Probing proton structure via physics-guided neural networks in holographic QCD*

Wei Kou (寇维)[1,2,3,4†] Xurong Chen (陈旭荣)[1,2,3,4‡]

[1]Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China
[2]Southern Center for Nuclear Science Theory (SCNT), Institute of Modern Physics, Chinese Academy of Sciences, Huizhou 516000, China
[3]School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China
[4]State Key Laboratory of Heavy Ion Science and Technology, Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China

Abstract: Describing the proton structure function $F_2(x,Q^2)$ across the nonperturbative and transition regimes of quantum chromodynamics remains a significant theoretical challenge. We construct a physics-guided phenomenological representation that combines fermion-resonance and Pomeron-like AdS/QCD basis functions with a neural parametrization. The fermion-resonance sector is constrained by the AdS eigenvalue problem and the physical proton mass, while the neural parametrization describes the effective kinematic mixing, the Pomeron normalization, and a regularized log-space correction. For 187 selected SLAC measurements, a fixed training–test partition containing 153 and 34 points, respectively, yields the point-normalized diagnostics $\chi^2/N_{\rm train}=0.542$, $\chi^2/N_{\rm test}=1.990$, and $\chi^2/N_{\rm all}=0.805$. Conditional on the chosen holographic basis, the fitted channel fractions show a continuous transition from Pomeron-like dominance at the lower end of the measured x range to fermion-resonance dominance at larger x. The midpoint of the mixing function occurs at $x\simeq0.187$, whereas equality of the final channel fractions occurs at $x\simeq0.226$. These two definitions identify a transition region of approximately $x\simeq0.19$–0.23, rather than a unique physical boundary. Although $\alpha_0\simeq1.08$ is compatible with the data, fixed-α_0 refits show that it is not separately identifiable from the flexible Pomeron normalization. The x-dependent holographic coordinates and associated effective excited scales should therefore be interpreted as basis-dependent effective quantities rather than direct measurements of an x-dependent proton spectrum. A comparison with a data-only neural baseline does not establish a predictive advantage for the holographic basis on the present dataset. The resulting framework thus provides a basis-conditioned phenomenological decomposition of the data, rather than a model-independent extraction of the underlying proton dynamics.

Keywords: structure function, deep inelastic scattering, AdS/CFT correspondence, deep learning

DOI: 10.1088/1674-1137/aea284 **CSTR:** 32044.14.ChinesePhysicsC.51013104

I. INTRODUCTION

Deep inelastic scattering (DIS) is a principal probe of the internal structure of the nucleon [1–4]. Fixed-target measurements performed at the Stanford Linear Accelerator Center (SLAC) provide important information on the proton structure function $F_2(x,Q^2)$ across the valence-dominated region and the transition between resonance-like and deep-inelastic behavior [3, 4]. In the low-to-moderate-Q^2 and moderate-to-large-x regime, however, a leading-twist perturbative description alone is generally insufficient. Target-mass effects, higher-twist contributions, confinement dynamics, and remnants of hadronic resonance structure may all become relevant [5–10]. A uniformly controlled first-principles calculation of F_2 throughout this kinematic region therefore remains challenging and motivates the use of nonperturbative effective descriptions.

Received 26 June 2026; Accepted 3 September 2026; Accepted manuscript online 4 September 2026

* Supported by the National Natural Science Foundation of China (12547118), the Research Program of State Key Laboratory of Heavy Ion Science and Technology, Institute of Modern Physics, Chinese Academy of Sciences (HIST2025CS08), and the National Key R&D Program of China (2024YFE0109800, 2024YFE0109802)

† E-mail: kouwei@impcas.ac.cn
‡ E-mail: xchen@impcas.ac.cn

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Article funded by SCOAP[3] and published under licence by Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd

Gauge/gravity duality [11–13] provides one such route to modeling strongly coupled hadronic dynamics. In particular, soft-wall AdS/QCD models reproduce approximately linear hadronic mass trajectories and supply normalizable bulk wave functions for hadronic states [14–16]. Within holographic descriptions of DIS, two complementary constructions have been used. At larger x, transitions of a bulk fermion into excited states provide an s-channel description of resonance-like contributions [17–22]. Toward smaller x, t-channel holographic Pomeron exchange provides a description of the smooth high-energy background [23–25]. These constructions describe different aspects and kinematic limits of holographic DIS. In the present work, they are adopted as two prescribed basis components within a common phenomenological representation rather than derived as sectors of a new unified dynamical equation. Moreover, within the SLAC kinematic region considered here, the Pomeron-like term should be understood as an effective smooth-background contribution, not as a precision determination of small-x Regge dynamics.

Machine-learning parametrizations have become useful in high-energy and nuclear physics [26–32], including global analyses of parton distribution functions [33–38], lattice-field-theory sampling [39], and physics-informed studies of differential equations [40–44]. Their flexibility is useful for representing multidimensional observables [45, 46], but it also introduces an identifiability problem. Because $F_2(x,Q^2)$ is a relatively smooth function over the selected kinematic domain, agreement with the data alone does not demonstrate that a particular physical basis has been uniquely selected, nor does it establish a unique physical interpretation of all fitted quantities. A physics-guided analysis must therefore distinguish the structure imposed by the theoretical basis from the information constrained by the data and examine the resulting description using held-out measurements, simpler data-driven parametrizations, and stability tests.

Motivated by these considerations, we construct a physics-guided phenomenological representation of the selected SLAC $F_2(x,Q^2)$ data. The fermion-resonance component is constrained by the AdS eigenvalue problem and the physical proton-mass condition, whereas the Pomeron-like component provides an effective smooth background. A neural parametrization determines the kinematic mixing $w_\theta(x,Q^2)$, the effective Pomeron normalization $g_{\mathbb{P}}(x,Q^2)$, and a regularized log-space correction $\delta_\theta(x,Q^2)$. These quantities control the relative weights and limited deformation of the prescribed holographic components, rather than defining an additional dynamical equation relating the two channels.

A series of training–test, replica, parametrization, kinematic-cut, and data-driven baseline studies indicates a qualitative transition from a Pomeron-like-dominated fitted contribution at lower x to a fermion-resonance-dominated contribution at larger x. The location of this transition is definition dependent, with the diagnostics considered here placing it in the approximate region $x\sim0.19$–0.23. The fitted value $\alpha_0\simeq1.08$ is compatible with conventional soft-Pomeron phenomenology [47], although it is not separately identifiable from the flexible Pomeron normalization. Accordingly, the framework is interpreted as a transparent, basis-conditioned phenomenological decomposition of the measured structure function, rather than as a unique determination of an x-dependent proton spectrum or a demonstration of predictive superiority over a data-driven parametrization.

The remainder of this paper is organized as follows. Section II reviews the two holographic contributions and the soft-wall background. Section III defines their phenomenological mixing, the constrained holographic quantities, and the regularized correction. Section IV describes the selected SLAC data, the fixed training–test partition, and the numerical diagnostics. Section V presents the global description of F_2, the channel fractions, the effective holographic manifold and excited scales, and the fixed-α_0 and profile-χ^2 analyses. The conclusions and limitations of the resulting interpretation are summarized in Section VI.

II. HOLOGRAPHIC BASIS FOR A DUAL-CHANNEL DESCRIPTION

This section summarizes the two holographic contributions adopted as prescribed basis functions for the phenomenological description of DIS. The corresponding physical picture is shown schematically in Fig. 1. The first contribution is associated with s-channel transitions between normalizable bulk-fermion states, whereas the

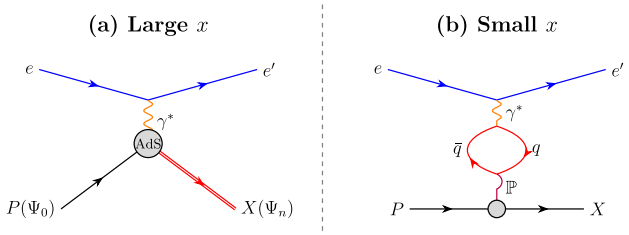


Fig. 1. (color online) Holographic dual-channel picture of deep inelastic scattering. (a) In the larger-x regime, the scattering is modeled by an s-channel fermion-resonance mechanism, in which the virtual photon couples to the ground-state bulk fermion field Ψ_0 through an AdS overlap integral and excites a tower of hadronic resonance states Ψ_n [18, 19, 48]. (b) Toward lower x, the smooth background is modeled by a t-channel holographic Pomeron exchange $\mathbb{P}$, corresponding to closed-string diffusion in the AdS bulk [49, 50]. In the present SLAC kinematic region, this contribution should be understood as an effective Pomeron-like background rather than as a precise small-x Regge determination.

second describes a smooth background through t-channel holographic Pomeron exchange [17–19, 50, 51]. These two contributions are not derived here as coupled sectors of a new dynamical equation; rather, they provide two existing holographic basis components whose relative contributions are determined phenomenologically in Sec. III. According to the AdS/CFT dictionary [11–13], the boundary electromagnetic current and hadronic states are represented by a bulk vector field and normalizable spinor modes in a five-dimensional asymptotically AdS_5 geometry [17, 52, 53].

A. Kinematics and the soft-wall background

Let P and q denote the proton and virtual-photon four-momenta, respectively, with $P^2 = M_p^2$ and $q^2 = -Q^2 < 0$. The Bjorken variable and the invariant mass of the inclusive hadronic final state are

$$x = \frac{Q^2}{2P\cdot q},$$
$$W^2 \equiv (P+q)^2 = M_p^2 + Q^2\frac{1-x}{x}. \tag{1}$$

The soft-wall construction introduces an infrared scale via the background profile

$$\Phi(z) = \kappa^2 z^2, \tag{2}$$

where z is the holographic coordinate and κ^2 is the dimensionful soft-wall or deformation scale [14–16, 54]. In the deformed AdS_5 realization used for the fermion sector, this scale enters the effective radial potential and governs the separation of the normalizable hadronic modes [18, 19].

In the underlying holographic construction, κ^2 and the five-dimensional fermion mass M_5 characterize the chosen background and bulk field. The bin-indexed quantities κ_i^2 and $M_{5,i}$ introduced in the phenomenological analysis should therefore not be interpreted as independent measurements of proton parameters at different values of Bjorken x. For x in the ith measured bin, they specify an effective point on a constrained family of holographic basis functions, with the ground-state proton mass held fixed.

B. The *s*-channel: bulk fermion and resonances

At the larger-x end of the selected kinematic region, the first basis component is constructed from transitions of the ground-state bulk fermion to excited normalizable states. The left- and right-chiral radial wave functions satisfy Schrödinger-type equations [16, 18, 19, 54]

$$-\psi''_{L,R}(z) + V_{L,R}(z;\kappa^2, M_5)\psi_{L,R}(z) = M_n^2\psi_{L,R}(z), \tag{3}$$

where M_n denotes the mass scale associated with the nth normalizable mode. The effective potentials $V_{L,R}$ incorporate the dependence on the deformed background and the bulk-fermion mass.

For each effective coordinate κ_i^2, the corresponding value of $M_{5,i}$ is fixed by the common ground-state condition

$$M_0(\kappa_i^2, M_{5,i}) = M_p, \qquad M_p = 0.938\ \text{GeV}. \tag{4}$$

Thus, $M_{5,i}$ is not an independently varying fit parameter. The excited scales $M_{n\geq 1}$ obtained at a selected point on this manifold are conditional properties of that effective basis and should not be interpreted as physical resonance pole masses that vary with Bjorken x.

Retaining the leading Dirac-type transition current, the relevant left- and right-chiral overlap integrals can be written as [48–52, 55]

$$\mathcal{I}_n^{L,R}(Q^2) = \int \mathrm{d}z\,\mathcal{J}(Q,z)\psi_{L,R,0}(z)\psi_{L,R,n}(z), \tag{5}$$

where $\mathcal{J}(Q,z)$ denotes the bulk-to-boundary propagator of the virtual photon. The dependence of the wave functions on κ_i^2 and $M_{5,i}$ is left implicit.

For x in the ith bin, the fermion-resonance basis used in the present analysis is the finite-state expression [17–19]

$$F_2^{\mathrm{F}}(x,Q^2) = \frac{G_{\mathrm{eff},i}Q^2}{8xM_p^2}\sum_{n=1}^{11}\left[\left|\mathcal{I}_n^L(Q^2)\right|^2 + \left|\mathcal{I}_n^R(Q^2)\right|^2\right], \tag{6}$$

where the fixed overall normalization factors have been absorbed into the effective fermion-sector normalization $G_{\mathrm{eff},i}$. The formal spectral derivation of the transition sum is reviewed in Appendix A.

Equation (6) retains the eleven excited states $n = 1,\cdots,11$. No Gaussian or Breit–Wigner resonance profile is introduced in the central construction. It should therefore be regarded as a phenomenological finite-state basis rather than a first-principles description of physical resonance widths.

The retained current contains the leading Dirac-type transition form factor, while the Pauli-type term and subleading tensor structures are omitted in the large-Q^2 reduction [52, 48, 55]. Because the selected data extend to $Q^2 = 0.583\ \text{GeV}^2$, this reduction is not uniformly controlled over the entire fitted domain. The lowest-Q^2 results consequently represent an application of the effective fermion-resonance basis rather than a complete low-Q^2 current calculation. The corresponding restricted-Q^2 stability diagnostics are discussed in Sec. V.

C. The *t*-channel: holographic pomeron exchange

At the lower end of the measured x range, a smooth energy-dependent background is described by holographic Pomeron exchange. In the corresponding bulk picture, the virtual photon generates a bulk electromagnetic profile near the ultraviolet boundary, which interacts with the proton profile via a t-channel closed-string or graviton exchange [23, 24, 49, 50, 56]. This contribution is used here as an effective Pomeron-like basis component; within the present SLAC kinematic regime, it is not interpreted as a precision determination of asymptotic small-x dynamics [57].

Using the optical theorem, the corresponding holographic integral may be written, up to an overall normalization convention, as [25, 50, 58]

$$\begin{aligned} I_2^{\mathrm{P}}(x,Q^2) &\propto \frac{Q^2}{4\pi^2\alpha_{\mathrm{em}}}\int \mathrm{d}z \int \mathrm{d}z'\, P_{13}(z,Q)P_{24}(z',M_p) \\ &\quad \times \mathrm{Im}\chi(\hat{s},z,z'), \\ \hat{s} &= \frac{Q^2}{x}, \end{aligned} \tag{7}$$

where P_{13} and P_{24} denote the virtual-photon and proton bulk profiles, respectively. The kernel $\chi(\hat{s},z,z')$ describes diffusion in the holographic direction and provides the smooth energy dependence of the exchanged closed-string contribution [25, 49, 50].

The Pomeron-like contribution entering the phenomenological combination is defined by

$$\begin{aligned} F_2^{\mathrm{P}}(x,Q^2) &= \mathcal{G}_{\mathrm{P}}(x,Q^2) I_2^{\mathrm{P}}(x,Q^2), \\ \mathcal{G}_{\mathrm{P}}(x,Q^2) &= g_P(x,Q^2)\left(\frac{Q^2}{xQ_0^2}\right)^{\alpha_0-1}, \qquad Q_0^2 = 3\ \mathrm{GeV}^2. \end{aligned} \tag{8}$$

The reference scale Q_0^2 renders the base of the noninteger power dimensionless. A change in this fixed reference scale can be absorbed into the effective normalization $g_P(x,Q^2)$ and therefore does not alter the predicted structure function.

Because $g_P(x,Q^2)$ is flexible, variations in α_0 can be compensated by corresponding changes in $g_P(x,Q^2)$. The quantity constrained by the present data is therefore the combined factor $\mathcal{G}_{\mathrm{P}}(x,Q^2)$, rather than $g_P(x,Q^2)$ and α_0 separately. Accordingly, α_0 is treated as an effective parameter whose compatibility with soft-Pomeron phenomenology can be assessed, not as an independently determined intercept [47, 57].

Equations (6) and (8) define the two prescribed holographic basis contributions used in the phenomenological combination below. Their labels refer to the holographic mechanisms from which they are constructed and should not be interpreted as a direct, model-independent separation of the measured structure function into valence-quark and sea-quark or gluon distributions.

III. PHYSICS-GUIDED PHENOMENOLOGICAL REPRESENTATION

This section defines the phenomenological representation used to combine the two holographic basis contributions introduced in Sec. II. The fermion-resonance sector provides $F_2^{\mathrm{F}}(x,Q^2)$, while the Pomeron-like sector provides $I_2^{\mathrm{P}}(x,Q^2)$ and the associated contribution $F_2^{\mathrm{P}}(x,Q^2)$ defined in Eq. (8). Their relative weighting and limited deformation are described by a neural parametrization.

The physical guidance enters through the prescribed holographic functional forms and the proton-mass constraint. The neural component does not introduce a new dynamical relation between the two channels, nor does the data fit solve the exact AdS eigenvalue problem at each optimization step. Instead, it provides a kinematic parametrization conditioned on the selected holographic basis. This use of prescribed physical structure together with data-trained components belongs to the broader class of theory-guided and hybrid physics–machine-learning models [59, 60]. The resulting construction is shown schematically in Fig. 2.

A. Kinematic parametrization and basis combination

The neural parametrization takes as input kinematic features constructed from x and Q^2 and returns three functions

$$\begin{aligned} &w_\theta(x,Q^2) \in (0,1), \\ &g_P(x,Q^2) > 0, \\ &\delta_\theta(x,Q^2) \in \mathbb{R}. \end{aligned} \tag{9}$$

Here, w_θ is a bounded gate that controls the relative weighting of the two basis components, g_{P} is the effective normalization entering the Pomeron-like contribution in Eq. (8), and δ_θ is a multiplicative correction expressed in logarithmic space. The exact input features, network architecture, and output transformations are specified in Sec. IV.

The two holographic contributions are first combined as follows

$$F_{2,\mathrm{mix}}(x,Q^2) = w_\theta(x,Q^2)F_2^{\mathrm{F}}(x,Q^2) + \left[1-w_\theta(x,Q^2)\right]F_2^{\mathrm{P}}(x,Q^2). \tag{10}$$

The final prediction is then defined by

$$F_2^{\mathrm{total}}(x,Q^2) = F_{2,\mathrm{mix}}(x,Q^2)\exp\left[\delta_\theta(x,Q^2)\right]. \tag{11}$$

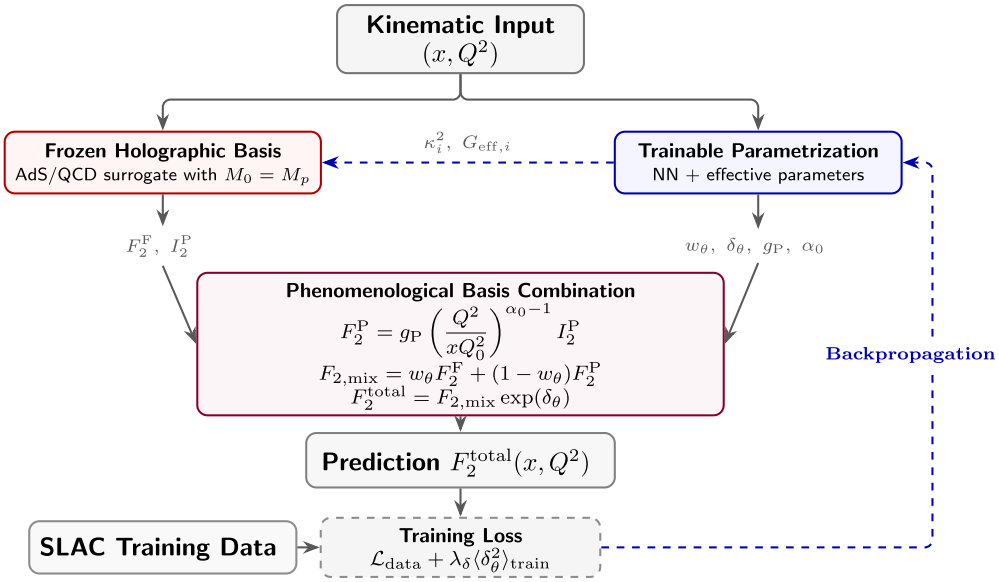


Fig. 2. (color online) Schematic representation of the basis-conditioned, physics-guided phenomenological fit. The exact holographic solver is used offline to generate a theory grid, which is represented during the SLAC fit by the frozen surrogate $\mathcal{S}_\varphi$. The data-trained block contains the neural-network outputs w_θ, g_{P}, and δ_θ, the bin-indexed effective quantities κ_i^2 and $G_{\mathrm{eff},i}$, and the global effective parameter α_0. For each κ_i^2, the associated bulk mass $M_{5,i}$ is fixed by the physical proton-mass condition $M_0(\kappa_i^2, M_{5,i}) = M_p$. The surrogate returns the fermion-resonance basis F_2^{F} and the Pomeron-like basis I_2^{P}. The latter is converted into F_2^{P} through the combined factor $\mathcal{G}_{\mathrm{P}}$ before the two prescribed basis contributions are combined phenomenologically. Only the training subset enters the optimization objective; the held-out points are used solely for diagnostics.

The exponential form makes δ_θ additive in $\log F_2$ and preserves the positivity of the prediction whenever the mixed basis contribution is positive.

Equations (10) and (11) constitute a phenomenological ansatz rather than a consequence of a common five-dimensional action. In particular, w_θ is not a microscopic transition probability, and the weighted combination is not a field-theoretic subtraction theorem for the possible overlap between the two holographic descriptions. Its role is to define a smooth, data-constrained interpolation within the selected basis.

B. Constrained holographic coordinates and regularization

For a data point belonging to the measured Bjorken-x_i bin, the fermion-resonance basis is evaluated using the effective coordinate κ_i^2 and the normalization $G_{\mathrm{eff},i}$. The associated bulk mass $M_{5,i}$ is not fitted independently; instead, it is fixed by the common ground-state condition in Eq. (4). The quantities κ_i^2 and $M_{5,i}$ therefore specify a point on the proton-mass-constrained holographic manifold. They are not interpreted as fundamental proton parameters that physically vary with Bjorken x.

The quantities fitted to the SLAC training data [3, 4] may be summarized as

$$\Theta_{\mathrm{fit}} = \left\{\theta_{\mathrm{NN}}, \left\{\kappa_i^2, G_{\mathrm{eff},i}\right\}_{i=1}^{N_x}, \alpha_0\right\}, \tag{12}$$

where θ_{NN} denotes the parameters governing w_θ, g_{P}, and δ_θ, and N_x is the number of measured Bjorken-x bins. The effective intercept α_0 is varied in the reference fit and held fixed in the profile-style refits described in Sec. V, following the general logic of profile-based identifiability diagnostics [61]. Neither $M_{5,i}$ nor the parameters of the frozen holographic surrogate belongs to Θ_{fit}.

The status of the quantities entering the representation is summarized in Table 1. The proton-mass condition removes $M_{5,i}$ as an independent fit quantity, but it does not, by itself, make the complete regression low-dimensional because the neural and bin-indexed parametrizations remain flexible.

The residual correction is controlled by a quadratic Tikhonov-type penalty [62] in the regularized objective.

$$\mathcal{L}(\Theta_{\mathrm{fit}}) = \mathcal{L}_{\mathrm{data}} + \lambda_\delta \langle \delta_\theta^2 \rangle_{\mathrm{train}}, \tag{13}$$

where the average in the second term is taken only over the training points. The exact weighted log-space definition of $\mathcal{L}_{\mathrm{data}}$, the central value of λ_δ, and the training–test construction are given in Sec. IV.

The penalty in Eq. (13) introduces an explicit tradeoff between adherence to the prescribed holographic basis and additional phenomenological flexibility. Consequently, a small fitted δ_θ cannot by itself be taken as evidence that the holographic basis is dynamically complete, because its magnitude is also influenced by λ_δ. Its

Table 1. Status and interpretation of the quantities entering the holographic-basis fit to the SLAC data.

Quantity	Status in the data fit	Role and interpretation
F_2^{F} and I_2^{P}	Prescribed basis functions	Holographic fermion-resonance and Pomeron-like basis components
θ_{NN}	Fitted to the training data	Controls $w_\theta(x,Q^2)$, $g_P(x,Q^2)$, and $\delta_\theta(x,Q^2)$
κ_i^2	Fitted effective coordinate	Selects a point in the constrained holographic basis manifold for the ith measured x bin
$G_{\mathrm{eff},i}$	Fitted effective normalization	Controls the normalization of the fermion-resonance basis in the ith measured x bin
α_0	Fitted in the reference analysis	Effective exponent entering the Pomeron-like contribution; not separately identifiable from the flexible normalization g_P
$M_{5,i}$	Constrained	Determined from κ_i^2 by the physical proton-mass condition
M_p and Q_0^2	Fixed	Physical proton mass and reference scale used in the dimensionless Pomeron factor

interpretation must therefore be accompanied by the residual-penalty scan and the data-only comparison.

C. Basis fractions and identifiability

The gate w_θ does not generally equal the physical fraction of the mixed basis associated with the fermion-resonance component. The two basis functions can have different magnitudes even when their gate coefficients are equal. We therefore define the normalized mixed-basis fractions by

$$\begin{aligned} C_{\mathrm{F}}(x,Q^2) &= \frac{w_\theta(x,Q^2)F_2^{\mathrm{F}}(x,Q^2)}{F_{2,\mathrm{mix}}(x,Q^2)}, \\ C_{\mathrm{P}}(x,Q^2) &= \frac{\left[1-w_\theta(x,Q^2)\right]F_2^{\mathrm{P}}(x,Q^2)}{F_{2,\mathrm{mix}}(x,Q^2)}, \\ C_{\mathrm{F}}(x,Q^2)+C_{\mathrm{P}}(x,Q^2) &= 1. \end{aligned} \tag{14}$$

Because the same factor $\exp[\delta_\theta(x,Q^2)]$ multiplies both components, it cancels in the normalized fractions. Thus, C_{F} and C_{P} characterize the composition of the prescribed holographic mixture, whereas δ_θ characterizes the subsequent common deformation of that mixture.

The relative magnitude of the residual deformation is quantified by

$$\begin{aligned} R_\delta(x,Q^2) &= \left|\frac{F_2^{\mathrm{total}}(x,Q^2)-F_{2,\mathrm{mix}}(x,Q^2)}{F_2^{\mathrm{total}}(x,Q^2)}\right| \\ &= \left|1-\exp\left[-\delta_\theta(x,Q^2)\right]\right|. \end{aligned} \tag{15}$$

Two distinct crossover criteria can therefore be considered. The gate criterion is defined by $w_\theta = 1/2$, whereas equality of the final mixed-basis contributions is defined by $C_{\mathrm{F}} = C_{\mathrm{P}}$. These conditions need not occur at the same value of x, because the latter also depends on the relative magnitudes of F_2^{F} and F_2^{P}. Neither criterion should be interpreted as defining a universal physical boundary.

A second identifiability limitation follows from Eq. (8). The data constrain the combined factor $\mathcal{G}_{\mathrm{P}}(x,Q^2)$ rather than α_0 and $g_{\mathrm{P}}(x,Q^2)$ separately. More generally, w_θ, g_{P}, and δ_θ all affect the same observable and can therefore partially compensate for one another. Such compensating directions are characteristic of parameter non-identifiability in flexible multiparameter models [61, 63]. The decomposition obtained from Eq. (14) is therefore conditional on the selected holographic basis, neural parametrization, and residual penalty. It should not be interpreted as a model-independent separation of the measured structure function into underlying partonic components.

IV. NUMERICAL SETUP, DATA PARTITION, AND DIAGNOSTICS

This section specifies the experimental sample, the fixed training–held-out partition, the frozen numerical representation of the holographic basis, and the optimization and reporting procedures. The implementation separates the offline evaluation of the holographic basis from the quantities calibrated using the SLAC data. The held-out sample and all robustness variations are used solely for diagnostic purposes and do not determine the reference fit.

A. SLAC Sample and fixed held-out partition

The experimental input consists of unpolarized proton structure-function data, $F_2^p(x,Q^2)$, obtained from the SLAC fixed-target analysis of Whitlow *et al.* [3, 4]. For each data point, the invariant hadronic mass is evaluated using Eq. (1). The nominal selection requires

$$x > 10^{-3}, \qquad W^2 > 1.15\ \mathrm{GeV}^2. \tag{16}$$

No additional Q^2 cut is imposed on the input table in the reference analysis. After selection, the retained sample contains ten measured Bjorken-x bins and $N_{\mathrm{data}} = 187$ points, with the realized kinematic coverage

$$0.080 \le x \le 0.850,$$
$$0.583\ \mathrm{GeV}^2 \le Q^2 \le 29.174\ \mathrm{GeV}^2,$$
$$2.665\ \mathrm{GeV}^2 \le W^2 \le 28.698\ \mathrm{GeV}^2. \tag{17}$$

Thus, although the nominal lower cut is $W^2 > 1.15\ \mathrm{GeV}^2$, the smallest value present in the fitted sample is

$$W^2_{\min} = 2.664844\ \mathrm{GeV}^2. \tag{18}$$

The quantity denoted by σ_i below is the total experimental uncertainty provided for the ith point in the input table. The present analysis uses these uncertainties as diagonal pointwise weights. A correlated experimental covariance matrix is not reconstructed; consequently, the reported goodness-of-fit quantities should not be interpreted as arising from a complete correlated likelihood.

A fixed partition is constructed independently within each measured Bjorken-x bin. For a bin containing N_j points, a pseudo-random permutation with seed 42 is generated and

$$N_{\mathrm{test},j} = \lfloor 0.20 N_j \rfloor . \tag{19}$$

Points are assigned to the held-out sample. The remaining points are assigned to training. This procedure yields

$$N_{\mathrm{train}} = 153, \qquad N_{\mathrm{test}} = 34. \tag{20}$$

The held-out points do not contribute to the optimized loss, learning-rate scheduler, stopping criterion, hyperparameter tuning, or selection of the reference fit. Because every measured x bin is represented in both subsets, the held-out statistic evaluates interpolation in Q^2 within already observed x bins. It does not evaluate extrapolation to an unseen Bjorken-x region or to a different experiment.

B. Offline holographic calculation and frozen surrogate

The holographic basis is first evaluated independently of the experimental fit. For each selected value of κ^2, the exact eigensolver determines M_5 through the proton-mass constraint in Eq. (4), after which the fermion-resonance and Pomeron-like basis quantities are computed on the theory grid.

The grid contains ten linearly spaced values of x, ten logarithmically spaced values of Q^2, eight values of κ^2, and five values of G_{eff}. Its domain is

$$0.01 \le x \le 0.90, \qquad 1\ \mathrm{GeV}^2 \le Q^2 \le 200\ \mathrm{GeV}^2,$$
$$0.10\ \mathrm{GeV}^2 \le \kappa^2 \le 0.80\ \mathrm{GeV}^2, \qquad 1 \le G_{\mathrm{eff}} \le 8. \tag{21}$$

The complete Cartesian grid therefore contains 4000 theory points.

The repeated solution of the exact eigenvalue problem during the experimental fit is replaced by a fixed numerical representation [64]

$$\mathcal{S}_\varphi : \left(x, \ln \frac{Q^2}{1\ \mathrm{GeV}^2}, \kappa_i^2, G_{\mathrm{eff},i} \right) \longmapsto \left(\ln F_2^{\mathrm{F}}, \ln I_2^{\mathrm{P}} \right). \tag{22}$$

The surrogate $\mathcal{S}_\varphi$ is a fully connected 4−64−128−2 network with tanh activations. It contains 8898 parameters and is trained for 500 epochs on the theory-generated grid using the Adam optimizer [65], with a learning rate of 5×10^{-3}, a batch size of 256, and a mean-squared error loss on the two logarithmic outputs. Its parameters are then frozen before the SLAC data are introduced.

The 8898 surrogate parameters are therefore not data-trained parameters of the phenomenological fit. During the subsequent fit, gradients propagate through the fixed map with respect to its effective inputs κ_i^2 and $G_{\mathrm{eff},i}$, but not through the exact AdS eigensolver. The resulting workflow is

$$\text{exact holographic solver} \longrightarrow \text{theory grid}$$
$$\longrightarrow \text{frozen surrogate} \longrightarrow \text{fit to SLAC data.} \tag{23}$$

The domain in Eq. (21) is the surrogate-training domain, not a calibrated emulator uncertainty band [66]. The archived reference implementation did not reserve an independent exact-solver test set from these 4000 theory points. Consequently, no formal surrogate uncertainty is propagated into the final observables, and the surrogate is not claimed to have uniform precision outside its training domain.

In particular, the 14 experimental points satisfying $0.583 \le Q^2 < 1\ \mathrm{GeV}^2$ lie below the theory-grid range and therefore require extrapolation in Q^2. The restricted-Q^2 diagnostics described below are used to test whether the main qualitative conclusions depend on these points, but they do not provide a formal emulator-error estimate.

C. Neural parametrization and output transformations

The data-trained neural parameterization uses three kinematic inputs

$$\mathbf{u}(x, Q^2) = \left(x, \ln \frac{1}{x}, \ln \frac{Q^2}{Q_0^2} \right), \qquad Q_0^2 = 3\ \mathrm{GeV}^2. \tag{24}$$

The reference network has architecture 3−16−8−3 with tanh activations in the hidden layers. Let h_δ, h_w, and h_g denote its three raw outputs. The transformations used in the reference calculation are

$$\begin{aligned}\delta_\theta(x,Q^2) &= 0.8\tanh h_\delta(x,Q^2),\\ w_\theta(x,Q^2) &= \frac{1}{1+\exp\left[-h_w(x,Q^2)-10(x-0.2)\right]},\\ \widetilde{g}_P(x,Q^2) &= \exp\left[\mathrm{clip}\left(h_g(x,Q^2),-2,2\right)\right].\end{aligned} \tag{25}$$

The fixed reference-scale conversion associated with Eq. (8) is absorbed through

$$g_P(x,Q^2) = \widetilde{g}_P(x,Q^2)\left(\frac{Q_0^2}{1\ \mathrm{GeV}^2}\right)^{\alpha_0-1}. \tag{26}$$

This rewriting leaves the Pomeron-like contribution numerically unchanged while rendering the noninteger power in Eq. (8) dimensionless.

Equation (25) makes explicit that the gate is not neutrally parameterized in x. The fixed term $10(x-0.2)$ favors an increasing fermion-resonance gate centered at $x=0.2$ before the data-dependent output h_w is added. Therefore, a solution of $w_\theta = 1/2$ is conditional on this architectural prior and cannot be interpreted as a transition scale identified independently by the data. The contribution-based fractions C_F and C_P defined in Eq. (14) remain distinct from this gate criterion.

For the ten measured x bins, the fit contains ten values of κ_i^2 and ten values of $G_{\mathrm{eff},i}$. Their reference initializations are

$$\kappa_i^2 = 0.45\ \mathrm{GeV}^2, \qquad G_{\mathrm{eff},i} = 4, \qquad \alpha_0 = 1.08. \tag{27}$$

During optimization, the numerical bounds

$$\begin{aligned}0.05\ \mathrm{GeV}^2 \le \kappa_i^2 \le 1.00\ \mathrm{GeV}^2,\\ 0.5 \le G_{\mathrm{eff},i} \le 15,\\ 1.01 \le \alpha_0 \le 1.35\end{aligned} \tag{28}$$

are imposed after each optimization step. These numerical bounds are broader than the theory-grid ranges in Eq. (21). A surrogate evaluation outside those ranges is therefore extrapolative and is not assigned a precision interpretation.

The global neural network contains

$$N_\mathrm{NN} = (3\times16+16)+(16\times8+8)+(8\times3+3) = 227, \tag{29}$$

parameters. Including the twenty bin-indexed quantities and the effective intercept gives

$$N_\mathrm{fit} = 227+20+1 = 248. \tag{30}$$

Quantities are adjusted using the 153 training points. The fit is therefore overparameterized in the conventional regression sense. The proton-mass constraint and residual regularization restrict its functional behavior, but they do not justify treating $N_\mathrm{data}-N_\mathrm{fit}$ as the conventional number of statistical degrees of freedom [67].

D. Optimization objective and reporting statistics

For each training point, the scaled logarithmic residual is

$$r_i^\mathrm{log} = \frac{\ln F_{2,i}^\mathrm{total} - \ln F_{2,i}^\mathrm{exp}}{s_i}, \qquad s_i = \max\left(\frac{\sigma_i}{F_{2,i}^\mathrm{exp}}, 0.01\right). \tag{31}$$

The one-percent floor prevents exceptionally small tabulated relative uncertainties from dominating the optimization. The Huber function used in the fit is [68]

$$\mathrm{Huber}_\Delta(r) = \begin{cases}\dfrac{1}{2}r^2, & |r|\le\Delta,\\ \Delta\left(|r|-\dfrac{\Delta}{2}\right), & |r|>\Delta.\end{cases} \tag{32}$$

The explicit realization of the objective introduced in Eq. (13) is

$$\mathcal{L} = \frac{1}{N_\mathrm{train}}\sum_{i\in\mathrm{train}}\mathrm{Huber}_{\Delta=2}\left(r_i^\mathrm{log}\right) + \lambda_\delta\langle\delta_\theta^2\rangle_\mathrm{train}, \qquad \lambda_\delta = 0.5. \tag{33}$$

The global-network parameters, κ_i^2 and $G_{\mathrm{eff},i}$, are optimized with Adam using a learning rate of 5×10^{-3} and a weight decay of 10^{-4}. The effective intercept α_0 is assigned a separate learning rate of 2×10^{-4} and no weight decay. A ReduceLROnPlateau scheduler operates on the training objective, reducing the learning rate by a factor of 0.5 after 150 epochs without improvement, with a minimum learning rate of 10^{-5}. The total gradient norm is clipped at 1 [69], and the reference optimization is run for 5000 epochs.

The Huber objective is not numerically identical to the diagnostic χ^2/N. After training, the latter is evaluated in linear F_2 space as

$$\frac{\chi^2_\mathcal{S}}{N_\mathcal{S}} = \frac{1}{N_\mathcal{S}}\sum_{i\in\mathcal{S}}\left(\frac{F_{2,i}^\mathrm{total}-F_{2,i}^\mathrm{exp}}{\sigma_i}\right)^2, \tag{34}$$

where $\mathcal{S}$ denotes the training, held-out, or complete sample. We also report [70]

$$\mathrm{MAPE}_{\mathcal{S}} = \frac{100}{N_{\mathcal{S}}} \sum_{i\in\mathcal{S}} \left| \frac{F_{2,i}^{\mathrm{total}} - F_{2,i}^{\mathrm{exp}}}{F_{2,i}^{\mathrm{exp}}} \right|. \tag{35}$$

The split-resolved diagnostics of the fixed-reference fit are summarized in Table 2. Because of the flexible parametrization and the absence of a full covariance matrix, χ^2/N is reported only as a pointwise descriptive statistic, rather than as χ^2/d.o.f. or as the basis for a formal goodness-of-fit probability.

Table 2. Split-resolved pointwise diagnostics for the fixed reference fit. The held-out sample is not used during optimization or model selection.

Sample	N	χ^2	χ^2/N	MAPE
Training	153	82.86	0.542	1.61%
Held-out	34	67.66	1.990	2.10%
Complete	187	150.52	0.805	1.70%

E. Robustness protocol

Several variations are evaluated independently of the reference fit. First, the free-α_0 calculation is repeated for ten parameter initializations. Architecture dependence is tested using 8−4 tanh, 16−8 tanh, 32−16 tanh, and 16−8 SiLU hidden layers [71], with five initializations for each architecture. The residual penalty is varied over $\lambda_\delta = 0$, 0.1, 0.5, 1, and 5. Twenty experimental-data replicas [33] are generated through independent Gaussian fluctuations with widths σ_i, and thirteen fixed-α_0 refits span $1.02 \le \alpha_0 \le 1.14$ [61].

A purely data-driven baseline uses the same three kinematic inputs, the same 16−8 hidden widths, the same fixed data partition, and matched initializations, but maps directly to $\ln F_2$ without the holographic basis. This comparison tests predictive interpolation only; it does not assess the physical interpretation provided by the holographic representation.

Kinematic sensitivity is examined through independent refits with $W^2 > 3\,\mathrm{GeV}^2$ and $W^2 > 4\,\mathrm{GeV}^2$. In addition, the frozen reference prediction is reevaluated after imposing $Q^2 \ge 1\,\mathrm{GeV}^2$ and $Q^2 \ge 2\,\mathrm{GeV}^2$. The latter are post-fit subset diagnostics rather than new parameter refits. The corresponding retained training and held-out counts are 143/30 and 118/21, respectively.

The spreads produced by different initializations, architectures, data replicas, penalties, and kinematic cuts quantify numerical and model sensitivity. They do not constitute a calibrated Bayesian posterior or a frequentist confidence interval and therefore are not labeled as a formal theoretical uncertainty band. Their numerical consequences are presented in Sec. V.

V. RESULTS AND DISCUSSION

This section presents the reference fit and the numerical variations defined in Sec. IV. The discussion distinguishes among three levels of inference: the description of the measured structure function, the conditional decomposition within the prescribed holographic basis, and the stability of effective fit quantities under changes in initialization, architecture, regularization, and kinematic selection. The held-out sample and all diagnostic refits remain separate from the optimization of the fixed reference result.

A. Reference fit and held-out performance

Figure 3 compares the reference prediction with the SLAC measurements in ten Bjorken-x bins [3, 4]. The upper panel displays $F_2(x, Q^2)$ with successive vertical offsets given by powers of two. Filled circles denote the training points, whereas open squares denote the held-out points. The lower panel shows the unshifted ratio of each measurement to the reference prediction.

The split-resolved statistics are presented in Table 2. For the reference fit,

$$\begin{aligned} \frac{\chi^2_{\mathrm{train}}}{N_{\mathrm{train}}} &= 0.542,\ \mathrm{MAPE}_{\mathrm{train}} = 1.61\%, \\ \frac{\chi^2_{\mathrm{test}}}{N_{\mathrm{test}}} &= 1.990,\ \mathrm{MAPE}_{\mathrm{test}} = 2.10\%, \\ \frac{\chi^2_{\mathrm{all}}}{N_{\mathrm{all}}} &= 0.805,\ \mathrm{MAPE}_{\mathrm{all}} = 1.70\%. \end{aligned} \tag{36}$$

The held-out χ^2/N is appreciably larger than its training counterpart, as expected for a flexible parametrization with more adjusted quantities than training points. Nevertheless, the held-out MAPE remains close to the training value, and Fig. 3 shows that the measured Q^2 dependence is reproduced across the observed x bins.

The difference between the training and held-out statistics is retained as a generalization diagnostic rather than being absorbed into a single full-sample number [33, 36]. Moreover, the spread from the diagnostic refits is not interpreted as a formal confidence interval. Initialization, data-replica, and model variations are instead reported explicitly below.

B. Conditional channel decomposition and residual sensitivity

The channel composition is evaluated using the mixed-basis fractions defined in Eq. (14). To avoid evaluating every x bin at a common Q^2 value outside its measured support, the reference scale for each bin is chosen as

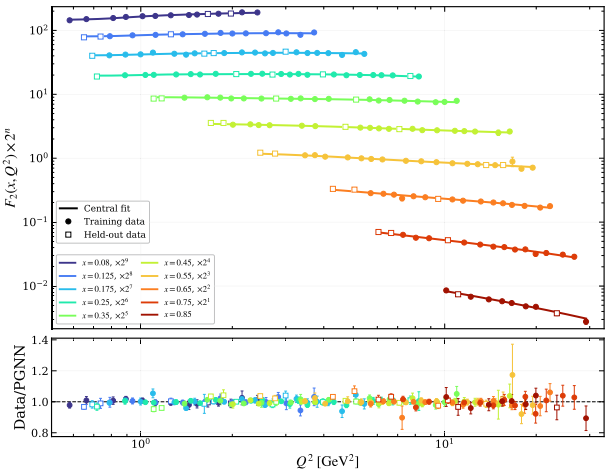


Fig. 3. (color online) Reference holographic-basis fit to the SLAC proton structure-function data [3, 4]. The upper panel shows $F_2(x,Q^2)$ for the ten measured Bjorken-x bins, with successive powers of two applied solely for visual separation. Filled circles and open squares denote the training and held-out points, respectively, while the solid curves denote the fixed reference prediction. The lower panel shows the corresponding data-to-prediction ratios. Experimental uncertainties are propagated to both panels. The held-out points are not used in the optimization.

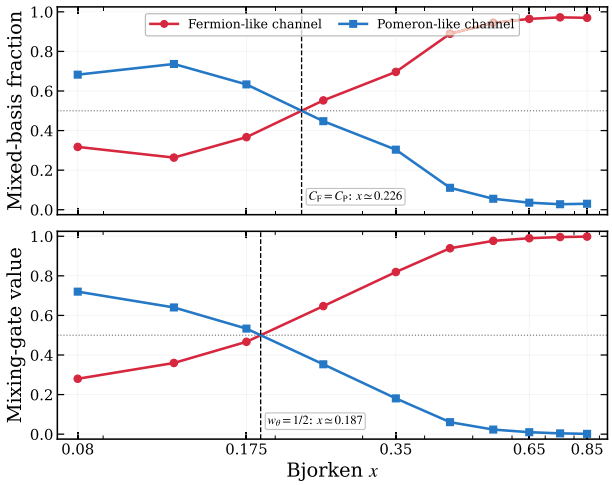


Fig. 4. (color online) Conditional decomposition of the prescribed holographic mixture along the median measured Q^2 in each Bjorken-x bin. The upper panel shows the mixed-basis fractions C_F and C_P defined in Eq. (14). Logarithmic interpolation between neighboring measured x bins yields $C_F = C_P$ near $x \simeq 0.226$. The lower panel shows the gate w_θ and its complement, $1 - w_\theta$; the corresponding condition $w_\theta = 1/2$ occurs near $x \simeq 0.187$. The gate is a diagnostic of the parametrization and is not itself a contribution fraction.

the median of the measured Q^2 values in that bin. The resulting path therefore follows the kinematic support of the SLAC sample rather than a fixed-Q^2 slice.

The upper panel of Fig. 4 shows C_F and C_P along this bin-median path. At the lowest measured point, $x = 0.08$ and $Q^2_{\rm med} = 1.3475\ {\rm GeV}^2$, the Pomeron-like fraction is

$$C_P = 0.683, \qquad C_F = 0.317. \tag{37}$$

At the largest measured value, $x = 0.85$ and $Q^2_{\rm med} = 16.032\ {\rm GeV}^2$, this ordering is reversed:

$$C_F = 0.970, \qquad C_P = 0.030. \tag{38}$$

The robust qualitative result is therefore a transition from a larger Pomeron-like contribution at the lower end of the measured x range [24, 25, 49, 50] to a fermion-resonance-dominated mixture at large x [17–19].

Logarithmic interpolation between neighboring measured x bins yields

$$\begin{aligned} w_\theta = \frac{1}{2}: &\quad x \simeq 0.187, \\ C_F = C_P: &\quad x \simeq 0.226. \end{aligned} \tag{39}$$

The separation between these values does not represent a statistical uncertainty associated with a single transition point. Rather, it arises from two different definitions: the first uses only the neural gate, whereas the second also incorporates the relative magnitudes of the two holographic basis functions.

In addition, the gate crossing is influenced by the fixed $10(x-0.2)$ term in Eq. (25). It therefore lies near the center encoded by the selected parametrization and cannot be regarded as a transition scale independently inferred from the data. The range $x \sim 0.19$–0.23 is quoted only as a descriptive, definition-dependent transition region within the measured sample. Neither endpoint defines a universal boundary between partonic regimes.

The residual deformation is quantified by Eq. (15). Its magnitude is controlled by λ_δ and therefore cannot be interpreted independently of the regularization [62]. Table 3 summarizes the five-initialization penalty scan.

Table 3. Median held-out diagnostics from the residual-penalty scan. Each row represents the median across five independent initializations and should not be confused with the single fixed reference fit in Table 2.

λ_δ	Median held-out $\chi^2/N_{\rm test}$	Median held-out $\lvert\delta_\theta\rvert$
0	1.741	0.183
0.1	2.018	0.124
0.5	2.143	0.071
1	2.038	0.050
5	1.981	0.020

Increasing λ_δ monotonically suppresses the logarithmic deformation, whereas the held-out statistic varies only moderately and non-monotonically. The reference value $\lambda_\delta = 0.5$ is therefore not selected because it minimizes the held-out error. Rather, it defines an interpretability–flexibility compromise, for which a typical held-out $|\delta_\theta|$ is approximately 0.071, corresponding to a multiplicative deformation of order 7%. The unregularized calculation attains a smaller median held-out statistic only by allowing a substantially larger residual.

A complementary comparison was performed with a purely data-driven network using the same inputs, hidden widths, data partition, and matched initializations. Across ten initializations [32, 59, 60], the median training and held-out statistics are

$$\begin{aligned} &\text{holographic guided:} && 0.516, && 2.174, \\ &\text{data only:} && 0.606, && 1.844, \end{aligned} \tag{40}$$

where the two columns denote the training and held-out χ^2/N, respectively. The data-only parametrization has a slightly smaller median held-out value, whereas the initialization distributions overlap. Therefore, the present dataset does not establish a predictive advantage for the holographic basis. Instead, the demonstrated role of that basis is to provide a conditional decomposition tied to the two prescribed holographic contributions and the proton-mass constraint.

C. Proton-mass manifold and effective excited scales

The proton-mass condition restricts each effective coordinate pair $(\kappa_i^2, M_{5,i})$ to the manifold $M_0(\kappa_i^2, M_{5,i}) = M_p$. Figure 5 shows this curve together with the ten points selected by the measured x bins. Variations in κ_i^2 therefore move the representation along a common mass-constrained manifold, rather than altering the physical ground-state proton mass [16, 18, 19].

The ground-state value remains fixed by construction, whereas the quantities M_1 and M_2 vary along the selected manifold. These higher eigenvalues provide effective mass scales for the fermion-resonance basis and influence its Q^2 dependence [48, 55]. However, because the underlying κ_i^2 values are bin-indexed fit coordinates, the resulting curves are conditional outputs of the parametrization. They do not constitute independent determinations of the soft-wall scale, the five-dimensional fermion mass, or the physical excited proton spectrum.

D. Optimization stability and effective pomeron intercept

The convergence and stability diagnostics are presented in Fig. 6. The upper-left panel shows the weighted training and held-out objectives during optimization, while the lower-left panel shows the trajectory of the global effective intercept. The held-out curve is evaluated solely for diagnostic purposes and does not feed back into the optimization.

The free reference fit gives

$$\alpha_0 = 1.0817. \tag{41}$$

For comparison, the Donnachie–Landshoff soft-Pomeron reference is $\alpha_{\mathbb{P}}(0) = 1.0808$ [47]. The proximity of the two values remains stable under the tested numerical variations. Ten initializations of the reference architecture yield

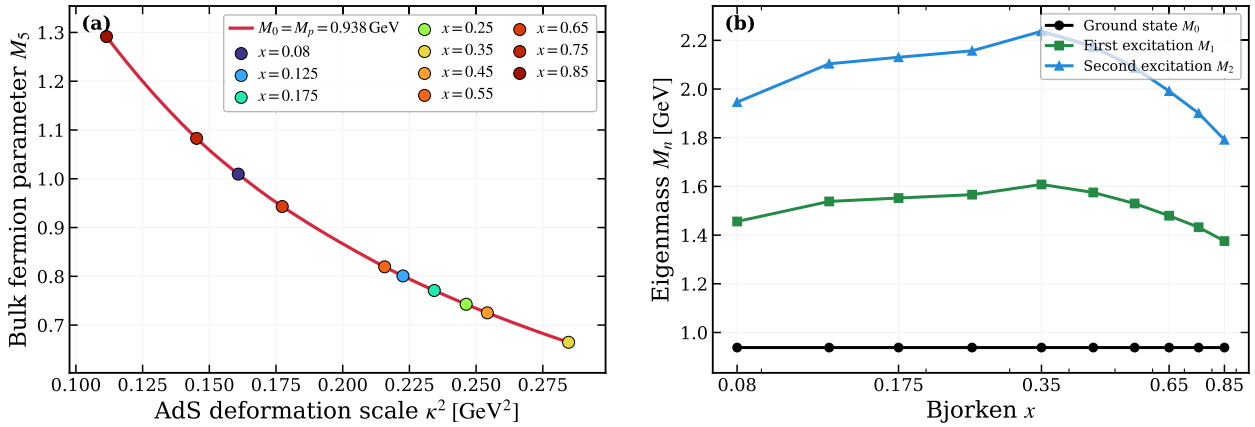


Fig. 5. (color online) Proton-mass-constrained holographic coordinates and their conditional outputs. Panel (a) shows the manifold on which $M_0 = M_p = 0.938$ GeV, together with the effective points $(\kappa_i^2, M_{5,i})$ selected for the ten measured Bjorken-x bins. Panel (b) shows the ground-state scale and the first two effective excited scales obtained at those points. The variation in M_1 and M_2 is a consequence of the bin-indexed coordinates used in the fit; these curves are not predictions of physical resonance pole masses or of a literal x-dependent proton spectrum.

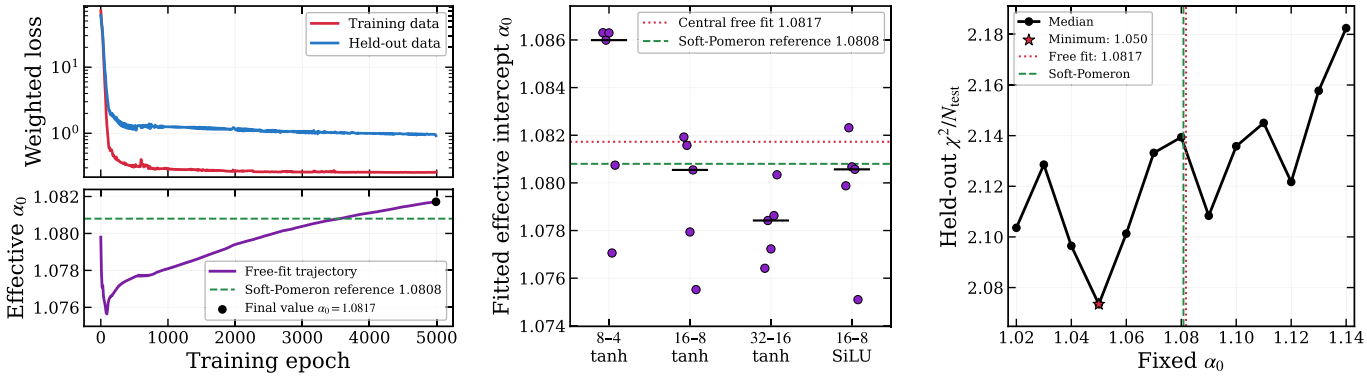


Fig. 6. (color online) Optimization and effective-intercept diagnostics. The upper-left panel shows the weighted training and held-out losses during the reference optimization, while the lower-left panel shows the free-fit trajectory of α_0, ending at $\alpha_0 = 1.0817$. The middle panel shows the fitted α_0 values obtained from four network architectures and five initializations per architecture; horizontal black segments denote the corresponding medians. The right panel shows the median held-out $\chi^2/N_{\rm test}$ from fixed-α_0 refits, with five initializations at each scanned value. The sampled minimum occurs at $\alpha_0 = 1.05$. The Donnachie–Landshoff reference, $\alpha_{\mathbb{P}}(0) = 1.0808$, is shown for comparison. The fixed-parameter scan is a robustness diagnostic, not a profile-likelihood confidence construction.

$$1.0755 \le \alpha_0 \le 1.0824, \tag{42}$$

while the complete scan over four architectures and twenty fits yields

$$1.0751 \le \alpha_0 \le 1.0863. \tag{43}$$

Twenty replicas of the experimental data [33] yield a median $\alpha_0 = 1.0804$, with individual values ranging from 1.0716 to 1.0846. These variations support only the limited conclusion that the free fits tend to remain near the soft-Pomeron reference value.

They do not, however, provide a precision determination of the intercept. When α_0 is fixed between 1.02 and 1.14 and all remaining quantities are reoptimized, the median held-out $\chi^2/N_{\rm test}$ changes only from 2.073 to 2.182. The lowest sampled median occurs at $\alpha_0 = 1.05$, rather than at the free-fit endpoint. This near-flatness is consistent with the identifiability limitation described in Sec. III C [61, 63]: the data constrain the combined Pomeron factor $\mathcal{G}_{\rm P}(x, Q^2)$, but do not separately determine α_0 and $g_P(x, Q^2)$. Consequently, $\alpha_0 \simeq 1.08$ is a soft-Pomeron-compatible effective value within this representation, not an independently extracted intercept [47, 50, 57].

E. Kinematic stability and scope

The effects of the higher-W^2 refits and the post-fit Q^2 restrictions are summarized in Table 4. The first block reports the medians over five independent refits for each W^2 selection. The second block reevaluates the fixed reference prediction after removing low-Q^2 points and therefore does not involve reoptimization.

The higher-W^2 refits leave the median effective intercept near 1.08, while the held-out statistic improves markedly as the low-W^2 points are removed. For the fixed reference prediction, imposing $Q^2 \ge 1\,\mathrm{GeV}^2$ leaves the held-out statistic essentially unchanged, whereas imposing $Q^2 \ge 2\,\mathrm{GeV}^2$ reduces it from 1.990 to 1.531. These diagnostics indicate that most of the residual tension is

Table IV. Kinematic-sensitivity diagnostics. Dashes indicate quantities not used in the corresponding comparison. The W^2 rows report medians over five refits, whereas the Q^2 rows correspond to subsets of the fixed reference prediction with $\alpha_0 = 1.0817$.

Procedure	Selection	$N_{\rm train}$	$N_{\rm test}$	$\chi^2_{\rm train}/N_{\rm train}$	$\chi^2_{\rm test}/N_{\rm test}$	α_0
		Independent W^2-cut refits				
Refit	$W^2 > 1.15\mathrm{GeV}^2$	153	34	–	2.143	1.0805
Refit	$W^2 > 3\mathrm{GeV}^2$	152	29	–	1.355	1.0807
Refit	$W^2 > 4\mathrm{GeV}^2$	135	22	–	0.367	1.0800
		Post-fit subsets of the fixed reference prediction				
Fixed prediction	Full sample	153	34	0.542	1.990	1.0817
Fixed prediction	$Q^2 \ge 1\mathrm{GeV}^2$	143	30	0.555	1.999	1.0817
Fixed prediction	$Q^2 \ge 2\mathrm{GeV}^2$	118	21	0.471	1.531	1.0817

associated with the low-W^2 and $Q^2 < 2\,\mathrm{GeV}^2$ portion of the sample. We nevertheless retain these points in the reference analysis to preserve coverage of the low-Q^2, low-W^2 nonperturbative and transition region.

These observations show that the reported qualitative channel ordering and effective-intercept compatibility are not driven exclusively by the lowest-W^2 or lowest-Q^2 points. They do not remove the limitations of the leading-tensor fermion basis at $Q^2 \simeq 0.58\,\mathrm{GeV}^2$ [7, 8, 52, 48, 55], nor do they provide a formal error estimate for the frozen surrogate below its theory-grid range [64, 66].

Taken together, the results support a conditional migration from a Pomeron-like-dominated mixture at the lower end of the measured x range to a fermion-resonance-dominated mixture at large x. The data do not determine a unique crossover, do not separately identify α_0 and g_P, and do not show that the holographic-basis fit predicts the held-out data more accurately than a matched data-only network. These qualifications define the scope within which the holographic decomposition and its effective parameters are interpreted.

VI. CONCLUSION AND OUTLOOK

We have developed a physics-guided neural representation of the proton structure function $F_2(x, Q^2)$ that combines two holographically motivated basis contributions with a trainable mixing gate, an effective Pomeron-like normalization, and a regularized logarithmic deformation. The fermionic coordinates are constrained to the proton-mass manifold, so that the ground-state mass remains fixed at $M_0 = M_p = 0.938\,\mathrm{GeV}$, while the effective coordinates vary across the measured Bjorken-x bins. This construction provides a structured phenomenological representation of the data rather than a first-principles extraction of unique holographic parameters.

For the fixed stratified partition of the 187 SLAC points, the reference fit gives $\chi^2_{\mathrm{train}}/N_{\mathrm{train}} = 0.542$, $\chi^2_{\mathrm{test}}/N_{\mathrm{test}} = 1.990$, and $\chi^2_{\mathrm{all}}/N_{\mathrm{all}} = 0.805$, with corresponding mean absolute percentage errors of 1.61%, 2.10%, and 1.70%. The held-out points do not participate in optimization or model selection. Their performance shows that the fitted representation interpolates the measured Q^2 dependence within the ten observed x bins. It does not constitute a test of extrapolation to an unseen Bjorken-x interval or to an independent experiment. Moreover, because the fit contains 248 adjusted quantities for 153 training points, the reported χ^2/N values are descriptive pointwise statistics and must not be interpreted as conventional χ^2/d.o.f. values.

Within the prescribed holographic decomposition, the fitted mixture exhibits a clear qualitative migration across the measured kinematic range. Along the median measured Q^2 in each x bin, the Pomeron-like fraction is $C_{\mathrm{P}} = 0.683$ at $x = 0.08$, whereas the fermion-resonance fraction reaches $C_{\mathrm{F}} = 0.970$ at $x = 0.85$. The contribution balance $C_{\mathrm{F}} = C_{\mathrm{P}}$ occurs near $x \simeq 0.226$, while the neural-gate condition $w_\theta = 1/2$ occurs near $x \simeq 0.187$. These two values represent different diagnostics and do not define the uncertainty interval of a single transition point. In particular, the gate midpoint is influenced by the fixed $10(x - 0.2)$ term in its parametrization. The data therefore support a conditional change in channel ordering, but not a unique or model-independent boundary between dynamical regimes.

The free reference fit gives an effective intercept $\alpha_0 = 1.0817$, close to the conventional soft-Pomeron reference value [47]. The fitted value remains near 1.08 under the tested changes in initialization, network architecture, experimental replicas, and kinematic selection. This numerical stability should not be confused with statistical identifiability. When α_0 is fixed over $1.02 \le \alpha_0 \le 1.14$ and the remaining quantities are reoptimized, the held-out statistic depends only weakly on the selected value and reaches its lowest sampled median at $\alpha_0 = 1.05$. Because the trainable function $g_P(x, Q^2)$ can compensate for changes in the intercept, the present data primarily constrain their combined Pomeron-like factor. Thus, $\alpha_0 \simeq 1.08$ is compatible with the fitted representation but is not an independently or precisely extracted Pomeron intercept.

The effective excited scales obtained along the proton-mass manifold must be interpreted with the same qualification. They control the kinematic behavior of the fermion-resonance basis, but their x dependence originates from the bin-indexed effective coordinates used in the fit. They are not predictions of physical resonance pole masses or evidence for a literal Bjorken-x-dependent proton spectrum. Likewise, the magnitude of the neural deformation is controlled by the penalty λ_δ. Its suppression under increasing regularization demonstrates that the residual can be controlled, but does not assign it a model-independent physical meaning.

The comparison with the matched data-only network further limits the scope of the conclusion. Across the tested initializations, the data-only model gives a median held-out $\chi^2/N_{\mathrm{test}} = 1.844$, compared with 2.174 for the holographically guided representation. The two initialization distributions overlap, but the present sample does not demonstrate a predictive advantage for the holographic basis. Its established role is instead to organize the fit into physically motivated, mass-constrained components whose conditional contributions can be examined separately.

Several limitations remain. The available experimental uncertainties are treated as pointwise total errors because a complete correlated covariance matrix is not included. The frozen surrogate was trained on a finite 4000-point theory grid, but no independent exact-solver validation sample was reserved, and no emulator uncertainty is propagated to the observables. Fourteen experimental

points lie below its nominal $Q^2 \geq 1\,\mathrm{GeV}^2$ theory-grid range, and the allowed effective parameter bounds also extend beyond the surrogate-training domain. In addition, the leading-tensor reduction used for the fermionic contribution is not uniformly controlled at the lowest measured values of Q^2. The kinematic-cut studies show that the main qualitative trends are not generated exclusively by these points, but they do not remove the underlying theoretical or numerical limitations. The spreads obtained from the performed variations should therefore be understood as sensitivity diagnostics rather than formal confidence or credible intervals.

The most immediate numerical improvement is to validate the surrogate on a new sample generated by the exact holographic solver and excluded from its training [64, 66]. The logarithmic root-mean-square error, median relative error, and 95th-percentile relative error should be reported separately for both surrogate outputs. Constraining the phenomenological fit to the validated interpolation domain, or propagating the resulting emulator error when that domain is exceeded, would then make the numerical accuracy of the holographic basis quantitatively assessable.

Further work should also test which conclusions genuinely require the holographic structure. Useful diagnostics include replacing the two holographic basis functions with trivial reference functions, removing one channel at a time, restricting the freedom of w_θ and g_P, eliminating the fixed gate-center term, and leaving out complete x bins rather than individual points. These tests would separate information carried by the holographic basis from that learned by the flexible neural functions. Replacing the independent bin-indexed coordinates with a smoother or hierarchical parametrization would additionally reduce the present freedom and permit more direct tests of parameter identifiability [61, 63].

Finally, incorporating correlated experimental systematics [72], extending the fermionic calculation beyond the leading large-Q^2 tensor structure [52, 48, 55], and testing the representation against independent DIS datasets or observables [73–76] would be necessary before assigning universal physical significance to the channel crossover or the effective holographic parameters. At the present stage, the central result is deliberately narrower: the holographic basis provides a consistent and interpretable coordinate system for describing the SLAC structure-function data, within which a stable qualitative change in channel ordering and a soft-Pomeron-compatible effective intercept are observed. Both conclusions remain conditional on the selected basis, parametrization, and kinematic domain.

ACKNOWLEDGMENTS

We are grateful to Dr. M. A. Martín Contreras for fruitful discussions on solving structure functions.

APPENDIX A: DERIVATION OF THE HADRONIC TENSOR AND STRUCTURE FUNCTIONS

In this appendix, we derive the leading-tensor spectral representation of the fermion-like contribution and specify the finite-state prescription used in the reference calculation. The derivation follows the holographic DIS formalisms of Refs. [17, 19]. The Pomeron-like t-channel contribution is defined separately in Sec. II.C and is therefore not repeated here.

The inclusive deep-inelastic-scattering cross section is governed by the hadronic tensor

$$W^{\mu\nu} = \frac{1}{4\pi}\int \mathrm{d}^4 y\, \mathrm{e}^{\mathrm{i}q\cdot y}\frac{1}{2}\sum_S \langle P,S|[J^\mu(y),J^\nu(0)]|P,S\rangle, \tag{A1}$$

where the factor $1/2$ accounts for the average over the initial proton spin. Lorentz covariance, parity invariance, and electromagnetic current conservation, $q_\mu W^{\mu\nu} = 0$, allow the unpolarized tensor to be decomposed as

$$W^{\mu\nu} = F_1(x,Q^2)P_1^{\mu\nu} + \frac{F_2(x,Q^2)}{P\cdot q}P_2^{\mu\nu}, \tag{A2}$$

where

$$P_1^{\mu\nu} = \eta^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}, \tag{A3}$$

and

$$P_2^{\mu\nu} = \left(P^\mu - \frac{P\cdot q}{q^2}q^\mu\right)\left(P^\nu - \frac{P\cdot q}{q^2}q^\nu\right). \tag{A4}$$

The kinematic variables satisfy

$$q^2 = -Q^2, \quad x = \frac{Q^2}{2P\cdot q}, \quad s = (P+q)^2 = M_p^2 + \frac{Q^2(1-x)}{x}. \tag{A5}$$

To evaluate the s-channel contribution, a complete set of holographic fermion states $|P_n,S_n\rangle$ is inserted between the two electromagnetic currents. With the standard relativistic normalization, the corresponding spectral representation is

$$W^{\mu\nu} = \frac{1}{2}\sum_{n,S,S_n}(2\pi)^4\delta^{(4)}(P+q-P_n)\langle P,S|J^\mu(0)|P_n,S_n\rangle\langle P_n,S_n|J^\nu(0)|P,S\rangle. \tag{A6}$$

Here, $M_n^2 = P_n^2$ denotes an eigenvalue of the bulk fermion problem for a fixed choice of holographic parameters.

The matrix element of the electromagnetic current is generated by the interaction between the bulk vector field $A_M(x,z)$, dual to the boundary current, and the initial and excited bulk spinor fields. The minimal interaction term can be written as

$$S_{\rm int} = g_V \int {\rm d}^5x\, \sqrt{-g}\, {\rm e}^{-\Phi(z)} \bar{\Psi}_n e_A^M \Gamma^A A_M \Psi_0, \tag{A7}$$

where g_V is an effective coupling, e_A^M is the vielbein, and Γ^A denotes the five-dimensional Dirac matrices.

A current-conserving decomposition of the four-dimensional transition vertex is

$$\begin{aligned} &\langle P_n, S_n|J^\mu(0)|P,S\rangle = \bar{u}(P_n,S_n)\Gamma_n^\mu(q)u(P,S), \\ &\Gamma_n^\mu(q) = \left(\gamma^\mu - \frac{q^\mu \not{q}}{q^2}\right) f_{1,n}(Q^2) + \frac{{\rm i}\sigma^{\mu\nu}q_\nu}{2M_p} f_{2,n}(Q^2). \end{aligned} \tag{A8}$$

The longitudinal completion in the first term guarantees $q_\mu \Gamma_n^\mu = 0$ and does not contribute when the vertex is contracted with the transverse projectors in Eqs. (A3) and (A4).

The Dirac-type transition form factor is determined by the overlap of the bulk-to-boundary propagator $\mathcal{J}(Q,z)$ with the normalizable fermion modes. Before transforming to Schrödinger-normalized wave functions, it is given by

$$\begin{aligned} f_{1,n}(Q^2) = &\int {\rm d}z\, \sqrt{-g}\, {\rm e}^{-\Phi(z)} z \mathcal{J}(Q,z) \\ &\times \left[\psi_{L,0}(z)\psi_{L,n}(z) + \psi_{R,0}(z)\psi_{R,n}(z)\right]. \end{aligned} \tag{A9}$$

The corresponding normalization factors can be absorbed into the effective fermion-channel coupling used in the phenomenological basis.

The reference calculation retains only the leading Dirac-type tensor structure and neglects the Pauli-type transition form factor $f_{2,n}(Q^2)$. This simplification is motivated by the large-Q^2 limit, in which the Pauli-type term is subleading. However, it is not uniformly controlled at the lowest momentum transfers in the fitted sample, which extends to $Q^2 = 0.583\ {\rm GeV}^2$. The post-fit diagnostics with $Q^2 \geq 1\ {\rm GeV}^2$ and $Q^2 \geq 2\ {\rm GeV}^2$, reported in Table 4, show that the principal qualitative results are not driven exclusively by the lowest-Q^2 points. However, such stability under these cuts does not provide a quantitative estimate of the omitted Pauli-type contribution.

After retaining the Dirac-type term, the spin summation can be written as a Dirac trace:

$$\begin{aligned} W^{\mu\nu} = &\frac{1}{2}\sum_n (2\pi)^4 \delta^{(4)}(P+q-P_n)\left|f_{1,n}(Q^2)\right|^2 \\ &\times {\rm Tr}\left[(\not{P}_n + M_n)\gamma^\mu(\not{P}+M_p)\gamma^\nu\right] \\ &+ \text{longitudinal terms}. \end{aligned} \tag{A10}$$

The terms proportional to q^μ or q^ν vanish under the transverse projection. The remaining trace is

$$\begin{aligned} {\rm Tr}\left[(\not{P}_n + M_n)\gamma^\mu(\not{P}+M_p)\gamma^\nu\right] = 4\Big[&P_n^\mu P^\nu + P_n^\nu P^\mu \\ &- \eta^{\mu\nu}(P_n\cdot P - M_n M_p)\Big]. \end{aligned} \tag{A11}$$

Using $P_n = P + q$ and matching the transverse tensor structures to Eq. (A2) yields the formal spectral form

$$\begin{aligned} F_{2,\rm spec}^{\rm F}(x,Q^2) &= CQ^2 \sum_n \delta(s-M_n^2)\mathcal{H}_n(Q^2), \\ \mathcal{H}_n(Q^2) &= \left|\mathcal{I}_{L,n}(Q^2)\right|^2 + \left|\mathcal{I}_{R,n}(Q^2)\right|^2, \end{aligned} \tag{A12}$$

where C contains the remaining normalization and phase-space factors. The chiral overlap integrals are defined, in the Schrödinger-normalized convention used in the numerical calculations, as

$$\begin{aligned} &\mathcal{I}_{\chi,n}\left(Q^2;\kappa_i^2,M_{5,i}\right) \\ &= \int {\rm d}z\, \mathcal{J}\left(Q,z;\kappa_i^2\right)\psi_{\chi,0}\left(z;\kappa_i^2,M_{5,i}\right)\psi_{\chi,n}\left(z;\kappa_i^2,M_{5,i}\right), \end{aligned} \tag{A13}$$

with $\chi = L,R$. The factors associated with the original curved-space measure are incorporated into the definition and normalization of these Schrödinger modes.

For completeness, in an asymptotic regime in which the excited states become densely spaced, a discrete spectral sum can formally be approximated by

$$\begin{aligned} \sum_n \delta(s-M_n^2)\mathcal{H}_n &\longrightarrow \int {\rm d}n\, \delta(s-M_n^2)\mathcal{H}_n \\ &= \frac{\partial n}{\partial M_n^2}\mathcal{H}_n\Big|_{M_n^2=s}. \end{aligned} \tag{A14}$$

For an asymptotically linear trajectory, $M_n^2 \simeq an + b$, the density of states approaches $\partial n/\partial M_n^2 \simeq 1/a$. This continuum relation explains the origin of the density-of-states expression frequently used in analytic holographic DIS calculations. It is not the prescription used in the numerical fit of the reference.

Instead, the fermion-like holographic basis employed in this work is defined directly through the finite overlap sum

$$F_2^{\mathrm{F}}\left(x,Q^2;\kappa_i^2,M_{5,i},G_{\mathrm{eff},i}\right)=\frac{\alpha_{\mathrm{F}}G_{\mathrm{eff},i}Q^2}{8xM_p^2}\sum_{n=1}^{11}\left[\left|\mathcal{I}_{L,n}\left(Q^2;\kappa_i^2,M_{5,i}\right)\right|^2+\left|\mathcal{I}_{R,n}\left(Q^2;\kappa_i^2,M_{5,i}\right)\right|^2\right], \tag{A15}$$

where $\alpha_{\mathrm{F}}=1$ is fixed in the reference implementation, and $G_{\mathrm{eff},i}$ is the effective fermion-channel normalization associated with the measured x_i bin. The ground state, $n=0$, represents the incoming proton and is not included in the excited-state sum.

Equation (A15) contains eleven discrete excited states, $n=1,\cdots,11$. Neither a Gaussian replacement nor a Breit–Wigner replacement of $\delta(s-M_n^2)$ is introduced in the numerical implementation, and no resonance-smearing width enters the reference calculation. The finite overlap sum should therefore be understood as a phenomenological discrete-state basis rather than as a first-principles calculation of physical resonance widths or line shapes.

For every fixed pair $(\kappa_i^2,M_{5,i})$, the quantities M_n are eigenvalues of the corresponding holographic problem. In the global fit, however, κ_i^2 is an effective parameter associated with the measured x_i bin, while $M_{5,i}$ is fixed by

$$M_0(\kappa_i^2,M_{5,i})=M_p. \tag{A16}$$

Consequently, the fitted values denoted by $M_n(x_i)$ are effective holographic scales evaluated at different points on the proton-mass manifold. They should not be interpreted as Bjorken-x-dependent physical resonance pole masses. By construction, only the ground-state mass is fixed to the physical proton mass.

The formal spectral representation in Eq. (A12) establishes the hadronic-tensor origin of the fermion tower, whereas Eq. (A15) provides the operational definition of the fermion-like basis used in the reference fit.

References

[1] E. D. Bloom, D. H. Coward, H. DeStaebler *et al.*, Phys. Rev. Lett. **23**, 930 (1969)
[2] M. Breidenbach, J. I. Friedman, H. W. Kendall *et al.*, Phys. Rev. Lett. **23**, 935 (1969)
[3] L. W. Whitlow, S. Rock, A. Bodek *et al.*, Phys. Lett. B **250**, 193 (1990)
[4] L. W. Whitlow, E. M. Riordan, S. Dasu *et al.*, Phys. Lett. B **282**, 475 (1992)
[5] S. J. Brodsky and G. P. Lepage, Adv. Ser. Direct. High Energy Phys. **5**, 93 (1989)
[6] A. v. Radyushkin, Fiz. Elem. Chast. Atom. Yadra **14**, 58 (1983)
[7] R. L. Jaffe and M. Soldate, Phys. Rev. D **26**, 49 (1982)
[8] R. K. Ellis, W. Furmanski, and R. Petronzio, Nucl. Phys. B **207**, 1 (1982)
[9] C. D. Roberts and A. G. Williams, Prog. Part. Nucl. Phys. **33**, 477 (1994), arXiv: hep-ph/9403224
[10] R. Alkofer and L. von Smekal, Phys. Rept. **353**, 281 (2001), arXiv: hep-ph/0007355
[11] J. M. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998), arXiv: hep-th/9711200
[12] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Phys. Lett. B **428**, 105 (1998), arXiv: hep-th/9802109
[13] E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998), arXiv: hep-th/9802150
[14] A. Karch, E. Katz, D. T. Son *et al.*, Phys. Rev. D **74**, 015005 (2006), arXiv: hep-ph/0602229
[15] O. Andreev and V. I. Zakharov, Phys. Rev. D **74**, 025023 (2006), arXiv: hep-ph/0604204
[16] T. Gutsche, V. E. Lyubovitskij, I. Schmidt *et al.*, Phys. Rev. D **85**, 076003 (2012), arXiv: 1108.0346[hep-ph]
[17] J. Polchinski and M. J. Strassler, JHEP **05**, 012 (2003), arXiv: hep-th/0209211
[18] N. R. F. Braga and A. Vega, Eur. Phys. J. C **72**, 2236 (2012), arXiv: 1110.2548[hep-ph]
[19] E. Folco Capossoli, M. A. Martín Contreras, D. Li *et al.*, Phys. Rev. D **102**, 086004 (2020), arXiv: 2007.09283[hep-ph]
[20] A. d. C. P. d. Nascimento, H. Boschi-Filho, and J. Noronha, *Proton Structure Functions from Holographic Einstein-Dilaton Models*, (2025), arXiv: 2505.13649[hep-th]
[21] J. Deng and D. Hou, Phys. Rev. D **112**, 036011 (2025), arXiv: 2502.00771[nucl-th]
[22] J. Deng and D. Hou, *Exploring Nucleon Structure and the Proton Mass Problem through Holographic QCD*, (2026), arXiv: 2603.04794[hep-ph]
[23] C. A. Ballon Bayona, H. Boschi-Filho, and N. R. F. Braga, JHEP **10**, 088 (2008), arXiv: 0712.3530[hep-th]
[24] E. Folco Capossoli and H. Boschi-Filho, Phys. Rev. D **92**, 126012 (2015), arXiv: 1509.01761[hep-th]
[25] A. Watanabe, T. Sawada, and M. Huang, Phys. Lett. B **805**, 135470 (2020), arXiv: 1910.10008[hep-ph]
[26] A. J. Larkoski, I. Moult, and B. Nachman, Phys. Rept. **841**, 1 (2020), arXiv: 1709.04464[hep-ph]
[27] D. Guest, K. Cranmer, and D. Whiteson, Ann. Rev. Nucl. Part. Sci. **68**, 161 (2018), arXiv: 1806.11484[hep-ex]
[28] A. Radovic, M. Williams, D. Rousseau *et al.*, Nature **560**, 41 (2018)
[29] K. Albertsson *et al.*, J. Phys. Conf. Ser. **1085**, 022008 (2018), arXiv: 1807.02876[physics.comp-ph]
[30] G. Carleo, I. Cirac, K. Cranmer *et al.*, Rev. Mod. Phys. **91**, 045002 (2019), arXiv: 1903.10563[physics.comp-ph]
[31] A. Boehnlein *et al.*, Rev. Mod. Phys. **94**, 031003 (2022), arXiv: 2112.02309[nucl-th]
[32] G. Aarts, K. Fukushima, T. Hatsuda *et al.*, Nature Rev. Phys. **7**, 154 (2025), arXiv: 2501.05580[hep-lat]
[33] L. Del Debbio *et al.* (NNPDF Collaboration), JHEP **03**, 080 (2005), arXiv: hep-ph/0501067
[34] R. D. Ball *et al.* (NNPDF Collaboration), Nucl. Phys. B **867**, 244 (2013), arXiv: 1207.1303[hep-ph]
[35] R. D. Ball *et al.* (NNPDF Collaboration), JHEP **04**, 040 (2015), arXiv: 1410.8849[hep-ph]

[36] R. D. Ball *et al*. (NNPDF Collaboration), Eur. Phys. J. C **77**, 663 (2017), arXiv: 1706.00428[hep-ph]

[37] S. Carrazza and J. Cruz-Martinez, Eur. Phys. J. C **79**, 676 (2019), arXiv: 1907.05075[hep-ph]

[38] R. D. Ball *et al*. (NNPDF Collaboration), Eur. Phys. J. C **82**, 428 (2022), arXiv: 2109.02653[hep-ph]

[39] G. Kanwar, M. S. Albergo, D. Boyda *et al*., Phys. Rev. Lett. **125**, 121601 (2020), arXiv: 2003.06413[hep-lat]

[40] M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **378**, 686 (2019), arXiv: 1711.10561[cs.AI]

[41] G. E. Karniadakis, I. G. Kevrekidis, L. Lu *et al*., Nature Rev. Phys. **3**, 422 (2021)

[42] W. Kou, X. Lin, B. Guo *et al*., Phys. Rev. D **112**, 056027 (2025), arXiv: 2506.03513[hep-ph]

[43] W. Kou and X. Chen, *Universality of Gluon Saturation from Physics-Informed Neural Networks*, (2026), arXiv: 2601.16391[hep-ph]

[44] S. W. Dai, F. P. Li, L. G. Pang *et al*., *Physics-Informed Global Extraction of the Universal Small-x Dipole Amplitude*, (2026), arXiv: 2603.08008[hep-ph]

[45] M. Gao, Z. B. Kang, J. Penttala *et al*., *Determination of the initial condition for the Balitsky-Kovchegov equation with transformers*, (2025), arXiv: 2510.26779[hep-ph]

[46] T. I. Baihaqi, C. Setyadi, Z. Akbar *et al*., *Determination of the HERA coherent diffractive J/ψ production cross section via artificial neural network*, (2025), arXiv: 2512.21704[hep-ph]

[47] A. Donnachie and P. V. Landshoff, Phys. Lett. B **296**, 227 (1992), arXiv: hep-ph/9209205

[48] T. Gutsche, V. E. Lyubovitskij, and I. Schmidt, Phys. Rev. D **97**, 054011 (2018), arXiv: 1712.08410[hep-ph]

[49] R. C. Brower, J. Polchinski, M. J. Strassler *et al*., JHEP **12**, 005 (2007), arXiv: hep-th/0603115

[50] R. C. Brower, M. Djuric, I. Sarcevic *et al*., JHEP **11**, 051 (2010), arXiv: 1007.2259[hep-ph]

[51] C. A. Ballon Bayona, H. Boschi-Filho, and N. R. F. Braga, JHEP **03**, 064 (2008), arXiv: 0711.0221[hep-th]

[52] Z. Abidin and C. E. Carlson, Phys. Rev. D **79**, 115003 (2009), arXiv: 0903.4818[hep-ph]

[53] J. H. Gao and B. W. Xiao, Phys. Rev. D **80**, 015025 (2009), arXiv: 0904.2870[hep-ph]

[54] S. J. Brodsky, G. F. de Teramond, H. G. Dosch *et al*., Phys. Rept. **584**, 1 (2015), arXiv: 1407.8131[hep-ph]

[55] S. Mamedov and S. Taghiyeva, J. Korean Phys. Soc. **85**, 541 (2024), arXiv: 2208.07115[hep-ph]

[56] Y. Hatta, E. Iancu, and A. H. Mueller, JHEP **01**, 026 (2008), arXiv: 0710.2148[hep-th]

[57] D. Jorrin and M. Schvellinger, Phys. Rev. D **106**, 066024 (2022), arXiv: 2207.02984[hep-ph]

[58] I. Borsa, D. Jorrin, R. Sassot *et al*., Phys. Rev. D **108**, 056024 (2023), arXiv: 2308.01975[hep-ph]

[59] A. Karpatne, G. Atluri, J. H. Faghmous *et al*., IEEE Trans. Knowl. Data Eng. **29**, 2318 (2017)

[60] J. Willard, X. Jia, S. Xu *et al*., ACM Computing Surveys **55**, 1 (2022)

[61] A. Raue, C. Kreutz, T. Maiwald *et al*., Bioinformatics **25**, 1923 (2009)

[62] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems* (V. H. Winston & Sons, Washington, DC, 1977).

[63] D. V. Raman, J. Anderson, and A. Papachristodoulou, Phys. Rev. E **95**, 032314 (2017), arXiv: 1607.07705[physics.data-an]

[64] M. F. Kasim, D. Watson-Parris, L. Deaconu, S. Oliver, P. Hatfield, D. H. Froula, G. Gregori, M. Jarvis, S. Khatiwala, J. Korenaga, *et al*., Machine Learning: Science and Technology **3**, 015013 (2022)

[65] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, in *International Conference on Learning Representations* (2014), arXiv: 1412.6980[cs.LG]

[66] M. C. Kennedy and A. O'Hagan, JRSS-B **63**, 425 (2001)

[67] J. Ye, J. Am. Stat. Assoc. **93**, 120 (1998)

[68] P. J. Huber, Robust estimation of a location parameter, in *Breakthroughs in statistics: Methodology and distribution* (Springer, 1992) p. 492

[69] R. Pascanu, T. Mikolov, and Y. Bengio, arXiv: 1211.5063[cs.LG]

[70] A. de Myttenaere, B. Golden, B. Le Grand *et al*., Neurocomputing **192**, 38 (2016)

[71] S. Elfwing, E. Uchibe, and K. Doya, arXiv: 1702.03118[cs.LG]

[72] H. Abramowicz *et al*. (H1, ZEUS), Eur. Phys. J. C **75**, 580 (2015), arXiv: 1506.06042[hep-ex]

[73] J. Dudek *et al*., Eur. Phys. J. A **48**, 187 (2012), arXiv: 1208.1244[hep-ex]

[74] R. Abdul Khalek *et al*., Nucl. Phys. A **1026**, 122447 (2022), arXiv: 2103.05419[physics.ins-det]

[75] D. P. Anderle *et al*., Front. Phys. (Beijing) **16**, 64701 (2021), arXiv: 2102.09222[nucl-ex]

[76] P. Agostini *et al*. (LHeC, FCC-he Study Group), J. Phys. G **48**, 110501 (2021), arXiv: 2007.14491[hep-ex]