PROCEEDINGS OF SCIENCE

Large scale dynamical Domain-Wall Fermion simulation on GPUs: Techniques and properties

Chulwoo Jung, for RBC/UKQCD collaborations [a,*]

[a] *Brookhaven National Laboratory,*
PO Box 5000, Upton, U.S.A.

E-mail: chulwoo@bnl.gov

We describe the latest dynamical DWF ensemble generation efforts by RBC/UKQCD collaboration, focusing on $96^3 \times 192$, $a \sim 0.07$fm, 2+1 flavor ensemble with Iwasaki gauge action at physical point. Basic properties of the ensemble as well as of the algorithms employed for the evolution and basic measurements are given.

The 38th International Symposium on Lattice Field Theory, LATTICE2021 26th-30th July, 2021
Zoom/Gather@Massachusetts Institute of Technology

*Speaker

PoS(LATTICE2021)165

1. Introduction

RBC/UKQCD collaborations has been focusing on Domain Wall Fermion(DWF) formulation, which includes Möbius as well as the original Shamir construction, as the discretization of Dirac operators for their dynamical ensemble generation efforts. DWF formulation allows the remnant symmetry breaking to be controlled separately from lattice spacing by increasing the extent of the 5th dimension (L_s) and the coupling between 4-dimensional slices. Also, in contrast to Wilson fermions where the discretized Dirac operator can have poles at or near the input mass, resulting in exceptional configurations, DWF formalism guarantees the Dirac operator is safe from such poles as long as the input valence mass is positive. Taking advantage of this, RBC/UKQCD has been focusing on generation of 2+1 flavor ensembles at or near physical point lattice spacings which eliminates the systematic uncertainty from the chiral extrapolation, and allows RBC/UKQCD collaborations to fully take advantage of the good chiral symmetry properties of DWF formulation for various quantities such as $K \to \pi\pi$ [1] and muon anomalous magnetic moment [2].

Figure 1 illustrates dynamical ensembles generated by RBC/UKQCD collaborations. In addition to Iwasaki gauge action, Dislocation Suppressing Determinant Ratio (DSDR), which suppresses the chiral symmetry breaking from dislocations on larger lattice spacing, allows generation of dynamical ensembles at or near physical point with large physical volume at moderate lattice volumes, as shown in Figure 1.

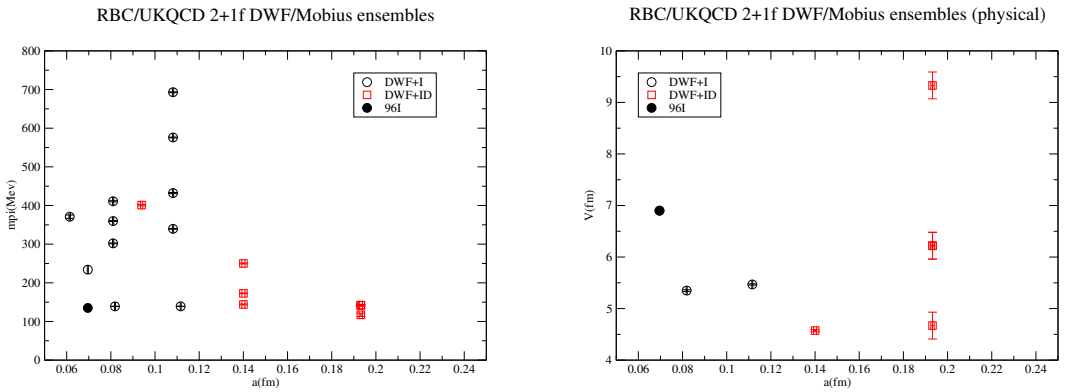


Figure 1: Parameters of 2+1-flavor dynamical DWF/Möbius ensembles generated by RBC/UKQCD. DWF+I denotes DWF with Iwasaki gauge action. DWF+ID denotes Iwasaki gauge action with Dislocation Suppressing Determinant Ratio(DSDR).

Here we describe the details of our most ambitious ensemble to date, a $96^3 \times 192$, $a \sim 0.07$fm 2+1-flavor dynamical ensemble, which we will refer to as 96I for the rest of the paper. In section 2, a description of hardware and software used for the generation of the configurations and basic measurements are given. Section 3 describes measurements of topological charge and Wilson flow scales of the 96I ensemble. Section 4 describes the preliminary results from pseudoscalar mesons and omega baryons measured on the ensemble. Discussion follows in section 5.

2. Simulation details

All of the gauge evolutions for the 96I ensemble have been done on Summit machine at Oak Ridge Leadership Computing Facility(OLCF). To save on the computing resource for thermalization, a thermalized $32^3 \times 64$ configurations, was duplicated in 4 direction. Input parameters to achieve physical point was estimated from previous studies with physical ensembles at larger lattice spacing and an $48^3 \times 96$ ensemble with the same lattice spacing as 96I, but at heavier quark masses [6].

	48I	64I	96I
Volume	$48^3 \times 96$	$64^3 \times 128$	$96^3 \times 192$
$L_s \times (b+c)$	24 ×2	12 ×2	12 ×2
β	2.13	2.25	2.31
am_l	0.00078	0.000678	0.00054
am_h	0.0362	0.02661	0.02144

Table 1: Basic parameters of 2+1 flavor DWF+I physical point ensembles

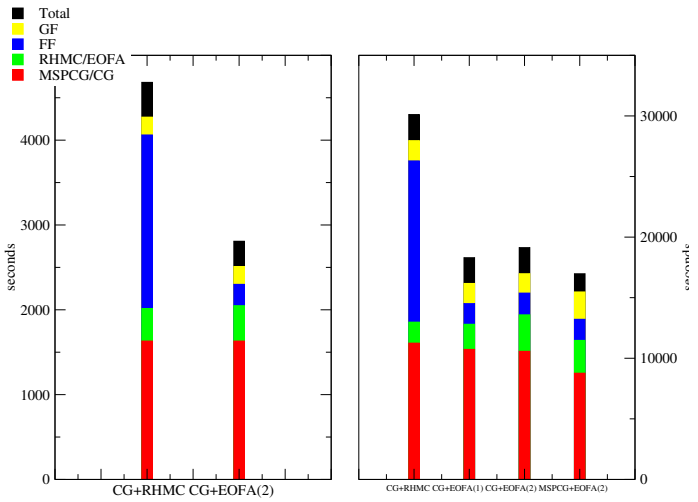
We use an exact hybrid Monte Carlo algorithm, with One Flavor Algorithm(EOFA) for strange quark, for 96I ensemble generation. 2 light quark and 1 strange quark comprised the outer level of 2-level Force Gradient integrator and Iwasaki gauge action comprised the inner. 5 Hasenbusch masses (0.0038,0.0145,0.045,0.25,0.51) were used for the light quark integrator. For a better control of spuriously large forces from the strange quark Hamiltonian, a Hasenbusch mass of 0.163 was used for the strange quark. The evolution was done with CPS[4], with QUDA[7] for optimized solvers for NVIDIA GPUs.

In comparison to algorithms and techniques used in previous physical ensembles, 2 recent developments are worth noting. One is the Multisplitting preconditions Conjugate Gradient(MSPCG)[5] which allows the QUDA inverter to minimize network traffic while maximizing the utilization of GPUs. The other is EOFA with Cayley Preconditioner[9]. It turned out that the reduction in memory footprint from utilizing EOFA for 1 flavor heavy dynamical quark instead of Rational HMC(RHMC) improves the overall simulation time of GPU based machines such as Summit even more significantly than expected due to the reduction of memory footprint, as illustrated by the reduction in time spent on 96I evolution with RHMC and EOFA in Figure 2.

Similar to previous studies, RBC/UKQCD basic measurement program on 96I took advantage of efficient generation of compressed eigenvectors, 5000 per configuration in the case of 96I. Following [3], 150 lowest eigenvectors for checkerboarded and normalized Dirac operators were generated with QUDA and CPS, after which a coarse grid Lanczos implemented with Grid[11] and GPT[12] was used to generate 5000 single precision, coarse grid eigenvectors. It turned out coarse grid with 4^4 blocking is sufficient to maintain the efficacy of low modes constructed from coarse grid eigenvectors, which resulted in a factor of ~ 50 compression for the data footprint of eigenvectors for each 96I configuration. 2 rounds of defect correction with deflation was enough to reach desired stopping condition for the light quark propagators when paired with half precision QUDA MSPCG inverter, while 1 was sufficient for the sloppy propagators.

With half precision solver used for inner solvers, deflation with the coarse grid eigenvectors reduce iteration count for light exact inversion by a factor of $20 \sim 25$. For each configurations,

h



1024 nodes
CG+RHMC/CG+EOFA(2)

128 nodes
CG+RHMC/CG+EOFA(1)/
CG+EOFA(2)/MSPCG+EOFA(2)

Figure 2: Timing breakdown of various parts of 96I evolution with different realizations of strange mass Hamiltonian.

12 wall source exact light and strange propagators were generated with 10^{-8} and 10^{-12} stopping condition respectively, while 10^{-4} and 10^{-6} was used for 48 sloppy propagators. The results of the section 4 is from measurements on configurations separated by 20 MD units, with some gaps to be filled by ongoing measurement. We employed All-mode-averaging with 5000 coarse grid eigenvectors described in section 2 for the light quark.

2.1 Mid-MD Checkpointing

Code development and testing of evolution code on Summit machine was aided by the collaboration with ECP VeloC[10] project, which developed portable and low overhead checkpointing library.

While traditional checkpointing, where after each Hybrid Monte Carlo trajectory the gauge configuration (U_μ) and Random Number Generator(RNG) state are checkpointed, has been quite adequate for most of ensemble generation so far, the job size and duration necessary for each trajectory is continuing to increase even with the rapid increase in computing power per node, especially after taking into account of the rapidly increasing autocorrelation time. Longer runtime makes it more vulnerable against hardware errors, which we have experienced in the early stages of Summit.

A portable, efficient and robust checkpointing with VeloC allows for more resiliency against queue policy and/or machine failure. While a more fine-grained checkpointing which would try to recover temporary variables from all the levels of nested MD integrator can be quite complicated, checkpointing at the outer most level with largest step size, which often are fermions, turned out

to be relatively straightforward. CPS now has native mid-MD checkpointing functioning with or without VeloC. The necessary steps are as follows:

- Register Gauge field, gauge momenta(H_μ) and Pseudofermions ($\psi(x)$)) for checkpointing when they are created.
- At the start of outermost integrator step, checks the version number of checkpointed files, and if they are higher than the current step number, update the step number and replace the content of registered fields with the checkpointed content. Otherwise, create a new checkpoint for registered fields.

3. Global topology and Wilson Flow scales on 96I ensemble

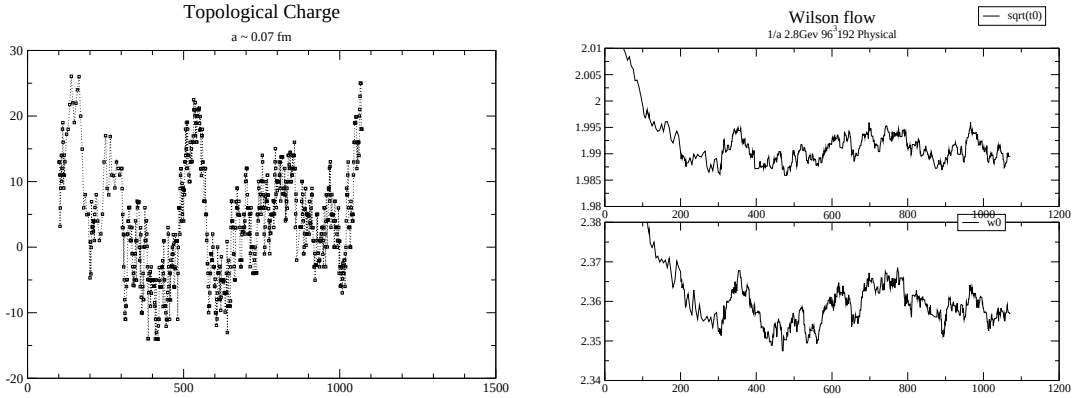


Figure 3: Evolution of topological charge Q_{top} and Wilson flow scale $t_0^{1/2}, w_0$ for 96I ensemble. Q_{top} was measured at wilson flow time 16.

Figures 3 shows the evolution of global topological charge and wilson flow scales $t_0^{1/2}$ and w_0 for 96I ensemble. An adaptive integrating scheme[13] was used for both wilson flow scale and smearing needed for the topological charge. 5Li[14] definition of topological charge was used. Measured autocorrelation time for $Q_{top}, t_0^{1/2}$, and w_0 are 54, 36, 49 MD units respectively.

4. Hadronic quantities

Following [8] (Eq.(35)), residual mass was measured by fitting the ratio between the pseudoscalar density measured at the midpoint of the 5th dimension and the physical pseudoscalar density. Similarly, Z_V was measured by fitting the temporal component of the pion electromagnetic form factors measured with different separations from 24 to 40 lattice spacing, between the source and the sink. Figure 4 shows the plateaus for both quantities.

For the Omega baryon, 3 different box sizes ($48^3, 32^3, 24^3$) with random Z3 source, described in [8], were employed to ensure the effect of excited states are under control. To increase the statistics without affected by autocorrelation, another set of Omega baryon propagators, propagated

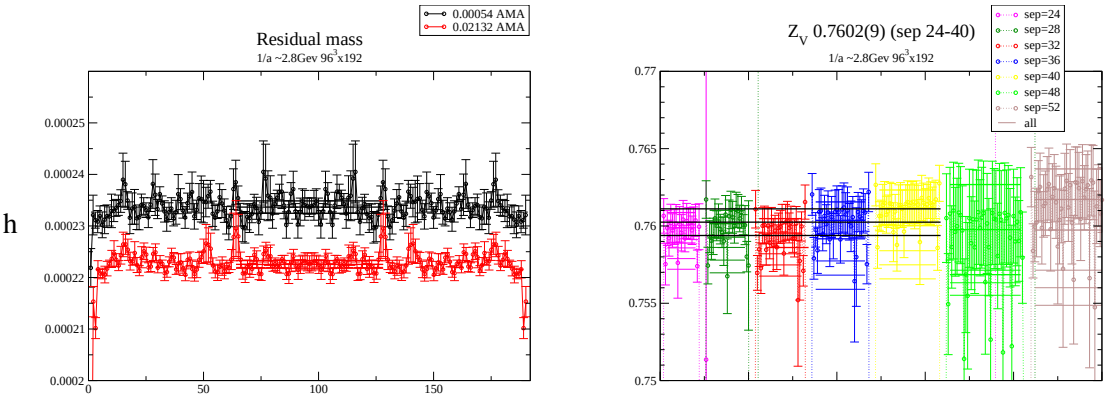
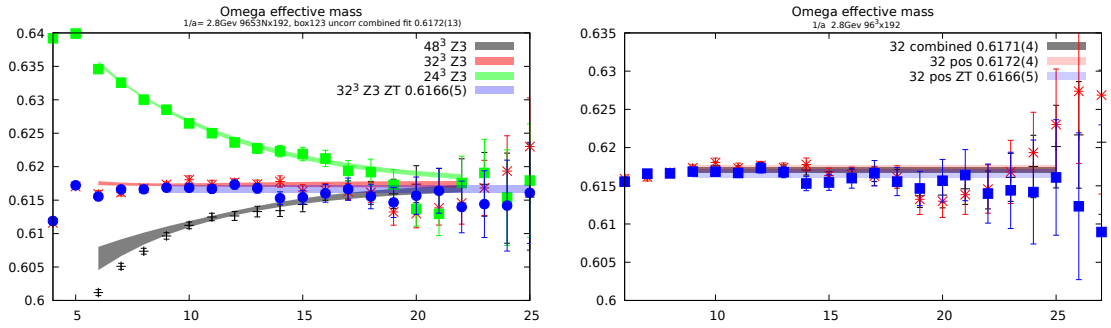


Figure 4: Residual mass and Z_V from 96I ensemble $m_{res}a \sim 2 \times 10^{-4}$



$$am_\Omega \sim 0.6171\ (4)$$

Figure 5: Effective mass plot of Ω baryon extracted from Z3 box sources on the 96I ensemble.

to z- direction instead of t- direction, with 32^3 Z3 box sources were generated. Figure 5 shows the combined fit of 4 Z3box sources to 2 exponentials. Further increase in statics with possibly also in x- and y- direction is being considered.

Figures 6 and 7 shows the effective masses of light-light and light-strange pseudoscalar correlators, PP^{LW}, PP^{WW}, AP^{LW} for each, where P, A, L, W denotes pseudoscalar operator, temporal component of Axial current, Local and Wall source/sink, respectively. Again following [8], combined fits of the 3 correlators for light-light and light-strange correlators each were used to extract m_π, f_π, m_K, f_K in Table 2.

Table 2 summarizes measured values of basic quantities and comparison with existing ensembles. It is encouraging to see the current measurement strategy, based on AMA with efficiently generated coarse-grid eigenvectors, produced comparable relative errors compared to existing ensembles, despite the relatively small number of configurations measured and the number of sloppy measurement per configurations, with a possible exception of Z_V where the relative error is somewhat larger than other ensembles. Other methods for extracting Z_V or Z_A are being studied . RBC/UKQCD has generated eigenvectors enough for a factor of 3 increase in statistics at least, and the measurement is ongoing.

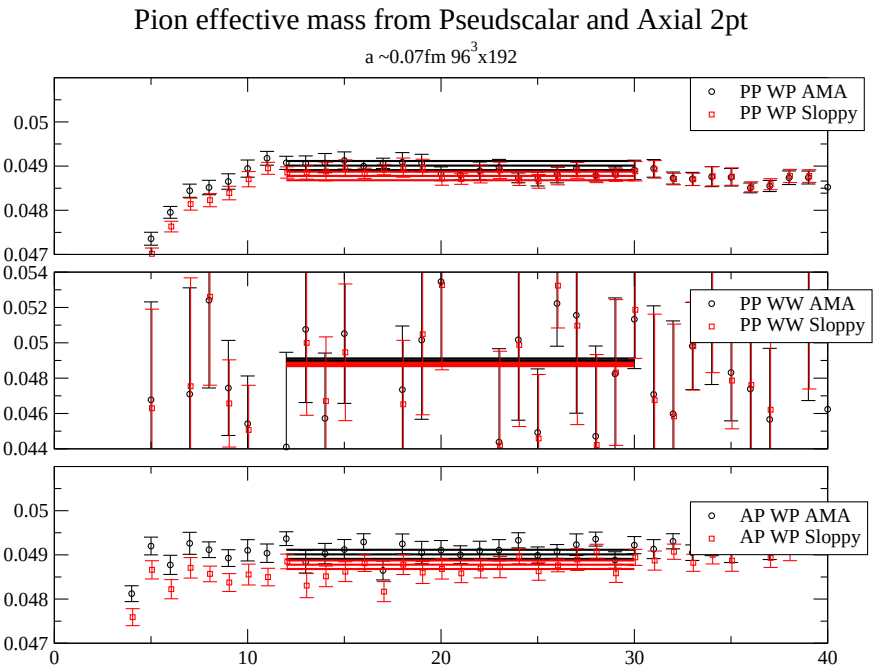


Figure 6: Fit for m_π and f_π form gauge fixed wall source propagators for 96I ensemble

	48I	64I	96I(Preliminary, stat only)
am_π	0.08049(13)	0.05903(13)	0.04901(15)
af_π	0.07580(8)	0.05550(10)	0.04817(14)
am_K	0.28853(14)	0.21531(17)	0.1808(1)
af_K	0.09040(9)	0.06653(10)	0.0575(1)
am_Ω	0.9702(10)	0.7181(7)	0.6171(4)
Z_V	0.71076(25)	0.74293(14)	0.7602(9)
${t_0}^{1/2}/a$	1.29659(28)	1.74496(62)	1.9897(7)
w_0/a	1.50125(94)	2.0495(15)	2.356(2)

Table 2: Comparison of basic quantities measured on the 96I ensemble with existing ensembles, from [8].

5. Summary and Discussion

The newly generated 96I DWF+I ensemble provides RBC/UKQCD collaborations with an ensemble with physical mass at the third lattice spacing. A description of algorithms employed in the generation of the ensemble and preliminary results of gluonic and Hadronic quantities were presented.

While the measured autocorrelation has been negligible for Hadronic quantities, factor of $\sim$ 3 increase in measurements will clarify whether autocorrelation is a concern. Master-field type studies which considers spatial correlation within a configuration as well as autocorrelation between different configurations on the same spatial position are ongoing, which hopefully will enable us to

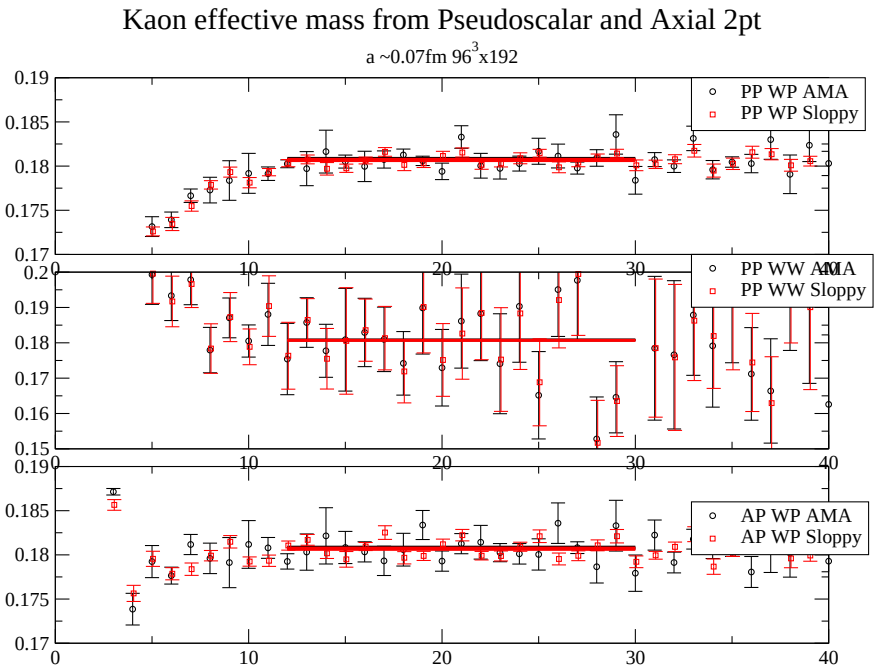


Figure 7: Fit for m_K and f_K form gauge fixed wall source propagators for 96I ensemble

get estimates of statistical uncertainty not limited by the number of configurations measured.

DWF ensemble generation on new and upcoming machines are likely to continue to be limited by the memory and internode bandwidth. While EOFA and MSPCG has helped mitigating these issues, evolution is more vulnerable compared to measurements, where various techniques such as exact deflation, AMA, A2A, and Split Grid are already available to mitigate, if not eliminate, bandwidth issues.

Current lattice spacing was tuned from previous studies to achieve reasonable topology tunneling within the typical number of configurations generated for each ensemble. Making the lattice spacing finer requires a significant increase in the computing resource required to generate similar number of independent configurations scale as $\sim$ (lattice spacing)$^{-10}$. While there are techniques such as open boundary conditions to control the steep increase of autocorrelation time in topological charge, the efficacy of these techniques on other quantities such as wilson flow scales is less clear. RBC/UKQCD has been pursuing various algorithmic methods to better control critical slowing down in QCD ensemble generation both as a part of US Exascale project, and independent of ECP. We refer to [16, 17] for these studies.

Acknowledgement

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. CJ was supported in part by the US DOE Contract No. DE-SC0012704.

References

[1] R. Abbott *et al.* [RBC and UKQCD], Phys. Rev. D **102**, no.5, 054509 (2020) doi:10.1103/PhysRevD.102.054509 [arXiv:2004.09440 [hep-lat]].

[2] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung and C. Lehner, Phys. Rev. Lett. **124**, no.13, 132002 (2020) doi:10.1103/PhysRevLett.124.132002 [arXiv:1911.08123 [hep-lat]].

[3] M. A. Clark, C. Jung and C. Lehner, EPJ Web Conf. **175**, 14023 (2018) [arXiv:1710.06884 [hep-lat]].

[4] C. Jung [RBC and UKQCD], PoS **LATTICE2013**, 417 (2014), https://github.com/RBC-UKQCD/CPS

[5] J. Tu, M. A. Clark, C. Jung and R. Mawhinney, doi:10.1145/3468267.3470613 [arXiv:2104.05615 [hep-lat]].

[6] P. Boyle, L. Del Debbio, F. Erben, A. Jüttner, T. Kaneko, M. Marshall, A. Portelli, J. T. Tsang and O. Witzel, [arXiv:2111.11287 [hep-lat]].

[7] https://github.com/lattice/quda

[8] T. Blum *et al.* [RBC and UKQCD], Phys. Rev. D **93**, no.7, 074505 (2016) doi:10.1103/PhysRevD.93.074505 [arXiv:1411.7017 [hep-lat]].

[9] C. Jung, C. Kelly, R. D. Mawhinney and D. J. Murphy, Phys. Rev. D **97**, no.5, 054503 (2018) [arXiv:1706.05843 [hep-lat]]; Y. C. Chen *et al.* [TWQCD], Phys. Lett. B **738**, 55-60 (2014) [arXiv:1403.1683 [hep-lat]].

[10] https://github.com/ECP-VeloC/VELOC

[11] P. Boyle, A. Yamaguchi, G. Cossu and A. Portelli, [arXiv:1512.03487 [hep-lat]]. https://github.com/paboyle/Grid

[12] *Grid Python Toolkit (GPT)*, C. Lehner, This conference. https://github.com/lehner/gpt

[13] P. Fritzsch and A. Ramos, JHEP **10**, 008 (2013) doi:10.1007/JHEP10(2013)008 [arXiv:1301.4388 [hep-lat]].

[14] P. de Forcrand, M. Garcia Perez and I. O. Stamatescu, Nucl. Phys. B **499**, 409-449 (1997) [arXiv:hep-lat/9701012 [hep-lat]].

[15] *Algorithms for domain wall Fermions*, P. Boyle, This conference.

[16] T. Nguyen, P. Boyle, N. Christ, Y. C. Jang and C. Jung, [arXiv:2112.04556 [hep-lat]].

[17] *Field-Transformation HMC algorithm*, L. Jin, This conference.