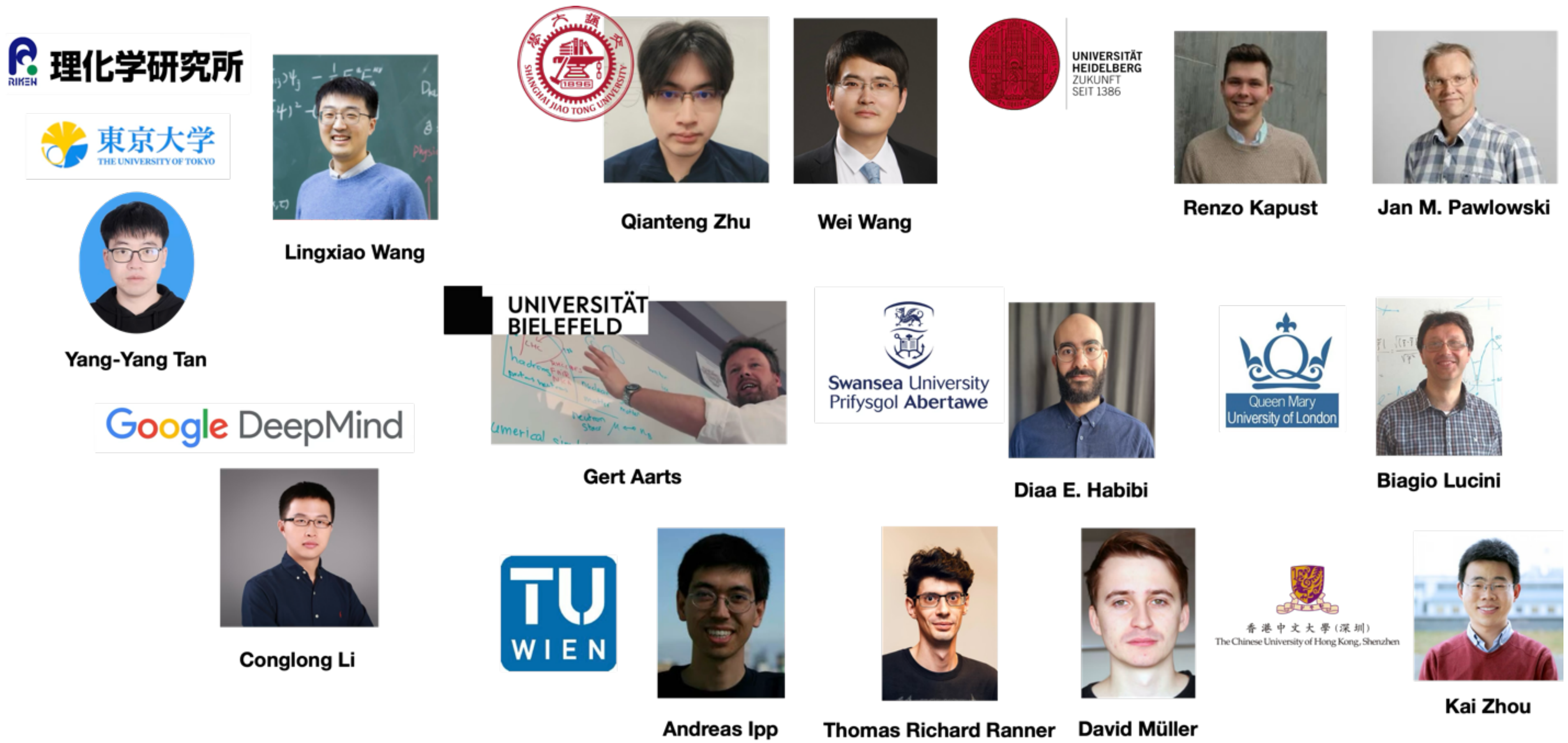
Diffusion Models as Stochastic Quantization (on the Lattice)

Lingxiao Wang (王 凌霄)
AI as Science team, RIKEN iTHEMS
Institute for Physics of Intelligence(iπ), UTokyo

with [DM-QFT collaboration]

13th International Conference on the Exact Renormalization Group 2026

Background and References



DM-QFT

- **LW**, G Aarts, K Zhou, *JHEP 05 (2024) 060* [2309.17082 [hep-lat]]
- G Aarts, D Habibi, **LW***, K Zhou, *Mach.Learn.Sci.Tech. 6 (2025) 2, 025004* and NeurIPS 2024 ML4PS [2410.21212 [hep-lat]] (Best ‘Physics for AI’ Paper)
- G Aarts, K Fukushima, T Hatsuda, A Ipp, S Shi, **LW***, K Zhou, *Nature Rev.Phys. 7 (2025) 3, 154-163*, [2501.05580 [hep-lat]]
- Q Zhu, G Aarts, W Wang, K Zhou, **LW***, *JHEP03(2026)111* [2502.05504 [hep-lat]] and NeurIPS 2024 ML4PS [2410.19602 [hep-lat]]
- G Aarts, D Habibi, A Ipp, D, Müller, T Ranner, **LW***, W Wang, Q Zhu [2601.19552 [hep-lat]]
- Y Tan, G Aarts, D Habibi, B Lucini, **LW*** [2607.08505 [hep-lat]]

Outline

What ‣ **What is Diffusion Model?**

Why ‣ **Physics of Diffusion Models**

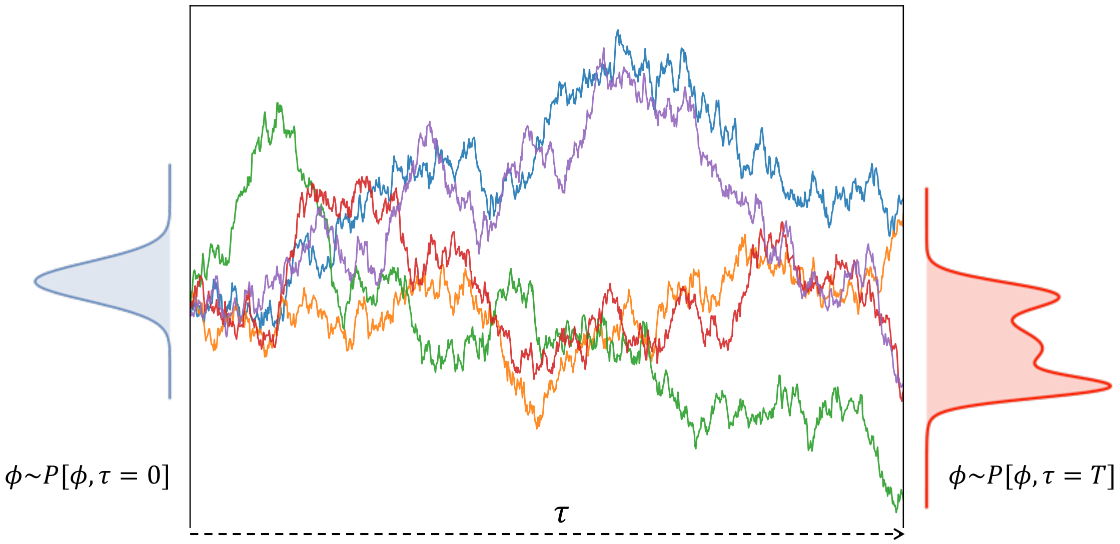
- Stochastic Quantization
- Effective Actions

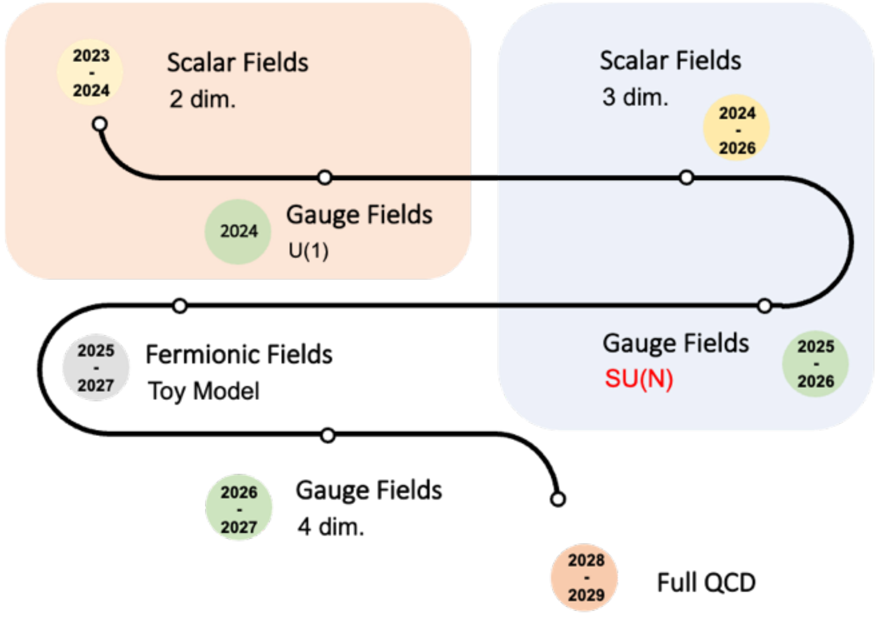
How ‣ **DM on the Lattice**

- **Generation of Lattice Fields**
- **Toward Criticality**
- **Gauge Field Simulations**

New ‣ **Optimal Stochastic Quantization**

‣ **Summary and Outlook**

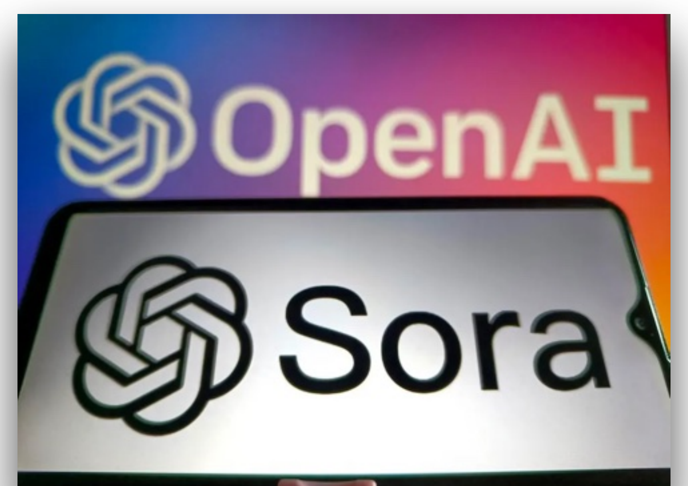




What is Diffusion Model?

A Diffusion Model is a type of **generative model** used in machine learning to **create (or generate) data** — such as images, audio, or text — by **gradually transforming random noise into structured data**.

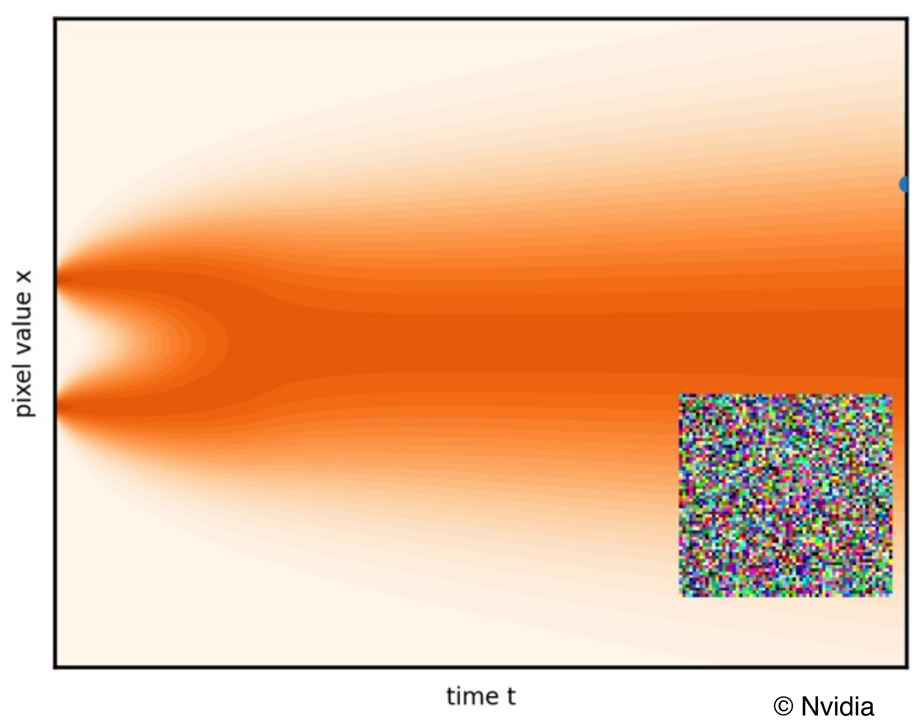
— GPT-5

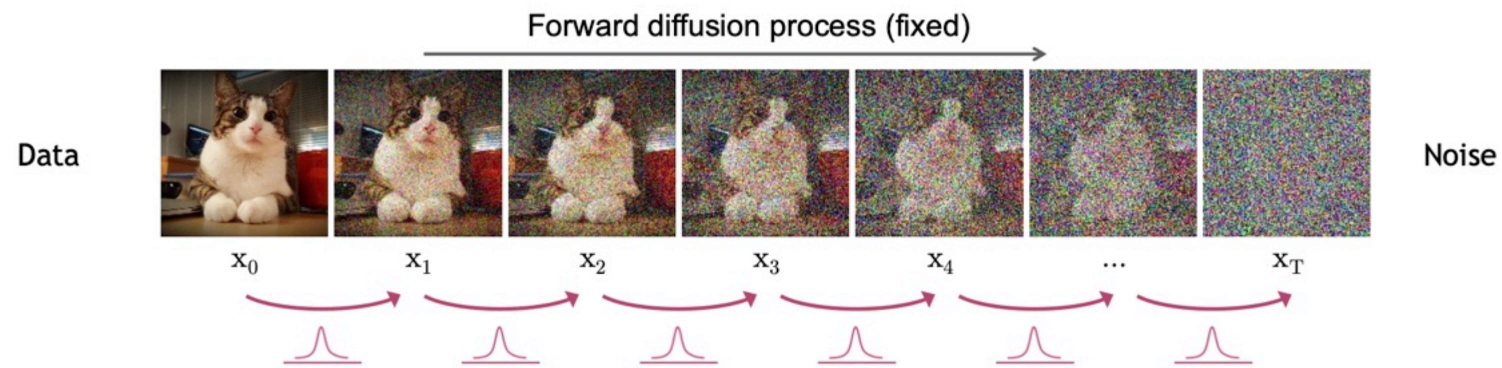


Diffusion Models

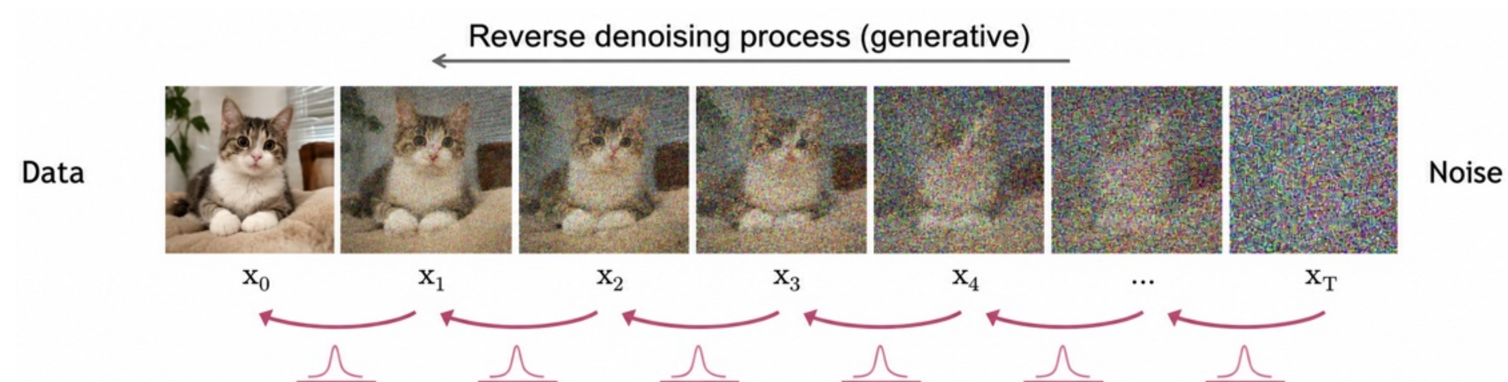
Practical Level

- Forward diffusion process **gradually adds noise** to input
- Reverse denoising process learns to **generate data by denoising**





Ho et al., Denoising Diffusion Probabilistic Models, NeurIPS 2020



Probabilistic models
learn how to **denoise** from **a (simple) prior distribution** to **a target distribution**

Diffusion Models

[Y Song et al., ICLR 2021]

[B Anderson, in Stochastic Processes and their Applications, 1982]

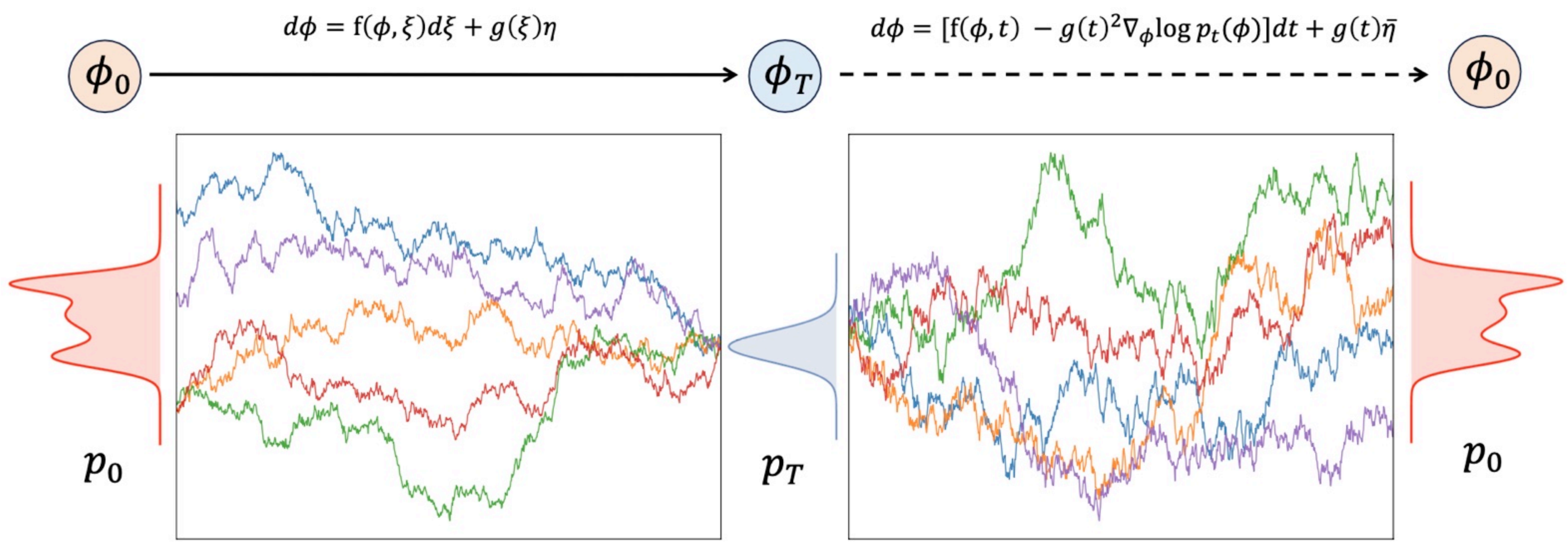
Reverse Diffusion

- **Forward Diffusion process is driven by a stochastic differential equation(SDE)**
  - **Drift term**: pulls towards mode
  - **Diffusion term**: injects noise

$$\frac{d\phi}{d\xi} = \boxed{f(\phi, \xi)} + \boxed{g(\xi)\eta(\xi)}$$

- **Reverse Diffusion SDE**
  - Drift term is adjusted with a "**Score Function**"

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



Diffusion Models

[Y Song et al., ICLR 2021]

[B Anderson, in Stochastic Processes and their Applications, 1982]

Reverse Diffusion

- **Reverse Diffusion SDE**
  - But how to get the score function?

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

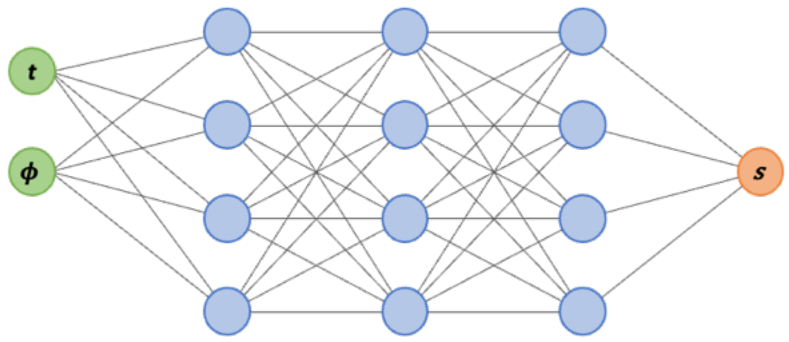
$$\phi_0 \sim p_{\text{data}} \xrightarrow{\phi_t \sim p_{0t}(\phi_t|\phi_0)} \phi_t \xrightarrow{\textbf{match conditional score}} s_\theta(\phi_t, t) \approx \nabla_{\phi_t} \log p_{0t}(\phi_t \,|\, \phi_0)$$

Vincent, Pascal. Neural computation 23.7 (2011): 1661-1674.

$$\boxed{\frac{d\phi}{dt} = \left[f(\phi,t) - g^2(t)\mathbf{s}_{\hat{\theta}}(\phi,t)\right] + g(t)\bar{\eta}(t)}$$

Universal Approximation Theorem (1989,1991)

A feed-forward network with a single hidden layer containing a finite number of neurons can approximate arbitrary continuous functions.



Stochastic Quantization

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

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

τ: **fictitious** time, α: diffusion constant

Parisi & Wu (1980); Damgaard & Hüffel (1987); Namiki (1993)

- **Fokker-Planck equation**

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

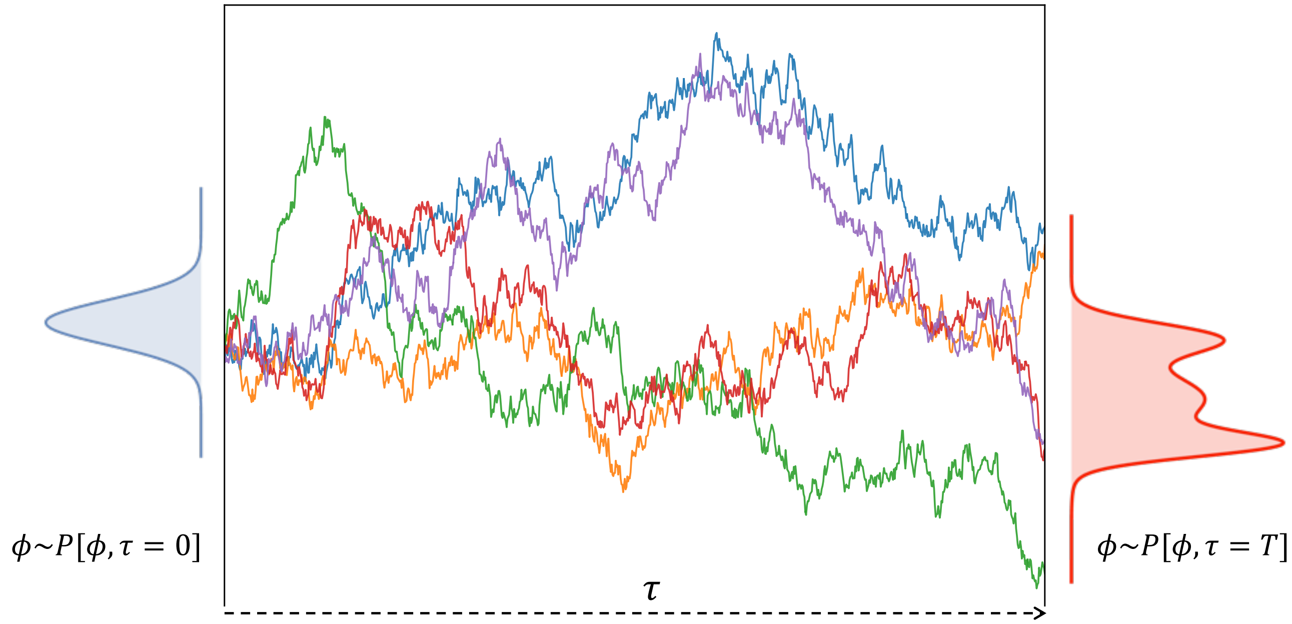
Equilibrium solution (long-time limit),

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

- **Quantum Diffusion with $\alpha = \hbar$**

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

Thermal Equilibrium Limit
→ Quantum Distribution



Trajectories

1. No need gauge-fixing!
2. Can handle Fermion fields naturally
→(Complex Langevin Method)

Recent review: Gert Aarts, Dénes Sexty, 2604.24290 [hep-lat]

Stochastic Quantization

Reverse Diffusion

LW, G Aarts, K Zhou, JHEP(2024)

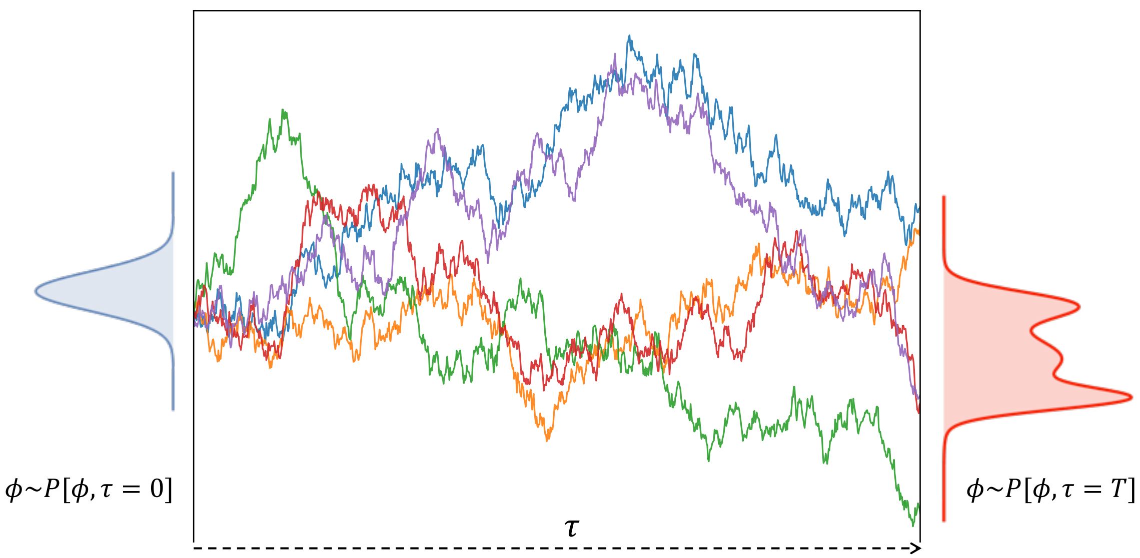
- **Diffusion Models (w/o drift term)**

$$\frac{d\phi}{dt} = -g(t)^2 \nabla_\phi \log p_t(\phi) + g(t)\bar{\eta}$$

redefine $\tau \equiv T - t (d\tau \equiv -dt)$

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

introducing **noise scale** $\langle \bar{\eta}^2 \rangle \equiv 2\bar{\alpha}$, **time scale** $g_\tau^2 \Delta\tau$



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

- **FP equation**

$$\frac{\partial p_\tau(\phi)}{\partial \tau} = \int d^n x \left\{ g_\tau^2 \bar{\alpha} \frac{\delta}{\delta\phi} \left(\frac{\delta}{\delta\phi} + \frac{1}{\bar{\alpha}} \nabla_\phi S_{DM} \right) \right\} p_\tau(\phi)$$

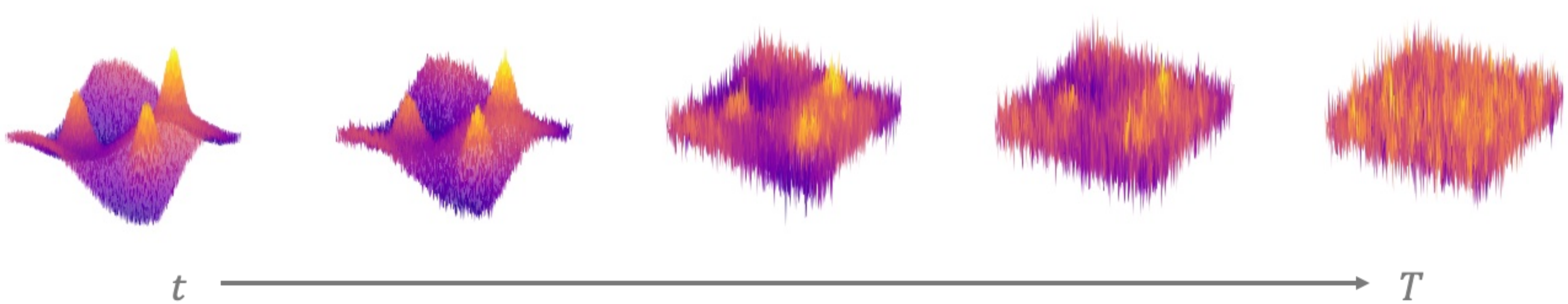
$$\boxed{\begin{gathered} \nabla_\phi S_{DM} \equiv -\bar{\alpha} \nabla_\phi \log q_\tau(\phi) \\ p_{eq}(\phi) \propto e^{-\frac{S_{DM}}{\bar{\alpha}}} \end{gathered}}$$

Reverse diffusion realizes **stochastic quantization** toward the data distribution.

DM as SQ

Aspect	Stochastic Quantization (SQ)	Diffusion Models (DM)
Drift / score	Fixed by the action $-\delta S[\phi]/\delta\phi$	Scale-dependent, learned from p_τ $s_\tau(\phi) = \nabla_\phi \log p_\tau(\phi)$
Noise kernel	Usually constant; generalized kernels possible	Scheduled and scale-dependent $g(\tau)$
Flow	Relaxation in fictitious time toward $P_{\rm eq}[\phi] \propto \exp(-S[\phi])$	Finite flow between noise and data distributions
Target	Sample the equilibrium distribution $\exp(-S[\phi])$	Learn $p_\tau(\phi)$ or equivalently its score
Equation	$\partial_\tau \phi(x,\tau) = -\dfrac{\delta S}{\delta \phi(x)} + \sqrt{2}\,\eta(x,\tau)$	Reverse-time SDE: $d\phi_\tau = \left[f(\phi_\tau,\tau) - g^2(\tau)\nabla_\phi \log p_\tau(\phi_\tau)\right] d\tau + g(\tau)\, d\bar{W}_\tau$ $\tau : T \to 0$

Fictitious Time as Noise Scale

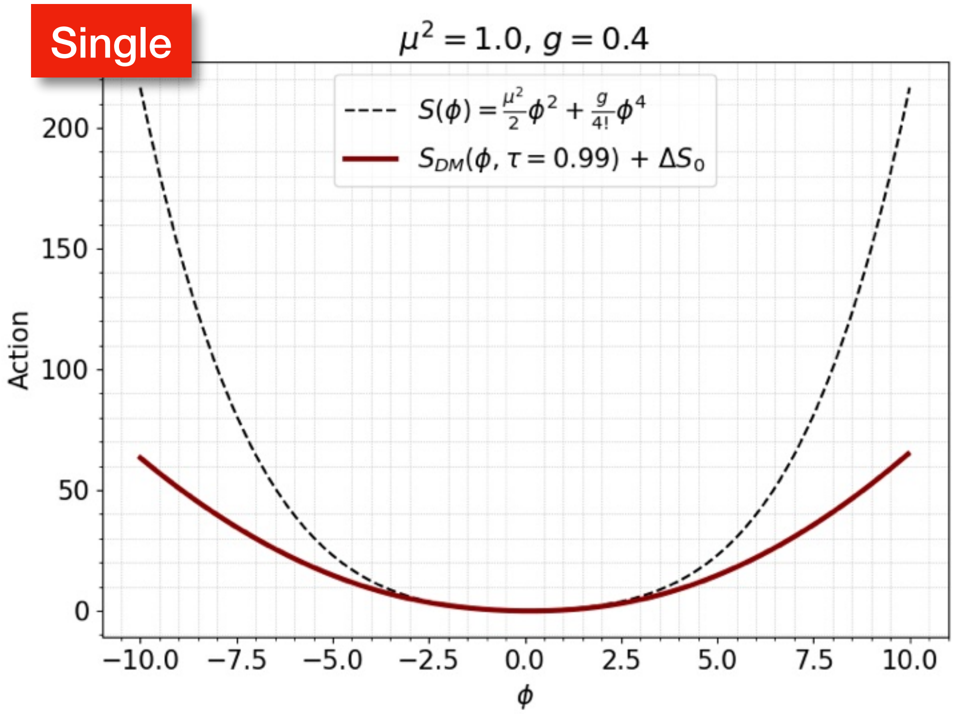


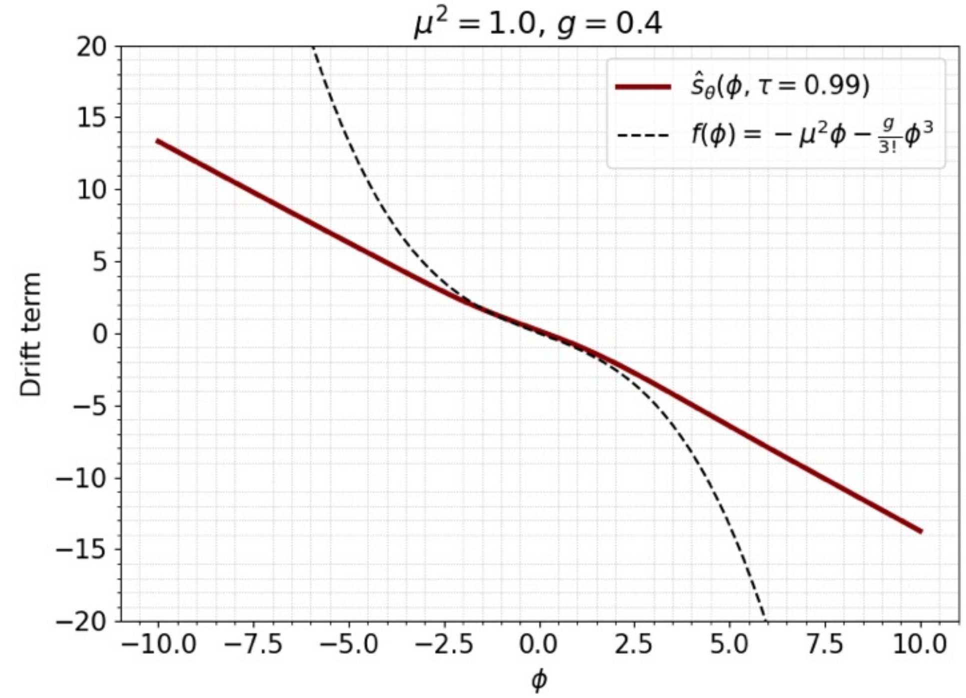
Effective Action

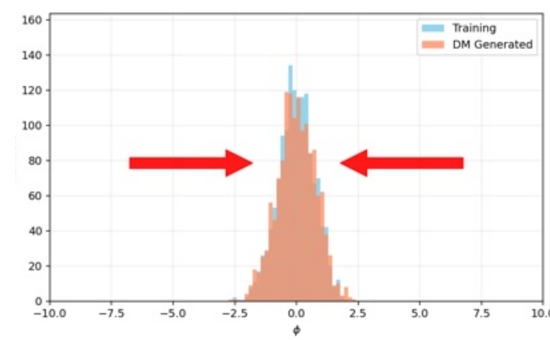
$$S(\phi) = \frac{\mu^2}{2}\phi^2 + \frac{g}{4!}\phi^4 \qquad f(\phi) = -\mu^2\phi - \frac{g}{3!}\phi^3$$

0+0 dimensional QFT (1 variable)

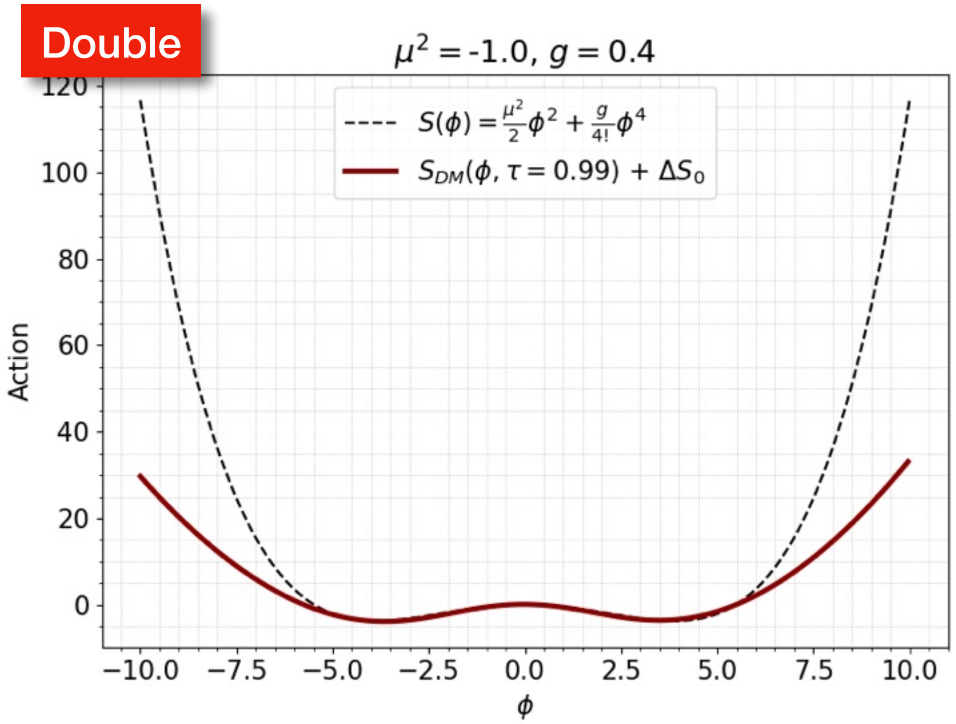
Single

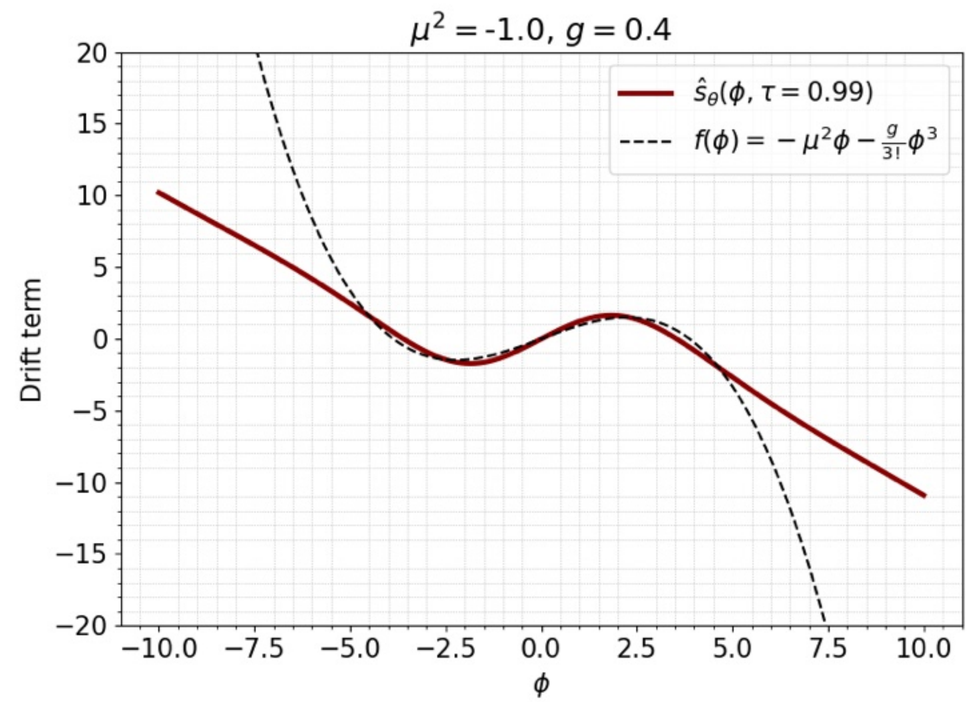


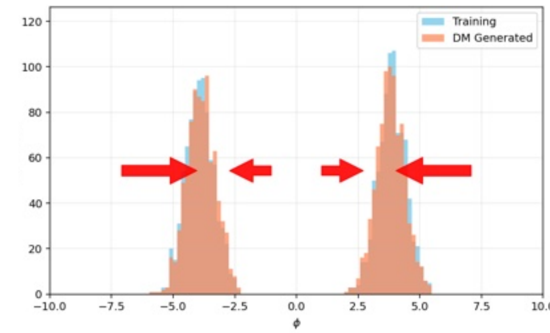




Double

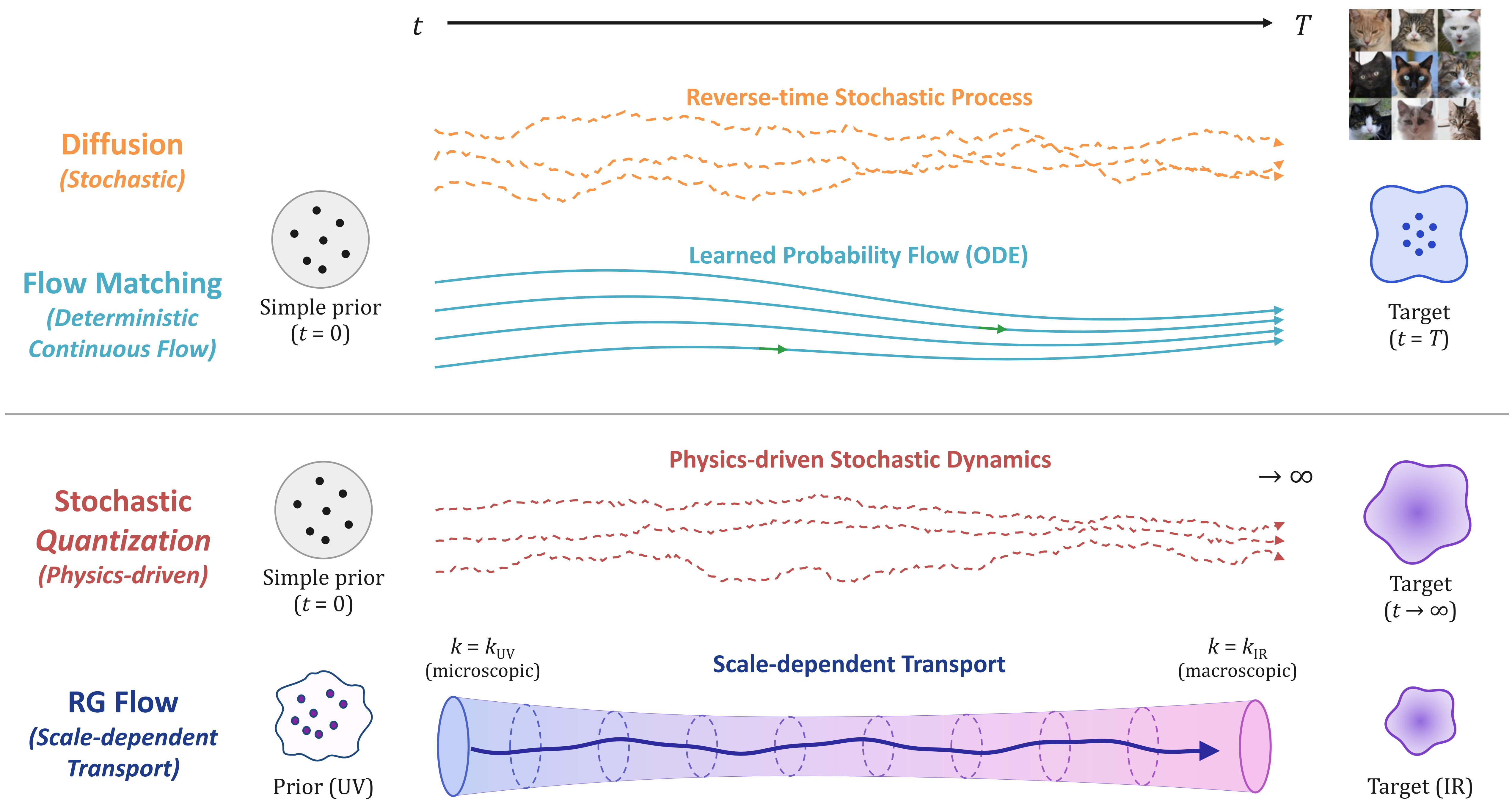






Bridge Two Distributions

[P30] Kanta's Poster
Yuto's Talk [04/09, 11:00]
Jan's Talk [04/09, 11:30]



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

Why Diffusion Models?

for Lattice simulations

Generative Models

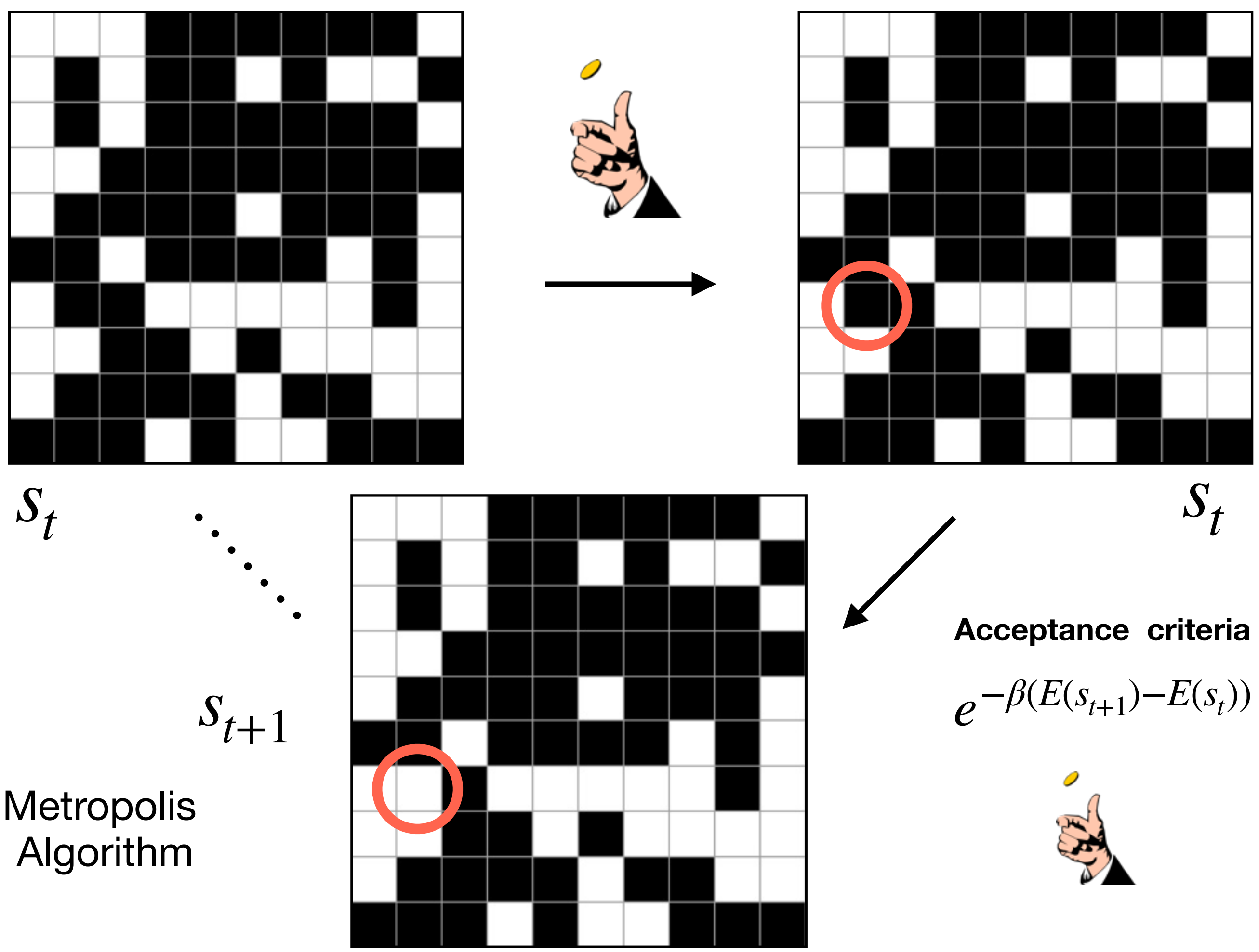
→ **Underlying Distributions in Data**

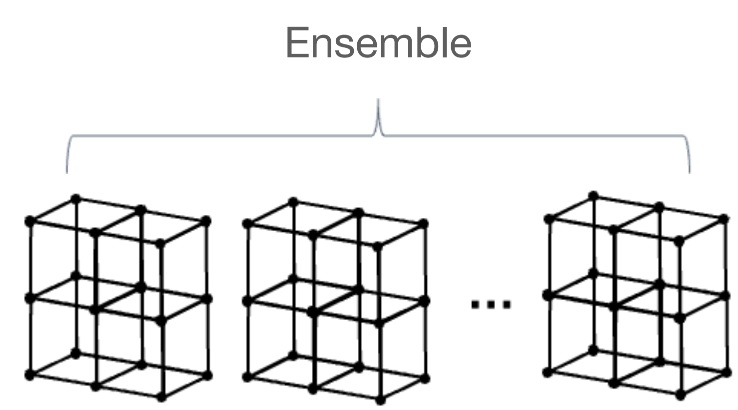
Lattice Simulations

→ **Physical Distributions, Sampling**

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

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



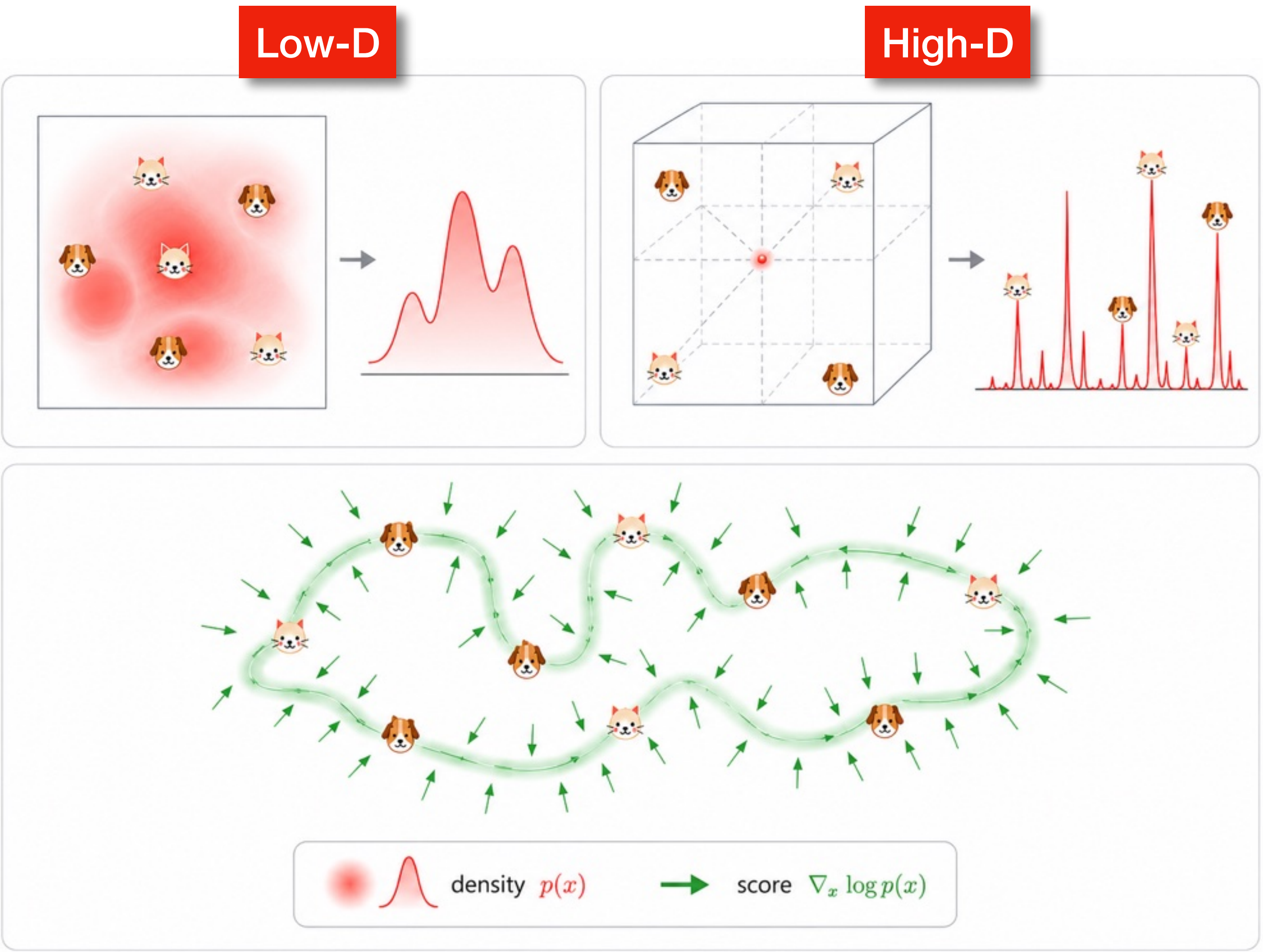


Universal and Accurate, but

- Local update, low-efficiency
- **Critical Slowing Down**
  [U. Wolff, Nucl. Phys. B 17, 93 (1990)]

Why Diffusion Models?

for Lattice simulations



Generative Models

$$q_\theta(\phi) \to p(\phi)$$

Physical Distributions, Sampling

- Implicit Likelihood Models
  Flexible but Poor Performance
  - VAEs and GANs
    [D. Giataganas, et al., NJP (2022)] [K. Zhou, et al., PRD (2019)] [J. Pawlowski and J. Urban, MLST (2020)] [J. Singh, et al., SciPost Phys. (2021)]
- Explicit Likelihood Models **Tractable, but Hard To Learn**
  - Autoregressive models
    [D. Wu, et al., PRL (2019)] [L. Wang, et al., CPL 39, 120502 (2022).] [P. Białas, P. Korcyl, and T. Stebel, CPC (2022)]
  - Normalizing flows
    [M. Albergo, G.Kanwar, and P.Shanahan, PRD (2019)] [G. Kanwar, et al., PRL (2020)] [K. A. Nicoli, et al., PRL (2021)] [L. Del Debbio, et al., PRD (2021)] [M. Caselle, et al., JHEP 2022, 15 (2022)] [S. Chen, et al., PRD (2023).] [K. Cranmer, G. Kanwar, S. Racanière, D. J. Rezende, and P. E. Shanahan, Nat Rev Phys 1 (2023)], …
- **Diffusion Models [This Talk]** **Implicit Density, Explicit Score**

Lattice Field Generations

DM-QFT

https://dm-qft.github.io/homepage

What I will (or not) introduce

What I will introduce

- Numerical results for lattice scalar field in 2-dim and 3-dim, and gauge field simulations in 2-dim and 4-dim
- Expandability of diffusion models for different lattice sizes and couplings

What I will not introduce, but can be found in Refs and GitHub pages

- Efficiency evaluations on acceptance rates and autocorrelation times
- Exactness of generations → Metropolis algorithms
- Predetermined (Forward) Diffusion Schemes

  Variance Expanding (e.g., Score-Based DM), Variance Preserving (e.g., DDPM), etc.
- Training Objective (Loss Function) and Set-Ups

  Score-Matching, $\mathscr{L}_\theta = \sum_{i=1}^{N} \sigma_i^2 \mathbb{E}_{p_0(\phi_0)} \mathbb{E}_{p_i(\phi_i|\phi_0)} \left[\left\| \mathbf{s}_\theta(\phi_i, \xi) - \nabla_{\phi_i} \log p_i(\phi_i | \phi_0) \right\|_2^2 \right]$
- Details of Neural Network for Score Functions

  CNNs, but modified for different field theories, e.g., L-CNN for gauge fields.

Lattice Scalar Fields

Scalar ϕ^4 field

- **Euclidean action** of bare fields

$$S_E = \int \mathrm{d}^d x \mathrm{d}\tau \left(\frac{1}{2}(\partial^2 \phi_0^2 + m^2 \phi_0^2) - \frac{\lambda_0}{4!}\phi_0^4 \right)$$

- Action on discrete lattice

$$S_E = \sum_x a^d [\sum_{\mu=1}^{d} \frac{(\phi_0(x+a\hat{\mu}) - \phi_0(x))^2}{a^2} + \frac{m_0^2}{2}\phi_0^2 + \frac{\lambda_0}{4!}\phi_0^4]$$

- Dimensionless form

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

$$a^{\frac{d-2}{2}}\phi_0 = (2\kappa)^{1/2}\phi$$

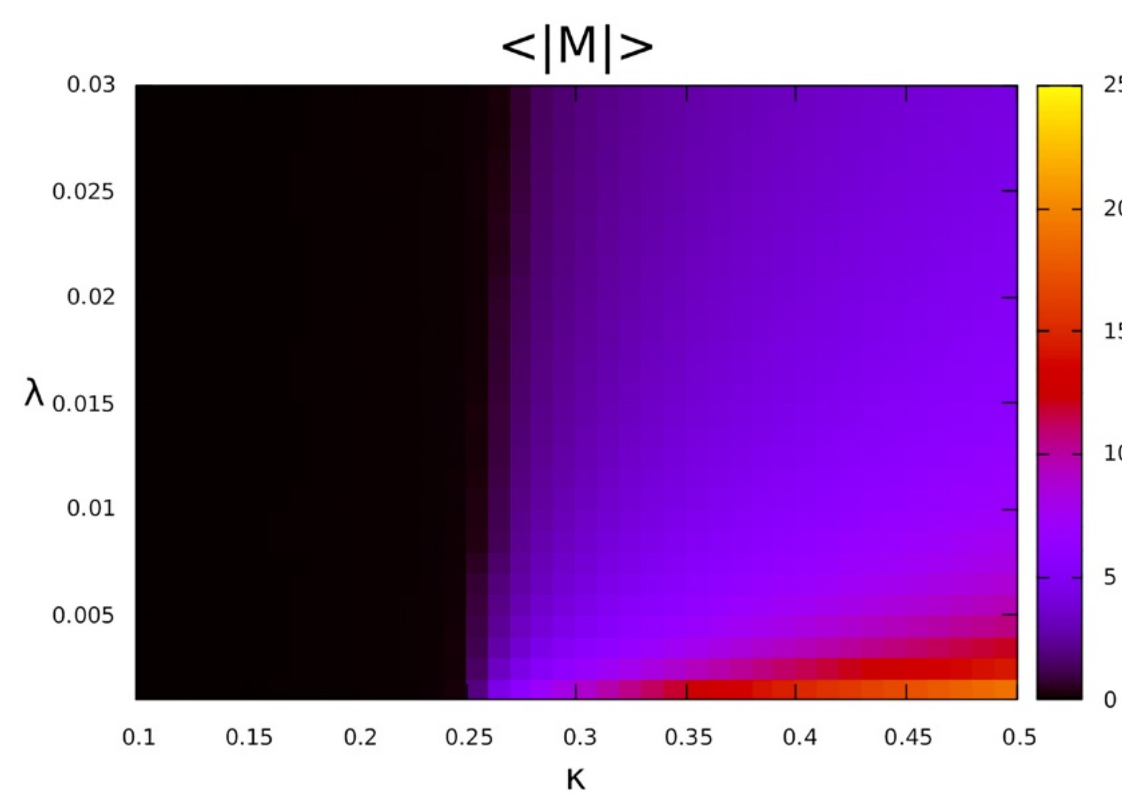
$$(am_0)^2 = \frac{1-2\lambda}{\kappa} - 2d, \quad a^{-d+4}\lambda_0 = \frac{6\lambda}{\kappa^2}$$

Hopping parameter κ, Coupling constant λ

Phase Transition

$\mathbb{Z}_2$ symmetry spontaneously broken above the critical point

Order parameter: magnetization



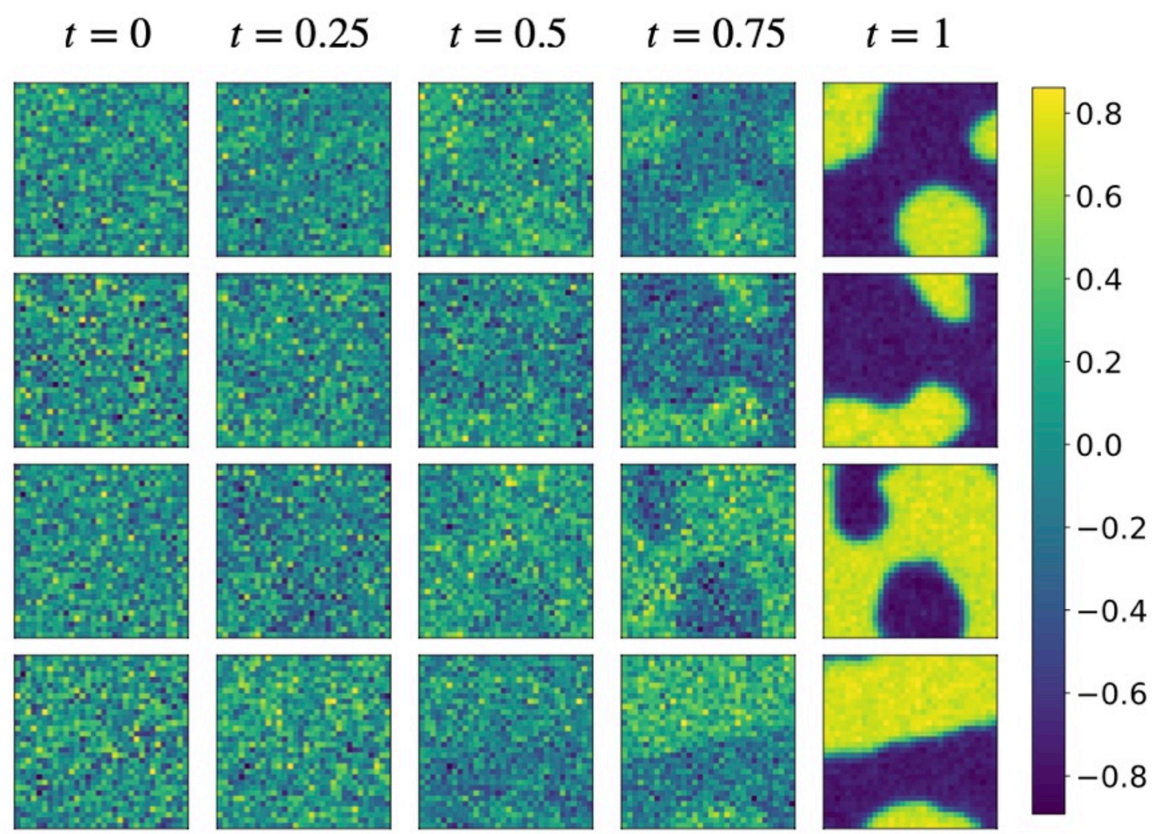
Phase diagram at $d = 2$ @Julian Urban

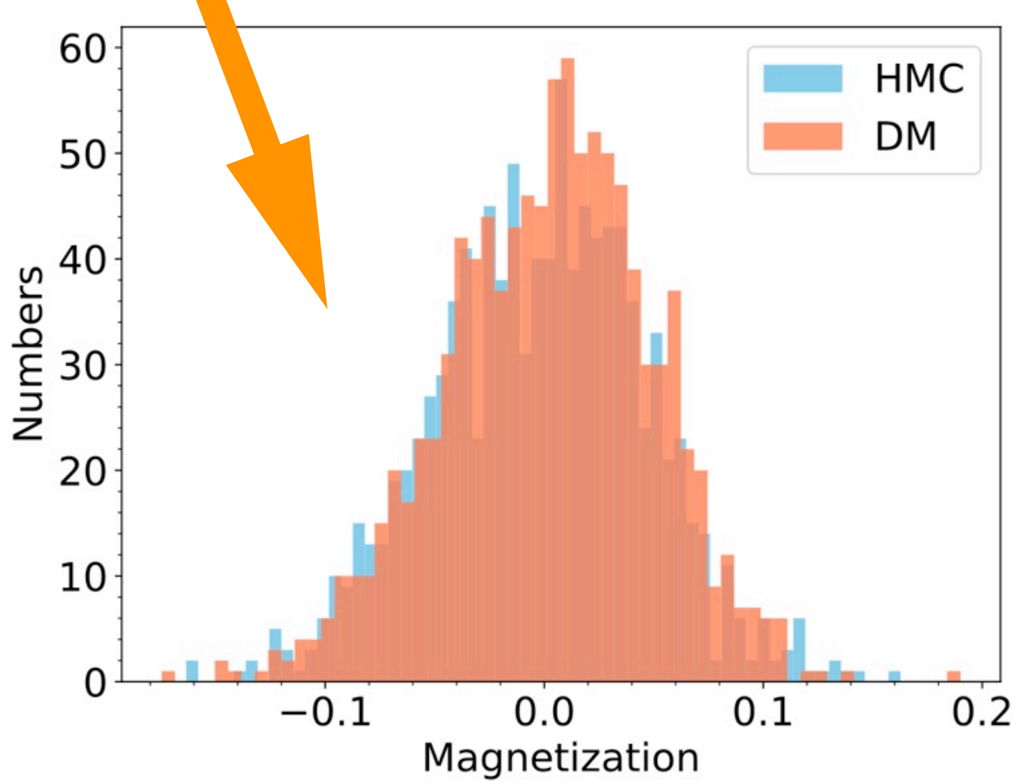
Lattice Scalar Fields

Different Phases in 2D ϕ^4

LW, G Aarts, K Zhou, JHEP(2024)

- **Diffusion Models**
  - Variance Expanding with $T = 1.0, \sigma = 25$
- **Data Generation**
  - 2D, 32×32 lattice; Hamiltonian Monte Carlo(HMC); 5120 configurations for training.
  - Broken phase: $\kappa = 1.0, \lambda = 0.022$
  - Symmetric phase: $\kappa = 0.21, \lambda = 0.022$

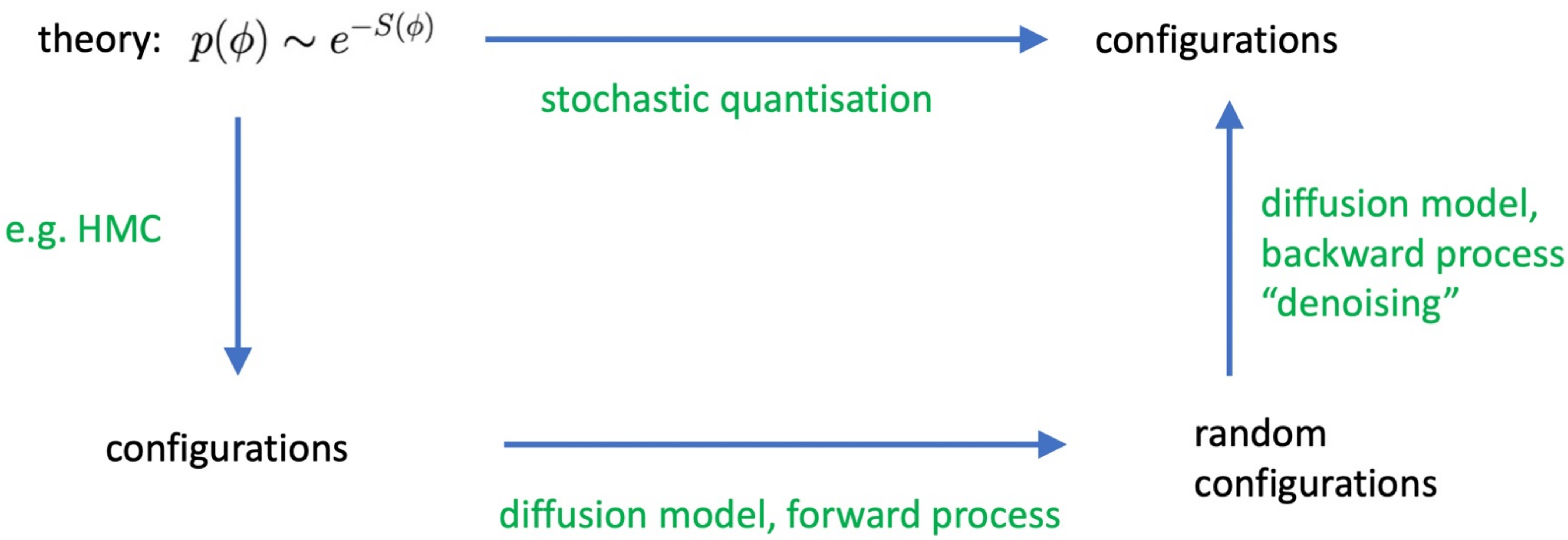




data-set	$\langle M \rangle$	χ_2	U_L
Training(HMC)	0.0012± 0.0007	2.5160 ± 0.0457	0.1042 ± 0.0367
Testing(HMC)	0.0018 ± 0.0015	2.4463 ± 0.1099	-0.0198 ± 0.1035
Generated(DM)	0.0017± 0.0015	2.4227 ± 0.1035	0.0484 ± 0.0959

Expandability

Learn the Sampler, Not Just Samples



Did we take a detour?

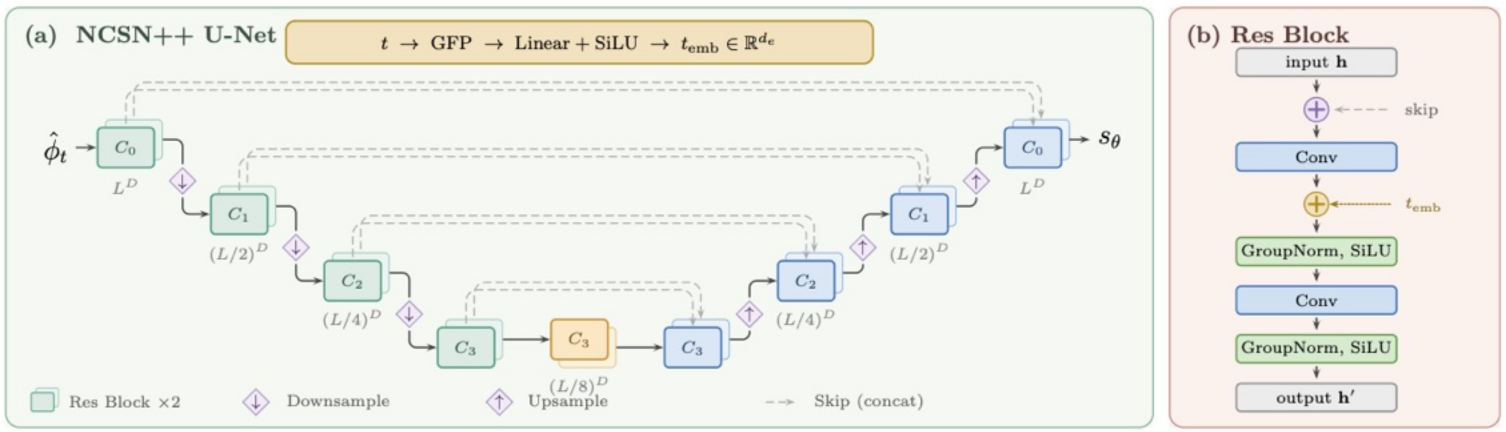
No! We learned a reusable transport map.

Train once, generate many configurations.
The learned score field can be reused or adapted across related lattice systems.

Near Critical Sampling

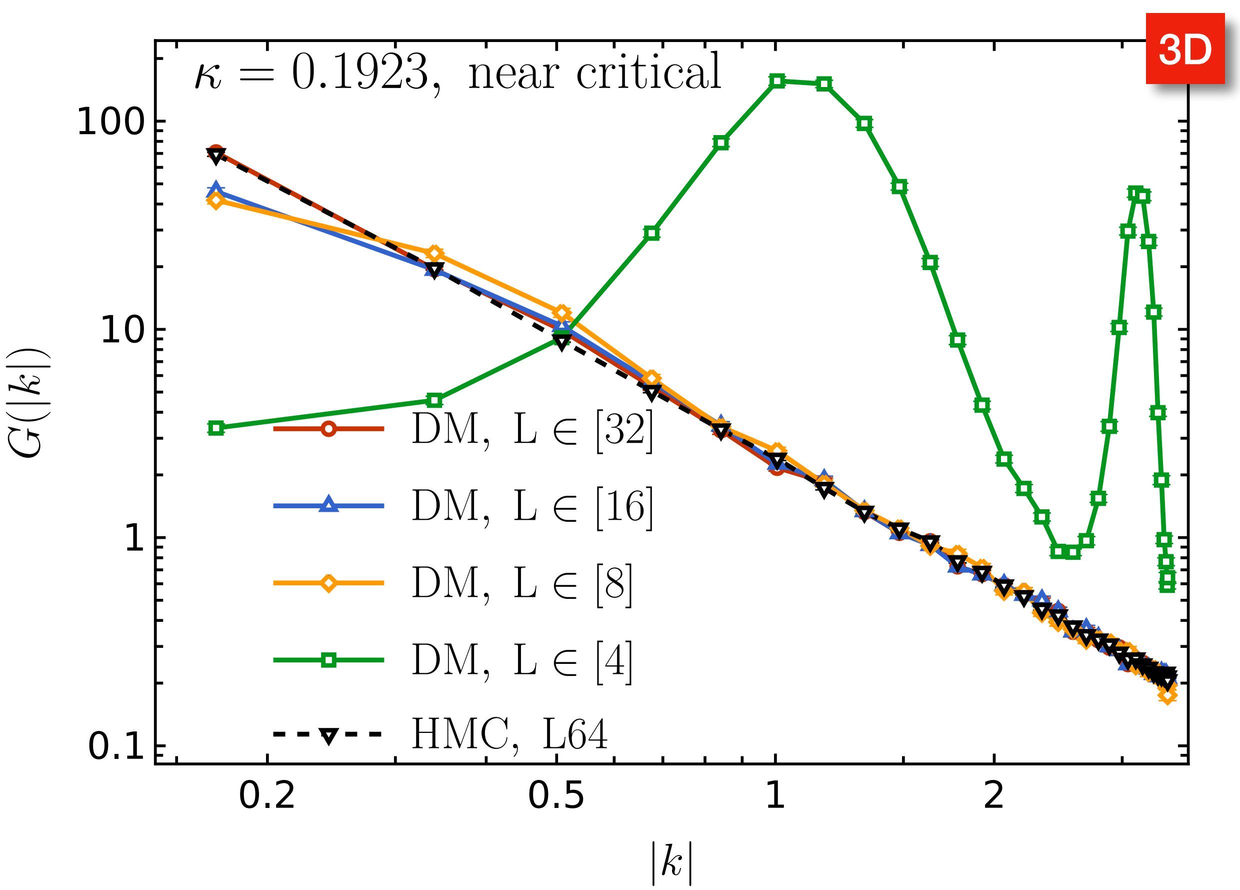
Y Tan, G Aarts, D Habibi, B Lucini, **LW*** [2607.08505 [hep-lat]]

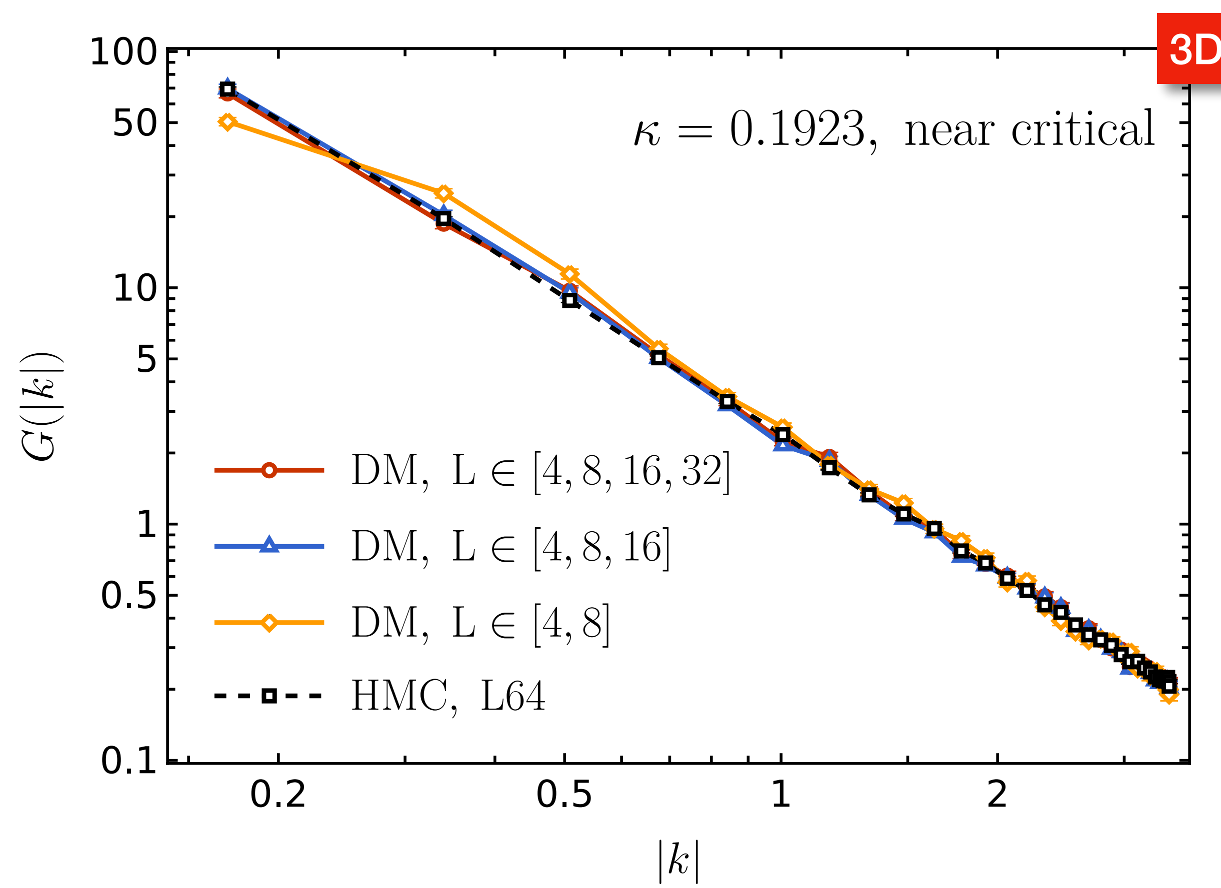
Yang-Yang Tan



Train on Small Volumes and Generate Large Ones

Local Operator $(f_\theta u)(x) = F_\theta\big(\{\, u(x+\delta) \mid \delta \in \Omega \,\}\big)$





Expandability

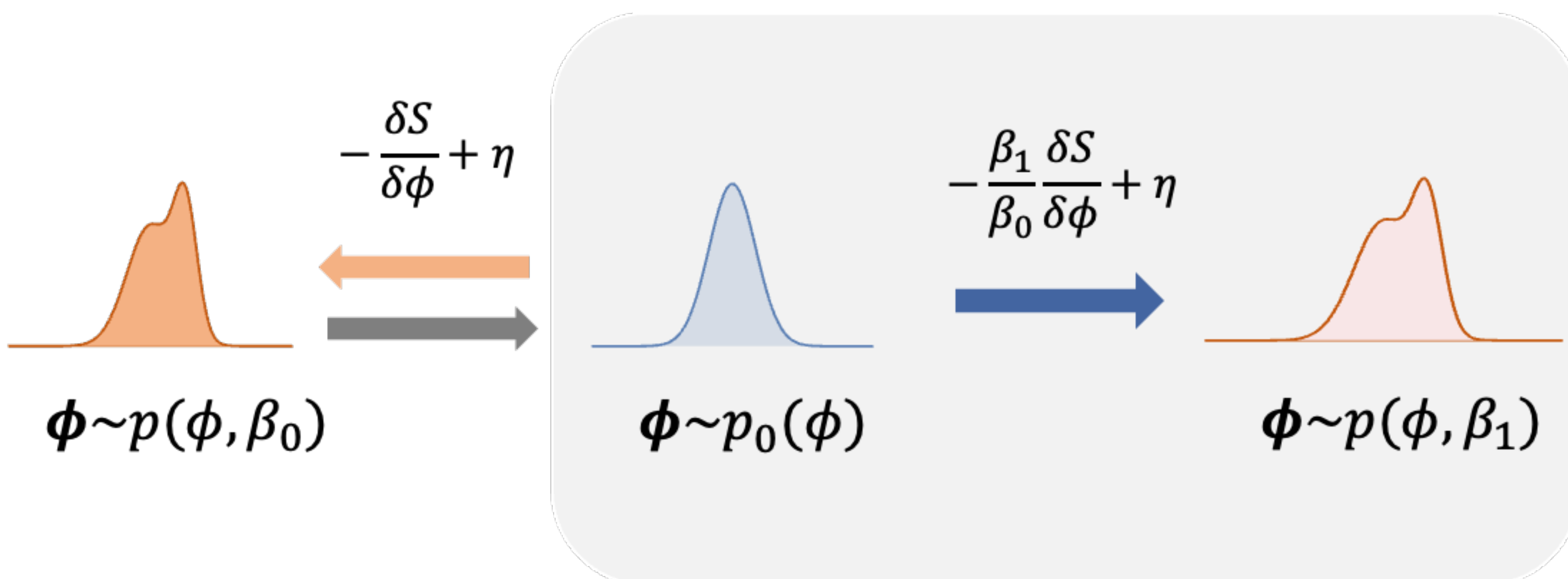
Learn the Sampler, Not Just Samples

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

Drift Term

$$\frac{d\phi}{dt} = -g^2(t)\boxed{\mathbf{s}_{\hat{\theta}}(\phi, t)} + g(t)\bar{\eta}(t)$$

Score Fucntion



e.g., pure gauge

$$S = \beta \sum_{\square} \left(1 - \mathrm{\mathbf{Re}}(U_\square)\right)$$

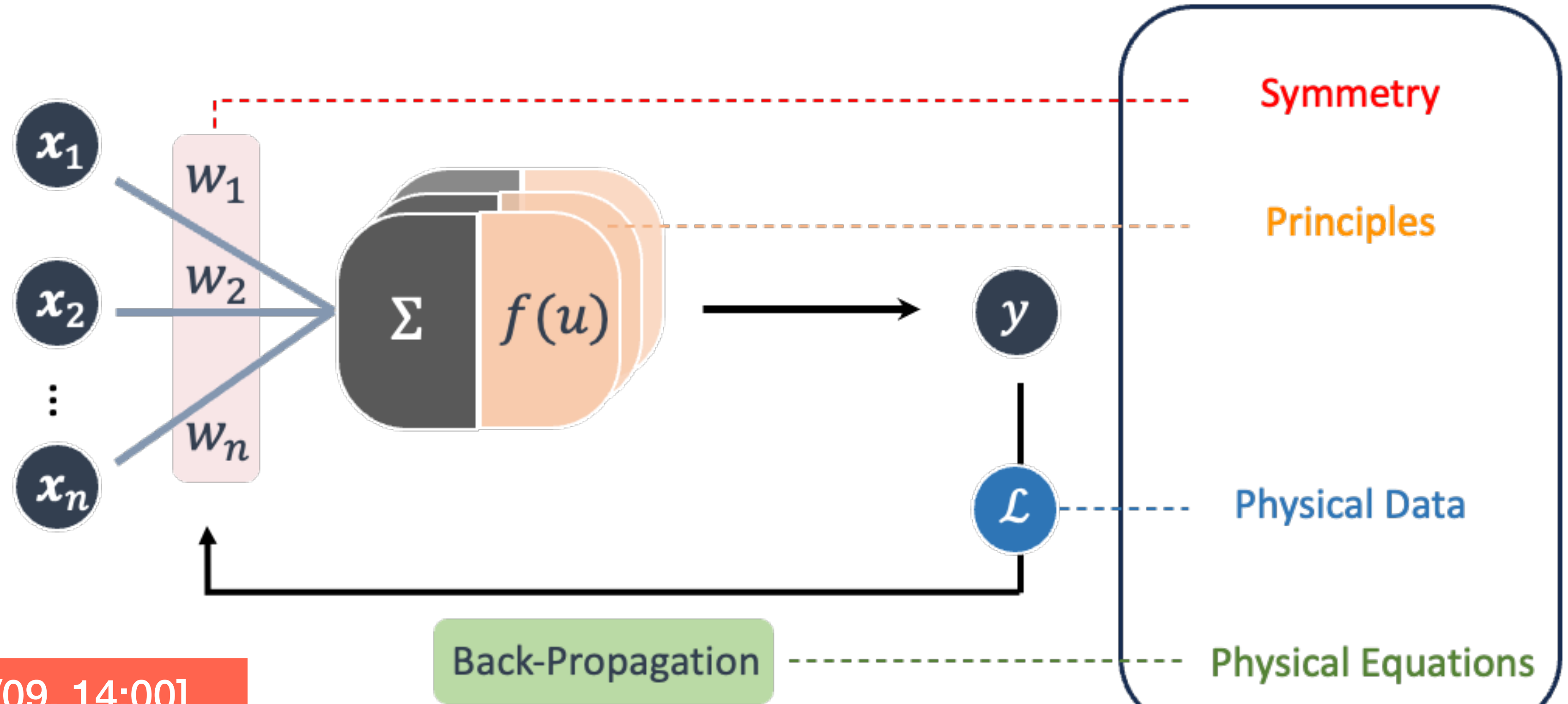
$$\tilde{\mathbf{s}}_{\hat{\theta}}(\phi, t) \equiv \beta \mathbf{s}_{\hat{\theta}}(\phi, t)$$

Expandability

Physics-Driven Learning

$$\tilde{\mathbf{s}}_{\hat{\theta}}(\phi, t) \equiv \beta \mathbf{s}_{\hat{\theta}}(\phi, t)$$

Different Lattice Sizes, Inverse Coupling Constants



Takeru's Talk [04/09, 14:00]
Yang-yang's Talk [04/09, 14:20]

G Aarts, K Fukushima, T Hatsuda, A Ipp, S Shi, **LW***, K Zhou,
Physics-Driven Learning for Inverse Problems in Quantum Chromodynamics
Nature Rev.Phys. 7 (2025) 3, 154-163, [2501.05580 [hep-lat]]

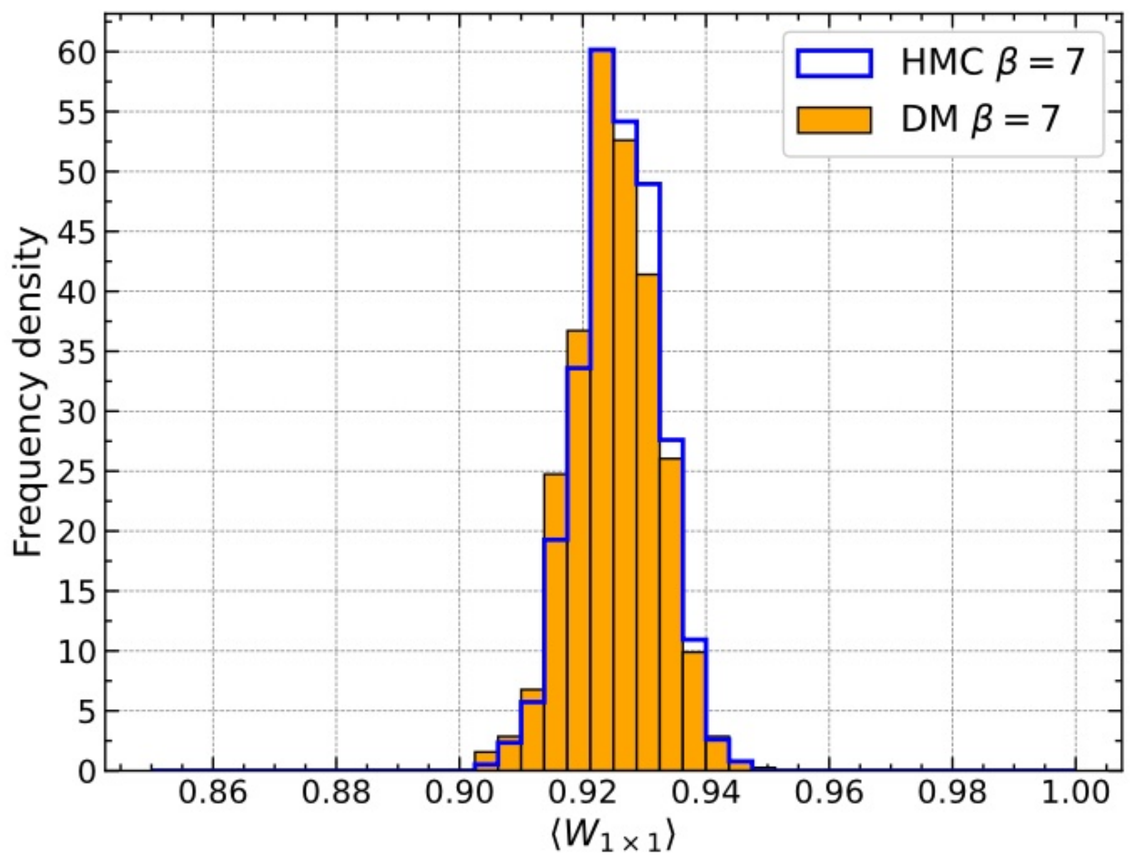
DM for U(1) Gauge Field

Q Zhu, G Aarts, W Wang, K Zhou, **LW**, JHEP(2026)

Comprehensive Comparison for 2D U(1)

Learned at $\beta = 1$ with **30k** cofigurations, L = 16

No Topological Freezing!



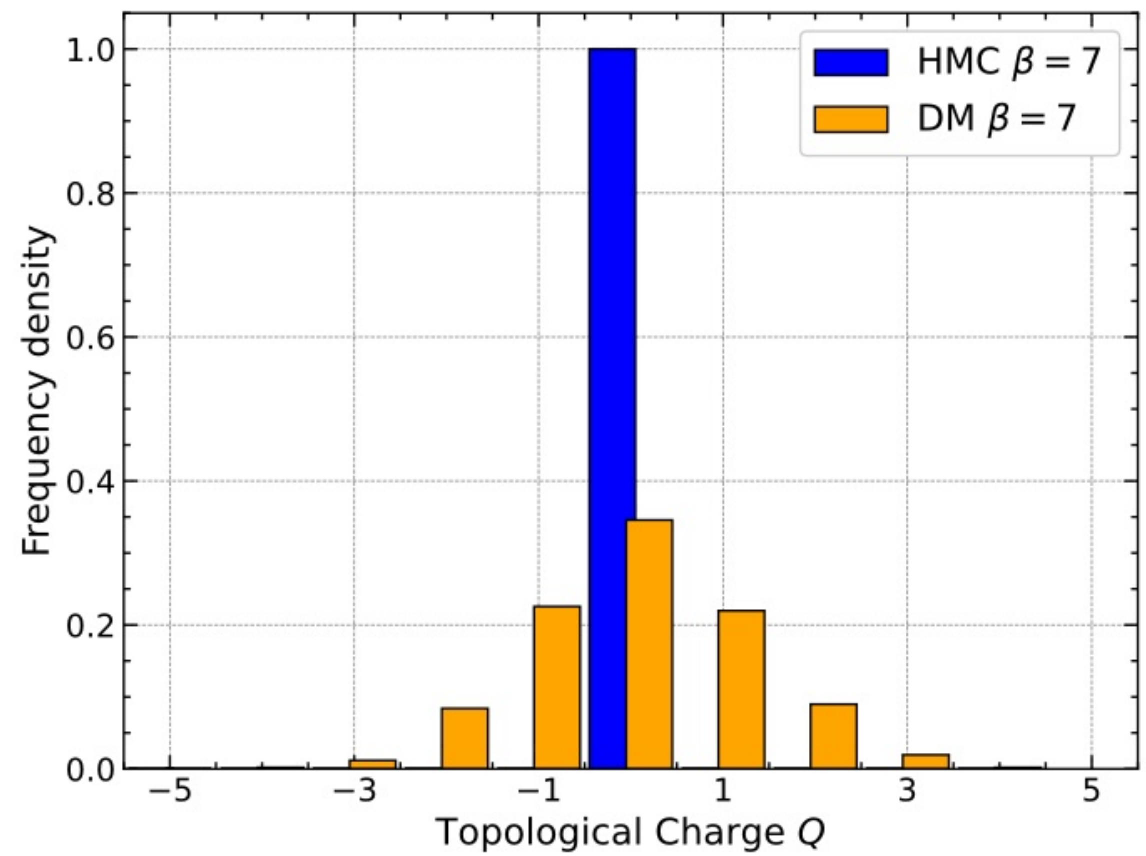


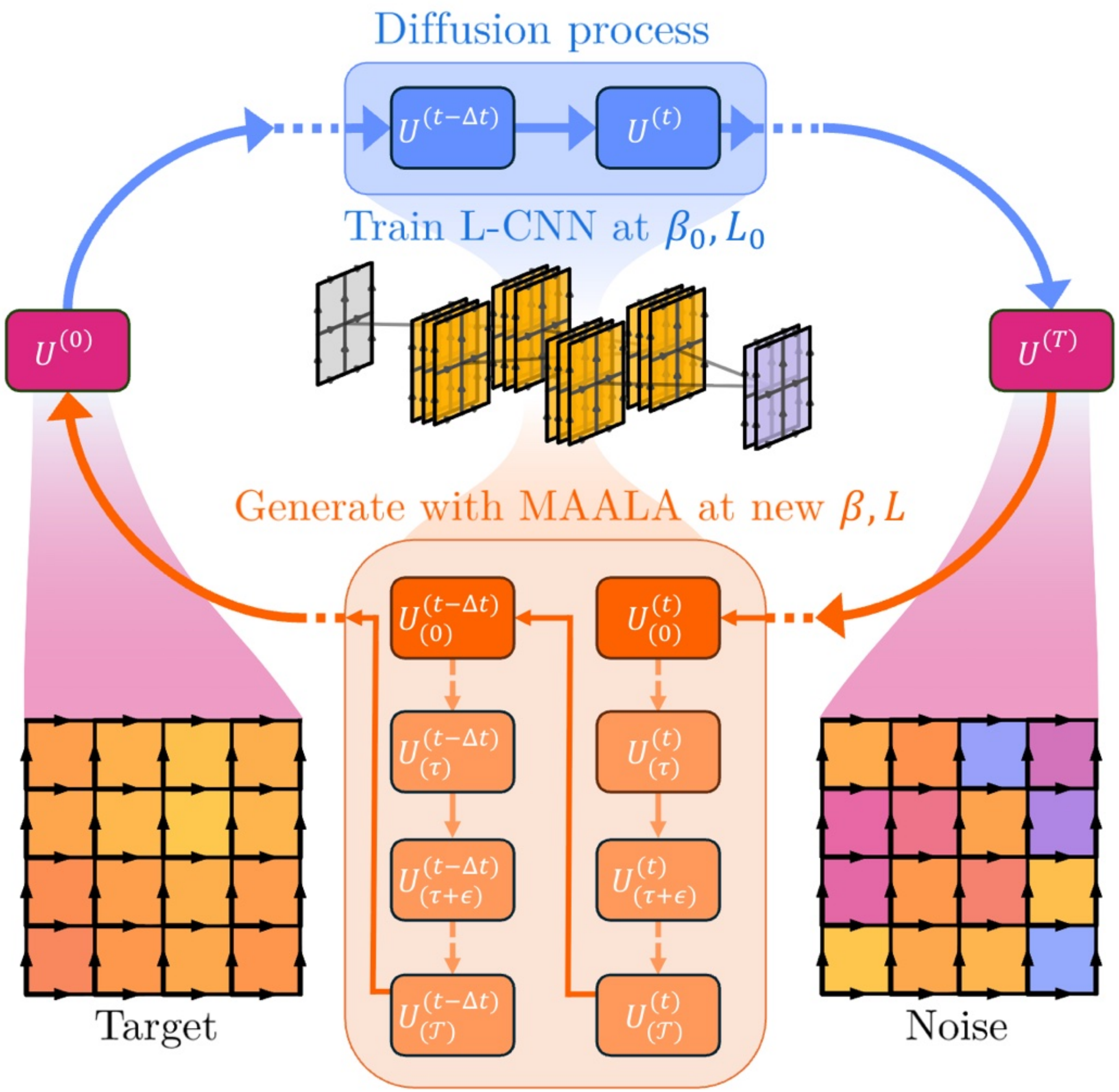
Table 2. Comparison of the $l \times l$ Wilson Loops for $L = 16, \beta = 7$

Loop size (l)	HMC	DM	Langevin	Exact
1	0.926(7)	0.926(7)	0.924(6)	0.926
2	0.737(31)	0.737(32)	0.730(34)	0.734
3	0.510(67)	0.496(72)	0.489(73)	0.498
4	0.311(97)	0.283(96)	0.283(106)	0.290

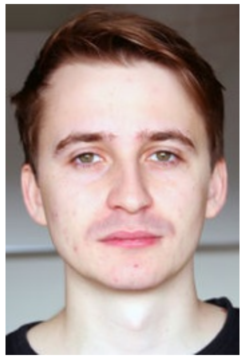
Table 3. Comparison of observables for $\beta = 7$ at different lattice sizes

Lattice Size (L)	1×1 Wilson Loop				Topological Susceptibility			
	HMC	DM	Langevin	Exact	HMC	DM	Langevin	Exact
8	0.927(13)	0.926(13)	0.921(13)	0.926	0.00006(3)	0.0040(12)	0.0143(5)	0.0040
16	0.926(7)	0.926(7)	0.924(6)	0.926	0.00013(2)	0.0045(5)	0.0131(5)	0.0039
32	0.926(3)	0.925(4)	0.924(4)	0.926	0.00013(2)	0.0040(4)	0.0137(6)	0.0039

Lattice U(N)/SU(N) Gauge Fields



Thomas R. Ranner



David Müller

Qianteng Zhu

G Aarts, D Habibi, A Ipp, D, Müller, T Ranner, **LW**, W Wang, Q Zhu, arXiv:2601.19552 [hep-lat]

DM for U(N)/SU(N) Gauge Fields

Group-Preserving Forward Process

T^a group generators, a color indices

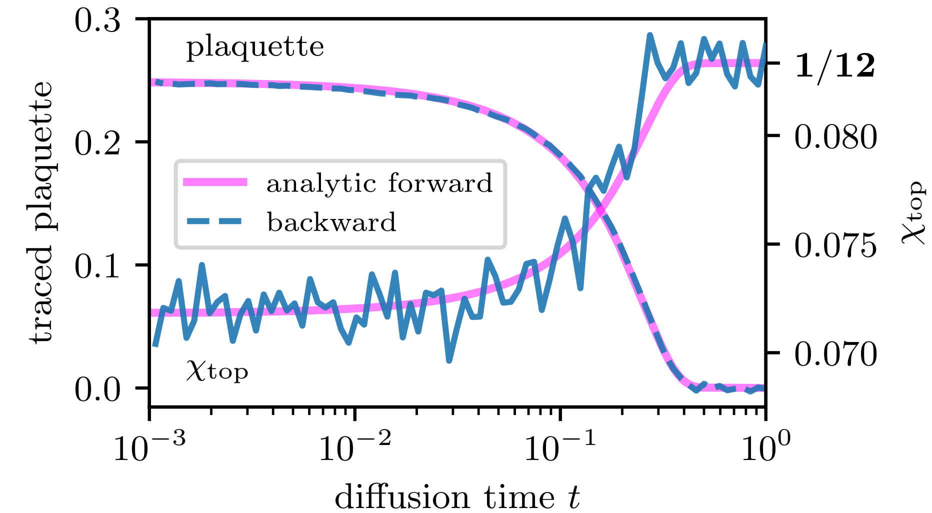
$$\mathrm{d}U^{(t)}_{x,\mu} = i\sum_a T^a\Big[K^a_{x,\mu}(\mathbf{U}^{(t)})\mathrm{d}t + g(t)\mathrm{d}W^{(t),a}_{x,\mu}\Big]\circ U^{(t)}_{x,\mu}$$

$$K^a_{x,\mu}=0 \quad\Big\downarrow\quad \eta^a_{x,\mu}\sim\mathcal{N}(0,1), \qquad \sigma(t)=\sqrt{(s^{2t}-1)/(2\ln s)}$$

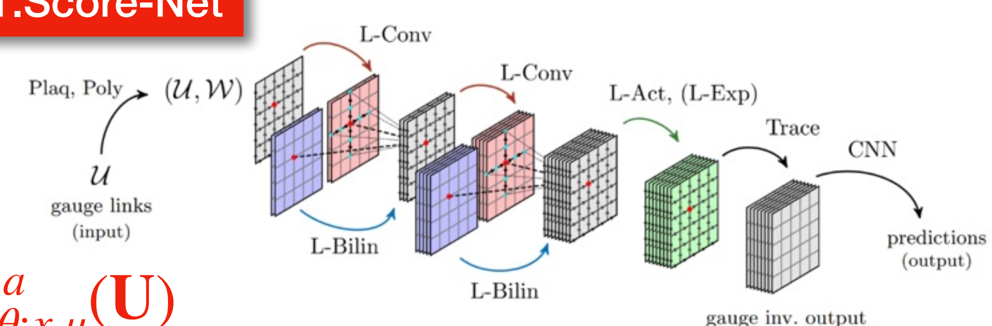
$$U^{(t)}_{x,\mu} = \exp\Big[i\sigma(t)\sum_a T^a\eta^a_{x,\mu}\Big]U^{(0)}_{x,\mu}$$

Loss Function

$$\int_0^T \mathrm{d}t \underset{\substack{p_0(\mathbf{U}^{(0)})\\ p_{0t}(\mathbf{U}^{(t)}\,|\,\mathbf{U}^{(0)})}}{\mathbb{E}} \left[\sum_{x,\mu,a}\left(s^a_{\theta;x,\mu}(\mathbf{U}^{(t)},t)+\frac{\eta^a_{x,\mu}}{\sigma(t)}\right)^2\right]$$



1.Score-Net



$s^a_{\theta;x,\mu}(\mathbf{U})$

Favoni et al., "Lattice Gauge Equivariant Convolutional Neural Networks"

2.MAALA

$$U^{(t)}_{(\tau+\epsilon)} = \exp\left[i\left(\epsilon\frac{\beta}{\beta_0}\hat{s}_\theta(\mathbf{U}^{(t)}_{(\tau)},t)+\sqrt{2\epsilon}\,\hat{\eta}\right)\right]U^{(t)}_{(\tau)}$$

$$p_{\textbf{accept}} = \min\left\{1,\frac{p(\tilde{\mathbf{U}}^{(t)}_{(\tau)})}{p(\mathbf{U}^{(t)}_{(\tau)})}\frac{q(\mathbf{U}^{(t)}_{(\tau)}\,|\,\tilde{\mathbf{U}}^{(t)}_{(\tau)})}{q(\tilde{\mathbf{U}}^{(t)}_{(\tau)}\,|\,\mathbf{U}^{(t)}_{(\tau)})}\right\}$$

DM for U(N)/SU(N) Gauge Fields

Group-Preserving Forward Process

T^a group generators, a color indices

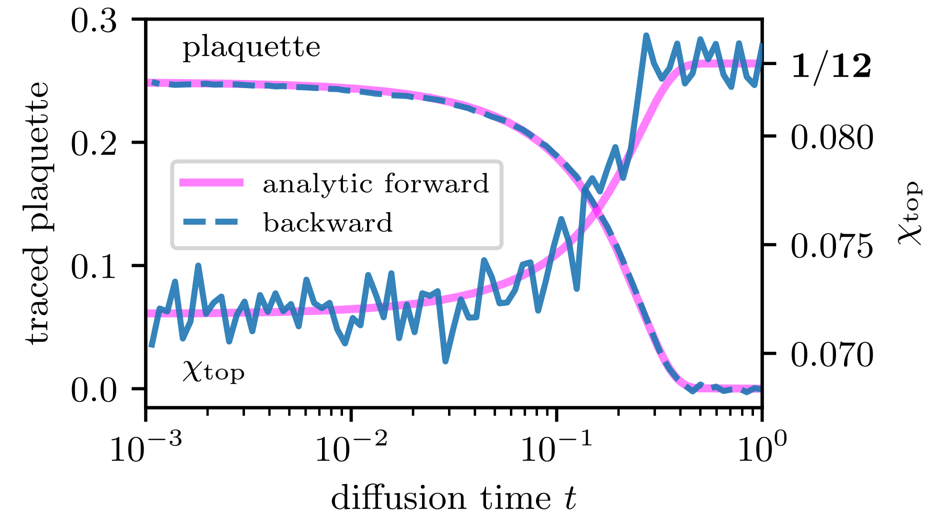
$$\mathrm{d}U_{x,\mu}^{(t)} = i\sum_a T^a\Big[K_{x,\mu}^a(\mathbf{U}^{(t)})\mathrm{d}t + g(t)\mathrm{d}W_{x,\mu}^{(t),a}\Big]\circ U_{x,\mu}^{(t)}$$

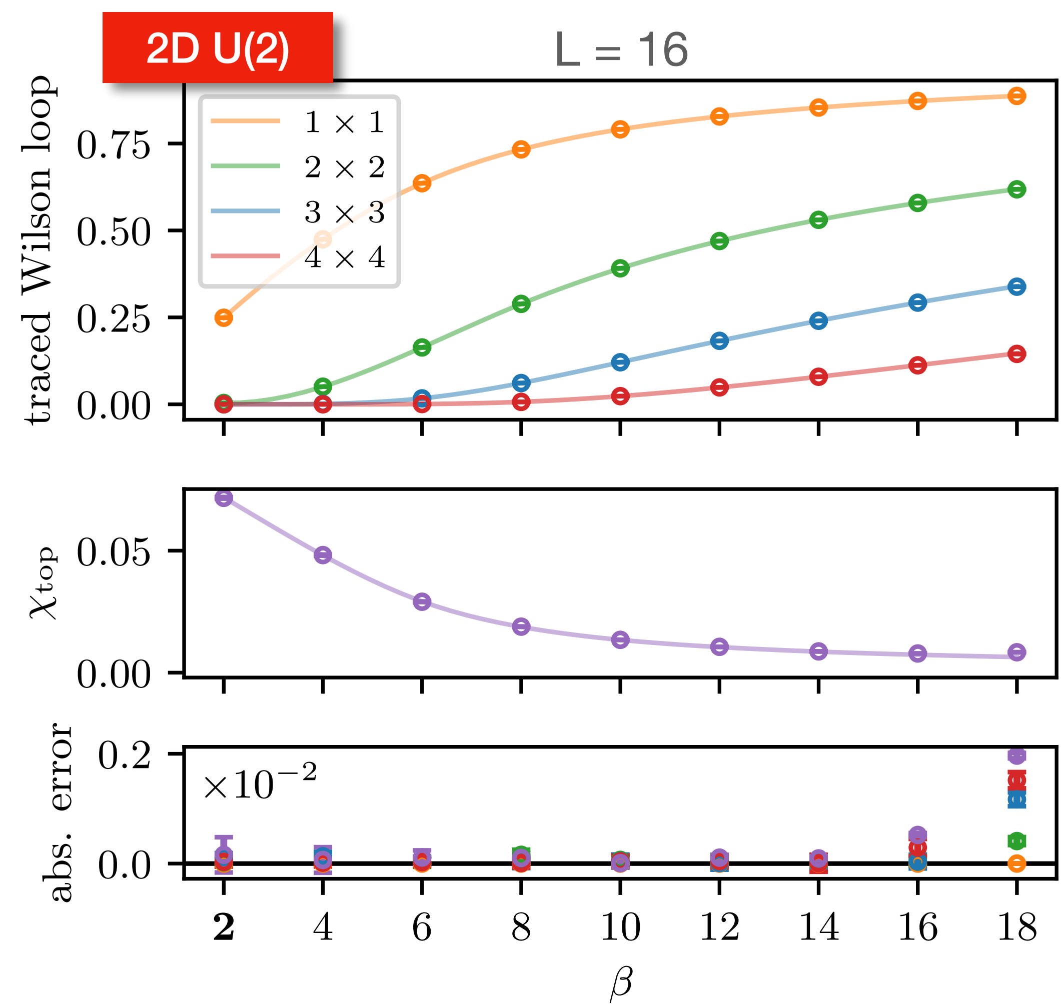
$$K_{x,\mu}^a = 0 \quad \Big\downarrow \quad \eta_{x,\mu}^a \sim \mathcal{N}(0,1), \quad \sigma(t) = \sqrt{(s^{2t}-1)/(2\ln s)}$$

$$U_{x,\mu}^{(t)} = \exp\Big[i\sigma(t)\sum_a T^a\eta_{x,\mu}^a\Big]U_{x,\mu}^{(0)}$$

Loss Function

$$\int_0^T \mathrm{d}t \underset{\substack{p_0(\mathbf{U}^{(0)})\\ p_{0t}(\mathbf{U}^{(t)}\,|\,\mathbf{U}^{(0)})}}{\mathbb{E}} \left[\sum_{x,\mu,a}\left(s_{\theta;x,\mu}^a(\mathbf{U}^{(t)},t)+\frac{\eta_{x,\mu}^a}{\sigma(t)}\right)^2\right]$$

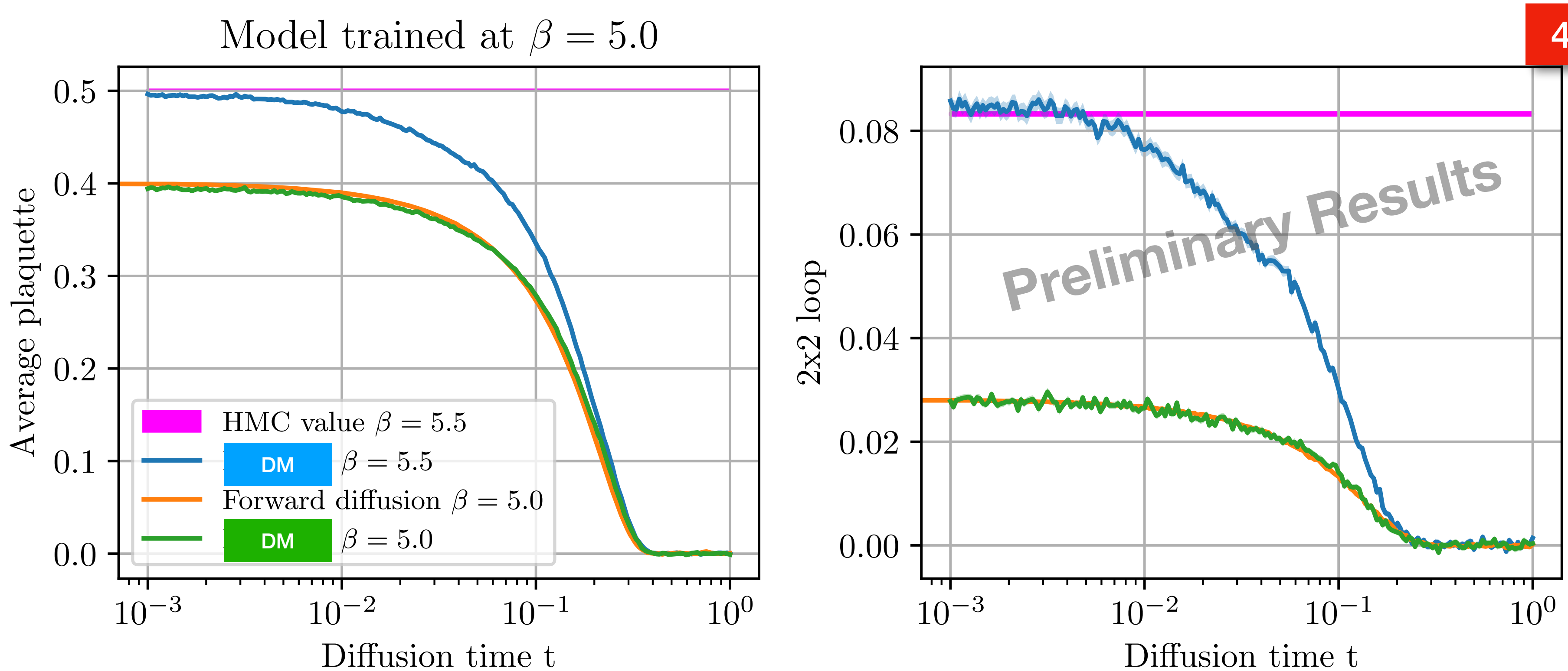




DM for U(N)/SU(N) Gauge Fields

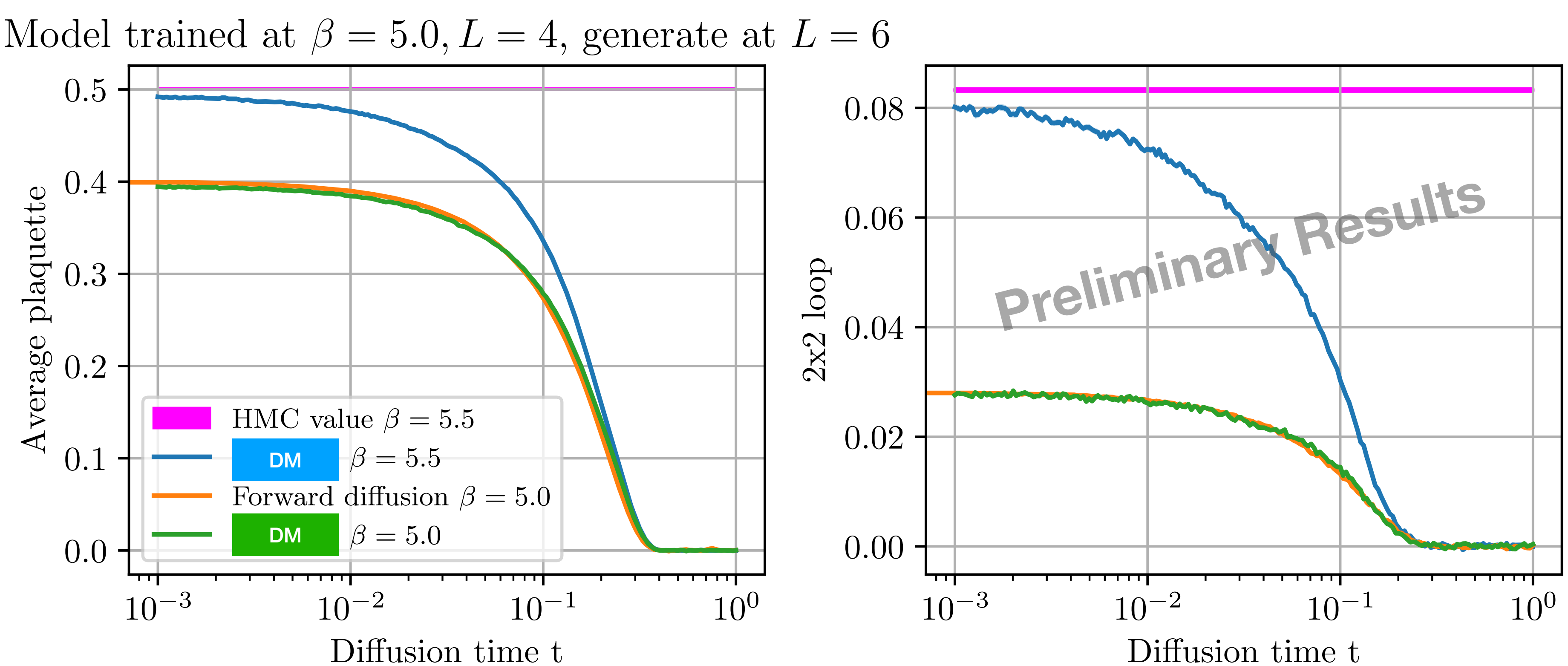
DM-QFT [260x.xxx [hep-lat]]

4D SU(3)



$V = 4^4, \beta_0 = 5$

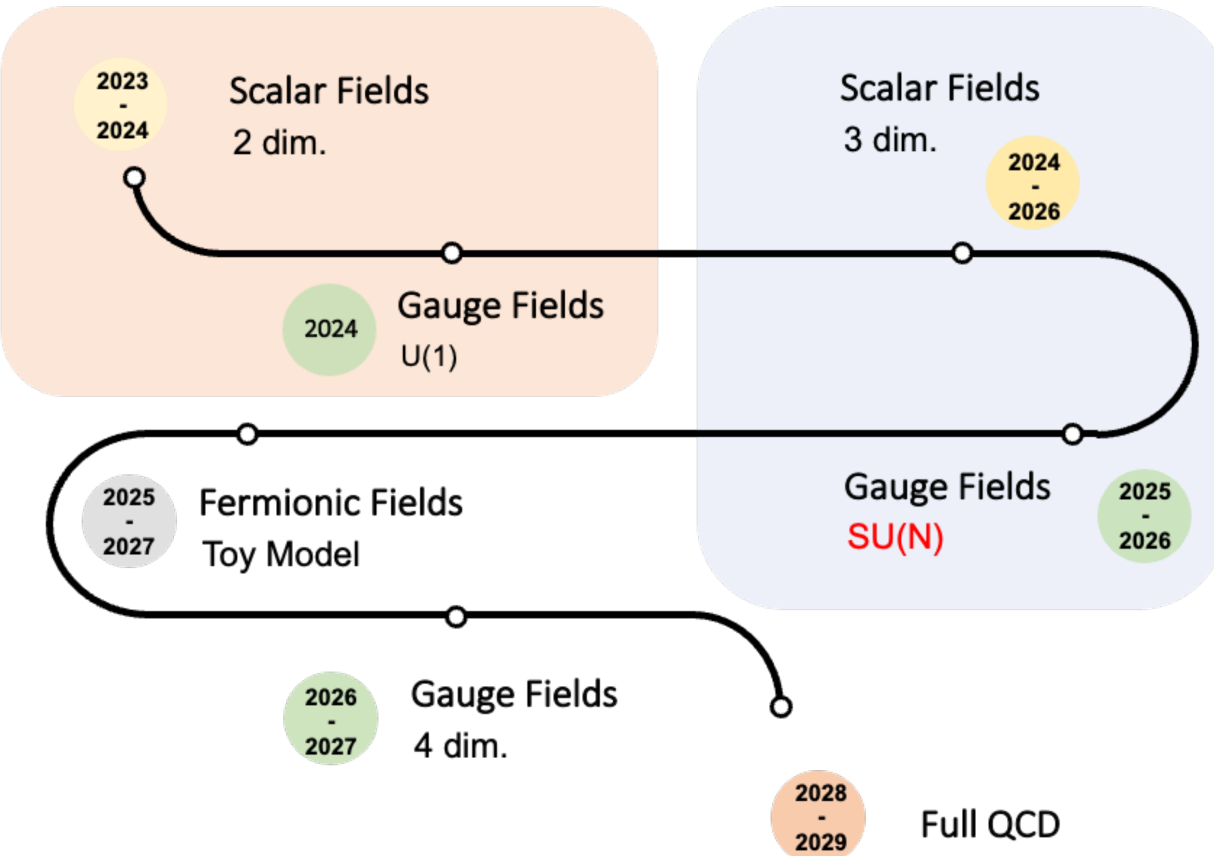
$V = 4^4, \beta = 5.5$

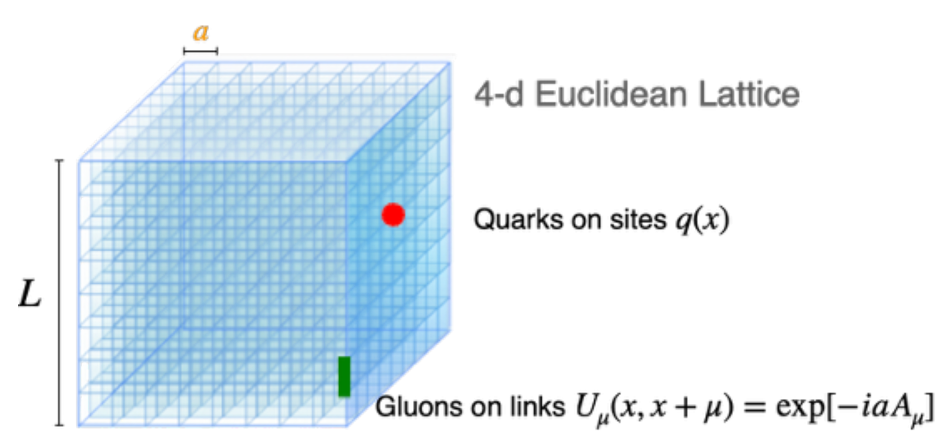


$V = 6^4, \beta = 5.5$

Roadmap toward full QCD

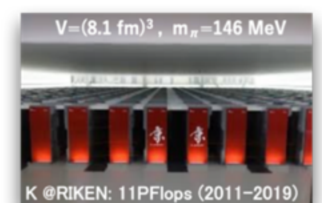
Where shall we go

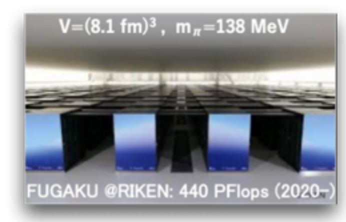




Faster and Reliable Simulation

- **Systematical 4D Gauge Field Simulations**
- **Evaluations and Improvements**
  - Faster Generation
- **Fermions and Sign Problem**
  - Complex Langevin Method, etc.





Simulations	Lattice QCD
Size	96×96×96×96
Supercomputer	Fugaku
Resource Consumption	**512 cores**
Time Consumption	**~ 1 day/configuration**

3-5 years to 1-2 days

Optimal Stochastic Quantization

LW [2607.21436 [hep-lat]]

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

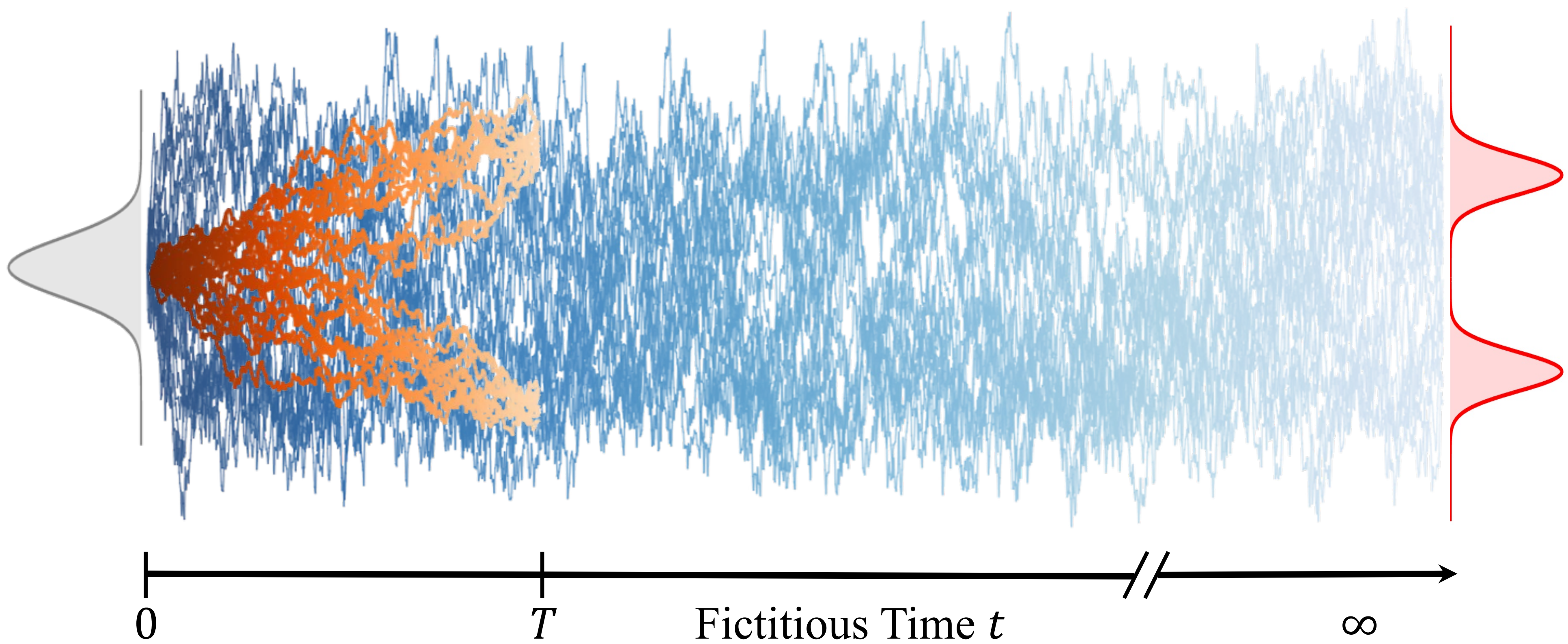
$$d\phi_t = u_\theta(\phi_t, t)dt + g dw_t, \quad t \in [0,T]$$

Physics-driven Decomposition

$$u_\theta(\phi_t, t) \equiv -\delta S_{ref}/\delta\phi + v_\theta(\phi_t, t)$$

Optimization Objective

$$\mathscr{L}(\theta) = \mathbb{E}_{\phi_0 \sim p_0}\left[\frac{1}{2}\int_0^T g^{-2}|v_\theta(\phi_t, t)|^2 dt + S_E(\phi_T)\right]$$



Optimal Stochastic Quantization

LW [2607.21436 [hep-lat]]

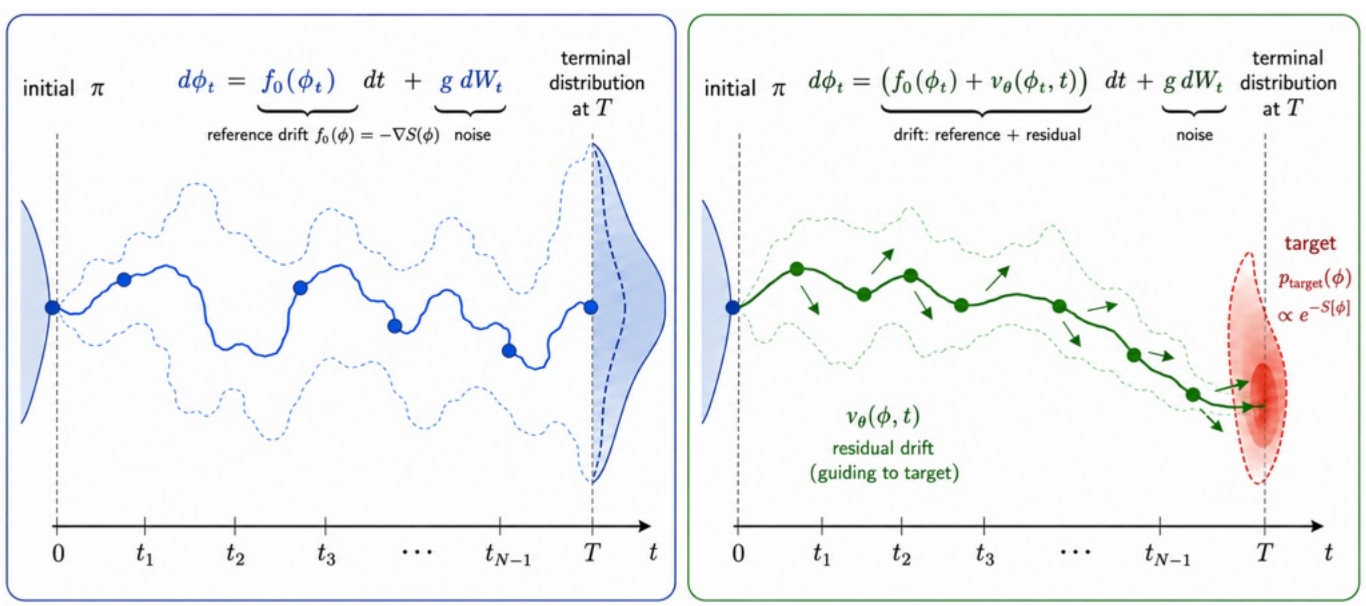
Learn a Residual Force for Finite-Time Equilibration

$$d\phi_t = u_\theta(\phi_t, t)dt + g dw_t, \quad t \in [0,T]$$

Physics-driven Decomposition

$$u_\theta(\phi_t, t) \equiv -\delta S_{ref}/\delta\phi + v_\theta(\phi_t, t)$$

$$\pi \xrightarrow{f_0 + v_\theta,\ 0 \le t \le T} q_T^\theta \approx p_{\text{target}}$$



Optimization Objective

$$\mathcal{L}(\theta) = \mathbb{E}_{\phi_0 \sim p_0}\left[\frac{1}{2}\int_0^T g^{-2}\,|v_\theta(\phi_t, t)|^2 dt + S_E(\phi_T)\right]$$

Running cost: keep close to solvable reference
Terminal cost: drive endpoint to low-action region

No density q_t^θ, **no score** $\nabla_\phi \log q_t^\theta$, **no divergence** $\nabla_\phi \cdot v_\theta$

$$v^*(\phi, t) = g^2 \nabla_\phi \log h_t(\phi)$$

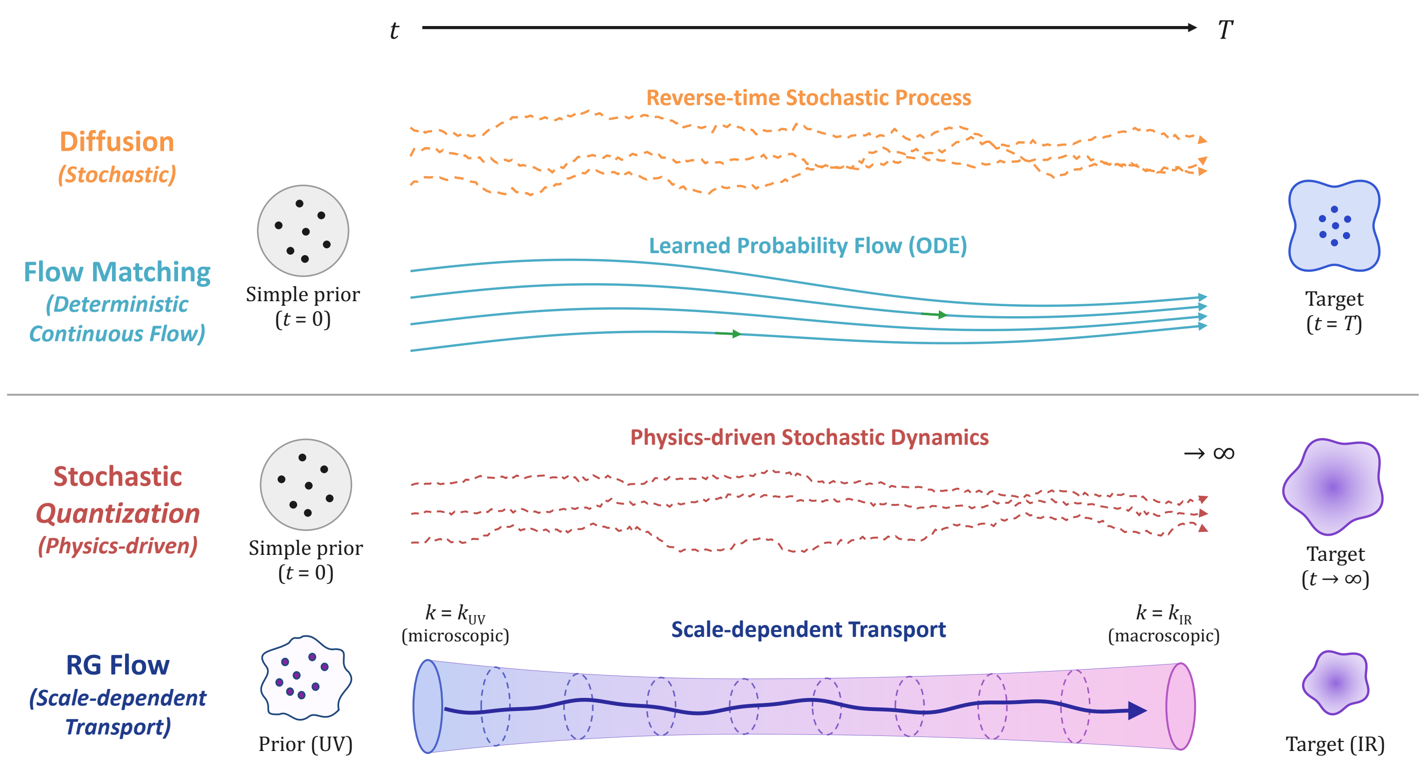
$$h_t(\phi) = \mathbb{E}_{\text{ref}}\left[e^{-S_E[\phi_T]} \mid \phi_t = \phi\right]$$

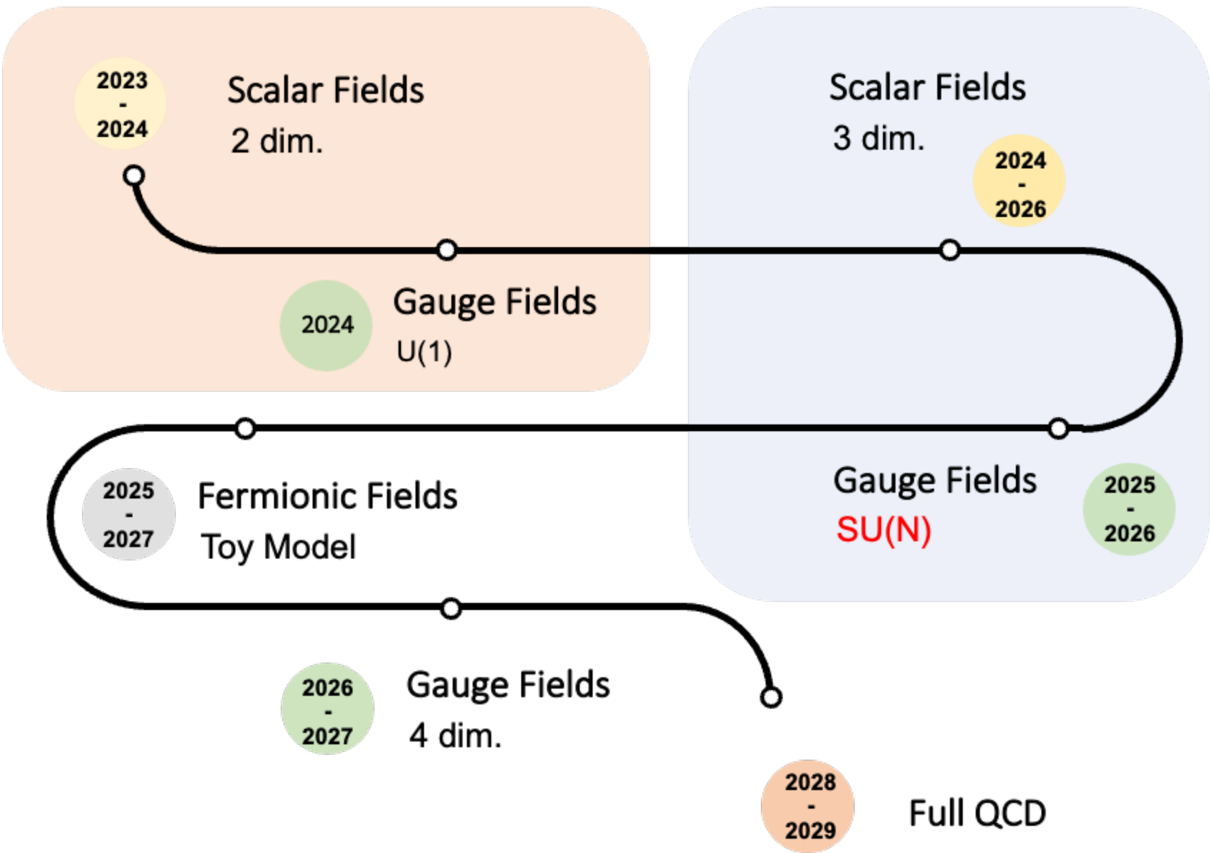
Optimal residual drift is a **Doob-transform force** which comes from expected future Boltzmann weight.

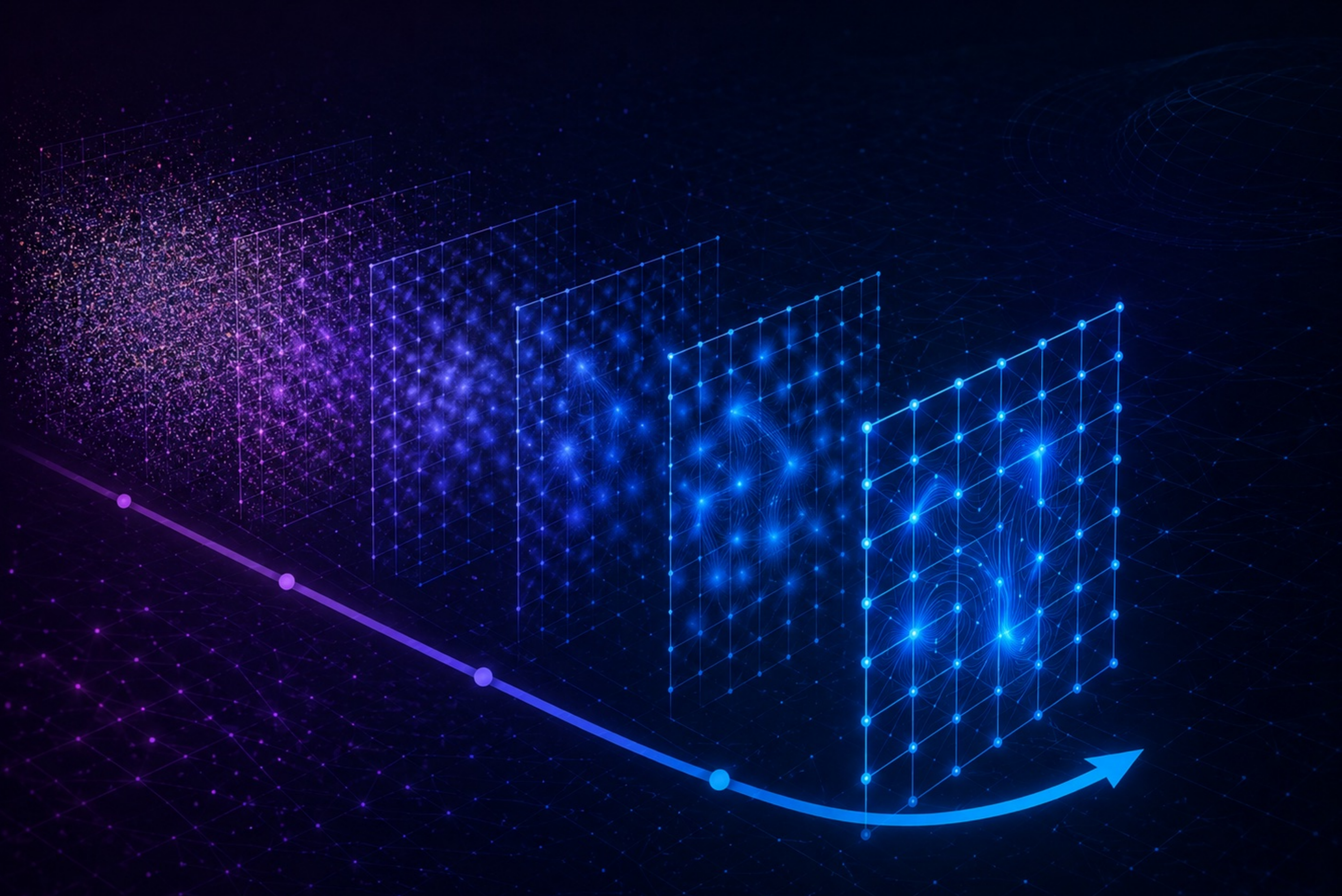
Summary

Take-home Messages

- **Diffusion Models**
  - Stochastic Quantization
  - Transport Flow
- **Simulate Lattice Fields**
  - Scalar and Gauge Fields
  - From 2D to 4D
- **Learn-To-Sample**
  - Learn From Existing Configs
  - Expandable, Exact, Efficient Sampling
- **Optimal Stochastic Quantization**







DM meets QFT, opportunities and challenges