# STOCHASTIC PROXIMAL LANGEVIN ALGORITHM

Adil Salim
Joint work with Dmitry Kovalev and Peter Richtárik

# SAMPLING PROBLEM

$$\mu^\star(\mathrm{d}x) \propto \exp(-U(x))\mathrm{d}x,$$

**where** $U : \mathbb{R}^d \to \mathbb{R}$ **convex.**

# LANGEVIN MONTE CARLO (LMC)

Assume $U$ smooth, $W^k$ i.i.d standard gaussian and $\gamma > 0$,

$$x^{k+1} = \underbrace{x^k - \gamma \nabla U(x^k)}_{\text{Gradient descent}} + \underbrace{\sqrt{2\gamma} W^{k+1}}_{\text{Gaussian noise}}.$$

Typical non asymptotic result: $\mathbf{KL}(\overline{\mu^k} | \mu^\star) = \mathcal{O}(1/\sqrt{k})$.

# FIRST INTUITION FOR LMC

LMC can be seen as a Euler discretization of the Langevin equation:

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t\,.$$

Non asymptotic results using this intuition in [Dalalyan 2017], [Durmus Moulines 2017].

# SECOND INTUITION FOR LMC

LMC can be seen as an (inexact) Gradient Descent for:

$$\mu^\star = \operatorname{argmin} \int U \mathrm{d}\mu(x) + \int \mu(x) \log(\mu(x)) \mathrm{d}x$$

$$\mu^\star = \operatorname{argmin} \mathrm{KL}(\mu \,|\, \mu^\star)\,.$$

Non asymptotic results using this intuition (+ extensions of LMC beyond GD) in [Durmus *et al.* 2018], [Wibisono 2018], [Bernton 2018].

# CONTRIBUTION: STOCHASTIC PROXIMAL LANGEVIN

**Case 1:** $U(x) = E_\xi(g(x,\xi))$

**Nonsmooth**

$$x^{k+1} = \underbrace{\mathbf{prox}_{\gamma g(\cdot,\xi^{k+1})}(x^k)}_{\text{Stochastic Prox}} + \sqrt{2\gamma}W^{k+1}.$$

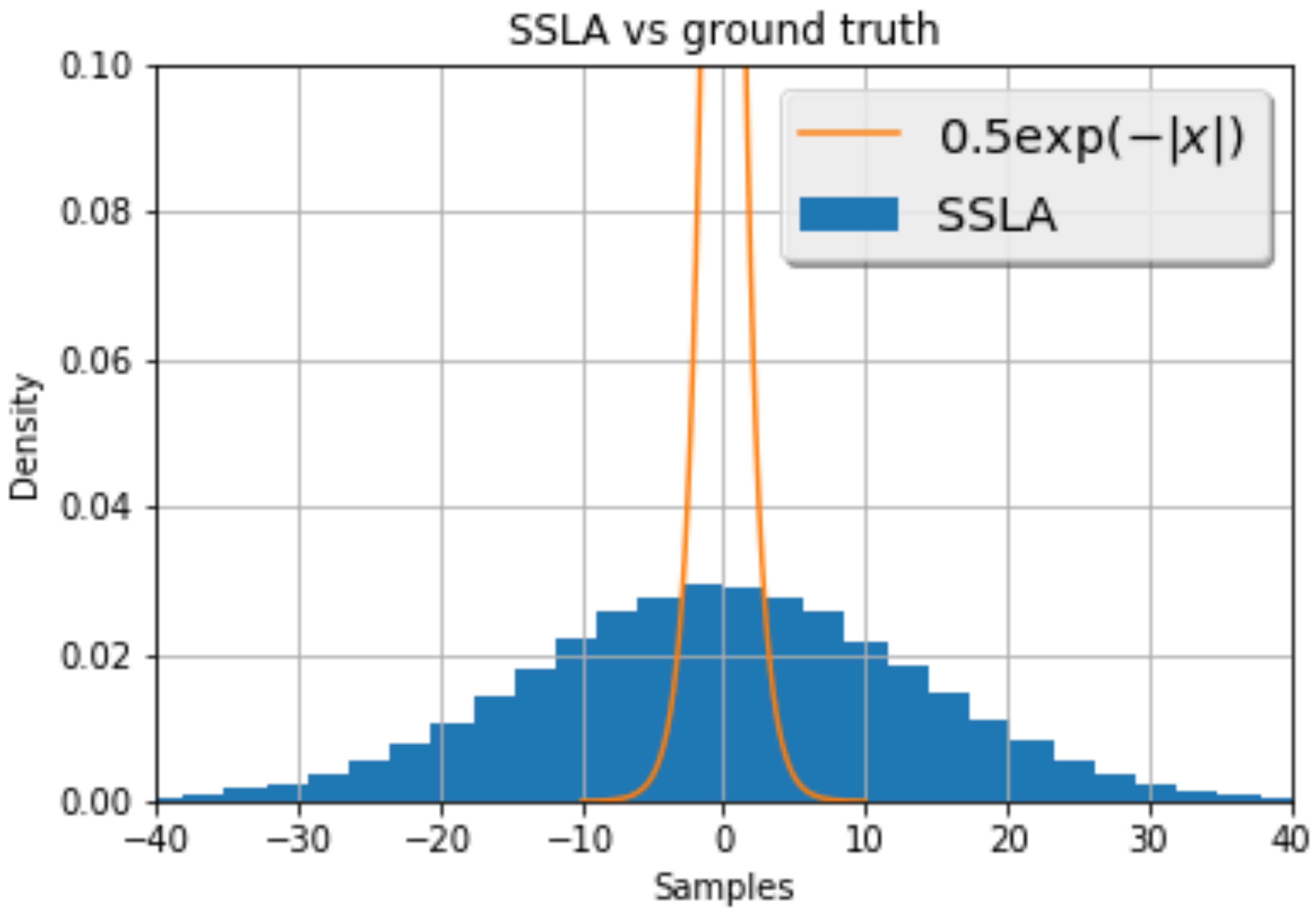
# CONTRIBUTION: STOCHASTIC PROXIMAL LANGEVIN

**Case 2:** $U(x) = \underbrace{E_\xi(f(x,\xi))}_{\text{Smooth}} + \underbrace{\sum_i E_\xi(g_i(x,\xi))}_{\text{Nonsmooth}}$

See our Poster #161.

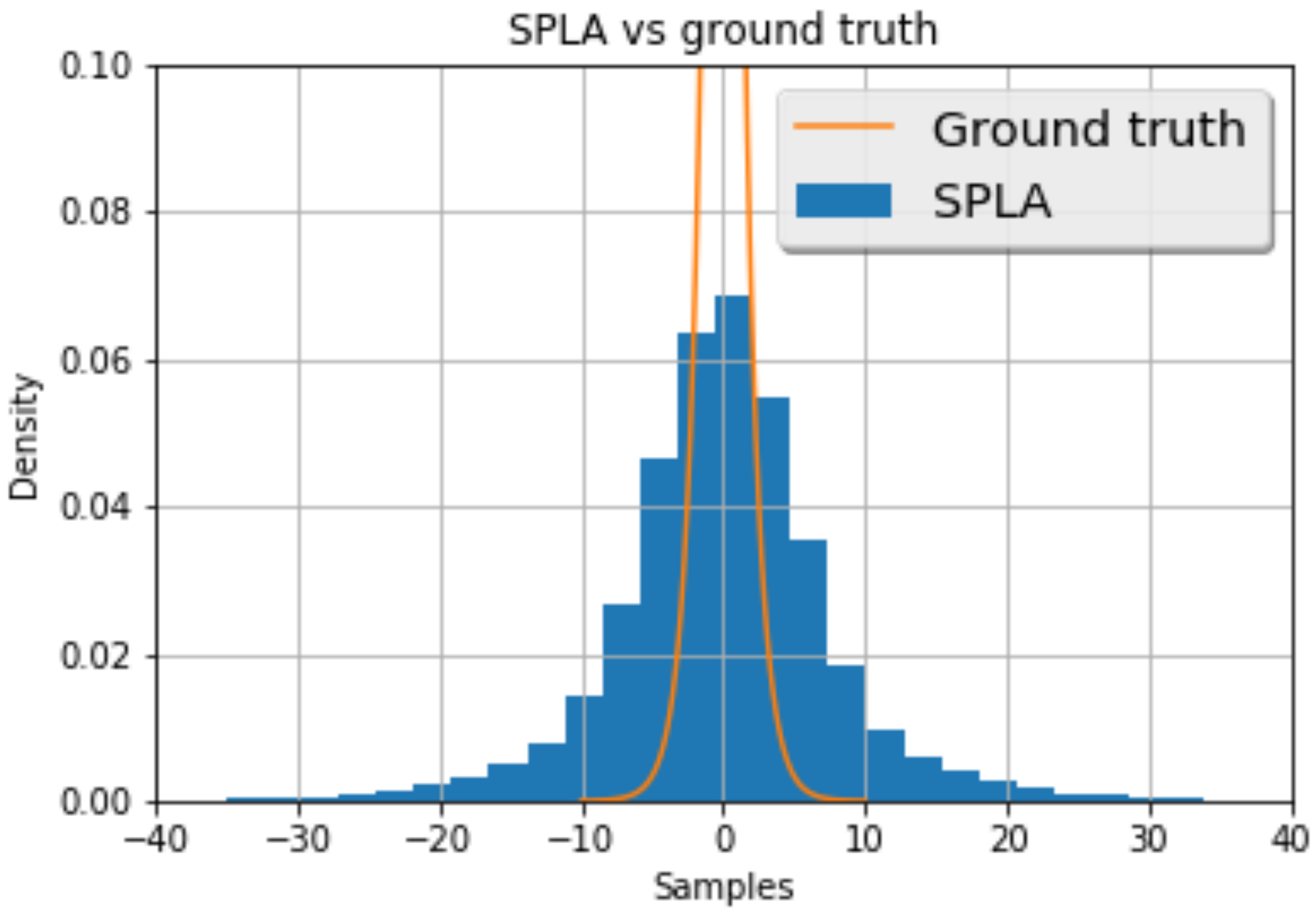
# STOCHASTIC SUBGRADIENT VS STOCHASTIC PROX

Sampling $\mu^\star(\mathrm{d}x) \propto \exp(-|x|)\mathrm{d}x$.




Stochastic subgradients

[Durmus *et al.* 2018]




Stochastic proximal

[Us]

Thanks for your attention.

See us at poster **#161**.