JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Towards indirect measurement of the gluonic structure of the proton

Jyväskylä Inverse Problems seminar

Henri Hänninen

November 18, 2025

Based on joint work with

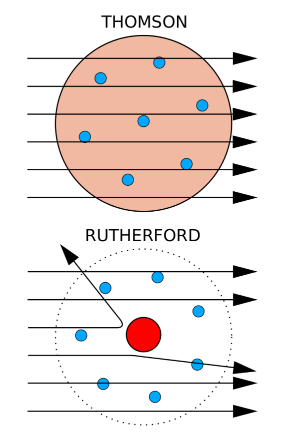
Antti Kykkänen & Hjørdis Schlüter

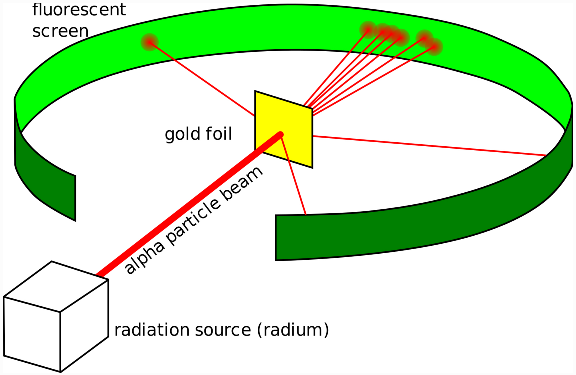
Phys.Rev.D 112 (2025) 9, 094026, arXiv:2509.05005 [hep-ph]

A brief history of the structure of the atom and proton

- Rutherford discovers the nucleus (1911)
- Rutherford discovers the proton (1919)
- Shrödinger and quantum mechanics: Hydrogen atom is understood as a quantum system. (1920s and 1930s)
- Particle zoo of hadronic particles and postulation of a new force and elementary particles: quarks (1960s)
- Proton's internal structure (1970–present)
- What we still don't know and *how inverse problems can help*

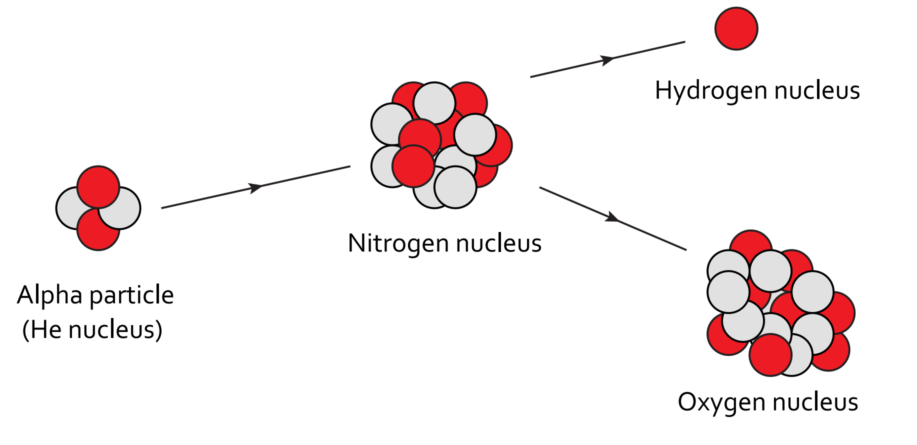
The gold foil experiment





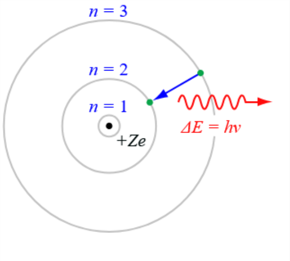
[4] E. Rutherford, "The Scattering of α and β Particles by Matter and the Structure of the Atom," Philos. Mag. 21, 669 (1911). From stanford.edu (link).

Discovery of the proton



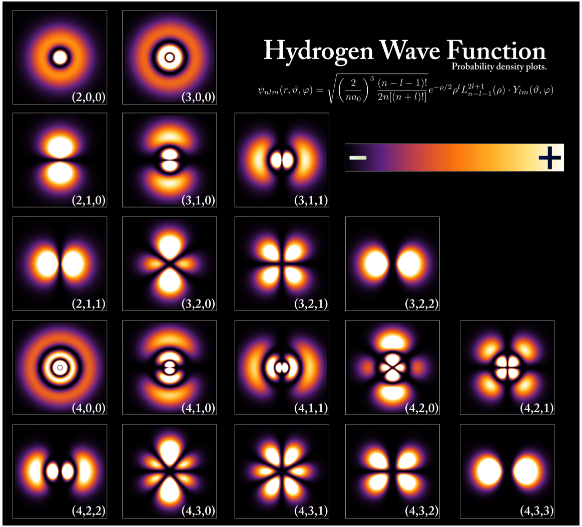
Rutherford, E. (1919-06-01). "LIV. Collision of α particles with light atoms. IV. An anomalous effect in nitrogen". The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 37 (222): 581–587. doi:10.1080/14786440608635919. ISSN 1941-5982.

Towards a quantum model of the hydrogen atom

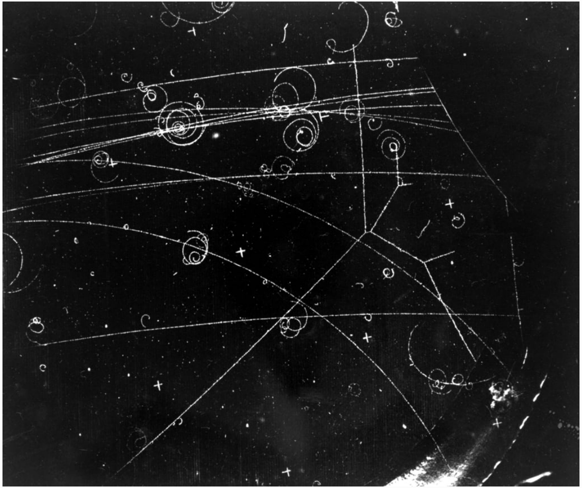


Bohr model of the atom 1910s.

Shrödinger equation and the model of a hydrogen atom



Particle physics in the 1950s and 1960s



Cloud chamber proton–proton scattering event, photograph of the experiment, 1957.

The ”particle zoo“

An incredible multitude of particles that look like the proton or neutron are seen in the cloud chamber experiments (with varying electric charges, masses, and quantum spins):

Baryon resonance particles

Nucleons			Δ particles			Λ particles			Σ particles			Ξ and Ω particles			Charmed particles			Bottom particles		
p	$1/2^+$	****	Δ(1232)	$3/2^+$	****	Λ	$1/2^+$	****	Σ^+	$1/2^+$	****	Ξ^0	$1/2^+$	****	Λ_c^+	$1/2^+$	****	Λ_b^0	$1/2^+$	***
n	$1/2^+$	****	Δ(1600)	$3/2^+$	****	Λ(1405)	$1/2^-$	****	Σ^0	$1/2^+$	****	Ξ^-	$1/2^+$	****	$\Lambda_c(2595)^+$	$1/2^-$	***	$\Lambda_b(5912)^0$	$1/2^-$	***
N(1440)	$1/2^+$	****	Δ(1620)	$1/2^-$	****	Λ(1520)	$3/2^-$	****	Σ^-	$1/2^+$	****	Ξ(1530)	$3/2^+$	****	$\Lambda_c(2625)^+$	$3/2^-$	***	$\Lambda_b(5920)^0$	$3/2^-$	***
N(1520)	$3/2^-$	****	Δ(1700)	$3/2^-$	****	Λ(1600)	$1/2^+$	***	Σ(1385)	$3/2^+$	****	Ξ(1620)		*	$\Lambda_c(2765)^+$		*	Σ_b	$1/2^+$	***
N(1535)	$1/2^-$	****	Δ(1750)	$1/2^+$	*	Λ(1670)	$1/2^-$	****	Σ(1480)		*	Ξ(1690)		***	$\Lambda_c(2860)^+$	$3/2^+$	***	Σ_b^*	$3/2^+$	***
N(1650)	$1/2^-$	****	Δ(1900)	$1/2^-$	***	Λ(1690)	$3/2^-$	****	Σ(1560)		**	Ξ(1820)	$3/2^-$	***	$\Lambda_c(2880)^+$	$5/2^+$	***	Ξ_b^0, Ξ_b^-	$1/2^+$	***
N(1675)	$5/2^-$	****	Δ(1905)	$5/2^+$	****	Λ(1710)	$1/2^+$	*	Σ(1580)	$3/2^-$	*	Ξ(1950)		***	$\Lambda_c(2940)^+$	$3/2^-$	***	$\Xi'_b(5935)^-$	$1/2^+$	***
N(1680)	$5/2^+$	****	Δ(1910)	$1/2^+$	****	Λ(1800)	$1/2^-$	***	Σ(1620)	$1/2^-$	*	Ξ(2030)	$\geq \frac{5}{2}$?	***	$\Sigma_c(2455)$	$1/2^+$	****	$\Xi_b(5945)^0$	$3/2^+$	***
N(1700)	$3/2^-$	***	Δ(1920)	$3/2^+$	***	Λ(1810)	$1/2^+$	***	Σ(1660)	$1/2^+$	***	Ξ(2120)		*	$\Sigma_c(2520)$	$3/2^+$	***	$\Xi_b(5955)^-$	$3/2^+$	***
N(1710)	$1/2^+$	****	Δ(1930)	$5/2^-$	***	Λ(1820)	$5/2^+$	****	Σ(1670)	$3/2^-$	****	Ξ(2250)		**	$\Sigma_c(2800)$		***	Ω_b^-	$1/2^+$	***
N(1720)	$3/2^+$	****	Δ(1940)	$3/2^-$	**	Λ(1830)	$5/2^-$	****	Σ(1690)		**	Ξ(2370)		**				$P_c(4380)^+$		*
N(1860)	$5/2^+$	**	Δ(1950)	$7/2^+$	****	Λ(1890)	$3/2^+$	****	Σ(1730)	$3/2^+$	*	Ξ(2500)		*	Ξ_c^+	$1/2^+$	***	$P_c(4450)^+$		*
N(1875)	$3/2^-$	***	Δ(2000)	$5/2^+$	**	Λ(2000)		*	Σ(1750)	$1/2^-$	***				Ξ_c^0	$1/2^+$	***			
N(1880)	$1/2^+$	***	Δ(2150)	$1/2^-$	*	Λ(2020)	$7/2^+$	*	Σ(1770)	$1/2^+$	*	Ω^-	$3/2^+$	****	$\Xi_c'^+$	$1/2^+$	***			
N(1895)	$1/2^-$	****	Δ(2200)	$7/2^-$	***	Λ(2050)	$3/2^-$	*	Σ(1775)	$5/2^-$	****	$\Omega(2250)^-$		***	$\Xi_c'^0$	$1/2^+$	***			
N(1900)	$3/2^+$	****	Δ(2300)	$9/2^+$	**	Λ(2100)	$7/2^-$	****	Σ(1840)	$3/2^+$	*	$\Omega(2380)^-$		**	$\Xi_c(2645)$	$3/2^+$	***			
N(1990)	$7/2^+$	**	Δ(2350)	$5/2^-$	*	Λ(2110)	$5/2^+$	***	Σ(1880)	$1/2^+$	**	$\Omega(2470)^-$		**	$\Xi_c(2790)$	$1/2^-$	***			
N(2000)	$5/2^+$	**	Δ(2390)	$7/2^+$	*	Λ(2325)	$3/2^-$	*	Σ(1900)	$1/2^-$	*				$\Xi_c(2815)$	$3/2^-$	***			
N(2040)	$3/2^+$	*	Δ(2400)	$9/2^-$	**	Λ(2350)	$9/2^+$	***	Σ(1915)	$5/2^+$	****				$\Xi_c(2930)$		*			
N(2060)	$5/2^-$	***	Δ(2420)	$11/2^+$	****	Λ(2585)		**	Σ(1940)	$3/2^+$	*				$\Xi_c(2980)$		***			
N(2100)	$1/2^+$	***	Δ(2750)	$13/2^-$	**				Σ(1940)	$3/2^-$	***				$\Xi_c(3055)$		***			
N(2120)	$3/2^-$	***	Δ(2950)	$15/2^+$	**				Σ(2000)	$1/2^-$	*				$\Xi_c(3080)$		***			
N(2190)	$7/2^-$	****							Σ(2030)	$7/2^+$	****				$\Xi_c(3123)$		*			
N(2220)	$9/2^+$	****							Σ(2070)	$5/2^+$	*									
N(2250)	$9/2^-$	****							Σ(2080)	$3/2^+$	**				Ω_c^0	$1/2^+$	***			
N(2300)	$1/2^+$	**							Σ(2100)	$7/2^-$	*				$\Omega_c(2770)^0$	$3/2^+$	***			
N(2570)	$5/2^-$	**							Σ(2250)		***				$\Omega_c(3000)^0$		***			
N(2600)	$11/2^-$	***							Σ(2455)		**				$\Omega_c(3050)^0$		***			
N(2700)	$13/2^+$	**							Σ(2620)		**				$\Omega_c(3065)^0$		***			
									Σ(3000)		*				$\Omega_c(3090)^0$		***			
															$\Omega_c(3120)^0$		***			
															Ξ_{cc}^+		*			
															Ξ_{cc}^{++}		***			

**** Existence is certain, and properties are at least fairly well explored.

*** Existence ranges from fairly certain to certain, but further confirmation is desirable, and/or quantum numbers, branching fractions, etc. are not well determined.

** Evidence of existence is only fair.

* Evidence of existence is poor.

Nomenclature of flavourless mesons

$q\bar{q}$ content	I	J^{PC} [i]			
		$0^{-+}, 2^{-+}, 4^{-+}, \ldots$	$1^{+-}, 3^{+-}, 5^{+-}, \ldots$	$1^{--}, 2^{--}, 3^{--}, \ldots$	$0^{++}, 1^{++}, 2^{++}, \ldots$
$u\bar{d}$ $\frac{u\bar{u}-d\bar{d}}{\sqrt{2}}$ $d\bar{u}$	1	π^+ π^0 π^-	b^+ b^0 b^-	ρ^+ ρ^0 ρ^-	a^+ a^0 a^-
Mix of $u\bar{u}$, $d\bar{d}$, $s\bar{s}$	0	η η'	h h'	ω ϕ	f f'
$c\bar{c}$	0	η_c	h_c	ψ [ii]	χ_c
$b\bar{b}$	0	η_b	h_b	Υ	χ_b
$t\bar{t}$	0	η_t	h_t	θ	χ_t

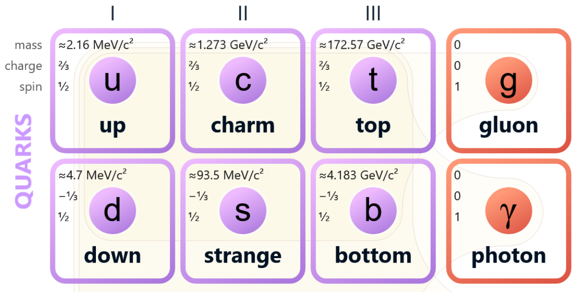
i. ^ C-parity is only relevant for neutral mesons.

ii. ^ For the special case $J^{PC} = 1^{--}$, the ψ is called the J/ψ

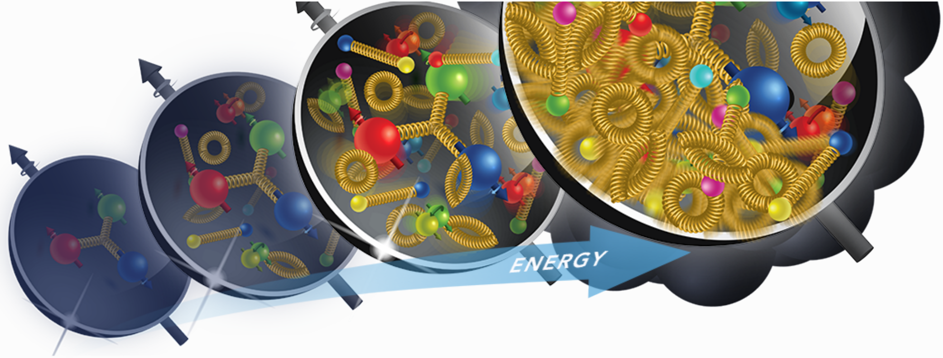
Nomenclature of flavoured mesons

Quark	Antiquark					
	up	down	charm	strange	top	bottom
up	—	[i]	$\bar{D}^0$	K^+	$\bar{T}^0$	B^+
down	[i]	—	D^-	K^0	T^-	B^0
charm	D^0	D^+	—	D_s^+	$\bar{T}_c^0$	B_c^+
strange	K^-	$\bar{K}^0$	D_s^-	—	T_s^-	B_s^0
top	T^0	T^+	T_c^0	T_s^+	—	T_b^+
bottom	B^-	$\bar{B}^0$	B_c^-	$\bar{B}_s^0$	T_b^-	—

”All of these dozens and dozens of particles cannot possibly be elementary (non-divisible)“
As an answer to this quandary, the existence of a new elementary force and particles were postulated, which lead to the quark model, and ultimately to quantum chromodynamics as a part of the Standard Model.



Proton is composed of quarks and gluons (1970–)



Parton distribution functions of the proton (Present day)

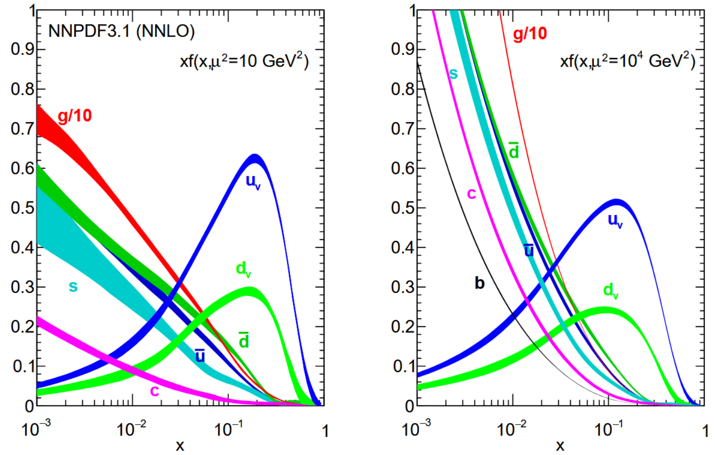


Figure 3.1: The NNPDF3.1 NNLO PDFs, evaluated at $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2 (right).

NNPDF collaboration, `https://doi.org/10.1140/epjc/s10052-017-5199-5` (2017)

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.
  - "*Proton spin puzzle*"

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.
  - "*Proton spin puzzle*"
- Quantification of quark and gluon densities in the proton

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.
  - "*Proton spin puzzle*"
- Quantification of quark and gluon densities in the proton
- Transverse density profile of the proton in a collision

On-going questions about the proton structure and how inverse problems can help?

- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.
  - "*Proton spin puzzle*"
- Quantification of quark and gluon densities in the proton
- Transverse density profile of the proton in a collision
- Gluonic structure and saturation within the proton

On-going questions about the proton structure and how inverse problems can help?

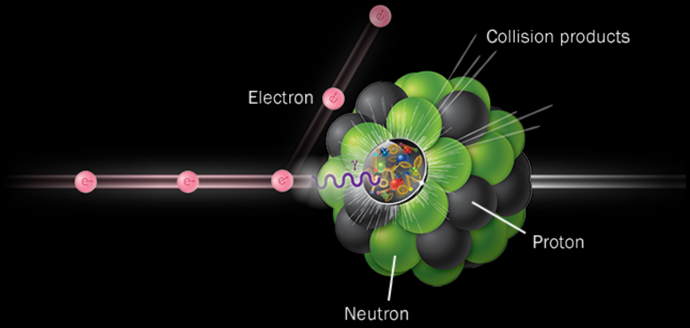
- Mass composition ($m_p = 1.67262192 \times 10^{-27}$ kg)
  - The uud quark masses make up 1% of the proton mass.
  - All the rest is potential energy of the strong nuclear force.
    - (Quark condensate $\sim 9\%$, Q energy $\sim 32\%$, gluon field $\sim 37\%$, anomalous gluonic contr. $\sim 23\%$)
- Proton has spin $\frac{1}{2}$
  - The uud quarks also have spin $\frac{1}{2}$
  - Naive guess based on quantum mechanics: uud alone produce the proton spin.
  - Estimate: quarks contribute $\sim 30 - 40\%$, gluons contribute $\sim 26 - 52\%$ of the spin.
  - "*Proton spin puzzle*"
- Quantification of quark and gluon densities in the proton
- Transverse density profile of the proton in a collision
- Gluonic structure and saturation within the proton
- Low-energy structure / quantum wavefunction of the proton

$\therefore$ Model agnostic indirect measurement techniques enabled by Inverse Problems to the rescue!

Idea: If you can't calculate it, measure it.

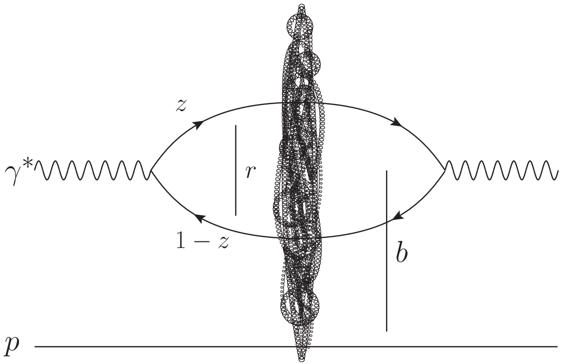
Deep inelastic scattering

- Proton and nucleus structures are studied in high-energy collisions.
- The same strong nuclear force that binds the quarks in the proton, also binds the protons and neutrons together in the atomic nucleus.
- Electron as an elementary particle is a nice clean probe to study the target; interacts with the target only electro-magnetically.



Deep inelastic scattering in the dipole picture

- The electro-magnetic interaction between the electron and target is mediated by a ”virtual photon“, γ^*, quantum state.
- Target proton is represented by a dense cloud of gluons in the dipole picture. (Probing the gluonic structure at high-energy)
- γ^* state has to fluctuate into a quantum state with color charge: quark-antiquark pair (*dipole*), which then interacts and scatters off the gluon cloud.
- This $q\bar{q}$–target scattering is theoretically described by a quantity known as the dipole amplitude $N(r, x)$.



r is the dist. btw. quark and antiquark. Impact parameter b is the separation between the interaction point and the center of the proton. z is the fraction of how momentum is distributed to q and $\bar{q}$.

End of introduction

Inverse Problems seminar introduction ends here, and the QCD seminar talk contents start.

Motivation

Reconstruction of the Dipole Amplitude in the Dipole Picture as a mathematical Inverse Problem, Phys. Rev. D (2025).

Abstract

We show that the inference problem of constraining the dipole amplitude with inclusive deep inelastic scattering data can be written into a discrete **linear inverse problem**, in an analogous manner as **can be done for computed tomography**. To this formulation of the problem, we apply standard inverse problems methods and algorithms to **reconstruct known dipole amplitudes from simulated reduced cross section data** with realistic precision. The main difference of this approach to previous works is that this implementation does not require any fit parametrization of the dipole amplitude. The **freedom from parametrization** also enables us for the first time to quantify the uncertainties of the inferred dipole amplitude in a novel more general framework. This mathematical approach to small-x phenomenology opens a path to parametrization bias free inference of the dipole amplitude from HERA and Electron–Ion Collider data.

A motivating prelude: diffractive meson electroproduction

S. Munier, A.M. Stasto, A.H. Mueller wanted to extract the dipole amplitude from vector meson production data without an assumed model parametrization [1]:

$$\int \frac{\mathrm{d}^2\Delta}{(2\pi)^2}\sqrt{\frac{\mathrm{d}\sigma}{\mathrm{d}t}}e^{-i\Delta\mathbf{b}} = \frac{1}{\sqrt{16\pi}}\sum_{h,\bar{h}}\int \mathrm{d}^2\mathbf{r}\mathrm{d}r\psi_{\gamma^*}^{h,\bar{h}}(z,r,Q)2[1-S(x,r,b)]\psi_V^{h,\bar{h}}(z,\mathbf{r}),$$

which, after defining a probability measure p using the light-cone wavefunctions and its normalization $N(Q)$, they invert into

$$\langle S(x,r,b)\rangle_p = 1 - \frac{1}{2N(Q)\pi^{3/2}}\int \mathrm{d}^2\Delta e^{-i\Delta\mathbf{b}}\sqrt{\frac{\mathrm{d}\sigma}{\mathrm{d}t}}.$$

This enabled the extraction of information about the r-averaged dipole amplitude from HERA data.

”Can one do something if the integral relation is not (known to be) an analytically invertible?“

[1] Nuclear Physics B Volume 603, Issues 1–2, 4 June 2001, Pages 427-445; https://doi.org/10.1016/S0550-3213(01)00168-7

Nomenclature: forward vs inverse problem

We will consider a linear equation:

$$\mathbf{b} = \mathbf{A}\mathbf{x},$$

where $\mathbf{b}$ is data acquired from an experiment, $\mathbf{A}$ is a so-called forward operator, here matrix valued, and $\mathbf{x}$ is the unknown we want to solve from the data (to indirectly measure).

Nomenclature: forward vs inverse problem

We will consider a linear equation:

$$\mathbf{b} = \mathbf{A}\mathbf{x},$$

where $\mathbf{b}$ is data acquired from an experiment, $\mathbf{A}$ is a so-called forward operator, here matrix valued, and $\mathbf{x}$ is the unknown we want to solve from the data (to indirectly measure).

- Forward problem $\hat{=}$ physical theory or model that enables calculation of predictions of data if the quantity $\mathbf{x}$ is known.
- Inverse problem $\hat{=}$ solving for $\mathbf{x}$ from a measurement $\mathbf{b}$.

Nomenclature: forward vs inverse problem

We will consider a linear equation:

$$\mathbf{b} = \mathbf{A}\mathbf{x},$$

where $\mathbf{b}$ is data acquired from an experiment, $\mathbf{A}$ is a so-called forward operator, here matrix valued, and $\mathbf{x}$ is the unknown we want to solve from the data (to indirectly measure).

- Forward problem $\hat{=}$ physical theory or model that enables calculation of predictions of data if the quantity $\mathbf{x}$ is known.
- Inverse problem $\hat{=}$ solving for $\mathbf{x}$ from a measurement $\mathbf{b}$.

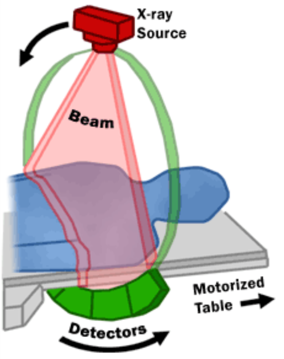
We will write DIS in the dipole picture in this form, where the data is $\sigma_r(Q, x)$, and the unknown is the dipole amplitude $N(r, x)$.

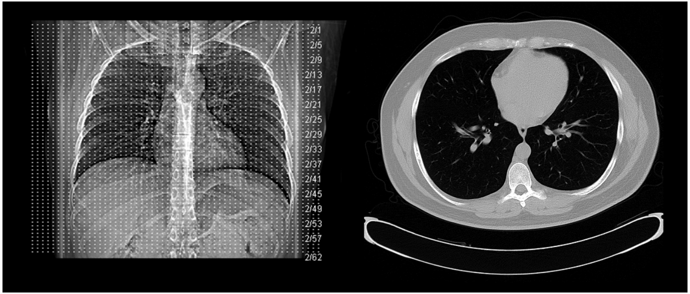
What do you mean computed tomography has something in common with deep inelastic scattering??

- It's *not* the physics (X-ray attenuation in the target)
- It's *not* the experiment (CT scanner, detector rotating around the target)
- It's *not* the mathematical relationship (Radon transform)

We need to look at CT to know how to throw out all of the above, and carefully replace them with the dipole picture of deep inelastic scattering.

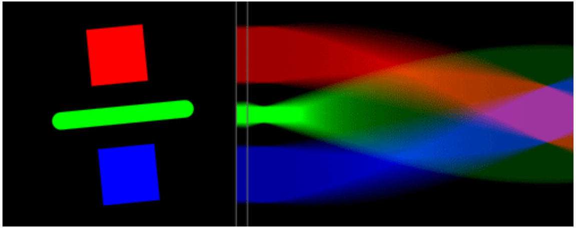
An instructive example: computed tomography





Tomographic reconstruction in medical imaging. We are interested about the mathematical theory, however.

CT measurement is called the *sinogram*



- X-ray source from the left exposes the target which is captured by the sensor on the right.
- The source–sensor pair is rotated around the target to produce the *sinogram* in the right panel.
- The mathematical inverse problem in CT is the recovery of the target structure from the measured sinogram.

Sinogram is the *Radon transform* of the target

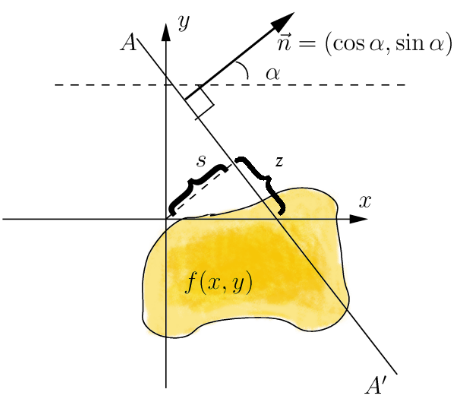
The *Radon transform* Rf is a function defined on the space of straight lines $L \subset \mathbb{R}^2$ by the line integral along each line as:

$$Rf(L) = \int_L f(x)|\mathrm{d}\mathbf{x}|.$$

Parametrizing the lines with path length z:

$$(x(z), y(z)) = \big((z \sin\alpha + s\cos\alpha), (-z\cos\alpha + s\sin\alpha)\big)$$

one has $Rf(\alpha, s) = \int_{-\infty}^{\infty} f(x(z), y(z))\mathrm{d}z$.



Sinogram is the *Radon transform* of the target

The *Radon transform* Rf is a function defined on the space of straight lines $L \subset \mathbb{R}^2$ by the line integral along each line as:

$$Rf(L) = \int_L f(x)|\mathrm{d}\mathbf{x}|.$$

Parametrizing the lines with path length z:

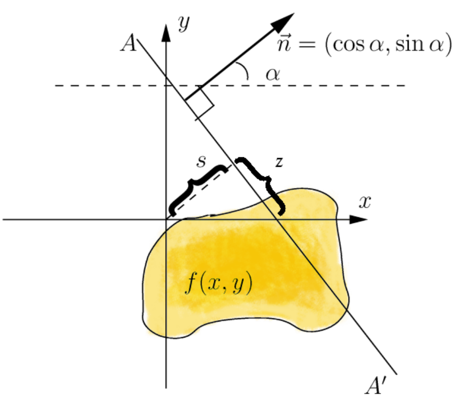
$$(x(z), y(z)) = \big((z \sin\alpha + s\cos\alpha), (-z\cos\alpha + s\sin\alpha)\big)$$

one has $Rf(\alpha, s) = \int_{-\infty}^{\infty} f(x(z), y(z))\mathrm{d}z$.

It is analytically invertible:

$$f(\mathbf{x}) = \int_0^\pi (Rf(\cdot,\theta) * h)\langle \mathbf{x}, \mathbf{n}_\theta\rangle)\mathrm{d}\theta,$$

with the convolution kernel h such that $\hat{h}(k) = |k|$.



Sinogram is the *Radon transform* of the target

The *Radon transform* Rf is a function defined on the space of straight lines $L \subset \mathbb{R}^2$ by the line integral along each line as:

$$Rf(L) = \int_L f(x)|\mathrm{d}\mathbf{x}|.$$

Parametrizing the lines with path length z:

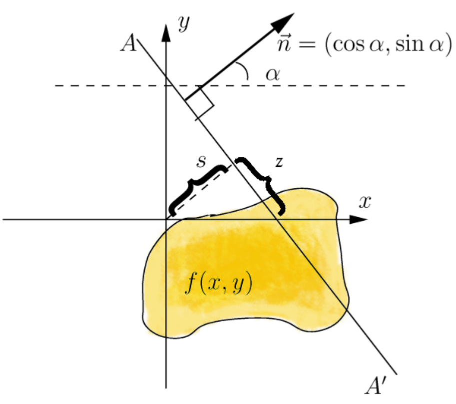
$$(x(z), y(z)) = ((z \sin\alpha + s\cos\alpha), (-z\cos\alpha + s\sin\alpha))$$

one has $Rf(\alpha, s) = \int_{-\infty}^{\infty} f(x(z), y(z))\mathrm{d}z$.

It is analytically invertible:

$$f(\mathbf{x}) = \int_0^{\pi} (Rf(\cdot, \theta) * h)\langle \mathbf{x}, \mathbf{n}_\theta \rangle)\mathrm{d}\theta,$$

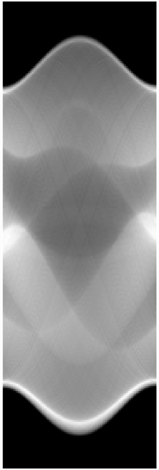
with the convolution kernel h such that $\hat{h}(k) = |k|$.



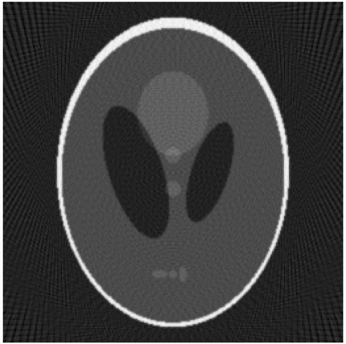
How do we apply this approach to our benefit in QCD?

Radon transform of the Shepp–Logan phantom

Shepp–Logan phatom: a standard test image.



Radon transform of the phantom.



Inverse Radon transform (reconstruction) from the sinogram.

Back to high-energy physics

Well how do we do anything remotely similar with the dipole picture of DIS?

Reduced cross section data to constrain the dipole amplitude

Reduced cross section σ_r is defined with the proton structure functions as:

$$\sigma_r(y, x, Q^2) = F_T(x, Q^2) + \frac{2(1-y)}{1+(1-y)^2} F_L(x, Q^2)$$

$$F_{L,T}(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\mathrm{em}}} \sigma_{L,T}(x, Q^2)$$

$$y = \frac{Q^2}{xs}$$

$$\sigma_{T,L}(x, Q) = \frac{\sigma_0}{2} \frac{1}{4\pi} \sum_f \int_{\mathbb{R}^2} \int_0^1 \left| \psi_{T,L}^{\gamma^* \to q_f \bar{q}_f}(\mathbf{r}, Q^2, z, f) \right|^2 N(\mathbf{r}, x) \mathrm{d}^2\mathbf{r} \mathrm{d}z,$$

which is an implicit inverse problem for N, conventionally solved by fitting a theoretically motivated model parametrization for N.

Mathematical observations

$$\sigma_r(y, x, Q^2) = F_T(x, Q^2) + \frac{2(1-y)}{1+(1-y)^2} F_L(x, Q^2)$$

$$F_{L,T}(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\mathrm{em}}} \sigma_{L,T}(x, Q^2)$$

$$y = \frac{Q^2}{xs}$$

$$\sigma_{T,L}(x, Q) = \frac{\sigma_0}{2} \frac{1}{4\pi} \sum_f \int_{\mathbb{R}^2} \int_0^1 \left| \psi_{T,L}^{\gamma^* \to q_f \bar{q}_f}(\mathbf{r}, Q^2, z, f) \right|^2 N(\mathbf{r}, x) \mathrm{d}^2\mathbf{r} \mathrm{d}z.$$

First simplifying assumptions:

- Keep x fixed (constant) in calculations. Reconstruct $N(x, r)$ at fixed x.
- Take $y := y(x, Q) = \frac{Q^2}{xs}$, i.e. keep s as constant.

Rewriting into an explicit inference problem

Define:

$$Z_{T,L}(r, Q^2) = \sum_f \int_0^1 \left| \psi_{T,L}^{\gamma^* \to q_f \bar{q}_f}(r, Q^2, z, f) \right|^2 \mathrm{d}z$$

and

$$Z(r, Q^2, y) := \frac{rQ^2}{4\pi\alpha_{\mathrm{em}}} \left[Z_T(r, Q^2) + \frac{2(1-y)}{1+(1-y)^2} Z_L(r, Q^2) \right].$$

Then

$$\begin{aligned}\sigma_r(x, Q^2) &= \frac{\sigma_0}{2} \frac{Q^2}{4\pi\alpha_{\mathrm{em}}} \int_0^\infty \left[Z_T(r, Q^2) + \frac{2(1-y)}{1+(1-y)^2} Z_L(r, Q^2) \right] N(r, x) r \mathrm{d}r \\ &= \int_0^\infty Z(r, Q^2, y) \frac{\sigma_0}{2} N(r, x) \mathrm{d}r,\end{aligned}$$

With this the reduced cross section is now an **integral transform** of the dipole amplitude.

Sidebar: Related integral transforms

The K-transform defined by

$$g(y, \nu) = \mathcal{K}_\nu[f(x); y] = \int_0^\infty (xy)^{\frac{1}{2}} K_\nu(xy) f(x) \mathrm{d}x$$

is analytically invertible. Identifying the measurement with $g(y) \sim \sigma_r(Q)$ and the dipole amplitude with $f(x) \sim N(r)$ bears a rough resemblance to the dipole amplitude inference problem.

Guess: the fact that our approach works could be related to the existence of a proper integral transform?

Discretizing the problem for numerical inversion

Taking r on a grid $\{r_0, r_1, \ldots, r_M\}$

$$\sigma_r^{\mathrm{d}}(x, Q^2) = \sum_{i=0}^{M} (r_{i+1} - r_i) Z(r_i, Q^2, y) \frac{\sigma_0}{2} N(r_i, x),$$

we finally write the problem as a proper linear equation:

$$\sigma_{r,j}^{\mathrm{d}} := \sigma_r^{\mathrm{d}}(Q_j^2, x) = \tilde{Z}_j^i \frac{\sigma_0}{2} N_i,$$

where we absorbed the interval length $(r_{i+1} - r_i)$ into the definition of $\tilde{Z}_j^i$.
This is now an explicit inference problem(?) / reconstruction problem(?), to which we may apply standard methods, such as Tikhonov regularization.
The forward operator $\tilde{Z}_j^i$ is now completely fixed by the (light-cone) perturbation theory, and all non-perturbative degrees of freedom are in $\frac{\sigma_0}{2} N_i$, to be solved by reconstruction.

Solving the inverse problem numerically

Generally the discrete problem $\sigma = \varsigma \mathbf{n}$ has

- $\mathbf{n} \in \mathbb{R}^M$
- $\sigma \in \mathbb{R}^m$
- $\varsigma \in \mathbb{R}^{m \times M}$,

where M is constrained by numerical accuracy of the computation, and m is constrained by the available data. If $m \equiv M$ we would have a full-rank problem (and ς could have a true inverse), but in practice we have a underdetermined problem with $m < M$.

- (i.e. there is fewer data points of σ_r available than we need to have discretization points for $N(r, x)$ to have accurate calculation of the cross sections.)

Regularization methods allow us to solve underdetermined problems.

Numerical reconstruction

We use 1st order derivative operator Tikhonov regularization, where we minimize the cost function

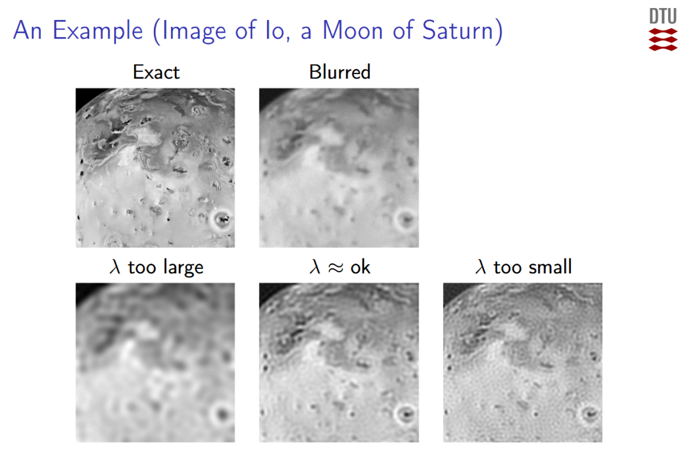
$$\underset{\mathbf{n}\in\mathbb{R}^M}{\arg\min}\,\{||\varsigma\mathbf{n}-\sigma_r||_2^2+\lambda||\frac{\mathrm{d}}{\mathrm{d}r}\mathbf{n}||_2^2\},$$

where λ is a numerical parameter which tunes the size of the penalization of a large value for the first derivative of $\mathbf{n}$.

- Standard Tikhonov regularization would just optimize for $||\mathbf{n}||_2^2$
- 2nd order preconditioned would optimize $||\frac{\mathrm{d}^2}{\mathrm{d}r^2}\mathbf{n}||_2^2$, i.e. limit large fluctuations in the 2nd order derivative of the dipole amplitude.

Implementations of regularization algorithms by AIR Tools II and regtools:
`https://github.com/jakobsj/AIRToolsII`,
`https://www2.compute.dtu.dk/~pcha/Regutools/`

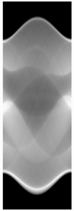
Heuristic idea of the role of the regularization parameter λ

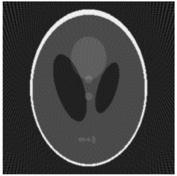


The regularization parameter λ ”tunes“ the sensitivity of the algorithm for noise: too large λ smoothens or blurs too much, and too small λ allows noise to pass through to the reconstructed solution. Choosing the right λ is a critical step, and these standard test images—and this closure test—are used to verify it is chosen as well as possible.

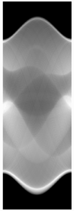
Since with real data we don't have a ”ground truth“ to refer to, we need to implement some unsupervised method, which we are looking into

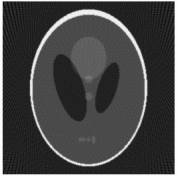
Choosing a "phantom" of the dipole amplitude

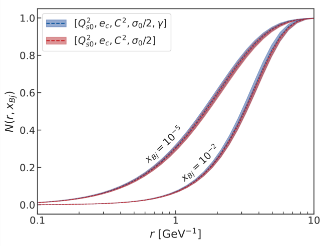




Choosing a ”phantom“ of the dipole amplitude

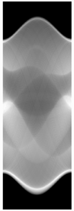


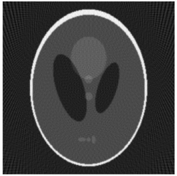


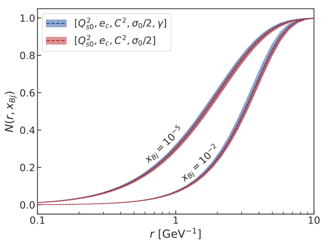


Casuga–Karhunen–Mäntysaari reference dipole

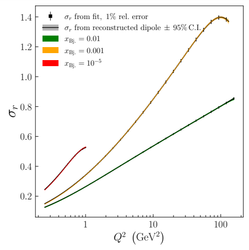
Choosing a ”phantom“ of the dipole amplitude





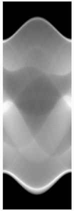


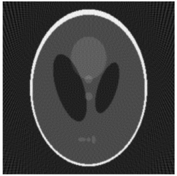
Casuga–Karhunen–Mäntysaari reference dipole

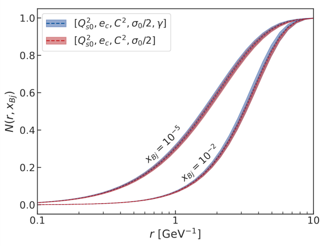


DIS reduced cross section from CKS dipole

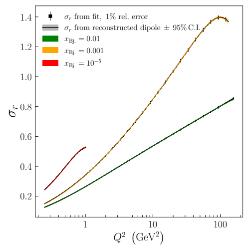
Choosing a ”phantom“ of the dipole amplitude



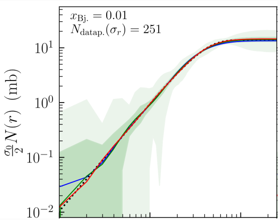




Casuga–Karhunen–Mäntysaari reference dipole

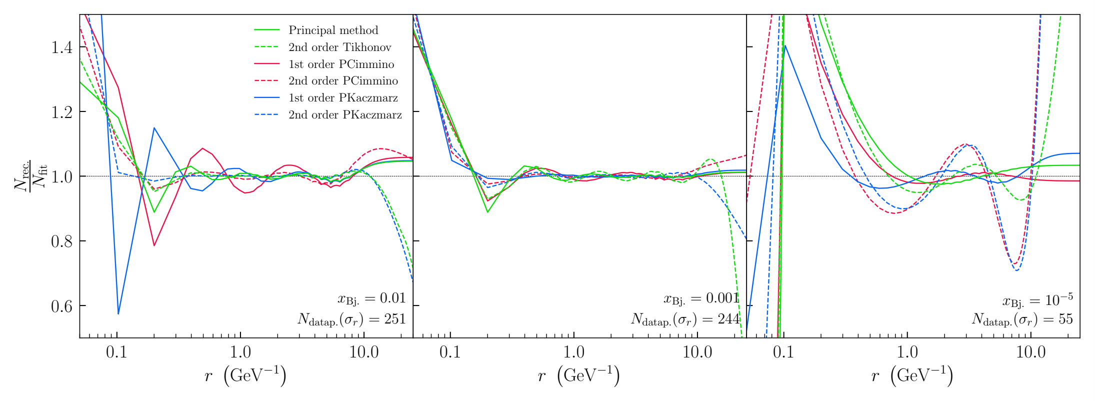


DIS reduced cross section from CKS dipole

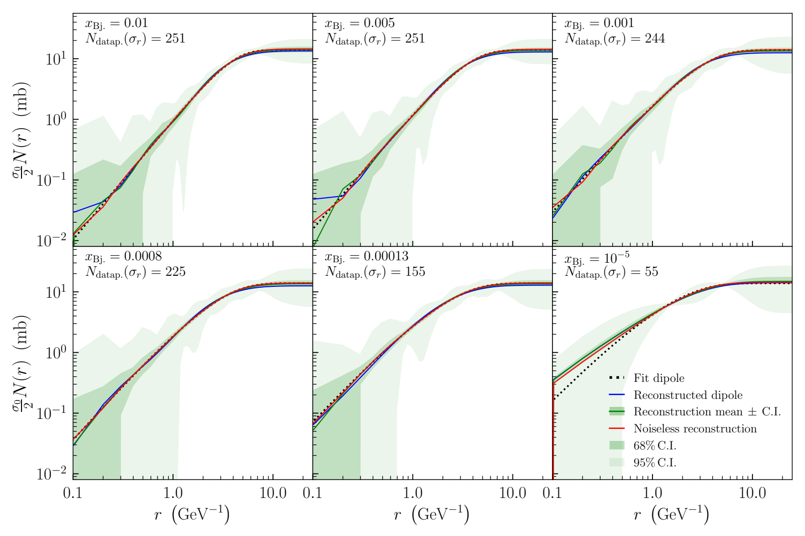


Reconstruction

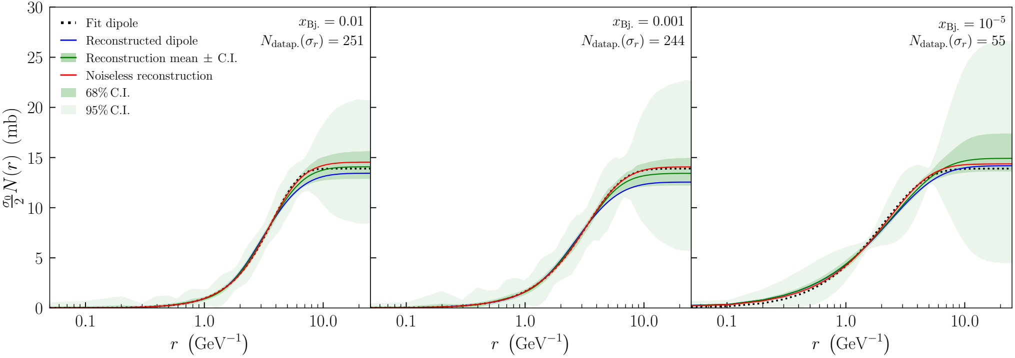
Comparison of regularization methods



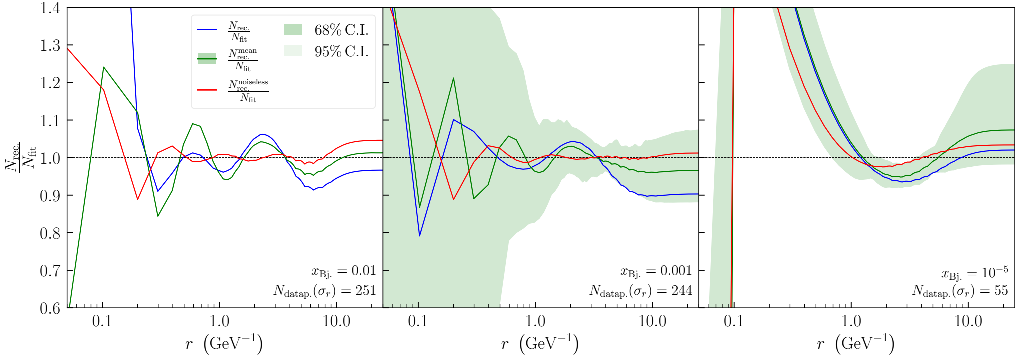
Reconstruction results: light quarks only, 4-par-CKM-dipole



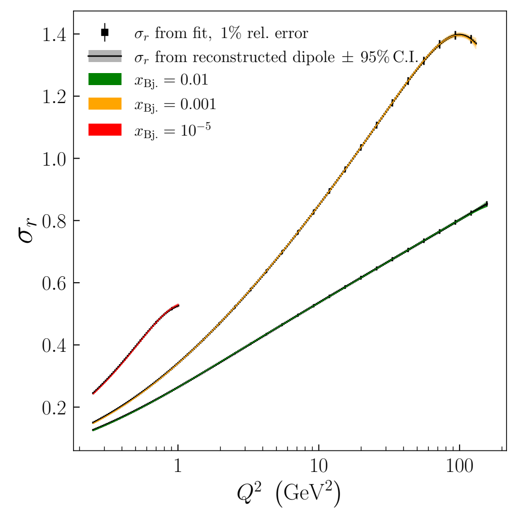
Reconstruction results: light quarks only, 4-par-CKM-dipole



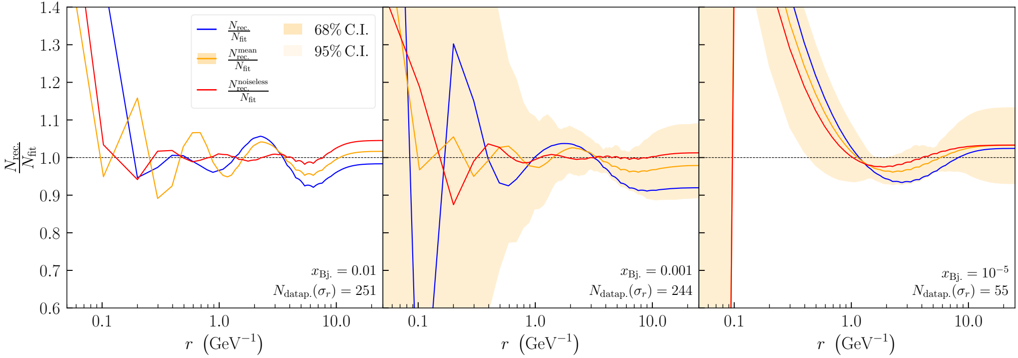
Relative accuracy of the reconstrution



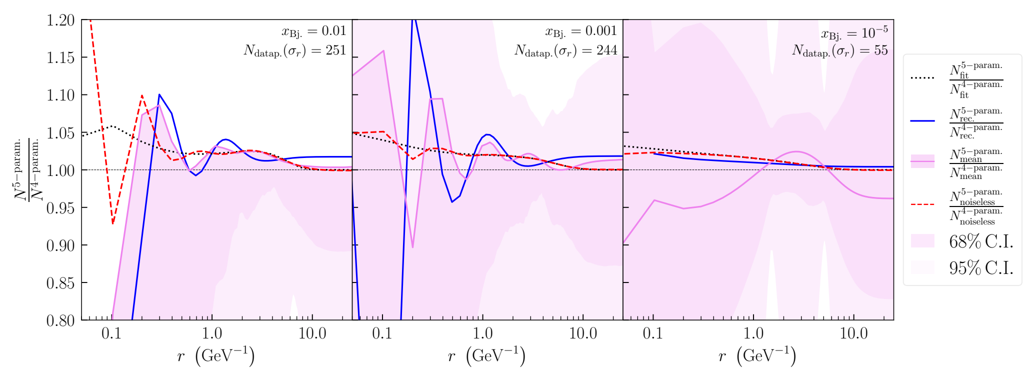
Cross sections from reference dipole and reconstruction



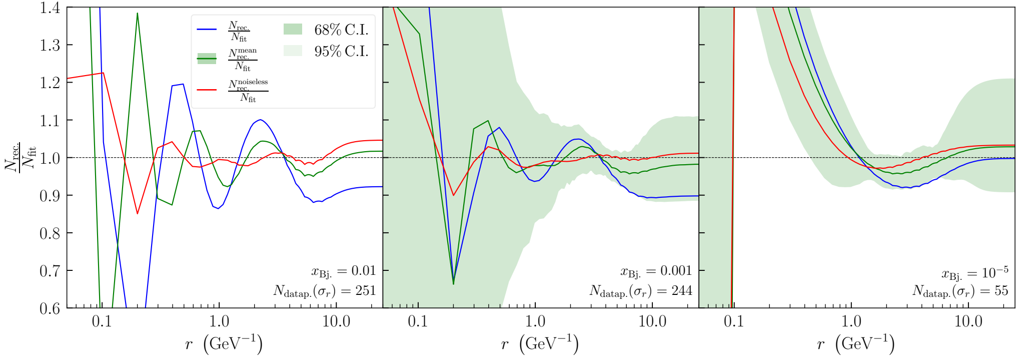
Reconstruction of the 5-parameter reference dipole



Can it differentiate between the two parametrizations?



Accuracy with light + charm quarks



Conclusions for the first paper

- Closure test results and conclusions would go here.
- Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

  Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris. Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non

HERA data

What about HERA data!?

Available HERA data

At $\sqrt{s} = 318\,\mathrm{GeV}$, $xbj <= 0.08$, all bins with ≥ 6 points in Q^2:

x	N
0.00013	8
0.0002	9
0.00032	11
0.0005	15
0.0008	19
0.0013	18
0.002	21
0.0032	24
0.005	24
0.008	23
0.013	30
0.02	35
0.032	33
0.05	30
0.08	30

$\sqrt{s}$	N
318.1	644
300.3	112
251.5	260
224.9	210

Data at lower $\sqrt{s}$ almost exclusively at large x, and bins are small.

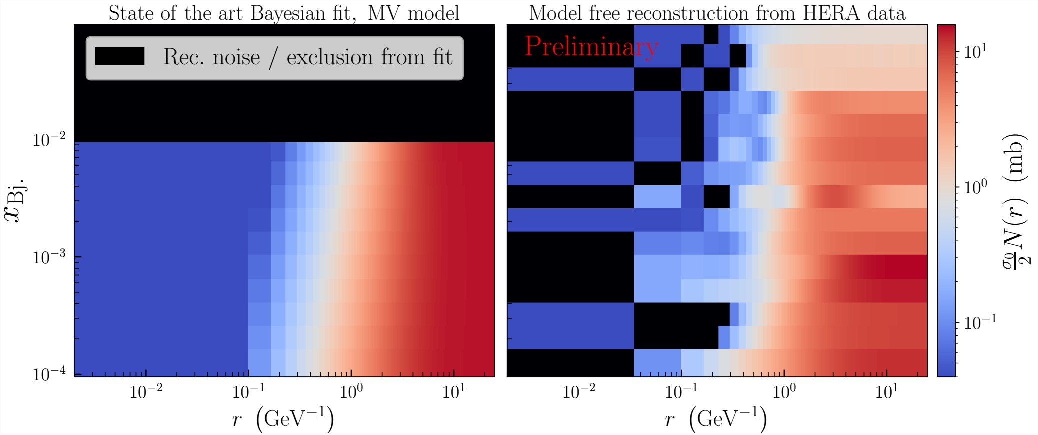
Towards global analysis as a multimodal inverse problem

How do we use all possible data simultaneously? And what about the charm production data?
We can stack the reconstruction problems and the algorithm doesn't care:

$$\begin{bmatrix} \sigma_r^{e^-p}(x, Q^2, s_1) \\ \sigma_r^{e^-p}(x, Q^2, s_2) \\ \sigma_r^{e^-p}(x, Q^2, s_3) \\ \sigma_r^{c\bar{c}}(x, Q^2, s_{c\bar{c}}) \end{bmatrix} = \begin{bmatrix} \varsigma_r^{\text{incl.}}(Q^2, s_1) \\ \varsigma_r^{\text{incl.}}(Q^2, s_2) \\ \varsigma_r^{\text{incl.}}(Q^2, s_3) \\ \varsigma_r^{c\bar{c}}(Q^2, s_{c\bar{c}}) \end{bmatrix} \mathbf{n}(r, x),$$

i.e. we can compute a new forward operator for each $s \in \mathbb{R}$, and include the corresponding data in the reconstruction. This also enables the inclusion of the charm production data, which is a strict subset of the total cross section data (only charm contribution in the forward problem.).
Work in progress.

(Pre-)preliminary HERA data reconstruction, only $\sqrt{s} = 318\,\mathrm{GeV}$



(Left) Smooth fit parametrization of the dipole amplitude $N(r, x)$.

(Right) Reconstruction of $N(r, x)$ from $\sigma_r(Q, x)$.

Model free inference can hopefully enable the discovery of new features from the data!

Conclusions

- We wrote the dipole picture of DIS into an integral transform of the dipole amplitude, and showed that it can be inverted with standard numerical inverse problems methods.

Conclusions

- We wrote the dipole picture of DIS into an integral transform of the dipole amplitude, and showed that it can be inverted with standard numerical inverse problems methods.
- Underlying feasibility connected to the K-transform or generalized Laplace transform? Analogously to the invertibility of the Radon transform, which is the foundation of computed tomography.
  - Tailored mathematics defined by the physical theory and measurement: Radon transform, ”dipole cross section transform“?.
  - Ultimate goal to enable generalized indirect measurement in QCD.

Conclusions

- We wrote the dipole picture of DIS into an integral transform of the dipole amplitude, and showed that it can be inverted with standard numerical inverse problems methods.
- Underlying feasibility connected to the K-transform or generalized Laplace transform? Analogously to the invertibility of the Radon transform, which is the foundation of computed tomography.
  - Tailored mathematics defined by the physical theory and measurement: Radon transform, ”dipole cross section transform“?.
  - Ultimate goal to enable generalized indirect measurement in QCD.
- Closure test shows promise for realistic application to real data, at least with the EIC, and with more advanced methods perhaps already with HERA data.

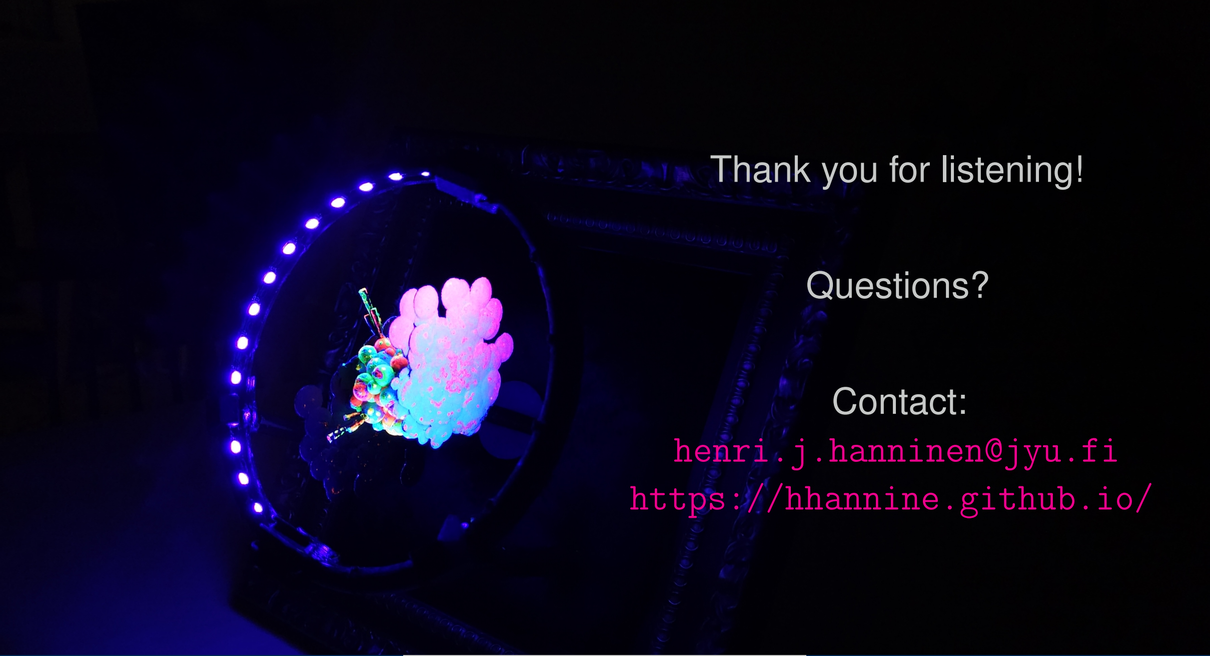
Conclusions

- We wrote the dipole picture of DIS into an integral transform of the dipole amplitude, and showed that it can be inverted with standard numerical inverse problems methods.
- Underlying feasibility connected to the K-transform or generalized Laplace transform? Analogously to the invertibility of the Radon transform, which is the foundation of computed tomography.
  - Tailored mathematics defined by the physical theory and measurement: Radon transform, ”dipole cross section transform“?.
  - Ultimate goal to enable generalized indirect measurement in QCD.
- Closure test shows promise for realistic application to real data, at least with the EIC, and with more advanced methods perhaps already with HERA data.
- Inverse problems mathematics and methods can open new paths towards indirect measurement in high-energy physics.

Conclusions

- We wrote the dipole picture of DIS into an integral transform of the dipole amplitude, and showed that it can be inverted with standard numerical inverse problems methods.
- Underlying feasibility connected to the K-transform or generalized Laplace transform? Analogously to the invertibility of the Radon transform, which is the foundation of computed tomography.
  - Tailored mathematics defined by the physical theory and measurement: Radon transform, ”dipole cross section transform“?.
  - Ultimate goal to enable generalized indirect measurement in QCD.
- Closure test shows promise for realistic application to real data, at least with the EIC, and with more advanced methods perhaps already with HERA data.
- Inverse problems mathematics and methods can open new paths towards indirect measurement in high-energy physics.

$\therefore$ If there is *data* and a *theory that can fit that data* $\implies$ basis for an inverse problem. (To develop a more general approach of indirect measurement than fitting model parameters.)



Thank you for listening!

Questions?

Contact:

henri.j.hanninen@jyu.fi

https://hhannine.github.io/