Article

Low-Rank and Sparse Matrix Recovery for Hyperspectral Image Reconstruction Using Bayesian Learning

Yanbin Zhang [1,2,*], Long-Ting Huang [3], Yangqing Li [1], Kai Zhang [1] and Changchuan Yin [1]

[1] Beijing Laboratory of Advanced Information Networks, Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China; liyq@bupt.edu.cn (Y.L.); kaizhang@bupt.edu.cn (K.Z.); ccyin@bupt.edu.cn (C.Y.)

[2] China Fire and Rescue Institute, Beijing 102202, China

[3] School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China; huanglt08@whut.edu.cn

[*] Correspondence: iyanbin4@bupt.edu.cn

Abstract: In order to reduce the amount of hyperspectral imaging (HSI) data transmission required through hyperspectral remote sensing (HRS), we propose a structured low-rank and joint-sparse (L&S) data compression and reconstruction method. The proposed method exploits spatial and spectral correlations in HSI data using sparse Bayesian learning and compressive sensing (CS). By utilizing a simultaneously L&S data model, we employ the information of the principal components and Bayesian learning to reconstruct the hyperspectral images. The simulation results demonstrate that the proposed method is superior to LRMR and SS&LR methods in terms of reconstruction accuracy and computational burden under the same signal-to-noise tatio (SNR) and compression ratio.

1. Introduction

HSI is a collection of hundreds of images that are usually acquired simultaneously in narrow and adjacent spectral bands by airborne sensors [1] or spaceborne spectrometers [2]. HSI combines the traditional two-dimensional remote sensing imaging technology and the optical imaging technology of spectroscopy [3]. Moreover, HSI has achieved the effect of acquiring images and spectra of objects at the same time, which has set off a revolution in the field of remote sensing. In recent years, with the rapid development of precision agriculture, hyperspectral imaging (HSI) technology in hyperspectral remote sensing [4] has been widely used. In precision agriculture, hyperspectral images can be used to monitor drought and flooding in farmlands, pests and diseases, and crop growth, as well as to predict farmland yield. According to the estimated variability from HSI, precision agriculture can improve resource-use efficiency, productivity, quality, profitability, and sustainability of agricultural production, such as by using less irrigation water, fewer pesticides and fertilizers, etc. HSI is also used in many applications besides precision agriculture, including forestry monitoring, natural resource investigation, vegetation observation, food safety, and geological mapping.

However, the acquisition and processing of HSI require a high-sensitivity detector, ultra-long-distance transmission, great computing power, and a huge data storage capacity in order to deal with such huge amounts of data. The HSI data need to be transmitted to ground stations. In this regard, one of the obstacles that researchers have had to face in the past 20 years is that of finding a way to program hyperspectral satellites to reduce the HSI data. Fortunately, HSI data are low-rank [5] and sparse in the transform domain (discrete cosine transform (DCT) or wavelet transform) [6–8]. These kinds of data have the features of being low-rank and sparse, which means that the data are redundant, and a

small number of datasets can be used to express the entirety of the data. So, we can use these features to reduce the amount of HSI data to be transmitted. This might be a good way to solve the problem.

Recently, many works have focused on modeling, compressing, and reconstructing HSI data using a structured framework. By dividing HSI into many blocks, Zhang et al. [5] introduced an HSI data reconstruction method based on low-rank matrix recovery (LRMR). In particular, for this kind of signal, many new compressive sensing (CS)-based methods [9–11] have been proposed [6,8,12–19]. Golbabaee et al. [6] simultaneously reconstructed HSI data with a low-rank and joint-sparse (L&S) structure by assuming that HSI data are low-rank and using a spatially joint-sparse wavelet representation. Zhang et al. [12] showed that the block sparse Bayesian learning (bSBL) algorithm has good recovery performance for data with a spatial block structure (such as an L&S structure). However, most of the existing research focused on low-rank structure reconstruction methods or HSI data denoising methods, and there are no methods for integrating the process of HSI acquisition and combining a sparse structure reconstruction method to reconstruct HSI data.

In this paper, a method based on the bSBL framework is proposed to compress, transmit, and reconstruct the entirety of HSI data. Here, each sub-matrix of multi-channel data is collected, compressed, and transmitted to the ground processing center through push-broom imaging. The proposed method not only combines the L&S structure priors on HSI data and the appropriate priors on hyperparameters, but also uses the information of the main components of the data. Firstly, we assume that the covariance matrix of the data is a diagonal matrix, and then obtain the initial value of the hyperparameter. Finally, an EM-like method is used to obtain the best reconstruction of the HSI data. The simulation results show that the proposed method has better performance than other existing methods.

We highlight the following main contributions:

(1) The proposed method gives the structure of the covariance matrix of the L&S signals, models HSI data with the L&S structure, and utilizes the CS and Bayesian learning methods to compress and reconstruct HSI data.

(2) In the reconstruction part, the proposed method makes use of the relationship between multiple dimensions of high-dimensional data and combines data reconstruction with HSI data acquisition. It can be used to realize the segmented acquisition, compression, and transmission of HSI data so as to reduce the amount of calculation and data transmission that must be performed in the satellite.

(3) We demonstrate the superior performance of the proposed method in comparison with state-of-the-art alternatives by conducting experiments on both synthetic signals and real signals.

The rest of this paper is arranged as follows: Section 2 introduces two 2D reconstruction methods for HSI data. Section 3 introduces the L&S data compression and reconstruction method combined with the acquisition mode. Section 3.1 introduces a model of the problem, and Section 3.2 proposes the HSI reconstruction method based on push-broom imaging. Section 3.3 introduces the simulation results. Discussion is given in Section 4. Finally, this work is concluded in Section 5.

Notation: $p(\mathbf{A}) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ denotes the probability density function of $\mathbf{A}$, which follows a Gaussian distribution with mean $\mathbf{0}$ and variance $\mathbf{\Sigma}$. $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$. $\|\mathbf{x}\|_2$ denotes the ℓ_2 norm of $\mathbf{x}$. $\text{vec}[\mathbf{A}]$ denotes the vectorization of the matrix $\mathbf{A}$ formed by stacking its columns into a single column vector. $\mathbf{A}^\top$ denotes the transpose of $\mathbf{A}$. $\text{tr}(\mathbf{A})$ denotes the trace of $\mathbf{A}$.

2. Two-Dimensional Reconstruction Methods for HSI Data

As described in the previous section, this section will combine the fast method proposed in our previous work [20] to compress and reconstruct HSI data from two different two-dimensional (2D) slice methods and will provide two 2D reconstruction algorithms for three-dimensional (3D) HSI data. These two methods are the method of expanding 3D data

HSI from different bands to 2D data by using a tensor and the method of slicing them into 2D data one by one according to the acquisition mode.

The HSI data experiment was carried out with the repetition of 100 real tests. Each experimental datum was selected from the Salinas Database (http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas-A_scene (accessed on 10 November 2021)) dataset in a 30×50 area, and 20 bands were used, that is, the data matrix was a $30 \times 50 \times 20$ 3D matrix with $N = 30$, $Q = 50$, $M = 20$. The two 2D reconstruction methods are first introduced in the following sections. The core part of this section shows the direct application of the fast L&S-bSBL method [20]. The same abbreviations as in [20] are used here.

(1)   Estimation of **x**.

After the Bayesian posterior probability is obtained with the Bayesian rule, the maximum a posteriori (MAP) is used to obtain the estimation of **x**:

$$\hat{\mathbf{x}} = \text{vec}(\hat{\mathbf{X}}^\top) \triangleq \boldsymbol{\mu}_\mathbf{x} = (\lambda \boldsymbol{\Sigma}_0^{-1} + \mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H} \boldsymbol{\Sigma}_0$$
$$= \boldsymbol{\Sigma}_0 \mathbf{H}^\top (\lambda \mathbf{I} + \mathbf{H} \boldsymbol{\Sigma}_0 \mathbf{H}^\top)^{-1} \mathbf{y}, \tag{1}$$

where $\boldsymbol{\Sigma}_0$ is a block matrix of **x**, and there are few elements that are non-zero. Moreover, the sparsity of blocks of $\hat{\mathbf{x}}$ is determined by $\gamma_i \gamma_j$. When $\gamma_k = 0$, the value of the kth related block in $\hat{\mathbf{x}}$ is zero.

(2)   Estimation of λ.

In order to obtain λ, we simplify the expression of $\boldsymbol{\Sigma}_0$ as

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Gamma} \otimes \mathbf{B}. \tag{2}$$

After maximizing the logarithm of the joint probability of **x** and **y**, we take the derivative of it with respect to λ:

$$\lambda \leftarrow \frac{\|\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_\mathbf{x}\|_2^2 + \lambda^{(pre)} \left[mn - \text{tr}(\boldsymbol{\Sigma}_\mathbf{x} \boldsymbol{\Sigma}_0^{-1}) \right]}{pn}. \tag{3}$$

(3)   Estimation of **B**.

$$\mathbf{B} = \arg\min_\mathbf{X} \text{tr} \left[\mathbf{B}_0^{-1} \left(\mathbf{X}\mathbf{X}^\top + \nabla_{\mathbf{B}_0^{-1}} \right) \right] + m \log |\mathbf{B}_0|$$
$$= \frac{1}{m} \left(\hat{\mathbf{X}} \hat{\mathbf{X}}^\top + \nabla_{\mathbf{B}_0^{-1}} \right). \tag{4}$$

where

$$\mathbf{B}_0 \leftarrow \frac{1}{m} \sum_{i=1}^m \frac{(\boldsymbol{\Sigma}_\mathbf{x}^i + \boldsymbol{\mu}_\mathbf{x}^i (\boldsymbol{\mu}_\mathbf{x}^i)^\top)}{\gamma_i^2}, \tag{5}$$

$$\nabla_{\mathbf{B}^{-1}} = \sum_{i=1}^m \mathbf{B} - \mathbf{B}\mathbf{H}_i^\top \left(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^\top + \lambda\mathbf{I} \right)^{-1} \mathbf{H}_i \mathbf{B}. \tag{6}$$

(4)   Estimation of $\boldsymbol{\Gamma}$.

The algorithm calculates the singular values of **X** through singular value decomposition (SVD) for each slice of the HSI data and sorts the singular values in descending order. Depending on the singular value of the sequence, we can obtain the value of the corresponding $\gamma_i (i = 1, \cdots, m)$. The γ_i corresponding to the larger singular value is 1, or is

otherwise 0. For HSI data, a large singular value distribution can be obtained by analyzing only one segment of its 2D slice signal in advance. The expression of $\boldsymbol{\Gamma}$ is

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_1\gamma_1 & \gamma_1\gamma_2 & \cdots & \gamma_1\gamma_m \\ \gamma_2\gamma_1 & \gamma_2\gamma_2 & \cdots & \gamma_2\gamma_m \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_m\gamma_1 & \gamma_m\gamma_2 & \cdots & \gamma_m\gamma_m \end{bmatrix}. \tag{7}$$

(5) Estimation of $\mathbf{S}_{opt}$.

$$\mathbf{X}^{opt} \leftarrow \mathbf{x}. \tag{8}$$

2.1. Two-Dimensional HSI Reconstruction Algorithm—L&S-bSBL (1)

Here, we consider a typical HSR scenario (push-broom imaging) in which there are M sensors with different wavelengths (called channels or bands; here, we have M bands) for acquiring HSI data $\mathcal{F} = [:, \mathbf{F}_q, :] \in \mathbb{R}^{MN \times Q}, q \in \{1, 2, \cdots, Q\}$ in an $N \times Q$ area following the Q-dimension, where $\mathbf{F}_q = [\mathbf{f}_1, \cdots, \mathbf{f}_M]^\top \in \mathbb{R}^{M \times N \times 1}$. Here, we simply express this as $\mathbf{F}_q = [\mathbf{f}_1, \cdots, \mathbf{f}_M]^\top \in \mathbb{R}^{M \times N}$. After all HSI data $\mathcal{F}$ are collected, the proposed algorithm performs the following steps on the HSI data, as shown in Figure 1.
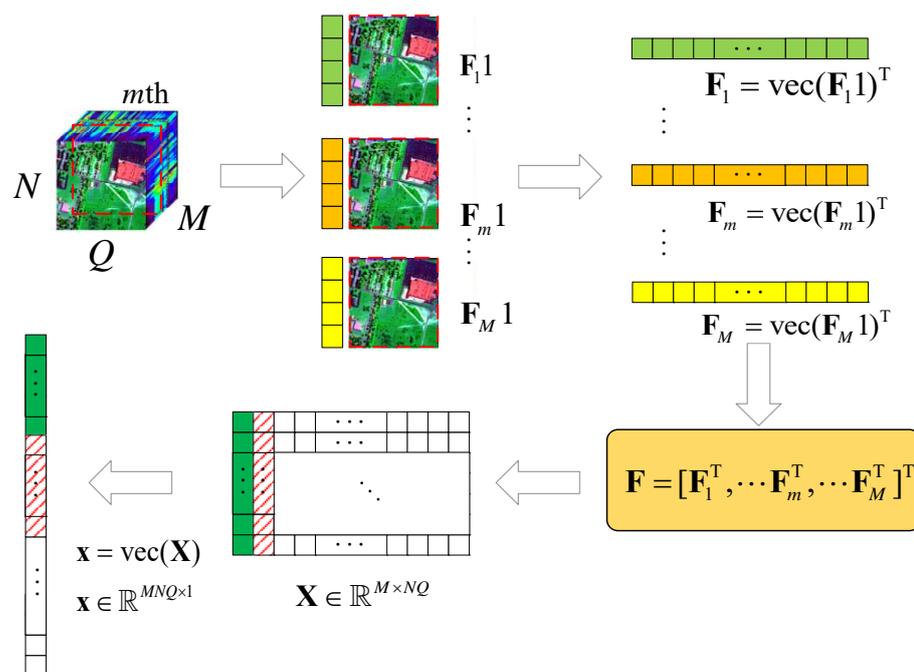
Figure 1. Schematic diagram of the HSI data tensor expanded into 2D data.

(1) According to the different bands M, we use the tensor to expand with $\mathbf{F}_1 1, \cdots, \mathbf{F}_m 1, \cdots, \mathbf{F}_M 1$;
(2) let $\mathbf{F}_m 1, m \in \{1, 2, \cdots, M\}$ turn into $\mathbf{F}_m, m \in \{1, 2, \cdots, M\}$ by using the vec operator;
(3) let $\mathbf{F} = [\mathbf{F}_1^\top, \mathbf{F}_2^\top, \cdots, \mathbf{F}_M^\top]$;
(4) obtain the corresponding value of $\mathbf{F}$ in the DCT domain $\mathbf{X} \in \mathbb{R}^{M \times NQ}$.

Here, we first conduct a principle component analysis (PCA) on the data $\mathbf{F}$. As shown in Figure 2, we can obtain 2D HSI tensor expansion data $\mathbf{F}$ in the sparse domain, which is mainly distributed in the first two columns, as well as the corresponding $\gamma_1 = 1$ and $\gamma_2 = 1$; the other $\gamma_i, i \in \{2, 3, \cdots, M\}$ is 0. Figure 3 shows a schematic diagram of the covariance matrix of $\mathbf{x}$.
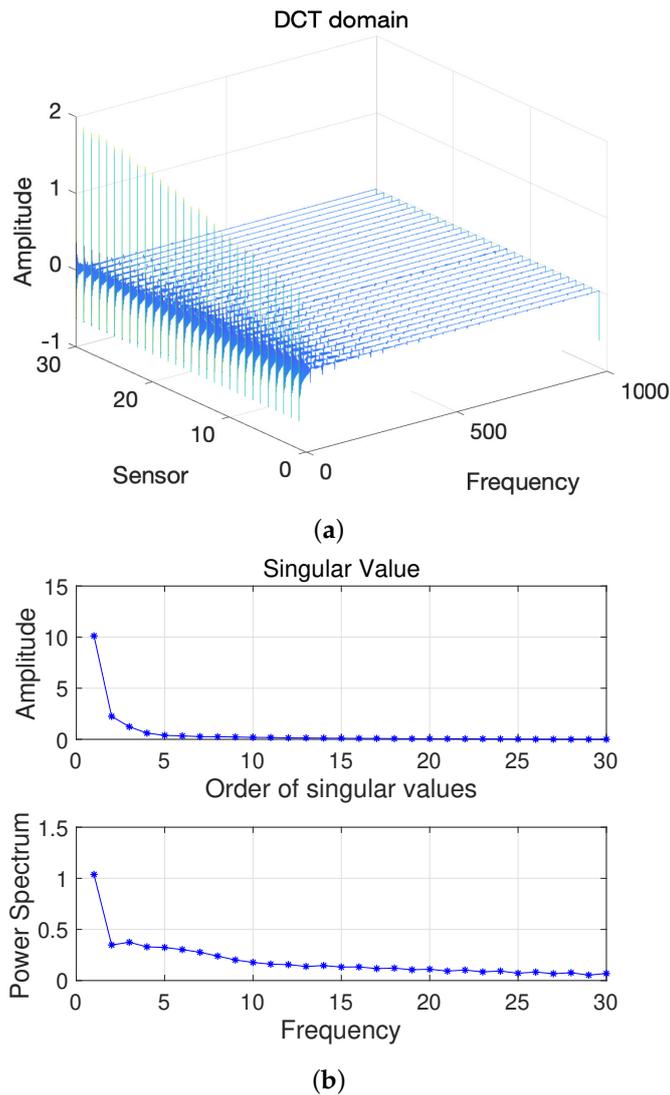
Figure 2. (**a**) Two-dimensional HSI in the DCT domain and (**b**) Singular value vs. Frequency.
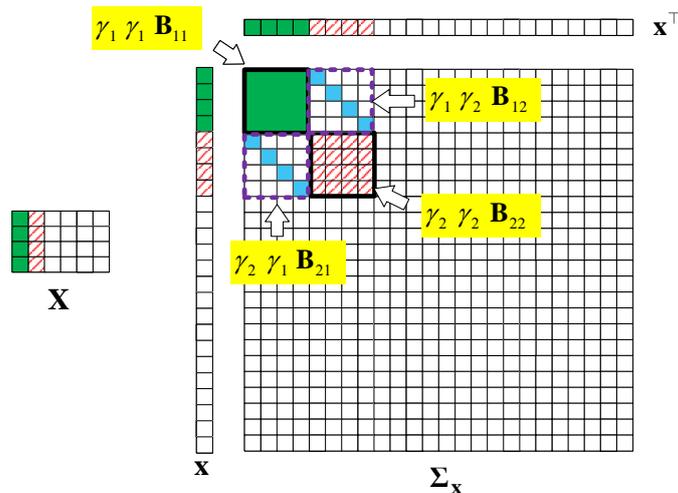
Figure 3. A covariance matrix of 2D HSI data **X**.

For the obtained $\mathbf{X} \in \mathbb{R}^{M \times NQ}$, the single measurement vector (SMV) model is used to solve the problem of data compression, transmission, and reconstruction. First, the

acquired data **X** are columnarized into **x**, encoded by the linear mapping matrix **H**, and then transmitted to the ground receiving station through a wireless channel. Here, we assume that the encoded data will be superimposed with noise **e** in the process of channel transmission, and the data received at the ground receiving station are denoted as **y**. Therefore, the mathematical model of the problem can be expressed as $\mathbf{y} = \mathbf{Hx} + \mathbf{e}$. The proposed algorithm uses the fast L&S-bSBL algorithm in Equations (1), (3), (4), (7) and (8) to iteratively reconstruct the original data **x**; then, we obtain **F**, $\mathcal{X}$. Most existing research regards the HSI signal as a low-rank signal to reconstruct, so the typical Bayesian low-rank reconstruction algorithm BARM [21] and the simultaneous low-rank and joint-sparse reconstruction algorithm SS&LR [22] are selected as the algorithms for comparison.

　　　Figure 4 shows a comparison of the reconstruction performance (including mean squared error (MSE) vs. signal-to-noise ratio (SNR) and runtime vs. SNR) of all algorithms in reconstructing 2D unfolded HSI data. In Figure 4a, we observe that the proposed algorithm L&S-bSBL (1) outperforms all of the other methods because the data expansion method of the proposed algorithm increases the amount of relevant data. As shown in Figure 4b, the proposed algorithm uses less runtime than BARM and almost the same amount as SS&LR. Comprehensive consideration of the two figures shows that the proposed algorithm is better than the other two algorithms in terms of reconstruction performance and computational resource consumption.

(**a**)

(**b**)
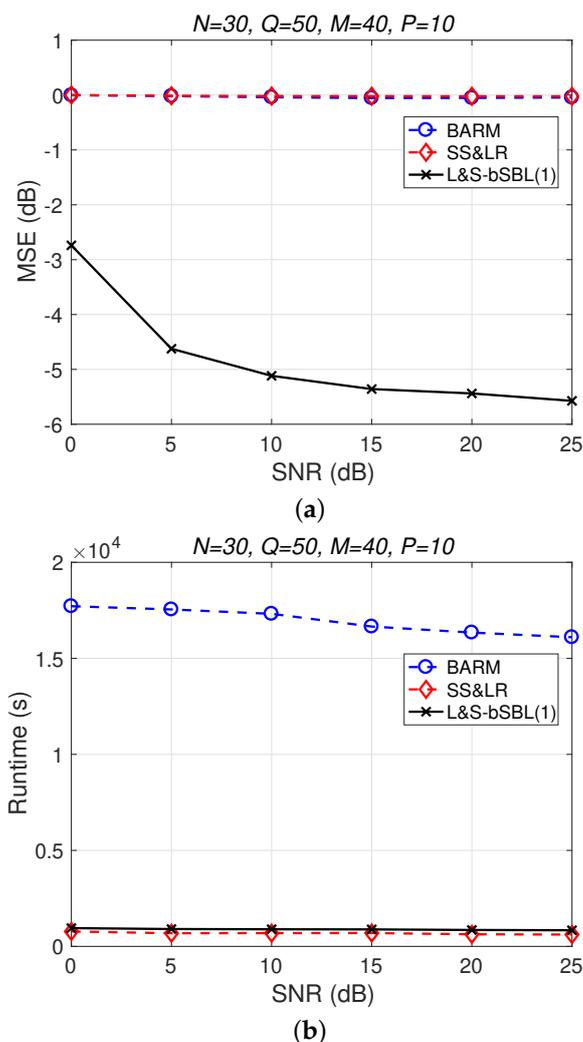
Figure 4. (**a**) MSE vs. SNR and (**b**) runtime vs. SNR for the reconstruction of 2D HSI data.

2.2. Two-Dimensional HSI Reconstruction Algorithm—L&S-bSBL (2)

Here, we also consider a typical HSR scenario (push-broom imaging) in which there are M sensors with different wavelengths (called channels or bands; here, we have M bands) to acquire HSI data $\mathcal{F} = [:, \mathbf{F}_q, :] \in \mathbb{R}^{MN \times Q}, q \in \{1, 2, \cdots, Q\}$ in an $N \times Q$ area following the Q-dimension, where $\mathbf{F}_q = [\mathbf{f}_1, \cdots, \mathbf{f}_M]^\top \in \mathbb{R}^{M \times N \times 1}$. After all HSI data $\mathcal{F}$ are collected, the proposed algorithm performs the following steps on the HSI data, as shown in Figure 5.

(1) Let slices $\mathbf{F}_q \in \mathbb{R}^{M \times N \times 1}$ and $q \in \{1, 2, \cdots, Q\}$ as $\mathbf{F}_q \in \mathbb{R}^{M \times N}$;
(2) obtain the corresponding value of $\mathbf{F}_q \in \mathbb{R}^{M \times N}$ in the DCT domain $\mathbf{X}_q, q \in \{1, 2, \cdots, Q\}$;
(3) let $\mathbf{X}_q$ turn into $\mathbf{x}_q, q \in \{1, 2, \cdots, Q\}$ by using the vec operator.
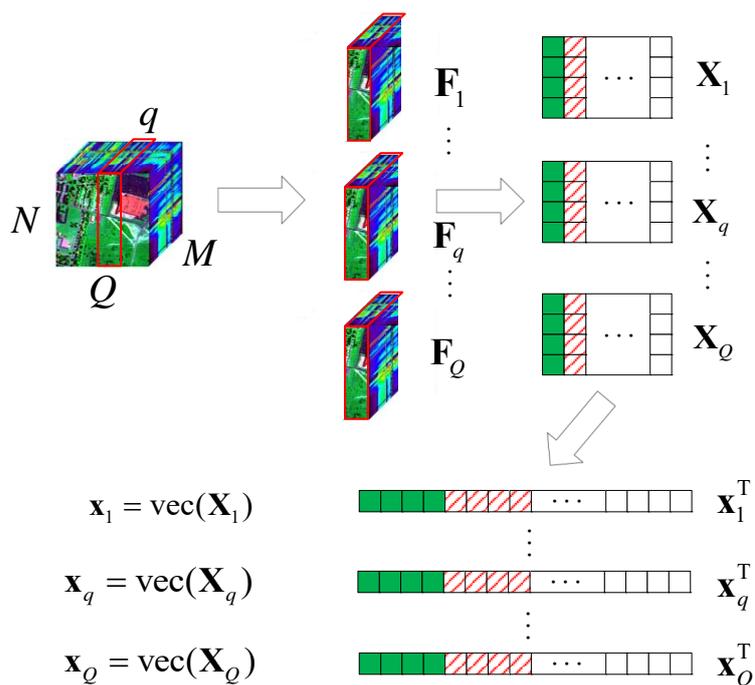
Figure 5. A schematic diagram of the HSI slice decomposition according to the push-broom imaging method following Q.

Here, we first conduct a PCA on the data $\mathbf{F}_q$. As shown in Figure 6, we can obtain 2D HSI slice data $\mathbf{F}_q$ in the sparse domain, which is mainly distributed in the first two columns, as well as the corresponding $\gamma_1 = 1$ and $\gamma_2 = 1$; the other $\gamma_i, i \in \{2, 3, \cdots, M\}$ is 0. The schematic diagram of the covariance matrix of $\mathbf{x}$ is similar to that shown in Figure 3.

For the obtained $\mathbf{X}_q \in \mathbb{R}^{M \times N}$, the SMV model is used to solve the problem of data compression, transmission, and reconstruction. Firstly, the acquired data $\mathbf{X}_q$ are columnarized into $\mathbf{x}_q$, encoded by the linear mapping matrix $\mathbf{H}$, and then transmitted to the ground receiving station through a wireless channel. Here, we assume that the encoded data will be superimposed with noise $\mathbf{e}_q$ in the process of channel transmission, and the data received at the ground receiving station are denoted as $\mathbf{y}_q$. Therefore, the mathematical model of the problem can be expressed as $\mathbf{y}_q = H\mathbf{x}_q + \mathbf{e}_q$. The proposed algorithm uses the fast L&S-bSBL algorithm in Equations (1), (3), (4), (7) and (8) to iteratively reconstruct the original data $\mathbf{x}$; then, we obtain $\mathbf{F}$, $\mathcal{X}$. Most existing research regards HSI signals as low-rank signals to recover, so BARM and SS&LR are selected as the algorithms for comparison.

(a)

(b)
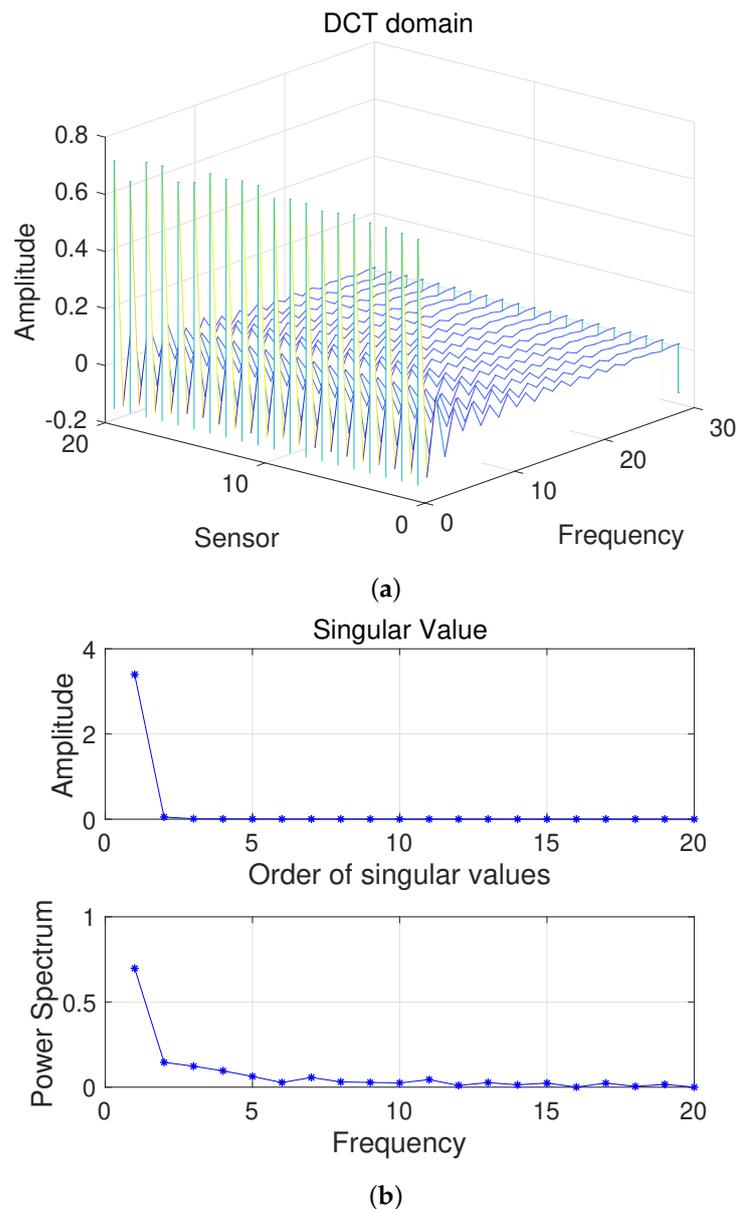
Figure 6. (**a**) Two-dimensional HSI slices in the DCT domain and (**b**) Singular value vs. Frequency.

Figure 7 shows a comparison of the reconstruction performance of all algorithms for all Q slices. In Figure 7a, we can see that the reconstruction of these slices is acceptable. However, L&S-bSBL (1) has better performance than L&S-bSBL (2) in reconstructing the same data, as shown in Figure 4. The reason is that L&S-bSBL (1) greatly utilizes the correlation between all data and improves the reconstruction performance, while L&S-bSBL (2) only uses the correlation within a certain band, ignoring the correlation between the bands at the same location.

Moreover, L&S-bSBL (2) only takes about 35 s compared with L&S-bSBL (1), which takes 150 s to recover the same data, as shown in Figure 4. This is because L&S-bSBL (1) has a large data dimension for a single operation and L&S-bSBL (2) has a small dimension, and the amount of calculation required for the calculation of the matrix inverse operation is not within an order of magnitude.

(a)

(b)
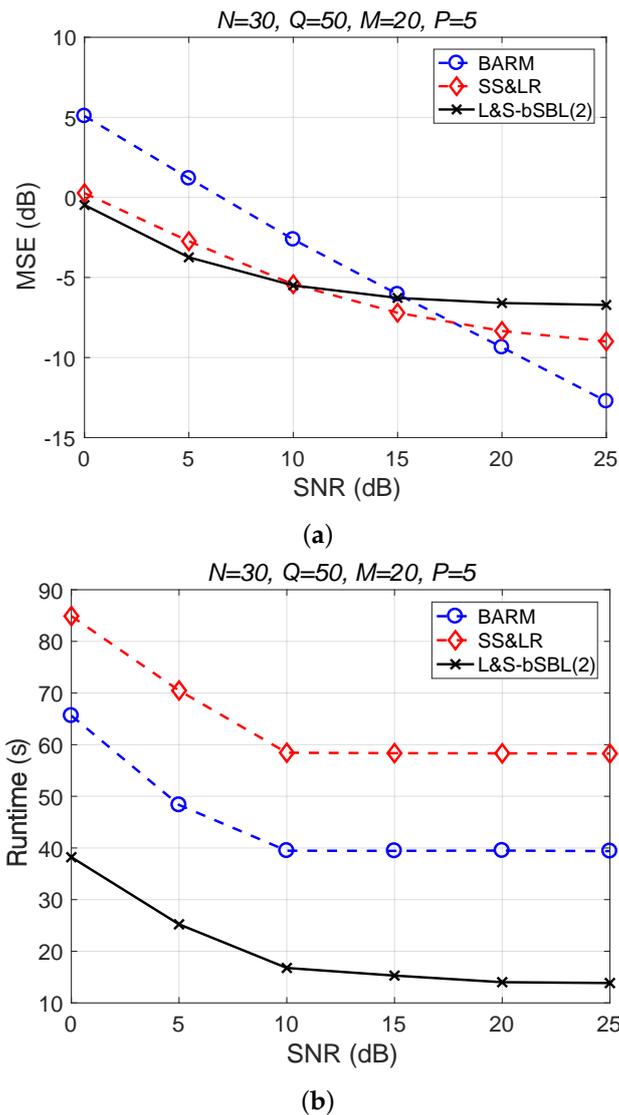
Figure 7. (**a**) MSE vs. SNR and (**b**) runtime vs. SNR for the reconstruction performance of 2D HSI push-brooming slices in all algorithms.

3. L&S Reconstruction Algorithm Combined with Acquisition Methods

3.1. Problem Formulation and Signal Model

We consider a typical HSR scenario (push-broom imaging) in which there are M sensors with different wavelengths (called channels or bands; here, we have M bands) to collect HSI data $\mathcal{F} = [:, \mathbf{F}_q, :] \in \mathbb{R}^{MN \times Q}, q \in \{1, 2, \cdots, Q\}$ in an $N \times Q$ area. Figure 8 shows that a spaceborne spectrometer (push-broom imaging) acquires multi-channel data $\mathbf{F}_q = [\mathbf{f}_1, \cdots, \mathbf{f}_M]^\top \in \mathbb{R}^{M \times N}$ with time synchronization among different regions following the Q-dimension, where $\mathbf{f}_m \in \mathbb{R}^{N \times 1}, m \in \{1, 2, \cdots, M\}$ stands for the data collected by the mth channel sensor and $\mathbf{F}_q$ is the spatially and temporally correlated data matrix. Then, $\mathbf{F}_q$ is encoded by linearly mixing with Ξ_q and transmitted to a ground receiving station, denoted as $\mathbf{Y}_q$, after superimposing noise $\mathbf{V}_q$. Finally, a novel CS method is used to decode all $\mathbf{F}_q$ at the same time at the ground processing center.

$$\mathbf{x}_q = \text{vec}[\mathbf{X}_q^\top]$$
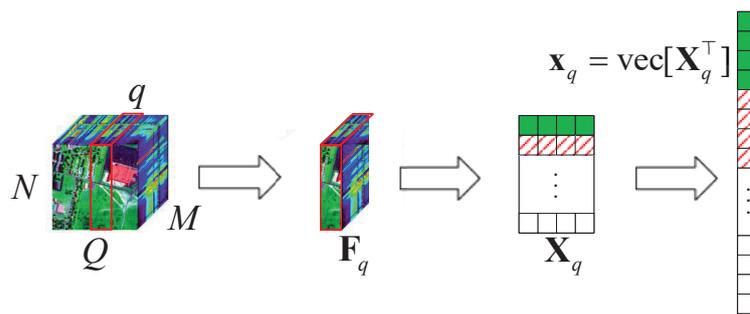
Figure 8. An example of hyperspectral image acquisition.

Fortunately, HSI data are highly correlated with the locations and bands of their non-zero elements in a sparse domain. So, we can get an approximately L&S matrix $\mathbf{X}_q \in \mathbb{R}^{M \times N}$ from $\mathbf{F}_q = \mathbf{\Psi}_q \mathbf{X}_q$, where $\mathbf{\Psi}_q \in \mathbb{R}^{M \times M}$ is a sparsifying basis (e.g., DCT matrix or wavelet matrix) [23], and $\mathcal{X} = [:, \mathbf{X}_q, :] \in \mathbb{R}^{MN \times Q}$. Thus, we have the following formulation:

$$\mathbf{Y}_q = \mathbf{\Phi}_q \mathbf{X}_q + \mathbf{V}_q, \tag{9}$$

where $\mathbf{\Phi}_q = \mathbf{\Xi}_q \mathbf{\Psi}_q$ is a known dictionary matrix. Here, this problem belongs to the multiple measurement vector (MMV) [24] problem.

Following our recent work [12], we now consider a bSBL framework [25] (i.e., L&S-bSBL) to reconstruct all $\mathbf{X}_q$. By letting $\mathbf{y}_q = \text{vec}[\mathbf{Y}_q^\top] \in \mathbb{R}^{P \times 1}$, $\mathbf{A}_q = \mathbf{\Phi}_q \otimes \mathbf{I}_n \in \mathbb{R}^{P \times NM}$, $\mathbf{x}_q = \text{vec}[\mathbf{X}_q^\top] \in \mathbb{R}^{NM \times 1}$, $\mathbf{v}_q = \text{vec}[\mathbf{V}_q^\top] \in \mathbb{R}^{P \times 1}$, after transforming the MMV problem to the block SMV [24] problem, we have

$$\mathbf{y}_q = \mathbf{A}_q \mathbf{x}_q + \mathbf{v}_q. \tag{10}$$

Then, the original problem becomes

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{11}$$

where $\mathbf{y} = [\mathbf{y}_1^\top, \cdots, \mathbf{y}_Q^\top]^\top \in \mathbb{R}^{QP \times 1}$, $\mathbf{A} = \mathbf{I}_Q \otimes \mathbf{A}_q (q \in \{1, 2, \cdots, Q\}) \in \mathbb{R}^{QP \times QNM}$, and $\mathbf{I}_Q \in \mathbb{R}^{Q \times Q}$ denotes a $Q \times Q$ identity matrix. $\mathbf{x} = [\mathbf{x}_1^\top, \cdots, \mathbf{x}_Q^\top]^\top \in \mathbb{R}^{QNM \times 1}$, $\mathbf{x}_q$ is the qth block in $\mathbf{x}$. The presence of K non-zero rows in $\mathbf{X}_q$ means that there are $K \times Q$ non-zero blocks in $\mathbf{x}$. Thus, $\mathbf{x}$ is a block sparse vector. $\mathbf{v} = [\mathbf{v}_1^\top, \cdots, \mathbf{v}_Q^\top]^\top \in \mathbb{R}^{QP \times 1}$.

We assume that the noise elements $\mathbf{v}_q, q \in \{1, \cdots, Q\}$ follow an identical and independent distribution with $p(\mathbf{v}_q) \sim \mathcal{N}(0, \lambda), \forall q$. We define the Gaussian likelihood for problem (11) as

$$p(\mathbf{y}|\mathbf{x}; \mathbf{A}, \lambda) \sim \mathcal{N}_{\mathbf{y}|\mathbf{x}}(\mathbf{A}\mathbf{x}, \lambda \mathbf{I}) \propto \exp[-\frac{1}{2\lambda}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2], \tag{12}$$

and the prior of $\mathbf{x}$ is given by

$$p(\mathbf{x}; \gamma_{qi}, \gamma_{qj}, \mathbf{B}_{qij}, \forall q, i, j) \sim \mathcal{N}_{\mathbf{x}}(\mathbf{0}, \mathbf{\Sigma}_0) \propto \exp[-\frac{1}{2}\mathbf{x}^\top \mathbf{\Sigma}_0^{-1}\mathbf{x}], \tag{13}$$

where $\mathbf{B}_{qij} \in \mathbb{R}^{N \times N}$ is a covariance matrix between $\mathbf{x}_{qi}$ and $\mathbf{x}_{qj}, i, j = 1, \cdots, M$. $\mathbf{\Sigma}_0 = \mathbf{I}_Q \otimes \mathbf{\Sigma}_q$, $\mathbf{\Sigma}_q = \mathbf{\Gamma}_q \otimes \mathbf{B}_{qij}$, $\mathbf{\Gamma}_q = \mathbf{\Gamma}_{q0}\mathbf{\Gamma}_{q0}^\top$. $\mathbf{\Gamma}_{q0} = [\gamma_{q1}, \cdots, \gamma_{qM}]^\top$ is the sparsity pattern vector of $\mathbf{X}_q$, where the support indicates $\gamma_{qi} \in \{0, 1\}, i = 1, \cdots, M$.

Typically, we can obtain

$$\Sigma_0 = \text{diag}(\Sigma_1, \cdots, \Sigma_q, \cdots, \Sigma_Q),$$

$$\Sigma_q = \begin{bmatrix} \gamma_{q1}\gamma_{q1}\mathbf{B}_{q11} & \gamma_{q1}\gamma_{q2}\mathbf{B}_{q12} & \cdots & \gamma_{q1}\gamma_{qM}\mathbf{B}_{q1M} \\ \gamma_{q2}\gamma_{q1}\mathbf{B}_{q21} & \gamma_{q2}\gamma_{q2}\mathbf{B}_{q22} & \cdots & \gamma_{q2}\gamma_{qM}\mathbf{B}_{q2M} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{qM}\gamma_{q1}\mathbf{B}_{qM1} & \gamma_M\gamma_2\mathbf{B}_{qM2} & \cdots & \gamma_{qM}\gamma_{qM}\mathbf{B}_{qMM} \end{bmatrix}. \tag{14}$$

Without losing generality, we assume that $\mathbf{X}_q$ is a matrix that is low-rank and joint-sparse in columns. Figure 9 illustrates an example of $\mathbf{X}_q$ and the structure of the covariance matrix Σ_q of $\mathbf{x}_q$ with $M = 4$. In the figure, we find that only the first two columns have values, so only $\gamma_{q1} = 1$ and $\gamma_{q2} = 1$. Thus, the parts related to γ_{q1}, γ_{q2} in Σ_q are valuable.
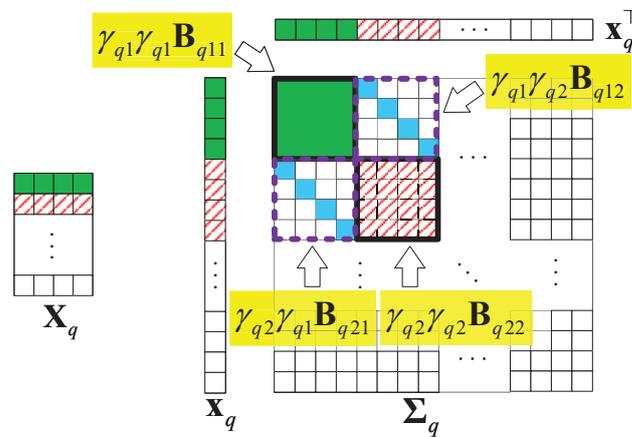
Figure 9. An example of the structure of the covariance matrix Σ_q of $\mathbf{x}_q$, $\mathbf{x}_q = \text{vec}[\mathbf{X}_q^\top]$, $M = 4$.

3.2. Proposed Method

Following our last work [12], we obtain the posterior density of $\mathbf{x}$ with the Bayesian rule:

$$p(\mathbf{x}|\mathbf{y}; \lambda, \gamma_q i, \gamma_q j, \mathbf{B}_{qij}, \forall q, i, j) \sim \mathcal{N}_\mathbf{x}(\boldsymbol{\mu}_\mathbf{x}, \Sigma_\mathbf{x}), \tag{15}$$

where the mean $\boldsymbol{\mu}_\mathbf{x}$ and the covariance $\Sigma_\mathbf{x}$ can be obtained by

$$\boldsymbol{\mu}_\mathbf{x} = \frac{1}{\lambda}\Sigma_\mathbf{x}\mathbf{A}^\top\mathbf{y}, \tag{16}$$

$$\begin{aligned} \Sigma_\mathbf{x} &= (\Sigma_0^{-1} + \frac{1}{\lambda}\mathbf{A}^\top\mathbf{A})^{-1} \\ &= \Sigma_0 - \Sigma_0\mathbf{A}^\top(\lambda\mathbf{I} + \mathbf{A}\Sigma_0\mathbf{A}^\top)^{-1}\mathbf{A}\Sigma_0. \end{aligned} \tag{17}$$

When all of the hyperparameters $\lambda, \gamma_{qi}, \gamma_{qj}, \mathbf{B}_{qij}, \forall q, i, j$ are given, the maximum a posteriori (MAP) estimate of $\mathbf{x}$ can be obtained by

$$\begin{aligned} \hat{\mathbf{x}} = \text{vec}[\hat{\mathcal{X}}^\top] \triangleq \boldsymbol{\mu}_\mathbf{x} &= (\lambda\Sigma_0^{-1} + \mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}\Sigma_0 \\ &= \Sigma_0\mathbf{A}^\top(\lambda\mathbf{I} + \mathbf{A}\Sigma_0\mathbf{A}^\top)^{-1}\mathbf{y}, \end{aligned} \tag{18}$$

where Σ_0 is the approximate diagonal matrix obtained by Equation (14), with most block matrices being zeros. It is clear that $\gamma_{qi}, \gamma_{qj}, \forall q, i, j$ control the sparsity of the blocks of $\hat{\mathbf{x}}$. When $\gamma_{qk} = 0$, the associated qkth block in $\hat{\mathbf{x}}$ becomes zero. In fact, $\gamma_{1i} = \cdots = \gamma_{qi} = \cdots = \gamma_{Qi}$.

Following the bSBL framework [25], to avoid overfitting, we use a common positive definite matrix Σ to model all of the covariance matrices Σ_q. In Σ_q, we also use a common positive definite matrix $\mathbf{B}$ instead of $\mathbf{B}_{qij}$ and we use a group of $\gamma_i \in \{0, 1\}$ instead of all $\gamma_{qi} \in \{0, 1\}$, $\mathbf{\Gamma}$ and instead of $\mathbf{\Gamma}_q$. From the previous analysis, we can find that the covariance

matrices induce the spatiotemporal correlation in the prior density. Thus, Equation (14) can be written as

$$\boldsymbol{\Sigma}_0 = \mathbf{I}_Q \otimes \boldsymbol{\Gamma} \otimes \mathbf{B}. \tag{19}$$

Using the Bayesian strategy, we maximize the marginal probability $p(\mathbf{x}, \mathbf{y})$ of $\mathbf{x}$:

$$\max_{\mathbf{B} \in \mathbf{N}^+, \boldsymbol{\Gamma} \geq 0} \int p(\mathbf{y}|\mathbf{x}; \mathbf{A}, \lambda) p(\mathbf{x}; \boldsymbol{\Gamma}, \mathbf{B}) d\mathbf{x}, \tag{20}$$

which is equivalent to minimizing the cost function of $-2 \log p(\mathbf{y}; \lambda, \boldsymbol{\Gamma}, \mathbf{B})$:

$$\mathcal{L}(\boldsymbol{\Gamma}, \mathbf{B}, \lambda) = \mathbf{y}^\top \boldsymbol{\Sigma}_y^{-1} \mathbf{y} + \log |\boldsymbol{\Sigma}_y|, \tag{21}$$

where $\mathbf{N}^+$ denotes a set of $N \times N$ positive definite matrices.

$$\boldsymbol{\Sigma}_y = \mathbf{A} \boldsymbol{\Sigma}_0 \mathbf{A}^\top + \lambda \mathbf{I}, \quad \boldsymbol{\Sigma}_0 = \mathbf{I}_Q \otimes \boldsymbol{\Gamma} \otimes \mathbf{B}. \tag{22}$$

Here, $\boldsymbol{\Sigma}_y$ denotes the covariance of $\mathbf{y}$ given $\boldsymbol{\Gamma}$ and $\mathbf{B}$.
Simply, $\boldsymbol{\Theta} \equiv \{\boldsymbol{\Gamma}, \mathbf{B}, \lambda\}$; thus, (21) turns into

$$\mathcal{L}(\boldsymbol{\Theta}) = \mathbf{y}^\top \boldsymbol{\Sigma}_y^{-1} \mathbf{y} + \log |\boldsymbol{\Sigma}_y|. \tag{23}$$

Firstly, $\mathbf{x}$ is treated as a hidden variable in the expectation maximization (EM) formulation that proceeds, and we maximize

$$\begin{aligned}
\mathcal{Q}(\boldsymbol{\Theta}) =& E_{\mathbf{x}|\mathbf{y}; \boldsymbol{\Theta}^{(pre)}} [\log(p(\mathbf{y}|\mathbf{x}; \lambda) p(\mathbf{x}; \boldsymbol{\Gamma}, \mathbf{B}))] \\
=& E_{\mathbf{x}|\mathbf{y}; \boldsymbol{\Theta}^{(pre)}} [\log p(\mathbf{y}|\mathbf{x}; \lambda)] \\
&+ E_{\mathbf{x}|\mathbf{y}; \boldsymbol{\Theta}^{(pre)}} [\log p(\mathbf{x}; \boldsymbol{\Gamma}, \mathbf{B})],
\end{aligned} \tag{24}$$

where $\boldsymbol{\Theta}^{(pre)}$ denotes the hyperparameters that are estimated in the previous iteration.

To estimate λ—only the first term in the $\mathcal{Q}$ function is correlated with λ—it can be simplified as

$$\begin{aligned}
\mathcal{Q}(\lambda) =\ & E_{\mathbf{y}|\mathbf{x}; \boldsymbol{\Theta}^{(pre)}} [\log p(\mathbf{y}|\mathbf{x}; \lambda)] \propto -\frac{QP}{2} \log \lambda \\
& -\frac{1}{2\lambda} \left[\|\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_\mathbf{x}\|_2^2 + \lambda^{(pre)} [QNM - \mathrm{tr}(\boldsymbol{\Sigma}_\mathbf{x} \boldsymbol{\Sigma}_0^{-1})] \right],
\end{aligned} \tag{25}$$

where $\lambda^{(pre)}$ denotes the estimation of λ in the previous iteration. When we calculate the derivative of Equation (25) over λ and set it equal to zero, λ is obtained:

$$\lambda \leftarrow \frac{\|\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_\mathbf{x}\|_2^2 + \lambda^{(pre)} \left[QNM - \mathrm{tr}(\boldsymbol{\Sigma}_\mathbf{x} \boldsymbol{\Sigma}_0^{-1}) \right]}{PN}. \tag{26}$$

To estimate $\boldsymbol{\Gamma}$ and $\mathbf{B}$, $\boldsymbol{\Gamma} = \mathrm{diag}(\gamma_1^2, \cdots, \gamma_M^2)$ is first assumed, where $\mathrm{diag}(\cdot)$ denotes a diagonal matrix operator. Notice that only the second term in Equation (24) is related to $\boldsymbol{\Gamma}$ and $\mathbf{B}$. So, we can simplify the $\mathcal{Q}$ function (24) to

$$\mathcal{Q}(\boldsymbol{\Gamma}, \mathbf{B}) = E_{x|y; \boldsymbol{\Theta}^{(pre)}} [\log p(\mathbf{x}; \boldsymbol{\Gamma}, \mathbf{B})], \tag{27}$$

Then, we have

$$\begin{aligned}
\mathcal{Q}(\boldsymbol{\Gamma}, \mathbf{B}) \propto -&\frac{QN}{2} \log(|\boldsymbol{\Gamma}|) - \frac{QM}{2} \log(|\mathbf{B}|) \\
-&\frac{1}{2} \mathrm{tr}[(\boldsymbol{\Gamma}^{-1} \otimes \mathbf{B}^{-1})(\boldsymbol{\Sigma}_\mathbf{x} + \boldsymbol{\mu}_\mathbf{x} \boldsymbol{\mu}_\mathbf{x}^\top)].
\end{aligned} \tag{28}$$

To obtain the values of γ_i, we use the feature given in [20]. Specifically, the values of γ_i corresponding to the larger singular values (the larger singular values are defined as singular values that are larger than 10% of the largest singular value) are 1; otherwise, they are 0. Figure 10 gives a schematic diagram of how γ_i is obtained from HSI data. From Figure 10, we can find that only the first singular value is larger, so $\gamma_1 = 1$.
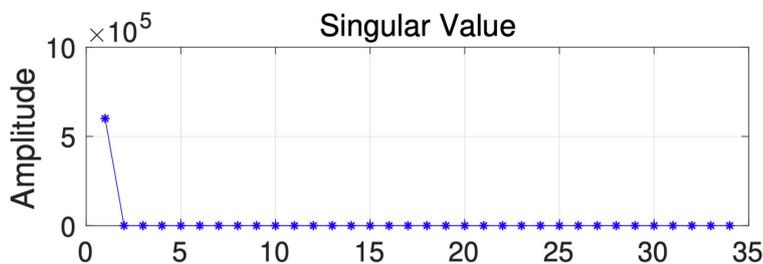
Figure 10. The schematic diagram of how γ_i is obtained.

To estimate $\mathbf{B}$, $\mu_{\mathbf{x}}$ and $\Sigma_{\mathbf{x}}$ are plugged into Equation (28). So, we can obtain the gradients of Equation (28) over $\mathbf{B}$, and then we can obtain $\mathbf{B}^{(pre)}$.

$$\mathbf{B}^{(pre)} \leftarrow \frac{1}{M} \sum_{i=1}^{M} \frac{\left(\Sigma_{\mathbf{x}}^i + \mu_{\mathbf{x}}^i (\mu_{\mathbf{x}}^i)^\top\right)}{\gamma_i^2}. \tag{29}$$

Thus, we will get $\Gamma^{(pre)}$. Using the same method, we can get $\lambda^{(pre)}$. Finally, we get $\Theta^{(pre)}$. Here, $\mathbf{A}^{(pre)}$ denotes an initial value of $\mathbf{A}$.

In order to get a closed form of Θ, we employ standard upper bounds to solve Equation (21), which is known as a non-convex optimization problem leading to an EM-like algorithm. For the first and second terms of $\mathcal{L}(\Gamma, \mathbf{B})$, we compute their respective bounds.

For the first term in Equation (21), we can obtain

$$\mathbf{y}^\top \Sigma_y^{-1} \mathbf{y} \leq \frac{1}{\lambda} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mathbf{x}^\top \Sigma_0^{-1} \mathbf{x}, \tag{30}$$

where equality is obtained when $\mathbf{x}$ satisfies Equation (18).

For the second term, we can obtain

$$\begin{aligned}
\log|\Sigma_y| &\equiv QM\log|\mathbf{B}| + \log|\lambda \mathbf{A}^\top \mathbf{A} + \Sigma_0^{-1}| \\
&\leq QM\log|\mathbf{B}| + \mathrm{tr}[\mathbf{B}^{-1}\nabla_{\mathbf{B}^{-1}}] + C,
\end{aligned} \tag{31}$$

where, for the second term $\log|\lambda \mathbf{A}^\top \mathbf{A} + \Sigma_0^{-1}|$, a first-order approximation is used to approximate it with a bias term C. The equality will hold when the gradient satisfies

$$\nabla_{\mathbf{B}^{-1}} = \sum_{m=1}^{M} \mathbf{B} - \mathbf{B}\mathbf{A}_m^\top (\mathbf{A}\Sigma_0 \mathbf{A}^\top + \lambda \mathbf{I})^{-1} \mathbf{A}_m \mathbf{B}, \tag{32}$$

where $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_M]$ and $\mathbf{A}_m \in \mathbb{R}^{QP \times QN}$. Finally, by using the upper bounds of Equations (30) and (31) and $\nabla_{\mathbf{B}^{-1}}$, we can obtain the optimal $\mathbf{B}$ in a closed form:

$$\mathbf{B}^{opt} = \arg \min_{\mathbf{x}} \mathbf{x}^\top \Sigma_0^{-1} \mathbf{x} + \mathrm{tr}[\mathbf{B}^{-1}\nabla_{\mathbf{B}^{-1}}] + QM\log|\mathbf{B}|. \tag{33}$$

Starting with $\mathbf{B} = \mathbf{B}^{(pre)}$, we iteratively compute Equations (18), (32) and (33), then obtain an estimation for $\mathbf{B}$ and a corresponding estimation for $\mathbf{x}$ given by Equation (18). Here, the proposed method is outlined in Algorithm 1.

Algorithm 1 Proposed method

Source data analysis
calculate singular values of $\mathbf{X}_q$ by using SVD;
obtain γ_i, $i = 1, \cdots, M$ by using singular values;
Input $\mathbf{y}, \mathbf{A}$; **Output** $\mathcal{X}$;
Initialize assume $\mathbf{\Gamma} = \text{diag}(\gamma_1^2, \cdots, \gamma_M^2)$;
iters $= 0$, $\delta = 10^{-6}$;
max iteration number $= 500$;
Set $\lambda, \mathbf{B}$ by $\lambda = 10^{-10}$, $\mathbf{B} = ones(N, N)$;
compute $\lambda, \mathbf{B}$ from Equations (26) and (33); $\mathbf{\Sigma}_0 \leftarrow \mathbf{I}_Q \otimes \mathbf{\Gamma} \otimes \mathbf{B}$;
While $\|\mathcal{X} - \hat{\mathcal{X}}\|_2^2 \geq \delta$
compute $\hat{\mathcal{X}}$ by Equation (18); compute $\nabla_{\mathbf{B}^{-1}}$ by Equation (32);
compute $\mathbf{B}^{opt}$ by Equation (33); *iters* $=$ *iters* $+ 1$;
if *iters* ≥ 500
    STOP;
end if
EndWhile
Get the best $\mathbf{B}^{opt}$ and $\mathcal{X}$.

3.3. Simulation Experiments

In this section, we present the results with HSI data in order to compare the performance of the proposed method with that of the prior state-of-the-art LRMR [5] and SS&LR [22] methods. For HSI data, 100 continuous-time trials were run. In each trial, data of a 30×50 two-dimensional area and 20 bands from the Salinas Database (available at http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas-A_scene (accessed on 10 November 2021)) and Indian Pines (available at http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Indian_Pines (accessed on 10 November 2021)) were used. Thus, the data matrix was a $30 \times 50 \times 20$ matrix.

Figures 11 and 12 plot the performance in terms of MSE versus SNR and runtime versus SNR, respectively. Here, $P = 10$, so the ratio of compression is $P/NM = 1/60$. In Figure 11a, we can observe that proposed method outperforms all of the other methods in terms of MSE; e.g., at SNR $= 20$ dB, we note that the proposed method achieves reconstruction gain at least 5 dB greater than those of the other methods. Figure 11b shows that the proposed method uses less runtime than the others. In Figure 12, we can observe that the effect of SS&LR is not good, but our proposed method works well and outperforms all of the other methods in terms of MSE; e.g., at SNR $= 15$ dB, we note that the proposed method achieves a reconstruction gain at least 6 dB greater than those of the other methods. The proposed method also uses less runtime than the others.

Figures 13 and 14 compare all of the methods for different values of compression P. Figure 13a shows that most of the MSE curves fall as P increases. As the compression ratio P/NM decreases (from 1/300 to 1/60), the reconstruction error becomes smaller, affecting the signal reconstruction performance. Fortunately, the proposed algorithm yields better performance in all cases and has a higher compression ratio with the same reconstruction performance. Figure 13b shows that the proposed method consumes less computation time at the same compression ratio. From Figure 14, we observe that the effects of SS&LR and LRMR are not good, but our proposed method outperforms all of the other methods in terms of MSE; e.g., at $P = 6$, we note that the proposed method achieves a reconstruction gain at least 7 dB greater than those of the other methods. The proposed method also uses less runtime than the others.

Figure 15 shows some visual reconstruction results obtained with the three methods. As expected, the proposed method preserves more fine details and better visual results than the others.
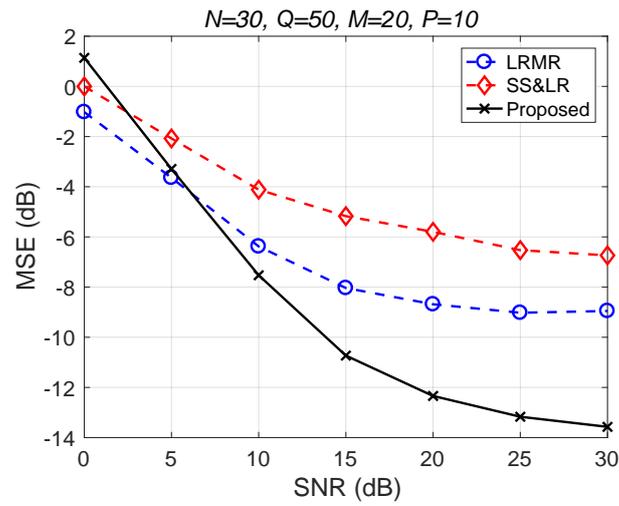
(**a**)
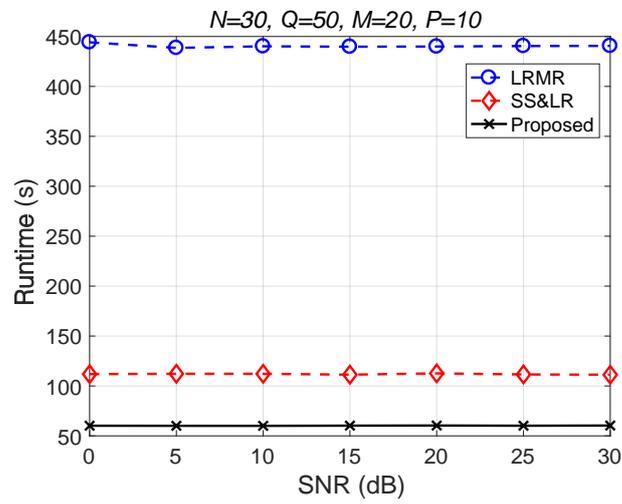
(**b**)

Figure 11. (**a**) MSE vs. SNR and (**b**) runtime vs. SNR for the reconstruction of the Salinas data.

(**a**)

Figure 12. *Cont.*

(**b**)
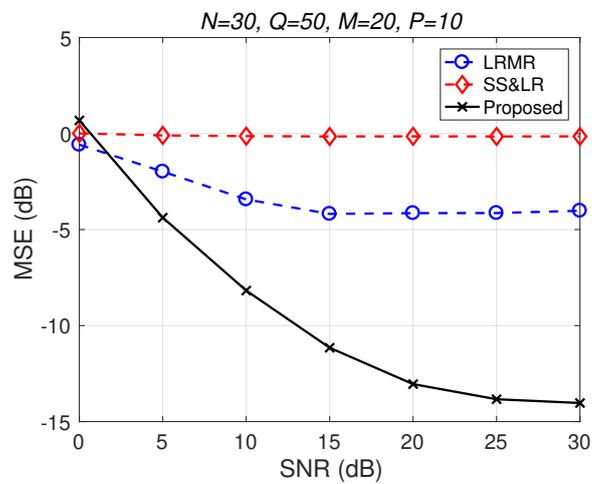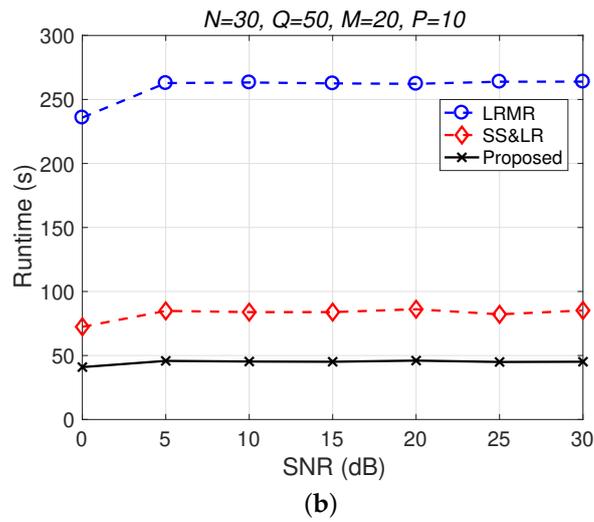
Figure 12. (**a**) MSE vs. SNR and (**b**) runtime vs. SNR for the reconstruction of the Indian Pines data.

(**a**)

(**b**)
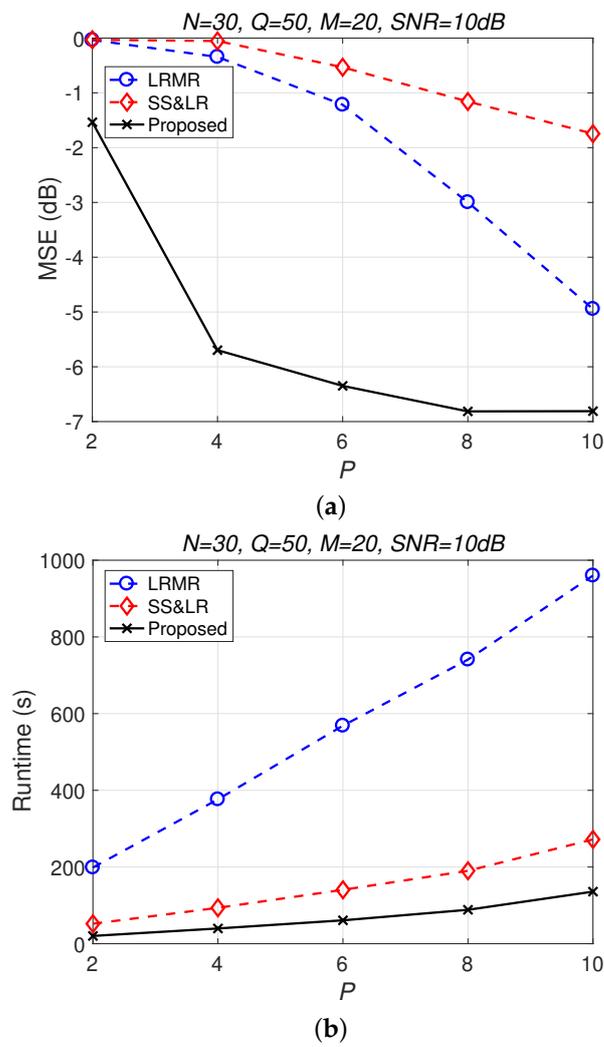
Figure 13. (**a**) MSE vs. P and (**b**) runtime vs. P with different compression ratios for the reconstruction of the Salinas data.

(a)

(b)
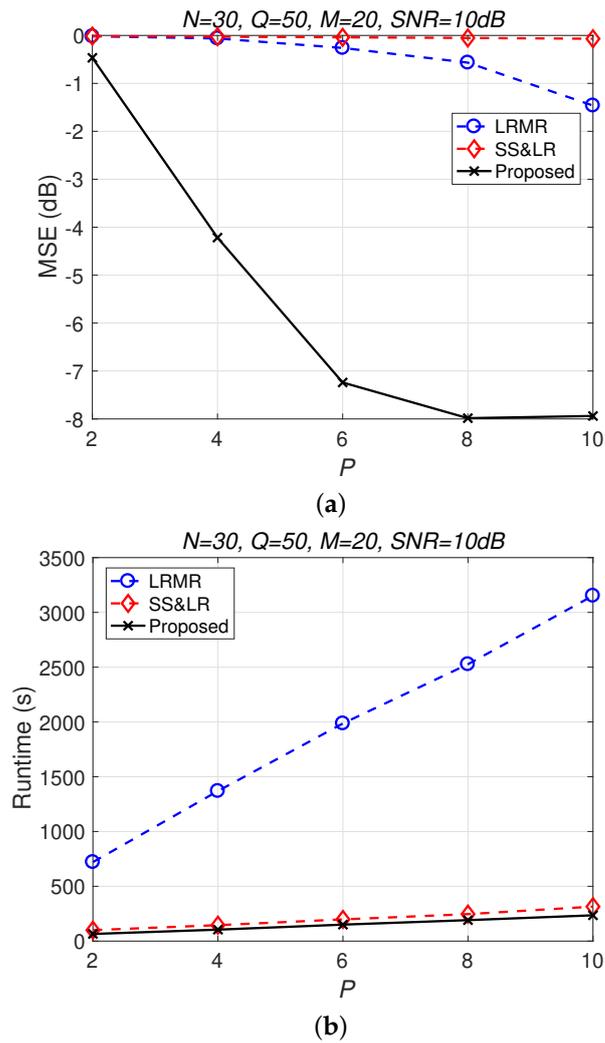
Figure 14. (**a**) MSE vs. *P* and (**b**) runtime vs. *P* with different compression ratios for the reconstruction of the Indian Pines data.
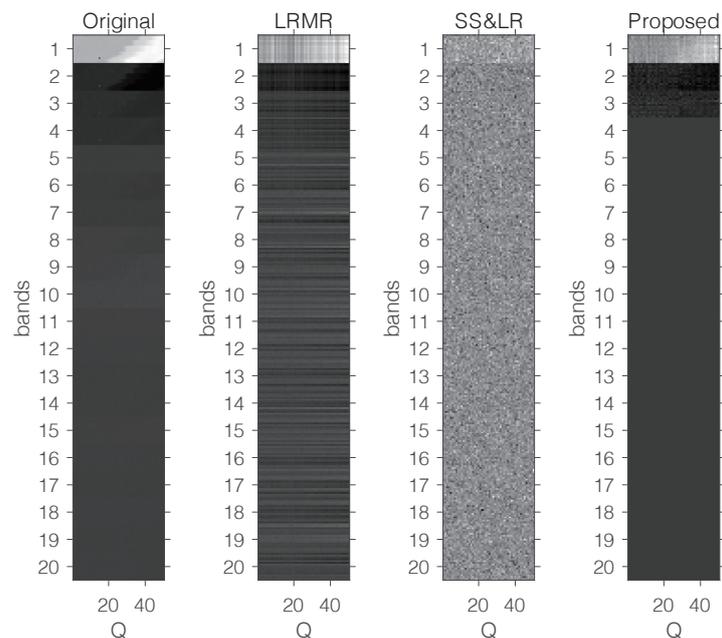
Figure 15. Visual comparison with a fixed compression ratio $P/NM = 1/30$ and 20 bands.

Figure 16 only shows a performance comparison chart of the various algorithms from the perspective of data recovery, and it does not start from the perspective of data collection methods, as in the beginning of this section. It can be seen in Figure 16 that the proposed 3D joint reconstruction algorithm has better performance and less time consumption than the other algorithms. What needs to be explained here is that the three methods—the proposed 3D HSI method, L&S-bSBL (1) method, and L&S-bSBL (2) method—are, in terms of the compression ratio: $P/(NM) = 1/60$, $P/(NQ) = 1/150$, and $P/N = 1/3$.

(**a**)

(**b**)
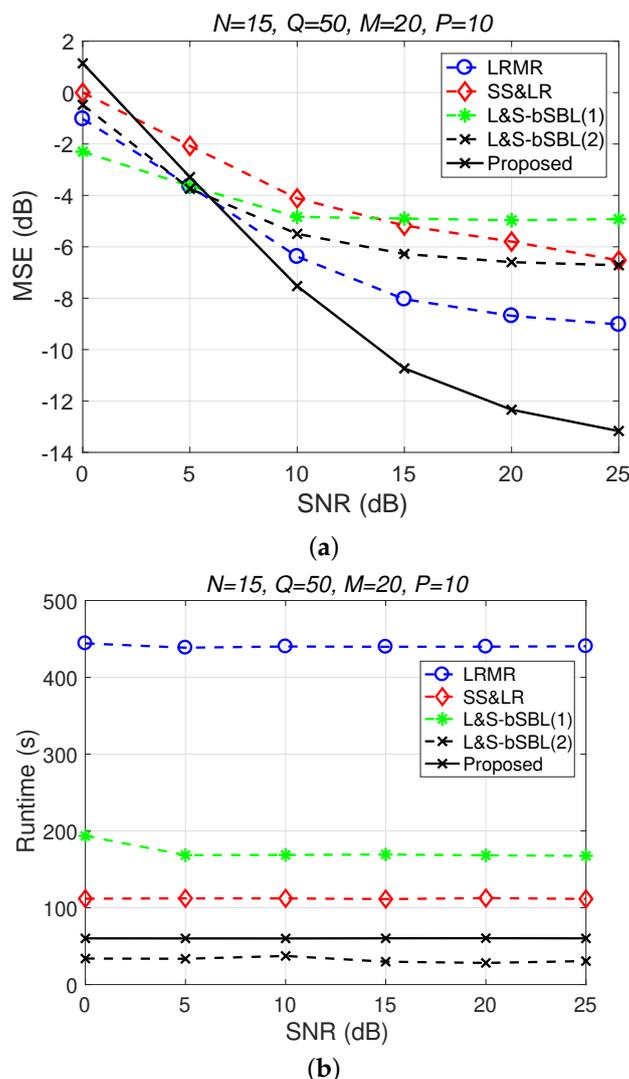
Figure 16. (**a**) MSE vs. SNR and (**b**) runtime vs. SNR for only the performance of each algorithm from the perspective of signal recovery, regardless of the HSI data acquisition method.

4. Discussion

From Section 3.3, the simulation results demonstrate that the proposed method is superior to LRMR and SS&LR methods in terms of reconstruction accuracy and computational burden under the same SNR and compression ratio. This is because LRMR only focuses on the reconstruction of low-rank signals, while SS&LR focuses on the reconstruction of low-rank and sparse signals, but it only uses the spatial relationship or interband relationship of HSI signals and does not combine the signal acquisition process.

Based on our present work, some issues that could be addressed in future research can be summarized as follows. One open problem is that of reducing, as much as possible, the computational complexity of the proposed method. In particular, unmanned aerial vehicles (UAVs) are used to monitor fields, such as in farmland monitoring, environmental

monitoring, etc. While our approach has significant reconstruction performance, there is still a long way to go before we can apply it to future portable hardware. We can also take advantage of the properties of HSI data with higher-dimensional structures to design an effective Bayesian learning-based reconstruction method; namely, tensor models could be considered to compress and reconstruct HSI data. Furthermore, we may also consider the design of a sensor selection method for HSI acquisition that reduces the number of sensors in different locations that are needed so as to extend the life and energy savings of sensors.

5. Conclusions

In this work, based on the fast method proposed in our previous work, two methods for 2D compression and reconstruction of 3D HSI data are given. Then, starting from the HSI acquisition method, we analyzed the push-brooming acquisition of data slices in the hyperspectral image acquisition process, and by assuming that the HSI data met the L&S model, a joint CS compression and reconstruction method for spatial and spectral correlation was studied and an L&S structure was proposed. The proposed method combines data reconstruction with HSI data acquisition. The simulation results show that the proposed method has better performance than the other two existing methods and the two proposed 2D reconstruction methods.

References

1. Chang, C.I. *Hyperspectral Data Exploitation: Theory and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
2. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–36. [CrossRef]
3. Goetz, A.F. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sens. Environ.* **2009**, *113*, S5–S16. [CrossRef]
4. Borengasser, M.; Hungate, W.S.; Watkins, R. *Hyperspectral Remote Sensing: Principles and Applications,* 1st ed.; CRC Press: Boca Raton, FL, USA, 2007.
5. Zhang, H.; He, W.; Zhang, L.; Shen, H.; Yuan, Q. Hyperspectral image restoration using low-rank matrix recovery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4729–4743. [CrossRef]
6. Golbabaee, M.; Vandergheynst, P. Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 2741–2744.
7. Li, C.; Ma, L.; Wang, Q.; Zhou, Y.; Wang, N. Construction of sparse basis by dictionary training for compressive sensing hyperspectral imaging. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium—IGARSS, Melbourne, Australia, 21–26 July 2013; pp. 1442–1445.
8. Hou, Y.; Zhang, Y. Effective hyperspectral image block compressed sensing using thress-dimensional wavelet transform. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec, QC, Canada, 13–18 July 2014; pp. 2973–2976.
9. Donoho, D. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]
10. Baraniuk, R.G. Compressive Sensing [Lecture Notes]. *IEEE Signal Process. Mag.* **2007**, *24*, 118–121. [CrossRef]
11. Candes, E.J.; Wakin, M.B. An Introduction To Compressive Sampling. *IEEE Signal Process. Mag.* **2008**, *25*, 21–30. [CrossRef]

12. Zhang, Y.; Li, Y.; Yin, C.; He, K. Low-rank and joint-sparse signal recovery for spatially and temporally correlated data using sparse Bayesian learning. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 4719–4723.
13. Wang, L.; Feng, Y.; Gao, Y.; Wang, Z.; He, M. Compressed Sensing Reconstruction of Hyperspectral Images Based on Spectral Unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1266–1284. [CrossRef]
14. Wang, Z.; Xiao, H.; He, M.; Wang, L.; Xu, K.; Nian, Y. Spatial-Spectral Joint Compressed Sensing for Hyperspectral Images. *IEEE Access* **2020**, *8*, 149661–149675. [CrossRef]
15. Xiao, H.; Wang, Z.; Cui, X. Distributed Compressed Sensing of Hyperspectral Images According to Spectral Library Matching. *IEEE Access* **2021**, *9*, 112994–113006. [CrossRef]
16. Yang, S.; Wang, M.; Li, P.; Jin, L.; Wu, B.; Jiao, L. Compressive Hyperspectral Imaging via Sparse Tensor and Nonlinear Compressed Sensing. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5943–5957. [CrossRef]
17. He, W.; Zhang, H.; Shen, H.; Zhang, L. Hyperspectral Image Denoising Using Local Low-Rank Matrix Recovery and Global Spatial–Spectral Total Variation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 713–729. [CrossRef]
18. Takeyama, S.; Ono, S. Joint Mixed-Noise Removal and Compressed Sensing Reconstruction of Hyperspectral Images via Convex Optimization. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1492–1495.
19. Fu, W.; Li, S. Semi-Tensor Compressed Sensing for Hyperspectral Image. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 2737–2740.
20. Zhang, Y.; Huang, L.; Li, Y.; Zhang, K.; Yin, C. Fast recovery of low-rank and joint-sparse signals in wireless body area networks. In Proceedings of the 2020 IEEE/CIC International Conference on Communications in China (ICCC), Chongqing, China, 9–11 August 2020; pp. 577–581.
21. Xin, B.; Wang, Y.; Gao, W.; Wipf, D. Exploring algorithmic limits of matrix rank minimization under affine constraints. *IEEE Trans. Signal Process.* **2016**, *64*, 4960–4974. [CrossRef]
22. Chen, W. Simultaneously sparse and low-rank matrix reconstruction via nonconvex and nonseparable regularization. *IEEE Trans. Signal Process.* **2018**, *66*, 5313–5323. [CrossRef]
23. Li, Y.; Chen, W.; Yin, C.; Han, Z. Approximate message passing for sparse recovering of spatially and temporally correlated data. In Proceedings of the 2015 IEEE/CIC International Conference on Communications in China (ICCC), Shenzhen, China, 2–4 November 2015; pp. 1–6.
24. Ziniel, J.; Schniter, P. Efficient high-dimensional inference in the multiple measurement vector problem. *IEEE Trans. Signal Process.* **2013**, *61*, 340–354. [CrossRef]
25. Zhang, Z.; Jung, T.P.; Makeig, S.; Rao, B.D. Compressed sensing of EEG for wireless telemonitoring with low energy consumption and inexpensive hardware. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 221–224. [CrossRef] [PubMed]