Stochastic RG and Gradient Flow

ANDREA CAROSSO

WITH ANNA HASENFRATZ AND ETHAN T. NEIL

UNIVERSITY OF COLORADO BOULDER

5/2/2019

(ARXIV:1806.01385, 1811.03182, 1904.13057)

Beyond Standard Model Physics and Gradient Flow

Many models of interest in BSM physics are strongly coupled

Anomalous dimensions of local operators near an IRFP are of special importance

Existing methods: MCRG, Dirac eigenmodes, hyperscaling, …

These have drawbacks: ensemble matching, discreteness, limited to few operators

Meanwhile: Gradient flow (GF) is a smoothing transformation that has been used to define renormalized quantities nonperturbatively on the lattice (Lüscher, 2009)

For scalars, “free” gradient flow is defined by a heat equation

$$\partial_t \phi_t(x) = \Delta \phi_t(x), \quad \phi_0(x) = \varphi(x)$$

The solution is a local average of nearby fields

$$\phi_t(x) = f_t \varphi(x), \quad f_t(z) = \frac{\mathrm{e}^{-z^2/4t}}{(4\pi t)^{d/2}}$$

Gradient Flow as RG?

Traditional spin-blocking RG: define blocked spins as local averages

$$\varphi_b(n/b) = \frac{b^{\Delta_\phi}}{b^d} \sum_\varepsilon \varphi(n+\varepsilon)$$

This field transformation suggests a natural effective action definition

$$\mathrm{e}^{-S_b(\phi)} = \int_\varphi \delta(\phi - \varphi_b) \mathrm{e}^{-S_0(\varphi)}$$

The lattice spacing of the blocked lattice is $a' = ba$

Observables in the blocked theory can be computed with MCRG

$$\langle \mathcal{O}(\phi) \rangle_{S_b} = \langle \mathcal{O}(\varphi_b) \rangle_{S_0}$$

GF suggests that a continuous version of a blocking transformation can be defined: integration of the bare field against the heat kernel looks like a *smoothened* version of the block-spin local average

… but there’s no mention of field rescaling, and it’s not clear what the definition of the effective action would be

Functional RG

Already in the early 70's, a non-perturbative definition of continuous RG transformations was provided by Wilson and Kogut

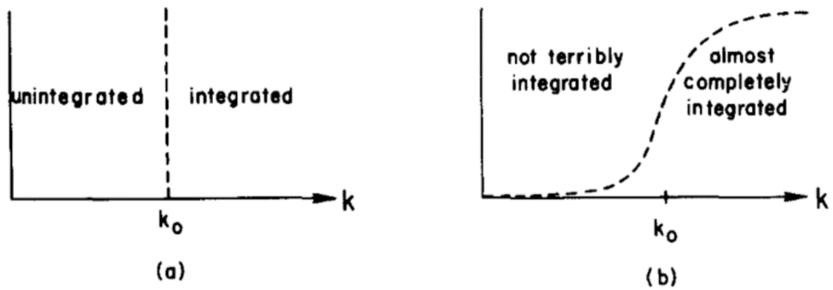
$$\mathrm{e}^{-S_t(\phi)} = \int_\varphi P_t(\phi,\varphi)\mathrm{e}^{-S_0(\varphi)}$$

The function P_t is a "constraint functional"

$$P_t(\phi,\varphi) = N_t \exp\Big[-\frac{1}{2}\int_p \frac{\nu(p)\big(\phi(p)-f_t(p)\varphi(p)\big)^2}{1-f_t^2(p)}\Big]$$

The effective Boltzmann factor satisfies a Fokker-Planck equation

$$\frac{\partial P_t(\phi)}{\partial t} = \frac{1}{2}\int_p \Big(K_0(p)\frac{\delta^2 P_t(\phi)}{\delta\phi(p)\delta\phi(-p)} + \omega(p)\phi(p)\frac{\delta P_t(\phi)}{\delta\phi(p)}\Big)$$



Adapted from
Wilson & Kogut (1973)

Stochastic RG (*arxiv: 1904.13057*)

Define the RG transformation as the solution of a simple Langevin equation

$$\partial_t \phi_t(p) = -p^2 \phi_t(p) + \eta_t(p)$$

The definition of the Fokker-Planck distribution is

$$P(\phi, t; \varphi, 0) = \mathbb{E}_{\mu_0}\left[\delta(\phi - \phi_t[\varphi; \eta])\right]$$

One may compute the distribution explicitly

$$P(\phi, t; \varphi, 0) = N_t \exp\left[-\frac{1}{2}\left(\phi - f_t\varphi, A_t^{-1}(\phi - f_t\varphi)\right)\right]$$

$$A_t(p, k) = (2\pi)^d \delta(p + k) K_0(k) \frac{1 - f_t^2(k)}{2k^2}$$

where $K_0(p) = \mathrm{e}^{-p^2/\Lambda_0^2}$ is a (Schwinger) cutoff function

Similar form as the constraint functionals in Wilson and Kogut (1973) or Wetterich (1991)

Effective Action and IRFP

The effective action can be written in terms of the bare theory's generator of connected Green functions

$$S_t(\phi) = F_t + \frac{1}{2}(\phi, A_t^{-1}\phi) - W_0^{(t)}(A_t^{-1} f_t \phi)$$

The tree-level 2-point function implies an effective (inverse) cutoff

$$\Lambda_t^{-2} = \Lambda_0^{-2} + 2t$$

To allow for the possibility of an IRFP, one must define rescaled variables

$$p = \Lambda_t \bar{p} \qquad \phi(p) = \Lambda_0^{d_\phi} b_t^{-\Delta_\phi} \Phi(\bar{p})$$

where the scale factor is defined by

$$b_t = \frac{\Lambda_0}{\Lambda_t}$$

An IRFP implies that rescaled observables can have nontrivial infinite time limits

$$\langle \Phi(\bar{p}_1) \cdots \Phi(\bar{p}_n) \rangle_{S_t} = \Lambda_0^{-d_\phi} b_t^{n\Delta_\phi} \langle \phi(p_1) \cdots \phi(p_n) \rangle_{S_t}$$

A study of the effective action in the case of phi4 in 3d demonstrates that an interacting IRFP exists, as expected

Stochastic MCRG

By writing expectation values of the effective theory in terms of the noise expectations, one finds an equivalence

$$\langle \mathcal{O}(\phi) \rangle_{S_t} = \langle \mathbb{E}_{\mu_0} \big[\mathcal{O}(\phi_t[\varphi;\eta]) \big] \rangle_{S_0}$$

MCRG in the sense that expectations in the effective theory may be computed without knowledge of the effective action

Numerically implementable: generate an ensemble of bare fields with usual lattice Monte Carlo, and integrate the Langevin equation on every configuration

If the form of b(t) is known, then we have access to the rescaled effective theory

$$\langle \Phi(\bar{p}_1) \cdots \Phi(\bar{p}_n) \rangle_{S_t} = \Lambda_0^{-d_\phi} b_t^{n\Delta_\phi} \langle \mathbb{E}_{\mu_0} \big[\phi_t(p_1) \cdots \phi_t(p_n) \big] \rangle_{S_0}$$

But not all observables require a full Langevin equation simulation: we'll see that long-distance quantities of the effective theory may be computed with gradient flow!

Effective Correlations and Gradient Flow

Connected N-point functions of the effective theory are related to gradient flowed n-points

$$\langle \phi(x)\phi(y)\rangle_{S_t}^{\text{conn}} = \langle f_t\varphi(x) f_t\varphi(y)\rangle_{S_0}^{\text{conn}} + A_t(x-y)$$

$$\langle \phi(x_1)\cdots\phi(x_n)\rangle_{S_t}^{\text{conn}} = \langle f_t\varphi(x_1)\cdots f_t\varphi(x_n)\rangle_{S_0}^{\text{conn}}$$

The function $A_t(x-y)$ is determined by the choice of LE but decays like a Gaussian at distances much greater than the effective cutoff

Connected correlators of composite operators are more complicated. For example, the phi2 operator 2-point function is given by

$$\langle \phi^2(x)\phi^2(y)\rangle_{S_t}^{\text{conn}} = \langle (f_t\varphi)^2(x)(f_t\varphi)^2(y)\rangle_{S_0}^{\text{conn}}$$

$$+A_t(x-y)\langle f_t\varphi(x) f_t\varphi(y)\rangle^{\text{conn}} + 2A_t(x-y)^2$$

Moral: The effective (connected) correlators are asymptotically equal to the corresponding gradient flow correlators. Short-distance expectations require the full LE simulation

Scaling Formulae

Any valid RG transformation will imply scaling formulae for the rescaled effective fields, and this can be checked for the stochastic RG transformation

For 2-point functions of scaling operators ($\bar{x} = x\Lambda_t$),

$$\langle \mathcal{O}(\bar{x}_1)\mathcal{O}(\bar{x}_2)\rangle_{S_{t+\epsilon}} \approx b_\epsilon(t)^{2\Delta_\mathcal{O}} \langle \mathcal{O}(\bar{y}_1)\mathcal{O}(\bar{y}_2)\rangle_{S_t}$$

where $b_\epsilon(t) = b_{t+\epsilon}/b_t$ is a *relative* scale factor

Recall that within expectation values of long-distance observables,

$$\Phi(\bar{x}) = b_t^{\Delta_\phi} \hat{\phi}(x) \approx b_t^{\Delta_\phi} f_t \hat{\varphi}(x)$$

This implies a *ratio formula*

$$\frac{\langle \mathcal{O}_{t+\epsilon}(x_1)\mathcal{O}_{t+\epsilon}(x_2)\rangle_{S_0}}{\langle \mathcal{O}_t(x_1)\mathcal{O}_t(x_2)\rangle_{S_0}} \approx b_\epsilon(t)^{2(\Delta_\mathcal{O} - m\Delta_\phi)}$$

Since $\Delta_\mathcal{O} - m\Delta_\phi = \gamma_\mathcal{O} - m\gamma_\phi$, one can use this to measure anomalous dimensions of scaling operators, if the form of b(t) is known

3d Scalar Field Theory

$$S(\hat{\varphi}) = \sum_n \Big[-\beta \sum_\mu \hat{\varphi}(n)\hat{\varphi}(n+\mu) + \hat{\varphi}^2(n) + \lambda(\hat{\varphi}^2(n) - 1)^2 - \lambda \Big]$$

RG transformations map the theory towards the IRFP (WFFP) when the system is tuned to the critical surface

The WFFP is described by a set of well-known exponents

$$\eta, \nu, \omega$$

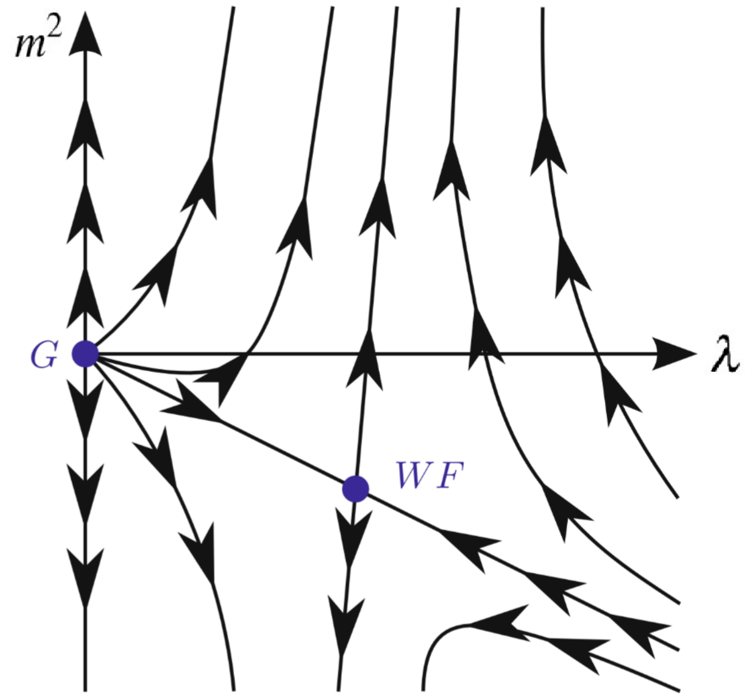
They're related to anomalous dimensions:

$$\gamma_\phi = \eta/2 = 0.0179(2)$$

$$\gamma_{\phi^2} = 2 - \nu^{-1} = 0.411(10)$$

$$\gamma_{\phi^3} = 1 + \gamma_\phi = 1.0179(2)$$

$$\gamma_{\phi^4} = 4 - d + \omega = 1.845(10)$$



Adapted from Kopietz *et al., Introduction to the Functional Renormalization Group (Springer 2010)*

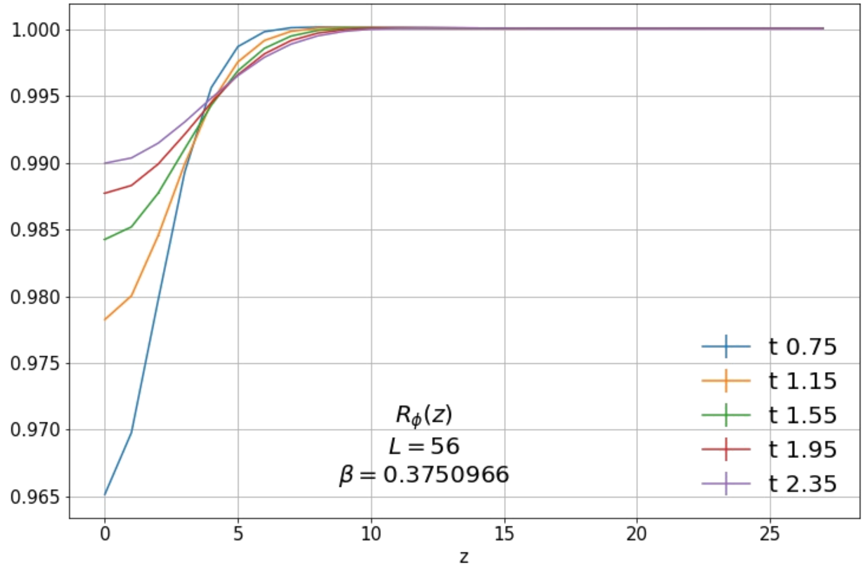
Phi Correlator

The scaling formula implies that the phi-phi ratios are approximately independent of time and distance at large distances:

$$R_\phi(t, z) = \frac{\langle \phi_{t+\epsilon}(z)\phi_{t+\epsilon}(0)\rangle}{\langle \phi_t(z)\phi_t(0)\rangle} \approx 1$$

so one cannot use this to measure the anomalous dimension of phi

The plateau “shrinks” as t increases, i.e. the blocking radius grows



Phi2 Correlator

This one should be time-dependent:

$$R_{\phi^2}(t,z) = \frac{\langle \phi^2_{t+\epsilon}(z)\phi^2_{t+\epsilon}(0)\rangle}{\langle \phi^2_t(z)\phi^2_t(0)\rangle} \approx b_\epsilon(t)^{\delta_2}$$

$$\delta_2 = 2(\gamma_2 - 2\gamma_1) \approx 0.752$$

Long-distance ratio shows clear movement

We fit the time dependence at fixed z0 using the ansatz

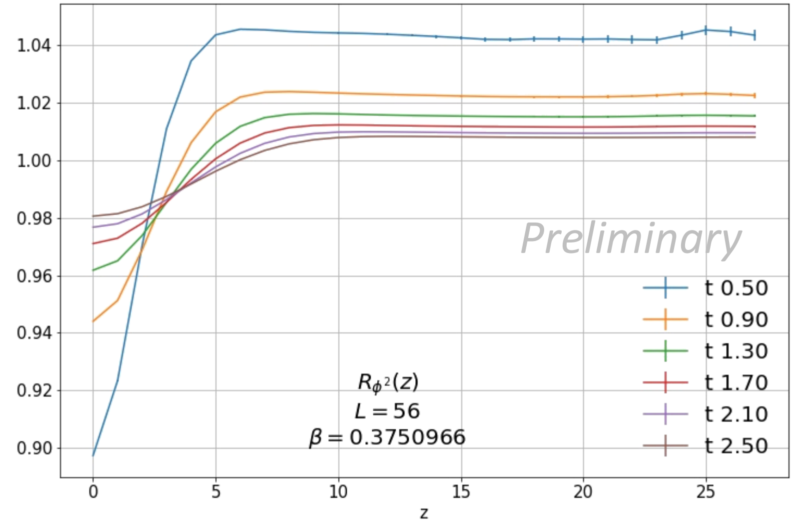
$$b_t = \sqrt{1+ct}$$

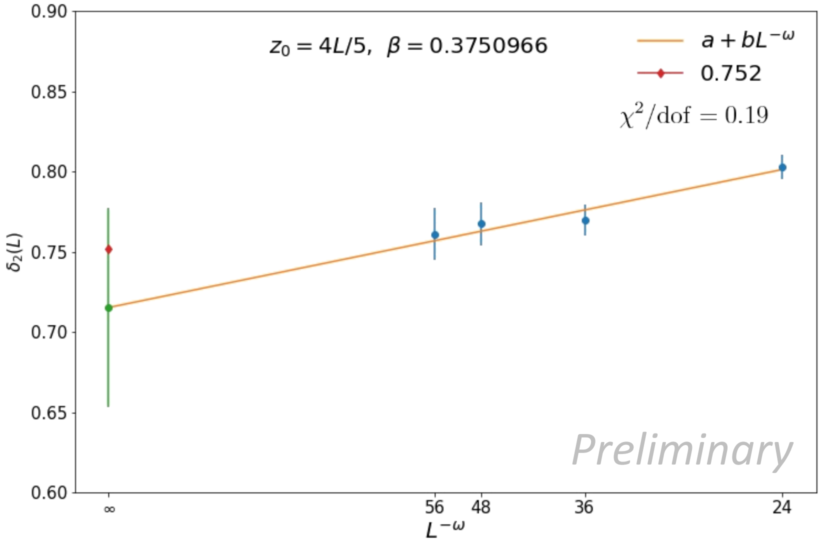
which implies a relative scale factor

$$b_\epsilon(t) = \left(1 + \frac{\epsilon}{c^{-1}+t}\right)^{1/2}$$

Try extrapolating to L = ∞ with inverse powers of L:

$$\delta_2(\infty) = 0.715(62)$$
$$b = 1.3(2.3)$$
$$\omega = 0.86(77)$$





Diagonalization Method

Ratios for higher operators like ϕ^3 and ϕ^4 do not give expected results

Likely due to operator mixing: higher ops are dominated by the contribution of the leading relevant ops, ϕ and ϕ^2 (confirmed by fits)

…need to isolate the scaling operators!

Recall that scaling ops are linear combinations of action ops

$$\mathcal{O}_a(\Phi) = \sum_i c_{ai} S_i(\Phi)$$

They scale simply under RG:

$$\mathcal{O}_{t+\epsilon} = b_\epsilon^{\Delta_{\mathcal{O}}} \mathcal{O}_t$$

And their mixed correlations vanish

$$\langle \mathcal{O}_a \mathcal{O}_b \rangle_{S_t} \propto \delta_{ab}$$

Can determine them numerically, in principle, by measuring mixed correlators of action ops and diagonalizing the matrix of correlations, with an appropriate rescaling:

$$O_a\left(b_t^{\Delta_\phi} \phi_t\right) = b_t^{n_a \Delta_\phi} \sum_i c_{ai} b_t^{(n_i - n_a)\Delta_\phi} S_i(\phi_t)$$

but results so far have been hindered by very poor signals at large distances

Generalization to Gauge-Fermion Systems

Define the RG transformations of the gauge and (staggered) fermion fields with the simplest diffusion equations that preserve their symmetry:

- Gauge fields evolve according to Wilson flow (Lüscher, 2009)

$$\partial_t U_\mu(x,t) = -g_0^2 (\partial S_W[U])(x,t) U_\mu(x,t)$$

- Fermions evolve with a gauge-covariant heat equation

$$\partial_t \psi(x,t) = \Delta[U] \psi(x,t)$$

At long distances, flowed-correlators should exhibit RG scaling of the fixed point if the system is tuned towards criticality

Nf=12, SU(3) gauge theory is expected to be conformal or near-conformal, so the ratio formula should be applicable

The mass and pseudoscalar anomalous dimensions are related: $\gamma_m = -\gamma_{\mathrm{ps}}$

$$P_t(x) = \overline{\psi}_t(x) \varepsilon(x) \psi_t(x)$$

Can also try measuring the baryon anomalous dimension, γ_N

$$B_t(x) = \epsilon_{abc} \psi_t^a(x) \psi_t^b(x) \psi_t^c(x)$$

Super Ratios

An issue with the ratio formula is that it includes the (usually unknown) anomalous dimension of the fundamental field, e.g.

$$R_P(t) \propto t^{\gamma_{\mathrm{ps}} - 2\gamma_\psi}$$

And we cannot measure γ_ψ directly from the ratio $R_\psi(t)$

Note: if an operator *A* has no anomalous dimension, then its ratio formula is

$$R_A(t) \propto t^{-n_A \gamma_\psi}$$

This could be used to measure γ_ψ, or to cancel it's effect in another ratio. Thus we may form the *super-ratio*

$$R_P(t) R_A(t)^{-n_P/n_A} \propto t^{\gamma_{\mathrm{ps}}}$$

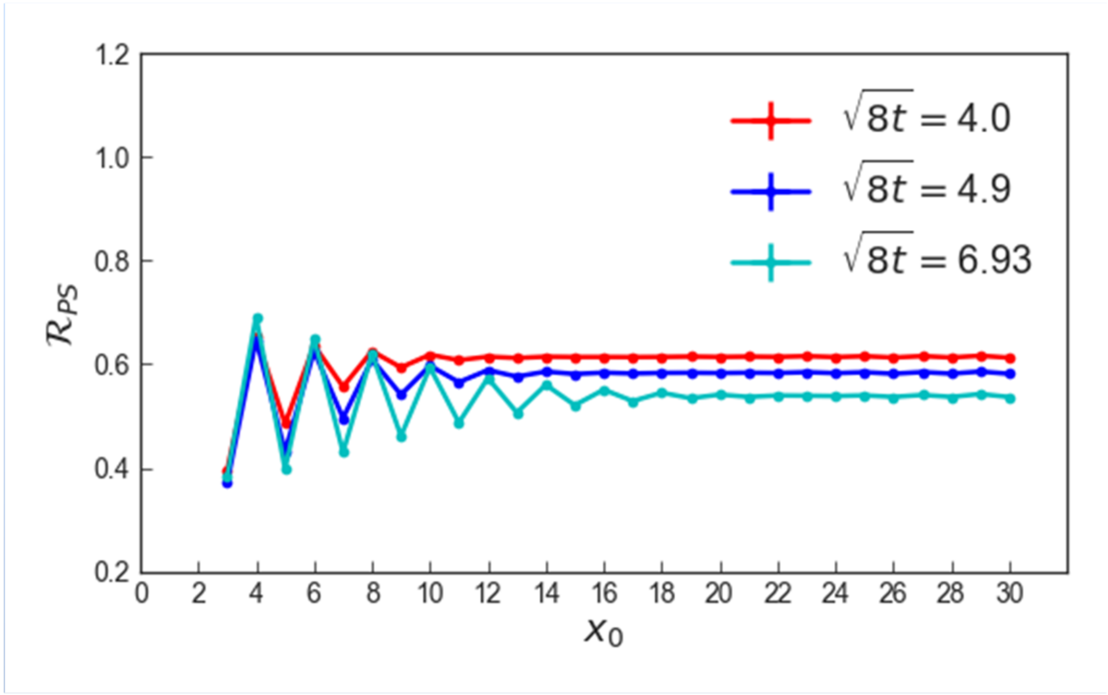
We choose the axial vector A_4 as our conserved operator

Pseudoscalar Ratios

$$R_P(t) \propto t^{\gamma_{ps} - 2\gamma_\psi}$$

The ratios of P-P correlators exhibit the expected plateaus at large distance

Short-distance smearing effects oscillate due to averaging nearby staggered fermions



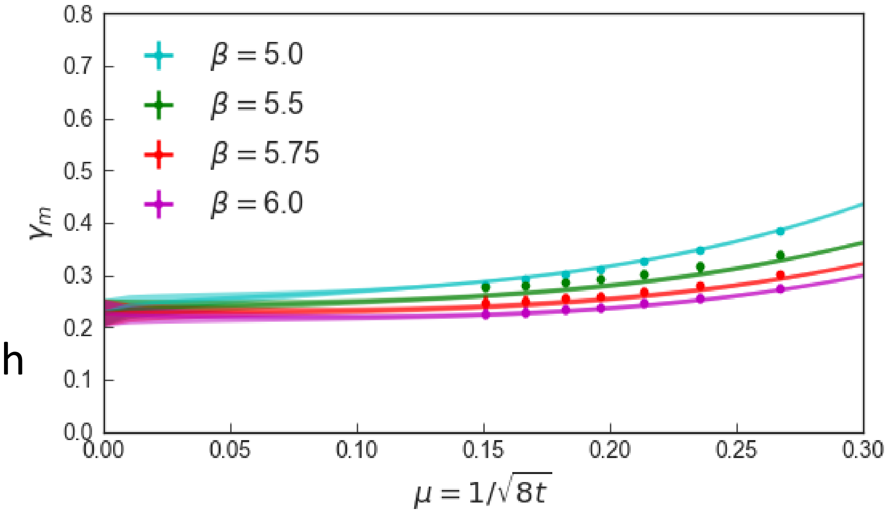
Carosso, et al., *Phys. Rev. Lett.* ***121***, *201601*

Anomalous Dimensions

Infinite volume, infinite time extrapolation yields

$$\gamma_m = 0.23(6)$$

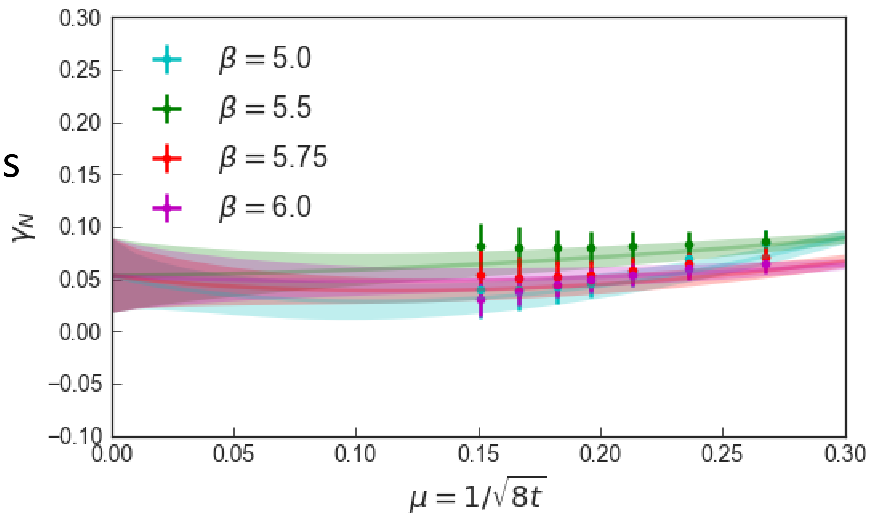
Consistent with several previous studies, both lattice and perturbative



Extrapolation of the nucleon anomalous dimension

$$\gamma_N = 0.05(5)$$

First non-perturbative prediction of γ_N for this system!



Carosso, et al., *Phys. Rev. Lett.* ***121***, *201601*

Conclusion and Future Work

Stochastic RG is a well-defined RG transformation: it can have nontrivial infrared fixed points, implies scaling formulae

It leads to a new type of MCRG on the lattice

Gradient flow can be used to study the effective theory and critical properties, using the equivalence of long-distance correlators and ratio formulae

Leading exponents in a given symmetry subspace are easiest to measure; higher exponents require diagonalization (which gets noisy)

Future work: Can potentially avoid noisiness by working with local observables, but this requires a full Langevin simulation. In particular, a continuous counterpart to the equations proposed by Swendsen (1980’s) seems possible

Binder Cumulant

Tuning to the critical surface:

$$U_4 = 1 - \frac{\langle M^4 \rangle}{3\langle M^2 \rangle^2}$$

extrapolates to a universal value as

$$U_4 = U_4^* + c_1(\lambda) L^{-\omega}$$

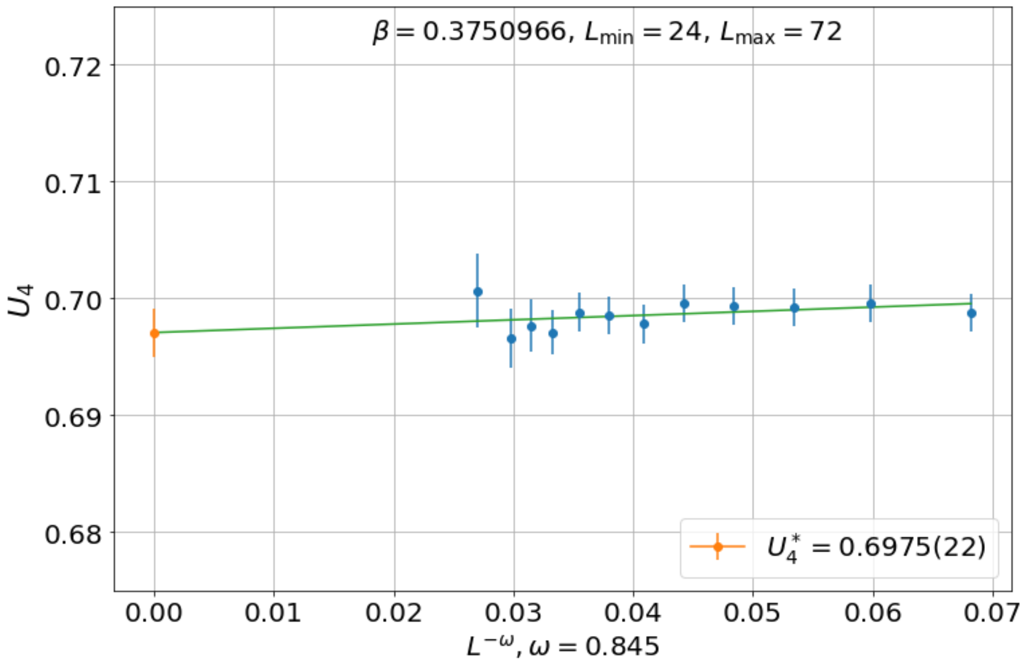
Hasenbusch found that $c_1(\lambda)$ was smallest at $\lambda = 1.1$

He estimated the critical value

$$U_4^* = 0.69819(12)$$

and at $\lambda = 1.1$,

$$\beta = 0.3750966$$



Simulation Details

Configurations were generated with the Wolff cluster method and Metropolis updates for the radial component of the field

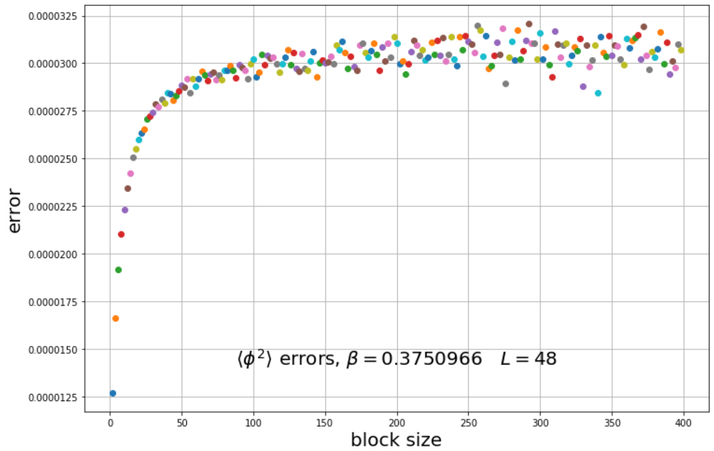
1 sweep = 5 radial updates + 1 cluster update

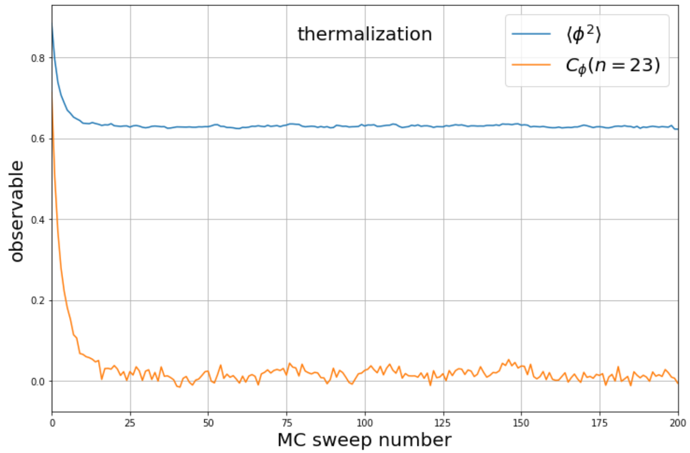
The radial update dominates the autocorrelation: $\tau_{\text{int}} \approx 4.19 - 5.34$

Binned errors plateau around size 100 – implies a consistent τ_int

Flow measurements made every 5 sweeps: flowed data bins of size 20

System appears thermalized by about 50 in both cases – conservatively took 10000 warms; total sweeps = 1 million so 200k measurements, and ~10k independent samples

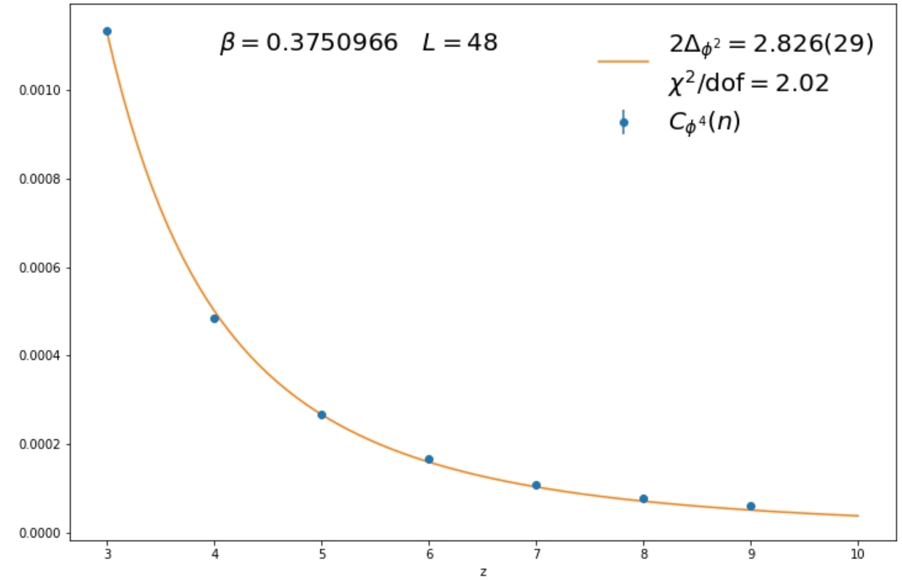


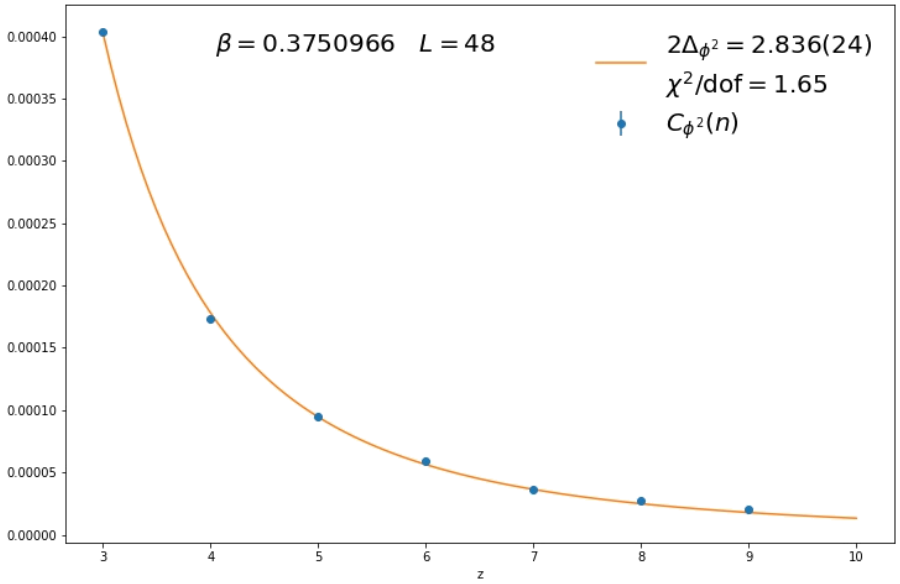


Power-Law Behavior

Power law fits are much better than exponentials, as expected near criticality in a conformal system

Same exponents: phi2 dominates the even subspace, strong operator mixing



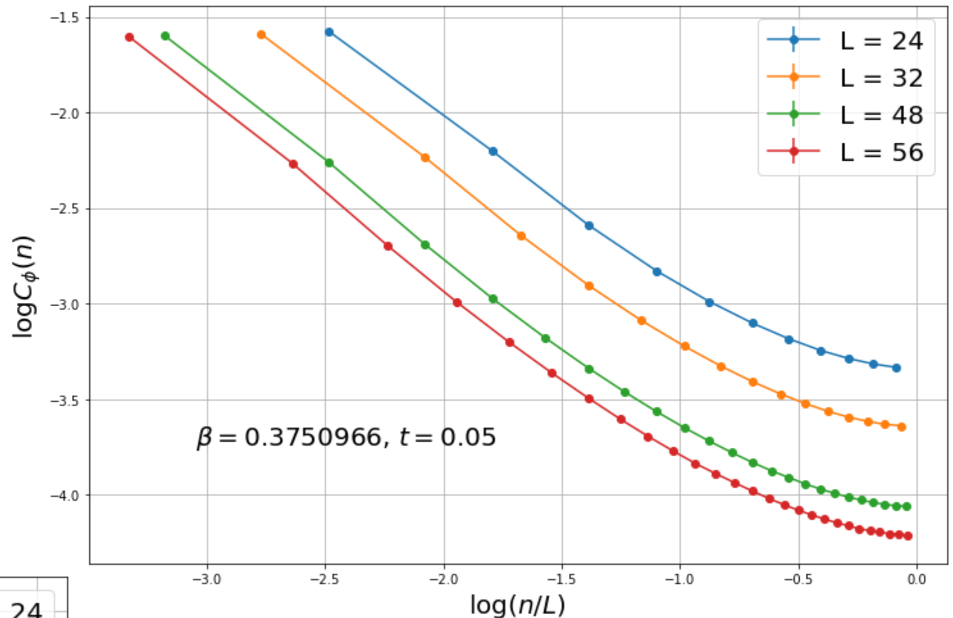


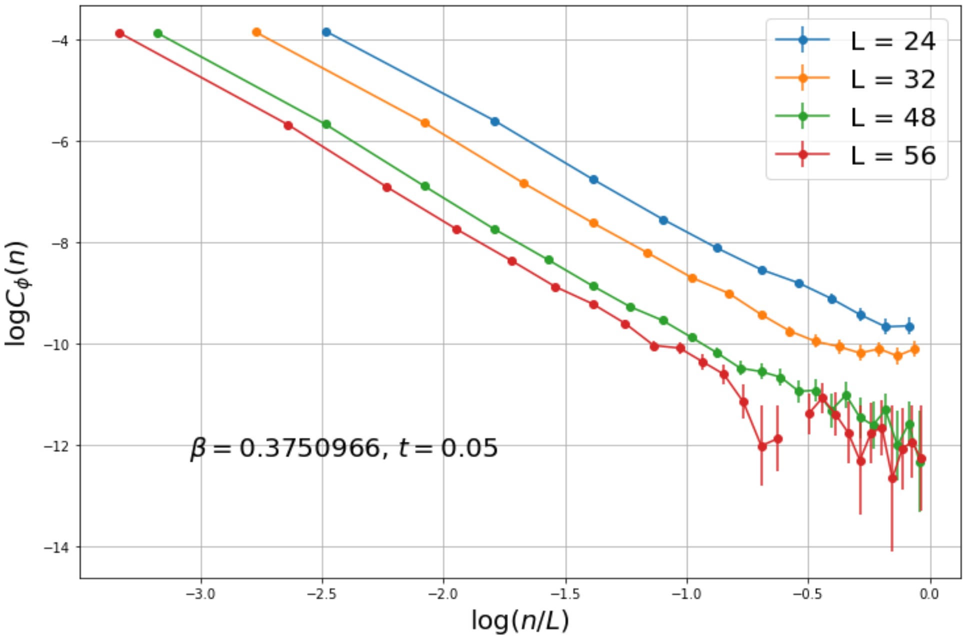
The same issue was observed for the phi and phi3 correlators

Correlator Noise

Phi-phi correlator is the cleanest (right); higher ops are noisier at large distances (below), due to the necessity of vacuum subtractions

This was 100k-sweep data, but the problem persists even for 1 million sweep ensemble





Phi3 results so far…

Diagonalization of the operator basis $\{\phi \,, \phi^3\}$

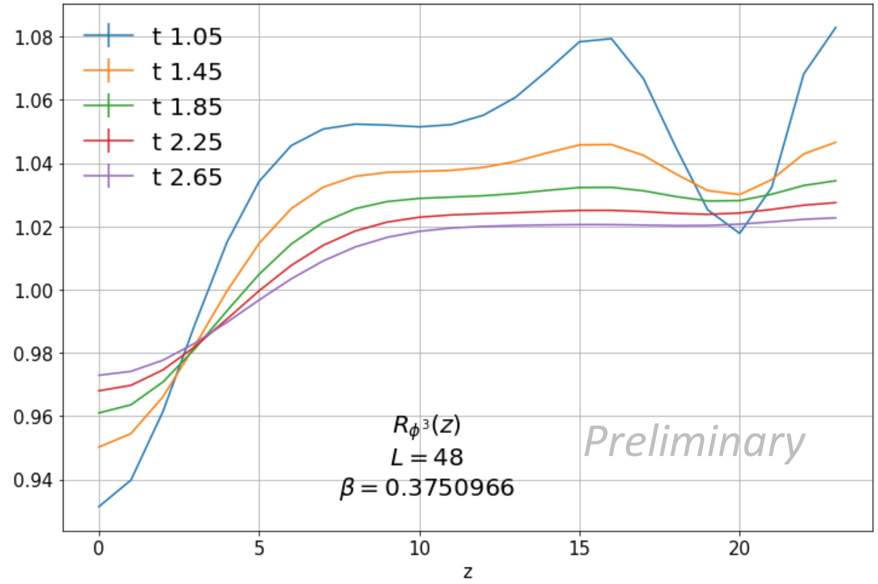
The phi3 scaling correlator has a power law form with exponent ~4.6 (expected ~5), so it seems to produce the right correlator

But the signal is very poor, noise fluctuations are apparent in the ratio plot

Some stable plateau regions can be found; fits to b(t) yield values for δ_3 that are consistent with the expected value 1.928, but insufficient data to perform an infinite volume extrapolation

Typical fit result in the range z0 = 9 – 12

on L = 48: $\delta_3 = 2.09(37)$



Scaling Formulae

The stochastic RG transformation is a time-homogeneous Markov process, and general observables satisfy an equation

$$\partial_t \langle \mathcal{O}(\phi) \rangle_{S_t} = \langle \mathcal{L}\mathcal{O}(\phi) \rangle_{S_t}$$

For discrete time steps, n-point functions satisfy

$$\langle \phi(x_1) \cdots \phi(x_n) \rangle_{S_{t+\epsilon}} = \langle \phi(x_1) \cdots \phi(x_n) \rangle_{S_t} + O\left(\epsilon \mathrm{e}^{-x_{ij}^2 \Lambda_t^2}\right)$$

This implies a scaling formula for rescaled n-point functions (analogue of spin-blocking scaling formulae)

$$\langle \Phi(\bar{x}_1) \cdots \Phi(\bar{x}_n) \rangle_{S_{t+\epsilon}} \approx b_\epsilon(t)^{n\Delta_\phi} \langle \Phi(\bar{y}_1) \cdots \Phi(\bar{y}_n) \rangle_{S_t}$$

where $b_\epsilon(t) = b_{t+\epsilon}/b_t$ is the *relative* scale factor and $\bar{x}$ is a dimless position:

$$\bar{x} = x\Lambda_t$$

The generalization to scaling operators is (2-point case)

$$\langle \mathcal{O}(\bar{x}_1)\mathcal{O}(\bar{x}_2) \rangle_{S_{t+\epsilon}} \approx b_\epsilon(t)^{2\Delta_\mathcal{O}} \langle \mathcal{O}(\bar{y}_1)\mathcal{O}(\bar{y}_2) \rangle_{S_t}$$