Exploring the nature of matter

Alberto Accardi Christopher Newport U. & Jefferson Lab

> Colloquium at Old Dominion University 29 October 2024

This work is partly supported by the DOE Office of Science

Outline

- Intro: what are we made of? (And how do we know?)
 - Quarks & gluons
 - Quantum Chromodynamics (QCD)
 - Color confinement
- Dynamical mass generation
 - How are massive particles created out of nearly massless quarks?
- Parton Distributions Functions (PDFs) in the proton
 - Quark and gluons at the "edge of confinement"
 - Global QCD analysis
 - \rightarrow Theory + Experiment + Data/Computer Science
- Perspectives

Acknowledgments

- My CTEQ-JLab (CJ) collaboration colleagues:
 - Matteo Cerutti (CNU), Xiaoxian Jing (Meta), Ishara Fernando (UVa),
 W.Melnitchouk (JLab), J.F.Owens (Florida State U.)
 - C.E. Keppel & Sanghwa Park (JLab), Shujie Li (LBL), P. Monaghan (CNU)

My other collaborators:

- C. Costa, A. Simonelli (JLab), A. Bacchetta (Pavia U., Italy), A. Krause (ODU),
 A. Signori (Torino U., Italy)
- [nCTEQ] F. Olness (Southern Methodist U.), R. Ruiz (Cracow, Poland), I.Schienbein (Grenoble)

• Grant funding:

- DE-SC0025004 "Large-x partons..."
- DE-AC05-06OR23177 (JLab theory)
- APS-Simons travel grant
- Artwork: Bernhard Musch (2009)
 Scattering discussion: R. Ruiz (2024)



General References

QCD global analysis from protons to nuclei:

- Accardi, <u>PoS DIS2015 (2015) 001</u>
- Jimenez-Delgado, Melnitchouk, Owens, <u>J.Phys.G40 (2013) 093102</u>
- Ethier, Nocera, Ann. Rev. Nucl. Part. Sci. (2020) 70, 1-34

QCD global analysis and statistical methods:

- Kovarik, Nadolsky, Soper, <u>*Rev.Mod.Phys.*</u> 92 (2020) 4, 045003
- Buckley, White and White, "<u>Practical Collider Physics</u>", IOP publishing 2021

References - technical

Inclusive Jet functions & dynamical mass generation

- Accardi, Costa, Signori, *Phys.Rev.D* 108 (2023) 114011
- Accardi, Signori, *Phys.Lett.B* 798 (2019) 134993 & *Eur.Phys.J.C* 80 (2020) 825
- Accardi, Bacchetta, *Phys.Lett.B* 773 (2017) 632

Large-x fits with nuclear corrections & applications

- <u>CJ15</u>: Accardi et al., <u>PRD 93 (2016) 114017</u>
 - Accardi, DNP 2020 / Fernando, GHP 2021 / Accardi, APS 2022
- F₂(n): Li, Accardi et al., <u>Phys. Rev. D 109 (2024) 074036</u>

Light quark asymmetry fits

- CJ15a: Accardi, Owens, Melnitchouk, *Phys.Lett.B* 801 (2020) 135143
- <u>CJ22</u>: Accardi, Jing, Owens, Park, <u>*Phys.Rev.D* 107 (2023) 113005</u>

PDF uncertainties

• Hunt-Smith, Accardi, Melnitchouk, Sato, Thomas, White, *Phys.Rev.D* 106 (2022) 036003

References - recent and in prep.

Inclusive Jet functions & dynamical mass generation

• Simonelli, Accardi, Cerutti, Costa, Signori

"Unveiling the Collins-Soper kernel in inclusive DIS at threshold"

• Rigorous QCD factorization with jet functions

Large-x fits with power and nuclear corrections & applications

• Accardi, Cerutti, Fernando, Li, Owens, Park

"Interplay of off-shell and higher-twist corrections in DIS at large x"

- Non negligible QCD analysis systematic effects
- Short version @ DIS 2024: *arXiv:2407.03589*
- Accardi, Krause

"Testing large-x QCD analysis in a spectator model"

Introduction: What are we made of?

What are we made of?

Quarks and gluons!





E. Rutherford, 1897, 1917

Quarks:

proposed: Gell-Mann & Zweig, 1964 discovered: SLAC (USA), 1968 Gluons: discovered: Desy (GER) 1979

accardi@jlab.org

What are we made of?

• Quarks and gluons!



How do we know?

• By using high-energy scattering as a microscope!

(actually, a femtoscope)

• And also looking at the spectrum of created particles

• Need a wavelength smaller than the proton's size:

$$\lambda pprox rac{hc}{E} < 10^{-15} ext{ m}$$

 $\Longrightarrow E > 1~{
m GeV}$

 $1 \text{ eV} = 10^{-19} \text{ J}$: energy needed to move an electric of charge $e = 10^{-19} \text{ C}$ through a potential difference of 1 V

Created by: A. Krause, 2021

© Andrew Krause, 2021







Electron-proton: elastic

• Exchanged photon wavelength: $\lambda_{\gamma} = h/p_{\gamma} \approx hc/E_{\gamma}$



Electron-proton: inelastic

• Exchanged photon wavelength: $\lambda_{\gamma} = h/p_{\gamma} \approx hc/E_{\gamma}$



Electron-proton: deeply inelastic scattering (DIS)

• Exchanged photon wavelength: $\lambda_{\gamma} = h/p_{\gamma} \approx hc/E_{\gamma}$



Cross section

• Roughly, the probability that the electron hits the target



accardi@jlab.org

Modern particle accelerators



Modern particle accelerators



accardi@jlab.org

Modern particle accelerators



Old Dominion U. - 29 Oct 2024

Quantum Chromodynamics (QCD)

- Simple rules
 - **Matter** = quarks + antiquarks:
 - \rightarrow 3 colors, 6 flavors, spin = ± $\frac{1}{2}$
 - Interaction = exchange of gluons

 $\mathcal{L} = \overline{\mathcal{F}}(i\mathcal{P} - m)\mathcal{F}$

4 Fur FMV

• "Lagrangian": contains the theory



matter:

Unlike any other fundamental theory!

- QCD is analogous to electromagnetism, but:
 - gluons can split... Ο
 - ...and bang on each other Ο
 - (no linear superposition principle!!)





- This happens so often that quarks are always surrounded by a dense network of gluons that holds together a group of several quarks
- Quarks form particles that are **always "color-neutral"** as a whole:

Color is *confined* inside hadrons!

Not your everyday charge... Ο



proton, neutron



Mysteries

• Gluon binding energy provides most of the proton mass:

 $egin{aligned} m_{ ext{proton}} &= 1 ext{ GeV}/c^2 \ ext{vs.} \ m_u + m_u + m_d &pprox 0.015 ext{ GeV}/c^2 \end{aligned}$

• Are we actually made of a force field??



• Why is the spin of proton = $\frac{1}{2}$?

An unlikely way of combining:

- the spin of (many) quark and gluons
- and their orbital angular momentum!



• In fact, why isn't there any free quark??

Color confinement

- Only "color neutral" particles have ever been (can ever be?) observed
- Why??
 - Mathematical explanation from QCD Lagrangian
 - \rightarrow I attempted one a long time ago (1996-1998) my first scientific love!
 - \rightarrow \$1M "millennium prize" from the Clay Institute!!
 - Lattice QCD computational explanation
 - \rightarrow But "black box"...



- Look at the effects of confinement
 - \rightarrow Dynamical mass generation
 - \rightarrow Parton Distribution Functions at the "edge of confinement"

my research!

Dynamical mass generation:

how do massive hadrons emerge from nearly massless quarks?

Parton propagation and fragmentation

Review: Accardi et al., Riv. Nuovo Cim. 032 (2010)

Nuclei as femtometer-scale detectors



Transverse momentum broadening

$$\Delta p_T^2 = \langle p_T^2 \rangle_A - \langle p_T^2 \rangle_D$$

Hadron attenuation

$$R_M = (N^h/N^e)_A/(N^h/N^e)_D$$

- At the Electron-ion collider
 - Higher energy
 "Jets" of particles aligned with struck quark Nucleus Jet π⁺ μ⁰
 Electron e⁻
 Li, Liu, Vltev, Phys. Lett. B 848 (2024) 138354

accardi@jlab.org

"Inclusive jets" in vacuum



Quantum mechanically

• Polarized electron+proton scattering:

 $\sigma(ec{e}+p^{\uparrow}
ightarrow X)\sim \mathcal{MM}^{*}\sim$



 q^\downarrow

 q^\downarrow

q

 q^{\uparrow}

 $\begin{array}{c} q^{\downarrow} \\ \hline \end{array} = \int \mu \, \rho_1(\mu^2) \, d\mu^2 \sim \langle \mu \rangle \\ \hline \end{array}$ "Spectral function": prob. distrib. of produced mass lattice

calculable: lattice QCD / Dyson-Schwinger

 q^+

Quantum mechanically

• Polarized electron+positron scattering:



- We are looking also for observables in proton-proton collisions
- Many observables → can think of a "global analysis" (see later)

Parton Distributions at the "edge of confinement"

What happens when a single quark contributes most of the proton's momentum?

Parton Distribution Functions

• **PDF** = probability distribution for a parton (quark or gluon) inside a proton to have momentum fraction $x = \frac{k}{p}$



• Large $x \rightarrow 1$

- The proton "is" a single quark up or down
- Confinement most directly determines the behavior of the d/u ratio



Parton Distribution Functions

• **PDF** = probability distribution for a parton (quark or gluon) inside a proton to have momentum fraction $x = \frac{k}{p}$



• Large $x \rightarrow 1$

- The proton "is" a single quark up or down
- Confinement most directly determines the behavior of the d/u ratio



How to measure the PDFs?



accardi@jlab.org

How to measure the PDFs?



Global QCD fits



• QCD factorization & universality:

can fit PDFs to a variety of hard scattering data

- Hadron-hadron collisions
 - \rightarrow Jets
 - → Electro-weak boson production
- Electron-proton DIS
- Electron-Deuteron DIS
- >1000's data points
- 40+ years of experience,
 - "High-energy" fitters:
 - \rightarrow CTEQ-TEA, MMHT, NNPDF, HERAPDF
 - Lower-energy / nuclear focus:
 - \rightarrow **CTEQ-JLab**, AKP, JAM



Global QCD fits



The CJ15 PDFs



- Fitted with $\chi^2 = 1.04$ / datum
- Propagation of exp. errors
 - Hessian analysis
 - Correlated errors used if available
- "PDF error band" for $\Delta \chi^2 = 2.71$
 - $\circ \rightarrow$ 90% c.l. in a perfect world
 - Many alternative methods
 - See review by Kovarik et al.
- Theoretical systematics more difficult
 - Recent effort by fitting community

Neutron structure function extraction

• There is no free neutron target

- Use (fitted) nuclear correction model
- Confine model dependence in deuteron / free p + n correction ratio



Apply data:

 \rightarrow e.g. to benchmark lattice QCD



(Some of the) challenges

- Very large data set:
 - Need to: modernize code, speed up calcs!
- Need fit flexibility → many parameters, highly correlated

 \rightarrow numerical instability in Hessian matrix

- Increase numerical robustness?
- Use neural networks?
 - → But: Cross Valid. leads to likelihood deformation!
- Better parametric methods
 - \rightarrow (iterative) bootstrap?
 - \rightarrow Fully Bayesian methods (MCMC, ...)?
- Visualization tools / information compression



(Some of the) challenges

• Incompatible data sets

- Increase "tolerance": blow up errors
- More solid (Bayesian) treatment
 - → E.g., multi-Gaussian mixtures <u>K. Mohan @ DIS 2023</u>
- How can AI/ML help?



To conclude...

Summary and Perspective

• What are we made of?

- Study the effects of confinement
- 2 intertwined research lines

• Global QCD analysis

- Many new data & better precision in near future
- Needs matching uncertainty determination
 - → Interdisciplinary challenge
 - → Expanded CTEQ-JLab collaboration

• Dynamical mass generation

- Theory: factorization with inclusive jet functions
- New observables
 - \rightarrow Novel global analysis
 - \rightarrow PDFs + i-JFN \rightarrow mass generation

Summary and Perspective

• What are we made of?

- Study the effects of confinement
- 2 intertwined research lines

• Global QCD analysis

- Many new data & better precision in near future
- Needs matching uncertainty determination
 - → Interdisciplinary challenge
 - → Expanded CTEQ-JLab collaboration

• Dynamical mass generation

- Theory: factorization with inclusive jet functions
- New observables
 - \rightarrow Novel global analysis
 - \rightarrow PDFs + i-JFN \rightarrow mass generation



Thank you!

Backup #1:

Evidence for quarks and gluons

Baryon spectroscopy – light sector (u, d, s), ground state

- J=3/2⁺: $|q_1^{\uparrow}, q_2^{\uparrow}, q_3^{\uparrow}\rangle$ totally symmetric w.fn.

- J=1/2⁺: $|q_1^{\uparrow}, q_2^{\uparrow}, q_3^{\downarrow}\rangle$



-

Baryon spectroscopy – light sector (u, d, s), ground state

 $\begin{array}{ll} & - J=3/2^+: \ | \mathbf{q}_1^{\uparrow}, \mathbf{q}_2^{\uparrow}, \mathbf{q}_3^{\uparrow} \rangle & \text{spin symmetric, color antisymmetric} \\ & - J=1/2^+: \ | \mathbf{q}_1^{\uparrow}, \mathbf{q}_2^{\uparrow}, \mathbf{q}_3^{\downarrow} \rangle & \text{spin antisymmetrics, color symmetric} \end{array}$



e+ + e⁻ annihilation into hadrons

– quark-mediated process $e^+ + e^- \rightarrow q + \bar{q} \rightarrow hadrons$







time

[http://www.quantumdiaries.org/author/richard-ruiz/]

e+ + e⁻ annihilation into hadrons

- quark-mediated process $e^+ + e^- \rightarrow q + \bar{q} \rightarrow hadrons$





- Jets in high-energy e+ + e⁻ collisions
 - Hadron produced in 2, 3, ... N, high-energy collimated "jets"
 - Evidence of common origin from a parton



Backup #2:

Experimental uncertainties or " Why do different global fits give different PDF uncertainties? "

Global fits are not created equal...



• Data choice and coverage, ...

- Can use SLAC, JLab only if considering TMC, HT corrections
- Highest *x* reach for d/u on proton if using reconstructed W asymmetries (vs. decay lepton asymmetries)

o ...

(*) CJ vs. CT comparison on "equal" footing: Accardi, Hobbs, Jing, Nadolsky, EPJC 81 (2021) 7

Bayesian estimators

• Bayes theorem $p(a|m) = \frac{1}{Z} p(m|a) p(a)$

with "evidence" $\mathcal{Z} = \int \mathrm{d} a \ p(m|a) \ p(a)$ and "likelihood" $p(m|a) = \mathcal{N} \exp\left[-\frac{1}{2}\chi^2(a,m)\right]$

Typical choice in PDF analyses

• Algorithms for sampling of likelihood \rightarrow probability density in $\{a_{\mu}\}$

- **HMC**: Hamiltonian Monte Carlo (an example of Markov-Chain MC methods)
- **NS**: Nested Sampling, primarily aimed at estimating the evidence
 - \rightarrow Samples the likelihood as a byproduct
- Expectation values

 $E_{\text{Bayes}} \{ \mathcal{O}(\boldsymbol{a}) \} = \frac{1}{n} \sum_{k=1}^{n} \mathcal{O}(\boldsymbol{a}_k) ,$

and variance

 $V_{\text{Bayes}}\{\mathcal{O}(a)\} = \frac{1}{n} \sum_{k=1}^{n} \left[\mathcal{O}(a_k) - E_{\text{Bayes}}\{\mathcal{O}(a)\}\right]^2$

Generalized Hessian Approximation

Hunt-Smith et al., PRD 106 (2022) 036003

• Start as usual:

- Find minimum of likelihood
- Diagonalize Hessian $\rightarrow e_k$ eigenvectors, w_k eigenvalues
- Change variables: $a(t) = a_0 + \sum_{k=1}^{n_{\text{par}}} t_k \frac{e_k}{\sqrt{w_k}}$, then $p(a|m) \to p(t|m)$
- Assume likelihood factorized along Hessian eigendirection, then

$$\begin{aligned} E_{\text{Hess}}\{\mathcal{O}(\boldsymbol{a})\} &= \int \mathrm{d}^{n} t \ p(\boldsymbol{t}|\boldsymbol{m}) \ \mathcal{O}(\boldsymbol{a}(\boldsymbol{t})) \ \approx \ \mathcal{O}(\boldsymbol{a}_{0}) \\ V_{\text{Hess}}\{\mathcal{O}(\boldsymbol{a})\} &\approx \sum_{k} T_{k}^{2} \left(\left. \frac{\partial \mathcal{O}(\boldsymbol{a}(\boldsymbol{t}))}{\partial t_{k}} \right|_{\boldsymbol{a}_{0}} \right)^{2} \end{aligned}$$

- Here $T_k^2 = \int dt_k \ p_k(t_k | \boldsymbol{m}) \ t_k^2$ is the "tolerance" :
 - $T_k = 1$ where likelihood is Gaussian;
 - Approximates well the likelihood in non-Gaussian directions
 - Maintains a "68%" or "1 σ " kind of meaning also when \neq 1

CT, MSTW \rightarrow T=5-10

• Often T_k determined "ad hoc" to account for statistical inconsistency of data

Data resampling

• Data Resampling (DR) approximates Bayes' posterior using frequentist logic

- Assume some prior (typically "flat")
- Reshuffle data within data uncertainty (Gaussian distribution)
- Maximize likelihood
- Repeat n_{rep} times $\rightarrow \{\boldsymbol{a}_k\}$
- Estimate

$$\begin{split} E_{\text{freq}}\{\mathcal{O}(\boldsymbol{a})\} &= \frac{1}{n_{\text{rep}}} \sum^{n_{\text{rep}}} \mathcal{O}(\boldsymbol{a}_{\text{rep}}) \,, \\ V_{\text{freq}}\{\mathcal{O}(\boldsymbol{a})\} &= \frac{1}{n_{\text{rep}}} \sum^{n_{\text{rep}}} \left[\mathcal{O}(\boldsymbol{a}_{\text{rep}}) - E_{\text{freq}}\{\mathcal{O}(\boldsymbol{a})\}\right]^2 \end{split}$$

• Good in parameter space region well constrained by data

Neural Networks and overfitting

• Neural networks provide:

- Efficient, very flexible parametrizations
- Hundreds of parameters
- Essentially a parameter free functional form ("nonparametric method")
- Use Data Resampling and aims at maximizing the same likelihood function

 $p(\boldsymbol{m}|\boldsymbol{a}) = \mathcal{N} \exp\left[-\frac{1}{2}\chi^2(\boldsymbol{a}, \boldsymbol{m})\right]$

- Without intervention, will overfit the data
 - The plot shows an extreme example



Cross-validation (CV) and stopping

- Needs a "stopping criterion"
 - to avoid fitting statistical noise instead of physics
- Randomly separate the data into 2 groups, say
 - \circ 70% \rightarrow training (T)
 - \circ 30 % \rightarrow validation (V)
- Fit the training, calculate $\chi^2(T)$ and $\chi^2(V)$
- Resample data, repeat
- "Stop" training when $\chi^2(V)$ is minimum:

 $\sigma = E[\sigma_{\rm fit}]$ $\delta \sigma = V[\sigma_{\rm fit}]$



Statistical uncertainties

- The method can effectively modify the likelihood!
 - Even with perfectly compatible (toy) data!

N. Hunt-Smith et al., PRD 106 (2022) 036003

• Bayesian Methods

(Markov Chain MC, Nested Sampling)

- Explore the likelihood function
- Well approximated by
 - \rightarrow Hessian, Data Resampling (**DR**)

• Cross Validation, NN-based fits

- Inflate the uncertainties
- Deform the likelihood



Statistical uncertainties

- The method can effectively modify the likelihood!
 - Even with perfectly compatible (toy) data!

N. Hunt-Smith et al., PRD 106 (2022) 036003

- Bayesian Methods
 - (Markov Chain MC, Nested Sampling)
 - Explore the likelihood function
 - Well approximated by
 - \rightarrow Hessian, Data Resampling (**DR**)

• Cross Validation & NN-based fits

- Inflate the uncertainties
- Deform the likelihood





Old Dominion U. – 29 Oct 2024