

# On Minimizing Arithmetic and Communication Complexity of Jacobi’s Eigenvalue Method: Review and Beyond

Yifu Wang, James Demmel, Hengrui Luo, Ryan Schneider

## Abstract

Jacobi’s method iteratively computes the eigenvalues and eigenvectors of a symmetric matrix. Remarkably simple to implement, Jacobi’s method is a compelling candidate for use on large-scale applications. On the other hand, matrix multiplication is fundamental in numerical linear algebra, often regarded as a building block for other matrix computations.

With these in mind, we establish theoretical bounds on the asymptotic complexity of Jacobi’s method in both arithmetic and communication, aiming for efficiency comparable to matrix multiplication.

We not only analyze the complexity of sequential and parallel Jacobi using classical  $O(n^3)$  matrix multiplication, but also introduce recursive Jacobi’s methods that leverage Strassen-like  $O(n^{\omega_0})$  matrix multiplication to achieve optimal arithmetic and communication lower bounds. We also offer rigorous proofs of convergence for the recursive algorithms. The main contributions are as follows:

1. Starting from a dense real symmetric matrix  $\mathbf{A} \in \mathbf{M}_n(\mathbb{R})$  (without loss of generality, we only consider the real case), the **Classical Jacobi’s method** sequentially rotates all off-diagonal entries of  $\mathbf{A}$  in some given ordering. We denote one *sweep* as rotating through all off-diagonal entries of  $\mathbf{A}$  once. Since Classical Jacobi almost always converges, we assume that the algorithm converges in  $O(1)$  sweeps and the corresponding total arithmetic cost is  $O(1) \cdot \Theta(n^3) = \Theta(n^3)$ .

For estimating the lower bound on the communication cost, assume for now that we could only change the ordering of rotations. We denote the size of fast memory by  $M$ . Then when  $M^{1/2} < n < M$ , we can attain a lower bound of  $\Omega(n^4/M)$  reads and writes to slow memory, asymptotically exceeding the  $O(n^3/\sqrt{M})$  cost of classical matrix multiplication. To attain the cost of matrix multiplication requires more changes to the algorithm.

2. Allowing ourselves more freedom than just choosing the ordering of to-be-rotated entries, we next consider the **Block Jacobi’s method**, in which we rotate  $2b$ -by- $2b$  blocks instead of one off-diagonal entry each time. We still assume  $O(1)$  sweeps for the algorithm to converge and choose  $b$  to be able to fit three  $2b$ -by- $2b$  sub-matrices into the fast memory, i.e.  $b = \Theta(\sqrt{M})$ . In this case, the algorithm attains the communication lower bound  $\Omega(n^3/\sqrt{M})$  with  $O(n^3)$  matrix multiplication.
3. The highlight of this paper is the **Recursive Jacobi’s method** we introduce, along with a series of its variations. To the best of our knowledge, this is the first work which can asymptotically attain the arithmetic and communication costs of Strassen-like matrix multiplication, including a convergence proof.

We first propose a “vanilla” recursive algorithm, in which we apply a divide-and-conquer strategy, where the algorithm recursively partitions the input  $n$ -by- $n$  matrix into smaller  $2b$ -by- $2b$  blocks, until the size of the to-be-rotated sub-matrices reach a certain threshold, where  $b = n^f$  and  $0 < f < 1$  is the block parameter. We show that under the assumption that

the outermost sweep is executed  $O(1)$  times, the arithmetic complexity is  $F(n) = O(\frac{\log \log n}{-\log f} \cdot n^{3(1-f)+\omega_0 f})$ , which asymptotically approaches  $O(n^{\omega_0})$  as  $f$  approaches 1.

Convergence analysis for Jacobi’s methods has been widely discussed, taking into account various pivoting strategies (such as rotation orderings and the choice between block and cyclic) as well as processing architectures (sequential or parallel). We refer readers to [7, 8, 10, 14] for further details. A key ingredient in [7] towards convergence of Classical Jacobi is to restrict the rotation angles of off-diagonal entries in a proper open subset of  $(-\frac{\pi}{2}, \frac{\pi}{2})$ . An analog for block Jacobi is uniformly bounded cosine transformations [6]. By reordering the columns of the orthogonal rotation matrix  $\mathbf{Q}$  via applying QR decomposition with column pivoting (QRCP for short) to the first-half leading rows of  $\mathbf{Q}$ , [6] successfully proved convergence for the block cyclic Jacobi. We leverage this idea and introduce our first variant of recursive Jacobi with convergence guarantee, the **Recursive Jacobi with QRCP**. With a slight trade-off between optimal arithmetic complexity lower bound and convergence guarantee, the recursive Jacobi with QRCP, within  $O(1)$  sweeps, can achieve

$$F(n) = \begin{cases} O(\frac{\log \log n}{-\log f} \cdot n^{3(1-f)+\omega_0 f}), & 0 < f \leq \frac{1}{4-\omega_0} \\ O(\frac{\log \log n}{-\log f} \cdot n^{2+f}), & \frac{1}{4-\omega_0} < f < 1 \end{cases}$$

due to the  $O(n^3)$  expensive complexity of QRCP.

The key of ensuring convergence in [6, 7] is to bound the cosines of rotation angles away from zero, which could also be done by applying LU decomposition with partial pivoting (LUPP for short). Unlike QRCP, LU decomposition can be implemented recursively with complexity of  $O(n^{\omega_0})$  [2, Section 4.2], and adding partial pivoting to the algorithm doesn’t increase the arithmetic complexity. By applying LUPP to the transpose of the first-half leading columns of  $\mathbf{Q}$ , we introduce the **Recursive Jacobi with LUPP** which enjoys both optimal arithmetic complexity and convergence.

In the sequential case, for  $2 < \omega_0 \leq 3$ , the recursive Jacobi is shown to analogously get close to attaining the expected communication lower bound  $\Omega(n^{\omega_0}/M^{\omega_0/2-1})$  [13]. In practical terms, recursive Jacobi should be considered as a “galactic algorithm” since the size  $n$  where the algorithm shows benefits grows rapidly as  $f$  approaches 1.

4. In addition to the sequential cases, we also studied **parallel block Jacobi** with  $O(n^3)$  matrix multiplication, in which the algorithm simultaneously rotates off-diagonal blocks in different columns and rows [1, 9, 12]. We store the  $n$ -by- $n$  matrix  $\mathbf{A}$  on a  $\sqrt{P} \times \sqrt{P}$  grid of  $P$  processors, with block sizes  $b = n/\sqrt{P}$ , which we assume to be an integer for simplicity. Under this scenario, the arithmetic complexity is  $O(n^3/P)$ , which demonstrates the optimal linear speedup, and the communication complexity is  $O(n^2/\sqrt{P})$  words and  $O(\sqrt{P} \log P)$  messages, which attains the communication lower bound (except for the  $\log P$  factor) for classical matrix multiplication using the minimum amount of memory.

One remark is that the above studies and estimates readily extend to the SVD due to its strong connection with Jacobi’s method [4, 5]. Furthermore, by not restricting ourselves to Jacobi-like methods, our recursive algorithm technique can also benefit non-Jacobi methods, for example combined with QDWH (QR-based dynamically weighted Halley algorithm) [11].

Additionally, since all our recursive algorithms follow a divide-and-conquer paradigm utilizing  $O(n^{\omega_0})$  matrix multiplication, it follows from the analysis in [2, 3] that all the proposed algorithms are backward stable.

In conclusion:

1. We have demonstrated an asymptotic approach to make the Jacobi's eigenvalue method and SVD nearly as fast as matrix multiplication, in terms of both arithmetic and communication complexity, across several scenarios. For  $O(n^3)$  matrix multiplication, we analyzed both sequential and parallel Jacobi's methods.

*A remaining open question is whether the (better) lower bound and communication complexity for matrix multiplication using more than the minimum memory is attainable for Jacobi.*

2. For  $O(n^{\omega_0})$  matrix multiplication, we introduced a series of recursive Jacobi's methods, focusing on minimizing arithmetic cost while also ensuring the convergence of the proposed algorithms.

*Another remaining open question is whether these asymptotically faster recursive Jacobi's methods can be parallelized and attain both the arithmetic and communication complexity lower bounds of matrix multiplication.*

## References

- [1] M. Berry and A. Sameh. An overview of parallel algorithms for the singular value and symmetric eigenvalue problems. *J. Comp. Appl. Math.*, 27:191–213, 1989.
- [2] James Demmel, Ioana Dumitriu, and Olga Holtz. Fast linear algebra is stable. *Numerische Mathematik*, 108(1):59–91, 2007.
- [3] James Demmel, Ioana Dumitriu, Olga Holtz, and Robert D. Kleinberg. Fast matrix multiplication is stable. *Numerische Mathematik*, 106:199–224, 2006.
- [4] K. Drmač and K. Veselić. New fast and accurate Jacobi SVD algorithm, I. *SIAM J. Mat. Anal. Appl.*, 29(4):1322–1342, 2008.
- [5] K. Drmač and K. Veselić. New fast and accurate Jacobi SVD algorithm, II. *SIAM J. Mat. Anal. Appl.*, 29(4):1343–1362, 2008.
- [6] Zlatko Drmač. A Global Convergence Proof for Cyclic Jacobi Methods with Block Rotations. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1329–1350, 2010.
- [7] G. E. Forsythe and P. Henrici. The Cyclic Jacobi Method for Computing the Principal Values of a Complex Matrix. *Transactions of the American Mathematical Society*, 94(1):1–23, 1960.
- [8] V. Hari. Convergence to diagonal form of block Jacobi-type methods. *Numer. Math.*, 129:449–481, 2015.
- [9] Franklin T. Luk and Haesun Park. A Proof of Convergence for Two Parallel Jacobi SVD Algorithms. *IEEE Trans. Computers*, 38:806–811, 1989.
- [10] W. Mascarenhas. Convergence of the Jacobi method for arbitrary orderings. *SIAM J. Mat. Anal. Appl.*, 16(4):1197–1209, Oct 1995.
- [11] Yuji Nakatsukasa and Nicholas J. Higham. Stable and Efficient Spectral Divide and Conquer Algorithms for the Symmetric Eigenvalue Decomposition and the SVD. *SIAM Journal on Scientific Computing*, 35(3):A1325–A1349, 2013.

- [12] A. Sameh. On Jacobi and Jacobi-like algorithms for parallel computers. *Math. Comp.*, 25(115):579–590, July 1971.
- [13] Jacob Scott. *An I/O-Complexity Lower Bound for All Recursive Matrix Multiplication Algorithms by Path-Routing*. PhD thesis, UC Berkeley Mathematics PhD thesis, 2015.
- [14] G. Shroff and R. Schreiber. On the convergence of the cyclic Jacobi method for parallel block orderings. *SIAM J. Mat. Anal. Appl.*, 10(3):326–346, 1989 1989.