The Research Council of Norway

University of Stavanger

Machine Learning assisted real-time simulations with Complex Langevin

Alexander Rothkopf
Faculty of Science and Technology
Department of Mathematics and Physics
University of Stavanger

NPACT
Norwegian Particle, Astroparticle & Cosmology Theory network

in collaboration with Daniel Alvestad & Rasmus Larsen

D. Alvestad, R. Larsen, A.R.
JHEP 08 (2021) 138 and
JHEP 04 (2023) 057

Overview

University of Stavanger

- **Reinforcement learning**
- **Real-time simulations & Complex Langevin (CL)**
- **Learning optimal kernels for correct convergence of CL**
- **Real-time dynamics in quantum mechanics & 1+1d field theory**
- **Summary**

Reinforcement learning – a ML success

University of Stavanger

- Agent with a set of **predefined actions** [e.g. move left, jump] in an **environment**

  Karakovskiy and Togelius , *IEEE Trans. on Com. Intel. and AI in Games* 4.1 (2012): 55-67

- **Policy/Cost function** that defines **success** [e.g. score on computer screen]

- Need to **encode** choice of actions and evaluate **gradients** to minimize cost

  Wang, Ziyu, et al. *Int. conf. on machine learning*. PMLR, 2016.

Need to handle **failure state** [e.g. falling into pits]

Improving the score: allow for more actions [e.g. move right]

Executive summary – ML strategy

University of Stavanger

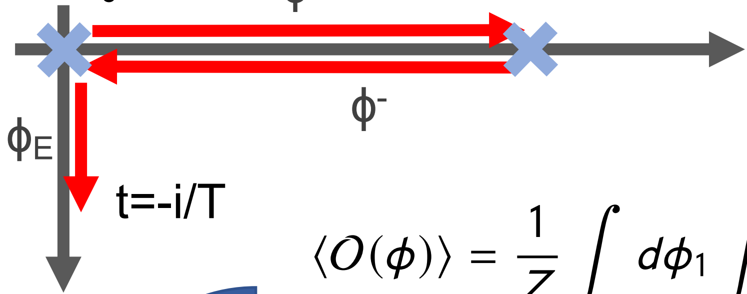
$$\frac{d\phi}{d\tau_L} = iK[\phi]\frac{\partial S}{\partial \phi} + \frac{\partial K[\phi]}{\partial \phi} + \sqrt{K[\phi]}\eta$$

- Environment: **space of distributions** explored by a stochastic process
- **Agent**: controller of the **non-neutral modification** represented by the **kernel K**. Limited actions – keep the kernel field & τ_L independent
- **Cost function**: deviation of late τ_L stationary distribution from **prior knowledge** (symmetries, known cumulants, etc.)
- Use auto differentiation or shadowing analysis to compute **robust gradients** of the inherently **chaotic dynamics**.
- We achieve **convergence to correct stationary distribution** for model systems in parameter regimes **previously inaccessible**.

Real-time quantum dynamics

University of Stavanger

- The path integral at finite temperature on the Schwinger-Keldysh contour



Goal: evaluation of real-time observables

$<O(t_0)O(t_1)> = \mathrm{Tr}[\ \rho\ O(t_0)O(t_1)\ ]$

$$\langle \mathcal{O}(\phi)\rangle = \frac{1}{Z}\underbrace{\int d\phi_1 \int d\phi_2\, \rho(\phi_1,\phi_2)}\ \underbrace{\int_{\phi_2}^{\phi_1} D\phi^+ D\phi^-\ \mathcal{O}(\phi)\ e^{iS[\phi_+]-iS[\phi_-]}}$$

sampling over statistically distributed initial conditions

quantum "sum over paths"

$$\langle \mathcal{O}(\phi)\rangle = \frac{1}{Z}\int D\phi_E\, e^{-S_E[\phi_E]} \int_{\phi_E(\beta)}^{\phi_E(0)} D\phi^+ D\phi^-\ \mathcal{O}(\phi)\ e^{iS[\phi_+]-iS[\phi_-]}$$

Real-valued Feynman weight:
Monte-Carlo methods applicable

Pure phase Feynman weight implies MC sign problem. One strategy:
Complex Langevin see C. Berger et.al. Phys.Rept. 892 (2021)

- Sign problem is NP-hard: **no generic solution** strategy is likely to exist
Troyer, Wiese PRL 94 170201 (2004)

Stochastic Quantization

University of Stavanger

- Langevin evolution in fictitious additional time to reproduce quantum fluctuations

for an in-depth review: M. Namiki et.al. Stochastic Quantization (Springer) 1992

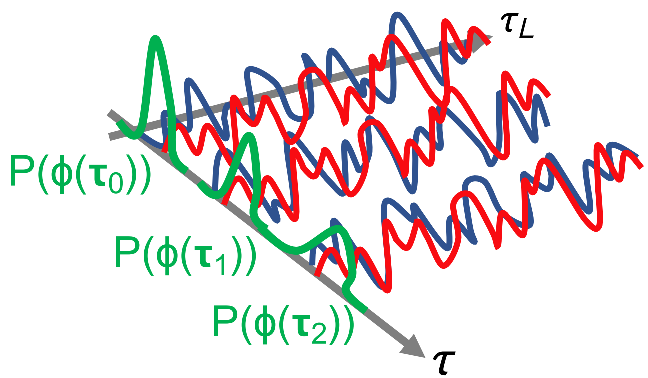
$$\frac{d\phi}{d\tau_L} = -\frac{\delta S_E[\phi]}{\delta\phi(x)} + \eta(x,\tau_L) \quad \text{with} \quad \langle\eta(x,\tau_L)\rangle = 0, \quad \langle\eta(x,\tau_L)\eta(x',\tau_L')\rangle = 2\delta(x-x')\delta(\tau_L-\tau_L')$$

Stochastic partial differential equation (SDE) with Gaussian noise

- Associated Fokker-Planck equation for P[ϕ]

$$\frac{\partial}{\partial\tau_L}\mathcal{P}(\phi) = \nabla_\phi\left[\left(S_E[\phi] + \nabla_\phi\right)\mathcal{P}(\phi)\right]$$

- Proof of convergence: $\lim_{\tau_L\to\infty} P[\phi,\tau_L] = e^{-S_E[\phi]}$



complexification: $-\frac{\delta S_E[\phi]}{\delta\phi(x)} \Rightarrow i\frac{\delta S[\phi]}{\delta\phi(x)}$ $\quad \phi(x,\tau_L) = \phi_R(x,\tau_L) + i\phi_I(x,\tau_L)$

$$\frac{d\phi_R}{d\tau_L} = \mathrm{Re}\left[i\frac{\delta S[\phi]}{\delta\phi(x)}\bigg|_{\phi=\phi_R+i\phi_I}\right] + \eta(x,\tau_L), \qquad \frac{d\phi_I}{d\tau_L} = \mathrm{Im}\left[i\frac{\delta S[\phi]}{\delta\phi(x)}\bigg|_{\phi=\phi_R+i\phi_I}\right]$$

$$\langle O[\phi]\rangle \leftrightarrow \lim_{T\to\infty}\frac{1}{T}\int_0^T d\tau_L O[\phi_R(x,\tau_L) + i\phi_I(x,\tau_L)]$$

Two challenges for Complex Langevin

University of Stavanger

$$\frac{d\phi_R}{d\tau_L} = \mathrm{Re}\left[i\frac{\delta S[\phi]}{\delta\phi(x)}\bigg|_{\phi=\phi_R+i\phi_I}\right] + \eta(x,\tau_L), \quad \frac{d\phi_I}{d\tau_L} = \mathrm{Im}\left[i\frac{\delta S[\phi]}{\delta\phi(x)}\bigg|_{\phi=\phi_R+i\phi_I}\right]$$

for a community overview see e.g. E. Seiler LATTICE17 EPJ Web Conf. 175, 2018

Divergent solutions (runaways)

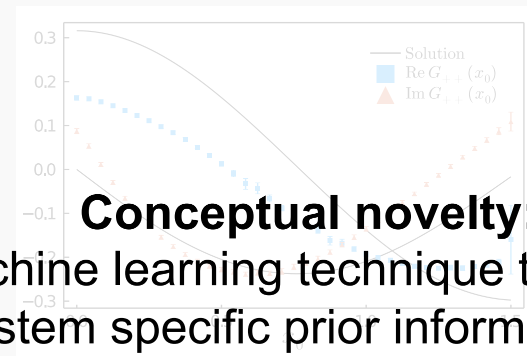
Technical novelty:
Implicit solvers render runaway problem moot

provide stability needed
to carry out **ML optimization**

In practice: use adaptive step size in attempt to keep solution finite

see e.g.: G. Aarts et.al. PLB 687(2-3), 154–159 (2010)

Convergence to incorrect solutions as real-time extent increases



Conceptual novelty:
Machine learning technique to infuse system specific prior information - loop-hole to beat the NP-hard sign problem

$$\int d\phi_R \int d\phi_I O(\phi_R + i\phi_I) P_{CL}(\phi_R,\phi_I) \neq \int d\phi O(\phi) e^{iS[\phi]}$$

Only a posteriori criterion available: tails in histogram of field d.o.f.

see e.g.: G. Aarts et.al. Eur. Phys. J. C71 (2011) 1756

Implicit solvers for Complex Langevin

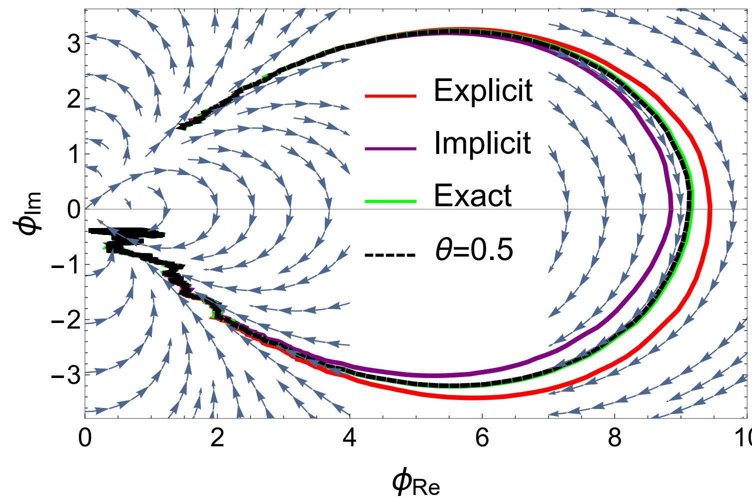
University of Stavanger

- Numerical solution of stochastic dynamics in the literature: explicit forward Euler

$$\frac{d\phi}{d\tau_L} = i\frac{\delta S[\phi]}{\delta\phi(x)} + \eta(x,\tau_L)$$

$$\phi_j^{\lambda+1} = \phi_j^{\lambda} + i\epsilon\frac{\partial S^{\lambda}}{\partial\phi_j} + \sqrt{\epsilon}\eta_j^{\lambda}$$

ϵ Langevin time step



  - Appearance of runaways indicates **stiff problem**: from PDEs we know implicit methods can help

$$\phi_j^{\lambda+1} = \phi_j^{\lambda} + i\epsilon\left[\theta\frac{\partial S^{\lambda+1}}{\partial\phi_j} + (1-\theta)\frac{\partial S^{\lambda}}{\partial\phi_j}\right] + \sqrt{\epsilon}\eta_j^{\lambda}$$

general Euler-Maruyama scheme

Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations, 1–50 (1992)

- Inherent regularization allows for the first time to simulate on untilted SK contour

D. Alvestad, R. Larsen, A.R. JHEP 08 (2021) 138

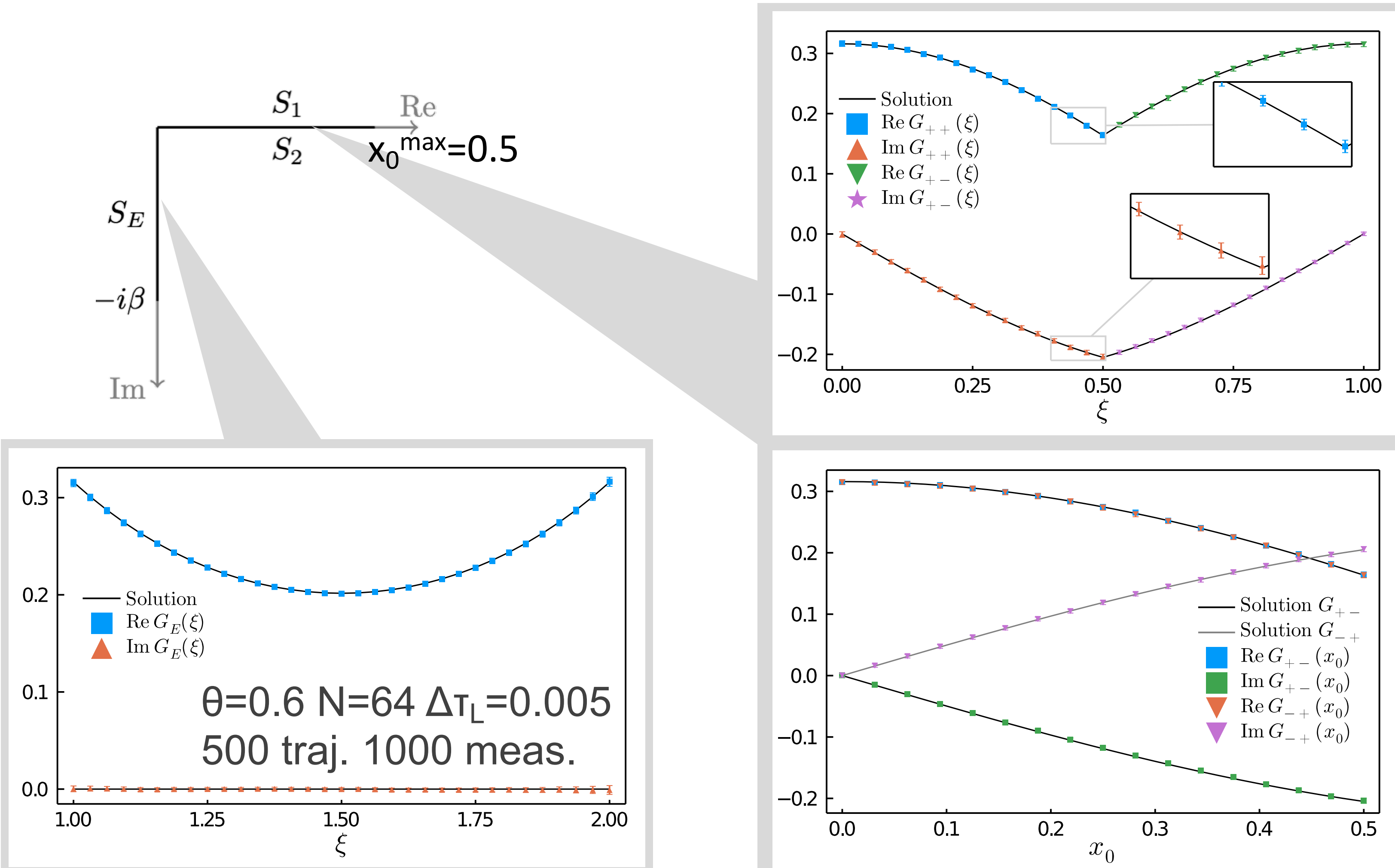
$$S_\theta = \frac{1}{2}\phi\left(M + i\epsilon\theta M^2\right)\phi = S_{\text{explicit}} + \frac{i\epsilon}{2}\theta\sum_j S_j^2$$

Numerical results at short real-times I

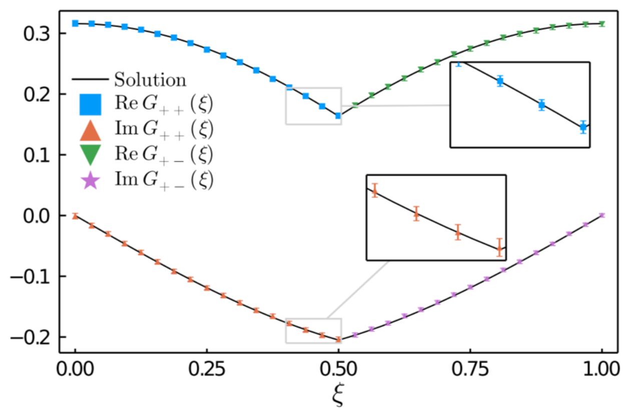
University of Stavanger

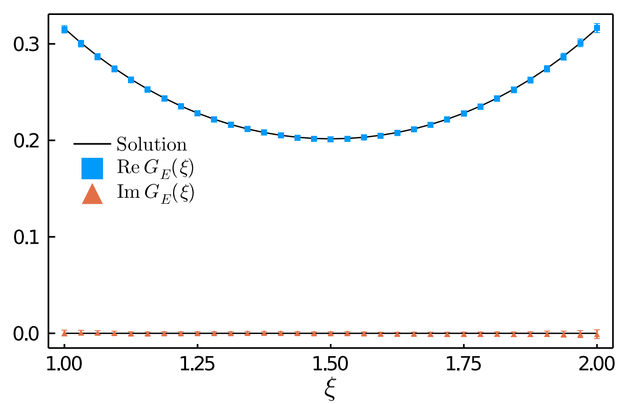
- Direct simulations **anharmonic oscillator** on the canonical SK contour in **thermal equilibrium** possible m=1 λ=24 βm=m/T=1

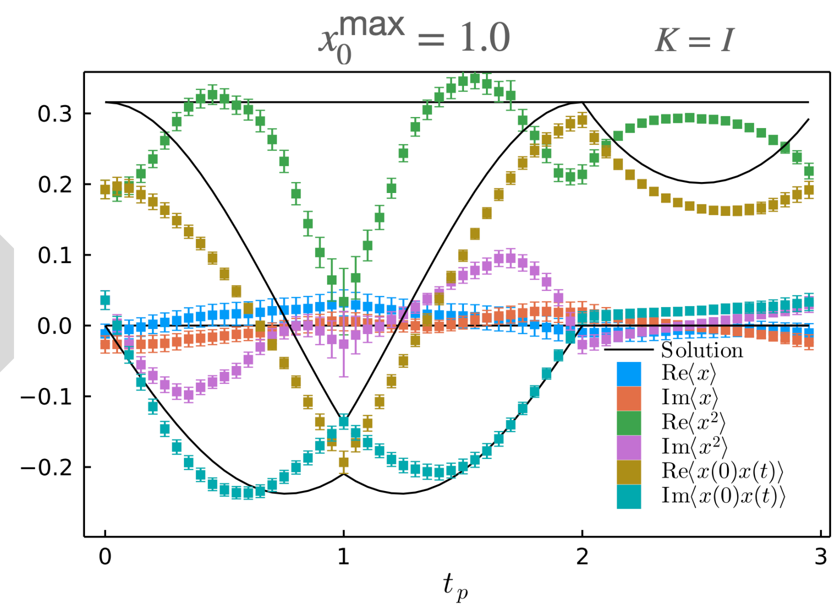
D. Alvestad, R. Larsen, A.R. JHEP 08 (2021) 138



What happens at later real-time?







- Convergence to incorrect solution without apparent pathologies

Convergence in CL

University of Stavanger

$$\lim_{\tau_L\to\infty}\int d\phi_R d\phi_I \mathcal{O}(\phi_R+i\phi_I)P_{\rm CL}[\phi_R,\phi_I,\tau_L] \stackrel{?}{=} \int d\phi\, \mathcal{O}(\phi)\, e^{iS}$$

- Necessary, while not sufficient criterion for correct convergence: absence of **boundary terms**

  developed in: G. Aarts et.al. Eur. Phys. J. C71 (2011) 1756
  see also: D. Alvestad, R. Larsen, A.R. JHEP 04 (2023) 057

- Strategies to minimize boundary: pull complexified d.o.f. back to a real manifold
  - **Gauge cooling**: exploit freedom to bring SL(N,C) links as close as possible to SU(N)

    Seiler, Sexty, Stamatescu, PLB 04 62 (2013)
  - **Dynamical stabilization**: modified drift term pulls towards the origin (non-holomorphic)

    Aarts, Attanasio, Jaeger, Sexty Acta Phys. Polon. Supp. 9, 621 (2016)

- Our idea for NP-hard sign problem: incorporate **system specific prior information** without affecting the proof of convergence

Kernelled complex Langevin

University of Stavanger

- Simultaneous modification of drift and noise allows to alter FP spectrum

$$\frac{d\phi}{d\tau_L} = iK[\phi]\frac{\partial S}{\partial \phi} + \frac{\partial K[\phi]}{\partial \phi} + \sqrt{K[\phi]}\eta$$

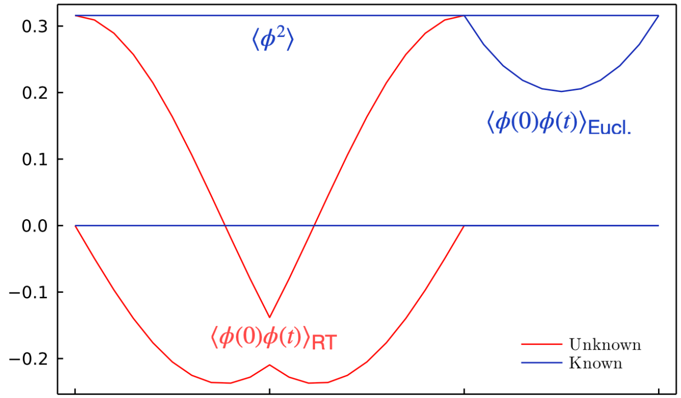
- **Optimality via prior information**: Symmetries, Euclidean correlator, Boundary

D. Alvestad, R. Larsen, A.R. JHEP 04 (2023) 057

$$L^{\rm sym} = \sum_t \left\{ \langle \phi_t \rangle^2 + \langle \phi_t^3 \rangle^2 + \left(\langle \phi_t^2 \rangle - \phi^2 \right) \right\}$$

$$L^{\rm bnd} = \sum_i \sum_k \left\{ \langle L_c[\phi_i] \mathcal{O}_k \rangle_Y \right)^2$$

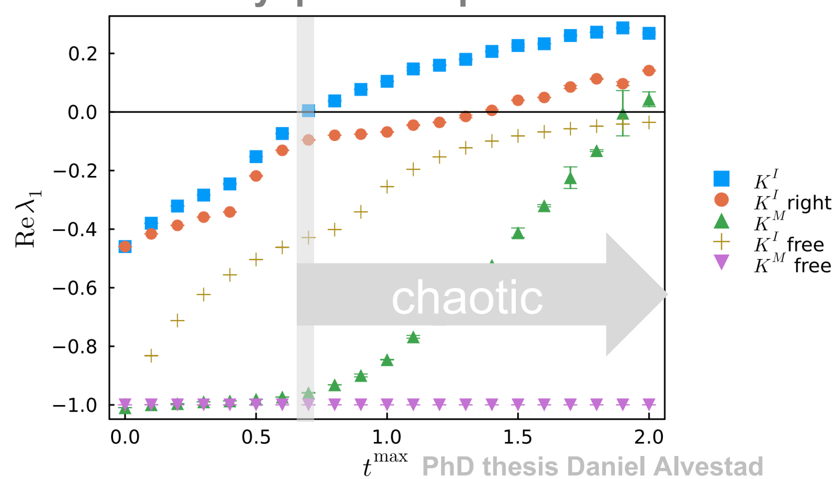
$$L^{\rm eucl} = \sum_i \left\{ \left(\langle \phi_0 \phi_i \rangle - D_i^E \right)^2 \right\}$$



Systematic learning of optimal kernels

- Autodifferentiation techniques to compute $\frac{\partial L^{\text{tot}}}{\partial K_{ij}}$ (derivative of stochastic process) [note: deterministic dynamics chaotic]

CL Lyapunov exponents



PhD thesis Daniel Alvestad

- In principle possible, in practice slow: cheaper optimization functional instead

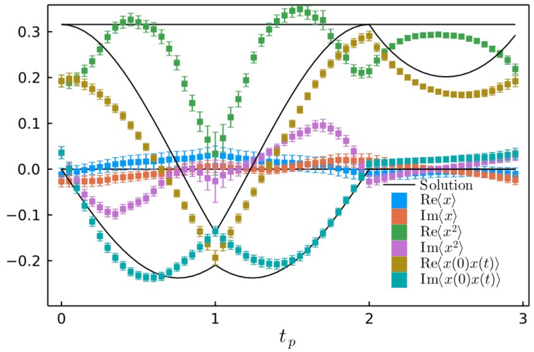
$$L^{\text{low cost}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\left| K\frac{\partial S}{\partial \phi}_t(-\phi_t) - \left|K\frac{\partial S}{\partial \phi}_t\right| |\phi_t| \right|^2$$

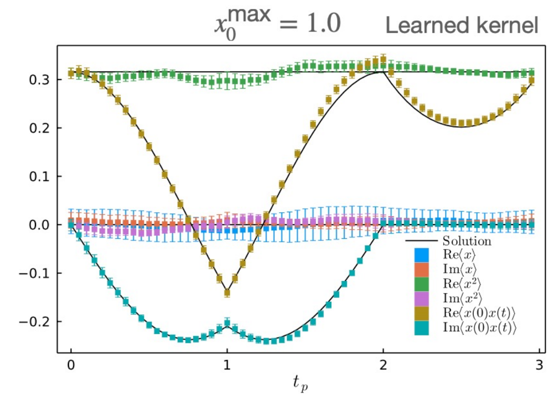
minimizes drift away from the origin (similar to dynamic stabilization but remains holomorphic)

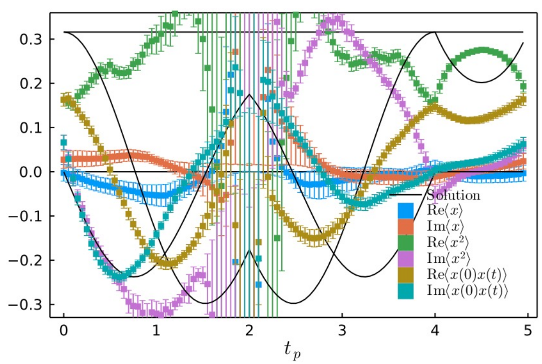
Performance in practice

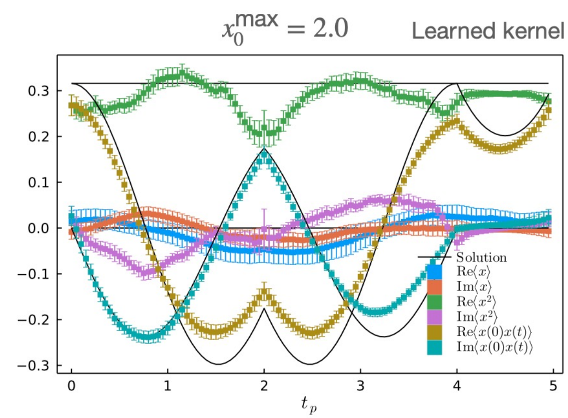
- Using a constant kernel $K = exp[A + iB]$ with A,B real matrices
- Optimize via low cost functional & check success via symmetries & Euclidean corr.

D. Alvestad, R. Larsen, A.R. JHEP 04 (2023) 057









- Achieve correct convergence up to **3x time extent** previously reported in literature

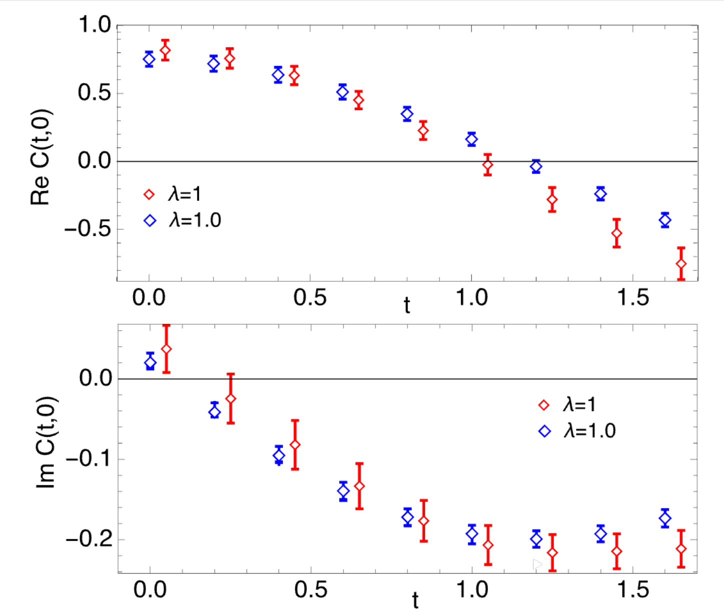
Optimal kernels & real-time field theory

University of Stavanger

- Promising scaling: CL with # of grid points – field independent kernel avoids need for Jacobian – implicit solvers performant even though Newton step
- Found that for 1+1d low-cost functional proposed by Graz group offers even better performance. N. Lampl and D. Sexty (in preparation)

$$L^{\text{low cost}} = \frac{1}{N_t}\sum_{i=1}^{N_t} \text{Im}[\phi]^2$$

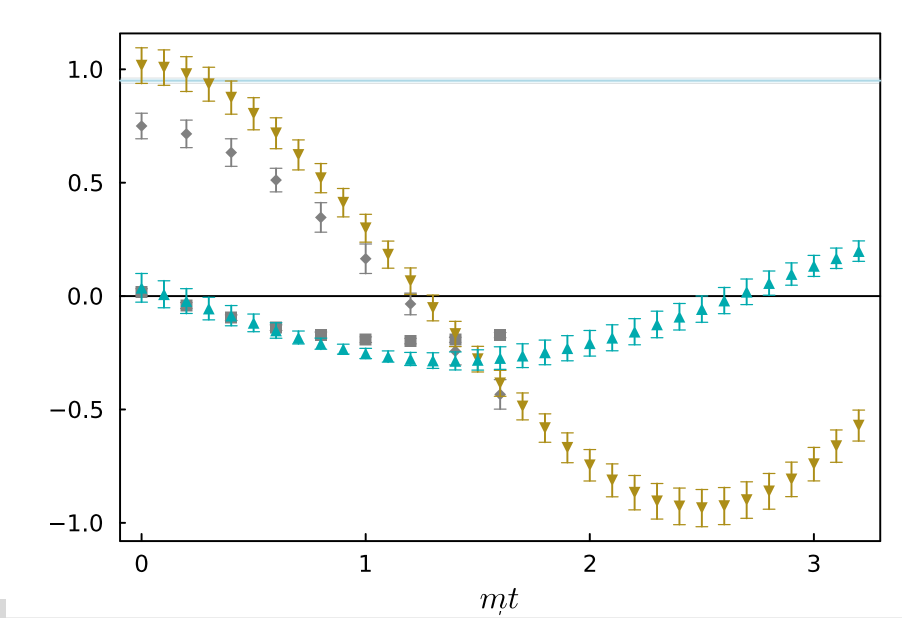
Current community benchmark



A. Alexandru et.al. PRD 95 (2017) 11, 114501

- based on contour deformations
- coarse grid on SK due to cost

Optimal learned CL kernels



(in preparation)

- avoid discretization artifacts with finer grids
- significantly extend simulation range

Conclusion & Outlook

University of Stavanger

- Overcoming **NP-hard sign problem** central to progress in theoretical physics
- **Complex Langevin** one possible path forward, but hampered by two major challenges: **instabilities** and **convergence to incorrect** solutions
- **Implicit solvers** render the runaway problem moot & allow stable optimization. Realize simulations on the **canonical** Schwinger-Keldysh contour.
- **ML strategy**: systematically incorporate **system specific prior information** (symmetries, Euclidean correlators, etc.) in simulation via **kernel** modification

Learned optimal field independent kernels

- Optimal **kernels in QM**: 3x extended range of validity of real-time CL
- Optimal **kernels in 1+1d**: new benchmark in accuracy & real-time extent (~2x)

- Next step: cost effective optimization strategies for **field dependent kernels** (adjoint sensitivity analysis, shadowing method (NILSS), etc.)

Backup slides

Limits to our current strategy

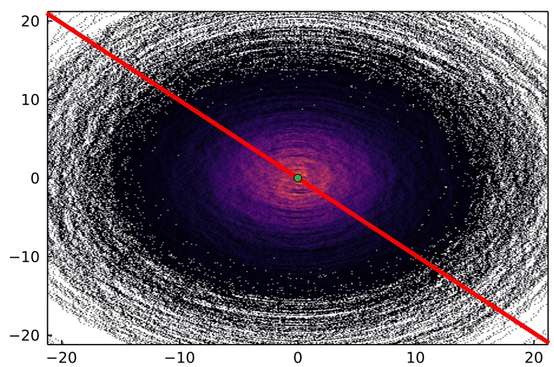
University of Stavanger

- Constant kernel works well in theories with single critical point at the origin

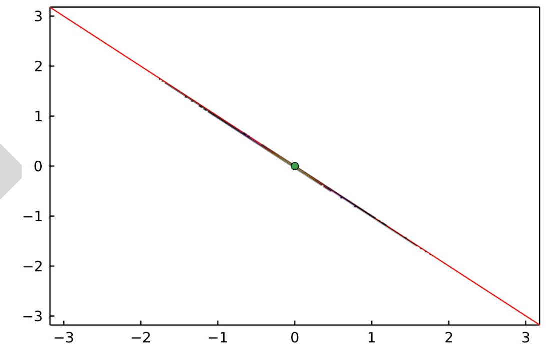
simple Gaussian model $S = \frac{1}{2} i x^2$

Lefschetz thimbles $\frac{d\phi}{d\tau} = \overline{\frac{dS}{d\phi}}$

naïve CL

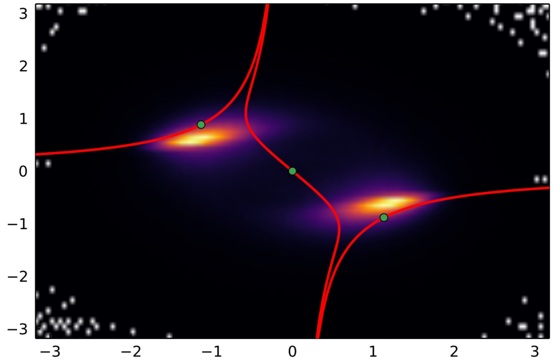


learned kernelled CL



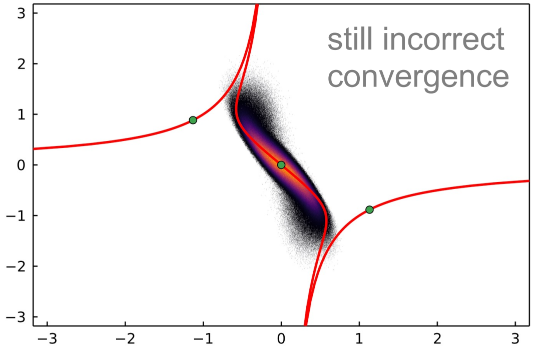
- Multiple critical points may require a field dependent kernel: $S = 2ix^2 + (1/2)x^4$

naïve CL



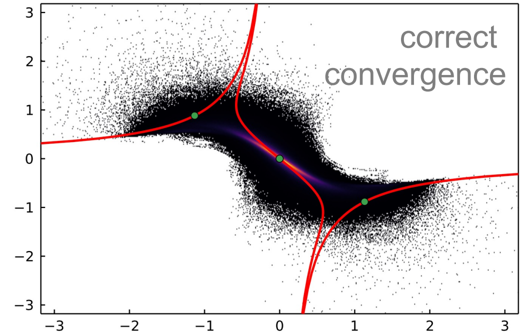
L^{tot}=0.888

learned const. kernelled CL



L^{tot}=0.486

field dependent kernel from Okamoto et.al. 1989



L^{tot}=0.023

Two convergence criteria

University of Stavanger

- Non-unique optima from the low-cost functional: allow to avoid boundary terms
- Correct convergence iff **in addition** Fokker-Planck EV in lower half plane

see D. Alvestad, R. Larsen, A.RJHEP 04 (2023) 057 and PhD thesis D. Alvestad

$$S = \frac{1}{2}\sigma x^2 + \frac{\lambda}{4}x^4 \qquad \sigma = 4i,\ \lambda = 2 \quad \Rightarrow \quad K = e^{i\theta}$$

single number, optimum found by scan of $L_{lowcost}$