Multiscale Monte Carlo Thermalization

Michael G. Endres
(MIT)

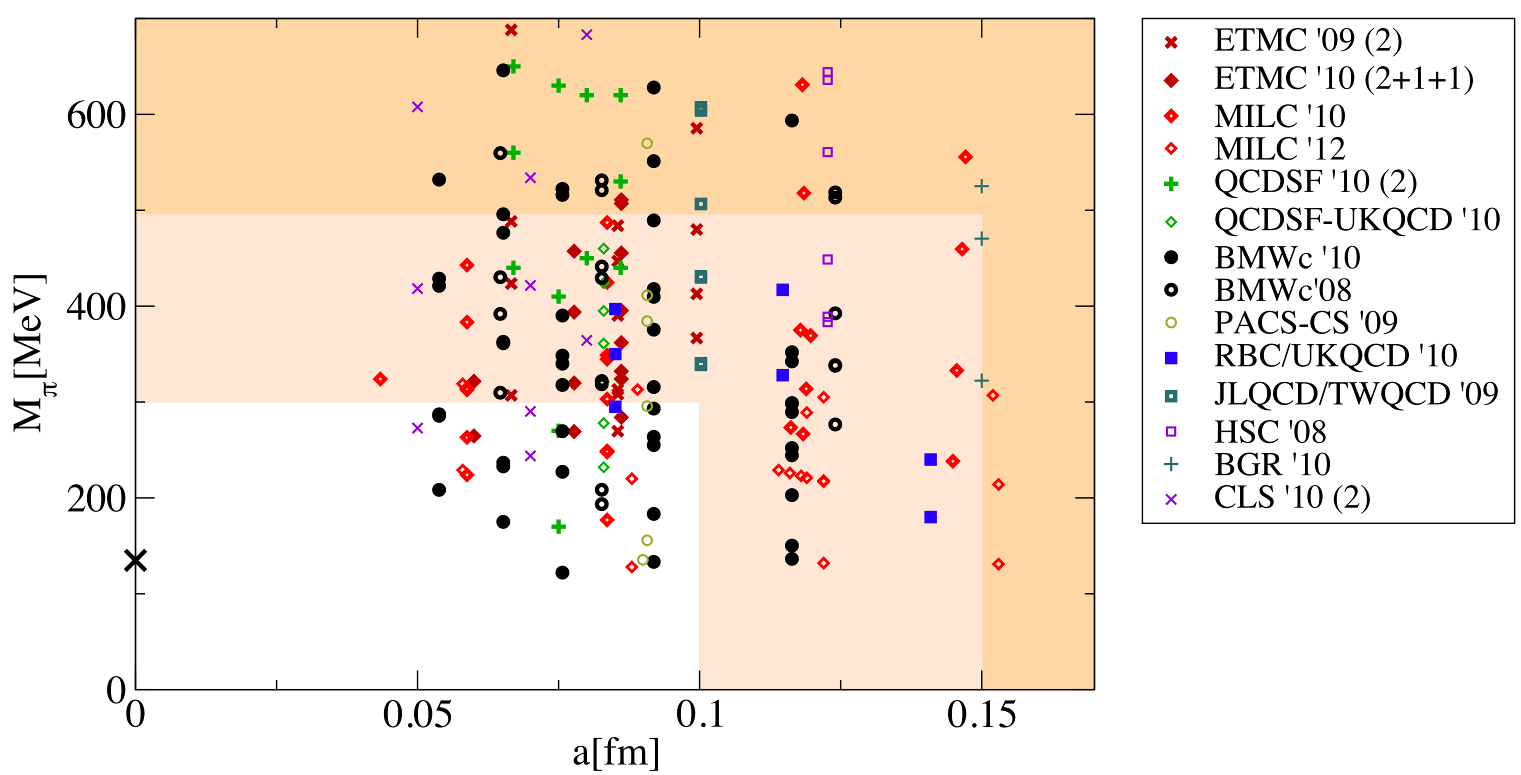
RBRC Workshop at BNL • Upton, NY
March 10, 2016

M. G. E., R. C. Brower, W. Detmold, K. Orginos and A.V. Pochinksy
Phys.Rev. D **92** (2015) 114516 [arXiv:1510.04675]

Motivation

- Critical slowing down
  - poor sampling of topology when a < 0.05 fm
  - physical pion masses are costly
- Efficient means of generating large physical volume lattices

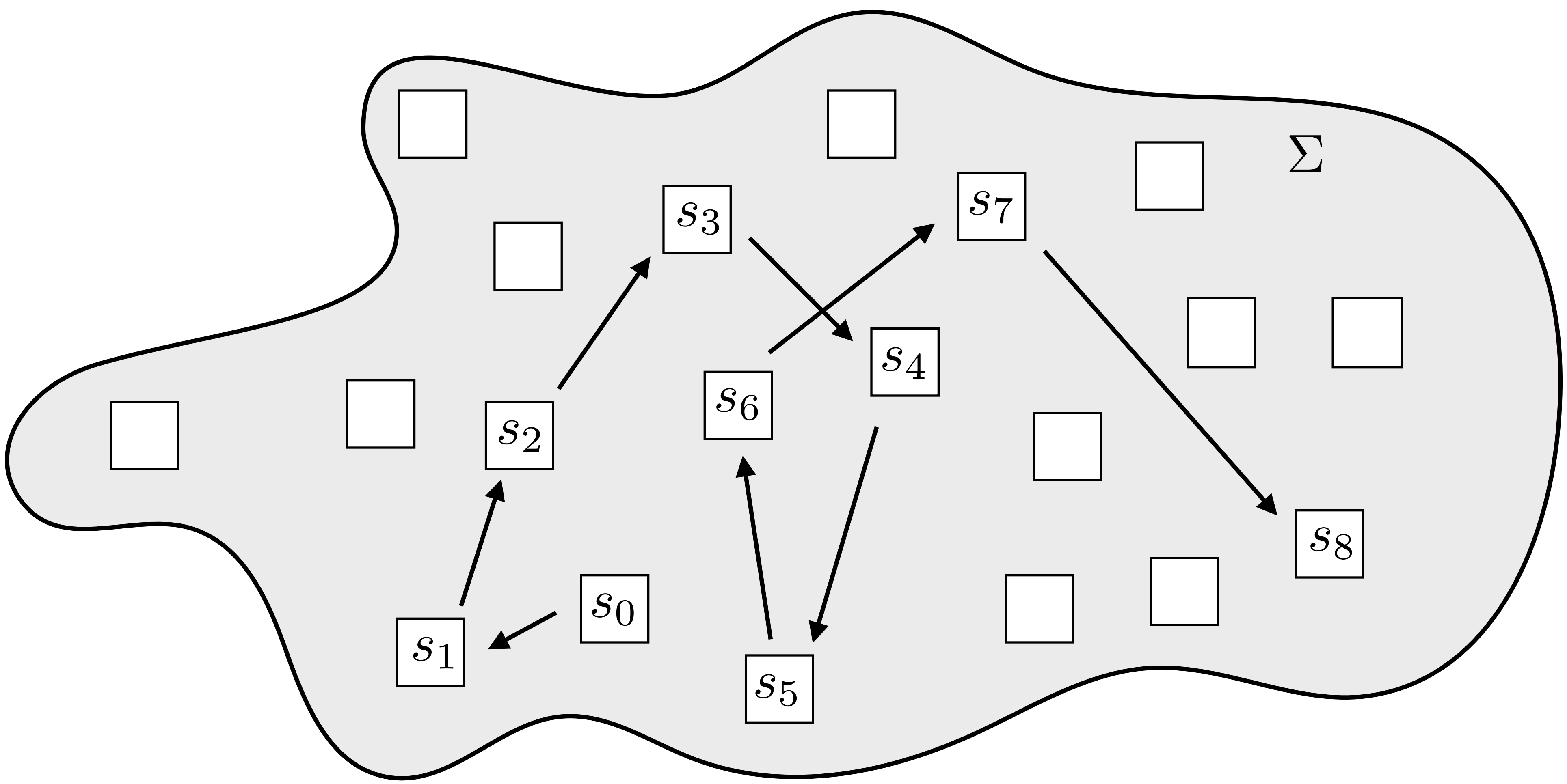
C. Hoelbling (2014) [arXiv:1410.3403]



Markov Chain Monte Carlo

$$\langle \mathcal{O} \rangle \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}(s_i) + \mathrm{O}\left(N^{-1/2}\right) \qquad s_i \text{ drawn from } \mathcal{P}(s) = \frac{e^{-\mathcal{A}(s)}}{Z}$$

Generation of s_i determined by a transition probability $\mathcal{M}(s', s)$ for $s \to s'$



Markov Chain Monte Carlo

$$\langle \mathcal{O} \rangle \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}(s_i) + \mathrm{O}\left(N^{-1/2}\right)$$

s_i drawn from $\mathcal{P}(s) = \dfrac{e^{-\mathcal{A}(s)}}{Z}$

Generation of s_i determined by a transition probability $\mathcal{M}(s', s)$ for $s \to s'$

Defines the "algorithm"

Correct sampling of path integral measure requires:

$$\sum_{s \in \Sigma} \mathcal{M}(s', s) \mathcal{P}(s) = \mathcal{P}(s')$$

stationary distribution

$$\sum_{s' \in \Sigma} \mathcal{M}(s', s) = 1$$

probability conservation

Markov Chain Monte Carlo — evolution

Spectral decomposition of a Markov process:

$$\mathcal{M} = \sum_{n\geq 0} e^{-1/\tau_n} |\chi_n\rangle\langle\chi_n| \quad \Longrightarrow \quad \mathcal{M}^\tau = \sum_{n\geq 0} e^{-\tau/\tau_n} |\chi_n\rangle\langle\chi_n|$$

under appropriate assumptions, bounded by unity

Evolution of probability distribution and expectation values:

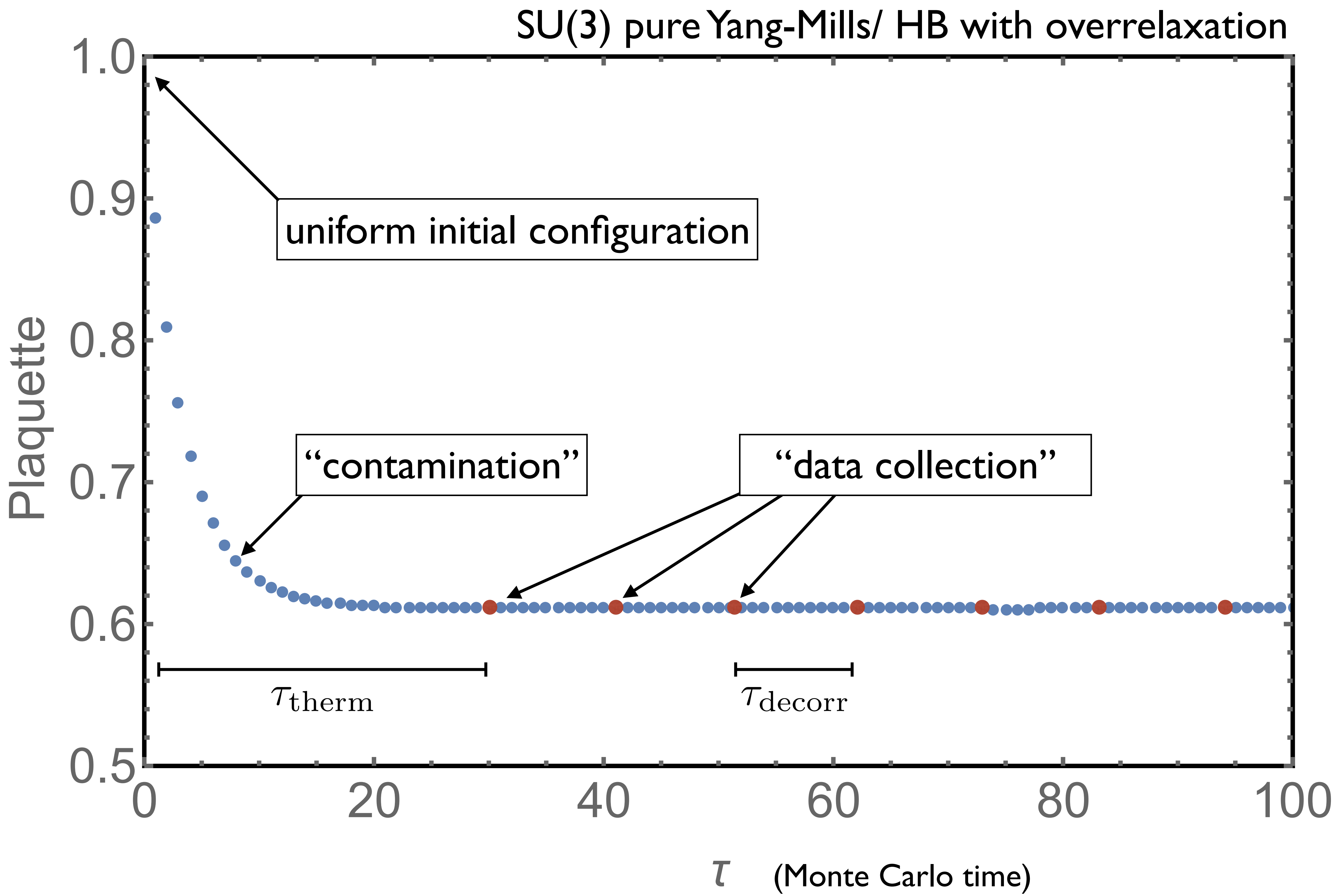
$$\mathcal{P}_\tau(s) = \mathcal{P}(s) + \sum_{n>0} \langle s|\chi_n\rangle \langle\chi_n|\mathcal{P}_0\rangle e^{-\tau/\tau_n}$$

time scales governing evolution

$$\langle\mathcal{O}\rangle_\tau = \langle\mathcal{O}\rangle + \sum_{n>0} \langle\mathcal{O}|\chi_n\rangle \langle\chi_n|\mathcal{P}_0\rangle e^{-\tau/\tau_n}$$

overlap of initial distribution $P_0(s)$ onto mode n

Markov Chain Monte Carlo — equilibration

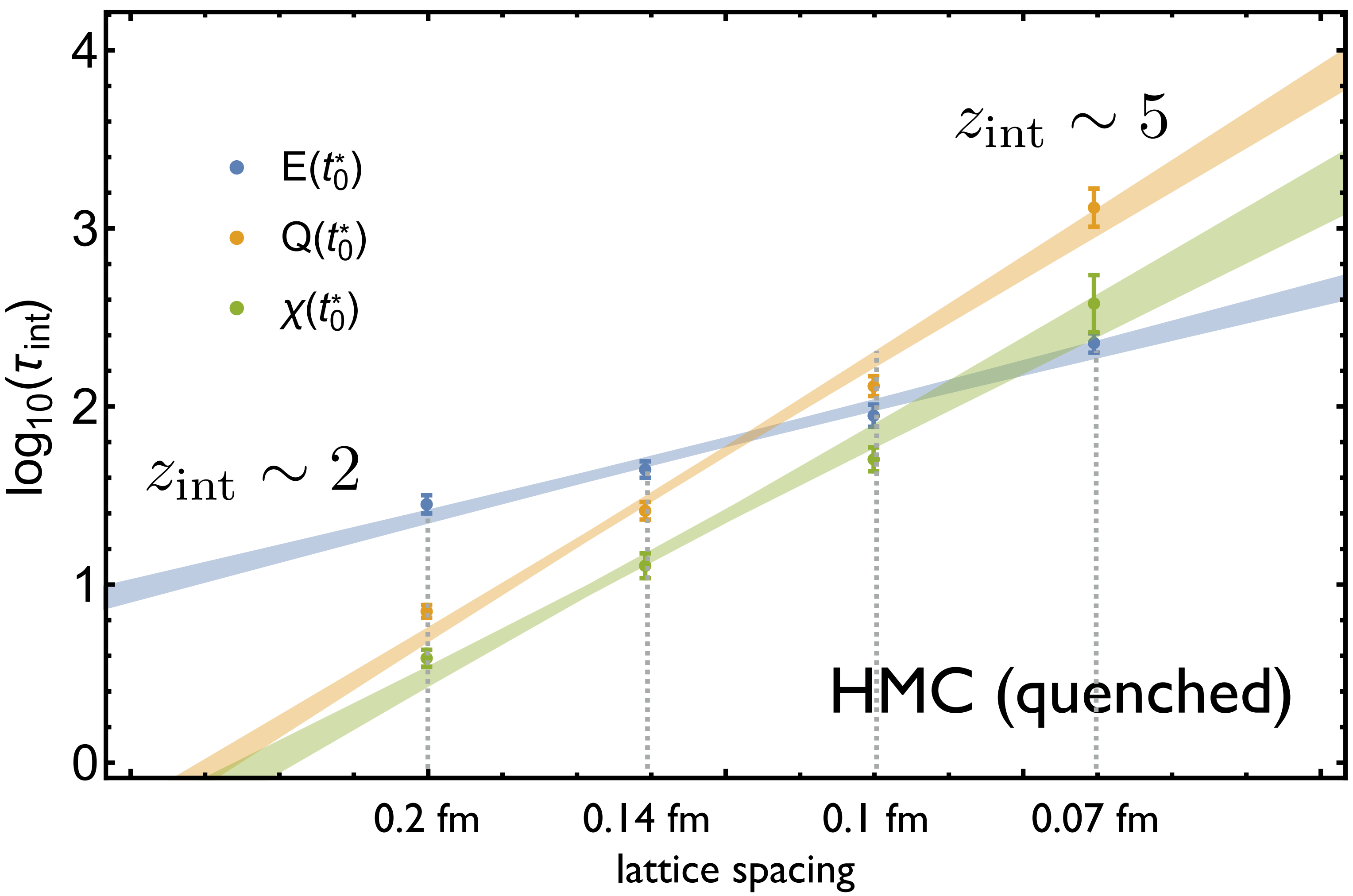


Markov Chain Monte Carlo — critical slowing down

Fine lattices decorrelate slower than coarse lattices

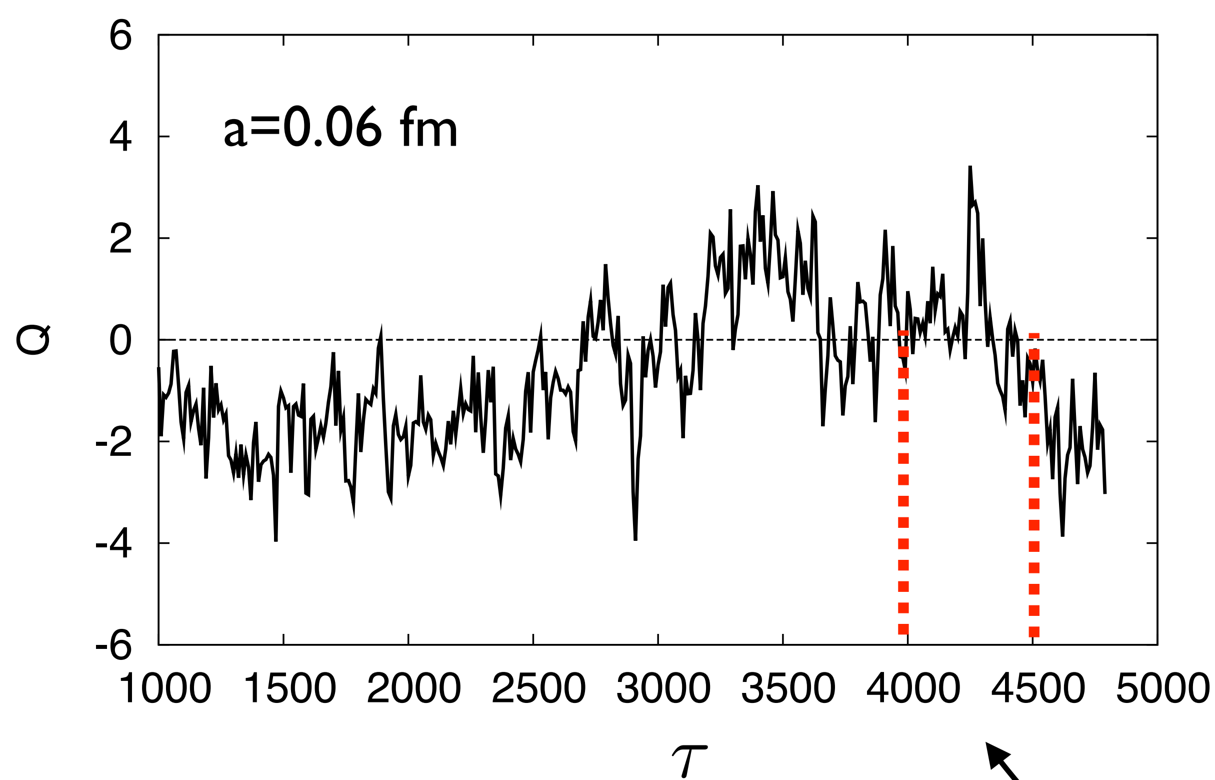
Integrated autocorrelation time

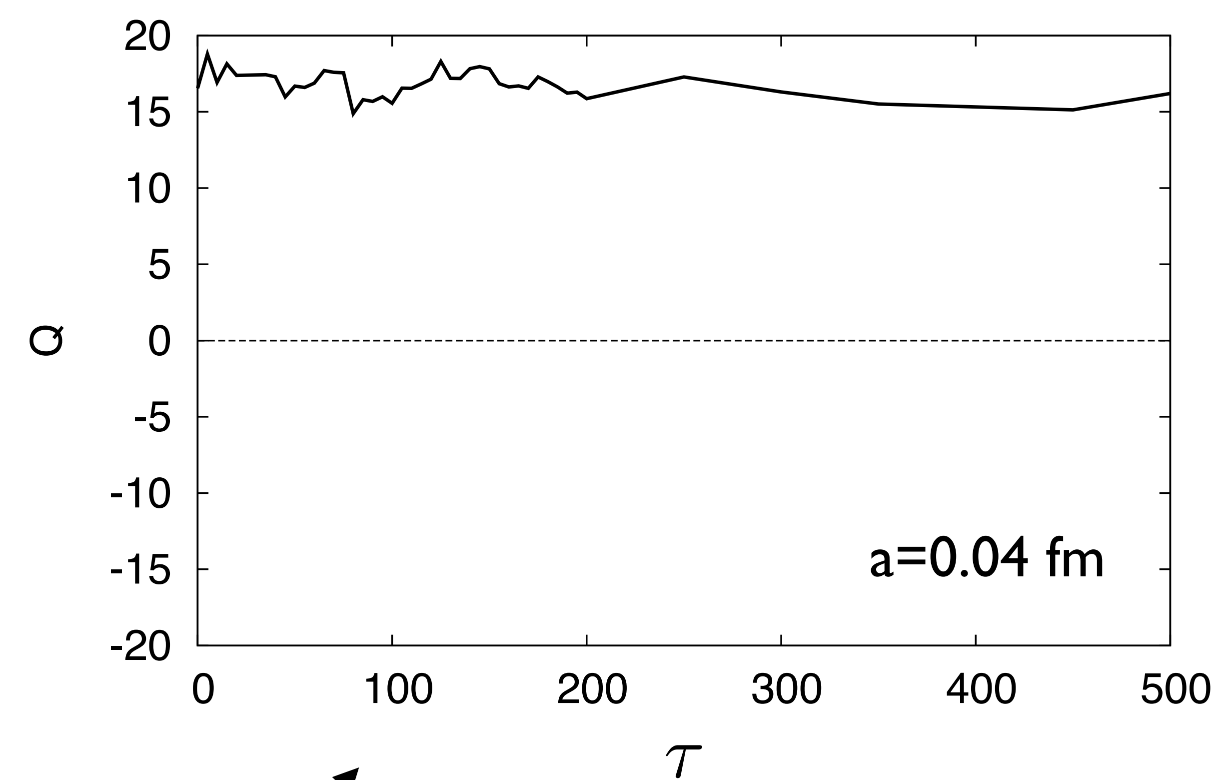
$$\tau_{\rm int}(\mathcal{O}) \sim \left(\frac{1}{a}\right)^{z_{\rm int}(\mathcal{O})} \quad \Longrightarrow \quad cost \sim \left(\frac{1}{a}\right)^{D+\max_{\mathcal{O}} z_{\rm int}(\mathcal{O})} \sim \left(\frac{1}{a}\right)^{9}$$



Markov Chain Monte Carlo — topological freezing

On a periodic lattice, topological charge fluctuations become *exponentially* suppressed in 1/a → Incorrect sampling





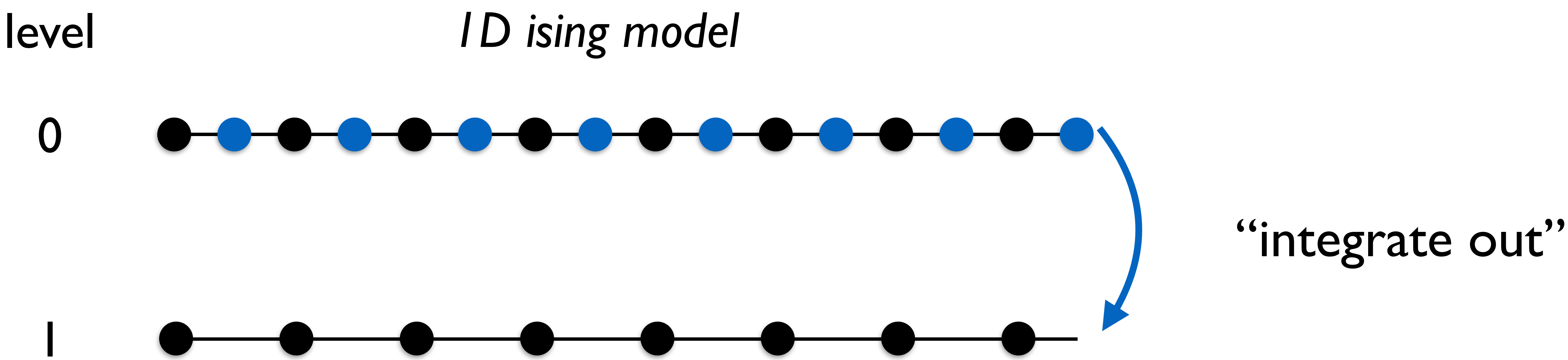
$m_\pi \sim 360 - 480$ MeV

S. Schaefer *et al.*, PoS **LAT2009** (2009) 032

Multiscale Monte Carlo

- **Ultimate goal:** an algorithm which allows for efficient updating of modes on multiple scales while retaining detailed balance
  - some progress for simple systems, remains challenging for gauge theories (QCD in particular)
- **More modest goal of this work:** realization of a multiscale *thermalization* algorithm; the strategy draws upon many ideas:
  - standard Monte Carlo techniques
  - multigrid concepts of *restriction* (coarse-graining) and *prolongation* (refinement)
  - real space renormalization

A multi-scale updating algorithm: Ising spin chain



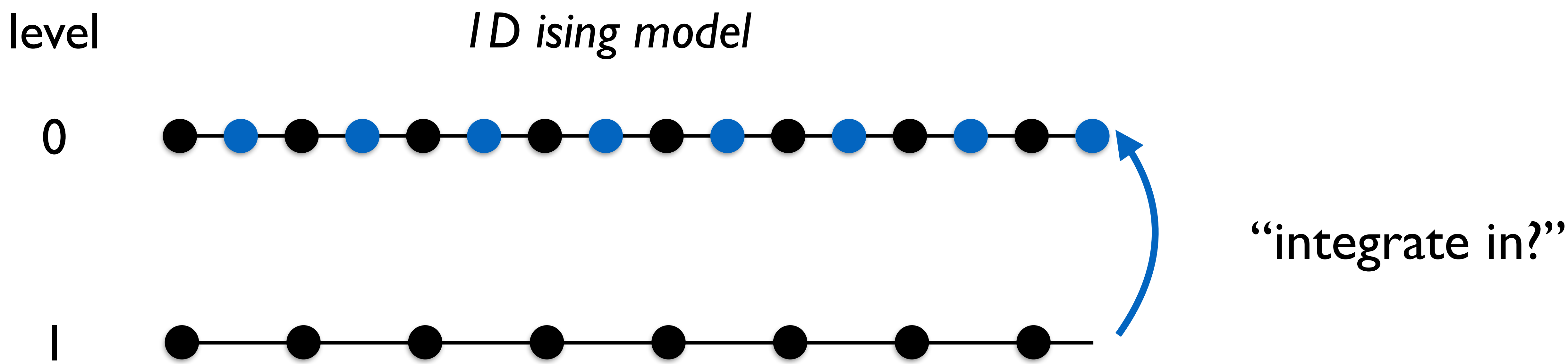
$$Z^{[\ell]} = \sum_{\{S\}} e^{-H^{[\ell]}}$$

$$H^{[0]} = J \sum_i S_{2i}(S_{2i+1} + S_{2i-1})$$

$$H^{[1]} = R(J) \sum_i S_{2i+1} S_{2i-1}$$

$$R(J) = \frac{1}{2} \cosh^{-1}\left(e^{2J}\right)$$

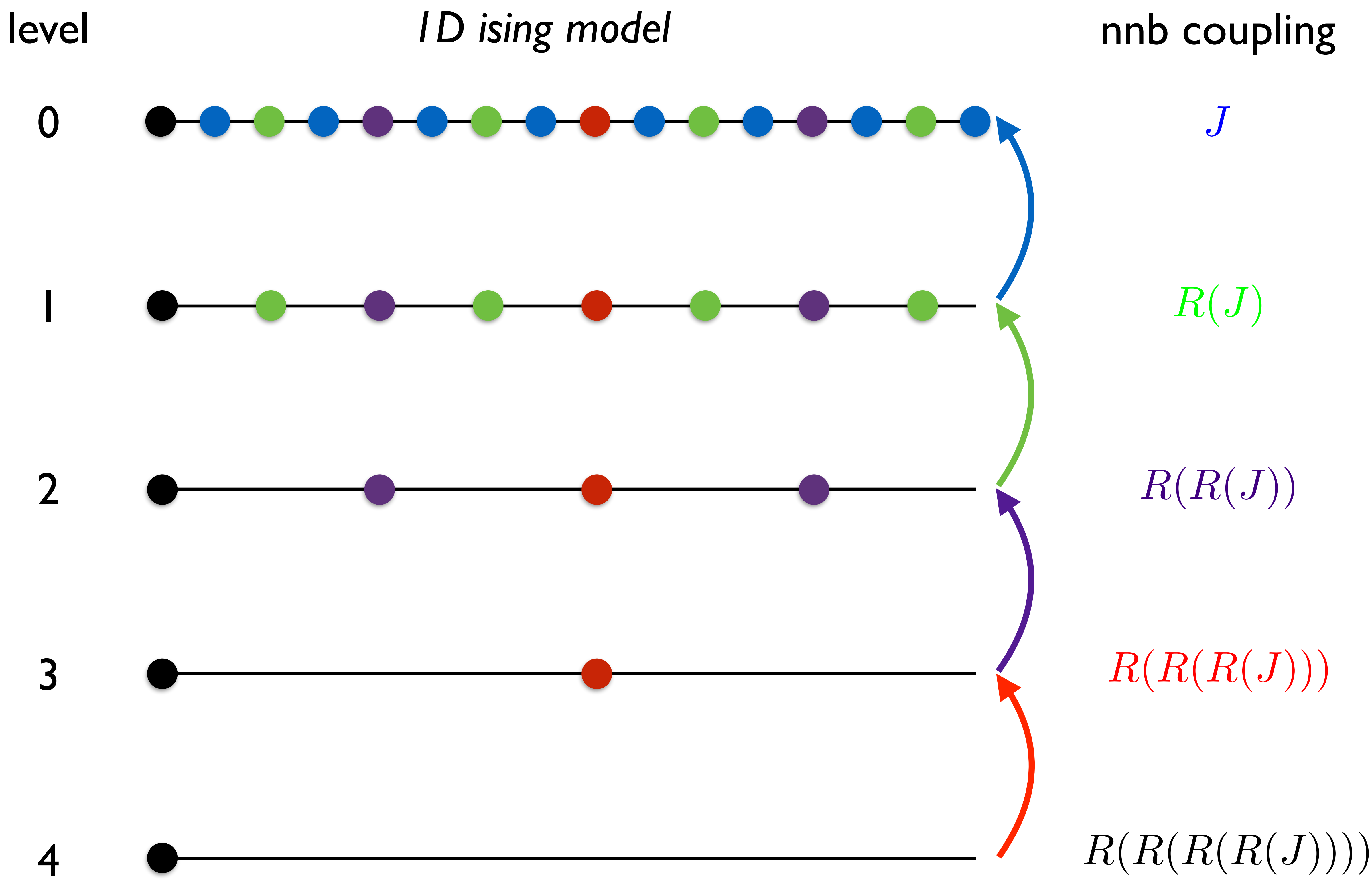
A multi-scale updating algorithm: Ising spin chain



"Integrating in" spins at the fine level (0) requires a single "heat bath" update per (undefined) site:

$$\mathcal{P}(S_{2i}) = \frac{e^{-JS_{2i}(S_{2i+1}+S_{2i-1})}}{\cosh\left(J(S_{2i+1}+S_{2i-1})\right)}$$

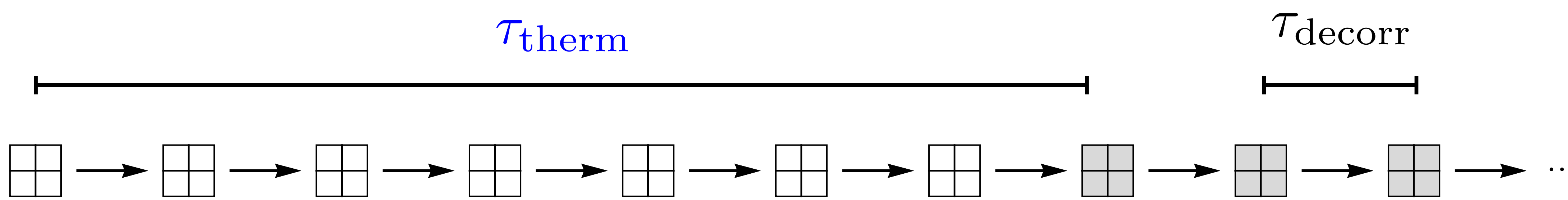
A multi-scale updating algorithm: Ising spin chain



Generalization to more complicated systems

- Generalization is achieved with approximations:
    - truncation of the coarse action; implies inexact RG matching
    - one-to-one refinement prescription based on interpolation, rather than exact prescription
- *Rethermalization* is crucial in order to correct for the errors induced by such approximations
- Effectiveness/use of approach depends on several factors
    - time scales associated with the conventional algorithm
    - refinement prescription
    - RG matching

Time scales



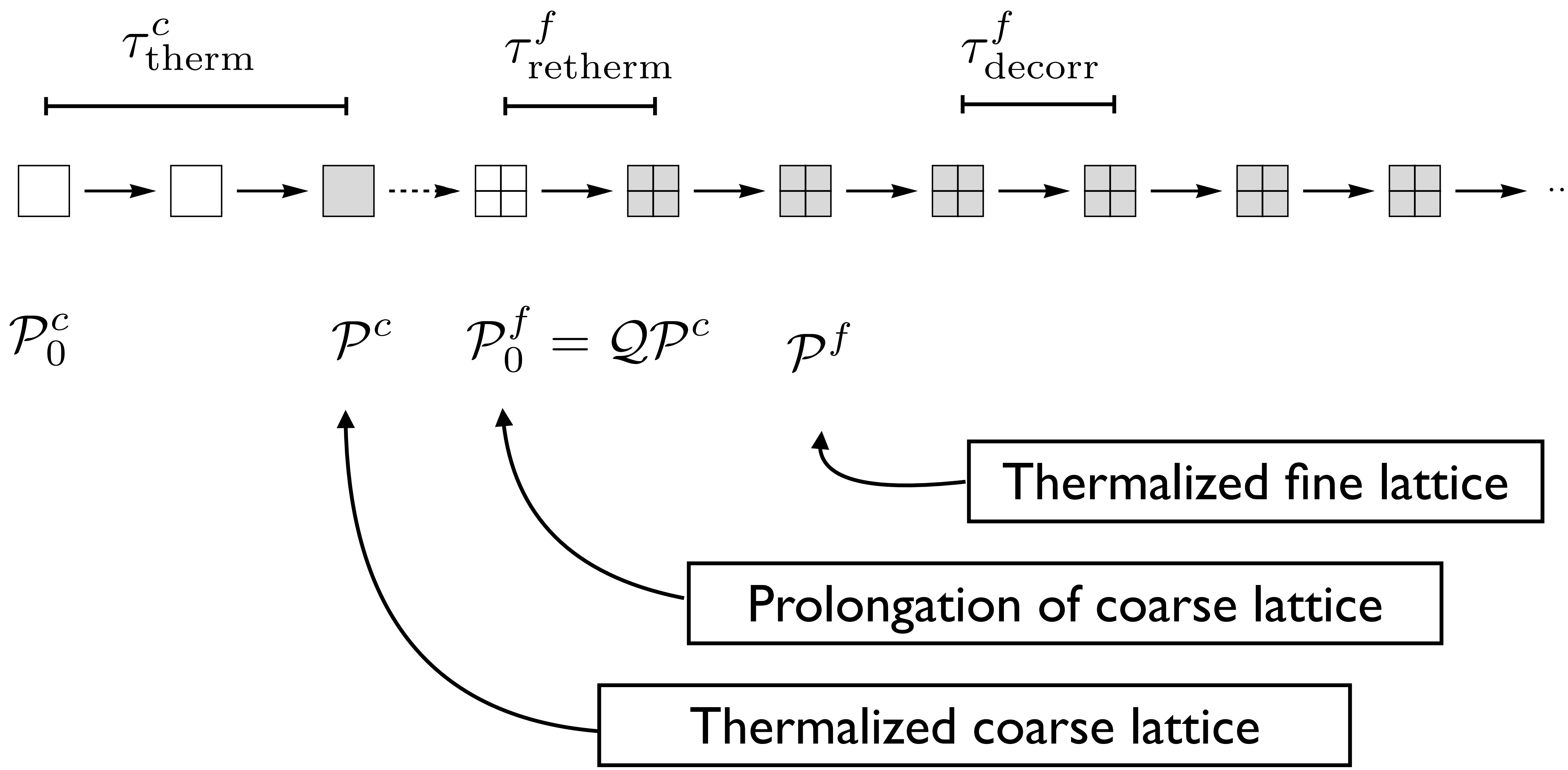
$$\mathcal{P}_\tau(s) = \mathcal{P}(s) + \sum_{n>0} \langle s|\chi_n\rangle \langle \chi_n|\mathcal{P}_0\rangle e^{-\tau/\tau_n}$$

$$\tau_{\text{decorr}} \lesssim 2\tau_1$$

$$\langle \mathcal{O}\rangle_\tau = \langle \mathcal{O}\rangle + \sum_{n>0} \langle \mathcal{O}|\chi_n\rangle \langle \chi_n|\mathcal{P}_0\rangle e^{-\tau/\tau_n}$$

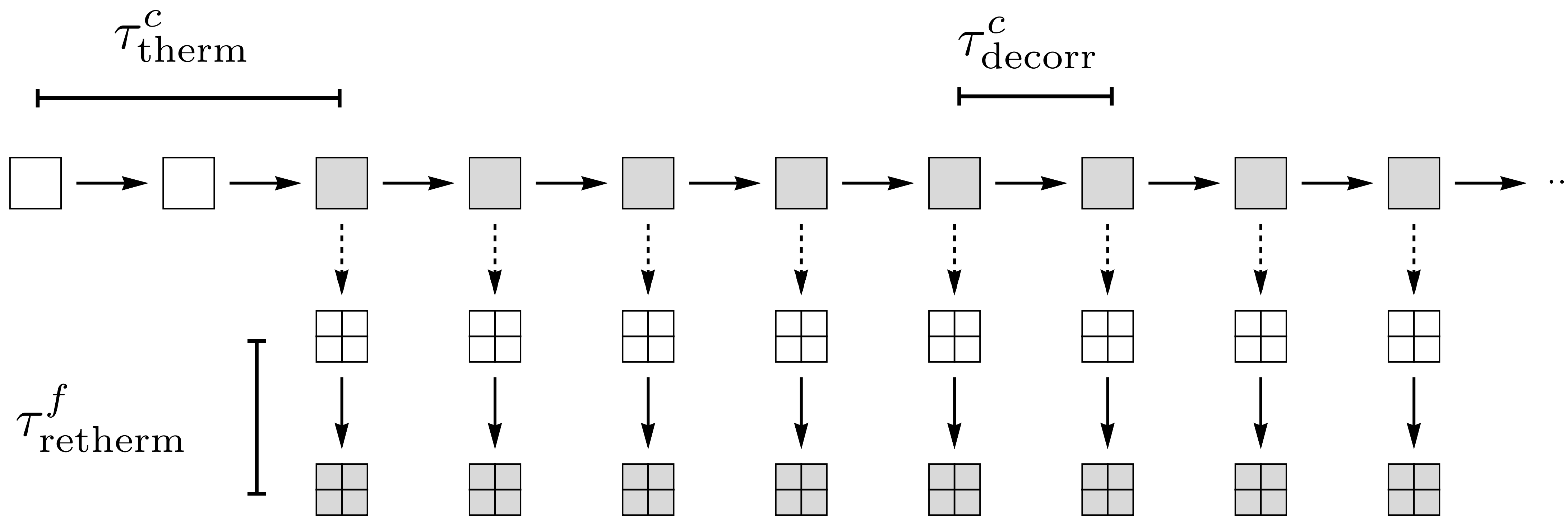
Goal: to find an initial probability distribution for which this overlap vanishes for small n

Time scales



Faster thermalization achieved if: $\tau^{c}_{\mathrm{therm}} + \tau^{f}_{\mathrm{retherm}} < \tau^{f}_{\mathrm{therm}}$

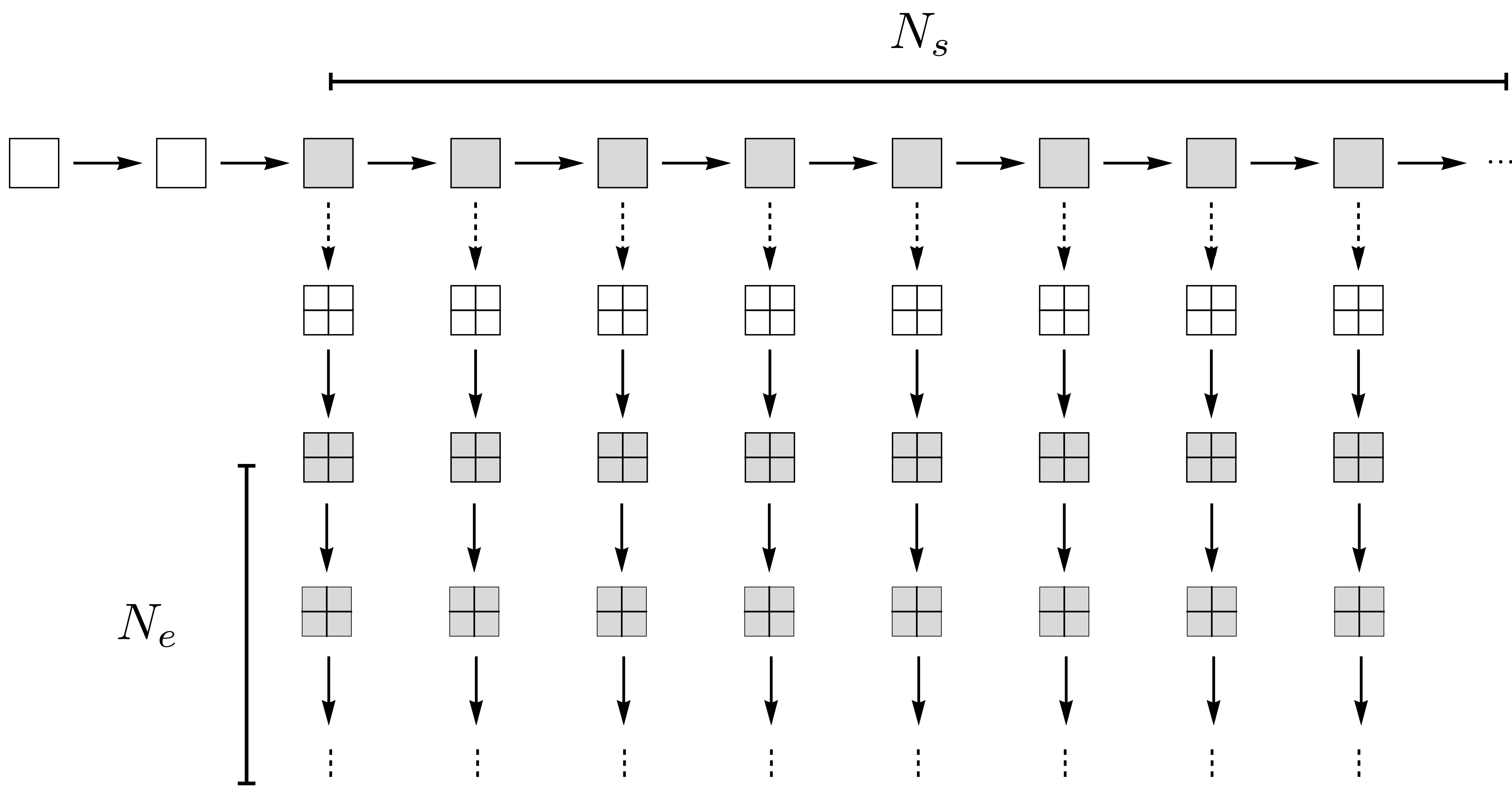
Time scales



Faster ensemble generation achieved if:

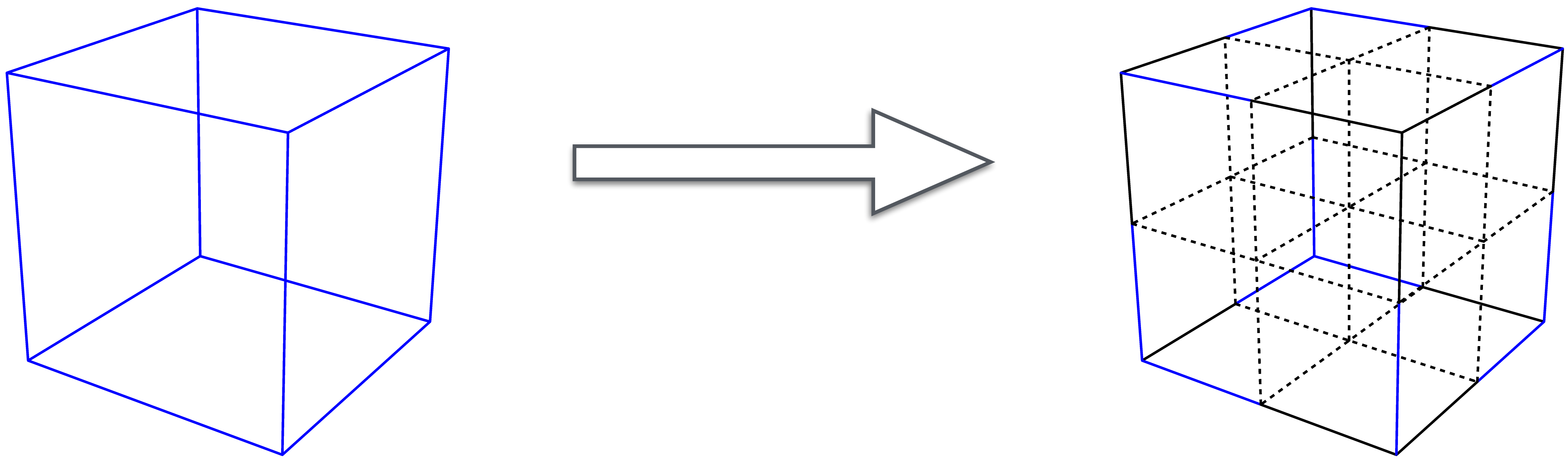
$$\tau^c_{\text{decorr}} + \tau^f_{\text{retherm}} \le \tau^f_{\text{decorr}}$$

Time scales



— more efficient use of computational resources
— greater statistical power due to *fully decorrelated streams*
— reduced critical slowing down; e.g., well sampled topology

Interpolation of gauge fields (à la 't Hooft)

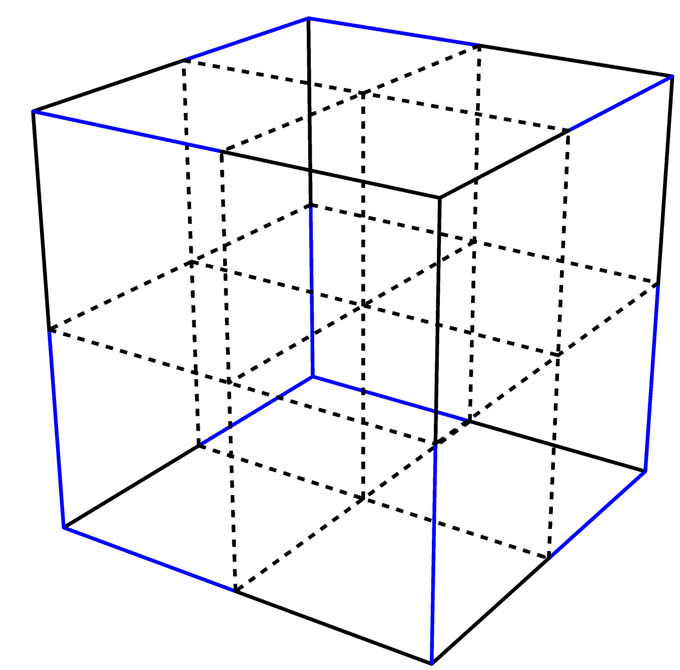


[1] Coarse lattice variables are transferred to the fine lattice

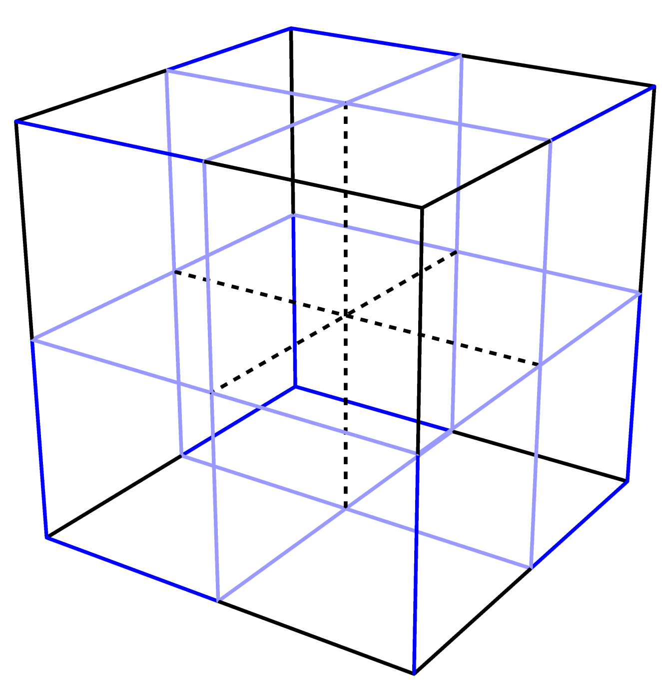
set to unity by a gauge choice

undefined bond variables
(set to unity)

Interpolation of gauge fields (à la ’t Hooft)



[2] Interior links are obtained by first minimization of action defined on 2x2 plaquettes

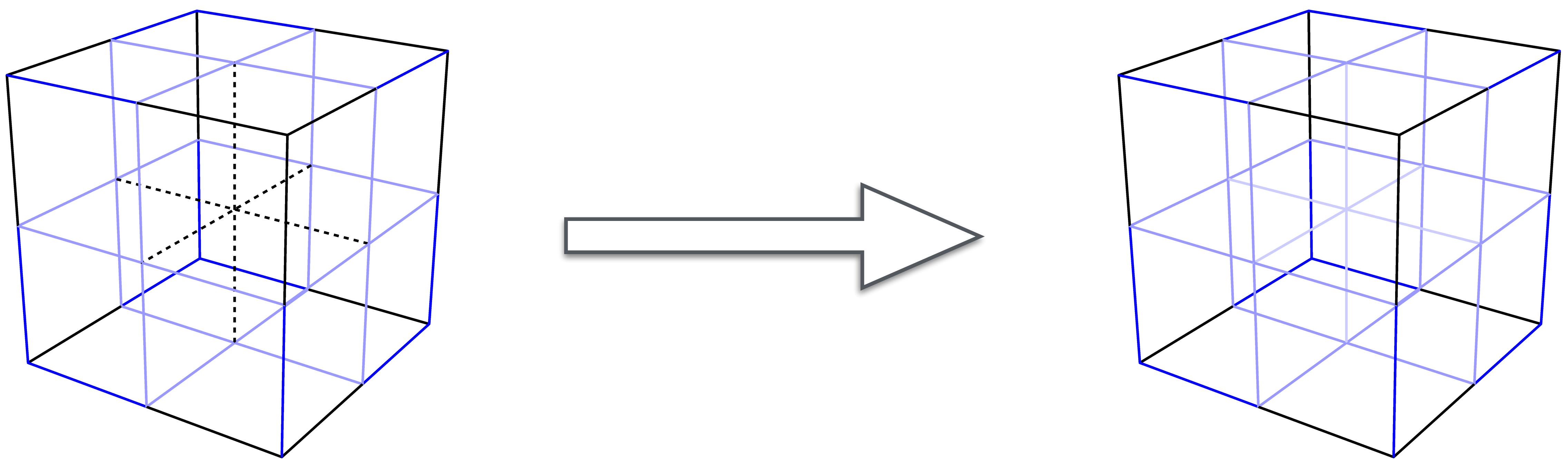


set to unity by a gauge choice

undefined bond variables (set to unity)

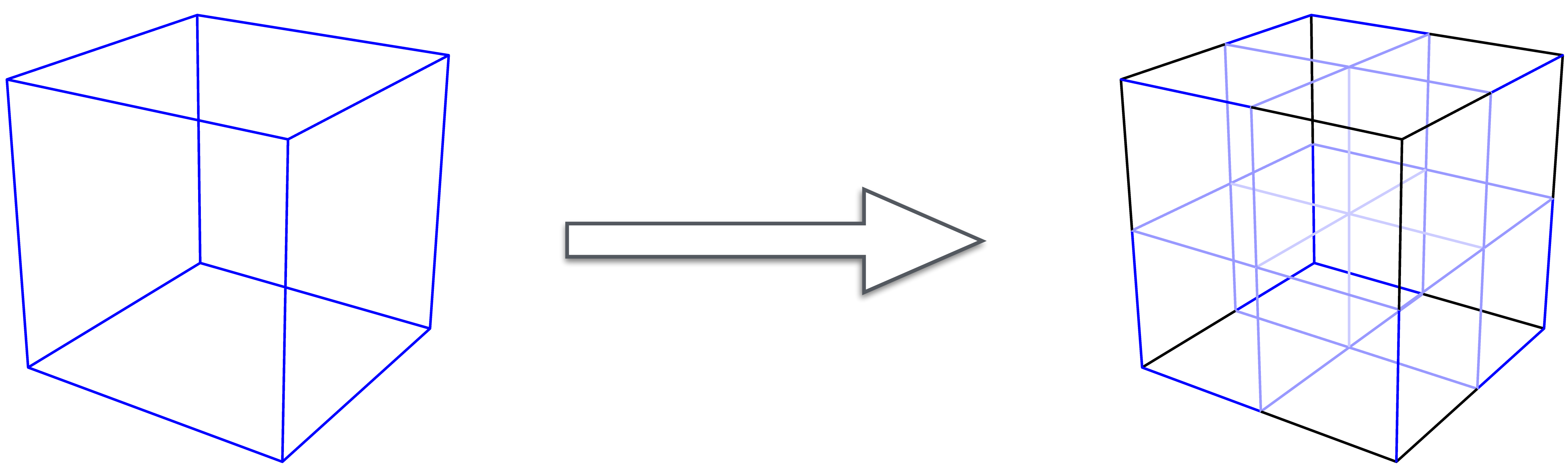
Interpolation of gauge fields (à la ’t Hooft)

[3] Minimization is repeated sequentially for interior cells



Properties of the interpolation

- Implementation is simple and efficient
- Can be performed locally
- Preserves long distance properties of coarse configuration
  - subset of even dimensional Wilson loops exactly
  - topological charge at sufficiently fine lattice spacing
  - discrete rotational invariance
- Breaks a subset of discrete translational symmetry
  - rapidly restored upon rethermalization

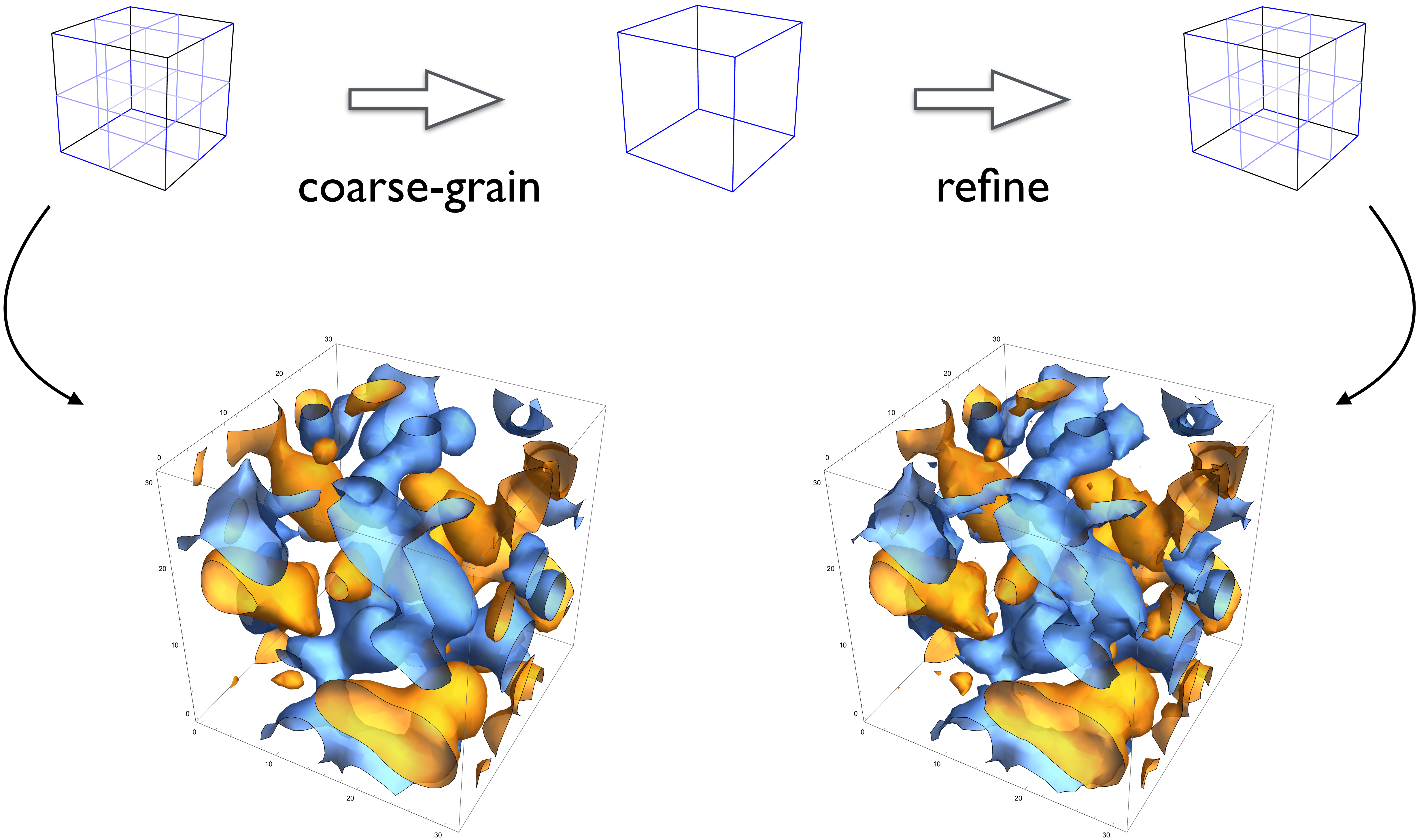


Numerical studies

Lattice	β	a [fm]	N
$12^3 \times 24$	5.626	0.1995(20) fm	385
$16^3 \times 36$	5.78	0.1423(5) fm	385
$24^3 \times 48$	5.96	0.0999(4) fm	185
$32^3 \times 72$	6.17	0.0710(3) fm	185

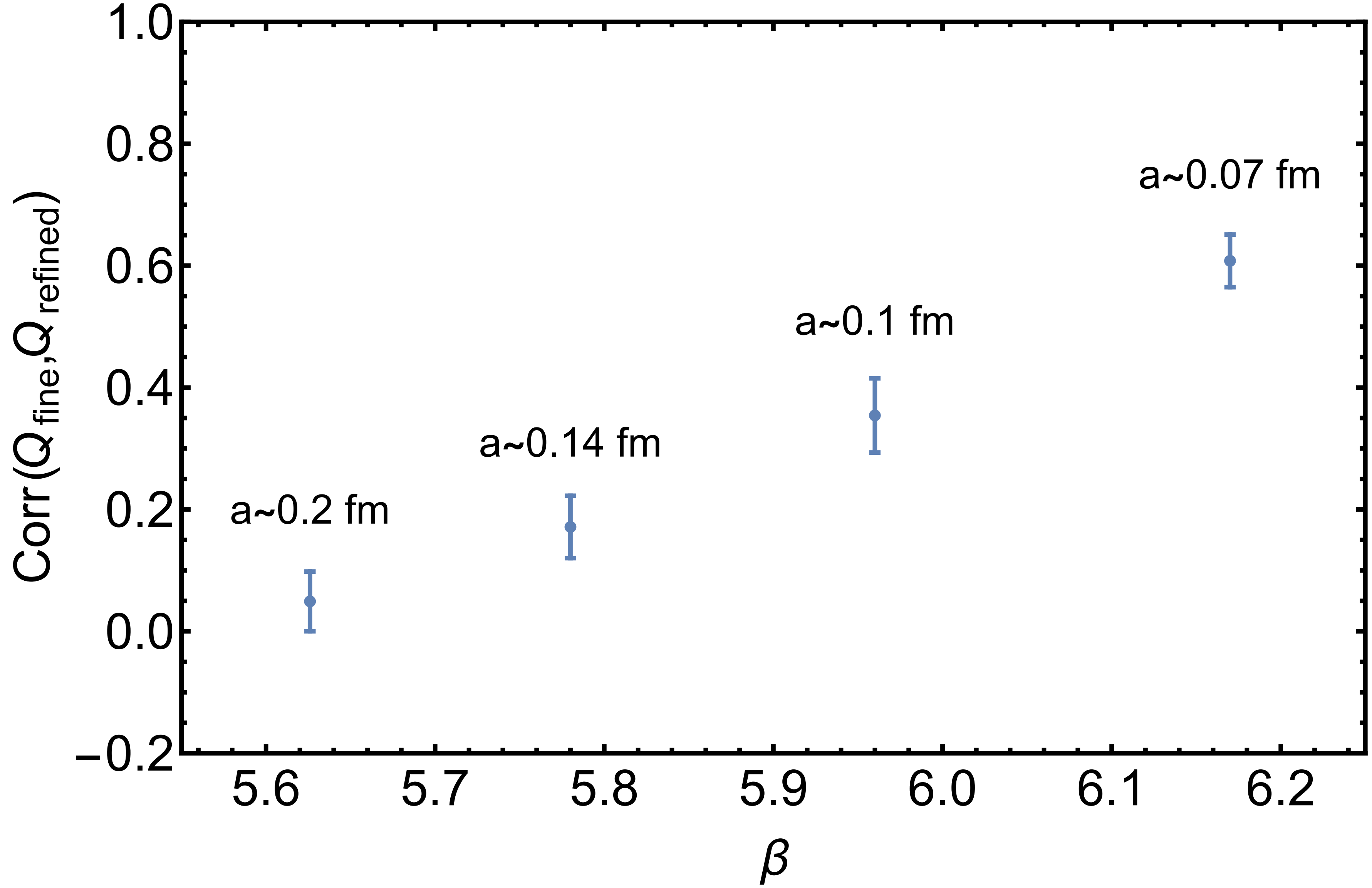
- Pure SU(3) gauge theory
- Two *pairs* of RG matched ensembles (plaquette action)
- All ensembles correspond to a fixed physical volumes ~ 2.3 fm
- Studied long-distance observables such as large Wilson loops and various quantities under "Wilson flow" (diffusion)
    - e.g., powers of the topological charge, action density

Interpolation — topological charge density



(single time slice)

Interpolation — topological charge

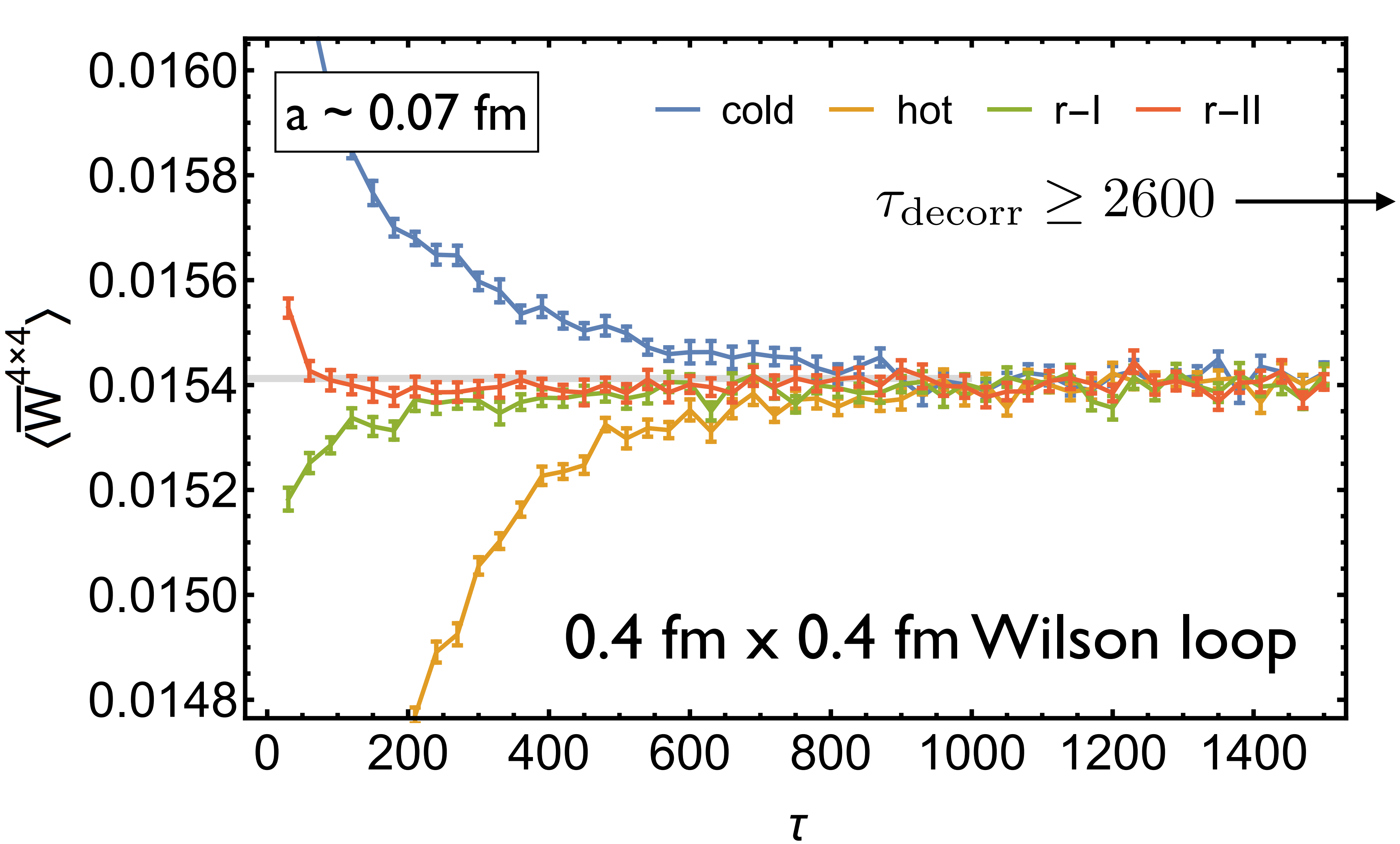


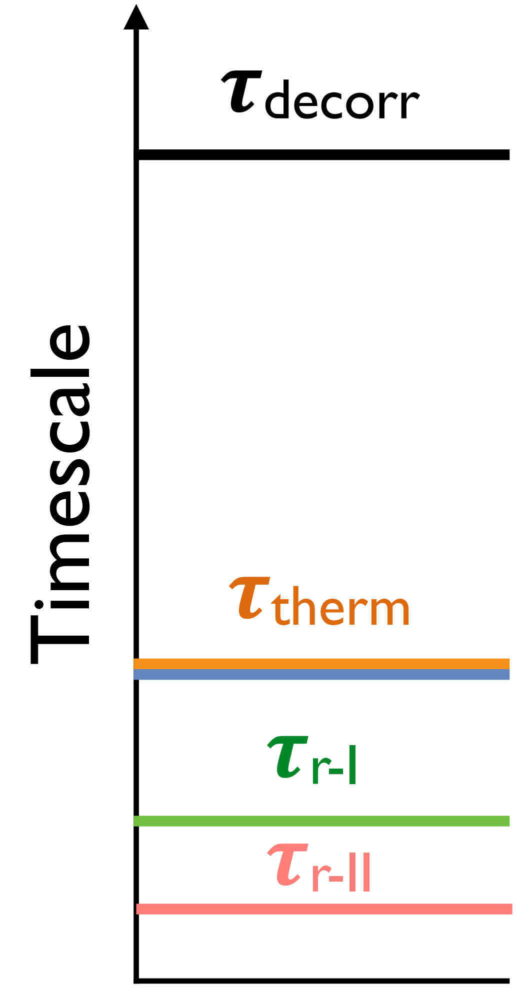
At sufficiently fine lattice spacing, topological charge of the coarse action is preserved configuration by configuration

Thermalization and rethermalization — HMC

- Ensembles of size N_s=24
- Thermalization times probed by long distance observables measured at various Wilson flow times: $\chi(t)$, $E(t)$
- Thermalization considered for four ensembles:
    - disordered (hot)
    - ordered (cold)
    - restriction followed by prolongation of fine lattices (r-I)
    - prolongation of an RG matched coarse ensemble, generated using a Wilson action (r-II)

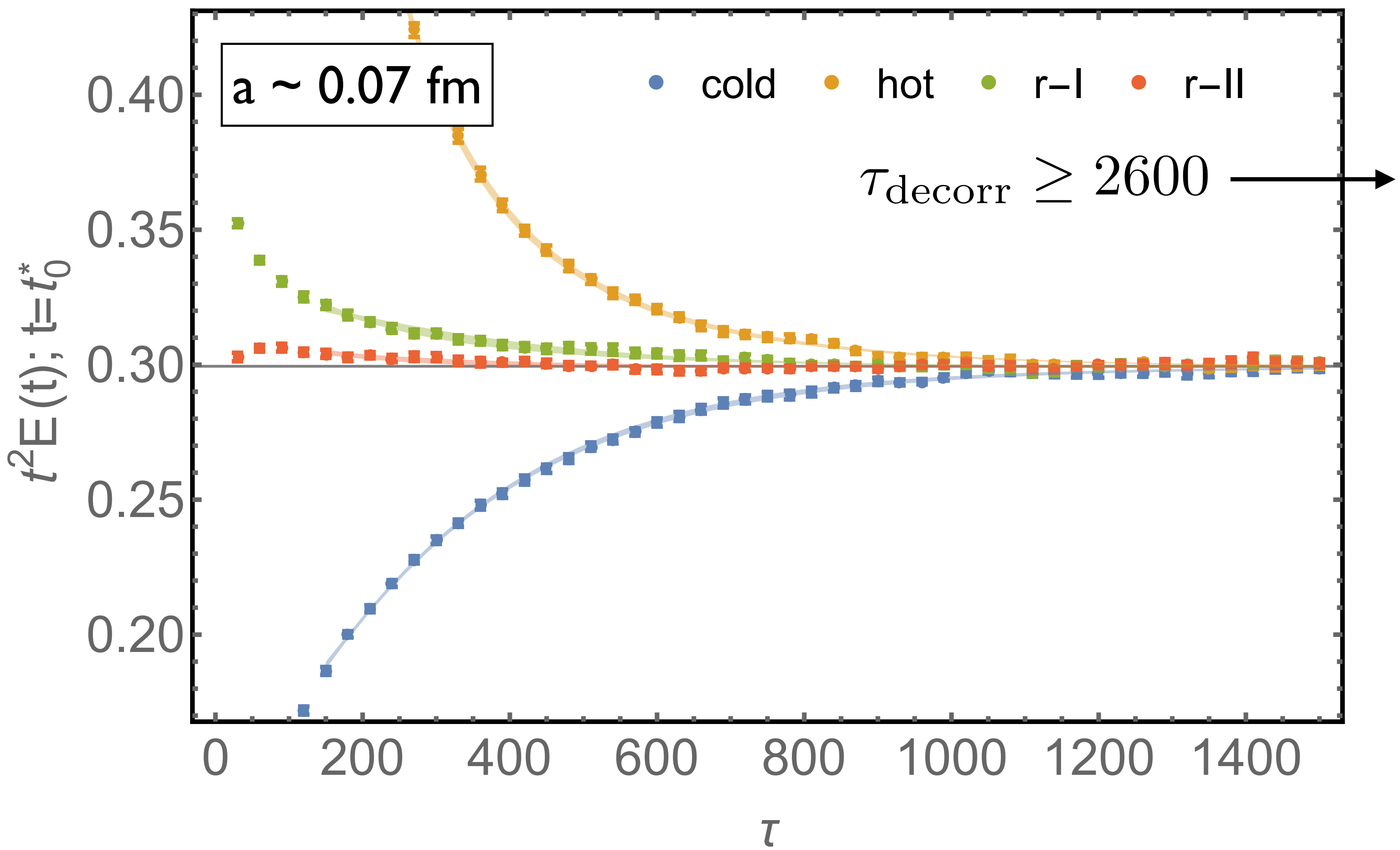
(Re)thermalization — Wilson loops





- Long-distance observables *rethermalize* on time scales shorter than
  - thermalization time for hot/cold starts (standard approach)
  - decorrelation time for fine evolution

(Re)thermalization — E(t)

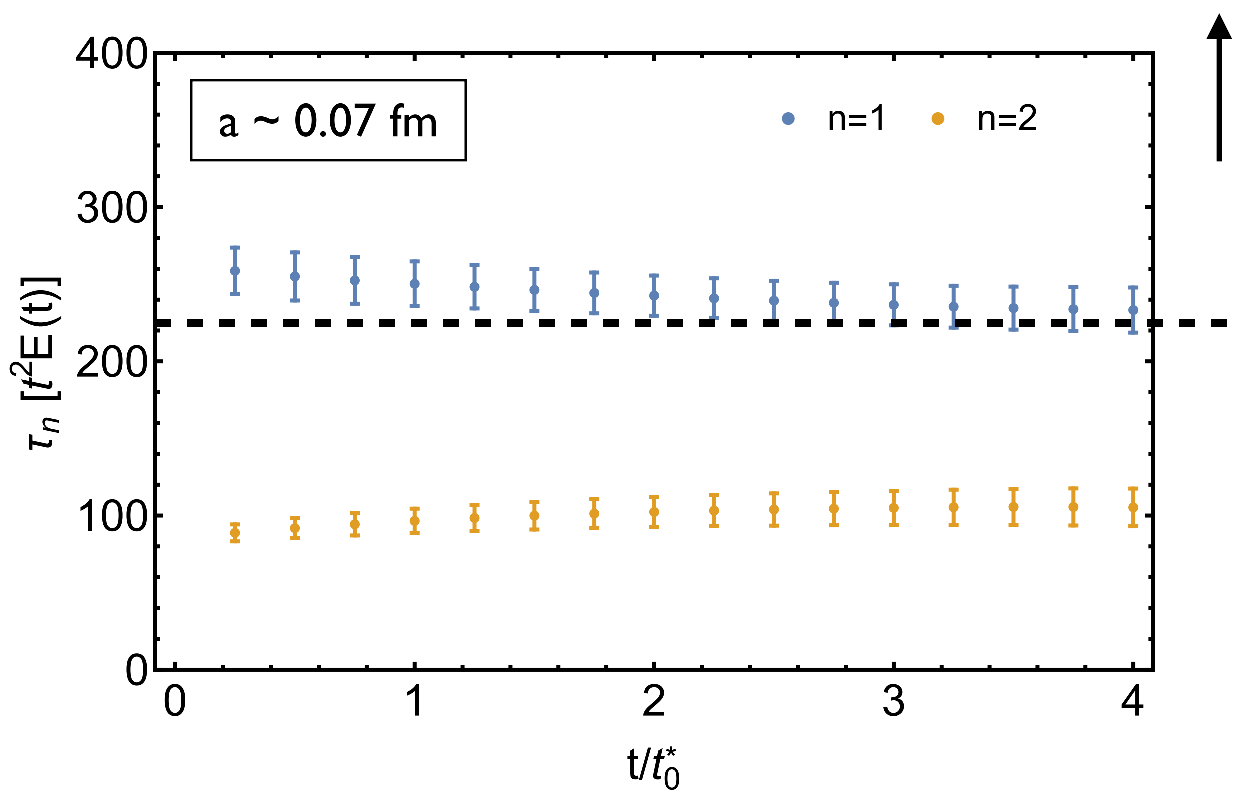


Least-squares fitting:

$$f^\alpha(\tau) = z_0 + z_1^\alpha e^{-\tau/\tau_1} + z_2^\alpha e^{-\tau/\tau_2}$$

$$z_0 = \langle \mathcal{O} \rangle \qquad z_n = \langle \mathcal{O} | \chi_n \rangle \langle \chi_n | \mathcal{P}_0 \rangle$$

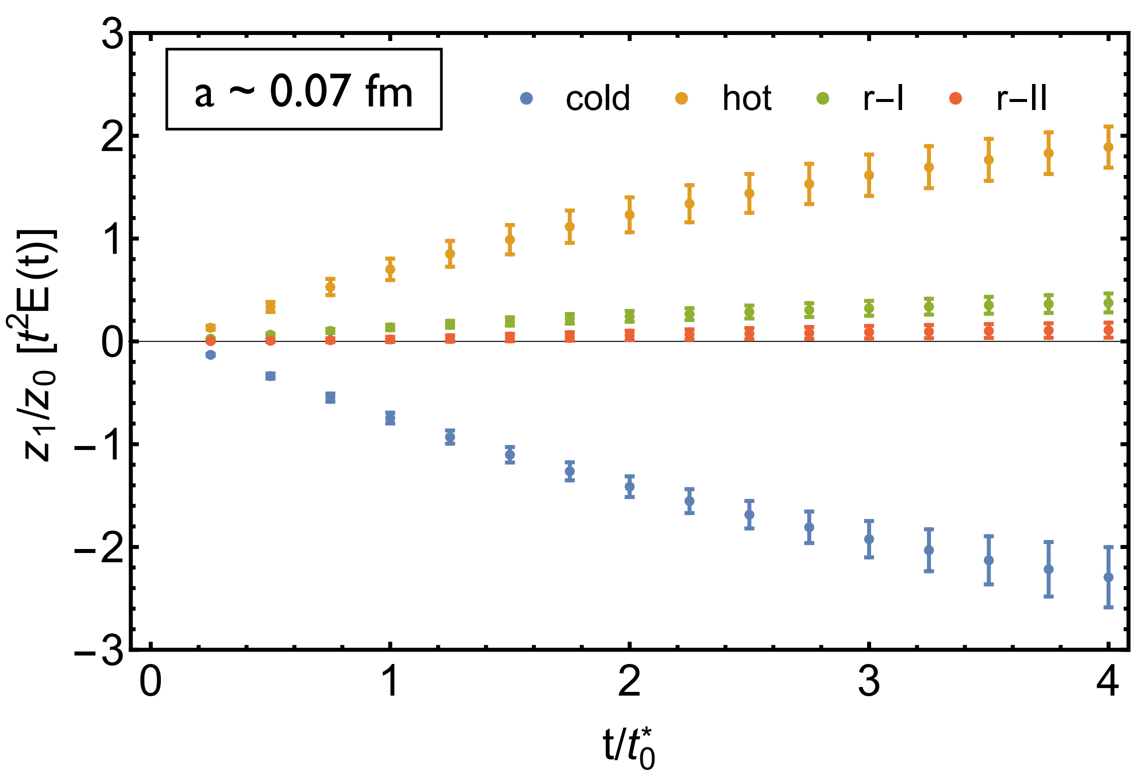
E(t) fit results as a function of flow time t/t_0

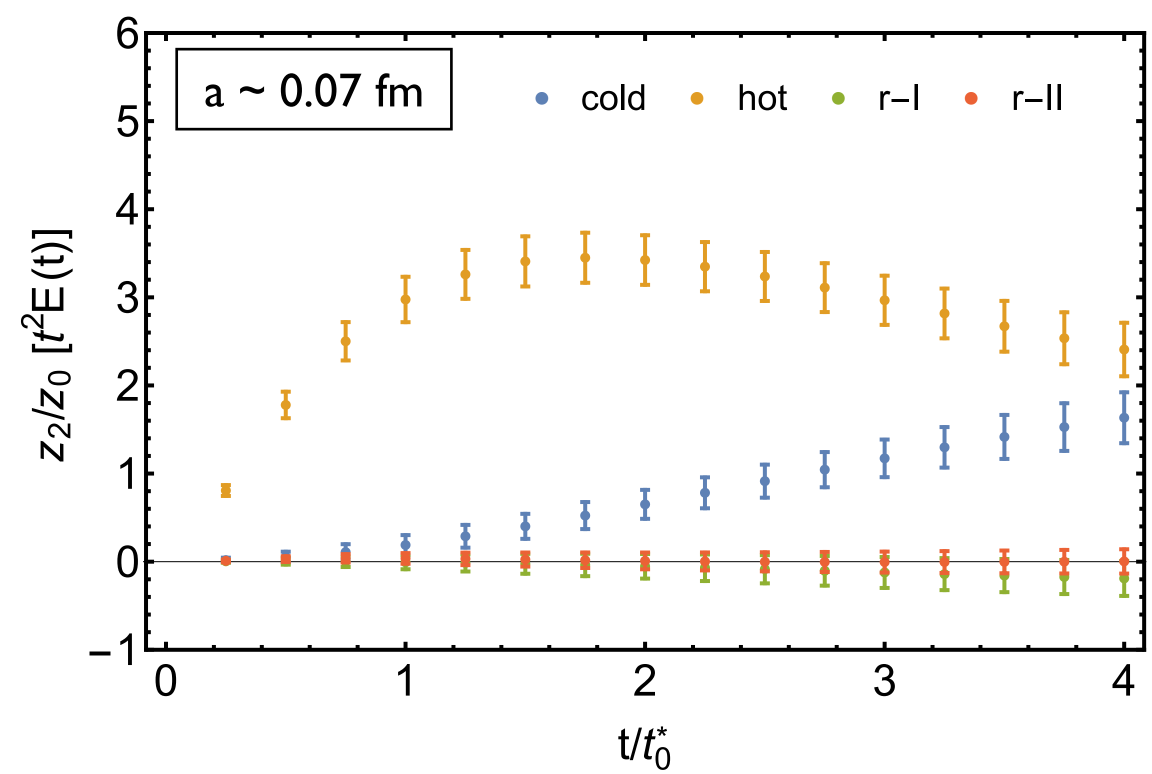


$\tau_{\mathrm{decorr}} \geq 2600$

τ_{int}

— *Able to remove at least the lowest two modes (n=1,2)*
— *Projection is independent of t_0*





Rapid thermalization

- Efficient multi-stream generation of uncorrelated gauge configurations
- Significantly reduces the problem of critical slowing down
    - enables numerical simulations at ultra-fine lattice spacings (a<0.05fm) with well-sampled topological charge
    - more efficient simulations expected for physical pion masses
- Alternatively, enables efficient numerical simulations at large volumes
- Approach successfully applied to Hybrid Monte Carlo simulations
    - next steps: inclusion of fermions