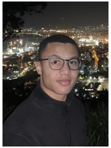
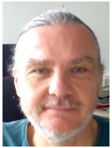
Parallel Tempered Metadynamics in full QCD

[2607.21575]

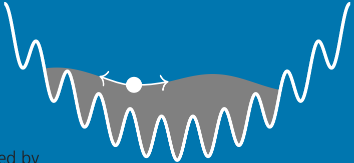
Gianluca Fuwa, Timo Eichhorn, Christian Hoelbling





Lattice 2026

27.07.2026



Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

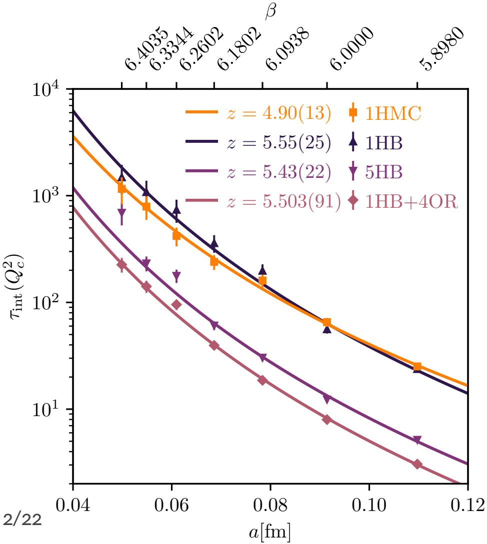
BERGISCHE UNIVERSITÄT WUPPERTAL

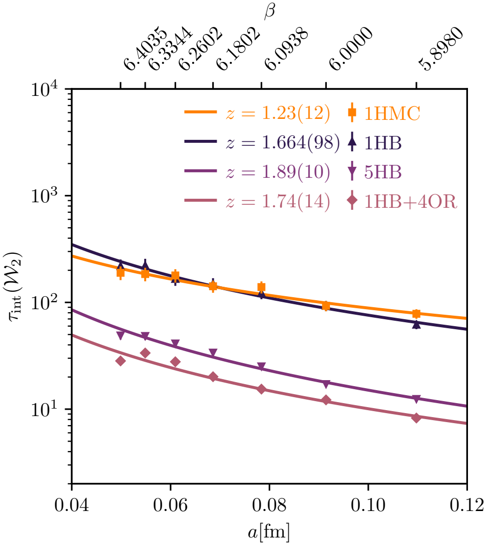
Cost of lattice QCD simulations

$$\text{Simulation cost} \propto \underbrace{V^{z_V}}_{\text{HMC}} \quad \underbrace{m_\pi^{-z_\pi}}_{\text{Solver}} \quad \underbrace{a^{-z_a}}_{\text{Autocorrelations}}$$

- $z_V \gtrsim 1$, depends on integrator, hard to imagine an algorithm where $z_V < 1$
  - HMC + second order integrator: $z_V = 1 + 1/4$
  - HMC + fourth order integrator: $z_V = 1 + 1/8$
- $z_\pi \sim 1$–2 after many algorithmic improvements, almost constant down to very small (below physical) quark masses
  - Hasenbusch mass preconditioning, domain decomposition, multiple pseudofermions
  - Multiple timescale integration
  - Multigrid solvers
  - Deflation
- $z_a \sim 2$ according to “naive” expectations
  $z_a \sim 5$–6+ in the presence of topological sectors $\Rightarrow$ ergodicty problems

Topological freezing - 4D SU(3) with Wilson action





Possible approaches to the problem

- Fixed topology simulations
- Master field simulations
- Modified boundary conditions
  - Open boundaries, P-periodic
  - Parallel tempering in boundary conditions
- Nested sampling
- Out-of-equilibrium algorithms
- Trivializing maps
- Machine learning (normalizing flows, diffusion models...)
- Multiscale thermalization
- Instanton updates
- **Metadynamics** (or more generally: **Biased sampling** methods)
- ...

Biased sampling

Introduce set of collective variables $\mathbf{s}(U)$ that capture relevant slow modes

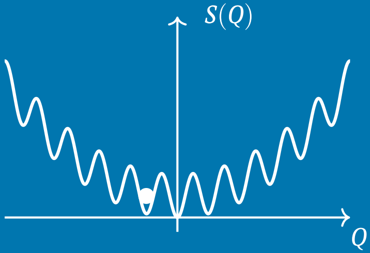
- **Physical distribution:** $p(U) \propto e^{-S(U)}$
- **Biased distribution:** $p_V(U) \propto e^{-S(U)-V(\mathbf{s}(U))}$
- **Target distribution (of CVs):** $p_{\text{tg}}(\mathbf{s}) = \,?$

  Possible choices:

  - Uniform: $p_{\text{tg}}(\mathbf{s}) \propto \text{const.}$
  - Well-tempered distribution: $p_{\text{tg}}(\mathbf{s}) \propto p(\mathbf{s})^{1/\gamma}, \ \gamma \in [1, \infty)$

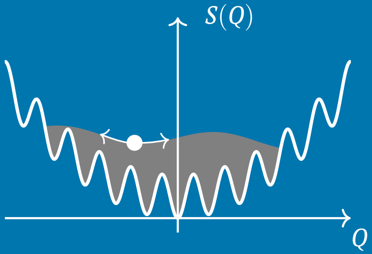
Remove barriers between sectors

Original action landscape



Stuck in sector (here $Q = -1$)

Modified action landscape



Able to move between sectors

Metadynamics in Lattice Gauge Theories

How can we construct appropriate potentials?

- Single collective variable Q_{CV}
  - Clover-based, stout smeared topological charge
  - Trade-off between
    - Staying in relevant regions of phase space (less smearing)
    - Having sufficient resolution (more smearing)
- Resulting modified action $S_t^M = S + V_t(Q)$

$\Rightarrow$ Reweighting necessary

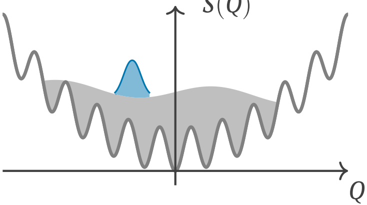
$\Rightarrow$ Stout force recursion (as with smeared fermions)

$$F_{\mathrm{CV}} = \frac{\partial V}{\partial Q_{\mathrm{CV}}} \frac{\partial Q_{\mathrm{CV}}}{\partial U^{(n)}} \frac{\partial U^{(n)}}{\partial U}$$

Construct time dependent potential from Gaussians

$$V_t(Q) = \sum_{t' \leq t} g(Q - Q(t'))$$

$$g(Q) = w \exp\left(-Q^2/(2\delta Q^2)\right)$$



Bias potential construction

Use multiple acceleration strategies to decrease needed trajectories to $\lesssim \mathcal{O}(1000)$:

1. Use prior knowledge/symmetries ($V(Q) = V(-Q)$)
2. Well-tempered Metadynamics (Gaussian weight decreases exponentially with $V_t(Q)$)
3. Multiple walkers $\rightarrow$ Buildup is trivially parallelizable with optimal weak scaling
4. Volume extrapolation of potentials at small to larger volumes via self-convolution of probability densities
5. Use CV values during the HMC trajectory (substep sampling/recycling HMC)

Metadynamics - Reweighting efficiency

Efficiency of reweighting?

- Effective sample size

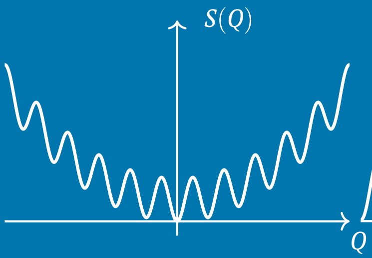
$$\mathrm{ESS} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2}$$

- Reweighting significantly reduces ESS down to $\mathcal{O}(10\,\%)$ or lower
- Two causes:
  - Oversampling of sectors with large $|Q|$
    $\Rightarrow$ Modify bias potential
  - Oversampling of configurations between sectors
    $\Rightarrow$ Combine with parallel tempering (PT-MetaD)

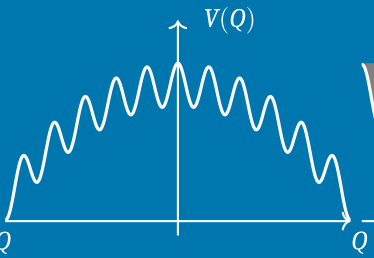
Metadynamics - bias potential modification

Suboptimal distribution overlap

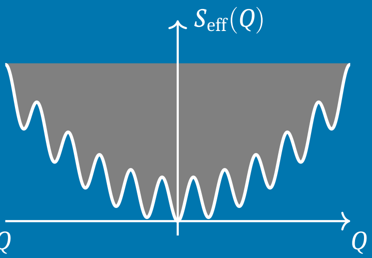
Original action landscape



Bias potential

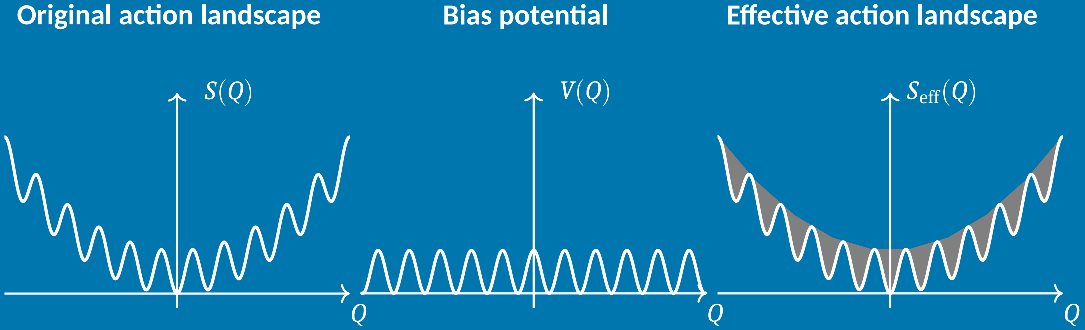


Effective action landscape



Metadynamics - bias potential modification

Improved distribution overlap



Parallel Tempered Metadynamics (PT-MetaD)

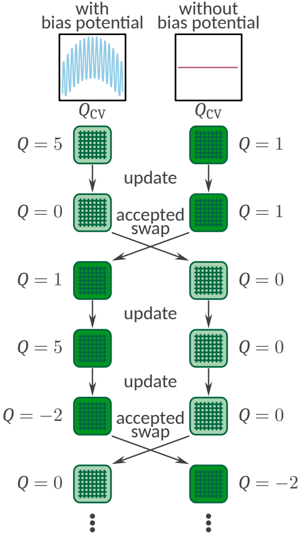
Address oversampling of configurations between sectors

- Two streams (inspired by [1706.04443]):
  - Stream 1 with bias potential $\Rightarrow$ tunneling
  - Stream 2 without bias potential $\Rightarrow$ measurements (no reweighting)
- Propose swaps like in standard parallel tempering
- Swaps **only depend on bias potential!**

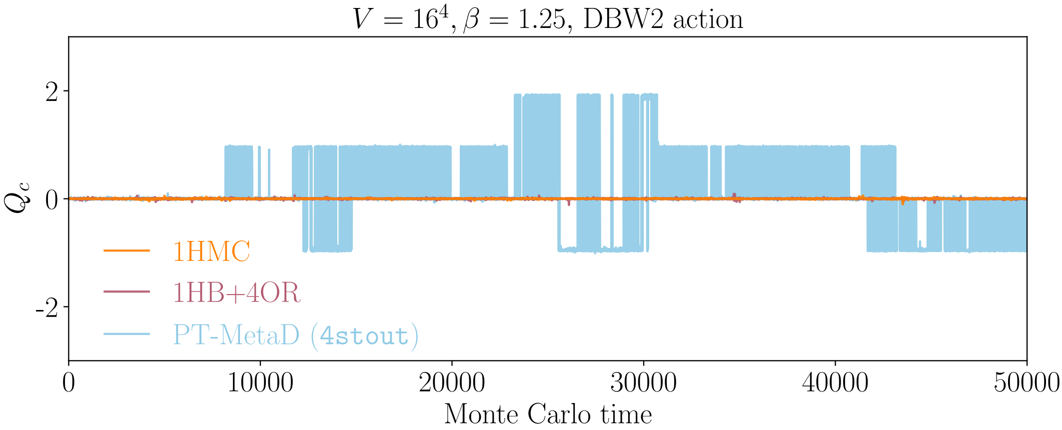
$$\begin{aligned}\Delta S_t^M &= \left[S_t^M(U_1) + S(U_2)\right] - \left[S_t^M(U_2) + S(U_1)\right] \\ &= V_t(Q_{\mathrm{CV},1}) - V_t(Q_{\mathrm{CV},2})\end{aligned}$$

$\Rightarrow$ No dirac equation solves required!

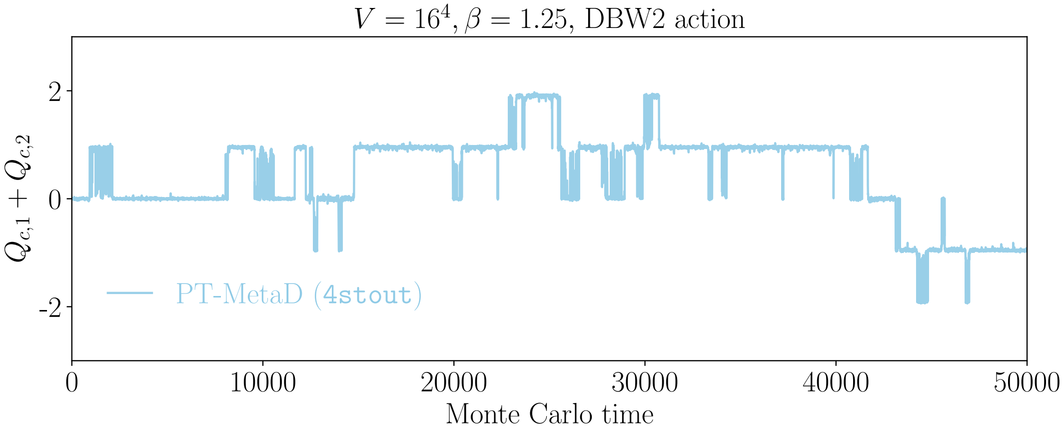
- Monitor τ_{int} of observables defined on product space



PT-MetaD - 4D SU(3)



PT-MetaD - 4D SU(3)

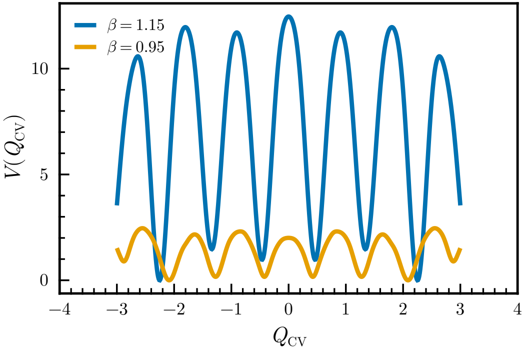


PT-Metad - From pure gauge to full QCD

- No conceptual difficulties (bias is a purely gluonic contribution)
- In some ways better suited to QCD simulations than pure gauge
  - HMC already required
  - Smearing already often used for fermions
  - $\Rightarrow$ Only little overhead from additional stout force recursion
- Using the aforementioned acceleration strategies makes bias buildup **very feasible in full QCD contexts!**

PT-MetaD - Full QCD

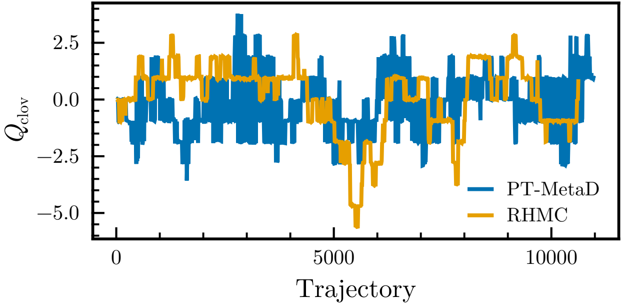
- Proof-of-concept in systems using the DBW2 gauge action
  topological freezing occurs at larger lattice spacing a
- $V = 24^4$
- $N_f = 2$
- $m = 0.02$
- Two different β-values:
  $\beta = 0.95 \rightarrow$ Frequent tunneling
  $\beta = 1.15 \rightarrow$ Effectively frozen
  $\Rightarrow m_q \sim 60-100$ MeV



$3000/4000$ trajectories (improvable!) with 4 walkers for $\beta = 0.95/1.15$ respectively + SSA

PT-MetaD - Full QCD

$\beta = 0.95$, $a = 0.0633(18)$ fm



$$\chi_{\mathrm{top}} V^{\mathrm{RHMC}} = 2.73(93)$$

$$\chi_{\mathrm{top}} V^{\mathrm{PT\text{-}MetaD}} = 2.19(35)$$

$$P^{\mathrm{RHMC}}_{\mathrm{smeared}} = 0.999\,501\,50(84)$$

$$P^{\mathrm{PT\text{-}MetaD}}_{\mathrm{smeared}} = 0.999\,501\,40(79)$$

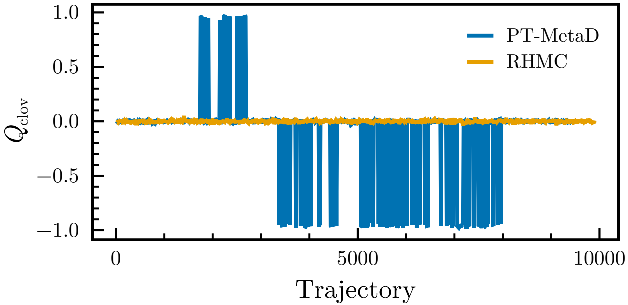
Efficiency for χ_{top} comparable to regular RHMC even without freezing
(including overhead from buildup and additional stream)

$$\tau^{\mathrm{RHMC}}_{\mathrm{int}}(Q^2) = 191(122)$$

$$\tau^{\mathrm{PT\text{-}MetaD}}_{\mathrm{int}}\left((Q_1 + Q_2)^2\right) = 120(67)$$

PT-MetaD - Full QCD

$\beta = 1.15$, $a = 0.0369(67)$ fm



$$\chi_{\mathrm{top}} V^{\mathrm{RHMC}} = 0.00(-)$$
$$\chi_{\mathrm{top}} V^{\mathrm{PT\text{-}MetaD}} = 0.127(33)$$
$$P^{\mathrm{RHMC}}_{\mathrm{smeared}} = 0.999\,720\,54(127)$$
$$P^{\mathrm{PT\text{-}MetaD}}_{\mathrm{smeared}} = 0.999\,734\,04(52)$$

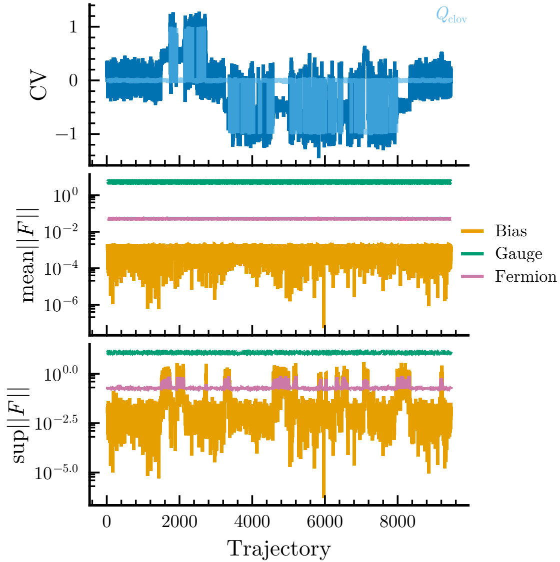
$\rightarrow$ Enables tunneling in completely frozen systems

$\sim$ 10σ difference in smeared Plaquette when RHMC is frozen!

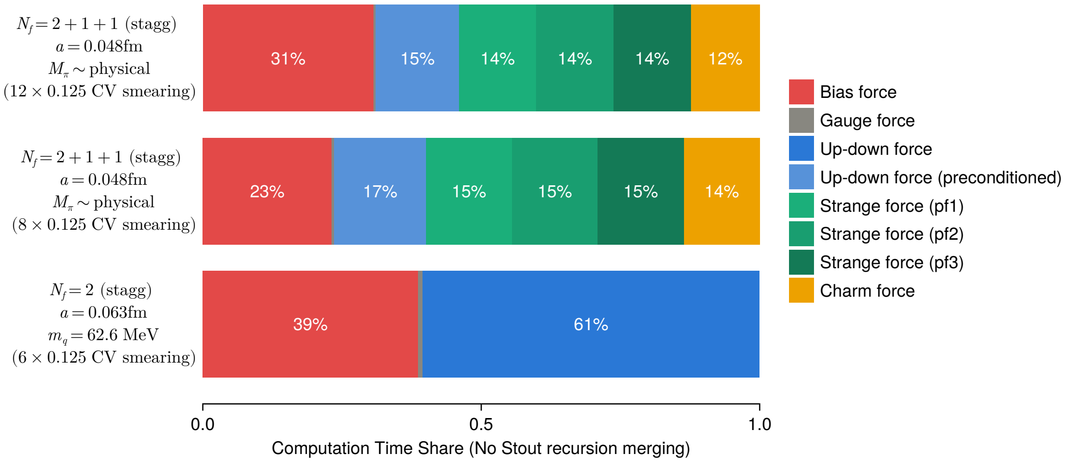
$$\tau^{\mathrm{RHMC}}_{\mathrm{int}}(Q^2) = \mathrm{N/A}$$
$$\tau^{\mathrm{PT\text{-}MetaD}}_{\mathrm{int}}\left((Q_1 + Q_2)^2\right) = 398(259)$$

PT-MetaD - Forces

- Bias force is stable and on same scale or lower as typical fermion forces
- More smearing increases the force caused by the bias
- Peaks when the system traverses inter-sector regions
- Seems to be highly concentrated in small sub-volumes (instantons?)



PT-MetaD - Overhead

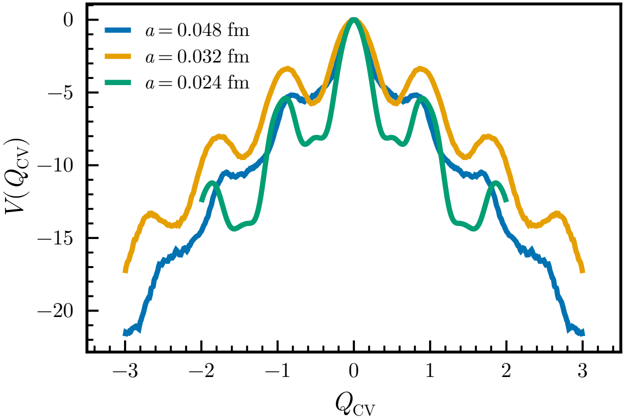


Studies in high T QCD

- **In progress: Determination of finite T topological susceptibility**
  - $N_f = 2+1+1$, `4stout` staggered
  - Physical pion masses (with $m_c/m_s = 11.85$)
  - LCP based on BMW ensembles see Table 6.1 in [F. Frech thesis]
  - CV: Q_c after 8×0.125 (2 coarsest lattices) or 12×0.125 (2 finest lattices) stout smearing steps

β	a [fm]	$L/a \times T/a$	am_s	m_s/m_l	CV smearing
4.1479	0.0483	28×8	0.019370	27.630	8×0.125
4.3513	0.0322	42×12	0.012335	27.630	8×0.125
4.4980	0.0242	56×16	0.009000	27.630	12×0.125
4.6144	0.0194	70×20	0.007059	27.630	12×0.125

Studies in high T QCD



- Bias potentials finished building for first three lattice spacings
- Reference and PT-MetaD ensembles currently being generated
- $a = 0.019$ fm bias still being built

Summary

- PT-MetaD successful in unfreezing topologically frozen full QCD systems
  - No reweighting required $\Rightarrow$ No reduction of effective sample size
  - Small relative overhead (with stout force recursion merging)
  - Improved scaling
    - With lattice spacing [2307.04742]
    - With volume (compared to standard parallel tempering in β)

  $\Rightarrow$ High T QCD runs in progress!
- Open questions:
  - Scaling behavior in dynamical simulations
  - Synergy with other approaches (only targets topological freezing, not critical slowing down in general)
- Dynamical/production code: [MetaQCD.jl] (actively maintained/developed)

Backup

Accelerating the bias potential thermalization

1. **Use prior knowledge/symmetries**
   - Charge conjugation symmetry of action ($Q \leftrightarrow -Q$)
     In practice: Whenever we update the bias potential at some Q_{CV} we also update it at $-Q_{\mathrm{CV}}$
     $\Rightarrow$ Approximately $2\times$ speed-up
   - Expect a (near) Gaussian distribution of Q with variance $\langle Q^2 \rangle \sim V$
     $\Rightarrow$ Extrapolate from smaller volumes?
   - Barrier shapes seem to be periodic
     $\Rightarrow$ Extrapolate from inner sectors?

Accelerating the bias potential thermalization

2. Well-tempered Metadynamics [Barducci' 08]

- Standard Metadynamics

$$V_{t+1}(Q) = V_t(Q) + w \exp\left(-\frac{(Q_t - Q)^2}{2\sigma^2}\right)$$

- Well-tempered Metadynamics

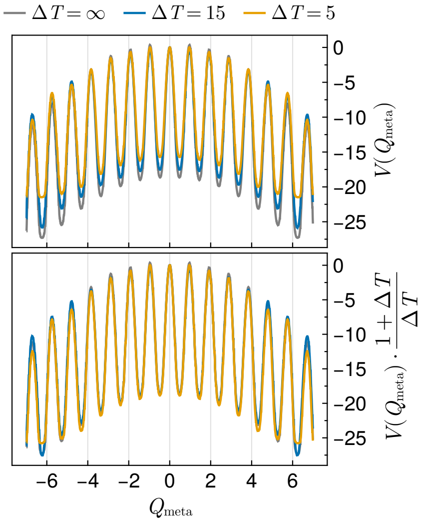
$$V_{t+1}(Q) = V_t(Q) + \exp\left(-\frac{V_t(Q)}{\Delta T}\right) w \exp\left(-\frac{(Q_t - Q)^2}{2\sigma^2}\right)$$

Tunable parameter ΔT:

- $\Delta T \to 0$: No Metadynamics
- $\Delta T \to \infty$: Standard Metadynamics

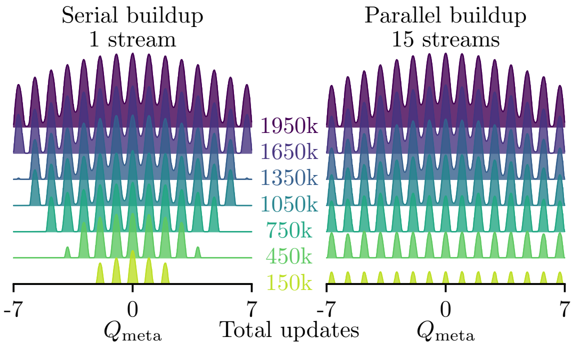
Accelerating the bias potential thermalization

2. **Well-tempered Metadynamics** [Barducci' 08]
   - Able to choose larger starting Gaussian weights while maintaining smoothness in the end
   - Quirk: Bias does not converge to $-S(Q)$ but $-\frac{\Delta T}{1+\Delta T} S(Q)$
     $\Rightarrow$ Too small ΔT may reintroduce ergodicity problems, as the barriers are not sufficiently canceled out



Accelerating the bias potential thermalization

3. **Multiple walkers** [Raiteri' 06]
   - Run N_{walkers} simulation streams in parallel, all working on the same potential
   - Trivially parallelizable (optimal weak scaling)
   - Minimal communication between processes required (e.g., a single `MPI.Allgather` call per iteration) $\Rightarrow$ speed-up of factor $\sim N_{\text{walkers}}$

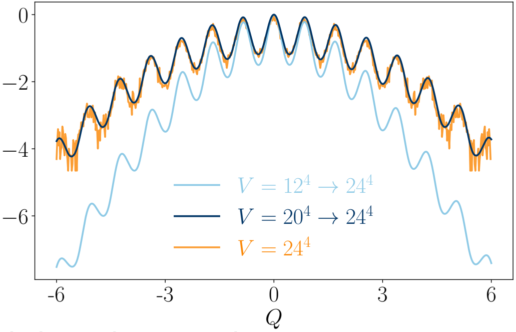


Accelerating the bias potential thermalization

4. Volume extrapolation

- Extrapolate bias potential from small to large volumes
- For $\Lambda = \Lambda_1 \cup \Lambda_2$ (assuming independent domains Λ_1, Λ_2):

$$p_\Lambda(q) = \int \mathrm{d}q' p_{\Lambda_1}(q')p_{\Lambda_2}(q-q')$$

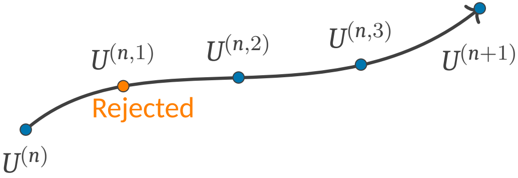


$\Rightarrow$ Convolution of probability density with itself

Finite size effects are maintained through the convolution procedure (seems to mostly affect sector weights, not barrier heights and positions!)

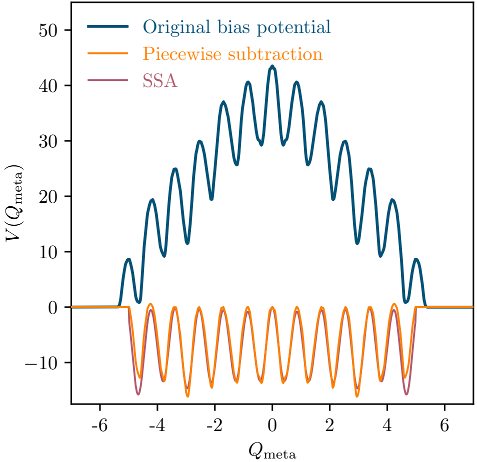
Accelerating the bias potential thermalization

5. **Substep sampling/Recycling HMC [1511.06925]**
   - Use configurations during HMC trajectory for measurements
   - **MD evolution unchanged**
   - Must perform accept-reject step against the initial configuration for each intermediate trajectory before measurement
     — for the bias build-up this requirement is not as stringent, as long as the acceptance rate is relatively high
   - Helpful in building up bias potential, possibly useful for variance reduction?

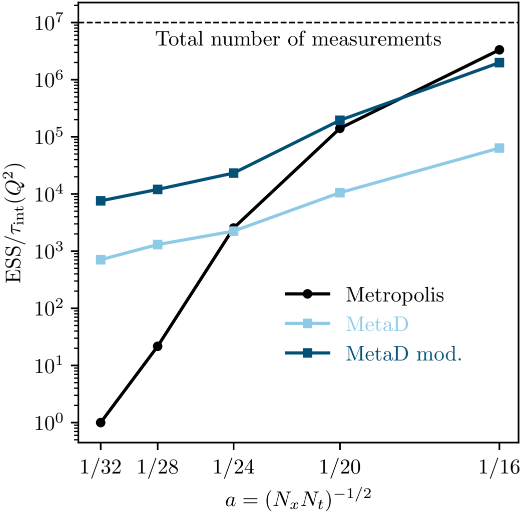


Metadynamics - bias potential modification

Oversampling of large $|Q|$

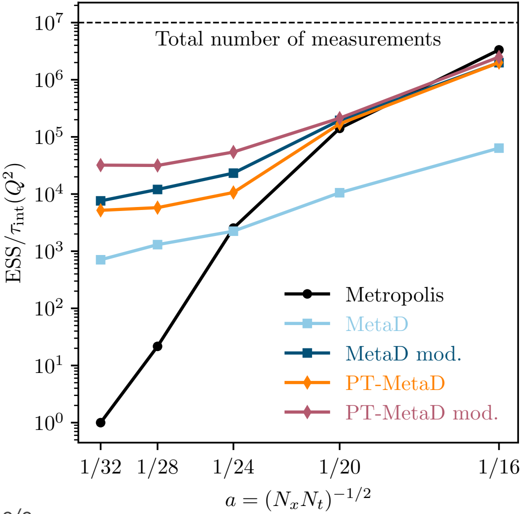


Scaling plot for 2D U(1)



Only remove barriers between sectors

Scaling of PT-MetaD in 2D U(1)



Putting everything together

- Further improvement compared to standard MetaD
- Best results when combined with modification of bias potential

$\Rightarrow$ Almost 5 orders of magnitude improvement for finest lattice spacing compared to conventional algorithm

- Scaling of all (PT-)MetaD variations much milder compared to conventional algorithms