A Case for Malleable Thread-Level Linear Algebra Libraries: The LU Factorization with Partial Pivoting

Sandra Catalán, José R. Herrero, Enrique S. Quintana-Ortí,

Rafael Rodríguez-Sánchez, Robert van de Geijn

The LU factorization (Right-looking variant)

Algorithm: $[A] := \text{LU\_BLK}(A)$

$$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$$

where A_{TL} is 0×0

while $n(A_{TL}) < n(A)$ do

Determine block size b

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

where A_{11} is $b \times b$

RL1. $\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} := \text{LU\_UNB}\left(\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}\right)$

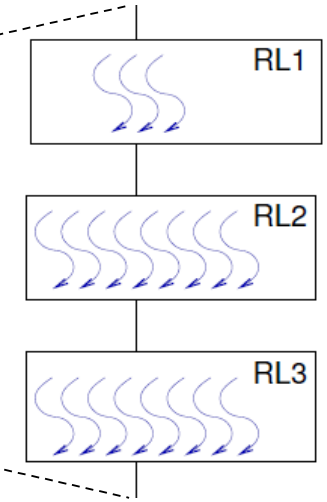
RL2. $A_{12} := \text{TRILU}(A_{11})^{-1} A_{12}$

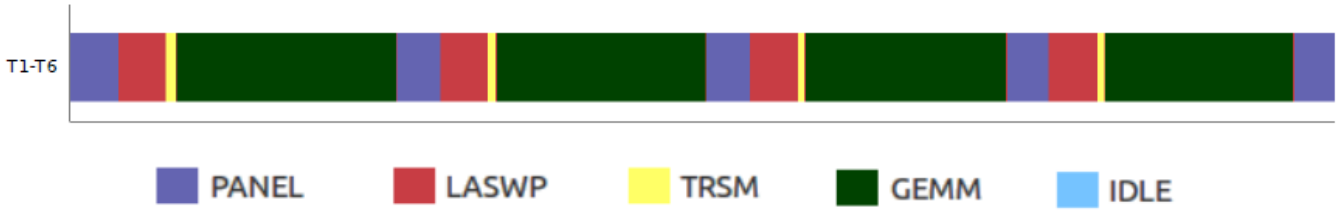
RL3. $A_{22} := A_{22} - A_{21} A_{12}$

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

endwhile

Calls to multi-threaded BLAS





Look-ahead to overlap RL1 with remaining kernels

Algorithm: $[A] := \text{LU\_LA\_BLK}(A)$

Determine block size b

$$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right), \quad A_{BR} \rightarrow \left(\begin{array}{c|c} A_{BR}^P & A_{BR}^R \end{array}\right)$$

where A_{TL} is 0×0, A_{BR}^P has b columns

$A_{BR}^P := \text{LU\_UNB}\left(A_{BR}^P\right)$

while $n(A_{TL}) < n(A)$ do

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

where A_{11} is $b \times b$

Determine block size b

% Partition into panel factorization and remainder

$$\left(\begin{array}{c} A_{12} \\ \hline A_{22} \end{array}\right) \rightarrow \left(\begin{array}{c|c} A_{12}^P & A_{12}^R \\ \hline A_{22}^P & A_{22}^R \end{array}\right)$$

where both A_{12}^P, A_{22}^P have b columns

% Panel factorization, T_{PF}

PF1. $A_{12}^P := \text{TRILU}(A_{11})^{-1} A_{12}^P$

PF2. $A_{22}^P := A_{22}^P - A_{21} A_{12}^P$

PF3. $A_{22}^P := \text{LU\_UNB}\left(A_{22}^P\right)$

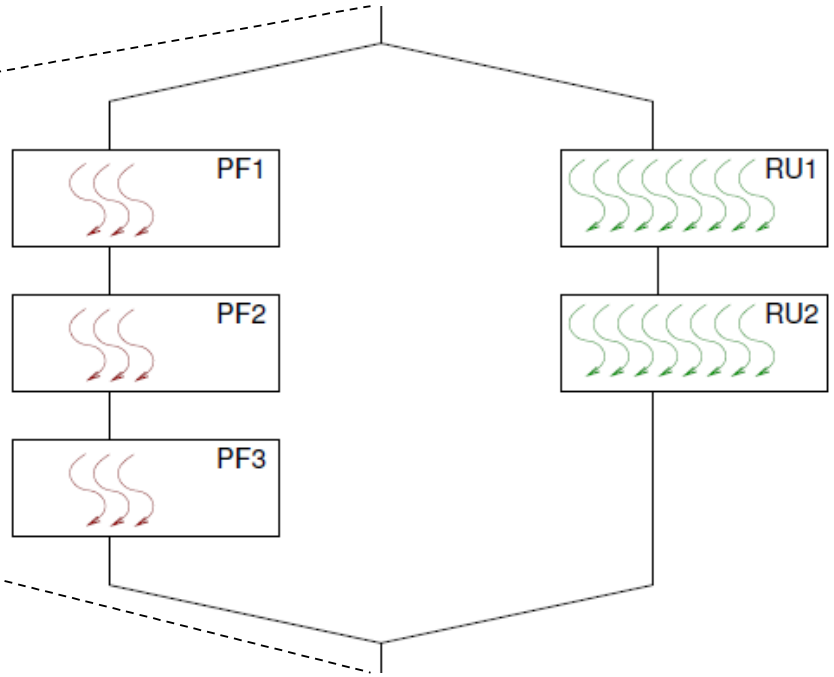
% Remainder update, T_{RU}

RU1. $A_{12}^R := \text{TRILU}(A_{11})^{-1} A_{12}^R$

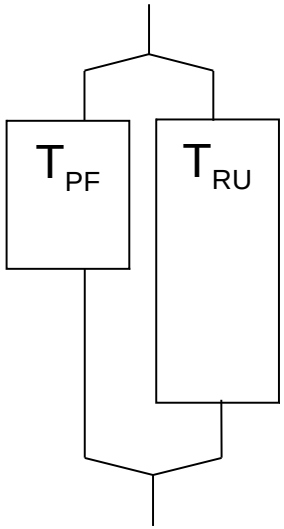
RU2. $A_{22}^R := A_{22}^R - A_{21} A_{12}^R$

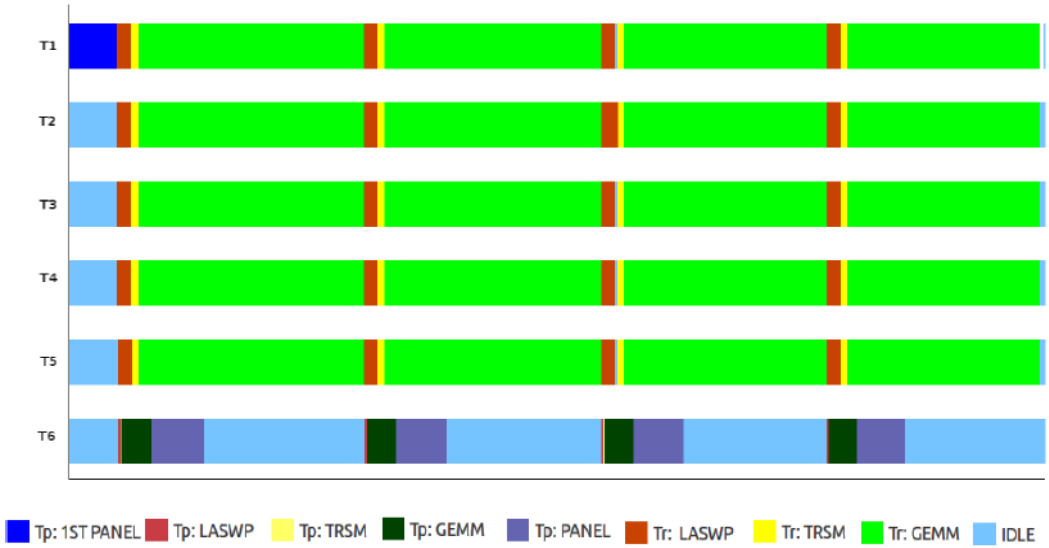
$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

endwhile

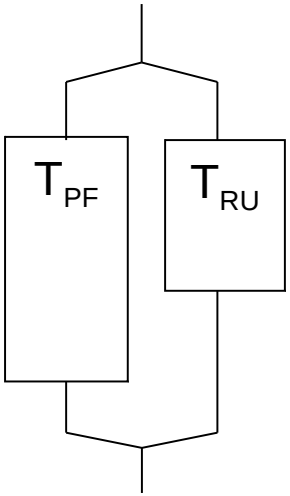


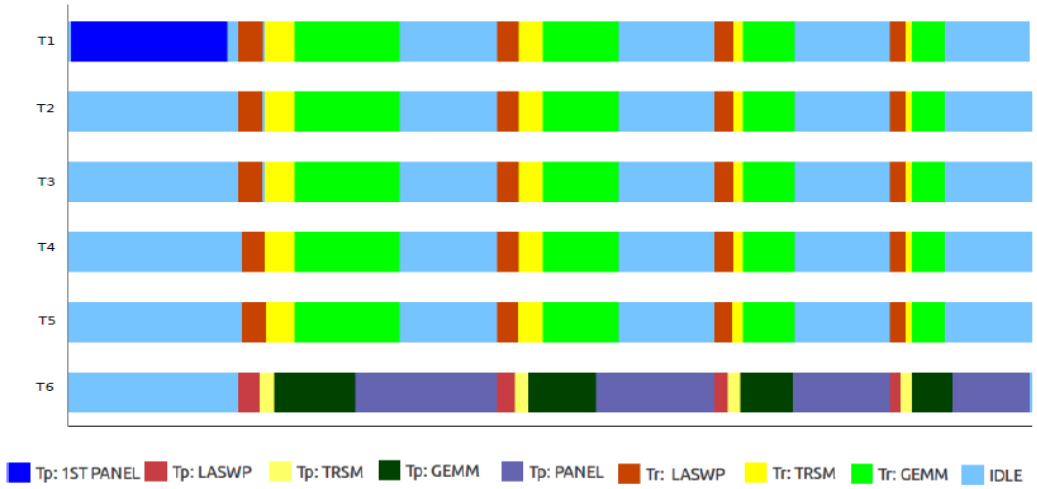
T_{RU}>T_{PF}: Malleable BLIS



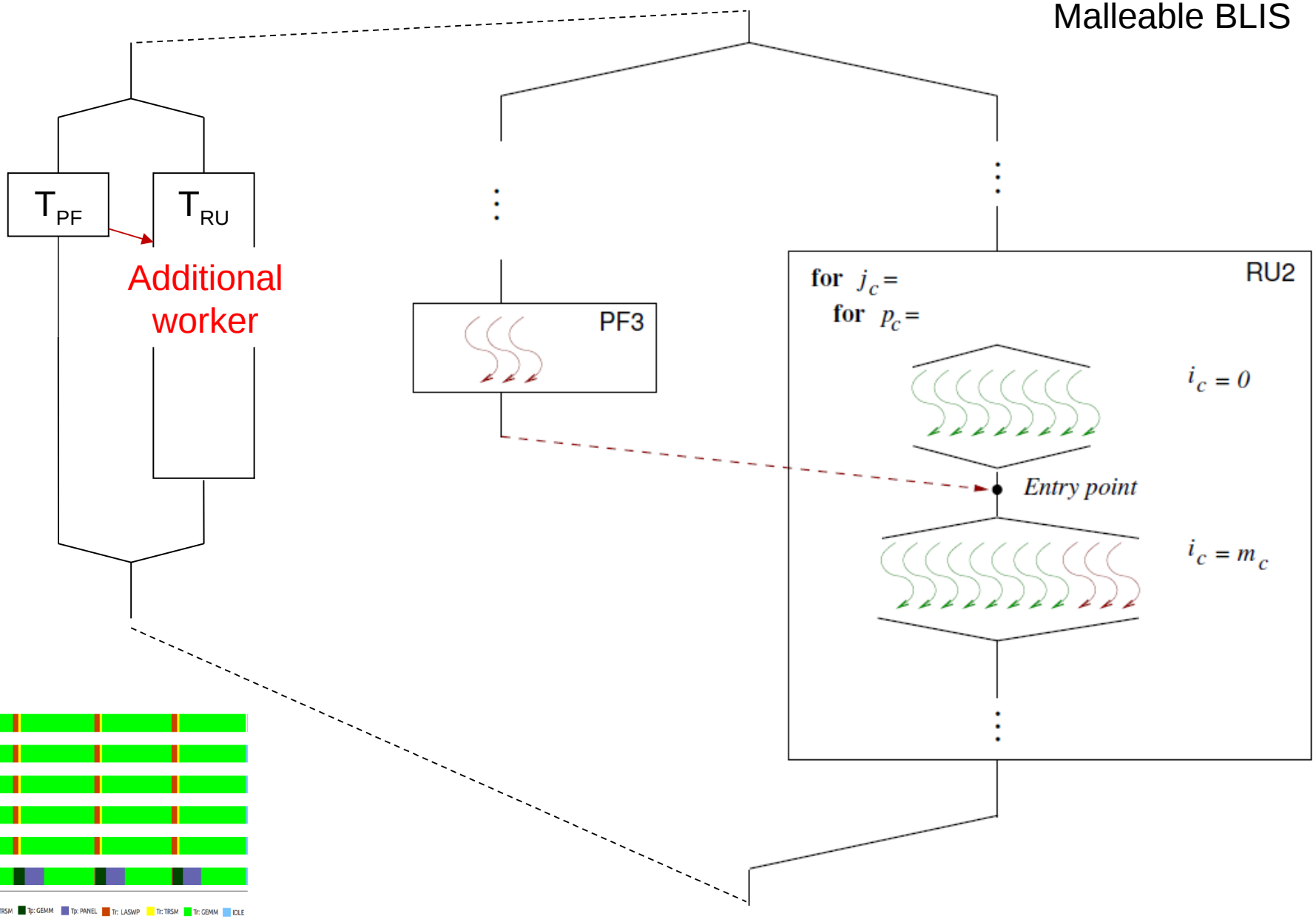


T_{PF}>T_{RU}: Early Termination (ET)

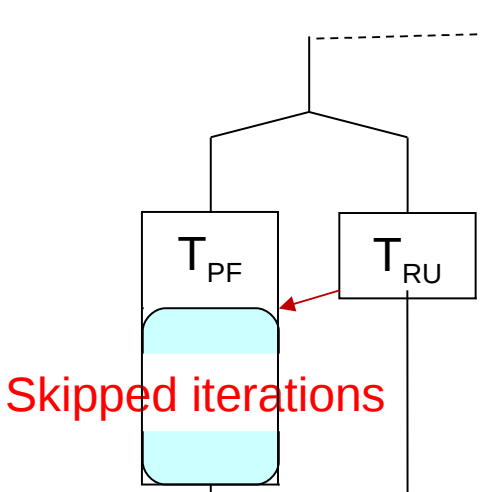


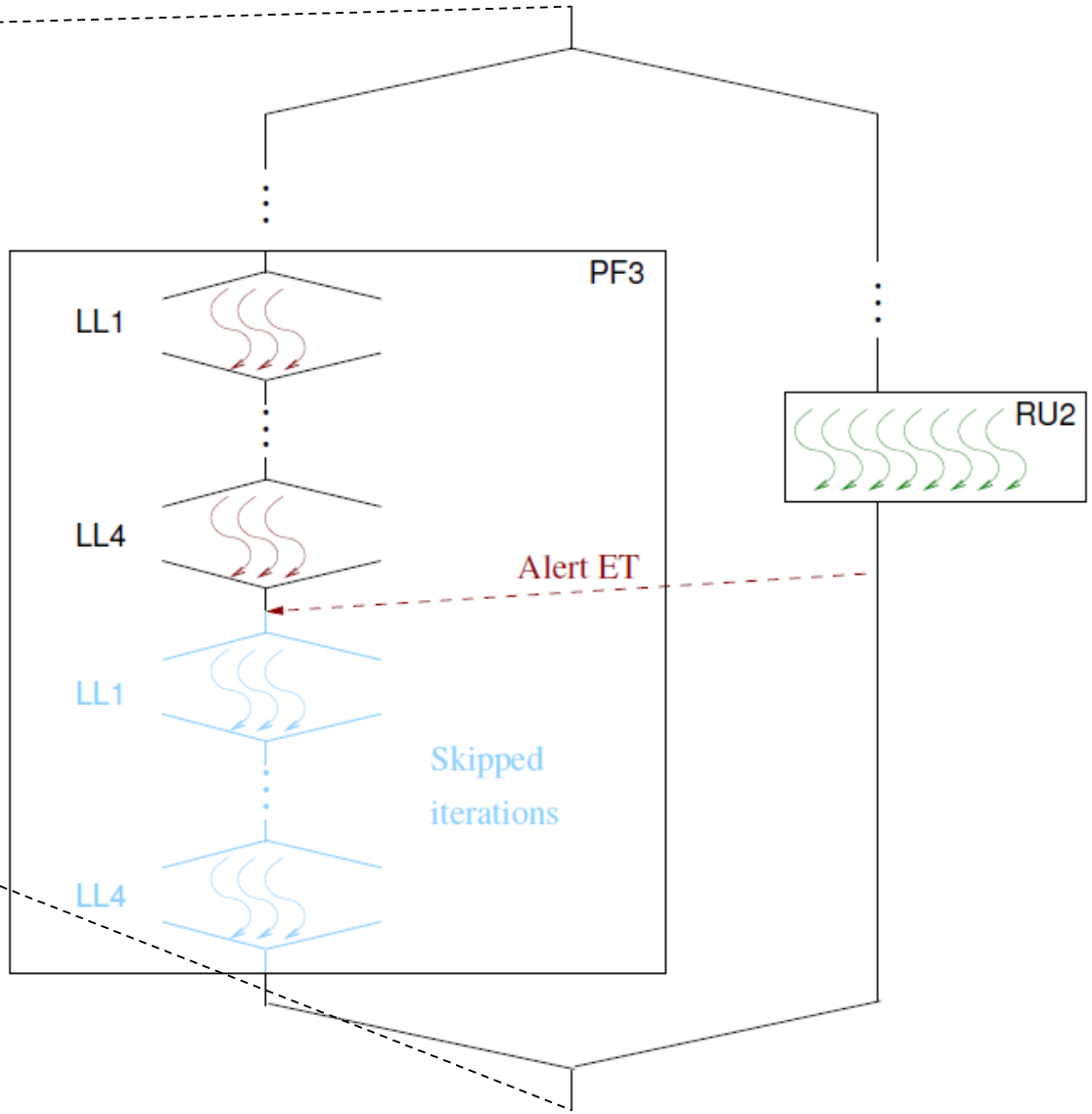


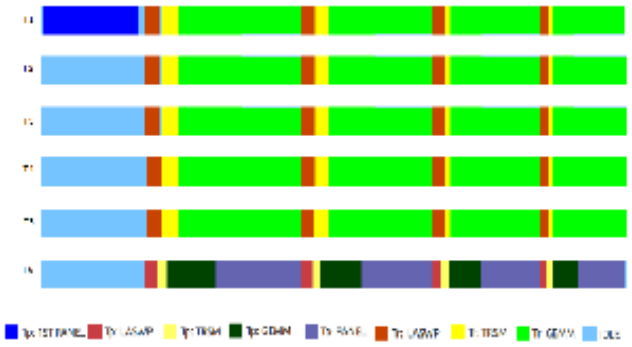
T_{RU}>T_{PF}: Malleable BLIS



T_{PF}>T_{RU}: Early Termination (ET)







Experimental evaluation