# Progress on parton pseudo distributions I

**Savvas Zafeiropoulos**

Universität Heidelberg

23.07.2018

Lattice 2018

In collaboration with J. Karpie (College of William & Mary), K. Orginos (College of William & Mary and JLAB), A. Radyushkin (ODU and JLAB) and A. Rothkopf (Stavanger U.)

# Light-like is a NO-GO

Hadronic Tensor Methods

- "Light-like" separated Hadronic Tensor K. F. Liu et al Phys.Rev.Lett. 72 (1994), A. J. Chambers et al Phys.Rev.Lett. 118 (2017)

Ioffe Time Pseudo Distribution Methods

- quasi-PDFs (X. Ji Phys.Rev.Lett. 110, (2013))
- pseudo-PDFs (A. Radyushkin Phys.Lett. B767 (2017))

Similarly to a global QCD analysis of high energy scattering data, PDFs can also be extracted from analyzing data generated by lattice-QCD calculation of good lattice cross-sections Y.-Q. Ma and J.-W. Qiu Phys. Rev. Lett. 120 (2018)

# Formalism

Computing PDFs in LQCD we start from the equal time hadronic matrix element with the quark and anti-quark fields separated by a finite distance. For non-singlet parton densities the matrix element

$$\mathcal{M}^{\alpha}(z,p) \equiv \langle p|\bar{\psi}(0)\,\gamma^{\alpha}\,\hat{E}(0,z;A)\tau_3\psi(z)|p\rangle$$

where $\hat{E}(0,z;A)$ is the $0 \to z$ straight-line gauge link in the fundamental representation, $\tau_3$ is the flavor Pauli matrix, and $\gamma^a$ is a gamma matrix. We can decompose the matrix element due to Lorentz invariance as

$$\mathcal{M}^{\alpha}(z,p) = 2p^{\alpha}\mathcal{M}_p(-(zp),-z^2) + z^{\alpha}\mathcal{M}_z(-(zp),-z^2)$$

# Formalism

- From the $\mathcal{M}_p(-(zp), -z^2)$ part the twist-2 contribution to PDFs can be obtained in the limit $z^2 \to 0$.
- By taking $z = (0, 0, 0, z_3)$, $\alpha$ in the temporal direction i.e. $\alpha = 0$, and the hadron momentum $p = (p^0, 0, 0, p)$ the $z^\alpha$-part drops out.
- The Lorentz invariant quantity $\nu = -(zp)$, is the "Ioffe time" (B. L. Ioffe, Phys. Lett. 30B, 123 (1969)) and

$$\langle p|\bar{\psi}(0)\,\gamma^0\,\hat{E}(0, z; A)\tau_3\psi(z)|p\rangle \ = 2p^0\mathcal{M}_p(\nu, z_3^2)$$

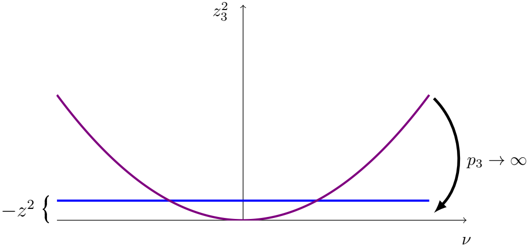
# Formalism

- The quasi-PDF $Q(x, p^2)$ is related to $\mathcal{M}_p(\nu, z_3^2)$ by

$$Q(x, p^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\nu\ e^{-ix\nu}\, \mathcal{M}_p(\nu, [\nu/p]^2)$$

  Quasi PDF mixes invariant scales until $p_z$ is effectively large enough
- While the pseudo-PDF has fixed invariant scale dependence

$$P(x, z_0^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\nu\ e^{-ix\nu}\, \mathcal{M}_p(\nu, z_0^2)$$



# Formalism

Ioffe time PDFs $\mathcal{M}(\nu, z_3^2)$ defined at a scale $\mu^2 = 4e^{-2\gamma_E}/z_3^2$ (at leading log level) are the Fourier transform of regular PDFs $f(x, \mu^2)$. (I.I. Balitsky and V.M. Braun, Nucl. Phys. B311, 541 (1988), V. Braun, et. al Phys. Rev. D 51, 6036 (1995)), A. Radyushkin Phys.Rev. D98 (2018) no.1, 014019

$$\mathcal{M}(\nu, z_3^2) = \int_{-1}^{1} dx\, f(x, 1/z_3^2) e^{ix\nu}$$

Scale dependence of the Ioffe time PDF derived from the DGLAP evolution of the regular PDFs.

Ioffe time PDFs evolution equation

$$\frac{d}{d\ln z_3^2}\,\mathcal{M}(\nu, z_3^2) = -\frac{\alpha_s}{2\pi}\, C_F \int_0^1 du\, B(u)\mathcal{M}(u\nu, z_3^2)$$

with $B(u) = \left[\frac{1+u^2}{1-u}\right]_+$, $C_F = 4/3$, and $B(u)$ is the LO evolution kernel for the non-singlet quark PDF (V. Braun, et. al Phys. Rev. D 51, 6036 (1995))

# Obtaining the Ioffe time PDF

$$z_3 \to 0 \Rightarrow \mathcal{M}_p(\nu, z_3^2) = \mathcal{M}(\nu, z_3^2) + \mathcal{O}(z_3^2)$$

But.... large $\mathcal{O}(z_3^2)$ corrections prohibit the extraction.
Conservation of the vector current implies $\mathcal{M}_p(0, z_3^2) = 1 + \mathcal{O}(z_3^2)$ ,
but in a ratio $z_3^2$ corrections (related to the transverse structure of the hadron) might cancel (A. Radyushkin Phys.Lett. B767 (2017))

$$\mathfrak{M}(\nu, z_3^2) \equiv \frac{\mathcal{M}_p(\nu, z_3^2)}{\mathcal{M}_p(0, z_3^2)}$$

- Much smaller $\mathcal{O}(z_3^2)$ corrections and therefore this ratio could be used to extract the Ioffe time PDFs
- A well defined continuum limit and does not require renormalization

# Matching to $\overline{MS}$

- In Phys.Rev. D98 (2018) no.1, 014019 it was shown by Radyushkin that at 1-loop evolution and matching to $\overline{MS}$ can be done simultaneously.
- This establishes a direct relation between the Ioffe time distribution function (ITDF) and pseudo-ITDF.
- Scales are needed as such that we are in a regime dominated by perturbative effects

$$
\begin{aligned}
\mathcal{I}(\nu,\mu^2) =& \mathfrak{M}(\nu,z_3^2) + \frac{\alpha_s}{\pi} C_F \int_0^1 dw\, \mathfrak{M}(w\nu, z_3^2) \\
& \times \Bigg\{ B(w) \ln\left[(1-w) z_3 \mu \frac{e^{\gamma_E + 1/2}}{2}\right] \\
& + \left[(w+1)\ln(1-w) - (1-w)\right]_+ \Bigg\}
\end{aligned}
$$

# Matching to $\overline{MS}$

- In Phys.Rev. D98 (2018) no.1, 014019 it was shown by Radyushkin that at 1-loop evolution and matching to $\overline{MS}$ can be done simultaneously.
- This establishes a direct relation between the Ioffe time distribution function (ITDF) and pseudo-ITDF.
- Scales are needed as such that we are in a regime dominated by perturbative effects

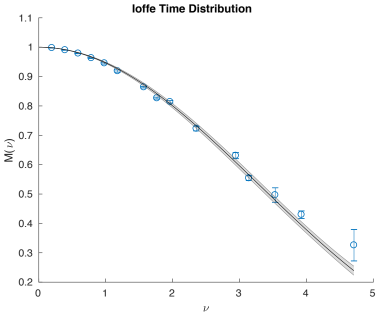
$$\mathcal{I}(\nu,\mu^2) = \mathfrak{M}(\nu, z_3^2) + \frac{\alpha_s}{\pi} C_F \int_0^1 dw\, \mathfrak{M}(w\nu, z_3^2) \times \left\{ B(w) \ln\left[(1-w) z_3 \mu \frac{e^{\gamma_E + 1/2}}{2}\right] + \left[(w+1)\ln(1-w) - (1-w)\right]_+ \right\}$$

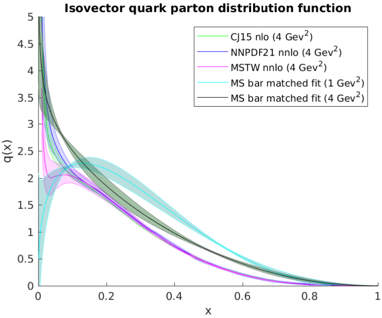
# Matching to $\overline{MS}$

- In Phys.Rev. D98 (2018) no.1, 014019 it was shown by Radyushkin that at 1-loop evolution and matching to $\overline{MS}$ can be done simultaneously.
- This establishes a direct relation between the Ioffe time distribution function (ITDF) and pseudo-ITDF.
- Scales are needed as such that we are in a regime dominated by perturbative effects

$$\begin{aligned}\mathcal{I}(\nu,\mu^2) =& \mathfrak{M}(\nu,z_3^2) + \frac{\alpha_s}{\pi}\,C_F\,\int_0^1 dw\,\mathfrak{M}(w\nu,z_3^2) \\ & \times\left\{B(w)\,\ln\left[(1-w)z_3\mu\frac{e^{\gamma_E+1/2}}{2}\right]\right. \\ & \left.+\left[(w+1)\ln(1-w)-(1-w)\right]_+\right\}\end{aligned}$$

# Comparison to global fits after converting to $\overline{MS}$



# Reconstruction

- Parton distribution functions (PDF) or distribution amplitudes (DA) may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position space matrix elements.
- One particular example are the Ioffe-time PDFs $\mathfrak{M}_R$, which are related to the physical PDF via the integral relation

$$\mathfrak{M}_R(\nu, \mu^2) \equiv \int_0^1 dx \, \cos(\nu x) \, q_v(x, \mu^2) \, .$$

- Here it is assumed that the lattice computed matrix element is converted to the $\overline{MS}$ Ioffe-time PDF at a scale $\mu^2$, using a perturbative kernel as discussed in Radyushkin (Phys.Rev. D98 (2018) no.1, 014019 ), Zhang et al Phys.Rev. D97 (2018) no.7, 074508
- The task at hand is then to reconstruct the PDF $q_v(x, \mu^2)$ given a limited set of simulated data for $\mathfrak{M}_R(\nu, \mu^2)$.

# Reconstruction

- Parton distribution functions (PDF) or distribution amplitudes (DA) may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position space matrix elements.
- One particular example are the Ioffe-time PDFs $\mathfrak{M}_R$, which are related to the physical PDF via the integral relation

$$\mathfrak{M}_R(\nu, \mu^2) \equiv \int_0^1 dx\ \cos(\nu x)\, q_v(x, \mu^2)\,.$$

- Here it is assumed that the lattice computed matrix element is converted to the $\overline{MS}$ Ioffe-time PDF at a scale $\mu^2$, using a perturbative kernel as discussed in Radyushkin (Phys.Rev. D98 (2018) no.1, 014019 ), Zhang et al Phys.Rev. D97 (2018) no.7, 074508
- The task at hand is then to reconstruct the PDF $q_v(x, \mu^2)$ given a limited set of simulated data for $\mathfrak{M}_R(\nu, \mu^2)$.

# Reconstruction

- Parton distribution functions (PDF) or distribution amplitudes (DA) may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position space matrix elements.
- One particular example are the Ioffe-time PDFs $\mathfrak{M}_R$, which are related to the physical PDF via the integral relation

$$\mathfrak{M}_R(\nu, \mu^2) \equiv \int_0^1 dx \, \cos(\nu x) \, q_v(x, \mu^2) \, .$$

- Here it is assumed that the lattice computed matrix element is converted to the $\overline{MS}$ Ioffe-time PDF at a scale $\mu^2$, using a perturbative kernel as discussed in Radyushkin (Phys.Rev. D98 (2018) no.1, 014019 ), Zhang et al Phys.Rev. D97 (2018) no.7, 074508
- The task at hand is then to reconstruct the PDF $q_v(x, \mu^2)$ given a limited set of simulated data for $\mathfrak{M}_R(\nu, \mu^2)$.

# Reconstruction

- Parton distribution functions (PDF) or distribution amplitudes (DA) may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position space matrix elements.
- One particular example are the Ioffe-time PDFs $\mathfrak{M}_R$, which are related to the physical PDF via the integral relation

$$\mathfrak{M}_R(\nu, \mu^2) \equiv \int_0^1 dx \, \cos(\nu x) \, q_v(x, \mu^2) \, .$$

- Here it is assumed that the lattice computed matrix element is converted to the $\overline{MS}$ Ioffe-time PDF at a scale $\mu^2$, using a perturbative kernel as discussed in Radyushkin (Phys.Rev. D98 (2018) no.1, 014019 ), Zhang et al Phys.Rev. D97 (2018) no.7, 074508
- The task at hand is then to reconstruct the PDF $q_v(x, \mu^2)$ given a limited set of simulated data for $\mathfrak{M}_R(\nu, \mu^2)$.

# Reconstruction

- There exist two challenges to this endeavor, the first being that the integral in question does *not extend over the full Brillouin zone*, the second that in practice only a *small number of points along* $\nu$ can be computed.
- As we will discuss in more detail below, taken together these issues render the extraction highly ill-posed and we explore different regularization strategies on how to nevertheless reliably estimate the PDF from the data at hand.
- Phenomenological investigations of PDFs have shown that their functional form may be reasonably well approximated by the following simple Ansatz

$$p(x) = \frac{\Gamma(a+b+2)}{\Gamma(a+1)\Gamma(b+1)} x^a (1-x)^b \, .$$

# Reconstruction

- There exist two challenges to this endeavor, the first being that the integral in question does *not extend over the full Brillouin zone*, the second that in practice only a *small number of points along* $\nu$ can be computed.
- As we will discuss in more detail below, taken together these issues render the extraction highly ill-posed and we explore different regularization strategies on how to nevertheless reliably estimate the PDF from the data at hand.
- Phenomenological investigations of PDFs have shown that their functional form may be reasonably well approximated by the following simple Ansatz

$$p(x) = \frac{\Gamma(a+b+2)}{\Gamma(a+1)\Gamma(b+1)} x^a (1-x)^b \,.$$

# Reconstruction

- There exist two challenges to this endeavor, the first being that the integral in question does *not extend over the full Brillouin zone*, the second that in practice only a *small number of points along* $\nu$ can be computed.
- As we will discuss in more detail below, taken together these issues render the extraction highly ill-posed and we explore different regularization strategies on how to nevertheless reliably estimate the PDF from the data at hand.
- Phenomenological investigations of PDFs have shown that their functional form may be reasonably well approximated by the following simple Ansatz

$$p(x) = \frac{\Gamma(a+b+2)}{\Gamma(a+1)\Gamma(b+1)} x^a (1-x)^b \,.$$

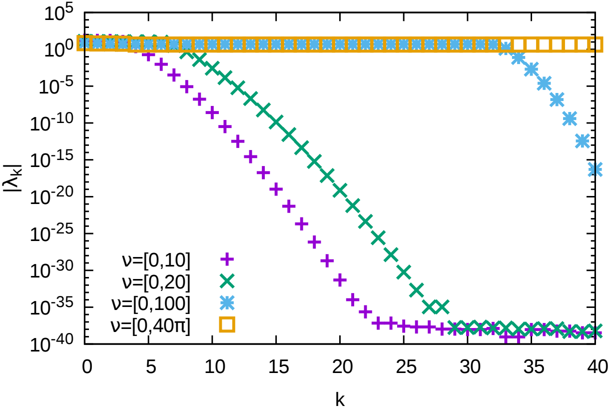
# Naive Reconstruction

- Discretize the integral, employing the trapezoid integration rule
- $\Delta x = \frac{1}{N_x}\,, \;\; x_k = k\Delta x = \frac{k}{N_x}$

  $$\mathfrak{M}_R(\nu) = \tfrac{1}{2}\cos(\nu x_0)\, q_v(x_0) + \sum_{k=1}^{N_x-1} \delta x\, \cos(\nu x_k)\, q_v(x_k) + \frac{1}{2}\cos(\nu x_{N_x})\, q_v(x_{N_x})$$

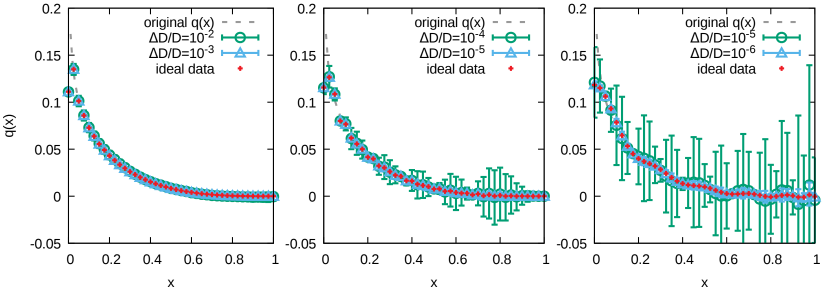
  We can determine the unknown values of the function $q_v(x_k)$ by solving a simple linear system of equations.
- Defining $\mathfrak{m}_k = \mathfrak{M}_R(\nu_k)$ where $\nu_k$ are the values of the Ioffe time for which data is available and $\mathfrak{q}$ be the vector with components the unknown values of $q_v(x_k)$ *i.e.* $\mathfrak{q}_k = q_v(x_k)$. Our problem is cast in a matrix equation $\mathfrak{m} = \mathfrak{C} \cdot \mathfrak{q}$,
- The conditioning of the problem is easily elucidated by considering the eigenvalues of the matrix $\mathfrak{C}$.

# Naive Reconstruction




**Eigenvalues $\lambda_k$ of the kernel matrix for various discretization intervals.** Only for the case corresponding to a genuine discrete Fourier transform $\nu = [0, 40\pi]$ all eigenvalues remain of order unity. The realistic case of $\nu = [0, 20]$ already shows a significant degradation of the spectrum.

# Naive Reconstruction




**Results for the direct inversion for different discretization intervals** (left $\nu = [0, 40\pi]$, center $\nu = [0, 100]$, right $\nu = [0, 20]$). Note the different size of the relative errors needed, to obtain a well behaved result (left $\Delta\mathfrak{M}_R/\mathfrak{M}_R = 10^{-2}$, center $\Delta\mathfrak{M}_R/\mathfrak{M}_R = 10^{-5}$, right $\Delta\mathfrak{M}_R/\mathfrak{M}_R = 10^{-6}$).

# Advanced PDF Reconstructions

- A versatile approach is Bayesian inference Y. Burnier and A. Rothkopf Phys.Rev.Lett. 111 (2013)
- It acknowledges the fact that the inverse problem is ill-defined and a unique answer may only provided, once further information about the system has been made available.
- Formulated in terms of probabilities, one finds in the form of Bayes theorem that

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

It states that the so called posterior probability $P[q|\mathfrak{M},I]$ for a test function $q$ to be the correct $x$-space PDF, given our simulated Ioffe-time PDF $\mathfrak{M}$ and additional prior information may be expressed in terms of three quantities.

# Advanced PDF Reconstructions

- A versatile approach is Bayesian inference Y. Burnier and A. Rothkopf Phys.Rev.Lett. 111 (2013)
- It acknowledges the fact that the inverse problem is ill-defined and a unique answer may only provided, once further information about the system has been made available.
- Formulated in terms of probabilities, one finds in the form of Bayes theorem that

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

It states that the so called posterior probability $P[q|\mathfrak{M},I]$ for a test function $q$ to be the correct $x$-space PDF, given our simulated Ioffe-time PDF $\mathfrak{M}$ and additional prior information may be expressed in terms of three quantities.

# Advanced PDF Reconstructions

- A versatile approach is Bayesian inference Y. Burnier and A. Rothkopf Phys.Rev.Lett. 111 (2013)
- It acknowledges the fact that the inverse problem is ill-defined and a unique answer may only provided, once further information about the system has been made available.
- Formulated in terms of probabilities, one finds in the form of Bayes theorem that

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

It states that the so called posterior probability $P[q|\mathfrak{M}, I]$ for a test function $q$ to be the correct $x$-space PDF, given our simulated Ioffe-time PDF $\mathfrak{M}$ and additional prior information may be expressed in terms of three quantities.

# Bayesian Reconstruction

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

- The likelihood probability $P[\mathfrak{M}|q, I]$ denotes how probable it is to find the data $\mathfrak{M}$ if $q$ were the correct PDF.
- Finding the most probable $q$ by maximizing the likelihood is nothing but a $\chi^2$ fit to the $\mathfrak{M}$ data, which as we saw from the direct inversion is by itself ill-defined.
- The prior probability $P[q|I]$, which quantifies, how compatible our test function $q$ is with respect to any prior information we have (e.g. appearance of non-analytic behavior of $q(x)$ at the boundaries of the $x$ interval).
- $P[\mathfrak{M}|I]$, the so called evidence is a $q$ independent normalization.

# Bayesian Reconstruction

- For sampled data, due to the central limit theorem, the likelihood probability may be written as the quadratic distance functional $P[\mathfrak{M}|q,I] = \exp[-L]$ with $L = \frac{1}{2}\sum_{k,l}(\mathfrak{M}_k - \mathfrak{M}_k^q)C_{kl}^{-1}(\mathfrak{M}_l - \mathfrak{M}_l^q)$.
- $\mathfrak{M}_k^q$ are the Ioffe-time data arising from inserting the test function $q$ into the cosine Fourier trafo and $C_{kl}$ denotes the covariance matrix of the $N_m$ measurements of simulation data $\mathfrak{M}_k^h$.
- We then specify an appropriate prior probability $P[q|I] = \exp[\alpha S[I]]$.
- Prior information enters in two ways here. On the one hand we deploy a particular functional form of the regularization functional

$$S_{BR}[q,m] = \sum_n \Delta x_n \Big(1 - \frac{q_n}{m_n} + \log\big(\frac{q_n}{m_n}\big)\Big)$$

which may be obtained by requiring positive definiteness of the resulting $q$, smoothness of $q$.

# Bayesian Reconstruction

- For sampled data, due to the central limit theorem, the likelihood probability may be written as the quadratic distance functional $P[\mathfrak{M}|q,I] = \exp[-L]$ with $L = \frac{1}{2}\sum_{k,l}(\mathfrak{M}_k - \mathfrak{M}_k^q)C_{kl}^{-1}(\mathfrak{M}_l - \mathfrak{M}_l^q)$.
- $\mathfrak{M}_k^q$ are the Ioffe-time data arising from inserting the test function $q$ into the cosine Fourier trafo and $C_{kl}$ denotes the covariance matrix of the $N_m$ measurements of simulation data $\mathfrak{M}_k^h$.
- We then specify an appropriate prior probability $P[q|I] = \exp[\alpha S[I]]$.
- Prior information enters in two ways here. On the one hand we deploy a particular functional form of the regularization functional

$$S_{BR}[q,m] = \sum_n \Delta x_n \Big(1 - \frac{q_n}{m_n} + \log\big(\frac{q_n}{m_n}\big)\Big)$$

which may be obtained by requiring positive definiteness of the resulting $q$, smoothness of $q$.

# Bayesian Reconstruction

- For sampled data, due to the central limit theorem, the likelihood probability may be written as the quadratic distance functional $P[\mathfrak{M}|q,I] = \exp[-L]$ with $L = \frac{1}{2}\sum_{k,l}(\mathfrak{M}_k - \mathfrak{M}_k^q)C_{kl}^{-1}(\mathfrak{M}_l - \mathfrak{M}_l^q)$.
- $\mathfrak{M}_k^q$ are the Ioffe-time data arising from inserting the test function $q$ into the cosine Fourier trafo and $C_{kl}$ denotes the covariance matrix of the $N_m$ measurements of simulation data $\mathfrak{M}_k^h$.
- We then specify an appropriate prior probability $P[q|I] = \exp[\alpha S[I]]$.
- Prior information enters in two ways here. On the one hand we deploy a particular functional form of the regularization functional

$$S_{BR}[q,m] = \sum_n \Delta x_n \Big(1 - \frac{q_n}{m_n} + \log\big(\frac{q_n}{m_n}\big)\Big)$$

which may be obtained by requiring positive definiteness of the resulting $q$, smoothness of $q$.

# Bayesian Reconstruction

- For sampled data, due to the central limit theorem, the likelihood probability may be written as the quadratic distance functional $P[\mathfrak{M}|q,I] = \exp[-L]$ with $L = \frac{1}{2}\sum_{k,l}(\mathfrak{M}_k - \mathfrak{M}_k^q)C_{kl}^{-1}(\mathfrak{M}_l - \mathfrak{M}_l^q)$.
- $\mathfrak{M}_k^q$ are the Ioffe-time data arising from inserting the test function $q$ into the cosine Fourier trafo and $C_{kl}$ denotes the covariance matrix of the $N_m$ measurements of simulation data $\mathfrak{M}_k^h$.
- We then specify an appropriate prior probability $P[q|I] = \exp[\alpha S[I]]$.
- Prior information enters in two ways here. On the one hand we deploy a particular functional form of the regularization functional

$$S_{BR}[q,m] = \sum_n \Delta x_n \Big(1 - \frac{q_n}{m_n} + \log\big(\frac{q_n}{m_n}\big)\Big)$$

which may be obtained by requiring positive definiteness of the resulting $q$, smoothness of $q$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.

# Bayesian Reconstruction

- What happens in the case of non-guaranteed positive definiteness?
- We need to change the regulator!
- Often the quadratic regulator is used

$$S_{QDR}[q,m] = \sum_n \Delta x_n \Big(q_n - m_n\Big)^2$$

- It is a comparatively strong regulator and usually imprints the form of the default model significantly onto the end result.
- Trying to keep the influence of the default model to a minimum, we extend the BR prior to non-positive functions.

$$S_{BRg}[q,m] = \sum_n \Delta x_n \Big( - \frac{|q_n - m_n|}{h_n} + \log\big(\frac{|q_n - m_n|}{h_n} - 1\big)\Big)$$

keeping the advantageous properties of the original BR prior at the price of having to introduce another default model related function $h$.

# Bayesian Reconstruction

- What happens in the case of non-guaranteed positive definiteness?
- We need to change the regulator!
- Often the quadratic regulator is used

$$S_{QDR}[q,m] = \sum_n \Delta x_n \Big( q_n - m_n \Big)^2$$

- It is a comparatively strong regulator and usually imprints the form of the default model significantly onto the end result.
- Trying to keep the influence of the default model to a minimum, we extend the BR prior to non-positive functions.

$$S_{BRg}[q,m] = \sum_n \Delta x_n \Big( -\frac{|q_n - m_n|}{h_n} + \log(\frac{|q_n - m_n|}{h_n} - 1) \Big)$$

keeping the advantageous properties of the original BR prior at the price of having to introduce another default model related function $h$.

# Bayesian Reconstruction

- What happens in the case of non-guaranteed positive definiteness?
- We need to change the regulator!
- Often the quadratic regulator is used

$$S_{QDR}[q,m] = \sum_n \Delta x_n \Big(q_n - m_n\Big)^2$$

- It is a comparatively strong regulator and usually imprints the form of the default model significantly onto the end result.
- Trying to keep the influence of the default model to a minimum, we extend the BR prior to non-positive functions.

$$S_{BRg}[q,m] = \sum_n \Delta x_n \Big( - \frac{|q_n - m_n|}{h_n} + \log(\frac{|q_n - m_n|}{h_n} - 1) \Big)$$

keeping the advantageous properties of the original BR prior at the price of having to introduce another default model related function $h$.

# Bayesian Reconstruction

- What happens in the case of non-guaranteed positive definiteness?
- We need to change the regulator!
- Often the quadratic regulator is used

$$S_{QDR}[q,m] = \sum_n \Delta x_n \Big( q_n - m_n \Big)^2$$

- It is a comparatively strong regulator and usually imprints the form of the default model significantly onto the end result.
- Trying to keep the influence of the default model to a minimum, we extend the BR prior to non-positive functions.

$$S_{BRg}[q,m] = \sum_n \Delta x_n \Big( -\frac{|q_n - m_n|}{h_n} + \log\big(\frac{|q_n - m_n|}{h_n} - 1\big) \Big)$$

keeping the advantageous properties of the original BR prior at the price of having to introduce another default model related function $h$.

# Bayesian Reconstruction

- What happens in the case of non-guaranteed positive definiteness?
- We need to change the regulator!
- Often the quadratic regulator is used

$$S_{QDR}[q,m] = \sum_n \Delta x_n \Big( q_n - m_n \Big)^2$$

- It is a comparatively strong regulator and usually imprints the form of the default model significantly onto the end result.
- Trying to keep the influence of the default model to a minimum, we extend the BR prior to non-positive functions.

$$S_{BRg}[q,m] = \sum_n \Delta x_n \Big( - \frac{|q_n - m_n|}{h_n} + \log\big(\frac{|q_n - m_n|}{h_n} - 1\big)\Big)$$

keeping the advantageous properties of the original BR prior at the price of having to introduce another default model related function $h$.

# Bayesian Reconstruction

- once $L$, $S$ and $m$ have been provided, the most probable PDF $q$, given simulation data and prior information is obtained by numerically finding the extremum of the posterior

$$\left.\frac{\delta P[q|\mathfrak{M},I]}{\delta q}\right|_{q=q_{\mathrm{Bayes}}} = 0.$$

- It has been proven that if the regulator is strictly concave, as is the case for all the regulators discussed above, there only exists a single unique extremum in the space of functions $q$ on a discrete $\nu$ interval.
- With positive definiteness is imposed on the end result, the space of admissible solutions is significantly reduced. Regulators admitting also $q$ functions with negative contributions have to distinguish between a multitude of oscillatory functions, which if integrated over, resemble a monotonous function to high precision. We will observe the emergence of ringing artefacts for the quadratic and generalized BR prior.

# Bayesian Reconstruction

- once $L$, $S$ and $m$ have been provided, the most probable PDF $q$, given simulation data and prior information is obtained by numerically finding the extremum of the posterior

$$\left.\frac{\delta P[q|\mathfrak{M},I]}{\delta q}\right|_{q=q_{\mathrm{Bayes}}} = 0.$$

- It has been proven that if the regulator is strictly concave, as is the case for all the regulators discussed above, there only exists a single unique extremum in the space of functions $q$ on a discrete $\nu$ interval.
- With positive definiteness is imposed on the end result, the space of admissible solutions is significantly reduced. Regulators admitting also $q$ functions with negative contributions have to distinguish between a multitude of oscillatory functions, which if integrated over, resemble a monotonous function to high precision. We will observe the emergence of ringing artefacts for the quadratic and generalized BR prior.

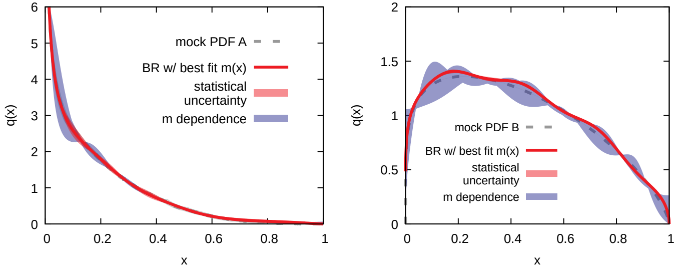
# Bayesian Reconstruction

- once $L$, $S$ and $m$ have been provided, the most probable PDF $q$, given simulation data and prior information is obtained by numerically finding the extremum of the posterior

$$\left.\frac{\delta P[q|\mathfrak{M}, I]}{\delta q}\right|_{q=q_{\text{Bayes}}} = 0.$$

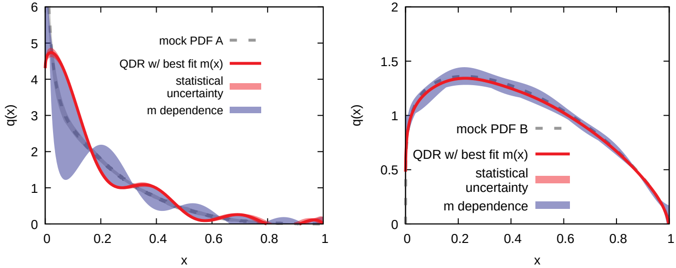
- It has been proven that if the regulator is strictly concave, as is the case for all the regulators discussed above, there only exists a single unique extremum in the space of functions $q$ on a discrete $\nu$ interval.
- With positive definiteness is imposed on the end result, the space of admissible solutions is significantly reduced. Regulators admitting also $q$ functions with negative contributions have to distinguish between a multitude of oscillatory functions, which if integrated over, resemble a monotonous function to high precision. We will observe the emergence of ringing artefacts for the quadratic and generalized BR prior.

# Bayesian Reconstruction

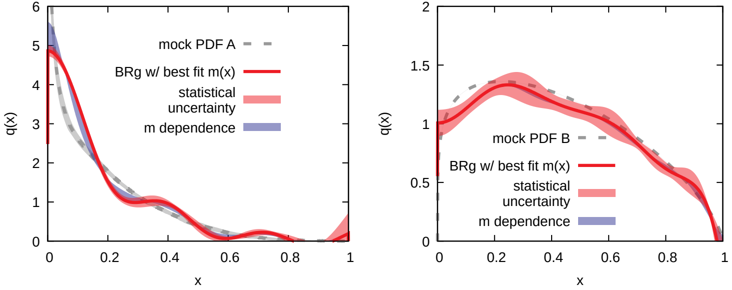



$x$-space PDF's reconstructed using the BR method from $N_\nu = 10$ Ioffe-time data points on the interval $\nu = [0, 20]$ The plots in the left column denote the results for mock data based on a phenomenological PDF (`NNPDF31_nnlo_as_0118`), while the right column arises from a scenario where $q(0)$ is finite.

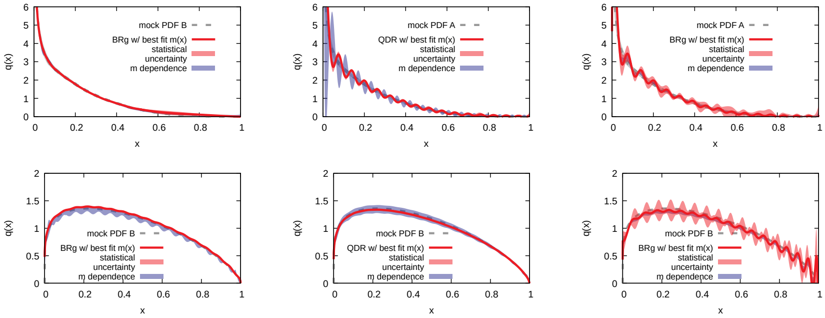
# Bayesian Reconstruction




$x$-space PDF's reconstructed using a quadratic prior Bayesian (QDR) method from $N_\nu = 10$ Ioffe-time data points on the interval $\nu = [0, 20]$. The plots in the left column denote the results for mock data based on a phenomenological PDF (`NNPDF31_nnlo_as_0118`), while the right column arises from a scenario where $q(0)$ is finite.




$x$-space PDF's reconstructed using the generalized Bayesian reconstruction (BRg) method from $N_\nu = 10$ Ioffe-time data points on the interval $\nu = [0, 20]$. The plots in the left column denote the results for mock data based on a phenomenological PDF (`NNPDF31_nnlo_as_0118`), while the right column arises from a scenario where $q(0)$ is finite.




$x$-space PDF's reconstructed in a best case scenario $(\nu = [0, 100], N_\nu = 100))$ using (left) the BR method (center) the quadratic prior and (right) the generalized BR method. The input data again is the one from a (top) $N_\nu = 100$ discretized Ioffe-time realistic PDF , while the bottom row arises from a scenario where $q(0)$ is finite.

# Backus-Gilbert Reconstruction

- The Backus-Gilbert (BG) method instead of imposing a smoothing condition on the resulting PDF $q(x)$ it imposes a minimization condition on the variance of the resulting function. G. Backus and F. Gilbert. Geophysical Journal of the Royal Astronomical Society, 16:169205, (1968)
- Let us define a rescaled kernel and rescaled PDF $h(x)$

$$K_j(x) \equiv cos(\nu_j x)p(x) \;\text{ and}\,,\, h(x) \equiv \frac{q_v(x)}{p(x)}$$

- where $p(x)$ corresponds to an appropriately chosen function that makes the problem easier to solve.
- We wish to incorporate into $p(x)$ most of the non-trivial structure of $q(x)$ apriorily, such that h(x) is a slowly varying function of $x$ and contains only the deviation of $q(x)$ from $p(x)$.

# Backus-Gilbert Reconstruction

- Starting from the preconditioned expression with a rescaled PDF $h(x)$ that is only a slowly varying function of $x$ our inverse problem becomes

$$d_j \equiv \mathfrak{M}_R(\nu_j) = \int_0^1 dx K_j(x) h(x) \,.$$

- BG introduces a function $\Delta(x - \bar{x}) = \sum_j q_j(\bar{x}) K_j(x)$, where $q_j(\bar{x})$ are unknown functions to be determined.
- It then estimates the unknown function as a linear combination of the data

$$\hat{h}(\bar{x}) = \sum_j q_j(\bar{x}) d_j \,, \text{or } \hat{q}_v(\bar{x}) = \sum_j q_j(\bar{x}) d_j p(\bar{x})$$

- If $\Delta(x - \bar{x})$ were to be a $\delta-$function then $\hat{h}(\bar{x}) = h(\bar{x})$. If $\Delta(x - \bar{x})$ approximates a $\delta$-function with a width $\sigma$, then the smaller $\sigma$ is the better the approximation of $\hat{h}(x)$ to $h(x)$.

# Backus-Gilbert Reconstruction

- In other words if $\hat{h}_\sigma(x)$ is the approximation resulting from $\Delta(x)$ with a width $\sigma$ then $\lim_{\sigma\to 0}\hat{h}_\sigma(x) = h(x)$.
- With this in mind BG minimizes the width $\sigma$ given by

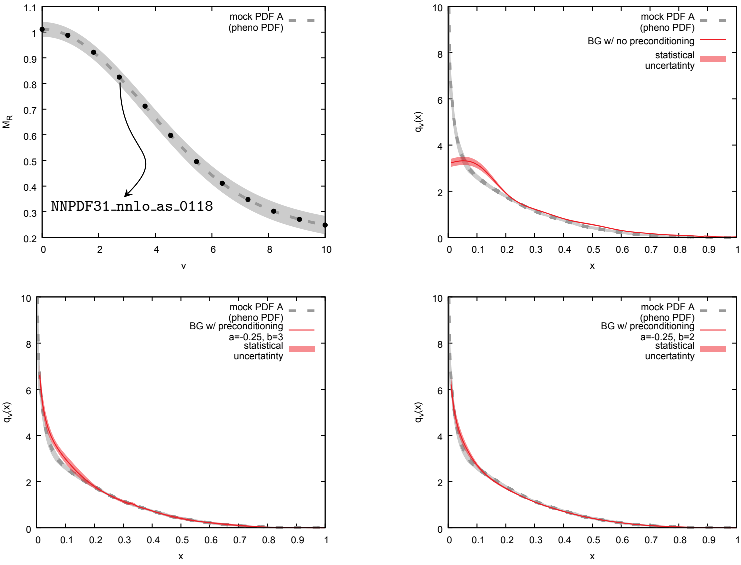
$$\sigma = \int_0^1 dx(x-\bar{x})^2\Delta(x-\bar{x})^2\,.$$

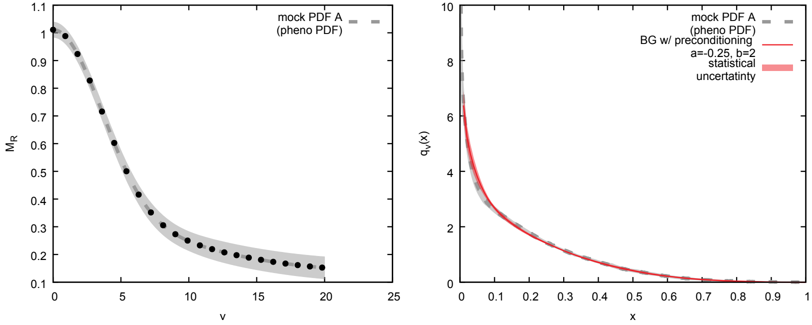
- Furthermore, if $\Delta(x)$ approximates a $\delta$-function then one has to impose the constraint $\int_0^1 dx\,\Delta(x-\bar{x}) = 1$. Using a Lagrange multiplier $\lambda$ one can minimize the width and impose the constraint by minimizing

$$\chi[q] = \int_0^1 dx(x-\bar{x})^2\sum_{j,k} q_j(\bar{x})K_j(x)K_k(x)q_k(\bar{x}) + \lambda\int_0^1 dx\sum_j K_j(x)q_j(\bar{x})\,.$$

- But let's see all this in practise ...

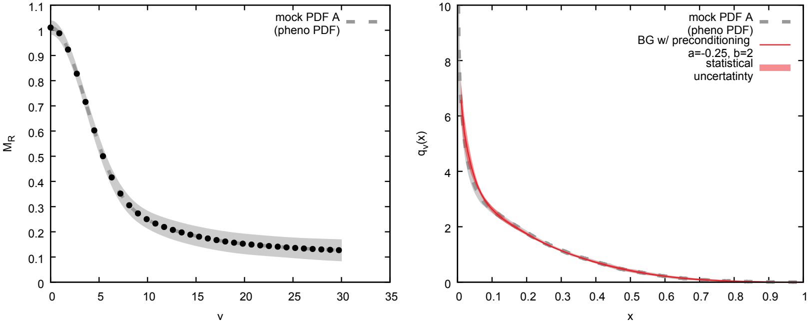
# Backus-Gilbert reconstruction





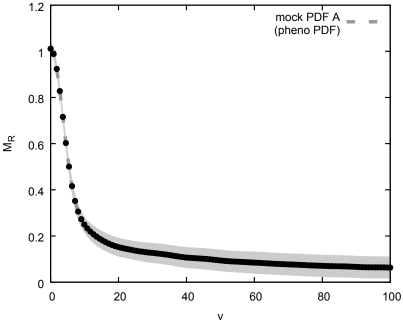

**Left**: The `NNPDF31_nnlo_as_0118` Ioffe time PDF data points used in this example, together with the dashed curve from which the data are chosen.
**Right**: The reconstructed Backus-Gilbert reconstructed PDF (red) together with the original PDF from the `NNPDF31_nnlo_as_0118` dataset (blue) with $b = 2$ and $\nu_{max} = 20$.

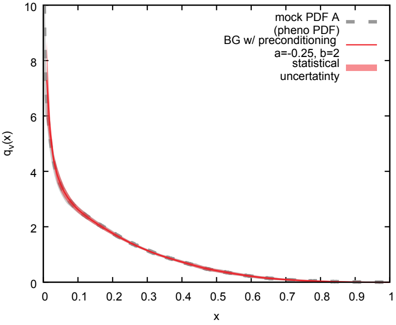
# Backus-Gilbert Reconstruction




**Left**: The NNPDF31_nnlo_as_0118 Ioffe time PDF data points used in this example, together with the dashed curve from which the data are chosen.
**Right**: The reconstructed Backus-Gilbert reconstructed PDF (red) together with the original PDF from the NNPDF31_nnlo_as_0118 dataset (blue) with $b = 2$ and $\nu_{max} = 30$.

# Backus-Gilbert Reconstruction

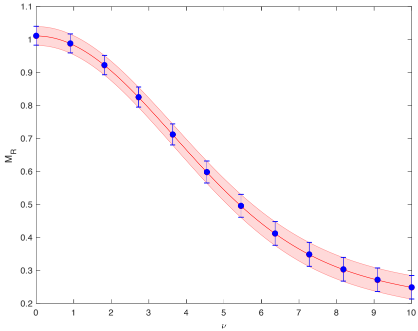


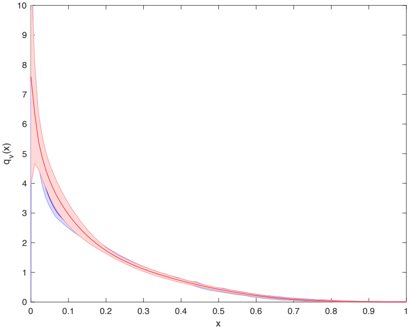




**Left**: The `NNPDF31_nnlo_as_0118` Ioffe time PDF data points used in this example, together with the dashed curve from which the data are chosen.
**Right**: The reconstructed Backus-Gilbert reconstructed PDF (red) together with the original PDF from the `NNPDF31_nnlo_as_0118` dataset (blue) with $b = 2$ and $\nu_{max} = 100$.

# Neural Network Reconstruction

VERY PRELIMINARY RESULTS!!!

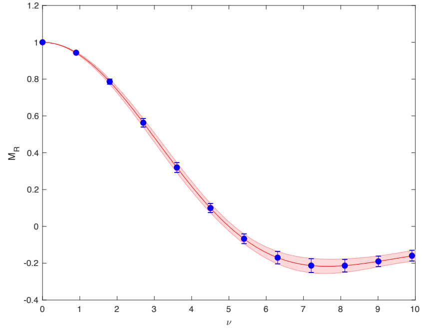


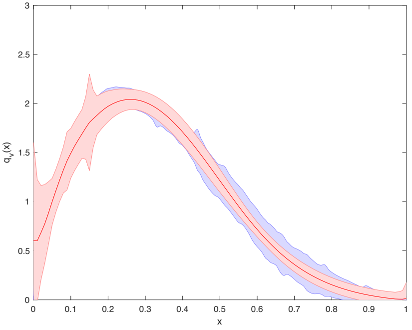




**Left**: Original data points (red) not visible. Red band representing errors on the original data points. Reconstructed data points (blue). **Right**: Original PDF (blue). Reconstructed PDF (red).

# Neural Network Reconstruction

VERY PRELIMINARY RESULTS!!!




**Left**: Original data points (red) not visible. Red band representing errors on the original data points. Reconstructed data points (blue). **Right**: Original PDF (blue). Reconstructed PDF (red).

# Conclusions-Outlook

- We studied in detail the problem of PDF reconstructions out of Ioffe time data
- An extremely ill-defined problem due to restricted range and number of $\nu$ data.
- We showed how methods of advanced reconstruction that have been successfully applied to different inverse problems in LQCD can also become handy for this task.
- We stressed the necessity of additional info in order to be able to provide a unique answer.
- These methods would be key ingredients of future studies.
- **Many thanks for your attention!!!**

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

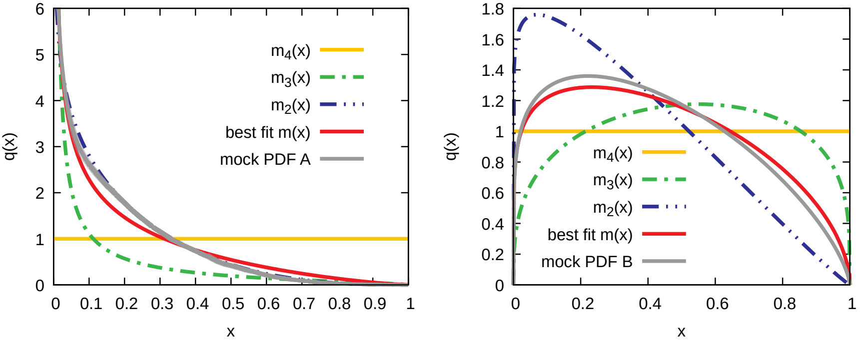
- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

# Bayesian Reconstruction

- The functional $S$ depends on the function $m$, the default model.
- By construction constitutes its unique extremum.
- In the Bayesian logic $m$ is the correct result for $q$ in the absence of any data.
- We select $m$ by a best fit of the Ioffe-PDF data and we will vary it to get a handle on systematics.
- In the definition of $P[q|I]$ we introduced a further parameter $\alpha$, a so called hyperparameter
- Weighs the influence of simulation data and prior information. It has to be taken care of self-consistently.
- In the Maximum Entropy Method $\alpha$ is selected, such that the evidence has an extremum. In the BR method we deploy here, we marginalize the parameter $\alpha$ apriori, i.e. we integrate the posterior w.r.t the hyperparameter, assuming complete ignorance of its values $P[\alpha] = 1$.

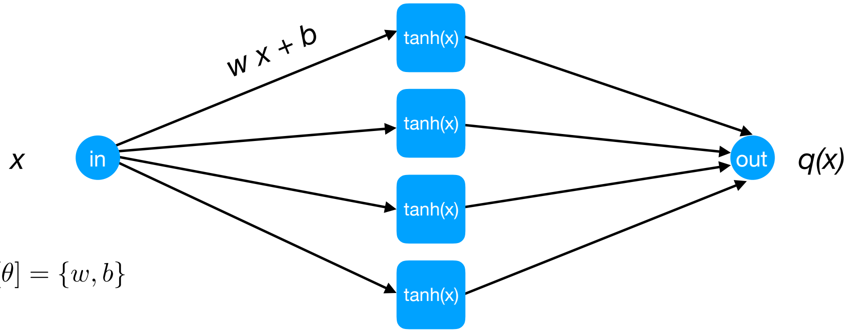



Best fit PDF (red solid line) from (left) realistic PDF data [mock scenario A] and (right) from a PDF Ansatz $q(x) = p(x, \frac{2}{10}, \frac{7}{10})$ [mock scenario B]. The actual mock PDF in both cases is given as gray solid line. To determine the dependence of our results on the choice of default model, three further choices for $m$ are plotted, two arising from varying the best fit parameters by factors of 2, one being the constant default model m=1.

# Neural Network Reconstruction

- The ensemble average of data is obtained in two steps
  - Starting from random $[w, b]$, minimize $\chi^2$ to find $[w, b]$.
  - Repeat 10 times to find 10 different Neural Nets (replicas).
- For each Neural Net, the minimizer is re-run for each jackknife sample to obtain a jackknife estimate $q(x)$ for each replica.
- The central value of $q(x)$ is estimated as the average over jackknife samples and replicas.
- The error is estimated by combining the fluctuations over the jackknife sample and replicas.

# Neural Network Reconstruction




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

$$\chi^2 = \sum_k \left( M(\nu_k) - \int_0^1 \mathrm{d}x\, q_{[\theta]}(x) \cos(\nu_k x) \right) \sigma_k^2 \left( M(\nu_k) - \int_0^1 \mathrm{d}x\, q_{[\theta]}(x) \cos(\nu_k x) \right)$$

$$\min_{[\theta]} [\chi^2] \to [w, b]$$