A Pixel-Based Bayesian Approach for Model Independent TMD Reconstruction

Marco Zaccheddu- Jefferson Lab

QCD Evolution 2026

The Inverse Problem

The Objective
Reconstruct TMDs in bT-space from data in kT-space

$\widetilde{W}(b_T)$ is a convolution of two TMDs

$$W(k_T) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}b_T\, b_T\, J_0(b_T\, k_T)\, \widetilde{W}(b_T)$$

Standard Approaches

$$\widetilde{\boldsymbol{W}} \equiv \widetilde{\boldsymbol{W}}(\boldsymbol{\theta})$$

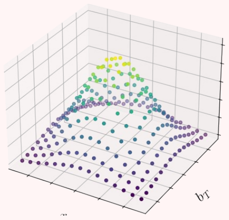
- Parametrization
- Neural Networks

$\boldsymbol{\theta}$ **Find best-fit (hyper)parameters**

Replica Method for Uncertainty Quantification

New Pixel-based Approach

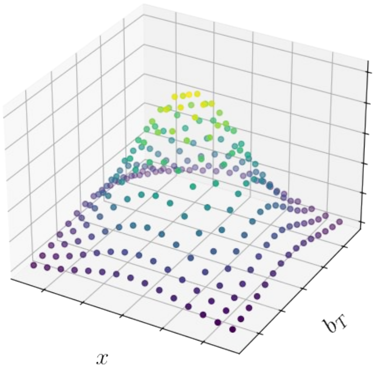
$\widetilde{W} \equiv \widetilde{W}[b_{T1}, \ldots, b_{TN}]$ Discretization or "Pixelization"



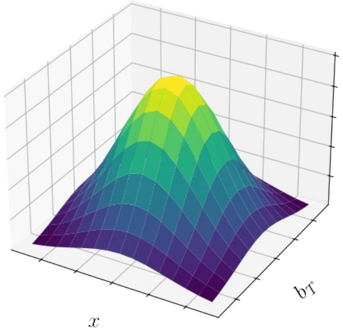
Advanced Methodology:

- Model Independent
- Generative AI for Pixel Distribution
- MH & Bayes Theorem for Uncertainty Quantification

The Inverse Problem: Pixel Based Approach



Function recovered via interpolation



Three Closure Tests

- Gaussian Distribution
- TMD PDF
- Fuu Structure Function

Outline

- Methodology
- Normalizing Flow + MH
- Solution Non-uniqueness
- Uncertainty Quantification
- Null TMDs

Based on: ArXiv 2605.06606

Gaussian Case: Data generation

Events generation: Inverse Transform Sampling

$$k_T = \sqrt{\sigma^2(\mu^2)} \cdot \text{erfinv}(u), \quad \sigma^2(\mu^2) = \alpha + g_{\text{evo}} \ln\left(\frac{\mu^2}{\mu_0^2}\right).$$

$$g_{evo} = 1.2\,\text{GeV}^2$$

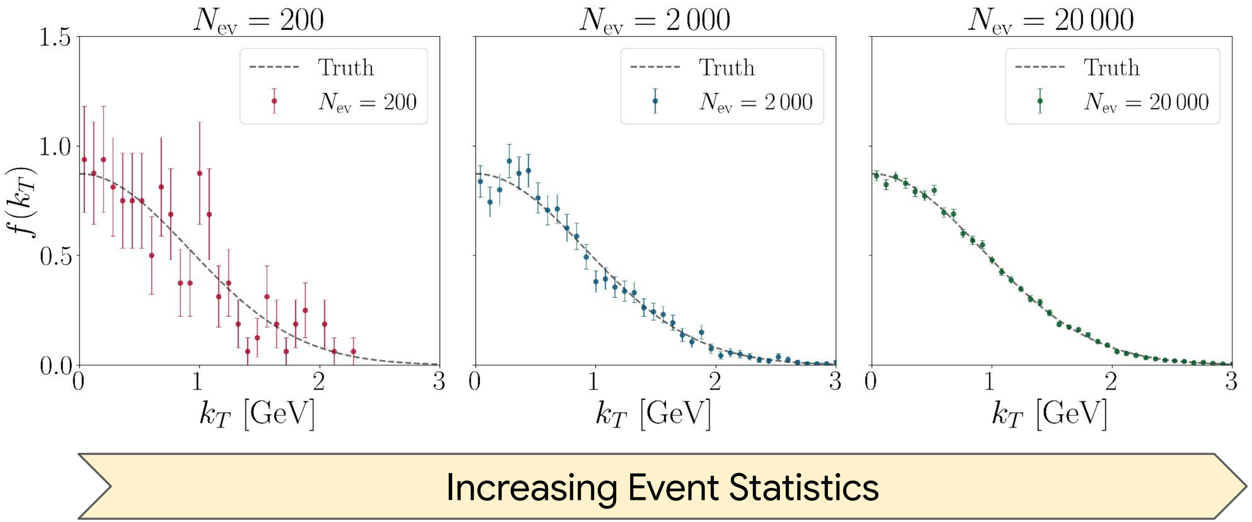
True Distributions

kT-space

$$f(k_T, \mu^2) = \frac{2}{\sqrt{\pi\sigma^2(\mu^2)}} \exp\left(-\frac{k_T^2}{\sigma^2(\mu^2)}\right),$$

bT-space

$$\tilde{f}(b_T, \mu^2) = 2\sqrt{\pi\sigma^2(\mu^2)} \exp\left(-\frac{b_T^2\,\sigma^2(\mu^2)}{4}\right).$$



- 3 scenarios for Number of events
- Fixed grid for binning
- Data generated at $\mu^2 = 2\,\text{GeV}^2$
- Errors shrink as Nev increaseas

$$\sigma \propto 1/\sqrt{N_{ev}}$$

Gaussian Case: Discretization and Bayes Theorem

$$f(k_{Tj}) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}b_T\, b_T J_0(b_T k_{Tj})\, \tilde{f}(b_T),$$

$$\tilde{f}(b_T) \approx \sum_{i=1}^{N} \tilde{f}_i\, h_i(b_T)\,,$$

Interpolation with Local Interpolators

$$[\mathcal{M}_{\mathrm{ref}}]_{ki} = \frac{1}{2\pi}\int_{b_{T\,\mathrm{min}}}^{b_{T\,\mathrm{max}}} \mathrm{d}b_T\, b_T\, J_0(b_T k^{\mathrm{ref}}_{T,k})\, h_i(b_T).$$

Transform as Matrix Multiplication

$$\boldsymbol{f} = \mathcal{M}\,\tilde{\boldsymbol{f}}$$

Forward Map

Many methods:

- Ogata
- Trapezoid Rule
- Gaussian Quadrature
- ...

$$\tilde{f}_i \equiv \tilde{f}(b_{Ti})$$

Interpolation matrices depending on data points

$$\mathcal{M}_{\mathrm{total}} = \mathcal{M}_{\mathrm{interp}}\,\mathcal{M}_{\mathrm{ref}}\,.$$

$$\tilde{\boldsymbol{f}} = \left[\tilde{f}(b_{T\,1}), \ldots, \tilde{f}(b_{T\,N})\right]^T$$ Pixel Vector

$$\boldsymbol{f} = \left[f(k_{T\,1}), \ldots, f(k_{T\,M})\right]^T$$ Data Vector

It cannot be directly inverted!
$\mathcal{M}$ is rank deficient
The solution is NOT Unique!

Bayesian Pixel-based approach + Gen AI

$$\tilde{\boldsymbol{f}} = \left[\tilde{f}(b_{T\,1}), \ldots, \tilde{f}(b_{T\,N})\right]^T$$ Pixel Vector

$\tilde{\boldsymbol{f}}$ Pixels are Random Variables with a Probabilty Distribution

Posterior and Bayes Theorem

$$P(\tilde{\boldsymbol{f}}|\boldsymbol{f}) \propto \mathcal{L}(\boldsymbol{f}|\tilde{\boldsymbol{f}})\, \pi(\tilde{\boldsymbol{f}}),$$

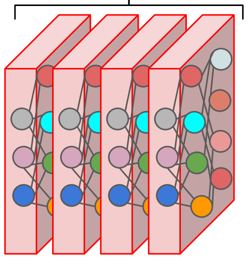
Likelihood: Chi-square

$$\chi^2(\boldsymbol{f}|\tilde{\boldsymbol{f}}) = \sum_{j=1}^{M} \left(\frac{\left[\mathcal{M}\tilde{\boldsymbol{f}}\right]_j - f_j}{\sigma_j} \right)^2 .$$

Prior

$$\ln \pi(\tilde{\boldsymbol{f}}) \propto -\sum_{k=0}^{2} \lambda_k\, \Omega_k(\tilde{\boldsymbol{f}}),$$

Affine Couplings

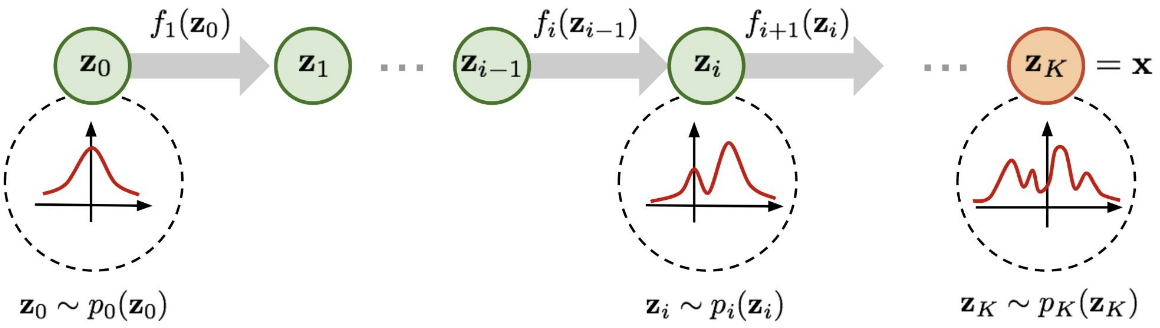


Generative AI for pixels reconstruction: *Normalizing Flow*

No Unique Solution:

The goal is to find all the possible solutions compatible with data and prior

Normalizing Flow



Based on Change of variable theorem

$$\int p(x)dx = \int \pi(z)dz = 1$$

$$p(x) = \pi(z) \left| \frac{dz}{dx} \right| = \pi(f^{-1}(x)) \left| \frac{df^{-1}}{dx} \right|$$

Generation of highly realistic pictures

We can approximate a target distribution through a chain of transformation

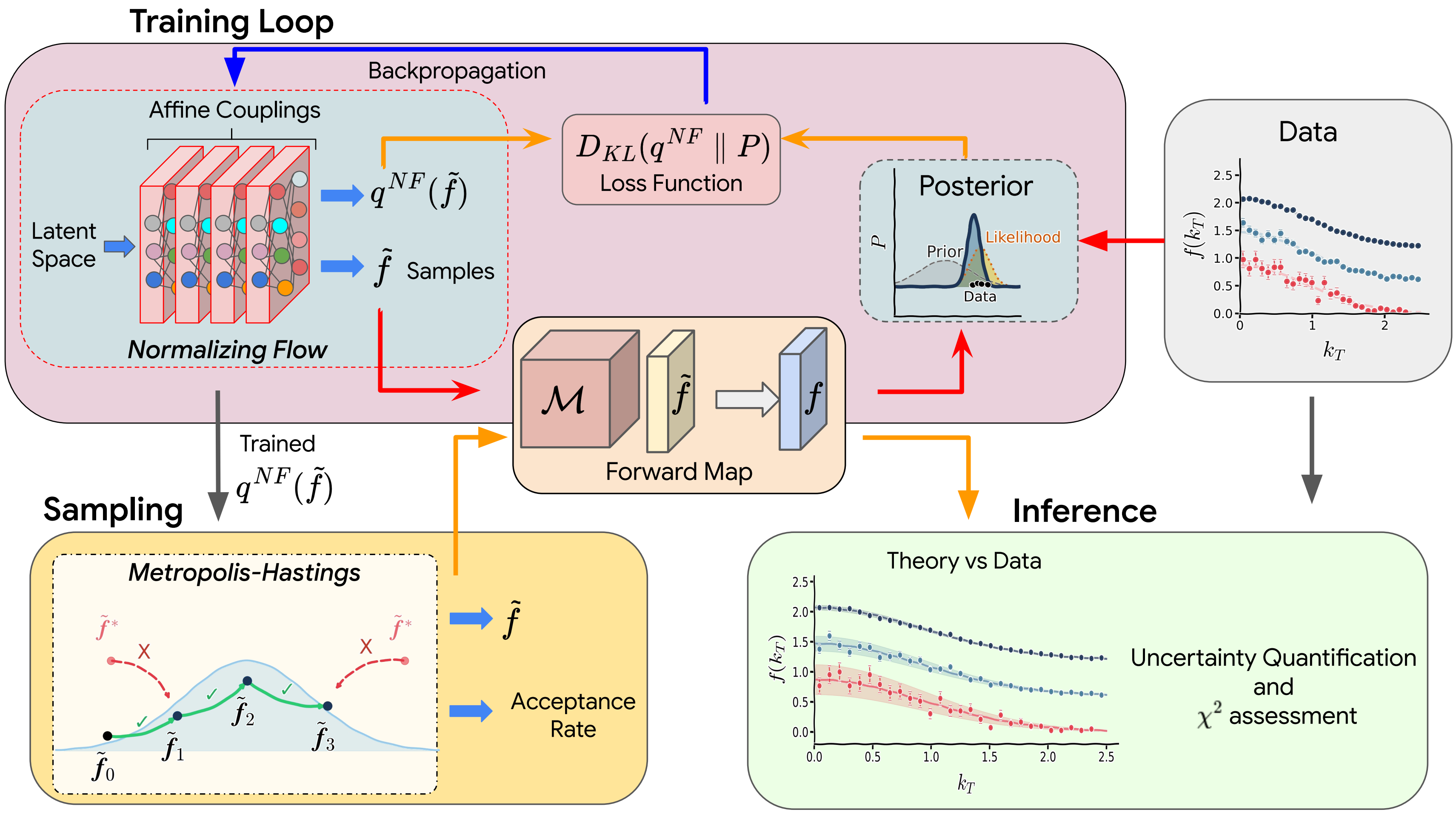
Target Distribution

Initial Distribution

Jacobian: Neural Networks

$$\log p(x) = \log \pi_0(z_0) - \sum_{i=1}^{K} \log \left| \det \frac{df_i}{dz_{i-1}} \right|$$

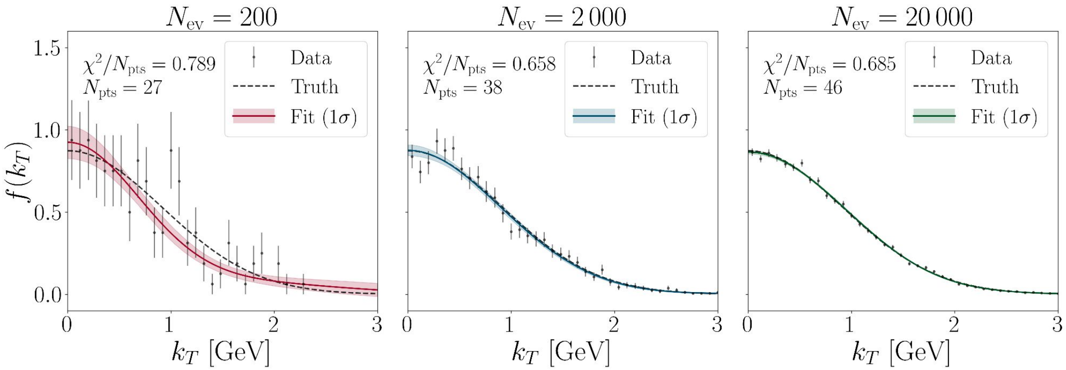
Density estimation using Real NVP Arxiv:1605.08803



Training Loop
Backpropagation
Affine Couplings
Latent Space
$q^{NF}(\tilde{f})$
$\tilde{f}$ Samples
Normalizing Flow
$D_{KL}(q^{NF} \parallel P)$
Loss Function
Posterior
P
Prior
Likelihood
Data
Data
$f(k_T)$
k_T
$\mathcal{M}$
$\tilde{f}$
f
Forward Map
Trained
$q^{NF}(\tilde{f})$
Sampling
Metropolis-Hastings
$\tilde{f}^*$
X
$\tilde{f}_0$
$\tilde{f}_1$
$\tilde{f}_2$
$\tilde{f}_3$
$\tilde{f}$
Acceptance Rate
Inference
Theory vs Data
Uncertainty Quantification and χ^2 assessment

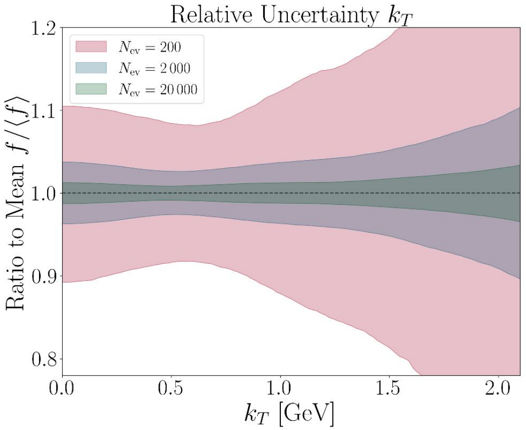
Gaussian Case: Results in kT space

Fit vs Data



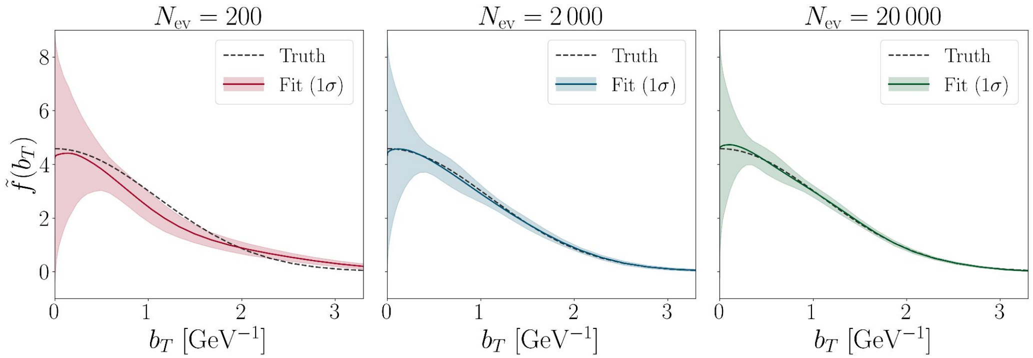
$N_{\rm ev}$	accept. rate	$N_{\rm pts}$	$\chi^2/N_{\rm pts}$
200	61%	27	0.79
2 000	76%	38	0.66
20 000	78%	46	0.69

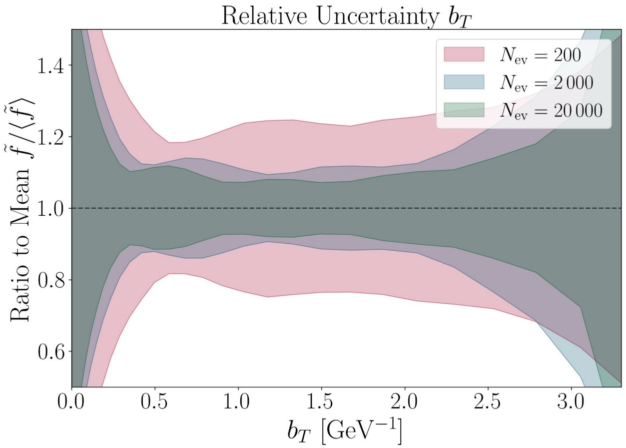
Error bands compatible and around the True model



- 3 Different Fits with same methodology
- Good description of Data
- Error bands scale as $1/\sqrt{N_{ev}}$

Gaussian Case: Results in bT space



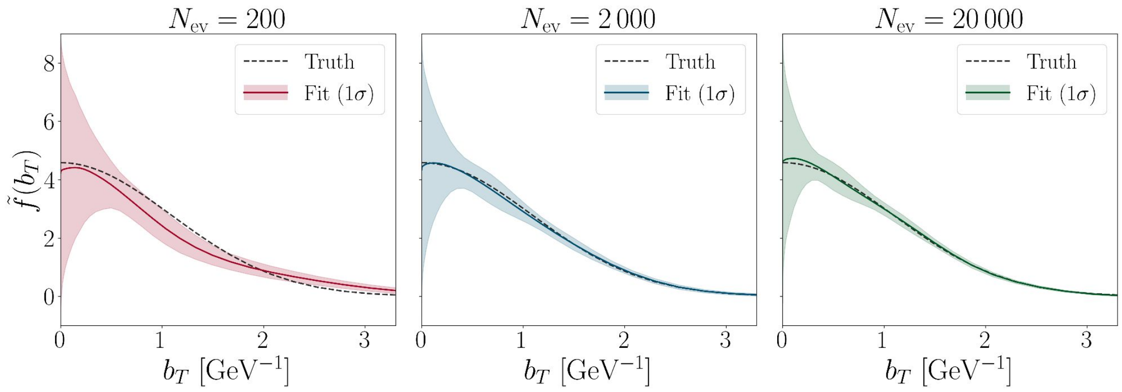


Error bands converge around the True distribution

Non Trivial Criticalities!

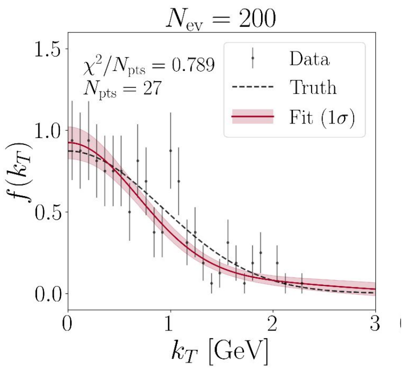
- Error bands DO NOT scale as $1/\sqrt{N_{ev}}$
- “Precision Floor”
- No real shrinking for small bt

Gaussian Case: Results in bT space



Non Trivial Criticalities!

- Error bands DO NOT scale as $1/\sqrt{N_{ev}}$
- “Precision Floor”
- No real shrinking for small bt

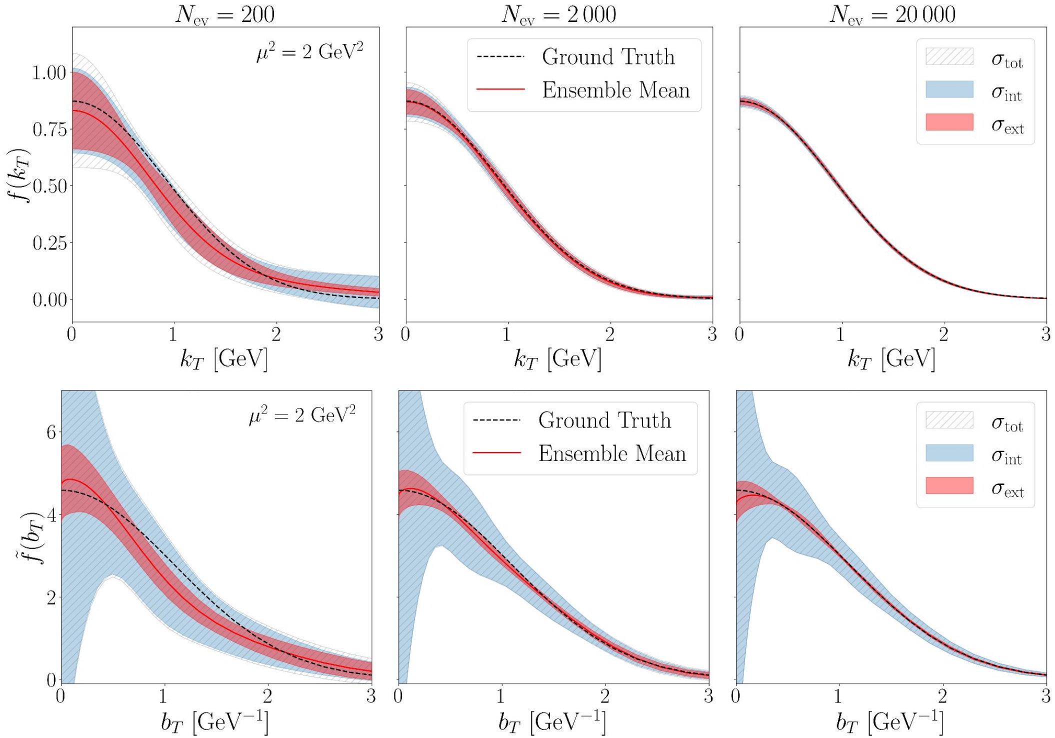


Fit not on top of the Truth

Is this an outcome of the method or a result of the statistics?

Gaussian Case: Ensamble Analysis

Generated 100 independent pseudo-data replicas and performed 100 independent fits.



The uncertainty can be decomposed in:

$$\sigma^2_{\text{tot}}(b_T) = \sigma^2_{\text{int}}(b_T) + \sigma^2_{\text{ext}}(b_T)$$

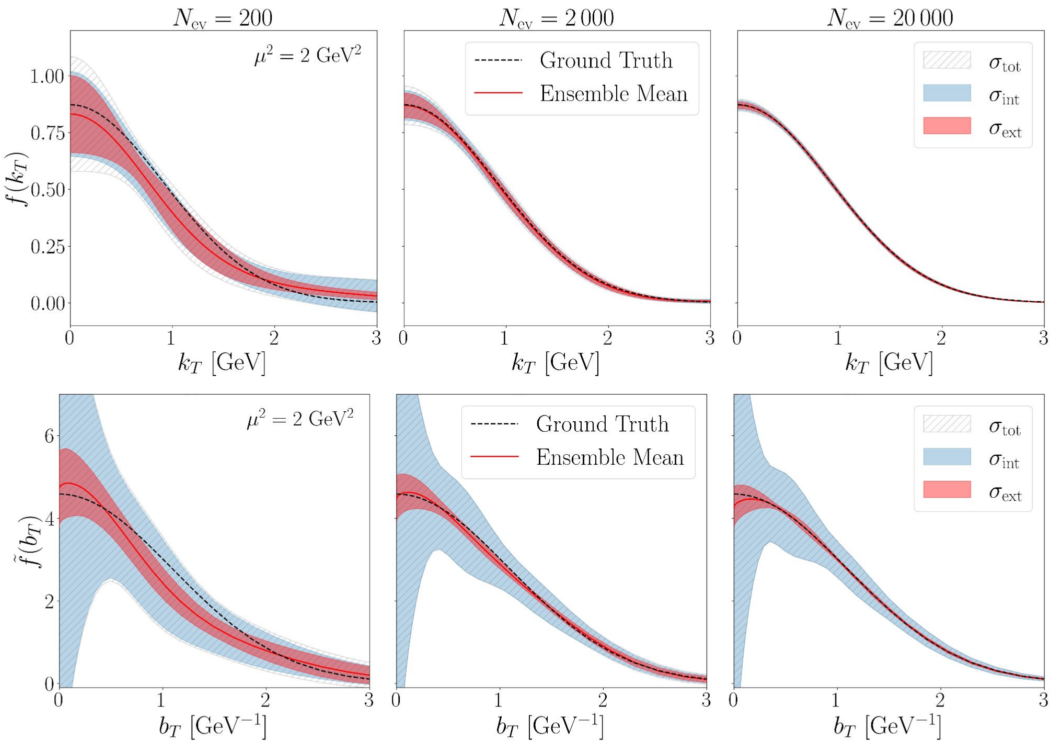
Average of the posterior variances obtained within each individual replica

Variance of the posterior means

- The total uncertainty band consistently encompasses the truth.
- σ_{ext} narrows significantly as the number of events increases.
- At small distances, the uncertainty remains substantial even in the high-statistics scenario.
- Non-uniform scaling and precision floor still present.

Gaussian Case: Ensamble Analysis

Generated 100 independent pseudo-data replicas and performed 100 independent fits.



The uncertainty can be decomposed in:

$$\sigma^2_{tot}(b_T) = \sigma^2_{int}(b_T) + \sigma^2_{ext}(b_T)$$

Average of the posterior variances obtained within each individual replica

Variance of the posterior means

- The total uncertainty band consistently

o.

SVD Analysis

Singular Value Decomposition - SVD

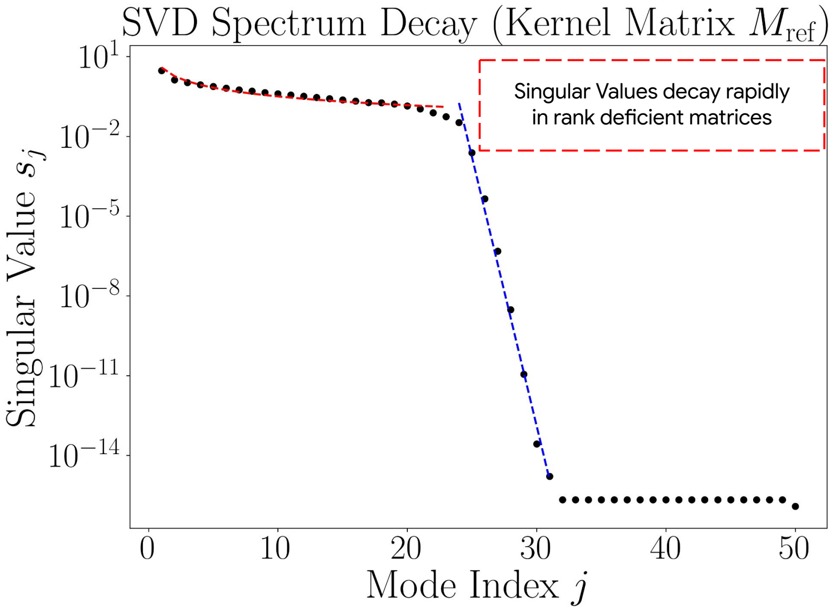
SVD Decomposition

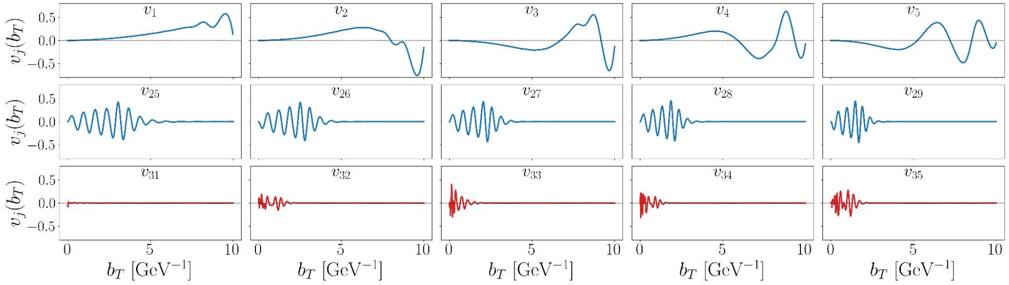
$$\mathcal{M} = U\Sigma V^T$$

U and V are orthonormal matrices

- U (M x M): Span Data Space
- V (N x N) : Span Model Space

Σ is a pseudo-diagonal matrix with Singular Values s_i





$$\mathcal{M}\mathbf{v}_i = s_i\mathbf{u}_i$$

Singular Value Decomposition - SVD

If only the first p singular values are non-zero:

$$V_{\text{obs}} = [\mathbf{v}_1, \ldots, \mathbf{v}_p]$$

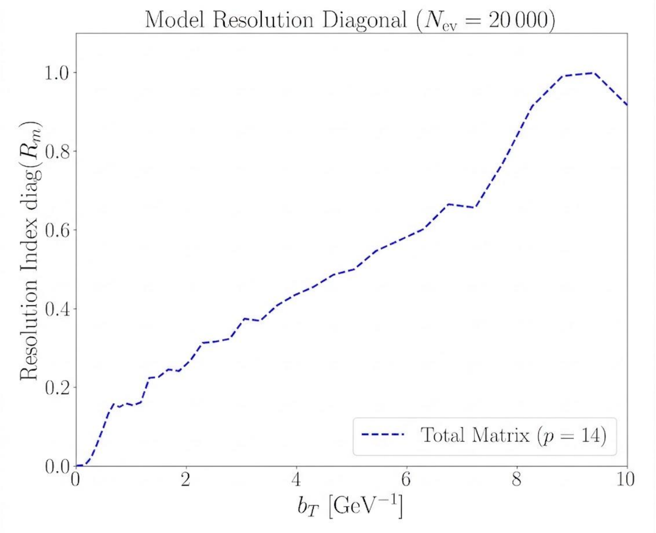
$$V_{\text{null}} = [\mathbf{v}_{p+1}, \ldots, \mathbf{v}_N]$$

$$P_{\text{obs}} = V_{\text{obs}} V_{\text{obs}}^T = \sum_{i=1}^{p} \mathbf{v}_i \mathbf{v}_i^T$$

$$P_{\text{null}} = V_{\text{null}} V_{\text{null}}^T = \sum_{i=p+1}^{N} \mathbf{v}_i \mathbf{v}_i^T$$

We can split the models into Observable and Null Space

$$\tilde{\boldsymbol{f}} = \tilde{\boldsymbol{f}}_{\text{obs}} + \tilde{\boldsymbol{f}}_{\text{null}} = P_{\text{obs}} \tilde{\boldsymbol{f}} + P_{\text{null}} \tilde{\boldsymbol{f}} \, ,$$

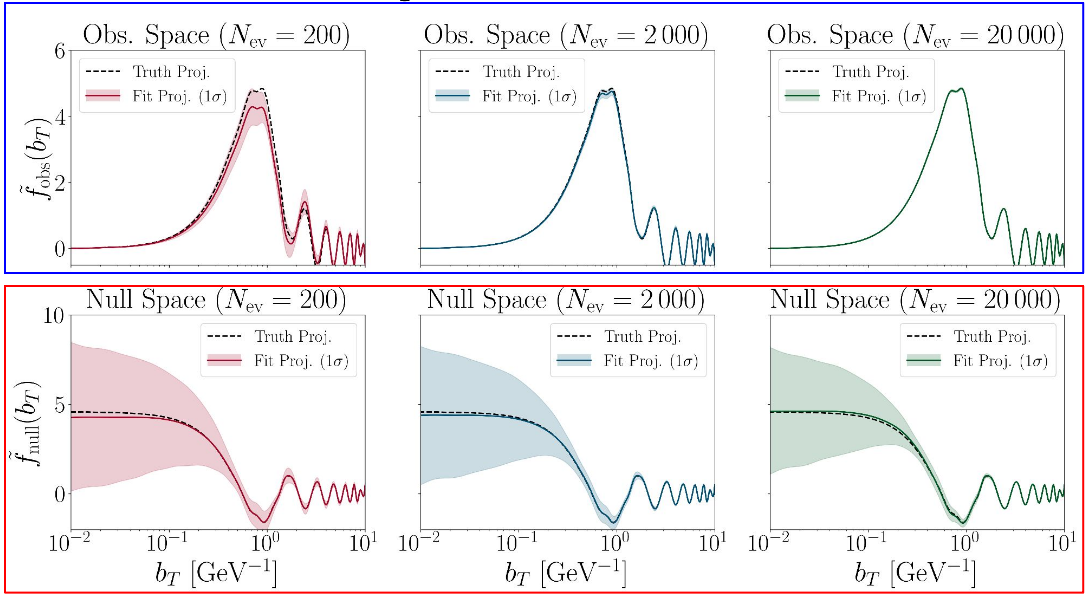


Only the Observable can be reconstructed from Data

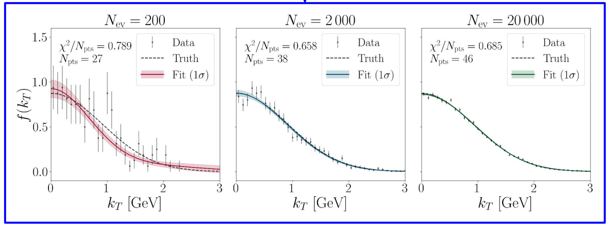
$$\mathcal{M}\tilde{\boldsymbol{f}} = \mathcal{M}\tilde{\boldsymbol{f}}_{\text{obs}} + \mathcal{M}\tilde{\boldsymbol{f}}_{\text{null}} \approx \mathcal{M}\tilde{\boldsymbol{f}}_{\text{obs}} \, ,$$

$$\chi^2(\boldsymbol{f}|\tilde{\boldsymbol{f}}) = \sum_{j=1}^{M} \left(\frac{[\mathcal{M}\tilde{\boldsymbol{f}}]_j - f_j}{\sigma_j} \right)^2 .$$

SVD: Fit Projections



Bands shrinks as the statistics increase
Enough to describe data



Bands DO NOT shrinks
Irrelevant for describing data

As for Shadow GPDs
adding terms prop. to to the Null component gives the same transform

$$\tilde{f} = \tilde{f}_{\text{obs}} + \tilde{f}_{\text{null}}$$

Null Models

Null Space is Matrix dependent:
kt points and range matter!

We can use the evolution for moving the Null space and shrink the error band

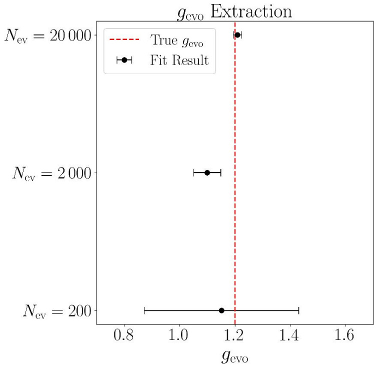
Multi-scale Analysis

4 different dataset with different energies and kT range

$$\mu_k^2 \in \{2, 5, 20, 50\}\ \mathrm{GeV}^2$$

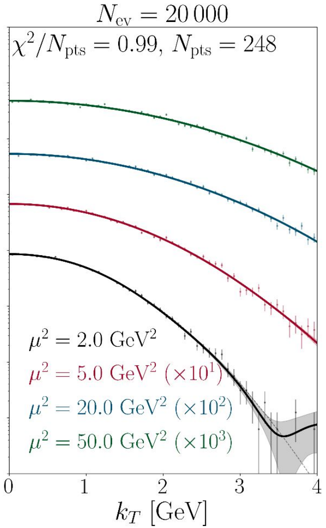
$$\sigma^2(\mu^2) = \alpha + g_{\mathrm{evo}} \ln\left(\frac{\mu^2}{\mu_0^2}\right)$$

N_{ev}	N_{pts}	χ^2/N_{pts}
200	148	0.99
2 000	207	1.05
20 000	248	0.99

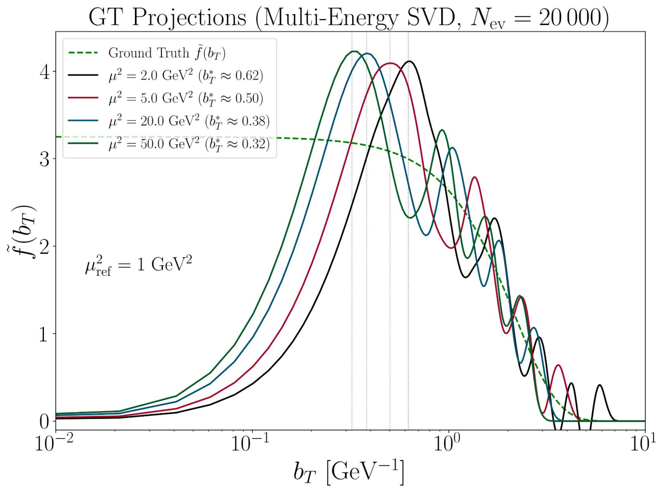


The method is capable of reconstructing g_{evo}

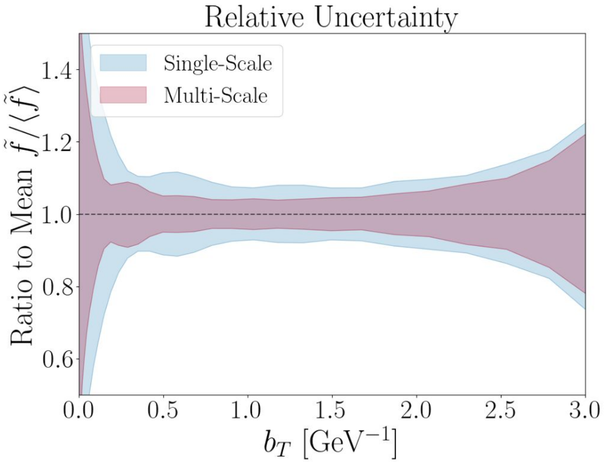
Fit vs Data



Increasing the kT range moves the Null Space



Single-scale vs Multi-scale



TMD Case: Data Generation and Fit

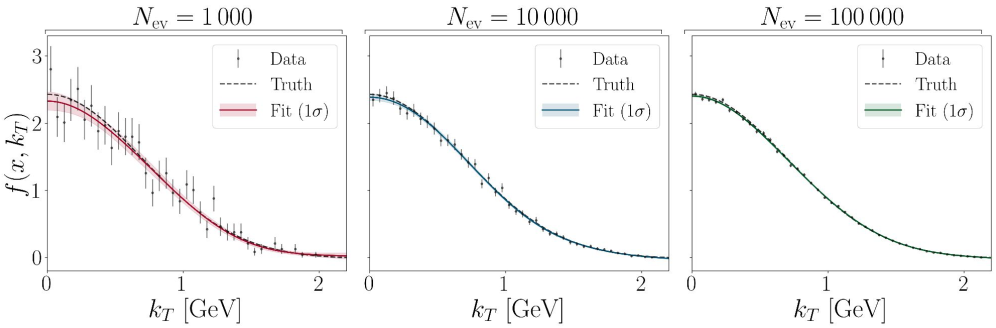
The TMD PDF is the observable

$$\tilde{f}_{f/P}(x, b_T; Q) = \sum_j \int_x^1 \frac{d\hat{x}}{\hat{x}} C_{f/j}(x/\hat{x}, b_*; \mu_b^2, \mu_b, g(\mu_b)) f_{j/P}(\hat{x}, \mu_b) \times \exp[S_{\rm pert}(b_T; \mu_b, Q)] M_f(x, b_T) \exp\left[-g_K(b_T) \ln \frac{Q}{Q_0}\right]$$

Non-perturbative Models

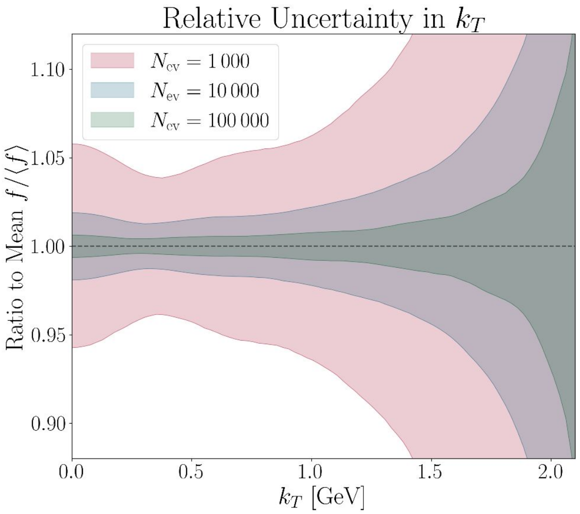
$$M_f(b_T) = R(b_T) + \tilde{p}(b_T)[1 - R(b_T)]$$

$$\tilde{p}_{\rm true}(b_T) = \exp\left(-\frac{b_T^2 w_f}{4}\right) \qquad g_K(b_T) = \frac{g_2}{2} b_T^2 [1 - R(b_T)]$$



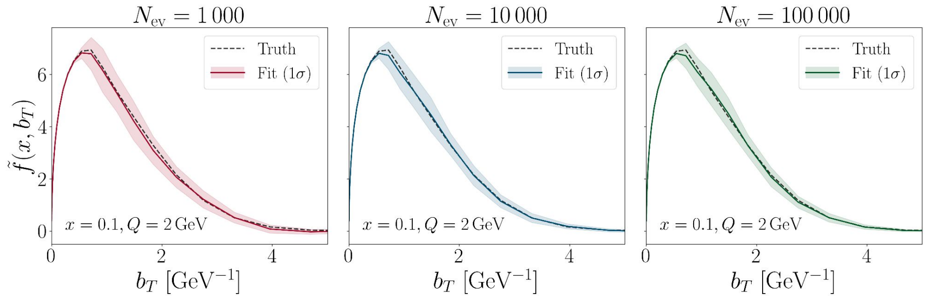
N_{ev}	accept. rate	N_{pts}	χ^2/N_{pts}
10^3	86%	37	0.83
10^4	83%	42	0.67
10^5	69%	43	0.94

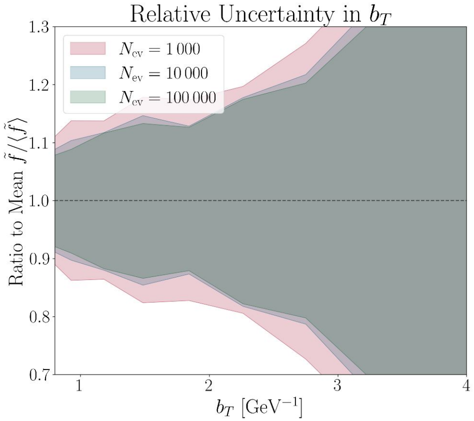
Good description of data
Uncertainty scales as expected



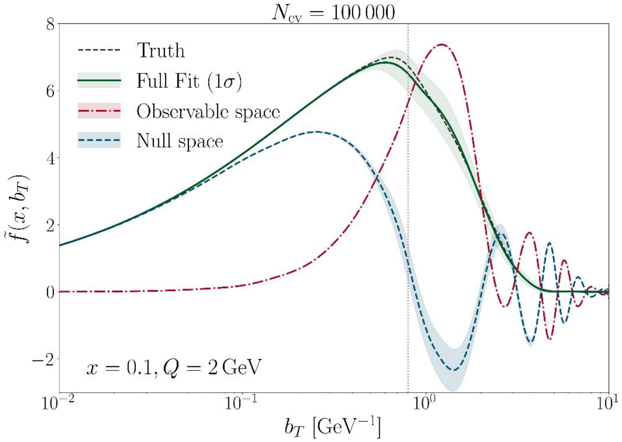
TMD Case: bT-space

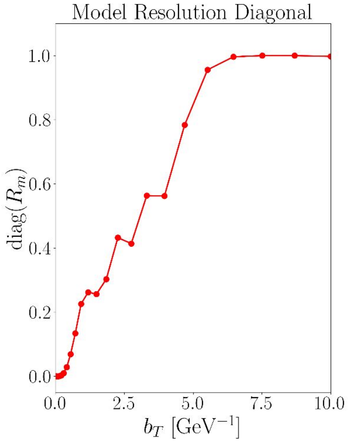
Error bands converge around the True distribution





TMD bands in bT-space do not shrink as the statistics increase





FIT = OBS + NULL

OBS shrinks as the statistics increases

NULL TMD: error bands stay the same as the statistics increase

The small bt region belongs to the NULL space!

$P_{\rm obs}$ is also referred as Model Resolution Matrix

It tells where we can resolve the function better

Fuu Case: Data vs Fit

TMD PDF fitted from the convolution

The TMD FF is fixed

$$F_{UU,T}(x,z,q_T,Q^2) = x\sum_a e_a^2 \frac{1}{2\pi}\int_0^\infty db_T\, b_T J_0(q_T b_T)\tilde{f}_1^a(x,b_T;Q)\tilde{D}_1^a(z,b_T;Q),$$

Only the up quark in considered
x = 0.1 and z=0.2

TMD FF

$$\tilde{D}_{H/f}(z,b_T;Q) = \sum_j \int_z^1 \frac{d\hat{z}}{\hat{z}^{3-2\epsilon}} \tilde{C}_{j/f}(z/\hat{z}, b_*; \mu_b^2, \mu_b, g(\mu_b)) d_{H/j}(\hat{z};\mu_b)$$
$$\times \exp\{S_{\rm pert}(b_T;\mu_b,Q)\}\, M_D(z,b_T) \exp\left\{-g_K(b_T)\ln\frac{Q}{Q_0}\right\}.$$

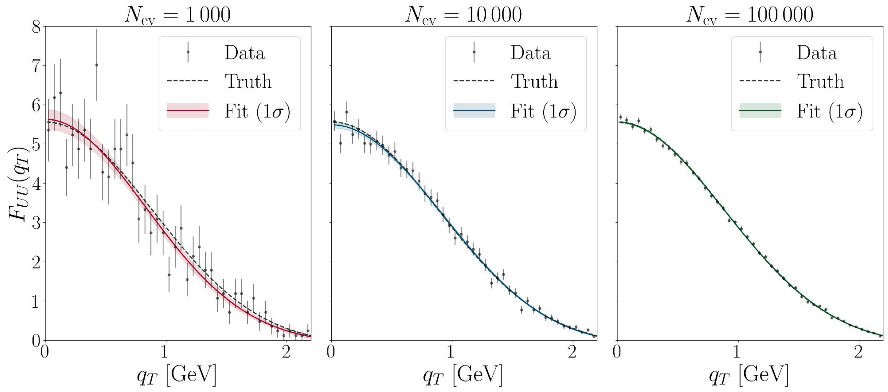
$$M_D(b_T) = R(b_T) + [1-R(b_T)]\exp\left(-\frac{b_T^2 w_D}{4}\right)$$

$$\mathbf{F}_{UU,T} = \mathcal{M}_{\rm eff}\tilde{\boldsymbol{f}}_1$$

TMD FF acts as a kernel

$$\mathcal{M}_{\rm eff} = \mathcal{M}_{\rm bridge}\cdot {\rm diag}\left(x\sum_a e_a^2 \tilde{\mathbf{D}}_1^a\right)$$

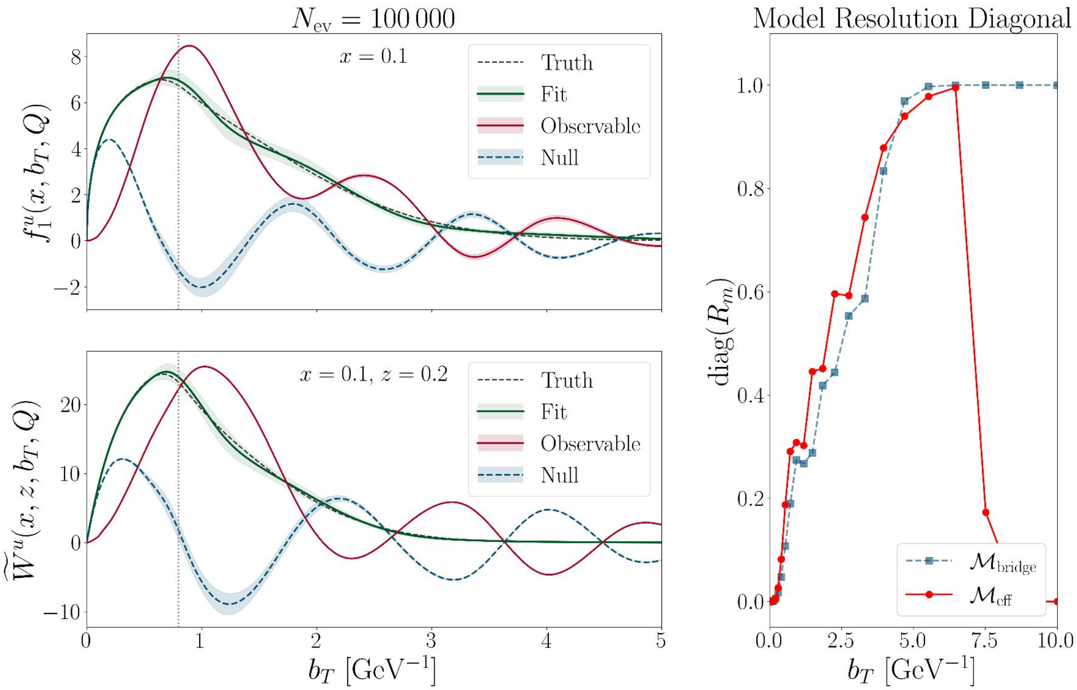
bT to kT Matrix



The fit tracks the data in qT as the statistics Nev increase.

$N_{\rm ev}$	accept. rate	$N_{\rm pts}$	$\chi^2/N_{\rm pts}$
10^3	89%	45	0.83
10^4	89%	48	0.92
10^5	84%	48	1.29

Fuu Case: Data vs Fit



Error bands are around the Ground Truth Model

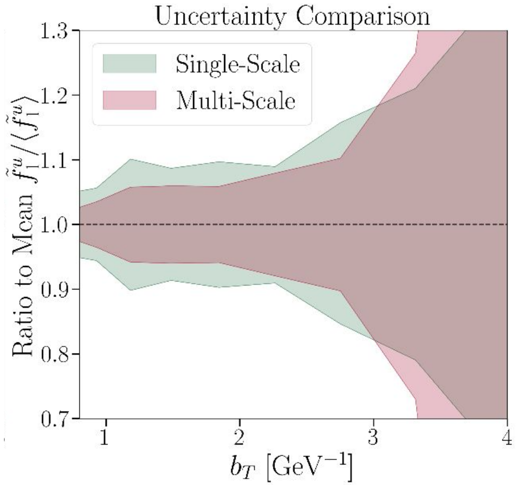
FIT = OBS + NULL
OBS shrinks as the statistics increases
NULL stay the same as the statistics increase

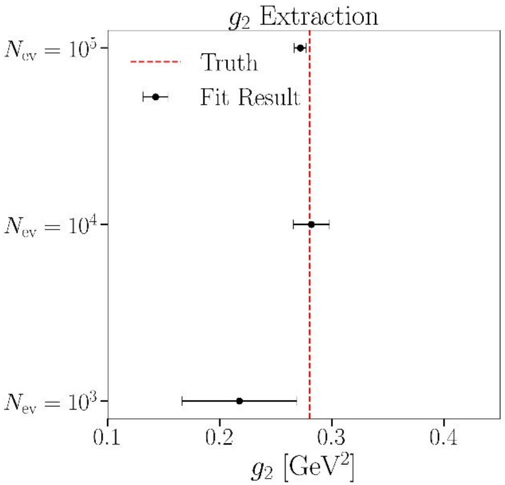
The small bt region belongs to the NULL space!

Effective Resolution Matrix: zero for large bT values because of TMD FF

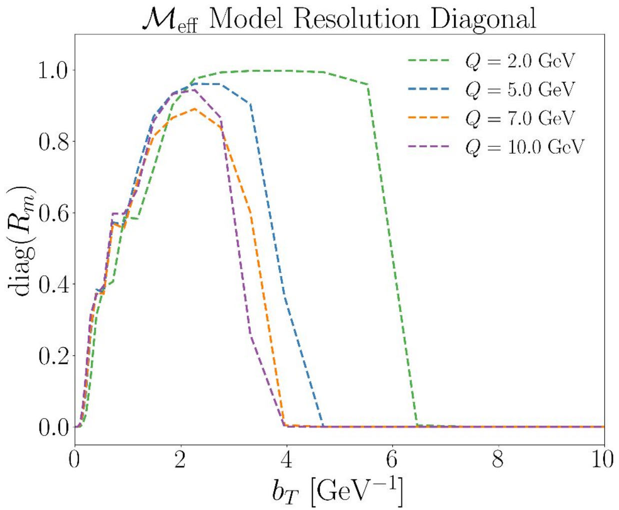
Fuu Case: Multi-Scale Fit vs Data

Multi-Scale analysis allows to shrink the error band compared to Single-Scale analysis



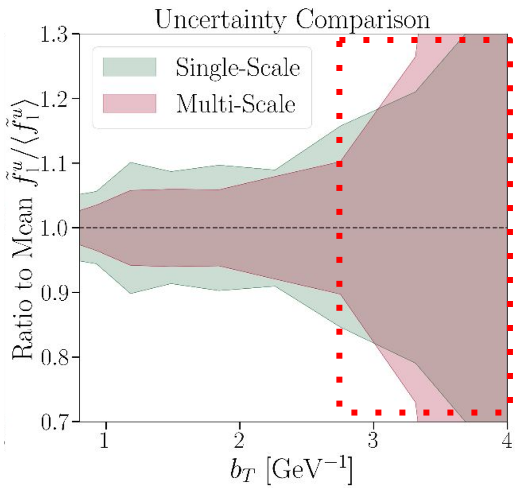


TMD FF is screening some bT regions

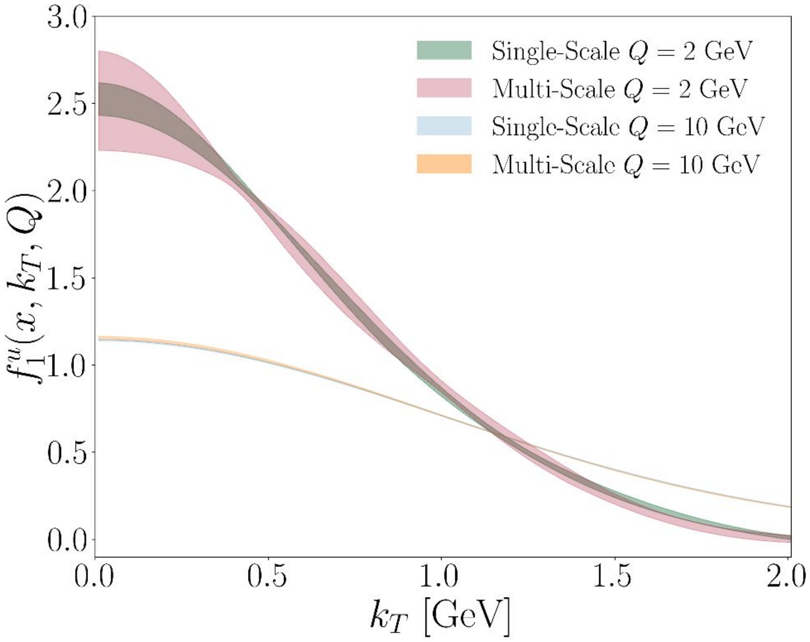


Fuu Case: Multi-Scale Fit vs Data

Multi-Scale error band is larger than Single-Scale case due to g2 extraction



At small energy this leads to Multi-Scale band larger than Single-Scale band



Conclusions

- Developed a pixel-based, model-independent Bayesian approach for TMD reconstruction,
- Utilizing Generative AI (Normalizing Flow) and MH for estimation and uncertainty quantification.
- The non-unique solution involves Observable and Null TMD components.
- A "Precision Floor" in bT space, attributed to the Null space,
- Highlights the necessity of Multi-Scale analysis to reduce uncertainty and overcome Null space limitations.

TMDs in the Lens of Generative AI:
A Pixel-Based Approach to Partonic Imaging
ArXiv 2605.06606