jh220009-NAHI (国際課題)
横田 理央 (東京工業大学)

Hierarchical low-rank approximation methods on distributed memory and GPUs

背景

H行列などの階層的低ランク近似法では、密行列を階層的に分割し
非対角ブロックを低ランク近似することで行列積や行列分解の
コストを$O(N^3)$から$O(N)$に低減できる

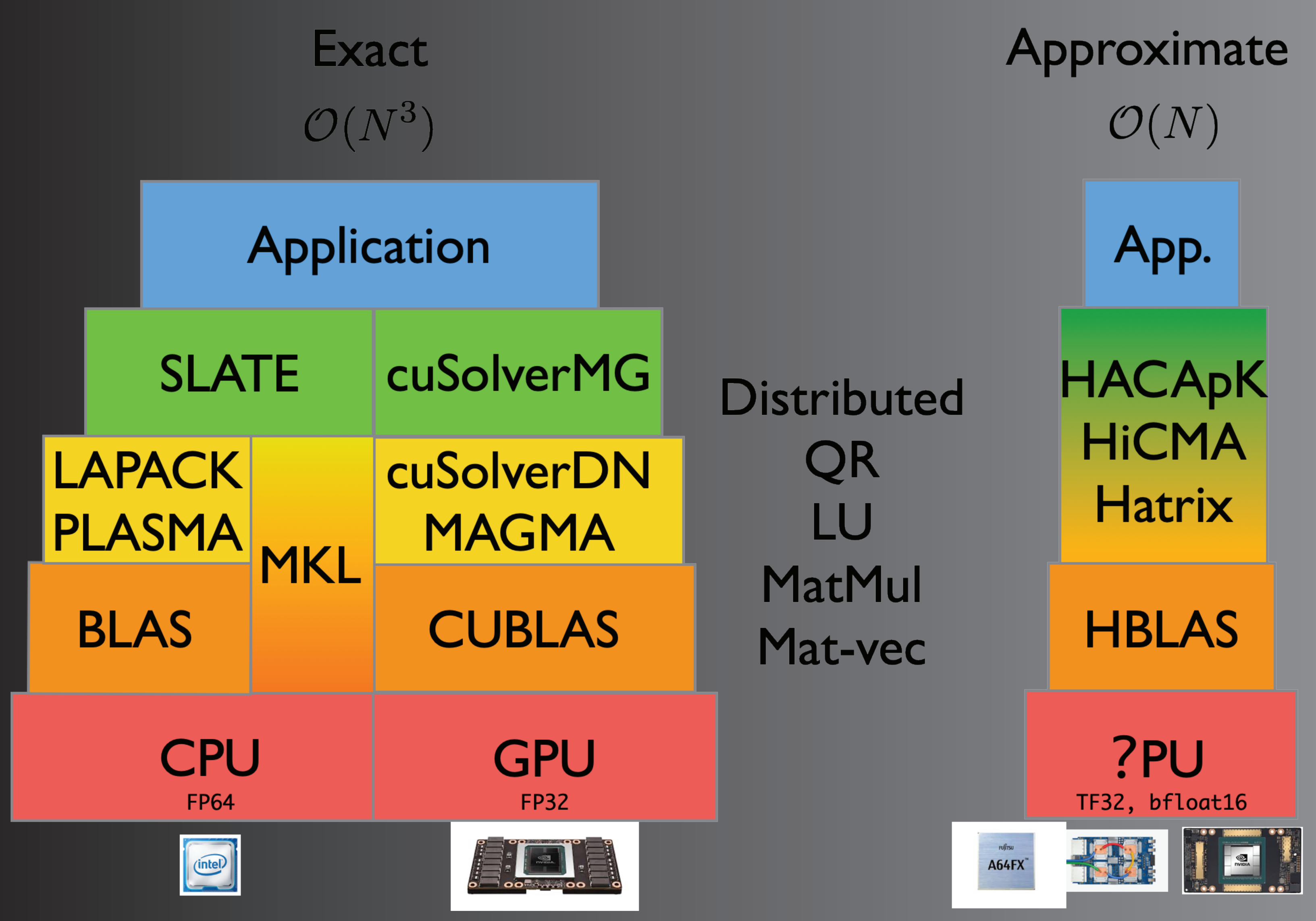
境界要素法の反復解法における前処理行列、疎行列のマルチフロンタル法
による解法におけるフロンタル行列、統計学における分散共分散行列の
行列式などの計算に現れる密行列はH行列で表すことができる

H行列はこれまで応用数学の分野の研究者が中心となって開発しており
高性能実装、並列化、GPU化などがほとんど行われていない。

H行列の内部では小さな行列積の計算が無数に行われるため、
TensorCoreなどを用いたBatched GEMMを活かすことができる

低精度演算を用いるとき、密行列を厳密に計算することは無駄であり、
H行列を用いれば低精度な行列演算は全て$O(N)$で行うことができる

線形代数ライブラリの置き換え



BLASやLAPACKは$O(N^3)$の関数からなるが、これらをH行列を用いて
$O(N)$の関数に置き換えることを目的とする

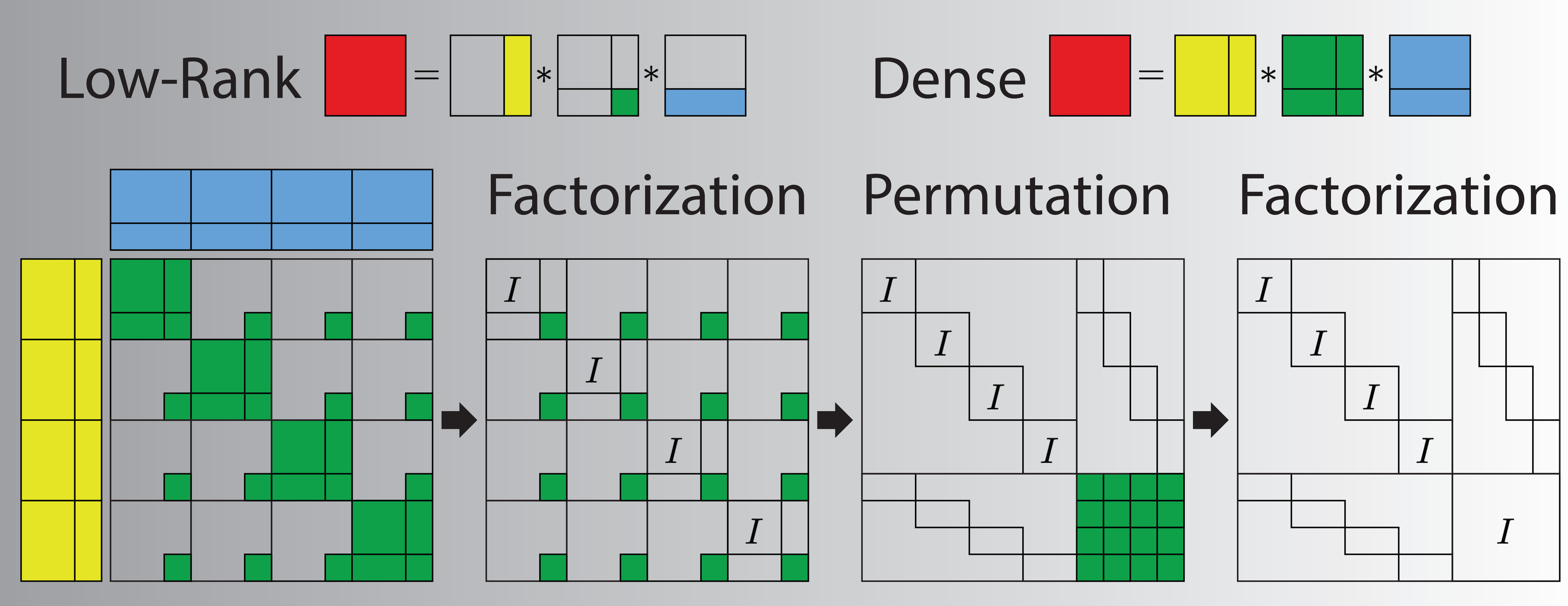
既存研究

Author/year	Matrix	Op	Type	LRA	OpenMP	MPI	GPU
2016Vogel	Dense	$Q\Lambda Q^T$	HSS	?	×	×	×
2017Akbudak	Dense	Cholesky	BLR	RSVD	○	×	×
2017Fernando	Dense	LU	HSS	ID	×	○	○
2017Ghysels	Sparse	LU	HSS	RSVD	○	○	×
2017Li	Sparse	LU	HSS	RRQR	×	○	×
2017Minden	Sparse	LU	H2	RRQR	×	×	×
2018Amestoy	Sparse	LU	MBLR	?	×	×	×
2018Börm	Dense	O(N)Cmp	H2	GCA	×	×	×
2018Cai	Dense	O(N)Cmp	HSS	RRQR	×	×	×
2018Kressner	Dense	QR	HODLR	?	×	×	×
2018Yu	Dense	O(N)Cmp	HSS	ID	○	○	×
2019Amestoy	Sparse	LU	BLR	?	○	○	×
2019Boukaram	Dense	MV	H2	RSVD	×	×	○
2019Cambier	Sparse	LU	H2	RRQR	×	×	×
2019Zaspel	Dense	MV	H	ACA	×	×	○

これまでのH行列の研究は応用数学分野の研究が多く
高性能計算分野におけるOpenMP, MPI, CUDA実装などが少ない

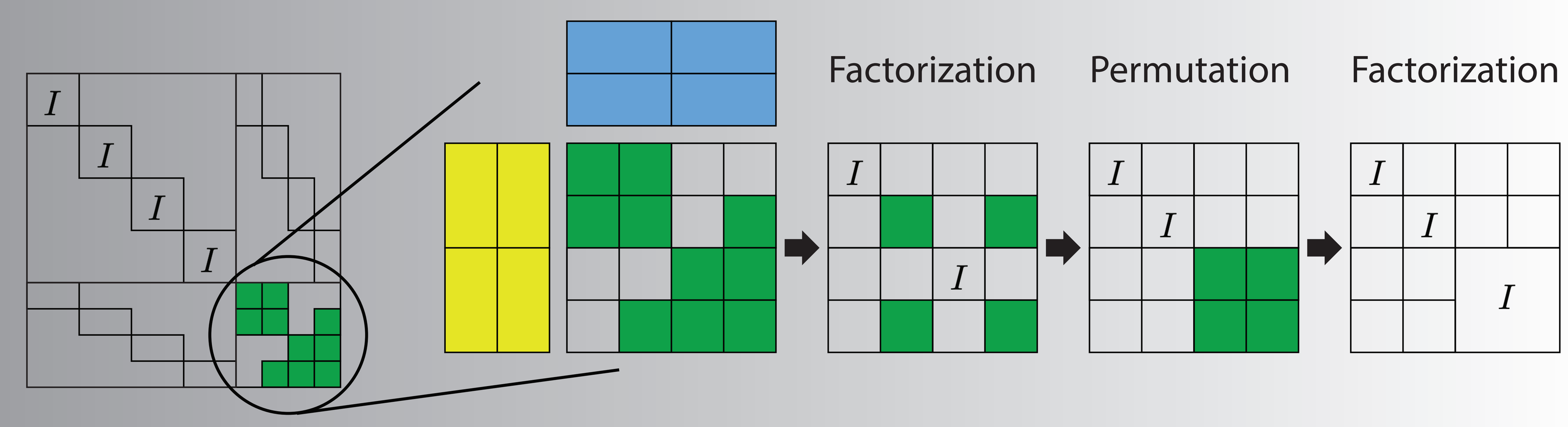
BLR²-ULV分解

$$A = [U^R\ U^S]\begin{bmatrix}0 & 0\\ 0 & S^{SS}\end{bmatrix}\begin{bmatrix}V^R\\ V^S\end{bmatrix} \qquad A = [U^R\ U^S]\begin{bmatrix}S^{RR} & S^{RS}\\ S^{SR} & S^{SS}\end{bmatrix}\begin{bmatrix}V^R\\ V^S\end{bmatrix}$$



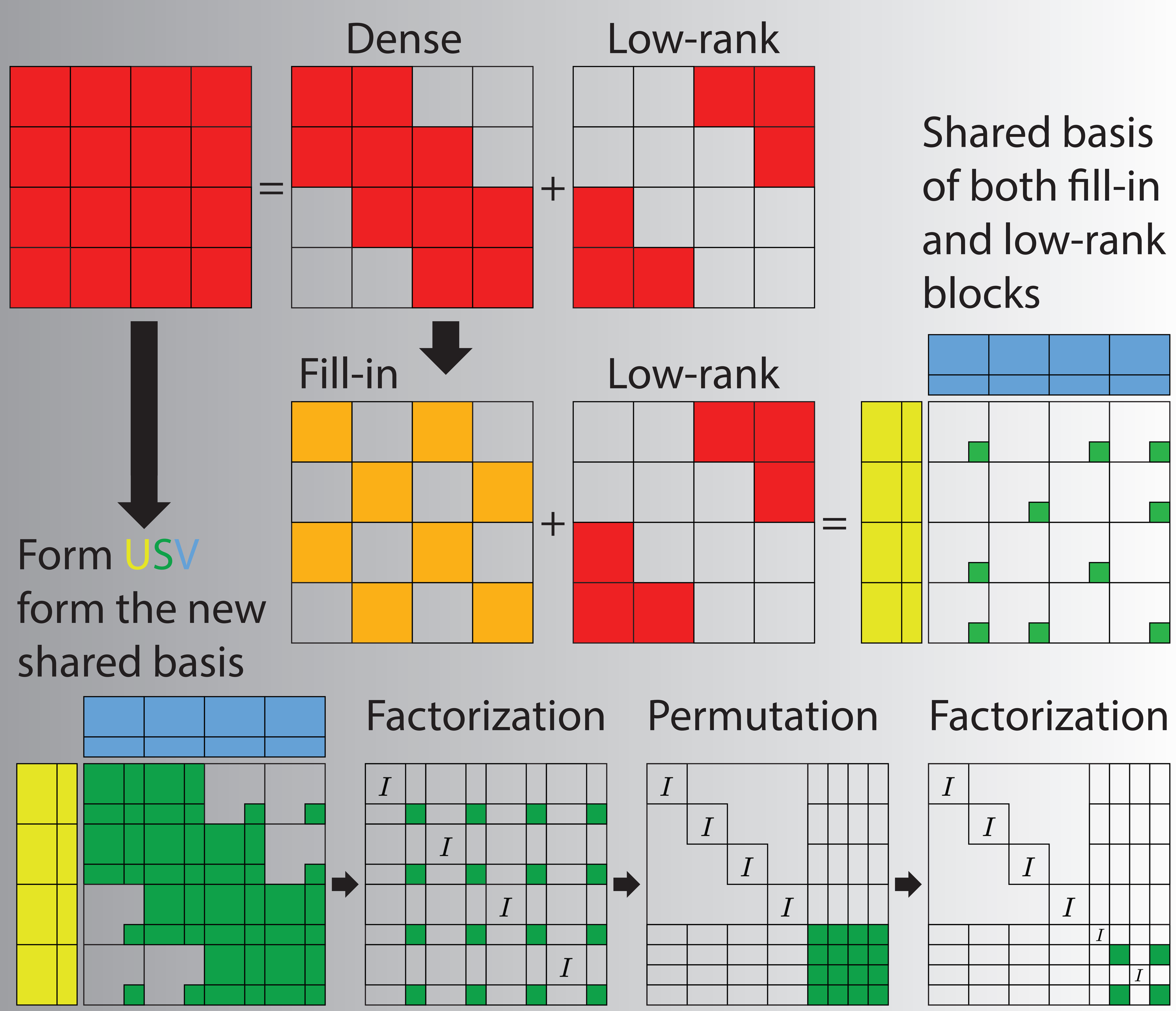
ULV分解を用いるとLU分解における対角ブロックの依存関係が
なくなり、並列に計算できるようになる

HSS-ULV分解



これを各階層で繰り返すことで密行列のLU分解を$O(N)$で実現
できる。ただし、HSSでは対角ブロックのみが密行列で表され
３次元の問題などでは非対角ブロックのランクがNに比例して
増加するため$O(N)$ではなくなってしまう。

H²-ULV分解



H²行列では非対角ブロックにも密行列が現れるため３次元の
問題などでも$O(N)$で解くことができる。ただし、fill-inが
起きるため、そのままでは対角ブロックの依存関係が再び
発生してしまう。本研究では、fill-inをULV分解の基底に
含めることにより、依存関係をなくすことができる。