Sharpening the numerical tools for precision calculations with Color Glass Condensate

Piotr Korcyl

in collaboration with H. Le, F. Cougoulic, F. Salazar, and T. Stebel

High energy QCD: from the LHC to the EIC
Benasque Science Center, August 12, 2025

NATIONAL SCIENCE CENTRE POLAND This work is supported by NCN grant nr 2022/46/E/ST2/00346.

CGC: Color Glass Condensate is an effective description of processes in pp and pA collisions at high energies $\Longleftrightarrow$ small x

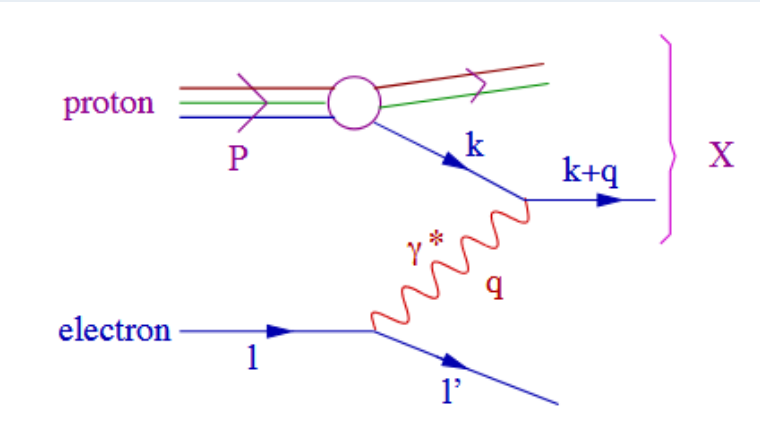
- access to unintegrated parton distribution functions
- nonperturbative initial condition - model
- perturbative evolution equation
- predictions of cross-sections at high-energies/small-x

Goal of the exercise

Test the universality of the CGC approach:

- so far, the fits were performed one observable at a time
- fitted parameters were used to calculate another observable and compared with experimental data
  Single inclusive particle production at high energy from HERA data to proton-nucleus collisions, T. Lappi, H. Mäntysaari, Phys.Rev.D 88 (2013) 114020
  Global Bayesian Analysis of J/ψ Photoproduction on Proton and Lead Targets, H. Mäntysaari, H. Roch, F. Salazar, B. Schenke, C. Shen, 2507.14087
- we plan to investigate if a simultaneous fit to a set of observables is possible
- on the way, we will develop numerical tools that, unexpectedly, can have more applications and open new opportunities

Virtual photon-proton cross section for transverse (T) and longitudinal (L) polarization of the virtual photon



$$x = \frac{-q^2}{(P+q)^2 - q^2 - M^2}$$

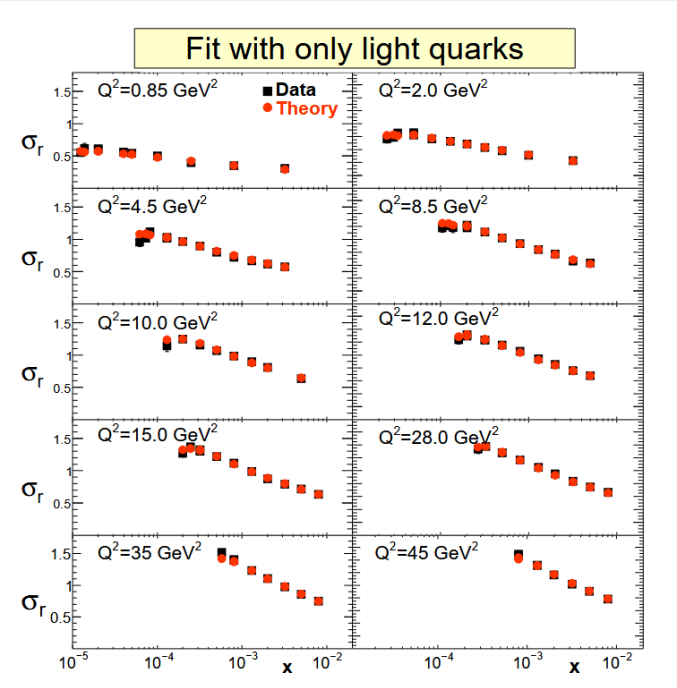
E. Iancu, *QCD in heavy ions collisions*

$$\sigma_{T,L}(x, Q^2) = \sigma_0 \sum_{f=u,d,s} \int_0^1 dz \int d\bar{\boldsymbol{r}}_\perp |\psi^f_{T,L}(e_f, m_f, z, Q^2, \bar{\boldsymbol{r}}_\perp)|^2 N_F(x, \bar{\boldsymbol{r}}_\perp)$$

Initial condition

$$N_F^{\mathrm{MV}}(x = x_0, \bar{\boldsymbol{r}}_\perp) = 1 - \exp\left[-\frac{\left(r^2 Q_0\right)^\gamma}{4} \ln\left(\frac{1}{r\Lambda} + e\right)\right]$$

Comparison with experimental data for the reduced cross sections in different Q^2 bins



$Q^2_{s0} = 0.164$ GeV^2 at $x_0 = 0.01$, $\sigma_0 = 32.324$, $\gamma = 1.123$, $C = 2.48$ and $m_l = 0.0182$

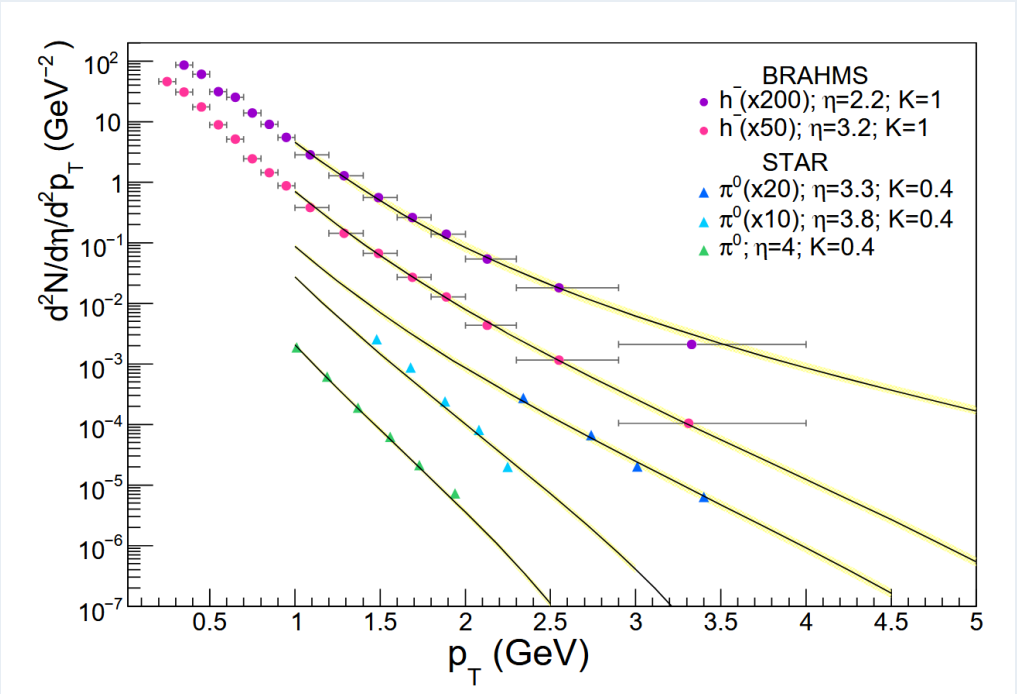
J. L. Albacete, N. Armesto, J.G. Milhano, P. Quiroga Arias, C.A. Salgado, *AAMQS: A non-linear QCD analysis of new HERA data at small-x including heavy quarks*, Eur. Phys. J. C 71, 1705 (2011)

Negatively charged hadron and π^0 yields in proton-proton collisions at $\sqrt{S_{NN}} = 200$ GeV

$$\frac{dN_h}{dy_h d^2 p_t} = \frac{\frac{1}{2}\sigma_0}{\sigma_{\text{inel}}} \frac{K}{(2\pi)^2} \int_{x_F}^1 \frac{dz}{z^2} \left\{ \sum_q \left[x_1 f_{q/p}(x_1, p_t^2) N_F(x_2, \frac{p_t}{z}) D_{h/q}(z, p_t^2) \right] + \right.$$
$$\left. + \left[x_1 f_{g/p}(x_1, p_t^2) N_A(x_2, \frac{p_t}{z}) D_{h/g}(z, p_t^2) \right] \right\}$$

J. Albacete, C. Marquet, *Single inclusive hadron production at RHIC and the LHC from the Color Glass Condensate*, Phys.Lett.B687:174-179,2010

Negatively charged hadron and π^0 yields in proton-proton collisions at $\sqrt{S_{NN}} = 200$ GeV



$Q_{s0}^2 = 0.4$ GeV^2 at $x_0 = 0.02$ was the only fitted parameter

J. Albacete, C. Marquet, *Single inclusive hadron production at RHIC and the LHC from the Color Glass Condensate*, Phys.Lett.B687:174-179,2010

Framework: requirements

- evolution equation
- initial condition
- access to gluon dipole amplitude in position and momentum spaces
- minimization algorithm $\Rightarrow$ levmar library
- access to fragmentation functions and PDFs $\Rightarrow$ LHAPDF library

Framework: features

- BK evolution equation with kinematical constraint
- Balitsky/daughter/mother dipole prescription for the running coupling
- Euler/Runge integration scheme
- uncertainty estimation
- parallelization $\Rightarrow$ shorter running time
- ...

Balitsky-Kovchegov evolution equation

The LO BK equation reads

$$\frac{\partial S_{\bar{\boldsymbol{x}}_\perp \bar{\boldsymbol{y}}_\perp}(\eta)}{\partial \eta} = \frac{\bar{\alpha}_s}{2\pi} \int d^2 \bar{\boldsymbol{z}}_\perp \, \mathscr{M}_{\bar{\boldsymbol{x}}_\perp \bar{\boldsymbol{y}}_\perp \bar{\boldsymbol{z}}_\perp} \left[S_{\bar{\boldsymbol{x}}_\perp \bar{\boldsymbol{z}}_\perp}(\eta) S_{\bar{\boldsymbol{z}}_\perp \bar{\boldsymbol{y}}_\perp}(\eta) - S_{\bar{\boldsymbol{x}}_\perp \bar{\boldsymbol{y}}_\perp}(\eta) \right],$$

where

$$\mathscr{M}_{\bar{\boldsymbol{x}}_\perp \bar{\boldsymbol{y}}_\perp \bar{\boldsymbol{z}}_\perp} = \frac{(\bar{\boldsymbol{x}}_\perp - \bar{\boldsymbol{y}}_\perp)^2}{(\bar{\boldsymbol{x}}_\perp - \bar{\boldsymbol{z}}_\perp)^2 (\bar{\boldsymbol{z}}_\perp - \bar{\boldsymbol{y}}_\perp)^2}.$$

Rewritten in radial variables

$$\frac{\partial S(r,\eta)}{\partial \eta} = \frac{\bar{\alpha}_s}{2\pi} \int d\phi \, dr_z \, r_z \, \frac{r^2}{r_z^2 (r^2 + r_z^2 - 2 r r_z \cos\phi)} \times \\ \times \left[S(r_z,\eta) \, S\left(\sqrt{r^2 + r_z^2 - 2 r r_z \cos\phi}, \eta \right) - S(r,\eta) \right].$$

Balitsky-Kovchegov evolution equation

Account for several additional physical effects, such as the running of the coupling constant with the energy scale, resummation of subleading corrections

$$\frac{\partial S(r,\eta)}{\partial \eta} = \int d\phi\, dr_z\, r_z \times$$
$$\times \left[\frac{\bar{\alpha}_s(r)}{2\pi r_z^2}\left(\frac{r^2}{r_{zy}^2+\varepsilon^2}+\frac{\bar{\alpha}_s(r_z)}{\bar{\alpha}_s(r_{zy})}-1+\frac{r_z^2}{r_{zy}^2+\varepsilon^2}\left(\frac{\bar{\alpha}_s(r_{zy})}{\bar{\alpha}_s(r_z)}-1\right)\right)\right]\times$$
$$\times \left[S(r_z,\eta-\delta_{r_z;r})S(r_{zy},\eta-\delta_{r_{zy};r})-S(r,\eta)\right],$$

where $r_{zy}=\sqrt{r^2+r_z^2-2rr_z\cos\phi}$. The shifts in η in the dipole amplitudes are given by $\delta_{r_z;r}=\max\left\{0,2\log\frac{r}{r_z}\right\}$ and similarly $\delta_{r_{zy};r}=\max\left\{0,2\log\frac{r}{r_{zy}}\right\}$.

B. Ducloué, E. Iancu, G. Soyez, D.N. Triantafyllopoulos, *HERA data and collinearly-improved BK dynamics*, Phys.Lett.B 803 (2020) 135305

Automatic Differentiation in a nutshell

Allows to evaluate 'analytic' derivatives of a computer program with respect to external parameters.

- numbers are promoted to vectors

$$x \to \begin{pmatrix} x \\ \partial_A \\ \partial_B \\ \partial_A^2 \\ \partial_A \partial_B \\ \vdots \end{pmatrix}$$

- all arithmetic operators are overloaded
- functions with derivatives have to be provided
- works for most algorithms

Automatic Differentiation for the Balitsky-Kovchegov evolution equation

$$S(r,\eta) \to \begin{pmatrix} S(r,\eta) \\ \partial_{Q_0} S(r,\eta) \\ \partial_r S(r,\eta) \\ \partial^2_{Q_0} S(r,\eta) \\ \partial^2_r S(r,\eta) \end{pmatrix}$$

Then

$$\frac{\partial S(r,\eta)}{\partial \eta} = \int d\phi\, dr_z\, r_z \times$$
$$\times \left[\frac{\bar{\alpha}_s(r)}{2\pi r_z^2} \left(\frac{r^2}{r_{zy}^2+\varepsilon^2} + \frac{\bar{\alpha}_s(r_z)}{\bar{\alpha}_s(r_{zy})} - 1 + \frac{r_z^2}{r_{zy}^2+\varepsilon^2} \left(\frac{\bar{\alpha}_s(r_{zy})}{\bar{\alpha}_s(r_z)} - 1 \right) \right) \right] \times$$
$$\times \left[S(r_z, \eta - \delta_{r_z;r}) S(r_{zy}, \eta - \delta_{r_{zy};r}) - S(r,\eta) \right],$$

gives $S(r,\eta)$ together with the evolved derivatives.

F. Cougoulic, P. Korcyl, T. Stebel, *Improving the solver for the Balitsky-Kovchegov evolution equation with Automatic Differentiation*, Comput.Phys.Commun. 313 (2025) 109616

Automatic Differentiation for the Balitsky-Kovchegov evolution equation

Benefits:

- faster convergence of the fit
  - fewer iterations
  - less computer time
  - can test more parameters in the initial condition
- access to the Hessian matrix allows easy estimation of uncertainties
- more reliable estimation of some TMD functions with long tails
- can tell how the initial condition is sensitive to the given experimental data

Costs:

- slower code, but less than naively expected

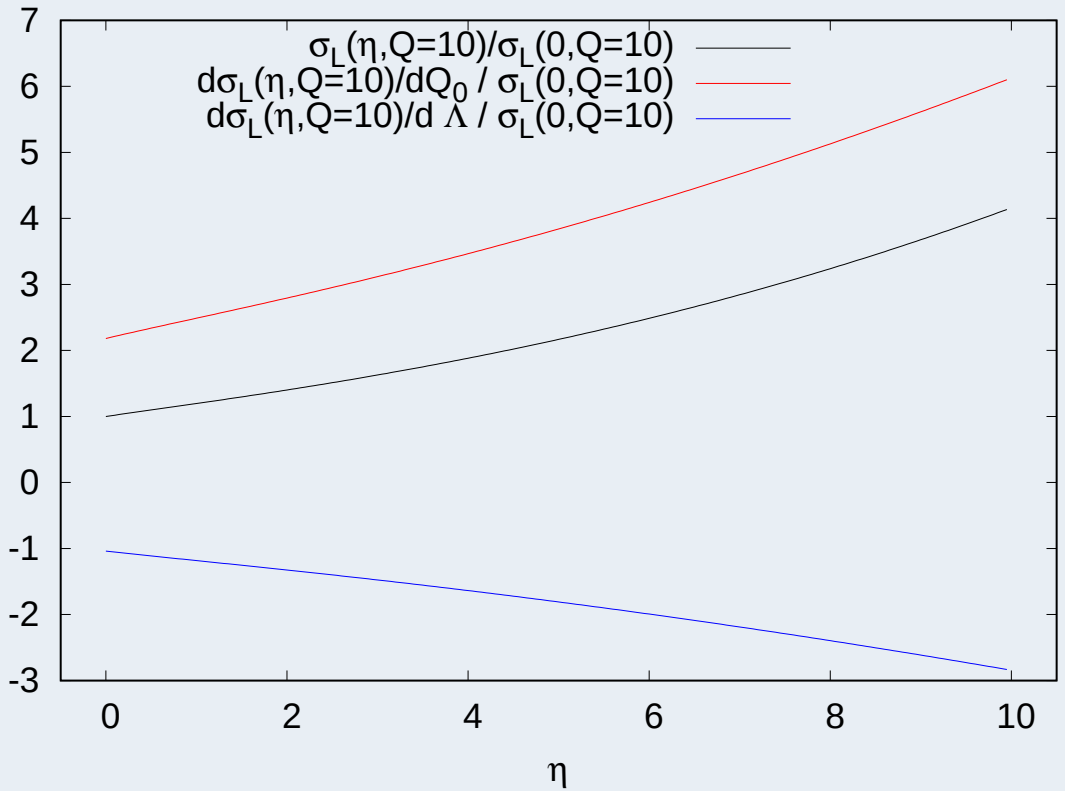
Timings

data type	size	without AD	r derivatives only	r derivatives + Hessian(2 params)
`double`	250	0m3(1)	0m5(0.5)	0m12(1)
`double`	500	0m11(1)	0m22(1)	0m50(1)
`double`	1000	0m46(1)	1m29(1)	3m22(2)
`long double`	1000	3m52(3)	6m11(2)	13m51(6)
`double`	2000	3m3(1)	5m50(4)	13m47(3)

data type	size	mode	time	—
`double`	1000	without AD	0m46(1)	
`double`	1000	r derivatives only	1m29(1)	
`double`	1000	r derivatives + Hessian(1 param)	2m12(1)	
`double`	1000	r derivatives + Hessian(2 params)	3m22(2)	
`double`	1000	r derivatives + Hessian(3 params)	5m4(2)	
`double`	1000	r derivatives + Hessian(4 params)	7m20(5)	
`double`	1000	r derivatives + Hessian(5 params)	9m42(3)	

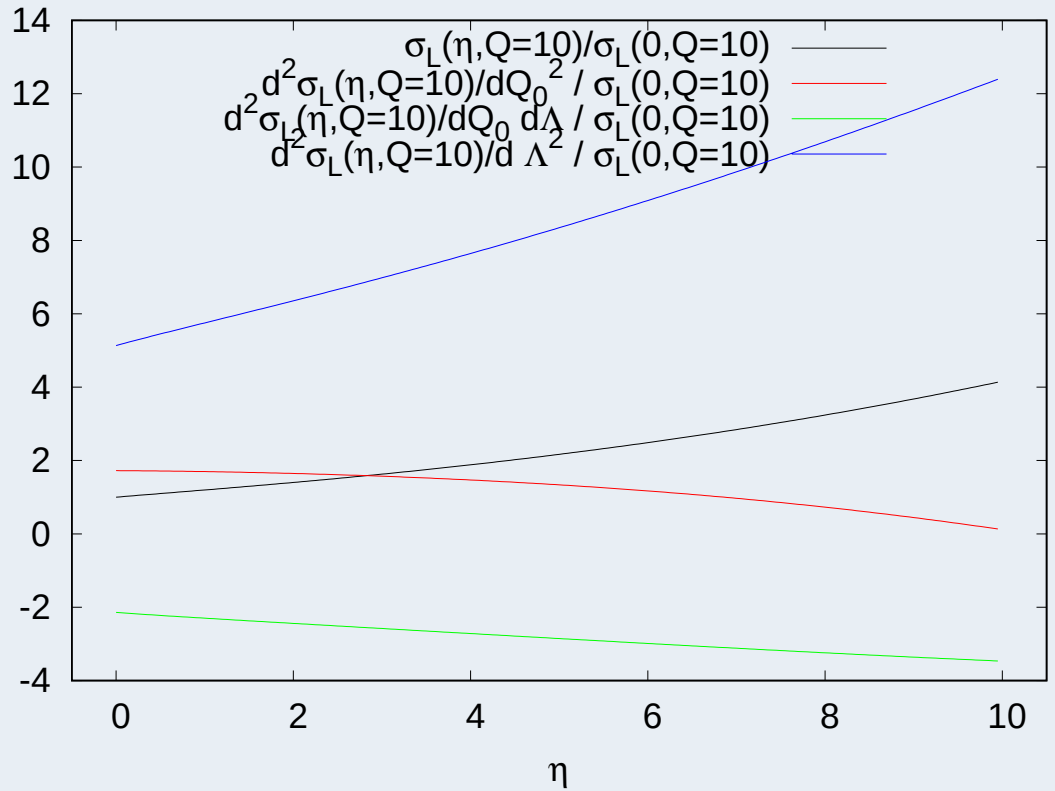
Timings on a two-socket 16-core procesor running with 64 openMP threads.

Advantage 1: Sensitivity of the observable to the parameters of the initial condition



First derivatives of the longitudinal cross-section $\sigma_L^{\gamma^*,p}(x, Q^2 = 10\,\mathrm{GeV}^2)$

Advantage 2: Immediate access to the exact Hessian matrix at the minimum



Second derivatives of the cross-section $\sigma_L^{\gamma^*,p}(x, Q^2 = 10\,\mathrm{GeV}^2)$

Advantage 2: Immediate access to the exact Hessian matrix at the minimum

Assume that $\chi^2_{\rm global}$ is quadratic about the global minimum

$$\Delta\chi^2_{\rm global} \equiv \chi^2_{\rm global} - \chi^2_{\rm min} = \sum_{i,j=1}^{n} H_{ij}\left(a_i - a_i^0\right)\left(a_j - a_j^0\right),$$

where

$$H_{ij} = \frac{1}{2}\frac{\partial^2 \chi^2_{\rm global}}{\partial a_i \partial a_j}\bigg|_{\rm min}$$

We can diagonalize the covariance matrix $C \equiv H^{-1}$,

$$\sum_{j=1}^{n} C_{ij} v_{jk} = \lambda_k v_{ik},$$

$$a_i - a_i^0 = \sum_{k=1}^{n} \left(\sqrt{\lambda_k} v_{ik}\right) z_k \quad \Rightarrow \quad \Delta\chi^2_{\rm global} = \sum_{k=1}^{n} z_k^2 \equiv T^2$$

Advantage 2: Immediate access to the exact Hessian matrix at the minimum

Comparison of the uncertainties obtained from the Hessian and Monte Carlo methods for the PDFs

(a) (b)

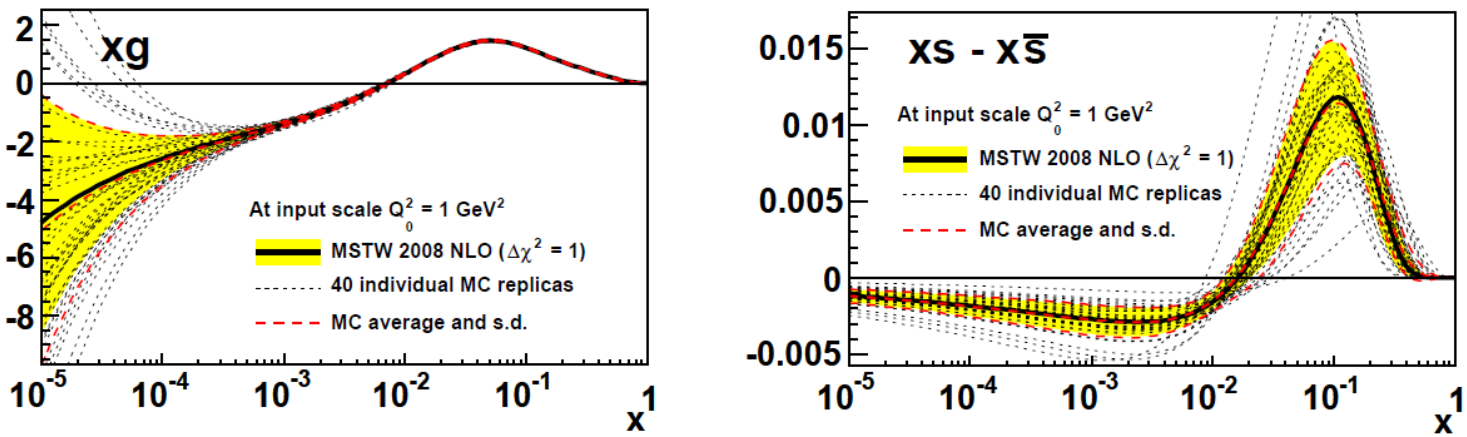
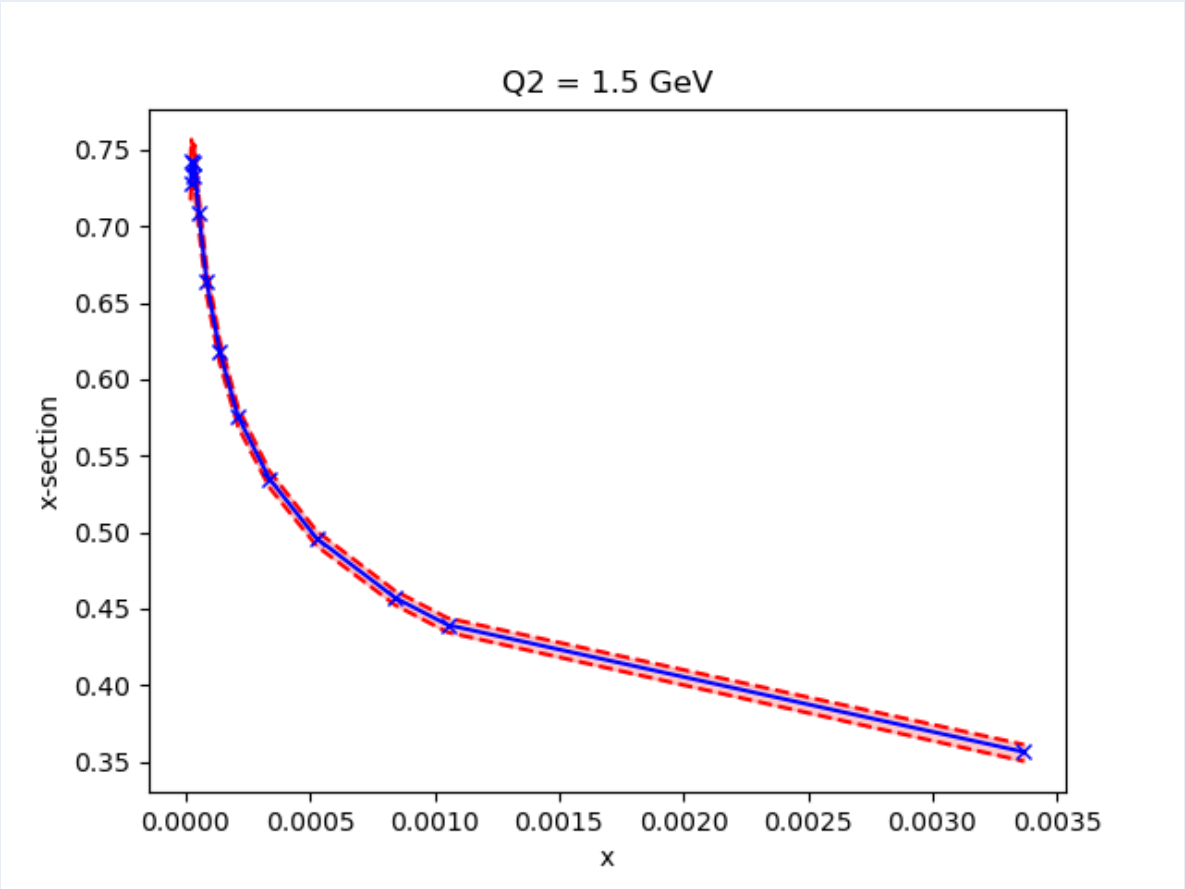


Figure 1. Comparison of Hessian and Monte Carlo results at the input scale of $Q_0^2 = 1$ GeV^2 for the (a) gluon distribution and (b) strange asymmetry. Both results allow $n = 20$ free PDF parameters and do not apply a tolerance (i.e. $T = 1$ in the Hessian case). The best-fit (solid curves) and Hessian uncertainty (shaded region) are in good agreement with the average and standard deviation (thick dashed curves) of the $N_{\text{rep}} = 40$ Monte Carlo replica PDF sets (thin dotted curves).

G. Watt, R. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW2008 PDFs*, JHEP 1208:052, 2012

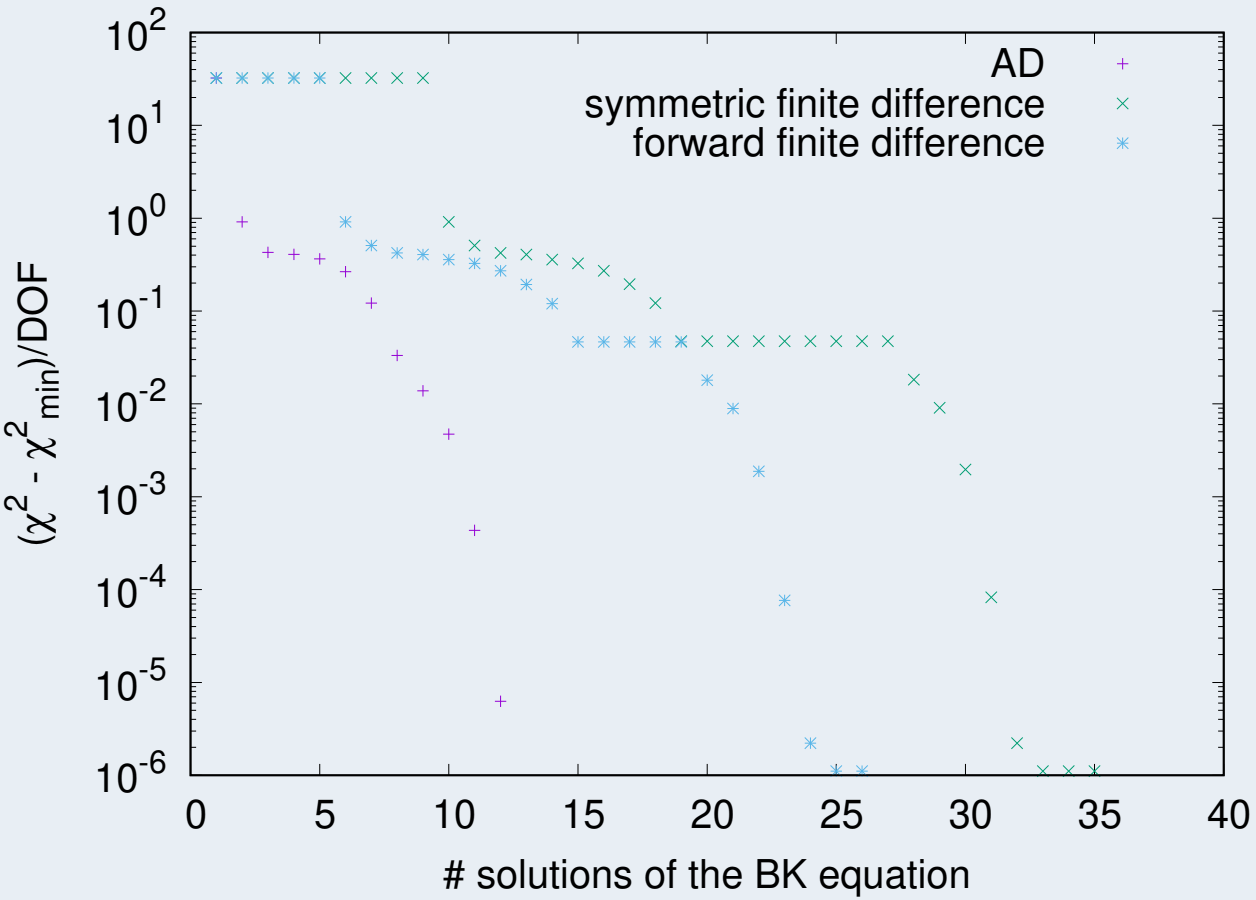
Advantage 2: Immediate access to the exact Hessian matrix at the minimum



Uncertainty of the DIS cross-section obtained with the Hessian method

Initial condition for the Balitsky-Kovchegov equation at next-to-leading order, C. Casuga, H. Hänninen, H Mäntysaari, Phys.Rev.D 112 (2025) 3, 034003

Advantage 3: Increased efficiency of the Levenberg-Marquard optimization algorithm



Fitting 4 parameters to the DIS HERA data

Logarithmic Fourier Transform

Popular in geophysics, cosmology, and signal processing.

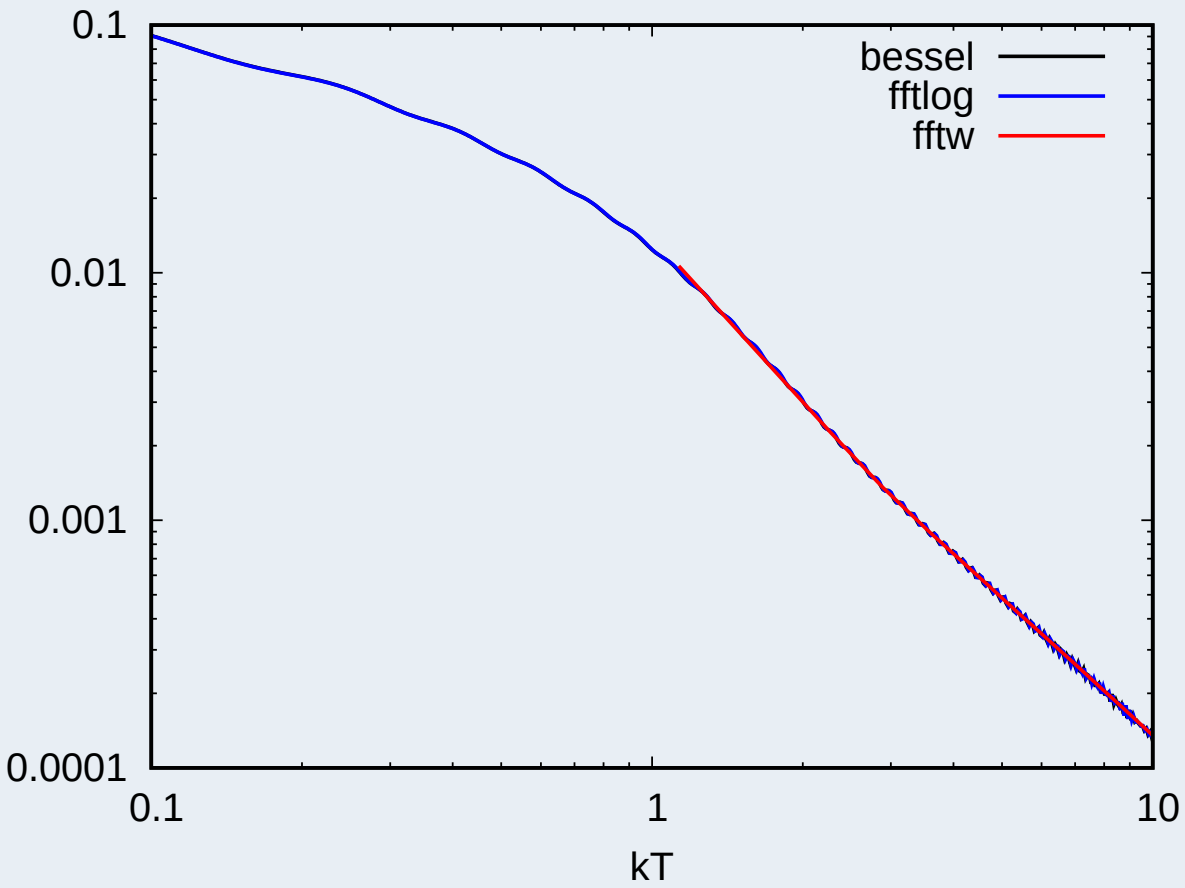
- It allows for the Fourier transform of data sampled on a logarithmic scale rather than linear,
- fits perfectly into our setup as we solve the BK equation on a logarithmic grid,
- more reliable than an ordinary 2D Fourier transform,
- Bessel function is not needed,
- order of magnitude more efficient in computer time.

Main idea:

$$\tilde{f}(k) = C(k)\ \mathrm{FT}_{1D}^{\tau\to k}\left[\ B(\tau)\ \mathrm{FT}_{1D}^{x\to\tau}\ \left[\ A(x)\ f(x)\ \right]\ \right]$$

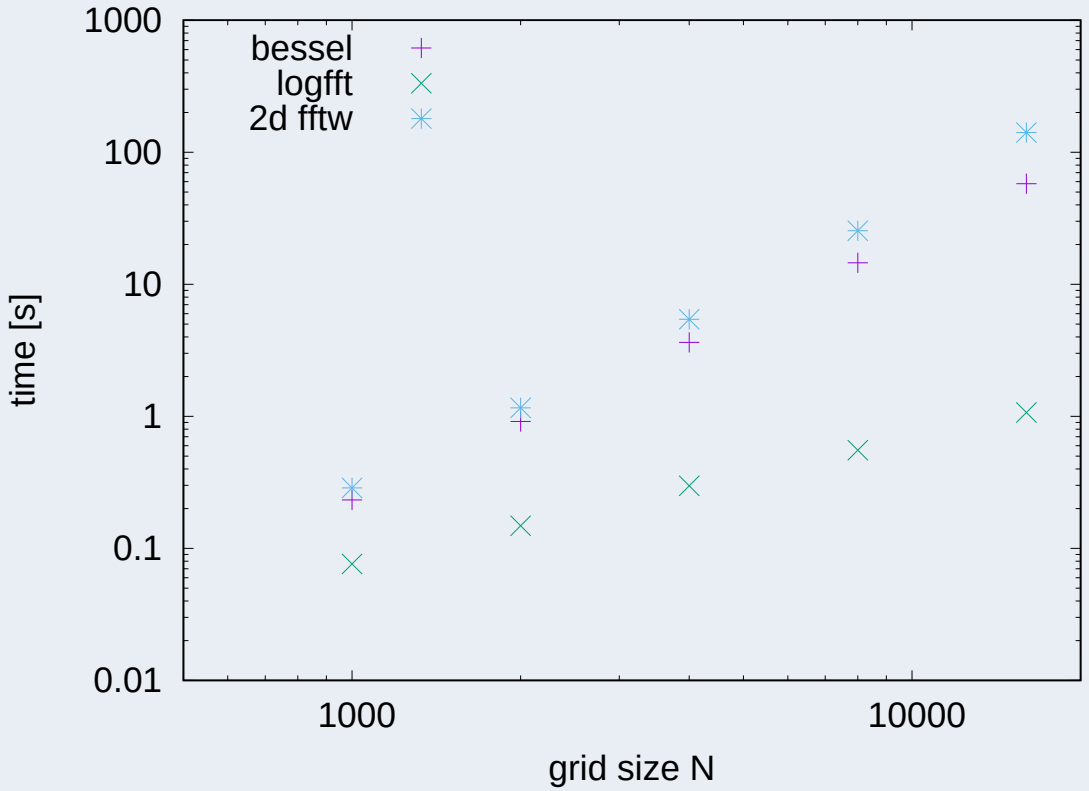
where $A(x)$, $B(\tau)$, and $C(k)$ are known functions that can be precomputed. $\mathrm{FT}_{1D}^{x\to k}$ is an ordinary, linear, one-dimensional FT.

Logarithmic Fourier Transform



The WW TMD structure function obtained with three different methods with MV model, $N = 2000$

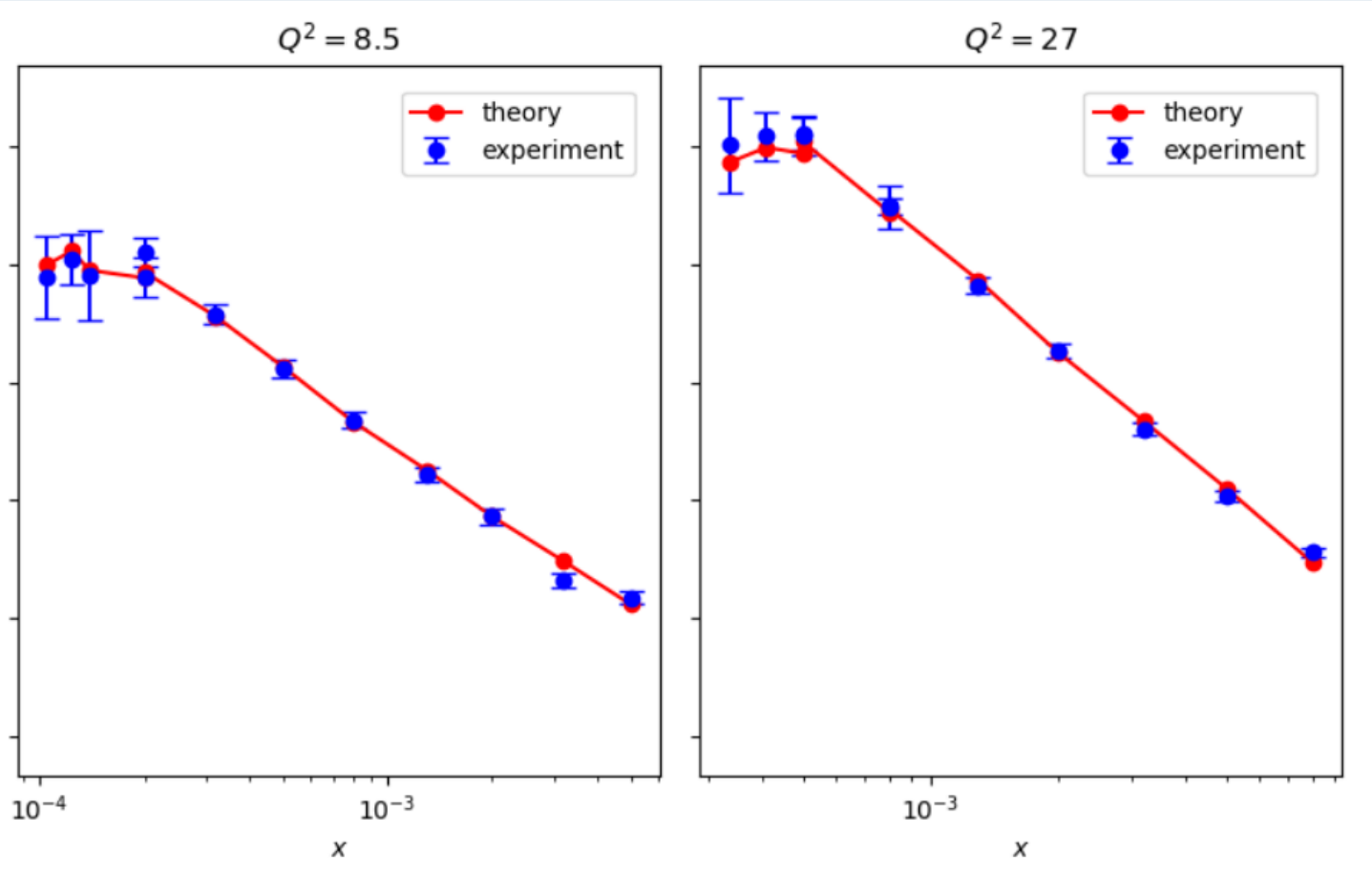
Logarithmic Fourier Transform



Time comparison: the logfft approach, numerical integration of the Bessel functions, and a 2D Fourier transform evaluated using the FFTW3 library.

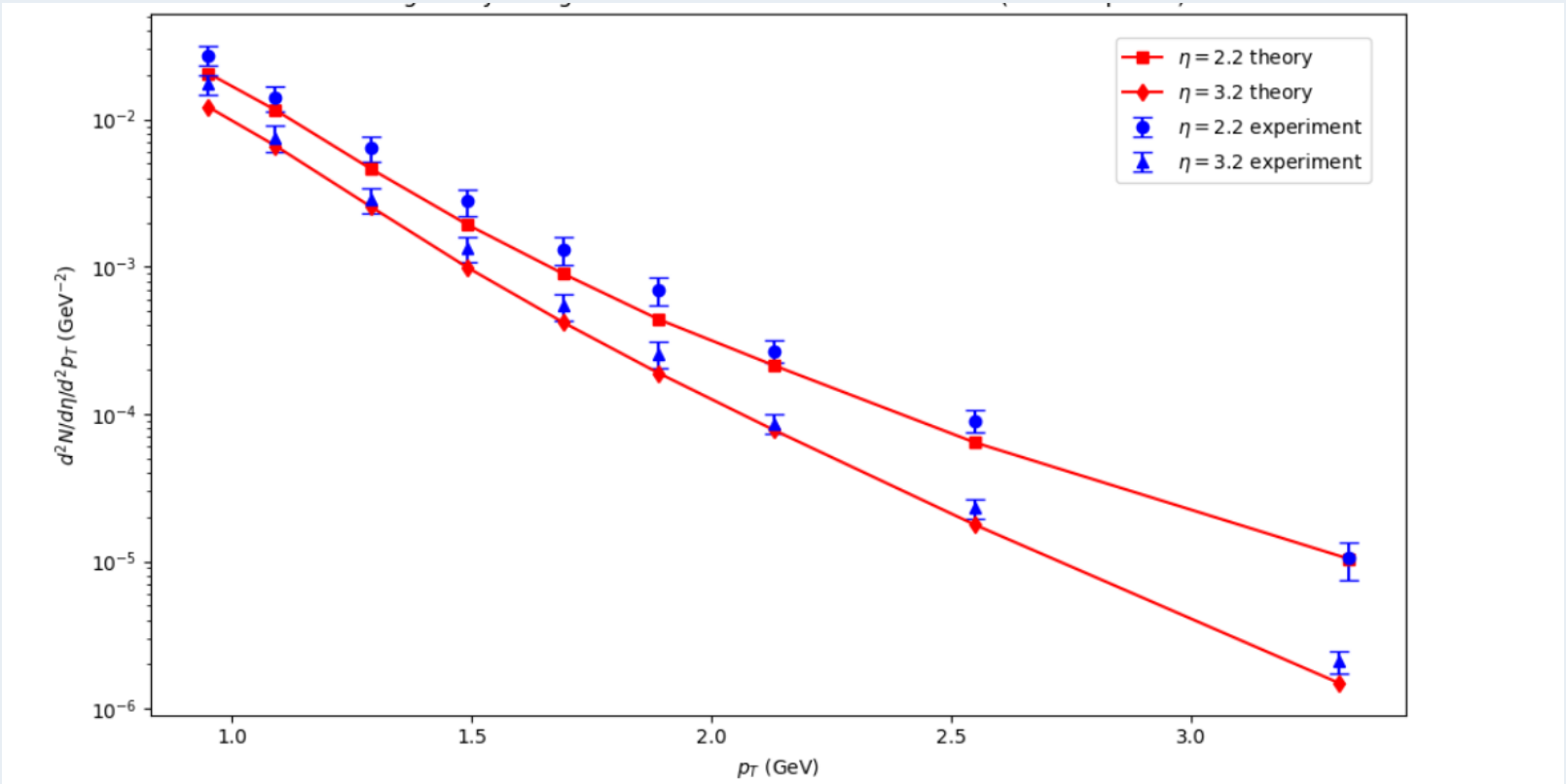
Efficient Fourier Transforms for Transverse Momentum Dependent Distributions, Z.-B. Kang, A. Prokudin, N. Sato, John Terry, Comput.Phys.Commun. 258 (2021) 107611

Preliminary results for combined fit



Back to DIS data: global fit with Balitsky running coupling prescription and kinematical constraint yielded $\chi^2/\mathrm{DOF} \approx 1.3$

Preliminary results for combined fit



Back to BRAMHS data: charged pions yields in proton-proton collisions at $\sqrt{S_{NN}} = 200$ GeV

Going further: dijet production

$$\frac{d\sigma(pA\to qgX)}{d^2P_t d^2k_t dy_1 dy_2} = \frac{\alpha_s^2}{2C_F}\frac{z(1-z)}{{P_t}^4}x_1 q(x_1,\mu^2)P_{gq}(z) \times\left\{\left[(1-z)^2-\frac{z^2}{N_c^2}\right]\mathcal{F}_{qg}^{(1)}(x_2,k_t)+\mathcal{F}_{qg}^{(2)}(x_2,k_t)\right\}, \tag{2.12}$$

$$\frac{d\sigma(pA\to q\bar{q}X)}{d^2P_t d^2k_t dy_1 dy_2} = \frac{\alpha_s^2}{2C_F}\frac{z(1-z)}{{P_t}^4}x_1 g(x_1,\mu^2)P_{qg}(z)\left\{[(1-z)^2+z^2]\mathcal{F}_{gg}^{(1)}(x_2,k_t) +2z(1-z)\mathrm{Re}\,\mathcal{F}_{gg}^{(2)}(x_2,k_t)-\frac{1}{N_c^2}\mathcal{F}_{gg}^{(3)}(x_2,k_t)\right\}, \tag{2.13}$$

$$\frac{d\sigma(pA\to ggX)}{d^2P_t d^2k_t dy_1 dy_2} = \frac{\alpha_s^2}{2C_F}\frac{z(1-z)}{{P_t}^4}x_1 g(x_1,\mu^2)P_{gg}(z) \times\left\{[(1-z)^2+z^2]\mathcal{F}_{gg}^{(1)}(x_2,k_t)+2z(1-z)\mathrm{Re}\,\mathcal{F}_{gg}^{(2)}(x_2,k_t)+\mathcal{F}_{gg}^{(6)}(x_2,k_t) +\frac{1}{N_c^2}\left[\mathcal{F}_{gg}^{(4)}(x_2,k_t)+\mathcal{F}_{gg}^{(5)}(x_2,k_t)-2\mathcal{F}_{gg}^{(3)}(x_2,k_t)\right]\right\}, \tag{2.14}$$

C. Marquet, E. Petreska, C. Roiesnel, *Transverse-momentum-dependent gluon distributions from JIMWLK evolution*, JHEP10 (2016) 065

TMD approximated in terms of the dipole amplitude S

$$\mathscr{F}_{qg}^{(1)}(\boldsymbol{k}_\perp, x) = \frac{N_c}{2\pi^2}\int \frac{r_\perp \mathrm{d}r_\perp}{2\pi} J_0(k_\perp r_\perp)\nabla_\perp^2 \left[1 - S(r_\perp, x)\right]$$

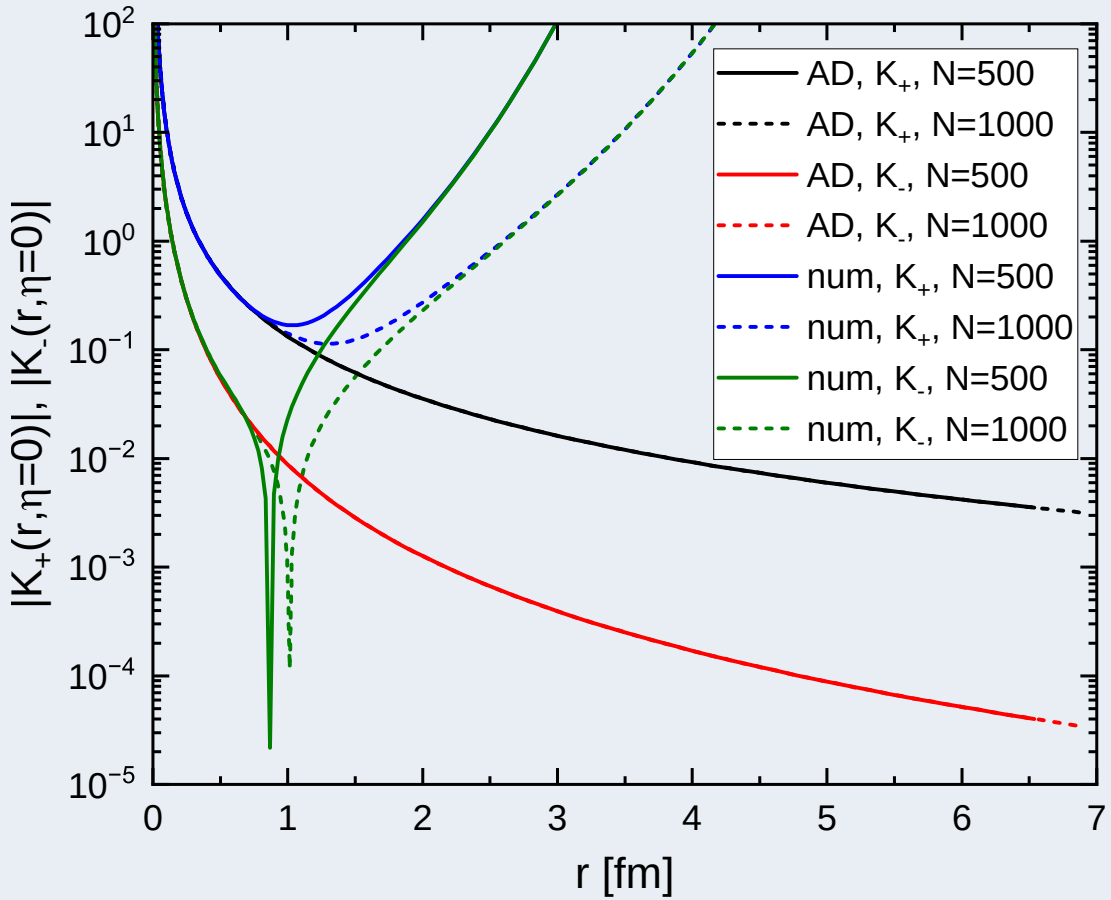
$$\mathscr{F}_{gg}^{(3)}(\boldsymbol{k}_\perp, x) = \frac{C_F}{2\pi^2}\int \frac{r_\perp \mathrm{d}r_\perp}{2\pi} J_0(k_\perp r_\perp)\mathscr{K}(r_\perp, x)\left[1 - \left(S(r_\perp, x)\right)^{N_c/C_F}\right] \times \times \left(S(r_\perp, x)\right)^2$$

$$\mathscr{F}_{WW}(\boldsymbol{k}_\perp, x) = \frac{C_F}{2\pi^2}\int \frac{r_\perp \mathrm{d}r_\perp}{2\pi} J_0(k_\perp r_\perp)\mathscr{K}(r_\perp, x)\left[1 - \left(S(r_\perp, x)\right)^{N_c/C_F}\right]$$

$$\nabla_\perp^2 = \frac{\partial^2}{\partial r_\perp^2} + \frac{1}{r_\perp}\frac{\partial}{\partial r_\perp}$$

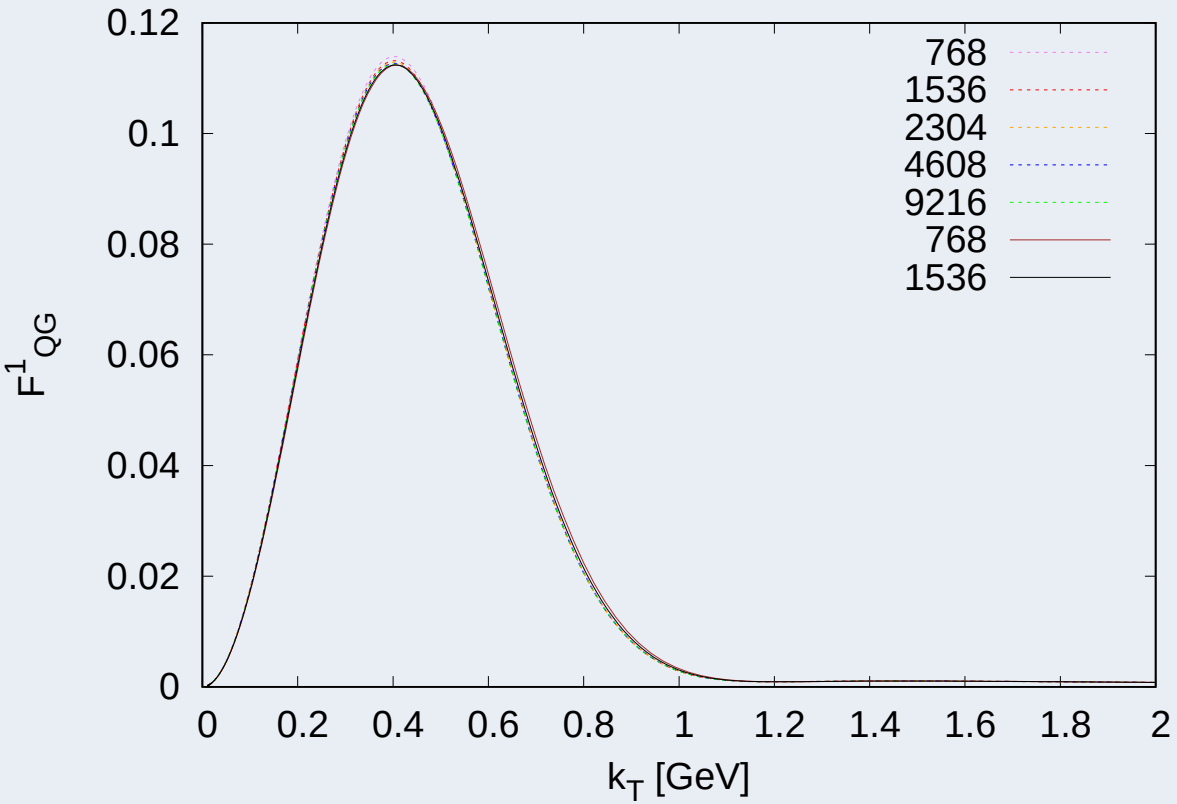
$$\mathscr{K}(r_\perp, x) = \frac{\nabla_\perp^2 \Gamma(r_\perp, x)}{\Gamma(r_\perp, x)} \quad \text{where} \quad \Gamma(r_\perp, x) = -\log\left[S(r_\perp, x)\right]$$

Back to Automatic Differentiation again



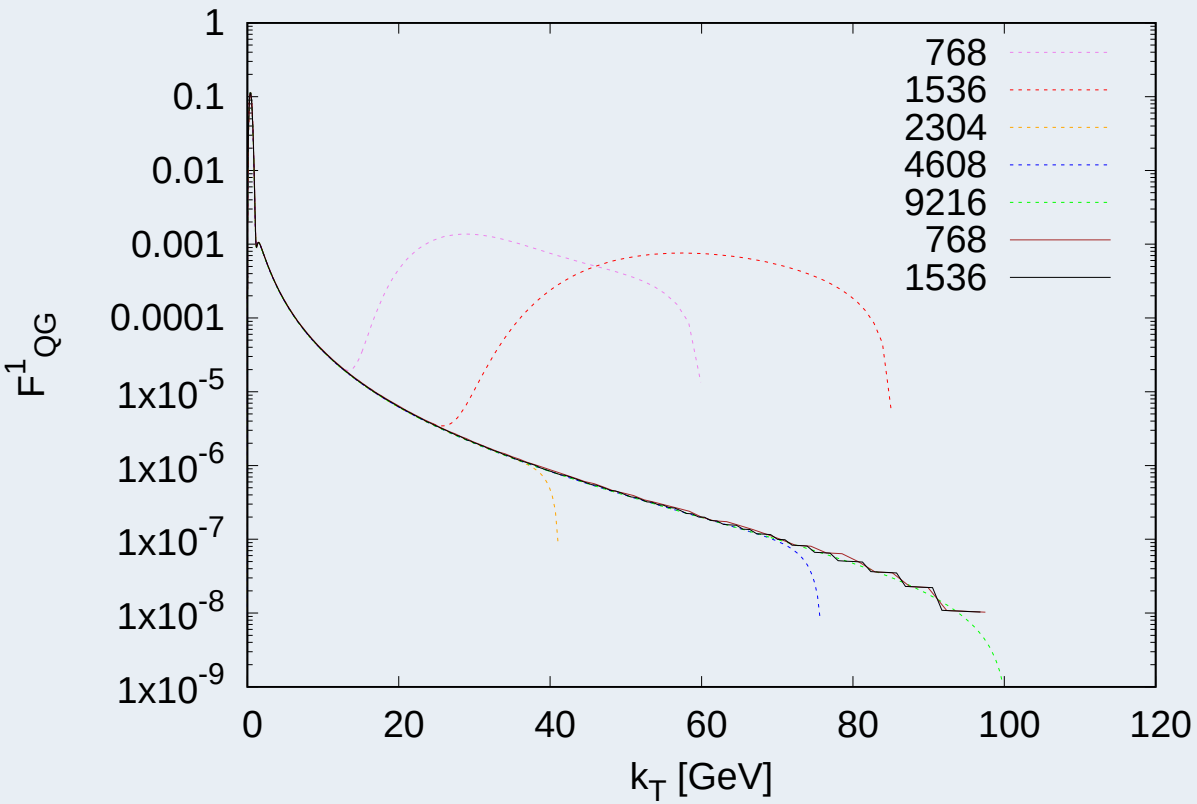
Functions $|\mathscr{K}_+(r_\perp)|$ and $|\mathscr{K}_-(r_\perp)|$ at initial condition $\eta = 0$

Automatic Differentiation combined with LogFFT



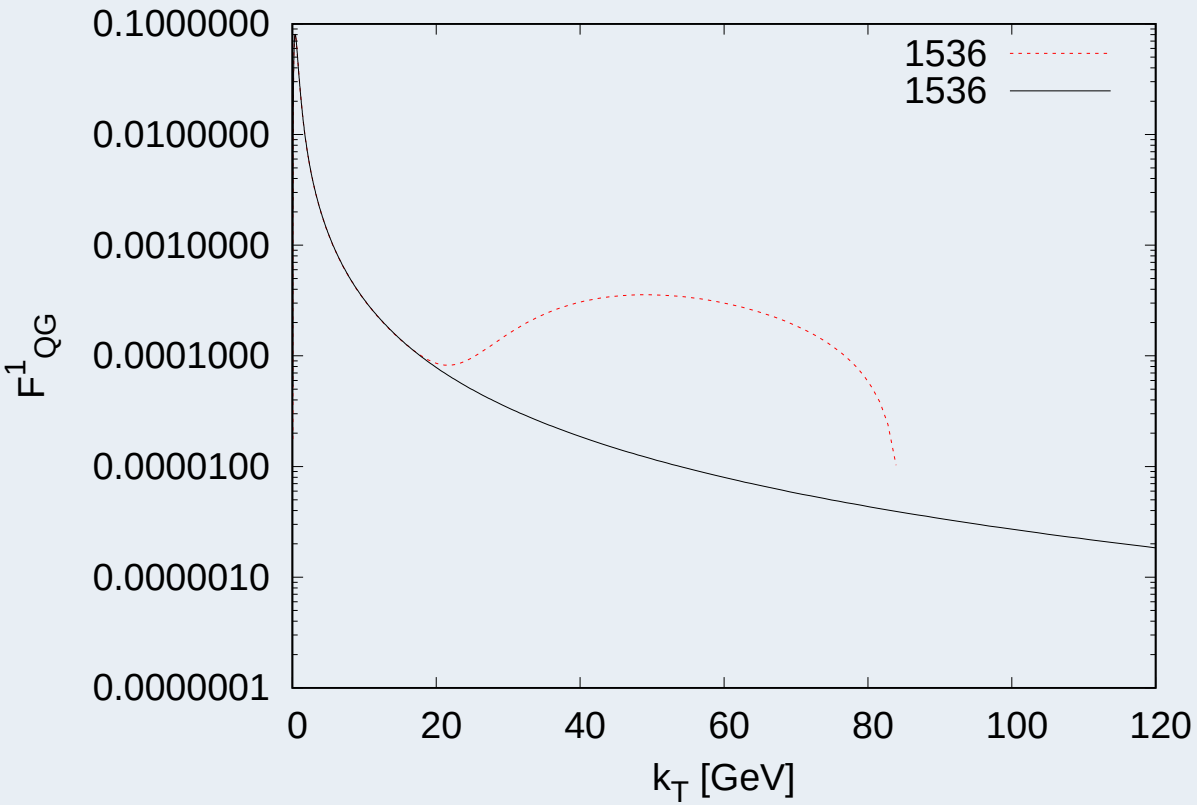
$F^1_{qg}(\eta = 0, k_\perp)$ in momentum space at initial condition $\eta = 0$ for different grid sizes

Automatic Differentiation combined with LogFFT



$F^1_{qg}(\eta = 0, k_\perp)$ in momentum space at initial condition $\eta = 0$ for different grid sizes

Automatic Differentiation combined with LogFFT



$F^1_{qg}(\eta = 3.0, k_\perp)$ in momentum space after BK evolution at $\eta = 3.0$ for different grid sizes

Next steps

- uncertainty analysis and model selection: Bayesian analysis based on the calculation of evidence; comparison of uncertainties from the Hessian method, Markov Chain Monte Carlo, and Nested Sampling algorithms
- testing the stability: impact of different running coupling prescriptions, different implementations of the kinematical constraint
- inclusion of other data/cross-section
- TMD functions from JIMWLK

JIMWLK evolution equation

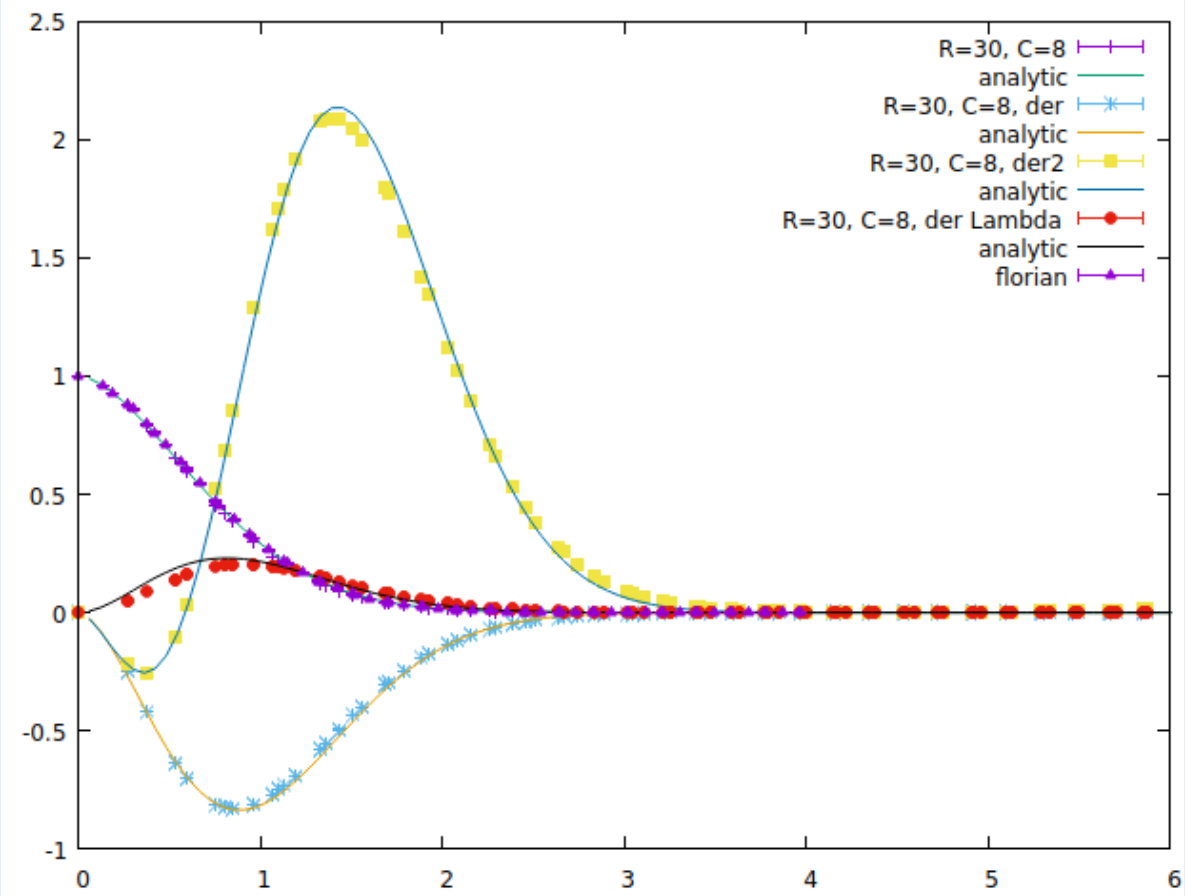
Basic facts

- JIMWLK equation describes the non-linear small-x evolution
- it uses Wilson lines as fundamental degrees of freedom
- two-point correlation function $\langle U^\dagger(x)U(y)\rangle$ gives the dipole amplitude
- two-point correlation functions with derivatives provide a basis for small-x TMD structure functions
- initial condition corresponds to a configuration of Wilson lines
- numerically useful reformulation as a Langevin equation

LO JIMWLK: Langevin formulation (Rummukainen, Weigert 2004, Lappi, Mantysaari 2014)

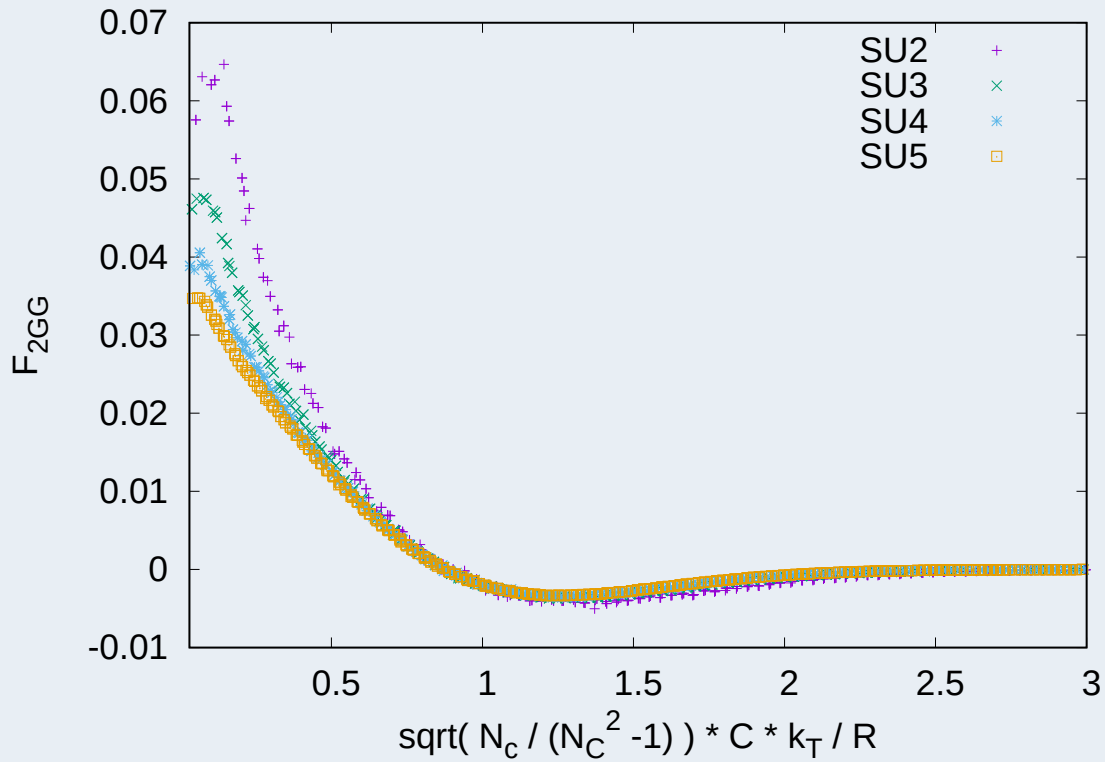
$$U(\mathrm{x}, s+\delta s) = \exp\left(-\sqrt{\delta s}\sum_{\mathrm{y}} U(\mathrm{y},s)(\mathrm{K}(\mathrm{x}-\mathrm{y})\cdot\xi(\mathrm{y}))U^\dagger(\mathrm{y},s)\right)\times$$
$$\times\, U(\mathrm{x},s)\times\exp\left(\sqrt{\delta s}\sum_{\mathrm{y}}\mathrm{K}(\mathrm{x}-\mathrm{y})\cdot\xi(\mathrm{y})\right).$$

Automatic Differentiation for JIMWLK



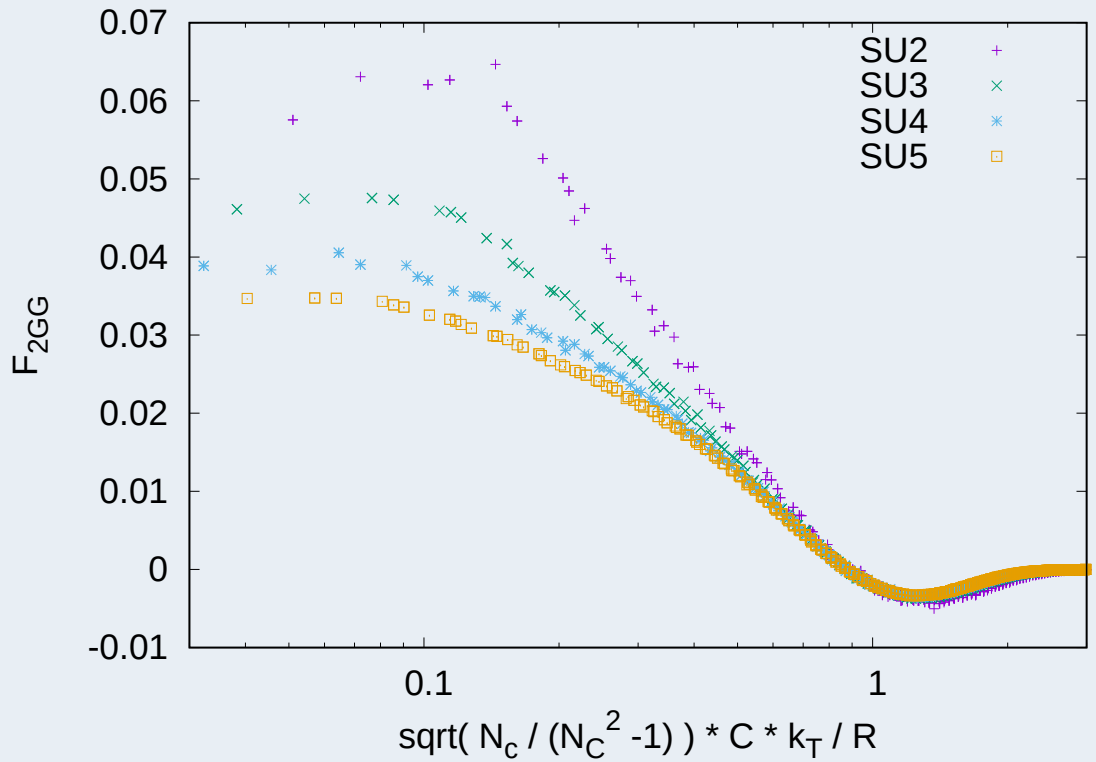
Gluon dipole amplitude obtained from JIMWLK, together with the first and second derivatives with respect to Q_0

SU(N) JIMWLK at $\eta = 0$



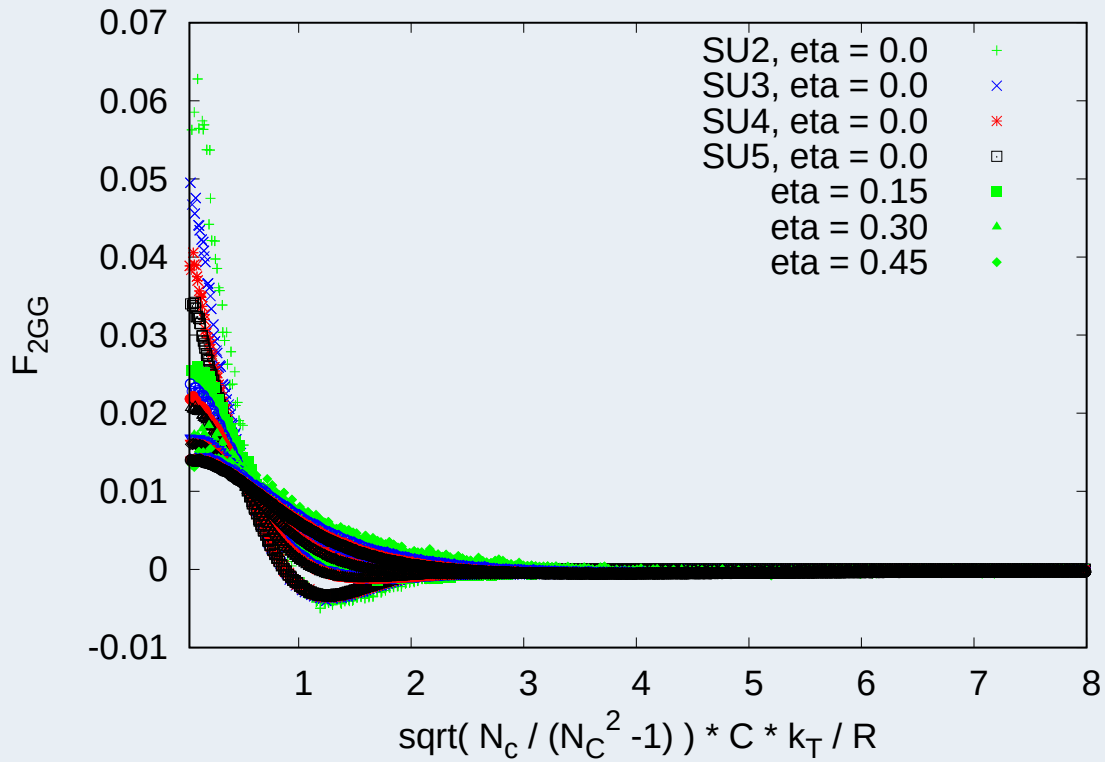
F^2_{gg} TMD structure function at the initial condition in the MV model.

SU(N) JIMWLK at $\eta = 0$



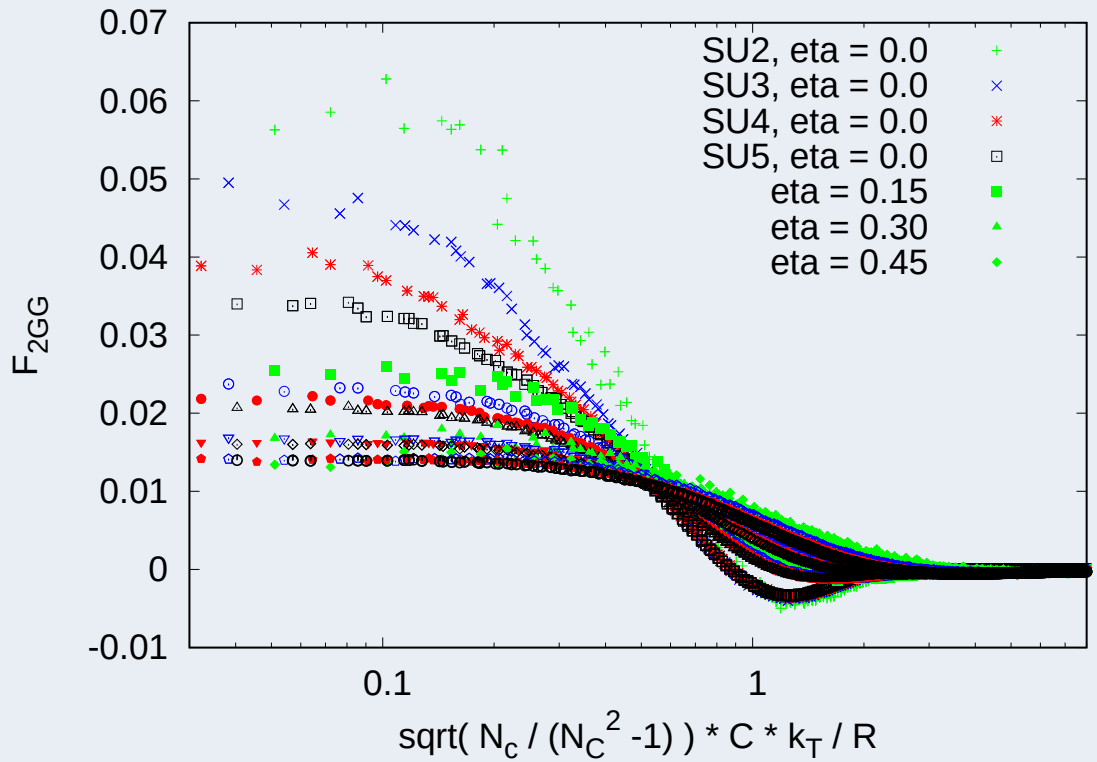
F_{gg}^2 TMD structure function at the initial condition in the MV model.

SU(N) JIMWLK with evolution



F_{gg}^2 TMD structure function evolved with JIMWLK in the MV model.

SU(N) JIMWLK with evolution



F_{gg}^2 TMD structure function evolved with JIMWLK in the MV model.

Summary

- I have presented elements of the framework that allow for the efficient fitting of several observables
- I have discussed the benefits of using automatic differentiation
- I have shown how to increase the performance by employing the logarithmic Fourier transform
- I have presented preliminary results of the fit to the DIS from HERA and single inclusive hadron production from BRAHMS
- I have highlighted future steps

Thank you very much for your attention!