The authors introduce an efficient implementation of the shared response model (SRM) analysis, and demonstrate its performance using simulations, and using public fMRI data from a naturalistic experiment. The implementation code is also publicly available.
This work has already a high quality and the results are convincing. However, I believe the manuscript could be improved further by addressing a few concerns, listed below.

1- I find a couple of statements in the introduction section slightly inaccurate. Specifically, the way that the authors define the fMRI modelling problem seems upside down to me. In my view, requiring a "design matrix" is a consequence of the formulation of an encoding problem. But an encoding model is not the only form of model that could be perceived. Even assuming having an encoding model, having a design matrix shared across subjects is not an ecological assumption, but rather it is by design derived from the stimulus and not the subject responses. That said, I find it hard to relate the design matrix with the shared space modelling of the responses. I think some rephrasing of the motivation paragraphs in the introduction section is needed.

2- At a few places throughout the manuscript the authors touch on the identifiability issue of the standard SRM modelling. I find it hard to easily understand this point without picking a pen and paper and writing down a few lines of equations myself. The authors point the readers to appendix D from a referenced paper, but I think it would be more accessible to the reader if the proposition 3 from Appendix D from that referenced paper was briefly repeated in this manuscript, especially since it is not more than a few lines.

3- The authors motivate the possibility of using a spatial decomposition by referring to the existence of cortical parcellations and atlases in the literature. A common property of such decompositions is that they cluster voxels into a few distinct homogenous clusters. I was wondering if there are such spatial constraint in the optimal spatial decomposition used here. It could be useful to have a discussion of functional consequences of presence or lack of such spatial constraints in the optimal spatial decomposition.

4- During reviewing previous works, the authors correctly mention that SRM is fundamentally similar to hyperalignment. Although the authors explicitly mention that the scope of the manuscript is limited to introducing a faster SRM implementation, it would be useful to discuss the difference between the two in terms of performance, computational demand, and functional inference. A curious reader could ask why bother using SRM or fastSRM if they do not provide an obvious advantage over hyperalignment?

5- I had a hard time understanding the notion of stability here, and relating it to identifiability (and relatedly, to understand what Fig. 4 shows). The paper could benefit from a better introduction and explanation of this relationship.

Minor concerns:

1- "…the number of voxels is about 300 000…" this is inaccurate. This could be the typical number of voxels in the whole FOV. However, analyses is usually performed on the masked brain data, which typically consist of ~50k voxels.
2- "In general probabilistic methods give much better results than their deterministic counterpart…" this statement should be backed by a reference of otherwise justified.
3- A reference or explanation for the Hungarian algorithm is needed.

Mo Shahdloo