Machine Learning (2025) 114:157
https://doi.org/10.1007/s10994-025-06793-2

Simultaneous outlier detection and elimination in hyperspectral unmixing via weighted non-negative matrix tri-factorization

Yasin Hashemi-Nazari[1] · Farid Saberi-Movahed[2] · Azita Tajaddini[1] · Catarina Moreira[3] · Xin Ning[4] · Prayag Tiwari[5]

Received: 8 February 2025 / Revised: 1 May 2025 / Accepted: 4 May 2025 /
Published online: 5 June 2025

Abstract
Hyperspectral unmixing (HU) involves separating mixed pixel spectra into pure endmember spectra and their corresponding abundance fractions. However, it faces significant challenges due to outliers in the hyperspectral data, which often appear as pixel and band anomalies. Outliers in pixels could result in incorrect classification and inaccurate quantification of materials, while outliers in bands could alter spectral characteristics, leading to misidentifying endmembers and incorrect estimates of abundance. To tackle these issues, this paper introduces a new approach, named simultaneous outlier detection and elimination via weighted non-negative matrix tri-factorization (SODE-WNMTF), which offers an efficient means of addressing the impact of outliers in the unmixing process. Leveraging the co-clustering property of NMTF, SODE-WNMTF introduces a novel weighting matrix, which involves simultaneous clustering of both pixels and spectral bands to effectively detect and mitigate the negative impact of both pixel and band outliers during the unmixing process. At the same time, the inherent structure of the hyperspectral image (HSI) is utilized through the examination of local and global connections among pixels and spectral bands, consequently improving the co-clustering procedure. In addition, SODE-WNMTF proposes a spatial weighting factor, which utilizes the similarity of adjacent pixels, to promote piecewise smoothness in abundance maps while mitigating the impact of outliers. Moreover, since pixels in regions dominated by a single endmember exhibit spectra closely resembling that endmember, SODE-WNMTF incorporates a sparse estimation technique for endmember signatures. Finally, to verify the performance of SODE-WNMTF, a series of experiments is conducted on both synthetic and real HSIs, with outcomes proving its superiority against other cutting-edge approaches. The source code is also available at https://github.com/yasinhashemi/SODE-WNMTF.

Keywords Hyperspectral unmixing (HU) · Non-negative matrix tri-factorization (NMTF) · Outlier detection · Co-clustering · Spatially weighted unmixing

Editor: Dino Ienco.

Extended author information available on the last page of the article

1 Introduction

Hyperspectral imaging has become a game-changer in remote sensing, offering detailed spectral information over a wide range of wavelengths for every pixel in an image (Ma et al., 2014). This rich spectral data helps us identify and analyze materials with incredible precision, which is essential for various applications like environmental monitoring, mineral exploration, and agriculture (Li et al., 2021a; Feng et al., 2022). However, a significant challenge arises with the mixed pixel problem, where a single pixel contains multiple materials, leading to complications in the use of HSIs for classification and object identification (Qu & Bao, 2020). This is where HU comes in, breaking down these mixed pixels into pure spectral signatures (endmembers) and their corresponding proportions (abundances). HU algorithms typically operate under the assumption of a specific mixing model, either linear or nonlinear (Bioucas-Dias et al., 2012). Linear models operate under the assumption that various endmembers function independently (Miao & Qi, 2007), whereas nonlinear models suggest that the observed spectrum results from a nonlinear relationship between the abundances of endmembers (Dobigeon et al., 2014).

The linear mixing model (LMM) is widely used in HU, assuming each pixel's spectrum is a linear combination of endmember spectra weighted by abundances (Miao & Qi, 2007). Although LMM has limitations, particularly when confronted with nonlinearly mixed pixels, it is acknowledged as a suitable model for numerous practical applications (Feng et al., 2022). Many diverse HU strategies have been suggested based on LMM, which can be divided into geometrical (Drumetz et al., 2020; Nascimento & Dias, 2005), statistical (Liu et al., 2023a, b; Mantripragada & Qureshi, 2024) and dictionary-based techniques (Yang et al., 2023; Cui et al., 2023). Among these, NMF-based methods (Lee & Seung, 1999; Haghir Chehreghani, 2020), a key subset of dictionary-based approaches, have gained prominence in HU. NMF aims to decompose a complex dataset into two simpler, nonnegative matrices. One matrix represents the basis vectors, which capture fundamental patterns in the data, while the other matrix contains coefficient vectors, indicating how much each pattern contributes to each data point (Wang et al., 2013). Within the framework of HU, the basis vectors represent endmember signatures, while the coefficient vectors denote the abundances of the respective endmembers. Nevertheless, traditional NMF algorithms frequently encounter challenges like local minima and non-unique solutions because of the non-convex nature of the objective function (Qu et al., 2024). To alleviate this issue, additional constraints need to be placed on NMF in order to obtain results that are meaningful in a physical sense. Several constrained NMF techniques have been developed for this purpose, and they can be classified into three groups: those imposing constraints on either endmembers or abundances, or on both.

Various endmember constraints for HU have been designed, with a particular focus on minimal volume methods, based on the assumption that image pixels typically lie within a simplex defined by endmember vertices (Rasti et al., 2022). Among these are two approaches: minimum volume constraint NMF (MVC-NMF) (Miao & Qi, 2007) and volume-regularized NMF (VRNMF) (Ang & Gillis, 2019), both of which aim to minimize the volume of the simplex to efficiently identify endmembers from highly mixed data. In contrast to employing simplex volume, the spectral-spatial constrained NMF (SS-NMF) (Zhang et al., 2021) technique was developed, incorporating the total distance between each endmember spectrum and the average spectral signature to ensure a compact simplex. Moreover, to overcome the challenge of incomplete spectral data in HU, researchers have developed innovative approaches. One such method, proposed by Wang et al. (2023),

utilizes a functional factorization model, which defines endmembers via a continuous representation acquired by implicit neural networks, allowing it to effectively capture the non-uniform spectral wavelength distribution.

Recent progress in NMF algorithms with constraints on abundance for HU has explored diverse constraints to enhance accuracy and robustness. These constraints involve utilizing total variation (TV) regularizers to enforce smoothness, aiming to capture the locally homogeneous distribution present in each abundance map (He et al., 2017; Fang et al., 2018b). An example of a model employing this strategy is presented by Song et al. (2024), where they introduce a relative TV regularization on the abundances to leverage the spatial characteristics of hyperspectral data. A different class of abundance constraints involves sparsity constraints, typically enforced through sparsity-inducing norms. For example, in NewSpr-NMF (Qu & Li, 2024) and L_1 sparsity-constrained archetypal analysis (L_1-AA) (Xu et al., 2022b), these two popular sparsity norms—namely $L_{\frac{1}{2}}$ and L_1—are utilized to address the inherent sparsity of endmember presence in individual pixels. Methods that incorporate spatial context, considering the relationships between abundances of neighboring pixels, have also become increasingly popular. Two prominent examples of this methodology are the spectral-spatial weighted sparse NMF (SSWNMF) (Zhang et al., 2022) and the spatial-spectral multi-scale sparse unmixing (S^2MSU) (Ince & Dobigeon, 2023), which integrate two weighting factors into the NMF framework to improve the sparsity of the solution while also preserving the spatially coherent pattern of the data. While these approaches aim to enhance HU performance, several challenges persist, such as the instability of sparsity constraints in noisy environments and the need for more comprehensive spatial structure details.

Dual-constrained NMF techniques for HU have emerged to address limitations of single-constraint approaches by simultaneously imposing constraints on both endmembers and abundances. These approaches combine diverse techniques, such as addressing sparsity in the abundance matrix and promoting minimal volume by shifting endmembers closer to the average of the dataset in robust collaborative NMF (R-CoNMF) (Li et al., 2016), and utilizing multilayer NMF enhanced by sparsity and graph regularization in double-constrained multilayer NMF (DCMLNMF) (Fang et al., 2018a) to investigate the inherent geometric structure of the data. Other specialized methods, such as multiple-priors ensemble constrained NMF (MPEC-NMF) (Qu & Bao, 2020), concurrently combine simplex volume minimization, sparsity promotion in abundance maps, and TV regularizers to enhance smoothness. Despite these advancements, challenges remain in effectively handling noise and outliers.

Most of the aforementioned HU models do not account for the impact of noise and outliers during the unmixing process. These irregular anomalies can greatly affect the precision of unmixing results, causing misclassification, incorrect material measurements, and spectral characteristic distortions (Kiani et al., 2024; Iwata et al., 2020). As a result, it can hinder endmember identification and abundance estimates, leading to inaccurate outcomes. Recent advances in NMF-based HU have focused on addressing mixed noise scenarios. For instance, approaches like sparsity regularized robust NMF (RNMF) (He et al., 2016) address mixed noise by individually modeling the sparse and Gaussian noise components. TV regularized reweighted sparse NMF (TV-RSNMF) (He et al., 2017) combines the TV regularization with reweighted sparsity for further noise robustness. Moreover, Zou et al. (2024) proposed a graph Laplacian weighted robust sparse unmixing algorithm based on superpixels combining local and global spatial information to reduce noise sensitivity while better reconstructing abundance details. CIM-NMF (Du et al., 2012) and EA-NMF

(Li & Chen, 2020) tackle noise by using a correntropy-induced metric instead of squared residuals. This approach is less sensitive to outliers and non-Gaussian noise, improving robustness. Furthermore, Huber (mHuber) NMF (Guo et al., 2021) tackles noise by incorporating a modified Huber loss function into NMF. This loss function is robust to outliers and noise, balancing between L_1 and L_2 norms to effectively handle both small and large errors in HSI. A different category of models tackles the problem of non-Gaussian noise. For example, Zhang et al. (2021) used an $L_{2,1}$-norm loss function to make their model less susceptible to outliers and this type of noise. Other types of models, such as higher-order graph regularization NMF with adaptive feature selection (HGNMF-FS) (Qu et al., 2023) and weighted residual NMF (WRNMF) (Ince & Dobigeon, 2022), denoise spectral band outliers by using a weighting matrix that penalizes bands with high residuals less and bands with low residuals more. Several models perform denoising by utilizing modified versions of standard NMF. For example, the target regularized tracking model (TRTM) (Li & Chen, 2023) employs NMTF to jointly analyze endmember regions, spatial pixels, and local patches, using a diagonal association matrix to capture spectral-spatial relationships. This reduces endmember variability, nonlinearity, and noise while improving extraction accuracy. Similarly, noise-free graph regularized model (NFGRM) (Li et al., 2021b) applies NMTF with a data-dependent diagonal matrix and weighted LASSO to suppress noise, while dual manifold constraints preserve structure.

Despite the proven efficiency of the previously robust HU methods, there remains a significant drawback: most of such methods do not effectively handle the impact of noisy signals in both the pixel and spectral band domains when unmixing. This lack of consideration can result in unstable and untrustworthy separation outcomes, putting the precision and correctness of the obtained data at risk (Lohrer et al., 2024). To get to the root of this problem, a novel model is introduced in this paper, which uses a weighted NMTF strategy to efficiently identify and eliminate the negative effects of outliers in both the pixel and band domains during the unmixing process at the same time. The suggested method distinguishes outliers by leveraging the co-clustering property of NMTF. It clusters pixels and spectral bands simultaneously, measuring the distance of each data point from its center. Points significantly further from their cluster center are considered outliers. A matrix of weights is created based on these distances, where lower weights indicate a higher likelihood of being outliers, reducing their influence on the HU results. The proposed approach further enhances the co-clustering accuracy by incorporating both local pixel neighborhood relationships and global spectral band similarities, leading to a more accurate representation of the underlying structure of data.

In addition, outliers can impede the effectiveness of spatial information in promoting sparsity within abundance maps, as they may be incorrectly weighted, resulting in a loss of smoothness in the abundance maps. To mitigate this problem, our proposed method utilizes a spatial weighting matrix based on the ℓ_1 sparse unmixing approach to reduce the negative effect of outliers by giving them lower weights, thereby promoting smoothness in the estimated abundance maps. This factor is determined based on the resemblance between contiguous pixels, aiming to enhance the piece-wise smoothness of the estimated abundance maps.

Another limitation seen in many NMF-based HU models is the lack of a regularization term, which could enhance the precision of estimating endmember signatures. Using the prior knowledge derived from pixel spectra in HSI, our innovative method presents a sparse estimation technique aimed at recognizing endmember signatures. In this method, each signature of an endmember can be extracted from merging solely the most representative pixel spectra, specifically those originating from areas where that endmember is detected. This method not

only incorporates the most characteristic pixel spectra in the HSI data but also mitigates the influence of mixed pixels where multiple endmember signatures are present.

These innovative concepts are integrated to create a new framework known as simultaneous outlier detection and elimination through weighted non-negative matrix tri-factorization (SODE-WNMTF), aimed at addressing the challenge of outliers in the unmixing procedure. The primary contributions of this research are as listed below:

1. A novel weighting matrix, leveraging the co-clustering property of NMTF, is introduced to effectively mitigate the negative impact of both pixel and band outliers during the unmixing process. Simultaneously, the underlying structure of HSI is exploited by considering both local and global relationships between pixels and spectral bands, thereby enhancing the co-clustering process.
2. The ℓ_1 sparse unmixing framework is enhanced by incorporating a spatial weighting factor that leverages the spatial correlation between adjacent pixels, promoting piecewise smoothness in the estimated abundances. Meanwhile, information extracted from the previous stage is utilized to mitigate the detrimental influence of outliers during the sparsification process.
3. The proposed method utilizes a sparse estimation technique which improves endmember estimation by focusing on combining the most representative pixel spectra from regions where each endmember is prevalent. This leads to clearer distinctions between endmember signatures. In order to assess the performance of the introduced method, a set of experiments is carried out on synthetic as well as real HSIs, demonstrating its superiority over other advanced methods.

The rest of this paper will be structured in the following manner. Section 2 provides a brief explanation of LMM and the fundamental NMF model. Sections 3 and 4 detail the proposed algorithm's model and the corresponding optimization algorithm. Section 5 discusses and evaluates the findings from both synthetic and real datasets. Section 6 presents the conclusion.

2 Background

This section begins with a concise overview of LMM. It then proceeds to a detailed examination of the standard NMF model and its well-known variant, NMTF.

2.1 Notations

In the subsequent sections, the matrices and vectors are denoted by the bold uppercase letters. For a matrix $\mathbf{B} \in \mathbb{R}^{I \times J}$, $\mathbf{B}_{i,:}$ and $\mathbf{B}_{:,j}$ represent its ith row and jth column, respectively. Moreover, the element in the ith row and jth column of $\mathbf{B}$ is denoted by $\mathbf{B}_{ij}$. The transpose of $\mathbf{B}$ is written as $\mathbf{B}^T$. A list of frequently used notations in this paper is provided in Table 1.

2.2 LMM

LMM is a fundamental tool for analyzing hyperspectral data, based on the assumption that the spectral signatures of different materials in a scene do not overlap (Miao & Qi, 2007). In this model, a mixed pixel, which contains contributions from multiple

Table 1 List of notations and corresponding definitions used in this paper

Notation	Description
$\mathbf{X}$	Hyperspectral image (HSI) matrix, $\mathbf{X} \in \mathbb{R}_+^{L \times N}$
L	Number of spectral bands in HSI
N	Total number of pixels in HSI
$\mathbf{M}$	Endmember matrix, $\mathbf{M} \in \mathbb{R}_+^{L \times K}$
$\mathbf{A}$	Spatial weighting matrix, $\mathbf{A} \in \mathbb{R}_+^{K \times N}$
K	Number of endmembers in the image
$\mathbf{E}$	Residual (error) matrix, $\mathbf{E} \in \mathbb{R}^{L \times N}$
$\mathbf{U}$	Row-coefficient matrix (spectral band clusters), $\mathbf{U} \in \mathbb{R}_+^{L \times Q}$
$\mathbf{S}$	Block matrix (feature-sample relationship), $\mathbf{S} \in \mathbb{R}_+^{Q \times K}$
$\mathbf{V}$	Column-coefficient matrix (pixel clusters), $\mathbf{V} \in \mathbb{R}_+^{K \times N}$
Q	Number of spectral band clusters
$\mathbf{T}$	Weighting matrix for outlier detection, $\mathbf{T} \in \mathbb{R}_+^{L \times N}$
$\mathbf{W}$	Sparse weight matrix for endmember estimation, $\mathbf{W} \in \mathbb{R}_+^{N \times K}$
$\mathbf{P}$	Similarity matrix for manifold learning, $\mathbf{P} \in \mathbb{R}_+^{N \times N}$
$\mathbf{L}$	Laplacian matrix, $\mathbf{L} = \mathbf{D} - \mathbf{P}$, $\mathbf{L} \in \mathbb{R}_+^{N \times N}$
$\mathbf{D}$	Degree matrix (diagonal), $\mathbf{D} \in \mathbb{R}_+^{N \times N}$
$\|\cdot\|_F$	Frobenius norm, $\|\mathbf{B}\|_F = \sqrt{\sum_{i=1}^{I} \sum_{j=1}^{J} \mathbf{B}_{ij}^2}$
$\|\cdot\|_2$	Euclidean norm, $\|\mathbf{B}_{i,:}\|_2 = \sqrt{\sum_{j=1}^{J} \mathbf{B}_{ij}^2}$
$\|\cdot\|_{1,1}$	$L_{1,1}$ norm, $\sum_{i=1}^{I} \sum_{j=1}^{J} \|\mathbf{B}_{ij}\|$
$\|\cdot\|_{\frac{1}{2}}$	$L_{\frac{1}{2}}$ norm, $\|\mathbf{B}\|_{\frac{1}{2}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sqrt{\|\mathbf{B}_{ij}\|}$
$\odot$	Element-wise multiplication
$\oslash$	Element-wise division
$\langle \cdot, \cdot \rangle$	Inner product operator
$\mathrm{Tr}(\cdot)$	Trace of a square matrix
$\mathbf{I}_Q$	$Q \times Q$ identity matrix
$\mathbb{R}_+$	Set of non-negative real numbers

materials, is represented by two key matrices: the endmember matrix, which holds the spectral signatures of the individual materials, and the abundance matrix, which reflects the proportion of each material present in the mixed pixel. Formally, LMM can be expressed as follows:

$$\mathbf{X} = \mathbf{MA} + \mathbf{E}, \tag{1}$$

where $\mathbf{M} = [\mathbf{M}_{:,1}, \mathbf{M}_{:,2}, \ldots, \mathbf{M}_{:,K}] \in {\mathbb{R}_+}^{L \times K}$ is the endmember matrix, and K is the number of endmembers in the image. The abundance matrix $\mathbf{A} = [\mathbf{A}_{:,1}, \mathbf{A}_{:,2}, \ldots, \mathbf{A}_{:,N}] \in {\mathbb{R}_+}^{K \times N}$ contains the abundance vectors, with each column corresponding to the abundance of materials in the ith pixel. Lastly, $\mathbf{E} \in \mathbb{R}^{L \times N}$ represents the residual matrix.

It is worth noting that two constraints must be applied to the abundance vectors to ensure they are physically meaningful: the abundance non-negativity constraint (ANC), which is described as $\mathbf{A}_{kn} \geq 0$, where $k = 1, 2, \ldots, K$ and $n = 1, 2, \ldots, N$, and the abundance sum-to-one constraint (ASC), which is given by $\sum_{k=1}^{K} \mathbf{A}_{kn} = 1$.

2.3 NMF

NMF has emerged as a powerful method for HU thanks to its ability to decompose complex data into meaningful components. It aligns well with LMM by breaking down the hyperspectral data into two non-negative matrices: one representing the endmembers and the other their abundances. Given the data matrix $\mathbf{X} \in \mathbb{R}_+^{L\times N}$, the goal of NMF is to discover two non-negative matrices, $\mathbf{M} \in \mathbb{R}_+^{L\times K}$ and $\mathbf{A} \in \mathbb{R}_+^{K\times N}$, that minimize the reconstruction error based on the Frobenius norm:

$$\min_{\mathbf{M}\geq 0,\mathbf{A}\geq 0} \|\mathbf{X}-\mathbf{M}\mathbf{A}\|_F^2, \tag{2}$$

in which the constraints $\mathbf{M} \geq 0$ and $\mathbf{A} \geq 0$ are applied to ensure the non-negativity of the factored matrices. One widely-used method for addressing the optimization problem of NMF is Multiplicative Update Rule (MUR), proposed in Lee and Seung (1999), by applying the following update rules:

$$\mathbf{M} \leftarrow \mathbf{M} \odot (\mathbf{X}\mathbf{A}^T) \oslash (\mathbf{M}\mathbf{A}\mathbf{A}^T), \tag{3}$$

$$\mathbf{A} \leftarrow \mathbf{A} \odot (\mathbf{M}^T\mathbf{X}) \oslash (\mathbf{M}^T\mathbf{M}\mathbf{A}). \tag{4}$$

However, minimizing (2) is challenging due to its simultaneous non-convexity with respect to both $\mathbf{M}$ and $\mathbf{A}$, making standard NMF solutions susceptible to becoming trapped in local minima. To address this, researchers have developed several extensions of NMF, such as weighted NMF, which often improve stability, promote sparsity, and more effectively handle complex data patterns by integrating additional constraints and regularization techniques.

2.4 NMTF

NMTF, an extension of traditional NMF, has gained prominence in data dimensionality reduction, particularly in the analysis of dyadic data, which explores relationships between two distinct sets of items. In the context of HU, NMTF is particularly valuable for its ability to perform co-clustering, which enables the simultaneous clustering of both pixels and spectral bands. This dual clustering capability offers a more detailed decomposition of hyperspectral data. In NMTF, the data matrix $\mathbf{X} \in \mathbb{R}_+^{L\times N}$ can be approximated by the product of three non-negative matrices in the form $\mathbf{X} \approx \mathbf{U}\mathbf{S}\mathbf{V}$, which is equivalent to writing that

$$\begin{aligned} &\min_{\mathbf{U},\mathbf{S},\mathbf{V}\geq 0} \|\mathbf{X}-\mathbf{U}\mathbf{S}\mathbf{V}\|_F^2 \\ &\text{s.t. } \mathbf{U}^T\mathbf{U}=\mathbf{I}_Q \ , \ \mathbf{V}\mathbf{V}^T=\mathbf{I}_K. \end{aligned} \tag{5}$$

Here, $\mathbf{U} \in \mathbb{R}_+^{L\times Q}$ represents the row-coefficient matrix providing information about the feature (band) clusters. $\mathbf{S} \in \mathbb{R}_+^{Q\times K}$ denotes the block matrix illustrating the relationships between the feature and data sample clusters, and $\mathbf{V} \in \mathbb{R}_+^{K\times N}$ stands for the column-coefficient matrix that conveys details about the sample (pixel) clusters (Ding et al., 2006; Dache et al., 2024). It is worthwhile to mention that the necessity of meaningful clusters relies on the orthogonality constraints. The orthogonality constraint on $\mathbf{U}$ guarantees that the columns of $\mathbf{U}$ are orthogonal to each other, resulting in unique clusterings of spectral bands. In the same way, imposing an orthogonal constraint on $\mathbf{V}$ causes its rows to become

orthogonal, leading to distinctive clusterings of pixels. This enhances the interpretability by preventing redundancy in cluster assignments and promoting clear separations between materials. However, a potential trade-off of these constraints is the reduction in flexibility. Specifically: 1. In some cases, enforcing strict orthogonality may limit the ability of the model to capture overlapping structures in hyperspectral data. Some materials may share spectral similarities, and relaxing the orthogonality constraint could allow for more nuanced representations. 2. The orthogonality constraints introduce additional complexity in optimization, which may lead to slower convergence or sub-optimal local minima if not carefully handled. 3. Strict orthogonality may make the model more sensitive to noise in real-world datasets, as it prioritizes distinct clusters over potentially useful shared features. To balance these trade-offs, we ensure that the constraints are applied in a way that maintains interpretability while preserving sufficient flexibility for meaningful decompositions.

Another important aspect is that in NMTF, the true reconstruction basis for each pixel–cluster is **US**, rather than **U** alone. This modification yields the following significant advantages.

1. By introducing the intermediate block matrix **S**, one can choose $Q \neq K$, allowing for different numbers of spectral–band clusters and pixel–clusters. This flexibility is critical in applications such as HU, where the intrinsic dimensionality of the spectral bands often differs from that of the spatial pixels. Therefore, the extra factor **S** provides additional degrees of freedom to achieve a low-rank approximation without sacrificing cluster purity, often improving reconstruction error compared to standard NMF with the same total parameter count.
2. Enforcing the orthogonality constraints $\mathbf{U}^T\mathbf{U} = \mathbf{I}_Q$ and $\mathbf{V}\mathbf{V}^T = \mathbf{I}_K$ ensures that **U** encodes a hard assignment of each spectral band to one of Q band clusters, while **V** assigns each pixel to one of K pixel clusters. The entries of **S** then quantify the strength of interaction between each band cluster and each pixel cluster. Consequently, **US** yields K prototypical spectra as linear combinations of band-cluster centroids, directly reflecting the block structure of **X**.

As a result, regarding **US** rather than **U** as the true basis in NMTF is a design that (i) decouples the dimensionalities of band and pixel clusters, (ii) explicitly captures inter-cluster affinities, and (iii) yields an immediately interpretable co-clustered representation of hyperspectral data.

3 Proposed model

This section begins by introducing the weighted NMTF method, which forms the foundation of our framework for concurrently detecting and eliminating outliers during the unmixing process. Subsequently, a detailed explanation of the regularization terms employed in the proposed model will be provided.

3.1 Weighted NMTF for outlier detection

In light of the distinctive co-clustering advantage of NMTF, a novel outlier-resistant unmixing framework is proposed in this section, that simultaneously detects and alleviates the adverse effects of outliers in both pixels and spectral bands during the unmixing

process. To accomplish this, the distance of each data point from its respective cluster center is calculated, with outliers being recognized as points that are significantly farther from their cluster center compared to others. These distance measures are then used to introduce a weight matrix $\mathbf{T}$ into the NMTF model. By incorporating this weight matrix, the influence of each pixel and spectral band is adjusted, thereby enhancing the robustness of the unmixing process. At the $(t+1)$th iteration of the algorithm, the elements of the weighting matrix $\mathbf{T} = [\mathbf{T}_{ln}] \in \mathbb{R}_+^{L\times N}$ are specified as follows:

$$\mathbf{T}_{ln}^{(t+1)} = \frac{\mu_1}{\|\mathbf{X}_{l,:} - (\mathbf{SV})_{\omega_l,:}^{(t)}\|_2} \times \frac{\mu_2}{\|\mathbf{X}_{:,n} - (\mathbf{US})_{:,\phi_n}^{(t)}\|_2}, \tag{6}$$

for $l = 1, 2, \ldots, L$ and $n = 1, 2, \ldots, N$. Here, $\mathbf{X}_{l,:}$ and $(\mathbf{SV})_{\omega_l,:}^{(t)}$ represent the lth band and its corresponding cluster center during the algorithm's tth iteration, respectively, while $\mathbf{X}_{:,n}$ and $(\mathbf{US})_{:,\phi_n}^{(t)}$ represent the nth pixel and its corresponding cluster center at the tth iteration of the algorithm. In addition, ω_l and ϕ_n are the cluster indices associated with the lth band and the nth pixel, respectively. The parameters μ_1 and μ_2 control the relative importance of band and pixel outliers. With this description, it can be inferred from (6) that the lower weights indicate a higher likelihood of being an outlier, leading to reduced impact on the HU results.

By integrating the suggested weighting matrix into the NMTF framework, the objective function of the weighted NMTF can be formulated as follows:

$$\begin{aligned} &\min_{\mathbf{U},\mathbf{S},\mathbf{V}\geq 0} \|\mathbf{T} \odot (\mathbf{X} - \mathbf{USV})\|_F^2 \\ &\text{s.t. } \mathbf{U}^T\mathbf{U} = \mathbf{I}_Q,\ \mathbf{VV}^T = \mathbf{I}_K. \end{aligned} \tag{7}$$

3.2 Improved endmember estimation via sparse pixel combination

Given the inherent characteristic of HSIs, where pixels in areas mainly affected by a single endmember tend to have spectra similar to that endmember, it can be suggested that each endmember signature can be effectively estimated using only a subset of pixel spectra from areas dominated by that specific endmember. This pixel selective approach not only enhances the accuracy of endmember extraction but also provides a more resilient estimate. This is achieved by reducing the influence of mixed pixels located in regions where multiple materials coexist. The mathematical formulation of this strategy can be demonstrated through the following expression:

$$(\mathbf{US})_{:,k} \approx \mathbf{XW}_{:,k}, \tag{8}$$

for $k = 1, 2, \ldots, K$. In (8), $(\mathbf{US})_{:,k}$ represents the kth endmember signature, and $\mathbf{W}_{:,k}$ represents the kth column of the sparse weight matrix $\mathbf{W} \in \mathbb{R}_+^{N\times K}$ that assigns the contribution of each pixel spectrum to the reconstruction of the kth endmember signature.

To further enhance the NMTF framework with the improved endmember estimation discussed in (8), two additional regularization terms are proposed as follows:

$$\mathbf{\Omega}_1(\mathbf{U}, \mathbf{S}, \mathbf{W}) = \|\mathbf{US} - \mathbf{XW}\|_F^2 + \|\mathbf{W}\|_{\frac{1}{2}}. \tag{9}$$

Here, the expression $\|\mathbf{W}\|_{\frac{1}{2}}$ based on the $L_{\frac{1}{2}}$ norm promotes the sparsity in $\mathbf{W}$. This sparsity is driven by the fact that each endmember spectral signature is associated with a limited number of distinct pixel signatures. Here, the $L_{\frac{1}{2}}$ norm was chosen over the L_1 norm due to its stronger sparsity-promoting property. While the L_1 norm encourages sparsity by linearly penalizing nonzero elements, the $L_{\frac{1}{2}}$ norm imposes a more aggressive penalty on small values, pushing more elements to exact zeros. It should be noted that, while this suggests an ideal scenario with many pure pixels, our model does not strictly require the presence of pure pixels in the scene. Instead, it leverages the sparsity-inducing properties of $\mathbf{W}$. In the absence of pure pixels, this sparsity constraint biases the algorithm toward selecting the least-mixed pixels available, those with the highest relative contribution from a single endmember.

3.3 Regularization strategies for weighted NMTF

As discussed above, the accuracy of the outlier detection and elimination process is highly dependent on the precision of the NMTF clustering procedure. Therefore, improving the estimation of $\mathbf{U}, \mathbf{S}$ and $\mathbf{V}$ matrices is essential in this context, which can be accomplished by leveraging the inherent structure of hyperspectral data. For further explanation, one approach to enhance the accuracy of the assessment of the endmember matrix $\mathbf{US}$ is by improving the precision of the weighting matrix $\mathbf{W}$ estimation, which involves maintaining the overall similarity structure of spectral bands in the lower-dimensional space. Therefore, considering that the rows of the endmember matrix $\mathbf{US}$ represent the lower-dimensional embeddings of the spectral bands, it can be deduced from the previous section that the rows of $\mathbf{XW}$ convey a similar idea. As a result, the spectral correlation between the rows in $\mathbf{XW}$ should reflect the correlation patterns observed in the corresponding rows of $\mathbf{X}$. This concept can be illustrated as:

$$\langle \mathbf{X}_{i,:}, \mathbf{X}_{j,:} \rangle \approx \langle (\mathbf{XW})_{i,:}, (\mathbf{XW})_{j,:} \rangle, \tag{10}$$

for $i, j = 1, 2, \ldots, L$. Here, the expression $\langle \mathbf{X}_{i,:}, \mathbf{X}_{j,:} \rangle$ serves as a metric for measuring the spectral similarity between two spectral bands. This spectral correlation idea can be incorporated into our weighted NMTF framework by introducing a novel regularization term as follows:

$$\mathbf{\Omega}_2(\mathbf{W}) = \|\mathbf{X}\mathbf{X}^T - \mathbf{X}\mathbf{W}\mathbf{W}^T\mathbf{X}^T\|_F^2. \tag{11}$$

In this case, $\mathbf{X}\mathbf{X}^T$ denotes the spectral correlation matrix of $\mathbf{X}$, representing the similarities between the spectral bands in the original space. Furthermore, $\mathbf{X}\mathbf{W}\mathbf{W}^T\mathbf{X}^T$ conveys the same concept in the embedded space.

In addition, a promising approach to improve the estimation of the abundance matrix $\mathbf{V}$ is through manifold learning, which seeks to preserve the inherent geometric structure of high-dimensional data when mapped to a lower-dimensional space. The core idea is to construct a nearest-neighbor graph that captures the local geometric relationships within the data (Lu et al., 2013). By incorporating this graph-based regularization technique, the unmixing process is guided to ensure that the abundance maps retain the local neighborhood structures observed in the original hyperspectral data. This nearest-neighbor graph can be expressed through a similarity matrix $\mathbf{P} = [\mathbf{P}_{ij}] \in \mathbb{R}_+^{N \times N}$, typically constructed using methods such as the heat kernel, formulated as:

$$\mathbf{P}_{ij} = \exp\left(-\frac{\|x_i - x_j\|_2^2}{\sigma}\right), \tag{12}$$

for $i,j = 1,2,\ldots,N$. It is evident that when $\mathbf{X}_{:,i}$ and $\mathbf{X}_{:,j}$ are in close proximity, the value of $\mathbf{P}_{ij}$ is relatively large. According to this rationale, the central assumption in manifold learning is that if $\mathbf{X}_{:,i}$ and $\mathbf{X}_{:,j}$ are neighbors in the original high-dimensional space, their corresponding low-dimensional embeddings, $\mathbf{V}_{:,i}$ and $\mathbf{V}_{:,j}$, should also be positioned close to each other in the lower-dimensional space. To enforce this assumption, a graph regularization term can be integrated into the proposed framework as follows:

$$\mathbf{\Omega}_3(\mathbf{V}) = \sum_{i,j=1}^{N} \|\mathbf{V}_{:,i} - \mathbf{V}_{:,j}\|_2^2 \mathbf{P}_{ij} = Tr(\mathbf{V}\mathbf{L}\mathbf{V}^T). \tag{13}$$

Here, $\mathbf{L} = \mathbf{D} - \mathbf{P}$ is the Laplacian matrix, where $\mathbf{D} \in \mathbb{R}_+^{N\times N}$ is the degree matrix, a diagonal matrix with entries $\mathbf{D}_{ii} = \sum_{j=1}^{N} \mathbf{P}_{ij}$ on the diagonal.

3.4 Robust spatial regularization for abundance maps

One of the main features of HSIs is the sparse nature of their abundance maps. This indicates that only the endmember signatures present in a pixel contribute to forming its spectral signature when the spectrum of pixel is expressed as a linear combination of endmember signatures. Given the importance of ensuring that the abundance matrix remains sparse, the objective function of weighted NMTF in (7) incorporates an orthogonality constraint on the abundance matrix $\mathbf{V}$ to encourage its sparsity. However, in traditional LMM, abundance vectors represent fractional contributions of endmembers, and these are not inherently orthogonal across endmembers. Furthermore, this approach does not account for the spatial relationships between neighboring pixels, which often exhibit similar abundance patterns. To address this limitation, we propose a new constraint for $\mathbf{V}$ that builds upon the spatial weighting matrices concept from prior research by Zhang et al. (2022). This new constraint employs a weighting factor derived from outlier probabilities, enabling the regularization to adapt to data quality in specific regions. Consequently, it reduces the influence of potential outliers while still leveraging spatial information when necessary. Accordingly, our alternative sparsity constraint is defined as follows:

$$\mathbf{\Omega}_4(\mathbf{V}) = \sum_{k=1}^{K}\sum_{n=1}^{N} \mathbf{A}_{kn}\mathbf{V}_{kn} = \|\mathbf{A} \odot \mathbf{V}\|_{1,1}. \tag{14}$$

In this context, $\mathbf{A} = [\mathbf{A}_{kn}] \in \mathbb{R}_+^{K\times N}$ represents the spatial weighting matrix. Each element $\mathbf{A}_{kn}$ corresponds to the abundance $\mathbf{V}_{kn}$ of the kth endmember within the nth pixel and is determined during the iteration $t+1$ of the algorithm as follows:

$$\mathbf{A}_{kn}^{(t+1)} = \left[\frac{1}{\sum_{p\in\mathcal{V}(n)} \tau_p^{(t)}} \sum_{p\in\mathcal{V}(n)} \tau_p^{(t)}\mathbf{V}_{kp}^{(t)} + \epsilon,\right]^{-1} \tag{15}$$

where ϵ is a small non-negative value to avoid numerical instability. Moreover, $\nu(n)$ denotes the set of indices corresponding to the spatial neighborhood of pixel n, while α_p is a coefficient that measures the probability that pixel p is an outlier. It is defined as follows:

$$\tau_p^{(t)} = \|\mathbf{T}_{:,p}^{(t)}\|_2, \tag{16}$$

where $\mathbf{T}_{:,p}^{(t)}$ is the pth column of the matrix $\mathbf{T}$ defined in (6) at the tth iteration. It is worth highlighting that the purpose of using the coefficients τ_p in the elements of the spatial weighting matrix $\mathbf{A}$ is to reduce the impact of neighboring pixels with high outlier probabilities. Consequently, the regularization term $\mathbf{\Omega}_4(\mathbf{V})$, as defined in (14), can adapt to the quality of local data, promoting spatial smoothness in stable regions while minimizing its effect in areas with potential anomalies.

It is important to note that the scene may contain small, scattered yet significant objects like rare minerals, which this method might overshadow. To preserve the spectral information of these rare minerals, we can lessen the impact of (14) by decreasing its associated hyperparameter in the final objective function. This modification enhances the model's sensitivity to sudden abundance variations, thereby maintaining the spectral signatures of specific localized features.

3.5 Final objective function

Our proposed model in this paper, SODE-WNMTF, is achieved through the unification of four regularization terms with the weighted NMTF framework: (1) $\mathbf{\Omega}_1$ (to improve endmember extraction), (2) $\mathbf{\Omega}_2$ (to preserve the overall similarity structure of the bands), (3) $\mathbf{\Omega}_3$ (for manifold learning), and (4) $\mathbf{\Omega}_4$ (to embed robustness within the ℓ_1 sparse unmixing model). The objective function for the SODE-WNMTF model is formally presented as follows:

$$\begin{aligned} \min_{\mathbf{U},\mathbf{S},\mathbf{V},\mathbf{W}} & \frac{1}{2}\|\mathbf{T} \odot (\mathbf{X} - \mathbf{USV})\|_F^2 + \frac{\alpha_1}{2}\|\mathbf{US} - \mathbf{XW}\|_F^2 + 2\alpha_2\|\mathbf{W}\|_{\frac{1}{2}} \\ & + \frac{\alpha_3}{4}\|\mathbf{XX}^T - \mathbf{XWW}^T\mathbf{X}^T\|_F^2 + \alpha_4\|\mathbf{A} \odot \mathbf{V}\|_{1,1} + \alpha_5 Tr(\mathbf{VLV}^T) \\ \text{s.t.} \quad & \mathbf{U}^T\mathbf{U} = \mathbf{I}_Q, \quad \mathbf{U}, \mathbf{S}, \mathbf{V}, \mathbf{W} \geq 0. \end{aligned} \tag{17}$$

The regularization term parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and α_5 are used to adjust the impact of each constraint item on the model. The more significant the values are, the more important their related constraint is in the objective function.

4 Optimization

The non-convex nature of the SODE-WNMTF optimization problem established in (17) with respect to $\mathbf{U}$, $\mathbf{S}$, $\mathbf{V}$, and $\mathbf{W}$ jointly indicates that finding a global optimal solution is impractical. Instead, the MUR method can be employed to iteratively optimize the objective function of SODE-WNMTF and achieve a local minimum. In this approach, MUR fixes all matrices except one during each iteration, optimizing the remaining matrix until convergence is reached. To this end, the associated Lagrange function Γ is defined as follows:

$$\begin{aligned}\Gamma = &\frac{1}{2}\|\mathbf{T}\odot(\mathbf{X}-\mathbf{USV})\|_F^2+\frac{\alpha_1}{2}\|\mathbf{US}-\mathbf{XW}\|_F^2+2\alpha_2\|\mathbf{W}\|_{\frac{1}{2}}\\&+\frac{\alpha_3}{4}\|\mathbf{XX}^T-\mathbf{XWW}^T\mathbf{X}^T\|_F^2+\alpha_4\|\mathbf{A}\odot\mathbf{V}\|_{1,1}\\&+\alpha_5\,Tr(\mathbf{VLV}^T)+\frac{\alpha_6}{2}\|\mathbf{U}^T\mathbf{U}-\mathbf{I}_Q\|_F^2+Tr(\boldsymbol{\beta}\mathbf{U}^T)\\&+Tr(\boldsymbol{\gamma}\mathbf{S}^T)+Tr(\boldsymbol{\rho}\mathbf{V}^T)+Tr(\boldsymbol{\lambda}\mathbf{W}^T),\end{aligned}\tag{18}$$

where $\boldsymbol{\beta}\in\mathbb{R}_+^{L\times Q},\boldsymbol{\gamma}\in\mathbb{R}_+^{Q\times K},\boldsymbol{\rho}\in\mathbb{R}_+^{K\times N}$ and $\boldsymbol{\lambda}\in\mathbb{R}_+^{N\times K}$ denote the corresponding Lagrange multipliers of $\mathbf{U},\mathbf{S},\mathbf{V}$ and $\mathbf{W}$, respectively. The update rules for $\mathbf{U},\mathbf{S},\mathbf{V}$ and $\mathbf{W}$ are formulated as follows:

1. Optimization of $\mathbf{W}$: The partial derivative of (18) w.r.t. $\mathbf{W}$ is calculated as:

$$\begin{aligned}\frac{\partial\Gamma}{\partial\mathbf{W}} = &-\alpha_1\mathbf{X}^T\mathbf{US}+\alpha_1\mathbf{X}^T\mathbf{XW}+\alpha_2\mathbf{W}^{-\frac{1}{2}}\\&-\alpha_3\mathbf{X}^T\mathbf{XX}^T\mathbf{XW}+\alpha_3\mathbf{X}^T\mathbf{XWW}^T\mathbf{X}^T\mathbf{XW}+\boldsymbol{\lambda}.\end{aligned}$$

By combining the condition $\frac{\partial\Gamma}{\partial\mathbf{W}}=0$ with the KKT conditions ($\boldsymbol{\lambda}\odot\mathbf{W}=0$), our conclusion is:

$$\begin{aligned}\mathbf{W}\leftarrow&\mathbf{W}\odot(\alpha_1\mathbf{X}^T\mathbf{US}+\alpha_3\mathbf{X}^T\mathbf{XX}^T\mathbf{XW})\\&\oslash(\alpha_1\mathbf{X}^T\mathbf{XW}+\alpha_2\mathbf{W}^{-\frac{1}{2}}+\alpha_3\mathbf{X}^T\mathbf{XWW}^T\mathbf{X}^T\mathbf{XW}).\end{aligned}\tag{19}$$

2. Optimization of $\mathbf{U}$: The partial derivative of (18) w.r.t. $\mathbf{U}$ can be obtained as:

$$\begin{aligned}\frac{\partial\Gamma}{\partial\mathbf{U}} = &-(\mathbf{T}\odot\mathbf{T}\odot\mathbf{X})\mathbf{V}^T\mathbf{S}^T+(\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV})\mathbf{V}^T\mathbf{S}^T\\&+\alpha_1\mathbf{USS}^T-\alpha_1\mathbf{XWS}^T+\alpha_6\mathbf{UU}^T\mathbf{U}-\alpha_6\mathbf{U}+\boldsymbol{\beta}.\end{aligned}$$

By using $\frac{\partial\Gamma}{\partial\mathbf{U}}=0$ and utilizing the KKT condition ($\boldsymbol{\beta}\odot\mathbf{U}=0$), it implies that:

$$\begin{aligned}\mathbf{U}\leftarrow&\mathbf{U}\odot((\mathbf{T}\odot\mathbf{T}\odot\mathbf{X})\mathbf{V}^T\mathbf{S}^T+\alpha_1\mathbf{XWS}^T+\alpha_6\mathbf{U})\\&\oslash((\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV})\mathbf{V}^T\mathbf{S}^T+\alpha_1\mathbf{USS}^T+\alpha_6\mathbf{UU}^T\mathbf{U}).\end{aligned}\tag{20}$$

3. Optimization of $\mathbf{S}$: The partial derivative of (18) w.r.t. $\mathbf{S}$ can be calculated as:

$$\begin{aligned}\frac{\partial\Gamma}{\partial\mathbf{S}} = &-\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{X})\mathbf{V}^T+\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV})\mathbf{V}^T\\&+\alpha_1\mathbf{U}^T\mathbf{US}-\alpha_1\mathbf{U}^T\mathbf{XW}+\boldsymbol{\gamma}.\end{aligned}$$

Through setting $\frac{\partial\Gamma}{\partial\mathbf{S}}=0$ and applying the KKT condition ($\boldsymbol{\gamma}\odot\mathbf{S}=0$), it suggests that:

$$\begin{aligned}\mathbf{S}\leftarrow&\mathbf{S}\odot(\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{X})\mathbf{V}^T+\alpha_1\mathbf{U}^T\mathbf{XW})\\&\oslash(\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV})\mathbf{V}^T+\alpha_1\mathbf{U}^T\mathbf{US}).\end{aligned}\tag{21}$$

4. Optimization of $\mathbf{V}$: The partial derivative of (18) w.r.t. $\mathbf{V}$ is denoted as:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \mathbf{V}} &= -\mathbf{S}^T\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{X}) + \mathbf{S}^T\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV}) \\ &\quad + \alpha_4\mathbf{A} + \alpha_5\mathbf{VD} - \alpha_5\mathbf{VP} + \boldsymbol{\rho}.\end{aligned} \tag{22}$$

Combining the condition $\frac{\partial \Gamma}{\partial \mathbf{V}} = 0$ with the KKT condition ($\boldsymbol{\rho}\odot\mathbf{V} = 0$), it is concluded that:

$$\begin{aligned}\mathbf{V} \leftarrow \mathbf{V} \odot (&\mathbf{S}^T\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{X}) + \alpha_5\mathbf{VP}) \\ \oslash (&\mathbf{S}^T\mathbf{U}^T(\mathbf{T}\odot\mathbf{T}\odot\mathbf{USV}) + \alpha_5\mathbf{VD} + \alpha_4\mathbf{A}).\end{aligned} \tag{23}$$

In the end, the optimization algorithm for the SODE-WNMTF method is summarized in Algorithm 1.

Algorithm 1 Iterative algorithm for SODE-WNMTF.

Input:
The data matrix $\mathbf{X} \in {\mathbb{R}_+}^{L\times N}$; the number of endmembers K; the parameters $\alpha_1, \ldots, \alpha_6$.

1: Initialize $\mathbf{US}$ and $\mathbf{V}$ using the VCA and FCLS methods [9], respectively; Construct $\mathbf{S}$ by the random initialization and then obtain $\mathbf{U}$ from $\mathbf{US} = \mathbf{M}$, where $\mathbf{M}$ is the matrix of endmembers derived from VCA.
2: **repeat**:
3: Update the weighting matrices $\mathbf{T}$, $\mathbf{A}$, and $\mathbf{W}$ by (6), (15) and (19), respectively;
4: Update the non-negative matrices $\mathbf{U}$, $\mathbf{S}$, and $\mathbf{V}$ by (20), (21), and (23), respectively;
5: **until** convergence

Output: The endmember matrix $\mathbf{US} \in {\mathbb{R}_+}^{L\times K}$ and the abundance matrix $\mathbf{V} \in {\mathbb{R}_+}^{K\times N}$.

4.1 Implementation issues

When implementing the unmixing algorithm, several problems arise. Here, various concerns related to the implementation of the unmixing algorithm will be addressed. According to the procedure of the proposed algorithm, the initial step is related to the techniques used to initialize the matrices $\mathbf{U}, \mathbf{S}$ and $\mathbf{V}$. Because it is unrealistic to achieve a global minimum by optimizing the non-convex objective function given in (17), various initial values will result in different outcomes. In this paper, vertex component analyze (VCA) (Nascimento & Dias, 2005) and fully constrained least squares (FCLS) are used to initialize $\mathbf{US}$ and $\mathbf{V}$, respectively, as it is a popular method in HU to provide robust endmember estimation, improved abundance estimation, and faster convergence.

The next important issue is to ensure ANC and ASC throughout the iterative process. If the initial values are non-negative, ANC will be strictly guaranteed based on the aforementioned updating rules. The approach used here involves the effective method from Heinz (2001) with regards to ASC, i.e, $\mathbf{1}_K^T\mathbf{V} = \mathbf{1}_N^T$. When the matrix $\mathbf{V}$ is updated using (23), the inputs $\mathbf{X}$ and $\mathbf{M} = \mathbf{US}$ will be replaced by the matrices $\bar{\mathbf{X}}$ and $\bar{\mathbf{M}}$, respectively, as defined below:

$$\bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \delta \mathbf{1}_N^T \end{bmatrix}, \quad \bar{\mathbf{M}} = \begin{bmatrix} \mathbf{M} \\ \delta \mathbf{1}_K^T \end{bmatrix}, \tag{24}$$

where δ modifies the impact of ASC. In reality, a higher δ results in a more precise outcome but at the cost of a slower convergence. For the experiments, a relatively small $\delta = 15$ is chosen to reach a desired trade-off. Another problem arises from the sparsity norm $L_{\frac{1}{2}}$, particularly the term $\mathbf{W}^{-\frac{1}{2}}$ in (19). When a zero entry appears in $\mathbf{W}$, a small positive number will be added to make the operation valid. Moreover, as stated in Qian et al. (2011), the elements in $\mathbf{W}$ that are below 10^{-4} will be adjusted using (19) excluding the term $\alpha_2 \mathbf{W}^{-\frac{1}{2}}$. By doing this, the robustness of the algorithm will be improved. The issue of stopping criteria is also worth mentioning. To do so, a predetermined value is established as the error tolerance for the suggested updating rules. If the error does not exceed the tolerance for ten successive times, the iteration will stop. The maximum number of iterations is also used as a stopping criterion. In the experiments, the number is fixed at 3000 to ensure convergence reliably. The final concern revolves around identifying the quantity of endmembers. The number of endmembers can be identified either by manual methods or by utilizing established approaches such as HySime (Bioucas-Dias & Nascimento, 2008).

4.2 Computational complexity analysis

In this context, we will examine the computational complexity of SODE-WNMTF. In order to improve clarity in the analysis, the time taken for floating point calculations has been documented for each iteration. Based on the aforementioned MURs, it is important to note that the matrix $\mathbf{P}$ is sparse. Typically, each row of $\mathbf{P}$ has only k non-zero elements. So, performing floating-point addition and multiplication in computing $\mathbf{VP}$ actually costs kKN times. The same applies to $\mathbf{VD}$. Moreover, calculating $W^{-\frac{1}{2}}$ requires a cost of $(NK)^2$. It is important to remember that an additional $O(N^2L)$ is required initially to construct the k-nearest neighbor graph. Therefore, if the iteration stops at the mth step, the total cost is approximately $O(N^2(m(K^2 + LK) + L))$.

5 Experimental results

This section presents empirical evaluations conducted on both synthetic and real-world datasets to assess the efficacy and validate the performance of the proposed model SODE-WNMTF which is compared with VCA-FCLS (Nascimento & Dias, 2005), $L_{\frac{1}{2}}$-NMF (Qian et al., 2011), GLNMF (Lu et al., 2013), projection-based NMF (PNMF) (Yuan et al., 2015), HGNMF-FS (Qu et al., 2023), SSWNMF (Zhang et al., 2022), WRNMF (Ince & Dobigeon, 2022), R-CoNMF (Li et al., 2016) and total variation regularized non-negative tensor factorization (MV-NTF-TV) (Xiong et al., 2019). Additionally, two recent deep learning-based methods, the global–local smoothing autoencoder (GLA) (Xu et al., 2022a) and the $L_{2,1}$ norm-based robust deep nonnegative matrix factorization ($L_{2,1}$-RDNMF) (Huang et al., 2023), are evaluated on real-world datasets to assess the effectiveness of SODE-WNMTF in comparison with deep learning models. There are two metrics used to assess the precision of the extracted endmembers and their abundances, which are the spectral angle distance (SAD) and the root mean square error (RMSE) (Wang et al., 2016), respectively. These two metrics are defined as follows:

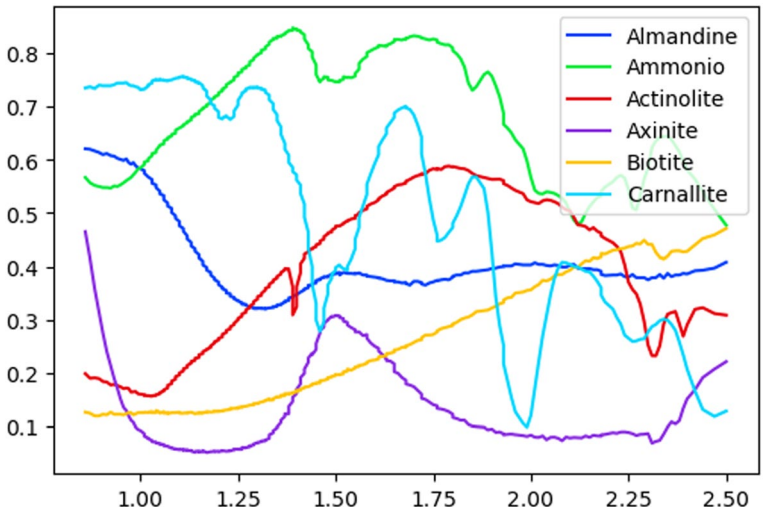

Fig. 1 Spectra of endmembers utilized in the synthetic dataset

$$\text{SAD} = \sum_{k=1}^{K} \arccos\left(\frac{\mathbf{M}_{:,i}^{T}\tilde{\mathbf{M}}_{:,i}}{\|\mathbf{M}_{:,i}\|_2\|\tilde{\mathbf{M}}_{:,i}\|_2}\right), \tag{25}$$

$$\text{RMSE} = \left(\frac{1}{N}\sum_{n=1}^{N}\|\mathbf{V}_{:,i} - \tilde{\mathbf{V}}_{:,i}\|_2^2\right)^{\frac{1}{2}}, \tag{26}$$

where $\mathbf{M}_{:,i}$ and $\tilde{\mathbf{M}}_{:,i}$ denote the estimated endmember vector and its corresponding ground-truth endmember for the ith estimate, respectively. Similarly, $\mathbf{V}_{:,i}$ and $\tilde{\mathbf{V}}_{:,i}$ represent the estimated abundance vector and its corresponding true abundance for the same estimate.

5.1 Experiments on synthetic data

In this experiment, six spectra from the U.S. geological survey (USGS) digital spectral library[1] are randomly selected to create synthetic data, as illustrated in Fig. 1. The process follows a similar method as Miao and Qi (2007): (1) Dividing a 64×64 image into 8×8 patches; (2) Filling each patch with a randomly chosen signature from the six options; (3) Using an 8×8 low-pass filter to create mixed pixels; (4) If an abundance of a pixel is over 0.8, it is replaced with a mixed combination of all endmembers, with each contributing an equal fraction of $\frac{1}{K}$; and (5) Synthesized data is mixed with Gaussian white noise, with a signal-to-noise ratio (SNR) of 20 dB being applied.

1. Hyperparameter Evaluation

[1] http://speclab.cr.usgs.gov/spectral.lib06.

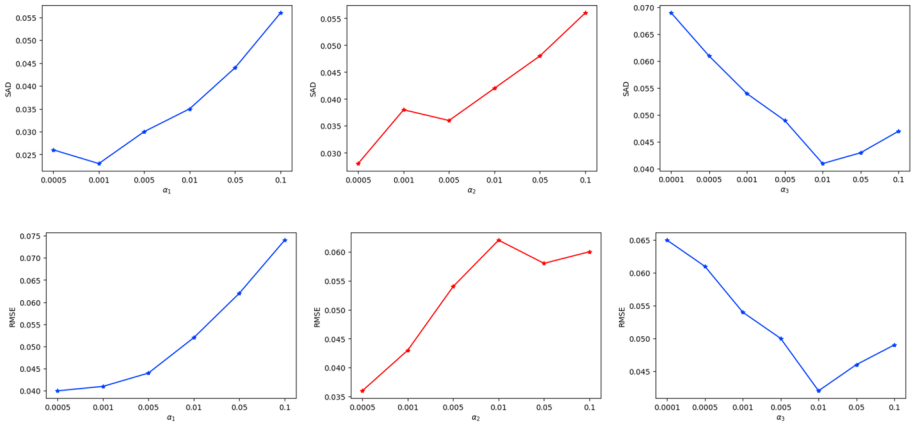

Fig. 2 Performance evaluation of the suggested approach regarding parameters μ_1 and μ_2 based on SAD and RMSE

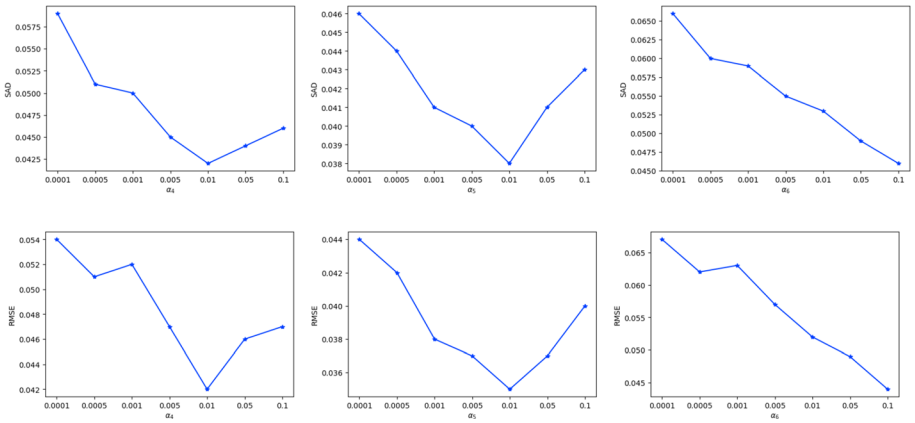

Fig. 3 Performance evaluation of the suggested approach regarding parameters α_1, α_2 and α_3 based on SAD and RMSE

In this experiment, the consideration is given to the six important hyperparameters $\alpha_1, \alpha_2, \ldots, \alpha_6$ used in the objective function (17) under the condition of SNR being equal to 20 dB. For simplicity, the fixed values of $\mu_1 = 0.5$ and $\mu_2 = 5$ are utilized throughout the experiment, based on Fig. 2. These values effectively mitigate the influence of outliers without requiring dataset-specific adjustments. The explanation focuses solely on the adjustments of parameters α_1 and α_2 The examination of the parameter α_1 is initially conducted while maintaining the values of the other tuning parameters constant. The relationship between α_1 and the performance metrics SAD and RMSE is depicted in Fig. 3, where it is observed that both metrics remain relatively stable for the small values of α_1 but increase rapidly for the values exceeding 0.005. To optimize performance, a value of α_1 should be selected that minimizes both SAD and RMSE. Accordingly, the preferable parameter value for α_1 seems to be 0.001, which offers the lowest SAD and maintains the RMSE very close to its minimum. On the other hand, a distinct trend is observed for α_2 from Fig. 3, where both SAD and RMSE consistently

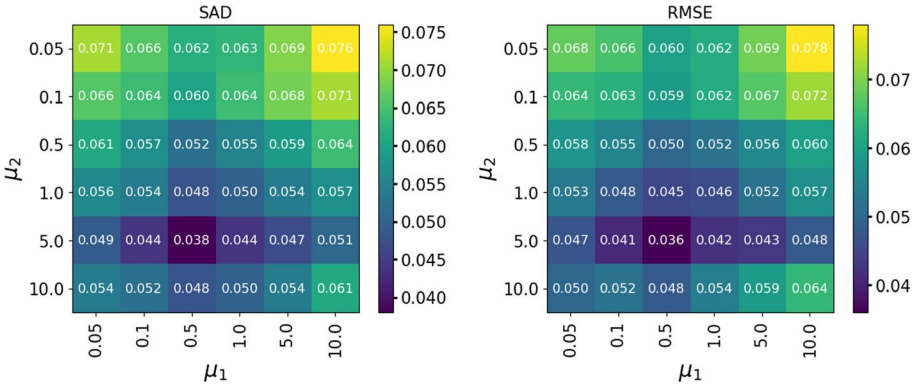

Fig. 4 Performance evaluation of the suggested approach regarding parameters α_4, α_5 and α_6 based on SAD and RMSE

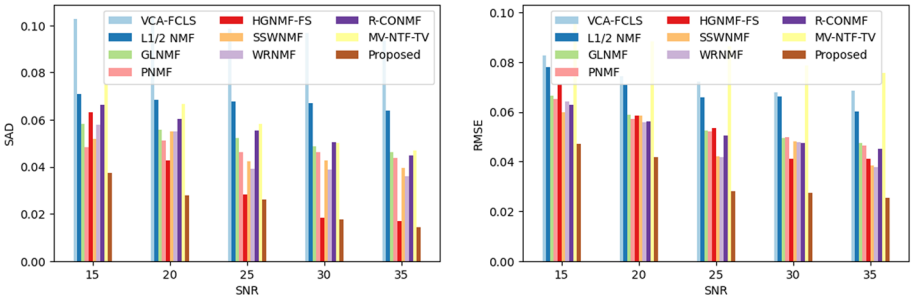

Fig. 5 Evaluating the algorithms at various noise levels based on SAD and RMSE

rise as α_2 increases from 0.0005 to 0.1. Unlike the previous scenario with α_1, there is no stable range observed here. The minimum values for both metrics are found at the smallest value of α_2 at 0.0005. As lower values of SAD and RMSE indicate superior performance, the optimal selection for α_2 based on these findings is 0.0005. A similar procedure was applied to the remaining parameters, with experimental results, depicted in Fig. 4, indicating that the optimal values for α_3, α_4, α_5, and α_6 were 0.01, 0.01, 0.01, and 0.1, respectively.

2. Noise Resilience Evaluation: The primary goal of this experiment is to evaluate the performance of SODE-WNMTF under different noise structures, particularly Gaussian, salt-and-pepper, and block noise. It is important to mention that all models in the experiment start with the same initial conditions and choose the same number of endmembers. Figures 5, 6, and 7 demonstrate the robustness of various models across different levels of Gaussian, salt-and-pepper, and block noise, respectively. Gaussian noise levels are represented by different SNR values, while salt-and-pepper noise is controlled by the parameter π, which denotes the probability of introducing "salt" (white pixels with a value of 255) into the image. In this study, the probabilities of adding salt and pepper (black pixels with a value of 0) are considered equal. For block noise, its intensity is determined by the size of the black window, which darkens the affected region of the HSI. As illustrated in Fig. 5, SODE-WNMTF demonstrates outstanding results in HU,

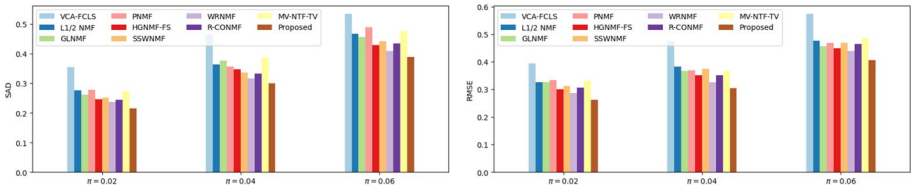

Fig. 6 Evaluating the algorithms at various salt and pepper noise levels based on SAD and RMSE

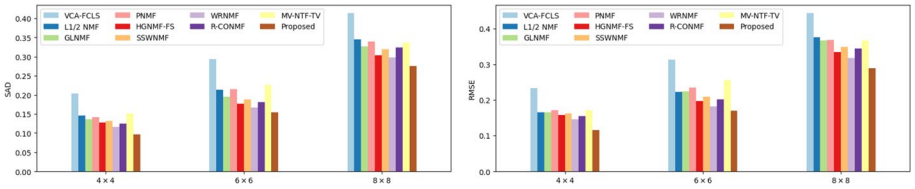

Fig. 7 Evaluating the algorithms at various block noise levels based on SAD and RMSE

consistently outperforming all other models across both SAD and RMSE metrics at all SNR levels. This finding applies to the other two types of noise as well, even in highly noisy conditions. This success can be attributed to its innovative features: adaptive outlier handling and robust endmember extraction. The former dynamically adjusts the influence of pixels and spectral bands, while the latter ensures accuracy even in high-noise scenarios. These attributes allow SODE-WNMTF to regularly surpass other algorithms, demonstrating its comprehensive strategy in tackling the complex challenges of HU.

3. Evaluating Outlier Detection Using the Weighting Matrix $\mathbf{T}$: To evaluate the outlier detection capability of SODE-WNMTF, we conducted an experiment by synthetically injecting outliers into a random band-pixel combination. In this experiment, the values of the 176th spectral band and the 11th pixel were deliberately perturbed by randomly sampling from a uniform distribution in the range $(0, 1)$. Our framework introduces a dynamic weighting matrix $\mathbf{T} \in \mathbb{R}_{+}^{L \times N}$, designed to attenuate the influence of such outliers during the unmixing process. This matrix is iteratively updated based on the distances of individual bands and pixels from their respective cluster centers, as defined in (6). Lower values in $\mathbf{T}$ signify a higher deviation from the expected data distribution and therefore indicate a higher likelihood of being an outlier. To visually assess the behavior of the $\mathbf{T}$ matrix, we plotted the weight distributions of selected bands and pixels after the unmixing process. The reshaped heatmaps shown in Fig. 8 correspond to the 176th and 99th rows of $\mathbf{T}$, representing the outlier-injected and clean bands, respectively. As observed, the 176th band exhibits significantly lower weights in the corresponding heatmap, aligning with its artificially induced anomaly. In contrast, the 99th band, free from contamination, displays uniformly higher values, reflecting its consistency with the data distribution. A similar trend was observed in the pixel domain. The weight map corresponding to the 11th pixel (reshaped to 7×32) demonstrated a pronounced drop in values, indicative of its outlier status, whereas a clean pixel such as the 37th maintained consistently higher weights. These visual and quantitative observations confirm that the proposed weighting mechanism effectively identifies and suppresses the influence of

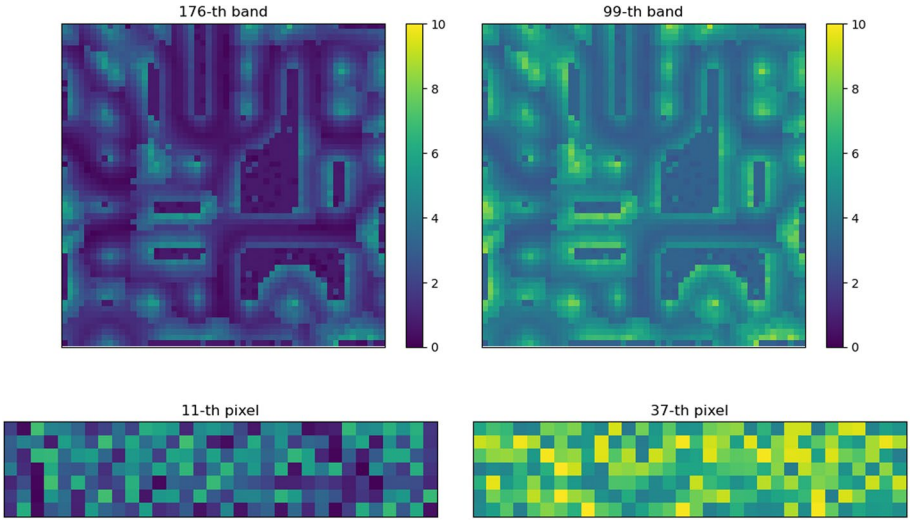

Fig. 8 Visualization of heatmaps generated from the weighting matrix $\mathbf{T}$, highlighting the differences between noisy and clean bands and pixels

anomalous bands and pixels. This contributes to enhanced robustness of the unmixing process, especially in real-world scenarios where data contamination is common.

4. Dimensions of Synthetic Data Image

   The aim of this experiment is to examine how SODE-WNMTF reacts to various image sizes. The sizes of the synthetic data utilized in this experiment range from 36×36 to 100×100. There are six endmembers in the synthetic data, with the SNR set at 20 dB. As shown in Fig. 9, SODE-WNMTF exhibits exceptional performance across various image sizes in HU, showcasing its versatility and stability. In fact, the proposed SODE-WNMTF framework adaptively handles data complexities, particularly as image sizes increase, while consistently achieving low SAD and RMSE values. This superior performance stems from some key attributes: effective spectral structure preservation, manifold-aware representation that adapts to varying spatial complexities, outlier-robust spatial regularization ensuring coherent abundance mapping, and sparse pixel selection for accurate endmember estimation.

5. Generalization to the Number of Endmembers

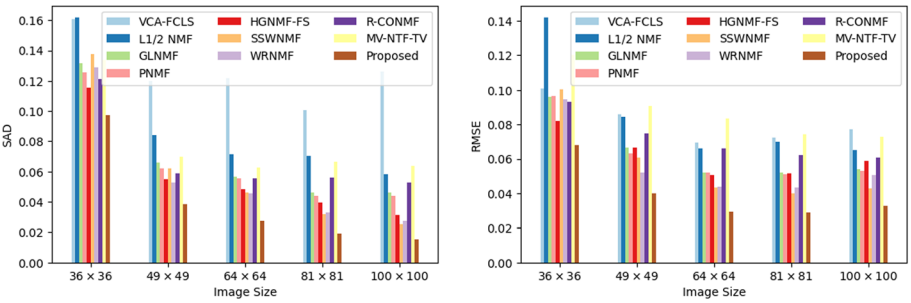

Fig. 9 Evaluating models at various image sizes based on SAD and RMSE metrics

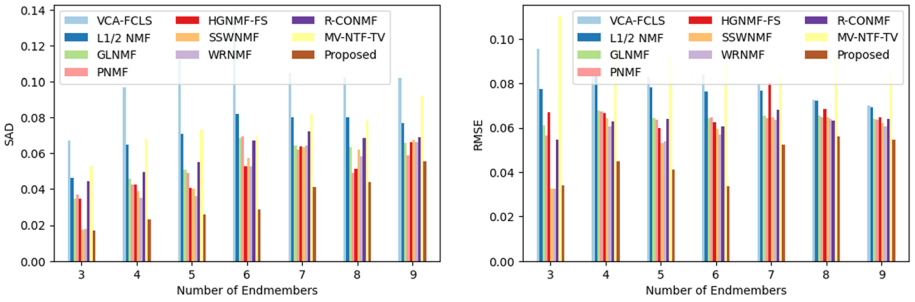

Fig. 10 Evaluating models at various endmember numbers based on SAD and RMSE

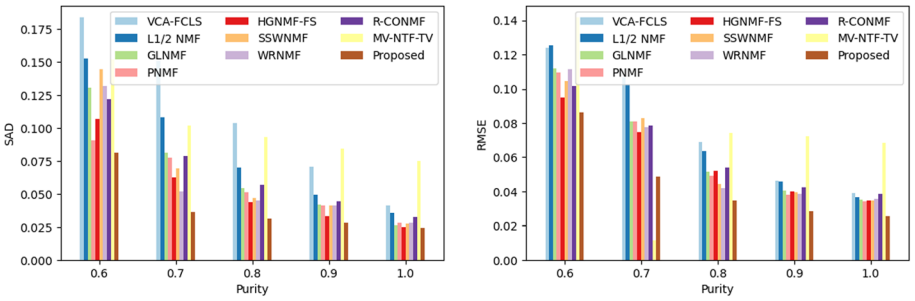

Fig. 11 Evaluating models at various mixing degrees based on SAD and RMSE

This study evaluates the performance of all methods with varying numbers of endmembers in synthetic data at an SNR setting of 20 dB, while the number of endmembers ranges from 3 to 9. From the analysis depicted in Fig. 10, it is evident that SODE-WNMTF consistently displays the lowest or one of the lowest error values across all endmember numbers in both metrics, demonstrating superior accuracy in estimating endmember signatures. Models like VCA-FCLS and $L_{\frac{1}{2}}$-NMF tend to exhibit higher error rates, whereas models such as GLNMF and PNMF show improved performance but still do not outperform SODE-WNMTF. This highlights the excellence of the proposed model in maintaining high precision and stability in HU tasks across varying numbers of endmembers. It is notable that the remarkable accuracy attained through the unique weighted NMTF framework is further enhanced by sparse endmember estimation. This approach effectively reduces the impact of mixed pixels when extracting endmember signatures from similar pixel spectra, thereby improving the endmember detection procedure.

6. Adaptation to Different Mixing Degree

When the pixels in the image are mixed to different degrees, the performance of the unmixing is affected. Therefore, it is crucial to consider the robustness of algorithm towards the endmember mixture. In this experiment, the parameter θ, which controls the level of mixing, gradually increases from 0.6 to 1, while the other variables are kept constant. As depicted in Fig. 11, SODE-WNMTF consistently exceeds the performance of other techniques in both SAD and RMSE, especially at lower purity levels (higher mixing degrees). The superior performance of SODE-WNMTF can be attributed to

several key factors. It employs a weighted approach combined with resilient outlier detection, effectively reducing outlier impact. Additionally, SODE-WNMTF adapts to spatial information while minimizing outlier influence and maintains spectral band similarities and local geometric structures. These features collectively enhance the unmixing process, resulting in greater accuracy and robustness compared to other methods.

7. Convergence Curves

   In this study, we provide a comparative examination of convergence analysis among various models using various sizes of synthetic datasets. The experimental setup required the selection of six endmembers and employed an SNR of 25 dB. The resulting residual error curves, illustrated in Fig. 12, demonstrate a marked declining trend in the main residual error, consistent with the anticipated research outcomes. As depicted in this figure, SODE-WNMTF exhibits a notably faster convergence rate compared to the other models across all dataset sizes, suggesting superior scalability for large datasets. Furthermore, the runtime for different models has been recorded in Table 2 for this dataset. With a runtime of 254.21 s, SODE-WNMTF outperforms several almost all methods. This indicates that the proposed approach balances computational cost and efficiency while maintaining superior factorization quality.

8. Ablation Analysis

   (a) The purpose of this experiment is to evaluate how the inclusion or exclusion of the weighting matrix **T** affects the efficiency of the proposed SODE-WNMTF model.

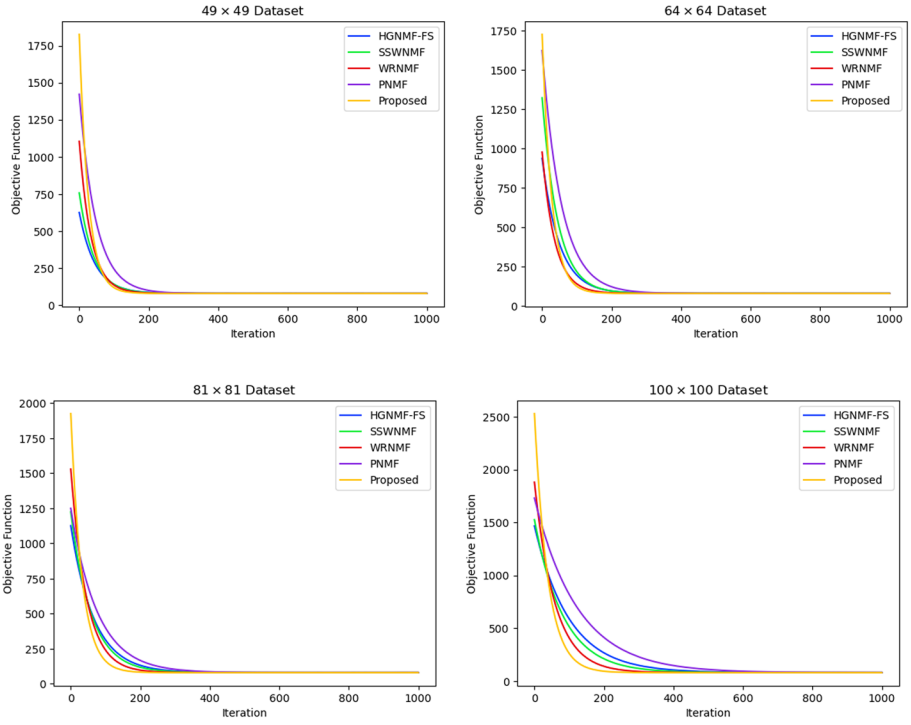

Fig. 12 Convergence curves of residual error

Table 2 Runtime comparison of various models on a 100×100 synthetic dataset

Algorithms	VCA-FCLS	$L_{\frac{1}{2}}$-NMF	GLNMF	PNMF	HGNMF-FS	SSWNMF	WRNMF	R-CONMF	MV-NTF-TV	Proposed
Runtime (s)	**46.4**	263.34	310.64	337.53	457.62	563.11	443.21	401.97	635.21	254.21

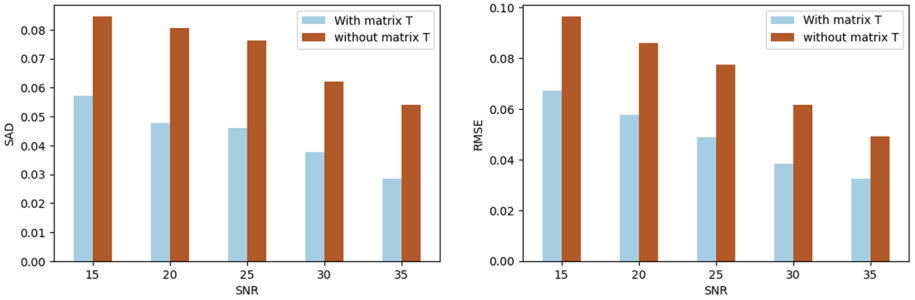

Fig. 13 Comparison of noise reduction capability by the proposed model using weight matrix T or without it, based on SAD and RMSE

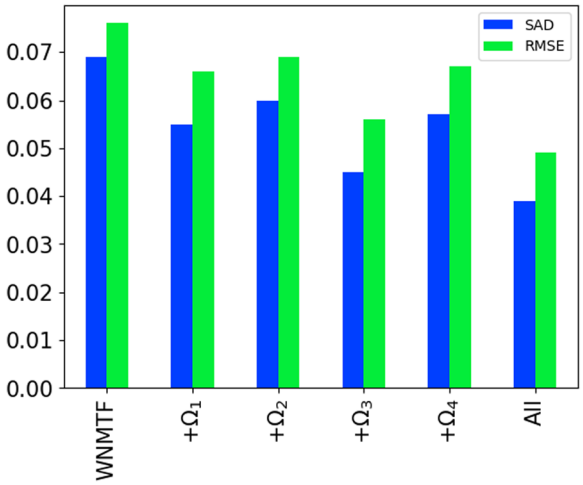

Fig. 14 Ablation study results showing the impact of individual regularization terms on SAD and RMSE

Additionally, the experiment examines how **T** mitigates the impact of noise and outliers in the HU process, with θ fixed at 0.7. It can be deduced from Fig. 13 that the model incorporating the matrix **T** consistently outperforms the model without it across all SNR levels. In particular, the difference in performance is large and remains significant as SNR values increase. This demonstrates that incorporating the matrix **T** in SODE-WNMTF effectively reduces the impact of noise and outliers on the HU process.

(b) In this experiment we provide a detailed assessment of the impact of different regularization terms on model performance. These results reveal the importance of each term in improving the ability of the model to extract meaningful spectral information and reconstruct data with higher accuracy. According to Fig. 14, results indicate that the core WNMTF model has the highest error values for both metrics. Introducing individual regularization terms reduces the errors, with varying degrees of improvement. $\mathbf{\Omega}_3$ shows the most significant reduction in SAD, while $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ also contribute positively. $\mathbf{\Omega}_4$ improves RMSE but less significantly affects SAD. The combination of all regularization terms achieves

the lowest error values, demonstrating their collective effectiveness in enhancing the HU performance.

5.2 Experiments on real hyperspectral data

To validate the efficacy of the proposed method, this section presents an experimental evaluation using three real-world hyperspectral datasets: Samson,[2] Jasper Ridge,[3] and Urban Hydice.[4]

1. Samson Dataset: The Samson dataset is frequently used for HU and includes 95×95 pixels, each comprising 156 spectral bands ranging from 401 to 889 nm. According to Salehani and Gazor (2017), this dataset encompasses three main resources: trees, water, and soil.

   According to Fig. 15, the proposed model shows outstanding results in predicting the signatures of endmembers. The estimated spectra closely resemble the library spectra for all three endmembers, with only small differences noticed. This high level of similarity suggests that SODE-WNMTF effectively captures the distinct spectral features of each endmember. Also, evaluating the Fig. 16 reveals a significant similarity between the estimates of the proposed model and the reference maps for all three endmembers. Furthermore, as shown in Table 3, SODE-WNMTF exhibits excellent performance overall, obtaining the lowest mean SAD value for all endmembers, demonstrating superior average accuracy. It shows strong performance for tree and water endmembers, achieving the lowest SAD values of 0.0464 and 0.0468, respectively. For the soil endmember, although not the top performer, SODE-WNMTF still demonstrates good competitive performance.
2. HYDICE Urban Dataset: In research conducted at HU, urban hyperspectral data is heavily employed. The dataset is composed of 307 by 307 pixels, with each pixel containing 210 spectral bands spanning the wavelength range from 0.4–2.4 µm (Liu et al., 2011). After removing noise channels, 162 channels remain. According to prior studies, this dataset involves six types of materials: asphalt road, roof #1, roof #2, grass, tree, and concrete road. The results are outlined in Table 4 and Figs. 17 and 18. As can be deduced from these results, examination of endmember estimation shows strong agreement between the signatures obtained from SODE-WNMTF and the reference library spectra throughout the wavelength range. The abundance maps of SODE-WNMTF closely match the reference distributions spatially, indicating an accurate representation of material prevalence. The SAD metric provides additional evidence supporting the superior performance of the model through quantitative evaluation. According to these results, SODE-WNMTF outperforms all the other models in achieving lower SAD values for important material categories like asphalt-road, grass, tree, and concrete-road. Thus, the capability of SODE-WNMTF to handle large and intricate HSIs is convincingly proven by its consistently superior results across multiple evaluation metrics.

[2] http://www.escience.cn/people/feiyunZHU/Dataset_GT.html.
[3] http://www.escience.cn/system/file?fileId=68574.
[4] http://www.agc.army.mil/.

Table 3 SAD results on the Samson dataset

Algorithms	VCA-FCLS	$L_{\frac{1}{2}}$-NMF	GLNMF	PNMF	HGNMF-FS	SSWNMF	WRNMF	R-CONMF	MV-NTF-TV	GLA	$L_{2,1}$-RDNMF	Proposed
Soil	0.0609	0.0320	0.0325	0.0271	0.0398	0.0133	0.0204	0.0684	0.0378	0.0311	**0.0129**	0.0316
Tree	0.0529	0.0698	0.0707	0.0524	0.0589	0.0710	0.0601	0.0653	0.0546	0.0502	0.0583	**0.0464**
Water	0.1257	0.0650	0.0662	0.1173	0.0960	0.0714	0.0539	0.1135	0.0875	0.0507	0.0583	**0.0468**
Mean	0.0798	0.0556	0.0565	0.0656	0.0649	0.0519	0.0448	0.0824	0.0600	0.044	0.0432	**0.0416**

The highest scores in each row are highlighted in bold

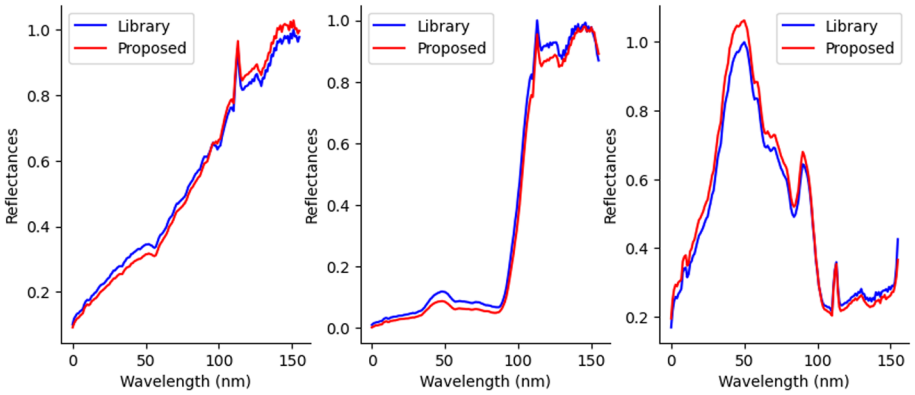

Fig. 15 Comparison between the reference spectra and the endmember signatures obtained using the proposed method on the Samson dataset. Followed by soil, tree and water

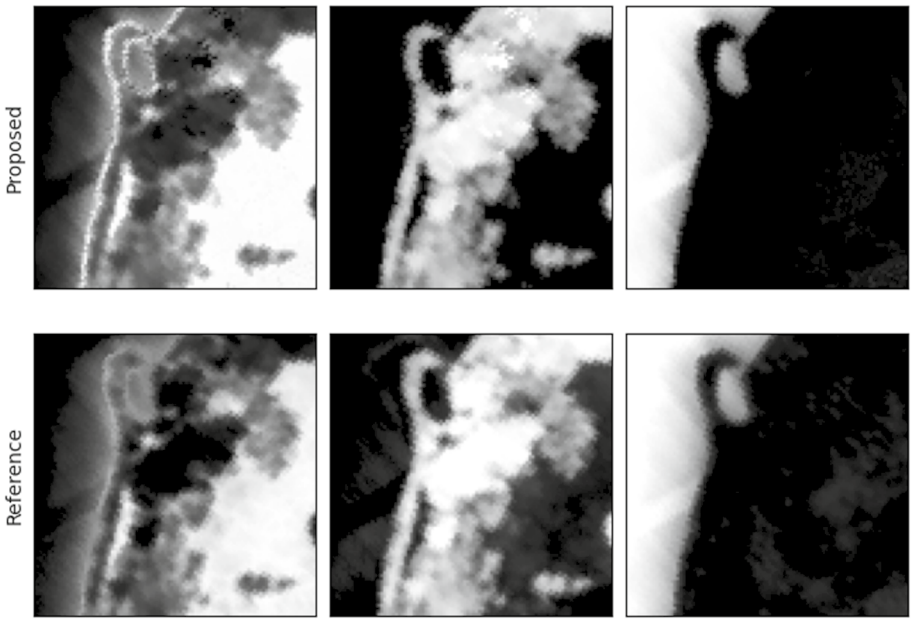

Fig. 16 Abundance maps of the various endmembers created using the method proposed on the Samson dataset. Soil, tree and water are arranged from left to right

3. Jasper Ridge Dataset: This dataset has seen significant application in HU research. Due to the intricate nature of Jasper Ridge data images, only a subset of 100×100 pixels is analyzed, with each pixel consisting of 198 bands following the removal of noise bands. According to Qian et al. (2017), four distinct endmembers-tree, water, soil, and road-are identified in this specific subscene. Figure 19 depicts the extracted endmember signatures using SODE-WNMTF and the reference endmember signatures. An accurate capture of spectral characteristics is suggested in Fig. 19 by a close match between the

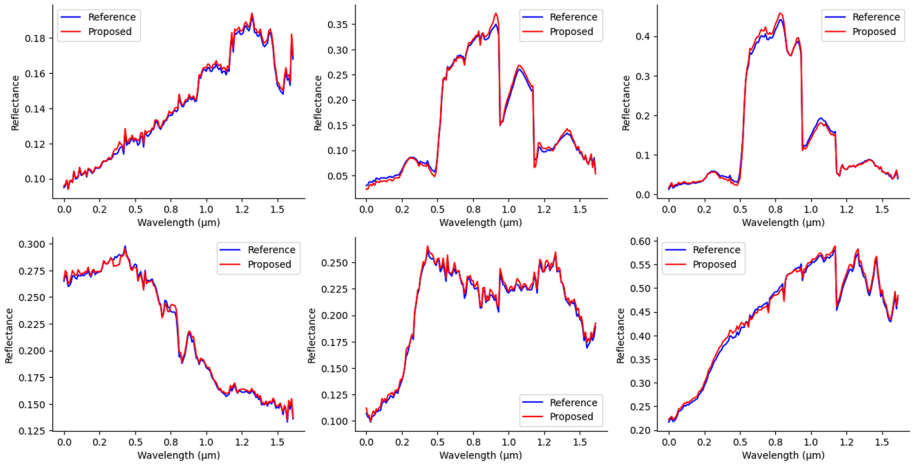

Fig. 17 Comparison between the reference spectra and the endmember signatures obtained using the proposed method on the Hydice dataset. Followed by roof #1, concrete-road, roof #2, tree, grass and asphalt-road

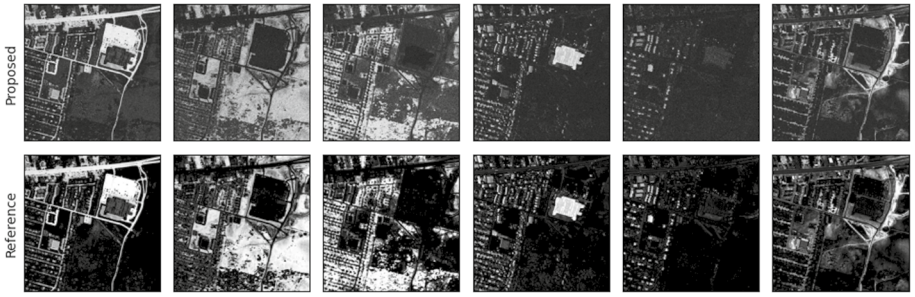

Fig. 18 Abundance maps of the various endmembers created using the method proposed on the Hydice dataset. roof #1, concrete-road, roof #2, tree, grass and asphalt-road are arranged from left to right

library and estimated spectra for all endmembers. Also, by delving deeper into Fig. 20, it can be deduced that the abundance maps generated by SODE-WNMTF closely resemble the reference maps, indicating precise estimation of spatial distribution and relative abundance of endmembers. Moreover, the SAD results shown in Table 5 reveal an exceptional performance of SODE-WNMTF, achieving the lowest overall mean SAD and best scores for soil, tree, and water endmembers. This consistent excellence across diverse endmember types highlights the versatility and robustness of the model.

Table 4 SAD results on the Urban Hydice dataset

Algorithms	VCA-FCLS	$L_{\frac{1}{2}}$-NMF	GLNMF	PNMF	HGNMF-FS	SSWNMF	WRNMF	R-CONMF	MV-NTF-TV	GLA	$L_{2,1}$-RDNMF	Proposed
Asphalt-road	0.0765	0.0425	0.0410	0.0564	0.0402	**0.0312**	0.0396	0.0415	0.0458	0.0310	0.0342	0.0318
Grass	0.1125	0.0830	0.0738	0.0843	0.0810	**0.0615**	0.0761	0.0769	0.0704	0.0687	0.0676	0.0693
Tree	0.1185	0.0758	0.1047	0.0713	0.0847	0.0667	0.0693	0.0844	0.0741	0.0698	0.0711	**0.0660**
Roof #1	0.1404	0.1220	0.1244	0.0956	0.0923	0.0980	0.0942	0.0926	0.1034	0.0924	0.0943	**0.0905**
Roof #2	0.1514	0.1145	0.1120	0.1019	0.0549	0.0970	0.0748	0.1034	0.0823	0.0502	0.0497	**0.0481**
Concrete-road	0.1882	0.0874	0.0854	0.0954	0.0801	0.0828	0.0684	0.0914	0.0712	0.0584	0.0592	**0.0571**
Mean	0.1312	0.0875	0.0902	0.0842	0.0722	0.0729	0.0704	0.0817	0.0745	0.0618	0.0627	**0.0605**

The highest scores in each row are highlighted in bold

Table 5 SAD results on the Jasper Ridge dataset

Algorithms	VCA-FCLS	$L_{\frac{1}{2}}$-NMF	GLNMF	PNMF	HGNMF-FS	SSWNMF	WRNMF	R-CONMF	MV-NTF-TV	GLA	$L_{2,1}$-RDNMF	Proposed
Soil	0.0644	0.0518	0.0570	0.0537	0.0543	0.0534	0.0607	0.0513	0.0578	**0.049**	0.0507	0.0497
Tree	0.4894	0.2925	0.2989	0.3163	0.2026	0.2724	0.1456	0.2654	0.3252	0.1303	0.1398	**0.1242**
Water	0.0649	0.0512	0.0494	0.0451	0.0411	0.0418	0.0454	0.0436	0.0489	0.0421	0.0409	**0.0398**
Road	0.0833	0.0612	0.0626	0.0676	0.0613	0.0622	**0.0576**	0.0686	0.0646	0.065	0.0672	0.0654
Mean	0.1505	0.1142	0.1170	0.1207	0.0898	0.1074	0.0773	0.1072	0.1241	0.0716	0.0747	**0.0698**

The highest scores in each row are highlighted in bold

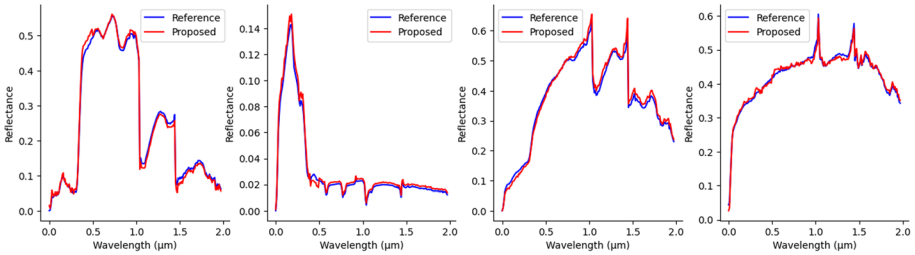

Fig. 19 Comparison between the reference spectra and the endmember signatures obtained using the proposed method on the Jasper Ridge dataset. Followed by soil, tree, water and road

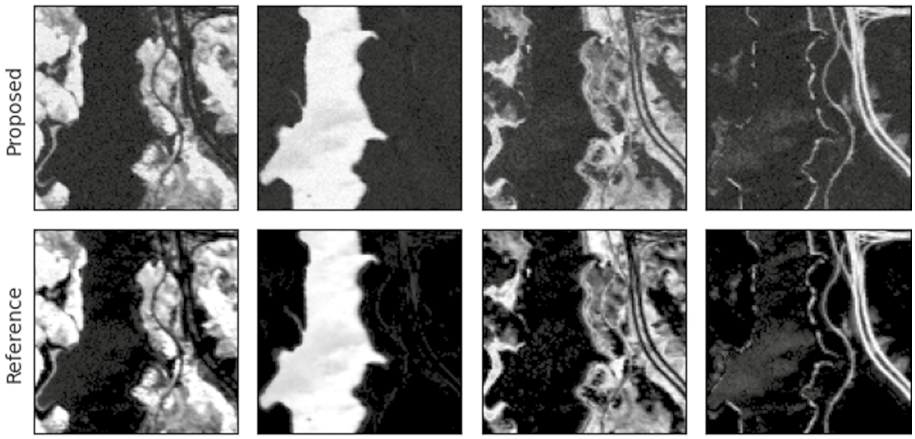

Fig. 20 Abundance maps of the various endmembers created using the method proposed on the Jasper Ridge dataset. soil, tree, water and road are arranged from left to right

6 Conclusion

This paper introduces a novel HU methodology, called SODE-WNMTF, which is designed to address the concurrent detection and denoising of pixel and spectral band outliers during the unmixing process. SODE-WNMTF successfully tackles some key challenges in HSI analysis by incorporating several innovative components. Firstly, it leverages the co-clustering property of NMTF to simultaneously detect outliers in both pixel and spectral band domains, enhancing the robustness of the unmixing process. Secondly, SODE-WNMTF employs a sparse endmember estimation technique that utilizes pixel spectrum information, enhancing the precision of endmember extraction, especially when dealing with pixels that are highly mixed. Thirdly, regularization terms based on the inherent structure of hyperspectral data are proposed to preserve the global structure of spectral bands and the intrinsic manifold of the data, thus improving the effectiveness of the clustering process. Finally, a spatial-spectral sparsity constraint is also applied to the abundance matrix, accounting for spatial relationships between adjacent pixels while adapting to potential anomalies. These components are integrated into a unified objective function, which is then optimized through the MUR method. Extensive experimental evaluations confirm the superiority of SODE-WNMTF, highlighting its improved performance in endmember extraction and

abundance estimation, especially in challenging conditions characterized by noise and outliers. Upcoming research will integrate a dual auto-encoder approach into the NMTF structure. This integration aims to better reveal the inherent patterns within the data, consequently refining the derived embedded representations utilized in clustering processes.

Author contributions Yasin Hashemi-Nazari: Writing – original draft, Methodology, Conceptualization. Azita Tajaddini: Writing – original draft, Methodology, Formal analysis, Conceptualization. Farid Saberi-Movahed: Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization. Catarina Moreira: Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. Xin Ning: Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. Prayag Tiwari: Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Funding Open access funding provided by Halmstad University.

Data availability The source code is also available at https://github.com/yasinhashemi/SODE-WNMTF.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Ang, A. M. S., & Gillis, N. (2019). Algorithms and comparisons of nonnegative matrix factorizations with volume regularization for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*, 4843–4853.

Bioucas-Dias, J. M., & Nascimento, J. M. P. (2008). Hyperspectral subspace identification. *IEEE Transactions on Geoscience and Remote Sensing, 46*, 2435–2445.

Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P. D., & Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*, 354–379.

Cui, C., Wang, X., Wang, S., Zhang, L., & Zhong, Y. (2023). Unrolling nonnegative matrix factorization with group sparsity for blind hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–12.

Dache, A., Vandaele, A., & Gillis, N. (2024). Orthogonal symmetric nonnegative matrix tri-factorization. In: *IEEE International Workshop on Machine Learning for Signal Processing*.

Ding, C., Li, T., Peng, W., & Park, H. (2006). Orthogonal nonnegative matrix T-factorizations for clustering. In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 126–135).

Dobigeon, N., Tourneret, J.-Y., Richard, C., Bermudez, J. C. M., McLaughlin, S., & Hero, A. O. (2014). Nonlinear unmixing of hyperspectral images: Models and algorithms. *IEEE Signal Processing Magazine, 31*, 82–94.

Drumetz, L., Chanussot, J., Jutten, C., Ma, W.-K., & Iwasaki, A. (2020). Spectral variability aware blind hyperspectral image unmixing based on convex geometry. *IEEE Transactions on Image Processing, 29*, 4568–4582.

Du, L., Li, X., & Shen, Y.-D. (2012). Robust nonnegative matrix factorization via half-quadratic minimization. In *2012 IEEE 12th International Conference on Data Mining* (pp. 201–210).

Fang, H., Li, A., Wang, T., & Xu, H. (2018a). Hyperspectral unmixing using double-constrained multilayer NMF. *Remote Sensing Letters, 10*, 224–233.

Fang, H., Li, A., Xu, H., & Wang, T. (2018b). Sparsity-constrained deep nonnegative matrix factorization for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters, 15*, 1105–1109.

Feng, X.-R., Li, H.-C., Wang, R., Du, Q., Jia, X., & Plaza, A. (2022). Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 4414–4436.

Guo, Z., Min, A., Yang, B., Chen, J., & Li, H. (2021). A modified Huber nonnegative matrix factorization algorithm for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 5559–5571.

Haghir Chehreghani, M. (2020). Unsupervised representation learning with minimax distance measures. *Machine Learning, 109*(11), 2063–2097.

Heinz, D. C., & Chang, C.-I. (2001). Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing, 39*, 529–545.

He, W., Zhang, H., & Zhang, L. (2016). Sparsity-regularized robust non-negative matrix factorization for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 4267–4279.

He, W., Zhang, H., & Zhang, L. (2017). Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 55*, 3909–3921.

Huang, R., Jiao, H., Li, X., Chen, S., & Xia, C. (2023). Hyperspectral unmixing using robust deep nonnegative matrix factorization. *Remote Sensing, 15*(11), 2900.

Ince, T., & Dobigeon, N. (2022). Weighted residual NMF with spatial regularization for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters, 19*, 1–5.

Ince, T., & Dobigeon, N. (2023). Spatial-spectral multiscale sparse unmixing for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters, 20*, 1–5.

Iwata, T., Toyoda, M., Tora, S., & Ueda, N. (2020). Anomaly detection with inexact labels. *Machine Learning, 109*(8), 1617–1633.

Kiani, R., Jin, W., & Sheng, V. S. (2024). Survey on extreme learning machines for outlier detection. *Machine Learning, 113*(8), 5495–5531.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*, 788–791.

Li, J., Bioucas-Dias, J. M., Plaza, A., & Liu, L. (2016). Robust collaborative nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 6076–6090.

Li, C., & Chen, X. (2020). Error approximation of hyperspectral unmixing via correntropy-induced metric. *IEEE Geoscience and Remote Sensing Letters, 17*(2), 282–286.

Li, C., & Chen, X. (2023). A staged approach with structural sparsity for hyperspectral unmixing. *IEEE Sensors Journal, 23*(12), 13248–13260.

Li, H., Feng, R., Wang, L., Zhong, Y., & Zhang, L. (2021a). Superpixel-based reweighted low-rank and total variation sparse unmixing for hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing, 59*, 629–647.

Li, C., Jiang, Y., & Chen, X. (2021b). Hyperspectral unmixing via noise-free model. *IEEE Transactions on Geoscience and Remote Sensing, 59*(4), 3277–3291.

Liu, H., Lu, Y., Wu, Z., Du, Q., Chanussot, J., & Wei, Z. (2023a). Spectral variability bayesian unmixing for hyperspectral sequence in wavelet domain. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–19.

Liu, W., Mei, X., Ma, Y., Huang, J., Chen, Q., & Li, H. (2023b). Graph l_1-Laplacians regularized GMM for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters, 20*, 1–5.

Liu, X., Xia, W., Wang, B., & Zhang, L. (2011). An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 49*(2), 757–772.

Lohrer, A., Kazempour, D., Hünemörder, M., & Kröger, P. (2024). Comadout-a robust outlier detection algorithm based on CoMAD. *Machine Learning, 113*(10), 8061–8135.

Lu, X., Wu, H., Yuan, Y., Yan, P., & Li, X. (2013). Manifold regularized sparse NMF for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 51*, 2815–2826.

Ma, W.-K., Bioucas-Dias, J. M., Chan, T.-H., Gillis, N., Gader, P., Plaza, A. J., Ambikapathi, A., & Chi, C.-Y. (2014). A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *IEEE Signal Processing Magazine, 31*(1), 67–81.

Mantripragada, K., & Qureshi, F. Z. (2024). Hyperspectral pixel unmixing with latent Dirichlet variational autoencoder. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1–12.

Miao, L., & Qi, H. (2007). Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing, 45*, 765–777.

Nascimento, J. M. P., & Dias, J. M. B. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 43*, 898–910.

Qian, Y., Jia, S., Zhou, J., & Robles-Kelly, A. (2011). Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing, 49*, 4282–4297.

Qian, Y., Xiong, F., Zeng, S., Zhou, J., & Tang, Y. Y. (2017). Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing, 55*(3), 1776–1792.

Qu, K., & Bao, W. (2020). Multiple-priors ensemble constrained nonnegative matrix factorization for spectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13*, 963–975.

Qu, K., & Li, Z. (2024). A fast sparse NMF optimization algorithm for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 1885–1902.

Qu, K., Li, Z., Luo, X., Bao, W., & Luo, F. (2024). Hyperspectral unmixing using reweighted unidirectional TV low-rank NTF with multiple-factor collaboration regularization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 9628–9644.

Qu, K., Li, Z., Wang, C., Luo, F., & Bao, W. (2023). Hyperspectral unmixing using higher-order graph regularized NMF with adaptive feature selection. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–15.

Rasti, B., Koirala, B., Scheunders, P., & Chanussot, J. (2022). MiSiCNet: minimum simplex convolutional network for deep hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–15.

Salehani, Y. E., & Gazor, S. (2017). Smooth and sparse regularization for NMF hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*, 3677–3692.

Song, F.-X., Kan, C., & Deng, S.-W. (2024). Mutual incoherence and relative total variation regularizations for blind hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1–15.

Wang, N., Du, B., & Zhang, L. (2013). An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*, 554–569.

Wang, T., Li, J., Ng, M. K., & Wang, C. (2023). Nonnegative matrix functional factorization for hyperspectral unmixing with non-uniform spectral sampling. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1–13.

Wang, W. H., Qian, Y., & Tang, Y. Y. (2016). Hypergraph-regularized sparse NMF for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 681–694.

Xiong, F., Qian, Y., Zhou, J., & Tang, Y. Y. (2019). Hyperspectral unmixing via total variation regularized nonnegative tensor factorization. *IEEE Transactions on Geoscience and Remote Sensing, 57*(4), 2341–2357.

Xu, X., Song, X., Li, T., Shi, Z., & Pan, B. (2022a). Deep autoencoder for hyperspectral unmixing via global-local smoothing. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–16.

Xu, M., Yang, Z., Ren, G., Sheng, H., Liu, S., Liu, W., & Ye, C. (2022b). l_1 sparsity-constrained archetypal analysis algorithm for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters, 19*, 1–5.

Yang, T., Song, M., Li, S., & Wang, Y. (2023). Spectral-spatial anti-interference NMF for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–17.

Yuan, Y., Feng, Y., & Lu, X. (2015). Projection-based NMF for hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*, 2632–2643.

Zhang, G., Mei, S., Feng, Y., & Du, Q. (2021). Spectral-spatial constrained nonnegative matrix factorization for spectral mixture analysis of hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 6766–6776.

Zhang, S., Zhang, G., Li, F., Deng, C., Wang, S., Plaza, A., & Li, J. (2022). Spectral-spatial hyperspectral unmixing using nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–13.

Zou, X., Xu, M., Liu, S., & Sheng, H. (2024). Superpixel-based graph Laplacian regularized and weighted robust sparse unmixing. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 1–15.

Authors and Affiliations

Yasin Hashemi-Nazari[1] · Farid Saberi-Movahed[2] · Azita Tajaddini[1] · Catarina Moreira[3] · Xin Ning[4] · Prayag Tiwari[5]

✉ Farid Saberi-Movahed
f.saberimovahed@kgut.ac.ir

✉ Prayag Tiwari
prayag.tiwari@ieee.org

Yasin Hashemi-Nazari
yhashemi@math.uk.ac.ir

Azita Tajaddini
atajadini@uk.ac.ir

Catarina Moreira
Catarina.PintoMoreira@uts.edu.au

Xin Ning
ningxin@semi.ac.cn

[1] Department of Applied Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

[2] Department of Applied Mathematics, Faculty of Sciences and Modern Technologies, Graduate University of Advanced Technology, Kerman, Iran

[3] Data Science Institute, University of Technology Sydney, Sydney, Australia

[4] AnnLab, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

[5] School of Information Technology, Halmstad University, Halmstad, Sweden