Entropic GANs meet VAEs: A Statistical Approach to Compute Sample Likelihoods in GANs

Yogesh Balaji

with Hamed Hassani, Rama Chellappa and Soheil Feizi

Generative Adversarial Networks

GANs are very successful at generating samples from a data distribution

BigGAN *(Brock et al., 2018)*

StyleGAN *(Karras et al., 2018)*

Generative models

Classical Approach:

Fitting an explicit model using maximum likelihood

Modern Approach:

Generative Adversarial Networks (GANs)

Lacks an explicit density model

Generative models

Classical Approach:

Fitting an explicit model using maximum likelihood

Modern Approach:

Generative Adversarial Networks (GANs)

Lacks an explicit density model

Generative models

Classical Approach:

Fitting an explicit model using maximum likelihood

Our contribution

Modern Approach:

Generative Adversarial Networks (GANs)

Lacks an explicit density model

Entropic GANs

Entropic GANs are Wasserstein GANs with entropy regularization

Primal

$$\min_{P_{Y,G(X)}} \mathbf{E}\left[l(Y, G(X))\right] - H\left(P_{Y,G(X)}\right)$$

Dual

$$\min_{G} \max_{D_1,D_2} \mathbf{E}\left[D_1(Y)\right] - \mathbf{E}\left[D_2(G(X))\right] - \lambda \mathbf{E}_{P_Y \times P_{\hat{Y}}}\left[\exp v(\mathbf{y}, \hat{\mathbf{y}})/\lambda\right]$$

where $v(\mathbf{y}, \hat{\mathbf{y}}) := D_1(\mathbf{y}) - D_2(\hat{\mathbf{y}}) - l(\mathbf{y}, \hat{\mathbf{y}})$

↓

Loss function between two samples

Notation

Y	Real data random variable
X	Noise random variable
$G : \mathbb{R}^r \rightarrow \mathbb{R}^d$	Generator function
$\hat{Y} := G(X)$	

An Explicit data model for Entropic GAN

We construct an explicit probability model for data distribution using GANs

$$f_{Y|X=x} = C \exp(-l(\mathbf{y}, G(\mathbf{x}))/\lambda)$$

Normalization constant

Loss function used in Entropic GANs

Main theorem

$$\underbrace{E_{P_Y}[\log f_Y(Y)]}_{\text{Avg. log likelihoods}} \geq -\frac{1}{\lambda}\underbrace{\{E_{P_{Y,\hat{Y}}}[l(Y,\hat{Y})] - \lambda H(P_{Y,\hat{Y}})\}}_{\text{Entropic GAN objective}} + constants$$

Entropic GAN objective is a variational lower-bound of log likelihood

Similar to the evidence lower-bound in Variational Auto-Encoders

Main theorem

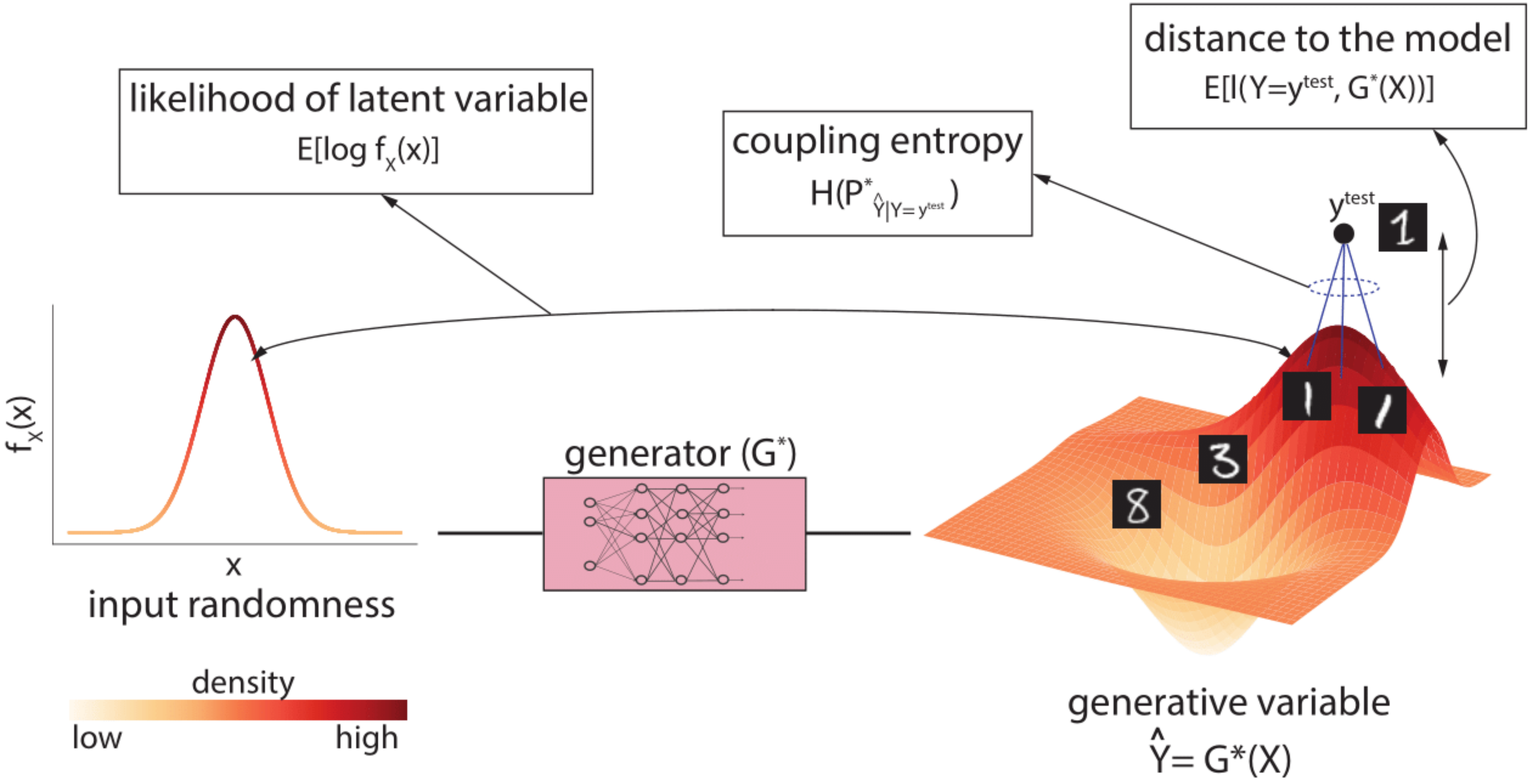
$$\underbrace{E_{P_Y}[\log f_Y(Y)]}_{\text{Avg. log likelihoods}} \geq -\frac{1}{\lambda}\{\underbrace{E_{P_{Y,\hat{Y}}}[l(Y,\hat{Y})] - \lambda H(P_{Y,\hat{Y}})}_{\text{Entropic GAN objective}}\} + constants$$

Entropic GANs meet VAEs

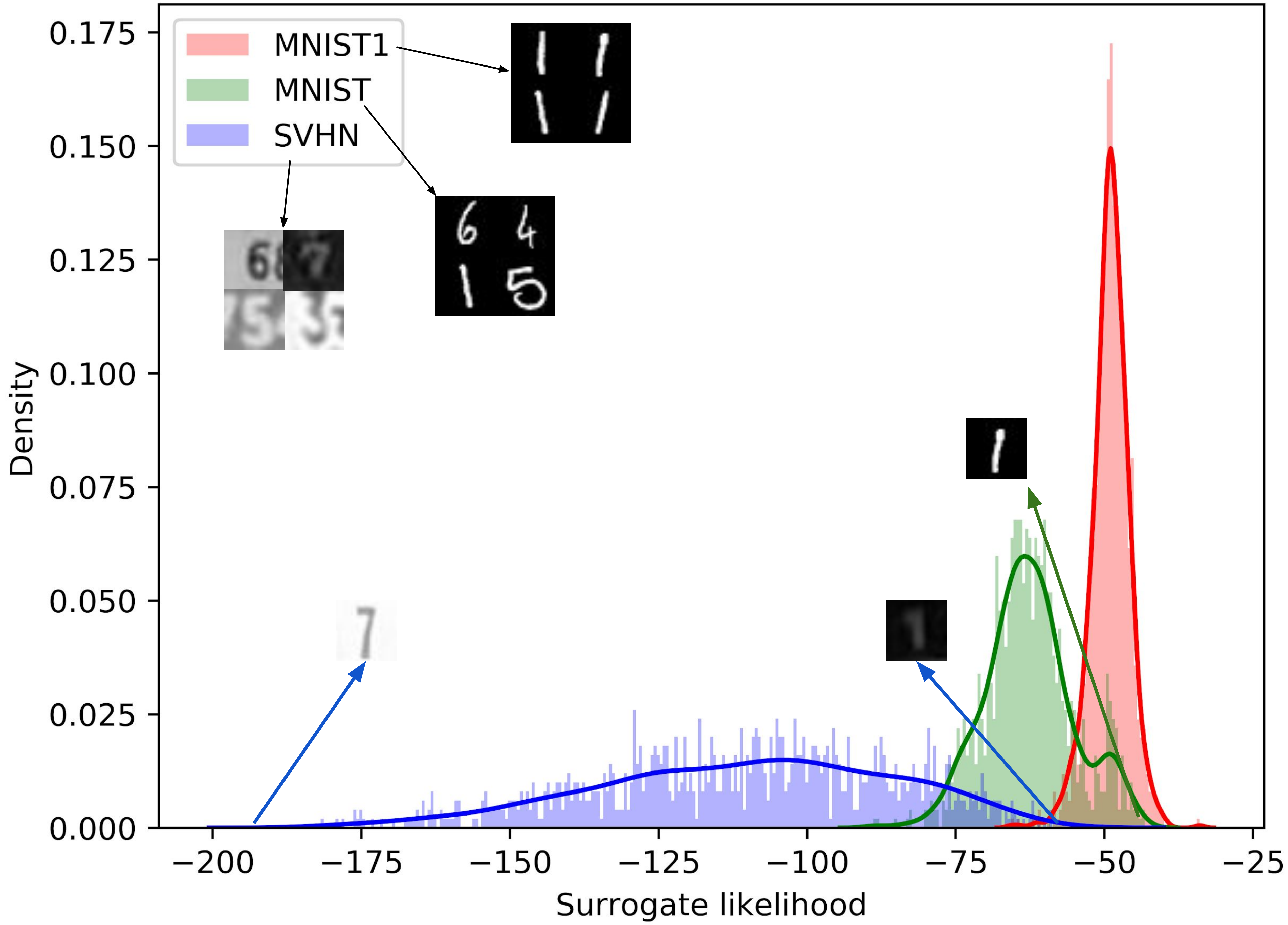
Is this bound useful?

- Provides a statistical interpretation to Entropic GAN's objective function
- Useful in computing sample likelihoods of test samples

Components of our surrogate likelihood



Likelihood computation



Given a dataset of MNIST-1 digits as source distribution, estimate the likelihood that MNIST and SVHN datasets are drawn from this distribution

Dissimilar datasets have low likelihood

Tightness of the lower-bound

We consider Gaussian input data distribution to compute the tightness of our variational lower-bound

Data dimension	Exact Log-Likelihood	Surrogate Log-Likelihood
5	-16.38	-17.94
10	-35.15	-43.60
20	-58.04	-66.58
30	-91.80	-100.69
64	-203.46	-217.52

Conclusion

Establish a connection between GANs and VAEs by deriving a variational lower-bound for GANs

Provide a principled framework for computing sample likelihoods using GANs

Please stop by Poster# 17

Code available at https://github.com/yogeshbalaji/EntropicGANs_meet_VAEs