QUARK-GLUON TOMOGRAPHY COLLABORATION

QuantOm Collaboration

Progress towards a machine learning extraction of GPDs from data

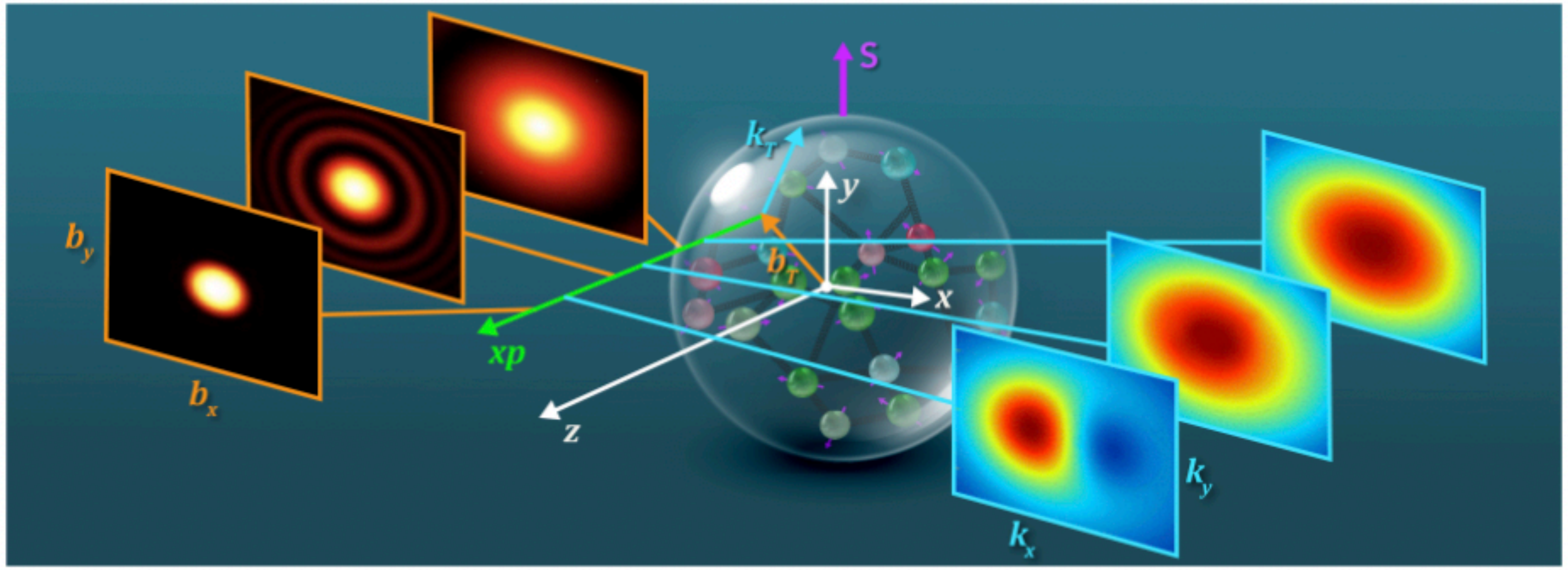
Eric Moffat

Argonne National Lab

GHP 3-16-24

Introduction

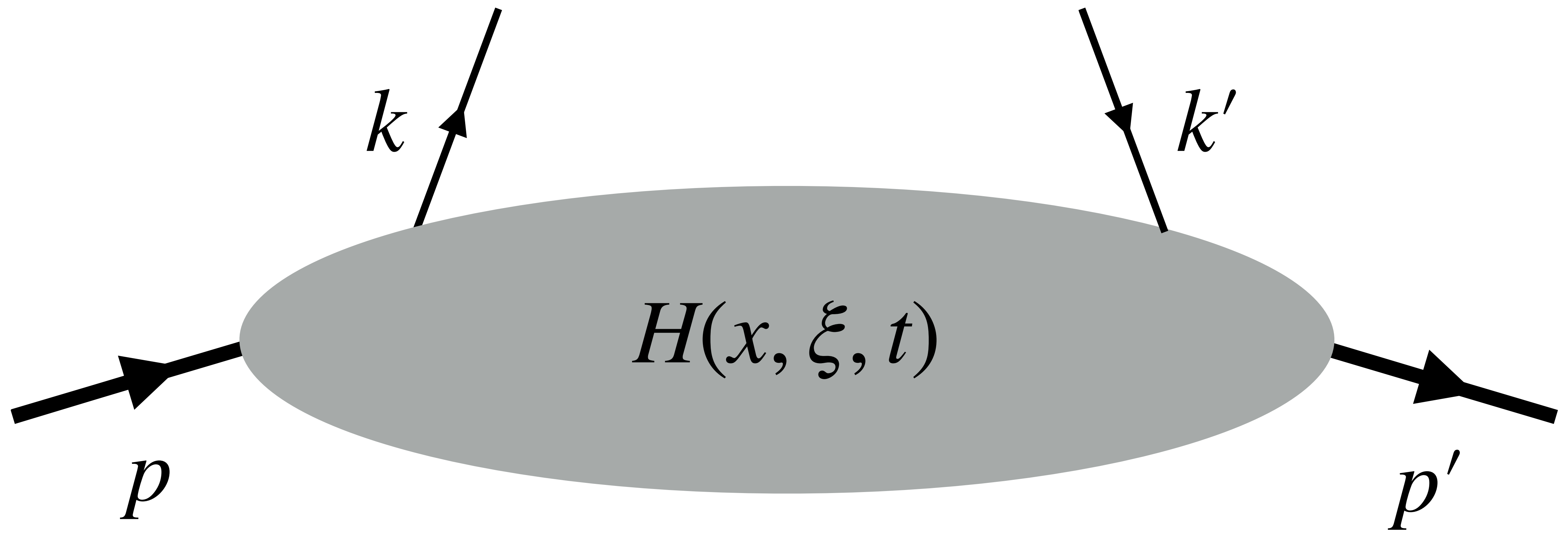
- Generalized Parton Distributions (GPDs) contain information about many hadron properties:
  - 3D structure
  - Spin sum
  - Pressure and shear force distributions
- The goal:
  - Perform a global analysis of Compton Form Factors (CFFs) and GPDs from available data



Phenomenological Challenges

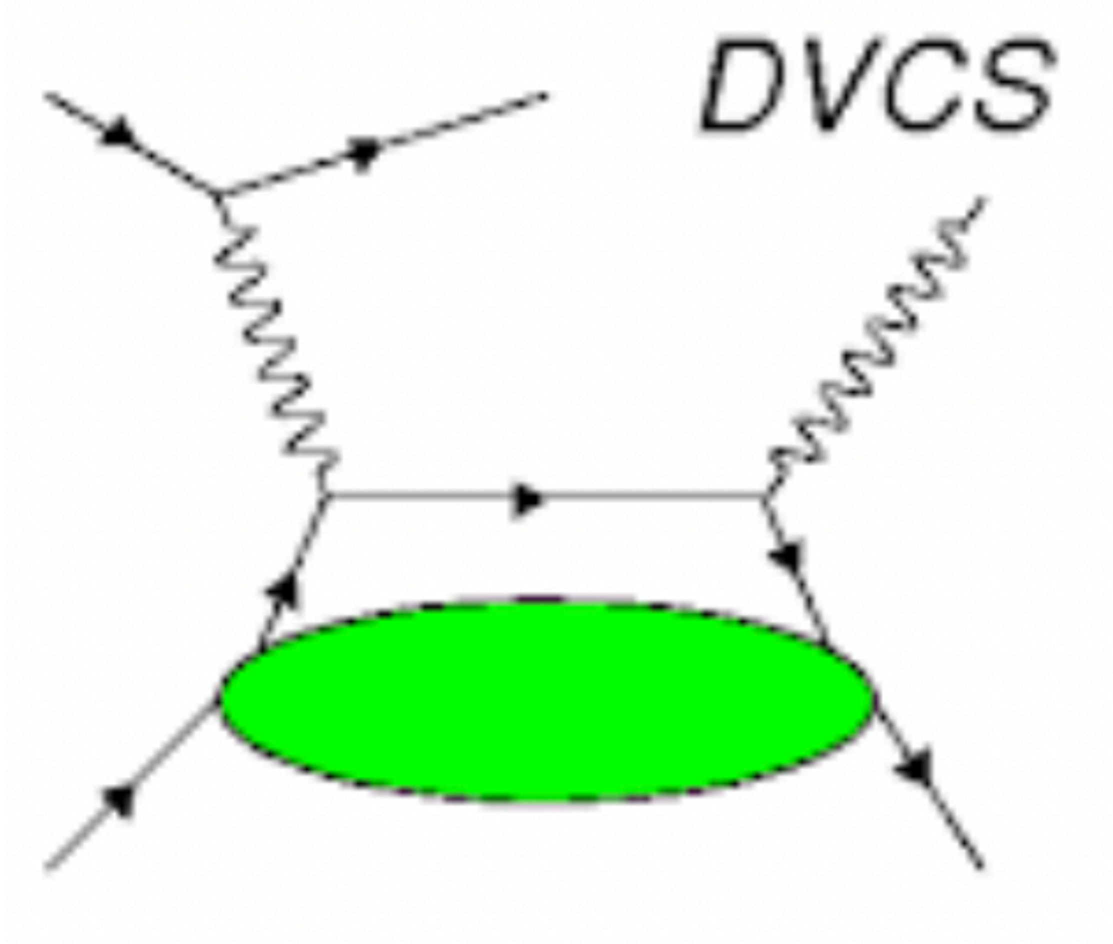
- Functions of x, ξ, and t:

$$x = \frac{k^+ + k'^+}{p^+ + p'^+} \quad \xi = \frac{p'^+ - p^+}{p^+ + p'^+} \quad t = (p' - p)^2$$



- Inverse Problem:
  - Shadow GPDs (SGPDS) (Bertone, et. al. Phys.Rev.D 103 (2021) 11, 114019):
    - There is an infinite number of functions that can give the same observable.

The Inverse Problem

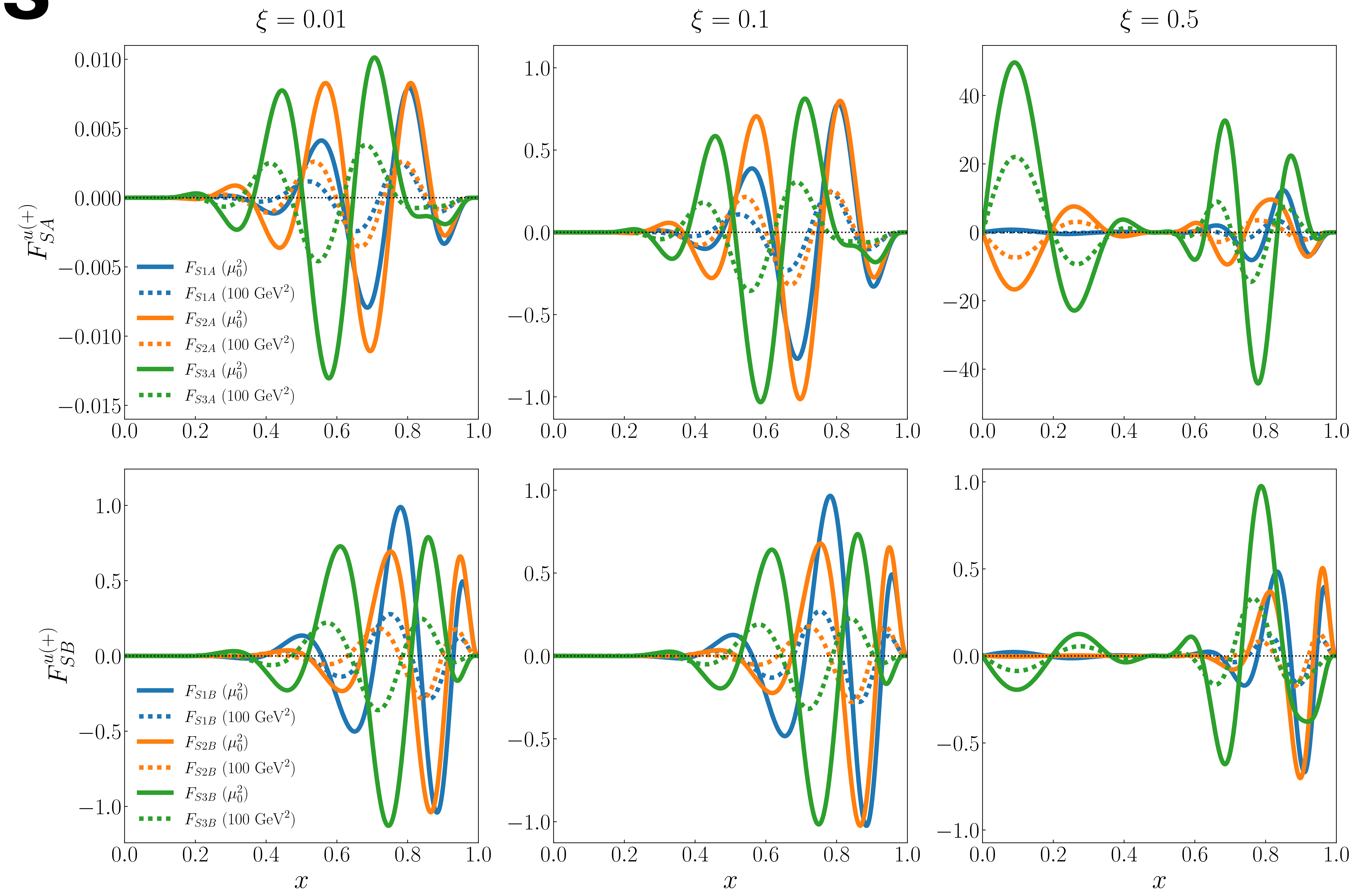


- Deeply virtual Compton scattering:
  - Compton Form Factors ($\mathscr{H}, \mathscr{E}, \tilde{\mathscr{H}}, \tilde{\mathscr{E}}$):

$$\mathscr{H}(\xi, t, Q^2) = \int_{-1}^{1} dx \sum_a C^a(x, \xi, Q^2, \mu^2)\, H^a(x, \xi, t; \mu^2)$$

- x-dependence is lost in the integration
- There is an infinite number of functions that can give the same CFF.
  - While a fit could obtain a GPD: Does the x-dependence represent the true GPD?

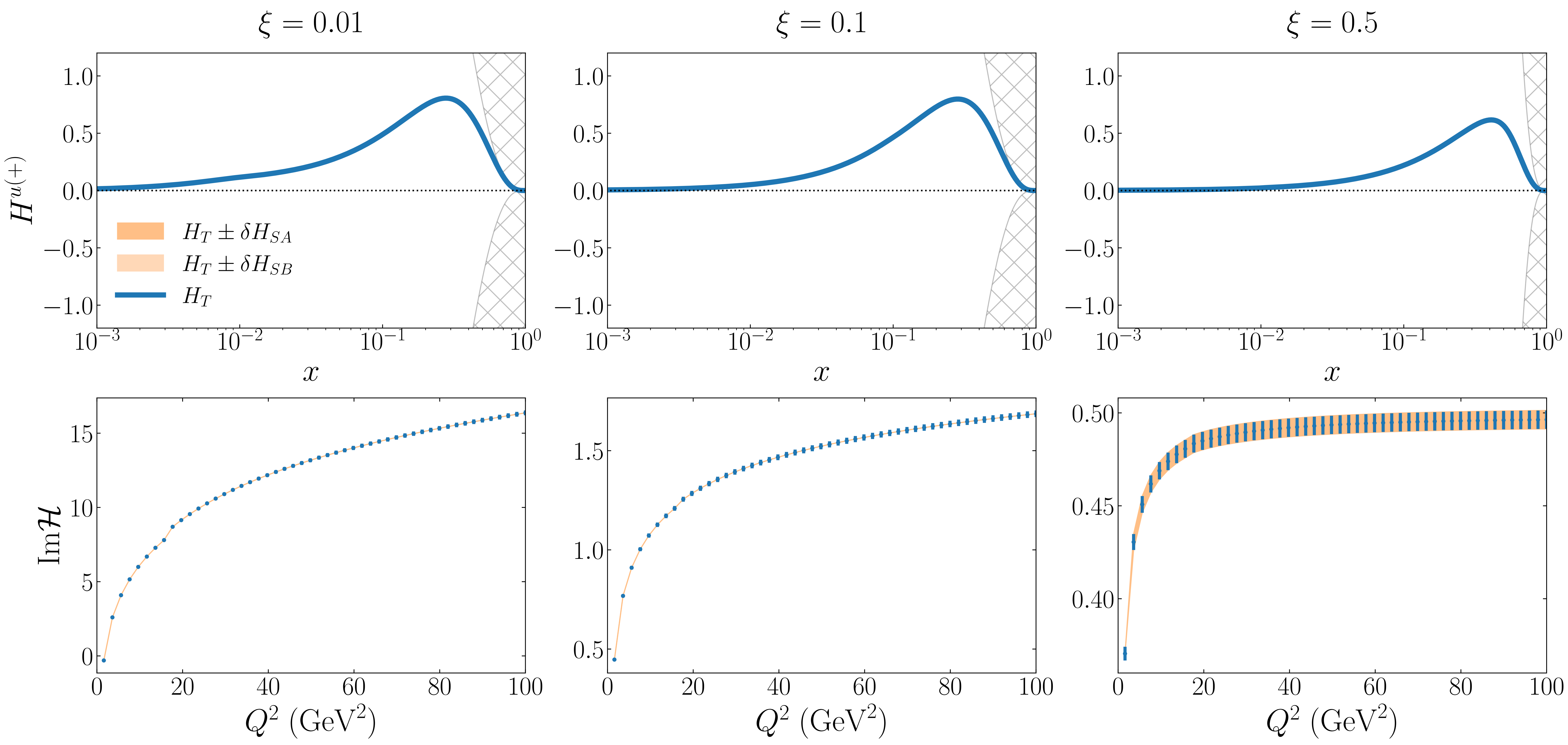
SGPDs



E. M. et al, Phys.Rev.D 108 (2023) 3, 036027

SGPDs

- Evidence that evolution can help constrain at least some SGPDs but unknown if this is true in general.

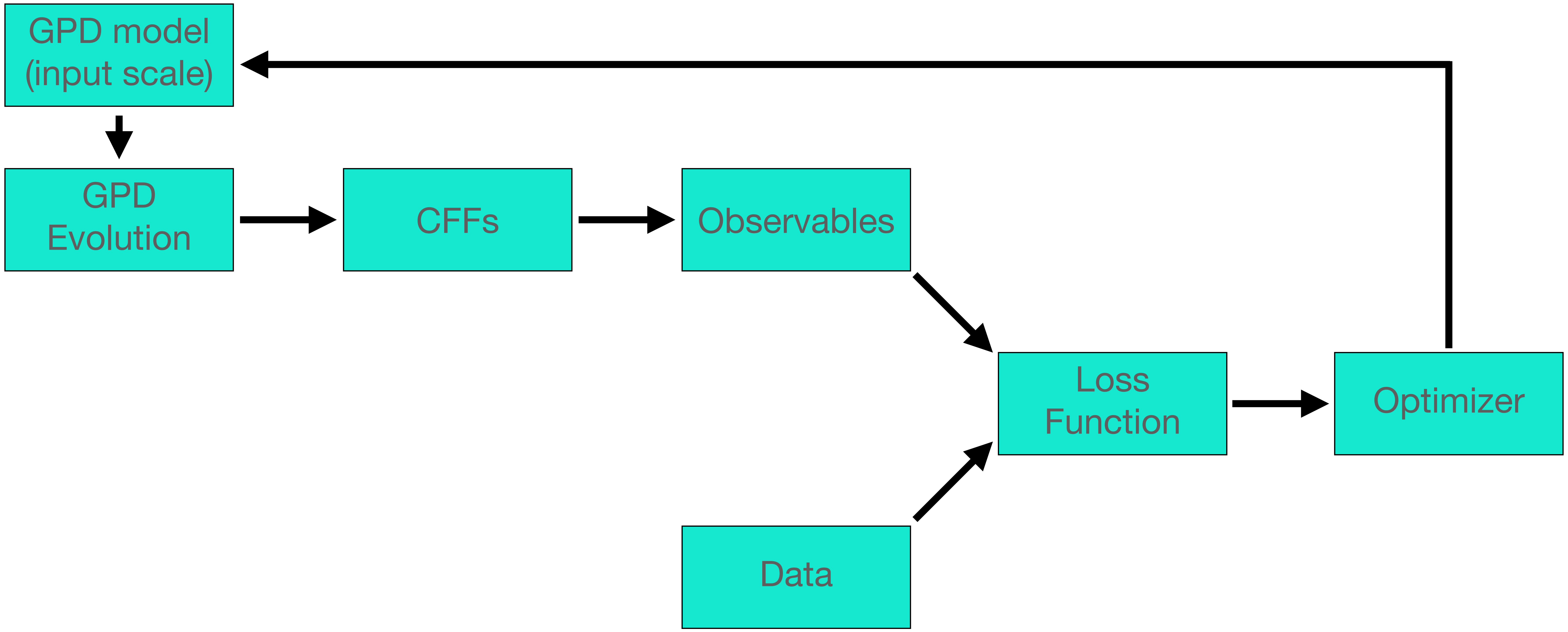


E. M. et al, Phys.Rev.D 108 (2023) 3, 036027

SGPDs

- Parametric models lack the flexibility to thoroughly sample SGPDs
    - Different parametric models can fit the data equally well but could yield significantly different results
- Need a highly flexible model to accurately account for uncertainties while minimizing bias:
    - Use Neural Networks (NNs)
- Developed a machine learning framework for GPD extraction
    - Utilizing Bayesian Monte Carlo approach consistent with methods used by the Jefferson Lab Angular Momentum (JAM) Collaboration

The machinery



- All pieces are backward differentiable to facilitate machine learning

The machinery

- GPD model:
  - Utilize double distributions to guarantee polynomiality

$$H^f(x,\xi,t;\mu_0^2) = \int_\Omega d\beta d\alpha\, \delta(x-\beta-\xi\alpha)\left[H^f_{DD}(\beta,\alpha,t;\mu_0^2) + \xi\delta(\beta)D^f(\alpha,t;\mu_0^2)\right]$$

$$E^f(x,\xi,t;\mu_0^2) = \int_\Omega d\beta d\alpha\, \delta(x-\beta-\xi\alpha)\left[E^f_{DD}(\beta,\alpha,t;\mu_0^2) - \xi\delta(\beta)D^f(\alpha,t;\mu_0^2)\right]$$

$$\tilde{H}^f(x,\xi,t;\mu_0^2) = \int_\Omega d\beta d\alpha\, \delta(x-\beta-\xi\alpha)\left[\tilde{H}^f_{DD}(\beta,\alpha,t;\mu_0^2)\right]$$

$$\tilde{E}^f(x,\xi,t;\mu_0^2) = \int_\Omega d\beta d\alpha\, \delta(x-\beta-\xi\alpha)\left[\tilde{E}^f_{DD}(\beta,\alpha,t;\mu_0^2)\right]$$

  - For H and $\tilde{H}$, use existing parton distribution functions for the forward limit

- Loss function:
  - Typical chi squared function

$$\sum \left(\frac{\text{data} - \text{theory}}{\text{uncertainty}}\right)^2$$

- Optimizer:
  - Use PyTorch Adam algorithm
    - Stochastic Gradient Descent

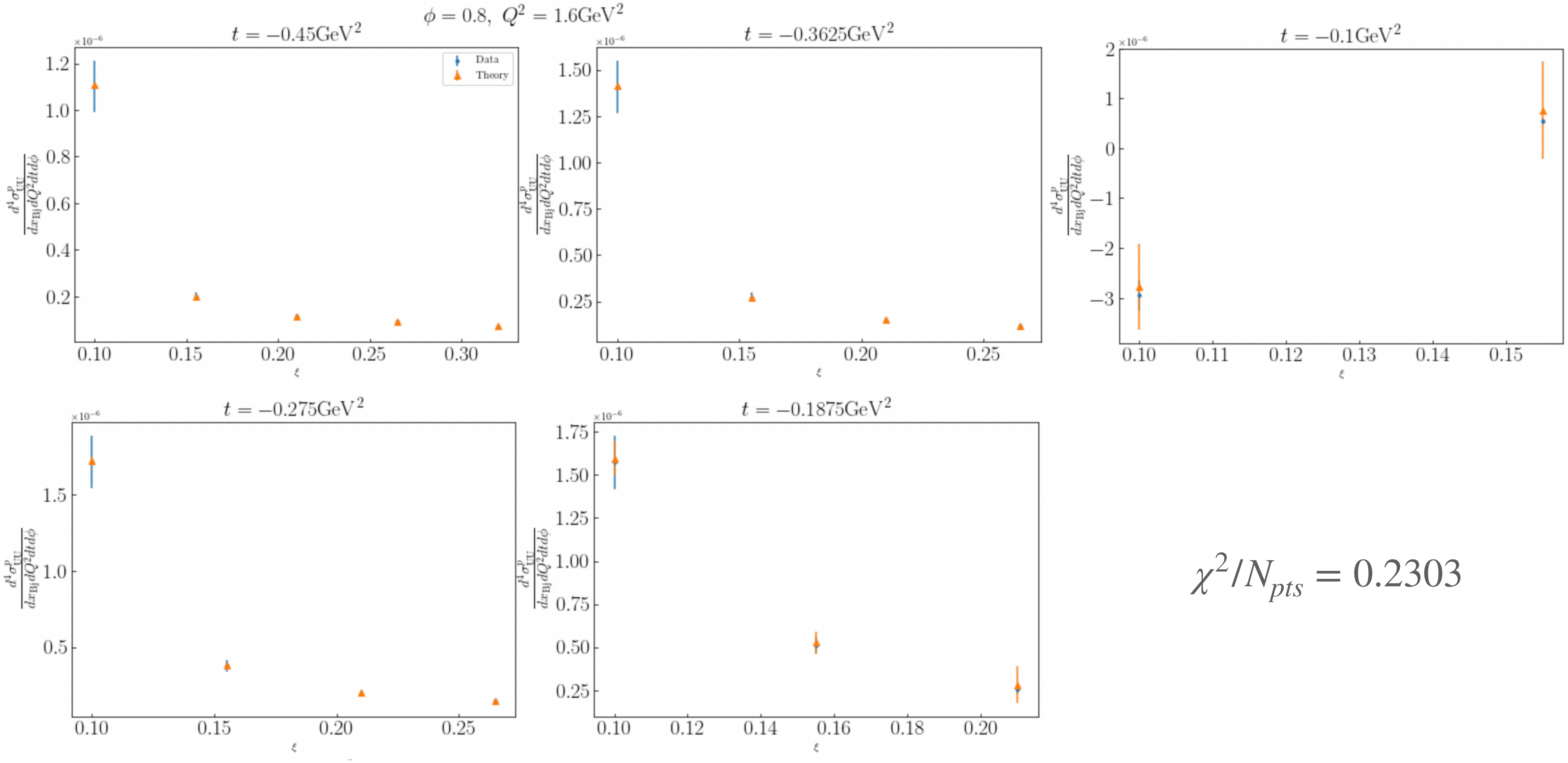
Closure test

- Generated pseudodata for various DVCS observables from model GPDs:
  - GPD model:
    - Double distributions:
      - Use GK model (Kroll, Moutarde, Sabatie, Eur. Phys. J. C (2013) 73:2278)
    - D term:
      - Use first three terms of a Gegenbauer series (Goeke, Polyakov, Vanderhaeghan, Prog. Part. Nucl. Phys. 47, 401 (2001)
  - Assume 10% uncertainty for all data points
- Fitted parameters (31 in total):
  - Fit uv and dv double distribution parameters:
    - For H and $\tilde{H}$, keep pdf parameters fixed and only fit profile function parameters
  - Fit the coefficients and the t dependence parameters in the D term for u and d

Closure test

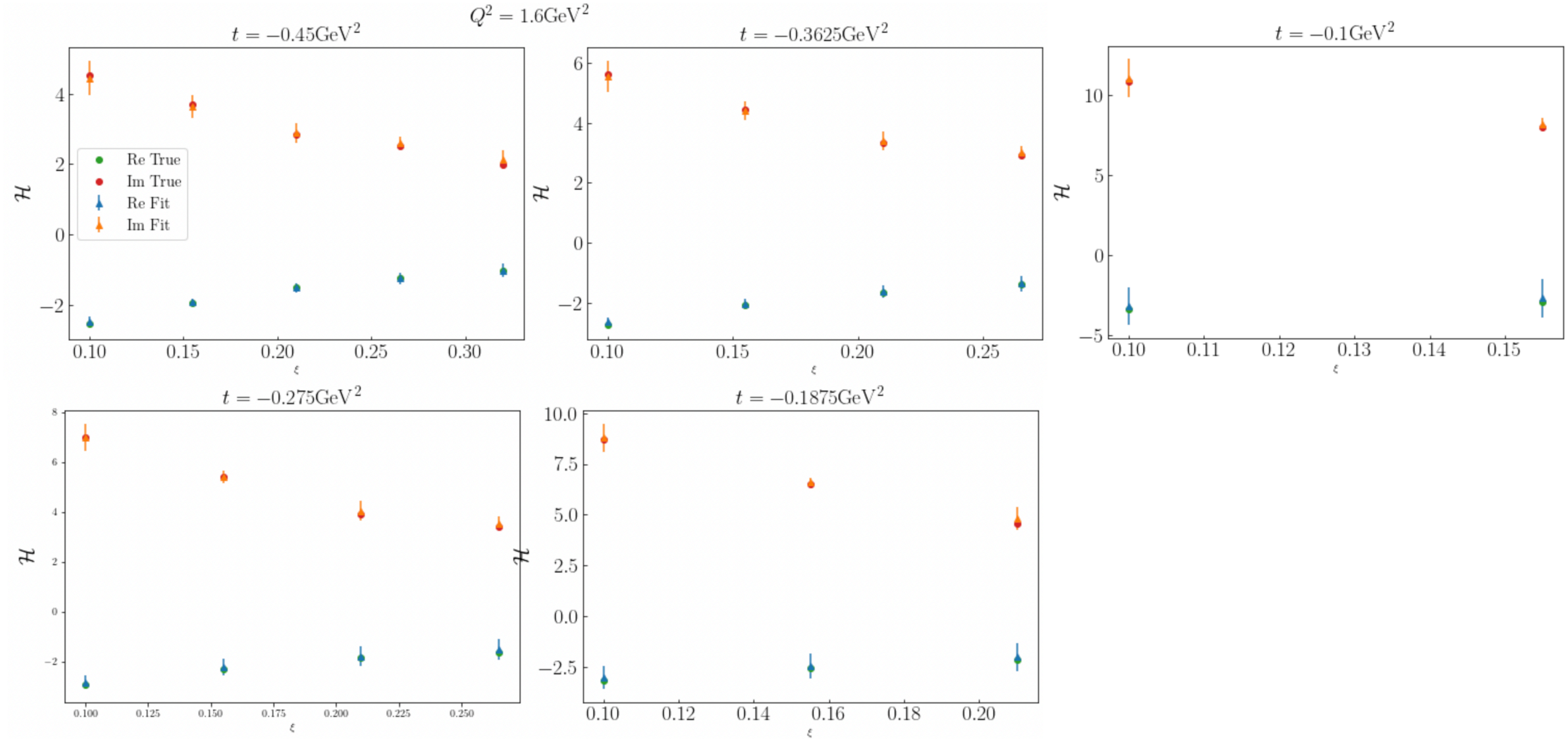
- Monte Carlo fit:
  - Conduct multiple fits (called replicas):
    - For each replica:
      - Starting parameters are randomly sampled
      - Data values sampled from Gaussian distribution
  - Calculate average and standard deviation of all replicas

Closure test results

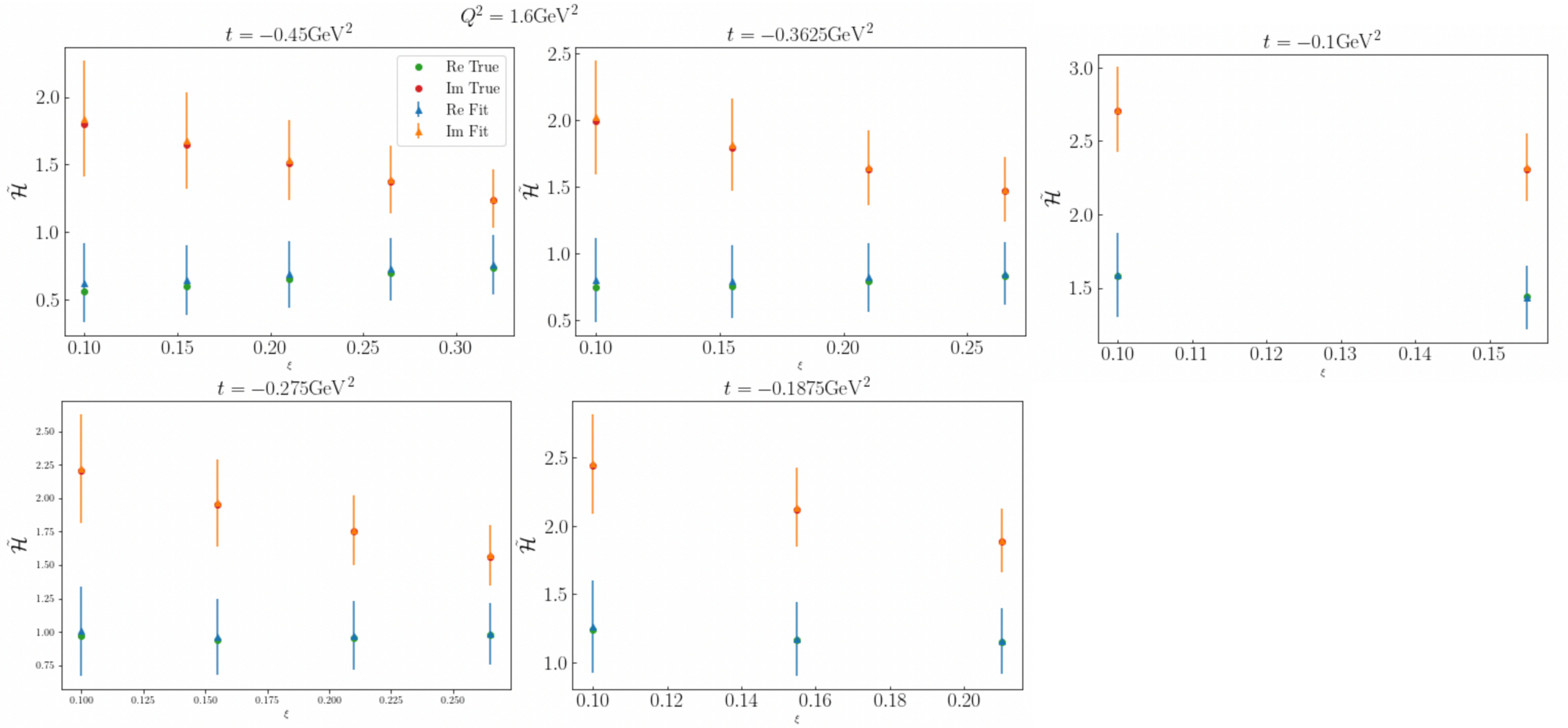


$$\chi^2/N_{pts} = 0.2303$$

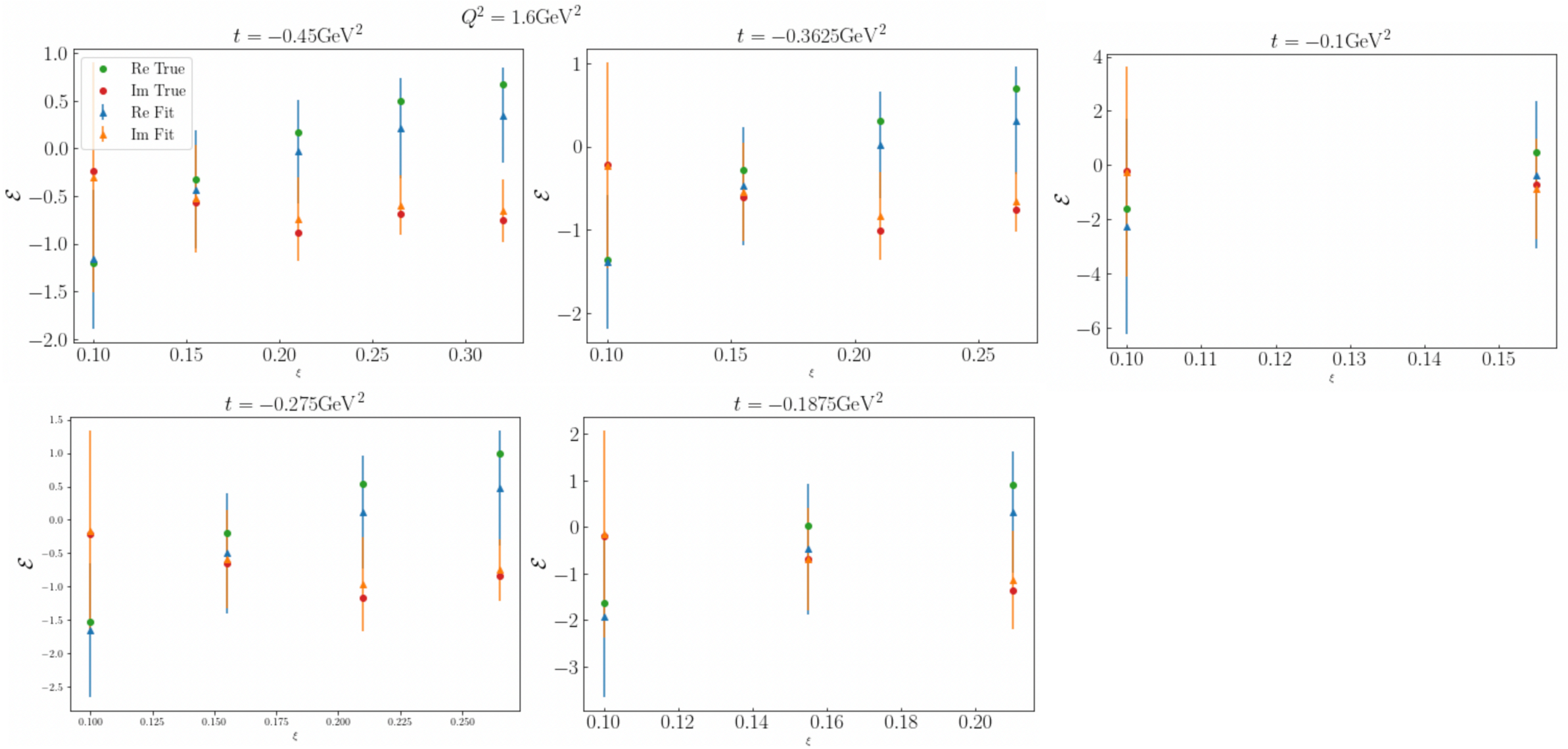
Closure test results



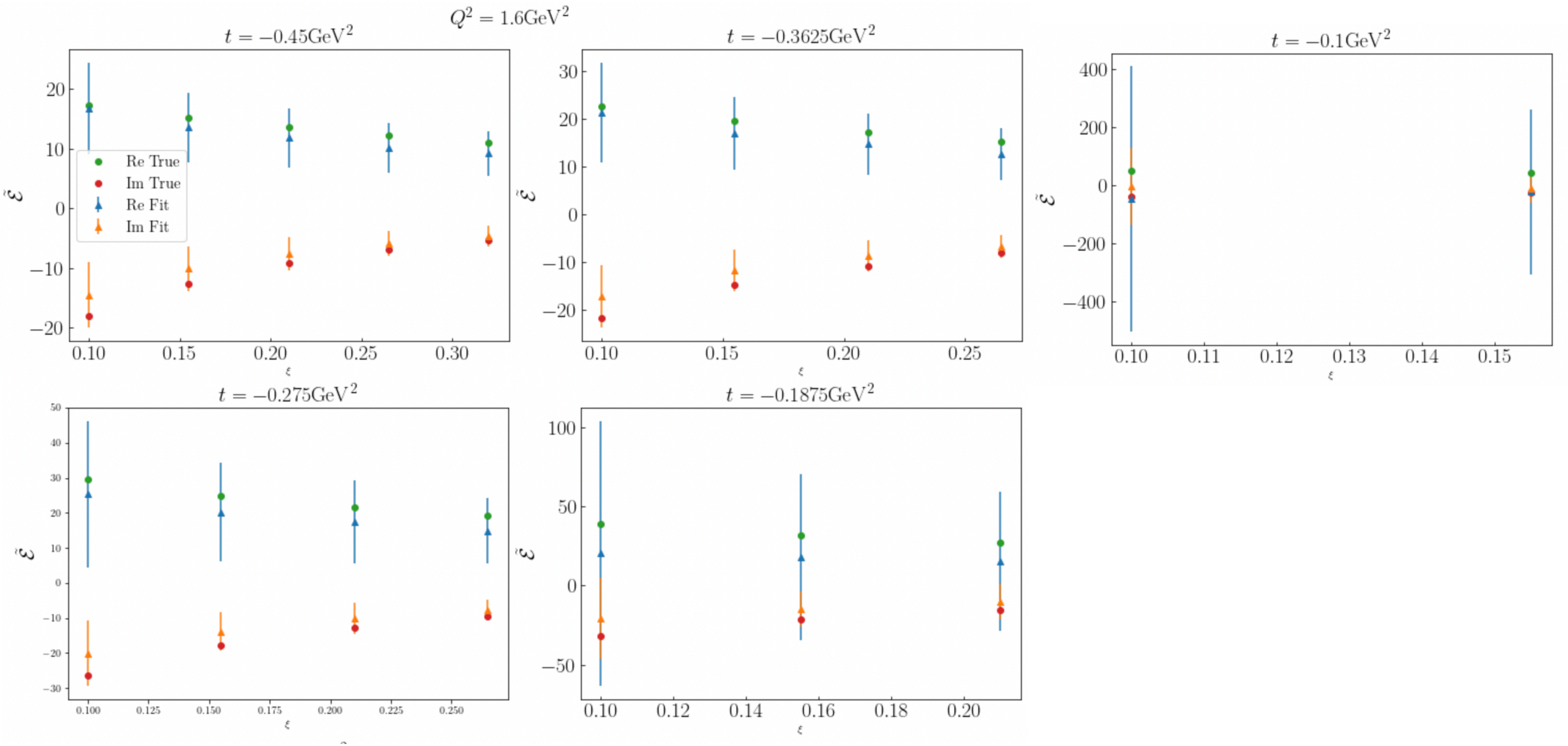
Closure test results



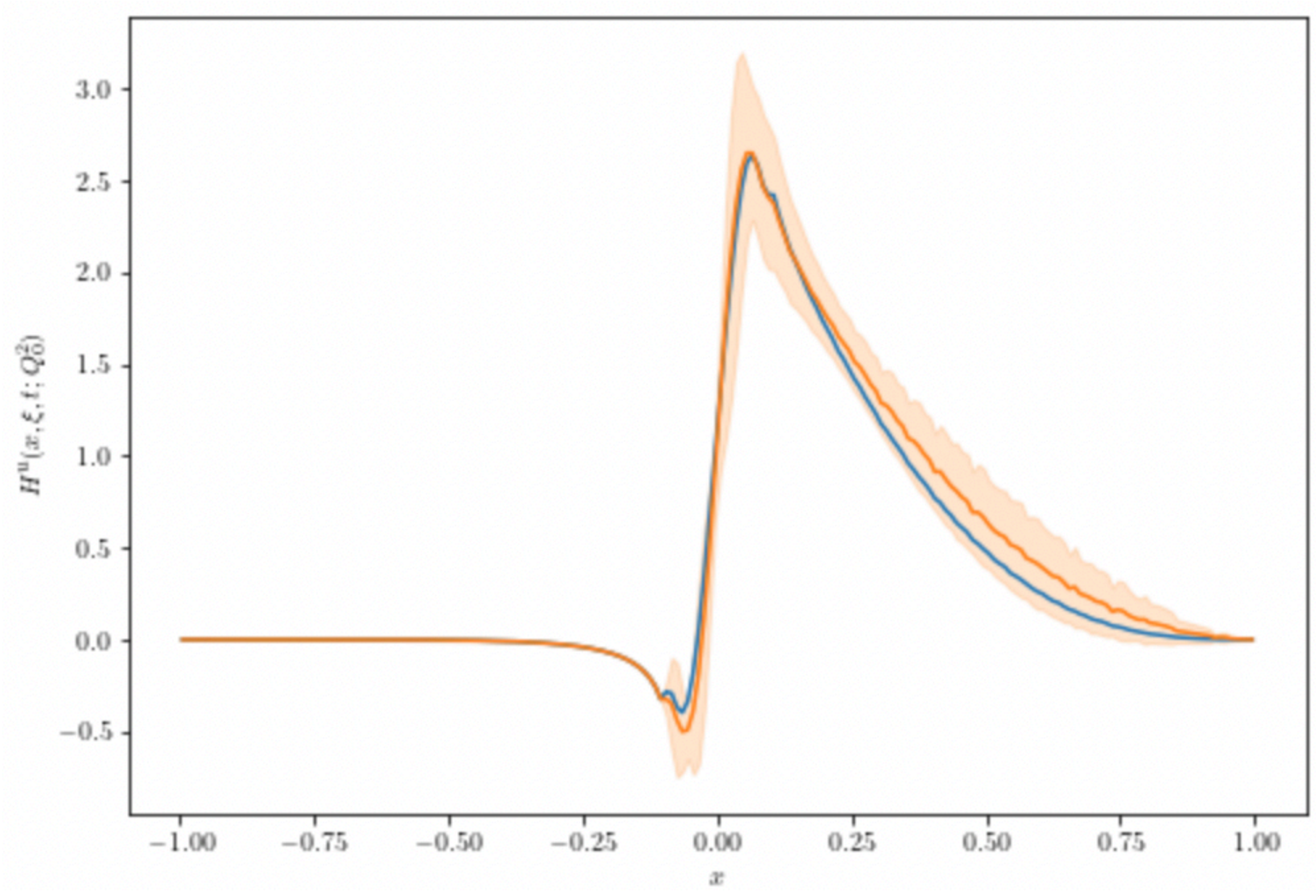
Closure test results

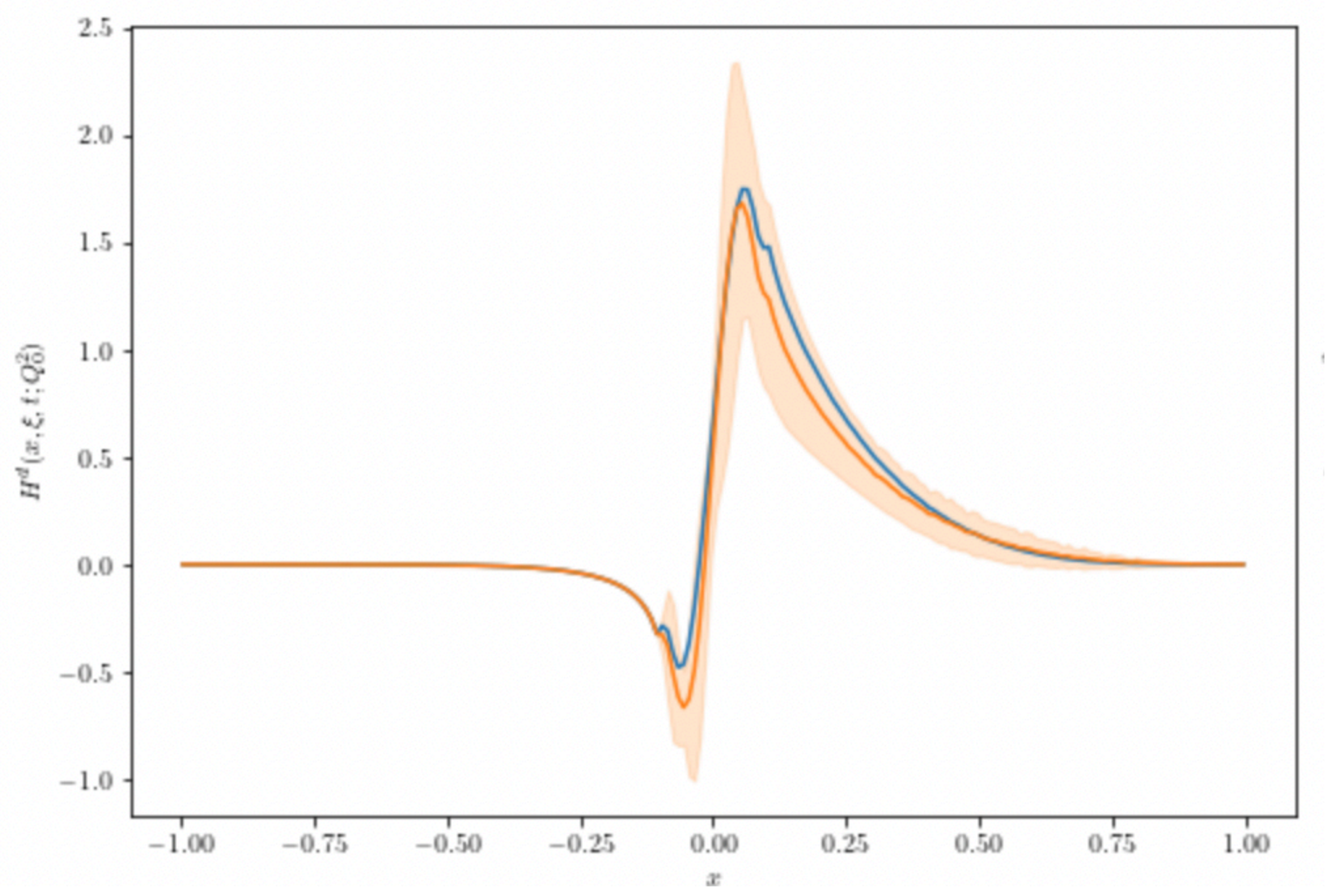


Closure test results

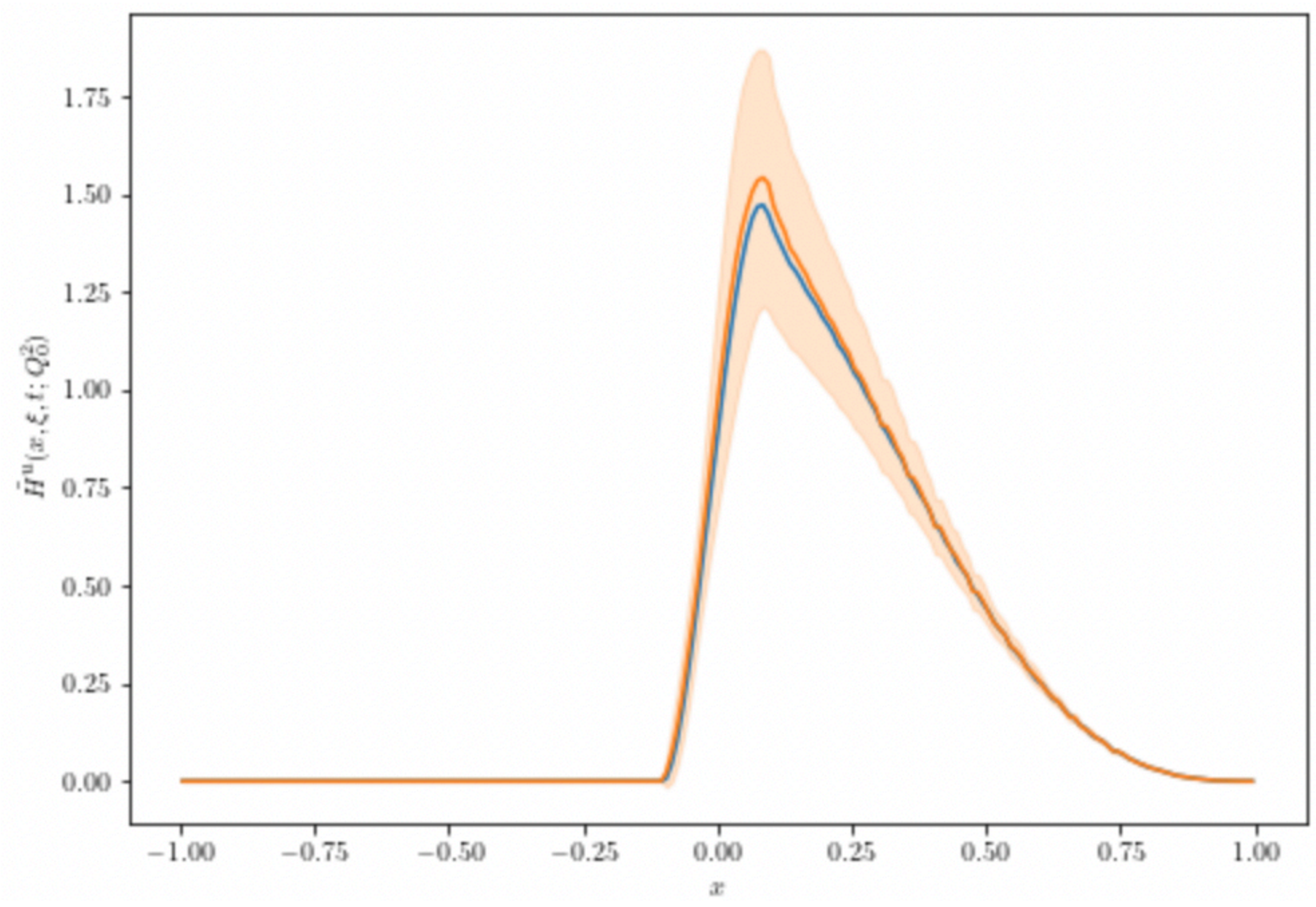


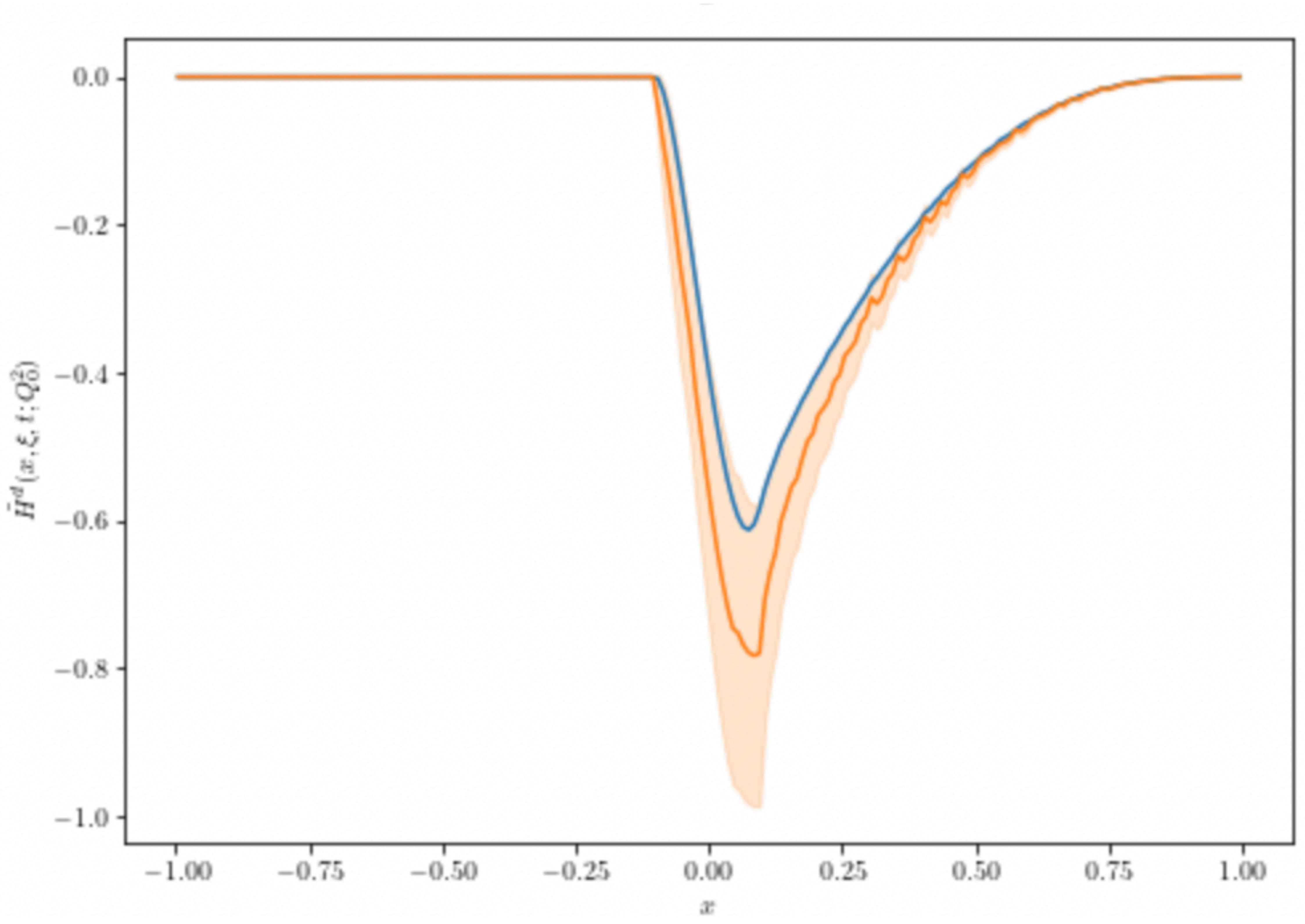
Closure test results



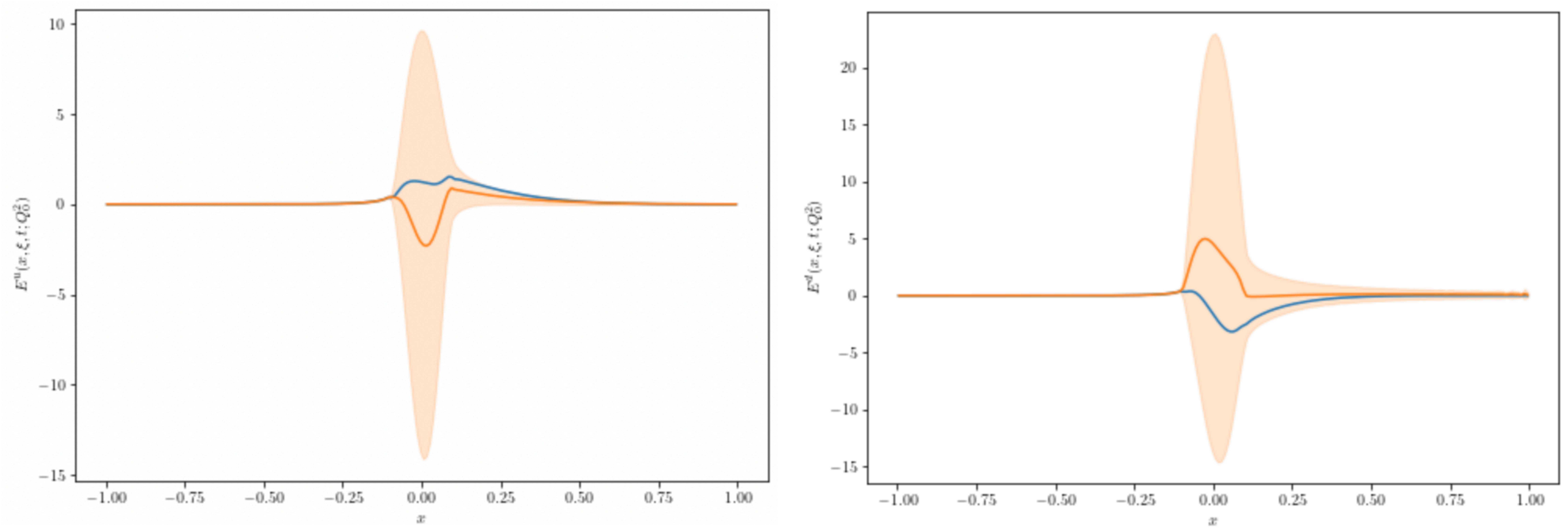


Closure test results

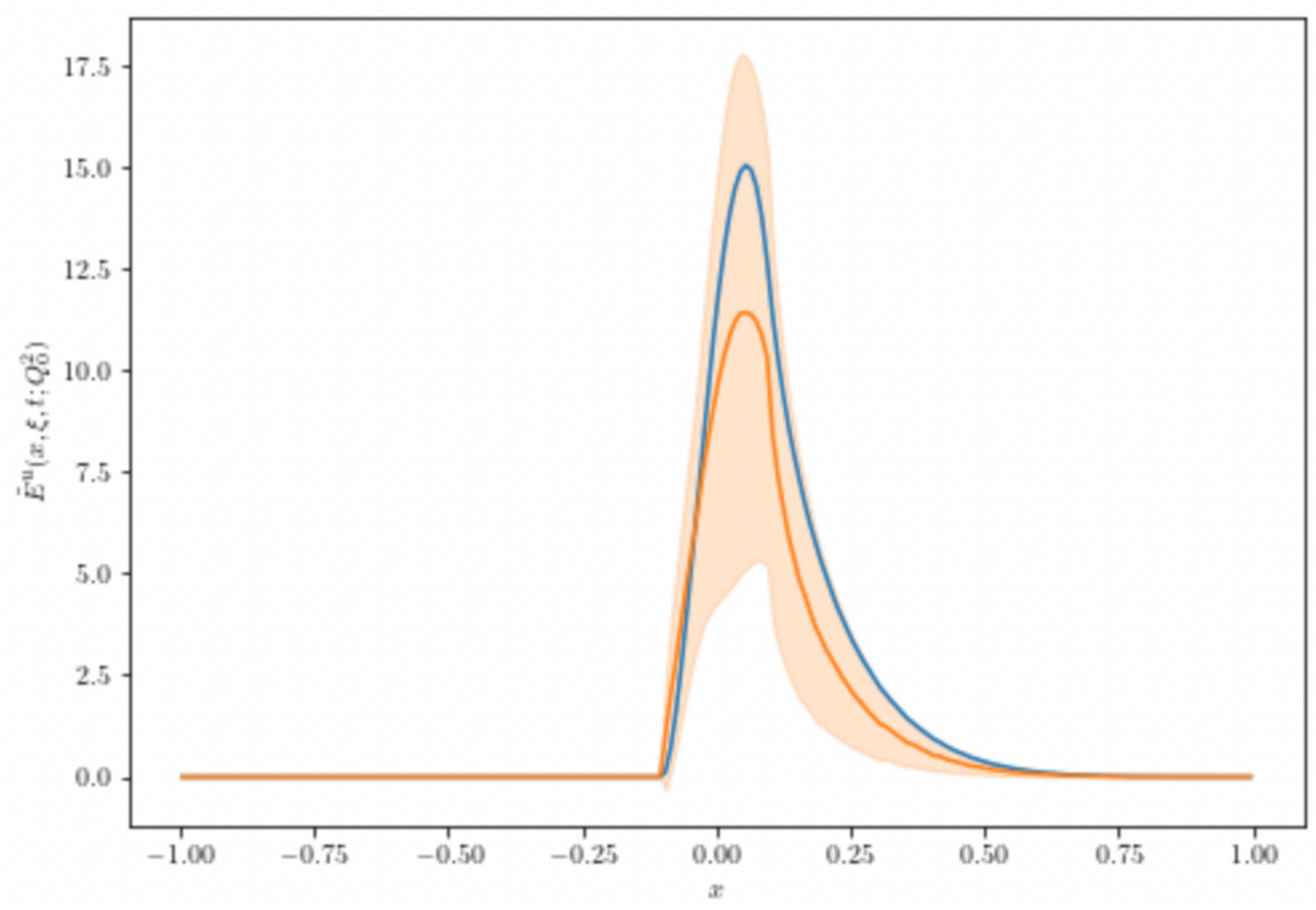


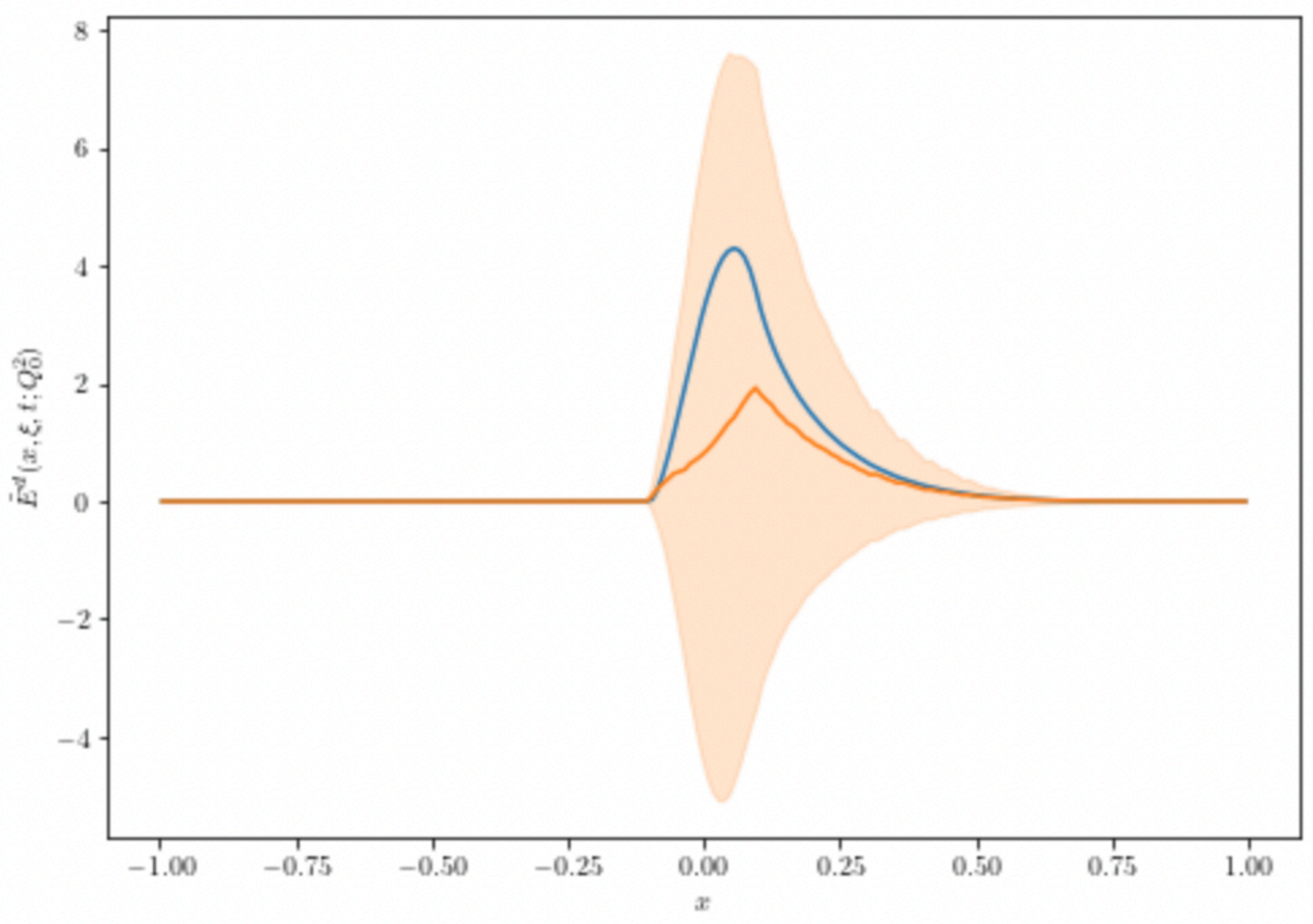


Closure test results



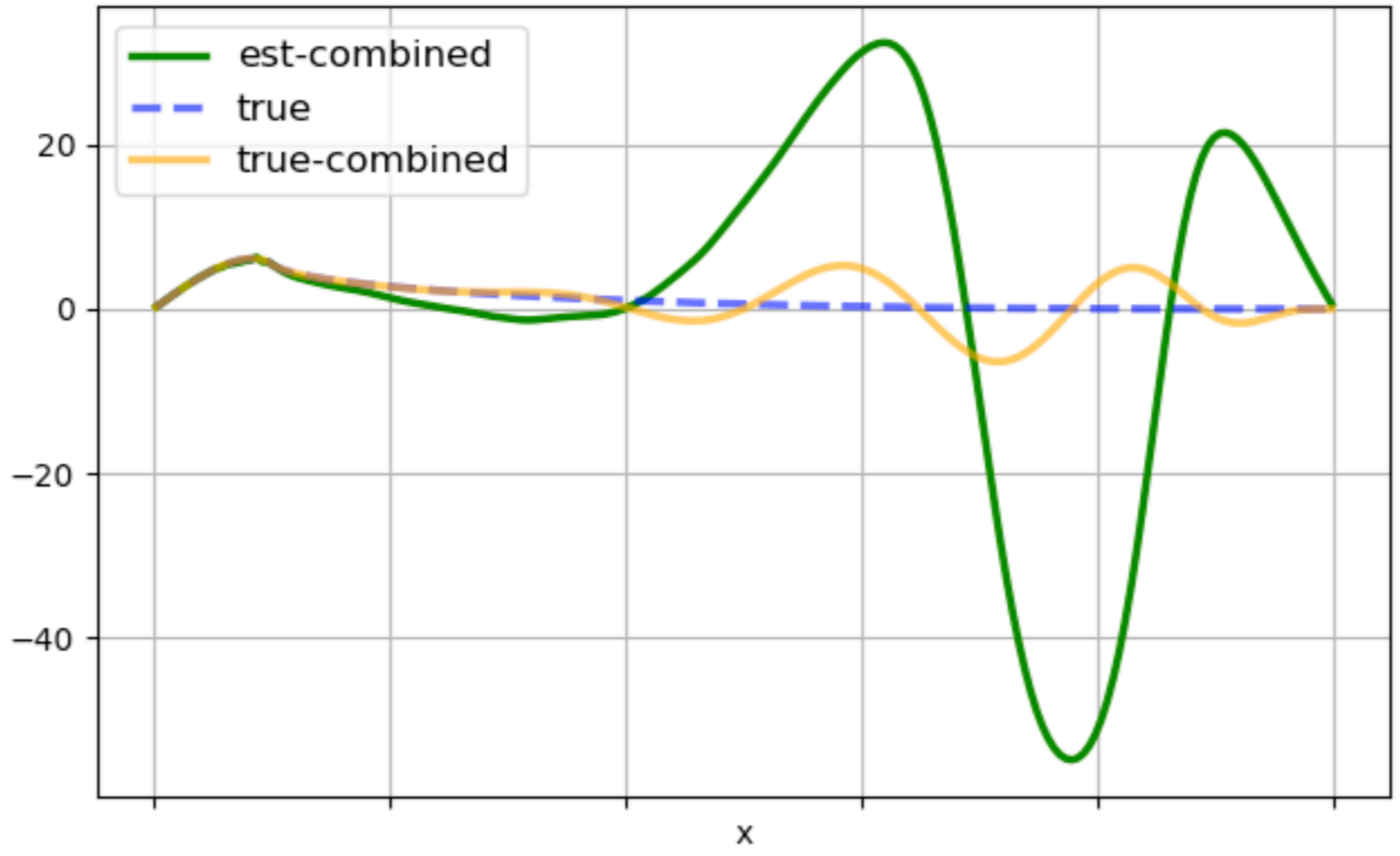
Closure test results





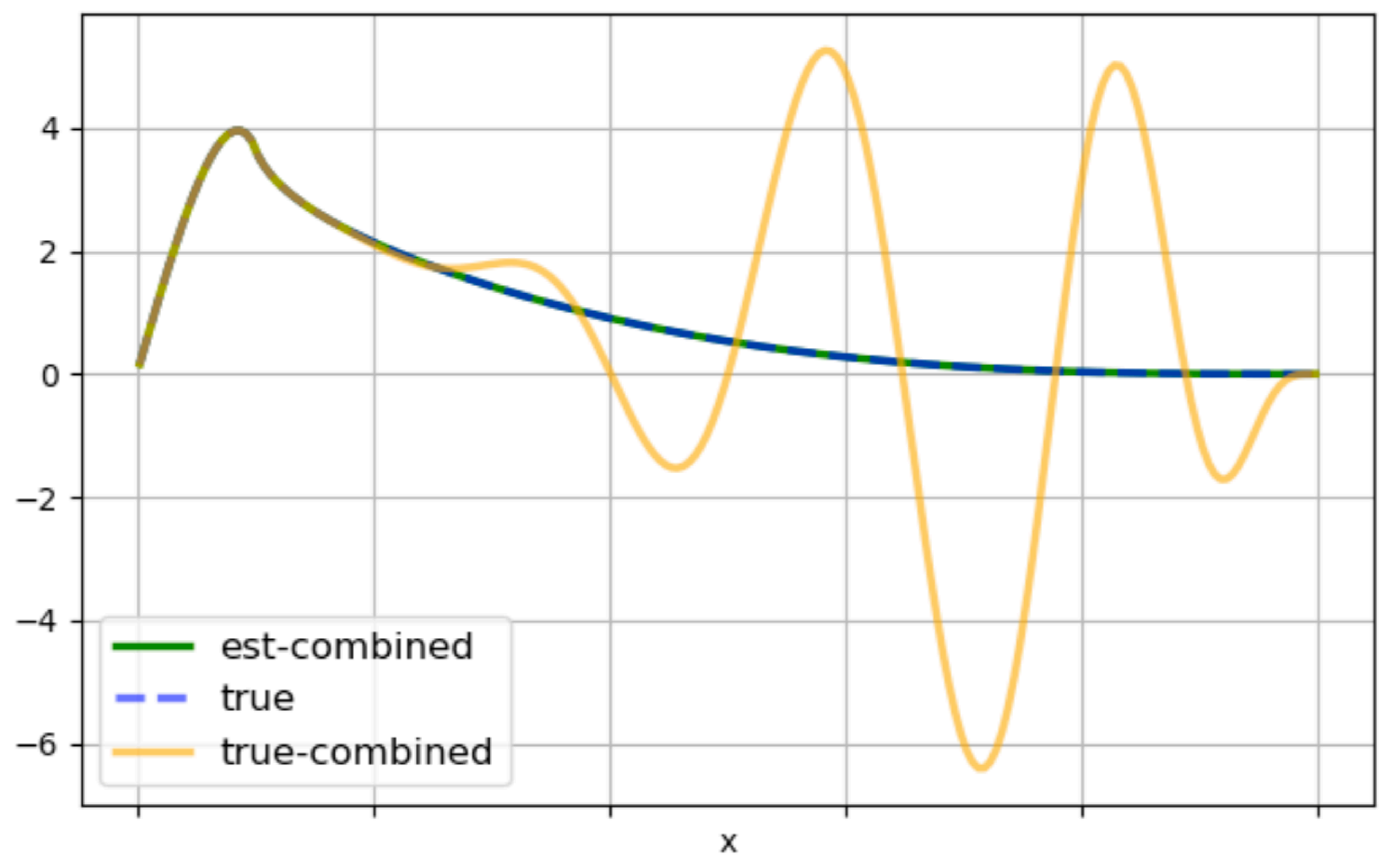
Progress in developing the Neural Network Model

- Utilize the NN to explore how much the data allows the GPD to deviate from a parametric starting point:
    - GPD = Parametric Model + NN
    - Use the same parametric model used in the closure tests
- Trained the model to CFF pseudodata at a fixed energy scale
- Setup is capable of capturing potential SGPDs



Effect of Evolution

- Can use the NN setup to explore the impact of evolution on SGPDs
- Take the NN initially trained at fixed energy and train to data at multiple energy scales
- NN model is able to match the truth when evolution is included.



Conclusion and Next Steps

- Summary:
  - Successful closure tests of fitting machinery with parametric model of the GPDs
  - Developed a NN model capable of capturing SGPDs
- Next Steps:
  - Parametric model:
    - Conduct an analysis with real data
  - NN model:
    - Utilize the model to generate a large sampling of different SGPDs and test the impact of evolution on constraining them
    - Conduct an analysis with real data using the results from the parametric fit as input