offer a new perspective.

Full gap vs. nonradial gap: accelerated convergence only on a subspace

Since L_α is normal, singular values are eigenvalue moduli.

Full centered space

$$\gamma_{\text{full}} = \min_{(k,m) \neq (0,0)} \sqrt{(2k + |m|)^2 + \alpha^2 m^2} = \min\{\sqrt{1 + \alpha^2}, 2\}.$$

Radial mode $|x|^2 - 2$ blocks global acceleration.

Nonradial observables

Consider the orthogonal complement of radial observables:

$$K_{\text{ang}} = \overline{\text{span}}\{\psi_{k,m} : m \neq 0\}.$$

Then

$$\gamma_{\text{ang}} = \sqrt{1 + \alpha^2}.$$

So for every $g \in K_{\text{ang}}$,

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{t\sqrt{1 + \alpha^2}} \|g\|^2.$$

It makes sense that rotation should only accelerate convergence for the empirical averages of nonradial observables.

Full gap vs. nonradial gap: accelerated convergence only on a subspace

Since L_α is normal, singular values are eigenvalue moduli.

Full centered space

$$\gamma_{\text{full}} = \min_{(k,m) \neq (0,0)} \sqrt{(2k + |m|)^2 + \alpha^2 m^2} = \min\{\sqrt{1 + \alpha^2}, 2\}.$$

Radial mode $|x|^2 - 2$ blocks global acceleration.

Nonradial observables

Consider the orthogonal complement of radial observables:

$$K_{\text{ang}} = \overline{\text{span}}\{\psi_{k,m} : m \neq 0\}.$$

Then

$$\gamma_{\text{ang}} = \sqrt{1 + \alpha^2}.$$

So for every $g \in K_{\text{ang}}$,

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{t\sqrt{1 + \alpha^2}} \|g\|^2.$$

It makes sense that rotation should only accelerate convergence for the empirical averages of nonradial observables.

Full gap vs. nonradial gap: accelerated convergence only on a subspace

Since L_α is normal, singular values are eigenvalue moduli.

Full centered space

$$\gamma_{\text{full}} = \min_{(k,m) \neq (0,0)} \sqrt{(2k + |m|)^2 + \alpha^2 m^2} = \min\{\sqrt{1 + \alpha^2}, 2\}.$$

Radial mode $|x|^2 - 2$ blocks global acceleration.

Nonradial observables

Consider the orthogonal complement of radial observables:

$$K_{\text{ang}} = \overline{\text{span}}\{\psi_{k,m} : m \neq 0\}.$$

Then

$$\gamma_{\text{ang}} = \sqrt{1 + \alpha^2}.$$

So for every $g \in K_{\text{ang}}$,

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{t\sqrt{1 + \alpha^2}} \|g\|^2.$$

It makes sense that rotation should only accelerate convergence for the empirical averages of nonradial observables.

Full gap vs. nonradial gap: accelerated convergence only on a subspace

Since L_α is normal, singular values are eigenvalue moduli.

Full centered space

$$\gamma_{\text{full}} = \min_{(k,m) \neq (0,0)} \sqrt{(2k + |m|)^2 + \alpha^2 m^2} = \min\{\sqrt{1 + \alpha^2}, 2\}.$$

Radial mode $|x|^2 - 2$ blocks global acceleration.

Nonradial observables

Consider the orthogonal complement of radial observables:

$$K_{\text{ang}} = \overline{\text{span}}\{\psi_{k,m} : m \neq 0\}.$$

Then

$$\gamma_{\text{ang}} = \sqrt{1 + \alpha^2}.$$

So for every $g \in K_{\text{ang}}$,

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{t\sqrt{1 + \alpha^2}} \|g\|^2.$$

It makes sense that rotation should only accelerate convergence for the empirical averages of nonradial observables.

Full gap vs. nonradial gap: accelerated convergence only on a subspace

Since L_α is normal, singular values are eigenvalue moduli.

Full centered space

$$\gamma_{\text{full}} = \min_{(k,m) \neq (0,0)} \sqrt{(2k + |m|)^2 + \alpha^2 m^2} = \min\{\sqrt{1 + \alpha^2}, 2\}.$$

Radial mode $|x|^2 - 2$ blocks global acceleration.

Nonradial observables

Consider the orthogonal complement of radial observables:

$$K_{\text{ang}} = \overline{\text{span}}\{\psi_{k,m} : m \neq 0\}.$$

Then

$$\gamma_{\text{ang}} = \sqrt{1 + \alpha^2}.$$

So for every $g \in K_{\text{ang}}$,

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{t\sqrt{1 + \alpha^2}} \|g\|^2.$$

It makes sense that rotation should only accelerate convergence for the empirical averages of nonradial observables.

Linear observables: asymptotic variance gives more precise information

For $L_\alpha \psi_{k,m} = (-n + i\omega)\psi_{k,m}$, $n = 2k + |m|$, $\omega = \alpha m$, one computes

$$\lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t \psi_{k,m}|^2 = \frac{2n}{n^2 + \omega^2}.$$

The n -th Gaussian chaos is

$$\mathcal{H}_n = \text{span}\{\psi_{k,m} : 2k + |m| = n\}.$$

$\mathcal{H}_1$: centered linear observables

$\mathcal{H}_1 = \text{span}\{\psi_{0,1}, \psi_{0,-1}\} = \text{span}\{x_1, x_2\}$, with

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2}.$$

Upper bound without asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{4}{\sqrt{1 + \alpha^2} t}$$

Upper bound with asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{2}{(1 + \alpha^2)t} + \frac{4}{(1 + \alpha^2)t^2}.$$

The asymptotic-variance refinement sees the sharper α^{-2} factor.

Linear observables: asymptotic variance gives more precise information

For $L_\alpha \psi_{k,m} = (-n + i\omega)\psi_{k,m}$, $n = 2k + |m|$, $\omega = \alpha m$, one computes

$$\lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t \psi_{k,m}|^2 = \frac{2n}{n^2 + \omega^2}.$$

The n -th Gaussian chaos is

$$\mathcal{H}_n = \text{span}\{\psi_{k,m} : 2k + |m| = n\}.$$

$\mathcal{H}_1$: centered linear observables

$\mathcal{H}_1 = \text{span}\{\psi_{0,1}, \psi_{0,-1}\} = \text{span}\{x_1, x_2\}$, with

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2}.$$

Upper bound without asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{4}{\sqrt{1 + \alpha^2} t}$$

Upper bound with asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{2}{(1 + \alpha^2)t} + \frac{4}{(1 + \alpha^2)t^2}.$$

The asymptotic-variance refinement sees the sharper α^{-2} factor.

Linear observables: asymptotic variance gives more precise information

For $L_\alpha \psi_{k,m} = (-n + i\omega)\psi_{k,m}$, $n = 2k + |m|$, $\omega = \alpha m$, one computes

$$\lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t \psi_{k,m}|^2 = \frac{2n}{n^2 + \omega^2}.$$

The n -th Gaussian chaos is

$$\mathcal{H}_n = \text{span}\{\psi_{k,m} : 2k + |m| = n\}.$$

$\mathcal{H}_1$: centered linear observables

$\mathcal{H}_1 = \text{span}\{\psi_{0,1}, \psi_{0,-1}\} = \text{span}\{x_1, x_2\}$, with

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2}.$$

Upper bound without asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{4}{\sqrt{1 + \alpha^2} t}$$

Upper bound with asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{2}{(1 + \alpha^2)t} + \frac{4}{(1 + \alpha^2)t^2}.$$

The asymptotic-variance refinement sees the sharper α^{-2} factor.

Linear observables: asymptotic variance gives more precise information

For $L_\alpha \psi_{k,m} = (-n + i\omega)\psi_{k,m}$, $n = 2k + |m|$, $\omega = \alpha m$, one computes

$$\lim_{t \rightarrow \infty} t \mathbb{E}_\mu |\mu_t \psi_{k,m}|^2 = \frac{2n}{n^2 + \omega^2}.$$

The n -th Gaussian chaos is

$$\mathcal{H}_n = \text{span}\{\psi_{k,m} : 2k + |m| = n\}.$$

$\mathcal{H}_1$: centered linear observables

$\mathcal{H}_1 = \text{span}\{\psi_{0,1}, \psi_{0,-1}\} = \text{span}\{x_1, x_2\}$, with

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2}.$$

Upper bound without asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{4}{\sqrt{1 + \alpha^2} t}$$

Upper bound with asymptotic variance:

$$\left(\Delta_t^{\mathcal{H}_1}\right)^2 \leq \frac{2}{(1 + \alpha^2)t} + \frac{4}{(1 + \alpha^2)t^2}.$$

The asymptotic-variance refinement sees the sharper α^{-2} factor.

Practical implication: variance reduction for observable classes

In dimension d , consider

$$L_{J,\alpha} = \Delta - x \cdot \nabla + \alpha(Jx) \cdot \nabla, \quad J^T = -J,$$

with invariant law $N(0, I_d)$.

On linear observables, we can compute

$$\gamma_{\mathcal{H}_1}(J)^2 = 1 + \alpha^2 \lambda_{\min}(J^T J).$$

When d is even, under the constraint $\|J\|_F = 1$, the best choice maximizing $\gamma_{\mathcal{H}_1}(J)$ is

$$J^T J = \frac{1}{d} I.$$

For instance, J can consist of $d/2$ blocks of 2-dimensional rotation. With this choice,

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2/d}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2/d}.$$

This perspective is in contrast with the conventional perspective on variance reduction, where the emphasis is on reducing asymptotic variance for a fixed observable.

Practical implication: variance reduction for observable classes

In dimension d , consider

$$L_{J,\alpha} = \Delta - x \cdot \nabla + \alpha(Jx) \cdot \nabla, \quad J^T = -J,$$

with invariant law $N(0, I_d)$.

On linear observables, we can compute

$$\gamma_{\mathcal{H}_1}(J)^2 = 1 + \alpha^2 \lambda_{\min}(J^T J).$$

When d is even, under the constraint $\|J\|_F = 1$, the best choice maximizing $\gamma_{\mathcal{H}_1}(J)$ is

$$J^T J = \frac{1}{d} I.$$

For instance, J can consist of $d/2$ blocks of 2-dimensional rotation. With this choice,

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2/d}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2/d}.$$

This perspective is in contrast with the conventional perspective on variance reduction, where the emphasis is on reducing asymptotic variance for a fixed observable.

Practical implication: variance reduction for observable classes

In dimension d , consider

$$L_{J,\alpha} = \Delta - x \cdot \nabla + \alpha(Jx) \cdot \nabla, \quad J^T = -J,$$

with invariant law $N(0, I_d)$.

On linear observables, we can compute

$$\gamma_{\mathcal{H}_1}(J)^2 = 1 + \alpha^2 \lambda_{\min}(J^T J).$$

When d is even, under the constraint $\|J\|_F = 1$, the best choice maximizing $\gamma_{\mathcal{H}_1}(J)$ is

$$J^T J = \frac{1}{d} I.$$

For instance, J can consist of $d/2$ blocks of 2-dimensional rotation. With this choice,

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2/d}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2/d}.$$

This perspective is in contrast with the conventional perspective on variance reduction, where the emphasis is on reducing asymptotic variance for a fixed observable.

Practical implication: variance reduction for observable classes

In dimension d , consider

$$L_{J,\alpha} = \Delta - x \cdot \nabla + \alpha(Jx) \cdot \nabla, \quad J^T = -J,$$

with invariant law $N(0, I_d)$.

On linear observables, we can compute

$$\gamma_{\mathcal{H}_1}(J)^2 = 1 + \alpha^2 \lambda_{\min}(J^T J).$$

When d is even, under the constraint $\|J\|_F = 1$, the best choice maximizing $\gamma_{\mathcal{H}_1}(J)$ is

$$J^T J = \frac{1}{d} I.$$

For instance, J can consist of $d/2$ blocks of 2-dimensional rotation. With this choice,

$$\gamma_{\mathcal{H}_1} = \sqrt{1 + \alpha^2/d}, \quad \sigma_{\mathcal{H}_1}^2 = \frac{2}{1 + \alpha^2/d}.$$

This perspective is in contrast with the conventional perspective on variance reduction, where the emphasis is on reducing asymptotic variance for a fixed observable.

Relations to other mixing quantities

Symmetrized gap

Let

$$L_s := \frac{L + L^*}{2}.$$

This is the generator of the additive reversibilized process.

Let γ_s be the spectral gap of L_s .

Symmetrized gap lower bounds singular-value gap

$$\gamma_s \leq \gamma.$$

Proof is one line:

$$\|Lf\| \|f\| \geq |\langle -Lf, f \rangle| \geq \operatorname{Re} \langle -Lf, f \rangle = \langle -L_s f, f \rangle \geq \gamma_s \|f\|^2.$$

But there is no general reverse inequality: Brownian motion with drift α on the circle has

$$\gamma_s = 1, \quad \gamma = \sqrt{1 + \alpha^2}.$$

Symmetrized gap

Let

$$L_s := \frac{L + L^*}{2}.$$

This is the generator of the additive reversibilized process.

Let γ_s be the spectral gap of L_s .

Symmetrized gap lower bounds singular-value gap

$$\gamma_s \leq \gamma.$$

Proof is one line:

$$\|Lf\| \|f\| \geq |\langle -Lf, f \rangle| \geq \operatorname{Re} \langle -Lf, f \rangle = \langle -L_s f, f \rangle \geq \gamma_s \|f\|^2.$$

But there is no general reverse inequality: Brownian motion with drift α on the circle has

$$\gamma_s = 1, \quad \gamma = \sqrt{1 + \alpha^2}.$$

Symmetrized gap

Let

$$L_s := \frac{L + L^*}{2}.$$

This is the generator of the additive reversibilized process.

Let γ_s be the spectral gap of L_s .

Symmetrized gap lower bounds singular-value gap

$$\gamma_s \leq \gamma.$$

Proof is one line:

$$\|Lf\| \|f\| \geq |\langle -Lf, f \rangle| \geq \operatorname{Re} \langle -Lf, f \rangle = \langle -L_s f, f \rangle \geq \gamma_s \|f\|^2.$$

But there is no general reverse inequality: Brownian motion with drift α on the circle has

$$\gamma_s = 1, \quad \gamma = \sqrt{1 + \alpha^2}.$$

Symmetrized gap

Let

$$L_s := \frac{L + L^*}{2}.$$

This is the generator of the additive reversibilized process.

Let γ_s be the spectral gap of L_s .

Symmetrized gap lower bounds singular-value gap

$$\gamma_s \leq \gamma.$$

Proof is one line:

$$\|Lf\| \|f\| \geq |\langle -Lf, f \rangle| \geq \operatorname{Re} \langle -Lf, f \rangle = \langle -L_s f, f \rangle \geq \gamma_s \|f\|^2.$$

But there is no general reverse inequality: Brownian motion with drift α on the circle has

$$\gamma_s = 1, \quad \gamma = \sqrt{1 + \alpha^2}.$$

Bounded-rate jump processes

For a jump generator

$$Lf(x) = \int (f(y) - f(x))q(x, dy), \quad q(x, S) \leq q_{\max},$$

we can bound the singular-value gap above with the symmetrized gap.

Upper bound on the singular-value gap for jump processes

$$\gamma \leq \sqrt{2q_{\max}\gamma_s}.$$

Key estimate:

$$\|Lf\|^2 \leq 2q_{\max} \langle f, -L_s f \rangle.$$

This is a nonreversible analog of bounding the generator by the Dirichlet form when jump rates are bounded.

Bounded-rate jump processes

For a jump generator

$$Lf(x) = \int (f(y) - f(x))q(x, dy), \quad q(x, S) \leq q_{\max},$$

we can bound the singular-value gap above with the symmetrized gap.

Upper bound on the singular-value gap for jump processes

$$\gamma \leq \sqrt{2q_{\max}\gamma_s}.$$

Key estimate:

$$\|Lf\|^2 \leq 2q_{\max} \langle f, -L_s f \rangle.$$

This is a nonreversible analog of bounding the generator by the Dirichlet form when jump rates are bounded.

Bounded-rate jump processes

For a jump generator

$$Lf(x) = \int (f(y) - f(x))q(x, dy), \quad q(x, S) \leq q_{\max},$$

we can bound the singular-value gap above with the symmetrized gap.

Upper bound on the singular-value gap for jump processes

$$\gamma \leq \sqrt{2q_{\max}\gamma_s}.$$

Key estimate:

$$\|Lf\|^2 \leq 2q_{\max} \langle f, -L_s f \rangle.$$

This is a nonreversible analog of bounding the generator by the Dirichlet form when jump rates are bounded.

Bounded-rate jump processes

For a jump generator

$$Lf(x) = \int (f(y) - f(x))q(x, dy), \quad q(x, S) \leq q_{\max},$$

we can bound the singular-value gap above with the symmetrized gap.

Upper bound on the singular-value gap for jump processes

$$\gamma \leq \sqrt{2q_{\max}\gamma_s}.$$

Key estimate:

$$\|Lf\|^2 \leq 2q_{\max} \langle f, -L_s f \rangle.$$

This is a nonreversible analog of bounding the generator by the Dirichlet form when jump rates are bounded.

Mixing-time lower bound

Define time- t total-variation distance to equilibrium and, respectively, ε -mixing time:

$$d(t) = \sup_{\nu \in \mathcal{P}(S)} \|\nu P_t - \mu\|_{\text{TV}} \quad t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

Lower bound on total-variation mixing time

If $d(t) \leq \varepsilon < 1/2$, then

$$\gamma \geq \frac{1 - \sqrt{2\varepsilon}}{t}.$$

Equivalently,

$$t_{\text{mix}}(\varepsilon) \geq \frac{1 - \sqrt{2\varepsilon}}{\gamma}.$$

Proof idea:

$$f - P_t f = \int_0^t P_s(-Lf) ds \implies \|f - P_t f\| \leq t \|Lf\|.$$

Small total variation gives

$$\|P_t f\| \leq \sqrt{2\varepsilon} \|f\|.$$

The same argument works in discrete time, arbitrary state space.

Mixing-time lower bound

Define time- t total-variation distance to equilibrium and, respectively, ε -mixing time:

$$d(t) = \sup_{\nu \in \mathcal{P}(S)} \|\nu P_t - \mu\|_{\text{TV}} \quad t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

Lower bound on total-variation mixing time

If $d(t) \leq \varepsilon < 1/2$, then

$$\gamma \geq \frac{1 - \sqrt{2\varepsilon}}{t}.$$

Equivalently,

$$t_{\text{mix}}(\varepsilon) \geq \frac{1 - \sqrt{2\varepsilon}}{\gamma}.$$

Proof idea:

$$f - P_t f = \int_0^t P_s(-Lf) ds \implies \|f - P_t f\| \leq t \|Lf\|.$$

Small total variation gives

$$\|P_t f\| \leq \sqrt{2\varepsilon} \|f\|.$$

The same argument works in discrete time, arbitrary state space.

Mixing-time lower bound

Define time- t total-variation distance to equilibrium and, respectively, ε -mixing time:

$$d(t) = \sup_{\nu \in \mathcal{P}(S)} \|\nu P_t - \mu\|_{\text{TV}} \quad t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

Lower bound on total-variation mixing time

If $d(t) \leq \varepsilon < 1/2$, then

$$\gamma \geq \frac{1 - \sqrt{2\varepsilon}}{t}.$$

Equivalently,

$$t_{\text{mix}}(\varepsilon) \geq \frac{1 - \sqrt{2\varepsilon}}{\gamma}.$$

Proof idea:

$$f - P_t f = \int_0^t P_s(-Lf) ds \implies \|f - P_t f\| \leq t \|Lf\|.$$

Small total variation gives

$$\|P_t f\| \leq \sqrt{2\varepsilon} \|f\|.$$

The same argument works in discrete time, arbitrary state space.

Mixing-time lower bound

Define time- t total-variation distance to equilibrium and, respectively, ε -mixing time:

$$d(t) = \sup_{\nu \in \mathcal{P}(S)} \|\nu P_t - \mu\|_{\text{TV}} \quad t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

Lower bound on total-variation mixing time

If $d(t) \leq \varepsilon < 1/2$, then

$$\gamma \geq \frac{1 - \sqrt{2\varepsilon}}{t}.$$

Equivalently,

$$t_{\text{mix}}(\varepsilon) \geq \frac{1 - \sqrt{2\varepsilon}}{\gamma}.$$

Proof idea:

$$f - P_t f = \int_0^t P_s(-Lf) ds \implies \|f - P_t f\| \leq t \|Lf\|.$$

Small total variation gives

$$\|P_t f\| \leq \sqrt{2\varepsilon} \|f\|.$$

The same argument works in discrete time, arbitrary state space.

Mixing-time lower bound

Define time- t total-variation distance to equilibrium and, respectively, ε -mixing time:

$$d(t) = \sup_{\nu \in \mathcal{P}(S)} \|\nu P_t - \mu\|_{\text{TV}} \quad t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 : d(t) \leq \varepsilon\}.$$

Lower bound on total-variation mixing time

If $d(t) \leq \varepsilon < 1/2$, then

$$\gamma \geq \frac{1 - \sqrt{2\varepsilon}}{t}.$$

Equivalently,

$$t_{\text{mix}}(\varepsilon) \geq \frac{1 - \sqrt{2\varepsilon}}{\gamma}.$$

Proof idea:

$$f - P_t f = \int_0^t P_s(-Lf) ds \implies \|f - P_t f\| \leq t \|Lf\|.$$

Small total variation gives

$$\|P_t f\| \leq \sqrt{2\varepsilon} \|f\|.$$

The same argument works in discrete time, arbitrary state space.

Mixing-time upper bound in finite state space

For finite-state bounded-rate jump processes,

$$\mu_{\min} := \min_x \mu(x) > 0.$$

Using

$$\gamma_s \geq \frac{\gamma^2}{2q_{\max}}$$

and L^2 -to-total-variation comparison,

$$\|P_t(x, \cdot) - \mu\|_{\text{TV}} \leq \frac{1}{2\sqrt{\mu_{\min}}} e^{-\gamma_s t}.$$

Proposition 5

$$t_{\text{mix}}(\varepsilon) \leq \frac{2q_{\max}}{\gamma^2} \log \left(\frac{1}{2\varepsilon\sqrt{\mu_{\min}}} \right).$$

The finite-state assumption is essential; on countable spaces, $\mu_{\min} = 0$ and worst-case TV mixing time could be infinite even with a positive reversible spectral gap.

Mixing-time upper bound in finite state space

For finite-state bounded-rate jump processes,

$$\mu_{\min} := \min_x \mu(x) > 0.$$

Using

$$\gamma_s \geq \frac{\gamma^2}{2q_{\max}}$$

and L^2 -to-total-variation comparison,

$$\|P_t(x, \cdot) - \mu\|_{\text{TV}} \leq \frac{1}{2\sqrt{\mu_{\min}}} e^{-\gamma_s t}.$$

Proposition 5

$$t_{\text{mix}}(\varepsilon) \leq \frac{2q_{\max}}{\gamma^2} \log \left(\frac{1}{2\varepsilon\sqrt{\mu_{\min}}} \right).$$

The finite-state assumption is essential; on countable spaces, $\mu_{\min} = 0$ and worst-case TV mixing time could be infinite even with a positive reversible spectral gap.

Mixing-time upper bound in finite state space

For finite-state bounded-rate jump processes,

$$\mu_{\min} := \min_x \mu(x) > 0.$$

Using

$$\gamma_s \geq \frac{\gamma^2}{2q_{\max}}$$

and L^2 -to-total-variation comparison,

$$\|P_t(x, \cdot) - \mu\|_{\text{TV}} \leq \frac{1}{2\sqrt{\mu_{\min}}} e^{-\gamma_s t}.$$

Proposition 5

$$t_{\text{mix}}(\varepsilon) \leq \frac{2q_{\max}}{\gamma^2} \log \left(\frac{1}{2\varepsilon\sqrt{\mu_{\min}}} \right).$$

The finite-state assumption is essential; on countable spaces, $\mu_{\min} = 0$ and worst-case TV mixing time could be infinite even with a positive reversible spectral gap.

Mixing-time upper bound in finite state space

For finite-state bounded-rate jump processes,

$$\mu_{\min} := \min_x \mu(x) > 0.$$

Using

$$\gamma_s \geq \frac{\gamma^2}{2q_{\max}}$$

and L^2 -to-total-variation comparison,

$$\|P_t(x, \cdot) - \mu\|_{\text{TV}} \leq \frac{1}{2\sqrt{\mu_{\min}}} e^{-\gamma_s t}.$$

Proposition 5

$$t_{\text{mix}}(\varepsilon) \leq \frac{2q_{\max}}{\gamma^2} \log \left(\frac{1}{2\varepsilon\sqrt{\mu_{\min}}} \right).$$

The finite-state assumption is essential; on countable spaces, $\mu_{\min} = 0$ and worst-case TV mixing time could be infinite even with a positive reversible spectral gap.

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_* and $\xi_* = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$y_A = \frac{1_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_* and $\xi_* = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$y_A = \frac{1_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_* and $\xi_* = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$y_A = \frac{1_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_s and $\xi_s = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$g_A = \frac{1_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_s and $\xi_s = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$g_A = \frac{1_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Comparison with the Cheeger constant

For a jump process with jump kernel q , the stationary jump flux is

$$Q(dx, dy) = \mu(dx)q(x, dy).$$

The Cheeger constant is defined as

$$\xi := \inf_{0 < \mu(A) \leq 1/2} \frac{Q(A, A^c)}{\mu(A)}.$$

Theorem 5: nonreversible Cheeger bounds

For a bounded-rate jump process on a standard Borel spaces,

$$\frac{\xi^2}{2q_{\max}} \leq \gamma \leq \frac{36}{5}\xi.$$

- Lower bound: apply reversible Cheeger inequality to γ_s and $\xi_s = \xi$.
- Key identity: stationarity implies flow balance across cuts,

$$Q(A, A^c) = Q(A^c, A).$$

- Upper bound: use empirical variance of a bottleneck observable

$$g_A = \frac{\mathbf{1}_A - \mu(A)}{\sqrt{\mu(A)(1 - \mu(A))}}.$$

Two applications

Shift-register chain on the binary de Bruijn graph

State space:

$$S_n = \{-1, 1\}^n,$$

with uniform measure.

Transition:

$$(x_1, \dots, x_n) \mapsto (x_2, \dots, x_n, \xi), \quad \xi \sim \text{Unif}\{-1, 1\}.$$

After $k < n$ steps,

$$X_k = (x_{k+1}, \dots, x_n, \xi_1, \dots, \xi_k).$$

The chain has forgotten only k coordinates.

Thus

$$\|P^k(x, \cdot) - \mu\|_{\text{TV}} = 1 - 2^{k-n}, \quad k < n,$$

and

$$P^n(x, \cdot) = \mu.$$

So $t_{\text{mix}}(\varepsilon) = n$ for fixed $\varepsilon < 1/2$.

Shift-register chain on the binary de Bruijn graph

State space:

$$S_n = \{-1, 1\}^n,$$

with uniform measure.

Transition:

$$(x_1, \dots, x_n) \mapsto (x_2, \dots, x_n, \xi), \quad \xi \sim \text{Unif}\{-1, 1\}.$$

After $k < n$ steps,

$$X_k = (x_{k+1}, \dots, x_n, \xi_1, \dots, \xi_k).$$

The chain has forgotten only k coordinates.

Thus

$$\|P^k(x, \cdot) - \mu\|_{\text{TV}} = 1 - 2^{k-n}, \quad k < n,$$

and

$$P^n(x, \cdot) = \mu.$$

So $t_{\text{mix}}(\varepsilon) = n$ for fixed $\varepsilon < 1/2$.

Shift-register chain on the binary de Bruijn graph

State space:

$$S_n = \{-1, 1\}^n,$$

with uniform measure.

Transition:

$$(x_1, \dots, x_n) \mapsto (x_2, \dots, x_n, \xi), \quad \xi \sim \text{Unif}\{-1, 1\}.$$

After $k < n$ steps,

$$X_k = (x_{k+1}, \dots, x_n, \xi_1, \dots, \xi_k).$$

The chain has forgotten only k coordinates.

Thus

$$\left\| P^k(x, \cdot) - \mu \right\|_{\text{TV}} = 1 - 2^{k-n}, \quad k < n,$$

and

$$P^n(x, \cdot) = \mu.$$

So $t_{\text{mix}}(\varepsilon) = n$ for fixed $\varepsilon < 1/2$.

Decomposition into nilpotent shift chains

Walsh characters:

$$\chi_A(x) = \prod_{i \in A} x_i, \quad A \subseteq [n].$$

They form an orthonormal basis of $L^2(\{-1, 1\}^n)$.

The transition operator satisfies

$$P\chi_A = \begin{cases} \chi_{A+1}, & n \notin A, \\ 0, & n \in A, \end{cases}$$

where

$$A + 1 = \{i + 1 : i \in A, i < n\}.$$

So P decomposes L_0^2 into orthogonal nilpotent shift chains.

Longest chain:

$$\chi_{\{1\}} \mapsto \chi_{\{2\}} \mapsto \cdots \mapsto \chi_{\{n\}} \mapsto 0.$$

Decomposition into nilpotent shift chains

Walsh characters:

$$\chi_A(x) = \prod_{i \in A} x_i, \quad A \subseteq [n].$$

They form an orthonormal basis of $L^2(\{-1, 1\}^n)$.

The transition operator satisfies

$$P\chi_A = \begin{cases} \chi_{A+1}, & n \notin A, \\ 0, & n \in A, \end{cases}$$

where

$$A + 1 = \{i + 1 : i \in A, i < n\}.$$

So P decomposes L_0^2 into orthogonal nilpotent shift chains.

Longest chain:

$$\chi_{\{1\}} \mapsto \chi_{\{2\}} \mapsto \cdots \mapsto \chi_{\{n\}} \mapsto 0.$$

Decomposition into nilpotent shift chains

Walsh characters:

$$\chi_A(x) = \prod_{i \in A} x_i, \quad A \subseteq [n].$$

They form an orthonormal basis of $L^2(\{-1, 1\}^n)$.

The transition operator satisfies

$$P\chi_A = \begin{cases} \chi_{A+1}, & n \notin A, \\ 0, & n \in A, \end{cases}$$

where

$$A + 1 = \{i + 1 : i \in A, i < n\}.$$

So P decomposes L_0^2 into orthogonal nilpotent shift chains.

Longest chain:

$$\chi_{\{1\}} \mapsto \chi_{\{2\}} \mapsto \cdots \mapsto \chi_{\{n\}} \mapsto 0.$$

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Singular-value gap correctly captures nonnormal memory

On a shift chain of length ℓ ,

$$P = S_\ell, \quad S_\ell e_1 = e_2, \dots, S_\ell e_\ell = 0.$$

For $L = P - I$,

$$\gamma = s_{\min}(I - S_n) = 2 \sin \left(\frac{\pi}{4n+2} \right) \asymp \frac{1}{n}.$$

Mixing time lower bound gives

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{1}{\gamma} \asymp n,$$

the correct order.

Different spectral quantities in this example

- Real-part spectral gap/Distance to spectrum: order 1, too optimistic.
- Symmetrized gap: order n^{-2} , too pessimistic.
- Singular-value gap: order n^{-1} , correct for mixing lower bound.

Here, the lower bound by γ^{-1} is of correct order and captures deterministic memory, but the mixing time in general could be larger.

Compact-resolvent criterion

In some situations, we can show the singular-value gap to be positive without explicitly understanding the spectrum of the generator.

Sufficient condition for a positive γ

Let L be a Markov generator on $L^2(\mu)$ with invariant probability measure μ . Suppose:

1. $\text{Ker } L$ consists only of constants;
2. L has compact resolvent.

Then

$$\gamma = \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|} > 0.$$

Proof idea:

- Compact resolvent implies spectrum is discrete.
- If 0 were in the spectrum of L restricted on $L_0^2(\mu)$, it would be an eigenvalue.
- Therefore, 0 is in the resolvent set and L^{-1} exists on $L_0^2(\mu)$.

Compact-resolvent criterion

In some situations, we can show the singular-value gap to be positive without explicitly understanding the spectrum of the generator.

Sufficient condition for a positive γ

Let L be a Markov generator on $L^2(\mu)$ with invariant probability measure μ . Suppose:

1. $\text{Ker } L$ consists only of constants;
2. L has compact resolvent.

Then

$$\gamma = \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|} > 0.$$

Proof idea:

- Compact resolvent implies spectrum is discrete.
- If 0 were in the spectrum of L restricted on $L_0^2(\mu)$, it would be an eigenvalue.
- Therefore, 0 is in the resolvent set and L^{-1} exists on $L_0^2(\mu)$.

Compact-resolvent criterion

In some situations, we can show the singular-value gap to be positive without explicitly understanding the spectrum of the generator.

Sufficient condition for a positive γ

Let L be a Markov generator on $L^2(\mu)$ with invariant probability measure μ . Suppose:

1. $\text{Ker } L$ consists only of constants;
2. L has compact resolvent.

Then

$$\gamma = \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|} > 0.$$

Proof idea:

- Compact resolvent implies spectrum is discrete.
- If 0 were in the spectrum of L restricted on $L_0^2(\mu)$, it would be an eigenvalue.
- Therefore, 0 is in the resolvent set and L^{-1} exists on $L_0^2(\mu)$.

Compact-resolvent criterion

In some situations, we can show the singular-value gap to be positive without explicitly understanding the spectrum of the generator.

Sufficient condition for a positive γ

Let L be a Markov generator on $L^2(\mu)$ with invariant probability measure μ . Suppose:

1. $\text{Ker } L$ consists only of constants;
2. L has compact resolvent.

Then

$$\gamma = \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|} > 0.$$

Proof idea:

- Compact resolvent implies spectrum is discrete.
- If 0 were in the spectrum of L restricted on $L_0^2(\mu)$, it would be an eigenvalue.
- Therefore, 0 is in the resolvent set and L^{-1} exists on $L_0^2(\mu)$.

Compact-resolvent criterion

In some situations, we can show the singular-value gap to be positive without explicitly understanding the spectrum of the generator.

Sufficient condition for a positive γ

Let L be a Markov generator on $L^2(\mu)$ with invariant probability measure μ . Suppose:

1. $\text{Ker } L$ consists only of constants;
2. L has compact resolvent.

Then

$$\gamma = \inf_{\substack{f \in \text{Dom}(L) \cap L_0^2(\mu) \\ f \neq 0}} \frac{\|Lf\|}{\|f\|} > 0.$$

Proof idea:

- Compact resolvent implies spectrum is discrete.
- If 0 were in the spectrum of L restricted on $L_0^2(\mu)$, it would be an eigenvalue.
- Therefore, 0 is in the resolvent set and L^{-1} exists on $L_0^2(\mu)$.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Hypoelliptic shear on the torus

Consider

$$L = \partial_y^2 + \sin(y)\partial_x$$

on $\mathbb{T}^2$ with normalized Lebesgue measure.

The symmetric part ∂_y^2 has no gap on $L_0^2(\mathbb{T}^2)$: functions of x alone are not dissipated.

Hörmander's bracket condition helps verify L is hypoelliptic, so we have

1. $\text{Ker } L$ consists only of constants.

We can also check compact embedding $(\text{Dom}(L), \|\cdot\| + \|L\cdot\|) \hookrightarrow L^2(\mathbb{T}^2)$, so

2. L has compact resolvent

By the compact-resolvent criterion,

$$\gamma > 0.$$

Therefore

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2, \quad g \in L_0^2(\mathbb{T}^2).$$

This converts qualitative hypoelliptic regularization into a concrete stationary empirical-variance bound. The drift coupling is enough to “traverse” the torus.

Past and future directions

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Where this fits

- **Chatterjee'25**: finite-state singular-value gap for nonreversible chains: control of empirical averages and connections to other mixing notions.
- **Dinwoodie'98, Huang-Li'25**: time-average deviations and Bernstein inequalities using singular-value gaps.
- **Fill'91, Paulin'18**: transition-operator singular values, multiplicative reversibilization, pseudo-spectral gaps.
- **Bhattacharya'82, Kipnis-Varadhan'86**: Poisson equations, martingale approximation, asymptotic variance, CLT under weaker assumptions.
- **Diaconis-Holmes-Neal'00, HHS'93'05, DLP'16**: accelerated convergence and variance reduction by irreversibility, with applications to sampling in mind.

This work's features

- most general setting
- stationary empirical averages, considered in $L^2(\mu)$
- uniform control of finite-time averages of observable classes

Main takeaways

1. Positive singular-value gap means the generator is invertible on $L_0^2(\mu)$

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

2. This gives robust finite-time empirical-variance bounds:

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

3. Local gaps give uniform bounds on classes of observables (e.g. Rotating OU).
4. γ^{-1} gives a uniform upper bound on asymptotic variance. When the asymptotic variance is nonzero for an observable g , it governs long-time behavior of $\mu_t g$:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{\sigma_g^2}{t} + O(t^{-2}).$$

5. The singular-value gap interacts meaningfully with symmetrization, mixing time, the Cheeger constant, and potentially other mixing quantities.

Main takeaways

1. Positive singular-value gap means the generator is invertible on $L_0^2(\mu)$

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

2. This gives robust finite-time empirical-variance bounds:

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

3. Local gaps give uniform bounds on classes of observables (e.g. Rotating OU).
4. γ^{-1} gives a uniform upper bound on asymptotic variance. When the asymptotic variance is nonzero for an observable g , it governs long-time behavior of $\mu_t g$:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{\sigma_g^2}{t} + O(t^{-2}).$$

5. The singular-value gap interacts meaningfully with symmetrization, mixing time, the Cheeger constant, and potentially other mixing quantities.

Main takeaways

1. Positive singular-value gap means the generator is invertible on $L_0^2(\mu)$

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

2. This gives robust finite-time empirical-variance bounds:

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

3. Local gaps give uniform bounds on classes of observables (e.g. Rotating OU).
4. γ^{-1} gives a uniform upper bound on asymptotic variance. When the asymptotic variance is nonzero for an observable g , it governs long-time behavior of $\mu_t g$:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{\sigma_g^2}{t} + O(t^{-2}).$$

5. The singular-value gap interacts meaningfully with symmetrization, mixing time, the Cheeger constant, and potentially other mixing quantities.

Main takeaways

1. Positive singular-value gap means the generator is invertible on $L_0^2(\mu)$

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

2. This gives robust finite-time empirical-variance bounds:

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

3. Local gaps give uniform bounds on classes of observables (e.g. Rotating OU).
4. γ^{-1} gives a uniform upper bound on asymptotic variance. When the asymptotic variance is nonzero for an observable g , it governs long-time behavior of $\mu_t g$:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{\sigma_g^2}{t} + O(t^{-2}).$$

5. The singular-value gap interacts meaningfully with symmetrization, mixing time, the Cheeger constant, and potentially other mixing quantities.

Main takeaways

1. Positive singular-value gap means the generator is invertible on $L_0^2(\mu)$

$$\| -L^{-1} \| \leq \gamma^{-1}.$$

2. This gives robust finite-time empirical-variance bounds:

$$\mathbb{E}_\mu |\mu_t g|^2 \leq \frac{4}{\gamma t} \|g\|^2.$$

3. Local gaps give uniform bounds on classes of observables (e.g. Rotating OU).
4. γ^{-1} gives a uniform upper bound on asymptotic variance. When the asymptotic variance is nonzero for an observable g , it governs long-time behavior of $\mu_t g$:

$$\mathbb{E}_\mu |\mu_t g|^2 = \frac{\sigma_g^2}{t} + O(t^{-2}).$$

5. The singular-value gap interacts meaningfully with symmetrization, mixing time, the Cheeger constant, and potentially other mixing quantities.

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Future directions

Directly related to current results:

- Other mechanisms for positive γ . Quantitative lower bounds for γ when the spectrum of the generator is not easily computable.
- Concentration estimates, or bounds with approximately stationary starts.
- Methods to identify subspaces with positive local gap. Optimization of irreversible perturbations for observable classes rather than single observables.
- Better understanding of when $\sigma_K^2 \asymp \gamma_K^{-1}$ and when oscillations force $\sigma_K^2 \ll \gamma_K^{-1}$. Similarly, understand the separation between the $t^{-1/2}$ and t^{-1} regimes for Δ_t . Note that the oscillatory examples we have seem to be edge cases.

Extensions:

- Relation to Kipnis-Varadhan-style CLT? What happens when finite-dimensional systems have singular-value gaps that tend to 0 in the thermodynamic limit?
- Extension to open quantum system with Lindbladian dynamics?

Thank you!