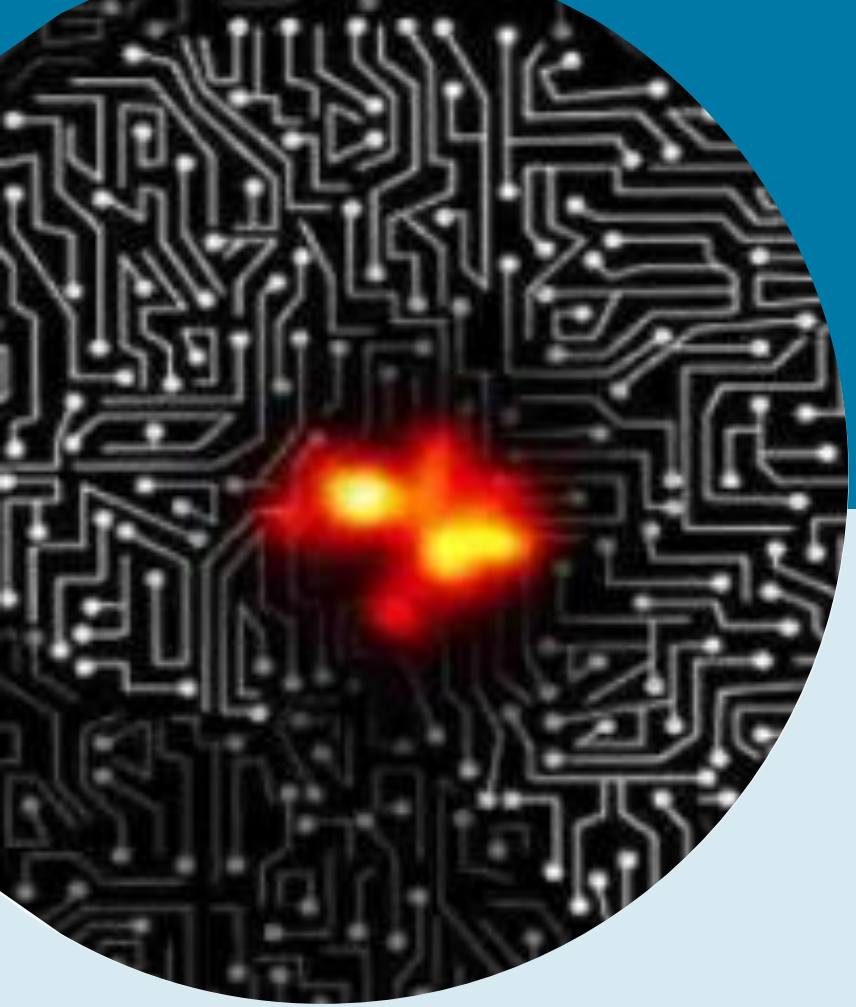


香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

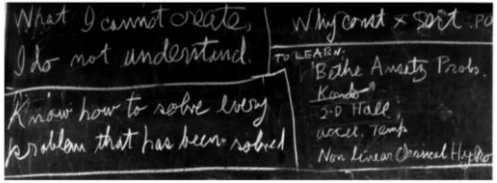
Generative Models meets Physics

周凯　香港中文大学（深圳）

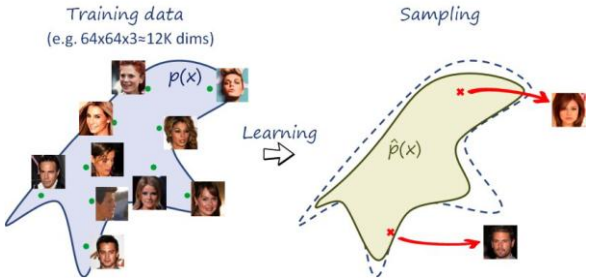
Workshop on QCD CEP and EFT
October 2026, Wuhan, CCNU

Generative AI





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



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

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

Suppose observation dataset :

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

We use parametric model to approach the data distribution :

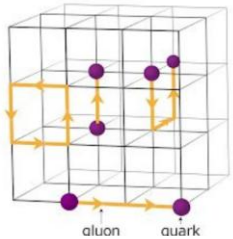
$$p_\theta(x) \rightarrow p_{data}(x)$$

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

Why GenAI matters for – high-dim distribution problems

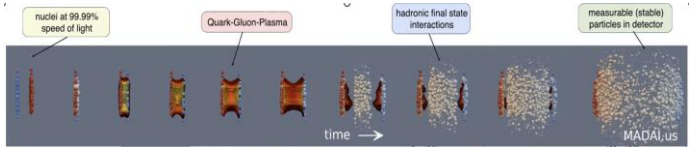
Lattice QFT

critical slowing down
topological freezing



HIC Simulation

Multi-stage EbE slow
Fluctuations/correlations



Inverse Inference

Recover physics from noisy data
Calibration and uncertainty



lattice sampler
$p(\phi) \propto e^{-S[\phi]}$

→

HIC surrogate
$p(\text{event} \mid c)$

→

posterior engine
$p(\theta \mid y)$

Bottlenecks differ, but all revolve around high-dimensional probability measures.

GenAI - Transport between probability measures

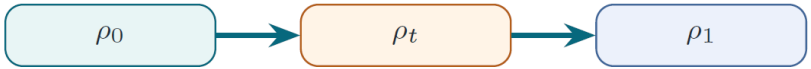
1. Generation = sampling

$$x_0 \sim \rho_0 \xrightarrow[x_1 = T_\theta(x_0)]{\text{learned dynamics}} x_1 \sim \rho_1.$$

- data-driven: $\rho_1 = p_{\text{data}}(x \mid y)$;
- action-driven: $\rho_1 = Z^{-1} e^{-S(x)}$;
- inference-driven: $\rho_1 = p(\theta \mid y)$.

2. A probability path is chosen

$\{\rho_t\}_{t\in[0,1]}$, ρ_0 simple, ρ_1 physical target.



path in distribution space

3. Different GenAI models realize the path differently

flow / FM:
$$dx_t = v_t(x_t)dt,$$
$$\partial_t \rho_t = -\nabla\cdot(\rho_t v_t),$$

diffusion:
$$dx_t = b_t(x_t)dt + g_t dW_t,$$
$$\partial_t \rho_t = -\nabla\cdot(b_t\rho_t) + \tfrac{1}{2}\nabla^2(g_t^2\rho_t).$$

Physical reading

- **state** = fields, particles, detector responses, or inferred parameters;
- **probability path** = evolution of uncertainty from prior/noise to physics ensemble;
- **dynamics** = transport that should respect action, symmetry, conservation laws, and response;
- **output** = an ensemble/posterior with calibrated observables, not a single plausible picture.



Learning from : Samples or Unnormalized Target

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

Known samples

$$x_i \sim p_{\mathrm{data}}(x), \qquad \theta^* = \arg\max_\theta \sum_i \log p_\theta(x_i).$$

HMC ensembles, hydrodynamic events, detector responses, experimental distributions.

Known unnormalized target

$$\tilde{p}_*(x) = e^{-E(x)}, \qquad p_*(x) = \tilde{p}_*(x)/Z.$$

$$D_{\mathrm{KL}}(q_\psi \| p_*) = \mathbb{E}_{q_\psi}[\log q_\psi(x) + E(x)] + \log Z.$$

Spin models, scalar/gauge fields, many-body Boltzmann weights, unnormalized Bayesian posteriors.

Message: The information boundary determines the objective: likelihood from samples, free energy from unnormalized targets, posterior density from observations.

Variational principle for data-less GenAI

NO TARGET TRAINING ENSEMBLE

$$\mathcal{F}[q] = \langle S \rangle_q + \langle \log q \rangle_q$$

action / energy — negative entropy

$$\mathcal{F}[q] + \log Z = D_{\mathrm{KL}}(q \| \pi) \geq 0$$

The trial measure is now a neural sampler.

Connection to the Gibbs–Bogoliubov–Feynman bound

$$-\log Z \leq -\log Z_0 + \langle S - S_0 \rangle_{q_0}$$

A richer trial measure is useful only if we can sample it and evaluate its density.

Difficulties in Applying the Variational Principle to Quantum Field Theories[1]

Richard P. Feynman

[1]transcript of his talk in 1987

Gibbs–Bogoliubov–Feynman principle;
Feynman, Phys. Rev. 97 (1955) 660;
Wu et al., PRL 122 (2019) 080602.

Given Hamiltonian/Action – Variational Free Energy

- Reverse KL divergence

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

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

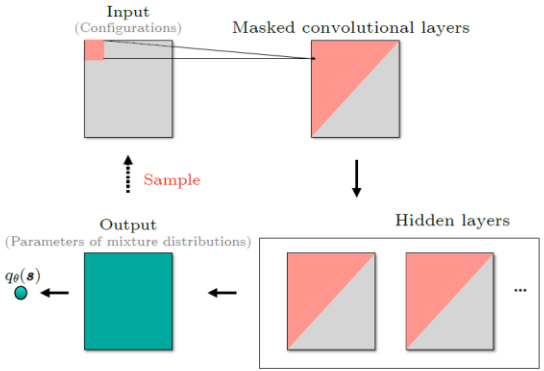
$$p(\mathbf{s}) = \frac{e^{-\beta E(\mathbf{s})}}{Z}$$

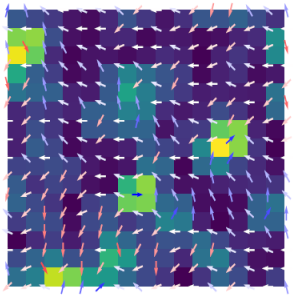
- Autoregressive $\quad q_\theta(\mathbf{s}) = \prod_{i=1}^{N} q_\theta(s_i \mid s_1, \ldots, s_{i-1})$

D. Wu, Lei Wang and P. Zhang, **PRL122,080602(2019)**

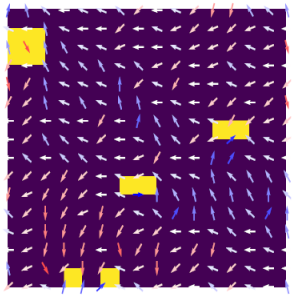
- Continuous Autoregressive for XY model

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





Probability distributions from CANs

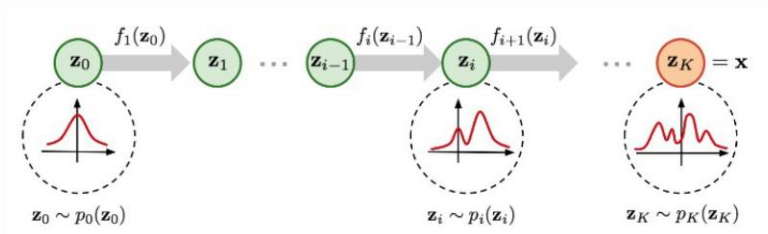


Vortices

Flow based GenAI – trivializing change of variables

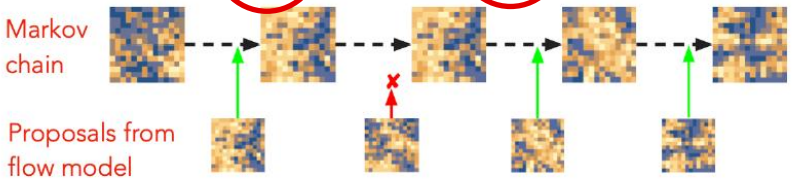
A series **(Flow)** of invertible/bijective transformations for p(z)

compose several invertible transformations to form the flow :



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

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

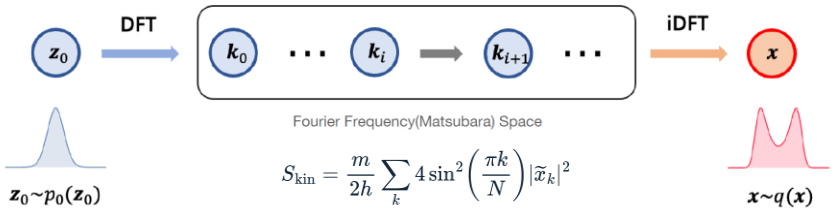


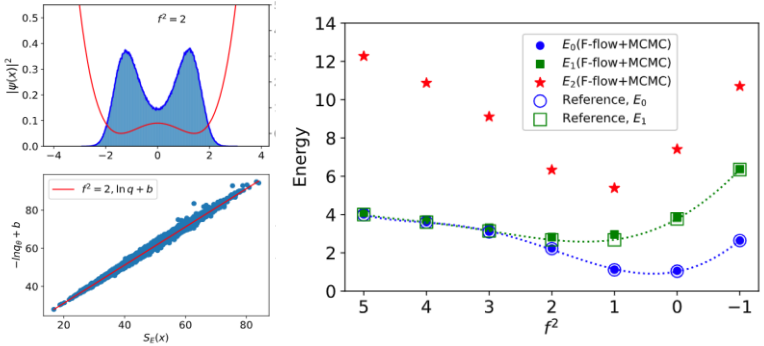
$$S_{\text{pull}}(z) = S(T_\theta(z)) - \log|\det DT_\theta(z)| \overset{\text{ideal}}{=} S_0(z) + C$$

Lüscher, 0907.5491; 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)

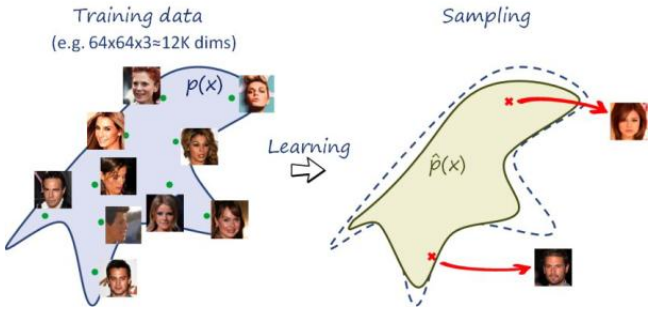
Fourier Flow Model



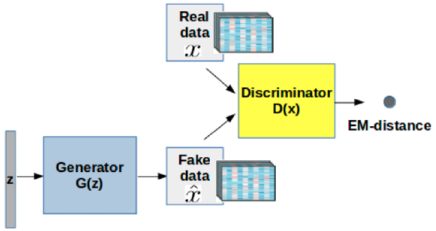


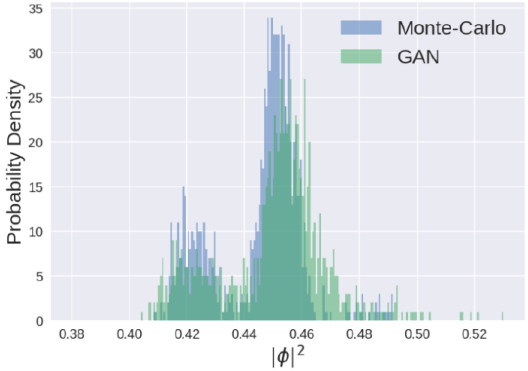
S. Chen, **O. Savchuk**, S. Z, B. C, H. S, L. W, **K. Zhou**, **PRD107, 056001(2023)**

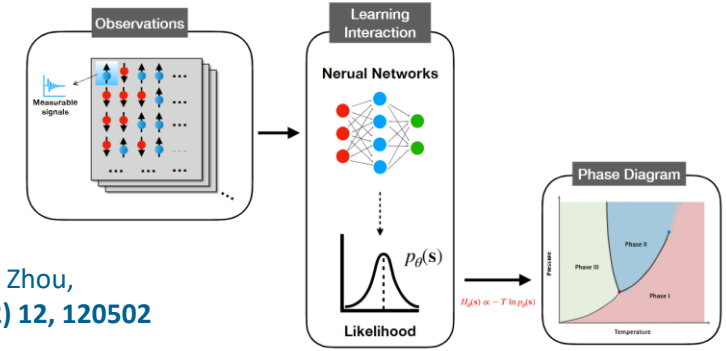
Data-driven GenAI – early proof of principle for IQFT/many-body physics



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

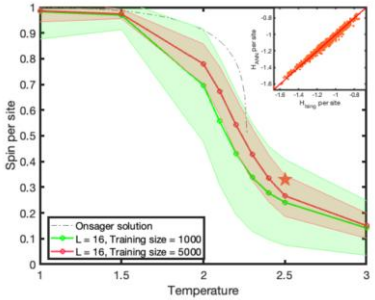


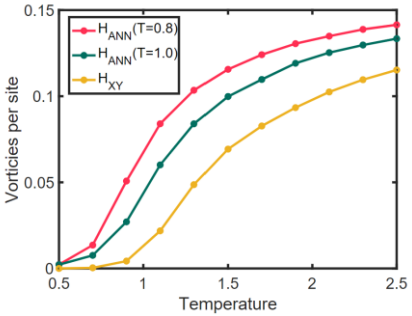




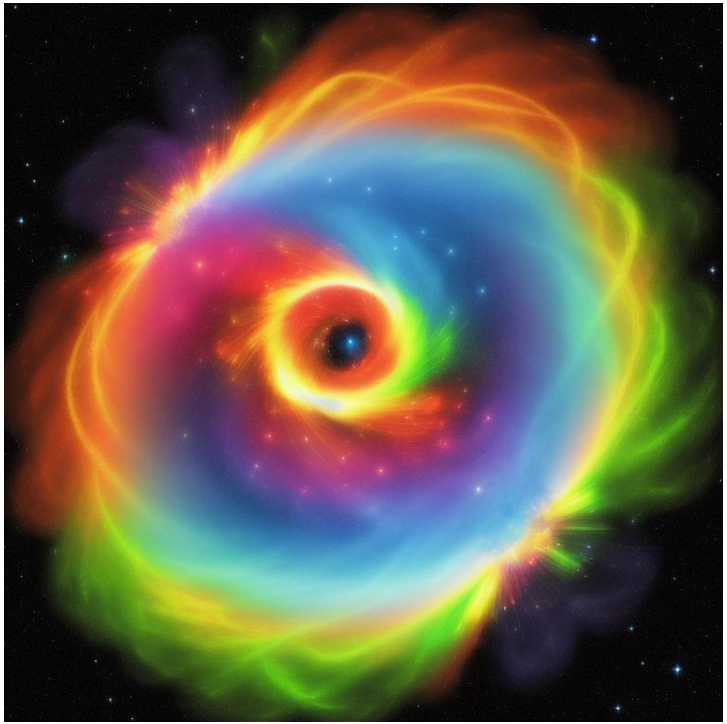
L. Wang, L. He, Y. Jiang, K. Zhou,
Chin. Phys. Lett. 39 (2022) 12, 120502

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





Diffusion Model



“A heavy quark move inside quark-gluon plasma”

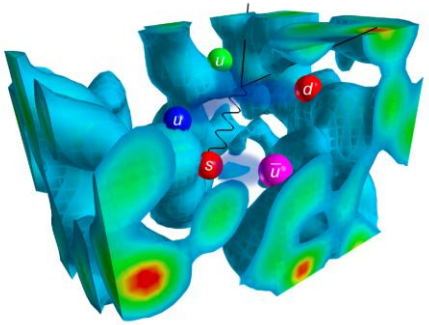


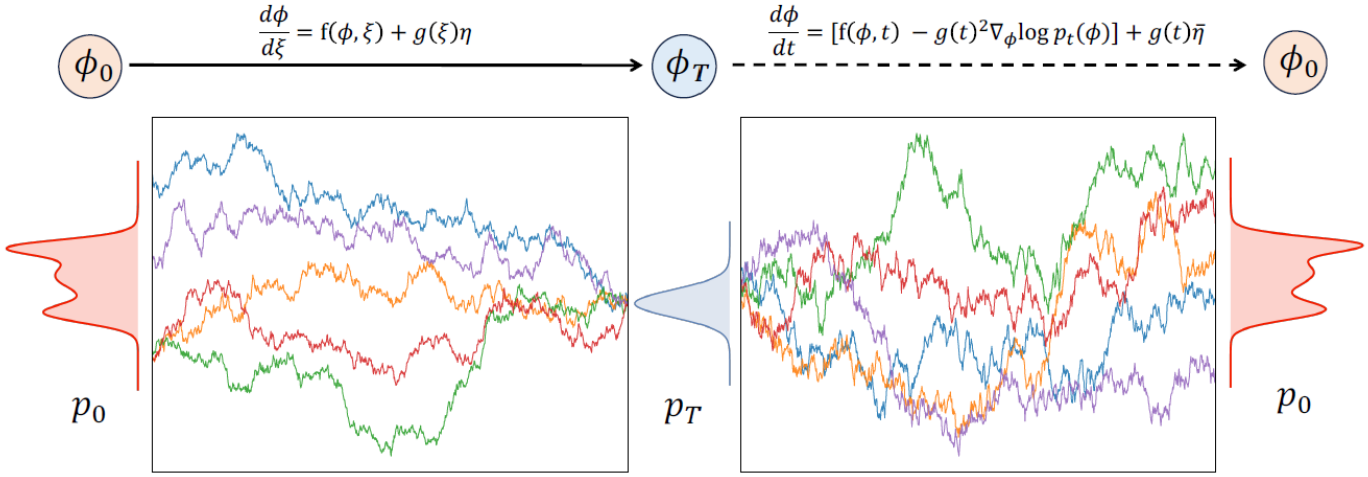


Diffusion Model meets lattice QFT

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

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





L. Wang, G. Aarts, K. Zhou, **JHEP 05 (2024) 060**
L. Wang, G. Aarts, K. Zhou, **2311.03578** (NeurIPS 2023 workshop "ML & Physical Sciences")
G. A, D. E. H, L. W, K. Zhou, **2410:21212** (NeurIPS 2024 workshop "ML & Physical Sciences) → **"Best Physics for AI Paper" Award**
Q. Zhu, G. Aarts, W. Wang, K. Zhou, L. Wang, **JHEP 2026,**
G. Aarts, D.E.H, L.W, K. Zhou, **MLST 6(2025)025004, 2510.01328**
S. Chen, M. Qian, G. Aarts, B. Lucini, K. Zhou**, JHEP 2026, 2606.13790 (training without HMC data preparation)**

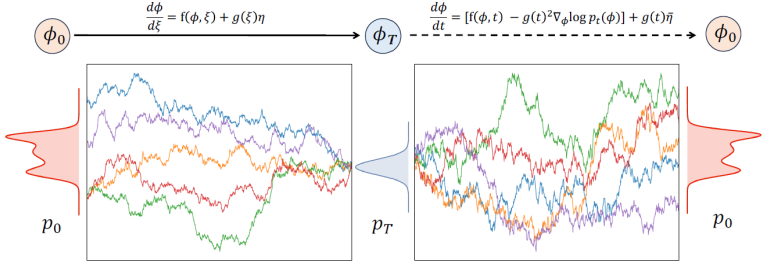
Diffusion Model as Stochastic Quantization

- Forward diffusion SDE $\dfrac{d\phi}{d\xi} = f(\phi,\xi) + g(\xi)\eta(\xi)$

$$\langle \eta(\xi)\eta(\xi')\rangle = 2\alpha\delta(\xi-\xi')$$

- Backward diffusion SDE

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



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

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

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

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

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

$$\frac{\partial \phi(x,\tau)}{\partial \tau} = -\nabla_\phi S(\phi) + \sqrt{2}\eta(x,\tau) \qquad \frac{\partial P[\phi,\tau]}{\partial \tau} = \alpha \int d^n x \left\{ \frac{\delta}{\delta\phi}\left(\frac{\delta}{\delta\phi} + \frac{\delta S_E}{\delta\phi} \right) \right\} P[\phi,\tau]$$

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

sampling from DM is ⇔ optimizing a stochastic trajectory to approach the "<u>equilibrium state</u>"

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

DM Relation to (inverse) RG

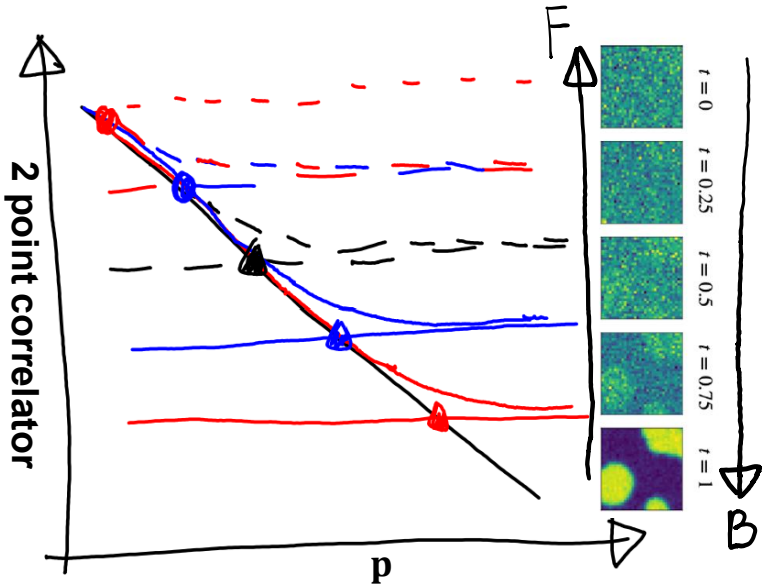
- Forward diffusion kernel: **gaussian smoothing**

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

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

- 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 first,

With decreasing cut scale because of the gradually increasing noise level !

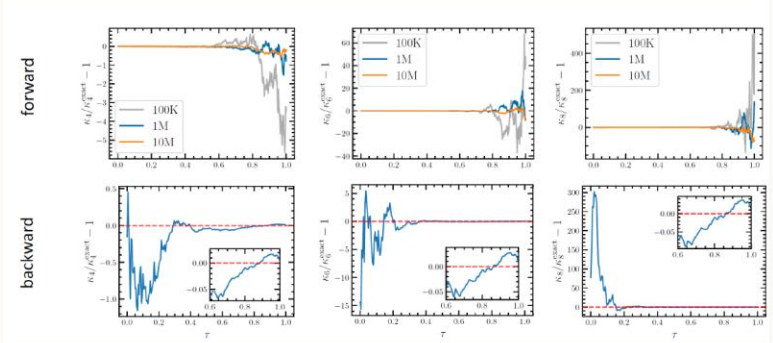


In FRG, the high frequency (short-distance) degrees of freedom is progressively integrated out !

Higher order cumulants

$$W[J] = \log Z[J], \qquad \kappa_n = \left.\frac{\delta^n W[J]}{\delta J^n}\right|_{J=0}$$

- Higher connected functions are fingerprints of non-Gaussian interactions.
- They test more than marginal histograms or two-point structure.
- A useful score must reconstruct hidden interaction information.



$$\kappa_2(\xi) = c_\xi^2 \kappa_2(0) + \mathrm{Var}(\eta_\xi)$$
$$\kappa_{n>2}(\xi) = c_\xi^n \kappa_n(0)$$

Cumulants turn diffusion validation into a field-theory question.

Demonstrated: higher connected cumulants can be learned by diffusion models
Bottleneck: checking interaction physics under denoising
Open: routine validation in harder theories

$$\kappa_{n>2}(t) = \left.\frac{d^n W[J]}{dJ(t)^n}\right|_{J=0} = \left.\frac{d^n}{dJ(t)^n} \log \mathbb{E}_{P_0}[e^{J(t)x_0 f(t,0)}]\right|_{J=0} = \kappa_n(0) f^n(t,0)$$

Why this matters for diffusion

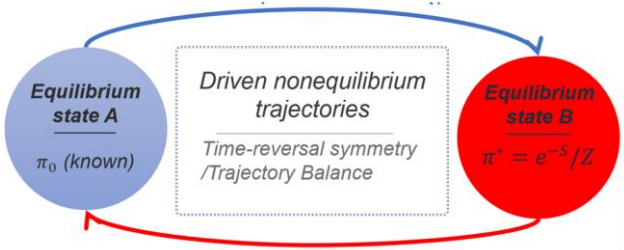
- forward noising quickly hides non-Gaussian structure behind added variance
- the learned score must therefore carry that information implicitly
- this is a much more physical diagnostic than image-like quality metrics

Aarts, Habibi, Wang, Zhou, MLST 6 (2025) 025004.

Stochastic Path Sampler without HMC data

Forward path

$$P_F(\tau) = \pi_0(x_0) \prod_{k=0}^{T-1} q_F^k(x_{k+1}|x_k).$$



Backward reference

$$\tilde{Q}_B(\tau) = e^{-\mathcal{E}(x_T)} \prod_{k=0}^{T-1} q_B^k(x_k|x_{k+1}).$$

$$\mathcal{L}_{\text{path}} \sim \mathbb{E}_{\mathcal{P}_F^\theta}\left[\log \frac{\mathcal{P}_F^\theta(\tau)}{Q_B^\theta(\tau)} + S(x_T)\right]$$

The learning signal is global trajectory consistency.

Forward protocol vs backward reference

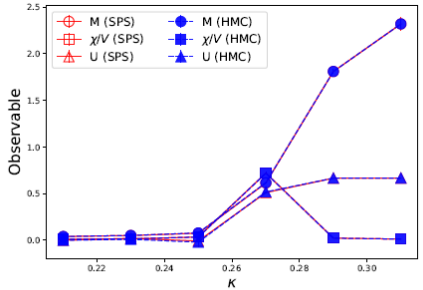
$$W(\tau) = \log \frac{P_F(\tau)}{\tilde{Q}_B(\tau)}, \qquad \Sigma(\tau) = \log \frac{P_F(\tau)}{Q_B(\tau)} = W(\tau) + \log Z.$$

$$\mathbb{E}_F[e^{-W}] = Z$$

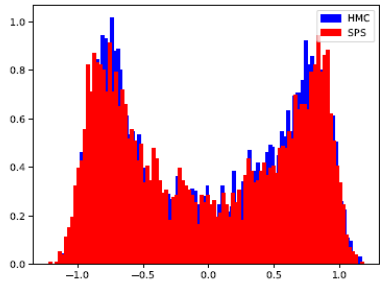
$$\langle O \rangle_\pi = \frac{\mathbb{E}_F[O(x_K)e^{-W}]}{\mathbb{E}_F[e^{-W}]}$$

Possible IMH on the path:

$$A = \min\{1, e^{W_{\text{old}} - W_{\text{new}}}\}$$



(b) $L = 32$



(b) $L = 32$

S. Chen, M. Qian, B. Lucini, G. Aarts, K. Zhou, **2606.13790 (JHEP accepted)**

GenAI – flow PDEs

General SDE ladder

假设各向同性扩散： a_t = α(t)I

1 SDE + Fokker-Planck

$$dX_t = b_t(X_t)\,dt + \sigma_t(X_t)\,dW_t, \quad a_t = \sigma_t\sigma_t^\top$$
$$\partial_t\rho_t = -\nabla\cdot(b_t\rho_t) + \tfrac{1}{2}\partial_i\partial_j(a_{ij,t}\rho_t)$$

↓

2 log transform: ρ_t = exp(-S_t)

$$\partial_t S_t = \nabla\cdot b_t - b_t\cdot\nabla S_t + \frac{\alpha(t)}{2}\left(\Delta S_t - \|\nabla S_t\|^2\right)$$

标量 action 比直接向量场少掉旋转 / gauge-like 冗余。

↓

3 score 与 probability-flow velocity

$$s_t = \nabla\log\rho_t = -\nabla S_t, \qquad v_t = b_t - \tfrac{\alpha(t)}{2}s_t$$
$$\partial_t s_t = \nabla\left[-\nabla\cdot b_t - b_t\cdot s_t + \tfrac{\alpha(t)}{2}(\nabla\cdot s_t + \|s_t\|^2)\right]$$

采样 ODE/SDE 只需 score；物理求解更适合从 S_t 或势函数出发。

VE DM case

b_t=0, α(t)=g(t)^2

forward heat semigroup

$$dX_t = g(t)dW_t, \quad \tau = \int_0^t g(r)^2 dr = \sigma(t)^2$$
$$\partial_\tau\rho_\tau = \tfrac{1}{2}\Delta\rho_\tau$$

effective action PDE = HJB / KPZ

$$\boxed{\partial_\tau S_\tau = \tfrac{1}{2}\left(\Delta S_\tau - \|\nabla S_\tau\|^2\right)}$$

Cole-Hopf: ρ=e^{-S} 让热方程变成非线性 action flow 。

score / Burgers / sampling velocity

$$s_\tau = -\nabla S_\tau, \quad v_\tau = -\tfrac{1}{2}s_\tau = \tfrac{1}{2}\nabla S_\tau$$
$$u_\tau := \nabla S_\tau \quad \Rightarrow \quad \partial_\tau u_\tau + (u_\tau\cdot\nabla)u_\tau = \tfrac{1}{2}\Delta u_\tau$$

Burgers 形式说明 score 场不是任意向量场，而是梯度流。

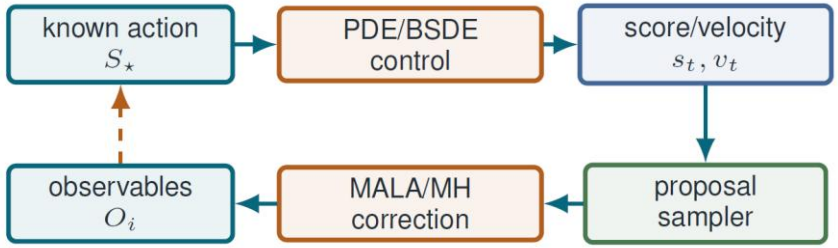
Outcome: path → PDE type → feasible solver

GenAI – flow PDEs

模型	path 动力学	scalar action S_t	score PDE	velocity PDE
VE	$dX = g(t)dW$	$\partial_\tau S = \frac{1}{2}(\Delta S - \|\nabla S\|^2)$	$\partial_\tau s = \frac{1}{2}\nabla(\nabla \cdot s + \|s\|^2)$	$v = -\frac{1}{2}s,\ \partial_\tau v + 2(v \cdot \nabla)v = \frac{1}{2}\Delta v$
DDPM/VP	$dX = -\frac{1}{2}\beta X dt + \sqrt{\beta}dW$	$\partial_t S = -\kappa d + \kappa x \cdot \nabla S + \kappa\Delta S - \kappa\|\nabla S\|^2$	$\partial_t s = \kappa[s + (x \cdot \nabla)s + \nabla(\nabla \cdot s + \|s\|^2)]$	$v = -\kappa(x + s)$
Flow matching	$X_t = (1 - t)X_0 + tX_1$	$\partial_t S = \nabla \cdot v - v \cdot \nabla S$	$\partial_t s = -\nabla(\nabla \cdot v + v \cdot s)$	$\partial_t v + (v \cdot \nabla)v = -\rho^{-1}(\rho\Sigma)$
Schrödinger bridge	$dX = \varepsilon\nabla\eta\, dt + \sqrt{\varepsilon}dW$	$\partial_t S = \Delta\psi - \nabla S \cdot \nabla\psi$	$\partial_t s = -\nabla(\Delta\psi + \nabla\psi \cdot s)$	$\partial_t v + (v \cdot \nabla)v = -\frac{\varepsilon^2}{2}\nabla(\Delta\sqrt{\rho}/\sqrt{\rho})$

From functional PDE to tractable control

- solve layerwise and center unknown normalizers;
- fit scalar potentials or local operator coefficients, not arbitrary high-dimensional vector fields;
- sample collocation points from the current typical set;
- use Hutchinson/JVP estimators for traces and Burgers terms;
- close with MALA/MH and physical observables.



PHYSICS-INFORMED KERNELS

$$\pi_t = Z_t^{-1} e^{-S_t}, \qquad \partial_t \pi_t = -\operatorname{div}(\pi_t v_t)$$

$$\operatorname{div} v_t - v_t \cdot \nabla S_t = \partial_t S_t - \langle \partial_t S_t \rangle_{\pi_t}$$

Choose the physical probability path, then Learn the velocity that can induce that path, or, conserves those probability.

Linear in v for specified action path; Reference-field subtraction removes the unknown time-only normalizer.

Ihssen, Kapust, Pawlowski, arXiv:2510.26678.

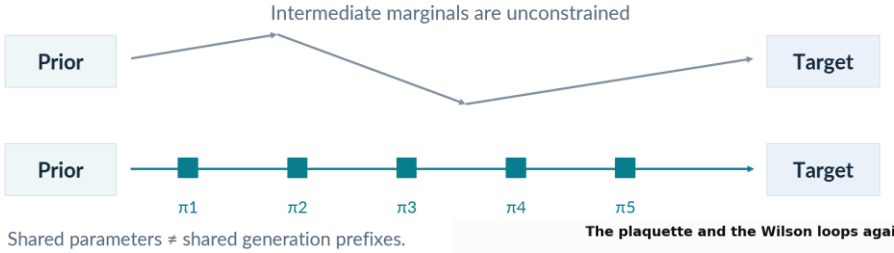
Stochastic path optimization for prescribed physical path

ENDPOINT OBJECTIVE

Any useful route to π_1

PRESCRIBED PHYSICAL PATH

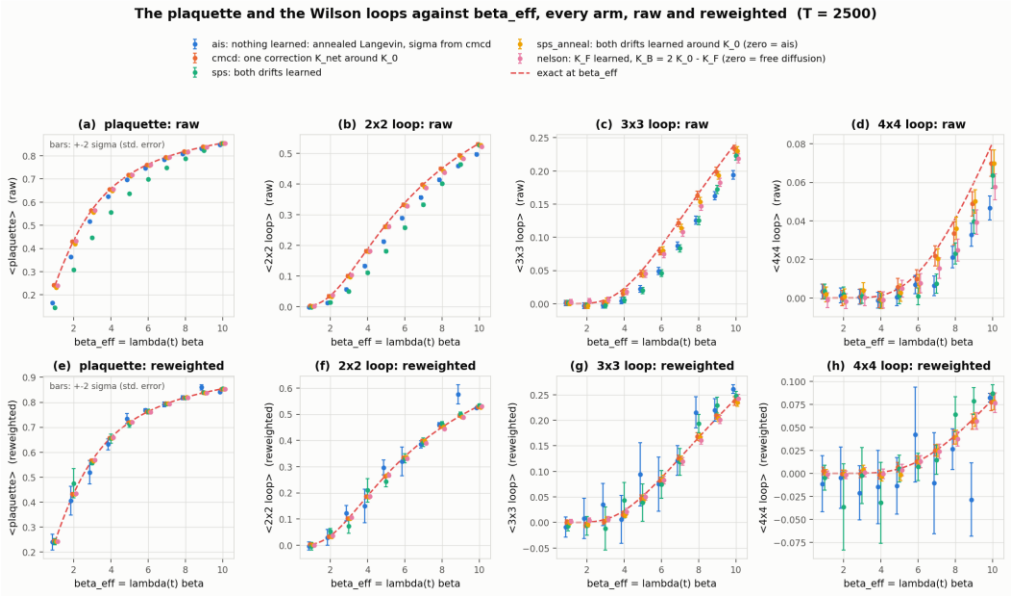
Specified πt at every time



Shared parameters ≠ shared generation prefixes.

To follow a specific physical path, one can set it via stochastic path consistency by constraining the relationship between forward and backward drifts →

A constrained stochastic path objective becomes a variational route to Wegner transport equation



R.Thakkar, K. Zhou, et.al,

to appear soon

← Preliminary

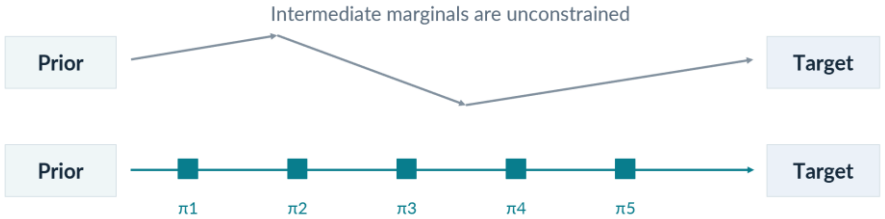
Stochastic path optimization for prescribed physical path

ENDPOINT OBJECTIVE

Any useful route to π_1

PRESCRIBED PHYSICAL PATH

Specified πt at every time



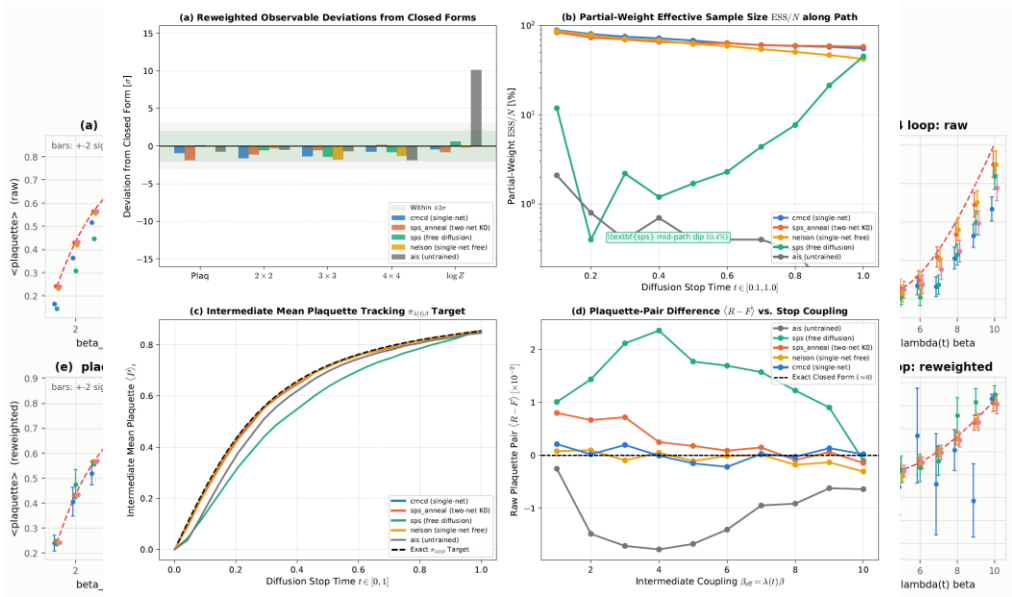
Shared parameters ≠ shared generation prefixes.

To follow a specific physical path, one can set it via stochastic path consistency by constraining the relationship between forward and backward drifts →

A constrained stochastic path objective becomes a variational route to Wegner transport equation

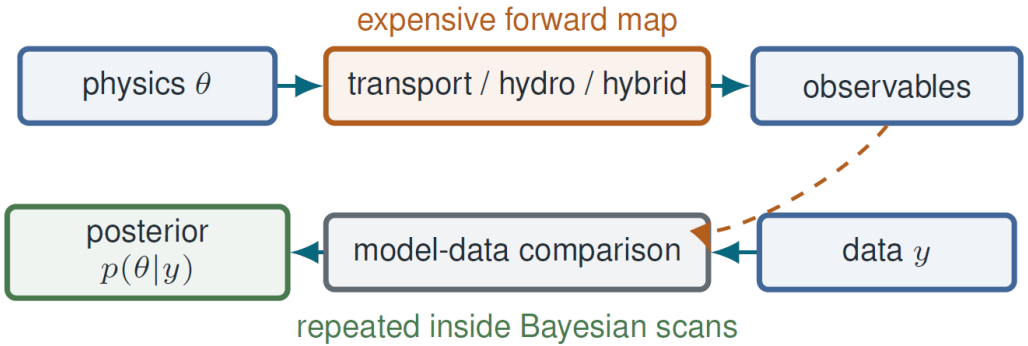
R.Thakkar, K. Zhou, et.al,

to appear soon



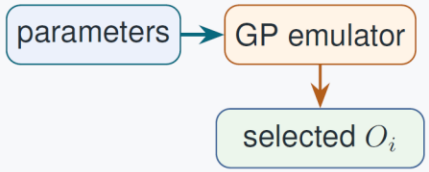
← Preliminary

GenAI for HIC physics bottleneck – Inference Loop



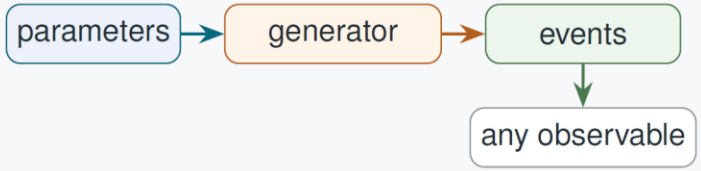
- Infer EoS, initial conditions, transport coefficients, and phase structure from observables.
- Forward map: transport, hydrodynamics, hybrid evolution, detector response.
- Bottleneck: repeated simulator calls inside likelihood or posterior loops.

Observable emulator



- Efficient for a small observable vector.
- Harder to scale when many differential observables, correlations, and cumulants are needed.

Event-level generator



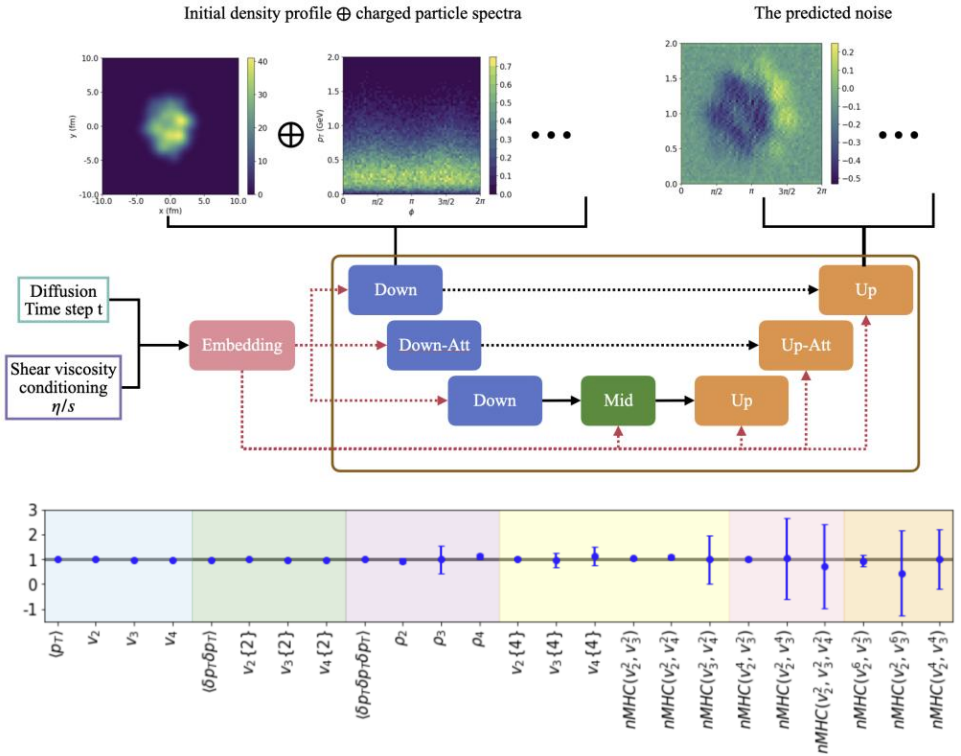
- Samples complete events; observables are computed afterwards.
- Better aligned with event-by-event Bayesian inference and multi-observable systematics.

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]

Phys.Rev.C 112 (2025) 5, L051903



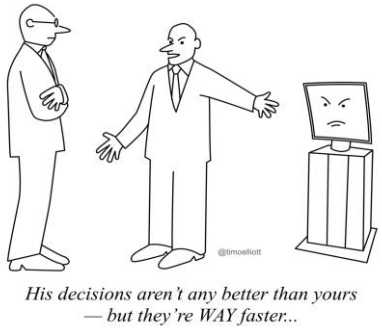
tor. We carried out (2+1)D minimum bias simulations of Pb-Pb collisions at 5.02 TeV, choosing the shear viscosity η/s to be one of three distinct values: 0.0, 0.1, and 0.2. For each value of η/s, we generate 12,000 pairs of initial entropy density profiles and final particle spectra, corresponding to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

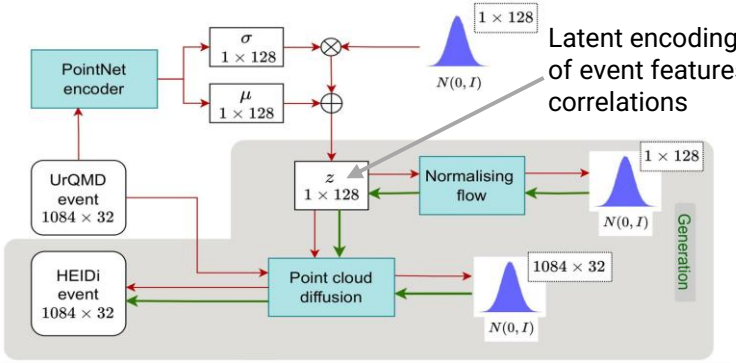
ble precision, as the traditional numerical simulation of hydrodynamics for one central event typically takes approximately 120 minutes (10^4 seconds) on a single CPU.

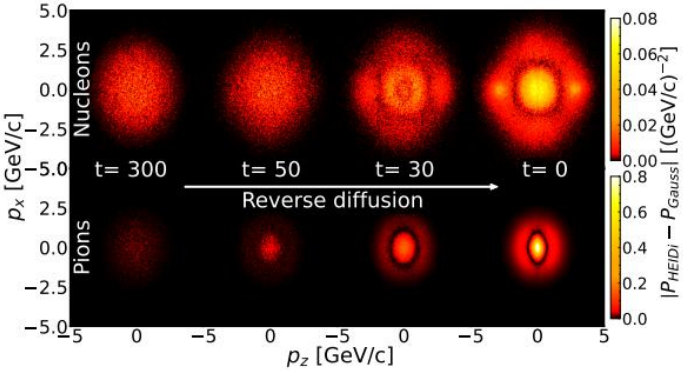
Point Cloud Diffusion Model for HICs – AI clone of simulation

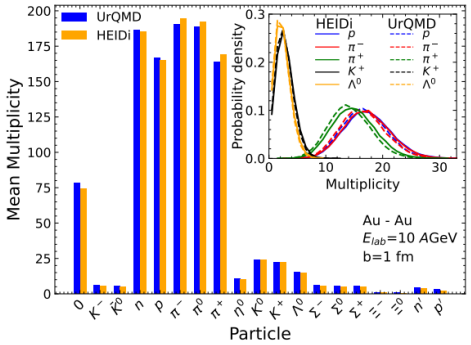


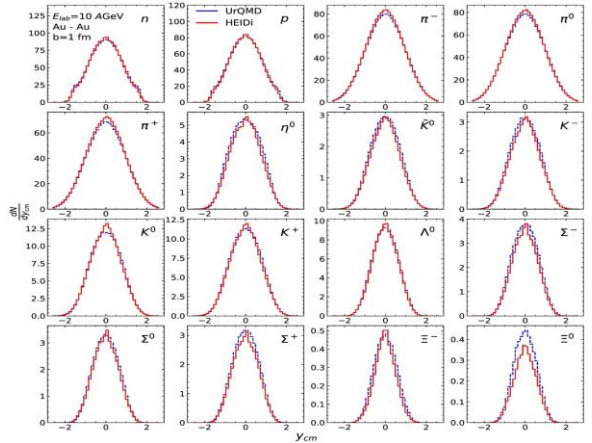
- 18k UrQMD simulation events for central Au-Au@10 AGeV collisions
- HEIDi:
  Heavy-ion Events through Intelligent Diffusion

PointNet encoder + Normalizing flow decoder + Pointcloud diffusion →



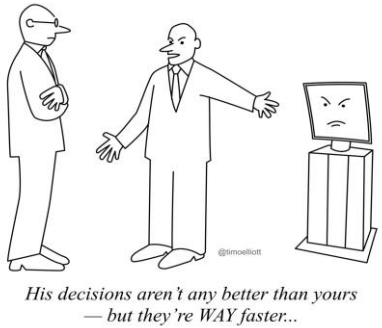






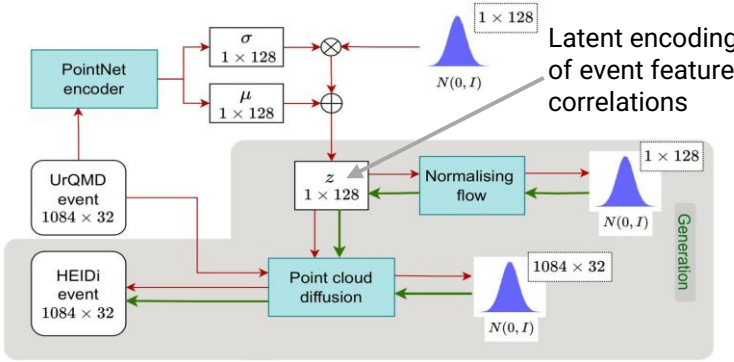
M. O.K, K. Zhou, J. S, H. S, **PRC 112 (2025), L051902; PRC 112 (2025), 054907**

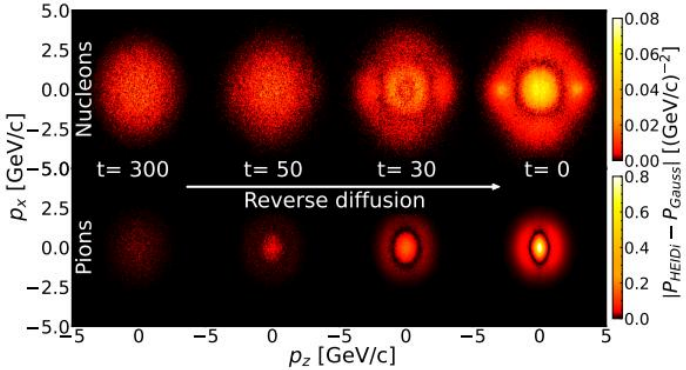
Point Cloud Diffusion Model for HICs – AI clone of simulation

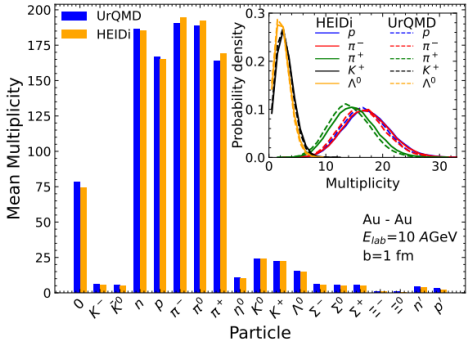


- 18k UrQMD simulation events for central Au-Au@10 AGeV collisions
- HEIDi:
  Heavy-ion Events through Intelligent Diffusion

PointNet encoder + Normalizing flow decoder + Pointcloud diffusion →





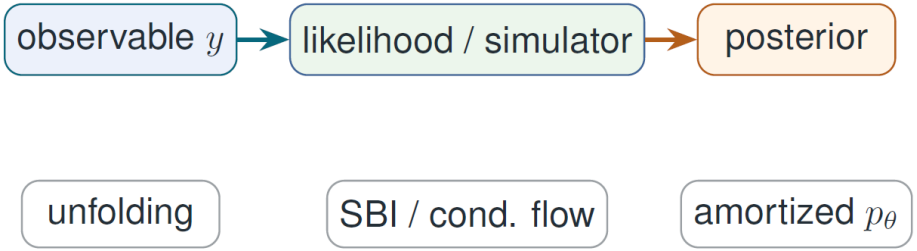


- **Running time of** UrQMD simulation
  cascade : ~ 3 sec/event;
  with potential : ~ 3 min/event;
  hybrid : ~ 1 hour/event
- HEIDI on A100: ~ 30 ms/event
- Speedup **2 ~ 5 orders** of magnitude

M. O.K, K. Zhou, J. S, H. S, **PRC 112 (2025), L051902; PRC 112 (2025), 054907**

Inverse Inference – GenAI as a posterior engine

$$p(\theta \mid y) \propto p(y \mid \theta)\, p(\theta)$$



- The physics signal may be hidden, smeared, or only indirectly observed.
- GenAI enters as learned unfolding, conditional density estimation, or simulation-based inference.
- The desired output is a **distribution over explanations**, not only a point estimate.

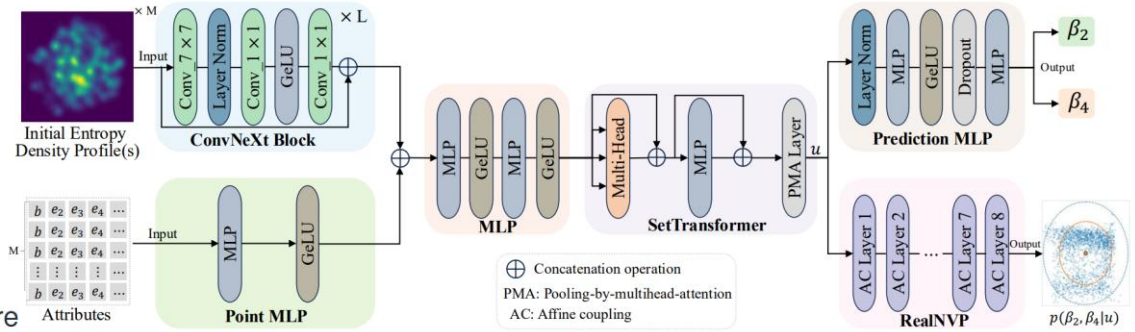
Physics interpretation

The model should not merely predict a hidden variable; it should quantify which physical explanations remain plausible.

Inverse problems turn GenAI from a generator into an uncertainty-aware inference tool.

Inference on nuclear structure in HICs

- initial entropy-density or nucleon-configuration features are encoded from event bags;
- conditional flow/SBI returns a calibrated posterior over deformation parameters;
- the GenAI component represents non-Gaussian uncertainty, not only regression.

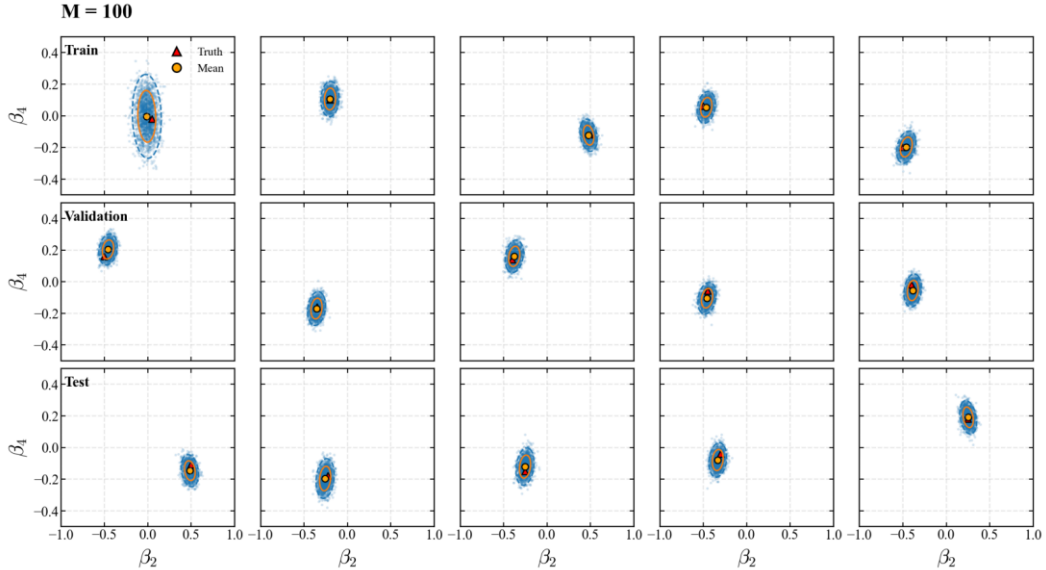


- Single events are too noisy: participant fluctuations obscure deformation information.
- Grouping many events with the same deformation parameters strongly improves identifiability.
- Conditional normalizing flows provide

$$p_\theta(\beta_2, \beta_4 \mid \text{event bag})$$

instead of a single estimate.

- GenAI acts as a many-event likelihood-to-posterior map.



J. Tao, Y. Liu, Y. Sha, X. Fan, Y. Tu, K. Zhou, H. Zheng, B. Zhang, **arXiv: 2603.24088**

Summary

Lattice QFT

sampler and PDE testbed

$$\pi[\phi] \propto e^{-S[\phi]}$$

transport must preserve action, symmetries, cumulants, topology, and path-space diagnostics

HIC simulation

conditional surrogate

$$p(\text{event} \mid c)$$

fast generators must preserve event structure relevant for Bayesian calibration and observables

Inverse problems

posterior engine

$$p(\theta \mid y)$$

models should infer hidden variables, uncertainties, and earlier dynamical stages

probability paths + action structure + physical validation

Unifying statement

GenAI is useful here when it becomes a controlled transport tool: a sampler, a surrogate, or a posterior map whose errors are judged by physics.