Statistically faithful sparsity and extended variational objective lead to improved inference amortization at reduced computational cost.

Program Graph Inverse Network Flow

\[\begin{align*}
  z_0 &\sim \mathcal{N}(0, 1) \\
  z_1 &\sim \mathcal{N}(-2, 1) \\
  z_2 &\sim \mathcal{N}(\tanh(z_0) + z_1 - 2.5; 0.1) \\
  z_3 &\sim \mathcal{N}(z_0 + z_1; 0.1) \\
  z_4 &\sim \mathcal{N}(7, 2) \\
  z_5 &\sim \mathcal{N}(\tanh(z_2) + z_4; 0.1) \\
  x_0 &\sim \mathcal{N}(z_0; 0.1) \\
  x_1 &\sim \mathcal{N}(z_5; 0.1)
\end{align*}\]

Sparse Neural ODE

Automated compiler pipeline for probabilistic programs (contributions [0])

1. **Compile** into graphical model
2. **Faithfully invert** graphical model using NaMI algorithm [1]
4. **Train** conditional continuous [3] normalizing flow with symmetric KL

Continuous Normalizing Flow

Learn neural network \(f_{\theta}\) for volume preserving, invertible particle flow

\[\frac{d}{dt} z_t = f_{\theta}(z_t, t; z)\]
\[\frac{d}{dt} \log q_{\phi}(z_t, t) = -\nabla_z \cdot f_{\theta}(z_t, t; z)\]

Sparsity Structure

Encode sparsity \(H\) of inverse graph into each layer of \(f_{\theta}\)

\[q_{\phi}(z_t, t) = h_{\phi}(z_t, t) \left( 1 + h_{\phi}(z_t, t) \right)\]

Improved Objective: Symmetric KL

In comparison to forward or backward KL (above), optimizing expected symmetric KL loss (below) improves amortized sample efficiency (model above).

Generative Model from Stochastic Inverse of MLP Classifier (MNIST)

Minimally faithful inverse sparsity for MNIST generator

Stochastic inverse of a convolution layer & output reconstructions

Effects of Sparsity

Comparable loss with < 50% parameters, compared to FFJORD [2] (model on the left)

Faithful sparsity is numerically stable (no increase in integration steps)

References