Team-Improving-VAEs: Midterm Presentation

Callihan Bertley, Claire Gan, Rishi Leburu, Malia Walewski
June 25, 2024 — Emory University
Mentor: Dr. Deepanshu Verma

VAE Skeleton

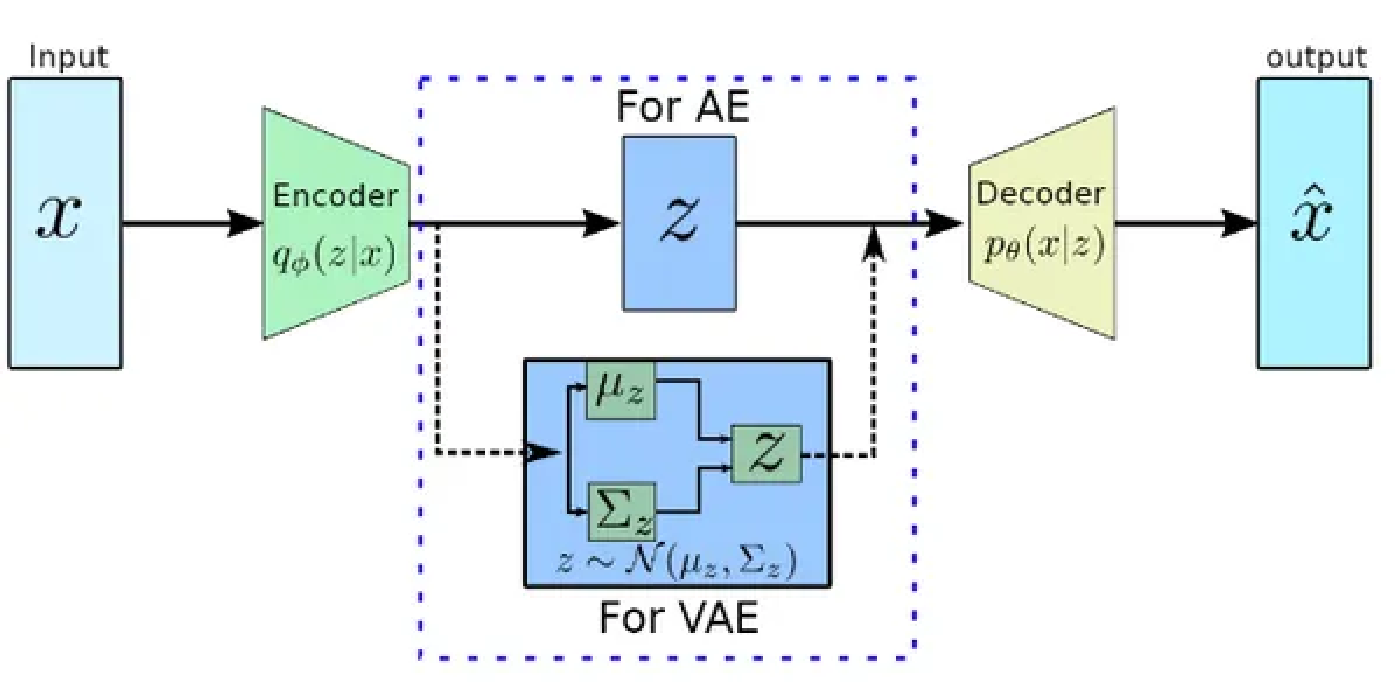


Image Source: "Generative AI with VAEs, GANs Transformers," Analytics Vidhya, 19 July 2023, [*https://www.analyticsvidhya.com/blog/2023/07/generative-ai-with-vaes-gans-transformers/*](https://www.analyticsvidhya.com/blog/2023/07/generative-ai-with-vaes-gans-transformers/)*. Accessed 24 June 2024.* [1]

Motivation

- VAEs default to a normal Gaussian to approximate the posterior
- The standard VAE makes weak assumptions about the latent space because it approximates it using a Gaussian. [2]

Objective

- Use Normalizing Flows to transform a Gaussian to better match the complex posterier distribution, $p_\theta(\mathbf{z}|\mathbf{x})$, of a VAE after training
- Experiment transforming the approximate posterior $q_\psi(\mathbf{z}|\mathbf{x})$ using different conditonal NFs
    - To improve generalization error
    - Create better generative model

Introduction to NFs

- **Normalizing Flows** transform a simple probability distribution into a more complex distribution
- This transformation is done through a sequence of invertible and differentiable mappings.

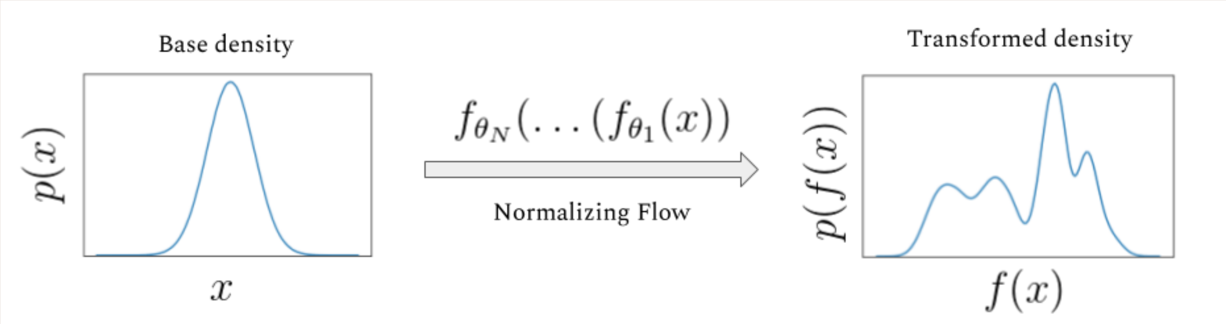
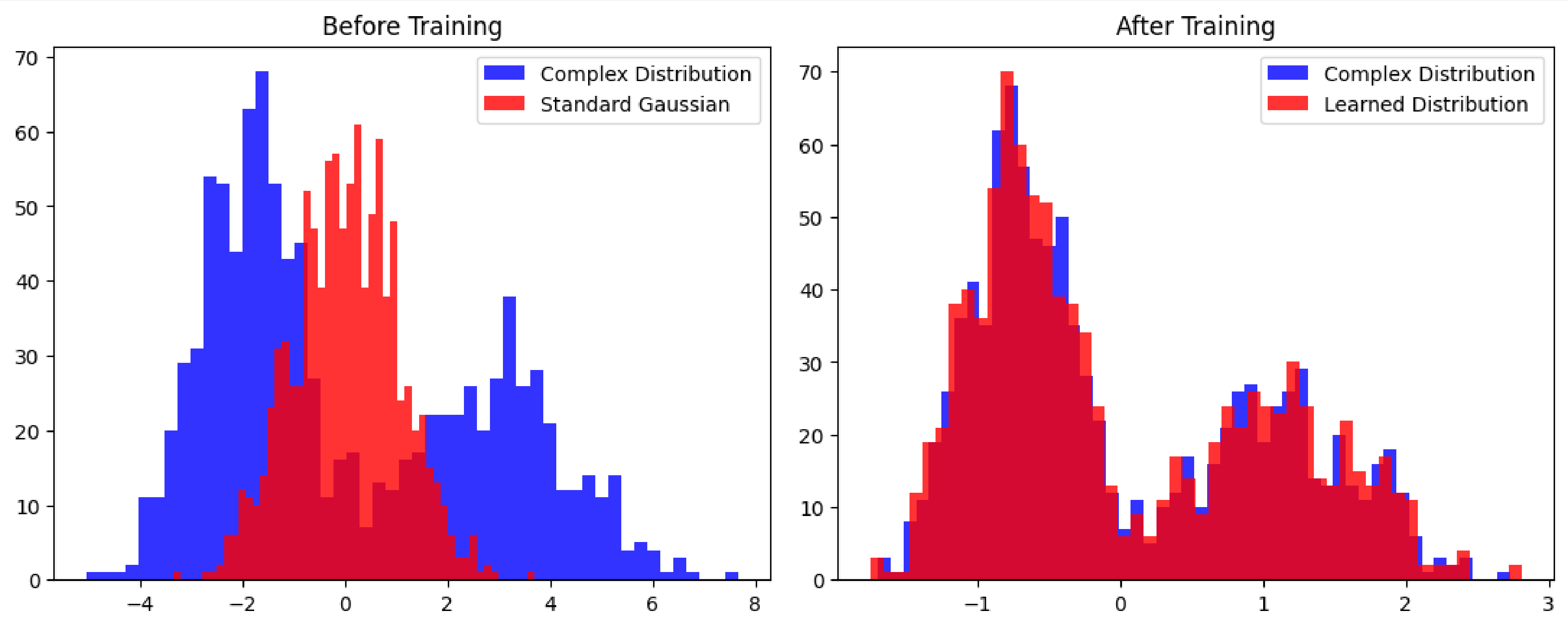


Image Source: "Normalizing Flows for Variational Inference," by B. M. Geboe, `https://gebob19.github.io/normalizing-flows/`. Accessed June 24, 2024

Practice NF Model



Inverse Autoregressive Transformations (IAT)

- When a random sample z_0 is passed through a function the pdf of the transformed sample is as below:

$$q_\psi(\mathbf{z}_1|\mathbf{x}) = q_\psi(\mathbf{z}_0|\mathbf{x}) \left|\det\left(\frac{d\mathbf{z}_0}{d\mathbf{z}_1}\right)\right| \tag{1}$$

$$\log(q_\psi(\mathbf{z}_1|\mathbf{x})) = \log(q_\psi(\mathbf{z}_0|\mathbf{x})) - \log\left(\left|\det\left(\frac{d\mathbf{z}_1}{d\mathbf{z}_0}\right)\right|\right) \tag{2}$$

- Applying T transformations sequentially:

$$\log(q_\psi(\mathbf{z}_T|\mathbf{x})) = \log(q_\psi(\mathbf{z}_0|\mathbf{x})) - \sum_{t=1}^{T} \log\left(\left|\det\left(\frac{d\mathbf{z}_t}{d\mathbf{z}_{t-1}}\right)\right|\right) \tag{3}$$

Inverse Autoregressive Transformations (IAT)

- Computational difficulties can be solved using Inverse Auto Regressive Transformations
- Multiply a sample by σ and add μ (outputs of an autoregressive neural network)
- Networks are built such that the parameters σ and μ for the j^{th} dimension of the transformed variable only depend on the $[0, j-1]$ inputs and the context vector, h. [2]

$$\implies \mathbf{z}_1^j = \mu^j(\mathbf{z}_0^{j-1}, \dots, \mathbf{z}_0^1, h) + \sigma^j(\mathbf{z}_0^{j-1}, \dots, \mathbf{z}_0^1, h)\mathbf{z}_0^j \quad (4)$$

Inverse Autoregressive Transformations (IAT)

$$\frac{d\mathbf{z}_1}{d\mathbf{z}_0} = \begin{bmatrix} \sigma^1 & 0 & 0 & 0 \\ \cdots & \sigma^2 & 0 & 0 \\ \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \sigma^K \end{bmatrix} \tag{5}$$

- If T transformations are applied sequentially:

$$\log(q_\psi(\mathbf{z}_T|\mathbf{x})) = \log(q_\psi(\mathbf{z}_0|\mathbf{x})) - \sum_{t=1}^{T}\sum_{k=1}^{K}\log\left(\sigma_t^k\right) \tag{6}$$

Inverse Autoregressive Flows (IAF)

- Inverse Autoregressive Flows (IAFs) apply a series of invertible, autoregressive transformations to a base distribution

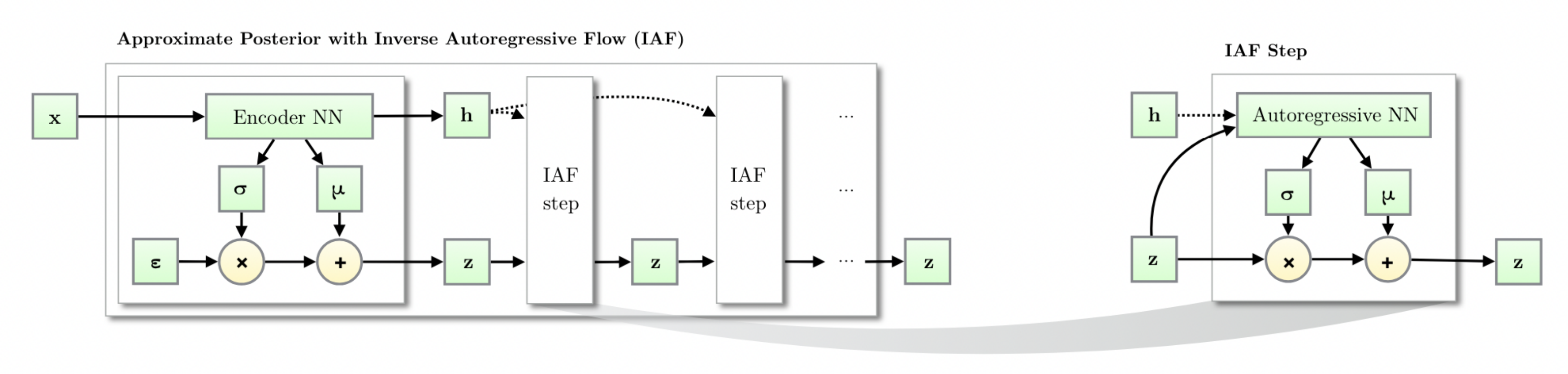


Image Source: L. Midgley, "Improving variational inference with inverse autoregressive flow," University of Cambridge, Tech. Rep., 2021. [Online]. Available: https://github.com/lollcat/Autoencoders-deep-dive/blob/Pytorch/Report.pdf.

Inverse Autoregressive Flows (IAF)

- The main advantage of IAFs/IATs are that they retain most of the properties of a standard Gaussian
- The are easy easy to compute (so easy to sample from as well) and parallelizable for high dimensional z

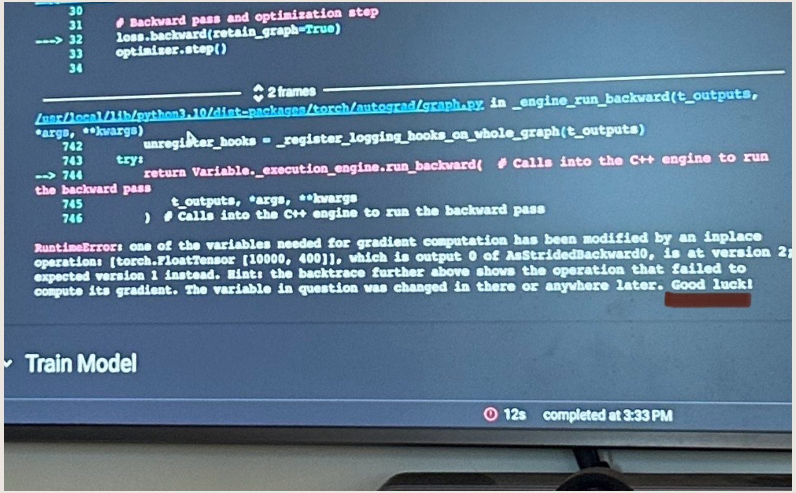
The KL Divergence & The ELBO

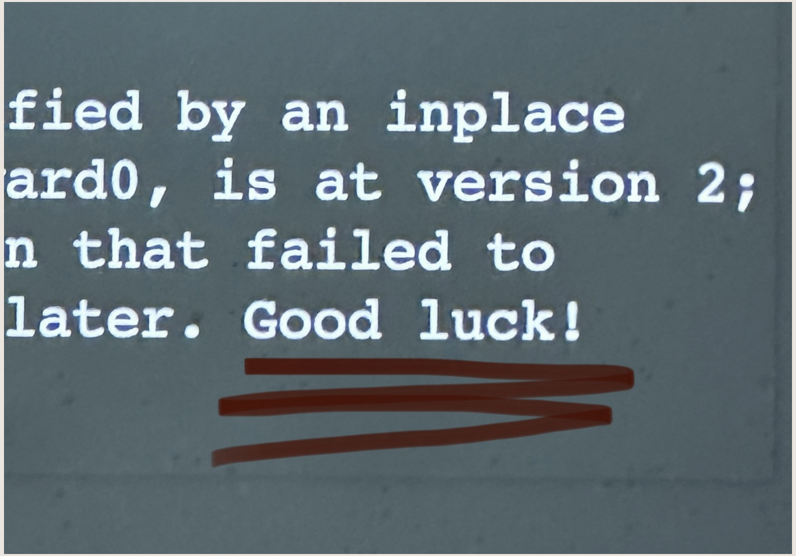
$$D_{kl}\left(q_\psi(\mathbf{z}|\mathbf{x}) \| (p_\theta(\mathbf{z}|\mathbf{x})\right) = -(\text{ELBO}) + \log p_\theta(\mathbf{x}) \tag{7}$$

$$\text{ELBO} = \underbrace{\mathbb{E}_{\mathbf{z} \sim q_\psi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x}|\mathbf{z})\right]}_{\text{reconstruction term}} - \underbrace{D_{kl}\left(q_\psi(\mathbf{z}|\mathbf{x}) \| p(\mathbf{z})\right)}_{\text{regularization term}} \tag{8}$$

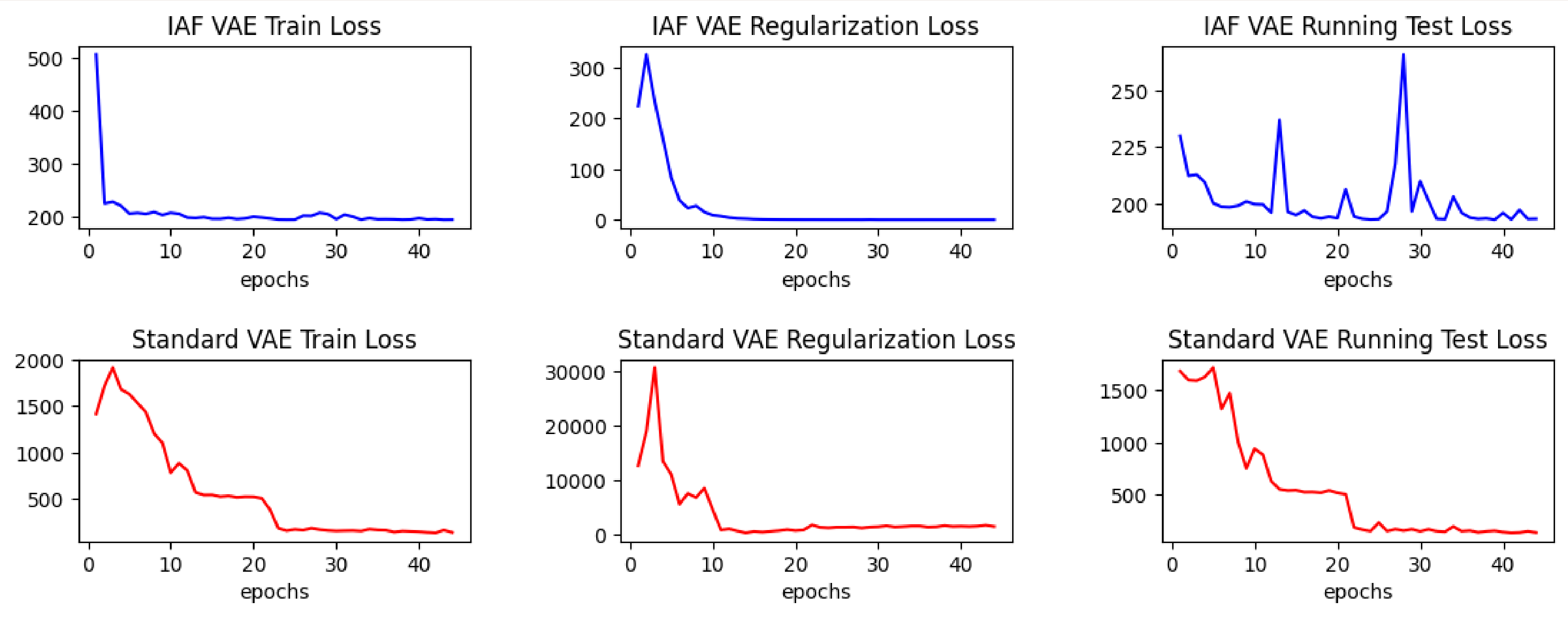
VAEs allow for efficient training of generative models by minimizing the negative evidence lower bound (ELBO). [3]

Coding is hard





IAF implementation into VAE



Next Steps

Implement Partially Convex Potential Maps (PCP-Map) [4]

- another type of architecture to transform simple distributions to complex ones
- experiment and compare results with IAF-VAE and standard VAE

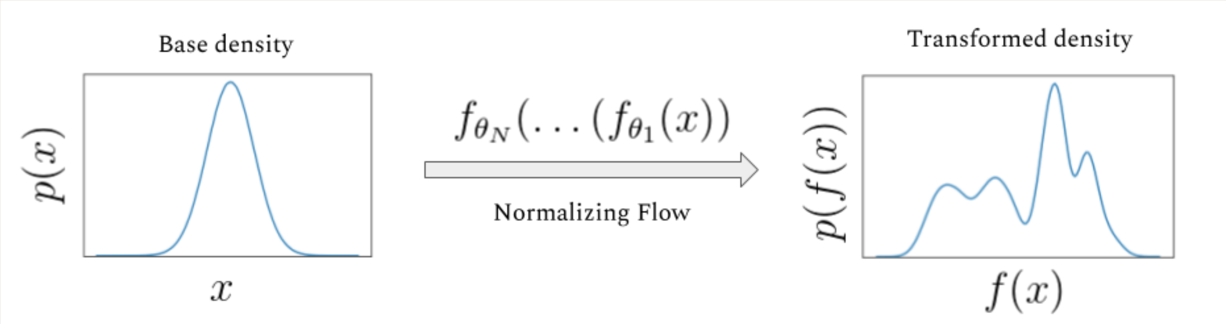


Image Source: "Normalizing Flows for Variational Inference," by B. M. Geboe, `https://gebob19.github.io/normalizing-flows/`. Accessed June 24, 2024

References

[1] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," 2014.

[2] L. Midgley, "Improving variational inference with inverse autoregressive flow," University of Cambridge, Tech. Rep., 2021. [Online]. Available: https://github.com/lollcat/Autoencoders-deep-dive/blob/Pytorch/Report.pdf

[3] L. Ruthotto and E. Haber, "An introduction to deep generative modeling," *GAMM-Mitt.*, vol. 44, no. 2, pp. Paper No. e202 100 008, 24, 2021. [Online]. Available: https://doi.org/10.1002/gamm.202100008

[4] Z. O. Wang, R. Baptista, Y. Marzouk, L. Ruthotto, and D. Verma, "Efficient neural network approaches for conditional optimal transport with applications in bayesian inference," 2023. [Online]. Available: https://arxiv.org/abs/2310.16975

- Special thanks to NSF funding awards DMS-2051019 and DMS-1751636

EMORY