The Continuum and Leading Twist Limits of Parton Distribution Functions in Lattice QCD

ANL Lattice Hadron Structure Journal Club
(Thank you to Yong Zhao and Ian Cloët for organizing)

Joe Karpie
(Columbia University)

Columbia University
In the City of New York

As part of the
HadStruc Collaboration
Along with

K. Orginos (W&M / JLab)
A. Radyushkin (Old Dominion U / JLab)
S. Zafeiropoulos (Aix-Marseille)

Based upon
arXiv:2105.13313

HadStruc Collaboration

- Chris Chamness, Colin Egerer, Tanjib Khan, Dan Kovner, Chris Monahan, Kostas Orginos, Raza Sufian (W&M)
- Robert Edwards, Christos Kallidonis, Nikhil Karthik, Jian-Wei Qiu, David Richards, Eloy Romero, Frank Winter (JLab)
- Wayne Morris, Anatoly Radyushkin (ODU)
- Bálint Joó (ORNL)
- Savvas Zafeiropoulos (Aix-Marseille)
- Joe Karpie (Columbia U)

LaMET and SDF

- Two related methods to analyze the space-like separated fields with Large Momentum Effective Theory or Short Distance Factorization to obtain PDFs
- LaMET/SDF and the PDF
  - LaMET: factorization relation and power expansion with respect to large momentum scale p_z^{-2}

    X. Ji (2013) 1305.1539
  - SDF: factorization relation and power expansion with respect to short distance scale z^2

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

- Wilson Line Operator matrix element

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

  - Lorentz Composition

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

B. Musch et al (2010) 1011.1213

LaMET and SDF

- Two related methods to analyze the space-like separated fields with Large Momentum Effective Theory or Short Distance Factorization to obtain PDFs
- LaMET and SDF
  - LaMET: factorization and power expansion with respect to large momentum scale p_z^{-2}
    X. Ji (2013) 1305.1539
  - SDF: factorization and power expansion with respect to short distance scale z^2

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

- SDF begins with the OPE with a short distance scale
  - Power corrections are ordered by twist K-F Liu (1999) 9910306
  - The SDF's leading twist kernel is related to LaMET's kernel by integral formula.

T. Izubuchi et al (2018) 1801.03917

$$\mathfrak{M}(\nu, z^2) = \int_0^1 du\, K(u, \mu^2 z^2) Q(u\nu, \mu^2) + O(z^2)$$

- Known to $O(\alpha_s)$ A. Radyushkin (2017) 1710.08813, J.-H. Zhang et. al. (2018) 1801.03023, T. Izubuchi et. al. (2018) 1801.03917; $O(\alpha_s^2)$ Z-Y Li, Y-Q Ma, J-Q Qiu 2006.12370

The Reduced distribution and normalization

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

- The pseudo-ITD is subject to many systematic errors
  - Lattice spacing, higher twist, incorrect pion mass, finite volume
- A ratio can remove renormalization constants and the low Ioffe time systematic errors
  - Avoids additional gauge fixed RI-Mom calculations
  - Is a renormalization group invariant quantity, guaranteeing finite continuum limit (no power divergences)

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

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

Systematic errors of Lattice PDFs

- Pion mass
  - Just use correct values (duh!)
  - Extrapolate PDF to physical pion mass

C. Alexandrou et al (2018) 1803.02685
J-W Chen et al (2018) 1803.04393

- Finite Volume
  - Calculate size of effects in a model theory

    R. Briceño et al (2018) 1805.01034
  - Parameterize unknown functional dependence

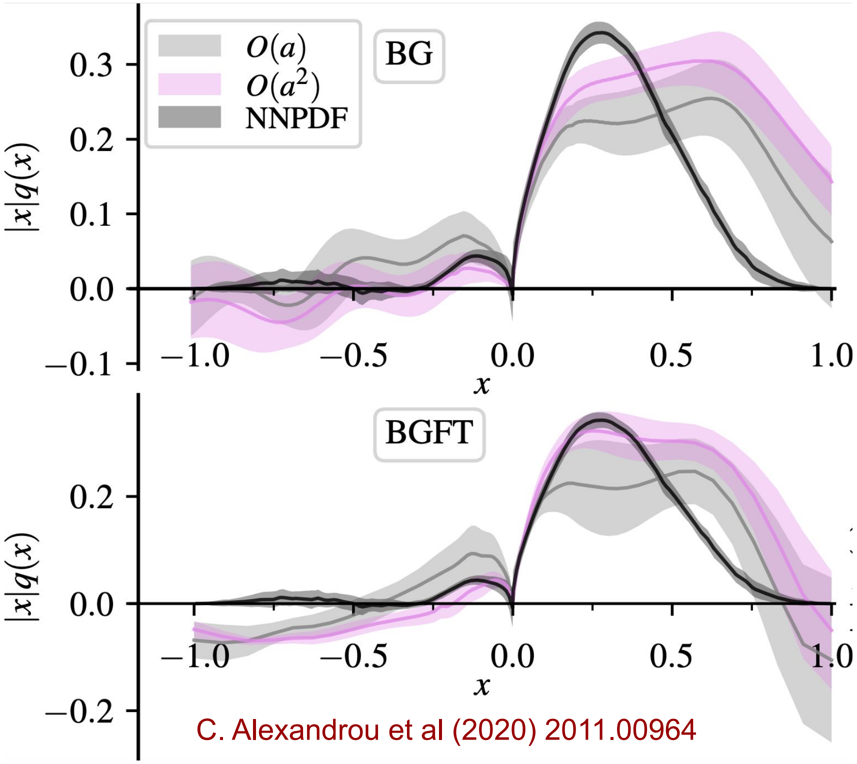
B. Joó, JK, K. Orginos, A. Radyushkin, D. Richards, R. Sufian, S. Zafeiropoulos (2019) 1908.09771
B. Joó, JK, K. Orginos, A. Radyushkin, D. Richards, R. Sufian, S. Zafeiropoulos (2019) 1909.08517
R. Sufian, C. Egerer, JK, R. Edwards, B. Joó, Y-Q Ma, K. Orginos, J-W Qiu, D. Richards (2020) 2001.04960
B. Joó, JK, K. Orginos, A. Radyushkin, D. Richards, R. Sufian, S. Zafeiropoulos (2020) 2004.01687

- Lattice Spacing
  - Parameterizing unknown functional dependence

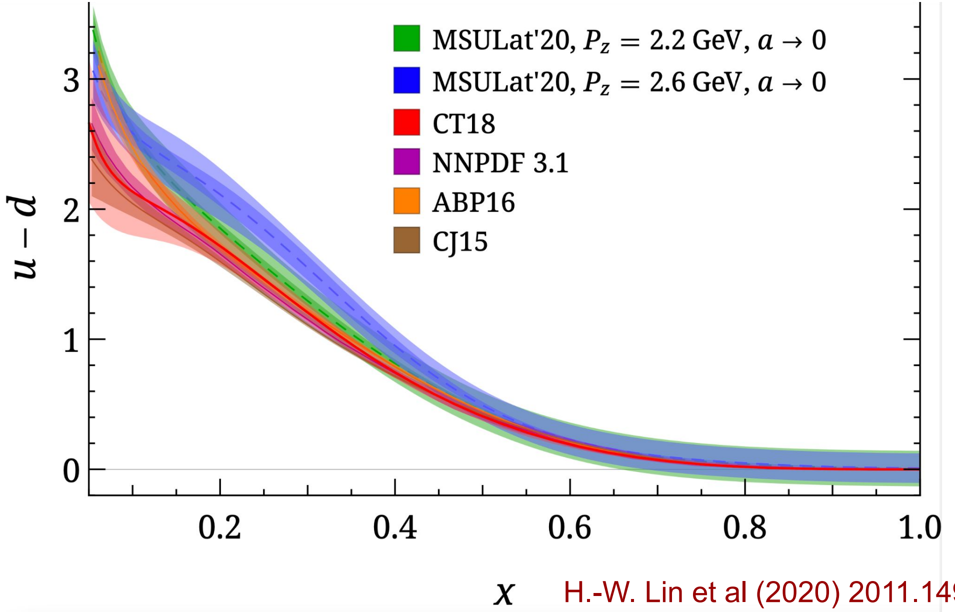
    X. Gao et al (2020) 2007.06590
  - Interpolate data at fixed hard scale and extrapolate continuum limit

    C. Alexandrou et al (2020) 2011.00964
    H.-W. Lin et al (2020) 2011.14971
- Power Corrections
  - LaMET p_z^{-2}
  - SDF and Lattice Cross Sections z^2
  - OPE without OPE and Hadronic Tensors Q^{-2}
- Inverse Problems
  - Get to these later

Continuum limits of other Lattice PDFs



C. Alexandrou et al (2020) 2011.00964



H.-W. Lin et al (2020) 2011.14971

Systematic Errors in pion pseudo-ITD fits

B. Joó, JK, K. Orginos, A. Radyushkin, D. Richards, R. Sufian, S. Zafeiropoulos (2019) 1909.08517

- Our data is subject to systematic errors from many sources
  - Higher twist, finite lattice spacing, unphysical pion mass, finite volume
- First matching is applied applied to the data to remove $\log(z^2)$ dependence

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

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

- Parameterizations of these corrections are fit to the ITD

$$\tau = \frac{\sqrt{1+\nu}-1}{\sqrt{1+\nu}+1}$$

  - Reduced pseudo-ITD corrections begin at $O(\nu^2)$

$$Q(\nu,\mu^2,z^2) = \sum_{k=0}^{k_{max}} \lambda_k \tau^k \left[1 + \nu^2(c_1 z^2 + c_2 z^2 \log\left(z^2\mu^2\frac{e^{2\gamma_E+1}}{4}\right) + c_3 e^{-m_\pi(L-z)}\right]$$

Systematic Errors in LCS in fits

R. Sufian, C. Egerer, JK, R. Edwards, B. Joó, Y-Q Ma, K. Orginos, J-W Qiu, D. Richards (2020) 2001.04960

- Our data is subject to systematic errors from many sources
  - Higher twist, finite lattice spacing, unphysical pion mass, finite volume
  - Expect to be larger than the pseudo-ITD due to lack of ratio

- Data does not have precision for identifying DGLAP $\log(z^2)$ behavior.

- Parameterizations of these corrections are fit to the LCS
  - LCS corrections begin at $O(\nu^0)$

$$\sigma_{\mathrm{VA}}(\nu, z^2) = \sum_{k=0}^{k_{\max}=4} \lambda_k \tau^k + b_1(m_\pi - m_{\pi,\mathrm{physical}}) + b_2 a^2 + b_3 z^2 + b_4 a^2 p_z^2 + b_5 e^{-m_\pi (L-z)}$$

Inverse Problem Solutions for Lattice PDFs

- Parametric — JK, K. Orginos, A. Rothkopf, S. Zafeiropoulos (2019) 1901.05408
  - Fit a phenomenologically motivated function
    - Method used by most pheno extractions
    - Potentially significant, but controllable model dependence
  - Fit to a neural network — S. Forte, L. Garrido, J. Latorre, A. Piccione (2002) 0204232
    - Machine learning is hip
    - Expensive tuning procedure

    K. Cichy, L. Del Debbio, T. Giani (2019) 1907.06037
    L. Del Debbio, T. Giani, JK, K. Orginos, A. Radyushkin, S. Zafeiropoulos (2020) 2010.03996
- Non-Parametric
  - Backus-Gilbert — J. Liang, K-F. Liu, Y-B. Yang (2017) 1710.11145
    - No model dependence, one tunable parameter
  - Bayesian Reconstruction — Y. Burnier and A. Rothkopf (2013) 1307.6106, J. Liang et al (2019) 1906.05312
    - Very general, Bayesian statistics has systematics included in meaningful way
  - Bayes-Gauss-Fourier transform — C. Alexandrou, G. Iannelli, K. Jansen, F. Manigrasso (2020) 2007.13800

Unknown functions

- Want to determine a continuous unknown function from the data
- Lattice systematic errors

J-W Chen et al (2017) 1710.01089

    - Lattice spacing is the only one used in this study

$$\mathfrak{M}(p, z, a) = \mathfrak{M}_{\text{cont}}(\nu, z^2) + \sum_{n=1} \left(\frac{a}{|z|}\right)^n P_n(\nu) + (a\Lambda_{\text{QCD}})^n R_n(\nu)$$

- Power Corrections

$$\mathfrak{M}_{\text{cont}}(\nu, z^2) = \mathfrak{M}_{\text{lt}}(\nu, z^2) + \sum_{n=1} (z^2 \Lambda_{\text{QCD}}^2)^n B_n(\nu)$$

- Factorization of the PDF

$$\text{Re/Im}\, \mathfrak{M}_{\text{lt}}(\nu, z^2) = \int_0^1 dx\, \mathcal{K}_{R/I}(x\nu, \mu^2 z^2) q_{\mp}(x, \mu^2)$$

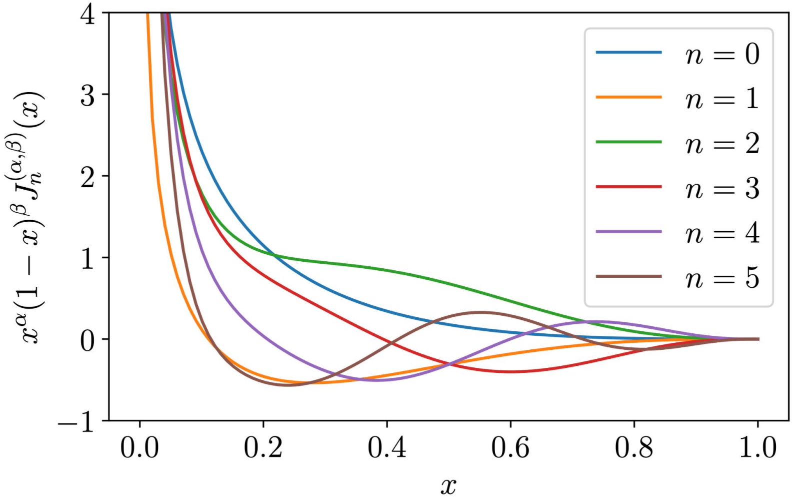
Jacobi Polynomials

- Orthogonal set of Polynomials
  - Textbook orthogonality relationship $\int_{-1}^{1} dz(1-z)^{\alpha}(1+z)^{\beta} j_n^{(\alpha,\beta)}(z) j_m^{(\alpha,\beta)}(z) = \tilde{N}_n^{(\alpha,\beta)} \delta_{n,m}$
- Change variables for more useful metric and integration range: $z = 1 - 2x$

$$\int_0^1 dx\, x^{\alpha}(1-x)^{\beta} J_n^{(\alpha,\beta)}(x) J_m^{(\alpha,\beta)}(x) = N_n^{(\alpha,\beta)} \delta_{n,m}$$

$$J_n^{(\alpha,\beta)}(x) = \sum_{j=0}^{n} \omega_{n,j}^{(\alpha,\beta)} x^j$$

$$\omega_{n,j}^{(\alpha,\beta)} = \binom{n}{j} \frac{(-1)^j}{n!} \frac{\Gamma(\alpha+n+1)\Gamma(\alpha+\beta+n+j+1)}{\Gamma(\alpha+\beta+n+1)\Gamma(\alpha+j+1)}$$



$\alpha = -0.5\,, \quad \beta = 3$

Jacobi Polynomial parameterizations

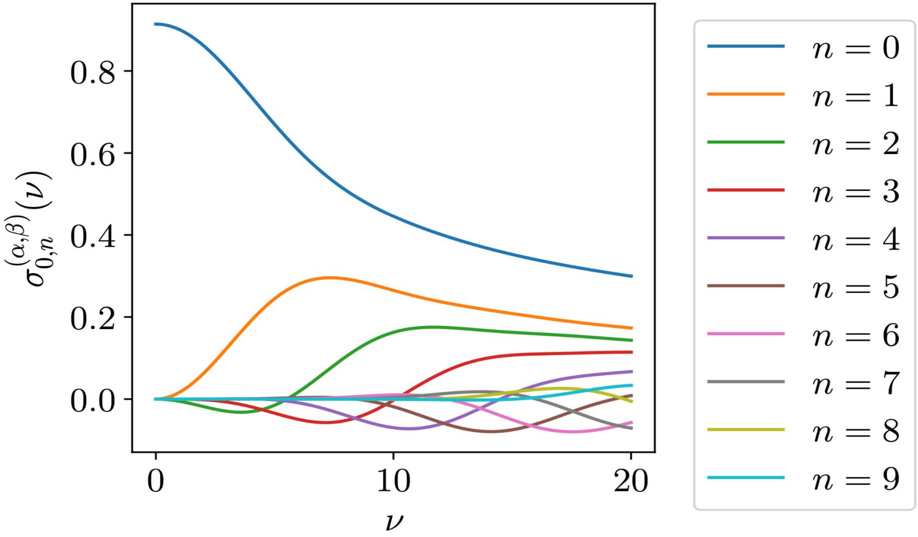
- Parameterize unknown functions
  - Example: PDFs $$q_\pm(x) = x^\alpha(1-x)^\beta \sum_{n=0} {}_\pm d_n^{(\alpha,\beta)} J_n^{(\alpha,\beta)}(x)$$
- How to Fourier transform of this parameterization

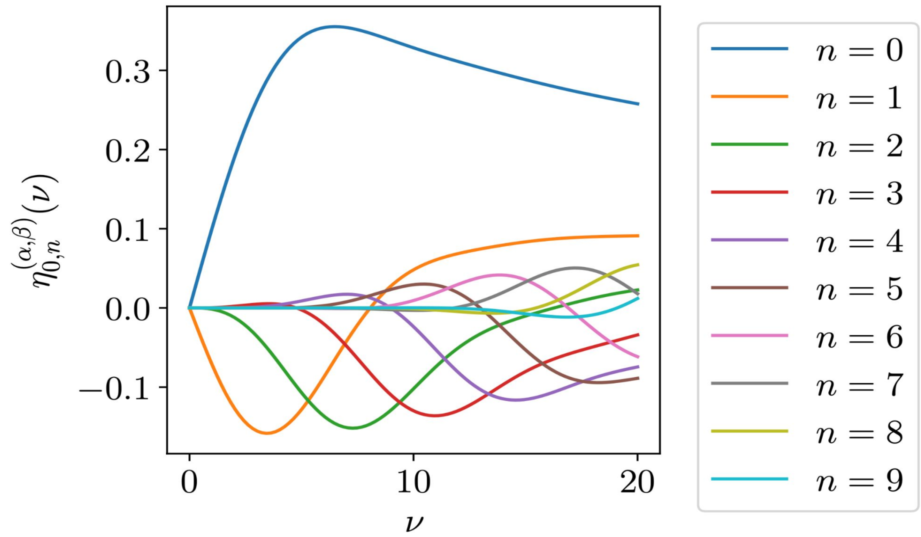
$$\sigma_{0,n}^{(\alpha,\beta)}(\nu) = \int_0^1 dx\, x^\alpha(1-x)^\beta \cos(\nu x) J_n^{(\alpha,\beta)}(x)$$

$$\eta_{0,n}^{(\alpha,\beta)}(\nu) = \int_0^1 dx\, x^\alpha(1-x)^\beta \sin(\nu x) J_n^{(\alpha,\beta)}(x)$$

$$\mathrm{Re}\, Q(\nu) = \sum_{n=0} \sigma_{0,n}^{(\alpha,\beta)}(\nu)_- d_n \qquad \mathrm{Im}\, Q(\nu) = \sum_{n=0} \eta_{0,n}^{(\alpha,\beta)}(\nu)_+ d_n$$

Jacobi Polynomial parameterizations





$$\alpha = -0.5\,, \quad \beta = 3$$

- Decays to 0 with Ioffe time
- Large n only at large Ioffe time if coefficients are small

Jacobi Polynomial parameterizations

$O(\alpha_s)$ A. Radyushkin (2017) 1710.08813
J.-H. Zhang et. al. (2018) 1801.03023
T. Izubuchi et. al. (2018) 1801.03917

$O(\alpha_s^2)$ Z-Y Li, Y-Q Ma, J-Q Qiu 2006.12370

- Including the factorization kernel

$$\sigma_n^{(\alpha,\beta)}(\nu,\mu^2 z^2) = \int_0^1 dx\, x^\alpha (1-x)^\beta \mathcal{K}_R(x\nu,\mu^2 z^2) J_n^{(\alpha,\beta)}(x)$$

$$\sigma_n^{(\alpha,\beta)}(\nu,\mu^2 z^2) = \sigma_{0,n}^{(\alpha,\beta)}(\nu) + \sigma_n^{(\mathrm{NLO})}(\nu,\mu^2 z^2) + O(\alpha_S^2)$$

$$\eta_n^{(\alpha,\beta)}(\nu,\mu^2 z^2) = \int_0^1 dx\, x^\alpha (1-x)^\beta \mathcal{K}_I(x\nu,\mu^2 z^2) J_n^{(\alpha,\beta)}(x)$$

$$\eta_n^{(\alpha,\beta)}(\nu,\mu^2 z^2) = \eta_{0,n}^{(\alpha,\beta)}(\nu) + \eta_n^{(\mathrm{NLO})}(\nu,\mu^2 z^2) + O(\alpha_S^2)$$

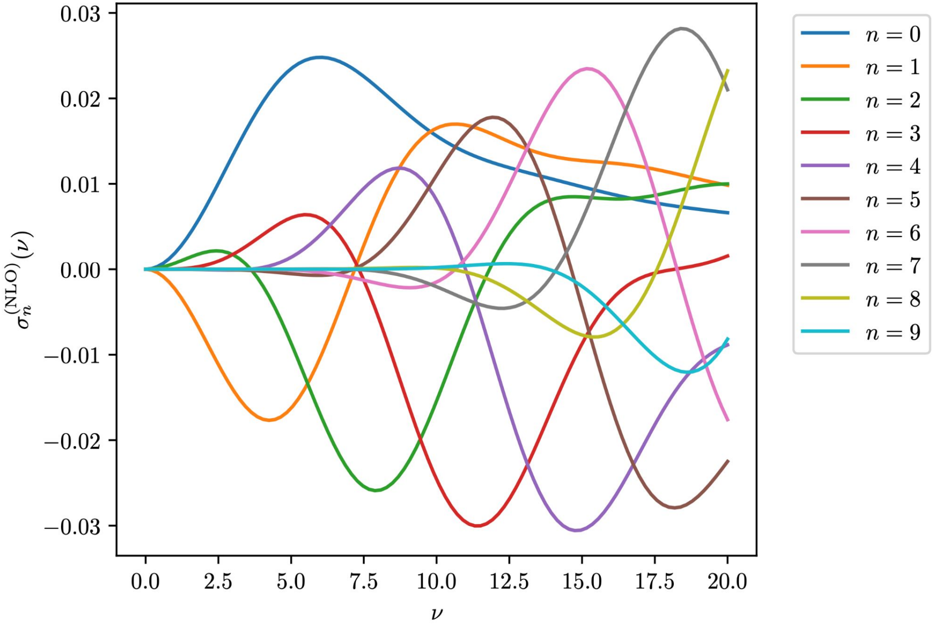
- Parameterize leading twist pseudo-ITD instead of ITD

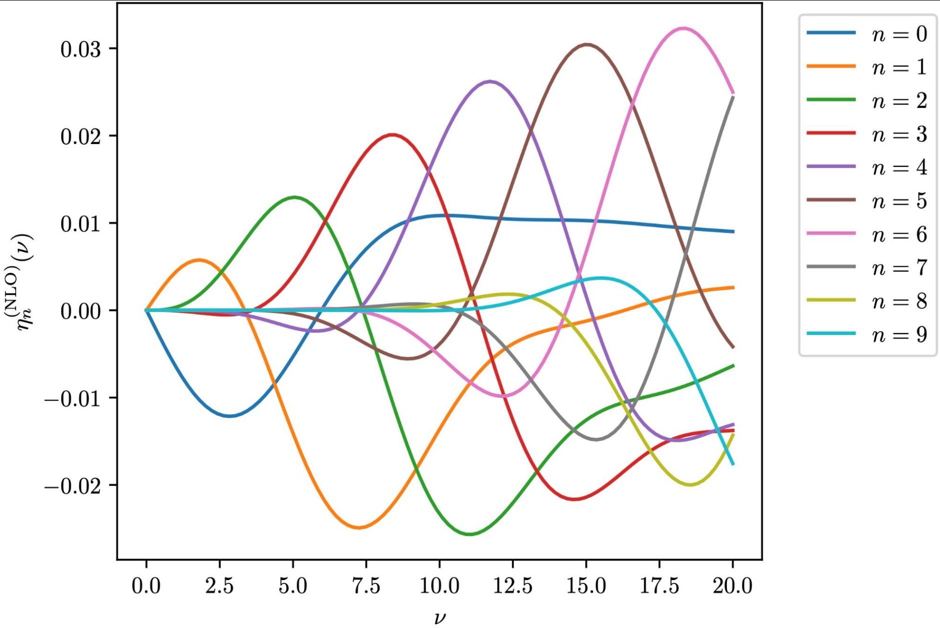
$$\mathrm{Re}\,\mathfrak{M}_{\mathrm{lt}}(\nu,z^2) = \sum_{n=0} \sigma_n^{(\alpha,\beta)}(\nu,\mu^2 z^2)_- d_n$$

$$\mathrm{Im}\,\mathfrak{M}_{\mathrm{lt}}(\nu,z^2) = \sum_{n=0} \eta_n^{(\alpha,\beta)}(\nu,\mu^2 z^2)_+ d_n$$

Jacobi Polynomial parameterizations

$z = 4a_{\mathrm{E5}}$





- Remains small function Ioffe time, generating small perturbative corrections at NLO
- Future work will expand to NNLO

$\alpha = -0.5\,, \quad \beta = 3$

Jacobi Polynomial parameterizations

- The normalization of the unknown functions is governed by the $n = 0$ coefficients
  - Nuisance terms will have no $n = 0$ terms
- With infinite terms, all α and β can parameterize the PDF
  - In that limit, α and β lose their meaning and cannot distinguish large or small x behavior
- At truncated number of terms, α and β can be fit to find optimal parameters for that truncation, given that it is common between all terms
- Relationship between linear coefficients and moments

$$\pm d_n^{(\alpha,\beta)} = \frac{1}{N_n^{(\alpha,\beta)}} \sum_{j=0}^{n} \omega_{n,j}^{(\alpha,\beta)} a_j^{\pm}$$

Jacobi Polynomial parameterizations

- Final functional form

ITD HT a□ a/z

$$\mathrm{Re}\,\mathfrak{M}(p,z,a) = \sum_{n=0}^{N_-+1} \sigma_n(\nu,\mu^2 z^2)\,{}_-d_n^{(\alpha,\beta)} + z^2\Lambda_{\mathrm{QCD}} \sum_{n=1}^{N_{R,b}} \sigma_{0,n}(\nu) b_{R,n}^{(\alpha,\beta)} + a\Lambda_{\mathrm{QCD}} \sum_{n=1}^{N_{R,r}} \sigma_{0,n}(\nu) r_{R,n}^{(\alpha,\beta)} + \frac{a}{|z|} \sum_{n=1}^{N_{R,p}} \sigma_{0,n}(\nu) p_{R,n}^{(\alpha,\beta)}$$

$$\mathrm{Im}\,\mathfrak{M}(p,z,a) = \sum_{n=0}^{N_+} \eta_n(\nu,\mu^2 z^2)\,{}_+d_n^{(\alpha,\beta)} + z^2\Lambda_{\mathrm{QCD}} \sum_{n=1}^{N_{I,b}} \eta_{0,n}(\nu) b_{I,n}^{(\alpha,\beta)} + a\Lambda_{\mathrm{QCD}} \sum_{n=1}^{N_{I,r}} \eta 0, n(\nu) r_{I,n}^{(\alpha,\beta)} + \frac{a}{|z|} \sum_{n=1}^{N_{I,p}} \eta_{0,n}(\nu) p_{I,n}^{(\alpha,\beta)}$$

- ${}_-d_0^{(\alpha,\beta)} = 1/B(\alpha+1,\beta+1)$

Bayesian Fits

$$P\left[\theta|\mathfrak{M}_L, I\right] = \frac{P\left[\mathfrak{M}_L|\theta\right] P\left[\theta|I\right]}{P\left[\mathfrak{M}_L|I\right]}$$

- Standard χ^2 minimization, but with modified function

$$P\left[\mathfrak{M}_L|\theta\right] = \frac{\exp[-\frac{\chi^2}{2}]}{Z_\chi} \qquad \chi^2 = \sum_{k,l}(\mathfrak{M}_k^L - \mathfrak{M}_k)C_{kl}^{-1}(\mathfrak{M}_l^L - \mathfrak{M}_l)$$

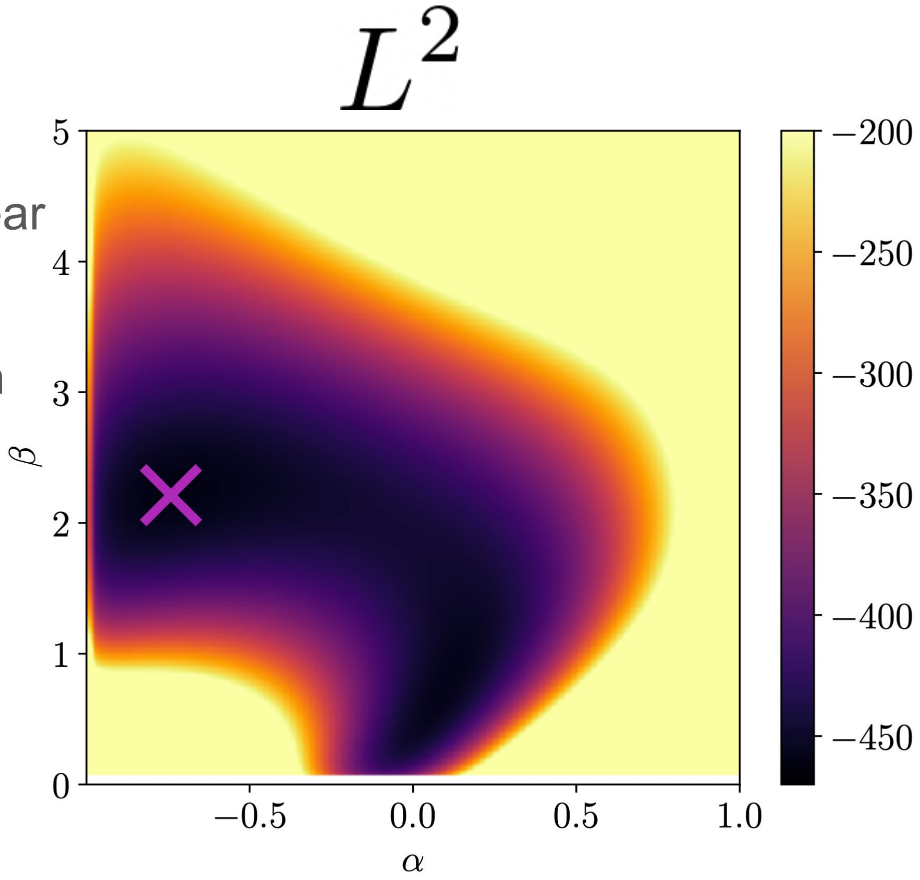
$$P\left[\theta|\mathfrak{M}_L, I\right] = \frac{\exp[-\frac{L^2}{2}]}{Z} \qquad L^2 = \chi^2 - 2\log\left(P[\theta|I]\right)$$

- Prior Distributions
    - Uniform distribution within bounds
    - Normal distribution
    - Log-Normal distribution

- Additional terms are designed to push weakly push the maximum probability to “reasonable” values

Variable Projection

- Fitting a linear combination of non-linear functions can be accelerated using Variable Projection (VarPro)
- Only Non-linear parameters needed in iterative non-linear minimization
- Linear parameters are minimized analytically
- After defining model only α and β are minimized
- Reducing number of parameters in non-linear fit dramatically improves stability
- Ratio with non-zero momenta cannot use VarPro



Lattice ensembles

ID	a(fm)	M_π(MeV)	β	c_{SW}	κ	$L^3 \times T$	N_{cfg}
$\tilde{A}5$	0.0749(8)	446(1)	5.2	2.01715	0.13585	$32^3 \times 64$	1904
E5	0.0652(6)	440(5)	5.3	1.90952	0.13625	$32^3 \times 64$	999
N5	0.0483(4)	443(4)	5.5	1.75150	0.13660	$48^3 \times 96$	477

- E5 and N5 were generated as part of CLS collaboration

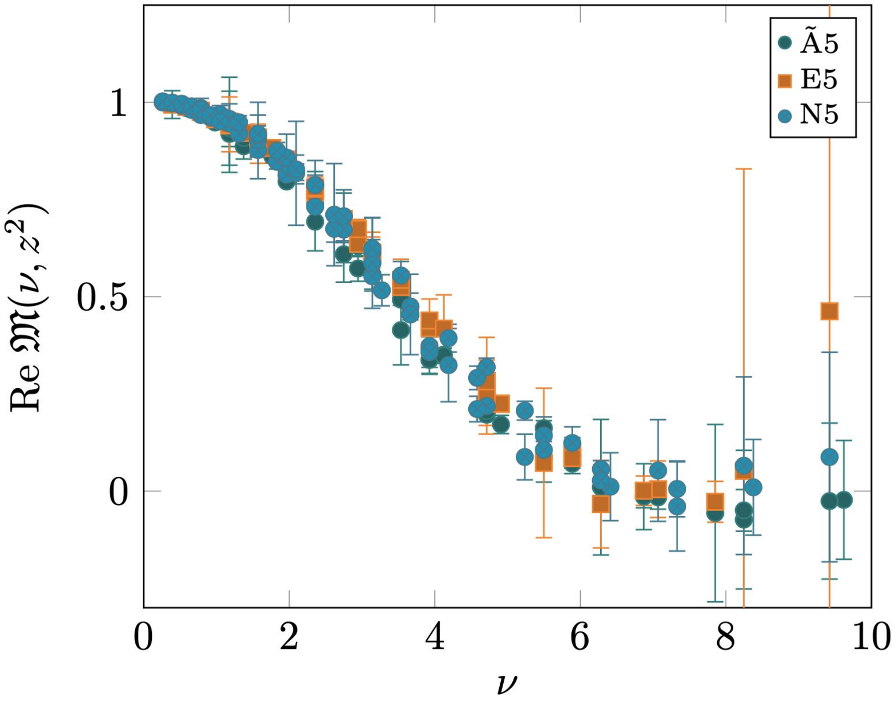
P. Fritzsch et al (2012) 1205.5380

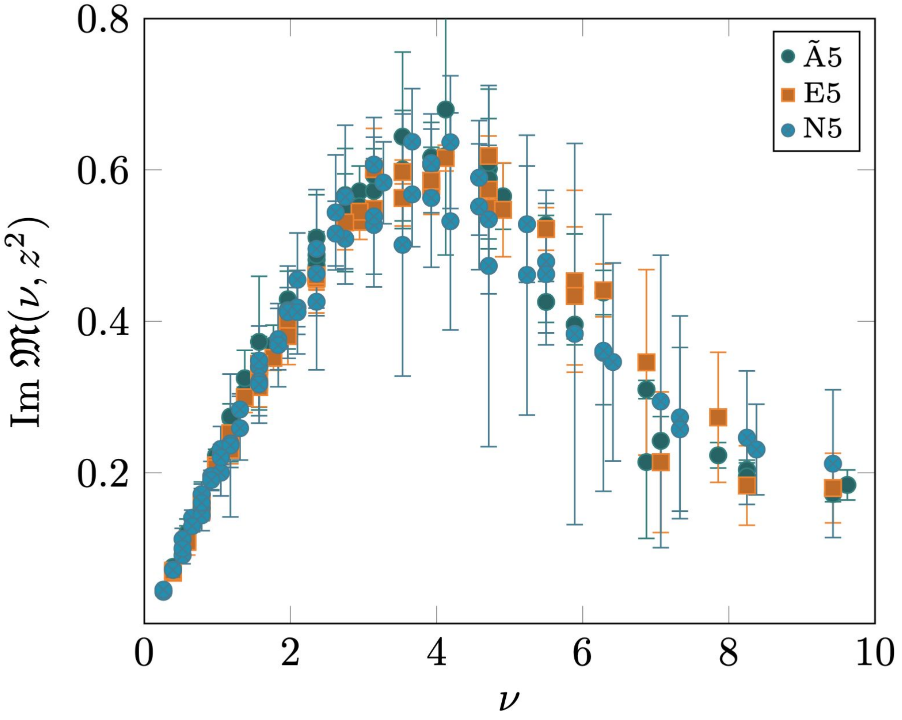
- Ã5 was generated for this study
- Three lattice spacings for lattice spacing dependence
- Fixed pion mass
- Will ignore the difference between physical volumes until future work

Obtaining Matrix Elements

- Used combination of summation and generalized eigenvalue problem methods (sGEVP) to control excited state contamination
  - 3 operators only gives slight improvement   J. Bulava (2011) 1108.3774
  - Used 3 momentum smearing parameters and 2 types of smearing the sink interpolator field (point and gaussian)
  - Whichever of the 6 correlators had sufficient signal were used within the fit, dropping the largest momentum smearing parameter results for fits to small momenta data and the smallest momentum smearing parameter results for fits to the large momenta data
- Estimated systematic error from fitting matrix elements by varying minimum Euclidean time

Reduced Matrix Elements





Prior Distribution parameters

- Non-linear parameters
  - Log-Normal distribution

$$\alpha_0 = 0\,, \quad \sigma_\alpha = 0.4$$

$$\beta_0 = 3\,, \quad \sigma_\beta = 1$$

- Linear PDF parameters
  - Normal distribution

$$d_0 = 0\,, \quad \sigma_d = 0.5$$

- Linear nuisance parameters
  - Normal distribution

$$c_0 = 0\,, \quad \sigma_c = 0.1$$

Chi squared of fits

$$N_{\pm} = 2 \quad N_{R/I,b/p/r} = 0, 1$$

model	Real L^2/d.o.f.	Real χ^2/d.o.f.	Imag L^2/d.o.f.	Imag χ^2/d.o.f.
Q only	3.173	3.094	3.146	3.095
Q and B_1	2.721	2.479	3.054	2.969
Q and R_1	3.028	2.748	3.068	2.871
Q and P_1	0.876	0.809	1.186	1.088
Q, B_1, and R_1	2.610	2.057	2.917	2.619
Q, B_1, and P_1	0.852	0.723	1.020	0.888
Q, R_1, and P_1	0.881	0.763	1.289	1.063
All terms	0.857	0.727	1.026	0.893

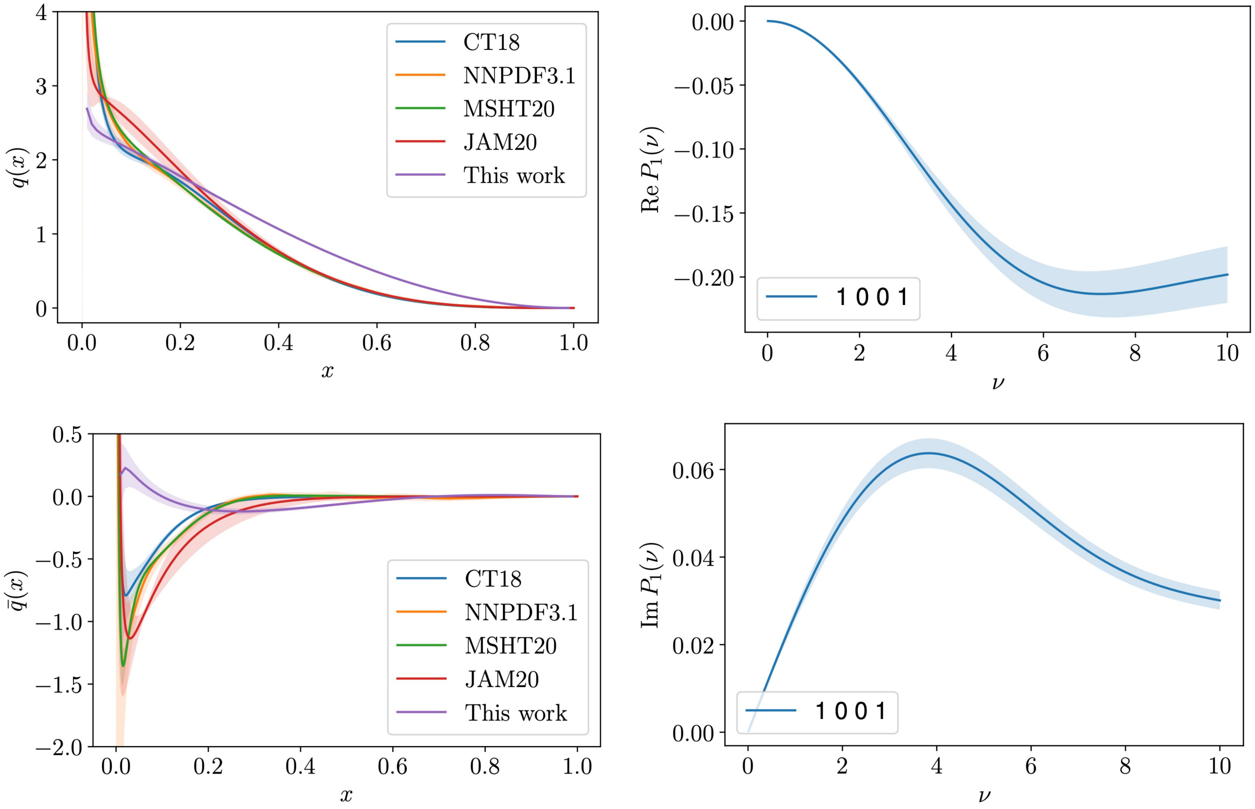
Minimalist parameterization

- Philosophy- The best models are those with the fewest number of parameters and steps should be taken to avoid overfitting

- Parameters should be added if they improve the χ^2 or L^2 significantly, and avoided otherwise.

- PDF terms $N_{\pm} = 1$
- $\frac{a}{|z|}$ terms $N_{R/I,p} = 1$
- No other nuisance terms $N_{R/I,b/r} = 0$

Minimalist parameterization



$$N_{\pm} = 1$$

$$N_{R/I,\,p} = 1$$

$$N_{R/I,\,b/r} = 0$$

$$L^2/\mathrm{d.o.f.} = 0.958$$

$$\chi^2/\mathrm{d.o.f.} = 0.856$$

Adding more parameters

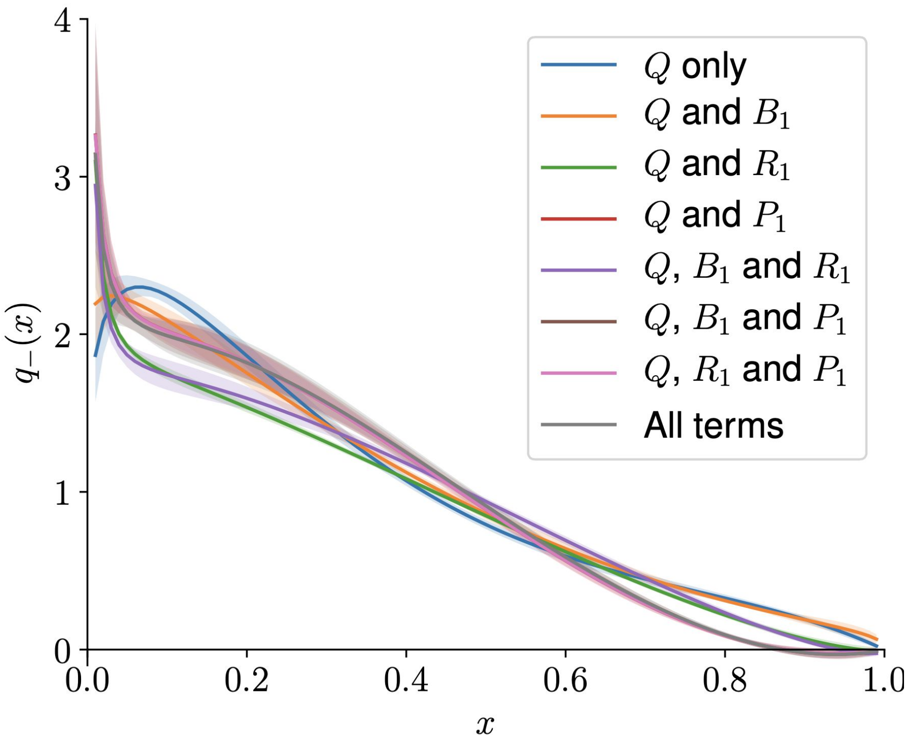
- Expectations of effects on χ^2 and L^2 of various data
  - Low Ioffe time data is generally more precise than large Ioffe time data
  - Large p_z data has large errors from signal-to-noise ratio of correlation functions
  - Large z data has large errors from requiring ratio to remove exponentially decaying renormalization constant
  - Ratio cancels nuisance terms in low Ioffe time limit, assuming their coefficients are small
- $\frac{a}{|z|}$ effects are strongest at small z, at low Ioffe time, which have more precision, but their size is suppressed by the ratio
  - Could have larger effects on χ^2 and L^2 due to precision
- z^2 effects are strong where data has less precision at larger Ioffe times
- $a\Lambda_{\mathrm{QCD}}$ effects could be strong in any regime

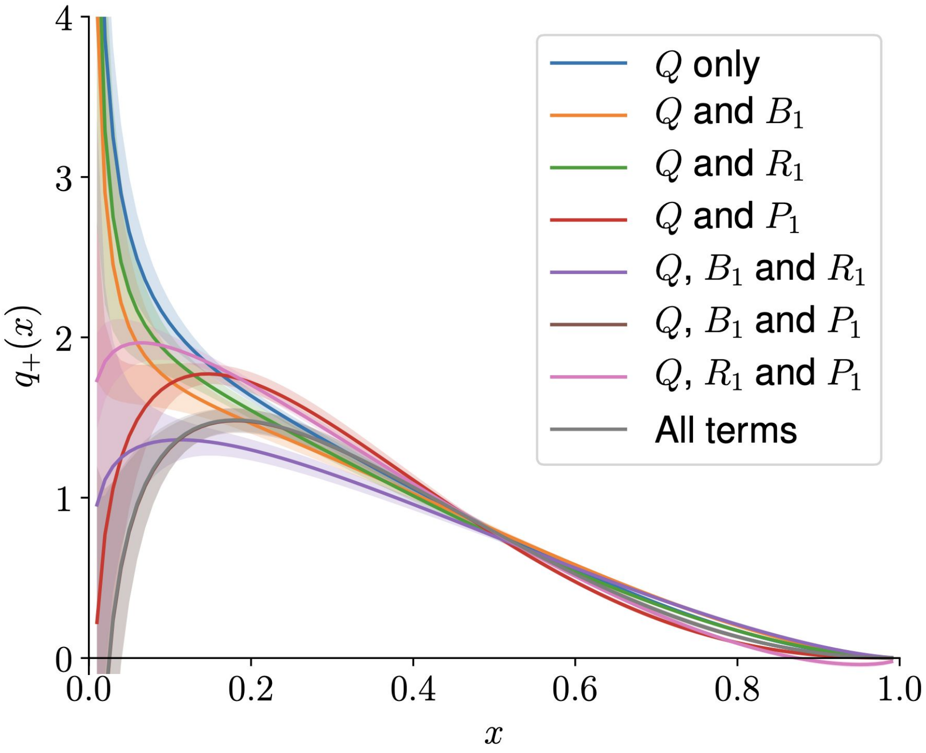
Chi squared of fits

$$N_{\pm} = 2 \quad N_{R/I,\, b/p/r} = 0,\, 1$$

model	Real L^2/d.o.f.	Real χ^2/d.o.f.	Imag L^2/d.o.f.	Imag χ^2/d.o.f.
Q only	3.173	3.094	3.146	3.095
Q and B_1	2.721	2.479	3.054	2.969
Q and R_1	3.028	2.748	3.068	2.871
Q and P_1	0.876	0.809	1.186	1.088
Q, B_1, and R_1	2.610	2.057	2.917	2.619
Q, B_1, and P_1	0.852	0.723	1.020	0.888
Q, R_1, and P_1	0.881	0.763	1.289	1.063
All terms	0.857	0.727	1.026	0.893

Study varying parameters





Studying nuisance terms

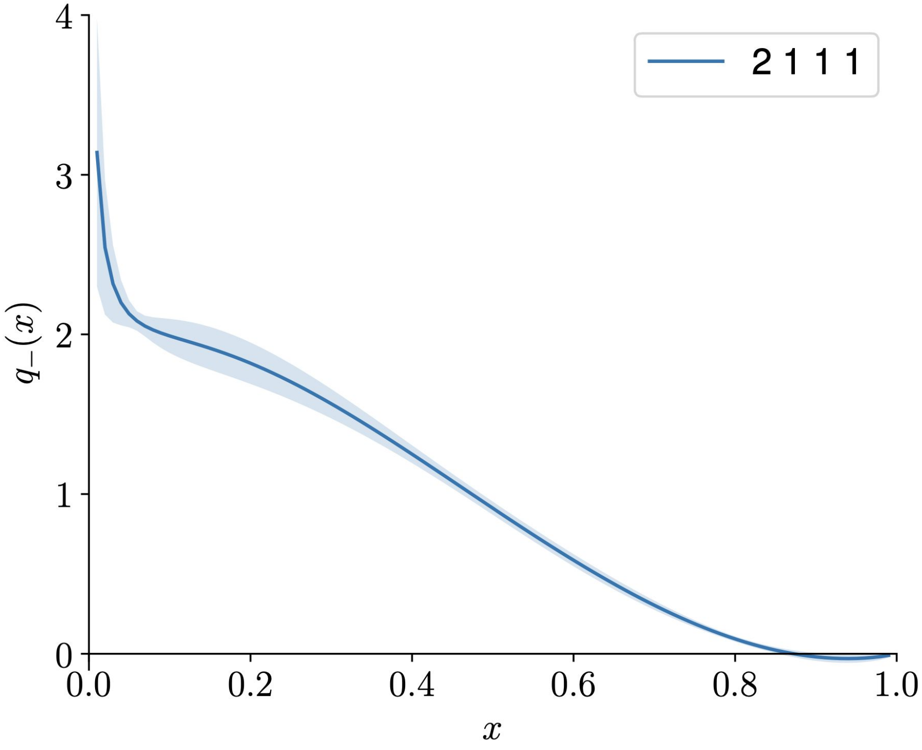
- Calculating all possible nuisance terms will allow us to see expected size of systematic errors

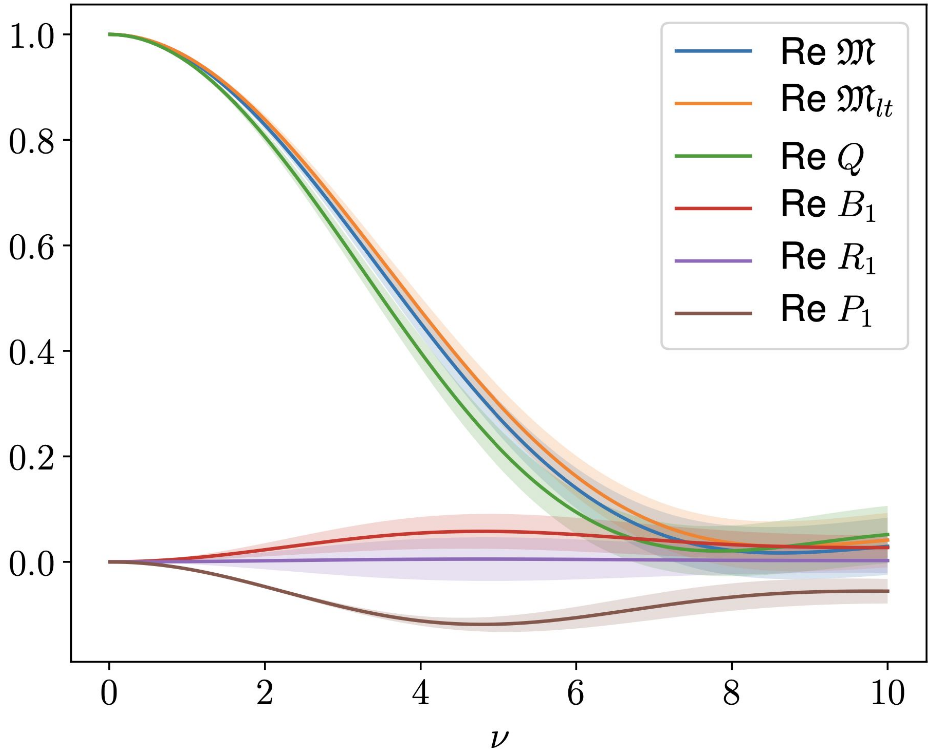
- Parameters whose distribution seems to match the prior distribution are not being controlled by the data
  - Removal of them should not effect χ^2only L^2

- If priors are uncorrelated, parameters which are dominated by priors will be uncorrelated to other parameters

Study nuisance terms

$$N_{\pm} = 2 \quad N_{R/I,b/p/r} = 1$$

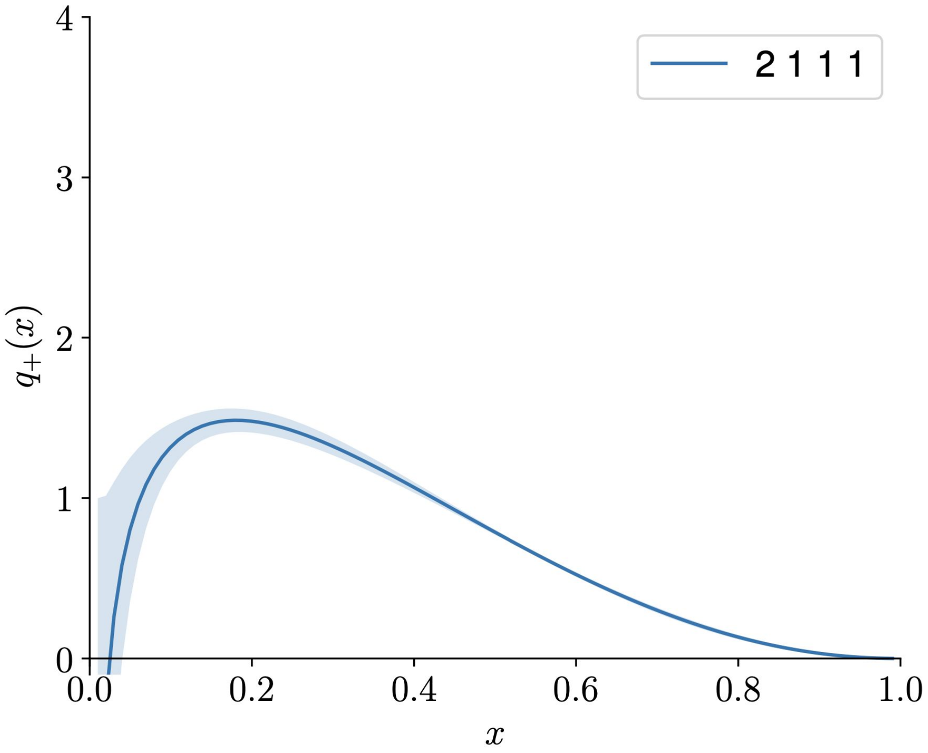


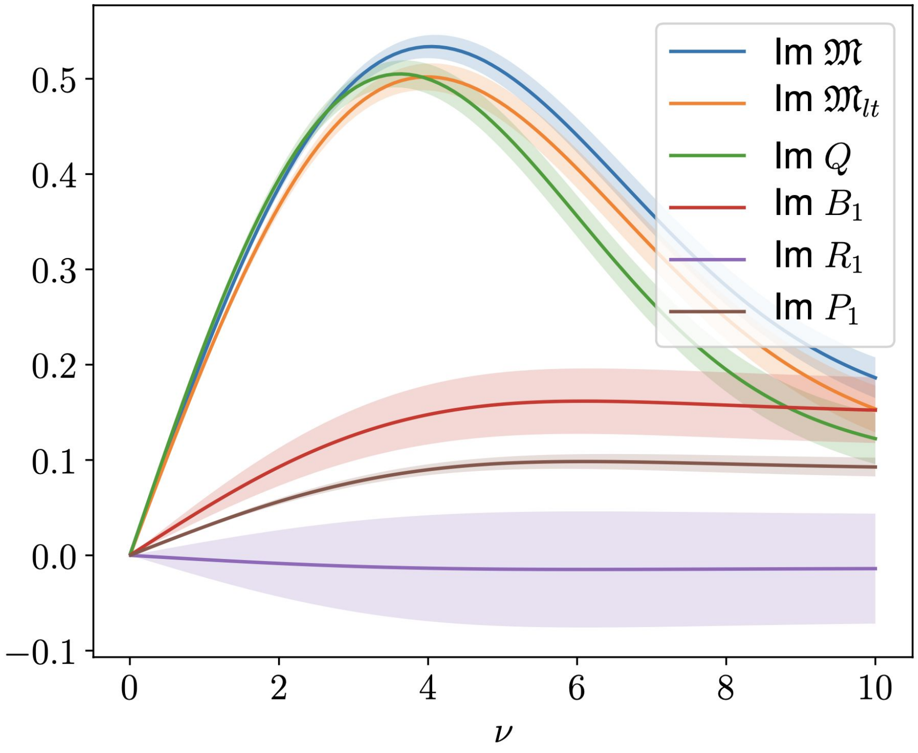


$$z = 4a_{\mathrm{E5}}$$

Study nuisance terms

$$N_{\pm} = 2 \quad N_{R/I,b/p/r} = 1$$





$$z = 4a_{\mathrm{E5}}$$

Fit results

$$N_{\pm} = 2 \quad N_{R/I,b/p/r} = 1$$

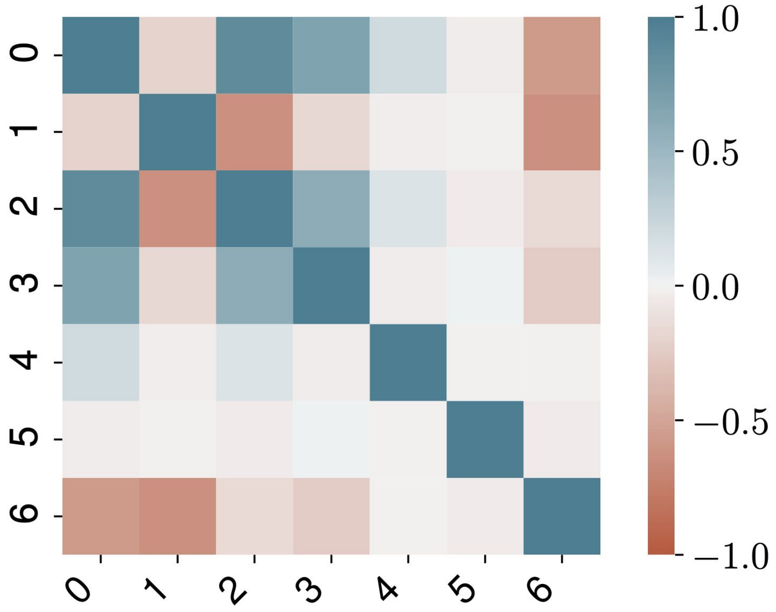
parameter	ID	value
α	0	-0.45(14)
β	1	0.93(20)
${}_{-}d_1^{(\alpha,\beta)}$	2	-0.29(31)
${}_{-}d_2^{(\alpha,\beta)}$	3	-0.77(6)
$b_{R,1}^{(\alpha,\beta)}$	4	0.13(6)
$r_{R,1}^{(\alpha,\beta)}$	5	0.01(10)
$p_{R,1}^{(\alpha,\beta)}$	6	-0.27(5)

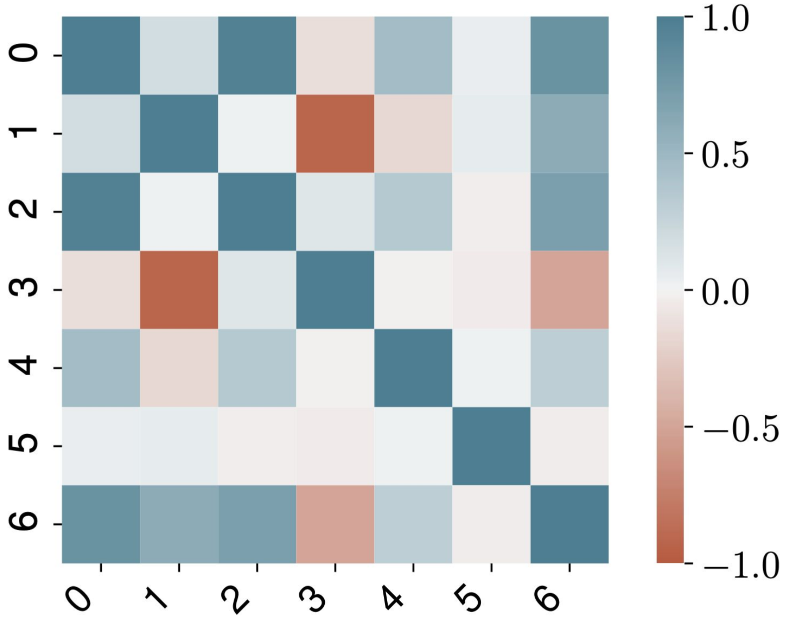
parameter	ID	value
α	0	-0.69(7)
β	1	2.11(13)
${}_{+}d_0^{(\alpha,\beta)}$	2	0.29(15)
${}_{+}d_1^{(\alpha,\beta)}$	3	-1.29(12)
$b_{I,1}^{(\alpha,\beta)}$	4	0.26(5)
$r_{I,1}^{(\alpha,\beta)}$	5	-0.02(10)
$p_{I,1}^{(\alpha,\beta)}$	6	0.16(2)

Correlations between parameters

$$N_{\pm} = 2 \quad N_{R/I,b/p/r} = 1$$

- Parameters which are only controlled by prior distributions will not be correlated to other parameters





AICc averaging

R. Zhang et al (2020) 2005.13955
H.-W. Lin et al (2020) 2011.14971

- Akaike Information Criteria (AIC)
  - Adds weight to disfavor models with too many parameters

$$a_i = 2k_i + 2L_i^2$$

- Corrected AIC (AICc)
  - Used when few number of datapoints compared to number of parameters

$$A_i = a_i + \frac{2k(k+1)}{n-k-1}$$

- Weighted average to determine expectation values of observables
  - Ideally, averages away model biases

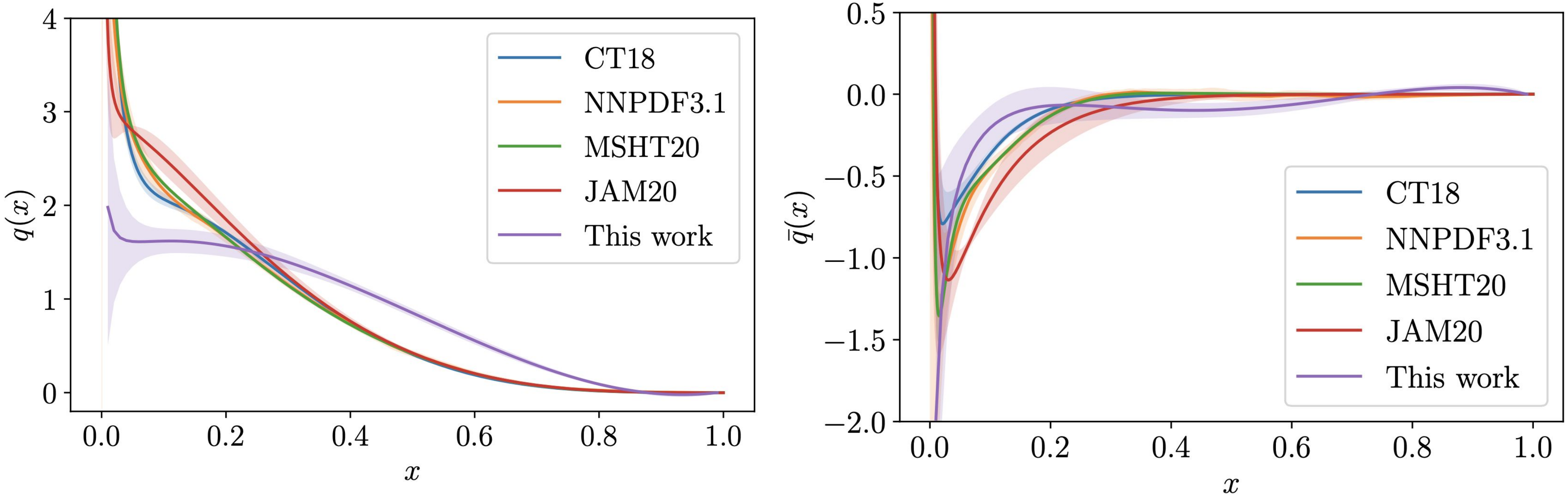
$$x = \sum_{i=1}^{N} w_i x_i \,, \qquad w_i = \frac{e^{-\frac{A_i}{2}}}{\sum_{i=1}^{N} e^{-\frac{A_i}{2}}}$$

AICc averaging

- Use a range of models for the AICc weighted average
- To average away model biases, sufficiently many distinct models are required
- Undesirable models could be removed, or their large AICc will exponentially suppress them in the weighted average
- For this study will use models with

$$N_{\pm} = 1,\ 2,\ 3$$
$$N_{R/I,b/p/r} = 0, \ldots, N_{\pm}$$

Averaged Results



Review

- Identify the systematic errors which your data is sensitive to
  - Higher twist, lattice spacing errors
- Define a function which parameterizes the systematic error
  - Define functions of series of Jacobi polynomials multiplying their orthogonality relationship's metric function
- Choose a set of prior distributions for your parameters
  - Normal distributions for the linear coefficients
  - Log-Normal distributions for the non-linear
- Vary parameterizations and priors to study model dependence and overfitting
- Use AICc to create a weighted average of believable models

Conclusions and Outlook

- Jacobi polynomial parameterizations allow for a systematically controlled determination of the PDF
  - With more ensembles, other systematics can be included in the same fashion
  - Pion mass dependence, finite volume, perturbative truncation
  - Can be used with different observables
- Parameterizing in x space allows for Ioffe time functions which decay to 0 at large Ioffe time
  - Also avoids intermediate matching between pseudo-ITD and ITD in our previous works
- We have studied a range of the number of parameters to attempt to handle model dependence with AIC/AICc averaging
  - Truly distinct models, parametric and non-parametric, are required to completely remove remaining model dependent biases

Extra slides

Studying Prior Distributions

- Prior distributions must be chosen to not introduce significant biases
- If prior distributions are creating significant biases, then χ^2 and L^2 will differ
- Increasing(decreasing) widths of prior distributions will generally increase(decrease) both χ^2 and L^2

Studying Prior Distributions

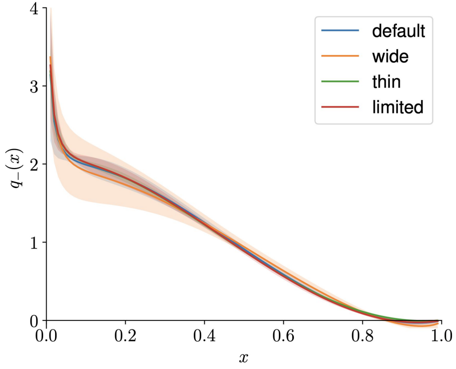
- Prior distributions must be chosen to not introduce significant biases

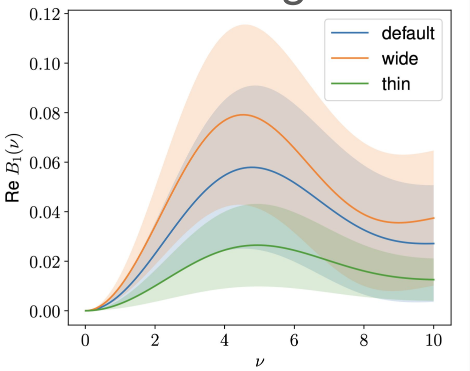
name	$N_{\pm}$	$N_{R/I,b}$	$N_{R/I,r}$	$N_{R/I,p}$	α_0	σ_α	β_0	σ_β	d_0	σ_d	c_0	σ_c
default	2	1	1	1	0	0.4	3	1	0	0.5	0	0.1
wide	2	1	1	1	0	0.8	3	2	0	1	0	0.5
thin	2	1	1	1	0	0.2	3	0.5	0	0.25	0	0.05
limited	2	0	0	1	0	0.4	3	1	0	0.5	0	0.1

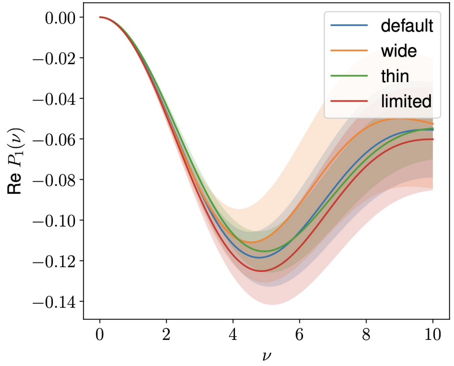
name	Real L^2/d.o.f.	Real χ^2/d.o.f.	Imag L^2/d.o.f.	Imag χ^2/d.o.f.
default	0.857	0.750	1.027	0.944
wide	0.726	0.708	0.899	0.893
thin	1.281	0.966	1.415	1.168
limited	0.876	0.809	1.187	1.148

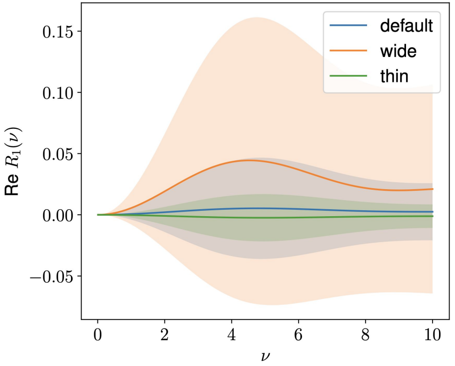
Studying Prior Distributions

- Prior distributions must be chosen to not introduce significant biases









Studying Prior Distributions

- Prior distributions must be chosen to not introduce significant biases