Multigrid Accelerated Simulations

results for Twisted Mass fermions at the physical point

Simone Bacchio[a,b], Constantia Alexandrou[a,c], Jacob Finkenrath[b], Bartosz Kostrzewa[d]

[a]University of Cyprus
[b]University of Wuppertal
[c]The Cyprus Institute
[d]University of Bonn

Lattice2017
35th annual International Symposium on Lattice Field Theory

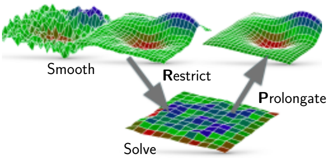
Overview

Tools:

- tmLQCD [K. Jansen, C. Urbach, 2009]
  - HMC for TM
  - Fermionic and gluonic observ.
  - Inverters
  - Rational HMC

European Twisted Mass Collaboration

- DD-αAMG [A. Frommer et al. 2013]



Covered topics:

- DD-αAMG in the HMC
  - optimizing setup time
  - stability in the traj.
  - adapting to ND operator
- $N_f = 2$ and $N_f = 2+1+1$
- down to the physical point

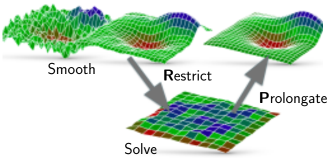
Simulations:

- $N_f = 2$
  - 64c128 phys.
- $N_f = 2+1+1$
  - 64c128 phys. [J. Finkenrath, Tue.]
  - 48c96 phys.
  - 32c64 non-phys.

Overview

Tools:

- tmLQCD [K. Jansen, C. Urbach, 2009]
  - HMC for TM
  - Fermionic and gluonic observ.
  - Inverters
  - Rational HMC
- DD-αAMG [A. Frommer et al. 2013]



Covered topics:

- DD-αAMG in the HMC
  - optimizing setup time
  - stability in the traj.
  - adapting to ND operator
- $N_f = 2$ and $N_f = 2+1+1$
- down to the physical point

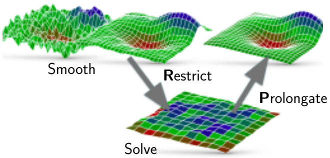
Simulations:

- $N_f = 2$
  - 64c128 phys.
- $N_f = 2+1+1$
  - 64c128 phys. [J. Finkenrath, Tue.]
  - 48c96 phys.
  - 32c64 non-phys.

Overview

Tools:

- tmLQCD [K. Jansen, C. Urbach, 2009]
  - HMC for TM
  - Fermionic and gluonic observ.
  - Inverters
  - Rational HMC
- DD-αAMG [A. Frommer et al. 2013]



Covered topics:

- DD-αAMG in the HMC
  - optimizing setup time
  - stability in the traj.
  - adapting to ND operator
- $N_f = 2$ and $N_f = 2+1+1$
- down to the physical point

Simulations:

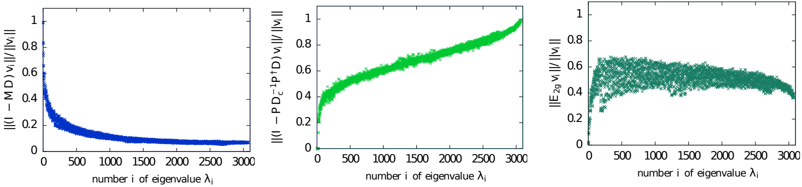
- $N_f = 2$
  - 64c128 phys.
- $N_f = 2+1+1$
  - 64c128 phys. [J. Finkenrath, Tue.]
  - 48c96 phys.
  - 32c64 non-phys.

DD-αAMG in a nutshell

It is a multigrid solver for Wilson-like fermions suitable for any $D = D(m, \mu, c_{sw})$.

- The coarse operator is $D_c = P^\dagger DP$.
- The prolongation operator P is constructed such that
  - preserves the sparsity of D,
  - projects into the coarse operator the IR modes of D,
  - preserves the Γ_5-hermiticity of D.
- P is constructed in an adaptive setup procedure.
- The smoother is Schwarz Alternating Procedure (SAP). [M. Lüscher, 2003]

Smoother (SAP) $+$ Coarse grid correction $=$ DD-αAMG



DD-αAMG in a nutshell

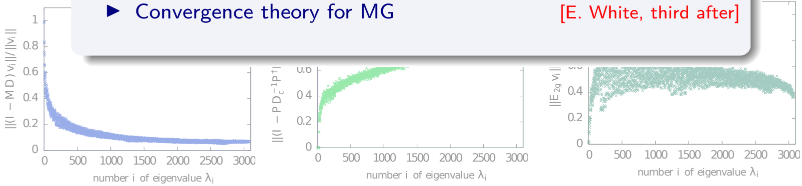
It is a multigrid solver for Wilson-like fermions suitable for any $D = D(m, \mu, c_{sw})$.

For more details

- DD-αAMG [A. Frommer et al. 2013]
- DD-αAMG for TM fermions [C. Alexandrou et al. 2016]
- DD-αAMG vs QUDA MG [S. Bacchio, Fri]

More about multigrid

- QUDA MG [K. Clark, just after]
- MG for staggered fermions [E. Weinberg, second after]
- Convergence theory for MG [E. White, third after]



HMC simulations for $N_f = 2$

The $N_f = 2$ TM operator is given by $D \pm i\Gamma_5\mu$ which represents the up and down flavors. Thus in the simulations the following weight is included

$$\det\left[D^\dagger D + \mu^2\right] = \frac{1}{\pi^N}\int \mathcal{D}(\phi)\exp\left\{-\phi^\dagger(D^\dagger D + \mu^2)^{-2}\phi\right\}.$$

We use Hasenbusch mass preconditioning [Hasenbusch, 2001]

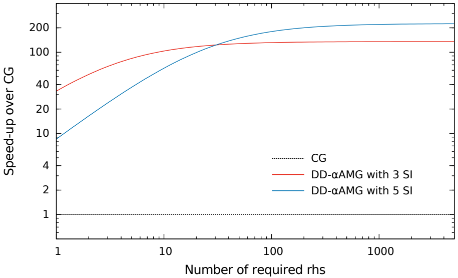
$$\det\left[Q^2 + \mu^2\right] = \det\left[\frac{Q^2 + \mu^2}{Q^2 + \mu^2 + \rho_1^2}\right]\det\left[\frac{Q^2 + \mu^2 + \rho_1^2}{Q^2 + \mu^2 + \rho_2^2}\right]\ldots\det\left[Q^2 + \mu^2 + \rho_n^2\right]$$

which helps to suppress IR-noise in the force terms by splitting up the determinant in determinant of ratios using operators with higher masses.

Each determinant can be placed on a different monomial and it can be integrated on different time-scales. This gives control on the large fluctuation of the force terms, avoiding instabilities during the HMC.

Speed-up at the physical point

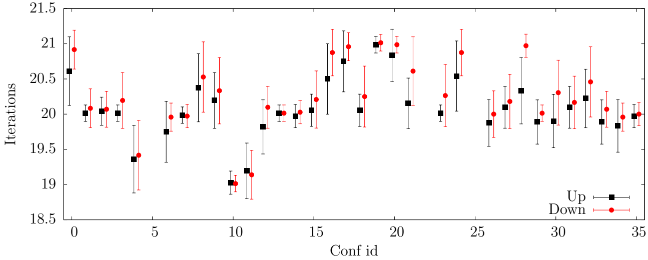
In the HMC we need a fast setup since we do few inversions per configuration.



Solver	Setup time [core-hrs]	Inversion time [core-hrs]	Solver iterations
CG	–	338.6	34 790
DD-αAMG with 3 setup iters	7.7	2.5	28
DD-αAMG with 5 setup iters	38.3	1.5	15

- Three initial setup iterations are enough for the best performance.

Stability of the solver - part 1

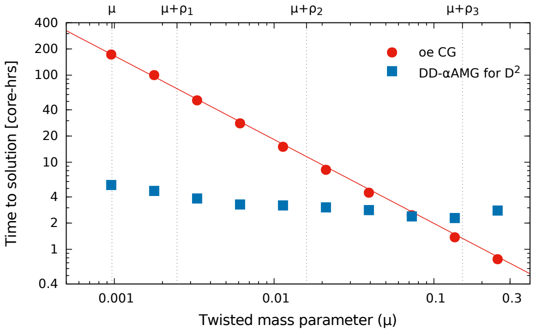


The setup is done for $+\mu$ and the same P is used for both signs

$$D_c(\mu) = P^\dagger(D_w + i\Gamma_5\mu)P \quad \text{and} \quad D_c(-\mu) = P^\dagger(D_w - i\Gamma_5\mu)P.$$

- performance is not affected,
- results are stable when 3-level are used,
- a 2-level approach has large fluctuations at the physical point.

Speed-up against oeCG at different masses

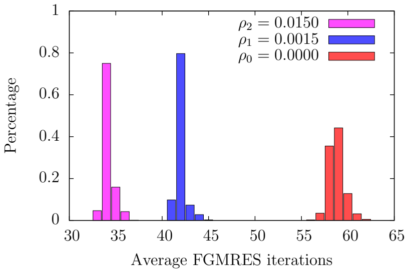


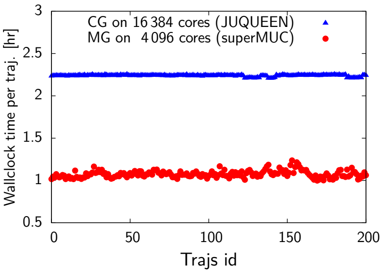
$$\det\left[\frac{Q^2+\mu^2}{Q^2+\mu^2+\rho_1^2}\right]\det\left[\frac{Q^2+\mu^2+\rho_1^2}{Q^2+\mu^2+\rho_2^2}\right]\det\left[\frac{Q^2+\mu^2+\rho_2^2}{Q^2+\mu^2+\rho_3^2}\right]\det\left[Q^2+\mu^2+\rho_3^2\right]$$

- The inversion of $Q^2+\mu^2+\rho_3^2$ is performed with oeCG.
- The other inversions are done by DD-αAMG.

μ	0.0009
ρ_1	0.0015
ρ_2	0.015
ρ_3	0.15

Stability of the solver - part 2





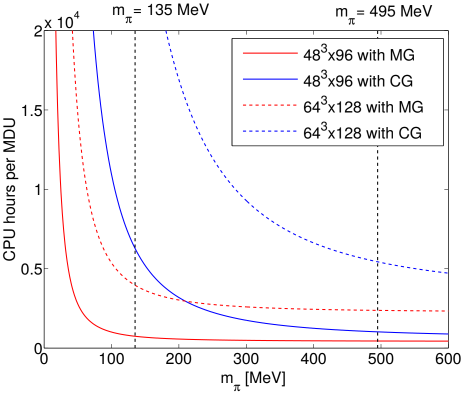
- The same P is used for all the operators of the kind $D \pm i\Gamma_5(\mu + \rho_i)$.
- The setup is updated when $Q^2 + \mu^2$ is inverted.
  Then one setup iteration is sufficient.
- For $\rho_i > 0$, P is not updated $\Rightarrow$ no performance lost observed.
- Small fluctuations in the traj. time (5 %) are due to the coarse grid
  $\Rightarrow$ fine grid iterations are stable.

HMC setup for $N_f = 2$

Software	tmLQCD interfaced to DDalphaAMG	[github.com/etmc/tmLQCD] [github.com/sbacchio/DDalphaAMG]
Integrator	nested OMF-scheme order 2	[I.P. Omelyan et al. 2003]
Monomials	4 (fermionic) + 1 (gauge) on 5 different timescales with int. scheme $N_{int} = \{20, 40, 80, 160, 320\}$	[A. Abdel-Rehim et al. 2015]
Solvers	DD-αAMG (x3) + oeCG (x1)	
MG setup	3 setup iterations at the beginning of the traj. 1 setup iteration on the outer monomial (42 times per traj.)	
Accuracy	10^{-9} DD-αAMG in the force term 10^{-7} oeCG in the force term 10^{-11} in the heat-bath and acceptance step	

HMC speed-up for $N_f = 2$

Multigrid in the HMC gives a speed-up of a factor 8 compared to oeCG by tmLQCD using TM fermions at the physical point with $N_f = 2$.

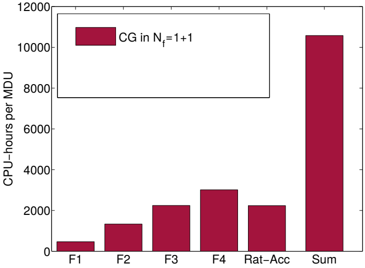


[432 Haswell Xeon Processors E5-2697 on SuperMUC]

Running simulation for $N_f = 2 + 1 + 1$

In the TM discretization the heavy sector, $N_f = 1 + 1$, is threated with a non-degenerate operator non-diagonal in flavor space. This is done in order to ensure a real determinant.

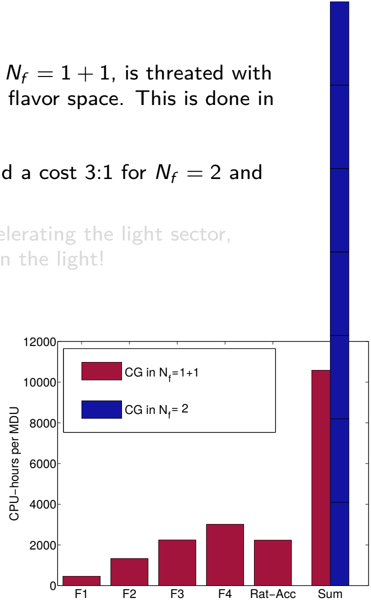
- Simulations performed with oeCG had a cost 3:1 for $N_f = 2$ and $N_f = 1 + 1$.
- Now that DD-αAMG is used for accelerating the light sector, the heavy sector costs 2.5x more than the light!



Running simulation for $N_f = 2 + 1 + 1$

In the TM discretization the heavy sector, $N_f = 1 + 1$, is threated with a non-degenerate operator non-diagonal in flavor space. This is done in order to ensure a real determinant.

- Simulations performed with oeCG had a cost 3:1 for $N_f = 2$ and $N_f = 1 + 1$.
- Now that DD-αAMG is used for accelerating the light sector, the heavy sector costs 2.5x more than the light!

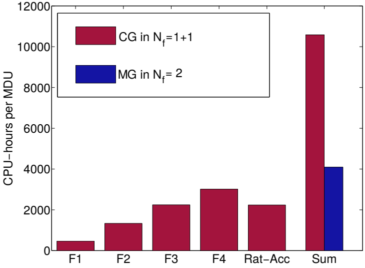


Running simulation for $N_f = 2 + 1 + 1$

In the TM discretization the heavy sector, $N_f = 1 + 1$, is threated with a non-degenerate operator non-diagonal in flavor space. This is done in order to ensure a real determinant.

- Simulations performed with oeCG had a cost 3:1 for $N_f = 2$ and $N_f = 1 + 1$.
- Now that DD-αAMG is used for accelerating the light sector, the heavy sector costs 2.5x more than the light!

Can MG speed-up the heavy sector?



DD-αAMG for non-degenerate operator

The non-degenerate TM operator is

$$D_{ND}(\bar{\mu},\bar{\epsilon}) = D_W \otimes \mathbb{1}_f + i\bar{\mu}\Gamma_5 \otimes \tau_3 - \bar{\epsilon}\,\mathbb{1} \otimes \tau_1 = \begin{bmatrix} D + i\bar{\mu}\Gamma_5 & -\bar{\epsilon}\,\mathbb{1} \\ -\bar{\epsilon}\,\mathbb{1} & D - i\bar{\mu}\Gamma_5 \end{bmatrix}$$

where $\mathbb{1}_f, \tau_3, \tau_1$ act in the flavor space. The operator is $(\Gamma_5 \otimes \tau_1)$-hermitian, having

$$Q_{ND}(\bar{\mu},\bar{\epsilon}) = \tau_1 D_{ND}(\bar{\mu},\bar{\epsilon})\Gamma_5 = (\tau_1 D_{ND}(\bar{\mu},\bar{\epsilon})\Gamma_5)^\dagger\,.$$

As we preserve the Γ_5-hermiticity of D, we want to construct a multigrid method which preserve the $(\Gamma_5 \otimes \tau_1)$-hermiticity of D_{ND}. This leads us to the definition of

$$\begin{aligned} D_{ND,c}(\bar{\mu},\bar{\epsilon}) &= \begin{bmatrix} P^\dagger & \\ & P^\dagger \end{bmatrix} \begin{bmatrix} D + i\bar{\mu}\Gamma_5 & -\bar{\epsilon}\,\mathbb{1} \\ -\bar{\epsilon}\,\mathbb{1} & D - i\bar{\mu}\Gamma_5 \end{bmatrix} \begin{bmatrix} P & \\ & P \end{bmatrix} \\ &= \begin{bmatrix} D_c + i\bar{\mu}\Gamma_{5,c} & -\bar{\epsilon}\,\mathbb{1}_c \\ -\bar{\epsilon}\,\mathbb{1}_c & D_c - i\bar{\mu}\Gamma_{5,c} \end{bmatrix} \end{aligned}$$

where P is the prolongation operator constructed for $D \pm i\bar{\mu}\Gamma_5$.

In the HMC we will use the P constructed for the light sector
$\Rightarrow$ motivated by the performance for $\mu + \rho_i$ with $\rho_i > 0$.

DD-αAMG for non-degenerate operator

The non-degenerate TM operator is

$$D_{ND}(\bar{\mu},\bar{\epsilon}) = D_W \otimes \mathbb{1}_f + i\bar{\mu}\Gamma_5 \otimes \tau_3 - \bar{\epsilon}\,\mathbb{1} \otimes \tau_1 = \begin{bmatrix} D + i\bar{\mu}\Gamma_5 & -\bar{\epsilon}\,\mathbb{1} \\ -\bar{\epsilon}\,\mathbb{1} & D - i\bar{\mu}\Gamma_5 \end{bmatrix}$$

where $\mathbb{1}_f, \tau_3, \tau_1$ act in the flavor space. The operator is $(\Gamma_5 \otimes \tau_1)$-hermitian, having

$$Q_{ND}(\bar{\mu},\bar{\epsilon}) = \tau_1 D_{ND}(\bar{\mu},\bar{\epsilon})\Gamma_5 = (\tau_1 D_{ND}(\bar{\mu},\bar{\epsilon})\Gamma_5)^\dagger \,.$$

As we preserve the Γ_5-hermiticity of D, we want to construct a multigrid method which preserve the $(\Gamma_5 \otimes \tau_1)$-hermiticity of D_{ND}. This leads us to the definition of

$$\begin{aligned} D_{ND,c}(\bar{\mu},\bar{\epsilon}) &= \begin{bmatrix} P^\dagger & \\ & P^\dagger \end{bmatrix} \begin{bmatrix} D + i\bar{\mu}\Gamma_5 & -\bar{\epsilon}\,\mathbb{1} \\ -\bar{\epsilon}\,\mathbb{1} & D - i\bar{\mu}\Gamma_5 \end{bmatrix} \begin{bmatrix} P & \\ & P \end{bmatrix} \\ &= \begin{bmatrix} D_c + i\bar{\mu}\Gamma_{5,c} & -\bar{\epsilon}\,\mathbb{1}_c \\ -\bar{\epsilon}\,\mathbb{1}_c & D_c - i\bar{\mu}\Gamma_{5,c} \end{bmatrix} \end{aligned}$$

where P is the prolongation operator constructed for $D \pm i\bar{\mu}\Gamma_5$.

In the HMC we will use the P constructed for the light sector
$\Rightarrow$ motivated by the performance for $\mu + \rho_i$ with $\rho_i > 0$.

HMC simulations for $N_f = 1 + 1$

Even if the $(\Gamma_5 \otimes \tau_1)$-hermiticity ensures that $\det[D_{ND}]$ is real, for the HMC an hermitian operator is necessary. Thus

$$\det[D_{ND}] = \det\left[\sqrt{Q_{ND} Q_{ND}}\right]$$

is commonly used in the simulations and a good approach for computing the square root is the the rational approximation [M. A. Clark, A. Kennedy, 2006].

The rational approximation reads as

$$\mathcal{R}(Q_{ND}^2) = \prod_{i=1}^{N} \frac{Q_{ND}^2 + \mu_i^2}{Q_{ND}^2 + \nu_i^2} \simeq \frac{1}{\sqrt{Q_{ND}^2}}$$

where the μ_i and ν_i are given by the Zolotarev solution for the optimal rational approximation to $1/\sqrt{y}$.

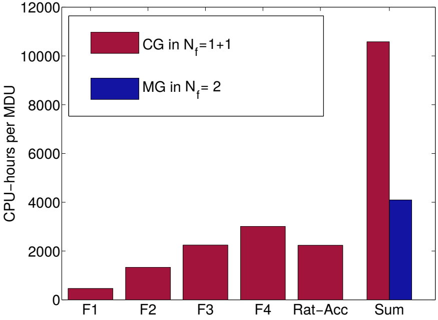
HMC speed-up for $N_f = 2 + 1 + 1$

In the considered case we have used a rational approximation of order 10 and

$$\det\left[\sqrt{Q_{ND}^2}\right] \simeq \prod_{i=1}^{10} \det\left[\frac{Q_{ND}^2 + \nu_i^2}{Q_{ND}^2 + \mu_i^2}\right]$$

has been split in 4 timescales (F1-4).

- For the timescale with smaller shift ν_i (F3-4) MG gives a speed-up. For the other timescales MMS-CG is used.
- In the correction term to the rational approximation (Rat) and in the acceptance step (Acc) a mixed approach MMS-CG + MG is used.



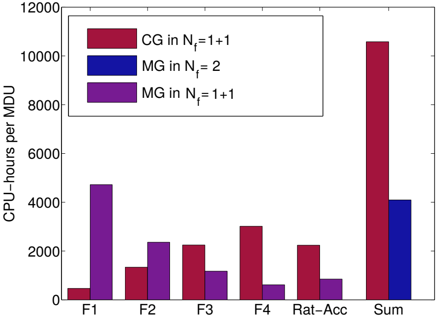
HMC speed-up for $N_f = 2 + 1 + 1$

In the considered case we have used a rational approximation of order 10 and

$$\det\left[\sqrt{Q_{ND}^2}\right] \simeq \prod_{i=1}^{10} \det\left[\frac{Q_{ND}^2 + \nu_i^2}{Q_{ND}^2 + \mu_i^2}\right]$$

has been split in 4 timescales (F1-4).

- For the timescale with smaller shift ν_i (F3-4) MG gives a speed-up. For the other timescales MMS-CG is used.
- In the correction term to the rational approximation (Rat) and in the acceptance step (Acc) a mixed approach MMS-CG + MG is used.



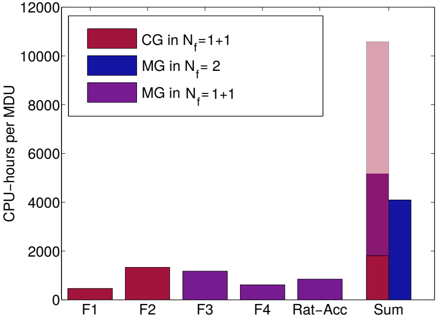
HMC speed-up for $N_f = 2 + 1 + 1$

In the considered case we have used a rational approximation of order 10 and

$$\det\left[\sqrt{Q_{ND}^2}\right] \simeq \prod_{i=1}^{10} \det\left[\frac{Q_{ND}^2 + \nu_i^2}{Q_{ND}^2 + \mu_i^2}\right]$$

has been split in 4 timescales (F1-4).

- For the timescale with smaller shift ν_i (F3-4) MG gives a speed-up. For the other timescales MMS-CG is used.
- In the correction term to the rational approximation (Rat) and in the acceptance step (Acc) a mixed approach MMS-CG + MG is used.



HMC speed-up for $N_f = 2 + 1 + 1$

In the considered case we have used a rational approximation of order 10 and

$$\det\left[\sqrt{Q_{ND}^2}\right] \simeq \prod_{i=1}^{10} \det\left[\frac{Q_{ND}^2 + \nu_i^2}{Q_{ND}^2 + \mu_i^2}\right]$$

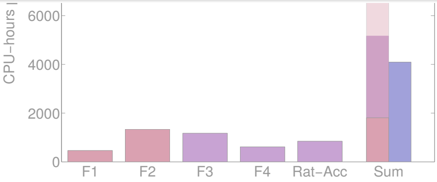
has been split in 4 timescales (F1-4).

Conclusions:

HMC simulations

- for $N_f = 2$ have been accelerated by a factor of 8,
- for $N_f = 1 + 1$ have been accelerated by a factor of 2,
- for $N_f = 2 + 1 + 1$ have been accelerated by a factor of 5.

A 3-level approach is needed for stable time in the HMC at the physical point.



Thank you for you attention!

and I kindly thank

My supervisors	Prof. C. Alexandrou and Prof. A. Frommer
Horizon 2020	for founding under the Marie Sklodowska-Curie grant agreement No 642069.
R. Frezzotti	for fruitful discussions.
C. Urbach	for helping with tmLQCD development.
Facilities:	SuperMUC at the Munich Supercomputing Centers.

Multigrid Accelerated Simulations

results for Twisted Mass fermions at the physical point

Simone Bacchio[a,b], Constantia Alexandrou[a,c], Jacob Finkenrath[b], Bartosz Kostrzewa[d]

[a]University of Cyprus
[b]University of Wuppertal
[c]The Cyprus Institute
[d]University of Bonn

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No' 642069.