Multigrid-Accelerated Low-Mode Averaging

Jakob Simeth[a], Gunnar Bali[a], Sara Collins[a], Andreas Frommer[b], Karsten Kahl[b], Issaku Kanamori[c], Matthias Rottmann[b]

[a]University of Regensburg – Institute for Theoretical Physics
[b]University of Wuppertal – Institute for Applied Mathematics
[c]Hiroshima University

July 16, 2015, Lattice Conference 2015, Kōbe

Universität Regensburg

Outline

Outline

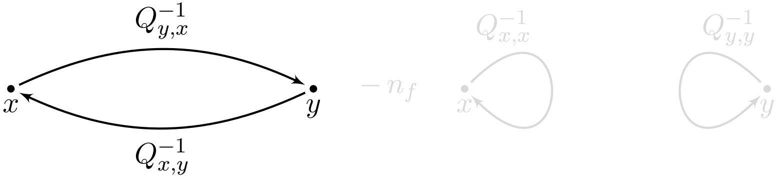
Introduction: Disconnected Contributions

- In Iso-singlet states quarks of equal flavour are coupled, e.g. in $n_f = 2$ for the η-meson

$$O^{\eta}(x) = \frac{1}{\sqrt{2}} \left(\bar{u}_x \gamma_5 u_x + \bar{d}_x \gamma_5 d_x \right).$$

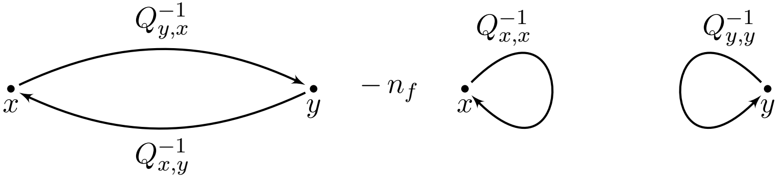
- This leads to additional terms in the correlation function, i.e. for degenerate light quark masses:

$$\begin{aligned} C_\eta(x,y) =& \langle O^\eta_x \bar{O}^\eta_y \rangle \\ =& \operatorname{tr}\left(\gamma_5 D^{-1}_{x,y} \gamma_5 D^{-1}_{x,y}\right) \\ & - n_f \operatorname{tr}\left(\gamma_5 D^{-1}_{x,x}\right) \operatorname{tr}\left(\gamma_5 D^{-1}_{y,y}\right) \end{aligned}$$



- The now appearing *all-to-all* propagators, $D^{-1}_{x,x}\gamma_5 = Q^{-1}_{x,x}$, are noisy and computationally challenging.
- Can be solved using stochastic sources, but this introduces additional noise on top of the gauge noise.
- *Noise reduction techniques* are mandatory.

Introduction: Disconnected Contributions



- The now appearing *all-to-all* propagators, $D^{-1}_{x,x}\gamma_5 = Q^{-1}_{x,x}$, are noisy and computationally challenging.
- Can be solved using stochastic sources, but this introduces additional noise on top of the gauge noise.
- *Noise reduction techniques* are mandatory.

Introduction: Low-Mode Averaging

- Split the (hermitianized) inverse Dirac Operator:

$$Q^{-1} = Q_{low}^{-1} + Q_{high}^{-1}$$

- Knowing the lowest N_{eig} eigenmodes of Q, we can write

$$Q_{low} = \sum_{i}^{N_{eig}} \lambda_i \mid u_i\rangle\langle u_i \mid .$$

- Then the low mode inverse is trivial

$$Q_{low}^{-1} = \sum_{i}^{N_{eig}} \frac{1}{\lambda_i} \mid u_i\rangle\langle u_i \mid$$

and we only have to invert the high mode contribution, e.g. using stochastic methods.

LMA works well and reduces the needed number of stochastic estimates, if the low modes dominate,

but: Calculation of the eigenmodes of large matrices is expensive by itself:

$$\text{cost} \propto V N_{eig}^2 .$$

Algorithmic developments needed:

- faster eigensolvers (see previous talk by M. Rottmann)
- (and) use approximate eigenmodes (this talk)

LMA works well and reduces the needed number of stochastic estimates, if the low modes dominate,

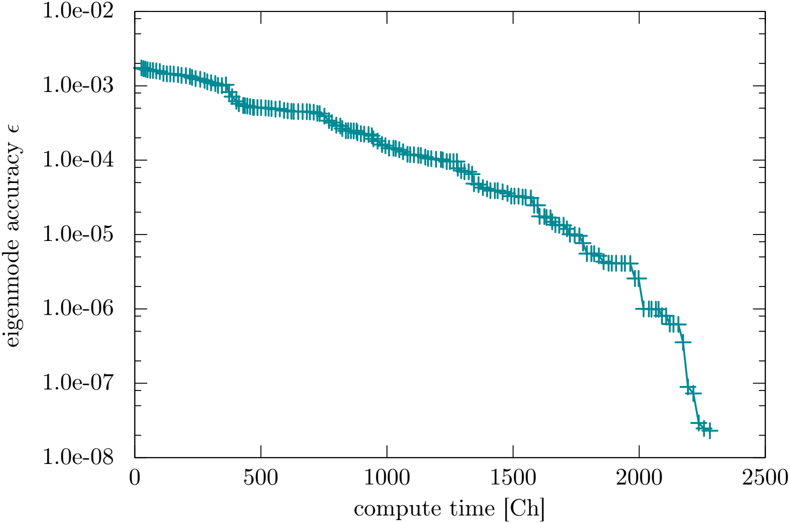
but: Calculation of the eigenmodes of large matrices is expensive by itself:

$$\text{cost} \propto V N_{eig}^2.$$

Algorithmic developments needed:

- faster eigensolvers (see previous talk by M. Rottmann)
- (and) use approximate eigenmodes (this talk)

Cost for the Calculation of Low Modes



Outline

Computational setup

- 64 configurations from the $n_f = 2$ QCDSF, $V = 40^3 \times 64$ ensemble at $m_\pi \approx 290\,\mathrm{MeV}$
- We use time dilution with a distance of $\Delta t = 4a$
- We use both APE Link and Wuppertal Quark Smearing of the Twopoint Functions.
- For the calculation of the connected part, we put one source per configuration at a random position.
- The high mode part, Q_{high}, is inverted on stochastic $\mathbb{Z}_2 + i\mathbb{Z}_2$ random sources.

From the multigrid eigensolver, we get

- N_{eig} *eigenvalues* λ_i and approximate orthonormal *eigenvectors* $\mid u_i\rangle$ of Q,
- We define the *accuracy* ϵ_i as

$$\epsilon_i = \|Q \mid u_i\rangle - \lambda_i \mid u_i\rangle\| \,.$$

- If $\epsilon \leq 10^{-8}$, we call the eigenmode *exact*. We use $N_{eig} = 20$ of these, here.
- We take $N_{eig} = 30$ test vectors of the Multigrid directly after the setup with an accuracy of typically $\epsilon \approx 10^{-3}$ as approximate eigenmodes.

Low-Mode averaging for the connected twopoint function:

$$C_{2pt}(\delta t) = C_{2pt}^{low}(\delta t) + \left(C_{2pt}^{p2a}(\delta t) - C_{2pt}^{low,p2a}(\delta t)\right),$$

where, in the exact case (with $B_{ij}(x) = {}_x\langle u_i \mid u_j\rangle_x$)

$$C_{2pt}^{p2a}(\delta t) = \frac{1}{N_t} \sum_x \mathrm{tr}\left[\left(Q^{-1}_{x,y_0}\right)^2\right]$$

$$C_{2pt}^{low}(\delta t) = \frac{1}{V} \sum_{i,j,\mathbf{x},\mathbf{y},t} \frac{1}{\lambda_i \lambda_j} \mathrm{tr}\left[B_{ij}(y) B_{ji}(x)\right]$$

$$C_{2pt}^{low,p2a}(\delta t) = \frac{1}{N_t} \sum_{i,j,\mathbf{x},t} \frac{1}{\lambda_i \lambda_j} \mathrm{tr}\left[B_{ij}(y_0) B_{ji}(x)\right]$$

(with $x = (\mathbf{x}, t)$, $y = (\mathbf{y}, t + \delta t)$ and $y_0 = (\mathbf{y}_0, t + \delta t)$)

Using approximate low modes for the connected part

Low-Mode averaging for the connected twopoint function:

$$C_{2pt}(\delta t) = C_{2pt}^{low}(\delta t) + \left(C_{2pt}^{p2a}(\delta t) - C_{2pt}^{low,p2a}(\delta t)\right),$$

where for the inexact case (with $B_{ij}(x) = {}_x\langle u_i \mid u_j \rangle_x$ and $A = \langle u_i \mid Q \mid u_j \rangle$)

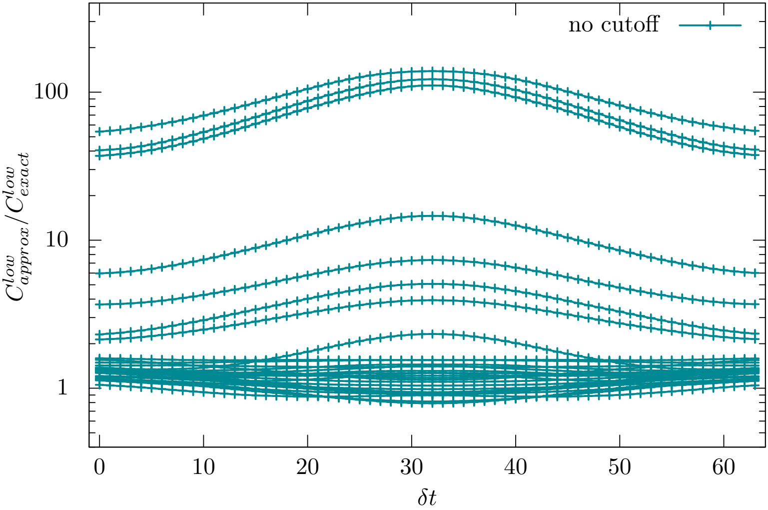
$$C_{2pt}^{p2a}(\delta t) = \frac{1}{N_t} \sum_x \mathrm{tr}\left[\left(Q_{x,y_0}^{-1}\right)^2\right]$$

$$C_{2pt}^{low}(\delta t) = \frac{1}{V} \sum_{\mathbf{x},\mathbf{y},t} \mathrm{tr}\left[A^{-1}B(y)A^{-1}B(x)\right]$$

$$C_{2pt}^{low,p2a}(\delta t) = \frac{1}{N_t} \sum_{\mathbf{x},t} \mathrm{tr}\left[A^{-1}B(y_0)A^{-1}B(x)\right]$$

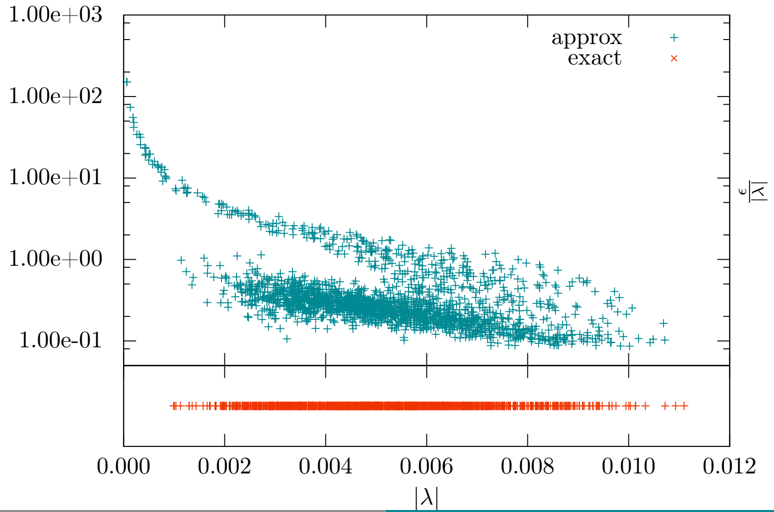
(with $x = (\mathbf{x}, t)$, $y = (\mathbf{y}, t + \delta t)$ and $y_0 = (\mathbf{y}_0, t + \delta t)$ and implied matrix multiplication of A^{-1} and B)

Naive Ratios of connected low-mode correlators



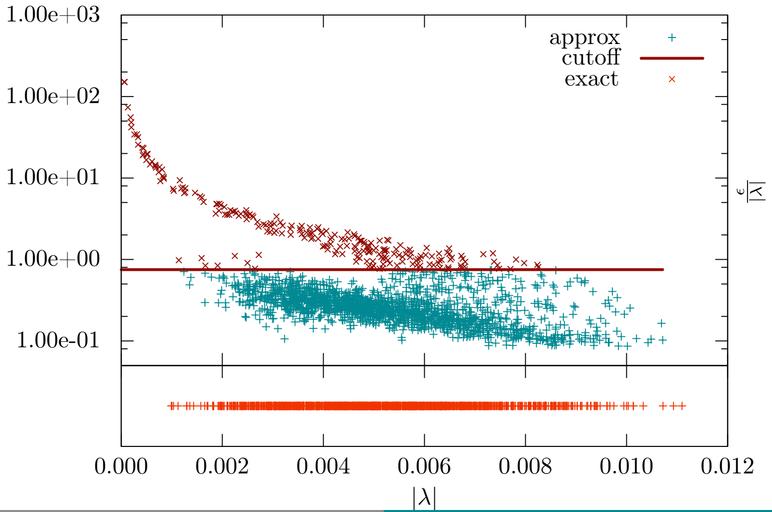
Eigenvalue spectrum

64 configurations, 30 lowest approximate eigenvalues each.



Eigenvalue spectrum

64 configurations, 30 lowest approximate eigenvalues each.



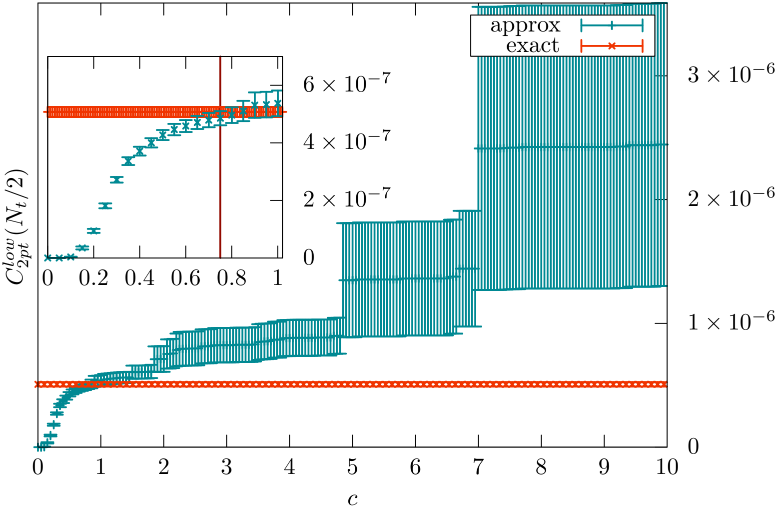
Choice of a cutoff

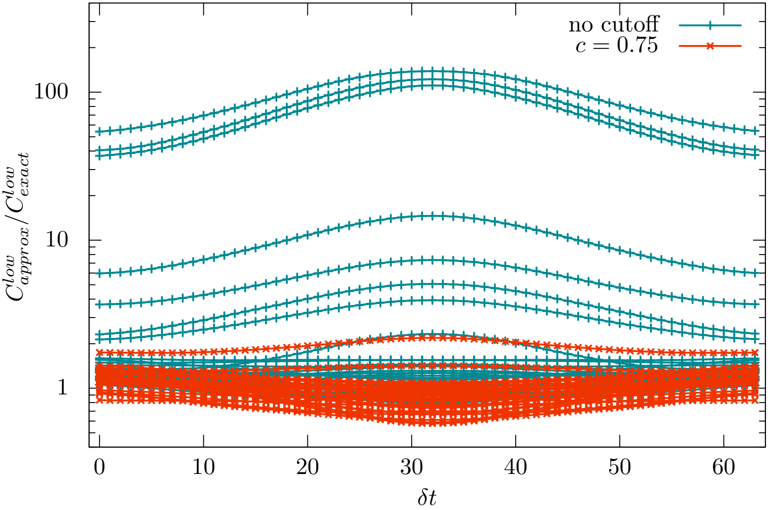
- cutoff criterion:

$$\{|\, u\rangle, \lambda, \epsilon\} \to \left\{|\, u\rangle, \lambda, \epsilon \quad \middle| \quad \frac{\epsilon}{\lambda} \leq c\right\},$$

  where we choose $c = 0.75$.
- This criterion is independent of lattice volume and quark mass.
- From the 30 test vectors per configuration, approximately 25 “survive”
- Other choices possible

Choice of a cutoff




no cutoff
$c = 0.75$
100
10
1
0
10
20
30
40
50
60
δt
$C^{low}_{approx}/C^{low}_{exact}$

Using approximate low modes for the connected part

Low-Mode averaging for the connected twopoint function:

$$C_{2pt}(\delta t) = C_{2pt}^{low}(\delta t) + \left(C_{2pt}^{p2a}(\delta t) - C_{2pt}^{low,p2a}(\delta t) \right),$$

where for the inexact case (with $B_{ij}(x) = {}_x\langle u_i \mid u_j \rangle_x$ and $A = \langle u_i \mid Q \mid u_j \rangle$)

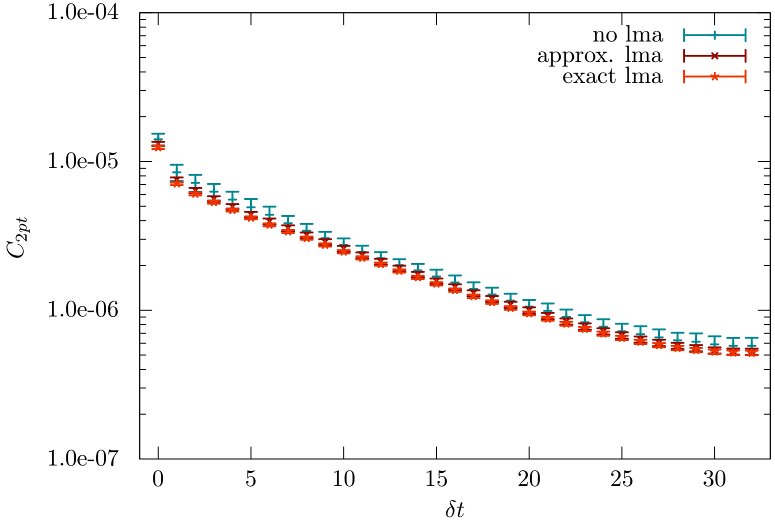
$$C_{2pt}^{p2a}(\delta t) = \frac{1}{N_t} \sum_x \mathrm{tr}\left[\left(Q_{x,y_0}^{-1} \right)^2 \right]$$

$$C_{2pt}^{low}(\delta t) = \frac{1}{V} \sum_{\mathbf{x},\mathbf{y},t} \mathrm{tr}\left[A^{-1} B(y) A^{-1} B(x) \right]$$

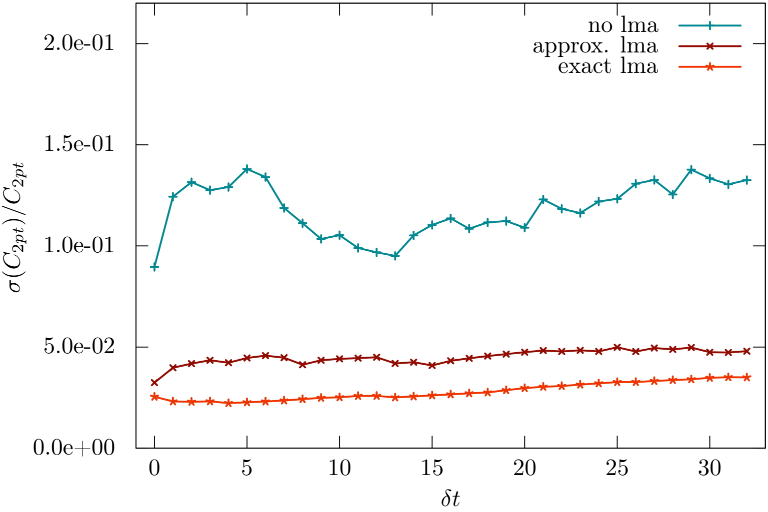
$$C_{2pt}^{low,p2a}(\delta t) = \frac{1}{N_t} \sum_{\mathbf{x},t} \mathrm{tr}\left[A^{-1} B(y_0) A^{-1} B(x) \right]$$

(with $x = (\mathbf{x}, t)$, $y = (\mathbf{y}, t + \delta t)$ and $y_0 = (\mathbf{y}_0, t + \delta t)$ and implied matrix multiplication of A^{-1} and B)

Using approximate low modes for the connected part



Using approximate low modes for the connected part



Using exact low modes for the disconnected part

- Correlator:

$$C_{dis}(\delta t) = \frac{1}{N_t} \sum_t L(t) L(t+\delta t)$$

- Disconnected loop:

$$L(t) = \sum_{\mathbf{x}} \mathrm{tr}\left[Q^{-1}_{x,x} \right] = L^{low}(t) + L^{high}(t)$$

- Calculation of the exact low and high part:

$$L^{low}(t) = \sum_{i,\mathbf{x}} \frac{1}{\lambda_i} \mathrm{tr}\left[{}_x\langle u_i \mid u_i \rangle_x \right]$$

$$L^{high}(t) = \sum_{\mathbf{x}} \mathrm{tr}\left[(\mathcal{P}_{exact} Q)^{-1}_{x,x} \right],$$

- with a suitable projector $\mathcal{P}_{exact}$:

$$\mathcal{P}_{exact} = 1 - \sum_i^{N_{eig}} \mid u_i \rangle \langle u_i \mid.$$

Using approximate low modes for the disconnected part

- Correlator:

$$C_{dis}(\delta t) = \frac{1}{N_t} \sum_t L(t) L(t + \delta t)$$

- Disconnected loop:

$$L(t) = \sum_{\mathbf{x}} \mathrm{tr}\left[Q^{-1}_{x,x}\right] = L^{low}(t) + L^{high}(t)$$

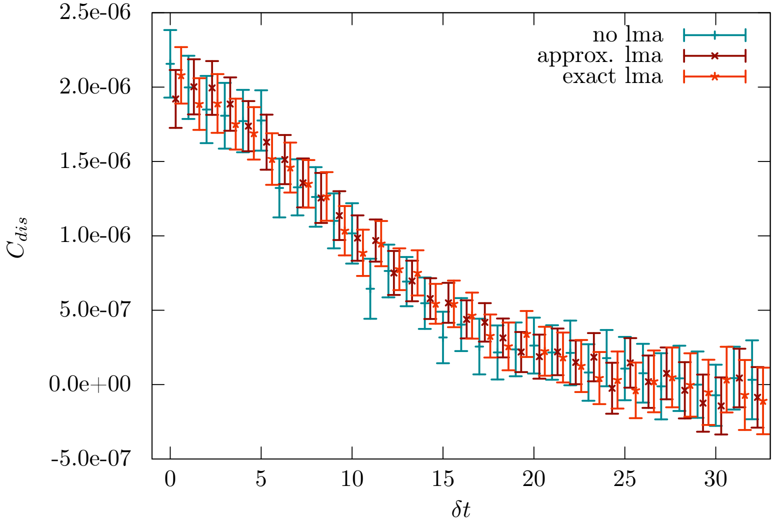
- The low and high part for approximate eigenmodes:

$$L^{low}(t) = \sum_{i,j,\mathbf{x}} \mathrm{tr}\left[{}_x\langle u_i \mid A^{-1}_{ij} \mid u_j\rangle_x\right]$$

$$L^{high}(t) = \sum_{\mathbf{x}} \mathrm{tr}\left[(\mathcal{R}_{approx} Q)^{-1}_{x,x}\right],$$

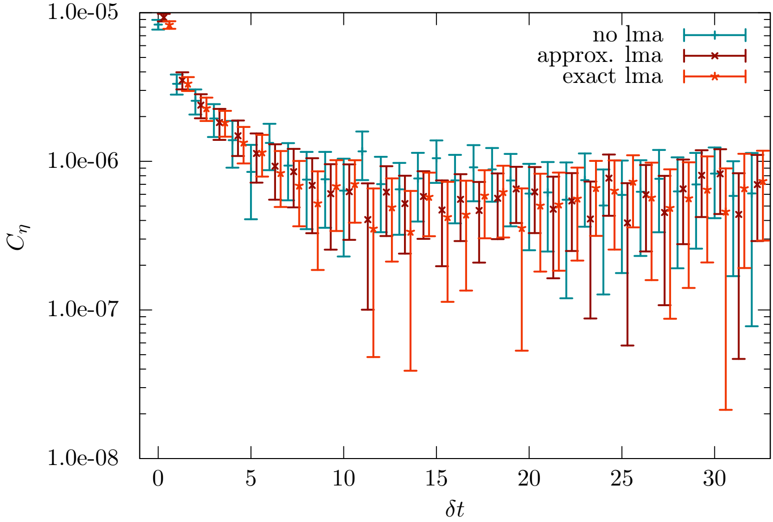
- with a suitable restrictor $\mathcal{R}_{approx}$

$$\mathcal{R}_{approx} = 1 - \sum_{i,j}^{N_{eig}} Q \mid u_i\rangle A^{-1}_{ij} \langle u_j \mid .$$



C_{dis}
2.5e-06
2.0e-06
1.5e-06
1.0e-06
5.0e-07
0.0e+00
-5.0e-07
0
5
10
15
20
25
30
δt
no lma
approx. lma
exact lma

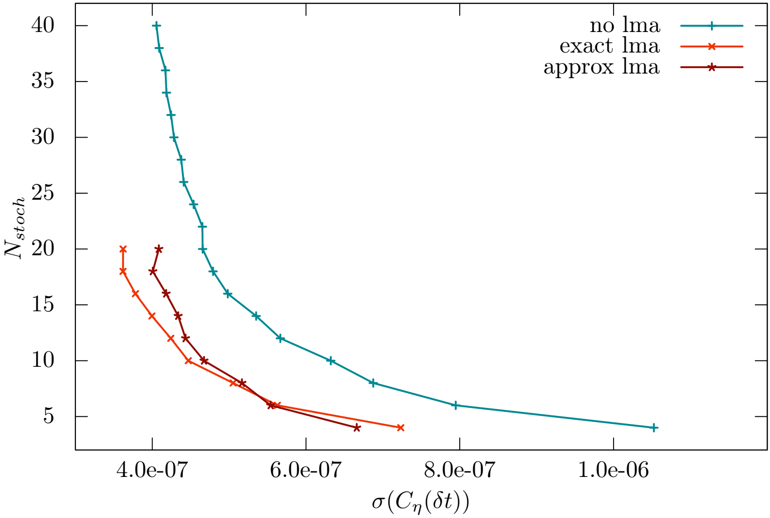
Full Correlator from (approximate) LMA



Outline

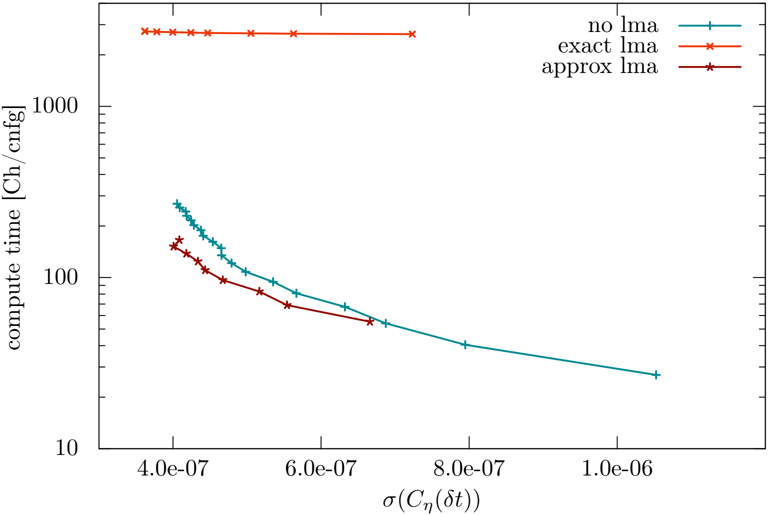
What it is all about: Cost Reduction

Error averaged over the first (last) 10 time slices.



What it is all about: Cost Reduction

Error averaged over the first (last) 10 time slices.



Outline

Conclusions and Outlook

This is work in progress:

- higher statistics
- Investigate the influence of the number of eigenmodes
- Integration of the Multgrid into our measurement program: Use the setup also for the inversion.
- Different observables, e.g. threepoint functions
- Results at physical pion mass (not shown here) looks promising, but needs larger statistics.

But...

- Developed improvement techniques which allow for the use of approximate eigenmodes
- Similar Errors than when using exact eigenmodes
- Approximate LMA gives a speedup of more than 10.