# Reparameterization Gradient for Non-differentiable Models

Wonyeol Lee, Hangyeol Yu, Hongseok Yang

School of Computing, KAIST

## Summary

- One of the key challenges in stochastic variational inference is to design a low-variance estimator of objective's gradient.
- The well-known reparameterization estimator has a low variance, but becomes biased for non-differentiable models.
- We generalize the reparameterization estimator so that it works for non-differentiable models as well.

## Variational Inference

- Let $p(\mathbf{x},\mathbf{z})$ be a probabilistic model about observed variable $\mathbf{x} \in \mathbb{R}^m$ and latent variable $\mathbf{z} \in \mathbb{R}^n$.
- We are interested in inferring the posterior density $p(\mathbf{z}|\mathbf{x}^0)$ given a particular value $\mathbf{x}^0$ of $\mathbf{x}$.
- Variational inference (VI) recasts the posterior inference problem as an optimization problem as follows.
- Given a collection of variational distributions $\{q_\theta(\mathbf{z})\}_{\theta\in\mathbb{R}^d}$, VI aims to find $\theta$ that maximizes the evidence lower bound (ELBO):

$$\mathrm{ELBO}_\theta \triangleq \mathbb{E}_{q_\theta(\mathbf{z})}\left[\log \frac{r(\mathbf{z})}{q_\theta(\mathbf{z})}\right], \quad \text{where } r(\mathbf{z}) \triangleq p(\mathbf{x}^0,\mathbf{z}).$$

- To solve the optimization problem efficiently, we need to estimate $\nabla_\theta \mathrm{ELBO}_\theta$ with a low variance.

## Standard Gradient Estimators

- Score estimator (or REINFORCE):

$$\nabla_\theta \mathrm{ELBO}_\theta = \mathbb{E}_{q_\theta(\mathbf{z})}\left[\nabla_\theta \log q_\theta(\mathbf{z}) \cdot \log \frac{r(\mathbf{z})}{q_\theta(\mathbf{z})}\right]$$

  - It has a high variance,
  - but can be applied even when $r(\mathbf{z})$ is non-differentiable.

- Reparameterization estimator:

$$\nabla_\theta \mathrm{ELBO}_\theta = \nabla_\theta \mathbb{E}_{q(\epsilon)}\left[\log \frac{r(f_\theta(\epsilon))}{q_\theta(f_\theta(\epsilon))}\right] = \mathbb{E}_{q(\epsilon)}\left[\nabla_\theta \log \frac{r(f_\theta(\epsilon))}{q_\theta(f_\theta(\epsilon))}\right]$$

  where $q(\cdot)$ and $f_\theta(\cdot)$ satisfy that $f_\theta(\epsilon)$ for $\boldsymbol{\epsilon} \sim q(\epsilon)$ has the distribution $q_\theta$.

  - It has a low variance,
  - but can be applied only when $r(\mathbf{z})$ is differentiable.

## Non-differentiable Models

- A probabilistic model can have non-differentiable density if
  - it uses both discrete and continuous random variable,
  - or it is specified using if-statements as in probabilistic programming.
- Assume $r(\mathbf{z})$ has the form:

$$r(\mathbf{z}) = \sum_{k=1}^{K} 1[\mathbf{z} \in R_k] \cdot r_k(\mathbf{z})$$

  where $r_k$ is differentiable, $\{R_k\}_{1\le k\le K}$ is a disjoint partition of $\mathbb{R}^n$, and $\partial R_k$ has Lebesgue measure zero.

- Example: Gaussian mixture model

$$p(z) = \mathcal{N}(z|0,1)$$

$$p(x|z) = 1[z>0]\mathcal{N}(x|5,1) + 1[z\le 0]\mathcal{N}(x|-2,1)$$

- For the above example with $x^0 = 0$ and $q_\theta(z) = \mathcal{N}(z|\theta,1)$,

$$\nabla_\theta \mathrm{ELBO}_\theta \neq \mathbb{E}_{q(\epsilon)}\left[\nabla_\theta \log \frac{r(f_\theta(\epsilon))}{q_\theta(f_\theta(\epsilon))}\right]$$

  where $q(\epsilon) = \mathcal{N}(\epsilon|0,1)$ and $f_\theta(\epsilon) = \epsilon + \theta$.

  This happens because the below equation does not hold in general if $g$ is non-differentiable in $\theta$:

$$\nabla_\theta \int g(\boldsymbol{\epsilon},\theta)d\boldsymbol{\epsilon} = \int \nabla_\theta g(\boldsymbol{\epsilon},\theta)d\boldsymbol{\epsilon}$$

- In sum, the standard reparameterization estimator is biased for non-differentiable models.

## Reparameterization for Non-differential Models

- Our unbiased reparameterization estimator:

$$\nabla_\theta \mathrm{ELBO}_\theta = \mathbb{E}_{q(\boldsymbol{\epsilon})}\left[\sum_{k=1}^{K} 1[f_\theta(\boldsymbol{\epsilon}) \in R_k] \cdot \nabla_\theta \log \frac{r_k(f_\theta(\boldsymbol{\epsilon}))}{q_\theta(f_\theta(\boldsymbol{\epsilon}))}\right] +$$

Reparam'n term

Correction term

$$\sum_{k=1}^{K} \int_{f_\theta^{-1}(\partial R_k)} \left( q(\boldsymbol{\epsilon}) \log \frac{r_k(f_\theta(\boldsymbol{\epsilon}))}{q_\theta(f_\theta(\boldsymbol{\epsilon}))} \cdot \mathbf{V}(\boldsymbol{\epsilon},\theta) \right) \cdot d\boldsymbol{\Sigma}$$

- Here $\mathbf{V}(\boldsymbol{\epsilon},\theta) \in \mathbb{R}^{d\times n}$ is the velocity of $f_\theta^{-1}$ defined as

$$\mathbf{V}(\boldsymbol{\epsilon},\theta)_{ij} \triangleq \left( \frac{\partial}{\partial\theta_i} f_\theta^{-1}(\mathbf{z}) \bigg|_{\mathbf{z}=f_\theta(\boldsymbol{\epsilon})} \right)_j$$

## Key Ingredients

- Differentiation under moving domains:

$$\nabla_\theta \int_{D_\theta} g(\boldsymbol{\epsilon},\theta)d\boldsymbol{\epsilon} = \int_{D_\theta} \left(\nabla_\theta g + \nabla_{\boldsymbol{\epsilon}} \cdot (g\mathbf{v})\right)(\boldsymbol{\epsilon},\theta)d\boldsymbol{\epsilon}$$

- Divergence theorem:

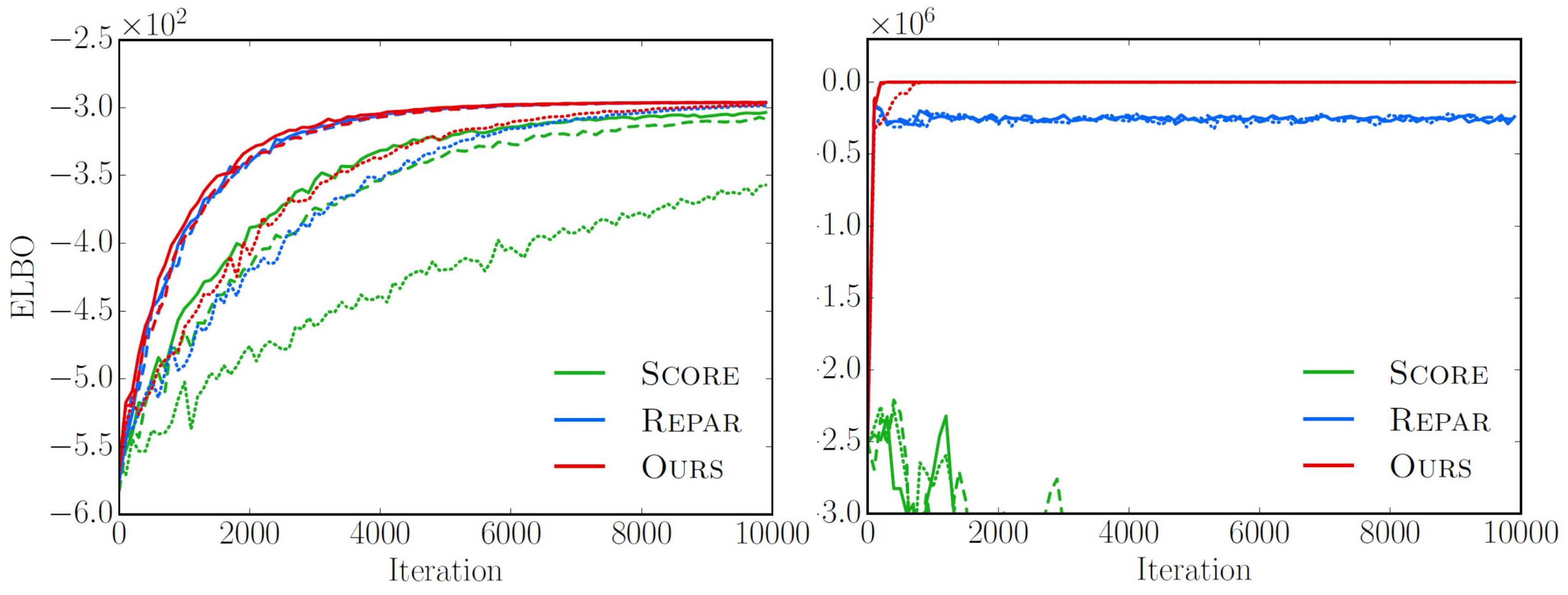
$$\int_V (\nabla \cdot \boldsymbol{G})dV = \int_{\partial V} \boldsymbol{G} \cdot d\boldsymbol{\Sigma}$$

- Estimation of surface integral:

  We assume that the boundaries $f_\theta^{-1}(\partial R_k)$ are affine.

## Experimental Evaluation

- ELBO objective as a function of iteration number:



textmsg (stepsize = 0.001) temperature (stepsize = 0.01)

- Ratio of {REPARAM,OURS}'s average variance to SCORE's

| Estimator | Type of Variance | temperature | textmsg | influenza |
|---|---|---|---|---|
| Reparam | $\mathrm{Avg}(\mathbb{V}(\cdot))$ | $\mathbf{4.45 \times 10^{-9}}$ | $2.91 \times 10^{-2}$ | $\mathbf{4.38 \times 10^{-3}}$ |
| | $\mathbb{V}(\lVert\cdot\rVert_2)$ | $\mathbf{2.45 \times 10^{-8}}$ | $2.92 \times 10^{-2}$ | $\mathbf{2.12 \times 10^{-3}}$ |
| Ours | $\mathrm{Avg}(\mathbb{V}(\cdot))$ | $1.85 \times 10^{-6}$ | $\mathbf{2.77 \times 10^{-2}}$ | $4.89 \times 10^{-3}$ |
| | $\mathbb{V}(\lVert\cdot\rVert_2)$ | $7.59 \times 10^{-5}$ | $\mathbf{2.46 \times 10^{-2}}$ | $2.36 \times 10^{-3}$ |

stepsize = 0.001

- Computation time (per iteration, in ms)

| Estimator | temperature | textmsg | influenza |
|---|---|---|---|
| Score | 21.7 | 4.9 | 18.7 |
| Reparam | 46.1 | 15.4 | 251.4 |
| Ours | 79.2 | 24.9 | 269.8 |

  - OURS subsamples the summation in the correction term.