Bridging ML and statistical field theory, ECT* Trento

Generative Modeling toward Critical Behaviors

Yang-yang Tan

IPI, The University of Tokyo

with **Gert Aarts**, **Diaa E. Habibi**, **Jinyang Li**, **Biagio Lucini** and **Lingxiao Wang**

September 8, 2026

DM-QFT

dm-qft.github.io/homepage

Outline

Critical phenomena and universality

Part I: critical statics from diffusion models

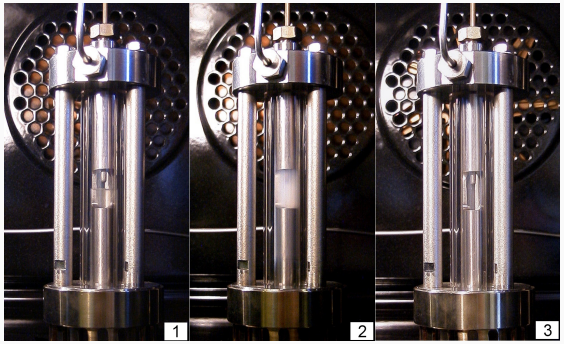
- Diffusion model on the lattice
- Critical scaling from the diffusion model
- Diagnosing the diffusion model
- Volume extrapolation

Part II: real-time critical dynamics

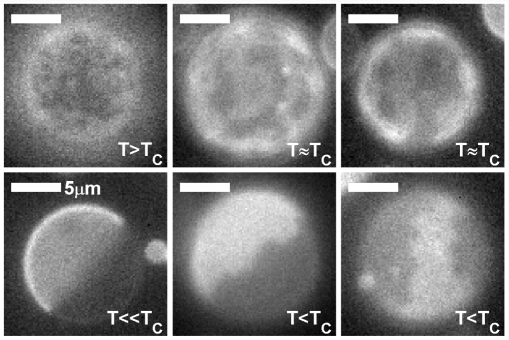
- Conserved dynamics on the lattice
- Dynamic critical scaling
- Phase separation

Critical phenomena and universality

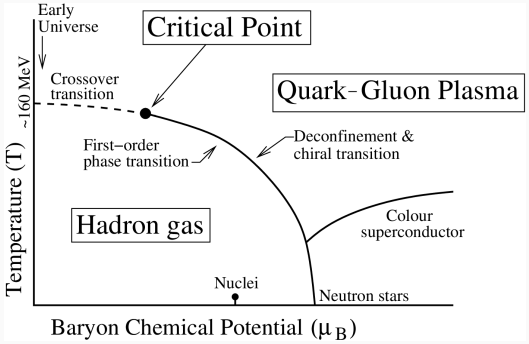
Critical phenomena across physics



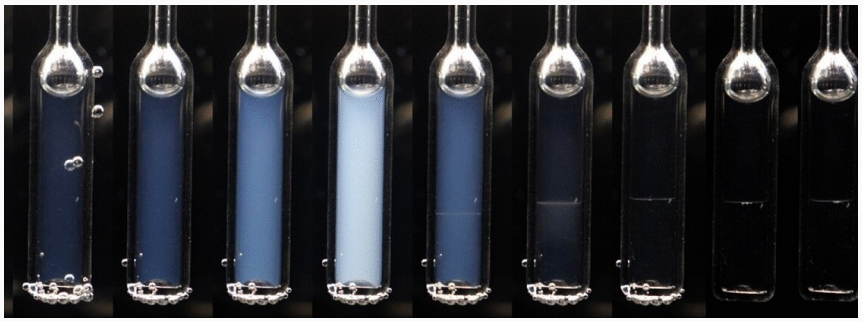
liquid–gas: ethane at T_c



cell membranes, $d = 2$



QCD: critical end point



binary mixture 3-picoline–D_2O at T_c

Universal features

no characteristic length, $\xi \to \infty$

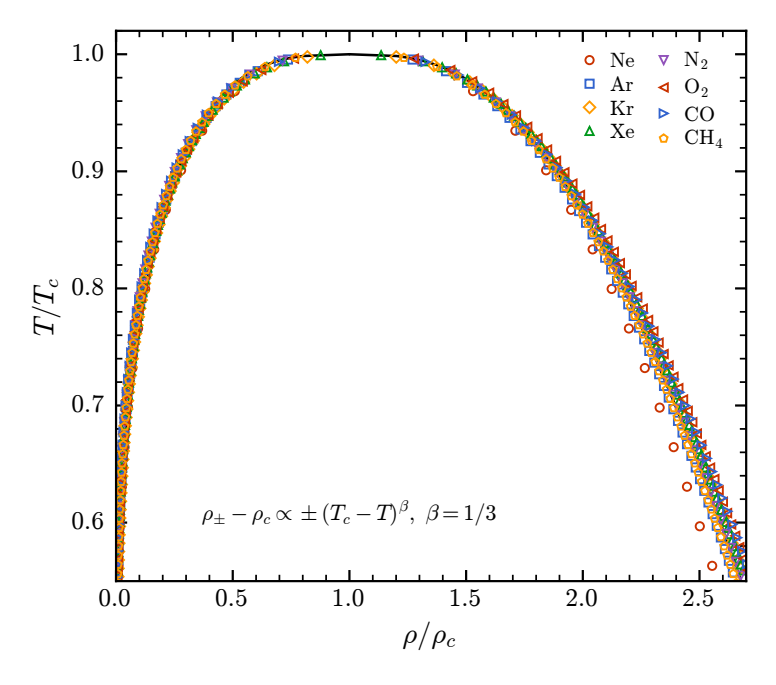
fluctuations on every scale

power-law singularities at T_c

$\chi \sim |T - T_c|^{-\gamma}$, $\xi \sim |T - T_c|^{-\nu}$

Ethane: Horstmann, Wikimedia Commons. **Membranes:** Gray, Karslake, Machta, Veatch, Biophys. J. 105, 2751 (2013). **Mixture:** Kalabiński, Drozd-Rzoska, Rzoska, EPJ E 45, 73 (2022). **QCD:** Bhalerao, doi:10.5170/CERN-2014-001.219; Stephanov, Rajagopal, Shuryak, PRL 81, 4816 (1998).

Universality



eight fluids, one coexistence curve

experimental data

	β	γ	ν
fluids	0.32–0.33	1.23–1.25	0.62–0.63
mixtures	0.32–0.33	1.23–1.25	0.62–0.64
magnets	0.32–0.33	1.24–1.25	0.62–0.64
3D Ising	0.3265	1.2372	0.6301

Only d and symmetry matter

fluids, mixtures, uniaxial magnets

and the QCD critical end point:

all 3D Ising

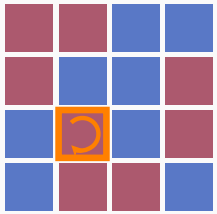
cell membranes: 2D Ising, same story

Lattice ϕ^4 is the minimal representative of the class

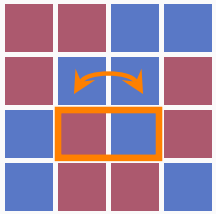
Coexistence curves: Guggenheim, J. Chem. Phys. 13, 253 (1945); data from the NIST Chemistry WebBook, line $\rho_\pm/\rho_c = 1 + \frac{3}{4}(1 - T/T_c) \pm \frac{7}{4}(1 - T/T_c)^{1/3}$. **Exponents:** Pelissetto, Vicari, Phys. Rept. 368, 549 (2002).

Dynamic universality

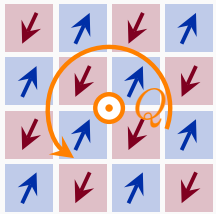
	conservation laws	examples	z
Model A	nonconserved ϕ	magnets, order–disorder alloys	$\simeq 2.0$
Model B	conserved ϕ	alloys, diffusive mixtures	$4-\eta$
Model G	nonconserved ϕ + conserved charges	antiferromagnets, chiral transition	$d/2$
Model H	conserved ϕ + conserved momentum	fluids, binary liquids, QCD end point	$\simeq 3.0$



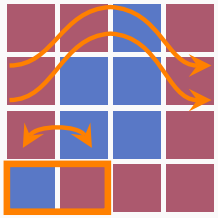
A: flip a site



B: swap a pair



G: precession with conserved Q



H: diffusion + advection

Same statics, different dynamics

the equilibrium weight e^{-S} is identical

relaxation time $\tau \sim \xi^z$, $\omega(k) \sim k^z$

conservation laws and mode couplings determine z

this talk: Model B

conserved ϕ, $z = 4 - \eta$

Classification: Hohenberg, Halperin, Rev. Mod. Phys. 49, 435 (1977). **Model B exponent:** Halperin, Hohenberg, Ma, PRB 10, 139 (1974). **Model G and chiral QCD:** Rajagopal, Wilczek, Nucl. Phys. B 399, 395 (1993). **Model H and QCD:** Son, Stephanov, PRD 70, 056001 (2004).

Part I: critical statics from diffusion models

Generative modeling for lattice field theory

Lattice ϕ^4 $S[\phi] = \sum_x \left[-2\kappa \sum_\mu \phi_x \phi_{x+\hat{\mu}} + \phi_x^2 + \lambda(\phi_x^2 - 1)^2 \right]$, Ising class

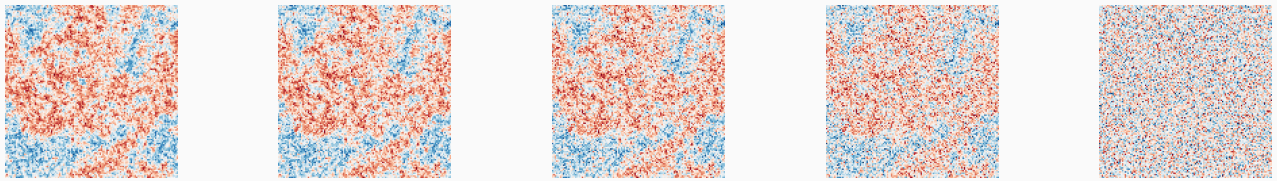
Flows: Albergo, Kanwar, Shanahan, PRD 100, 034515 (2019). **Physics-informed kernels:** Ihssen, Kapust, Pawlowski, arXiv:2510.26678. **Diffusion models on the lattice:** Wang, Aarts, Zhou, JHEP 05, 060 (2024); Aarts et al., MLST 6, 025004 (2025); Zhu et al., JHEP 03, 111 (2026). **Critical slowing down in DMs:** Del Bono et al., arXiv:2605.12597.

Generative modeling for lattice field theory

Lattice ϕ^4 $S[\phi] = \sum_x \left[-2\kappa \sum_\mu \phi_x \phi_{x+\hat{\mu}} + \phi_x^2 + \lambda(\phi_x^2 - 1)^2 \right]$, Ising class

Learn the distribution from data

forward: noise erases the configuration, $\phi \to$ Gaussian



reverse: the learned drift $\nabla \log p_t$ turns Gaussian noise into a sample of e^{-S}

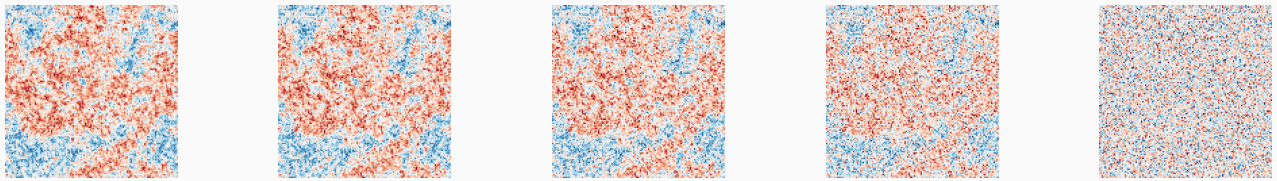
Flows: Albergo, Kanwar, Shanahan, PRD 100, 034515 (2019). **Physics-informed kernels:** Ihssen, Kapust, Pawlowski, arXiv:2510.26678. **Diffusion models on the lattice:** Wang, Aarts, Zhou, JHEP 05, 060 (2024); Aarts et al., MLST 6, 025004 (2025); Zhu et al., JHEP 03, 111 (2026). **Critical slowing down in DMs:** Del Bono et al., arXiv:2605.12597.

Generative modeling for lattice field theory

Lattice ϕ^4 $S[\phi] = \sum_x \left[-2\kappa \sum_\mu \phi_x \phi_{x+\hat{\mu}} + \phi_x^2 + \lambda(\phi_x^2 - 1)^2 \right]$, Ising class

Learn the distribution from data

forward: noise erases the configuration, $\phi \to$ Gaussian



reverse: the learned drift $\nabla \log p_t$ turns Gaussian noise into a sample of e^{-S}

This talk statics: sampling the transition, and volumes never seen
dynamics: Model B from a learned transition kernel

Also here RG: Anna's, Elia's talks self learning: Akio's talk
flows: Nils's, Lena's, Simran's talks annealed DM: Gurtej's talk
PIKs: Jan's, Friederike's, Renzo's talks

Flows: Albergo, Kanwar, Shanahan, PRD 100, 034515 (2019). **Physics-informed kernels:** Ihssen, Kapust, Pawlowski, arXiv:2510.26678. **Diffusion models on the lattice:** Wang, Aarts, Zhou, JHEP 05, 060 (2024); Aarts et al., MLST 6, 025004 (2025); Zhu et al., JHEP 03, 111 (2026). **Critical slowing down in DMs:** Del Bono et al., arXiv:2605.12597.

Score-based diffusion is learned stochastic quantization

Forward VE process $\Sigma(1) \gtrsim$ data spread, so $p_{t=1}$ is a broad Gaussian

$$\mathrm{d}\phi = \sigma^t \, \mathrm{d}w, \qquad \phi_t = \phi_0 + \Sigma(t)\, \eta, \qquad \Sigma^2(t) = \frac{\sigma^{2t} - 1}{2 \ln \sigma}$$

Score matching minimizer $s_\theta = \nabla_\phi \log p_t$ RGDM: Yuto's talk

$$\mathcal{L}(\theta) = \Big\langle \big\| \Sigma(t)\, s_\theta(\phi_t,\, t) + \eta \big\|^2 \Big\rangle_{t, \phi_0, \eta}$$

Score-based diffusion is learned stochastic quantization

Forward VE process $\Sigma(1) \gtrsim$ data spread, so $p_{t=1}$ is a broad Gaussian

$$\mathrm{d}\phi = \sigma^t \, \mathrm{d}w, \qquad \phi_t = \phi_0 + \Sigma(t)\, \eta, \qquad \Sigma^2(t) = \frac{\sigma^{2t} - 1}{2 \ln \sigma}$$

Score matching minimizer $s_\theta = \nabla_\phi \log p_t$ RGDM: Yuto's talk

$$\mathcal{L}(\theta) = \left\langle \left\| \Sigma(t)\, s_\theta(\phi_t,\, t) + \eta \right\|^2 \right\rangle_{t, \phi_0, \eta}$$

Reverse SDE and ODE same marginals p_t; Euler–Maruyama sampler

$$\mathrm{d}\phi = -g^2(t)\, s_\theta(\phi, t)\, \mathrm{d}t + g(t)\, \mathrm{d}\bar{w}, \qquad \frac{\mathrm{d}\phi}{\mathrm{d}t} = -\tfrac{1}{2} g^2(t)\, s_\theta(\phi, t)$$

Score-based diffusion is learned stochastic quantization

Forward VE process $\Sigma(1) \gtrsim$ data spread, so $p_{t=1}$ is a broad Gaussian

$$\mathrm{d}\phi = \sigma^t \, \mathrm{d}w, \qquad \phi_t = \phi_0 + \Sigma(t)\, \eta, \qquad \Sigma^2(t) = \frac{\sigma^{2t} - 1}{2 \ln \sigma}$$

Score matching minimizer $s_\theta = \nabla_\phi \log p_t$ RGDM: Yuto's talk

$$\mathcal{L}(\theta) = \left\langle \left\| \Sigma(t)\, s_\theta(\phi_t,\, t) + \eta \right\|^2 \right\rangle_{t, \phi_0, \eta}$$

Reverse SDE and ODE same marginals p_t; Euler–Maruyama sampler

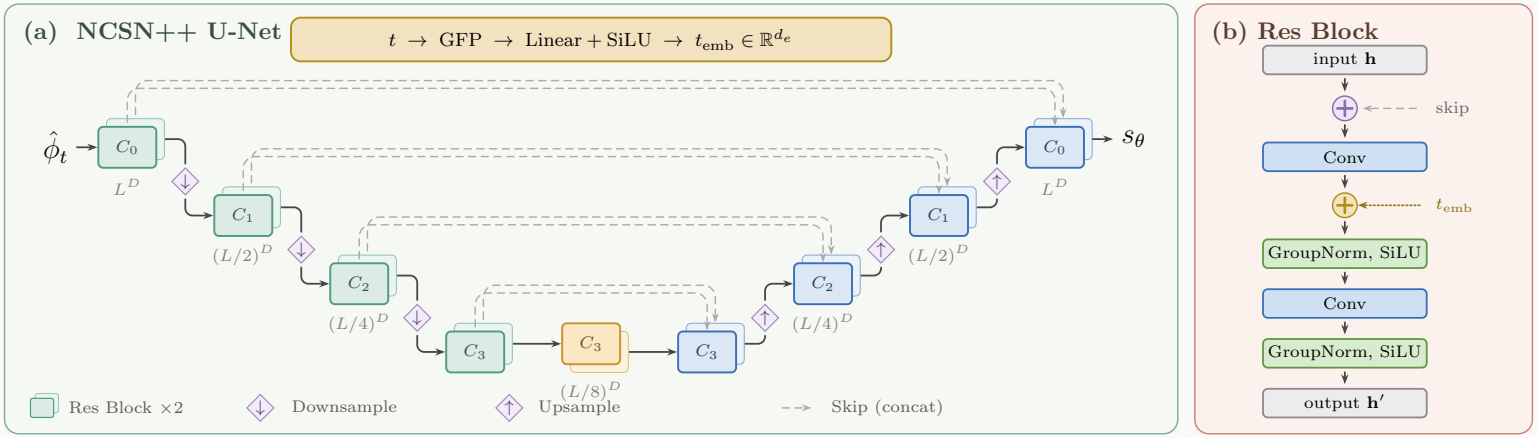
$$\mathrm{d}\phi = -g^2(t)\, s_\theta(\phi, t)\, \mathrm{d}t + g(t)\, \mathrm{d}\bar{w}, \qquad \frac{\mathrm{d}\phi}{\mathrm{d}t} = -\tfrac{1}{2} g^2(t)\, s_\theta(\phi, t)$$

Small-time limit the drift of stochastic quantization: is it reproduced?

$$s_\theta(\phi, t \to 0) \;\to\; \nabla_\phi \log p_{\mathrm{data}} = -\nabla_\phi S[\phi]$$

Score SDE: Song et al., ICLR (2021). **DSM:** Vincent, Neural Comput. 23, 1661 (2011).
Stochastic quantization: Parisi, Wu, Sci. Sin. 24, 483 (1981). **DM as stochastic quantization:** Wang, Aarts, Zhou, JHEP 05, 060 (2024); Fukushima, Kamata, Hirono, JPSJ 94, 031010 (2025).

Score network and reference configurations



Physics built in

periodic lattice: circular padding

translation symmetry: convolutions

nothing else, not even $\phi \to -\phi$

gauge equivariance: Thomas's talk

Architecture

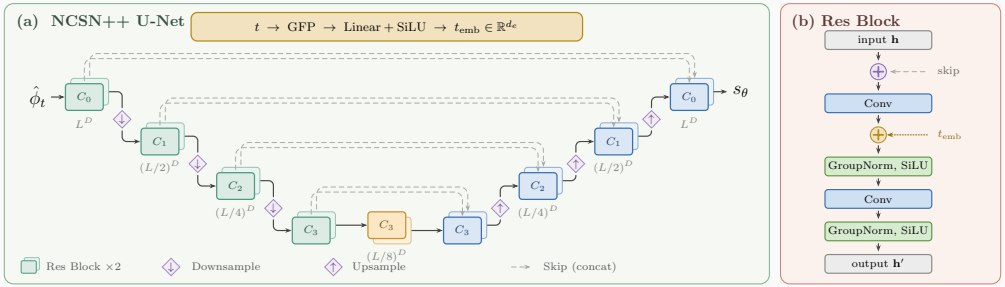
U-Net, $L \to L/8$: reaches the infrared

skip connections keep the ultraviolet

no dense layer: any L, same weights

NCSN++: Song et al., ICLR (2021). The noise level t enters every Res Block through an embedding, panel (b). $\kappa_c = 0.27088$ (2D, $L = 128$), 0.19225 (3D, $L = 64$) from Binder crossings; 3D: Hasenbusch, J. Phys. A 32, 4851 (1999).

Score network and reference configurations



Physics built in

periodic lattice: circular padding

translation symmetry: convolutions

nothing else, not even $\phi \to -\phi$

gauge equivariance: Thomas's talk

Physics learned

the action $S[\phi]$, never given

the $\phi \to -\phi$ symmetry

Jinyang's talk

critical correlations, which phase

Architecture

U-Net, $L \to L/8$: reaches the infrared

skip connections keep the ultraviolet

no dense layer: any L, same weights

Reference configurations

HMC with Wolff reflections

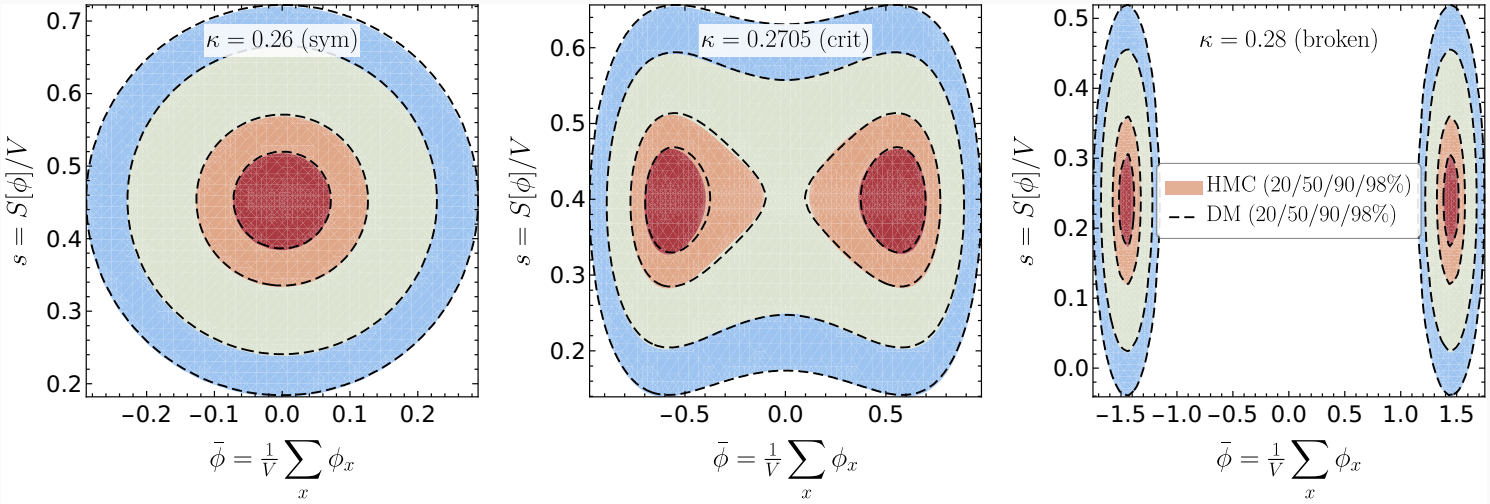
κ below, at and above κ_c

10240 / 2560 configs, all it ever sees

data-free: Biagio's, Shinichi's talks

NCSN++: Song et al., ICLR (2021). The noise level t enters every Res Block through an embedding, panel (b). $\kappa_c = 0.27088$ (2D, $L = 128$), 0.19225 (3D, $L = 64$) from Binder crossings; 3D: Hasenbusch, J. Phys. A 32, 4851 (1999).

Magnetization and action density



2D, $L = 128$: joint distribution $P(\bar{\phi}, s)$ of mean field and action density

Symmetric	**Critical**	**Broken**
a single peak at $\bar{\phi} = 0$	two finite-volume peaks	peaks at the right $\pm v$
contours coincide	shape and χ agree	too wide: χ 30% too large

all three phases right, the sole defect: the ordered peak is too wide

Cumulants in diffusion models: Aarts, Habibi, Wang, Zhou, MLST 6, 025004 (2025).
$\chi = V(\langle\bar{\phi}^2\rangle - \langle|\bar{\phi}|\rangle^2)$; contours enclose $20/50/90/98\%$ of the probability mass.

The propagator converges with training

Resolve the scales

$G(k) = \langle |\tilde{\phi}(k)|^2 \rangle$

a separate number for each k,

the loss averages them

2D, critical

short distances first

the infrared last

3D, critical

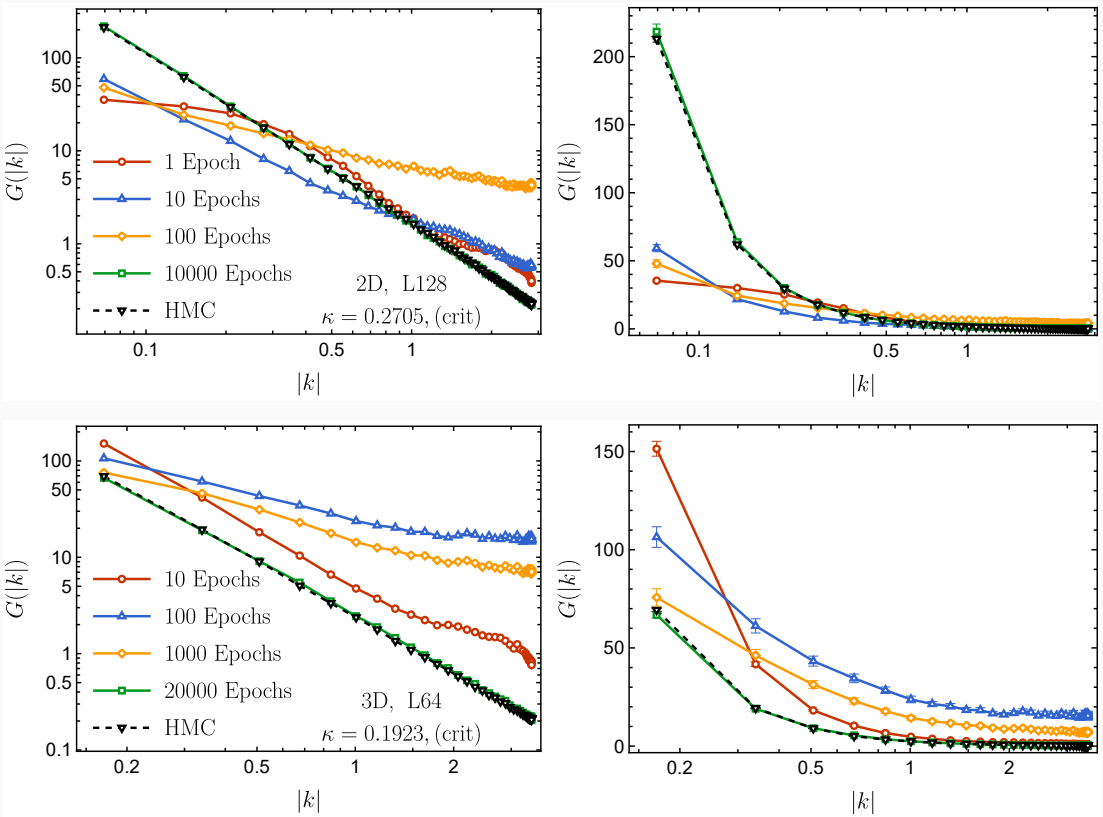
early overshoot at all k

infrared settles first

After training

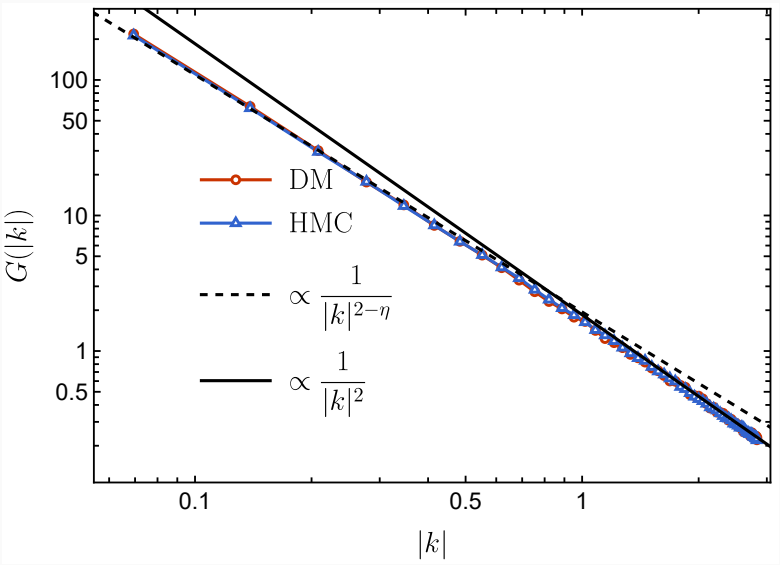
all six couplings

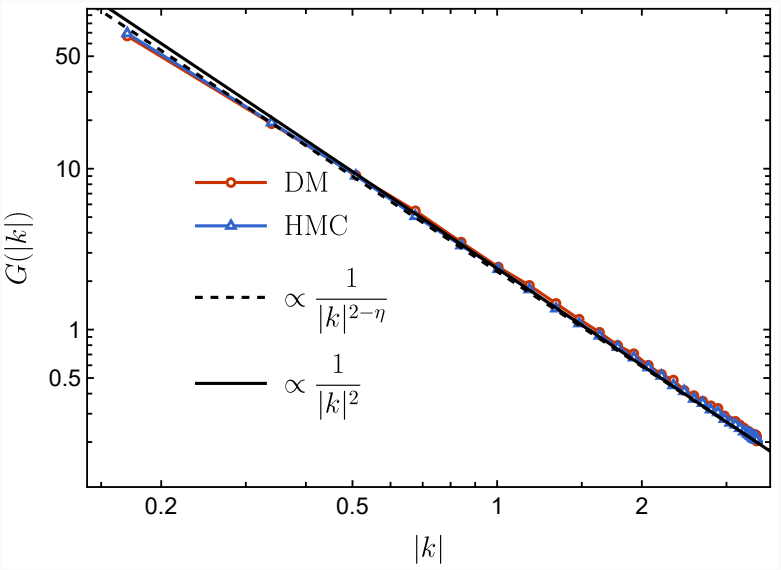
match HMC at every k



Left panels log–log, right panels log–linear, zero mode excluded. Each curve is a checkpoint of the same training run, sampled with the reverse SDE.

Anomalous dimension at the critical point





2D, $L = 128$, $\kappa = 0.2705$, $\eta = 1/4$

3D, $L = 64$, $\kappa = 0.1923$, $\eta \simeq 0.036$

Critical scaling

$G(k) \propto k^{-(2-\eta)}$

$\eta = 0$ for a free field

2D Ising

$\eta = 1/4$ resolved

off the free $1/k^2$

3D Ising

$\eta \simeq 0.036$, lines close

DM and HMC coincide

the interacting fixed point leaves its fingerprint in η, and the samples carry it

Exponents: Pelissetto, Vicari, Phys. Rept. 368, 549 (2002); Hasenbusch, J. Phys. A 32, 4851 (1999). Converged models, trained and sampled at the same L.

Diagnostic 1: a Metropolis test of the learned drift

A step with the learned drift

$$\psi = \phi + h\, s_\theta(\phi, t \to 0) + \sqrt{2h}\,\eta$$

$$\alpha = \frac{p(\psi)\; q(\phi \,|\, \psi)}{p(\phi)\; q(\psi \,|\, \phi)}, \qquad p \propto e^{-S}$$

Acceptance as a local test

e^{-S} judges the proposal

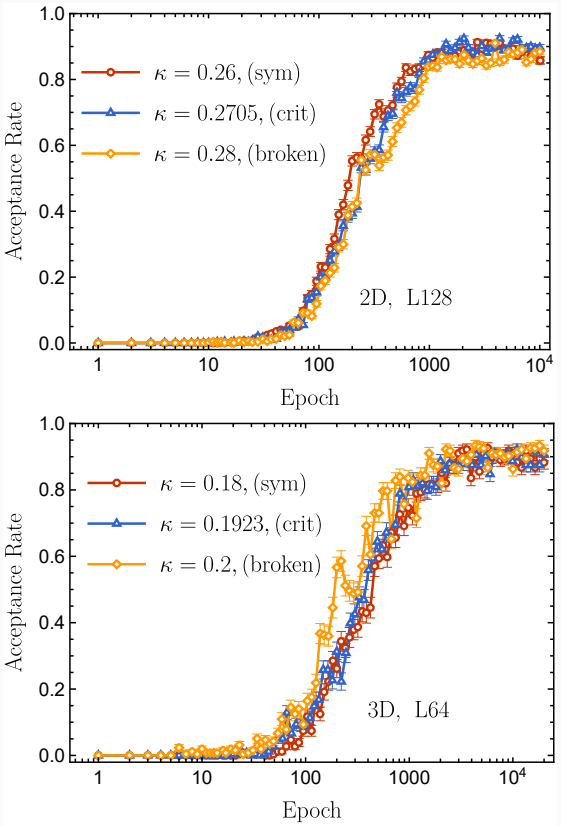
near 1: the local drift is right

no reverse trajectory involved

Result

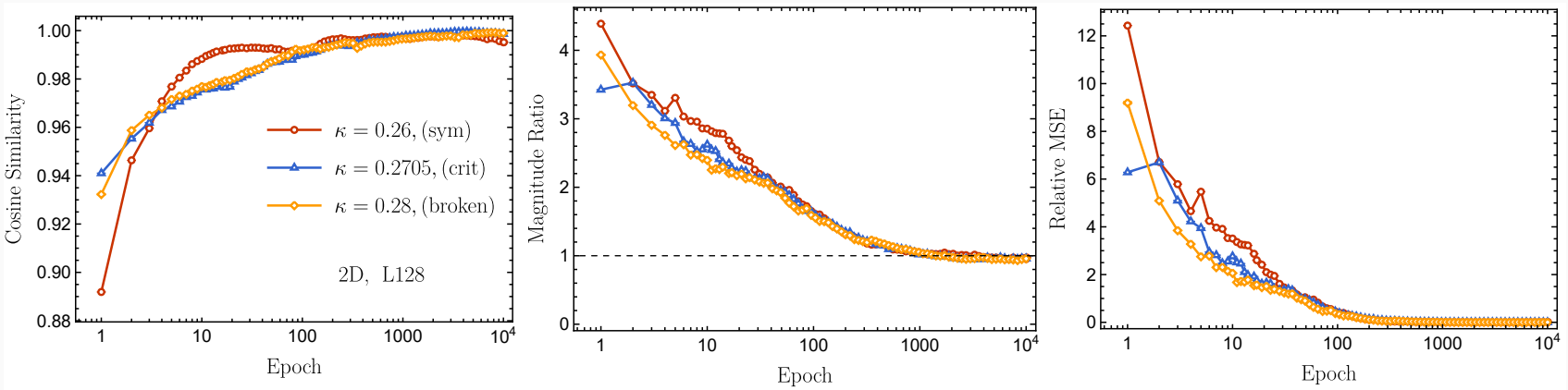
above 65% at every $L \geq 8$, every κ

0.85–0.91 on the largest lattices



MALA scaling: Roberts, Rosenthal, J. R. Statist. Soc. B 60, 255 (1998). **Metropolis-adjusted diffusion:** Lam et al., arXiv:2605.09654; Zhu et al., JHEP 03, 111 (2026). Step size $h \propto L^{-d}$, chosen where early and late checkpoints differ most.

Diagnostic 2: is the learned drift the force $-\nabla S$?



$s_\theta(\phi, t \to 0)$ against $-\nabla S[\phi]$ on HMC configurations, S never used in training

Direction

$$\frac{\langle s_\theta, -\nabla S \rangle}{|s_\theta|\,|\nabla S|}$$

cosine $0.9 \to 0.998$

within $\sim 10^2$ epochs

same in every phase

Magnitude

$$\frac{|s_\theta|}{|\nabla S|}$$

3–4× too strong

$\to 1$ by 5×10^3 epochs

$-\nabla S$ reconstructed

from samples alone

Overall error

$$\frac{|s_\theta + \nabla S|^2}{|\nabla S|^2}$$

falls $10 \to 10^{-2}$

zero only if direction and magnitude agree

2D, $L = 128$; the same trends in 3D. The test isolates the network as a function: errors that build up along the reverse trajectory are not seen here.

Diagnostic 3: what is a DM sample worth?

No autocorrelation to measure

DM samples are independent, $\tau_{\text{int}} = \frac{1}{2}$ even when biased

Error-matched sample size

$$N_{\text{eff},O} = \frac{\sigma^2_{p,O}}{\delta^2_O + \text{Var}(\bar{O}_{\text{DM}})}$$

exact samples with the same total error, bias δ_O included

Bias from the HMC reference

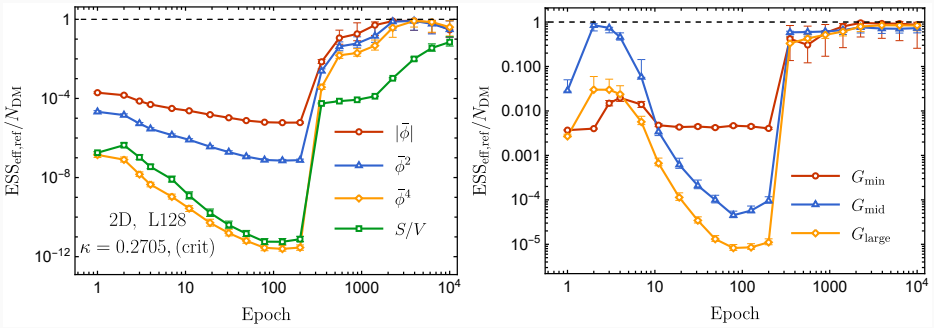
δ_O: difference of the means

an N_{eff} for each observable

Another probe

a configurational thermometer:

Diaa's talk



2D, $L = 128$, critical: $N_{\text{eff}}/N_{\text{DM}}$ along training

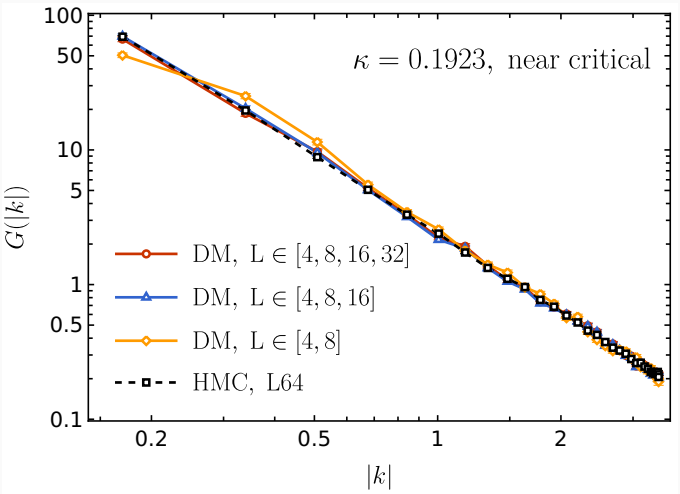
Early: bias dominated, ratios $\ll 1$

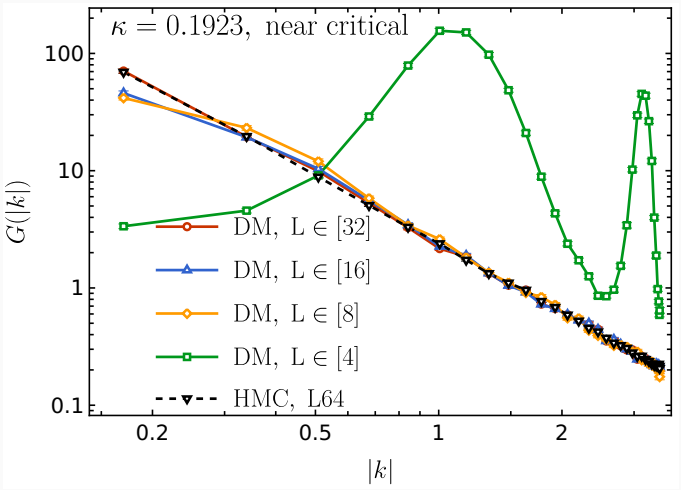
Late: ratios close to 1 for scalars and for the propagator at all scales

3D critical: all but S/V reach 1, the action bias remains

Propagator bands: modes averaged over the infrared, intermediate and ultraviolet ranges. Errors from bootstrap; the HMC reference variance includes its autocorrelation.

Train on small lattices, sample a lattice never seen





multi-L training $\rightarrow L = 64$

3D, $\kappa = 0.1923$: $L_{\text{train}} \leq 32 \implies L = 64$

$\{4, 8, 16, 32\}$ tracks HMC down to $k_{\min}$

$\{4, 8\}$ misses the infrared rise

single L: 32 closest, 4 fails

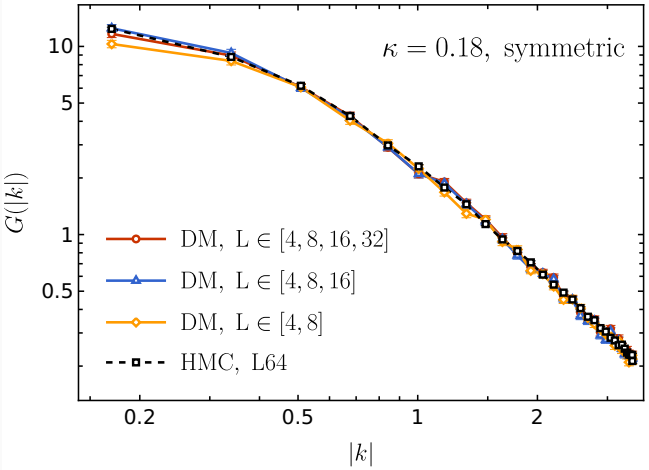
several volumes teach the infrared

single-L baselines $\rightarrow L = 64$

training L	acceptance
$\{4, 8\}$	$0.578(8)$
$\{4, 8, 16\}$	$0.830(6)$
$\{4, 8, 16, 32\}$	$0.877(6)$
$\{64\}$, in distribution	$0.911(9)$

Multi-lattice generative models: Máté, Fleuret, MLST 5, 045053 (2024); Bauer, Kapust, Pawlowski, Temmen, SciPost Phys. 19, 077 (2025); Singha et al., ICLR (2025); Zhu et al., JHEP 03, 111 (2026). **Cost of large-volume training:** Abbott et al., EPJA 59, 257 (2023). No $L = 64$ information enters training or sampling.

Unseen lattice: symmetric and broken phases

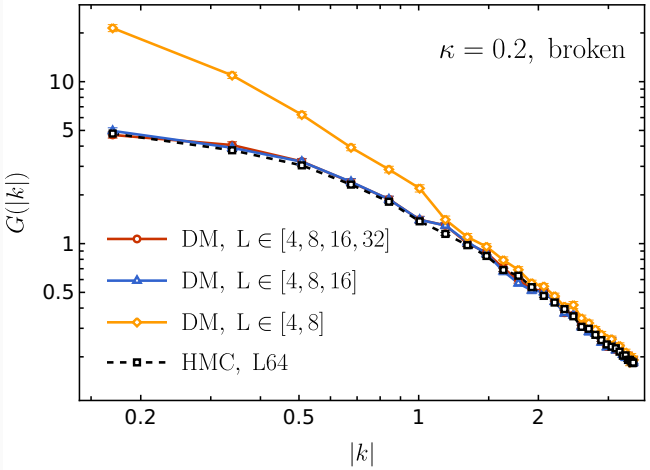


Symmetric phase, $\kappa = 0.18$

ξ is a few spacings, so every training lattice sees all of it

$\{4, 8\}$ undershoots $k_{\min}$ by 15%

$\{4, 8, 16\}$ and up: on top of HMC



Broken phase, $\kappa = 0.2$

condensate, not ξ, sets the infrared

$L = 4, 8$ still flip sign: no ordered lattice in the set, $\{4, 8\}$ invents $4\times$ too much infrared power

$\{4, 8, 16\}$ and up: on top of HMC

Rule of thumb largest training lattice within $2\times$ of the target

$k_{\min} = 2\pi/64$ is below every training momentum, $k = 0$ is not shown

3D, $\lambda = 0.9$, $L = 64$, same training sets as on the previous slide.

Part II: real-time critical dynamics

From statics to dynamics

Statics

Boltzmann weight $e^{-S[\phi]}$

however it was reached

no time

Dynamics

equation of motion $\partial_t \phi = \mathcal{F}[\phi, \xi]$

stochastic trajectories $\phi(t)$

$\mathcal{F}$ does not depend on t

From statics to dynamics

Statics

Boltzmann weight $e^{-S[\phi]}$

however it was reached

no time

Dynamics

equation of motion $\partial_t \phi = \mathcal{F}[\phi, \xi]$

stochastic trajectories $\phi(t)$

$\mathcal{F}$ does not depend on t

Transition kernel $p_{\Delta t}(\phi' \,|\, \phi)$, the probability to go from ϕ to ϕ' in Δt

$$\phi(t) \xrightarrow{p_{\Delta t}} \phi(t+\Delta t) \xrightarrow{p_{\Delta t}} \phi(t+2\Delta t) \xrightarrow{p_{\Delta t}} \cdots$$

From statics to dynamics

Statics

Boltzmann weight $e^{-S[\phi]}$

however it was reached

no time

Dynamics

equation of motion $\partial_t \phi = \mathcal{F}[\phi, \xi]$

stochastic trajectories $\phi(t)$

$\mathcal{F}$ does not depend on t

Transition kernel $p_{\Delta t}(\phi' \,|\, \phi)$, the probability to go from ϕ to ϕ' in Δt

$$\phi(t) \xrightarrow{p_{\Delta t}} \phi(t+\Delta t) \xrightarrow{p_{\Delta t}} \phi(t+2\Delta t) \xrightarrow{p_{\Delta t}} \cdots$$

Learned transition kernel

pairs $(\phi_t, \phi_{t+\Delta t})$ suffice, from a simulation or an experiment

each kernel step spans $\Delta t \gg \delta t$, the step of a direct integrator

Chapman–Kolmogorov: $p_{2\Delta t}(\phi'|\phi) = \int \mathrm{d}\phi''\, p_{\Delta t}(\phi'|\phi'')\, p_{\Delta t}(\phi''|\phi)$. Colored noise: enlarge the state, $X = (\phi, \eta)$, and the joint kernel composes again.

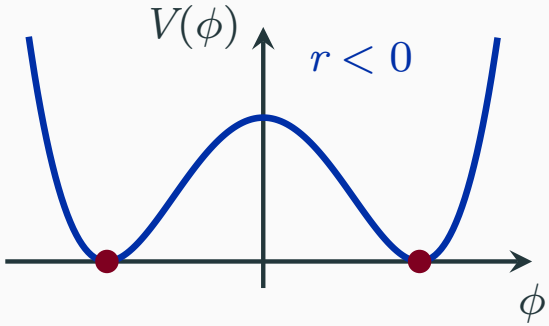
Model B: conserved dynamics

2D conserved ϕ^4 field

$$\partial_t \phi(\boldsymbol{x}, t) = \Gamma \nabla^2 \frac{\delta F[\phi]}{\delta \phi(\boldsymbol{x}, t)} + \zeta(\boldsymbol{x}, t),$$

$$\langle \zeta(\boldsymbol{x}, t) \zeta(\boldsymbol{x}', t') \rangle = -2\, T \Gamma \nabla^2 \delta^{(d)}(\boldsymbol{x} - \boldsymbol{x}') \delta(t - t'),$$

$$F[\phi] = \int d^d x \left[\tfrac{1}{2} (\nabla \phi)^2 + \tfrac{r}{2} \phi^2 + \tfrac{u}{4} \phi^4 \right].$$



Model B: Hohenberg, Halperin, Rev. Mod. Phys. 49, 435 (1977). Reference trajectories from a direct stochastic integrator (SOSRA) of these equations, on the lattice action of Part I. **Kernel:** $-\Gamma\nabla^2$ in front of $\delta F/\delta\phi$ leaves the equilibrium $e^{-F/T}$ unchanged and sets the dynamics.

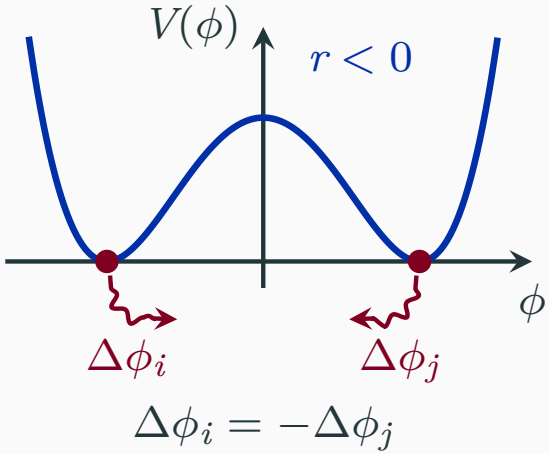
Model B: conserved dynamics

2D conserved ϕ^4 field

$$\partial_t \phi(\boldsymbol{x}, t) = \Gamma \nabla^2 \frac{\delta F[\phi]}{\delta \phi(\boldsymbol{x}, t)} + \zeta(\boldsymbol{x}, t),$$

$$\langle \zeta(\boldsymbol{x}, t) \zeta(\boldsymbol{x}', t') \rangle = -2\, T \Gamma \nabla^2 \delta^{(d)}(\boldsymbol{x} - \boldsymbol{x}') \delta(t - t'),$$

$$F[\phi] = \int d^d x \left[\tfrac{1}{2} (\nabla \phi)^2 + \tfrac{r}{2} \phi^2 + \tfrac{u}{4} \phi^4 \right].$$



Model B: Hohenberg, Halperin, Rev. Mod. Phys. 49, 435 (1977). Reference trajectories from a direct stochastic integrator (SOSRA) of these equations, on the lattice action of Part I. **Kernel:** $-\Gamma\nabla^2$ in front of $\delta F/\delta\phi$ leaves the equilibrium $e^{-F/T}$ unchanged and sets the dynamics.

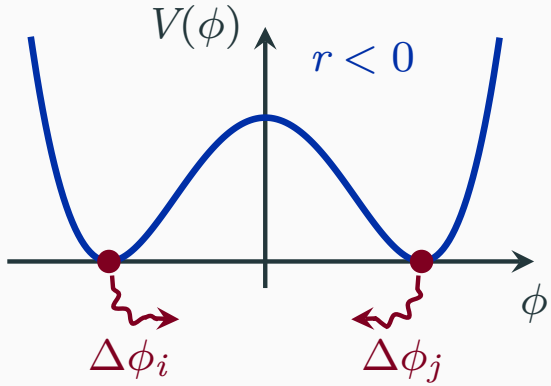
Model B: conserved dynamics

2D conserved ϕ^4 field

$$\partial_t \phi(\boldsymbol{x}, t) = \Gamma \nabla^2 \frac{\delta F[\phi]}{\delta \phi(\boldsymbol{x}, t)} + \zeta(\boldsymbol{x}, t),$$

$$\langle \zeta(\boldsymbol{x}, t) \zeta(\boldsymbol{x}', t') \rangle = -2\, T \Gamma \nabla^2 \delta^{(d)}(\boldsymbol{x} - \boldsymbol{x}') \delta(t - t'),$$

$$F[\phi] = \int d^d x \left[\tfrac{1}{2} (\nabla \phi)^2 + \tfrac{r}{2} \phi^2 + \tfrac{u}{4} \phi^4 \right].$$



$$\Delta\phi_i = -\Delta\phi_j$$

Conserved mean

$$M(t) \equiv V^{-1} \int d^d x\, \phi(\boldsymbol{x}, t)$$

$$M(t) = M(0)$$

Critical dynamics

$$\tau \propto \xi^z \propto k^{-z}$$

$$z = 4 - \eta$$

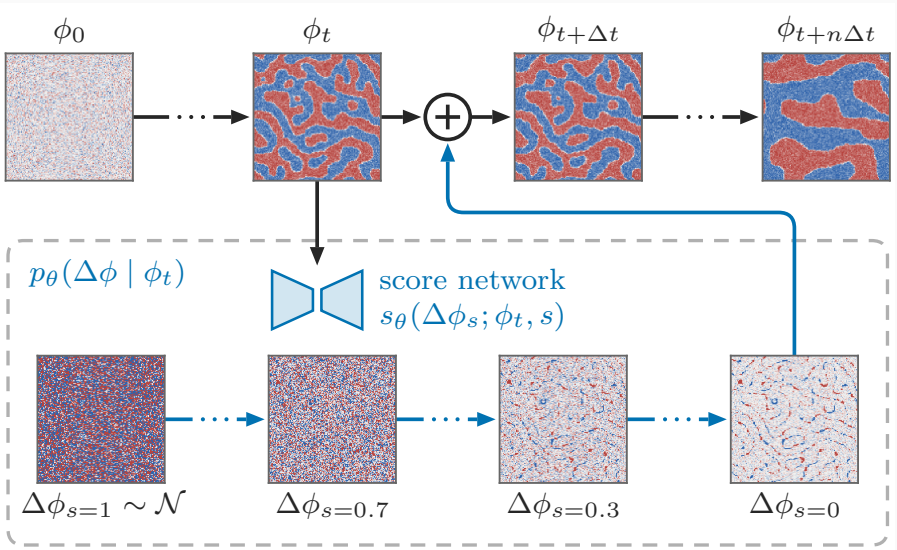
Kernel interval

$$\Delta t = 10$$

$$\gg \text{ a solver step}$$

Model B: Hohenberg, Halperin, Rev. Mod. Phys. 49, 435 (1977). Reference trajectories from a direct stochastic integrator (SOSRA) of these equations, on the lattice action of Part I. **Kernel:** $-\Gamma\nabla^2$ in front of $\delta F/\delta\phi$ leaves the equilibrium $e^{-F/T}$ unchanged and sets the dynamics.

Conditional diffusion as the transition kernel



$$\Delta\phi \sim p_\theta(\,\cdot \mid \phi_t)$$
$$\phi_{t+\Delta t} = \phi_t + \Delta\phi$$

Condition

the present field ϕ_t

Target

the increment $\Delta\phi$, diffused and denoised as in Part I

Physics learned

$p(\Delta\phi \,|\, \phi_t)$ from pairs alone

no drift, noise or free energy

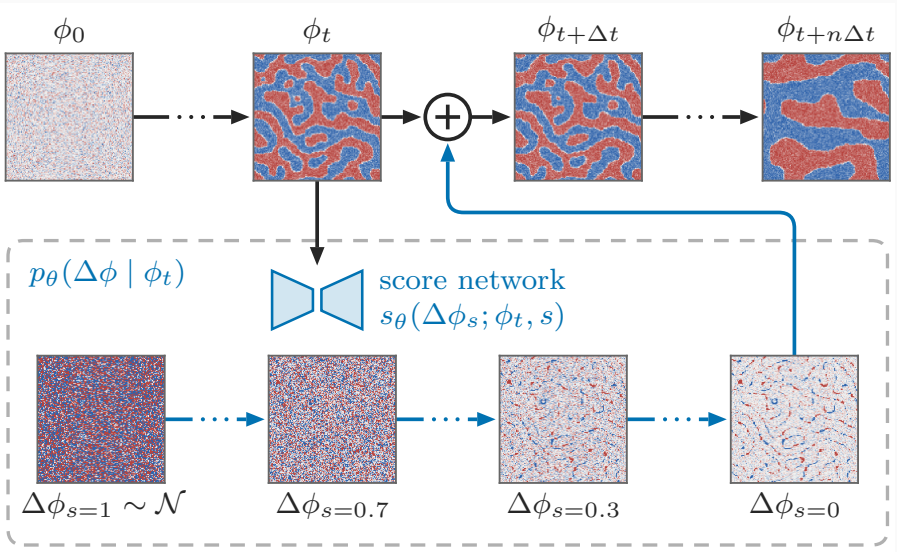
Physics built in

$\sum_x \Delta\phi_x = 0$ by projection

at every step, M conserved exactly

Fast ODE solver: Zheng et al., DPM-Solver-v3, NeurIPS (2023). Same U-Net as in Part I with ϕ_t as a second input channel; each kernel step costs a few dozen network calls.

Conditional diffusion as the transition kernel



$$\Delta\phi \sim p_\theta(\,\cdot \mid \phi_t)$$
$$\phi_{t+\Delta t} = \phi_t + \Delta\phi$$

Condition

the present field ϕ_t

Target

the increment $\Delta\phi$, diffused and denoised as in Part I

Physics learned

$p(\Delta\phi \,|\, \phi_t)$ from pairs alone

no drift, noise or free energy

Physics built in

$\sum_x \Delta\phi_x = 0$ by projection

at every step, M conserved exactly

Projection, three times $P\,v = v - \frac{1}{V}\sum_x v_x$, no retraining

$$\mathrm{d}\Delta\phi_s = -g^2(s)\, P\, s_\theta(\Delta\phi_s; \phi_t, s)\,\mathrm{d}s + g(s)\, P\,\mathrm{d}\bar{w}, \qquad \Delta\phi_{s=1} = P\,\Sigma(1)\,\eta$$

initial noise, score, noise increments: $\sum_x \Delta\phi_x = 0$ exactly at every step

Fast ODE solver: Zheng et al., DPM-Solver-v3, NeurIPS (2023). Same U-Net as in Part I with ϕ_t as a second input channel; each kernel step costs a few dozen network calls.

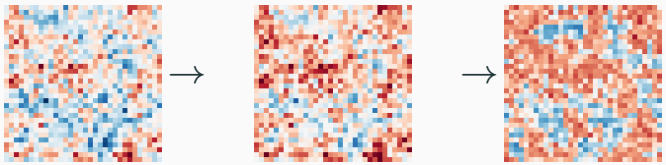
Training data for the kernel

Critical kernel $\kappa = \kappa_c,\ L = 32$

Gaussian field χ with propagator

$$\langle \chi_k \chi_{-k} \rangle = \frac{1}{Zk^2 + m^2}, \quad \phi_0 = M + \chi$$

Z sets the amplitude, m^2 the correlation length, M the mean, all random over a broad range



χ $\quad$ $\phi_0 = M + \chi$ $\quad$ $\phi_{\Delta t}$

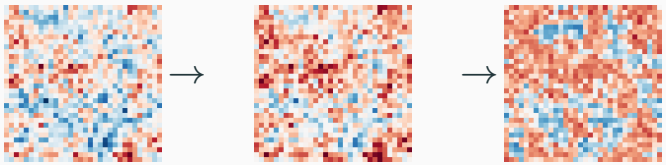
Training data for the kernel

Critical kernel $\kappa = \kappa_c,\ L = 32$

Gaussian field χ with propagator

$$\langle \chi_k \chi_{-k} \rangle = \frac{1}{Zk^2 + m^2}, \quad \phi_0 = M + \chi$$

Z sets the amplitude, m^2 the correlation length, M the mean, all random over a broad range



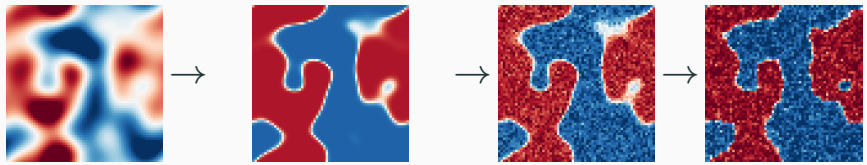
Coarsening kernel $\kappa = 0.85,\ L \le 64$

g Gaussian with $\langle g(\boldsymbol{x})\, g(0) \rangle = e^{-\boldsymbol{x}^2/2\ell^2}$

$$\phi_0 = A \tanh\left(\tfrac{\ell}{w}\, g\right) + \eta\, \xi$$

domains $\pm A$ of size ℓ, walls of width w, white noise ξ of level η

$\ell,\ w,\ A,\ \eta$ random over a broad range



Target Model B (SOSRA) run from ϕ_0 over $\Delta t = 10$, $\Delta\phi = \phi_{\Delta t} - \phi_0$

The point the kernel learns the rule of motion, not trajectories

every test is an extrapolation: other ensembles, longer times, larger lattices

Z log-uniform in $[0.05, 10]$, $m^2 \in [0, 0.8]$, $M \in [-0.9, 0.9]$; ℓ log-uniform in $[2, L/3]$, $w \in [0.7, 3]$, $A \in [0.8, 1.8]$, $\eta \in [0.05, 0.35]$, $1/8$ pure noise; 51200 pairs per L.

Equilibrium and nonequilibrium evolution

Critical point

2D, $L = 32$

strongly coupled

Initial ensembles

HMC equilibrium

Gaussian quench

Composition

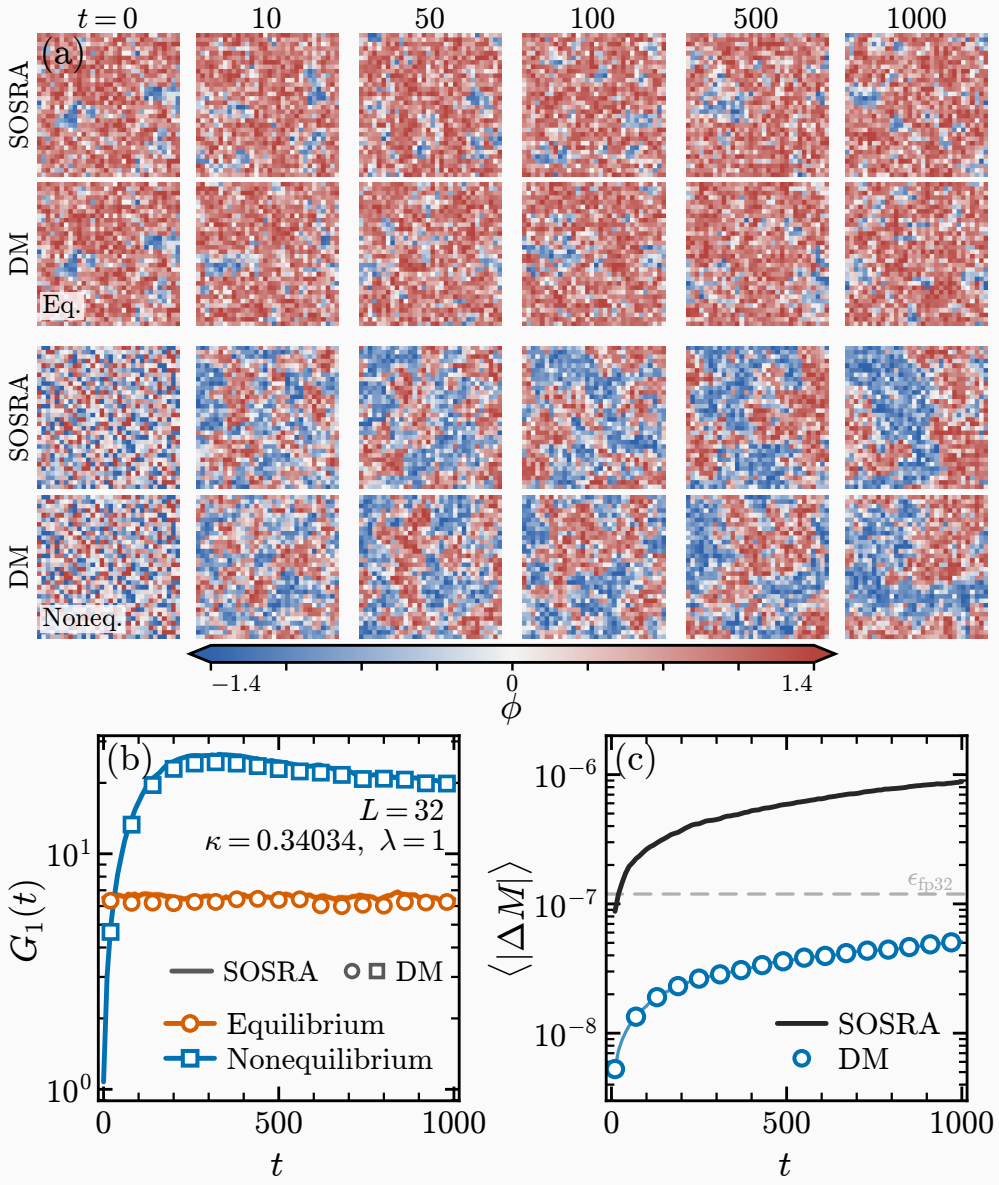
200 kernel steps

$t = 2000$

Against SOSRA

$G_1(t)$ agrees

$M(t) = M(0)$



Mode-resolved two-time correlations

Two-time correlation $\rho_n(\Delta t)$

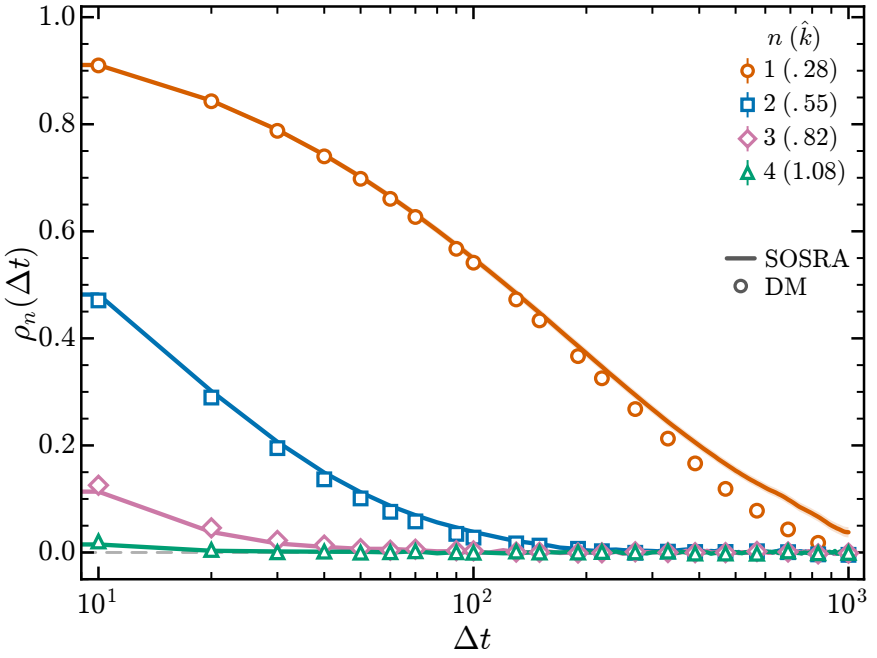
mode k_n at t against $t + \Delta t$,
averaged over t along the chain

Conserved dynamics is slow

$\tau_n \propto k_n^{-z}$ with $z \simeq 4$:
the softest mode outlives $n = 4$
by two orders of magnitude

Iterated kernel against SOSRA

lines: direct integration
markers: kernel, $\Delta t = 10$ per step
the same curve for every mode
then read off τ_n, test $\tau_n \propto k_n^{-z}$



2D, $L = 32$, critical, diagonal modes

$n = 1$–4

Time origins along the $t = 2000$ equilibrium chain of the previous slide. Errors: leave-one-trajectory-out jackknife.

The critical exponent emerges

Mode autocorrelation

$$\rho_n(\Delta t) = \frac{\langle \mathrm{Re}[\phi_n(t+\Delta t)\,\phi_n^*(t)]\rangle_t}{\sqrt{\langle G_n(t+\Delta t)\rangle_t \langle G_n(t)\rangle_t}}$$

$$\phi_n \equiv \phi(\boldsymbol{k}_n), \quad G_n = V^{-1}\langle |\phi_n|^2 \rangle$$

Dynamic scaling

$$\rho_n(\Delta t) = f(\Delta t/\tau_n), \qquad \tau_n \propto \widehat{k}_n^{-z}$$

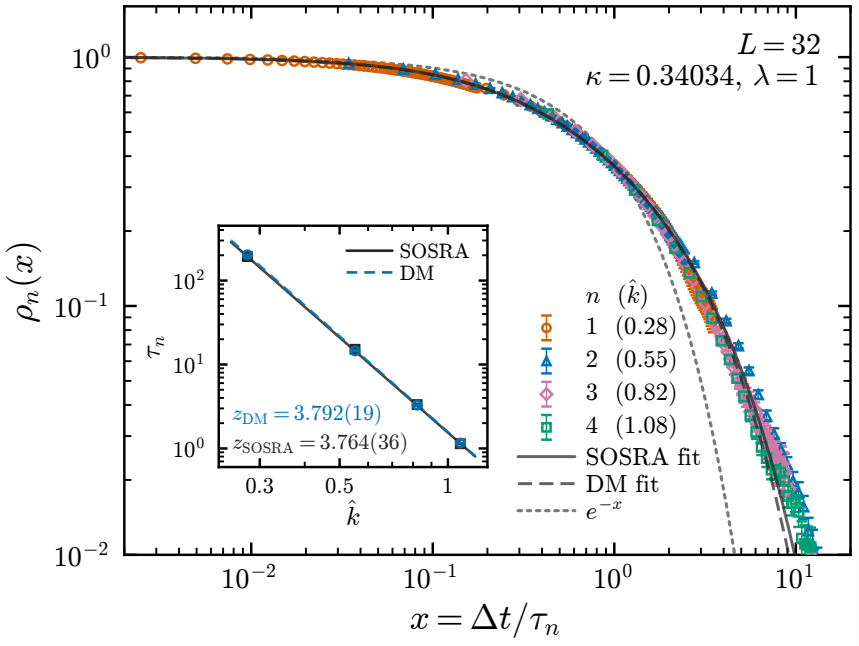
all modes on a common curve

Fits at $L = 32$

$z_{\mathrm{DM}} = \mathbf{3.792(19)}$

$z_{\mathrm{SOSRA}} = \mathbf{3.764(36)}$

exact: $z = 4 - \eta = 3.75$



$z = 4 - \eta = 15/4 = 3.75$. **Exact.** Halperin et al., PRB 10, 139 (1974).

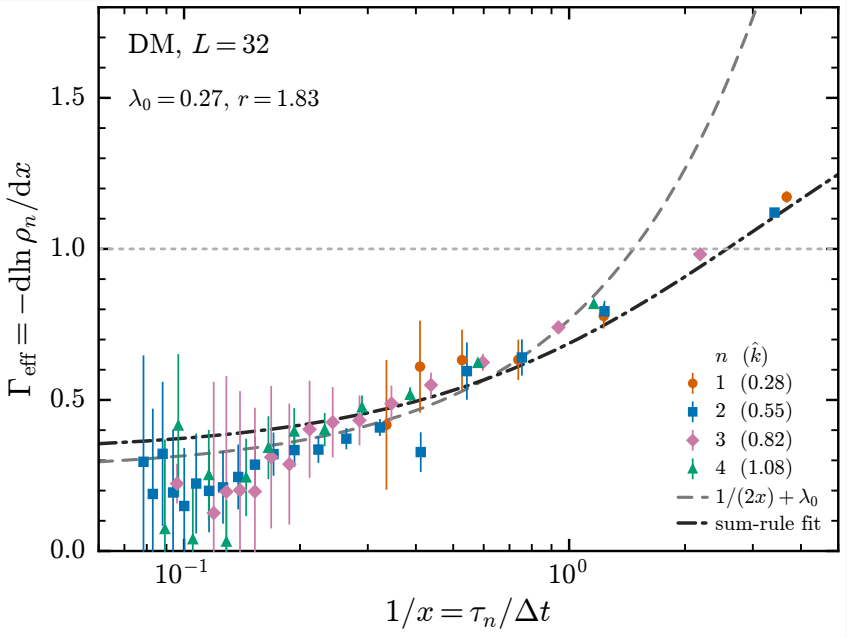
$z = 3.83(10),\ 3.716(17)$. **MC.** Schweitzer et al., NPB 984, 115944 (2022).

$z = 3.95(10)$. **Kawasaki MC.** Zheng, PLA 277, 257 (2000), erratum 282, 132 (2001).

$z = 3.79 \pm 0.05$. **MC RG.** Moseley et al., J. Stat. Phys. 57, 421 (1989).

$z = 3.80$. **MC RG.** Yalabik and Gunton, PRB 25, 534 (1982).

Beyond the exponent: the scaling function



decay rate of the learned kernel's ρ_n,

$$x = \Delta t/\tau_n$$

Effective decay rate

$\Gamma_{\rm eff}(x) = -\,d\ln\rho_n/dx$

a pure exponential: a flat line

Short times: set by the statics

fluctuation–dissipation relation

$\Gamma_0 = \widehat{k}_n^2/G_n, \quad r = \Gamma_0\tau_n \simeq 1.8$

Long times: a slower single rate

$\lambda_0 \simeq 0.27$, seven times below r

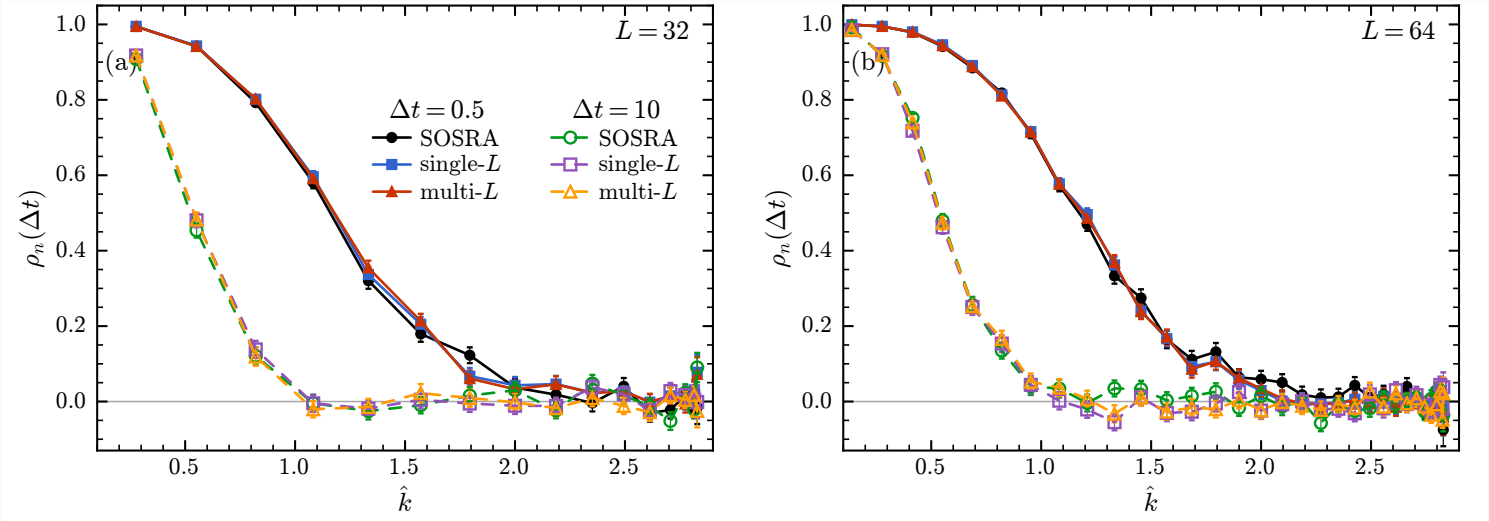
detailed balance: ρ_n is a positive sum of exponentials, the rate can only fall, from r to the gap λ_0

All four modes on a common curve

the whole function is reproduced, not only z. SOSRA has the tail too

Dashed: $\lambda_0 + 1/(2x)$. Dash-dotted: fit $f(x) = e^{-\lambda_0 x}(1+ax)^{-b}$ with the sum rule $\lambda_0 + ab = r$, $b \simeq \frac{1}{2}$. The direct SOSRA dynamics gives the same curve within errors.

Single-step test of the transition kernel



$\rho_n(\Delta t)$ after one kernel step from equilibrium fields, against $\widehat{k}$

The kernel alone

no error accumulates

all modes in a single plot

Two intervals

$\Delta t = 0.5$: fast modes

$\Delta t = 10$: slow modes

the interval is a choice

Two sizes

kernels from $L = 32$

and from $L = 8, 16, 32$

both right at $L = 64$,

a size neither has seen

Circles SOSRA, squares the single-size kernel, triangles the multi-size kernel. Filled markers $\Delta t = 0.5$, open markers $\Delta t = 10$.

Prediction on unseen $L = 64$

Size extrapolation

$$L_{\text{train}} = 8, 16, 32 \implies \boldsymbol{L = 64}$$

no $L = 64$ trajectory ever seen

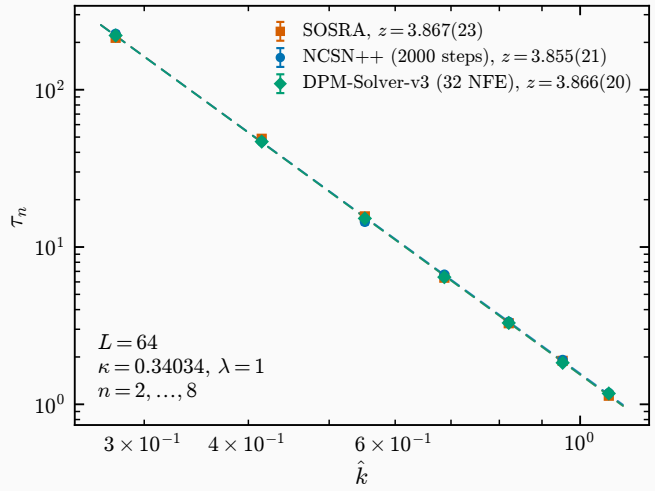
Fits at $L = 64$

$$\boldsymbol{z_{\text{DM}} = 3.855(21)}$$

$$\boldsymbol{z_{\text{Model B}} = 3.867(23)}$$

As fast as the best direct solver

512 trajectories, $t = 100$, GH200, matched accuracy



Method	Wall time (s) equil.	noneq.
DPM-Solver-v3 (This work)	3.24	3.26
SOSRA (This work)	1.30	3.28
SOSRA (StochasticDiffEq.jl)	12.36	30.67
ShARK (Diffrax)	5.56	17.76
EM (Diffrax)	19.45	38.75
EM (other GPU/CPU solvers)	> 100	> 200

Phase separation on unseen $L = 128$

Sign-field correlation

$$\sigma(\boldsymbol{x}, t) = \operatorname{sgn} \phi(\boldsymbol{x}, t)$$

$$C(r, t) = \langle \sigma(\boldsymbol{x}, t)\, \sigma(\boldsymbol{x} + \boldsymbol{r}, t) \rangle$$

$L(t)$: first zero of $C(r, t)$

Phase-ordering scaling

$$C(r, t) = g\big(r / L(t)\big)$$

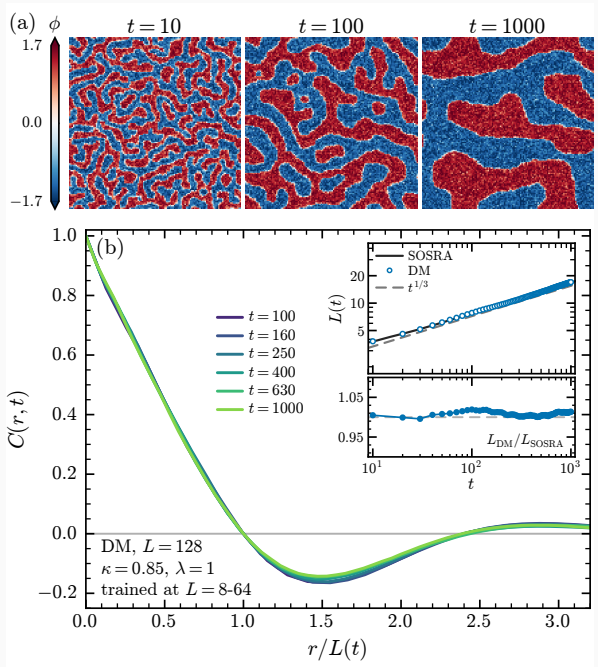
$$L(t) \propto t^{1/3}$$

Size extrapolation

$L_{\text{train}} \leq 64 \implies \boldsymbol{L = 128}$

$\boldsymbol{L(t) \propto t^{1/3}}$

$L_{\text{DM}} / L_{\text{SOSRA}} = 1$ within 2%



$t^{1/3}$. **LSW theory.** Lifshitz and Slyozov, JPCS 19, 35 (1961).

Asymptotic $t^{1/3}$, corrections. **Conserved Ising.** Huse, PRB 34, 7845 (1986).

$t^{1/3}$, scaling. **2D critical Model B.** Rogers et al., PRB 37, 9638 (1988).

$t^{1/3}$ scaling. **Review.** Bray, Adv. Phys. 43, 357 (1994).

$t^{1/3}$. **Polymer experiment.** Krausch et al., PRL 71, 3669 (1993).

Critical $\alpha = 1/3$. **2D CH.** König et al., PCCP 23, 24823 (2021).

$\alpha_L = 0.34, 0.37$ from perimeter. **Membrane.** Zhiliakov et al., BBA 1863, 183446 (2021).

$t^{1/3}$. **Passive membrane, CH.** Arnold et al., PRL 131, 128402 (2023).

$t^{1/3}$. **HP1 Model B.** Tortora et al., PNAS 120, e2211855120 (2023).

Long-time $t^{1/3}$. **Diffusive polymer Model B.** Garcia et al., JCP 159, 214904 (2023).

Summary

Learned scores $\longrightarrow$ critical statics and critical dynamics

Statics	Dynamics
$s_\theta \to -\nabla S$	$z = 3.792(19), \quad L(t) \propto t^{1/3}$
ϕ^4 in 2D and 3D, all three regimes	Model B from pairs alone
unseen L: $\{4, 8, 16, 32\} \Rightarrow 64$	lattices beyond the training size

Thank you!