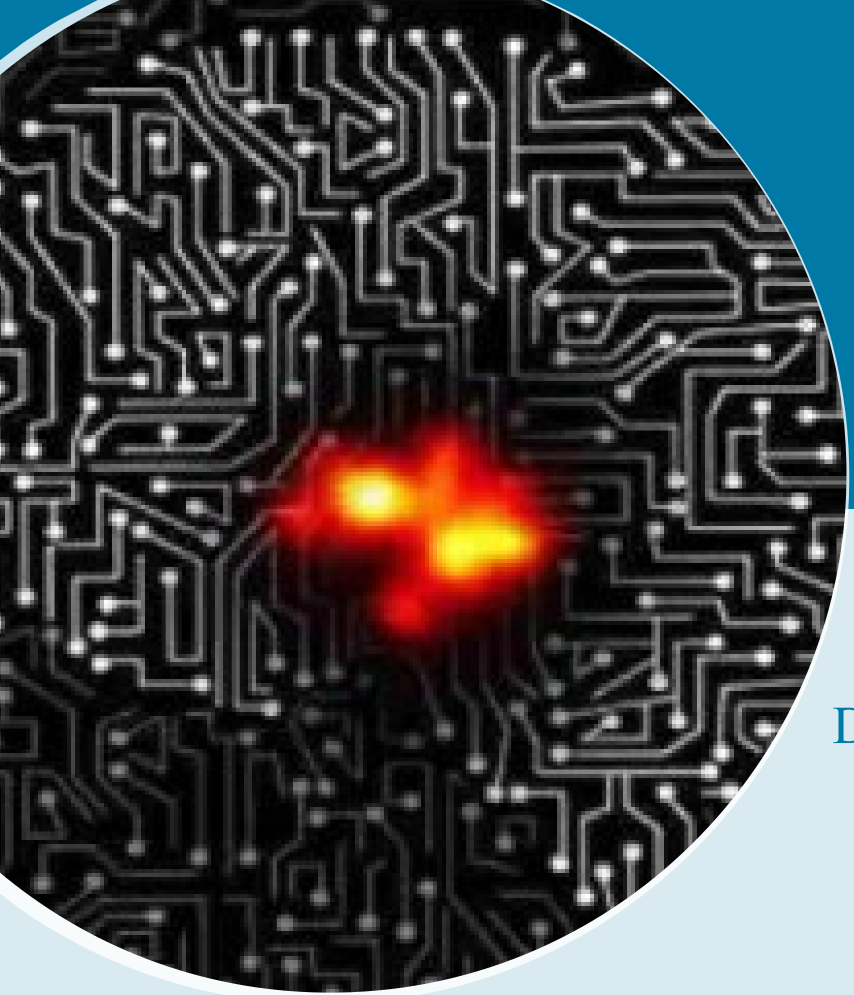


香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

High Energy Nuclear Physics meets Discriminative and Generative AI

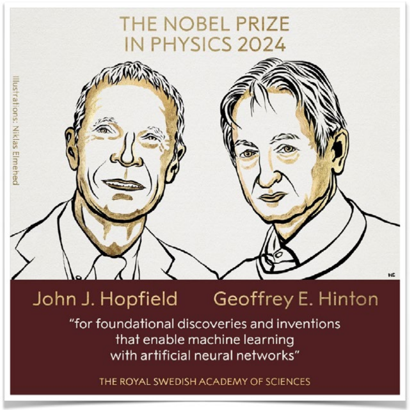
Kai Zhou (CUHK-Shenzhen)

第4届核物理与核数据中的机器学习应用研讨会， 南华大学，衡阳，湖南 2025

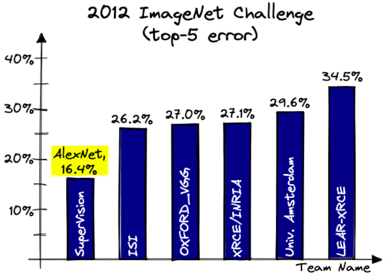
Overview : Nuclear Physics meets Machine Learning

- **2012** : Discovery of Higgs boson

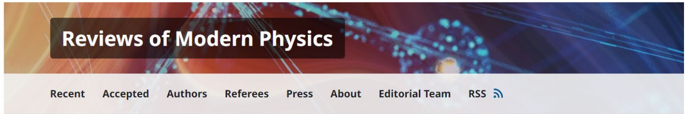
- **2024** : Nobel Prize in Physics



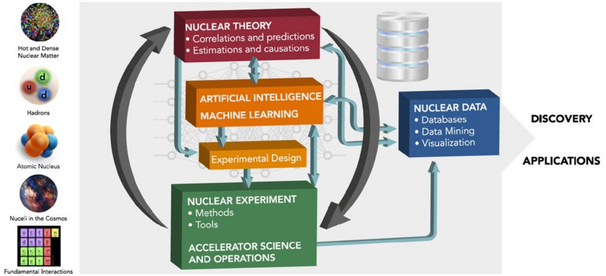
- AlexNet - Birth of Deep Learning



- ML4Physics, AI4Science

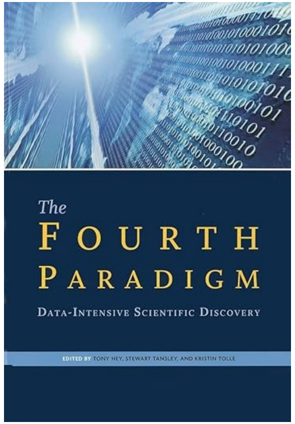


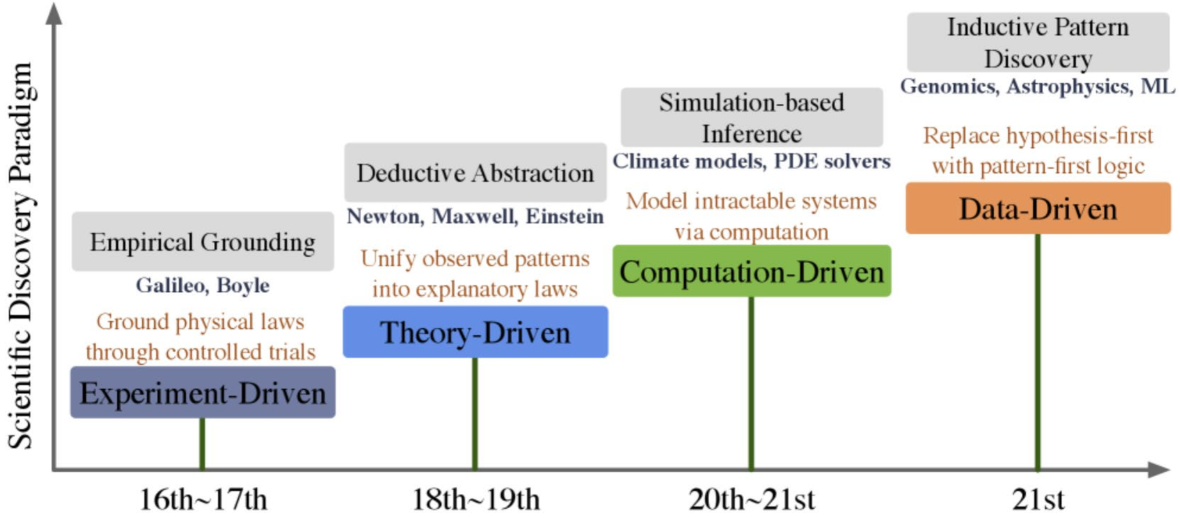
Colloquium: Machine learning in nuclear physics

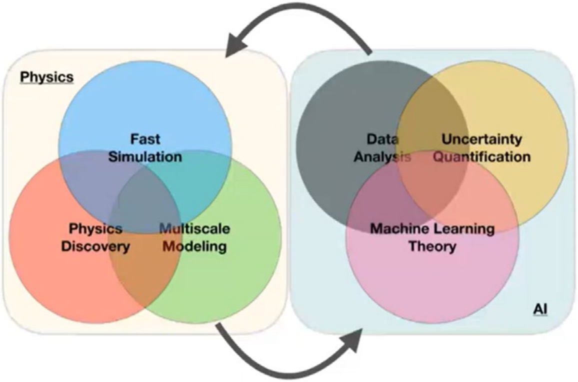


Overview : Fifth paradigm for scientific discovery

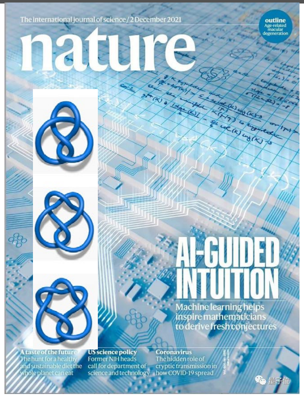
- Scientific discovery were driven by series of methodological paradigm evolution
- Machine Learning & Synthetic data forms the fifth methodological paradigm

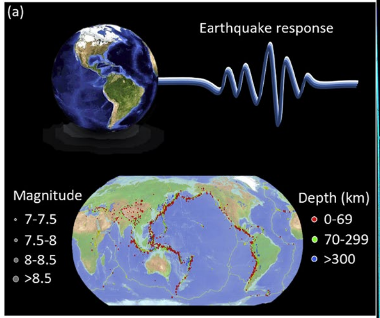


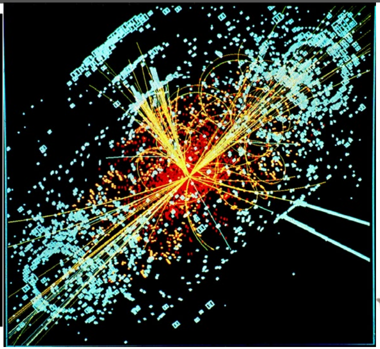


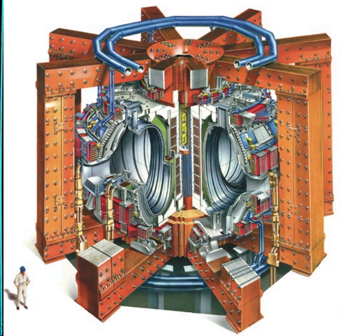


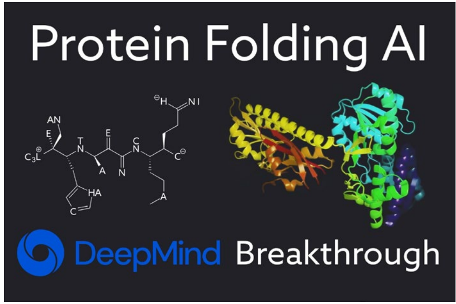
Overview : Machine- and Deep-Learning

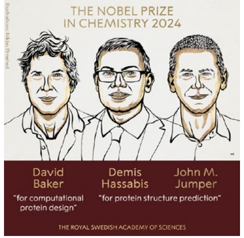








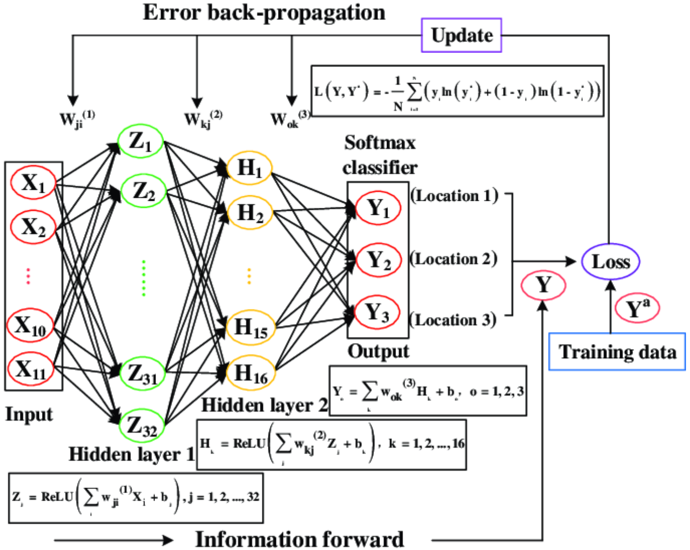




Find and Decode the mapping/representations into Deep Neural Network

➡ **Function approximator**

Universal approximator (Hastad et al 86 & 91)



Differentiable programming

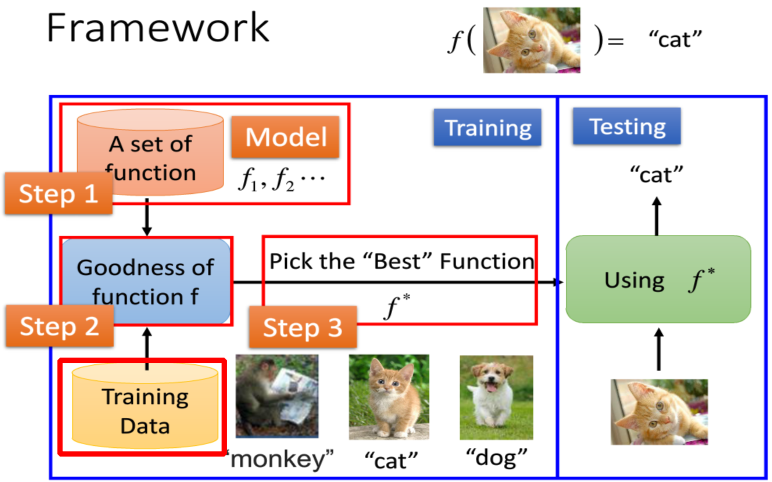
Backward Propagation

Gradient Descent Algorithm

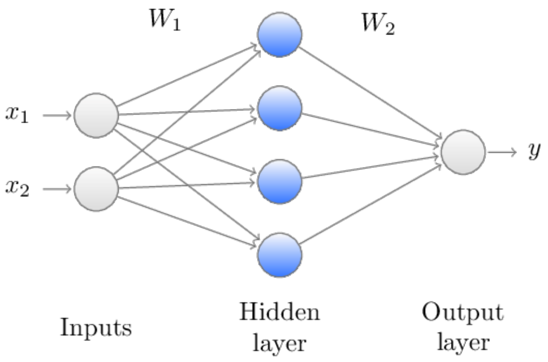
Overview: Machine learning, Deep Neural Networks, Representation learning

Framework

Image Recognition:



modified from Hung-Yi.Lee



Composing :

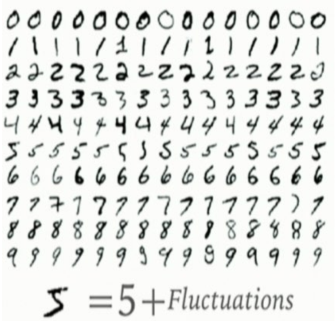
Linear affine transformation

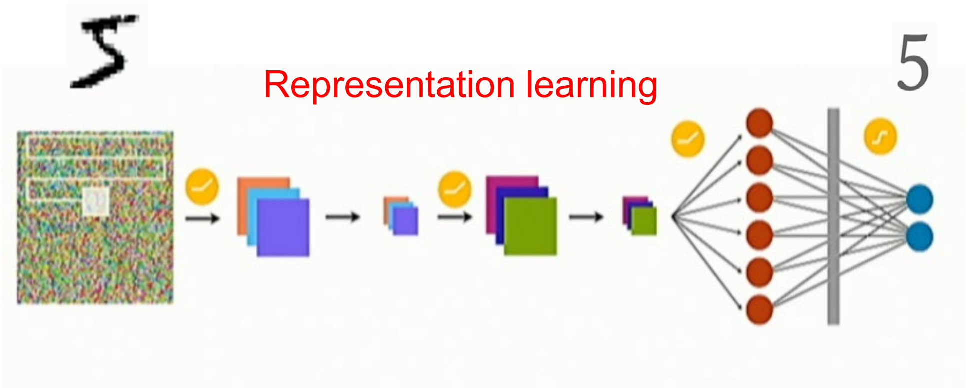
+

Non-linear activation

Layer by layer

$$f_{NN}(x;\theta) = h_2(w_2 h_1(w_1 x + b_1) + b_2)$$





Discriminative / Generative ML

- Discriminative Learning : **prediction**

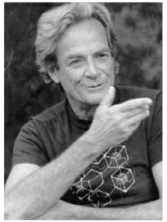
function fitting $y = f(x)$

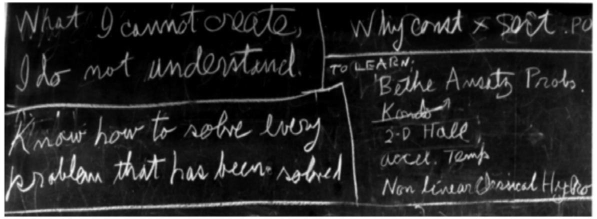
conditional probability $p_\theta(y|x) \to p(y|x)$

- Generative Modelling : **understand**

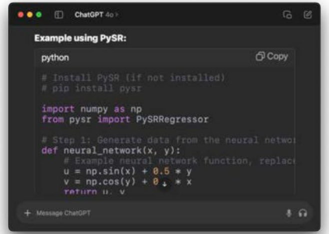
Joint probability distribution $p_\theta(x, y) \to p(x, y)$

$p(X)$ → pixels, words, atoms, ...





DaLL-E

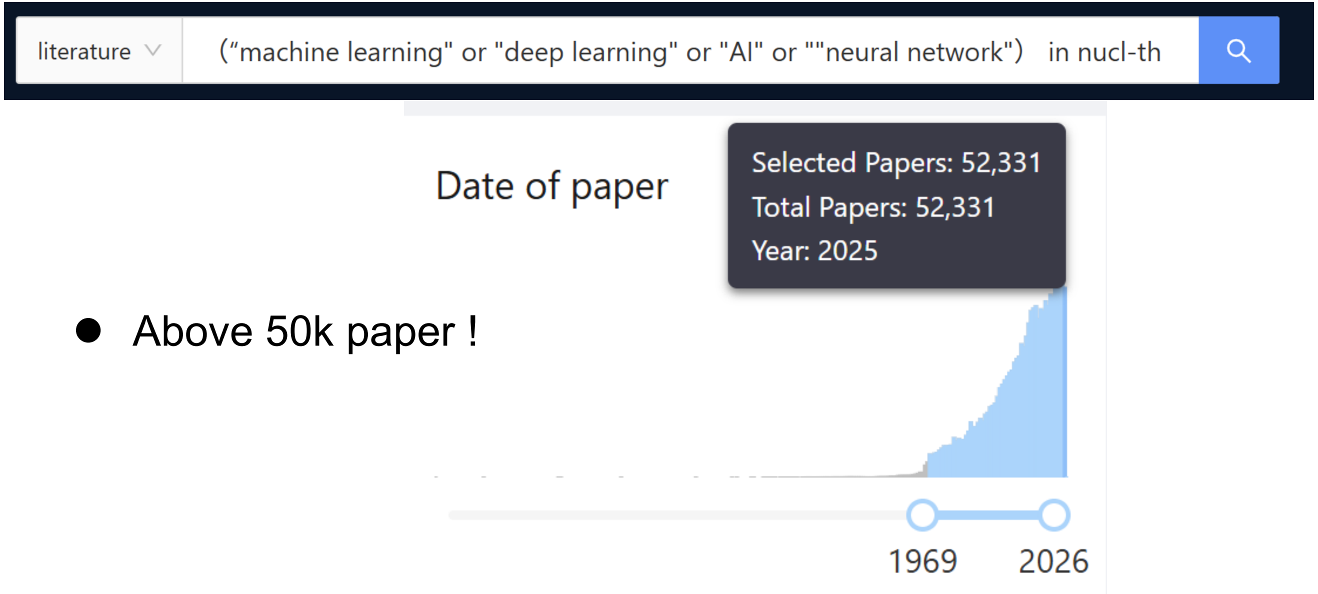


ChatGPT

AlphaFold3

"What I can not create, I do not understand"

Overview : super biased introduction in this talk in following

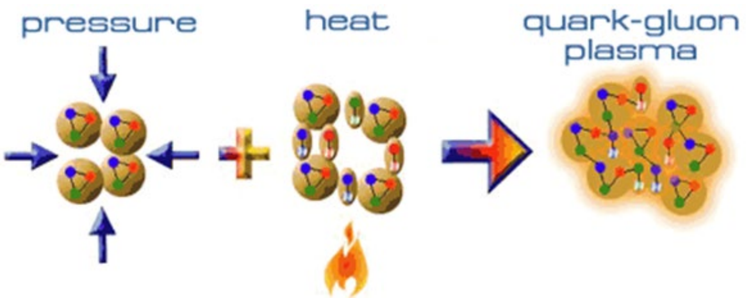


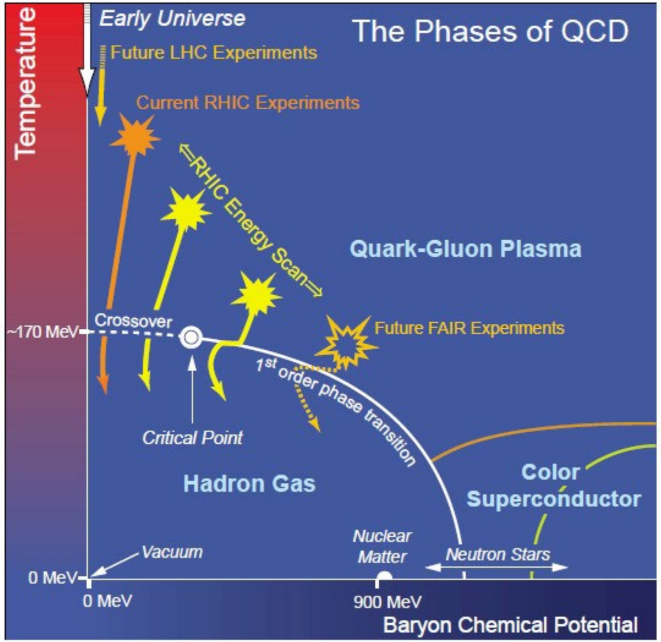
- Above 50k paper !

- This talk will selectively (biasedly) focus on high energy nuclear physics studies with:
- Discriminative ML to Physics Unfolding and Generative ML for Physics obs Generation

- **Phases** of matter : solid, liquid, gas, plasma

- Matter in extreme conditions reveals its **constituents** :
  nuclear matter → quark matter





To study **the most elementary particle matter** :

- **Nuclear Collisions** : heat & compress matter
- **Neutron Star** : dense matter, astronomy constraints
- **Lattice Field Theory / fQCD / Effective models**

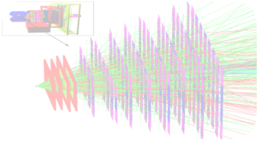
Outline

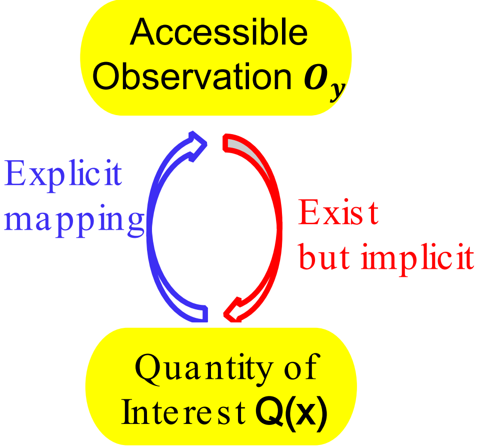
- **Discriminative AI** for unfolding physics in high-energy nuclear physics

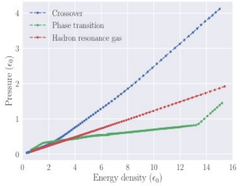
  physics priors need to be embedded in solving inverse problems

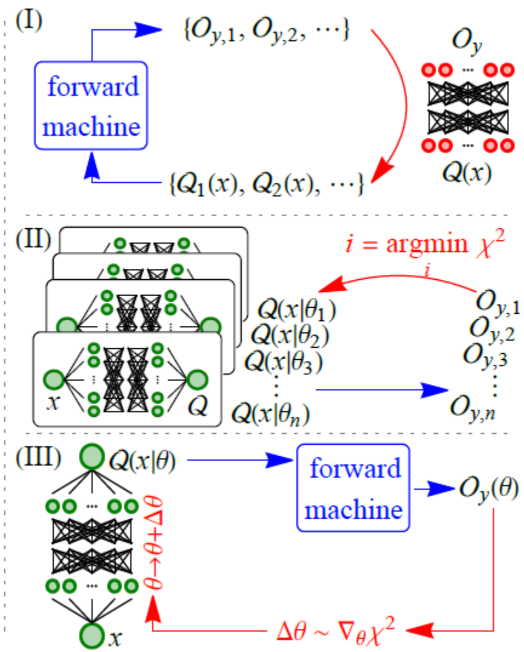
- **Generative AI** for speeding up physics simulations (IQFT, HIC modelling)

Inverse Problems Solving with discriminative ML







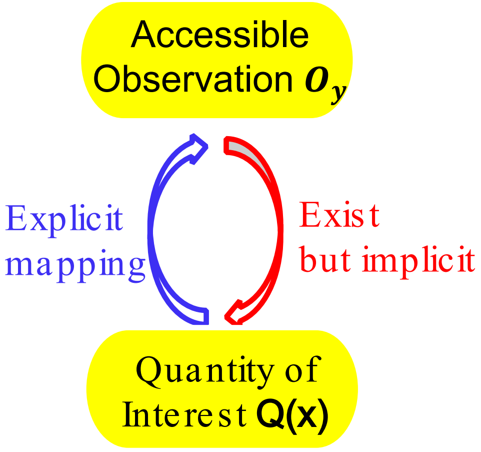


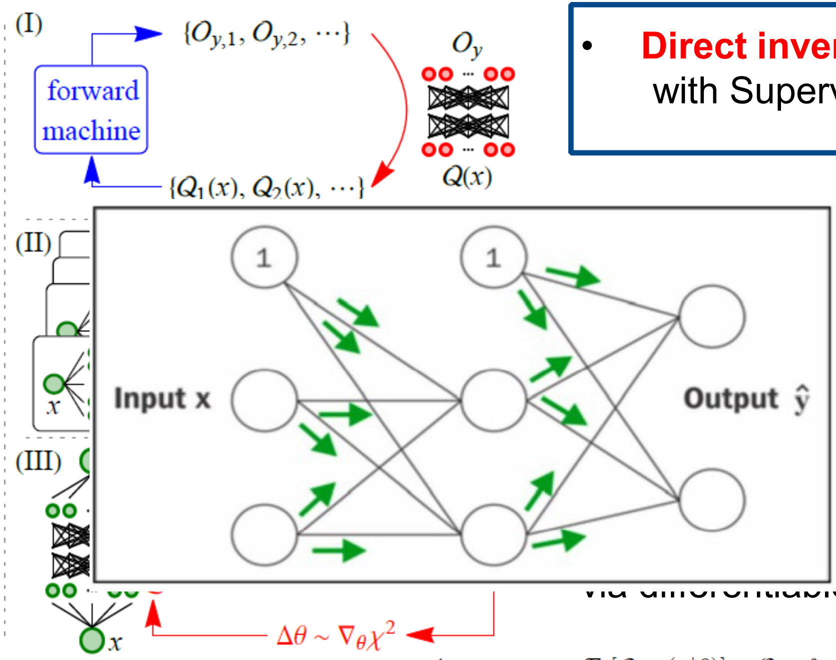
- **Direct inverse mapping capturing :** with Supervised Learning
- **Statistical approach to χ^2 fitting :** Bayesian Reconstruction for posterior or Heuristic (Generic) Algorithm to min.

$$\chi^2 = \sum_y \left(\frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{\Delta \mathcal{O}_y} \right)^2$$

- **Automatic Differentiation :** fuse physical prior into reconstruction via differentiable programming strategy

$$\frac{1}{2}\nabla_{\boldsymbol{\theta}}\chi^2 = \sum_y \frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{(\Delta \mathcal{O}_y)^2} \int \mathrm{d}x \frac{\delta \mathcal{F}_y[\mathcal{Q}(x)]}{\delta \mathcal{Q}(x)}\bigg|_{\mathcal{Q}(x)=\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}} \mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})$$





- **Direct inverse mapping capturing :** with Supervised Learning

approach to χ^2 fitting :
Reconstruction for posterior
neric) Algorithm to min.

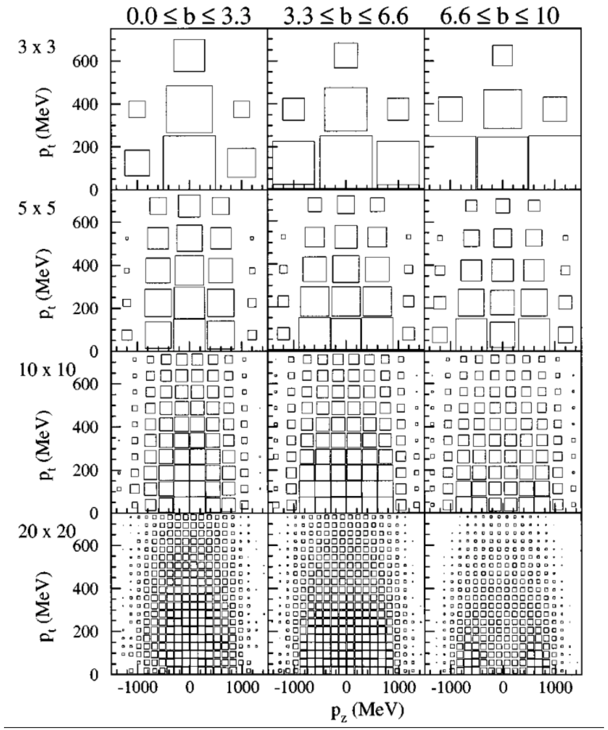
$$[\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})] - \mathcal{O}_y \over \Delta\mathcal{O}_y)^2$$

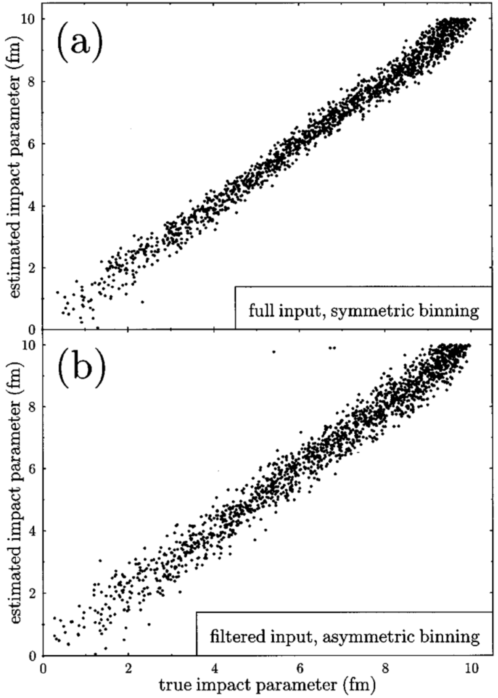
Differentiation :
al prior into reconstruction
via differentiable programming strategy

$$\frac{1}{2}\nabla_{\boldsymbol{\theta}}\chi^2 = \sum_y \frac{\mathcal{F}_y[\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{(\Delta\mathcal{O}_y)^2} \int \mathrm{d}x \frac{\delta\mathcal{F}_y[\mathcal{Q}(x)]}{\delta\mathcal{Q}(x)}\bigg|_{\mathcal{Q}(x)=\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}}\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})$$

Early attempts : impact parameter determination

Simple DNN
Trained on
QMC data
Input 5X5



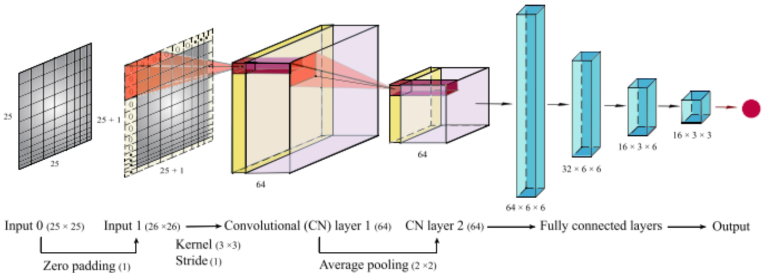


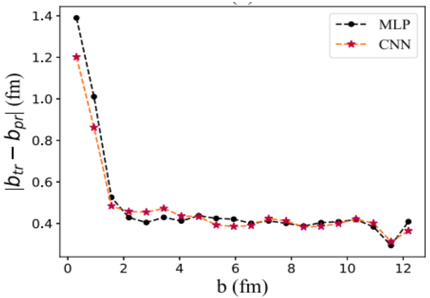
S. A. Bass, A. Bischoff, J. A. Maruhn, H. Stöcker, and W. Greiner, Phys. Rev. C 53, 2358 (1996)

Further dev for impact parameter determination

P. Xiang, Y. Zhao, X. Huang,
Chi. Phys. C 53, 2358 (2022)

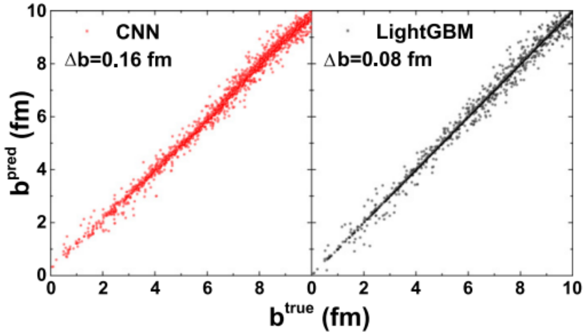
MLP and **CNN**
(on AMPT event)

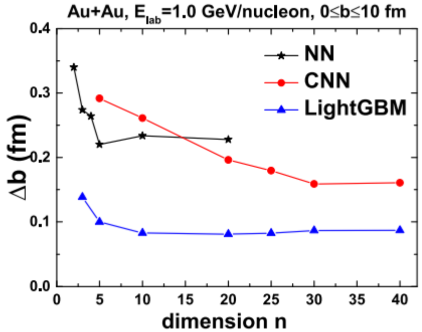




F. Li, Y. Wang, H. Lue, P. Li,
Q. Li, F. Liu, **JPG** 47, 115104 (2020)

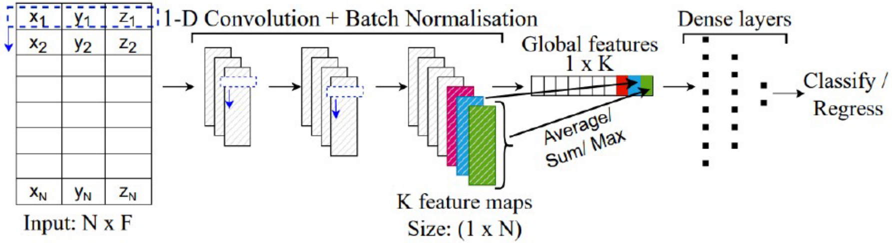
CNN and **LightGBM**
(on UrQMD event)

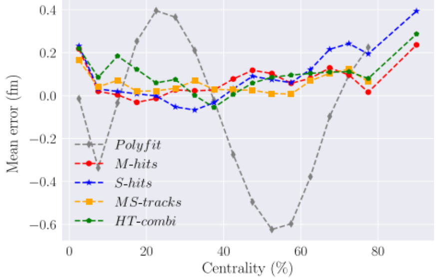




M. OK, J. S, K. Z, H. S,
PLB 811,135872 (2020)

PointCloud Network
(on UrQMD + CBMRoot event)
End-to-end b estimation





Manjunath O.K. and Kai Zhou, etc. *Phys.Lett.B* 811 (2020) 135872; JHEP10(2021)184.

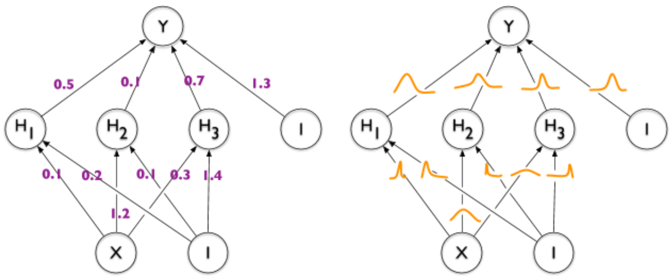
Initial clustering structure identification in HICs

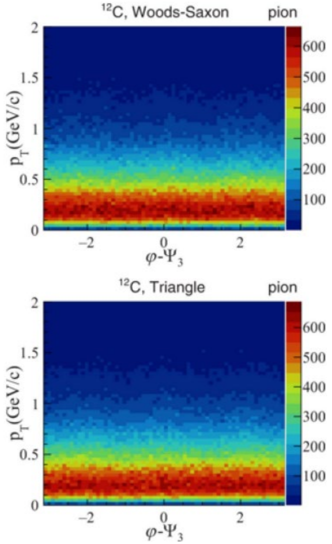
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

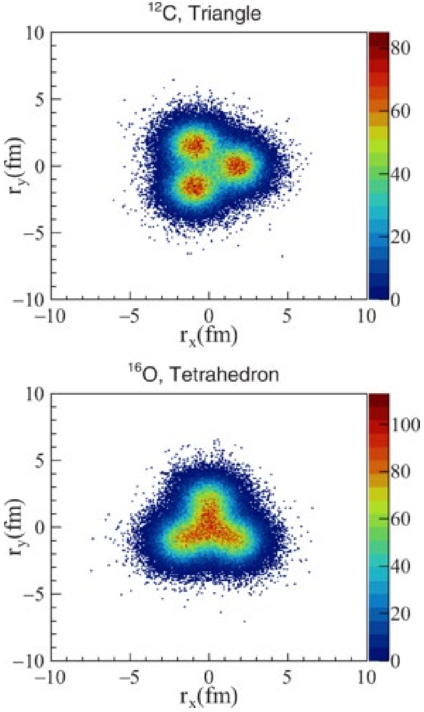
PHYSICAL REVIEW C 104, 044902 (2021)

Machine-learning-based identification for initial clustering structure in relativistic heavy-ion collisions

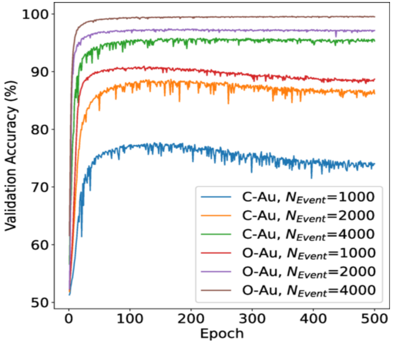
Junjie He (何俊杰),[1,2] Wan-Bing He (何万兵),[3,*] Yu-Gang Ma (马余刚),[3,†] and Song Zhang (张松)[3]







Bayesian CNN
on AMPT events (multiple-event basis)
Charged pions (phi, pT) from 12C/16O + 197Au collisions at 200 GeV



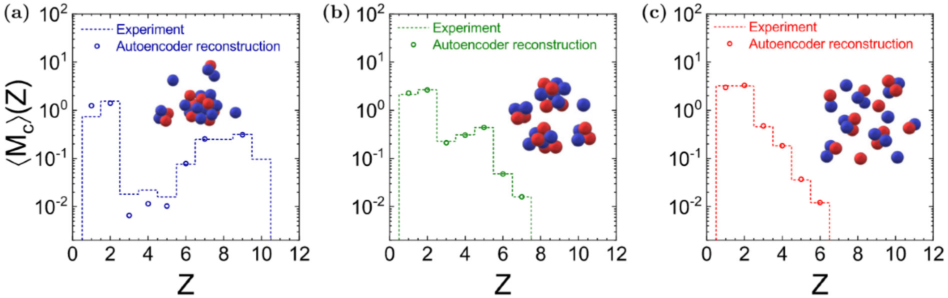
Multiple event basis

Nuclear liquid-gas phase transition recognition from autoencoder

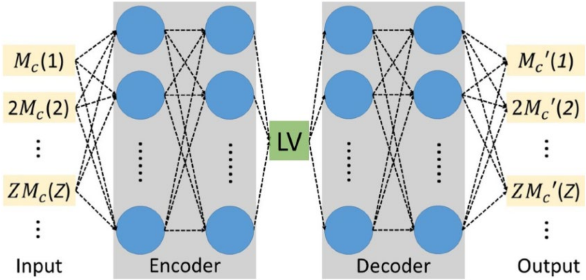
PHYSICAL REVIEW RESEARCH **2**, 043202 (2020)

Nuclear liquid-gas phase transition with machine learning

Rui Wang,[1,2,*] Yu-Gang Ma,[1,2,†] R. Wada,[3] Lie-Wen Chen,[4] Wan-Bing He,[1] Huan-Ling Liu,[2] and Kai-Jia Sun[3,5]



EbE charge-weighted charge multiplicity distribution of quasi-projectile as input →

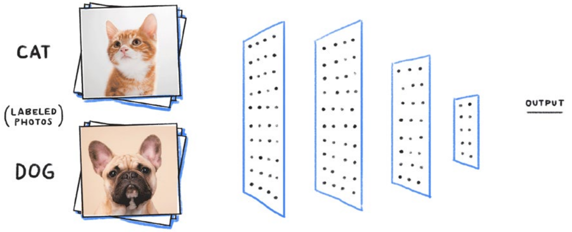


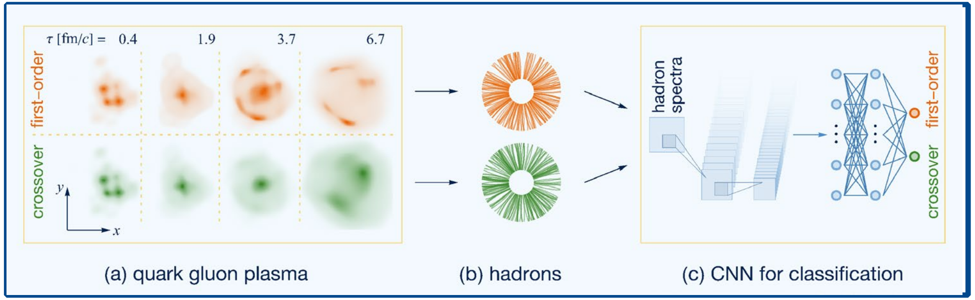
<u>**Autoencoder + confusion scheme**</u>
(on NIMROD experiment)

Direct inverse mapping with CNN for identifying QCD transition

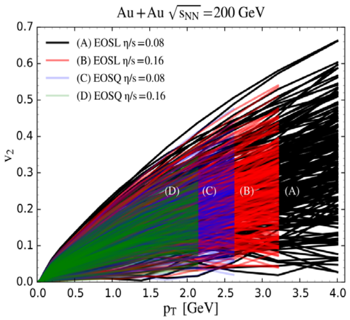
Data-driven
Inverse Mapping

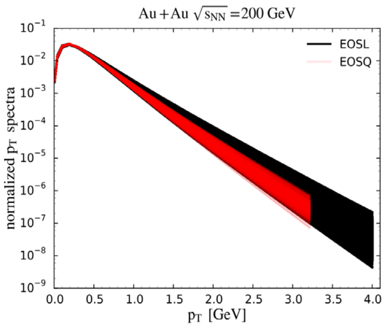
Physics Simulation
provide the Prior





(a) quark gluon plasma (b) hadrons (c) CNN for classification



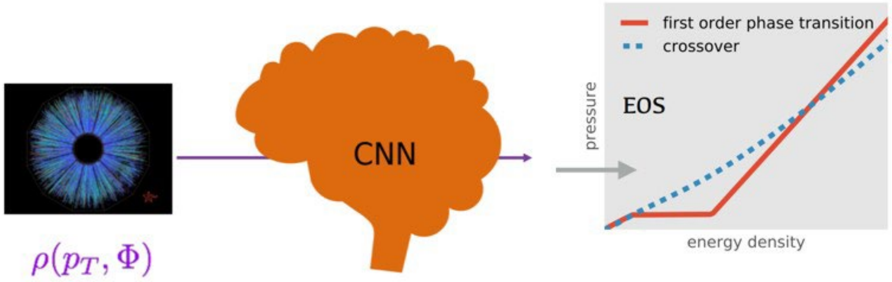


- Conventional obs. hard to distinguish
- Strongly influence from initial fluctuations and other uncertainties
- CNN : 95% event-by-event accuracy!
- Robust to initial conditions, eta/s

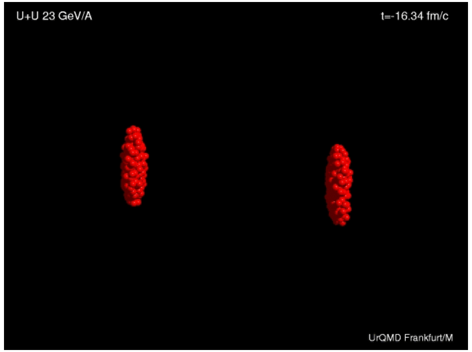
Conclusion : Information of early dynamics can **survive** to the end of hydrodynamics and encoded within the final state raw spectra, immune to evolution's uncertainties, **with deep CNN we can decode it back**.

L. Pang, K. Zhou, N. Su, H. Stoecker, H. Petersen, X. Wang, **Nature Commu.**9 (2018), no.1, 210

Into more realistic situations

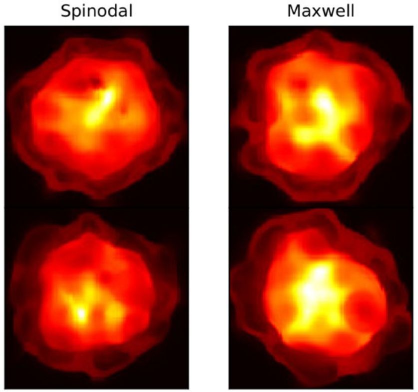


Hadronic cascade
(UrQMD considered)



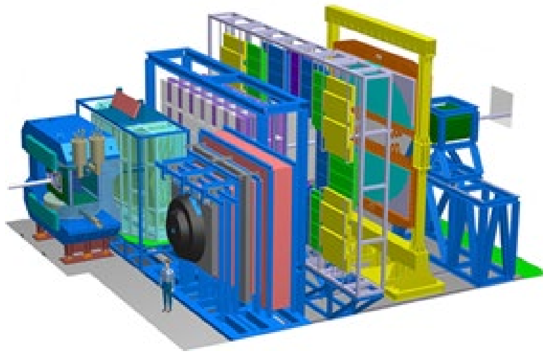
Eur. Phys. J. C 80 (2020) no.6,516

Non-equilibrium transition
(Baryon Clumping, spinodal)



JHEP 12,122(2019)
Phys. Rev. D 103,116023 (2021)

Detector effects
(Hits/Tracks, Point-Net)



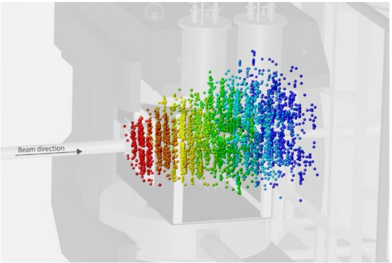
Phys. Lett. B 811, 135872
JHEP 21 (2021) 184

Point Cloud Network for Physics online analysis for HICs

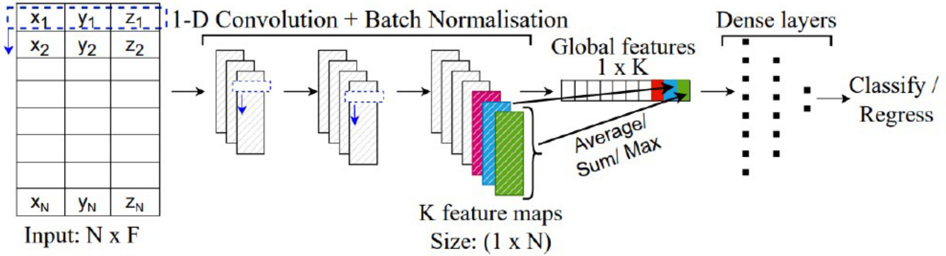
- Experimental data has inherent **point cloud structure**
  - **collection of particles** as *2D array* :

x1	y1	z1
x2	y2	z2
. . . .	. . . .	. . . .
xn	yn	zn

E	Px	Py	Pz	pid
6.84	1.07	4.5	6.83	211
40.4	0.06	0.54	40	321
…	…	…	…	…

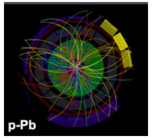


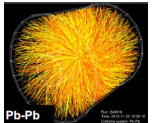
- PointNet based models learn directly from point clouds.
  - respects the **order invariance** of point clouds
  - direct processing of experimental data from detector ⇒ **ideal online analysis algorithm**
  - optimal for higher dimensional data

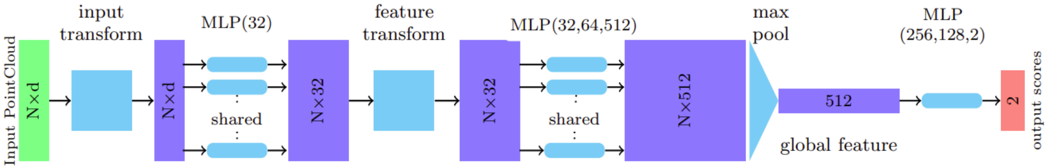


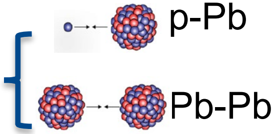
- **Collision Centrality Regression**
  M. OK, J. S, K. Zhou, H. S, **Phys.Lett.B** 811 (2020) 135872
- **EoS Classification**
  M. OK, K. Zhou, J. S, H. S, **JHEP** 10(2021) 184
- **Small/ Large-system Identification**

Manjunath O.K. and Kai Zhou, etc. *Phys.Lett.B* 811 (2020) 135872; JHEP10(2021)184.







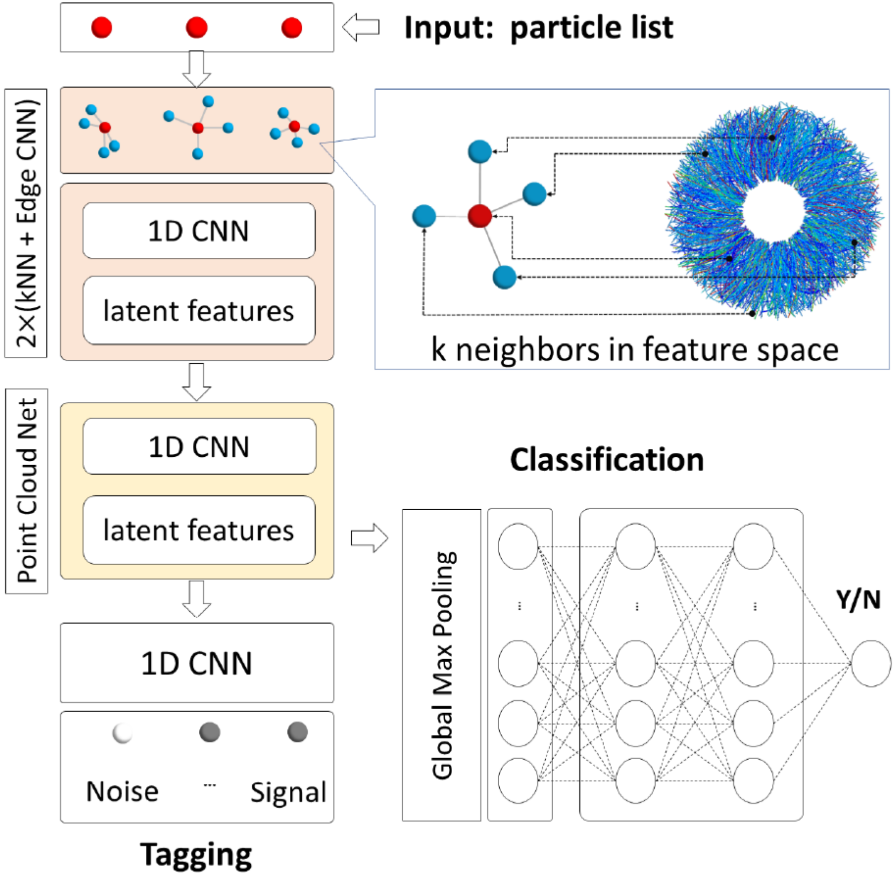


S.Guo, H. Wang, K. Zhou, G. Ma, **Phy.Rev.C** 2024

Modern dynamical edge CNN + PCN for self similarity searching

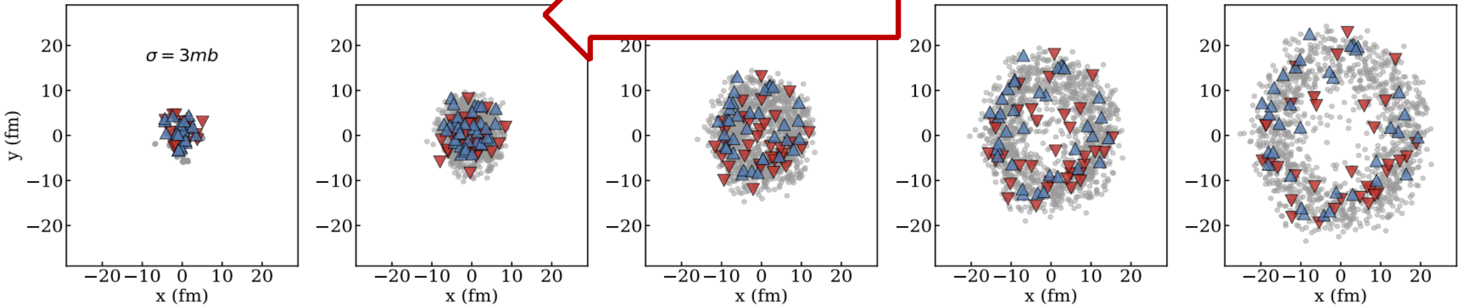
dynamical edge convolution network followed by a point cloud net is used to identify self-similarity and critical fluctuations in HIC

Repeating the **KNN** and **edge convolution** blocks twice helps to find long-range multi-particle correlations that are the key to searching for critical fluctuations.

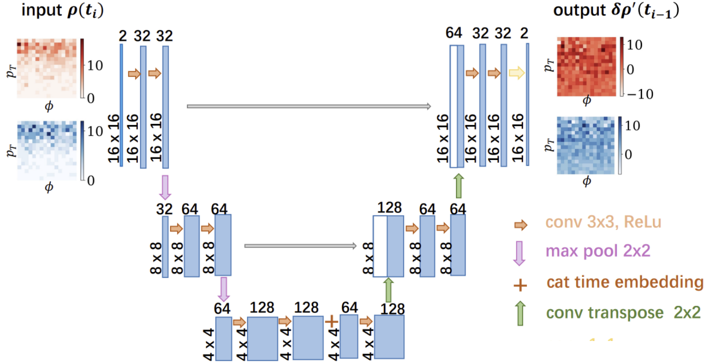


Y.-G. Huang, L.-G. Pang, X.F. Luo and X.-N. Wang, PLB 827(2022) 137001

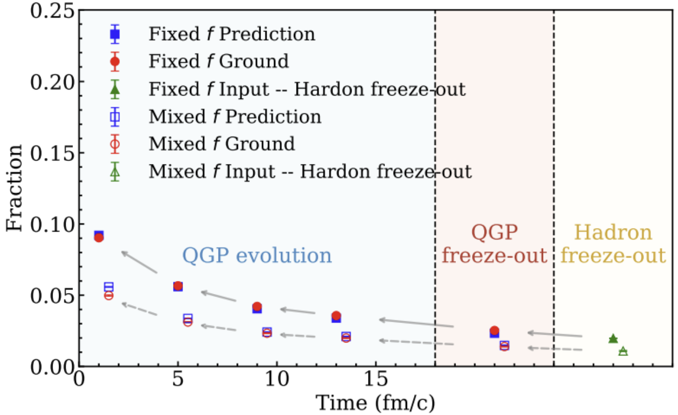
Unfold CME in HICs with time-embedded UNet



$$f = \frac{N^{\pm}_{\uparrow(\downarrow)} - N^{\pm}_{\downarrow(\uparrow)}}{N^{\pm}_{\uparrow(\downarrow)} + N^{\pm}_{\downarrow(\uparrow)}}$$



$$\delta\rho(t_{i-1}) = \rho(t_{i-1}) - \rho(t_i)$$
$$\rho'(t_{i-1}) = \delta\rho'(t_{i-1}) + \rho(t_i)$$



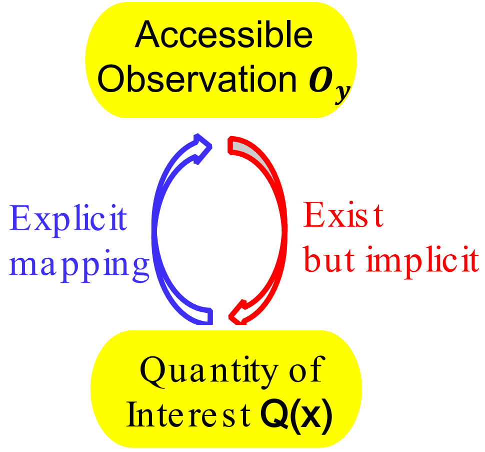
See talk of Shuang Guo this afternoon, 14:50pm

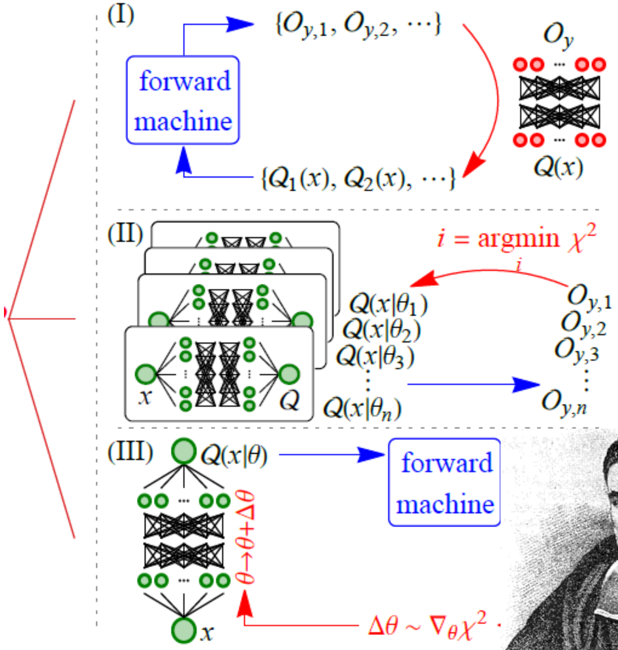
S.Guo, L. Wang, K. Zhou, G. Ma, **Chin. Phys. Lett. 42, 110101** (2025)

Lessons

- Low-level features (raw data) might work better in DL era !
- Point Cloud representation for nuclear collision analysis ! Theory-Data
- Physics prior info is important !

- Need to have well-defined problem : classification, regression
- Benchmark datasets release from our community?
- Interpretable ? How to connect to fundamental physics?

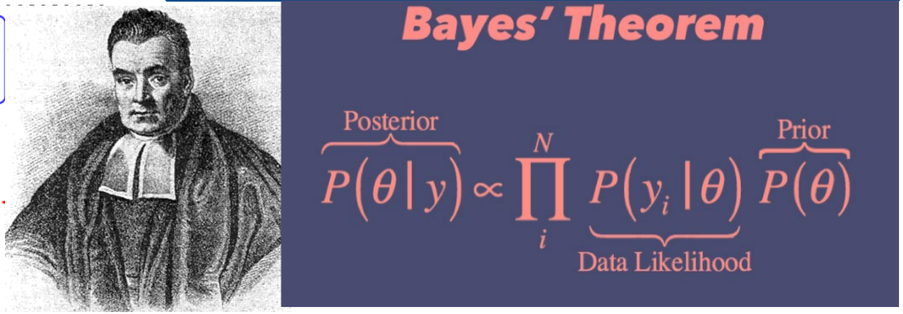




- **Direct inverse mapping capturing :** with Supervised Learning

- **Statistical approach to χ^2 fitting :** Bayesian Reconstruction for posterior or Heuristic (Generic) Algorithm to min.

$$\chi^2 = \sum_y \left(\frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\theta)] - \mathcal{O}_y}{\Delta \mathcal{O}_y} \right)^2$$



Bayesian (Statistical) Inference of hot matter EoS from HIC data

PRL 114, 202301 (2015) PHYSICAL REVIEW LETTERS week ending 22 MAY 2015

Constraining the Equation of State of Superhadronic Matter from Heavy-Ion Collisions

Scott Pratt,[1] Evan Sangaline,[1] Paul Sorensen,[2] and Hui Wang[2]
[1]*Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory Michigan State University, East Lansing, Michigan 48824, USA*
[2]*Brookhaven National Laboratory, Upton, New York 11973, USA*
(Received 19 January 2015; published 19 May 2015)

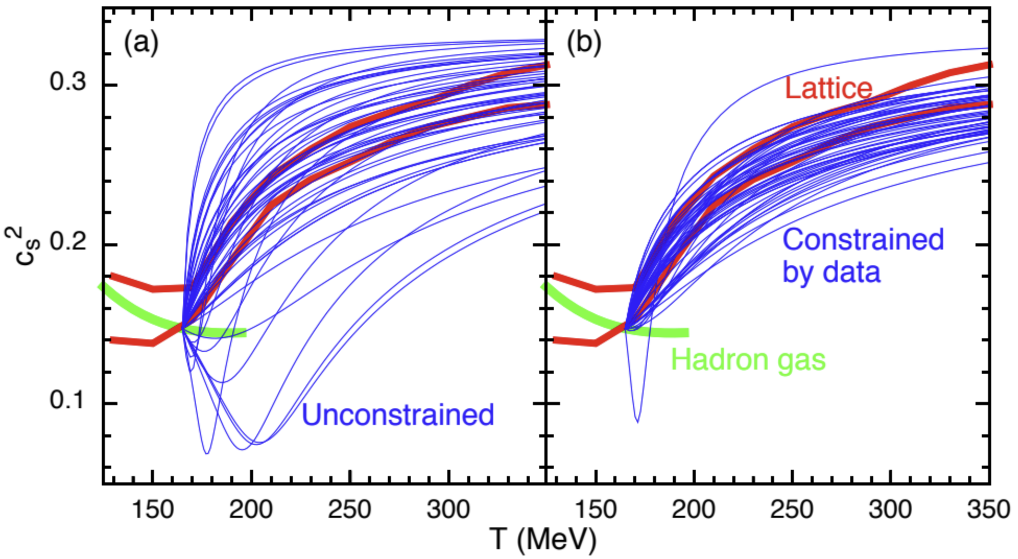
$$c_s^2(\epsilon) = c_s^2(\epsilon_h) + \left(\frac{1}{3} - c_s^2(\epsilon_h)\right)\frac{X_0 x + x^2}{X_0 x + x^2 + X'^2},$$

$$X_0 = X' R c_s(\epsilon)\sqrt{12}, \quad x \equiv \ln \epsilon/\epsilon_h,$$

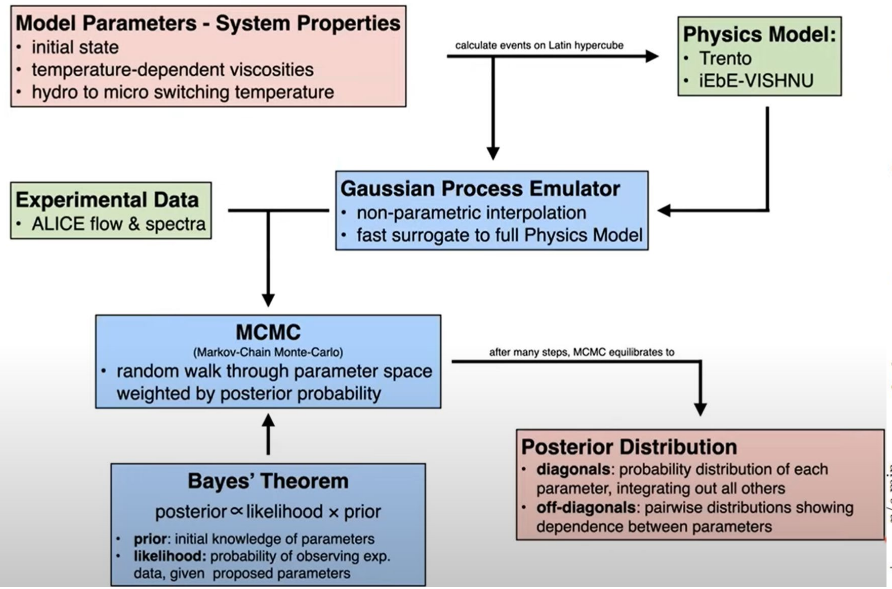
$$P(D|\theta) = \Pi_i \exp(-(z_i(\theta) - z_{i,\exp})^2/2),$$

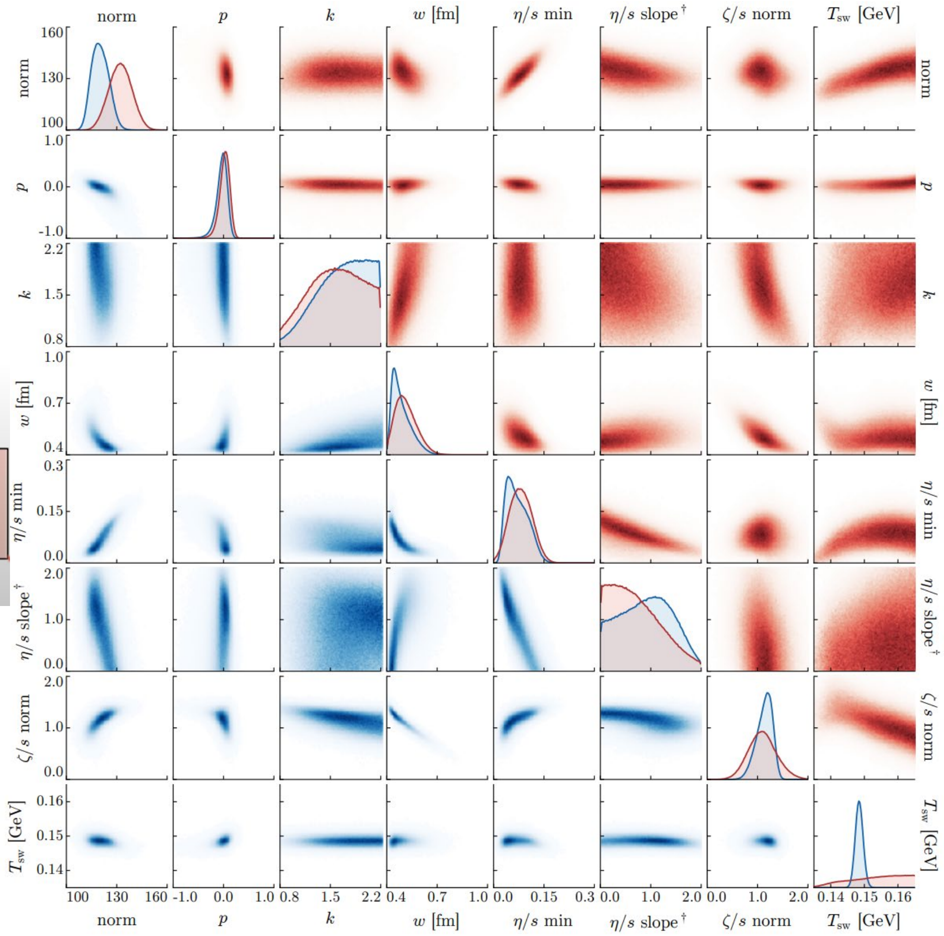
14 model parameters

speed of sound squared slightly
softer than lattice EoS
But significantly overlap



Bayesian (Statistical) global fit on HICs





Trento + IEBE-VishNew + UrQMD

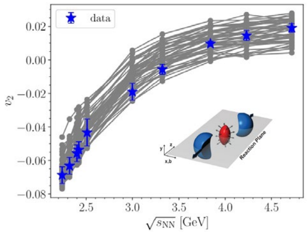
J. Bernhard, J. Morel, S. Bass, **Nat. Phy.** 15, 1113 (2019)

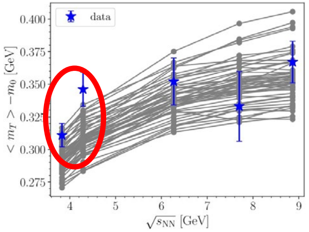
G. Nijs, W. Schee, U. Guersoy, R. Snellings, **PRC**103,054909;
JETSCAPE, **PRL**126,242301; U. Heinz+, 2302.14184 (VAH)
M. R. Heffernan, C. Gale, S. Jeon, J. Pauet, **PRC**109,065207;
… … … …

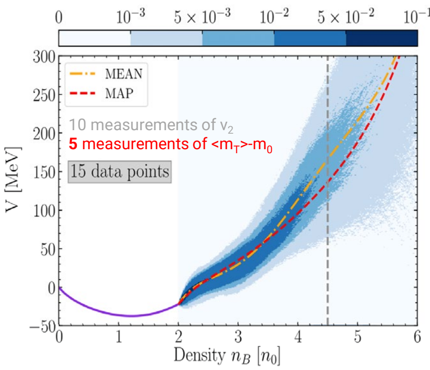
Jet quenching/diffusion:
Y. He, L. Pang, X. Wang, **PRL** 122 (25) 252302
M. Xie, W. Ke, H. Zhang, X. Wang, **PRC108** (2023) L011901; … …

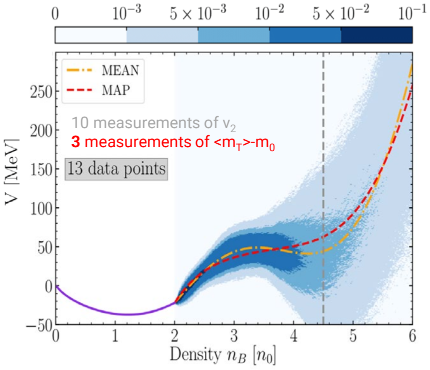
Bayesian Inference Dense Matter EoS from HIC and Holography

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen









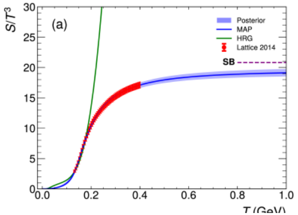
- Comprehensive Bayesian inference necessary for unambiguous solution
- Tension between data-data or model (**UrQMD**)-data
- Next-gen experiments will provide immense amount of high precision data

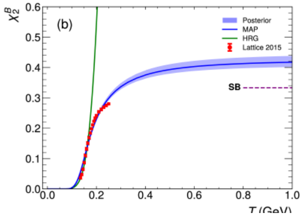
M.OK, J. Steinheimer, K. Zhou,
H. Stoecker, **PRL131,202303(2023)**

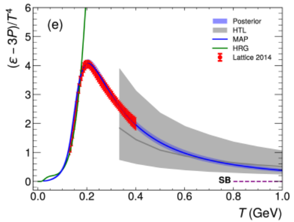
$$S_E = \frac{1}{16\pi G_5}\int d^5x\sqrt{-g}\left[R - \frac{f(\phi)}{4}F^2 - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right]$$

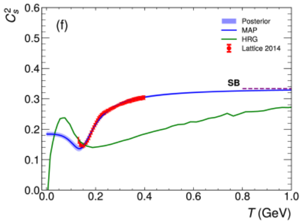
$$A(z) = d\ln(az^2+1) + d\ln(bz^4+1),\ f(z) = e^{cz^2 - A(z) + k}$$

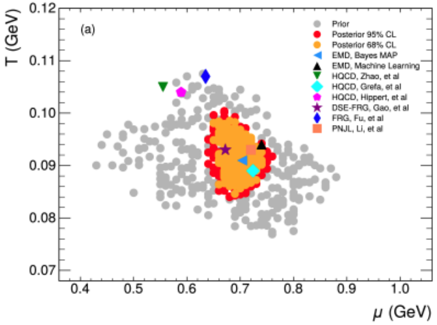
See talk of Li-qiang Zhu tomorrow 16:40pm

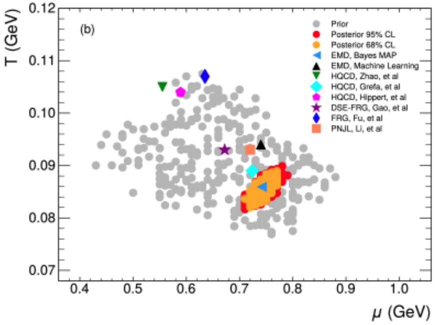












- Critical endpoint from **holography (EMD)** via Bayesian Inference

L. Zhu, X. Chen, K. Zhou,
H. Zhang, M, Huang,
Phys. Rev. D112(2025) 026019

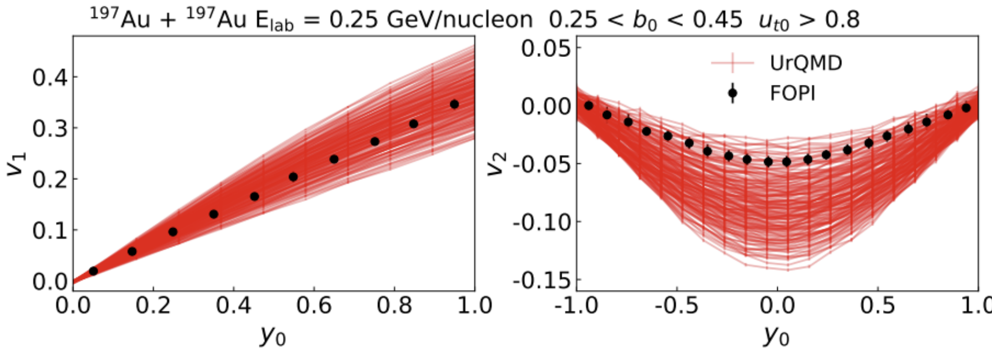
Incompressibility K from Bayesian Inference with low-energy HIC

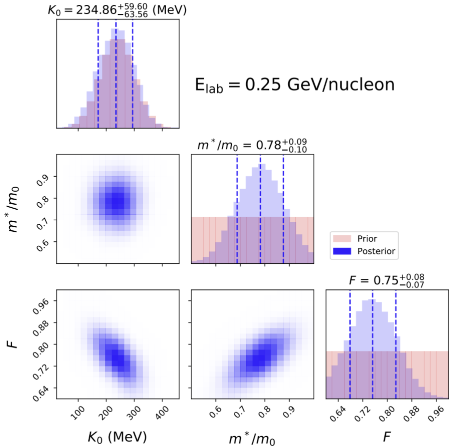
Bayesian analysis of properties of nuclear matter with the FOPI experimental data

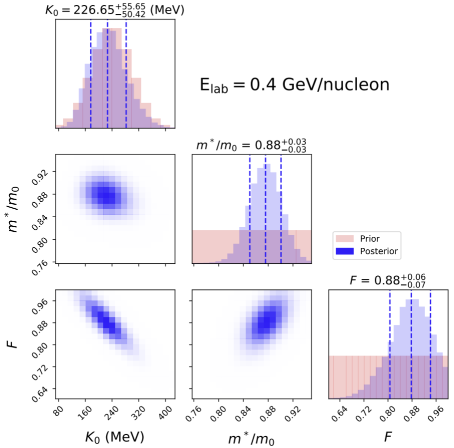
Guojun Wei,[1,2] Manzi Nan,[1,3] Pengcheng Li,[1] Yongjia Wang,[1,4,*] Qingfeng Li,[1,†] Gaochan Yong,[3] and Fuhu Liu[2]

arXiv:2509.03406

UrQMD simulation of proton v1 and v2
@Au+Au at E=0.25, 0.4 GeV/Nucleon







See talk of Qing-feng Li
this morning 11:00am

Incompressibility K from Bayesian Inference with low-energy HIC

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

Bayesian inference of nuclear incompressibility from proton elliptic flow in central Au+Au collisions at 400 MeV/nucleon

J. M. Wang (汪金梅),[1,2] X. G. Deng (邓先概),[1,2,*] W. J. Xie (谢文杰),[3] B. A. Li (李宝安),[4,†] and Y. G. Ma (马余刚)[1,2,‡]

Chinese Physics C 49, 124105 (2025)

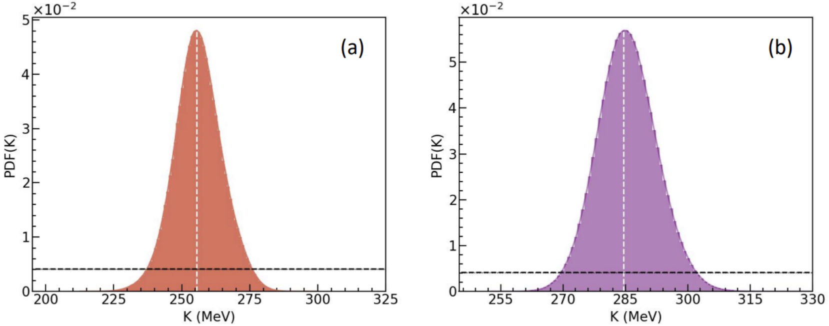


FIG. 4. Without considering the MDI: the posterior PDFs of K. Left: using the observables $-v_2(y_0)$ and $-v_2(u_{t0})$, right: observables $-v_2(y_0)$, $-v_2(u_{t0})$ and $v_1\left(p_t^{(0)}\right)$.

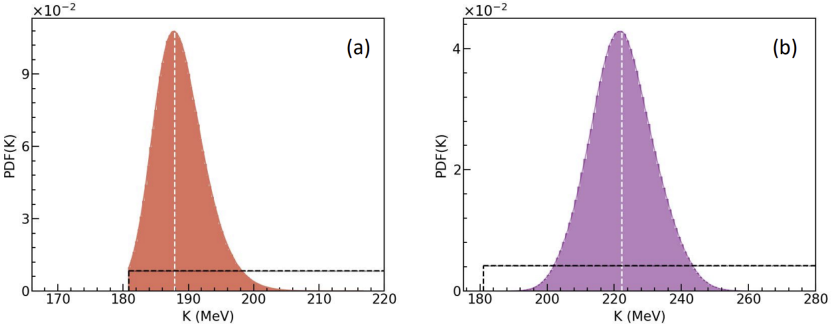


FIG. 6. Considering the MDI: the posterior PDFs of K. Left: observables $-v_2(y_0)$ and $-v_2(u_{t0})$, right: observables $-v_2(y_0)$, $-v_2(u_{t0})$ and $v_1\left(p_t^{(0)}\right)$.

IQMD simulation of
proton v2 Au+Au at E=400 MeV/Nucleon

MDI: momentum dependent Interaction

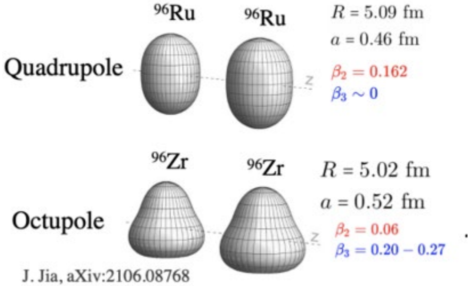
$$E/A = \frac{\alpha}{2}\frac{\rho}{\rho_0} + \frac{\beta}{\gamma+1}\left(\frac{\rho}{\rho_0}\right)^{\gamma} + \frac{3}{10m}\left(\frac{3\pi^2\hbar^3\rho}{2}\right)^{2/3} + \frac{1}{2}t_4\frac{\rho}{\rho_0}\int f(\vec{p})\ln^2\left[1+t_5\left(\vec{p}-\langle\vec{p}'\rangle\right)^2\right]d^3p, \quad (1)$$

$$P = \rho^2\frac{\partial E/A}{\partial\rho} = \frac{\alpha}{2}\frac{\rho^2}{\rho_0} + \frac{\beta\gamma\rho}{\gamma+1}\left(\frac{\rho}{\rho_0}\right)^{\gamma} + \frac{1}{5m}\left(\frac{3}{2}\pi^2\hbar^3\right)^{\frac{2}{3}}\rho^{\frac{5}{3}} + \frac{t_4}{2}\frac{\rho^2}{\rho_0}\ln^2\left(1+t_5P_F^2\right), \quad (2)$$

$$K = 9\rho^2\frac{\partial^2 E/A}{\partial\rho^2}\Big|_{\rho_0} = -\frac{3}{5m}\left(\frac{3\pi^2\hbar^3\rho_0}{2}\right)^{2/3} + \frac{9\beta\gamma(\gamma-1)}{\gamma+1} + \ln\left(1+t_5P_F^2\right)\frac{6t_4t_5P_F^2}{1+t_5P_F^2}, \quad (3)$$

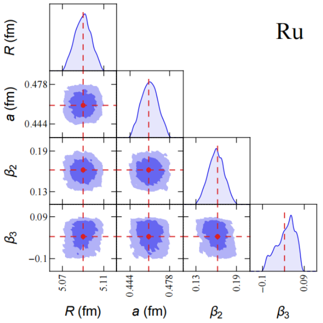
Bayesian Imaging for Nuclear Structure in Isobar Collisions

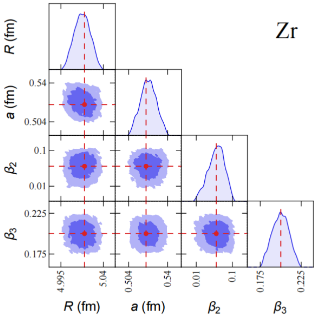
- Nuclear Structure imaging for single system ? (caveat: model dependent)
- Simultaneous inference for isobar systems with ratio?
- **Bayesian Inference:** Gaussian Process emulator + PCA dim reduction + MCMC
  Data: MC-Glauber + Matching (linear response approximation)



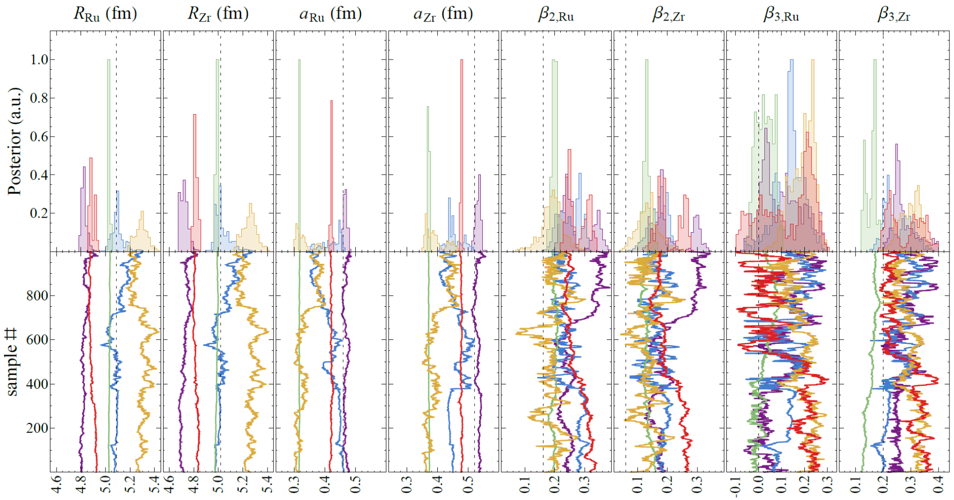
J. Jia, aXiv:2106.08768

$$\boldsymbol{y}_{\rm Ru} \equiv \{P_a^{\rm Ru}, \varepsilon_{2,a}^{\rm Ru}, \varepsilon_{3,a}^{\rm Ru}, d_{\perp,a}^{\rm Ru}\}_{a=1,\cdots,40}$$





Single system works good



With purely the Isobar-Ratios:

MCMC can not converge to a stationary inference ofthe nuclear structure

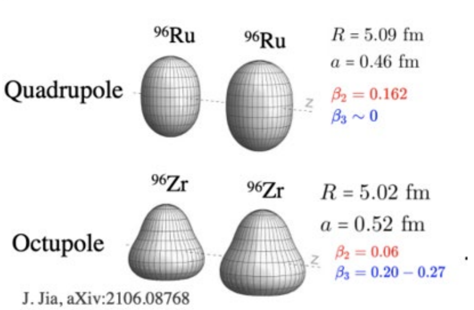
$$\boldsymbol{y}_{\rm r,1} \equiv \{R_{P,a}, R_{\varepsilon_2,a}, R_{\varepsilon_3,a}, R_{d_\perp,a}\}_{a=1,\cdots,40}$$

- No unique solution : only ratios $\not\Rightarrow$ *nuclear structure !*

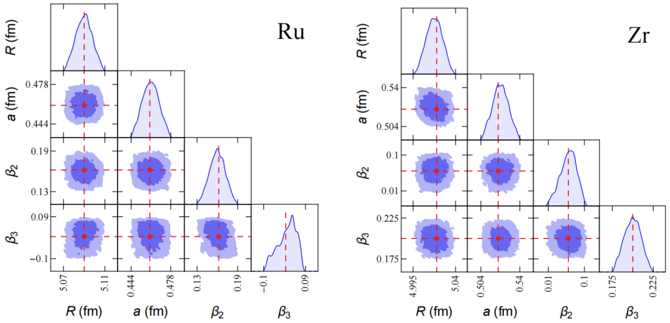
Y.Cheng, S.Shi, Y. Ma, H. S., K. Zhou,
Phys. Rev. C 107 (2023) 064909

Bayesian Imaging for Nuclear Structure in Isobar Collisions

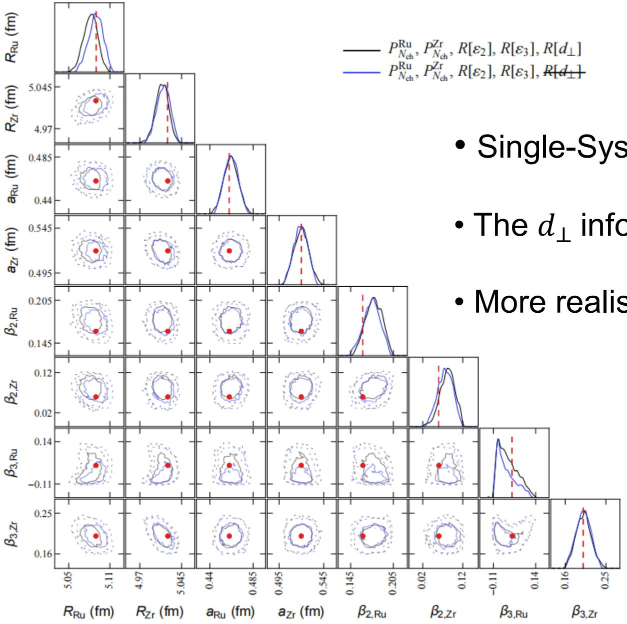
- Nuclear Structure imaging for single system ? (caveat: model dependent)
- Simultaneous inference for isobar systems with ratio?
- **Bayesian Inference:** Gaussian Process emulator + PCA dim reduction + MCMC
  Data: MC-Glauber + Matching (linear response approximation)



$$\boldsymbol{y}_{\mathrm{Ru}} \equiv \{P_a^{\mathrm{Ru}}, \varepsilon_{2,a}^{\mathrm{Ru}}, \varepsilon_{3,a}^{\mathrm{Ru}}, d_{\perp,a}^{\mathrm{Ru}}\}_{a=1,\cdots,40}$$



Single system works good



$P_{N_{ch}}^{\mathrm{Ru}}, P_{N_{ch}}^{\mathrm{Zr}}, R[\varepsilon_2], R[\varepsilon_3], R[d_\perp]$

$P_{N_{ch}}^{\mathrm{Ru}}, P_{N_{ch}}^{\mathrm{Zr}}, R[\varepsilon_2], R[\varepsilon_3]$, ~~$R[d_\perp]$~~

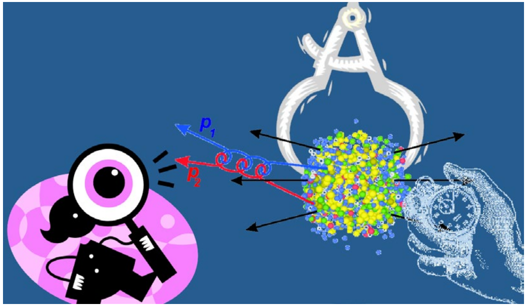
- Single-System Multiplicity makes it possible
- The $d_\perp$ information is redundant
- More realistic analysis with AMPT in progress

$$\boldsymbol{y}_{\mathrm{r},2} \equiv \{P_a^{\mathrm{Ru}}, P_a^{\mathrm{Zr}}, R_{\varepsilon_2,a}, R_{\varepsilon_3,a}, R_{d_\perp,a}\}_{a=1,\cdots,40}$$

Y.Cheng, S.Shi, Y. Ma, H. S., K. Zhou,
Phys. Rev. C 107 (2023) 064909

Bayesian reconstruction for h-h interaction from femtoscopy – mock test

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen



$$C(k) = \int S(\mathbf{r}) \, |\, \psi_k(\mathbf{r}) \,|^2 \, \mathrm{d}^3 r,$$

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi + V\psi = E\psi,$$

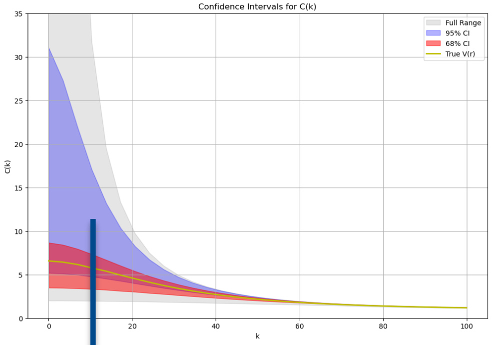
$$S(r) = \frac{1}{(4\pi r_0^2)^{3/2}} \exp\left(-\frac{r^2}{4r_0^2}\right),$$

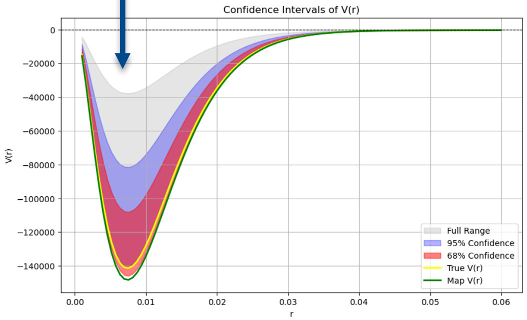
$$r_0 = 1.0 \text{ fm}$$

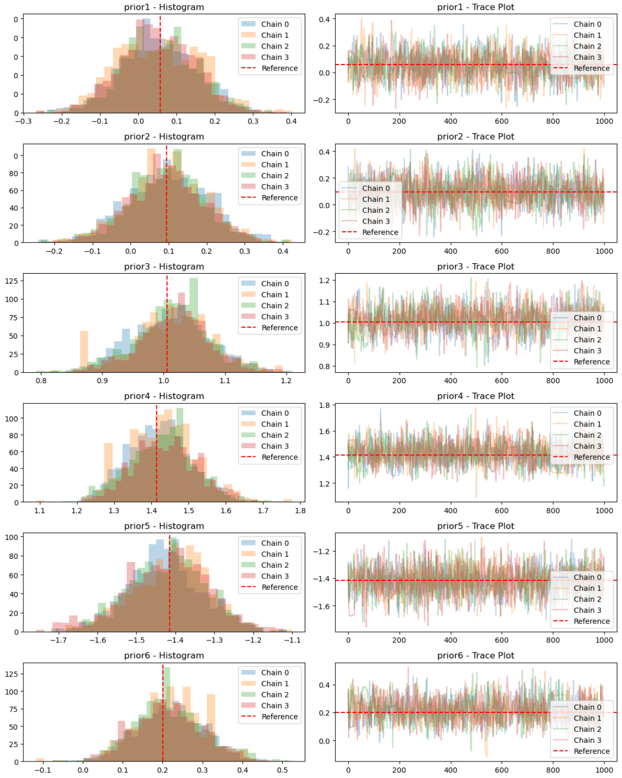
$$V(r) = \sum_{i=1,2} a_i e^{-(r/b_i)^2} + a_3 m_\pi^4 f(r, b_3) \frac{e^{-2m_\pi r}}{r^2}$$

$$f(r, b_3) = \left(1 - e^{-(r/b_3)^2}\right)^2$$

There are six parameters are a_i, b_i $(i = 1,2,3)$ for this potential !







DNN emulator + PCA for correlation + PyMC
With O. L, J. Z, X. C, etc., in preparation

OPEN ACCESS
IOP Publishing
Journal of Physics G: Nuclear and Particle Physics
J. Phys. G: Nucl. Part. Phys. **51** (2024) 103001 (43pp)
https://doi.org/10.1088/1361-6471/ad6a2b

Topical Review

Applications of emulation and Bayesian methods in heavy-ion physics

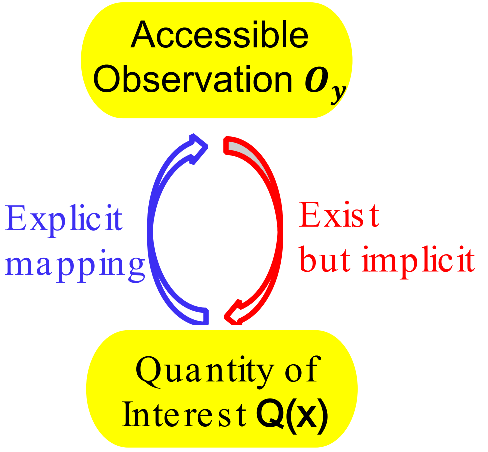
Jean-François Paquet

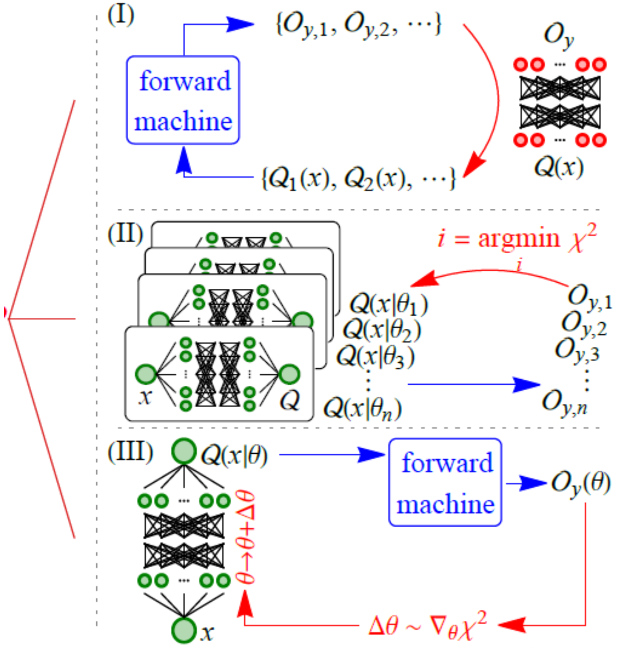
Different analyses = different constraints

- Use different data sets
- Different modelling assumptions:
  - Initial conditions
  - Cooper-Frye
  - Allowed parametrization of transport coefficient
- Treatment of correlations in experimental uncertainties

References	Pre-hydro	Hydro	Cooper-Frye	Data	Covariance
Bernhard *et al* [28]	Trento+f.s.	DNMR	P.T.B.; σ meson production	Pb–Pb @ 2.76 TeV and 5.02 TeV	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +non-diag. $\Sigma^{\rm syst}_{\rm expt}$ + $\Sigma_{\rm extra}$
Moreland *et al* [89]	Trento w/ subnucleonic d.o. f.+f.s.	DNMR	P.T.B.; σ meson production	p-Pb & Pb–Pb @ 5.02 TeV	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +non-diag. $\Sigma^{\rm syst}_{\rm expt}$
JETSCAPE [15, 53]	Trento+f.s.	DNMR	Grad, Chapman–Enskog, P.T.B.	Au–Au @ 0.2 TeV and Pb–Pb @ 2.76 TeV	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +diag. $\Sigma^{\rm syst}_{\rm expt}$
Nijs *et al* [111, 112]	Trento w/ subnucleonic d.o. f. + modified streaming	DNMR	P.T.B.; σ meson production	Pb–Pb @ 2.76 TeV; p-Pb and Pb–Pb @ 5.02 TeV; added differential observables	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +non-diag. $\Sigma^{\rm syst}_{\rm expt}$
Parkkila *et al* [52, 113]	Trento+f.s.	DNMR	P.T.B. σ meson production	Pb–Pb @ 2.76 TeV and 5.02 TeV; added event-plane correlations	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +non-diag. $\Sigma^{\rm syst}_{\rm expt}$ + $\Sigma_{\rm extra}$
Liyanage *et al* [35]	Trento + anisotropic hydro parameters	Viscous anisotropic hydro	P.T.M.A.	Pb–Pb @ 2.76 TeV	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +diag. $\Sigma^{\rm syst}_{\rm expt}$
Heffernan *et al* [24, 26]	IP-Glasma	DNMR	Grad, Chapman–Enskog	Pb–Pb @ 2.76 TeV; added event-plane correlations	$\Sigma_{\rm emul}$ +diag. $\Sigma^{\rm stat}_{\rm expt}$ +diag. $\Sigma^{\rm syst}_{\rm expt}$

Inverse Problems Solving with ML





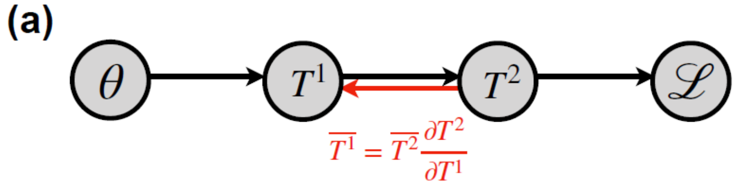
- **Direct inverse mapping capturing :** with Supervised Learning
- **Statistical approach to χ^2 fitting :** Bayesian Reconstruction for posterior or Heuristic (Generic) Algorithm to min.

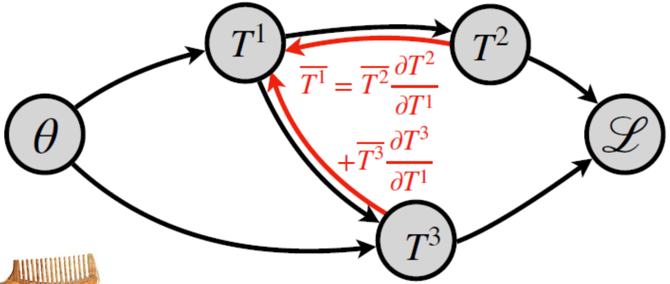
$$\chi^2 = \sum_y \left(\frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{\Delta \mathcal{O}_y} \right)^2$$

- **Automatic Differentiation :** fuse physical prior into reconstruction via differentiable programming strategy

$$\frac{1}{2}\nabla_{\boldsymbol{\theta}}\chi^2 = \sum_y \frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{(\Delta\mathcal{O}_y)^2} \int \mathrm{d}x \frac{\delta \mathcal{F}_y[\mathcal{Q}(x)]}{\delta \mathcal{Q}(x)}\bigg|_{\mathcal{Q}(x)=\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}}\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})$$

Deep Learning composes differentiable components to a program, e.g. DNN, then optimizes it with gradients






"comb graph"

Chain rule for gradients : $\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial T^n}\frac{\partial T^n}{\partial T^{n-1}}\cdots\frac{\partial T^2}{\partial T^1}\frac{\partial T^1}{\partial \theta}$

Defining adjoint variables : $\overline{T} = \partial \mathcal{L}/\partial T$

$$\overline{T^i} = \overline{T^{i+1}}\frac{\partial T^{i+1}}{\partial T^i}$$

$$\overline{\theta} = \overline{T^1}\frac{\partial T^1}{\partial \theta}$$

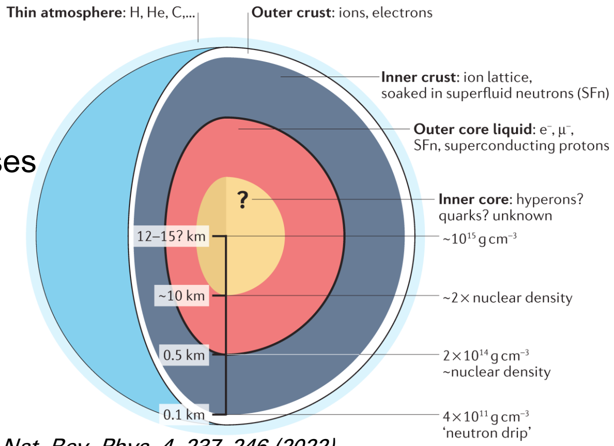
$$\overline{T^i} = \sum_{j:\,\text{child of}\,i} \overline{T^j}\frac{\partial T^j}{\partial T^i}$$

Differentiable programming tools

HIPS/autograd, TensorFlow, flux, PyTorch, MindSpore, SciML, NiLang

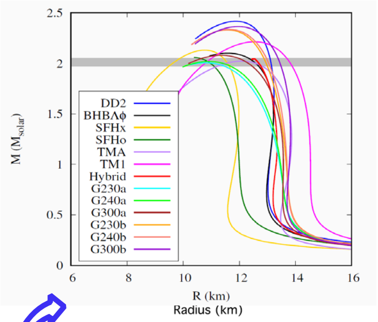
From EoS to NS Stellar Structure (MR)

- Mass ~ 2 solar masses
- Radii ~ 10 km
- Densities 5-8 ρ_0



Nat. Rev. Phys. 4, 237–246 (2022)

- Noisy/Limited NS Observables to EoS ?



TOV equations

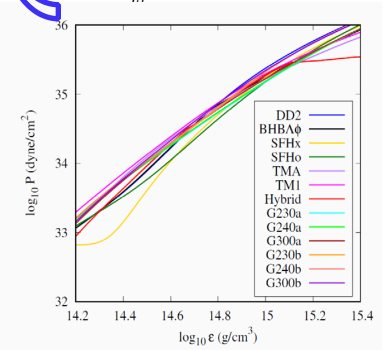
$$-\frac{dP}{dr} = \frac{[\epsilon(r)+P(r)]\,[M(r)+4\pi r^3 P(r)]}{r[r-2M(r)]}$$

$$\frac{dM(r)}{dr} = 4\pi r^2 \epsilon(r),$$

- Gravity ← → Pressure

$$\frac{dP}{dr} = -\frac{G}{r^2}\left(\rho+\frac{P}{c^2}\right)\left(m+4\pi r^3\frac{P}{c^2}\right)\left(1-\frac{2Gm}{c^2 r}\right)^{-1}$$

$$M = m(R) = \int_0^R 4\pi r^2 \rho\, dr$$

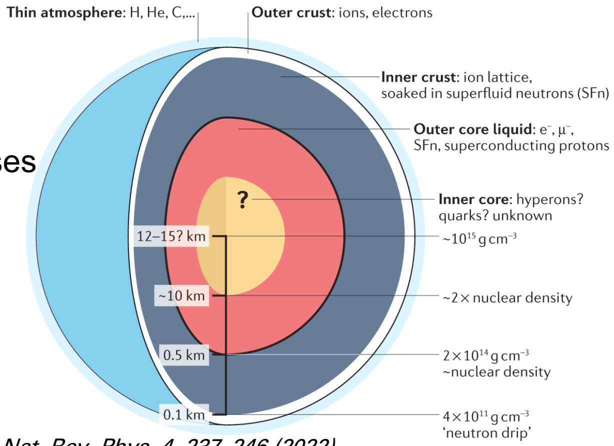


- Dense matter Equation of State $P(\rho)$

From EoS to NS Stellar Structure (MR) -- Inverse ?

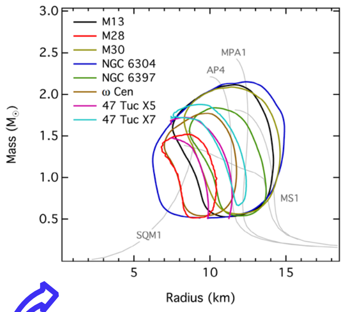
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

- Mass ~ 2 solar masses
- Radii ~ 10 km
- Densities 5-8 ρ_0



Nat. Rev. Phys. 4, 237–246 (2022)

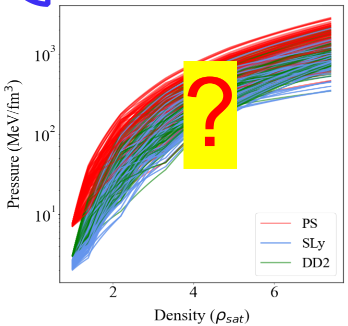
- Noisy/Limited NS Observables to EoS ?



TOV equations

$$-\frac{dP}{dr} = \frac{[\epsilon(r)+P(r)]\,[M(r)+4\pi r^3 P(r)]}{r[r-2M(r)]}$$

$$\frac{dM(r)}{dr} = 4\pi r^2 \epsilon(r),$$



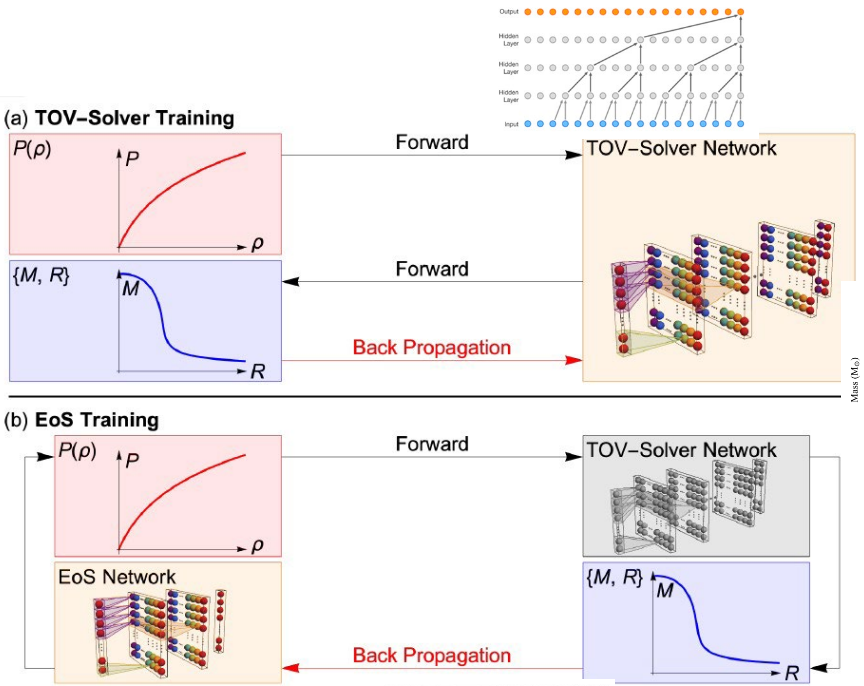
- Gravity ← → Pressure

$$\frac{dP}{dr} = -\frac{G}{r^2}\left(\rho + \frac{P}{c^2}\right)\left(m + 4\pi r^3 \frac{P}{c^2}\right)\left(1 - \frac{2Gm}{c^2 r}\right)^{-1}$$

$$M = m(R) = \int_0^R 4\pi r^2 \rho \, dr$$

- Dense matter Equation of State $P(\rho)$

Auto-diff framework and Results



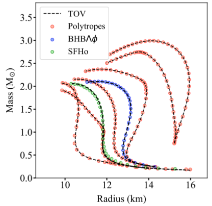
$$\mathscr{L} = \chi^2 = \sum_{i=1}^{N_{obs}} \frac{(M_i - M_{obs,i})^2}{\Delta M_i^2} + \frac{(R_i - R_{obs,i})^2}{\Delta R_i^2}$$

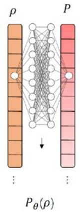
Loss function

$$g_t = \frac{\delta \chi^2}{\delta \theta_t} = \frac{\delta \chi^2}{\delta \mathbf{z}} \frac{\delta \mathbf{z}}{\delta x} \frac{\delta x}{\delta \theta_t}$$

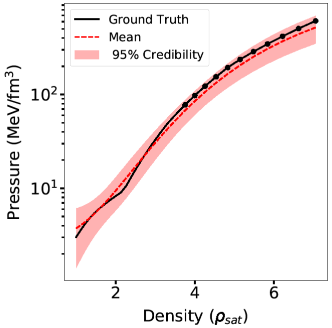
$\mathbf{z} = (M_i, R_i), x = P_i(\rho_i)$

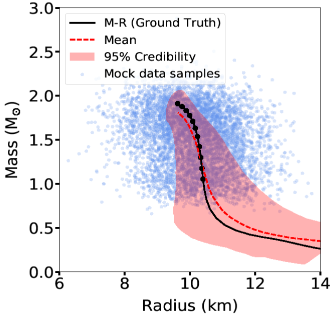
Gradients



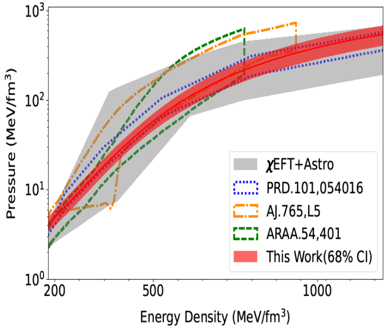


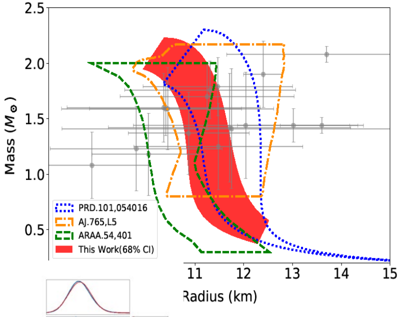
- Well validated through **Mock Tests**

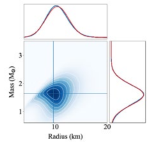




- With **real observable** we reconstruct the NS EoS also



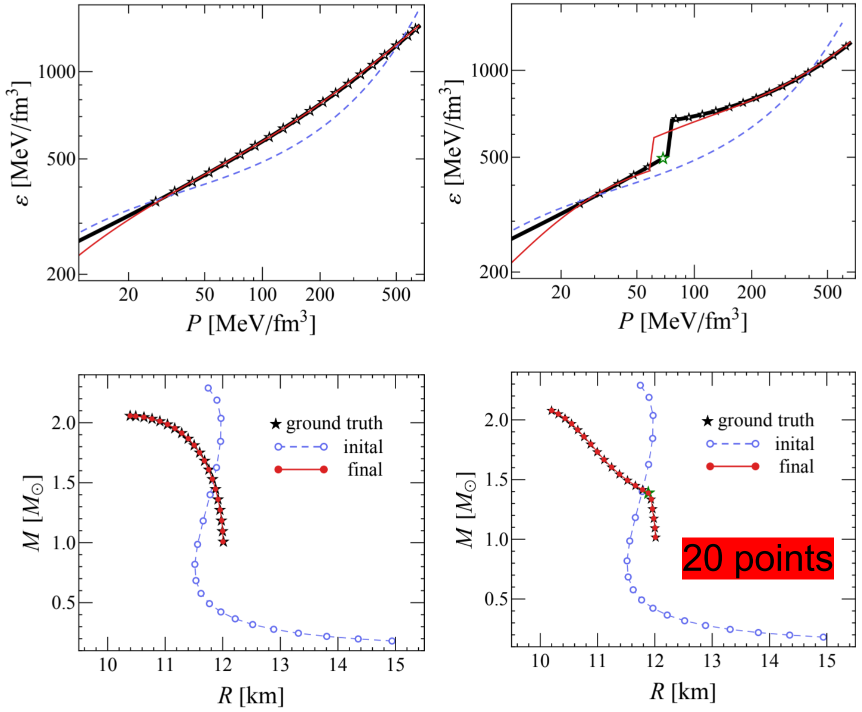




S. Soma, L. Wang, S. Shi, H. Stoecker, **K. Zhou**, **JCAP** 98 (2022) 071
S. Soma, L. Wang, S. Shi, H. Stoecker, **K. Zhou**, **Phys. Rev. D** 107 (2023)083028

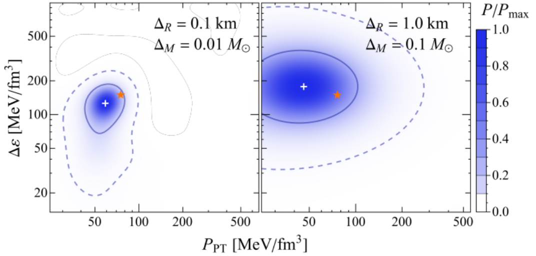
Extend to First-order phase transition reconstruction with AutoDiff

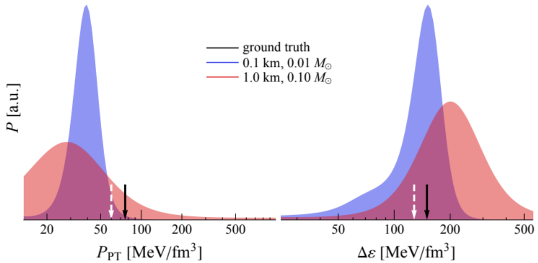
Linear response analysis get the gradients! Then use DNN :



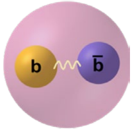
We parameterize the inverse speed of sound squared containing both regular parts and Dirac-δ functions corresponding to possible first-order phase transitions,

We adopt SFHo as the baseline EoS and introduce a PT with latent heat $\Delta\varepsilon = 150\ \mathrm{MeV/fm^3}$ at pressure $P_{\rm PT} = 76\ \mathrm{MeV/fm^3}$. Above the PT point, we take the stiffest (causal) limit that $c_s = 1$. We employ twenty



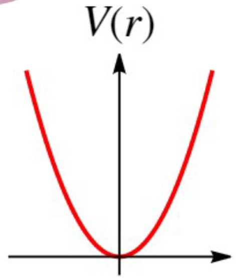


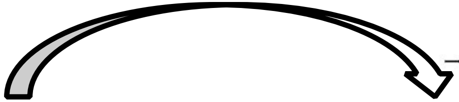
R. Li, S. Han, Z. Lin, L. Wang, K. Zhou,
S. Shi, **Phys. Rev. D 111(2025)074026**

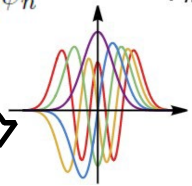


$$\hat{H}\psi_n = -\frac{\nabla^2}{2m_\mu}\psi_n + V(r)\psi_n = E_n\psi_n$$

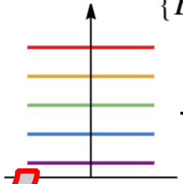
$\{\psi_n(r)\}$







$\{E_n\}$

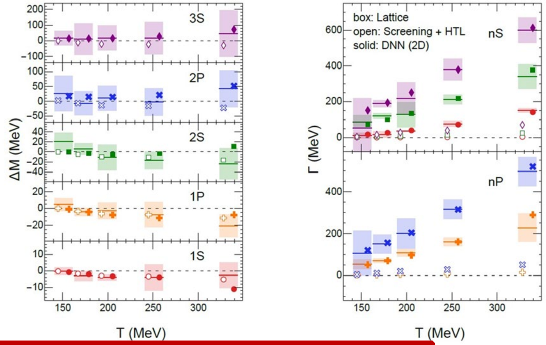


$$\begin{cases} \mathrm{Re}[E_n] = m - 2m_b \\ \mathrm{Im}[E_n] = -\Gamma \end{cases}$$

$$V(T,r) = V_R(T,r) + i \cdot V_I(T,r)$$

New lQCD results cannot be explained by Perturbative HTL-inspired potentials !

R. Larsen, et.al, PRD(2019), PLB(2020), PRD(2020)

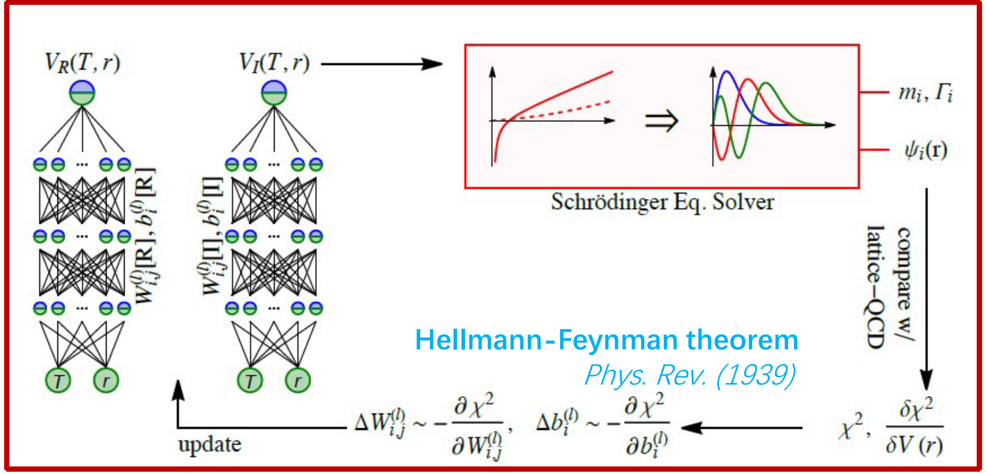


How to extract effective potential given limited spectroscopy ? →

$$\chi^2 = \sum_{T,i} \frac{(m_{T,i} - m^{\text{lattice}}_{T,i})^2}{(\delta m^{\text{lattice}}_{T,i})^2} + \frac{(\Gamma_{T,i} - \Gamma^{\text{lattice}}_{T,i})^2}{(\delta \Gamma^{\text{lattice}}_{T,i})^2}$$

$$T \in \{0, 151, 173, 199, 251, 334\} \text{ MeV}$$

$$i \in \{1S, 2S, 3S, 1P, 2P\}$$

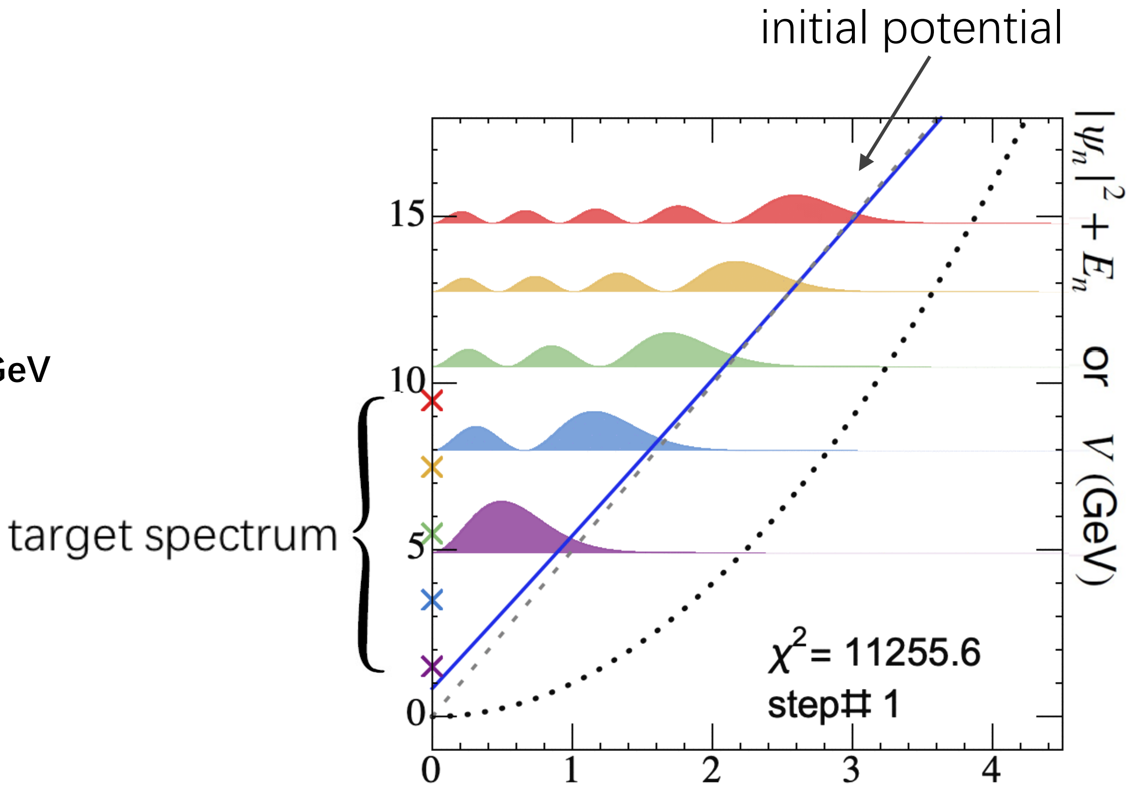


S.S, K. Z, J.Z, S.M., P. Z, Phys. Rev. D 105 (2022) 1, 1

limited spectrum { En } to continuous interaction V(r) ?

Learn V(r) from 5 eigenvalues :

{ En } = {3/2, 7/2, 11/2, 15/12, 19/2} GeV



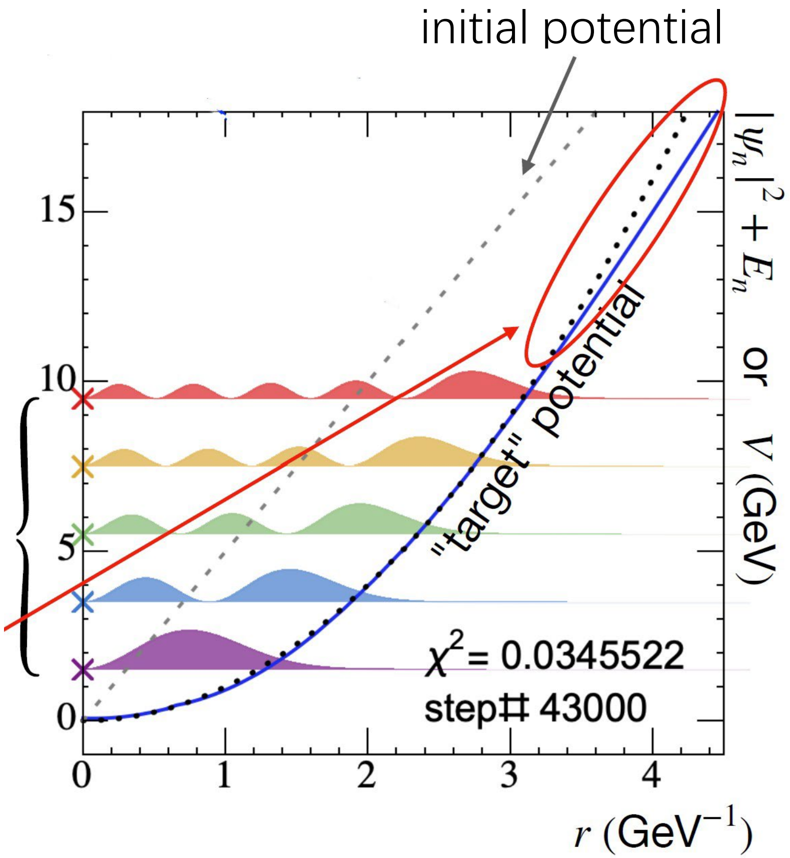
limited spectrum { En } to continuous interaction V(r) ?

-- Yes! But to some range decided by the used states.

Learn V(r) from 5 eigenvalues :

{ En } = {3/2, 7/2, 11/2, 15/12, 19/2} GeV

target spectrum

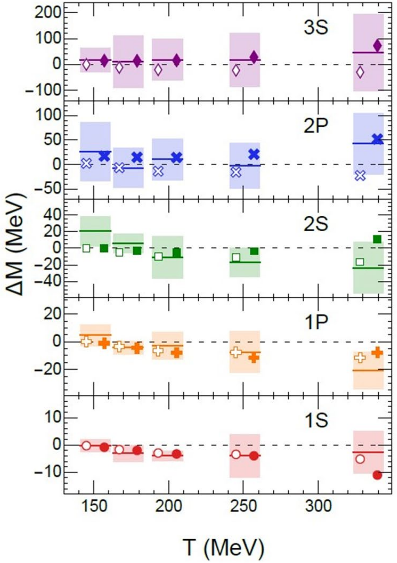


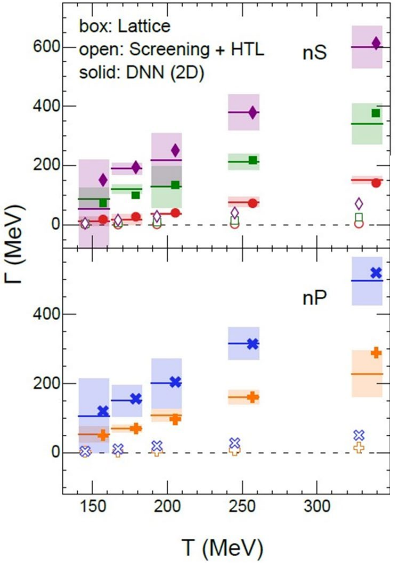
Deviation @ given states' wavefunction vanishes

$$\delta E_n = \langle \psi_n | \delta V(r) | \psi_n \rangle$$

Results with lattice data for mass/width and the reconstructed HQ Potential

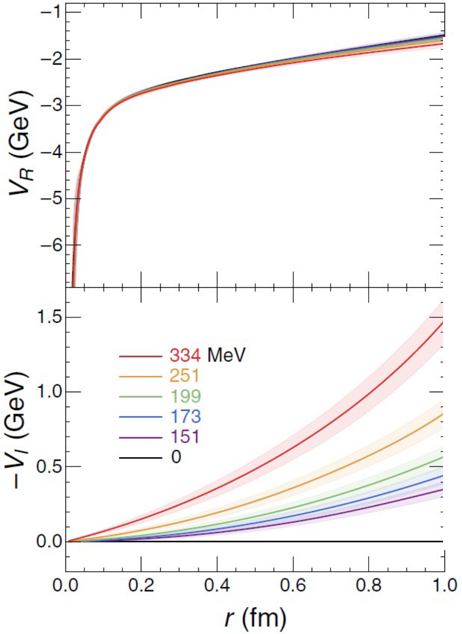
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

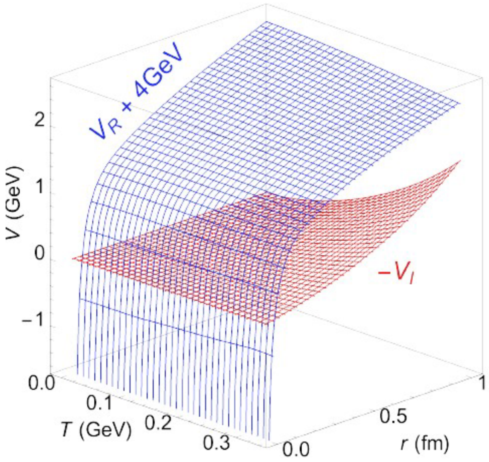




Chi2-per-data=16.5/30

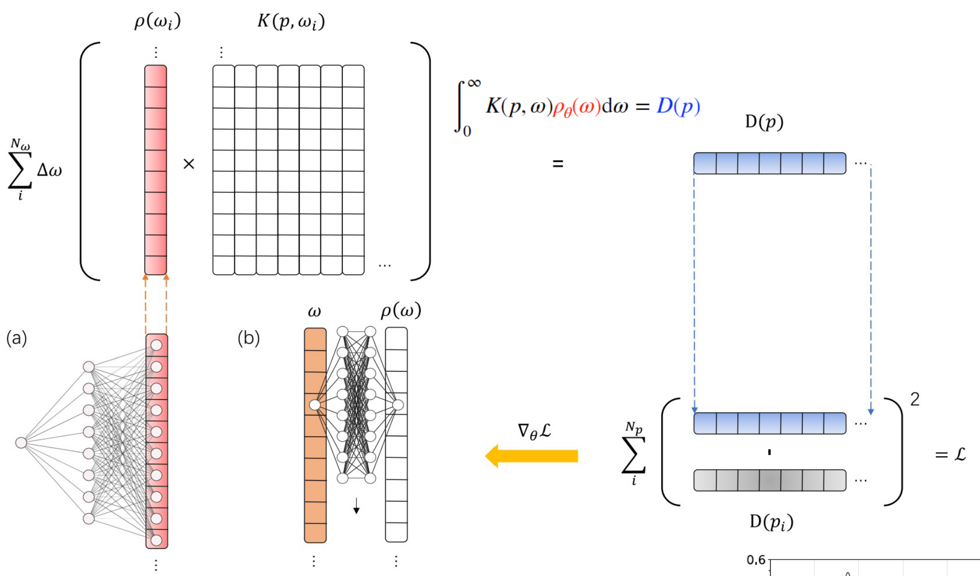
The reconstructed T, r dependent potential

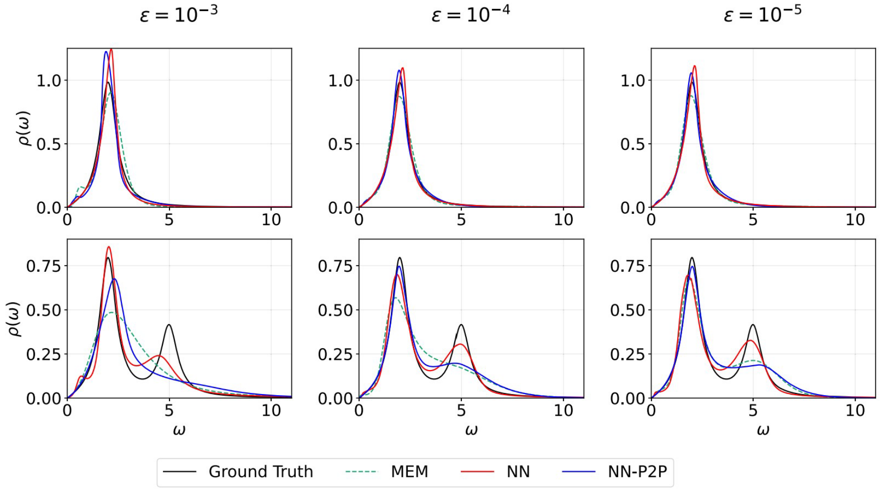


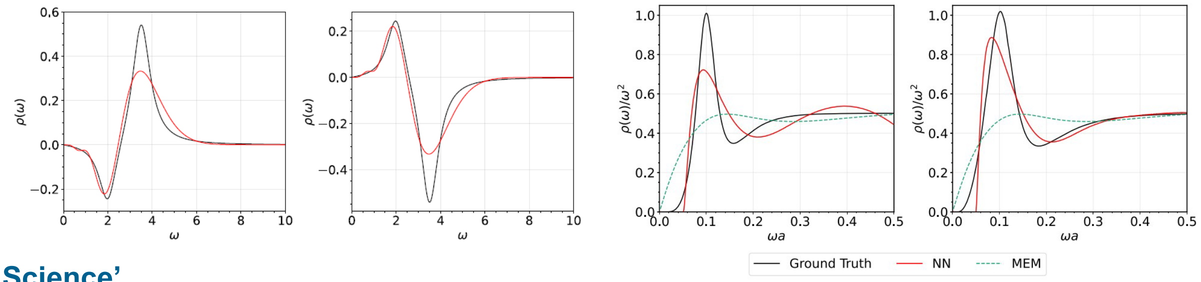


S.S, K. Z, J.Z, S.M., P. Z, Phys. Rev. D 105 (2022) 1, 1

Spectral function reconstruction from Euclidean correlator







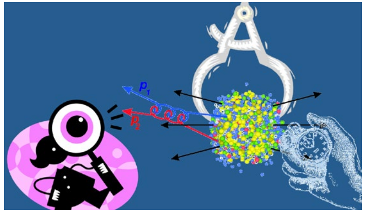
$$G(\tau, T) = \int_0^\infty \frac{d\omega}{2\pi} K(\omega, \tau, T)\rho(\omega, T)$$

$$K(\omega, \tau, T) = \frac{\cosh \omega(\tau - \frac{1}{2T})}{\sinh \frac{\omega}{2T}}$$

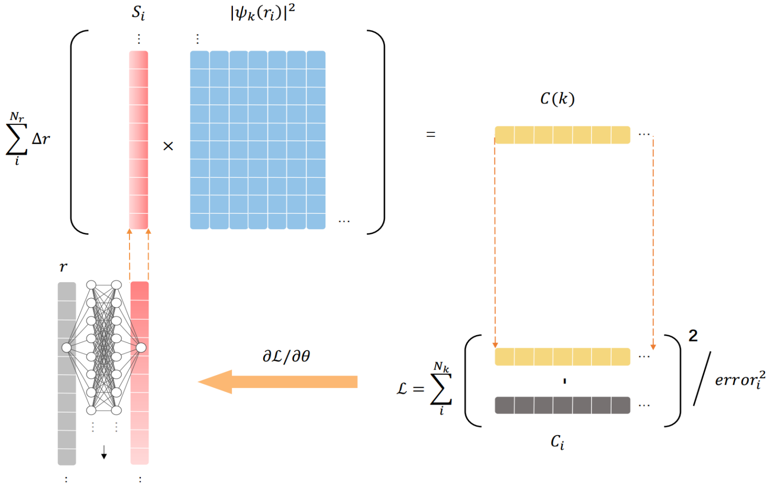
L. Wang, S. Shi and **K. Zhou,**

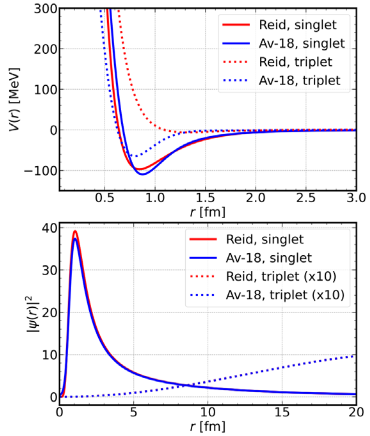
NeurIPS2021 'Machine learning and the Physical Science',
Phys. Rev. D 106, L051502 (Letter),
Computer Physics Communications (2022) 108547,

Hadron emission source reconstruction via femtoscopy



$$C(k) = \int S(\mathbf{r}) \mid \psi_k(\mathbf{r}) \mid^2 d^3 r,$$

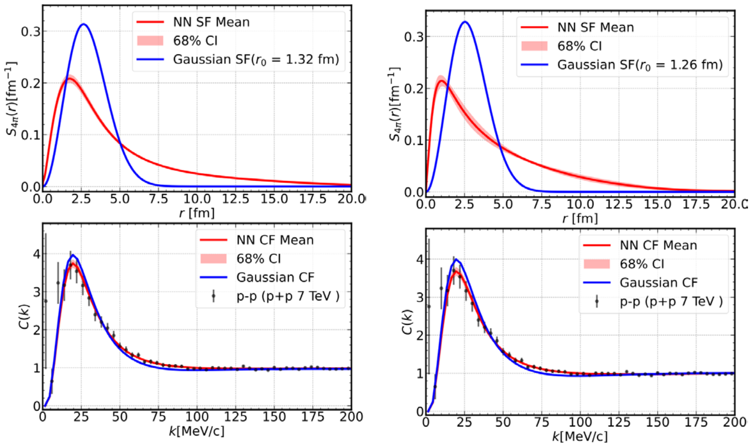




Learning Hadron Emitting Sources with Deep Neural Networks

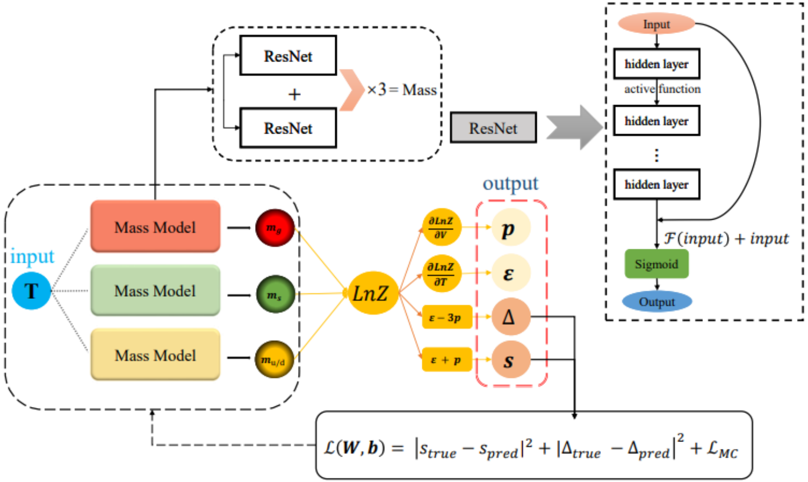
Lingxiao Wang[1] and Jiaxing Zhao[2,3,*]

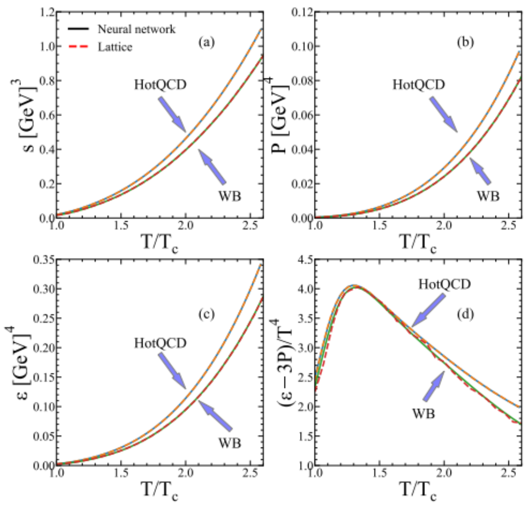
arXiv:2411.16343

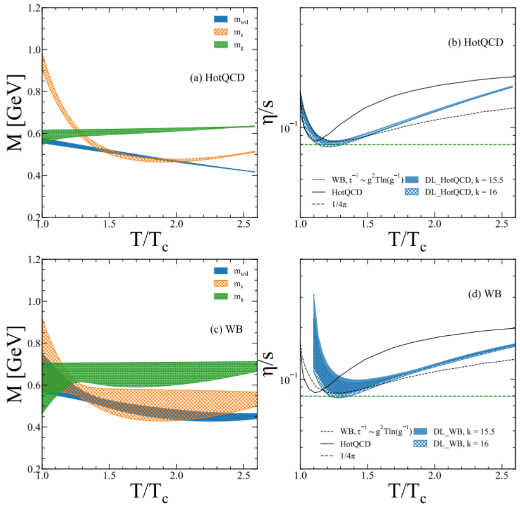


Quasi-particle analysis of lQCD thermodynamics

香 港 中 文 大 學 (深 圳)
The Chinese University of Hong Kong, Shenzhen







$$\ln Z_g(T) = -\frac{16V}{2\pi^2}\int_0^\infty p^2 dp \ln\left[1 - \exp\left(-\frac{1}{T}\sqrt{p^2 + m_g^2(T)}\right)\right], \quad (2)$$

$$\ln Z_{q_i}(T) = +\frac{12V}{2\pi^2}\int_0^\infty p^2 dp \ln\left[1 + \exp\left(-\frac{1}{T}\sqrt{p^2 + m_{q_i}^2(T)}\right)\right], \quad (3)$$

$$P(T) = T\left(\frac{\partial \ln Z(T)}{\partial V}\right)_T, \quad (5)$$

$$\epsilon(T) = \frac{T^2}{V}\left(\frac{\partial \ln Z(T)}{\partial T}\right)_V, \quad (6)$$

See talk of Fu-peng Li tomorrow morning 10:20am

$$\chi_i^B = \left.\frac{\partial P(T, \hat{\mu}_B)/T^4}{\partial \hat{\mu}_B^i}\right|_{\hat{\mu}_B=0}, \hat{\mu}_B = \mu_B/T.$$

$$\mathcal{L}(\theta_1, \theta_2, \theta_3) = |s_{NN} - s_{input}| + \left|\frac{\Delta_{NN} - \Delta_{input}}{T}\right| + |\chi^B_{2,NN} - \chi^B_{2,input}| + |\chi^B_{4,NN} - \chi^B_{4,input}| + \mathcal{L}_{MC}$$

F. Li, H. Lue, L. Pang and G. Qin, **Phys. Lett. B** 2023, arXiv:2211.07994

Recently generalized to finite baryon chemical potential (use lQCD res with Taylor expansion) → **arXiv:2501.10012**

Neural-net Quantum State + VMC for Nuclear Many-Body Problem

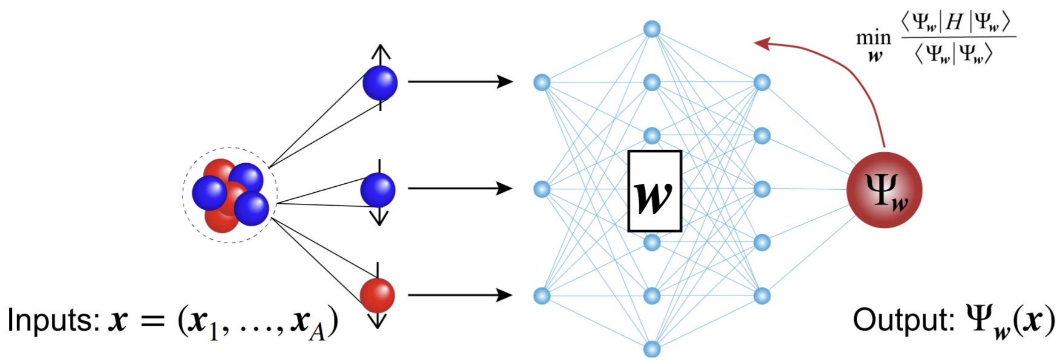
$$H = \sum_{i=1}^{A} \frac{p_i^2}{2M_i} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk}$$

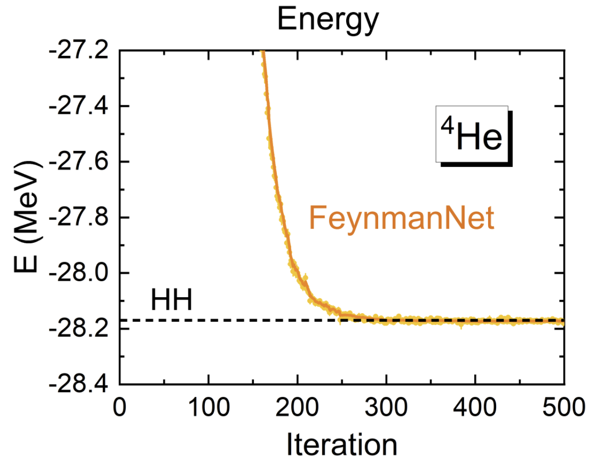
$$H\Psi(x_1, x_2, \ldots, x_A) = E\Psi(x_1, x_2, \ldots, x_A)$$

See talk of Yi-Long Yang this afternoon, 15:40pm

$$\psi_T = \psi_T(R, \vec{\theta})$$

$$E_T = \frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \geq E_0$$





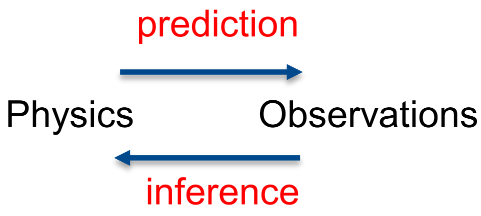
FeynmanNet: YL Yang, PWZ*, **PRC 107, 034320 (2023)**

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

- Merge ML/DL with accumulated Physics Knowledge via AD
- Differentiable Physics Programming to directly invert the simulation

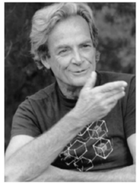
Key words: inverse control, inverse design

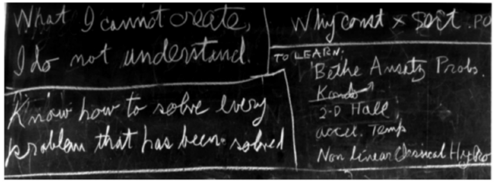
- Physics priors still important in DNN representation constraints
- Uncertainties can be obtained via Bayesian perspective training : Langevin



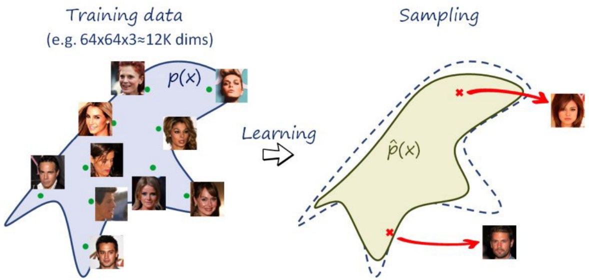
Generative Models

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen





"What I can not create, I do not understand"



Want to **model** the observed data's underlying but unknown **distribution**, to further :

- Understand/Inference the data (inherent structure, properties, features...)
- Sample according to the distribution

Suppose observation dataset :

$$\mathbf{X} = \{x^{(1)}, x^{(2)}, ..., x^{(N)}\} \overset{i.i.d}{\sim} p_{data}(x)$$

We use parametric model to approach the data distribution :

$$p_\theta(x) \to p_{data}(x)$$

Often use NN to parametrize transformation $\log p_\theta(\mathbf{x}) = \log p_\theta(f_\theta(\mathbf{z}_0))$

- Maximize Likelihood Estimation : **(given training samples)**

$$\theta^* = \arg\max_\theta \log p_\theta(\mathbf{X}) = \arg\max_\theta \frac{1}{N}\sum_{i=1}^{N} \log p_\theta(x^{(i)})$$

Reverse KL Divergence : Sample many $\mathbf{z}_0 \sim p_0(\mathbf{z}_0)$
(given unnormalized target distribution, e.g., Action)

$$\theta^* = \arg\min_\theta \frac{1}{N}\sum_{i=1}^{N}[\log p_\theta(f_\theta(\mathbf{z}_0)) - \log \tilde{p}_{target}(f_\theta(\mathbf{z}_0))]$$

- Variational free energy minimization - Reverse KL divergence

$$p(\mathbf{s}) = \frac{e^{-\beta E(\mathbf{s})}}{Z}$$

$$D_{\mathrm{KL}}(q_\theta \,\|\, p) = \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \ln\left(\frac{q_\theta(\mathbf{s})}{p(\mathbf{s})}\right) = \beta(F_q - F)$$

$$F_q = \frac{1}{\beta}\sum_{\mathbf{s}} q_\theta(\mathbf{s})\left[\beta E(\mathbf{s}) + \ln q_\theta(\mathbf{s})\right]$$

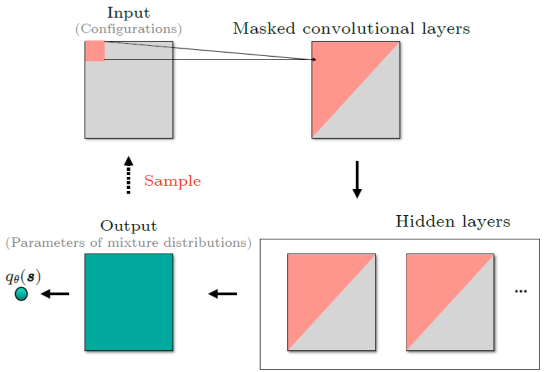
$$\mathop{\mathbb{E}}_{X\sim p(X)}\left[E(X) + k_B T \ln p(X)\right]$$

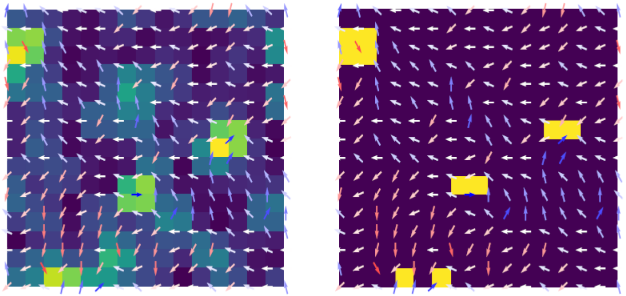
- **Autoregressive** $q_\theta(\mathbf{s}) = \prod_{i=1}^{N} q_\theta(s_i \mid s_1, \ldots, s_{i-1})$

D. Wu, Lei Wang and P. Zhang, **PRL**122,080602(2019)

- **Continuous** Autoregressive Net for XY model

L. Wang, Y. Jiang, L. He, K. Zhou, **CPL** 39, 120502 (2022)





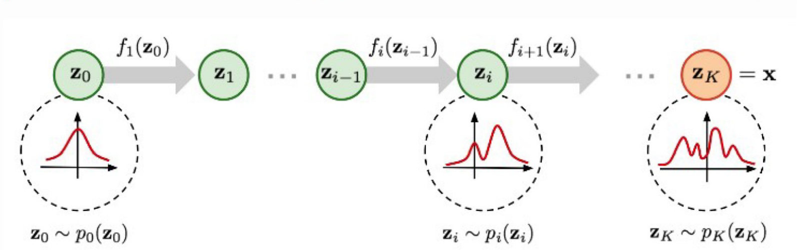
Probability distributions from CANs

Vortices

Flow based generative model given unnormalized distribution

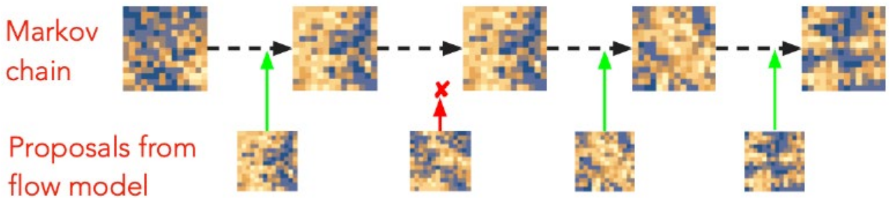
A series **(Flow)** of invertible/bijective transformations for p(z)

compose several invertible transformations to form the flow :



$$p_i(\mathbf{z}_i) = p_{i-1}(f_i^{-1}(\mathbf{z}_i))|\det J_{f_i^{-1}}| = p_{i-1}(\mathbf{z}_{i-1})|\det J_{f_i}|^{-1}$$

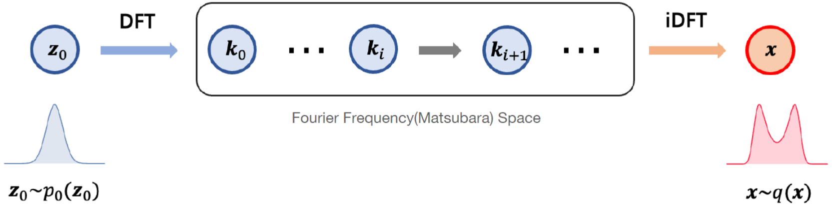
$$\rightarrow \log p(\mathbf{x}) = \log p_0(f^{-1}(\mathbf{x})) + \sum^{K} \log|\det J_{f_i^{-1}}| = \log p_0(\mathbf{z}_0) - \sum^{K} \log|\det J_{f_i}|$$

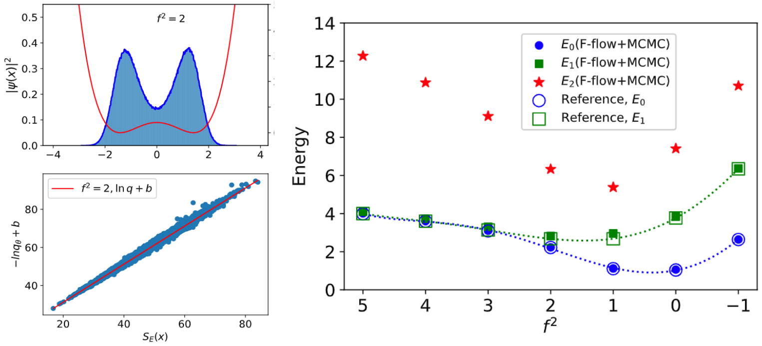


Albergo +, 1904.12072; Boyda +, 2008.05456; Favoni +, 2012.12901; Abbott +, 2208.03832; Abbott +, 2211.07541; Abbott +, 2305.02402; Bulgarelli+ 2412.00200 (SU(3)); Abbott +, arXiv:2502.00263

K.C, G. K., S. R., D. R., P. S., **Nature Reviews Physics** 5, 526-535 (2023)

Fourier Flow Model



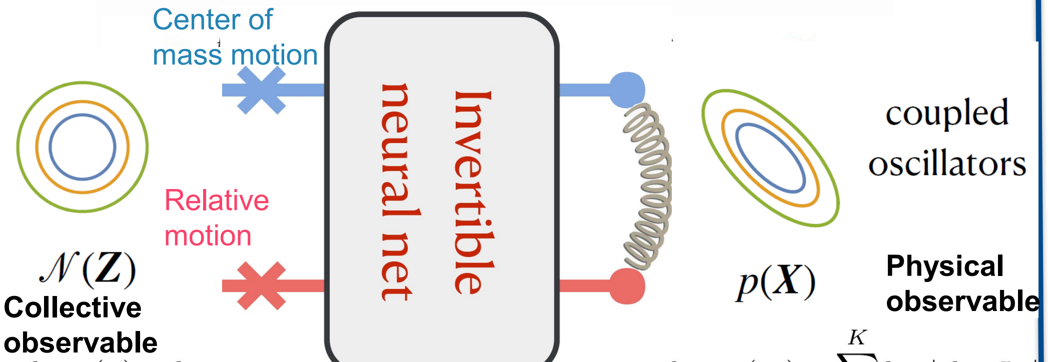


S.C, O. S, S. Z, B. C, H. S, L. W, **K. Zhou**,
Phys. Rev. D 107, 056001(2023)

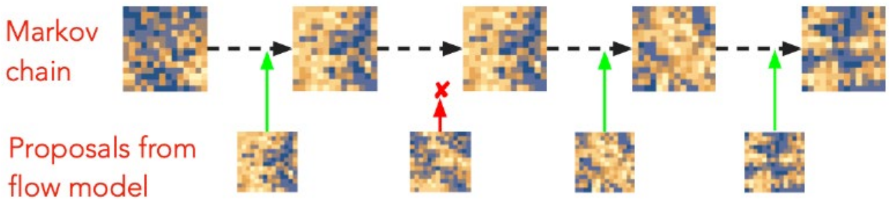
Flow based generative model given unnormalized distribution

A series **(Flow)** of invertible/bijective transformations for p(z)

<u>compose several invertible transformations to form the flow :</u>



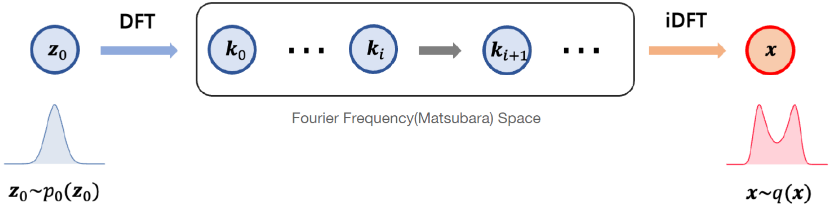
$$\to \log p(\mathbf{x}) = \log p_0(\mathbf{z}_0) - \sum^{K} \log |\det J_{f_i}|$$

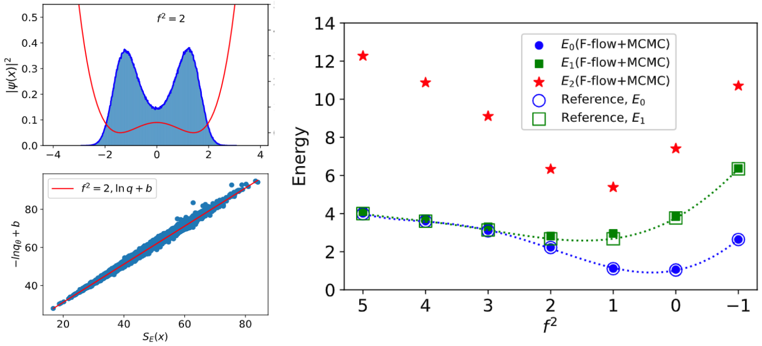


Albergo +, 1904.12072; Boyda +, 2008.05456; Favoni +, 2012.12901; Abbott +, 2208.03832; Abbott +, 2211.07541; Abbott +, 2305.02402; Bulgarelli+ 2412.00200 (SU(3)); Abbott +, arXiv:2502.00263

K.C, G. K., S. R., D. R., P. S., **Nature Reviews Physics** 5, 526-535 (2023)

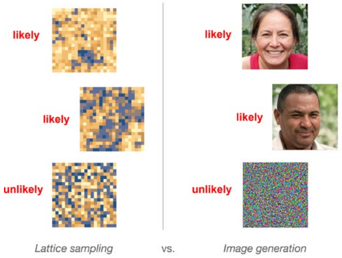
Fourier Flow Model



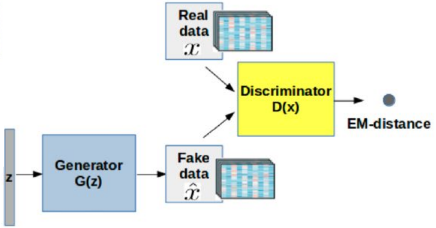


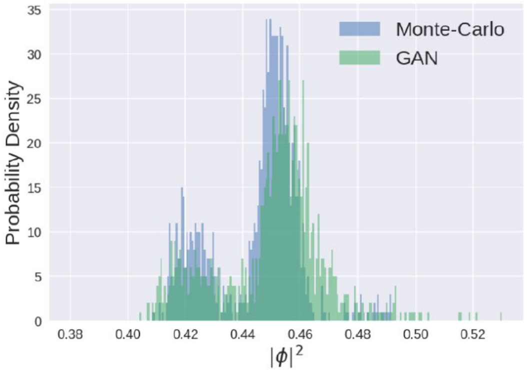
S.C, O. S, S. Z, B. C, H. S, L. W, **K. Zhou**,
Phys. Rev. D 107, 056001(2023)

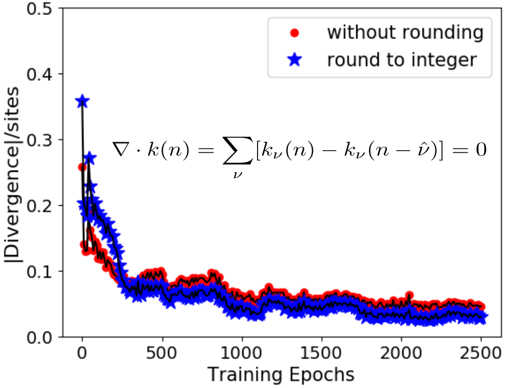
Given an ensemble of data from the target distribution

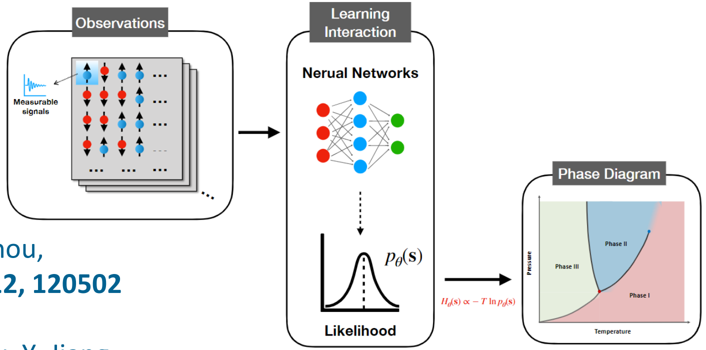


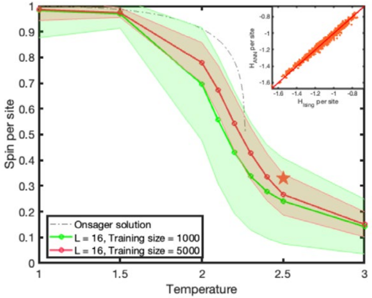
K. Zhou, G. Endrődi, L.P, and H. S
Phys. Rev. D 100, 011501 (2019)

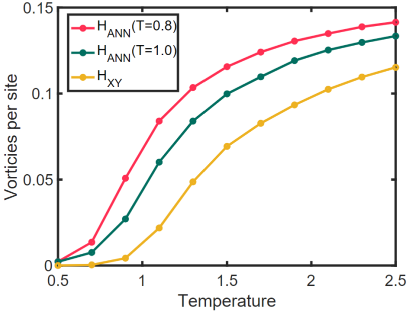












L. Wang, L. He, Y. Jiang, K. Zhou,
Chin. Phys. Lett. 39 (2022) 12, 120502

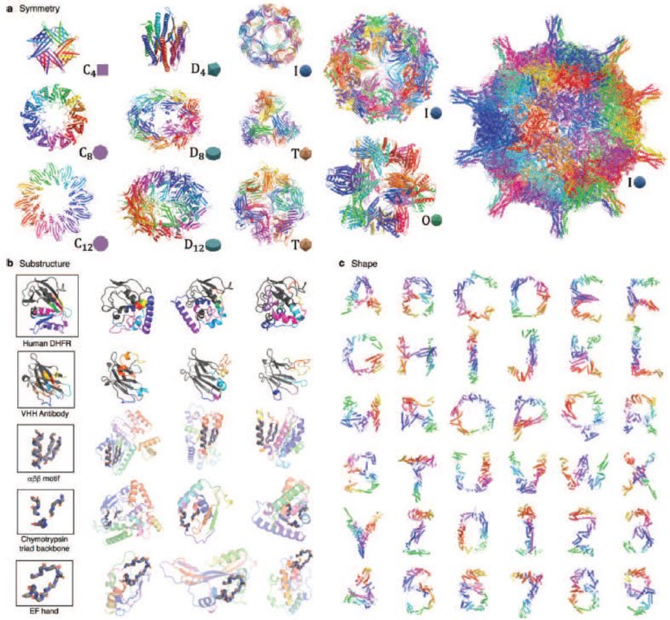
T. Xu, L. Wang, L. He, K. Zhou, Y. Jiang,
Chin. Phys. C 48 (2024) 10, 103101

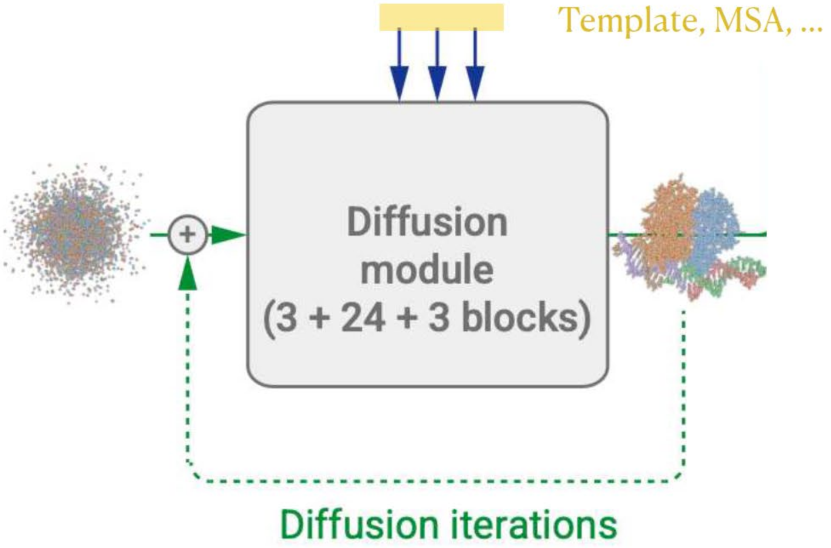
Diffusion Model

"A heavy quark move inside quark-gluon plasma"

Diffusion Model

protein structure prediction and design





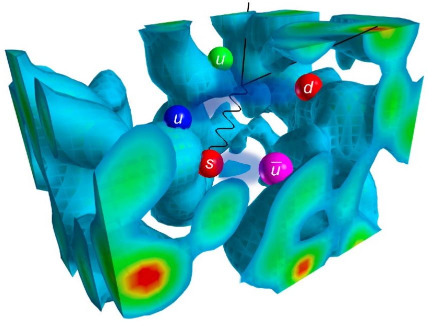
Ingraham et al, Chroma, Nature 2023
https://generatebiomedicines.com/chroma

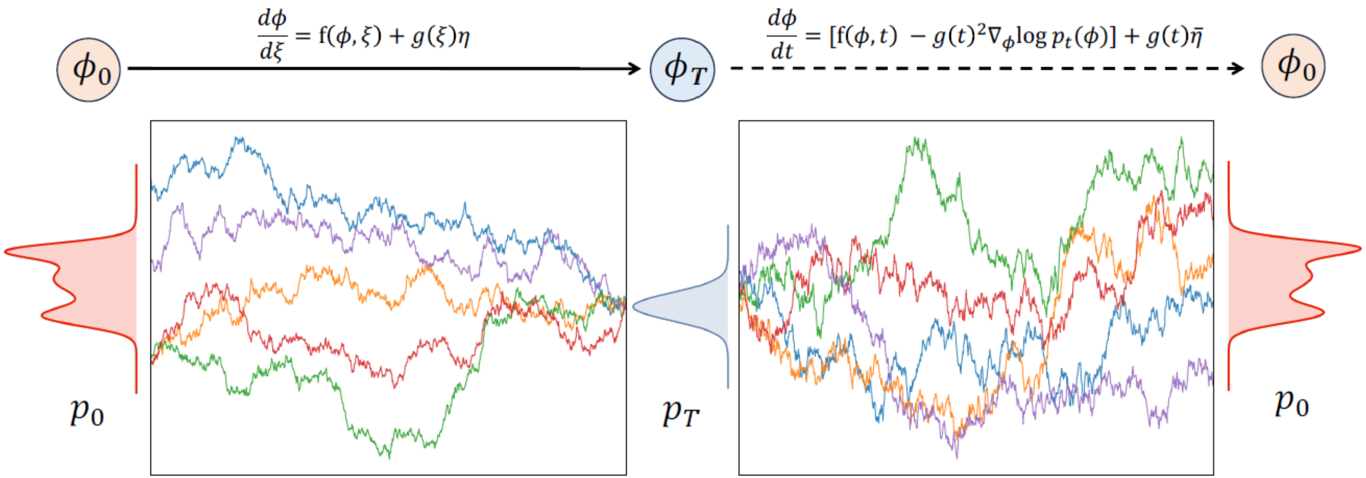
Abramson et al, AlphaFold3, Nature 2024
https://deepmind.google/technologies/alphafold/

Diffusion Model meets lattice QFT

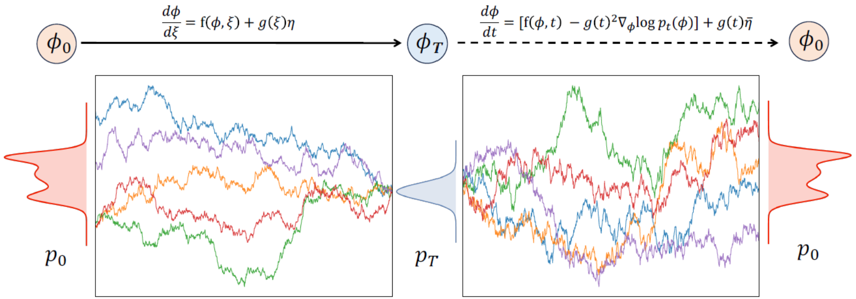
$$p(\phi) = e^{-S(\phi)}/Z$$

$$\langle O \rangle \approx \frac{1}{N} \sum_i O_i$$





L. Wang, G. Aarts, K. Zhou, **JHEP 05 (2024) 060**
L. Wang, G. Aarts, K. Zhou, **arXiv:2311.03578** (NeurIPS 2023 workshop "ML&Physical Sciences")
G. A, D. E. H, L. W, K. Zhou, **arXiv:2410:21212** (NeurIPS 2024 workshop "ML&Physical Sciences") → **"Best Physics for AI Paper" Award**
Q. Zhu, G. Aarts, W. Wang, K. Zhou, L. Wang, **arXiv:2410.19602** (NeurIPS 2024 workshop "ML&Physical Sciences)
G.Aarts, D.E.H, L.W, K. Zhou, **arXiv:2510.01328**



- Forward diffusion SDE

$$\frac{d\phi}{d\xi} = f(\phi,\xi) + g(\xi)\eta(\xi) \qquad \langle \eta(\xi)\eta(\xi')\rangle = 2\alpha\delta(\xi-\xi')$$

- Backward diffusion SDE

$$\frac{d\phi}{dt} = [f(\phi,t) - g^2(t)\nabla_\phi \log p_t(\phi)] + g(t)\bar{\eta}(t) \qquad t \equiv T - \xi$$

- Score matching Training $\mathcal{L}_\theta = \sum_{i=1}^{N} \sigma_i^2 \mathbb{E}_{p_0(\phi_0)} \mathbb{E}_{p_i(\phi_i|\phi_0)} \left[\| s_\theta(\phi_i,\xi) - \nabla_{\phi_i} \log p_i(\phi_i|\phi_0) \|_2^2 \right]$

- Sample generation SDE in **variance exploding scheme :** $\frac{d\phi}{d\tau} = g_\tau^2 \nabla_\phi \log q_\tau(\phi) + g_\tau \bar{\eta}(\tau) \qquad \tau \equiv T - t$

$$\frac{\partial p_\tau(\phi)}{\partial \tau} = \int d^n x \left\{ g_\tau^2 \frac{\delta}{\delta\phi} \left(\bar{\alpha}\frac{\delta}{\delta\phi} + \nabla_\phi S_{\rm DM} \right) \right\} p_\tau(\phi),$$

$$p_{\rm eq}(\phi) \propto e^{-S_{\rm DM}/\bar{\alpha}}$$

$$\nabla_\phi S_{\rm DM} \equiv -\nabla_\phi \log q_\tau(\phi)$$

- A flow of **effective action** will be learned in DMs

sampling from a DM is equivalent to optimizing a stochastic trajectory to approach the "equilibrium state"

L. Wang, G. Arts, K. Zhou, **JHEP 05(2024) 060**

Diffusion Model for field configurations

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

- Forward diffusion SDE

$$\frac{d\phi}{d\xi} = f(\phi,\xi) + g(\xi)\eta(\xi) \qquad \langle \eta(\xi)\eta(\xi')\rangle = 2\alpha\delta(\xi-\xi')$$

- Backward diffusion SDE

$$\frac{d\phi}{dt} = [f(\phi,t) - g^2(t)\nabla_\phi \log p_t(\phi)] + g(t)\bar{\eta}(t) \qquad t \equiv T-\xi$$

- Physical motivation
  - Destroy structure in data through a diffusive process.
  - Carefully record the destruction.
  - Use deep networks to reverse time and create structure from noise.

- Score matching Training $\mathcal{L}_\theta = \sum_{i=1}^{N} \sigma_i^2 \mathbb{E}_{p_0(\phi_0)} \mathbb{E}_{p_i(\phi_i|\phi_0)} \left[\| s_\theta(\phi_i,\xi) - \nabla_{\phi_i} \log p_i(\phi_i|\phi_0) \|_2^2 \right]$

- Sample generation SDE in **variance exploding scheme :** $\frac{d\phi}{d\tau} = g_\tau^2 \nabla_\phi \log q_\tau(\phi) + g_\tau \bar{\eta}(\tau) \qquad \tau \equiv T - t$

$$\frac{\partial p_\tau(\phi)}{\partial \tau} = \int d^n x \left\{ g_\tau^2 \frac{\delta}{\delta\phi} \left(\bar{\alpha} \frac{\delta}{\delta\phi} + \nabla_\phi S_{\rm DM} \right) \right\} p_\tau(\phi),$$

$$\boxed{p_{\rm eq}(\phi) \propto e^{-S_{\rm DM}/\bar{\alpha}}}$$

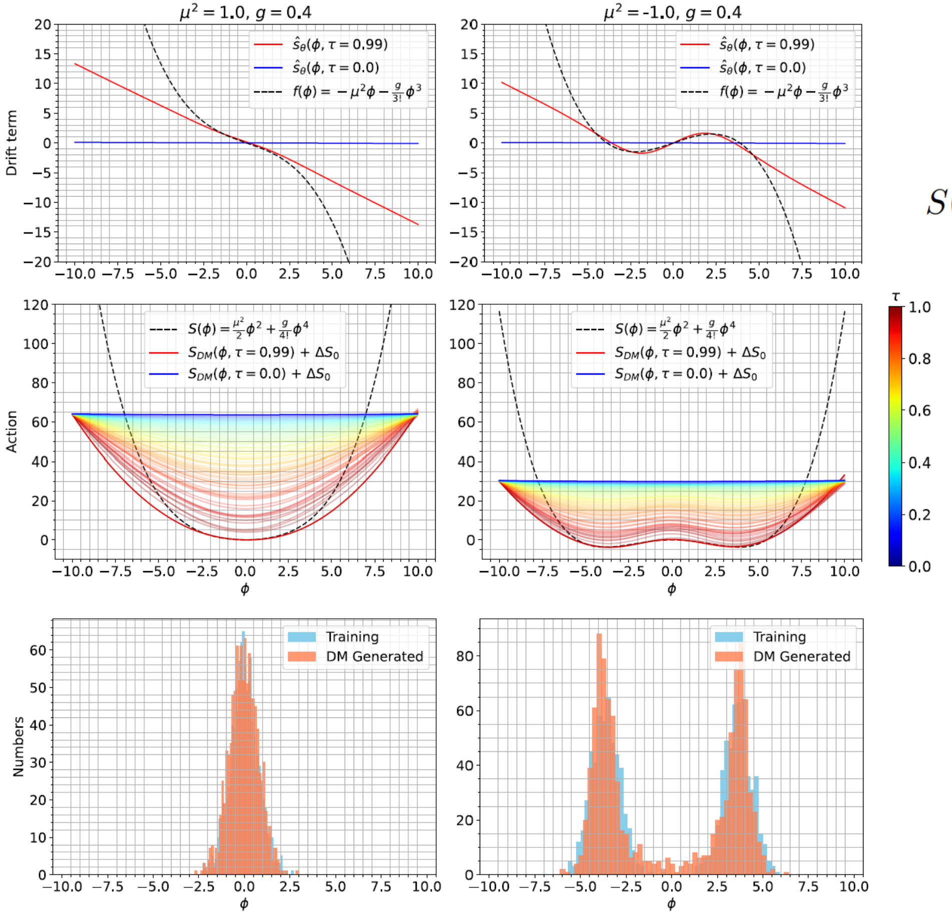
$$\nabla_\phi S_{\rm DM} \equiv -\nabla_\phi \log q_\tau(\phi)$$

- A flow of **effective action** will be learned in DMs

sampling from a DM is equivalent to optimizing a stochastic trajectory to approach the "equilibrium state"

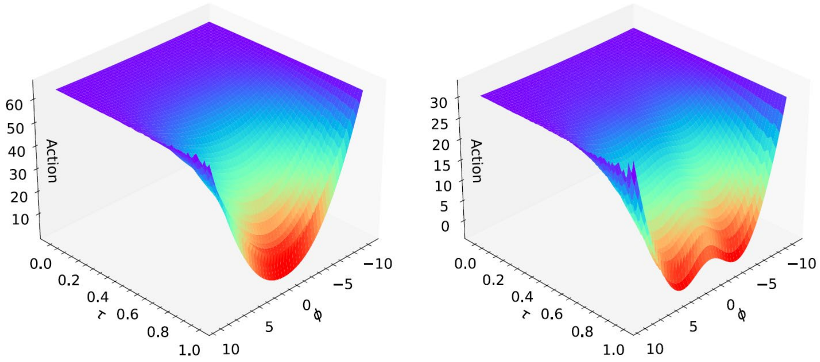
L. Wang, G. Arts, K. Zhou, **JHEP 05(2024) 060**

Effective Action on A Toy model



- Flow of the effective action

$$S(\phi) = \frac{\mu^2}{2}\phi^2 + \frac{g}{4!}\phi^4, \qquad f(\phi) = -\frac{\partial S(\phi)}{\partial \phi} = -\mu^2\phi - \frac{g}{3!}\phi^3$$



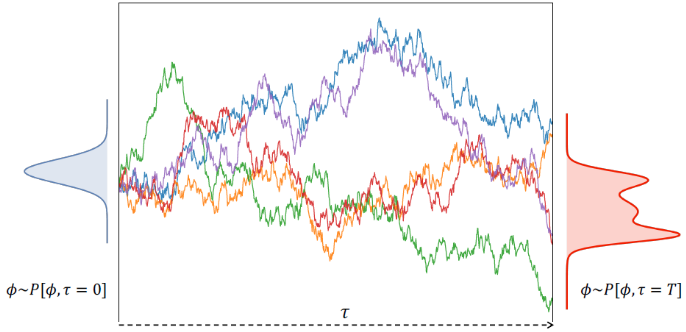
$$S_{\mathrm{DM}}(\phi, \tau) = \int^{\phi} \hat{s}_\theta(\tilde{\phi}, \tau) d\tilde{\phi}$$

$$\frac{\partial\phi(x,\tau)}{\partial\tau} = -\frac{\delta S_E[\phi]}{\delta\phi(x,\tau)} + \eta(x,\tau)$$

$$\langle\eta(x,\tau)\rangle = 0, \quad \langle\eta(x,\tau)\eta(x',\tau')\rangle = 2\alpha\delta(x-x')\delta(\tau-\tau')$$

τ: fictitious time, α: diffusion constant

One can construct stochastic process to reproduce the quantum path integral with its equilibrium:



Thermal equilibrium limit → Quantum distribution

- **Fokker-Planck equation**

$$\frac{\partial P[\phi,\tau]}{\partial\tau} = \alpha\int d^n x \left\{\frac{\delta}{\delta\phi}\left(\frac{\delta}{\delta\phi} + \frac{\delta S_E}{\delta\phi}\right)\right\} P[\phi,\tau]$$

Equilibrium solution (long-time limit),

$$P_{\text{eq}}[\phi] \propto e^{-\frac{1}{\alpha}S_E[\phi]}$$

- Set the diffusion constant as $\alpha = \hbar$

$$P_{\text{eq}}[\phi] \sim e^{-\frac{1}{\hbar}S_E[\phi]} = P_{\text{quantum}}[\phi]$$

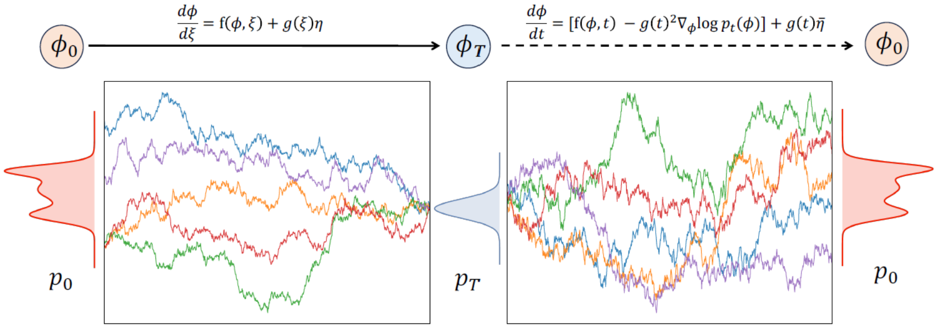
Parisi G. and Wu Y. S., **Sci. China, A 24**, ASITP-80-004 (1980).

Diffusion Model as SQ

- **DM <u>generation SDE</u> and <u>Stochastic Quantization :</u>**

$$\frac{\partial \phi(x,\tau)}{\partial \tau} = g^2(\tau)\nabla_\phi \log P(\phi;\tau) + g(\tau)\eta(x,\tau)$$

$$\frac{\partial \phi(x,\tau)}{\partial \tau} = -\nabla_\phi S(\phi) + \sqrt{2}\eta(x,\tau)$$



- **Similarities and differences:**

✓ SQ: fixed drift, determined from known action
constant noise variance (but can be generalised using kernels)
thermalisation followed by long-term evolution in equilibrium

✓ DM: drift and noise variance time-dependent, learn from data
evolution between $0 \leq \tau \leq T = 1$, many short runs

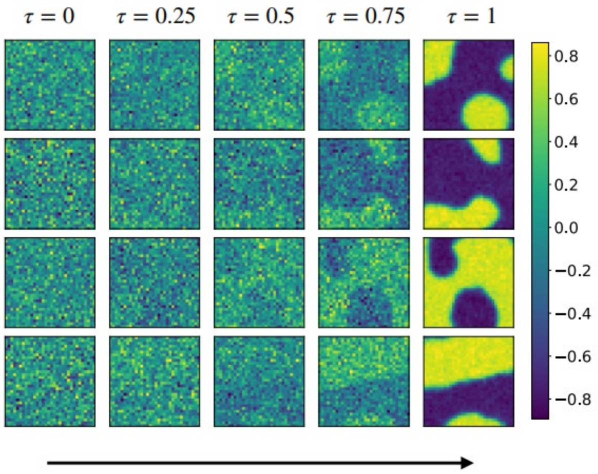
L. Wang, G. Arts, K. Zhou, **JHEP 05(2024) 060**

DM on 2d scalar ϕ^4 model

- 32x32 lattice, HMC generated 5120 configurations as training set

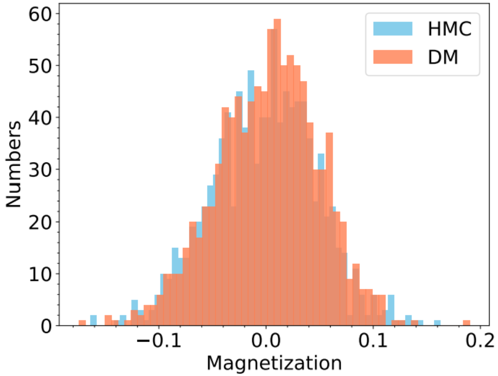
$$S_E = \sum_x [-2\kappa \sum_{\mu=1}^{d} \phi(x)\phi(x+\hat{\mu}) + (1-2\lambda)\phi(x)^2 + \lambda\phi(x)^4].$$

Broken phase :



numerous "bulk" patterns emerge

symmetric phase :

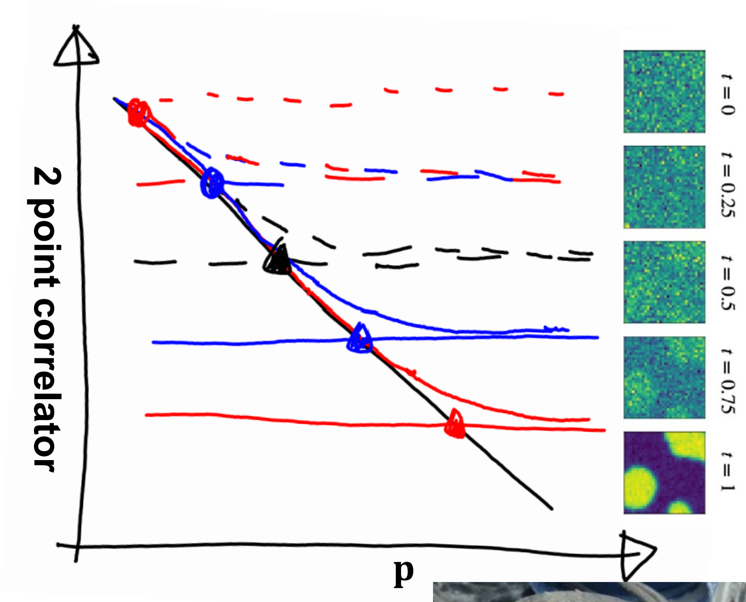


data-set	$\langle M \rangle$	χ_2	U_L
Training (HMC)	0.0012± 0.0007	2.5160 ± 0.0457	0.1042 ± 0.0367
Testing (HMC)	0.0018 ± 0.0015	2.4463 ± 0.1099	-0.0198 ± 0.1035
Generated (DM)	0.0017± 0.0015	2.4227 ± 0.1035	0.0484 ± 0.0959

- Forward diffusion kernel: **gaussian smoothing**

$$p_\xi(\phi_\xi|\phi_0) = \mathcal{N}\left(\phi_\xi; \phi_0, \frac{1}{2\log\sigma}(\sigma^{2\xi}-1)\mathbf{I}\right)$$

$$\phi_\tau(\mathbf{x}) = \phi_0(\mathbf{x}) + \sqrt{\frac{\sigma^{2\tau}-1}{2\log\sigma}}\epsilon(\mathbf{x}) \text{ with } \epsilon \sim \mathcal{N}(0,\mathbf{I})$$



- In **Fourier space**:

$$\phi_\tau(p) = \phi_0(p) + \sqrt{\frac{\sigma^{2\tau}-1}{2\log\sigma}}\epsilon(p)$$

- ! the above evolution will perturb (smear) higher momentum modes first,
- With decreasing cut scale because of the gradually increasing noise level !

↕

In FRG, the high frequency (short-distance) degrees of freedom is progressively integrated out !

Generative model to speed up HIC simulation

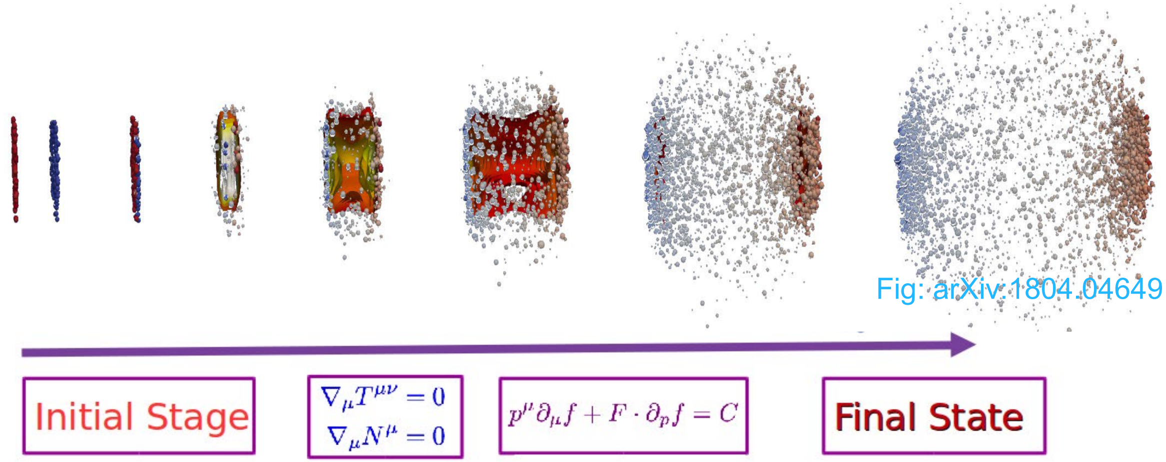


Fig: arXiv:1804.04649

Impact parameter
Nuclear structure

Bulk Matter properties:
Phase Transition
EOS
Shear/Bulk viscosities
Hard Probes

Event Generation
Final observables

Applications to Lattice QFT

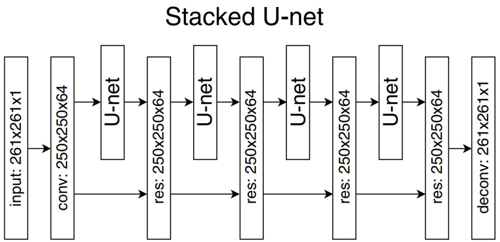
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

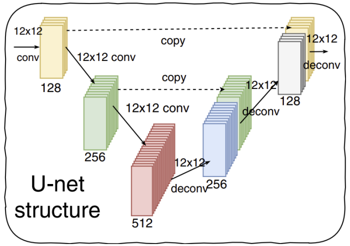
U-net Emulator for relativistic hydrodynamics

PHYSICAL REVIEW RESEARCH **3**, 023256 (2021)

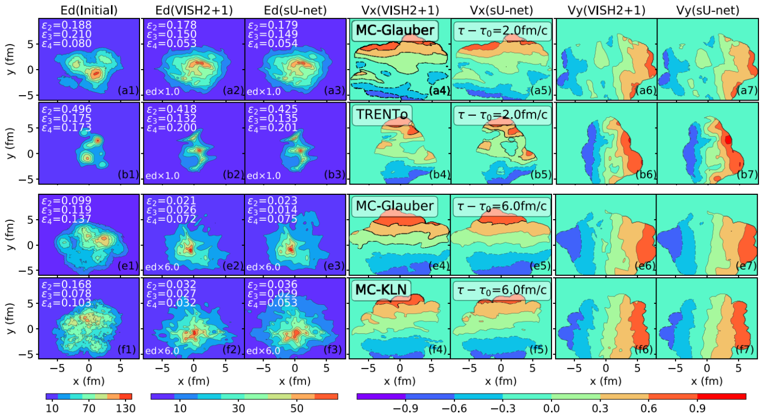
Applications of deep learning to relativistic hydrodynamics

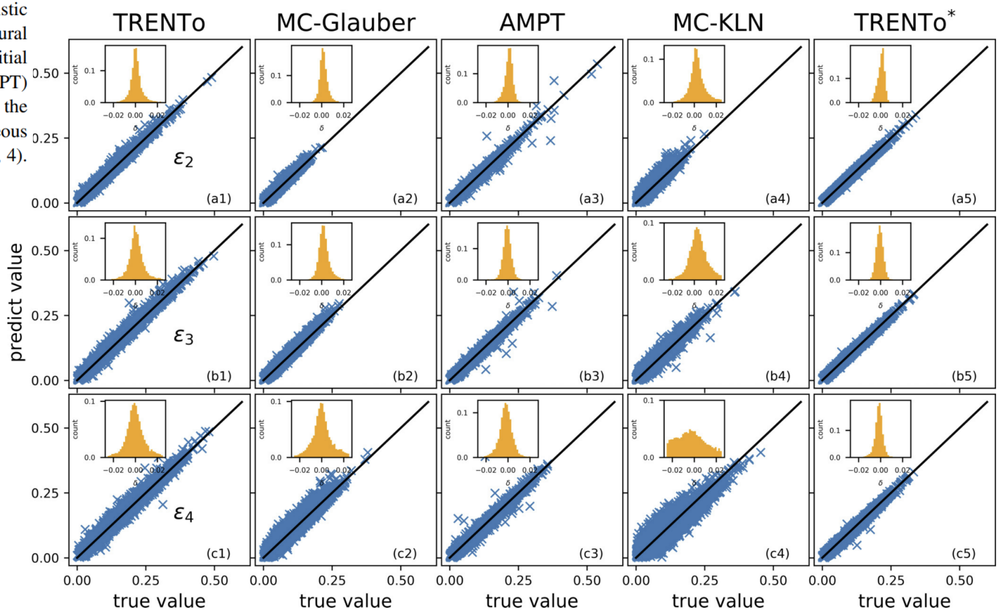
Hengfeng Huang,[1,2] Bowen Xiao,[3] Ziming Liu,[1] Zeming Wu,[1,2] Yadong Mu,[3,4] and Huichao Song[1,2,5]





tic heavy-ion collisions. Using 10 000 initial and final profiles generated from (2+1)-dimensional relativistic hydrodynamics VISH2+1 with Monte Carlo Glauber (MC-Glauber) initial conditions, we train a deep neural network based on the stacked U-net, and use it to predict the final profiles associated with various initial conditions, including MC-Glauber, MC Kharzeev-Levin-Nardi (MC-KLN), a multiphase transport (AMPT) model, and the reduced thickness event-by-event nuclear topology (TRENTo) model. A comparison with the VISH2+1 results shows that the network predictions can nicely capture the magnitude and inhomogeneous structures of the final profiles, and creditably describe the related eccentricity distributions $P(\varepsilon_n)$ $(n = 2, 3, 4)$.





CNN Emulator to hydrodynamic results of heavy-ion collisions

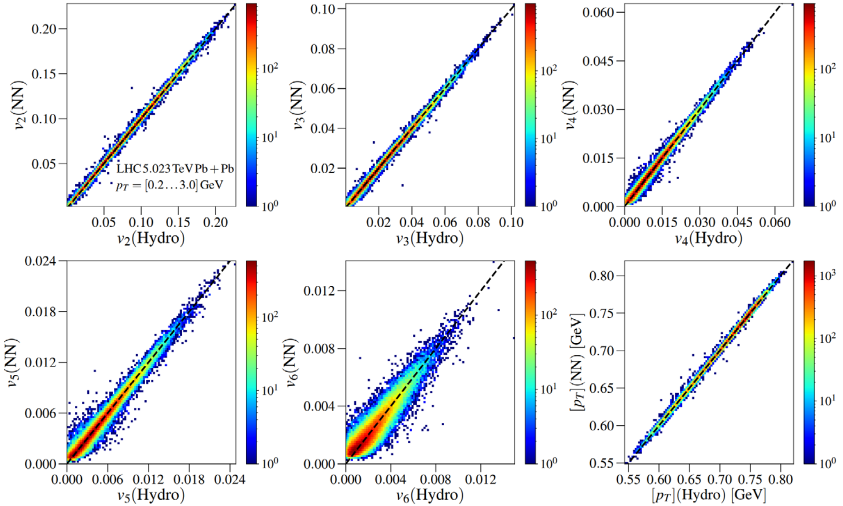
PHYSICAL REVIEW C **108**, 034905 (2023)

Deep learning for flow observables in ultrarelativistic heavy-ion collisions

H. Hirvonen, K. J. Eskola, and H. Niemi

As an input, the DenseNet model uses discretized initial energy density in the transverse-coordinate (x, y) plane calculated from the EKRT model with a grid size 269×269 and a resolution of 0.07 fm. The DenseNet model is trained to reproduce a set of final state p_T integrated observables v_n, average transverse momentum $[p_T]$, and charged particle multiplicity $dN_{ch}/d\eta$ for each event. The input energy density is normalized in such a

Block	Output size	Layers
Convolution	134x134x64	7x7 conv, stride 2
Pooling	67x67x64	3x3 max pool, stride 2
Dense Block	67x67x256	1x1 conv, 3x3 conv x 6
Transition Layer	67x67x128	1x1 conv
	33x33x128	2x2 average pooling, stride 2
Dense Block	33x33x512	1x1 conv, 3x3 conv x 12
Transition Layer	33x33x256	1x1 conv
	16x16x256	2x2 average pooling, stride 2
Dense Block	16x16x896	1x1 conv, 3x3 conv x 20
Transition Layer	16x16x448	1x1 conv
	8x8x448	2x2 average pooling, stride 2
Dense Block	8x8x1216	1x1 conv, 3x3 conv x 24
Output Layer	1x1x1216	8x8 global average pooling
	N_{out}	Fully connected layer with ReLU activation



10^7 events using the neural network, which takes around 20 h with the GPU. This is a very substantial difference compared to doing full hydrodynamic simulations using CPU, which would take about 5×10^6 CPU hours.

Similar study for intermediate energies with IBUU +DNN emulation see: **B. Li+, arXiv:2406.18421**

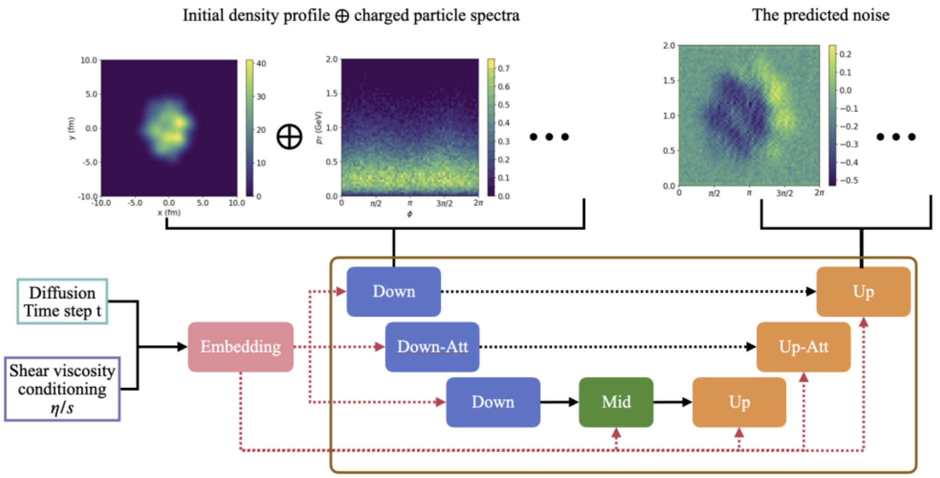
Generative diffusion model to heavy-ion collisions

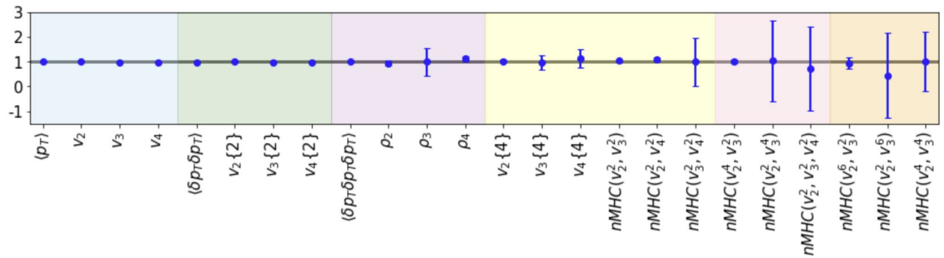
See the talk of
Jing-An Sun
tomorrow 10:40am

An end-to-end generative diffusion model for heavy-ion collisions

Jing-An Sun,[1,2] Li Yan,[1,3] Charles Gale,[2] and Sangyong Jeon[2]

Phys. Rev. C (Letter) 2025
arXiv:2410.13069





tor. We carried out (2+1)D minimum bias simulations of Pb-Pb collisions at 5.02 TeV, choosing the shear viscosity η/s to be one of three distinct values: 0.0, 0.1, and 0.2. For each value of η/s, we generate 12,000 pairs of initial entropy density profiles and final particle spectra, corresponding to 12,000 simulated events, as the training dataset. 70% of the total events are used for training and the rest are used for validation.

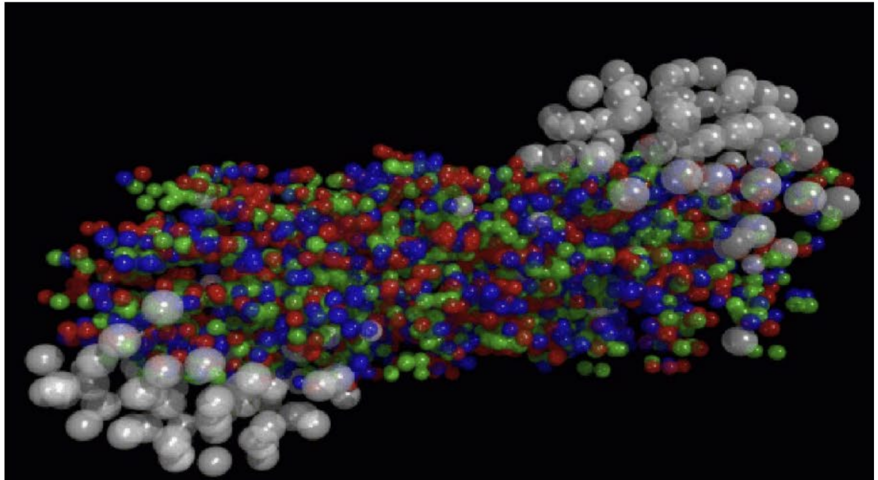
Considering that the spectra $\boldsymbol{S}_0$ depend on the initial entropy density profiles $\boldsymbol{I}$ and the shear viscosity η/s, we train a conditional reverse diffusion process $p(\boldsymbol{S}_0|\boldsymbol{I}, \eta/s)$ without modifying the forward process.

one single central collision event in just 10^{-1} seconds on a GeForce GTX 4090 GPU.

ble precision, as the traditional numerical simulation of hydrodynamics for one central event typically takes approximately 120 minutes (10^4 seconds) on a single CPU.

Point Cloud Diffusion Model for HICs – UrQMD cascade model

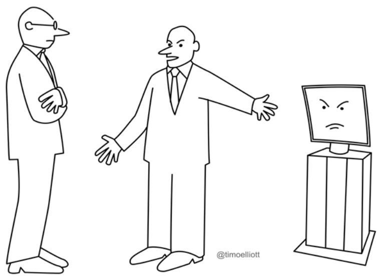
- Event-by-event collision output
- Microscopic non-equilibrium description
- hadrons on classical trajectories
  - stochastic binary scatterings
  - color string formation
  - resonance excitation and decays


https://itp.uni-frankfurt.de/~bleicher/urqmd-cern.gif

- interactions based on scattering cross sections
- default setup effective EoS: Hadron Resonance Gas
- Non-trivial interactions can be added through QMD approach

Can we emulate UrQMD with DL?

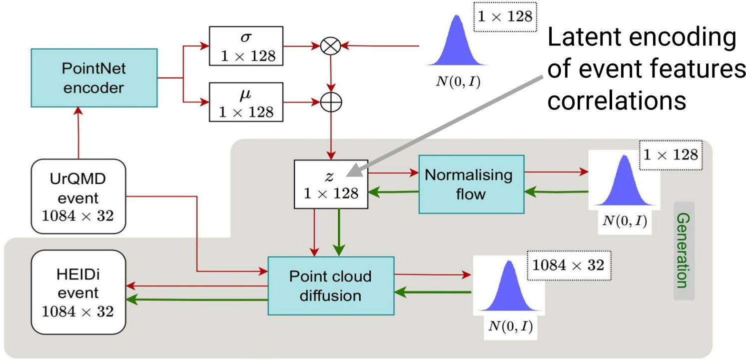
M. O.K, **K. Zhou**, J. S, H. S, **arXiv:2412.10352 , arXiv:2502.16330**

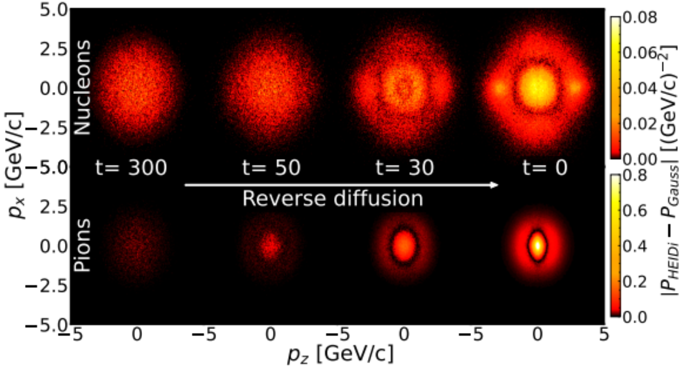


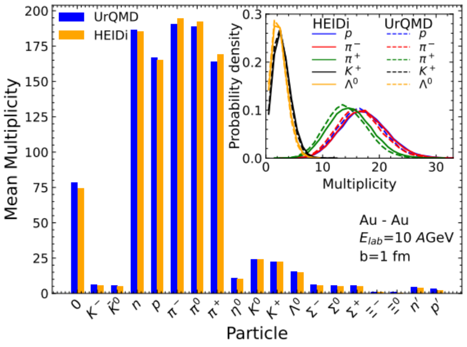
His decisions aren't any better than yours — but they're WAY faster...

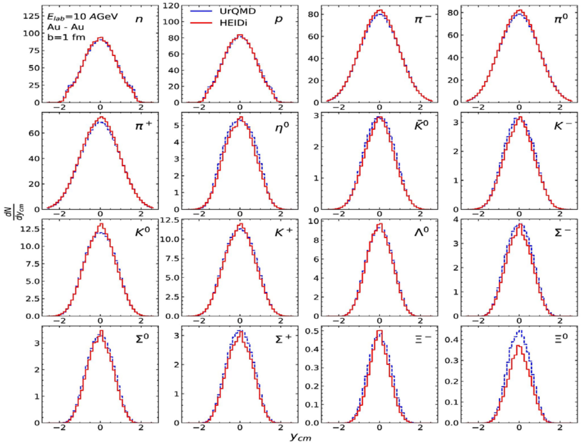
- 18k UrQMD simulation events for central Au-Au@10 AGeV collisions
- HEIDi:
  Heavy-ion Events through Intelligent Diffusion

PointNet encoder + Normalizing flow decoder + Pointcloud diffusion →



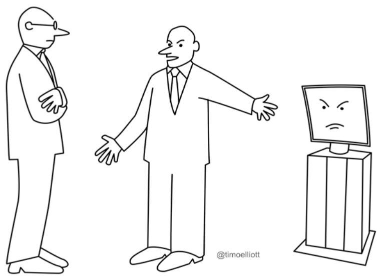






M. O.K, **K. Zhou**, J. S, H. S, arXiv:**2502.16330**, arXiv:**2412.10352**

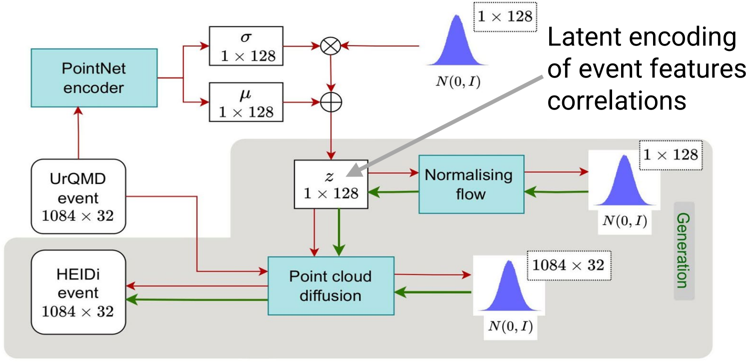
Point Cloud Diffusion Model for HICs – AI clone of simulation

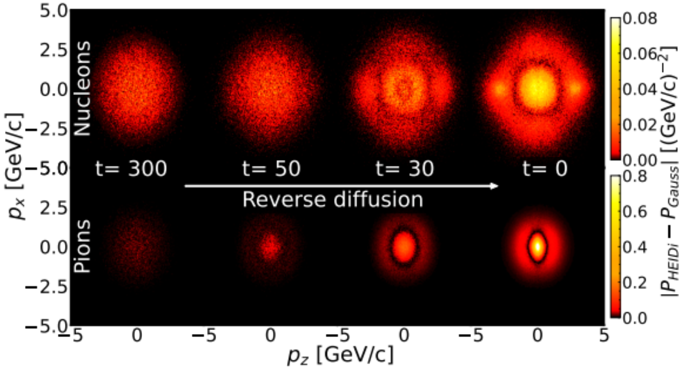


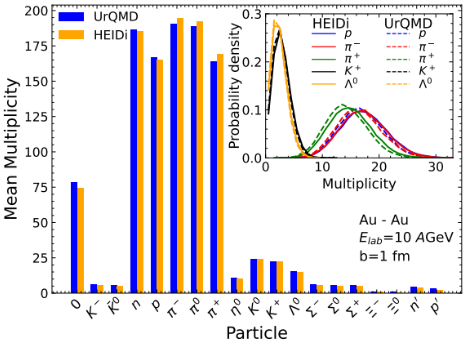
His decisions aren't any better than yours — but they're WAY faster...

- 18k UrQMD simulation events for central Au-Au@10 AGeV collisions
- HEIDi:
  Heavy-ion Events through Intelligent Diffusion

PointNet encoder + Normalizing flow decoder + Pointcloud diffusion →







- **Running time of** UrQMD simulation
  cascade : ~ 3 sec/event;
  with potential : ~ 3 min/event;
  hybrid : ~ 1 hour/event
- HEIDI on A100: ~ 30 ms/event
- Speedup **2 ~ 5 orders** of magnitude

M. O.K, **K. Zhou**, J. S, H. S, arXiv:**2502.16330**, arXiv:**2412.10352**

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Should artificial intelligence be interpretable to humans?

Matthew D. Schwartz

Nature Reviews Physics **4**, 741–742 (2022) | Cite this article

Machine vs. Biological intelligence

- Biological intelligence grows by a factor of 2 in one million years
- Machine intelligence grows by a factor of 10 in 1 year

- **Factor of 10^6 difference in *exponent***

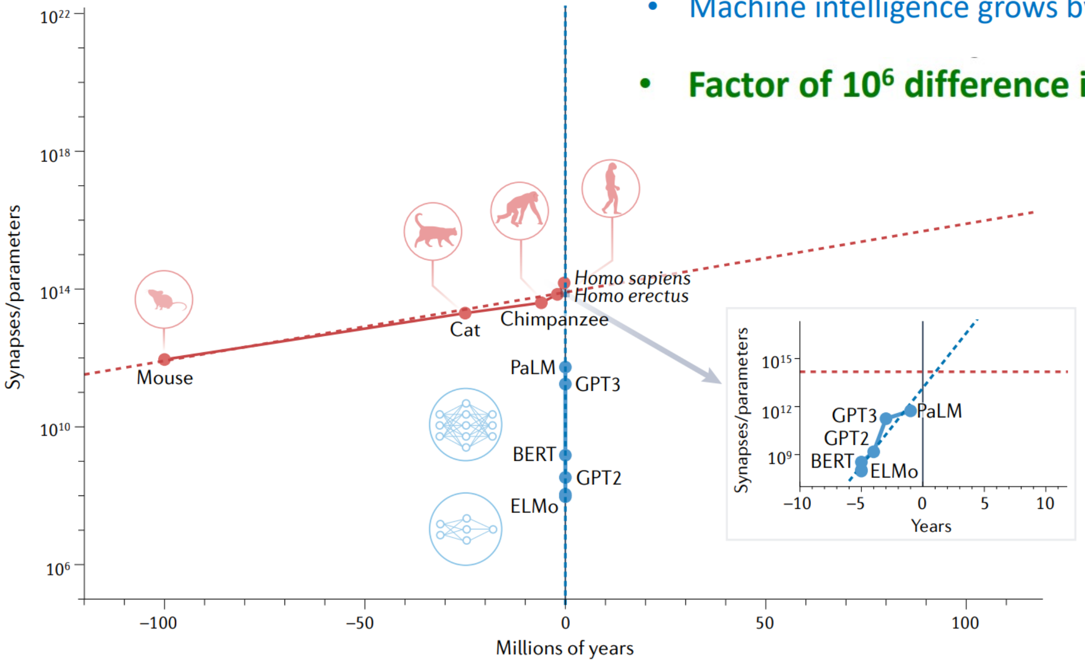


Fig. 1 | The evolution of biological and artificial intelligence takes place on dramatically different timescales. Any hope of interpreting and understanding AI will exponentially fade. Some example data points are highlighted in the evolution of biological (red) and artificial (blue) intelligence. The dashed lines represent the linear regression to these points. The acronyms in the figure are: Pathways Language Model (PaLM), Embeddings from Language Model (ELMo), Bidirectional Encoder Representations from Transformers (BERT), Generative Pre-trained Transformer (GPT).

Chat-GPT

↓

Chat-GPh.T?
Chat-GPh.D?

Summary: Machine Learning and HENP

- **Deep Learning** help bridging HIC experiment with theory/model for physics exploration/inversion   caveat: model dependency

- **Bayesian Inference** for EoS from different beam energy experiment data (v2 and mT) perform well - consistent with dv1/dy measurements and BNSM constraint
  sensitivity check reveals tension: measurement uncertainty or model limitation

- **Auto-diff** help physics extraction taking advantange of GPU and DNN

→ Combined global fit of EoS from HIC and NS obs. ? (need to take care of isospin dependence)

→ More Physics Priors, and, Generative AI with Discriminative AI together as the fifth paradigm for nuclear physics

Nature Review Physics (2025)

Prog. Part. Nucl. Phys. 135 (2024) 104084

Nucl. Sci. Tech. 34 (2023) 6, 88

Summary: Machine Learning and HENP

- **Deep Learning** help bridging HIC experiment with theory/model for physics exploration/inversion caveat: model dependency

- **Bayesian Inference** for EoS from different beam energy experiment data (v2 and mT) perform well - consistent with dv1/dy measurements and BNSM constraint
  sensitivity check reveals tension: measurement uncertainty or model limitation

- **Auto-diff** help physics extraction taking advantange of GPU and DNN

→ Combined global fit of EoS from HIC and NS obs. ? (need to take care of isospin dependence)

→ More Physics Priors, and, Generative AI with Discriminative AI together as the fifth paradigm for nuclear physics

Thanks !

Nature Review Physics (2025)

Prog. Part. Nucl. Phys. 135 (2024) 104084

Nucl. Sci. Tech. 34 (2023) 6, 88

- forward process: $$\dot{x}(t) = K(x(t), t) + g(t)\eta(t)$$ $0 \le t \le T$

noise profile $g(t) = \sigma^{t/T}$

- backward process:

$$x'(\tau) = -K(x(\tau), T-\tau) + g^2(T-\tau)\partial_x \log P(x, T-\tau) + g(T-\tau)\eta(\tau)$$

score

$\tau = T - t$

two main schemes:

- variance-expanding (VE): no drift $K(x,t) = 0$
- variance-preserving (VP) or denoising diffusion probabilistic models (DDPMs):

linear drift $K(x(t), t) = -\frac{1}{2}k(t)x(t)$

- initial data from target ensemble $x_0 \sim P_0(x_0)$
- solution: $$x(t) = x_0 f(t, 0) + \int_0^t ds\, f(t,s) g(s)\eta(s)$$
- with $f(t,s) = e^{-\frac{1}{2}\int_s^t ds'\, k(s')}$

G. Aarts, D. E.H, L. W, and K. Zhou,
Mach. Learn.: Sci. Tecnol. 6(2025)025004

Correlations evolution in forward process in DM

- moments $\mu_n(t) = \mathbb{E}[x^n(t)]$ and cumulants or connected n-point functions $\kappa_n(t)$ $\kappa_n = \mu_n - \sum_{m=2}^{n-2} \binom{n-1}{m-1} \kappa_m \mu_{n-m}$

- second moment/cumulant: (assume: first moment vanishes: $x_0 \to x_0 - \mathbb{E}_{P_0}[x_0]$)

$$\kappa_2(t) = \mu_2(t) = \mu_2(0) f^2(t,0) + \Xi(t)$$

$$\Xi(t) = \int_0^t ds \int_0^t ds'\, f(t,s) f(t,s') g(s) g(s') \mathbb{E}_\eta[\eta(s)\eta(s')] = \int_0^t ds\, f^2(t,s) g^2(s)$$

- higher-order moment and cumulants:

$$\kappa_3(t) = \mu_3(t) = \kappa_3(0) f^3(t,0)$$

$$\mu_4(t) = \mu_4(0) f^4(t,0) + 6\mu_2(0) f^2(t,0)\Xi(t) + 3\Xi^2(t)$$

$$\kappa_4(t) = \mu_4(t) - 3\mu_2^2(t)$$

$$\kappa_4(t) = \left[\mu_4(0) - 3\mu_2^2(0)\right] f^4(t,0) = \kappa_4(0) f^4(t,0)$$

variance-expanding scheme: no drift

$$f(t,0) = 1$$

higher cumulants conserved!

→ $$\boxed{\kappa_{n>2}(t) = \kappa_n(0) f^n(t,0)}$$

- in variance-expanding scheme ($f(t,0) = 1$, no drift): distribution at end of forward process as correlated as target distribution

Generating functions – simple structures

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

- proof to all orders: generating functionals $Z[J] = \mathbb{E}[e^{J(t)x(t)}]$ $\quad W[J] = \log Z[J]$

- average over both noise and target distribution $$Z_\eta[J] = \mathbb{E}_\eta[e^{J(t)x(t)}] = \frac{\int D\eta\, e^{-\frac{1}{2}\int_0^t ds\, \eta^2(s) + J(t)\left[x_0 f(t,0) + \int_0^t ds\, f(t,s) g(s)\eta(s)\right]}}{\int D\eta\, e^{-\frac{1}{2}\int_0^t ds\, \eta^2(s)}}$$

- noise average: $$Z_\eta[J] = e^{J(t)x_0 f(t,0) + \frac{1}{2}J^2(t)\Xi(t)}$$

- total average: $$Z[J] = \mathbb{E}[e^{J(t)x(t)}] = e^{\frac{1}{2}J^2(t)\Xi(t)} \int dx_0\, P_0(x_0) e^{J(t)x_0 f(t,0)}$$

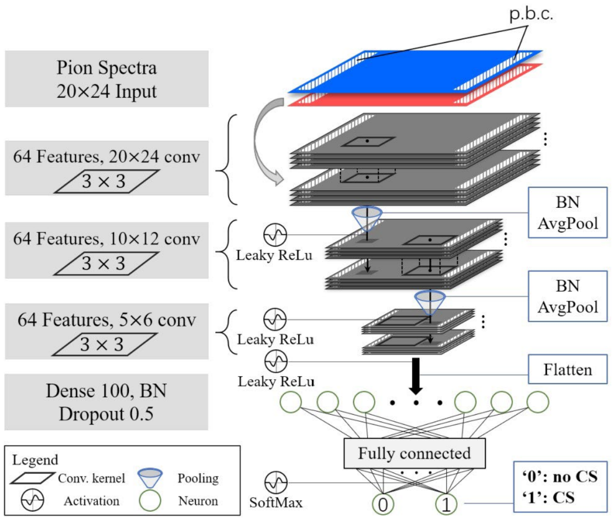
- cumulants: $$W[J] = \log Z[J] = \frac{1}{2}J^2(t)\Xi(t) + \log \int dx_0\, P_0(x_0) e^{J(t)x_0 f(t,0)}$$

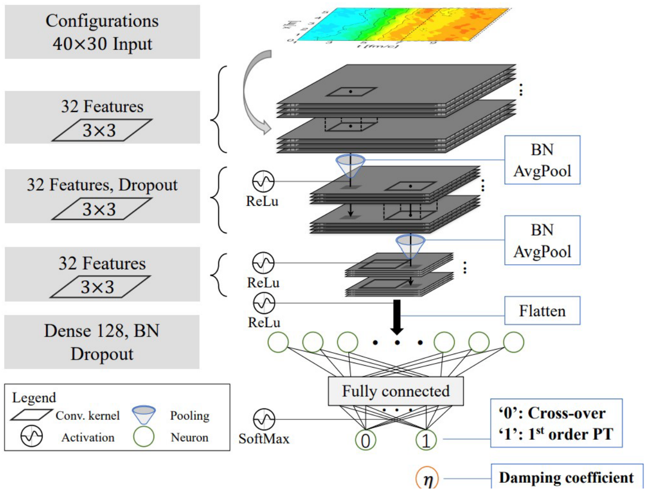
- 2nd cumulant: $$\kappa_2(t) = \frac{d^2 W[J]}{dJ(t)^2}\bigg|_{J=0} = \Xi(t) + \mathbb{E}_{P_0}[x_0^2] f^2(t,0) \quad ✓$$

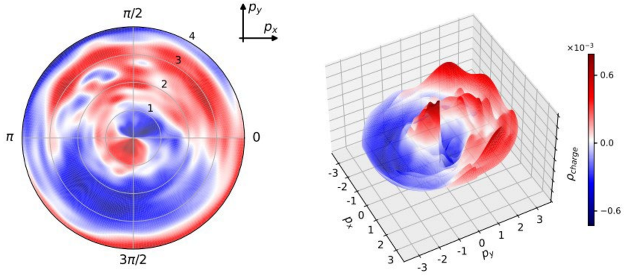
- higher-order cumulants: $$\kappa_{n>2}(t) = \frac{d^n W[J]}{dJ(t)^n}\bigg|_{J=0} = \frac{d^n}{dJ(t)^n} \log \mathbb{E}_{P_0}[e^{J(t)x_0 f(t,0)}]\bigg|_{J=0} = \kappa_n(0) f^n(t,0) \quad ✓$$

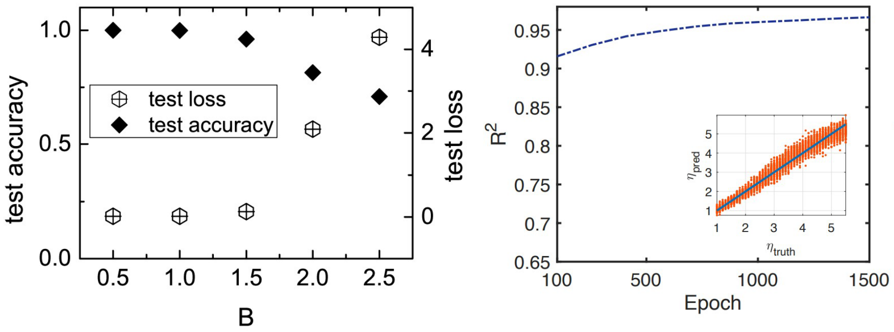
G. Aarts, D. E.H, L. W, and K. Zhou,
Mach. Learn.: Sci. Tecnol. 6(2025)025004

CNN to detect CME, and regress stochastic dynamics in HICs





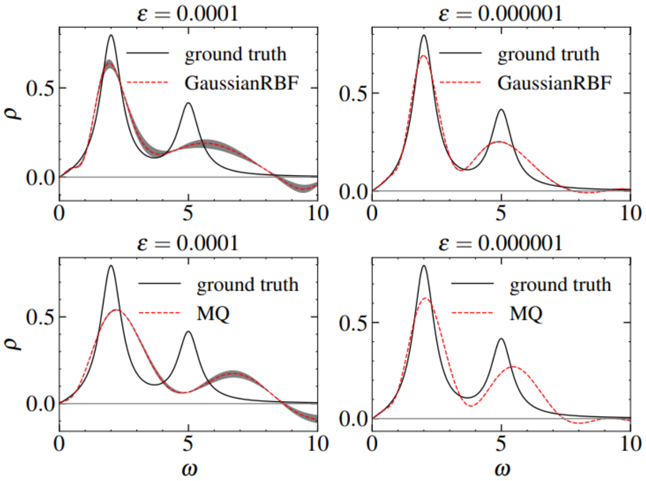


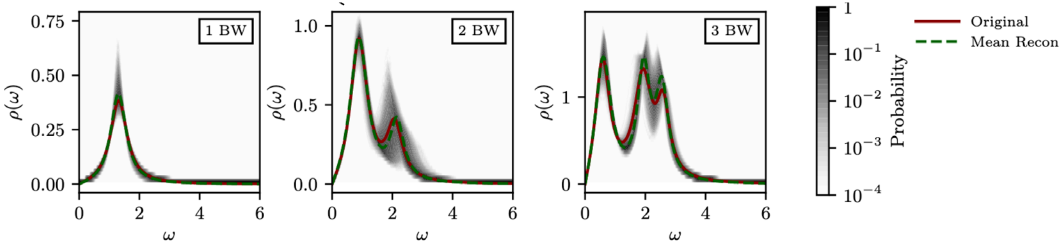


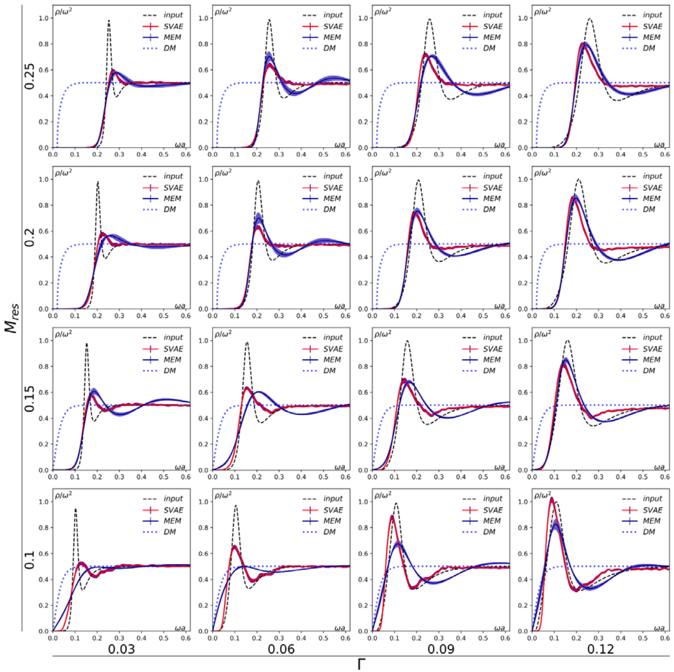
Y. Zhao, L. Wang, K. Zhou, X. Huang, ***PRC 106, L051901(Letter)***

L. Jiang, L. Wang, K. Zhou, ***PRD 103, 116023***

Spectral function reconstruction from Euclidean correlator



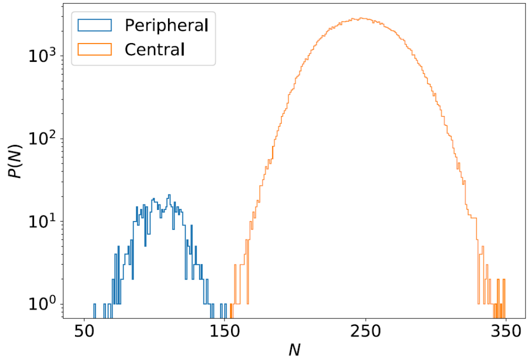


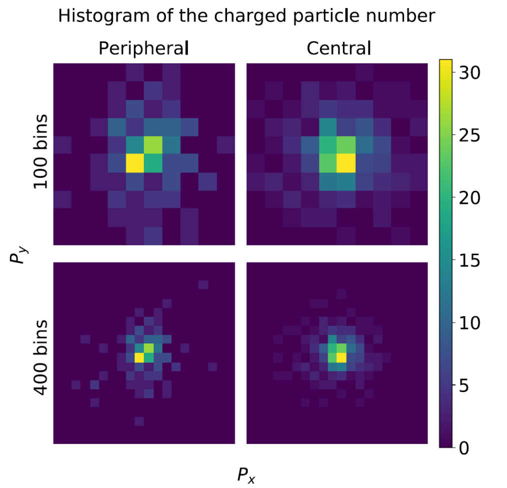


Offler+ 1912.12900; [Kades+ PRD102 (2020) 096001;
Chen+ 2110.13521; Zhou+ PRD104(2021)076011
Horak+ PRD105 (2022) 036014

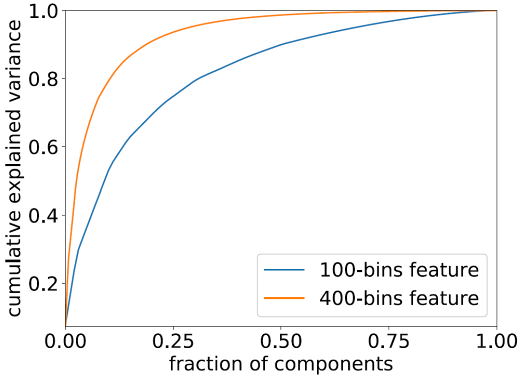
Outlier Detection for HIC

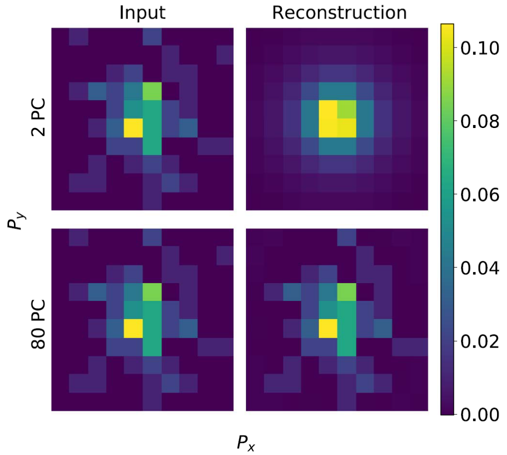
- Use **centrality misclassification** as example





- PCA to reduce dim while keep some **reconstruction** →





Phys. Scripta 96(2021)6,064003

A series **(Flow)** of invertible/bijective transformations for p(z)

Normalizing : keep the probability to be normalized →

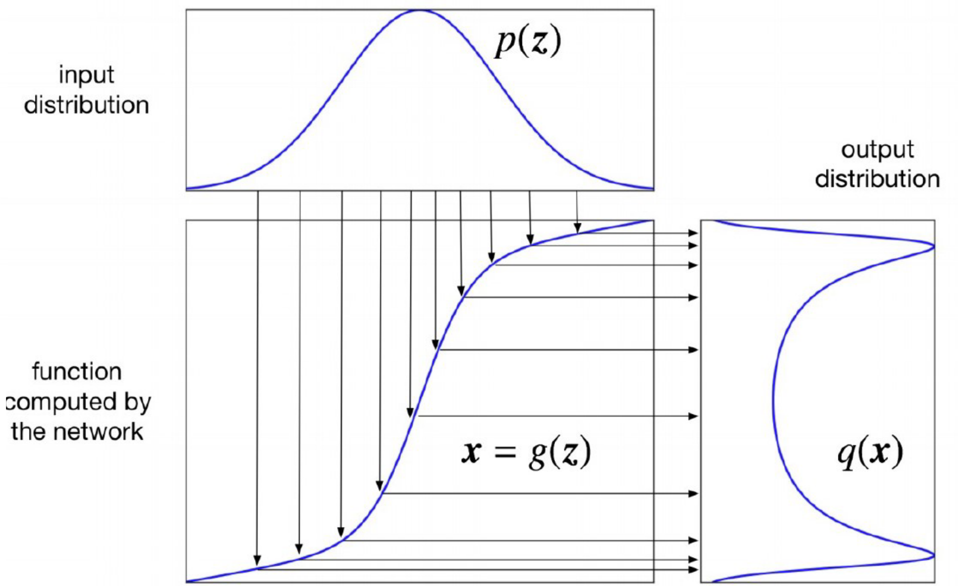
- Change of variable theorem

$$z \sim p(z)$$

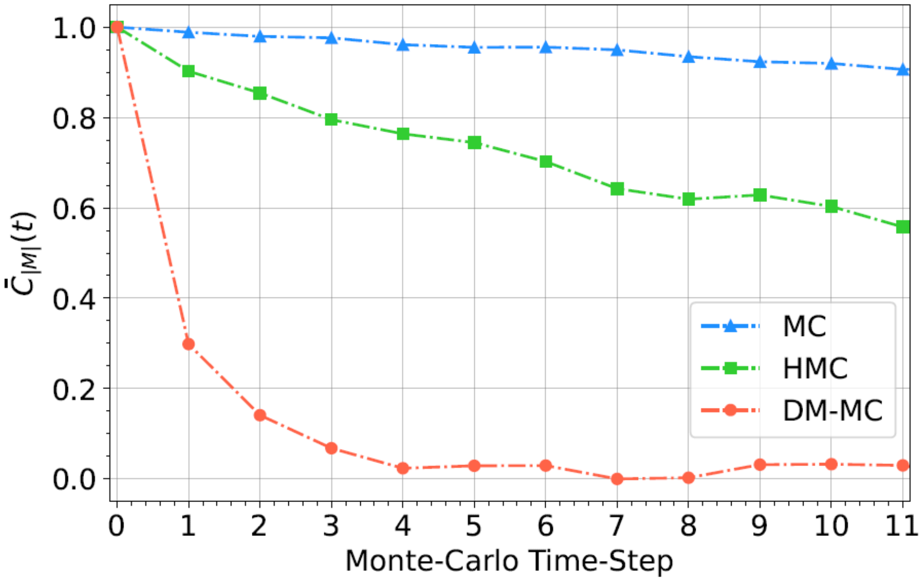
$$x = g(z)$$

$$q(x) = p(z)\left|\frac{dz}{dx}\right|$$

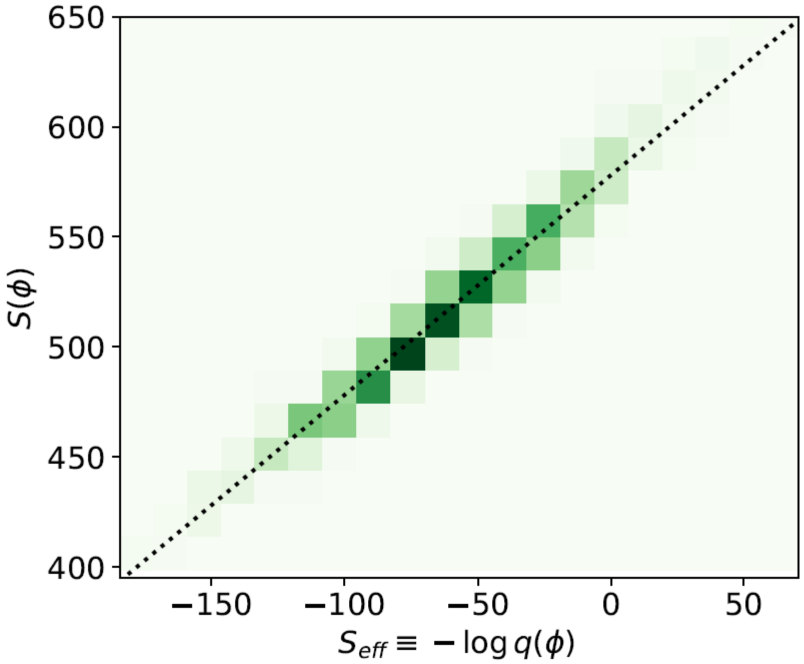
$$= p(g^{-1}(x))\left|\frac{dg^{-1}(x)}{dx}\right|$$



Autocorrelation time and finally captured effective Action



validation R2 ~0.96



$$C_O(t) = \langle (O_{t_0} - \langle O_{t_0} \rangle)(O_{t_0+t} - \langle O_{t_0+t} \rangle) \rangle = \langle O_{t_0} O_{t_0+t} \rangle - \langle O_{t_0} \rangle \langle O_{t_0+t} \rangle$$

A collision event output

- UrQMD outputs a list of final state hadrons along their momentum info
- Pointclouds: ideal representation

- Consider Au-Au 10 AGeV, impact parameter b=1 fm
  - An event= 1084 X 32
  - Empty rows=0,0,0,0,...
  - p_x, p_y, p_z, One hot encoded PID
  - 26 hadron species, spectator nucleons, empty particles

$$\mathbf{X}^{(0)} = \{\mathbf{x}_i^{(0)}\}_{i=1}^{1084}$$

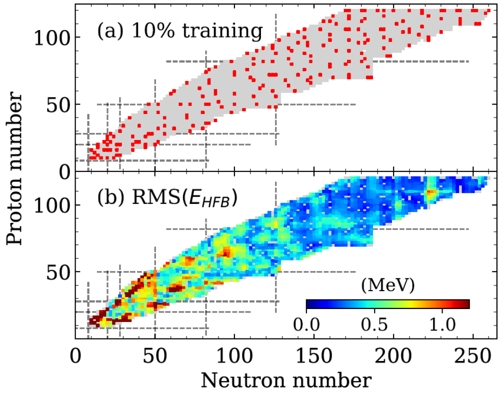
$$\mathbf{x}_i^{(0)} = \{\mathbf{p}_i^{(0)}, \mathrm{ID}_i^{(0)}\},$$

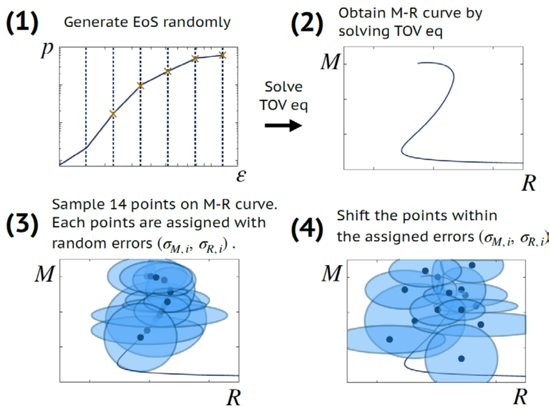
$$\mathbf{p}_i^{(0)} = (p_{x_i}^{(0)}, p_{y_i}^{(0)}, p_{z_i}^{(0)})$$

M. O.K, **K. Zhou**, J. S, H. S, **arXiv:2412.10352, arXiv:2502.16330**

Regressive and Generative AI for High Energy Nuclear Physics

- Nuclear properties prediction
  - Dripline locations, atomic masses, separation energies, superheavy nuclei location…
  - ANN application since 1992, later to beta-decay etc., → BNN
  - Different ML methods, SVM, Gradient boost, BDT,…
- Interpo-/extrapolation of nuclear data, augment nuclear model
  - Nuclear masses, nuclear charge radii, alpha-decay rate,
  - Fission yield constrain, fusion cross-section estimation, isotopic cross-sectio prediction
  - Within nuclear DFT, Energy density functional (EDF) need to be adjusted to exp data – with ML
- Nuclear matter equation of state and Neutron Star properties
  - Inverse problems in heavy ion collisions and EoS extraction
  - Experimental global analysis, for QGP properties and PDF
  - Neutron Star analysis with Bayesian, DNN, Auto-diff…
  - Fast Simulation (Emulator) for HICs
- In lattice QCD
  - Inverse problem: spectral function or interaction or PDF reconstruction
  - Bayesian inference and DNN, and also auto-diff
  - Configurations Generation via Generative models



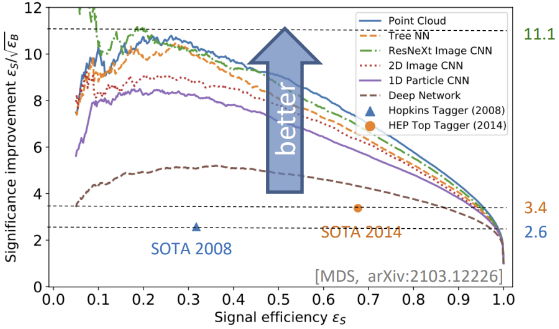


(particle phy) Supervised Learning – Regressive Tasks

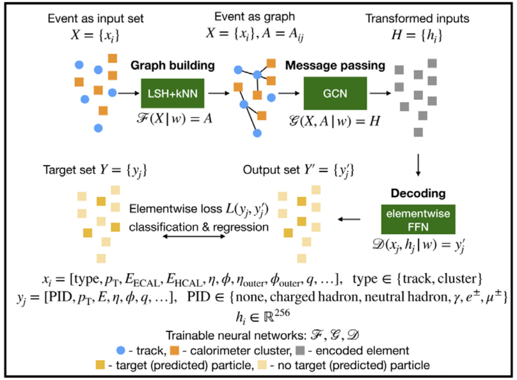
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

- Jet **Tagging**, PID, - BDT, CNN, GNN, PCN (CMS-DeepJet, ATLAS)
  - B-tagging (identifying jets originating from b-quarks)
  - Tau-jet : Lepton/photon vs. hadron separation
  - Heavy-flavor jets (specific particle decays)
  - Pion, kaon, proton identification, medium-like/vacuum-like jets

Top Tagging (2008 – 2022)



- **Reconstruction** – GNN, CNN, Self-Supervised (ML4ParticleFlow)
  - Convert raw detector signals to physical variables (4 mom, vertex)
  - Calibrating reconstructed energies in calorimeter
  - Correcting measured momenta from tracking detectors

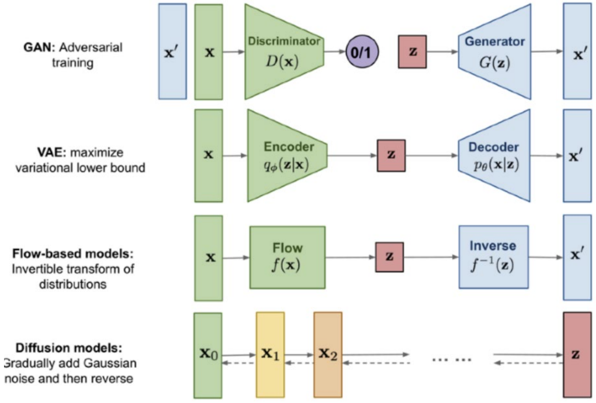


- Real-Time **Trigger** / **Filtering** system – CNN, RL, Q-learning
  - Ultra-low latency classification of collision event / signals
  - Implementing ML inference on specialized hardware (FPGAs)
  - Online distillation reducing raw data flow from Terabytes/second to manageable levels

- **Inference** – cINN, flows
  - Learn param of theory from high-d exp data – simulation based inference
  - Inverse problem solving

(particle phy) Unsupervised Learning – Generative or Detection tasks

- **Simulation** – GAN, VAE, flow, Diffusion (CaloGAN)
  - ➢ Fast simulation of collision events and detector responses.
  - ➢ Use classical simulation or collider data as input, train surrogate
  - ➢ 3D voxel image or Point Cloud
  - ➢ Replace time-consuming full GEANT4 simulations to accelerate experimental analyses.



- **Unfolding** – GNN, CNN, Self-Supervised (ML4ParticleFlow)
  - ➢ Recover theoretical-level physical distributions (e.g., transverse momentum) from detector-level data.
  - ➢ Calibrating reconstructed energies in calorimeter
  - ➢ Correct for detector effects such as resolution and efficiency.
  - ➢ NN based direct unfolding, or Generative based probabilistic unfolding

Fast Calorimeter Simulation Challenge 2022

View on GitHub

Welcome to the home of the first-ever Fast Calorimeter Simulation Challenge!

The purpose of this challenge is to spur the development and benchmarking of fast and high-fidelity calorimeter shower generation using deep learning methods. Currently, generating calorimeter showers of interacting particles (electrons, photons, pions, ...) using GEANT4 is a major computational bottleneck at the LHC, and it is forecast to overwhelm the computing budget of the LHC experiments in the near future. Therefore there is an urgent need to develop GEANT4 emulators that are both fast (computationally lightweight) and accurate. The LHC collaborations have been developing fast simulation methods for some time, and the hope of this challenge is to directly compare new deep learning approaches on common benchmarks. It is expected that participants will make use of cutting-edge techniques in generative modeling with deep learning, e.g. GANs, VAEs and normalizing flows.

This challenge is modeled after two previous, highly successful data challenges in HEP – the top tagging community challenge and the LHC Olympics 2020 anomaly detection challenge.

- **Anomaly** Detection – Classifier, PCA, AutoEncoder,
  - ➢ Detect physics phenomena beyond standard models.
  - ➢ Search for rare events, such as dark matter signals or new resonances.