



Generative AI for Inverse Problems in Nuclear and Particle Physics

Yaohang Li
Department of Computer Science
Old Dominion University

yaohang@cs.odu.edu

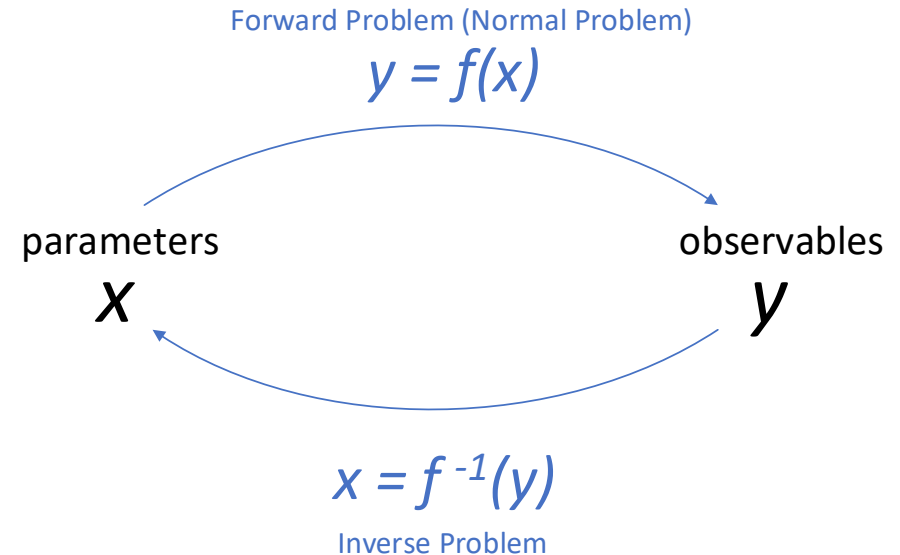
11/6/2025 @ University of Genoa, Italy

Agenda

- The Inverse Problem
- The Machine Learning Architectures
 - Variational Autoencoder Inverse Mapper
 - Diffusion Model
- Applications in Nuclear and Particle Physics
 - QCF Parameterization from DIS
 - Extraction of Crompton Form Factors as an Inverse Problem
 - Event-level Analysis

The Inverse Problem

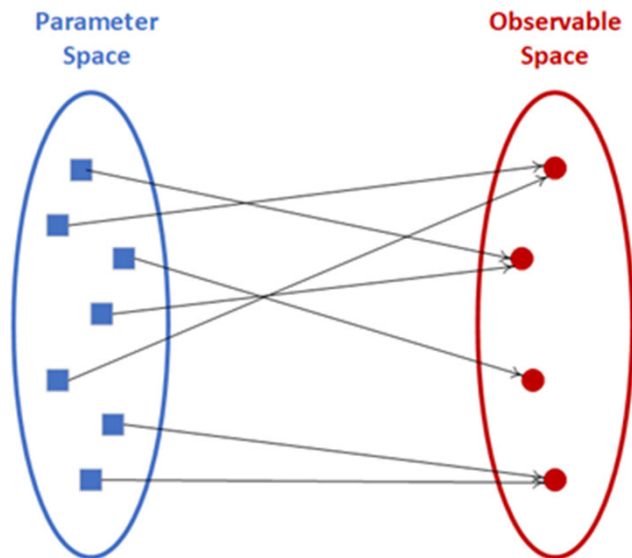
- The Forward Problem (Normal Problem)
 - Use the model parameters to calculate the observables
- The Inverse Problem
 - Use the results of actual observables to infer the values of the parameters characterizing the system
- Challenges
 - Ill-posedness
 - Different values of the model parameters may be consistent with the observables
 - Curse of Dimensionality
 - Need to explore a huge, high-dimensional parameter space



Ill-posedness of Inverse Problems

Forward Mapper

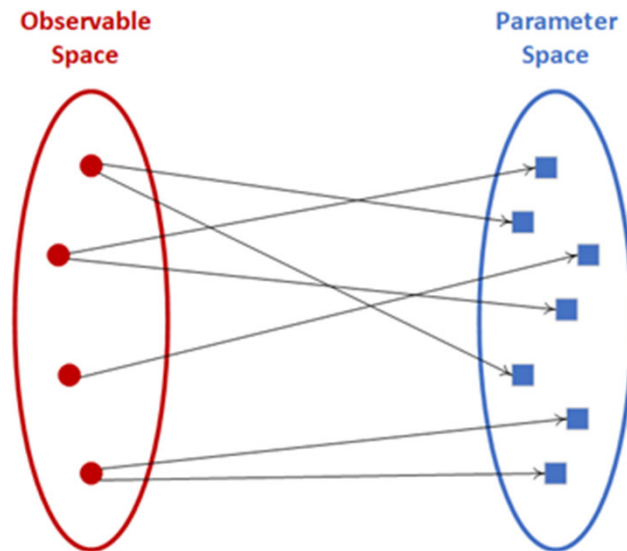
well-posed



$$\mathbf{y} = f(\mathbf{x})$$

Backward Mapper

ill-posed



$$\mathbf{x} = f^{-1}(\mathbf{y})$$

Information Loss

Agenda

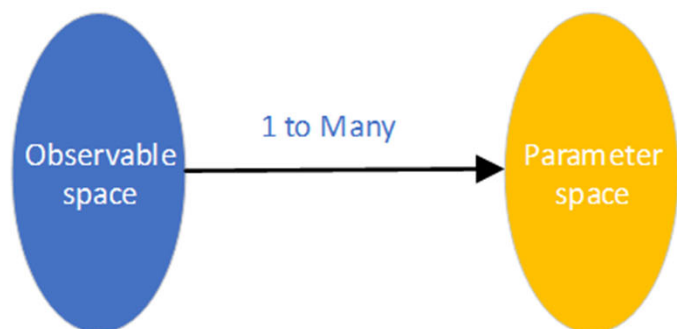
- The Inverse Problem
- The Machine Learning Architectures
 - Variational Autoencoder Inverse Mapper
 - Diffusion Model
- Applications in Nuclear and Particle Physics
 - Extraction of Crompton Form Factors as an Inverse Problem
 - Event-level Analysis

Fundamental Idea

Variational Autoencoder **Inverse Mapper**

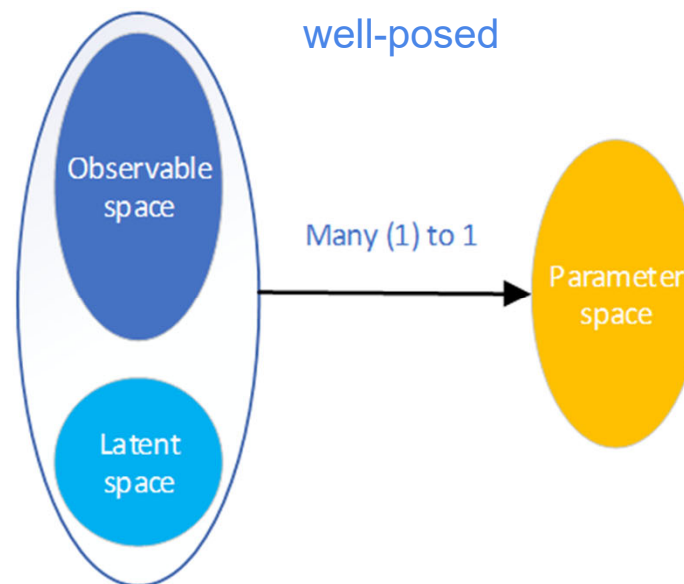
Inverse Problem

ill-posed



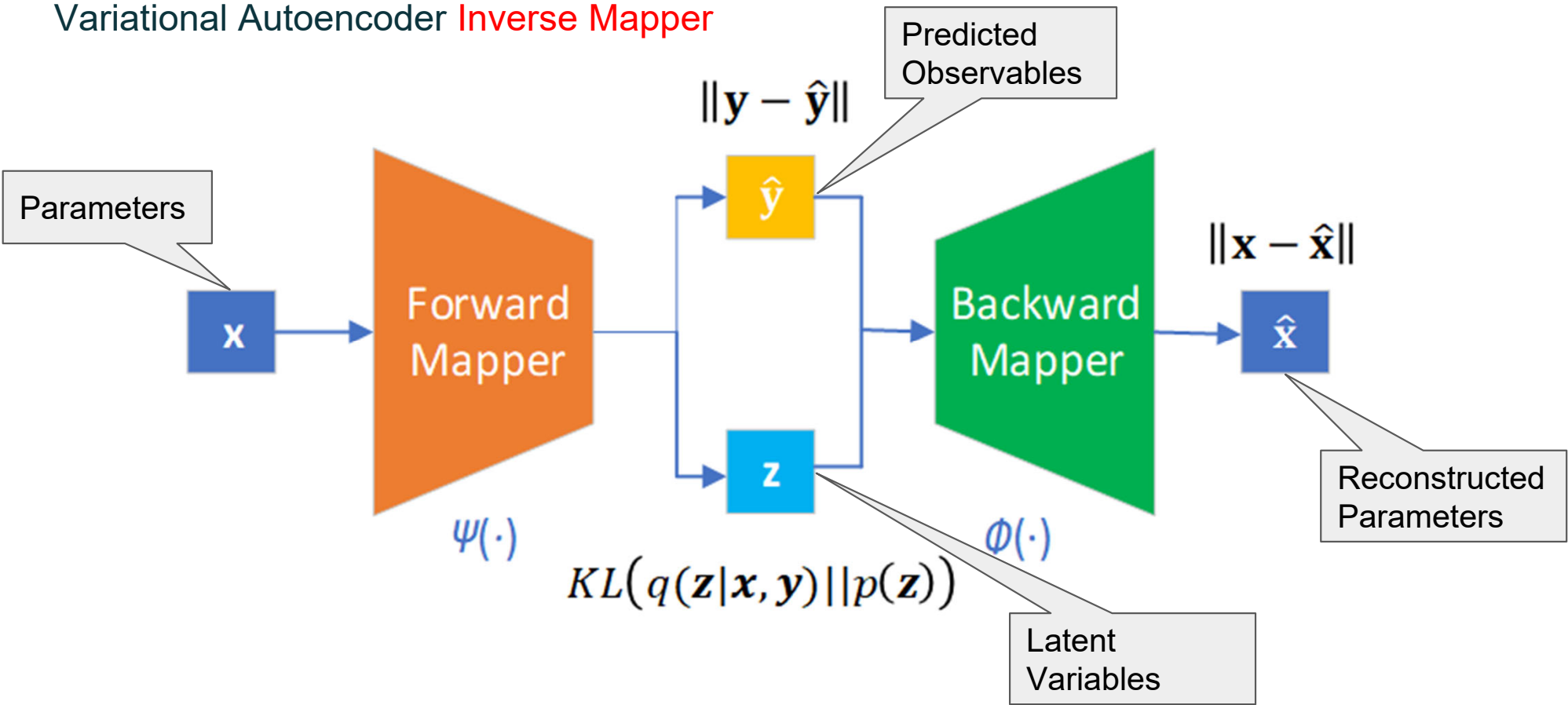
VAIM

well-posed

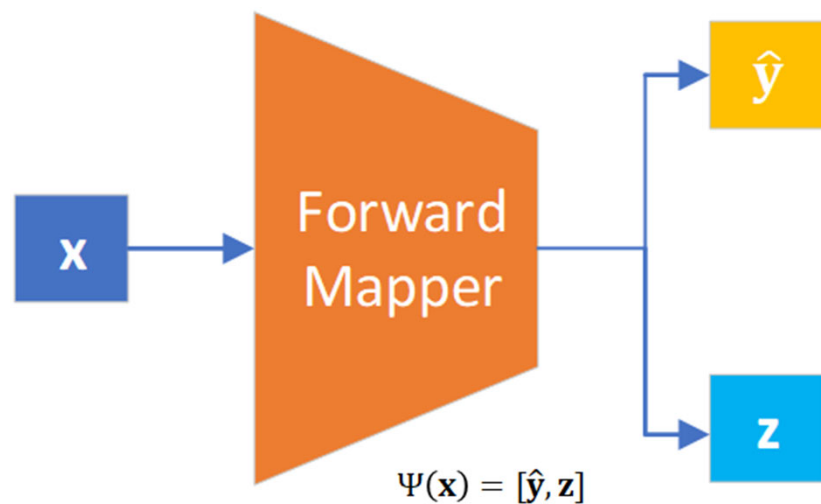


Variational Autoencoder Inverse Mapper Architecture

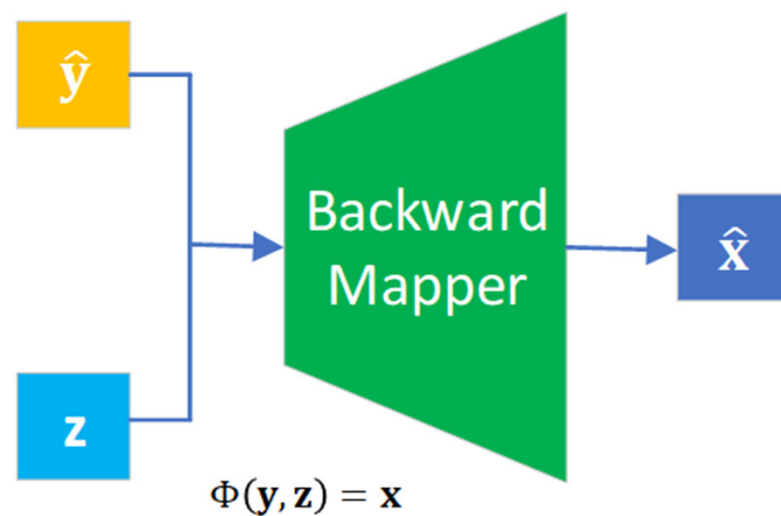
Variational Autoencoder Inverse Mapper



Forward Mapper and Backward Mapper



Learn posterior distribution
 $p(\mathbf{z}|\mathbf{x}, \mathbf{y})$



Learn likelihood distribution
 $p(\mathbf{x}, \mathbf{y}|\mathbf{z})$

Math behind Variational Autoencoder **Inverse Mapper**

- Approximate

- True posterior distribution $p(\mathbf{z}|\mathbf{x}, \mathbf{y})$

- Variational Inference

- Learn an approximate distribution $q(\mathbf{z}|\mathbf{x}, \mathbf{y})$ such that
$$q(\mathbf{z}|\mathbf{x}, \mathbf{y}) \sim p(\mathbf{z}|\mathbf{x}, \mathbf{y})$$
- Minimize the Kullback-Leibler (KL) divergence

$$\min KL(q(\mathbf{z}|\mathbf{x}, \mathbf{y}) || p(\mathbf{z}|\mathbf{x}, \mathbf{y}))$$

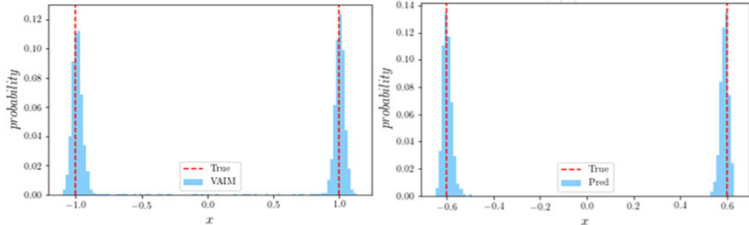
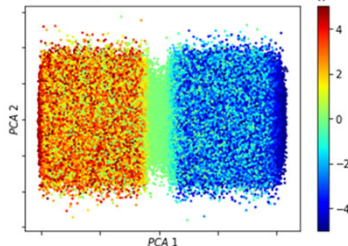
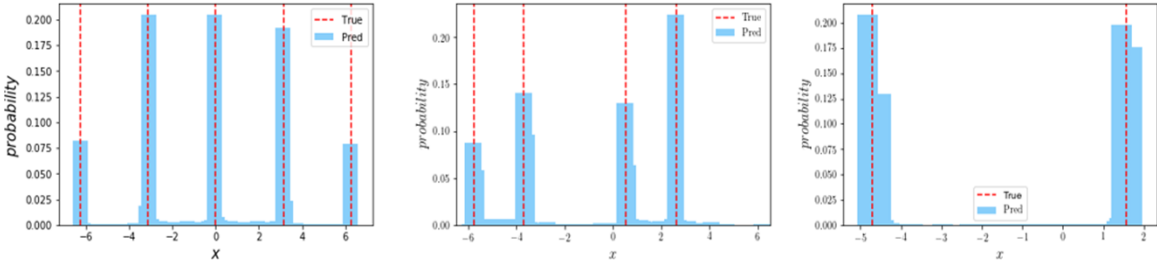
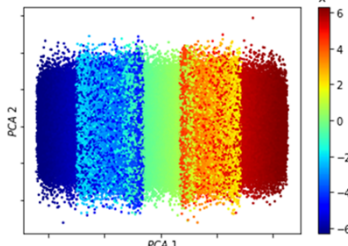
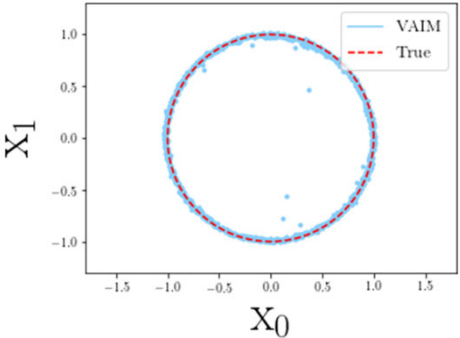
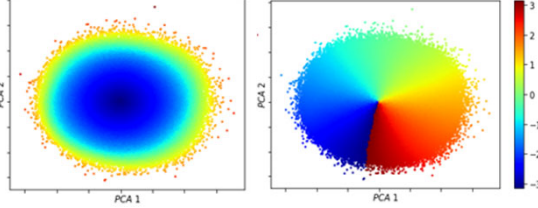
- Variational Autoencoder Theory

$\min KL(q(\mathbf{z}|\mathbf{x}, \mathbf{y}) || p(\mathbf{z}|\mathbf{x}, \mathbf{y}))$
equivalent to

$$\min \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2 + \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 + KL(q(\mathbf{z} | \mathbf{x}, \mathbf{y}) || p(\mathbf{z}))$$

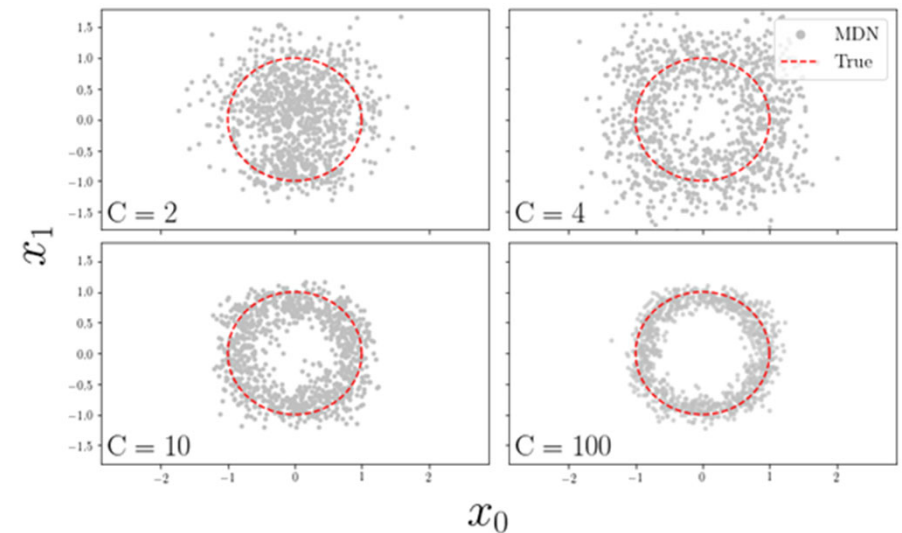
- True prior distribution $p(\mathbf{z})$
 - Select tractable distribution easy to generate
 - Gaussian
 - Uniform

VAIM on Toy Inverse Problems

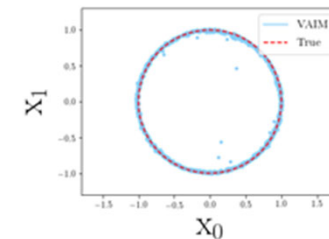
	Multiple Solutions	Latent Space Distribution
$f(x) = x^2$	 Parameter x distribution for $f(x) = 1.0$ Parameter x distribution for $f(x) = 0.36$	Sign Information 
$f(x) = \sin(x)$ $x \in [-2\pi, 2\pi]$	 $f(x) = 1.0$ $f(x) = -0.5$ $f(x) = 0.0$	Period Information 
$f(x_0, x_1) = x_0^2 + x_1^2$	$f(x_0, x_1) = 1.0$ 	Radius and Polar Angle Information 

Comparison with Mixture Density Network (MDN)

- Fundamental Idea of MDN
 - Construct a conditional probability $p(\mathbf{y}|\mathbf{x})$
 - Approximated with mixing Gaussian components
 - Assumption
 - (Finite) Gaussian Mixture
 - Poor approximation when the inverse problem is significantly non-Gaussian
- Advantage of VAIM
 - No Gaussian Assumption



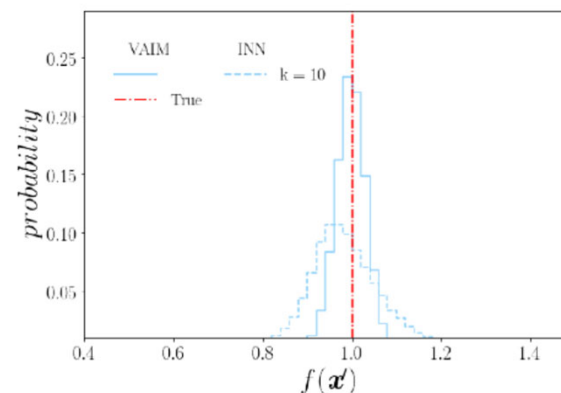
MDN predictions for toy problem $f(x) = x_0^2 + x_1^2$



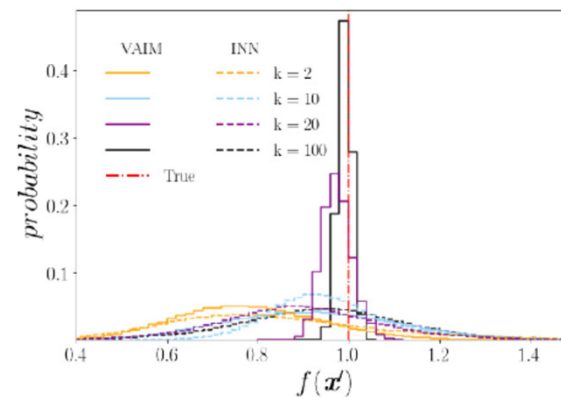
VAIM predictions for toy problem $f(x) = x_0^2 + x_1^2$

Comparison with Invertible Neural Networks (INN)

- Invertible Neural Networks (INN)
 - Maximum Mean Discrepancy (MMD)
 - Degrade polynomially at best as dimension increases
- VAIM
 - KL-divergence
 - Degrade constantly as dimension increases
 - Forward and Backward Mappers can adopt different NN Architectures



(a)



(b)

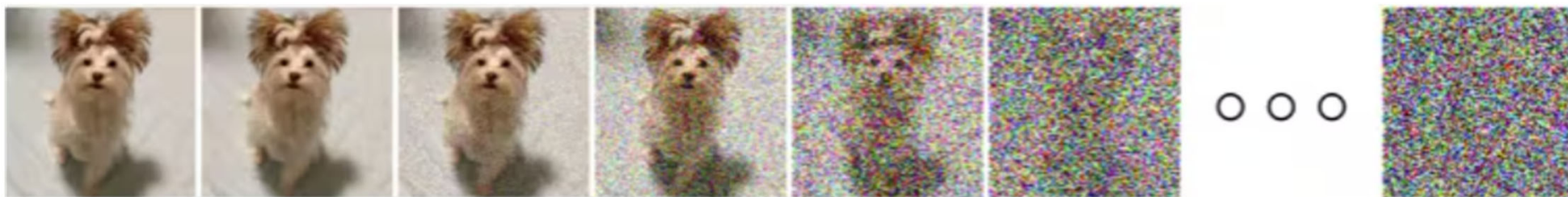
Fig. 9: Comparison of the solution distributions of $f(\mathbf{x}')$ obtained by VAIM and INN in the toy problem $f(\mathbf{x}) = \sum_i x_i^2$, when $f(\mathbf{x}) = 1$ is given, for the (a) 2D and (b) 10D cases.

Diffusion Model for Inverse Problems

Forward Diffusion

Corrupt data by adding noise

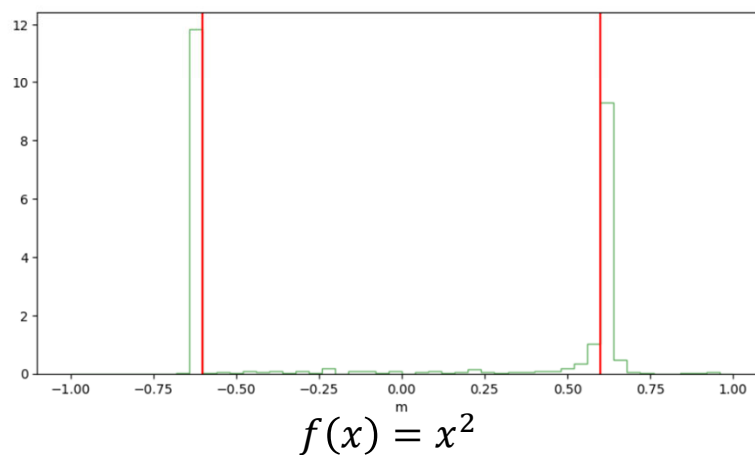
Data



Noise

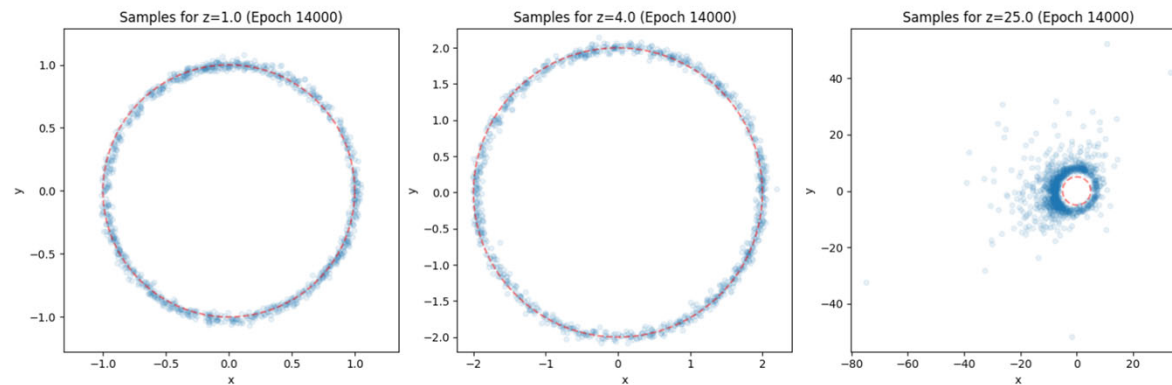
Generate sample by denoising

Reverse Diffusion

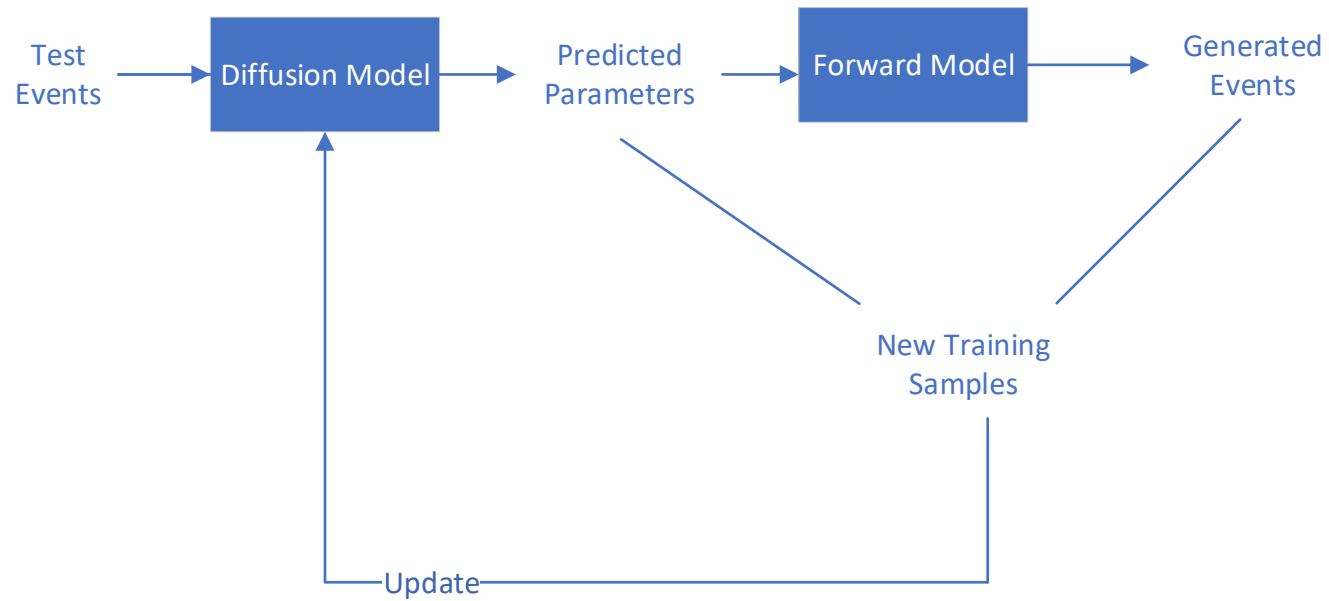


Toy Example

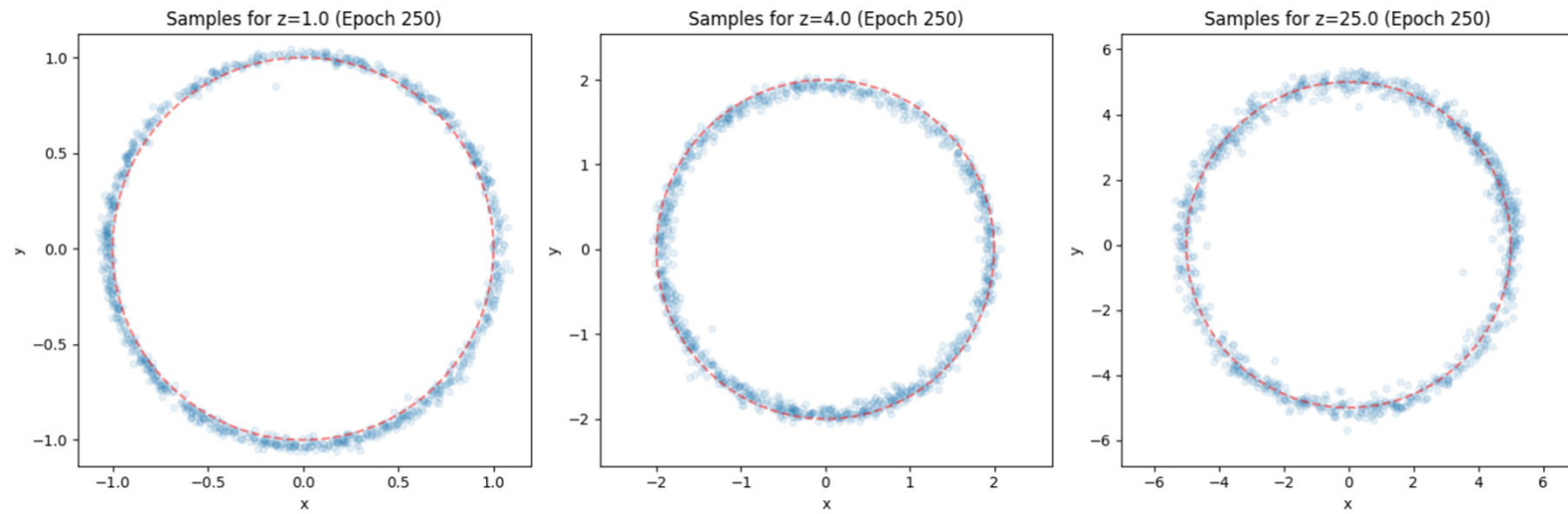
- $x^2 + y^2 = z$
- Training samples: $z \in [0, 4]$



Active Learning



Toy Example: Active Learning



Agenda

- The Inverse Problem
- The Machine Learning Architectures
 - Variational Autoencoder Inverse Mapper
 - Diffusion Model
- Applications in Nuclear and Particle Physics
 - QCF Parameterization from DIS
 - Extraction of Crompton Form Factors as an Inverse Problem
 - Event-level Analysis

Extraction of Quantum Correlation Functions from Deep-inelastic Lepton-Nucleon Scattering (DIS) Data

- Observables

$$\sigma_p(x, Q^2) = 4u(x, Q^2) + d(x, Q^2),$$

$$\sigma_n(x, Q^2) = 4d(x, Q^2) + u(x, Q^2).$$

- PDFs with DGLAP like behavior

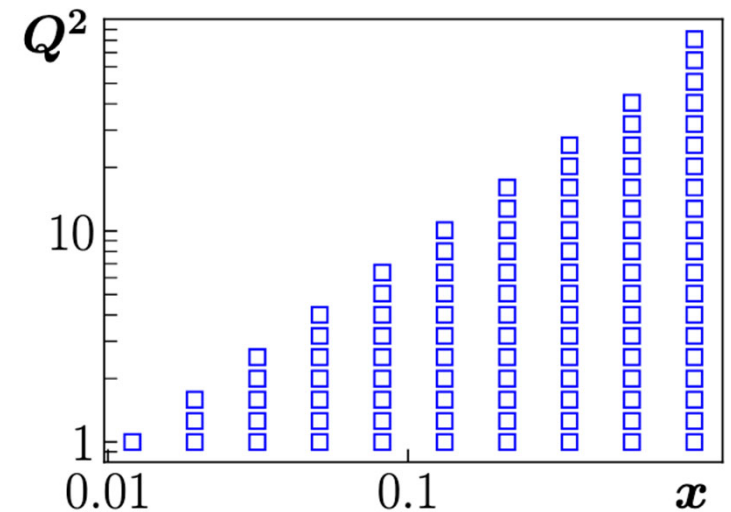
$$u(x, Q^2) = N_u(Q^2) x^{\alpha_u(Q^2)} (1 - x)^{\beta_u(Q^2)} (1 + \gamma_u(Q^2) \sqrt{x} + \delta_u(Q^2) x),$$

$$d(x, Q^2) = N_d(Q^2) x^{\alpha_d(Q^2)} (1 - x)^{\beta_d(Q^2)} (1 + \gamma_d(Q^2) \sqrt{x} + \delta_d(Q^2) x),$$

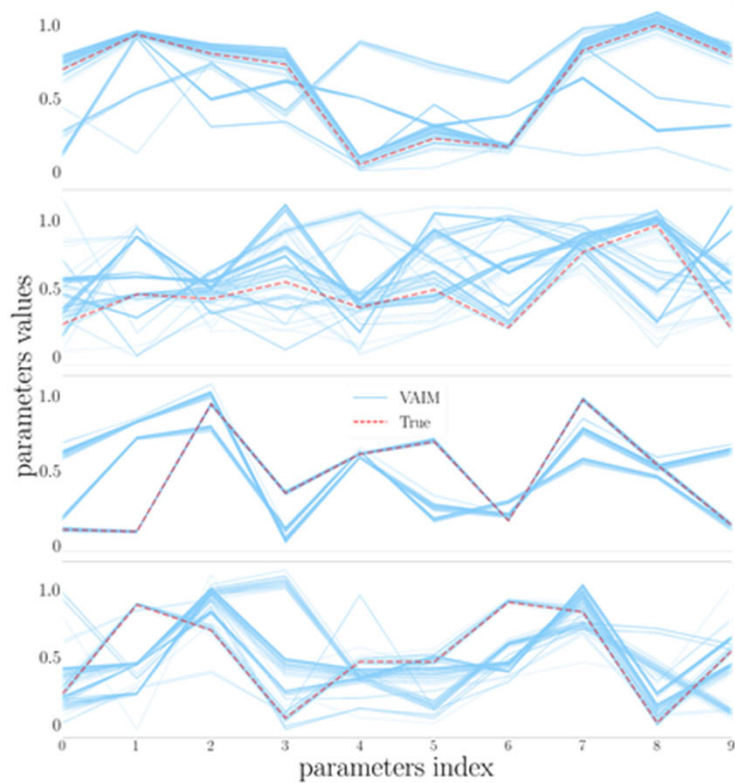
- Shape parameters $p = \{N_{u,d}, \alpha_{u,d}, \beta_{u,d}, \gamma_{u,d}, \delta_{u,d}\}$ are parameterized as

$$p(Q^2) = p^{(0)} + p^{(1)} s(Q^2), \quad s(Q^2) = \log \left(\frac{\log(Q^2/\Lambda_{\text{QCD}}^2)}{\log(Q_0^2/\Lambda_{\text{QCD}}^2)} \right)$$

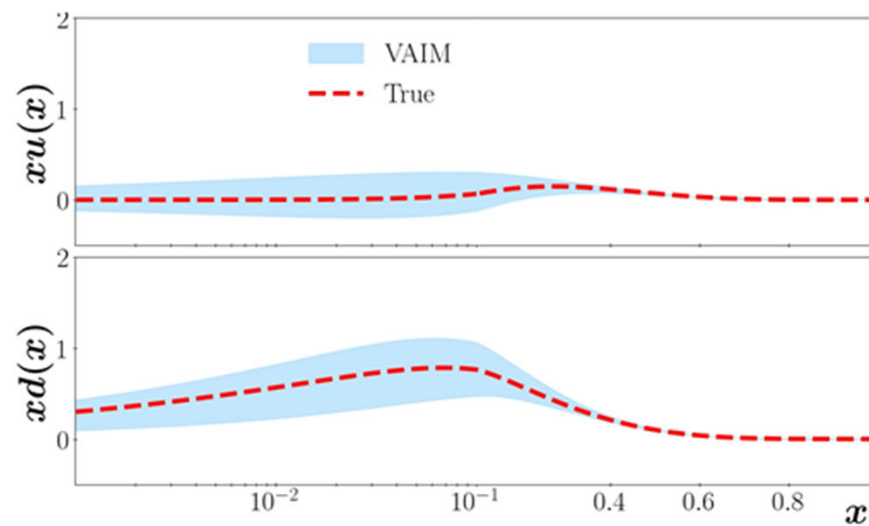
Phase Space



VAIM Results in Toy DIS Problem



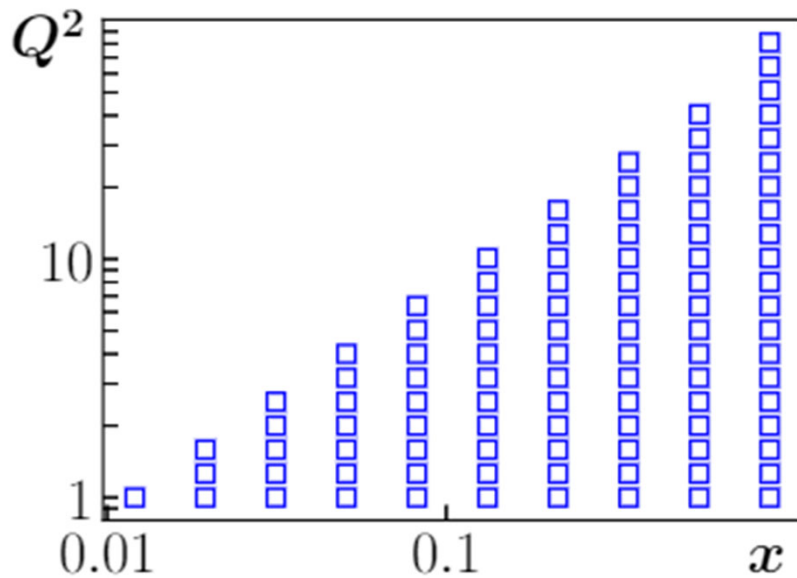
Parameter distributions generated by VAIM in four control samples



Reconstructed PDF using a control sample

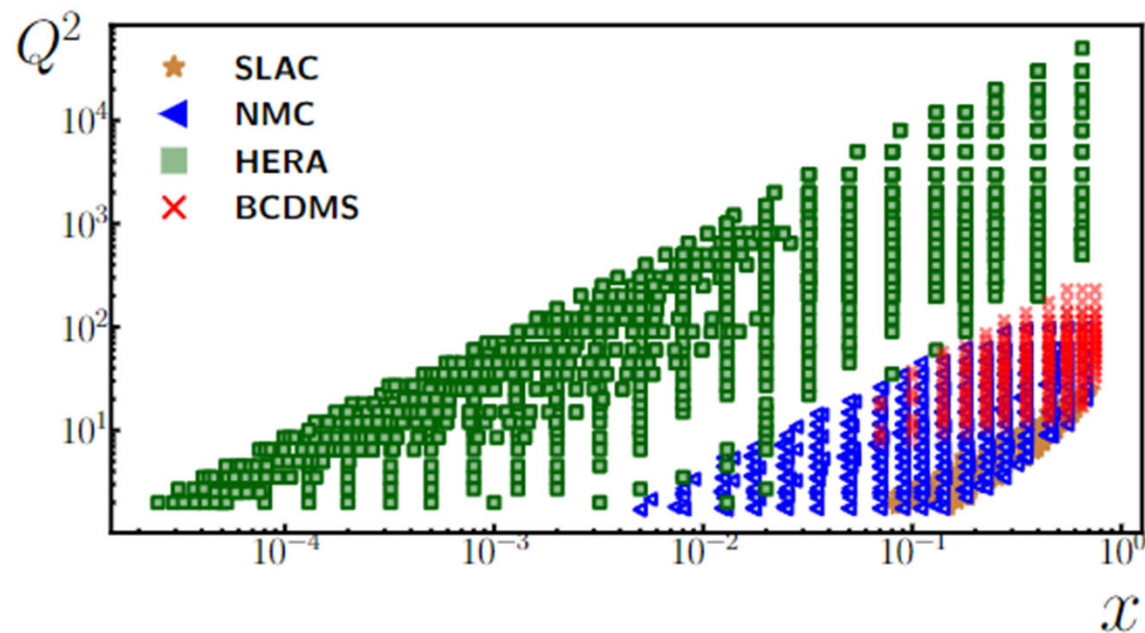
Point Cloud-based VAIM (PC-VAIM)

Limitation of VAIM: Observables across regular, discretized kinematic bins



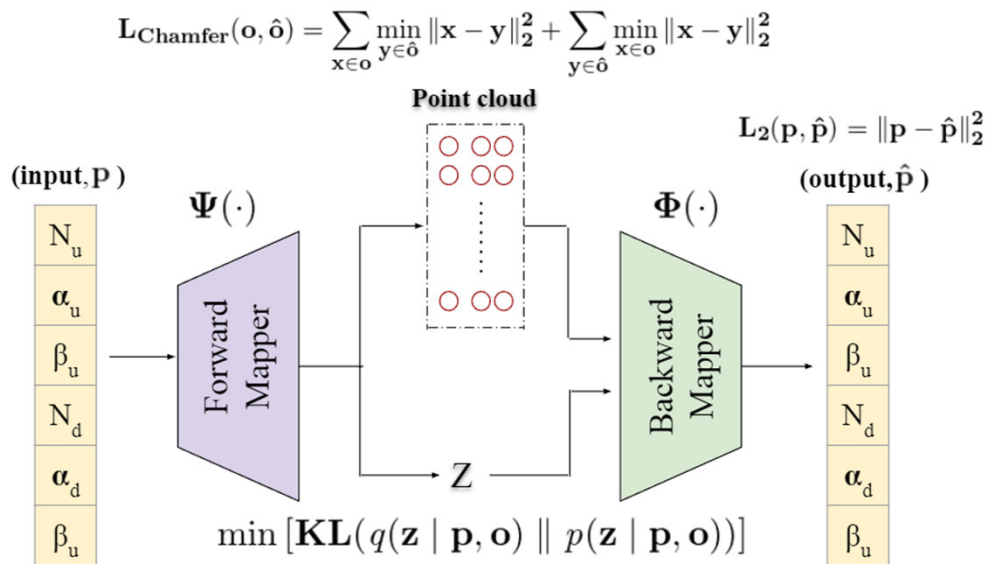
Regular, discretized kinematic bins of x and Q^2 where cross sections are evaluated

Reality: Observables are ill-defined
Data in different experiments are observed on different kinematic bins

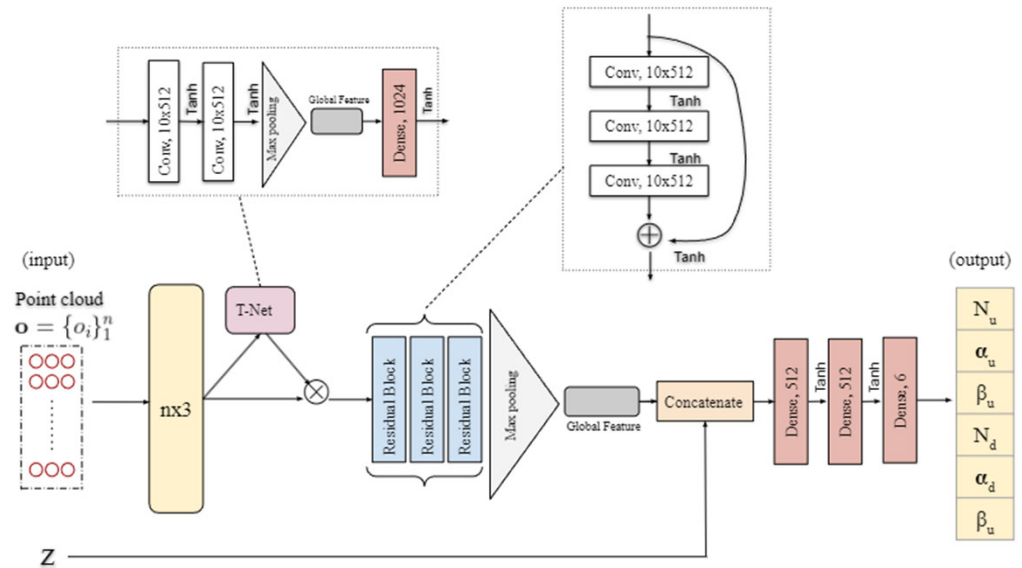


Kinematic bins of SLAC, NMC, HERA, and BCDMS experiments

PC-VAIM Architecture

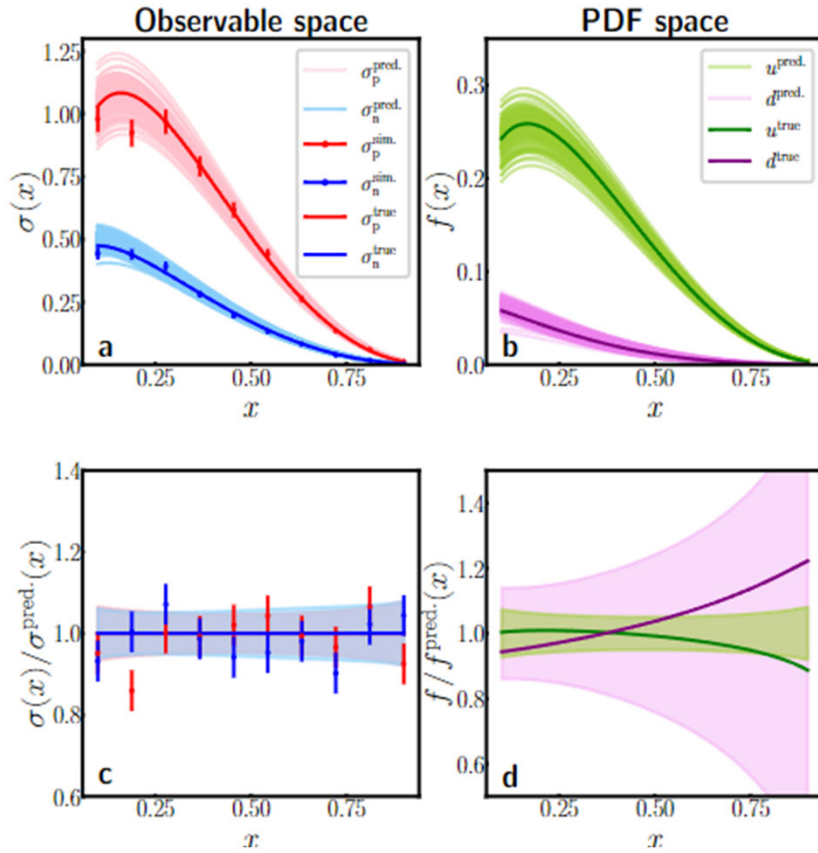


Overall Architecture

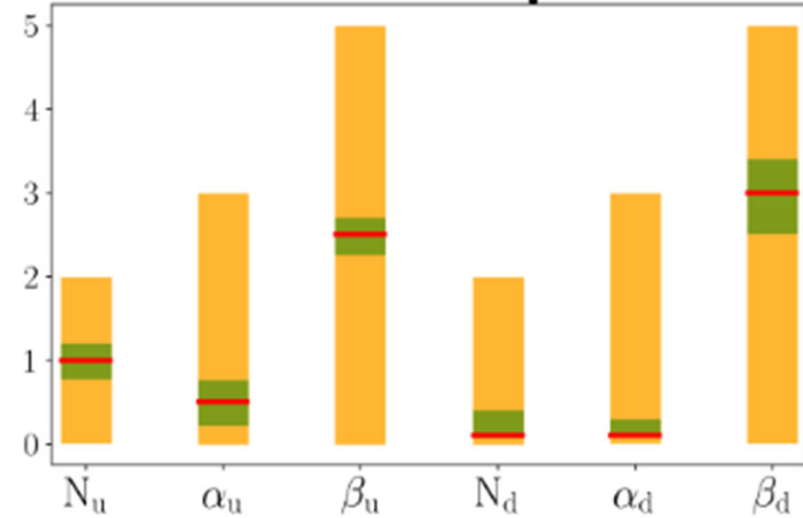


Backward Mapper: A PointNet-based architecture is used to handle the point cloud observable input

PC-VAIM on Extracting QCFs from Grid-Independent DIS data



Parameter space



Parameter regions from which the training samples are drawn (orange), the control data (red), and PC-VAIM predictions (green band)

PC-VAIM predictions on 1,000 simulated control samples given by σ_p^{sim} and σ_n^{sim} . (a) Reconstructed observable σ_p and σ_n for the PC-VAIM predicted parameters using physics theory model. (b) PDFs for the "up" and "down" quarks produced by physics theory model corresponding to the predicted parameters. (c) Ratio of the reconstructed observables over the true observables. (d) Ratio of the reconstructed PDFs over the true. PC-VAIM correctly learns the mapping between the observable space and the PDF space.

Agenda

- The Inverse Problem
- The Machine Learning Architectures
 - Variational Autoencoder Inverse Mapper
 - Diffusion Model
- Applications in Nuclear and Particle Physics
 - QCF Parameterization from DIS
 - Extraction of Crompton Form Factors as an Inverse Problem
 - Event-level Analysis

Extraction of Compton Form factors

- Generalized Parton Distributions (GPDs)
 - Multi-dimensional descriptions of proton structure
- Deeply virtual exclusive scattering processes
 - **Golden channel** for the extraction of information on partonic 3D dynamics in the nucleon
- Compton Form Factors (CFFs)
 - 2D Slices of GPDs
 - Measured in Deeply Virtual Compton Scattering (DVCS)
 - Contain potentially new information on hadronic structure

Extraction of CFFs as an **Inverse Problem**

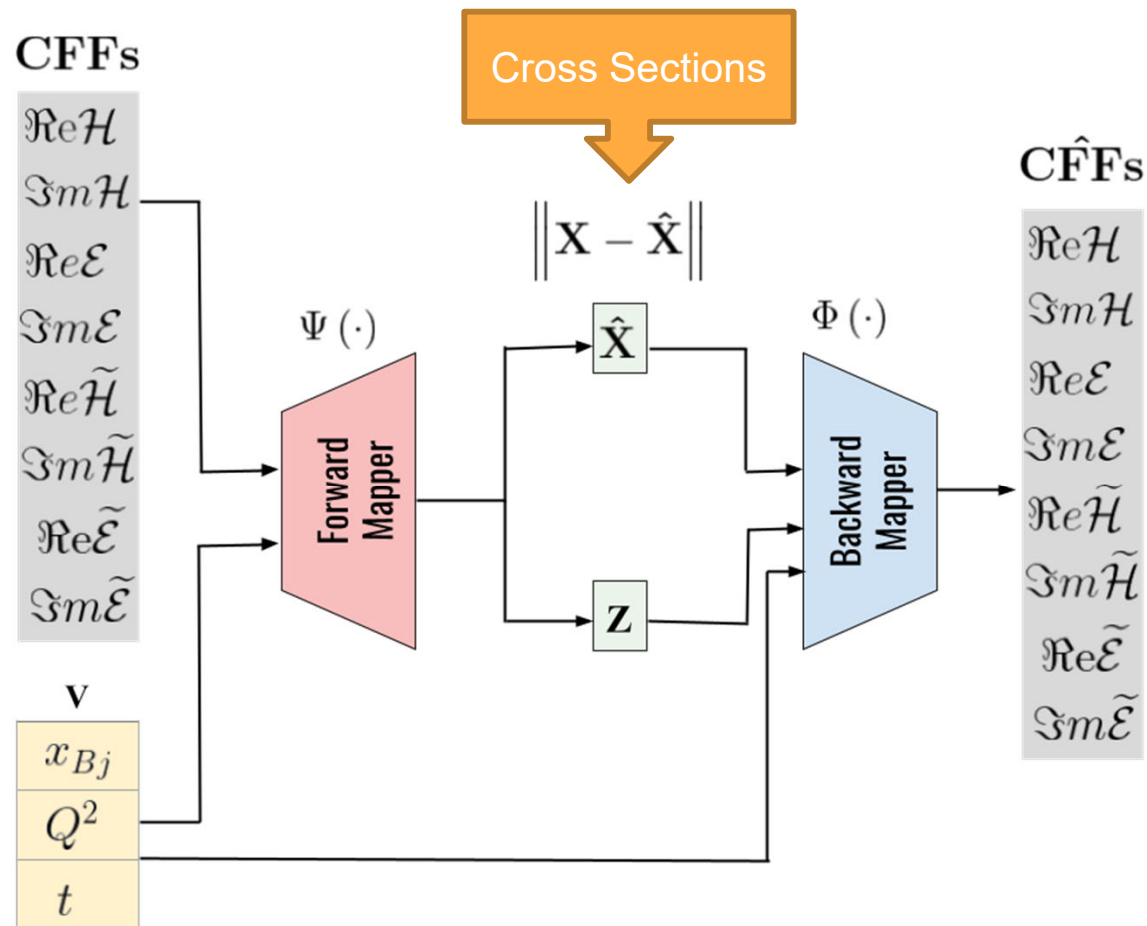
- Extraction of 8 CFFs from a single polarization observable

- An inverse problem of extracting 8 unknowns from a single equation

$$Re(H), Im(H), Re(\tilde{H}), Im(\tilde{H}), Re(E), Im(E), Re(\tilde{E}), Im(\tilde{E})$$

- Quantification of information extracted from experiments

Conditional VAIM Architecture for CFF Extraction



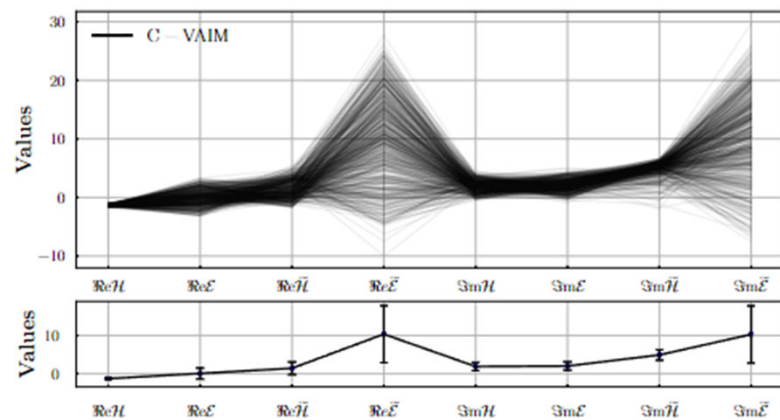
Training of Conditional VAIM

- Training Data:
 - Kinematics values

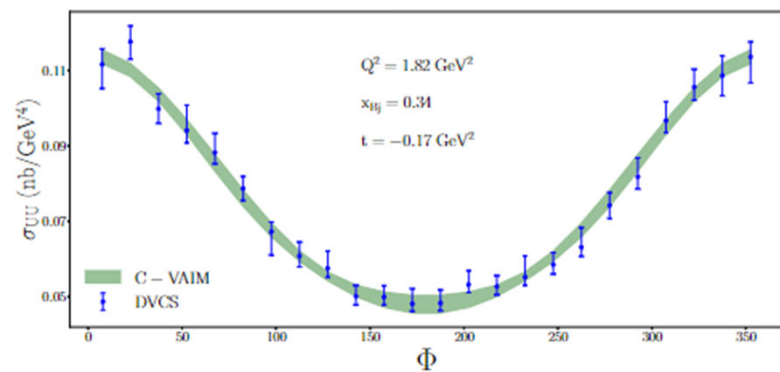
Bin	x_{bj}	t (GeV ²)	Q^2 (GeV ²)
1	0.343	-0.172	1.820
2	0.368	-0.232	1.933
3	0.375	-0.278	1.964
4	0.379	-0.323	1.986
5	0.381	-0.371	1.999

- Generate uniformly distributed CFFs
 - $ReH \in [-4, 4]$
 - $ReE \in [-4, 4]$
 - $Re\tilde{H} \in [-10, 10]$
 - $Re\tilde{E} \in [-10, 30]$
 - $ImH \in [-1, 5]$
 - $ImE \in [-1, 5]$
 - $Im\tilde{H} \in [-1, 20]$
 - $Im\tilde{E} \in [-10, 30]$
- Compute cross sections

Prediction of CFFs

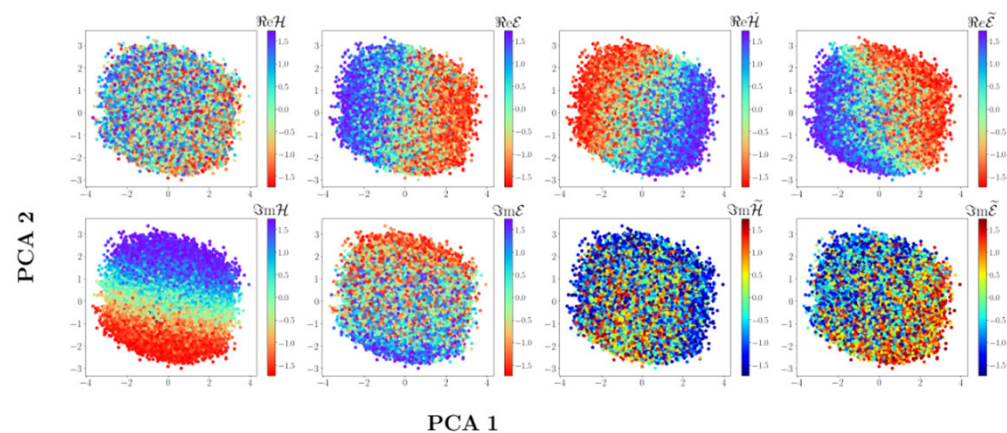


CFFs



● Predicted CFFs from VAIM-CFF

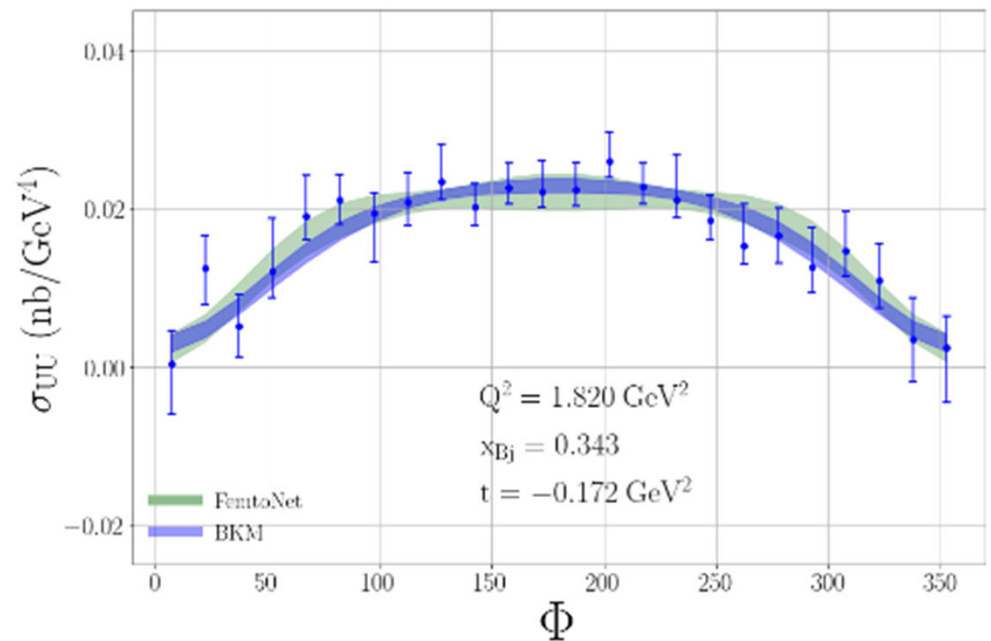
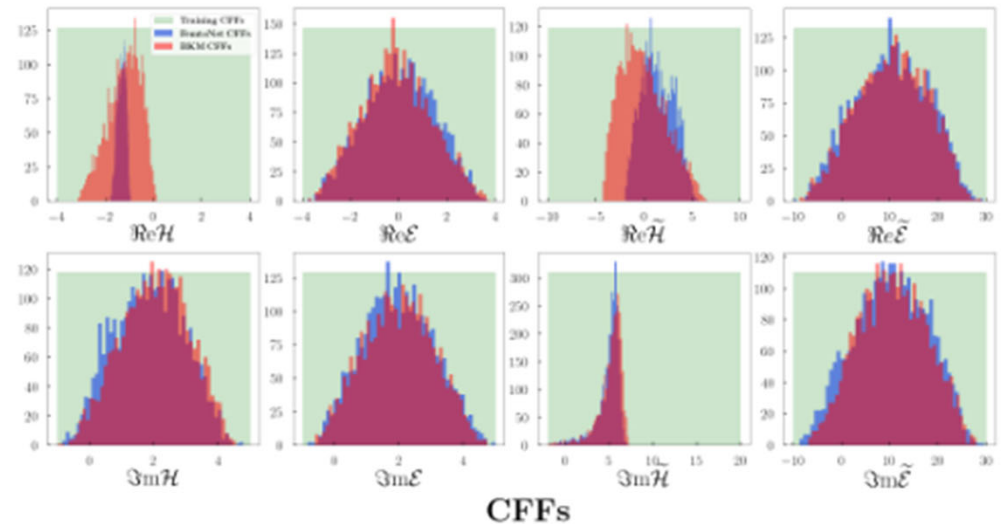
- $X_{Bj} = 0.343$
- $t = -0.172 \text{ GeV}^2$
- $Q^2 = 1.820 \text{ GeV}^2$



Latent Space Analysis

Training C-VAIM on Different Cross-section Formulations

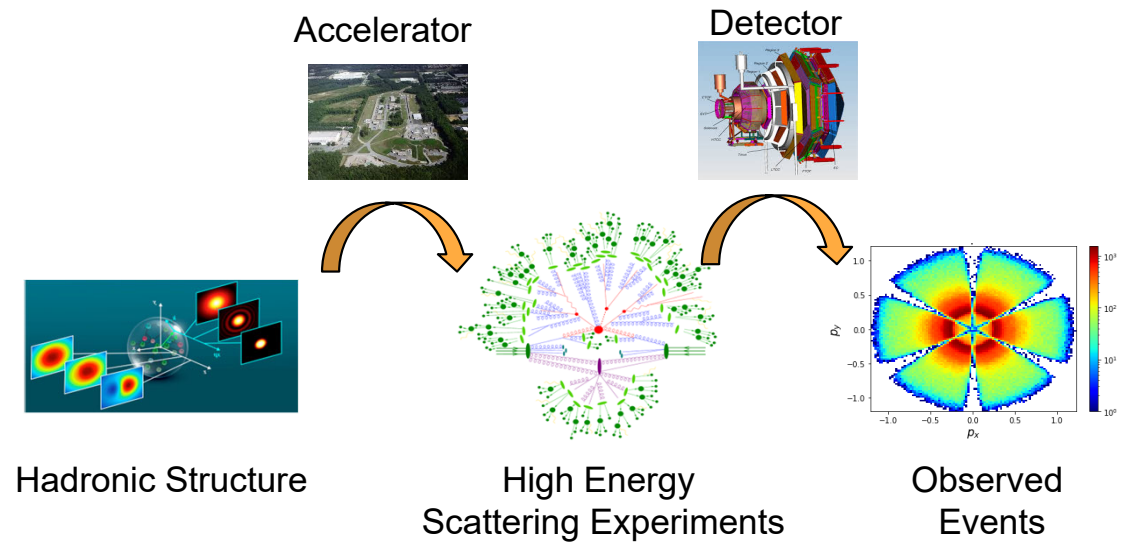
- Two Cross-section Formulations
 - FemtoNet (UVA)
 - BKM (Belitksy, Kirchner, Mueller)
- Kinematics
 - $X_{Bj} = 0.35$
 - $t = 0.172$
 - $Q^2 = 1.9 \text{ GeV}^2$
 - $Eb = 5.75 \text{ GeV}$



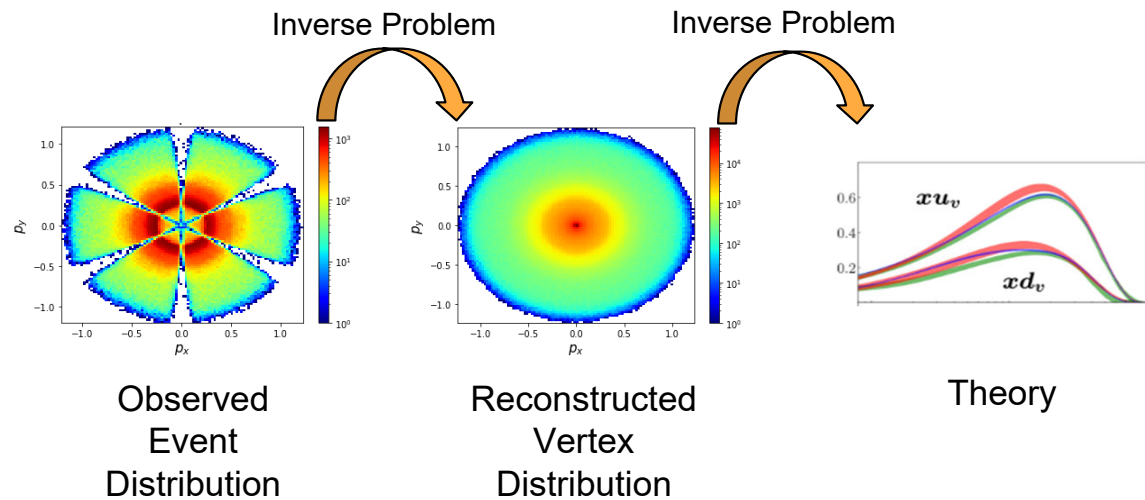
Agenda

- The Inverse Problem
- The Machine Learning Architectures
 - Variational Autoencoder Inverse Mapper
 - Diffusion Model
- Applications in Nuclear and Particle Physics
 - QCF Parameterization from DIS
 - Extraction of Crompton Form Factors as an Inverse Problem
 - Event-level Analysis

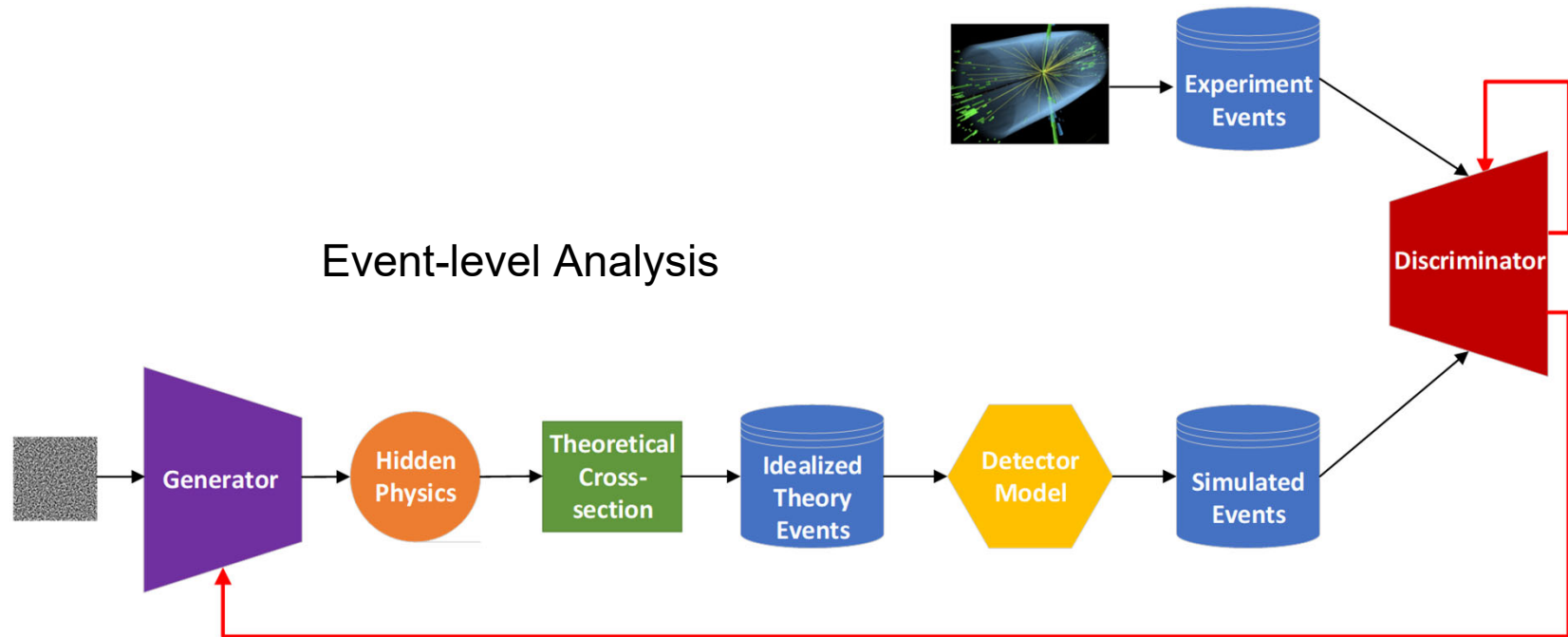
Experiment



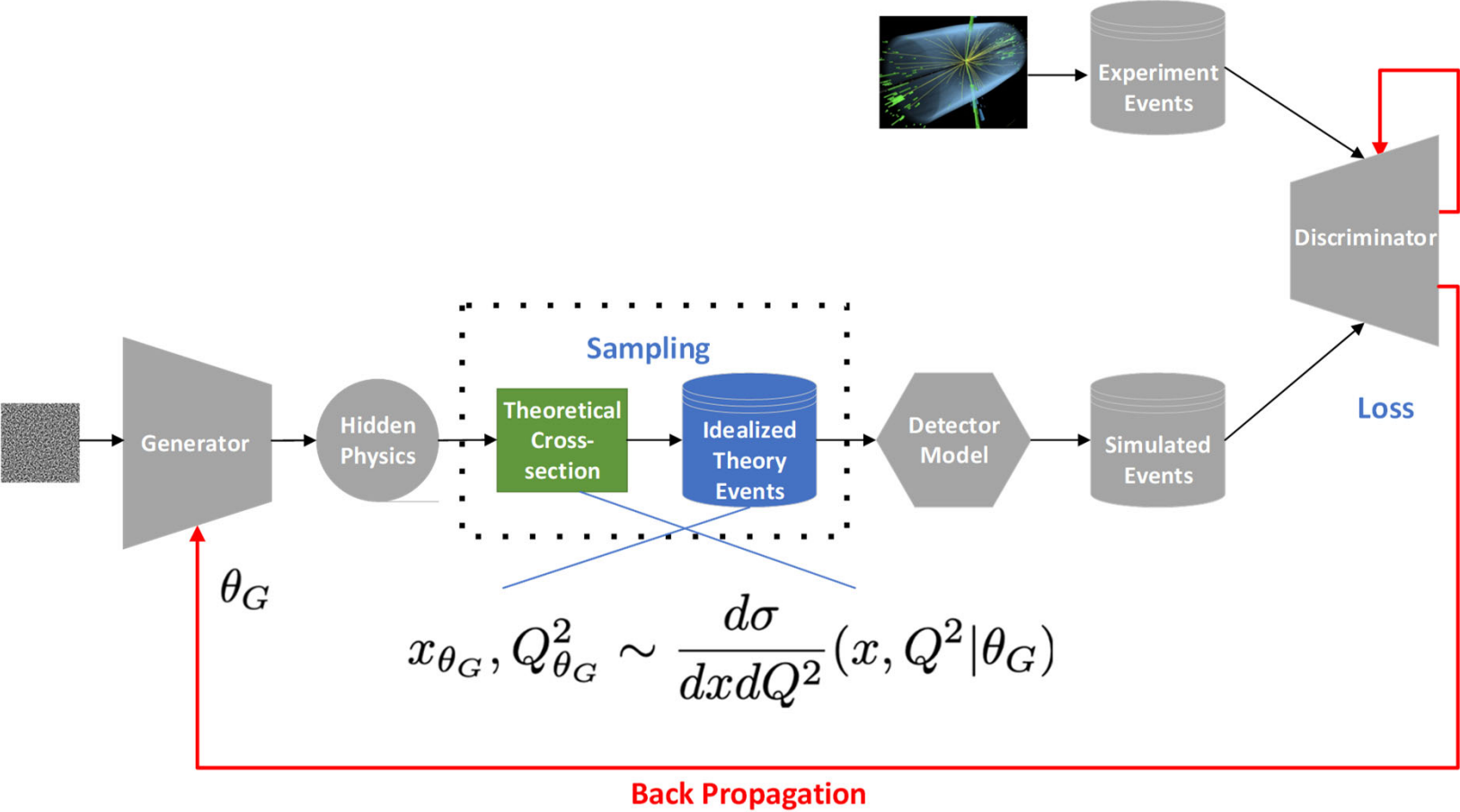
Inverse Inference



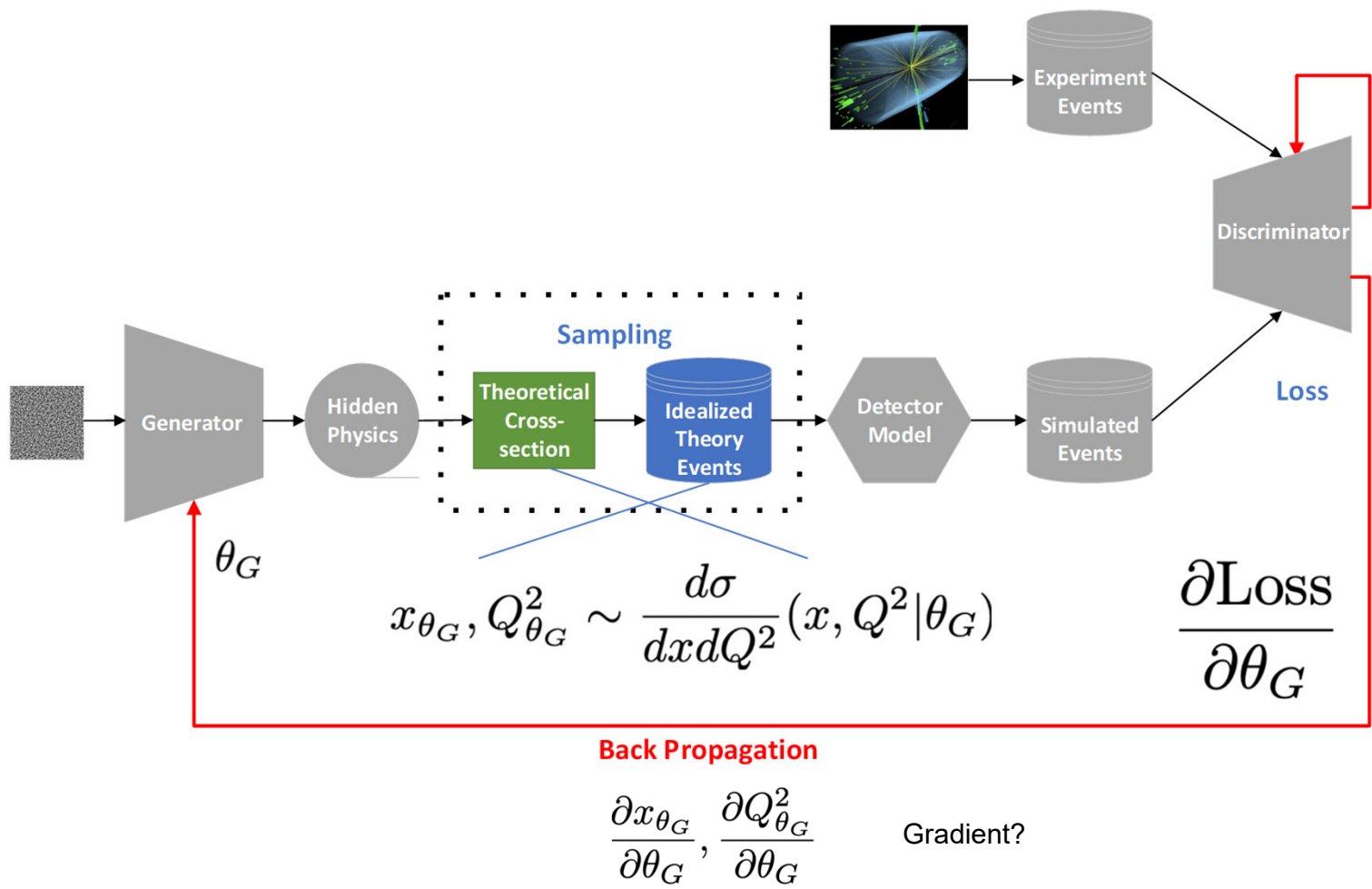
GAN-based Workflow for Event-level Analysis



Sampling

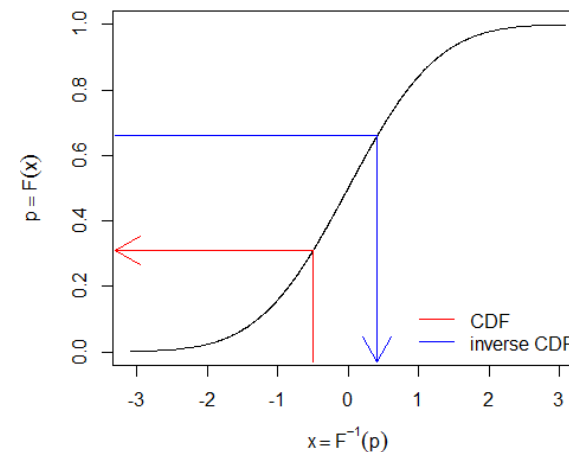


A Critical Obstacle in Backpropagation: Gradient over Sampling



Sampling Algorithms

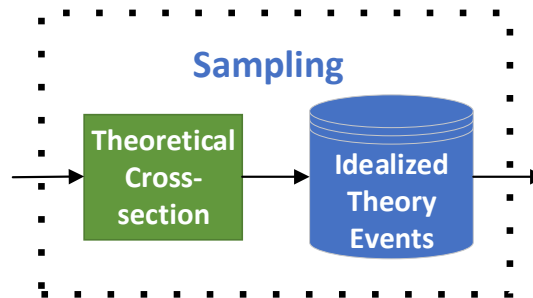
- **Inverse CDF**
 - Can provide meaningful gradient
 - Only works in 1D



- **Markov Chain Monte Carlo (Gaussian proposal function)**
 - Random walk
 - Can work in higher dimensionality
 - No useful gradient

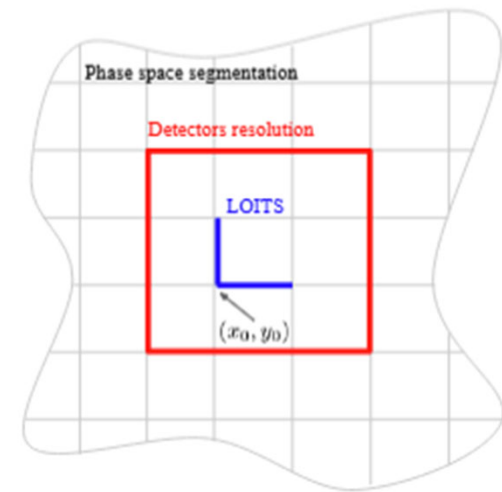
Two Possible Solutions

- **Make the Sampling Method Differentiable**
 - LOITS (by Nobuo Sato)
- **Build a Surrogate Model**
 - Neural Network
 - Approximate the sampling process
 - Differentiable
 - Easy to incorporate in the event-based analysis workflow



LOITS

- **Procedure (proposed by Nobuo Sato)**
 - Partition space into grid cells
 - Compute the distribution value at each grid point
 - Sample each grid cell
 - Compute marginal distribution in each dimension
 - 1D inverse CDF in each dimension
 - Assemble the samples (Approximate)
 - MCMC Correction (Precise)
 - $\text{Ratio} = \frac{p(prop)p_{LOITS}(curr)}{p(curr)p_{LOITS}(prop)}$



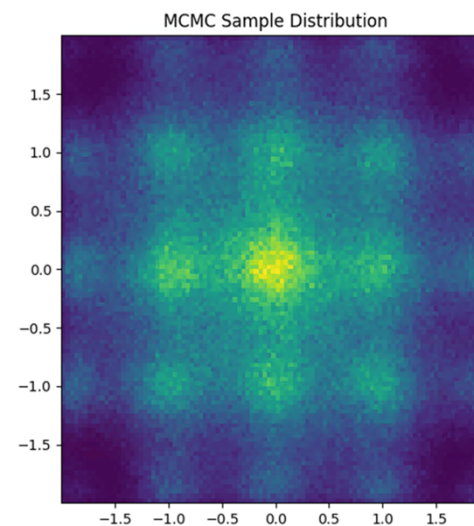
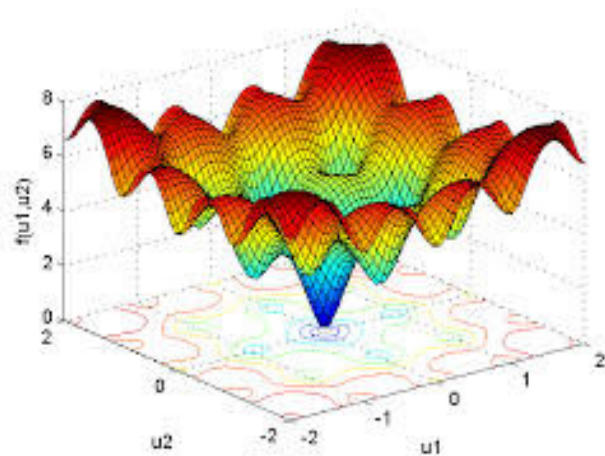
Normalizing Flow Guided Sampling

- **Procedure**

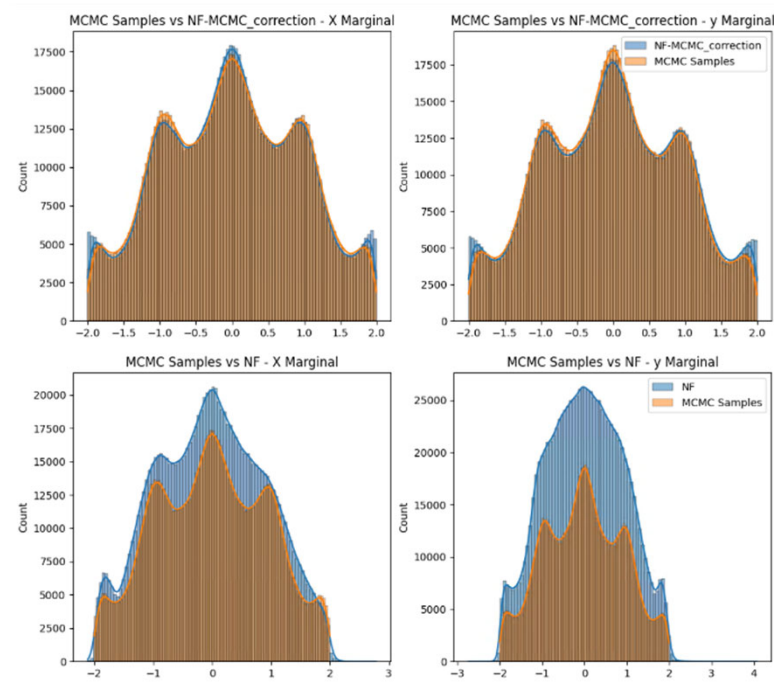
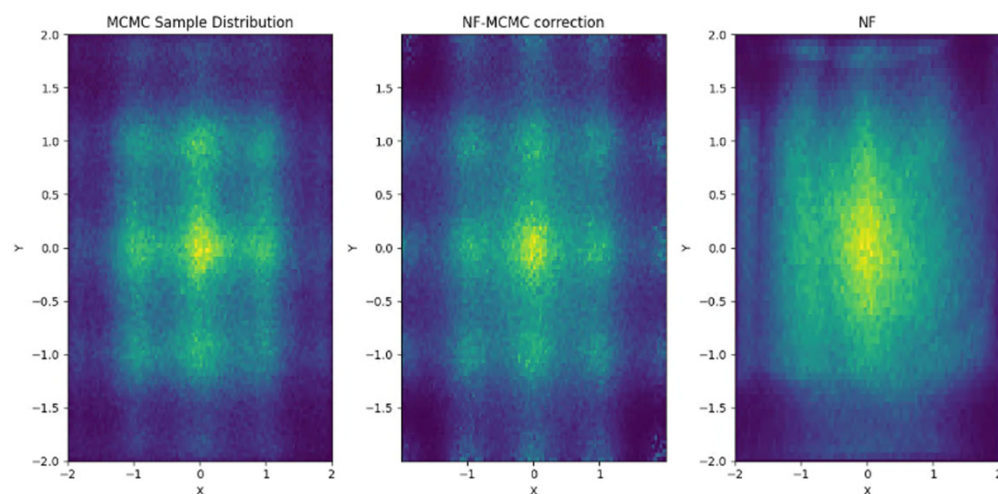
- Learn a normalizing flow to approximate the target distribution
 - Normalizing Flow not only gives us the samples, but also the corresponding PDF value
- Use MCMC correction
 - $\text{Ratio} = \frac{p(prop)p_{NF}(curr)}{p(curr)p_{NF}(prop)}$

Example: Ackley's function

- **Rough Landscape**
 - Many local minima around one deep global minimum



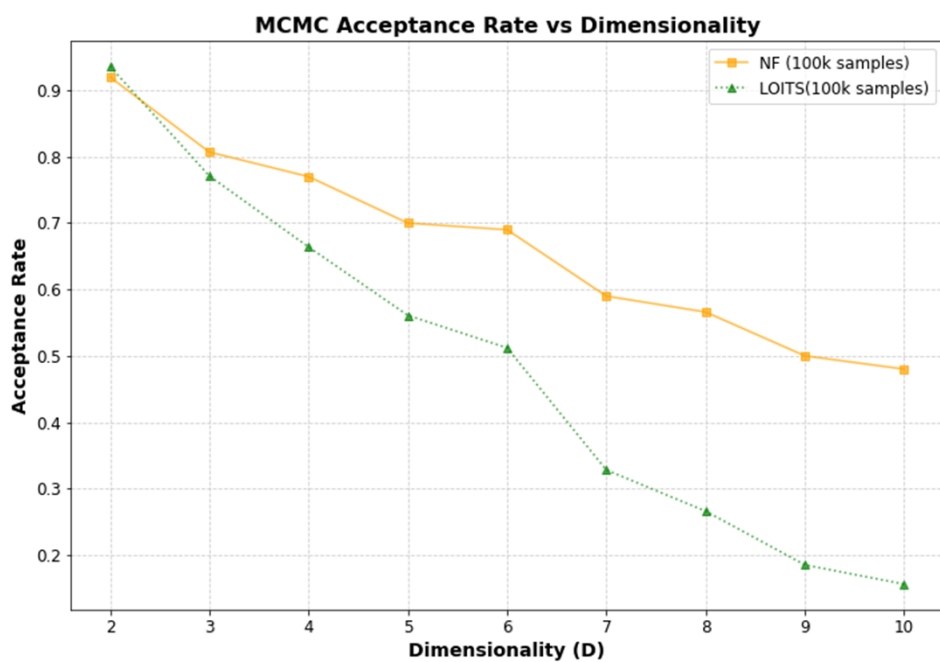
NF-guided Sampling



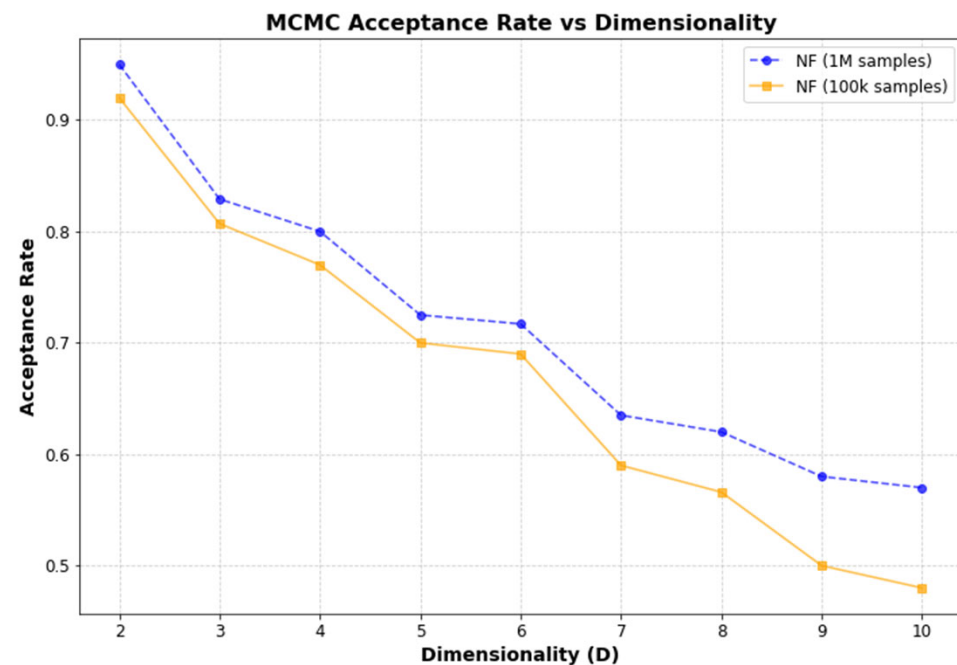
First row shows the truth(MCMC samples) vs the NF-MCMC-correction

Second row shows the the truth(MCMC samples) vs NF

Dimensionality Analysis (NF-guided vs LOITS)



Sampling with a better trained NF



Simulation-based Inference

■ An Inverse Problem

- Extract Quantum Correlation Functions (QCFs) from experimental events
 - Forward Simulation Model
 - sample QCFs to generate events
 - Statistical Inference Model
 - compare the generated events with the target ones to determine the parameters

■ Surrogate Models

- Approximate the forward sampler
- Good fidelity of the simulation
- Provide gradient for parameter inference in gradient-based optimization
- Fast computation
 - Only neural network evaluations

QCF Proxy Model

- **Measurable observables**

- σ_p^o and σ_n^o
 - Cross-section distributions from proton and neutron targets, respectively.
 - Driven by underlying quark distributions $u(x; p)$ and $d(x; p)$, which act as proxies for QCFs.

- **Forward model**

$$\sigma_1(x; p) = 4u(x; p) + d(x; p)$$

$$\sigma_2(x; p) = u(x; p) + 4d(x; p)$$

$$u(x; p) = N_u x^{a_u} (1 - x)^{b_u}$$

$$d(x; p) = N_d x^{a_d} (1 - x)^{b_d}$$

- **Inverse problem**

- Given event-level samples σ_p^o and σ_n^o , infer QCF parameters
 $p = \{N_u, a_u, b_u, N_d, a_d, b_d\}$

Computational Experiments

- **1D QCF Proxy Model**

- We can use 1D inverse CDF as the truth

- **1D inverse CDF**

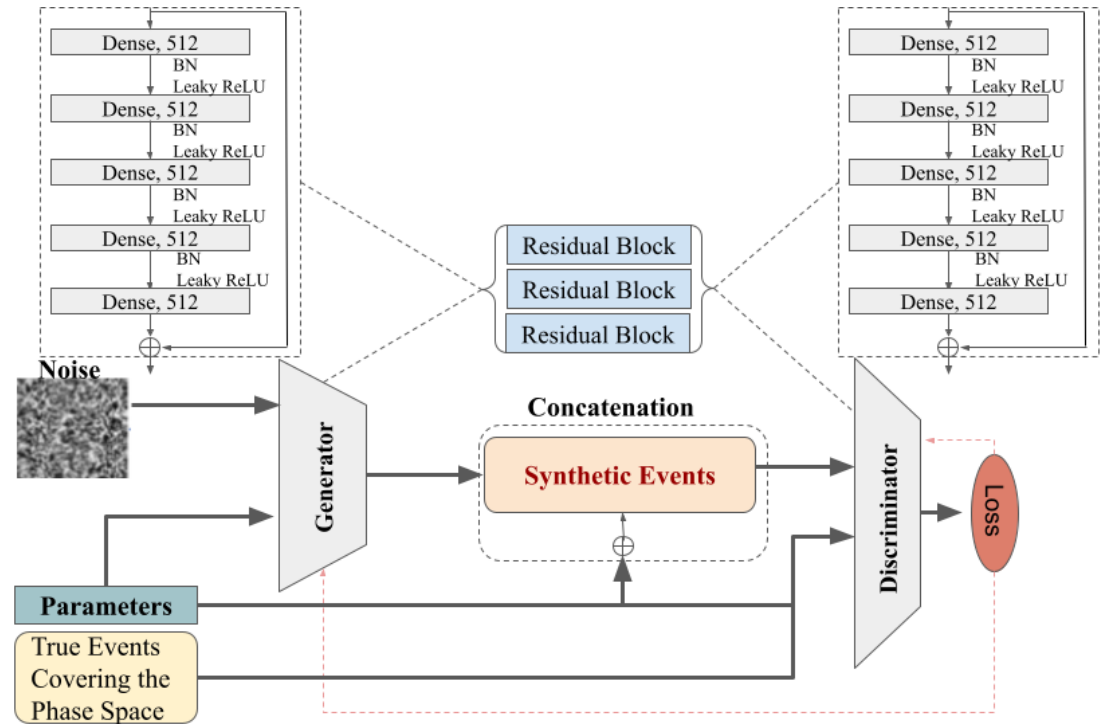
- “Perfect” sampling
- Need to evaluate PDF for every sample

- **Generative AI-based Surrogates**

- Conditional GAN
- Conditional Diffusion Model (DM)
- Both trained on 20M parameter-event pairs

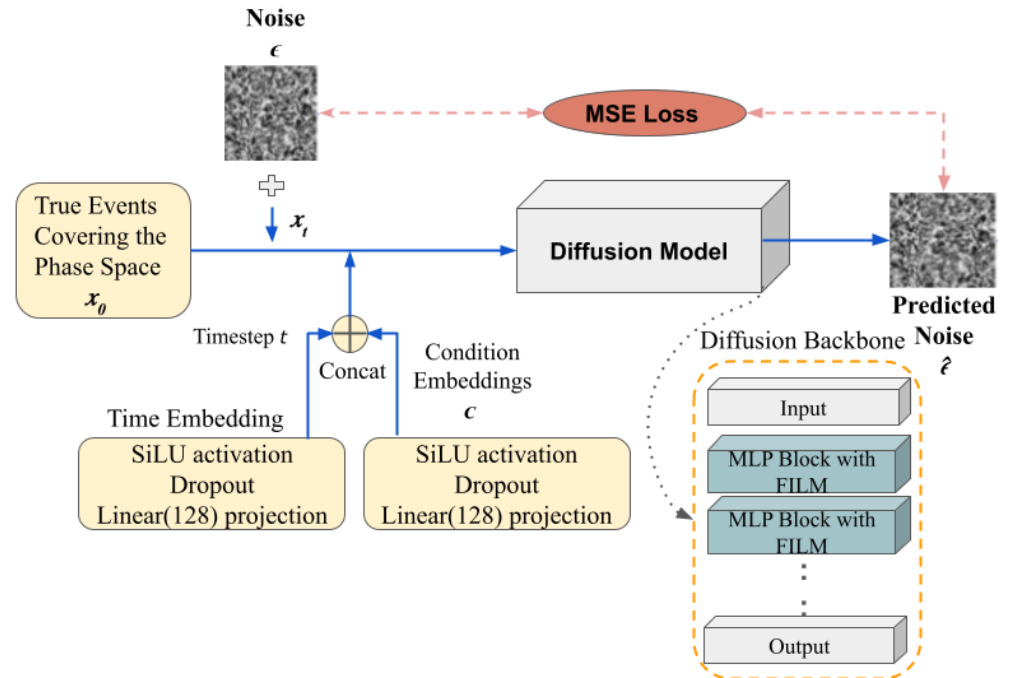
Conditional GAN as the Surrogate

- **Generator**
 - Residual blocks
 - 5 dense layers
 - 512 neurons each
- **Discriminator**
 - Similar architecture with 5% dropout
- **Loss**
 - Least Squares

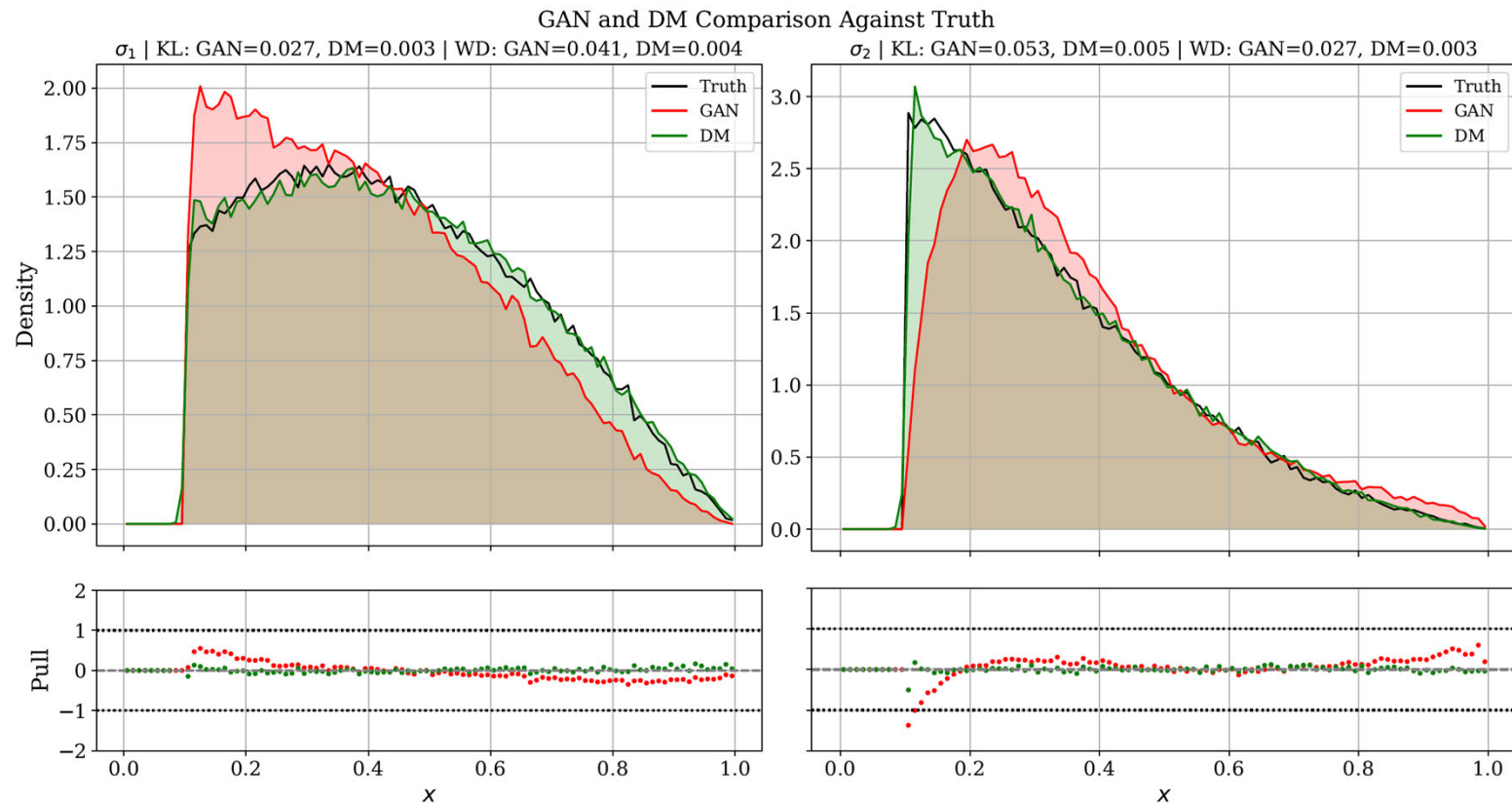


Diffusion Model as the Surrogate

- **Denoising Network**
 - 6 MLP blocks (256 hidden units)
 - FiLM (Feature-wise Linear Modulation) conditioning
- **Noise schedule**
 - $T = 1,000$ steps
 - $\beta_{\min} = 10^{-4}$
 - $\beta_{\max} = 0.02$
- **Training**
 - MSE loss between predicted and true noise

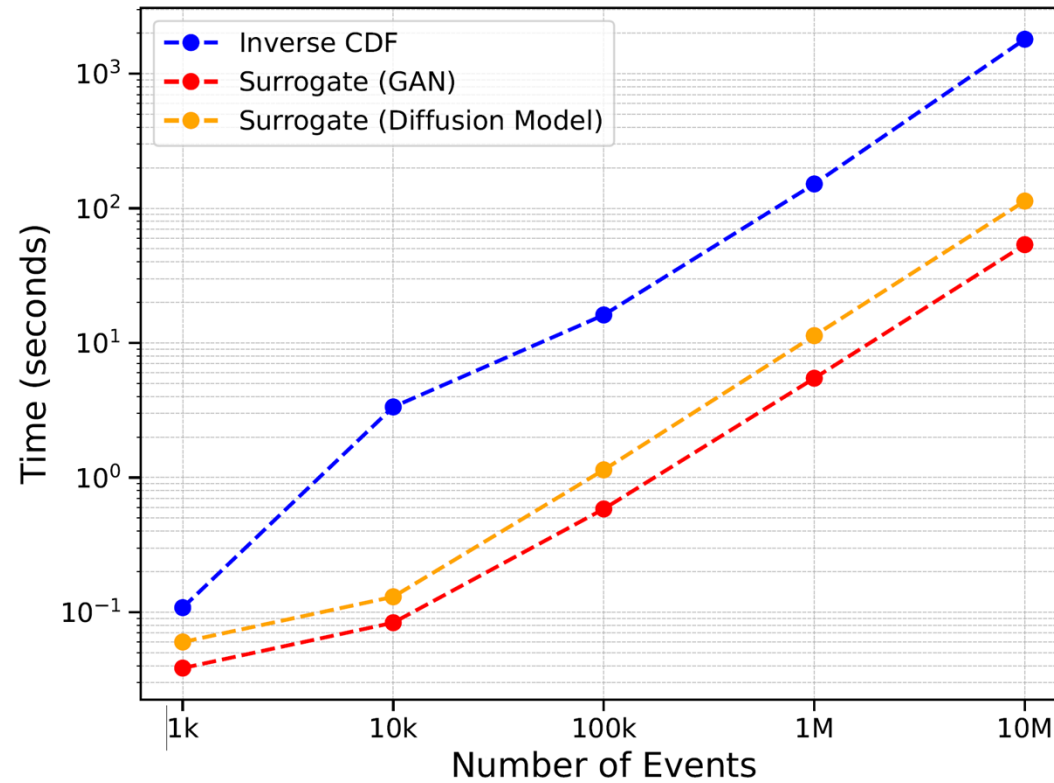


Fidelity Comparison of Generative AI Models

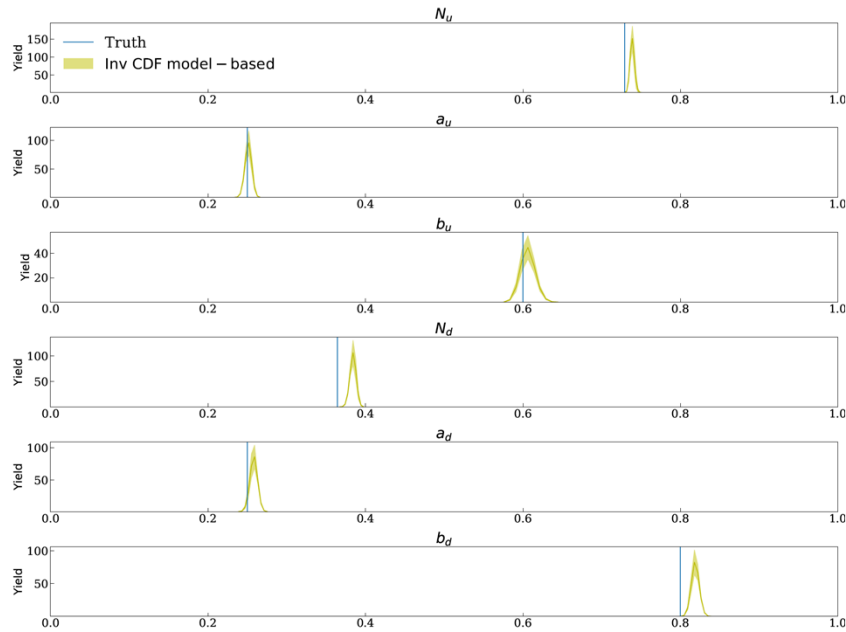


Diffusion Model yields better fidelity compared to GAN.

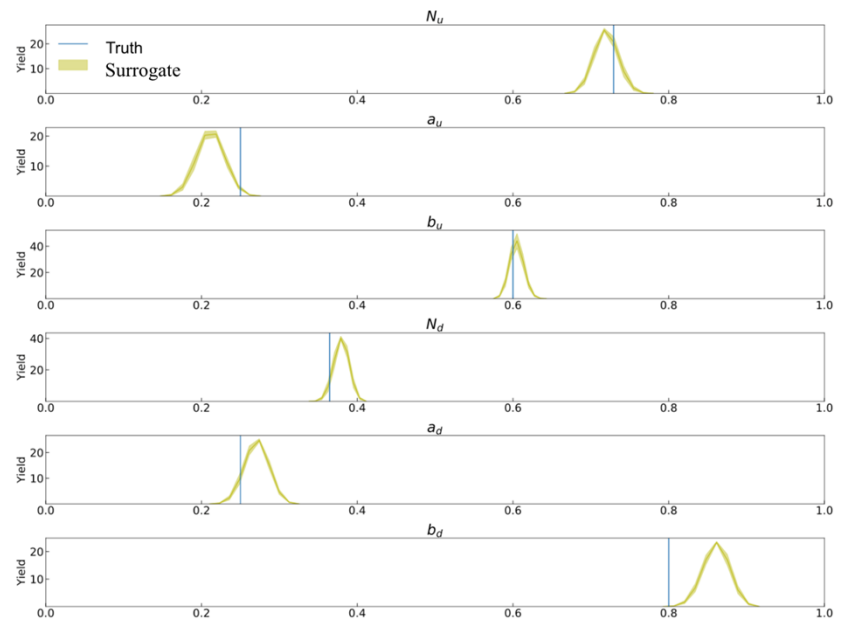
Comparison of Computational Time



QCF Parameter Inference

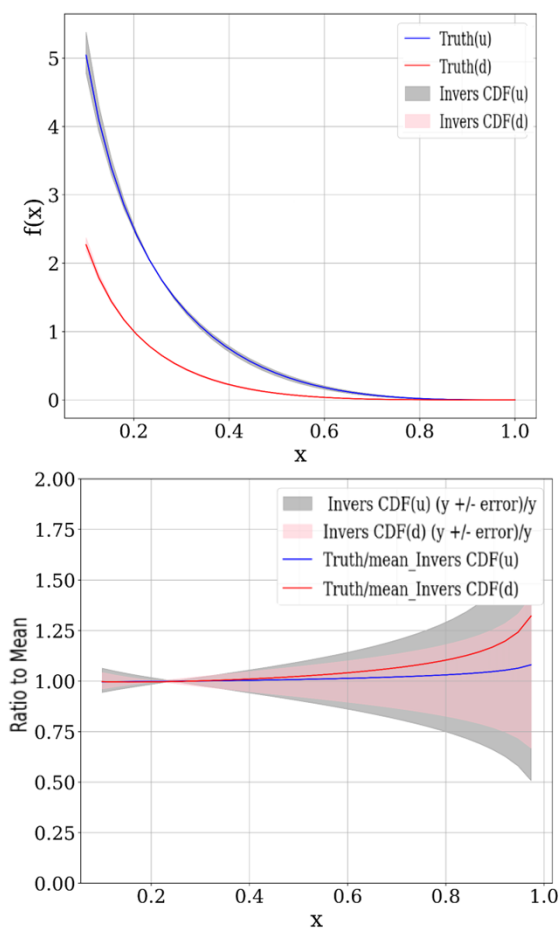


Parameter distributions obtained using an inverse-CDF sampler

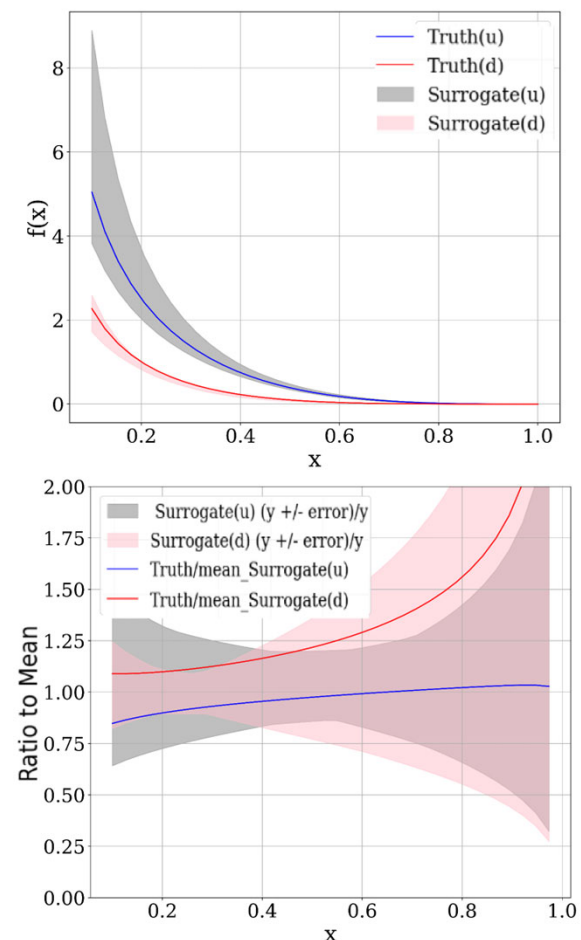


Parameter distributions obtained using the surrogate sampler

Reconstructed QCFs

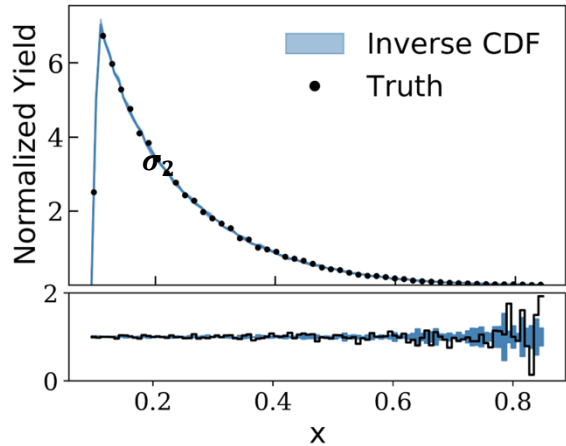
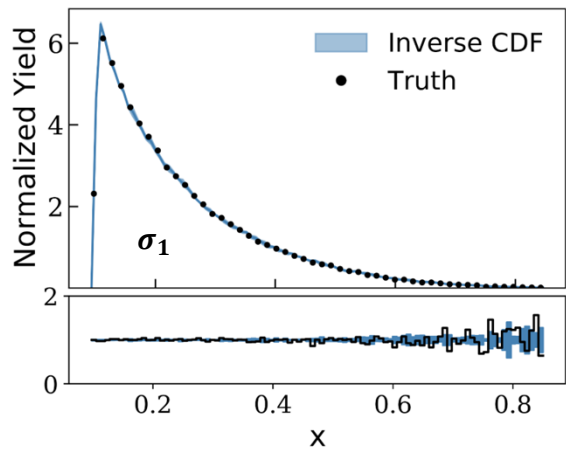


Inverse CDF

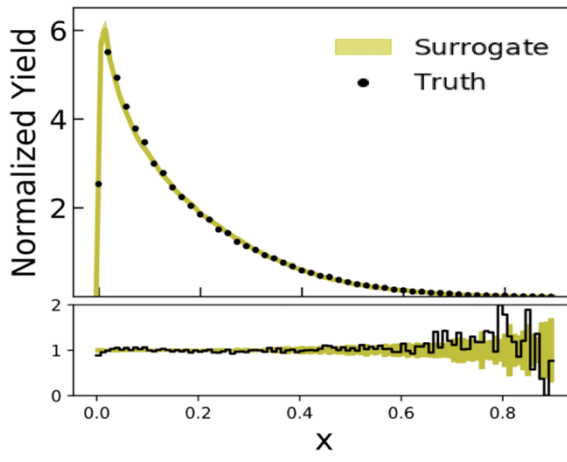
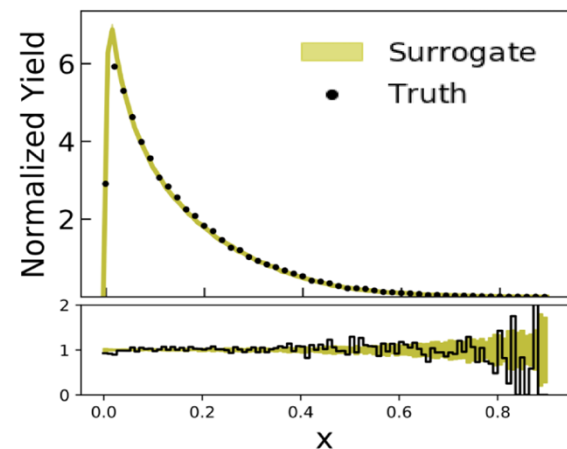


Generative AI-based Surrogate

Reconstructed Cross-sections



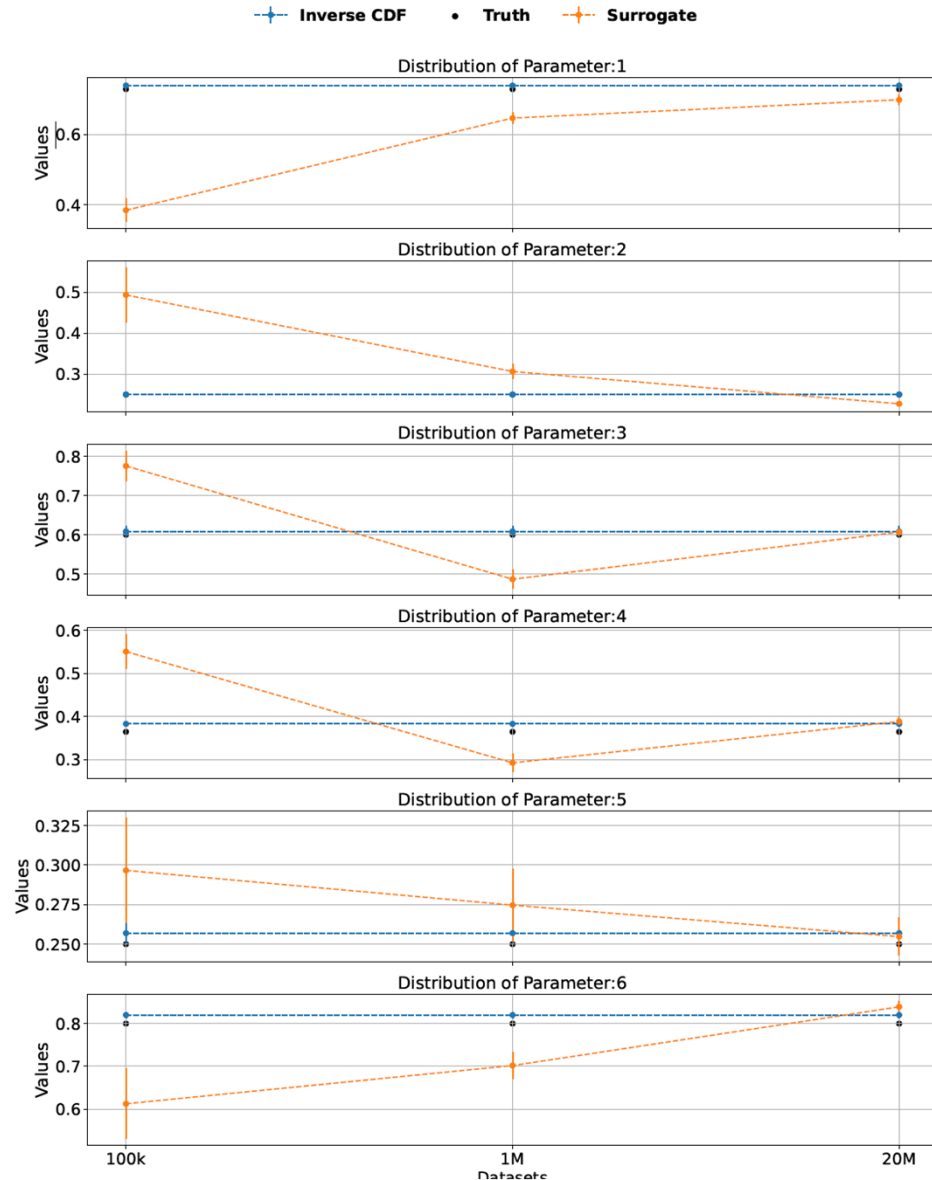
Reconstructed cross sections from parameters using an inverse-CDF sampler



Reconstructed cross sections from parameters using a Surrogate

Training Samples for Surrogate Model

Comparison of parameter distributions for six parameters across datasets of different sizes using the inverse CDF, surrogate, and ground truth values. Error bars indicate the standard deviation of the predicted parameters.



Comparison

■ Differentiable Samplers

○ Pros

- Accurate
- No pre-defined range of parameters
- No training needed

○ Cons

- Computational Costly
- Difficulty in implementation to ensure differentiability

■ Generative AI-based Surrogates

○ Pros

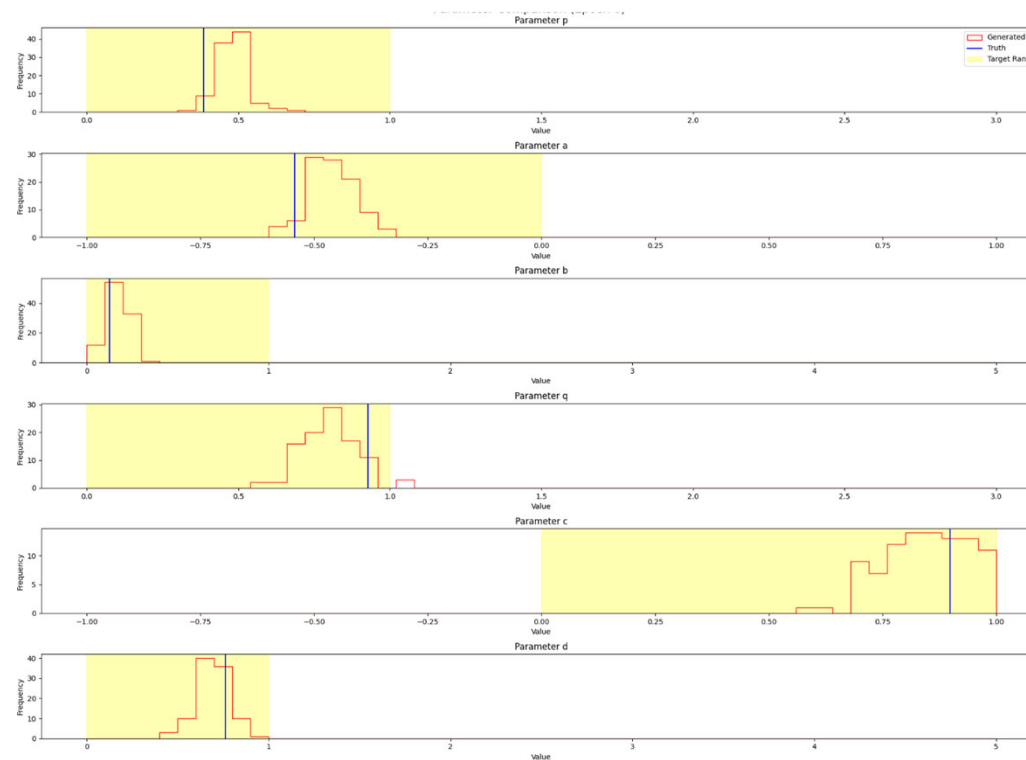
- Fast (neural network evaluation)
- Differentiable
- Easy to incorporate into inference framework

○ Cons

- Approximation error
- Need to train with a lot of samples
- Pre-defined range of parameters

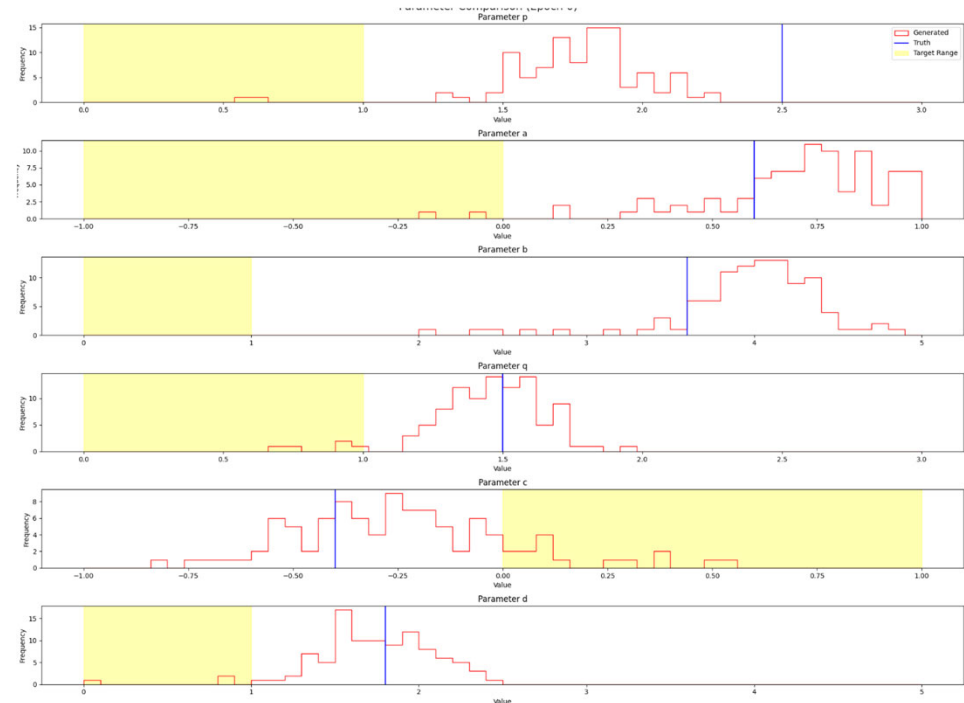
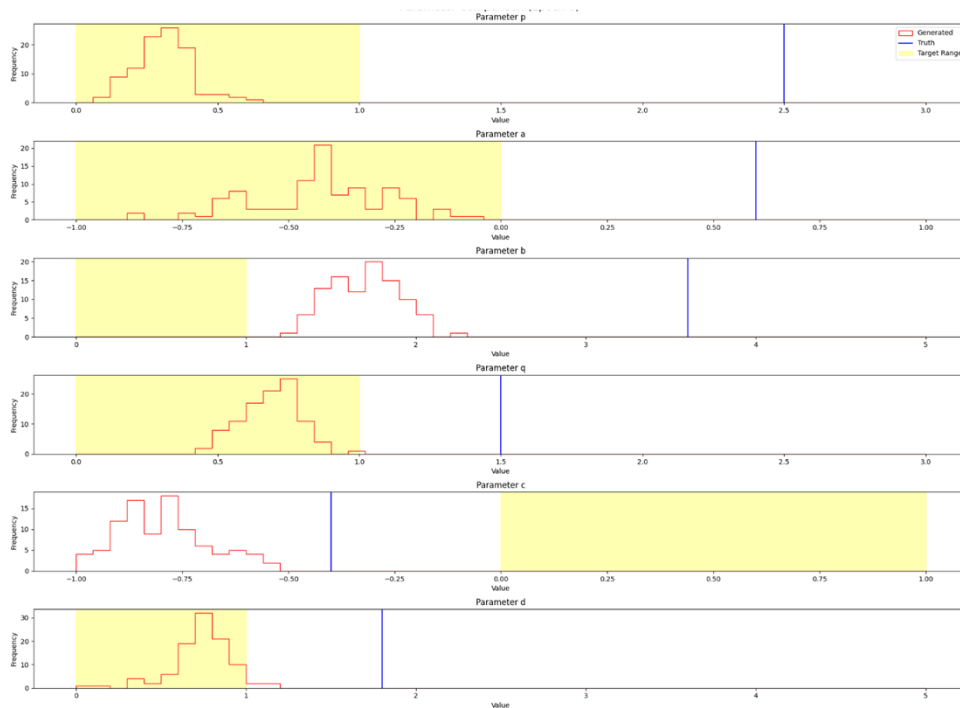
Training Range Problem

- DM model trained on parameters sampled in parameter ranges (shaped)
- Applied to test events generated by parameters within range



Active Learning to Address the Training Range Problem

- Pre-trained DM model on events generated by out-of-range
- After a few steps of adaptive learning



Preliminary Results (Unpublished)

Summary

- Inverse Problem

- Ill-posedness
- Generative AI is natural to solving inverse problem
 - Generate multiple solutions
 - Diffusion Model yields better fidelity than GAN and Variational Autoencoder
- Training Range

- Challenges

- Stability
- Robustness
- Uncertainty Quantification
- Programming physics into the deep learning framework
- Extracting physics from inverse inferencing

References

- M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, M. P. Kuchera, Y. Li., “Variational Autoencoder Inverse Mapper: An End-to-End Deep Learning Framework for Inverse Problems.” Proceedings of International Joint Conference on Neural Networks (IJCNN2021), 2021.
- M. Almaeen, Y. Alanazi, N. Sato, W. Melnitchouk, Y. Li, “Point Cloud-based Variational Autoencoder Inverse Mappers (PC-VAIM) - An Application on Quantum Chromodynamics Global Analysis,” Proceedings of IEEE International Conference on Machine Learning and Applications (ICMLA2022), 2022.
- K. Braga, M. Diefenthaler, S. Goldenberg, D. Lersch, Y. Li, J. W. Qiu, K. Rajput, F. Ringer, N. Sato, M. Schram, “Toward an event-level analysis of hadron structure using differential programming,” arXiv:2507.15768, 2025.
- M. Almaeen, T. Alghamdi, B. Kriesten, D. Adams, Y. Li, H. W. Lin, S. Liuti, “VAIM - CFF: A variational autoencoder inverse mapper solution to Compton form factor extraction from deeply virtual exclusive reactions,” The European Physical Journal C, 85(5): 499, 2025.
- T. Alghamdi, J. Xu, N. Ramachandra, N. Sato, Y. Li, “Towards an event-level analysis in hadronic physics using generative AI-based surrogates,” IEEE International Conference on Tools with Artificial Intelligence (ICTAI2025), 2025.

Acknowledgements

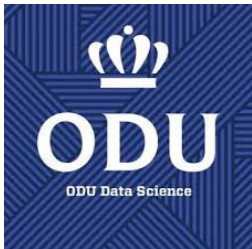
- DOE SciDAC Program



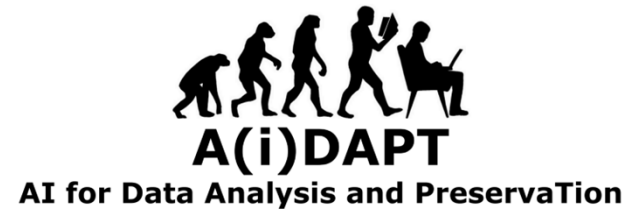
- Center for Nuclear Femtography (CNF)



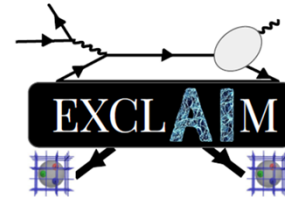
- ODU Data Science Seed Fundings 2023



- A(I)DAPT Collaboration



- EXCLAIM Collaboration



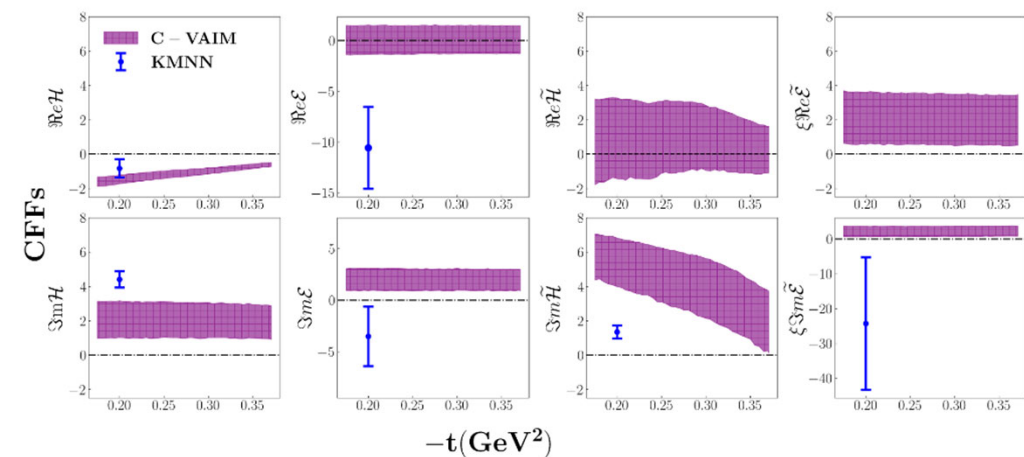
- QuantOm Collaboration



Prediction of CFFs w.r.t. t

- Prediction of CFFs as a function of the Mandelstam variable t for Jlab 6GeV Kinematics
 - $X_{\text{Bj}} = 0.35$
 - $Q^2 = 1.9 \text{ GeV}^2$

All 8 CFFs



Fixing $\text{Re}\tilde{\mathcal{H}} = 0$, $\text{Re}\tilde{\mathcal{E}} = 0$

