First dynamical symulations with minimally doubled fermions

Réka Á. Vig

in collaboration with
S. Borsányi, Z. Fodor, D. Godzieba, P. Parotto, C. H. Wong

Lattice conference 2023
Aug 1, 2023

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Motivation

For simulating two light flavors dynamically at $T \neq 0,\ \mu \neq 0$ one can choose from several discretizations:

- Wilson fermions
  Explicit breaking of chiral simmetry $\rightarrow$ at $T = 0$ OK.
- Overlap fermions
  Computationally costly.
- Staggered fermions
  Rooting is ambiguous for $\mu \neq 0$. (For more details see C. H. Wong's talk)
- Minimally doubled fermions
  A remnant chiral simmetry is present, rooting is not needed.
  Do we have O(a) discretization errors?

Karsten-Wilczek fermions

Our choice for a minimally doubled action is the so-called Karsten-Wilczek (KW) action [L.H. Karsten 1981, F. Wilczek 1987].

$$S_F^{KW} = S_F^N + \sum_x \sum_{j=1}^{3} \bar{\psi}(x) \frac{i\zeta}{2} \gamma^\alpha \left(2\psi(x) - U_j(x)\psi(x+\hat{j}) - U_j^\dagger(x)\psi(x-\hat{j}) \right),$$

Where S_F^N is the naive fermion action, the $U_\mu(x)$ are the links in diretion μ at lattice site x. ζ is the Wilczek parameter, with $|\zeta| > 1/2$. S_F^{KW} breaks hypercubic symmetry, thus we need three counter-terms for renormalization (coefficients were determined perturbatively [S. Capitani et.al. (2009)])

$$S^{3f} = c \sum_x \bar{\psi}(x) i\gamma^\alpha \psi(x),$$

$$S^{4f} = d \sum_x \bar{\psi}(x) \frac{1}{2} \gamma^\alpha \left(U_\alpha(x)\psi(x+\hat{\alpha}) - U_\alpha^\dagger(x-\hat{\alpha})\psi(x-\hat{\alpha}) \right),$$

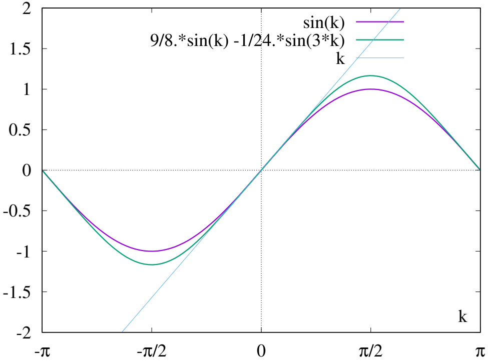
$$S^{4g} = d_G \sum_x \sum_{\mu \neq \alpha} \mathrm{Re}\,\mathrm{Tr}\left(1 - \mathcal{P}_{\mu\alpha}(x)\right),$$

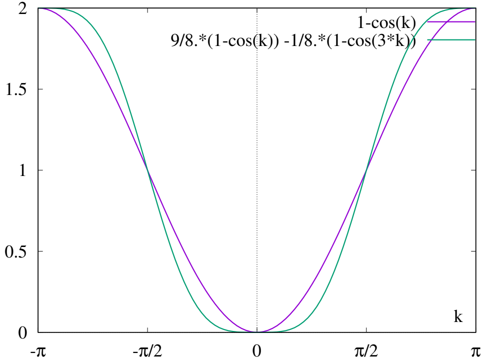
In our case $\alpha \equiv t$.

Karsten-Wilczek Dirac operator in momentum space

In momentum space the Dirac operator reads

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





Karsten-Wilczek Dirac operator in position space

The KW Dirac operator in position space acts on a fermion fied the following way

$$D^{KW}\psi[s] \equiv \sum_{\mu}\Big[c_\mu(s)\ \Gamma^\mu\ U_\mu(s)\ \psi[s+\mu] - c_\mu^{-1}(s-\mu)\ \Gamma^{\mu\dagger} U_\mu^\dagger(s-\mu)\psi[s-\mu]\Big]$$
$$+ (2m + 2i(3\zeta + c)\gamma^\alpha)\psi[s]$$

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

Where $(1+d)$ is structurally the same as an anisotropy parameter, thus we will use the notation $\xi \equiv (1+d)$ and also $\xi_\beta \equiv (1+d_G)$.

$c_\mu(s) = e^{\frac{\mu_q}{N_t}\ \delta_{\mu,t}}\Phi(s)$, with boundary condition $\Phi(s) = \pm 1$,

($\mu \in \{0,1,2,3\}$ which are equivalent to directions $\{t,z,y,x\}$.)

Naik improvement

The KW term is first order in the lattice spacing a. Thus for a better approach of the continuum limit we introduced improvement in the Dirac operator.

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

The $sin(ak_\mu)$ term corresponds to the Nabla and the $(1 - cos(ak_j))$ corresponds to the Laplacian.
We can elimininate the $\mathcal{O}(k^2)$ term in this expression by realizing improvement of the Laplacian

$$(1 - \cos\, ak_j) \to c_1(1 - \cos\, ak_j) + c_3(1 - \cos\, 3ak_j)$$

with the special choice $c_1 = 9/8$, $c_3 = -1/8$.

Naik improvement

In position space we have the following improved expression

$$D^{KW}\psi[s] \equiv \sum_\mu \xi_\mu \Big[c_\mu(s)\ \Gamma^\mu_{(1)}\ U_\mu(s)\ \psi[s+\mu] - c_\mu^{-1}(s-\mu)\ \Gamma^{\mu\dagger}_{(1)} U^\dagger_\mu(s-\mu)\psi[s-\mu]$$
$$+ c_\mu(s)c_\mu(s+\mu)c_\mu(s+2\mu)\ \Gamma^\mu_{(3)}\ U_\mu(s)U_\mu(s+\mu)U_\mu(s+2\mu)\ \psi[s+3\mu]$$
$$-c_\mu^{-1}(s-\mu)c_\mu^{-1}(s-2\mu)c_\mu^{-1}(s-3\mu)\ \Gamma^{\mu\dagger}_{(3)} U^\dagger_\mu(s\mu)U^\dagger_\mu(s-2\mu)U^\dagger_\mu(s-3\mu)\psi[s-3\mu]\Big]$$
$$+ (2m + 2i(3\zeta + c)\gamma^0)\ \psi[s]$$

Where

$$\Gamma^0_{(1)} = \gamma^0, \qquad \Gamma^\mu_{(1)} = \bar{s}_1\gamma^\mu - i\bar{c}_1\gamma^0 \quad (\mu \neq 0)$$
$$\Gamma^0_{(3)} = 0, \qquad \Gamma^\mu_{(3)} = \bar{s}_3\gamma^\mu - i\bar{c}_3\gamma^0 \quad (\mu \neq 0)$$
$$1 = \bar{s}_1 + 3\bar{s}_3, \quad \zeta = \bar{c}_1 + \bar{c}_3$$

The relation to the momentum space c_1 and c_3 is $c_1 = \zeta\bar{c}_1,\ c_3 = \zeta\bar{c}_3$.

Half vector trick

- We want to use only 2 Dirac components instead of 4 $\rightarrow$ half as many operations.
- If $\bar{s}_1 = \bar{c}_1$ and $\bar{s}_3 = \bar{c}_3$, then $\Gamma^{\mu}_{(1),(3)} = \gamma_j + i\gamma_0$ are projectors. This allows the half vector trick.
- By the following choice of parameters criteria for the half vector trick and Laplacian improvement are satisfied

$$\bar{s}_1 = \bar{c}_1 = 1.5$$
$$\bar{s}_3 = \bar{c}_3 = -1/6$$
$$\zeta = 4/3$$

- Because of the Laplacian improvement the action is $\mathcal{O}(a)$ improved even in this case.

Properties of the Dirac operator

The KW Dirac operator $D^{KW} = m + \not{D}$ has the following properties

$$\not{D}^\dagger = \gamma_5 \not{D} \gamma_5 \qquad \gamma_5 - \text{hermiticity}$$

$$D^{KW\dagger}\gamma_5 = \gamma_5 D^{KW}$$

$$\gamma_5 \not{D} \gamma_5 = -\not{D} \qquad \text{chiral symmetry}$$

$$D^{KW\dagger} = m - \not{D}$$

$$D^{KW\dagger} D^{KW} = (m - \not{D})(m + \not{D}) = m^2 - \not{D}^2 = D^{KW} D^{KW\dagger} \qquad \mathrm{D}^{\mathrm{KW}} \text{ is normal}$$

$$D^{KW\dagger} D^{KW} \gamma_5 = D^{KW\dagger} \gamma_5 D^{KW\dagger} = \gamma_5 D^{KW} D^{KW\dagger} = \gamma_5 D^{KW\dagger} D^{KW}$$

Consequences:
The eigenvalues of $\not{D}$ consist of complex conjugate pairs $\pm i\lambda$.
$D^{KW\dagger} D^{KW}$ is positive definite with eigenvalues $m^2 + \lambda^2$ and its eigenvectors are left(L)- or right(R)-handed.

Simulating with HMC

Properties of D^{KW} makes it possible to use Hybrid Monte Carlo simulations without the rational approximation.

$$M = \begin{pmatrix} M_{LL} & M_{LR} \\ M_{RL} & M_{RR} \end{pmatrix} = \left(\begin{array}{c|c} m\mathbb{1} & \not{D}_{LR} \\ \hline \not{D}_{RL} & m\mathbb{1} \end{array}\right)$$

$$M^+M = \begin{pmatrix} m^2 - \not{D}_{LR}\not{D}_{RL} & 0 \\ 0 & m^2 - \not{D}_{RL}\not{D}_{LR} \end{pmatrix}$$

Then the action takes the form $S_f^{KW} = \chi_L^+[(M^+M)_{LL}]^{-N_f/2}\chi_L$ $(N_f = 2)$ with $\chi_L = P_L D^{KW}\eta$, where $P_L = \frac{1}{2}(1+\gamma_5)$ and η is a random Gaussian field.

Dynamical simulations with KW action

We produced gauge configurations dynamically for the first time with Karsten-Wilczek action. We tuned the three coefficients of the counter-terms c, ξ, ξ_β were tuned non-perturbatively.
The setup of the simulations was:

HMC and Rational HMC algorithm was used with Hasenbusch preconditioning.
For the preconditioned field we used a Multishift Conjugate Gradient (MCG) solver, for the preconditioner CG solver was used on $\mathrm{D}^{\mathrm{KW}\dagger}\mathrm{D}^{\mathrm{KW}}$.
Temporal and spatial size of the lattices : $N_t = 64,\ N_s = 16$
Inverse gauge coupling : $\beta = 3.5$ corresponding to $a_s = 0.180(3)\ \mathtt{fm}$
For each different setup of c, ξ and ξ_β the number of configurations : $\mathrm{n_{cf}} \approx 200$

Tuning the parameter c

For a non-perturbatve tuning of the c parameter we measured the mesonic correlation function in the γ_0 channel

$$C_0(x,y) \sim \langle \bar{\psi}(x)\gamma_0\psi(x)\bar{\psi}(y)\gamma_0\psi(y)\rangle$$

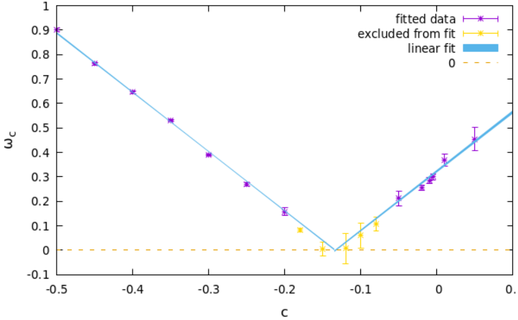
For the correlator parallel to the KW direction oscillation happens if c is not tuned properly [J. Weber (2017)]

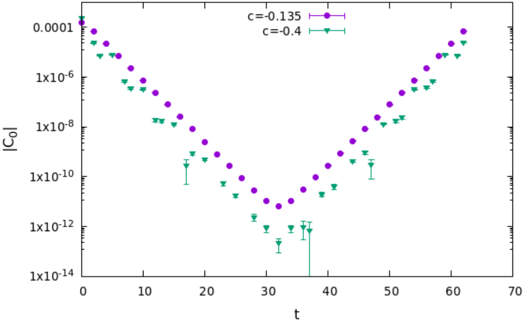
$$C_0(t) \approx A\left[\cos(t\omega+\phi)\exp(-mt)+\cos((N_t-t)\omega+\phi)\exp(-m(N_t-t))\right]$$

We have to set c such that $\omega_c \equiv \omega - \pi = 0$.

We fitted the correlation functions measured on tiled configurations.
(For more details about tiling and the tuning of c see the next talk presented by D. Godzieba.)

We fixed $\xi = \xi_\beta = 1$ and did a scan in the c parameter. By fitting $C_0(t)$ (right plot) at each c we extracted the oscillation frequency ω_c (left plot). We performed a linear fit on the data and interpolated the point where $\omega_c = 0$: $c = -0.1336(13)$.





Tuning the parameters ξ and ξ_β

Fortunately ω_c remained small by keeping c at its previously tuned value even for different settings of ξ and ξ_β. Thus we could perform a 2D scan in the ξ-ξ_β plane by keeping c fixed.

To find the tuned values we checked two criteria for lattice isotropy:

A1 We measured the parallel and perpendicular components of γ_5 channel of meson (pion) correlation function

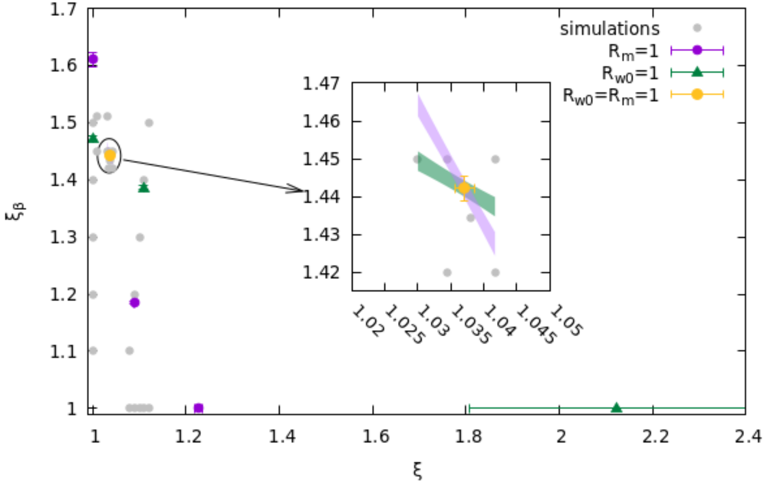
$$C_5(x,y) \sim \langle \bar{\psi}(x)\gamma_5\psi(x)\bar{\psi}(y)\gamma_5\psi(y)\rangle$$

After a fit was performed on $C_{5||}$ and $C_{5\perp}$ we searched for parameter setups where the mass ratio $R_m = \frac{m_{||}}{m_\perp} = 1$.

A2 We measured the lattice spacing in the parallel (a_t) and perpendicular (a_s) directions using the w_0 scale [S.Borsanyi et.al.(2012)]. Then we searched for setups where $R_a = \frac{w_0 a_t}{w_0 a_s} = 1$.

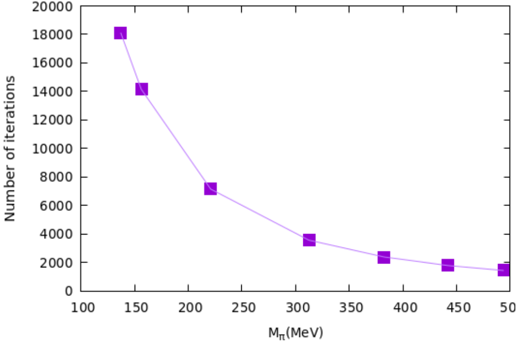
2D scan in ξ and β_ξ

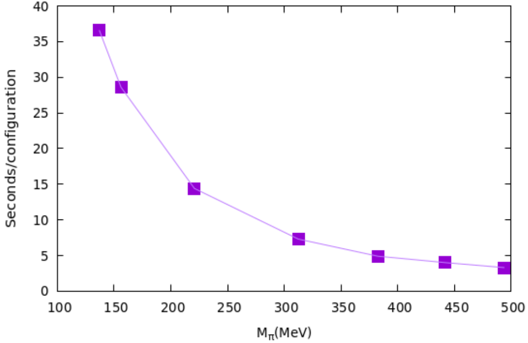
Scanning the 2D plane at different points we could interpolate points where the criteria A_1 and A_2 are satisfied. This defined two lines in the parameter plane. They cross each other at $\xi_\beta = 1.442(3)$, $\xi = 1.0370(15)$.



Cost at the physical point

A mass scan was done starting from pion mass $M_\pi = 486(8)$MeV to $M_\pi = 135(2)$MeV with larger volumes $N_s = 24$, $N_t = 48$ at $\xi_\beta = 1.4345$, $\xi = 1.0381$ for which R_a and R_{w_0} are both 0.1% away from ratio 1.



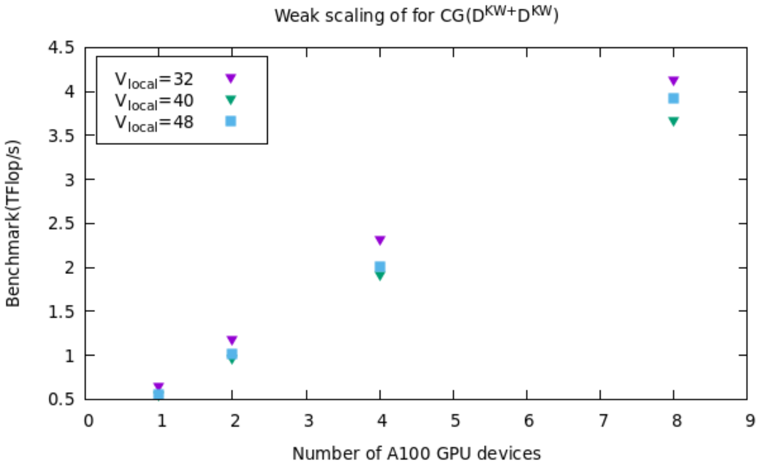


Summary

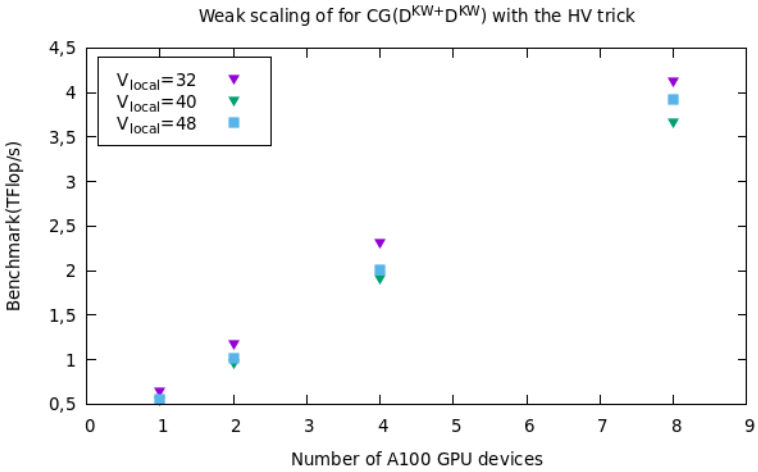
- We did dynamical simulations with Karsten-Wilczek fermions for the first time.
- We eliminated the $\mathcal{O}(a)$ errors by introducing an improvement of the Laplacian.
- We tuned the coefficients of the counter terms c, ξ and β_ξ at $a = 0.180(3)\mathrm{fm}$ nonperturbatively.
- We checked the computational cost going toward the physical pont.

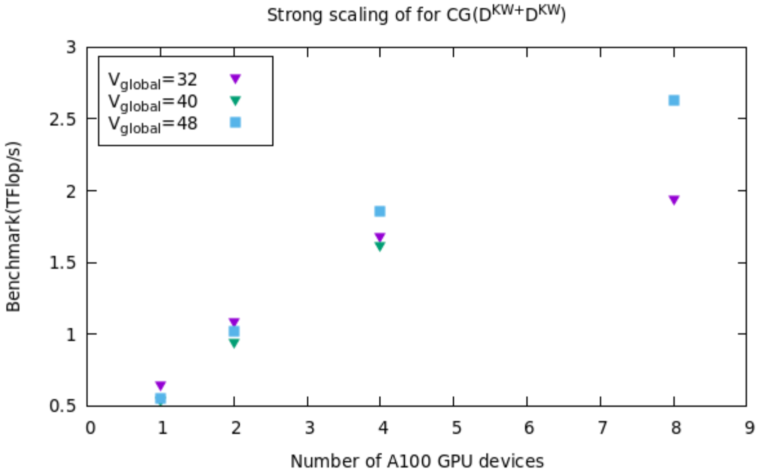
- Outlook: With S. Capitani we are working on the construction of a renormalized action with beyond tree level improvement by using perturbation theory. For this we need to determine values of $c_1 = 9/8 + g_0^2 c_1^{(1)}$ and $c_3 = -1/8 + g_0^2 c_3^{(1)}$.

Backup

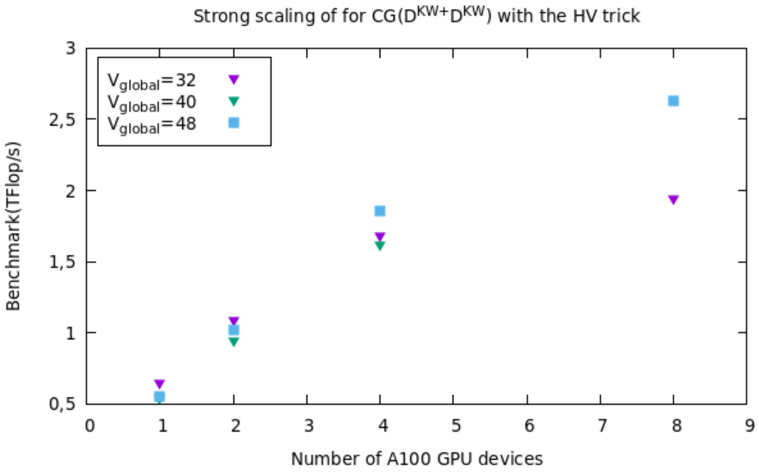


Weak scaling of for CG(D^KW+D^KW)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
Benchmark(TFlop/s)
V_local=32
V_local=40
V_local=48
0
1
2
3
4
5
6
7
8
9
Number of A100 GPU devices

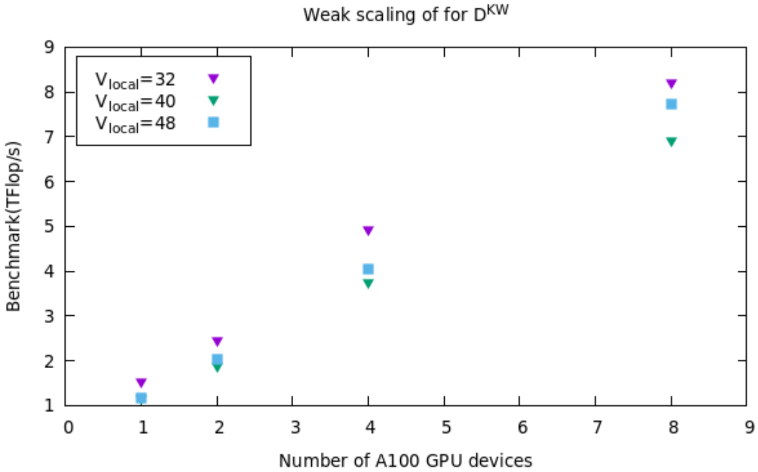

Weak scaling of for CG(D^KW+D^KW) with the HV trick
Benchmark(TFlop/s)
Number of A100 GPU devices
V_local=32
V_local=40
V_local=48
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
1
2
3
4
5
6
7
8
9

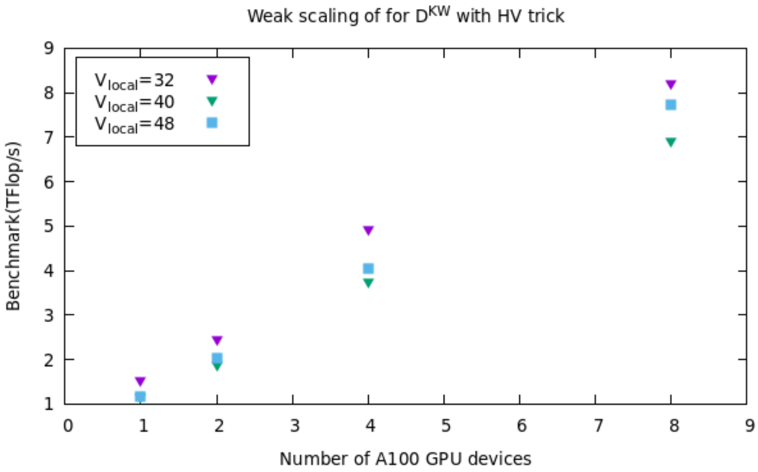


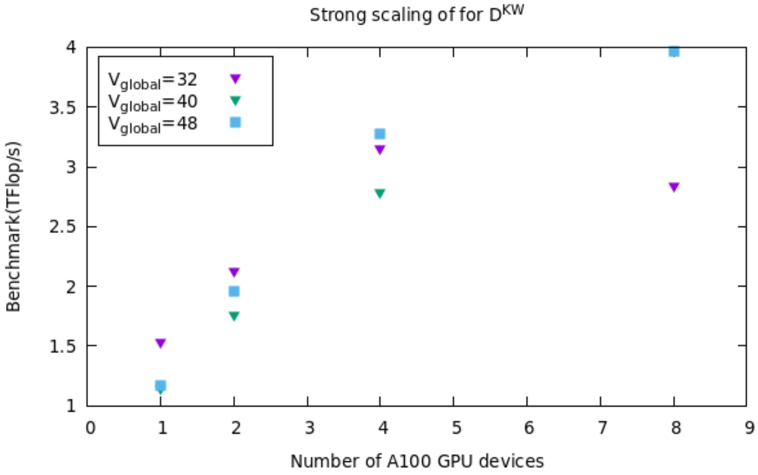
Strong scaling of for CG(D^KW+D^KW)
V_global=32
V_global=40
V_global=48
3
2.5
2
1.5
1
0.5
0
1
2
3
4
5
6
7
8
9
Benchmark(TFlop/s)
Number of A100 GPU devices



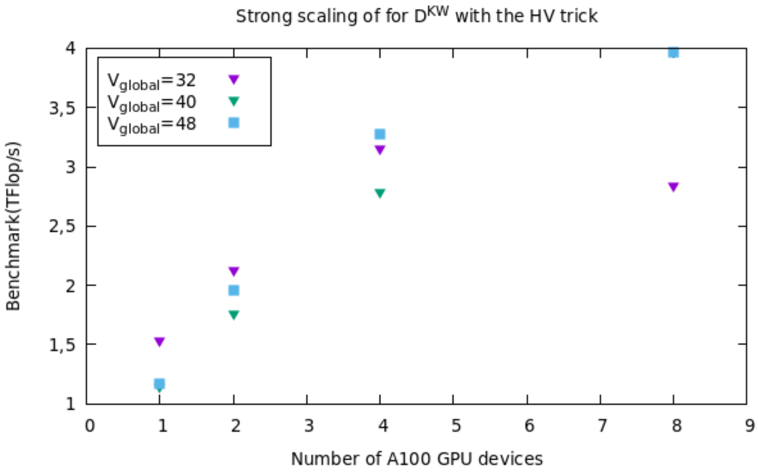
Strong scaling of for CG(D^KW+D^KW) with the HV trick
V_global=32
V_global=40
V_global=48
Benchmark(TFlop/s)
Number of A100 GPU devices
0,5
1
1,5
2
2,5
3
0
1
2
3
4
5
6
7
8
9


Weak scaling of for D^KW
V_local=32
V_local=40
V_local=48
Benchmark(TFlop/s)
1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8
9
Number of A100 GPU devices

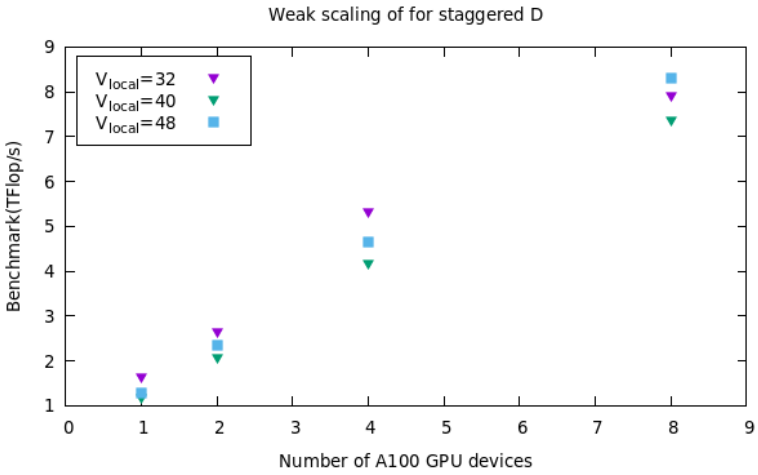

Weak scaling of for D^KW with HV trick
Benchmark(TFlop/s)
Number of A100 GPU devices
V_local=32
V_local=40
V_local=48
9
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9

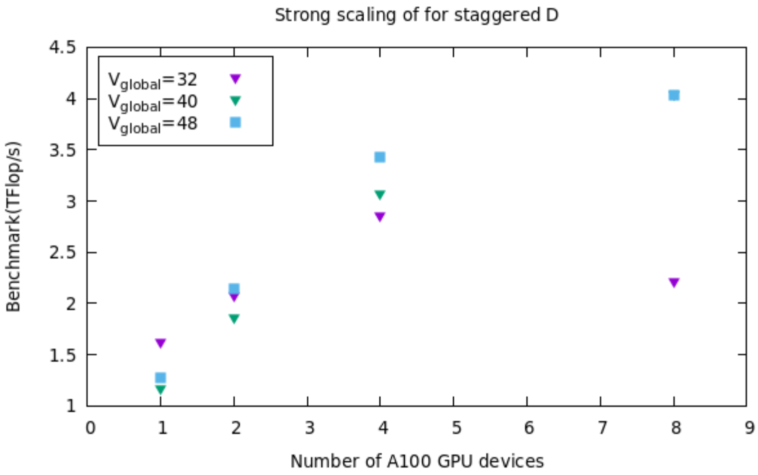


Strong scaling of for D^KW
Benchmark(TFlop/s)
Number of A100 GPU devices
Vglobal=32
Vglobal=40
Vglobal=48
4
3.5
3
2.5
2
1.5
1
0
1
2
3
4
5
6
7
8
9



Strong scaling of for D^{KW} with the HV trick
V_{global}=32
V_{global}=40
V_{global}=48
Benchmark(TFlop/s)
4
3,5
3
2,5
2
1,5
1
0
1
2
3
4
5
6
7
8
9
Number of A100 GPU devices


Weak scaling of for staggered D
Benchmark(TFlop/s)
Number of A100 GPU devices
V_{local}=32
V_{local}=40
V_{local}=48
1
2
3
4
5
6
7
8
9
0



Strong scaling of for staggered D
Benchmark(TFlop/s)
Number of A100 GPU devices
V_global=32
V_global=40
V_global=48
4.5
4
3.5
3
2.5
2
1.5
1
0
1
2
3
4
5
6
7
8
9