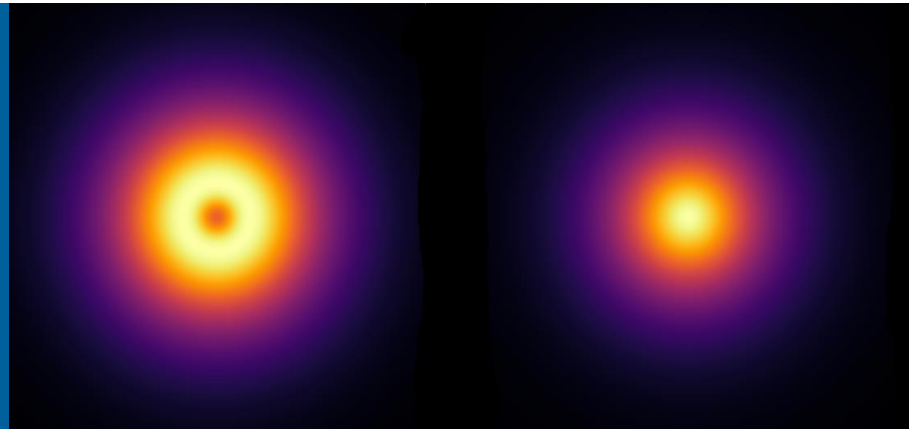
QuantOm Collaboration

www.anl.gov/phy/QuantOm

The authors acknowledge support by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, contract numbers DE-AC02-06CH11357 and DE-AC05-06OR23177, and Office of Advanced Scientific Computing Research, contract no. DE-SC0023472.

THE QUANTOM EVENT-LEVEL INFERENCE FRAMEWORK



DANIEL LERSCH

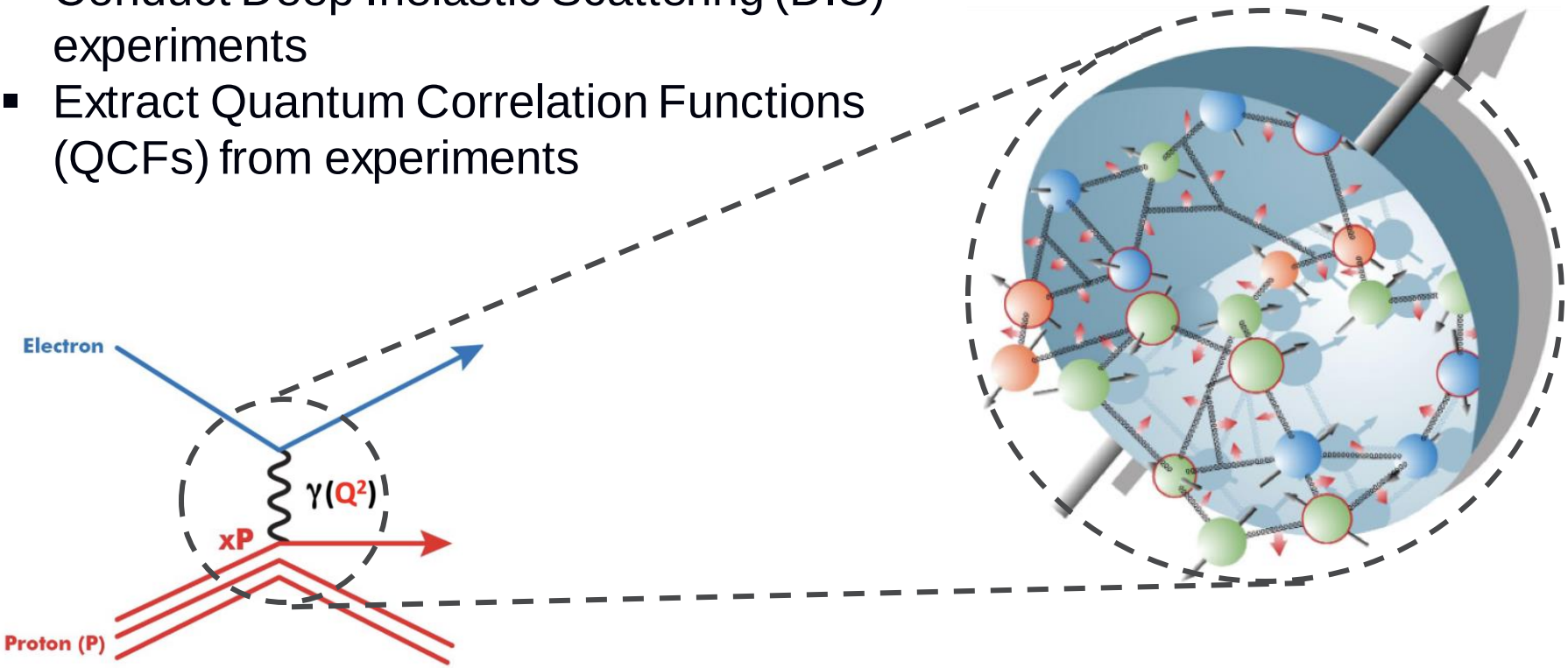
for the QuantOM Collaboration

Argonne NATIONAL LABORATORY
Jefferson Lab
VIRGINIA TECH
NREL NATIONAL RENEWABLE ENERGY LABORATORY
OLD DOMINION UNIVERSITY
THE OHIO STATE UNIVERSITY

26 September 2023

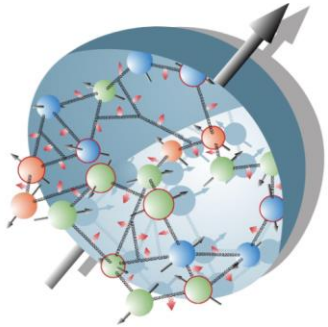
EXPLORING QUARK-GLUON STRUCTURE

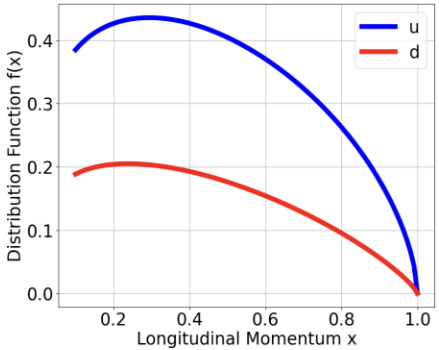
- Want to understand Quark-Gluon system
- Conduct Deep Inelastic Scattering (DIS) experiments
- Extract Quantum Correlation Functions (QCFs) from experiments



COMMONLY USED WORKFLOW

- Input from theory
- QCD factorization / evolution
- Fit experimental data
- Inspire / motivate new experiments ?



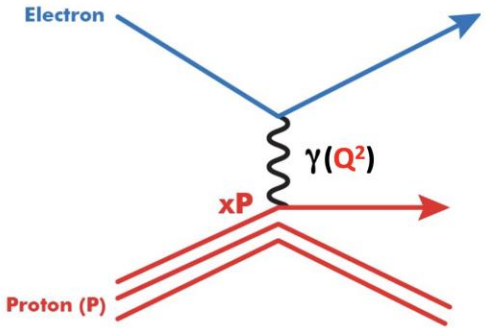


QCD

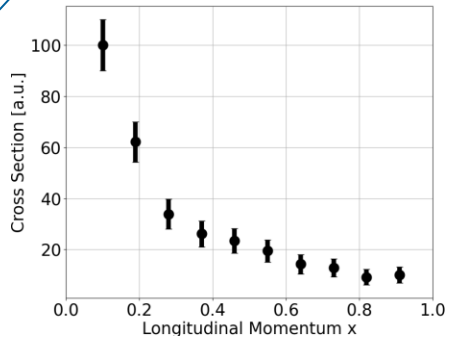
DIS Experiments

Analysis

Extract QCFs

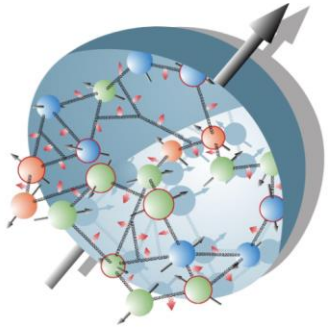


- Event selection
- Histogram observables
- Acceptance corrections
- Radiative corrections
- ...



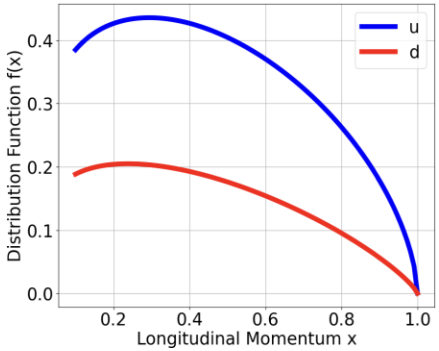
COMMONLY USED WORKFLOW

- Input from theory
- QCD factorization / evolution
- Fit experimental data
- Inspire / motivate new experiments ?



EIC²

Jefferson Lab
Thomas Jefferson National Accelerator Facility

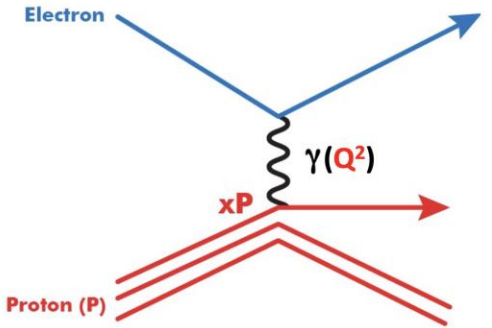


QCD

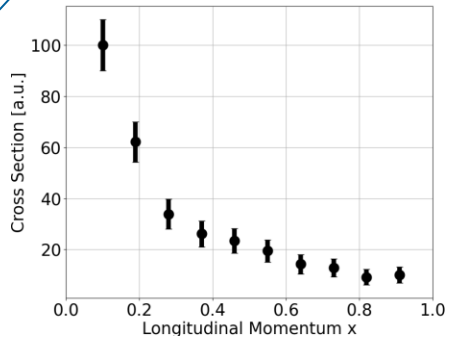
Extract QCFs

DIS Experiments

Analysis

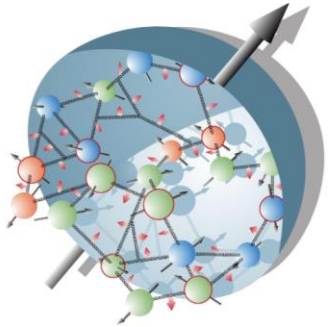


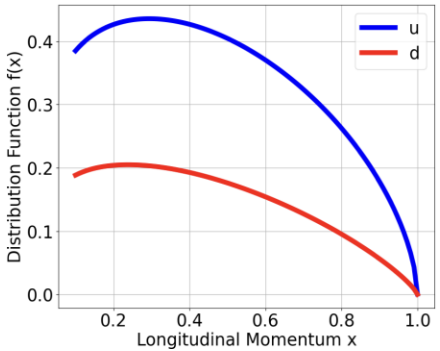
A lot of work goes in here!

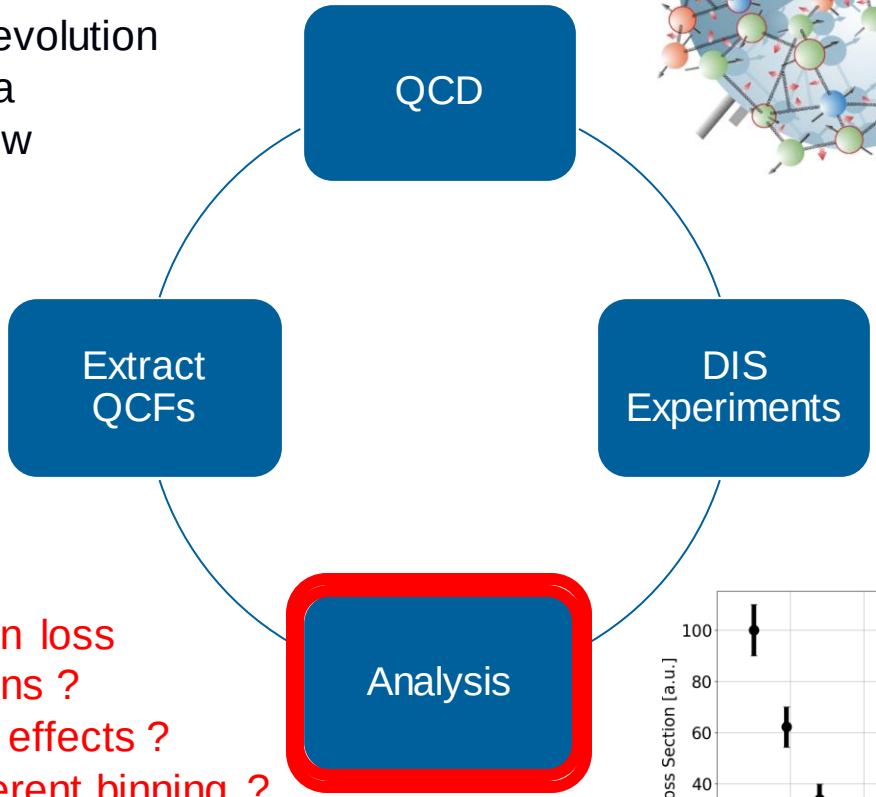


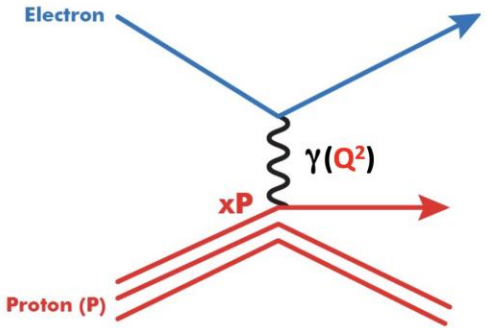
COMMONLY USED WORKFLOW

- Input from theory
- QCD factorization / evolution
- Fit experimental data
- Inspire / motivate new experiments ?

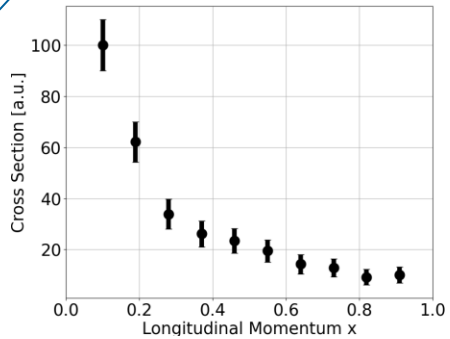




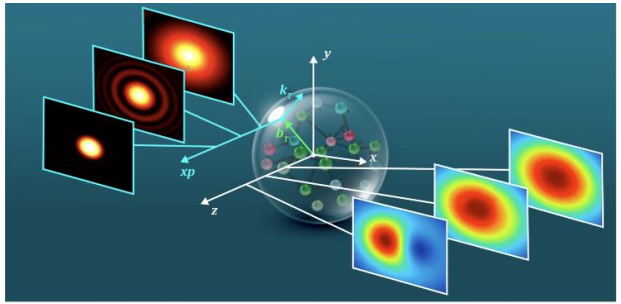


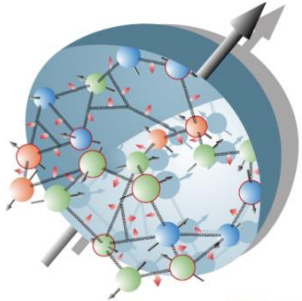


- Possible information loss
- Handling correlations ?
- Binning / smearing effects ?
- Re-iterate with different binning ?
- ...



TOWARDS AN EVENT-LEVEL ANALYSIS





QCD

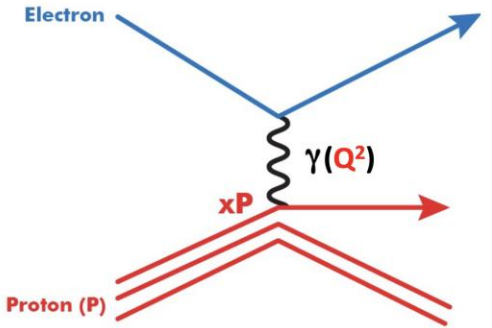
3D Imaging of QCFs

DIS Experiments

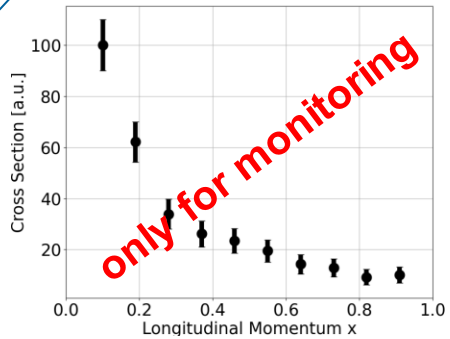
Event-Level Analysis

EIC²

Jefferson Lab
Thomas Jefferson National Accelerator Facility



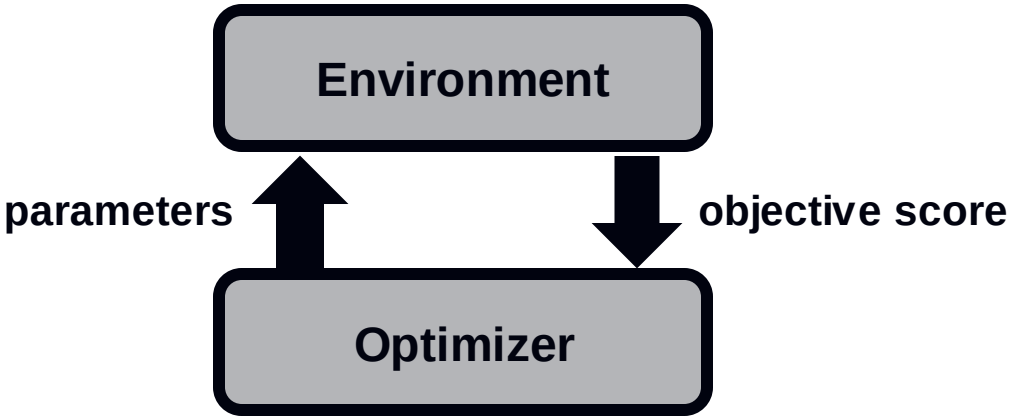
- Utilize all available information (i.e. theory, experiment,..)
- Explore entire feature space in experimental data
- Analyze experimental data "on the fly"



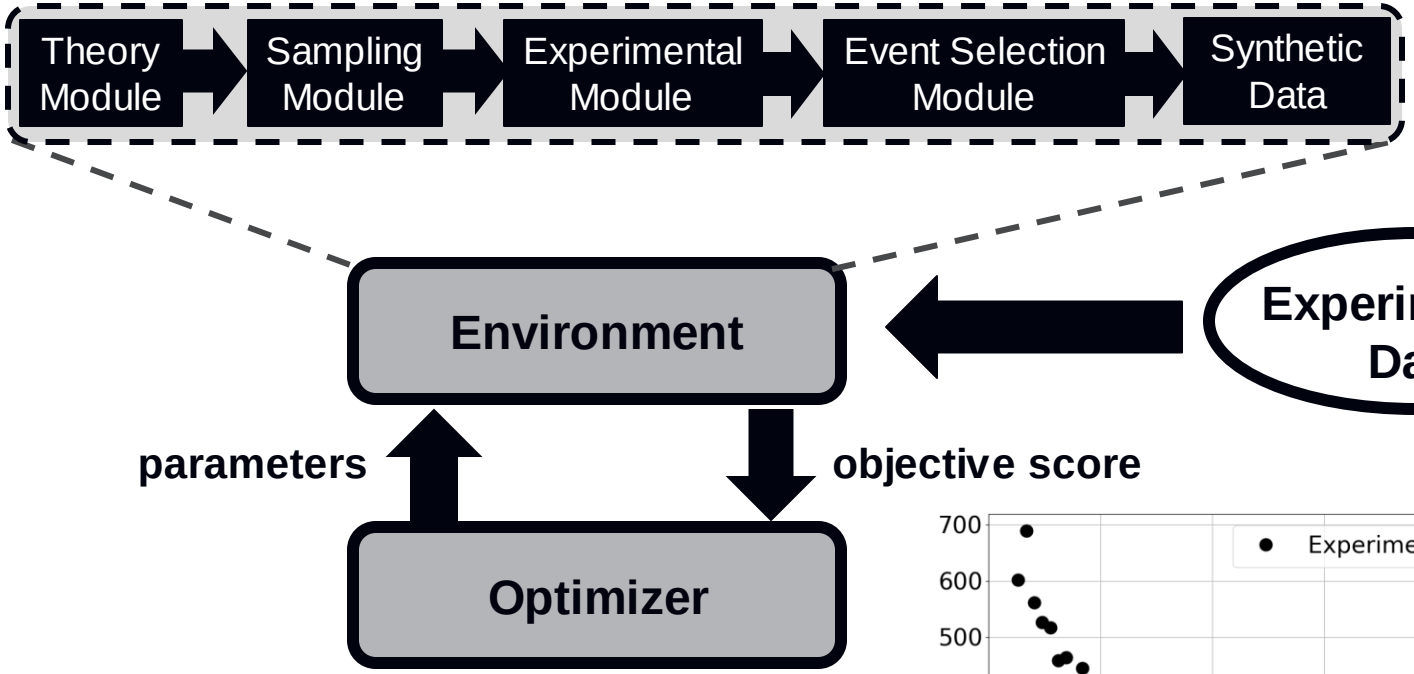
THE QUANTUM NUCLEAR TOMOGRAPHY (QUANTOM) COLLABORATION

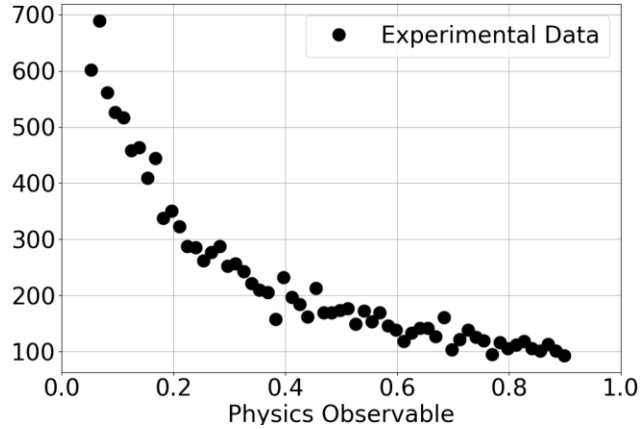
- Part of **Sci**entific **D**iscovery through **A**dvanced **C**omputing (**SciDAC**)
- Interdisciplinary research
  - Applied mathematics
  - Computer and data science
  - Theoretical and experimental nuclear physics
  - High performance computing
- Collaboration between multiple national research institutions
  - Jefferson Lab
  - Argonne National Laboratory
  - Old Dominion University
  - Virginia Tech

THE QUANTOM WORKFLOW

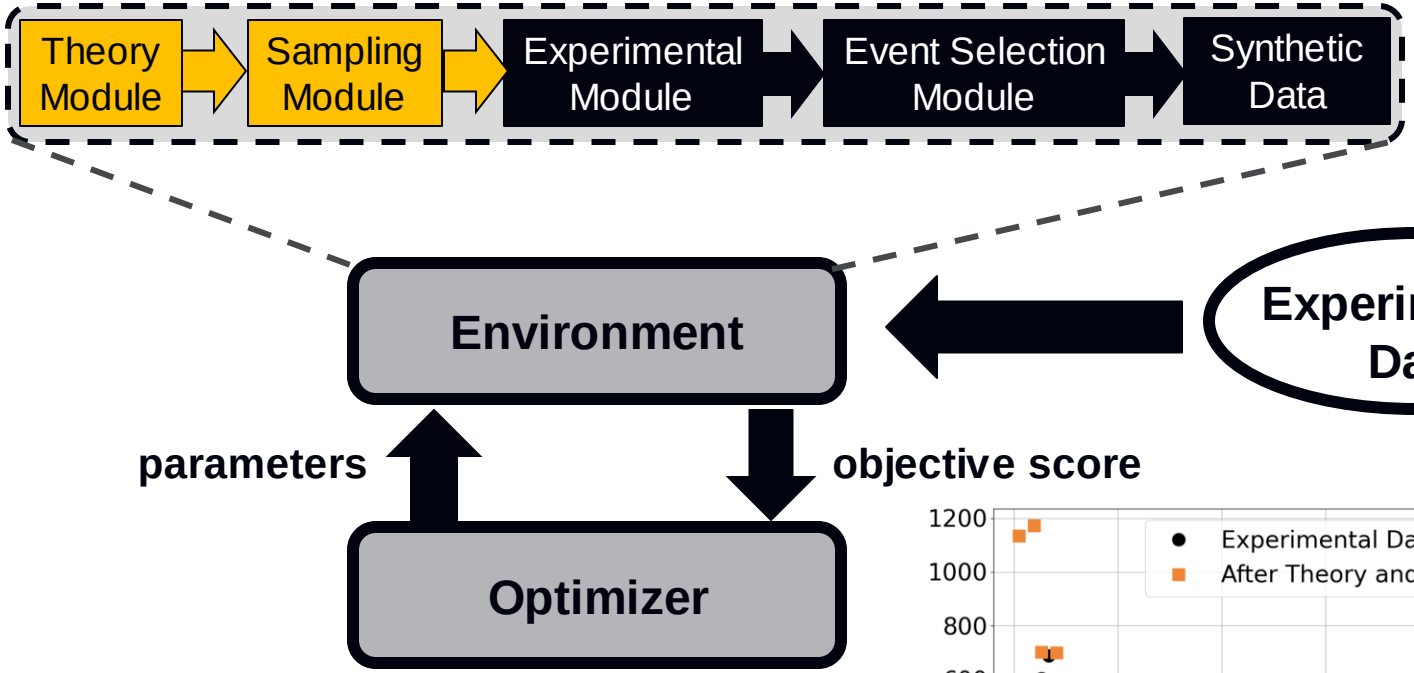


THE QUANTOM WORKFLOW

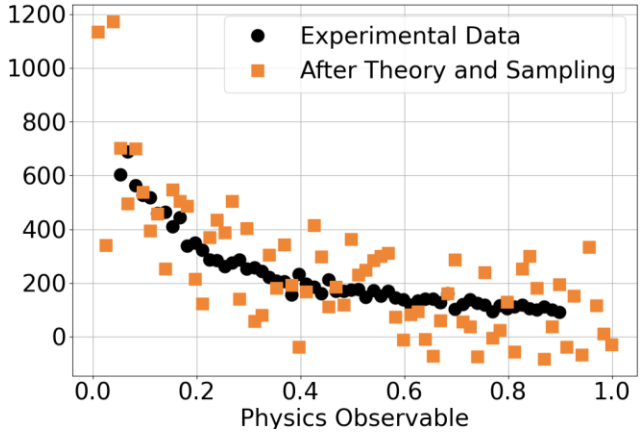




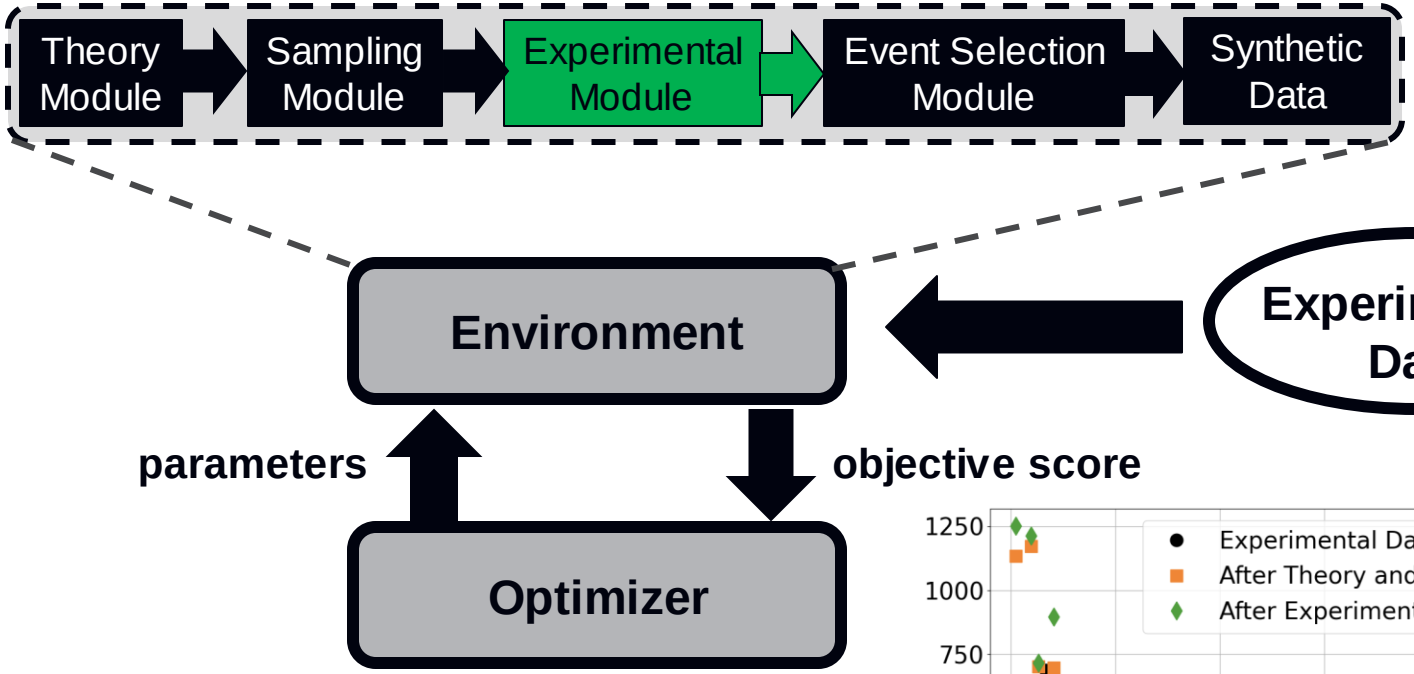
THE QUANTOM WORKFLOW



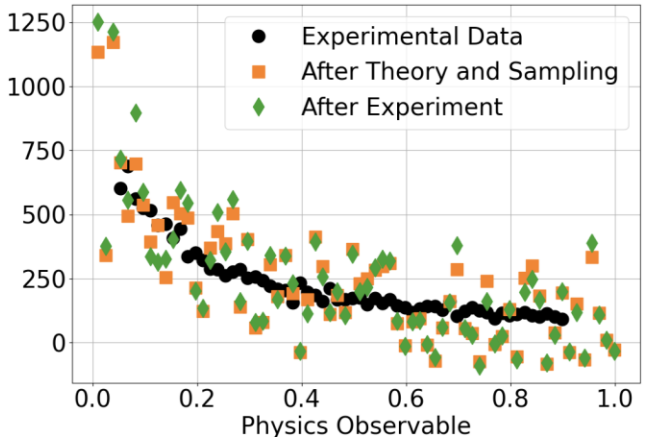
- Convert parameters to parton density functions (PDFs)
- Include higher order and radiate corrections
- Sample events from PDFs (e.g. via MCMC)



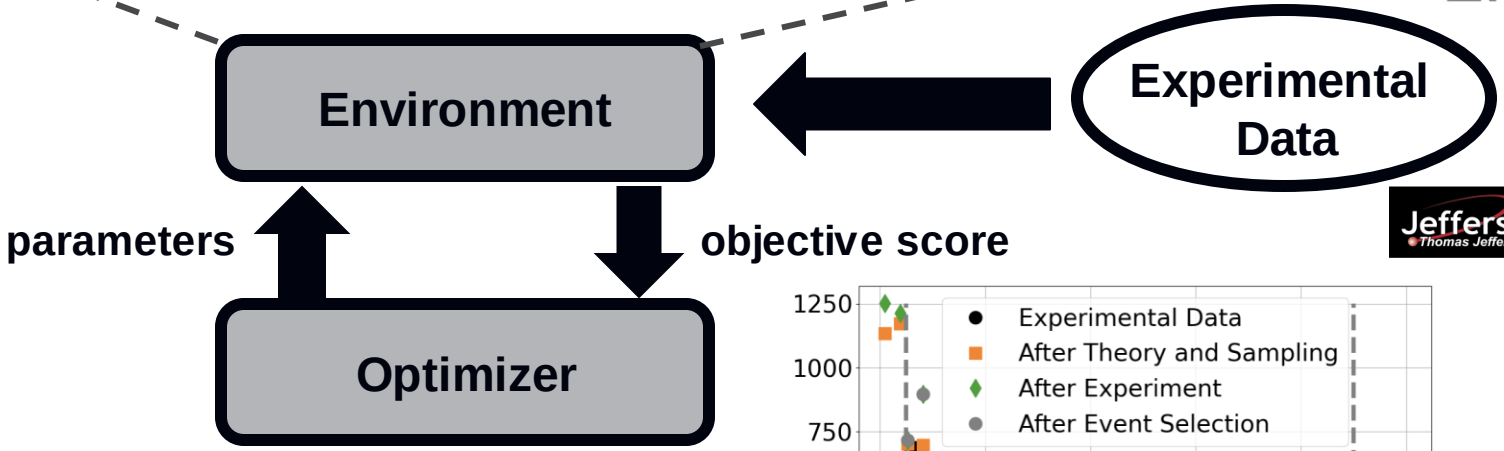
THE QUANTOM WORKFLOW



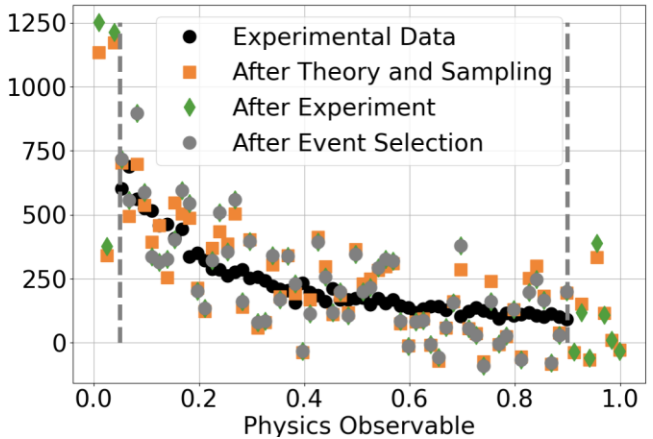
- Apply experimental effects (e.g. resolution, acceptance)
- Handle background contributions
- Use surrogate for detector



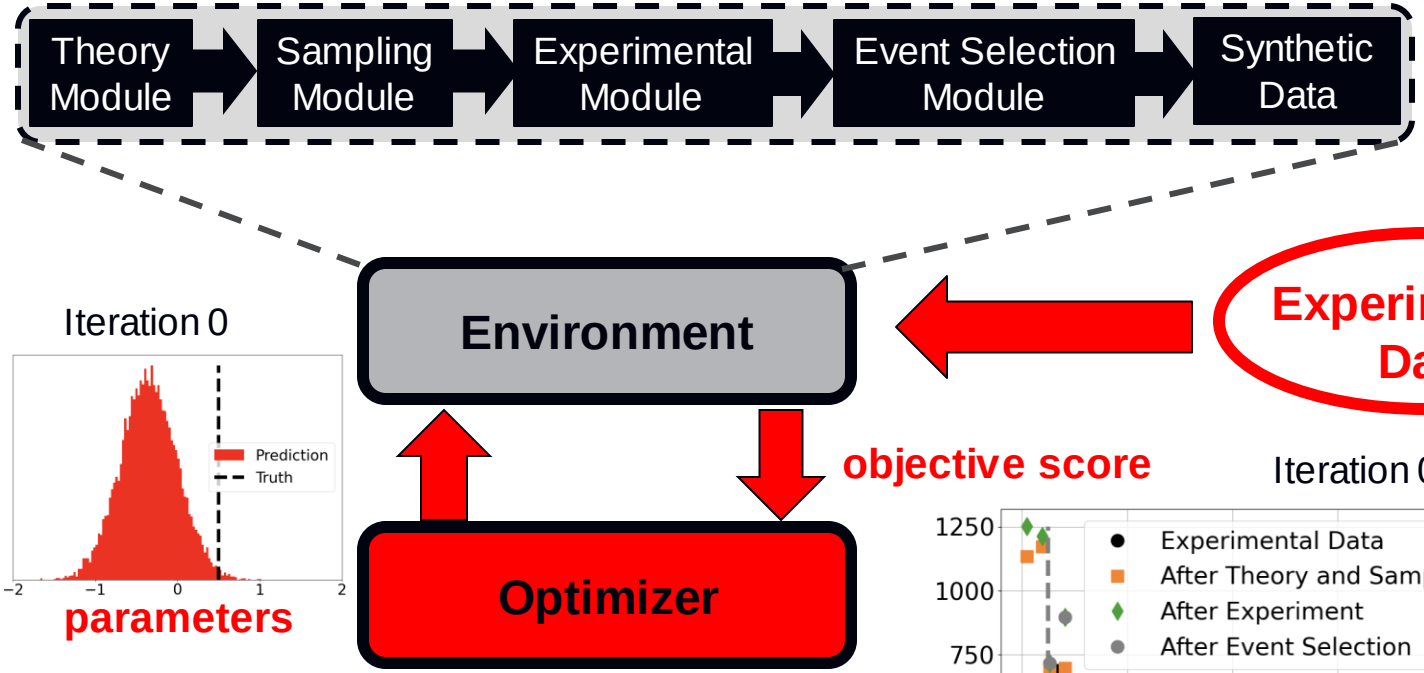
THE QUANTOM WORKFLOW



- Exclude un-physical data points
- Match experimental and synthetic data



THE QUANTOM WORKFLOW

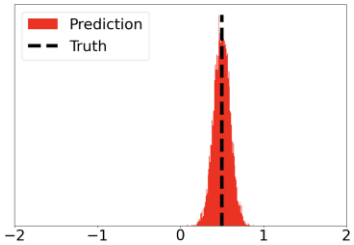


- Iteratively minimize objective score
- Optimizer: GAN generator, RL agent, BO, …
- Objective score: loss, reward, likelihood,…

THE QUANTOM WORKFLOW

Theory Module → Sampling Module → Experimental Module → Event Selection Module → Synthetic Data

Iteration 100



parameters

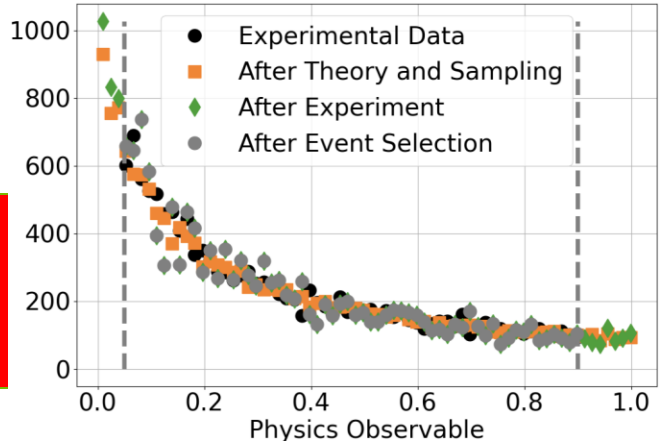
Environment

Optimizer

Experimental Data

objective score

Iteration 100



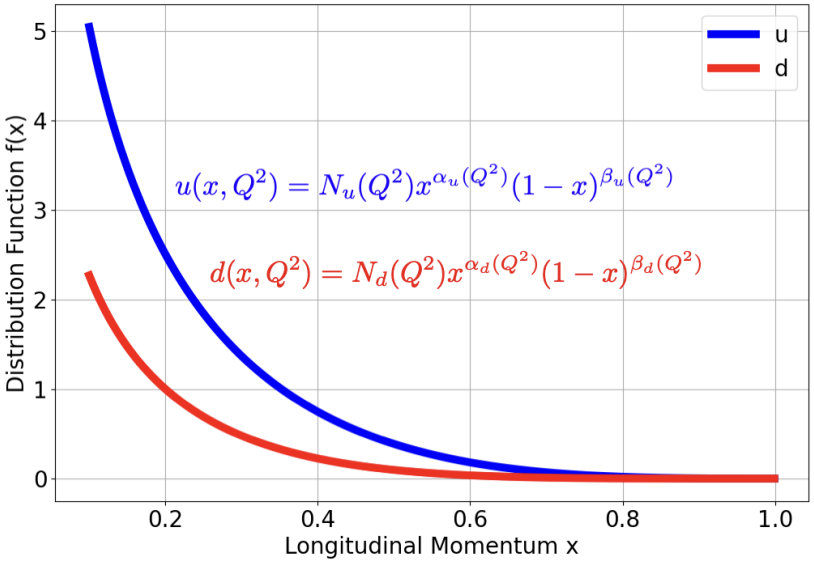
- Iteratively minimize objective score
- Optimizer: GAN generator, RL agent, BO, …
- Objective score: loss, reward, likelihood,…

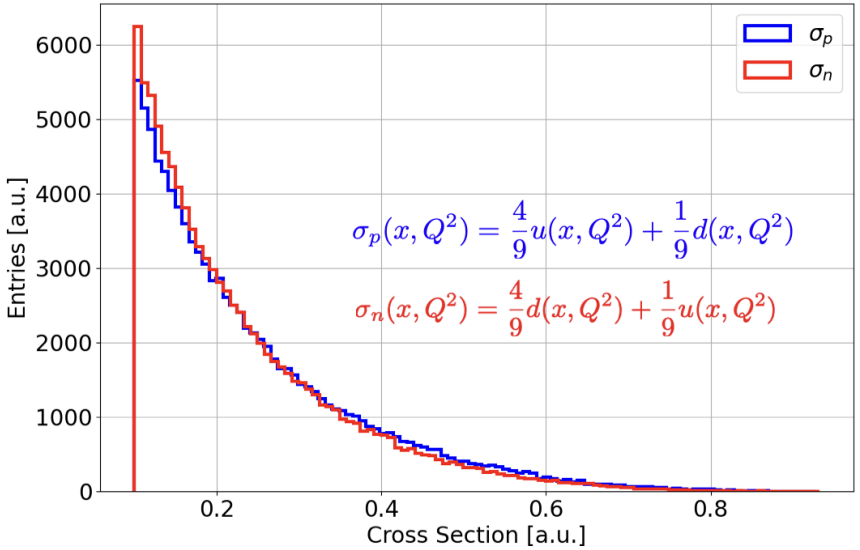
WORKFLOW SPECS

- Written in python
- Supports Tensorflow, Keras, PyTorch,…
- Runs on CPU / GPU
- Modular
  - Change / update / add individual modules
  - Customize entire pipeline
- Each module has its dedicated working group (e.g. theory, experiment,…)
- Fit multiple experiments simultaneously <--> Each experiment has its dedicated module

THE 1D PROXY APP

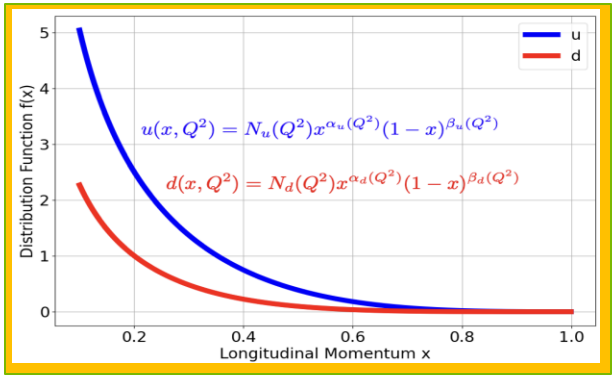
- **Test, debug and benchmark workflow**
- **Given:** Toy data set consisting of cross section "measurements"
- **Goal:** Extract the underlying PDFs that determine the cross sections

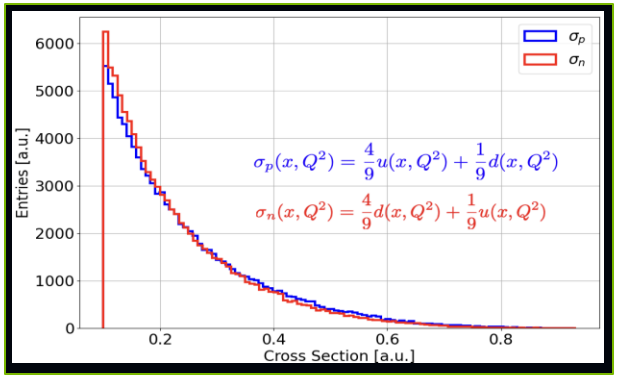




- **Approach:** Use workflow to find PDFs
- Truth, i.e. inputs, is known here
- Perform loop closure tests

THE 1D PROXY APP - ENVIRONMENT





Theory Module → Sampling Module → Experimental Module → Event Selection Module → Synthetic Data

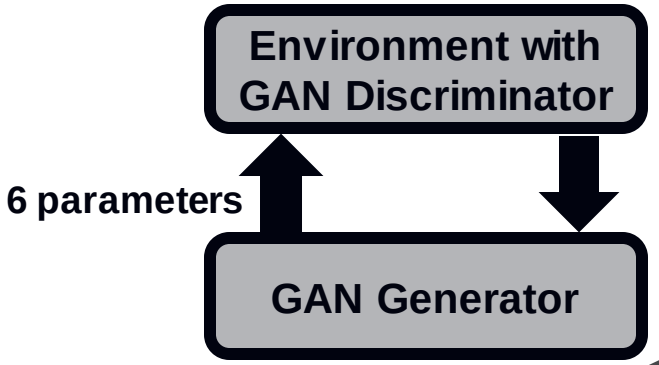
- 6 parameters define toy PDFs
- No quark-gluon interaction
- Omit Q-dependence
- Use 1D inverse CDF to sample cross sections from PDFs

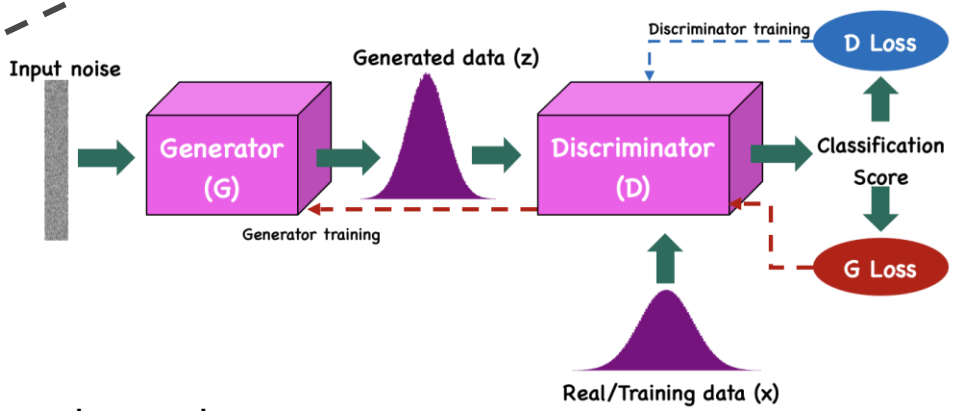
- Simplified detector
- Turn resolution effects on / off

$$D = \begin{pmatrix} \mathcal{N}(1,\sigma_0) & \epsilon_0 \\ \epsilon_1 & \mathcal{N}(1,\sigma_1) \end{pmatrix}$$

Turned off

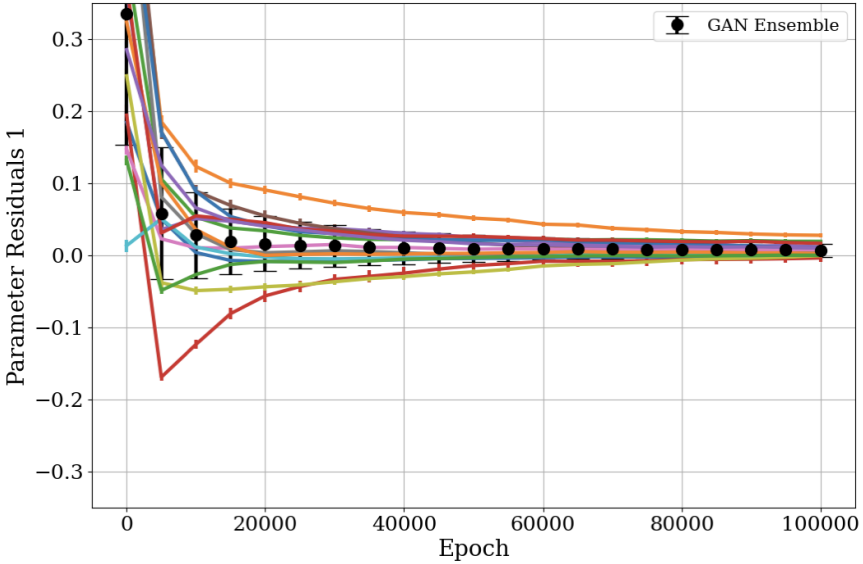
THE 1D PROXY APP - ENVIRONMENT



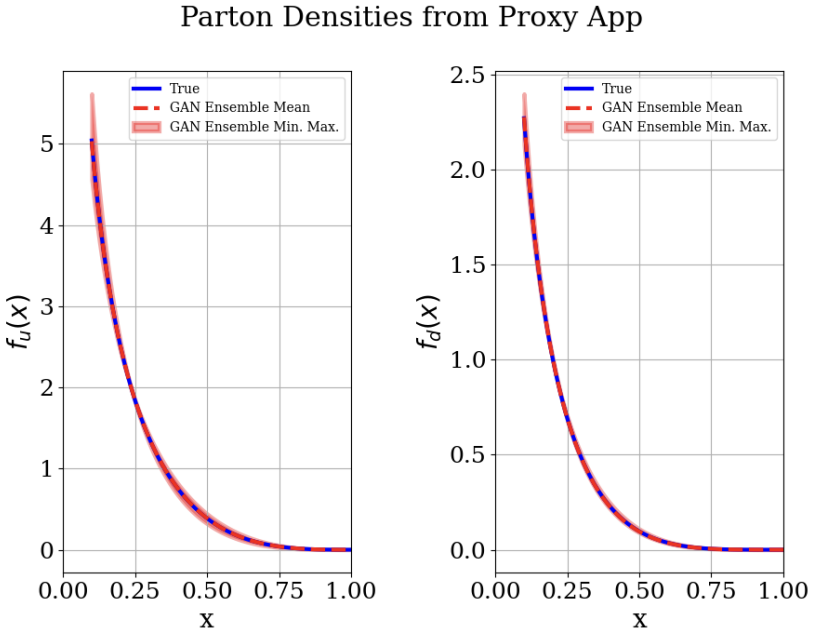


1. Generator predicts 6 parameters
2. Parameters are translated to synthetic events by environment
3. Discriminator (part of environment) is trained on synthetic and toy data
4. Use discriminator loss on synthetic data to update generator
5. Repeat steps 1 – 4 until convergence

ENSEMBLE ANALYSIS ON 1D PROXY APP

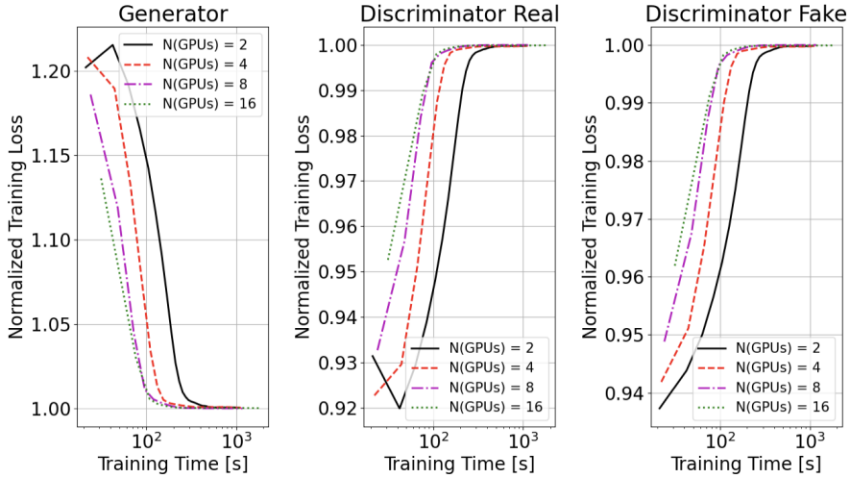


- Used ensemble with 15 GANs
- Benefit from individual parameter initialization
- Ensemble converges earlier to expected solution than individual GAN



- Ran loop closure tests with / without resolution effects
- Reproduced input PDFs
- Need to include uncertainty quantification

FIRST SCALING STUDIES WITH 1D PROXY APP

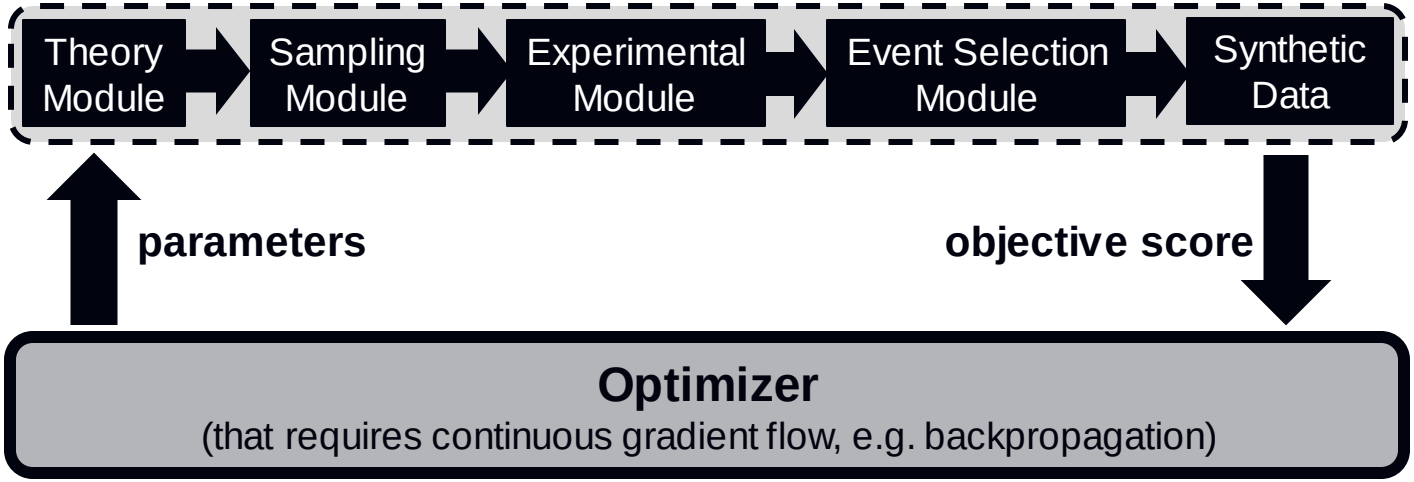


- **Goal:** Run workflow on HPC systems ==> Scaling
- Utilized horovod for distributed learning
- Ran proxy analysis with single GAN on Polaris @ Argonne National Lab
- First results indicate poor scaling behavior
- Identified several bottle necks
- Overhauled workflow ==> Need to repeat scaling studies

MANAGING THE GRADIENT FLOW

Forward pass

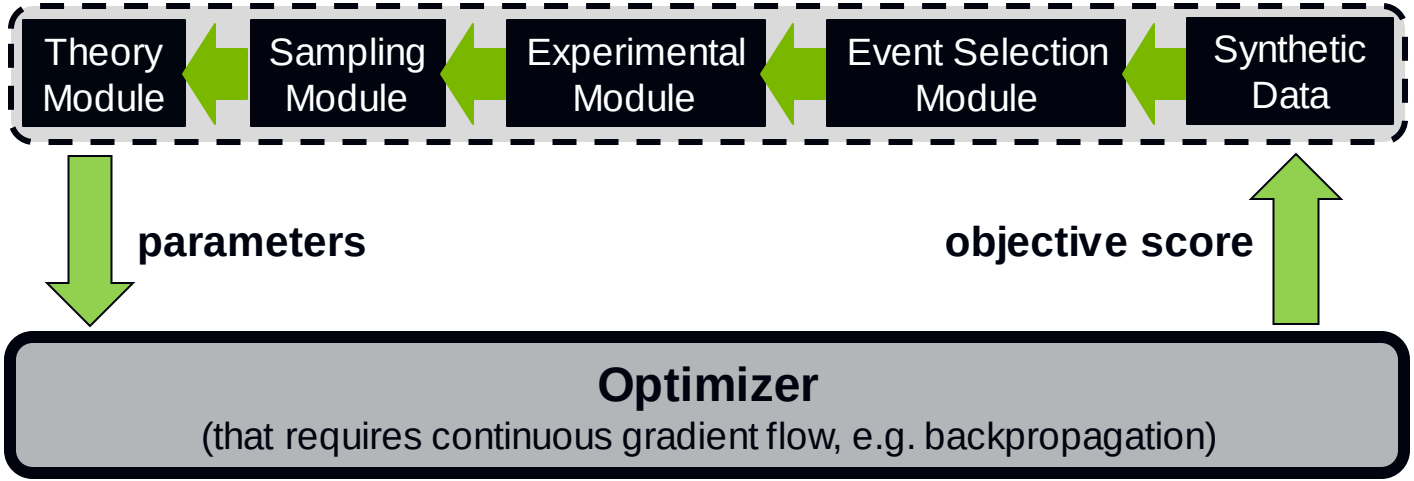
- Optimizer predicts parameters
- Parameters are translated to synthetic data
- Synthetic data defines, together with experimental data, an objective score



MANAGING THE GRADIENT FLOW

Backward pass

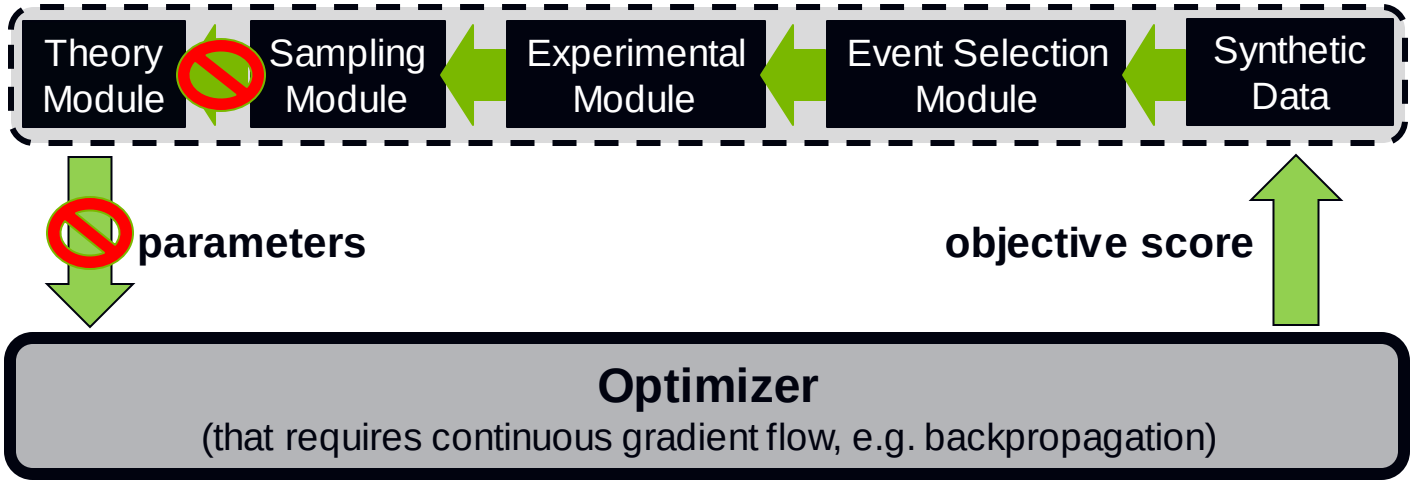
- Rely on chain rule to update optimizer state
- Propagate gradients back through entire pipeline
- Every module needs to be differentiable



MANAGING THE GRADIENT FLOW

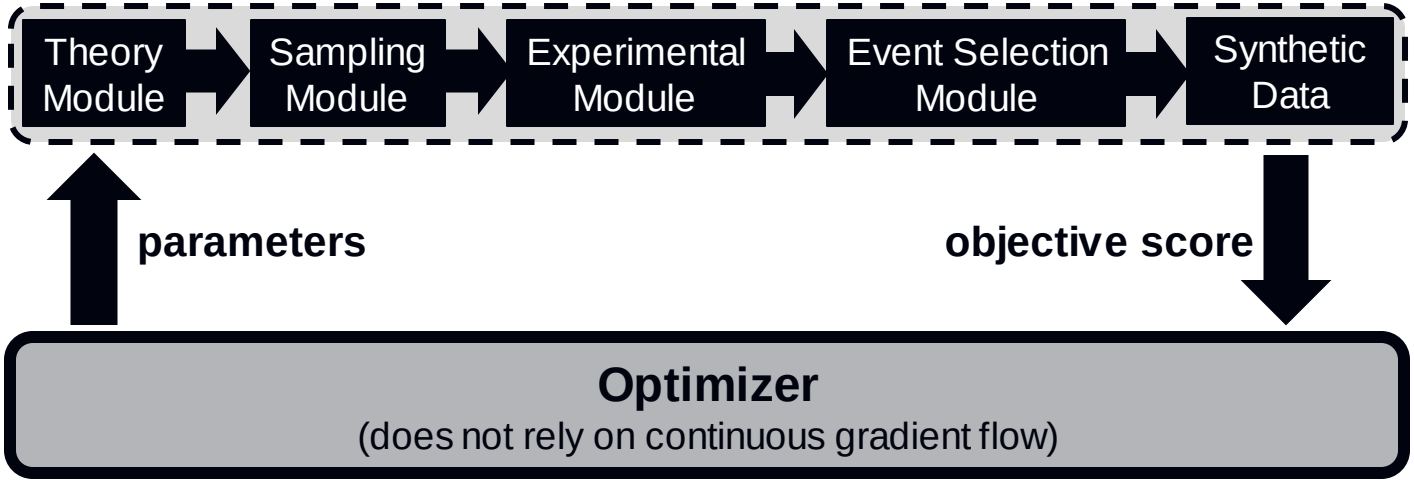
Trouble

- Rely on chain rule to update optimizer state
- Gradient flow is disturbed
- At least one module is not differentiable (e.g. sampler)

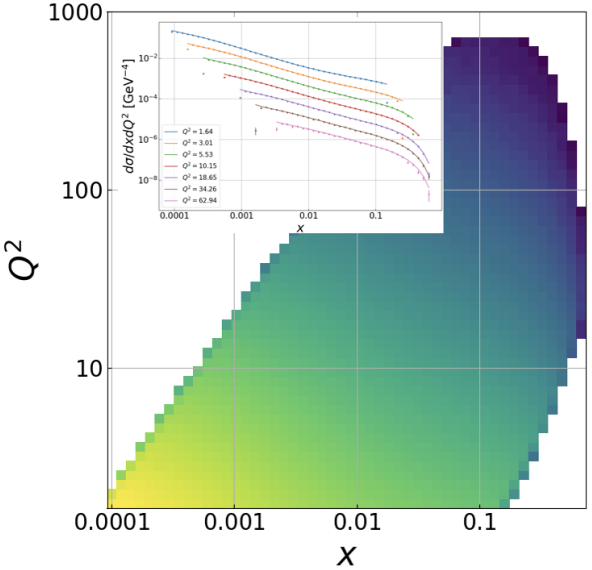


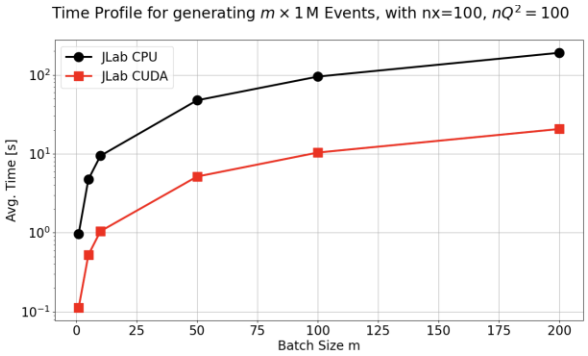
ALTERNATIVE: AVOID GRADIENT FLOW

- Do not care about differentiability
- Minimize / Maximize objective score w.r.t predicted parameters
- Currently exploring: Reinforcement Learning, Genetic Algorithms, Simulated Annealing,...

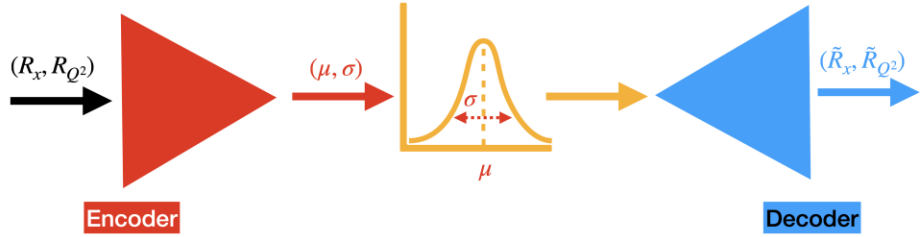


TOWARDS A 2D PROXY APP AND DIS ANALYSIS





- DIS Theory following Duke & Owens
- **L**ocal **O**rthogonal **I**nverse **T**ransform **S**ampler (LOITS), developed by Nobuo Sato
- Generate DIS events in x and Q
- LOITS is differentiable
- Use VAE as a detector surrogate ==> Model residuals in x and Q
- Enable analysis of real measured DIS data
- Workflow already set up for 2D proxy analysis
- Currently in testing and debugging phase



Loss $\sim \|(R_x, R_{Q^2}) - (\tilde{R}_x, \tilde{R}_{Q^2})\|$ + **KL-Divergence**

$$R_i \equiv \frac{i - i'}{i}, i = x, Q^2$$

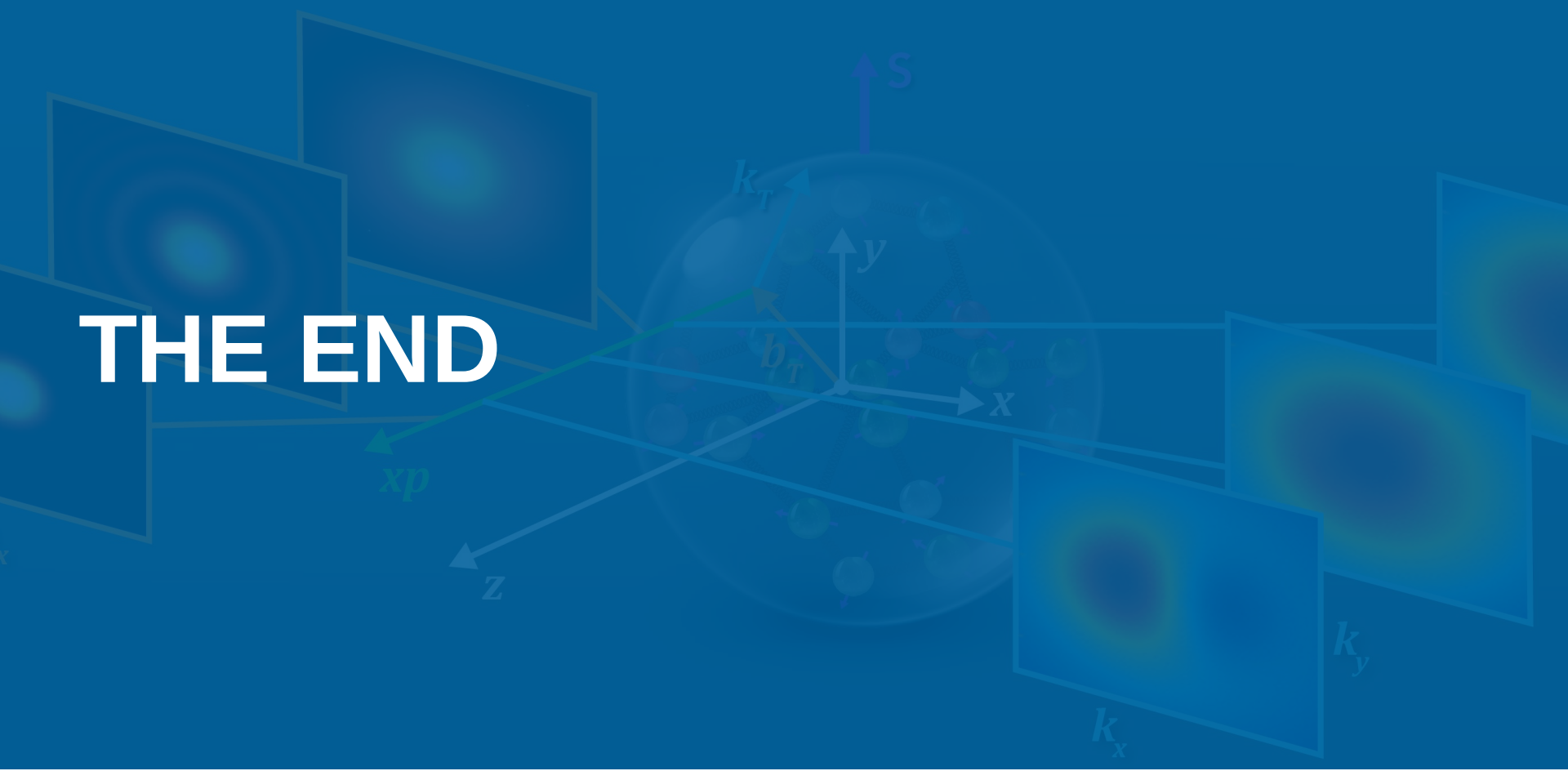
QuantOm Collaboration

SUMMARY AND OUTLOOK

- Composable event level workflow for inverse problem solving
- Results and lessons learned from 1D proxy GAN analysis
  - Successful loop-closure tests
  - Enable faster convergence with ensemble analysis
  - Studied workflow behavior on Polaris @ Argonne
  - Work on scalability for HPC machines
  - Formulate proper uncertainty quantification (UQ)
  - Define proper convergence metric (truth is unknown in general)
- Explore non-gradient based optimizers
  - Can not always guarantee differentiability
  - Certain methods (e.g. RL) support asynchronous learning
- 2D proxy analysis in the pipeline
  - Close to "real" DIS measurements
  - Prepare workflow for measured data analysis
- Extend experimental module ==> Include effects other than detector resolution / acceptance (e.g. background)
- Deploy workflow to HPC machines, e.g. Aurora, Sunspot,...

JOIN OUR TEAM !

- Jefferson Lab data science department is expanding
- Various projects with / without connection to nuclear physics
  - QuantOM SciDAC (todays' talk)
  - Controls in accelerator science (BNL, CEBAF, industry,…)
  - Support of experimental halls at Jefferson Lab
  - Detector read out and optimization
  - Environmental science and health
- Two positions available
  1. Temporary staff position with possibility to permanent **https://misportal.jlab.org/hr/recruiting/postings/13251**
  2. Two postdoc positions (one for QuantOM SciDAC) **https://misportal.jlab.org/hr/recruiting/postings/13231**
- For further questions / details please contact: Malachi Schram (schram@jlab.org)



THE END