ERG2026, University of Sussex

Solving Functional Renormalization Group Equations with Neural Networks

Phys. Rev. D **114**, 036026 (2026)

Yang-yang Tan

IPI, The University of Tokyo

with **Wei-jie Fu**, **Lianyi He** and **Lingxiao Wang**

September 4, 2026

Introduction

Wetterich equation

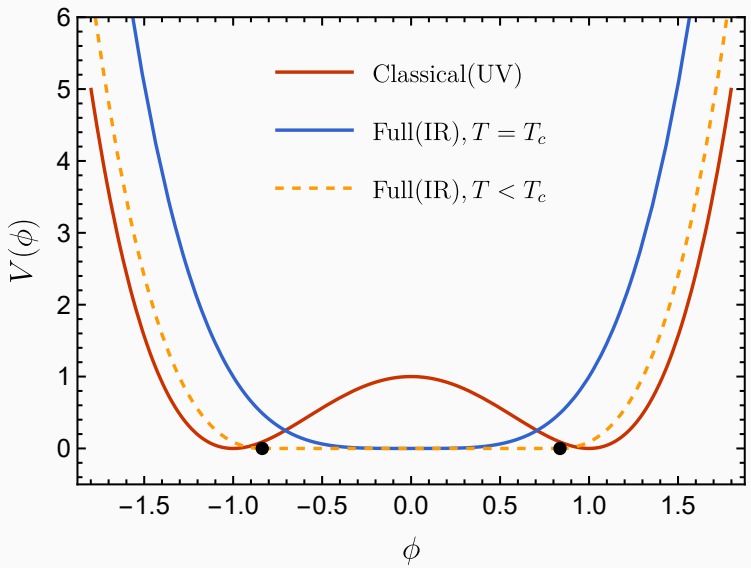
$$\partial_t \Gamma_k = \frac{1}{2} \mathrm{Tr} \left[\partial_t R_k \left(\Gamma_k^{(2)} + R_k \right)^{-1} \right]$$

LPA flow of $V_k(\rho)$, a stiff PDE

Taylor, grids, Chebyshev, DG/FEM, FV

Bottleneck

stiffness, curse of dimensionality



$$\Gamma_{k=\Lambda} = S \longrightarrow \Gamma_{k\to 0} = \Gamma$$

Wetterich: PLB 301, 90 (1993). **Taylor:** Litim, NPB 631, 128 (2002). **Chebyshev:** Borchardt, Knorr, PRD 91, 105011 (2015); PRD 94, 025027 (2016); Chen, Wen, Fu, PRD 104, 054009 (2021). **DG / FEM:** Grossi, Wink, SciPost Phys. Core 6, 071 (2023); Sattler, Pawlowski (DiFfRG), arXiv:2412.13043. **FV:** Zorbach, Koenigstein, Braun, arXiv:2412.16053.

Introduction

Wetterich equation

$$\partial_t \Gamma_k = \frac{1}{2} \mathrm{Tr} \left[\partial_t R_k \left(\Gamma_k^{(2)} + R_k \right)^{-1} \right]$$

LPA flow of $V_k(\rho)$, a stiff PDE

Taylor, grids, Chebyshev, DG/FEM, FV

Bottleneck

stiffness, curse of dimensionality

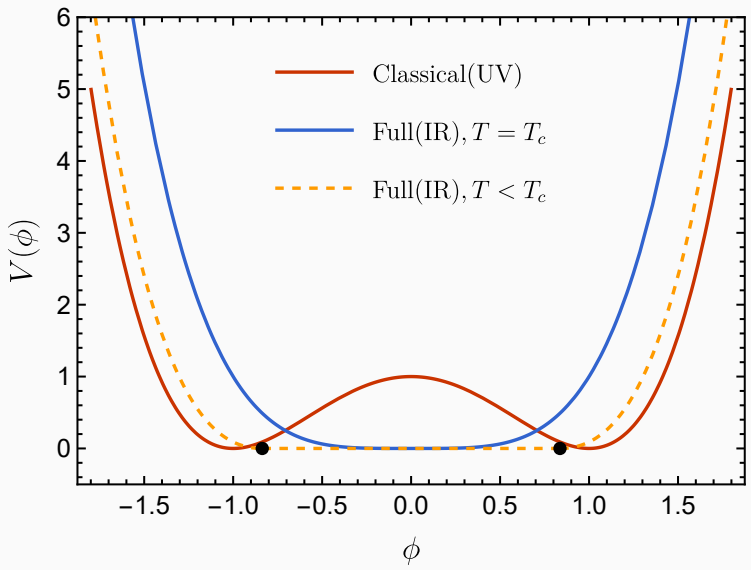
This talk

Neural Networks for the LPA flow

$V_k'(\rho)$ = a neural network

loss = flow-equation residual

learn only the finite-N correction



$$\Gamma_{k=\Lambda} = S \longrightarrow \Gamma_{k\to 0} = \Gamma$$

PINN for fRG on a lattice

Yokota, PRB 109, 214205 (2024)

Takeru's talk

Wetterich: PLB 301, 90 (1993). **Taylor:** Litim, NPB 631, 128 (2002). **Chebyshev:** Borchardt, Knorr, PRD 91, 105011 (2015); PRD 94, 025027 (2016); Chen, Wen, Fu, PRD 104, 054009 (2021). **DG / FEM:** Grossi, Wink, SciPost Phys. Core 6, 071 (2023); Sattler, Pawlowski (DiFfRG), arXiv:2412.13043. **FV:** Zorbach, Koenigstein, Braun, arXiv:2412.16053.

Why a network instead of a grid?

Continuous and differentiable

smooth in $(t, \hat{\rho})$, no interpolation

Flexibility

boundary condition,

physics-driven constraints

Transfer learning

one model seeds a scan in T, μ, N

Physics-driven learning: Aarts, Fukushima, Hatsuda, Ipp, Shi, Wang, Zhou, Nat. Rev. Phys. 7, 154 (2025). **Vertex expansion in QCD:** Mitter, Pawlowski, Strodthoff, PRD 91, 054035 (2015); Fu, Huang, Pawlowski, Tan, Zhou, PRD 112, 054047 (2025). **Operator learning:** DeepONet, Lu et al., Nat. Mach. Intell. 3, 218 (2021); FNO, Li et al., arXiv:2010.08895. **Physics-informed kernels:** Ihssen, Pawlowski, Ann. Phys. 481, 170177 (2025); arXiv:2507.13011; Ihssen, Kapust, Pawlowski, arXiv:2510.26678; arXiv:2603.03159.

Why a network instead of a grid?

Continuous and differentiable

smooth in $(t, \hat{\rho})$, no interpolation

Flexibility

boundary condition,

physics-driven constraints

Transfer learning

one model seeds a scan in T, μ, N

No curse of dimensionality

grid cost $\sim N^{n_{\text{fields/momenta}}}$,

a network only adds inputs

Examples

multi-field V: $2+1$ flavour QCD

vertex expansion: $\Gamma^{(4)}(p_1, p_2, p_3)$,

$32^6 \sim 10^9$ grid points

Physics-driven learning: Aarts, Fukushima, Hatsuda, Ipp, Shi, Wang, Zhou, Nat. Rev. Phys. 7, 154 (2025). **Vertex expansion in QCD:** Mitter, Pawlowski, Strodthoff, PRD 91, 054035 (2015); Fu, Huang, Pawlowski, Tan, Zhou, PRD 112, 054047 (2025). **Operator learning:** DeepONet, Lu et al., Nat. Mach. Intell. 3, 218 (2021); FNO, Li et al., arXiv:2010.08895. **Physics-informed kernels:** Ihssen, Pawlowski, Ann. Phys. 481, 170177 (2025); arXiv:2507.13011; Ihssen, Kapust, Pawlowski, arXiv:2510.26678; arXiv:2603.03159.

Why a network instead of a grid?

Continuous and differentiable

smooth in $(t, \hat{\rho})$, no interpolation

Flexibility

boundary condition,

physics-driven constraints

Transfer learning

one model seeds a scan in T, μ, N

No curse of dimensionality

grid cost $\sim N^{n_{\text{fields/momenta}}}$,

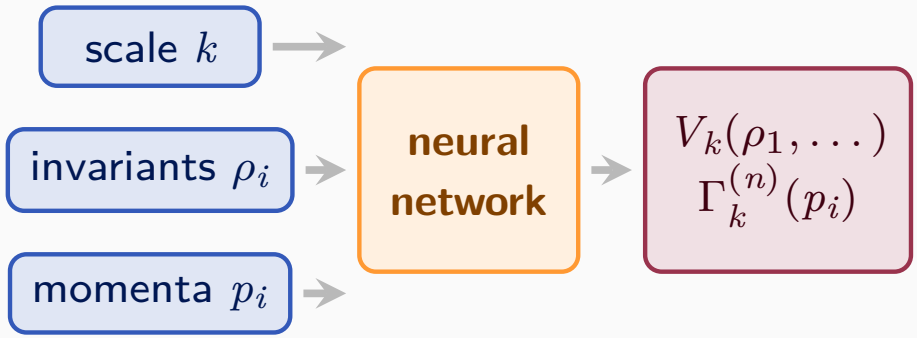
a network only adds inputs

Examples

multi-field V: $2+1$ flavour QCD

vertex expansion: $\Gamma^{(4)}(p_1, p_2, p_3)$,

$32^6 \sim 10^9$ grid points



universal approximators, strong representation power

ML for RG, RG for ML: Jan's talk physics-informed RG: Friederike's talk

Physics-driven learning: Aarts, Fukushima, Hatsuda, Ipp, Shi, Wang, Zhou, Nat. Rev. Phys. 7, 154 (2025). **Vertex expansion in QCD:** Mitter, Pawlowski, Strodthoff, PRD 91, 054035 (2015); Fu, Huang, Pawlowski, Tan, Zhou, PRD 112, 054047 (2025). **Operator learning:** DeepONet, Lu et al., Nat. Mach. Intell. 3, 218 (2021); FNO, Li et al., arXiv:2010.08895. **Physics-informed kernels:** Ihssen, Pawlowski, Ann. Phys. 481, 170177 (2025); arXiv:2507.13011; Ihssen, Kapust, Pawlowski, arXiv:2510.26678; arXiv:2603.03159.

The flow is stiff

$O(N)$ model, LPA (Model A)

$$\partial_t V_k = \mathscr{C}\, T k^d \left[\frac{1}{1+\bar{m}^2_{\sigma,k}} + \frac{N-1}{1+\bar{m}^2_{\pi,k}}\right]$$

$$\bar{m}^2_{\sigma,k} = \frac{V'_k + 2\rho\, V''_k}{k^2}, \qquad \bar{m}^2_{\pi,k} = \frac{V'_k}{k^2}$$

Convexity restoration

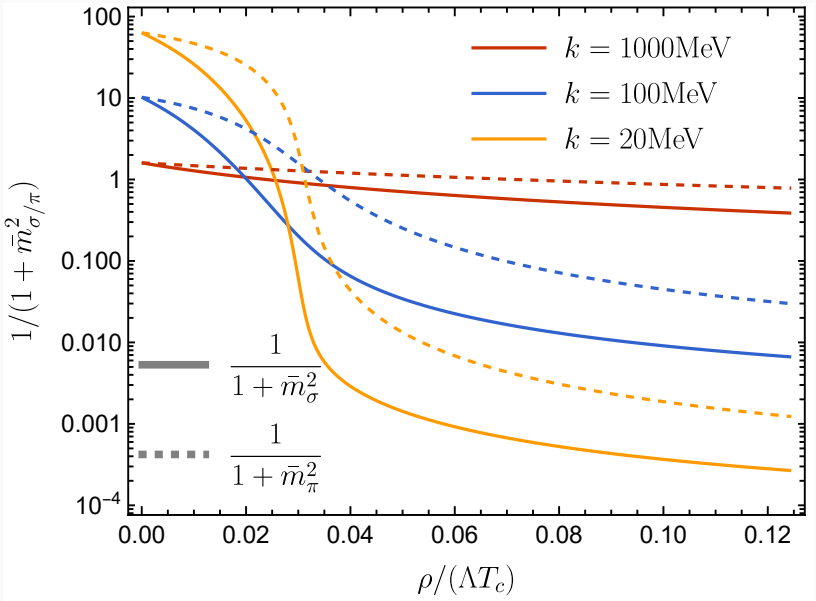
$1 + \bar{m}^2_{\sigma/\pi} \geq 0$ at all k

$\rho < \rho_0$: $V'_k \to -k^2$, so $1 + \bar{m}^2 \to 0^+$

Stiffness in a narrow field window

propagators span $\mathbf{10^{-4}}$–$\mathbf{10^{2}}$,

$\partial_t V'$ spans 10^2–10^7 (k=20 MeV)



$O(4)$, $T = 100\,\mathrm{MeV}$, broken phase.

$\rho = \frac{1}{2}\phi_a\phi_a$, $t = \ln(k/\Lambda)$, $\hat{\rho} = \rho/(\Lambda\, T_c)$. **Model A:** Hohenberg, Halperin, RMP 49, 435 (1977). **Flow equation from real-time fRG:** Tan, Chen, Fu, SciPost Phys. 12, 026 (2022); Chen, Tan, Fu, PRD 109, 094044 (2024).

Learn only the finite-N correction

Decomposition

$$\hat{V}'(t,\hat{\rho}) = \underbrace{\hat{V}'_{\mathrm{LN}}(t,\hat{\rho})}_{\text{analytic}} + \underbrace{\Delta\,\hat{V}'_{\theta}(t,\hat{\rho})}_{\text{network}}$$

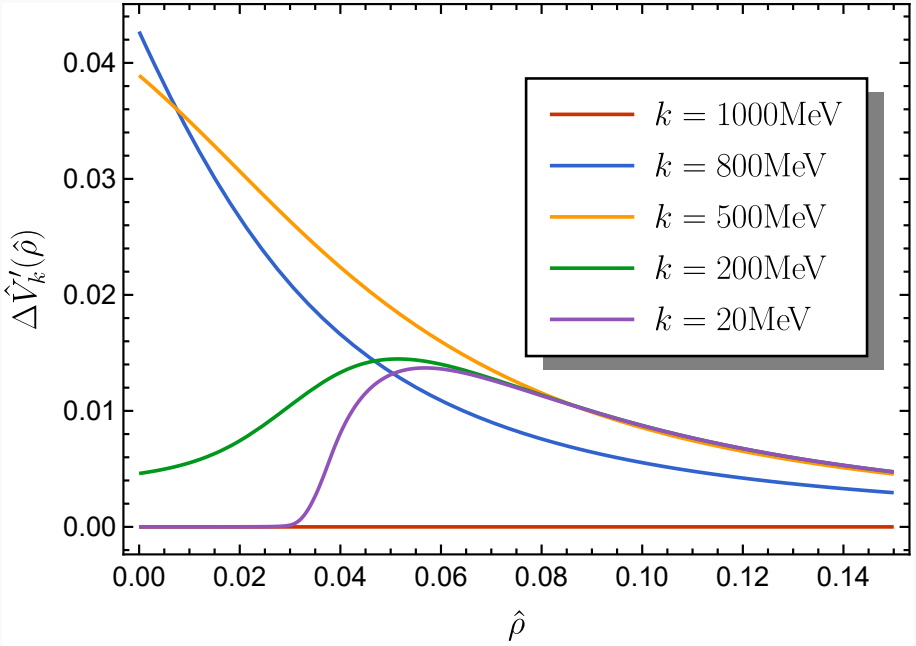
Large-N flow

$$\partial_t V_k = \mathscr{C}\, T k^d \,\frac{N-1}{1+\bar{m}^2_{\pi,k}}$$

solvable by characteristics:
V' constant along them,
ρ_k in closed form $({}_2F_1)$

Why it helps

large N restores convexity;
the rest is smooth and localized



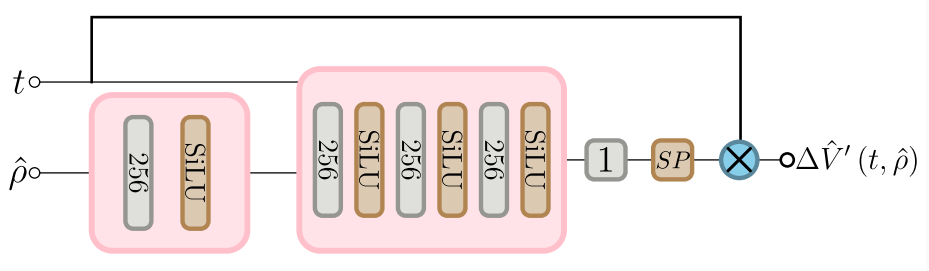
Learned $\Delta\,\hat{V}'_k(\hat{\rho})$, $O(4)$, $T = 100\,\mathrm{MeV}$.

Zero at $k{=}\Lambda$, largest at $k \sim 0.2$–$0.6\,\mathrm{GeV}$.

Initial and boundary conditions

$\Delta\,\hat{V}'_{\theta}(0,\hat{\rho}) = 0$
$\Delta\,\hat{V}'_{\theta}(t,0)$, $\Delta\,\hat{V}'_{\theta}(t,\hat{\rho}_{\max})$: not specified

Solving the flow equation as an optimization problem

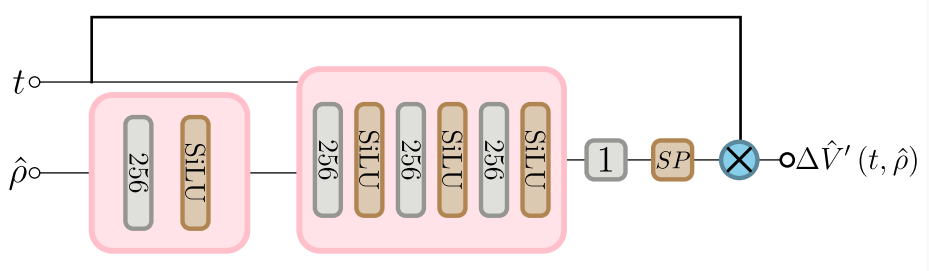


inputs $(t, \hat{\rho}) \rightarrow$ simple nonlinear layers $\rightarrow$ output $\Delta\,\hat{V}'_\theta$

1. Residual insert $\hat{V}'_{\mathrm{LN}} + \Delta\,\hat{V}'_\theta$ into the flow equation on the grid: $\mathcal{R}_{ij}(\theta) = \partial_t\,\hat{V}'_\theta - \mathcal{F}[\,\hat{V}'_\theta,\,\hat{V}''_\theta,\,\hat{V}'''_\theta]\big|_{(\hat{\rho}_i,t_j)}$, zero for the exact solution

Details: 3-point finite differences on a 201×501 grid, 3–$4\times$ less memory than autodiff; Adam; float32. **PINN:** Raissi, Perdikaris, Karniadakis, JCP 378, 686 (2019); Karniadakis et al., Nat. Rev. Phys. 3, 422 (2021). **PINNs for fRG / FDE / DSE:** Yokota, PRB 109, 214205 (2024); Miyagawa, Yokota, NeurIPS 2024; Terin, SciPost Phys. Core 8, 054 (2025). **FD derivatives in PINNs:** Chiu et al., CMAME 395, 114909 (2022).

Solving the flow equation as an optimization problem



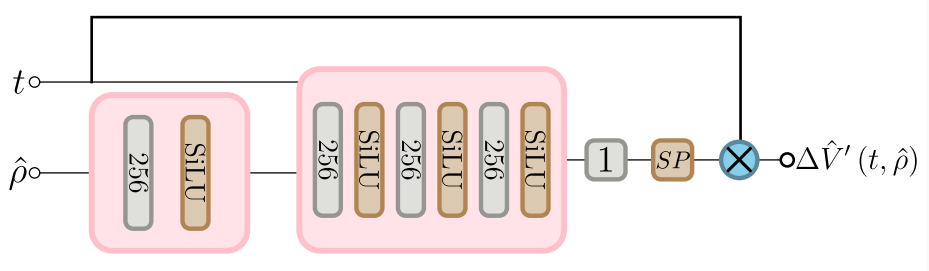
inputs $(t, \hat{\rho}) \to$ simple nonlinear layers $\to$ output $\Delta\,\hat{V}'_\theta$

1. Residual insert $\hat{V}'_{\mathrm{LN}} + \Delta\,\hat{V}'_\theta$ into the flow equation on the grid:
$\mathcal{R}_{ij}(\theta) = \partial_t \hat{V}'_\theta - \mathcal{F}[\hat{V}'_\theta, \hat{V}''_\theta, \hat{V}'''_\theta]\big|_{(\hat{\rho}_i, t_j)}$, zero for the exact solution

2. Solve = minimize $L(\theta) = \frac{1}{N_{\mathrm{grid}}} \sum_{ij} |\mathcal{R}_{ij}(\theta)|^2$ over θ
gradient descent; the equation is the only input, no training data

Details: 3-point finite differences on a 201×501 grid, 3–4$\times$ less memory than autodiff; Adam; float32. **PINN:** Raissi, Perdikaris, Karniadakis, JCP 378, 686 (2019); Karniadakis et al., Nat. Rev. Phys. 3, 422 (2021). **PINNs for fRG / FDE / DSE:** Yokota, PRB 109, 214205 (2024); Miyagawa, Yokota, NeurIPS 2024; Terin, SciPost Phys. Core 8, 054 (2025). **FD derivatives in PINNs:** Chiu et al., CMAME 395, 114909 (2022).

Solving the flow equation as an optimization problem



inputs $(t, \hat{\rho}) \to$ simple nonlinear layers $\to$ output $\Delta\,\hat{V}'_\theta$

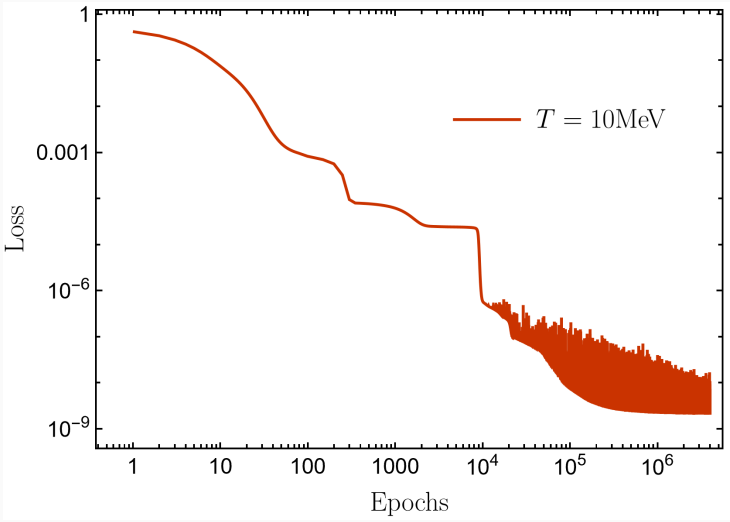
1. Residual insert $\hat{V}'_{\mathrm{LN}} + \Delta\,\hat{V}'_\theta$ into the flow equation on the grid: $\mathcal{R}_{ij}(\theta) = \partial_t \hat{V}'_\theta - \mathcal{F}[\hat{V}'_\theta, \hat{V}''_\theta, \hat{V}'''_\theta]\big|_{(\hat{\rho}_i, t_j)}$, zero for the exact solution

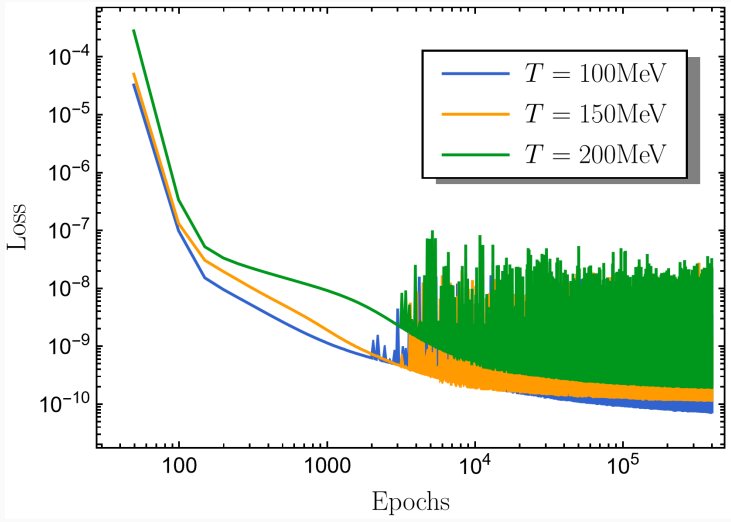
2. Solve = minimize $L(\theta) = \frac{1}{N_{\mathrm{grid}}} \sum_{ij} |\mathcal{R}_{ij}(\theta)|^2$ over θ
gradient descent; the equation is the only input, no training data

Built in initial condition (output $\propto t$); no field boundary condition

Details: 3-point finite differences on a 201×501 grid, 3–4$\times$ less memory than autodiff; Adam; float32. **PINN:** Raissi, Perdikaris, Karniadakis, JCP 378, 686 (2019); Karniadakis et al., Nat. Rev. Phys. 3, 422 (2021). **PINNs for fRG / FDE / DSE:** Yokota, PRB 109, 214205 (2024); Miyagawa, Yokota, NeurIPS 2024; Terin, SciPost Phys. Core 8, 054 (2025). **FD derivatives in PINNs:** Chiu et al., CMAME 395, 114909 (2022).

Training





From scratch, $T{=}10\,\mathrm{MeV}$, $4{\times}10^6$ epochs

Transfer from $T{=}120\,\mathrm{MeV}$, $5{\times}10^5$ epochs

Optimizer

Adam, lr 5×10^{-4}

float32, then float64

Domain

$k \in [20\,\mathrm{MeV}, 1\,\mathrm{GeV}]$

$\hat{\rho} \in [0, 0.15]$

grid 201×501

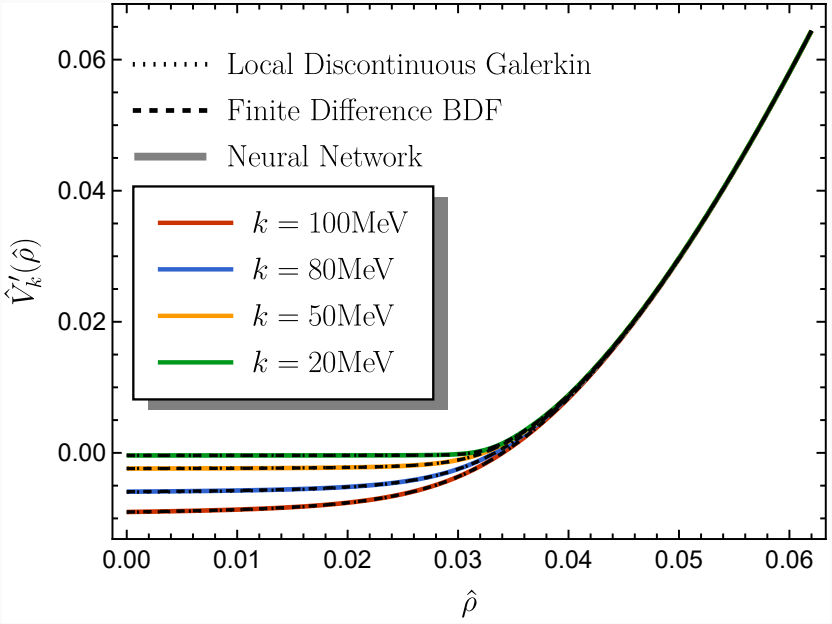
Convergence

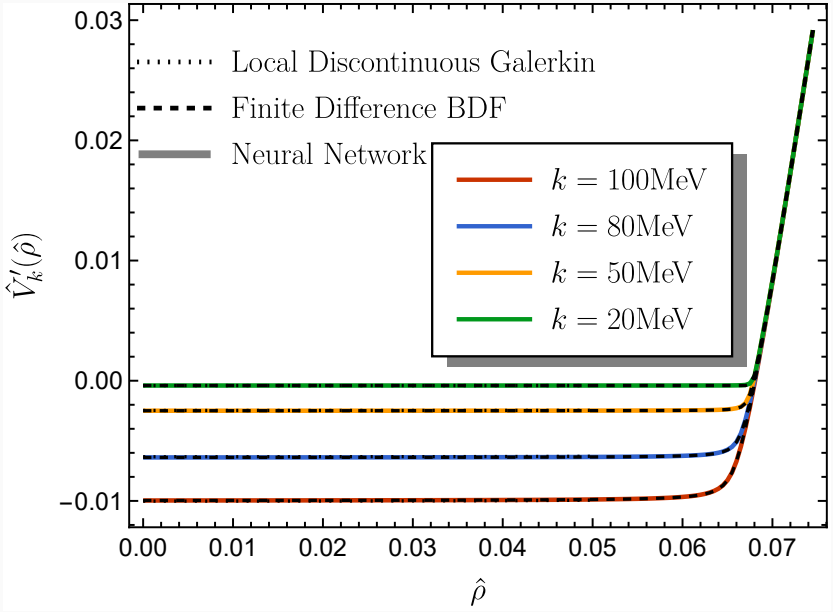
staircase descent,

plateau at $\sim 10^{-9}$

Grid 2001×501 and float64 refinement for the $T = 10\,\mathrm{MeV}$ run. Denominators regularized by $\epsilon = 10^{-13}$. PyTorch, PhysicsNeMo, Lightning.

Broken phase: $\hat{V}_k'(\hat{\rho})$ vs. LDG and finite differences





$T = 100\,\text{MeV},\ k = 100, 80, 50, 20\,\text{MeV}$

$T = 10\,\text{MeV}$, deeply broken, same k

Benchmarks

LDG and FD-BDF

Accuracy

absolute error

$< 4 \times 10^{-5}$

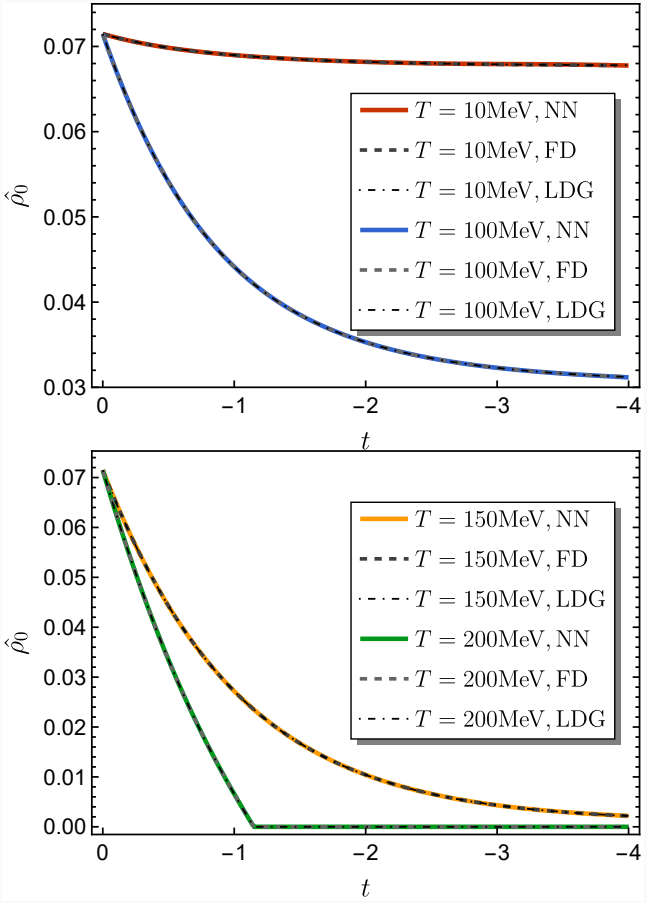
at all k and $\hat{\rho}$

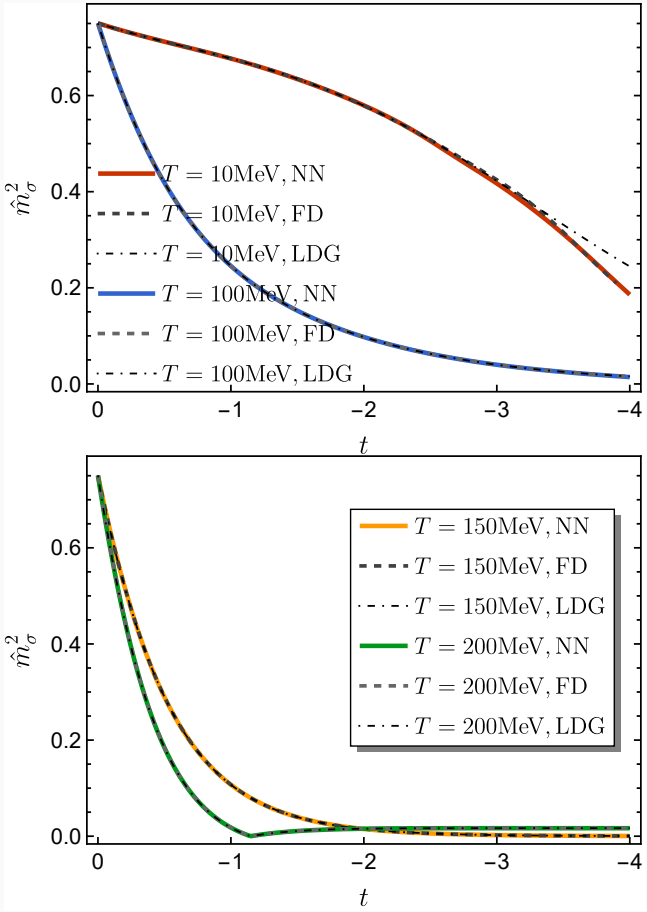
Stiff region

no oscillations at the crossover $\hat{\rho}_0(k)$

LDG benchmark (DiFfRG): Sattler, Pawlowski, arXiv:2412.13043.

Observables across the phase transition





Order parameter $\hat{V}'(\hat{\rho}_0) = 0$

$T < T_c$: $\hat{\rho}_0$ finite, $T \geq T_c$: $\hat{\rho}_0 \to 0$

Curvature mass

$$\hat{m}_\sigma^2 = \hat{V}'(\hat{\rho}_0) + 2\hat{\rho}_0 \hat{V}''(\hat{\rho}_0)$$

Same idea for the Wilson–Fisher fixed point

Fixed-point equation, $d = 3$, LPA

$$0 = -3u + \bar{\rho}u' + \mathscr{C}\left[\frac{1}{1+u'+2\bar{\rho}u''} + \frac{N-1}{1+u'}\right]$$

Shooting is delicate

fine tuning to isolate WF from Gaussian and multicritical branches

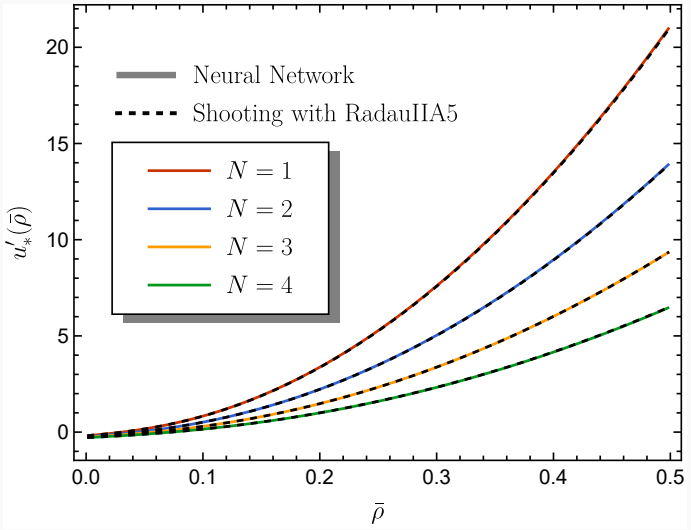
Network

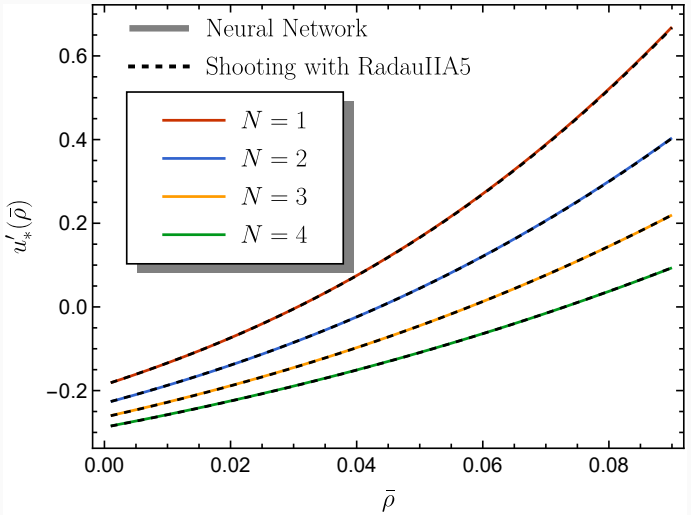
$u' = u'_{\mathrm{LN}} + \Delta u'_\theta$

u'_{LN} in closed form $({}_2F_1)$

4-layer MLP, ODE residual loss

no shooting, no Gaussian collapse





Shooting benchmark: Tan, Huang, Chen, Fu, EPJC 84, 897 (2024).

The large-N limit is not essential

Baselines from the equation itself

large field: $u'_* \to A\bar{\rho}^2,\ A = 3\gamma \simeq 84$

$$u' = A\bar{\rho}^2 + \Delta u'_\theta$$

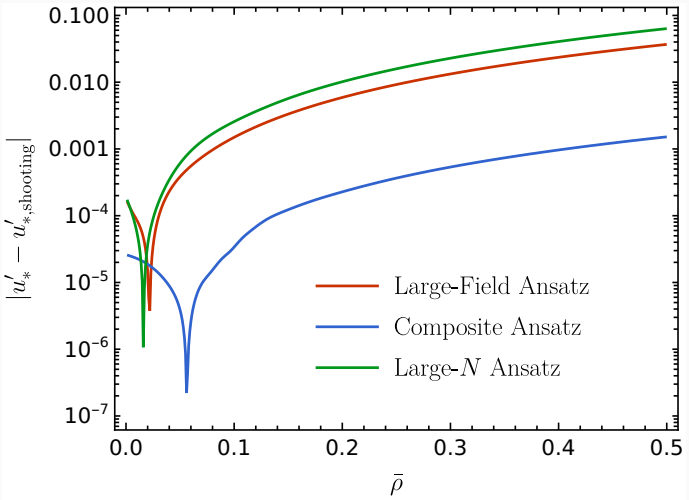
Composite ansatz $(O(1),\ d = 3)$

$$u'_{\text{base}} = u'_{\text{poly}}\,[1 - w] + A\bar{\rho}^2\, w$$

$$w = \tfrac{1}{2}\{1 + \tanh\left[(\bar{\rho} - 0.1)/0.02\right]\}$$

u'_{poly}: 10th-order Taylor series

$$u' = u'_{\text{base}} + \Delta u'_\theta$$



Large-field asymptotics, $A = 3\gamma$: Tan, Huang, Chen, Fu, EPJC 84, 897 (2024).

The large-N limit is not essential

Baselines from the equation itself

large field: $u'_* \to A\bar{\rho}^2,\ A = 3\gamma \simeq 84$

$$u' = A\bar{\rho}^2 + \Delta u'_\theta$$

Composite ansatz $(O(1),\ d = 3)$

$$u'_{\text{base}} = u'_{\text{poly}}\,[1 - w] + A\bar{\rho}^2\, w$$

$$w = \tfrac{1}{2}\{1 + \tanh\left[(\bar{\rho} - 0.1)/0.02\right]\}$$

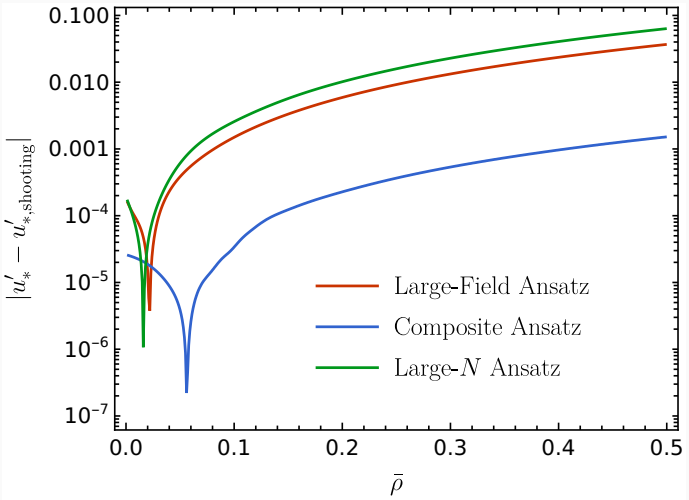
u'_{poly}: 10th-order Taylor series

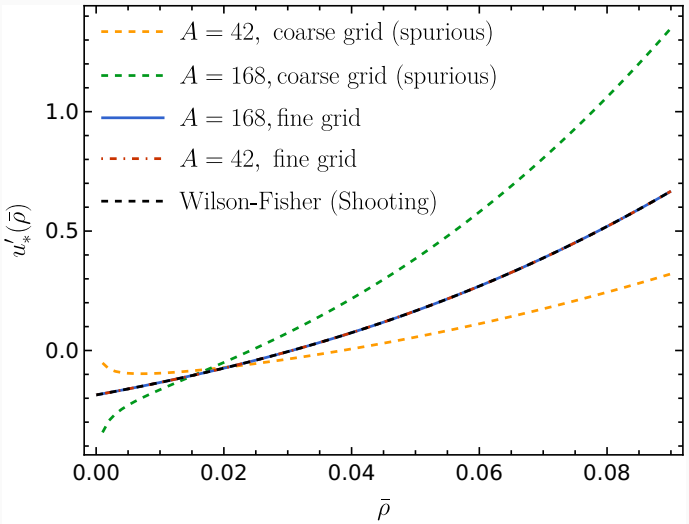
$$u' = u'_{\text{base}} + \Delta u'_\theta$$

Mistuned amplitude $A/2$ or $2A$

coarse grid: a spurious solution

fine grid: WF recovered, A to 0.005%





Large-field asymptotics, $A = 3\gamma$: Tan, Huang, Chen, Fu, EPJC 84, 897 (2024).

Summary

Flow equation in the loss $\longrightarrow$ $V_k(\rho)$ as a neural network

Finite-T $O(4)$ flows	Key ingredient	Wilson–Fisher, $N \leq 4$
all three regimes	$\hat{V}' = \hat{V}'_{\mathrm{LN}} + \Delta\,\hat{V}'_{\theta}$	no shooting
symmetric, broken, critical	stiffness absorbed	any baseline with the
error $< 4 \times 10^{-5}$	analytically	right asymptotics

Outlook: multi-field potentials, vertex expansion, operator learning, links to physics-informed kernels

Thank you!