Supplementary Material

Christos Chatzis[1,2], Carla Schenker[1], Max Pfeffer[3], and Evrim Acar[1]

[1]Simula Metropolitan Center for Digital Engineering
[2]Oslo Metropolitan University, Faculty of Art, Design and Technology
[3]Institute for Numerical and Applied Mathematics Georg-August-Universität Göttingen

1 Sensitivity analysis of hyperparameter λ_B

In this section, we explore the performance of tPARAFAC2 in terms of the smoothness hyperparameter in the fully observed experimental setting presented in section 4.2.1. Figure S.1 shows that for smaller values, the regularization is ineffective, while excessively high values result in over-smoothing of the patterns, which in turn degrades accuracy. On the lowest noise level considered ($\eta = 0.5$), the performance is not sensitive to the smoothness hyperparameter, and only the threshold above which oversmoothing happens is visible. In contrast, at higher noise levels, stronger temporal penalties are necessary to achieve improved factor recovery and the improvement is more clear. The best-performing value for each case is dependant on the level of the noise, with higher noise levels requiring stronger regularization.

2 Additional information on metabolomics application

2.1 Selection of the number of components

We select the number of components ($R = 2$) based on the replicability of the extracted components, *i.e.,* the ability to extract similar patterns (in terms of their FMS) from random subsamples of the data. This approach has been previously used for the selection of the number of components in CP models [1]. An R-component model is considered to be replicable if 95% of all computed FMS values are higher than 0.9. We then choose the highest number of components that produces a replicable model.

In Figure S.2, we show the replicability results for $R = 2$ and $R = 3$ components for the PARAFAC2 model. The horizontal line indicates the 95% highest FMS values. For $R = 2$, we can clearly see that all FMS values are above 0.9, meaning that this model is replicable. For $R = 3$, on the other hand, the line is below 0.9, indicating a non-replicable model.

2.2 Correlations with meta variables

Here, we show the correlation between the subject scores from the component of interest and meta variables for the CP model [1], the PARAFAC2 model, the PARAFAC2 model with ridge ($\lambda = 1, 10$) and tPARAFAC2 with $\lambda_B = 1, 10, 100$ and ridge $\lambda = 10$. Descriptions of meta variables are as follows: HOMAIR: Homeostatic model assessment for Insulin Resistance; MuscleFatRatio: Muscle to fat ratio; FatPercent: Body fat percentage; MuscleMass: Amount of muscle in the body (kg); Weight (kg); BMI: Body Mass Index; Waist: Waist circumferance (cm); WaistHeightRatio: Waist measurement divided by height (cm); FatMass: Amount of body fat (kg); FatMassIndex; FFMI: Fat Free Mass Index.

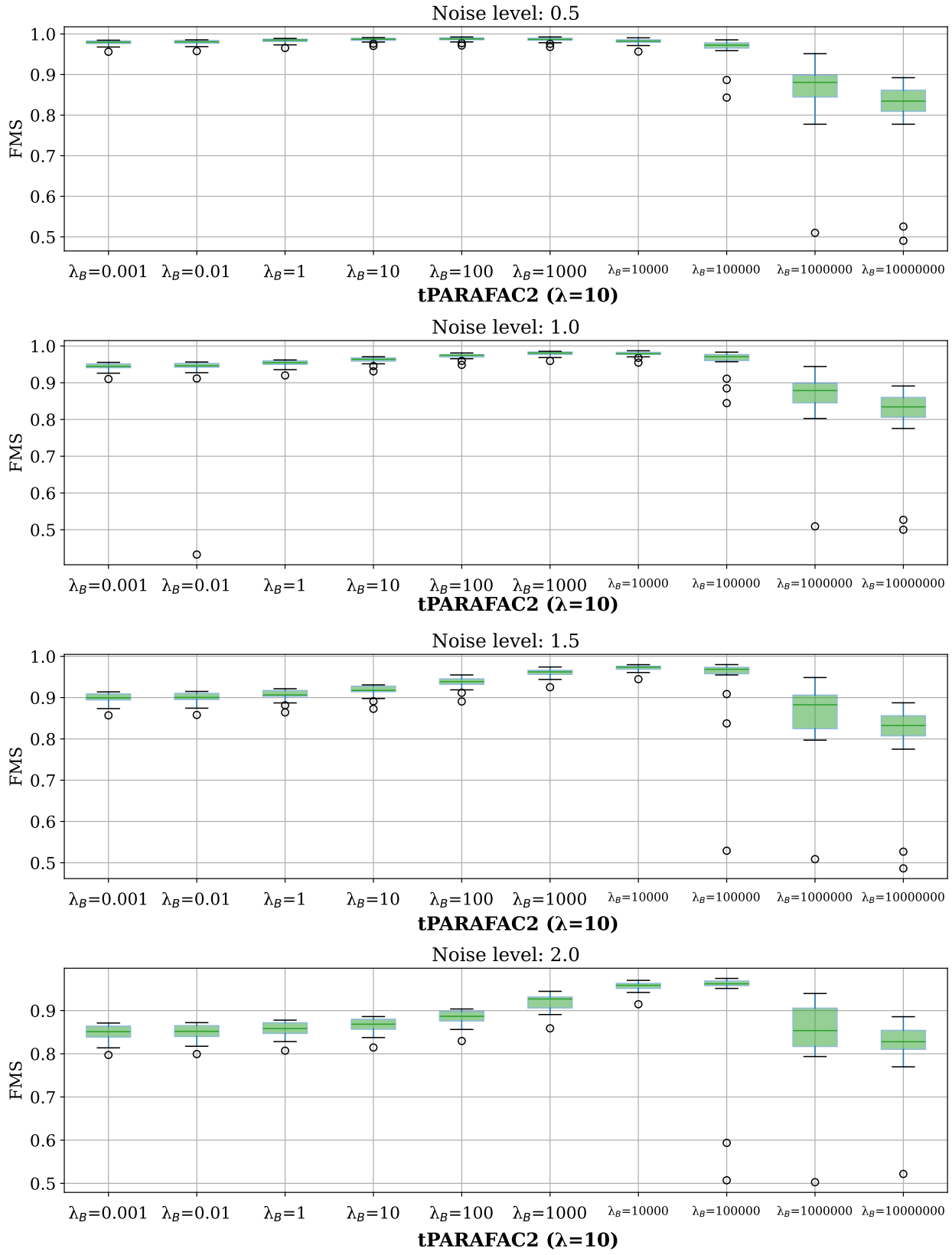

Figure S.1: FMS of tPARAFAC2 factors of the best run for each dataset with the ground truth factors for the four noise levels considered. Each boxplot contains 20 points, one for each dataset of the setting.

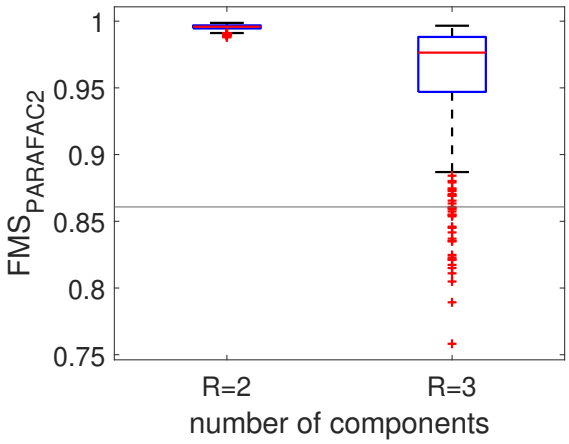

Figure S.2: Replicability of PARAFAC2 components for different number of components, i.e., $R = 2$ and $R = 3$ components.

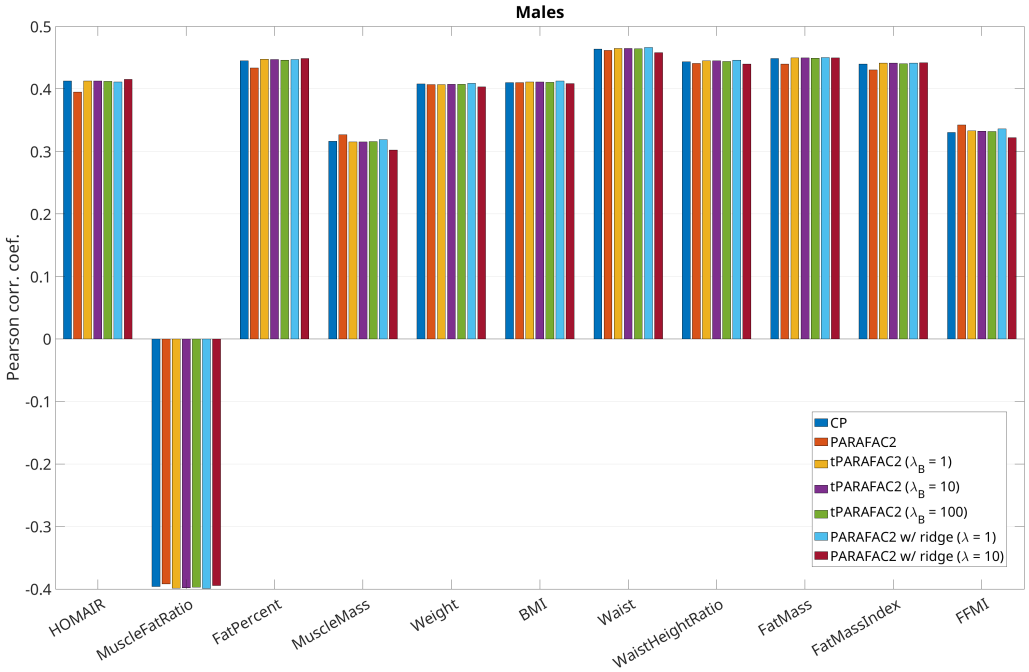

Figure S.3: Correlation coefficients of subject scores with meta variables for different models.

3 AO-ADMM exit conditions

As also mentioned in the main text, outer AO-ADMM exit once either the absolute or the relative change in function value is smaller than the pre-defined tolerance:

$$\left\| f^{(n+1)} - f^{(n)} \right\|_F < \epsilon_{abs} \tag{S.1}$$

$$\frac{\left\| f^{(n+1)} - f^{(n)} \right\|_F}{\| f^{(n)} \|_F} < \epsilon_{rel} \tag{S.2}$$

where the function value at the n-th iteration is given by:

$$f^{(n)} = \left\{ \sum_{k=1}^{K} \left\| \mathbf{X}_k - \mathbf{A}^{(n)} \mathbf{D}_k^{(n)} \mathbf{B}_k^{(n)T} \right\|_F^2 + \lambda_A \left\| \mathbf{A}^{(n)} \right\|_F^2 + \sum_{k=1}^{K} \left(\lambda_{\mathbf{D}} \left\| \mathbf{D}_k^{(n)} \right\|_F^2 \right) \right\} \tag{S.3}$$

with $\mathbf{A}^{(n)}$, $\mathbf{B}_k^{(n)}$ and $\mathbf{D}_k^{(n)}$ denote the computed factor at the n-th iteration. A maximum number of iterations is pre-set as well. Additionally, the solution has to have adequately small feasibility gaps:

$$\frac{\| \mathbf{M} - \mathbf{Z}_M \|_F}{\| \mathbf{M} \|_F} < \epsilon_{feasibility} \tag{S.4}$$

Furthermore, for the inner AO-ADMM loops (e.g. Algorithm 1 in the main text), the stopping conditions are:

$$\frac{\left\| \mathbf{M}^{(l)} - \mathbf{Z}_M^{(l)} \right\|_F}{\left\| \mathbf{M}^{(l)} \right\|_F} < \epsilon_{inner} \tag{S.5}$$

$$\frac{\left\| \mathbf{M}^{(l)} - \mathbf{Z}_M^{(l-1)} \right\|_F}{\left\| \mathbf{Z}_M^{(l)} \right\|_F} < \epsilon_{inner} \tag{S.6}$$

where l denotes the inner iteration number and $\mathbf{M}$ and $\mathbf{Z}_M$ denote any of the factor matrices (that involve regularization) and the respective auxiliary variable introduced. The maximum number of inner iterations is set to 5 in all of our experiments for the inner loops.

4 Convergence

In this section we briefly discuss the convergence of tPARAFAC2. Solving for the PARAFAC2 factors requires solving a non-convex problem, which includes the PARAFAC2 constraint (Equation (2)). While the AO-ADMM framework for PARAFAC2 [2] allows for applying different constraints on the evolving mode, something we leverage when fitting tPARAFAC2, it does not have any convergence guarantees for the full problem. Nevertheless, existing experiments, confirmed also by our results, indicate that the method in practice converges.

4.1 Fully observed data

To exemplify, we show the convergence behavior of tPARAFAC2 on one of the datasets in the fully observed data (see in the GitHub repo, `synthetic_data/noisy/input/incrementals[4]/0.5/noisy_dataset.pkl`), where the noise level $\eta = 0.5$ and the model hyperparameters are chosen as $\lambda = 10$ and $\lambda_B = 10000$. The performance of all thirty initializations is shown in Figure S.4 in terms of total regularized loss (given by Equation (6)) and the feasibility of the solutions in terms of the PARAFAC2 constraint and the temporal smoothness imposed. The solid line shows the median across runs per iteration, and the shaded area spans values all the non-median runs. The non-negativity feasibility is exactly 0, therefore, the solution is indeed feasible.

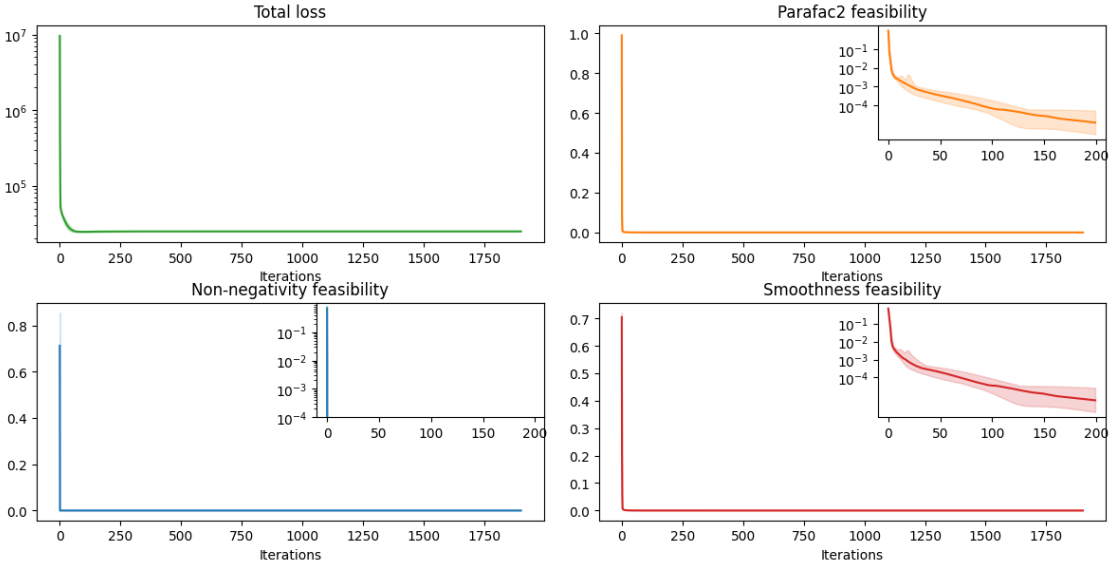

Figure S.4: Convergence diagnostics for dataset 4, with noise level $\eta = 0.5$. All thirty initializations used are shown with the 'median' run highlighted with a solid line.

The following can be inferred from Figure S.4:

- The total loss is constantly decreasing and the algorithm is converging before reaching the maximum number of iterations of 10000.
- The solutions given by the algorithm do satisfy the set constraints within the given feasibility gap.

This behavior is consistent across all datasets of this case.

4.2 Data with unobserved entries

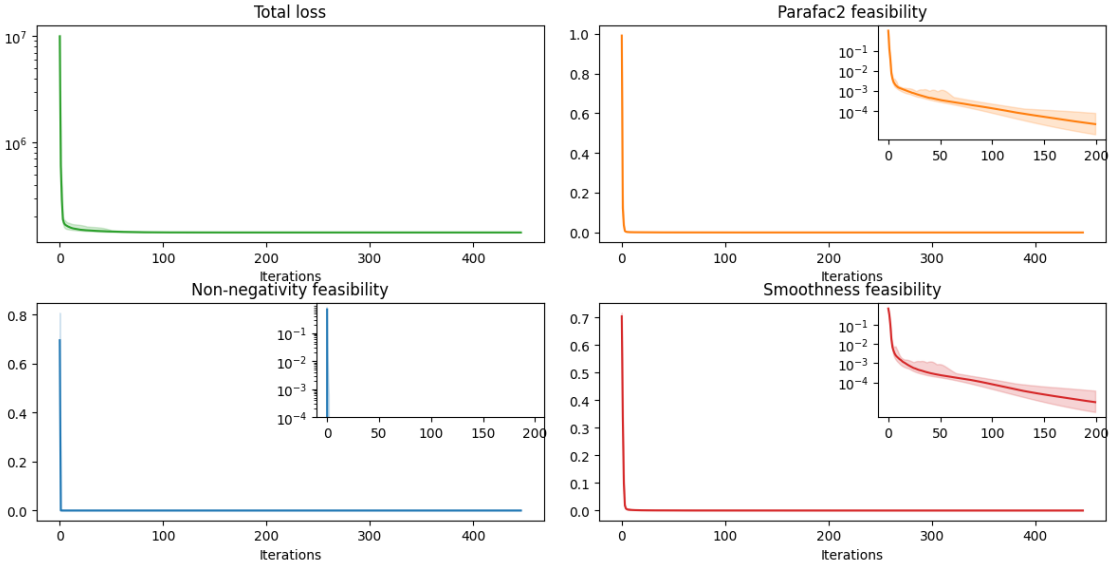

Figure S.5: Convergence diagnostics for dataset 11, when `mask9` is chosen. All thirty initializations are shown with the 'median' run highlighted with a solid line.

We also observe similar behavior in terms of convergence when part of the input is unobserved. In Figure S.5 we can see the same plot as the one in the fully observed data for `synthetic_data/missing_random/input/missing_incrementals[11]/50/`, when `mask9` is chosen as the indicator tensor. Similar behavior was observed also on experiments with structured unobserved entries (fibers).

References

[1] S. Yan, L. Li, D. Horner, P. Ebrahimi, B. Chawes, L. O. Dragsted, M. A. Rasmussen, A. K. Smilde, and E. Acar, "Characterizing human postprandial metabolic response using multiway data analysis," *Metabolomics*, vol. 20, no. 3, p. 50, 2024.

[2] M. Roald, C. Schenker, V. D. Calhoun, T. Adali, R. Bro, J. E. Cohen, and E. Acar, "An ao-admm approach to constraining parafac2 on all modes," *SIAM Journal on Mathematics of Data Science*, vol. 4, no. 3, pp. 1191–1222, 2022.