Quantitative convergence properties of simple slice sampling

Viacheslav Natarovskii & Daniel Rudolf

Problem

For $G \subseteq \mathbb{R}^d$ and $\varrho\colon G \to [0,\infty)$ we want to sample with respect to

$$\pi(A) = \frac{\int_A \varrho(x)\mathrm{d}x}{\int_G \varrho(x)\mathrm{d}x}, \qquad A \subseteq G,$$

by the simple slice sampler.

Qualitative convergence results, see e.g. [1].

Simple slice sampler

The simple slice sampler generates a Markov chain $(X_n)_{n\in\mathbb{N}}$ as follows: Assume that $X_n = x_n$, then the next state $X_{n+1} = x_{n+1}$ is generated by the following steps:

1. Draw $T_n \sim$ Unif$[0, \varrho(x_n)]$, call the result t_n;
2. Draw $X_{n+1} \sim$ Unif$(G(t_n))$, call the result x_{n+1}.

Notations and basics

- The level set of ϱ (of level t):
$$G(t) := \{x \in G | \varrho(x) \geq t\}, \qquad t \in [0, \|\varrho\|_\infty];$$
- The transition kernel of $(X_n)_{n\in\mathbb{N}}$:
$$U(x,A) = \frac{1}{\varrho(x)} \int_0^{\varrho(x)} \frac{\mathrm{vol}_d(A \cap G(t))}{\mathrm{vol}_d(G(t))}\mathrm{d}t, \qquad A \subseteq G;$$
- The transition kernel of $(T_n)_{n\in\mathbb{N}}$:
$$Q(t,B) = \frac{1}{\mathrm{vol}_d(G(t))} \int_{G(t)} \frac{\mathrm{vol}_1\left(B \cap [0,\varrho(x)]\right)}{\varrho(x)}\mathrm{d}x, \qquad B \subseteq (0,\infty);$$
- The probability measure μ on $(0,\infty)$:
$$\mu(B) := \frac{\int_B \mathrm{vol}_d(G(s))\mathrm{d}s}{\int_0^\infty \mathrm{vol}_d(G(t))\mathrm{d}t}, \qquad B \subseteq (0,\infty);$$
- The Markov operators U and Q:
$$Uf(x) := \int_G f(y)U(x,\mathrm{d}y), \qquad Qg(t) := \int_0^\infty g(s)Q(t,\mathrm{d}s).$$

Spectral gap

The spectral gaps of U and Q are defined by

$$\mathrm{gap}(U) := 1 - \|U - \mathrm{E}_\pi\|_{L_2(\pi)\to L_2(\pi)}, \quad \mathrm{gap}(Q) := 1 - \left\|Q - \mathrm{E}_\mu\right\|_{L_2(\mu)\to L_2(\mu)}.$$

If $\mathrm{gap}(U) > 0$, then

- $\|\nu U^n(\cdot) - \pi(\cdot)\|_{\mathrm{tv}} \leq (1-\mathrm{gap}(U))^n \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\pi} - 1\right\|_2$;
- a CLT holds for the ergodic average for all $f \in L_2(\pi)$, see [2];
- there is a mean squared error bound for the ergodic average for $f \in L_p(\pi)$, $p > 2$, see [3].

Main tool: Wasserstein metric

- Wasserstein distance between two measures on G:
$$W(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{G\times G} |x-y|\mathrm{d}\gamma(x,y),$$
where $\Gamma(\mu,\nu)$ is the set of all couplings of μ and ν;
- By [4, Proposition 30] we have
$$W(U(x,\cdot),U(\widetilde{x},\cdot)) \leq \alpha\,|\widetilde{x}-x|, \quad \forall x,\widetilde{x} \in G \quad \Longrightarrow \quad \mathrm{gap}(U) \geq 1-\alpha.$$

Main results

Theorem 1 (Natarovskii, Rudolf, 2018). Let $\varrho(x) = e^{-\varphi(|x|)}$, where $\varphi : [0,\infty) \to [0,\infty)$ satisfies the following conditions:

1. $\varphi \in C^1[0,\infty)$,
2. $\varphi'(0) \geq 0$ and $\varphi'(s) > 0$ for all $s > 0$,
3. $\varphi'(s)$ is non-decreasing.

Then

$$W(U(x,\cdot),U(\widetilde{x},\cdot)) \leq \frac{d}{d+1}|\widetilde{x}-x|, \qquad \forall x,\widetilde{x} \in G.$$

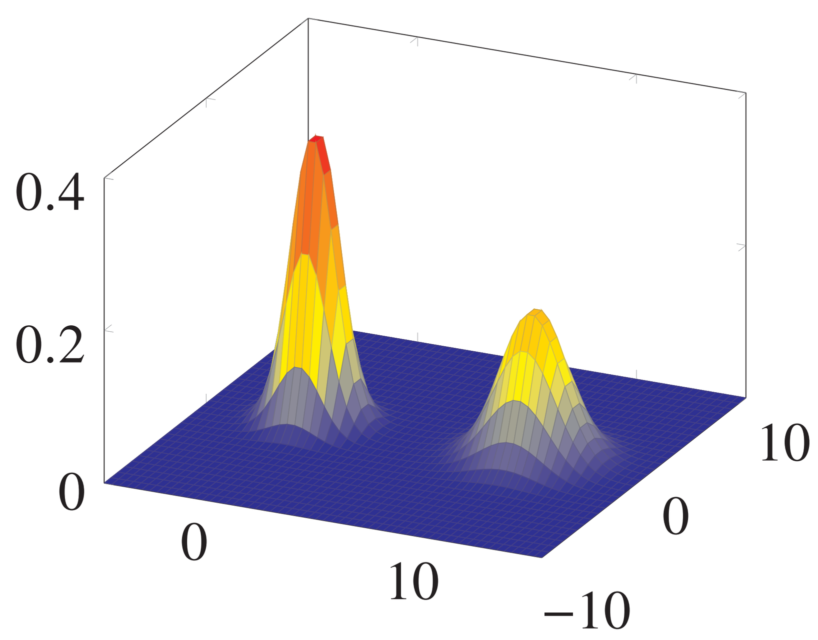
Lemma 2 (Natarovskii, Rudolf, 2018). $\mathrm{gap}(U) = \mathrm{gap}(Q)$.

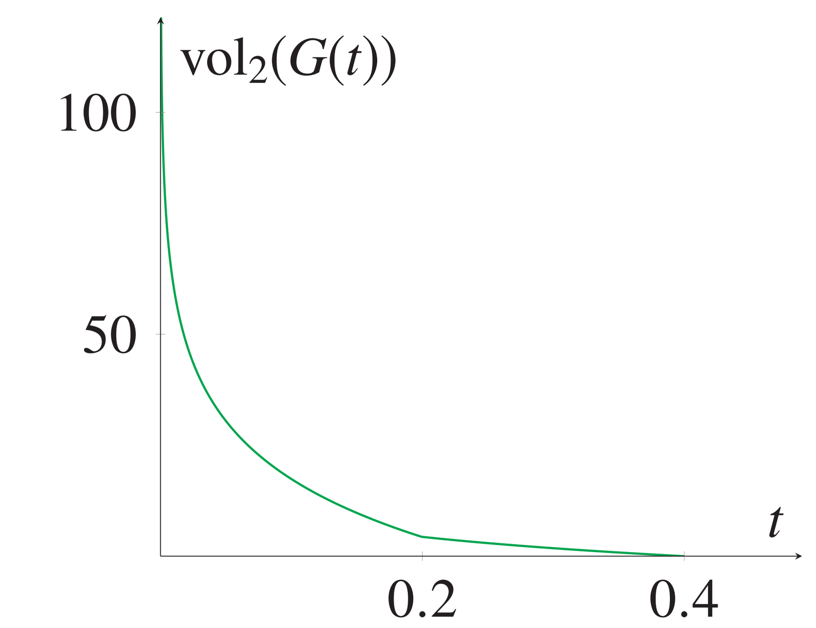
Theorem 3 (Natarovskii, Rudolf, 2018). Let $\varrho(x)$ be such that for some $k \in \mathbb{N}$ the function $\varphi(s) := -\log\left(\inf\left\{t : \mathrm{vol}_d(G(t)) \leq \mathrm{vol}_k(B_k)s^k\right\}\right)$ satisfies conditions **1.-3.** from Theorem 2. Then

$$\mathrm{gap}(U) \geq 1 - \frac{k}{k+1} = \frac{1}{k+1}.$$

Illustration of the technique

Let $\varrho(x)$ be the maximum of two Gaussian densities:





Examples of densities with the same level-set function as $\varrho(x)$:

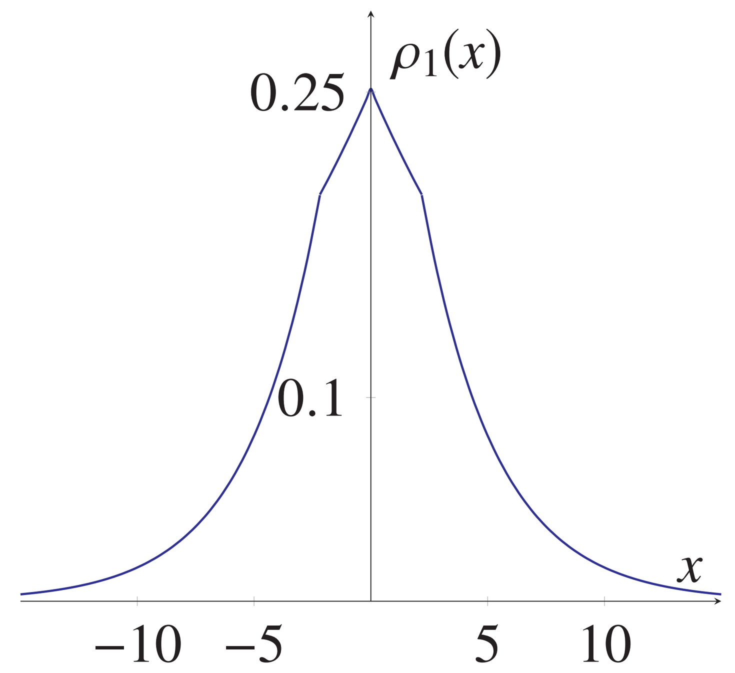


Figure: Density $\varrho_1(x)$ in $\mathbb{R}^1$

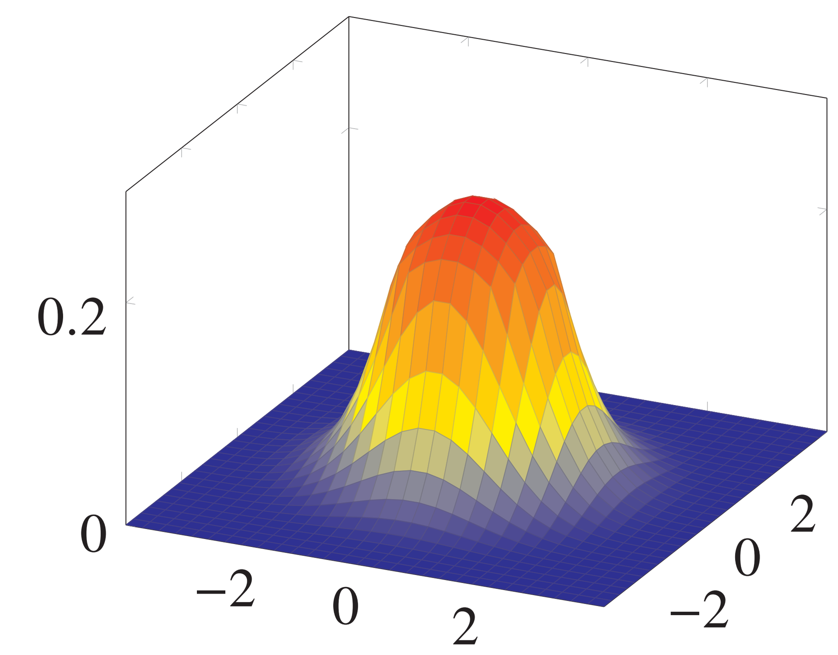


Figure: Density $\varrho_2(x)$ in $\mathbb{R}^2$

Future research

- Weakening assumptions of Theorem 1;
- Spectral gap for elliptical slice sampling, see [5];
- Spectral gap for hybrid slice sampling, see [6].

References

1. G. Roberts and J. Rosenthal, *Convergence of slice sampler Markov chains*, J. R. Stat. Soc. Ser. B Stat. Methodol. 61 (1999), no. 3, 643–660.
2. Kipnis, Varadhan, *Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions*, Communications in Mathematical Physics 104 (1986), no. 1, 1–19.
3. D. Rudolf, *Explicit error bounds for Markov chain Monte Carlo*, Dissertationes Math. 485 (2012), 93 pp.
4. Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. 256 (2009), no. 3, 810–864.
5. I. Murray, R. Adams, and D. MacKay, *Elliptical slice sampling*, Journal of Machine Learning Research: W&CP 9 (2010), 541–548.
6. K. Łatuszyński and D. Rudolf, *Convergence of hybrid slice sampling via spectral gap*, Arxiv (2014), submitted.

Institute for Mathematical Stochastics, University of Göttingen vnataro@uni-goettingen.de