The Implicit Metropolis-Hastings Algorithm

Kirill Neklyudov, Evgenii Egorov, Dmitry Vetrov

Skoltech — Skolkovo Institute of Science and Technology

SAMSUNG AI Center - Moscow

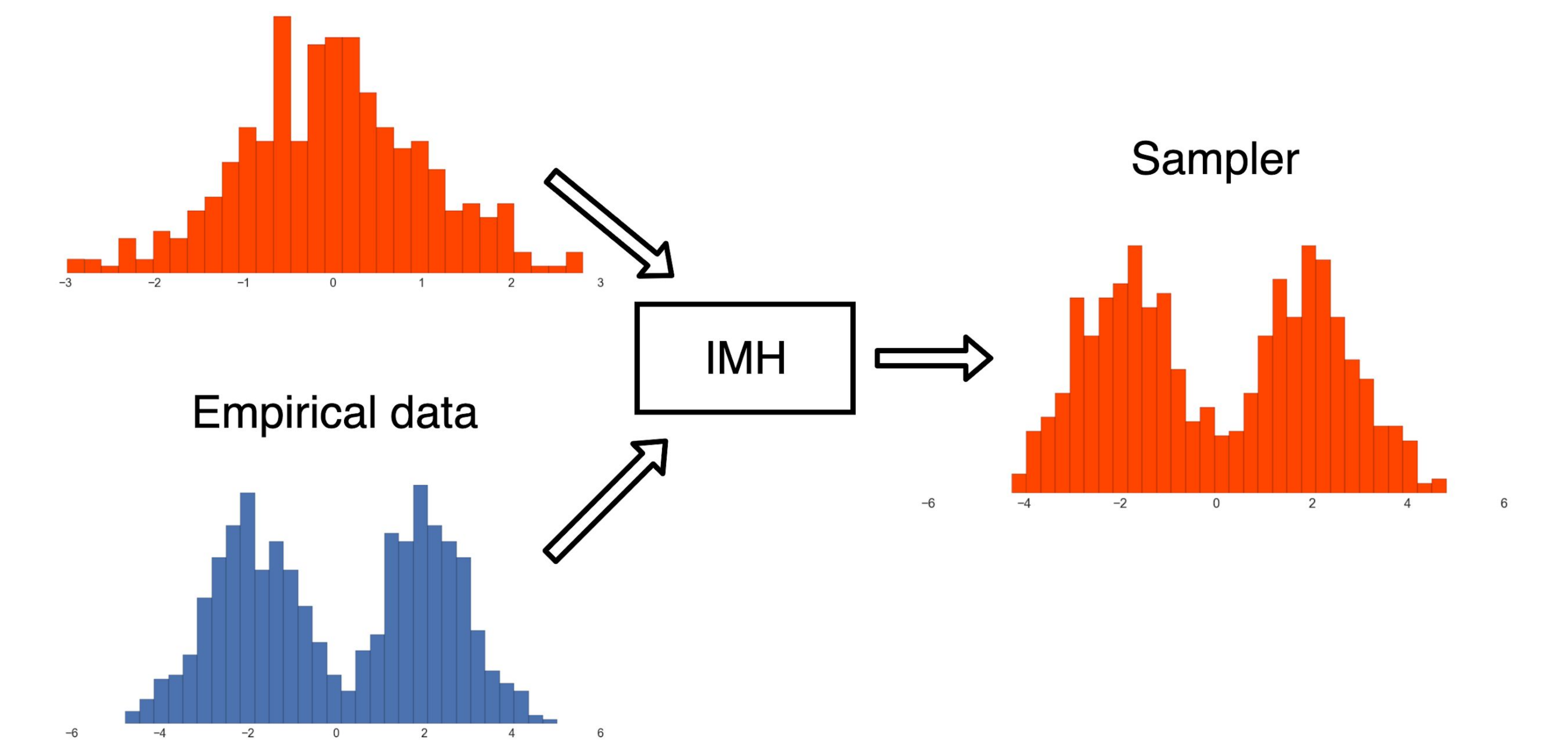
p(B|A)yesgroup.ru

ADASE

Paper

Code

Motivation



Metropolis-Hastings (MH)

input density of target distribution $\hat{p}(x) \propto p(x)$
input proposal distribution $q(x \,|\, y)$
  $y \leftarrow$ random init
  for $i = 0 \dots n$ **do**
    sample proposal point $x \sim q(x \,|\, y)$
    $P = \min\{1, \frac{\hat{p}(x)q(y \,|\, x)}{\hat{p}(y)q(x \,|\, y)}\}$
    $x_i = \begin{cases} x, & \text{with probability } P \\ y, & \text{with probability } (1-P) \end{cases}$
    $y \leftarrow x_i$
  end for
output $\{x_0, \dots, x_n\}$

Implicit Metropolis-Hastings (IMH)

input target dataset $\mathcal{D}$
input implicit model $q(x \,|\, y)$
input learned discriminator $d(\cdot,\cdot)$
  $y \sim \mathcal{D}$ initialize from the dataset
  for $i = 0 \dots n$ **do**
    sample proposal point $x \sim q(x \,|\, y)$
    $P = \min\{1, \frac{d(x,y)}{d(y,x)}\}$
    $x_i = \begin{cases} x, & \text{with probability } P \\ y, & \text{with probability } (1-P) \end{cases}$
    $y \leftarrow x_i$
  end for
output $\{x_0, \dots, x_n\}$

	MH	IMH
Target Distribution	Density Function	Set of Samples
Proposal Generator	Explicit (Restrictive)	Implicit (Rich Family)
Acceptance Probability	Depends on Exact Density Ratio	Depends on Estimated Density Ratio
Guarantees	Holds	See Our Results ↗

Sampling algorithm = Transitional kernel

$$t(\,|\,) = \underbrace{q(\,|\,)}_{\text{proposal generator}} \overbrace{P(\,|\,)}^{\text{acceptance probability}} + \delta(\,-\,) \underbrace{\int q(\,|\,)\left(1 - P(\,|\,)\right)}_{\text{rejection probability}}$$

Our Result (Intuition)

We study **Fixed Point Equation** for t_{IMH}

For **MH**:

$$t^i(\;) = t\left(t^{i-1}(\;)\right)$$

$$t^i(\;) = \;\; = t^\infty, \;\; \forall i$$

For **IMH**:

$$t^i(\;) \neq \;$$

Under mild conditions:

$$\|t^i(\;) - \;\|_{TV}$$

$$\leq \frac{1}{1-\varepsilon} \|t\;(\;) - \;\|_{TV}$$

$$\|t^\infty - \;\| \leq \text{Loss}(d(\cdot,\cdot)) + \text{Const.}$$

We obtain **Losses** to enforce **F**ixed **P**oint **E**quation

Losses over $d(\cdot,\cdot)$ for kernels:

with **Markov Proposal**

Upper bound (UB)

$$\int dxdy\; p(x)q(y \,|\, x)\left[\log \frac{d(y,x)}{d(x,y)} + \frac{d(y,x)}{d(x,y)}\right]$$

Markov cross-entropy (MCE)

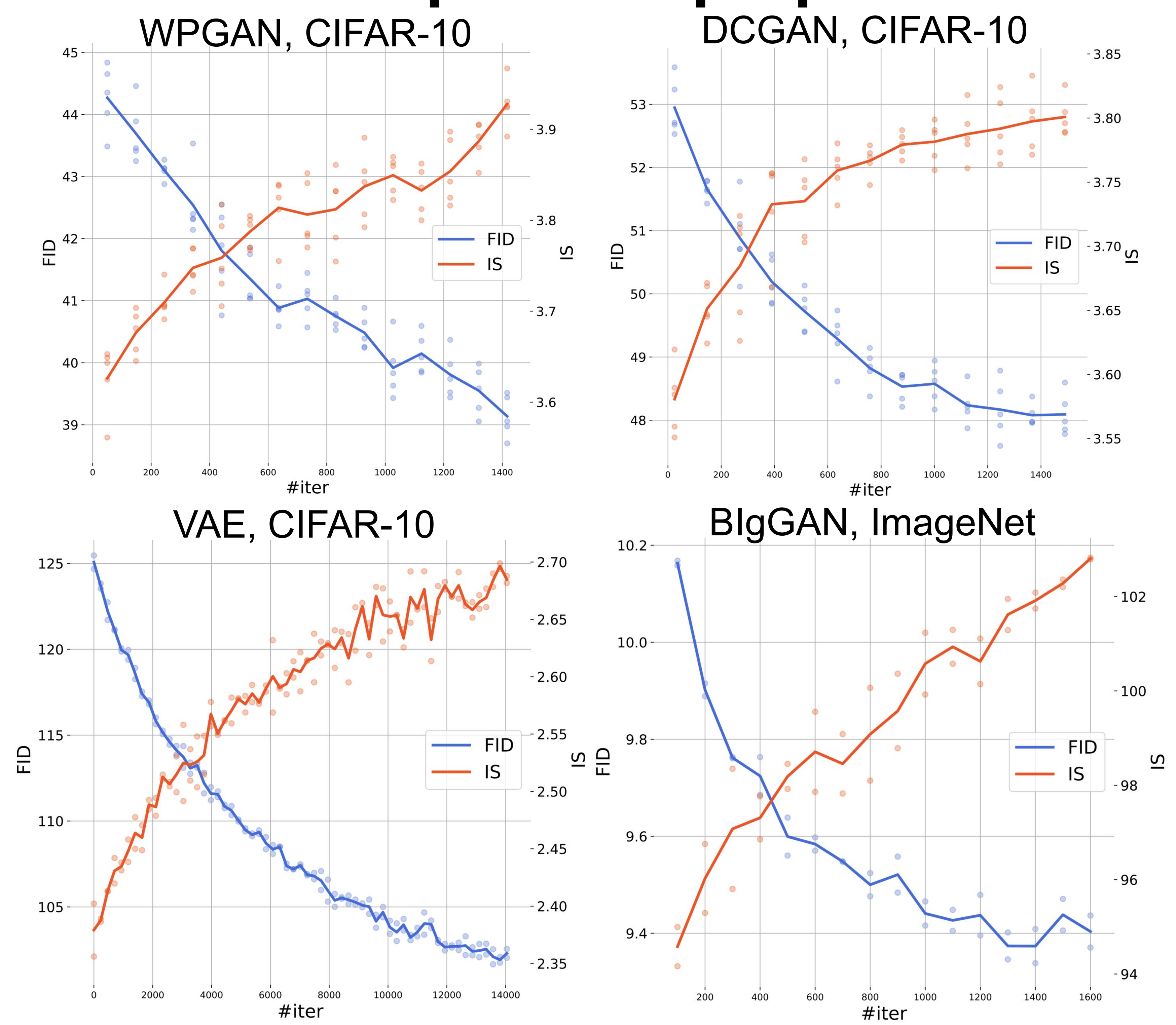
$$\int dxdy\; p(x)q(y \,|\, x)[-\log d(x,y) - \log(1 - d(y,x))]$$

with **Independent Proposal**

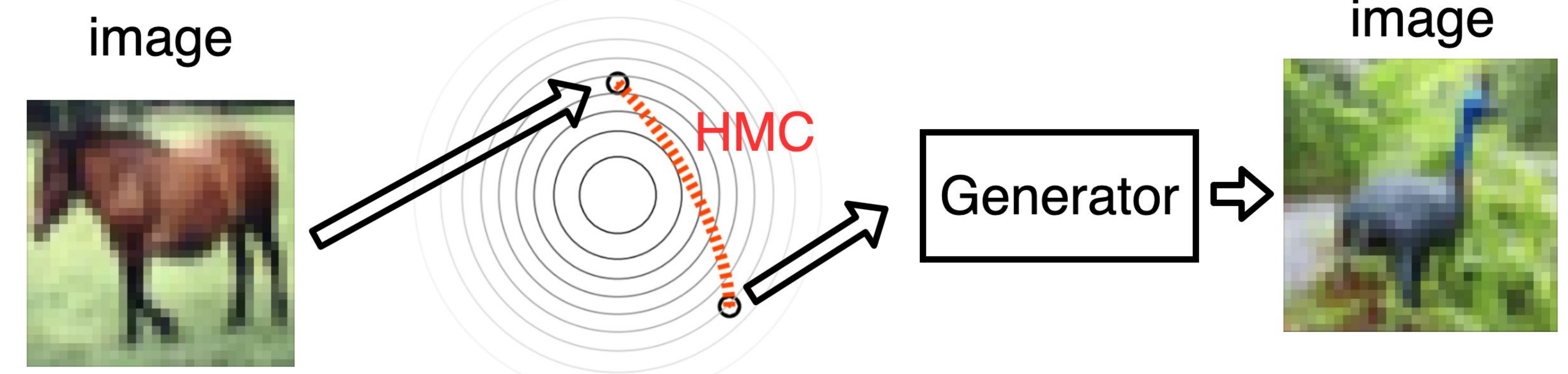
Conventional cross-entropy (CCE)

$$\int dxdy\; p(x)q(y)[-\log d(x)(1 - d(y))]$$

Independent proposal



Markov proposal



Reconstructions (finding latent vectors of the dataset)