Efficient Serial and Parallel Block-Jacobi EVD/SVD Algorithms

Martin Bečka, Gabriel Okša

Institute of Mathematics, Department of Informatics,
Slovak Academy of Sciences, Bratislava, Slovakia

65 years of IM SAS, October 13–15, 2024, Smolenice

Outline

Our task

Compute in parallel the Singular Value Decomposition (SVD) of a complex matrix A of the size $m \times n$, $m \geq n$:

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^H,$$

where $U(m \times m)$ and $V(n \times n)$ are orthogonal and $\Sigma = \mathrm{diag}(\sigma_i)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.
Numerically stable way of computation:

- one- or two-sided block-Jacobi methods;
- large degree of parallelism.

Target architecture:

- distributed memory machines (parallel supercomputers and clusters) with Message Passing Interface (MPI).

Applications

- EVD of symmetric matrices: quantum energy of atoms and molecules.
- Latent Semantic Indexing: processing and searching documents.
- Civil Engineering: eigenfrequencies and eigenmodes of buildings.

Our international partners:

- University of Electro-Communications, Tokyo, Japan
- Institute of Informatics, AS CR, Prague
- University of Salzburg, Austria
- University of Zagreb, Croatia

Recent VEGA grants

- VEGA project no. 2/004/17 “Parallel block algorithms for some canonical matrix decompositions”, 2017–2019.
- VEGA project no. 2/0015/20 “Convergence of block algorithms for canonical matrix decompositions”, 2020–2022.
- VEGA project no. 2/0001/23 “Efficient block Jacobi algorithms for the matrix EVD/SVD and their numerical properties”, 2023–2025.

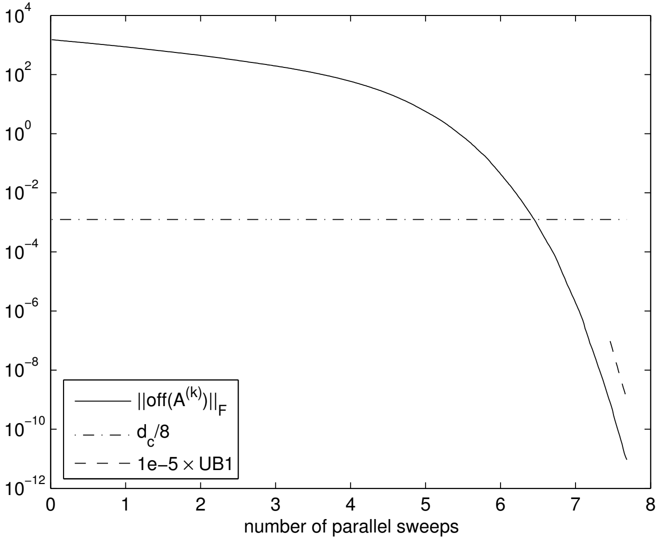
Asymptotic quadratic convergence

Having p processors, the blocking factor is $w = 2p$ and a matrix is partitioned into the $w \times w$ block structure.
Then, under some additional assumptions, there exists such an integer constant W, $w - 1 \leq W < 2w(\log w + 1)$, that after W parallel iteration steps one observes the AQC of the off-diagonal Frobenius norm of matrix A:

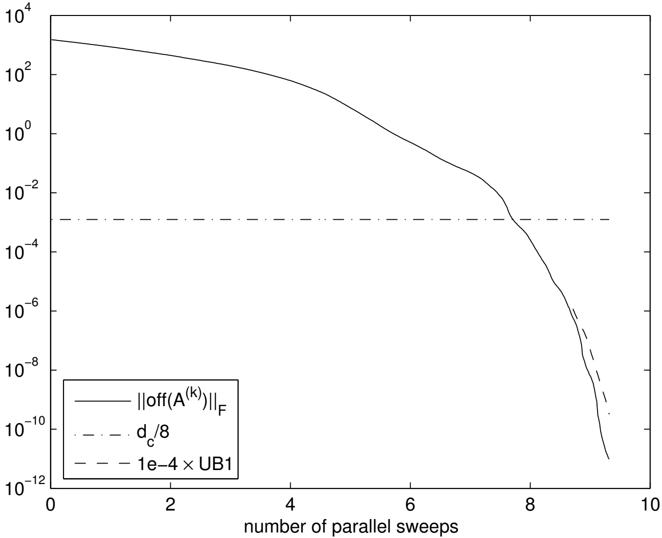
$$\|\mathrm{off}(A^{(W)})\|_F \leq \sqrt{12(w-2)}\,\frac{\|\mathrm{off}(A^{(0)})\|_F^2}{\delta},$$

where $\delta = \sqrt{2}d_c/4$ and d_c is the minimal gap between the centres of the clusters of singular values.

AQC present


10^4
10^2
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
0
1
2
3
4
5
6
7
8
number of parallel sweeps
||off(A^(k))||_F
d_c/8
1e–5 × UB1

AQC not present



Preconditioning

- Main idea: Find a ‘cheap’ orthogonal matrix matrix P such that AP will be close to $U\Sigma$ in the sense of the proximity of column vector spaces.
- First approach: Compute the EVD of the symmetric matrix $A^T A$ and use its eigenvectors W as a preconditioner: $A \to AW$, AW is the input into the block-Jacobi SVD algorithm.
- Suitable for well conditioned matrices when the condition number of $A^T A$ is not too large.
- Second approach: Compute the (partial) polar decomposition of $A = U_p H$ by the Halley iterations (cubic convergence!), then the EVD of the Hermitian factor H, and use its eigenvectors W for precond.
- Since the Gram matrix $A^T A$ is not needed at all, this approach is suitable for very ill conditioned matrices.

W based on $A^T A$, in parallel

Table: $w = 2p$, $n = 8000$, $\texttt{mode} = 3$, $\kappa(A) = 10^2$

Algorithm		p	
		10	20
PDGESVD	T [s]	934	645
PP_OSBJA	T [s]	178	98
	G+EVD+MM [s]	14+67+23	7+39+12
	Jacobi [s]	74	40
	# it	10	22

W based on $A^T A$, in parallel

Table: $w = 2p$, $n = 8000$, $\texttt{mode} = 3$, $\kappa(A) = 10^8$

Algorithm		p	
		10	20
PDGESVD	T [s]	859	499
PP_OSBJA	T [s]	408	202
	G+EVD+MM [s]	14+61+23	7+36+12
	Jacobi [s]	310	147
	# it	33	72

W based on HI, serial algorithm

- 6 iterations in the Halley algorithm are sufficient in the double precision, here we use only one iteration for the partial polar decomposition.
- 3 variants V1–V3 of serial algorithm for the partial polar decomposition were developed, they differ in the QR decomposition of a highly structured matrix of size 8000×4000 (not shown).

Table: $n = 4000$, $\texttt{mode} = 3$, $\kappa(A) = 10^{10}$

Precond. W	V1	V2	V3	Gram
$maxw$	5.4e-4	8.2e-4	1.3e-3	1.0
$iter(J)$	41	38	39	52
$T(J)$[s]	45.1	42.4	43.9	72.6
T_{tot}[s]	**65.9**	136.3	98.2	81.6

Recent publications

1 YAMAMOTO, Y., OKŠA, G., VAJTERŠIC, M., *On convergence to eigenvalues and eigenvectors in the block-Jacobi EVD algorithm with dynamic ordering*, Lin. Alg. and Its Appl. , 622 (2021) 19-45.

2 M. BEČKA, G. OKŠA, *Preconditioned Jacobi SVD algorithm outperforms PDGESVD*, In: Proc. of PPAM 2019, LNCS 12043 (2020), 555-566, Springer Nature Switzerland AG.

3 G. OKŠA, Y. YAMAMOTO, M. VAJTERŠIC, *Convergence to singular triplets in the two-sided block-Jacobi SVD algorithm with dynamic ordering*, SIAM Journal on Matrix Anal. and Appl. 43 (2022) 1238-1262.

4 G. OKŠA, M. BEČKA, *On relative accuracy of the one-sided block-Jacobi SVD algorithm*, In: Proc. of PPAM 2022, LNCS 12043 (2023), 464-475, Springer Nature Switzerland AG.