Oscillating autocorrelation and the HMC algorithm

Falk Zimmermann
fzimmermann@hiskp.uni-bonn.de

Helmholtz-Institut für Strahlen- und Kernphysik

39th International Symposium on Lattice Field Theory
August 2022

Motivation: fluctuation of correlators

Algorithm:

1. remove τ average and IR filter

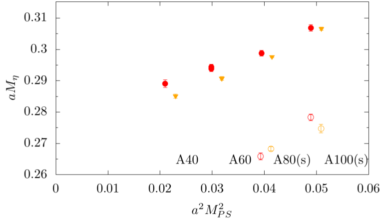
$$c_\lambda(t,\tau) \equiv \sum_{\tau'=1}^{N} [c(t,\tau') - \langle c\rangle_\tau(t)] \frac{1}{\sqrt{2\pi\lambda^2}} e^{-\frac{(\tau-\tau')^2}{2\lambda^2}}$$

2. normalise each time slice, respectively

$$\bar{c}(t,\tau) \equiv \frac{c_\lambda(t,\tau) - \langle c_\lambda\rangle_\tau(t)}{\sigma_\tau[c](t)}$$

3. remove all t means, respectively

$$\bar{\bar{c}}_\lambda(t,\tau) \equiv \bar{c}_\lambda(t,\tau) - \langle \bar{c}_\lambda\rangle_t(\tau)$$



t-shift first: $c(t) \to c(t+1) - c(t)$

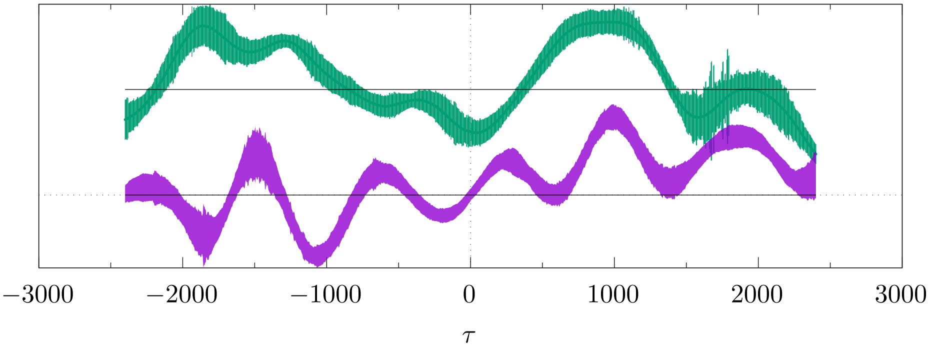
remove excited states first via cosh fit

Meta correlation

The fluctuation fields $\bar{c}(\tau, t)$ & $\bar{\bar{c}}(\tau, t)$ are t-isotropic and one may study their autocorrelations

- generalised 2d-correlation: $\Gamma[c_a, c_b](\Delta\tau, t)$
- standard AC for time slice t: $\Gamma_t[c_a, c_b](\Delta\tau)$
- and the t-averaged AC: $\langle\Gamma_t[c_a, c_b]\rangle_t(\Delta\tau)$

for fields $c_a = \bar{c}[\eta_l^{\text{connected}}]$ and the topology correlator $c[qq_{10}]$ with $q \propto tr[F_{\mu\nu}F^{\mu\nu}]$, $N_{\text{HYP}} = 10$.



Oscillating autocorrelation

Challenge:

Complex and/or hard to estimate autocorrelation $\Gamma(\tau)$

Prime examples:

- long AC with noisy tail
  $\rightarrow$ prevents precise estimation of the integrated AC time τ_{int}. **[Wolff:2003sm]**
  $\rightarrow$ and variance estimation $\sigma^2 = 2\tau_{\mathrm{int}}\Gamma(0)/N$
- non-stationary stochastic process: dependence on iteration number τ
  $\Gamma(\Delta\tau) \ := \ \langle\hat{o}(0)\hat{o}(\Delta\tau)\rangle_{\text{independent runs}} \neq \langle\hat{o}(\tau)\hat{o}(\Delta\tau+\tau)\rangle_{\text{independent runs}}$ for $\tau > 0$

Madras-Sokal formula

expected variance of the normalised AC $\rho(\tau) \equiv \Gamma(\tau)/\Gamma(0)$: **[Luscher:2004pav]**

$$\langle\delta\rho^2(\tau)\rangle \approx \frac{1}{N}\sum_{k=1}^{\infty}\{\rho(k+\tau)+\rho(k-\tau)-2\rho(k)\rho(\tau)\}$$

is not valid for non-stationary stochastic processes!

An effective number of independent DOF

Idea 1: entropy $\mathcal{H}$ of a distribution is a measure for the amount of contained information:

$$N_{\mathrm{eff}}[c] \;\propto\; \mathcal{H}[p[c_a](t,\tau)]$$

Idea 2: Interested in the ratio given by the entropy of the autocorrelated distribution $\mathcal{H}[Corr]$ divided by $\mathcal{H}[\mathrm{diag}(Corr)]$. Advantages:

- No normalisation needed: don't care about units of information (bits, nats)

It can be shown, that the Shannon entropy $\mathcal{H}[p(x)] \;\equiv\; -\int_x p(x)\log p(x)\mathrm{d}x$ for multiple (n) normal p-distributions with correlation $Corr$:

$$\mathcal{H}[c] \;=\; \log\left(\sqrt{2\pi}\,\det(Corr)\right) + n/2$$

$$N_{\mathrm{eff}}[c] \;\equiv\; \frac{\mathcal{H}[p, Corr]}{\mathcal{H}[p, \mathbb{1}]} \;=\; \frac{\log\left(\sqrt{2\pi\det^2(Corr)}\right) + N_{\mathrm{conf}}/2}{\log\left(\sqrt{(2\pi\sigma_f^2)^{N_{\mathrm{conf}}}}\right) + N_{\mathrm{conf}}/2}$$

Entropy of continuous distributions

$$\mathcal{H}(p) \;\equiv\; -\sum_x p(x)\log p(x) \;\rightarrow\; -\int p(x)\log p(x)\,dx$$

Naive generalisation of discrete $\mathcal{H}$ for continuous p distributions is flawed:

- if p is approximated by a histogram h with n_{classes}
  $\rightarrow \mathcal{H}$ strongly depends on n_{classes}
- consider a simple rectangular probability distribution, represented by:

$$h_1(c) = 1 \text{ with } n_{\text{classes}} = 1 \;\;\&\;\; h_2(c) = 1/2 \text{ with } n_{\text{classes}} = 2$$

$$\rightarrow \mathcal{H}[h_1] = 0 \;\neq\; \mathcal{H}[h_2] = \log 2$$

- naive definition widely in use: $\mathcal{H}[\mathcal{N}(\mu,\sigma)] \;=\; \frac{1}{2}\log(2\pi\sigma^2) + \frac{1}{2}$
  ... $\mathcal{H}[\mathcal{N}]$ becomes negative $\Leftrightarrow \sigma^2 < 1/2\pi e$

Limiting density:

$\mathcal{H} \;\equiv -\int_x p(x)\log\frac{p(x)}{m(x)}\,dx$, $m(x)$ uniform distribution: “resolution” **[Jaynes:1963]**

Determining the resolution m(x)

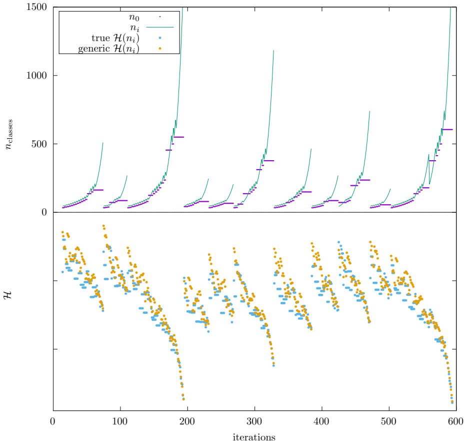
The resolution has a large impact on the amount of information needed to store an invent drawn from $p(x) \rightarrow$ Need to compute the resolution for each element of $Corr(\tau, t)$:

Idea: increasing the resolution m$=n_{\text{classes}}^{-1}$ of a histogram h increases $\mathcal{H}[h]$ slower if there is more than uniform noise.

Algorithm:

input: list L of data pts

1. generate histogram $h[L]$, $n_{\text{classes}} = n_0$
2. draw N_{samples} lists G_s of length$(L) \sim h[L]$
3. $n_i := n_0$, iterate:
   at resolution $n_{\text{classes}} := n_i$ generate
   - generic histogram samples $h_s[G_s]$
   - the true histogram $h_{\text{True}}[L]$
   - if $\langle \mathcal{H}[h_s] \rangle_s - \mathcal{H}[h_{\text{True}}] > \sigma_c$ $\rightarrow$ $n_0 := n_i$, go to 1.
   - increase $n_i := max(1.1 n_i, n_i + 1)$



Gives lower boundary of m.

Toeplitz matrix

Main task is the computation of the determinant of the $N_{\text{conf}} \times N_{\text{conf}}$ autocorrelation matrix $Corr$:

$$Corr(\tau_1, \tau_2) \equiv \sqrt{\Gamma_a(|\tau_1 - \tau_2|) * \Gamma_b(|\tau_1 - \tau_2|)} \quad \text{for correlators } a,\ b$$

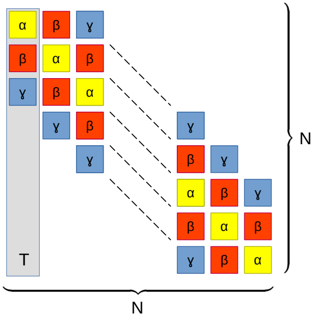
$\rightarrow$ simple structure: constant diagonals!

- so called "Toeplitz matrix"
- characterized by a N_{conf} vector T
- computation of is not trivial!
- error of the AC: repeated for each point τ

Algorithm:

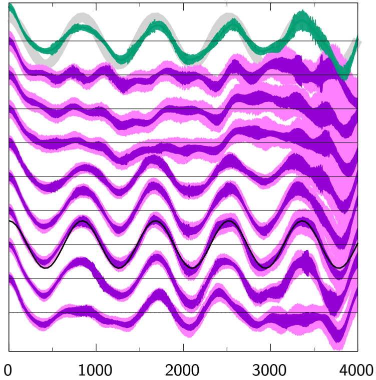
"A fast elementary algorithm for computing the determinant of toeplitz matrices" **[Cinkir:2011]**

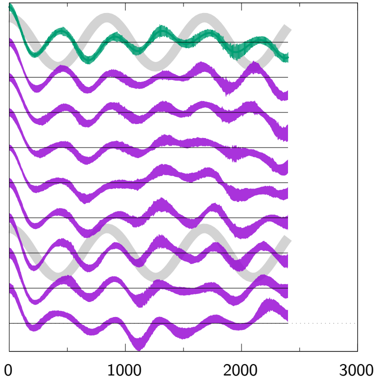
- complexity $\mathcal{O}((N_{\text{cut-off}} - 1)^3 \cdot log(N))$
- $N_{\text{cut-off}} \approx \frac{1}{2} N_{\text{conf}} \leftarrow$ AC oscillations
  $N_{\text{cut-off}} \ll N_{\text{conf}}$ may even result in $\det(Corr) < 0$!
- was generalized for block toeplitz matrices



Result1: Oscillating autocorrelation

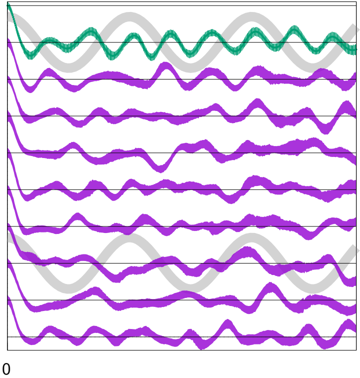
Example1: AC of connected eta_l (LHS) and η_s correlators (RHS) for $\lambda = 68$



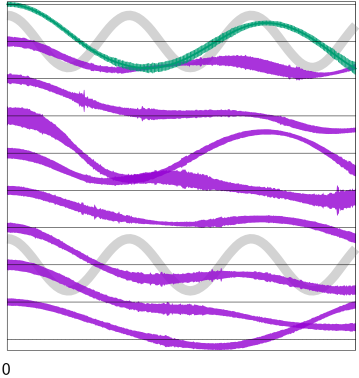


Result1: Oscillating autocorrelation

Example2: topological charge q via an improved version of $tr[F_{\mu\nu}F^{\mu\nu}]$ with $N_{\text{HYP}} = 10$ smearing steps.



$\lambda = 36$



$\lambda = 304$

Amplitudes:

How much of the signal is left?

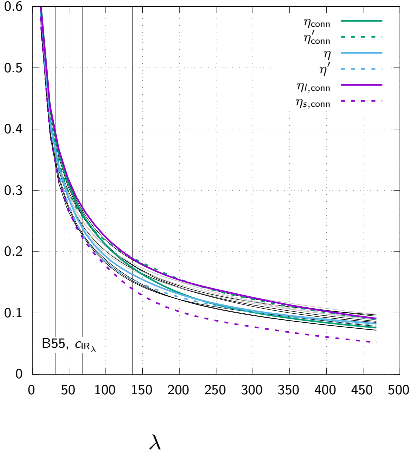
- define relative amplitude via the sd ratio of fields c_λ & $c_{\lambda=0}$:

$$A_\lambda \equiv \sigma_{t,\tau}\left[c_\lambda(t,\tau)/\sigma_\tau[c_{\lambda=0}](t)\right]$$

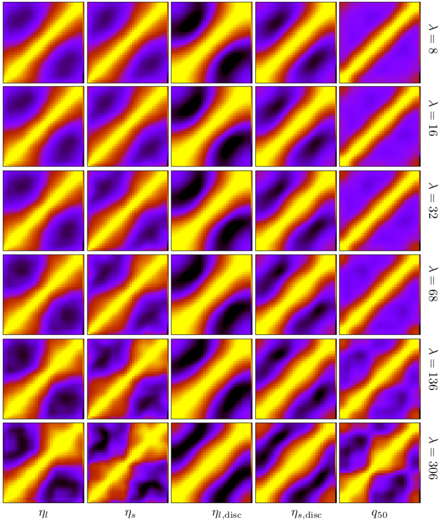
$\rightarrow$ between 20% and 40% left after IR-filtering

- smoothing introduces AC such that N/λ configurations remain independent
  - $\rightarrow$ fraction $O(1/\lambda)$ of the original information
  - neglecting the (small) AC of the UV-contributions

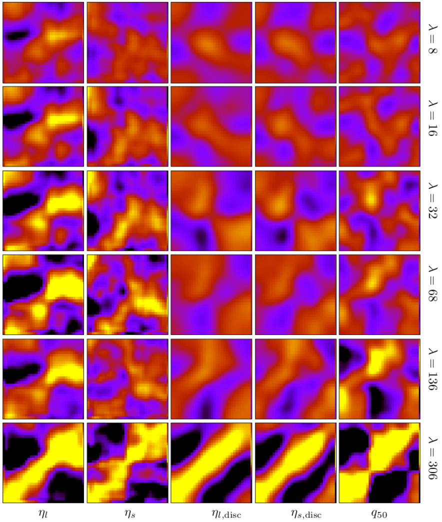
The latter effect should be dominant and explains why observables usually don't show oscillations.



Result2: coupling to (space)-time modes

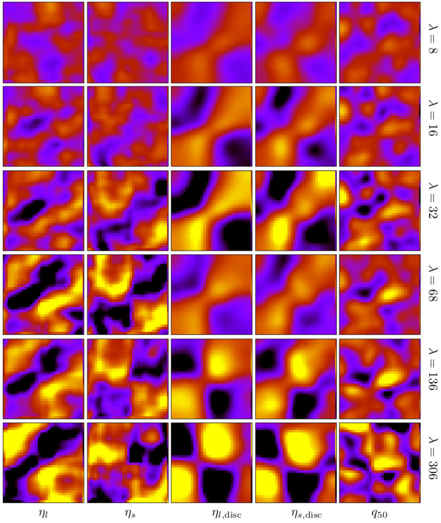


$\Delta\tau = 0$

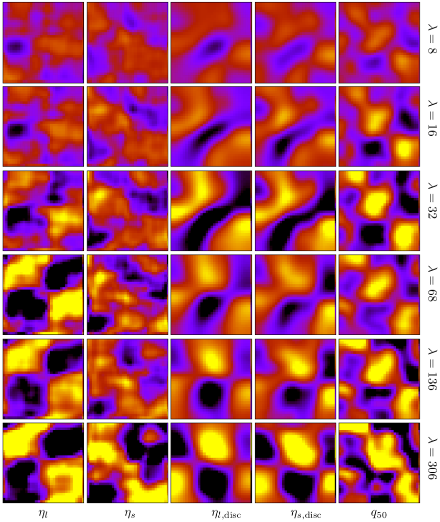


$\Delta\tau = 99$

Result2: coupling to (space)-time modes



$\Delta\tau = 540$



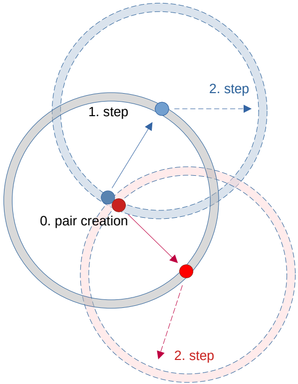
$\Delta\tau = 650$

HMC, a 5D-theory

The Hamiltonian used for Molecular Dynamics part of the HMC introduces canonical momenta P for the fields U and essentially defines a 5d-theory.
The momenta are defined by the choice of the kinetic term in H_{MD} and the Hamiltonian EOM.

$$H_{\text{MD}} \equiv \frac{1}{2}\sum_{x,\mu} P_{x,\mu}^2 + S_{\text{G}}[U] + S_{\text{PF}}[U,\phi,\phi^\dagger]$$

- AC oscillations originate from 5d-theory.
- theory includes QCD $\rightarrow$ potentially complex
- similar modes seen with Fourier Acceleration **[Sheta:2021hsd]**:
  - bco weak coupling limit & landau gauge fixing?
  - fields decouple & perform simple harmonic motions
  - showed oscillations for vec potentials A_μ
- not weakly coupled, here.



measure for pair annihilation after 2 updates
<< other configuration space

Randomization of momenta $P_{x,\mu} \propto e^{-P^2/2}$ at the beginning of each trajectory shall ensure ergodicity $\rightarrow$ identify bypass?

1. 2nd law of thermodynamics introduces arrow of time, nevertheless.
2. fixed trajectory length: system left in equivalent dynamical states \w lots of potential energy V & grad$V \neq 0$ @end of trajectory ...

Bibliography

See [Sheta:2021hsd] for related observations from Fourier Acceleration.

- Wolff, Ulli (2004). "Monte Carlo errors with less errors". In: *Comput. Phys. Commun.* 156. [Erratum: Comput.Phys.Commun. 176, 383 (2007)], pp. 143–153. DOI: 10.1016/S0010-4655(03)00467-3. arXiv: hep-lat/0306017.
- Luscher, Martin (2005). "Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD". In: *Comput. Phys. Commun.* 165, pp. 199–220. DOI: 10.1016/j.cpc.2004.10.004. arXiv: hep-lat/0409106.
- Jaynes, Edwin T. (1963). "Information Theory and Statistical Mechanics". In: *Brandeis University Summer Institute: In K. Ford (ed.). Statistical Physics (PDF). Benjamin, New York* 3, pp. 181–218. URL: http://bayes.wustl.edu/etj/articles/brandeis.pdf.
- Cinkir, Zubeyir (2011). *A fast elementary algorithm for computing the determinant of toeplitz matrices*. DOI: 10.48550/ARXIV.1102.0453. URL: https://arxiv.org/abs/1102.0453.
- Sheta, Ahmed, Yidi Zhao, and Norman H. Christ (2022). "Gauge-Fixed Fourier Acceleration". In: *PoS* LATTICE2021, p. 084. DOI: 10.22323/1.396.0084. arXiv: 2108.05486 [hep-lat].