Metropolis-Hastings with Approximate Acceptance Ratio Calculation

Yury Svirschevsky [1]   Sergey Samsonov [1]

[1]HDI Lab, HSE

Approximate Metropolis-Hastings

Input: target density $\pi(x)$, existing sample $X_1, \ldots, X_n$

Train a generative model $\mathcal{M}$ on X
$\hat{p}_{\mathcal{M}} \leftarrow$ unbiased estimator of marginal likelihood of $\mathcal{M}$
$Y_0 \leftarrow X_0$
for i=1 **to** n **do**
  Draw sample X_i from $\mathcal{M}$
  Compute acceptance rate

$$\alpha(Y_{i-1}, X_i) = \frac{\pi(X_i)\hat{p}_{\mathcal{M}}(Y_{i-1})}{\pi(Y_{i-1})\hat{p}_{\mathcal{M}}(X_i)} \wedge 1$$

  Get next sample

$$Y_i \leftarrow \begin{cases} X_i & \text{with probability } \alpha(Y_{i-1}, X_i), \\ Y_{i-1} & \text{with probability } 1 - \alpha(Y_{i-1}, X_i) \end{cases}$$

end for

- Generalization of the Metropolis-Hastings Algorithm, a popular MCMC method
- Generative model with **intractable marginal likelihood** used to model the proposal distribution
- **Estimate** of the model's marginal likelihood used in acceptance probability calculations instead of the exact value
- Range of applicability of the algorithm significantly increased

Marginal Likelihood Estimation

The marginal likelihood of a **Variational Autoencoder** is intractable, but can be approximated using an L-sample **importance weighted** estimate:

$$\hat{p}_L(x) = \frac{1}{L}\sum_{i=1}^{L} \frac{p_\theta(x, Z_i)}{q_\phi(Z_i|x)},$$

where $Z_1, \ldots, Z_L \sim q_\phi(\cdot|x)$ are sampled independently from the encoder and $p_\theta(x, z)$ is the joint distribution defined by the decoder.

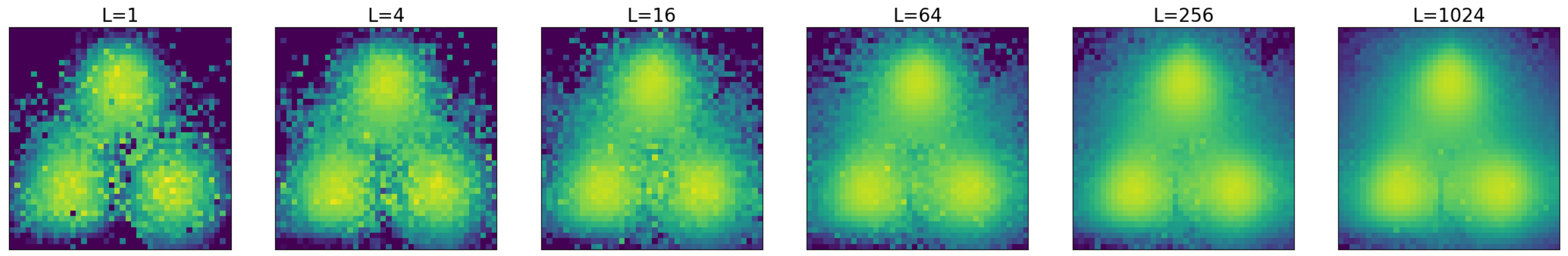

Figure 1. Importance Weighted Marginal Likelihood Estimates Based on Different Numbers of Samples

Performance of Intractable Generative Models can be Improved Using MCMC

Sample Quality is Improved

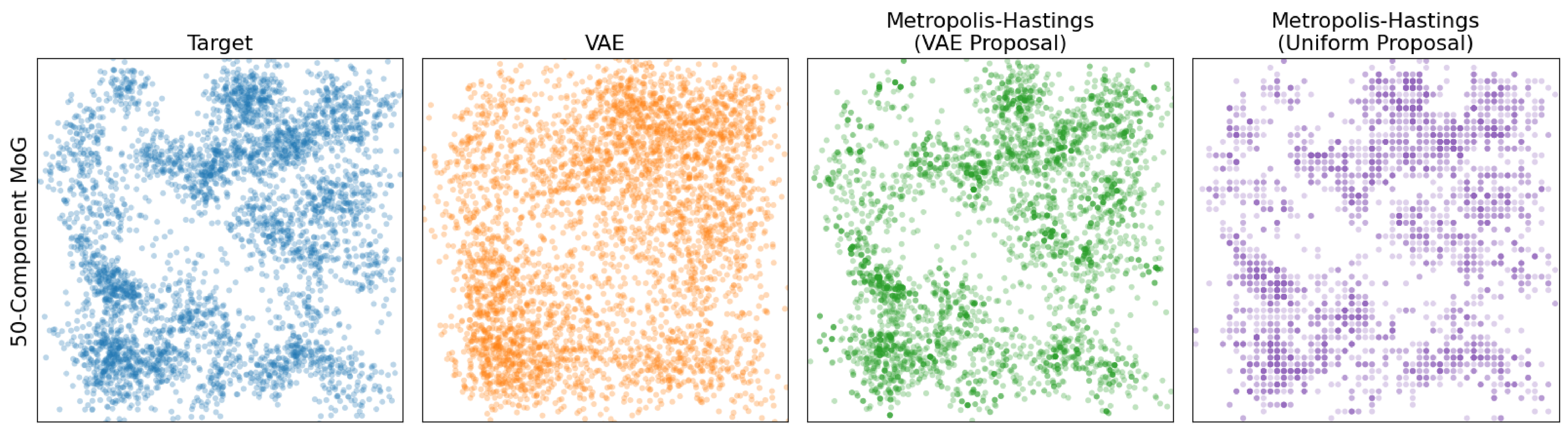

Figure 2. Demonstration on a 2D Mixture-of-Gaussians Target

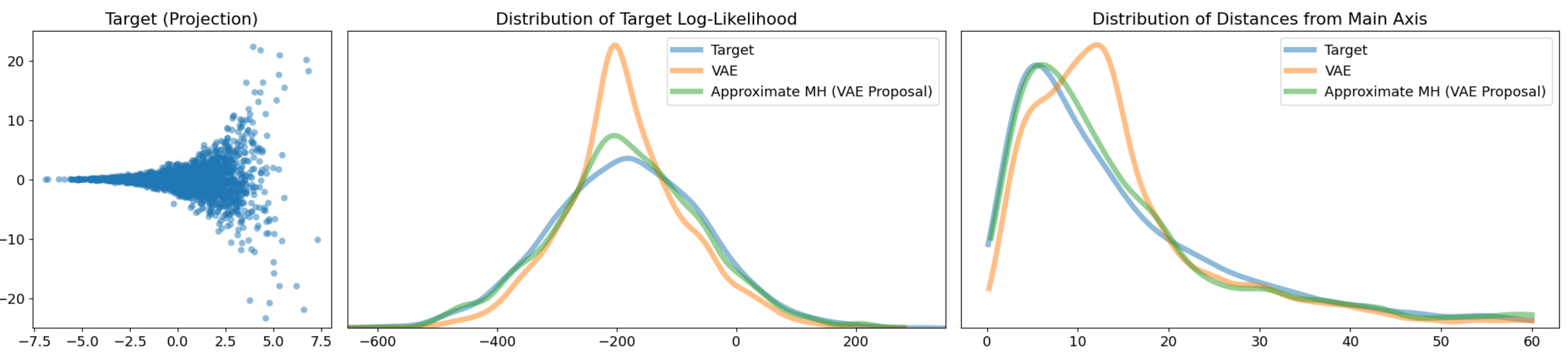

Figure 3. Approximate Metropolis-Hastings Improves Feature Distributions for a 128D Funnel

Better Than Classic Metropolis-Hastings

VAE-proposal Approximate Metropolis-Hastings can outperform RealNVP-proposal Metropolis-Hastings due to the greater power of the proposal model.

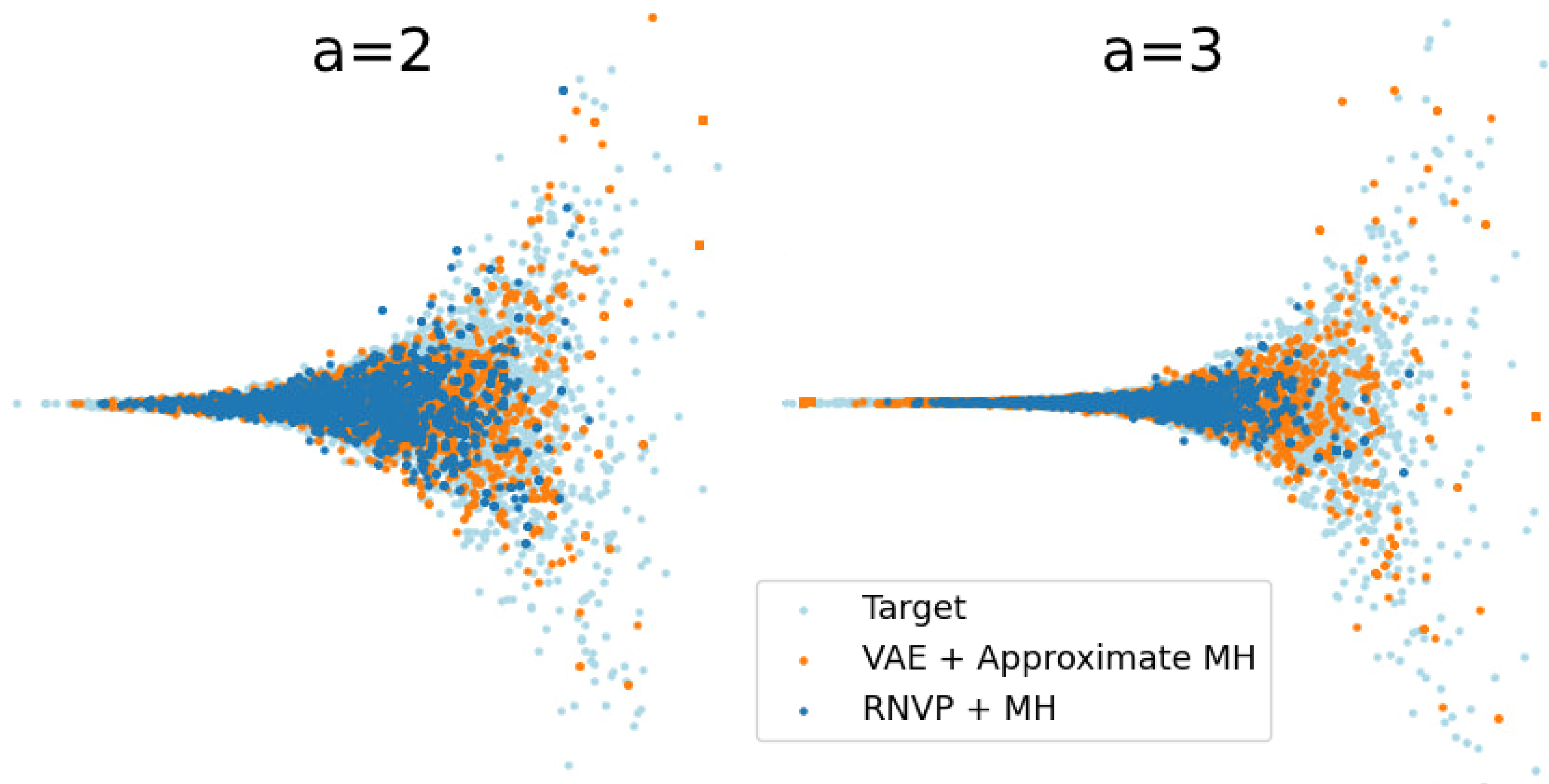

Figure 4. VAE-Proposal Approximate M-H Beats Flow-Proposal M-H on 128D Funnels