Bayesian and Neural Network Approaches to PDF Reconstruction

Joe Karpie
(Columbia University)

As part of the

HadStruc Collaboration

Along with

C. Carlson, C. Egerer, C. Kallidonis, T. Khan,
C. Monahan, K. Orginos, R. Sufian (W&M / JLab)
R. Edwards, B. Joó, J.W. Qiu, D. Richards, E. Romero, F. Winter (JLab)
W. Morris, A. Radyushkin (Old Dominion U / JLab)
A. Rothkopf (Stavanger U)
S. Zafeiropoulos (CPT Marseille)

Lattice “Structure Functions” and Inverse problems

- All modern approaches to calculate the PDF require an inverse problem
  - Experiments, physical or computational, will only have a limited range of data
  - The results of these experiments will be integrals of the PDF, not the PDF directly
- Lattice calculable matrix elements can be Lorentz invariant functions which are factorized into the PDF in analogy to the cross sections and structure functions of experiments.
- Worse yet, Lattice calculations are naturally done in coordinate space in terms of Ioffe time not momentum space in terms of the prefered momentum fraction
  - Leads to Fourier oscillatory or Laplace exponentially decaying natures of the inverse problem

Ioffe Time Pseudo-Distributions

- A general matrix element of interest
    - Analogy to the PDF's matrix element definition

$$M^\alpha(p,z) = \langle p|\bar{\psi}(z)\gamma^\alpha W(z;0)\psi(0)|p\rangle$$

- Lorentz decomposition
    - Physicists love to use of symmetries
    - Choice of p, z, and α can remove higher twist term

$$M^\alpha(p,z) = 2p^\alpha \mathcal{M}(\nu, z^2) + 2z^\alpha \mathcal{N}(\nu, z^2)$$

- Factorizable Relation to PDF
    - Perturbatively calculable Wilson coefficients for each parton with Short distance factorization

$$\mathfrak{M}(\nu, z^2) = \int_{-1}^{1} dx\, K(x\nu, z^2\mu^2) f(x,\mu^2) + \mathcal{O}(z^2)$$

V. Braun and D. Müller (2007) 0709.1348
A.Radyushkin (2017) 1705.01488
Y. Q. Ma and J. W. Qiu (2017) 1709.03018

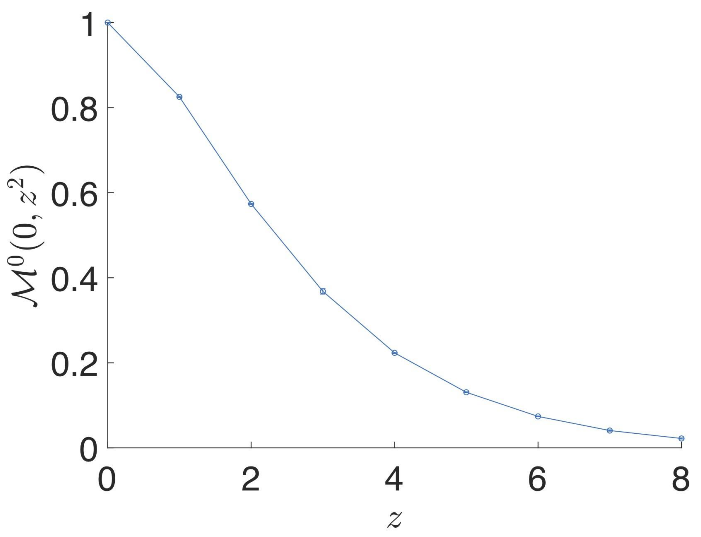
The Reduced distribution and normalization

A.Radyushkin (2017) 1705.01488
T. Izubuchi et. al. (2020) 2007.06590

- The pseudo-ITD usually subject to many systematic errors
  - Lattice spacing, higher twist, incorrect pion mass, finite volume
- A ratio can remove renormalization constants and the low Ioffe time systematic errors
  - In style of ratios from older Lattice calculations of g_V/g_A
  - Avoids additional gauge fixed RI-Mom calculations
  - Is a renormalization group invariant quantity, guaranteeing finite continuum limit

$$\mathfrak{M}(\nu, z^2) = \frac{M^0(p,z)/M^0(p,0)}{M^0(0,z)/M^0(0,0)}$$

- New ratio method with non-zero momentum could remove different HT errors



Pseudo Distribution to MS-bar distribution

- Matching between reduced pseudo-ITD and MS bar scheme ITD via factorization of IR divergences.
- At 1-loop, scale evolution and matching can be simultaneous
- Allows for a direct relationship between ITD/PDF and pseudo-ITD
  - No more need for extrapolations in the scale
  - Does require scale to be in regime dominated by perturbative effects
- Go directly from pseudo-ITD to PDF is numerical unstable
- Only perturbative correction proportional to α_s (around 10%)

$$\mathfrak{M}(\nu, z^2) = Q(\nu, \mu^2) - \frac{\alpha_s C_F}{2\pi} \int_0^1 du \left[B(u) \log(\mu^2 z^2 \frac{e^{2\gamma_E}}{4}) + L(u) \right] Q(u\nu, \mu^2)$$

$$B(u) = \left[\frac{1+u^2}{1-u} \right]_+ \qquad L(u) = \left[4\frac{\ln(1-u)}{1-u} - 2(1-u) \right]_+$$

A. Radyushkin (2017) 1710.08813
J.-H. Zhang (2018) 1801.03023
T. Izubuchi (2018) 1801.03917

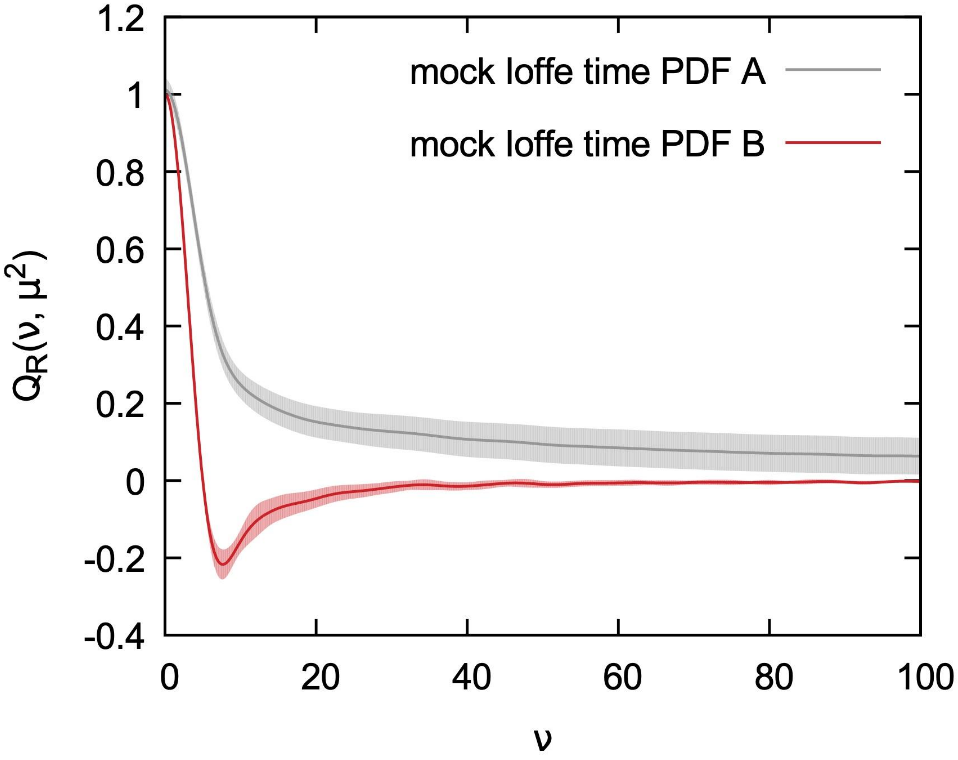
We know how to get data.
What do we do when we have it?

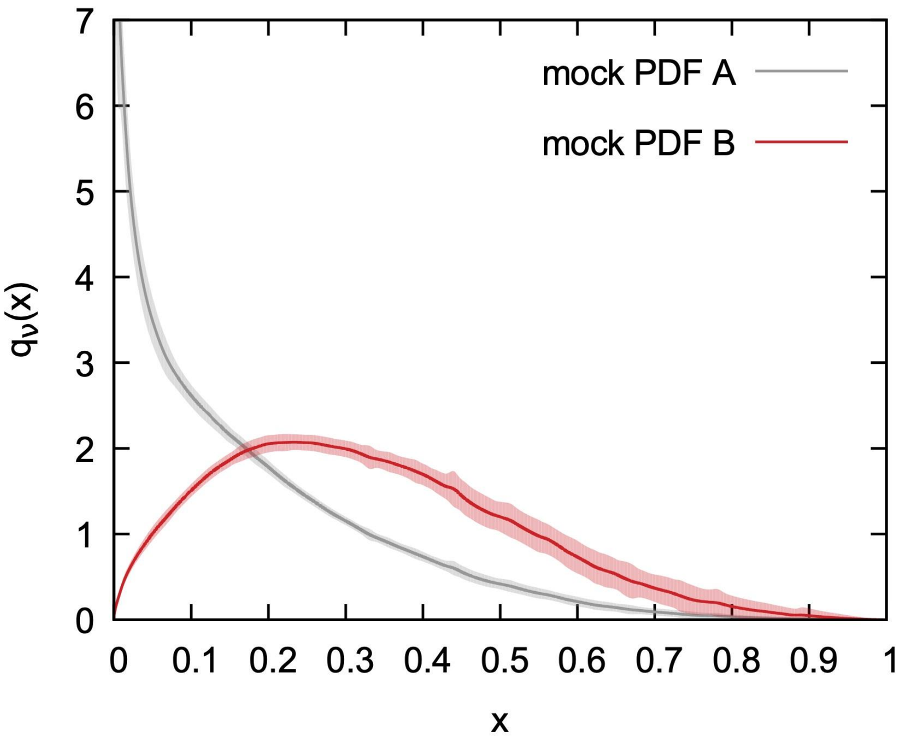
- To study the methods mock data will be used
- Attempts will be made to apply these methods to real data

Mock Trials

JK, K. Orginos, A. Rothkopf, S. Zafeiropoulos (2019) 1901.05408

- Mock Data made from NNPDF31_nnlo_as0118 at scale 2 GeV





Numerical Studies

Dynamical Tree level tadpole Symanzik improved gauge action

- a127m415 : $\beta = 6.1 \quad m_\pi = 415\,\mathrm{MeV} \quad 24^3 \times 64 \quad a = 0.127\,\mathrm{fm}$
- a127m415L : $\beta = 6.1 \quad m_\pi = 415\,\mathrm{MeV} \quad 32^3 \times 96 \quad a = 0.127\,\mathrm{fm}$
- a094m360 : $\beta = 6.3 \quad m_\pi = 358\,\mathrm{MeV} \quad 32^3 \times 64 \quad a = 0.094\,\mathrm{fm}$
- a094m280 : $\beta = 6.3 \quad m_\pi = 278\,\mathrm{MeV} \quad 32^3 \times 64 \quad a = 0.094\,\mathrm{fm}$
- a091m170 : $\beta = 6.3 \quad m_\pi = 172\,\mathrm{MeV} \quad 64^3 \times 128 \quad a = 0.091\,\mathrm{fm}$

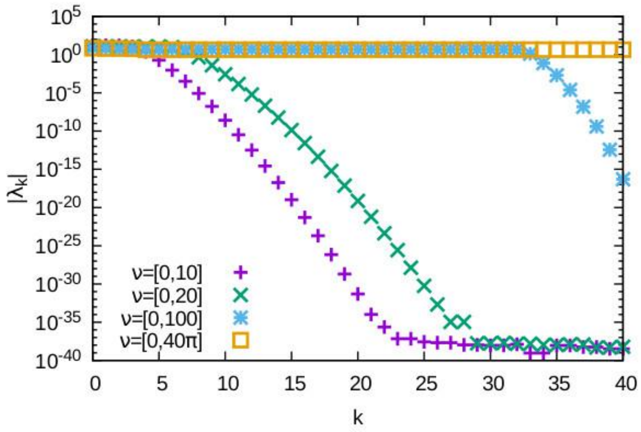
Inverse Solutions for Lattice PDFs

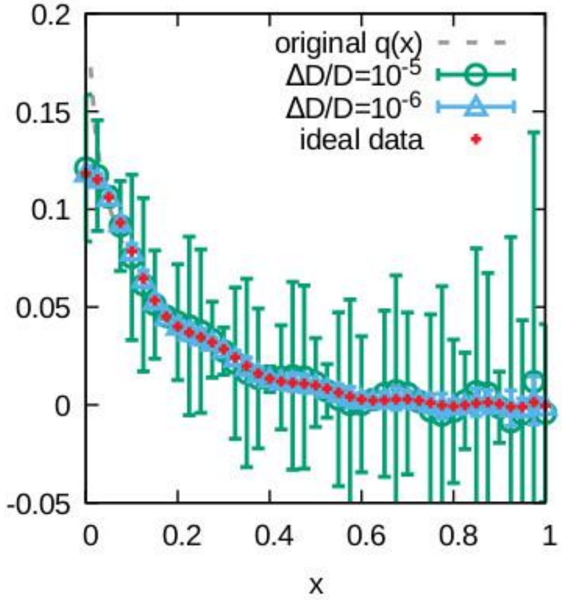
JK, K. Orginos, A. Rothkopf, S. Zafeiropoulos (2019) 1901.05408

- Discrete Fourier Transform
  - The DFT adds additional information that all data outside range is 0
  - For any available lattice data this bad information creates statistically significant oscillations
- Parametric
  - Fit a phenomenologically motivated function
    - Method used by most pheno extractions
    - Potentially significant, but controllable model dependence
  - Fit to a neural network S. Forte, L. Garrido, J. Latorre, A. Piccione (2002) 0204232
    - Machine learning is hip
    - Expensive tuning procedure
- Non-Parametric
  - Backus-Gilbert J. Liang, K-F. Liu, Y-B. Yang (2017) 1710.11145 C. Alexandrou et al (2020) 2008.10573
    - No model dependence, one tunable parameter
  - Bayesian Reconstruction Y. Burnier and A. Rothkopf (2013) 1307.6106 J. Liang, et. al. (2019) 1906.05312
    - Very general, Bayesian statistics has systematics included in meaningful way
  - Bayes-Gauss-Fourier transform C. Alexandrou, G. Iannelli, K. Jansen, F. Manigrasso (2020) 2007.13800

Discrete Fourier Transform Issues

- Additional information that all data outside region is precisely 0 and the points in between are interpolated based on integrator
- Truncated discretized Fourier (cosine) transform is unreliable for realistic lattice data
  - Ill posed inverse problem
  - Consider problem as matrix equation $\mathcal{M}_i = \sum_j C_{ij} q_j$
- Mock test to reconstruct PDF from 40 evenly spacing Ioffe time PDF points given Gaussian noise.
  - Noisy data requires either unreasonably large ranges Ioffe Time unreasonably precise data to reproduce model PDFs.





Backus Gilbert Reconstruction

G. Backus and F. Gilbert, Geophysical Journal International 16, 169 (1968)

J. Liang, K-F. Liu, Y-B. Yang (2017) 1710.11145

C. Alexandrou et al (2020) 2008.10573

- Finds "most stable", i.e. lowest variance, solution to inverse
- Used in wide range of engineering and physics applications
- Used to extract PDF from Lattice calculation of Hadronic Tensor
- Create "Delta function" $\Delta(x-\bar{x}) = \sum_j q_j(\bar{x}) K_j(x)$ and minimize its width

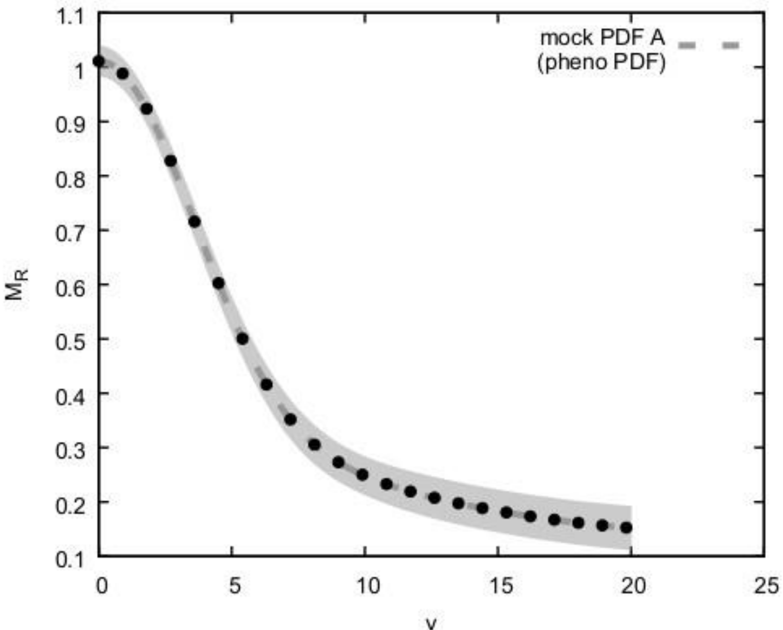
  $$\sigma = \int dx (x-\bar{x})^2 \Delta(x-\bar{x})$$

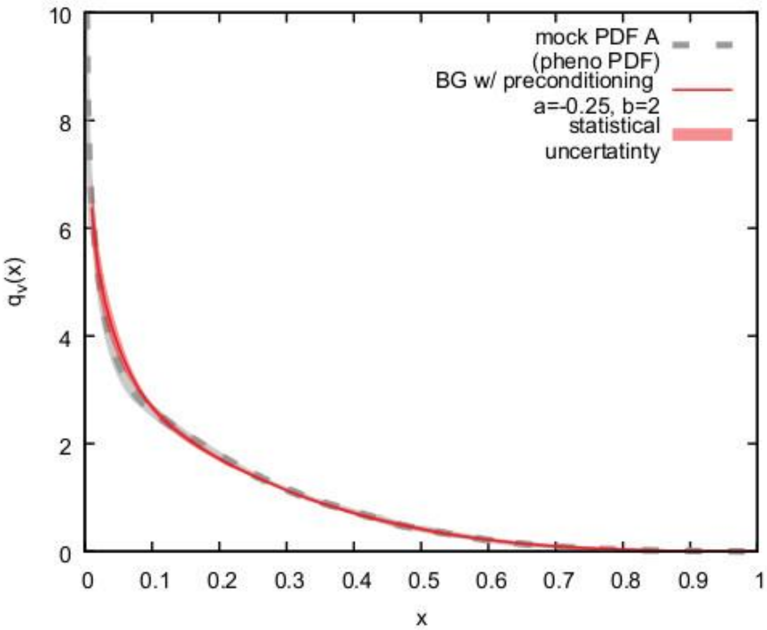
- In limit of width to 0, the Backus Gilbert method would reconstruct exact unknown function

See talks of
Jian Liang and
Aurora Scapellato,
Tuesday

Mock Tests of Backus Gilbert Reconstruction

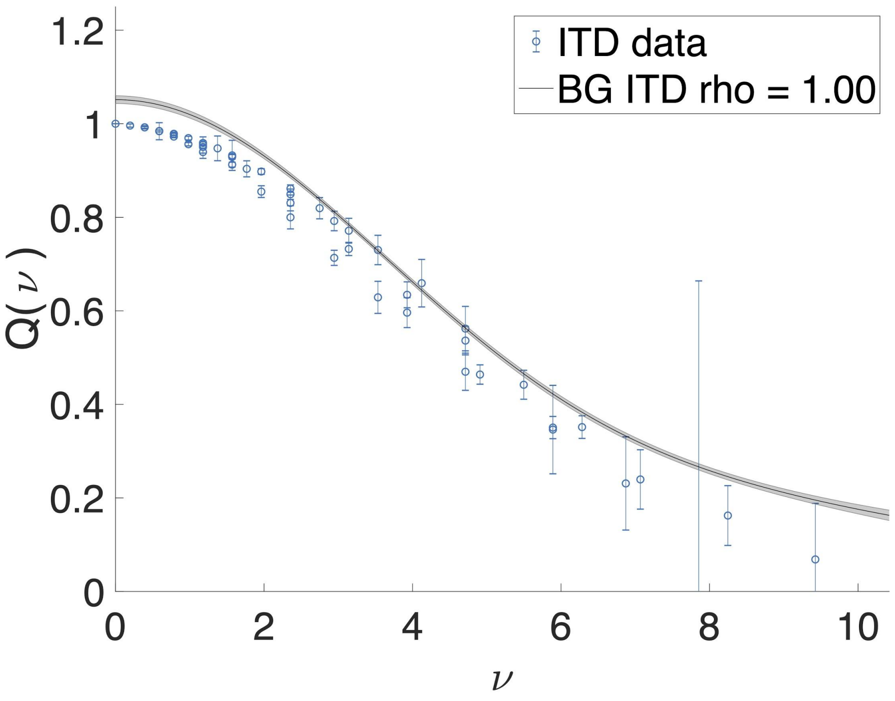
- Tests use NNPDF31_nnlo_as_0118 data set with artificial errors.
- Reconstructions are more stable and reliable than direct inversion or fits.
  Backus Gilbert

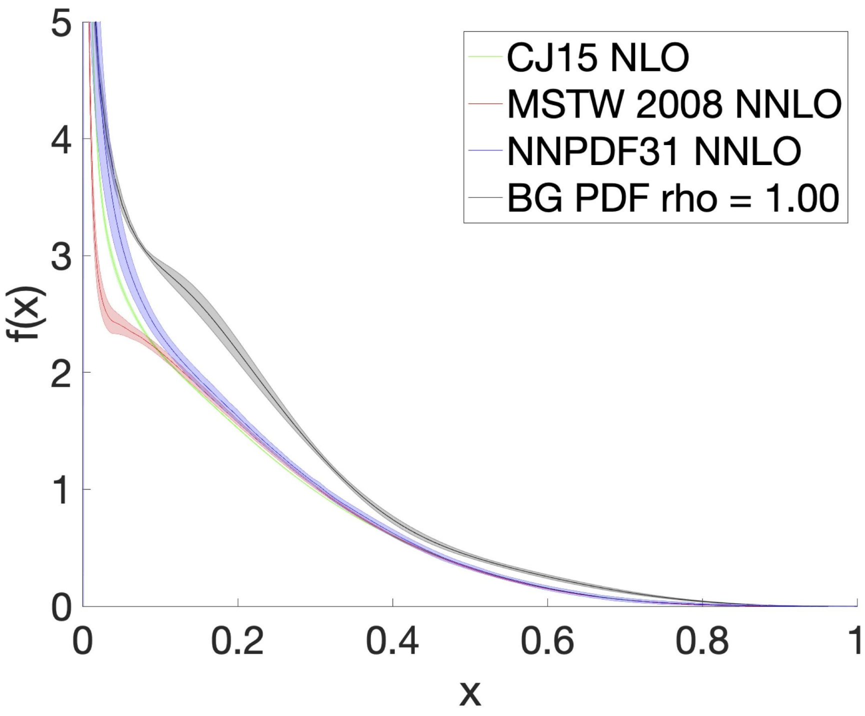




Results from Backus Gilbert

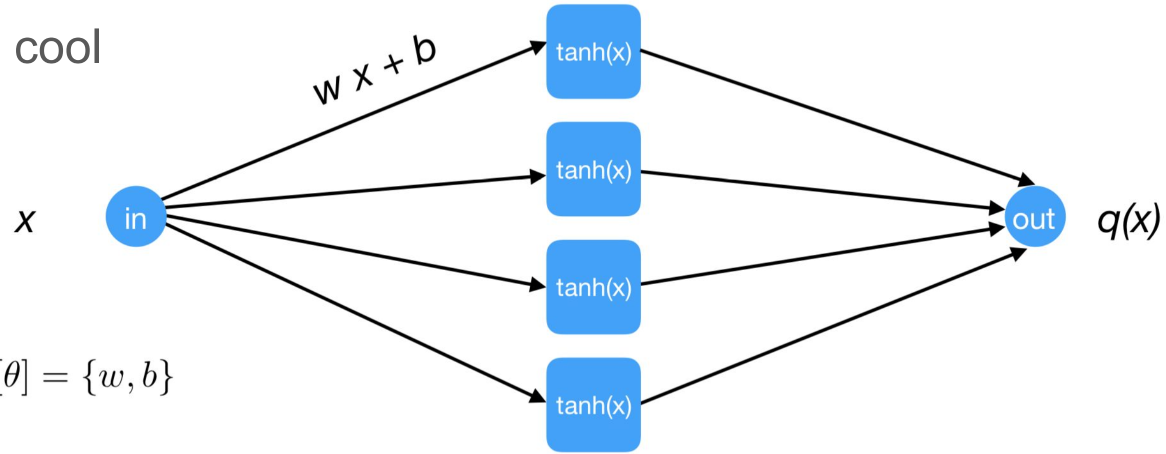
$$\mathrm{a127m415L}: \quad \beta = 6.1 \quad m_\pi = 415\,\mathrm{MeV} \quad 32^3 \times 96 \quad a = 0.127\,\mathrm{fm}$$





Neural Network Reconstruction

- In the style of NNPDF, a series of neural networks can be constructed to represent the ill posed inverse transformation.
    - Many choices of Network geometry and activation functions need to be explored
- Even a small Neural net can be used to reconstruct PDF to accuracy of other methods.
- Machine Learning is cool



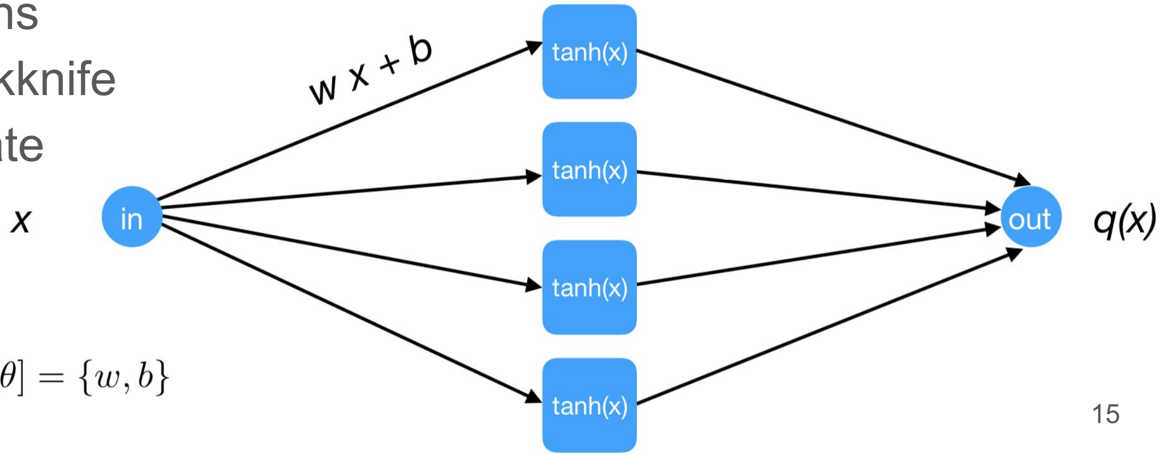
$$[\theta] = \{w, b\}$$

Neural Network Procedure

- Choose network geometry and activation function
- Using the full dataset, minimize network with respect to

$$\chi^2 = (M_k - \int_0^1 dx cos(\nu_k x) q_{[\theta]}(x)) C^{-1} (M_k - \int_0^1 dx cos(\nu_k x) q_{[\theta]}(x))^T$$

  several times, removing networks with largest value
- Repeat for a few generations
- Retrain each replica on jackknife samples to get error estimate

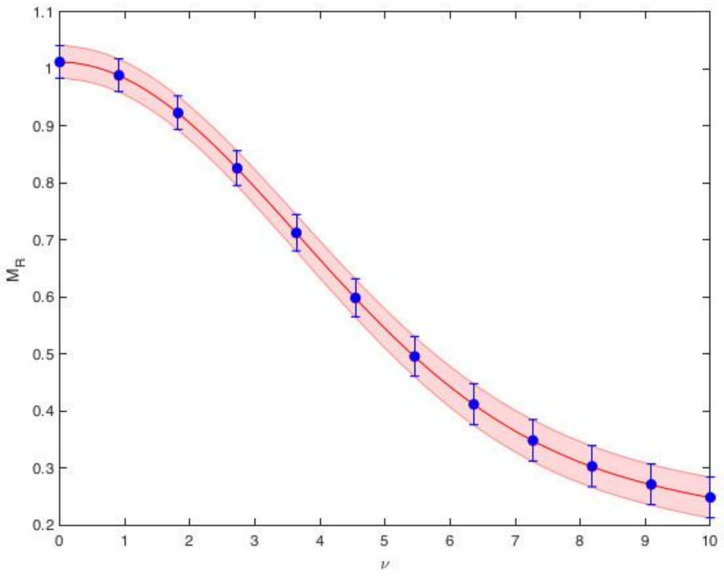


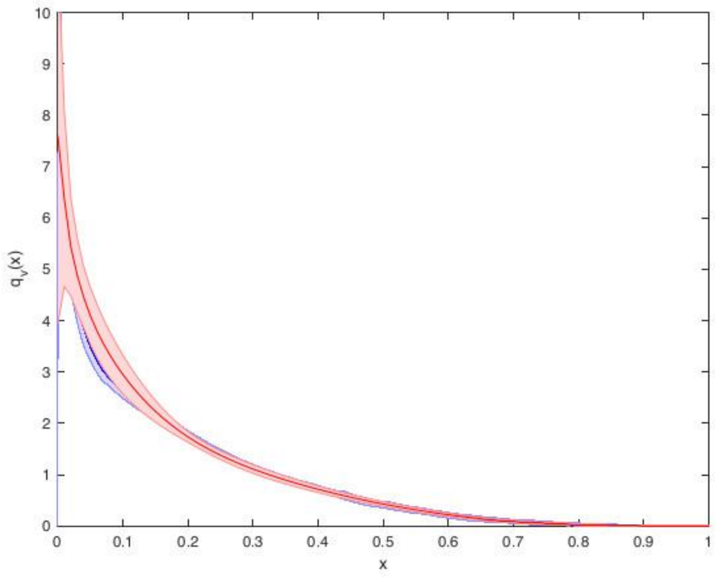
$$[\theta] = \{w, b\}$$

Mock Tests of Neural Network Reconstructions

- Tests use modified NNPDF data set with artificial errors.
- Reconstructions are more stable and reliable than direct inversion or fits.

  Neural Network

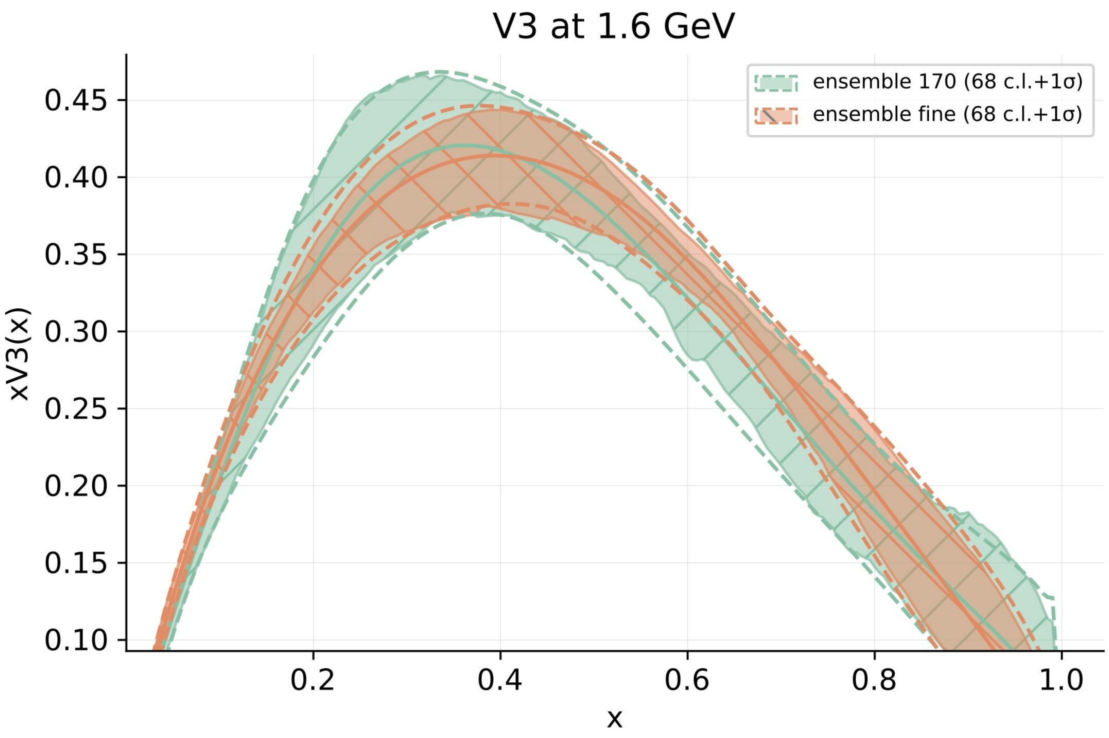




Results from NNPDF Framework

$$\text{a094m360}: \quad \beta = 6.3 \quad m_\pi = 358\,\text{MeV} \quad 32^3 \times 64 \quad a = 0.094\,\text{fm}$$

$$\text{a091m170}: \quad \beta = 6.3 \quad m_\pi = 172\,\text{MeV} \quad 64^3 \times 128 \quad a = 0.091\,\text{fm}$$



See talk of Tommaso Giani, Wednesday

L. Del Debbio, T. Gianni, JK, K. Orginos, A. Radyushkin, S. Zafeiropoulos (2020) 2010.03996

Bayesian Reconstruction

Y. Burnier and A. Rothkopf (2013) 1307.6106

J. Liang, et. al. (2019) 1906.05312

- Technique based upon Bayes Theorem

$$P[q|M,I] = \frac{P[M|q,I]P[q|I]}{P[M|I]}$$

- Acknowledging the ill posed nature of the problem and that a unique solution require addition of further information
- Parameterize the probabilities and extremize the posterior probability
- Developed for extraction of quark spectral function which is a much harder application

$$P[M|q,I] = exp[-L]$$

$$P[q|I] = exp[-S] \qquad S = \sum_n \Delta x_n \left(1 - \frac{q_n}{m_n} + \log \frac{q_n}{m_n}\right)$$

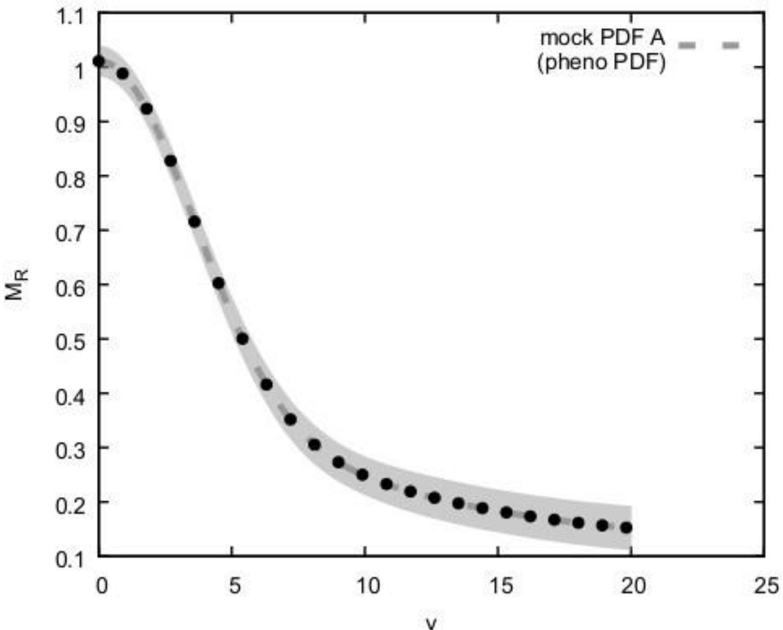
$$P[M|I] = N$$

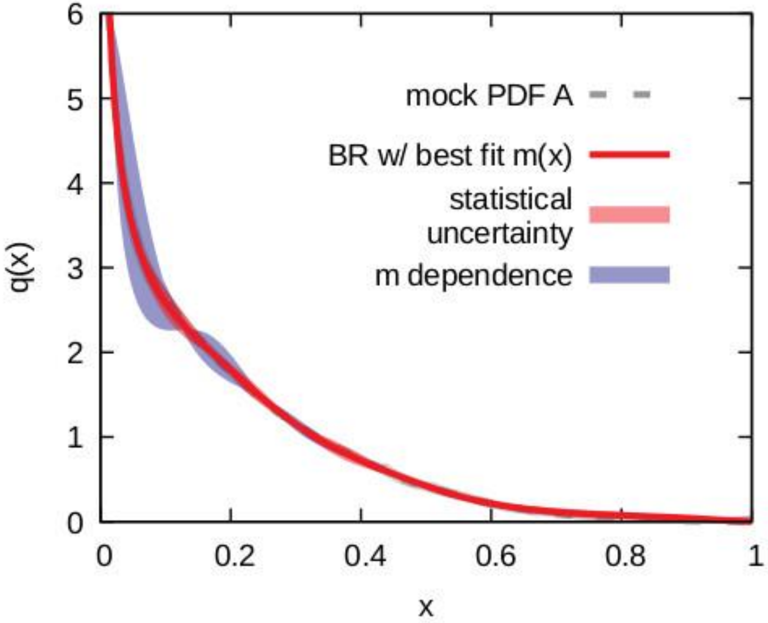
See talk of Jian Liang, Tuesday

Mock Tests of Bayesian Reconstruction

- Tests use NNPDF31_nnlo_as_0118 data set with artificial errors.
- Reconstructions are more stable and reliable than direct inversion or fits.

  Bayesian Reconstruction





Summary and Outlook

- Much work is needed to control systematic errors from inverse problem
- Methods of combining results from different solutions are required to reduce or remove biases that they all contain
- Future applications of non-parametric inversions could remove potential biases from current fits
- The inverse problem is potentially the largest hinderance to a systematically controlled PDF calculation from Lattice QCD