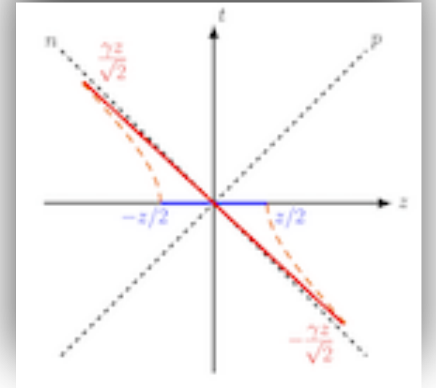
Controlling Systematic Uncertainties in Lattice QCD Calculations of Proton Unpolarized GPDs



Presented by -

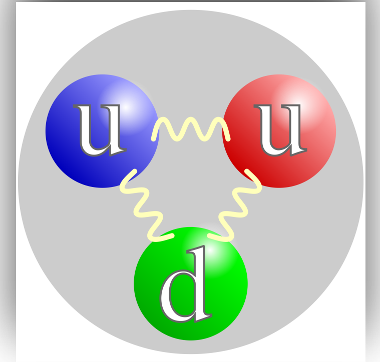
Arif Sarkar

PhD student (2nd Year)

Adam Mickiewicz University in Poznań

Supervised by - **Prof. Krzysztof Cichy**

In collaboration with:
Dr. Min-Huan Chu and **Manuel Colaço**



Outline:

1. Core Framework & Lattice Setup
    A. Generalized Parton Distributions (GPDs) and Mellin Moments
    B. Lattice extraction of Bare Nucleon Matrix Elements

2. Renormalization & Matching
    A. Non-perturbative Renormalization: The Double Ratio Scheme
    B. Short-Distance Factorization & Light-Cone Matching
3. t-dependent Parameterization via Artificial Neural Networks
4. Control of Systematics: Excited State Contamination

A work in progress…

Generalized Parton Distributions:

GPDs are important in the context of the study of strong interactions. These are non perturbative in nature.

GPDs for quarks are defined on light cone coordinates ($z^+ = 0 \ \& \ z_\perp = 0$) as:

$$\tilde{F}(x, t, \xi) = \frac{1}{2}\int \frac{dz^-}{2\pi} e^{-ixP^+z^-} \langle p_2 \,|\, \bar{\psi}^q(-\frac{z^-}{2})\gamma^+ \mathscr{W}(-\frac{z^-}{2}, \frac{z^-}{2})\psi^q(\frac{z^-}{2}) \,|\, p_1 \rangle$$

$$= \frac{1}{2P^+}\bar{u}(p_2)\left(\gamma^+ H^q(x, t, \xi) + \frac{i\sigma^{+\nu}\Delta_\nu}{2M} E^q(x, t, \xi)\right) u(p_1)$$

With P = $\frac{p_2 + p_1}{2}$, $\Delta = \overrightarrow{p_f} - \overrightarrow{p_i}$, $\xi = -\Delta^+/2P^+$ and $z^\pm = \frac{1}{\sqrt{2}}(z^0 \pm z^3)$

Ref. C. Mezrag, Few-Body Syst. 63, 62 (2022).

Local OPE of Non-Local Quark Bilinears

Leading twist (twist 2) expansion of non-local matrix elements —

local operators

$$\bar{q}\left(-\frac{z}{2}\right)\gamma^{\rho}\mathscr{W}\left(-\frac{z}{2},\frac{z}{2}\right)q\left(\frac{z}{2}\right)=\sum_{j=0}^{\infty}\frac{1}{j!}z^{\mu_1}z^{\mu_2}\dots z^{\mu_j}\left[\bar{q}(0)\gamma^{\rho}\overleftrightarrow{D}_{\mu_1}\overleftrightarrow{D}_{\mu_2}\dots\overleftrightarrow{D}_{\mu_j}q(0)\right]$$

The Mellin Moments: $\mathscr{M}(\xi,t)=\int_{-1}^{1}dx\,x^{m}H(x,\xi,t)$

The Mellin Moment are related to the local operator matrix elements via:

Robert Hjalmar Mellin
1854 - 1933

Local OPE of Non-Local Quark Bilinears

Leading twist (twist 2) expansion of non-local matrix elements —

$$\bar{q}\left(-\frac{z}{2}\right)\gamma^\rho \mathscr{W}\left(-\frac{z}{2},\frac{z}{2}\right) q\left(\frac{z}{2}\right) = \sum_{j=0}^{\infty}\frac{1}{j!} z^{\mu_1} z^{\mu_2}\dots z^{\mu_j}\overbrace{\left[\bar{q}(0)\gamma^\rho \overleftrightarrow{D}_{\mu_1}\overleftrightarrow{D}_{\mu_2}\dots\overleftrightarrow{D}_{\mu_j} q(0)\right]}^{\text{local operators}}$$

The Mellin Moment are related to the local operator matrix elements via:

$$\left\langle p_2 \middle| \bar{\psi}^q(0)\gamma^\rho \overleftrightarrow{D}^{\mu_1}\dots\overleftrightarrow{D}^{\mu_j}\psi^q(0) \middle| p_1 \right\rangle = P^{\{\rho}\sum_{i=1}^{m+1} P^{\mu_1}\dots P^{\mu_{i-1}}\Delta^{\mu_i}\dots\Delta^{\mu_m\}} A^q_{i,m}(t) + \Delta^\rho \Delta^{\mu_1}\dots\Delta^{\mu_m} C^q_{m+1}(t)$$

Generalized Form Factors

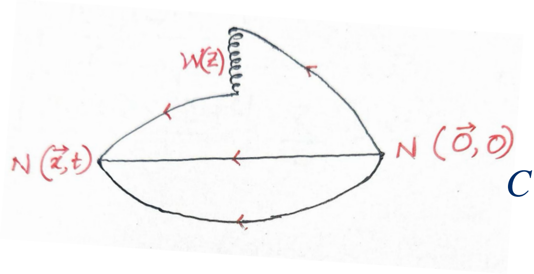
Polynomiality Relation:

$$\int_{-1}^{1} dx\, x^m H^q(x,\xi=0,t) = \sum_{\substack{k=0\\ \text{even}}}^{m} A^q_{m+1,k}(t)\ , \qquad \int_{-1}^{1} dx\, x^m E^q(x,\xi=0,t) = \sum_{\substack{k=0\\ \text{even}}}^{m} B^q_{m+1,k}(t)$$

Ref. C. Mezrag, Few-Body Syst. 63, 62 (2022).

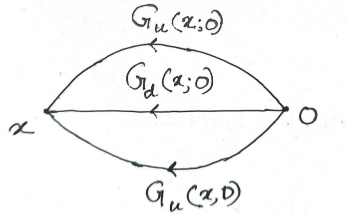
Towards Lattice calculations of (bare) GPDs:

The nucleon two- and three- point functions can be calculated from lattice:



$$C^{2pt}(\Gamma, t) = \Gamma_{\alpha,\beta} \sum_{\vec{x}} e^{-i\vec{P}\cdot\vec{x}} \langle 0 | N_\alpha(\vec{x}, t) \bar{N}_\beta(\vec{x}, t) | 0 \rangle$$

$$C^{3pt}(\Gamma, z, p_f, p_i, t, \tau) = \Gamma_{\alpha\beta} \sum_{\vec{x},\vec{y}} \langle 0 | N_\alpha(\vec{x}, t) \mathcal{O}(\vec{y}, \tau; z) \bar{N}_\beta(\vec{0}, 0) | 0 \rangle$$



Ref: K. Cichy and M. Constantinou, Adv. High Energy Phys. (2019) 3036904

The Matrix Elements(MEs) are calculated by using ratio:

$$R_\mu(\Gamma_\kappa, z, P_f, P_i; t_s, \tau) = \frac{C^{3\mathrm{pt}}_\mu(\Gamma_\kappa, z, P_f, P_i; t_s, \tau)}{C^{2\mathrm{pt}}(\Gamma_0, P_f; t_s)} \sqrt{\frac{C^{2\mathrm{pt}}(\Gamma_0, P_i, t_s - \tau) C^{2\mathrm{pt}}(\Gamma_0, P_f, \tau) C^{2\mathrm{pt}}(\Gamma_0, P_f, t_s)}{C^{2\mathrm{pt}}(\Gamma_0, P_f, t_s - \tau) C^{2\mathrm{pt}}(\Gamma_0, P_i, \tau) C^{2\mathrm{pt}}(\Gamma_0, P_i, t_s)}},$$

1. **Ground State Dominant** MEs $\rightarrow 0 \ll \tau \ll t_s$
2. From here MEs of GPD are calculated via Lorentz invariant Decompositions as mentioned in Eqn. 73 to 88 of: *S. Bhattacharya et al., PRD 106 (2022) 114512.*

Lattice calculations of GPDs - Asymmetric Frame

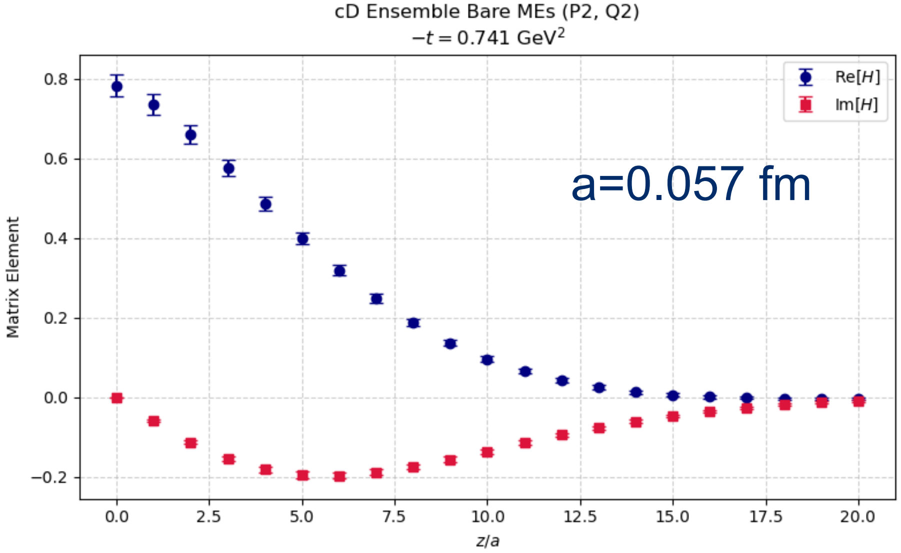
The MEs are calculated in asymmetric frame, defined as:

$$p_f = \overrightarrow{P} = (0,\ 0,\ P_3) \qquad p_i = \overrightarrow{P} - \overrightarrow{\Delta} = (\Delta_1\ \ \Delta_2,\ P_3)$$

Zero-skewness → $\Delta_3 = 0$

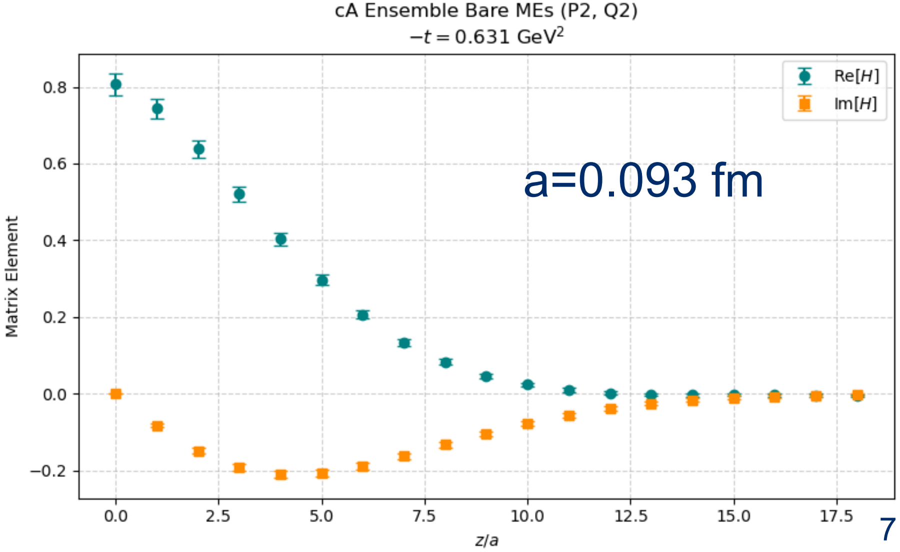
MEs for several *t*-values can be computed at once → computational advantage over symmetric frame.

$$\boxed{-t = \overrightarrow{\Delta}^2 - [E(p_f) - E(p_i)]^2}$$



cD Ensemble Bare MEs (P2, Q2)
$-t = 0.741\ \mathrm{GeV}^2$

a=0.057 fm



cA Ensemble Bare MEs (P2, Q2)
$-t = 0.631\ \mathrm{GeV}^2$

a=0.093 fm

Double Ratio Scheme Renormalization:

bare MEs exhibit UV divergences:

$$h^B(z, P_3, a) = e^{\delta m|z|} Z(a) h^R(z, P_3, a)$$

Ref: X. Ji et al., PRL 120 (2018) 112001

The **RG invariant**, double ratio:

$$\mathscr{M}(z, P_3, P_3^0) = \frac{h^B(z, P_3, a)}{h^B(z, P_3^0, a)} \frac{h^B(0, P_3^0, a)}{h^B(0, P_3, a)}$$

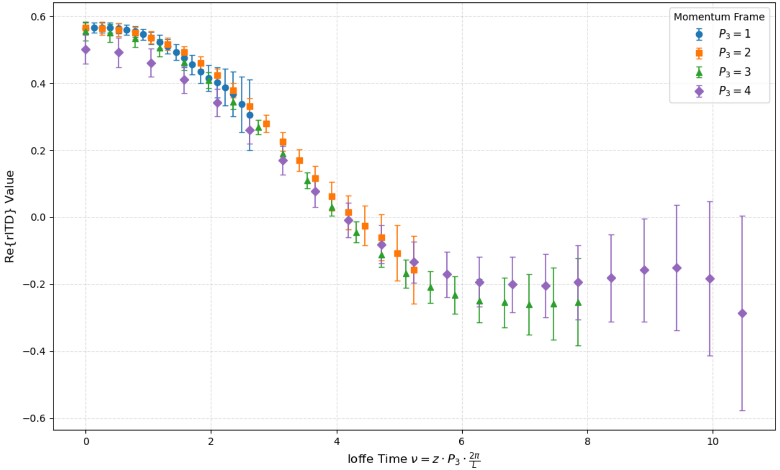
This Double Ratio ME is named as reduced **ioffee-time** distribution(rITD) —

- Lorentz invariant → $\mathscr{M}(z, P_3, P_3^0) \sim \mathscr{M}(z^2, zP_3, zP_3^0) = \mathscr{M}(z^2, \nu, \nu_0)$
- Cancels some **discretization effects** and **higher twist effects**.

Refs
1. H. Dutrieux et al. (HadStruc Collaboration), JHEP. 2024, 162 (2024)
2. X. Gao et al., PRD 102 (2020) 094513

rITDs at different lattice spacing (H_u-d)

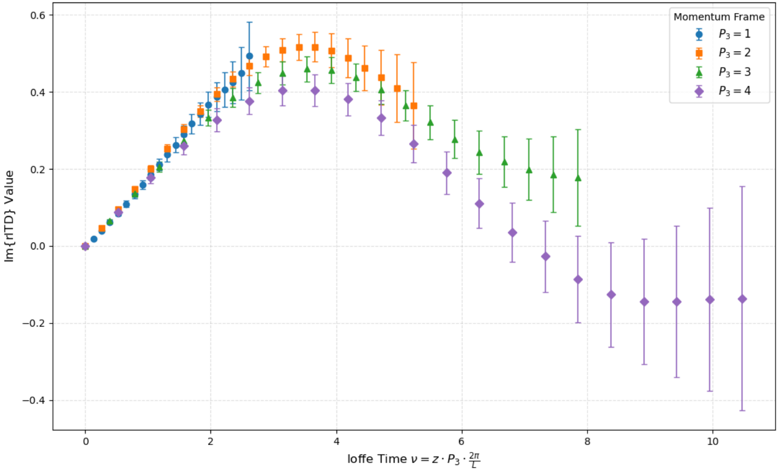
Real Part: Reduced ITD Data vs. Ioffe Time ν ($H[u-d]$ @ Q2)



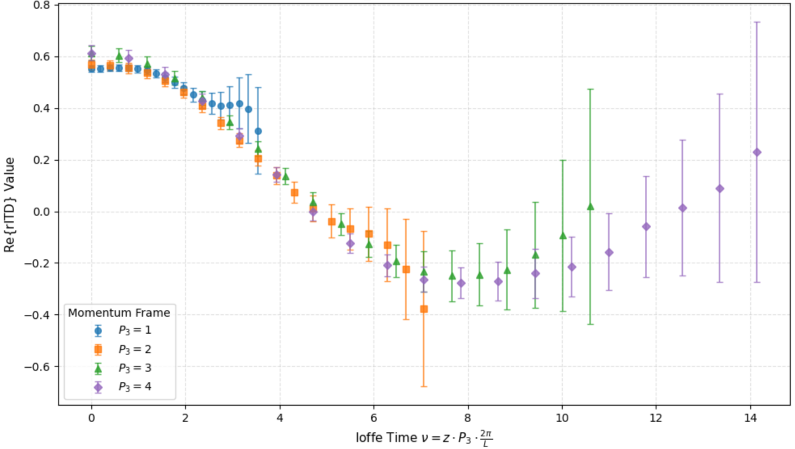
a = 0.057 fm

P_3
0.4448 GeV
0.8896 GeV
1.3344 GeV
1.7793 GeV

Imaginary Part: Reduced ITD Data vs. Ioffe Time ν ($H[u-d]$ @ Q2)



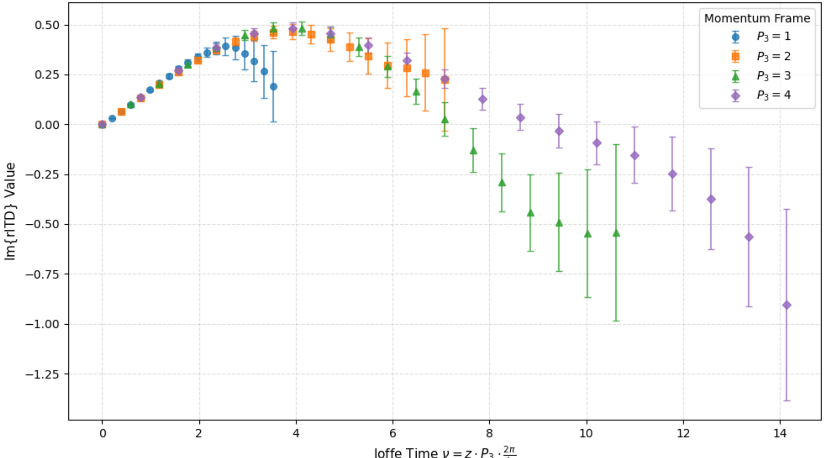
Real Part: H_u-d Data Distribution @ Q2



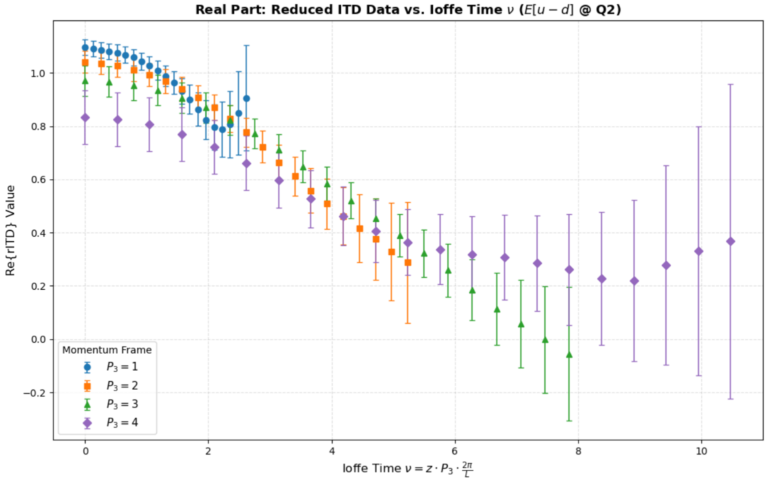
a = 0.093 fm

P_3
0.4089 GeV
0.8179 GeV
1.2268 GeV
1.6358 GeV

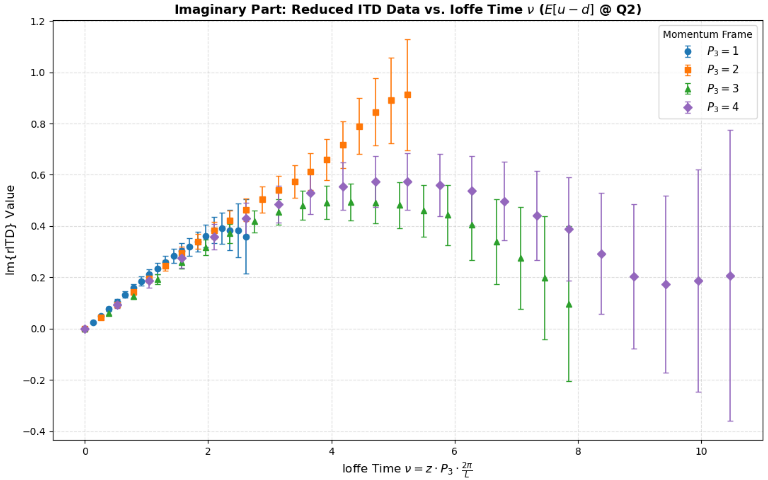
Imaginary Part: H_u-d Data Distribution @ Q2

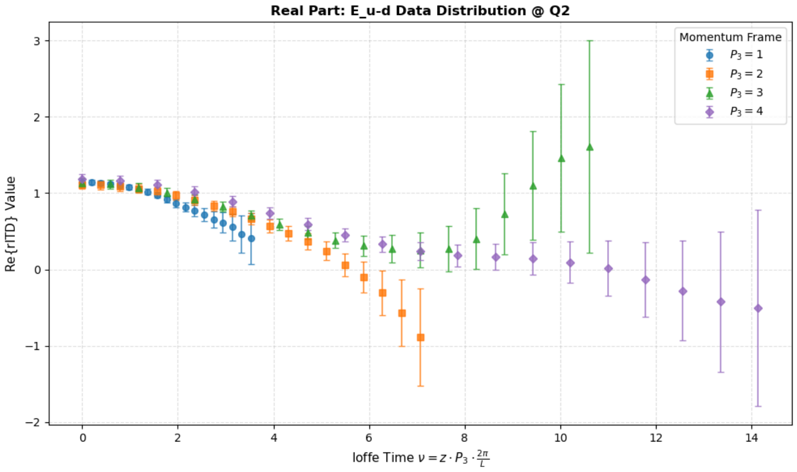


rITDs at different lattice spacing (E_u-d)

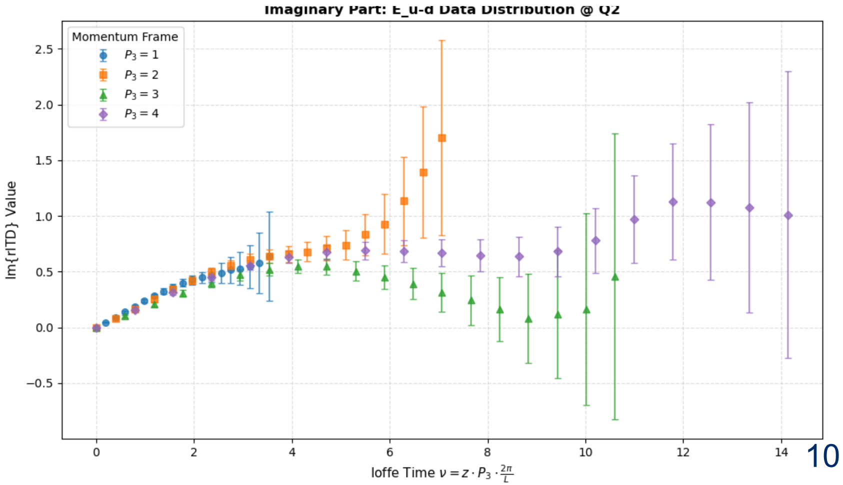

Real Part: Reduced ITD Data vs. Ioffe Time ν ($E[u-d]$ @ Q2)

A = 0.057 fm

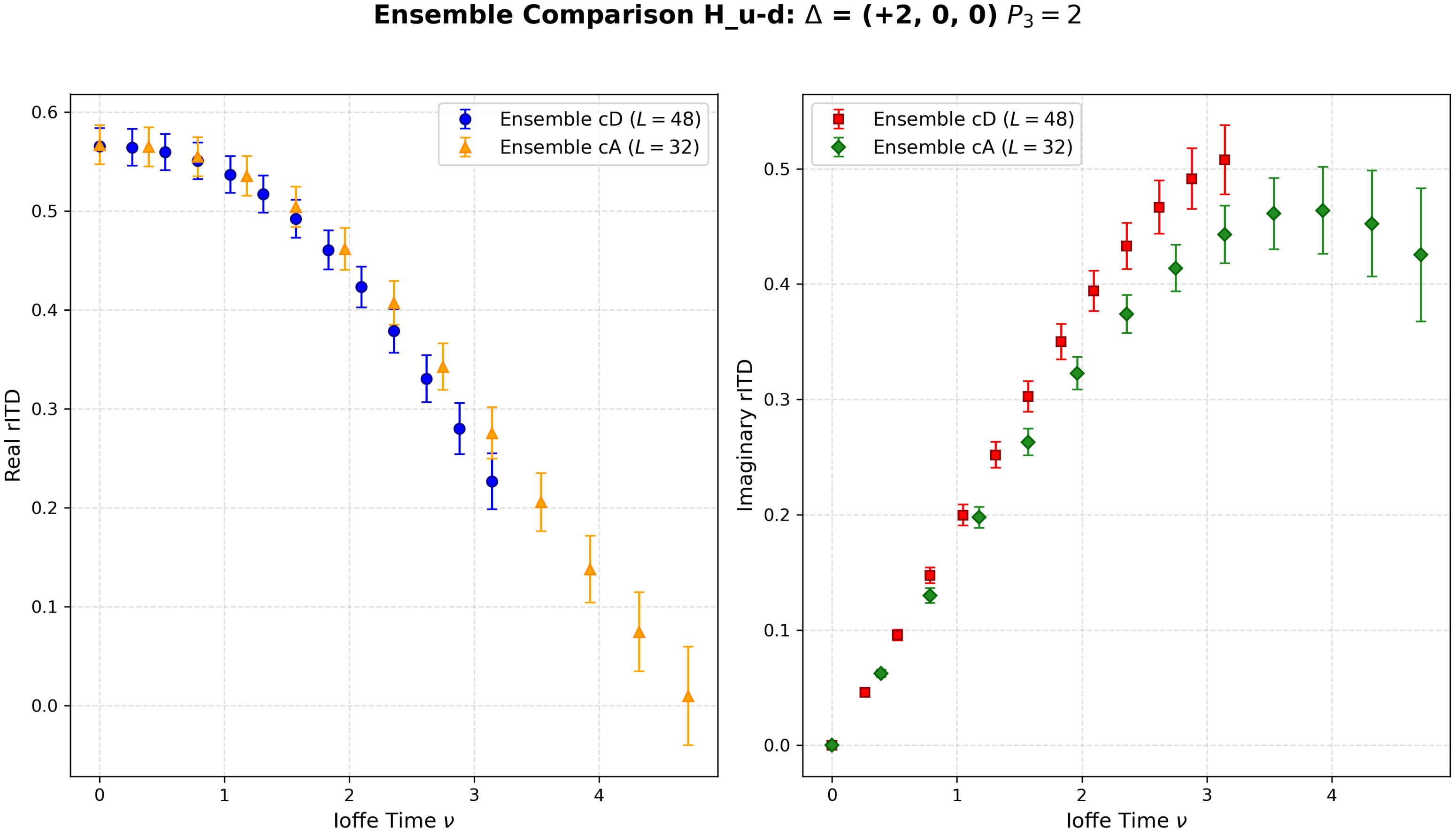

Imaginary Part: Reduced ITD Data vs. Ioffe Time ν ($E[u-d]$ @ Q2)


Real Part: E_u-d Data Distribution @ Q2

A = 0.093 fm

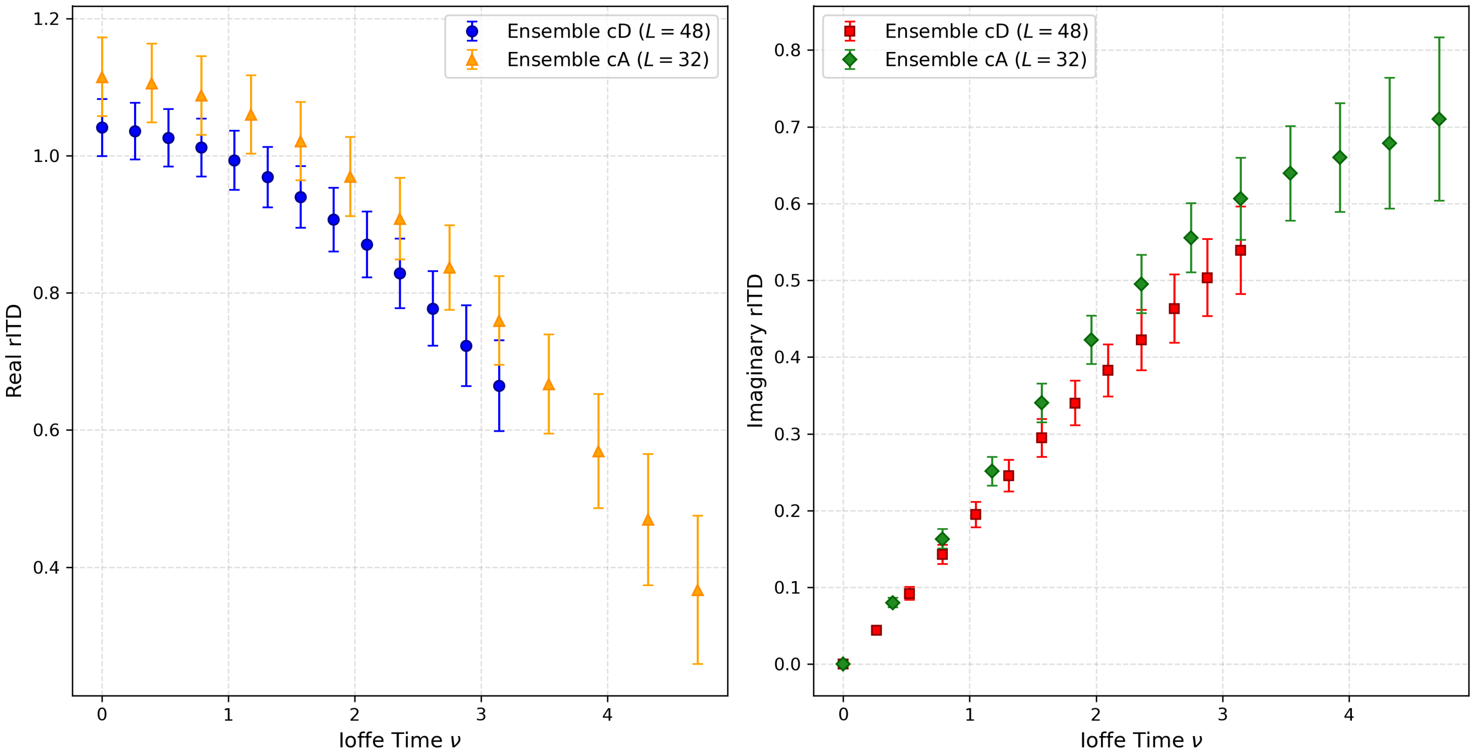

Imaginary Part: E_u-d Data Distribution @ Q2

Comparison of rITDs at different lattice spacing (H_u-d)

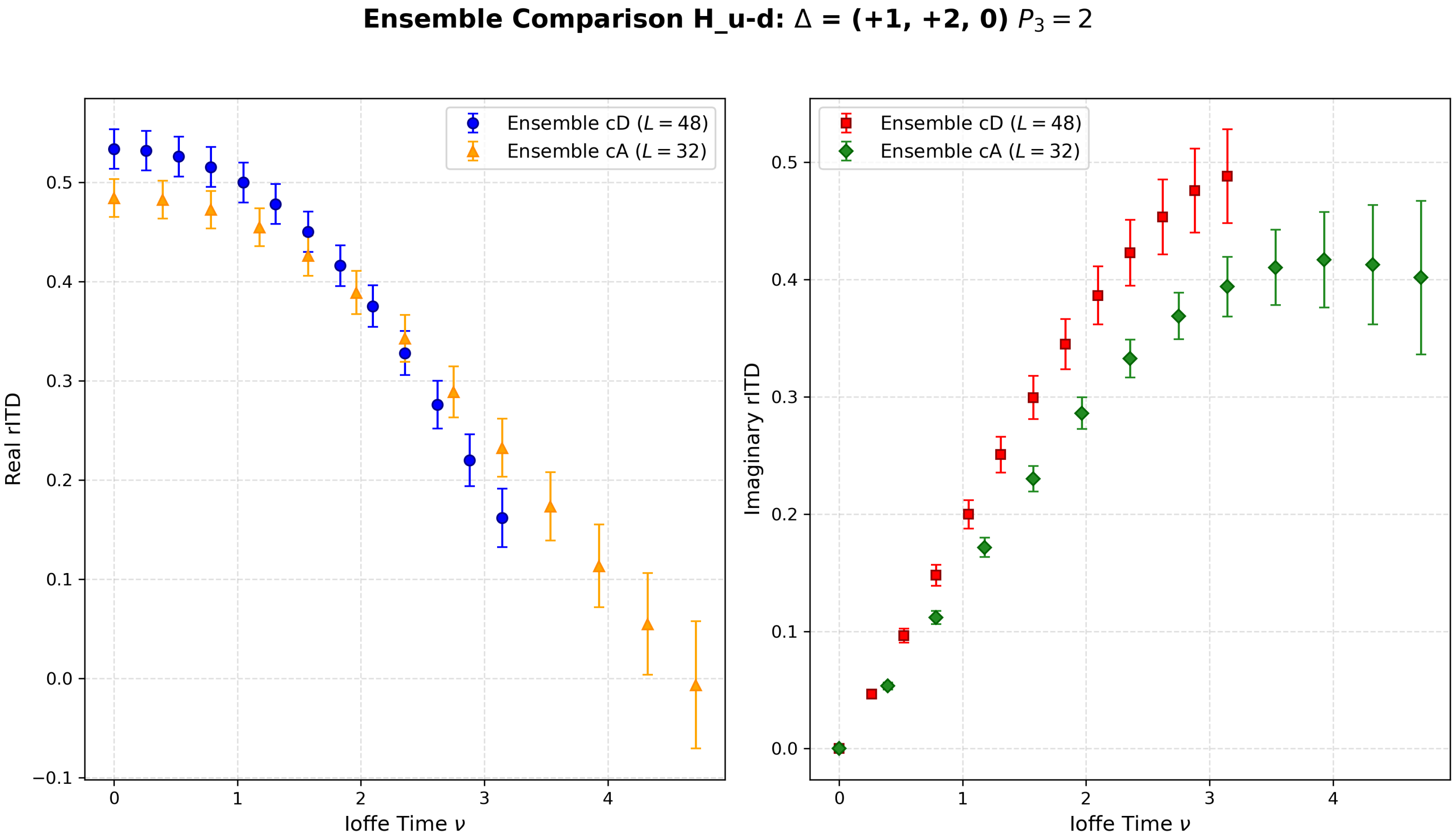


Comparison of rITDs at different lattice spacing (E_u-d)

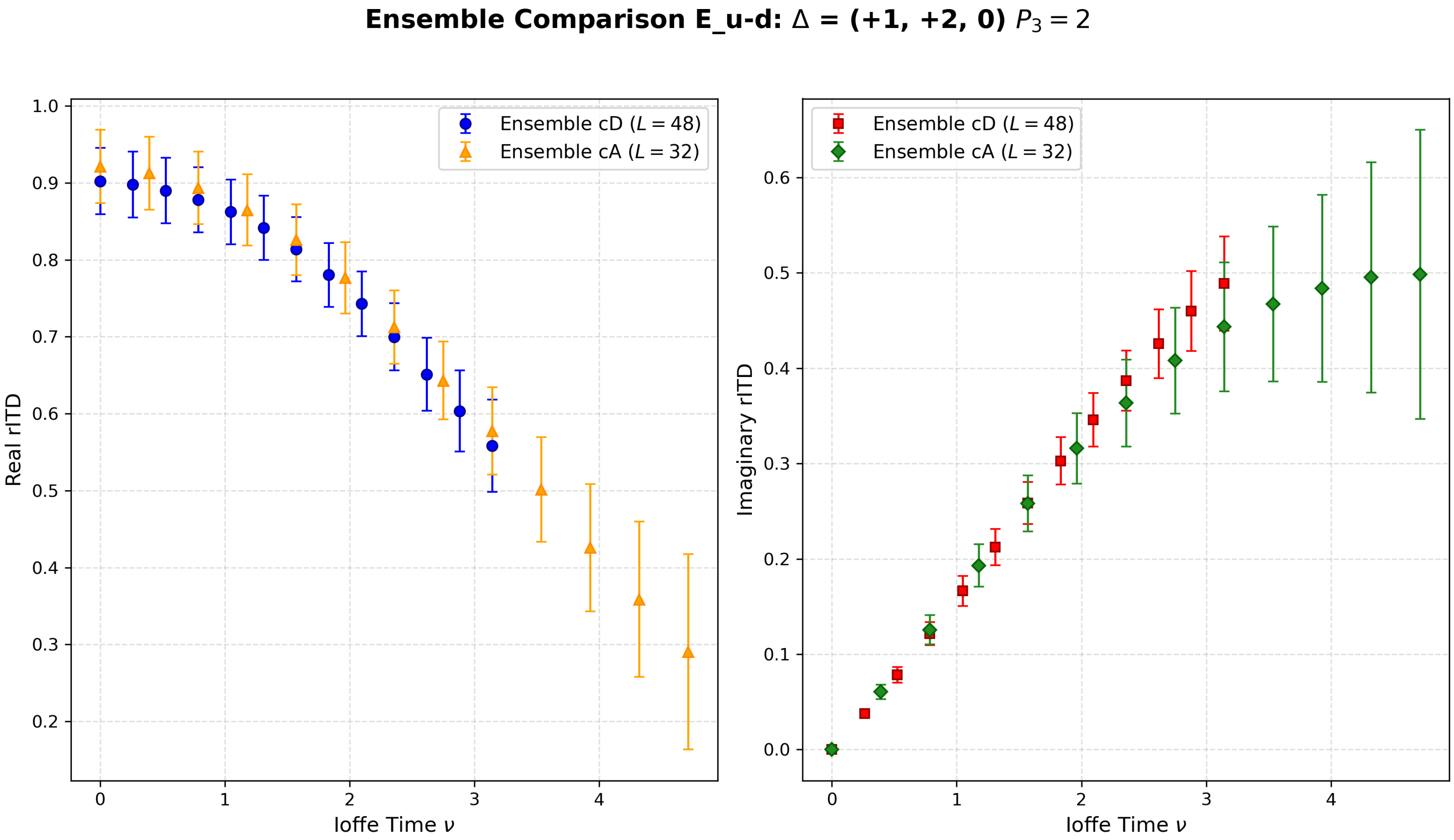
Ensemble Comparison E_u-d: Δ **= (+2, 0, 0)** $P_3 = 2$



Comparison of rITDs at different lattice spacing (H_u-d)



Comparison of rITDs at different lattice spacing (E_u-d)


Ensemble Comparison E_u-d: Δ **= (+1, +2, 0)** $P_3 = 2$

Matching to Light-Cone Coordinates:

Matching to the light-cone is performed at the level of moments. the rITD moments are:

$$\mathscr{M}(\nu, \xi, t, z^2) = \sum_{n=0}^{\infty} \frac{(i\nu)^n}{n!} \mathscr{M}_{n+1}(\xi, t, z^2)$$

The rITD moments are matched to the matched ITD moments via perturbative kernel $C_{n+1,n-k}(\alpha_s, z^2\mu^2)$:

$$(-i)^n \mathscr{M}_{n+1}(\xi, t, z^2) = (-i)^n \sum_{k=0}^{n} \bar{F}_{k+1}(\xi, t, \mu) \xi^{n-k} C_{n+1,n-k}(\alpha_s, z^2\mu^2) + \mathcal{O}(z^2\Lambda^2_{QCD})$$

—known as **Short Distance Factorization**.

- My calculation involves $\xi = 0$ and fixed (next leading) order matching.
- Target scale = 2 GeV and α_s = 0.393945

Refs
1. H. Dutrieux et al. (HadStruc Collaboration), JHEP. 2024, 162 (2024)
2. M. Bhat et al., PRD 103 (2021) 034510.

Can we construct Matched ITD Moments as a Function of *t* ?

The moments of Matched ITDs in my calculations are represented using **Artificial Neural Networks (ANNs)**.

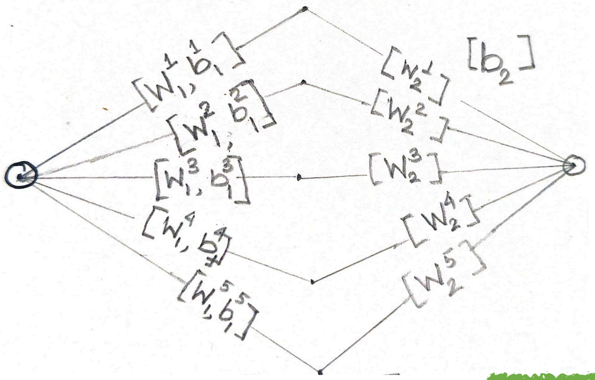
Ref: M.-H. Chu et al., JHEP 2026 (2026) 210

Qxy = $\left(x \cdot \frac{2\pi}{aL}, y \cdot \frac{2\pi}{aL}, 0\right)$

Pz = $z \cdot \frac{2\pi}{aL}$

$$\begin{bmatrix} t_{11}(Q1,P1) & t_{12}(Q2,P1) & \cdots & t_{18}(Q22,P1) \\ t_{21}(Q1,P2) & t_{22}(Q2,P2) & \cdots & \\ t_{31}(Q1,P3) & \vdots & & \\ t_{41}(Q1,P4) & & & t_{48}(Q22,P4) \end{bmatrix}$$

input



output

hidden
(5 neurones)

$$\mathrm{Re}\left[\overline{\mathscr{M}}\left(\nu, \xi, t, \mu^2\right)\right] = \bar{F}_1\left(\nu, \xi, t; \mu^2\right) + \frac{\nu^2}{2}\bar{F}_3\left(\nu, \xi, t; \mu^2\right)$$

$$\mathrm{Im}\left[\overline{\mathscr{M}}\left(\nu, \xi, t, \mu^2\right)\right] = \nu\bar{F}_2\left(\nu, \xi, t; \mu^2\right) - \frac{\nu^3}{6}\bar{F}_4\left(\nu, \xi, t; \mu^2\right)$$

4 such networks corresponding to $\bar{F}_1(t), \bar{F}_2(t), \bar{F}_3(t), \bar{F}_4(t)$.

How does ANN work? Why Should I use it?

The Architecture: $ANN(t) = \sum_{j=1}^{5} W_2^j \cdot \tanh\left[W^i t + b_1^i\right]^j + b_2$

Random guess values of [W1, b1, W2, b2] and t → ANN prediction → calculate χ^2:

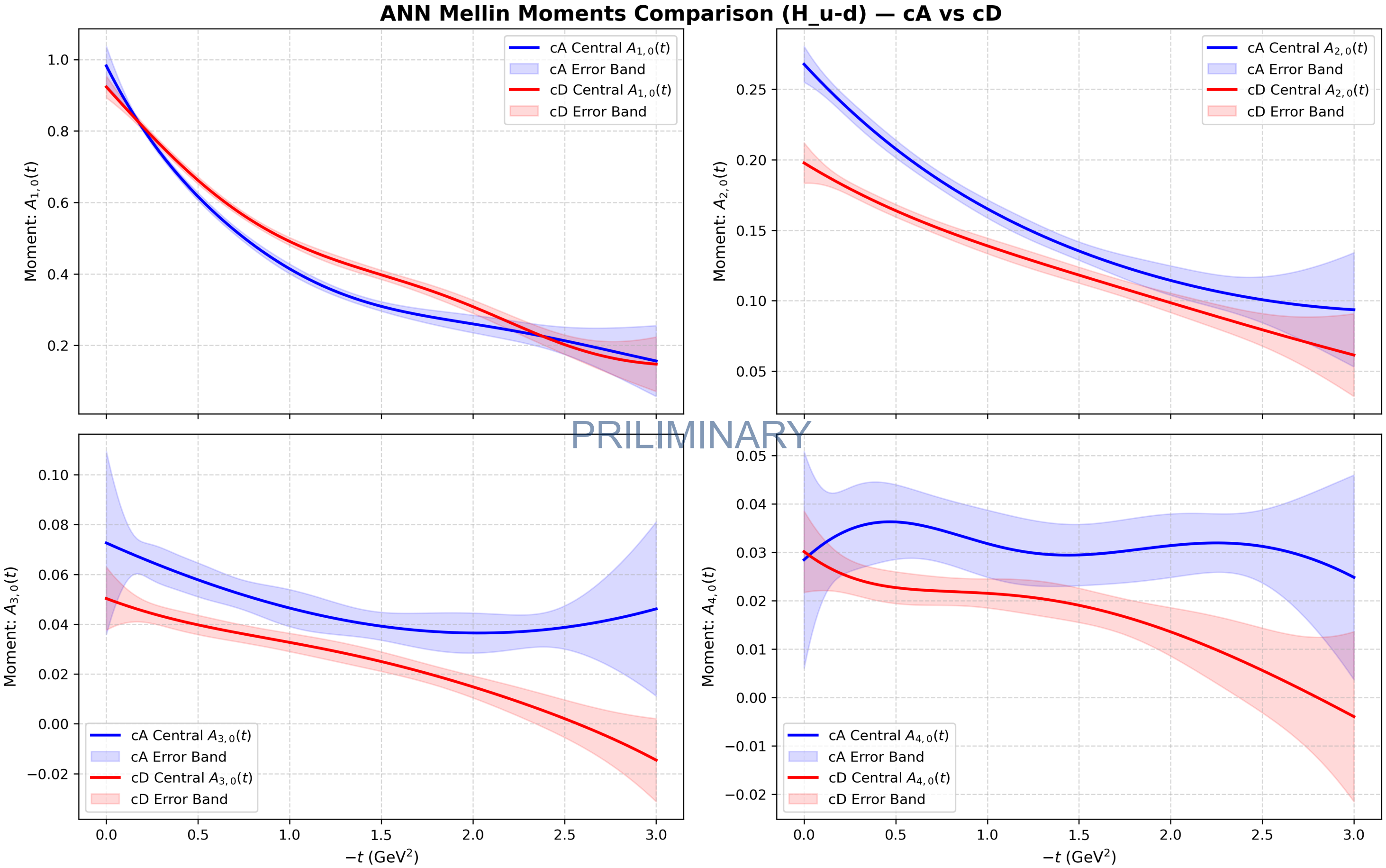
$$\chi^2(W_1, b_1, W_2, b_2) = \sum_{n=0}^{n_{data}} \frac{\left[\mathscr{M}_{ANN}(t \,|\, W_1, b_1, W_2, b_2) - \mathscr{M}_{data}(t)\right]^2}{\sigma_{\mathscr{M}}^2}$$

Gradient Based Optimizers (Adam) is used to train the parameters such that χ^2 is minimum.

$$\theta_{n+1} \to \theta_n - \eta \nabla \chi^2(\theta_n)$$

1. Once the parameters are trained, ANN becomes a “**continuous**” function of *t*.
2. Takes **full t-dependence into account**, (→Manuel's talk).
3. **Model Independent** — highly suitable for taking forward limit.

Comparison of Moments for different lattice spacings

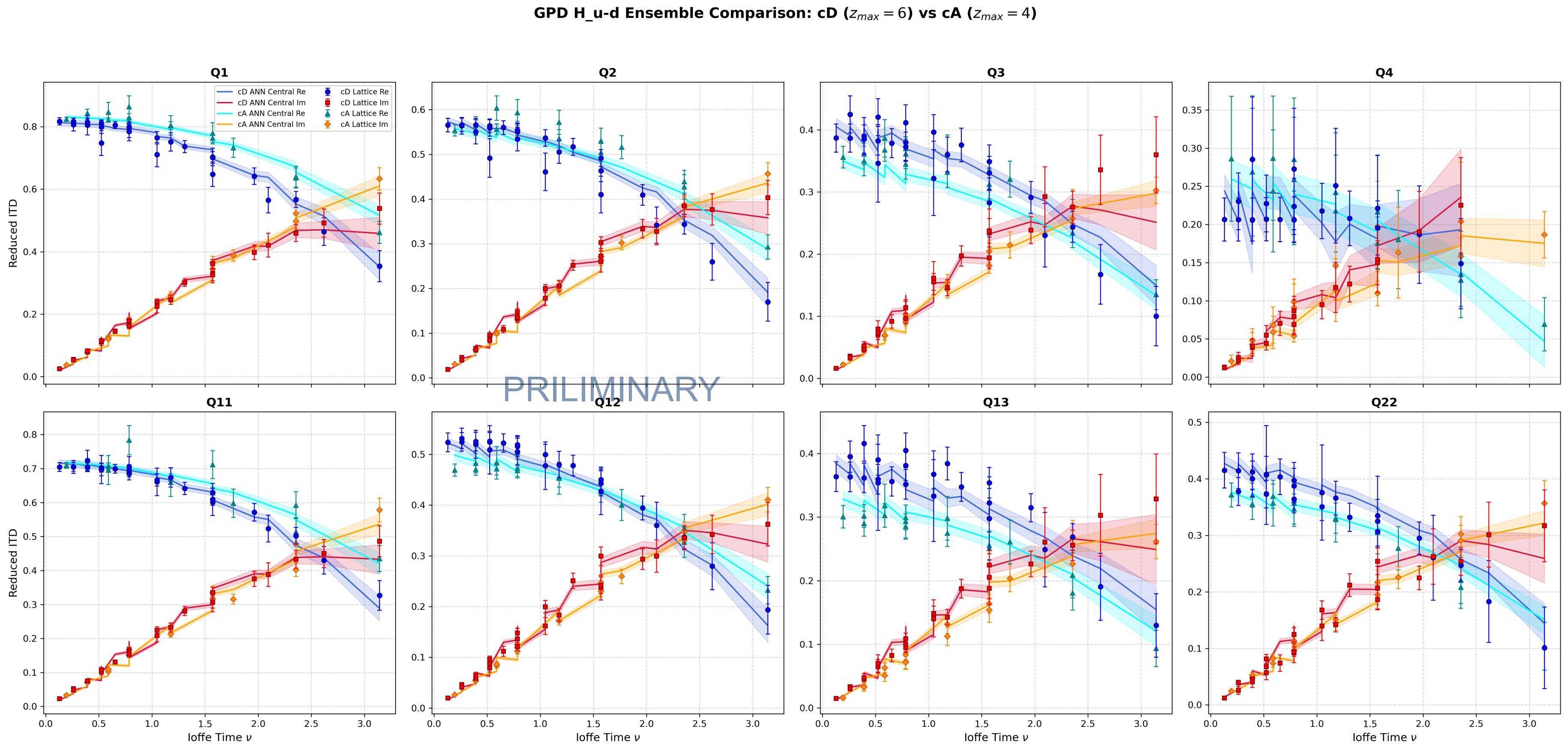


Forward Limit of H GPD, u-d flavor

	$A_{10}(0)$	$A_{20}(0)$
[5]	0.982 (47)	0.267 (13)
0.093 fm (cA)	0.982 (54)	0.268 (12)
0.057 fm (cD)	0.923 (30)	0.198 (14)

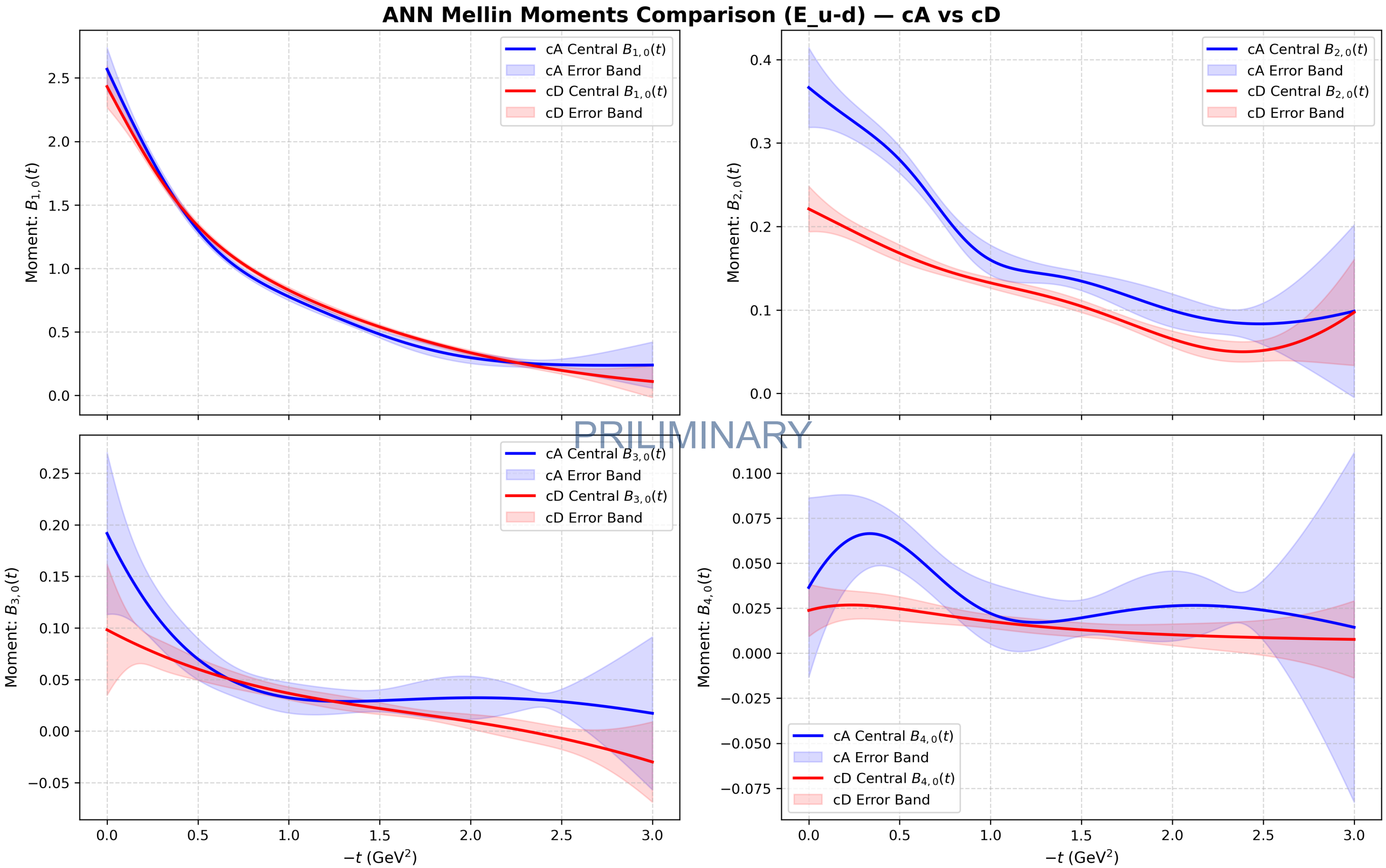
ETMC' 23: S. Bhattacharya et al., PRD 108 (2023) 014507

Fitting data with ANNs (lattice spacings [0.093fm, 0.057fm])



GPD H_u-d Ensemble Comparison: cD ($z_{max} = 6$) vs cA ($z_{max} = 4$)

Comparison of Moments for different lattice spacings



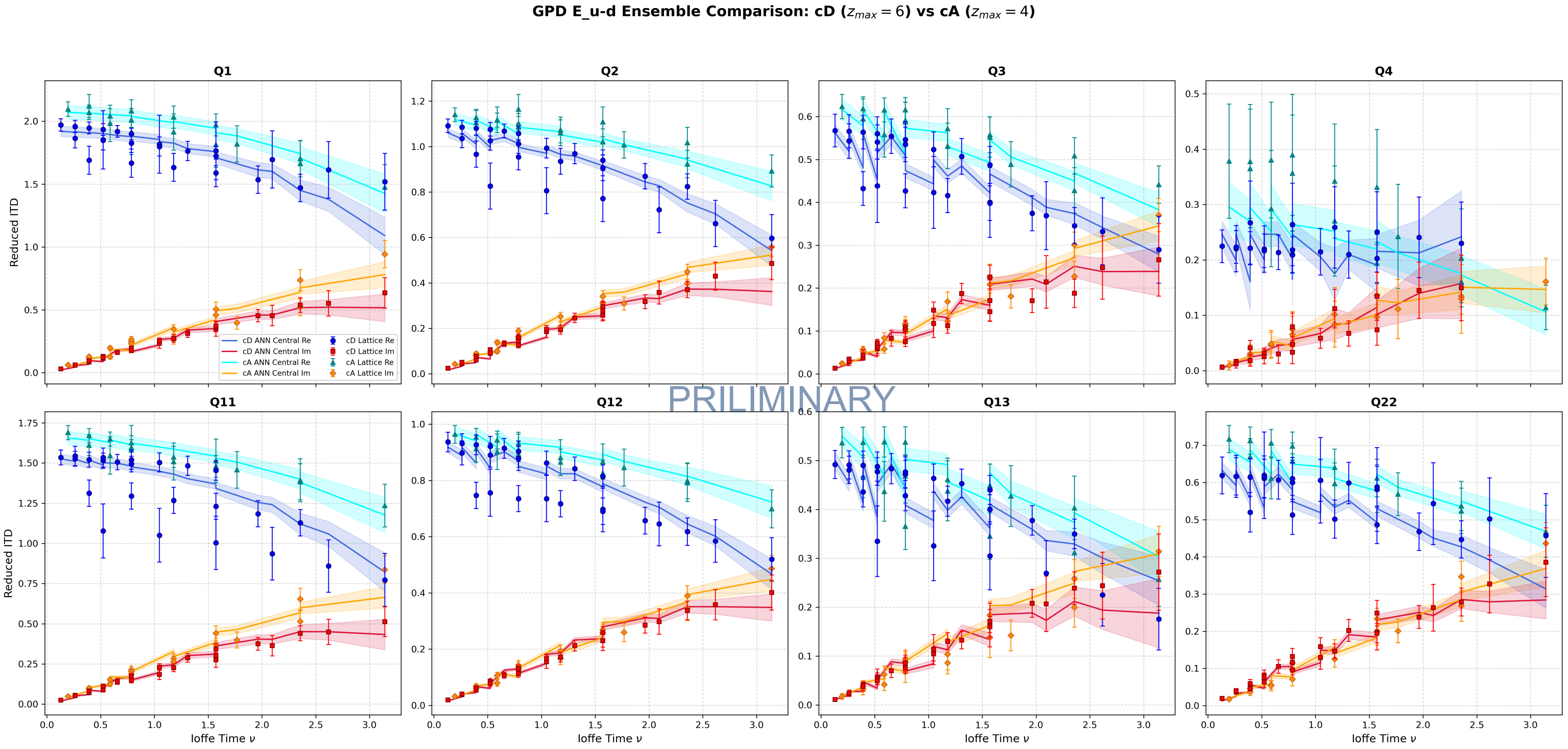
PRILIMINARY

Forward Limit of E GPD,
u-d flavor

	$B_{10}(0)$	$B_{20}(0)$
[5]	2.540 (221)	0.295 (38)
0.093 fm (cA)	2.569 (169)	0.366 (47)
0.057 fm (cD)	2.432 (164)	0.221 (27)

ETMC' 23: S. Bhattacharya et al.,
PRD 108 (2023) 014507

Fitting data with ANNs (lattice spacings [0.093fm, 0.057fm])



— Excited State Analysis —

Fitting Ansatze for Excited State Analysis:

The functional forms of two- and three-point correlations —

$$C^{2pt}(\overrightarrow{P}, t) = |Z_0|^2 e^{-E_0 t}\left(1 + |f_{10}|^2 e^{-\Delta E}\right) + \text{terms(n} \geq 2)$$

$$C^{3pt}(\overrightarrow{P}, t) = |Z_0|^2 e^{-E_0 t}\left(\langle 0|\mathcal{O}|0\rangle + f_{10}\langle 1|\mathcal{O}|0\rangle e^{-\Delta E(t-\tau)} + f_{10}^{\dagger}\langle 1|\mathcal{O}|0\rangle e^{-\Delta E\tau} + |f_{10}|^2\langle 1|\mathcal{O}|1\rangle e^{-\Delta E t}\right) + \text{terms(n} \geq 2)$$

The lattice calculated ***PDF***/GPD matrix elements are complex numbers. Hence, the fitting function for ratio has also real and imaginary parts:

$$\text{Re}\left[\text{Ratio}(P_3, z, t_s)\right] = \frac{\text{Re}(A_{00}) + \text{Re}(B)\left[e^{-\Delta E(t_s-\tau)} + e^{-\Delta E\tau}\right] + |f_{10}|^2\,\text{Re}(A_{11})e^{-\Delta E t_s}}{1 + |f_{10}|^2 e^{-\Delta E t_s}}$$

$$\text{Im}\left[\text{Ratio}(P_3, z, t_s)\right] = \frac{\text{Im}(A_{00}) + \text{Im}(B)\left[e^{-\Delta E(t_s-\tau)} - e^{-\Delta E\tau}\right] + |f_{10}|^2\,\text{Im}(A_{11})e^{-\Delta E t_s}}{1 + |f_{10}|^2 e^{-\Delta E t_s}}$$

With $\text{Re}(A_{00}) + i\,\text{Im}(A_{00}) = \langle 0|\mathcal{O}|0\rangle$, $\text{Re(B)} + i\,\text{Im(B)} = f_{10}\langle 1|\mathcal{O}|0\rangle$, $\text{Re}(A_{11}) + i\,\text{Im}(A_{11}) = \langle 1|\mathcal{O}|1\rangle$

Extracting information out of C^{2pt}

Fitting Window Determination: it is essential to determine the interval $[t_{min}, t_{max}]$, where contributions are dominated by the nucleon ground state and the first excited state.

Perform Correlated Fits

Extracting information out of C^{2pt} - *The Correlated Fits:*

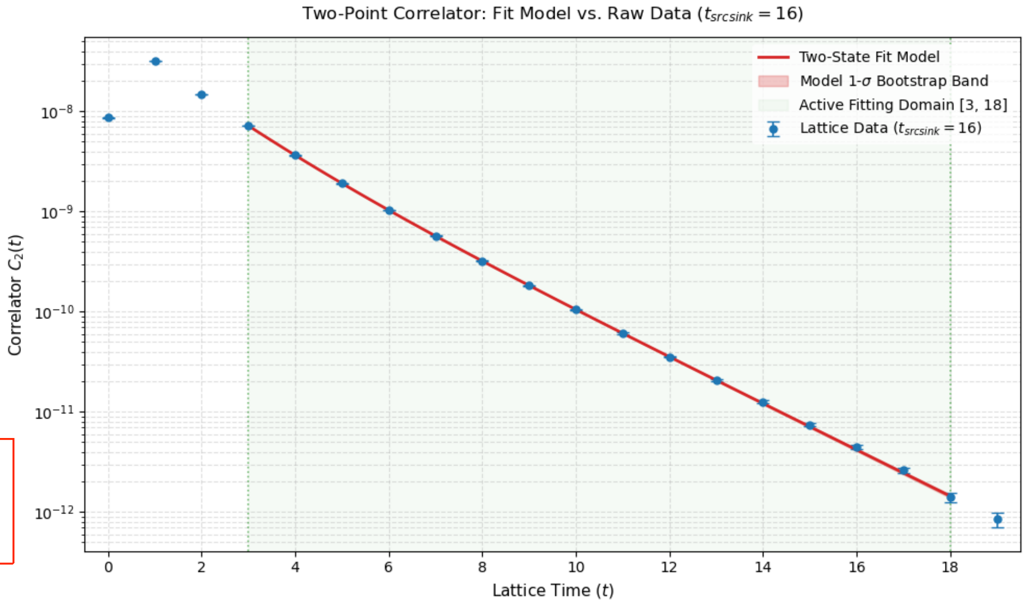
Step 1: Generalized χ^2: $\chi^2(\theta) = [C^{2pt}_{\text{ansatz}}(t|\theta) - C^{2pt}_{data}(t)]^T[Cov^{2pt}]^{-1}[C^{2pt}_{\text{ansatz}}(t|\theta) - C^{2pt}_{data}(t)]$

$$Cov^{2pt} = \frac{1}{n-1}\sum_{i=1}^{n}(\vec{c}_i - E[\vec{c}\,])(\vec{c}_i - E[\vec{c}\,])^T$$

Ref: M. Bruno and R. Sommer, Comput. Phys. Commun. 285, 108643 (2023)

Includes correlation between the data from t_i to t_j for $i \neq j$ and $i = j$.

Step 2: Generalized Chi2, is minimised with an optimizer (such as Nelder-Mead).



Reduced χ^2 normalizes to ~1,
for appropriate fitting windows, such as t $\in [3,18]$

Towards *Two State Fits* of the Ratio:

Parameters estimated from fitting C^{2pt} should be unique and stable —

$$|Z_0|^2,\ E_0,\ |f_{10}|^2,\ \Delta E = 2.00744 \times 10^{-8} \pm 1.10384 \times 10^{-9},\ 0.53013 \pm 0.00506,\ 2.40937 \pm 0.14965,\ 0.38644 \pm 0.01328$$

(measured in lattice unites, a = 0.057 fm)

$E_0 = 0.53836 \pm 0.00450$ from $E_{eff}(t)\ vs\ t$

In **sequential fit** procedure, the parameters ($|f_{10}|^2$ and ΔE) are substituted into the ratio ansatz —

$$Re[\text{Ratio}(P_3, t, z)] = \frac{1}{D(t_s)}\text{Re}(A_{00}) + \left[\frac{e^{-\Delta E(t_s-\tau)} + e^{-\Delta E\tau}}{D(t_s)}\right]\text{Re}(B) + \frac{|f_{10}|^2 e^{-\Delta E t_s}}{D(t_s)}\text{Re}(A_{11}) \text{ With } D(t_s) = \left(1 + e^{-\Delta t_s}\right)$$

My calculations involved $t_s = 10a,\ 12a,\ 14a$ and $16a$ in fm, for a = 0.057 fm.

The data that takes contributions from nucleon ground state and the first excited state lies in the range: $t_s \in [3,\ t_s - 3]$

Ref: C. Alexandrou et al., PRD 99 (2019) 114504

More on *Two State Fits* of the Ratio:

The model becomes linear in terms of parameters $\theta \in$ [Re(A_{00}), Re(B), Re(A_{11})]:

$$\text{Generalized } \chi^2(\theta) = \left[Y - X_{\text{coeff}} \cdot \theta\right]^T \left[Cov(Y)\right]^{-1} \left[Y - X_{\text{coeff}} \cdot \theta\right]$$

Generalized χ^2 can be minimized as:

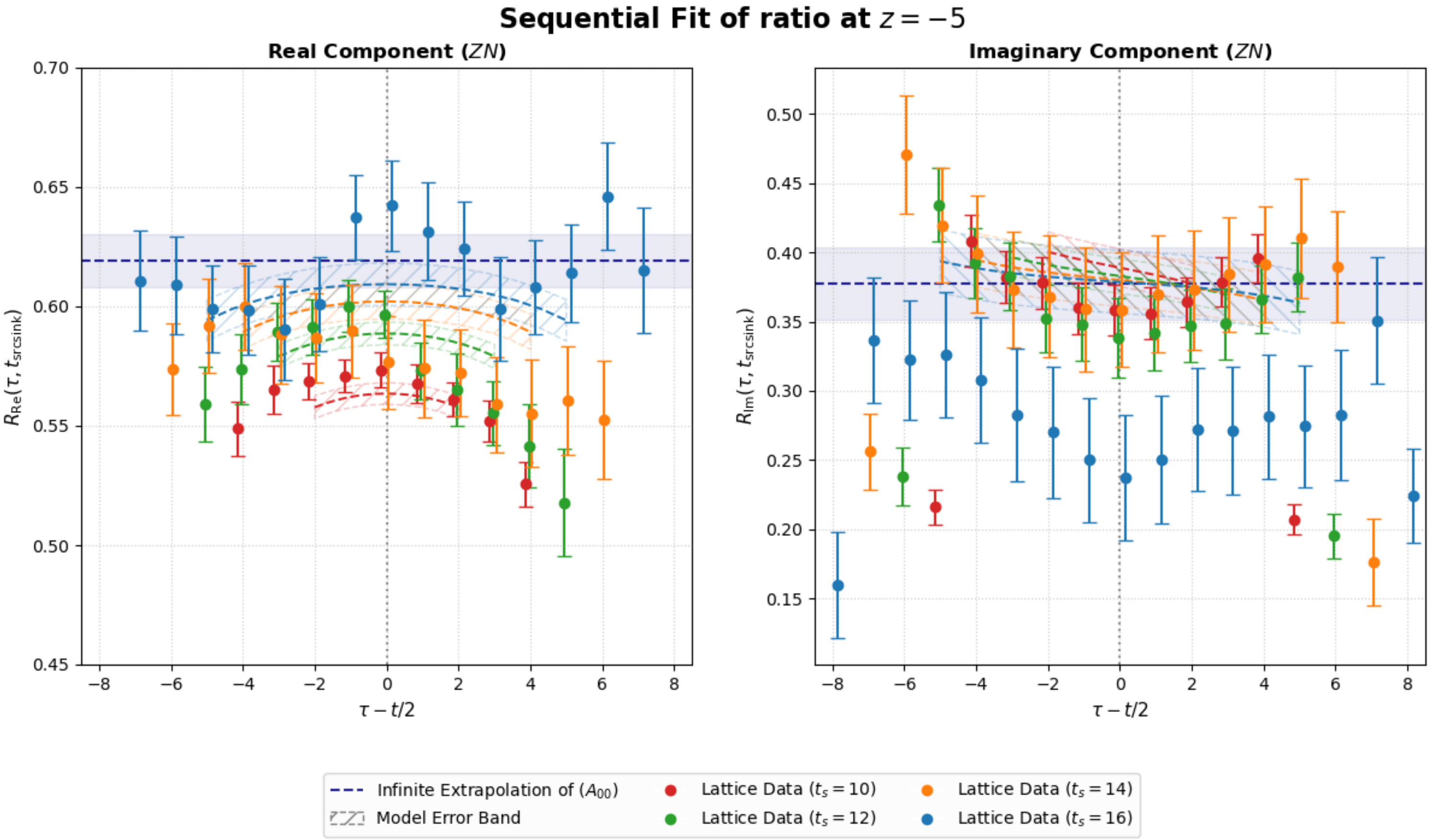
$$\boxed{\nabla_{\theta^T} \chi^2(\theta) = 0} \longrightarrow \boxed{\theta_{sol} = \left[X^T_{\text{coeff}}\, Cov(Y)\, X_{\text{coeff}}\right]^{-1} X^T_{\text{coeff}} W Y}$$

Sequential Fit —

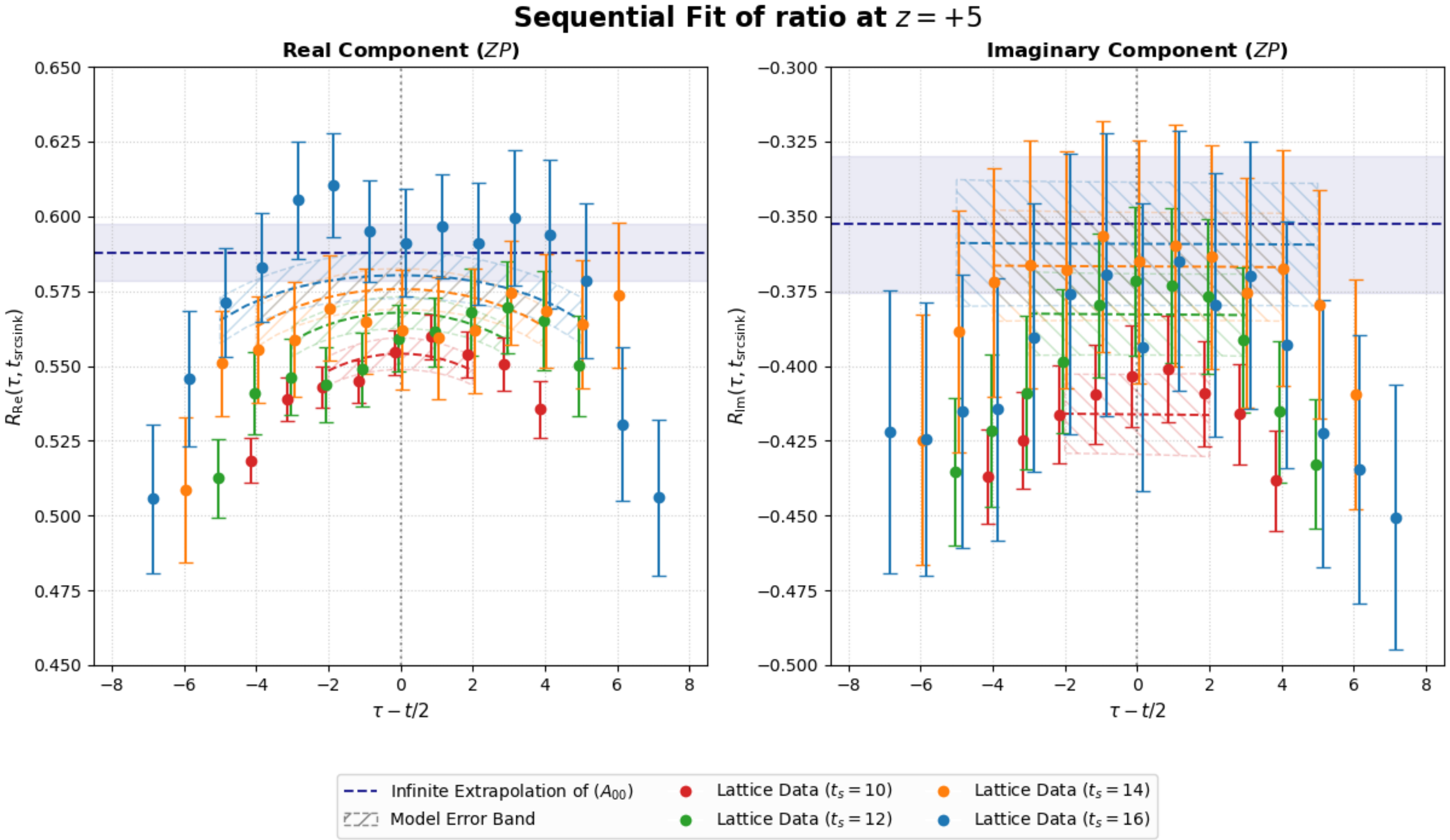
1. Solves the problem exactly and **uniquely** determines the parameters.
2. The analysis is exactly the same for real and imaginary parts.

Ref: C. Alexandrou et al., PRD 99 (2019) 114504

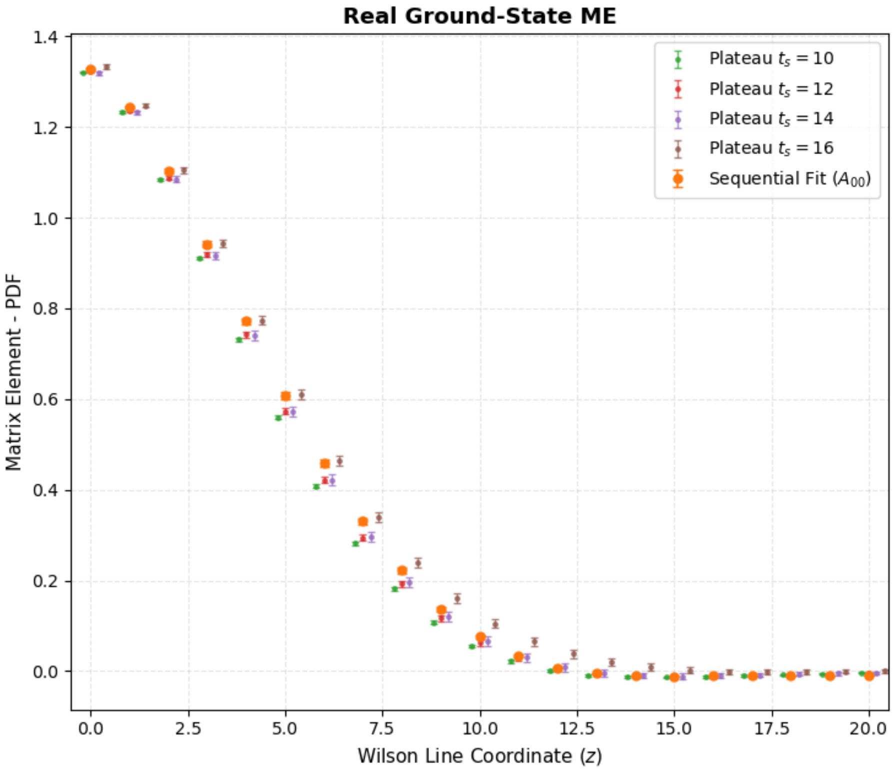
Results from *Two State Sequential Fits:*

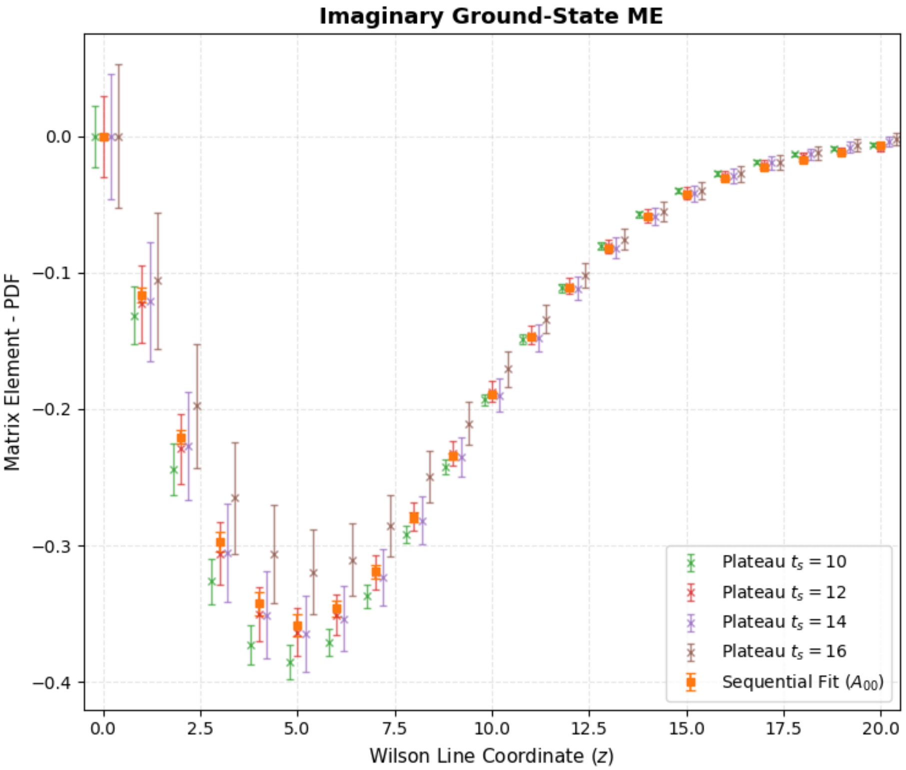


Results from *Two State Sequential Fits:*

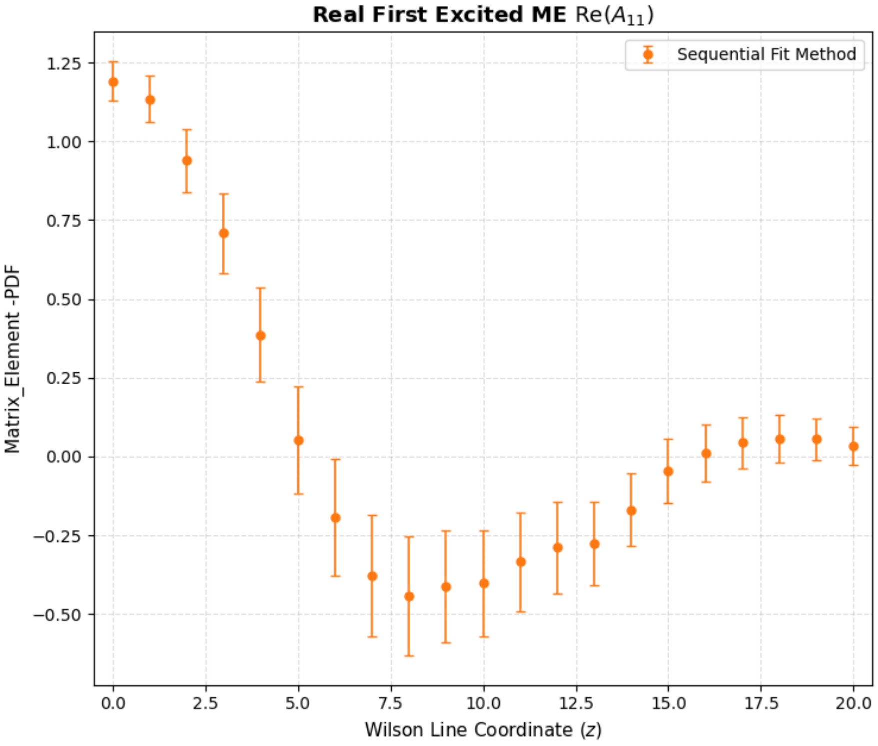


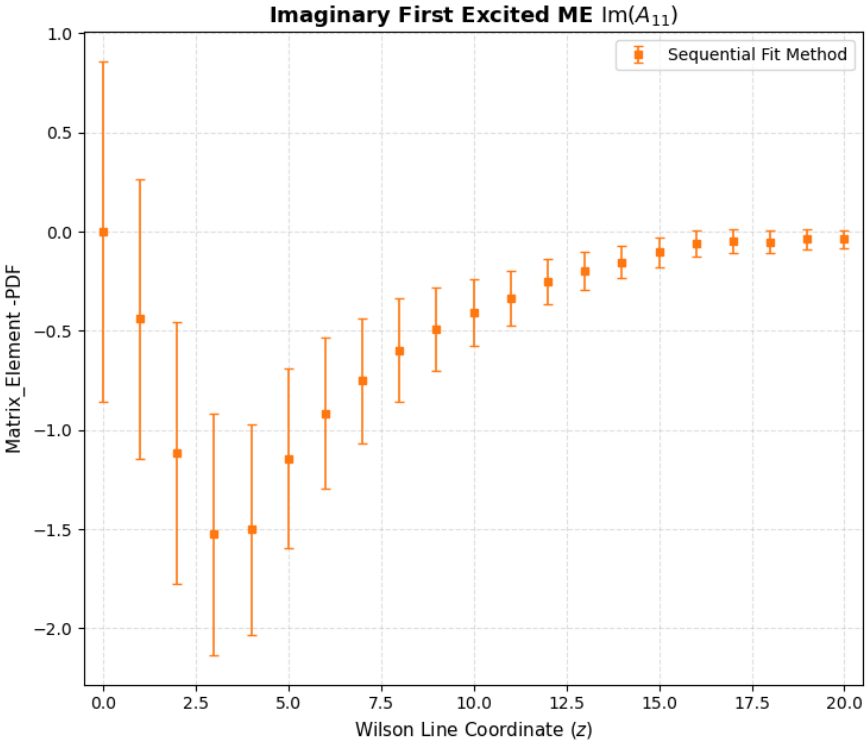
Ground State Matrix Elements from Sequential and Plateau fit:


Real Ground-State ME


Imaginary Ground-State ME

First Excited State Matrix Elements from Sequential fit:



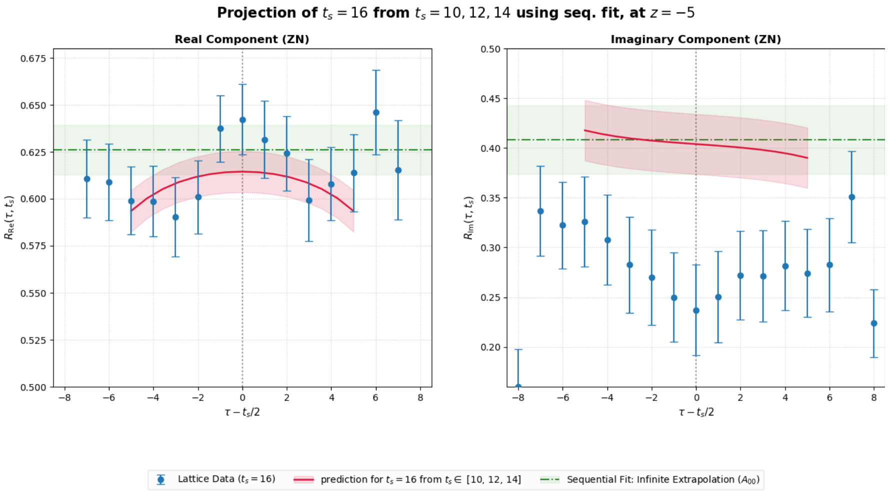


Why do we need Two-State Fits?

The signal-to-noise: $$\frac{R(t)}{\sigma(t)} \simeq \sqrt{N_{cfg}}exp\left(-\left[E_n(P)-\frac{3}{2}m_\pi\right]t\right)$$

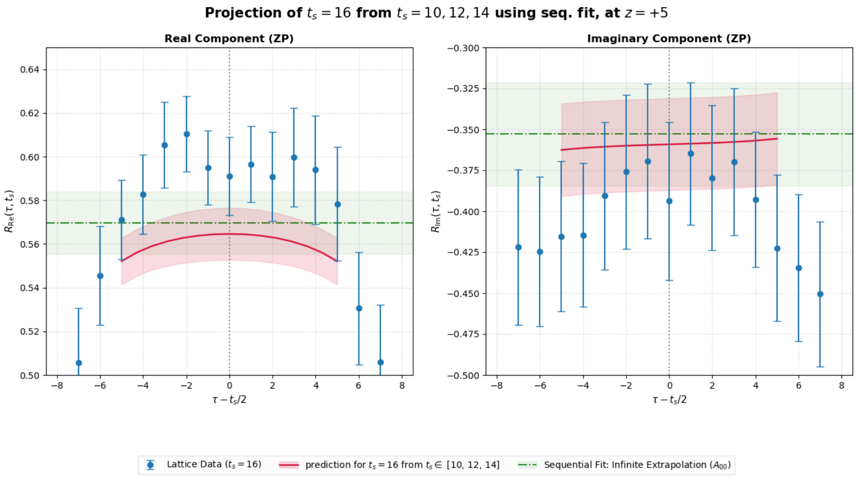
Ref: T. DeGrand and C. E. Detar, Lattice methods for quantum chromodynamics (World Scientific, 2006).

1. restricts from extracting the Matrix Elements at higher momentum boots (say ~ 2GeV), for sufficiently high value of t_s.
2. Let us see if lower values of $t_s \in$ [10a, 12a, 14a] can reproduce the results for $t_s = 16a$


Projection of $t_s = 16$ from $t_s = 10, 12, 14$ using seq. fit, at $z = -5$

Real Component (ZN)

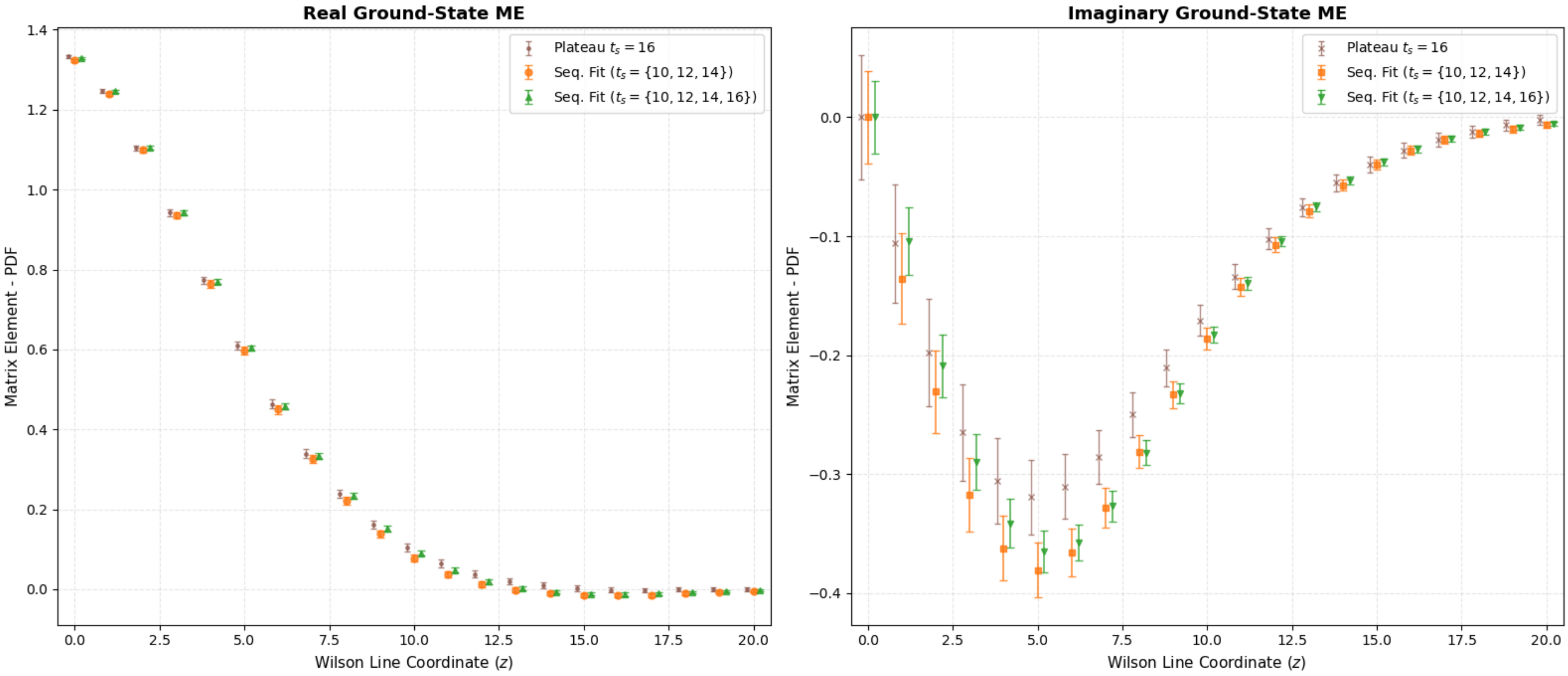
Imaginary Component (ZN)


Projection of $t_s = 16$ from $t_s = 10, 12, 14$ using seq. fit, at $z = +5$

Real Component (ZP)

Imaginary Component (ZP)

The comparison that t_s = 10a, 12a, 14a can reproduce 16a.



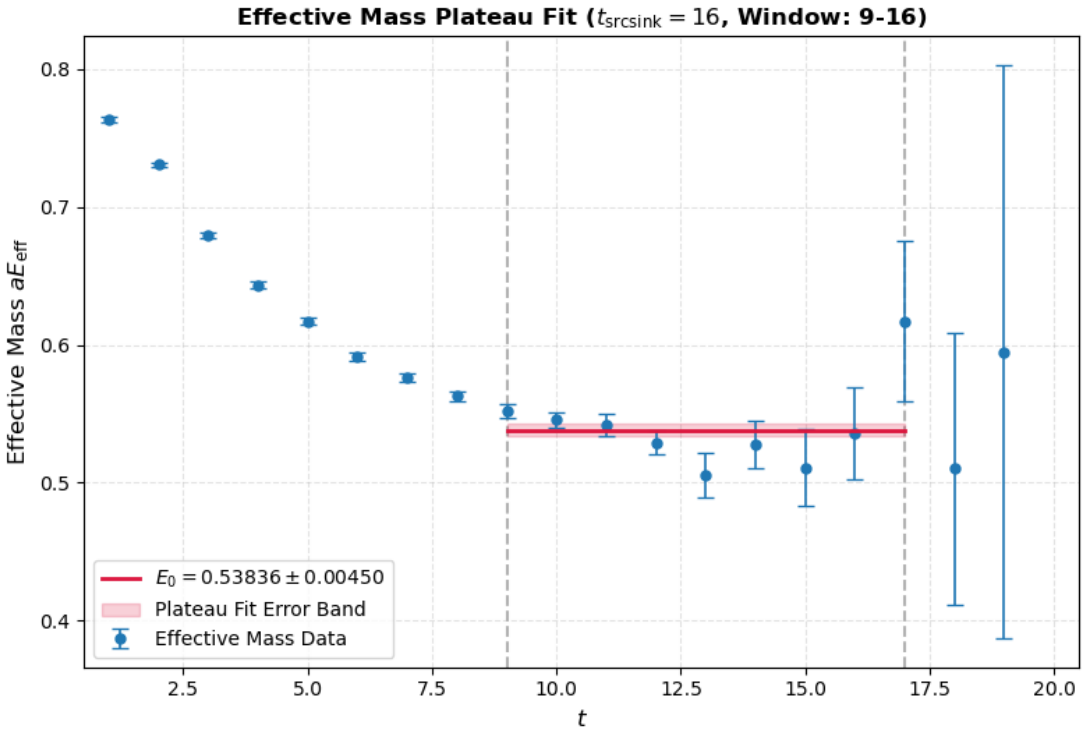
Conclusions and What's Next?

- ANN captures full t-dependence of Mellin moments model-independently. It serves as the baseline for the continuum extrapolations in future.
- Quantifying ANN-related systematic errors is important, specifically on overfitting and underfitting via regularization techniques.
- We'll perform the excited-state analysis using the Direct Fit method and extend the entire framework for GPDs.

Questions, please?

Back-up slide [1] —

In sequential fit procedure, fitting C2pt also evaluates E_0 - which needs ti be consistent with the value predicted from E_eff.



Back-up slide [2] —

A few words about the direct fits:

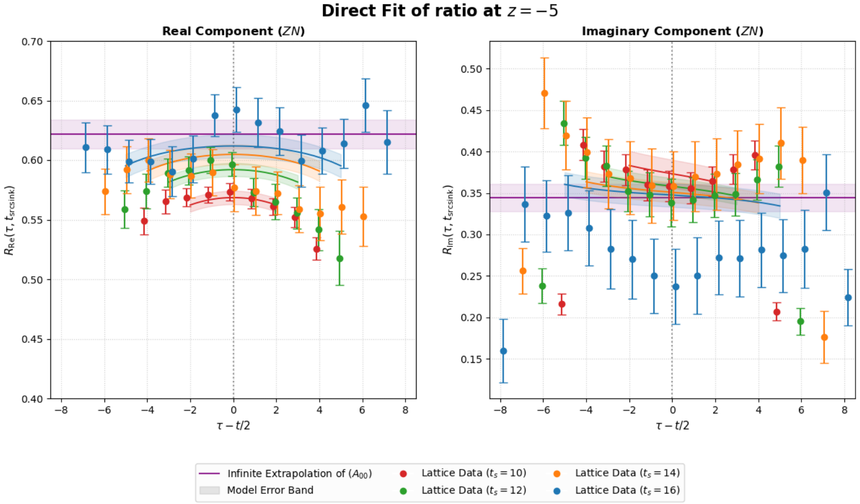
A. In this procedure, the Re(Ratio) ansatz is fitted against the data and the parameters: $Re(A_{00})$, Re(B), $|f_{10}|^2$, ΔE and Re(A_{11}) are determined. Similarly for the imaginary part.

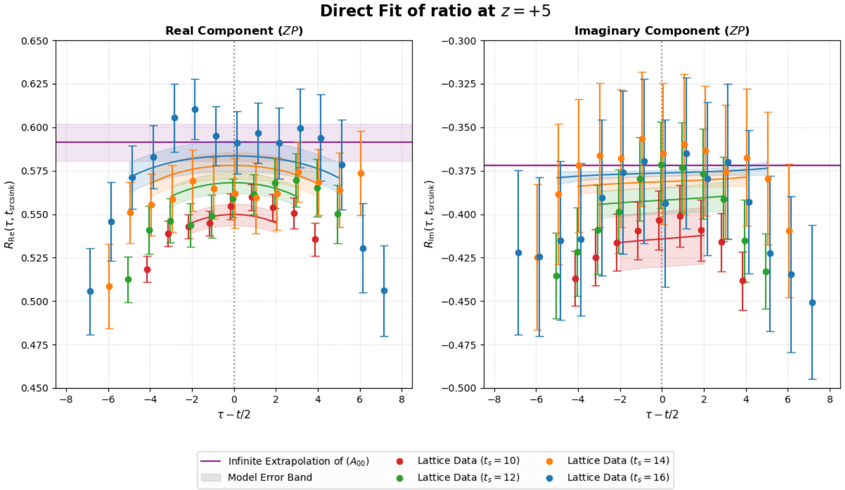
B. Correlated fitting is performed such that, we can appropriately *capture the fitting window*, For example the reduced χ^2 turns out ~0.9 and ~1.2 for training sample 0 data, for window $\tau \in [3, t_s - 3]$, otherwise χ^2 doesn't get normalized.

C. Procedure is stable, yet dependent on guess values. χ^2 is minimized in a controlled/constrained way such that that some parameters converges to an appropriate physical values.

Back-up slide [3] —

Some results from direct fits:


Direct Fit of ratio at $z = -5$


Direct Fit of ratio at $z = +5$