High Performance Implementation of LTL^T Decomposition of a Skew-Symmetric Matrix

Chao Yin, Ishna Satyarth, Ruqing Xu, Devin Matthews

Sep. 26th, 2024

Austin, TX, USA

SMU

Background

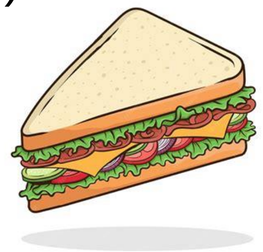
Blocked algorithms

Blocked Left Looking Algorithm (BLL)

$$\begin{pmatrix} x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix} := \begin{pmatrix} x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix} - \begin{pmatrix} L_{20} & l_{21} \\ \hline l_{30}^T & \lambda_{31} \\ \hline L_{40} & l_{41} \end{pmatrix} \begin{pmatrix} X_{00} & \star \\ \hline x_{10}^T & 0 \end{pmatrix} \begin{pmatrix} l_{10} & L_{20}^T \\ \hline 1 & l_{21}^T \end{pmatrix}$$

Skew-Triangular-Matrix-Multiplication (SK-TRI-GEMM)

$$\left[\begin{pmatrix} \chi_{11} & \star \\ \hline x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix}, \begin{pmatrix} L_{22} & 0 \\ \hline l_{32}^T & 1 \\ \hline L_{42} & l_{43} \end{pmatrix} \right] := \text{LTLT_UNB}\left(\begin{pmatrix} \chi_{11} & \star \\ \hline x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix} \right)$$



Blocked Right Looking Algorithm (BRL)

$$\left[\begin{pmatrix} \chi_{11} & \star \\ \hline x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix}, \begin{pmatrix} L_{22} & 0 \\ \hline l_{32}^T & 1 \\ \hline L_{42} & l_{43} \end{pmatrix} \right] := \text{LTLT_UNB_0}\left(\begin{pmatrix} \chi_{11} & \star \\ \hline x_{21} & X_{22} \\ \hline \chi_{31} & x_{32}^T \\ \hline x_{41} & X_{42} \end{pmatrix} \right)$$

$$\begin{pmatrix} \chi_{33} & \star \\ \hline x_{43} & X_{44} \end{pmatrix} := \begin{pmatrix} \chi_{33} & \star \\ \hline x_{43} & X_{44} \end{pmatrix} - \begin{pmatrix} l_{32}^T & 1 \\ \hline L_{42} & l_{43} \end{pmatrix} \begin{pmatrix} X_{22} & \star \\ \hline x_{32}^T & 0 \end{pmatrix} \begin{pmatrix} l_{32} & L_{42}^T \\ \hline 1 & l_{43}^T \end{pmatrix}$$

Skew-Triangular-Matrix-Multiplication (Triangular) (SK-TRI-GEMMT)

$$X_{44} := X_{44} + (l_{43} x_{43}^T - x_{43} l_{43}^T)$$

Skew-Symmetric Rank-2 (SKR2)

https://doi.org/10.48550/arXiv.2311.10700

SMU

Background

Unblocked algorithms

Unblocked Left Looking Algorithm (UBLL)

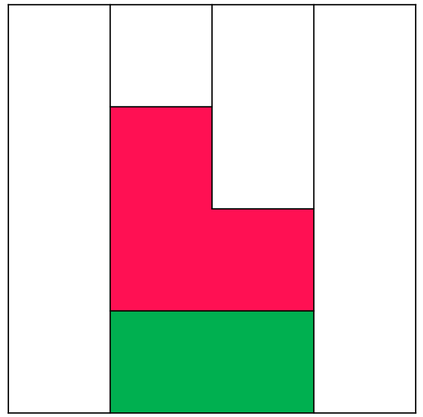
$$\left(\begin{array}{c} \chi_{21} \\ \hline x_{31} \end{array}\right) := \left(\begin{array}{c} \chi_{21} \\ \hline x_{31} \end{array}\right) - \left(\begin{array}{c|c} l_{20}^T & \lambda_{21} \\ \hline L_{30} & l_{31} \end{array}\right)\left(\begin{array}{c|c} X_{00} & -x_{10} \\ \hline x_{10}^T & 0 \end{array}\right)\left(\begin{array}{c} l_{10} \\ \hline 1 \end{array}\right)$$

Skew-Triangular-Matrix-Vector-Multiplication (SKTRI-GEMV)

$$l_{32} := x_{31}/\chi_{21}$$
$$x_{31} := 0$$

Unblocked Right Looking Algorithm (UBRL)

$$l_{32} := x_{31}/\chi_{21}$$
$$x_{31} := 0$$
$$X_{33} := X_{33} + (l_{32}x_{32}^T - x_{32}l_{32}^T)$$



Skew-Symmetric Rank-2 (SKR2)

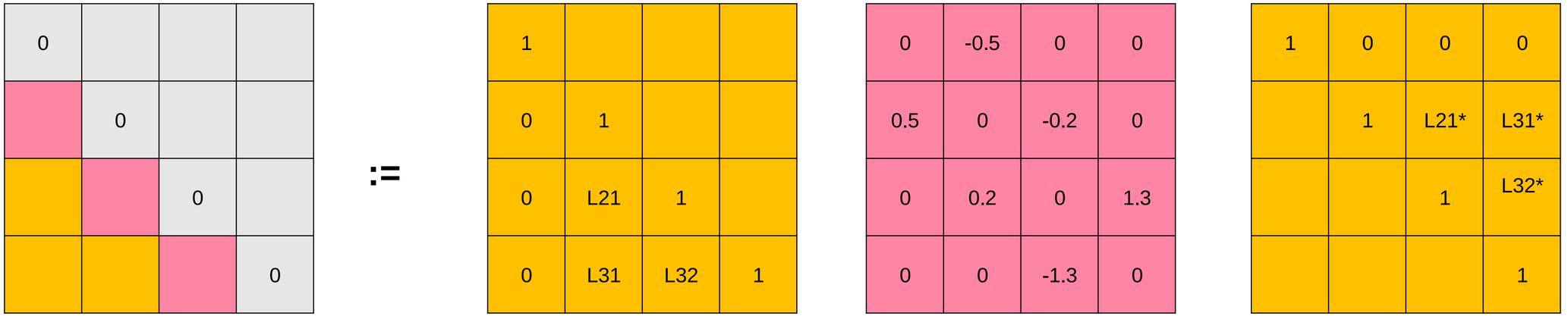
General Matrix rank-2 update (GER2)

https://doi.org/10.48550/arXiv.2311.10700

SMU

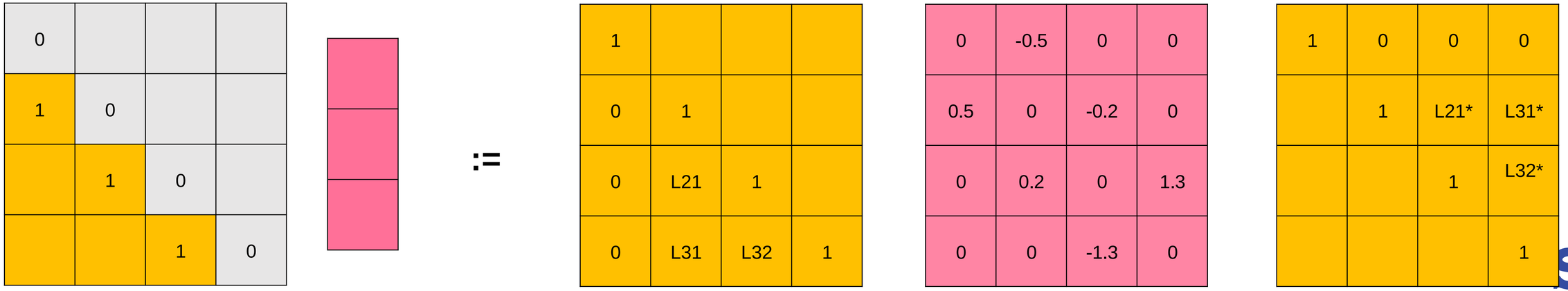
Implementation (in-place or out-of-place)

T Matrix in-place



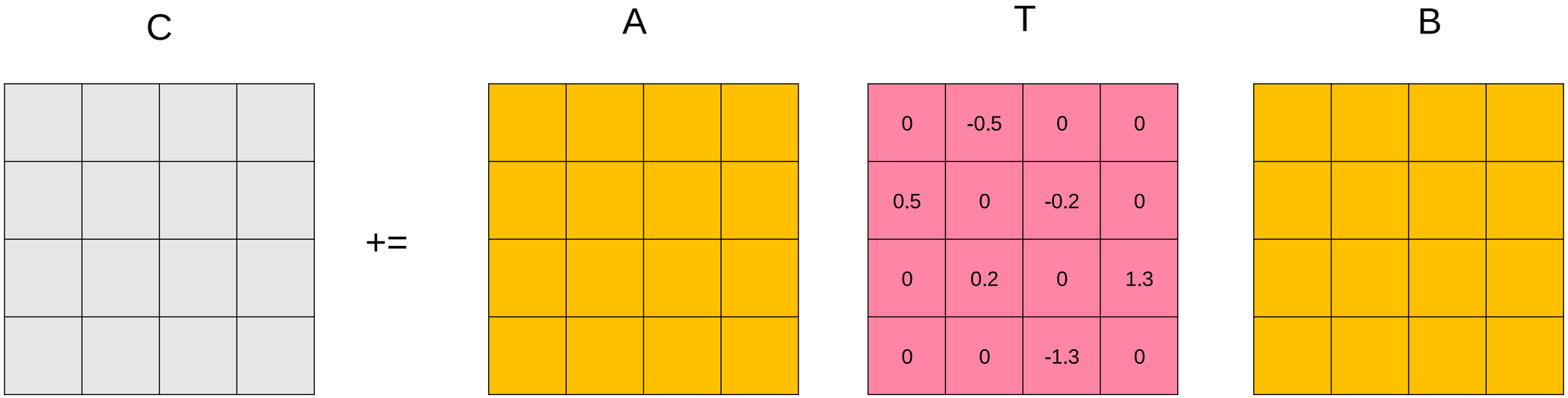
T Matrix out-of-place

$$\left(\begin{array}{c|c} \chi_{33} & \star \\ \hline x_{43} & X_{44} \end{array}\right) := \left(\begin{array}{c|c} \chi_{33} & \star \\ \hline x_{43} & X_{44} \end{array}\right) - \left(\begin{array}{c|c} l_{32}^T & 1 \\ \hline L_{42} & l_{43} \end{array}\right)\left(\begin{array}{c|c} X_{22} & \star \\ \hline x_{32}^T & 0 \end{array}\right)\left(\begin{array}{c|c} l_{32} & L_{42}^T \\ \hline 1 & l_{43}^T \end{array}\right)$$

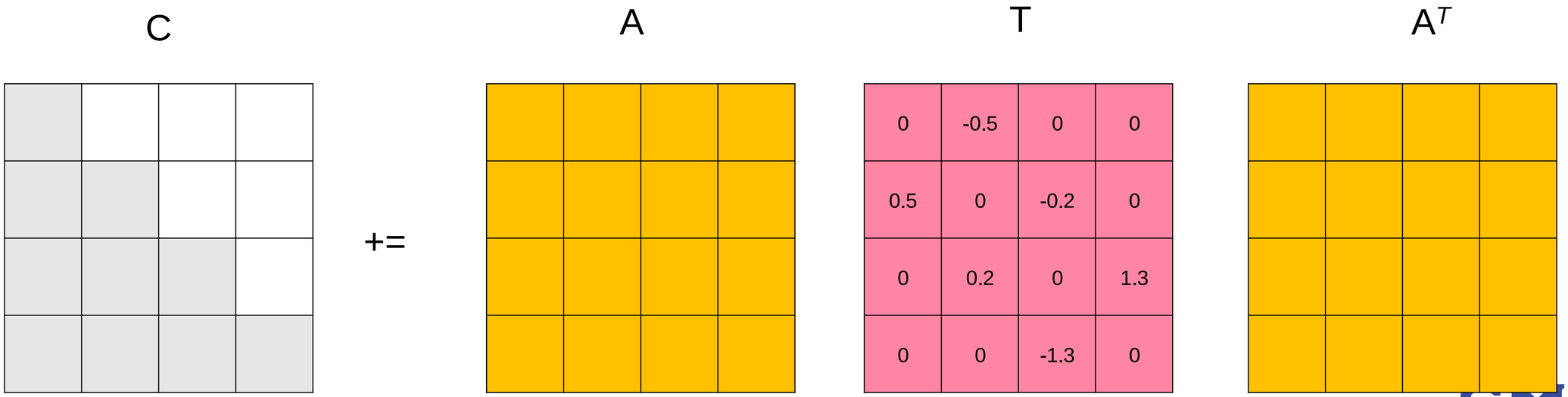


Implementation (Level-3 BLAS)

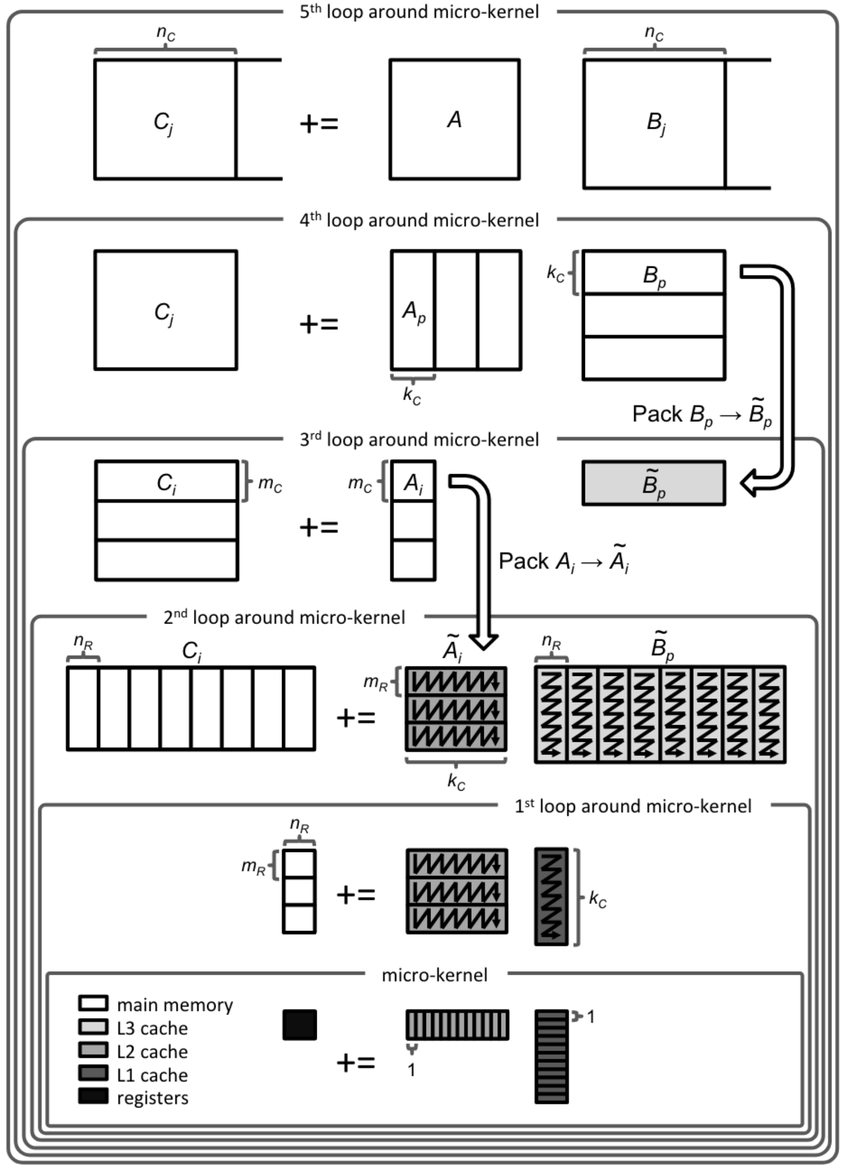
SKTRI-GEMM

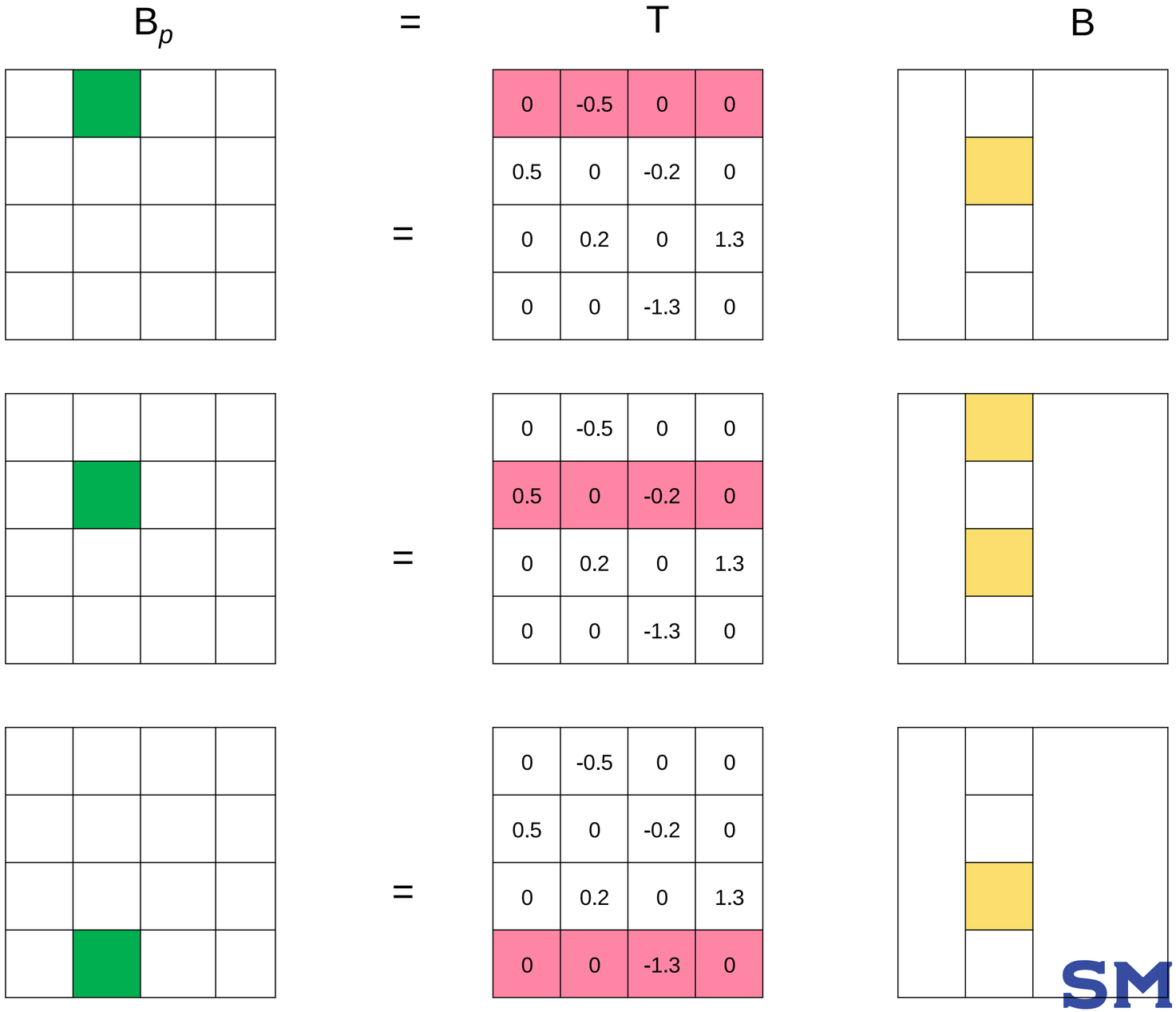


SKTRI-GEMMT

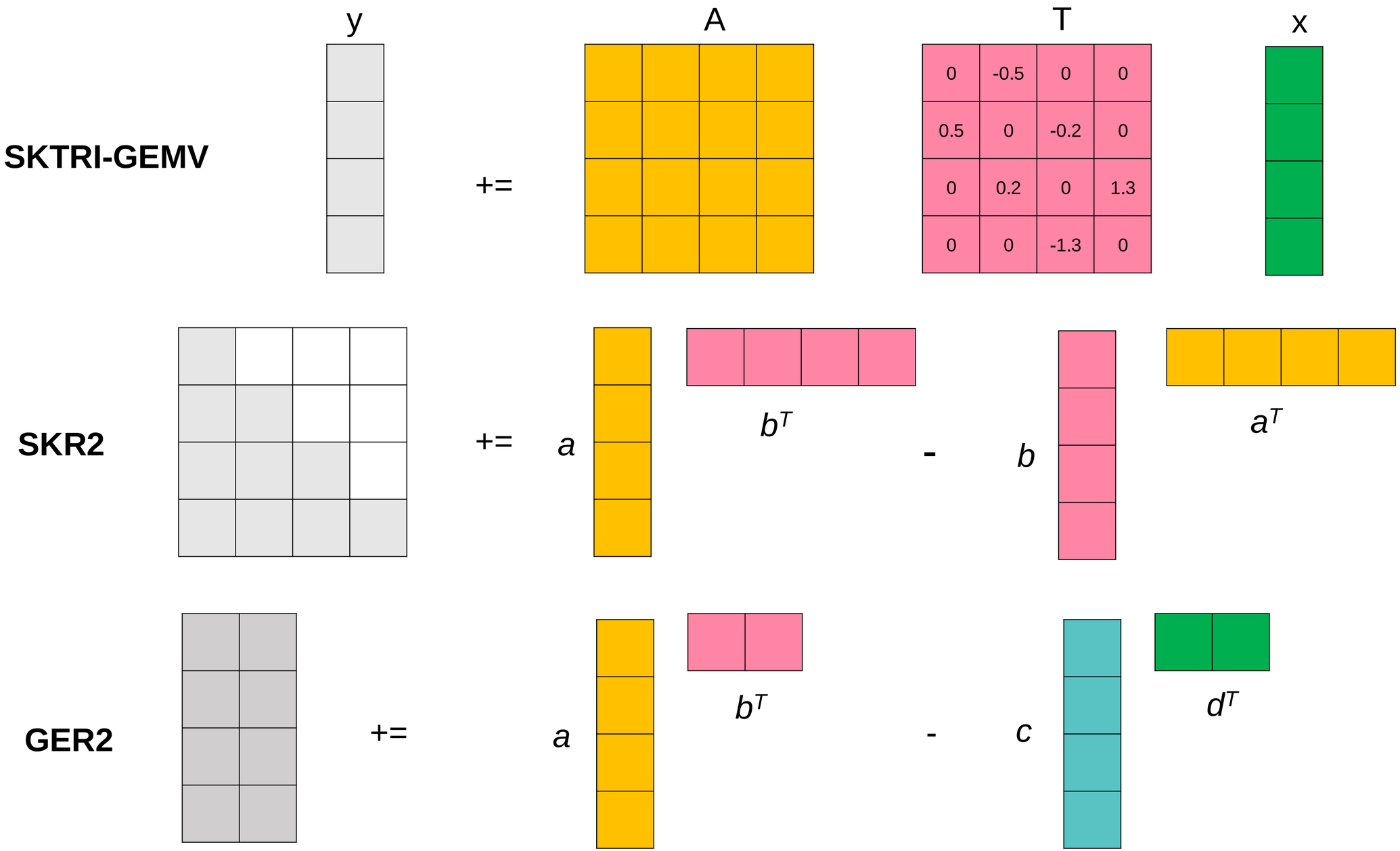


Implementation (Level-3 BLAS)

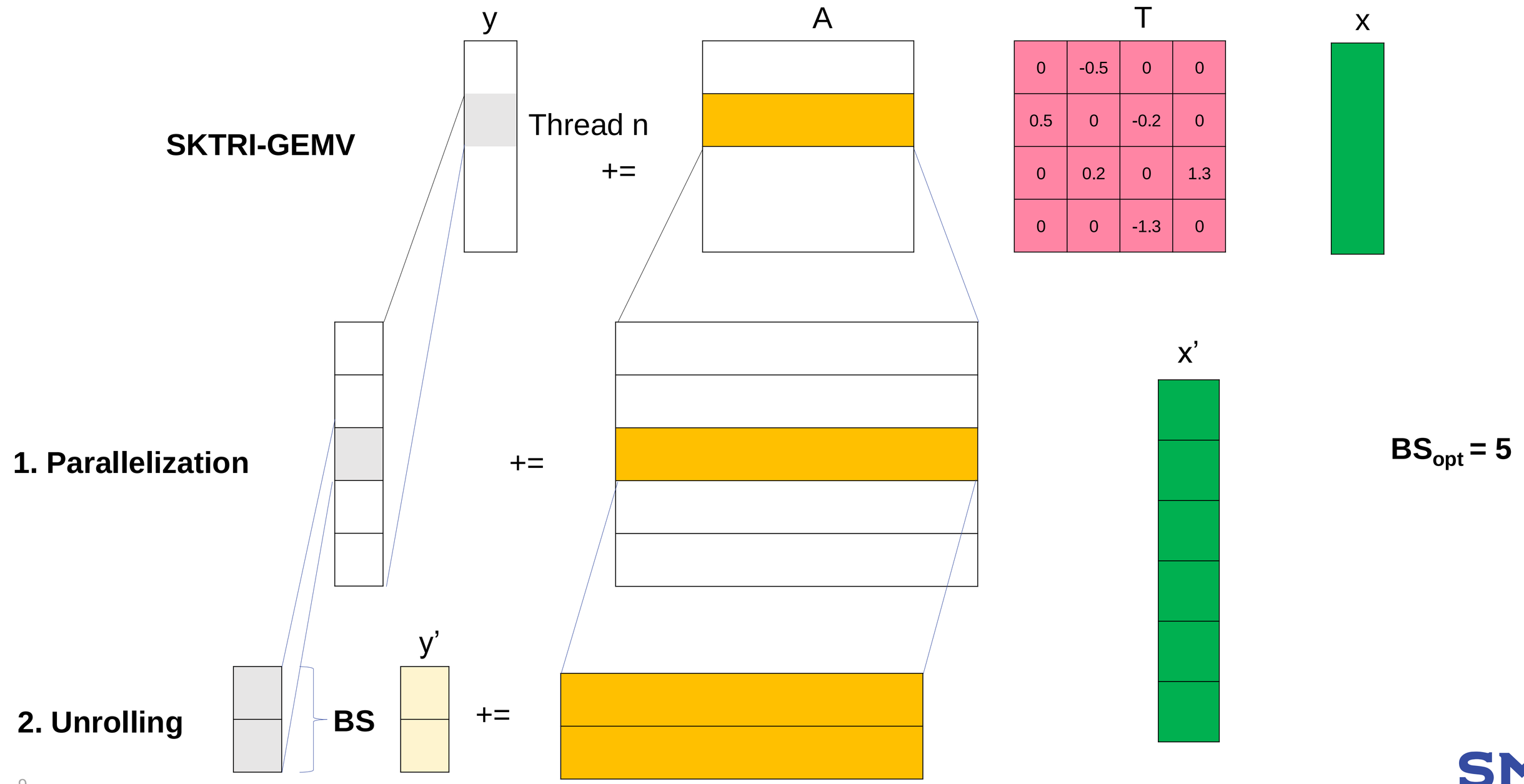




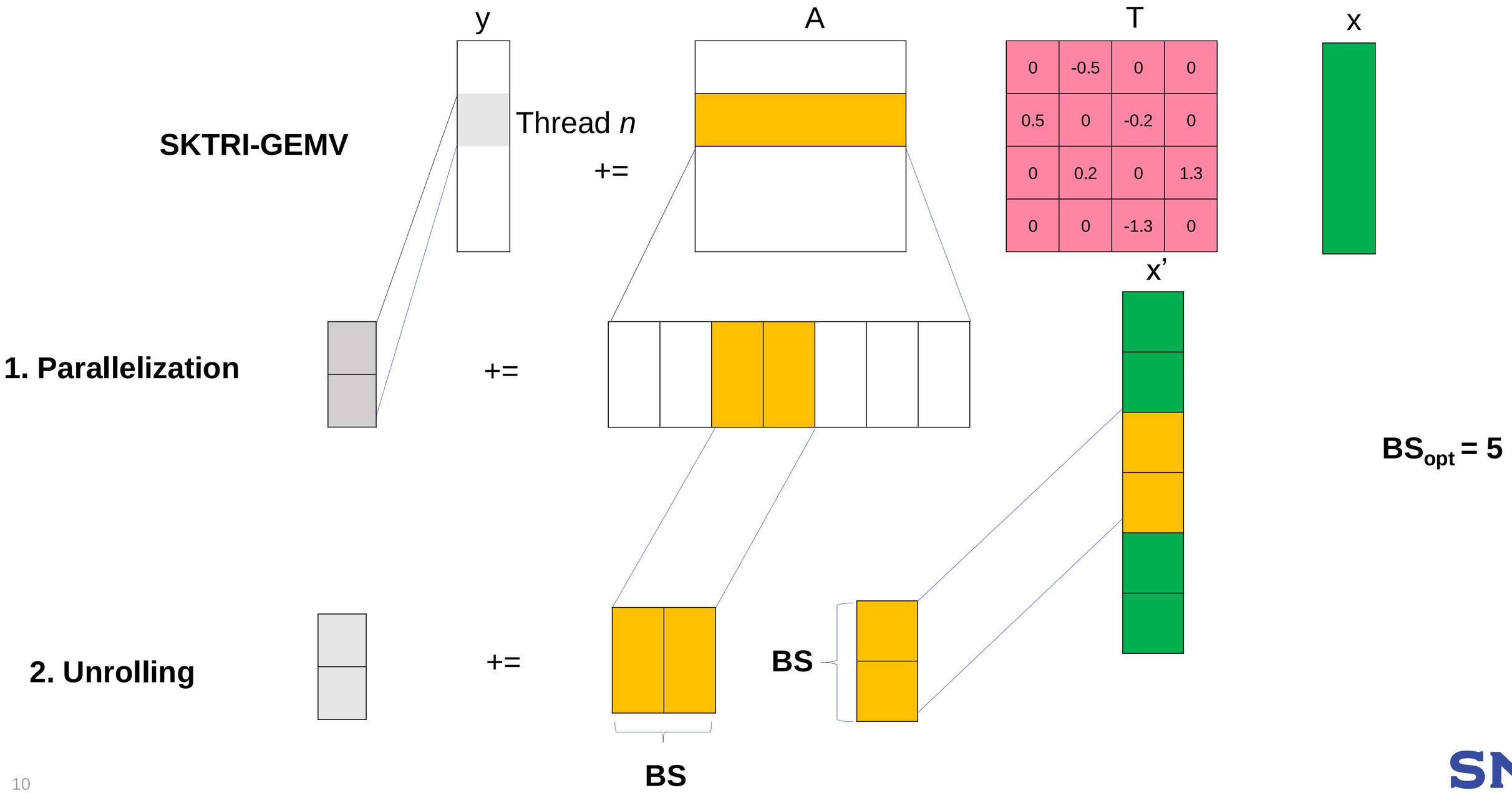
Implementation (Level-2 BLAS)



Implementation (SK-TRI-GEMV, Row-Major)

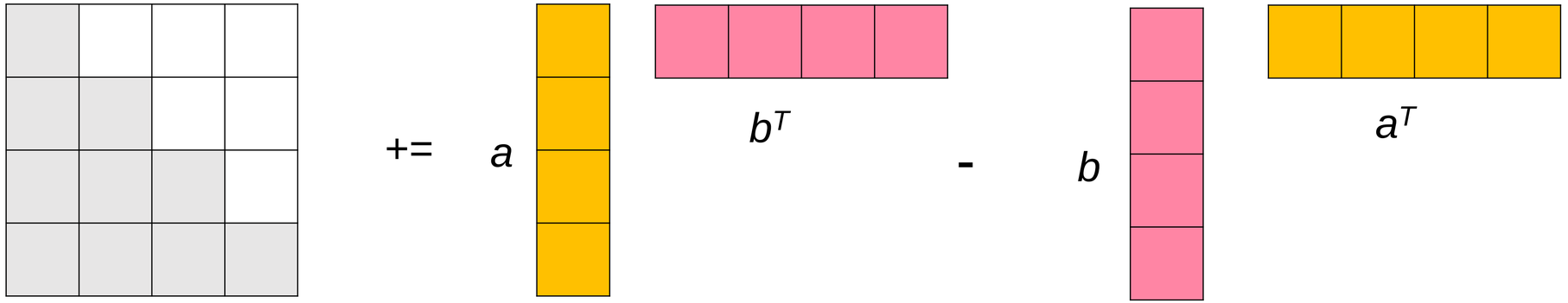


Implementation (SK-TRI-GEMV, Column-Major)



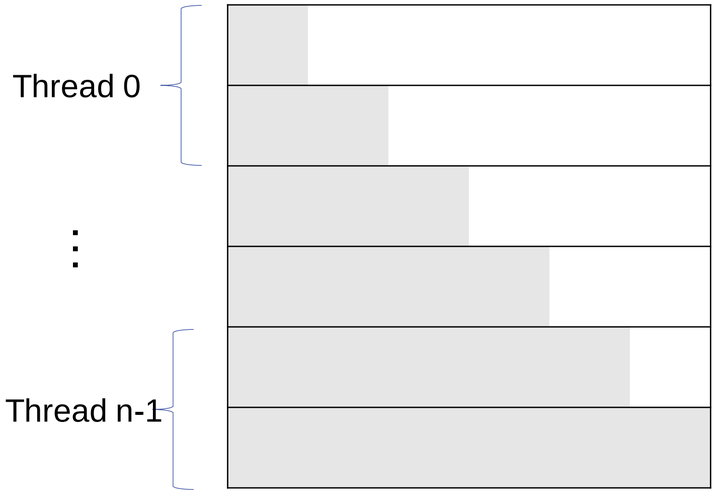
Implementation (SKR2)

SKR2

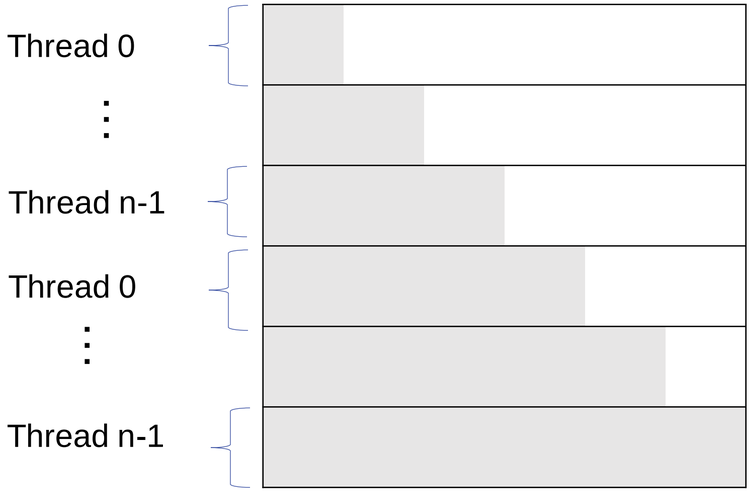


Parallelization

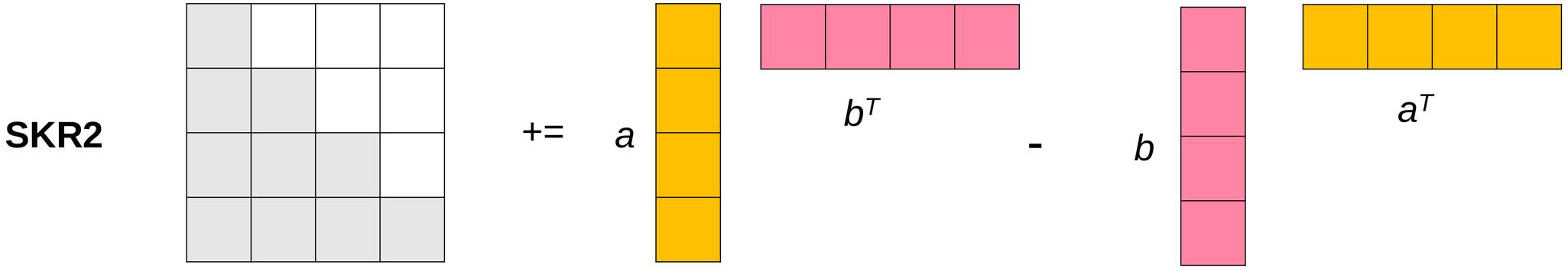
1)

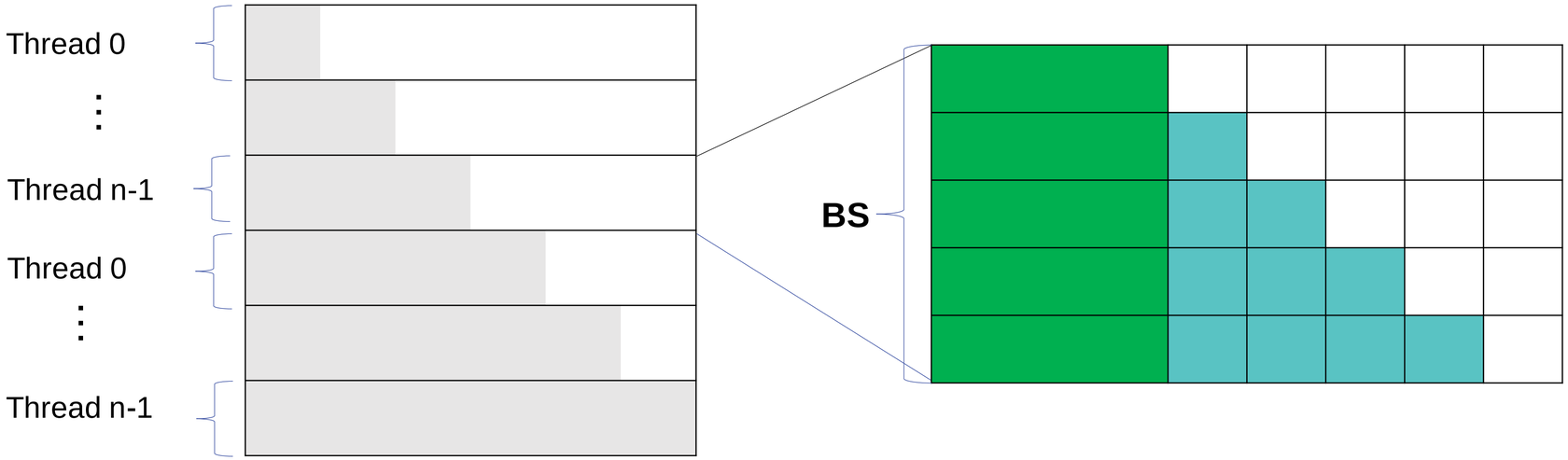


2)



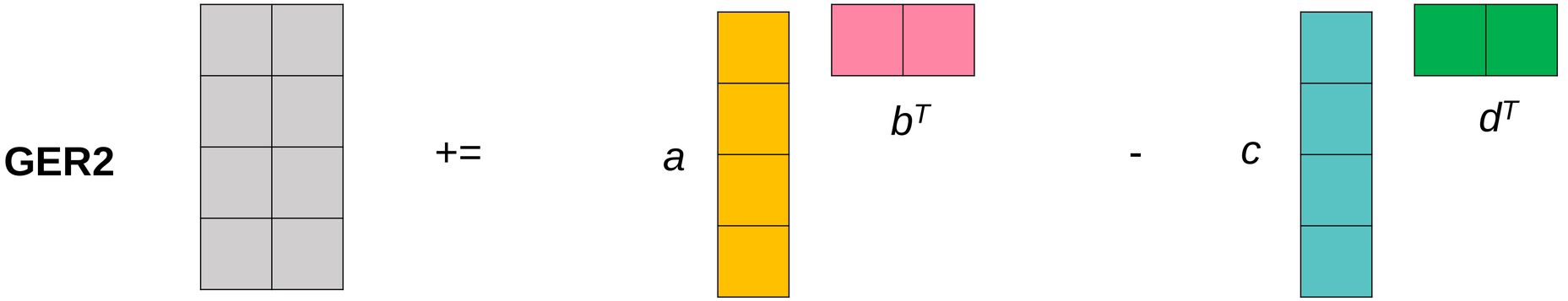
Implementation (SKR2)



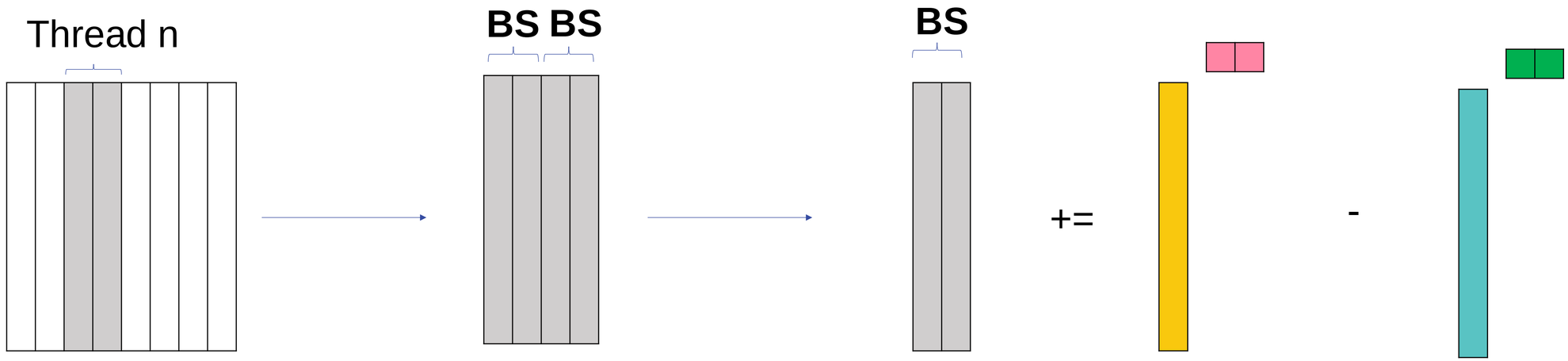


$BS_{opt} = 2$

Implementation (GER2, Column-Major)



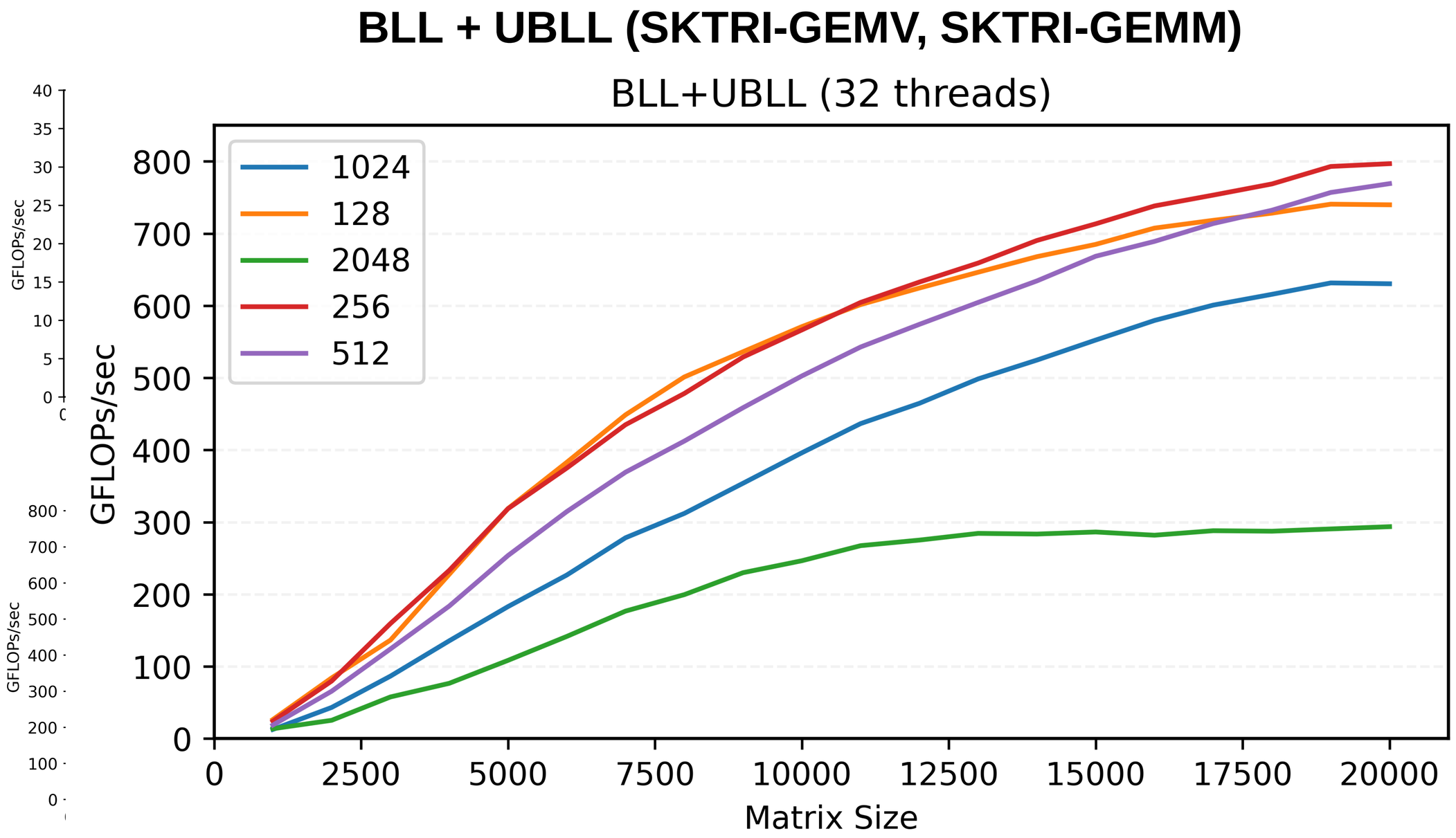
BS_{opt} = 5



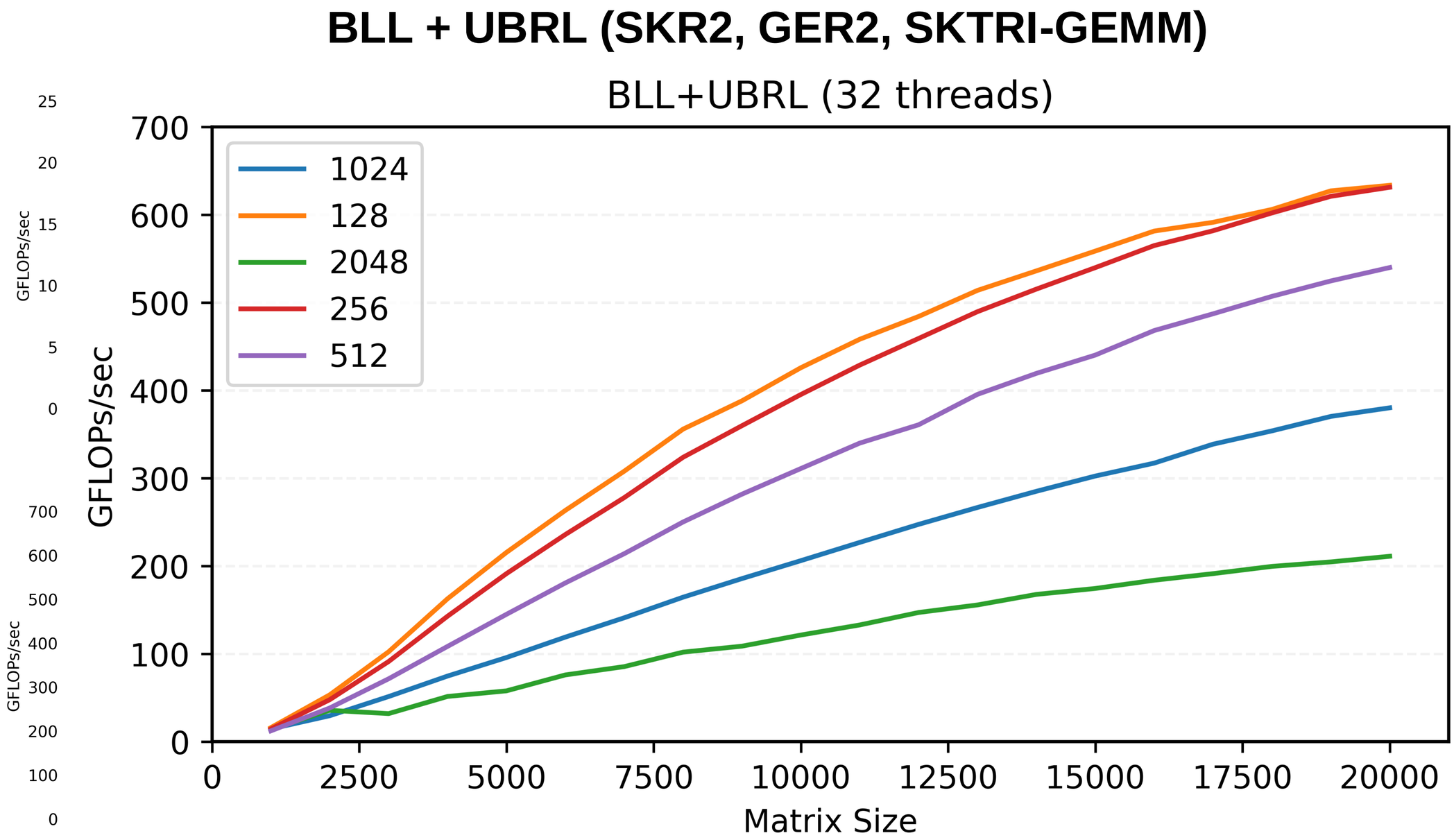
1. Parallelization

2. Unrolling

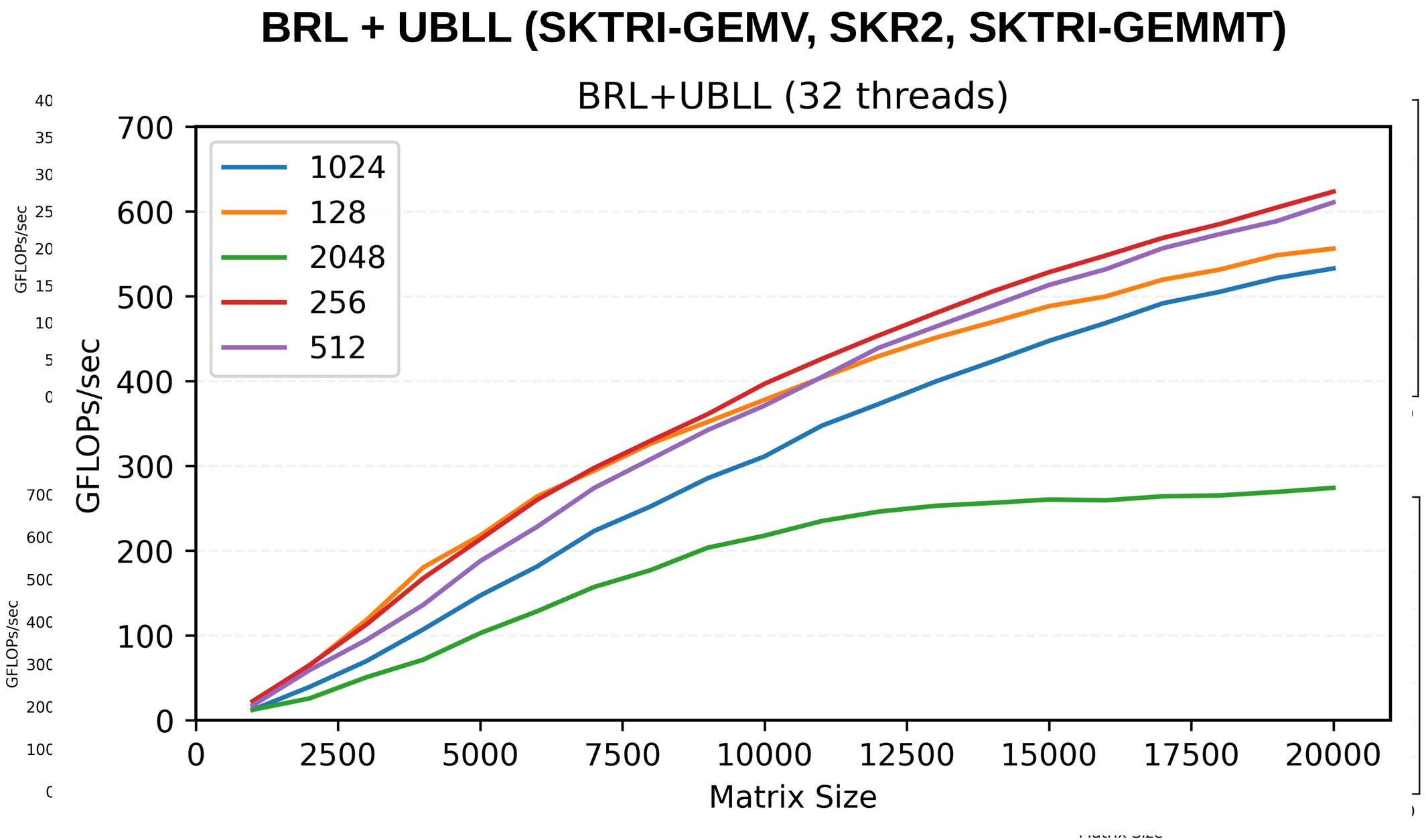
Result (Different block size)



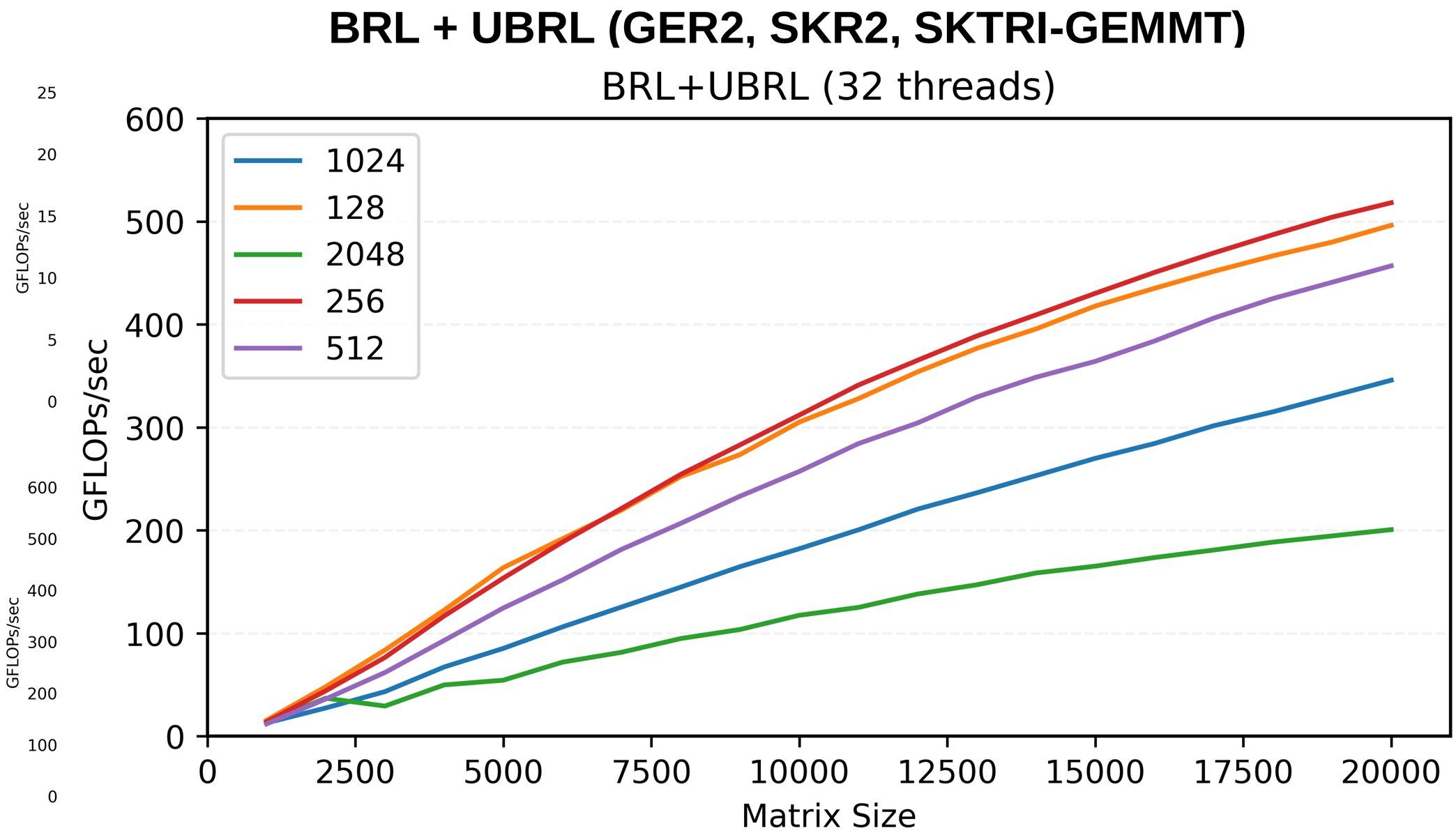
Result (Different block size)



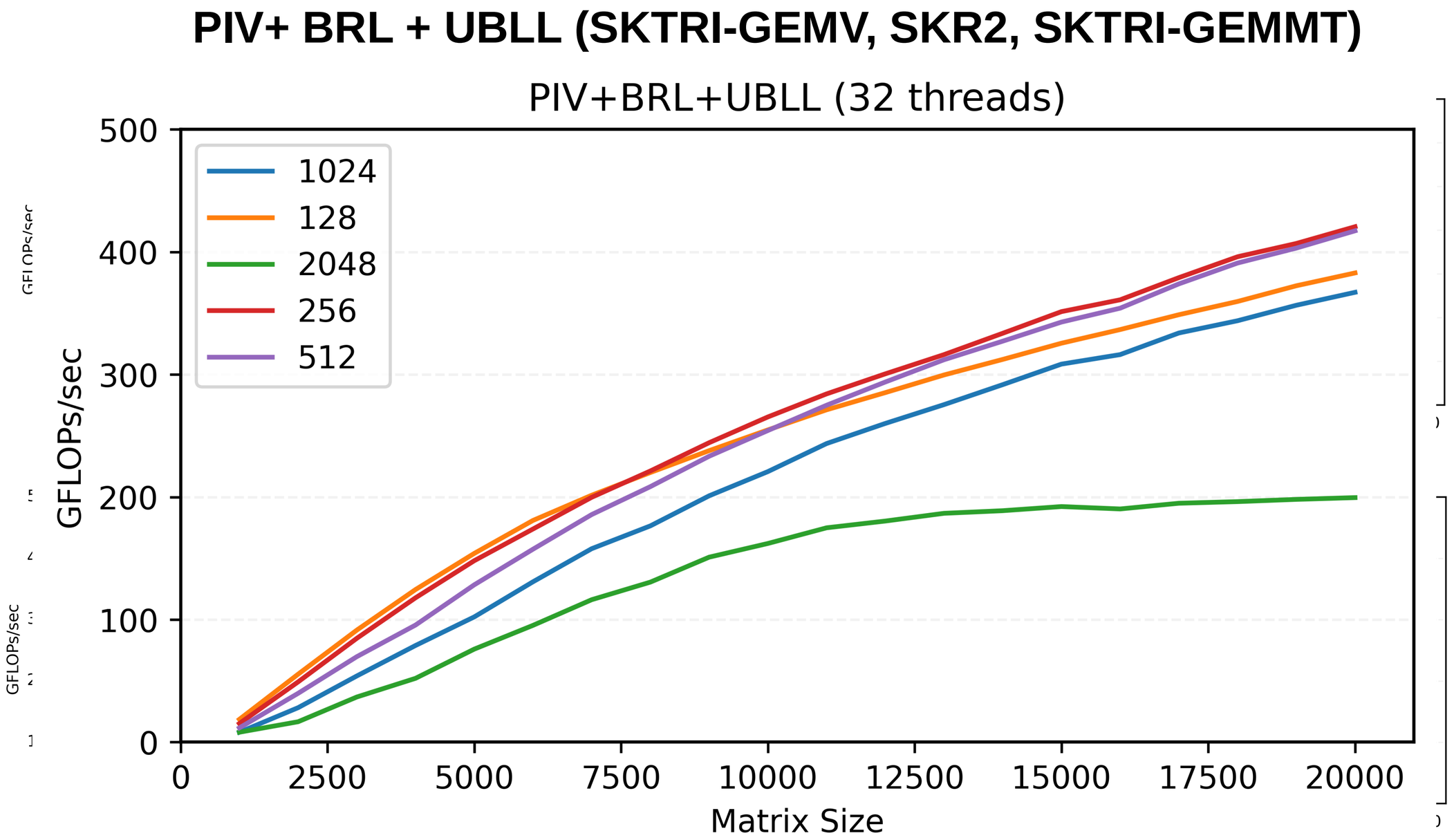
Result (Different block size)



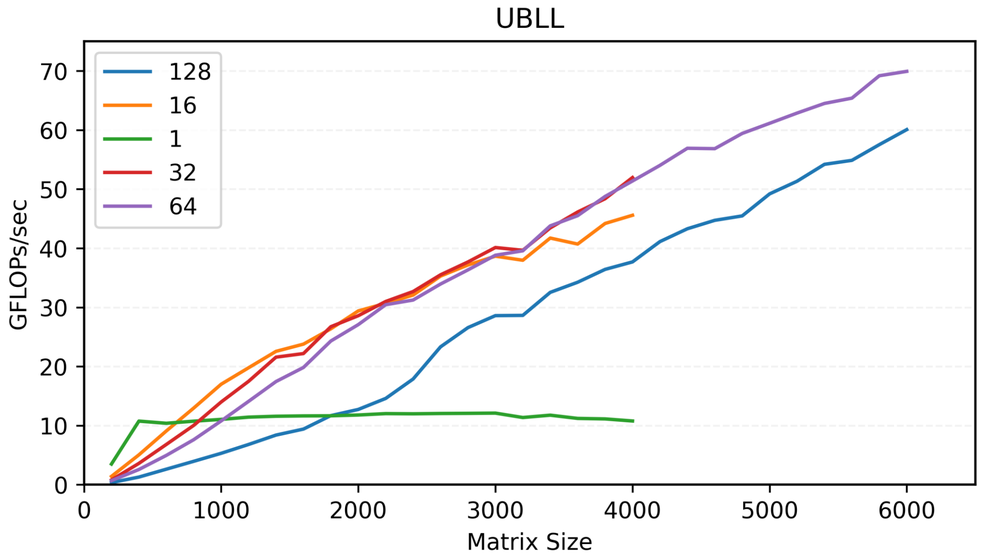
Result (Different block size)

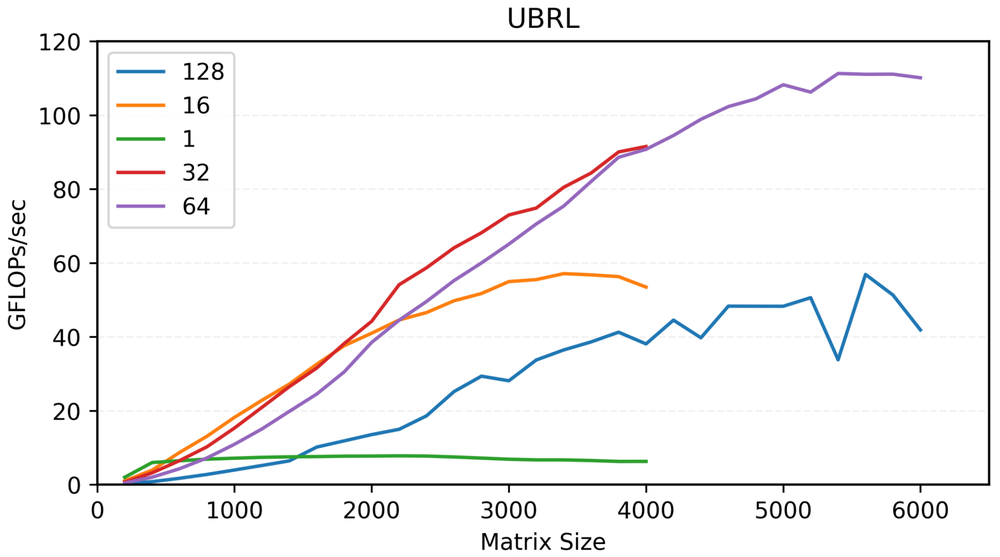


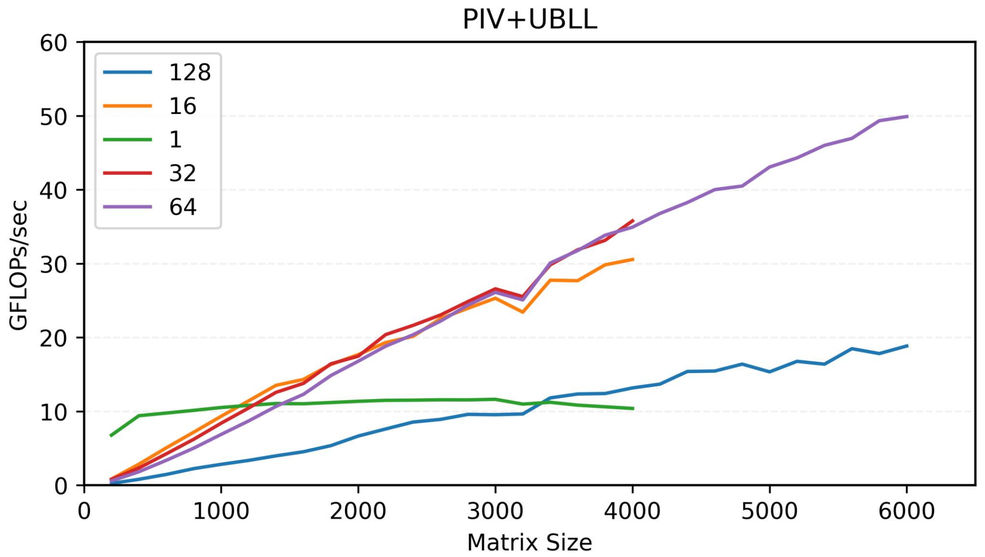
Result (Different block size)

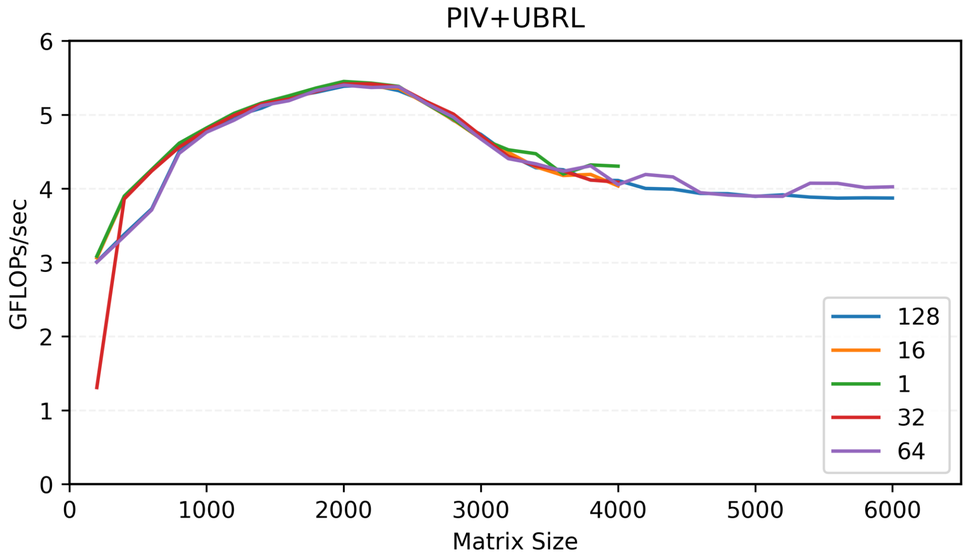


Result (Unblocked Algorithms with multi-threading)



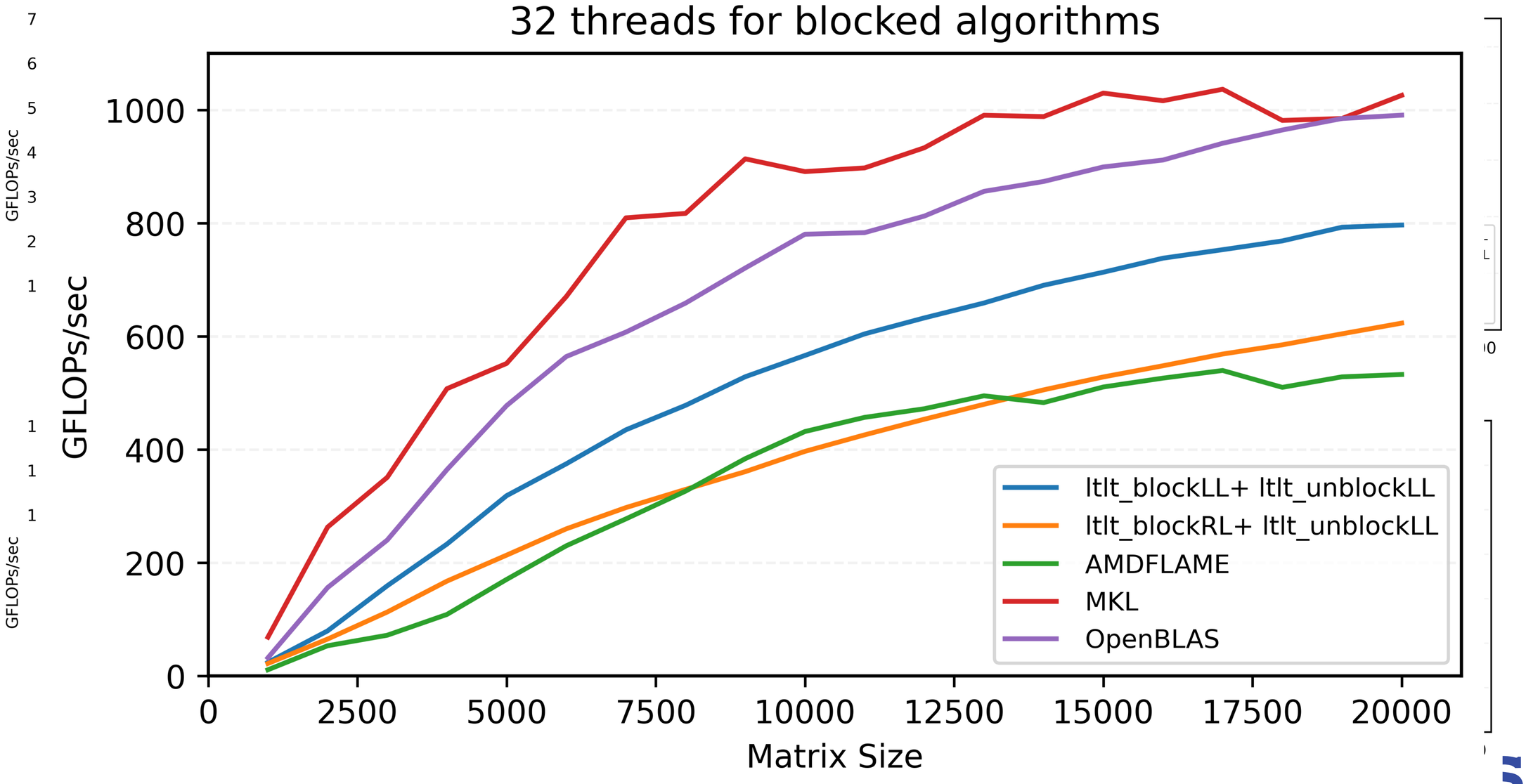






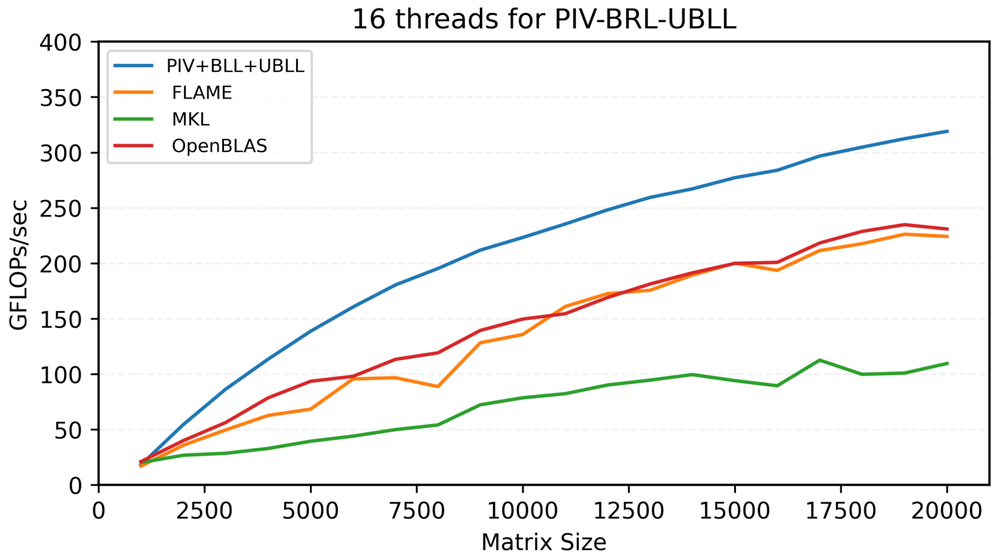
Comparison with other vendors (POTRF)

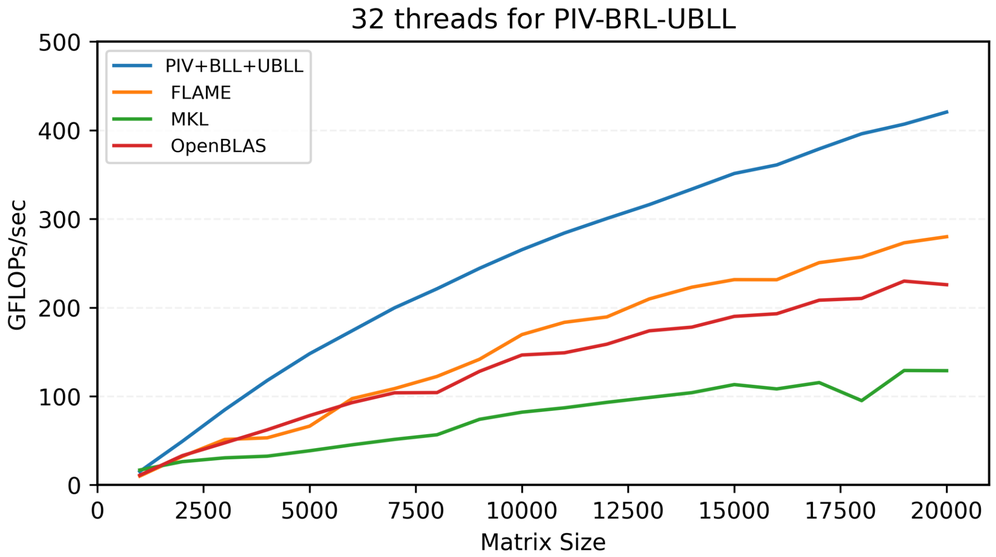
POTRF : computes the Cholesky factorization of a real symmetric positive definite matrix A



Comparison with other vendors (PSTRF)

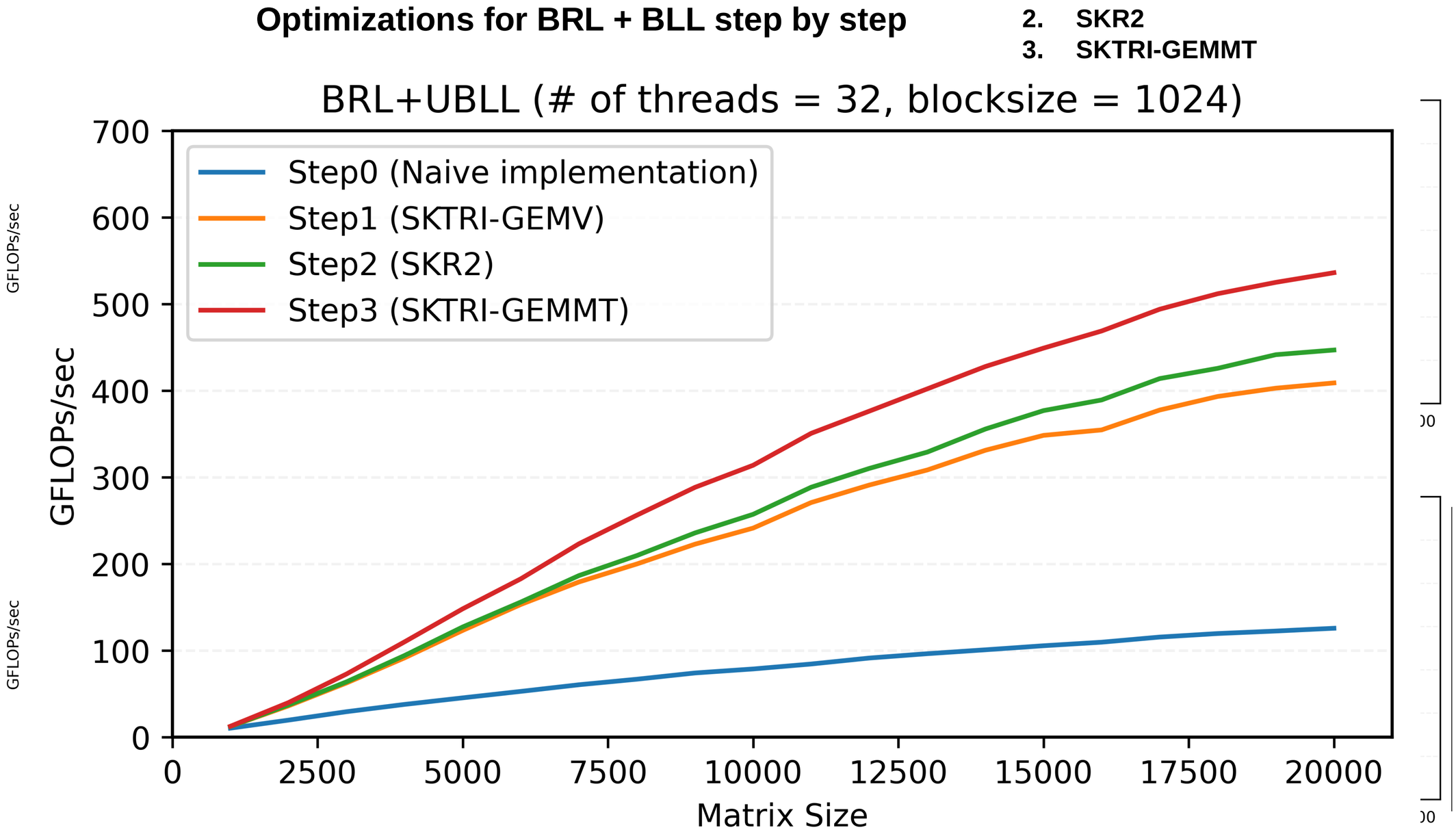
PSTRF : computes the Cholesky factorization with complete pivoting of a real symmetric positive semidefinite matrix





Optimization step by step

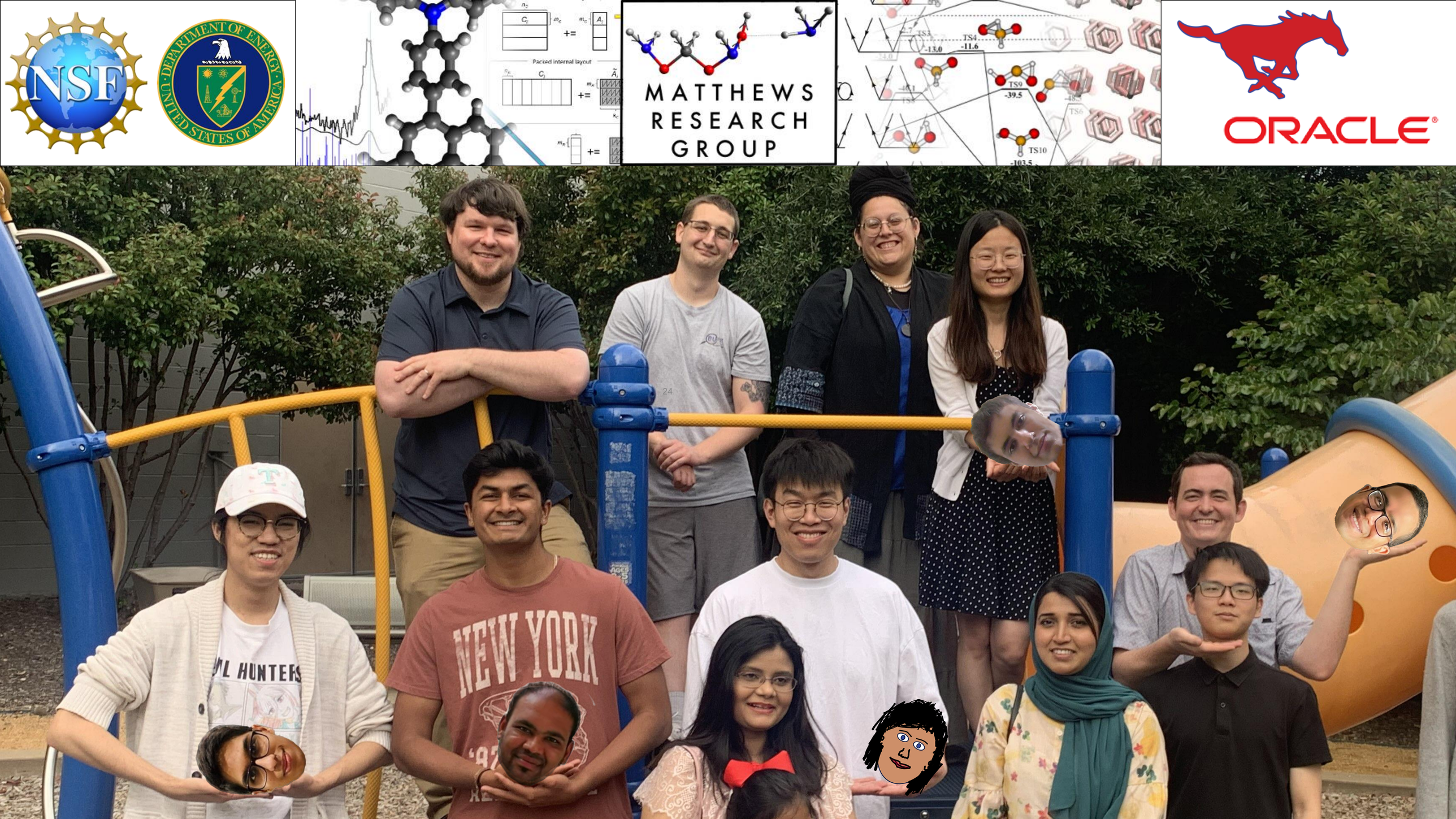
1. SKTRI-GEMV
2. SKR2
3. SKTRI-GEMMT



Summary

- By using FLAME methodology, we could systematically derive algorithms for LTLT decomposition of a Skew- symmetrical matrix
- Some new operations, which are not included in current BLAS, may probably needs to be optimized for those new algorithms to get high performance and BLIS could leverage those challenges.
- In this work, we optimize SKR2, GER2 and SKTRI-GEMV Level-2 operations, and two level-3 operations (SKTRI-GEMM, SKTRI-GEMMT)
- We profile the performance for 5 blocked algorithms and 4 unblocked algorithms including pivoting and unpivoting (BLL + UBLL is the best)
- BLIS-2.0 gives us opportunities to optimize those Level-2 operations.

SMU


MATTHEWS
RESEARCH
GROUP
ORACLE
NEW YORK

ADVERTISEMENT

I’m looking for a job!