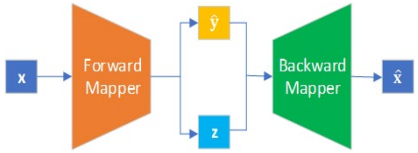


Decoding inverse problems in QCD with ML algorithms

Yaohang Li
Department of Computer Science
Old Dominion University

yaohang@cs.odu.edu

11/28/2023 @AI4EIC Workshop

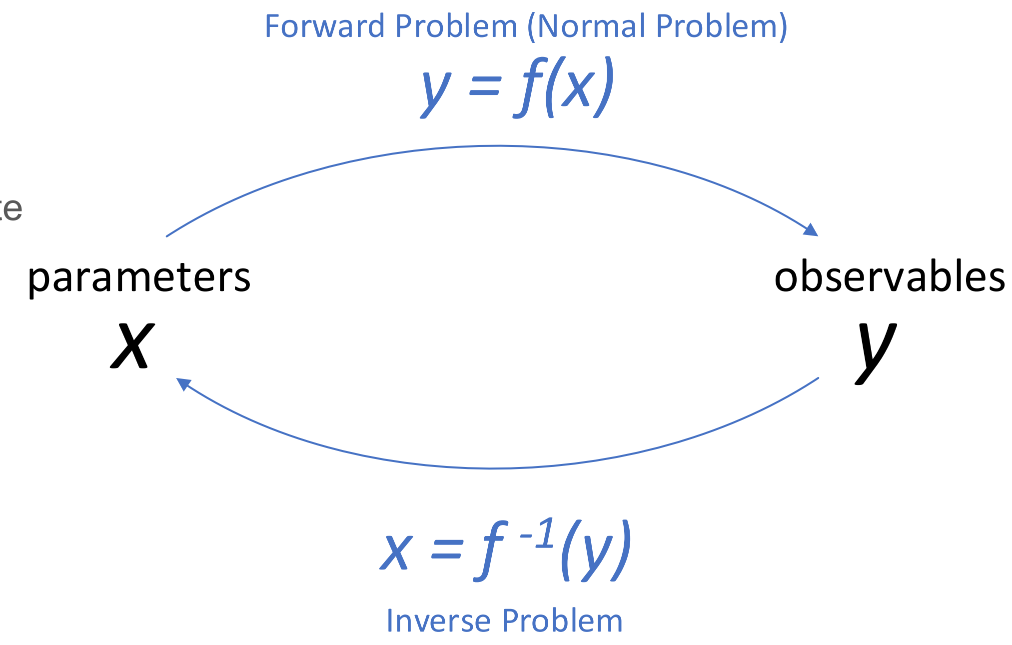
Agenda

- The Inverse Problem
- Variational Autoencoder Inverse Mapper (VAIM) – an End-to-End Framework for Inverse Problem
  - Fundamental Idea
  - Toy Inverse Problems
  - Extracting QCF from DIS Data
- Point Cloud-based VAIM
  - Extracting QCF from DIS Data
- Extraction of Crompton Form Factor as an Inverse Problem

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



The Challenges of Solving the Inverse Problem

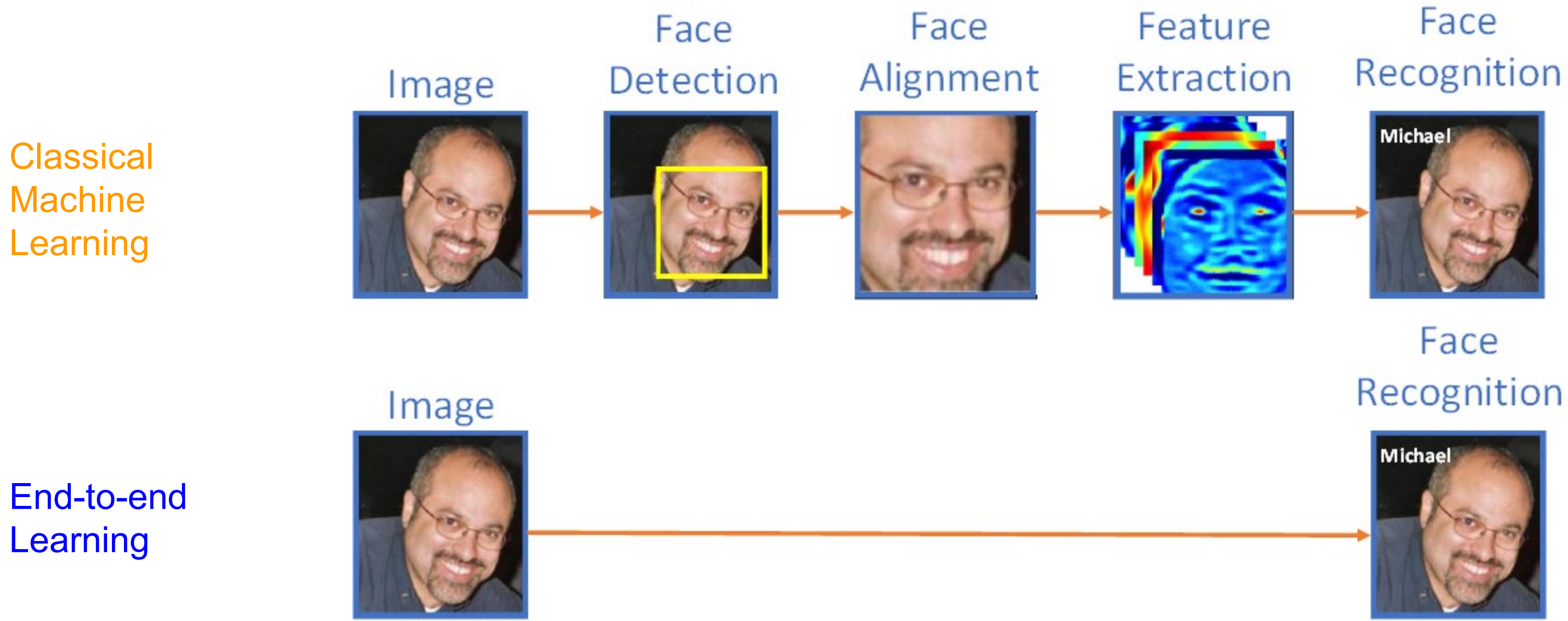
- Ill-posedness
  - Different values of the model parameters may be consistent with the observables
- Curse of Dimensionality
  - Need to explore a huge, high-dimensional parameter space
  - Finding a needle in a haystack

End-to-End Learning for Inverse Problems

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

End-to-End Learning

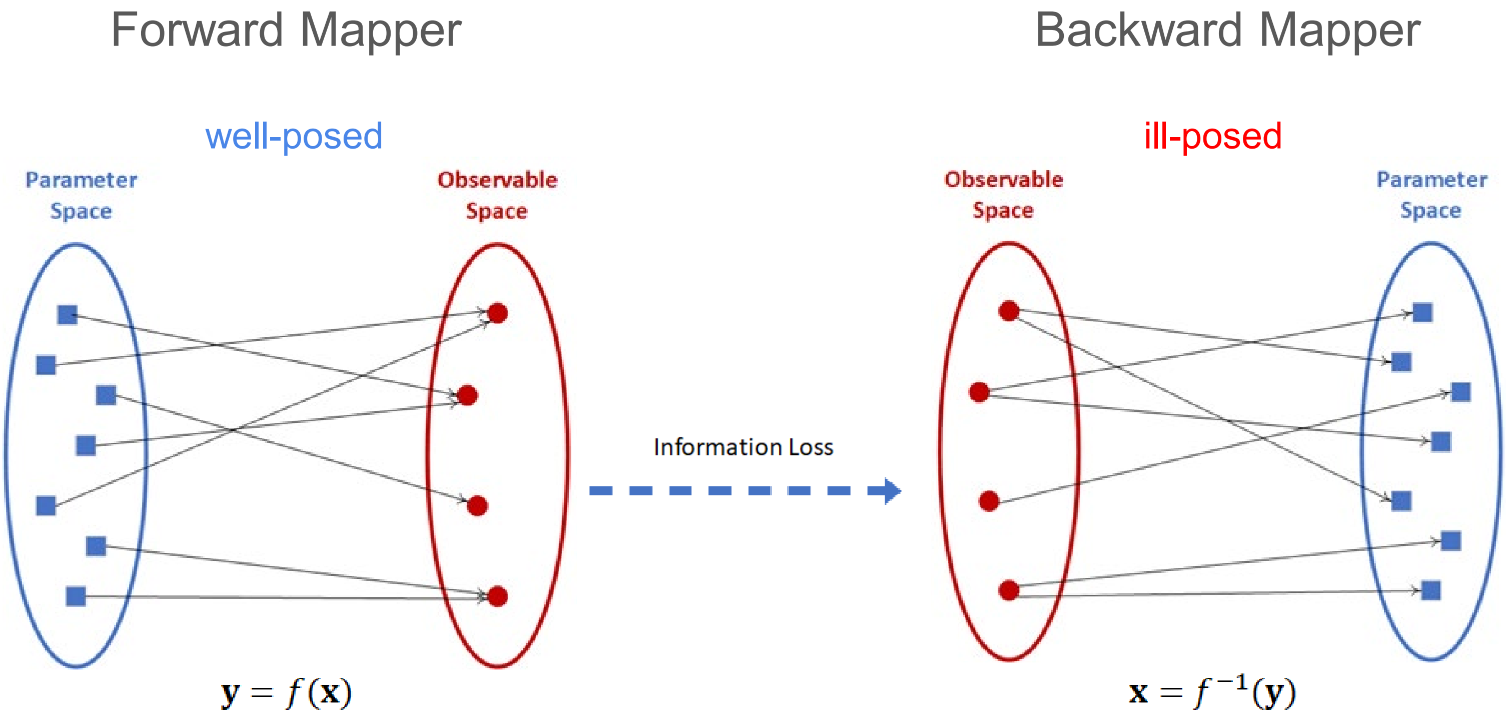
- End-to-End Learning
  - Machine learning model automatically learns all features
  - Directly convert input data into output prediction



Agenda

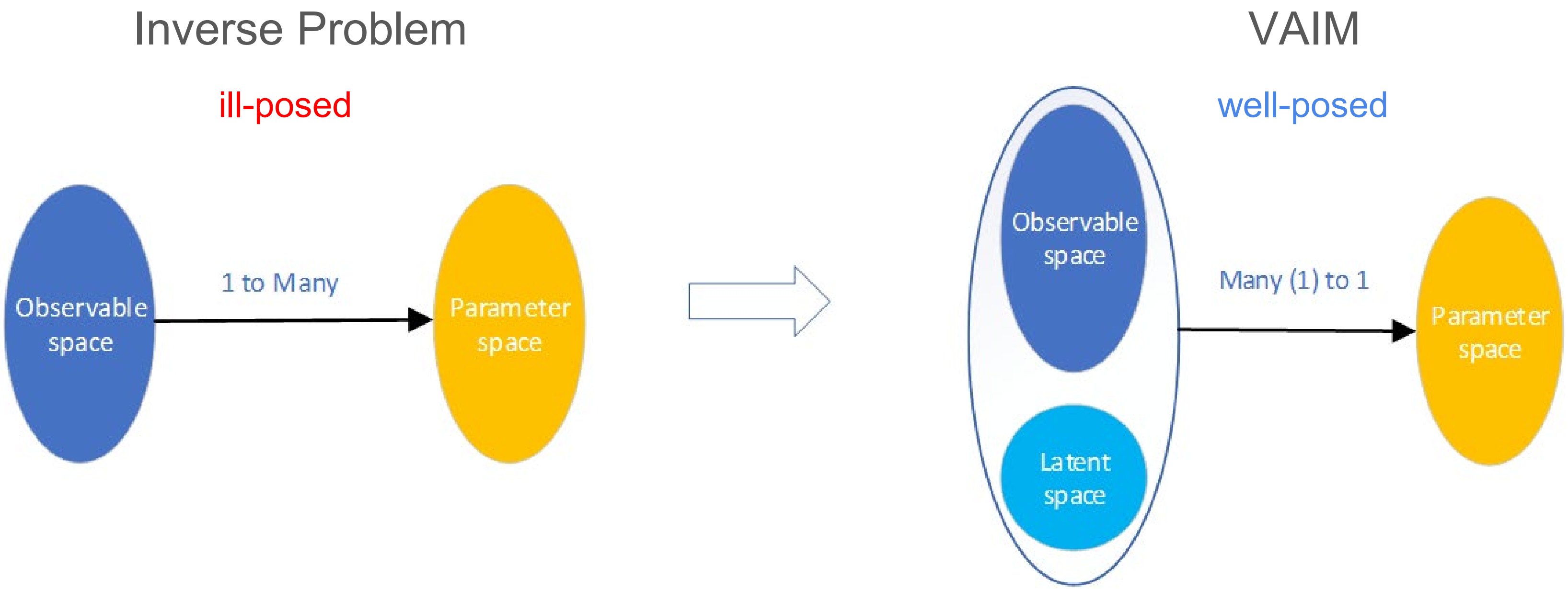
- The Inverse Problem
- Variational Autoencoder Inverse Mapper (VAIM) – an End-to-End Framework for Inverse Problem
  - Fundamental Idea
  - Toy Inverse Problems
  - Extracting QCF from DIS Data
- Point Cloud-based VAIM
  - Extracting QCF from DIS Data
- Extraction of Crompton Form Factor as an Inverse Problem

Ill-posedness of Inverse Problems



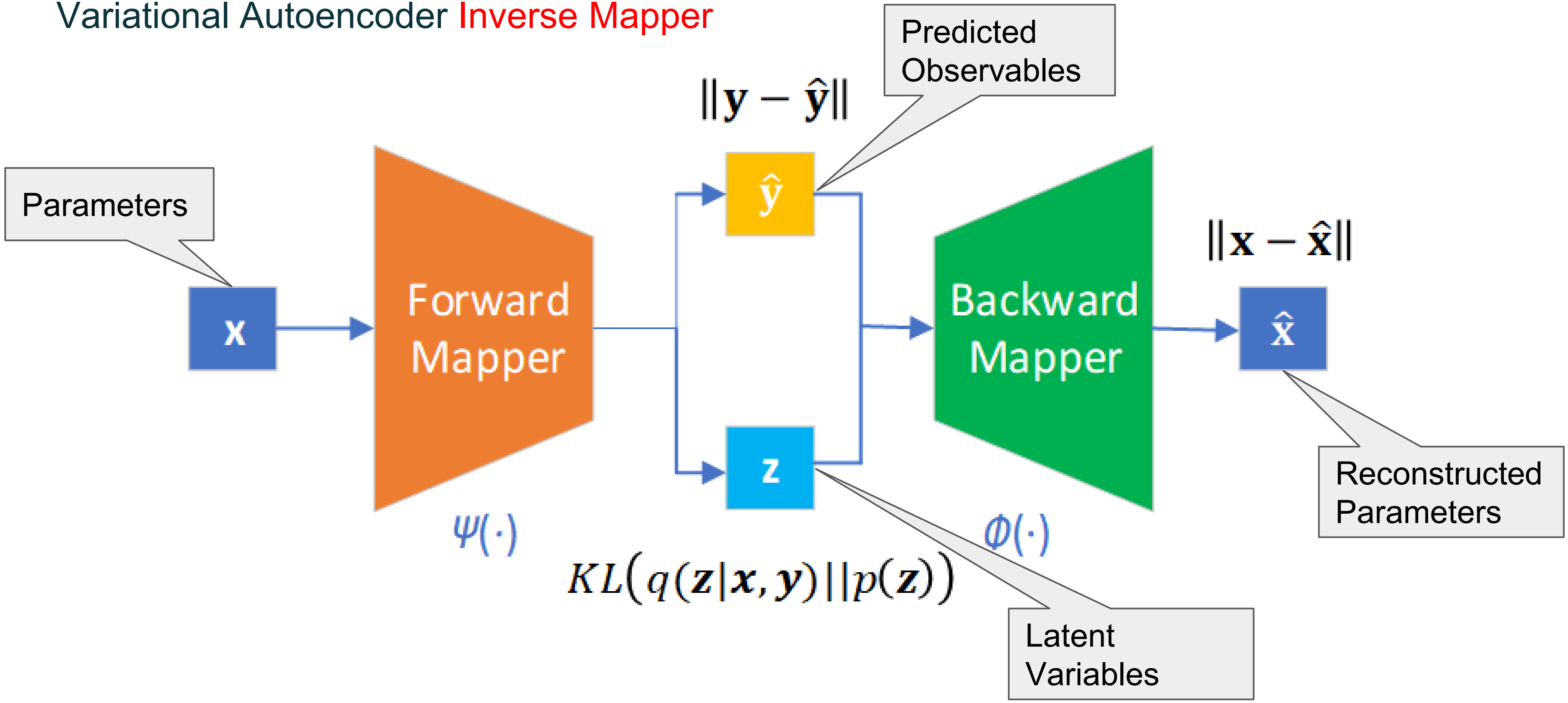
Fundamental Idea

Variational Autoencoder Inverse Mapper

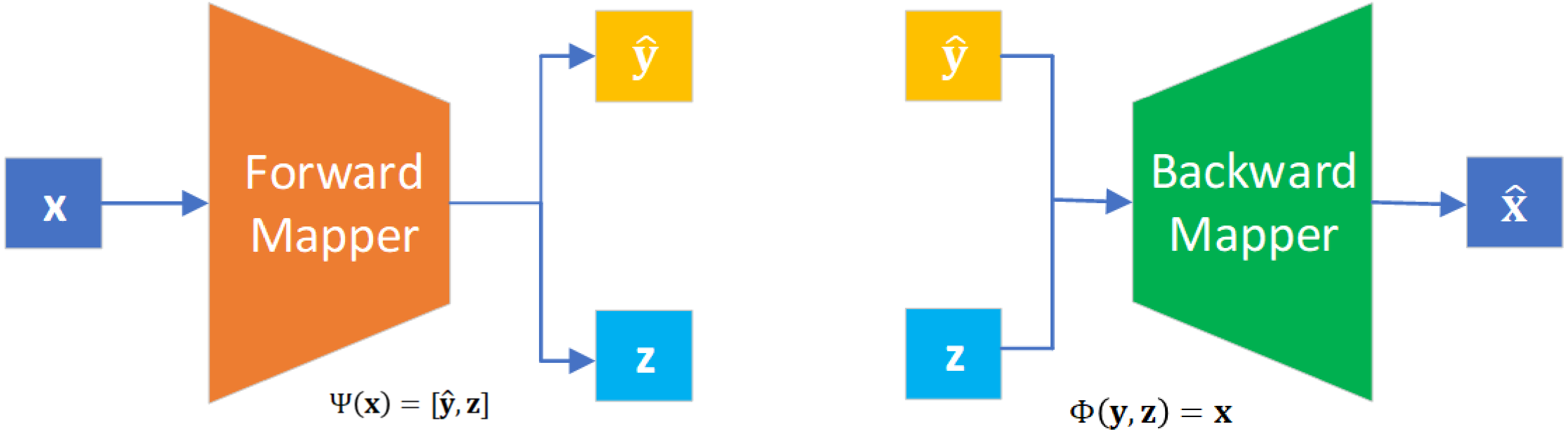


Machine Learning Architecture

Variational Autoencoder Inverse Mapper



Forward Mapper and Backward Mapper



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

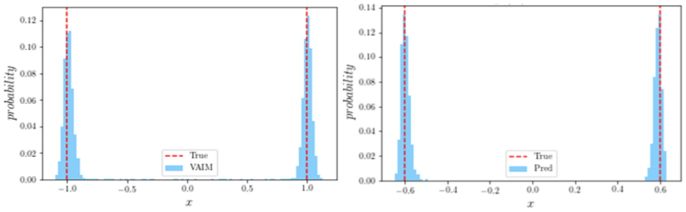
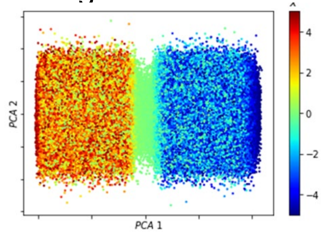
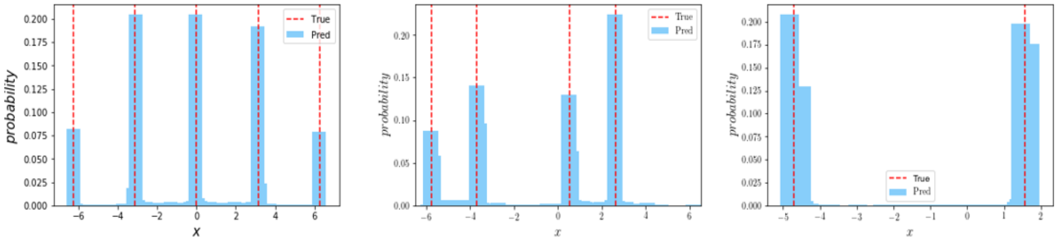
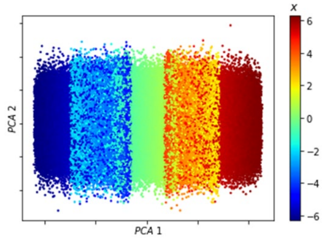
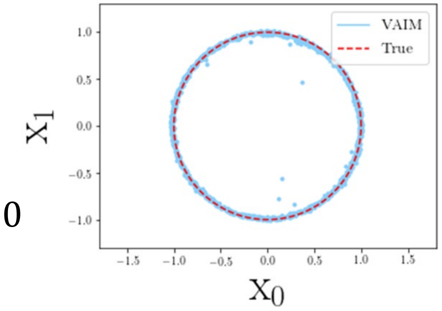
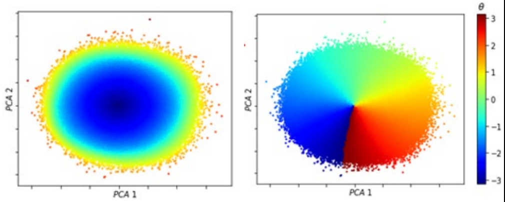
Learn likelihood distribution
$p(\mathbf{x},\mathbf{y}|\mathbf{z})$

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

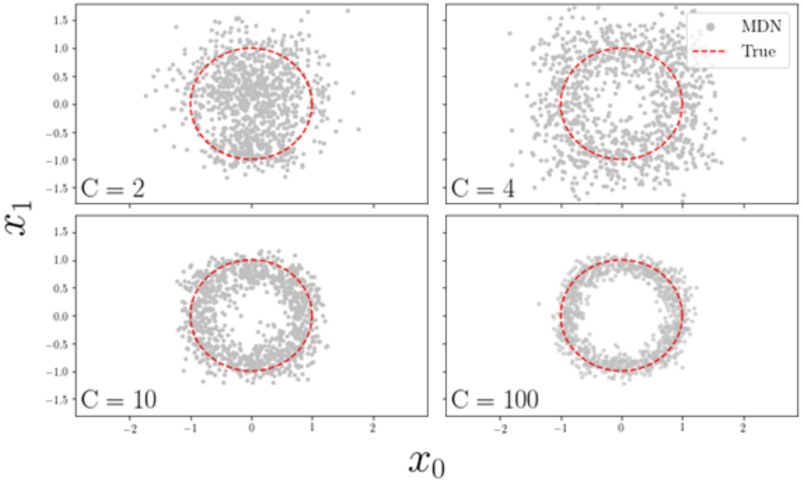
- Variational Autoencoder Theory
  $$\min KL(q(\mathbf{z}|\mathbf{x},\mathbf{y})||p(\mathbf{z}|\mathbf{x},\mathbf{y}))$$
  equivalent to
  $$\min \|\mathbf{y}-\hat{\mathbf{y}}\|_2^2 + \|\mathbf{x}-\hat{\mathbf{x}}\|_2^2 + KL(q(\mathbf{z} \mid \mathbf{x},\mathbf{y}) \,||\, p(\mathbf{z}))$$
  - True prior distribution $p(\mathbf{z})$
    - Select tractable distribution easy to generate
      - Gaussian
      - Uniform

VAIM on Toy Inverse Problems

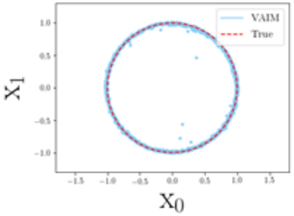
	Multiple Solutions	Latent Space Distribution
$f(x) = x^2$	 Parameter x distribution for $f(x) = 1.0$; Parameter x distribution for $f(x) = 0.36$	Sign Information 
$f(x) = \sin(x)$ $x \in [-2\pi, 2\pi]$	 $f(x) = 1.0$; $f(x) = -0.5$; $f(x) = 0.0$	Period Information 
$f(x_0, x_1) = {x_0}^2 + {x_1}^2$	$f(x_0,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



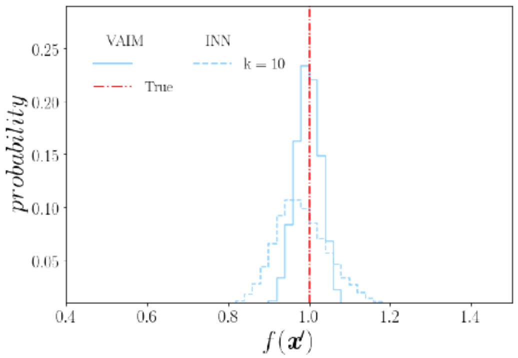
MDN predictions for toy problem $f(x) = {x_0}^2 + {x_1}^2$



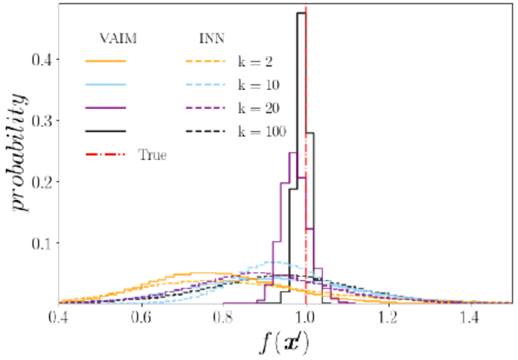
VAIM predictions for toy problem $f(x) = {x_0}^2 + {x_1}^2$

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



(a)



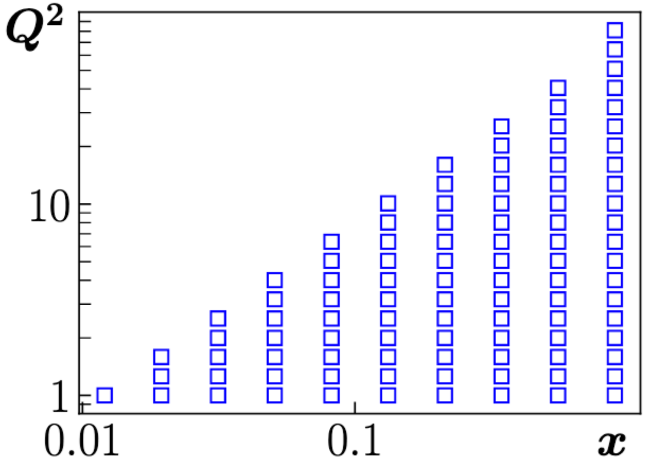
(b)

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

Ardizzone et al., arXiv:188.04730

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

Phase Space



- 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

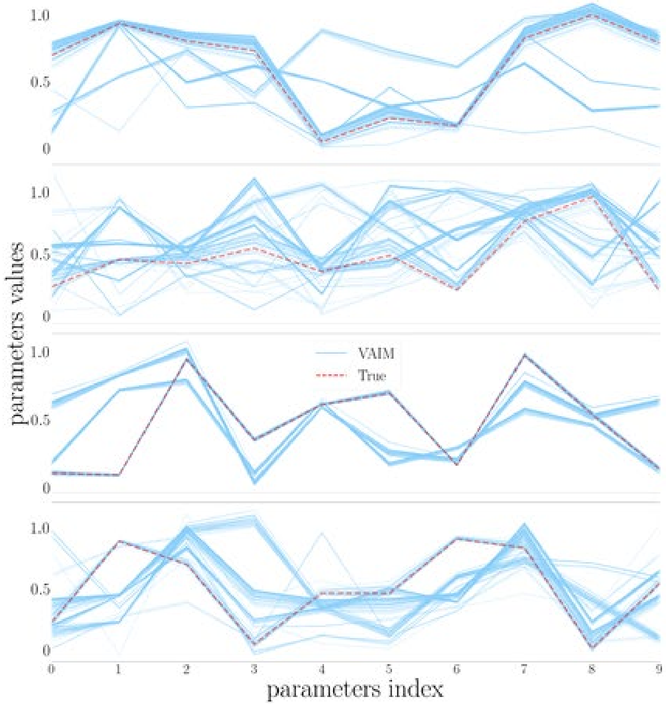
$$u(x,Q^2) = N_u(Q^2)\, x^{\alpha_u(Q^2)}(1-x)^{\beta_u(Q^2)}(1+\gamma_u(Q^2)\sqrt{x}+\delta_u(Q^2)\, x),$$

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

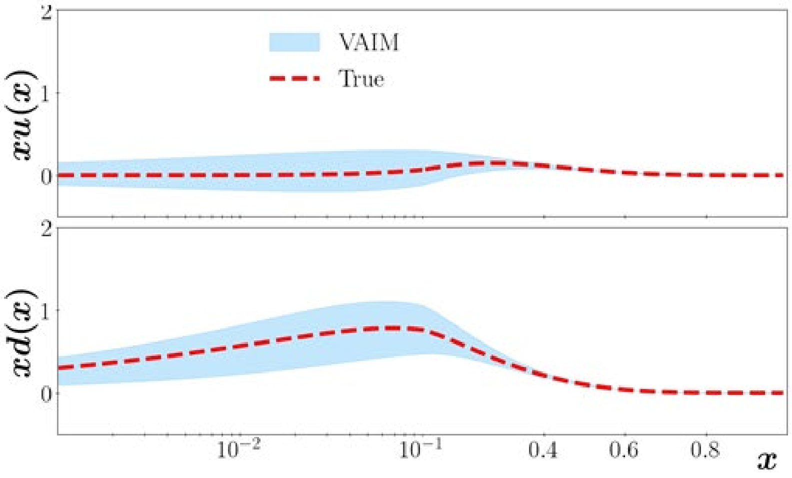
- Shape parameters $p = \{N_{u,d},\, \alpha_{u,d},\, \beta_{u,d},\, \gamma_{u,d},\, \delta_{u,d}\}$ are parameterized as

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

VAIM Results in Toy DIS Problem



Parameter distributions generated by VAIM in four control samples



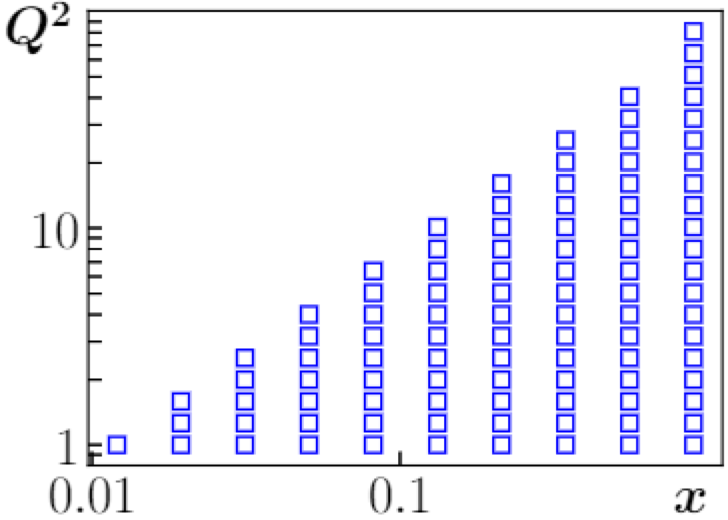
Reconstructed PDF using a control sample

Agenda

- The Inverse Problem
- Variational Autoencoder Inverse Mapper (VAIM) – an End-to-End Framework for Inverse Problem
  - Fundamental Idea
  - Toy Inverse Problems
  - Extracting QCF from DIS Data
- Point Cloud-based VAIM
  - Extracting QCF from DIS Data
- Extraction of Crompton Form Factor as an Inverse Problem

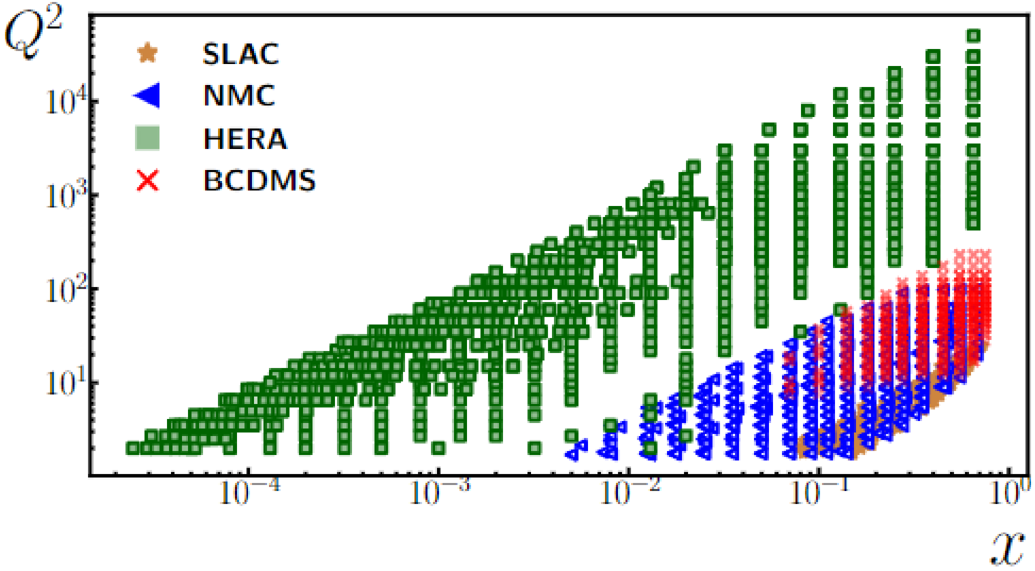
Point Cloud-based VAIM (PC-VAIM)

Limitation of VAIM: Observables across regular, discretized kinematic bins



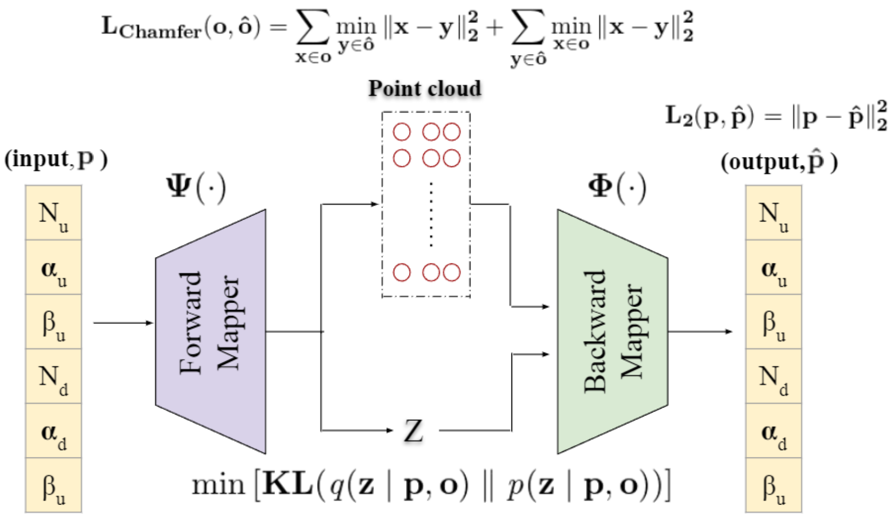
Regular, discretized kinematic bins of *x* and Q^2 where cross sections are evaluated

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

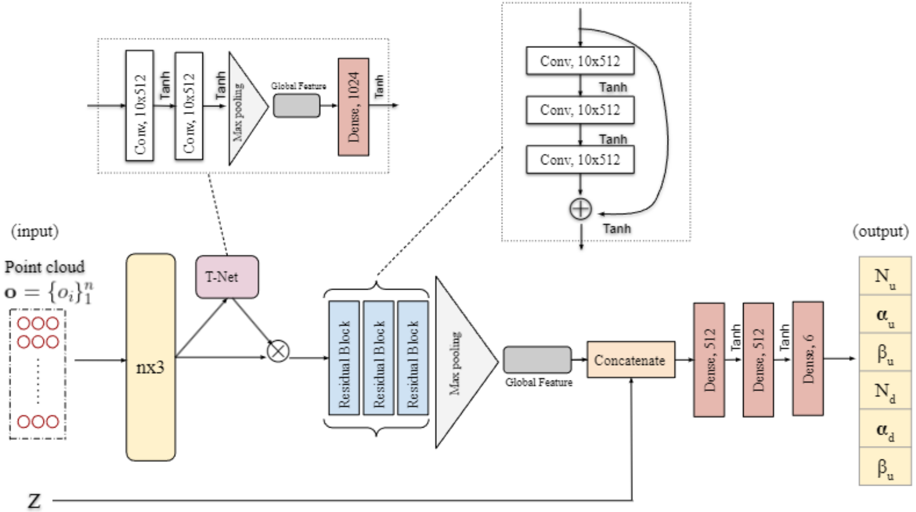


Kinematic bins of SLAC, NMC, HERA, and BCDMS experiments

PC-VAIM Architecture

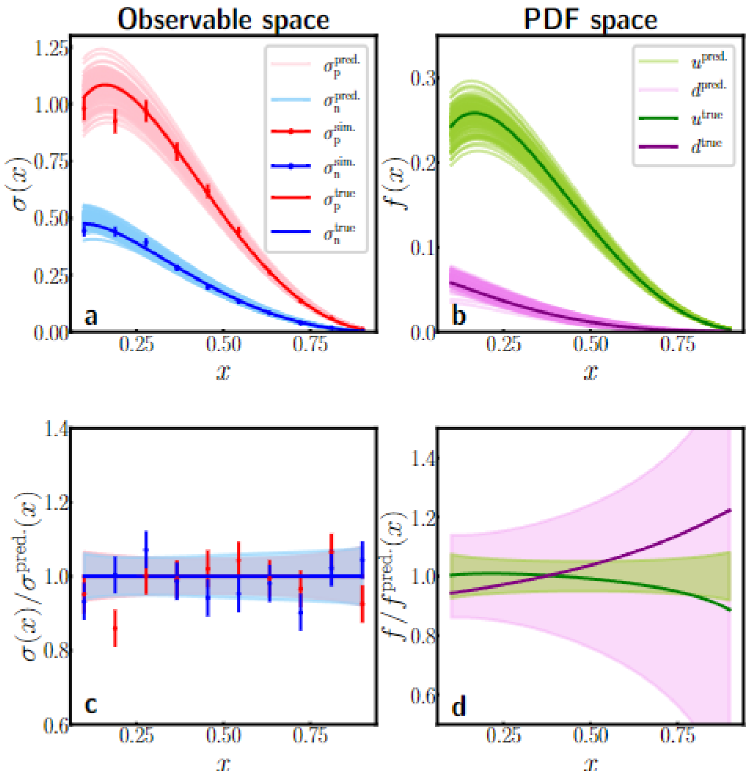


Overall Architecture

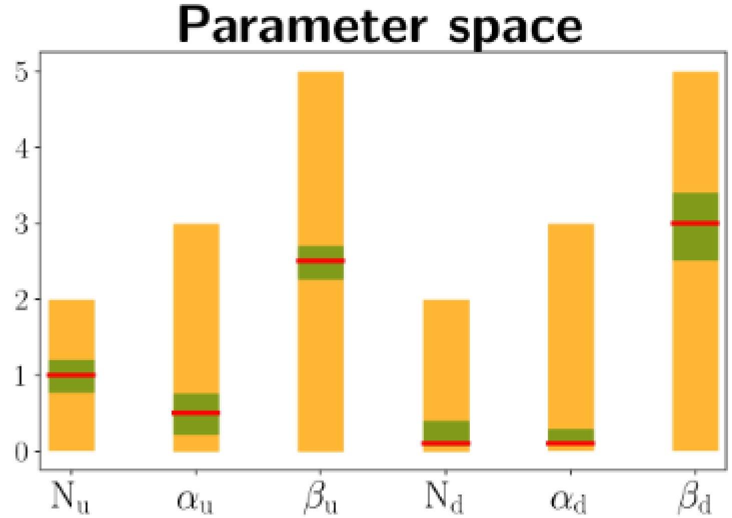


Backward Mapper: A PointNet-based architecture is used to handle the point cloud observable input

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



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.



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

Agenda

- The Inverse Problem
- Variational Autoencoder Inverse Mapper (VAIM) – an End-to-End Framework for Inverse Problem
  - Fundamental Idea
  - Toy Inverse Problems
  - Extracting QCF from DIS Data
- Point Cloud-based VAIM
  - Extracting QCF from DIS Data
- Extraction of Crompton Form Factor as an Inverse Problem

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

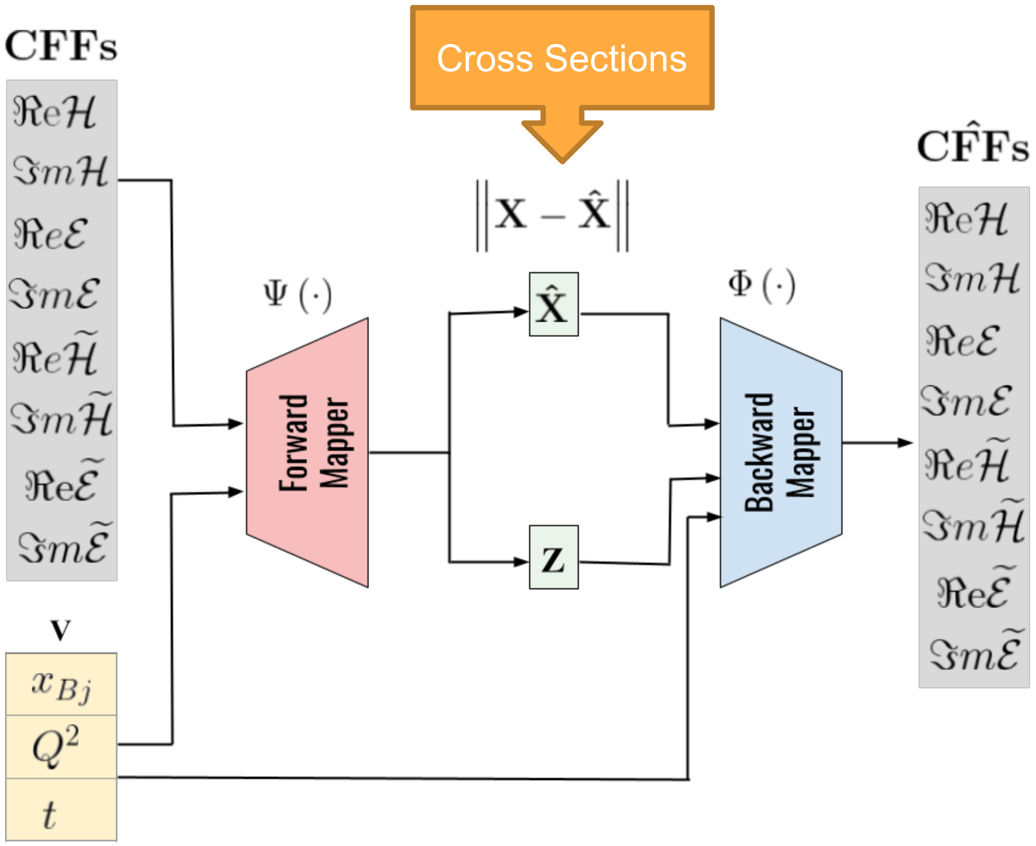
Extraction of CFFs as an Inverse Problem

- Extraction of 8 CFFs from a single polarization observable
  - An inverse problem of extracting 8 unknowns from a single equation

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

  - Quantification of information extracted from experiments

Conditional VAIM Architecture for CFF Extraction



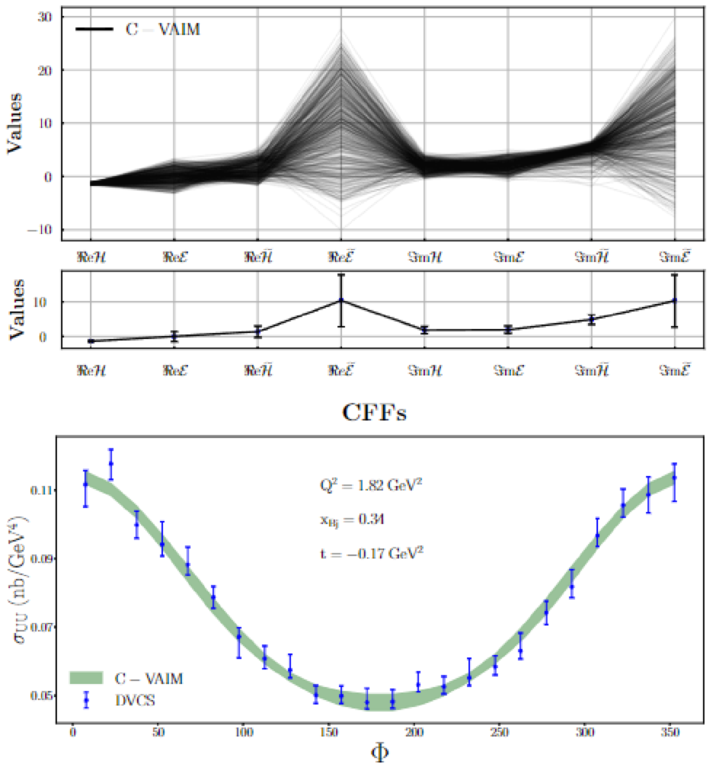
Training of Conditional VAIM

- Training Data:
  - Kinematics values

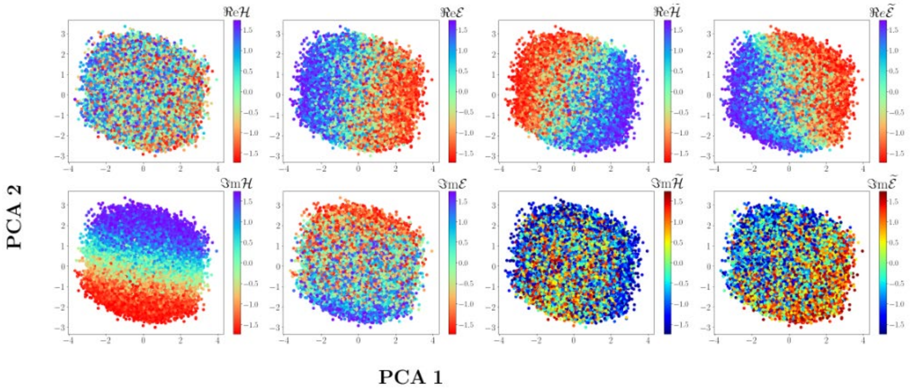
Bin	x_{bj}	t (GeV^2)	Q^2 (GeV^2)
1	0.343	-0.172	1.820
2	0.368	-0.232	1.933
3	0.375	-0.278	1.964
4	0.379	-0.323	1.986
5	0.381	-0.371	1.999

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

Prediction of CFFs



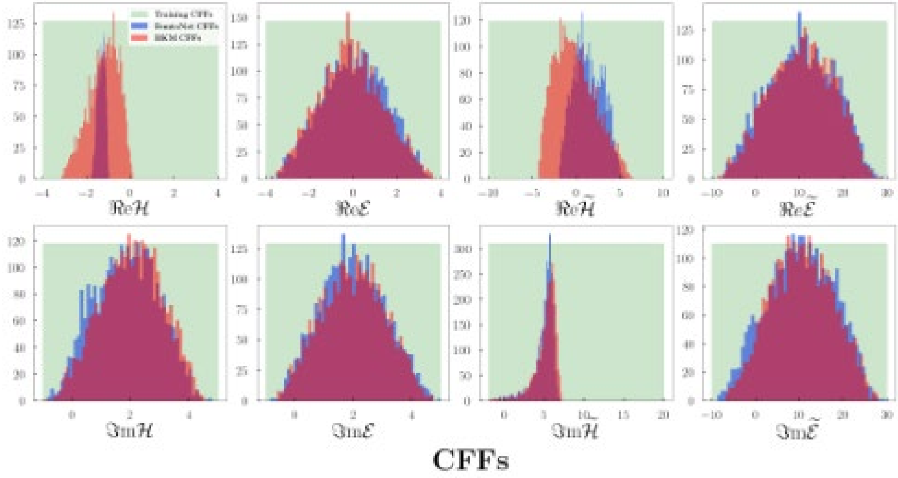
- Predicted CFFs from VAIM-CFF
  - X_{Bj} = 0.343
  - t = -0.172 GeV^2
  - Q^2 = 1.820 GeV^2

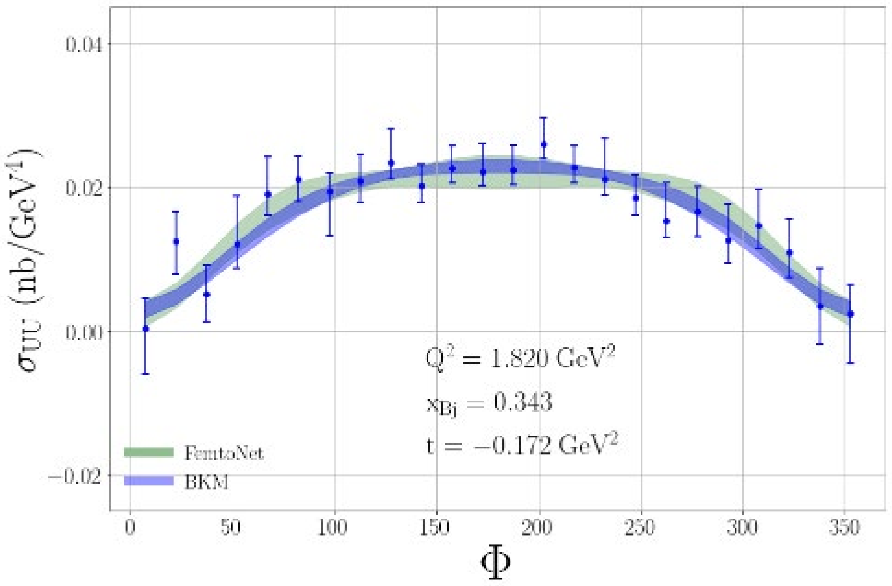


Latent Space Analysis

Training C-VAIM on Different Cross-section Formulations

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

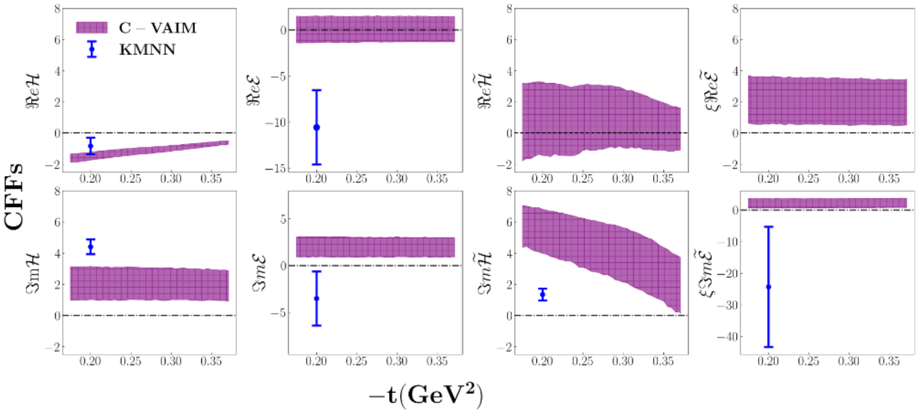




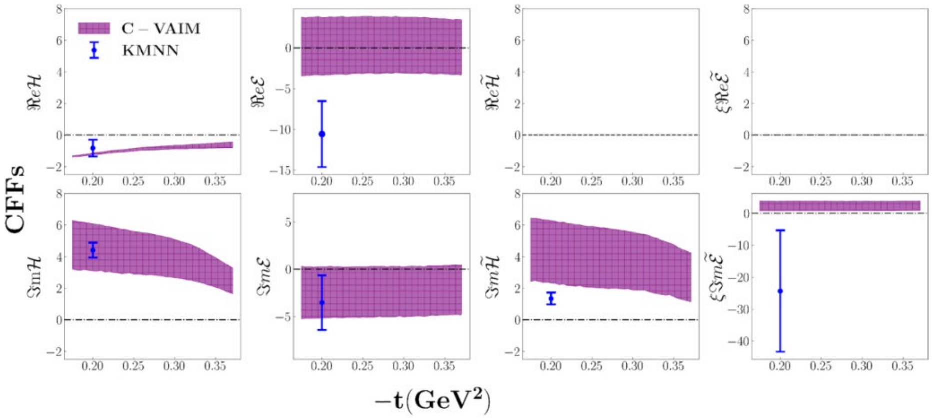
Prediction of CFFs w.r.t. *t*

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

All 8 CFFs



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



Summary

- VAIM
  - An end-to-end deep learning model works well in inverse problems with very different solution patterns
    - A hybrid of supervised and unsupervised learning
  - PC-VAIM is able to handle irregular observable inputs
  - Applications in Extracting QCFs from DIS Data
  - Applications in Extracting CFFs
- Challenges
  - Stability
  - Robustness
  - Programming physics into the deep learning framework

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.
- M. Almaeen, J. Grigsby, J. Hoskins, B. Kriesten, Y. Li, H-W. Lin, S. Liuti, “Benchmarks for a Global Extraction of Information from Deeply Virtual Exclusive Scattering,” arXiv:2207.10766, 2022.
- M. Almaeen, J. Hoskins, B. Kriesten, 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,” in preparation, 2022.

Acknowledgements

- **Collaborators**
  - VAIM, PC-VAIM, Extraction from DIS
    - Nobuo Sato (JLab)
    - Wally Melnitchouk (JLab)
    - Manal Almaeen (ODU)
    - Yasir Alanazi (JLab/ODU)
  - CFFs Extraction
    - Simonetta Liuti (UVA)
    - Brandon Kriesten (CNF/UVA)
    - Joshua Hoskins (UVA)
    - H-W. Lin (MSU)
- **We thank Jianwei Qiu and Xiangdong Ji for helpful discussions.**
- **This work was supported by the Center for Nuclear Femtography (CNF) 2020, 2021, and 2022 funds, and CNF Graduate Fellowship to Manal Almaeen.**