RECONSTRUCTING DOUBLE DISTRIBUTIONS USING FEM AND NEURAL FIELDS

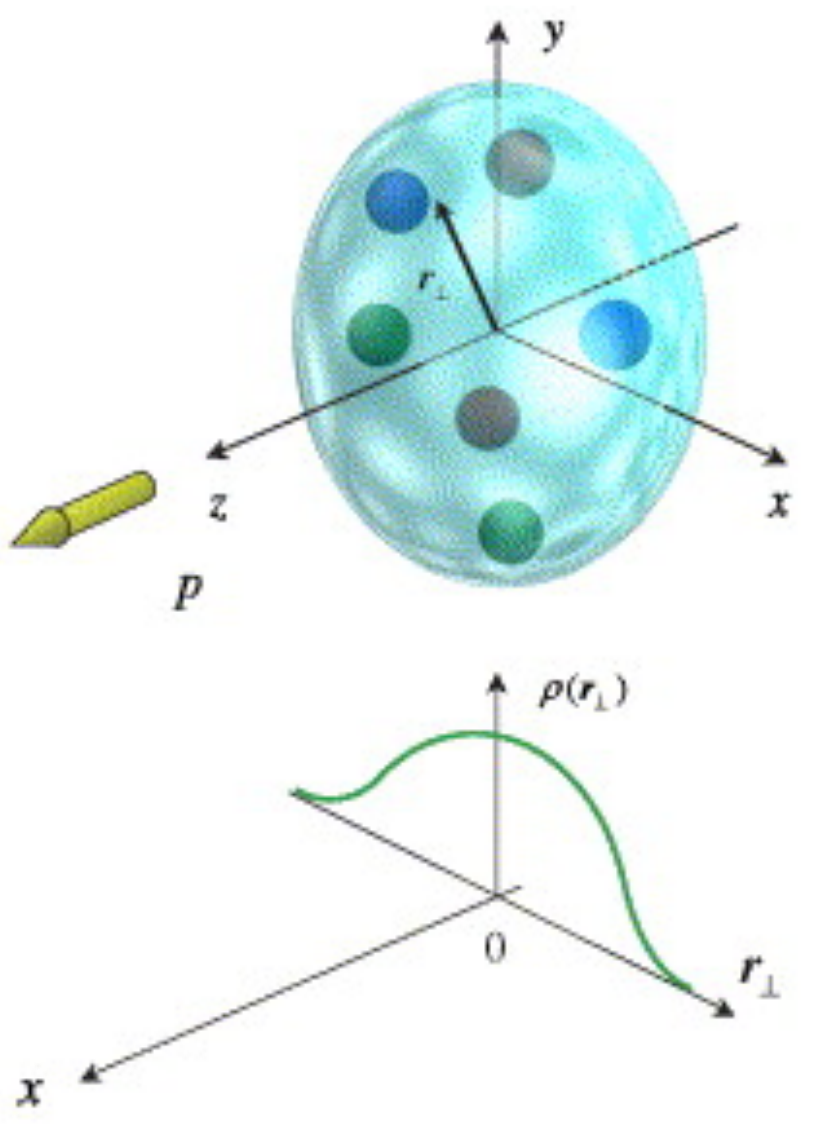
MARIJA ČUIĆ
IN COLLABORATION WITH CÉDRIC MEZRAG AND JÉRÔME BOBIN

Progress in algorithms and numerical tools for QCD
June 9 - 10, Orsay, France

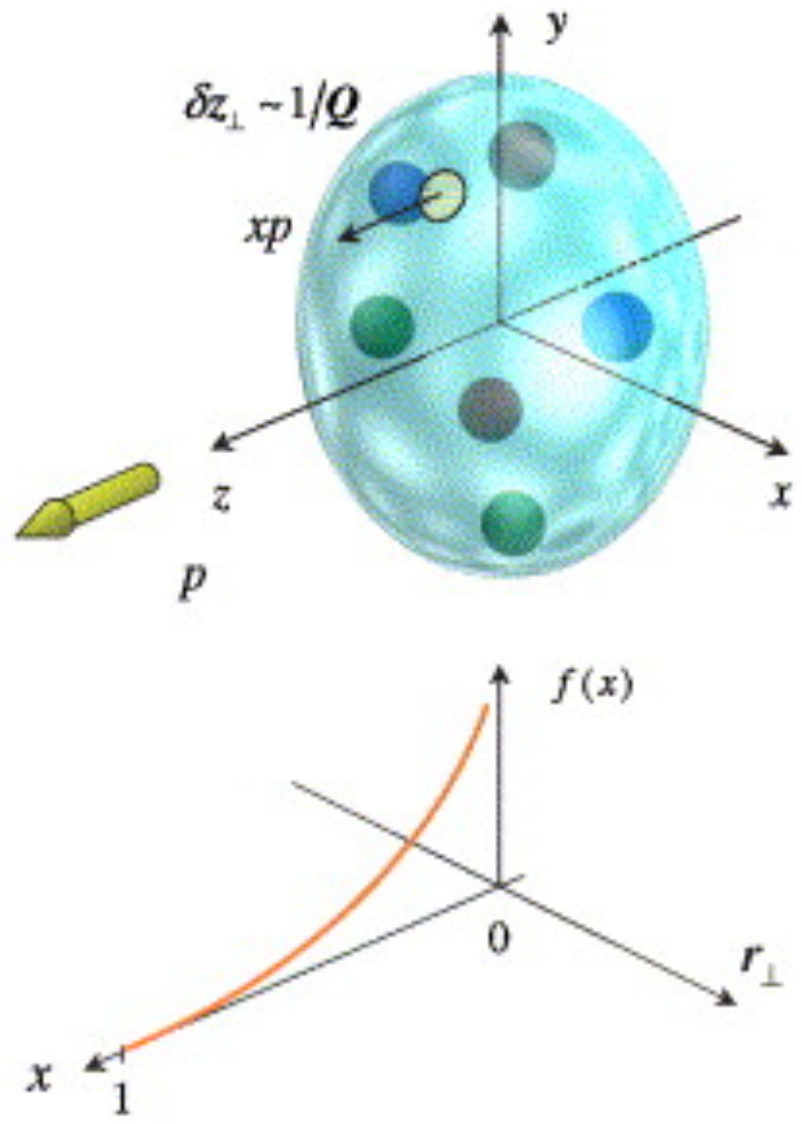
HADRON FEMTOGRAPHY

We want to uncover the inner structure of nucleons:

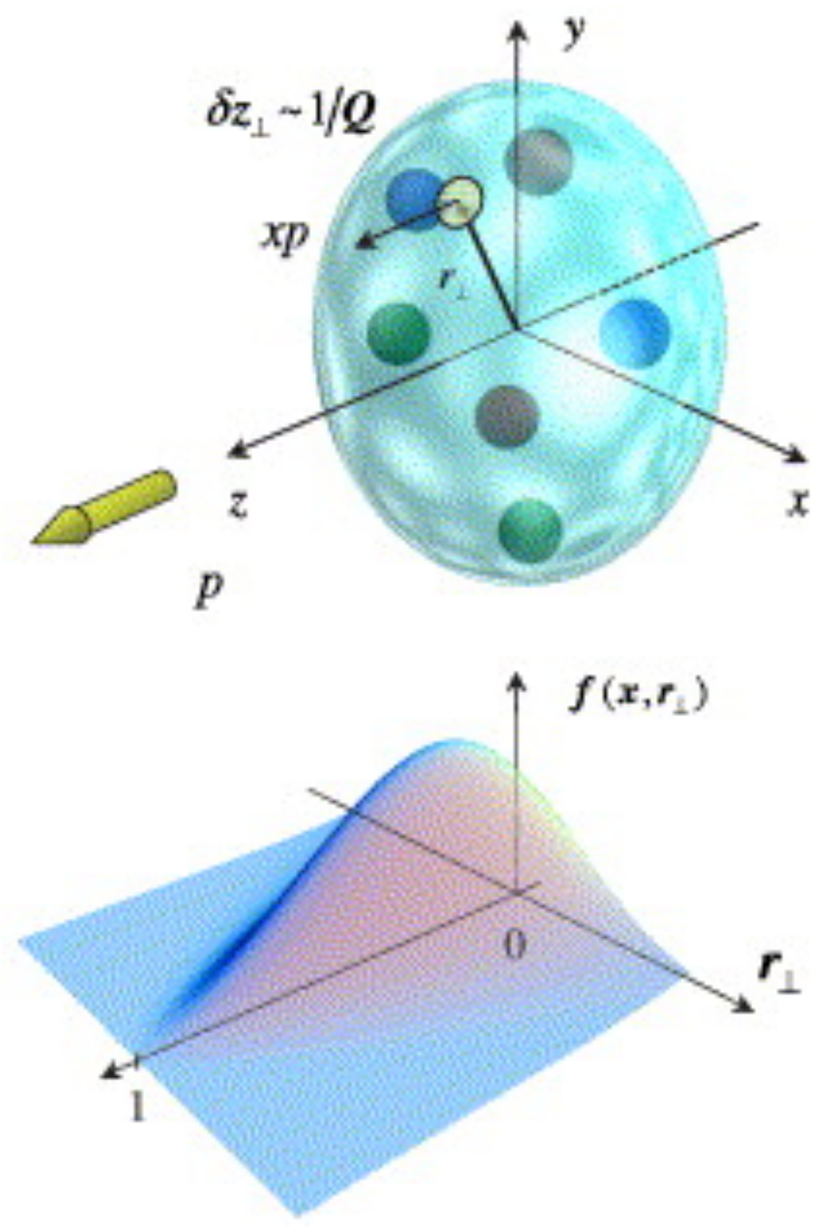
- **3D position and momenta of partons (quarks and gluons)**
- **spin and angular momentum distribution**
- **mechanical properties such as pressure**



Form factors, transverse position of partons



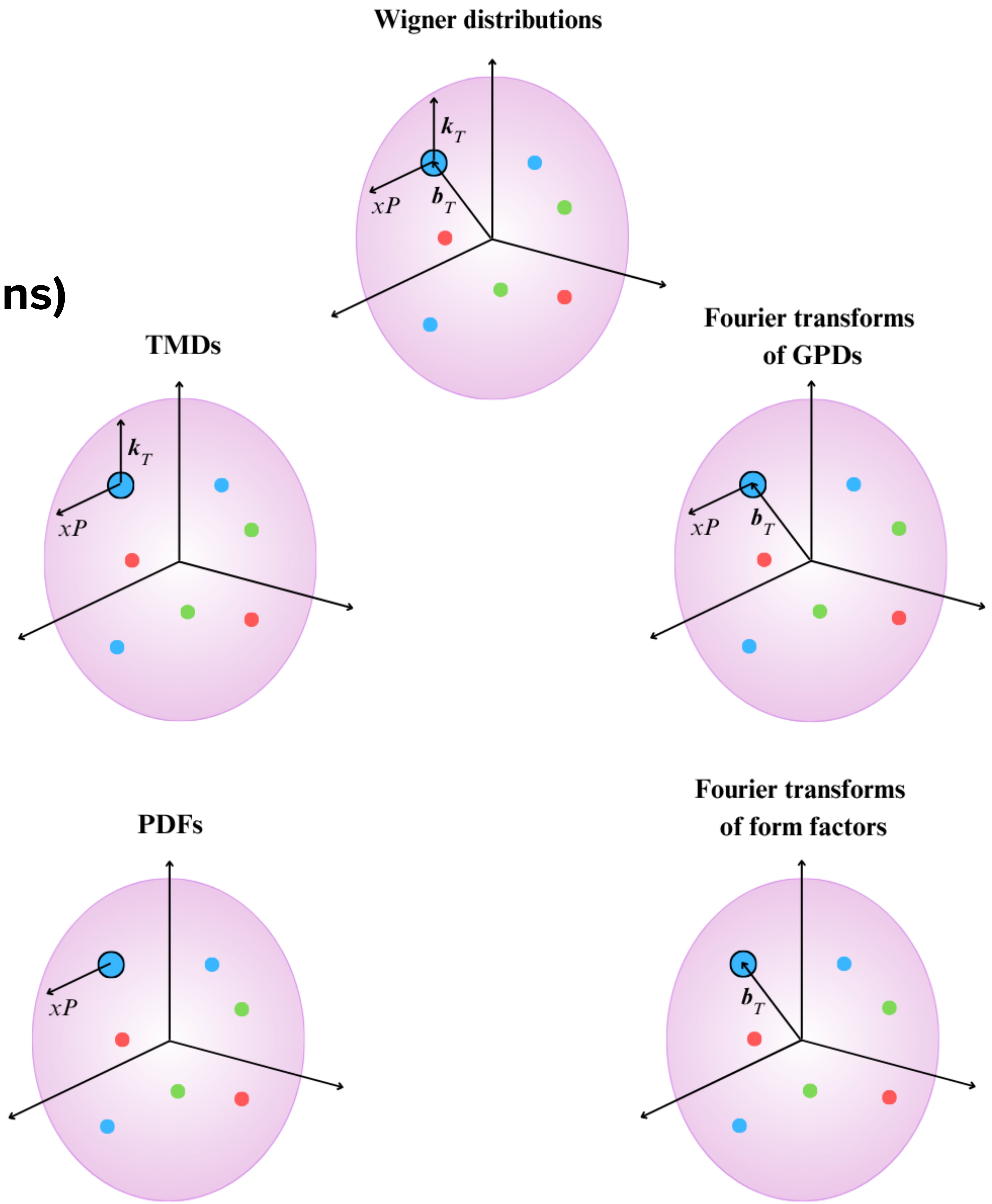
PDFs, longitudinal momentum of partons



GPDs, transverse position and longitudinal momentum of partons

Hofstadter, et. al., Phys. Rev. 91, 422
Hofstadter, Rev. Mod. Phys. 28, 214

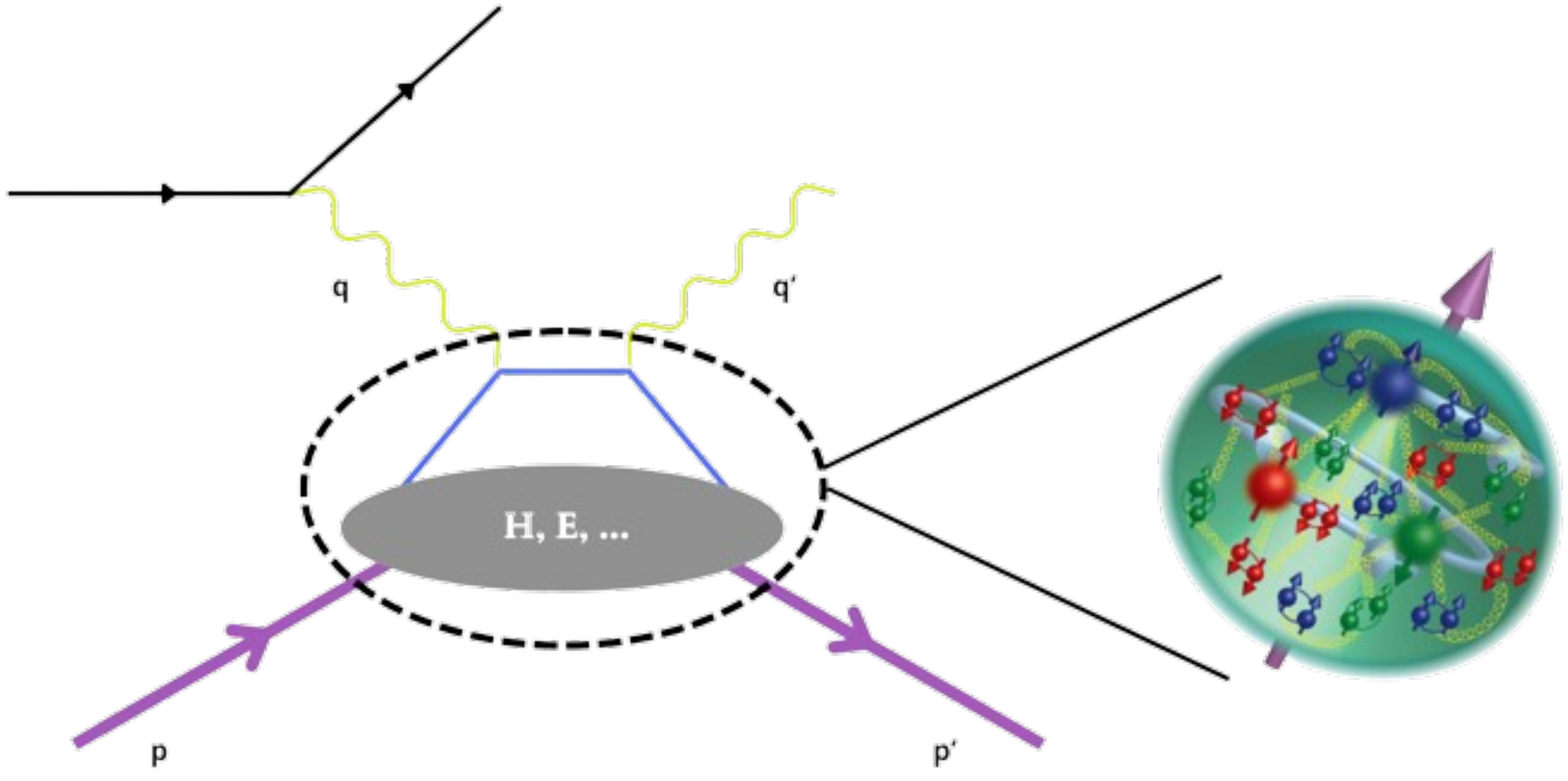
Feynman, Phys. Rev. Lett. 23, 1415



Hierarchy of distributions

INVERSE PROBLEM

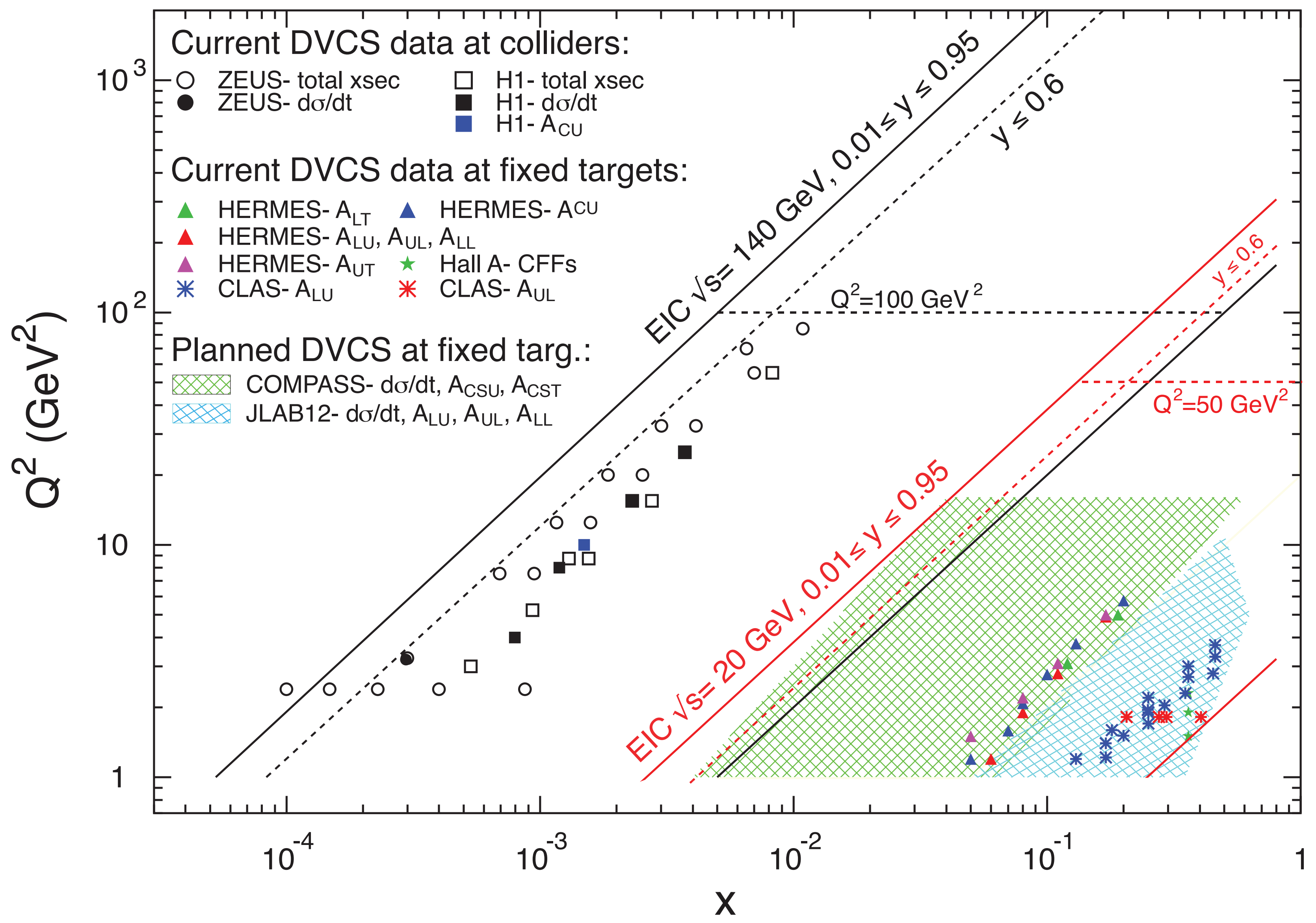
DVCS: $\ell + p \to \ell + \gamma + p$



$$\frac{d^5\sigma_{DVCS}}{dx_B dy dt d\phi d\varphi} \propto 4\left(1 - x_B\right)\left(\boxed{|\mathscr{H}|^2} + |\widetilde{\mathscr{H}}|^2\right) + \ldots$$

$$\mathscr{H}^A\left(\xi, \Delta^2, Q^2\right) = \int_{-1}^{1} \frac{dx}{2\xi} \underbrace{{}^A T\left(x, \xi \mid \alpha_s\left(\mu_R\right), \left\{\frac{Q^2}{\mu^2}\right\}\right)}_{\text{hard scale}} \underbrace{\boxed{H^A}\left(x, \xi, \Delta^2, \mu_F^2\right)}_{\text{soft scale}}$$

Limited experimental coverage:



Inverse problem: observables → GPDs inside convolution
→ ill-posed problem!

Additionally, experiment dominantly probes the DGLAP region, specifically at LO:

$$\mathfrak{Im}\mathscr{H}\left(\xi, t, Q^2\right) = \pi \sum_{q=u,d,\ldots} e_q^2 \left[H^q\left(x = \xi, \xi, t, \mu_F^2\right) - H^q\left(x = -\xi, \xi, t, \mu_F^2\right)\right].$$

Any extraction beyond that is unreliable. How do we access the ERBL region?

DOUBLE DISTRIBUTION REPRESENTATION OF GPDS

GPD definition: $$F^q_{\Lambda_2,\Lambda_1} = \frac{1}{2}\int \frac{\mathrm{d}z^-}{2\pi} e^{ixP^+z^-} \langle P_2, \Lambda_2 \,|\, \bar{\psi}^q\left(-\frac{z}{2}\right)\gamma^+\psi^q\left(\frac{z}{2}\right)\,|\, P_1, \Lambda_1\rangle \Big|_{\substack{z^+_\perp = 0 \\ z_\perp = 0}}$$

$$= \frac{1}{2P^+}\left[H^q(x,\xi,t)\bar{u}\left(P_2,\Lambda_2\right)\gamma^+ u\left(P_1,\Lambda_1\right) + E^q(x,\xi,t)\bar{u}\left(P_2,\Lambda_2\right)\frac{i\sigma^{+\mu}\Delta_\mu}{2M_H}u\left(P_1,\Lambda_1\right)\right]$$

Alternative parametrization of the matrix element:

$$\left\langle P+\frac{\Delta}{2}\right| \bar{\psi}^q\left(-\frac{z^-}{2}\right)\gamma^+\psi^q\left(\frac{z^-}{2}\right)\left|P-\frac{\Delta}{2}\right\rangle = \iint_\Omega \mathrm{d}\beta\,\mathrm{d}\alpha e^{-i\beta P^+z^- + i\alpha\frac{1}{2}\Delta^+z^-}\left(2P^+ f(\beta,\alpha) - \Delta^+ g(\beta,\alpha)\right)$$

double distributions

Radon transform: $H(x,\xi) = \iint_\Omega \mathrm{d}\beta\,\mathrm{d}\alpha\delta(x-\beta-\xi\alpha)(f(\beta,\alpha)+\xi g(\beta,\alpha))$

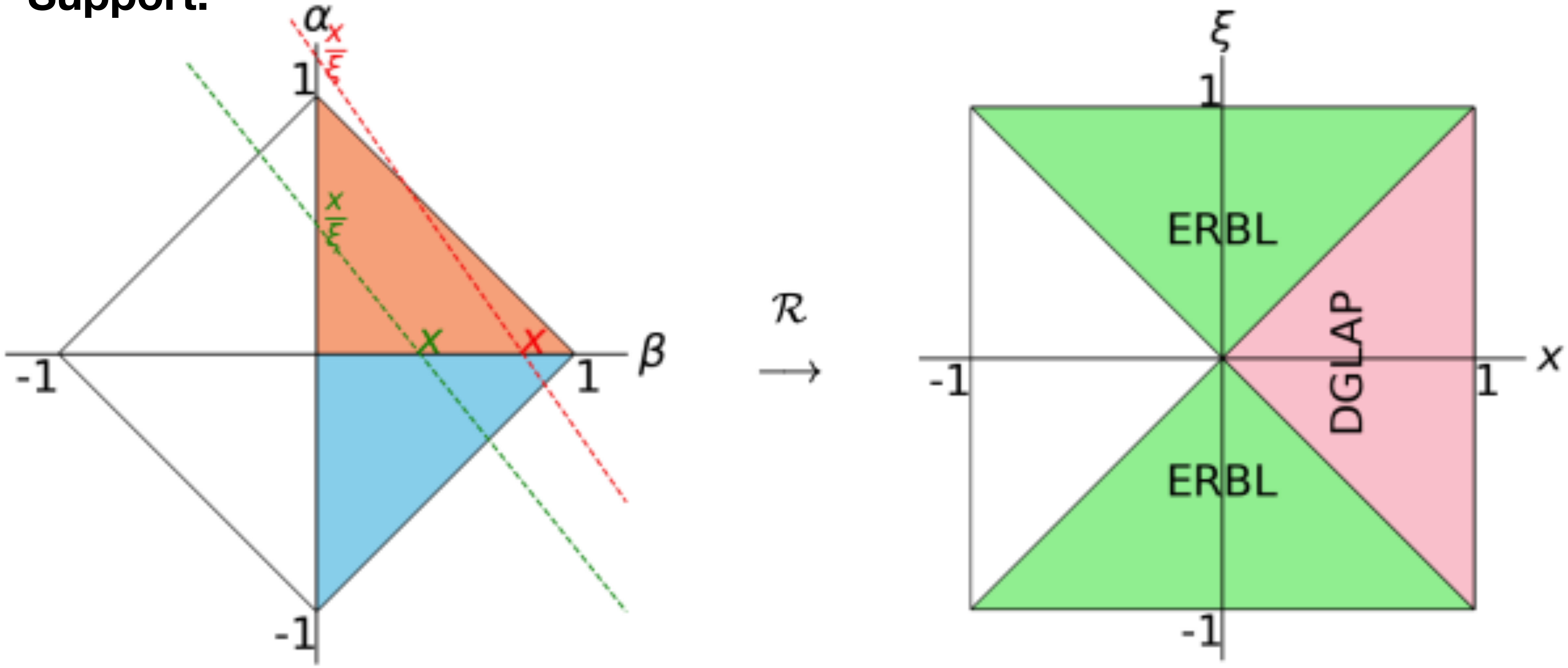
Useful property - polynomiality (Lorentz covariance):

$$H^m(\xi) = \int_{-1}^{1} x^m H(x,\xi)\mathrm{d}x = \sum_{\substack{k=0 \\ k\ \text{even}}}^{m+1} H^m_k \xi^k \Rightarrow \int_{-1}^{1} x^m H(x,\xi)\mathrm{d}x = \iint_\Omega \mathrm{d}\beta\,\mathrm{d}\alpha(\beta+\xi\alpha)^m(f(\beta,\alpha)+\xi g(\beta,\alpha))$$

[arXiv:hep-ph/9704207]

RADON TRANSFORM

Support:



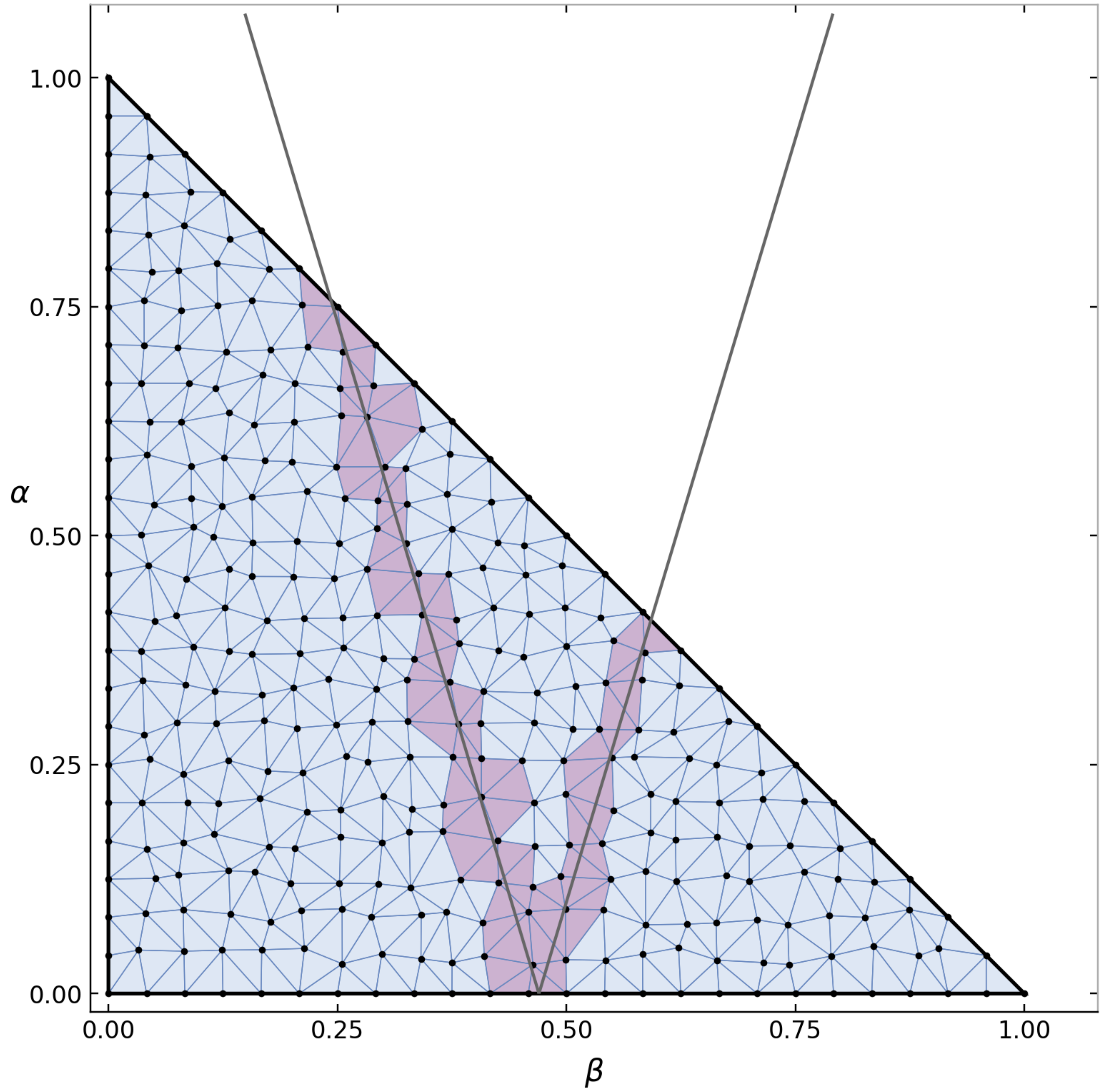
We obtain a 2D function by performing integrals over the lines $x - \beta - \xi\alpha = 0$. The negative α domain is equivalent to sampling the positive domain with the reflected lines $x - \beta + \xi\alpha = 0$. The $\beta < 0$ is probed by the antiquark distributions, which we do not consider at this point.

Symmetries: $H(x, -\xi, t) = H(x, \xi, t) \Rightarrow f(\beta, -\alpha) = f(\beta, \alpha), \quad g(\beta, -\alpha) = -g(\beta, \alpha)$

Goal: extract GPDs in the DGLAP region, obtain the corresponding double distribution, extend it to the ERBL region, invert the Radon transform and obtain the GPD over the whole domain. Analytically hard to perform, we rely on numerics. [arXiv:2401.12013]

FINITE ELEMENTS METHOD

$$H(x,\xi) = \mathscr{R}[h(\beta,\alpha)] \Rightarrow \begin{pmatrix} \vdots \\ H_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \cdots & \\ \vdots & R_{ij} \end{pmatrix} \begin{pmatrix} \vdots \\ h_j \\ \vdots \end{pmatrix}$$



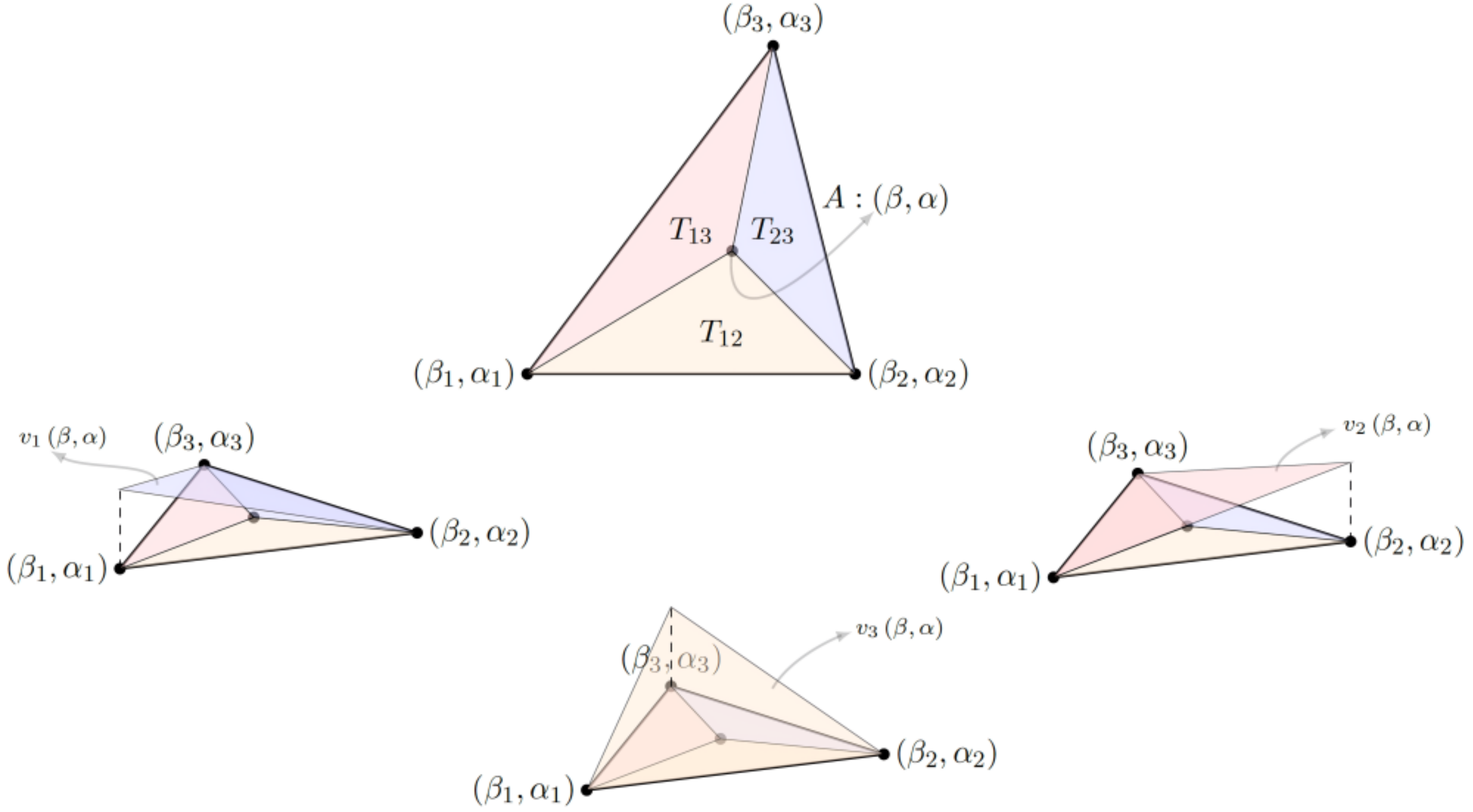
Radon matrix is a sparse matrix, we need to oversample to reach full rank!

1. **Discretize the domain with Delaunay triangulation:**
   $\Omega^+ = \bigcup_e \Omega_e^+$
2. **Interpolate the discrete double distribution**
   $h(\beta,\alpha) \to P(\beta,\alpha) = \sum_e P_e(\beta,\alpha)\theta\left(\Omega_e^+\right)$
   - **A. On each element of the triangulation define a simple function, such as linear polynomial**
     $P_e(\beta,\alpha) = c_e + b_e\beta + a_e\alpha$
   - **B. Calculate the polynomials using barycentric coordinates**
     $$P_e(\beta,\alpha) = \begin{vmatrix} \beta & \alpha & 1 \\ \beta_2 & \alpha_2 & 1 \\ \beta_3 & \alpha_3 & 1 \end{vmatrix} h_1^e + \begin{vmatrix} \beta_1 & \alpha_1 & 1 \\ \beta & \alpha & 1 \\ \beta_3 & \alpha_3 & 1 \end{vmatrix} h_2^e + \begin{vmatrix} \beta_1 & \alpha_2 & 1 \\ \beta_2 & \alpha_2 & 1 \\ \beta & \alpha & 1 \end{vmatrix} h_3^e$$
3. **Choose sampling lines (x_i,ξ_i), usually oversample**
4. **Find the intersection points of each line in each mesh element**
5. **Calculate the Radon integral for each node in the mesh**
6. **Add up all sampled elements**

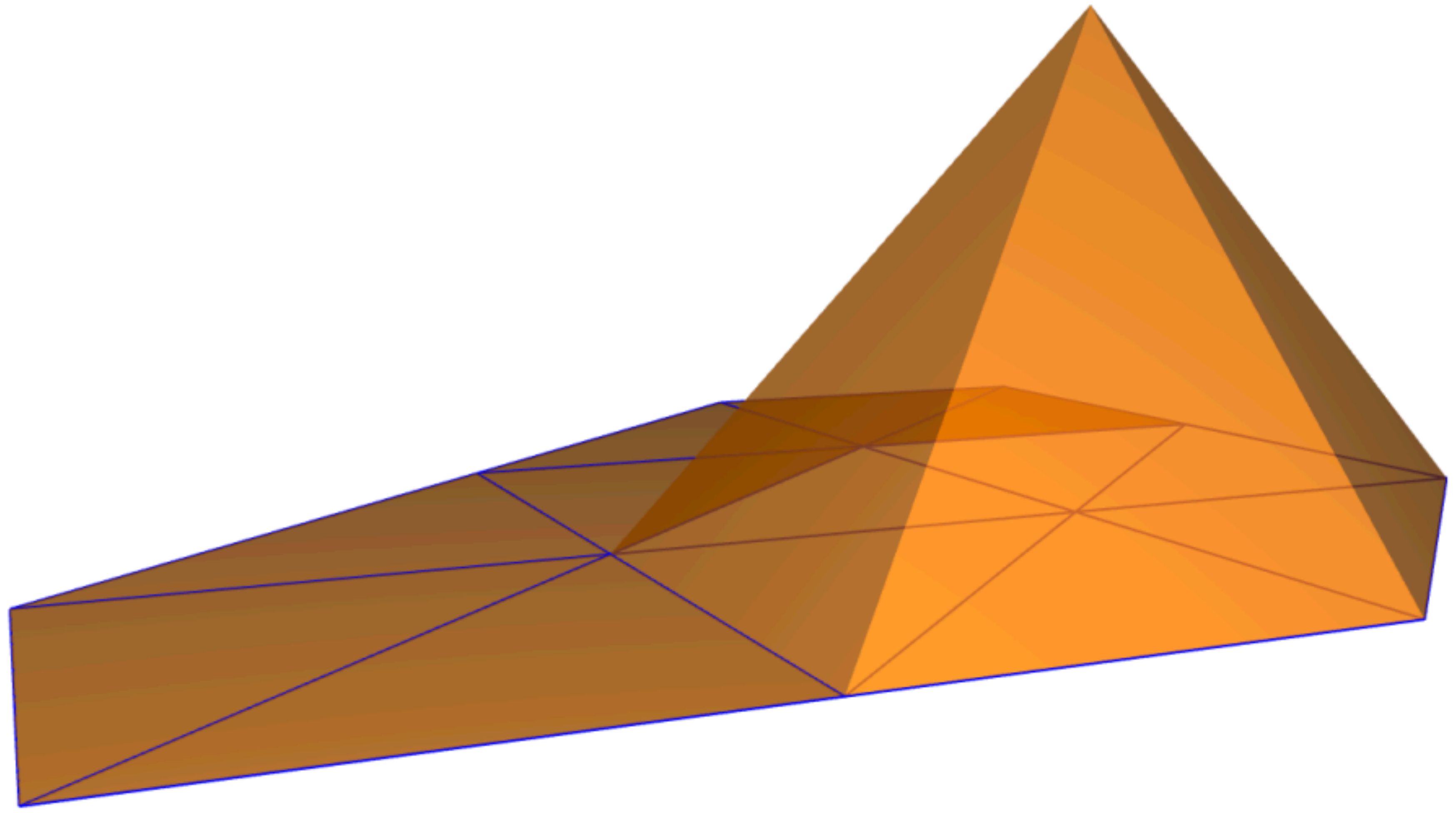
$$H(x_i,\xi_i) = \sum_j h_j \int_0^1 \mathrm{d}\beta \int_0^1 \mathrm{d}\alpha\delta(x_i - \beta - \xi_i\alpha)v_j(\beta,\alpha)\Theta(1-\alpha-\beta) + \text{reflected}$$

$$= \sum_{\text{sampled } \Omega_e^+} \sum_{j\in\Omega_e^+} h_j \left\{ \left(c_j + b_j x_i\right)(\alpha_{in,e} - \alpha_{out,e}) + \left(a_j - b_j\xi_i\right) \frac{\alpha_{in,e}^2 - \alpha_{out,e}^2}{2} \right\} + \text{reflected}$$

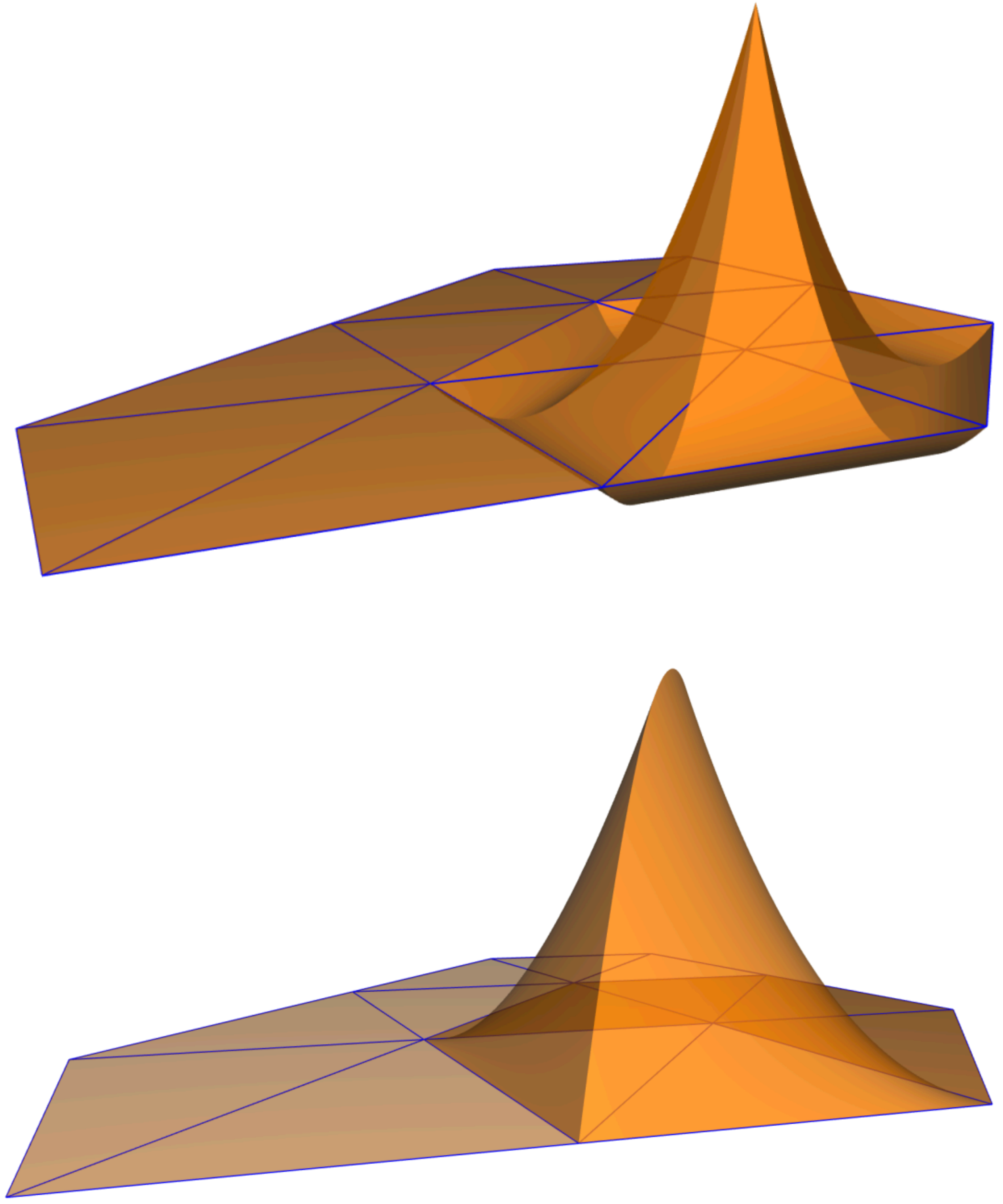
Linear polynomials on each element:



P1 Lagrange polynomials

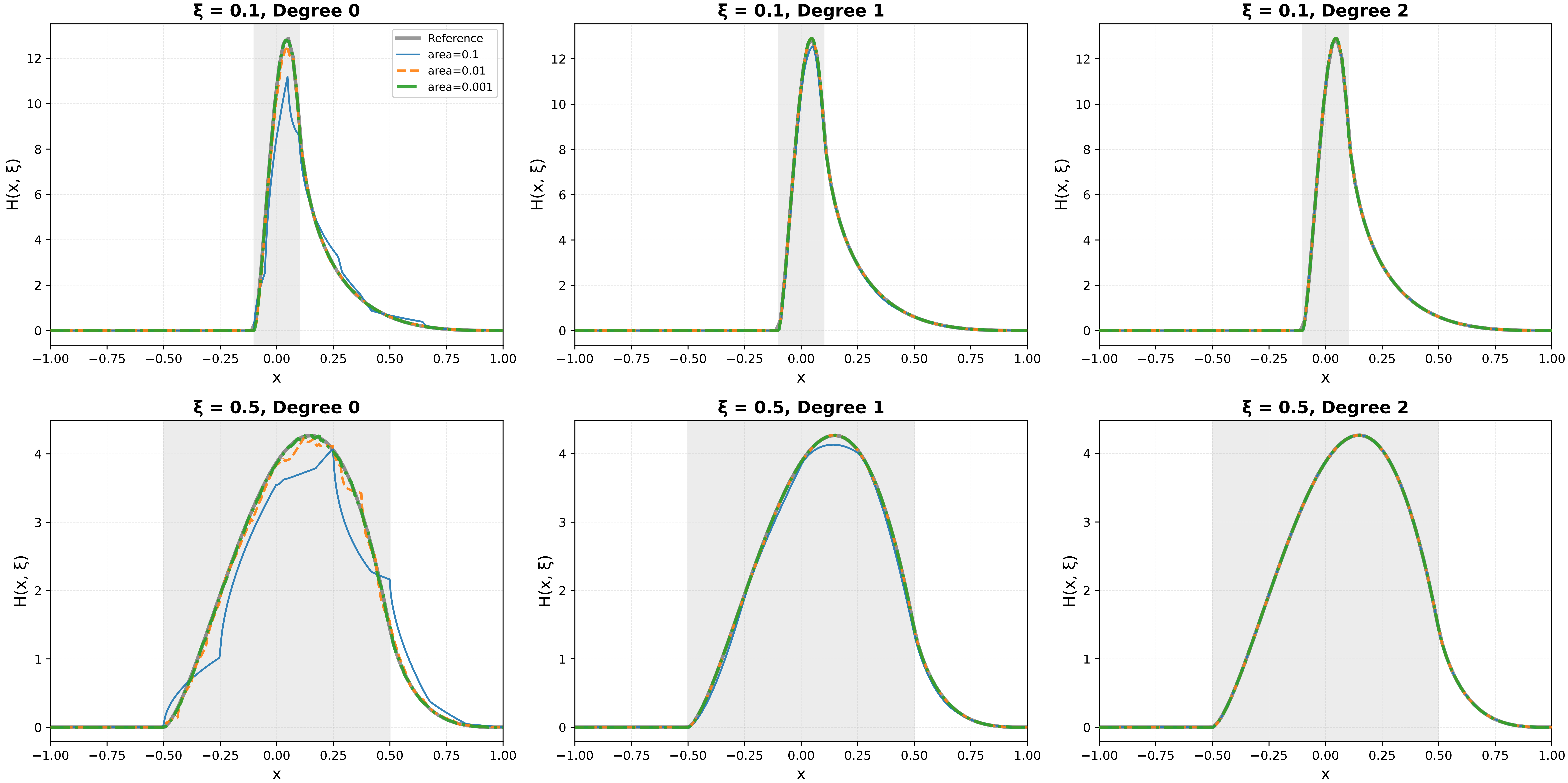


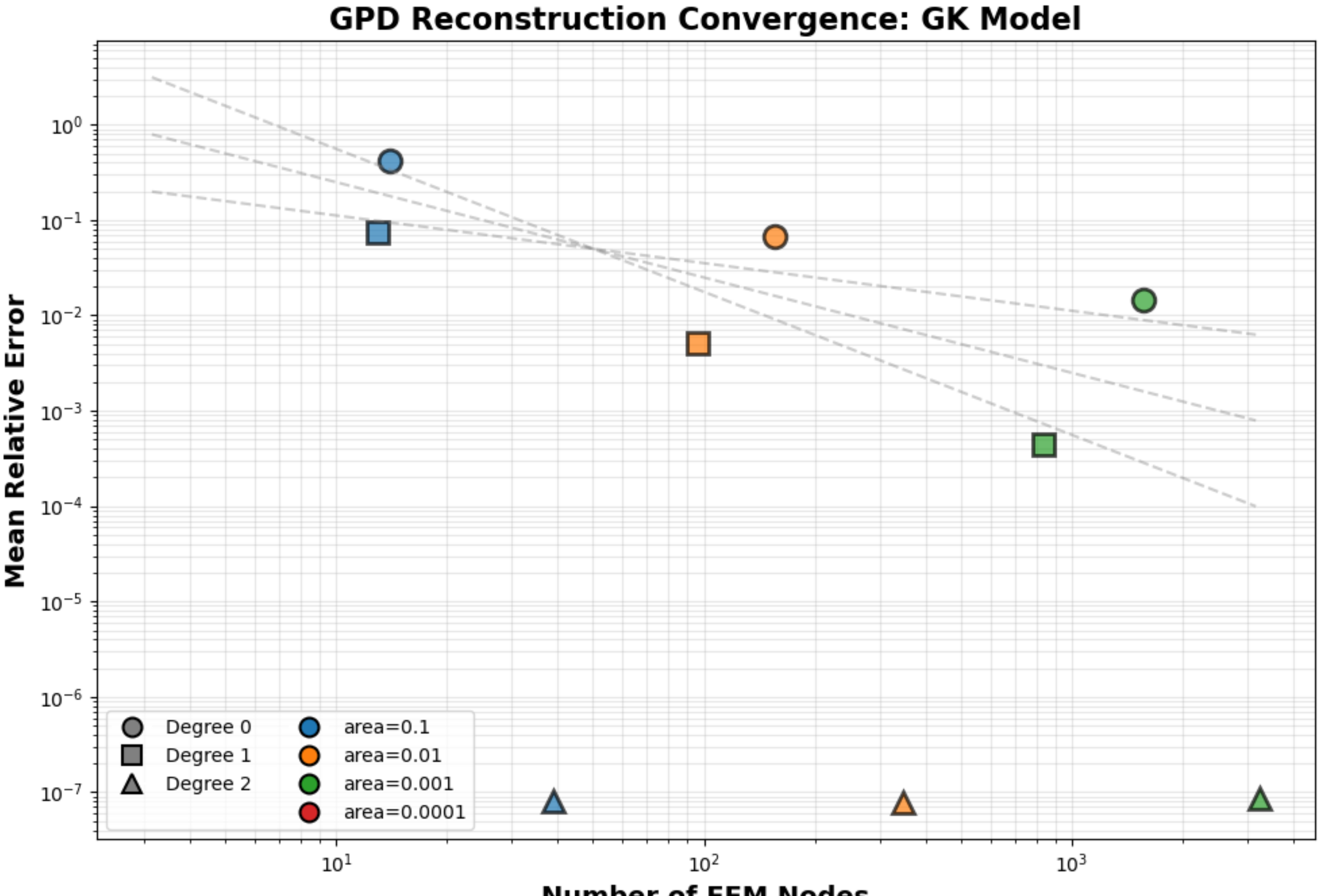
P2 Lagrange polynomials



RESULTS FOR GK MODEL

$$F(\beta,\alpha) = 2\,N_u\,\beta^{-\alpha_0}\,(1-\beta)^{b_u}\,\frac{\Gamma(2b+2)}{2^{2b+1}\,\Gamma(b+1)^2}\cdot\frac{\left[(1-\beta)^2-\alpha^2\right]^b}{(1-\beta)^{2b+1}}$$




GPD Reconstruction Convergence: GK Model
Mean Relative Error
Number of FEM Nodes
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^1
10^2
10^3
Degree 0
Degree 1
Degree 2
area=0.1
area=0.01
area=0.001
area=0.0001

WHAT IS THE SHADOW GPD CONTRIBUTION?

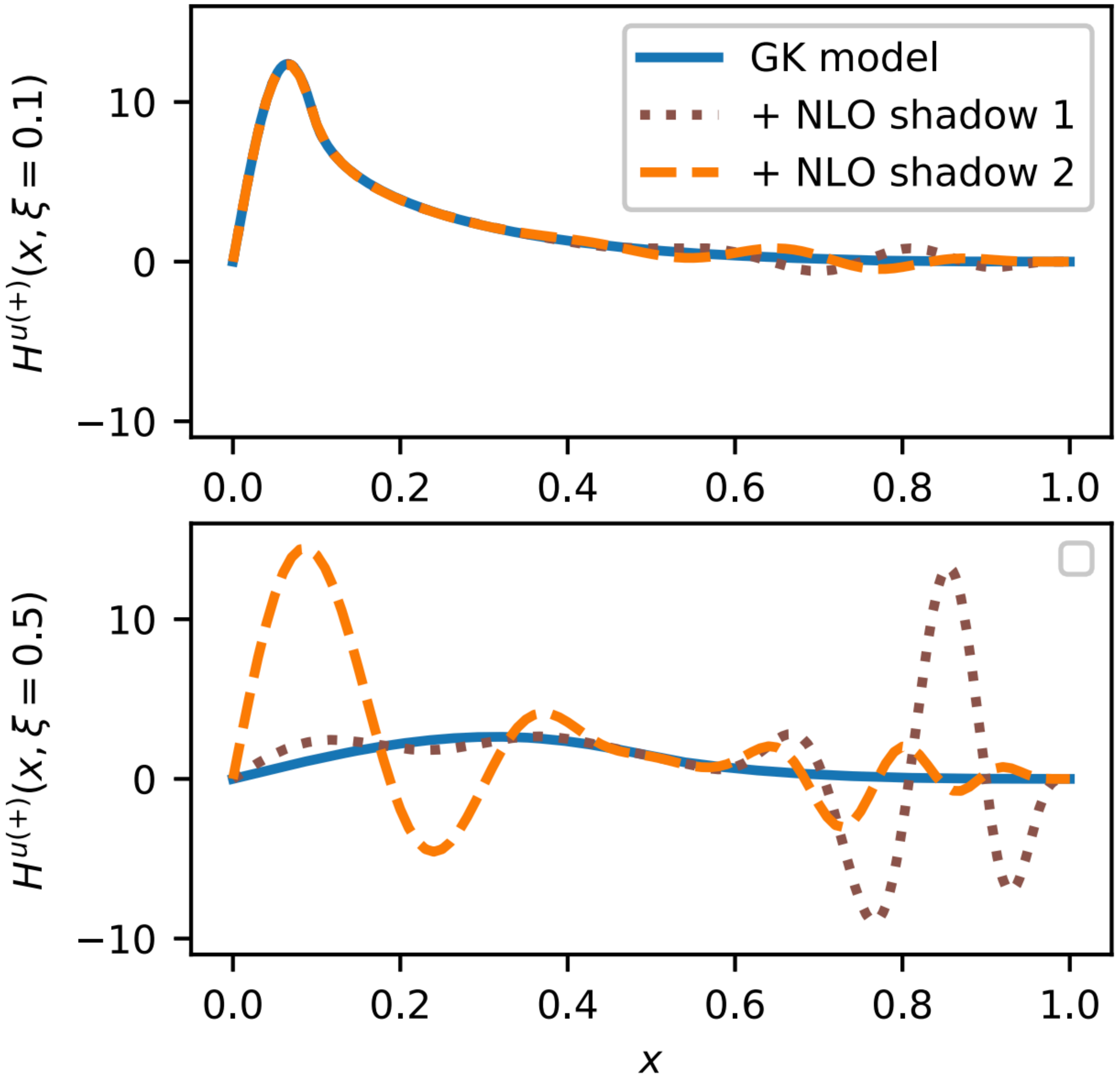
SHADOW GPDS

Result of the inverse problem: $F_S^A(x,\xi,t) = F_E^A(x,\xi,t) - F_T^A(x,\xi,t)$

Must satisfy all properties of the GPD and vanish in physical observables:

- **Polynomiality**
- **Zero contribution to CFF:** $\sum_A {}^A T\left(x, \xi \mid \alpha_s(\mu_R), \left\{\frac{Q^2}{\mu^2}\right\}\right) \otimes H^A\left(x, \xi, \Delta^2, \mu_F^2\right)$
- **Zero contribution to PDF:** $H_S^A(x,0,0) = 0$

[arXiv:2107.11312]



We test polynomial shadow DDs:

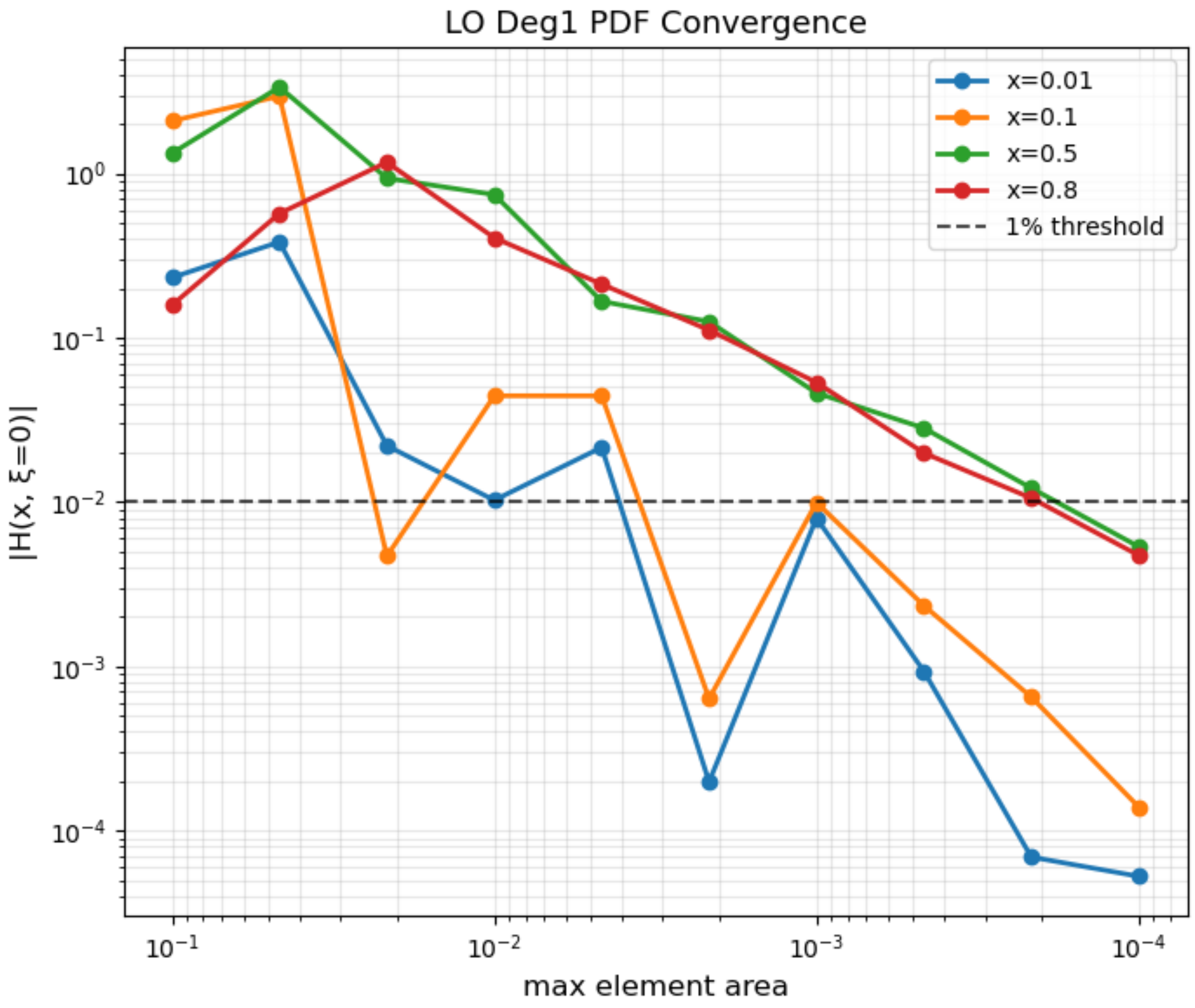
- **LO shadow DD** [arXiv:2104.03836v1]**:**

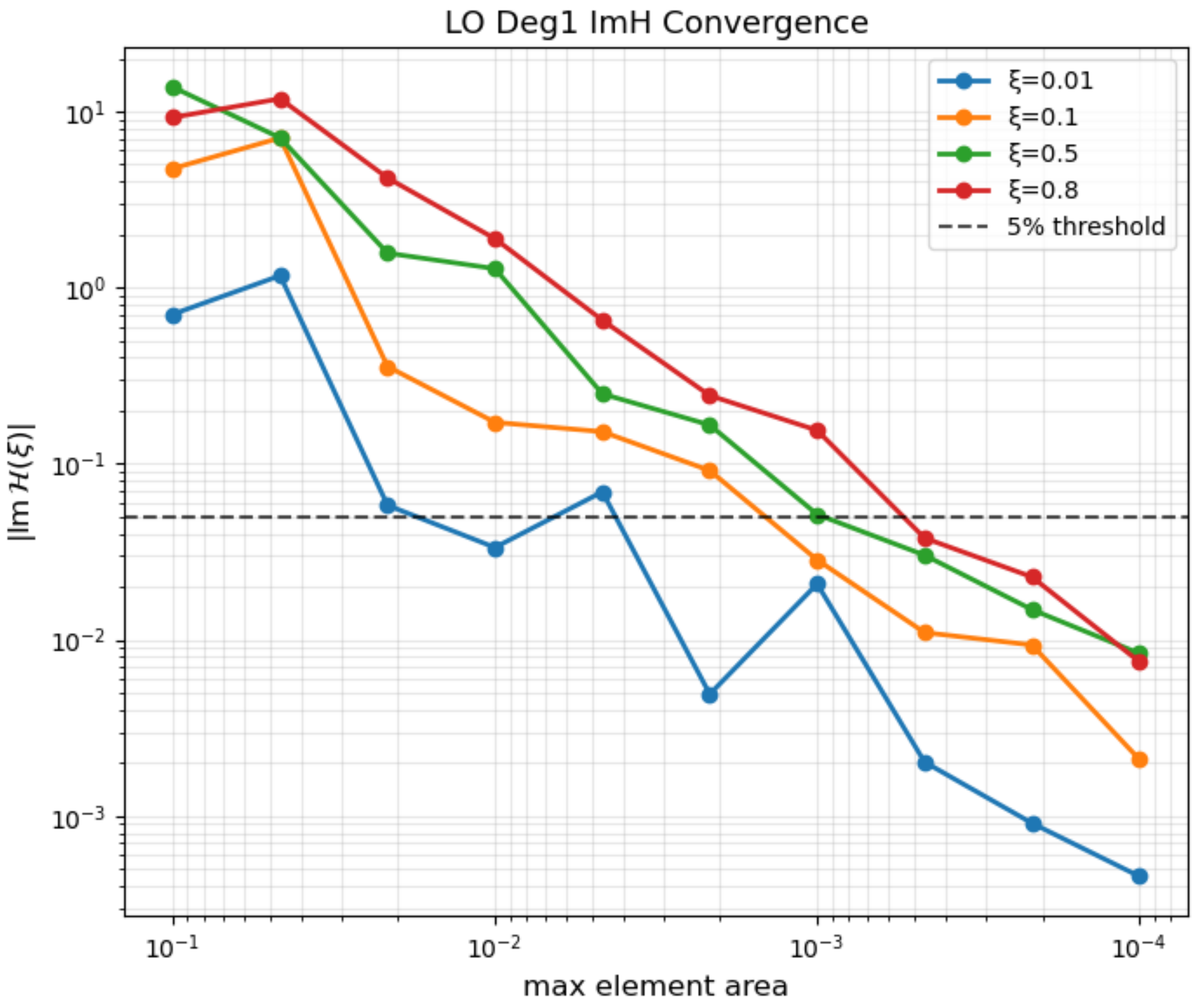
$$F_N^{q(+)}(\beta,\alpha) = \beta^{N-8}\left[\alpha^8 - \frac{28}{9}\alpha^6\left(\frac{N^2-3N+20}{(N+1)N} + \beta^2\right) + \frac{10}{3}\alpha^4\left(\frac{N^2-7N+40}{(N+1)N} + \frac{2\left(N^2-3N+44\right)}{3(N+1)N}\beta^2 + \beta^4\right)\right.$$
$$\left. - \frac{4}{3}\alpha^2\left(\frac{N^2-11N+60}{(N+1)N} - \frac{N-8}{N}\beta^2 - \frac{N^2-3N-28}{(N+1)N}\beta^4 + \beta^6\right) + \frac{1}{9}\left(1-\beta^2\right)^2\left(\frac{N^2-15N+80}{(N+1)N} - \frac{2(N-8)}{N}\beta^2 + \beta^4\right)\right], N \geq 9, \textbf{odd}$$

- **NLO shadow DD: order 27 polynomial**

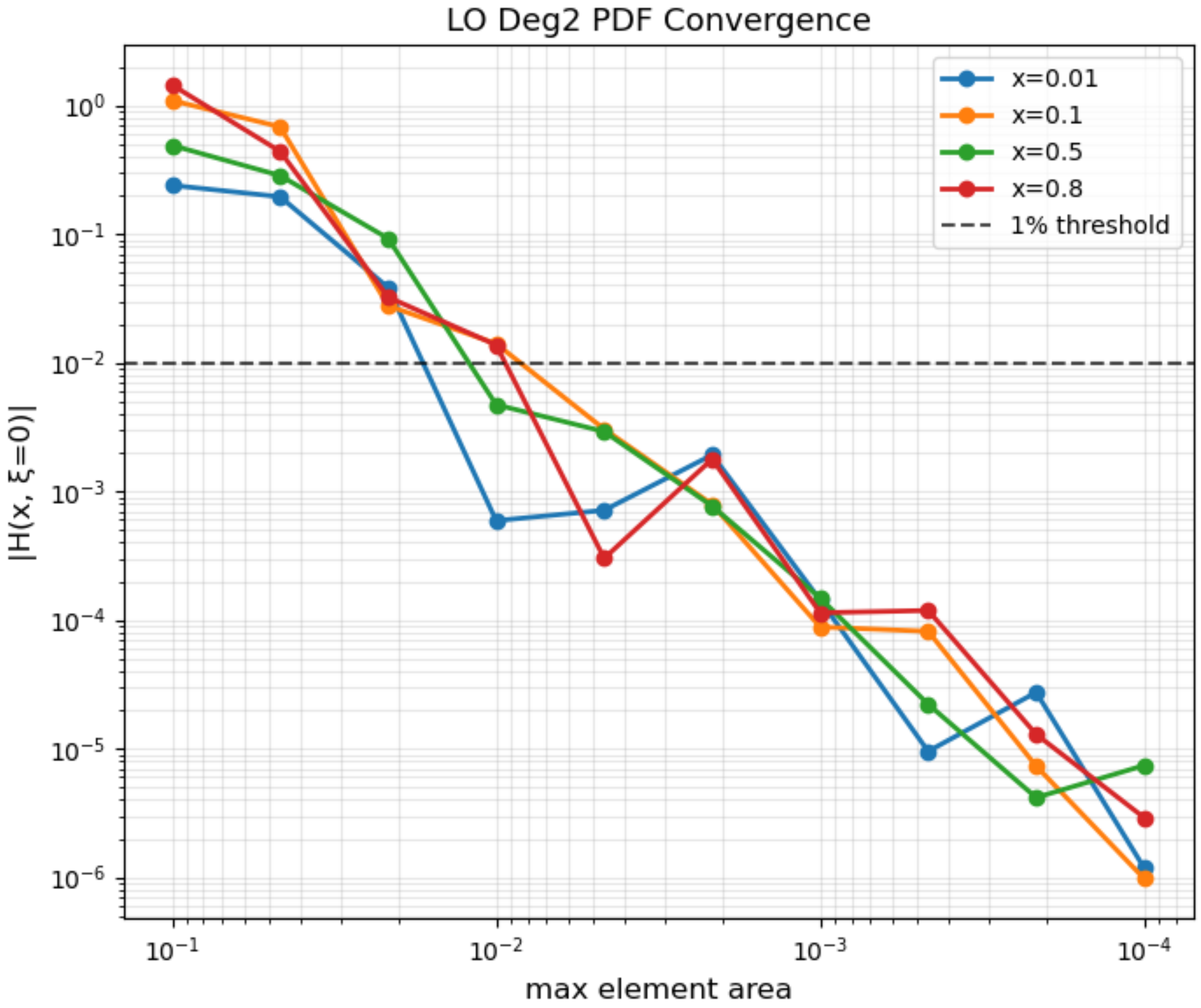
LO POLYNOMIAL SHADOW DD

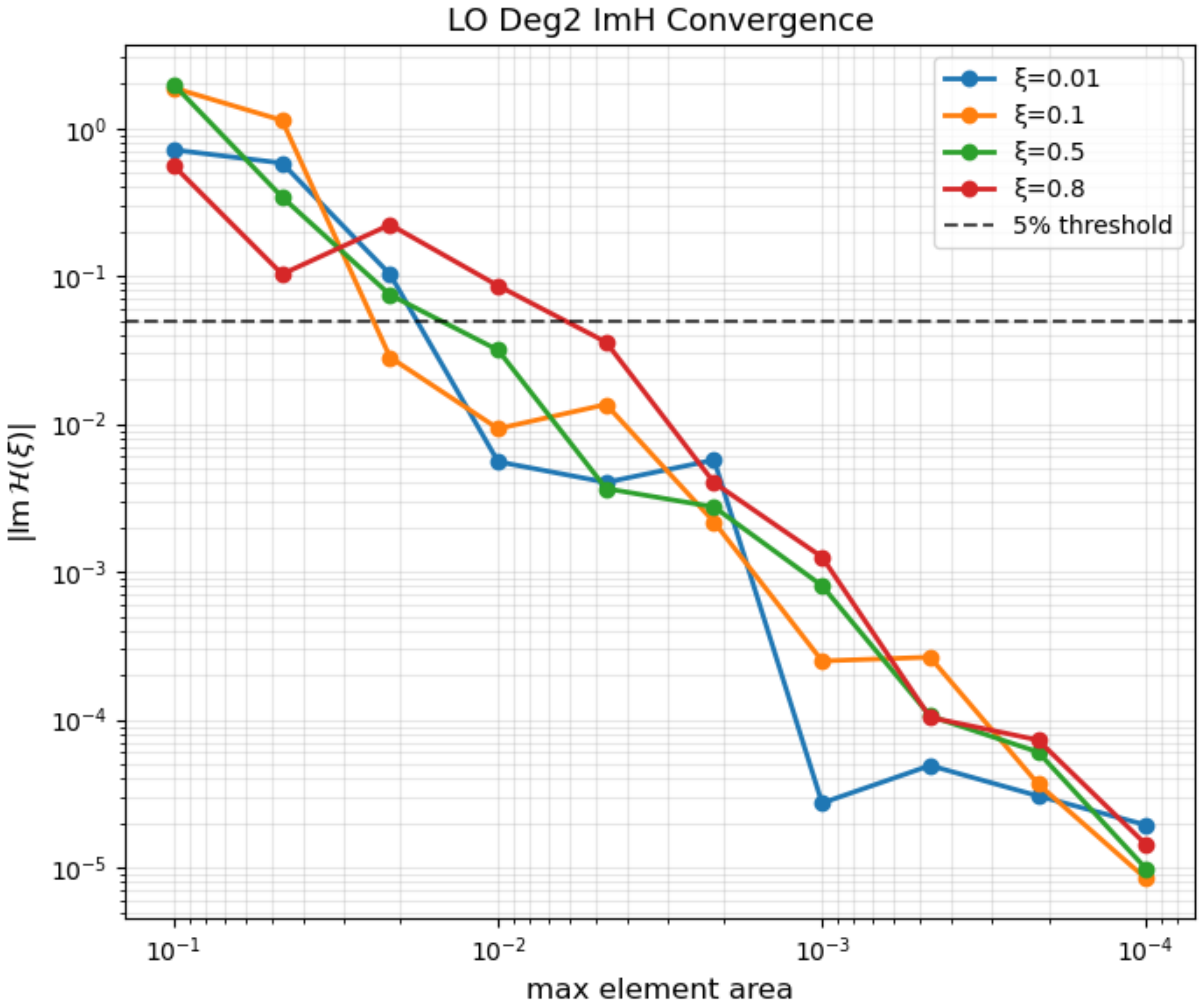
LO shadow DD, degree 1 polynomial interpolation





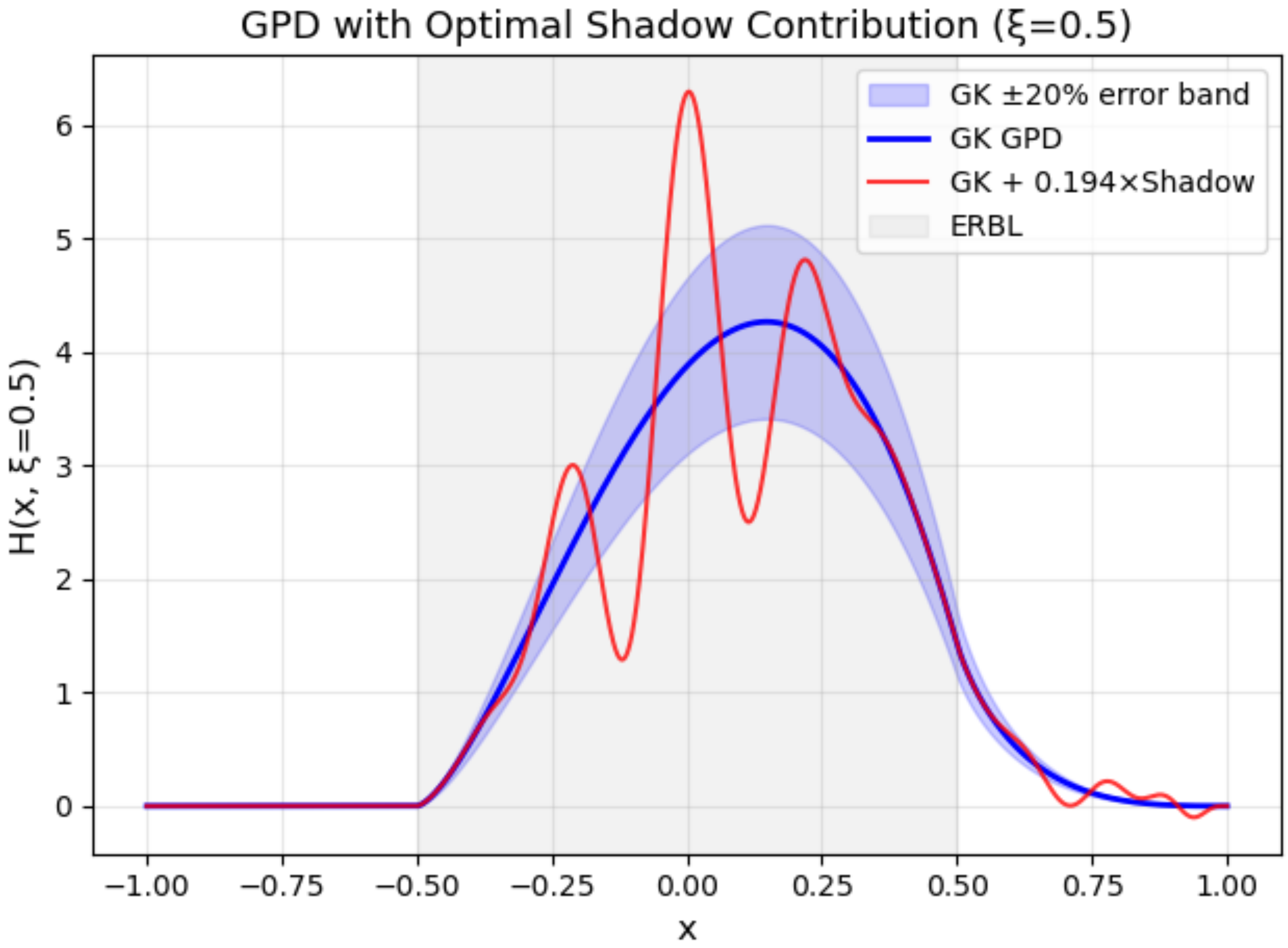
LO shadow DD, degree 2 polynomial interpolation



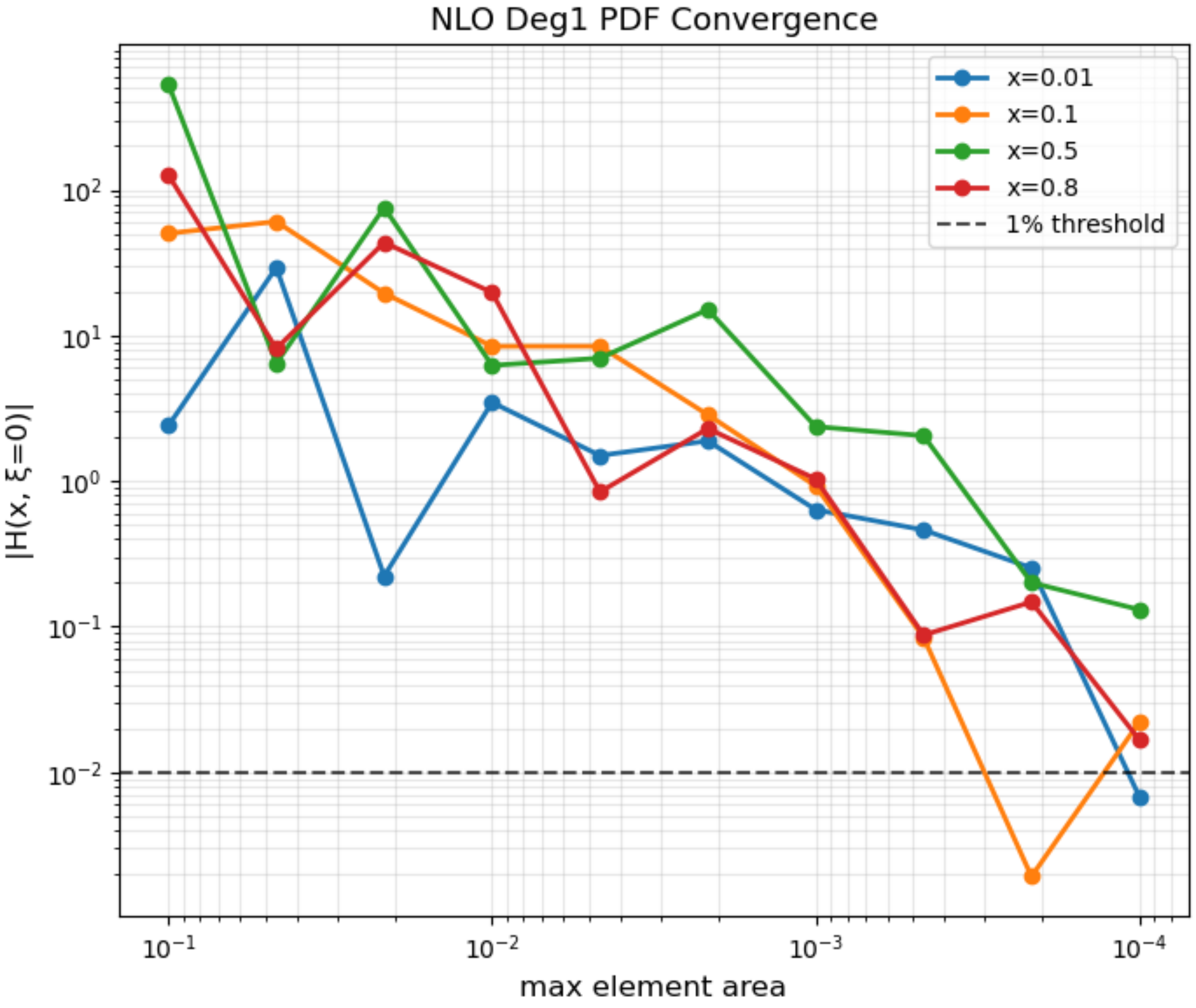


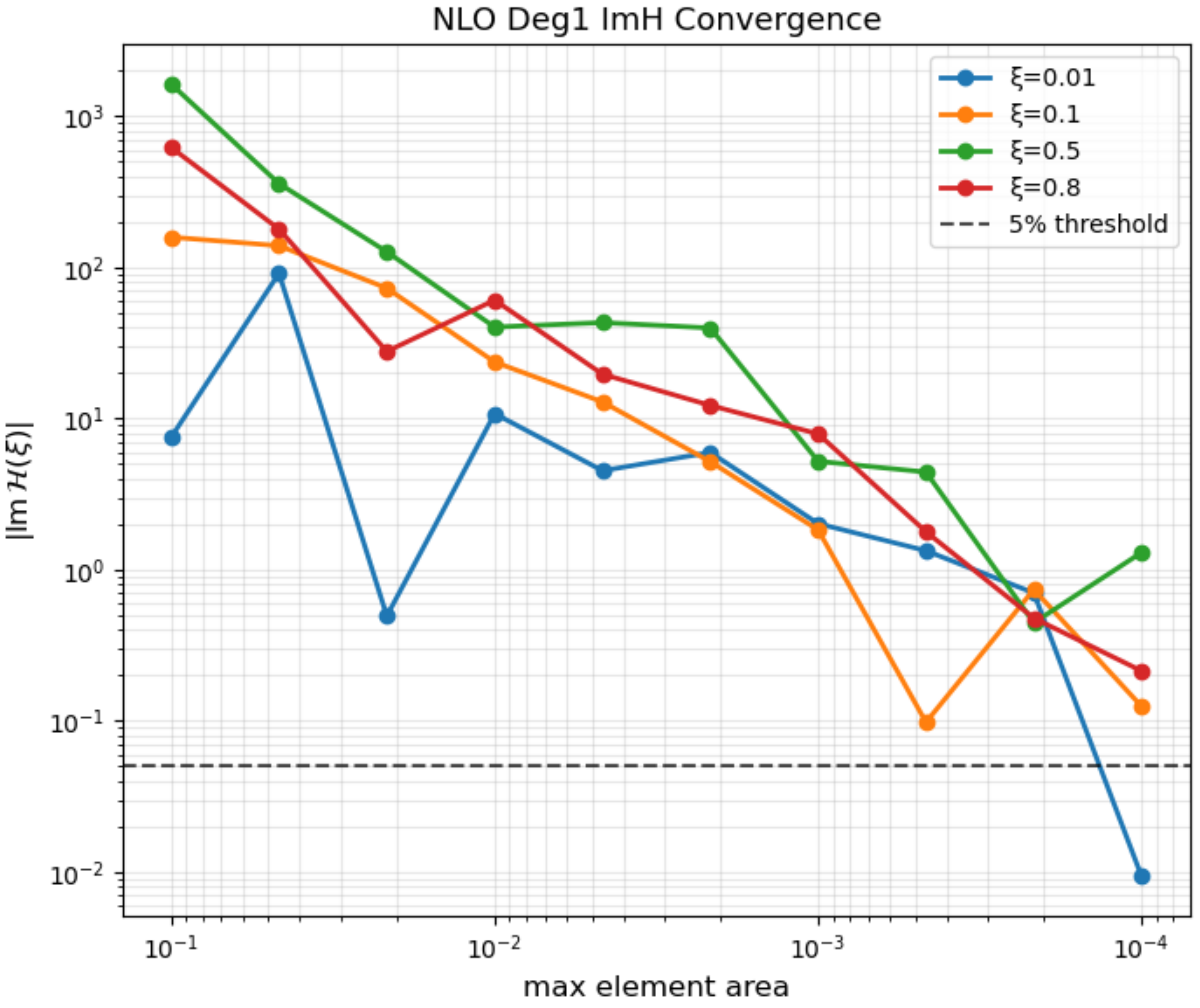
NLO POLYNOMIAL SHADOW DD

Shadow DD defined so that GK+shadow GPD is 67% of the time inside 20% GK error.



NLO shadow DD, degree 1 polynomial interpolation



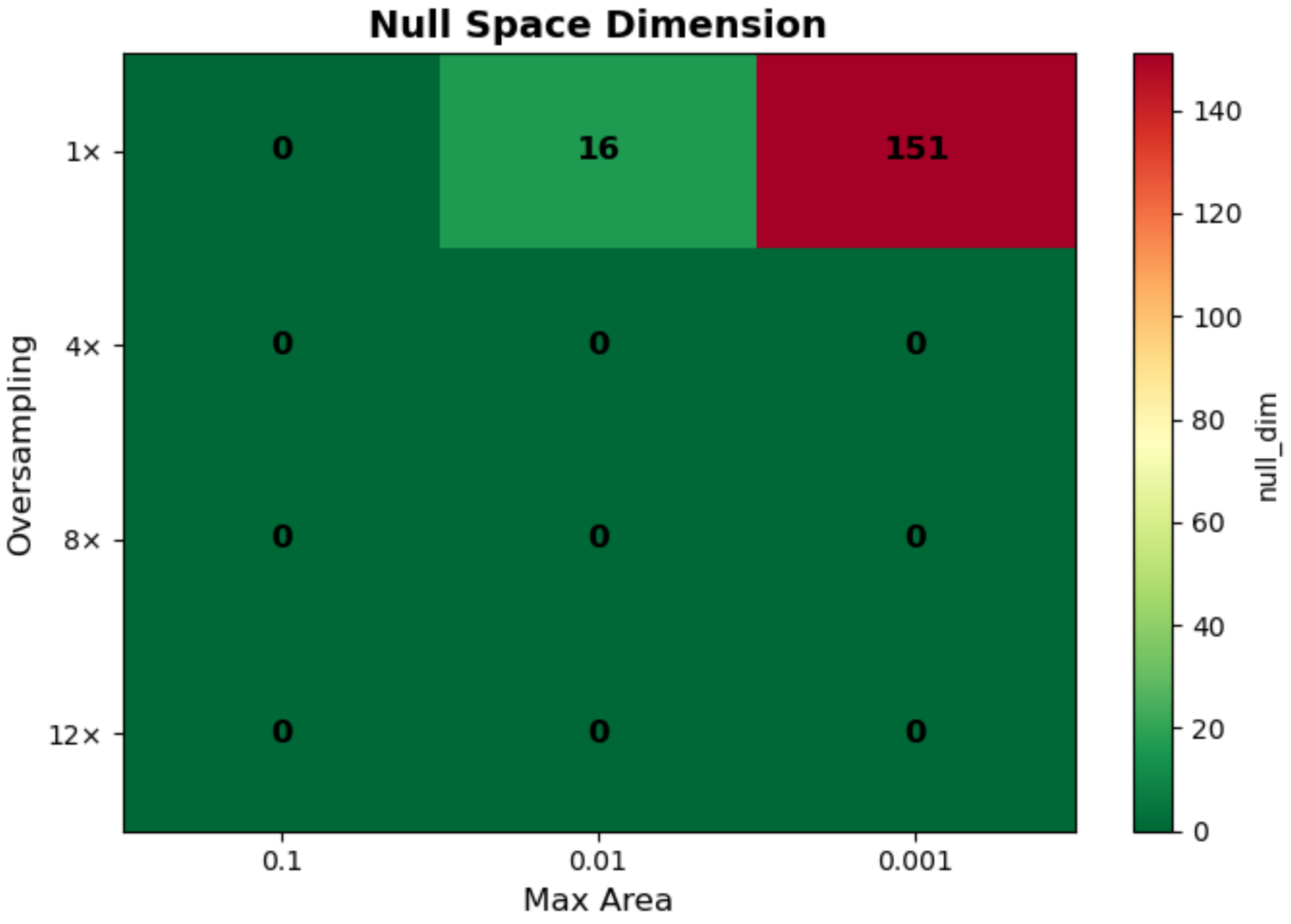


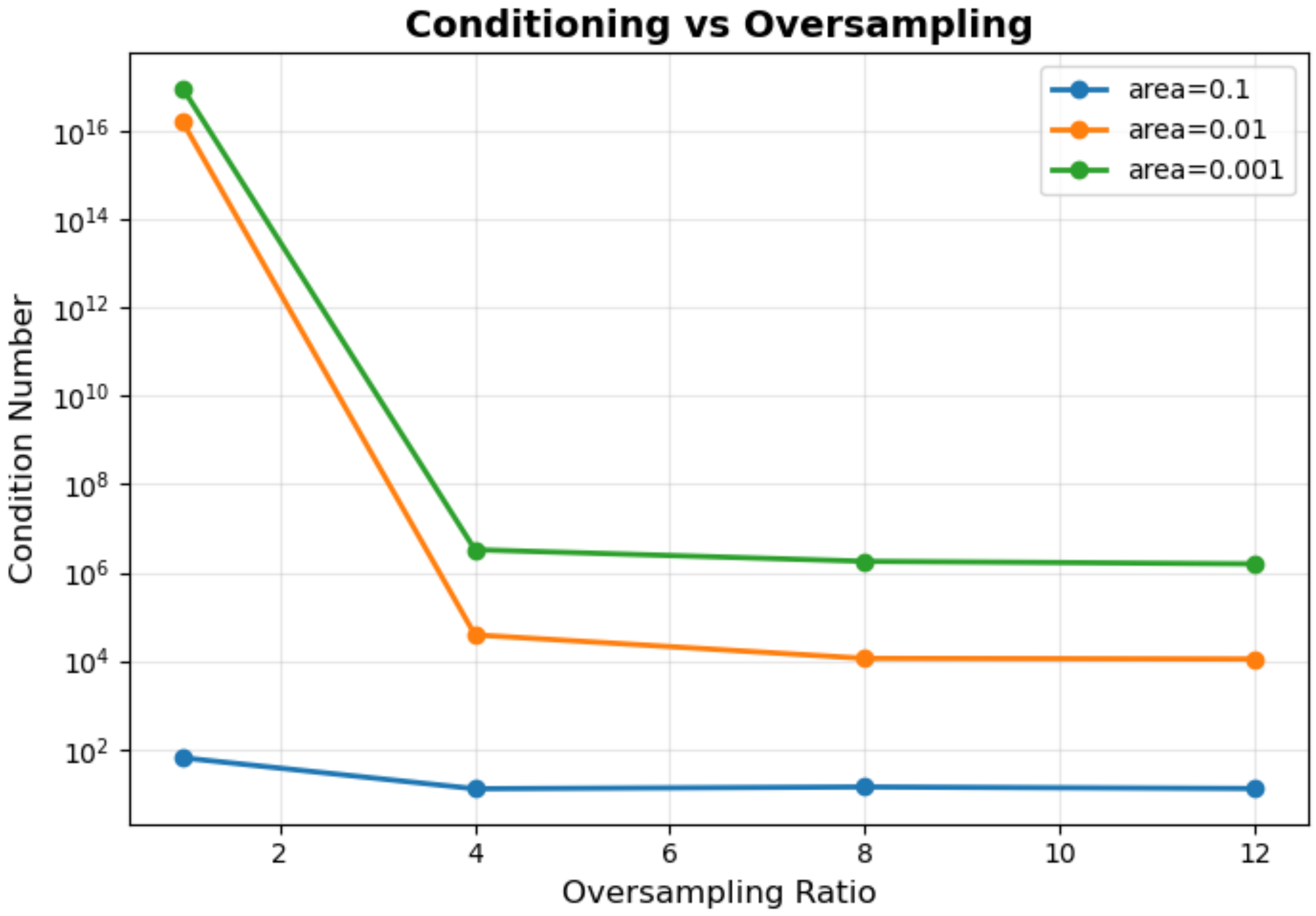
NLO shadow DD, degree 2 polynomial interpolation

The FEM allows us to see shadow contributions and understand how they behave with the precision of the FEM, i.e. it quantifies the relationship between experimental precision in terms of x and ξ and the theoretical precision of accessing the GPD without shadow contaminations!

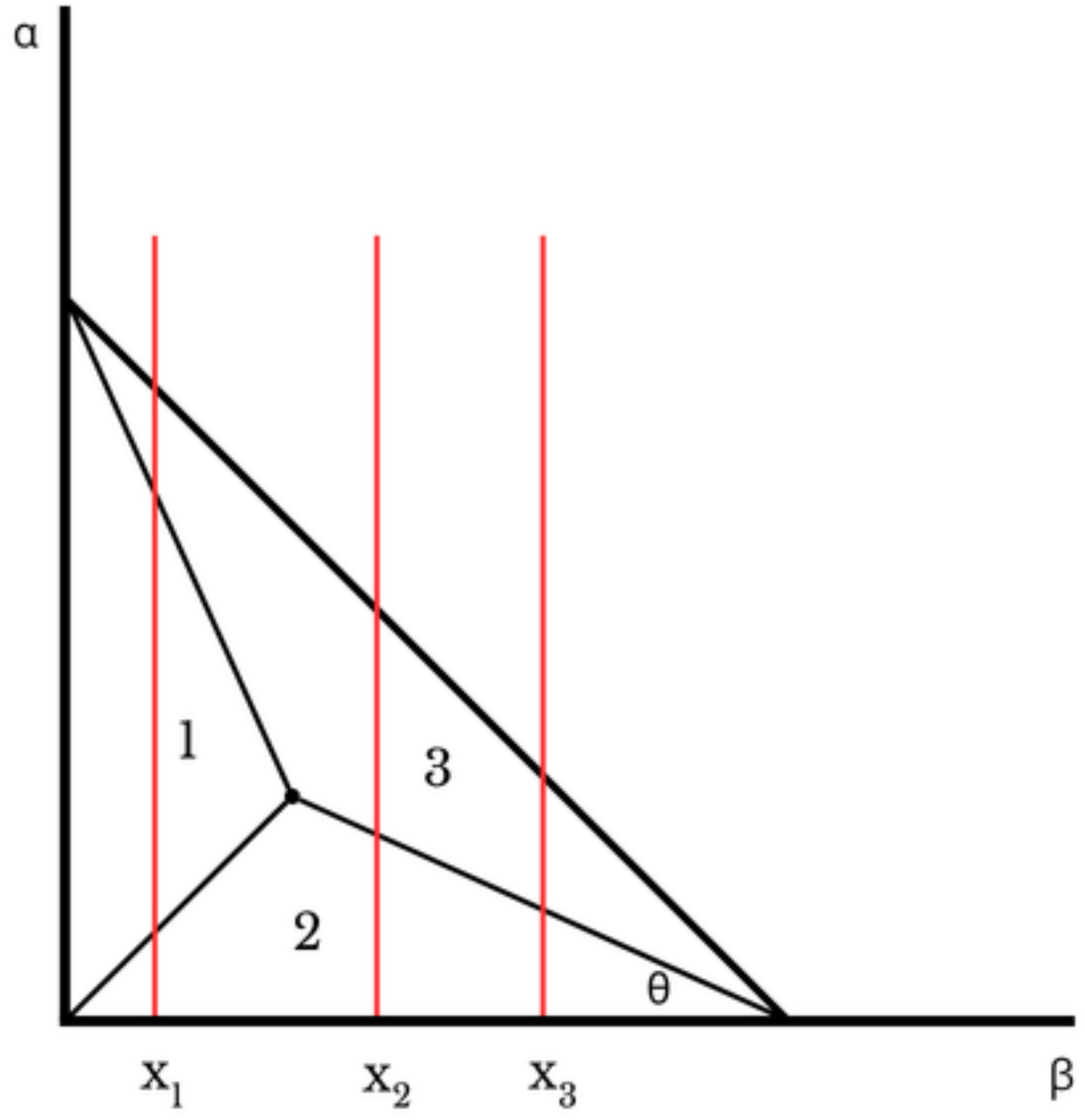
DO WE HAVE DISCRETE SHADOW DDS?

$$\left.\begin{aligned} H_S(x,0) &= Rh = 0 \rightarrow h = 0,\ \forall x \\ H_S(\xi,\xi) &= Rh = 0 \rightarrow h = 0,\ \forall \xi \end{aligned}\right\} \quad R \textbf{ has to be invertible}$$

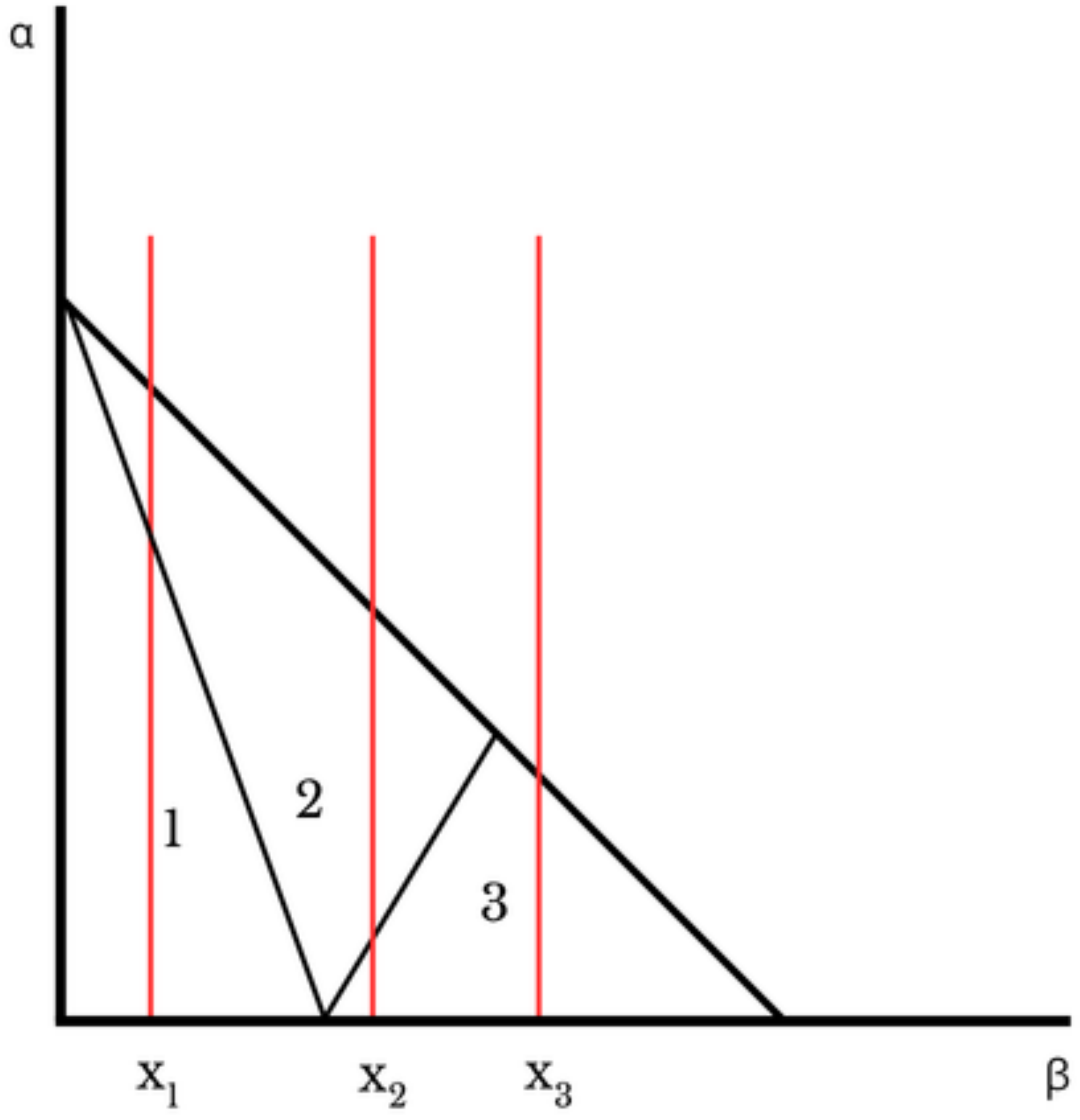




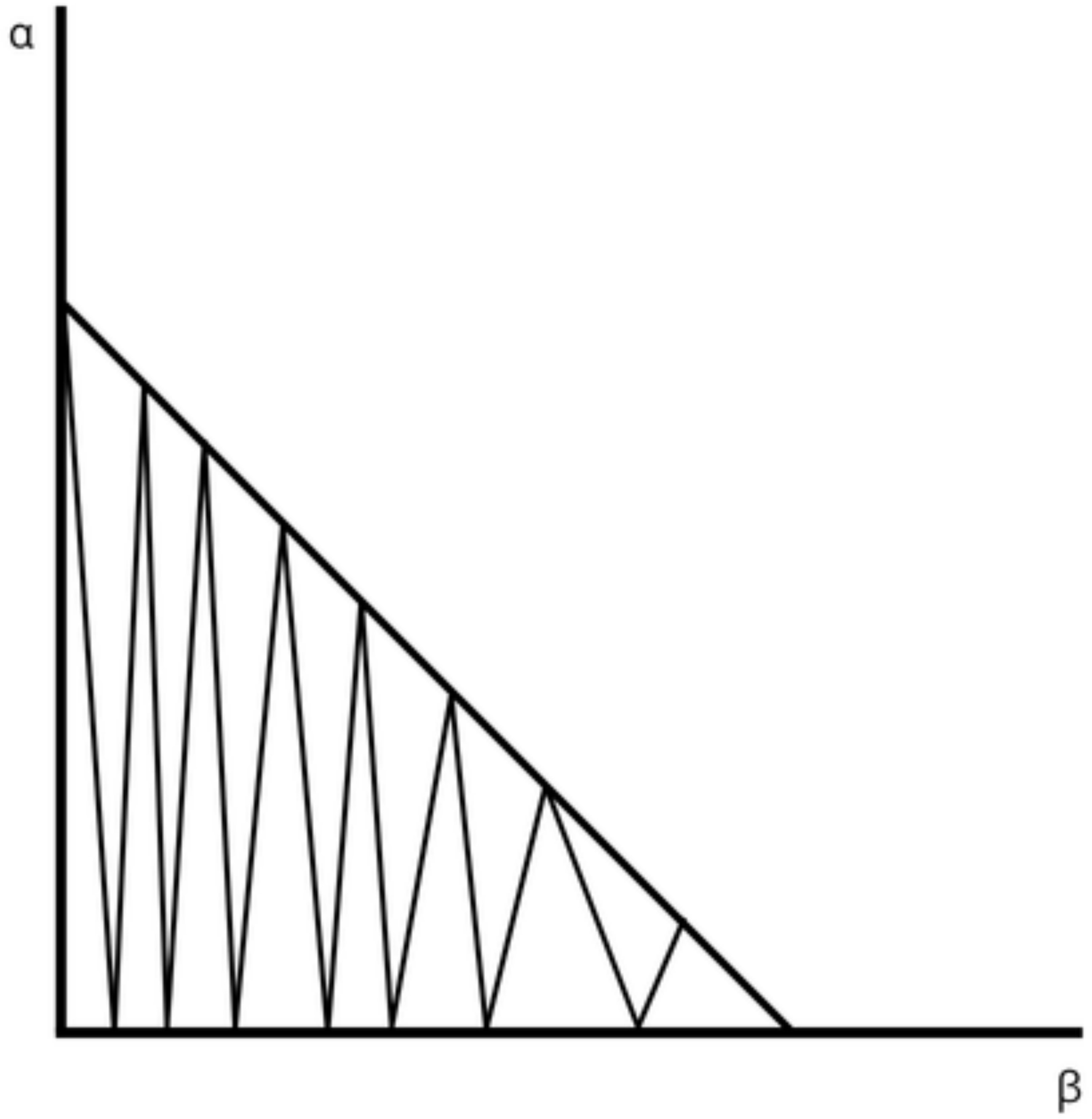
GEOMETRY OF PDF SAMPLING



$$\det R = \begin{vmatrix} l_1^1 & l_1^2 & l_1^3 \\ 0 & \tan\theta(1 - x_2) & \frac{1}{\sqrt{2}}(1 - x_2) \\ 0 & \tan\theta(1 - x_3) & \frac{1}{\sqrt{2}}(1 - x_3) \end{vmatrix} = 0$$



$$\det R = \begin{vmatrix} l_1^1 & l_1^2 & 0 \\ 0 & l_2^2 & l_2^3 \\ 0 & 0 & l_3^3 \end{vmatrix} \neq 0$$



Sliver triangles, ill-conditioned, erasure of α information

Proof of no discrete DDs highly depends on the geometry and ordering of elements and sampling lines, random geometries do not reach full rank!

NEURAL FIELDS REPRESENTATION

NEURAL FIELD FOR THE DD

A **NEURAL FIELD** IS A NEURAL NETWORK THAT REPRESENTS A CONTINUOUS PHYSICAL FIELD AS A FUNCTION OF COORDINATES. RATHER THAN STORING DATA ON A DISCRETE GRID, THE NETWORK WEIGHTS ENCODE THE FIELD IMPLICITLY — IT CAN BE EVALUATED AT ANY POINT IN THE DOMAIN. HERE, THE DOUBLE DISTRIBUTION $F(\beta,\alpha)$ IS REPRESENTED AS A NEURAL FIELD: THE NETWORK TAKES (β,α) COORDINATES AS INPUT AND RETURNS THE FIELD VALUE THERE.

Goloskokov-Kroll DD:

$$F(\beta,\alpha) = 2\,N_u\,\beta^{-\alpha_0}(1-\beta)^{b_u}\frac{\Gamma(2b+2)}{2^{2b+1}\,\Gamma(b+1)^2}\cdot\frac{\left[(1-\beta)^2-\alpha^2\right]^b}{(1-\beta)^{2b+1}} \rightarrow F(\beta,\alpha) = \exp\Big(\underbrace{\mathbf{MLP}}_{\text{deals with the singularity}}\big(\log_{\text{norm}}\beta, \overset{\text{symmetry}}{|\alpha|}\big)\Big)\cdot\big(\underbrace{1-|\beta|-|\alpha|}_{\text{boundary vanishing}}\big)^n\cdot\overbrace{\mathbf{1}[\beta>0]}^{\text{valence support}}$$

FOURIER FEATURES REPLACE THE RAW INPUT COORDINATES WITH A HIGH-DIMENSIONAL ENCODING $[\sin(2\pi Bx), \cos(2\pi Bx)]$, WHERE B IS A RANDOM MATRIX WITH ENTRIES DRAWN FROM $\mathcal{N}(0,\sigma^2)$. THIS OVERCOMES THE SPECTRAL BIAS OF STANDARD MLPS — THE TENDENCY TO LEARN LOW-FREQUENCY FUNCTIONS FIRST — ALLOWING THE NETWORK TO REPRESENT SHARP, PEAKED FUNCTIONS LIKE THE GK DOUBLE DISTRIBUTION. HERE WE USE 128 RANDOM FFS. [ARXIV:2006.10739]

The architecture is physics informed through the loss function:

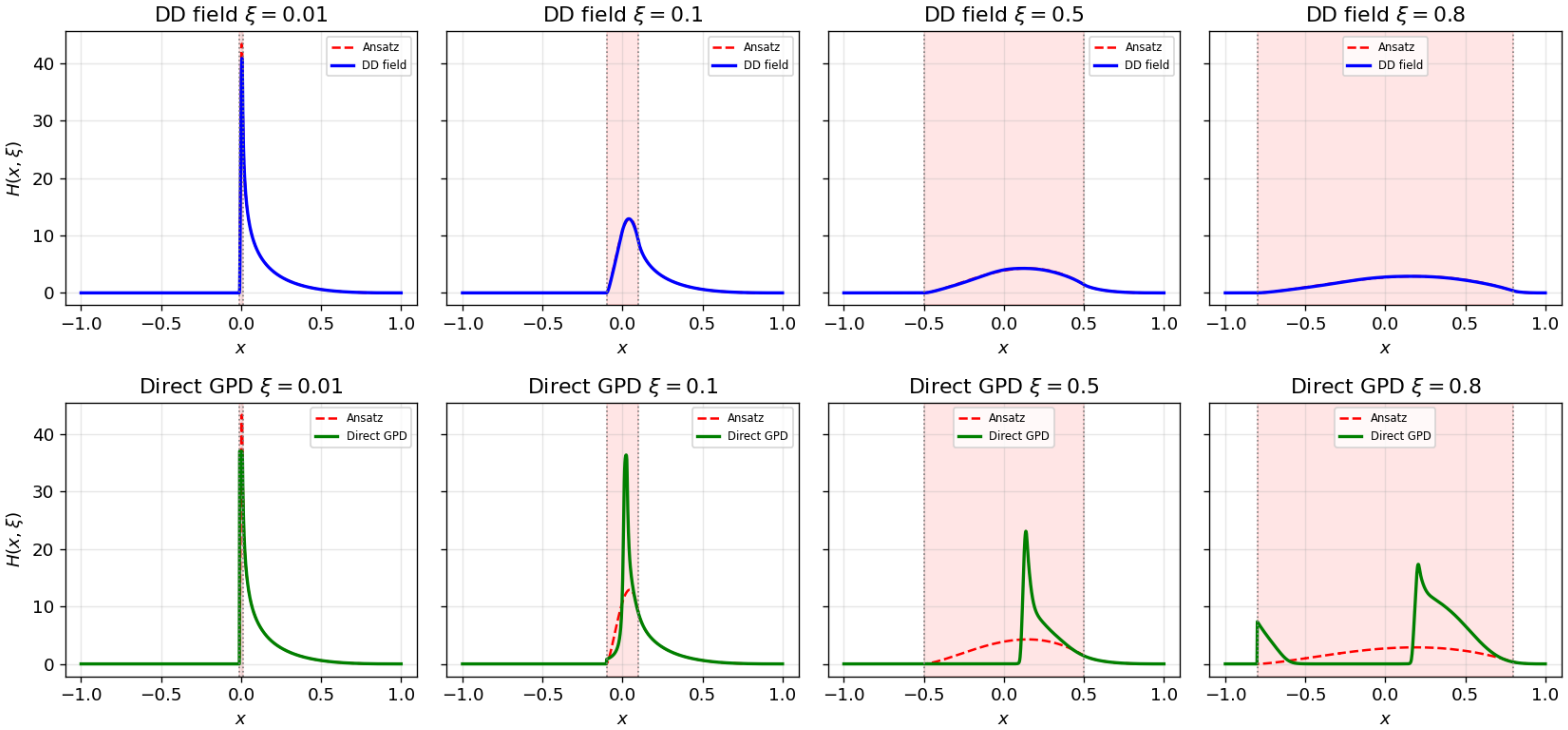
$$\mathcal{L} = \mathcal{L}_{\text{DGLAP}} + \lambda_{\text{PDF}}\,\mathcal{L}_{\text{PDF}} + \lambda_{\text{FF}}\,\mathcal{L}_{\text{FF}}$$

$$\mathcal{L}_{\text{DGLAP}} = \frac{1}{N_{\text{d}}}\sum_i\left(\frac{H(x_i,\xi_i)-H_i^{\text{data}}}{\max(|H_i^{\text{data}}|, 0.01)}\right)^2, \quad \mathcal{L}_{\text{PDF}} = \frac{1}{N_{\text{p}}}\sum_j\left(\frac{H(x_j,\xi\to 0)-q(x_j)}{\max(|q(x_j)|, 0.01)}\right)^2, \quad \mathcal{L}_{\text{FF}} = \text{Var}_\xi\left[F_1(\xi)\right] + 0.1\left(\langle F_1(\xi)\rangle_\xi - F_1^{\text{target}}\right)^2$$

ARCHITECTURE

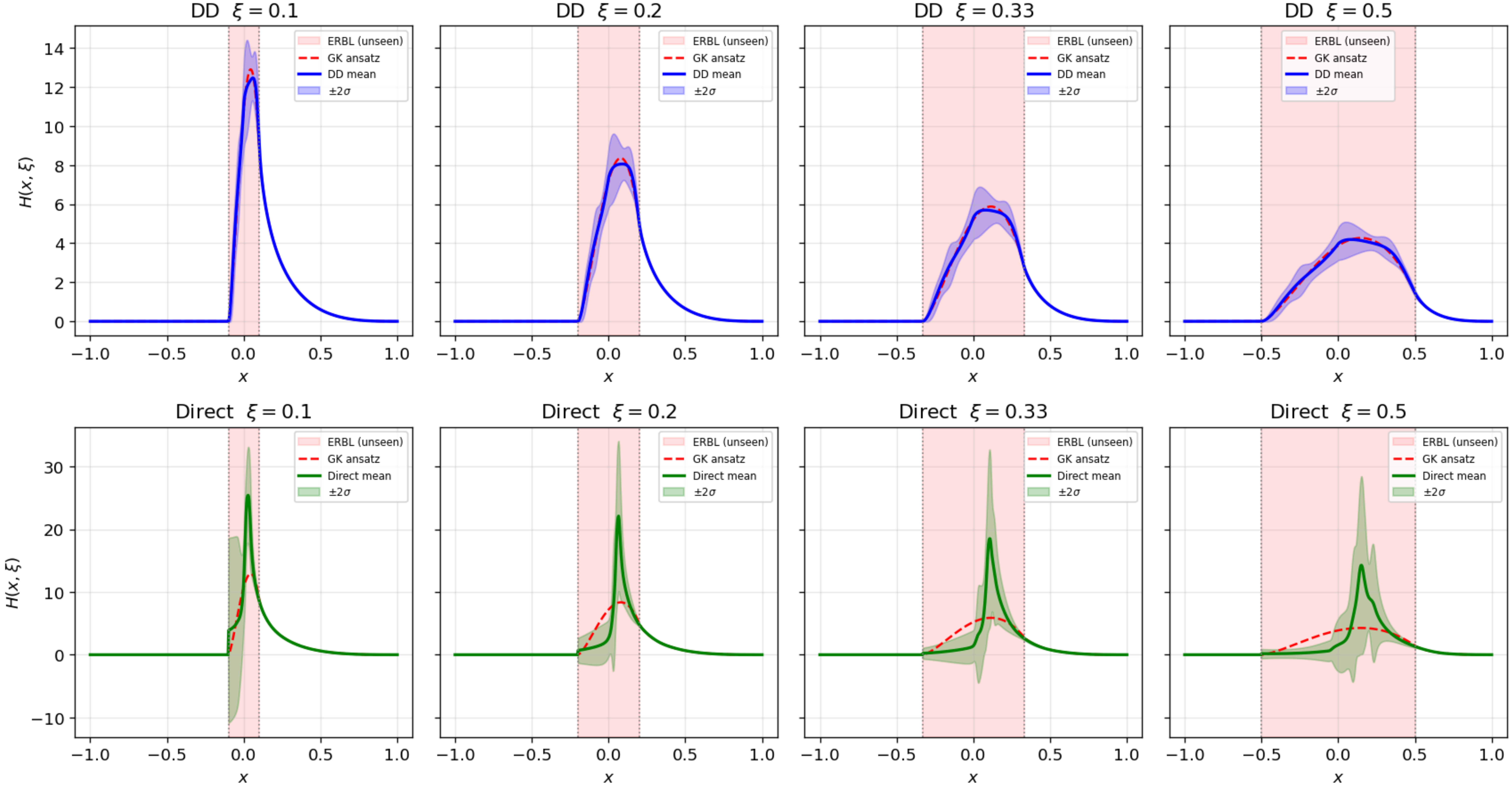
COMPARISON OF TRAINING WITH AND WITHOUT DDS

1389 DGLAP training points $\xi \in \langle 10^{-4}, 0.8 \rangle$, 500 PDF labels, 16 FF points, Adam with learning rate = 3×10^{-4}, cosine annealing with warm restarts $(T_0 = 500,\ T_{mult} = 2)$, 6 000 epochs, batch size 512, no uncertainties, $\lambda_{PDF} = 0.5,\ \lambda_{FF} = 0.1$



SIMULATING DATA COVERAGE AND UNCERTAINTIES

Toy model: 10 ensembles for systematics and statistics, trained on MC replicas of data with 5% noise, simulated experimental coverage (300 training datapoints, clustered in various kinematic regions, 40 PDF labels)



CONCLUSIONS

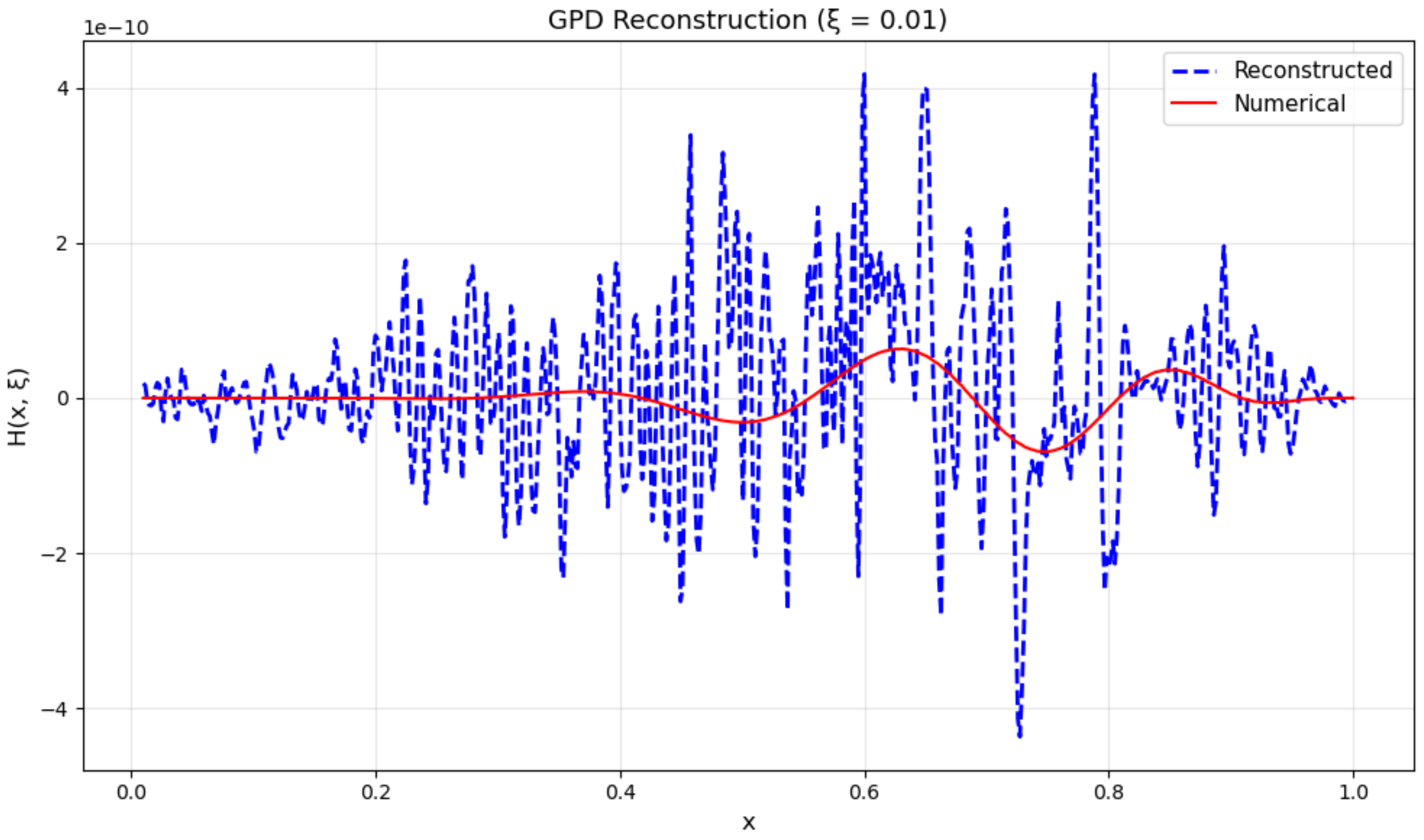
- Double distributions obtained from DGLAP GPDs contain the ERBL region
- FEM precision in terms of x and ξ connected to theoretical precision in terms of shadow contamination
- Degree-2 polynomial interpolation sufficient for reasonable precision for LO and NLO studies
- Exclusion of discrete DDs depends on the geometry of the mesh
- Ill-posed inverse problem: does the data allow for reliable extraction of GPDs and TMDs?
- How to quantify uncertainties in areas with no experimental coverage?
- What do uncertainty bands mean?

SUPPLEMENTAL SLIDES

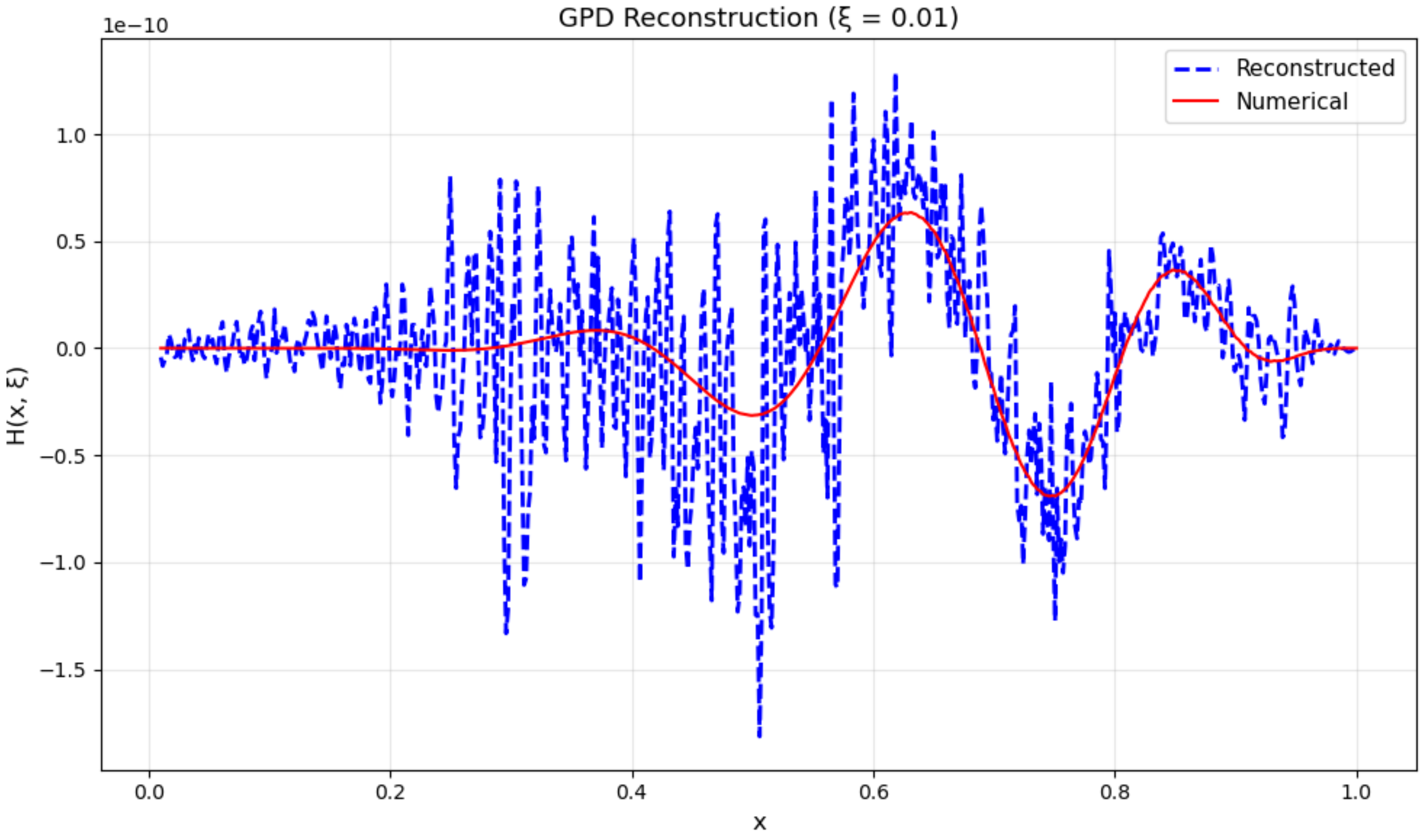
DGALP PRECISION AT LOW ξ

NLO shadow GPD

Max area = 10^{-4}



Max area = 5×10^{-5}



Double distribution model