Inverse problems in nuclear tomography

Nobuo Sato

EIC Software AI WG (Jun 2022)
INT Program 22-1

Machine Learning for Nuclear Theory

March 28 - April 22, 2022

Organizers:
- Gaute Hagen, Oak Ridge National Laboratory, hageng@ornl.gov
- Nobuo Sato, Thomas Jefferson National Accelerator Facility, nsato@jlab.org
- Phiala Shanahan, Massachusetts Institute of Technology, pshana@mit.edu

Diversity Coordinator:
- Gaute Hagen, Oak Ridge National Laboratory, hageng@ornl.gov

Program Coordinator:
- Megan Baunsgard, mjb47@uw.edu, (206) 685-4286
Parton showers and GANs

• The generator sequentially generates partons \( n \rightarrow n + 1 \)

\[
(p_1, \ldots, p_n) \xrightarrow{p_k} \text{time indep. NN} \xrightarrow{\text{softmax}} z_i, \phi_i \xrightarrow{p_k} (p_1, \ldots, p_{n+1}) \xrightarrow{\text{time dep. NN}} \xrightarrow{\text{softmax}} \theta_i \xrightarrow{p_{n+1}} \text{...}
\]

Shower history

Individual splitting

Postprocessing step

Lai, Ploskon, Neill, Ringer ‘20
2. Neutron Star: EoS $\leftrightarrow$ mass-radius relation

TOV equations:
\[
\frac{dP}{dr} = \frac{(m + 4\pi r^3 P)(P + \epsilon)}{r^2 - 2m r}, \\
\frac{dm}{dr} = 4\pi r^2 \epsilon, \\
\epsilon = \epsilon(P),
\]

Discrete \{M,R\} observations?
Lindblom's algorithm if the whole M(R) is known
Parton showers and GANs

2. Neutron Star: EoS $\leftrightarrow$ mass-radius relation

Nuclear mass models based on microscopic calculations

Advantage of the nuclear energy density functional

- Include as much as possible physical ingredients
- Several parameters that can be adjusted to nuclear data
- Reasonable computational time

Mass models

- Based on skyrme: HFB1 to HFB32 $\rightarrow$ best HFB27
- Gogny interaction: D1M $\rightarrow$ rms = 0.788 MeV
- All based on HFB (Hartree-Fock-Bogoliubov) oscillator basis

Objective

We want to go further by developing a mass model calculation.

Active learning method

1st method,

Committee of Multi-layer neural network

Additional proposed method:
Quantum correlation functions (QCFs) in Nuclear femtography

Parton distribution functions (PDFs)

Transverse momentum distributions (TMDs)

Generalized parton distributions (GPDs)
What do we mean by “hadron structure”? 

\[ \xi = \frac{k^+}{P^+} \]

Parton momentum fraction relative to parent hadron

\[ f_i(\xi) = \int \frac{d\omega^-}{4\pi} e^{-i\xi p^+ w^-} \langle N | \bar{\psi}_i(0, w^-, 0_T) \gamma^+ \psi_i(0) | N \rangle \]

in non-interacting QCD

\[ \psi_i(x) = \sum_{k, \alpha} b_{k, \alpha}(x^+) u_{k, \alpha} e^{-ik^+x^+ + ik_T \cdot x_T} + d_{k, \alpha}^\dagger(x^+) u_{k, -\alpha} e^{ik^+x^- - ik_T \cdot x_T} \]

\[ f_i(\xi) \sim \sum_{\alpha} \int d^2k_T \langle N | b_{k, \alpha}^\dagger b_{k, \alpha}(\xi p^+, k_T, \alpha) | N \rangle \]

parton distribution function (PDF)
How quarks and gluons are distributed?

\[ k \]

\[ \xi = \frac{k^+}{P^+} \]  
Momentum fraction

\[ f_i(\xi) \sim \sum_{\alpha} \int d^2k_T \langle N| b_{k,\alpha}^\dagger b_{k,\alpha}(\xi p^+, k_T, \alpha) |N \rangle \]

number operator
What do we mean by “hadronization”? 

\[ \zeta = \frac{p_h^+}{k^+} \]  

hadron momentum fraction relative to parent parton

\[ d_{h/j}(\zeta) = \frac{\text{Tr}_{\text{color,Dirac}}}{4N_{c,j}} \sum_X \zeta \int \frac{dw^+}{2\pi} e^{i(p_h^- / \zeta)w^+} \times \gamma^- \langle 0 | \bar{\psi}_j(0, w^+, 0_T) | p_h, X \rangle \langle p_h, X | \psi_j(0) | 0 \rangle \]

Fragmentation functions (FFs)

\[ X = \text{all states except detected hadron } h \]
How hadrons emerges from quarks and gluons

\[ \zeta = \frac{p_h^+}{k^+} \]

Number density of hadrons from parent parton

\[ d_{h/j}(\zeta) \]
Hadron structure in interacting theory

Definition of PDFs in field theory requires renormalization.

PDFs will depend on renormalization scale and its RGEs are the famous DGLAP equations.

Renormalization

$$f_i(\xi) = \int \frac{d\omega^+}{4\pi} e^{-i\xi p^+ w^-} \langle N | \bar{\psi}_i(0, w^-; 0_T) \gamma^+ \psi_i(0) | N \rangle$$

Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP)

$$\frac{df_i(\xi, \mu^2)}{d\ln \mu^2} = \sum_j \int_{\xi}^{1} \frac{dy}{y} P_{ij}(\xi, \mu^2) f_j \left( \frac{y}{\xi}, \mu^2 \right)$$

aka DGLAP
Spin structures

Unpol pdfs

Helicity distribution

Transversity

\[ f = f_\rightarrow + f_\leftarrow \]

\[ \Delta f = f_\rightarrow - f_\leftarrow \]

\[ \delta_T f = f_\uparrow - f_\downarrow \]

\[ \langle N|\bar{\psi}_i(0, \omega^-, 0_T)\gamma^+\gamma_5\psi_i(0)|N\rangle \]

\[ \langle N|\bar{\psi}_i(0, \omega^-, 0_T)\gamma^+\gamma_5\gamma^\perp\gamma_5\psi_i(0)|N\rangle \]

\[ \langle N|\bar{\psi}_i(0, \omega^-, 0_T)\gamma^+\gamma_5\psi_i(0)|N\rangle \]
Extensions to 3D

\[ f(\xi, b_T) \]

Impact parameter distribution $\rightarrow$ GPDs

\[ f(\xi) \]

PDFs

\[ f(\xi, k_T) \]

Transverse momentum distribution $\rightarrow$ TMDs
So how do we get hadron structure from experimental data?

Want to see internal structure

But we only see debris

The inverse problem

Factorization
Factorization in deep-inelastic scattering

\[ e^- \rightarrow e^- \quad e^- \rightarrow N \quad E' \frac{d\sigma}{d^3 l'} \]

Collision dependent factor

Internal structure

\[ \sum_i \int_x^1 \frac{d\xi}{\xi} H_i(\xi) f_i \left( \frac{x}{\xi} \right) + O \left( \frac{m^2}{Q^2} \right) \]

Error of approximations
Factorization in other reactions

SIA

$e^- \rightarrow \pi \rightarrow e^+$

$\sum_i H^\text{SIA}_i \otimes d_i$

hadronization

SIDIS

$e^- \rightarrow N \rightarrow \pi$

$\sum_{ij} H^\text{SIDIS}_{ij} \otimes f_i \otimes d_j$

structure + hadronization

..and many more
Universality

cross sections described by **universal non-perturbative** functions, e.g. PDFs, FFs
The Bayesian inference

Experiments = theory + errors

\[ d\sigma_{\text{DIS}} = \sum_i H_{i}^{\text{DIS}} \otimes f_i \]
\[ d\sigma_{\text{DY}} = \sum_{ij} H_{ij}^{\text{DY}} \otimes f_i \otimes f_j \]
\[ d\sigma_{\text{SIA}} = \sum_i H_{i}^{\text{SIA}} \otimes d_i \]
\[ d\sigma_{\text{SIDIS}} = \sum_{ij} H_{ij}^{\text{SIDIS}} \otimes f_i \otimes d_j \]

RGE boundary conditions (QCF modeling)

\[ f_i(\xi, \mu_0^2) = N_i \xi^{a_i} (1 - \xi)^{b_i} (1 + ...) \]
\[ d_i(\xi, \mu_0^2) = N_i \xi^{a_i} (1 - \xi)^{b_i} (1 + ...) \]
\[ a = (N_i, a_i, b_i, ...) \]

\[ \rho(a|\text{data}) \sim \mathcal{L}(a, \text{data}) \pi(a) \]

\[ \mathcal{L}(a, \text{data}) = \exp \left[ -\frac{1}{2} \chi^2(a, \text{data}) \right] \]

\[ \chi^2(a) = \sum_{i,e} \left( \frac{d_{i,e} - \sum_k r_{i,e}^{k} \beta_{i,e}^{k} - T_{i,e}(a)/N_e}{\alpha_{i,e}} \right)^2 \]

\[ E[f_i(\xi, \mu^2)] = \int d^n a \rho(a|\text{data}) f_i(\xi, \mu^2; a) \]
\[ V[f_i(\xi, \mu^2)] = \int d^n a \rho(a|\text{data}) [f_i(\xi, \mu^2; a) - E[f_i(\xi, \mu^2)]]^2 \]
How do we deal with the posterior?
\[ E[f_i(\xi, \mu^2)] = \int d^n a \, \rho(a|data) f_i(\xi, \mu^2; a) \]
\[ V[f_i(\xi, \mu^2)] = \int d^n a \, \rho(a|data) \left[ f_i(\xi, \mu^2; a) - E[f_i(\xi, \mu^2)] \right]^2 \]

Maximum likelihood  
+ Hessian  
+ Lagrange  

MC methods  
+ Data resampling  
+ Markovian approaches
Maximum likelihood (+Hessian)

\[
E\{\mathcal{O}(a)\} = \int d^n t \, p(t|m) \, \mathcal{O}(a(t)) \approx \mathcal{O}(a_0).
\]

\[
V\{\mathcal{O}(a)\} = \int d^n t \, p(t|m) \left[ \mathcal{O}(a(t)) - E\{\mathcal{O}(a)\} \right]^2
\]

\[
\approx \prod_k \int dt_k \, p\left(t_k \frac{e_k}{\sqrt{w_k}} \mid m\right) \left( \sum_l \left. \frac{\partial \mathcal{O}(a(t))}{\partial t_l} \right|_{a_0} t_l \right)^2
\]

\[
= \sum_k T_k^2 \left( \left. \frac{\partial \mathcal{O}(a(t))}{\partial t_k} \right|_{a_0} \right)^2, \quad \rightarrow \quad T_k^2 = \int dt_k \, p_k(t_k \mid m) \, t_k^2.
\]

\[
V\{\mathcal{O}(a)\} \approx \sum_k \frac{1}{4} \left[ \mathcal{O}\left(a_0 + T_k \frac{e_k}{\sqrt{w_k}}\right) - \mathcal{O}\left(a_0 - T_k \frac{e_k}{\sqrt{w_k}}\right) \right]^2
\]

Hunt-Smith, Accardi, Melnitchouk, NS, Thomas, White (in prep)
Data resampling

\[ d_{k,i}^{\text{(pseudo)}} = d_{i}^{\text{(original)}} + \alpha_i R_{k,i} \]

\[ E_{\text{freq}}\{\mathcal{O}(a)\} = \frac{1}{n_{\text{rep}}} \sum_{i=1}^{n_{\text{rep}}} \mathcal{O}(a_{\text{rep}}), \]

\[ V_{\text{freq}}\{\mathcal{O}(a)\} = \frac{1}{n_{\text{rep}}} \sum_{i=1}^{n_{\text{rep}}} \left[ \mathcal{O}(a_{\text{rep}}) - E_{\text{freq}}\{\mathcal{O}(a)\} \right]^2. \]
New combined analysis of PDFs and FFs including unidentified charged hadron SIDIS and SIA data. The update analysis from JAM19 indicates again the strong nucleon suppression.

First global analysis with lattice off-the-light cone matrix elements for polarized and unpolarized PDFs. Polarized lattice data compatible with experimental data.

Support for EIC yellow report including unpolarized and polarized nucleon PDFs, electroweak parameters, meson structure and TMD.

First global QCD analysis of polarized PDFs using small x evolution. The constrained small x indicates a strong preference for negative $g_1p$ at small x. Provides important guidance for EIC simulations.

PDF analysis with the inclusion of collider W/Z data and the MARATHON d/p, Helium, Triton DIS data. Evidence for iso-vector effects illuminating nuclear effects in light nuclei.

First global analysis of all SSA in TMD+CT3 framework. New constraints on nucleon tensor charges.

Including of RHIC W+/− data and Seaquest DY data. New constraints on antimatter asymmetry in the nucleon.

New results on pion PDFs providing the effective large x asymptotic and its theory uncertainties.
Current paradigm

Detector level events → Detector level histograms → Unfolded histograms → Deconvolution “femtography”

Inverse problem → Inverse problem

1D QCFs
Complexity

Parton distribution functions (PDFs)-1D

Transverse momentum distributions (TMDs)-2D

Generalized parton distributions (GPDs)-3D

3 dimensional histograms

Inclusive observables

Semi-inclusive observables

6 dimensional histograms

9 dimensional histograms

Exclusive observables

Spin degrees of freedom

Flavor degrees of freedom
Challenges

Increasingly difficult in higher dimensional observables

Detector level events → Detector level histograms → Unfolded histograms

- Subjected to theory bias
- Requires to remove detector effects
- Subjected to parametrization bias
- Deconvolution relies on an approximation, needs validation

Arbitrary choice of binning

Experimental domain

Theory domain

Detector level histograms

Unfolded histograms

Deconvolution “femtography”
Event-based analysis?

Can we compare real vs synthetic events?

Why?
- Maximize physics extraction
- Preserve correlations
- Avoid unfolding and use direct simulation

Optimize physics parameters
GANs
Application to inclusive DIS

\[ k + p \rightarrow k' + X \]

- A GANs to train a detector emulator
- Train particle generator using GAN detector
- Change of variables to improve discriminator

\[
\nu_1 = \ln \left( \frac{(k'_0 - k'_z)}{1\text{ GeV}} \right), \\
\nu_2 = \ln \left( \frac{(2E_e - k'_0 - k'_z)}{1\text{ GeV}} \right)
\]

Case 1: no detector effects
Detector GAN proxy

- Use simple detector parametrization (EICSmear)
- Train detector GAN proxy using EICSmear
Future: GAN+theory

Parameters Generator

ML posterior distribution sampler

Experimental Events

Event level Discriminator

Simulated Events

Optimize physics parameters
Summary/Outlook

- Even-level interpolators can be constructed using generative models
- Discriminators have the unique feature to compare data at the event-level
- ML unifies theory and experiment by solving inverse problems in hadronic physics at the event-level