JÜLICH
FORSCHUNGSZENTRUM

Project Talk: Optimizing the ChASE eigensolver for Bethe-Salpeter computations

JLESC NCSA 2017 | **Jan Winkelmann**, E. Di Napoli, A. Schleife

Outline

Topic

Motivation and Problem Statement

ChASE: The “New” Eigensolver

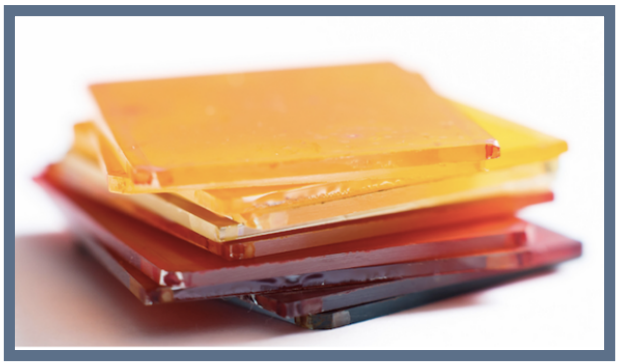
Integrating BSE and ChASE
- The BSE Software
- Double vs Single Precision
- File I/O

Outlook and Conclusion

Theoretical Spectroscopy: Oxides, Perovskites

Optoelectronics and semiconductor technology:

- Lasers and light-emitting diodes
- Transparent electronics
- Tunable optical properties



$CH_3NH_3PbI_3$

Energy-related applications:

- Photocatalytic water splitting
- Photovoltaic absorbers
- Transparent electrodes for PV

Excitonic effects: solution of the Bethe-Salpeter equation

- Leads to eigenvalue problem (excitonic Hamiltonian)
- Huge matrix: Rank typically > 50,000
- Time-propagation approach to calculate the dielectric function

- Excellent description of the optical properties of the oxides

→ predictive power (e.g. for perovskites)

Andre Schleife @ MatSE @ UIUC • Email: schleife@illinois.edu • Web: http://schleife.matse.illinois.edu • @aschleife

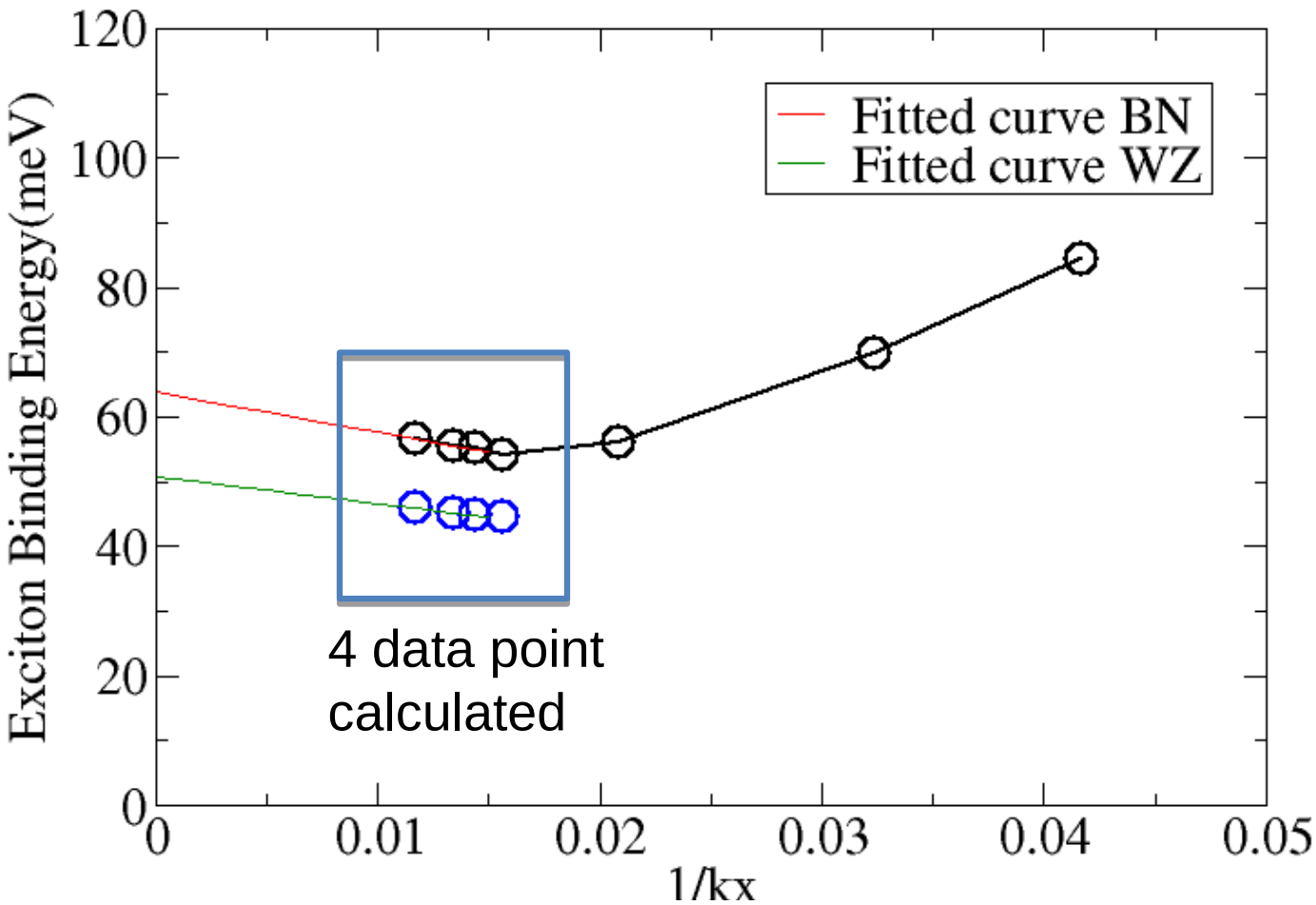
Computational Problem

Extremal Hermitian Eigenvalue Problem

- Obtain smallest $\sim$100 eigenpairs of
- $H^\dagger = H \in \mathbb{C}^{n \times n}$, dense
- $20,000 \leq n \leq 200,000$ (hopefully soon $1,000,000$)
- Single Precision Complex

- Sequence of problems with increasing size and similar eigenvalues
- Weak scaling of particular interest
- Example: Zinc Oxide BN

Exciton binding energies: BN ZnO



- Standard approach: Linear extrapolation after “turnaround point”
- Larger energy cutoff desirable (but unaffordable)
- More complicated oxides desirable, e.g. for materials design

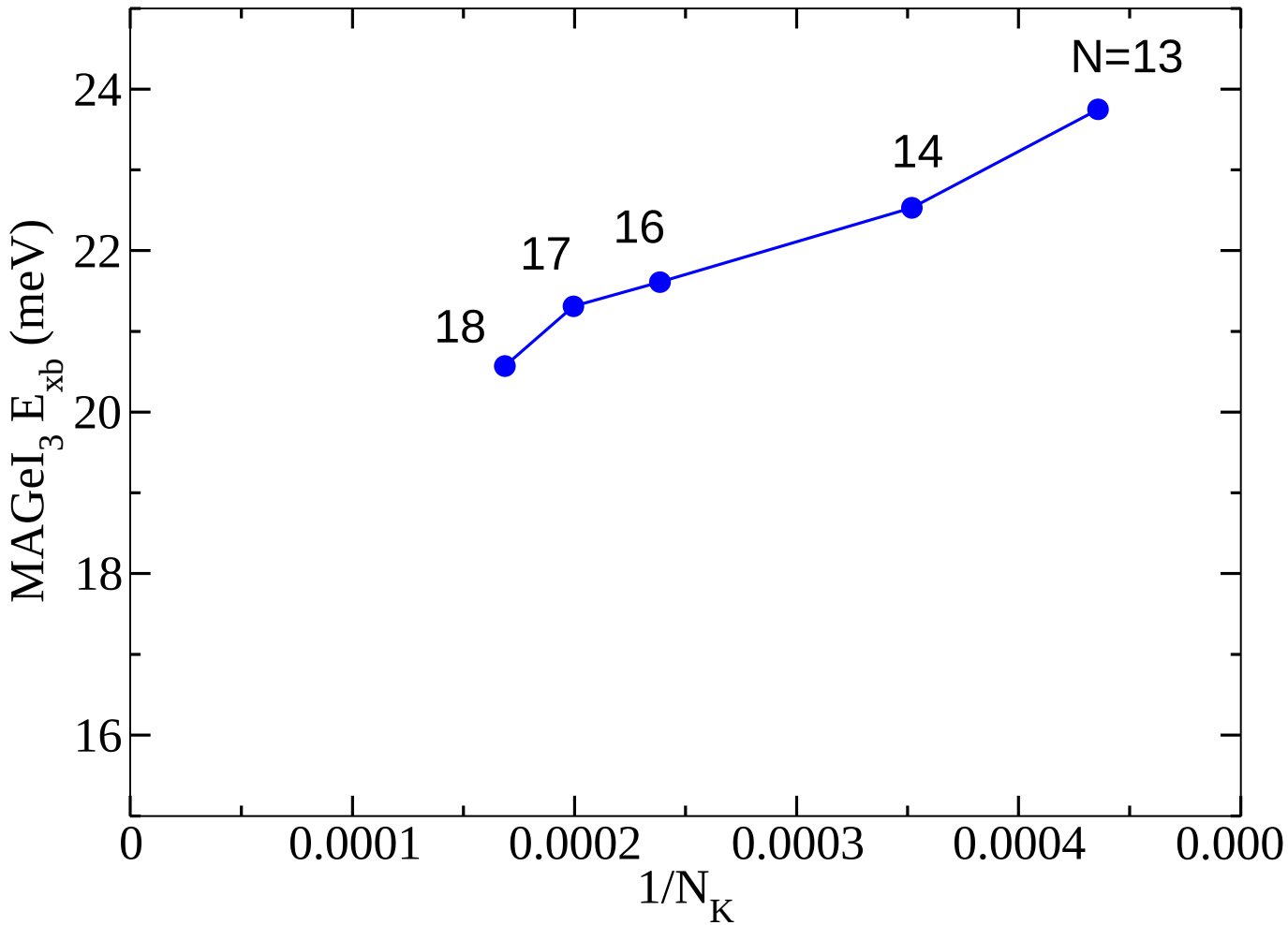
Exciton binding energies: BN ZnO

- Layered meta-stable BN structure of ZnO reported for ultra-thin films
- Computation of (converged) binding energies computationally expensive
- Particularly difficult: Convergence with respect to ***k***-point sampling

# k-points	Rank of matrix	Total # of matrix elements	Time consumed	Memory required	Nodes used
10945	82499	6.8×10^{9}	1.5 hours	50.7 GB	8
12713	96399	9.3×10^{9}	2 hours	69.2 GB	8
16299	124281	1.5×10^{10}	2 hours	115.1 GB	16
25367	195281	3.8×10^{10}	5.5 hours	284.1 GB	16

- Cost increases enormously as k-points number increases
- Convergence barely achieved, despite very simple system (4 atoms per unit cell)

Exciton binding energies: $MAGeI_3$



- Turnaround-point not reached yet (higher k-point density needed)
- Explore cutoff dependence more in detail
- We hope that ChASE will enable us to do these simulations

Exciton binding energies: $MAGeI_3$

- $MAGeI_3$ is a hybdrid perovskite: methyl-ammonia germanium iodide
- Here used as a test system (simpler than $MAPbI_3$)
- Unit cell: 12 atoms (cubic phase)
- Hybrid k-point grids used: 5x5x5 outer grid, 2x2x2 inner boxes sampled using NxNxN points

N	Num k-points	BSE Matrix Dim.	Eigenval. Cal. Time
13	2295	30898	0.150 hrs.
14	2842	38823	0.235 hrs.
16	4194	58549	0.550 hrs.
17	5011	70570	0.780 hrs.
18	5930	84043	1.147 hrs

- Wave function size additional limiting factor (addressing this currently)
- K-point sampling challenging

Topic

Motivation and Problem Statement

ChASE: The “New” Eigensolver

Integrating BSE and ChASE
- The BSE Software
- Double vs Single Precision
- File I/O

Outlook and Conclusion

ChASE

Subspace Iteration with Rayleigh-Ritz projection

- Choose an initial system of vectors $X^0 = [x_1, \ldots, x_m]$.
- Perform successive multiplication $X^k := AX^{k-1}$.
- Every once in a while orthonormalize column-vectors in X^k.
- Compute Rayleigh quotient
- Solve reduced problem

ChASE Eigensolver

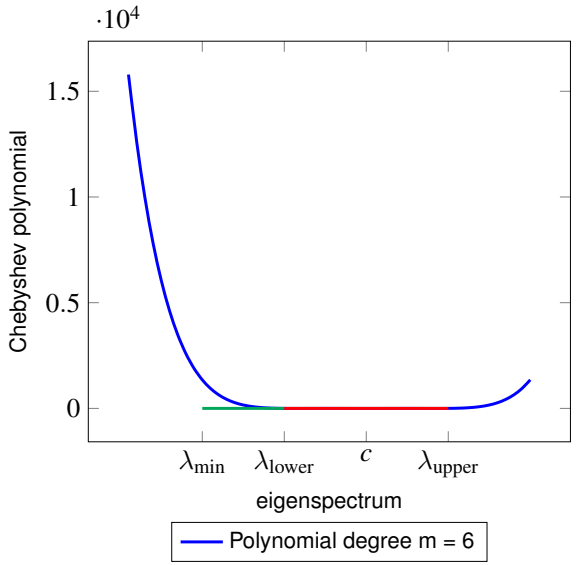
- Substitute $A^k X \longrightarrow p(A)X$.
- Chebyshev filter improves the rate of convergence.

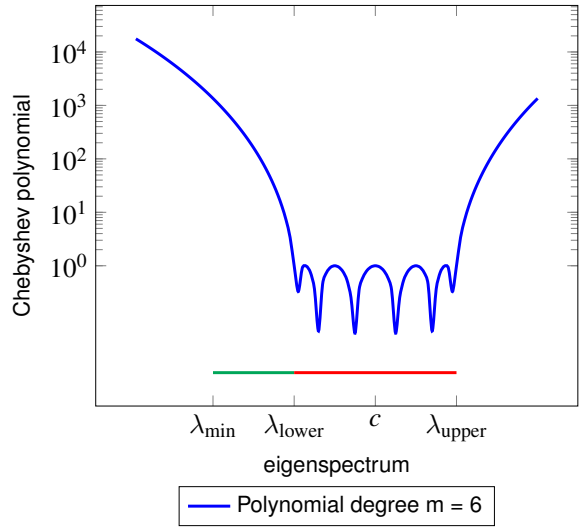
The core of the algorithm: Chebyshev filter

Chebyshev polynomials

A generic vector $v = \sum_{i=1}^{n} s_i x_i$ is very quickly aligned in the direction of the eigenvector corresponding to the extremal eigenvalue λ_1

$$
\begin{aligned}
v^m = p_m(H)v &= \sum_{i=1}^{n} s_i\, p_m(H) x_i = \sum_{i=1}^{n} s_i\, p_m(\lambda_i) x_i \\
&= s_1 x_1 + \sum_{i=2}^{n} s_i \frac{C_m\left(\frac{\lambda_i - c}{e}\right)}{C_m\left(\frac{\lambda_1 - c}{e}\right)} x_i \sim \boxed{s_1 x_1}
\end{aligned}
$$





The core of the algorithm: Chebyshev filter

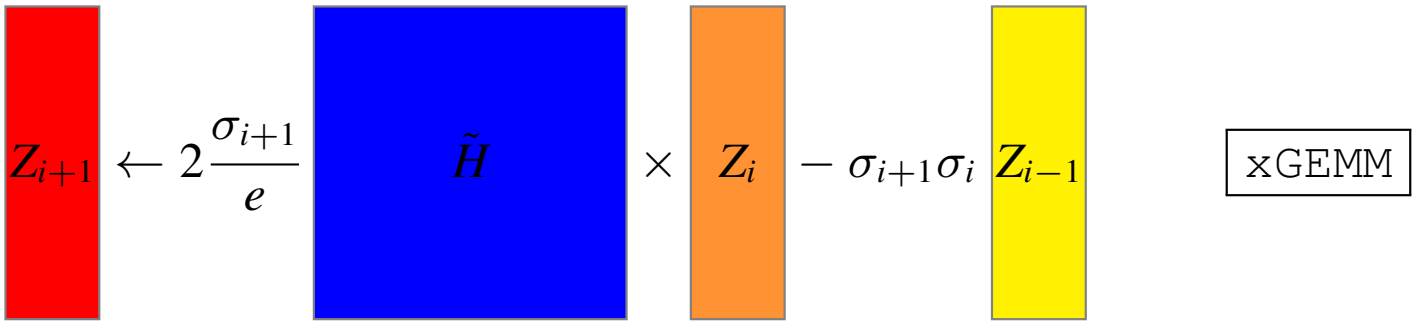
In practice

Three-terms recurrence relation

$$C_{m+1}(t) = 2xC_m(t) - C_{m-1}(t); \qquad m \in \mathbb{N}, \quad C_0(t) = 1, \quad C_1(t) = x$$

$$Z_m \doteq p_m(\tilde{H})\, Z_0 \quad \text{with} \quad \tilde{H} = H - cI_n$$

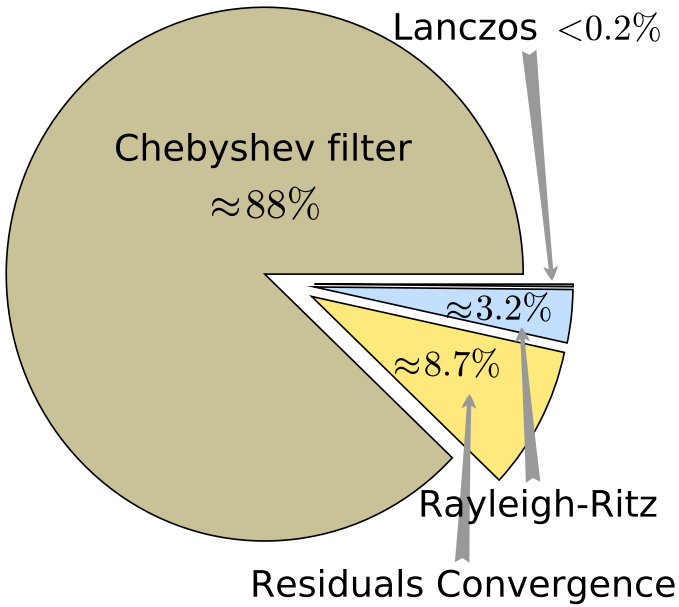
FOR: $i = 1 \rightarrow \text{DEG} - 1$



END FOR.

Parallelization Approach

$Au_{98}Ag_{10}$ - $n = 8{,}970$ - 32 cores.



- `xGEMM` most expensive part
- Parallelizes easily over
  - MPI
  - GPUs
- Good weak scaling
- Recall:
  Matrix dimensions skewed

Topic

Motivation and Problem Statement

ChASE: The “New” Eigensolver

Integrating BSE and ChASE
- The BSE Software
- Double vs Single Precision
- File I/O

Outlook and Conclusion

BSE: Solving the Bethe-Salpeter Hamiltonian(s)

# k-points	Rank of matrix	Total # of matrix elements	Time consumed	Memory required	Nodes used
10945	82499	6.8×10^{9}	1.5 hours	50.7 GB	8
12713	96399	9.3×10^{9}	2 hours	69.2 GB	8
16299	124281	1.5×10^{10}	2 hours	115.1 GB	16
25367	195281	3.8×10^{10}	5.5 hours	284.1 GB	16

- Two separate parts:
  1. Generate and write largest considered Hamiltonian
  2. Read Hamiltonian for given k-points and “solve”
- “Legacy” Software

The question of precision

- SLQM provides solvers *and* necessary experience
- Accuracy requirement: 6 digits for eigenvalues
- Can we do this in single precision?

The question of precision

- SLQM provides solvers *and* necessary experience
- Accuracy requirement: 6 digits for eigenvalues
- Can we do this in single precision?
- Yes! (not obvious)

  Recall: solvers check for $||Z^{\dagger}HZ - \Lambda||$

The question of precision

- SLQM provides solvers *and* necessary experience
- Accuracy requirement: 6 digits for eigenvalues
- Can we do this in single precision?
- Yes! (not obvious)

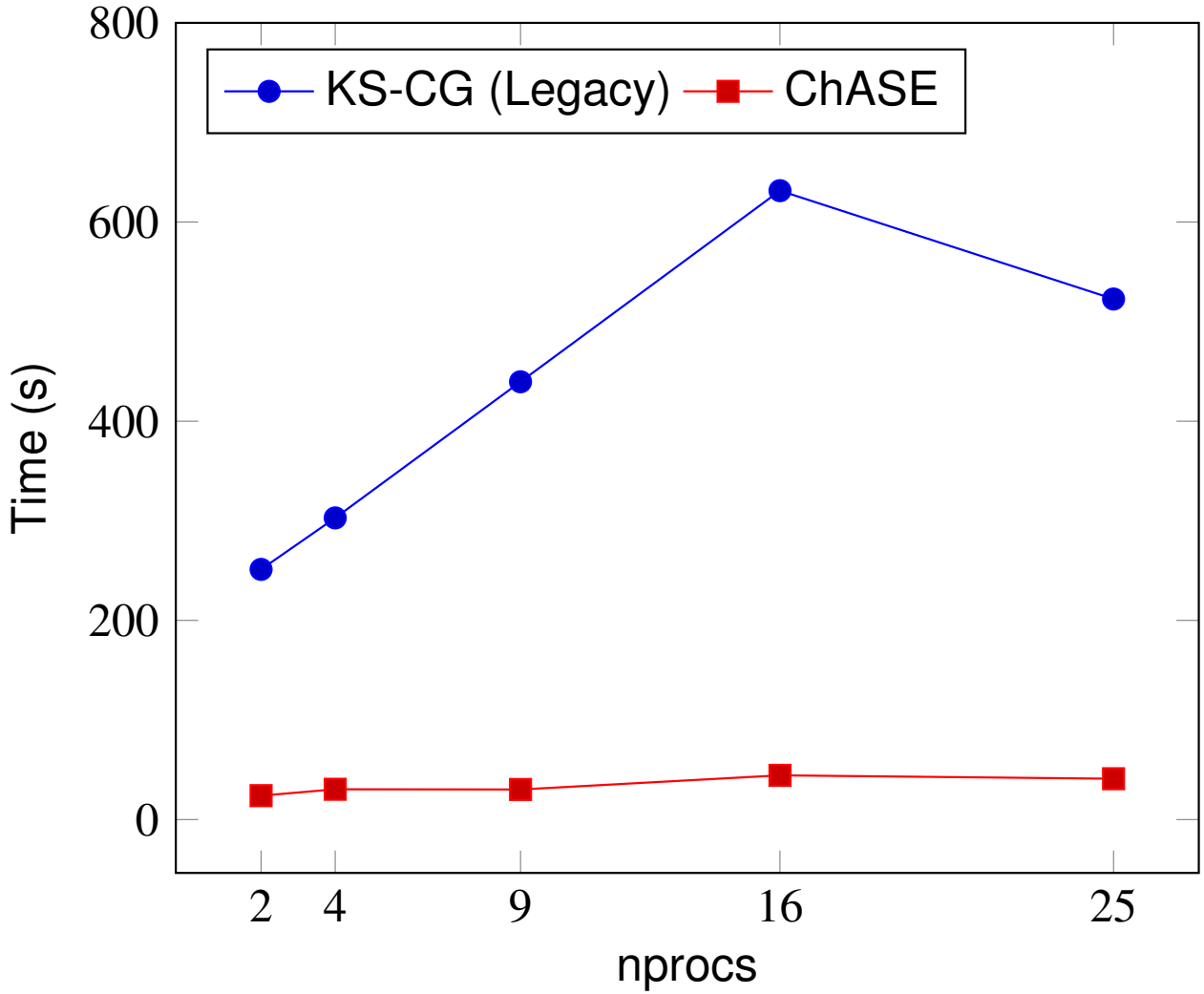
  Recall: solvers check for $||Z^{\dagger}HZ - \Lambda||$
- No! (also not obvious)

Single Precision cGEMM

CPU vs GPU

- In some cases ChASE does not converge to desired tolerance
- However, only when cuBLAS is used
- MKL works fine (for all cases tested so far)
- Initial investigation indicates lower accuracy of cuBLAS in some cases
- These are extreme cases, but warrant further investigation

Single Precision: Weak Scaling



CPU based `CGEMM` Obtained on JURECA Hamiltonian: Indium Oxide

File I/O

- Read the Hamiltonian for # k-points and solve for eigenpairs
- BSE distributes Hamiltonian in columns (for MPI)

$\Rightarrow$ Not optimal for ChASE `xGEMM`

- Stop-gap:

File I/O

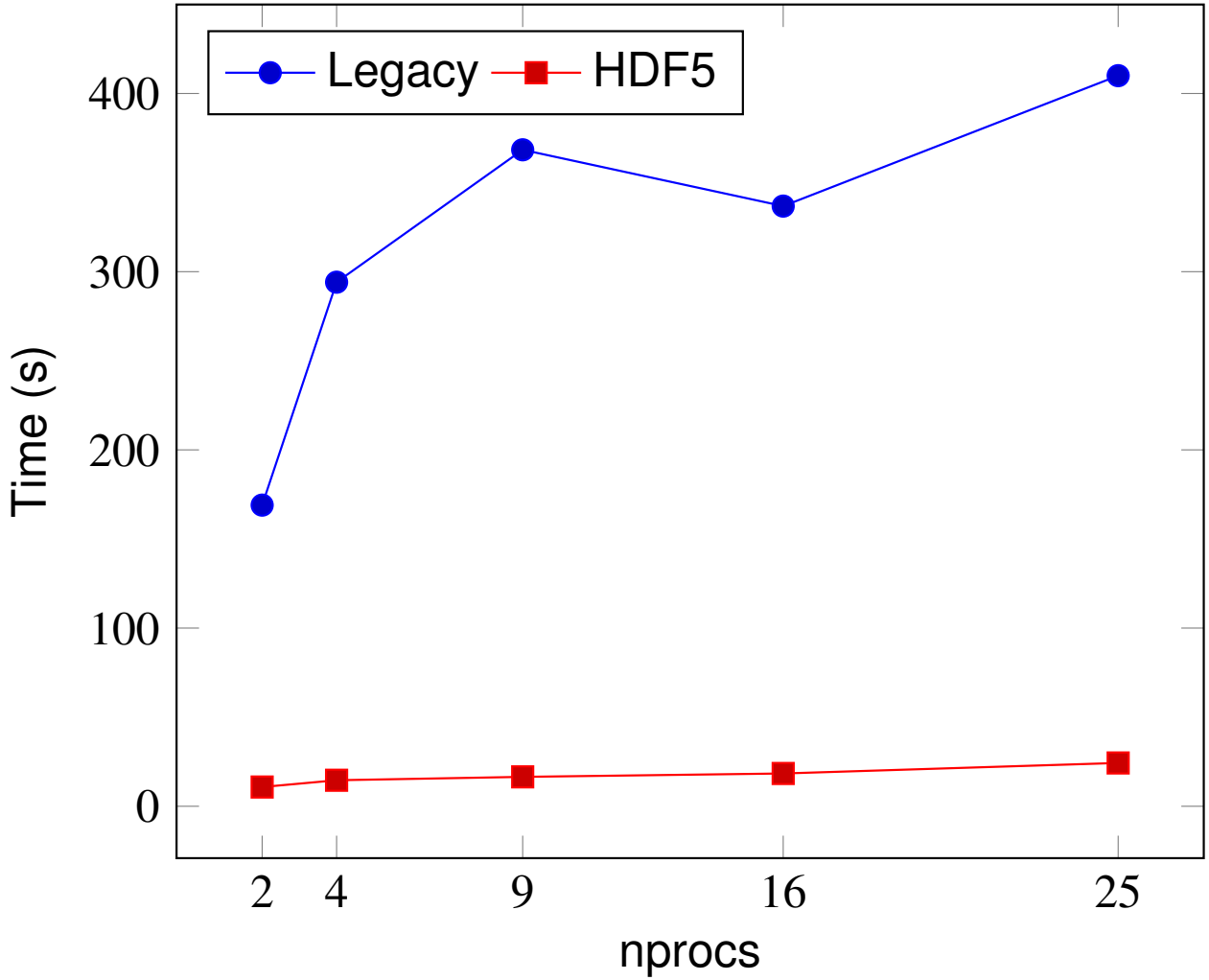
- Read the Hamiltonian for # k-points and solve for eigenpairs
- BSE distributes Hamiltonian in columns (for MPI)

$\Rightarrow$ Not optimal for ChASE `xGEMM`

- Stop-gap:

1. Read Hamiltonian
2. Write entire Hamiltonian to HDF5 file
3. Re-read with appropriate data layout
4. Then solve with ChASE

File I/O: Weak Scaling



CPU based `CGEMM` Obtained on JURECA Hamiltonian: Indium Oxide

Topic

Motivation and Problem Statement

ChASE: The “New” Eigensolver

Integrating BSE and ChASE
- The BSE Software
- Double vs Single Precision
- File I/O

Outlook and Conclusion

Project Status

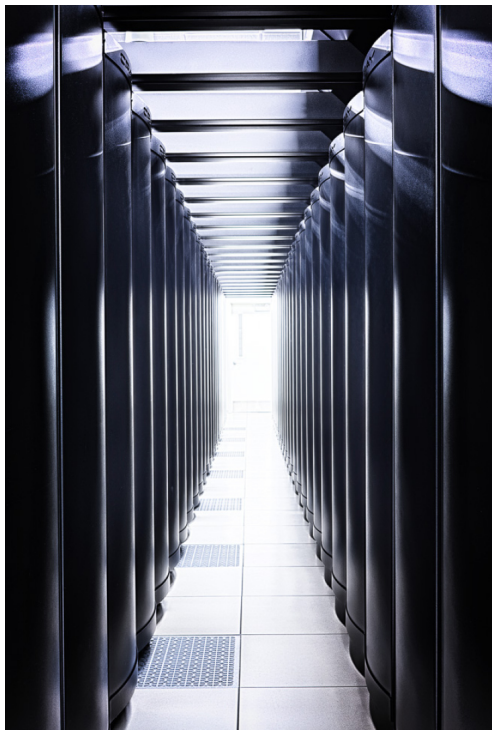
What happened since Kobe

- Project Meeting at NCSA
- Integration of BSE and ChASE
  - Accuracy and algorithmic hurdles
  - Single vs double precision
  - HDF5 for parallel I/O
- Promising results on JURECA

Future Work

- Porting ChASE (CPU and GPU) to Blue Waters
- Tuning ChASE for Blue Waters and the BSE application
- Large scale (weak) scaling on Blue Waters
- Accuracy of Single Precision GPU `CGEMM`

The Blue Waters Supercomputer



Specifications:

- Cray: 22640 XE6 and 4224 XK7 nodes
- 64 GB memory (XE6) and 32 GB (XK7)
- XE6: 2 AMD Interlagos chips (8 cores each) per node
- XE7: 1 AMD Interlagos and 1 Kepler accelerator

System performance:

- Peak performance: 13+ Petaflops
- 1.5 PB of total memory
- 25 PB online disk storage and 300+ PB near-line storage

Access and Services:

- Operated by NCSA (National Center for Supercomputing Applications) at UIUC
- Part of the XSEDE program of the National Science Foundation
- Visualization support

Acknowledgments

- **Xiao Zhang (Oxide semiconductors)**
- **Josh Leveillee (Hybrid perovskites)**

CBET-1437230
DMR-1555153

OCI-0725070
ACI-1238993

Acknowledgements (cont'd)

For more information

e.di.napoli@fz-juelich.de

fz-juelich.de/ias/jsc/slqm

schleife@illinois.edu

schleife.matse.illinois.edu

Special thanks to the JSC for providing travel funding.

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

Experiment Results

nprocs	N	Reading	KS-CG	HDF5	ChASE
2	22360	169.0	251.2	10.8	24.0
4	30477	294.1	302.9	14.6	30.4
9	44487	368.4	439.6	16.5	30.2
16	59725	336.8	631.5	18.4	44.4
25	73636	410.0	522.7	24.3	41.0