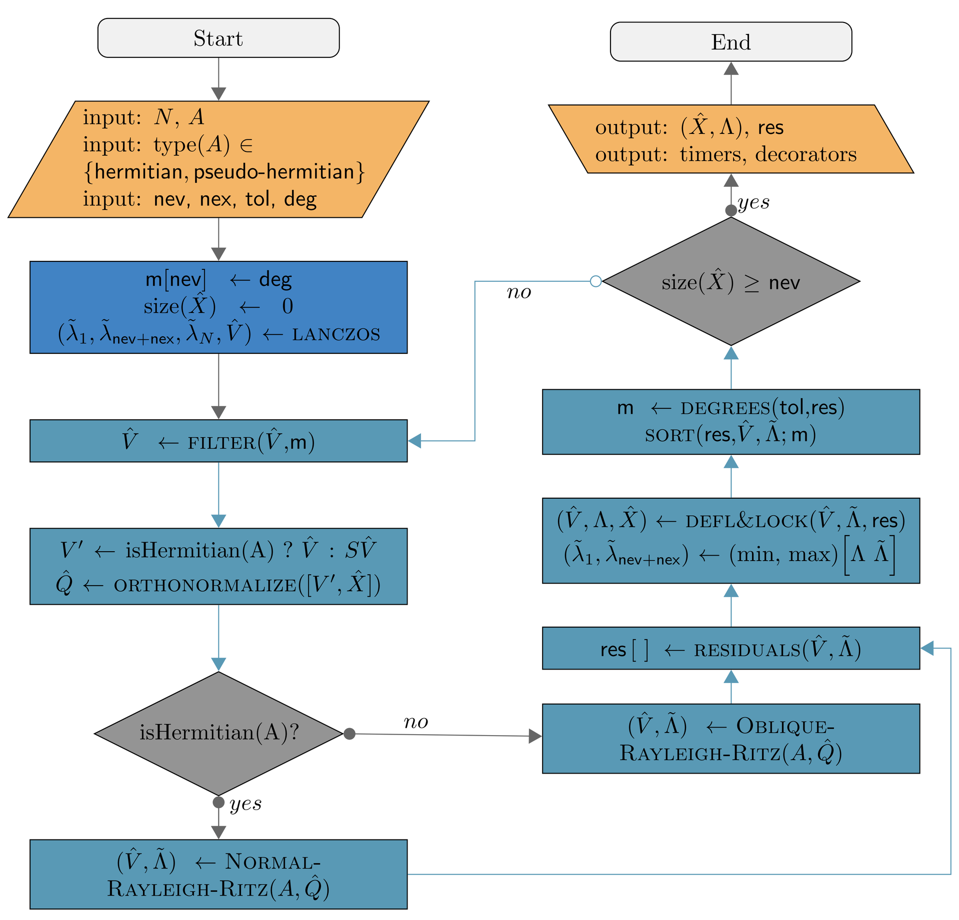
Chebyshev Accelerated Subspsace Eigensolvers for Pseudo-hermitian Hamiltonians

Clément Richefort (c.richefort@fz-juelich.de), Xinzhe Wu, Edoardo Di Napoli
Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany

ChASE Algorithm



ChASE — CHEBYSHEV ACCELERATED SUBSPACE EIGENSOLVER

- GitHub: `https://github.com/ChASE-library/ChASE`
- License: Open source – BSD 3-Clause
- Latest release: v1.7.0-rc2 – January, 2026
- Docs: `chase-library.github.io/ChASE/quick-start.html`
- Easy-to-integrate: ready-to-use C++ to Fortran interface
- accessible via ELSI (Electronic Structure Infrastructure)

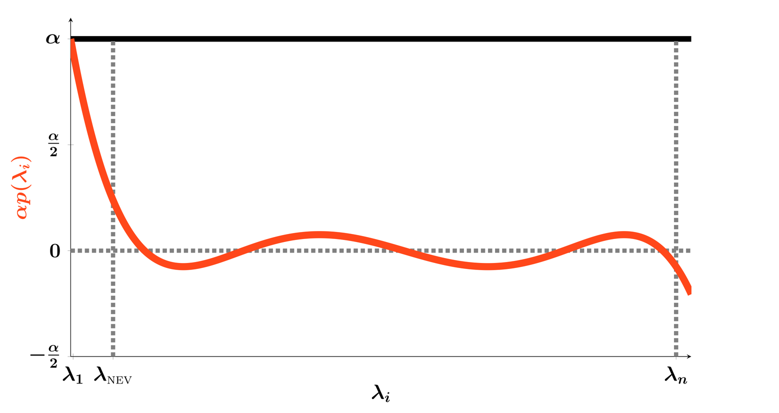


Figure: Illustration of the filter on an uniformly distributed random vector

Problem Definition

The Hamiltonian H derived by the BSE and by the DFT is of the form

$$H := \begin{bmatrix} A & B \\ -\bar{B} & -\bar{A} \end{bmatrix} \quad \text{with} \quad A = A^* \quad \text{and} \quad B = B^T.$$

The matrix H is said "pseudo-hermitian" as it satisfies

$$SH = H^*S \quad \text{with} \quad S := \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}.$$

This poster focuses on the most common case in which the eigenproblem

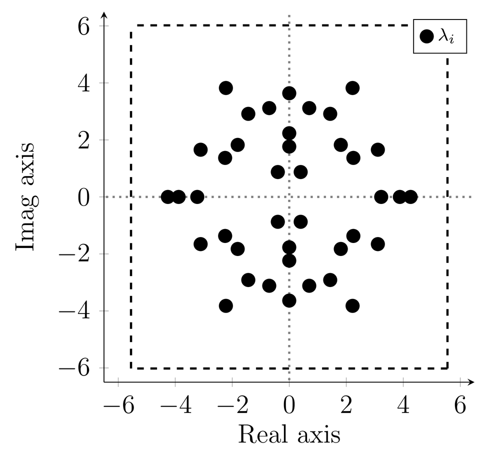
$$H\boldsymbol{v} = \lambda\boldsymbol{v} \quad \Leftrightarrow \quad SH\boldsymbol{v} = \lambda S\boldsymbol{v} \quad \text{with} \quad \boldsymbol{v} \in \mathbb{C}^n \quad \text{and} \quad \lambda \in \mathbb{R}$$

is "definite", meaning that SH is hermitian positive definite (HPD), i.e.,

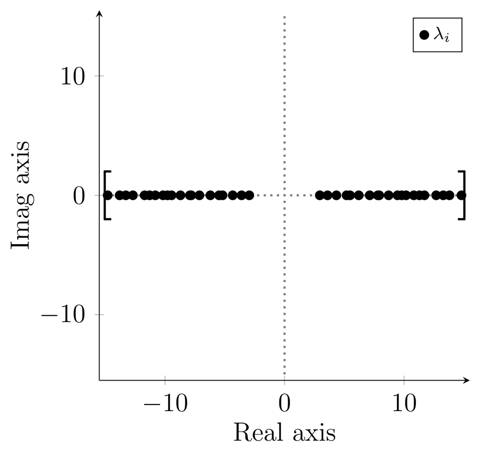
$$SH = \begin{bmatrix} A & B \\ \bar{B} & \bar{A} \end{bmatrix} \succ 0.$$

Summary of the key properties :

- The spectrum is real, i.e.,
$$\lambda(H) \in \mathbb{R}.$$
- Eigenvalues appear in positive-negative pairs, such that
$$H\boldsymbol{v} = \lambda\boldsymbol{v} \quad \Leftrightarrow \quad HK\bar{v} = -\bar{\lambda}K\bar{\boldsymbol{v}} \quad \text{with} \quad K := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$
- The left and right eigenvectors are related such that
$$U = SV.$$
- The set of right eigenvectors is not orthonormal, but satisfies
$$V^*SV = D \quad \text{with } D \text{ a non-singular real diagonal matrix.}$$



(a) $\lambda_i(H)$ - SH is indefinite



(b) $\lambda_i(H)$ - SH is HPD

Major Algorithmic Innovations

Rayleigh-Ritz for pseudo-hermitian Hamiltonian

Orthonormal vs. Oblique Rayleigh-Ritz

Hermitian case :

- Orthonormal Projection :
$$\tilde{\boldsymbol{y}}^*\left(A\tilde{\boldsymbol{x}} - \tilde{\omega}\tilde{\boldsymbol{x}}\right) = 0, \quad \forall\tilde{\boldsymbol{y}} \in \text{range}(Q).$$
- Quadratic Convergence :
$$|\omega - \tilde{\omega}| \leq |\sigma_{\boldsymbol{v}}|^2 \|A - \omega I\|_2,$$
- Single basis :
$$\lim_{Q \to V} Q^*AQ = \Omega.$$

Non-hermitian case :

- Oblique Projection :
$$\tilde{\boldsymbol{u}}^*\left(H\tilde{\boldsymbol{v}} - \tilde{\lambda}\tilde{\boldsymbol{v}}\right) = 0, \quad \text{with} \quad \tilde{\boldsymbol{v}} \in \text{range}(Q) \text{ and } \forall\tilde{\boldsymbol{u}} \in \text{range}(Q_L),$$
- Non-immediate Quadratic Convergence :
$$\left|\lambda - \tilde{\lambda}\right| \leq \left|\tilde{\delta}\right|^{-1} \cdot \left|\bar{\sigma}_{\boldsymbol{u}}\sigma_{\boldsymbol{v}}\right| \cdot \left\|H - \lambda I\right\|_2, \quad \tilde{\delta} := \tilde{\boldsymbol{u}}^*\tilde{\boldsymbol{v}}$$
- Dual basis, such that $Q_L^*Q = I$:
$$\lim_{Q \to V,\, Q_L \to U} Q_L^*HQ = \Lambda.$$

General Form of Dual Basis

Let M be a non-singular matrix. A dual basis Q_L is defined as

$$Q_L := [SQ - Q(Q^*SQ - M)]\, M^{-1}.$$

Since Q is orthonormal ($Q^*Q = I$), we have

$$\begin{aligned} Q_{\mathcal{L}}^*Q &= \left([SQ - Q(Q^*SQ - M)]M^{-1}\right)^*Q \\ &= (M^{-1})^*(Q^*S^*Q - Q^*S^*Q + M^*) \\ &= (M^{-1})^*M^* = I. \end{aligned}$$

The Rayleigh-Quotient $G := Q_L^*HQ$ is not hermitian.

Solving a spectrally equivalent hermitian eigenvalue problem

Setting $M := Q^*SQ$ leads to

$$Q_L := SQ(Q^*SQ)^{-1} \quad \Rightarrow \quad G = (Q^*SQ)^{-1}Q^*SHQ$$

which is equivalent to solving the hermitian eigenvalue problem

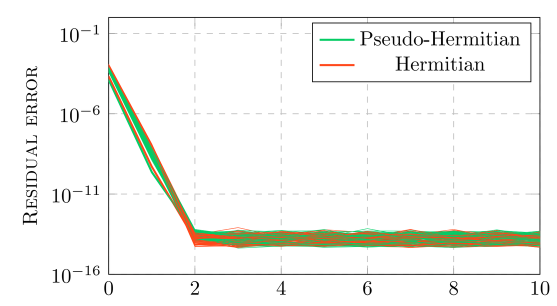
$$L^*(Q^*SQ)^{-1}L\boldsymbol{y} = \tilde{\lambda}\boldsymbol{y}, \quad \text{with} \quad \boldsymbol{y} = L^*\boldsymbol{w} \quad \text{and} \quad \boldsymbol{z} = L\boldsymbol{y}.$$

with L the Cholesky factor of the decomposition

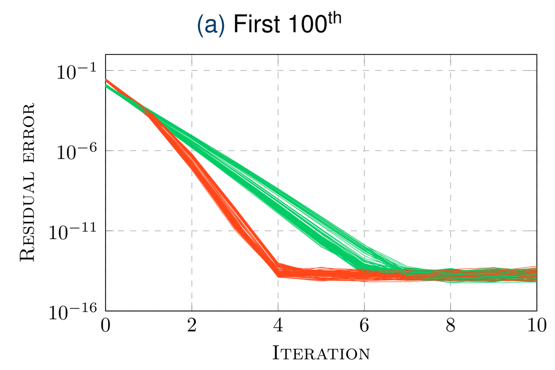
$$Q^*SHQ = LL^* \quad \text{since } SH \text{ is HPD.}$$

Construction of the hermitian Rayleigh Quotient with implicit dual space

1: Compute $T \leftarrow HQ$ ▷ BLAS GEMM n^2k
2: Flip the sign of the lower half $T \leftarrow ST$ ▷ INTERNAL FLIP $\frac{n}{2}k$
3: Compute $W \leftarrow Q^*T$ ▷ BLAS GEMM nk^2
4: Factorize $L \leftarrow$ CHOLESKY(W) ▷ LLAPACK PORTF k^3
5: Compute $M \leftarrow I - 2Q_2^*Q_2 \quad (= Q^*SQ)$ ▷ BLAS GEMM $\frac{n^2}{4}k$
6: Copy $G \leftarrow M$ ▷ BLAS COPY k^2
7: Solve $G \leftarrow L^{-1}(GL^{-*})$ ▷ 2 LLAPACK TRSM $2k^2$
8: Solve $GY = \tilde{\Lambda}Y$ ▷ LLAPACK HEEVD k^3
9: Back-transform & Return $\tilde{\Lambda}$, $Y \leftarrow \tilde{\Lambda}^{-1}, L^{-1}Y$ ▷ LLAPACK TRSM k^2



(a) First 100[th]



(b) Last 100[th]

Figure: Evolution of the residuals over nev = 1000 for a Silicon pseudo-hermitian H of size $n = 23552$ generated with Yambo

Key benefits of the Hermitian Rayleigh-Quotient

- Equivalent quality for left and right eigenspaces
$$\|(I - \Pi(Q_{\mathcal{L}}))\,\boldsymbol{u}\|_2 = \|(I - \Pi(Q))\,\boldsymbol{v}\|_2 \quad \Rightarrow \quad \mathcal{O}(\sigma_{\boldsymbol{v}}) = \mathcal{O}(\sigma_{\boldsymbol{u}})$$
- Bounded $\left|\tilde{\delta}\right|^{-1}$, such that
$$\left|\tilde{\delta}\right|^{-1} \leq \frac{\sqrt{\text{cond}(H)}}{\left|\lambda_{\min}(Q^*SQ)\right|}$$

These two features lead to **quadratic convergence** :

$$\left|\lambda - \tilde{\lambda}\right| \leq \kappa \cdot \mathcal{O}(\sigma_{\boldsymbol{v}}^2) \quad \text{with} \quad \kappa := \frac{\sqrt{\text{cond}(H)} \cdot \left\|H - \lambda I\right\|_2}{\left|\lambda_{\min}(Q^*SQ)\right|}.$$

Alternative Rayleigh-Ritz

Setting $M := \text{diag}(Q^*SQ)$ leads to

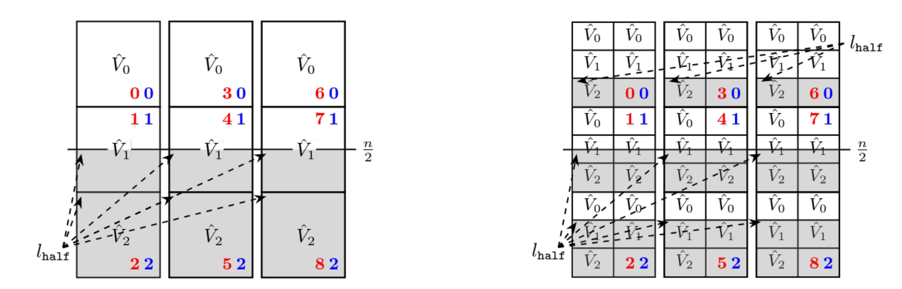
$$Q_L := [SQ - Q(Q^*SQ - \text{diag}(Q^*SQ))]\, \text{diag}^{-1}(Q^*SQ),$$

and the Rayleigh-Quotient G is given by

$$G = \text{diag}^{-1}(Q^*SQ)\,[Q^*SHQ - (Q^*SQ - \text{diag}(Q^*SQ))\, Q^*HQ]\,.$$

This alternative Rayleigh-Ritz requires a **non-hermitian eigensolver** (GEEV).

Flip-sign functions that emulate the multiplication by S



(a) Block-Block Distribution
(b) Block-Cyclic Distribution

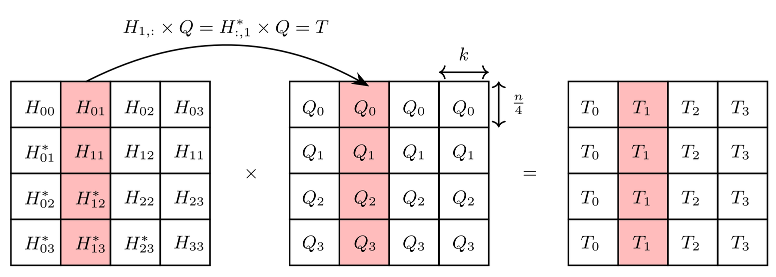
Figure: Flip sign function depending on the data distribution - (gray) sign flipped area, (red) MPI world rank, (blue) MPI column rank

Communication avoiding parallel GEMMs

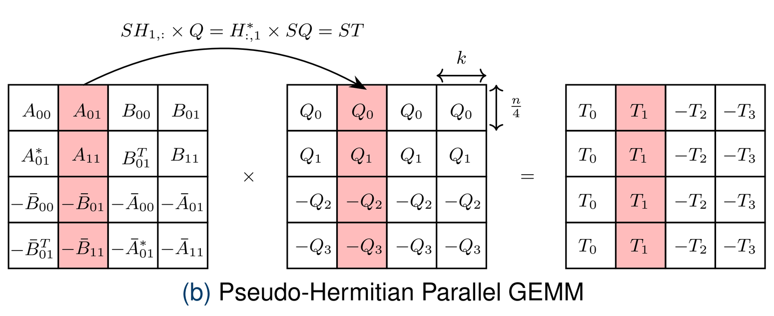
Chebyshev three-term recurrence relation

$$V_{i+1} = \alpha_i AV_i + \beta_i V_{i-1}, \quad V_1 = AV_0, \quad V_0 = \text{RAND}$$

```
V = random(N,nevex)
GEMM('N','N',alpha,A,V,0,W)
for d <= degree:
  GEMM('N','N',alpha,A,W,beta,V)
  V,W = W,V
```



(a) Hermitian Parallel GEMM



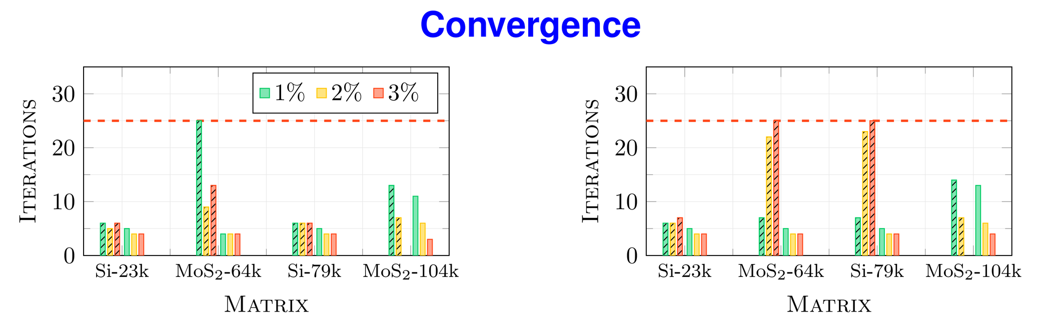
(b) Pseudo-Hermitian Parallel GEMM

Numerical and Parallel Performance Evaluation

Testing Platforms

Component	JUWELS Booster @JSC	Fugaku @Riken-RCCS
Nodes	936 nodes	158,976 nodes
CPU	2× AMD EPYC 7402, 96 threads	FUJITSU A64FX, 48 cores @ 2.2 GHz
CPU Memory	512 GB DDR4 RAM	32 GB HBM2 @ 1024 GB/s
GPU	4× A100 (40 GB) / NVLink 3	
MPI grid	1 node / 4 GPUs / 4 ranks	1 node / 4 ranks / 48 threads

Convergence



(a) tol $= 10^{-8}$
(b) tol $= 10^{-9}$

Figure: Number of iterations with respect to the matrix size - GEEV (dashed bars) vs. HEEVD (plain bars)

Percentage of TPP

	JUWELS Booster - 16 GPUs			Fugaku - 64 ranks *small* & *med* vs. 256 ranks *large*		
n	nev $= n \times 1\%$	nev $= n \times 2\%$	nev $= n \times 3\%$	nev $= n \times 1\%$	nev $= n \times 2\%$	nev $= n \times 3\%$
23552	36.2	43.8	**48.4**	29.872	31.642	**32.874**
64512	**78.2**	78.1	78.2	**42.9**	41.9	41.3
104832	**86.8**	84.4	81.3	**42.2**	38.7	33.7

Table: Percentage of TPP occupied by ChASE for each matrix size with respect to nev

Benchmark on JUWELS Booster

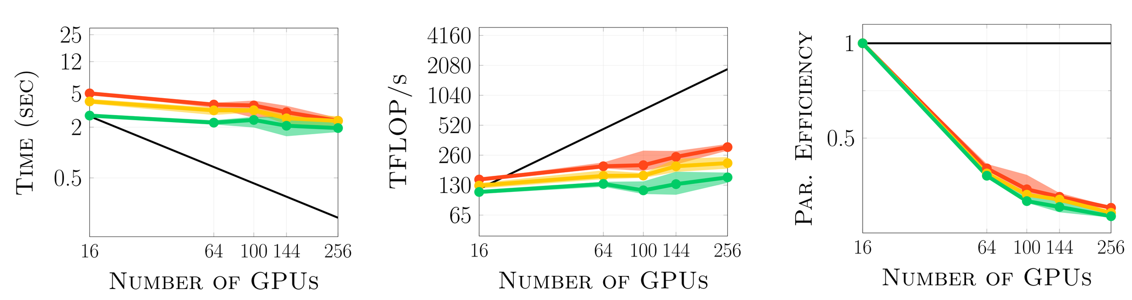


Figure: Performance benchmark on a *small* matrix of size $n = 23{,}552$ - nev = 235, 471, 706

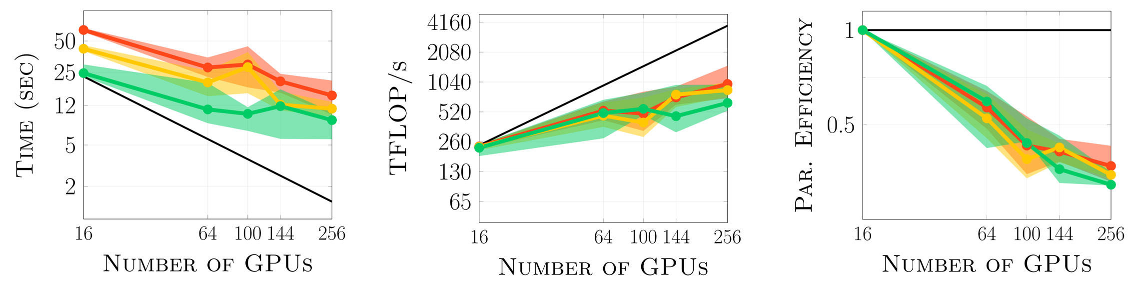


Figure: Performance benchmark on a *medium* matrix of size $n = 64{,}512$ - nev = 645, 1290, 1935

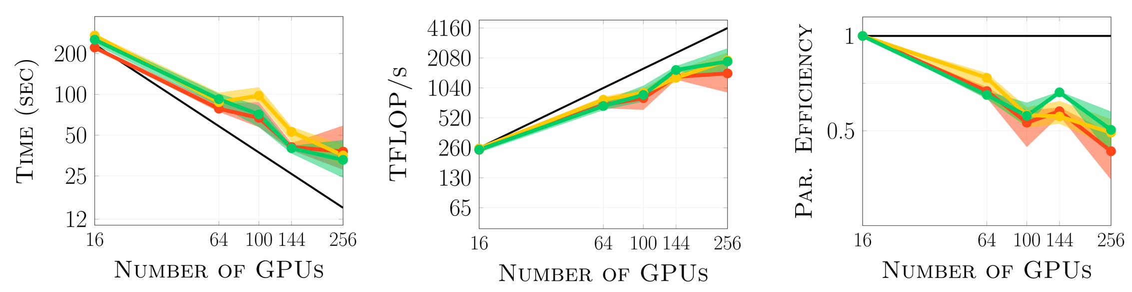


Figure: Performance benchmark on a *large* matrix of size $n = 104{,}832$ - nev = 1048, 2096, 3144

Benchmark on Fugaku

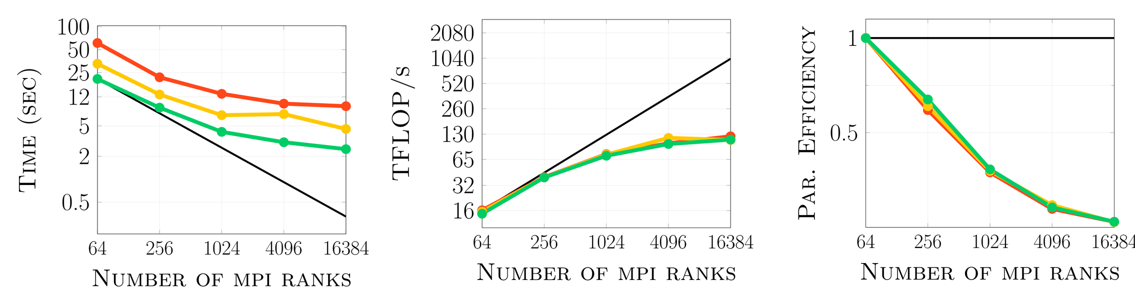


Figure: Performance benchmark on a *small* matrix of size $n = 23{,}552$ - nev = 235, 471, 706

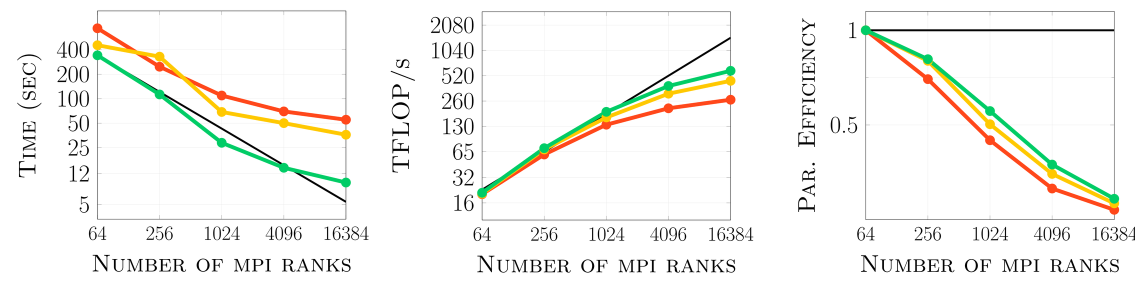


Figure: Performance benchmark on a *medium* matrix of size $n = 64{,}512$ - nev = 645, 1290, 1935

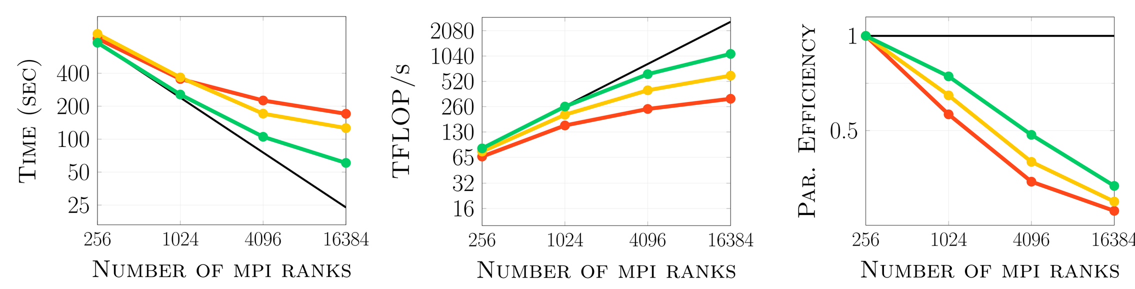


Figure: Performance benchmark on a *large* matrix of size $n = 104{,}832$ - nev = 1048, 2096, 3144

[1] J. Winkelmann, P. Springer, and E. Di Napoli. *ChASE: a Chebyshev Accelerated Subspace iteration Eigensolver for sequences of Hermitian eigenvalue problems.* ACM TOMS 45(2), Art.21 (2019).
[2] M. Berljafa, D. Wortmann, and E. Di Napoli. *An Optimized and Scalable Eigensolver for Sequences of Eigenvalue Problems.* Concurrency Computat.: Pract. Exper. 27 (2015), pp. 905–922.
[3] X. Wu, D. Davidović, S. Achilles, E. Di Napoli. *ChASE: a distributed hybrid CPU-GPU eigensolver for large-scale hermitian eigenvalue problems.* PASC22.
[4] X. Wu, E. Di Napoli. *Advancing the distributed Multi-GPU ChASE library through algorithm optimization and NCCL.* SC'23 Workshops, pp. 1688–1696.