Classically perfect gradient flows

from machine-learned fixed-point actions

Urs Wenger
Institute for Theoretical Physics
Albert Einstein Center for Fundamental Physics

in collaboration with Kieran Holland (University of Pacific), Andreas Ipp and David Müller (TU Wien)

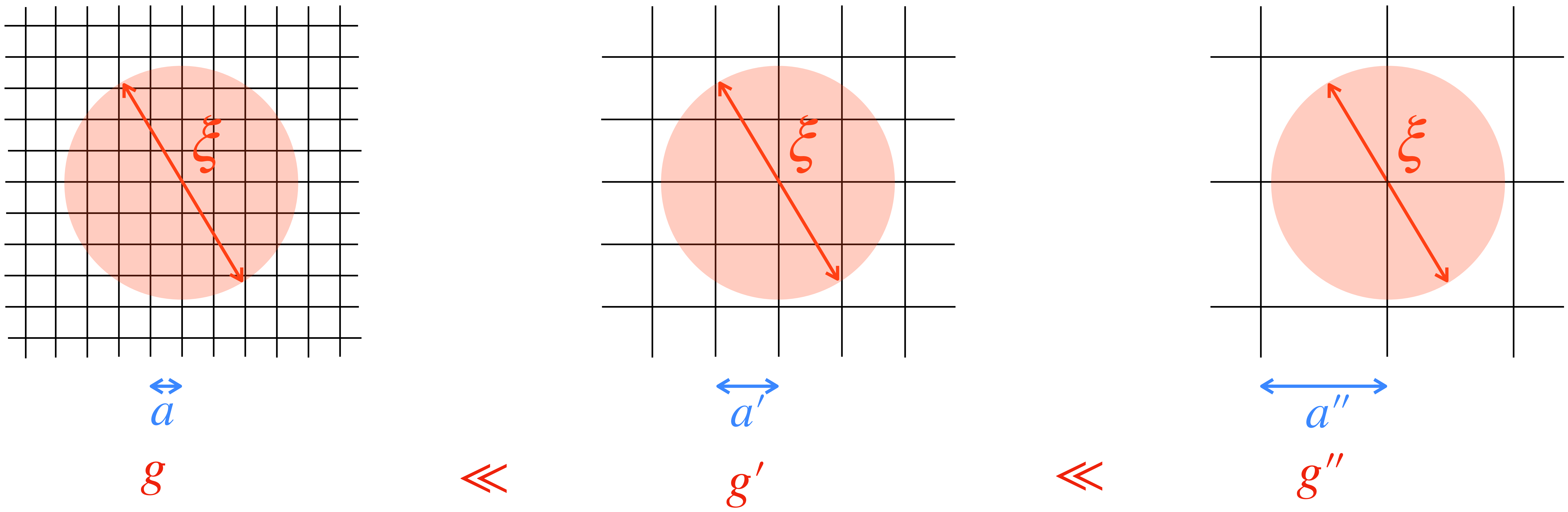
arXiv:2311.17816 [hep-lat], arXiv:2401.06481 [hep-lat], arXiv:2502.03315 [hep-lat]

Gradient Flow Workshop, 13 February 2025 - University of Zurich, Switzerland

Introduction

The lattice spacing a is determined by the gauge coupling: $\beta = \dfrac{2N_c}{g^2}$

continuum limit (2nd order phase transition $\xi/a \to \infty$)

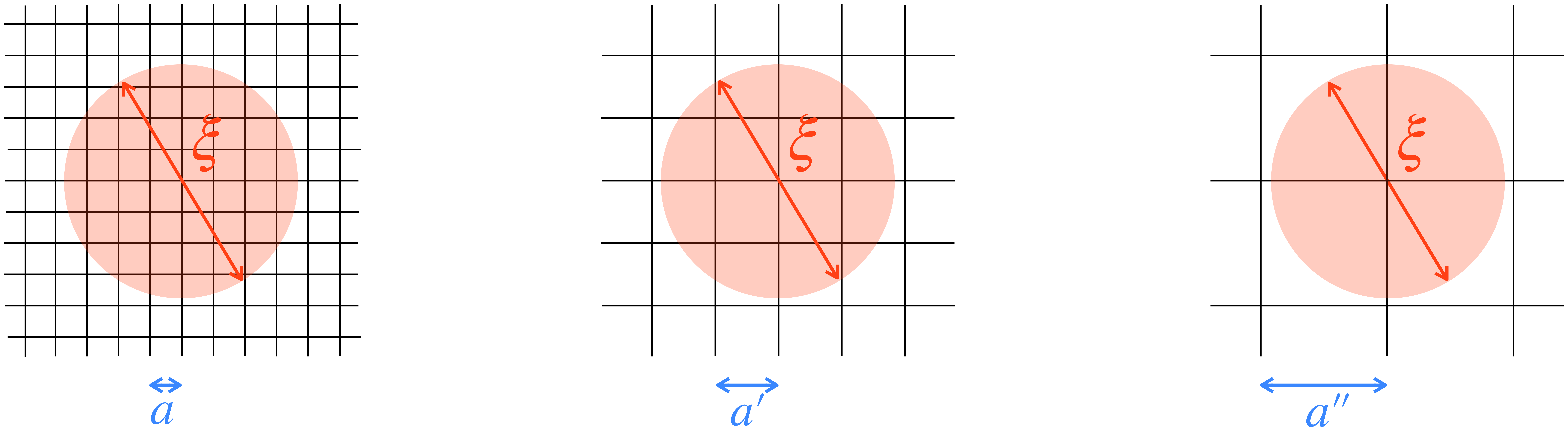


lattice spacing a from dimensionless g: $\Rightarrow$ dimensional transmutation

Introduction

The lattice spacing a is determined by the gauge coupling: $\beta = \dfrac{2N_c}{g^2}$

continuum limit (2nd order phase transition $\xi/a \to \infty$)



critical slowing down
(topological freezing) ?

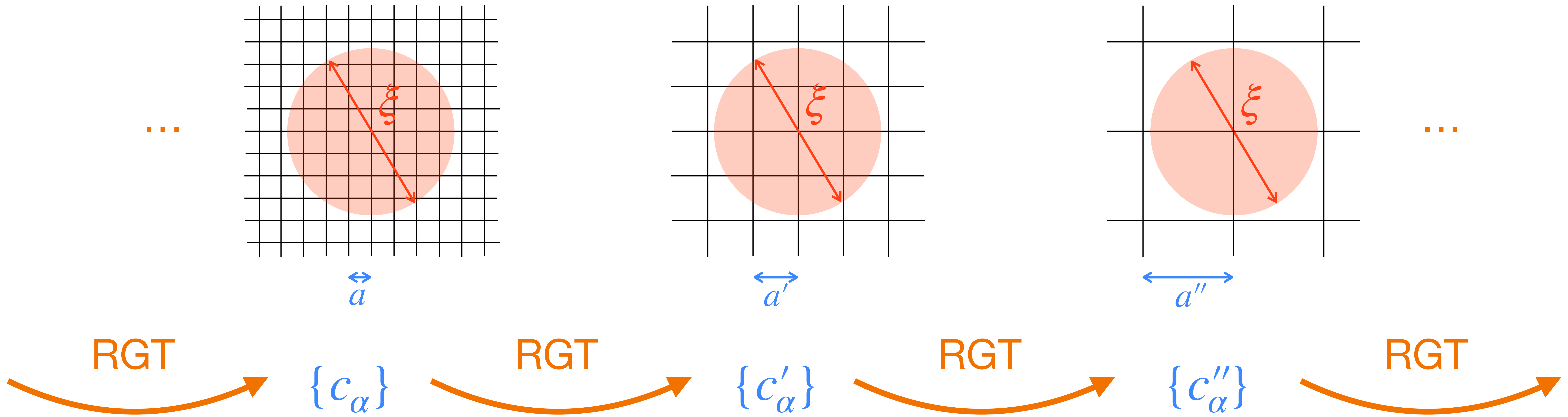
large lattice artefacts?

Renormalization group transformation

Introduce (coordinate space) renormalization group transformation (RGT):

$$\exp\left\{-\beta' A'[V]\right\} = \int \mathscr{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

The effective action $\beta' A'[V]$ is described by infinitely many couplings $\{c'_\alpha\}$:



Renormalization group transformation

Introduce (coordinate space) renormalization group transformation (RGT):

$$\exp\left\{-\beta' A'[V]\right\} = \int \mathcal{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

where $T[U,V]$ is a blocking kernel relating the fine gauge links U to the coarse gauge links $V \equiv U'$:

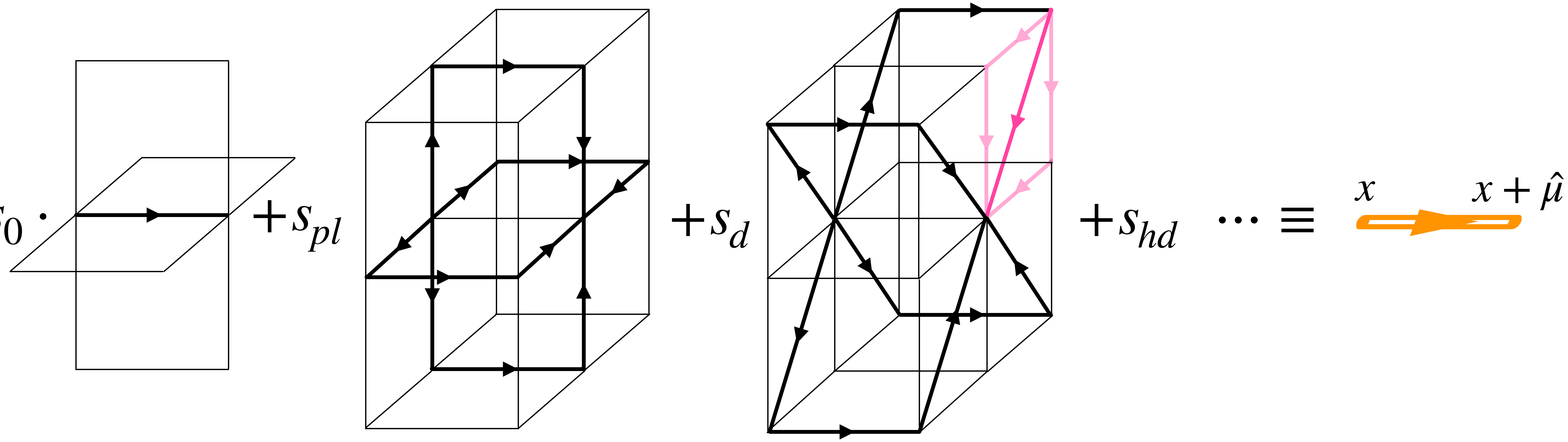
$$T[U,V] = -\frac{\kappa}{N_c}\sum_{x_B,\mu}\left\{\mathrm{ReTr}\left(V_\mu(x_B)\cdot Q_\mu^\dagger(x_B)\right) - \mathcal{N}_\mu^\beta\right\}$$

($\mathcal{N}_\mu^\beta$ is a normalization factor guaranteeing $Z(\beta') = Z(\beta)$, i.e., unchanged long distance physics)

RGT blocking kernel

$$T[U,V] = -\frac{\kappa}{N_c}\sum_{x_B,\mu}\left\{\mathrm{ReTr}\left(V_\mu(x_B)\cdot Q_\mu^\dagger(x_B)\right) - \mathcal{N}_\mu^\beta\right\}$$

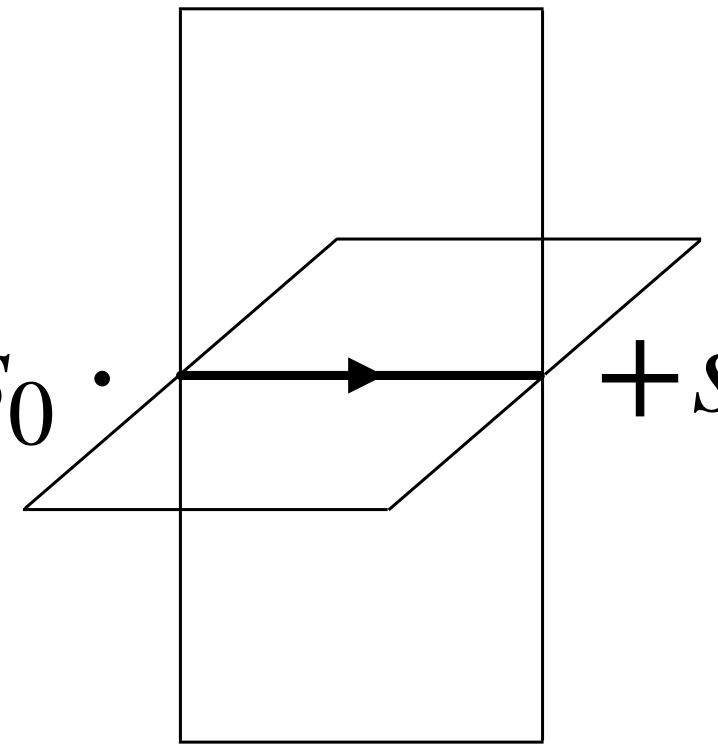
$$S_\mu^{\text{smeared}}[U] = s_0 \cdot \quad + s_{pl} \quad + s_d \quad + s_{hd} \quad \cdots \equiv$$



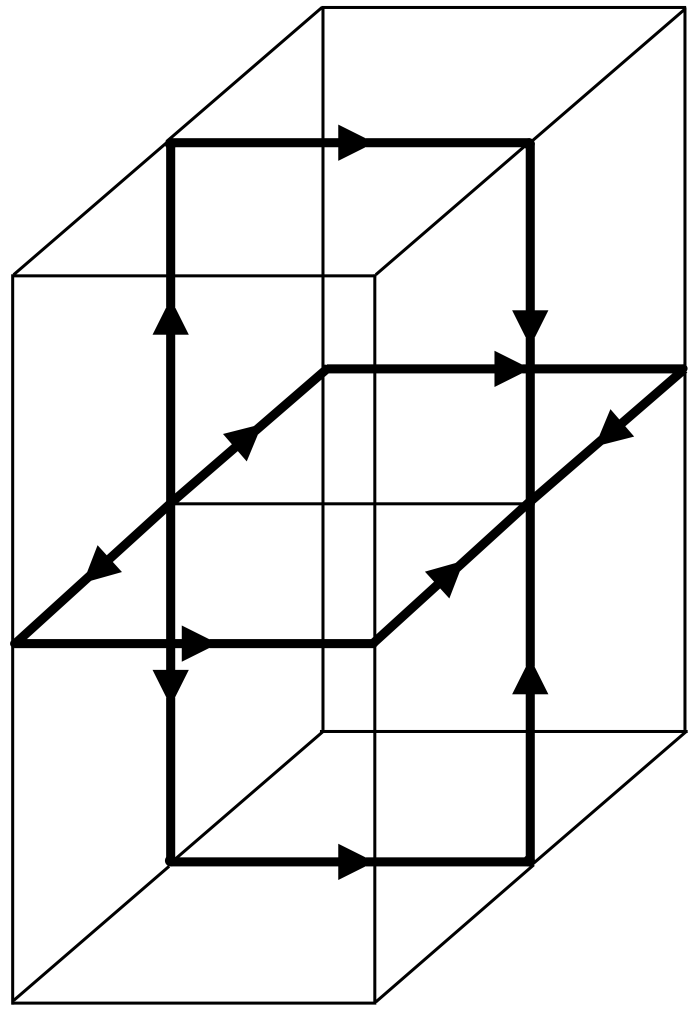
RGT blocking kernel

$$T[U,V] = -\frac{\kappa}{N_c}\sum_{x_B,\mu}\left\{\mathrm{ReTr}\left(V_\mu(x_B)\cdot Q_\mu^\dagger(x_B)\right) - \mathcal{N}_\mu^\beta\right\}$$

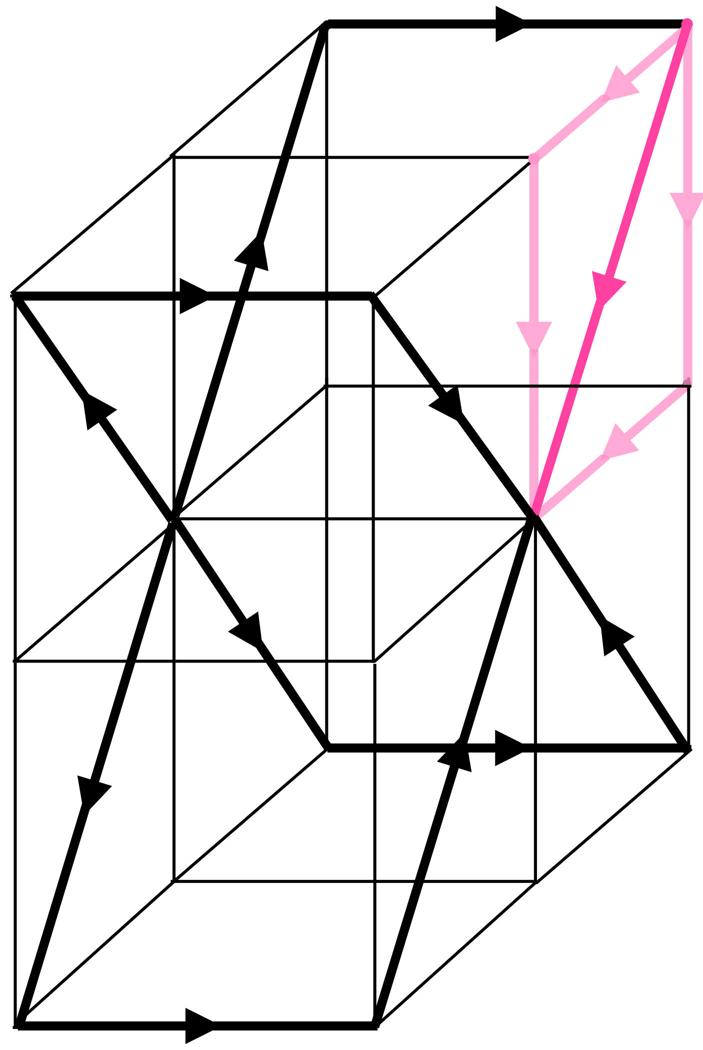
$S_\mu^{\text{smeared}}[U] = s_0 \cdot$



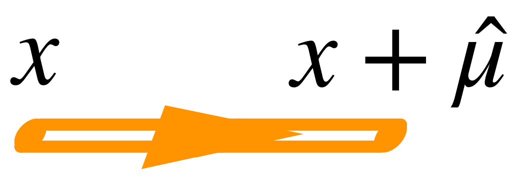
$+ s_{pl}$



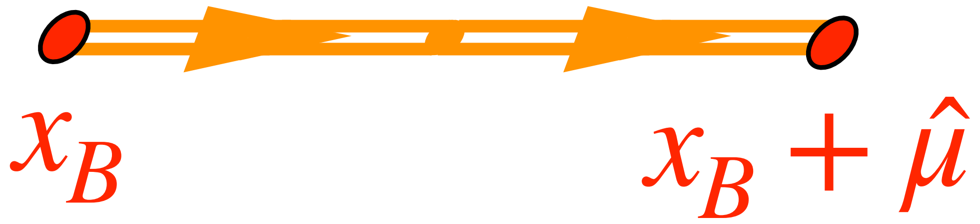
$+ s_d$



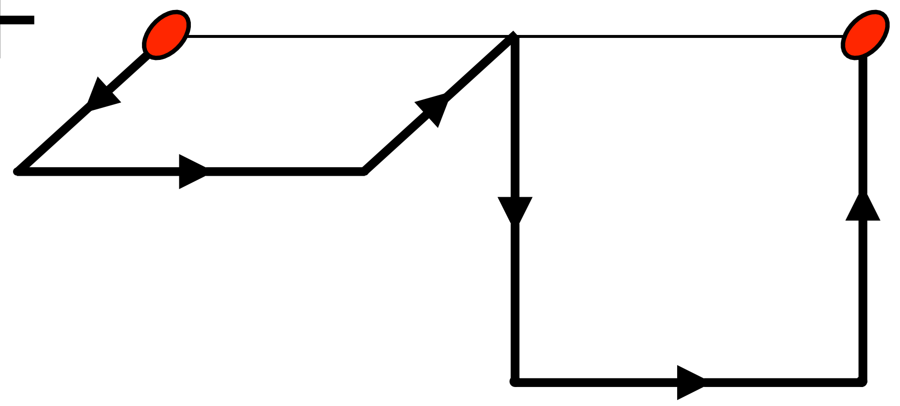
$+ s_{hd} \ \cdots \equiv$



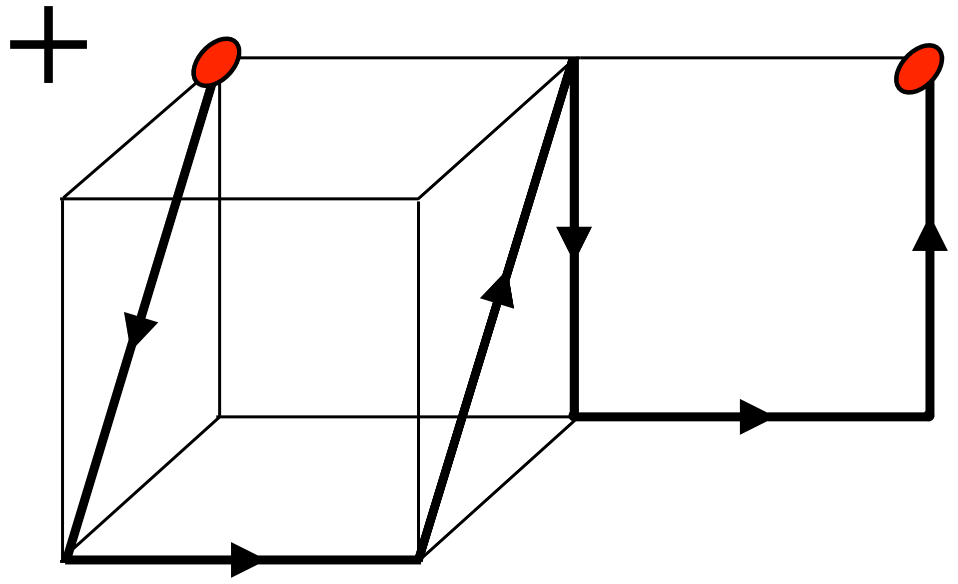
$Q_\mu(x_B) =$



$= \ \ldots +$



$+$



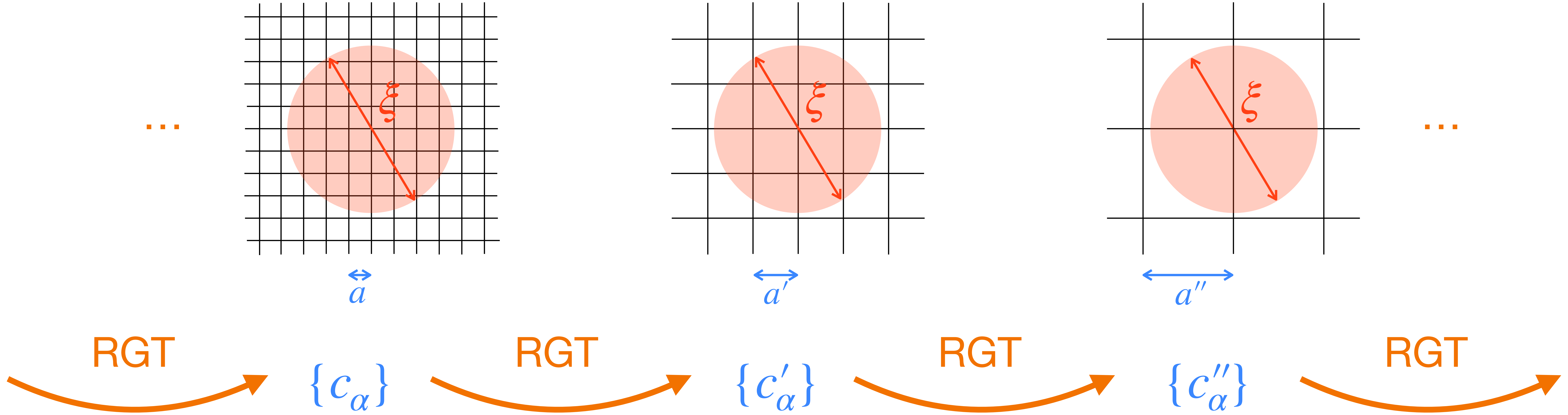
$+ \ldots$

Renormalization group transformation

Introduce (coordinate space) renormalization group transformation (RGT):

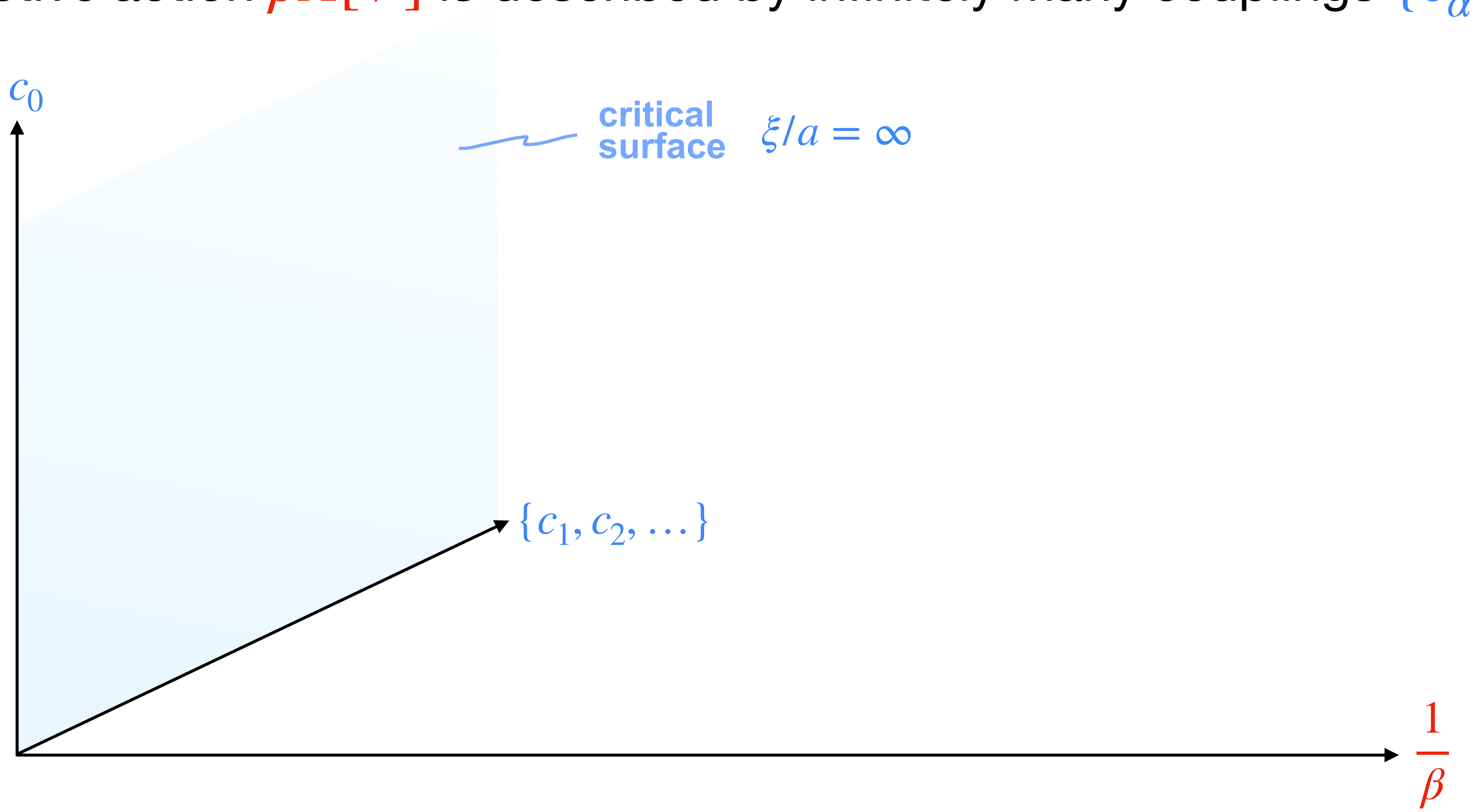
$$\exp\left\{-\beta' A'[V]\right\} = \int \mathscr{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

The effective action $\beta' A'[V]$ is described by infinitely many couplings $\{c'_\alpha\}$:



Renormalization group transformation

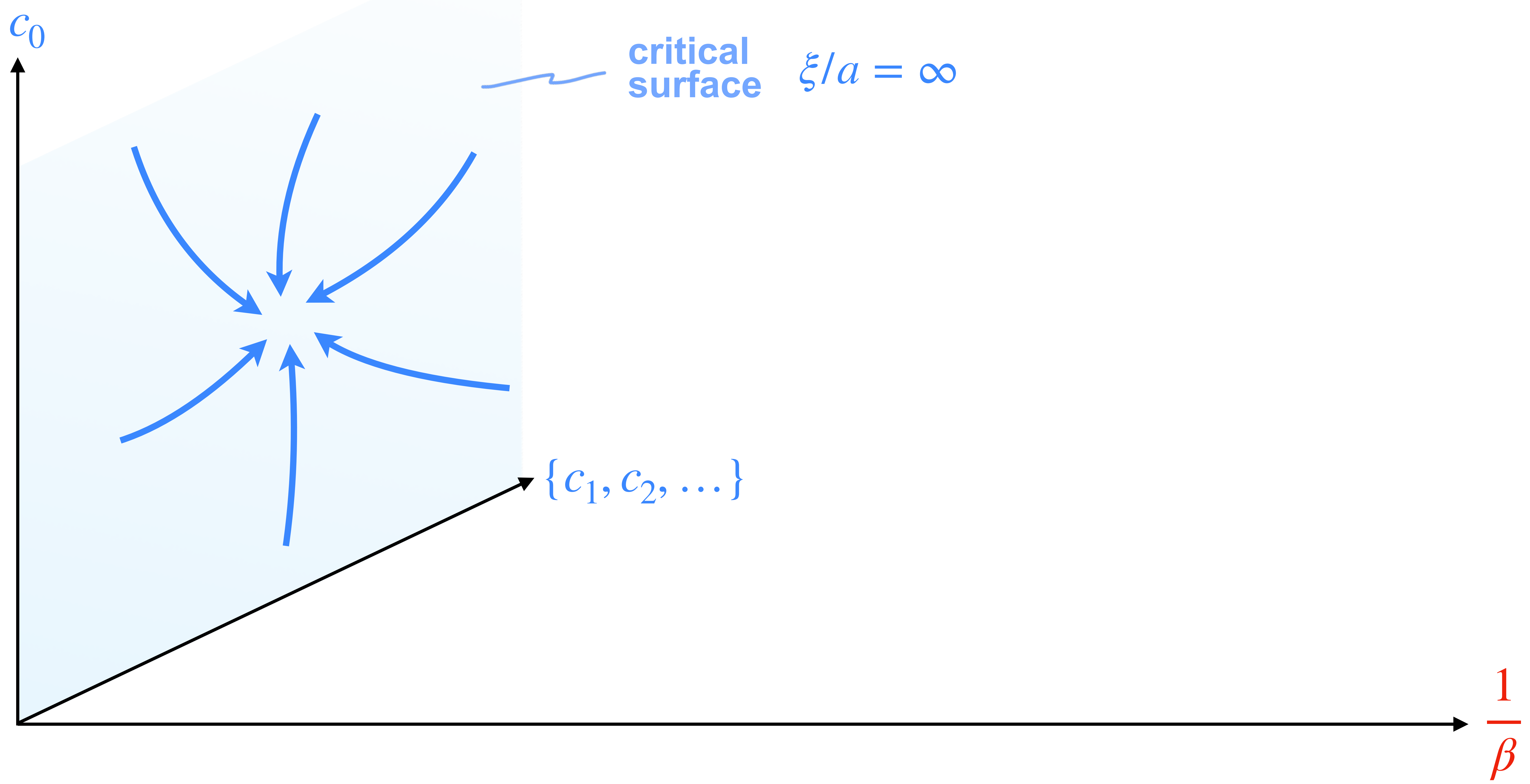
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ for asymptotically free gauge theories, there is one relevant direction $1/\beta$

Renormalization group transformation

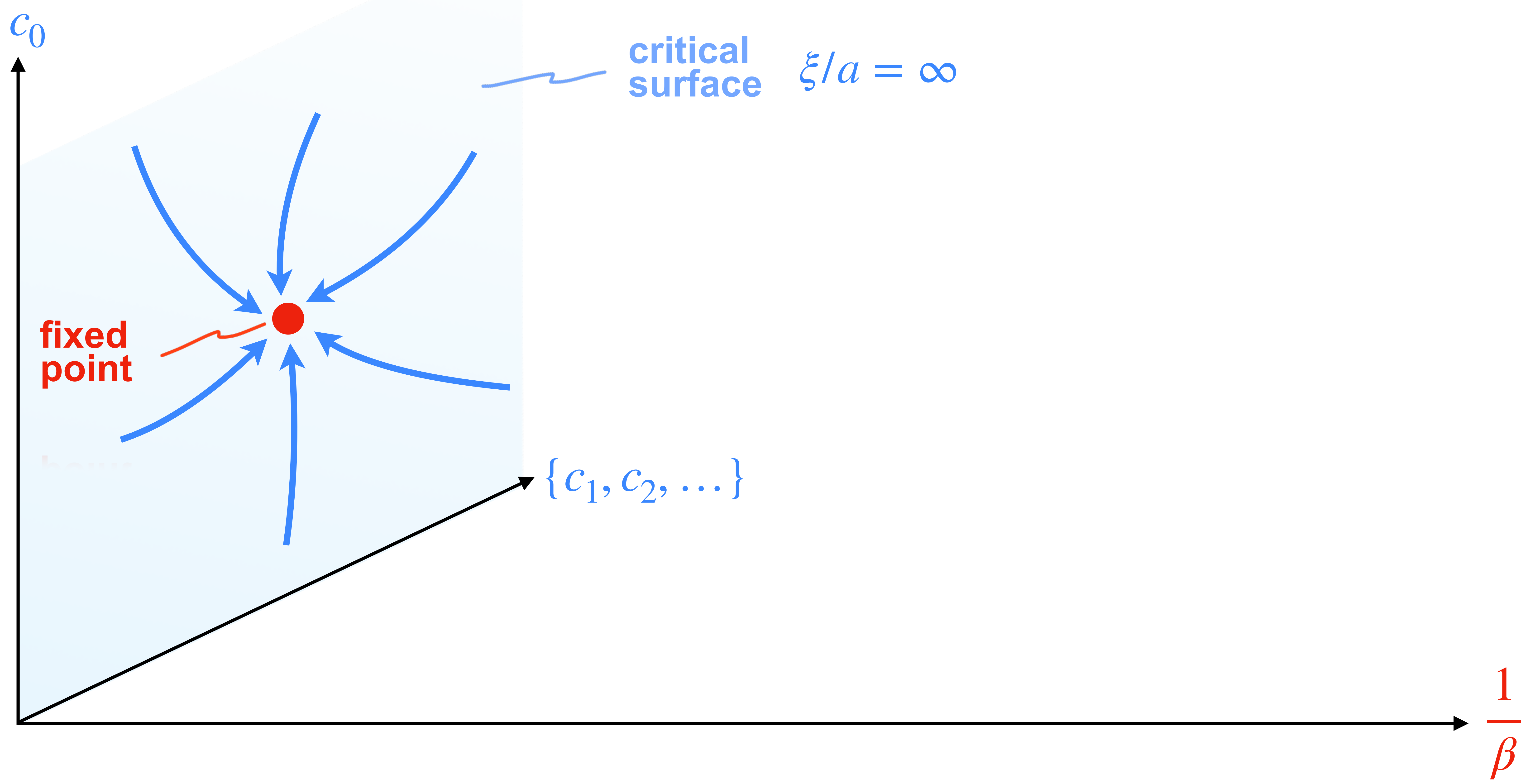
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ for asymptotically free gauge theories, there is one relevant direction $1/\beta$

Renormalization group transformation

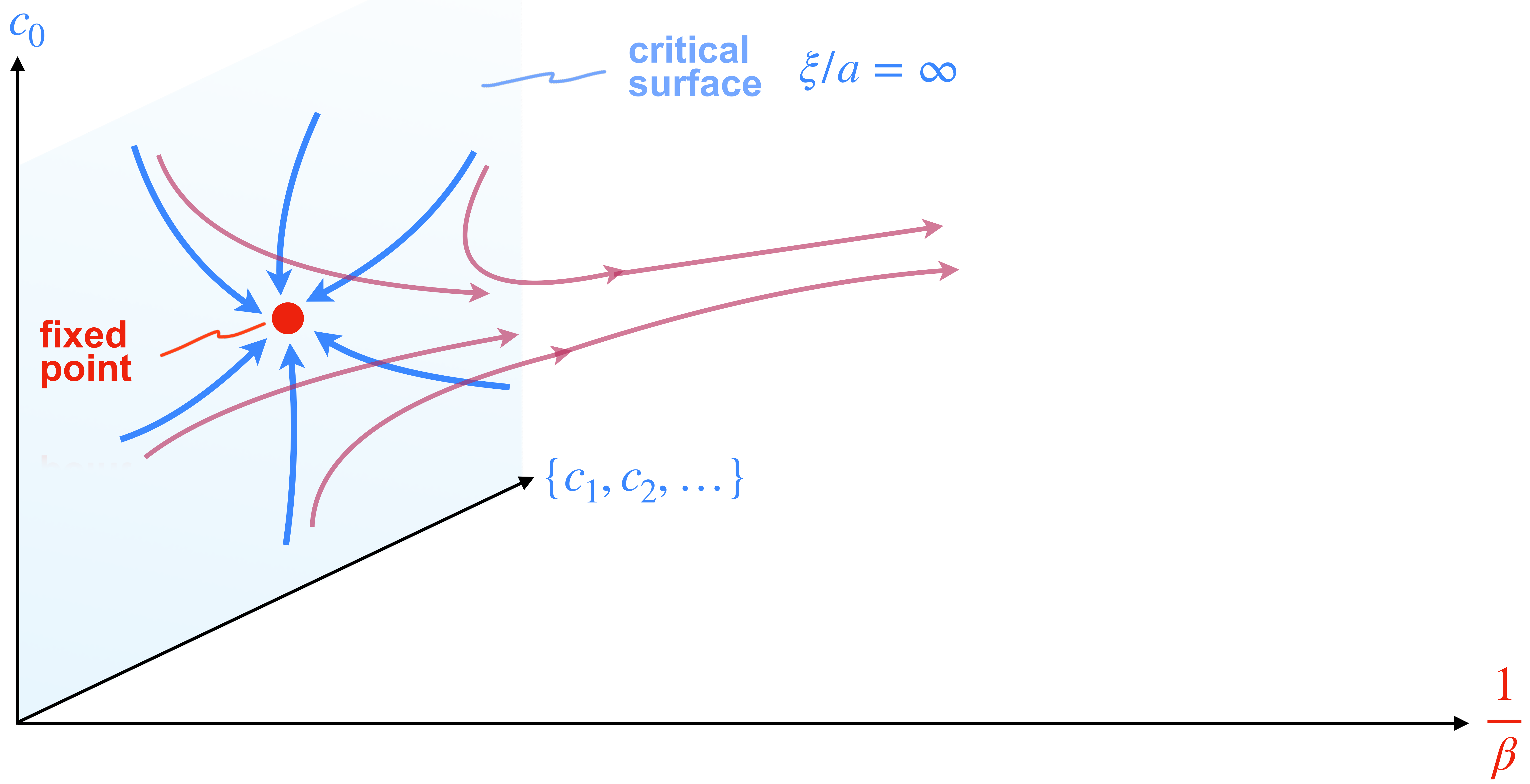
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ fixed point of RGT iterations (when $\xi/a \to \infty$): $\{c_\alpha^{FP}\} \xrightarrow{\text{RGT}} \{c_\alpha^{FP}\}$

Renormalization group transformation

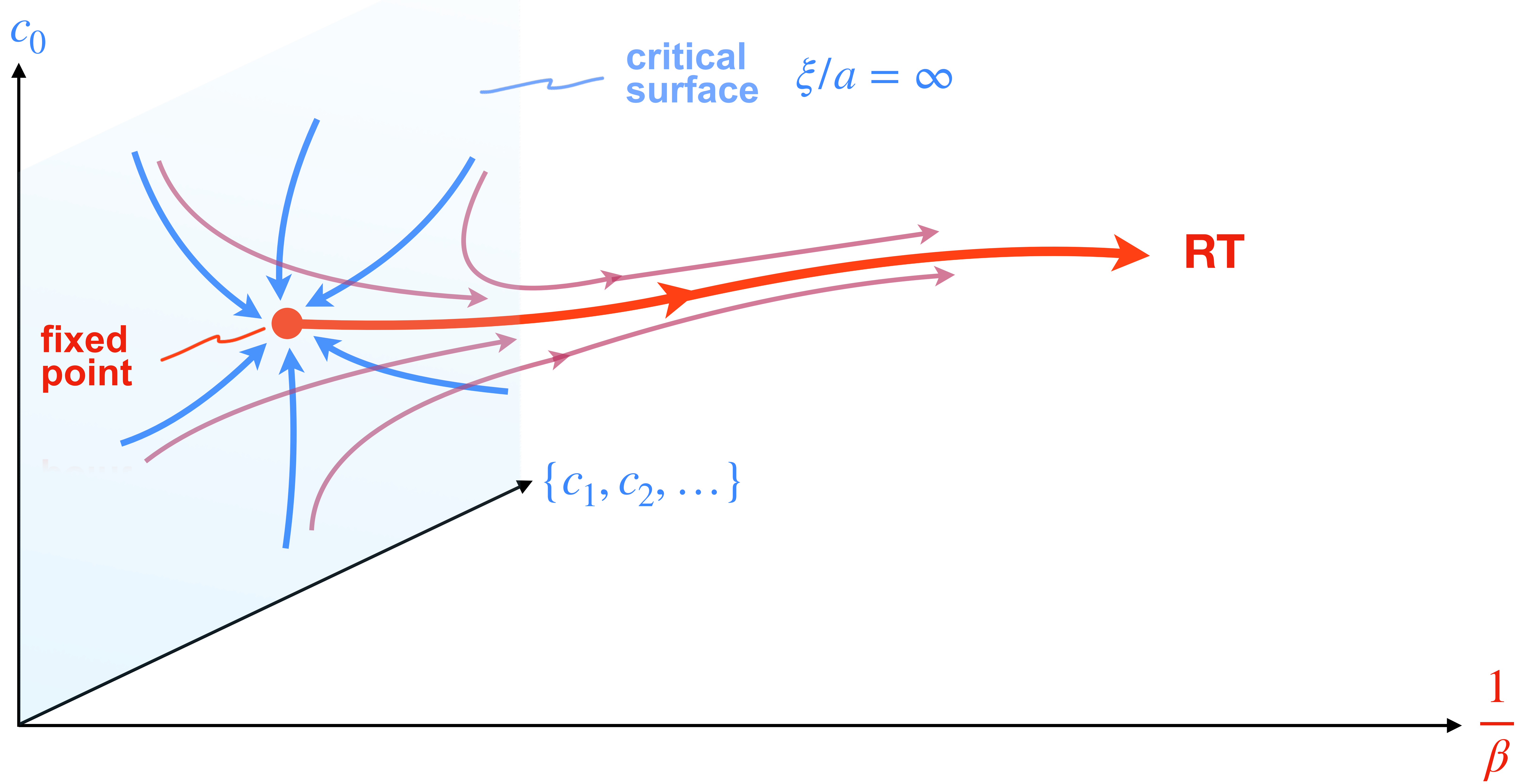
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ fixed point of RGT iterations (when $\xi/a \to \infty$): $\{c_\alpha^{FP}\} \xrightarrow{\text{RGT}} \{c_\alpha^{FP}\}$

Renormalization group transformation

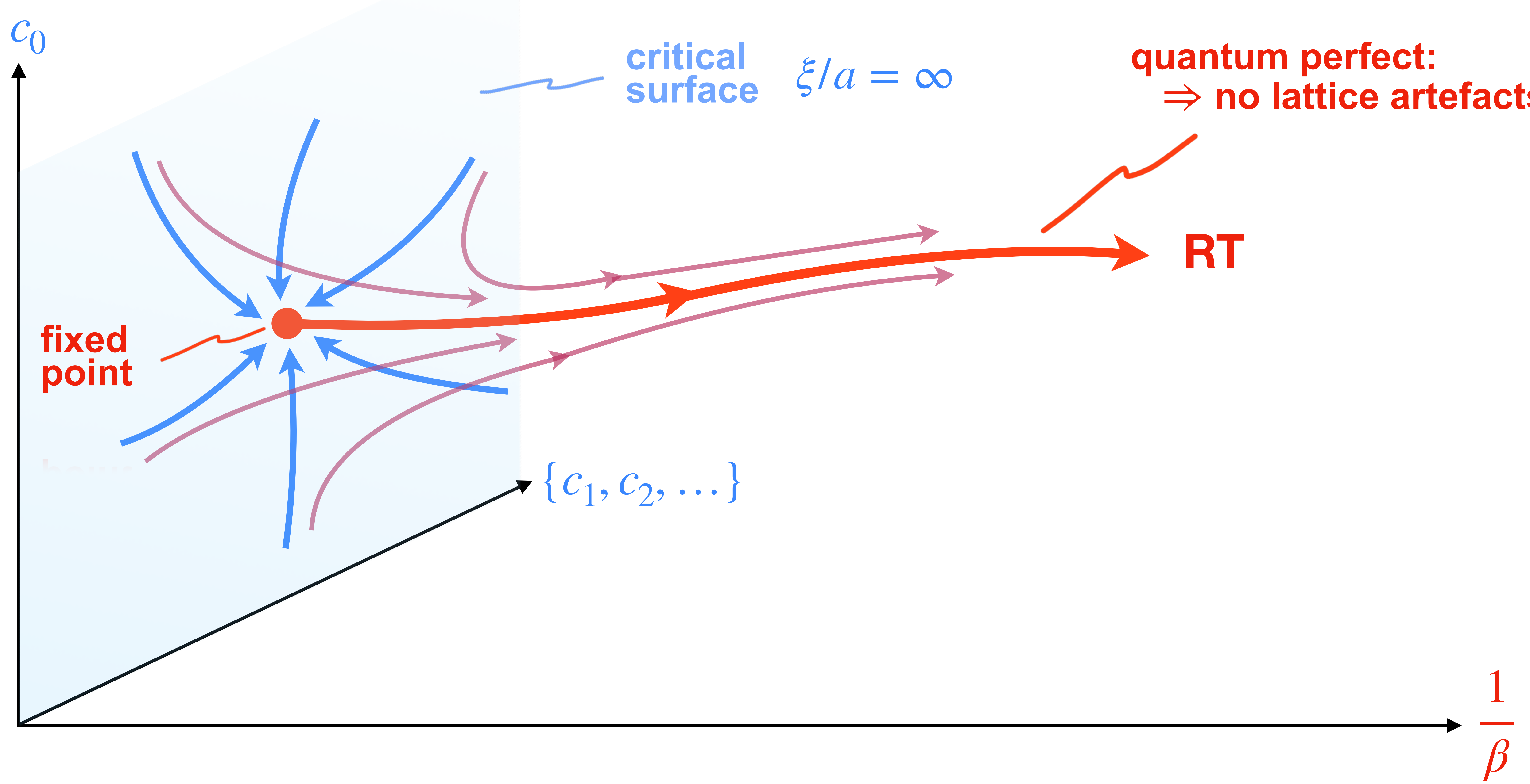
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ fixed point of RGT iterations (when $\xi/a \to \infty$): $\{c_\alpha^{FP}\} \xrightarrow{\text{RGT}} \{c_\alpha^{FP}\}$

Renormalization group transformation

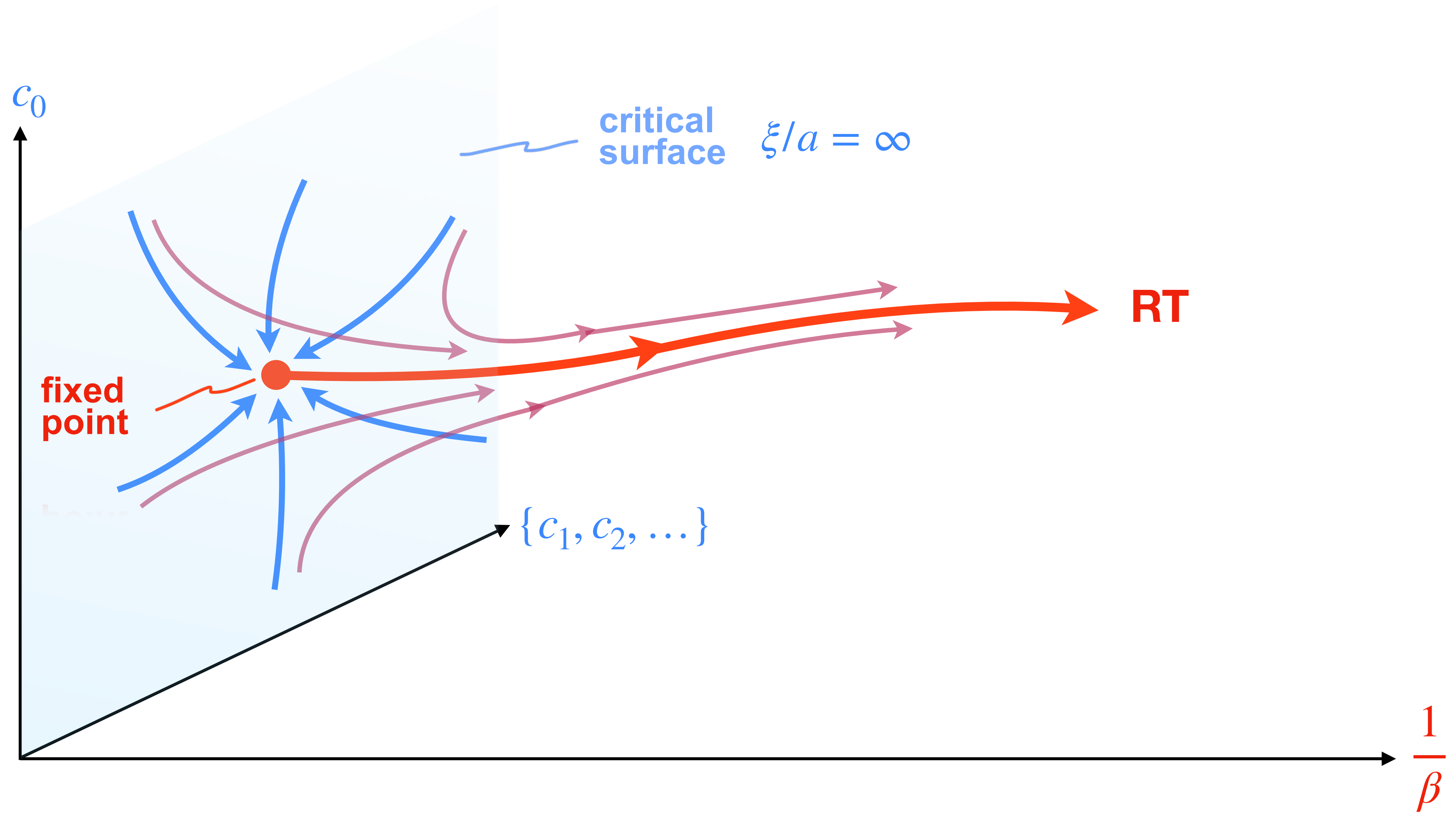
The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$\Rightarrow$ fixed point of RGT iterations (when $\xi/a \to \infty$): $\{c_\alpha^{FP}\} \xrightarrow{\text{RGT}} \{c_\alpha^{FP}\}$

Renormalization group transformation

The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



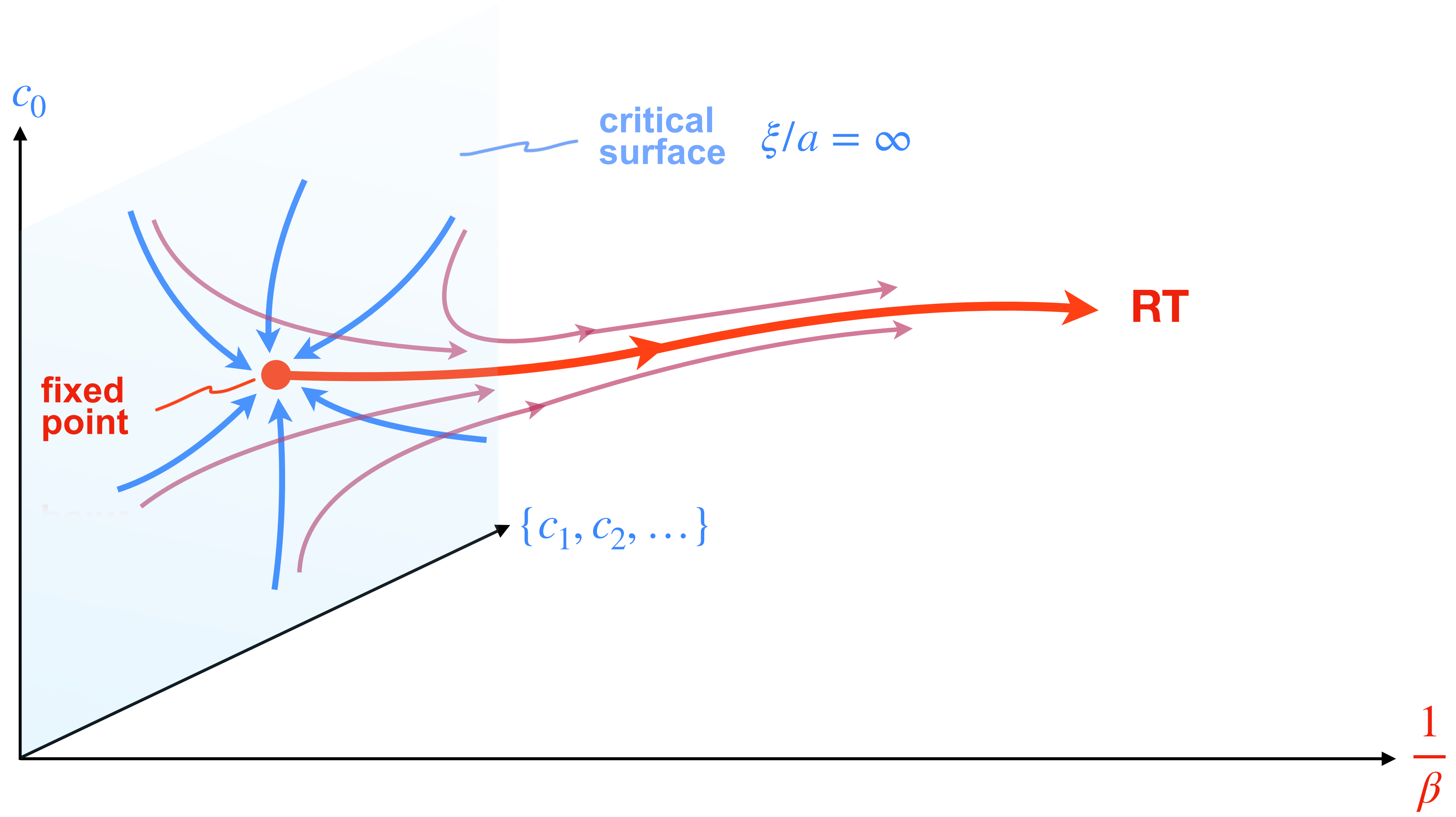
$$\exp\left\{-\beta' A'[V]\right\} = \int \mathscr{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

Two practical problems:

- how to parametrize RT, i.e., which set $\{c_\alpha\}$?
- how to determine $\{c_\alpha^{\mathrm{RT}}\}$ or $\{c_\alpha^{\mathrm{FP}}\}$?

Renormalization group transformation

The effective action $\beta A[V]$ is described by infinitely many couplings $\{c_\alpha\}$:



$$\exp\left\{-\beta' A'[V]\right\} = \int \mathscr{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

Two practical problems:

- how to parametrize RT, i.e., which set $\{c_\alpha\}$?
- how to determine $\{c_\alpha^{\mathrm{RT}}\}$ or $\{c_\alpha^{\mathrm{FP}}\}$?

P. Hasenfratz, F. Niedermayer [Nucl. Phys. B414 (1994) 785, hep-lat/9308004]

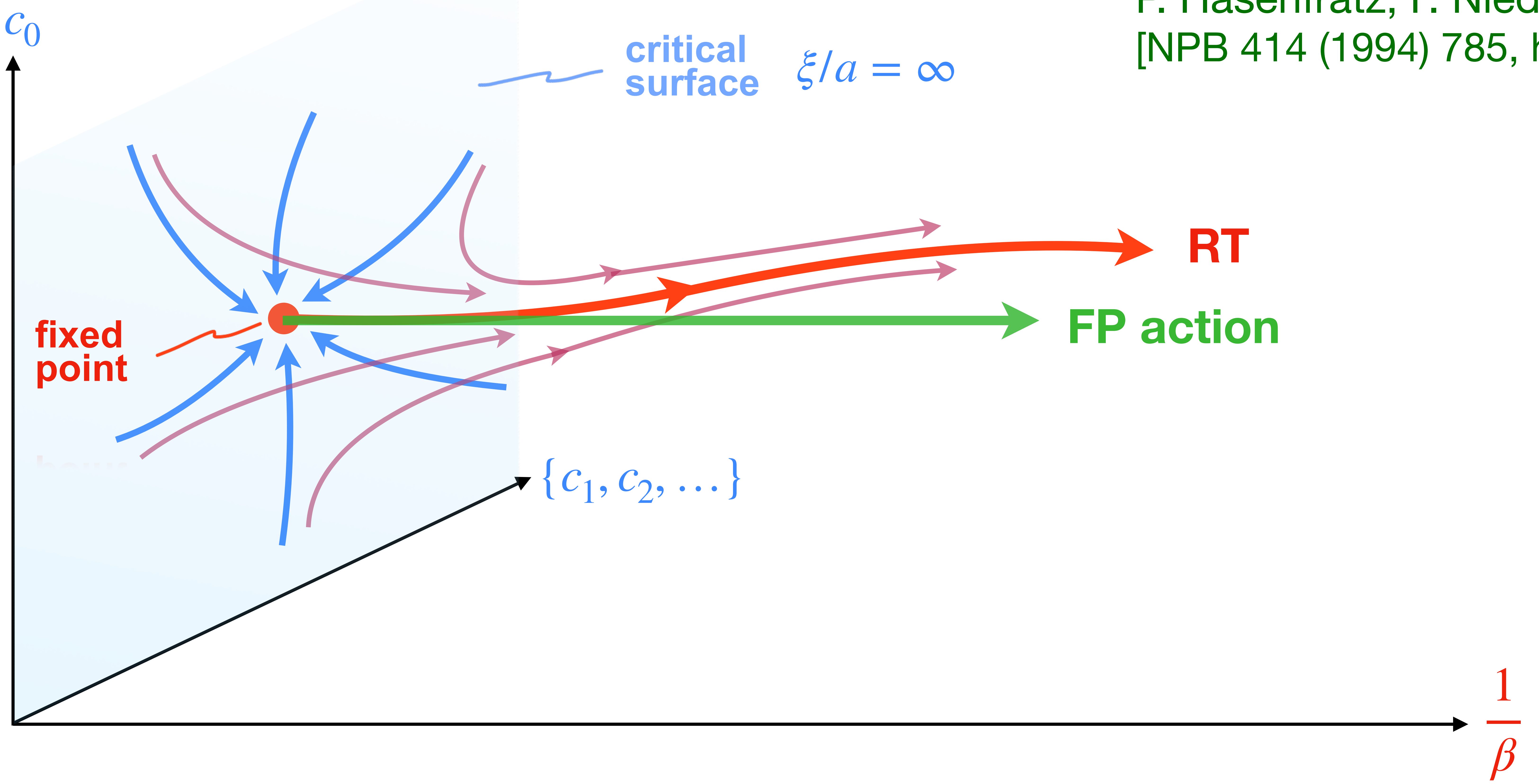
for $\beta \to \infty$ (on critical surface) the RGT becomes a classical saddle point problem:

$$A^{\mathrm{FP}}[V] = \min_{\{U\}} \left\{A^{\mathrm{FP}}[U] + T[U,V]\right\}$$

Classically perfect FP actions

The classical FP action A^{FP} defines an action for all β:

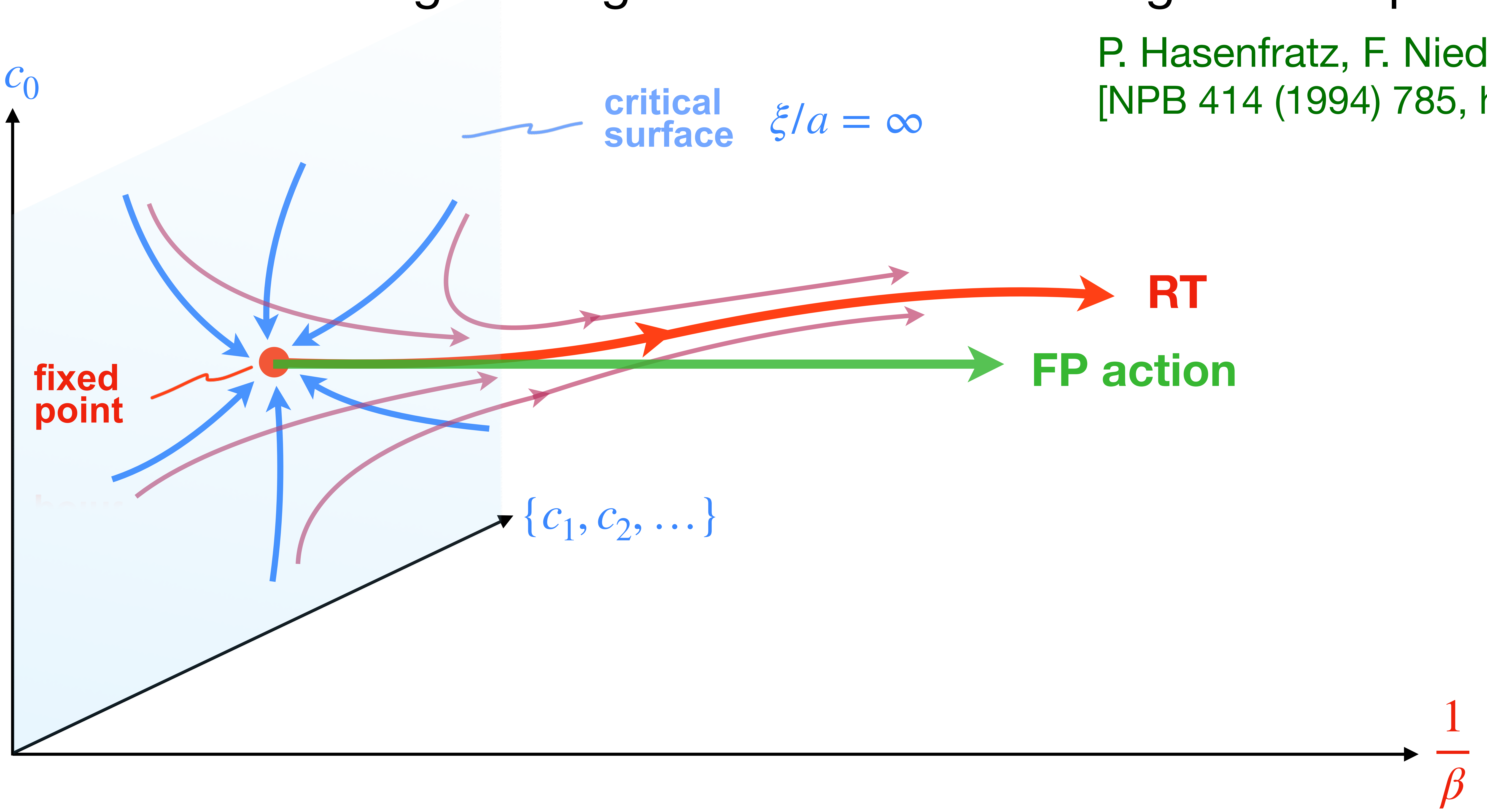
P. Hasenfratz, F. Niedermayer
[NPB 414 (1994) 785, hep-lat/9308004]



Classically perfect FP actions

The FP action values for rough configurations defined through an inception procedure:

P. Hasenfratz, F. Niedermayer
[NPB 414 (1994) 785, hep-lat/9308004]

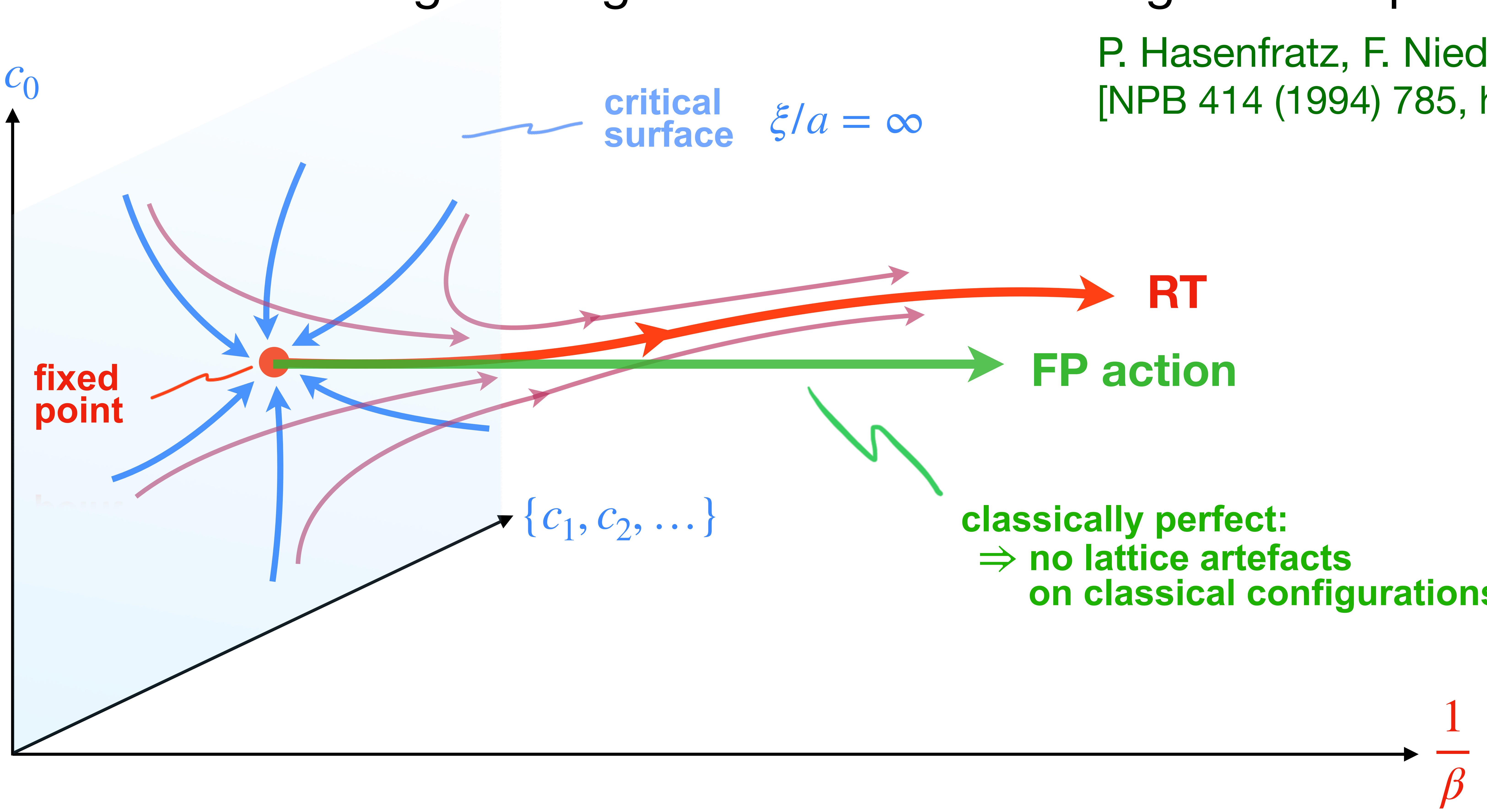


$$A^{\mathrm{FP}}[V] = \min_{\{U\}} \{A^{\mathrm{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\mathrm{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

Classically perfect FP actions

The FP action values for rough configurations defined through an inception procedure:

P. Hasenfratz, F. Niedermayer
[NPB 414 (1994) 785, hep-lat/9308004]



$$A^{\text{FP}}[V] = \min_{\{U\}} \{A^{\text{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\text{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

Classically perfect FP actions

The classical FP equation can be iterated:

$$A^{\mathsf{FP}}[V] = \min_{\{U\}} \{A^{\mathsf{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\mathsf{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

- There are no lattice artefacts on classical configurations:

$$\frac{\delta A^{\mathsf{FP}}[V]}{\delta V} = 0 \quad \Rightarrow \quad \frac{\delta A^{\mathsf{FP}}[U]}{\delta U} = 0$$

Classically perfect FP actions

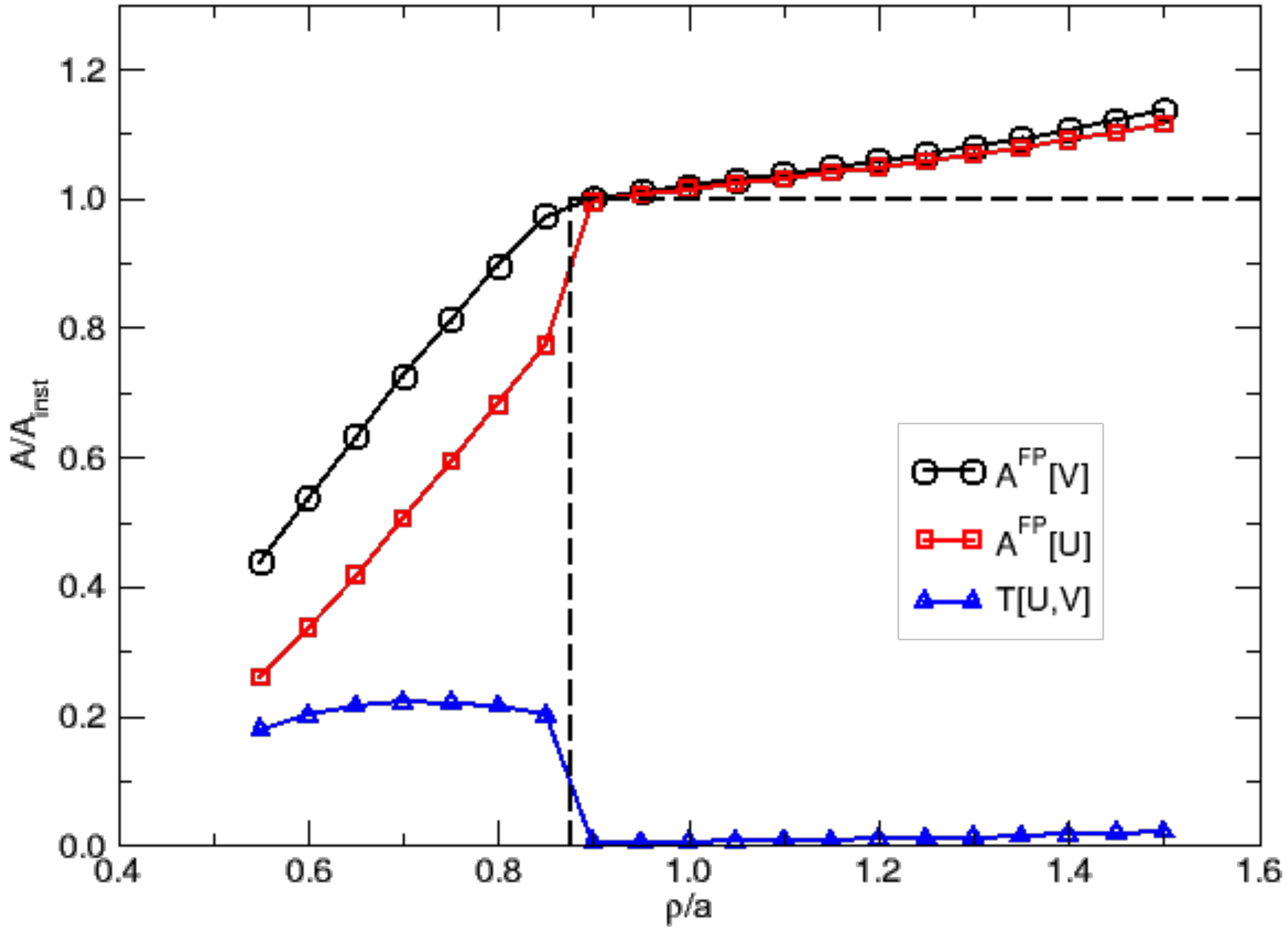
The classical FP equation can be iterated:

$$A^{\mathrm{FP}}[V] = \min_{\{U\}} \{A^{\mathrm{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\mathrm{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

- There are no lattice artefacts on classical configurations:

$$\frac{\delta A^{\mathrm{FP}}[V]}{\delta V} = 0 \quad \Rightarrow \quad \frac{\delta A^{\mathrm{FP}}[U]}{\delta U} = 0$$

$\Rightarrow A^{\mathrm{FP}}$ has scale invariant instanton solutions



Classically perfect FP actions

The classical FP equation can be iterated:

$$A^{\mathrm{FP}}[V] = \min_{\{U\}} \{A^{\mathrm{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\mathrm{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

- There are no lattice artefacts on classical configurations:

$$\frac{\delta A^{\mathrm{FP}}[V]}{\delta V} = 0 \quad \Rightarrow \quad \frac{\delta A^{\mathrm{FP}}[U]}{\delta U} = 0$$

$\Rightarrow$ $A^{\mathrm{FP}}[V]$ has scale invariant instanton solutions

- For large $\beta, A^{\mathrm{FP}}[V]$ is very close to $A^{\mathrm{RT}}[V]$:

$\Rightarrow$ lattice artefacts expected to be substantially reduced:

$$\cancel{\mathcal{O}(a^{2n})}, \mathcal{O}(g^2 a^{2n}) \quad n = 1,2,\ldots$$

Classically perfect FP actions

The classical FP equation can be iterated:

$$A^{\text{FP}}[V] = \min_{\{U\}} \{A^{\text{FP}}[U] + T[U,V]\} = \min_{\{U',U\}} \{A^{\text{FP}}[U'] + T[U',U] + T[U,V]\} = \ldots$$

- There are no lattice artefacts on classical configurations:

$$\frac{\delta A^{\text{FP}}[V]}{\delta V} = 0 \quad \Rightarrow \quad \frac{\delta A^{\text{FP}}[U]}{\delta U} = 0$$

$\Rightarrow A^{\text{FP}}[V]$ has scale invariant instanton solutions

- For large $\beta, A^{\text{FP}}[V]$ is very close to $A^{\text{RT}}[V]$:

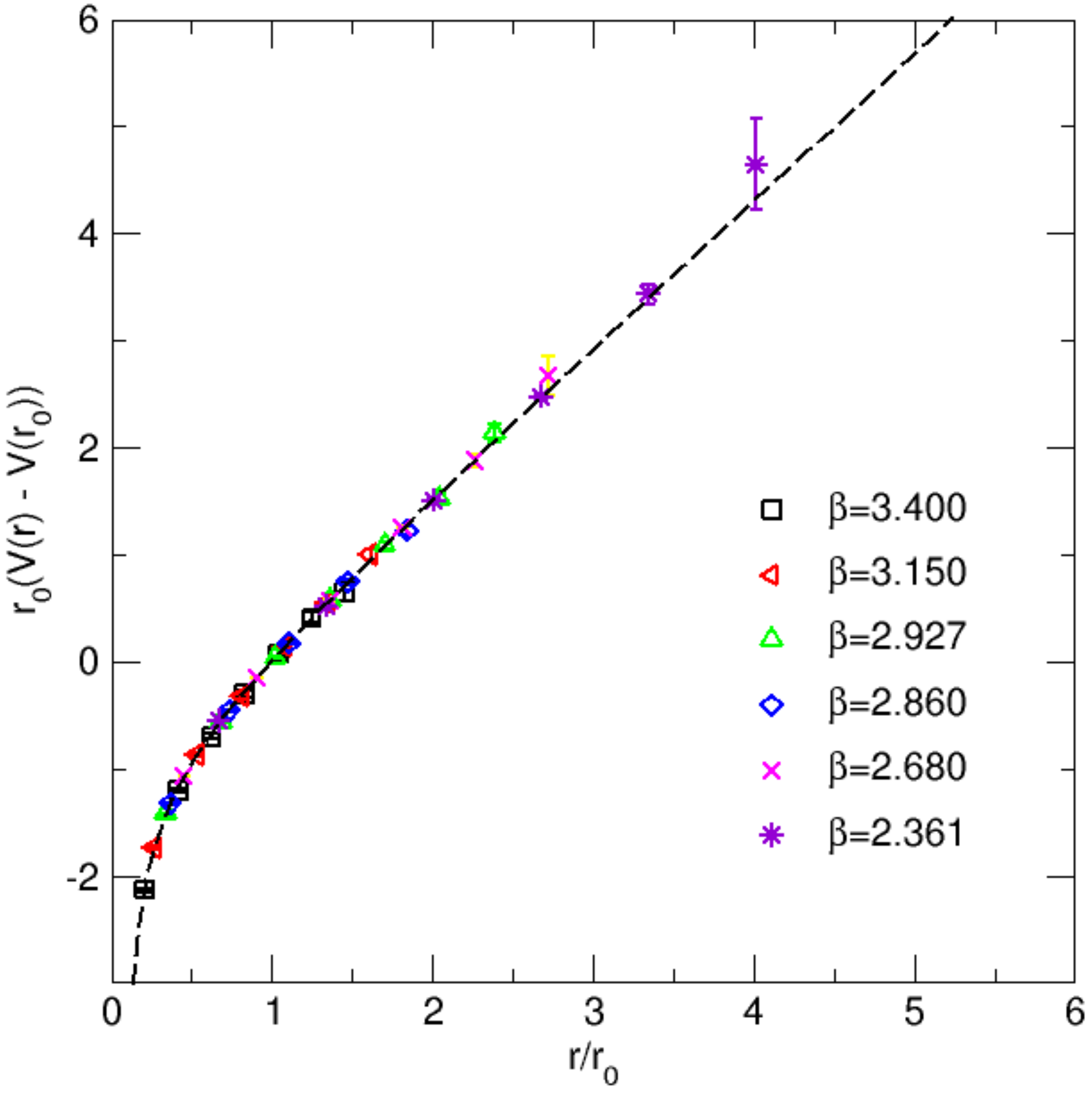
$\Rightarrow$ lattice artefacts expected to be substantially reduced:

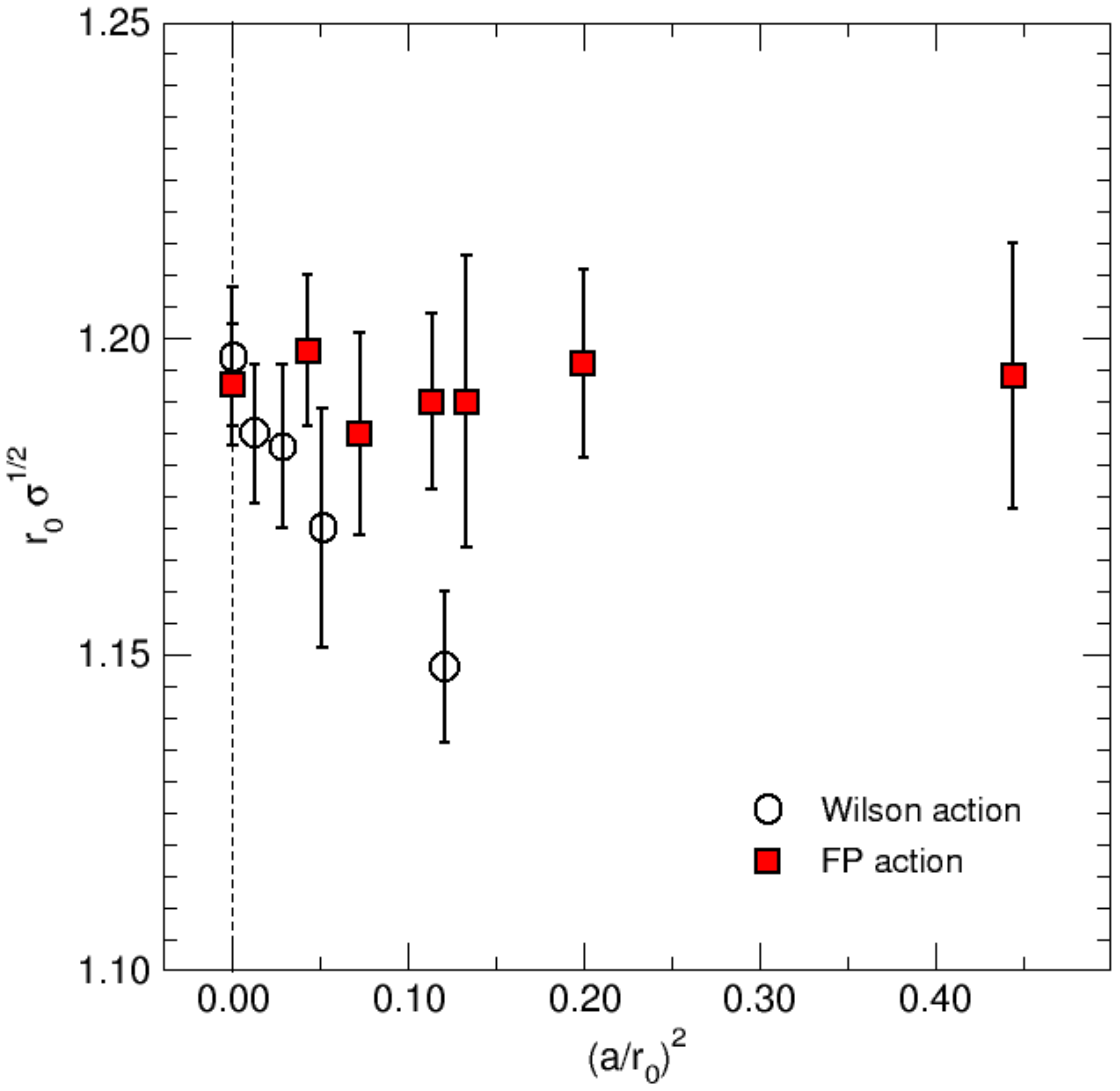
$$\cancel{\mathcal{O}(a^{2n})}, \mathcal{O}(g^2 a^{2n}) \quad n = 1,2,\ldots$$

$\Rightarrow$ initiated a large activity, culminating in the discovery of **GW fermions**!

FP action in action

Static quark-antiquark potential, lattice spacings between $a = 0.33$ fm, $\cdots$, 0.10 fm :



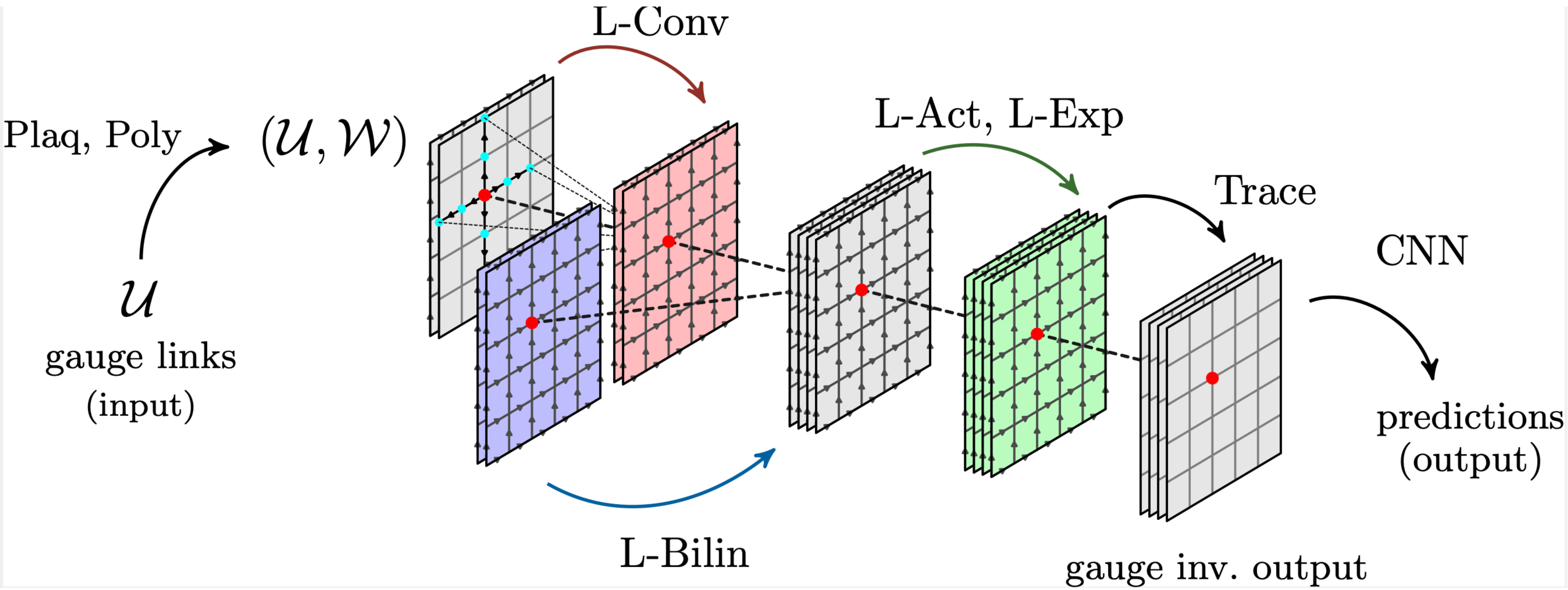


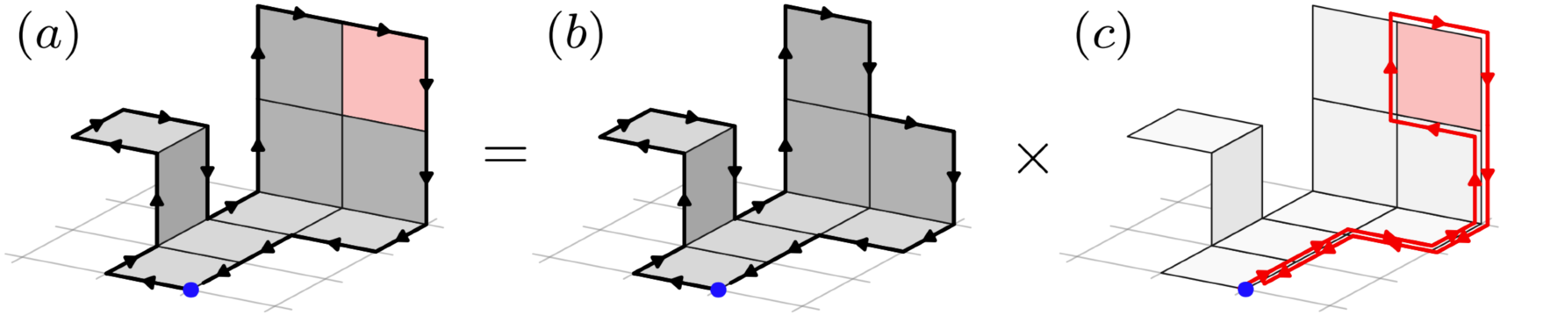
[Niedermayer, Rüfenacht, UW, Nucl.Phys.B 597 (2001) 413, hep-lat/0007007]

Machine learning the FP action

ML architecture: Lattice gauge equivariant Convolutional Neural Network (L-CNN)

[Favoni, Ipp, Müller, Schuh, PRL 128 (2022) 3, 2012.12901]

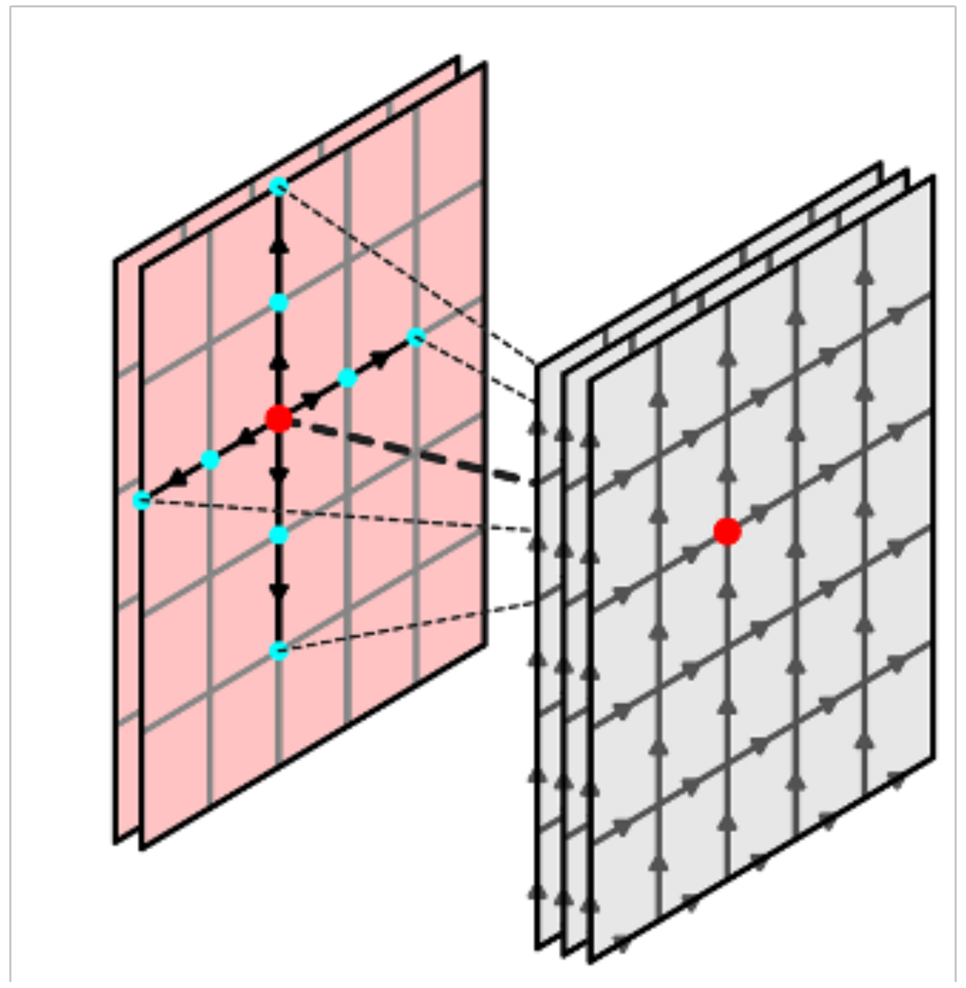




Machine learning the FP action

ML architecture: Lattice gauge equivariant Convolutional Neural Network (L-CNN)

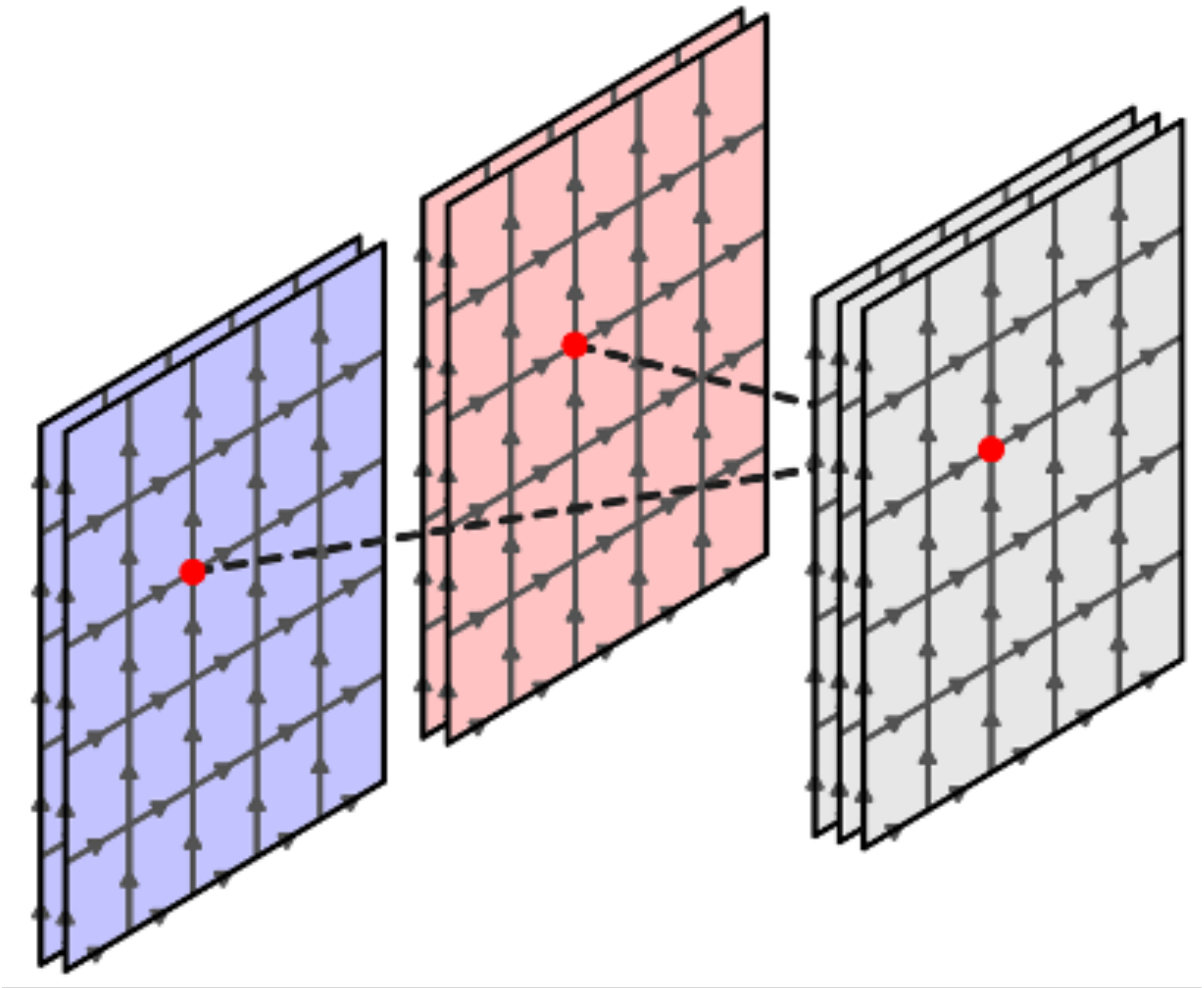
L-Conv:



$$(\mathcal{U},\mathcal{W}) \to (\mathcal{U},\mathcal{W}')$$

$$W'_{x+k\cdot\mu,j} = U_{x,k\cdot\mu} W_{x+k\cdot\mu,j} U^{\dagger}_{x,k\cdot\mu},$$

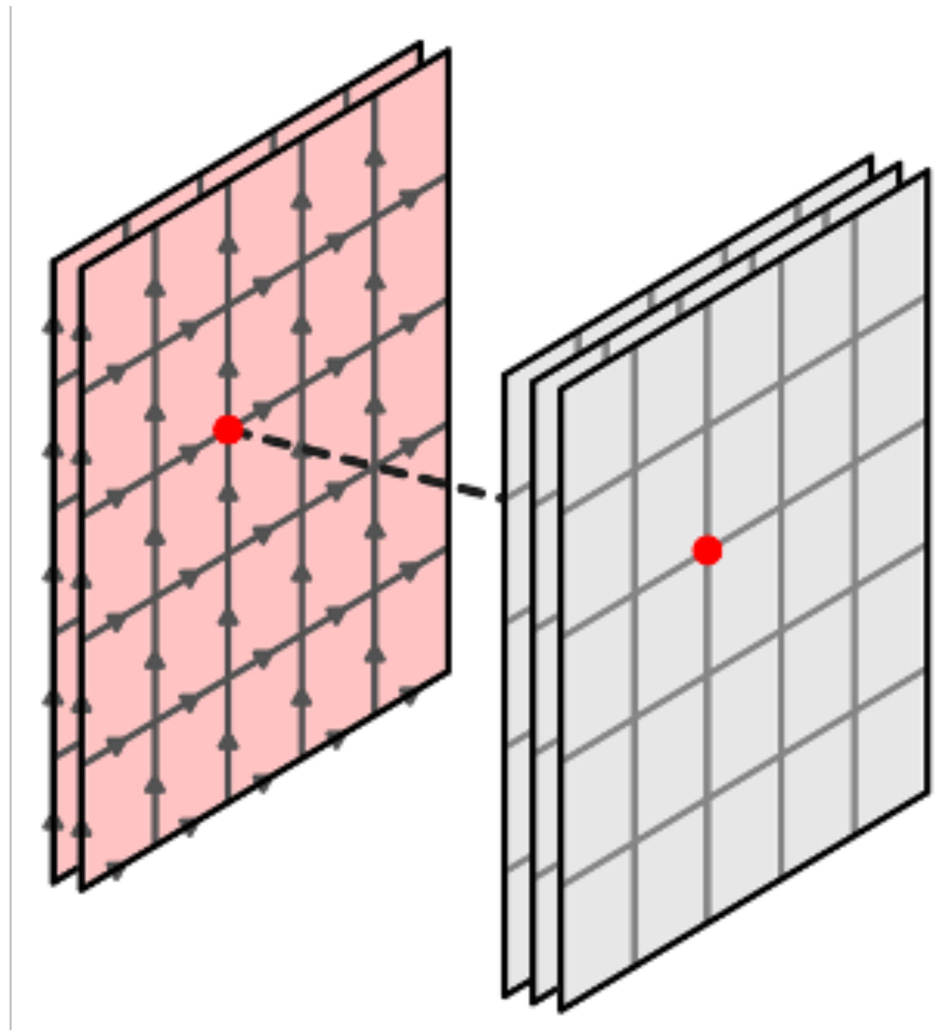
L-Bilin:



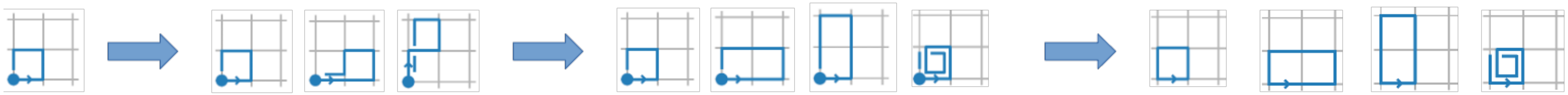
$$(\mathcal{U},\mathcal{W}) \times (\mathcal{U},\mathcal{W}') \to (\mathcal{U},\mathcal{W}'')$$

$$W_{x,i} \to \sum_{j,j',k} \alpha_{i,j,j',k} W_{x,j} W'_{x+k\cdot\mu,j'}.$$

Trace:



$$w_{\mathbf{x},i} = \mathrm{Tr}\, W_{\mathbf{x},i} \in \mathbb{C}$$



Machine learning the FP action: FP data

Use the exact FP action values for training, plus the derivatives of the FP action:

$$\frac{\delta A^{\mathrm{FP}}[V]}{\delta V^a_{x,\mu}} = \frac{\delta T[U,V]}{\delta V^a_{x,\mu}} = -\kappa\,\mathrm{Re\ Tr}(it^a\, V_{x,\mu} Q^\dagger_{x,\mu}) \qquad Q^\dagger_{x,\mu} = Q^\dagger_{x,\mu}[U]$$

⇒ yields 4 x 8 x Volume (link) (color) (position) data per configuration

Gauge invariance of A^{FP} yields conserved local quantity via Noether's theorem:

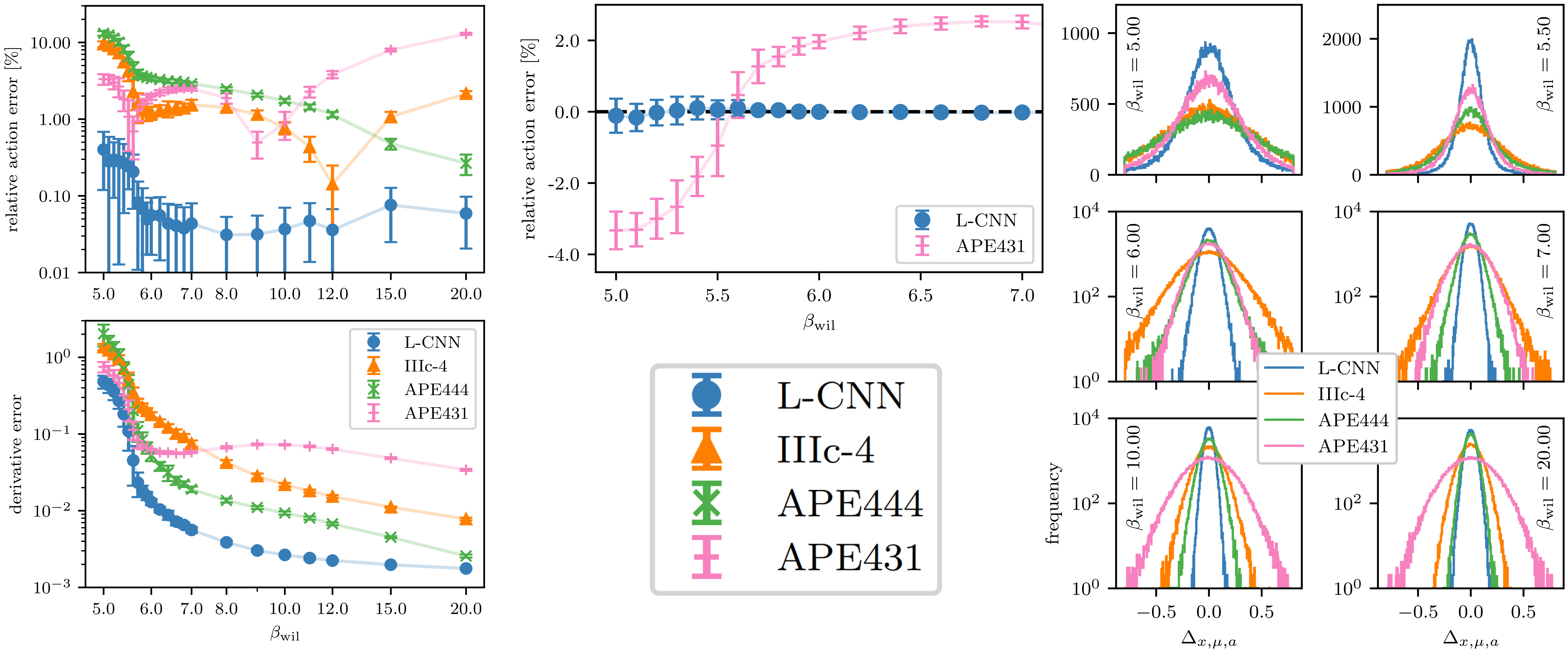
$$D^{FP}_{x,\mu} = \sum_a t^a \frac{\delta A^{FP}[V]}{\delta V^a_{x,\mu}} \implies \sum_\mu \mathscr{D}^B_\mu D^{FP}_{x,\mu}[V] = 0$$

⇒ consistency check satisfied up to the accuracy in minimization

$$\left.\begin{array}{l}\bullet\ \text{FP action values}\\ \bullet\ \text{FP action derivatives}\end{array}\right\} \Rightarrow \quad \text{data set for supervised ML}$$

Machine learning the FP action: Results

Superiority of L-CNN over old parameterizations of FP action:



Scaling properties of FP actions

Use renormalized GF coupling as scaling quantity:

$$\frac{dA_\mu(t)}{dt} = \frac{\delta S_{YM}}{\delta A_\mu} \qquad \langle t^2 E(t) \rangle = \frac{3(N^2-1)g^2}{128\pi^2}\left(1 + O(g^2)\right) \equiv \frac{3g_{GF}^2(t)}{16\pi^2}$$

where g is the renormalised $\overline{\text{MS}}$ coupling at RG scale $\mu = 1/\sqrt{8t}$, with the corresponding β-function:

$$\mu^2 \frac{dg_{GF}^2}{d(\mu^2)} = -\,t\frac{dg_{GF}^2}{dt}$$

Scaling properties of FP actions

Use renormalized GF coupling as scaling quantity:

$$\frac{dA_\mu(t)}{dt} = \frac{\delta S_{YM}}{\delta A_\mu} \qquad \langle t^2 E(t) \rangle = \frac{3(N^2-1)g^2}{128\pi^2}\left(1 + O(g^2)\right) \equiv \frac{3g_{GF}^2(t)}{16\pi^2}$$

where g is the renormalised $\overline{\text{MS}}$ coupling at RG scale $\mu = 1/\sqrt{8t}$, with the corresponding β-function:

$$\mu^2 \frac{dg_{GF}^2}{d(\mu^2)} = -t\frac{dg_{GF}^2}{dt}$$

DANGER: observable itself may introduce lattice artifacts…

Scaling properties of FP actions

Use renormalized GF coupling as scaling quantity:

$$\frac{dA_\mu(t)}{dt} = \frac{\delta S_{YM}}{\delta A_\mu} \qquad \langle t^2 E(t) \rangle = \frac{3(N^2-1)g^2}{128\pi^2}\left(1 + O(g^2)\right) \equiv \frac{3g_{GF}^2(t)}{16\pi^2}$$

where g is the renormalised $\overline{\text{MS}}$ coupling at RG scale $\mu = 1/\sqrt{8t}$, with the corresponding β-function:

$$\mu^2 \frac{dg_{GF}^2}{d(\mu^2)} = -\,t\frac{dg_{GF}^2}{dt}$$

DANGER: observable itself may introduce lattice artifacts…

⇒ turns out that GF with FP actions is classically perfect!

Gradient flow on the lattice

On the lattice, the flow of the gauge links is with a lattice flow action S^f: $\quad \frac{dU_\mu}{dt} = -i\frac{\delta S^f}{\delta U_\mu}U_\mu$

In addition, separate choice of lattice action S^e for E and the simulated gauge action S^g contributing to lattice artifacts:

$$t^2\langle E(t)\rangle = \frac{3(N^2-1)g_0^2}{128\pi^2}\left[C(a^2/t) + \mathcal{O}(g_0^2)\right]$$

$$C(a^2/t) = \frac{64\pi^2 t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\mathrm{Tr}\left[e^{-t(S^f+\mathcal{G})}(S^g+\mathcal{G})^{-1}e^{-t(S^f+\mathcal{G})}S^e\right]$$

where $C(a^2/t) = 1 + \mathcal{O}(a^2/t)$ contains the tree-level lattice artifacts.

Gradient flow on the lattice

On the lattice, the flow of the gauge links is with a lattice flow action S^f: $\quad \dfrac{dU_\mu}{dt} = -i\dfrac{\delta S^f}{\delta U_\mu}U_\mu$

In addition, separate choice of lattice action S^e for E and the simulated gauge action S^g contribute to lattice artifacts:

$$t^2\langle E(t)\rangle = \frac{3(N^2-1)g_0^2}{128\pi^2}\left[C(a^2/t) + \mathcal{O}(g_0^2)\right]$$

$$C(a^2/t) = \frac{64\pi^2 t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\mathrm{Tr}\left[e^{-t(S^f+\mathcal{G})}(S^g+\mathcal{G})^{-1}e^{-t(S^f+\mathcal{G})}S^e\right]$$

where $C(a^2/t) = 1 + \mathcal{O}(a^2/t)$ contains the tree-level lattice artifacts.

Calculation in momentum space to quadratic order in A_μ:

$$A_\mu(p)S_{\mu\nu}(p)A_\nu(-p)$$

(with gauge fixing term $\mathcal{G}$)

Gradient flow on the lattice

Choosing the same action for all three $S^f = S^e = S^g = S$:

$$C(a^2/t) = \frac{64\pi^2 t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\mathrm{Tr}\left[e^{-t(S+\mathcal{G})}(S+\mathcal{G})^{-1}e^{-t(S+\mathcal{G})}S\right]$$

Gradient flow on the lattice

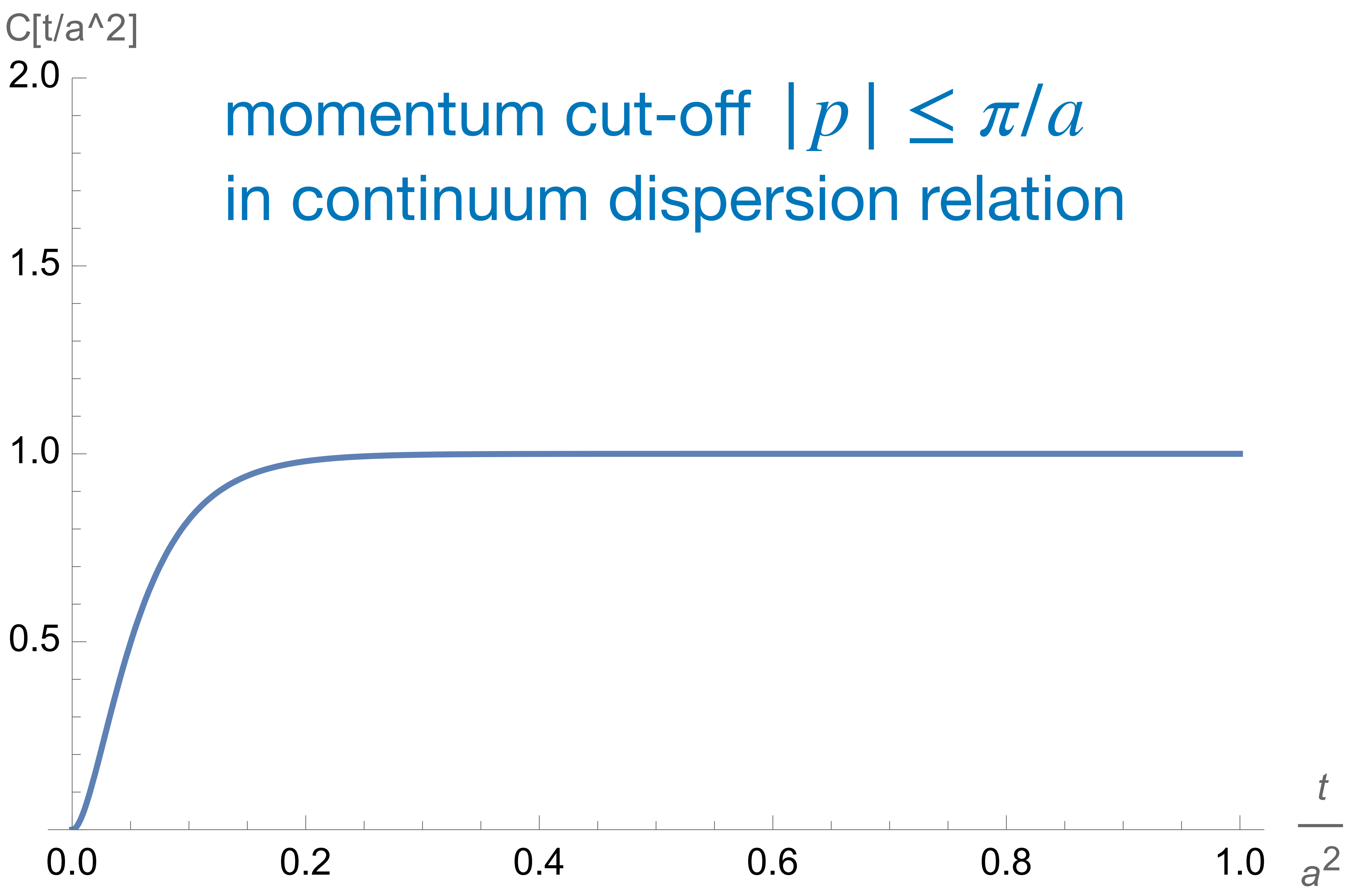
Choosing the same action for all three $S^f = S^e = S^g = S$:

$$C(a^2/t) = \frac{64\pi^2 t^2}{3} \int_{-\pi/a}^{\pi/a} \frac{d^4 p}{(2\pi)^4} \mathrm{Tr}\left[e^{-2t(S+\mathscr{G})}\right]$$

Gradient flow on the lattice

Choosing the same action for all three $S^f = S^e = S^g = S$:

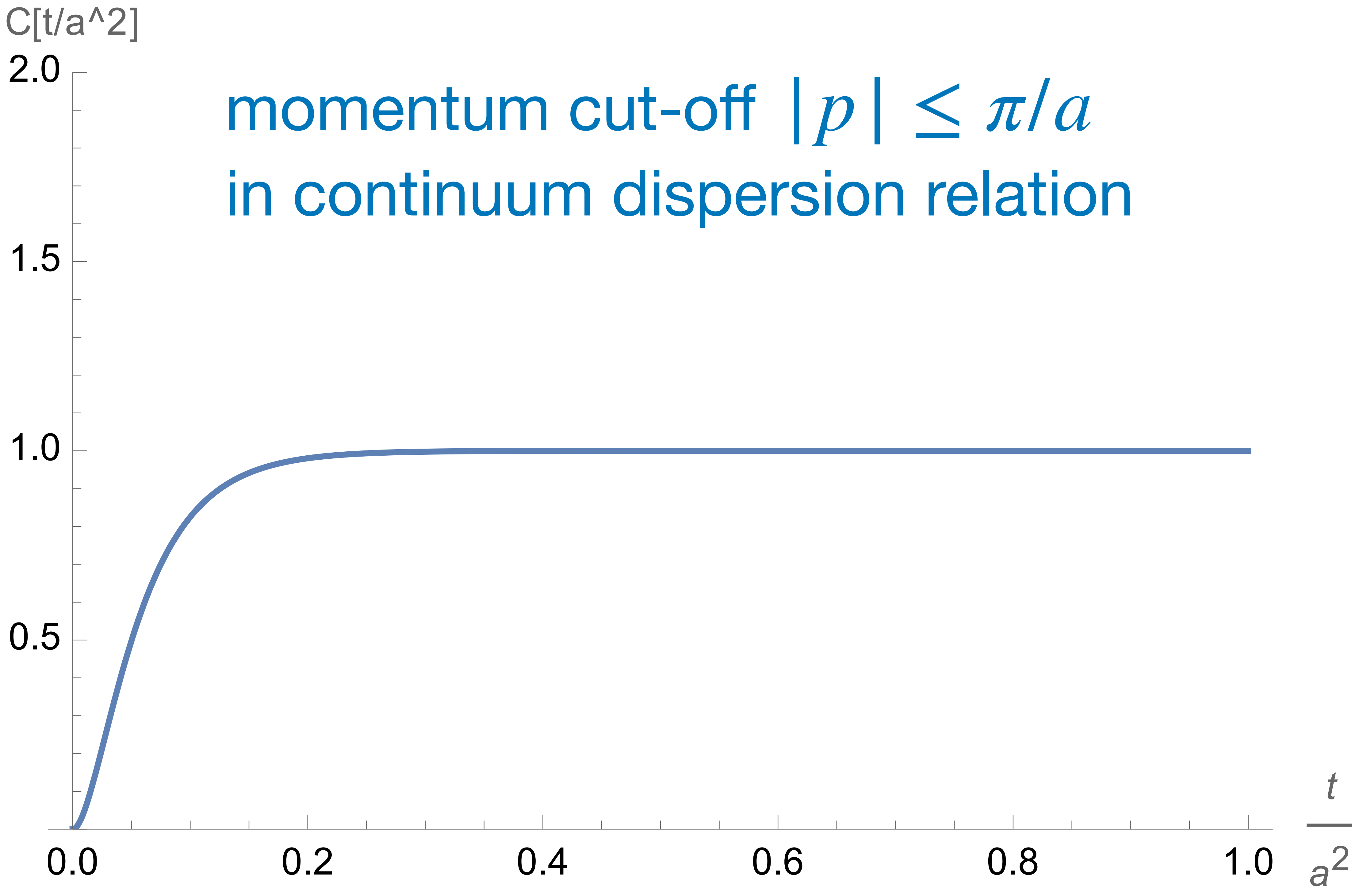
$$C(a^2/t) = \frac{64\pi^2 t^2}{3} \int_{-\pi/a}^{\pi/a} \frac{d^4 p}{(2\pi)^4} \mathrm{Tr}\left[e^{-2t(S+\mathcal{G})}\right]$$

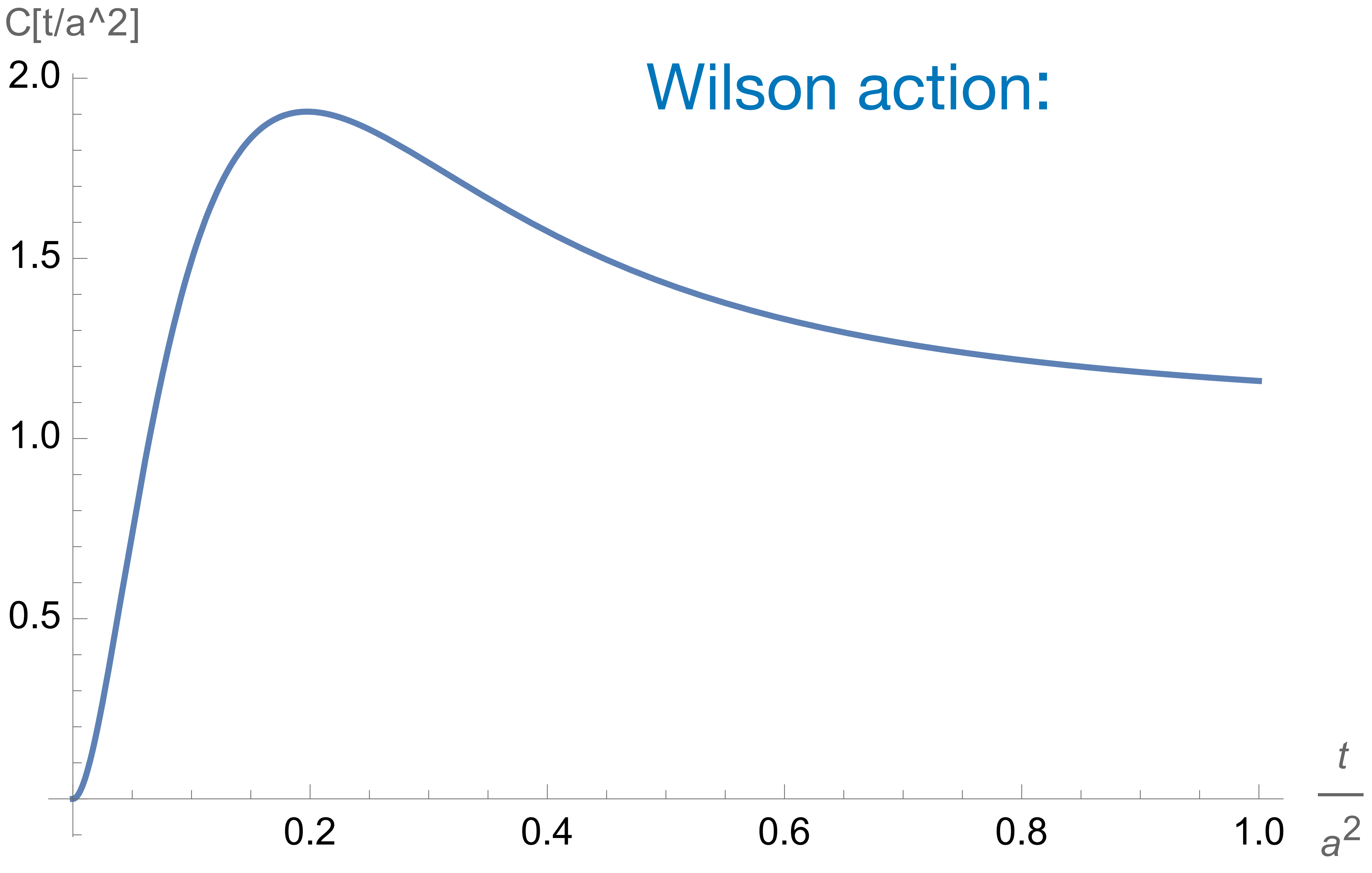


Gradient flow on the lattice

Choosing the same action for all three $S^f = S^e = S^g = S$:

$$C(a^2/t) = \frac{64\pi^2 t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\mathrm{Tr}\left[e^{-2t(S+\mathscr{G})}\right]$$





Classically perfect gradient flow:

Lattice momenta restricted as usual: $-\pi/a \leq p_\mu \leq \pi/a$

but iterated RG transformations generate additional poles in the propagator:

$$(p + 2\pi l)^2 \qquad \text{for } l = 0,1,2,\ldots$$

extending momentum range to $-\infty \leq p_\mu \leq \infty$ and yielding the continuum dispersion relation:

$$C^{FP}(a^2/t) = \frac{64\pi^2 t^2}{3} \cdot 3 \left(\int_{-\infty}^{+\infty} \frac{dp}{2\pi} e^{-2tp^2} \right)^4 = \frac{64\pi^2 t^2}{(\sqrt{8\pi t})^4} = 1$$

$\Rightarrow$ gradient flow with FP actions is classically perfect!

Scaling of gradient-flow scales

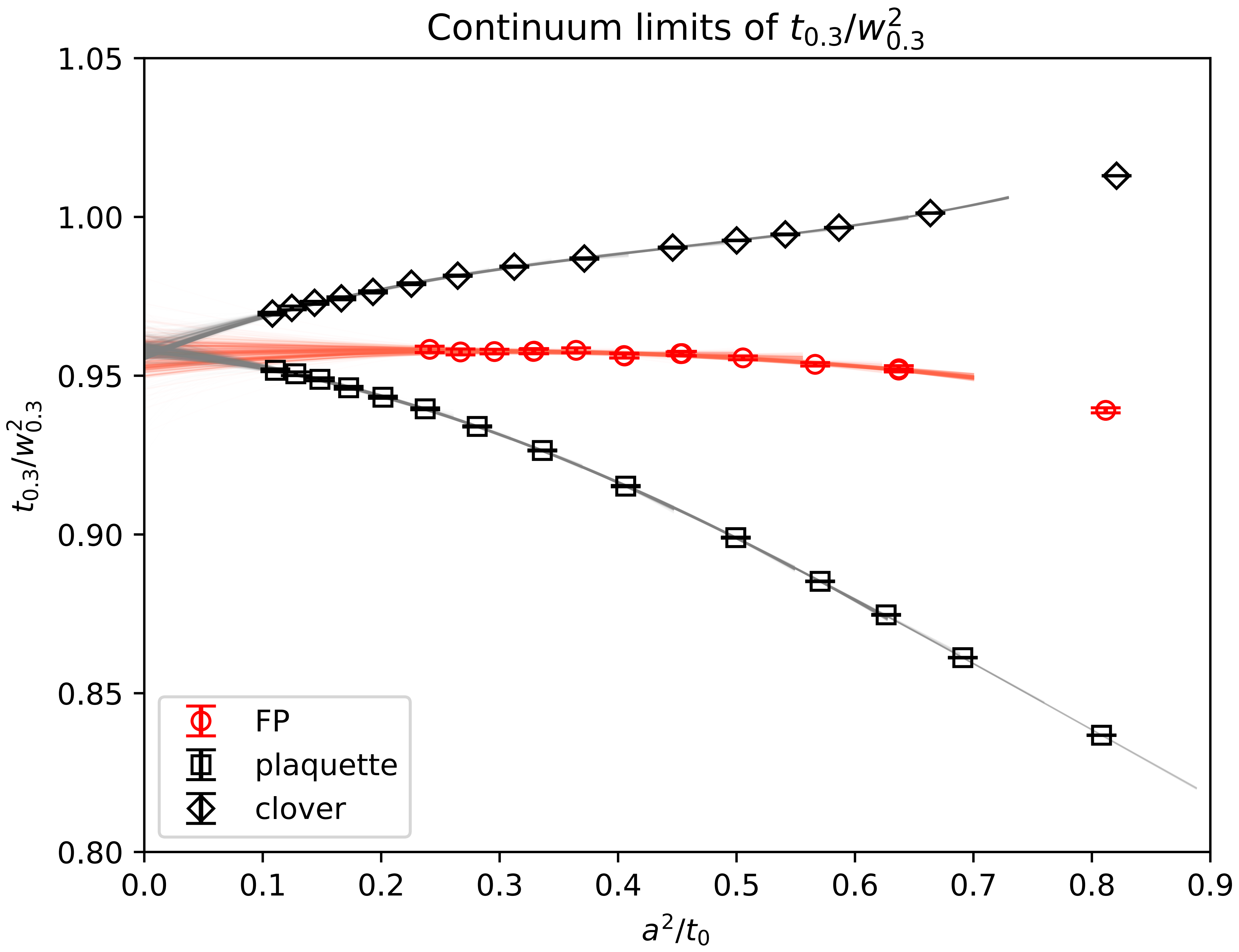
Physical reference scales defined through $t^2\langle E\rangle\,|_{t=t_x} = x,\qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

yielding dimensionless ratios t_x/w_x^2 as scaling quantities vs. $a^2/t_{0.3}$:

Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle\big|_{t=t_x} = x, \qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

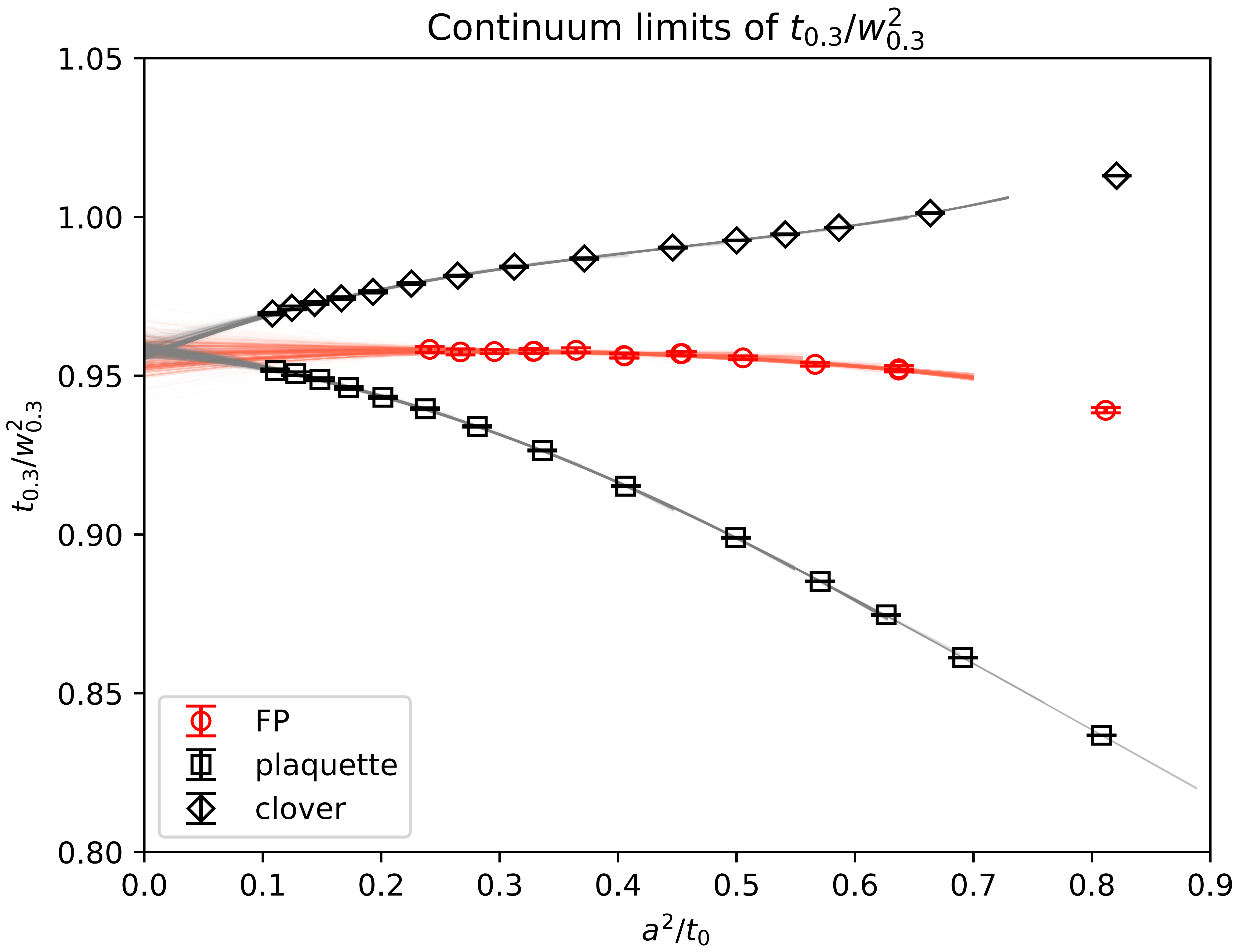
yielding dimensionless ratios t_x/w_x^2 as scaling quantities vs. $a^2/t_{0.3}$:

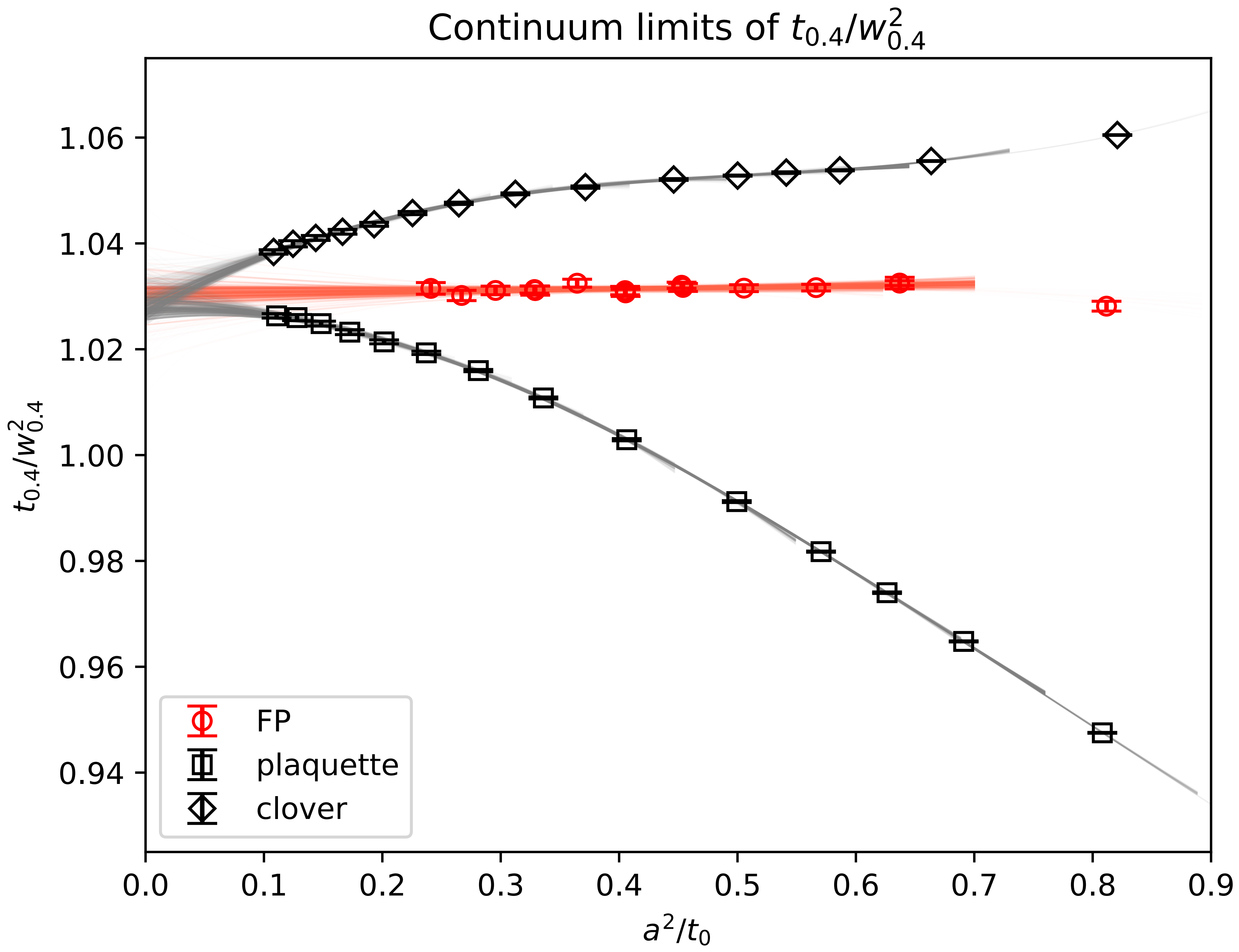


Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle\,|_{t=t_x} = x, \qquad t\dfrac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

yielding dimensionless ratios t_x/w_x^2 as scaling quantities vs. $a^2/t_{0.3}$:

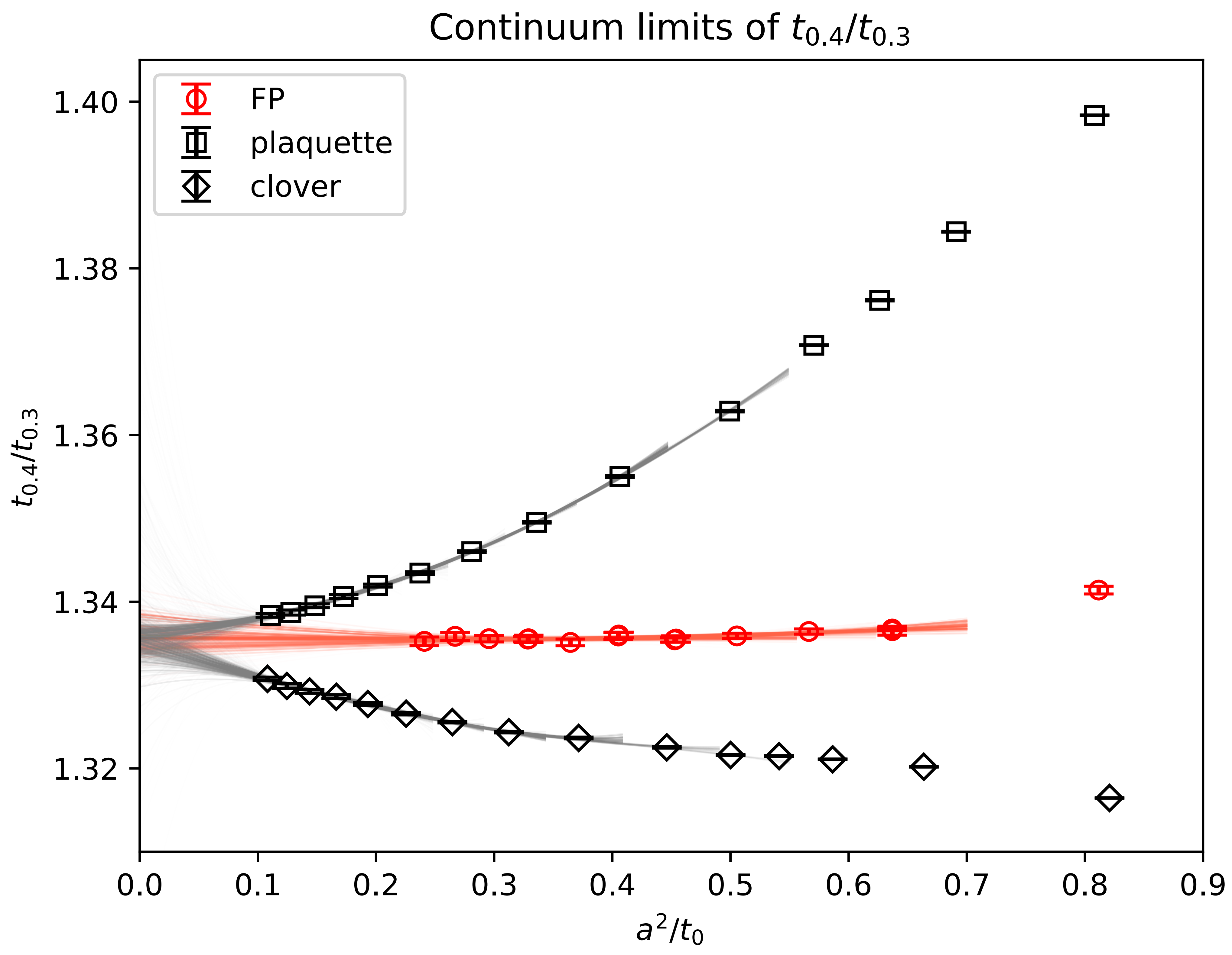


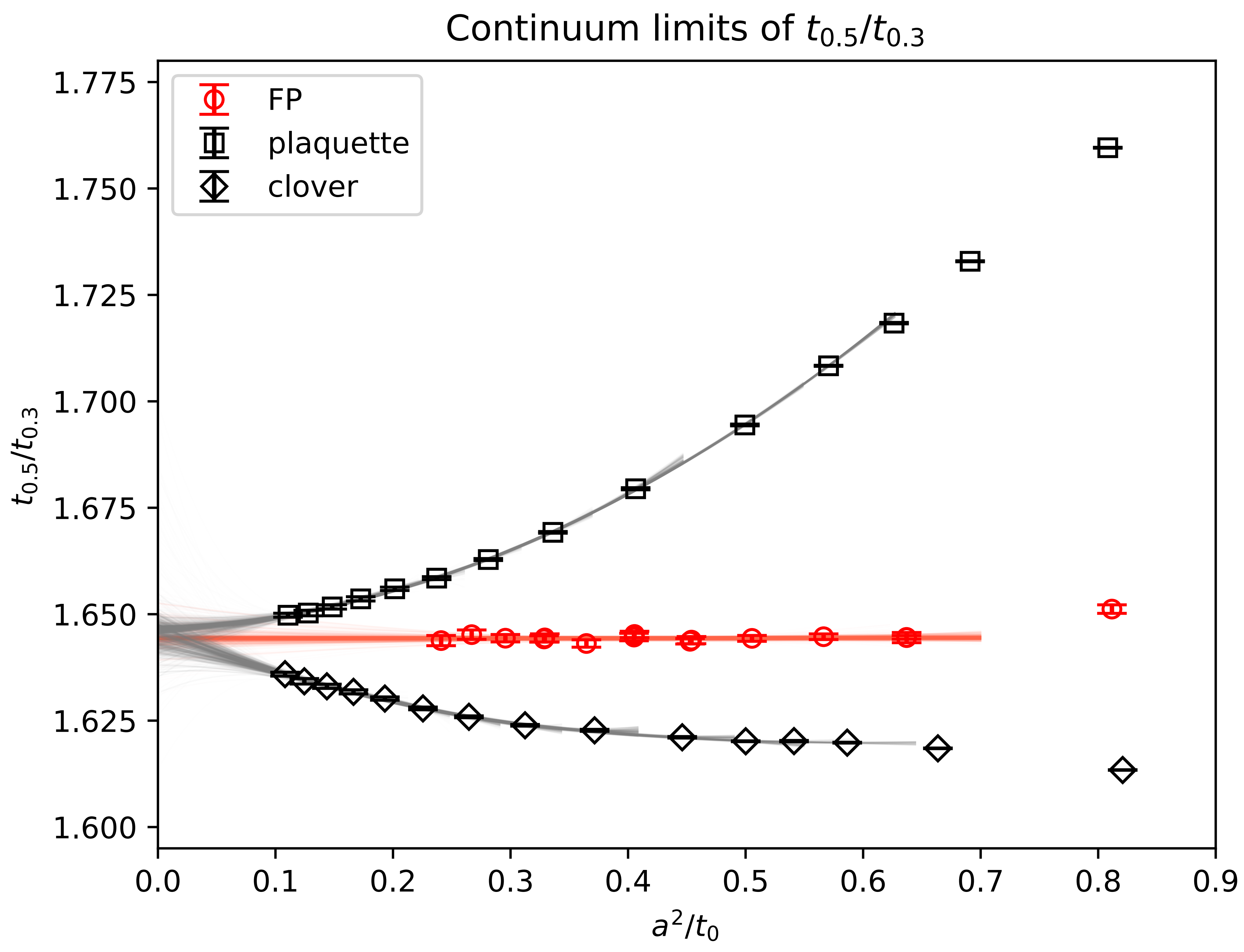


Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle\,|_{t=t_x} = x, \qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

yielding dimensionless ratios t_x/w_x^2 as scaling quantities vs. $a^2/t_{0.3}$:

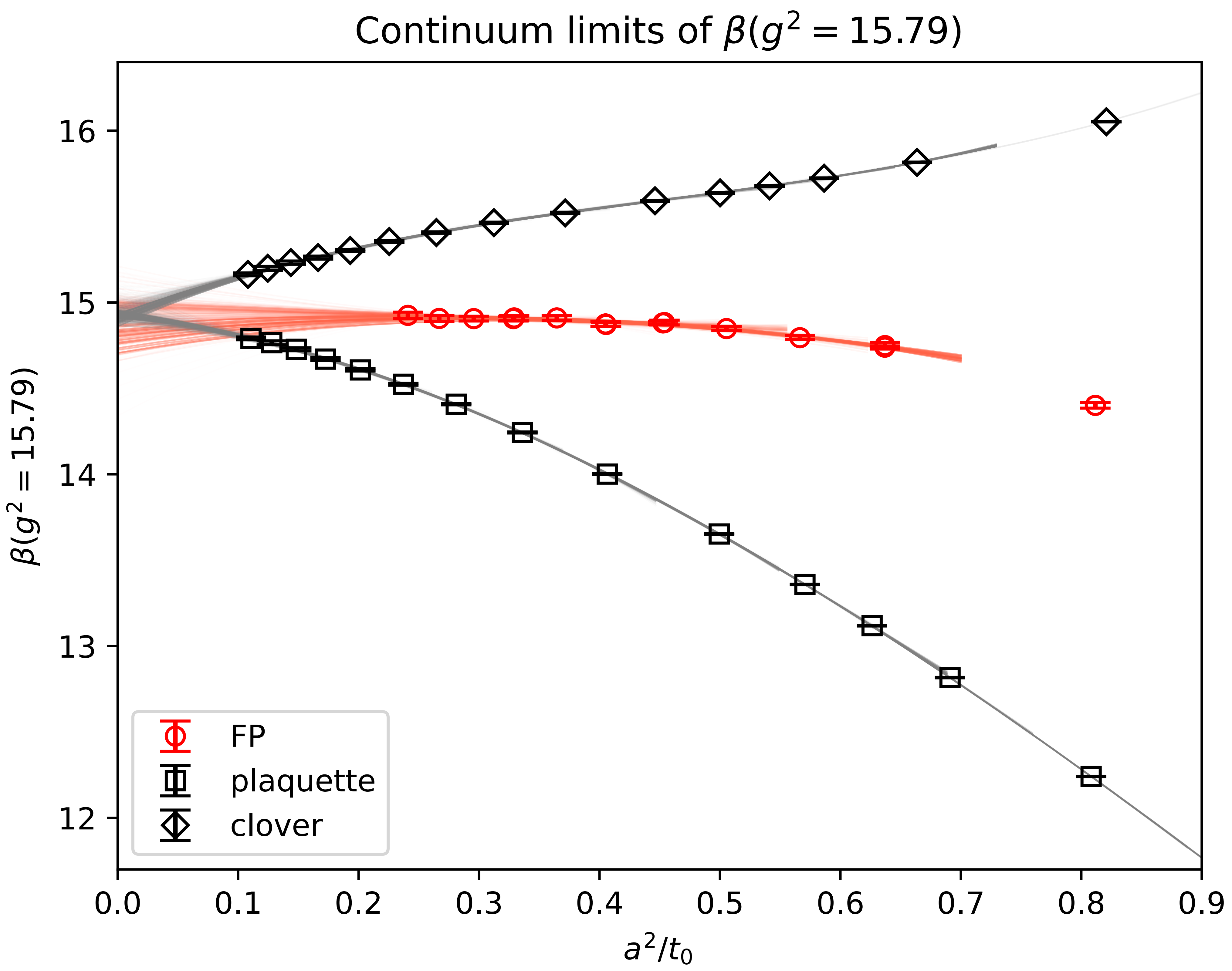




Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle|_{t=t_x} = x, \qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

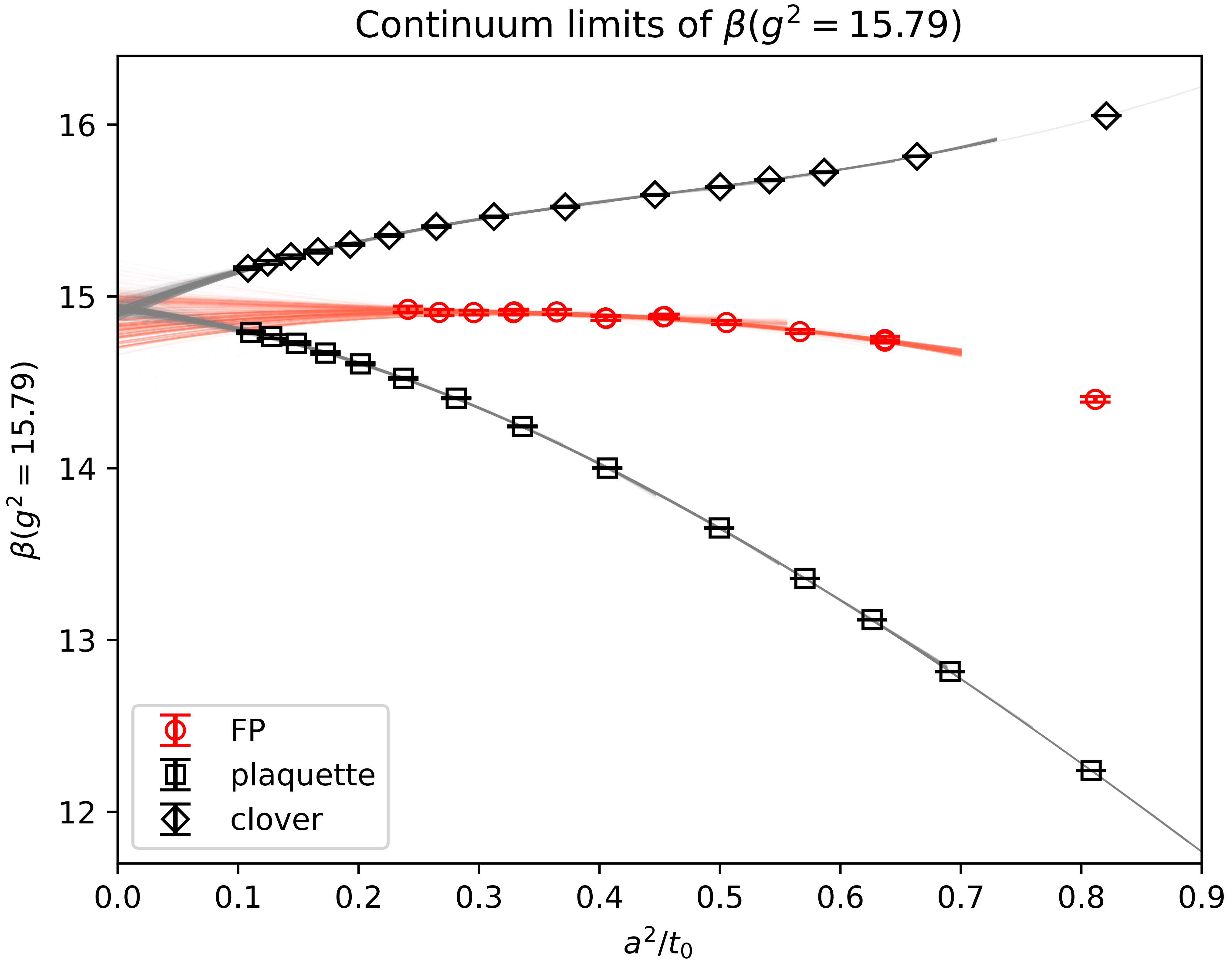
β-function at $g_{GF}^2 = 15.79$:

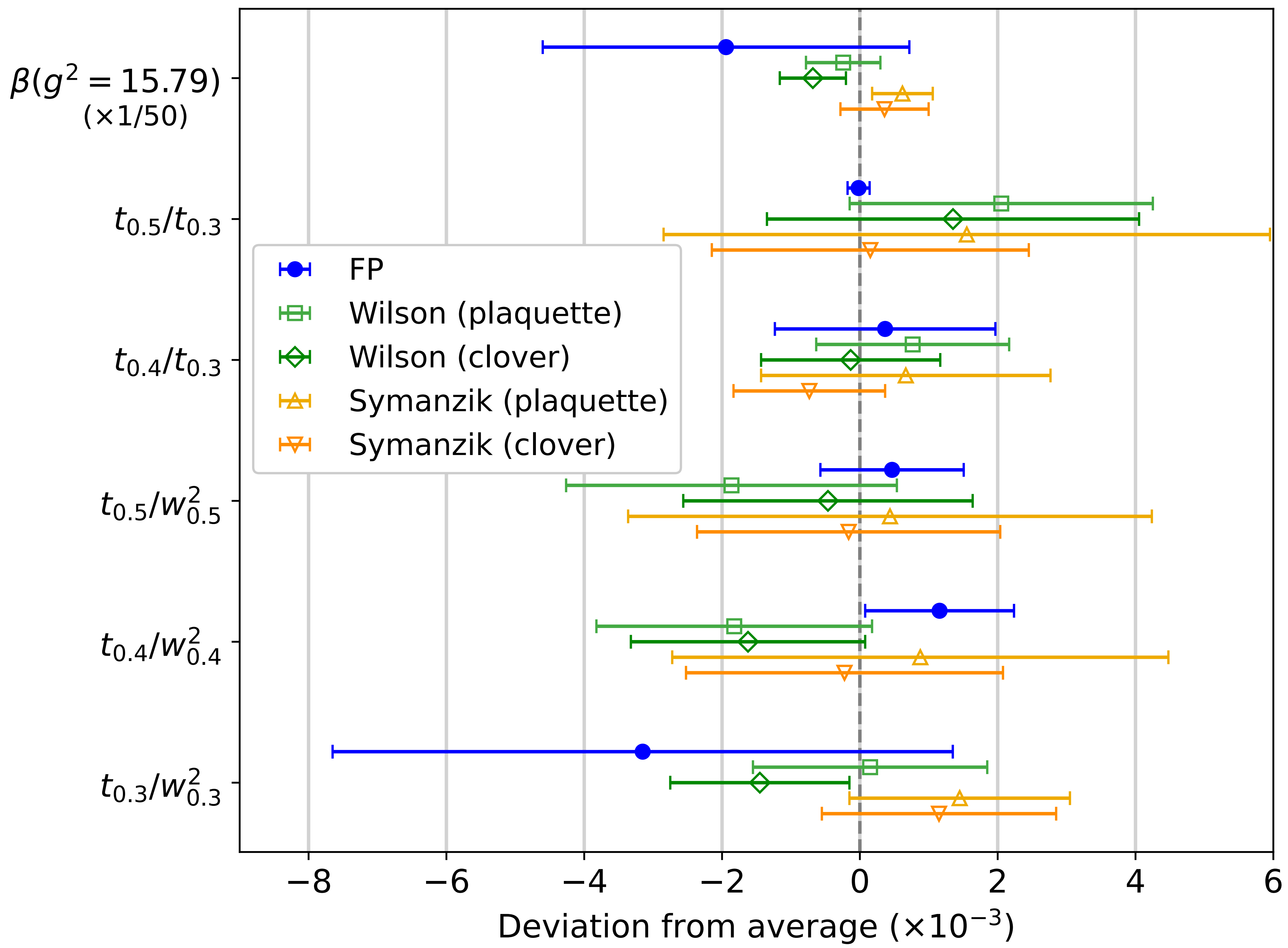


Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle\,|_{t=t_x} = x, \qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

β-function at $g_{GF}^2 = 15.79$:

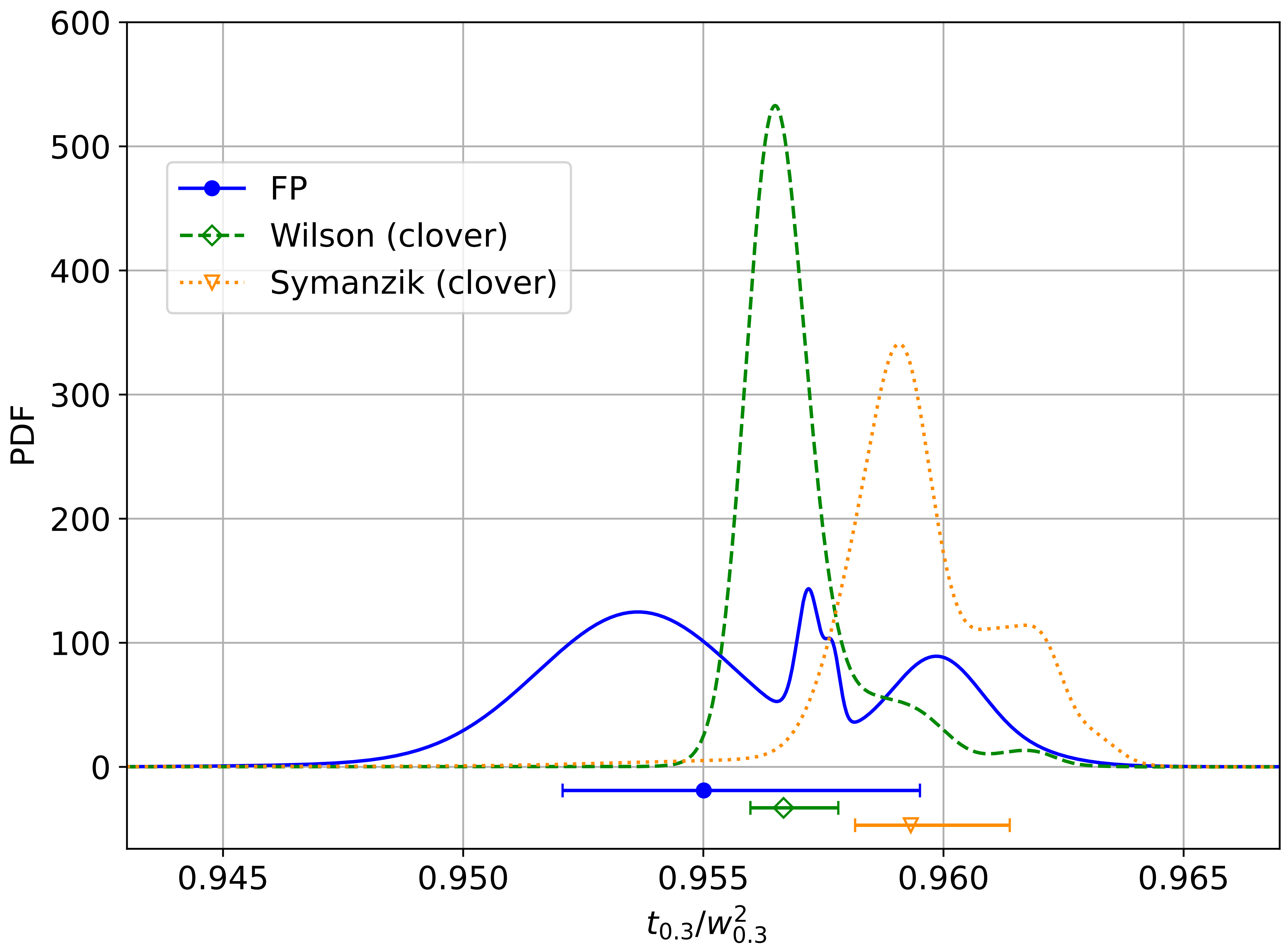


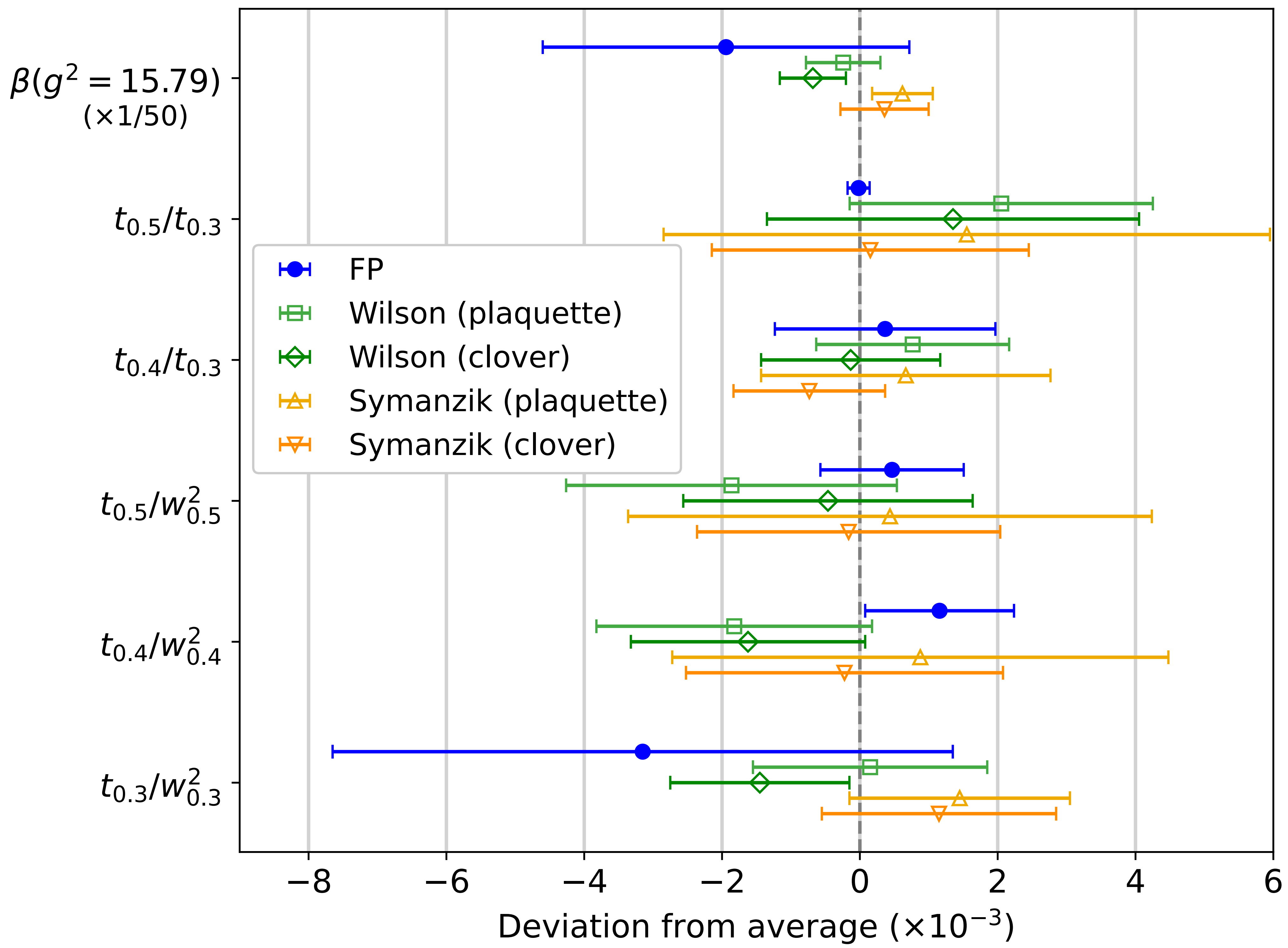


Scaling of gradient-flow scales

Physical reference scales defined through $t^2\langle E\rangle\,|_{t=t_x} = x, \qquad t\frac{d}{dt}\left(t^2\langle E\rangle\right)\Big|_{t=w_x^2} = x$

FP errors often systematics dominated:





FP action with L-CNN: Conclusions

Three questions were addressed:

- can the FP action be parametrised sufficiently well? ✓
- is the FP action sufficiently local for truncations to work? ✓
- how good are the scaling properties of the L-CNN FP action? ✓

This provides a solution to critical slowing down and topological freezing…

Availability of derivatives from the L-CNN is the stepping stone for:

- HMC, Langevin, gradient flow
- application of exact RGT step(s)

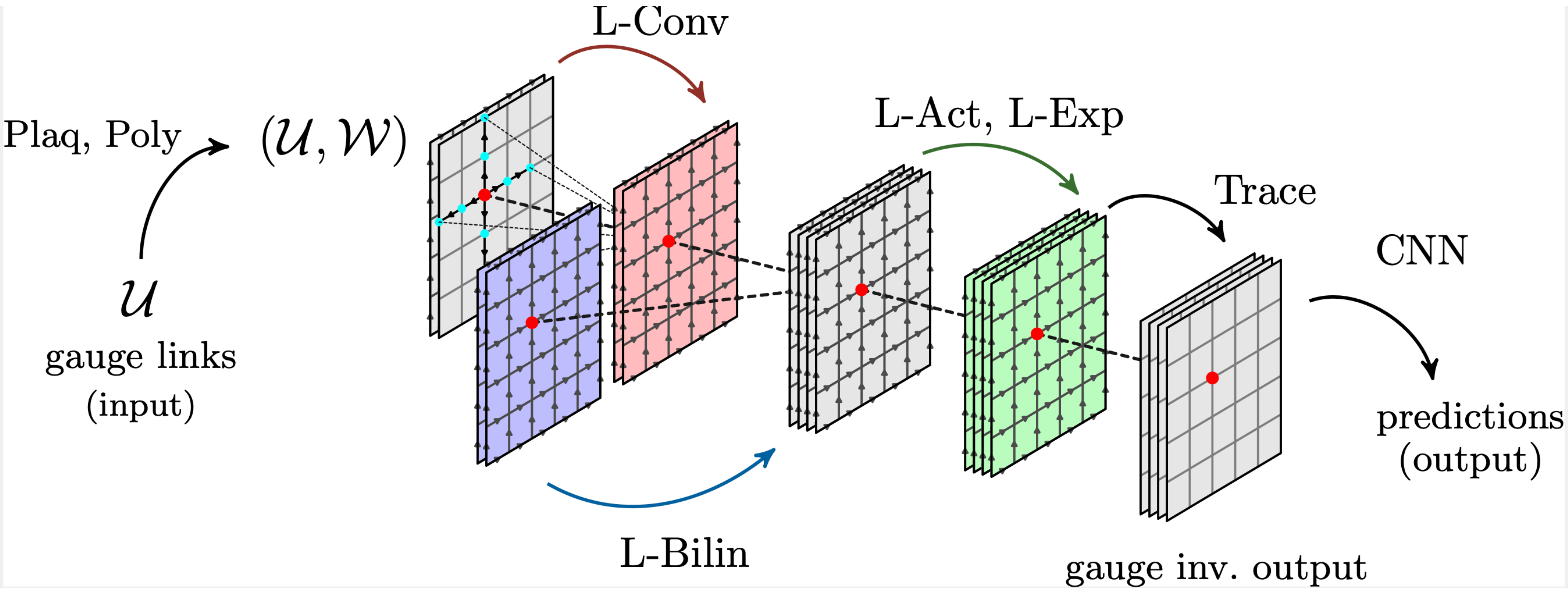
The gradient flow with FP actions is classically perfect!

Backup slides

Machine learning the FP action

ML architecture: Lattice gauge equivariant Convolutional Neural Network (L-CNN)

[Favoni, Ipp, Müller, Schuh, PRL 128 (2022) 3, 2012.12901]



L-Conv:

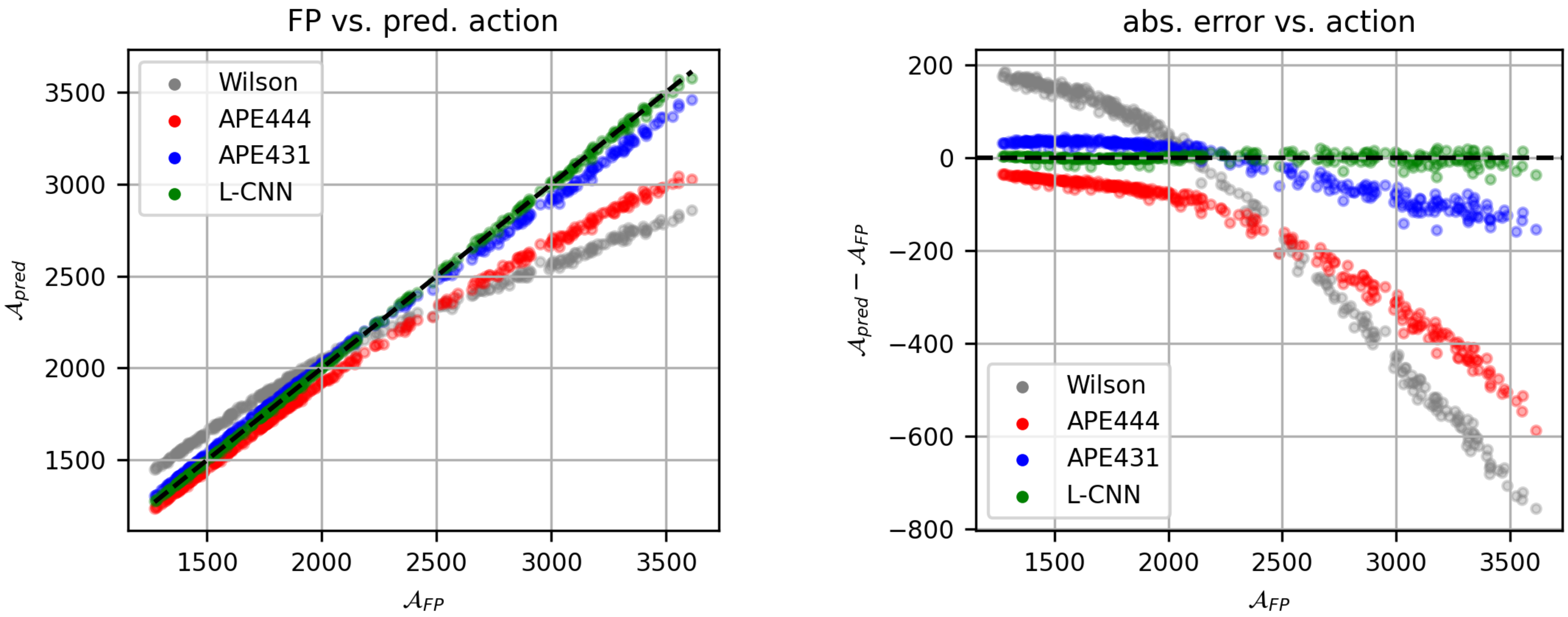
$$W'_{x+k\cdot\mu,j} = U_{x,k\cdot\mu} W_{x+k\cdot\mu,j} U^{\dagger}_{x,k\cdot\mu},$$

L-Bilin:

$$W_{x,i} \to \sum_{j,j',k} \alpha_{i,j,j',k} W_{x,j} W'_{x+k\cdot\mu,j'}.$$

Machine learning the FP action: Results

Superiority of L-CNN over old parameterization of FP action:



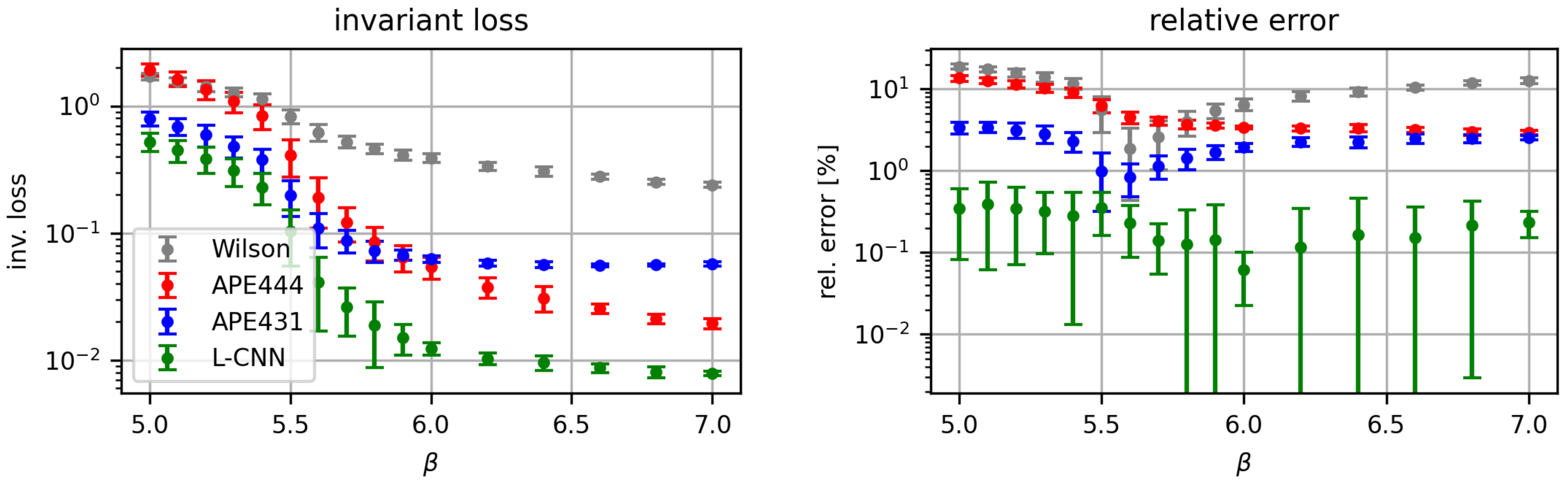
Machine learning the FP action

Training example: L-CNN model with

- 3 layers with 12, 24, 24 channels each
- parallel transport in ± 1 in first 2 layers
- local in 3rd layer

	L-CNN	APE431	APE444	Wilson
L1 (A/V)	0.02148	0.19690	0.62189	0.90898
rel. err.	0.226%	2.1965%	6.1356%	9.7577%
inv. loss (DA)	0.13480	0.23799	0.49264	0.73533

Older parametrizations of FP action as baselines:



Classically perfect gradient flow:

Starting from Wilson propagator: $D^{(0)}_{\mu\nu}(p) = \frac{\delta_{\mu\nu}}{|\hat{p}|^2} + \alpha \frac{\hat{p}_\mu \hat{p}^*_\nu}{|\hat{p}|^4}, \qquad \hat{p}_\mu = \frac{2}{a}\sin\left(\frac{ap_\mu}{2}\right)$

the propagator maintains its form after an arbitrary number of RG iterations:

$$D_{\mu\nu}(p) = G_{\mu\nu}(p) + \alpha f(p)\hat{p}_\mu \hat{p}^*_\nu$$

with

$$G'_{\mu\nu}(p_B) = \frac{1}{16}\sum_{l=0}^{1}\left[\omega(\frac{p_B}{2}+\pi l)G(\frac{p_B}{2}+\pi l)\omega^\dagger(\frac{p_B}{2}+\pi l)\right]_{\mu\nu} + \frac{1}{\kappa}\delta_{\mu\nu}$$

and

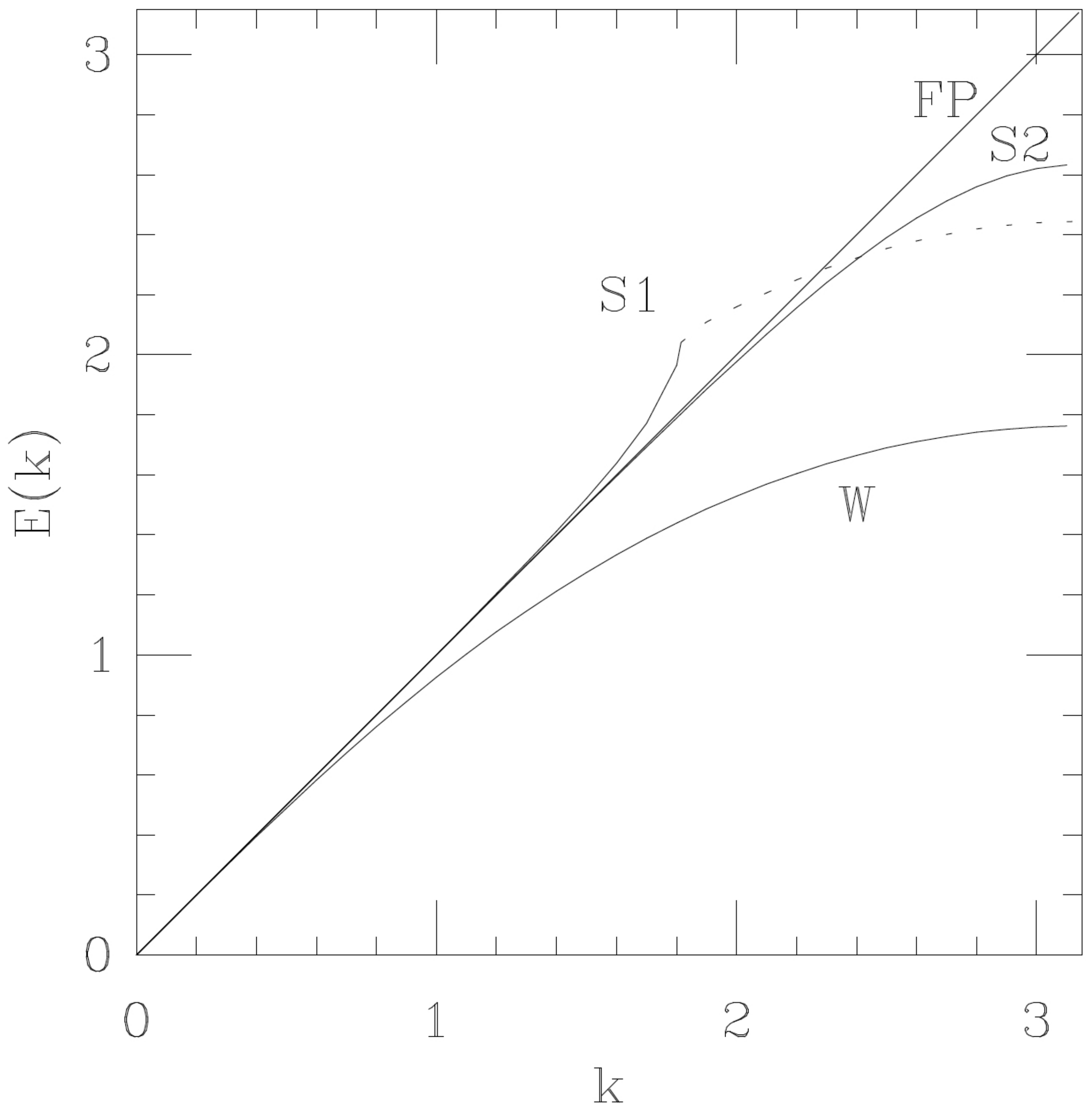
$$f'(p_B) = \frac{1}{16}\sum_{l=0}^{1} f(\frac{p_B}{2}+\pi l)$$

$\Rightarrow$ gradient flow with FP actions is classically perfect!

Classically perfect gradient flow:

After n RGT steps the part $\propto \delta_{\mu\nu}$ reads: $$\left[\Omega^{(n)}\left(\frac{p+2\pi l}{2^n}\right)\Omega^{(n)\dagger}\left(\frac{p+2\pi l}{2^n}\right)\right]_{\mu\nu}\frac{1}{(p+2\pi l)^2}$$

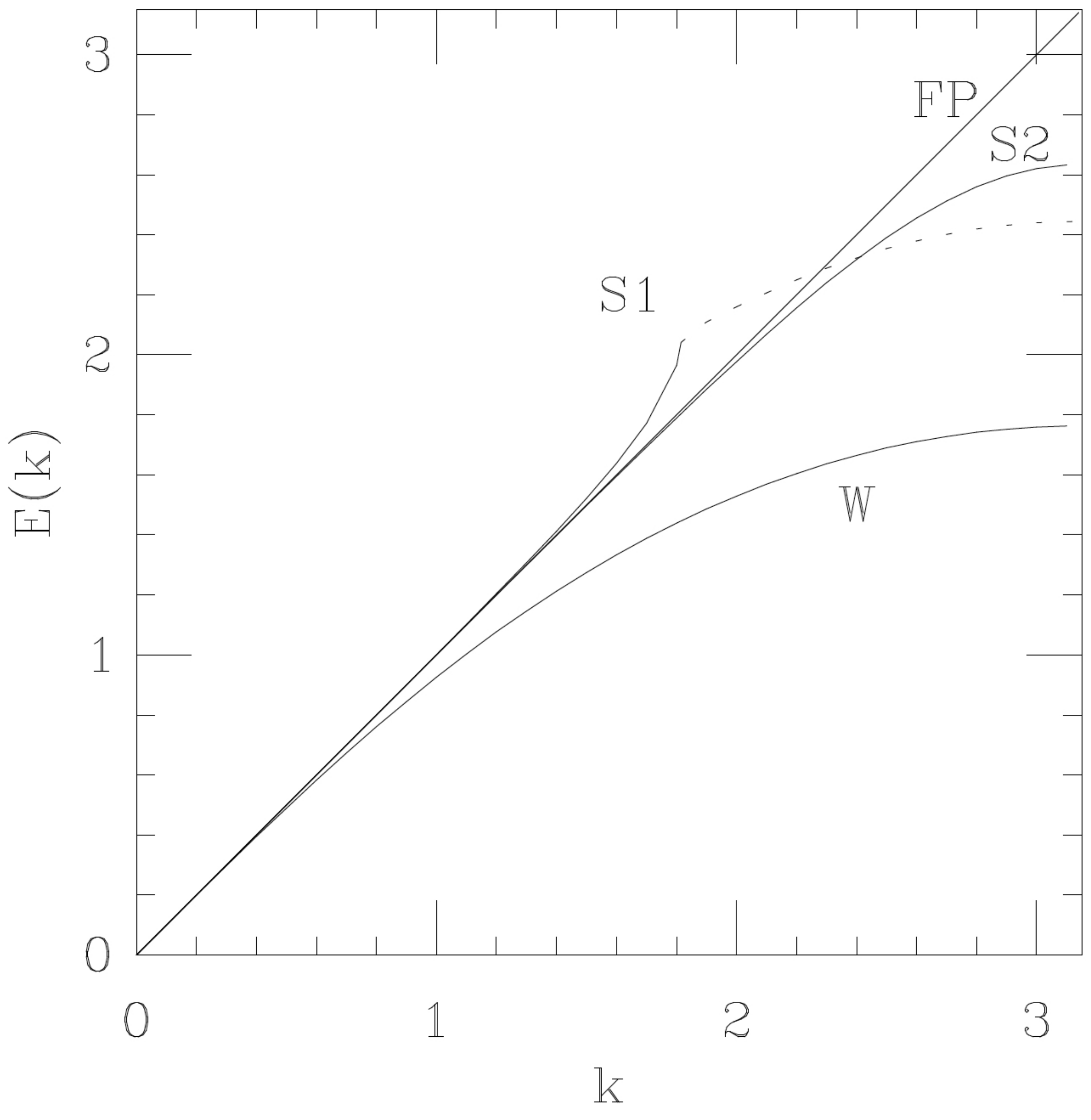
with the dispersion relation from the poles:



Classically perfect gradient flow:

After n RGT steps the part $\propto \delta_{\mu\nu}$ reads: $\left[\Omega^{(n)}\left(\frac{p+2\pi l}{2^n}\right)\Omega^{(n)\dagger}\left(\frac{p+2\pi l}{2^n}\right)\right]_{\mu\nu}\frac{1}{(p+2\pi l)^2}$

with the dispersion relation and the momentum range extended to $-\infty \leq p_\mu \leq \infty$:



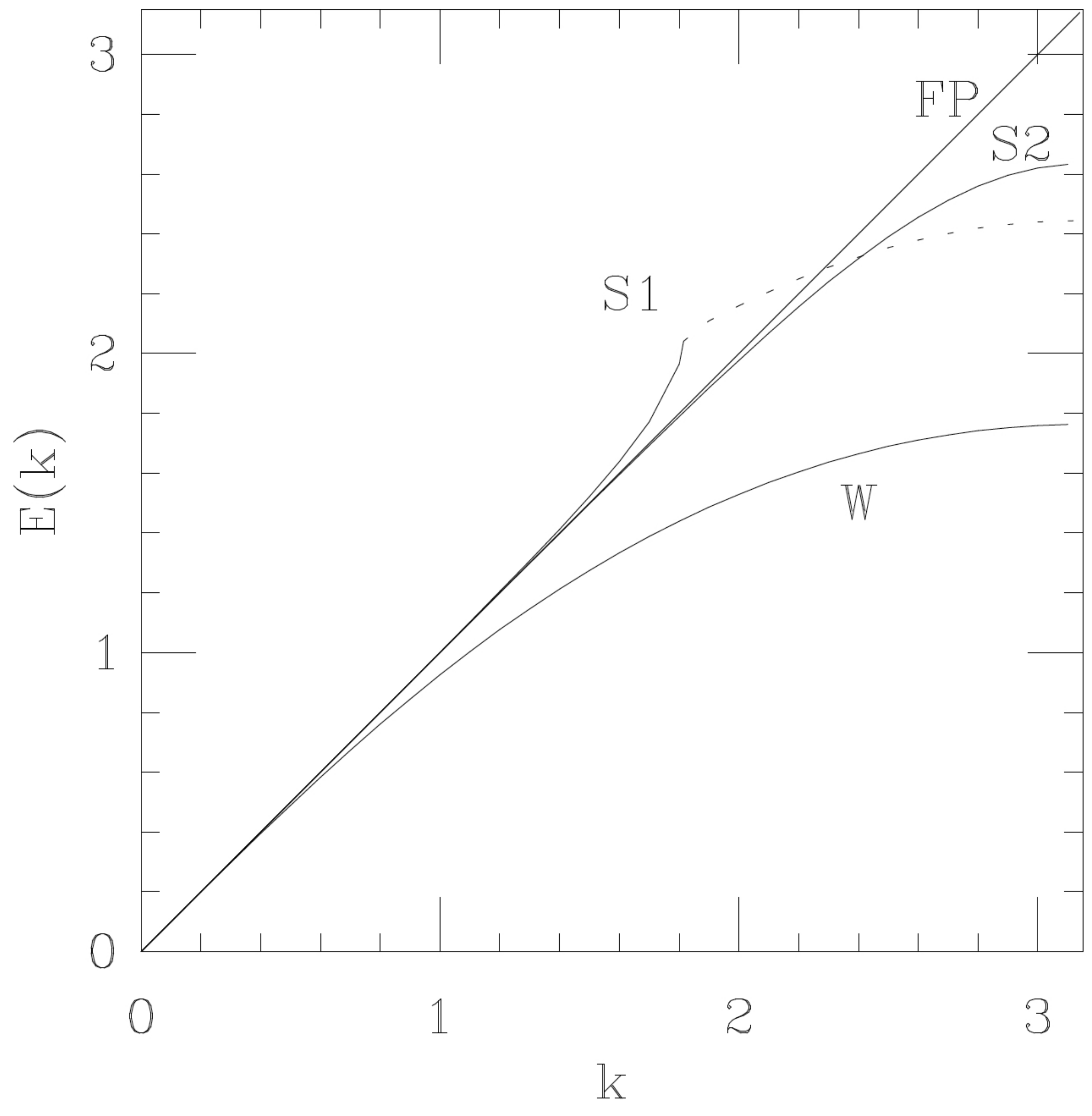
$(p+2\pi l)^2$ for $l = 0,1,2,\ldots$

through the iterated RGTs.

Classically perfect gradient flow:

After n RGT steps the part $\propto \delta_{\mu\nu}$ reads: $$\left[\Omega^{(n)}\left(\frac{p+2\pi l}{2^n}\right)\Omega^{(n)\dagger}\left(\frac{p+2\pi l}{2^n}\right)\right]_{\mu\nu}\frac{1}{(p+2\pi l)^2}$$

with the dispersion relation and the momentum range extended to $-\infty \leq p_\mu \leq \infty$:



$$(p+2\pi l)^2 \qquad \text{for } l = 0,1,2,\ldots$$

through the iterated RGTs.

$\Rightarrow$ gradient flow with FP actions is classically perfect!