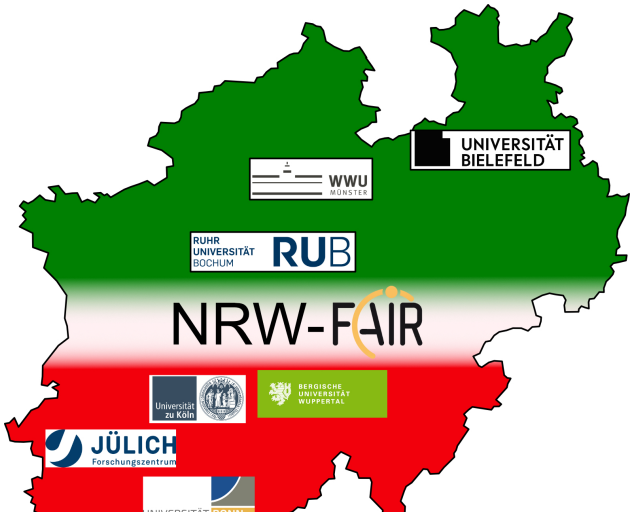
Experiences with the Karsten-Wilczek solver

Szabolcs Borsanyi, Chik Him Wong

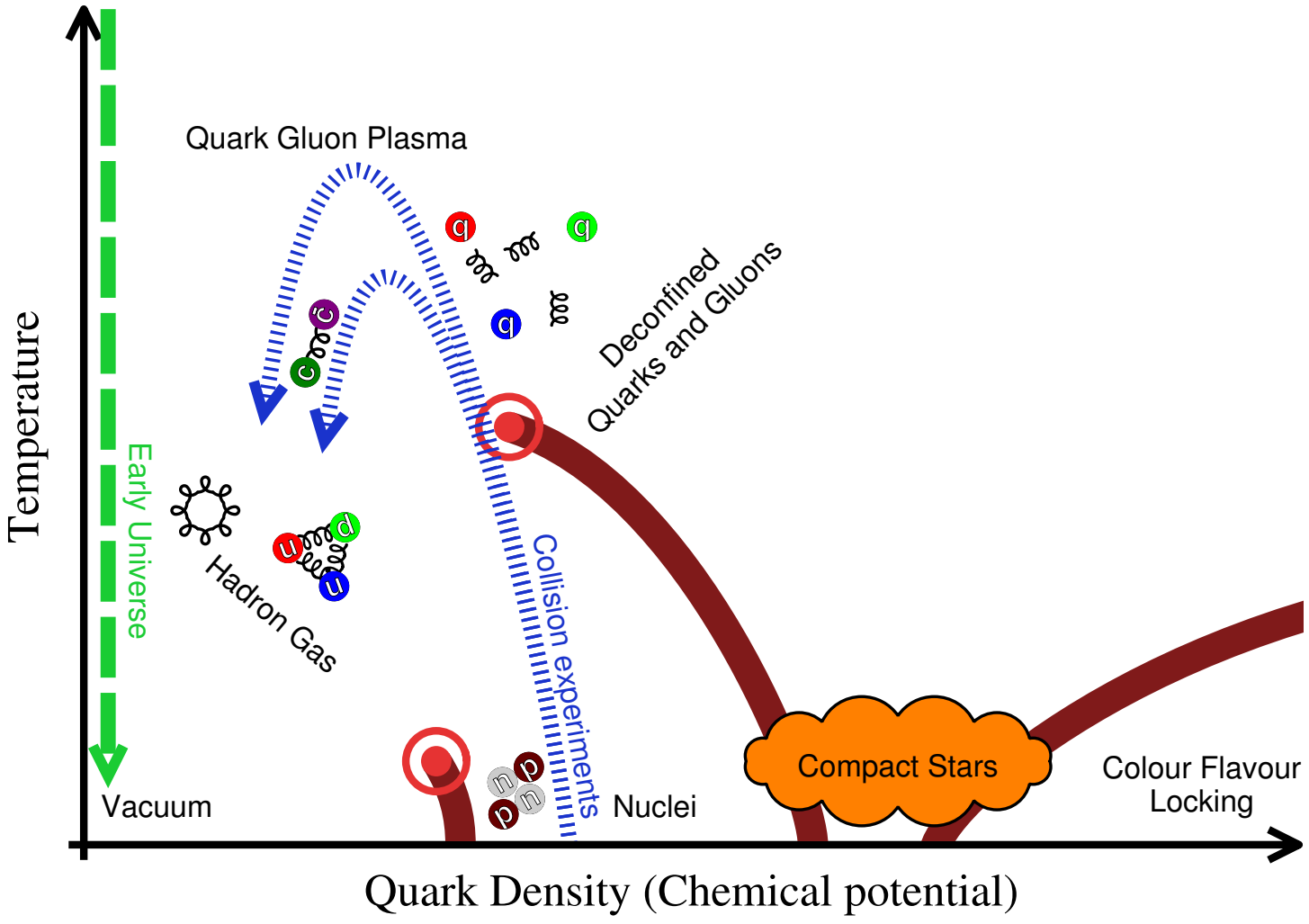
Attila Pásztor, Réka Víg

University of Wuppertal

Mainz, March 8, 2023
Novel Lattice Fermions and their Suitability for High-Performance Computing and Perturbation Theory



The QCD phase diagram

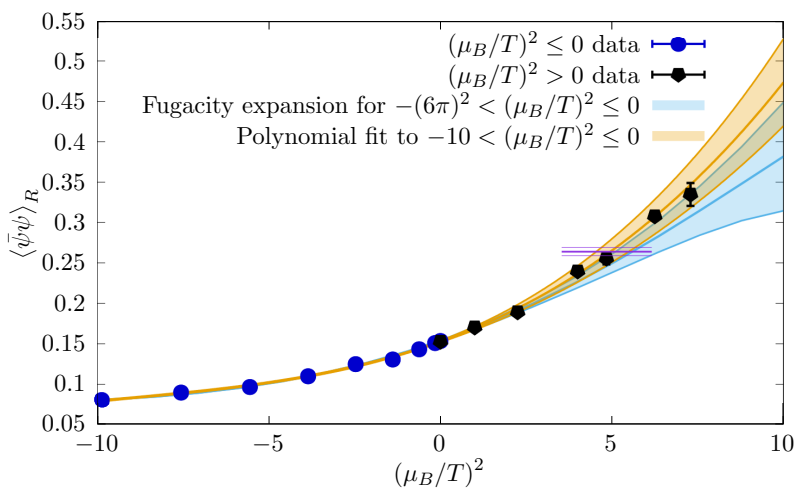


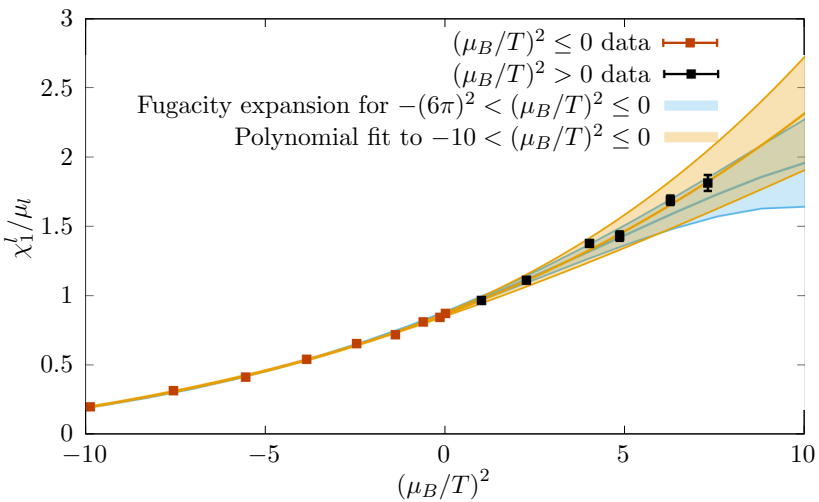
Direct (reweighted) simulations at $\mu_B > 0$

How far can we go in the chemical potential?

We compare in these plots for 140 MeV

- Taylor expansion from imaginary μ_B
- Fugacity expansion from imaginary μ_B
- Direct finite density simulations at $0 < \mu_B \leq 380$ MeV
  Direct = reweigting from sign quenched





The direct result has the smallest errors.

[Wuppertal-Budapest 2108.09213]

Which action should we use?

What do we need?

- Two flavors with a chemical potential: up+down
- What we do with strange? *the strange barrier comes much later*
- Broken symmetries should not distort chiral transition
- Low costs to fight off the sign+overlap problems
- No time-like multi-hops, reduced matrix with few space-like points

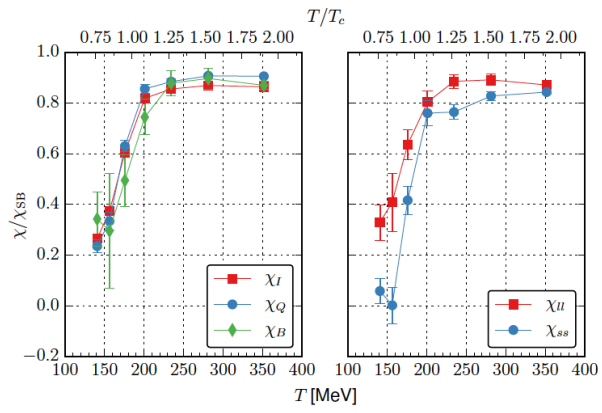
What choices do we have?

- **Rooted staggered**: square root of a complex determinant is ambiguous. Is it analytic in μ_B?
- **Wilson fermions**: conceptually great, but the chiral transition requires much finer lattices than on staggered.
- **Overlap construction**: do we really want to drive μ through a step function?

Non-staggered finite temperature results

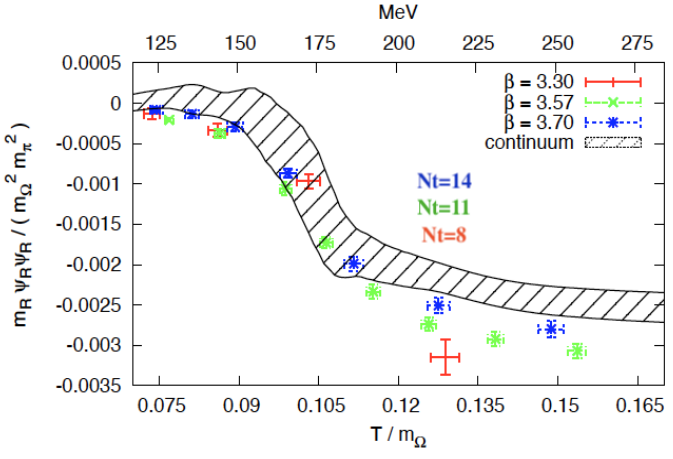
Quark number susceptibilities

anisotropic wilson $m_\pi = 392$ MeV



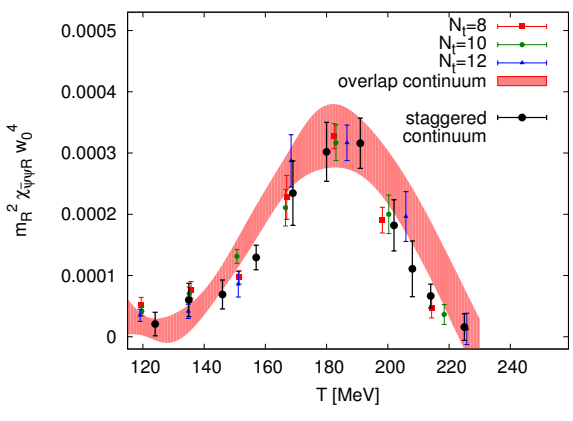
[FASTSUM 1309.6253,1412.6411]

isotropic wilson in continuum 285 MeV.



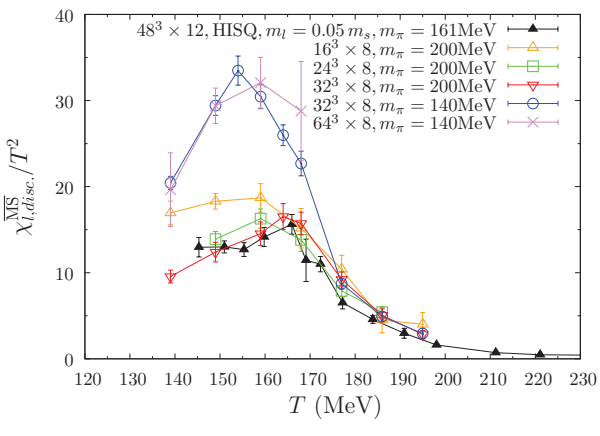
[WB 1205.0440,1504.03676]

Chiral susceptibility



Overlap fermions ($m_\pi = 350$ MeV)

[WB 1204.4089].



Domain wall ($m_\pi = 135$ MeV)

[HotQCD 1402.5175]

Why the Karsten-Wilczek action?

What do we need?

- Two flavors with a chemical potential: up+down
- What we do with strange? *the strange barrier comes much later*
- Broken symmetries should not ditort chiral transition
- Low costs to fight off the sign+overlap problems

Pro Karsten-Wilczek:

- Two flavors
- Anisotropy is not a bug, it is a feature in thermodynamics
- Exact remnant chiral symmetry
- No nested inversion

Contra Karsten-Wilczek:

- Non-trivial tuning
- Are there $\mathcal{O}(a)$ corrections?

Boriçi-Creutz does not lend itself so naturally for finite T

Study of the dispersion relation

$$D_{\mathrm{KW}}(k) = \frac{i}{a}\left[\sum_{\mu=0}^{3}\gamma_\mu \xi \sin\, ak_\mu + \zeta\gamma_0 \sum_{j=1}^{3}(1-\cos\, ak_j)\right]$$

What is the discretization error on the pressure?

$$p = \frac{T}{V}\sum_k \log(D(k)) = \frac{\xi}{N_x^3 N_t}\sum_k \log(D(k))$$

This is divergent, but subtracting the vacuum you get

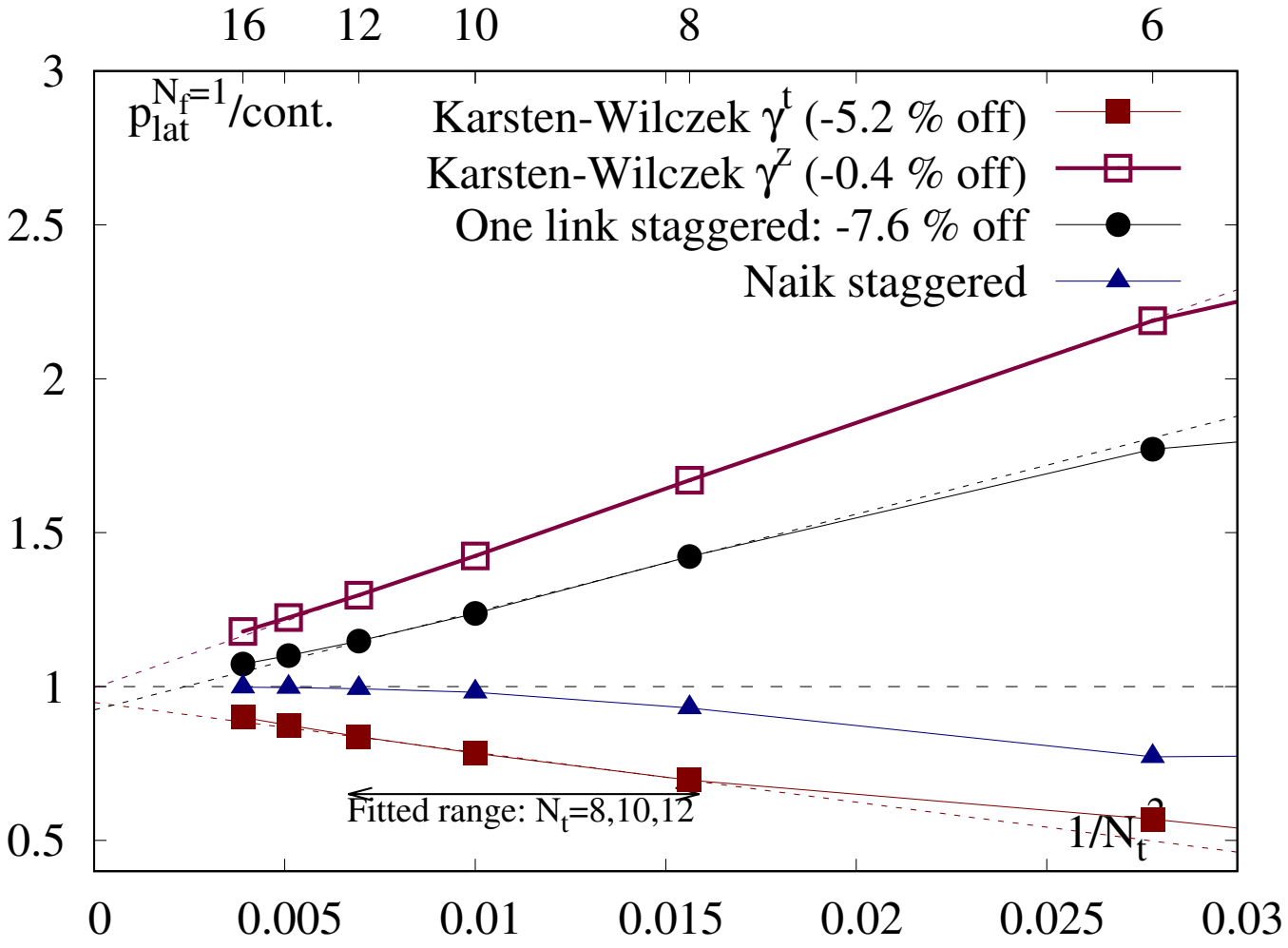
$$\left.\frac{p}{T^4}\right|_{\mathrm{lat}} = N_t^4 \frac{(p(T)-p(T=0))}{\xi^4}$$

In the continuum you get for one flavor (K-W gives this with factor 2):

$$\left.\frac{p}{T^4}\right|_{\mathrm{cont}} = \frac{7\pi^2}{180}$$

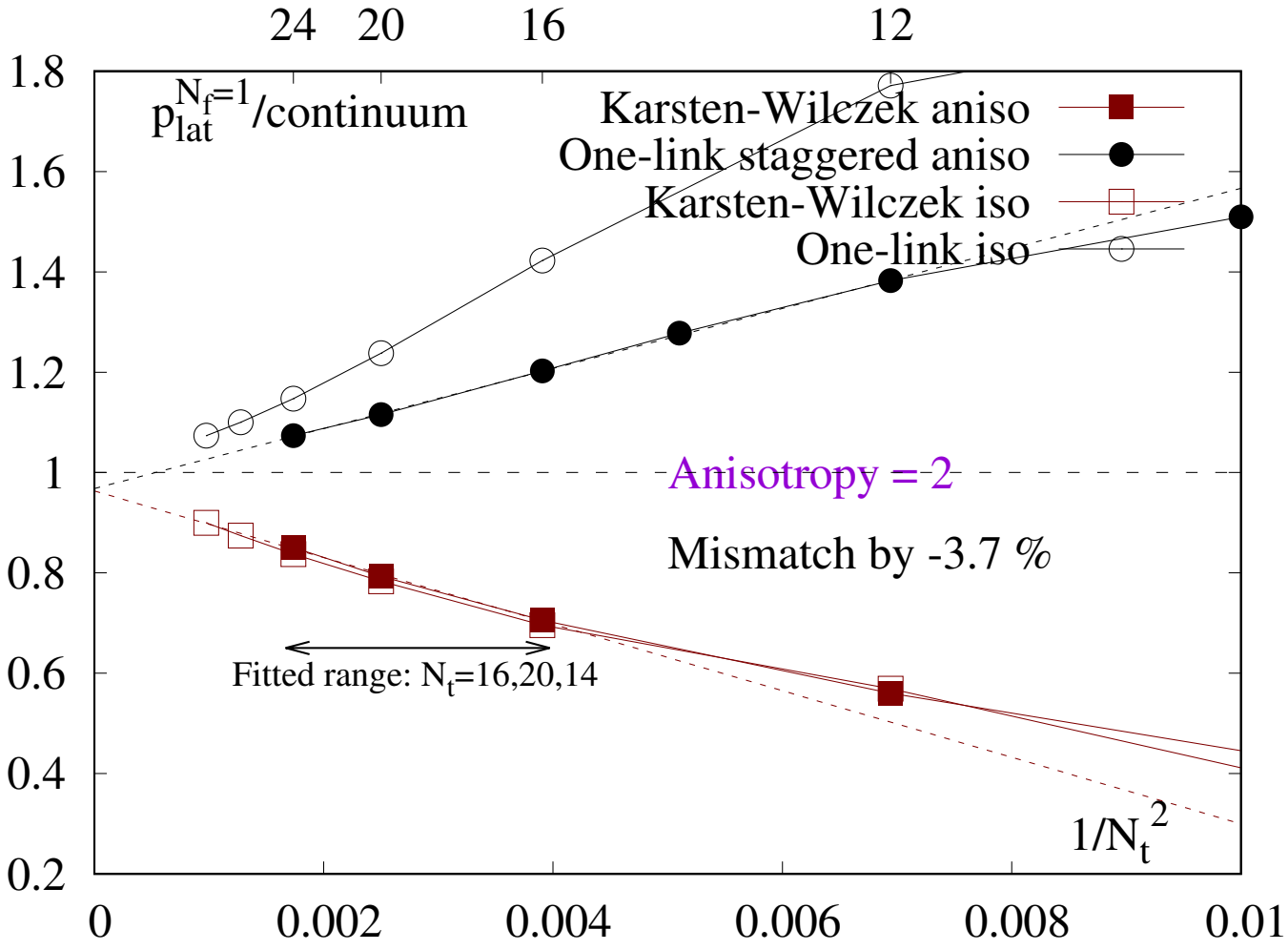
How do your lattices need to be so that a continuum extrapolation will give you back this continuum number? [Phd Thesis A Peikert, Bielefeld 2000]

Pressure discretized / contniuum



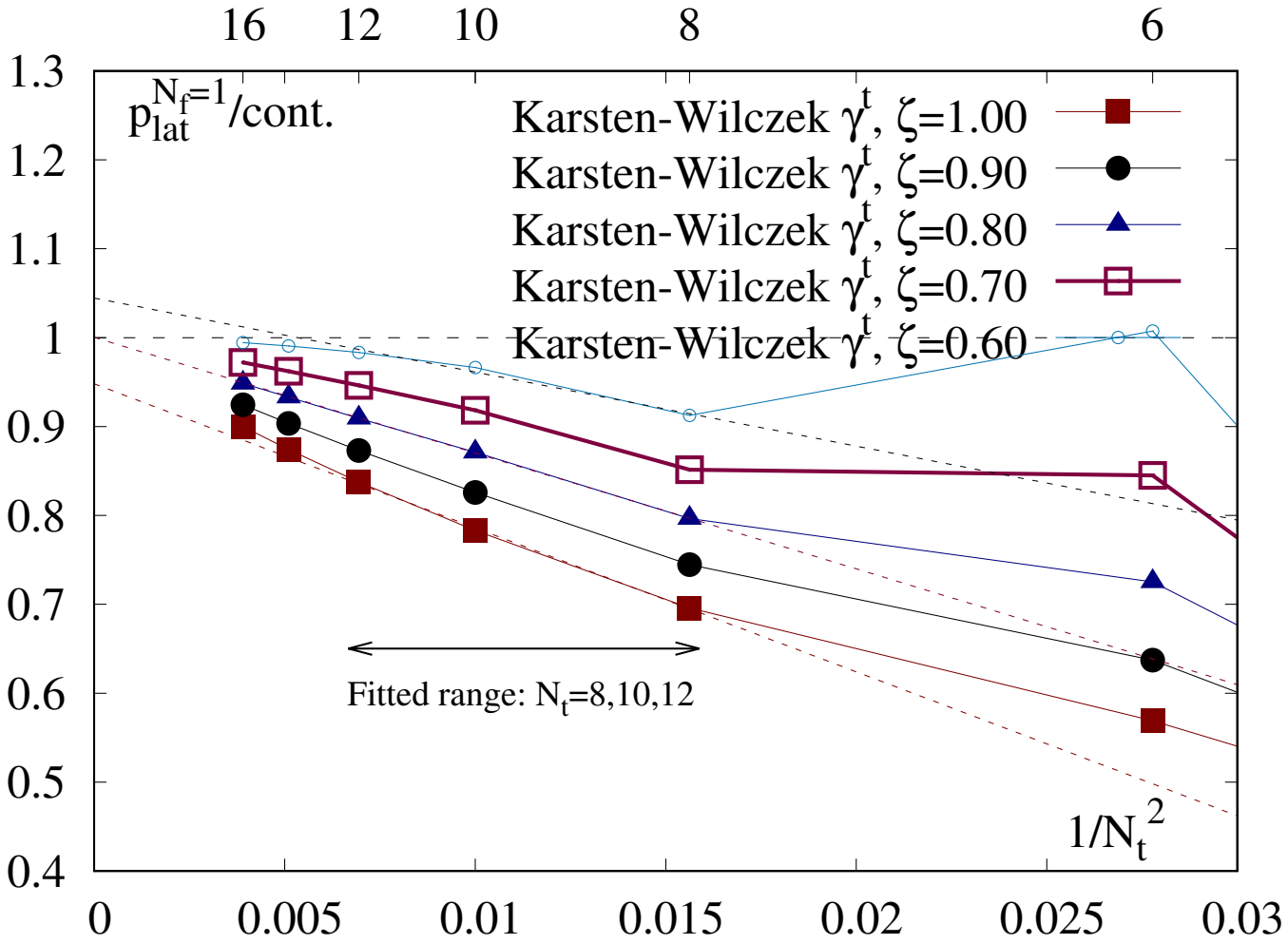
Pressure discretized / contniuum

Let's introduce an anisotropy of $\xi = 2$.



Pressure discretized / contniuum

Let's play with the Wilczek parameter ζ



Tuning the counterterms

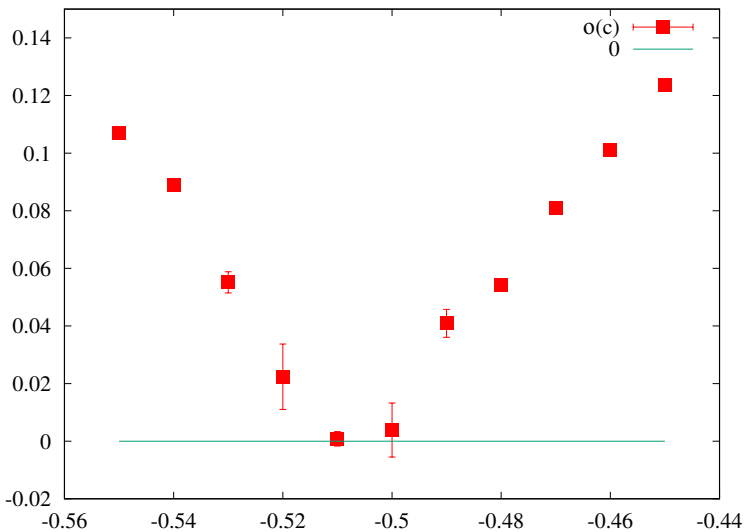
Dimension 4 and gauge counterterms are just the standard ξ_g and ξ_f for anisotropic actions.

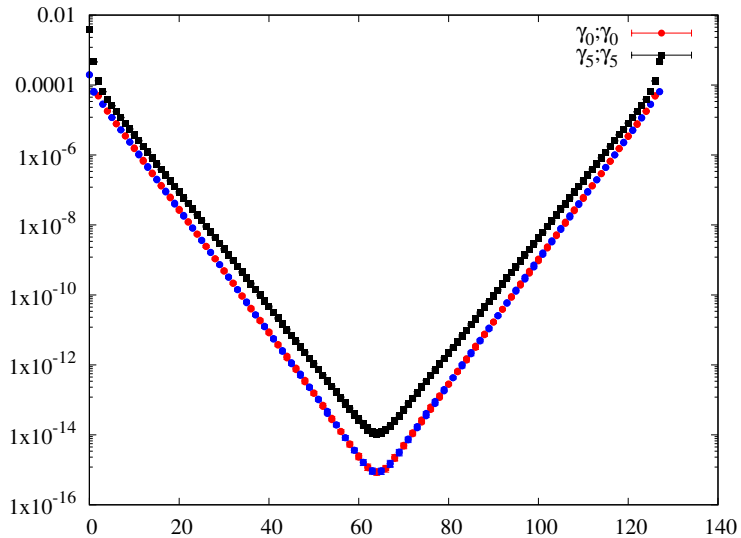
The only new term c: dimension 3.

J. Weber: Look at the oscillations in the parallel correlator of $\bar{\psi}\gamma^0\psi$.

$$A\left[\cos(t\omega+\phi)\exp(-mt)+\cos((N_t-t)\omega+\phi)\exp(-m(N_t-t))\right]$$

The correct c es defined through $\omega-\pi=0$.





The two masses differ by 8%, with 600 MeV mass

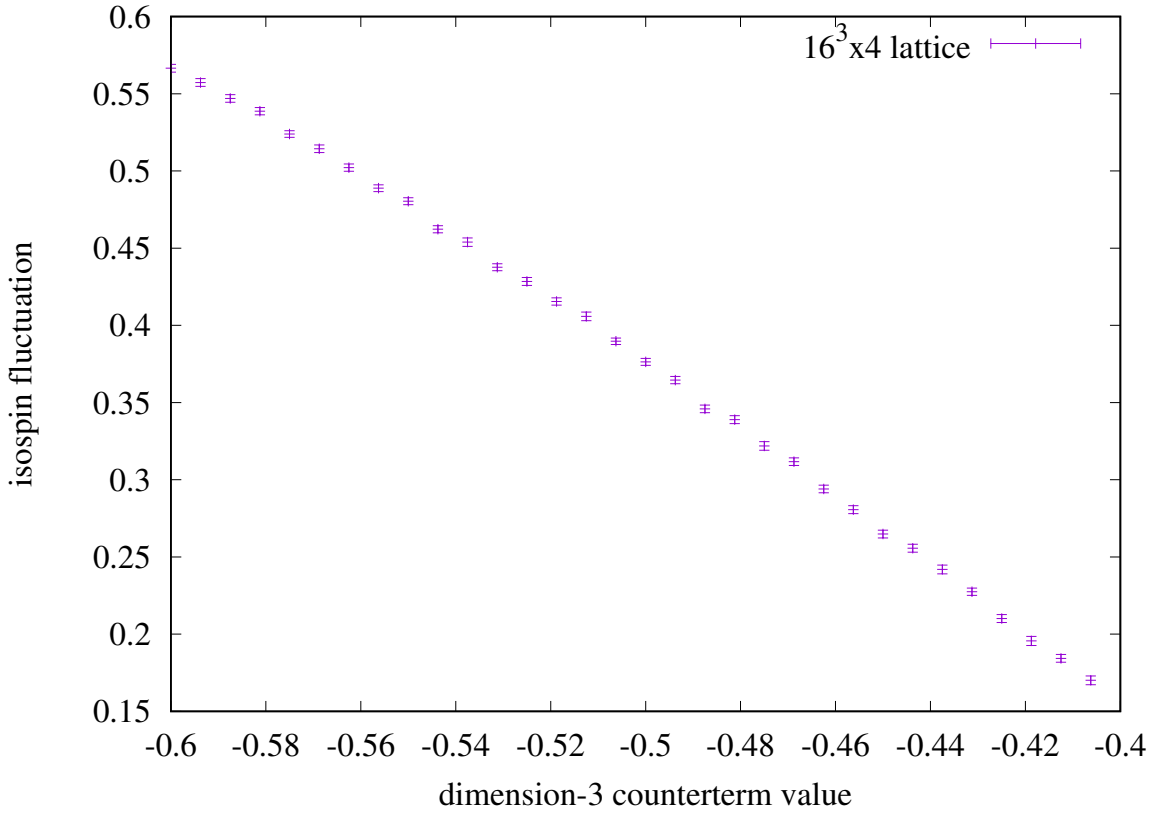
$\rightarrow \sqrt{M_{00}^2-M_{55}^2}=250MeV$, $a=0.17$ fm.

How accurately do we need this counterterm?

Take a relevant observable, here

bulk isospin fluctuations

and observe it as a function of the counterterm.

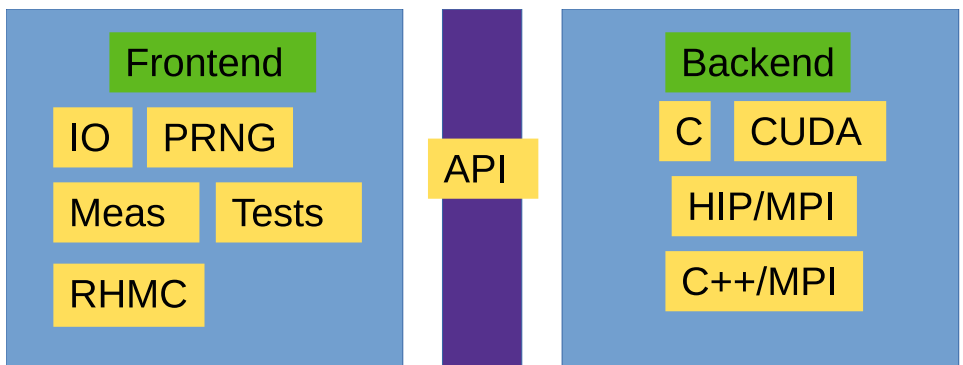


The simulation code – Outline

- Overview of the simulation code
- Implementation of K-W action
- Comparison of Scaling among Staggered, Wilson and K-W actions
- Conclusion

Overview of the simulation code

- The simulation code is written in C, C++, HIP and CUDA
- Composed of different backends (corresponding to different architectures or implementation details) , sharing modules in the frontend
- The frontend connects to the backend via an API composed of "kernel functions" (not GPU kernels) , which are the critical routines provided by the backends
- All local operations on each site are written as backend-independent macros that are shared in the frontend
- NVSHMEM package is used to avoid most of host-device copying
- Under active development, currently supports Staggered, Wilson and Karsten-Wilczek (K-W) actions, in C and CUDA backends. C++/MPI and HIP/MPI backends are planned.



Implementation of K-W action

- $D \equiv 2(m + \not{D})$:

$$
\begin{aligned}
D\psi[s] \equiv \sum_{\mu} \xi_{\mu} \big[& c_{\mu}(s)\ \Gamma^{\mu}\ U_{\mu}(s)\ \psi[s+\mu] \\
& - c_{\mu}^{-1}(s-\mu)\ \Gamma^{\mu\dagger} U_{\mu}^{\dagger}(s-\mu)\psi[s-\mu]\big] \\
& + (2m + 2i(3\zeta + c)\gamma^{\alpha})\psi[s]
\end{aligned}
$$

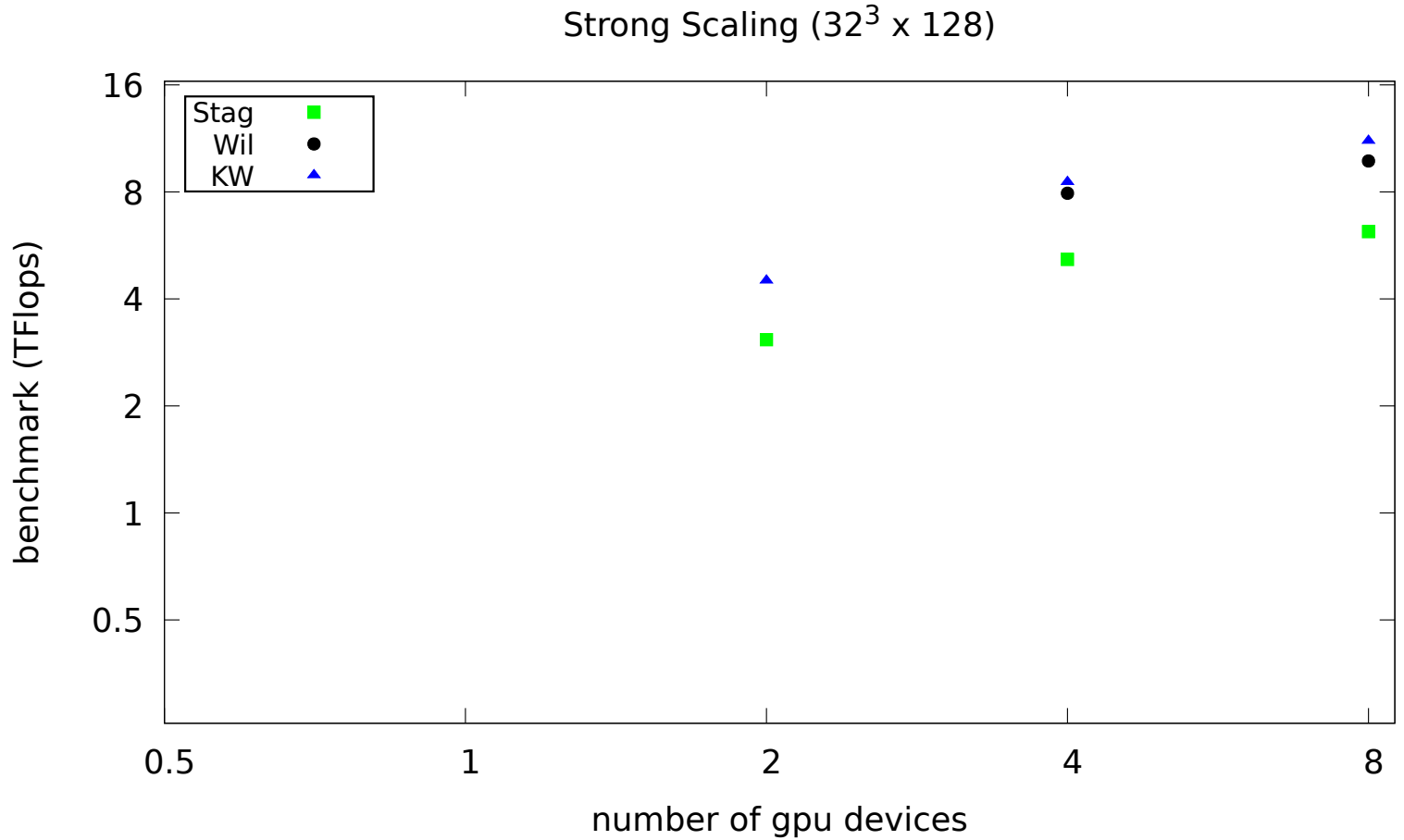
$$
\Gamma^{\mu} \equiv \begin{cases} \gamma^{\mu} - i\zeta\gamma^{\alpha} & , \mu \neq \alpha \\ (1+d)\gamma^{\mu} & , \mu = \alpha \end{cases}
$$

  $c_{\mu}(s) = e^{\frac{\mu_q}{N_t}\ \delta_{\mu,t}} \Phi(s)$,
  $\Phi(s) = \pm 1$ according to b.c. ,
  chemical potential μ_q being a complex number in general,
  $\xi_{\mu} = \xi^{\delta_{\mu t}}$ is the anisotropy, always understood in the Euclidean time direction
  d is absorbed by ξ_{μ} if α is t

- α is general but a compile-time decision
- Critical routines are combined into larger CUDA kernels for better performance

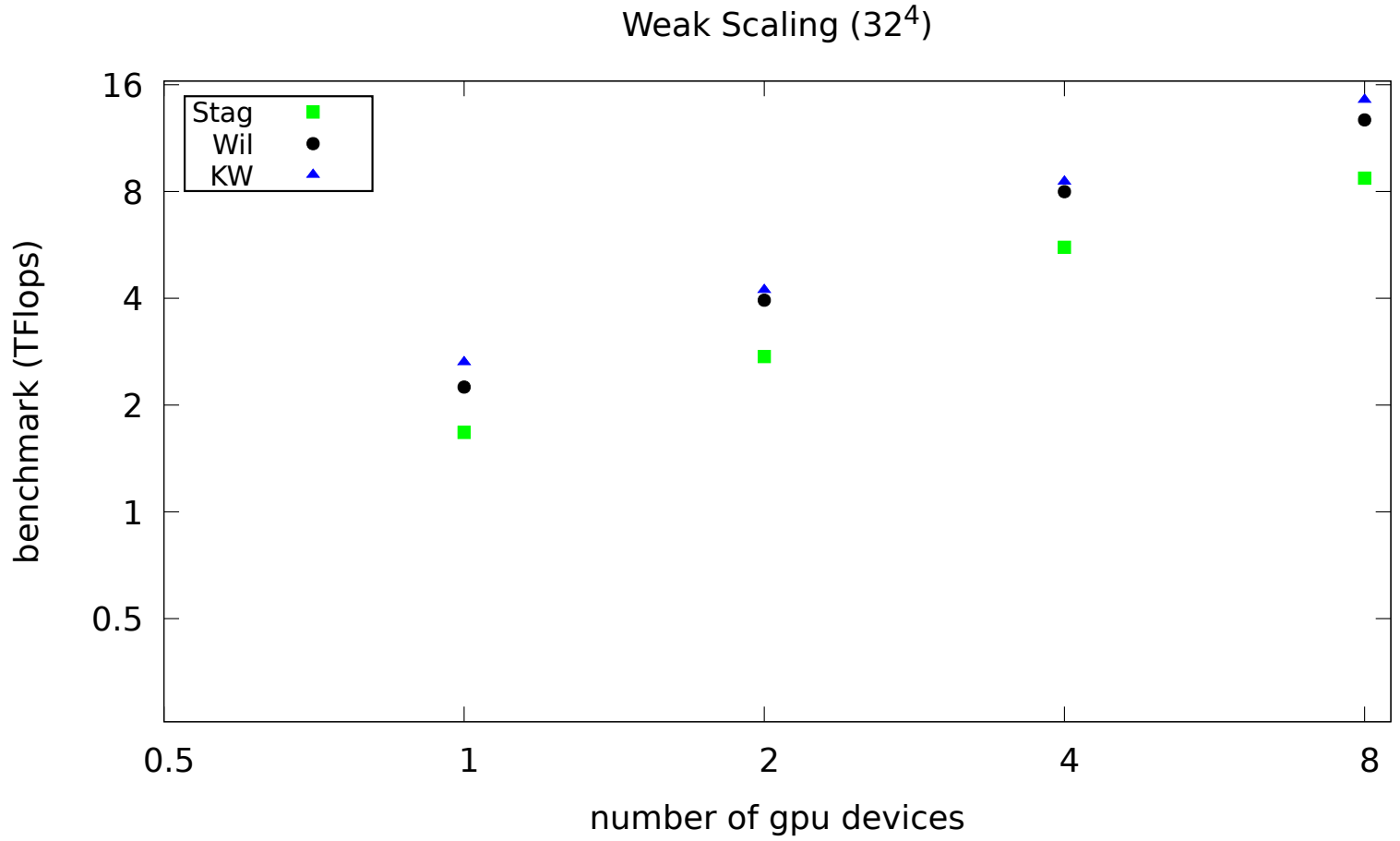
Comparison of Scaling among Staggered, Wilson and K-W actions

Test on A100 GPUs in JUWELS BOOSTER of Juliech Supercomputing Center (JSC)



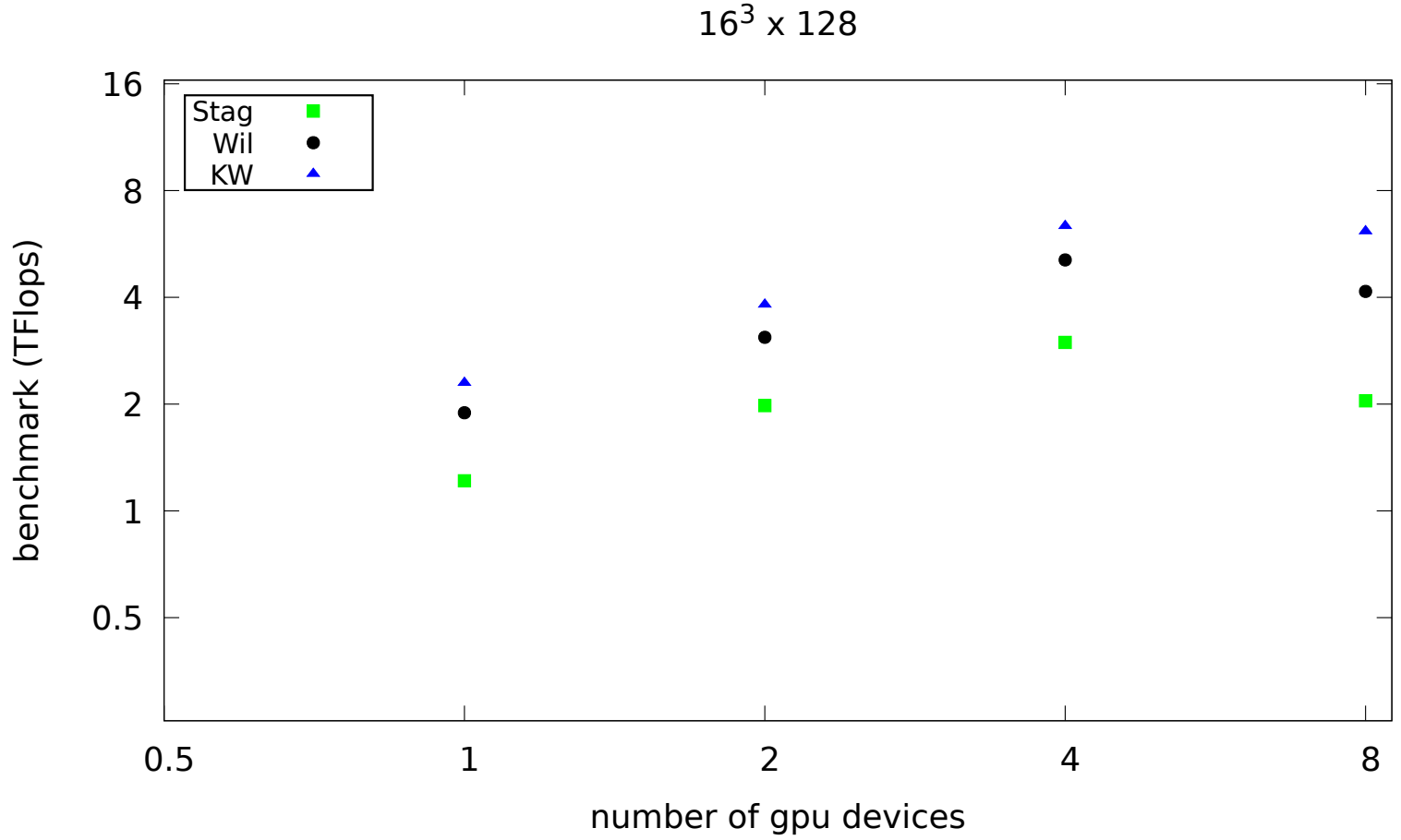
Comparison of Scaling among Staggered, Wilson and K-W actions

Test on A100 GPUs in JUWELS BOOSTER of Juliech Supercomputing Center (JSC)



Comparison of Scaling among Staggered, Wilson and K-W actions

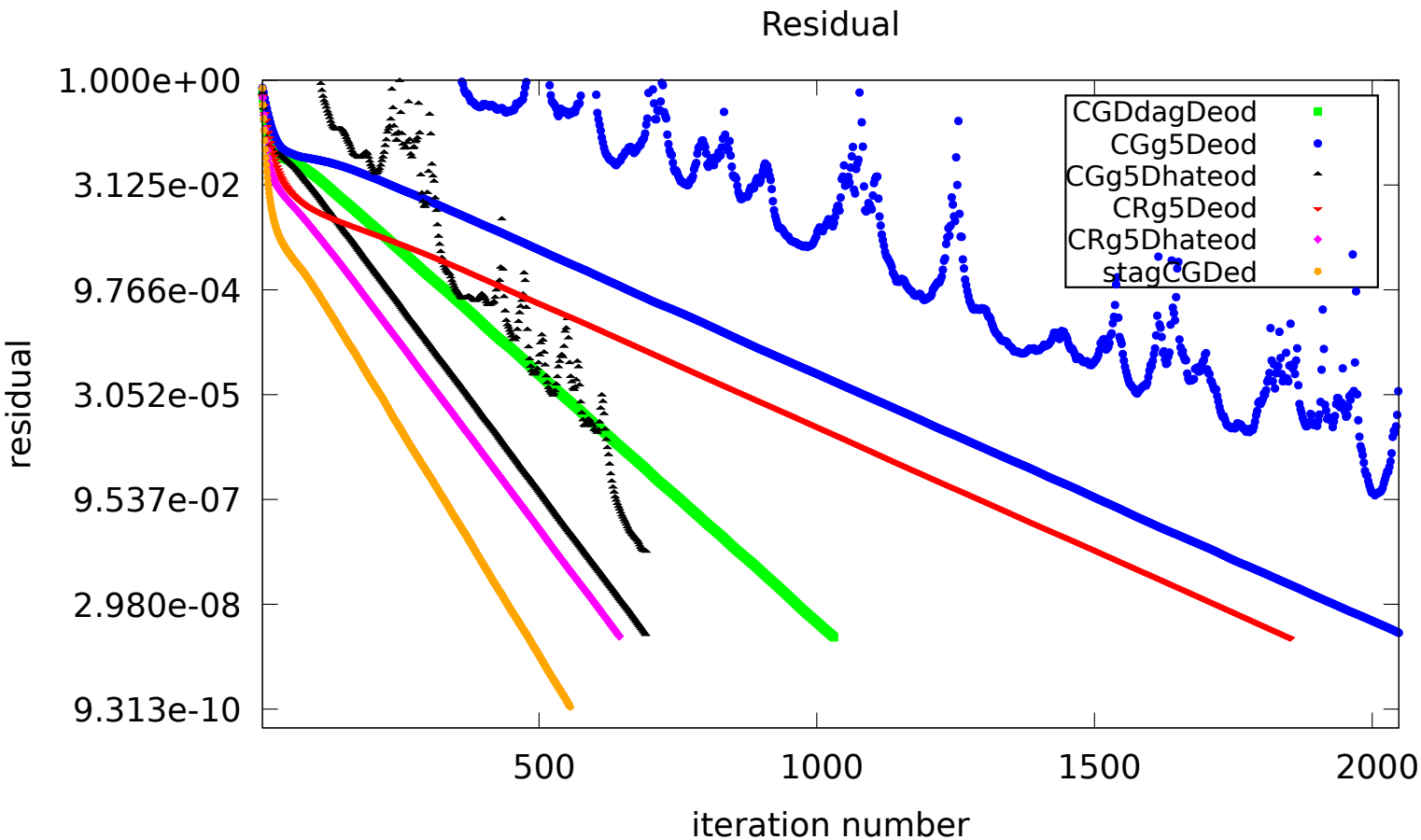
Test on A100 GPUs in JUWELS BOOSTER of Juliech Supercomputing Center (JSC)



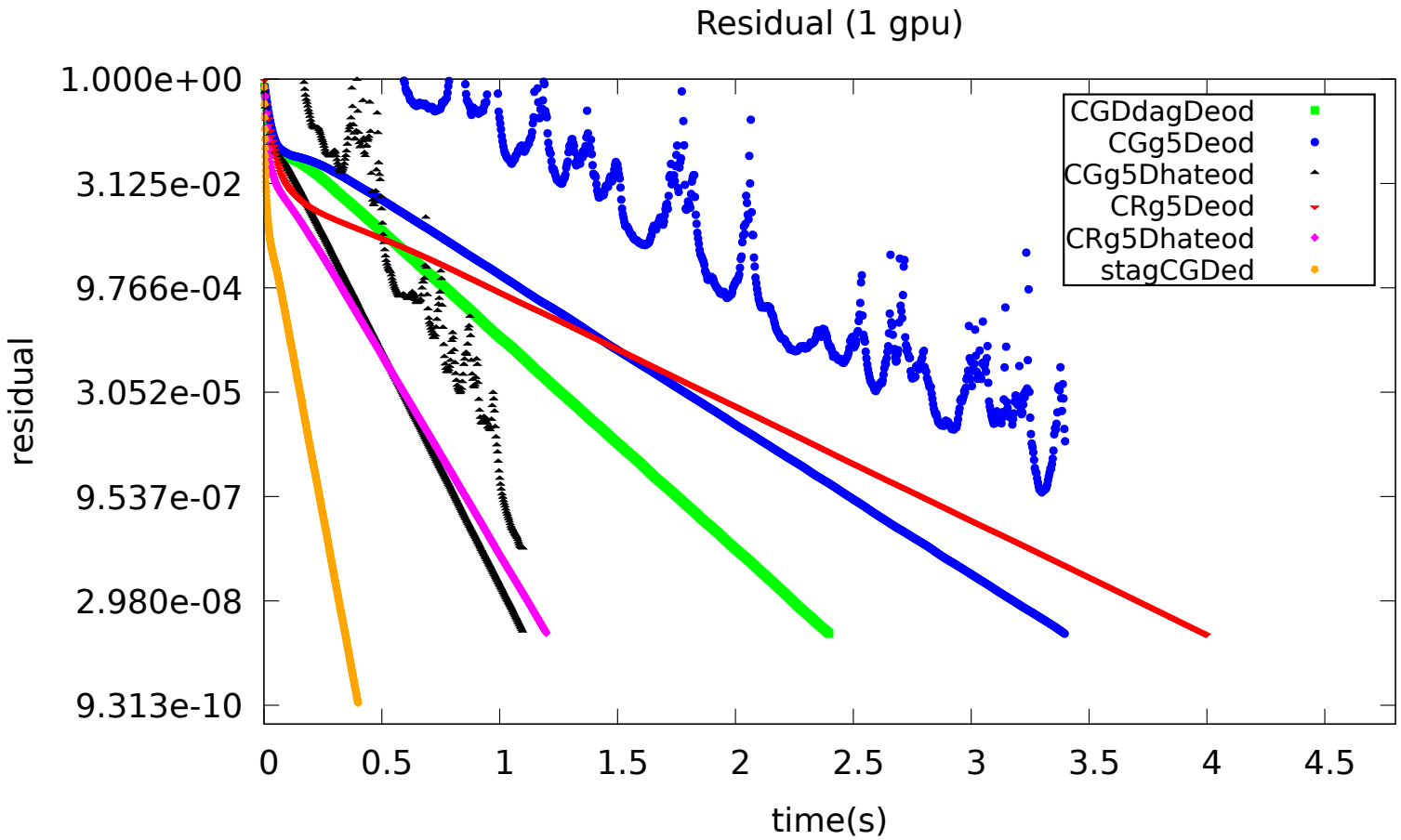
Comparison of Scaling among Staggered, Wilson and K-W actions

- Test on A100 GPUs with a quenched $16^3 \times 128$ config with approximately tuned parameter values :
  $\beta = 4.2095$, $\xi_g = 1.811336$, $c = -0.49$, $m = 0.05$, 2-stout with $\rho = 0.15$ (ξ_f or d left untuned)
- So far we have tried:
  - CG on $D^\dagger D$(CGDdagD), $\gamma^5 D$(CGg5D), $\gamma^5 D$ with e/o preconditioning (CGg5Dhat),
  - CR on $\gamma^5 D$(CRg5D), CR on $\gamma^5 D$ with e/o preconditioning(CRg5Dhat)
- Comparison: CG in staggered(stagCG) on the same config :
  $\beta = 4.2095$, $\xi_g = 1.811336$, $\xi_f = 1.84$, $m = 0.0463$, 2-stout with $\rho = 0.15$

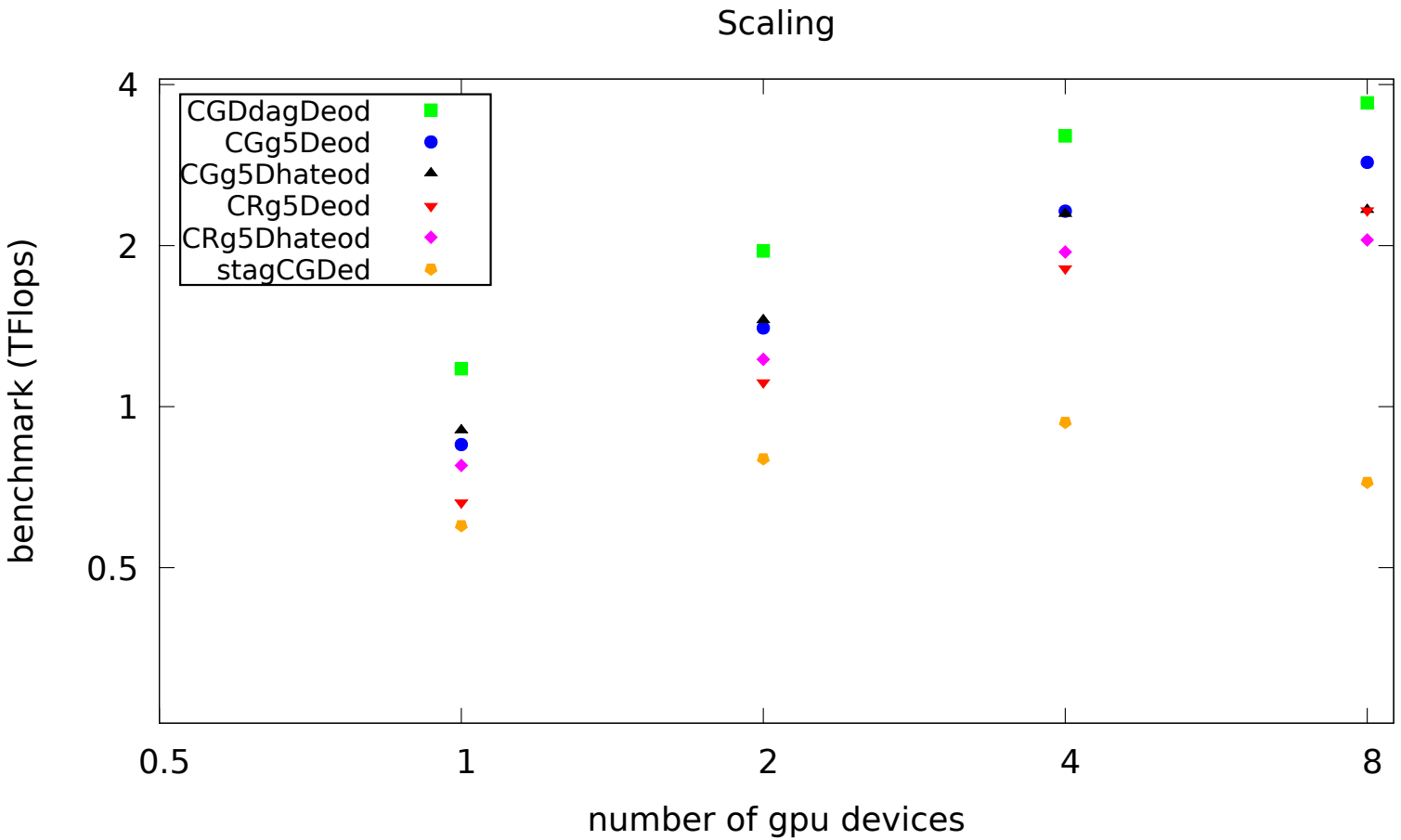
Comparison of Scaling among Staggered, Wilson and K-W actions



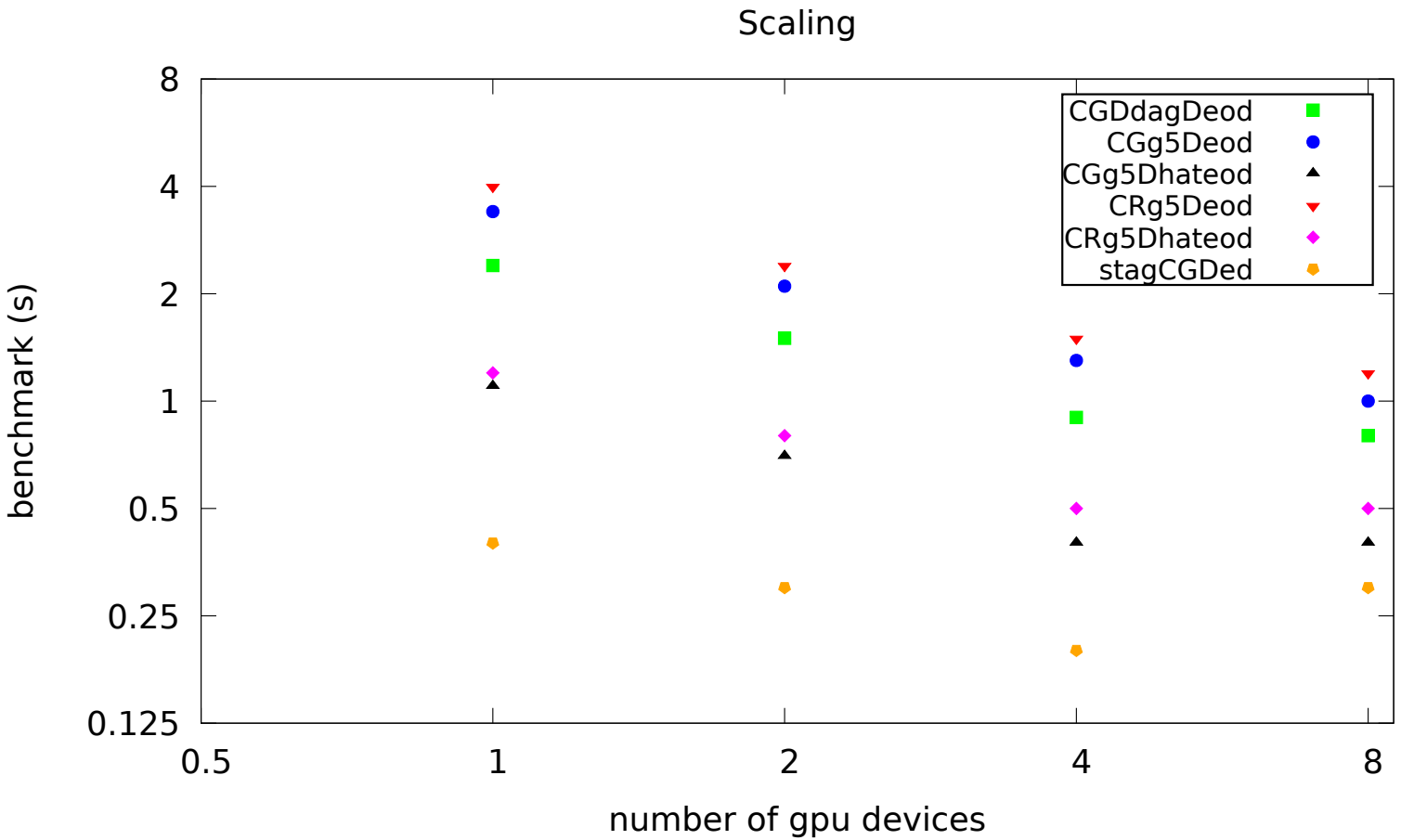
Comparison of Scaling among Staggered, Wilson and K-W actions



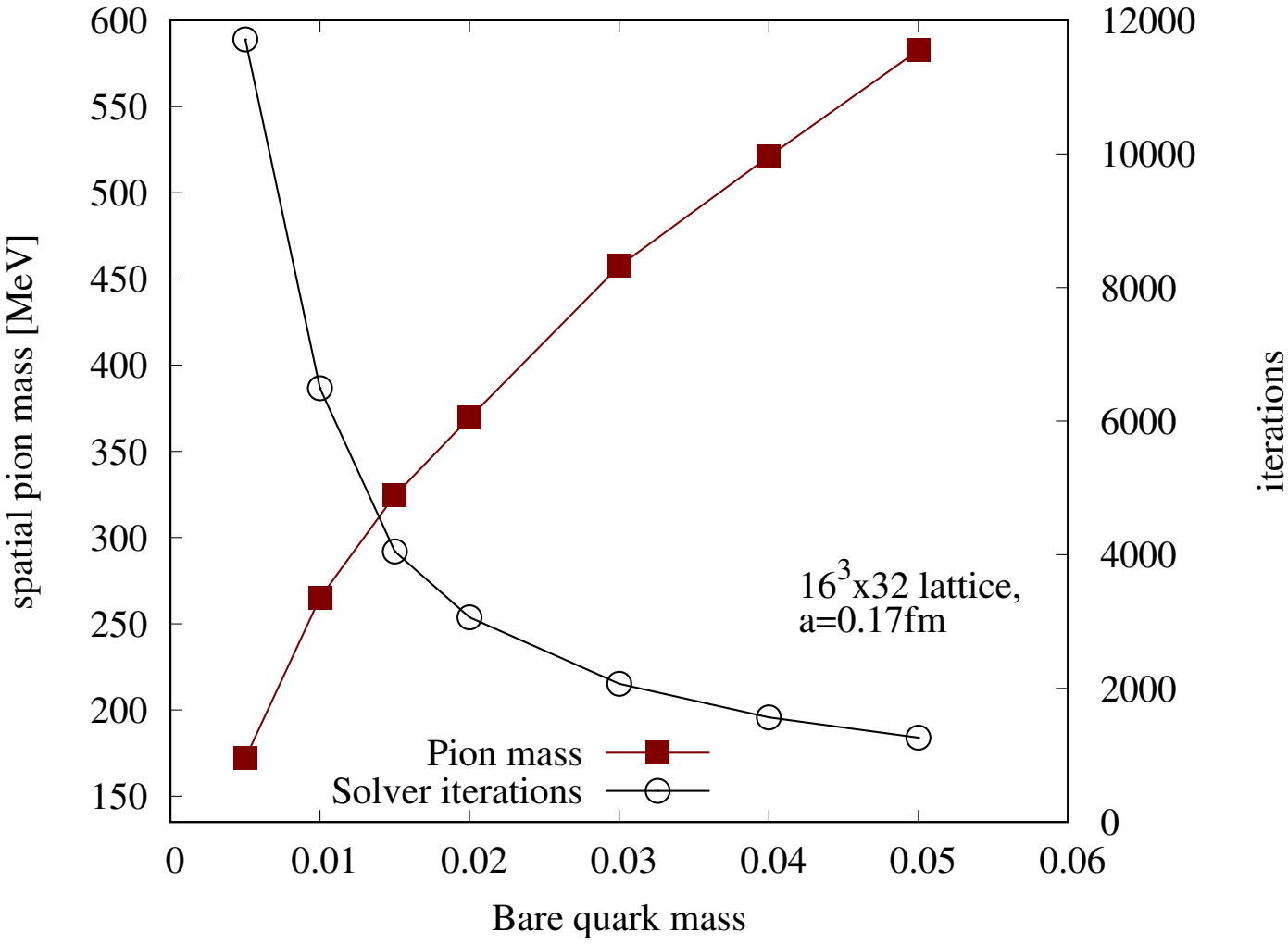
Comparison of Scaling among Staggered, Wilson and K-W actions



Comparison of Scaling among Staggered, Wilson and K-W actions



Towards physical pion mass



Conclusion

- We need to move away from Staggered; the K-W action is attractive
- We plan to conduct dynamical simulations with the K-W action to study 2-flavor QCD at finite real μ
- We are currently tuning the renormalization constants
- The K-W action is implemented in C and CUDA and the fermionic force routine will be implemented
- We have tried different solvers at relevant parameter values and configs. So far CR or CG on $\gamma^5 D$ with e/o preconditioning gives best performance. There will be tests on other solvers e.g. BiCGStab, CGNE.
- It is observed that the cost of inversion for K-W action is about ~ 3 times that of staggered action
- Dim-5 improvement terms may also be implemented and used
- Improvement of the dispersion relation is considered