O(*a*)**-improved QCD+QED Wilson Dirac operator on GPUs** openQxD with QUDA

Roman Gruber

RC* collaboration: Anian Altherr Isabel Campos-Plasencia, Alessandro Cotellucci, Alessandro De Santis, Roman Gruber, Tim Harris, Javad Komijani, Marina Marinkovic, Francesca Margari, Letizia Parato, Agostino Patella, Gaurav Ray, Sara Rosso, Nazario Tantalo, and Paola Tavella

July 30, 2024

Intro / Overview

- 1. [Motivation](#page-2-0)
- 2. [Interfacing openQxD with QUDA](#page-10-0)
- 3. [Solver interface](#page-24-0)
- 4. [Performance](#page-28-0)
- 5. [Conclusion](#page-31-0)

Motivation

openQxD [\[5\]](#page-34-0)

- Simulations of QCD and QCD+QED *O*(*a*) improved Wilson-Clover fermions
- **Based on openQCD v1.6 [\[1,](#page-34-1) [2\]](#page-34-2)**
- Variety of BCs; open/SF/periodic in time, C* boundaries [\[3\]](#page-34-3) or periodic boundaries in space
- **Powerful solvers: CGNE, GCR with Schwarz-alternating procedure and** inexact deflation [\[4\]](#page-34-4)
- Pure-MPI parallelisation, C89 standard (next release will be C99)
- Actively developed and maintained by RC^{*} collaboration

Requirement

C [⋆] boundaries and QCD+QED Wilson-Clover fermions

Main Goal

Offload solves to GPU (target system: new Alps machine and Lumi-G)

Main Goal

Offload solves to GPU (target system: new Alps machine and Lumi-G)

\rightarrow OpenMP offloading

- $+$ Easy and rapid porting
- − Disappointing results (more efforts required)

Main Goal

Offload solves to GPU (target system: new Alps machine and Lumi-G)

\rightarrow OpenMP offloading

- $+$ Easy and rapid porting
- − Disappointing results (more efforts required)

\rightarrow Coupling to existing solver suite

- − Operator on GPU still a problem
- − General solvers not on eye level with state-of-the-art lattice solvers

Main Goal

Offload solves to GPU (target system: new Alps machine and Lumi-G)

\rightarrow OpenMP offloading

- $+$ Easy and rapid porting
- − Disappointing results (more efforts required)

\rightarrow Coupling to existing solver suite

- − Operator on GPU still a problem
- − General solvers not on eye level with state-of-the-art lattice solvers

\rightarrow Own CUDA/HIP implementation in openQxD

- $+$ Cleanest solution (no external dependencies)
- − Insane effort (lots of core changes, breaking changes, ...)

Main Goal

Offload solves to GPU (target system: new Alps machine and Lumi-G)

\rightarrow OpenMP offloading

- $+$ Easy and rapid porting
- − Disappointing results (more efforts required)

\rightarrow Coupling to existing solver suite

- − Operator on GPU still a problem
- − General solvers not on eye level with state-of-the-art lattice solvers

\rightarrow Own CUDA/HIP implementation in openQxD

- $+$ Cleanest solution (no external dependencies)
- − Insane effort (lots of core changes, breaking changes, ...)

\rightarrow Coupling to QUDA

- + No need to reinvent the wheel
- $+$ Get all features of QUDA (solver suite, eigensolvers, ...)
- − Only real additional efforts: (1) Interface, (2) C[⋆] boundaries, (3) QCD+QED Wilson-Clover

- **Plug and play library to offload Dirac solves**
- Supports many lattice discretisations (Wilson, staggered, Domain-wall, ...)
- Powerful solvers: BiCGstab, GCR with multigrid [\[6,](#page-34-5) [7\]](#page-34-6), ...
- \blacksquare C++-14 standard
- Supports NVIDIA, AMD, Intel and CPU threading
- Actively developed and maintained by NVIDIA + many others
- NVIDIA licence (similar to MIT)

Interfacing openQxD with QUDA

openQxD: memory layout I

Figure: Complex double struct

Figure: Clover field struct

Figure: Gauge field struct

■ Gauge field d.o.f: 4*V* (V = lattice volume, 8 directions)

■ Clover field d.o.f: 2*V* (V, 2 chiralities, 6x6 matrix (complex, Hermitian))

openQxD: memory layout II

Figure: SU(3) vector struct

Figure: Spinor field struct

Spinor field d.o.f: *V* (*V* = lattice volume, 4 spin, 3 color) $→$ array of structs

Different gauge field layouts

openQxD

- ▶ stores 8 (forward and backward) directed gauge fields for all odd-parity points
- ▶ locally stores gauge fields on the boundaries only for odd-parity points and not for even-parity points
- OUDA
	- ▶ 4 gauge fields for each space-time point (one for each positive direction

Figure: 2D example $(4 \times 4 \text{ local lattice})$ of how and which gauge fields are stored in memory in openQxD (left) and QUDA (right). Filled lattice points are even, unfilled odd lattice points.

Interface C [⋆] boundaries QCD+QED Wilson-Clover

STATUS

Ø Interface C [⋆] boundaries QCD+QED Wilson-Clover C [⋆] boundaries

Figure: 2D example of a 6×6 lattice with C^* boundary conditions on both directions. We have the (doubled) x-direction (horizontal) and a direction with C^* boundaries (vertical). Left is the physical, right the mirror lattice. The union is the extended lattice

.

- Analogue to the implementation in openQCD
- Doubling the lattice as it comes from openQxD (i.e. additional index: physical, mirror)
- Communication grid topology struct now contains a member property cstar \longrightarrow number of spatial C* directions
- comm_rank_displaced(): calculates the neighbouring rank number given one of (positive or negative) 8 directions \rightarrow implements the shifted boundaries

STATUS

Ø Interface C [⋆] boundaries QCD+QED Wilson-Clover

STATUS

QCD+QED

In addition to the $SU(3)$ -valued gauge field $U_{\mu}(x)$, we have the $U(1)$ -valued gauge field $A_{\mu}(x)$

Combined: $\bm{\mathsf{U}}(3)$ -valued field $e^{iqA_{\mu}(x)}\bm{\mathsf{U}}_{\mu}(x)$ with q_f the charge of a quark

In QUDA, we just use

- ▶ QUDA_RECONSTRUCT_9
- ▶ QUDA_RECONSTRUCT_13
- ▶ QUDA_RECONSTRUCT_NO

■ We have an $U(1)$ SW-term,

$$
D_{\rm w} \to D_{\rm w} + q c_{\rm sw}^{U(1)} \frac{i}{4} \sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu} \hat{A}_{\mu\nu} , \qquad (1)
$$

where q is the charge and the $\mathit{U}(1)$ and $\hat{A}_{\mu\nu}(\mathit{x})$ is the field strength tensor.

QCD+QED: Implementation in QUDA

- Resulting term has the same properties as the *SU*(3) SW-term (Hermitian, diagonal w.r.t chiralities)
- Clover field reorder class: openQxD (row-major):

 $\int u_0 u_6 + i u_7 u_8 + i u_9 u_{10} + i u_{11} u_{12} + i u_{13} u_{14} + i u_{15}$ $\overline{}$ · *u*₁ *u*₁₆ + *iu*₁₇ *u*₁₈ + *iu*₁₉ *u*₂₀ + *iu*₂₁ *u*₂₂ + *iu*₂₃ · · · · *u*₂ *u*₂₄ + *iu*₂₅ *u*₂₆ + *iu*₂₇ *u*₂₈ + *iu*₂₉ · · · *u*³ *u*³⁰ + *iu*³¹ *u*³² + *iu*³³ · · · · · · · · · · · · · · · **u**₄ $u_{34} + iu_{35}$ · · · · · *u*⁵ \setminus $\overline{}$. (2)

QUDA (column-major):

$$
\begin{pmatrix}\n u_0 & \cdots & \cdots & \cdots \\
 u_6 + i u_7 & u_1 & & \cdots & \vdots \\
 u_8 + i u_9 & u_{16} + i u_{17} & u_2 & \cdots & \vdots \\
 u_{10} + i u_{11} & u_{18} + i u_{19} & u_{24} + i u_{25} & u_3 & \cdots \\
 u_{12} + i u_{13} & u_{20} + i u_{21} & u_{26} + i u_{27} & u_{30} + i u_{31} & u_4 \\
 u_{14} + i u_{15} & u_{22} + i u_{23} & u_{28} + i u_{29} & u_{32} + i u_{33} & u_{34} + i u_{35} & u_5\n\end{pmatrix}.
$$
\n(3)

STATUS

STATUS

Ø Interface \bullet C* boundaries **◆ QCD+QED Wilson-Clover**

Solver interface

Solver interface in openQxD

- Solvers are called by means of their function, i.e. cgne(), sap_gcr(), dfl_sap_gcr()
- **Usual utility:**
	- \blacktriangleright input file parsing
	- \blacktriangleright solver setup
	- \blacktriangleright call solver

Figure: Example solver sections

Additional solver type

Add solver type QUDA

■ All options from QudaInvertParam and QudaMultigridParam

Figure: Example QUDA solver section

- No doubling of the gauge field
- Calculate U(1) SW-term in QUDA (no transfer)
- Offload smearing, contractions
- Spinor field memory management (field unification)
- **Partitioning**
- multiple RHS

Performance

Tested system

- Tödi testing system at CSCS, Switzerland
- 4x NVIDIA[®] Grace[™] CPU, 120GB RAM, 72 Neoverse V2 Armv9 cores
- 4x NVIDIA[®] H100 GPU, 96GB RAM
- NVLink[®] provides all-to-all cache-coherent memory between all host and device memory

[Wikipedia,](https://commons.wikimedia.org/w/index.php?curid=170969) Niklausschreiber2, [CC BY-SA 3.0](https://creativecommons.org/licenses/by-sa/3.0/)

Figure: Tödi: highest mountain in the Glarus Alps (3612 m)

Inverter scaling

Figure: Strong scaling of one inversion of the Dirac operator; $T \times L^3 = 128 \times 64^3$, m_{π} = 300 MeV, C^{*}-boundaries in all 3 spatial directions.

- GDR not yet available on Alps
- NVSHMEM not yet available on Alps

Conclusion

CONCLUSIONS

- **O** Up and running interface to QUDA
- \bullet C* boundaries in QUDA
- **O** QCD+QED Wilson-Clover in QUDA
- **O** Offloaded Dirac solves and eigensolver
- **Q** Contractions
- **D** Smearing
- **O** Field memory manager

Thanks for listening!

References I

- [1] M. Lüscher et al., *Openqcd, simulation programs for lattice qcd***,** (2012)
- [2] M. Lüscher and S. Schaefer, **"Lattice QCD with open boundary conditions and twisted-mass reweighting",** [Comput. Phys. Commun.](https://doi.org/10.1016/j.cpc.2012.10.003) **184**, 519–528 (2013), arXiv:[1206.2809 \[hep-lat\]](https://arxiv.org/abs/1206.2809).
- [3] A. S. Kronfeld and U. J. Wiese, **"SU(N) gauge theories with C-periodic boundary conditions (I). Topological structure",** [Nuclear Physics B](https://doi.org/10.1016/0550-3213(91)90479-H) **357**, 521–533 (1991).
- [4] M. Lüscher, **"Local coherence and deflation of the low quark modes in lattice qcd",** [Journal of High Energy Physics](https://doi.org/10.1088/1126-6708/2007/07/081) **2007**, 081 (2007), eprint: <0706.2298>,
- [5] I. Campos et al., **"openQ*D code: a versatile tool for QCD+QED simulations",** [The](https://doi.org/10.1140/epjc/s10052-020-7617-3) [European Physical Journal C](https://doi.org/10.1140/epjc/s10052-020-7617-3) **80**, 1–24 (2020), eprint: <1908.11673>.
- [6] R. Babich et al., **"Adaptive multigrid algorithm for the lattice Wilson-Dirac operator",** Phys. Rev. Lett. **105**[, 201602 \(2010\),](https://doi.org/10.1103/PhysRevLett.105.201602) arXiv:[1005.3043 \[hep-lat\]](https://arxiv.org/abs/1005.3043).
- [7] J. Espinoza-Valverde, A. Frommer, G. Ramirez-Hidalgo, and M. Rottmann, **"Coarsest-level improvements in multigrid for lattice QCD on large-scale computers",** [Comput. Phys. Commun.](https://doi.org/10.1016/j.cpc.2023.108869) **292**, 108869 (2023), arXiv:[2205.09104](https://arxiv.org/abs/2205.09104) [\[math.NA\]](https://arxiv.org/abs/2205.09104).

References II

[8] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, **"Solving lattice qcd systems of equations using mixed precision solvers on gpus",** [Computer Physics](https://doi.org/10.1016/j.cpc.2010.05.002) Communications **181**[, 1517–1528 \(2010\),](https://doi.org/10.1016/j.cpc.2010.05.002) eprint: <0911.3191>.

Appendix

openQxD: spacetime ordering I

txyz-convention, i.e. 4-vector $x = (x_0, x_1, x_2, x_3)$

Example 1 Lexicographical index $(L_u = \text{rank-local lattice extent})$:

$$
\Lambda(x, L) := L_3 L_2 L_1 x_0 + L_3 L_2 x_1 + L_3 x_2 + x_3. \tag{4}
$$

- openQxD orders indices in cache-blocks: decomposition of the rank-local lattice into equal blocks of extent *B*_u
	- ▶ Within a block: Λ(*b*, *B*), where *b* = block-local Euclidean 4-vector
	- \blacktriangleright Block themselves: $\Lambda(n, N_B)$, where $N_{B,u} = L_u/B_u$ and $n_u = |x_u/B_u|$
- Even-odd ordering in the block (but not the blocks themselves)

$$
\hat{x} = \left[\frac{1}{2} \left(V_B \Lambda(n, N_B) + \Lambda(b, B) \right) \right] + P(x) \frac{V}{2}, \tag{5}
$$

where $V_B = B_0 B_1 B_2 B_3$ is the volume of a block, $P(x) = \frac{1}{2}(1 - (-1)^{\sum_{\mu} x_{\mu}})$ gives the parity and $V = L_3L_2L_1L_0$.

openQxD: spacetime ordering II

Figure: 2D example (8 × 8 local lattice) of the rank-local unique lattice index in openQxD (in time first convention (txyz)). The blue rectangles denote cache blocks of size 4×4 . Gray sites are odd, white sites are even lattice points.

The QCD+QED C^* Wilson-Clover Dirac operator in QCD simulations applied onto a spinor field $\psi(x)$ is (the lattice spacing is set to $a = 1$)

$$
D_{\rm w}\psi(x) = (4+m_0)\psi(x)
$$

\n
$$
-\frac{1}{2}\sum_{\mu=0}^{3} \left\{ H_{\mu}(x)(1-\gamma_{\mu})\psi(x+\hat{\mu}) + H_{\mu}(x-\hat{\mu})^{-1}(1+\gamma_{\mu})\psi(x-\hat{\mu}) \right\}
$$

\n
$$
+c_{\rm sw}^{SU(3)}\frac{i}{4}\sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi(x) + qc_{\rm sw}^{U(1)}\frac{i}{4}\sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu}\hat{A}_{\mu\nu}\psi(x),
$$
\n(6)

where the gauge field $H_\mu(x)$ is the $U(3)$ -valued link between extended lattice point *x* and $x + \hat{\mu}$, the γ_{μ} are the Dirac matrices obeying the Euclidean Clifford algebra, $\{\gamma_\mu,\gamma_\nu\}=2\delta_{\mu\nu}$ and $\sigma_{\mu\nu}=\frac{i}{2}\,[\gamma_\mu,\gamma_\nu].$

The $SU(3)$ field strength tensor \hat{F} is defined as

$$
\hat{F}_{\mu\nu}(x) = \frac{1}{8} \left\{ Q_{\mu\nu}(x) - Q_{\nu\mu}(x) \right\}, \nQ_{\mu\nu}(x) = U_{\mu}(x) U_{\nu}(x + \hat{\mu}) U_{\mu}(x + \hat{\nu})^{-1} U_{\nu}(x)^{-1} \n+ U_{\nu}(x) U_{\mu}(x - \hat{\mu} + \hat{\nu})^{-1} U_{\nu}(x - \hat{\mu})^{-1} U_{\mu}(x - \hat{\mu}) \n+ U_{\mu}(x - \hat{\mu})^{-1} U_{\nu}(x - \hat{\mu} - \hat{\nu})^{-1} U_{\mu}(x - \hat{\mu} - \hat{\nu}) U_{\nu}(x - \hat{\nu}) \n+ U_{\nu}(x - \hat{\nu})^{-1} U_{\mu}(x - \hat{\nu}) U_{\nu}(x + \hat{\mu} - \hat{\nu}) U_{\mu}(x)^{-1}
$$

where the gauge field $U_{\mu}(x)$ is $SU(3)$ -valued

We add the $U(1)$ SW-term,

$$
D_{\rm w} \to D_{\rm w} + q c_{\rm sw}^{U(1)} \frac{i}{4} \sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu} \hat{A}_{\mu\nu} , \qquad (7)
$$

where q is the charge and the $\bm{\mathsf{U}}(\bm{\mathsf{1}})$ field strength tensor $\hat A_{\mu\nu}(\bm{\mathsf{x}})$ is defined as

$$
\hat{A}_{\mu\nu}(x) = \frac{i}{4q_{el}} Im \{ z_{\mu\nu}(x) + z_{\mu\nu}(x - \hat{\mu})
$$

$$
+ z_{\mu\nu}(x - \hat{\nu}) + z_{\mu\nu}(x - \hat{\mu} - \hat{\nu}) \}
$$

$$
z_{\mu\nu}(x) = e^{i \{ A_{\mu}(x) + A_{\nu}(x + \hat{\mu}) - A_{\mu}(x + \hat{\nu}) - A_{\nu}(x) \}}
$$

C [⋆] boundary conditions

The implementation of the C* boundary conditions for the fields is the following (orbifold construction):

$$
A_{\mu}(x + L_{\hat{k}}\hat{k}) = -A_{\mu},
$$

\n
$$
\psi_f(x + L_{\hat{k}}\hat{k}) = C^{-1}\overline{\psi}_f^T(x),
$$

\n
$$
\overline{\psi}_f(x + L_{\hat{k}}\hat{k}) = -\psi_f^T(x)C,
$$

\n
$$
U_{\mu}(x + L_{\hat{k}}\hat{k}) = U^* \mu(x),
$$
\n(8)

where L_k is the size of the lattice in direction \hat{k} , U^* denotes complex conjugation. The charge–conjugation matrix C satisfies

$$
C^{T} = -C, \quad C^{\dagger} = C^{-1}, \quad C^{-1} \gamma_{\mu} C = -\gamma_{\mu}^{T}.
$$
 (9)

The gauge action is

$$
S_{g, SU(3)} = \frac{1}{g_0^2} \sum_{C \in S_0} \text{tr} [1 - U(C)], \qquad (10)
$$

$$
S_{g,U(1)} = \frac{1}{2q_{el}^2 e_0^2} \sum_{C \in S_0} \text{tr} [1 - z(C)], \qquad (11)
$$

where the bare coupling constants are $q_0, \rho_0, q_{el} = 1/6$. Given a path C on a lattice, $U(C)$ and $Z(C)$ denote the $SU(3)$ and $U(1)$ parallel transport along C.

- On the extended lattice, points x and $x + L_k\hat{k}$ do not coincide!
- Admissible fields are given by the boundary conditions
- Admissible gauge fields on mirror lattice are completely determined by their value on the physical lattice
- \blacksquare On physical lattice: ψ and $\bar{\psi}$ are independent Grassmann variables
- \blacksquare On extended lattice: $\bar{\psi}$ is completely determined by ψ
- Integration measure for fermion field:

$$
[\mathrm{d}\psi]_{\Lambda_{phys}} [\mathrm{d}\bar{\psi}]_{\Lambda_{phys}} = \prod_{x \in \Lambda_{phys}} \mathrm{d}\psi(x)\bar{\psi}(x) = \prod_{x \in \Lambda_{exended}} \mathrm{d}\psi(x) = [\mathrm{d}\psi]_{\Lambda_{extended}} \qquad (12)
$$

 \implies We need the doubled lattice for the fermion field!

Dirac operator scaling I

Figure: C [⋆] Wilson-Clover Dirac operator strong scaling

- GDR not yet available on Alps
- NVSHMEM not yet available on Alps

Dirac operator scaling II

Figure: C [⋆] Wilson-Clover Dirac operator weak scaling

- GDR not yet available on Alps
- NVSHMEM not yet available on Alps

Unification of fields

- Initial code: all functions implemented in CPU \rightarrow no transfers needed
- **I** Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \rightarrow we'll probably never reach that
- Intermediate phase: some functions are ported to GPU, but not all of them \rightarrow needs transfers
- **Initial code:** all functions implemented in CPU \rightarrow no transfers needed
- **I** Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \rightarrow we'll probably never reach that
- Intermediate phase: some functions are ported to GPU, but not all of them \rightarrow needs transfers

Requirement 1

We don't want to rewrite every program, when a new function is ported to GPU!

- **Initial code:** all functions implemented in CPU \rightarrow no transfers needed
- **I** Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \rightarrow we'll probably never reach that
- Intermediate phase: some functions are ported to GPU, but not all of them \rightarrow needs transfers

Requirement 1

We don't want to rewrite every program, when a new function is ported to GPU!

Requirement 2

Fully backwards compatible with openQxD's memory layout

openQxD: overloading of functions I

Figure: Example overloading of functionA.

openQxD: overloading of functions II

```
1 #if (defined AVX)
2 // implementation using AVX intrinsics
3 void functionA(spinor dble *s) { ... }
4 #elif (defined x64)
5 // implementation using SSE2 intrinsics
6 void functionA(spinor dble \stars) { ... }
7 #elif (defined GPU_OFFLOADING)
8 // GPU overloading of the function
9 void functionA(spinor dble *s) { ... }
10 \#else
11 // default implementation
12 void functionA(spinor dble *s) { ... }
13 #endif
```
Figure: Example overloading of functionA.

Figure: Each field with openQxD corresponds to a field within QUDA.

- openQxD operates on base pointers of struct-arrays
- Establish a 1-1 correspondence between CPU/GPU fields
- \implies Everytime (de-)allocating a field \rightarrow (de-)allocate on both devices
- \implies Maintain consistency among the two fields (CPU/GPU manipulates field)

Figure: Current field allocation scheme.

Figure: New field allocation scheme (spinor_info struct *after* the data).

Information held by the spinor_info struct:

- Field status: CPU_NEWER, GPU_NEWER, IN_SYNC
- GPU pointer: pointer to field on the GPU (i.e. pointer to ColorSpinorField instance)
- Other information: eg. field size in bytes, stats, ...
- \implies Only changes in the (de-)allocation functions: alloc wsd(). reserve_wsd(), release_wsd() + their single precision variants

PROCEDURE

Functions within openQxD still operate on base pointers (in the same way as before!) \implies they all still work (no change needed)

GPU-offloaded functions now take the same CPU base pointer

- 1. Navigate to the spinor info struct
- 2. Check if field needs to be transferred
- 3. Transfer if needed
- 4. Obtain GPU field pointer from info struct
- 5. Update status field in info struct
- 6. Continue function body with GPU field

■ openQxD functions take the usual CPU base pointer

- 1. Navigate to the spinor info struct
- 2. Check if field needs to be transferred
- 3. Transfer if needed
- 4. Update status field in info struct
- 5. Continue function body with CPU field