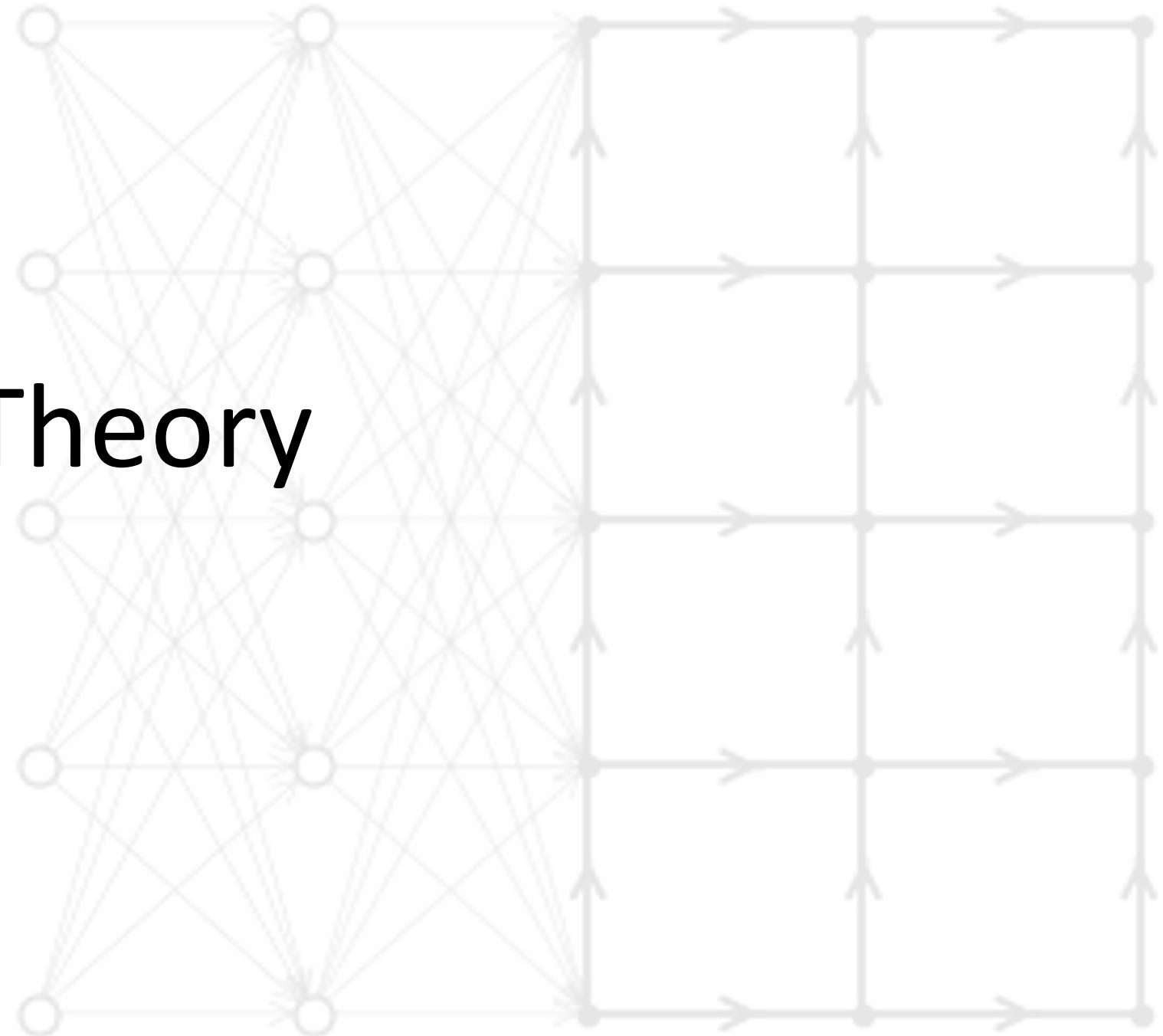


AI/ML Tools for Nuclear Theory

Dan Hackett (MIT / IAI FI)

QNP 2022

September 5, 2022



Overview

Machine learning & its role in theoretical physics

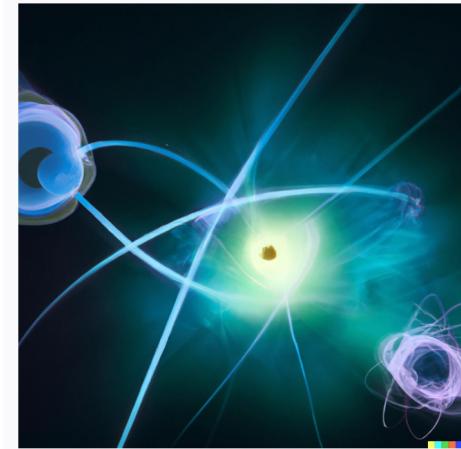
Ex 1: empirical parametrizations

Ex 2: variational methods

Lattice gauge theory & lattice QCD

Ex 3: accelerating observable measurements

Ex 4: generative models for LQCD



Disclaimer: not a review, lack of mention \neq criticism

Emphasis: deep learning / neural networks

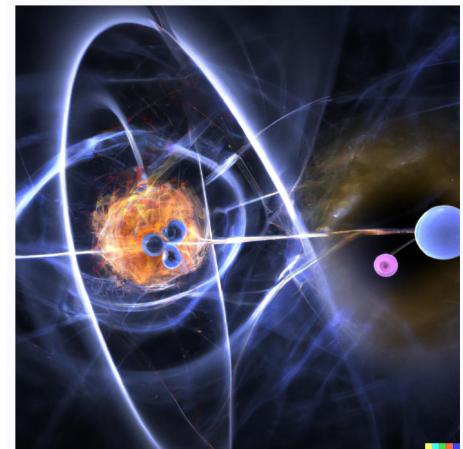
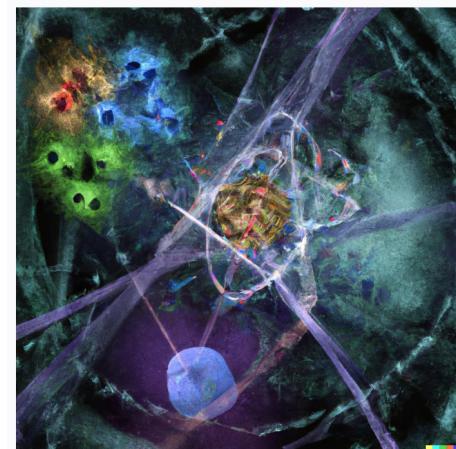
Not discussed (but also interesting!):

"Classical" ML methods

Bayesian methods

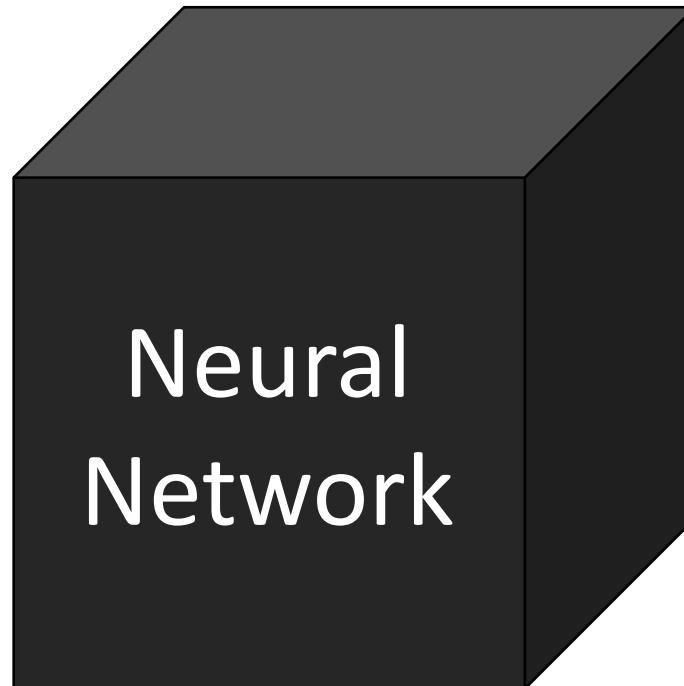
Calculation emulators

Theoretical physics for ML



DALL-E 2: "High-quality, digital art, the interior of a proton showing quarks and gluons interacting"

Machine learning



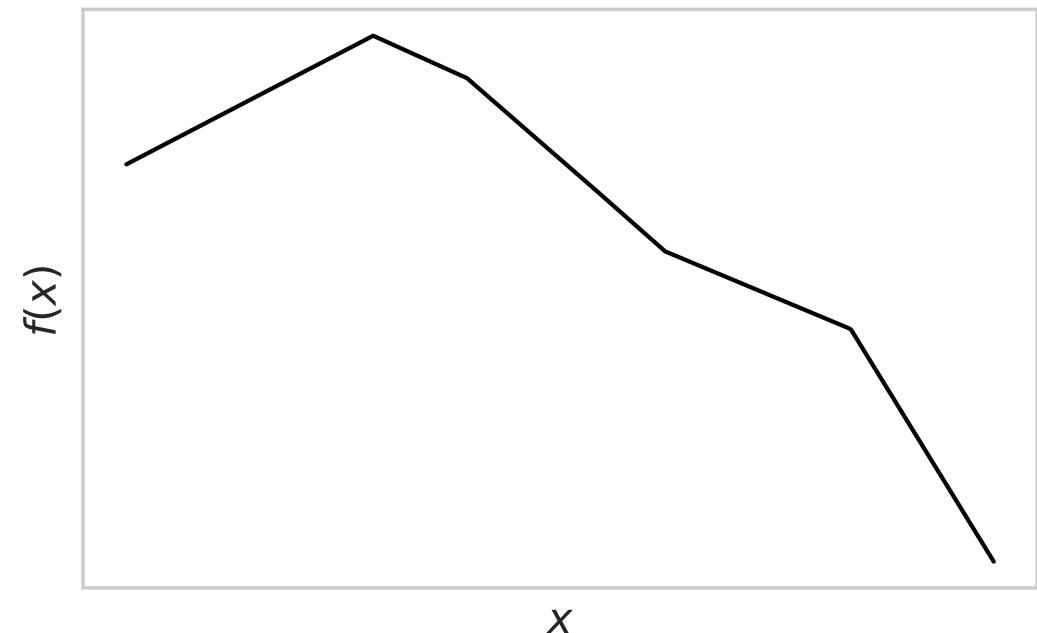
Neural networks are just functions

Feed-forward neural network:

- Apply linear transformation $x'_i = W_{ij} x_j + b_j$
 W_{ij}, b_j are free parameters
- Apply elementwise nonlinear activation
e.g. tanh, ReLU
- Repeat

Result: expressive, tunable function ansatz

**1d example: NN w/ ReLU activations
= piecewise linear function**



Example 1: empirical parametrizations

Lacking some physical model, can use NNs to model functional dependence (cf. polynomials, splines, etc.)

e.g. PDFs: $x f(x) = x^\alpha (1 - x)^\beta NN(x)$

[Collected works of NNPDF]

[Karpie, Orginos, Rothkopf, Zafeiropoulos 1901.05408]

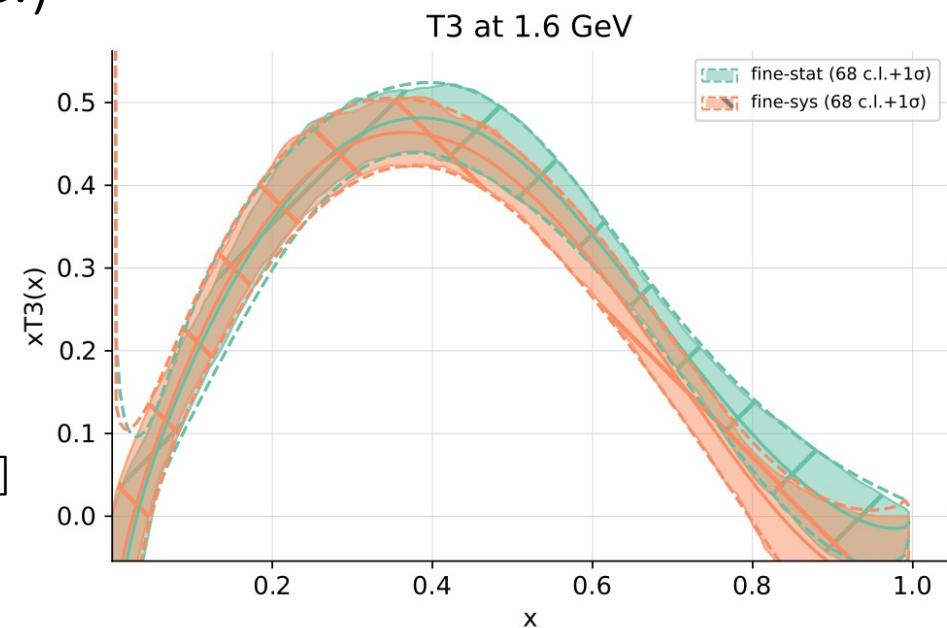
[Cichy, Del Debbio, Giani 1907.06037]

[Del Debbio, Giani, Karpie, Orginos, Radyushkin, Zafeiropoulos 2010.03996]

e.g. QCD EOS vs. neutron star mass-radius curve

$$P(\rho) = NN(\rho) \xrightarrow{TOV} M(R)$$

[Soma, Wang, Shi, Stöcker, Zhou 2201.01756]



Reproduced from [Del Debbio, Giani, Karpie, Orginos, Radyushkin, Zafeiropoulos 2010.03996]

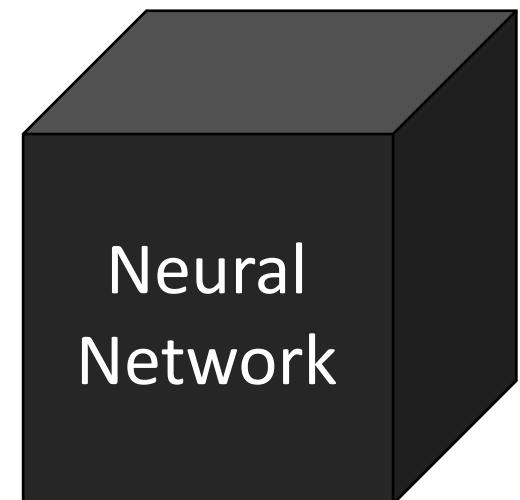
Neural networks as approximations

$$NN_{\text{untrained}}(x) \xrightarrow{\text{Some cleverly engineered training procedure}} NN_{\text{trained}}(x) \approx f(x)$$

For some f in our physics analysis/computation/calculation, ML methods can often provide a useful, high-quality approximation

However: approximation is *uncontrolled*

How can we use black boxes in theoretical physics?



Example 2: Variational methods

Wavefunctions: $\langle x|\psi \rangle : (\text{system dof } x) \rightarrow \mathbb{C}$

One possible ansatz: $\langle x|\psi \rangle = NN(x)$

Variational principle: $\frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_{gs}$

Wavefunctions for nuclear many-body systems

Need fermion antisymmetry \rightarrow extra structure

e.g. [Keeble, Rios 1911.13092]

e.g. [Adams, Carleo, Lovato, Rocco 2007.14282]

e.g. [Sun, Detmold, Luo, Shanahan 2202.03530]

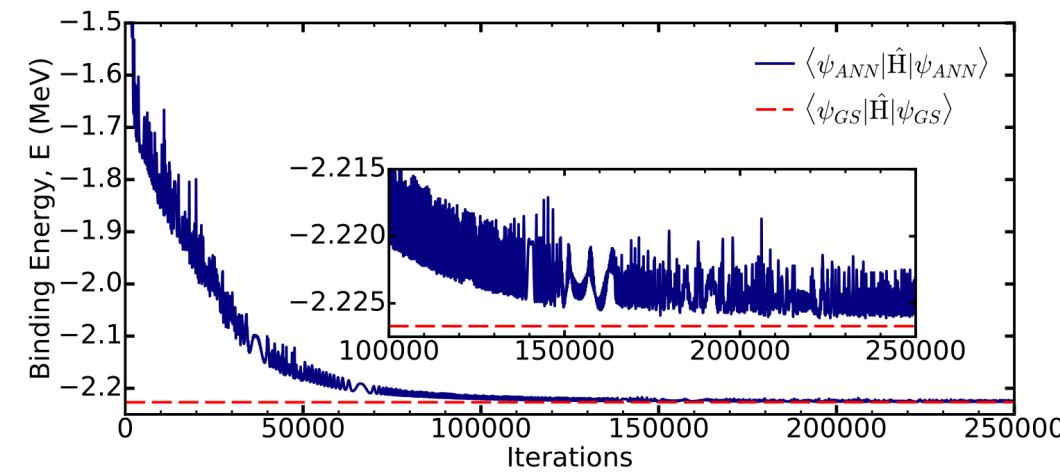
(many others)

Wavefunctionals for lattice gauge theories

Need gauge symmetry \rightarrow extra structure

[Luo, Carleo, Clark, Stokes 2012.05232]

Energy of variational state \sim loss
Minimization \sim training



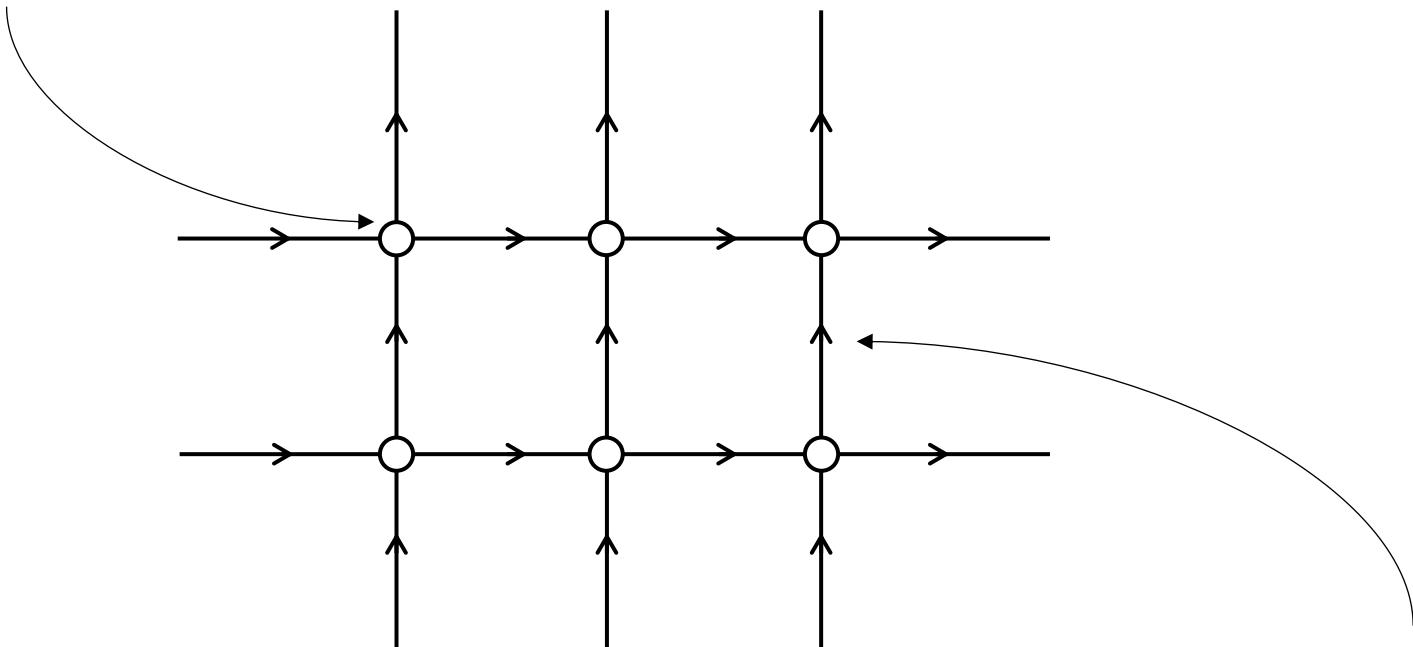
Reproduced from [Keeble, Rios 1911.13092]

Lattice gauge theories

Discretize field dof onto a spacetime lattice (so that it fits on a computer)

Matter fields on sites, e.g. per site:

- a real or complex number (scalar)
- a Grassmann variable (fermion)



Gauge fields on links, e.g. per link:

- a $U(1)$ phase (photons)
- an $SU(3)$ matrix (gluons)

Lattice QCD

Compute discretized QCD path integral on supercomputers w/ Monte Carlo

$$\langle \mathcal{O} \rangle = \int (\Pi_{\mathbf{x}} d\phi_{\mathbf{x}}) \frac{1}{Z} e^{-S_E(\phi)} \mathcal{O}(\phi)$$

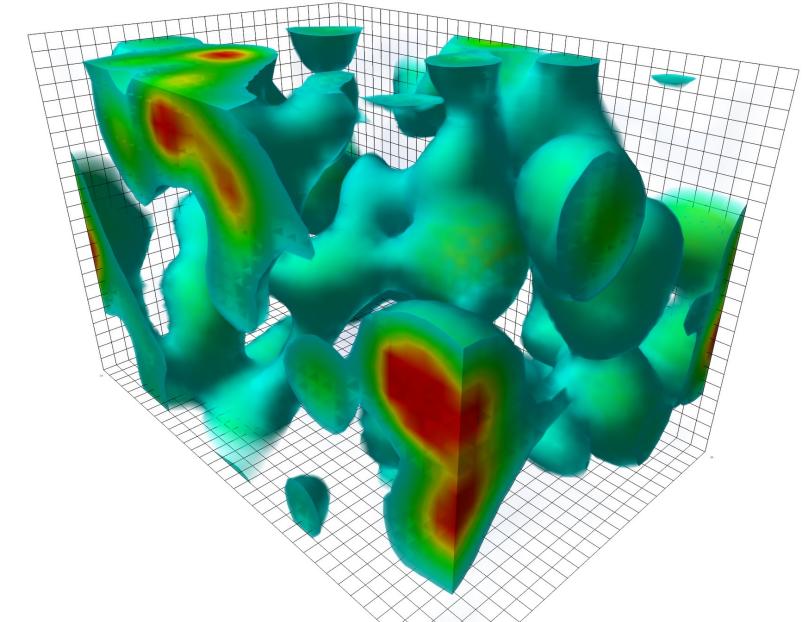
Only known controlled, systematically improvable way to investigate non-perturbative dynamics of QCD at low energies

Better algorithms → better physics

e.g. muon $g - 2$

e.g. larger nuclei

e.g. PDFs / GPDs / etc



[Visualizations of Quantum Chromodynamics](#)

LQCD Calculation Workflow

1. Generate gauge configurations

Use MCMC to draw samples U from $p(U) \propto e^{-S_E(U)}$

2. Measure observables

Evaluate observables O on each config U

$$\langle O \rangle \approx \frac{1}{N} \sum_{i=1}^N O(U_i)$$

MCMC expectations

Asymptotically unbiased
⇒ Exactness guaranteed

3. Analyze to extract physics

ML applications being explored throughout all parts of LQCD workflow

[Boyda, Calì, Foreman, Funcke, DH, Lin, Aarts, Alexandru, Jin, Lucini, Shanahan 2202.05838]

Example 3: Accelerating LQFT observable measurements

e.g. learned observable approximations

$$\langle O \rangle = \langle \tilde{O} \rangle_{UD} + \langle o - \tilde{o} \rangle_{BC}$$

- Learn cheap approximation \tilde{O} of expensive O on training data
- Measure O and \tilde{O} on common configs BC to compute bias correction
- Compute inexpensive \tilde{O} on unlabelled data UD

[Yoon, Bhattacharya, Gupta 1807.05971]

[Zhang, Fan, Li, Lin 1909.10990]

Example 3': Accelerating LQFT observable measurements

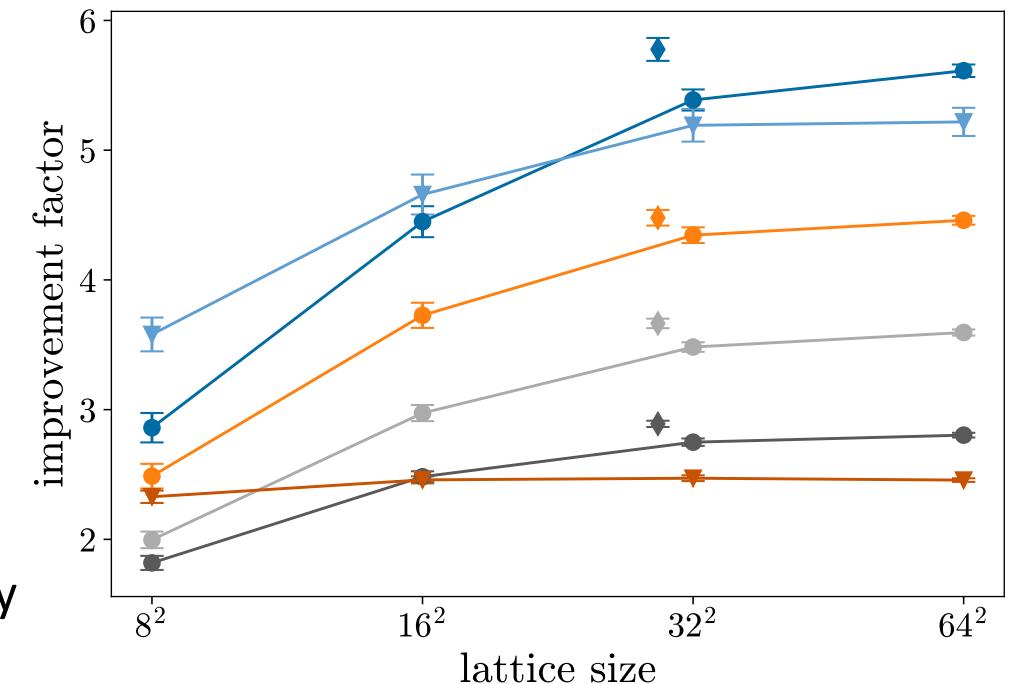
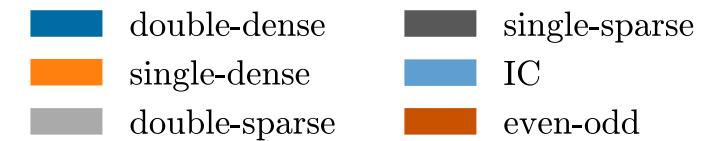
e.g. ML preconditioners for the conjugate gradient algorithm

Dominant cost in LQCD: solve $Ax = b$

Preconditioning: solving $M\bar{A}x = M\bar{b}$ cheaper for well-chosen M

Can use ML to produce preconditioners M

[Calì, DH, Lin, Shanahan, Xiao 2208.02728]

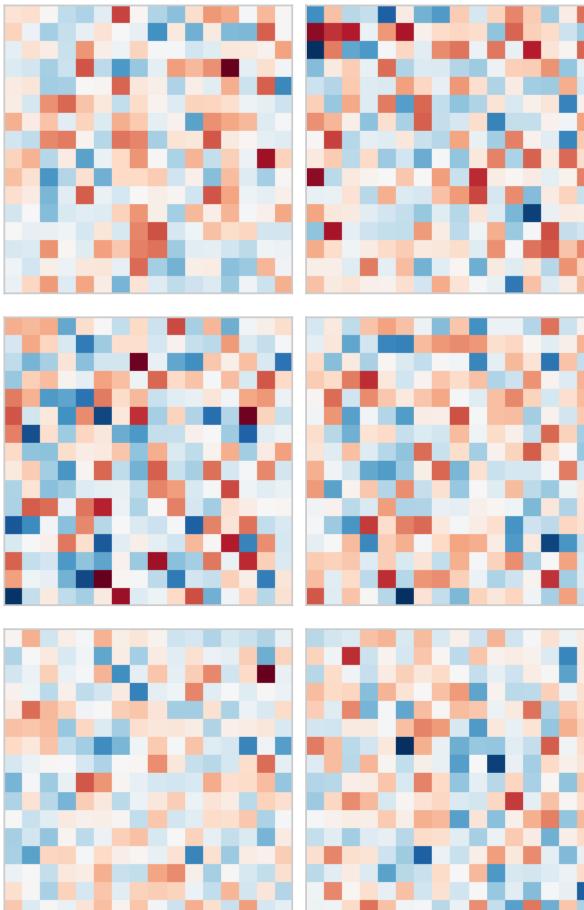


Result: learning on small volumes
transfers to larger volumes → scalability

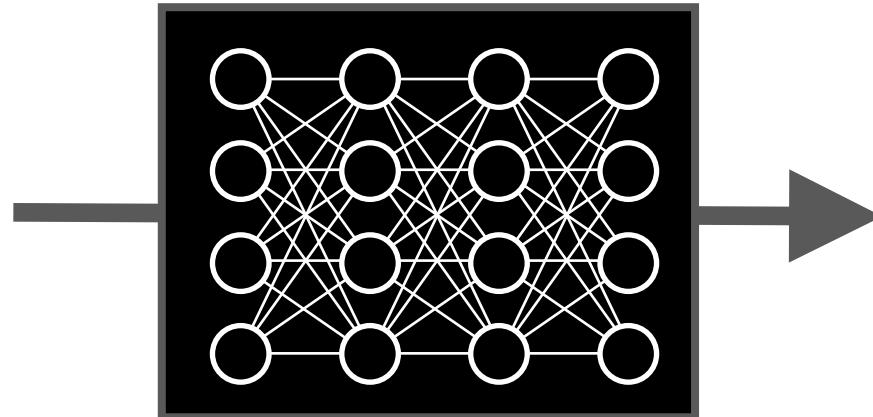
Example 4: Generative models for LQFT sampling

$z \sim \text{Gaussian noise}$

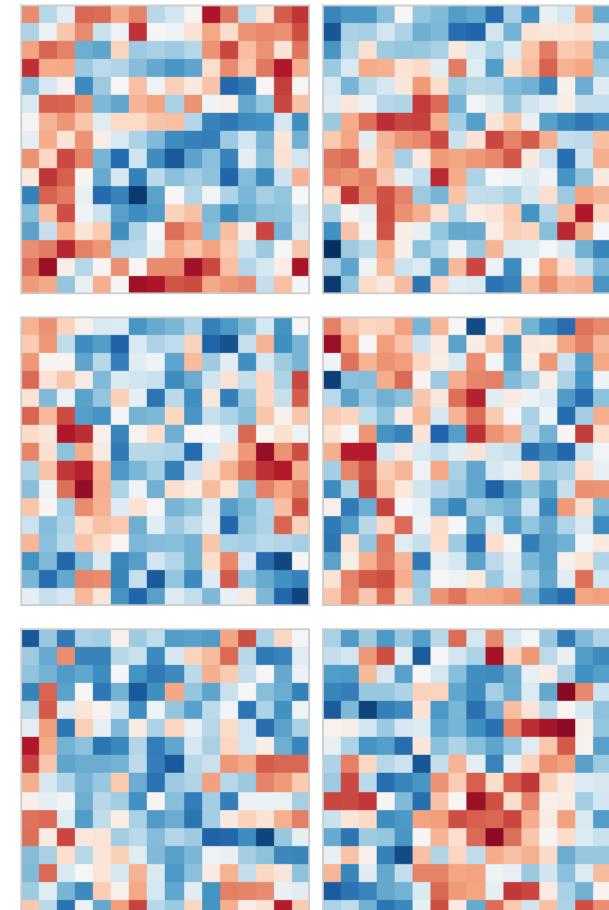
$$r[z] = \prod_{\mathbf{x}} \frac{1}{\sqrt{2\pi}} e^{-z(\mathbf{x})^2/2}$$



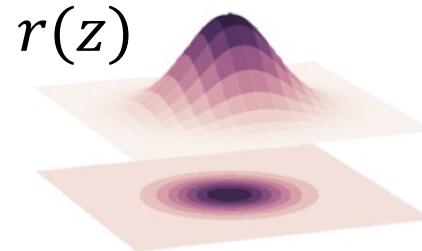
Learned transformation
 $\phi = f(z)$



$\phi \sim \text{model for distribution of lattice field configurations}$
i.e. $\phi \sim q \approx p = e^{-S(\phi)} / Z$



Normalizing flows



Simple **base distribution** r

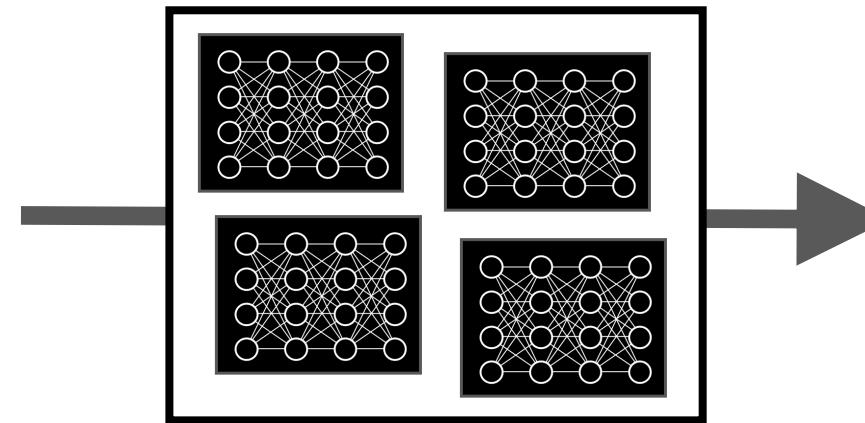
- Tractable density $r(z)$
- $r(z) > 0$ everywhere

"Flow" transformation

- *Parametrized* by NNs

By construction:

- Tractable Jacobian determinant
- Invertible



Learned **model distribution** q

$$q(\phi) = \left| \det \frac{\partial \phi}{\partial z} \right|^{-1} r(z)$$

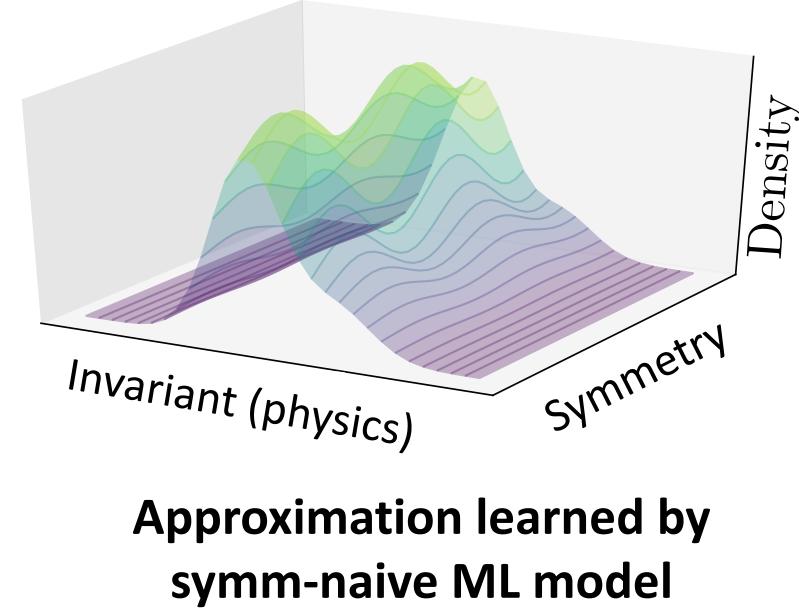
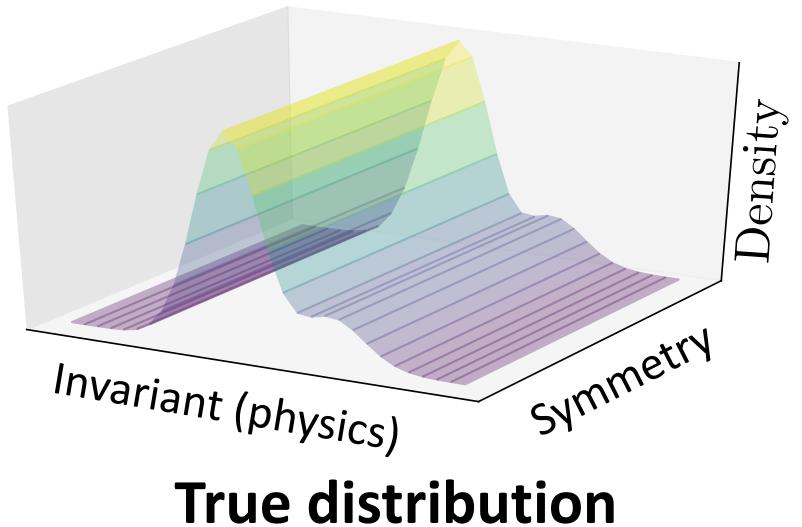
Can compute "reweighting factors" $\frac{p(\phi)}{q(\phi)}$ to correct for approximation and recover exactness!

Equivariant flows and invariant models

Invariance under $\phi \rightarrow \phi'$ means $p(\phi') = p(\phi)$

General approach:

1. Invariant base distribution $r(\phi') = r(\phi)$
2. Equivariant flow $f(\phi') = f(\phi)'$
i.e. flow and symmetry commute



The road to QCD

Flows for LQFT (scalar field theories)

[Albergo, Kanwar, Shanahan 1904.12072]

[DH, Hsieh, Albergo, Boyda, JW Chen, KF Chen, Cranmer, Kanwar, Shanahan 2107.00734]

Gauge-equivariant flows

U(1) [Kanwar, Albergo, Boyda, Cranmer, DH, Racanière, Rezende, Shanahan 2003.06413]

SU(N) [Kanwar, Albergo, Boyda, Cranmer, DH, Racanière, Rezende, Shanahan 2003.06413]

Flows for fermionic theories

Yukawa model [Albergo, Kanwar, Racanière, Rezende, Urban, Boyda, Cranmer, DH, Shanahan 2106.05934]

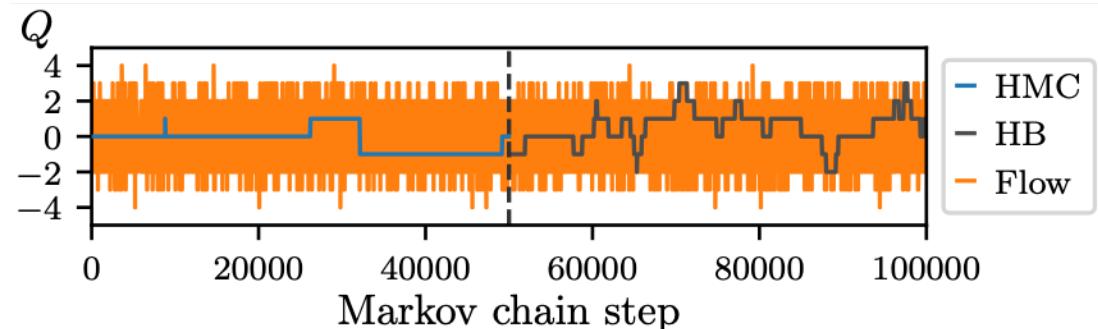
Schwinger model [Albergo, Boyda, Cranmer, DH, Kanwar, Racanière, Rezende, Romero-López, Shanahan, Urban 2202.11712]

Stochastic methods for fermionic gauge theories [Abbott, Albergo, Boyda, Cranmer, DH, Kanwar, Racanière, Rezende, Romero-López, Shanahan, Tian, Urban 2207.08945]

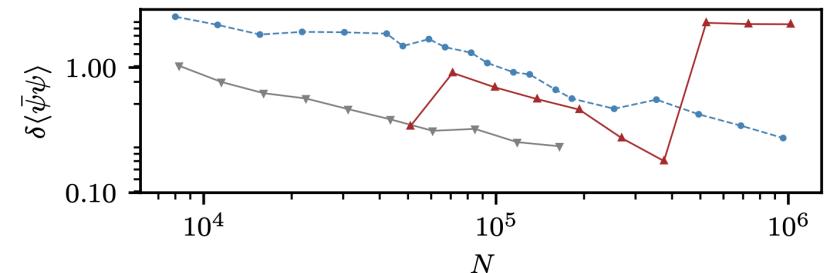
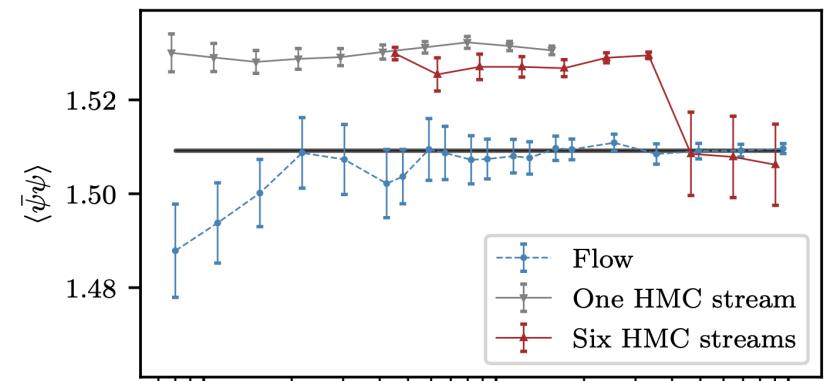
QCD!

Early demonstration [Abbott, Albergo, Botev, Boyda, Cranmer, DH, Kanwar, Matthews, Racanière, Razavi, Rezende, Romero-López, Shanahan, Urban 2208.03832]

Result: Improved sampling in (1+1)d U(1) gauge theory



Result: controlled results in Schwinger model where traditional algorithms fail



Access to new kinds of physics?

e.g. unbiased estimates of partition functions w/ flows

$$Z \approx \frac{1}{N} \sum_i \frac{e^{-S(U_i)}}{q(U_i)}$$

Example use case: *direct* estimation of thermodynamic observables / EOS

via $F = -T \log Z$

[\[Nicoli, Anders, Funcke, Hartung, Jansen, Kessel, Nakajima, Stornati 2007.07115\]](#)

cf. non-equilibrium thermodynamics [\[Caselle, Cellini, Nada, Panero 2201.08862\]](#)

e.g. contour deformations for numerical sign problems

Access to finite density / real time dynamics?

Better signal for nuclear correlation functions?

[\[Alexandru, Başar, Bedaque, Warrington 2007.05436\]](#)

[\[Lawrence, Yamauchi 2101.05755\]](#)

[\[Detmold, Kanwar, Lamm, Wagman, Warrington 2101.12668\]](#)

Conclusion

ML involves black boxes, but those can be useful for theory

Many different applications possible & under exploration

Reviews: 2006.05422 2112.02309

ML may enable new kinds of calculations & new physics results

Creativity necessary to find more opportunities!