

Résumé

Enjeu : évaluation rapide d'un opérateur linéaire.
Factorisation butterfly : pour la compression de matrices.

$$\mathbf{A} \approx \begin{matrix} \text{[Matrix 1]} \\ \times \\ \text{[Matrix 2]} \\ \times \\ \text{[Matrix 3]} \\ \times \\ \text{[Matrix 4]} \end{matrix}$$

Flexibilité : approcher \mathbf{A} à permutation près des lignes et colonnes.
Problème : comment approcher \mathbf{A} par $\hat{\mathbf{A}} := \mathbf{Q}^\top \mathbf{X}^{(1)} \dots \mathbf{X}^{(L)} \mathbf{P}$?

Problème d'approximation butterfly

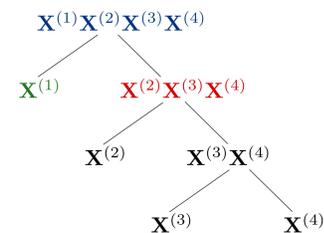
Pour $\mathbf{A} \in \mathbb{C}^{N \times N}$ avec $N = 2^L$, en notant $\mathbf{S}_{\text{bf}}^{(\ell)} := \mathbf{I}_{2^{\ell-1}} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \otimes \mathbf{I}_{N/2^\ell}$,

$$\min_{(\mathbf{X}^{(\ell)})_{\ell=1}^L, \mathbf{P}, \mathbf{Q}} \|\mathbf{A} - \mathbf{Q}^\top \mathbf{X}^{(1)} \dots \mathbf{X}^{(L)} \mathbf{P}\|_F, \text{ supp}(\mathbf{X}^{(\ell)}) \subseteq \text{supp}(\mathbf{S}_{\text{bf}}^{(\ell)}).$$

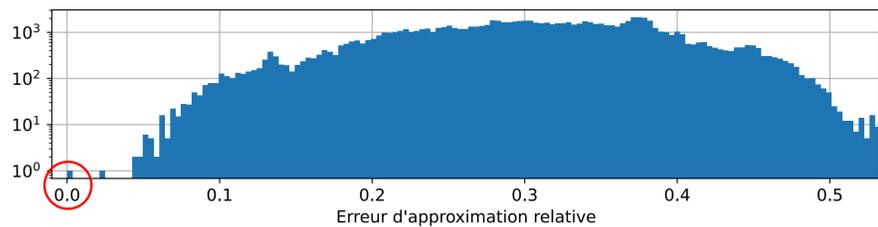
Exemple : factorisation *exacte* pour les matrices Hadamard, DFT [1].

P, Q connus : facile.

Il existe un algorithme de factorisation hiérarchique, en complexité $\mathcal{O}(N^2)$, avec garanties de reconstruction [2, 3].



P, Q inconnus : difficile. Il faut trouver de bonnes permutations.



Histogramme : approximation de $\mathbf{A} := \tilde{\mathbf{Q}}^\top \tilde{\mathbf{A}} \tilde{\mathbf{P}}$ par énumération de \mathbf{P}, \mathbf{Q}

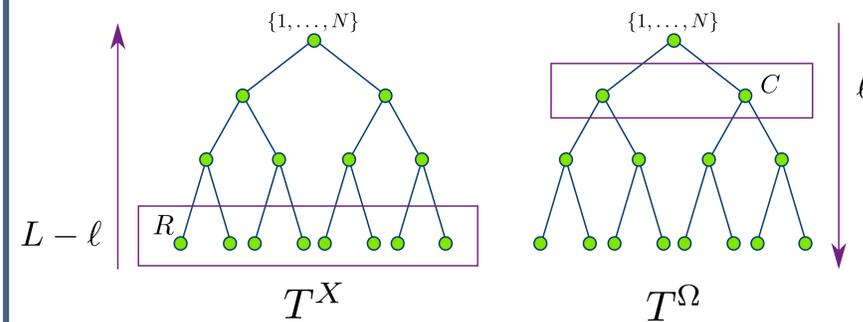
Approche : identification de blocs de rang un

$\mathbf{A} := \tilde{\mathbf{Q}}^\top \tilde{\mathbf{A}} \tilde{\mathbf{P}} \implies \mathbf{A}$ est de « rang faible complémentaire » [4, 5, 6] pour des arbres *inconnus* T^X, T^Ω .

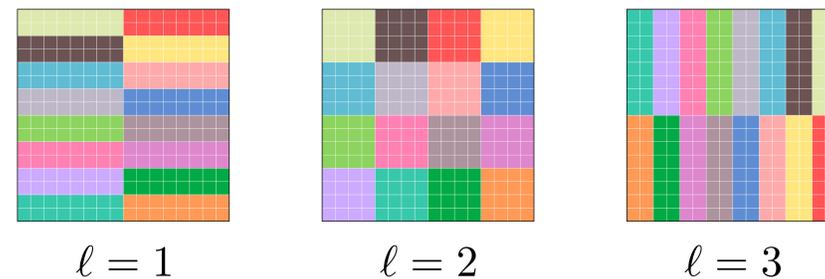
1. Identifier les arbres T^X et T^Ω .
2. Fixer \mathbf{P}, \mathbf{Q} telles que \mathbf{QAP}^\top soit de rang faible complémentaire pour les arbres $T_{\text{bf}}^X, T_{\text{bf}}^\Omega$.
3. Appliquer l'algorithme hiérarchique connaissant \mathbf{P}, \mathbf{Q} .

Propriété de rang faible des matrices butterfly

Théorème : $\mathbf{A} = \mathbf{X}^{(1)} \mathbf{X}^{(2)} \dots \mathbf{X}^{(L)}$ où $\text{supp}(\mathbf{X}^{(\ell)}) \subseteq \text{supp}(\mathbf{S}_{\text{bf}}^{(\ell)})$ ssi \mathbf{A} est de rang faible complémentaire pour des arbres connus $T_{\text{bf}}^X, T_{\text{bf}}^\Omega$.



Définition : rang faible complémentaire (T^X, T^Ω) [6]. À chaque niveau $\ell \in [L-1]$, $\text{rang}(\mathbf{A}_{R,C}) \leq 1$ pour chaque nœud R du niveau $L-\ell$ de T^X , et pour chaque nœud C du niveau ℓ dans T^Ω .



Reconstruction des arbres de partitionnement

Approche : pour chaque $\ell \in [L-1]$, reconstruire le $(L-\ell)$ -ème niveau de T^X et le ℓ -ème niveau de T^Ω via

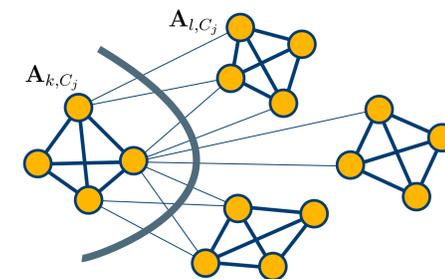
$$\min_{\{R_i\}_i, \{C_j\}_j} \sum_{i,j} \min_{\mathbf{x}, \mathbf{y}} \|\mathbf{A}_{R_i, C_j} - \mathbf{x}\mathbf{y}^*\|_F^2.$$

À l'issue, recoller les niveaux si possible.

Partitionnement spectral [7] : supposons la partition $\{C_j\}_j$ des colonnes connue. Pour chaque groupe C_j , définissons un graphe \mathcal{G}_j .

Corrélation entre les lignes restreintes à C_j :

$$\mathbf{W}_{k,l}^{(j)} := \left(\frac{|\mathbf{A}_{k,C_j}^* \mathbf{A}_{l,C_j}|}{\|\mathbf{A}_{k,C_j}\|_2 \|\mathbf{A}_{l,C_j}\|_2} \right)^\alpha.$$



Puis, partitionnement spectral des lignes via $\mathbf{W} := \sum_j \mathbf{W}^{(j)}$.

Optimisation alternée : minimisation selon la partition des lignes en fixant la partition des colonnes, puis vice-versa.

Validation numérique

Matrice cible : $\mathbf{A} := \tilde{\mathbf{Q}}^\top \tilde{\mathbf{A}} \tilde{\mathbf{P}} + \varepsilon(\|\tilde{\mathbf{A}}\|_F / \|\mathbf{N}\|_F) \mathbf{N}$

- matrice de la transformée de Fourier discrète
- matrice butterfly orthogonale aléatoire

Résultats : 100 % de taux de réussite pour $N \leq 64$. Pour $N = 128$:

Niveau de bruit ε	0	0.01	0.03	0.1
Butterfly orthogonale aléatoire ($N = 128$)	100 %	100 %	100 %	100 %
Transformée de Fourier discrète ($N = 128$)	100 %	95 %	90 %	50 %

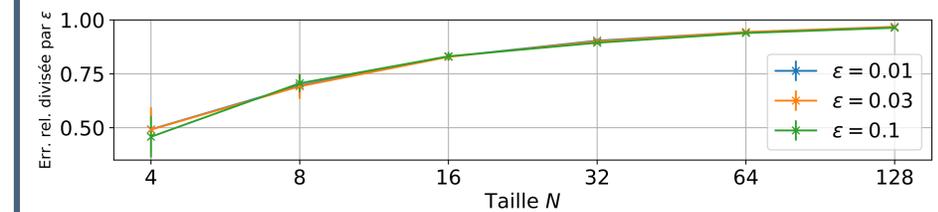


Figure : Erreur relative divisée par le niveau de bruit ε après reconstruction des arbres T^X et T^Ω (matrice butterfly orthogonale aléatoire).

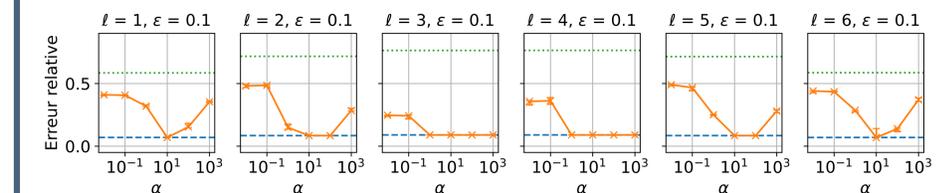


Figure : Erreur relative du partitionnement spectral alterné (matrice butterfly orthogonale aléatoire bruitée, avec niveau de bruit ε , taille $N = 128$). Bleu, tiret : erreur avec T^X, T^Ω connus. Vert, pointillé : erreur minimale sur 1000 tirages aléatoires de partitions.

Perspectives

Conclusion : heuristique pour identifier les partitions d'une matrice en blocs de rang faible permettant une factorisation butterfly.

Limites : complexité $\mathcal{O}(N^3)$, manque de robustesse pour $N \geq 128$.

Applications : vérifier algorithmiquement qu'un opérateur possède une propriété de rang faible complémentaire donnant lieu à une factorisation butterfly (e.g. transformée de Fourier sur graphe).

Références

- [1] T. Dao, A. Gu, M. Eichhorn, A. Rudra, and C. Ré. Learning fast algorithms for linear transforms using butterfly factorizations. In *ICML*, 2019.
- [2] L. Zheng, E. Riccietti, and R. Gribonval. Efficient identification of butterfly sparse matrix factorizations. *SIAM Journal on Mathematics of Data Science*, 2023.
- [3] Q.-T. Le, L. Zheng, E. Riccietti, and R. Gribonval. Fast learning of fast transforms, with guarantees. In *ICASSP*, 2022.
- [4] M. O'Neil, F. Woolfe, and V. Rokhlin. An algorithm for the rapid evaluation of special function transforms. *Applied and Computational Harmonic Analysis*, 2010.
- [5] E. Candes, L. Demanet, and L. Ying. A fast butterfly algorithm for the computation of Fourier integral operators. *Multiscale Modeling & Simulation*, 2009.
- [6] Y. Li, H. Yang, E. R. Martin, K. L. Ho, and L. Ying. Butterfly factorization. *Multiscale Modeling & Simulation*, 2015.
- [7] R. Vidal, Y. Ma, and S.S. Sastry. *Generalized Principal Component Analysis*. Springer New York, 2016.