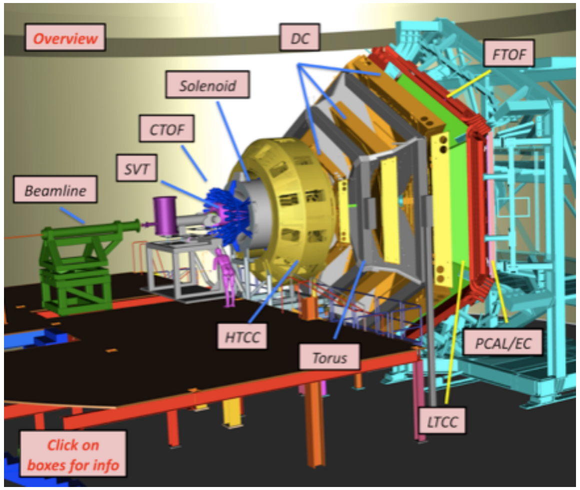
Experiment, Theory and Computation: opening a new window on hadron structure

David Richards

Jefferson Lab

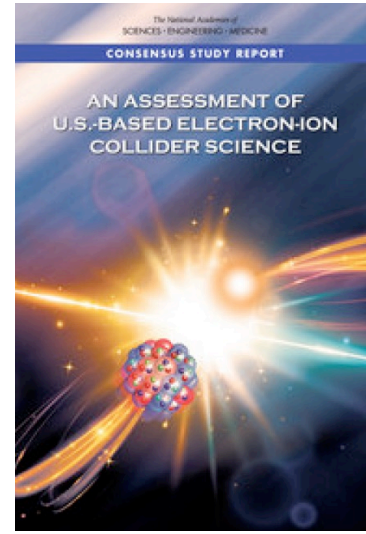
CLAS Collaboration Meeting, Nov 12-15, 2019

A New Opportunity in Hadron Structure

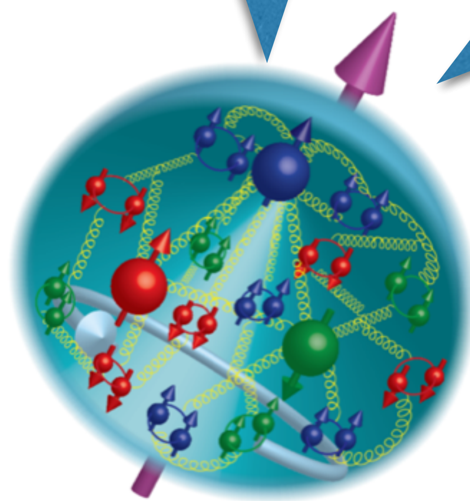


JLab@12GeV

Lattice QCD



Future Electron-Ion Collider



3D Image of nucleon and nuclei at the femtoscale

Virginia Center for Nuclear Femtography

- funded by Commonwealth to "*…..to facilitate the application of modern developments in **data science** to the problem of imaging and visualization of sub-femtometer scale structure of protons, neutrons, and atomic nuclei*"
- Multi-disciplinary, bringing together *nuclear theorists and experimentalists, mathematicians, computer scientists,… … and architects and artists*

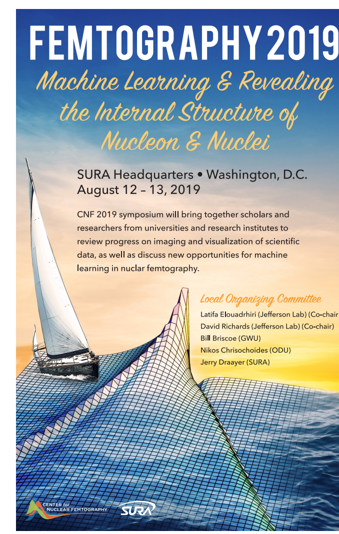
FEMTOGRAPHY 2018

Symposium on Imaging and Visualization in Science

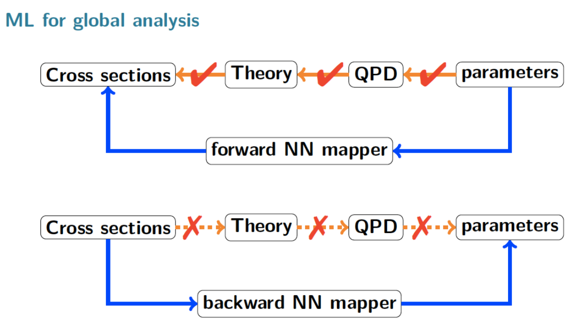
December 10-11, 2018, University of Virginia

The Symposium on Imaging and Visualization in Science will be held at the University of Virginia December 10-11, 2018. This symposium will bring together scholars and researchers from Virginia universities and research institutes to discuss recent developments and future opportunities in the imaging and visualization of scientific data.

CENTER *for* NUCLEAR FEMTOGRAPHY



Center for Nuclear Femtography

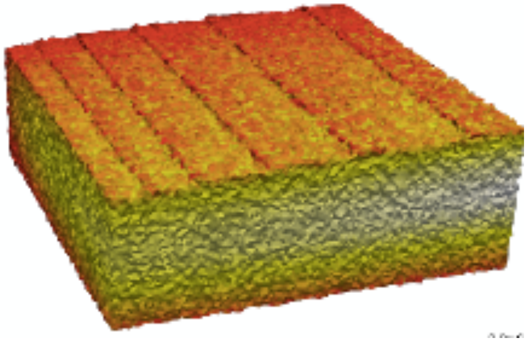


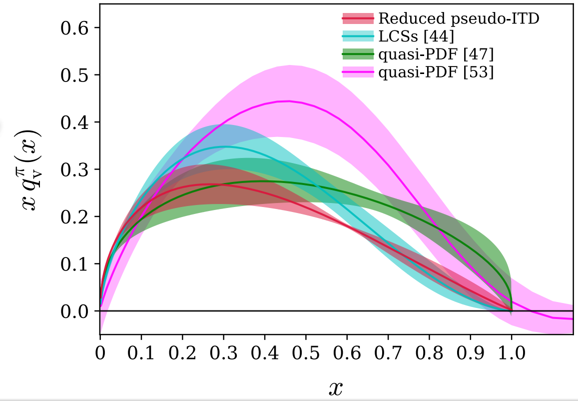
Machine Learning and Statistics

Visualization

Mesh-based Data Representation

Capability and Capacity Computing





Next-generation imaging filters and mesh-based data representation for phase-space calculations in nuclear femtography

PI: Nikos Chrisochoides (ODU)
Co-PIs: Gagik Gavalian (JLAB) and Christian Weiss (JLAB)

Executive Summary:

We propose an interdisciplinary project to leverage advanced computational methods from medical imaging for processing phase space distributions in nuclear femtography experiments, with the aim to enable next-generation process simulations, data analyses, and physics model comparisons. The objectives are: (**O1**) Implement an n-dimensional Exact Signed Euclidean Distance Transform for image-to-mesh conversion of phase space data; (**O2**) Implement an n-dimensional mesh-based representation of phase space data using tessellation methods and test it in low-dimensional scenarios (n=3); (**O3**) Explore potential physics model comparisons enabled by the new technologies.

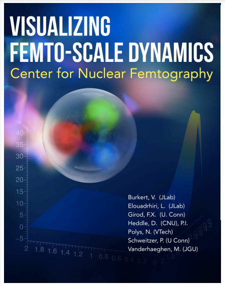
QCD theory and machine learning for global analysis

Nobuo Sato[1,2]*, Ian Cloët[3], Michelle Kuchera[4], Yaohang Li[1], Wally Melnitchouk[2], Andreas Metz[5]

[1] *Old Dominion University*, [2] *Jefferson Lab*, [3] *Argonne National Laboratory*, [4] *Davidson College*, [5] *Temple University*

Executive Summary

The goal of this project is to build the next generation of global QCD analysis tools using machine learning techniques to study the quantum probability distributions characterizing the internal structure of the nucleon. In concert, QCD-inspired models will be developed and used to calculate Wigner distributions and their projections onto generalized parton distributions (GPDs) and transverse momentum dependent (TMD) distributions. The QCD theory will be used to help train and optimize machine learning algorithms by putting physical constraints on the mapping between observables and the quantum probability distributions. This project is multi-disciplinary in nature, requiring collaboration between *nuclear physicists*, *computer scientists* and *information technology specialists*. The resulting product will be a critical tool for the nuclear physics community, opening up new possibilities for collaboration with computer science in the exploration and visualization of the inner structure of hadrons and nuclei. This proposal directly addresses project areas 1, 2, and 4 as identified in the proposal call.



Visualizing Femto-Scale Dynamics

Center for Nuclear Femtography

D. Heddle (CNU)

Executive Summary

Protons and neutrons constitute more than 99 percent of the mass in the visible universe. Through experiments using particle accelerators, we have learned much about these particles in the century since their discoveries. Yet very little is known about their internal mechanical and dynamical properties. With the advancement of accelerator, detector and associated electronics technologies, a new generation of experiments became possible at Jefferson Lab in Virginia. These experiments are generating the most comprehensive and the most precise data ever on their internal substructure, with data rates measured in Terabytes per hour. In the future, the proposed Electron Ion Collider (EIC) will allow scientists to explore these questions using precisely defined and controlled collisions of very-high energy ions and electrons at high luminosity.

Data Visualization and New Initiatives in Doubly Virtual Compton Scattering

Carl Carlson (WM), Marc Vanderhaeghen (Mainz)

Executive Summary

The flow of data from Jefferson Lab and other labs is already significant and the increase in the future will also be significant. Currently, much data that is presented as a many row by many column grid of small two-dimensional plots. Related problems are addressed, and to some extent solved, in the medical profession, where one often has a set of scans which can be presented as a set of fixed direction small two-dimensional images, looking rather like a grid of postage stamps. Computer processing turns these into interactive images that can be presented with arbitrary centers, with arbitrary scanning planes, and zoomable. We would aim, and be able to do, similar processing of nuclear science data.

A second and separate project concerns doubly virtual Compton scattering, $\gamma^* p \to \gamma^* p$, with the incoming γ spacelike and the outgoing γ timelike. This is an extension of approved experiments at JLab. Crucial applications of information obtained from such data are in evaluating two-photon exchange (TPE) corrections to radiative corrections in ep elastic scattering, TPE corrections to the Lamb shift in muonic hydrogen (critical for the proton radius measurements and puzzle), and for calculating the electromagnetic contribution to the neutron-proton mass difference (using the Cottingham formula). We would calculate the Bethe-Heitler amplitudes which interfere with the purely Compton amplitudes, and give expressions allowing extractions of information of the individual structures parameterizing the Compton amplitude.

Next-generation Visual Analysis Workspace for Multidimensional Nuclear Femtography Data

PI Nicholas Polys (VT, npolys@vt.edu)
CO-PIs Srijith Rajamohan (VT), Markus Diefenthaler (JLAB), Dmitry Romanov (JLAB)

Executive Summary

The experimental data in five or more kinematics dimensions that allows to constrain GPDs and TMDs is a multidimensional data science challenge. We propose to apply recent advances in scientific visualization to gain more insights into the multidimensional datasets at the forefront of Nuclear Femtography. In the initial phase of our project, we will explore Semantic Interactions as a visualization technique to analyze an ensemble of scientific data sets. This is motivated by our ongoing R&D and the direct connection between Semantic Interactions and machine learning.

Title: Parton Distribution Functions from Lattice QCD

PI: Konstantinos Orginos (W&M/JLab)
co-PI: Andreas Stathopoulos (W&M)
Grad Students: Joseph Karpie(W&M)
Post-doc: Eloy Romero(W&M)

Executive Summary:

In the last few years, a major achievement in hadron structure has been the development of new methods that allow for lattice QCD computations of parton distribution functions. This is a groundbreaking development as it allows for the first time to determine the full longitudinal momentum fraction dependence of the PDFs from lattice QCD, and thus opens up a new window for the theoretical study of the structure of fundamental building blocks of matter such as the pion and the nucleon. Experimentally hadron structure studies are a central part of DOE's nuclear physics programs both with current experimental facilities such as the 12 GeV upgrade of Jefferson Lab, as well as at the future electron-ion collider (EIC).

Principal Investigator: **Simonetta Liuti** (UVA Physics)

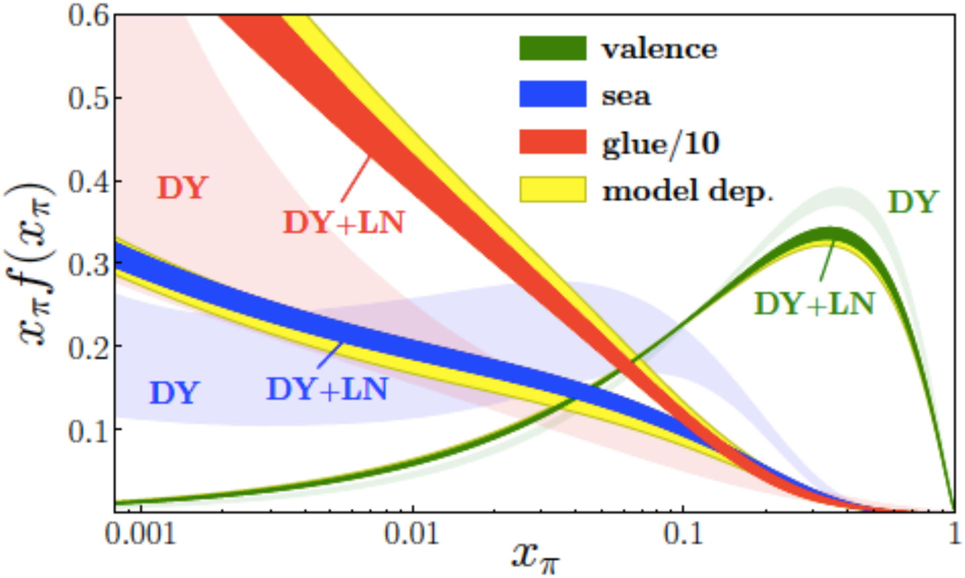
Wigner Imaging

Co-Principal Investigators: **Peter Alonzi (UVA School of Data Science)**
Matthias Burkardt (NMSU Physics)
Dustin Keller (UVA Physics)
Olivier Pfister (UVA Physics)
Petra Reinke (UVA School of Engineering and Applied Science)

The science of Nuclear Femtography probed by deeply virtual exclusive reactions has revolutionized our approach to exploring the internal structure of the nucleon. A new generation of current and planned experiments at the future EIC could in principle allow us to incorporate all the information from data and phenomenology into a tomographic image connecting the deepest part of the quantum world with what we see as everyday matter around us. However, to harness and organize information from experiment and increase the reach of this emergent field will require going beyond the standard computational toolbox. The proposed pilot project is an effort in this direction. It consists of two parts: computational and visualization on one side, and theoretical, on the other. In order to carry out our program we will: **(1)** examine and evaluate the use of new state of the art computational methods and techniques, including visualization to address the many layers of analysis which are necessary to extract the signal in its complex background after the large experimental data sets are acquired; **(2)** simultaneously develop a flexible model of the Wigner distribution which underlies the theoretical description of the data.

Inverse Problem

In deciphering the structure of hadrons, we are always dealing with *sparse* and *incomplete* data.



Barry et al., arXiv:1804.01965

Assuming *factorization*

$$\sigma = \int H \otimes f_a \otimes f_b$$

Incomplete Cross section

PQCD Kernel

PDF

This is an ill-posed problem

➡ *Need additional information*

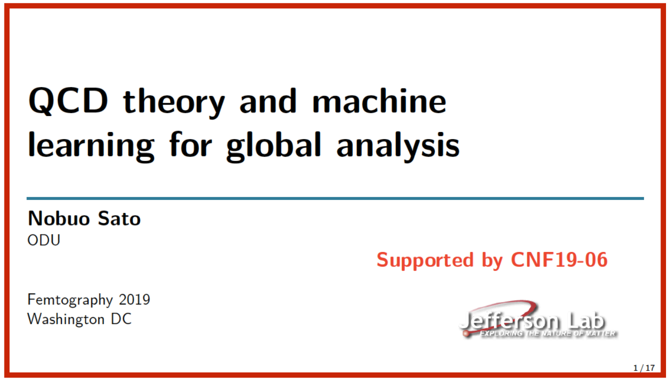
e.g. we assume a functional form:

$$q(x) = N x^{\alpha} (a - x)^{\beta} P(x)$$

There are no model-independent PDFs

This is common amongst many areas of science!

Machine Learning


QCD theory and machine learning for global analysis

Nobuo Sato
ODU

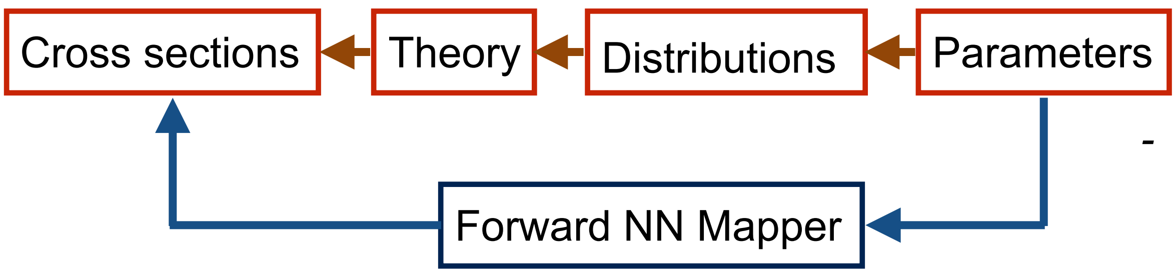
Supported by CNF19-06

Femtography 2019
Washington DC

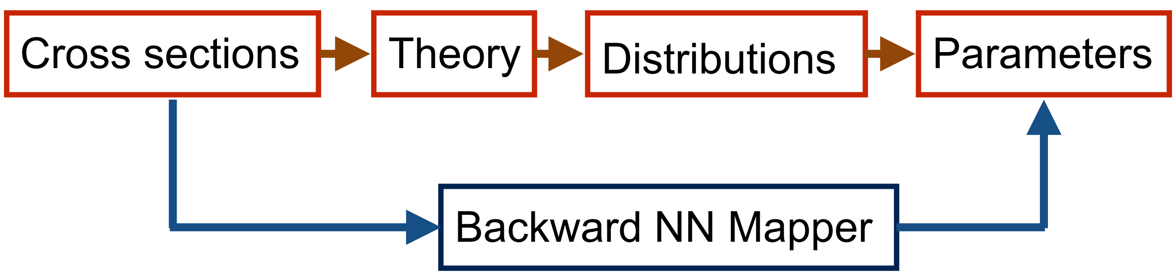
1 / 17

"To build the next generation of global QCD analysis tools using machine learning (ML) techniques to study the quantum probability distributions (QPD) characterizing the internal structure of the nucleon."

Slide from Nobuo Sato, CNF 2019



- *Predictions for Future Experiments*

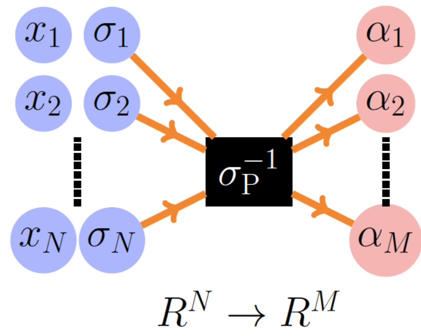


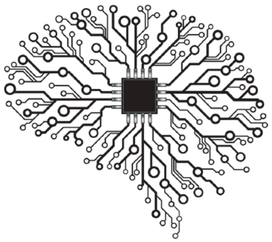
- *What datasets/points constrain distributions?*
- *How do distributions depend on parametrization?*
- *What additional measurements to contain distributions?*

- We know how to go from $\boldsymbol{a}$ to cross sections e.g.

$$\frac{d\sigma}{dx dQ^2} = \sum_q \int_x^1 \frac{d\xi}{\xi} H(\xi) f_q\left(\frac{x}{\xi}, \mu; \boldsymbol{a}\right)$$

- We **DON'T** have the inverse function to go from cross sections to $\boldsymbol{a}$

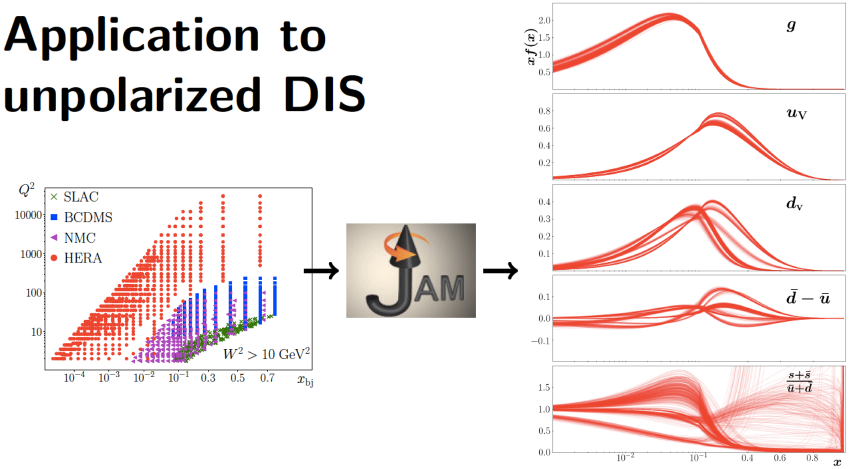




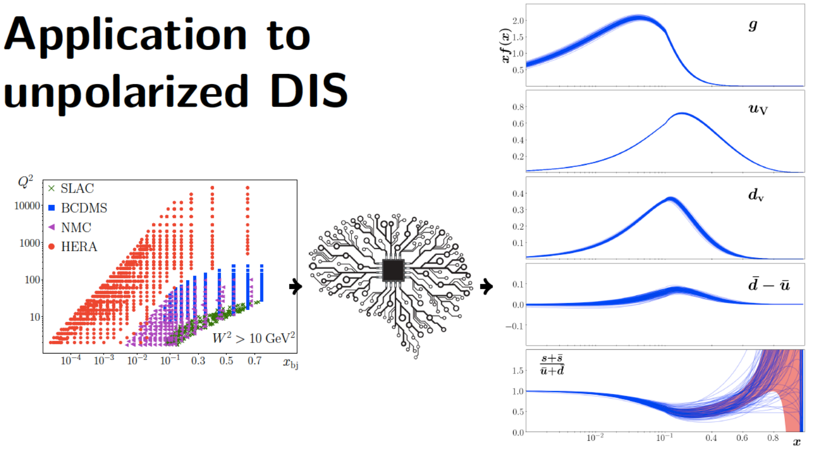
Can we use Machine Learning?

N.Sato, EINN2019

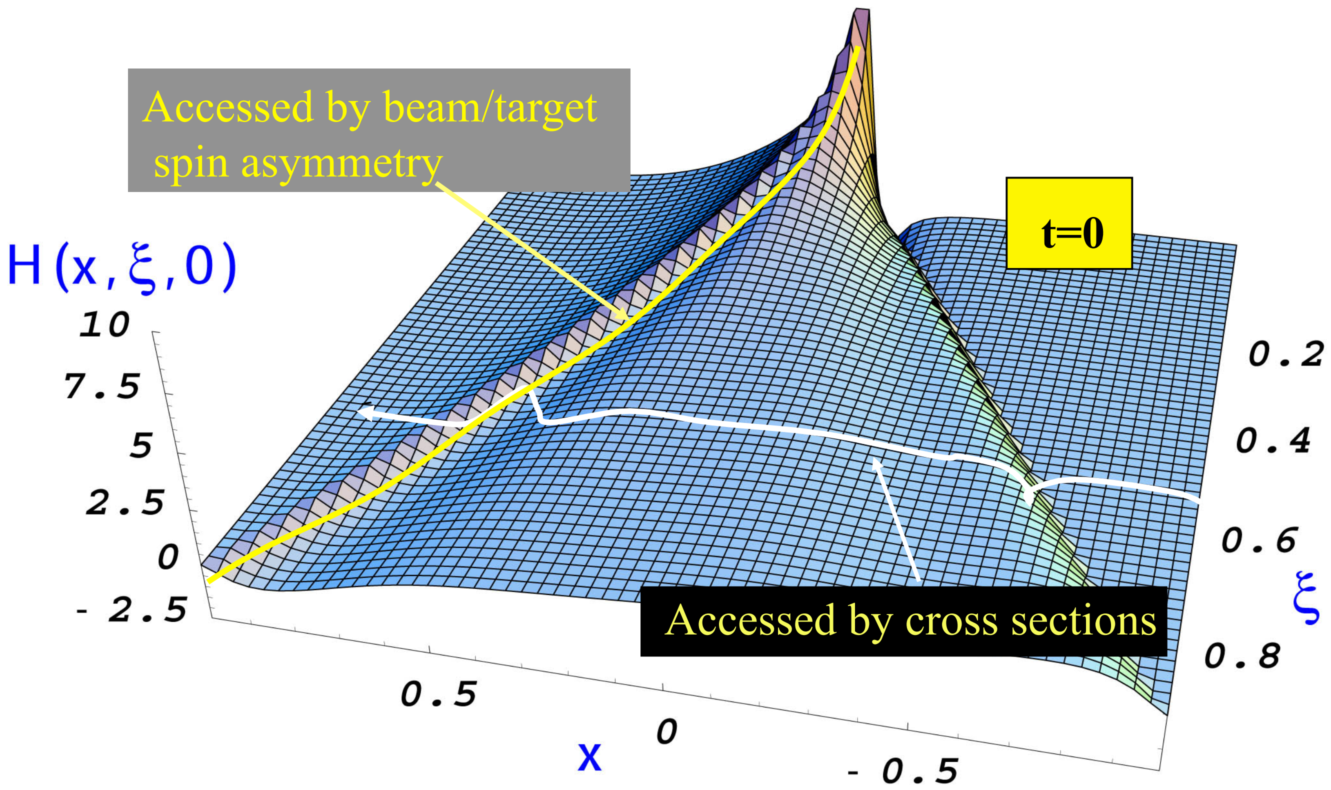
Application to unpolarized DIS



Application to unpolarized DIS



GPDS KINEMATICS



For 3D imaging, GPDs and PDFs, an even bigger challenge - less guidance on parametrization

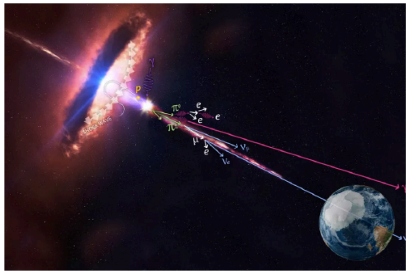
Visualization

- Guiding analysis and error checking
- Understanding
- "Eye Candy"

There are more things in Heaven and Earth…

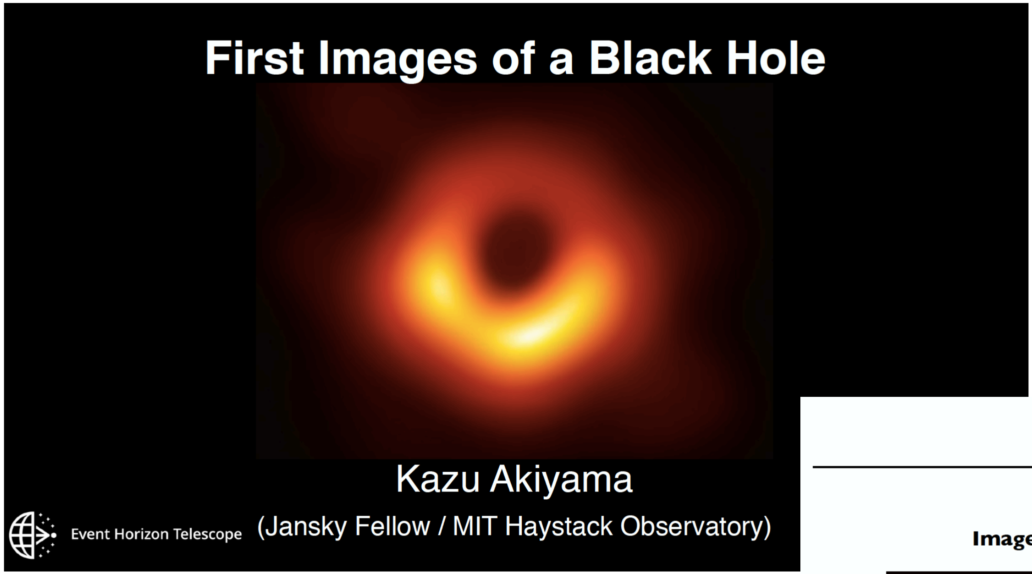
"Extensive background radiation studies by IBM in the 1990s suggest that computers typically experience about one cosmic-ray-induced error per 256 megabytes of RAM per month. If so, a superstorm, with its unprecedented radiation fluxes, could cause widespread computer failures. Fortunately, in such instances most users could simply reboot" (*Supplement to the feature "Bracing the Satellite Infrastructure for a Solar Superstorm," August 2008 issue,* Scientific American.)

"While double bit flips were deemed unlikely, the density of DIMMs at Oak Ridge National Lab's Cray XT5 causes them ***to occur on a daily basis*** (at a rate of one per day for 75,000+ DIMMs)" (Fiala+, 2012)

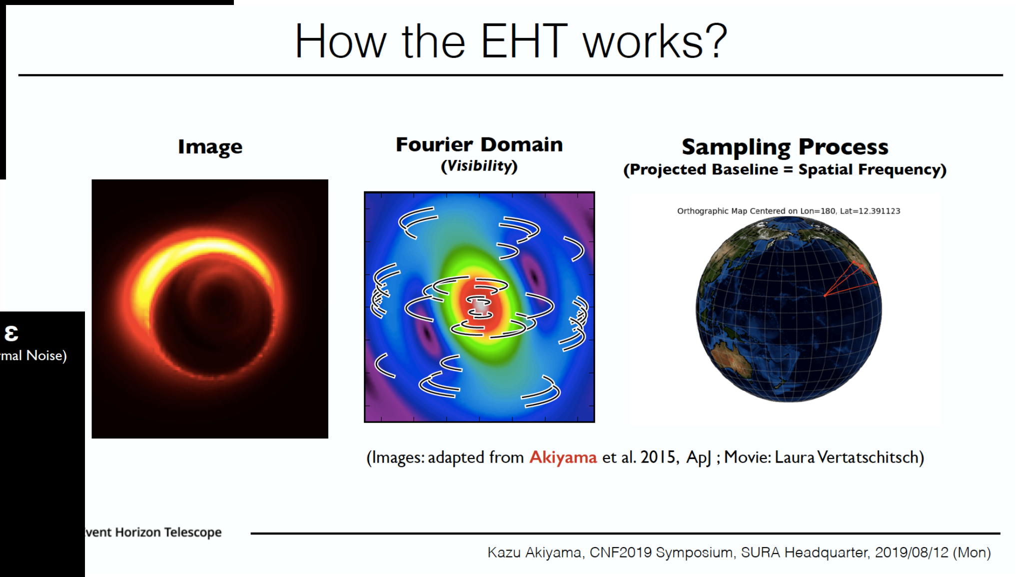


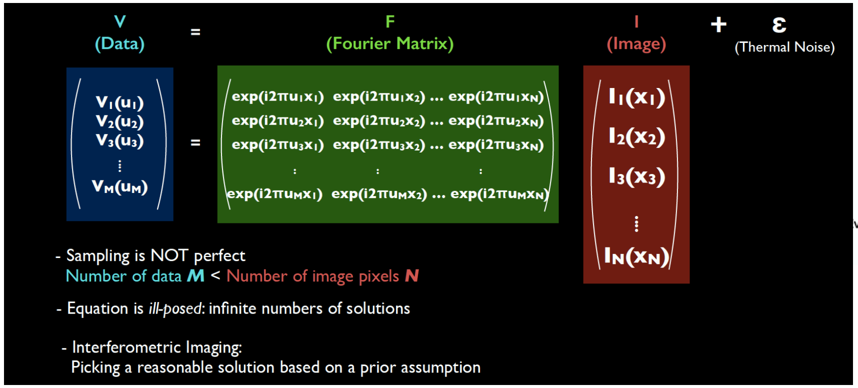
Computers can be pretty good cosmic-ray detectors - visualization for debugging

Bronson Messer, ORNL, CNF2019



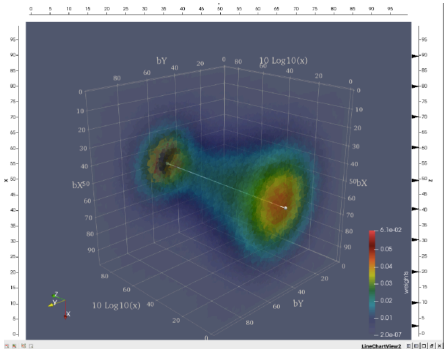
slide from *Kazu Akiyama*
- Femtography 2019





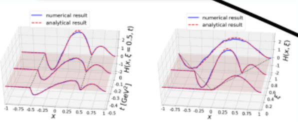
Next-Generation Imaging Filters and Mesh-Based Data Representation for Phase-Space Calculations in Nuclear Femtography (CNF19-04)

Gagik Gavalian (Jefferson Lab)
Nikos Chrisochoides (ODU)
Christian Weiss (Jefferson Lab)
Pawel Sznajder (NCBJ Warsaw)
Christos Tsolakis (ODU),
Angelos Angelopoulos (ODU)



FULL Chain

PARTONS



PARTONS Generate:
- **quark distribution functions**
- **GPD's (H,E,Ht,Et)**
- **Beam Spin Asymmetries**
- **Cross Sections (BH, DVCS, INT)**

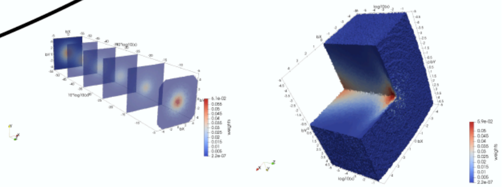
ANY USER PROGRAM

OTHER CNF Proposals

CNF_I2M:
- convert GRID to NRRD/DICOM
- generate tessellated file (VTK)
- software is available through docker

CNF_I2M Tessellation ODU Docker

ParaView



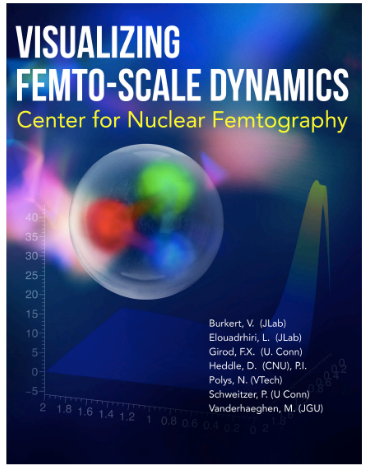
ParaView:
- plot slices of tessellated image
- boolean operations with planes and objects
- calculate integrals over lines or planes
- custom Python views with fully programmable data analysis routines.
- provides most customizable data visualization plugins (user programmable).

Future:
- user programs can use this chain
- other proposals can take advantage the software.

CNF19-09 Visualizing Femto-Scale Dynamics
Mid-Term Report

V. Burkert
Jefferson Lab (Newport News)

N. Polys
Virginia Tech (Blacksburg)

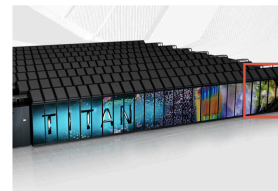


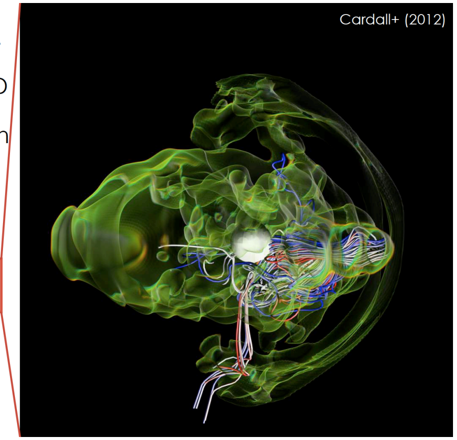
CNF19-09 *FEMTOGRAPHY2019 - Symposium, SURA Washington DC, 8/12-13, 2019* 1

Bronson Messer, ORNL, CNF2019

OK, some eye candy…

- This image from an MHD version of the SASI graced the front of Titan for >7 years.





OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

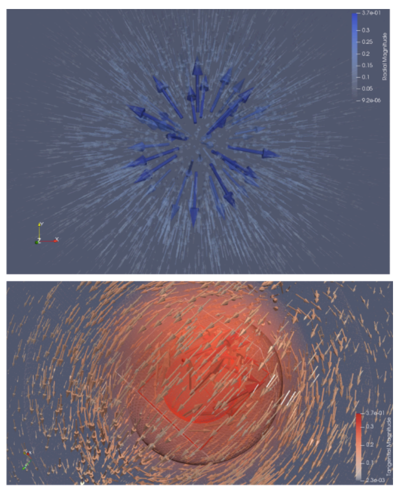
↓

Nucleon Standard Model

Fast Monte Carlo DVCS:

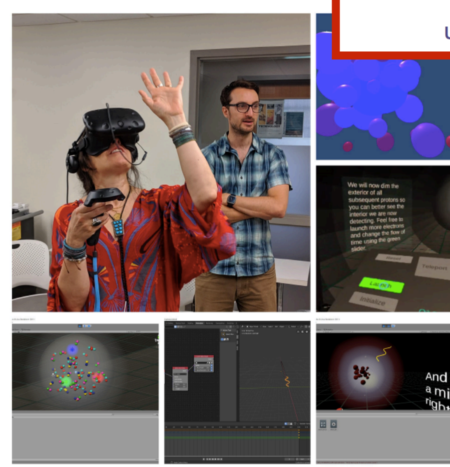
- Radial forces
- Tangential forces

Arrow color and length by force magnitude



Summer Institute on Wigner Imaging and Femtography

Simonetta Liuti
University of Virginia

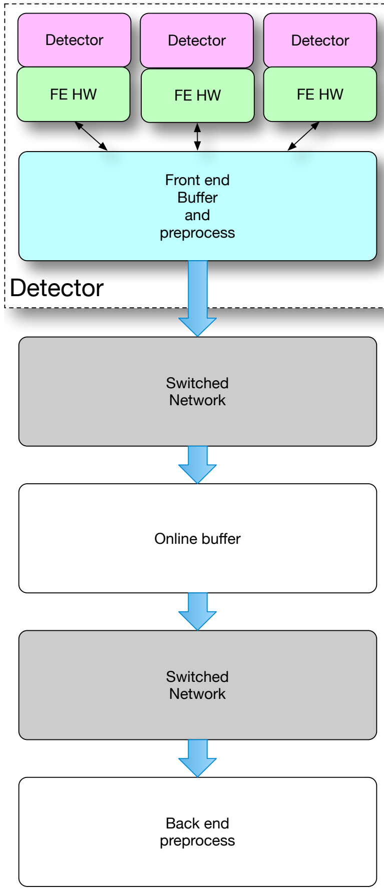


Computation

New technology driving change

Slide: Graham Hayes, SC19

- Existing DAQ designs at JLab are based on several assumptions:
  - *Experiments generate data at a bandwidth that is too high for an affordable system to acquire.*
  - *Even if the data could be acquired it could not be stored.*
  - *At these rates the data could not be processed by software in real, or near-real, time.*
- In recent years it has become clear to several groups, both at JLab and outside, that these assumptions are no longer true.
- It is now possible to acquire data with minimal filtering in parallel streams to short term storage and process in near real time to reduce it to a volume that can be permanently archived.
- This approach is known as *streaming readout.*
  - Much of what was formerly done online in custom electronics, firmware and embedded software is moved near/off-line.



Exploit HPC

Opposite of Edge Computing?

Lattice QCD

Capability Computing - Gauge Generation

e.g. Summit at ORNL

$$P[U] \propto \det M[U] e^{-S_G[U]}$$

Several *V, a, T, m_π*

~ 10% Leadership-Class Resources

Capacity Computing - Observable Calculation

e.g. GPU/KNL cluster at JLab, BNL, FNAL

$$\langle \mathcal{O} \rangle = \frac{1}{N} \sum_{n=1}^{N} \mathcal{O}(U^n, G[U^n])$$

$$\text{e.g. } C(t) = \sum_{\vec{x}} \langle N(\vec{x}, t) \bar{N}(0) \rangle$$

"Desktop" Computing - Physical Parameters

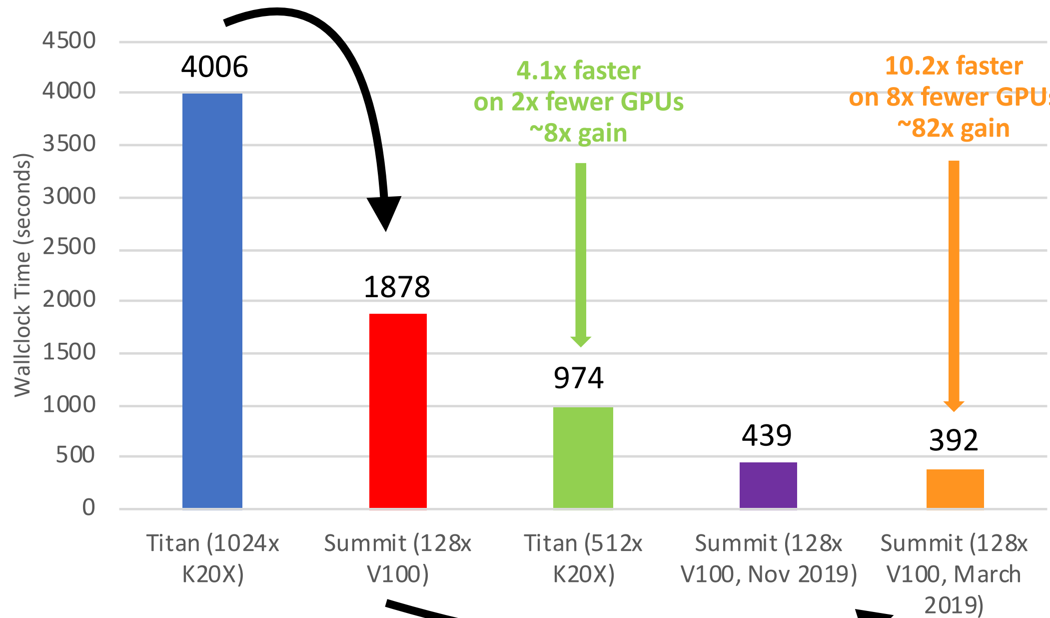
e.g. Mac at your desk

$$C(t) = \sum_n A_n e^{-E_n t}$$

$$M_N(a, m_\pi, V)$$

LQCD: Moving toward the exascale

Hardware: 2.13x wall-time on 8x fewer GPUs = **17x**



Algorithms, Software and Tuning: **4.79x**

The USQCD Clover Gauge Generation
ECP FoM since 2016
using Chroma, QDP-JIT + QUDA on Summit

81.6x
overall gain

LQCD: Theory Advances

- Euclidean lattice precludes calculation of light-cone/time-separated correlation functions

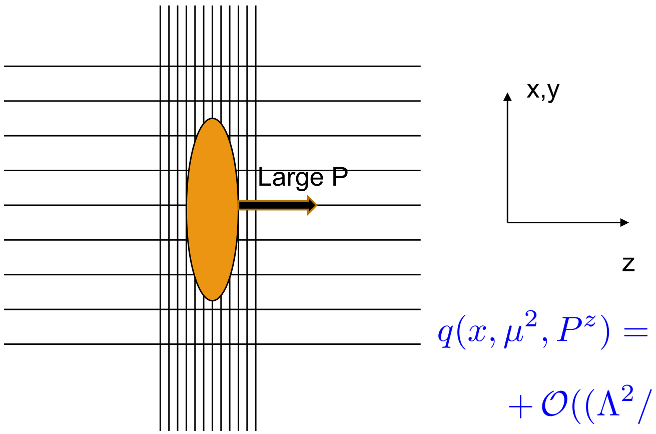
$$q(x,\mu) = \int \frac{d\xi^-}{4\pi} e^{-ix\xi^- P^+} \langle P \mid \bar{\psi}(\xi^-)\gamma^+ e^{-ig\int_0^{\xi^-} d\eta^- A^+(\eta^-)} \psi(0) \mid P\rangle$$

So…. …Use *Operator-Product-Expansion to formulate in terms of Mellin Moments* with respect to Bjorken x.

$$\langle P \mid \bar{\psi}\gamma_{\mu_1}(\gamma_5)D_{\mu_2}\dots D_{\mu_n}\psi \mid P\rangle \rightarrow P_{\mu_1}\dots P_{\mu_n} a^{(n)}$$

- Discretised lattice: power-divergent mixing for higher moments

Large-Momentum Effective Theory (LaMET)



"Equal time" correlator

X. Ji, Phys. Rev. Lett. 110, 262002 (2013).
X. Ji, J. Zhang, and Y. Zhao, Phys. Rev. Lett. 111, 112002 (2013).
J. W. Qiu and Y. Q. Ma, arXiv:1404.686.

$$q(x,\mu^2,P^z) = \int \frac{dz}{4\pi} e^{izk^z} \langle P \mid \bar{\psi}(z)\gamma^z e^{-ig\int_0^z dz' A^z(z')} \psi(0) \mid P > + \mathcal{O}((\Lambda^2/(P^z)^2), M^2/(P^z)^2))$$

$$q(x,\mu^2,P^z) = \int_x^1 \frac{dy}{y} Z\left(\frac{x}{y},\frac{\mu}{P^z}\right) q(y,\mu^2) + \mathcal{O}(\Lambda^2/(P^z)^2, M^2/(P^z)^2)$$

Pseudo-PDFs

Pseudo-PDF (pPDF) recognizing generalization of PDFs in terms of *Ioffe Time* $\nu = p \cdot z$.

A.Radyushkin, Phys. Rev. D 96, 034025 (2017)

B.Ioffe, PL39B, 123 (1969); V.Braun *et al*, PRD51, 6036 (1995)

$$M^\alpha(p,z) = \langle p \mid \bar{\psi}\gamma^\alpha U(z;0)\psi(0) \mid p\rangle$$

Ioffe-Time Distribution

$$M^\alpha(z,p) = 2p^\alpha \mathcal{M}(\nu, z^2) + 2z^\alpha \mathcal{N}(\nu, z^2)$$

$z = (0, z_-, 0_T)$

Ioffe-time pseudo-Distribution (pseudo-ITD) generalization to *space-like z*

$$\Downarrow$$

$$\mathcal{M}(\nu, z^2) = \int_{-1}^{1} dx\, e^{i\nu x}\mathcal{P}(x, z^2)$$

← *pseudo-PDF*

$$\Downarrow$$

$$f(x) = \mathcal{P}(x,0) \underset{z_3^2 \to 0}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} d\nu e^{-i\nu x}\mathcal{M}(\nu, -z_3^2)$$

pPDFs - II

To deal with UV divergences, introduce reduced distribution $\mathfrak{M} = \frac{\mathcal{M}(\nu, z^2)}{\mathcal{M}(0, z^2)}$

$$\mathfrak{M}(\nu, z^2) = \int_0^1 du\, K(u, z^2\mu^2, \alpha_s) Q(u\nu, \mu^2)$$

Computed on lattice

Perturbatively calculable

Ioffe-time Distribution

$$Q(\nu,\mu) = \mathfrak{M}(\nu, z^2) - \frac{\alpha_s C_F}{2\pi} \int_0^1 du \left[\ln\left(z^2\mu^2 \frac{e^{2\gamma_E+1}}{4}\right) B(u) + L(u)\right] \mathfrak{M}(u\nu, z^2).$$

K. Orginos et al., PRD96 (2017), 094503

Match data at different *z*

⇓

$$Q(\nu) = \int_{-1}^{1} dx\, q(x) e^{i\nu x}$$

$$q(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\nu\, e^{-i\nu x} Q(\nu)$$

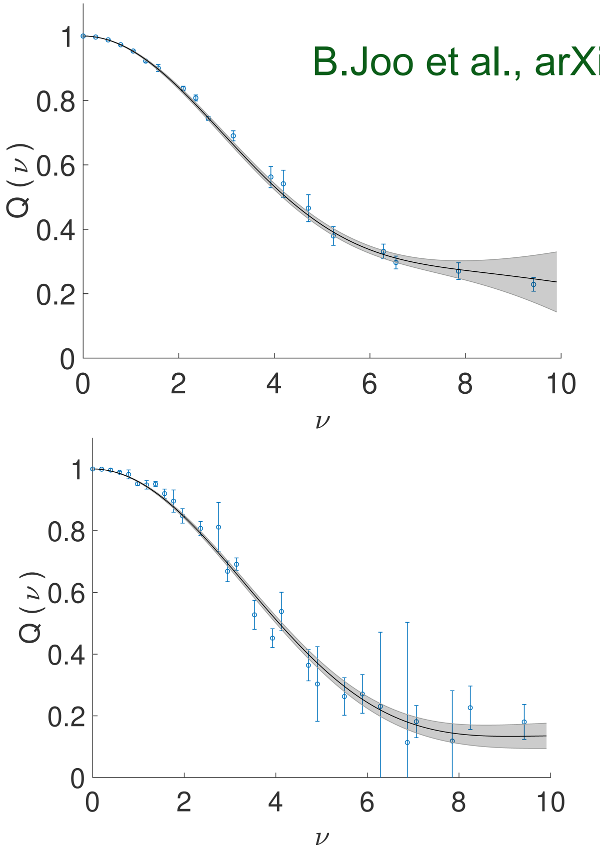
Need data for all ν, *or additional physics input*

Inverse problem

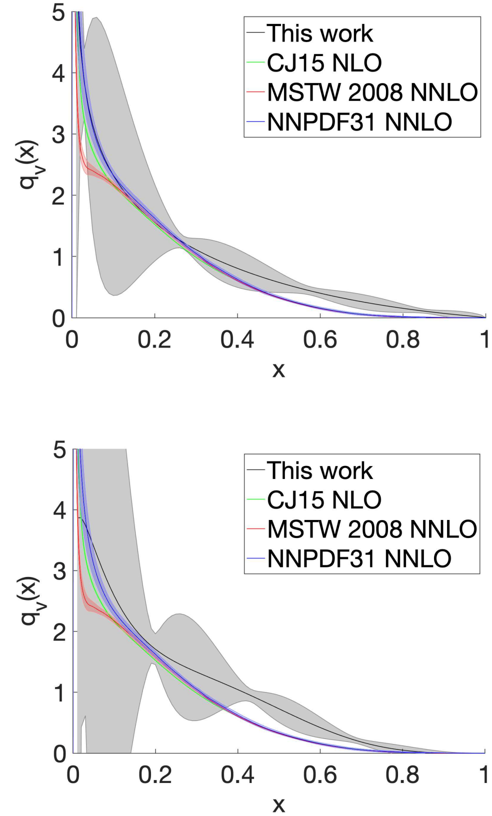
Nucleon PDF

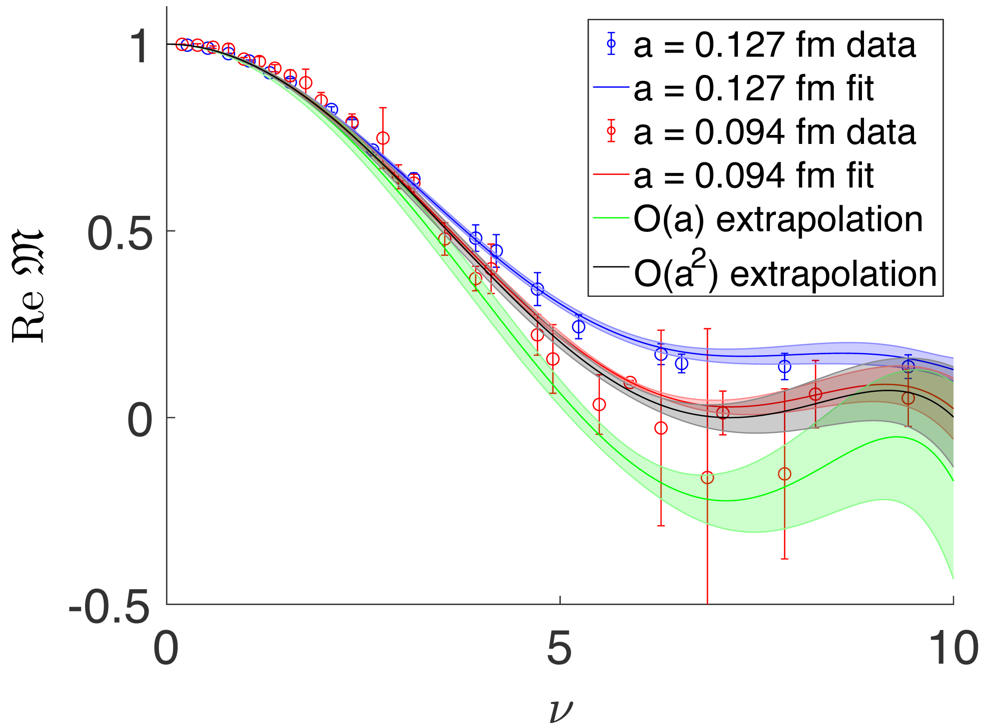
Ground-breaking quenched calculation: K. Orginos et al., PRD96 (2017), 094503

ID	a(fm)	M_π(MeV)	β	c_{SW}	am_l	am_s	$L^3 \times T$	N_{cfg}
$a127m415$	0.127(2)	415(23)	6.1	1.24930971	-0.2800	-0.2450	$24^3 \times 64$	2147
$a127m415L$	0.127(2)	415(23)	6.1	1.24930971	-0.2800	-0.2450	$32^3 \times 96$	2560
$a094m390$	0.094(1)	390(71)	6.3	1.20536588	-0.2350	-0.2050	$32^3 \times 64$	417



⟹





Large Range in ν

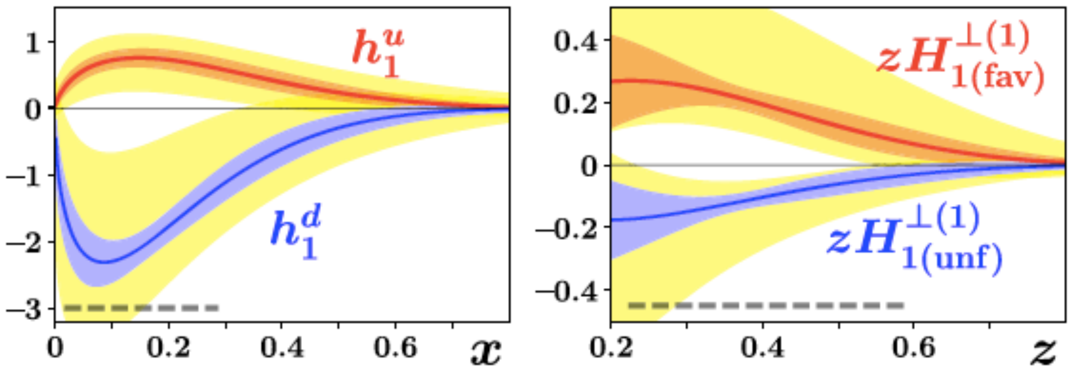
For precision calculations

- small lattice spacings
- Large Spatial Volumes
- Calculations at physical pion mass

We now know how to do the calculations. Exascale computing means we can!

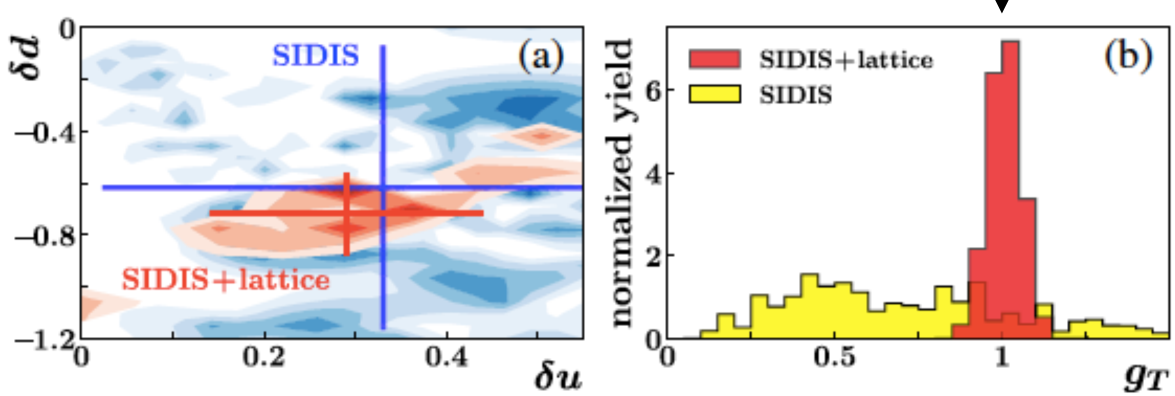
Experiment and Computation = Femtography

H-W Lin *et al.*, Phys. Rev. Lett. 120, 152502 (2018)

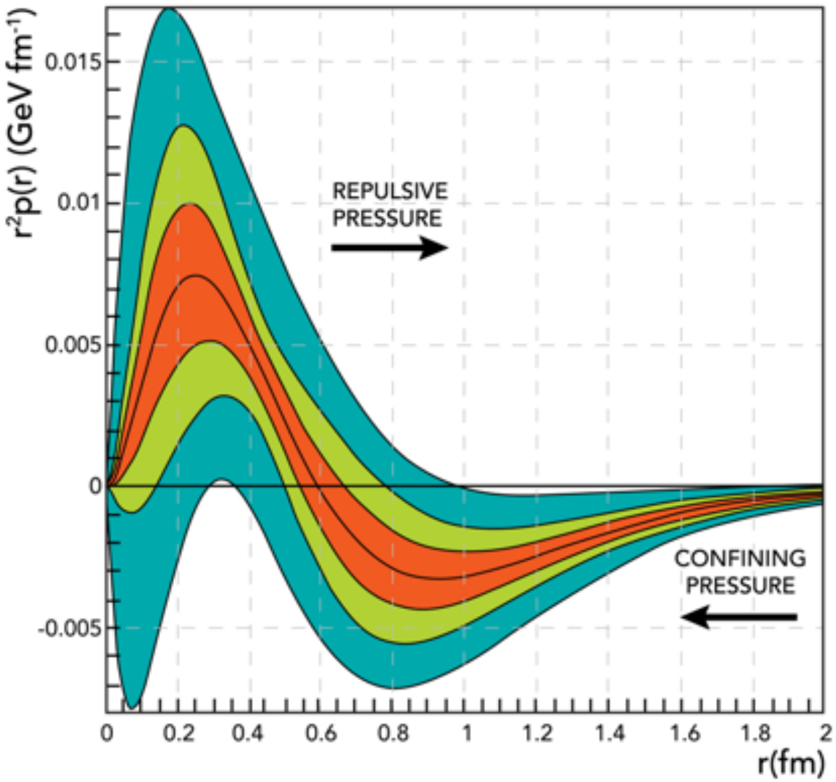


Transversity

Precisely calculated in LQCD



LQCD: not *testing* but *understanding* QCD



nature
International weekly journal of science

Nature 557 (2018) no.7705, 396-399

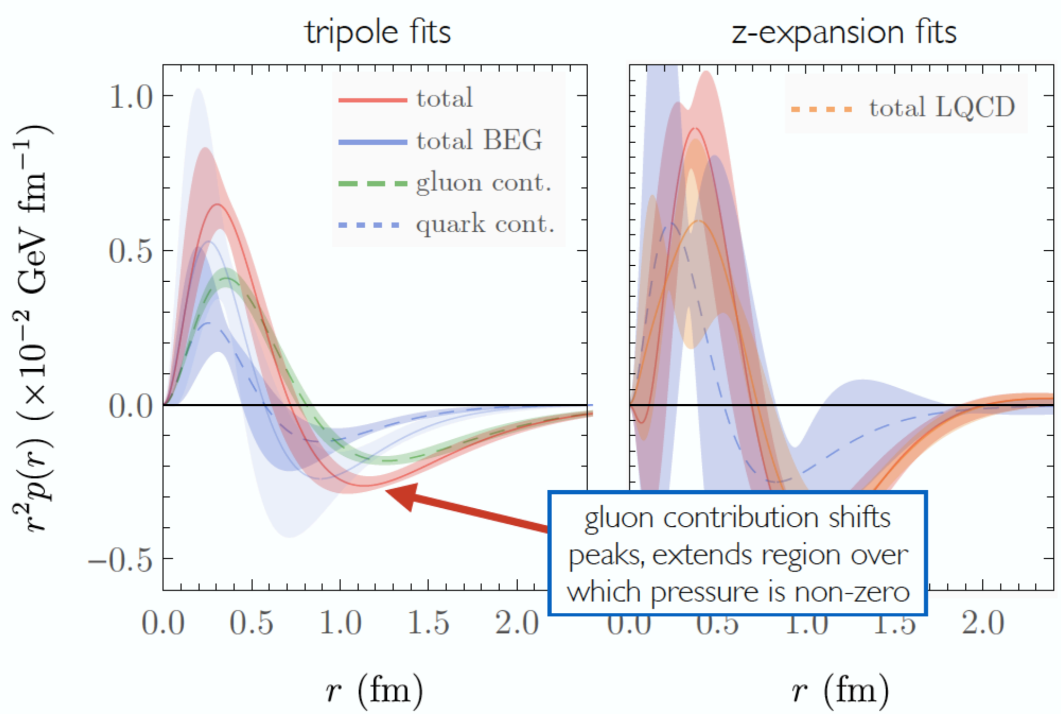


V. Burkert, L. Elouadrhiri, F.X. Girod

Expt + Lattice

P. Shanahan, EINN 2019

Nucleon pressure using LQCD results for gluon GFF, JLab results for quark GFF



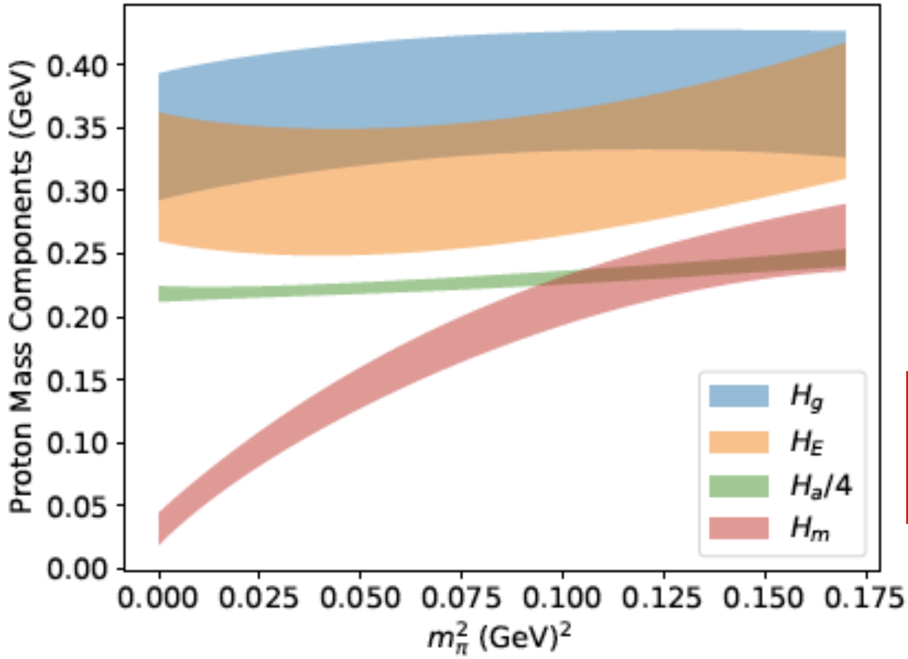
Gluon GFFs: Shanahan, Detmold, PRD99, 014511 & PRL122, 072003 (2019)
Quark GFFs: P. Hägler et al. (LHPC), PRD77, 094502 (2008)
Expt quark GFFs (BEG): Burkert et al, Nature 557, 396 (2018)

LQCD can predict and complement: gluon structure, pion,. . .

Energy-Momentum Tensor

$$T_{\mu\nu} = \frac{1}{4}\bar{\psi}\gamma_{(\mu}D_{\nu)}\psi + G_{\mu\alpha}G_{\nu\alpha} - \frac{1}{4}\delta_{\mu\nu}G^2; \langle P \mid T_{\mu\nu} \mid P\rangle = P_\mu P_\nu / M$$

$$\text{Trace Anomaly: } T_{\mu\mu} = -(1+\gamma_m)\bar{\psi}\psi + \frac{\beta(g)}{2g}G^2$$



Yang et al., Phys. Rev. Lett. 121, 212001 (2018)

How does mass decomposition change with quark mass?

Summary

- New era in hadron structure calculations driven by
  - New and upcoming experimental facilities
  - Theoretical advances
  - Approach to exascale computing
- To capitalize on this we need a coordinated effort of experiment, theory and computation
  - Exploit developments in machine learning
  - Visualization to analyse and to learn
  - Development of new algorithms and methods for computation.