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, RO-MAN GRUBER, TIM HARRIS, JAVAD KOMIJANI, MARINA MARINKOVIC, FRANCESCA MARGARI, LETIZIA PARATO, AGOSTINO PATELLA, GAURAV RAY, SARA ROSSO, NAZARIO TANTALO, AND PAOLA TAVELLA

DECEMBER 11, 2024

INTRO / OVERVIEW

- 1. Motivation
- 2. Interfacing openQxD with QUDA
- 3. Solver interface
- 4. Performance
- 5. Conclusion

MOTIVATION

OPENQXD [5]





- Simulations of QCD and QCD+QED *O*(*a*) improved Wilson-Clover fermions
- Based on openQCD v1.6 [1, 2]
- Variety of BCs; open/SF/periodic in time, C* boundaries [3] or periodic boundaries in space
- Powerful solvers: CGNE, GCR with Schwarz-alternating procedure and inexact deflation [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)

➔ 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)

➔ OpenMP offloading

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

→ 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)

➔ OpenMP offloading

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

→ Coupling to existing solver suite

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

→ 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)

➔ OpenMP offloading

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

→ Coupling to existing solver suite

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

→ Own CUDA/HIP implementation in openQxD

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

➔ 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, 7], ...
- 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

<pre>1 /* Complex double struct */</pre>	1 /* Clover field struct */
2 typedef struct	2 typedef struct
3 {	3 {
4 double re,im;	4 double u[36];
<pre>5 } complex_dble;</pre>	<pre>5 } pauli_dble;</pre>

Figure: Complex double struct

Figure: Clover field struct



Figure: Gauge field struct

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

Clover field d.o.f: 2V (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
- QUDA
 - ▶ 4 gauge fields for each space-time point (one for each positive direction



Figure: 2D example (4×4 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



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 → number of spatial C* directions
- comm_rank_displaced(): calculates the neighbouring rank number given one of (positive or negative) 8 directions → implements the shifted boundaries

10

STATUS





C* boundaries QCD+QED Wilson-Clover

STATUS



QCD+QED Wilson-Clover

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: U(3)-valued field $e^{iqA_{\mu}(x)}U_{\mu}(x)$ 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,

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

where q is the charge and the U(1) and $\hat{A}_{\mu\nu}(x)$ 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):

 $\begin{pmatrix} u_{0} & u_{6} + iu_{7} & u_{8} + iu_{9} & u_{10} + iu_{11} & u_{12} + iu_{13} & u_{14} + iu_{15} \\ \cdot & u_{1} & u_{16} + iu_{17} & u_{18} + iu_{19} & u_{20} + iu_{21} & u_{22} + iu_{23} \\ \cdot & \cdot & u_{2} & u_{24} + iu_{25} & u_{26} + iu_{27} & u_{28} + iu_{29} \\ \cdot & \cdot & \cdot & u_{3} & u_{30} + iu_{31} & u_{32} + iu_{33} \\ \cdot & \cdot & \cdot & \cdot & u_{4} & u_{34} + iu_{35} \\ \cdot & \cdot & \cdot & \cdot & \cdot & u_{5} \end{pmatrix} .$ (2)

QUDA (column-major):

$$\begin{pmatrix} u_{0} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{6} + iu_{7} & u_{1} & \cdot & \cdot & \cdot & \cdot \\ u_{8} + iu_{9} & u_{16} + iu_{17} & u_{2} & \cdot & \cdot & \cdot \\ u_{10} + iu_{11} & u_{18} + iu_{19} & u_{24} + iu_{25} & u_{3} & \cdot & \cdot \\ u_{12} + iu_{13} & u_{20} + iu_{21} & u_{26} + iu_{27} & u_{30} + iu_{31} & u_{4} & \cdot \\ u_{14} + iu_{15} & u_{22} + iu_{23} & u_{28} + iu_{29} & u_{32} + iu_{33} & u_{34} + iu_{35} & u_{5} \end{pmatrix}$$

$$(3)$$

STATUS



QCD+QED Wilson-Clover

STATUS

- InterfaceC* 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:
 - input file parsing
 - solver setup
 - call solver

1 [Solver 0] 2 solver CGNE 3 nmx 256 4 res 1.0e-12	1 [Solver 1] 2 solver SAP_GCR 3 nkv 16 4 isolv 1 5 nmr 4 6 ncy 5	1 [Solver 2] 2 solver DFL_SAP_GCR 3 idfl 0 4 nkv 16 5 isolv 1 6 nmr 4
	7 NMX 24	7 ncy 5
	8 res 1.0e-8	8 nmx 24

Figure: Example solver sections

1.0<u>e-8</u>

ADDITIONAL SOLVER TYPE

Add solver type QUDA

All options from QudaInvertParam and QudaMultigridParam

2 solver QUDA 3 gcrNkrylov 16 4 tol 1e-12 5 inv_type QUDA_GCR_INVERTER 6 inv_type_precondition QUDA_MG_INVERTER 7		[Solver 3]	
3 gcrNkrylov 16 4 tol 1e-12 5 inv_type QUDA_GCR_INVERTER 6 inv_type_precondition QUDA_MG_INVERTER 7		solver	QUDA
4 tol 1e-12 5 inv_type QUDA_GCR_INVERTER 6 inv_type_precondition QUDA_MG_INVERTER 7 ···		gcrNkrylov	16
5 inv_type QUDA_GCR_INVERTER 6 inv_type_precondition QUDA_MG_INVERTER 7 ···		tol	10-12
<pre>6 inv_type_precondition QUDA_MG_INVERTER 7 ···</pre>		inv_type	QUDA_GCR_INVERTER
		inv_type_precondition	QUDA_MG_INVERTER
9 [Solver 3 Multigrid]		[Solver 3 Multigrid]	
10 n_level 2	10	n_level 2	
11	11		
12	12		
13 [Solver 3 Multigrid Level ⊙]	13	[Solver 3 Multigrid Le	evel o]
15	15		
16 [Solver 3 Multigrid Level 1]		[Solver 3 Multigrid Le	evel 1]
17 • • •	17		

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, Niklausschreiber2, CC BY-SA 3.0

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

18

INVERTER SCALING



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

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

19



CONCLUSIONS

- Up and running interface to QUDA
- C* boundaries in QUDA
- O QCD+QED Wilson-Clover in QUDA
- Offloaded Dirac solves and eigensolver
- O Contractions
- **O** 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. 184, 519–528 (2013), arXiv:1206.2809 [hep-lat].
- [3] A. S. KRONFELD AND U. J. WIESE, "SU(N) GAUGE THEORIES WITH C-PERIODIC BOUNDARY CONDITIONS (I). TOPOLOGICAL STRUCTURE", Nuclear Physics B 357, 521–533 (1991).
- [4] M. LÜSCHER, "LOCAL COHERENCE AND DEFLATION OF THE LOW QUARK MODES IN LATTICE QCD", Journal of High Energy Physics 2007, 081 (2007), eprint: 0706.2298,
- [5] I. CAMPOS ET AL., "OPENQ*D CODE: A VERSATILE TOOL FOR QCD+QED SIMULATIONS", The European Physical Journal C 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), arXiv:1005.3043 [hep-lat].
- [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. 292, 108869 (2023), arXiv:2205.09104 [math.NA].

[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 Communications 181, 1517–1528 (2010), eprint: 0911.3191.

APPENDIX

OPENQXD: SPACETIME ORDERING I

- txyz-convention, i.e. 4-vector $x = (x_0, x_1, x_2, x_3)$
- **Lexicographical** index (L_{μ} = 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_μ
 - Within a block: $\Lambda(b, B)$, where b = block-local Euclidean 4-vector
 - ► Block themselves: $\Lambda(n, N_B)$, where $N_{B,\mu} = L_{\mu}/B_{\mu}$ and $n_{\mu} = \lfloor x_{\mu}/B_{\mu} \rfloor$
- Even-odd ordering in the block (but not the blocks themselves)

$$\hat{x} = \left\lfloor \frac{1}{2} \left(V_B \Lambda(n, N_B) + \Lambda(b, B) \right) \right\rfloor + P(x) \frac{V}{2},$$
(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 $V = L_3 L_2 L_1 L_0$.

OPENQXD: SPACETIME ORDERING II

41	11	45	15	57	27	61	31
9	43	13	47	25	59	29	63
40	10	44	14	56	26	60	30
8	42	12	46	24	58	28	62
33	3	37	7	49	19	53	23
1	35	5	39	17	51	21	55
32	2	36	6	48	18	52	22
0	34	4	38	16	50	20	54
	\rightarrow t						

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)

$$D_{w}\psi(x) = (4 + m_{o})\psi(x) -\frac{1}{2}\sum_{\mu=0}^{3} \left\{ H_{\mu}(x)(1 - \gamma_{\mu})\psi(x + \hat{\mu}) + H_{\mu}(x - \hat{\mu})^{-1}(1 + \gamma_{\mu})\psi(x - \hat{\mu}) \right\} + c_{sw}^{SU(3)}\frac{i}{4}\sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi(x) + qc_{sw}^{U(1)}\frac{i}{4}\sum_{\mu,\nu=0}^{3} \sigma_{\mu\nu}\hat{A}_{\mu\nu}\psi(x) ,$$
(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

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

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

We add the U(1) SW-term,

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

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

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

C^{\star} boundary conditions

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

$$\begin{aligned} A_{\mu}(\mathbf{x} + L_{k}\hat{k}) &= -A_{\mu}, \\ \psi_{f}(\mathbf{x} + L_{k}\hat{k}) &= C^{-1}\overline{\psi}_{f}^{\mathsf{T}}(\mathbf{x}), \\ \overline{\psi}_{f}(\mathbf{x} + L_{k}\hat{k}) &= -\psi_{f}^{\mathsf{T}}(\mathbf{x})\mathcal{C}, \\ U_{\mu}(\mathbf{x} + L_{k}\hat{k}) &= U^{*}\mu(\mathbf{x}), \end{aligned}$$
(8)

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

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

The gauge action is

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

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

where the bare coupling constants are g_0 , e_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:

$$\left[\mathrm{d}\psi\right]_{\Lambda_{phys}}\left[\mathrm{d}\bar{\psi}\right]_{\Lambda_{phys}} = \prod_{\mathbf{x}\in\Lambda_{phys}}\mathrm{d}\psi(\mathbf{x})\bar{\psi}(\mathbf{x}) = \prod_{\mathbf{x}\in\Lambda_{exended}}\mathrm{d}\psi(\mathbf{x}) = \left[\mathrm{d}\psi\right]_{\Lambda_{extended}}$$
(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
- Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \rightarrow we'll probably never reach that
- \blacksquare Intermediate phase: some functions are ported to GPU, but not all of them \rightarrow needs transfers

- \blacksquare Initial code: all functions implemented in CPU \rightarrow no transfers needed
- \blacksquare Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \longrightarrow we'll probably never reach that
- \blacksquare 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!

- \blacksquare Initial code: all functions implemented in CPU \rightarrow no transfers needed
- \blacksquare Ideal final code: all functions implemented in GPU \rightarrow no transfers needed \longrightarrow we'll probably never reach that
- \blacksquare 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 *s) { ... }
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, ...
- 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!) ⇒ 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