Quantitative bounds for Markov chain based Monte Carlo methods in high dimensions

Andreas Eberle

September 18, 2012

- Metropolis-Hastings algorithms in $\mathbb{R}^d$, d large
- Sequential Monte Carlo Samplers in high dimensions

- Coupling approach $\leftrightsquigarrow$ Convergence in Wasserstein distance
- Functional inequalities $\leftrightsquigarrow$ Convergence in L^p sense

1 INTRODUCTION

$$U(x) = \frac{1}{2}|x|^2 + V(x)\,, \qquad x \in \mathbb{R}^d, \qquad V \in C^4(\mathbb{R}^d),$$

$$\mu(dx) = \frac{1}{Z}\, e^{-U(x)}\, \lambda^d(dx) = \frac{(2\pi)^{d/2}}{Z}\, e^{-V(x)}\, \gamma^d(dx),$$

$\gamma_d = N(0, I_d)$ standard normal distribution in $\mathbb{R}^d$.

AIM :

- Approximate Sampling and MC integral estimation w.r.t. μ.
- Rigorous error and complexity estimates, $d \to \infty$.

A PROTOTYPICAL EXAMPLE: TRANSITION PATH SAMPLING

$$dY_t \;=\; dB_t - \nabla H(Y_t)\,dt\,, \qquad Y_0 = y_0 \in \mathbb{R}^n,$$

μ = conditional distribution on $C([0,T],\mathbb{R}^n)$ of $(Y_t)_{t\in[0,T]}$ given $Y_T = y_T$.

By Girsanov's Theorem:

$$\mu(dy) \;=\; Z^{-1}\,\exp(-V(y))\,\gamma(dy),$$

γ = distribution of Brownian bridge from y_0 to y_T,

$$V(y) \;=\; \int_0^T \left(\frac{1}{2}\Delta H(y_t) + |\nabla H(y_t)|^2\right) dt.$$

Finite dimensional approx. via Karhunen-Loève or Wiener-Lévy expansion:

$$\gamma(dy) \;\to\; \gamma^d(dx)\,, \qquad V(y) \;\to\; V_d(x) \qquad \rightsquigarrow \text{ setup above}$$

POSSIBLE APPROACHES:

- *Metropolis-Hastings*, Gibbs Sampler
- Parallel Tempering, Equi-Energy Sampler
- *Sequential Monte Carlo Sampler*

2 Metropolis-Hastings methods with Gaussian proposals

MARKOV CHAIN MONTE CARLO APPROACH

- Simulate an ergodic Markov process (X_n) with stationary distribution μ.
- n large: $P \circ X_n^{-1} \approx \mu$
- Continuous time: *(over-damped) Langevin diffusion*

$$dX_t = -\frac{1}{2} X_t \, dt - \frac{1}{2} \nabla V(X_t) \, dt + dB_t$$

- Discrete time: *Metropolis-Hastings Algorithms*

METROPOLIS-HASTINGS ALGORITHM

(Metropolis et al 1953, Hastings 1970)

$\mu(x) := Z^{-1}\exp(-U(x))$ density of μ w.r.t. λ^d,

$p(x,y)$ stochastic kernel on $\mathbb{R}^d$ proposal density, > 0,

ALGORITHM

1. Choose an initial state X_0.
2. For $n := 0, 1, 2, \ldots$ do
   - Sample $Y_n \sim p(X_n, y)dy,\ U_n \sim \text{Unif}(0,1)$ independently.
   - If $U_n < \alpha(X_n, Y_n)$ then accept the proposal and set $X_{n+1} := Y_n$; else reject the proposal and set $X_{n+1} := X_n$.

METROPOLIS-HASTINGS ACCEPTANCE PROBABILITY

$$\alpha(x,y) \; = \; \min\left(\frac{\mu(y)p(y,x)}{\mu(x)p(x,y)},1\right) \; = \; \exp\left(-G(x,y)^+\right), \quad x,y\in\mathbb{R}^d,$$

$$G(x,y) = \log\frac{\mu(x)p(x,y)}{\mu(y)p(y,x)} = U(y) - U(x) + \log\frac{p(x,y)}{p(y,x)} = V(y) - V(x) + \log\frac{\gamma^d(x)p(x,y)}{\gamma^d(y)p(y,x)}$$

- (X_n) is a time-homogeneous Markov chain with transition kernel

$$q(x,dy) = \alpha(x,y)p(x,y)dy + q(x)\delta_x(dy), \;\; q(x) = 1 - q(x,\mathbb{R}^d\setminus\{x\}).$$

- *Detailed Balance:*

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

PROPOSAL DISTRIBUTIONS FOR METROPOLIS-HASTINGS

$x \mapsto Y_h(x)$ proposed move, $h > 0$ step size,

$$p_h(x, dy) = P[Y_h(x) \in dy] \text{ proposal distribution,}$$
$$\alpha_h(x, y) = \exp(-G_h(x, y)^+) \text{ acceptance probability.}$$

- **Random Walk Proposals** ($\rightsquigarrow$ Random **W**alk **M**etropolis)

$$\begin{aligned} Y_h(x) &= x + \sqrt{h} \cdot Z, \qquad Z \sim \gamma^d, \\ p_h(x, dy) &= N(x, h \cdot I_d), \\ G_h(x, y) &= U(y) - U(x). \end{aligned}$$

- **Ornstein-Uhlenbeck Proposals** ($\rightsquigarrow$ Preconditioned **RWM**)

$$\begin{aligned} Y_h(x) &= \left(1 - \frac{h}{2}\right) x + \sqrt{h - \frac{h^2}{4}} \cdot Z, \qquad Z \sim \gamma^d, \\ p_h(x, dy) &= N((1 - h/2)x, (h - h^2/4) \cdot I_d), \qquad \text{det. balance w.r.t. } \gamma^d \\ G_h(x, y) &= V(y) - V(x). \end{aligned}$$

- **Euler Proposals** ($\rightsquigarrow$ **M**etropolis **A**djusted **L**angevin **A**lgorithm)

$$Y_h(x) \;=\; \left(1-\frac{h}{2}\right)x-\frac{h}{2}\nabla V(x)+\sqrt{h}\cdot Z, \qquad Z\sim\gamma^d.$$

(Euler step for Langevin equation $dX_t = -\frac{1}{2}X_t\,dt - \frac{1}{2}\nabla V(X_t)\,dt + dB_t$)

$$\begin{aligned} p_h(x,dy) \;&=\; N((1-\frac{h}{2})x-\frac{h}{2}\nabla V(x), h\cdot I_d),\\ G_h(x,y) \;&=\; V(y)-V(x)-(y-x)\cdot(\nabla V(y)+\nabla V(x))/2\\ &\quad +h(|\nabla U(y)|^2-|\nabla U(x)|^2)/4. \end{aligned}$$

REMARK. Even for $V\equiv 0$, γ^d is not a stationary distribution for p_h^{Euler}. Stationarity only holds asymptotically as $h\to 0$. This causes substantial problems in high dimensions.

- **Semi-implicit Euler Proposals** ($\rightsquigarrow$ Preconditioned **MALA**)

  [Beskos,Roberts, Stuart, Voss 2008]

$$Y_h(x) = \left(1-\frac{h}{2}\right)x - \frac{h}{2}\nabla V(x) + \sqrt{h-\frac{h^2}{4}}\cdot Z, \qquad Z \sim \gamma^d,$$

$$p_h(x,dy) = N((1-\frac{h}{2})x - \frac{h}{2}\nabla V(x), (h-\frac{h^2}{4})\cdot I_d) \qquad (=p_h^{OU} \text{ if } V\equiv 0\ )$$

$$\begin{aligned} G_h(x,y) &= V(y)-V(x)-(y-x)\cdot(\nabla V(y)+\nabla V(x))/2 \\ &\quad + \frac{h}{8-2h}\left((y+x)\cdot(\nabla V(y)-\nabla V(x)) + |\nabla V(y)|^2 - |\nabla V(x)|^2\right). \end{aligned}$$

KNOWN RESULTS FOR METROPOLIS-HASTINGS IN HIGH DIMENSIONS

- Scaling of acceptance probabilities and mean square jumps as $d \to \infty$
- Diffusion limits as $d \to \infty$
- Ergodicity, Geometric Ergodicity
- Quantitative bounds for mixing times, rigorous complexity estimates

Optimal Scaling and diffusion limits as $d \to \infty$

- *Roberts, Gelman, Gilks 1997*: Diffusion limit for RWM with product target, $h = O(d^{-1})$
- *Roberts, Rosenthal 1998*: Diffusion limit for MALA with product target, $h = O(d^{-1/3})$
- *Beskos, Roberts, Stuart, Voss 2008*: Preconditioned MALA applied to Transition Path Sampling, Scaling $h = O(1)$
- *Mattingly, Pillai, Stuart 2010*: Diffusion limit for RWM with non-product target, $h = O(d^{-1})$
- *Pillai, Stuart, Thiéry 2011a*: Diffusion limit for MALA with non-product target, $h = O(d^{-1/3})$
- *Pillai, Stuart, Thiéry 2011b*: Preconditioned RWM, Scaling $h = O(1)$, Diffusion limit as $h \downarrow 0$ independent of the dimension

Geometric ergodicity for MALA in $\mathbb{R}^d$ (d fixed)

- *Roberts, Tweedie 1996*: Geometric convergence holds if ∇U is globally Lipschitz but fails in general
- *Bou Rabee, van den Eijnden 2009*: Strong accuracy for truncated MALA
- *Bou Rabee, Hairer, van den Eijnden 2010*: Convergence to equilibrium for MALA at exponential rate up to term exponentially small in time step size

BOUNDS FOR MIXING TIME, COMPLEXITY

Metropolis with ball walk proposals

- *Dyer, Frieze, Kannan 1991*: $\mu = Unif(K)$, $K \subset \mathbb{R}^d$ convex
  $\Rightarrow$ Total variation mixing time is polynomial in d and $diam(K)$
- *Applegate, Kannan 1991, ... , Lovasz, Vempala 2006*: $U : K \to \mathbb{R}$ concave, $K \subset \mathbb{R}^d$ convex
  $\Rightarrow$ Total variation mixing time is polynomial in d and $diam(K)$

Langevin diffusions

- If μ is strictly log-concave, i.e.,

$$\exists\, \kappa > 0 : \quad \partial^2 U(x) \geq \kappa \cdot I_d \qquad \forall\, x \in \mathbb{R}^d$$

then Wasserstein contractivity holds:

$$\mathcal{W}(\,\text{law}(X_t)\,,\,\mu\,) \;\leq\; e^{-\kappa t}\,\mathcal{W}(\,\text{law}(X_0)\,,\,\mu\,),$$

where $\mathcal{W}(\nu, \mu) = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}[d(X, Y)]$ is L^1 Wasserstein distance.

- Bound is independent of dimension, sharp !

- Under additional conditions, a corresponding result holds for the Euler discretization.

- Extension to non log-concave measures: *A.E., Reflection coupling and Wasserstein contractivity without convexity, C.R.Acad.Sci.Paris 2011*.

- These results suggest that comparable bounds might hold for MALA, or even for Ornstein-Uhlenbeck proposals.

Metropolis-Hastings with Ornstein-Uhlenbeck proposals

- *Hairer, Stuart, Vollmer 2011:* Dimension independent contractivity in modified Wasserstein distance

Metropolis-adjusted Langevin algorithm

- No rigorous complexity estimates so far

3 Quantitative Wasserstein bounds for preconditioned MALA

A.E., Metropolis-Hastings algorithms for perturbations of Gaussian measures in high dimensions: Contraction properties and error bounds in the log-concave case, Preprint 2012.

Preconditioned MALA: Coupling of proposal distributions $p_h(x, dy)$, $x \in \mathbb{R}^d$:

$$Y_h(x) \;=\; \left(1 - \frac{h}{2}\right) x - \frac{h}{2}\,\nabla V(x) + \sqrt{h - \frac{h^2}{4}} \cdot Z, \qquad Z \sim \gamma^d,\; h > 0,$$

$\rightsquigarrow$ Coupling of MALA transition kernels $q_h(x, dy)$, $x \in \mathbb{R}^d$:

$$W_h(x) \;=\; \begin{cases} Y_h(x) & \text{if } U \le \alpha_h(x, Y_h(x)) \\ x & \text{if } U > \alpha_h(x, Y_h(x)) \end{cases}, \quad U \sim Unif(0,1) \text{ independent of } Z,$$

We fix a radius $R \in (0, \infty)$ and a norm $\| \cdot \|_- = \langle \cdot, \cdot \rangle^{1/2}$ on $\mathbb{R}^d$ such that

$$\|x\|_- \leq |x| \qquad \text{for any } x \in \mathbb{R}^d,$$

and we set

$$d_R(x, \tilde{x}) \;:=\; \min(\|x - \tilde{x}\|_-, 2R), \qquad B_R^- \;:=\; \{x \in \mathbb{R}^d \;:\; \|x\|_- < R\}.$$

EXAMPLE: Transition Path Sampling

- $|x|_{\mathbb{R}^d}$ is finite dimensional projection of Cameron-Martin norm/ H^1 norm

$$|x|_{CM} \;=\; \left(\int_0^T \left| \frac{dx}{dt} \right|^2 dt \right)^{1/2}.$$

- $\|x\|_-$ is finite dimensional approximation of H^α norm, $\alpha \in (0, 1/2)$.

ASSUMPTIONS:

(A1) There exist finite constants $C_n, p_n \in [0,\infty)$ such that

$$|(\partial^n_{\xi_1,\ldots,\xi_n} V)(x)| \;\le\; C_n \max(1, \|x\|_-)^{p_n} \|\xi_1\|_- \cdots \|\xi_n\|_-$$

for any $x \in \mathbb{R}^d$, $\xi_1, \ldots, \xi_n \in \mathbb{R}^d$, and $n = 2, 3, 4$.

(A2) There exists a constant $K > 0$ such that

$$\langle \eta, \nabla^2 U(x) \cdot \eta\rangle \;\ge\; K\,\langle \eta, \eta\rangle \qquad \forall\, x \in B_R^-,\; \eta \in \mathbb{R}^d.$$

THEOREM (AE 2012). If (A1) and (A2) are satisfied then

$$E\left[\|W_h(x) - W_h(\tilde{x})\|_-\right] \;\le\; \left(1 - \frac{1}{2}Kh + C(R)h^{3/2}\right)\|x - \tilde{x}\|_- \quad \forall\, x, \tilde{x} \in B_R^-,\; h \in (0,1)$$

with an explicit constant $C(R) \in (0,\infty)$ that does depend on the dimension only through the moments

$$m_k \;:=\; \int_{\mathbb{R}^d} \|x\|_-^k \,\gamma^d(dx)\,, \qquad k \in \mathbb{N}.$$

REMARKS.

- $h \downarrow 0$: approaches optimal contraction rate $1 - Kh/2$
- $h^{-1} = O(R^q)$: contraction rate $\geq 1 - Kh/4$
- For Ornstein-Uhlenbeck proposals, the contraction term is $O(h)$ instead of $O(h^{3/2})$
- The corresponding bounds for standard MALA and RWM are dimension dependent.

CONTRACTIVITY IN WASSERSTEIN DISTANCE

q_h = transition kernel of preconditioned MALA

COROLLARY. If (A1) and (A2) are satisfied, then there exist explicit constants $C, D, q \in (0, \infty)$ that do not depend on the dimension such that

$$\mathcal{W}_{2R}(\pi q_h^n, \nu q_h^n) \le (1 - \frac{K}{4} h)^n \mathcal{W}_{2R}(\pi, \nu) + DR \exp(-KR^2/8) \, nh$$

for any $n \in \mathbb{N}, h, R \in (0, \infty)$ such that $h^{-1} \ge C(1 + R)^q$, and for any initial distributions π, ν with support in B_R^-.

Approximation of quasi-stationary distribution

$\mu_R(A) := \mu(A|B_R^-)$.

COROLLARY. If (A1) and (A2) are satisfied, then there exist explicit constants $C, \bar{D}, q \in (0,\infty)$ that do not depend on the dimension such that

$$\mathcal{W}_{2R}(\nu q_h^n, \mu_R) \;\leq\; 58\, R(1 - \frac{K}{4} h)^n \;+\; \bar{D} R \exp(-KR^2/33)\, nh$$

whenever $h^{-1} \geq C(1+R)^q$ and the initial distribution ν has support in $B_{R/2}^-$.

REMARK.

- To attain a given error bound ε for the Wasserstein distance, h has to be chosen sufficiently small (roughly $h^{-1} \sim O((\log \varepsilon^{-1})^{q/2})$, but in a dimension-independent way!
- There is a best possible error bound $\varepsilon > 0$ that can be attained, since after a long time the chain will exit from the metastable state B_R^-.

KEY INGREDIENTS IN PROOF:

Dimension independent bounds that quantify

- Rejection probabilities
- Dependence of rejection event on the current state

THEOREM. Suppose that Assumption (A1) is satisfied. Then there exist polynomials $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}_+$ of degree $\max(p_3 + 3, 2p_2 + 2)$ and $\mathcal{Q} : \mathbb{R}^2 \to \mathbb{R}_+$ of degree $\max(p_4 + 2, p_3 + p_2 + 2, 3p_2 + 1)$ such that

$$E[1 - \alpha_h(x, Y_h(x))] \;\leq\; E[G_h(x, Y_h(x))^+] \;\leq\; \mathcal{P}(\|x\|_-, \|\nabla U(x)\|_-) \cdot h^{3/2}$$

$$E\left[\|\nabla_x G_h(x, Y_h(x))\|_+\right] \;\leq\; \mathcal{Q}(\|x\|_-, \|\nabla U(x)\|_-) \cdot h^{3/2}$$

for all $x \in \mathbb{R}^d$, $h \in (0, 2)$, where

$$\|\eta\|_+ \;:=\; \sup\{\xi \cdot \eta \;:\; \|\xi\|_- \leq 1\}.$$

REMARK.

- The polynomials $\mathcal{P}$ and $\mathcal{Q}$ are explicit. They depend only on the values $C_2, C_3, C_4, p_2, p_3, p_4$ and on the moments

$$m_k = E[\|Z\|_-^k]$$

  but they do not depend on the dimension d.

- For MALA with explicit Euler proposals, corresponding estimates hold with m_k replaced by $\tilde{m}_k = E[|Z|^k]$. Note, however, that $\tilde{m}_k \to \infty$ as $d \to \infty$.

4 Sequential MCMC, SMC Sampler

A.E., C. Marinelli, Quantitative approximations of evolving probability measures and sequential MCMC methods, PTRF 2012, Online First.

$$\mu_t(dx) \;=\; Z_t^{-1} \exp\left(-U_t(x)\right)\; \gamma(dx), \qquad t \in [0, t_0], \qquad \mu_{t_0} = \mu$$

probability measures on state space S.

$$H_t(x) \;:=\; -\frac{\partial}{\partial t} \log \frac{d\mu_t}{d\gamma}(x) \;=\; \frac{\partial}{\partial t} U_t(x) - \left\langle \frac{\partial}{\partial t} U_t,\, \mu_t \right\rangle.$$

$$\mu_t(dx) \;\propto\; \exp\left(-\int_0^t H_s(x)\, ds\right)\; \gamma(dx)$$

Let $\mathcal{L}_t,\; t \geq 0,$ be generators of a time-inhomogeneous Markov process on S such that $\mathcal{L}_t$ satisfies the detailed balance condition w.r.t. μ_t. In particular,

$$\mathcal{L}_t^* \mu_t \;=\; 0 \quad \text{(infinitesimal stationarity).}$$

Fix constants $\lambda_t \geq 0$.

SMC SAMPLER IN CONTINUOUS TIME

$X_t^N = (X_{t,1}^N, \ldots, X_{t,N}^N)$ Markov process on S^N with generator

$$\begin{aligned}\mathcal{L}_t^N \varphi(x_1, \ldots, x_N) &= \lambda_t \sum_{i=1}^N \mathcal{L}_t^{(i)} \varphi(x_1, \ldots, x_N) \\ &+ \frac{1}{N} \sum_{i,j=1}^N (H_t(x_i) - H_t(x_j))^+ \cdot \left(\varphi(x^{i \to j}) - \varphi(x)\right),\end{aligned}$$

$\mathcal{L}_t^{(i)}$ action of $\mathcal{L}_t$ on i th component.

- Independent Markov chain moves with generator $\lambda_t \cdot \mathcal{L}_t$
- $X_{t,i}^N$ replaced by $X_{t,j}^N$ with rate $\frac{1}{N}(H_t(X_{t,i}^N) - H_t(X_{t,j}^N))^+$

ESTIMATORS FOR μ_t: $X_{0,i}^N$ i.i.d. $\sim \mu_0$

$$\eta_t^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_{t,i}^N}, \qquad \nu_t^N := \exp\left(-\int_0^t \langle H_s, \eta_s^N\rangle \, ds\right) \eta_t^N .$$

PERFORMANCE IN HIGH DIMENSIONS ?

Possible test cases:

1. Product models
2. Models with dimension-independent global mixing properties
3. Disconnected unions of such models
4. Models with a disconnectivity tree structure
5. Models with a phase transition
6. Disordered systems

5 Quantitative error bounds and dimension dependence

$$\varepsilon_t^{N,p} := \sup \left\{ \mathbb{E}\left[\left| \langle f, \nu_s^N \rangle - \langle f, \mu_s \rangle \right|^2 \right] \,:\, s \in [0,t],\ \|f\|_{L^p(\mu_s)} \le 1 \right\},\ p \in [2,\infty].$$

GOAL:

- Bounds for $\varepsilon_t^{N,p}$ for a **fixed** number N of replicas.
- Explicit dependence on the dimension for test models.

ERROR BOUNDS AND DIMENSION DEPENDENCE UNDER GLOBAL MIXING CONDITIONS

Fix $t_0 \in (0, \infty)$ (length of time interval), $p \in (6, \infty)$, $q \in (p, \infty)$, and let

$$\omega = \sup_{t \in [0,t_0]} \operatorname{osc}(H_t) \; ; \quad K_t = \int_0^t \|H_s\|_{L^q(\mu_s)} \, ds$$

$$C_t = \sup_{\langle f, \mu_t \rangle = 0} \frac{\int f^2 \, d\mu_t}{\mathcal{E}_t(f,f)}$$ Poincaré constant (inverse spectral gap)

$$\gamma_t = \sup_{\langle f^2, \mu_t \rangle = 1} \frac{\int f^2 \log |f| \, d\mu_t}{\mathcal{E}_t(f,f)}$$ Log-Sobolev constant

where

$$\mathcal{E}_t(f,f) = -(f, \mathcal{L}_t f)_{L^2(\mu_t)}$$

is the *Dirichlet form* of $\mathcal{L}_t$ on $L^2(\mu_t)$.

THEOREM (A.E., C. Marinelli 2012) Suppose that

$$\begin{aligned} N \quad &\geq \quad 40 \cdot \max\left(K_{t_0},\, 1\right), \qquad \text{and} \\ \lambda_t \quad &\geq \quad \omega \cdot \max\left(\frac{p}{4} \cdot \left(1 + t \cdot \frac{p+3}{4}\right) \cdot C_t \;,\; a(p,q) \cdot \gamma_t\right) \qquad \forall t \in [0, t_0]. \end{aligned}$$

Then

$$\varepsilon_t^{N,p} \;\leq\; \frac{2 + 8\,K_t}{N} \cdot \left(1 + \frac{16\,K_t}{N}\right) \qquad \forall\; t \in [0, t_0].$$

Here $a(p,q)$ is an explicit constant depending only on p and q.

EXAMPLE 1: Product measures

$$S = \prod_{k=1}^{d} S_k \,, \qquad \mu_t = \bigotimes_{k=1}^{d} \mu_t^{(k)}$$

$$\Rightarrow \quad H_t(x) = -\frac{d}{dt} \log \mu_t(x) = \sum_{k=1}^{d} H_t^{(k)}(x_k)$$

$$\Rightarrow \quad \omega = \sup_{t,x,y} |H_t(x) - H_t(y)| \leq \sum_{k=1}^{d} \omega^{(k)} \,.$$

$$\mathcal{L}_t(x,y) = \sum_{k=1}^{d} \mathcal{L}_t^{(k)}(x,y) \qquad \text{product dynamics}$$

$$\Rightarrow \quad C_t = \max_k C_t^{(k)}, \quad \gamma_t = \max_k \gamma_t^{(k)} \,.$$

EXAMPLE 1: Product measures

$$S = \prod_{k=1}^{d} S_k\,, \qquad \mu_t = \bigotimes_{k=1}^{d} \mu_t^{(k)}$$

Assumption:

$$\omega^{(k)} \leq 1 \;\; \forall\, k, \qquad C_t^{(k)},\ \gamma_t^{(k)} \text{independent of } k.$$

$\Rightarrow$ $\omega = O(d)$, $C_t = O(1)$, $\gamma_t = O(1)$

$\Rightarrow$ $N = O(d^{1/2})$ and $\lambda_s = O(d)$ are sufficient for a given precision

$\Rightarrow$ total effort of order $O(d^3)$ (resp. $O(d^{2.5})$) is sufficient

EXAMPLE 1: Product measures

Bound independent of d holds provided there are

- $O(d)$ resampling steps
- $O(d)$ MCMC steps between each resampling step
- $O(d^{1/2})$ particles

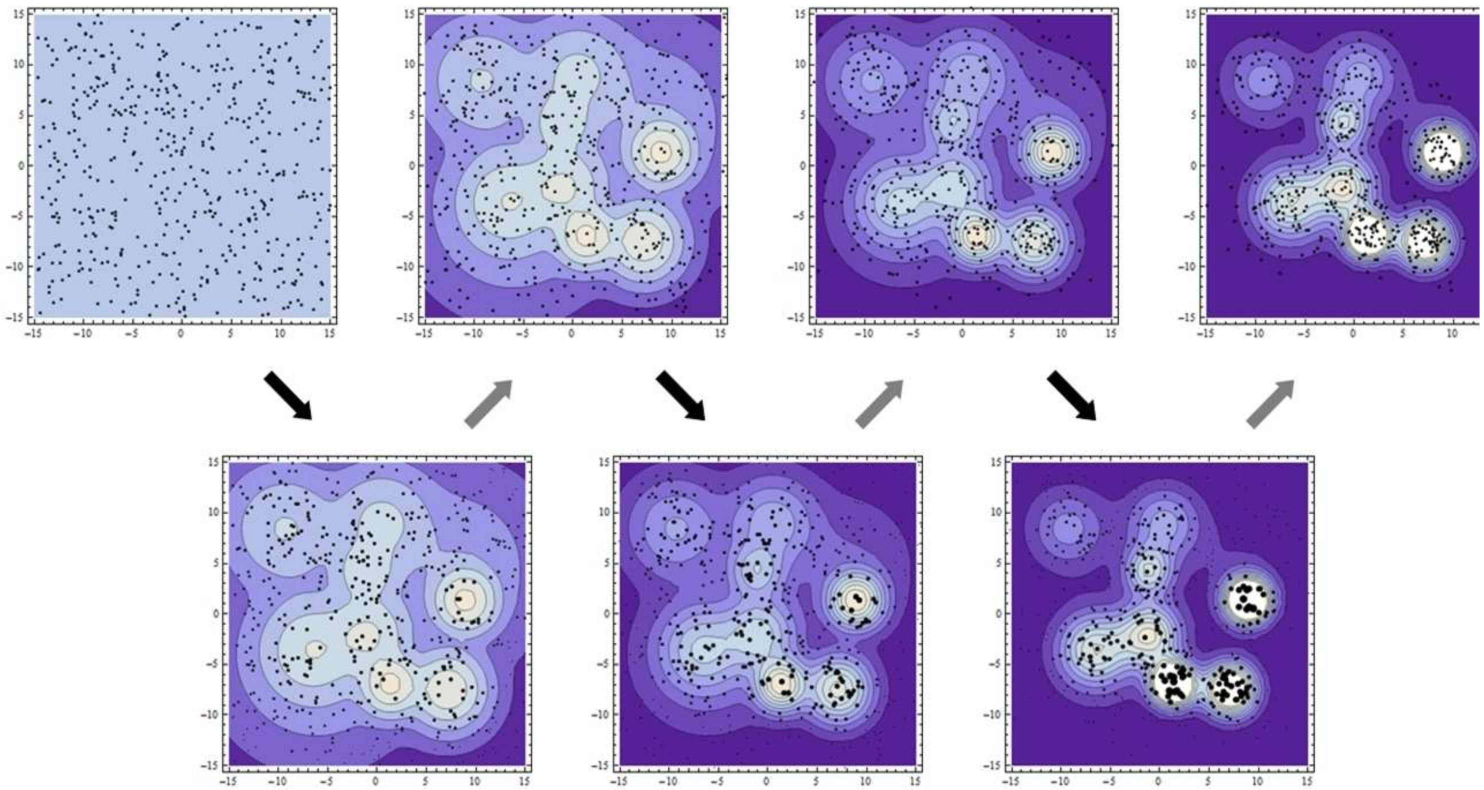
EXAMPLE 2: Log Sobolev and spectral gap independent of the dimension

$\rightsquigarrow$ similar bounds as in Example 1.

REMARK. [Beskos, Crisan, Jasra, Whiteley 2011]

- In the product case, $O(1)$ resampling steps are sufficient.
- This holds true because strong mixing properties make up even for huge errors and degeneracy due to resampling.
- One can not expect equally strong results in more general scenarios.

ERROR BOUNDS AND DIMENSION DEPENDENCE WITHOUT GLOBAL MIXING



NON-ASYMPTOTIC BOUNDS FOR DISCONNECTED UNIONS

$S = \bigcup S_i$ disjoint decomposition of state space. Suppose that

$$\begin{aligned} \mathcal{L}_t(x,y) &= 0 \;\forall\, t \geq 0,\; x \in S_i,\; y \in S_j\; (i \neq j), \text{ and let} \\ \mu_t^i &= \mu_t(\,\cdot\,|S_i), \qquad \|f\|^{\sim}_{L^p(\mu_t)} := \max_i \|f\|_{L^p(\mu_t^i)}\;, \\ \tilde{\varepsilon}_t^{N,p} &:= \sup\left\{\mathbb{E}\left[\left|\langle f, \nu_s^N\rangle - \langle f, \mu_s\rangle\right|^2\right] \,:\, s \in [0,t],\, \|f\|^{\sim}_{L^p(\mu_s)} \leq 1\right\}. \end{aligned}$$

THEOREM. Suppose conditions as above hold with C_t, γ_t replaced by

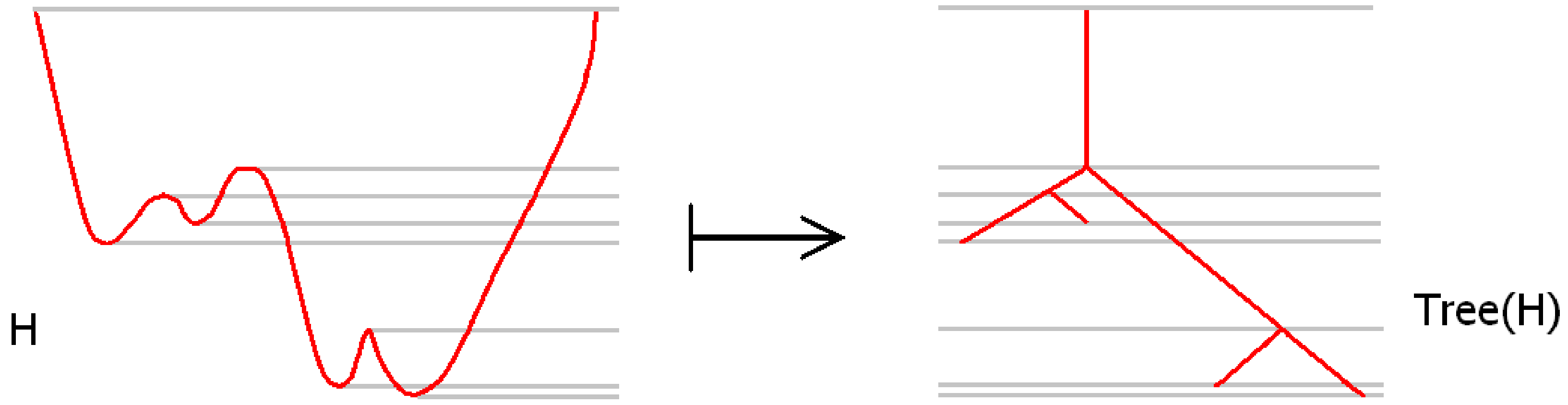
$$\tilde{C}_t = \max_i C_t^i\,, \qquad \tilde{\gamma}_t = \max_i \gamma_t^i\,.$$

Then

$$\tilde{\varepsilon}_t^{N,p} \leq \frac{2 + 8\,K_t\,\tilde{M}_t^2}{N} \cdot \left(1 + \frac{16\,\tilde{K}_t\,\tilde{M}_t^2}{N}\right)$$

where

$$\tilde{M}_t = \max_i \sup_{0 \leq r \leq s \leq t} \frac{\mu_s(S_i)}{\mu_r(S_i)}\,.$$



EXAMPLE 3: Disjoint union of i.i.d. product models

Dimension dependence as above holds in particular if

$$\liminf_{d\to\infty} \min_i \mu_0(S_i) \; > \; 0.$$

EXAMPLE 4: Disconnectivity tree

see talk of Nikolaus Schweizer