# Parton pseudo distribution functions from Lattice QCD

**Savvas Zafeiropoulos**

CPT Marseille

31.10.2019

EINN 2019

Distribution functions: Lattice QCD meets phenomenology

Paphos, Cyprus

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

# PDFs from the lattice: Pseudo-PDFs Formalism

Starting point: the equal time hadronic matrix element with the quark and anti-quark fields separated by a finite distance Radyushkin (2017)

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

$$z = (0,0,0,z_3) \quad p = (p^0,0,0,p) \quad \alpha = 0$$

$$\xrightarrow{\text{Lorentz inv.}} \mathcal{M}^\alpha(z,p) = 2p^\alpha \underbrace{\mathcal{M}_p(-(zp),-z^2)}_{\textbf{Leading twist}} + \cancel{z^\alpha \underbrace{\mathcal{M}_z(-(zp),-z^2)}_{\textbf{Higher twist}}}$$

- The Lorentz invariant quantity $\nu = -(zp)$, is the "Ioffe time"
- 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 $f(x,\mu^2)$ Balitsky, Braun (1988), Braun et al. (1995)

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

# 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 Radyushkin (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
- All UV singularities are exactly cancelled and when computed in lattice QCD it can be extrapolated to the continuum limit at fixed $\nu$ and $z^2$.

# Numerical implementation

First case study in an unphysical setup Karpie, Orginos, Radyushkin SZ, Phys.Rev. D96 (2017) no.9, 094503

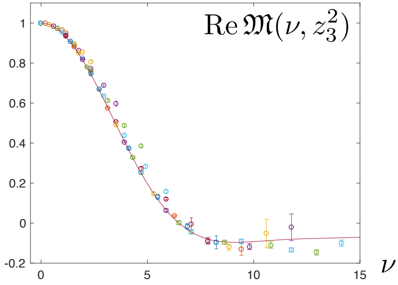
- Quenched approximation
- $32^3 \times 64$ lattices with $a = 0.093$fm.
- $m_\pi = 601$MeV and $m_N = 1411$MeV

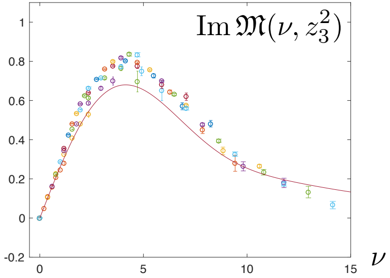
Now employing dynamical ensembles

| $a$(fm) | $M_\pi$(MeV) | $\beta$ | $L^3 \times T$ |
|---|---|---|---|
| 0.127(2) | 415 | 6.1 | $24^3 \times 64$ |
| 0.127(2) | 415 | 6.1 | $32^3 \times 96$ |
| 0.094(1) | 390 | 6.3 | $32^3 \times 64$ |
| 0.094(1) | 280 | 6.3 | $32^3 \times 64$ |
| 0.094(1) | 172 | 6.3 | $64^3 \times 128$ |

Table: Parameters for the lattices generated by the JLab/W&M collaboration using 2+1 flavors of clover Wilson fermions and a tree-level tadpole-improved Symanzik gauge action. The lattice spacings, $a$, are estimated using the Wilson flow scale $w_0$. Stout smearing implemented in the fermion action makes the tadpole corrected tree-level clover coefficient $c_{\mathrm{SW}}$ used, to be very close to the value determined non-pertubatively with the Schrödinger functional method

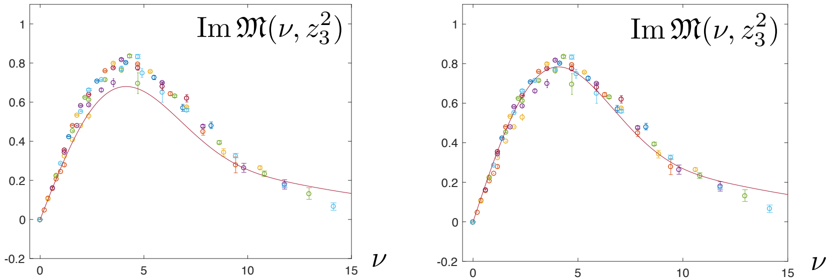
# Results for the Re and Im parts of $\mathfrak{M}(\nu, z_3^2)$



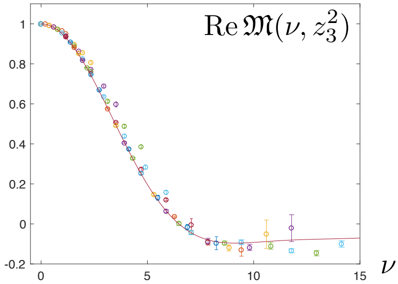

- Curves represent Re and Im Fourier transforms of $q_v(x) = \frac{315}{32}\sqrt{x}(1-x)^3$.
- Considering CP even and odd combinations
  - even: $q_-(x) = f(x) + f(-x) = q(x) - \bar{q}(x) = q_v(x)$
  - odd: $q_+(x) = f(x) = f(-x) = q(x) + \bar{q}(x) = q_v(x) + 2\bar{q}(x)$

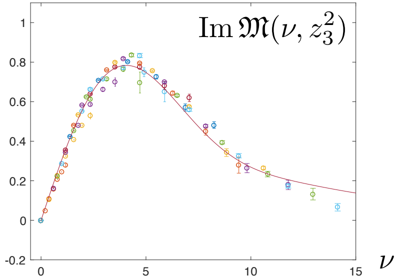
# Results for the Im part of $\mathfrak{M}(\nu, z_3^2)$




- Curves represent the Im Fourier transforms of $q_v(x) = q(x) - \bar{q}(x)$ and $q_+(x) = q(x) + \bar{q}(x) = q_v(x) + 2\bar{q}(x)$ respectively.
- The agreement with the data is strongly improved if we use a non-vanishing antiquark contribution, namely $\bar{q}(x) = \bar{u}(x) + \bar{d}(x) = 0.07[20x(1-x)^3]$.

# Results for the Re and Im parts of $\mathfrak{M}(\nu, z_3^2)$




- Data as function of the Ioffe time. A residual $z_3$-dependence can be seen.
- This is more visible when, for a particular $\nu$ we have several data points corresponding to different values of $z_3$.
- Different values of $z_3^2$ for the same $\nu$ correspond to the Ioffe time distribution at different scales.

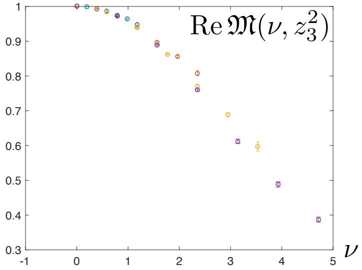
# Residual $z_3$-dependence

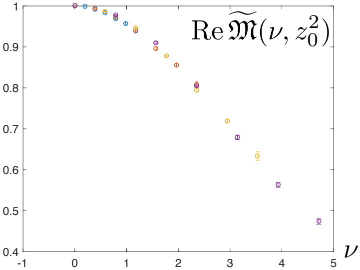
- Is the residual scatter in the data points consistent with evolution? By solving the evolution equation at LO, the Ioffe time PDF at $z_3'$ is related to the one at $z_3$ by

$$\mathfrak{M}(\nu, {z'}_3^2)=\mathfrak{M}(\nu, z_3^2) \ - \frac{2}{3}\frac{\alpha_s}{\pi}\ln({z_3'}^2/z_3^2)\int_0^1 du\, B(u)\, \mathfrak{M}\,(u\nu, z_3^2)$$

- Only applicable at small $z_3$

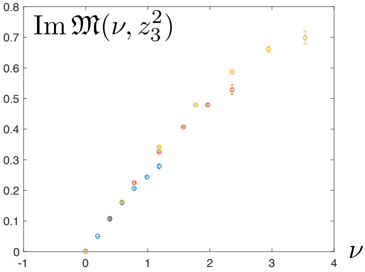
# Before and after evolution

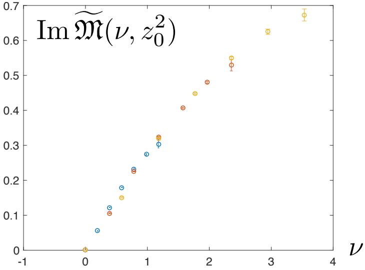


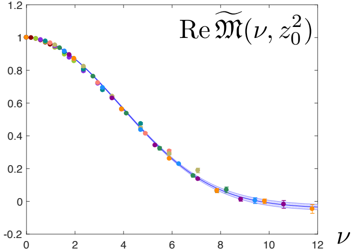

The ratio $\mathfrak{M}(\nu, z_3^2)$ for for $z_3/a = 1, 2, 3,$ and 4. **LHS:** Data before evolution. **RHS:** Data after evolution. The reduction in scatter indicates that evolution collapses all data to the same universal curve.

# Before and after evolution




The ratio $\mathfrak{M}(\nu, z_3^2)$ for for $z_3/a = 1, 2, 3,$ and 4. **LHS:** Data before evolution. **RHS:** Data after evolution. The reduction in scatter indicates that evolution collapses all data to the same universal curve.




- Evolved points fitted with cosine FT of

$$q_v(x) = N(a,b)\; x^a\; (1-x)^b$$

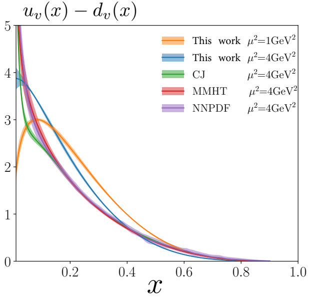
$$a = 0.36(6), \qquad b = 3.95(22)$$

Karpie, Orginos, Radyushkin, S.Z. (2017)

- Evolved data can be exploited to build

$$u_v(x) - d_v(x)$$

- Results compared with predictions from global fits

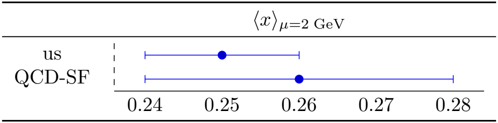


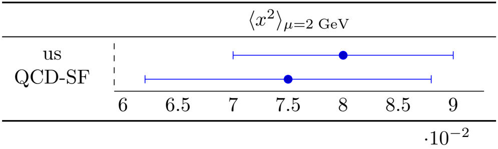
# Sanity checks *vs* other lattice results

- Extract lowest PDF moments from our data Karpie, Orginos, S.Z., JHEP 1811 (2018) and compare with the lattice literature QCD-SF collaboration (1996)
- $\overline{MS}$ moments up to $\mathcal{O}(\alpha_s^2, z^2)$ directly from the reduced function $\mathfrak{M}(\nu, z^2)$

$$a_{n+1}(\mu) = (-i)^n \frac{1}{c_n(z^2\mu^2)} \left. \frac{\partial^n \mathfrak{M}(\nu, z^2)}{\partial \nu^n} \right|_{\nu=0} + \mathcal{O}(z^2, \alpha_s^2)$$

- Our method avoids mixing and allows the extraction of any moment



# The pertinent systematics in PDF extraction

- Parton distribution functions or distribution amplitudes may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position-space matrix elements
- One example are the Ioffe-time PDFs, $\mathfrak{M}_R$, related to the physical PDF $q_v(x, \mu^2)$ via the integral relation

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

Karpie, Orginos, Rothkopf, S.Z. JHEP 1904 (2019) 057

# The pertinent systematics in PDF extraction

- Parton distribution functions or distribution amplitudes may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position-space matrix elements
- One example are the Ioffe-time PDFs, $\mathfrak{M}_R$, related to the physical PDF $q_v(x, \mu^2)$ via the integral relation

**Only a handful of lattice data**

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

Karpie, Orginos, Rothkopf, S.Z. JHEP 1904 (2019) 057

# The pertinent systematics in PDF extraction

- Parton distribution functions or distribution amplitudes may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position-space matrix elements
- One example are the Ioffe-time PDFs, $\mathfrak{M}_R$, related to the physical PDF $q_v(x,\mu^2)$ via the integral relation

**Only a handful of lattice data**

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

**Cosine not orthogonal in $[\mathbf{0},\mathbf{1}]$**

Karpie, Orginos, Rothkopf, S.Z. JHEP 1904 (2019) 057

# The pertinent systematics in PDF extraction

- Parton distribution functions or distribution amplitudes may be defined in lattice QCD by inverting the quasi-Fourier transform of a certain class of hadronic position-space matrix elements
- One example are the Ioffe-time PDFs, $\mathfrak{M}_R$, related to the physical PDF $q_v(x, \mu^2)$ via the integral relation

**Only a handful of lattice data**

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

**Cosine not orthogonal in $[\mathbf{0}, \mathbf{1}]$**

- The task at hand is then to reconstruct the PDF $q_v(x, \mu^2)$ given a limited set of simulated data for $\mathfrak{M}_R(\nu, \mu^2)$.
- The extraction is highly ill-posed, so one has to resort to regularization strategies in order to find a way to reliably estimate the PDF from the data at hand

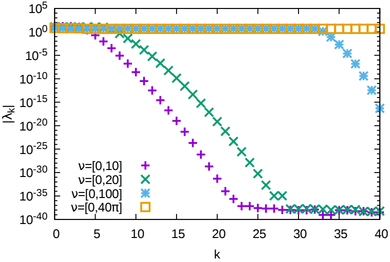
Karpie, Orginos, Rothkopf, S.Z. JHEP 1904 (2019) 057

# Naive Reconstruction

- Discretize the integral, employing the trapezoid rule

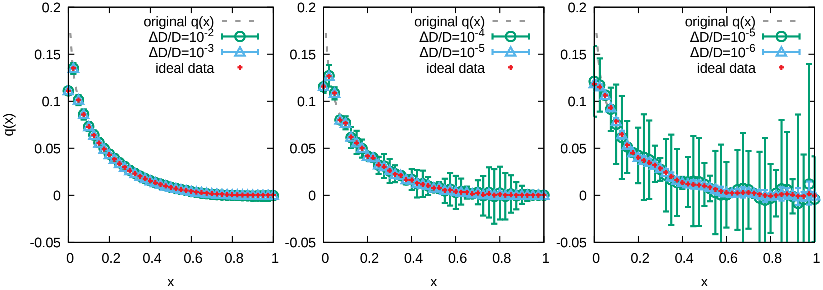
$$\mathfrak{M}_R(\nu) = \frac{1}{2}\cos(\nu x_0)\, q_v(x_0) + \sum_{k=1}^{N_x-1} \delta x\, \cos(\nu x_k)\, q_v(x_k) + \frac{1}{2}\cos(\nu x_{N_x})\, q_v(x_{N_x})$$

- Casting our problem 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 $\mathfrak{C}$.




Karpie, Orginos, Rothkopf, S.Z. - arXiv:1901.05408 - JHEP 1904 (2019) 057

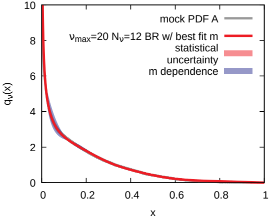
# Naive Reconstruction



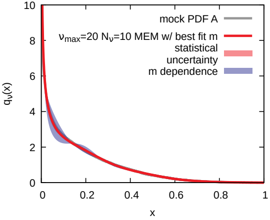

**Results for the direct inversion for different discretization intervals** (left $\nu = [0, 40\pi]$, center $\nu = [0, 100]$, right $\nu = [0, 20]$).

# Advanced PDF Reconstructions

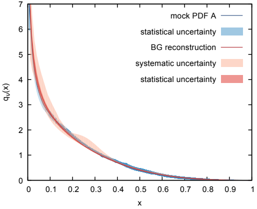
**Bayesian Reconstruction**



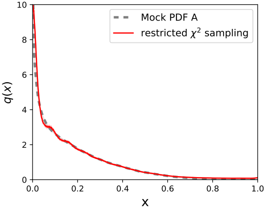

**Max. Entropy Method**




**Backus-Gilbert algorithm**

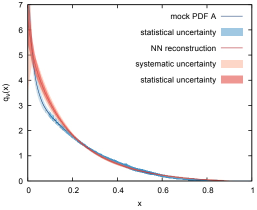



**HMC $\chi^2$ evaluation**



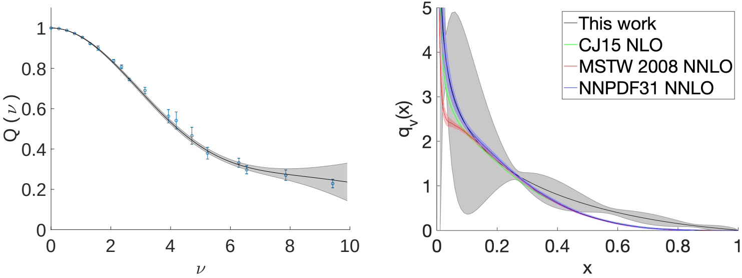

Capitalize of the good scanning in Ioffe time and use advanced reconstruction methods to extract the maximum amount of information also for the small-$x$ region.

**Neural Network**

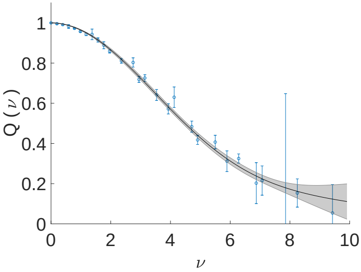


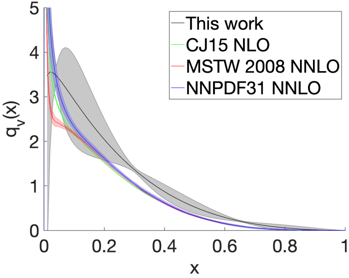

Karpie, Orginos, Rothkopf, S.Z. JHEP 1904 (2019) 057

# New results with $N_\mathrm{f} = 2 + 1$ fermions for the nucleon



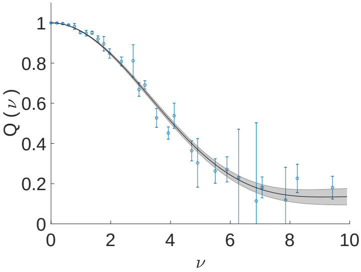

The nucleon valence distribution obtained from the ensemble $a127m415$ fit to the form used by the JAM collaboration. The $\chi^2/\mathrm{d.o.f.}$ for the fit with all the data is 2.5(1.5). The uncertainty band is obtained from the fits to the Jackknife samples of the data.

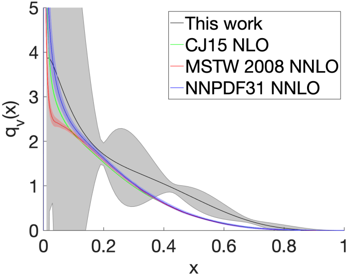



The nucleon valence distribution obtained from the ensemble $a127m415L$ fit to the form used by the JAM collaboration. The $\chi^2/\mathrm{d.o.f.}$ for the fit with all the data is 2.1(6). The uncertainty band is obtained from the fits to the Jackknife samples of the data.

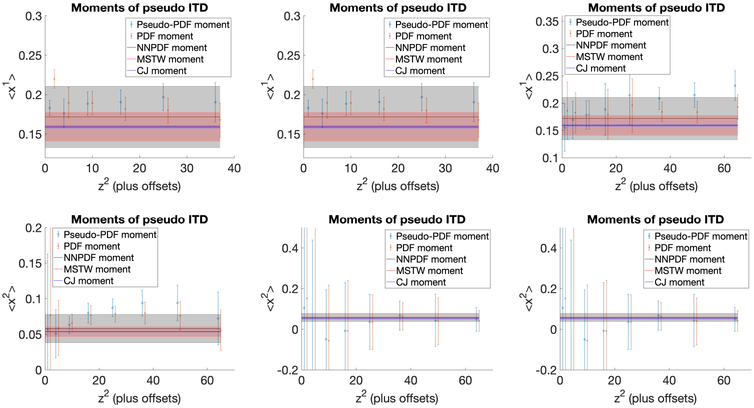
# New results with $N_\mathrm{f} = 2 + 1$ dynamical fermions



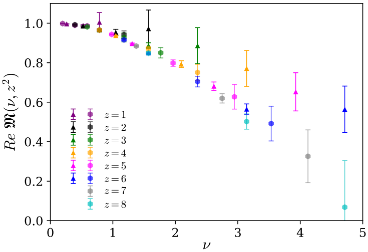



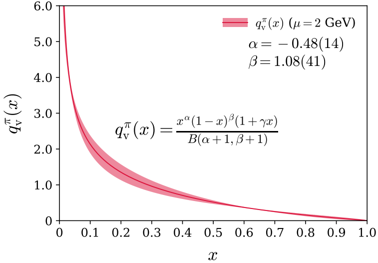

The nucleon valence distribution obtained from the ensemble $a094m390$ fit to the form used by the JAM collaboration. The $\chi^2/\mathrm{d.o.f.}$ for the fit with all the data is 2.0(5). The uncertainty band is obtained from the fits to the Jackknife samples of the data.

# New results with $N_f = 2 + 1$ fermions for the nucleon



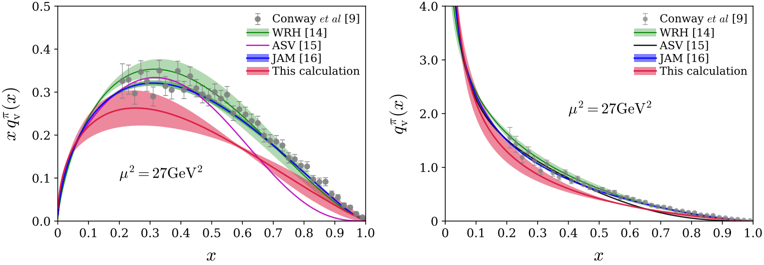
# Results with $N_\mathrm{f} = 2+1$ fermions for the pion



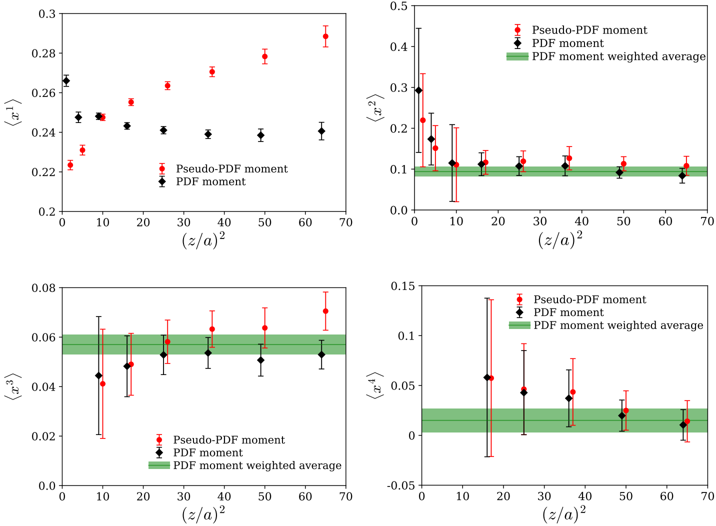

ITD obtained from ensembles $a127m415$ and $a127m415L$ after 1-loop perturbative matching at $\mu = 2$ GeV. The circle ($\circ$) symbols indicate the reduced pseudo-ITD matrix elements $M^0$ extracted from the $a127m415$ ensemble and the diamond ($\diamond$) symbols denote those for the $a127m415L$ ensemble. The red band is obtained from a simultaneous fit to the matched ITDs on these two ensembles in the limit of infinite volume (LHS). The pion valence distribution (RHS).

# Results with $N_\mathrm{f} = 2 + 1$ flavors for the pion



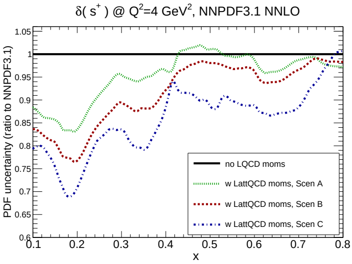

(LHS) Comparison of the pion $xq_\mathrm{v}^\pi(x)$-distribution with the LO extraction from DY data (gray data points), NLO fits (green, maroon, and blue). This lattice QCD calculation of $q_\mathrm{v}^\pi(x)$ is evolved from an initial scale $\mu_0^2 = 4$ GeV$^2$ at NLO. All the results are evolved to an evolution scale of $\mu^2 = 27$ GeV$^2$. Similar comparison of the pion $q_\mathrm{v}^\pi(x)$-distribution (RHS).

# Results with $N_\mathrm{f} = 2 + 1$ flavors for the pion

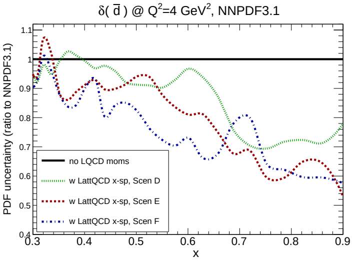


# Lattice impact in the precision of global fits

**$s^+$ quark PDFs, percentage PDF uncertainty**




**$\bar{d}$ quark PDFs, percentage PDF uncertainty**




- The comparison is with the results of including lattice-QCD pseudo-data for moments of PDFs (left) and info from lattice-QCD pseudo-data on x-space PDFs (right). Lin et al. (2018), Cichy, Del Debbio and Giani JHEP 1910 (2019) 137

# Take home messages

- Realization that $\mathcal{M}(\nu, z_3^2)$ should be treated as a function of Lorentz invariants $\nu$ and $z^2$.
- Only the $\nu$-dependence is responsible for the $x$-dependence of the PDFs.
- $\mu^2$ serves the same role as $z^2$ in the pseudo-PDF
- Don't mix $z^2$ and $\nu$.
- The ITpD $\mathcal{M}(\nu, z_3^2)$ can be directly matched to the light cone ITD $\mathcal{I}(\nu, \mu^2)$ with no intermediaries needed.
- Knowing $\mathcal{I}(\nu, \mu^2)$ means that we know $f(x, \mu^2)$.
- Small difference between the pseudo-ITD and the light cone ITD

# Take home messages

- Realization that $\mathcal{M}(\nu, z_3^2)$ should be treated as a function of Lorentz invariants $\nu$ and $z^2$.
- Only the $\nu$-dependence is responsible for the $x$-dependence of the PDFs.
- $\mu^2$ serves the same role as $z^2$ in the pseudo-PDF
- Don't mix $z^2$ and $\nu$.
- The ITpD $\mathcal{M}(\nu, z_3^2)$ can be directly matched to the light cone ITD $\mathcal{I}(\nu, \mu^2)$ with no intermediaries needed.
- Knowing $\mathcal{I}(\nu, \mu^2)$ means that we know $f(x, \mu^2)$.
- Small difference between the pseudo-ITD and the light cone ITD

# Take home messages

- Issues with the Fourier transform
- In the pseudo-approach the Fourier transform is in the end over $\nu$
- However a FT over $z$ could be a source of systematics.
- Extrapolating from $z_{\mathrm{m}ax}$ to $\infty$ and then carry out the integral?
- The renormalized m.e. have growing and growing errors at $z_{\mathrm{m}ax}$ saying that this is equal to zero is not a precision statement.
- pPDFs allow for a good control of higher twist effects
- We study the residual $z^2$ effects for the moments as well as the matched ITD and have a concrete idea of the size of these effects.
- If one carries the FT over $z$ instead you can not clearly control the higher twist effects because the integral over $z$ mixes the leading and higher twist regions.

# Take home messages

- Issues with the Fourier transform
- In the pseudo-approach the Fourier transform is in the end over $\nu$
- However a FT over $z$ could be a source of systematics.
- Extrapolating from $z_{\mathrm{m}ax}$ to $\infty$ and then carry out the integral?
- The renormalized m.e. have growing and growing errors at $z_{\mathrm{m}ax}$ saying that this is equal to zero is not a precision statement.
- pPDFs allow for a good control of higher twist effects
- We study the residual $z^2$ effects for the moments as well as the matched ITD and have a concrete idea of the size of these effects.
- If one carries the FT over $z$ instead you can not clearly control the higher twist effects because the integral over $z$ mixes the leading and higher twist regions.

# Take home messages

- No need for infinite $p$ for the pseudo-PDFs.
- Large $p$ is useful for good scanning in $\nu$ but no $p \to \infty$ is required.
- The ratio is an RGI quantity agnostic to the regulator
- Finite moments which make our goal twofold (extraction of moments beyond the $x$-dependence).

# Conclusions and outlook

- PDFs are needed as theoretical inputs to all hadron scattering experiments and in some cases are the largest theory uncertainty.
- The lattice community is by now able to provide ab-initio determinations of PDFs without theoretical obstructions.
- The interplay between lattice QCD and global fits is very important
- Also important in the search of New Physics Gao, Harland-Lang, Rojo (2018)
- What next? Polarized, Transversity, gluon PDFs and GPDs eventually
- **Many thanks for your attention!!!**

# Formalism

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

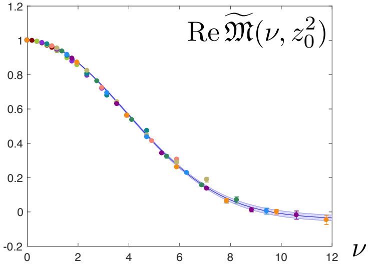
$$Q(x, p^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\nu \; e^{-ix\nu} \, \mathcal{M}_p(\nu, [\nu/p]^2)$$

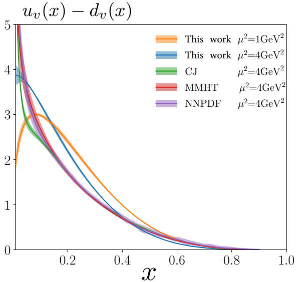
  Quasi PDF mixes invariant scales until $p_z$ is effectively large enough

- While the pseudo-PDF has fixed invariant scale dependence

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

# Comparison to global fits

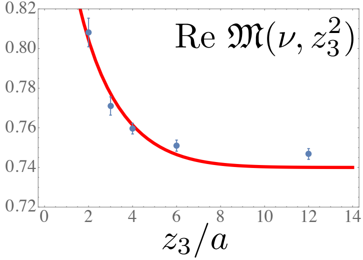



LHS: Data points for Re $\mathfrak{M}\,(\nu, z_3^2)$ with $z_3 \leq 10a$ evolved to $z_3 = 2a$. By fitting these evolved points with a cosine FT of $q_v(x) = N(a,b)x^a(1-x)^b$ we obtain $a = 0.36(6)$ and $b = 3.95(22)$ (statistical errors). RHS: Curve for $u_v(x) - d_v(x)$ built from the evolved data shown in the left panel and treated as corresponding to the $\mu^2 = 1$ GeV$^2$ scale; then evolved to the reference point $\mu^2 = 4$ GeV$^2$ of the global fits. 1-loop matching to $\overline{\mathrm{MS}}$ still to be done on our data

A. Radyushkin 1710.08813, Zhang et al 1801.03023, Izubuchi et al 1801.03917

# More on evolution




- LO evolution cannot be extended to very low scales.
- It is known that evolution stops below a certain scale (by observing our data we infer that this is the case for $z_3 \geq 6a$.)
- Adopt an evolution that leaves the PDF unchanged for length scales above $z_3 = 6a$ and use the leading perturbative evolution formula to evolve to smaller $z_3$ scales.

# Numerical implementation

Following C. Bouchard et.al Phys. Rev. D 96, no. 1, 014504 (2017) , we compute a regular nucleon two point function

$$C_p(t) = \langle \mathcal{N}_p(t)\overline{\mathcal{N}}_p(0)\rangle \,,$$

$$C_p^{\mathcal{O}^0(z)}(t) = \sum_\tau \langle \mathcal{N}_p(t)\mathcal{O}^0(z,\tau)\overline{\mathcal{N}}_p(0)\rangle$$

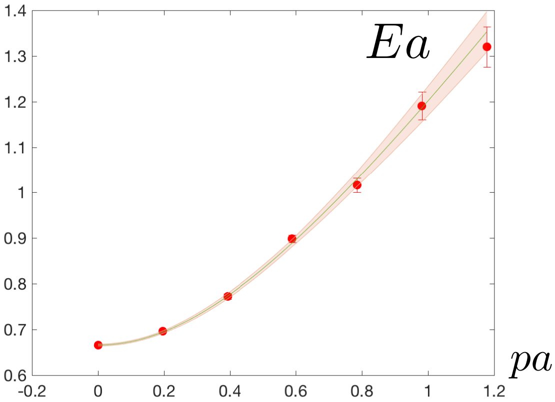
with $\quad \mathcal{O}^0(z,t) = \overline{\psi}(0,t)\gamma^0\tau_3\hat{E}(0,z;A)\psi(z,t)$

Proton momentum and displacement of the quark fields along the $\hat{z}$ axis

$$\mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t) = \frac{C_p^{\mathcal{O}^0(z)}(t+1)}{C_p(t+1)} - \frac{C_p^{\mathcal{O}^0(z)}(t)}{C_p(t)}$$

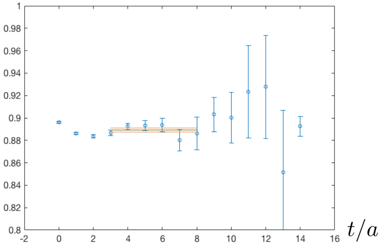
Extract the desired ME $\mathcal{J}$ at large Euclidean time separation as $\frac{\mathcal{J}(z_3 p, z_3^2)}{2p^0} = \lim_{t\to\infty} \mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t)$ , where $p^0$ is the energy of the nucleon.

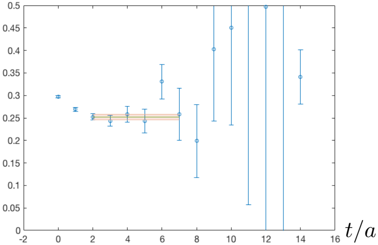
# Results for the nucleon dispersion relation




Energies and momenta are in lattice units. The solid line is the continuum dispersion relation (not a fit) while the errorband is an indication of the statistical error of the lattice nucleon energies

# Results

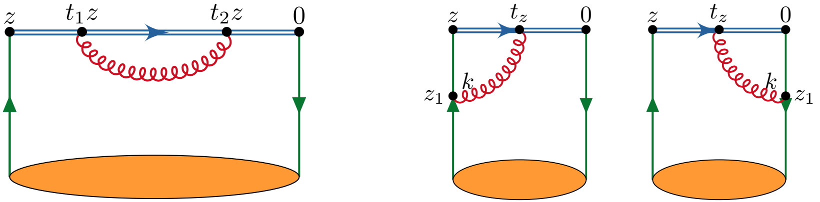



Typical fits used to extract the reduced matrix element (here $p = 2\pi/L \cdot 2$ and $z = 4$ (LHS) and $p = 2\pi/L \cdot 3$ and $z = 8$ (RHS)). The average $\chi^2$ per degree of freedom was $\mathcal{O}(1)$. All fits are performed with the full covariance matrix and the error bars are determined with the jackknife method.

# Renormalization

- In a series of articles Dotsenko Nucl.Phys. B169 (1980) 527, Ishikawa et al. Phys. Rev. D 96, 094019 (2017), Chen et al. Nucl.Phys. B915 (2017) and A. V. Radyushkin Phys.Lett. B781 (2018) 433-442 the one loop renormalizability of $\mathcal{M}^{\alpha}(z,p,a)$ has been discussed
- by analyzing the pertinent diagrams one can see that there is a linear divergence from the link self-energy contribution and a logarithmic divergence associated to the anomalous dimension $2\gamma_{\mathrm{end}}$ due to two end-points of the link.



# Renormalization

- $\mathcal{M}$ has been shown to renormalize multiplicatively as $\mathcal{M}_R(\nu, z^2, \mu) = Z_j^{-1} Z_{\bar{j}}^{-1} e^{-\delta m|z|} \mathcal{M}_B(\nu, z^2, a)$, where $\delta m = C_F \frac{\alpha_s}{2\pi} \frac{\pi}{a}$, is an effective mass counterterm removing power divergences in the Wilson line and $Z_j^{-1}, Z_{\bar{j}}^{-1}$ are renormalization constants (RCs) associated with the endpoints of the Wilson line independent of $z, p$.
- The entire renormalization is independent of the external momentum
- Forming the ratio, the RCs cancel and thus the reduced Ioffe time distribution has a great potential to reduce systematic effects related to renormalization.The UV divergences generated by the link-related and quark-self-energy diagrams cancel in the ratio.

# Numerical implementation

- Renormalization of the ME?
- For $z_3 = 0$ $\mathcal{M}(z_3 p, z_3^2) \rightarrow$ the local iso-vector current, should be $= 1$ (but ...) lattice artifacts...
- Introduce an RC $Z_p = \frac{1}{\mathcal{J}(z_3 p, z_3^2)\big|_{z_3=0}}$
- $Z_p$ has to be independent from $p$. But lattice artifacts or potential fitting systematics ...
- renormalize the ME for each momentum with its own $Z_p \rightarrow$ maximal statistical correlations to reduce statistical errors, and cancellation of lattice artifacts in the ratio

# Numerical implementation

- in practise use the double ratio

$$\mathfrak{M}(\nu, z_3^2) = \lim_{t\to\infty} \frac{\mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t)}{\mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t)\big|_{z_3=0}} \times \frac{\mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t)\big|_{p=0, z_3=0}}{\mathcal{M}_{\text{eff}}(z_3 p, z_3^2; t)\big|_{p=0}} \; ,$$

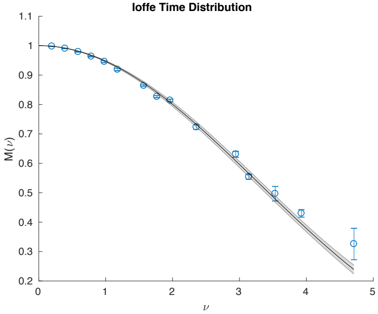
- infinite $t$ limit is obtained with a fit to a constant for a suitable choice of a fitting range.

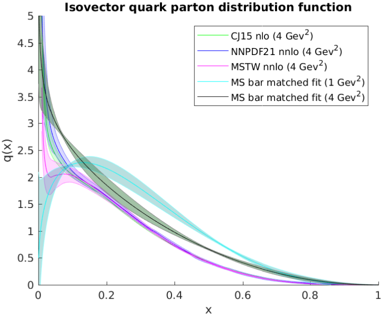
# Matching to $\overline{MS}$

- In 1801.02427 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 the $\overline{MS}$ scheme



# 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

- 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.

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

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

- 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

- 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 $P[\alpha] = 1$.

# 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

- 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

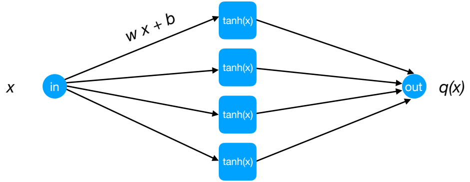
- once $L$, $S$ and $m$ have been provided, the most probable PDF $q$, given simulation data and prior information is obtained by numerically finding the extremum of the posterior

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

- 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.

# 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.
- Central value of $q(x)$ is the average over jackknife samples and replicas.
- Error by combining the fluctuations over the jackknife sample and replicas.




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

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

# Lattice QCD requirements

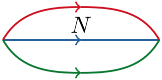
- Largest momentum on the lattice $aP_{\max} = \pi/2 \propto \mathcal{O}(1)$
- $a = 0.1$fm $\rightarrow P_{\max} = 10\Lambda$ where $\Lambda = 300$ MeV
- $a = 0.05$fm $\rightarrow P_{\max} = 20\Lambda$

Large momentum is required to suppress high twist effects (quasi-PDFs) and to provide a wide coverage of the Ioffe time $\nu$

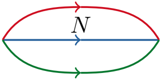
$P_{\max} = 3$ GeV easily achievable with moderate values of the lattice spacing but still demanding due to statistical noise

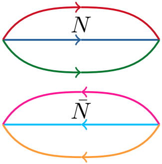
$P_{\max} = 6$ GeV exponentially harder requiring very fine values of the lattice spacing

# Signal to Noise

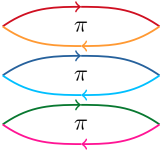



$$\langle C_N(t) \rangle \sim e^{-m_N t}$$




$$\langle |C_N(t)|^2 \rangle \sim e^{-3 m_\pi t}$$

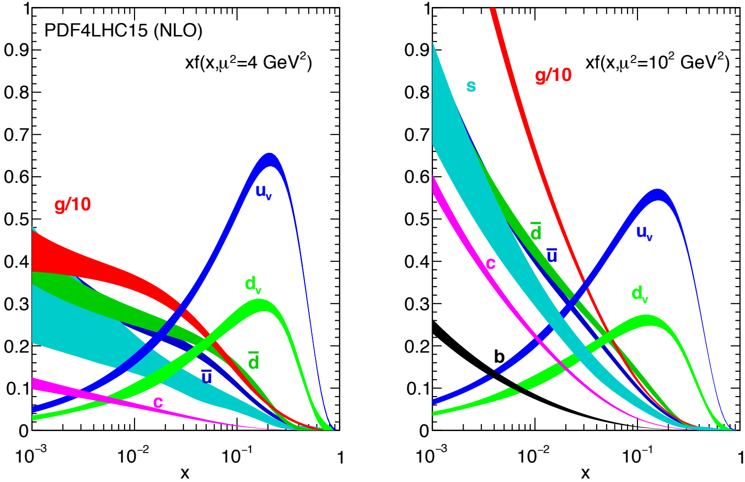



Statistical accuracy drops exponentially with increasing momentum $P$

$$\mathrm{StN}(O) = \frac{\langle O \rangle}{\sqrt{\mathrm{var}(O)}} \propto e^{-[E_N(P) - 3/2 m_\pi] t}$$

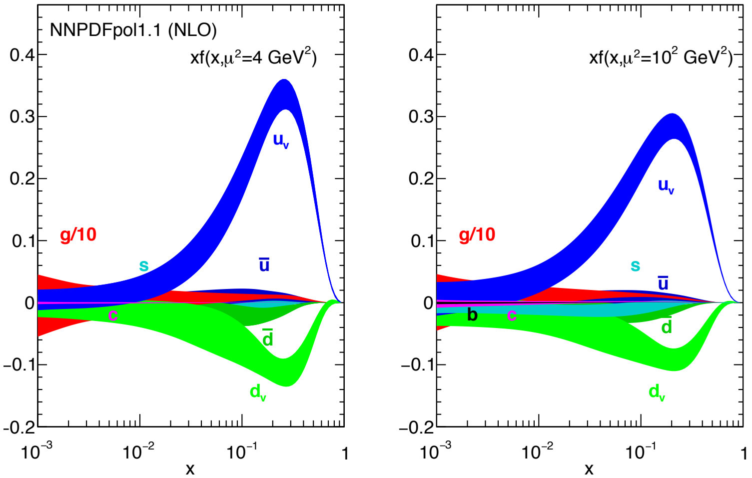
G. Parisi (1984) P. Lepage (1989)

# Determination of PDFs from Experiment




Global fits to experimental data Parton distributions and lattice QCD calculations: a community white paper arXiv: 1711.07916

# Determination of PDFs from Experiment




Global fits to experimental data Parton distributions and lattice QCD calculations: a community white paper arXiv: 1711.07916

# 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 ...