Response to Reviewer #3 Comments

John Waczak and David Lary

10-31-2024

We sincerely appreciate the time and effort that you dedicated to reviewing our manuscript and are grateful for your insightful comments and suggested improvements. Thank you!

Point 1 — The proposed Generative Simplex Mapping (GSM) model offers a fresh approach to non-linear endmember extraction and spectral unmixing in hyperspectral imagery, which is a significant contribution to the field. Highlighting the model's flexibility to handle both linear and non-linear mixing, as well as its probabilistic nature, is commendable. However, it would be helpful if the authors more clearly articulated how the GSM directly advances the state of the art compared to other existing methods.

Response — Thank you for your kind comment. We have updated the introduction to explicitly include a list of the novel contributions made by our approach. The updated text now includes the following:

> In summary, the key innovations introduced by the GSM are:
>
> - the GSM can model linear and nonlinear spectral mixing
> - the GSM does not assume the presence of pure pixels in the dataset
> - the probabilistic formulation of the GSM accounts for spectral variability
> - the simplex used for the latent space structure of the GSM is directly interpretable and forces abundances to satisfy both the abundance sum-to-one and abundance non-negativity constraints
> - the fiting procedure introduced for the GSM maintains non-negativity of endmember spectra.

Point 2 — The description of the model, especially the non-linear mapping function and the (n−1)-simplex latent space, is interesting but requires more elaboration. Providing additional mathematical explanations or diagrams could help clarify the mechanics of the latent space and the transition between linear and non-linear regimes. Some readers may struggle with the abstract nature of the model description without further illustrative examples.

Response — Thank you for your comment. We have updated Section 2 to include additional descriptions for the model components. Specifically, the descriptions of Equations (1) and (2) have been extended with additional details for the variables. The

description of the non-linear acitvation functions has also been updated and now reads as:

> A visual representation of the non-linear activation functions in Equation 4 is shown below in Figure 1 for a 2-component GSM. In this form, the first N_v columns of $\boldsymbol{W}$ correspond to endmember spectra, while the remaining columns account for additional non-linear effects. Equation 4 is specifically chosen so that no non-linear contributions are possible for pure spectra at the vertices of the simplex. At all other points, the output of ψ involves both linear and non-linear contributions. If only linear mixing is present, the GSM training algorithm should therefore drive W_{dm} to 0 for $m \geq N_v$.

We have also added a new figure to illustrate the non-linear activation function for a 2-dimension GSM.

Point 3 — The comparison with three varieties of non-negative matrix factorization (NMF) on synthetic data is valuable. However, it would strengthen the evaluation if the authors included additional benchmark models, such as other widely used non-linear unmixing algorithms, to more comprehensively demonstrate GSM's performance advantages. Are there specific scenarios where GSM is expected to significantly outperform standard methods, and if so, could these be highlighted?

Response — Thank you for your comment. We agree that evaluating the GSM against standard models is valuable. Our goal with the initial experiment on the *linearly* mixed synthetic dataset was primarly to demonstrate that the GSM can solve linear mixing tasks without introducing unnecesary complexity by forcing spurious nonlinear contributions. We believe this is a key advantage over other non-linear mixing approaches such as autoencoder models, which include non-linear mixing by design even if the underlying data is only linearly mixed.

To highlight this, we have updated the text to now read as follows:

> To illustrate the effectiveness of the GSM, we first demonstrate its ability to model linear mixing. This serves as an important limiting case since linearly mixed spectra should not lead to the spurous introduction of non-linear contributions by the GSM. The goal of this first test is therefore to demonstrate that the GSM drives non-linear weights to zero for linearly mixed data while providing a fair test to compare the GSM to a well-established linear mixing model. This ability clealy distinguishes the GSM from other non-nonlinear unmixing approaches such as autoencoders, which

> by their design, include non-linear mixing even when it is not present in the underlying data.

Additionally, we have added the following line after describing the varietites of NMF we considered:

> We note that the goal of this test is not to prove the GSM is superior to other models for linear mixing, but rather to demonstrate that the GSM can model linearly mixed data without introducing unnecessary complexity.

Point 4 — The synthetic data experiment and the real-world case study over a pond in North Texas are good choices for demonstrating the capabilities of GSM. Nonetheless, it would be beneficial to include more details on the real dataset, such as the resolution, spectral characteristics, and preprocessing steps, to help assess the generalizability of the method. Additionally, further validation on different types of real-world datasets with varying levels of complexity would strengthen the claims of the model's robustness.

Response — Thank you for your comment. Details for the *real* HSI dataset are described in Section *3.2 Non-linear Mixing: Water Contaminent Identification*. Regarding the resolution and spectral characteristics, we wrote

> Each HSI pixel included 462 wavelength bins ranging from 391 to 1011 nm.

For the preprocessing steps, captured HSI were converted from radiance into reflectance cubes using the downwelling irradiance spectrum according to Equation (20).

The remaining processing steps leading to the final dataset are outlined as follows:

> From the collected HSI, a water-only pixel mask was generated by identifying all pixel spectra with a normalized difference water index (NDWI) greater than 0.25 as defined in ref. [47]. Of these water pixels, a combined data set of 15,000 spectra was sampled for GSM training. As a final processing step, reflectance spectra were limited to $\lambda \leq 900$ nm as wavelengths above this threshold showed significant noise.

Many papers which develop non-linear mixing models rely on synthetic datasets with explicitly defined non-linear effects such as bilinear mixing or polynomial post-nonlinear mixing. However, these models can often be fit by linear mixing models with extra *virtual* sources for the higher order terms. We therefore felt that a fairer demonstration for the GSM would be to use real HSI captured for water where we can expected more complicated non-linear mixing effects.

We did try searching for other benchmark datasets with *real* HSI for additional evaluation. The best we were able to find was the DLR HySU datasets from Cerra et al. (https://doi.org/10.3390/rs13132559), however the *ground truth* values for the abundances provided by the authors assume a linear mixing model as the sources corresponeded to flat tarps of different materials placed on the ground. We would welcome any suggestions for additional benchmark non-linear unmixing data-sets with ground truth abundance values.

Point 5 — The probabilistic treatment of spectral variability using a precision parameter is an interesting aspect of the model. However, the authors could provide more details on how this precision parameter is estimated, and how it affects the unmixing results. A deeper discussion of its impact on the overall performance of the GSM model, compared to deterministic approaches, would be valuable.

Response — Thank you for your comment. We agree that further clarification regarding the precision parameter will help strengthen the paper. The parameter is first introduced in Equation (1) where it parameterizes the normal distribution assumed for the (spectral) data space. We have augmented the description to now include the following:

> As Equation 1 indicates, the precision parameter corresponds to a standard deviation of $\sqrt{\beta^{-1}}$.

The precision parameter is updated during each iteration of the EM routine during the maximization step according to Equation (13). As a clear example, Figure (7) shows the extracted endmembers for the simulated datasets at an SNR of 20 with their associate spectral variability from the fitted β parameter. The figure includes the following description:

> Colored bands are included around each spectrum corresponding to the spectral variability estimated by the GSM precision parameter β where the band width is $2\sqrt{\beta^{-1}}$ corresponding to 2 standard deviations.

We also comment on these results, noting

The SNR of 20 added to this example corresponds to zero-mean Gaussian noise with a standard deviation of $\sigma = 0.0493$. After training, the GSM found $\sqrt{\beta^{-1}} = 0.0495$, accurately identifying this introduced noise. The ability to assess the spectral variability of extracted endmembers is a key advantage of the GSM resulting from its probabilistic formulation.

Point 6 — While the model's performance is highlighted, the computational cost of GSM relative to other models (e.g., NMF) is not fully discussed. Since hyperspectral data can be large and computationally demanding, it would be useful for the authors to provide an analysis of the algorithm's scalability and its computational requirements, especially for large datasets or real-time applications.

Response — Thank you for your comment, this is an excellent point. A key factor impacting the computational efficiency of the model is the number of nodes utilized the the latent space. Together with the total number of data points, this contributes to the size of the responsability matrix which is evaluated during each expectation step. To address this we described an alternitive approach in which points in the latent space are sampled using a uniform Dirichlet distribution in order to acheive a specified total of K-many nodes (referred to as the *big* GSM model in Figure (6)). We also refer to this limitation in the discussion section which now includes:

The main limitation of the GSM is the curse of dimensionality encountered when generating a grid on the $(N_v - 1)$-simplex for large numbers of endmembers, N_v. This can be mitigated by instead randomly sampling points within the interior of the simplex using a uniform Dirichlet distribution to obtain a pre-determined number of nodes across the latent space. As the mixing coefficients π_k are adapted during training, variability in spacing between nodes should not significantly affect the performance of the model. This was confirmed for the simulated data set as shown in Figure 6. In terms of computational efficiency, each expectation step involves O(K × N) operations to update the entries of the responsibility matrix leading to extended training times for considerably large data sets. This can be addressed by augmenting the EM procedure to use mini-batches of training samples as outlined by Bishop et al. for the GTM in ref. [48]. Rather than updating the full responsibility matrix during each iteration, a subset of $\boldsymbol{R}$ corresponding to a single batch of training data can be evaluated with all other entries kept constant. The GSM may also be extended in other ways, for example, by replacing the precision parameter β with a vector β_λ of values to model wavelength-dependent spectral variability common to many hyperspectral imaging platforms.