

Central Limit Theorem for ergodic averages of Markov chains & the comparison of sampling algorithms for heavy-tailed distributions

Aleksandar Mijatović
Department of Statistics



Joint work with Miha Brešar and Gareth Roberts

11 December 2025

Part I: Markov chains with heavy-tailed invariant distributions and applications to ULA-type samplers

EPSRC

Engineering and Physical Sciences
Research Council

EP/W006227/1 & EP/V009478/1

Motivation: sampling from heavy-tailed targets on $\mathbb{R}^d$

Let density $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for large $|x|$ and some $\nu \in (0, \infty)$

Objective: **estimate** $\pi(g) = \int_{\mathbb{R}^d} g(x)\pi(x)dx$ for $g : \mathbb{R}^d \rightarrow \mathbb{R} \in L^1(\pi)$

Motivation: sampling from heavy-tailed targets on $\mathbb{R}^d$

Let density $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for large $|x|$ and some $\nu \in (0, \infty)$

Objective: **estimate** $\pi(g) = \int_{\mathbb{R}^d} g(x)\pi(x)dx$ for $g : \mathbb{R}^d \rightarrow \mathbb{R} \in L^1(\pi)$

This problem appears in **practical** applications [Gel+08; GLM18]
including logistic regression with t -distribution or Cauchy priors.

Motivation: sampling from heavy-tailed targets on $\mathbb{R}^d$

Let density $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for large $|x|$ and some $\nu \in (0, \infty)$

Objective: **estimate** $\pi(g) = \int_{\mathbb{R}^d} g(x)\pi(x)dx$ for $g : \mathbb{R}^d \rightarrow \mathbb{R} \in L^1(\pi)$

This problem appears in **practical** applications [Gel+08; GLM18] including logistic regression with t -distribution or Cauchy priors.

Run **Sampling algorithm** $X = (X_n)_{n \in \mathbb{N}}$ (i.e. **Markov chain with invariant law π**):

$$\text{estimator} \quad S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \longrightarrow \pi(g) \quad \text{converges a.s. as } N \uparrow \infty.$$

Motivation: sampling from heavy-tailed targets on $\mathbb{R}^d$

Let density $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for large $|x|$ and some $\nu \in (0, \infty)$

Objective: estimate $\pi(g) = \int_{\mathbb{R}^d} g(x)\pi(x)dx$ for $g : \mathbb{R}^d \rightarrow \mathbb{R} \in L^1(\pi)$

This problem appears in **practical** applications [Gel+08; GLM18] including logistic regression with t -distribution or Cauchy priors.

Run **Sampling algorithm** $X = (X_n)_{n \in \mathbb{N}}$ (i.e. **Markov chain with invariant law π**):

$$\text{estimator} \quad S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \longrightarrow \pi(g) \quad \text{converges a.s. as } N \uparrow \infty.$$

What is the **confidence** interval for $S_N(g)$ when $g \in L^2(\pi)$?

How fast does the law of X_n **converge** to π as $n \rightarrow \infty$?

Trouble starts in dimension $d = 1$

Let $\pi(x) \approx (1 + |x|)^{-(1+\nu)}$ for large $|x|$ and some $\nu > 0$

$$S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \quad \text{for } g(x) = \mathbb{1}\{x \geq 2\}$$

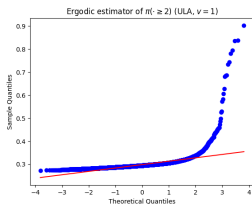
QQ-plots ($\sqrt{N}(S_N(g) - \pi(g))$ vs. normal) for $S_N(g)$ estimating the probability $\pi(\cdot \geq 2)$:

Trouble starts in dimension $d = 1$

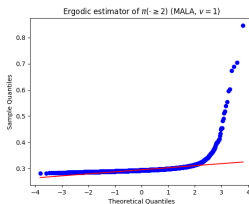
Let $\pi(x) \approx (1 + |x|)^{-(1+\nu)}$ for large $|x|$ and some $\nu > 0$

$$S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \quad \text{for } g(x) = \mathbb{1}\{x \geq 2\}$$

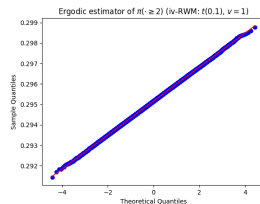
QQ-plots ($\sqrt{N}(S_N(g) - \pi(g))$ vs. normal) for $S_N(g)$ estimating the probability $\pi(\cdot \geq 2)$:



(a) ULA



(b) MALA



(c) iv-RWM

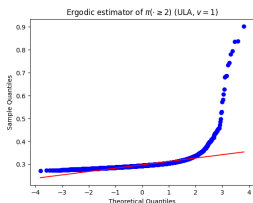
CLT for $S_N(g)$ **fails** in (a) and (b), making sampling from heavy-tailed targets notoriously **difficult** ...

Trouble starts in dimension $d = 1$

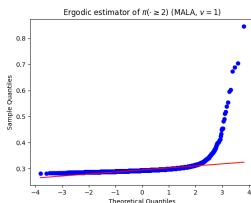
Let $\pi(x) \approx (1 + |x|)^{-(1+\nu)}$ for large $|x|$ and some $\nu > 0$

$$S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \quad \text{for } g(x) = \mathbb{1}\{x \geq 2\}$$

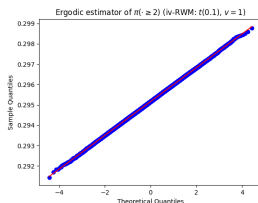
QQ-plots ($\sqrt{N}(S_N(g) - \pi(g))$ vs. normal) for $S_N(g)$ estimating the probability $\pi(\cdot \geq 2)$:



(a) ULA



(b) MALA



(c) iv-RWM

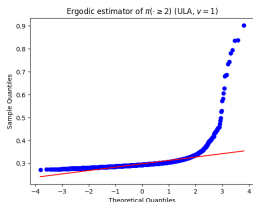
Empirical evidence for the failure of the CLT dates back to [FM03].
However, **necessary** conditions for CLT of $S_N(g)$ are largely **unavailable**.

Trouble starts in dimension $d = 1$

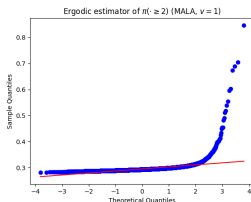
Let $\pi(x) \approx (1 + |x|)^{-(1+\nu)}$ for large $|x|$ and some $\nu > 0$

$$S_N(g) = \frac{1}{N} \sum_{k=1}^N g(X_k) \quad \text{for } g(x) = \mathbb{1}\{x \geq 2\}$$

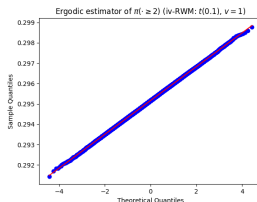
QQ-plots ($\sqrt{N}(S_N(g) - \pi(g))$ vs. normal) for $S_N(g)$ estimating the probability $\pi(\cdot \geq 2)$:



(a) ULA



(b) MALA



(c) iv-RWM

Empirical evidence for the failure of the CLT dates back to [FM03]. However, **necessary** conditions for CLT of $S_N(g)$ are largely **unavailable**. This talk presents a **novel** approach: it provides **characterisation** of CLT, **enabling algorithm comparison** and **guiding algorithm selection**.

Known results

Sufficient conditions for CLT and upper bounds on convergence rates are well-known [KV86; MT93; DGM08; Dou+04; Dou+18].

CLTs, the corresponding asymptotic variance and ergodicity remains extensively studied [Dou+25; OM24; DL18].

Sufficient conditions for CLT and Upper bounds: summary

Let X be irreducible Markov chain (sampling algorithm) with transition kernel P on a locally compact space $\mathcal{X}$ (in our applications $\mathcal{X} = \mathbb{R}^d$).

Sufficient conditions for CLT and Upper bounds: summary

Let X be irreducible Markov chain (sampling algorithm) with transition kernel P on a locally compact space $\mathcal{X}$ (in our applications $\mathcal{X} = \mathbb{R}^d$).

Let $V : \mathcal{X} \rightarrow [1, \infty)$ continuous, $\phi : [1, \infty) \rightarrow [1, \infty)$ increasing, concave, $A \in \mathcal{B}(\mathcal{X})$ small (compact) set and $b \in \mathbb{R}_+$, such that

$$PV - V \leq -\phi \circ V, \quad \text{on the complement of some compact set.}$$

Sufficient conditions for CLT and Upper bounds: summary

Let X be irreducible Markov chain (sampling algorithm) with transition kernel P on a locally compact space $\mathcal{X}$ (in our applications $\mathcal{X} = \mathbb{R}^d$).

Let $V : \mathcal{X} \rightarrow [1, \infty)$ continuous, $\phi : [1, \infty) \rightarrow [1, \infty)$ increasing, concave, $A \in \mathcal{B}(\mathcal{X})$ small (compact) set and $b \in \mathbb{R}_+$, such that

$$PV - V \leq -\phi \circ V, \quad \text{on the complement of some compact set.}$$

By [Dou+18; Dou+04; Jon04; DGM08; MT93]:

$$(a) \quad \|\mathbb{P}_x(X_n \in \cdot) - \pi(\cdot)\|_{\text{TV}} \leq \frac{CV(x)}{\phi \circ H_\phi^{-1}(n)}, \text{ where } H_\phi(u) = \int_1^u \frac{ds}{\phi(s)}$$

Sufficient conditions for CLT and Upper bounds: summary

Let X be irreducible Markov chain (sampling algorithm) with transition kernel P on a locally compact space $\mathcal{X}$ (in our applications $\mathcal{X} = \mathbb{R}^d$).

Let $V : \mathcal{X} \rightarrow [1, \infty)$ continuous, $\phi : [1, \infty) \rightarrow [1, \infty)$ increasing, concave, $A \in \mathcal{B}(\mathcal{X})$ small (compact) set and $b \in \mathbb{R}_+$, such that

$$PV - V \leq -\phi \circ V, \quad \text{on the complement of some compact set.}$$

By [Dou+18; Dou+04; Jon04; DGM08; MT93]:

- (a) $\|\mathbb{P}_x(X_n \in \cdot) - \pi(\cdot)\|_{\text{TV}} \leq \frac{C V(x)}{\phi \circ H_\phi^{-1}(n)}$, where $H_\phi(u) = \int_1^u \frac{ds}{\phi(s)}$
- (b) If $\pi(V \cdot \phi \circ V) < \infty$ and $|g| \leq \phi \circ V$ then $\exists \sigma_g^2 < \infty$ such that

$$\sqrt{N}(S_N(g) - \pi(g)) = \frac{1}{\sqrt{N}} \sum_{k=1}^N (g(S_k) - \pi(g)) \xrightarrow{d} N(0, \sigma_g^2), \quad \text{as } N \rightarrow \infty.$$

Sufficient conditions for CLT and Upper bounds: summary

Let X be irreducible Markov chain (sampling algorithm) with transition kernel P on a locally compact space $\mathcal{X}$ (in our applications $\mathcal{X} = \mathbb{R}^d$).

Let $V : \mathcal{X} \rightarrow [1, \infty)$ continuous, $\phi : [1, \infty) \rightarrow [1, \infty)$ increasing, concave, $A \in \mathcal{B}(\mathcal{X})$ small (compact) set and $b \in \mathbb{R}_+$, such that

$$PV - V \leq -\phi \circ V, \quad \text{on the complement of some compact set.}$$

By [Dou+18; Dou+04; Jon04; DGM08; MT93]:

- (a) $\|\mathbb{P}_x(X_n \in \cdot) - \pi(\cdot)\|_{\text{TV}} \leq \frac{CV(x)}{\phi \circ H_\phi^{-1}(n)}$, where $H_\phi(u) = \int_1^u \frac{ds}{\phi(s)}$
- (b) If $\pi(V \cdot \phi \circ V) < \infty$ and $|g| \leq \phi \circ V$ then $\exists \sigma_g^2 < \infty$ such that

$$\sqrt{N}(S_N(g) - \pi(g)) = \frac{1}{\sqrt{N}} \sum_{k=1}^N (g(S_k) - \pi(g)) \xrightarrow{d} N(0, \sigma_g^2), \quad \text{as } N \rightarrow \infty.$$

How about the necessary conditions and lower bounds?

L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ (following [BM24])

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ (following [BM24])

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

(i) $\varphi : (0, 1] \rightarrow \mathbb{R}_+$ continuous, $r\varphi(1/r) \downarrow 0$ as $r \uparrow \infty$ and

$$P(1/V)(x) - 1/V(x) \leq \varphi(1/V)(x) \quad \text{for } |x| \text{ large.}$$

L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ (following [BM24])

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

(i) $\varphi : (0, 1] \rightarrow \mathbb{R}_+$ continuous, $r\varphi(1/r) \downarrow 0$ as $r \uparrow \infty$ and

$$P(1/V)(x) - 1/V(x) \leq \varphi(1/V)(x) \quad \text{for } |x| \text{ large.}$$

(ii) For $\Psi : [1, \infty) \rightarrow [1, \infty)$ submultip., any $\ell > \ell_0$, $\exists C_\ell > 0$ s.t.

$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell / \Psi(r)$ for all $\ell + 1 < V(x) < r < \infty$ where

$T^{(r)} := \inf\{t \geq 0 : V(X_t) > r\}$, $S_{(\ell)} := \inf\{t \geq 0 : V(X_t) < \ell\}$.

L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

(i) $\varphi : (0, 1] \rightarrow \mathbb{R}_+$ continuous, $r\varphi(1/r) \downarrow 0$ as $r \uparrow \infty$ and

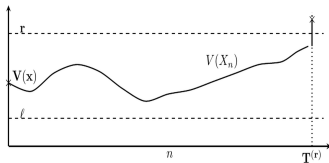
$$P(1/V)(x) - 1/V(x) \leq \varphi(1/V)(x) \quad \text{for } |x| \text{ large}$$

(ii) For $\Psi : [1, \infty) \rightarrow [1, \infty)$ submultip., any $\ell > \ell_0$, $\exists C_\ell > 0$ s.t.

$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell / \Psi(r)$ for all $\ell + 1 < V(x) < r < \infty$ where

$T^{(r)} := \inf\{n \in \mathbb{N} : V(X_n) > r\}$ & $S_{(\ell)} := \inf\{n \in \mathbb{N} : V(X_n) < \ell\}$

Intuition: • Ψ quantifies the **tendency** for X to explore the tails:



L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

(i) $\varphi : (0, 1] \rightarrow \mathbb{R}_+$ continuous, $r\varphi(1/r) \downarrow 0$ as $r \uparrow \infty$ and

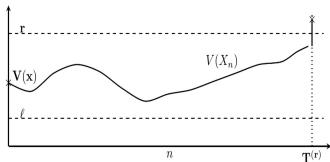
$$P(1/V)(x) - 1/V(x) \leq \varphi(1/V)(x) \quad \text{for } |x| \text{ large}$$

(ii) For $\Psi : [1, \infty) \rightarrow [1, \infty)$ submultip., any $\ell > \ell_0$, $\exists C_\ell > 0$ s.t.

$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell / \Psi(r)$ for all $\ell + 1 < V(x) < r < \infty$ where

$T^{(r)} := \inf\{n \in \mathbb{N} : V(X_n) > r\}$ & $S_{(\ell)} := \inf\{n \in \mathbb{N} : V(X_n) < \ell\}$

Intuition: • Ψ quantifies the **tendency** for X to explore the tails:



• in practice (ii) implied by *drift cond* $P(\Psi \circ V)(x) \geq \Psi \circ V(x)$ for $|x|$ big

L-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$

$V : \mathcal{X} \rightarrow [1, \infty)$ cts, $\limsup_{n \rightarrow \infty} V(X_n) = \infty$ and $b, \ell_0 \in [0, \infty)$

(i) $\varphi : (0, 1] \rightarrow \mathbb{R}_+$ continuous, $r\varphi(1/r) \downarrow 0$ as $r \uparrow \infty$ and

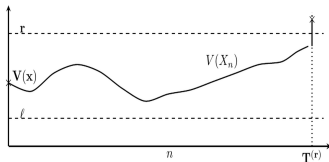
$$P(1/V)(x) - 1/V(x) \leq \varphi(1/V)(x) \quad \text{for } |x| \text{ large}$$

(ii) For $\Psi : [1, \infty) \rightarrow [1, \infty)$ submultip., any $\ell > \ell_0$, $\exists C_\ell > 0$ s.t.

$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell / \Psi(r)$ for all $\ell + 1 < V(x) < r < \infty$ where

$T^{(r)} := \inf\{n \in \mathbb{N} : V(X_n) > r\}$ & $S_{(\ell)} := \inf\{n \in \mathbb{N} : V(X_n) < \ell\}$

Intuition: • Ψ quantifies the **tendency** for X to explore the tails:



- in practice (ii) implied by *drift cond* $P(\Psi \circ V)(x) \geq \Psi \circ V(x)$ for $|x|$ big
- φ quantifies the **stickiness** of X in the tail and **duration** of its return

Necessary conditions for the CLT of an ergodic average

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold and recall $S_N(g) = \sum_{k=1}^N g(X_k)/N$. For $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ non-decreasing at infinity and any $q \in (0, 1)$, define

$$G_h(r) := q(1 - q) \frac{h(r)}{r\varphi(1/r)}, \quad r \in [1, \infty),$$

Necessary conditions for the CLT of an ergodic average

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold and recall $S_N(g) = \sum_{k=1}^N g(X_k)/N$. For $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ non-decreasing at infinity and any $q \in (0, 1)$, define

$$G_h(r) := q(1 - q) \frac{h(r)}{r\varphi(1/r)}, \quad r \in [1, \infty),$$

- ① Let G_h^{-1} be the inverse of G_h on a neighbourhood of infinity. If $g \in L^2(\pi)$ satisfies $g \geq h \circ \mathbf{V}$ and $\exists q \in (0, 1)$ such that

$$t \mapsto 1/\Psi \left(G_h^{-1}(\sqrt{t})/(1 - q)^2 \right) \notin L_{\text{loc}}^1(\infty), \quad (1)$$

then the CLT **fails** for the ergodic average $S_n(g)$.

Necessary conditions for the CLT of an ergodic average

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold and recall $S_N(g) = \sum_{k=1}^N g(X_k)/N$. For $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ non-decreasing at infinity and any $q \in (0, 1)$, define

$$G_h(r) := q(1 - q) \frac{h(r)}{r\varphi(1/r)}, \quad r \in [1, \infty),$$

- ① Let G_h^{-1} be the inverse of G_h on a neighbourhood of infinity. If $g \in L^2(\pi)$ satisfies $g \geq h \circ \mathbf{V}$ and $\exists q \in (0, 1)$ such that

$$t \mapsto 1/\Psi \left(G_h^{-1}(\sqrt{t})/(1 - q)^2 \right) \notin L_{\text{loc}}^1(\infty), \quad (1)$$

then the CLT **fails** for the ergodic average $S_n(g)$.

- ② If (1) holds for $h \equiv 1$, then the CLT **fails** for ergodic average $S_n(g)$ for any bounded function $g : \mathcal{X} \rightarrow \mathbb{R}$.

Tails of the invariant measure π of the chain X

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold. For any $q \in (0, 1)$, define

$$L_q(r) := r\varphi(1/r)\Psi(r/(1-q)^2)D(r), \quad r \in [1, \infty),$$

where $D : [1, \infty) \rightarrow \mathbb{R}_+$ tends to infinity arbitrarily slowly.^a

Then there exists a constant $c_q \in (0, 1)$ such that

$$c_q/L_q(r) \leq \pi(\{\mathbf{V} \geq r\}) \quad \text{for all } r \in [1, \infty).$$

^aE.g., $D(r) = \log \circ \dots \circ \log(r)$

Remarks • Comparison of the invariant law of a biased sampler (e.g., ULA) with target law requires lower bound on the tail of its invariant law.

Tails of the invariant measure π of the chain X

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold. For any $q \in (0, 1)$, define

$$L_q(r) := r\varphi(1/r)\Psi(r/(1-q)^2)D(r), \quad r \in [1, \infty),$$

where $D : [1, \infty) \rightarrow \mathbb{R}_+$ tends to infinity arbitrarily slowly.^a

Then there exists a constant $c_q \in (0, 1)$ such that

$$c_q/L_q(r) \leq \pi(\{\mathbf{V} \geq r\}) \quad \text{for all } r \in [1, \infty).$$

^aE.g., $D(r) = \log \circ \dots \circ \log(r)$

- Remarks** • Comparison of the invariant law of a biased sampler (e.g., ULA) with target law requires lower bound on the tail of its invariant law.
- Proof controls **modulated moments**: for $x \in \mathcal{X}$ and non-decreasing h ,

$$\mathbb{P}_x \left(\sum_{k=0}^{\tau_B(1)} h \circ \mathbf{V}(X_k) \geq r \right) \geq \frac{qC_{\ell,m}P(x, \{\mathbf{V} \geq \ell + m\})}{\Psi(G_h^{-1}(r)/(1-q)^2)} \quad \forall r \in (G_h(\ell + m), \infty).$$

Lower bounds on L^p -Wasserstein metric and f -variation

Theorem

Let $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold and $f_* : [1, \infty) \rightarrow (0, \infty)$ smooth. Let a cts $h : [1, \infty) \rightarrow [1, \infty)$ satisfy: $h(r)/f_*(r) \uparrow \infty$ as $r \uparrow \infty$ and $\exists v : \mathcal{X} \times \mathbb{N} \rightarrow [1, \infty)$ increasing in the second argument such that

$$P^n(h \circ V)(x) \leq v(x, n) \quad \text{for all } x \in \mathcal{X} \text{ and } n \in \mathbb{N}.$$

For $g := h/f_*$, assume $a_g(r) \leq f_*(g^{-1}(r))c_q/L_q(g^{-1}(r))$ ($\forall r \geq 1$) and $A_g(r) := ra_g(r) \uparrow \infty$ as $r \uparrow \infty$. Then $\forall x \in \mathcal{X}$ we have:

① **f -variation** (with $f = f_* \circ V$):

$$\frac{1}{2}(a_g \circ A_g^{-1} \circ (2v))(x, n) \leq \|P^n(x, \cdot) - \pi\|_f \quad \forall n \in \mathbb{N} \setminus \{0\}$$

② **L^p -Wasserstein** ($p \geq 1$): for $\mathcal{X}$ metric & V L_V -Lipschitz,

$$\frac{1}{2L_V}(a_g \circ A_g^{-1} \circ (2^p v))(x, n)^{1/p} \leq \mathcal{W}_p(P^n(x, \cdot), \pi) \quad \forall n \in \mathbb{N} \setminus \{0\}$$

Unadjusted Langevin Algorithm (ULA)

Let $\pi \in C^2(\mathbb{R}^d)$. The Euler discretisation of Langevin diffusion

$$dY_t = \nabla \log \pi(Y_t)dt + \sqrt{2}dB_t$$

(B is d -dimensional Brownian motion and ∇ gradient) is **ULA chain** $X^{(h)}$:

$$X_{k+1}^{(h)} = X_k^{(h)} + h\nabla \log \pi(X_k^{(h)}) + \sqrt{2h}Z_{k+1}, \quad \text{for } k \in \mathbb{N},$$

where $(Z_k)_{k \in \mathbb{N}}$ is an i.i.d. sequence with $Z_k \sim N(0, \text{Id})$.

Unadjusted Langevin Algorithm (ULA)

Let $\pi \in C^2(\mathbb{R}^d)$. The Euler discretisation of Langevin diffusion

$$dY_t = \nabla \log \pi(Y_t)dt + \sqrt{2}dB_t$$

(B is d -dimensional Brownian motion and ∇ gradient) is **ULA chain** $X^{(h)}$:

$$X_{k+1}^{(h)} = X_k^{(h)} + h\nabla \log \pi(X_k^{(h)}) + \sqrt{2h}Z_{k+1}, \quad \text{for } k \in \mathbb{N},$$

where $(Z_k)_{k \in \mathbb{N}}$ is an i.i.d. sequence with $Z_k \sim N(0, \text{Id})$.

Advantage of ULA over MALA: process moves at every time step, allowing for rapid exploration of the state space.

Unadjusted Langevin Algorithm (ULA)

Let $\pi \in C^2(\mathbb{R}^d)$. The Euler discretisation of Langevin diffusion

$$dY_t = \nabla \log \pi(Y_t)dt + \sqrt{2}dB_t$$

(B is d -dimensional Brownian motion and ∇ gradient) is **ULA chain** $X^{(h)}$:

$$X_{k+1}^{(h)} = X_k^{(h)} + h\nabla \log \pi(X_k^{(h)}) + \sqrt{2h}Z_{k+1}, \quad \text{for } k \in \mathbb{N},$$

where $(Z_k)_{k \in \mathbb{N}}$ is an i.i.d. sequence with $Z_k \sim N(0, \text{Id})$.

Advantage of ULA over MALA: process moves at every time step, allowing for rapid exploration of the state space.

Drawback: if $X^{(h)}$ is ergodic, then invariant distribution $\pi_h \neq \pi$ (discretisation introduces bias!)

Unadjusted Langevin Algorithm (ULA)

Let $\pi \in C^2(\mathbb{R}^d)$. The Euler discretisation of Langevin diffusion

$$dY_t = \nabla \log \pi(Y_t)dt + \sqrt{2}dB_t$$

(B is d -dimensional Brownian motion and ∇ gradient) is **ULA chain** $X^{(h)}$:

$$X_{k+1}^{(h)} = X_k^{(h)} + h\nabla \log \pi(X_k^{(h)}) + \sqrt{2h}Z_{k+1}, \quad \text{for } k \in \mathbb{N},$$

where $(Z_k)_{k \in \mathbb{N}}$ is an i.i.d. sequence with $Z_k \sim N(0, \text{Id})$.

Advantage of ULA over MALA: process moves at every time step, allowing for rapid exploration of the state space.

Drawback: if $X^{(h)}$ is ergodic, then invariant distribution $\pi_h \neq \pi$ (discretisation introduces bias!)

The **L**-drift conditions **L**(**V**, φ , Ψ) for ULA chain $X^{(h)}$ targeting $\pi(|\cdot| \geq r) \approx r^{-\nu}$ (for some $\nu \in (0, \infty)$) hold with:

$$\mathbf{V}(x) = |x|, \quad \varphi(1/r) = 1/r^3 \quad \text{and} \quad \Psi(r) = r^{\nu+2+\varepsilon}.$$

Unadjusted Langevin Algorithm (ULA)

Let $\pi \in C^2(\mathbb{R}^d)$. The Euler discretisation of Langevin diffusion

$$dY_t = \nabla \log \pi(Y_t)dt + \sqrt{2}dB_t$$

(B is d -dimensional Brownian motion and ∇ gradient) is **ULA chain** $X^{(h)}$:

$$X_{k+1}^{(h)} = X_k^{(h)} + h\nabla \log \pi(X_k^{(h)}) + \sqrt{2h}Z_{k+1}, \quad \text{for } k \in \mathbb{N},$$

where $(Z_k)_{k \in \mathbb{N}}$ is an i.i.d. sequence with $Z_k \sim N(0, \text{Id})$.

Advantage of ULA over MALA: process moves at every time step, allowing for rapid exploration of the state space.

Drawback: if $X^{(h)}$ is ergodic, then invariant distribution $\pi_h \neq \pi$ (discretisation introduces bias!)

The **L-drift** conditions **L**(**V**, φ , Ψ) for ULA chain $X^{(h)}$ targeting π ($|\cdot| \geq r \approx r^{-\nu}$ (for some $\nu \in (0, \infty)$)) hold with:

$$\mathbf{V}(x) = |x|, \quad \varphi(1/r) = 1/r^3 \quad \text{and} \quad \Psi(r) = r^{\nu+2+\varepsilon}.$$

Note $1/\pi(\cdot \geq r) \approx r\varphi(1/r)\Psi(r)$. Not a coincidence!!

ULA chain $X^{(h)}$ targeting $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$, $|x|$ large
 $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold with: $\mathbf{V}(x) = |x|$, $\varphi(1/r) = 1/r^3$, $\Psi(r) = r^{\nu+2+\varepsilon}$.

Theorem

ULA chain $X^{(h)}$ with invariant law π_h & $S_n(g) = \frac{1}{n} \sum_{k=1}^n g(X_k^{(h)})$

① **[CLT-necessary condition]** Fix measurable $g : \mathbb{R}^d \rightarrow \mathbb{R}$.

- (Ia) $g(x) \geq |x|^s$ (all $|x|$ large) for some $2s \in (0, \nu)$: if $2s \in (\nu - 1, \nu)$, asymptotic variance for $S_n(g)$ is **infinite**.
- (Ib) g bounded: if $\nu < 2$, asymptotic var. for $S_n(g)$ **infinite**.

ULA chain $X^{(h)}$ targeting $\pi(x) \approx (1 + |x|)^{-(d+v)}$, $|x|$ large
 $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold with: $\mathbf{V}(x) = |x|$, $\varphi(1/r) = 1/r^3$, $\Psi(r) = r^{v+2+\varepsilon}$.

Theorem

ULA chain $X^{(h)}$ with invariant law π_h & $S_n(g) = \frac{1}{n} \sum_{k=1}^n g(X_k^{(h)})$

① **[CLT-necessary condition]** Fix measurable $g : \mathbb{R}^d \rightarrow \mathbb{R}$.

(Ia) $g(x) \geq |x|^s$ (all $|x|$ large) for some $2s \in (0, v)$: if
 $2s \in (v-1, v)$, asymptotic variance for $S_n(g)$ is **infinite**.

(Ib) g bounded: if $v < 2$, asymptotic var. for $S_n(g)$ **infinite**.

② **[Lower bounds]** Fix $x \in \mathbb{R}^d$. For $p \geq 0$, let $f_p(y) = 1 \vee |y|^p$.

(IIa) For any $p \in [0, v)$ and every $\varepsilon > 0$, $\exists c_p \in (0, \infty)$ such that

$$c_p n^{-\frac{v-p}{2}-\varepsilon} \leq \|Q_h^n(x, \cdot) - \pi_h\|_{f_p} \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

(IIb) For any $p \in [1, v)$ and every $\varepsilon > 0$, $\exists c_{\mathcal{W}, p} \in (0, \infty)$ s.t.

$$c_{\mathcal{W}, p} n^{-\frac{v-p}{2p}-\varepsilon} \leq \mathcal{W}_p(Q_h^n(x, \cdot), \pi_h) \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

ULA chain $X^{(h)}$ targeting $\pi(x) \approx (1 + |x|)^{-(d+v)}$, $|x|$ large
 $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold with: $\mathbf{V}(x) = |x|$, $\varphi(1/r) = 1/r^3$, $\Psi(r) = r^{v+2+\varepsilon}$.

Theorem

ULA chain $X^{(h)}$ with invariant law π_h & $S_n(g) = \frac{1}{n} \sum_{k=1}^n g(X_k^{(h)})$

① **[CLT-necessary condition]** Fix measurable $g : \mathbb{R}^d \rightarrow \mathbb{R}$.

(Ia) $g(x) \geq |x|^s$ (all $|x|$ large) for some $2s \in (0, v)$: if
 $2s \in (v-1, v)$, asymptotic variance for $S_n(g)$ is **infinite**.

(Ib) g bounded: if $v < 2$, asymptotic var. for $S_n(g)$ **infinite**.

② **[Lower bounds]** Fix $x \in \mathbb{R}^d$. For $p \geq 0$, let $f_p(y) = 1 \vee |y|^p$.

(IIa) For any $p \in [0, v)$ and every $\varepsilon > 0$, $\exists c_p \in (0, \infty)$ such that

$$c_p n^{-\frac{v-p}{2}-\varepsilon} \leq \|Q_h^n(x, \cdot) - \pi_h\|_{f_p} \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

(IIb) For any $p \in [1, v)$ and every $\varepsilon > 0$, $\exists c_{\mathcal{W}, p} \in (0, \infty)$ s.t.

$$c_{\mathcal{W}, p} n^{-\frac{v-p}{2p}-\varepsilon} \leq \mathcal{W}_p(Q_h^n(x, \cdot), \pi_h) \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

③ For any $\varepsilon > 0$, there exist $c, C \in (0, \infty)$ satisfying

$$cr^{-v-\varepsilon} \leq \pi_h(|\cdot| \geq r) \leq Cr^{-v+\varepsilon}, \quad \text{for all } r \in [1, \infty).$$

ULA chain $X^{(h)}$ targeting $\pi(x) \approx (1 + |x|)^{-(d+v)}$, $|x|$ large
 $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ hold with: $\mathbf{V}(x) = |x|$, $\varphi(1/r) = 1/r^3$, $\Psi(r) = r^{v+2+\varepsilon}$.

Theorem

ULA chain $X^{(h)}$ with invariant law π_h & $S_n(g) = \frac{1}{n} \sum_{k=1}^n g(X_k^{(h)})$

① **[CLT-necessary condition]** Fix measurable $g : \mathbb{R}^d \rightarrow \mathbb{R}$.

(Ia) $g(x) \geq |x|^s$ (all $|x|$ large) for some $2s \in (0, v)$: if
 $2s \in (v-1, v)$, asymptotic variance for $S_n(g)$ is **infinite**.

(Ib) g bounded: if $v < 2$, asymptotic var. for $S_n(g)$ **infinite**.

② **[Lower bounds]** Fix $x \in \mathbb{R}^d$. For $p \geq 0$, let $f_p(y) = 1 \vee |y|^p$.

(IIa) For any $p \in [0, v)$ and every $\varepsilon > 0$, $\exists c_p \in (0, \infty)$ such that

$$c_p n^{-\frac{v-p}{2}-\varepsilon} \leq \|Q_h^n(x, \cdot) - \pi_h\|_{f_p} \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

(IIb) For any $p \in [1, v)$ and every $\varepsilon > 0$, $\exists c_{\mathcal{W}, p} \in (0, \infty)$ s.t.

$$c_{\mathcal{W}, p} n^{-\frac{v-p}{2p}-\varepsilon} \leq \mathcal{W}_p(Q_h^n(x, \cdot), \pi_h) \quad \text{for all } n \in \mathbb{N} \setminus \{0\}.$$

③ For any $\varepsilon > 0$, there exist $c, C \in (0, \infty)$ satisfying

$$cr^{-v-\varepsilon} \leq \pi_h(|\cdot| \geq r) \leq Cr^{-v+\varepsilon}, \quad \text{for all } r \in [1, \infty).$$

For any $h > 0$, tail decays of π_h and target π **match!**

ULA chain $X^{(h)}$ simulations

Target measure $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for $|x|$ large and some $\nu > 0$
Consider an estimator $S_N(\mathbb{1}_A)$ for tail probability $\pi(A)$.

ULA chain $X^{(h)}$ simulations

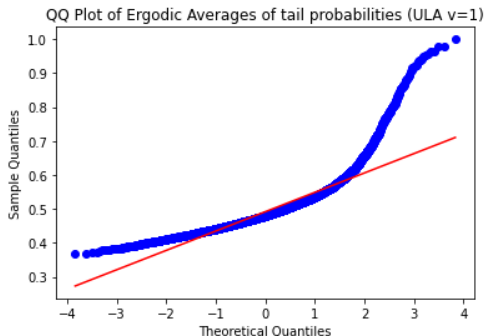
Target measure $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for $|x|$ large and some $\nu > 0$
Consider an estimator $S_N(\mathbb{1}_A)$ for tail probability $\pi(A)$.

By above result, asymptotic variance for $S_n(\mathbb{1}_A)$ is infinite if $\nu < 2$

ULA chain $X^{(h)}$ simulations

Target measure $\pi(x) \approx (1 + |x|)^{-(d+\nu)}$ for $|x|$ large and some $\nu > 0$
Consider an estimator $S_N(\mathbb{1}_A)$ for tail probability $\pi(A)$.

By above result, asymptotic variance for $S_n(\mathbb{1}_A)$ is infinite if $\nu < 2$



QQ-plot ($d = 1$) for ULA suggests **infinite** asymptotic variance ($g = \mathbb{1}_A$)

Heavy-tailed ULA: a large family of potential algorithms

Fix $h > 0$ and cts. **vector field** $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and let chain X evolve as:

$$X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}, \quad \text{for } k \in \mathbb{N}, \quad (2)$$

where $(Z_k^{(h)})_{k \in \mathbb{N}}$ IID in $\mathbb{R}^d$ with **zero mean** and density $p_{\alpha,h}$ satisfying

$$r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0, \quad \text{for some } \alpha \in (1, 2), \quad (3)$$

and $p_{\alpha,h} \xrightarrow{h \downarrow 0}$ Delta measure at $0 \in \mathbb{R}^d$.

Heavy-tailed ULA: a large family of potential algorithms

Fix $h > 0$ and cts. **vector field** $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and let chain X evolve as:

$$X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}, \quad \text{for } k \in \mathbb{N}, \quad (2)$$

where $(Z_k^{(h)})_{k \in \mathbb{N}}$ IID in $\mathbb{R}^d$ with **zero mean** and density $p_{\alpha,h}$ satisfying

$$r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0, \quad \text{for some } \alpha \in (1, 2), \quad (3)$$

and $p_{\alpha,h} \xrightarrow{h \downarrow 0}$ Delta measure at $0 \in \mathbb{R}^d$.

Important Example: a drift $b_\pi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ (in terms of π !) makes Y ,

$dY_t = b_\pi(Y_t)dt + dL_t$, where L is an isotropic **α -stable** Lévy process in $\mathbb{R}^d$,

exponentially ergodic see e.g. [HMW21].

Heavy-tailed ULA: a large family of potential algorithms

Fix $h > 0$ and cts. **vector field** $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and let chain X evolve as:

$$X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}, \quad \text{for } k \in \mathbb{N}, \quad (2)$$

where $(Z_k^{(h)})_{k \in \mathbb{N}}$ IID in $\mathbb{R}^d$ with **zero mean** and density $p_{\alpha,h}$ satisfying

$$r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0, \quad \text{for some } \alpha \in (1, 2), \quad (3)$$

and $p_{\alpha,h} \xrightarrow{h \downarrow 0} \text{Delta measure at } 0 \in \mathbb{R}^d$.

Important Example: a drift $b_\pi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ (in terms of π !) makes Y ,

$dY_t = b_\pi(Y_t)dt + dL_t$, where L is an isotropic **α -stable** Lévy process in $\mathbb{R}^d$,

exponentially ergodic see e.g. [HMW21]. Euler scheme

$$X_{k+1} = X_k + hb(X_k) + h^{1/\alpha} Z_{k+1}, \quad \text{for } k \in \mathbb{N}, \text{ with } p_{\alpha,h} = h^{1/\alpha} q_\alpha$$

satisfies (2)-(3), where q_α is an α -stable density on $\mathbb{R}^d$.

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Theorem

Let $Z_k^{(h)} \sim p_{\alpha,h}$ and $r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0$ (for $1 < \alpha < 2$). For any cts. vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, we have:

If drift b is unbounded, Lipschitz and inward pointing, assumption in Part (1) holds (e.g. Lévy-driven OU process, arising in queueing [APS19])

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Theorem

Let $Z_k^{(h)} \sim p_{\alpha,h}$ and $r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0$ (for $1 < \alpha < 2$). For any cts. vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, we have:

- 1 Assume $|b(x)| \leq C(|x| + 1) \forall x \in \mathbb{R}^d$ and constant $C > 0$. If X has invariant distrib. π_h , then $\forall \varepsilon > 0 \exists C_{\pi,h} \in (0, \infty)$ s.t.

$$\pi_h(|\cdot| > r) \geq C_{\pi,h} r^{-\alpha-\varepsilon} \quad \text{for all } r \in [1, \infty).$$

If drift b is unbounded, Lipschitz and inward pointing, assumption in Part (1) holds (e.g. Lévy-driven OU process, arising in queueing [APS19])

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Theorem

Let $Z_k^{(h)} \sim p_{\alpha,h}$ and $r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0$ (for $1 < \alpha < 2$). For any cts. vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, we have:

- 1 Assume $|b(x)| \leq C(|x| + 1) \forall x \in \mathbb{R}^d$ and constant $C > 0$. If X has invariant distrib. π_h , then $\forall \varepsilon > 0 \exists C_{\pi,h} \in (0, \infty)$ s.t.

$$\pi_h(|\cdot| > r) \geq C_{\pi,h} r^{-\alpha-\varepsilon} \quad \text{for all } r \in [1, \infty).$$

- 2 If $\liminf_{|x| \rightarrow \infty} |b(x)|/|x| = \infty$, then chain X is transient:

$$\mathbb{P}_x \left(\lim_{k \rightarrow \infty} |X_k| = \infty \right) = 1 \quad \text{for all } x \in \mathbb{R}^d.$$

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Theorem

Let $Z_k^{(h)} \sim p_{\alpha,h}$ and $r^\alpha \int_{\{|x|>r\}} p_{\alpha,h}(x) dx \xrightarrow{r \uparrow \infty} c > 0$ (for $1 < \alpha < 2$). For any cts. vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, we have:

- 1 Assume $|b(x)| \leq C(|x| + 1) \forall x \in \mathbb{R}^d$ and constant $C > 0$. If X has invariant distrib. π_h , then $\forall \varepsilon > 0 \exists C_{\pi,h} \in (0, \infty)$ s.t.

$$\pi_h(|\cdot| > r) \geq C_{\pi,h} r^{-\alpha-\varepsilon} \quad \text{for all } r \in [1, \infty).$$

- 2 If $\liminf_{|x| \rightarrow \infty} |b(x)|/|x| = \infty$, then chain X is transient:

$$\mathbb{P}_x \left(\lim_{k \rightarrow \infty} |X_k| = \infty \right) = 1 \quad \text{for all } x \in \mathbb{R}^d.$$

If drift b is unbounded, Lipschitz and inward pointing, assumption in Part (1) holds (e.g. Lévy-driven OU process, arising in queueing [APS19])

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): **L**-drift condition **L**(**V**, φ , Ψ) holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): **L**-drift condition **L**(**V**, φ , Ψ) holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of Part (2): show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): **L**-drift condition **L**(**V**, φ , Ψ) holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of Part (2): show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Continuous time. Langevin diffusion with Lévy noise and appropriate drift: **exponential** ergodicity to **heavy-tailed** measure [HMW21; ZZ23].

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): $\mathbf{L}$ -drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of Part (2): show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Continuous time. Langevin diffusion with Lévy noise and appropriate drift: **exponential** ergodicity to **heavy-tailed** measure [HMW21; ZZ23]. By [HMW21, Thm 1.1], isotropic α -stable driven SDE converges **exponentially** fast to π with $\alpha \in (1, 2)$ moments.

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): **L**-drift condition $\mathbf{L}(\mathbf{V}, \varphi, \Psi)$ holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of Part (2): show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Continuous time. Langevin diffusion with Lévy noise and appropriate drift: **exponential** ergodicity to **heavy-tailed** measure [HMW21; ZZ23]. By [HMW21, Thm 1.1], isotropic α -stable driven SDE converges **exponentially** fast to π with $\alpha \in (1, 2)$ moments.

Euler scheme for **Lévy-driven SDE cannot work** for $\pi(|x|^2) < \infty$!
Chain X is either convergent to π_h **without** a finite $(\alpha + \varepsilon)$ -moment (X in Ψ -fast φ -slow regime) or is a **transient** (no convergence).

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of Part (1): **L**-drift condition **L**(**V**, φ , Ψ) holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of Part (2): show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Continuous time. Langevin diffusion with Lévy noise and appropriate drift: **exponential** ergodicity to **heavy-tailed** measure [HMW21; ZZ23]. By [HMW21, Thm 1.1], isotropic α -stable driven SDE converges **exponentially** fast to π with $\alpha \in (1, 2)$ moments.

Euler scheme for **Lévy-driven SDE cannot work** for $\pi(|x|^2) < \infty$!
Chain X is either convergent to π_h **without** a finite $(\alpha + \varepsilon)$ -moment (X in Ψ -fast φ -slow regime) or is a **transient** (no convergence).

This is a **heavy-tailed analogue** of famous example [RT96, Thm 3.2]: Euler scheme for Langevin diffusion targeting π with tails **lighter than Gaussian** is **necessarily transient**.

Heavy-tailed ULA chain: $X_{k+1} = X_k + hb(X_k) + Z_{k+1}^{(h)}$

Proof of **Part (1)**: **L**-drift condition **L**(**V**, φ , Ψ) holds with $V(x) = \log(|x| \vee e)$, $\Psi(r) = \exp(\alpha r)$ and $\varphi(1/r) \propto 1/r^2$.

Proof of **Part (2)**: show $1/V(X)$ is a supermartingale and transience follows by standard arguments [MT93].

Continuous time. Langevin diffusion with Lévy noise and appropriate drift: **exponential** ergodicity to **heavy-tailed** measure [HMW21; ZZ23]. By [HMW21, Thm 1.1], isotropic α -stable driven SDE converges **exponentially** fast to π with $\alpha \in (1, 2)$ moments.

Euler scheme for **Lévy-driven SDE cannot work** for $\pi(|x|^2) < \infty$! Chain X is either convergent to π_h **without** a finite $(\alpha + \varepsilon)$ -moment (X in Ψ -fast φ -slow regime) or is a **transient** (no convergence).

This is a **heavy-tailed analogue** of famous example [RT96, Thm 3.2]: Euler scheme for Langevin diffusion targeting π with tails **lighter than Gaussian** is **necessarily transient**.

New **tamed schemes** with heavy-tailed increments are needed to **preserve** exponential ergodicity of the underlying continuous-time dynamics!

We now apply our theory to the following families of MCMC algorithms:

((SIM-ALG)) **RWM** with **finite** (**fv-RWM**) and **infinite** (**iv-RWM**) variance proposals, **MALA**, Unadjusted Langevin algorithm (**ULA**) and the Stereographic Projection Sampler (**SPS**).

See Part II of the presentation

- [APS19] Ari Arapostathis, Guodong Pang, and Nikola Sandrić. "Ergodicity of a Lévy-driven SDE arising from multiclass many-server queues". *Ann. Appl. Probab.* 29 (2019).
- [BM24] Miha Brešar and Aleksandar Mijatović. "Subexponential lower bounds for f -ergodic Markov processes". *PTRF* (2024), 58pp. arXiv: 2403.14826 [math.PR].
- [DGM08] Randal Douc, Arnaud Guillin, and Eric Moulines. "Bounds on regeneration times and limit theorems for subgeometric Markov chains". *Ann. Inst. Henri Poincaré Probab. Stat.* 44 (2008).
- [DL18] George Deligiannidis and Anthony Lee. "Which ergodic averages have finite asymptotic variance?" *Ann. Appl. Probab.* 28.4 (2018).
- [Dou+04] Randal Douc, Gersende Fort, Eric Moulines, and Philippe Soulier. "Practical drift conditions for subgeometric rates of convergence". *Ann. Appl. Probab.* 14 (2004).
- [Dou+18] Randal Douc, Eric Moulines, Pierre Priouret, and Philippe Soulier. *Markov chains. Springer Series in Operations Research and Financial Engineering*. Springer, Cham, 2018.
- [Dou+25] Randal Douc, Pierre E Jacob, Anthony Lee, and Dootika Vats. "Solving the Poisson equation using coupled Markov chains". arXiv preprint arXiv:2206.05691, to appear in *Annals of Statistics* (2025).
- [FM03] G. Fort and E. Moulines. "Polynomial ergodicity of Markov transition kernels". *Stochastic Process. Appl.* 103 (2003).
- [Gel+08] Andrew Gelman, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. "A weakly informative default prior distribution for logistic and other regression models". *Ann. Appl. Stat.* 2 (2008).
- [GLM18] Joyee Ghosh, Yingbo Li, and Robin Mitra. "On the use of Cauchy prior distributions for Bayesian logistic regression". *Bayesian Anal.* 13 (2018).
- [HMW21] Lu-Jing Huang, Mateusz B. Majka, and Jian Wang. "Approximation of heavy-tailed distributions via stable-driven SDEs". *Bernoulli* 27 (2021).
- [Jon04] Galin L. Jones. "On the Markov chain central limit theorem". *Probab. Surv.* 1 (2004), pp. 299–320.
- [KV86] C. Kipnis and S. R. S. Varadhan. "Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions". *Comm. Math. Phys.* 104.1 (1986), pp. 1–19. ISSN: 0010-3616.
- [MT93] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability. Communications and Control Engineering Series*. Springer-Verlag London, Ltd., London, 1993.
- [OM24] Alain Oliviero-Durmus and Éric Moulines. "On geometric convergence for the Metropolis-adjusted Langevin algorithm under simple conditions". *Biometrika* 111 (2024).
- [RT96] Gareth O. Roberts and Richard L. Tweedie. "Exponential convergence of Langevin distributions and their discrete approximations". *Bernoulli* 2 (1996).
- [ZZ23] Xiaolong Zhang and Xicheng Zhang. "Ergodicity of supercritical SDEs driven by α -stable processes and heavy-tailed sampling". *Bernoulli* 29 (2023).

Thank you for your attention!!