Performance benchmark of the latest EigenExa on Fugaku

Toshiyuki Imamura*, Takeshi Terao, Takuya Ina (RIKEN R-CCS), Yusuke Hirota (University of Fukui), Katsuhisa Ozaki, Yuki Uchino (Shibaura Institute of Technology)

1. Introduction

Eigensolver is one of the strongly demanded tools in modern simulations. RIKEN has been developing high-performance and reliable eigensolvers since 2012 for the K computer, 2020 for the supercomputer Fugaku. 'EigenExa' [3] refers to the heart of our eigensolver project and the signature software for large-scale and highly parallel computations. We have just released version 2.11 in 2021 as a standard library on Fugaku, which employs the novel one-stage scheme. Also, the traditional one-stage scheme (the green path in Figure 1) is available as most of the state-of-the-arts eigensolvers such as ScaLAPACK and ELPA.

2. EigenExa

The latest RC version of EigenExa version 2.12 supports (only FP64);

- Real-symmetric Standard eigenvalue problem
- Real-symmetric Generalized eigenvalue problem
- Hermitian Standard eigenvalue problem, and
- Hermitian Generalized eigenvalue problem.

Along with investigations for modern parallel systems, a band-matrix version of the one-stage scheme (the red path in Figure 1), communication avoiding algorithms, and kd-tree based load balancing are introduced.

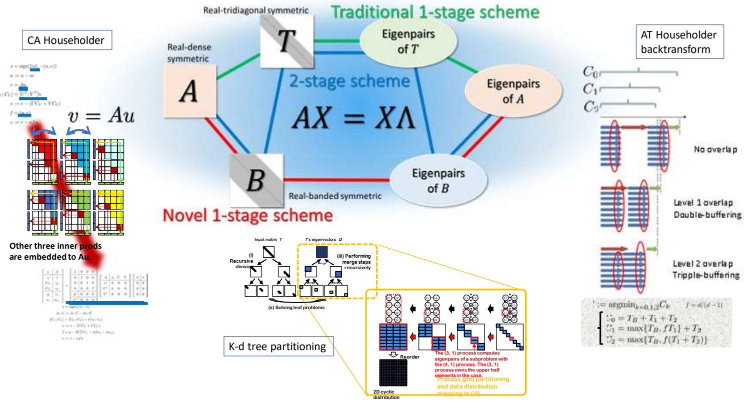


Figure 1: Schematics of EigenExa schemes and parallel implementations

☐ Reproducibility:

The difficulties to eliminate the nondeterminism of parallel computation and to ensure perfect reproducibility are common issues in the HPC field. EigenExa is coded elaborately to guarantee as much reproducible as possible, while ScaLAPACK is not permanently reproducible even for consecutive calls.

☐ Iterative Refinement approach:

Like most eigensolvers, EigenExa exhibits typical accuracy behavior, dominated by the condition number and the gap of the eigenvalues. The idea to refine the accuracy is another approach, and we are currently attempting to include an accuracy booster according to Ogita-Aishima's [2] scheme into version 2.12 or later.

☐ Verification:

Numerical verification is also investigated with a new approach using only floating-point arithmetic with nearest-point rounding [4]. The method features high-precision matrix multiplication to handle large matrices and to suppress the overestimation of the error upper bound even for ill-conditioned matrices.

- Algorithm of refinement for eigenvectors [2]
- Assuming that all approximate eingenvectors are given.

$$R \leftarrow I - \hat{X}^T\hat{X}$$
$$S \leftarrow \hat{X}^T A\hat{X}$$
$$\lambda_j \leftarrow s_{jj}/(1-r_{jj})$$
$$D' \leftarrow \mathrm{diag}(\lambda)$$
$$\delta \leftarrow 2(\|S-D'\| - \|A\|\|R\|)$$
$$e_{ij} \leftarrow \begin{cases} \dfrac{s_{ij}+\lambda_j r_{ij}}{\lambda_j-\lambda_i} & \text{if } |\lambda_j-\lambda_i|>\delta \\ r_{ij}/2 & \text{otherwise} \end{cases}$$
$$X' \leftarrow \hat{X} + \hat{X}E$$

- Algorithm of verification for all eigenvalues for standard eigenvalue problem [4]
- Assuming that all approximate eigenpairs are given.
- This scheme produces the upper bound of errors for all the approximate eigenvalues.

$$\mathbf{R} \leftarrow I - \hat{X}^T\hat{X}, \quad I - \hat{X}^T\hat{X} \in \mathbf{R}$$
$$\mathbf{T} \leftarrow A\hat{X} - \hat{X}\hat{D}, \quad A\hat{X} - \hat{X}\hat{D} \in \mathbf{T}$$

Culculate v_1 and v_2 such that

$$v_1 \geq |R|\cdot e \quad \text{and} \quad v_2 \geq |T^T|(|T|\cdot e), \quad e=(1,\dots,1)^T,$$

for $\forall R \in \mathbf{R}$ and $\forall T \in \mathbf{T}$.

$$\epsilon \leftarrow (1+4u)\sqrt{\frac{\max v_2}{1-\max v_1}}, \quad u \text{ is the unit roundoff}$$

Figure 2: Iterative Refinement (left) and verification (right) schemes

3. Benchmark on Fugaku

The preliminary performance comparison among P{DSY|ZHE}EVD of ScaLAPACK and eigen_{FS|h} of EigenExa in Figure 3, which perform the real-symmetric/Hermitian standard eigenproblem for all spectrum. The benchmark was done on Fugaku housed at RIKEN using up to 16,384 nodes (2.2GHz boost mode).

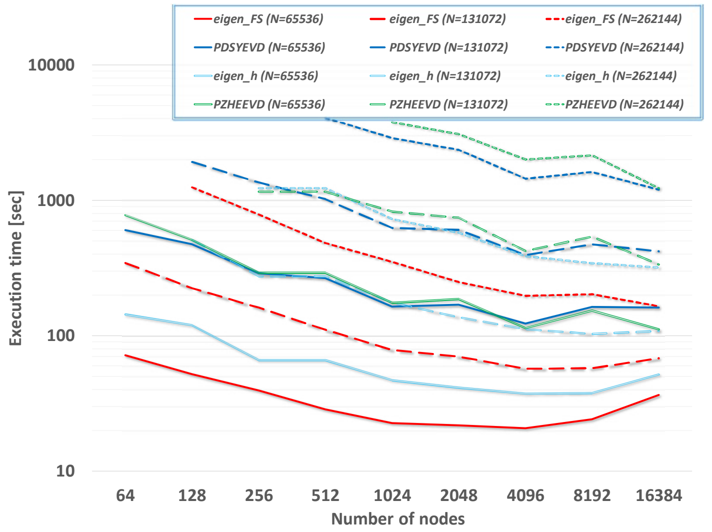
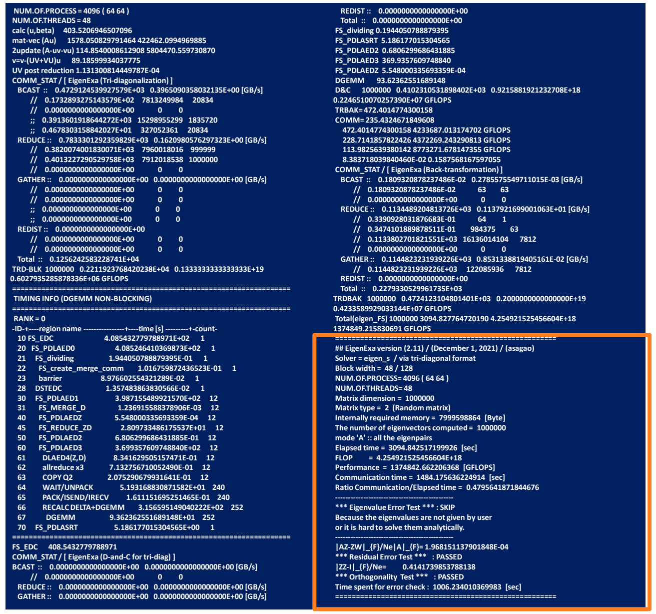


Figure 3: Strong scaling benchmark on Fugaku

Observations reveal greater parallel scaling of EigenExa, with a 3-8x performance improvement over ScaLAPACK in the real symmetric case. The performance profile of Hermitian on EigenExa exhibits a similar trend (1.5-4x improvement) to the real symmetric case. However, PDSYEVD and PZHEEVD show almost equivalent or upset performance, which is rather unnatural in terms of 2x memory traffic and 2x computational intensity between real and complex, expected 2x or much faster performance under normal circumstances, thus, to be clarified in the future.

For N=10^6, a remarkable benchmark of 4,096 nodes (theoretical peak 16.7PFLOPS) has recorded 1.37 PFLOPS in 3,100 seconds, which is comparable or superior to those ever done on the K computer and a Fujitsu FX10 [1].



JOBID: 9751855 (the benchmark shot from Sun Dec 5 13:00:17 JST 2021 to Sun Dec 5 17:00:33 JST 2021)

4. Summary

EigenExa, a high performance eigensolver on Fugaku, performs highly parallel in the range of N=10^4 to 10^6. On the considerably huge benchmark with N=10^6, the sustained performance exceeded one PFLOPS. The latest version 2.11 can handle Hermitian matrices, which is expected to broaden the field of applications. The next version 2.12 will be the first massively parallel eigenvalue library in the world to provide an accuracy booster that supports accuracy verification and accuracy improvement without using interval arithmetic, consequently, contributing to a reliable computing and real-world applications related to Society 5.0

5. Acknowledgements

The research was supported by the Fugaku trial phase program and used computational resources of the supercomputer Fugaku provided by RIKEN R-CCS (ra000006, hp200263, and hp200313).

References

[1] H. Imachi and T. Hoshi. 2016. Hybrid Numerical Solvers for Massively Parallel Eigenvalue Computations and Their Benchmark with Electronic Structure Calculations. Journal of Information Processing 24, 1 (2016), 164–172.

[2] T. Ogita and K. Aishima. 2018. Iterative Refinement for Symmetric Eigenvalue Decomposition. Japan Journal of Industrial and Applied Mathematics 35, 3 (2018), 1007–1035.

[3] RIKEN R-CCS 2021. EigenExa homepage. https://www.r-ccs.riken.jp/labs/lpnctrt/projects/eigenexa/.

[4] T. Terao, K. Ozaki, T. Ogita, and T. Imamura. 2021. Verified Numerical Computations for a Standard Eigenvalue Problem without Directed Rounding. The 40th JSST Annual International Conference on Simulation Technology.

*Contact: imamura.toshiyuki@riken.jp

https://www.r-ccs.riken.jp/labs/lpnctrt/projects/eigenexa/