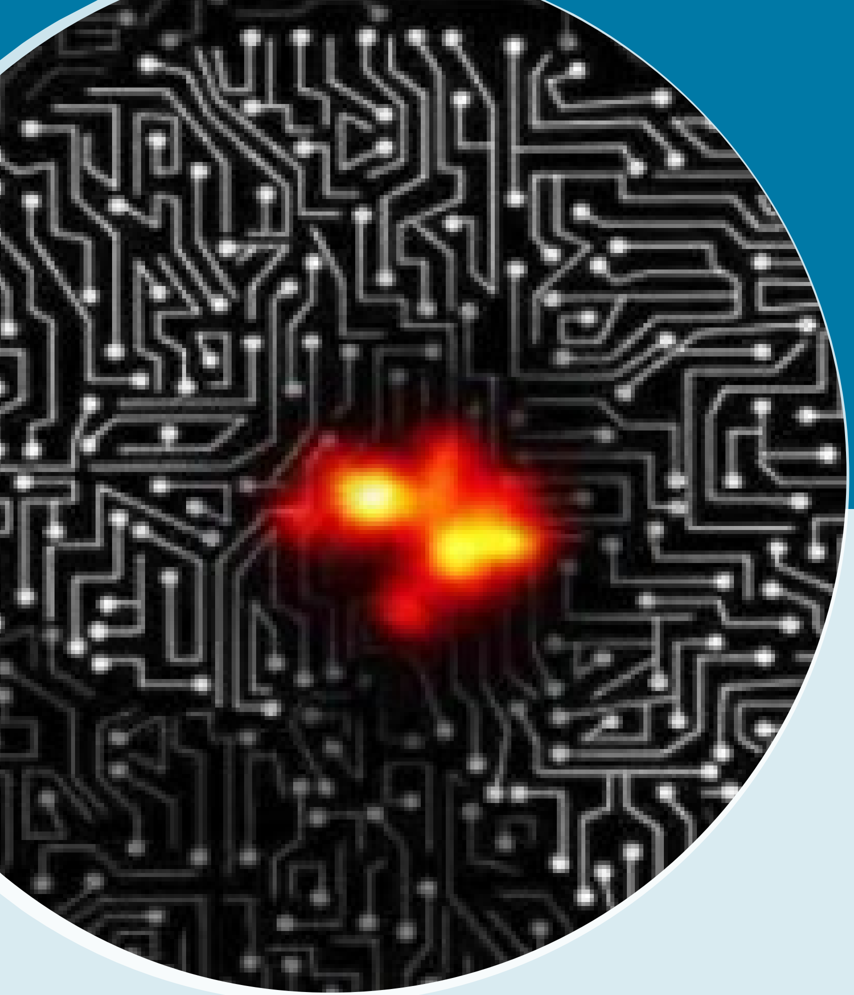


香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

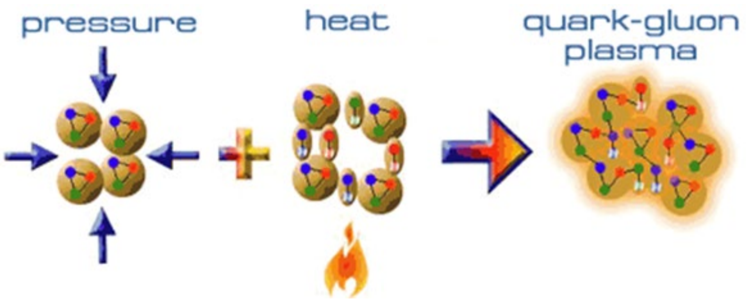
Extreme QCD Matter Exploration meets Machine Learning

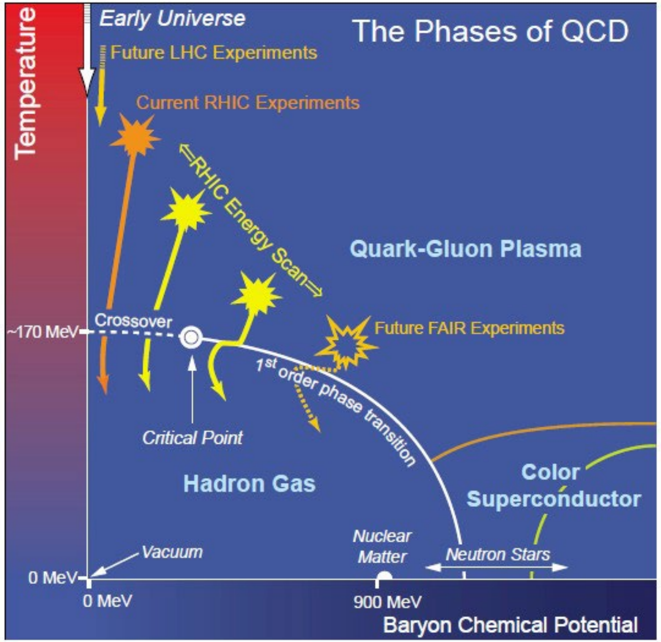
Kai Zhou (CUHK- Shenzhen)

Precision Frontier of QCD Matter:
Inference and Uncertainty Quantification
01-13 Sep, CCNU, Wuhan

Overview : Golden Age of QCD matter in extreme

- **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
- **Lattice Field Theory / fQCD / Effective models**
- **Neutron Star** : dense matter, astronomy constraints

Overview : Nuclear Physics meets Machine Learning

- **2012** : Discovery of Higgs boson


The New York Times

Physicists Find Elusive Particle Seen as Key to Universe

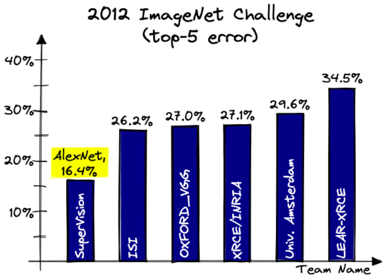
- AlexNet - Birth of Deep Learning


The New York Times

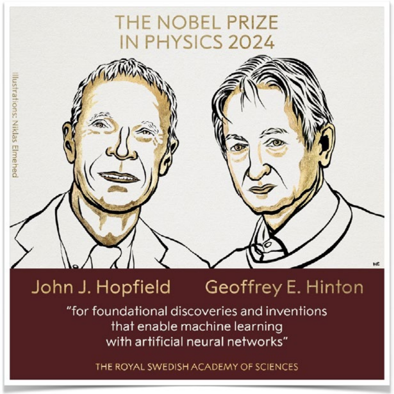
Chatbots › OpenAI Unveils GPT-4 What GPT-4 Can and Can't Do Funding Frenzy Escalates

Scientists See Promise in Deep-Learning Programs

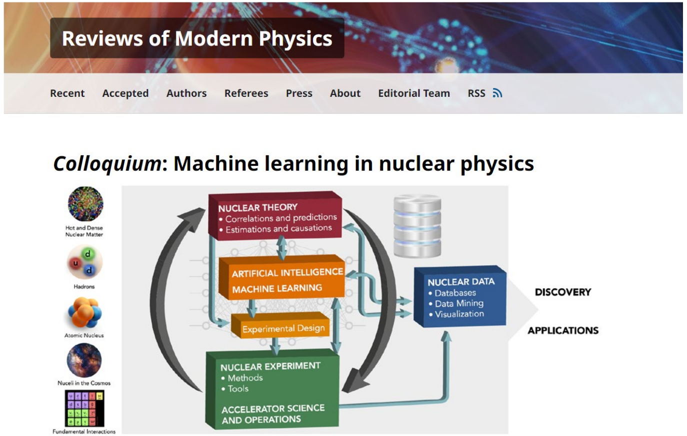
Give this article



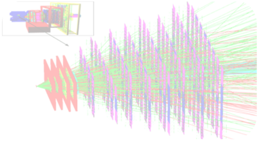
- **2024** : Nobel Prize in Physics

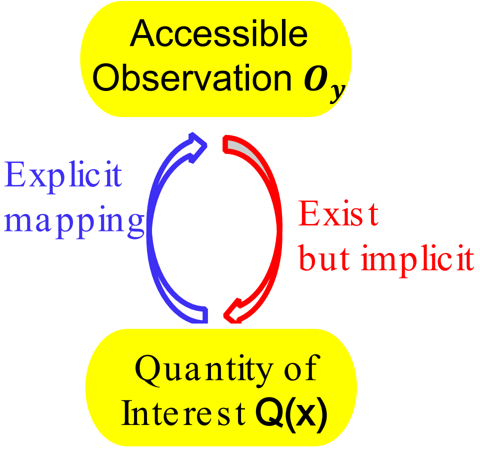


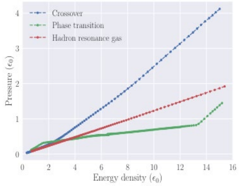
- ML4Physics, AI4Science

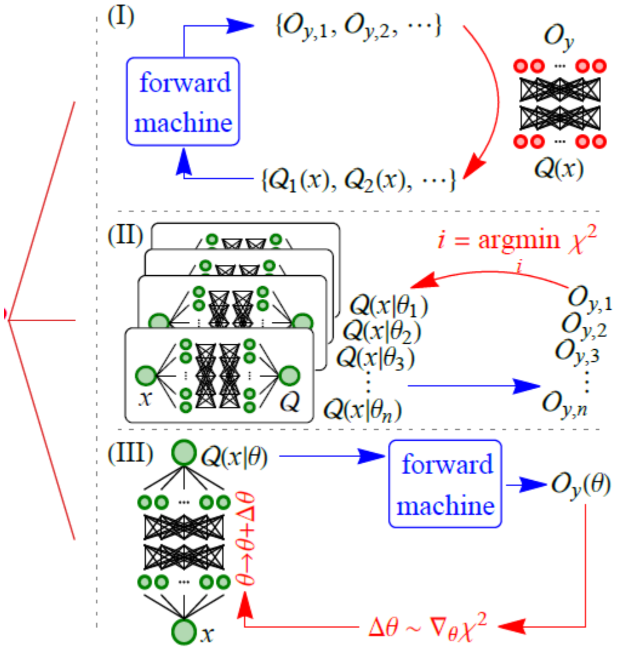


Inverse Problems Solving with ML









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

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

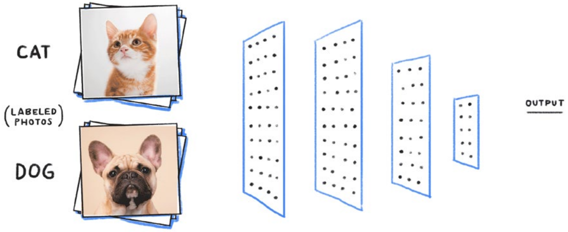
- **Automatic Differentiation :** fuse physical prior into reconstruction via differentiable programming strategy

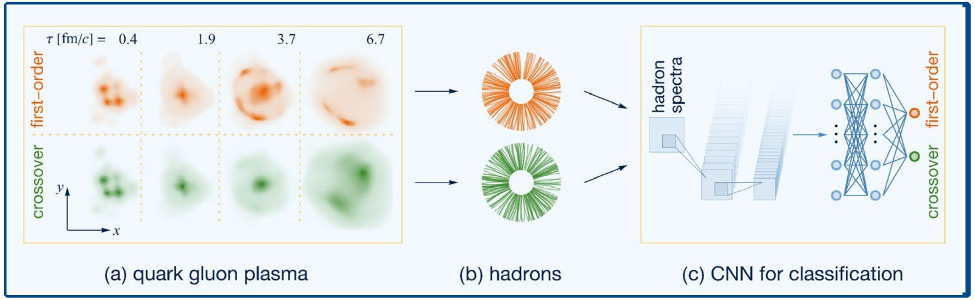
$$\frac{1}{2}\nabla_{\boldsymbol{\theta}}\chi^2 = \sum_y \frac{\mathcal{F}_y[\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})] - \mathcal{O}_y}{(\Delta\mathcal{O}_y)^2} \int \mathrm{d}x \frac{\delta\mathcal{F}_y[\mathcal{Q}(x)]}{\delta\mathcal{Q}(x)}\bigg|_{\mathcal{Q}(x)=\mathcal{Q}_{\mathrm{NN}}(x|\boldsymbol{\theta})}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

Direct inverse mapping with CNN for identifying QCD transition

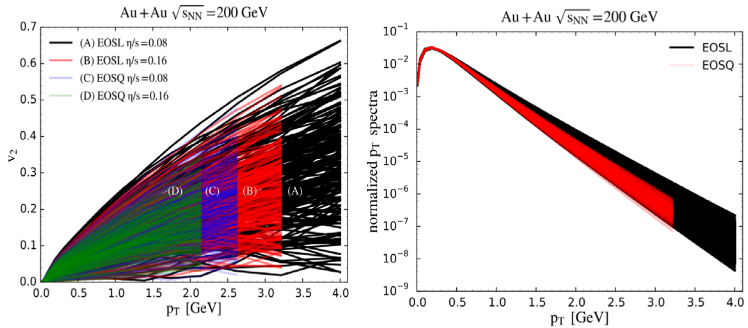
Data-driven
Inverse Mapping

Physics Simulation
provide the Prior





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

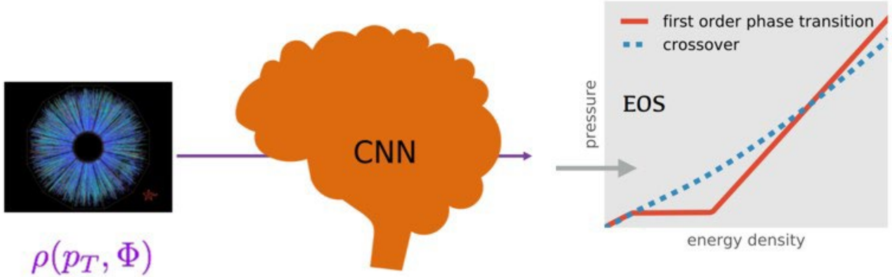


- Conventional obs. hard to distinguish
- Strongly influence from initial fluctuations and other uncertainties
- CNN : 95% event-by-event accuracy!
- Robust to initial conditions, eta/s

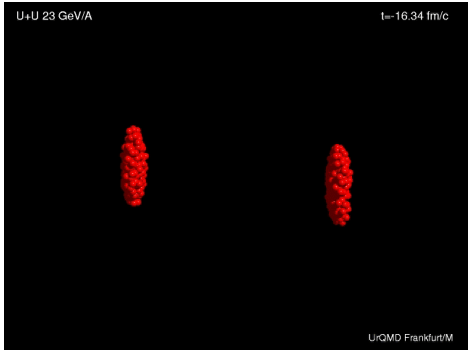
Conclusion : Information of early dynamics can **survive** to the end of hydrodynamics and encoded within the final state raw spectra, immune to evolution's uncertainties, **with deep CNN we can decode it back**.

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

Into more realistic situations

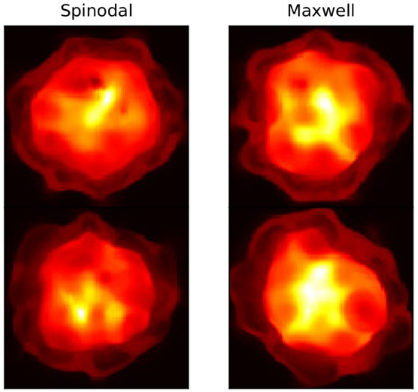


Hadronic cascade
(UrQMD considered)



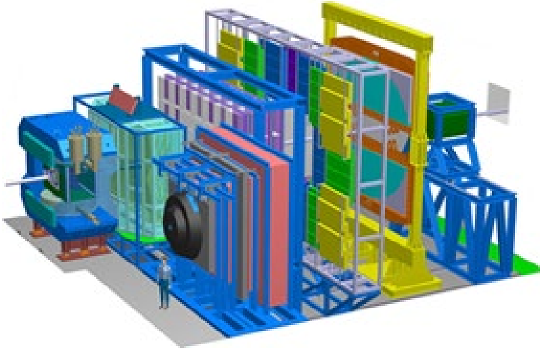
Eur. Phys. J. C 80 (2020) no.6,516

Non-equilibrium transition
(Baryon Clumping, spinodal)



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

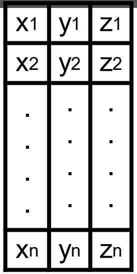
Detector effects
(Hits/Tracks, Point-Net)



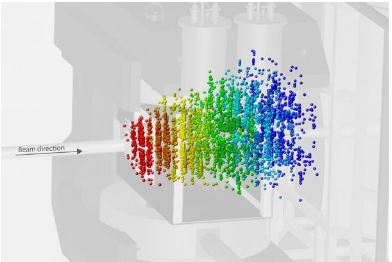
Phys. Lett. B 811, 135872
JHEP 21 (2021) 184

Point Cloud Network for Physics online analysis for HICs

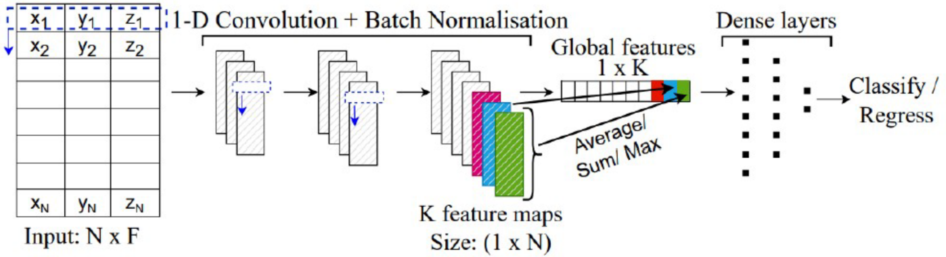
- Sensor data has inherent **point cloud structure**
  - **collection of particles** as *2D array* :



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

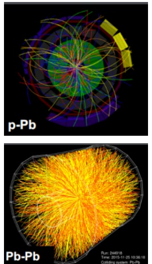


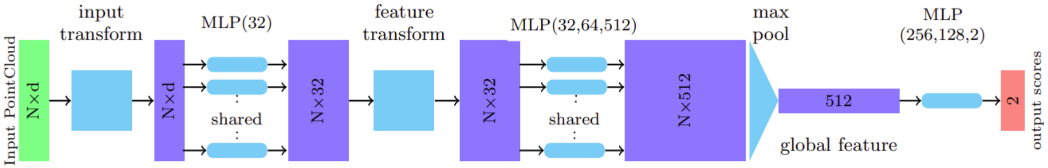
- PointNet based models learn directly from point clouds.
  - respects the **order invariance** of point clouds
  - direct process ⇒ **ideal online analysis algorithm**

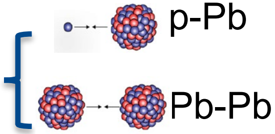


Manjunath O.K. and Kai Zhou, etc. *Phys.Lett.B* 811 (2020) 135872; JHEP10(2021)184.

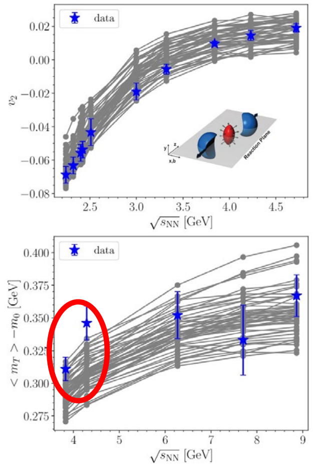
- **Collision Centrality Regression**
  M. OK, J. S, K. Zhou, H. S, **Phys.Lett.B** 811 (2020) 135872
- **EoS Classification**
  M. OK, K. Zhou, J. S, H. S, **JHEP** 10(2021) 184
- **Small/ Large-system Identification**
  S.Guo, H. Wang, K. Zhou, G. Ma, **Phy.Rev.C 110 (2024)2**

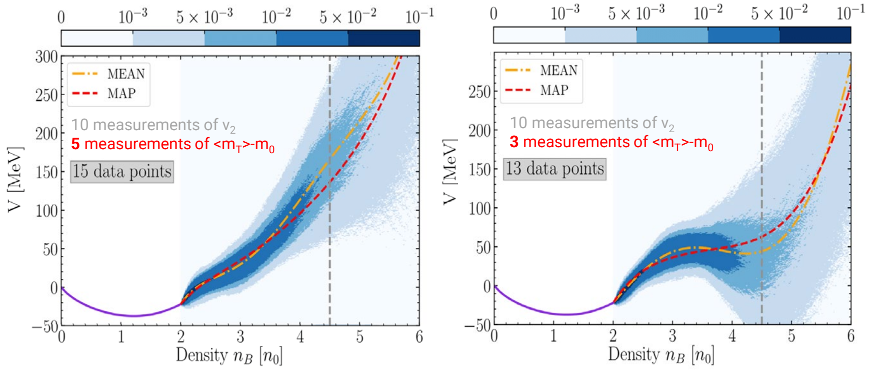






Bayesian Inference Dense Matter EoS from HIC and Holography





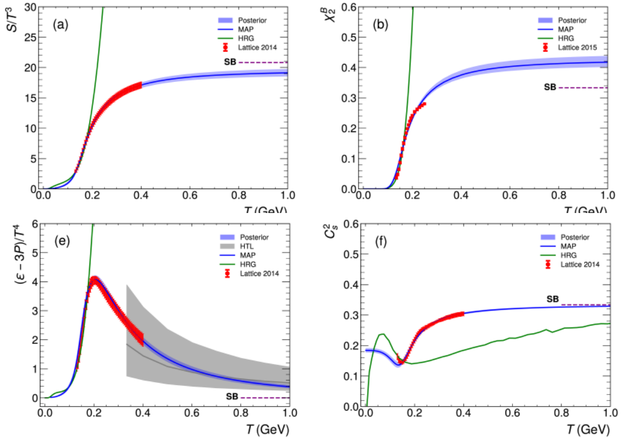
- Comprehensive Bayesian inference necessary for unambiguous solution
- Tension between data-data or model (**UrQMD**)-data
- Next-gen experiments will provide immense amount of high precision data

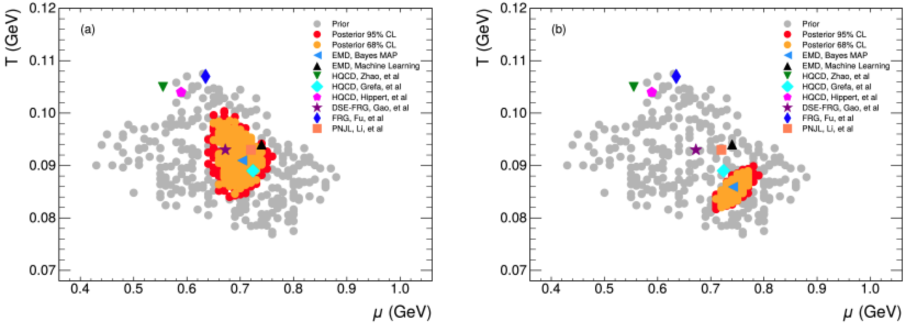
M.OK, J. Steinheimer, K. Zhou, H. Stoecker, **PRL131,202303(2023)**

See talk by Manjunath. O.K (this afternoon)

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

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



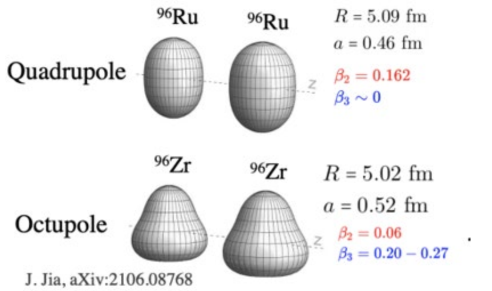


- Critical endpoint from **holography (EMD)** via Bayesian Inference

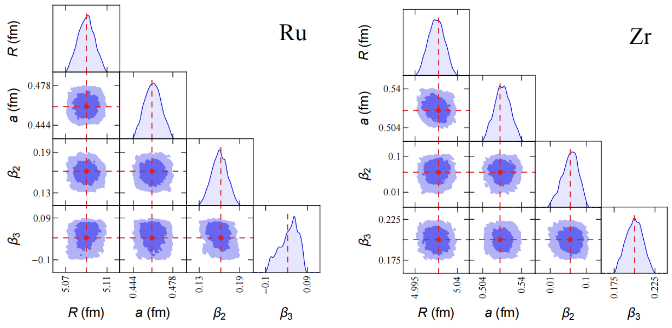
L. Zhu, X. Chen, K. Zhou, H. Zhang, M, Huang, **arXiv:2501.17763**

Bayesian Imaging for Nuclear Structure in Isobar Collisions

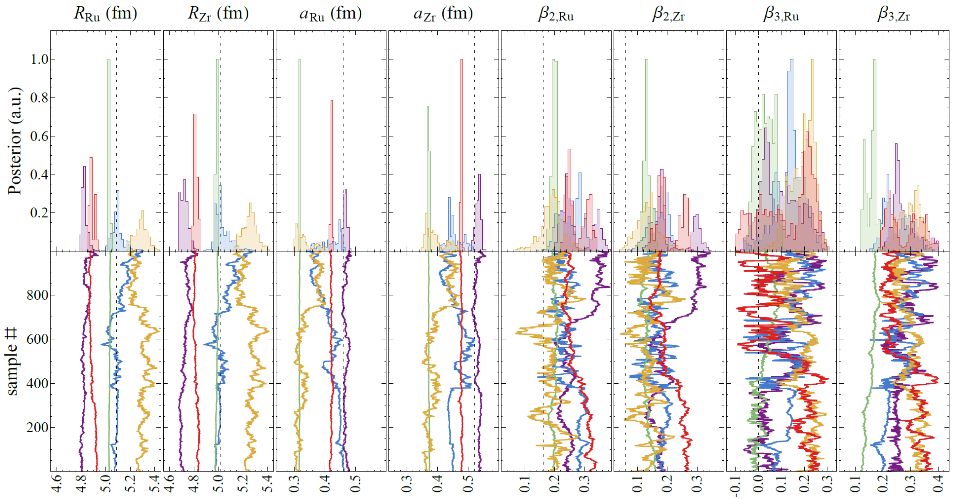
- Nuclear Structure imaging for single system ? (caveat: model dependent)
- Simultaneous inference for isobar systems with ratio?
- **Bayesian Inference:** Gaussian Process emulator + PCA dim reduction + MCMC
  Data: MC-Glauber + Matching (linear response approximation)



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



Single system works good



With purely the Isobar-Ratios:

MCMC can not converge to a stationary inference ofthe nuclear structure

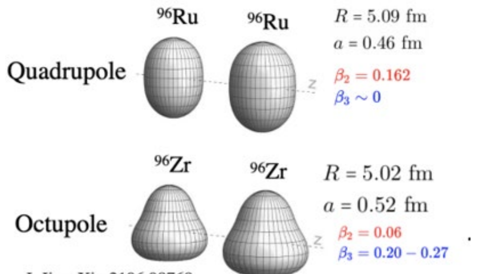
$$\boldsymbol{y}_{r,1} \equiv \{R_{P,a}, R_{\varepsilon_2,a}, R_{\varepsilon_3,a}, R_{d_\perp,a}\}_{a=1,\cdots,40}$$

- No unique solution : only ratios $\nRightarrow$ *nuclear structure !*

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

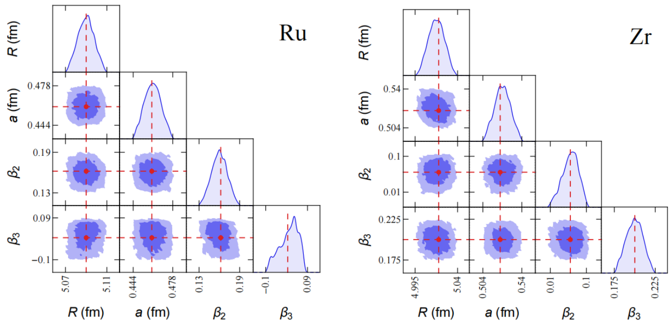
Bayesian Imaging for Nuclear Structure in Isobar Collisions

- 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)

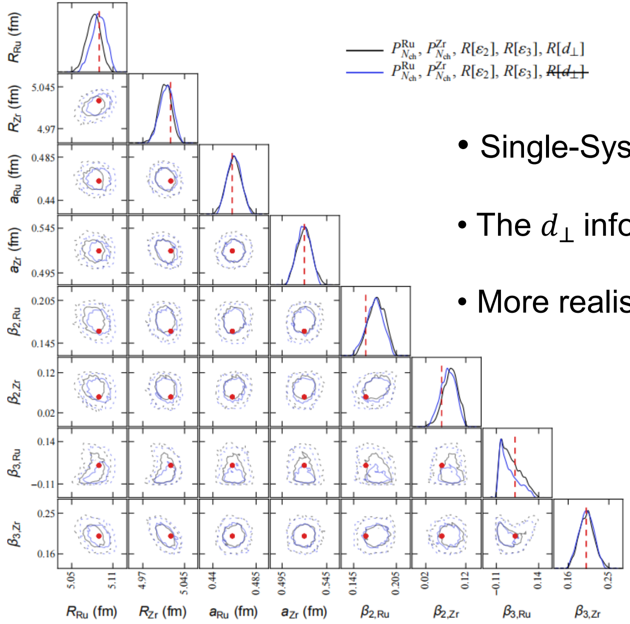


J. Jia, aXiv:2106.08768

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



Single system works good



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

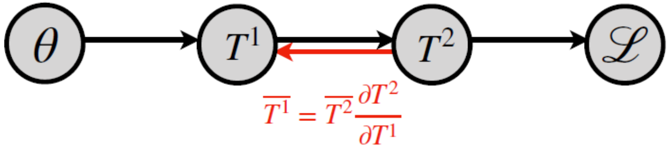
$$\boldsymbol{y}_{\mathrm{r},2} \equiv \{P_a^{\mathrm{Ru}}, P_a^{\mathrm{Zr}}, R_{\varepsilon_2,a}, R_{\varepsilon_3,a}, R_{d_\perp,a}\}_{a=1,\cdots,40}$$

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

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

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

(a)

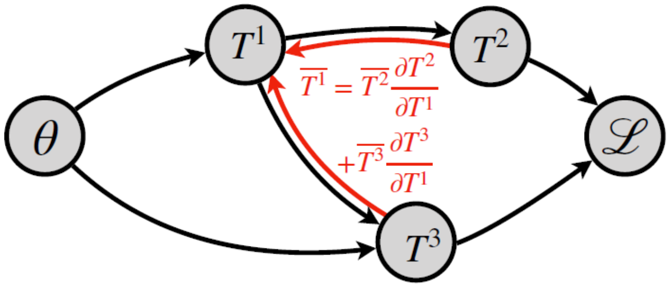


Defining adjoint variables : $\overline{T} = \partial \mathcal{L} / \partial T$

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

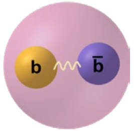
$$\overline{\theta} = \overline{T^1} \frac{\partial T^1}{\partial \theta}$$

(b)



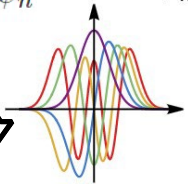
$$\overline{T^i} = \sum_{j:\,\text{child of } i} \overline{T^j} \frac{\partial T^j}{\partial T^i}$$

HQ Potential Model, Inverse Shroedinger Eq.

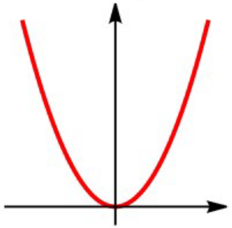


$$\hat{H}\psi_n = -\frac{\nabla^2}{2m_\mu}\psi_n + V(r)\psi_n = E_n\psi_n$$

$\{\psi_n(r)\}$

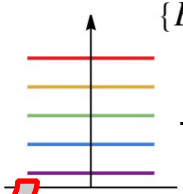


$V(r)$



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

$\{E_n\}$

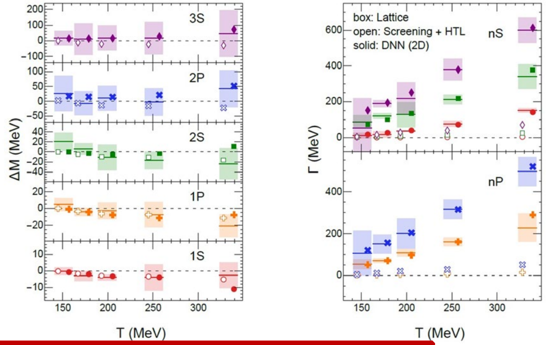


$$\begin{cases} \mathrm{Re}[E_n] &= m - 2m_b \\ \mathrm{Im}[E_n] &= -\Gamma \end{cases}$$

?

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

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

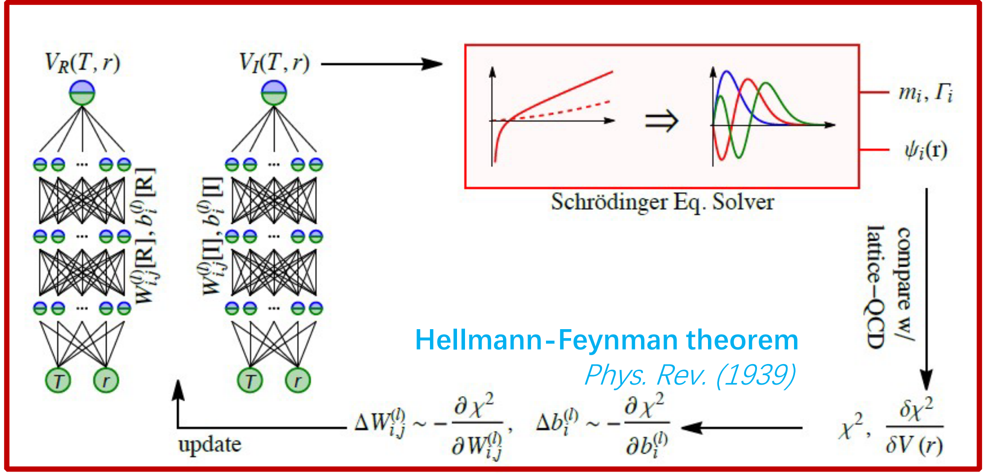


How to extract effective potential given limited spectroscopy ? →

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

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

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



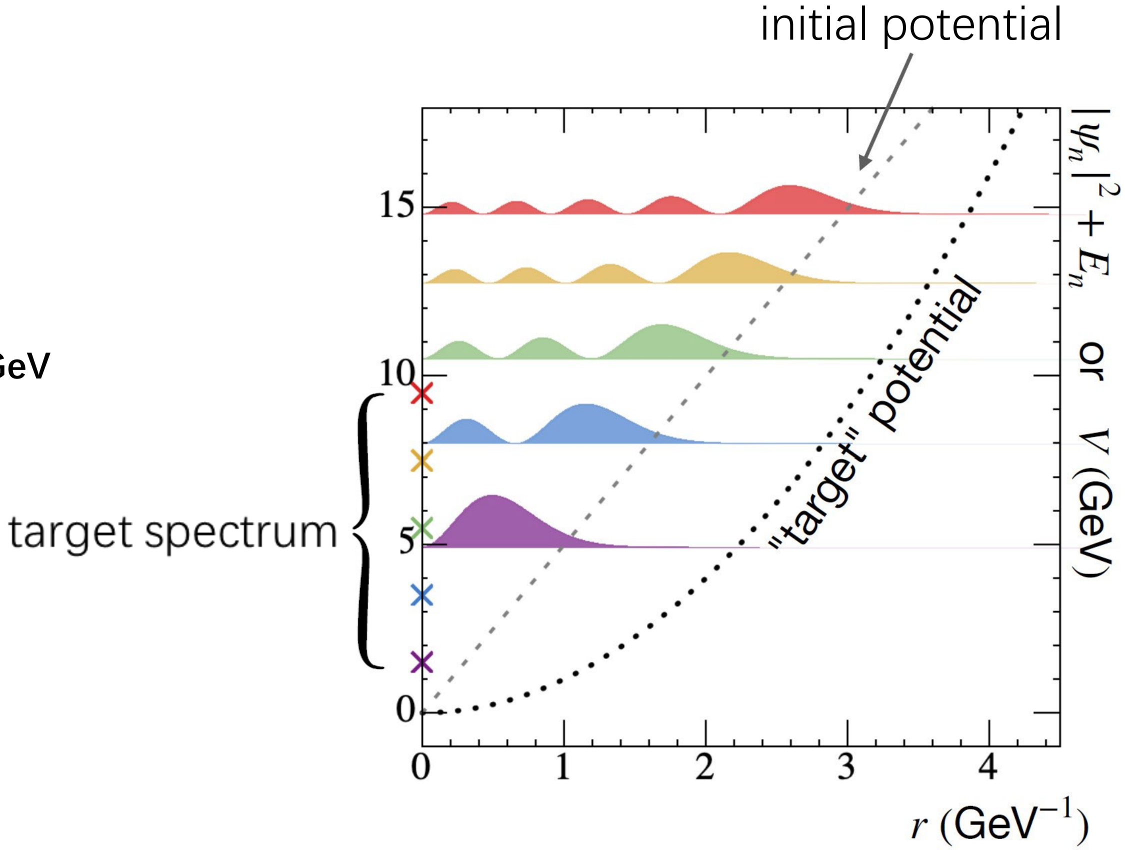
S.S, K. Z, J.Z, S.M., P. Z, Phys. Rev. D 105 (2022) 1, 1

Proof of Concept

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



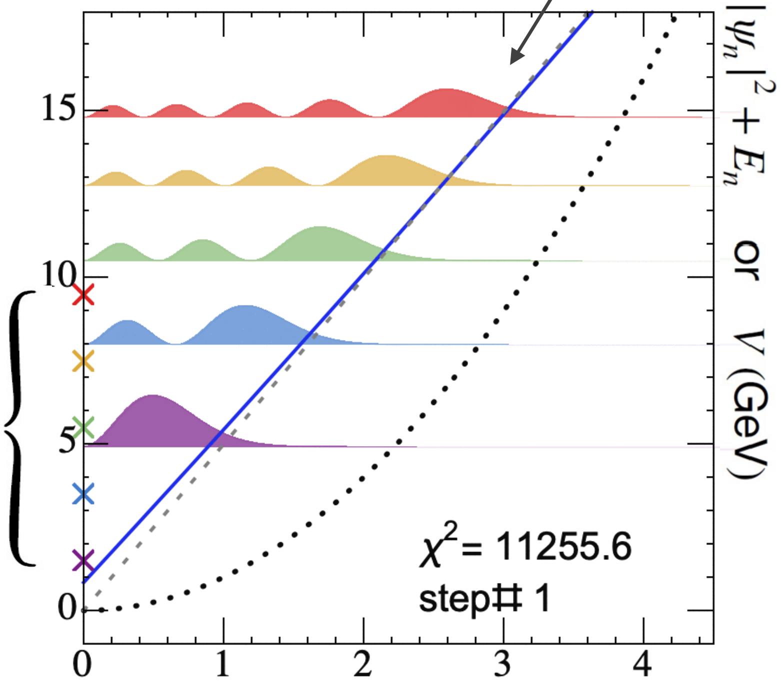
limited spectrum { En } to continuous interaction V(r) ?

Learn V(r) from 5 eigenvalues :

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

target spectrum

initial potential



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

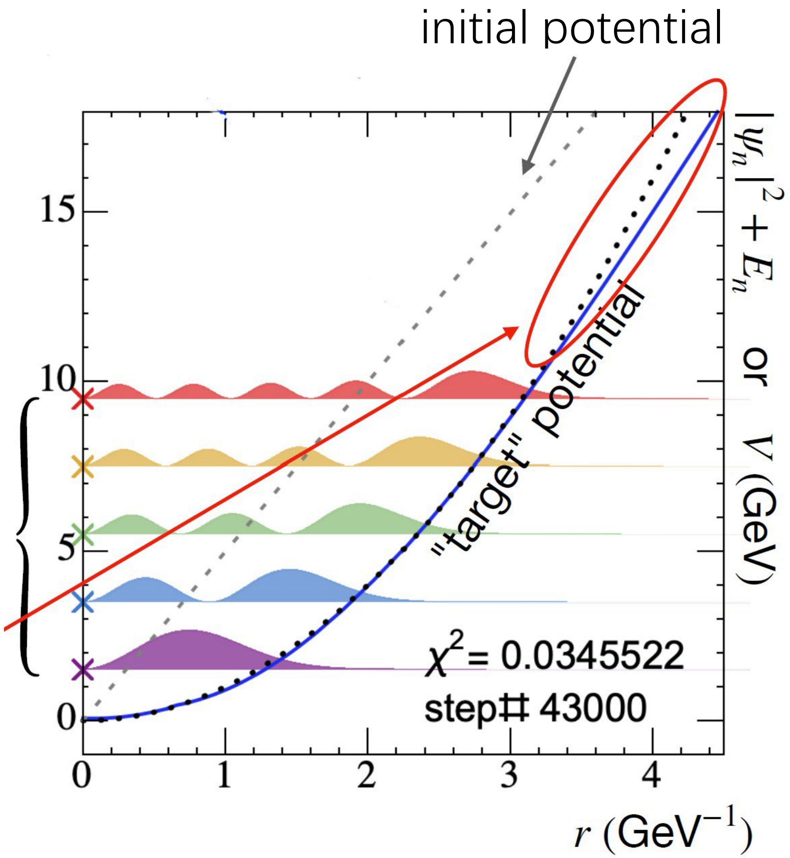
-- Yes! But to some range decided by the used states.

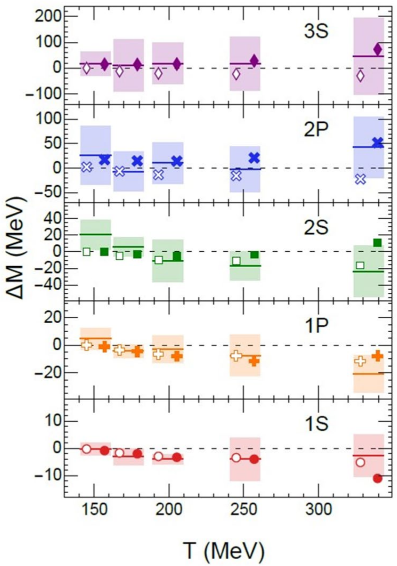
Learn V(r) from 5 eigenvalues :

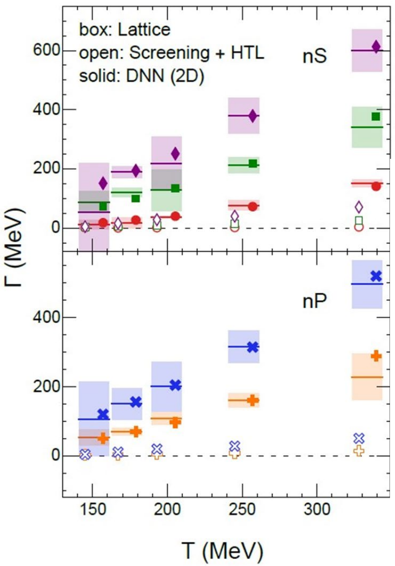
{ En } = {3/2, 7/2, 11/2, 15/12, 19/2} GeV

Deviation @ given states' wavefunction vanishes

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

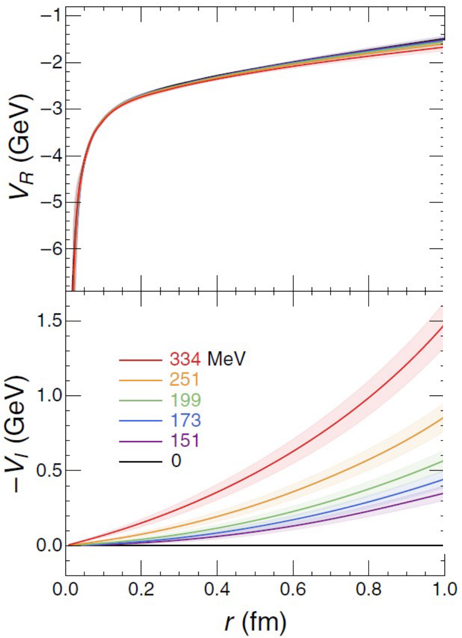


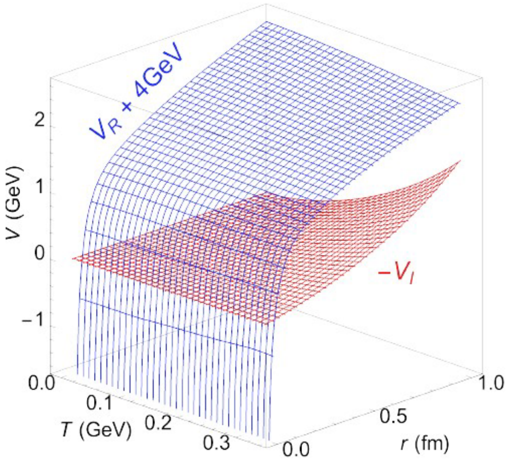




Chi2-per-data=16.5/30

The reconstructed T, r dependent potential



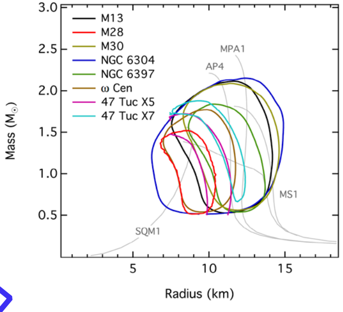


S.S, K. Z, J.Z, S.M., P. Z, Phys. Rev. D 105 (2022) 1, 1

From NS Stellar Structure (MR) to Interior EoS - AutoDiff

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

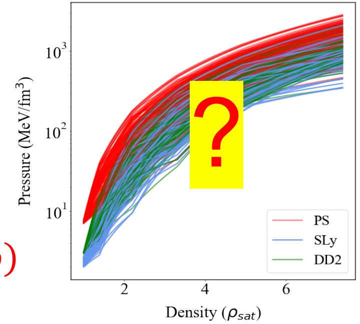
Noisy/Limited NS Observables to EoS ?



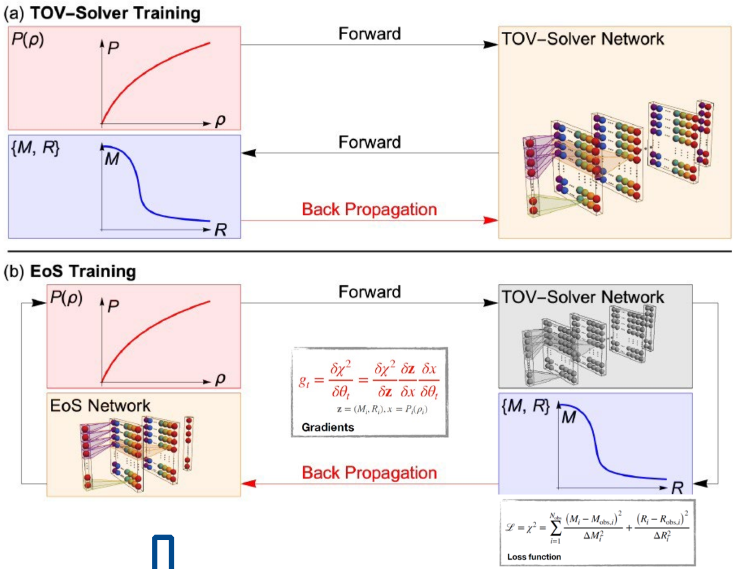
TOV eqs

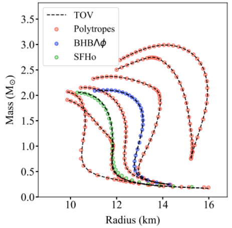
$$\frac{dP}{dr} = -\frac{G}{r^2}\left(\rho + \frac{P}{c^2}\right)\left(m + 4\pi r^3 \frac{P}{c^2}\right)\left(1 - \frac{2Gm}{c^2 r}\right)^{-1}$$

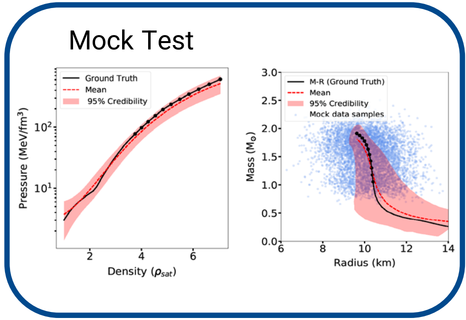
$$M = m(R) = \int_0^R 4\pi r^2 \rho \, dr$$

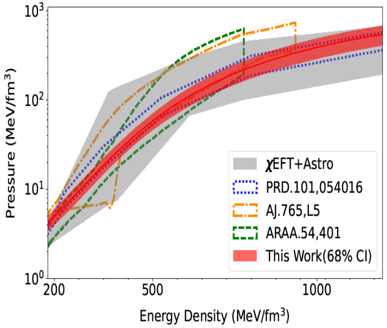


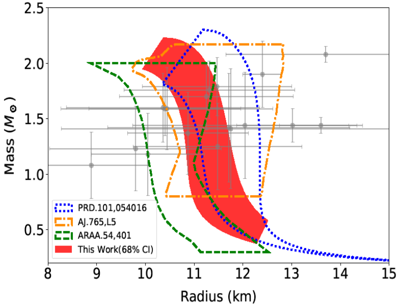
$P(\rho)$





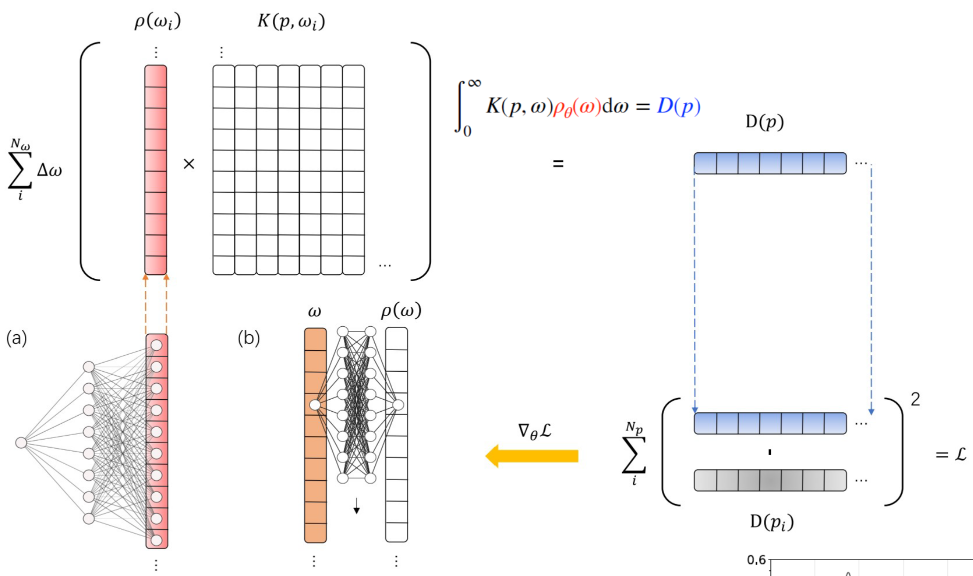


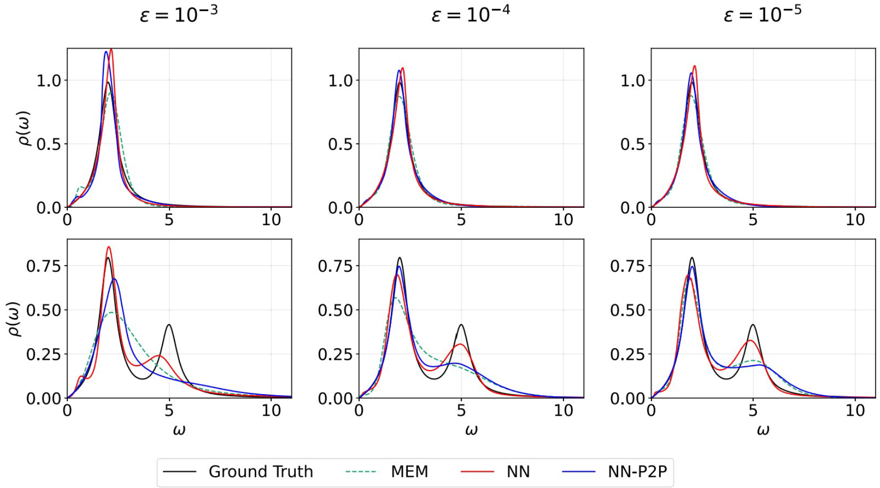


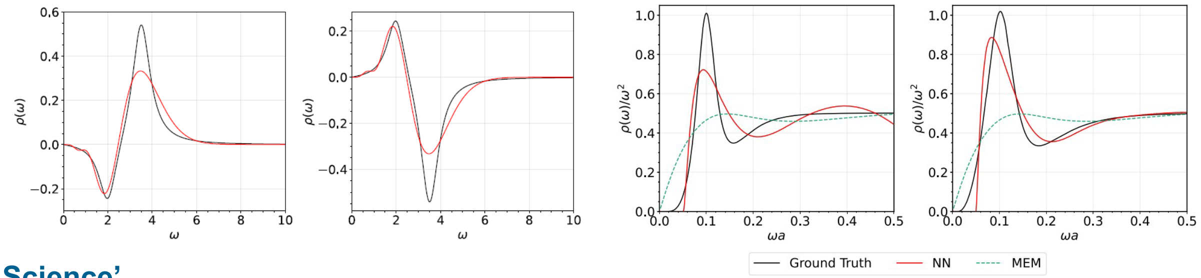


S. S, L. W, S. S, H. S, K. Z, **JCAP** 98(2022)071;**PRD**107(2023)083028

Spectral function reconstruction from Euclidean correlator via AD







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

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

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,

Discriminative / Generative model

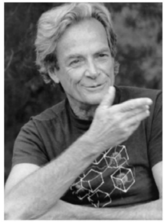
- Discriminative Learning : **prediction**

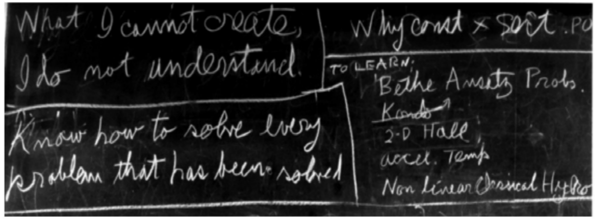
  function fitting $y = f(x)$

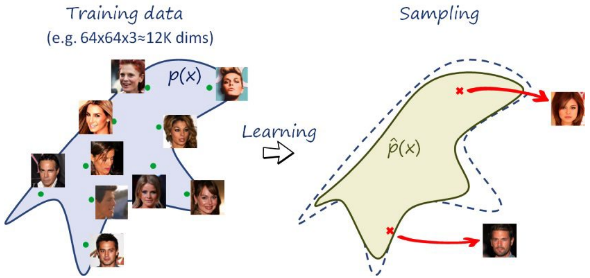
  conditional probability $p_\theta(y|x) \rightarrow p(y|x)$

- Generative Modelling : **understand**

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







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

Variational Free Energy Learning with autoregressive generative model

- Variational free energy minimization - Reverse KL divergence

$$D_{\mathrm{KL}}(q_\theta \,\|\, p) = \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \ln \left(\frac{q_\theta(\mathbf{s})}{p(\mathbf{s})} \right) = \beta(F_q - F) \qquad F_q = \frac{1}{\beta} \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \left[\beta E(\mathbf{s}) + \ln q_\theta(\mathbf{s}) \right]$$

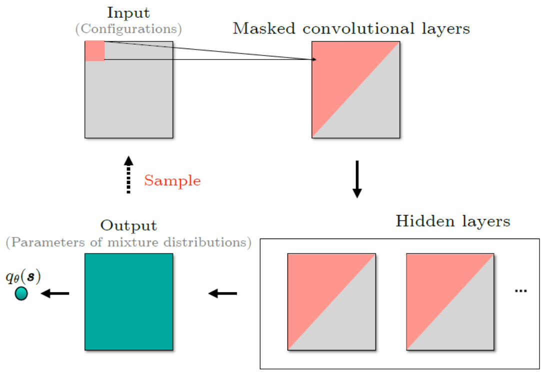
$$p(\mathbf{s}) = \frac{e^{-\beta E(\mathbf{s})}}{Z}$$

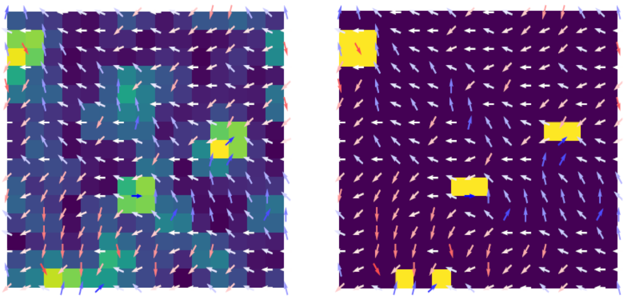
- **Autoregressive** $q_\theta(\mathbf{s}) = \prod_{i=1}^{N} q_\theta(s_i \mid s_1, \ldots, s_{i-1})$

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)





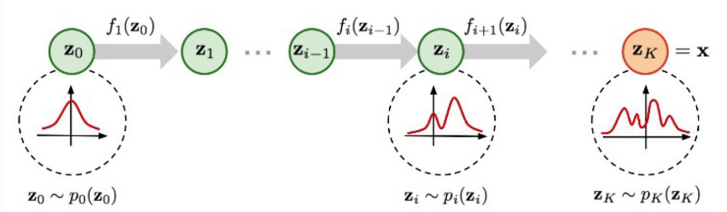
Probability distributions from CANs

Vortices

Flow based generative model to QFT

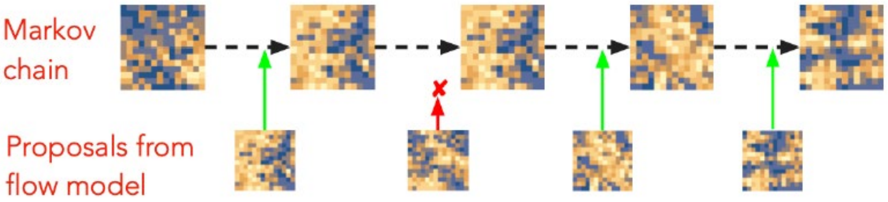
A series **(Flow)** of invertible/bijective transformations for p(z)

compose several invertible transformations to form the flow :



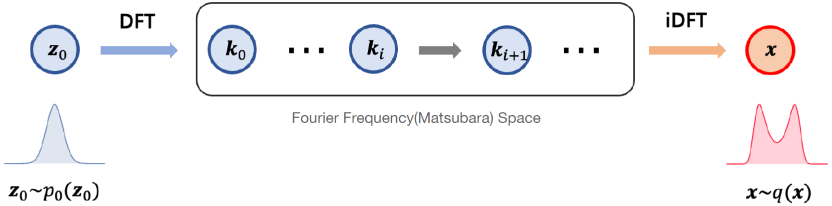
$$p_i(\mathbf{z}_i) = p_{i-1}(f_i^{-1}(\mathbf{z}_i))|\det J_{f_i^{-1}}| = p_{i-1}(\mathbf{z}_{i-1})|\det J_{f_i}|^{-1}$$

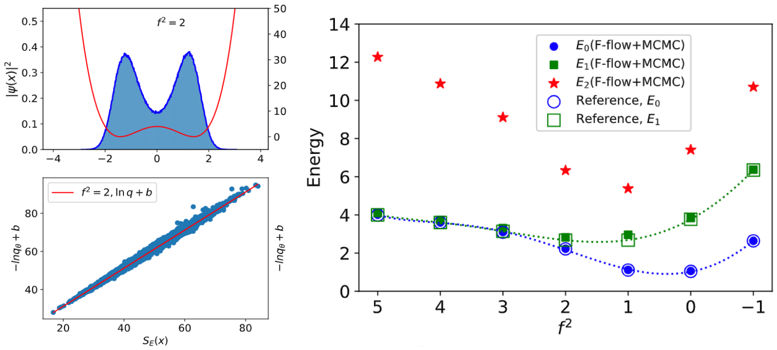
$$\rightarrow \log p(\mathbf{x}) = \log p_0(f^{-1}(\mathbf{x})) + \sum_{i=1}^{K} \log|\det J_{f_i^{-1}}| = \log p_0(\mathbf{z}_0) - \sum_{i=1}^{K} \log|\det J_{f_i}|$$



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



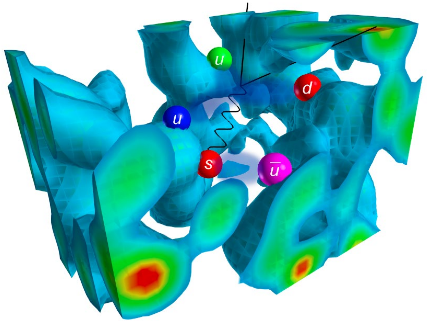


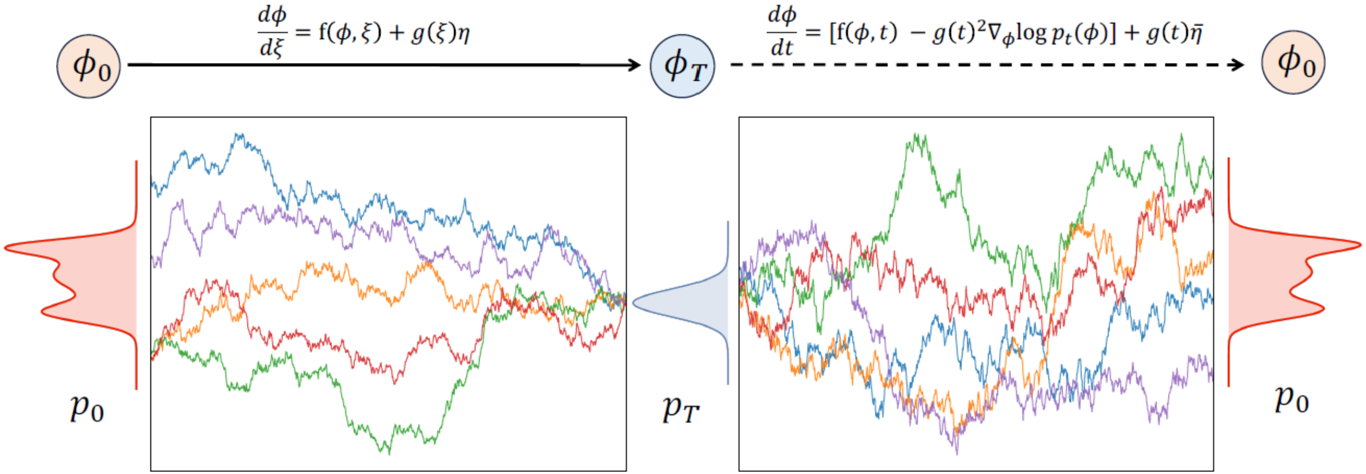
S.Chen, O. Savchuk, S. Zheng, B. Chen, H. Stoecker, L. Wang, **K. Zhou**, **PRD107, 056001(2023)**

Diffusion Model on lattice QFT configurations

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

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





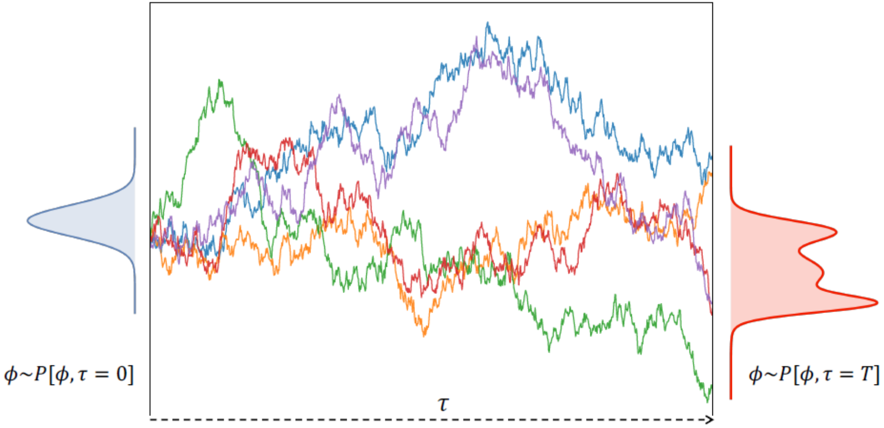
L. Wang, G. Arts, K. Zhou, JHEP 05 (2024) 060
L. Wang, G. Arts, K. Zhou, arXiv:2311.03578 (NeurIPS 2023 workshop “ML&Physical Sciences”)
G. A, D. E. H, L. W, K. Z, arXiv:2410:21212 (NeurIPS 2024 workshop “ML&Physical Sciences) “**Best Physics for AI Paper” Award**
Q. Zhu, G. Aarts, W. Wang, K. Zhou,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

Stochastic Quantization

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

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

τ: fictitious time, α: diffusion constant



- **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),

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

Thermal equilibrium limit → Quantum distribution

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

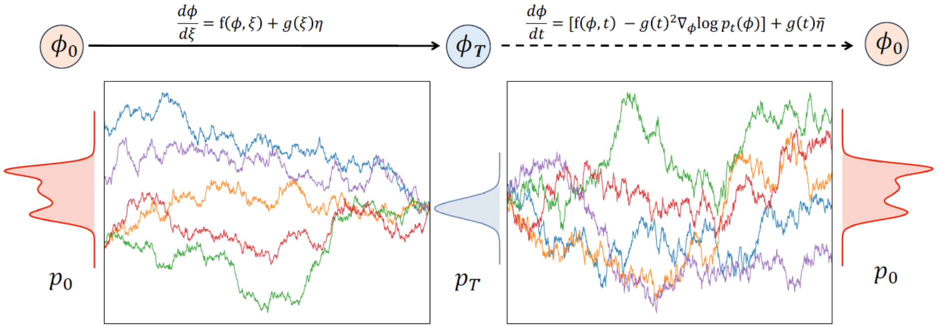
$$P_{\text{eq}}[\phi] \sim e^{-\frac{1}{\hbar} S_E[\phi]} = P_{\text{quantum}}[\phi]$$

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

- **DM generation SDE and Stochastic Quantization**

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

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



- **Similarities and differences:**

$p_{\tau=T}(\phi) \rightarrow P[\phi, T]$

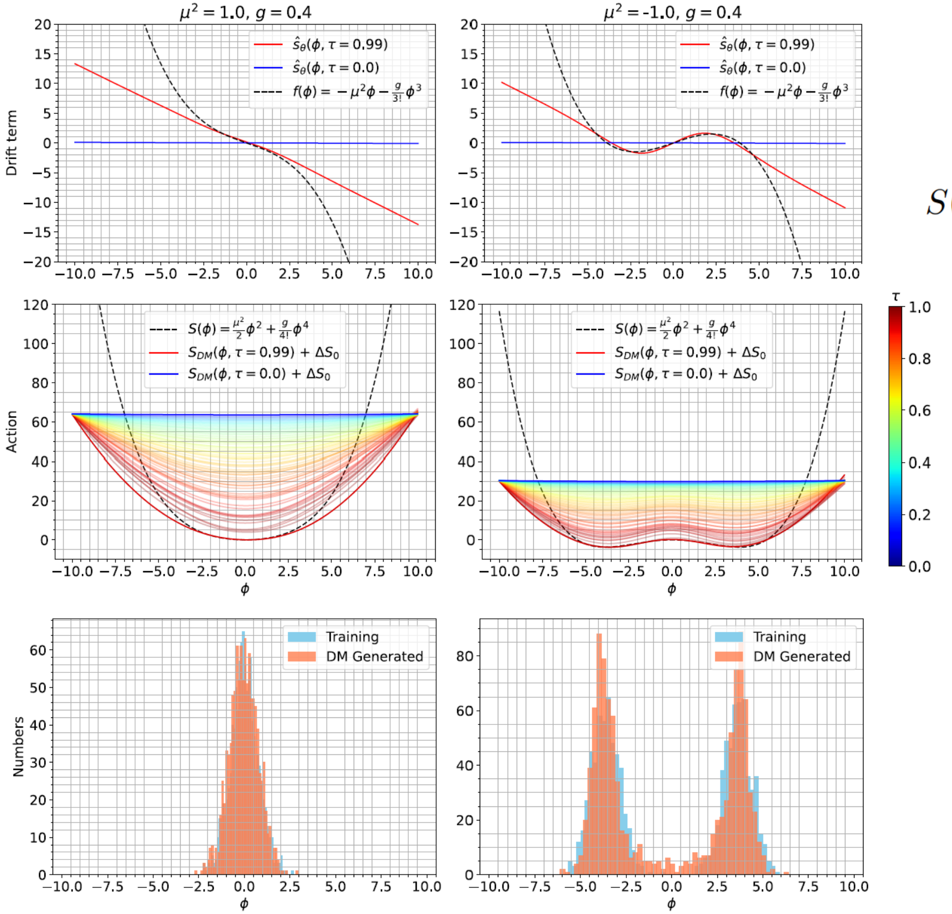
$O(\bar{\alpha}) \sim O(\hbar)$

✓ 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

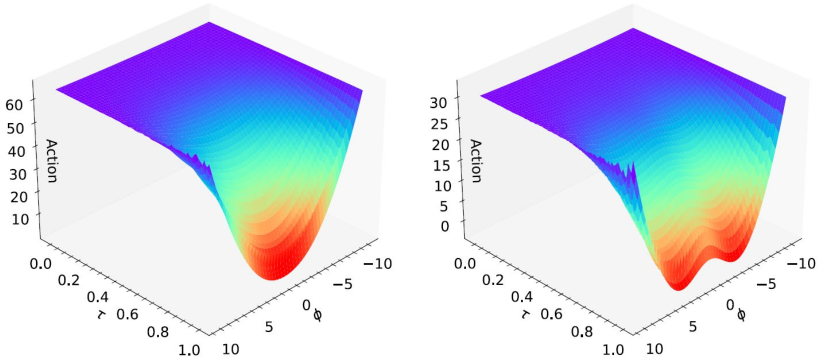
L. Wang, G. Arts, K. Zhou, JHEP 05(2024) 060

Effective Action on A Toy model



- Flow of the effective action

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



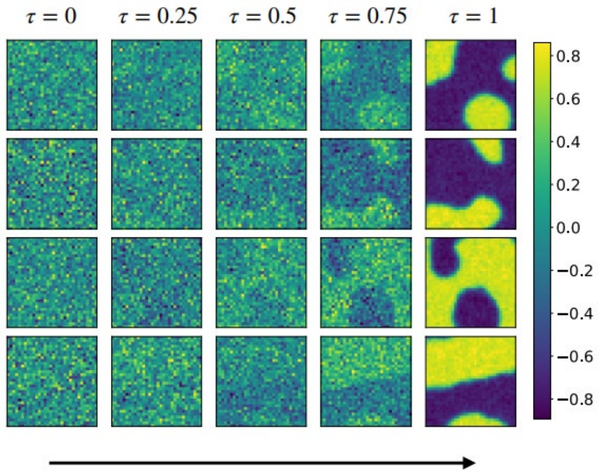
$$S_{\rm DM}(\phi, \tau) = \int^{\phi} \hat{s}_\theta(\tilde{\phi}, \tau) d\tilde{\phi}$$

DM on 2d scalar ϕ^4 model

- 32x32 lattice, HMC generated 5120 configurations for training

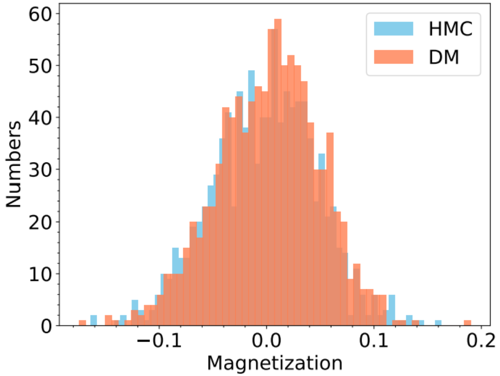
$$S_E = \sum_x [-2\kappa \sum_{\mu=1}^{d} \phi(x)\phi(x+\hat{\mu}) + (1-2\lambda)\phi(x)^2 + \lambda\phi(x)^4].$$

Broken phase :



"bulk" patterns emerge from DM

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

Further progresses on gauge fields and with more efficient generation recently

Relation to (inverse) RG

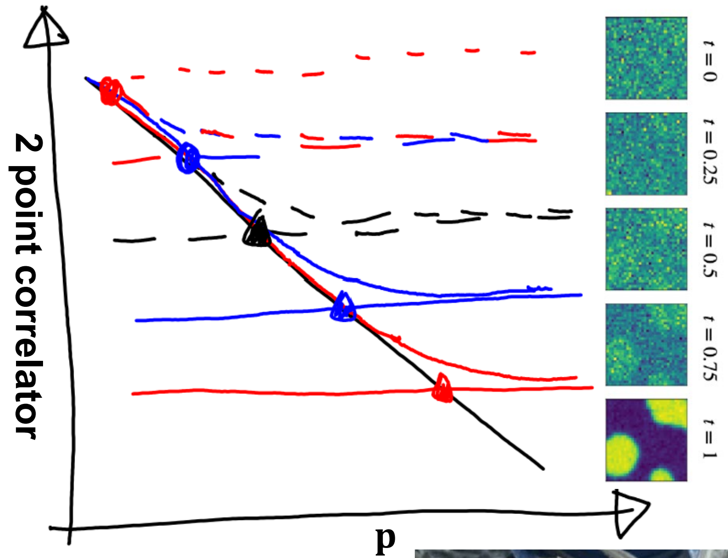
- Forward diffusion kernel: <u>**gaussian smoothing**</u>

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

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

- 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 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 !

Initial state + Bulk matter + Generative model

Fig: arXiv:1804.04649

Initial Stage

$$\nabla_\mu T^{\mu\nu} = 0$$
$$\nabla_\mu N^\mu = 0$$

$$p^\mu \partial_\mu f + F \cdot \partial_p f = C$$

Final State

Impact parameter
Nuclear structure

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

Event Generation
Final observables

Applications to Lattice QFT

U-net Emulator for relativistic hydrodynamics

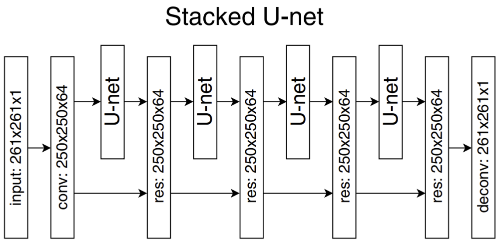
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

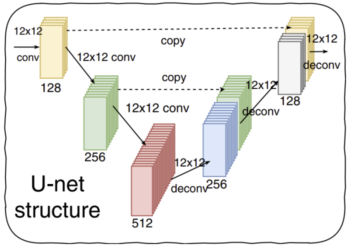
PHYSICAL REVIEW RESEARCH **3**, 023256 (2021)

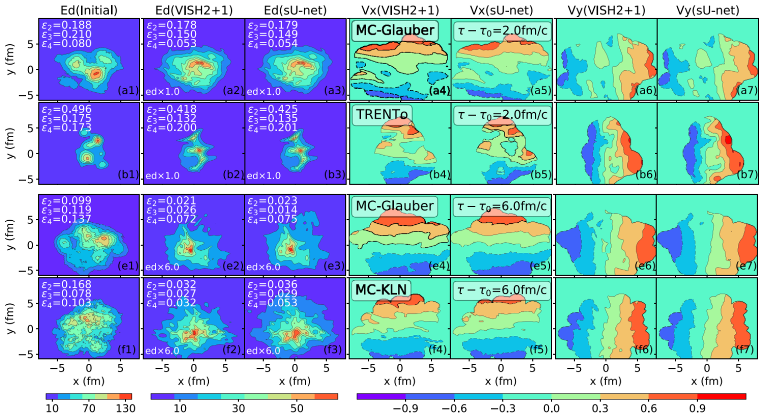
Applications of deep learning to relativistic hydrodynamics

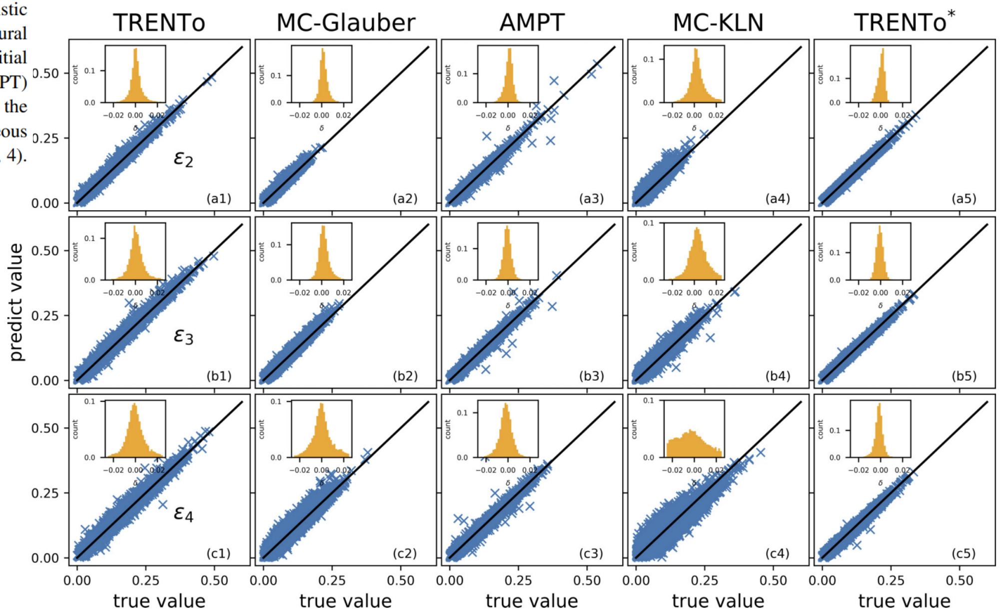
Hengfeng Huang,[1,2] Bowen Xiao,[3] Ziming Liu,[1] Zeming Wu,[1,2] Yadong Mu,[3,4] and Huichao Song[1,2,5]

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

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

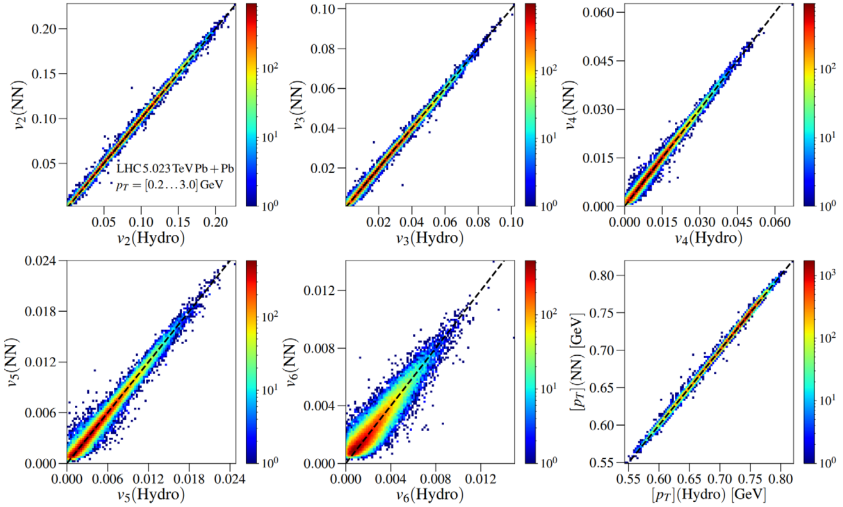
PHYSICAL REVIEW C **108**, 034905 (2023)

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.

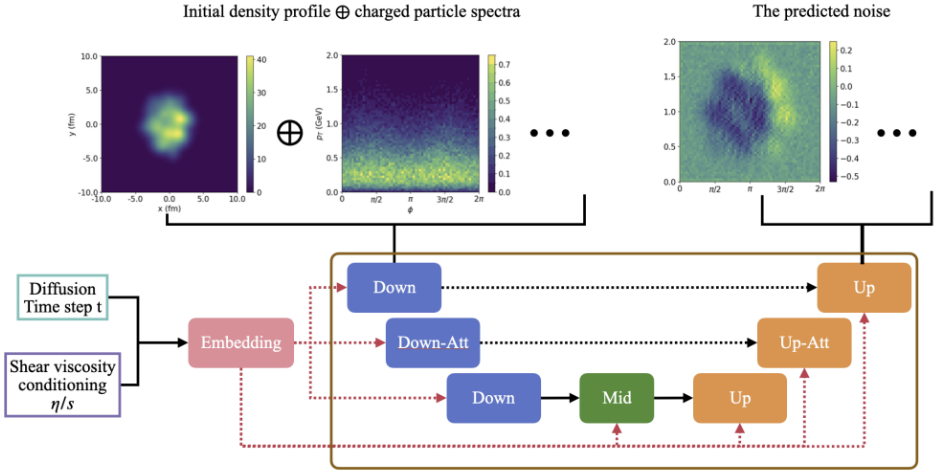
Similar study for intermediate energies with IBUU +DNN emulation see: **B. Li, etc., arXiv:2406.18421**

Generative diffusion model to heavy-ion collisions

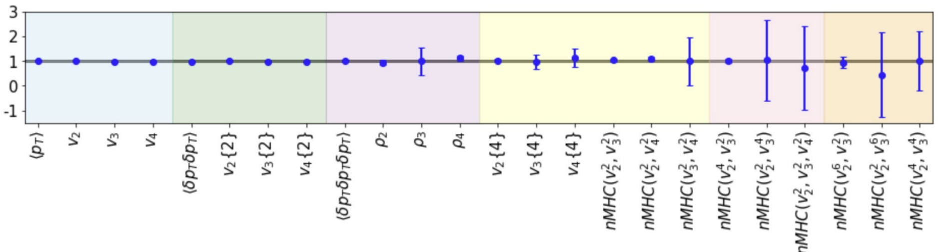
An end-to-end generative diffusion model for heavy-ion collisions

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

arXiv:2410.13069



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



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.

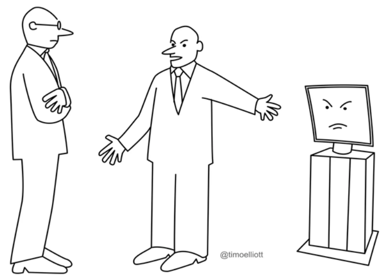
Considering that the spectra $\boldsymbol{S}_0$ depend on the initial entropy density profiles $\boldsymbol{I}$ and the shear viscosity η/s, we train a conditional reverse diffusion process $p(\boldsymbol{S}_0|\boldsymbol{I}, \eta/s)$ without modifying the forward process.

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.

Point Cloud Diffusion Model for HICs – AI clone of simulation

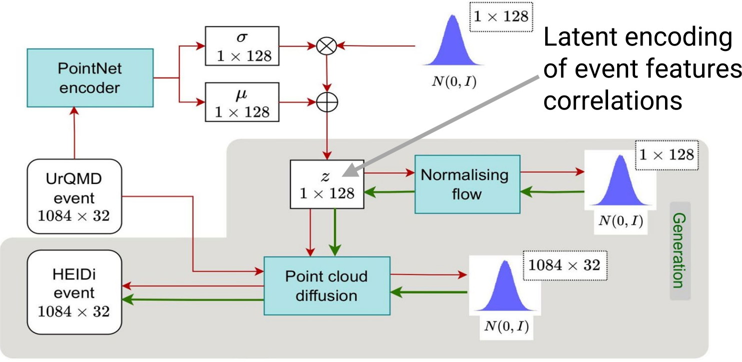
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

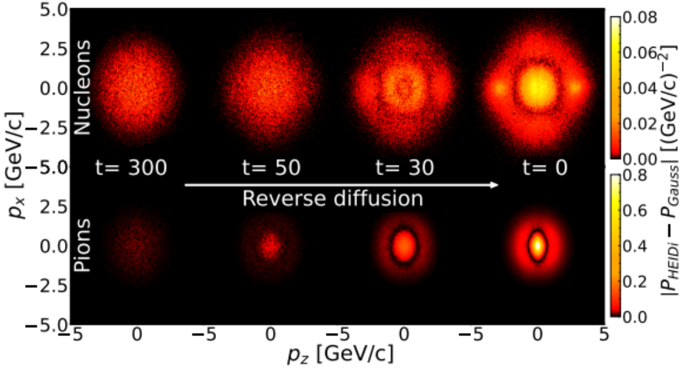


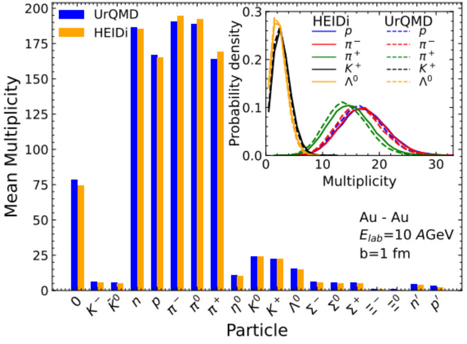
His decisions aren't any better than yours — but they're WAY faster...

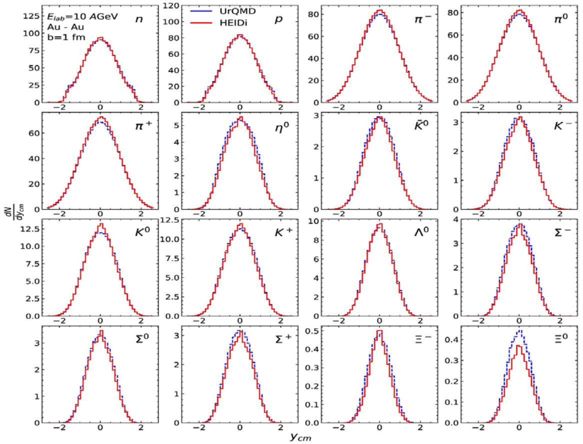
- 18k UrQMD simulation events for central Au-Au@10 AGeV collisions
- HEIDi:
  Heavy-ion Events through Intelligent Diffusion

PointNet encoder + Normalizing flow decoder + Pointcloud diffusion →





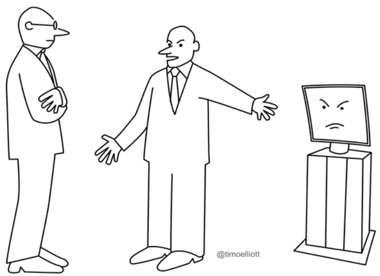




See talk by Manjunath. O.K (this afternoon)

M. O.K, K. Z, J. S, H. S, PRC2025, arXiv:**2412.10352**

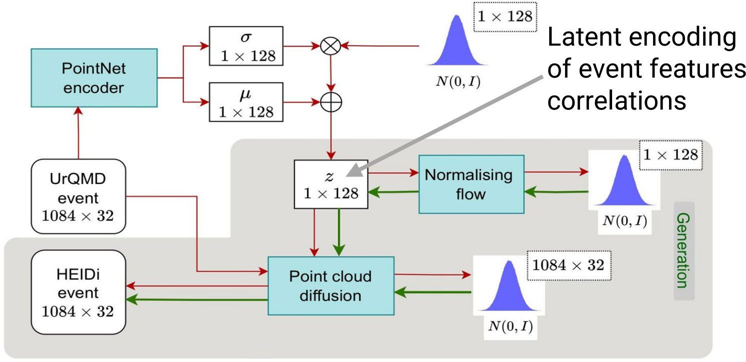
Point Cloud Diffusion Model for HICs – AI clone of simulation

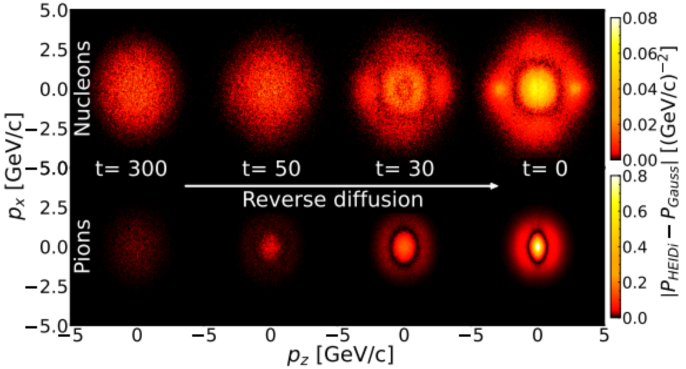


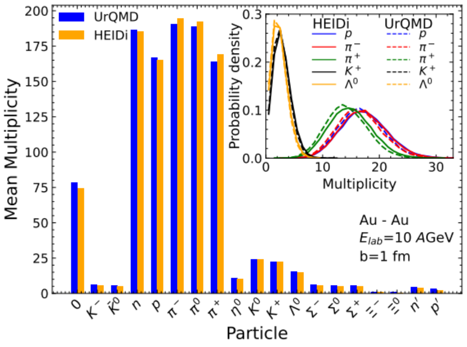
His decisions aren't any better than yours — but they're WAY faster...

- 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,** UrQMD
  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. Z, J. S, H. S, PRC2025, arXiv:**2412.10352**

Summary: Machine Learning and HENP

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

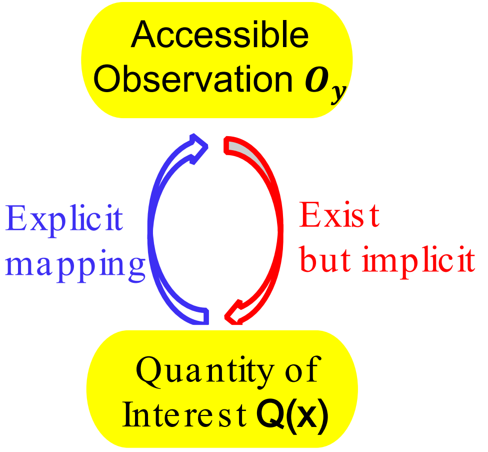
- **Bayesian Inference** for EoS from different beam energy experiment data (v2 and mT) perform well - consistent with dv1/dy measurements and BNSM constraint
  sensitivity check reveals tension: measurement uncertainty or model limitation

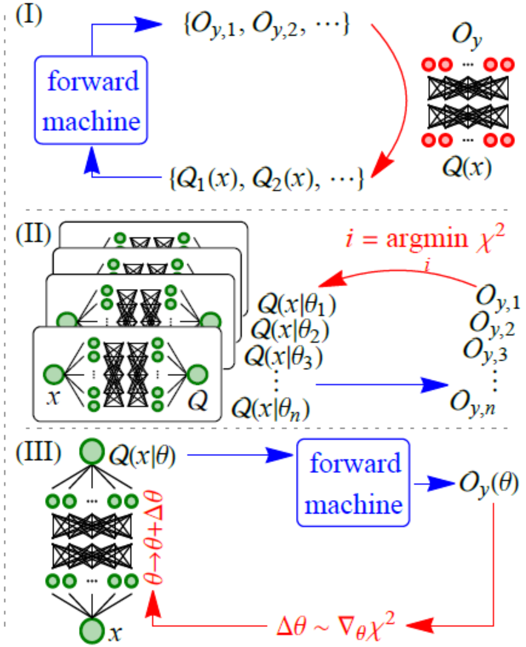
- **Auto-diff** help physics extraction taking advantange of GPU and DNN

- **Flexible DNN represented** EoS of Neutron Star can be inferred from astro-obs. (M-R) via **auto-diff based inference** with uncertainty well estimated

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

Summary : 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

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

Thanks !

Nature Review Physics (2025)

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

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