Sampling crash course

Andreas Eberle, Univ. Bonn 31.8.21
wt.iam.uni-bonn.de/eberle

Additional references:

- AE, N. Bou-Rabee: Lecture notes on MCMC methods (see homepage)
- Montenegro/Tetali: Mathem. aspects of mixing times of Markov chains
- Levin/Peres/Wilmer: Markov chains and mixing times

$$\boxed{\mu(dx) = \underbrace{\frac{1}{Z} e^{-U(x)}}_{=:\mu(x)} \underbrace{\lambda^d}_{\text{Lebesgue measure}}(dx)}$$ prob. measure on $\mathbb{R}^d$

GOAL Generation of approximate samples from μ

REMARK (From sampling to optimization)

$$\mu_\beta(dx) = \frac{1}{Z_\beta} e^{-\beta U(x)} \lambda^d(dx), \quad \beta \uparrow \infty$$

$\Rightarrow$ μ_β concentrates near minima of U

APPROACHES

- **Direct methods** e.g. in $\mathbb{R}^1$: $x = \underset{\substack{\uparrow \\ \text{generalized inverse of cdf}}}{F^{-1}}(u)$, $u \sim \mathrm{Unif}(0,1)$

- Acceptance-Rejection (AR)

$\nu(dx)$ prob. measure on $\mathbb{R}^d$ "proposal distribution"

$$\mu(dx) \propto \varrho(x)\, \nu(dx)$$

Assumption: $\exists\, C \in (0,\infty) : \varrho \leq C$

AR algorithm: repeat

$x \leftarrow \text{Sample}(\nu)$

$u \leftarrow \text{Sample}(\text{Unif}(0,1))$

until $u \leq \varrho(x)/C$

- Generates exact sample from μ
- Running time $T \sim \text{Geom}(p)$, $p = \frac{1}{C} \int \varrho\, d\nu$
- Average running time $\mathbb{E}[T] = \frac{1}{p}$ often too large !

- Markov Chain Monte Carlo (MCMC)

$\pi(x, dy)$ probability kernel on $\mathbb{R}^d$ "transition probab."

Assumption: μ is invariant measure (stationary distrib.) for π

$$\boxed{\mu(B) = (\mu\pi)(B)} := \int_{\mathbb{R}^d} \mu(dx)\,\pi(x,B) \quad (B \in \mathcal{B}(\mathbb{R}^d))$$

or at least "$\approx$"

$\nu(dx)$ prob. measure on $\mathbb{R}^d$ "initial distribution"

$X_0, X_1, X_2, \ldots$ Markov chain (ν, π)

$\nu_t(B) := \mathbb{P}[X_t \in B] = (\nu\pi^t)(B)$ "Law at time t"

IDEA: t sufficiently large ? $\overset{?}{\Longrightarrow}$ $\nu_t \overset{?}{\approx} \mu$

Assumptions

KEY QUESTIONS:

- Upper and lower bounds for $\mathcal{W}(\nu_t, \mu)$

  $\mathcal{W}$ distance function on prob. measures / divergence

  e.g. TV, Wasserstein   e.g. χ^2, KL

- Bounds for mixing / relaxation times

$$t^{W}_{mix}(\varepsilon) := \min\{t>0 : W(\nu_t,\mu) \leq \varepsilon W(\nu,\mu) \ \forall \nu \in Prob(\mathbb{R}^d)\}$$

$$t^{TV}_{mix}(\varepsilon, K) := \min\{t>0 : TV(\nu_t,\mu) \leq \varepsilon \ \forall \nu \in Prob(K)\}$$

- Bounds for ergodic averages (→ integral estimation)

$$\mathbb{P}\left[\left|\frac{1}{t}\sum_{i=b}^{b+t-1} f(X_i) - \int f d\mu\right| \geq \varepsilon\right] \leq \cdots$$

MATHEMATICAL APPROACHES

A. Coupling methods

- Couplings of transition kernels / Wasserstein contractions
- Couplings of Markov processes / Coupling lemma
- Synchronous and reflection couplings of diffusions
- Weak Harris theorem

B. Functional inequalities

- Conductance
- Spectral gap
- Logarithmic Sobolev inequalities
- Hypocoercivity

1. METROPOLIS-HASTINGS METHODS

$$p(x,dy) = \underbrace{p(x,y)}_{>0} \lambda^d(dy)$$ "proposal kernel"

Transition step $x \rightsquigarrow x'$ of Markov chain given by:

- Sample $y \sim p(x,dy)$
- Sample $u \sim \mathrm{Unif}(0,1)$
- If $u \leq \frac{\mu(y)\, p(y,x)}{\mu(x)\, p(x,y)}$ then accept: $x' \leftarrow y$
  else reject: $x' \leftarrow x$

Defines Markov chain with transition kernel

$$\pi(x,dy) = a(x,y)\, p(x,dy) + r(x)\, \delta_x(dy)$$

$$a(x,y) = \min\left(\frac{\mu(y)\, p(y,x)}{\mu(x)\, p(x,y)}, 1 \right)$$ Acceptance prob.

$$r(x) = 1 - \int a(x,y)\, p(x,dy)$$ Rejection prob.

THEOREM μ is invariant for π.

Proof. $\mu(dx)\pi(x,dy) = \mu(dy)\pi(y,dx)$ Detailled balance

Integration over x $\Rightarrow$ $\mu\pi = \mu$ $\square$

VARIANTS 1) **Independence Sampler**: $p(x, dy) = \nu(dy)$
independent of x

Similar to Accept-Reject.

2) **Random Walk Metropolis** $p(x, \cdot) = N(x, h I_d)$, $h > 0$ fixed

$p(x,y) = p(y,x) \Rightarrow$ $\boxed{a(x,y) = e^{-(U(y)-U(x))^+}}$

Proposal $Y = x + \sqrt{h}\, Z$, $Z \sim N(0, I_d)$

3) **Metropolis adjusted Langevin algorithm (MALA)**

Proposal $Y = x - h \nabla U(x) + \sqrt{2h}\, Z$, $Z \sim N(0, I_d)$

see further below

EXAMPLE: RWM with Gaussian target

$\mu = N(0, I_d)$, $U(x) = \frac{1}{2}|x|^2$, $Y = x + \sqrt{h}\, Z$

$$U(Y)-U(x) = \underbrace{\sqrt{h}\, x^T Z}_{\in \mathcal{O}(|x|)} + \frac{h}{2}\underbrace{\sum_{i=1}^{d}\left(Z_i^2 - 1\right)}_{\in \mathcal{O}(\sqrt{d}) \text{ by CLT}} + \frac{hd}{2}$$

hd large $\Rightarrow$ $a(x,Y)$ typically very small

$$\mathbb{P}\left[a(x,Y) \geq e^{-\frac{hd}{4}}\right] \leq e^{-\frac{1}{128} hd \cdot \frac{d}{|x|^2}} + e^{-\underset{\text{constant} > 0}{c}d}$$

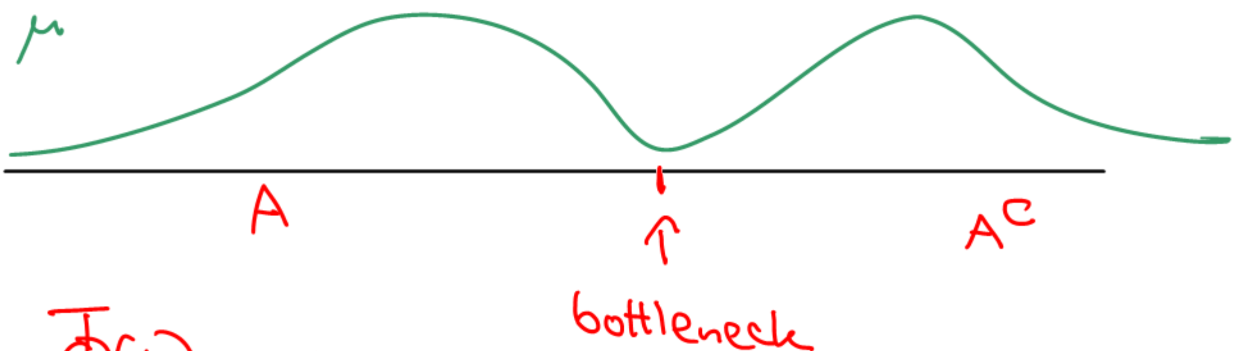
Lower bound for mixing time

$$Q(A,B) := \int_A \mu(dx)\, \pi(x,B)$$

Equilibrium flow from A to B

$$\Phi(A) := \frac{Q(A,A^c)}{\min(\mu(A), \mu(A^c))}$$

Conductance of A



$\Phi(A)$ small

THEOREM $\mu(A) \in [0, \frac{1}{2}] \Rightarrow t_{mix}^{TV}(\frac{1}{4}, A) \geq \frac{1}{4\Phi(A)}$

EXAMPLE RWM with $\mu = N(0, I_d)$:

$$t_{mix}\left(\frac{1}{4}, B(0,r)\right) \geq \frac{1}{8} e^{\frac{hd}{4}} \quad \forall h < \frac{1}{20}, r > 0$$

$\rightsquigarrow$ need $h \in O(d^{-1})$!

REMARK By Cheeger's inequality, conductance also provides upper bounds on L^2 relaxation times and mixing times from warm start, see e.g. Montenegro / Tetali.

Improved proposals? Take into account ∇U.

2. OVERDAMPED LANGEVIN DYNAMICS

$\mu(x) \propto e^{-U(x)}$

Continuous time Markov process with inv. measure μ?

$$(*) \quad dX_t = b(X_t)\,dt + \sqrt{2}\,dB_t, \quad X_0 \sim \nu$$

$$b = -\nabla U + \beta \quad \in C^1(\mathbb{R}^d, \mathbb{R}^d)$$

THEOREM Suppose $\limsup\limits_{|x|\to\infty} \frac{x^T b(x)}{|x|^2} < \infty$. Then:

1) $\exists!$ non-explosive solution for every initial law ν.

2) $(X_t, \mathbb{P}_\nu)$ is a strong Markov process with

$$(P_t f)(x) := \int p_t(x, dy)\, f(y) = \mathbb{E}_{\delta_x}[f(X_t)] \quad \text{transition semigroup}$$

$$\mathcal{L}f = \lim_{t \downarrow 0} \frac{P_t f - f}{t} = \Delta f + b^T \nabla f \quad \forall f \in C_0^\infty \quad \text{generator}$$

e.g. in sup-norm

3) μ invariant for $(p_t)_{t\geq 0} \iff \operatorname{div}(e^{-U}\beta) = 0$

4) Detailled balance holds $\iff \beta = 0$ Kolmogorov

PROOF 1) Drift condition $\Rightarrow$ non-explosive

2) Itō's formula $\Rightarrow f(X_t) - \int_0^t (\mathcal{L}f)(X_s)\,ds$ martingale

3) $\mu p_t = \mu \quad \forall t \geq 0$

$$\iff \int p_t f \, d\mu = \int f \, d\mu \quad \forall f \geq 0 \text{ meas.}$$

C_0^∞ operator core

$$\overset{(\dots)}{\iff} \int \mathcal{L}f \, d\mu = 0 \qquad \forall f \in C_0^\infty(\mathbb{R}^d)$$

integr. by parts

$$\iff \operatorname{div}(e^{-U}\beta) = 0$$

4) $b = -\nabla U$

$$\Rightarrow -\int \mathcal{L}f \, g \, d\mu \overset{\text{integr. by parts}}{=} \underbrace{\int \nabla f \cdot \nabla g \, d\mu}_{\text{symmetric Dirichlet form}} \quad \forall f,g \in C_0^\infty$$

C_0^∞ core

$$\Rightarrow \int p_t f \, g \, d\mu = \int f \, p_t g \, d\mu \quad \forall f,g \in \mathcal{L}^2(\mu)$$

$\square$

Additional ref.:

Royer: An initiation to log. Sobolev inequalities

AE: Lecture notes "Markov Processes", Homepage

Couplings and convergence to equilibrium

$$W^p(\nu,\mu) := \inf_{\substack{X\sim\nu \\ Y\sim\mu}} \mathbb{E}\left[|X-Y|^p\right]^{1/p} \qquad p\in[1,\infty)$$

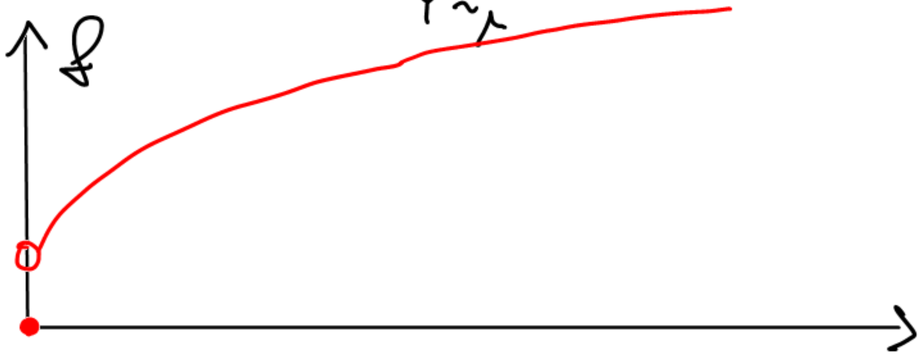
Coupling of μ,ν

L^p Wasserstein distance

$$W_f(\nu,\mu) := \inf_{\substack{X\sim\nu \\ Y\sim\mu}} \mathbb{E}\left[f(|X-Y|)\right] \qquad f:[0,\infty)\to[0,\infty)$$

non-decr., concave

$f(0)=0$



EXAMPLE TV-distace: $f(t)=\mathbb{1}_{t>0}$

$$W_f(\nu,\mu) = \inf_{\substack{X\sim\nu \\ Y\sim\mu}} \mathbb{P}[X\neq Y] = \|\nu-\mu\|_{TV}$$

REMARK (X_t,Y_t) coupling of $MP(\nu,p_t)$ and $MP(\mu,p_t)$

$$\Longrightarrow \quad W^p(\nu p_t, \overbrace{\mu p_t}^{=\mu}) \leq \mathbb{E}\left[|X_t-Y_t|^p\right]^{1/p}$$

$$W_f(\nu p_t, \mu p_t) \leq \mathbb{E}\left[f(|X_t-Y_t|)\right]$$

$$(**) \quad \begin{aligned} dX_t &= b(X_t)\,dt + \sqrt{2}\,dB_t, \quad X_0 \sim \nu \\ dY_t &= b(Y_t)\,dt + \sqrt{2}\,d\tilde{B}_t, \quad Y_0 \sim \mu \end{aligned}$$

$B, \tilde{B}$ Brownian motions

a) Synchronous coupling $\tilde{B} = B$

$$\frac{d}{dt}(X_t - Y_t) = b(X_t) - b(Y_t)$$

$$A(K) \quad (x-y)^T(b(x)-b(y)) \leq -K|x-y|^2 \quad \forall x,y \in \mathbb{R}^d$$

if $b = -\nabla U$: $\Leftrightarrow$ U K-strongly convex

THEOREM If $A(K)$ holds for some $K>0$ then

1) $|X_t - Y_t| \leq e^{-Kt}|X_0 - Y_0| \quad \forall t \geq 0$

2) $\mathcal{W}^p(\nu p_t, \mu p_t) \leq e^{-Kt}\,\mathcal{W}^p(\nu, \mu) \quad \forall p \in [1,\infty),\ t \geq 0$

Wasserstein contractivity

PROOF 1) ✓

2) (X_0, Y_0) coupling of $\nu, \mu \Rightarrow (X_t, Y_t)$ coupling of $\nu p_t, \mu p_t$

$\Rightarrow W^p(\nu p_t, \mu p_t) \le \mathbb{E}[|X_t - Y_t|^p]^{1/p} \le e^{-kt}\, \mathbb{E}[|X_0 - Y_0|^p]^{1/p}$

Now minimize over all couplings of μ and ν. $\square$

REMARKS ⊕ Result is sharp:

$\sup_{x \neq y} \frac{(x-y)^T (b(x)-b(y))}{|x-y|^2} \ge 0 \Rightarrow$ Strict W^p contractivity does not hold

⊕ dimension-free

⊕ robust

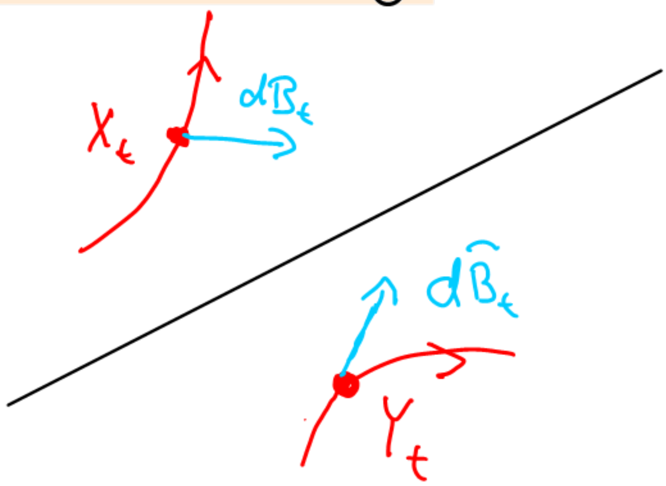
⊖ only applies in strongly convex case

⊖ noise is not used at all

⊖ does not provide TV bounds

b) Reflection coupling

(Lindvall & Rogers '86)



$$d\tilde{B}_t = \left(I_d - 2\,\mathbb{1}_{X_t \neq Y_t}\, e_t e_t^T\right) dB_t$$

$$e_t = \frac{X_t - Y_t}{|X_t - Y_t|}$$

$\tilde{B}$ is a Brownian motion by Lévy's characterization

$\exists\ K, L > 0,\ R \in [0, \infty)$ such that

$A(K,L,R)$

$$(x-y)^T\big(b(x)-b(y)\big) \leq \begin{cases} L\,|x-y|^2 & \forall x, y \in \mathbb{R}^d \\ -K\,|x-y|^2 & \text{if } |x-y| > R \end{cases}$$

THEOREM If $A(K,L,R)$ holds then $\exists\ c, a, m, R_1 \in (0,\infty)$,

$f : [0,\infty) \to [0,\infty)$ increasing, concave, C^1 on $(0,\infty)$ s.t.

$$\mathbb{E}\left[f(|X_t - Y_t|)\right] \leq e^{-ct}\, \mathbb{E}\left[f(|X_0 - Y_0|)\right]$$

Proof via Itô's formula

COROLLARY $\boxed{\omega_f(\nu p_t, \mu p_t) \leq e^{-ct}\, \omega_f(\nu, \mu)}$

REMARKS

- All constants and function f are explicit and depend only on L, K, R.
- Dimension free for L, K, R fixed.
- Provides TV bounds: $\| \nu p_t - \mu p_t \|_{TV} \leq \frac{1}{a}\, \omega_f(\nu p_t, \mu p_t)$
- Arbitrary initial laws

Additional ref.

Lindvall / Rogers: Coupling of multidim. diffusions by reflection
AoP 1986

AE: Reflection couplings and contraction rates for diffusions
PTRF 2016

Relaxation time and entropy decay rate

$$\chi^2(\nu|\mu) := \begin{cases} \int (\varrho-1)^2 d\mu & \text{if } \nu \ll \mu \text{ with density } \varrho \\ \infty & \text{otherwise} \end{cases}$$

$$H(\nu|\mu) := \begin{cases} \int \varrho \log \varrho \, d\mu & \text{if } \nu \ll \mu \text{ with density } \varrho \\ \infty & \text{otherwise} \end{cases}$$

Kullback-Leibler divergence, Relative entropy

FACT μ invariant $\Rightarrow \chi^2(\nu p_t|\mu),\ H(\nu p_t|\mu)$ are non-increasing in t

$$t_{rel} := \inf \{ t \geq 0 : \chi^2(\nu p_t|\mu) \leq \frac{1}{e} \chi^2(\nu|\mu) \ \forall \nu \in \mathrm{Prob}(\mathbb{R}^d) \}$$

L^2 relaxation time

$$t_{ent} := \inf \{ t \geq 0 : H(\nu p_t|\mu) \leq \frac{1}{e} \underbrace{H(\nu|\mu)} \ \forall \nu \in \mathrm{Prob}(\mathbb{R}^d) \}$$

Entropy relaxation time

can be infinite, e.g. for $\nu = \delta_x$

$$t_{mix}(\varepsilon, M) := \inf \{ t \geq 0 : \|\nu p_t - \mu\|_{TV} \leq \varepsilon \ \forall \nu \text{ with } \frac{d\nu}{d\mu} \leq M \}$$

Mixing time from M-warm start

THEOREM 1) $t_{mix}(\varepsilon, M) \leq \left(\log(M-1) + 2\log\varepsilon^{-1}\right) t_{rel}$

2) $t_{mix}(\varepsilon, M) \leq \left(\log\log M + 2\log\varepsilon^{-1} + 1 - \log 2\right) t_{ent}$

Proof: Exercise, Hint: Pinsker's inequality

$$\|\nu p_t - \mu\|_{TV}^2 \leq \frac{1}{2} H(\nu p_t | \mu)$$

Upper bounds on t_{rel} and t_{ent} by functional inequalities

$$\mathcal{E}(f,g) = -\int f \mathcal{L} g \, d\mu = \int_{\mathbb{R}^d} \nabla f \cdot \nabla g \, d\mu$$

Dirichlet form

Assume reversibility, i.e., $b = -\nabla U$.

THEOREM 1) The following statements are equivalent:

(i) $\int \left(f - \int f d\mu\right)^2 d\mu \leq C\, \mathcal{E}(f,f) \quad \forall f \in C_0^\infty(\mathbb{R}^d)$

Poincaré inequality ($\Longleftrightarrow$ Spectral gap $\lambda = \frac{1}{C}$)

(ii) $\chi^2(\nu p_t | \mu)^{1/2} \leq e^{-t/C} \chi^2(\nu|\mu)^{1/2} \quad \forall \nu \in \mathrm{Prob}(\mathbb{R}^d),\ t \geq 0$

Exponential decay of χ^2 divergence

2) The following statements are equivalent:

(i) $\int f^2 \log \frac{f^2}{\int f^2 d\mu} \, d\mu \leq 2C\,\mathcal{E}(f,f) \quad \forall f \in C_0^\infty(\mathbb{R}^d)$

Logarithmic Sobolev inequality

(ii) $H(\nu P_t | \mu) \leq e^{-\frac{2t}{C}} H(\nu|\mu) \quad \forall \nu \in \mathrm{Prob}(\mathbb{R}^d),\ t \geq 0$

Exponential decay of relative entropy

Informal proof of 2) Suppose $\nu P_t \ll \mu$ with smooth density $\varrho_t \geq \varepsilon > 0$.

$$\frac{d}{dt} H(\nu P_t | \mu) = \frac{d}{dt} \int \varrho_t \log \varrho_t \, d\mu$$

$$\overset{\frac{d}{dt}\varrho_t = \mathcal{L}^* \varrho_t}{=} 2\int (1+\log \varrho_t)\, \mathcal{L}^* \varrho_t \, d\mu$$

$$= -\mathcal{E}(\varrho_t, \log \varrho_t) = -\int \nabla \varrho_t \cdot \nabla \log \varrho_t \, d\mu$$

$$= -4\int |\nabla \sqrt{\varrho_t}|^2 d\mu = -4\,\mathcal{E}(\sqrt{\varrho_t}, \sqrt{\varrho_t})$$

$$\overset{LSI}{\leq} -\frac{2}{C} H(\nu P_t | \mu)$$

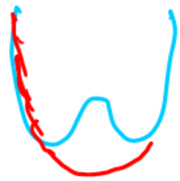
$\square$

COROLLARY

"=" in reversible case

$$t_{rel} \le C_{Poin}/2$$

$$t_{ent} \le C_{LS}/2$$



THEOREM (Bakry-Emery + Holley-Stroock)

Suppose $\boxed{U = V + W}$, $V, W \in C^2$, $\boxed{\nabla^2 V \geq K \cdot I_d}$, $\boxed{\sup W - \inf W \leq M}$.

strongly convex   bounded perturbation

$\Longrightarrow$ Log. Sobolev inequality with $\boxed{C_{LS} = e^M K}$

Additional refs.

Bakry, Gentil, Ledoux: Analysis and geometry of Markov diffusion operators

Royer: An initiation to LSI

3. (OVERDAMPED) LANGEVIN ALGORITHMS

$$dX_t = -\nabla U(X_t)\,dt + \sqrt{2}\,dB_t$$

Euler discretization: Markov chain with transition $x \rightsquigarrow X'$

$$X' = x - h\nabla U(x) + \sqrt{2h}\,Z, \quad Z \sim N(0, I_d),\ h > 0$$

$$p(x,y) \propto e^{-\frac{1}{4h}|y-x+h\nabla U(x)|^2}$$ trans. density

$\longrightarrow$ Unadjusted Langevin Algorithm (ULA)

EXAMPLE $\mu = N(0, I_d)$, $\nabla U(x) = x$ $\rightarrow$ AR(1) process

Invariant measure of p: $\mu_h = N\left(0, \frac{1}{1-\frac{h}{2}} I_d\right) \neq \mu$

Asymptotic bias !

$$W^2(\mu_h, \mu) = \mathbb{E}\left[\left|\frac{1}{\sqrt{1-\frac{h}{2}}}Z - Z\right|^2\right]^{1/2}$$

$$= \left|\frac{1}{\sqrt{1-\frac{h}{2}}} - 1\right| d^{1/2} \sim \frac{1}{4} h d^{1/2} \overset{!}{\leq} \varepsilon$$

$\longrightarrow$ need $h \in O(d^{-1/2}\varepsilon^{-1})$ (at least)

THEOREM Assume

(A1) $\exists\, \kappa > 0 : \operatorname{Hess} U \geq \kappa I_d$

(A2) $\exists\, L < \infty : |\nabla U(x) - \nabla U(y)| \leq L|x-y| \quad \forall x,y$

(A3) $\exists\, M < \infty : \sup_{\substack{n\in\mathbb{N} \\ nh\leq 1}} \| X_n^{(h)} - X_{nh} \|_{L^2(\mathbb{P})} \leq M h \quad \forall h \leq \frac{1}{2},\ X_0 \sim \mu$

Euler discr. OLD

Then $$W^2(\mu, \mu_h) \leq \max\left(4, \frac{8}{\kappa}\right) M h$$

$$W^2(\nu p^n, \mu) \leq e^{-\frac{\kappa n h}{2}} W^2(\nu, \mu_h) + \max\left(4, \frac{8}{\kappa}\right) M h$$

REMARK 1) (A1) can be relaxed substantially, see [Durmus/Moulines], [De Bortoli/Durmus].

2) $M \in O(d^{1/2})$ for "nice" models, in general $M \in O(d)$, see [Durmus/AE], Arxiv.

Alternatively: use Euler step as MH-proposal

$\rightarrow$ Metropolis-adjusted Langevin algorithm

THEOREM For MALA with $\mu = N(0, I_d)$,

$$t_{mix}\left(\tfrac{1}{4}, B(0,r)\right) \geq \frac{1}{8} e^{h^2 d/4} \quad \forall h < \tfrac{1}{3},\ r > 0$$

$\Longrightarrow$ For cold start need to choose $h \in O(d^{-1/2})$

Recent converse result:

[Chewi, Lu, Ahn, Cheng, Le Gouic, Rigollet]

Mixing from warm start in $O(d^{\frac{1}{2}} \log d)$ steps.

4. MCMC WITH DEGENERATE NOISE

Too much randomness $\Rightarrow$ Diffusive behaviour
Slow mixing

a) LANGEVIN DYNAMICS

position

velocity

State space $\mathbb{R}^d \times \mathbb{R}^d \ni (x, v)$

$$dX_t = V_t \, dt$$

$$dV_t = \underbrace{-\nabla U(X_t)\, dt}_{\text{Hamiltonian dynamics}} \underbrace{- \gamma V_t \, dt + \sqrt{2\gamma}\, dB_t}_{\text{random collisions and damping}}$$

$\gamma > 0$ friction

- $\exists\,!$ solution for given initial law, Markov process
- Invariant measure is $\mu_{BG} = \mu \otimes N(0, I_d)$

$$\mu_{BG}(dx\,dv) = \frac{1}{Z} e^{-H(x,v)} \lambda^{2d}(dx\,dv)$$

$$H(x,v) = U(x) + \frac{1}{2}|v|^2$$ Hamiltonian

- Transition function p_t

EXAMPLE

$$\mu = N(0,C),\quad C \in \mathbb{R}^{d\times d} \text{ symm. pos. def.}$$

$$U(x) = \frac{1}{2} x^T C^{-1} x$$

$$\boxed{d\begin{pmatrix} X \\ V \end{pmatrix} = -A\begin{pmatrix} X \\ V \end{pmatrix}dt + \begin{pmatrix} 0 \\ \sqrt{2\gamma}\, dB \end{pmatrix}}$$

$$A = \begin{pmatrix} 0 & -I_d \\ C^{-1} & \gamma I_d \end{pmatrix}$$ Gaussian process

Wasserstein contraction coefficient:

Gaussian measures with same covariance

$$\alpha_2(p_t) := \sup_{\mu \neq \nu} \frac{W^2(\mu p_t, \nu p_t)}{W^2(\mu,\nu)} = \sup_{x\neq y} \frac{W^2(p_t(x,\cdot), p_t(y,\cdot))}{|x-y|}$$

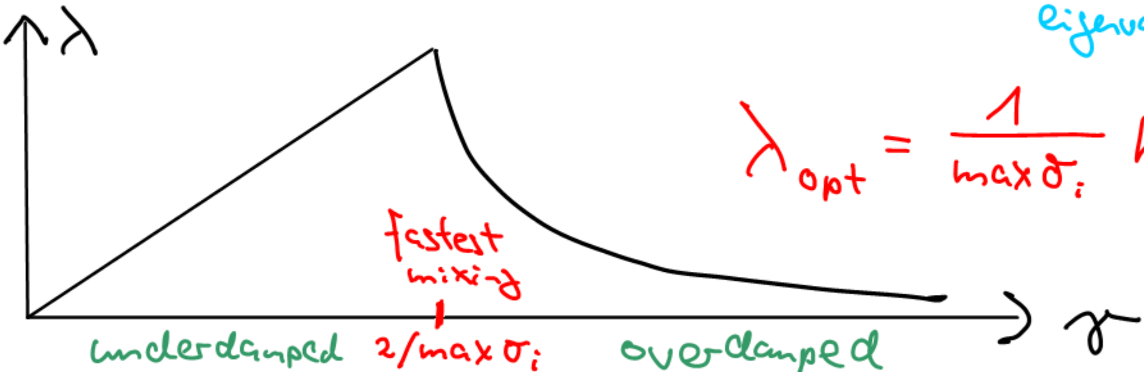
synchr. coupling is optimal

$$= \|e^{-tA}\|_{op} \sim e^{-t\lambda} \quad \text{as } t\uparrow\infty$$

$$\lambda = \inf \operatorname{Re}\operatorname{Spec}(A) = \frac{\gamma}{2}\left(1 - \sqrt{\left(1-4\gamma^{-2}/\max\sigma_i^{2}\right)^+}\right)$$

eigenvalues of C

$$\lambda_{opt} = \frac{1}{\max \sigma_i} \text{ kinetic}$$



Quantitative bounds for general potentials

- Cheng et al, Dalalyan / Riou-Durand (Bernoulli '20)

  strongly convex case, synchronous coupling

- AE, Guillin, Zimmer (AOP '19)

  non-convex cases, reflection coupling to hyperplane

- Cao, Lu, Wang (Arxiv)

  general case (warm start), space-time Poincaré inequality à la Armstrong-Mourrat

b) HAMILTONIAN MONTE CARLO (HMC)

See talk by Nisheeth Vishnoi this afternoon / night

Discussion on Zoom: Tomorrow 8.30-9.20 am

Meeting-ID: 974 2684 7158   Password: 988212

https://uni-bonn.zoom.us/j/97426847158?pwd=YXFkejg4OUxBKzVuOGVmL2JuOTBzUT09