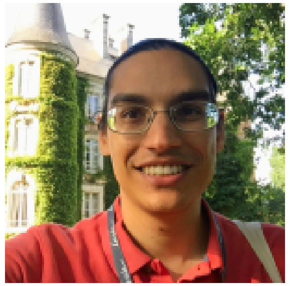
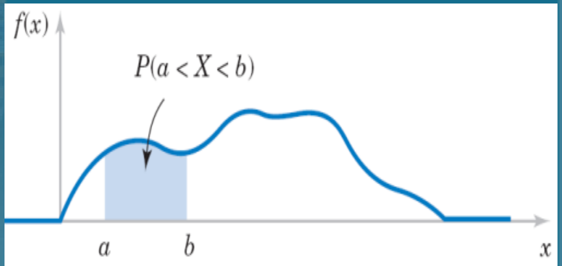
SAND2022-9305C

Sandia National Laboratories

A Study of Bias-Variance in Variational Inferencing Using Delta Method





Niladri Das & Thomas A. Catanach (P.I.)

July 12, 2022

2022 SIAM Annual Meeting, Pittsburg, Pennsylvania, USA

U.S. DEPARTMENT OF ENERGY NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

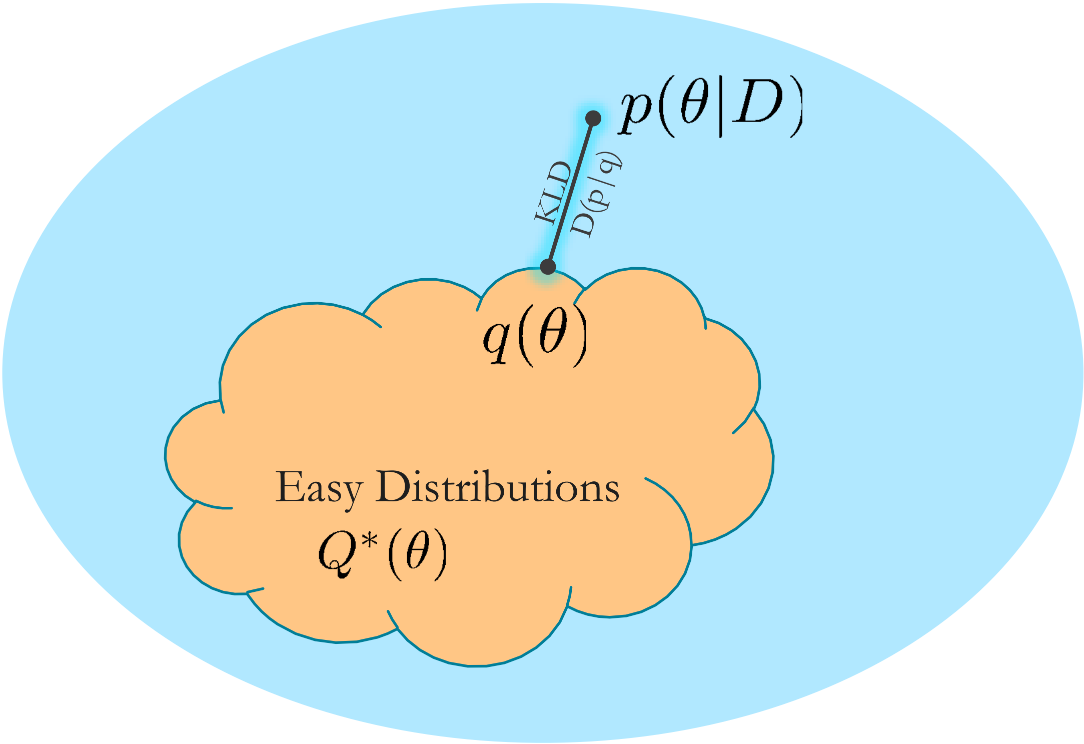
A Study of Bias-Variance in **Variational Inferencing** Using Delta Method

Bayes Rule

Posterior — Prior — Likelihood

$$p(\theta|D) = \frac{p(\theta)p(D|\theta)}{\int_\theta p(\theta)p(D|\theta)d\theta}$$

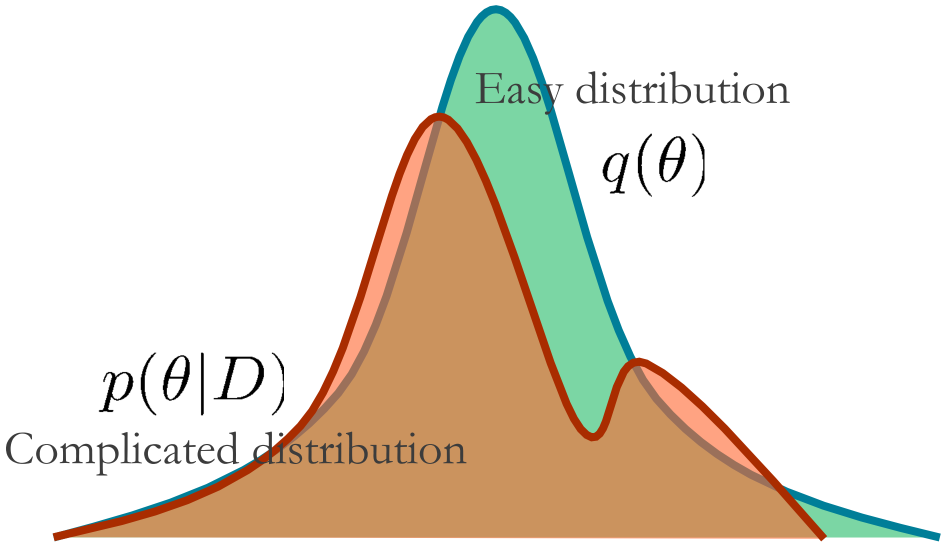
$\mathbb{R}^{100\dots}$



At even moderately high dimensions of the θ amount numerical operations explode.

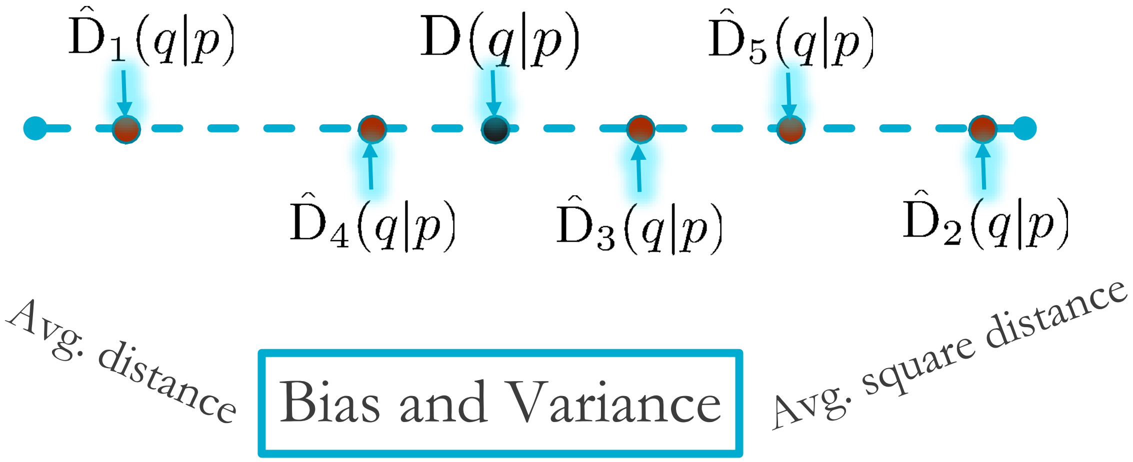
$q(\theta)$: Variational Distribution

A Study of **Bias-Variance** in Variational Inferencing Using **Delta Method**



D(p | q) = KLD

Sample Estimate



$$D(q|p) = \int q\left(\theta \mid \phi\right) \log \frac{q\left(\theta \mid \phi\right)}{p\left(\theta\right)} d\theta - \int q\left(\theta \mid \phi\right) \log p\left(\mathcal{D} \mid \theta\right) d\theta + \log \int p\left(\mathcal{D} \mid \theta\right) p\left(\theta\right) d\theta$$

Minimize

Log Evidence , independent of VI params. ϕ

$$\text{ELBO} = -\int q\left(\theta \mid \phi\right) \log \frac{q\left(\theta \mid \phi\right)}{p\left(\theta\right)} d\theta + \int q\left(\theta \mid \phi\right) \log p\left(\mathcal{D} \mid \theta\right) d\theta$$

Maximize

$$= -\frac{1}{N}\Sigma\left(\ldots\right) \qquad + \frac{1}{N}\Sigma\left(\ldots\right)$$

A Study of **Bias-Variance** in Variational Inferencing Using **Delta Method**

$$D(q|p) = \int q(\theta \mid \phi) \log \frac{q(\theta \mid \phi)}{p(\theta)} d\theta - \int q(\theta \mid \phi) \log p(\mathcal{D} \mid \theta) d\theta + \log \int p(\mathcal{D} \mid \theta)\, p(\theta)\, d\theta$$

Minimize

Accurate Calculation

$$= \frac{1}{N}\Sigma(\ldots) \quad \mathcal{X}$$

$$= \frac{1}{N}\Sigma(\ldots) \quad \log(\mathcal{Y})$$

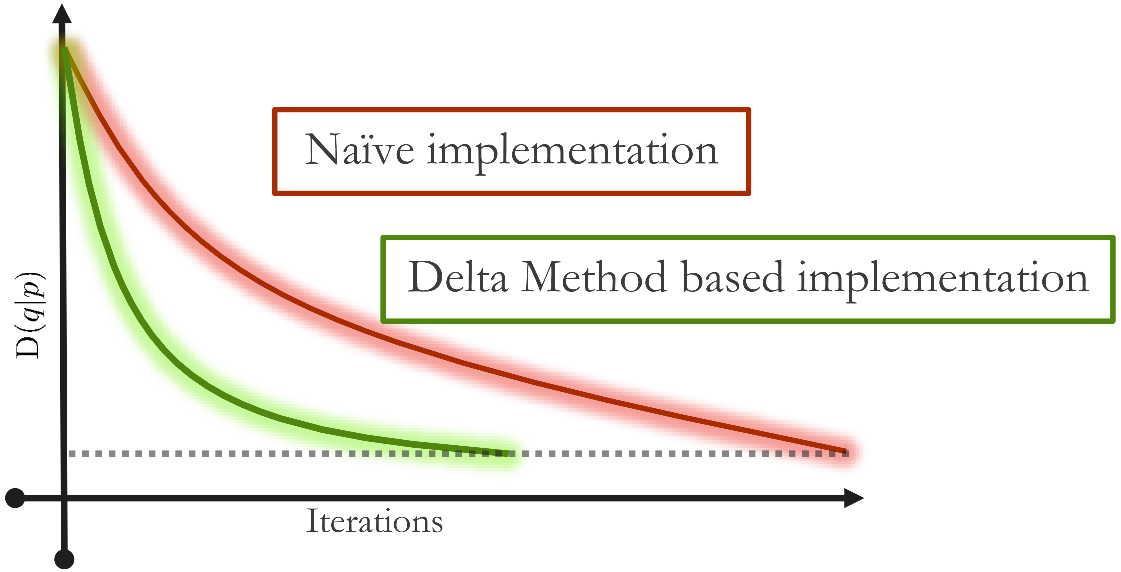
$$\mathcal{V}ar(\hat{\mathcal{X}} + \log(\hat{\mathcal{Y}})) = \frac{\Sigma\mathcal{X}}{N} + \frac{\Sigma\mathcal{Y}}{N\mathcal{Y}^2} + 2\frac{\Sigma\mathcal{X}\mathcal{Y}}{N\mathcal{Y}}$$

A Study of **Bias-Variance** in Variational Inferencing Using **Delta Method**

Delta Method : Approx. probability distribution of $\mathrm{D}(q|p)$

Technique : Reduce variance in approx. of $\mathrm{D}(q|p)$

Performance : Faster convergence to $q(\theta)$



A Study of **Bias-Variance** in Variational Inferencing Using **Delta Method**

$$D(q|p) = \int q(\theta \mid \phi) \log \frac{q(\theta \mid \phi)}{p(\theta)} d\theta - \int q(\theta \mid \phi) \log p(\mathcal{D} \mid \theta) d\theta + \log \int \frac{p(\mathcal{D} \mid \theta)\, p(\theta)}{q(\theta \mid \phi)} q(\theta \mid \phi)\, d\theta$$

$$= f(\phi) - \int \boxed{\frac{q(\theta(\zeta) \mid \phi)}{r(\theta(\zeta) \mid \varphi)}} \log p(\mathcal{D} \mid \theta(\zeta)) p(\zeta) d\zeta + \log \int \boxed{\frac{p(\mathcal{D} \mid \theta(\zeta))\, p(\theta(\zeta))}{r(\theta(\zeta) \mid \varphi)}} p(\zeta) d\zeta$$

Importance Distribution

$r(\theta(\zeta) \mid \varphi)$

Re-parameterization params.

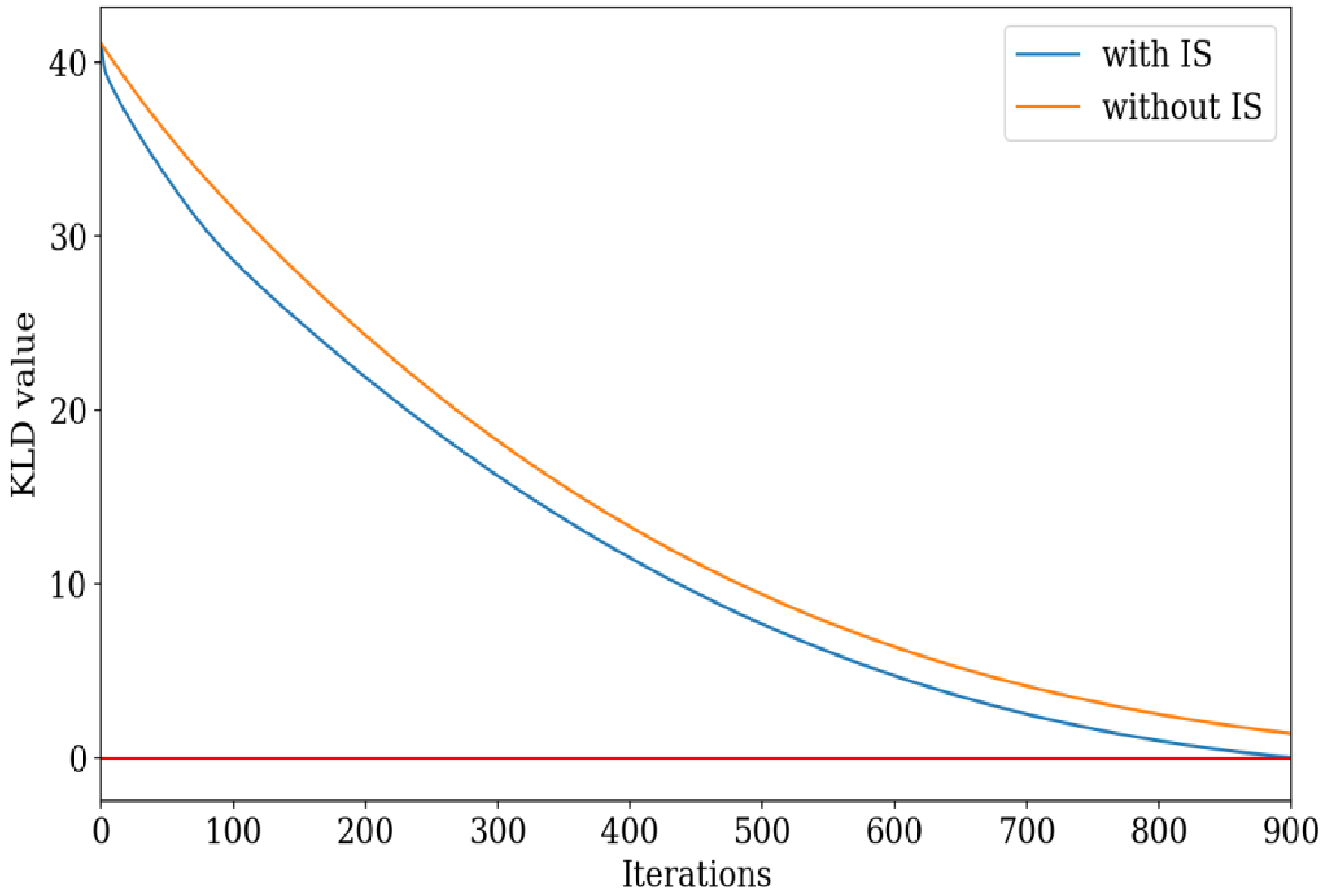
ζ

Two Step Optimization:

- **VI params. Optimization** ϕ
- **IS params. Optimization** φ

A Study of **Bias-Variance** in Variational Inferencing Using **Delta Method**

- Dimension of the problem : 5
- Linear model := $\mathcal{Y} = X + v$
- v : Zero mean Gaussian noise
- Importance distribution is Gaussian, initialized as the initial variational distribution.
- The VI parameters (μ, L) where $LL^T = \Sigma$
- Sample size 1000 for both VI and IS.



Contact Information

Niladri Das | Postdoctoral Appointee | ndas@sandia.gov

Thomas A. Catanach | P.I. | tacatan@sandia.gov

08739 | Extreme-scale DS & Analytics

Thank You