Stabilizing complex Langevin for real-time YM theory

Based on arXiv:2212.08602

Paul Hotzy

in collaboration with Kirill Boguslavski and David I. Müller

Institute for Theoretical Physics, TU Wien

February 16, 2023

∫dkΠ
Doktoratskolleg
Particles and Interactions

TU TECHNISCHE
WIEN UNIVERSITÄT
WIEN
Vienna | Austria

FWF

Outline

Motivation

- Field theoretic and many-body systems are not solvable analytically
  - Heavy ion collisions, neutron stars, condensed matter, cold atoms, ...

- Calculations often have to be done have to be done numerically or approximations are used (or both)

- Standard numerical integration methods fail for oscillatory integrals
  - Based on sampling according to some weight function $\rightarrow$ not applicable for complex weights
  - Improving accuracy leads to exponentially many samples needed $\rightarrow$ computationally not feasible

- We miss methods that calculate observables from first principles!

Schwinger-Keldysh formalism for Yang-Mills theory

Expectation values in terms of path integrals

$$\langle \mathscr{O}[A] \rangle = \frac{1}{Z} \int \mathcal{D}A_E \, e^{-S_E[A_E]} \int \mathcal{D}A_+ \, \mathcal{D}A_- \, e^{iS[A_+,A_-]} \, \mathcal{O}(x)$$

- Expectation values of non-local observables in thermal equilibrium
  - $\langle \mathcal{O}_1(t)\mathcal{O}_2(t') \rangle = \text{Tr}\left[e^{-\beta H} \mathcal{O}_1(t)\mathcal{O}_2(t') \right]$

  - Only dependent on $t - t'$ due to time translation invariance

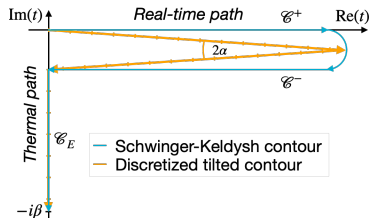 - Transport coefficients, viscosities, spectral functions, ...

  - Periodic boundary conditions:
    $A_\mu^a(t = 0) = A_\mu^a(t = -i\beta)$

  - Yang-Mills action:
    $S_{\text{YM}} = -\frac{1}{4} \int_{\mathcal{C}} d^4x F_a^{\mu\nu} F_{\mu\nu}^a$

Im(t) — Real-time path — $\mathscr{C}^+$ — Re(t)
2α
$\mathscr{C}^-$
Thermal path
$\mathscr{C}_E$
$-i\beta$

— Schwinger-Keldysh contour
— Discretized tilted contour

Sign problem of real-time YM theory

Complex weight function leads to the sign problem

- Real-time contour leads to complex weight function

$$\langle \mathcal{O}[A] \rangle = \frac{1}{Z} \int \mathcal{D}A_E \, e^{-S_E[A_E]} \int \mathcal{D}A_+ \, \mathcal{D}A_- \, \mathbf{e^{iS[A_+,A_-]}} \, \mathcal{O}(x)$$

- Direct calculation of oscillatory integrals is not feasible
- Numerical integration methods exponentially costly

Alternative integration methods needed

- Reweighting, Contour deformation, Analytic continuation, Taylor expansion, Lefschetz thimbles, **Complex Langevin method**, . . .

Reweighting not applicable

- Ideas of reweighting:
  - Split the action in imaginary $S_R + iS_I = -iS$
  - Reinterpret the observable by multiplying the phase $\mathcal{O} \rightsquigarrow \mathcal{O}e^{-iS_I}$

$$\langle \mathcal{O}_1(t)\mathcal{O}_2(t') \rangle = \frac{\int \mathcal{D}x e^{-S_R}e^{-iS_I}\mathcal{O}_1(t)\mathcal{O}_2(t)}{\int \mathcal{D}x e^{-S_R}e^{-iS_I}}$$

$$= \frac{\int \mathcal{D}x e^{-S_R}e^{-iS_I}\mathcal{O}_1(t)\mathcal{O}_2(t)}{\int \mathcal{D}x e^{-S_R}} \frac{\int \mathcal{D}x e^{-S_R}}{\int \mathcal{D}x e^{-S_R}e^{-iS_I}}$$

$$= \frac{\langle \mathcal{O}_1(t)\mathcal{O}_2(t')e^{-iS_I} \rangle_{S_R}}{\langle e^{-iS_I} \rangle_{S_R}}$$

- Enumerator and denominator can be estimated by MC
  - for the SK integral the average phase vanishes identically, reweighting is not applicable
- Assumption: $\langle e^{-iS_I} \rangle_{S_R} \neq 0$ (in practice sufficiently large)
  - measure of how "hard" the sign problem is

Outline

Introduction of the real Langevin method

Correspondence of Langevin and Fokker-Planck equations

Equivalence by same stationary solution: $P(\mathbf{x}, \theta \to \infty) = \exp\left[-S(\mathbf{x})\right]$

$$\overbrace{d\mathbf{x} = \mathbf{K}d\theta + d\mathbf{w}}^{SDE} \overset{\text{Itô's lemma}}{\Longleftrightarrow} \overbrace{\partial_\theta P(\mathbf{x}, \theta) = \nabla(\nabla - \mathbf{K})P(\mathbf{x}, \theta)}^{PDE} \qquad (1)$$

- Langevin time θ
- Wiener increment/noies term: $d\mathbf{w}$
- Drift term: $\mathbf{K} = -\nabla S$
- Averages of some random variable X depending on x reduces to sampling from stochastic process

$$\langle X \rangle_P \approx \lim_{\theta_0 \to \infty} \frac{1}{T} \int_{\theta_0}^{\theta_0 + T} d\theta \, X(\theta)$$

Introduction of the complex Langevin method

Idea: Generalization of correspondence of Langevin and Fokker-Planck equations

Complexification of degrees of freedom: $\mathcal{M} \ni \mathbf{x} \to \mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathcal{M}_c$

$$\left. \begin{array}{l} d\mathbf{x} = \mathbf{K}_x d\theta + d\mathbf{w} \\ d\mathbf{y} = \mathbf{K}_y d\theta \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} \partial_\theta P(\mathbf{x}, \mathbf{y}, \theta) = L^T P(\mathbf{x}, \mathbf{y}, \theta) \\ L = \nabla_x (\nabla_x + \mathbf{K}_x) + \nabla_y \mathbf{K}_y \end{array} \right.$$

$$\overset{?}{\Longleftrightarrow} \left\{ \begin{array}{l} \partial_\theta \rho(\mathbf{x}, \theta) = L_c^T \rho(\mathbf{x}, \theta) \\ L_c = \nabla_x (\nabla_x + \mathbf{K}) \end{array} \right.$$

Complex action leads to complex weight function $\rho(\mathbf{x}, \theta) \to \exp[-S(x)]$

- Holomorphic drift term
- Spectrum of L_c is in the left half plane and 0 is non-degenerate
- Vanishing boundary terms for all (!) moments imply

$$\langle \mathcal{O} \rangle = \int dx dy \mathcal{O}(x + iy) P(x, y, \theta \to \infty) = \int dx \mathcal{O}(x) \rho(x, \theta \to \infty)$$

Complex Langevin method for gauge fields

- CL equation for gauge fields

$$\partial_\theta A_\mu^a(\theta, x) = i \frac{\delta S_{\mathrm{YM}}}{\delta A_\mu^a(t, x)} + \eta_\mu^a(\theta, x)$$

- Complexification of the Lie algebra $\mathfrak{su}(N_c, \mathbb{C}) \to \mathfrak{sl}(N_c, \mathbb{C})$
- Gaussian distributed noise term

$$\langle \eta_\mu^a(\theta, t, \mathbf{x}) \rangle = 0,$$
$$\langle \eta_\mu^a(\theta, t, \mathbf{x}) \eta_\nu^b(\theta', t', \mathbf{x}') \rangle = 2\delta(\theta - \theta')\delta(t - t')\delta^{(d-1)}(\mathbf{x} - \mathbf{x}')\delta^{ab}\delta_{\mu\nu}$$

Goal: Overcoming the sign problem

- CL bypasses the sign problem by sampling at late θ

$$\langle \mathcal{O}[A] \rangle \approx \lim_{\theta_0 \to \infty} \frac{1}{T} \int_{\theta_0}^{\theta_0 + T} d\theta \, \mathcal{O}[A(\theta)]$$

Outline

Three major Problems of CL for real-time YM

1. **Discretized path integral along Schwinger-Keldysh contour is non-analytic w.r.t. lattice spacing**

   $\rightarrow$ we tilt the contour to regularize [see paper by N. Matsumoto [1]]

2. **Complex contours give rise to unambiguities in the CL formulation because of the noise correlator (Dirac delta)**

   $\rightarrow$ we introduce parametrization dependent CL equation

3. **CL suffers from instabilities and wrong convergence issues**

   $\rightarrow$ stabilization techniques needed - *we introduce an anisotropic kernel*

> **Path integral for Wilson action on a Minkowski time contour has an ill-defined continuum limit!**
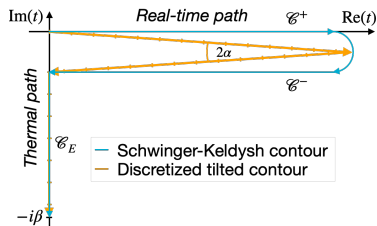
- Contour deformation resolves non-analyticity

$$\beta_0 \to e^{+i\alpha}\beta_0, \quad \beta_s \to e^{-i\alpha}\beta_s$$

- For a SK contour we effectively tilt the real-time paths

$$a_{t,k} \to e^{-i\alpha}a_{t,k} \qquad \text{for} \quad t_k \in \mathscr{C}^+,$$
$$a_{t,k} \to e^{+i\alpha}a_{t,k} \qquad \text{for} \quad t_k \in \mathscr{C}^-$$

- **Order of limits is crucial:**
  - ▸ **First $a_{t,k} \to 0$ then $\alpha \to 0$.**

Figure labels: $\mathrm{Im}(t)$, Real-time path, $\mathscr{C}^+$, $\mathrm{Re}(t)$, 2α, $\mathscr{C}^-$, Thermal path, $\mathscr{C}_E$, $-i\beta$

— Schwinger-Keldysh contour
— Discretized tilted contour

Parametrization dependent CL equation

Introduce curve parameter λ for $t(\lambda)$ leads to

$$\partial_\theta A_\mu^a(\theta, \lambda, \mathbf{x}) = i \left. \frac{\delta S_{\text{YM}}}{\delta A_\mu^a(\lambda, \mathbf{x})} \right|_{A=A(\theta)} + \eta_\mu^a(\theta, \lambda, \mathbf{x}),$$

$$\langle \eta_\mu^a(\theta, \lambda, \mathbf{x}) \eta_\nu^b(\theta', \lambda', \mathbf{x}') \rangle = 2\delta_{\mu\nu} \delta^{ab} \delta(\theta - \theta') \delta(\lambda - \lambda') \delta^{(3)}(\mathbf{x} - \mathbf{x}')$$

- Complex time contours leads to ambiguous noise correlator expression (δ-distribution for complex arguments)
  - Ambiguities are resolved by a parametrization dependent CL formulation

  $$t : [a, b] \mapsto \mathbb{C}, \qquad t(a) = 0, \quad t(b) = -i\beta$$

  - Noise correlator in terms of λ

  $$\delta(t(\lambda) - t(\lambda')) \rightsquigarrow \delta(\lambda - \lambda')$$

- Solution is however independent of the chosen parametrization
- Limiting cases (Minkowski and Euclidean contour) are consistent with new formulation

Before we come to Problem 3: Some lattice QCD...

- Link variables and plaquette variables

$$U_{x,\mu} \simeq \exp\left[iga_\mu A_\mu\left(x + \hat{\mu}/2\right)\right] \in \mathrm{SU}(N_c) \rightsquigarrow \mathrm{SL}(N_c, \mathbb{C}),$$
$$U_{x,\mu\nu} = U_{x,\mu} U_{x+\hat{\mu},\nu} U_{x+\hat{\nu},\mu}^{-1} U_{x,\nu}^{-1}$$

- Wilson plaquette action

$$S_\mathrm{W}[U] = \frac{1}{2N_c} \sum_{k,\mathbf{x},\mu\neq\nu} \beta_{\mu\nu} \mathrm{Tr}\left[U_{x,\mu\nu} - 1\right],$$

- Coupling constants

$$\beta_{0i} = -\frac{2N_c}{g^2}\frac{a_s}{a_{t,k}}, \qquad \beta_{ij} = +\frac{2N_c}{g^2}\frac{\bar{a}_{t,k}}{a_s}$$

- Time reversibility is retained by averaged spacing (turning points of the SK contour!)

$$\bar{a}_{t,k} = (a_{t,k} + a_{t,k+1})/2$$

Euler-Maruyama for numerical solution of the CLE

CL update step for link variables on a SK-contour

- Discretization of contour parameter
- Additional lattice spacing factors compared to common CL-step [2]

$$U_{x,t}(\theta + \epsilon) = \exp\left(it^a \left[i\epsilon \frac{\mathbf{a}_{\lambda,\mathbf{k}}}{\mathbf{a_s}} \left. \frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,t}} \right|_\theta + \sqrt{\epsilon} \sqrt{\frac{\mathbf{a}_{\lambda,\mathbf{k}}}{\mathbf{a_s}}} \eta^a_{x,\lambda}(\theta) \right] \right) U_{x,t}(\theta)$$

$$U_{x,i}(\theta + \epsilon) = \exp\left(it^a \left[i\epsilon \frac{\mathbf{a_s}}{\bar{\mathbf{a}}_{\lambda,\mathbf{k}}} \left. \frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,i}} \right|_\theta + \sqrt{\epsilon} \sqrt{\frac{\mathbf{a_s}}{\bar{\mathbf{a}}_{\lambda,\mathbf{k}}}} \eta^a_{x,i}(\theta) \right] \right) U_{x,i}(\theta)$$

- Langevin time step $\epsilon \to 0$
- Noise $\eta^a_{x,i}$ is approximated by a Gaussian distribution
- $\sqrt{\frac{a_s}{\bar{a}_{\lambda,k}}}$ leads to issues in the $N_t \to \infty$ limit

CL suffers from instabilities

- Runaway instabilities

  $\rightarrow$ numerical blow-up due to excursions into the non-compact manifold

- Wrong-convergence issues

  $\rightarrow$ distorted expectation values due to violation of the criterion of correctness

Bad: We are facing both types of instabilities for YM on a SK contour
Even worse: Severity of the instabilities drastically increases with shrinking tilt-angles $\rightarrow$ limit to SK contour difficult

Existing stabilization techniques

- **Gauge cooling [3]:** Exploit gauge freedom to minimize non-unitarity measured by a functional $F[U]$

$$U_{x,\mu} \mapsto U_{x,\mu}^V = V_x U_{x,\mu} V_{x+\mu}^{-1}, \qquad F[U] \geq F[U^V]$$

- **Adaptive stepsize [4]:** Regulate large drift terms which lead to instabilities

$$\epsilon \mapsto \tilde{\epsilon} = \epsilon \min\left(1, \frac{B}{\max\limits_{x,\mu,a} |K_{x,\mu}^a|}\right)$$

- **Dynamical stabiliziation [5]:** Penalize drift terms depending on the local non-unitary of the configuration

$$K_{x,\mu}^a \mapsto \tilde{K}_{x,\mu}^a = K_{x,\mu}^a + i\alpha_{\mathrm{DS}} M_x^a,$$

$$M_x^a = b_x^a \left(\sum_c b_x^c b_x^c\right), \quad b_x^a = \sum_\mu \mathrm{Tr}[t^a U_{x,\mu} U_{x,\mu}^\dagger]$$

Kernel freedom of the CL method

Field-independent kernel freedom

- *Kerneled Langevin equation* yields same limiting density function

$$\partial_\theta A^a_\mu(\theta, x) = i \int dx' \, \Gamma^{ab}_{\mu\nu}(x, x') \frac{\delta S}{\delta A^b_\nu(x')} + \int dx' \, \tilde{\Gamma}^{ab}_{\mu\nu}(x, x') \eta^b_\nu(\theta, x')$$

- $\Gamma^{ab}_{\mu\nu}(x, x')$ is required to be factorizable

$$\Gamma^{ab}_{\mu\nu}(x, x') = \int dx'' \, \tilde{\Gamma}^{ac}_{\mu\sigma}(x, x'') \tilde{\Gamma}^{bc}_{\nu\sigma}(x', x'')$$

- Kerneled Langevins equations correspond to the same Fokker-Planck equations (equivalence classes)
- *In practice:* Kernels may aggravate or mitigate instabilities

Kerneled CL update step

- Rescaling of Langevin time for temporal and spatial d.o.f.

$$U_{x,t}(\theta + \epsilon) = \exp\left(it^a\left[i\epsilon\left(\frac{a_{\lambda,k}}{a_s}\right)^2 \left.\frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,t}}\right|_\theta + \sqrt{\epsilon}\frac{a_{\lambda,k}}{a_s}\eta^a_{x,\lambda}(\theta)\right]\right) U_{x,t}(\theta)$$

$$U_{x,i}(\theta + \epsilon) = \exp\left(it^a\left[i\epsilon\left.\frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,i}}\right|_\theta + \sqrt{\epsilon}\eta^a_{x,i}(\theta)\right]\right) U_{x,i}(\theta)$$

Empirical motivation for chosen kernel:

- Noise of spatial update blows up for $a_{\lambda,k} \to 0 \quad \rightsquigarrow \quad \theta \mapsto \frac{\bar{a}_{\lambda,k}}{a_s}\theta$
- Slow down dynamics of temporal plaquettes $\quad \rightsquigarrow \quad \theta \mapsto \frac{a_{\lambda,k}}{a_s}\theta$

Comparision of update schemes

Commonly used CL update step [2]

$$U_{x,\mu}(\theta + \epsilon) = \exp\left(it^a\left[i\epsilon\,\frac{\delta S_{\mathrm{W}}}{\delta A^a_{x,\mu}}\bigg|_\theta + \sqrt{\epsilon}\eta^a_{x,\mu}(\theta)\right]\right)U_{x,\mu}(\theta)$$

New kerneled CL update step

$$U_{x,t}(\theta + \epsilon) = \exp\left(it^a\left[i\epsilon\left(\frac{a_{\lambda,k}}{a_s}\right)^2\frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,t}}\bigg|_\theta + \sqrt{\epsilon}\frac{a_{\lambda,k}}{a_s}\eta^a_{x,\lambda}(\theta)\right]\right)U_{x,t}(\theta)$$

$$U_{x,i}(\theta + \epsilon) = \exp\left(it^a\left[i\epsilon\,\frac{\delta S_{\mathrm{W}}}{\delta \tilde{A}^a_{x,i}}\bigg|_\theta + \sqrt{\epsilon}\eta^a_{x,i}(\theta)\right]\right)U_{x,i}(\theta)$$

We see great interconnection of the three proposed solutions:

- Path integral regulaization is done by contour deformation (tilting)
- Tilted contours (non-real, non-imaginary) can be dealt with by parametrized CL formulation
- Calculation of physical observables requires a double-extrapolation

$$\Rightarrow \text{first } a_t \to 0, \text{ second } \alpha \to 0$$

- Kernel systematically improves stability for smaller a_t

> **Instabilities get worse for $\alpha \to 0$ and are counteracted by our kernel for $a_t \to 0$!**

Outline

- We simulate $SU(2)$ gauge theory on a $N_t \times 4^3$ lattice

- Kernel is controlled by number of temporal lattice points N_t

- Inverse temperature $\beta = 4.0$, coupling constant $g = 1$

    ▶ We are in the deconfined phase ($\langle |P| \rangle \neq 0$).

- Simulations are initialized with identity matrices

- Langevin time is rescaled by the autocorrelation time
    $\rightarrow$ scale for number of independent samples

Correct expectation values of one-point functions

Trace of average spatial plaquette
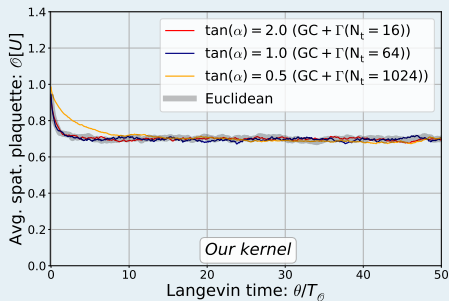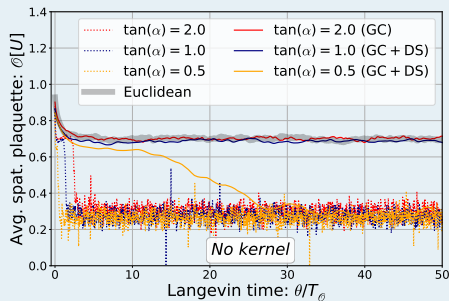
- Avg. spatial plaquette (considered in earlyier studies [2]):

$$\mathscr{O}[U] = \frac{1}{N_t N_s^3} \sum_x \frac{1}{3} \sum_{i<j} \frac{1}{N_c} \mathrm{ReTr}\, U_{x,ij}$$

- Existing methods not enough for stabilizing simulations
- **Kernels successfully stabilizes even small tilt angles**

Expectation values from CL

**Results for average spatial plaquette $\mathscr{O}$ via CL:**

$\tan(\alpha)$	Stabilization techniques	N_t	$\langle \mathscr{O} \rangle$
(Euclidean)	None	16	0.6992(3)
2.0	AS, GC	16	0.6981(2)
1.0	AS, GC, DS	16	0.6858(1)
0.5	AS, GC, DS	16	0.2801(3)
2.0	AS, GC, Γ	16	0.6987(3)
1.0	AS, GC, Γ	64	0.6977(5)
0.5	AS, GC, Γ	1024	0.6973(8)
0.4	AS, GC, Γ	8192	0.6968(6)

Systematic improvement of CL instabilities

Anisotropic kernel systematically stabilizes CL

- Our kernel allows simulations even for $\max \operatorname{Re}(t) > \beta$
- Stability is systematically improved for partial continuum limit
  - $\Rightarrow$ **Calculation of real-time observables might be feasible!**

- Stability region grows faster then auto-correlation time w.r.t. N_t
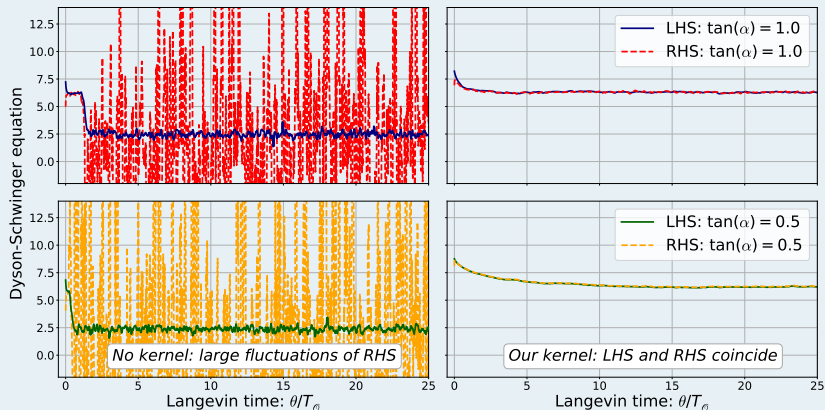
- Computational cost grows linearly with N_t
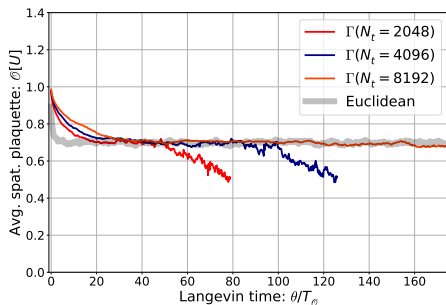
Tilt angle: $\tan(\alpha) = 0.4$

Dyson-Schwinger equations

DSE for spatial plaquettes
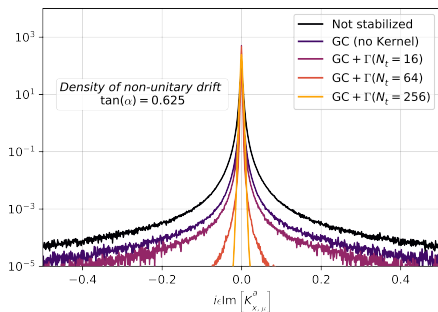
- Self-consistency check of link configuration

$$\frac{2(N_c^2 - 1)}{N_c} \left\langle \mathrm{ReTr}(U_{x,ij}) \right\rangle = \frac{i}{2N_c} \sum_{|\rho| \neq i} \beta_{i\rho} \left\langle \mathrm{ReTr} \left[(U_{x,i\rho} - U_{x,i\rho}^{-1})U_{x,ij} \right] \right\rangle$$

Localized histogram of the drift terms

Anisotropic kernel and the criterion of correctness

- We implicitly assume that $K_y PO$ decays sufficiently fast
- Localised histograms of the K_y should lead to vanishing boundary terms

- Gauge cooling helps, but skirts are still present

- No skirts of histograms using our kernel with sufficiently large N_t

Density of non-unitary drift
$\tan(\alpha) = 0.625$

Legend:
- Not stabilized
- GC (no Kernel)
- GC + $\Gamma(N_t = 16)$
- GC + $\Gamma(N_t = 64)$
- GC + $\Gamma(N_t = 256)$

x-axis: $\iota\varepsilon \mathrm{Im} \left[K^a_{x,\mu} \right]$

Outline

Conclusion & Outlook

Conclusion

- Stabilization techniques extend the applicability of CL
- We found an approach to systematically improve stability
- Anisotropic kernel enables simulations of $\max \operatorname{Re}(t) > \beta$
  - ⤳ Extrapolation to Schwinger-Keldysh contour might be possible

- What I have shown you
  - ▸ Average spatial plaquette
  - ▸ Dyson-Schwinger equations
  - ▸ Histograms of the drift term

- What I have *not* shown you
  - ▸ Larger Wilson loops ($W_{n \times n}$)
  - ▸ Polyakov loop
  - ▸ Unitarity norm
  - ▸ Euclidean correlators

Teaser of our most recent results

Outlook

- Checks of mathematical features of correlation function
  - Independence $\frac{t+t'}{2}$
  - Relations of Wightman functions and Propagators
  - Fluctuation-dissipation relation

- Calculation of physical lattice spacing (onging)
  - Renormalization needed for physical observables

- Correlation functions (onging)
  - We carry out the double extrapolation
  - We are able to simulate even larger real-time extents at weak couplings

- Far-future prospect: Non-thermal quantum systems

References

📄 N. Matsumoto, *Comment on the subtlety of defining real-time path integral in lattice gauge theories*, *PTEP* **2022** (2022), no. 9 093B03, [`arXiv:2206.00865`].

📄 J. Berges, S. Borsanyi, D. Sexty, and I. O. Stamatescu, *Lattice simulations of real-time quantum fields*, *Phys. Rev. D* **75** (2007) 045007, [`hep-lat/0609058`].

📄 E. Seiler, D. Sexty, and I.-O. Stamatescu, *Gauge cooling in complex Langevin for QCD with heavy quarks*, *Phys. Lett. B* **723** (2013) 213–216, [`arXiv:1211.3709`].

📄 G. Aarts, F. A. James, E. Seiler, and I.-O. Stamatescu, *Adaptive stepsize and instabilities in complex Langevin dynamics*, *Phys. Lett. B* **687** (2010) 154–159, [`arXiv:0912.0617`].

📄 F. Attanasio and B. Jäger, *Dynamical stabilisation of complex Langevin simulations of QCD*, *Eur. Phys. J. C* **79** (2019), no. 1 16, [`arXiv:1808.04400`].