PHD RESEARCH TOPIC - ANR ADAPTD

ParaTuck Decomposition for Advanced Multidimensional Data Processing

Host institution: University of Toulon, LIS UMR 7020, SIIM team
Project: ANR ADAPTD - Advancing Multidimensional Data Processing with ParaTuck Decomposition
Supervisors: Yassine Zniyed, André L. F. de Almeida, Nadège Thirion-Moreau, and Eric Moreau

Keywords: Signal processing, tensor decomposition, numerical optimization, applied mathematics.

Application: Please send your CV, transcripts, and a cover letter to zniyed@univ-tln.fr.

Modern data are often indexed by several variables simultaneously: space, time, frequency, experimental condition, population, sensor, or operating regime. Tensors preserve this multidimensional organization instead of merging heterogeneous axes into a matrix, and their low-rank decompositions can reveal latent components that remain interpretable in each mode. Canonical polyadic decomposition (CPD) provides parsimonious and often essentially unique representations, Tucker models allow richer interactions through a core tensor, and tensor networks address very high orders. These tools have become central to signal processing and machine learning because they combine compression, data fusion, identifiability, and computational structure [1, 2, 3].

Many inverse problems, however, are governed by a more specific interaction: two latent families are independently modulated and coupled through an intermediate operator. A third-order ParaTuck-2 decomposition of $I \times J \times K$ tensor $\boldsymbol{\mathcal{X}}$ is defined by its mode-3 slices as

$$\boldsymbol{\mathcal{X}}_{:,:,k} = \mathbf{A}\mathbf{D}_{\mathbf{f}}^{(k)}\mathbf{H}\mathbf{D}_{\mathbf{g}}^{(k)}\mathbf{B}^{\mathsf{T}}, \qquad k = 1, \ldots, K, \tag{1}$$

where the two diagonal matrices $\mathbf{D}_{\mathbf{f}}^{(k)}$ and $\mathbf{D}_{\mathbf{g}}^{(k)}$ describe slice-dependent modulations and $\mathbf{H}$ describes interactions between the latent families. This double-diagonalization mechanism is more flexible than CPD while retaining the prospect of essential uniqueness under appropriate assumptions; it therefore provides an attractive compromise between PARAFAC-type identifiability and Tucker-type interaction flexibility [4, 5, 6]. DEDICOM is a tied special case in which the loading and modulation factors coincide. Conversely, higher-order ParaTuck models, such as ParaTuck-(2,4), introduce additional observation modes and are needed when data are indexed by several spatial, temporal, spectral, or experimental variables.

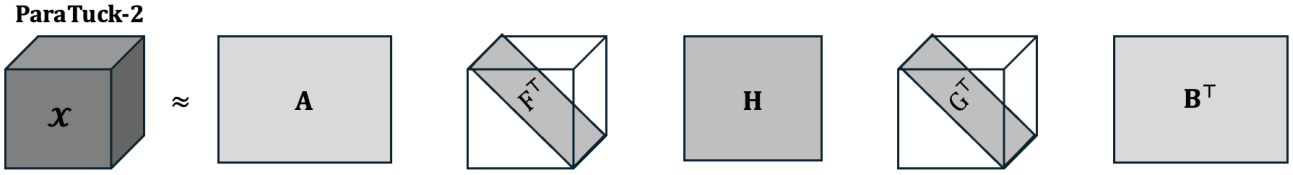


Figure 1: Schematic representation of the ParaTuck-2 decomposition.

Despite this modeling potential, ParaTuck estimation remains much less mature than CPD or Tucker estimation. Alternating least squares (ALS) is the standard baseline, but the fully unknown ParaTuck objective is strongly nonconvex, ill-conditioned, and affected by interacting scale ambiguities. In practice, successful ALS-based models often rely on at least one known or strongly constrained factor, for example training, coding, or allocation structures in communication systems [7, 8]. Classical DEDICOM solvers, including Kiers' ALS methods and ASALSAN, provide useful baselines but do not solve the general untied ParaTuck problem [9, 10]. Stochastic and population-based ALS strategies improve the probability of escaping poor attraction basins, yet their performance remains most convincing at low or moderate ranks and their repeated local searches can be extremely costly [11, 12]. The algebraic lifting approach of Usevich is another important alternative, but it depends on restrictive Harshman-Lundy-type conditions and its lifting and factor extraction steps can be sensitive to noise [13, 5]. There is consequently no method that is simultaneously robust, scalable, computationally efficient, and supported by convergence guarantees for general ParaTuck models.

The thesis will address this gap through four tightly connected objectives. First, it will develop efficient and robust algorithms for ParaTuck-2 when all factors are unknown, with reduced sensitivity to initialization, noise, ill-conditioning, rank, and moderate model mismatch. Candidate strategies include damped Gauss-Newton and Levenberg-Marquardt methods, trust-region or nonlinear conjugate-gradient schemes, variable projection, proximal block optimization, and hybrid deterministic-stochastic procedures. Their complexity will be evaluated in arithmetic cost, memory, runtime, and number of initializations needed for reliable recovery.

Second, the algorithms will be generalized to large dimensions and higher orders, in particular ParaTuck-(2,4), using tensor contractions, matrix-free Jacobian products, randomized linear algebra, and parallel implementations rather than explicit Kronecker matrices. The aim is to preserve the model structure and physical meaning of the factors while controlling the rapid growth in memory and computational cost.

Third, the thesis will analyze convergence and stability. Beyond empirical comparisons, the work will characterize stationary points, local rates, conditioning, attraction basins, perturbation sensitivity, and the influence of initialization. It will seek conditions under which deterministic, stochastic, or population-based iterations converge to meaningful solutions, and relate the numerical behavior to identifiability and to the geometry induced by the double-diagonal operator.

Fourth, physical and statistical constraints will be introduced directly into the algorithms: nonnegativity, sparsity, smoothness, symmetry, positive definiteness, Vandermonde or Toeplitz structure, shared factors, and partial calibration information. These priors are not only application-specific regularizers. They can remove continuous ambiguities, improve Jacobian conditioning, prevent degenerate iterates, and enlarge useful attraction basins; when inconsistent with the data, they may instead create biased or additional stationary points. A central contribution will therefore be to quantify their role in convergence, stability, identifiability, and data requirements.

The methodology will remain general and will be assessed in three complementary application domains. In covariance-domain blind source separation, a single dataset with nonstationary correlated sources leads to a symmetric DEDICOM model, whereas multiple datasets generate a general ParaTuck-2 model with two loading matrices and two modulation families; the objective is to recover both the mixtures and their correlation structure beyond conventional joint diagonalization [14, 15]. In telecommunications, ParaTuck-2 and ParaTuck-(2,4) arise in multidimensional parameter estimation, including time-varying MIMO channels and sensing components of integrated sensing and communication systems; these examples will mainly test high-order scalability and physical constraints [16, 8]. Finally, Jacobians of multivariate maps can be organized as ParaTuck-2 tensors whose factors encode neural-network weights and activation derivatives. The link established by Usevich et al. opens a route to identifiable and interpretable learning, for which the algorithms developed in the thesis could provide more reliable training procedures [17].

The expected outcome is a unified methodology for double-diagonal tensor estimation: robust solvers for fully unknown ParaTuck models, scalable extensions to large and high-order data, theoretical convergence and stability results, and a principled account of how physical constraints improve or hinder recovery.

Selected references

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[2] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.

[3] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation.* Chichester, UK: John Wiley & Sons, 2009.

[4] R. A. Harshman and M. E. Lundy, "PARAFAC: Parallel factor analysis," *Computational Statistics and Data Analysis*, vol. 18, no. 1, pp. 39–72, 1994.

[5] ——, "Uniqueness proof for a family of models sharing features of tucker's three-mode factor analysis and PARAFAC/CANDECOMP," *Psychometrika*, vol. 61, no. 1, pp. 133–154, 1996.

[6] G. Favier and A. L. F. de Almeida, "Overview of constrained PARAFAC models," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 142, 2014.

[7] P. Comon, X. Luciani, and A. L. F. de Almeida, "Tensor decompositions, alternating least squares and other tales," *Journal of Chemometrics*, vol. 23, no. 7–8, pp. 393–405, 2009.

[8] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Space-time spreading-multiplexing for MIMO wireless communication systems using the PARATUCK-2 tensor model," *Signal Processing*, vol. 89, no. 11, pp. 2103–2116, 2009.

[9] H. A. L. Kiers, "An alternating least squares algorithm for fitting the two- and three-way DEDICOM model and the IDIOSCAL model," *Psychometrika*, vol. 54, no. 3, pp. 515–521, 1989.

[10] B. W. Bader, R. A. Harshman, and T. G. Kolda, "Temporal analysis of semantic graphs using ASALSAN," in *Proceedings of the Seventh IEEE International Conference on Data Mining*, 2007, pp. 33–42.

[11] Y. Zniyed and A. L. F. de Almeida, "A stochastic algorithm for the ParaTuck decomposition," *Digital Signal Processing*, vol. 156, p. 104767, 2025.

[12] Y. Zniyed, "A population-based ALS for ParaTuck-2 decomposition with local convergence guarantees," *Digital Signal Processing*, vol. 181, p. 106261, 2026.

[13] K. Usevich, "A lifting approach to ParaTuck-2 tensor decompositions," 2025.

[14] E. Moreau, "A generalization of joint-diagonalization criteria for source separation," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 530–541, 2001.

[15] D.-T. Pham and J.-F. Cardoso, "Blind separation of instantaneous mixtures of nonstationary sources," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1837–1848, 2001.

[16] G. Favier, M. N. da Costa, A. L. F. de Almeida, and J. M. T. Romano, "Tensor space-time coding for MIMO wireless communication systems," *Signal Processing*, vol. 92, no. 4, pp. 1079–1092, 2012.

[17] K. Usevich, Y. Zniyed, M. Ishteva, P. Dreesen, and A. L. F. de Almeida, "Tensor-based two-layer decoupling of multivariate polynomial maps," in *31st European Signal Processing Conference (EUSIPCO)*, 2023, pp. 655–659.