

Comparison of topology changing update algorithms

Timo Eichhorn* and Christian Hoelbling

Department of Physics, University of Wuppertal, Gaußstraße 20, D-42119 Wuppertal, Germany

E-mail: timo.eichhorn@protonmail.com, hch@uni-wuppertal.de

In modern lattice simulations, conventional update algorithms do not allow for tunneling between topological sectors at fine lattice spacings. We compare the viability of multiple less commonly used algorithms (metadynamics, instanton updates, and multiscale thermalization) with respect to proper sampling of all topological sectors in the Schwinger model. We briefly comment on the prospects of applying these methods to 4-dimensional SU(3) simulations.

*The 38th International Symposium on Lattice Field Theory, LATTICE2021 26th-30th July, 2021
Zoom/Gather@Massachusetts Institute of Technology*

*Speaker

1. Introduction

An important property that has to be taken into consideration when analyzing data generated by Markov chain Monte Carlo methods is the correlation of subsequent states in the Markov chain. A particularly problematic form of these autocorrelations in the context of physical simulations is critical slowing down: Near critical points, many autocorrelation times associated with observables measured on the generated configurations diverge.

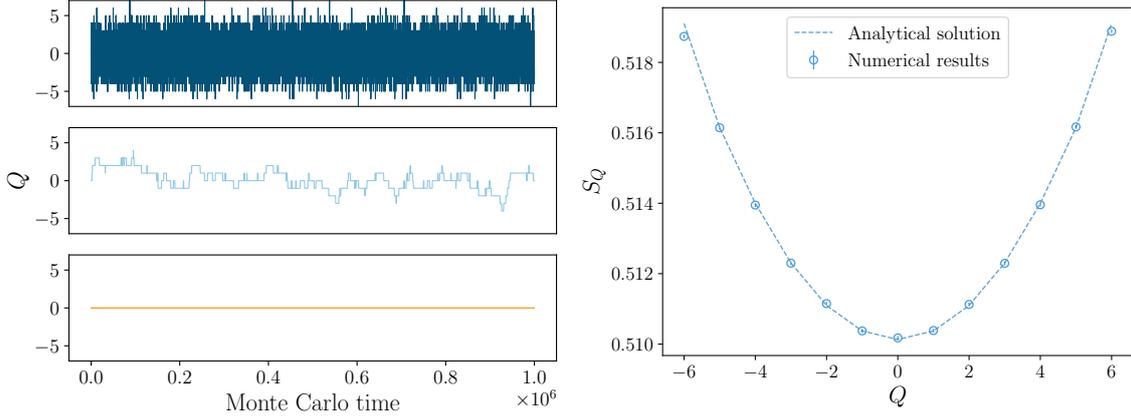


Figure 1: Left: Time series of the topological charge in the Schwinger model for three simulations on a line of constant physics with the parameters $V = \{16^2, 24^2, 32^2\}$ and $\beta = \{3.2, 7.2, 12.8\}$ (from top to bottom). **Right:** Expectation values of the gauge action in the Schwinger model in different topological sectors for a 32^2 lattice at $\beta = 12.8$ obtained through simulation and an analytical solution [1, 2, 3, 4].

In lattice QCD and lattice gauge theory in general, one prominent example of critical slowing down is related to the topological charge when taking the continuum limit [5, 6, 7]. The autocorrelation time associated to the topological charge increases with the inverse lattice spacing, and after a certain point, the Markov chain may even be completely frozen in a topological sector for all practically achievable simulation timescales. The left plot in figure 1 illustrates this phenomenon in the Schwinger model for three different lattice spacings. Whereas transitions between different topological sectors happen frequently on the coarsest lattice, the Markov chain is confined to the $Q = 0$ sector throughout the entire simulation for the finest lattice spacing. This not only affects the estimation of statistical errors, but also has direct consequences on the expectation values of observables - even those not directly related to the topological charge. For instance, the average action increases for sectors with a larger absolute value of the topological charge, as shown in the right plot in figure 1. Therefore, it is more accurate to describe the time evolution of the Markov chain in terms of modes that couple differently to observables, instead of simply associating a single autocorrelation time to each observable. In this case, the topological charge is strongly coupled to slow modes, whereas the action is not coupled as strongly to those modes.

While topological freezing appears in both the Schwinger model and four-dimensional $SU(3)$ (with and without dynamical fermions), the results presented here are mostly limited to the former. We study three previously suggested algorithmic approaches to alleviate the problem: Metadynamics [8, 9], instanton updates [10, 11, 12], and multiscale thermalization [13, 14, 15].

2. Metadynamics

Metadynamics is based upon introducing a set of collective variables (CVs) to measure which parts of the configuration space have already been visited by the Markov chain. The space spanned by the CVs can thus be understood as a low-dimensional projection of the complete configuration space. Using the CVs, a history-dependent metapotential V_{meta} is built up, which has been shown to be an estimate of the negative free energy/action of the system [16]. During each accept-reject step of the update algorithm, the metapotential is added to the action of the configuration. Therefore, once the metapotential has reached equilibrium, the probability distribution will be approximately constant as a function of the CVs.

Following the approach of [9] we use a linearly interpolating histogram instead of the originally proposed sum of Gaussians, which allows us to keep the memory footprint constant in time. The metapotential is defined by four parameters: Q_{min} and Q_{max} specify the interval $[Q_{\text{min}}, Q_{\text{max}})$ in which the metapotential is updated, δQ is the bin width, and w is a weight that controls how fast the metapotential is updated. As CV we use the imaginary part of all plaquettes, which yields a non-integer definition of the topological charge:

$$Q_{\text{meta}} = \frac{1}{2\pi} \text{Im} \left(\sum_{\mathbf{n} \in V(\Lambda)} P_{tx}(\mathbf{n}) \right). \quad (1)$$

$P_{tx}(\mathbf{n})$ denotes the plaquette in t - x -direction at the point \mathbf{n} , and $V(\Lambda)$ refers to the set of all lattice points. At Monte Carlo time t with a configuration of topological charge $Q_{\text{meta}}(t)$, the metapotential is updated in the following way:

$$\begin{aligned} V_{\text{meta},i}(t) &= V_{\text{meta},i}(t-1) + w \left(1 - \frac{Q_{\text{meta}}(t) - (Q_{\text{min}} + i \cdot \delta Q)}{\delta Q} \right), \\ V_{\text{meta},i+1}(t) &= V_{\text{meta},i+1}(t-1) + w \frac{Q_{\text{meta}}(t) - (Q_{\text{min}} + i \cdot \delta Q)}{\delta Q}. \end{aligned} \quad (2)$$

The index i is defined as

$$i = \lfloor \frac{Q_{\text{meta}}(t) - Q_{\text{min}}}{\delta Q} \rfloor. \quad (3)$$

In the end, the observables need to be reweighted back to the original probability distribution:

$$\langle O \rangle = \frac{\sum_t O(t) e^{\bar{V}_{\text{meta}}(Q_{\text{meta}}(t))}}{\sum_t e^{\bar{V}_{\text{meta}}(Q_{\text{meta}}(t))}}. \quad (4)$$

Here, \bar{V}_{meta} refers to an estimate of the negative action, which can for instance be obtained via the time average of the metapotential (after excluding equilibration) or simply $V_{\text{meta}}(t)$ for large t .

In practice, it is advisable to introduce an additional penalty potential to prevent the Markov chain from getting stuck in regions outside the interval $(Q_{\text{min}}, Q_{\text{max}})$ where the metapotential is not updated. The exact form of the penalty potential is not important - in our case we used a penalty potential that scales quadratically with the distance and depends on three parameters k , $Q_{\text{thr, min}}$, and $Q_{\text{thr, max}}$:

$$V_{\text{pen}} = \begin{cases} k \cdot \min\{(Q - Q_{\text{thr, min}})^2, (Q - Q_{\text{thr, max}})^2\} & Q \notin (Q_{\text{thr, min}}, Q_{\text{thr, max}}), \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

3. Instanton updates

Instanton updates, which resemble another recently proposed update algorithm [17], generate new configurations by multiplying the previous configuration with an instanton of charge $Q_j = \pm 1$, where the sign is chosen randomly with equal probability (for recent applications see [18, 19]). In $D < 4$ dimensions, instantons are maximally delocalized, i.e., the field strength is constant over the entire lattice. In the temporal gauge on a $N_x \times N_t$ lattice with periodic boundaries, instantons of charge Q_j can be constructed by setting the link variables to [20]:

$$\begin{aligned} U_t^I(Q_j; t, x) &= \exp\left(-2\pi i x \frac{Q_j}{N_x N_t}\right), \\ U_x^I(Q_j; t, x) &= \exp\left(2\pi i t \frac{Q_j}{N_t} \delta_{x, N_x}\right). \end{aligned} \quad (6)$$

The update procedure then assigns new values by setting

$$U_\mu(t, x) \rightarrow U'_\mu(t, x) = U_\mu(t, x) U_\mu^I(Q_j; t, x) \quad \forall U_\mu(t, x) \in E(\Lambda), \quad (7)$$

where $E(\Lambda)$ denotes the set of all link variables. Detailed balance can be achieved by adding a Metropolis accept-reject step; the inclusion of dynamical fermion would then necessitate the additional calculation of the determinant ratio of the new and old configuration. Finally, since this update procedure by itself is not ergodic, it must be combined with an ergodic update algorithm, e.g., a regular Metropolis link update.

4. Multiscale thermalization

Multiscale thermalization is a multigrid and real-space renormalization group (RG) inspired approach to accelerate the thermalization of configurations and to enable the generation of an ensemble of uncorrelated configurations more efficiently than through conventional approaches. Initially, configurations can be generated on a coarse lattice where the thermalization and autocorrelation times are shorter, before they are fine-grained onto a finer lattice. Since the fine-graining procedure is generally not an exact RG transformation, it is necessary to introduce rethermalization updates on the fine lattice. Even if the distribution of configurations is not exactly preserved by the fine-graining of the coarse ensemble, the rethermalization can in theory correct any deviations. Therefore the multiscale thermalization strategy is more efficient than conventional thermalization provided that the sum of the integrated autocorrelation time on the coarse lattice τ_{int}^c and the rethermalization time τ_r is shorter than the integrated autocorrelation time on the fine lattice τ_{int}^f (neglecting the initial thermalization times on both lattices).

Of course, this procedure is not limited to only two different lattice spacings, and the generalization to multiple levels with different lattice spacings is straightforward as long as the action can be matched between the different levels. Here, we only consider the simplest case of a single fine-graining transformation from a coarse lattice with spacing $2a$ to a fine lattice with spacing a . The fine-graining transformation used here attempts to interpolate the fields by first assigning values to those fine lattice links that can be matched to the coarse lattice links:

$$U_\mu^f(\mathbf{n}) = U_\mu^c(\mathbf{n}/2), \quad U_\mu^f(\mathbf{n} + \hat{\mu}) = \mathbb{1} \quad \forall \mathbf{n} \in V(\Lambda) : \sum_\mu (n_\mu \bmod 2) = 0. \quad (8)$$

Here $\hat{\mu}$ denotes the orthonormal basis vector in μ -direction, and n_μ is the μ -component of the vector \mathbf{n} . Evidently, this transformation leaves all observables defined on the lattice corresponding to the coarse lattice invariant. Afterwards, the remaining links on the fine lattice are set to unity and smeared to reduce the occurring short range fluctuations of the field. Finally, a number of rethermalization updates are performed using conventional update algorithms.

5. Results

Figure 2 and table 1 show the expectation values of the action, the topological charge, and the topological susceptibility obtained using the different approaches and combinations thereof. With the exception of the two metadynamics approaches, all update schemes include overrelaxation updates. Additionally, the results are compared with analytical results [1, 2, 3, 4].

Clearly, the standard Metropolis link update is unable to correctly sample different topological sectors at this lattice spacing, and the action as well as the topological susceptibility are underestimated, since all generated configurations were in the $Q = 0$ sector. On the other hand, all three approaches presented in the previous sections are able to unfreeze the topological charge and enable tunneling between the different topological sectors. However, both the metadynamics and the multiscale thermalization approach require several parameters to be tuned, whereas the instanton update can be applied directly.

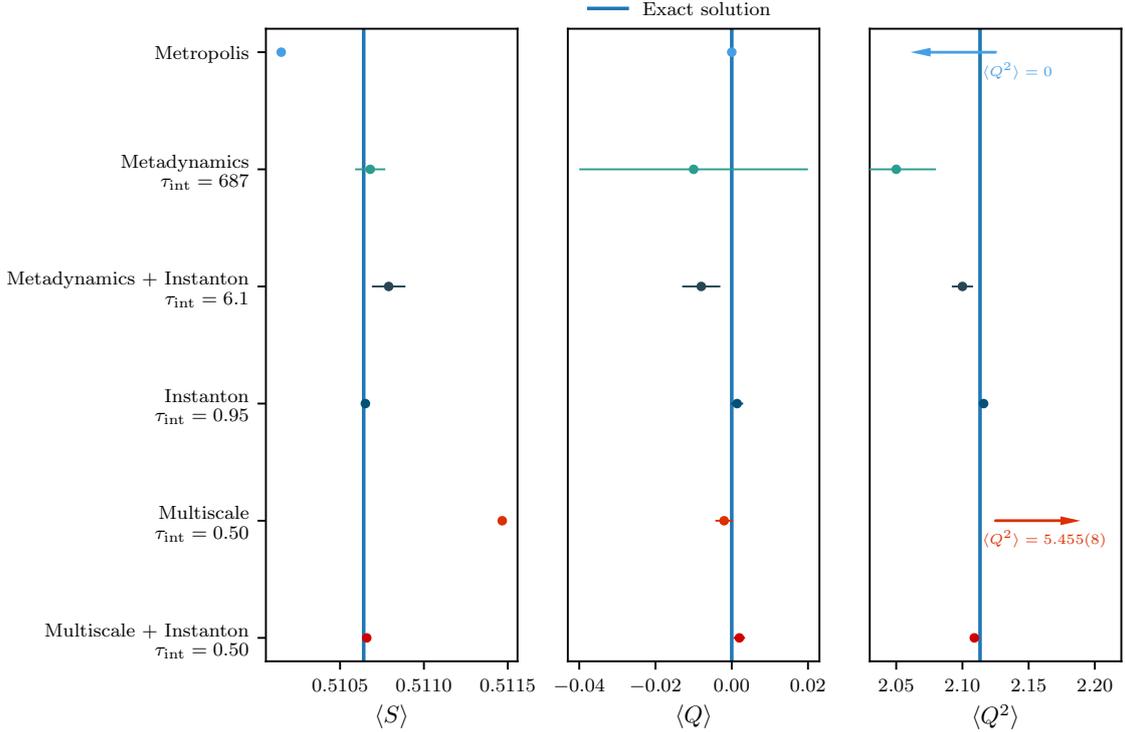


Figure 2: Comparison of expectation values obtained using the different approaches and combinations thereof. All simulations were performed on a 32^2 lattice at $\beta = 12.8$, where conventional update algorithms are unable to sample different topological sectors properly. Additionally, analytical results [1, 2, 3, 4] are indicated by the blue line.

Approach	$\langle S \rangle$	$\langle Q \rangle$	$\langle Q^2 \rangle$	τ_{int}
Analytical solution	0.510 641	0	2.113 304	–
Metropolis	0.510 149(23)	0	0	–
Metadynamics	0.510 68(9)	−0.01(3)	2.05(3)	687
Metadynamics + Instanton	0.510 79(10)	−0.008(5)	2.100(8)	6.1
Instanton	0.510 651(23)	0.0014(16)	2.116(3)	0.95
Multiscale	0.511 467(23)	−0.0020(23)	5.425(8)	0.50
Multiscale + Instanton	0.510 659(22)	0.0020(15)	2.109(3)	0.50

Table 1: Comparison of expectation values obtained using the different approaches and combinations thereof, as well as an analytical solution. The results presented here are the same as in [figure 2](#).

We found that the best results with metadynamics were achieved by first building up the metapotential, and then simulating on a static potential in a separate run. This allows us to estimate the statistical errors in the usual way (for instance with block bootstrap resampling). The range of the metapotential defined by $Q_{\text{thr, min}}$ and $Q_{\text{thr, max}}$ requires some fine-tuning however. If the range is chosen too small, not all relevant topological sectors are properly sampled, but if the range is chosen too large, the number of effectively relevant configurations decreases, since the contribution of sectors with higher absolute values of topological charge to the expectation values are exponentially suppressed. Therefore, a lower bound for the range should be estimated by first simulating on a coarser, physically equivalent lattice where the topological sectors are still well-sampled. The metapotentials for a 32^2 system at $\beta = 12.8$ obtained after 2×10^6 update sweeps using two different update schemes, namely multi-hit Metropolis updates with and without instanton updates, are shown in [figure 3](#).

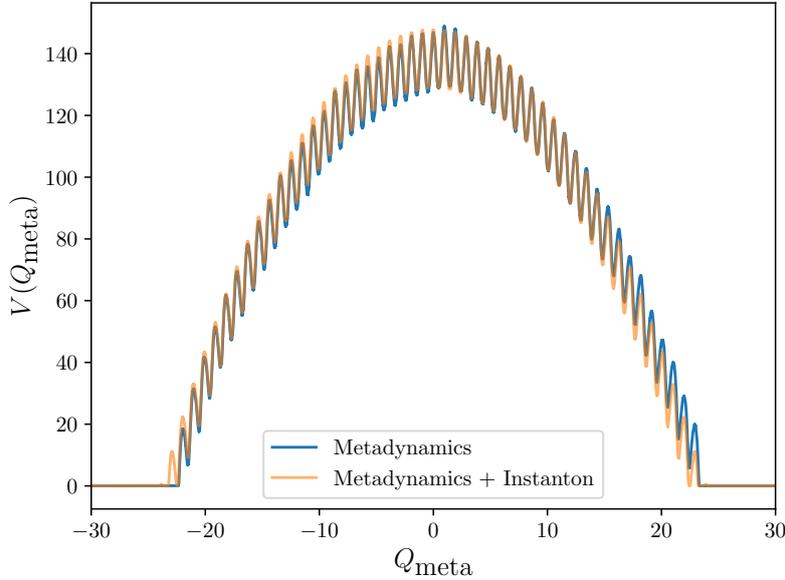


Figure 3: Comparison of the metapotentials in the range $[-25, 25)$ obtained using either multi-hit Metropolis updates or a combination of multi-hit Metropolis and instanton updates for a 32^2 lattice at $\beta = 12.8$.

Both metapotentials clearly show the action barriers between topological sectors, as well as the overall increasing average action for larger absolute values of Q_{meta} . Compared with using only multi-hit Metropolis updates, the potential obtained using combined multi-hit Metropolis and instanton updates is more symmetrical around the origin.¹

For the results in figure 2 and table 1, the metapotential parameters were chosen as follows:

$$Q_{\text{min}} = -10, Q_{\text{max}} = 10, \delta Q = 0.001, w = 0.0001. \quad (9)$$

The penalty potential parameters were set to the following values:

$$k = 1000, Q_{\text{thr, min}} = -8.0, Q_{\text{thr, max}} = 8.0. \quad (10)$$

The metapotentials were first generated using 2×10^6 update sweeps. Afterwards, the expectation values were calculated from 1×10^6 configurations separated by 10 update sweeps using a static metapotential (obtained using the final values of the metapotentials from the previous runs). All observables are compatible with the analytical results within ~ 2 standard deviations. However, the inclusion of instanton updates decreased the autocorrelation time by about two orders of magnitude and also reduced the statistical uncertainties of topological observables.

The application of the instanton update is straightforward, as no parameters have to be tuned. As shown in figure 4, the acceptance rate is only weakly dependent on the lattice spacing and lies between 0.72 and 0.75 for all parameters examined here.

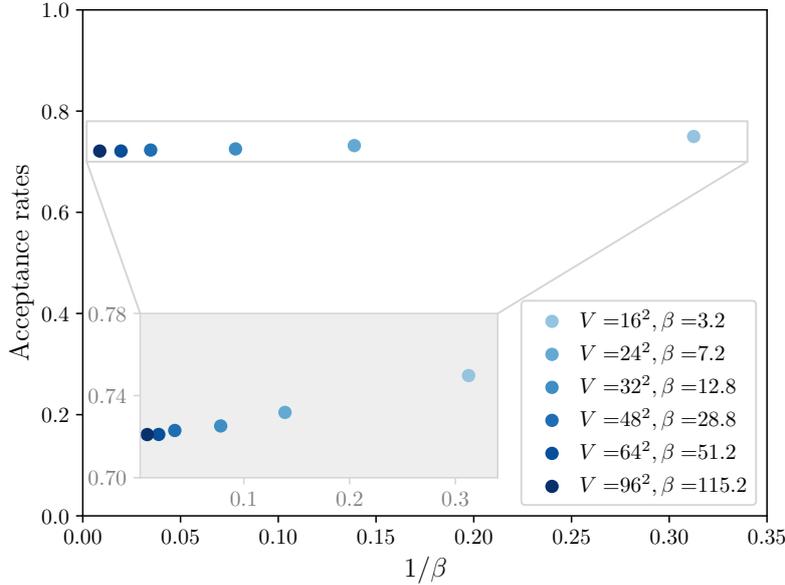


Figure 4: Acceptance rates of the instanton update in dependence of the gauge coupling. For all parameters examined here, the acceptance rate is nearly constant and lies between 0.72 and 0.75.

The expectation values shown in figure 2 and table 1 were calculated from 1×10^6 configurations separated by 10 update sweeps. All expectation values are compatible with the analytical solution within one standard deviation, and consecutive configurations are nearly uncorrelated.

¹Of course, the metapotential can be symmetrized around $Q_{\text{meta}} = 0$ [21] at the cost of the observable $\langle Q \rangle$.

For the multiscale thermalization, we performed 1×10^8 updates on the coarse lattice. Every 100th configuration was fine-grained, and an additional 500 rethermalization updates were performed before measuring observables. This approach lead to correlation times slightly shorter than the instanton update, but in our case it did not sample the probability distribution of the fine lattice correctly. An important crosscheck is given by the correlation of the topological charge measured on the coarse lattice and on the fine lattice after fine-graining and rethermalization as seen in figure 5.

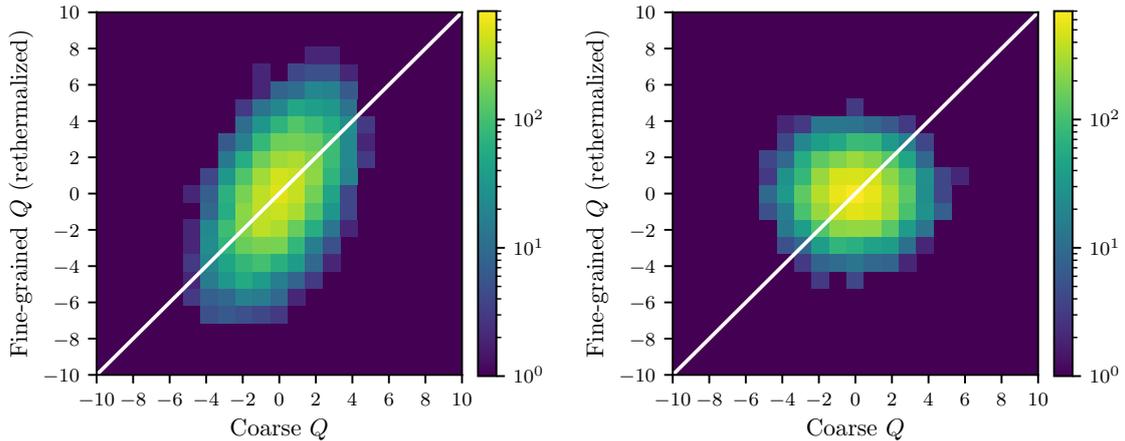


Figure 5: Charge correlation of configurations on the coarse lattice and configurations on the fine lattice after fine-graining and 100 rethermalization updates. The heatmap on the left side shows runs where the rethermalization was performed using multi-hit Metropolis and overrelaxation updates, whereas the heatmap on the right side shows runs where the rethermalization updates were performed using a combination of multi-hit Metropolis, overrelaxation, and instanton updates. The white lines are drawn to guide the eye.

While the charges are clearly correlated, the charge is not exactly preserved, and instead spread out over a larger range on the fine lattice. In theory, this can be corrected by a longer rethermalization time, but since the topological charge is effectively frozen on the fine lattice, the required number of rethermalization updates makes this strategy unfeasible. The situation can be improved by including instanton updates. Despite the correlation between the two charges being less pronounced than before, the instanton update allows for tunneling between topological sectors and thereby corrects some of the deviations from the correct probability distribution during rethermalization. With this combined approach, all observables are compatible with the analytical solution within less than two standard deviations. However, it should be noted that runs with only 100 instead of 500 rethermalization updates overestimated the action by almost five standard deviations.

6. Conclusion

In these proceedings, we examined three algorithmic approaches towards alleviating the problem of topological freezing in the Schwinger model. The first approach - metadynamics - leads to higher autocorrelation times than the other approaches and requires some parameters to be tuned, but it allows one to control which parts of the configuration space are sampled. Instanton updates

were found to produce the results most compatible with an analytical solution out of all update schemes studied here, and they do not require any parameters to be tuned. Finally, we found that the multiscale thermalization approach was unable to correctly sample the probability distribution of the fine lattice on its own for the parameters studied here. While that indicates this approach might not be suitable to properly sample all topological sectors at fine lattice spacings, it is still useful to reduce the thermalization and autocorrelation times of non-topological observables. We would also like to mention that a machine learning based approach to efficiently sample topological sectors in the Schwinger model has been presented at this conference [22].

The extension of all three approaches to 4-dimensional SU(3) is straightforward, but does not always lead to similar levels of improvement as in the Schwinger model. Metadynamics or methods similar in spirit have been applied to high temperature QCD [23] and pure SU(3) gauge theory [24]. The instanton update suffers from high action penalties, which effectively means that the acceptance is close to zero. Initial estimates show that the action penalty is of order ~ 3000 for a 8^4 lattice at $\beta = 6.0$, and of order ~ 12000 for a 16^4 lattice at $\beta = 6.0$ with Wilson gauge action. A possible remedy to this problem could be the use of gradient flow [25]: By first flowing to positive flow time, applying the instanton update, and then flowing back in a symmetric way to respect detailed balance, the acceptance rates could possibly be enhanced. Finally, multiscale thermalization has already been successfully applied to pure SU(3) gauge theory and two-color QCD [13, 14]. Interestingly, recent studies suggest that rethermalization updates may not be necessary for very fine lattices [15], which stands in contrast to our results in the Schwinger model.

References

- [1] T.G. Kovacs, E.T. Tomboulis and Z. Schram, *Topology on the lattice: 2-d Yang-Mills theories with a theta term*, *Nucl. Phys. B* **454** (1995) 45 [[hep-th/9505005](#)].
- [2] S. Elser, *The Local bosonic algorithm applied to the massive Schwinger model*, other thesis, 3, 2001, [[hep-lat/0103035](#)].
- [3] C. Bonati and P. Rossi, *Topological susceptibility of two-dimensional $U(N)$ gauge theories*, *Phys. Rev. D* **99** (2019) 054503 [[1901.09830](#)].
- [4] C. Bonati and P. Rossi, *Topological effects in continuum two-dimensional $U(N)$ gauge theories*, *Phys. Rev. D* **100** (2019) 054502 [[1908.07476](#)].
- [5] B. Alles, G. Boyd, M. D’Elia, A. Di Giacomo and E. Vicari, *Hybrid Monte Carlo and topological modes of full QCD*, *Phys. Lett. B* **389** (1996) 107 [[hep-lat/9607049](#)].
- [6] L. Del Debbio, H. Panagopoulos and E. Vicari, *theta dependence of $SU(N)$ gauge theories*, *JHEP* **08** (2002) 044 [[hep-th/0204125](#)].
- [7] S. Schaefer, R. Sommer and F. Virota, *Investigating the critical slowing down of QCD simulations*, *PoS LAT2009* (2009) 032 [[0910.1465](#)].
- [8] A. Laio and M. Parrinello, *Escaping free-energy minima*, *Proceedings of the National Academy of Sciences* **99** (2002) 12562–12566.

- [9] A. Laio, G. Martinelli and F. Sanfilippo, *Metadynamics surfing on topology barriers: the CP^{N-1} case*, *JHEP* **07** (2016) 089 [1508.07270].
- [10] F. Fucito and S. Solomon, *DOES STANDARD MONTE CARLO GIVE JUSTICE TO INSTANTONS?*, *Phys. Lett. B* **134** (1984) 230.
- [11] H. Dilger, *Topological zero modes in Monte Carlo simulations*, *Int. J. Mod. Phys. C* **6** (1995) 123 [hep-lat/9408017].
- [12] S. Durr, *Physics of η' with rooted staggered quarks*, *Phys. Rev. D* **85** (2012) 114503 [1203.2560].
- [13] M.G. Endres, R.C. Brower, W. Detmold, K. Orginos and A.V. Pochinsky, *Multiscale Monte Carlo equilibration: Pure Yang-Mills theory*, *Phys. Rev. D* **92** (2015) 114516 [1510.04675].
- [14] W. Detmold and M.G. Endres, *Multiscale Monte Carlo equilibration: Two-color QCD with two fermion flavors*, *Phys. Rev. D* **94** (2016) 114502 [1605.09650].
- [15] W. Detmold and M.G. Endres, *Scaling properties of multiscale equilibration*, *Phys. Rev. D* **97** (2018) 074507 [1801.06132].
- [16] G. Bussi, A. Laio and M. Parrinello, *Equilibrium free energies from nonequilibrium metadynamics*, *Phys. Rev. Lett.* **96** (2006) 090601.
- [17] D. Albande, P. Hernández, A. Ramos and F. Romero-López, *Topological sampling through windings*, *Eur. Phys. J. C* **81** (2021) 873 [2106.14234].
- [18] S. Durr and J.H. Weber, *Minimally doubled fermions and topology in 2D*, in *38th International Symposium on Lattice Field Theory*, 8, 2021 [2108.11766].
- [19] F.J. Frech, *Hadronic vacuum polarization from step scaling in the Schwinger model*, in *38th International Symposium on Lattice Field Theory*, vol. LAT2021, p. 218, 2021.
- [20] J. Smit and J.C. Vink, *Remnants of the Index Theorem on the Lattice*, *Nucl. Phys. B* **286** (1987) 485.
- [21] D.B. Leinweber, A.G. Williams, J.-b. Zhang and F.X. Lee, *Topological charge barrier in the Markov chain of QCD*, *Phys. Lett. B* **585** (2004) 187 [hep-lat/0312035].
- [22] S. Foreman, X.-Y. Jin and J.C. Osborn, *LeapfrogLayers: A Trainable Framework for Effective Topological Sampling*, in *38th International Symposium on Lattice Field Theory*, 12, 2021 [2112.01582].
- [23] C. Bonati, M. D'Elia, G. Martinelli, F. Negro, F. Sanfilippo and A. Todaro, *Topology in full QCD at high temperature: a multicanonical approach*, *JHEP* **11** (2018) 170 [1807.07954].
- [24] G. Cossu, D. Lancastera, B. Lucini, R. Pellegrini and A. Rago, *Ergodic sampling of the topological charge using the density of states*, *Eur. Phys. J. C* **81** (2021) 375 [2102.03630].
- [25] M. Luscher, *Trivializing maps, the Wilson flow and the HMC algorithm*, *Commun. Math. Phys.* **293** (2010) 899 [0907.5491].