Applications of Deep Generative Modeling in Finance

Bloomberg

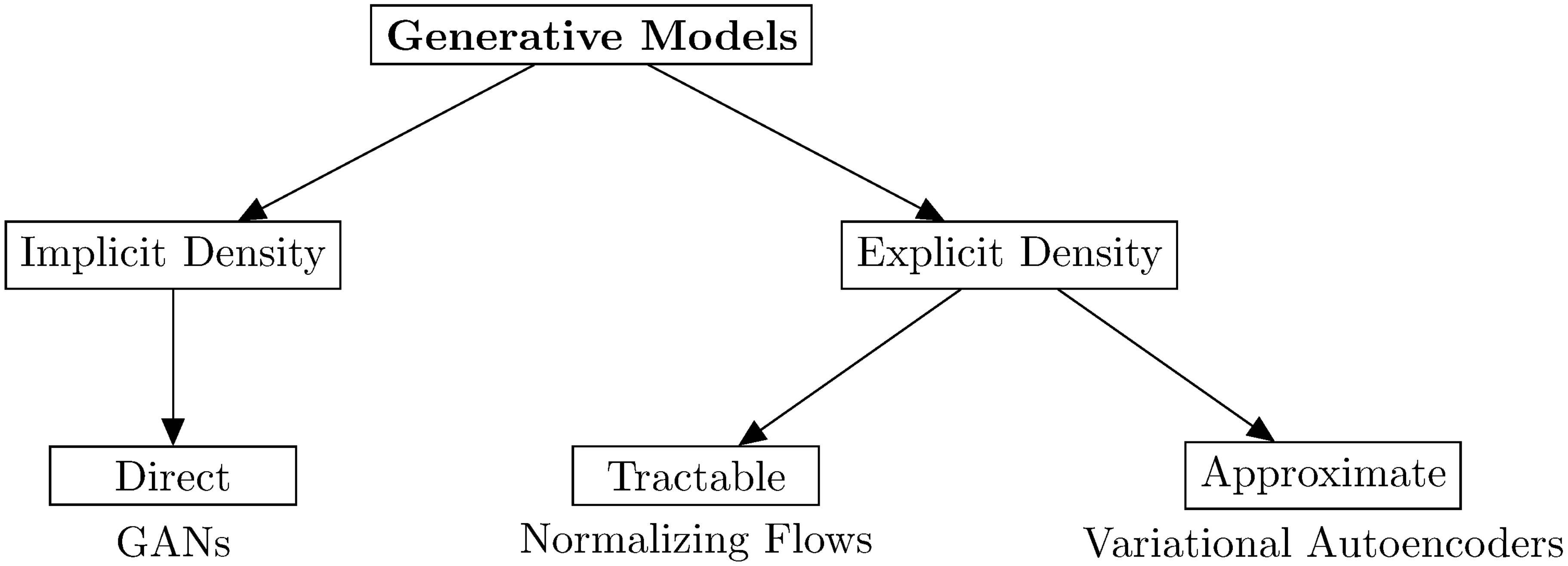
Machine Learning in Finance Workshop 2020
August 7, 2020

Achintya Gopal
Quant Researcher

TechAtBloomberg.com

What is Generative Modeling?

Goal: Any model from which we can sample from



Amortized Inference

Example: Carbon Modeling

Goal: Estimate a company’s carbon emissions

Data

- Annually Reported
- About 1,000 companies report

Estimation

- Distributional Estimates per Company per Year

Traditional Approach

- Partition (bucket) the companies into similar subsets
- Estimate Gamma GLM on subsets

Implicit Assumptions

- Every bucket is independent of every other bucket

Problems

- What about Country? Company size? Similar industries?

Amortized Inference Approach

Model

- Train one model on ALL data

$$\mathrm{Gamma}(k_\phi(X), \theta_\phi(X))$$

Loss

- Maximum Likelihood Estimation (MLE) instead of using Mean Squared Error

Why?

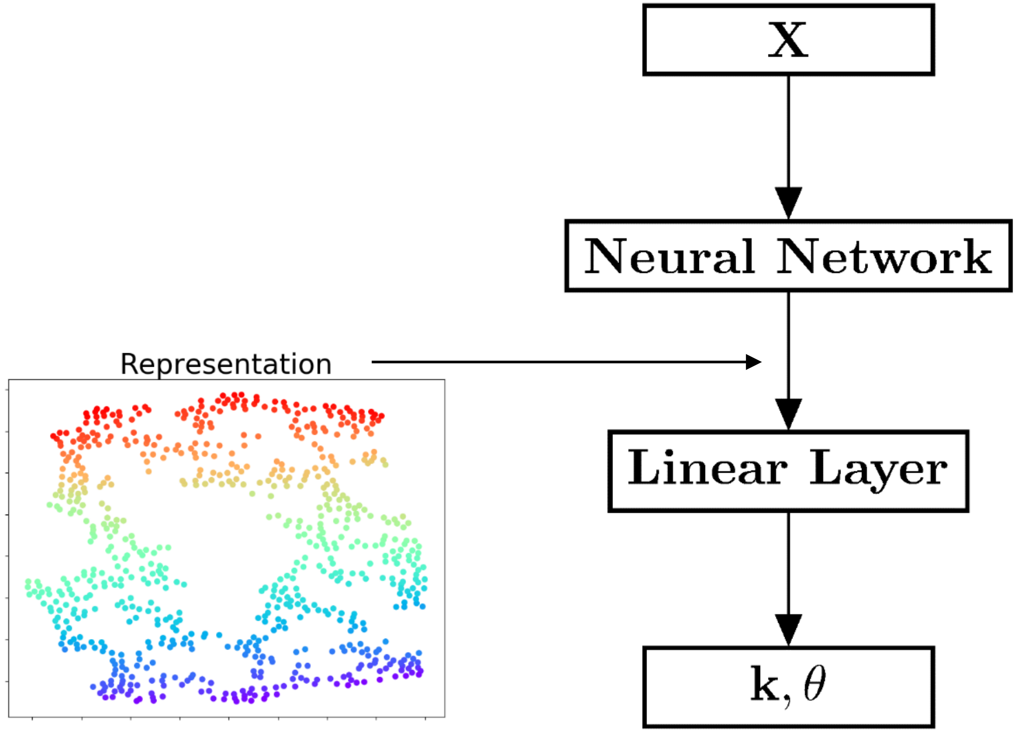
- Each company has hundreds of features
- Linear relationship is not justified for all features

Amortized Inference Benefits

1. Shares statistical strength across data

2. Implicitly buckets the data

3. Allows for more complex distributions
   — For example, multivariate distributions

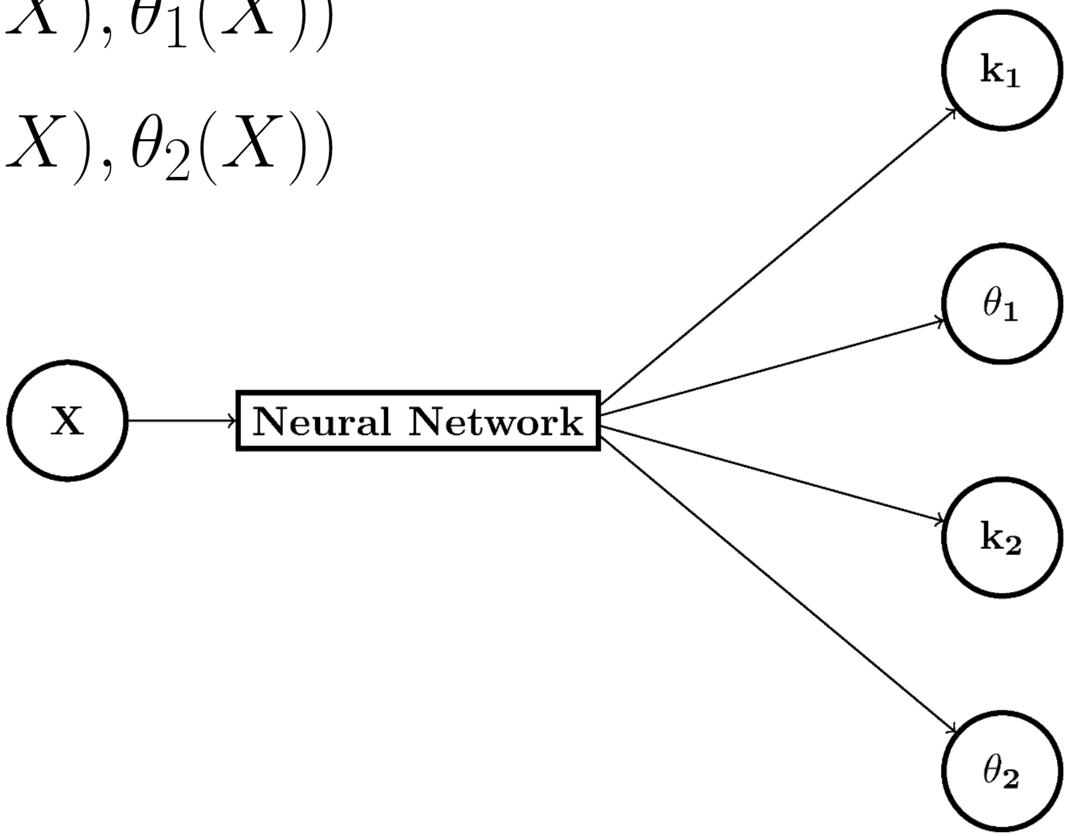


Example: Direct and Indirect Carbon Emissions

Generative Process:

$$\text{Direct Carbon Emissions} \sim \text{Gamma}(k_1(X), \theta_1(X))$$

$$\text{Indirect Carbon Emissions} \sim \text{Gamma}(k_2(X), \theta_2(X))$$



Example: LQA

Goal

- Estimate Trade Cost

$$C = \frac{P_{\text{trade}} - P_{\text{true}}}{P_{\text{true}}}$$

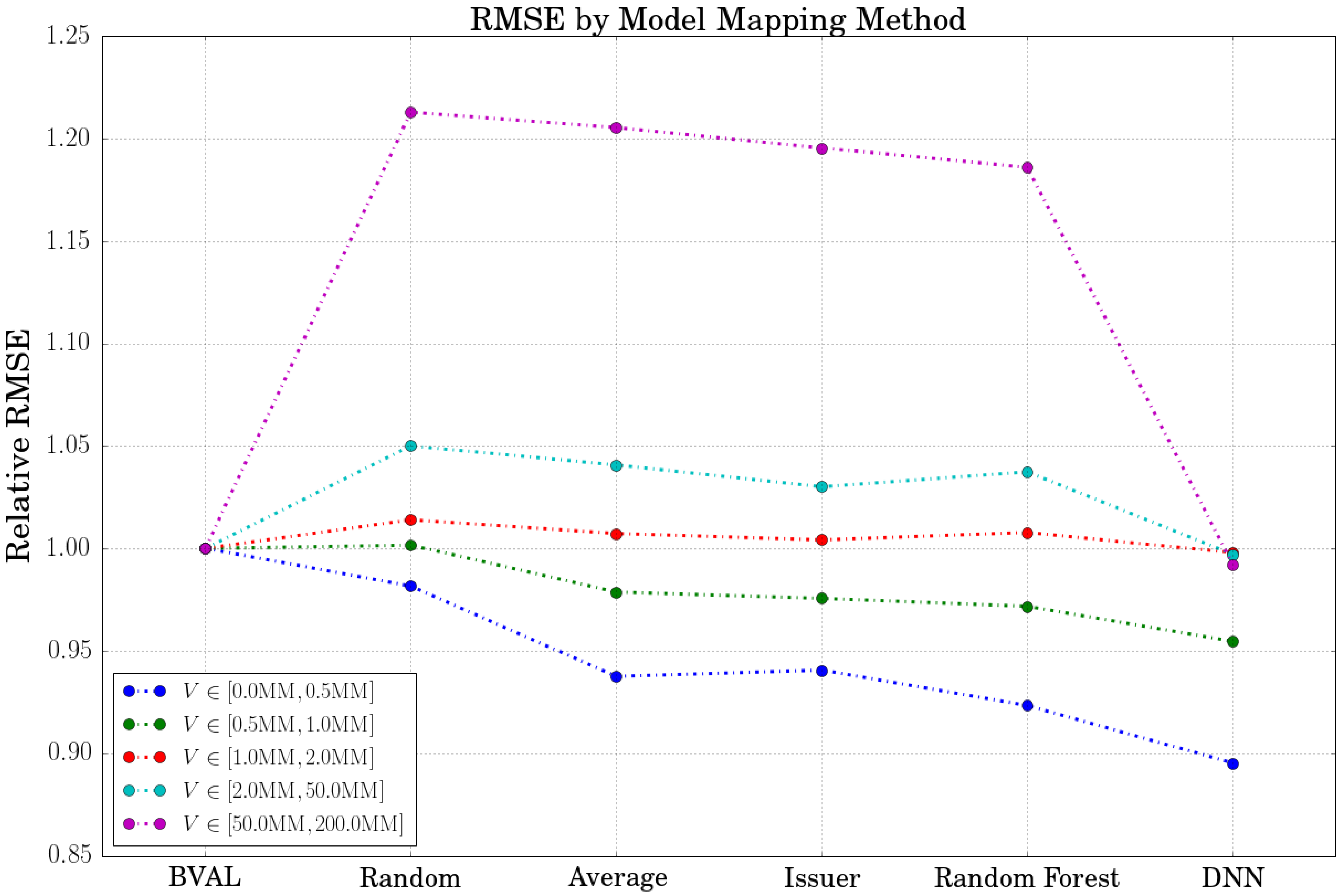
Data

- <1% of securities have more than 5 trades

Estimation

- Distributional Estimates

DNN Results



RMSE by Model Mapping Method
Relative RMSE
1.25
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
BVAL
Random
Average
Issuer
Random Forest
DNN
$V \in [0.0\text{MM}, 0.5\text{MM}]$
$V \in [0.5\text{MM}, 1.0\text{MM}]$
$V \in [1.0\text{MM}, 2.0\text{MM}]$
$V \in [2.0\text{MM}, 50.0\text{MM}]$
$V \in [50.0\text{MM}, 200.0\text{MM}]$

Interpretability

- What if we have a known relationship between some features and the target?

Example

- Carbon Emissions is proportional to Company Size (S)

Model

$$\mathrm{Gamma}(a_\phi(X) + S * b_\phi(X), \theta_\phi(X))$$

Generative Modeling: GANs

TechAtBloomberg.com

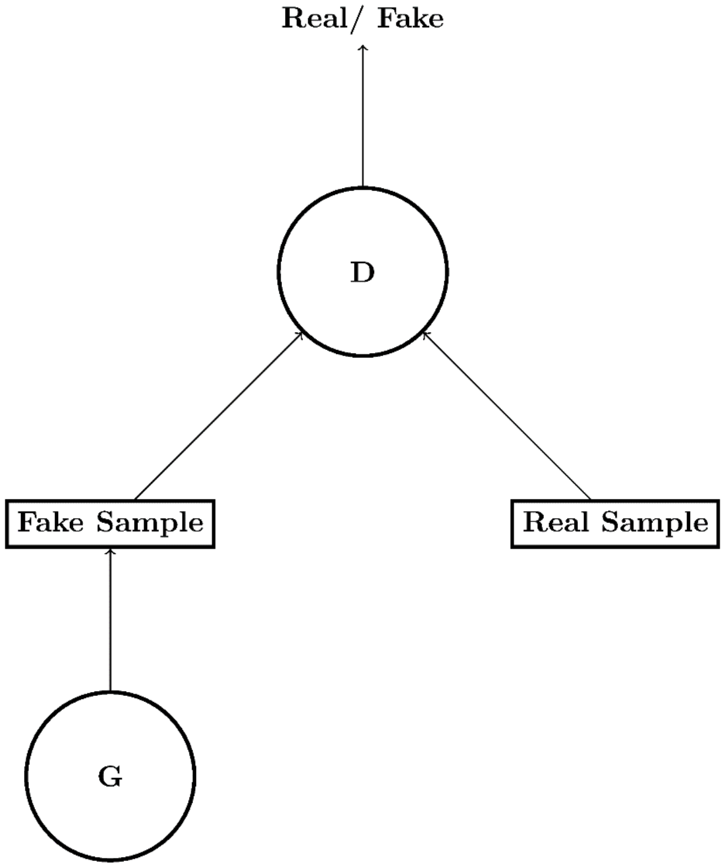
Likelihood Free Estimation

Limitations

- To perform MLE, we require explicit likelihood
- Many distributions are sums/mixtures of random variables without closed form likelihoods
- However, we CAN sample from the Generative Process

Generative Adversarial Networks (GANs)

- (G)enerator is trying to trick the (D)iscriminator
- Effectively minimizes JS Divergence



$$\min_G \max_D \left[\mathbb{E}_{x \sim p_{real}}[\log D(x))] + \mathbb{E}_{x \sim p_{fake}}[\log(1 - D(x)))] \right]$$

Example: Directionless Cost Estimation

Goal

- Trade cost estimation

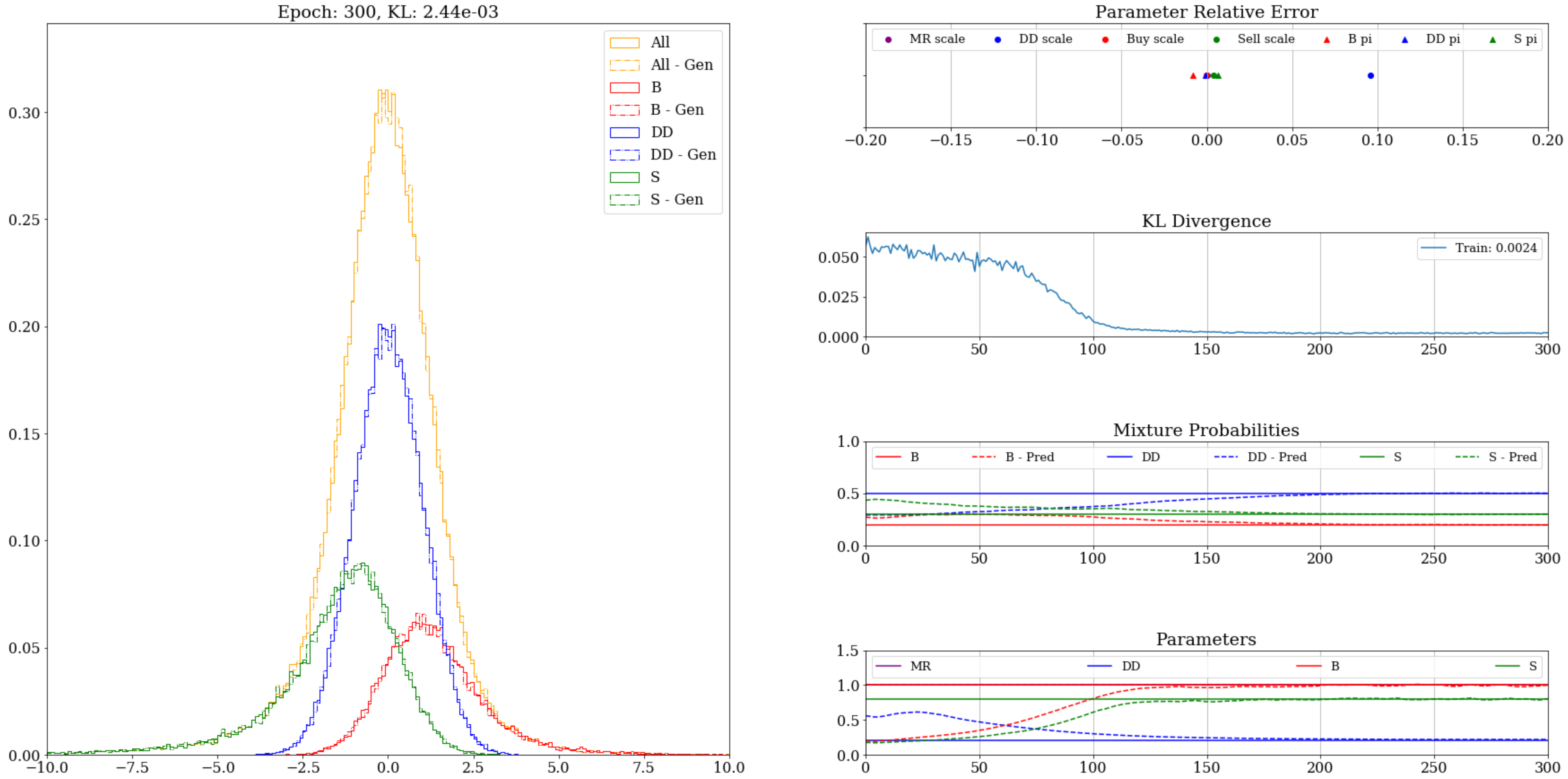
Data

- Missing Trade Direction

Estimation

- Mixture of Three Components
  - Mixture Weights: (π)
  - Dealer to Dealer: Normal(0, σ_d) + Normal(0, σ_{mr})
  - Dealer to Client (Buy): LogNormal(0, σ_b) + Normal(0, σ_{mr})
  - Client to Dealer (Sell): -LogNormal(0, σ_s) + Normal(0, σ_{mr})

Analytic GAN Example



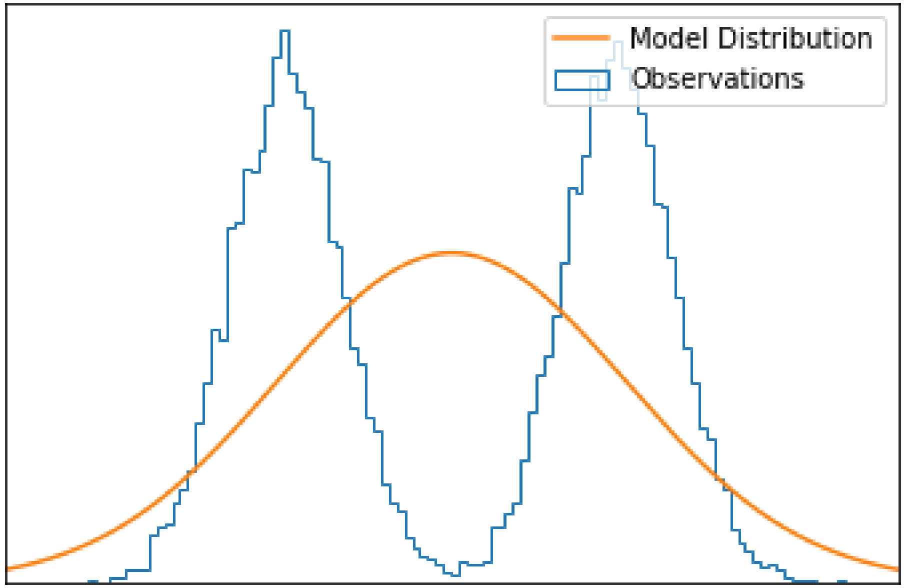
Generative Modeling: Normalizing Flows

TechAtBloomberg.com

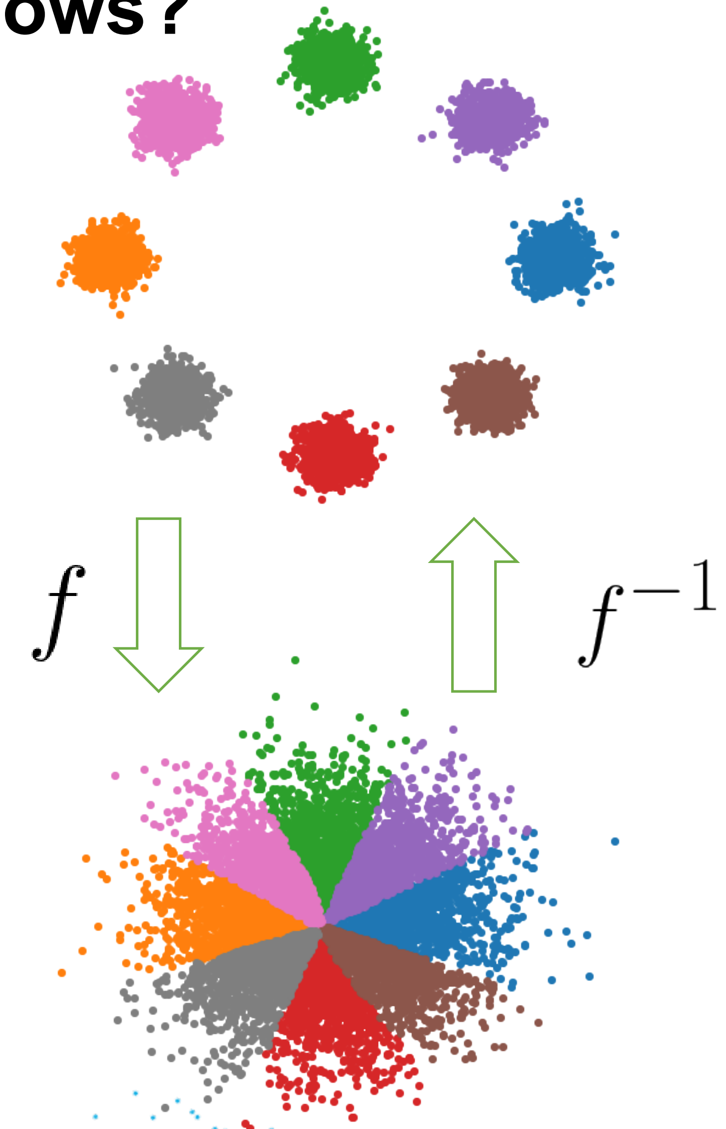
More Flexible Distributions

Limitations

- The output distribution limits the flexibility of the model to describe the data
- If we have no theoretical assumptions for distribution, we should be able to use any free form distribution
- Normalizing Flows are deep learning method for modeling complex distributions

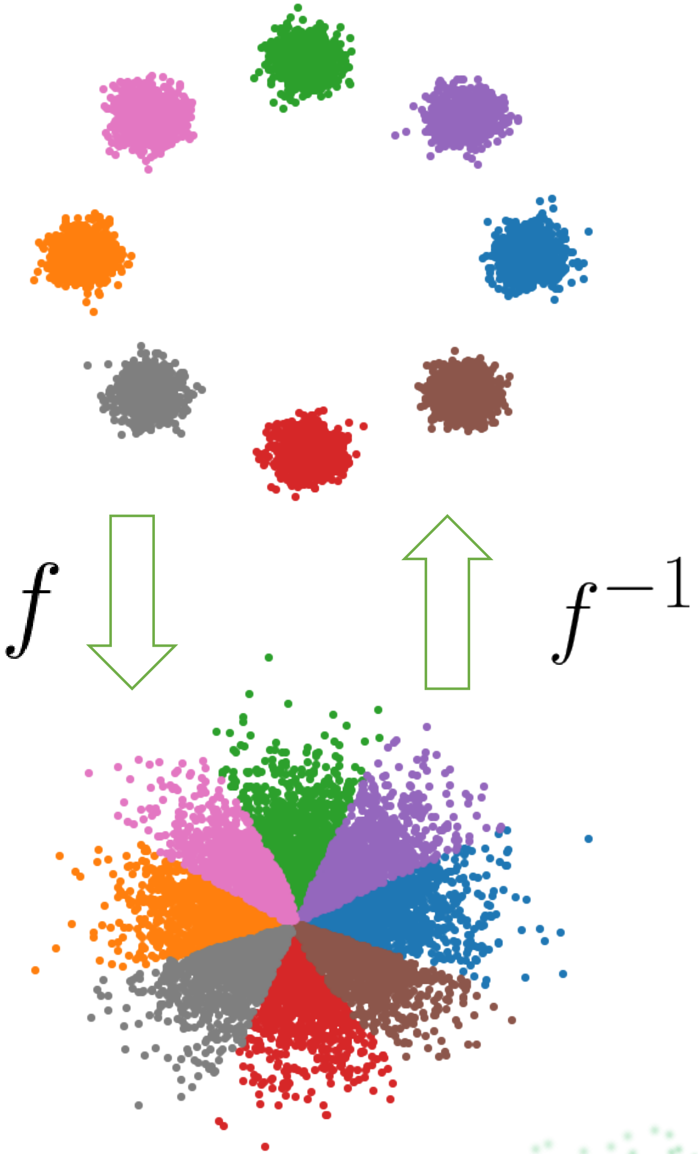


What are Normalizing Flows?

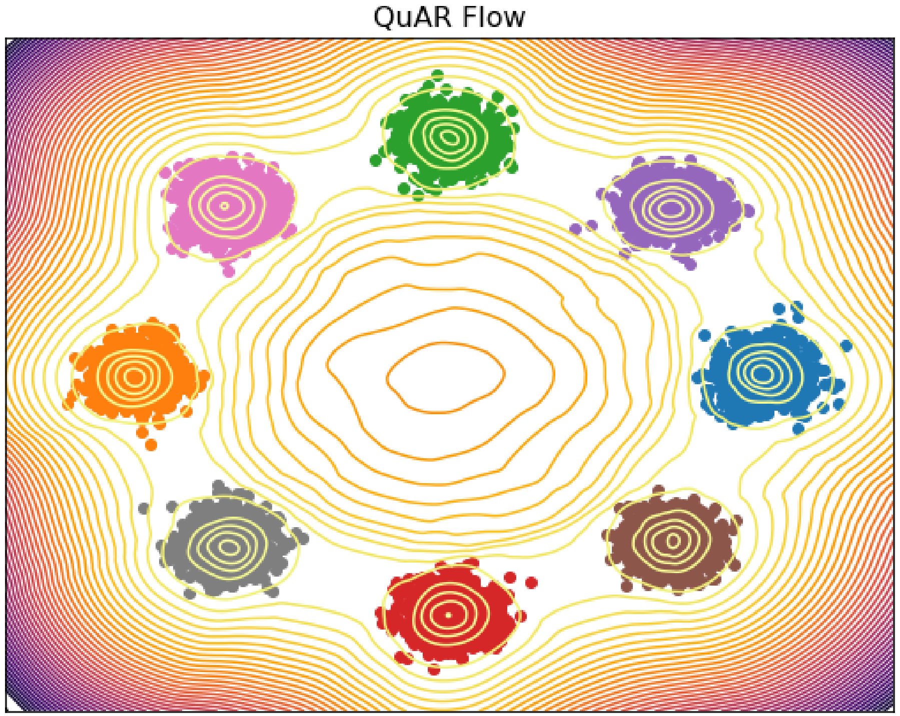


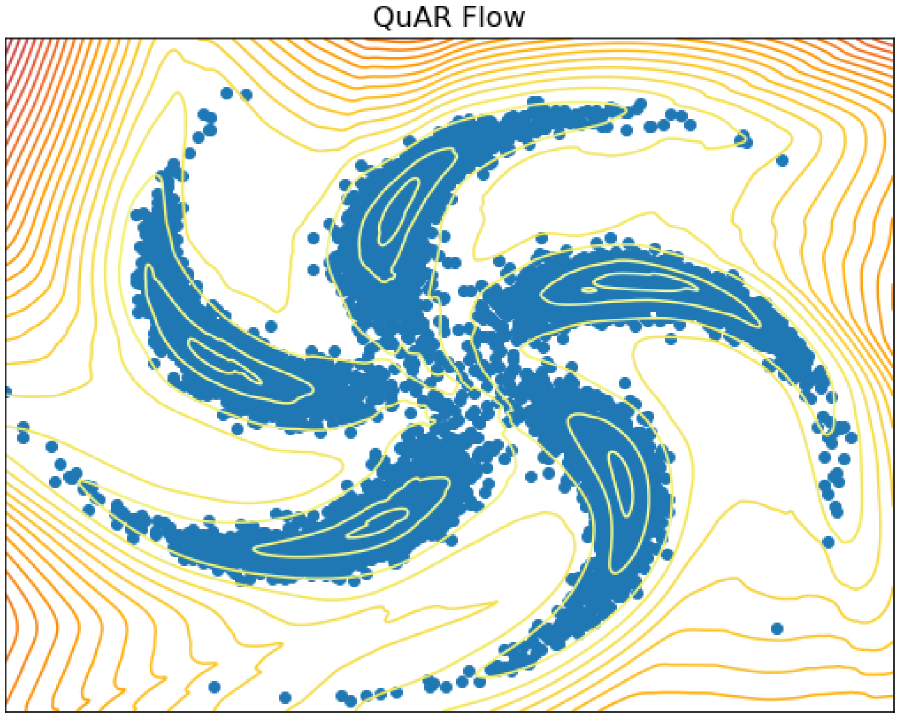
What are Normalizing Flows?

$$\log p(x) = \log p(f(x)) + \log \left| \frac{\partial f(x))}{\partial x} \right|$$



Example Distributions



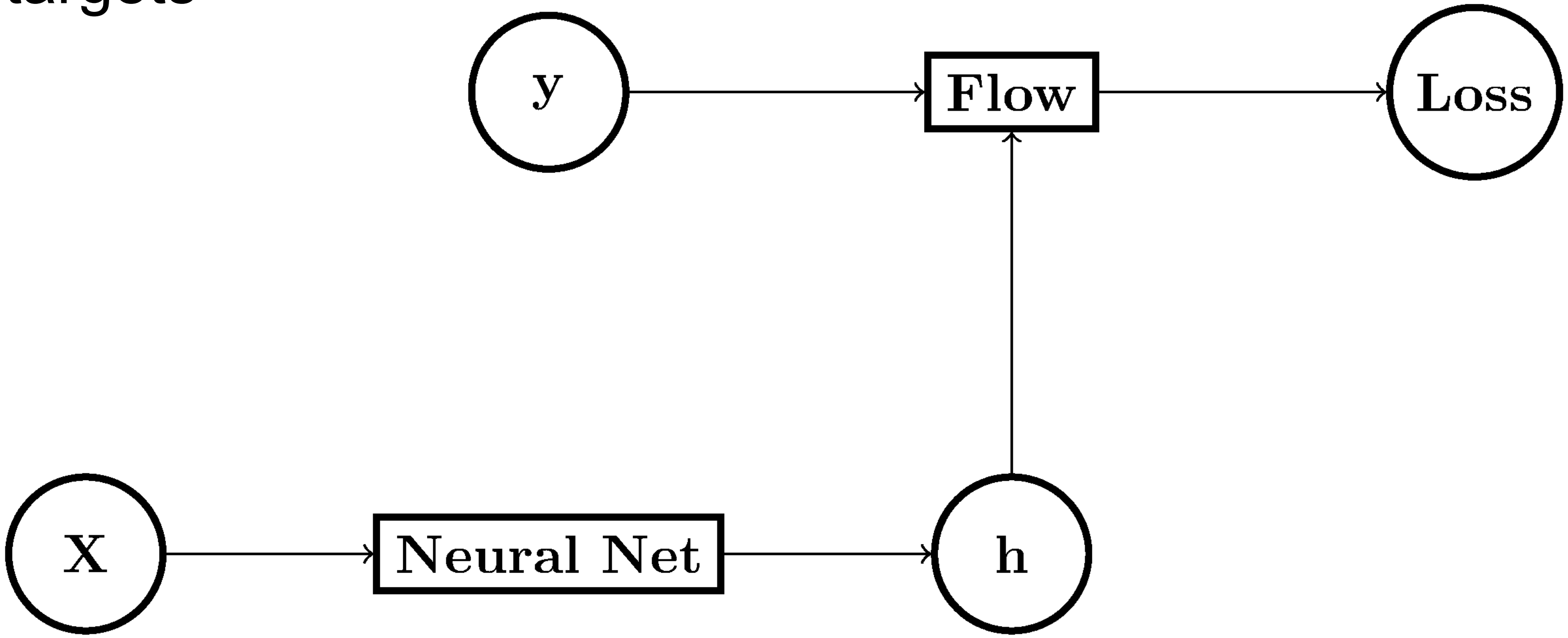


Reference: QuAR Flows

Conditional Flow

Benefits

- More Data Driven Output Distribution
- Model high dimensional targets
  —Model complex correlations



Generative Modeling: Variational Autoencoders

TechAtBloomberg.com

Variational Inference

Goal: Approximate posterior to perform MLE

Traditional Technique

1. Approximate posterior for all data points
2. Optimize parameters to maximize approximate likelihood

Example: Mixture of Gaussians

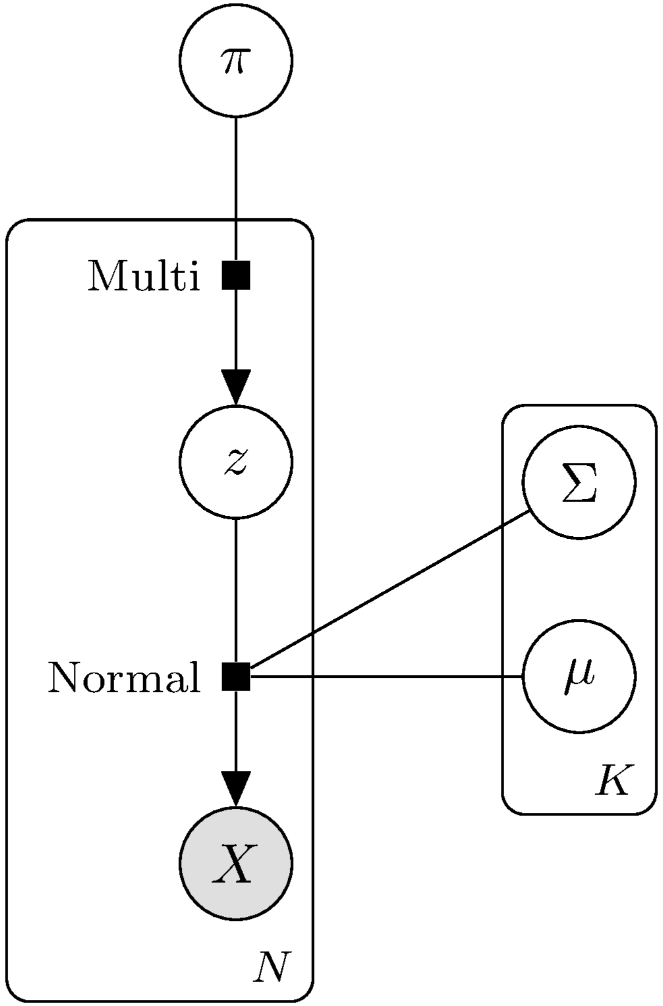
Generative Process:

$$z \sim \text{Categorical}(\pi)$$

$$X|z \sim \text{Normal}(\mu_z, \Sigma_z)$$

Approximate Posterior:

$$z \sim \text{Categorical}(\theta(X))$$



$$\log p(x) \geq \mathbb{E}_{z \sim q(z|x)} \left[\log p(x|z) + \log p(z) - \log q(z|x)\right]$$

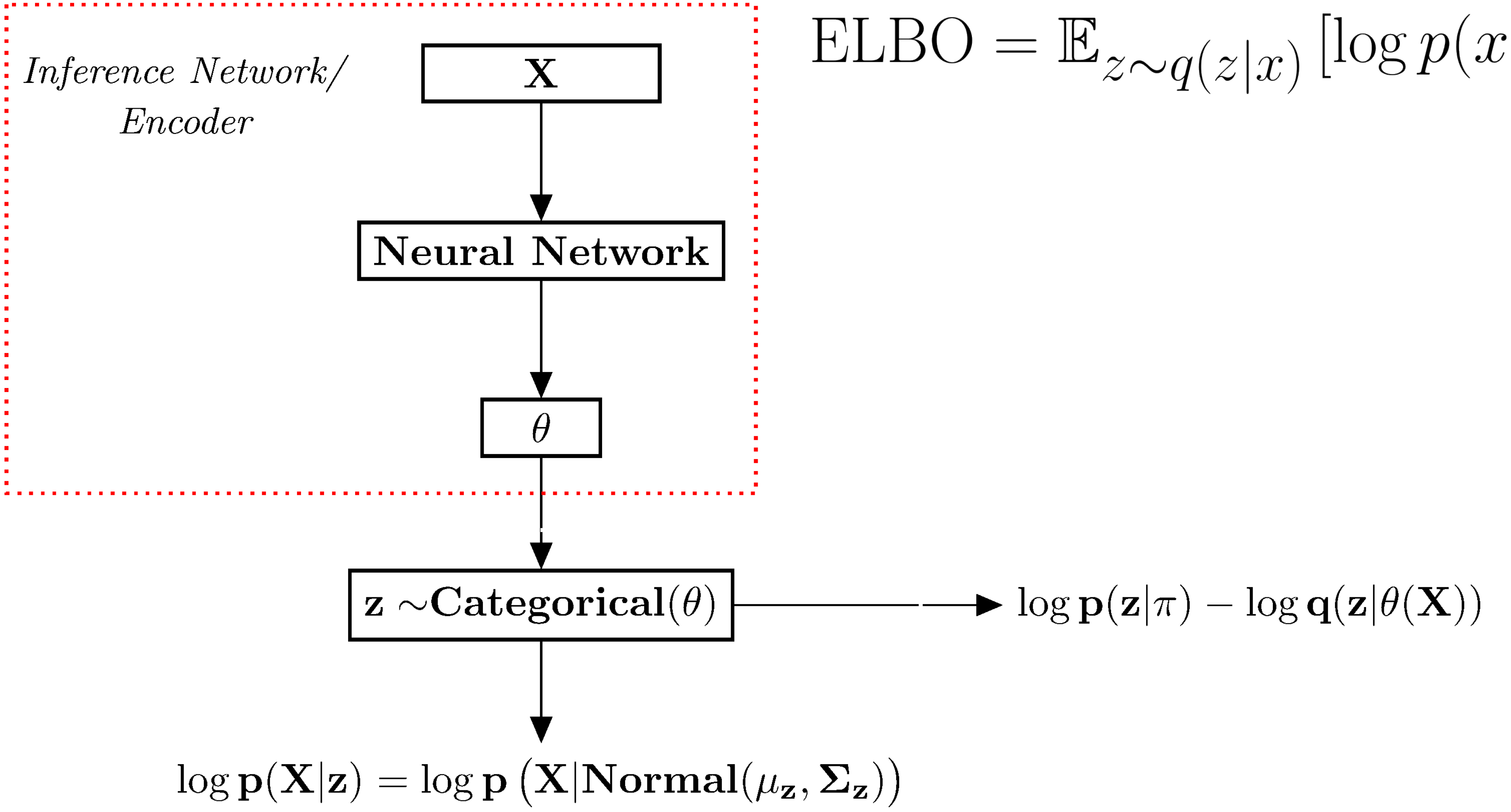
Variational Autoencoder

$$\text{ELBO} = \mathbb{E}_{z \sim q(z|x)}\left[\log p(x|z) + \log p(z) - \log q(z|x)\right]$$

$$\boxed{\mathbf{X}}$$
$$\downarrow$$
$$\boxed{\textbf{Neural Network}}$$
$$\downarrow$$
$$\boxed{\theta}$$
$$\downarrow$$
$$\boxed{\mathbf{z} \sim \textbf{Categorical}(\theta)} \longrightarrow \log \mathbf{p}(\mathbf{z}|\pi) - \log \mathbf{q}(\mathbf{z}|\theta(\mathbf{X}))$$
$$\downarrow$$
$$\log \mathbf{p}(\mathbf{X}|\mathbf{z}) = \log \mathbf{p}\left(\mathbf{X}|\textbf{Normal}(\mu_{\mathbf{z}}, \boldsymbol{\Sigma}_{\mathbf{z}})\right)$$

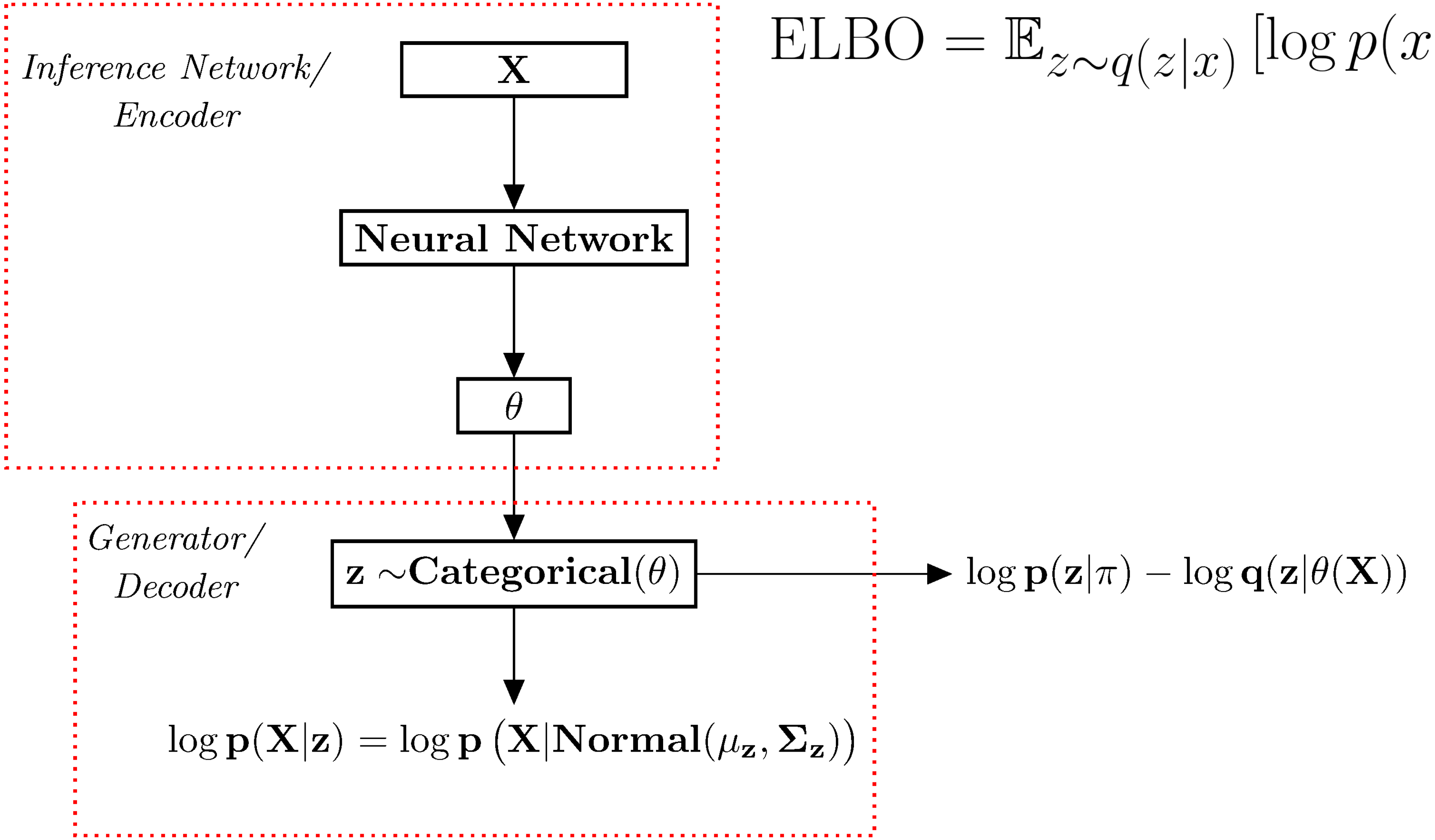
Variational Autoencoder

$$\text{ELBO} = \mathbb{E}_{z \sim q(z|x)} \left[\log p(x|z) + \log p(z) - \log q(z|x) \right]$$

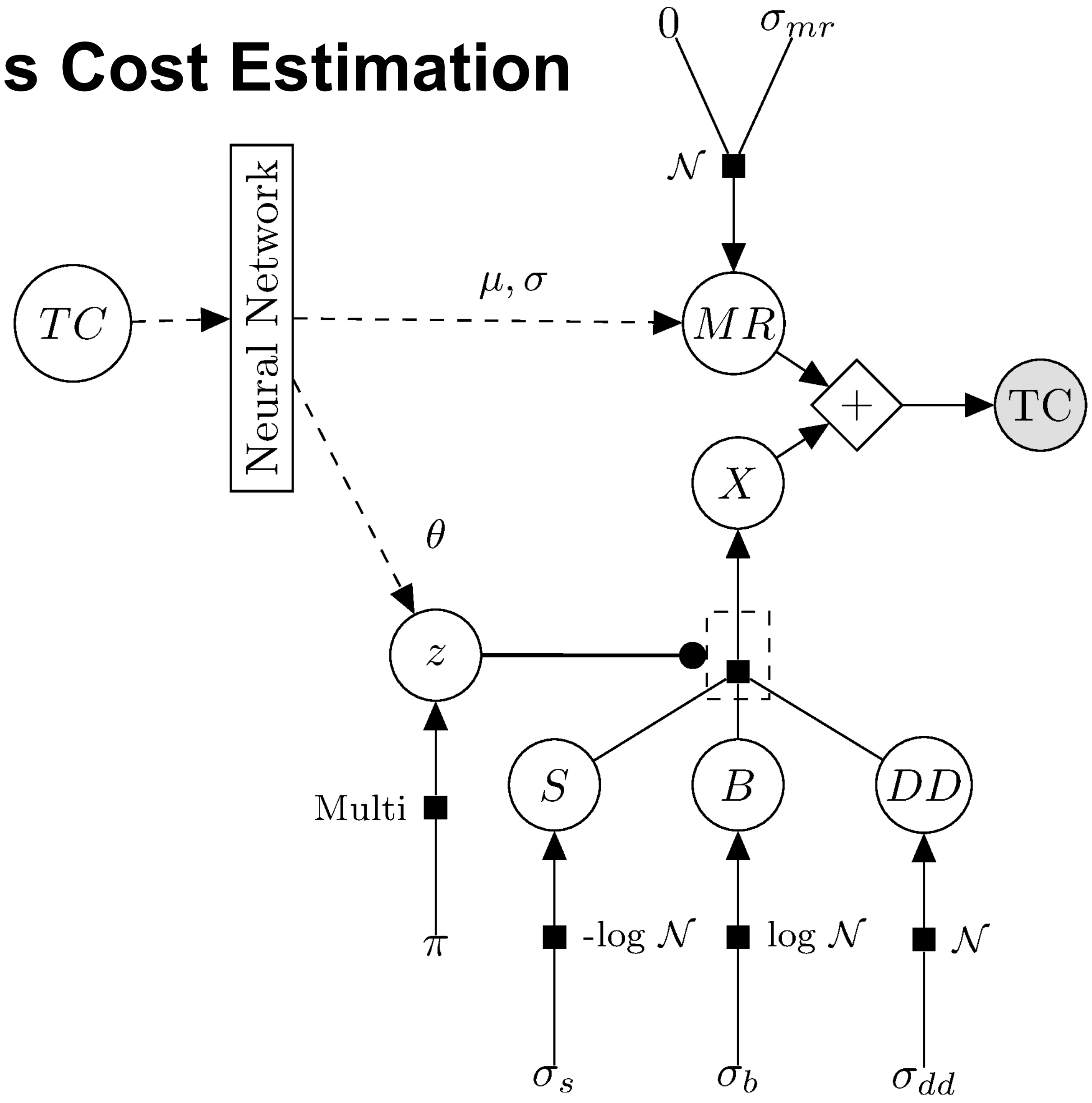


Variational Autoencoder

$$\mathrm{ELBO} = \mathbb{E}_{z \sim q(z|x)} \left[\log p(x|z) + \log p(z) - \log q(z|x)\right]$$



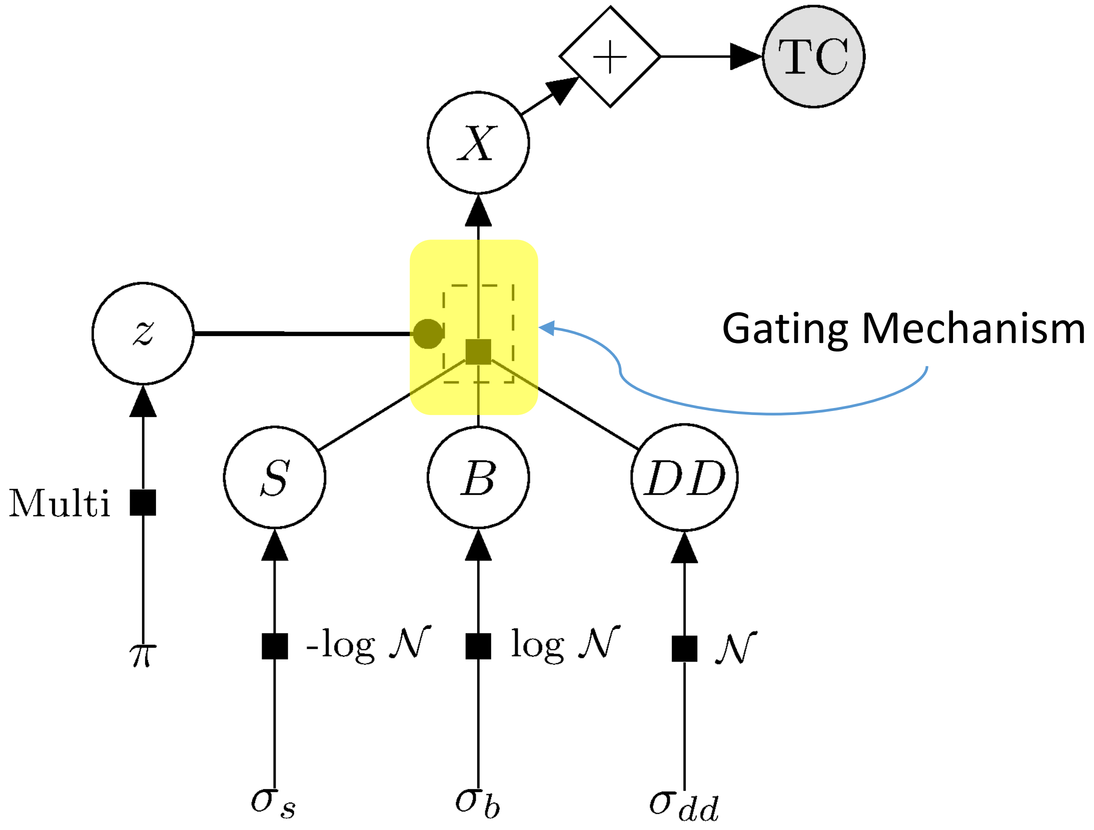
Example: Directionless Cost Estimation



Example: Directionless Cost Estimatio

Mixture of Three Compon

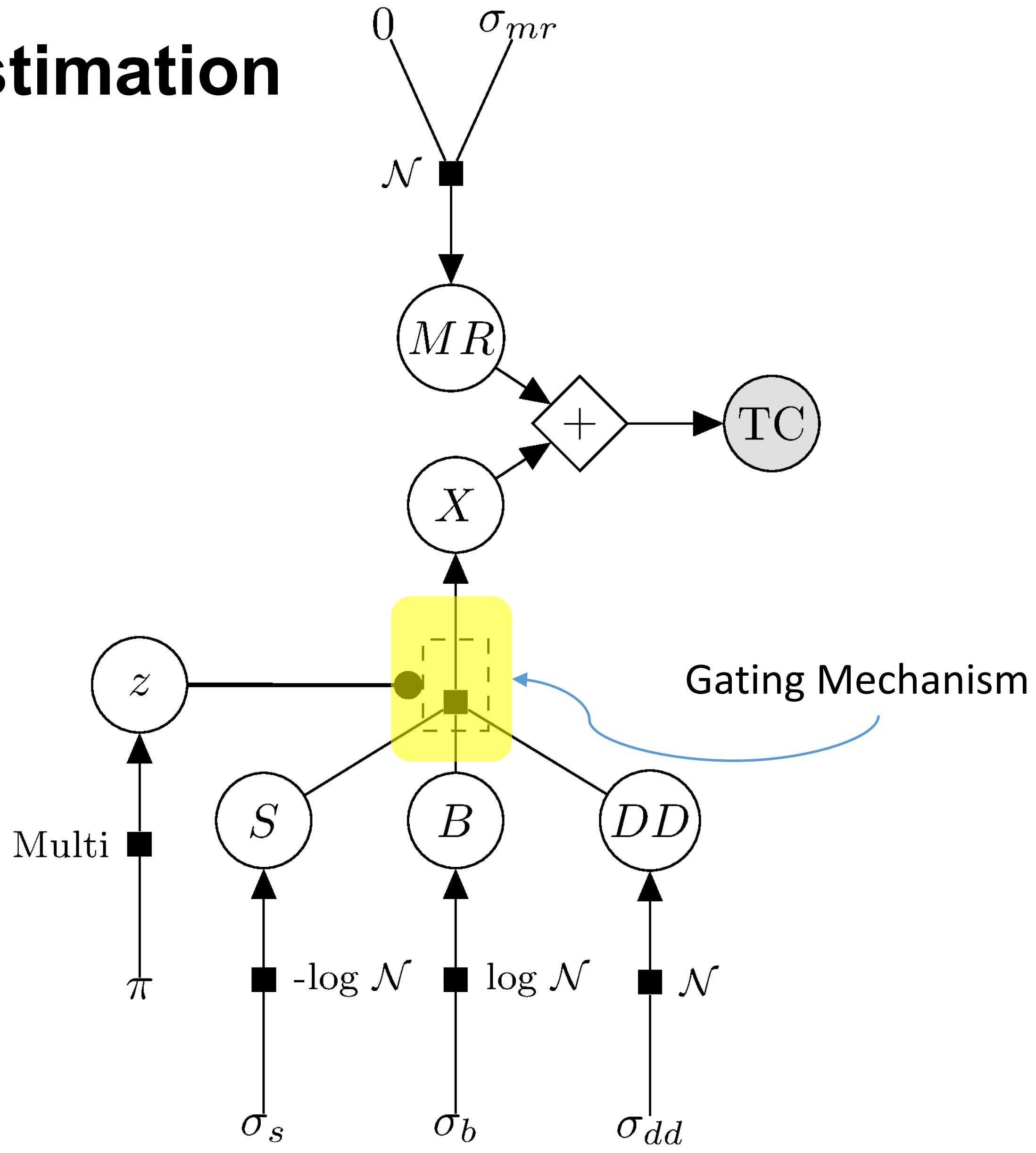
- Dealer to Dealer (DD)
- Dealer to Client (B)
- Client to Dealer (S)



Example: Directionless Cost Estimation

Mixture of Three Compon
- Dealer to Dealer (DD)
- Dealer to Client (B)
- Client to Dealer (S)

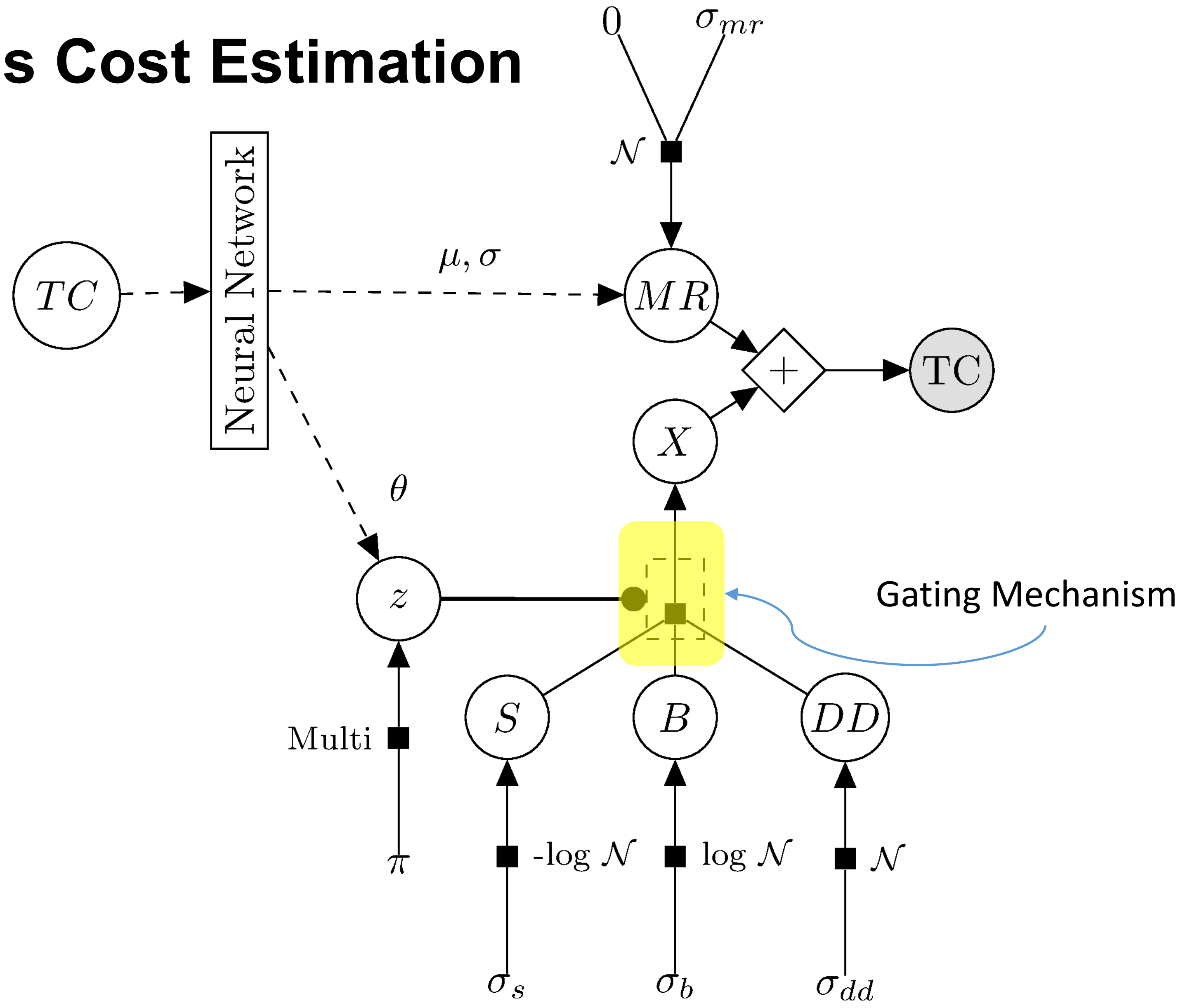
Market Risk Component (M



Example: Directionless Cost Estimation

Inference Network

- Categorical for z
- Normal for MR



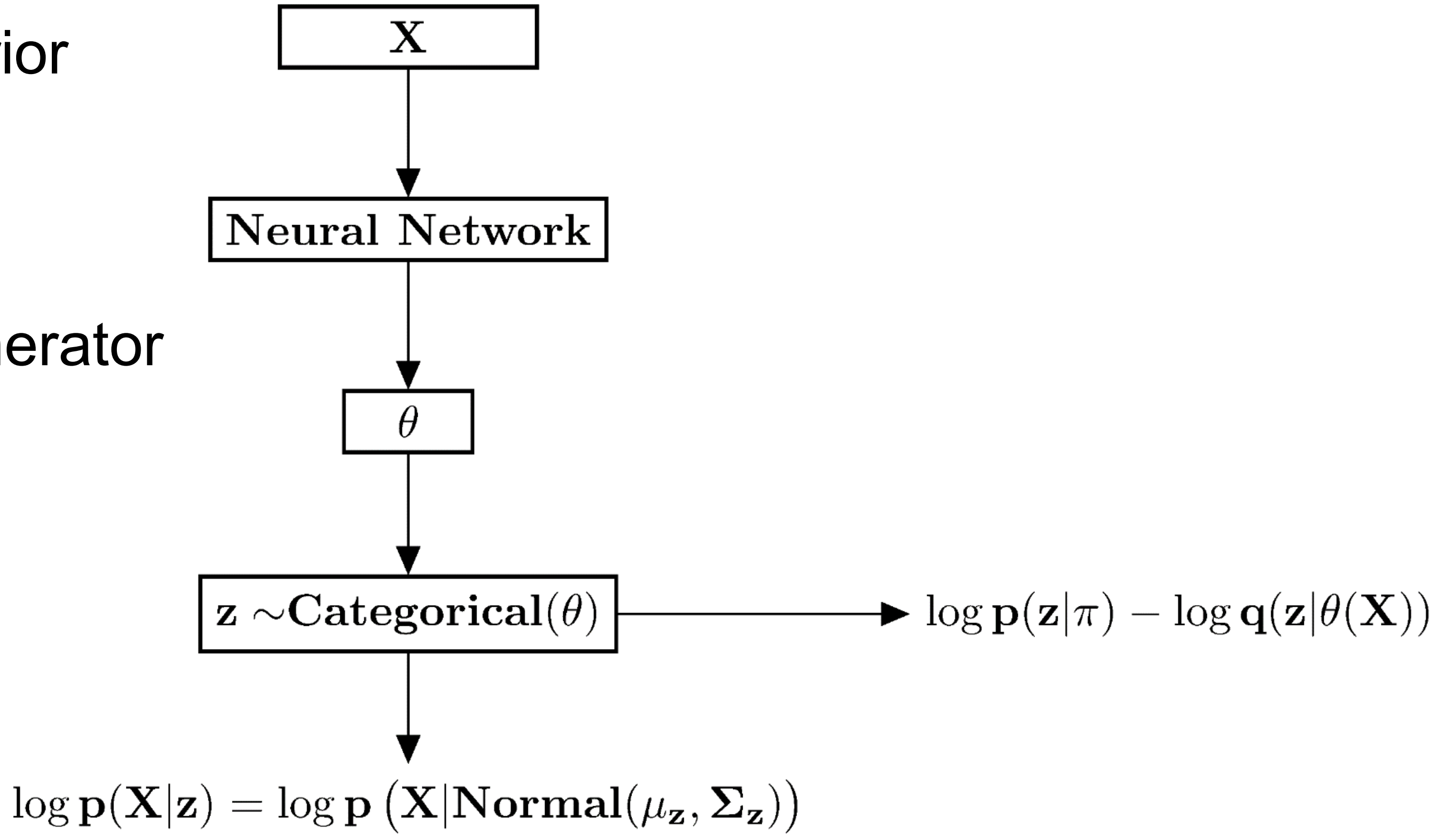
Variational Autoencoder

Benefits

- Uses less memory for posterior
- Utilize other VAE research

Observation

- Capacity is controlled by generator



Summary

- Amortized Inference is a lens to view how deep learning can be applied to generative processes

- Distilling Deep Learning Research
  - GAN: Likelihood Free Optimization of Generative Processes
  - Normalizing Flows: Deep Learning Approach to Learning Complex Distributions
  - VAEs: Deep Learning Approach to Enhance Variational Inference

Thank you!

https://www.bloomberg.com/careers

TechAtBloomberg.com