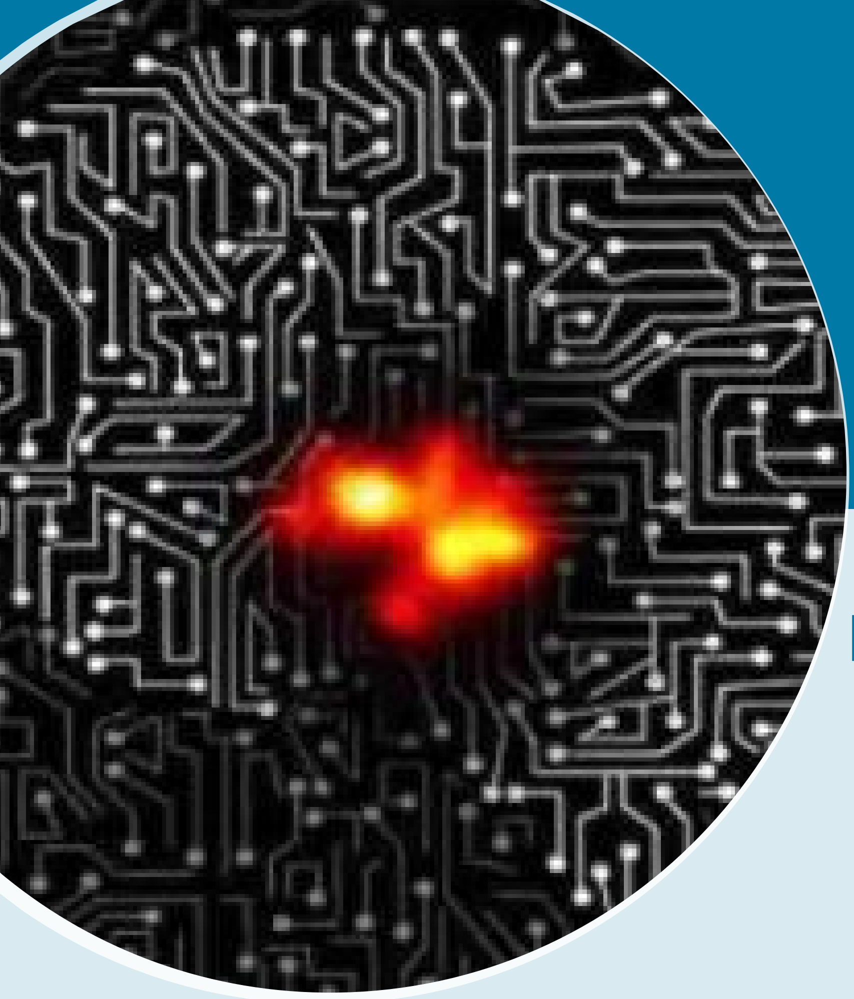


香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Machine Learning in High Energy Nuclear Physics

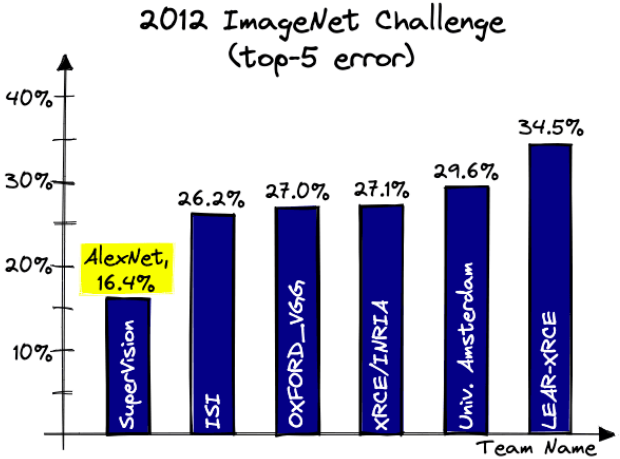
Kai Zhou (CUHK-Shenzhen)

HENPIC Seminar 2025

Overview : 13 years ago

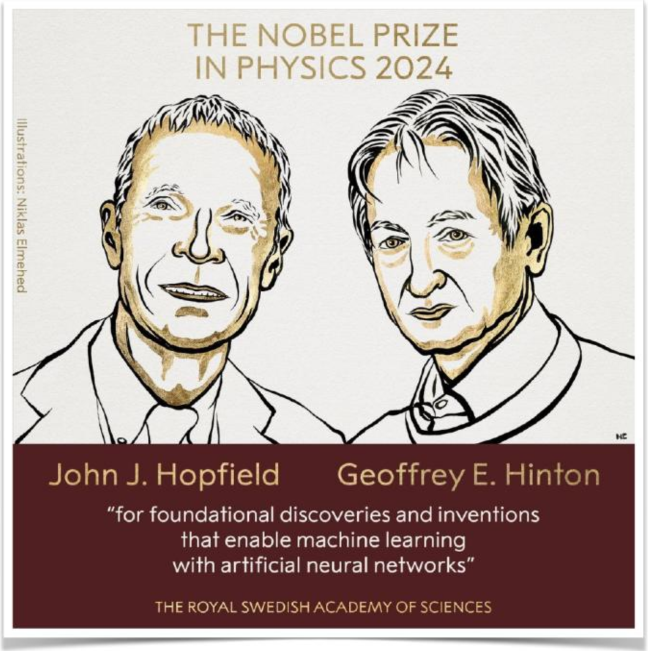
- The discovery of the Higgs boson in 2012

- AlexNet-ImageNet : birth of Deep Learning

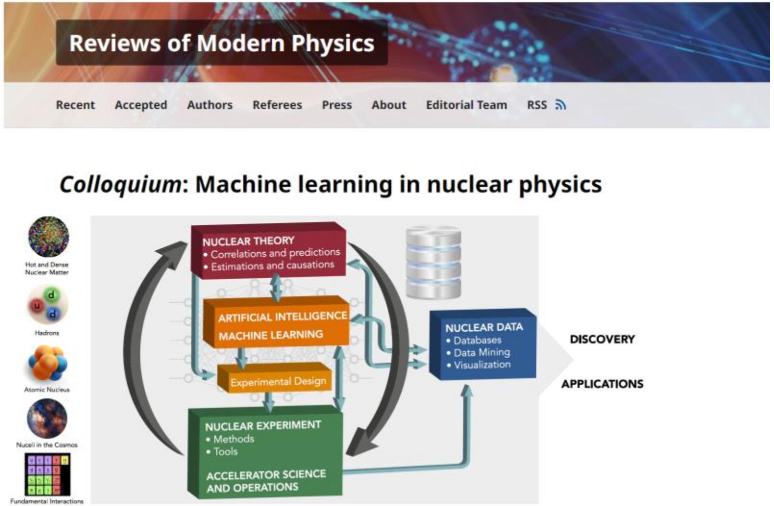


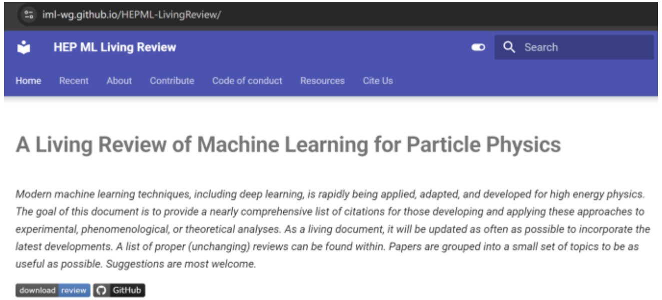
Overview : last year

- Physics Nobel Prize 2024



- ML for Nuclear Physics




iml-wg.github.io/HEPML-LivingReview/

HEP ML Living Review

Search

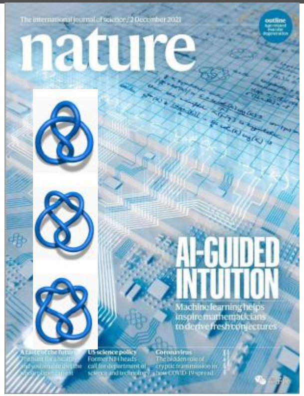
Home Recent About Contribute Code of conduct Resources Cite Us

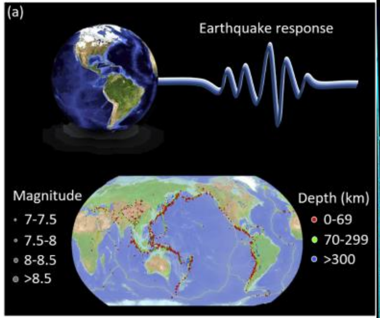
A Living Review of Machine Learning for Particle Physics

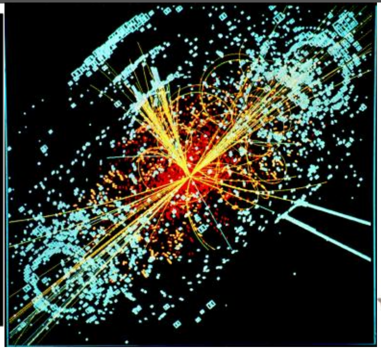
Modern machine learning techniques, including deep learning, is rapidly being applied, adapted, and developed for high energy physics. The goal of this document is to provide a nearly comprehensive list of citations for those developing and applying these approaches to experimental, phenomenological, or theoretical analyses. As a living document, it will be updated as often as possible to incorporate the latest developments. A list of proper (unchanging) reviews can be found within. Papers are grouped into a small set of topics to be as useful as possible. Suggestions are most welcome.

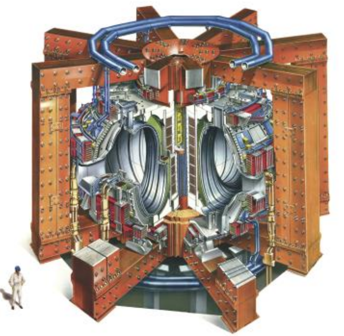
download review GitHub

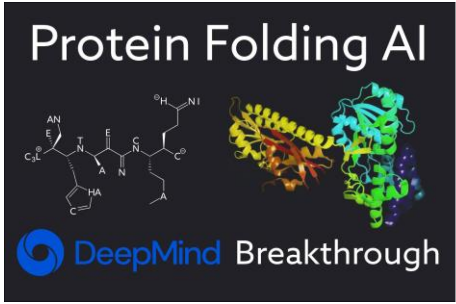
Overview : Machine- and Deep-Learning

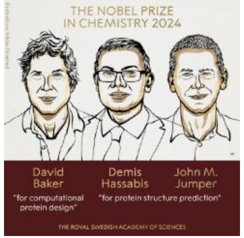








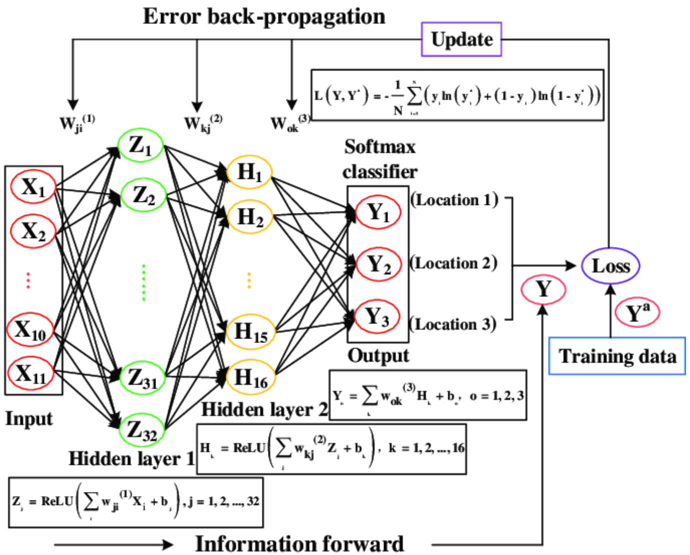




Find and Decode the mapping/representations into **Deep Neural Network**

➡ **Function approximator**

Universal approximator (Hastad et al 86 & 91)

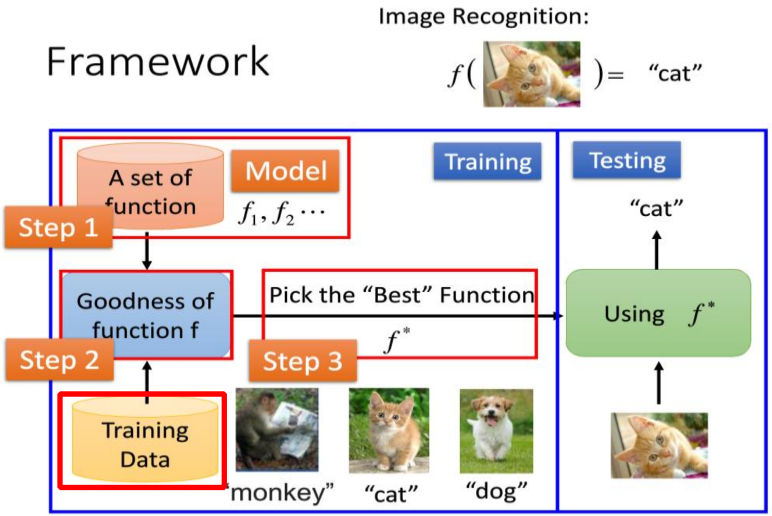


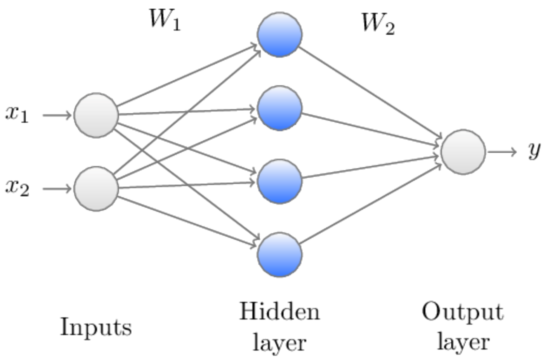
Differentiable programming

Backward Propagation

Gradient Descent Algorithm

Overview: Machine learning, Deep Neural Networks, Representation learning





Composing :

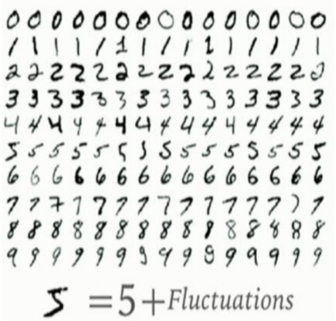
Linear affine transformation

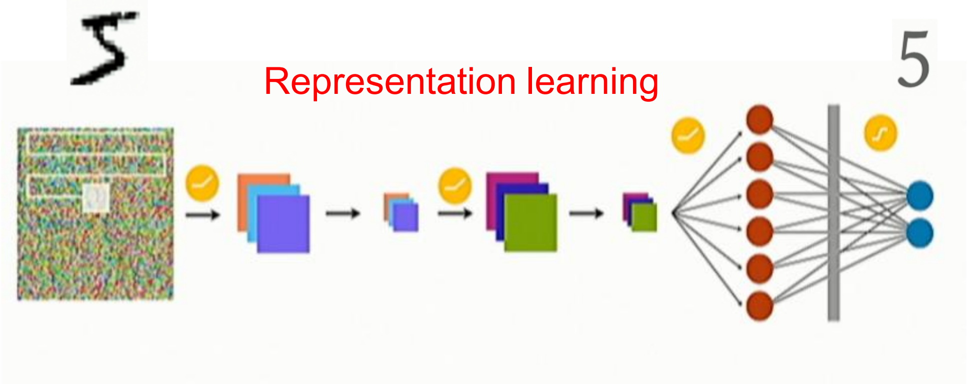
+

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

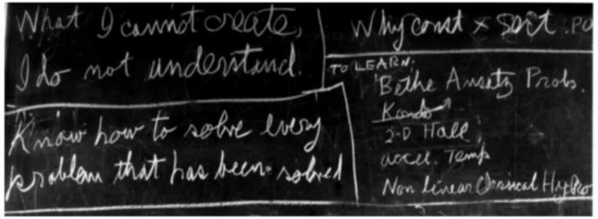
- Discriminative Learning : **prediction**

  function fitting $y = f(x)$

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

- Generative Modelling : **understand**

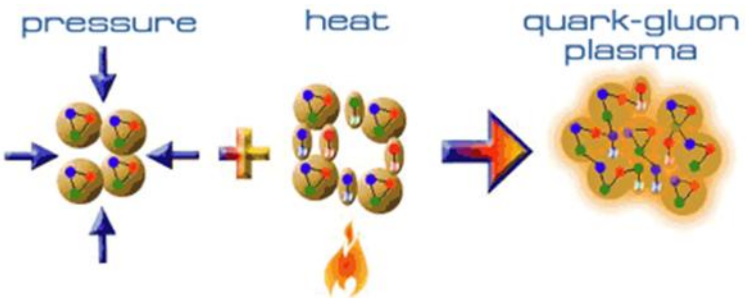
  Joint probability distribution $p_\theta(x, y) \rightarrow p(x, y)$

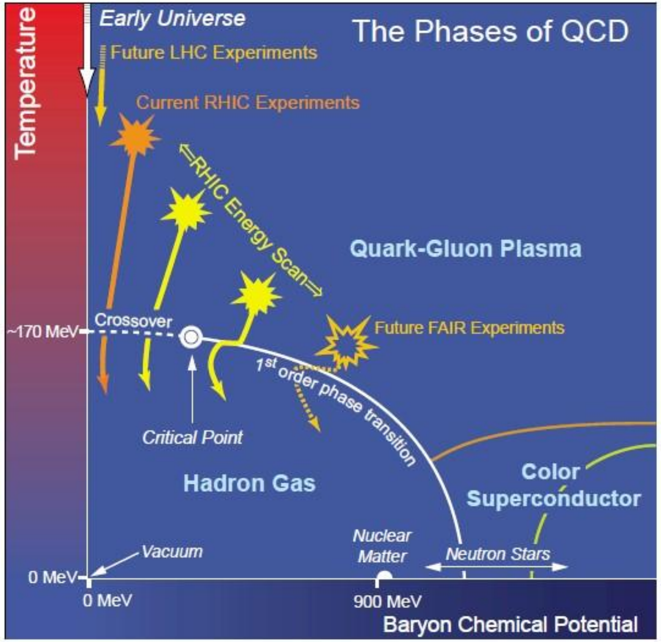


"What I can not create, I do not understand"

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

- Matter in extreme conditions reveals its **constituents** : <u>nuclear matter</u> → <u>quark matter</u>





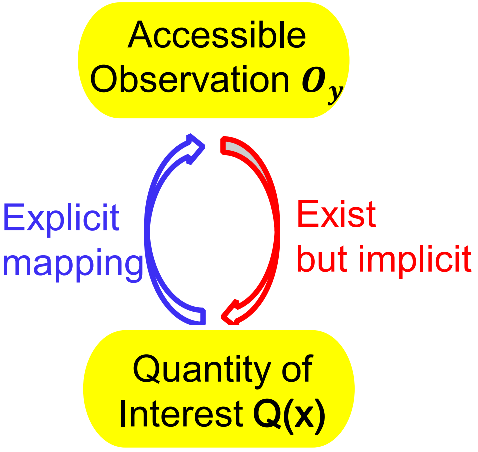
<u>To study **the most elementary particle matter**</u> :

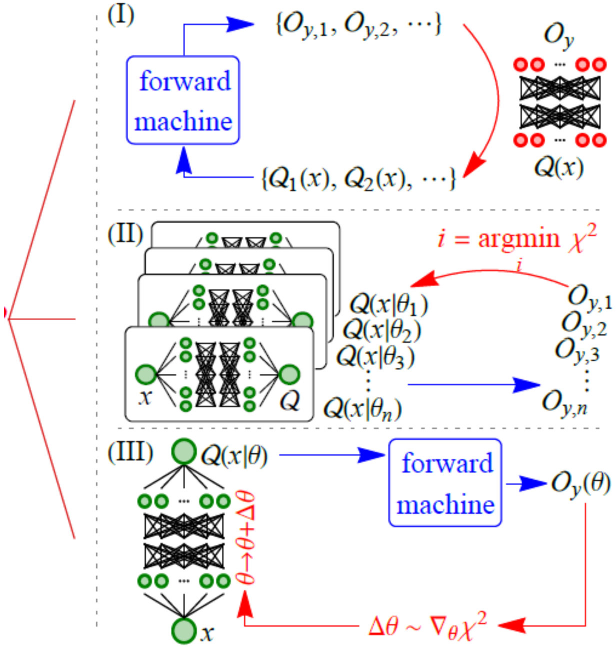
- **Nuclear Collisions** : heat & compress matter

- **Neutron Star** : dense matter, astronomy constraints

- **Lattice Field Theory / fQCD / Effective models**

Inverse Problems Solving with ML



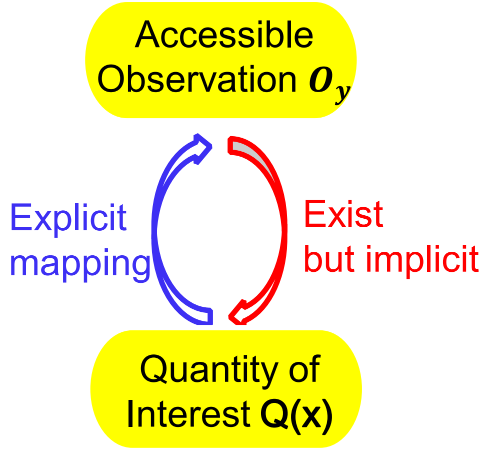


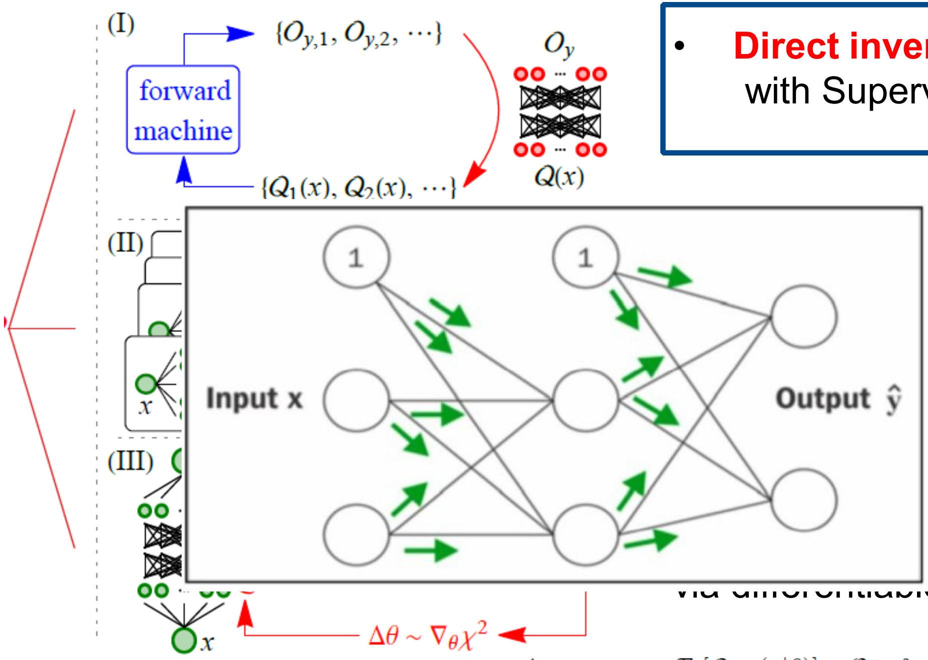
- **Direct inverse mapping capturing :** with Supervised Learning
- **Statistical approach to χ^2 fitting :** Bayesian Reconstruction for posterior or Heuristic (Generic) Algorithm to min.

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

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

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





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

approach to χ^2 fitting :

econstruction for posterior

neric) Algorithm to min.

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

Differentiation :

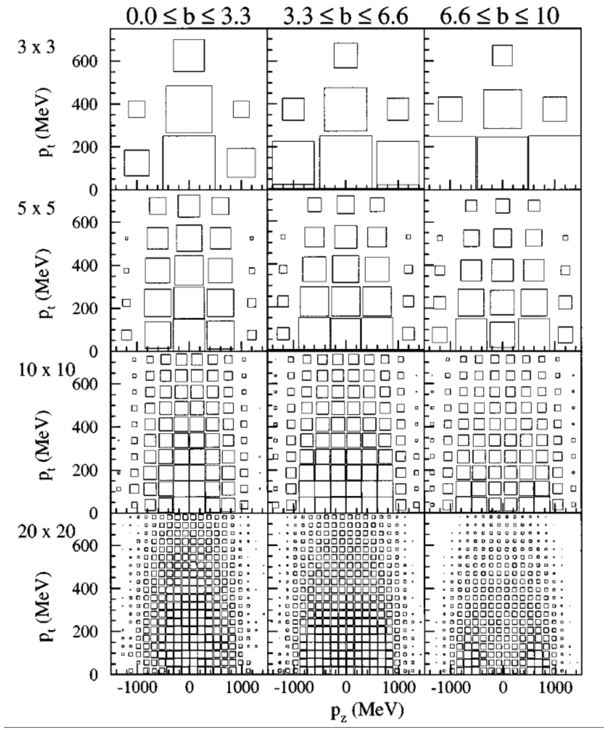
al prior into reconstruction

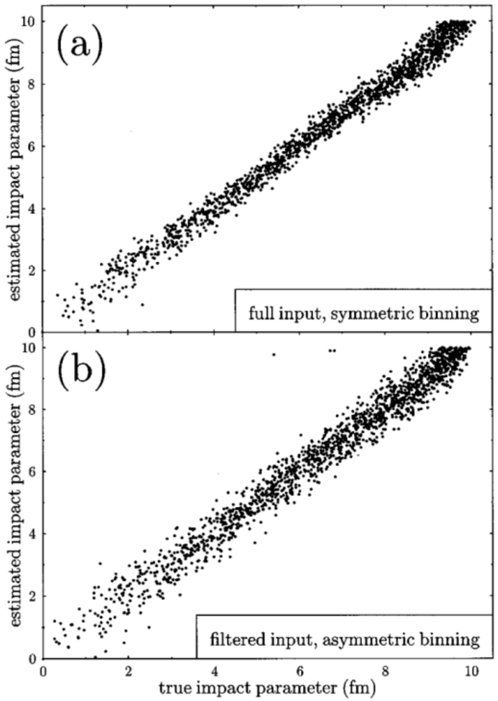
via differentiable programming strategy

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

Early attempts : impact parameter determination

Simple DNN
Trained on
QMD data
Input 5X5



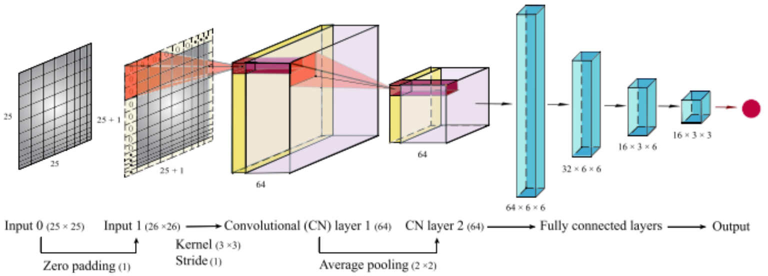


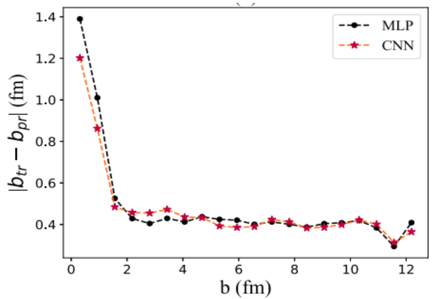
S. A. Bass, A. Bischoff, J. A. Maruhn, H. Stöcker, and W. Greiner, Phys. Rev. C 53, 2358 (1996)

Further dev for impact parameter determination

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

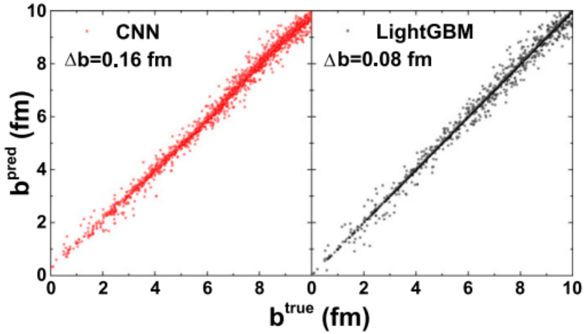
MLP and **CNN**
(on AMPT event)

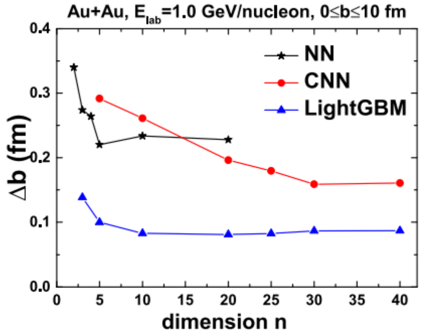




F. Li, Y. Wang, H. Lue, P. Li,
Q. Li, F. Liu, **JPG** 47, 115104 (2020)

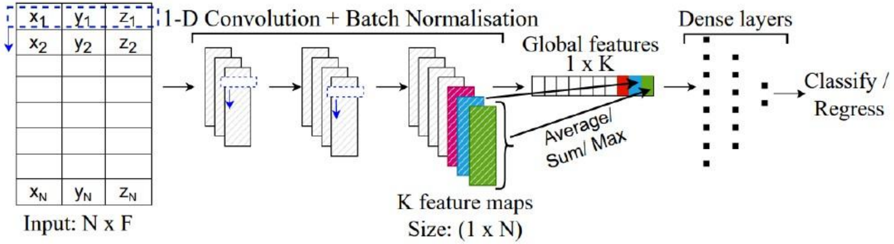
CNN and **LightGBM**
(on UrQMD event)

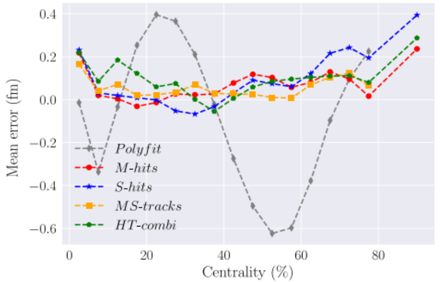




M. OK, J. S, K. Z, H. S,
PLB 811,135872 (2020)

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





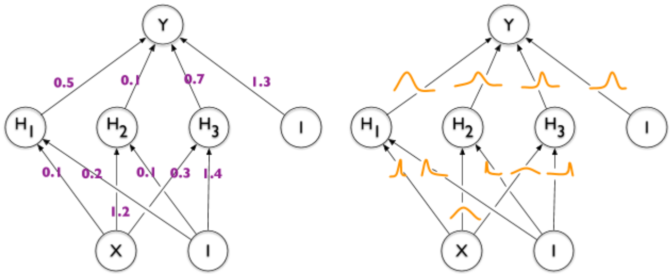
Manjunath O.K. and Kai Zhou, etc. *Phys.Lett.B* 811 (2020) 135872; JHEP10(2021)184.

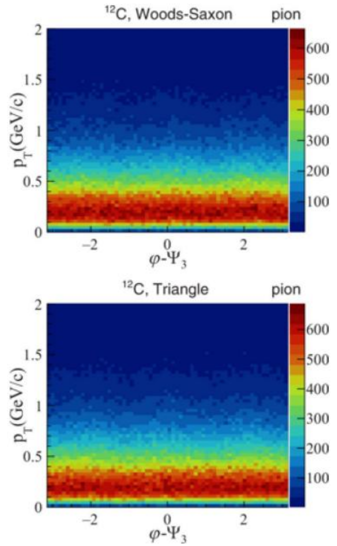
Initial clustering structure identification in HICs

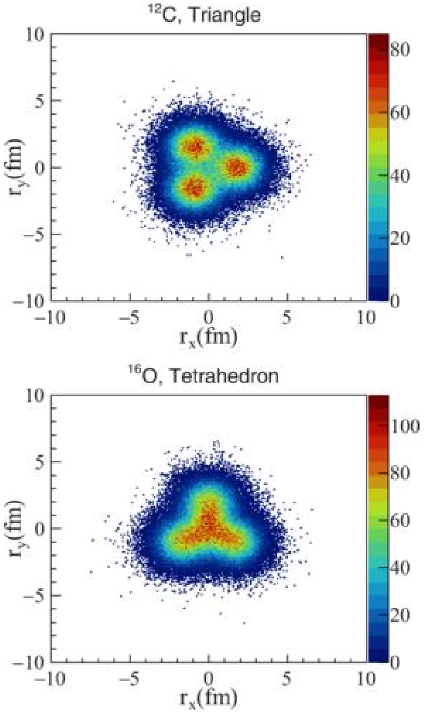
PHYSICAL REVIEW C 104, 044902 (2021)

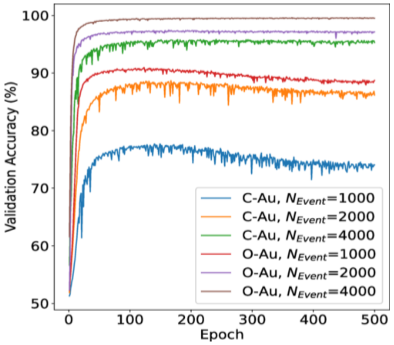
Machine-learning-based identification for initial clustering structure in relativistic heavy-ion collisions

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)

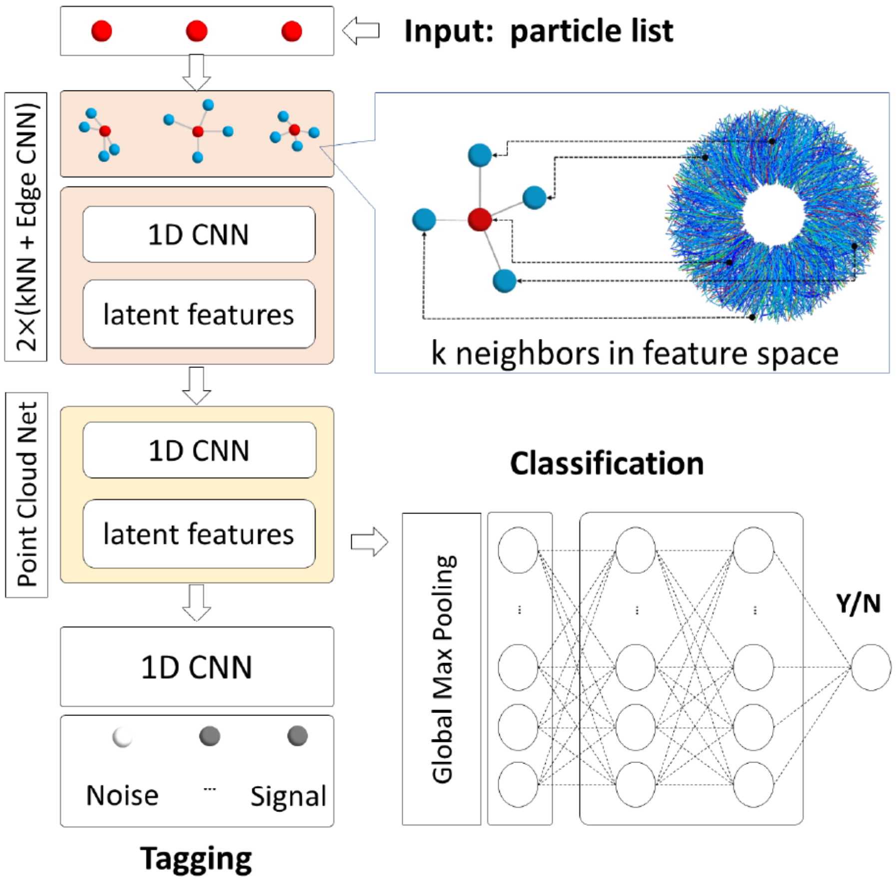
Charged pions (phi, pT) from 12C/16O + 197Au collisions at 200 GeV

Multiple event basis

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.



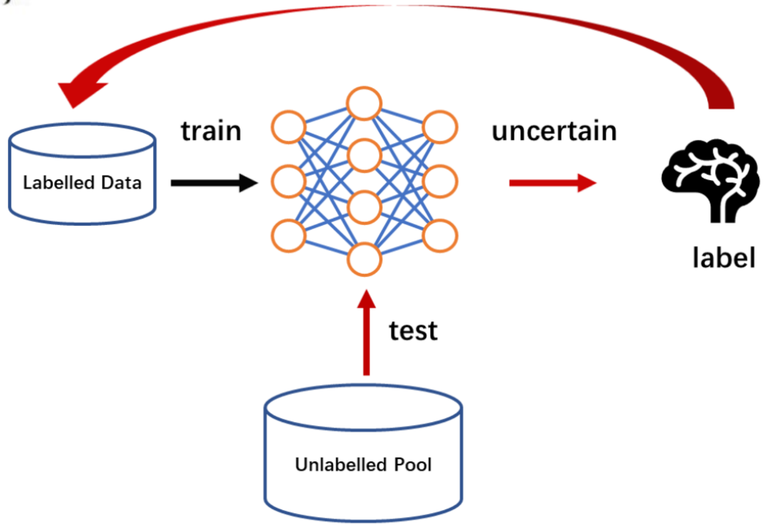
PLB *827(2022) 137001*,
Y.-G. Huang, L.-G. Pang,
X.F. Luo and X.-N. Wang.,

Active Learning map out unphysical QCD EoS

$$(\mu_{BC}, \alpha_{\mathrm{diff}}, w, \rho) \mapsto P(T, \mu_B) \mapsto \{\text{acceptable, unstable, acausal}\}.$$

Stable : $P, s, \varepsilon, n_B, \chi_2^B, \left(\frac{\partial S}{\partial T}\right)_{n_B} > 0$

Causal : $0 \leq c_s^2 \leq 1.$



Classify EoS physical relevance
based on 3D Ising model parameters

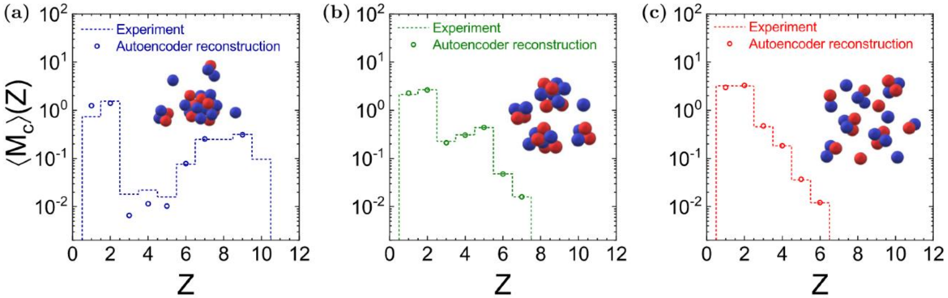
D. Mroczek, M. Hjorth-Jensen, J. Noronha-Hostler, P. Parotto, C. Ratti, and R. Vilalta, **PRC**107, 054911

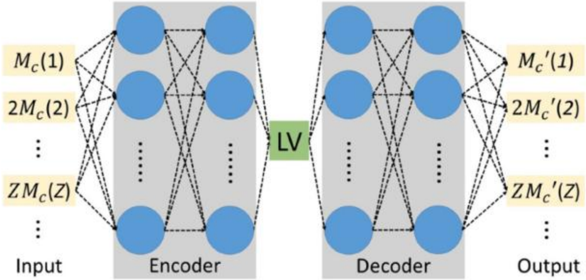
Nuclear liquid-gas phase transition recognition from autoencoder

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]

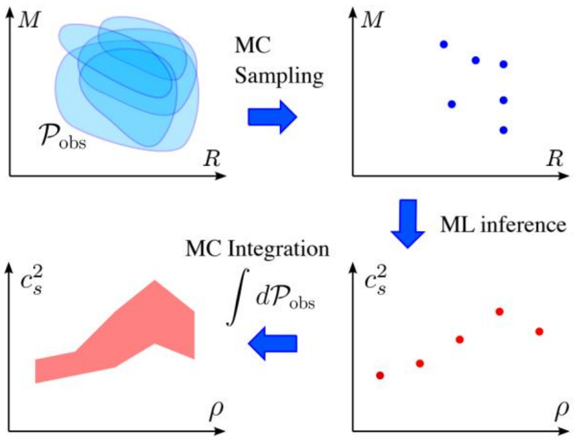


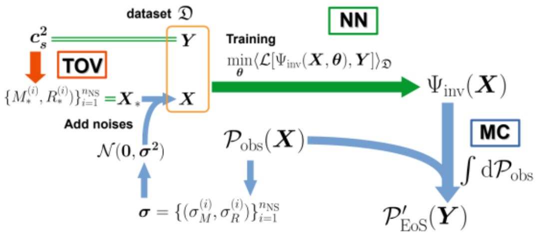


EbE charge-weighted charge multiplicity distribution of quasi-projectile as input →

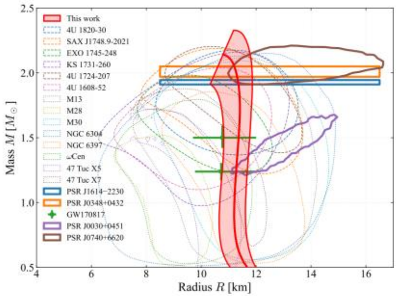
Autoencoder + confusion scheme
(on NIMROD experiment)

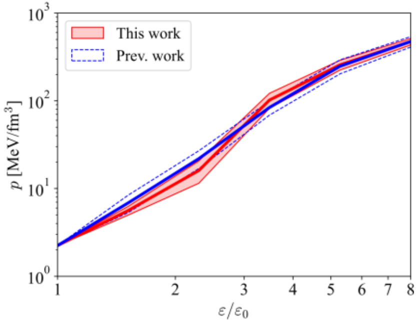
Nuclear EoS, Nuclear matter properties from NS observables





Y. Fujimoto, K. Fukushima, etc.,
PRD98(2018)023019
PRD101(2020)054016;
PRD110 (2024)034035;



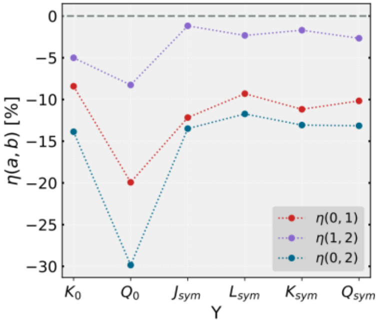


From neutron star observations to nuclear matter properties: A machine learning approach

Valéria Carvalho*, Márcio Ferreira†, and Constança Providência‡

Show more

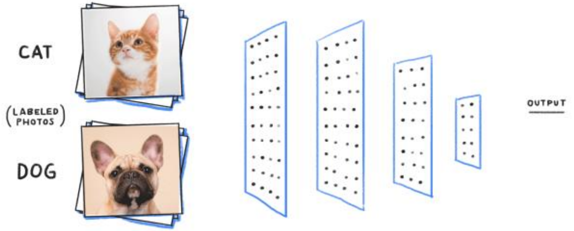
Phys. Rev. D **109**, 123038 – **Published 26 June, 2024**

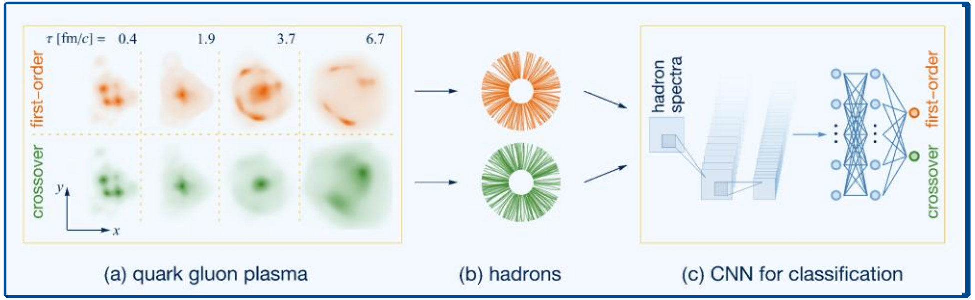


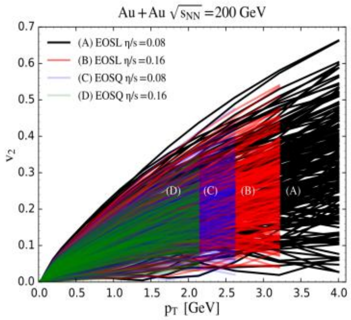
Direct inverse mapping with CNN for identifying QCD transition

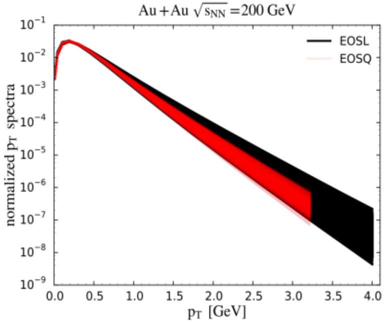
Data-driven
Inverse Mapping

Physics Simulation
provide the Prior







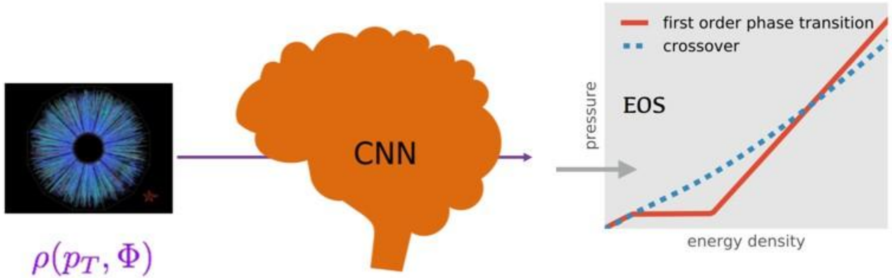


- Conventional obs. hard to distinguish
- Strongly influence from initial fluctuations and other uncertainties
- CNN : 95% event-by-event accuracy!
- Robust to initial conditions, eta/s

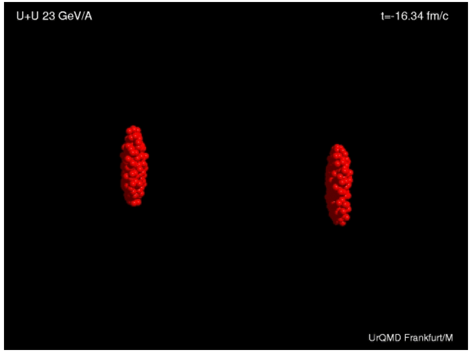
Conclusion : Information of early dynamics can **survive** to the end of hydrodynamics and encoded within the final state raw spectra, immune to evolution's uncertainties, **with deep CNN we can decode it back**.

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

Into more realistic situations

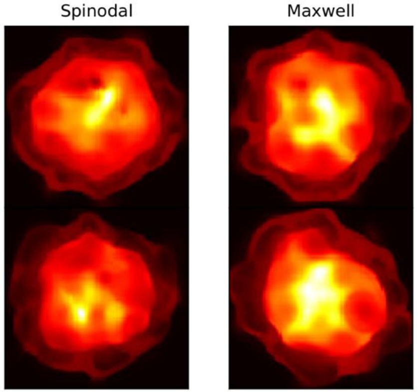


Hadronic cascade
(UrQMD considered)



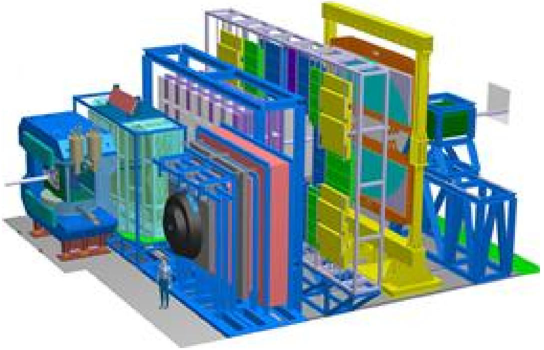
Eur. Phys. J. C 80 (2020) no.6,516

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

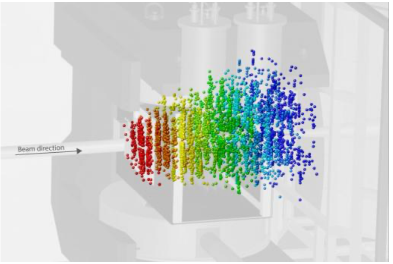
Point Cloud Network for Physics online analysis for HICs

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

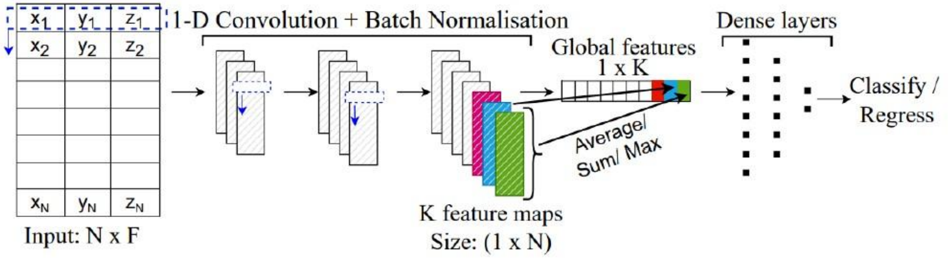
- Experimental data has inherent **point cloud structure**
  - **collection of particles** as *2D array* :

x_1	y_1	z_1
x_2	y_2	z_2
.	.	.
.	.	.
.	.	.
x_n	y_n	z_n

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

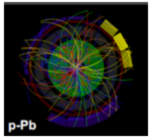


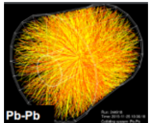
- PointNet based models learn directly from point clouds.
  - respects the **order invariance** of point clouds
  - direct processing of experimental data from detector ⇒ **ideal online analysis algorithm**
  - optimal for higher dimensional data

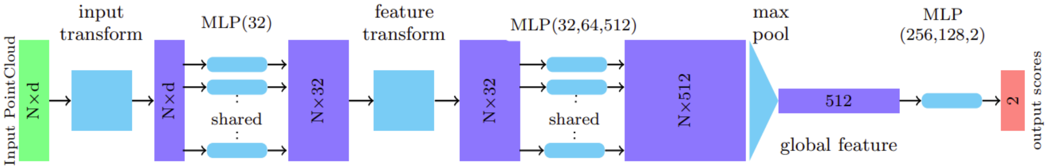


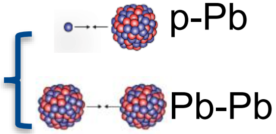
Manjunath O.K. and Kai Zhou, etc. *Phys.Lett.B* 811 (2020) 135872; JHEP10(2021)184.

- **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**

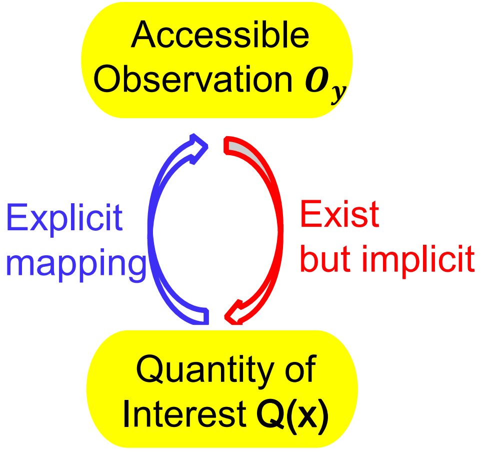


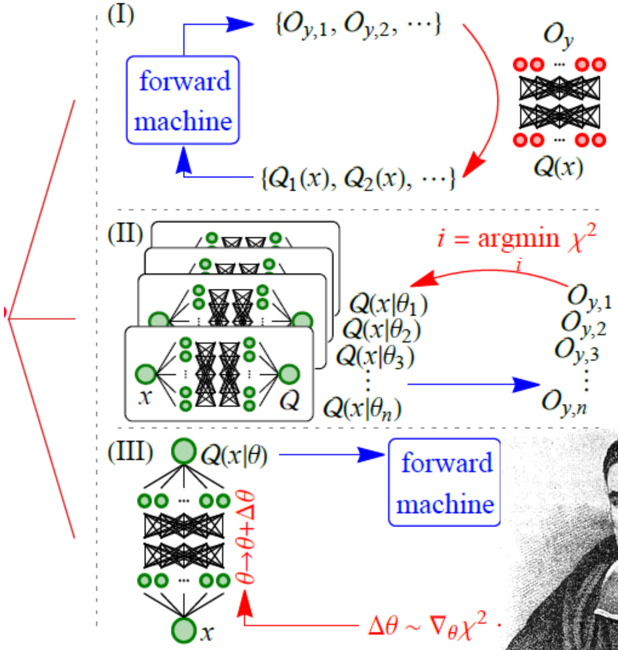






S.Guo, H. Wang, K. Zhou, G. Ma, **Phy.Rev.C** 2024

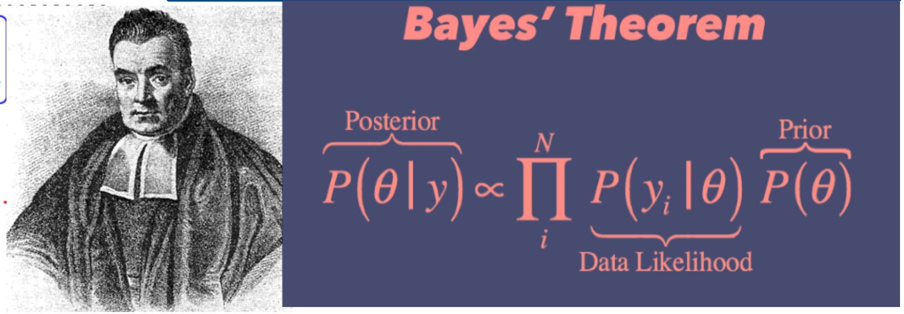




- **Direct inverse mapping capturing :** with Supervised Learning

- **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)

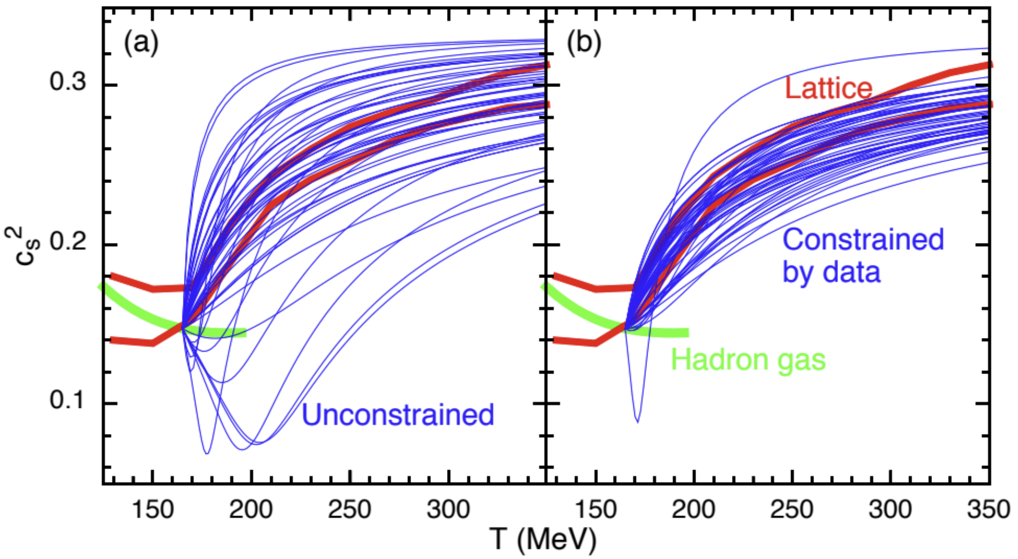
$$c_s^2(\epsilon) = c_s^2(\epsilon_h) + \left(\frac{1}{3} - c_s^2(\epsilon_h)\right)\frac{X_0 x + x^2}{X_0 x + x^2 + X'^2},$$

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

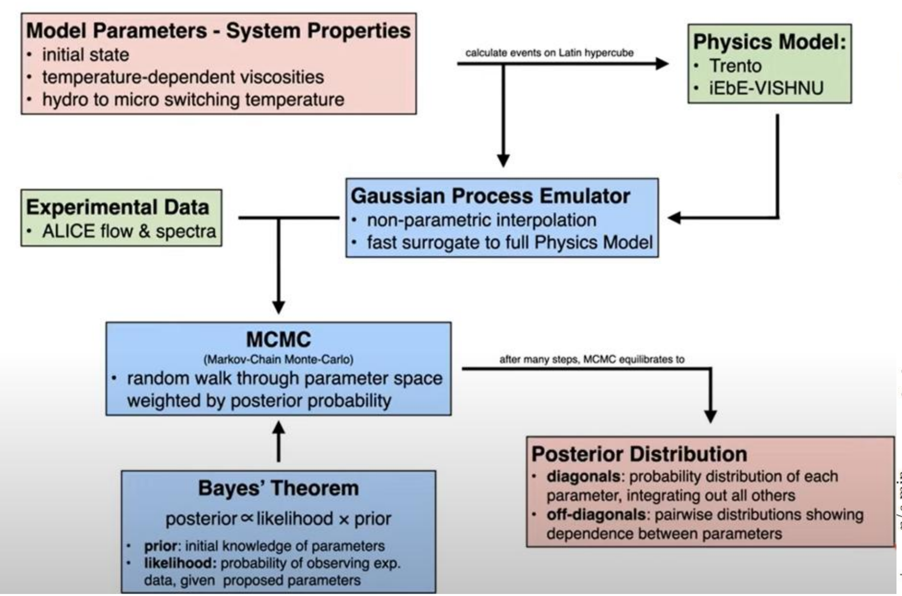
$$P(D|\theta) = \Pi_i \exp(-(z_i(\theta) - z_{i,\mathrm{exp}})^2/2),$$

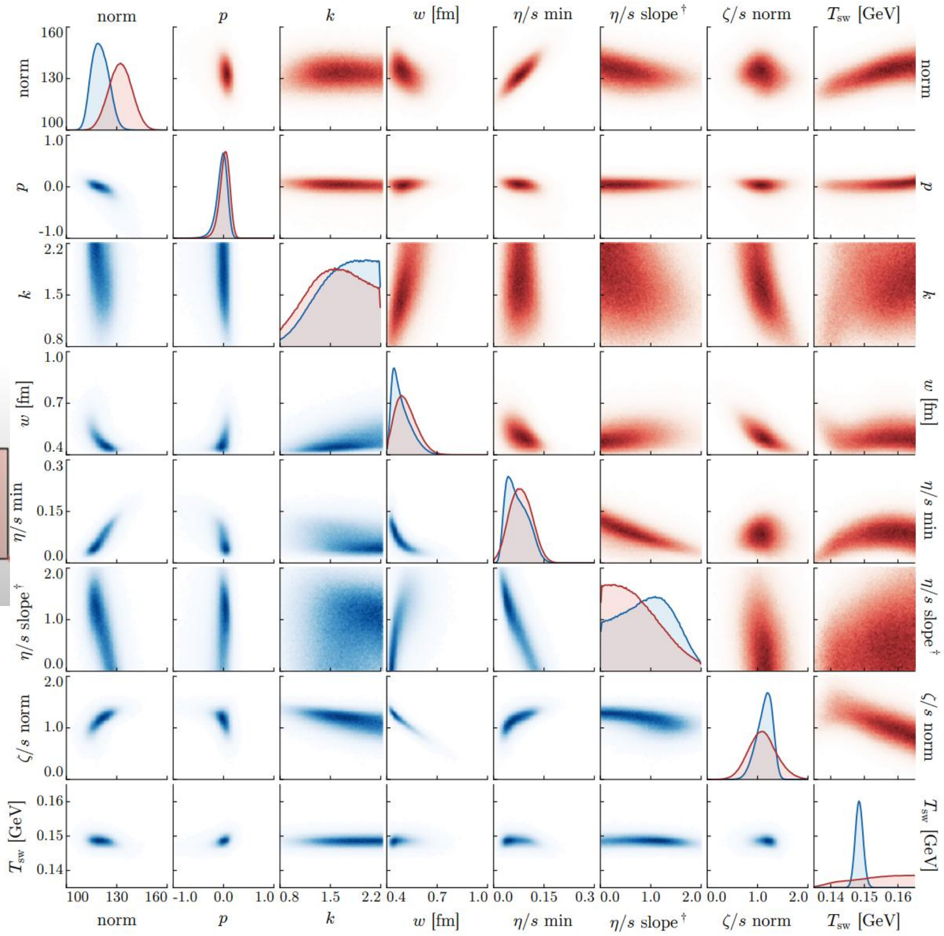
14 model parameters

speed of sound squared slightly
<u>softer</u> than lattice EoS
<u>But significantly overlap</u>



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 (Statistical) global fit on HICs

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

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	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +non-diag. $\Sigma^{\text{syst}}_{\text{expt}}$ + Σ_{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	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +non-diag. $\Sigma^{\text{syst}}_{\text{expt}}$
JETSCAPE [15, 53]	Trento+f.s.	DNMR	Grad, Chapman–Enskog, P.T.B.	Au–Au @ 0.2 TeV and Pb–Pb @ 2.76 TeV	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +diag. $\Sigma^{\text{syst}}_{\text{expt}}$
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	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +non-diag. $\Sigma^{\text{syst}}_{\text{expt}}$
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	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +non-diag. $\Sigma^{\text{syst}}_{\text{expt}}$ + Σ_{extra}
Liyanage *et al* [35]	Trento + anisotropic hydro parameters	Viscous anisotropic hydro	P.T.M.A.	Pb–Pb @ 2.76 TeV	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +diag. $\Sigma^{\text{syst}}_{\text{expt}}$
Heffernan *et al* [24, 26]	IP-Glasma	DNMR	Grad, Chapman–Enskog	Pb–Pb @ 2.76 TeV; added event-plane correlations	Σ_{emul} +diag. $\Sigma^{\text{stat}}_{\text{expt}}$ +diag. $\Sigma^{\text{syst}}_{\text{expt}}$

Bayesian Reconstruction for dense matter EoS from HICs

- Hadronic cascade dominant
- **UrQMD model** adapted to any density dependent EoS →

  via density dependent potential

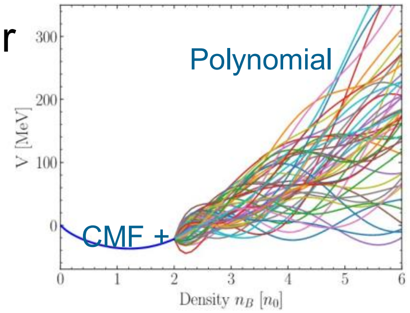
  Eur.Phys.J.C82(2022)5,417

$$\dot{\mathbf{r}}_i = \frac{\partial \mathrm{H}}{\partial \mathbf{p}_i}, \quad \dot{\mathbf{p}}_i = -\frac{\partial \mathrm{H}}{\partial \mathbf{r}_i}$$

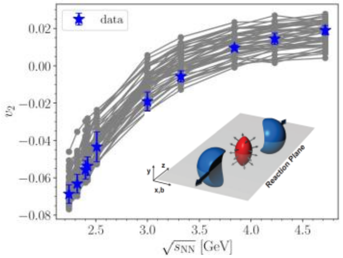
$$\dot{\mathbf{p}}_i = -\frac{\partial \mathrm{H}}{\partial \mathbf{r}_i} = -\frac{\partial \mathrm{V}}{\partial \mathbf{r}_i}$$

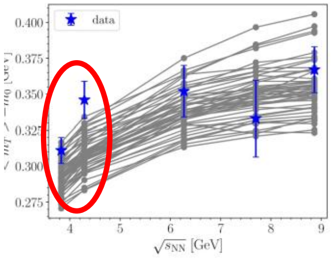
$$= -\left(\frac{\partial V_i}{\partial n_i} \cdot \frac{\partial n_i}{\partial \mathbf{r}_i}\right) - \left(\sum_{j \neq i} \frac{\partial V_j}{\partial n_j} \cdot \frac{\partial n_j}{\partial \mathbf{r}_i}\right)$$

- Prior

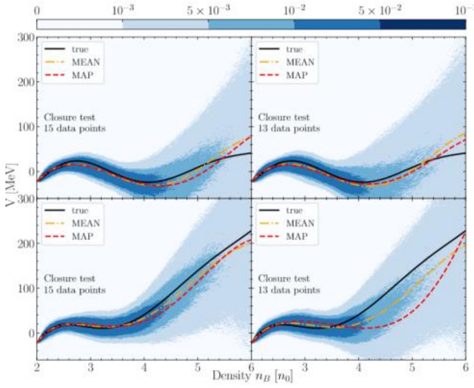


Evidence : p's v2 and transverse kinetic energy

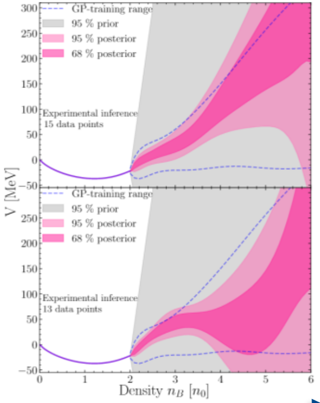


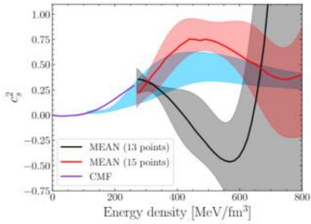


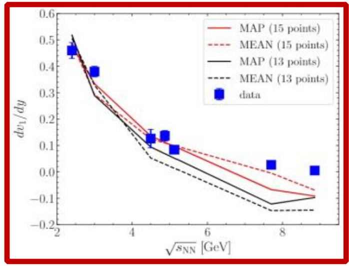
Closure Test:



Use real exp. Data:







Test the extracted EoS on <u>different observables</u>
(not used in Bayesian analysis)

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

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,‡]

arXiv:2406.07051

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

MDI: momentum dependent Interaction

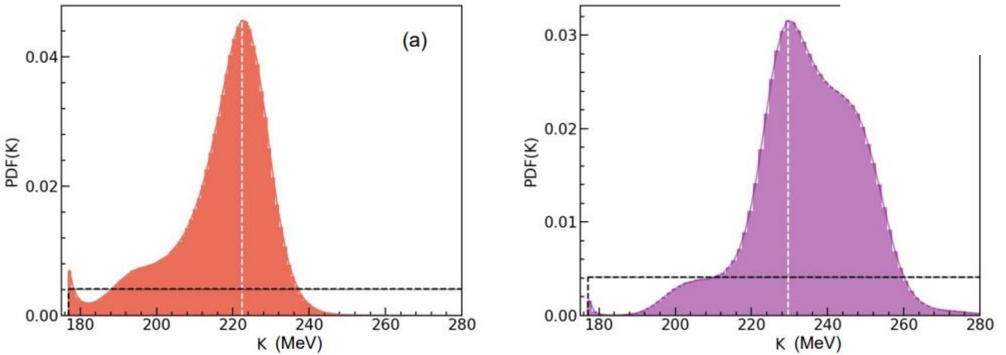


FIG. 5. Without considering the MDI: the posterior PDFs of K. Left: using only the observable $-v_2(y_0)$, right: using the two observables $-v_2(y_0)$ and $-v_2(u_{t0})$.

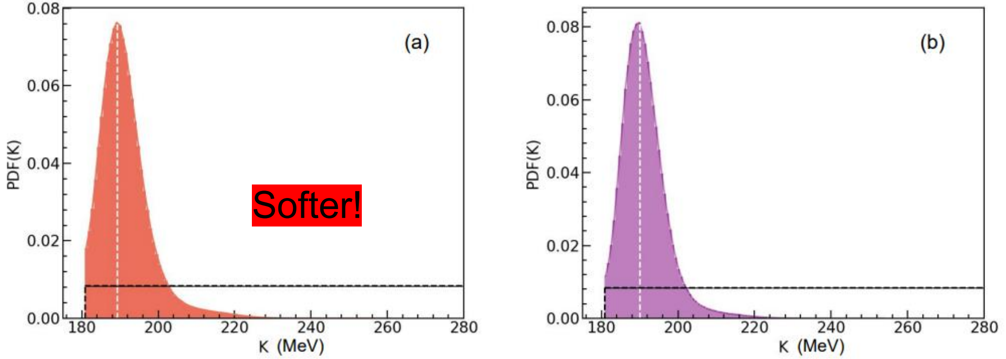


FIG. 7. Considering the MDI: the posterior PDFs of K. Left: observable only $-v_2(y_0)$, right: observables $-v_2(y_0)$ and $-v_2(u_{t0})$.

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

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

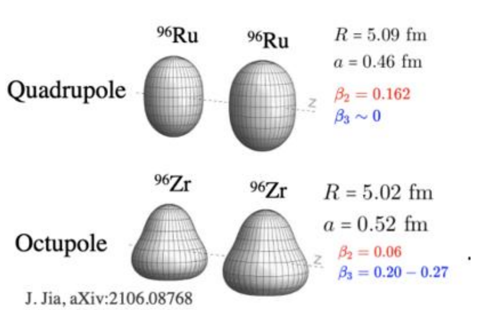
$$K = 9\rho^2\frac{\partial^2 E/A}{\partial\rho^2}\Big|_{\rho_0} = -\frac{3}{5m}\left(\frac{3\pi^2\hbar^3\rho_0}{2}\right)^{2/3} + \frac{9\beta\gamma(\gamma-1)}{\gamma+1} + \ln\left(1+t_5P_F^2\right)\frac{6t_4t_5P_F^2}{1+t_5P_F^2}, \quad (3)$$

Bayesian Imaging for Nuclear Structure in Isobar Collisions

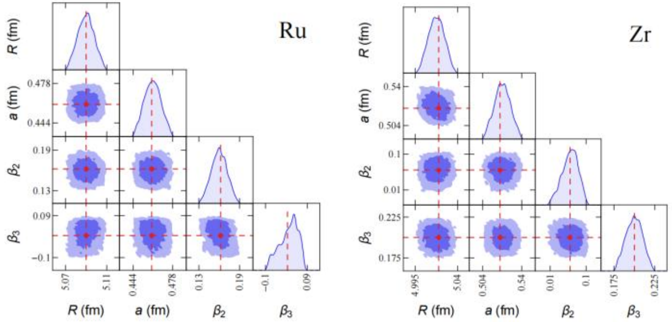
- Nuclear Structure imaging for single system ? (caveat: model dependent)

- Simultaneous inference for isobar systems with ratio?
- **Bayesian Inference:** Gaussian Process emulator + PCA dim reduction + MCMC

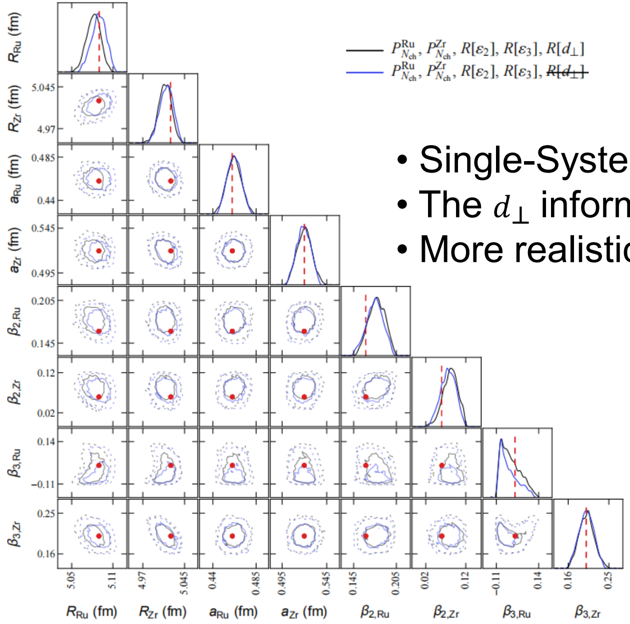
Data: MC-Glauber + Matching (linear response approximation)



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



Single system works good

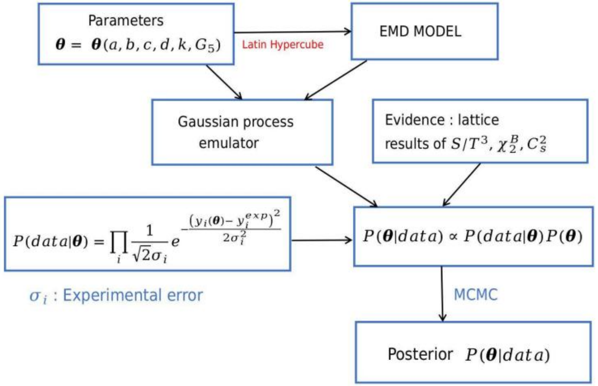


- Single-System Multiplicity makes it possible
- The $d_\perp$ information is redundant
- More realistic analysis with AMPT in progress

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

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

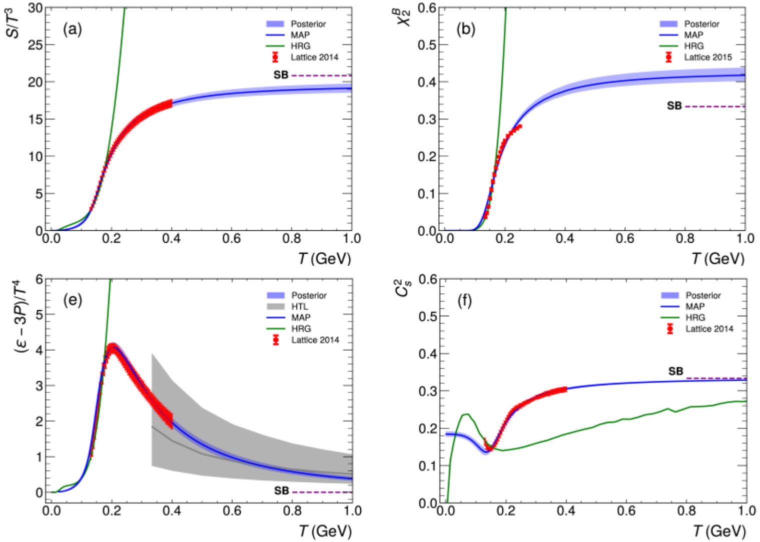
Critical Endpoint from Holographic QCD via Bayesian Inference

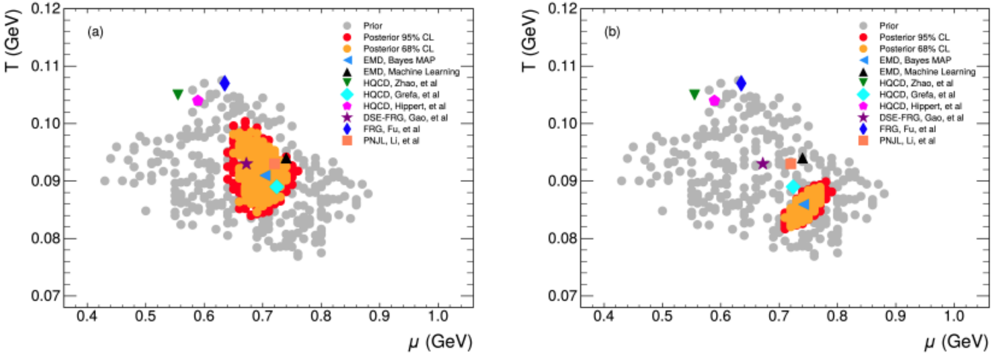


$$S_E = \frac{1}{16\pi G_5} \int d^5x \sqrt{-g} \left[R - \frac{f(\phi)}{4} F^2 - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right]$$

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

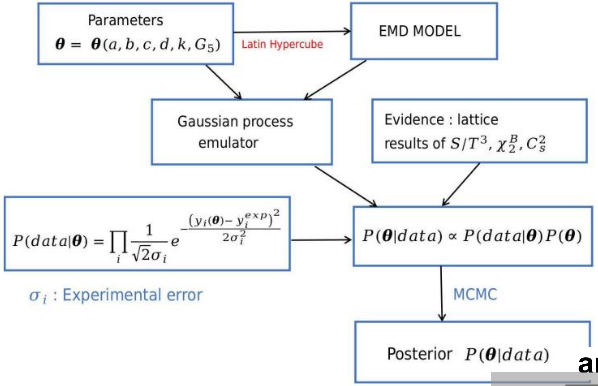
$$s = \frac{e^{3A(z_h)}}{4G_5 z_h^3}. \qquad \chi_2^B = \frac{1}{T^2} \frac{\partial \rho}{\partial \mu}.$$





L. Zhu, X. Chen, K. Zhou, H. Zhang, M, Huang,
arXiv:2501.15810

Bayesian inference of the critical endpoint from holographic QCD

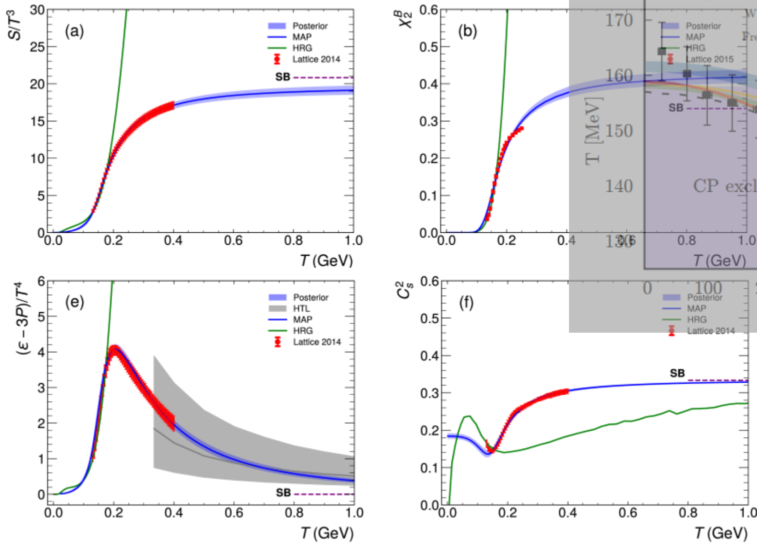


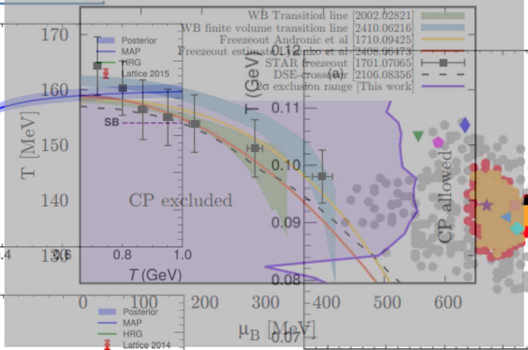
$$S_E = \frac{1}{16\pi G_5} \int d^5x \sqrt{-g} \left[R - \frac{f(\phi)}{4} F^2 - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right]$$

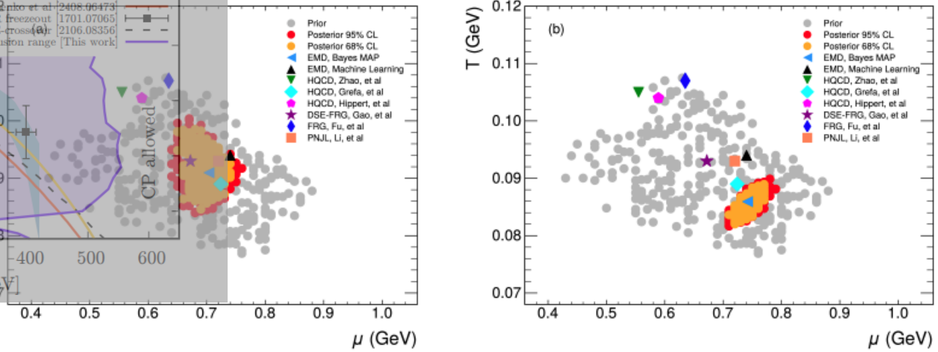
$$A(z) = d\ln(az^2 + 1) + d\ln(bz^4 + 1), \; f(z) = e^{cz^2 - A(z) + k}$$

$$s = \frac{e^{3A(z_h)}}{4G_5 z_h^3}. \qquad \chi_2^B = \frac{1}{T^2} \frac{\partial \rho}{\partial \mu}.$$

arXiv:2502.10267

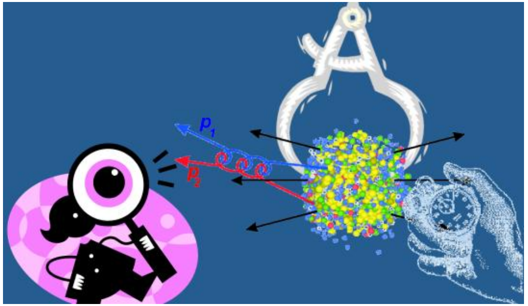






L. Zhu, X. Chen, K. Zhou, H. Zhang, M, Huang,
arXiv:2501.15810

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



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

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

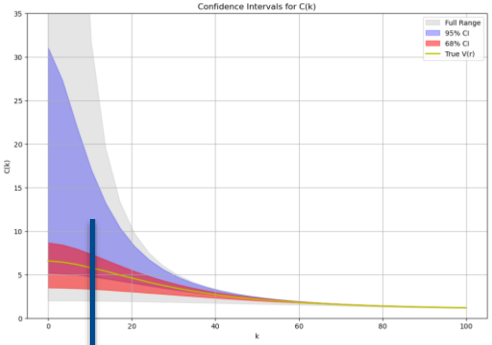
$$S(r) = \frac{1}{(4\pi r_0^2)^{3/2}} \exp\left(-\frac{r^2}{4r_0^2}\right),$$

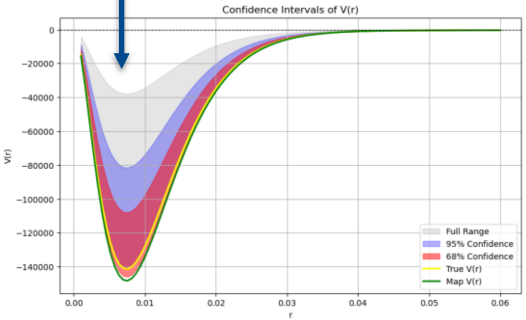
$$r_0 = 1.0 \text{ fm}$$

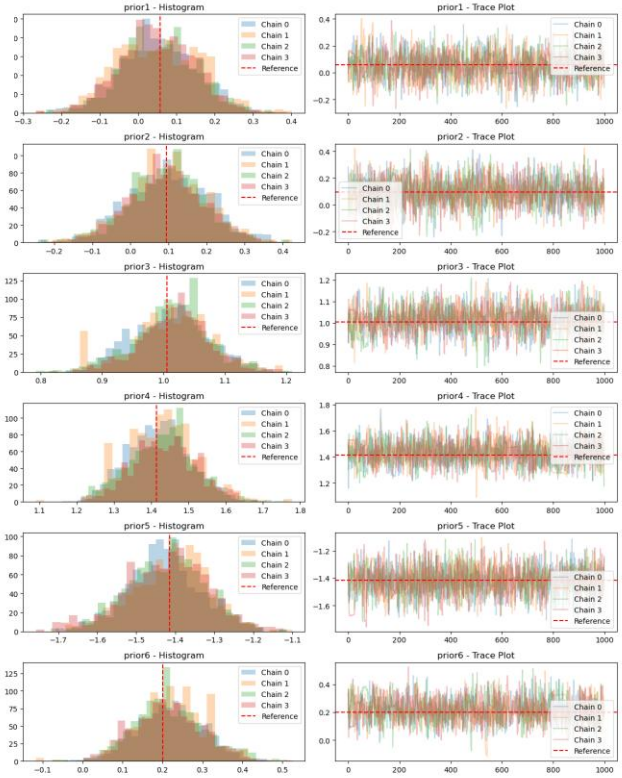
$$V(r) = \sum_{i=1,2} a_i e^{-(r/b_i)^2} + a_3 m_\pi^4 f(r, b_3) \frac{e^{-2m_\pi r}}{r^2}$$

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

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



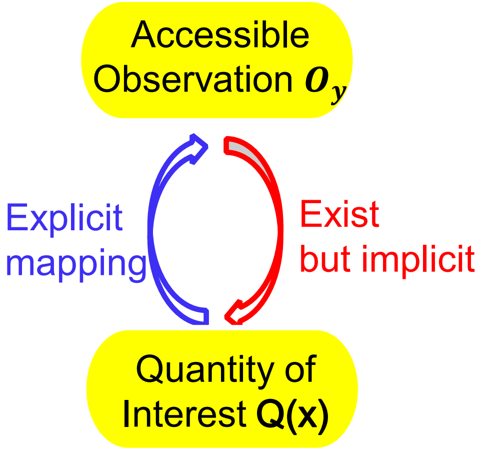


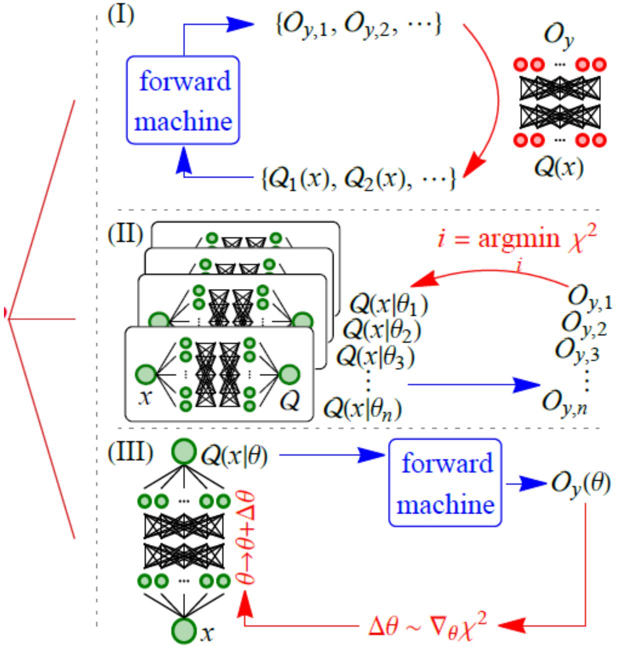


DNN emulator + PCA for correlation + PyMC

With O. L, J. Z, X. C, etc., <u>in preparation</u>

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.

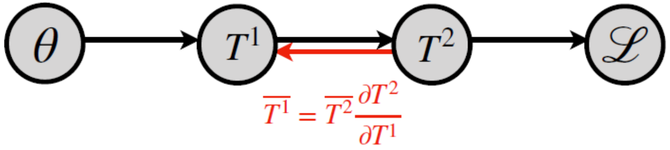
$$\chi^2 = \sum_y \left(\frac{\mathcal{F}_y[\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{\Delta \mathcal{O}_y} \right)^2$$

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

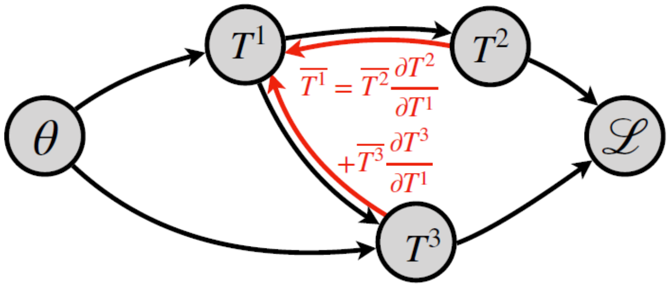
$$\frac{1}{2}\nabla_{\boldsymbol{\theta}}\chi^2 = \sum_y \frac{\mathcal{F}_y[\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{(\Delta\mathcal{O}_y)^2} \int \mathrm{d}x \frac{\delta \mathcal{F}_y[\mathcal{Q}(x)]}{\delta \mathcal{Q}(x)}\bigg|_{\mathcal{Q}(x)=\mathcal{Q}_{\rm NN}(x|\boldsymbol{\theta})}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

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

(a)



(b)



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

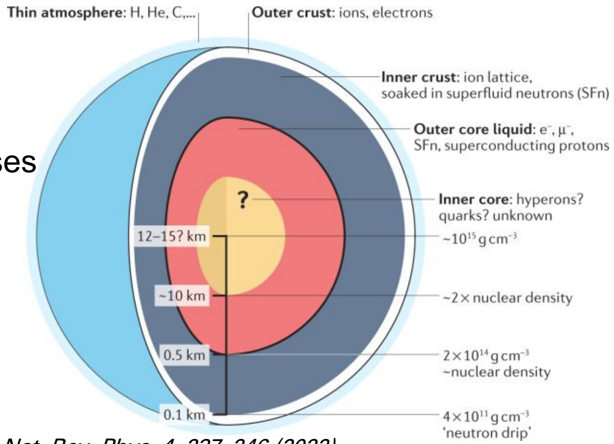
$$\boxed{\overline{T^i} = \overline{T^{i+1}}\frac{\partial T^{i+1}}{\partial T^i}}$$

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

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

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)

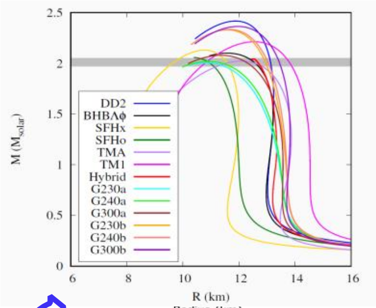
- Gravity ← → Pressure

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

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

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

- <u>Noisy/Limited NS Observables</u> to <u>EoS</u> ?

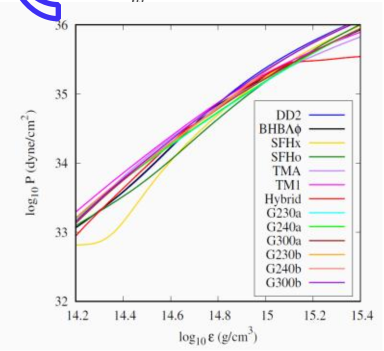


Radius (km)

TOV equations

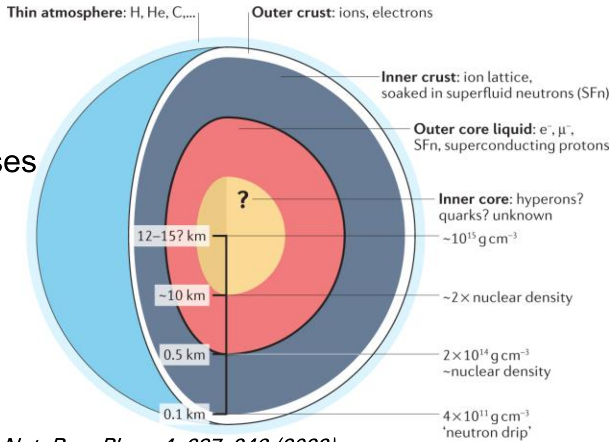
$$-\frac{dP}{dr} = \frac{[\epsilon(r) + P(r)]\,[M(r) + 4\pi r^3 P(r)]}{r[r - 2M(r)]}$$

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



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

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



Nat. Rev. Phys. 4, 237–246 (2022)

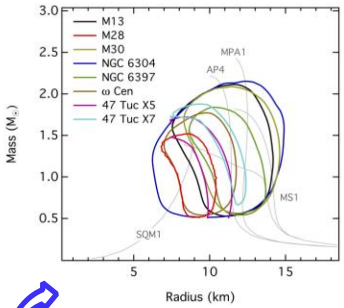
- Gravity ← → Pressure

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

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

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

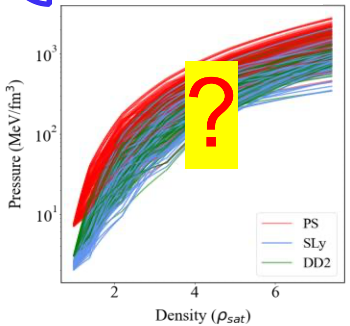
- <u>Noisy/Limited NS Observables</u> to <u>EoS</u> ?



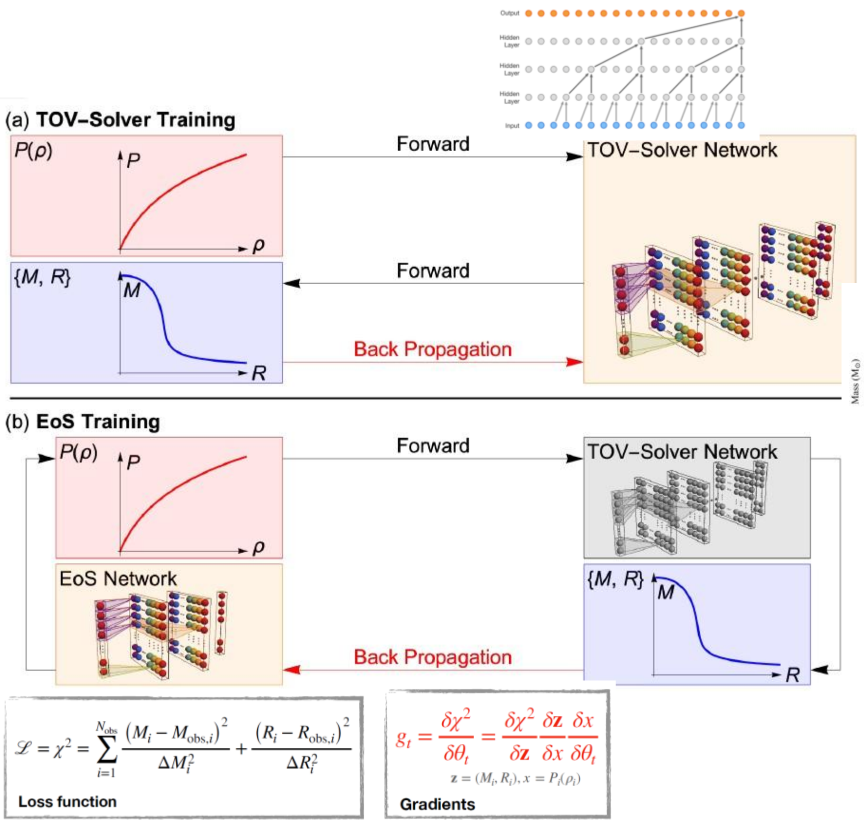
TOV equations

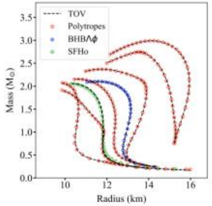
$$-\frac{dP}{dr} = \frac{[\epsilon(r) + P(r)]\,[M(r) + 4\pi r^3 P(r)]}{r[r - 2M(r)]}$$

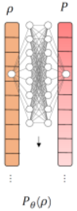
$$\frac{dM(r)}{dr} = 4\pi r^2 \epsilon(r),$$



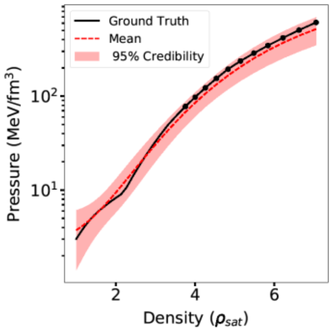
Auto-diff framework and Results

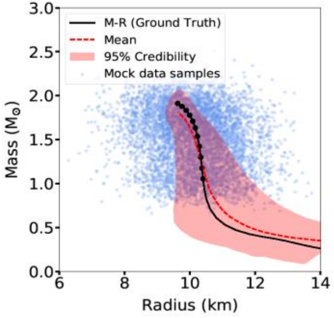




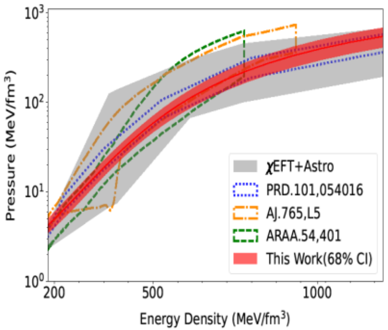


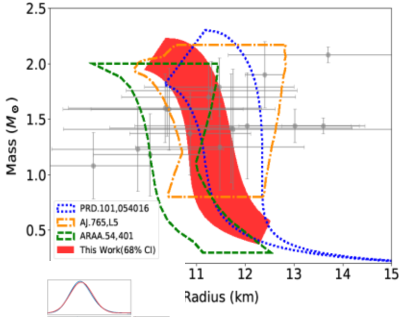
- Well validated through **Mock Tests**

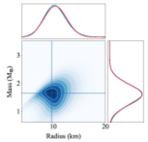




- With **real observable** we reconstruct the NS EoS also

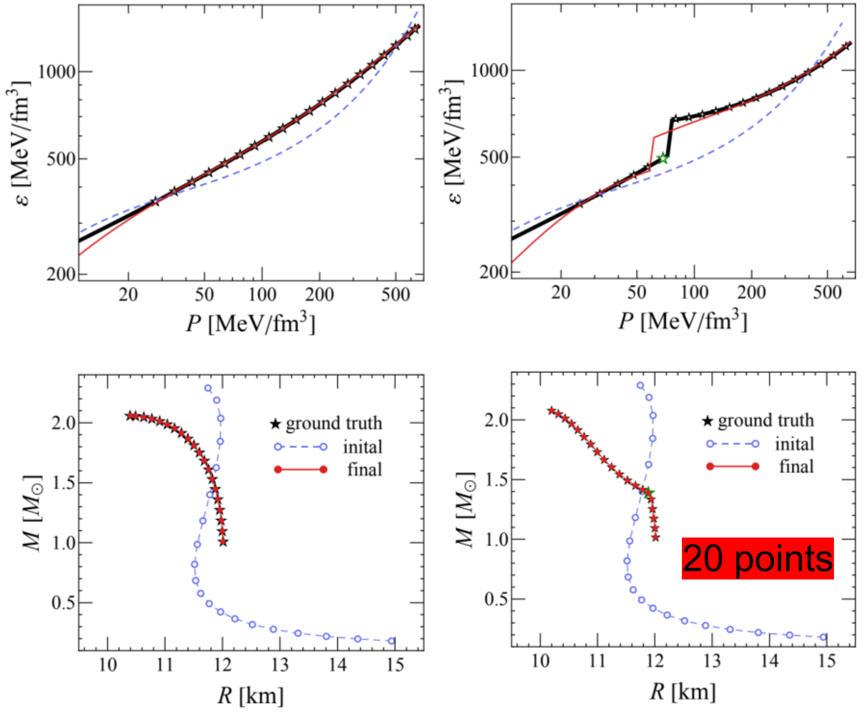






S. Soma, L. Wang, S. Shi, H. Stoecker, **K. Zhou**, **JCAP** 98 (2022) 071
S. Soma, L. Wang, S. Shi, H. Stoecker, **K. Zhou**, **Phys. Rev. D** 107 (2023)083028

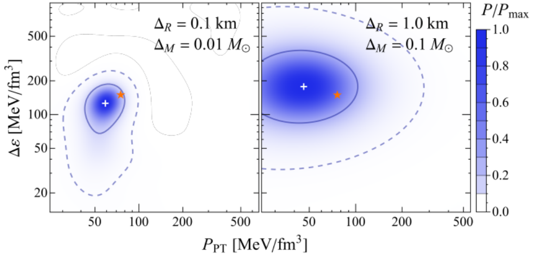
First-order phase transition reconstruction from NS obs. via auto-diff

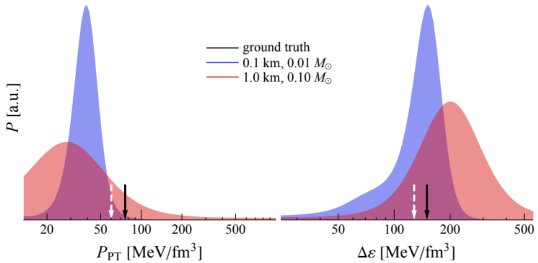


<u>Linear response analysis</u> get the gradients! Then use DNN :

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

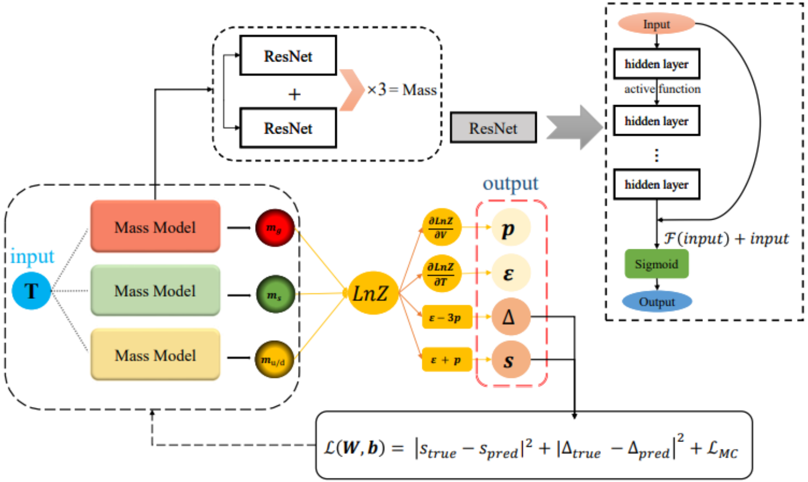
We adopt `SFHo` as the baseline EoS and introduce a PT with latent heat $\Delta\varepsilon = 150\ \mathrm{MeV/fm^3}$ at pressure $P_{\mathrm{PT}} = 76\ \mathrm{MeV/fm^3}$. Above the PT point, we take the stiffest (causal) limit that $c_s = 1$. We employ twenty

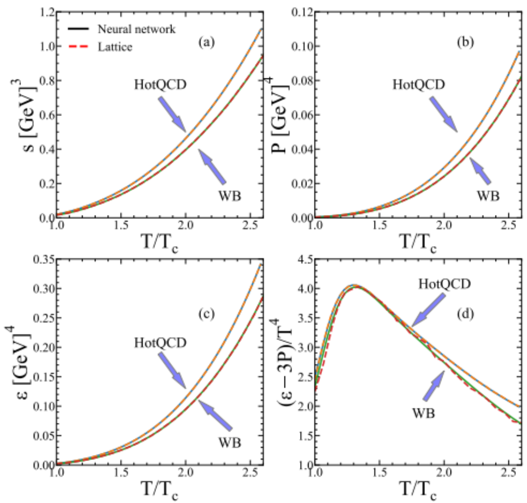


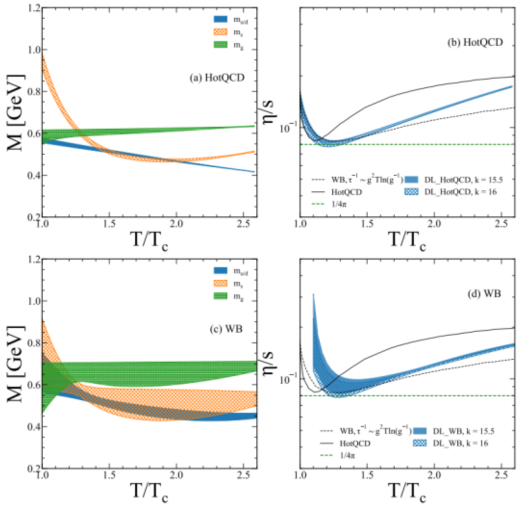


R. Li, S. Han, Z. Lin, L. Wang, K. Zhou, S. Shi, **arXiv:2501.15819**

Quasi-particle analysis of lQCD thermodynamics







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

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

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

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

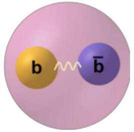
$$\chi_i^B = \left.\frac{\partial P(T, \hat{\mu}_B)/T^4}{\partial \hat{\mu}_B^i}\right|_{\hat{\mu}_B = 0}, \hat{\mu}_B = \mu_B/T.$$

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

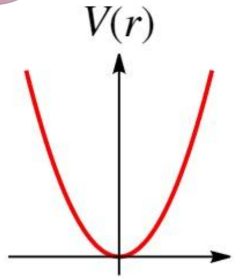
F. Li, H. Lue, L. Pang and G. Qin, **Phys. Lett. B** 2023, arXiv:2211.07994

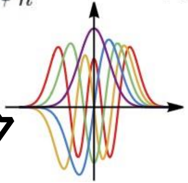
Recently generalized to finite baryon chemical potential (use lQCD res with Taylor expansion) → **arXiv:2501.10012**

HQ Potential Model, Inverse Shroedinger Eq.

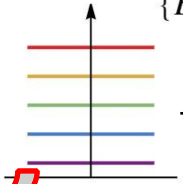


$$\hat{H}\psi_n = -\frac{\nabla^2}{2m_\mu}\psi_n + V(r)\psi_n = E_n\psi_n \qquad \{\psi_n(r)\}$$





$\{E_n\}$



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

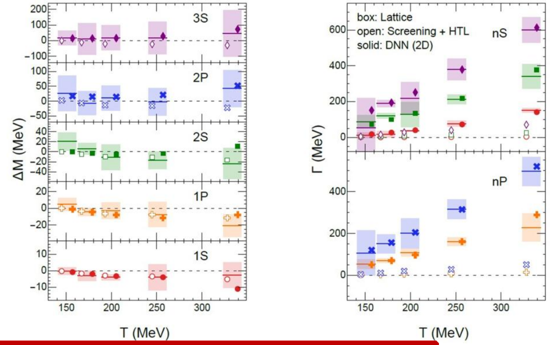
$$V(T,r) = V_R(T,r) + i \cdot V_I(T,r)$$

?

How to extract effective potential given limited spectroscopy ? →

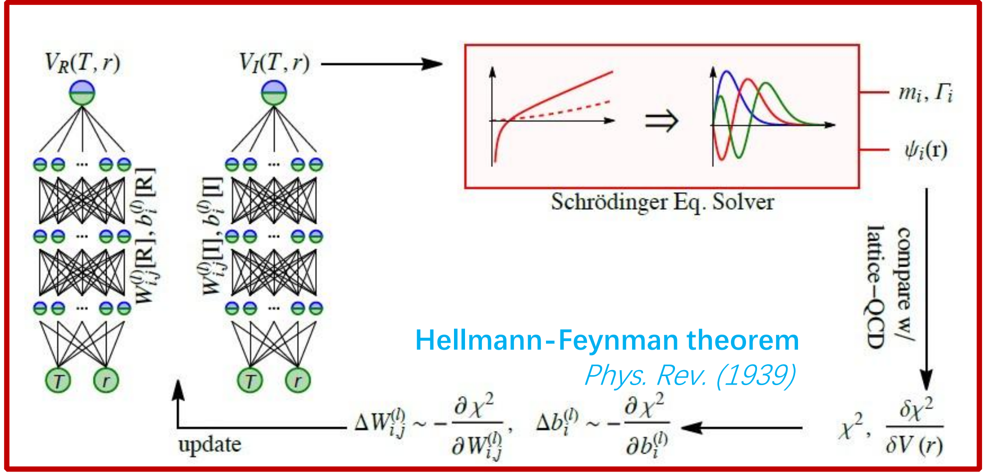
New lQCD results cannot be explained by Perturbative HTL-inspired potentials !

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



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

$$T \in \{0, 151, 173, 199, 251, 334\}\ \mathrm{MeV}$$
$$i \in \{1S, 2S, 3S, 1P, 2P\}$$



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) ?

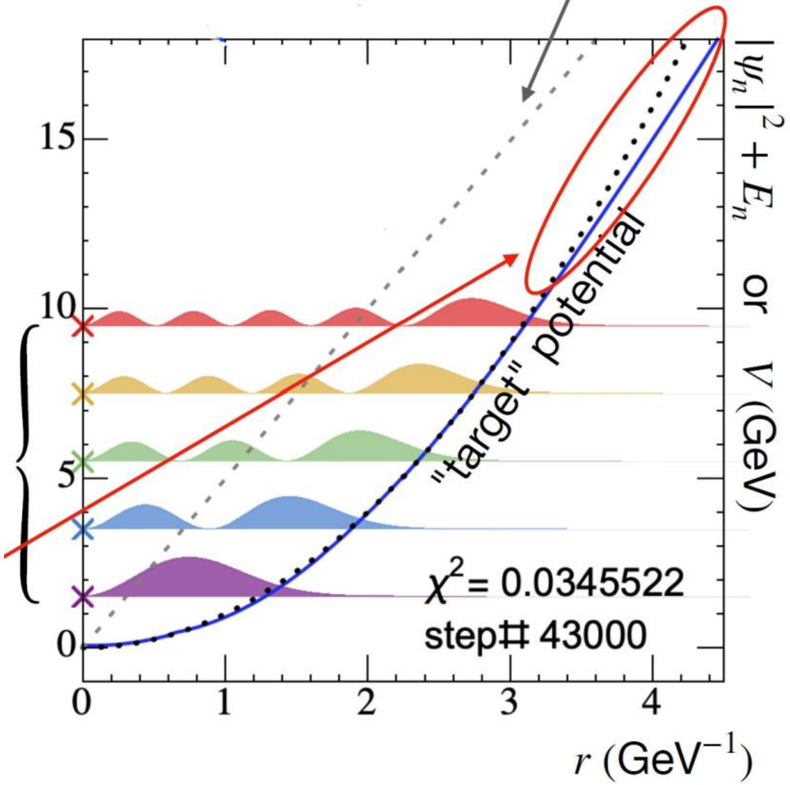
-- Yes! But to some range decided by the used states.

initial potential

Learn V(r) from 5 eigenvalues :

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

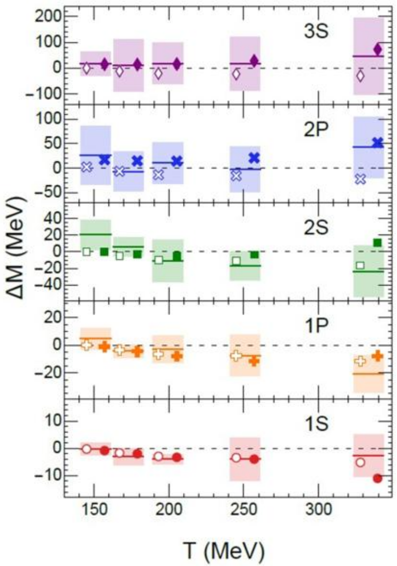
target spectrum

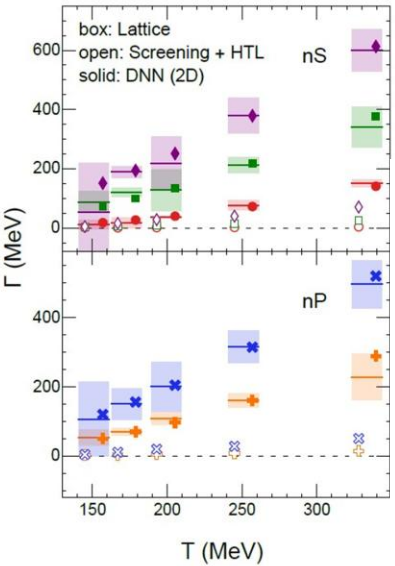


Deviation @ given states' wavefunction vanishes

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

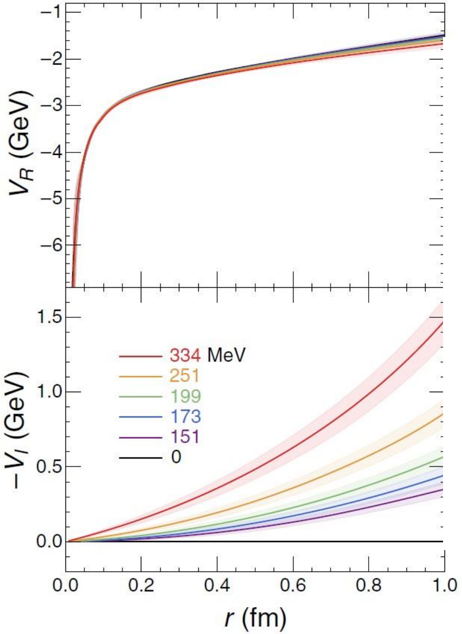
Results with lattice data for mass/width and the reconstructed HQ Potential

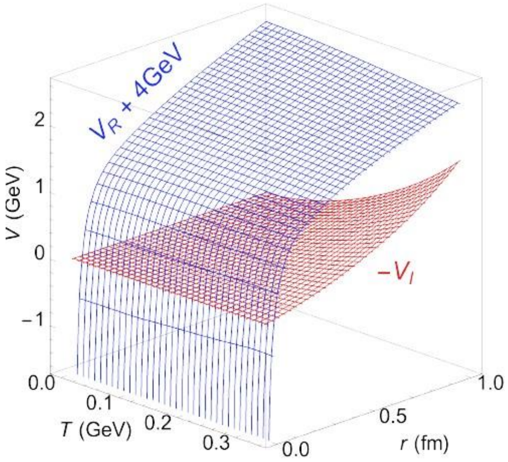




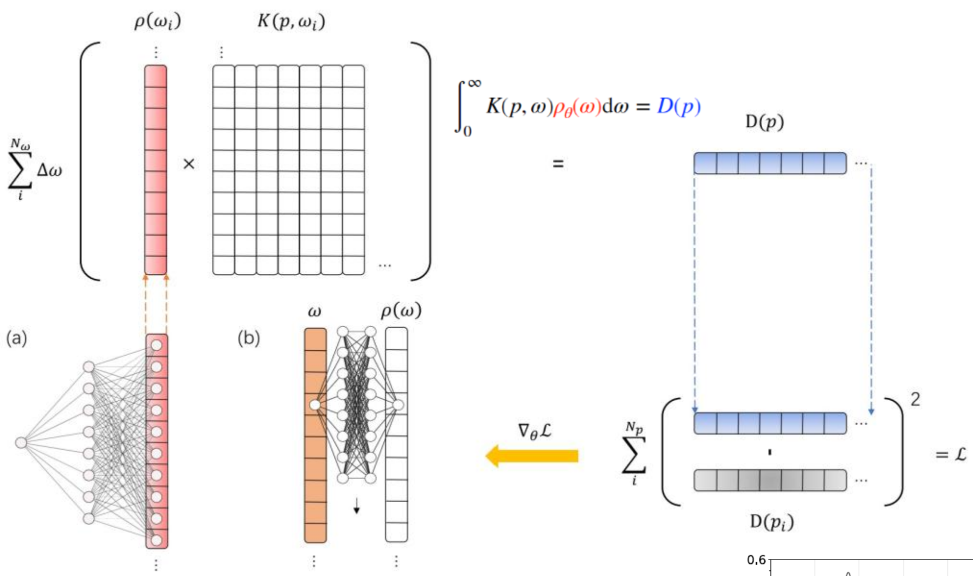
Chi2-per-data=16.5/30

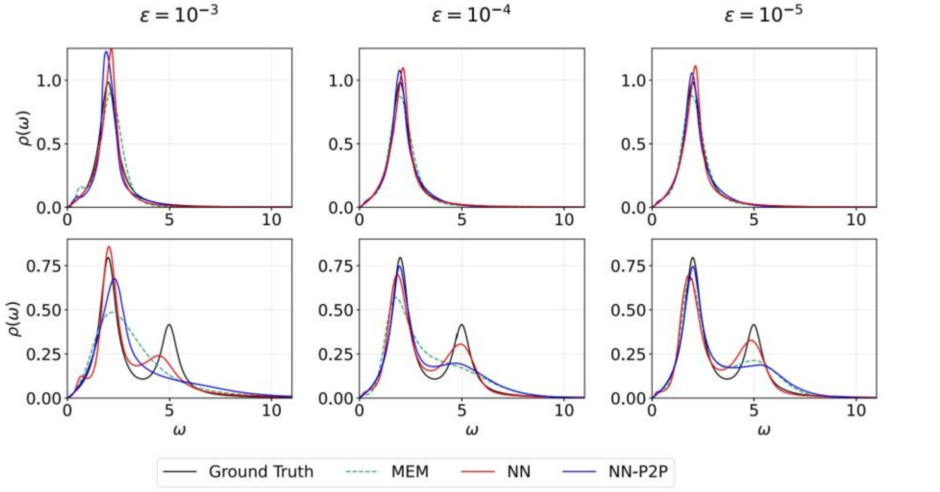
The reconstructed T, r dependent potential

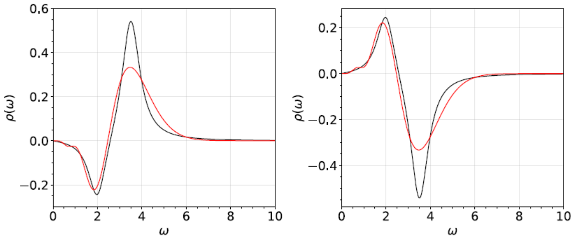


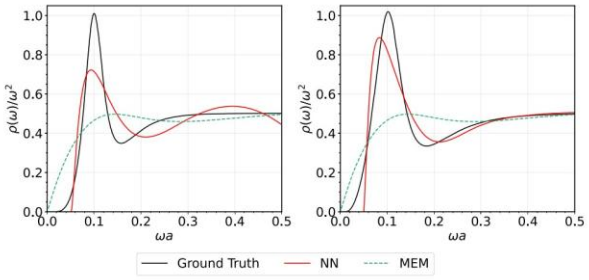


S.S, K. Z, J.Z, S.M., P. Z, Phys. Rev. D 105 (2022) 1, 1









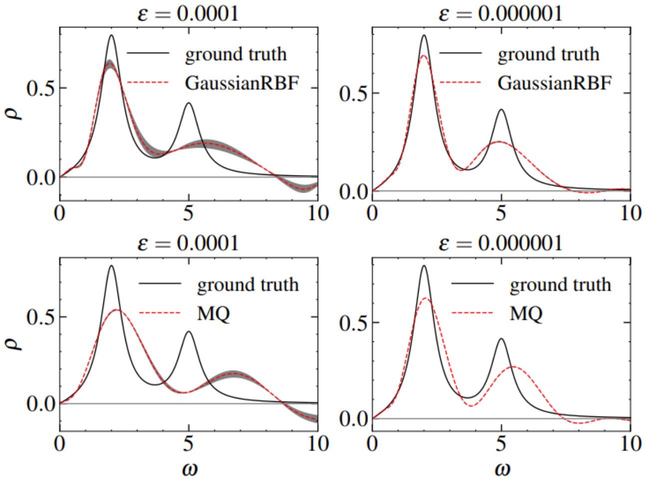
$$G(\tau,T) = \int_0^\infty \frac{d\omega}{2\pi} K(\omega,\tau,T)\rho(\omega,T)$$

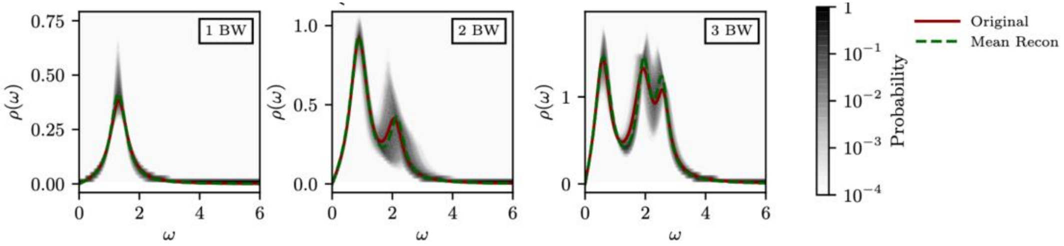
$$K(\omega,\tau,T) = \frac{\cosh \omega(\tau - \frac{1}{2T})}{\sinh \frac{\omega}{2T}}$$

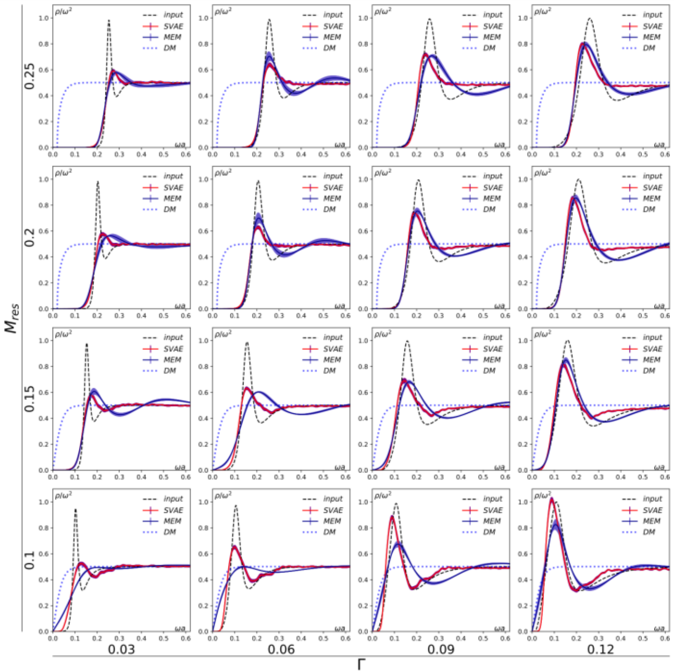
L. Wang, S. Shi and K. Zhou

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

Spectral function reconstruction from Euclidean correlator

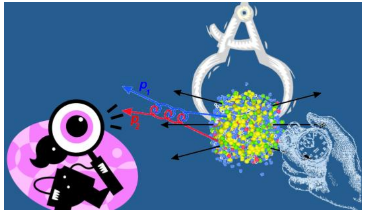




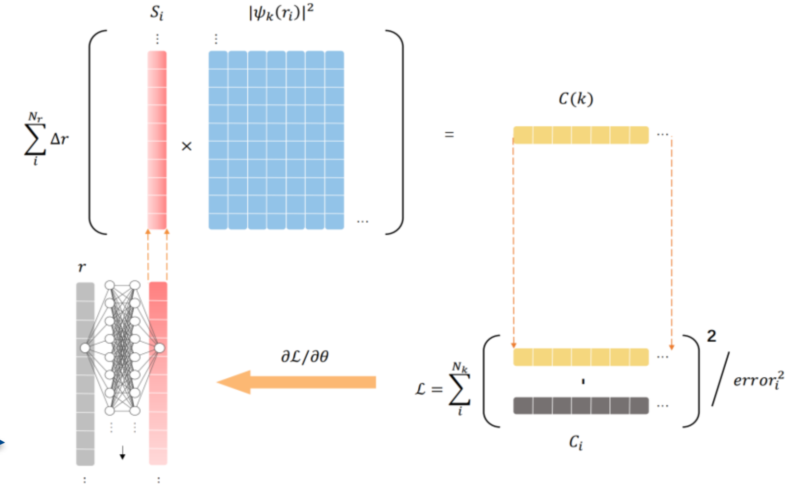


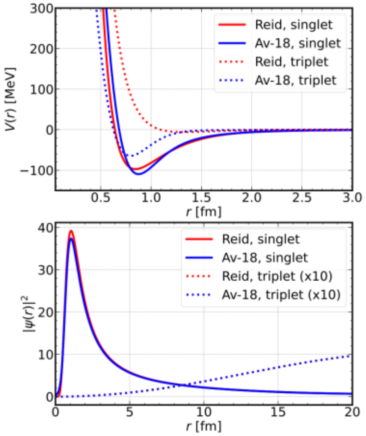
Offler+ 1912.12900; [Kades+ PRD102 (2020) 096001;
Chen+ 2110.13521; Zhou+ PRD104(2021)076011
Horak+ PRD105 (2022) 036014

Hadron emission source reconstruction via femtoscopy



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

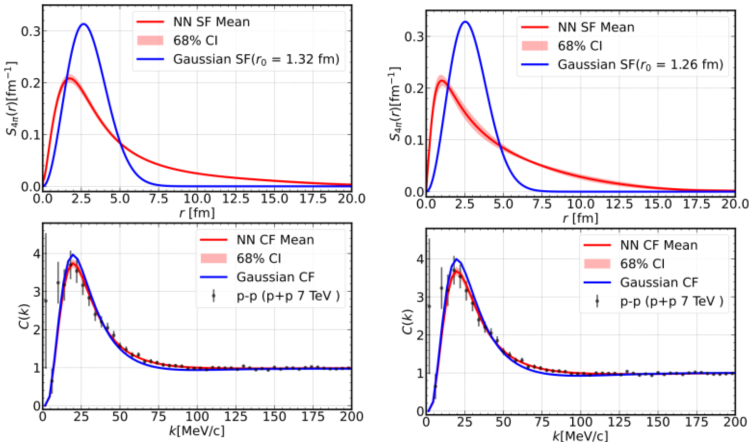




Learning Hadron Emitting Sources with Deep Neural Networks

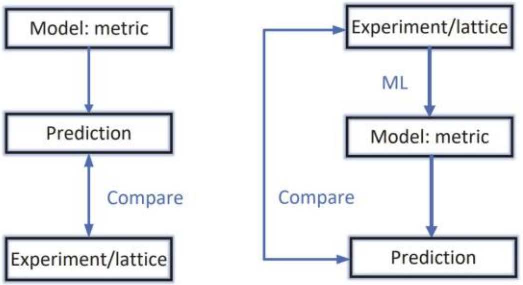
Lingxiao Wang[1] and Jiaxing Zhao[2,3,*]

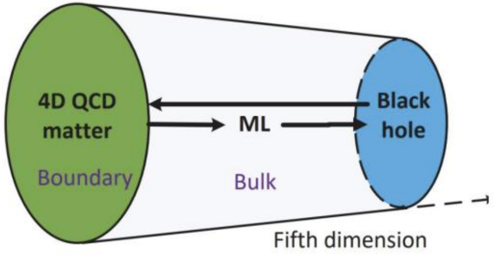
arXiv:2411.16343

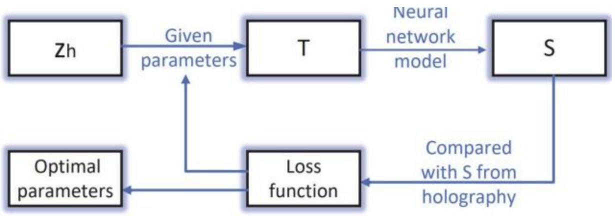


ML Holography (EMD model) with lattice QCD reference

Conventional Holographic model: **ML Holographic model:**







Einstein-Maxwell-Dilation model

O. DeWolfe, S. S. Gubser, and C. Rosen, Phys. Rev. D 83, 086005 (2011), arXiv:1012.1864.

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

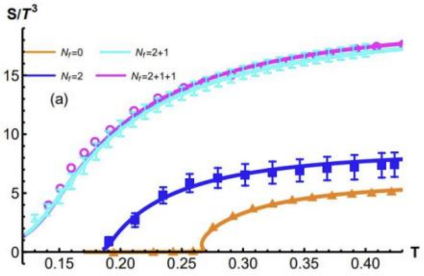
Non-conformal

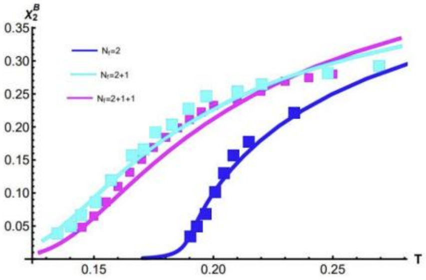
ϕ is dilaton $\quad\quad$ F is the tensor of gauge field

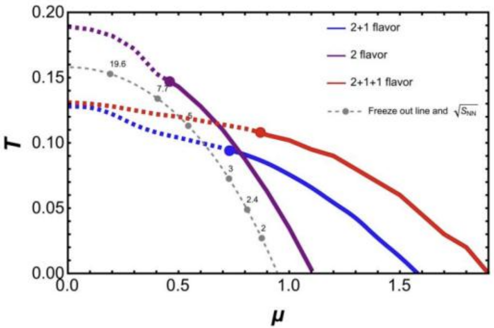
Metric ansatz: $$ds^2 = \frac{e^{2A(z)}}{z^2}\left[-g(z)dt^2 + \frac{dz^2}{g(z)} + d\vec{x}^2\right]$$

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

$$s = \frac{e^{3A(z_h)}}{4G_5 z_h^3}. \qquad \chi_2^B = \frac{1}{T^2}\frac{\partial\rho}{\partial\mu}.$$





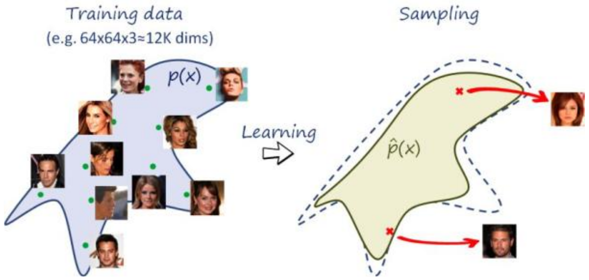


X. Chen, M. Huang, **Phys.Rev.D** 109 (2024) L051902; **JHEP**02 (2025)123

Generative Models Problem Set

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



I. Goodfellow, arXiv:1701.00160 (2017)

Suppose observation dataset :

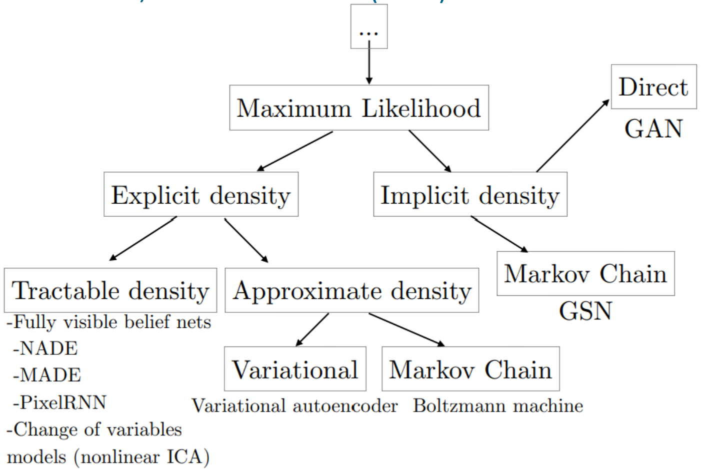
$$\mathbf{X} = \{x^{(1)}, x^{(2)}, ..., x^{(N)}\} \overset{i.i.d}{\sim} p_{data}(x)$$

We use parametric model to approach the data distribution :

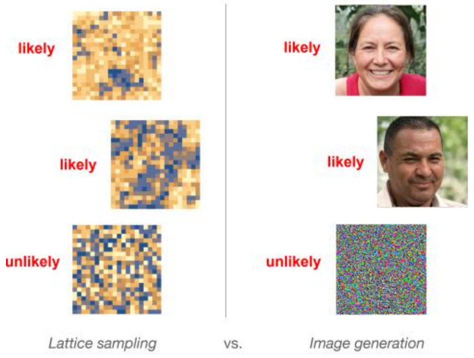
$$p_\theta(x) \to p_{data}(x)$$

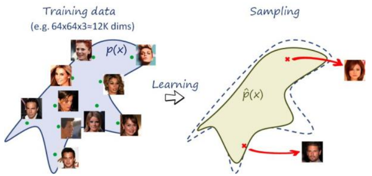
- Maximize Likelihood Estimation :

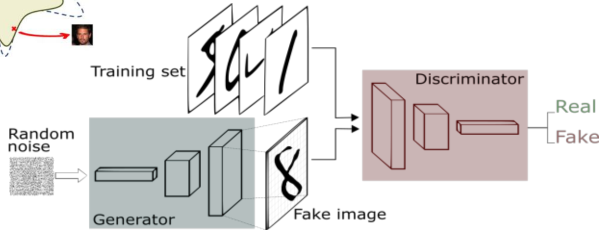
$$\theta^* = \arg\max_\theta \log p_\theta(\mathbf{X}) = \arg\max_\theta \frac{1}{N}\sum_{i=1}^{N} \log p_\theta(x^{(i)})$$

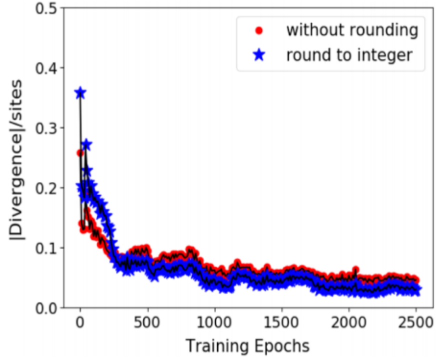


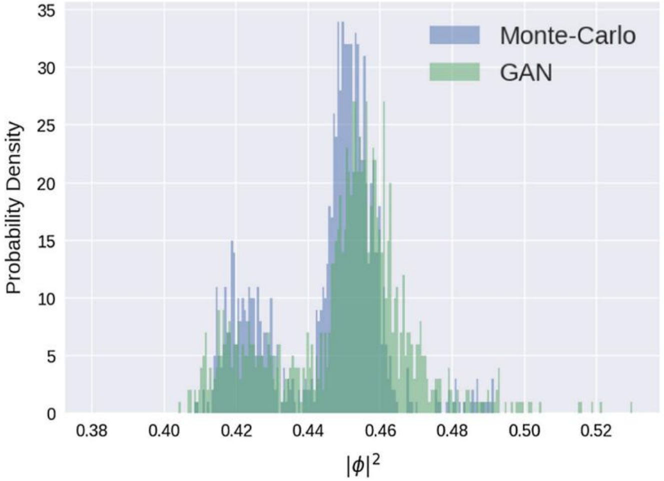
GAN to generate complex scalar field configurations

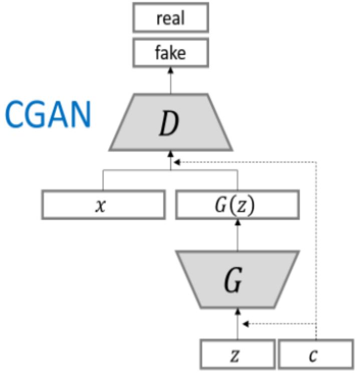


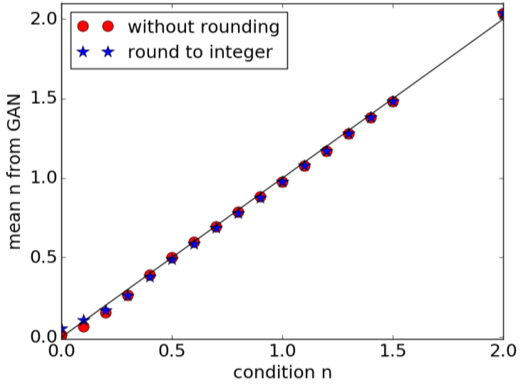






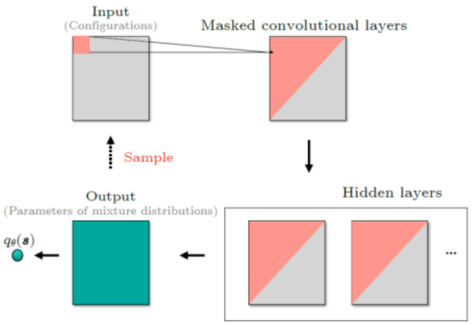






K. Zhou, G. Endrődi, L.-G. Pang, and H. Stöcker,
PRD 100, 011501 (2019)

Autoregressive network for variational many-body physics calculation

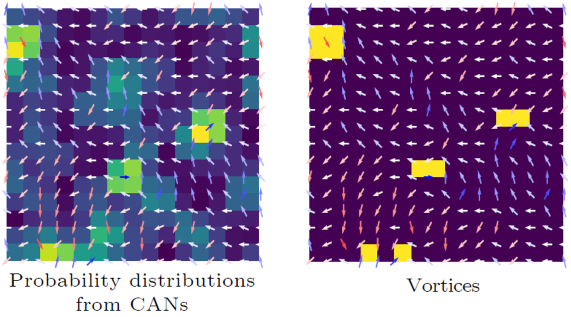


$$\boxed{p(\mathbf{s}) = \frac{e^{-\beta E(\mathbf{s})}}{Z}}$$

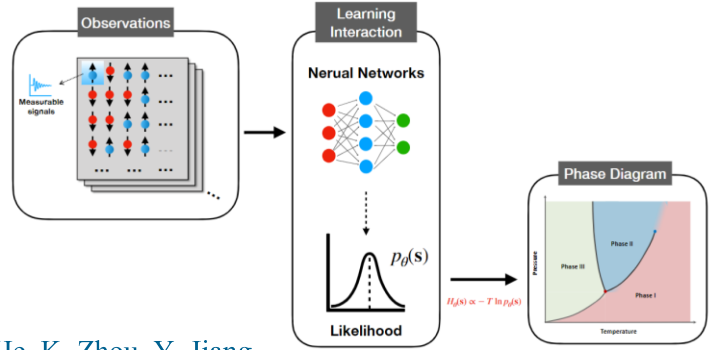
$$q_\theta(\mathbf{s}) = \prod_{i=1}^{N} q_\theta(s_i \mid s_1, \ldots, s_{i-1})$$

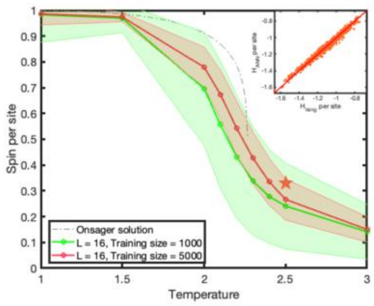
$$D_{\mathrm{KL}}(q_\theta \,\|\, p) = \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \ln\left(\frac{q_\theta(\mathbf{s})}{p(\mathbf{s})}\right) = \beta(F_q - F)$$

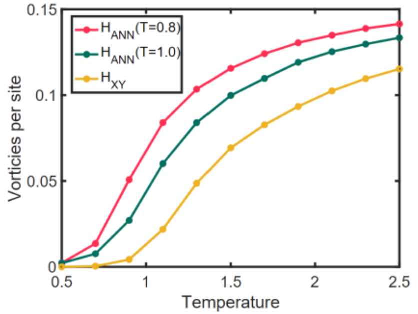
$$F_q = \frac{1}{\beta}\sum_{\mathbf{s}} q_\theta(\mathbf{s})\left[\beta E(\mathbf{s}) + \ln q_\theta(\mathbf{s})\right]$$



L. Wang, L. He, Y. Jiang, K. Zhou,
Chi. Phys. Lett. 2022, arXiv:2007.01037





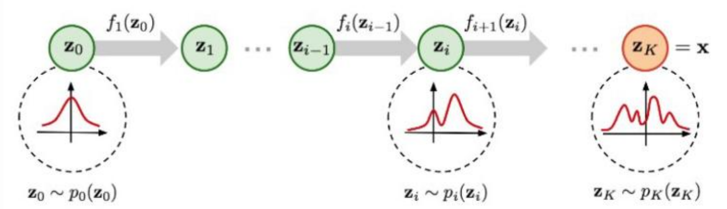


T. Xu, L. Wang, L. He, K. Zhou, Y. Jiang,
Chi. Phys. C 2024

Flow based generative model (given unnormalized distribution)

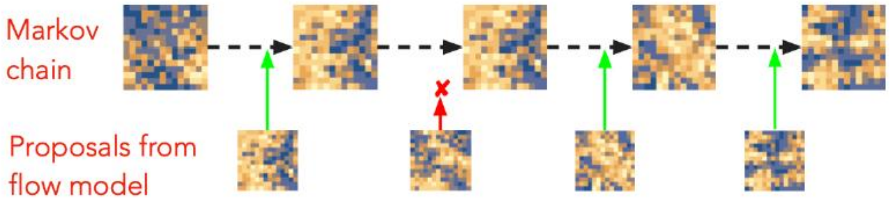
A series **(Flow)** of invertible/bijective transformations for p(z)

compose several invertible transformations to form the flow :



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

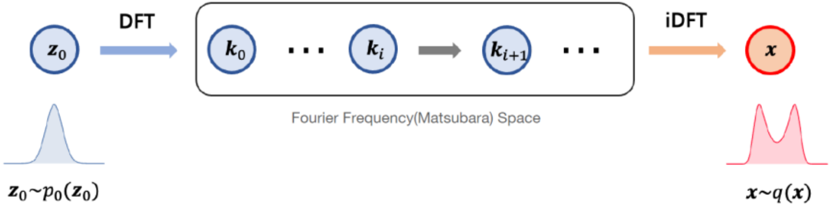
$$\rightarrow \log p(\mathbf{x}) = \log p_0(f^{-1}(\mathbf{x})) + \sum_{i=1}^{K} \log|\det J_{f_i^{-1}}| = \log p_0(\mathbf{z}_0) - \sum_{i=1}^{K} \log|\det J_{f_i}|$$

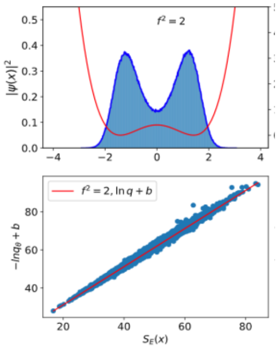


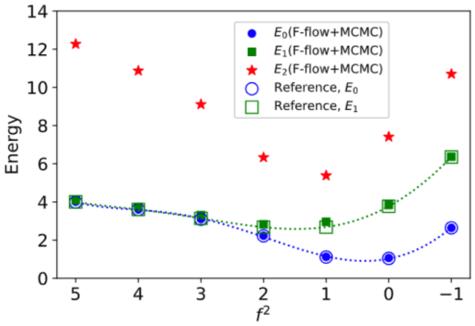
Albergo +, 1904.12072; Boyda +, 2008.05456; Favoni +, 2012.12901; Abbott +, 2208.03832; Abbott +, 2211.07541; Abbott +, 2305.02402; Bulgarelli+ 2412.00200 (SU(3)); Abbott +, arXiv:2502.00263
K.C, G. K., S. R., D. R., P. S., **Nature Reviews Physics** 5, 526-535 (2023)

Fourier Flow Model

S.Chen, O. Savchuk, S. Zheng, B. Chen, H. Stoecker, L. Wang, K. Zhou, **PRD107, 056001(2023)**



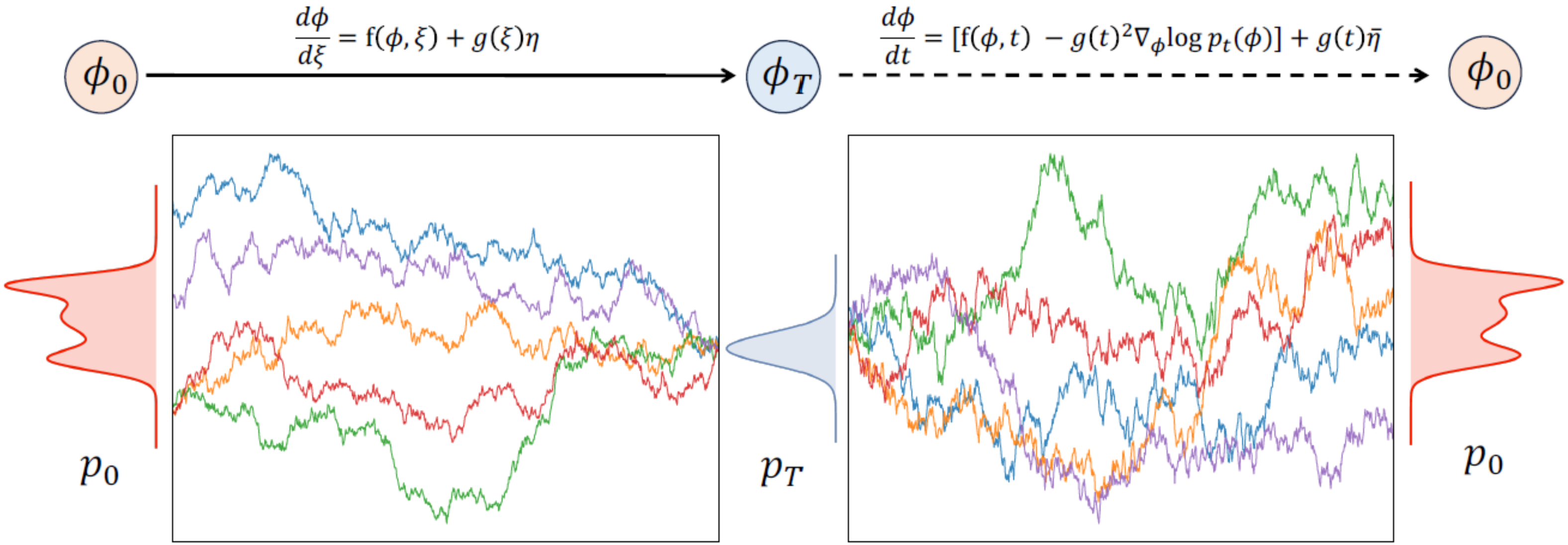




Diffusion Model

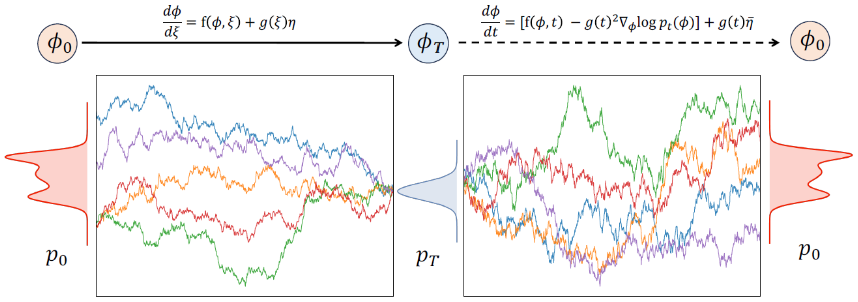
“A heavy quark move inside quark-gluon plasma”

Diffusion Model on lattice QFT configurations



L. Wang, G. Arts, K. Zhou, JHEP 05 (2024) 060
L. Wang, G. Arts, K. Zhou, arXiv:2311.03578 (NeurIPS 2023 workshop “ML&Physical Sciences”)
G. A, D. E. H, L. W, K. Z, 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)

Diffusion Model for field configurations



- Forward diffusion SDE

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

- Backward diffusion SDE

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

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

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

- **Sample generation SDE** $\frac{d\phi}{dt} = [f(\phi,t) - g^2(t)s_{\hat{\theta}}(\phi,t)] + g(t)\bar{\eta}(t).$

$$\tau \equiv T - t \qquad \frac{d\phi}{d\tau} = g_\tau^2 \nabla_\phi \log q_\tau(\phi) + g_\tau \bar{\eta}(\tau)$$

$$\frac{\partial p_\tau(\phi)}{\partial \tau} = \int d^n x \left\{ g_\tau^2 \frac{\delta}{\delta\phi} \left(\bar{\alpha}\frac{\delta}{\delta\phi} + \nabla_\phi S_{\mathrm{DM}} \right) \right\} p_\tau(\phi). \qquad \nabla_\phi S_{\mathrm{DM}} \equiv -\nabla_\phi \log q_\tau(\phi) \quad p_{\mathrm{eq}}(\phi) \propto e^{-S_{\mathrm{DM}}/\bar{\alpha}} \qquad p_{\tau=T}(\phi) \to P[\phi, T]$$

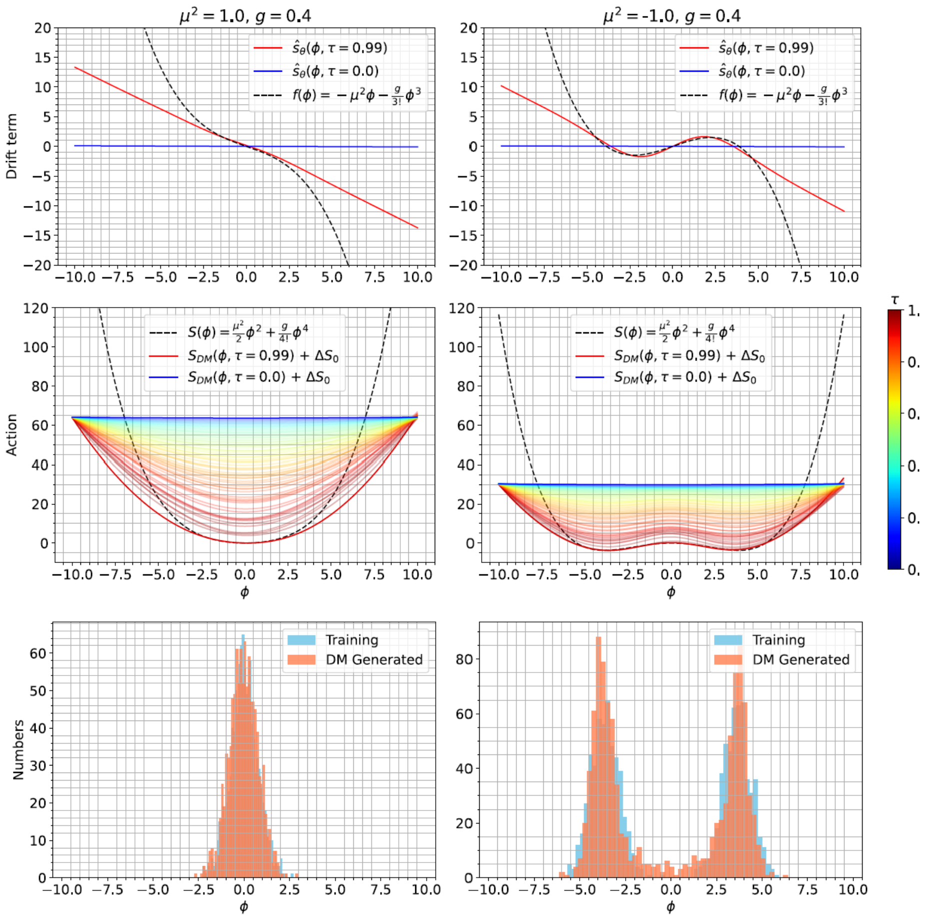
$$O(\bar{\alpha}) \sim O(\hbar)$$

- A flow of <u>**effective action**</u> will be learned in DMs

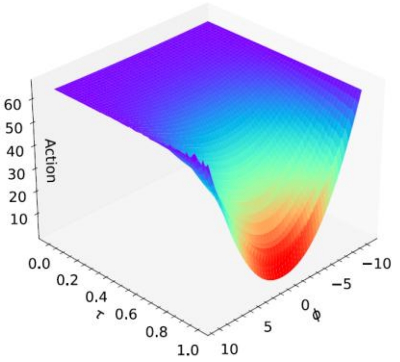
L. Wang, G. Arts, K. Zhou, JHEP 05(2024) 060

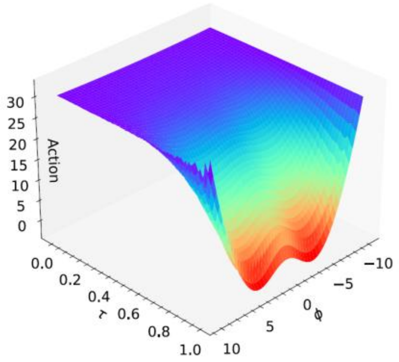
sampling from a DM is equivalent to optimizing a stochastic trajectory to approach the “equilibrium state”

Effective Action on toy model



- Flow of an effective action



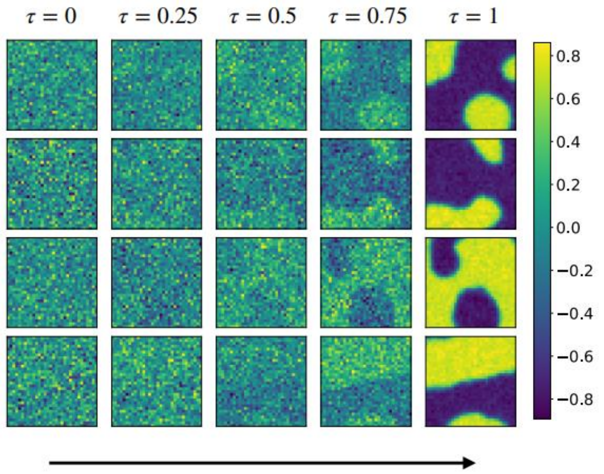


DM on 2d scalar ϕ^4 model

- **32x32 lattice, HMC generated <u>5120 configurations</u> for training**

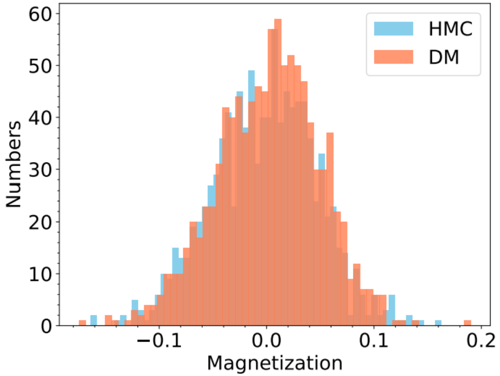
$$S_E = \sum_x [-2\kappa \sum_{\mu=1}^{d} \phi(x)\phi(x+\hat{\mu}) + (1-2\lambda)\phi(x)^2 + \lambda\phi(x)^4].$$

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

Relation to (inverse) RG

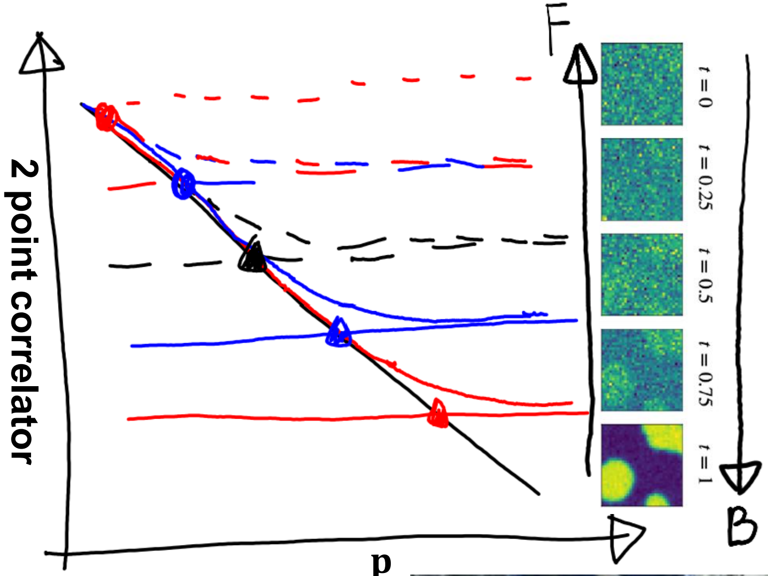
- Forward diffusion kernel: **gaussian smoothing**

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

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

- 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 faster because of the gradually increasing noise level

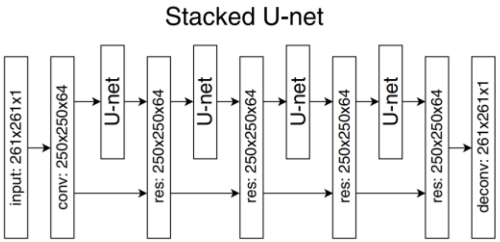
In FRG, the high frequency (short-distance) degrees of freedom is progressively integrated out !
See Semon's and Mathis's talk!

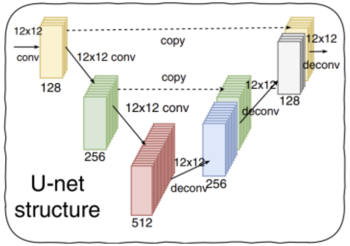
U-net Emulator for relativistic hydrodynamics

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

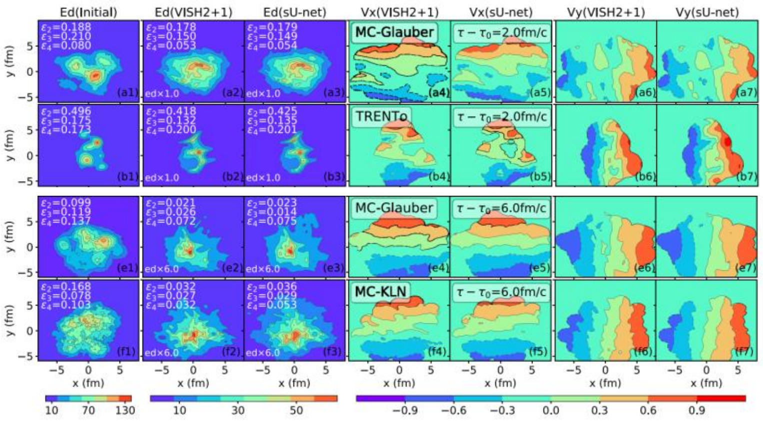
Applications of deep learning to relativistic hydrodynamics

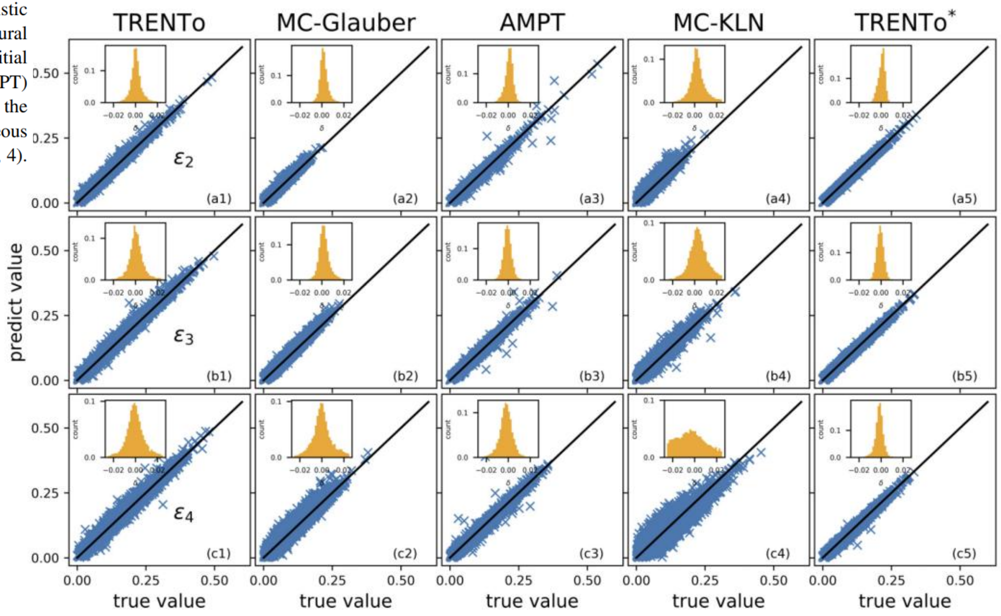
Hengfeng Huang,[1,2] Bowen Xiao,[3] Ziming Liu,[1] Zeming Wu,[1,2] Yadong Mu,[3,4] and Huichao Song[1,2,5]





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

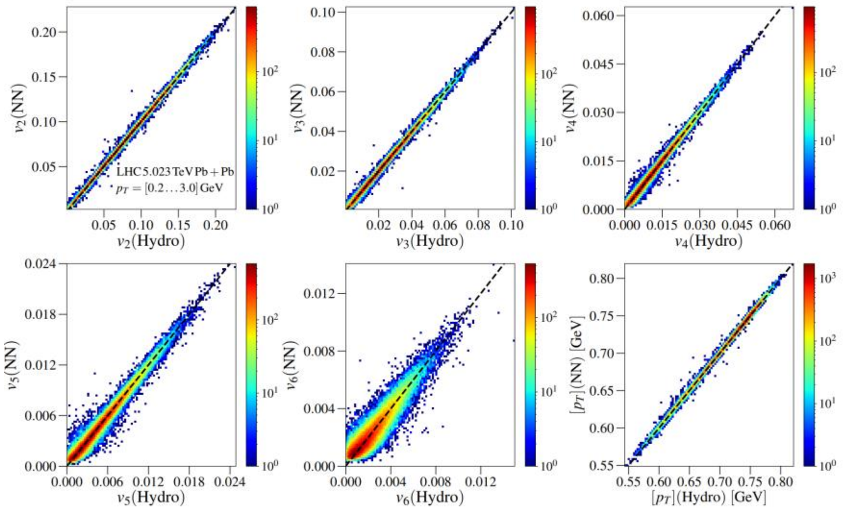
PHYSICAL REVIEW C **108**, 034905 (2023)

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.

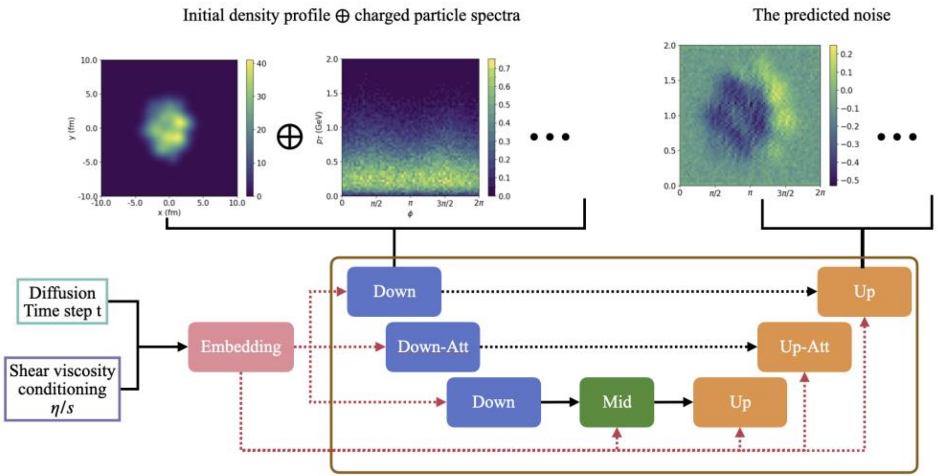
Similar study for intermediate energies with IBUU +DNN emulation see: **B. Li+, arXiv:2406.18421**

Generative diffusion model to heavy-ion collisions

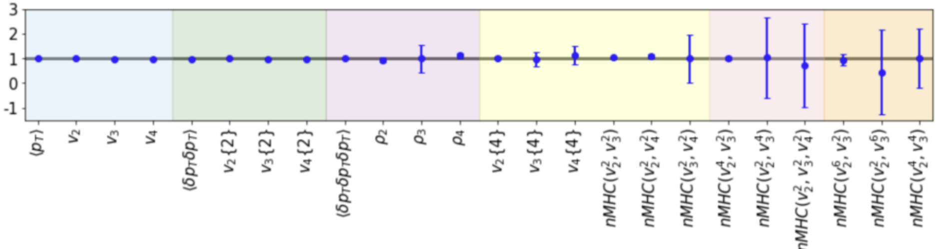
An end-to-end generative diffusion model for heavy-ion collisions

arXiv:2410.13069

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



$$S_t = \sqrt{\bar{\alpha}_t} S_0 + \sqrt{1-\bar{\alpha}_t}\,\varepsilon$$



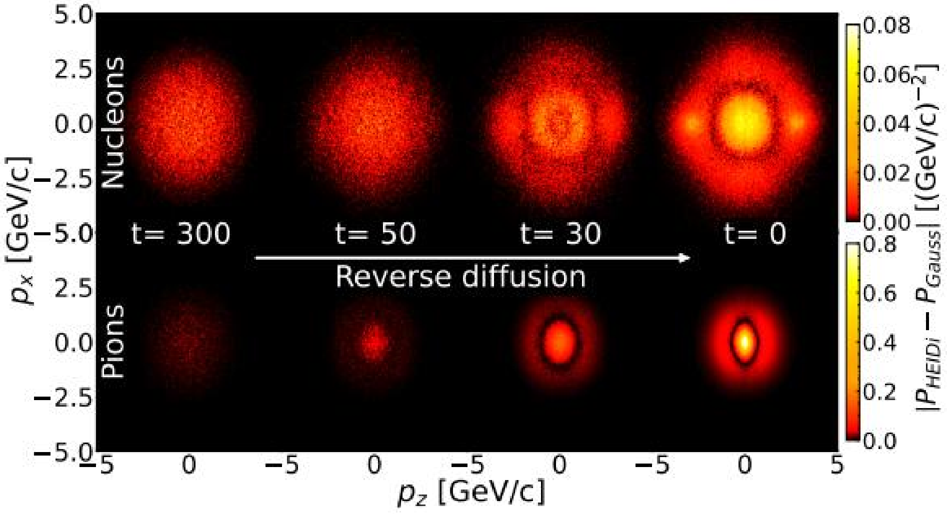
tor. We carried out (2+1)D minimum bias simulations of Pb-Pb collisions at 5.02 TeV, choosing the shear viscosity η/s to be one of three distinct values: 0.0, 0.1, and 0.2. For each value of η/s, we generate 12,000 pairs of initial entropy density profiles and final particle spectra, corresponding to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

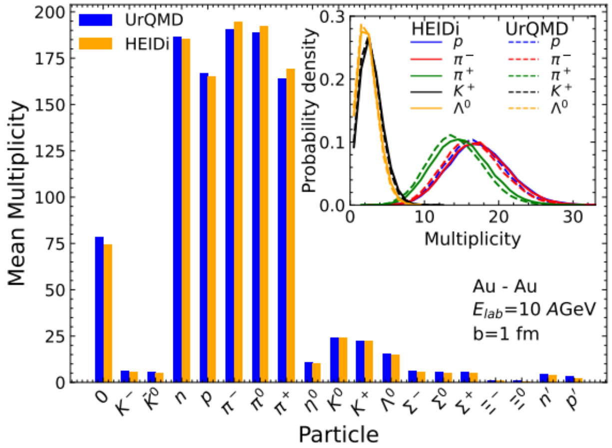
Considering that the spectra $\boldsymbol{S}_0$ depend on the initial entropy density profiles $\boldsymbol{I}$ and the shear viscosity η/s, we train a conditional reverse diffusion process $p(\boldsymbol{S}_0|\boldsymbol{I},\eta/s)$ without modifying the forward process.

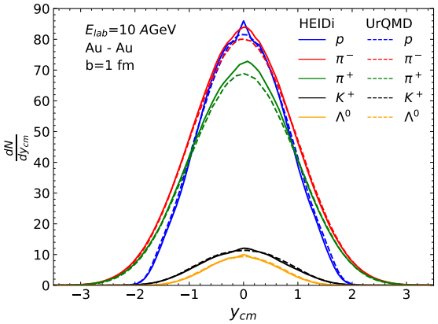
one single central collision event in just 10^{-1} seconds on a GeForce GTX 4090 GPU.

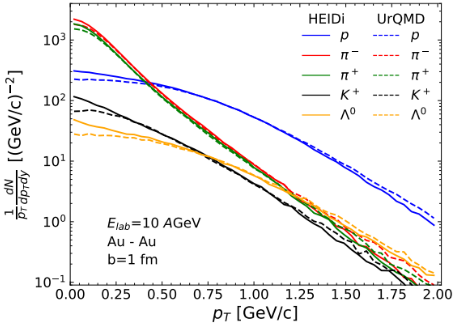
ble precision, as the traditional numerical simulation of hydrodynamics for one central event typically takes approximately 120 minutes (10^4 seconds) on a single CPU.

- **18k UrQMD simulation events for Au-Au@10 AGeV**
- **HEIDi: Heavy-ion Events through Intelligent Diffusion**
- **Point-cloud representation: momentum + ID**









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

What’s the future? – Generative AI and LLM maybe the future

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***

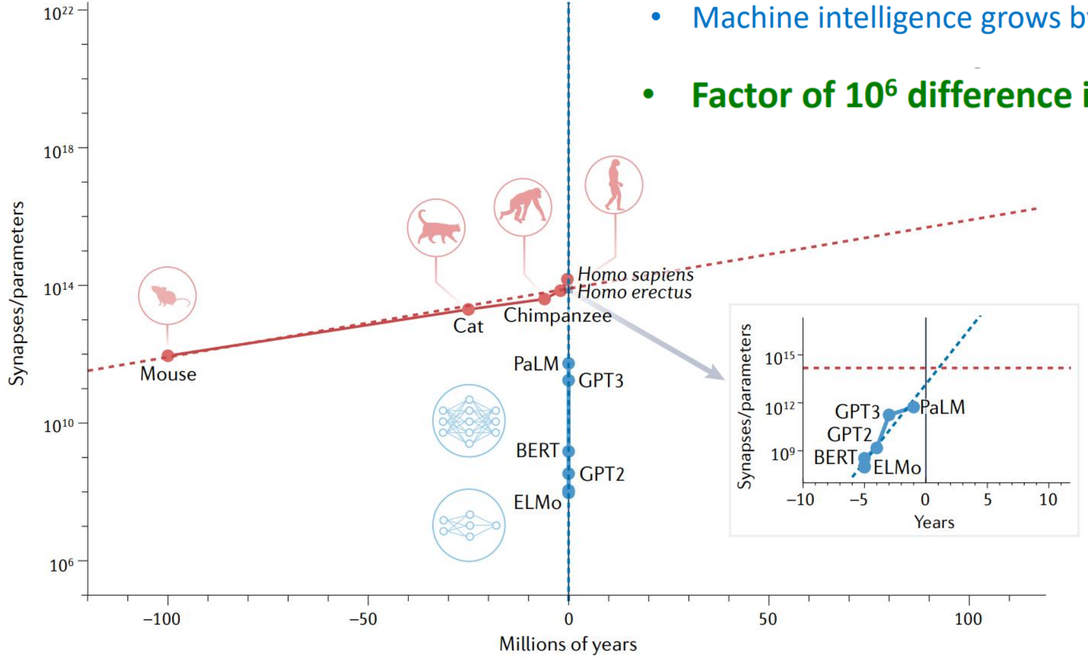


Fig. 1 | The evolution of biological and artificial intelligence takes place on dramatically different timescales. Any hope of interpreting and understanding AI will exponentially fade. Some example data points are highlighted in the evolution of biological (red) and artificial (blue) intelligence. The 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

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

nature reviews physics
https://doi.org/10.1038/s42254-024-00798-x

Perspective

Check for updates

Physics-driven learning for inverse problems in quantum chromodynamics

Gert Aarts [1], Kenji Fukushima [2], Tetsuo Hatsuda [3], Andreas Ipp [4], Shuzhe Shi [5], Lingxiao Wang [3] & Kai Zhou [6,7]

Abstract

The integration of deep learning techniques and physics-driven designs is reforming the way we address inverse problems, in which accurate physical properties are extracted from complex observations. This is particularly relevant for quantum chromodynamics (QCD) – the theory of strong interactions – with its inherent challenges in interpreting

Sections

- Introduction
- Physics-driven learning
- QCD physics
- Conclusions and outlook

Nature Review Physics (2025)

Thanks!

Progress in Particle and Nuclear Physics 135 (2024) 104084

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

Contents lists available at ScienceDirect

ELSEVIER

Progress in Particle and Nuclear Physics

journal homepage: www.elsevier.com/locate/ppnp

Review

Check for updates

Exploring QCD matter in extreme conditions with Machine Learning

Kai Zhou [a,b,*], Lingxiao Wang [a,*], Long-Gang Pang [c,*], Shuzhe Shi [d,e,*]

[a] Frankfurt Institute for Advanced Studies (FIAS), D-60438 Frankfurt am Main, Germany
[b] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China
[c] Institute of Particle Physics and Key Laboratory of Quark and Lepton Physics (MOE), Central China Normal University, Wuhan, 430079, China
[d] Department of Physics, Tsinghua University, Beijing 100084, China

Nuclear Science and Techniques (2023) 34:88
https://doi.org/10.1007/s41365-023-01233-z

REVIEW ARTICLE

Nucl. Sci. Tech. 34 (2023) 6

High-energy nuclear physics meets machine learning

Wan-Bing He [1,2] · Yu-Gang Ma [1,2] · Long-Gang Pang [3] · Hui-Chao Song [4] · Kai Zhou [5]

Received: 10 March 2023 / Revised: 13 April 2023 / Accepted: 18 April 2023 / Published online: 21 June 2023

Abstract

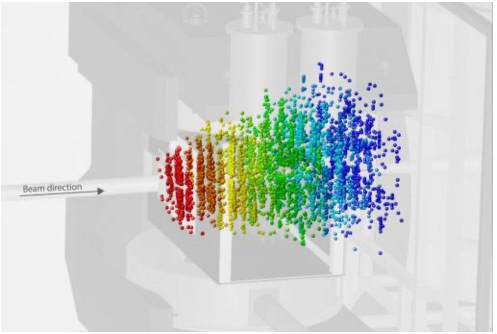
Although seemingly disparate, high-energy nuclear physics (HENP) and machine learning (ML) have begun to merge last few years, yielding interesting results. It is worthy to raise the profile of utilizing this novel mindset from ML in H to help interested readers see the breadth of activities around this intersection. The aim of this mini-review is to infor community of the current status and present an overview of the application of ML to HENP. From different aspects and examples, we examine how scientific questions involving HENP can be answered using ML.

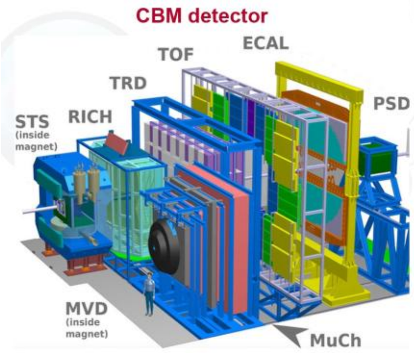
Keywords Heavy-ion collisions · Machine learning · Initial state · Bulk properties · Medium effects · Hard probes ·

point cloud data structure, and CBM challenge

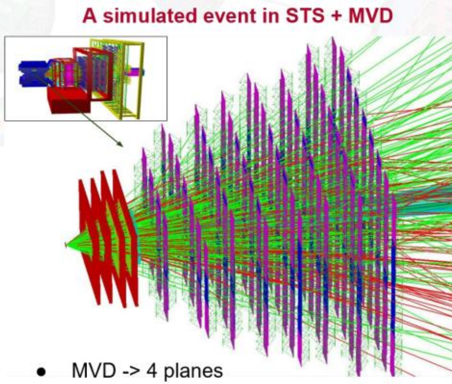
- **Experimental data has inherent point cloud structure**
  - Point clouds: collection of points in space
- Point clouds are represented as 2D array.
  - each row= a point in the point cloud
  - each column =a dimension of the point cloud
- PointNet based models learn directly from point clouds.
  - respects the order invariance of point clouds
  - **direct processing of experimental data**
- Advantages:
  - less processing time ⇒**ideal online algorithm**
  - optimal for higher dimensional data
- We consider the **CBM experiment** as a use case
  - Au-Au collisions
  - 10 AGeV
  - **CBM Challenges** →

x_1	y_1	z_1
x_2	y_2	z_2
⋮	⋮	⋮
x_n	y_n	z_n



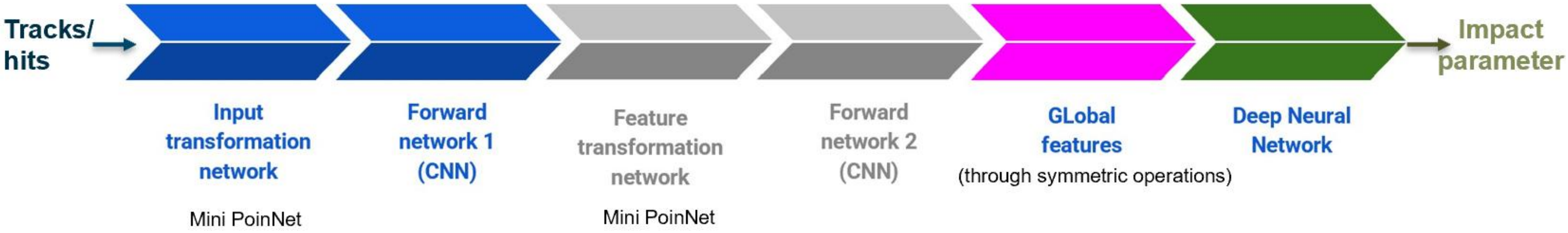


- Upto 45 AGeV collisions
- 10^7 collisions/ Second
- 1000 tracks per collision
- 1 TBytes/Second raw data



- MVD -> 4 planes
  Position resolution: 3.5-6 μm
  Secondary vertex resolution: 50 μm
- STS-> 8 planes
  Momentum resolution: 1 %

models

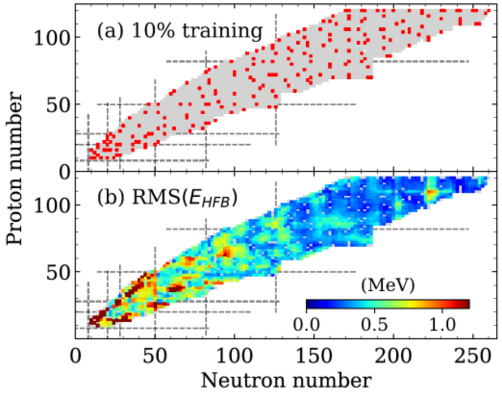


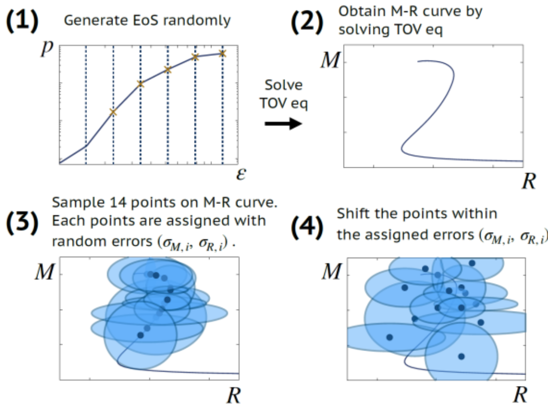
- Works on direct exp. output
- Event-by-event online possible
- Training data generated by **'UrQMD → CbmRoot'** 10^5 Au+Au 10 AGeV events b ~ (0-16)fm

01	**M-hits**	• (x,y,z) of all Hits in MVD
02	**S-hits**	• (x,y,z) of all hits in STS
03	**MS-tracks**	• (x,y,z,dx/dz,dy/dz,q/P) of all tracks in first and last plane from MVD+ STS
04	**HT-combi**	• Hits from MVD and tracks from MVD+STS
05	**Polyfit (non-ML baseline)**	• Third order polynomial fit to multiplicity vs. impact parameter curve

Regressive and Generative AI for High Energy Nuclear Physics

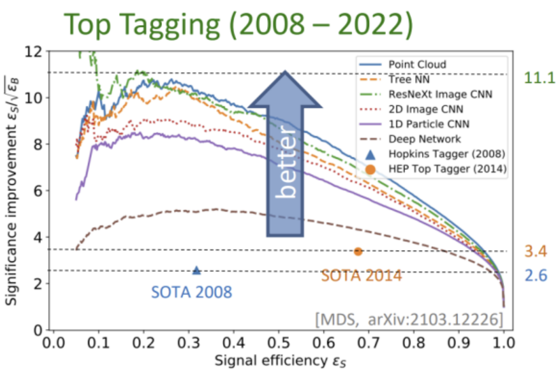
- Nuclear properties prediction
  - Dripline locations, atomic masses, separation energies, superheavy nuclei location…
  - ANN application since 1992, later to beta-decay etc., → BNN
  - Different ML methods, SVM, Gradient boost, BDT,…
- Interpo-/extrapolation of nuclear data, augment nuclear model
  - 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



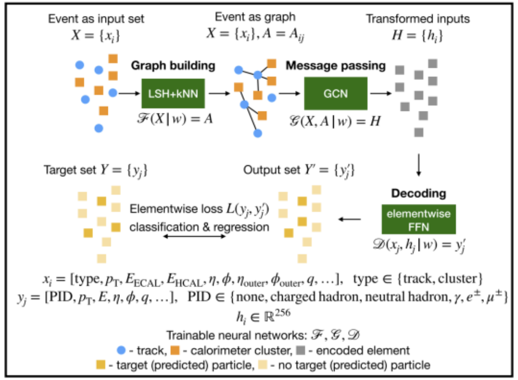


(particle phy) Supervised Learning – Regressive Tasks

- Jet **Tagging**, PID, - <u>BDT, CNN, GNN, PCN </u>(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



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



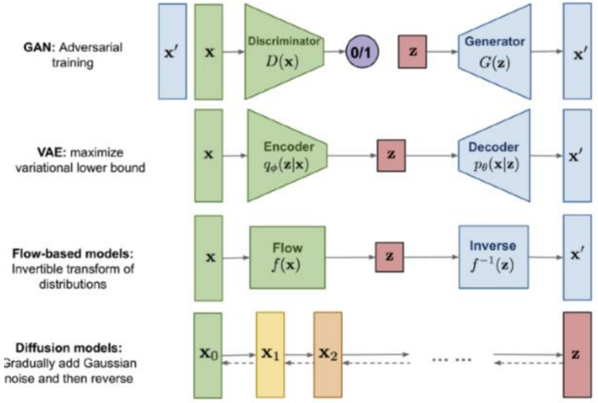
- Real-Time **Trigger** / **Filtering** system – <u>CNN, RL, Q-learning</u>
  - ➢ Ultra-low latency classification of collision event / signals
  - ➢ Implementing ML inference on specialized hardware (FPGAs)
  - ➢ Online distillation reducing raw data flow from Terabytes/second to manageable levels

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

- **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

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



Fast Calorimeter Simulation Challenge 2022

View on GitHub

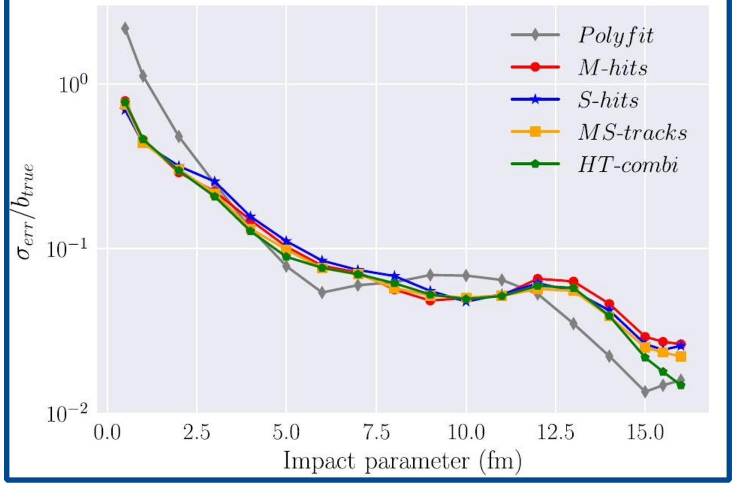
Welcome to the home of the first-ever Fast Calorimeter Simulation Challenge!

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.

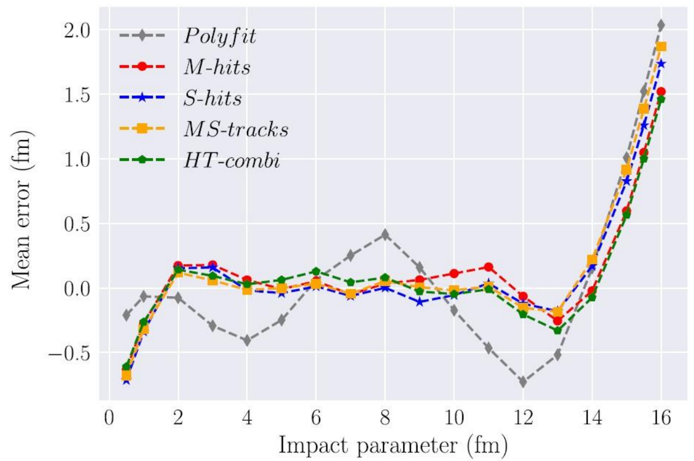
This challenge is modeled after two previous, highly successful data challenges in HEP – the top tagging community challenge and the LHC Olympics 2020 anomaly detection challenge.

Test results

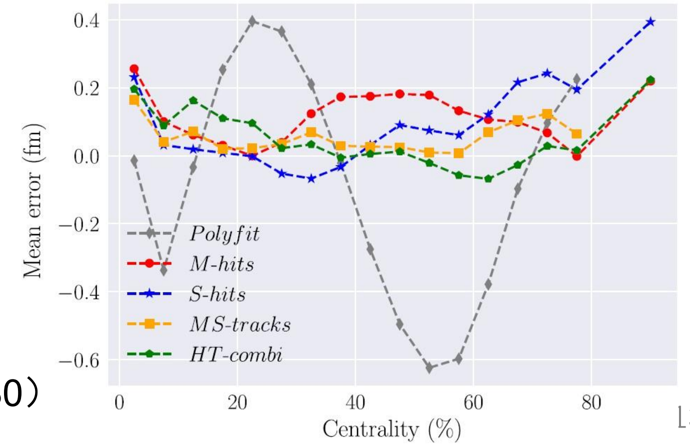
- Quantifies precision in predictions
- Polyfit fails for central events!
- Similar precision for b>3 fm

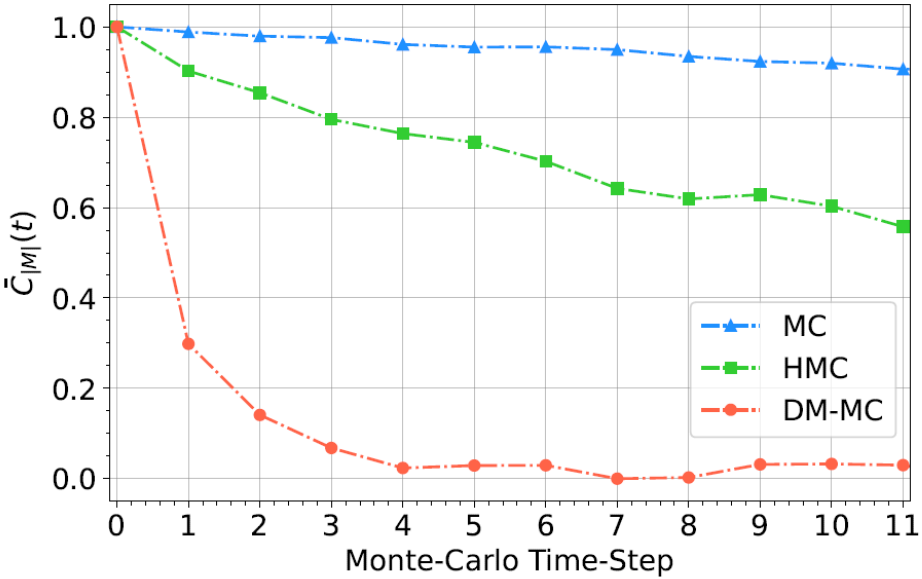


- Quantifies accuracy
- DL: -0.3 - 0.2 fm for b = 2-14fm
- Polyfit fluctuating

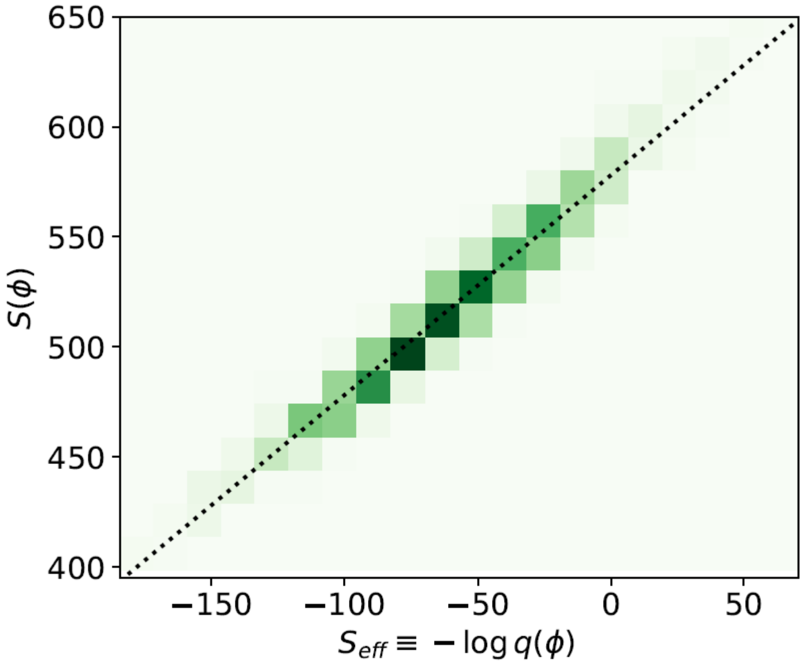


- For realistic bdb distributed testing collision events →
- Fast online event selection ~ 1k events/s on 1 game GPU（RT2080）

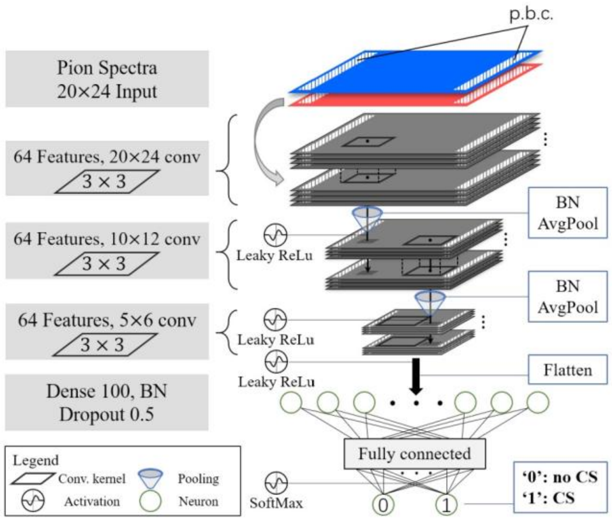


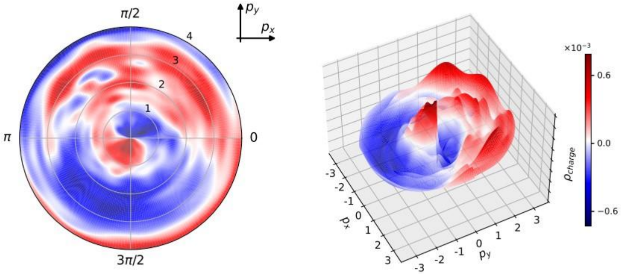


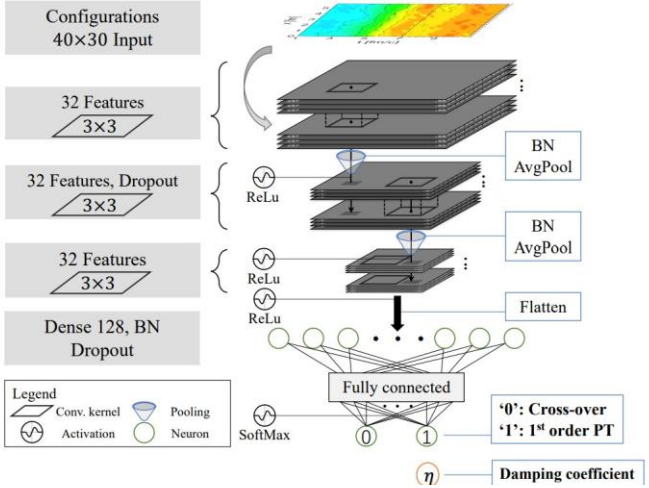
validation R2 ~0.96

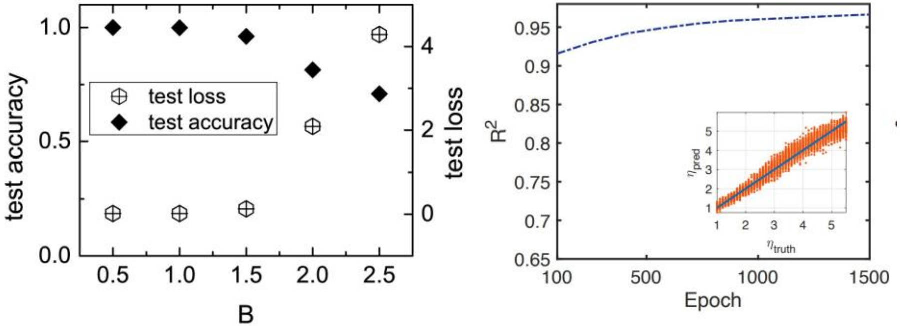


CNN to detect CME, and regress stochastic dynamics in HICs





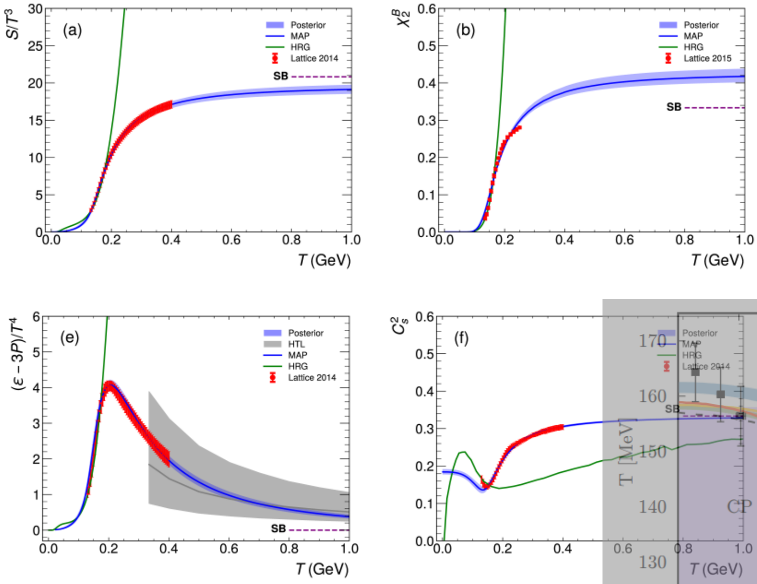




Y. Zhao, L. Wang, K. Zhou, X. Huang, PRC 106, L051901(Letter)

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

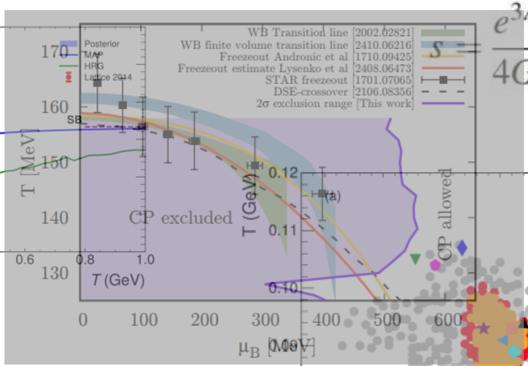
Critical Endpoint from Holographic QCD via Bayesian Inference

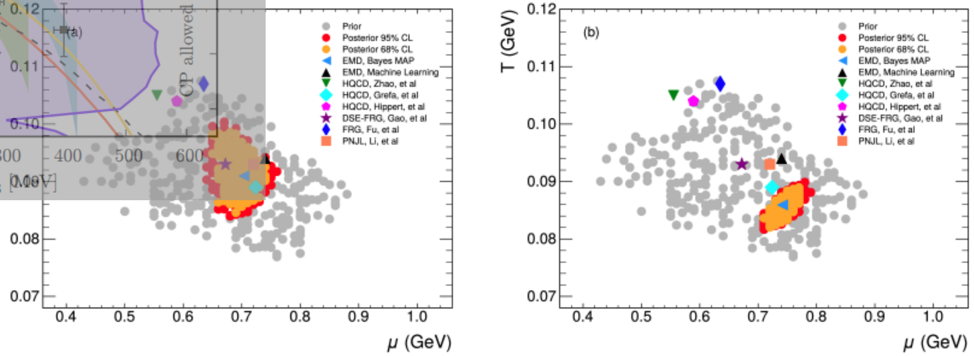


$$S_E = \frac{1}{16\pi G_5}\int d^5x\sqrt{-g}\left[R - \frac{f(\phi)}{4}F^2 - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right]$$

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

$$S = \frac{e^{3A(z_h)}}{4G_5 z_h^3}.\qquad \chi_2^B = \frac{1}{T^2}\frac{\partial\rho}{\partial\mu}.$$





L. Zhu, X. Chen, K. Zhou, H. Zhang, M, Huang,
arXiv:2501.15810