A Framelet Sparse Reconstruction Method for Pansharpening with Guaranteed Convergence

Zhong-Cheng Wu, Ting-Zhu Huang, Liang-Jian Deng [*†], Gemine Vivone [‡]

Abstract

Pansharpening refers to the super resolution of a low-resolution multispectral (LR-MS) image in virtue of an aligned panchromatic (PAN) image. Such an inverse problem mainly requires a proper use of the spatial information from the auxiliary PAN image. In this paper, we suggest a nonconvex regularization model for pansharpening via framelet sparse reconstruction, called NC-FSRM, which investigates the coefficient similarity among the underlying high-resolution MS (HR-MS) and PAN images on the framelet domain, then characterizes the strong statistical sparsity of their error using ℓ_0 norm. Compared with previous methods, NC-FSRM can more precisely and concisely establish the relation between the underlying HR-MS and PAN images. In particular, the piece-wise smoothness prior of the former can simultaneously be captured without adding additional regularizers. For solving the suggested nonconvex model, we further develop an efficient proximal alternating minimization (PAM) based algorithm, which is theoretically proven to converge to the coordinatewise minimizers under some mild assumptions. Numerical experiments conducted on different datasets demonstrate the superiority of the suggested NC-FSRM compared with other state-of-the-art pansharpening methods. The source code is publicly available at *https://github.com/zhongchengwu/code_ncFSRM*.

Keywords Framelet Sparse Reconstruction Method (FSRM); ℓ_0 Norm; Super-resolution; Pansharpening.

1 Introduction

In image processing, super resolution (SR) reconstruction is a classic and widely studied topic, e.g., [49, 51, 61], and particularly, pansharpening can be considered as a special case of SR. Due to inherent limitations on hardware systems, existing imaging devices, such as Pléiades, Gaofen-2 (GF-2), WorldView-4 (WV-4), and WorldView-3 (WV-3), only measure a lower spatial-resolution multispectral (LR-MS) image, while capturing the finer spatial information into a gray-scaled panchromatic (PAN) image alternatively. In general, such a compromise solution imposes limitations on many subsequent applications relied on high-resolution multispectral (HR-MS) images. To tackle this dilemma, pansharpening has emerged as an algorithmic tool for increasing the spatial resolution of the LR-MS image (i.e., $\mathcal{Y} \in \mathbb{R}^{h\times w\times S}$) by referring to the corresponding PAN image (i.e., $\boldsymbol{P} \in \mathbb{R}^{H\times W}$), aiming to generate an underlying HR-MS image (i.e., $\mathcal{X} \in \mathbb{R}^{H\times W\times S}$). Especially, $H = h \times r$, $W = w \times r$ holds, where the integer $r \geq 2$ is the scale factor.

Recently, numerous pansharpening researches divided into different categories have been developed, e.g., the component substitution (CS) based methods [3, 16, 26, 27, 35, 46, 53], the multi-resolution analysis (MRA) based methods [2, 36, 42, 50, 55], the deep learning based methods [29, 31, 39, 60, 62, 65], and the regularization

[*]The research is supported by NSFC (Grant Nos. 12171072, 12271083), Key Projects of Applied Basic Research in Sichuan Province (Grant No. 2020YJ0216), Natural Science Foundation of Sichuan Province (Grant No. 2022NSFSC0501), and National Key Research and Development Program of China (Grant No. 2020YFA0714001). *(Corresponding authors: Ting-Zhu Huang; Liang-Jian Deng.)*

[†]Zhong-Cheng Wu, Ting-Zhu Huang and Liang-Jian Deng are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731, China (e-mails: wuzhch97@163.com; tingzhuhuang@126.com; liangjian.deng@uestc.edu.cn).

[‡]Gemine Vivone is with the Institute of Methodologies for Environmental Analysis (CNR-IMAA), Tito Scalo, 85050, Italy, and with the National Biodiversity Future Center (NBFC), Palermo, 90133, Italy (e-mail: gemine.vivone@imaa.cnr.it).

based methods [7, 9, 20, 25, 37, 56, 58]. Among them, the regularization category has garnered considerable attention because of its capability to settle the ill-posed inverse problems, whose local or global minima correspond to the pansharpened results. Moreover, a more solid mathematical foundation, suitable flexibility, and higher model interpretability are demonstrated. In this paper, the proposed NC-FSRM is one such method.

As a general inverse imaging problem, the primary objective of the regularized pansharpening methods is to model the image degradation from the underlying HR-MS image to the degraded LR-MS image. Similar to the traditional SR issue [28, 38], there also exists an acknowledged and widely-used degradation model for pansharpening, which is given by matrix form as follows,

$$\boldsymbol{Y}_{(3)} = \boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} + \boldsymbol{N}, \tag{1.1}$$

where $\boldsymbol{X}_{(3)} \in \mathbb{R}^{S\times HW}$ and $\boldsymbol{Y}_{(3)} \in \mathbb{R}^{S\times hw}$ are the mode-3 matricization (see Section 2.1) of the HR-MS and the LR-MS images, respectively, $\boldsymbol{B} \in \mathbb{R}^{HW\times HW}$ and $\boldsymbol{S} \in \mathbb{R}^{HW\times hw}$ imply the blurring and down-sampling matrices, respectively, and $\boldsymbol{N} \in \mathbb{R}^{S\times hw}$ indicates an additive Gaussian white noise. Based on degradation model (1.1), the constrained minimization model of pansharpening can typically be formulated by

$$\min_{\mathcal{X}} \ J_1(\mathcal{X}, \boldsymbol{P}) + \lambda J_2(\mathcal{X}), \quad \text{s.t. } \boldsymbol{Y}_{(3)} = \boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} + \boldsymbol{N}, \tag{1.2}$$

where $\boldsymbol{P} \in \mathbb{R}^{H\times W}$ is the known PAN image, J_1 is an energy function, J_2 is a regularizer, and $\lambda > 0$ is a regularization parameter. Fortunately, pansharpening model (1.2) can also be regarded as a dual-objective fusion problem, thus its ill-posedness degrades significantly. This indicates that function J_1 instead of regularizer J_2 actually plays a prominent role for model (1.2), and the latter only provides incremental improvements. Consequently, most existing pansharpening methods, e.g., [37, 56], mainly aim to characterize a more accurate relationship between the underlying HR-MS image and the PAN image, thereby constructing the corresponding energy function J_1 via some assumptions or statistical analysis, etc.

More specifically, Ballester et al. [9] propose a pioneering variational method for pansharpening, named P+XS. This method mainly relies upon the assumption that the PAN image can be represented as a linear combination of the underlying HR-MS image bands. Mathematically, function J_1 of P+XS is given by

$$\int_{\Omega} (\mathcal{X} \times_3 \boldsymbol{v} - \boldsymbol{P})^2 \, d\boldsymbol{l}, \tag{1.3}$$

where $\boldsymbol{l} = (l_1, l_2) \in \Omega$ is the related image domain, $\boldsymbol{v} \geq 0$ is a sum-to-one coefficient vector, and $\times_k$ denotes the tensor k-mode multiplication [33] (see Section 2.1). Afterwards, some methods, e.g., [20], instead relax the sum-to-one constraint and develop a data-dependent least-squares regression to adaptively estimate the coefficient vector, aiming to facilitate its flexibility. However, because the linearization assumption is essentially biased and even unrealistic [25, 52], such methods usually result in local distortion.

To improve the biases, subsequent methods, e.g., [7, 23, 56], further assume that the previous linear relationship can globally be established in the high-frequency domain. Accordingly, (1.3) is modified to

$$\int_{\Omega} \Big(\mathcal{H}(\mathcal{X}) \times_3 \boldsymbol{v} - \mathcal{H}(\boldsymbol{P})\Big)^2 d\boldsymbol{l}, \tag{1.4}$$

where $\mathcal{H}(\cdot)$ is the high-pass operator. Under this framework, various instances of $\mathcal{H}$, such as gradient [24] and Laplacian [56] operators, generally distinguish different methods. Additionally, instead of using ℓ_2 norm in (1.4), Jiang et al. [32] also characterize the relation $\mathcal{H}(\mathcal{X}) \times_3 \boldsymbol{v} - \mathcal{H}(\boldsymbol{P})$ via $\ell_{1/2}$ norm, aiming to satisfy the statistical hyper-Laplacian distribution [34]. Nonetheless, $\mathcal{H}(\mathcal{X}) \times_3 \boldsymbol{v} - \mathcal{H}(\boldsymbol{P})$ is actually equivalent to $\mathcal{H}(\mathcal{X} \times_3 \boldsymbol{v} - \boldsymbol{P})$, and thus, these so-called methods still cannot evade the linearization assumption, leading to inflexibility and limited improvements on different local image patches.

To tackle the issues, many more flexible methods have recently been developed, e.g., [24, 25, 40, 58, 59]. Most of them can be uniformly summarized into band-independent modeling, i.e.,

$$\int_{\Omega} \sum_{i=1}^{S} \omega_i \Big(\mathcal{H}_i(\boldsymbol{X}_{[i]}) - \kappa_i \mathcal{H}_i(\boldsymbol{P}) + \boldsymbol{C}_{[i]}\Big)^2 d\boldsymbol{l}, \tag{1.5}$$

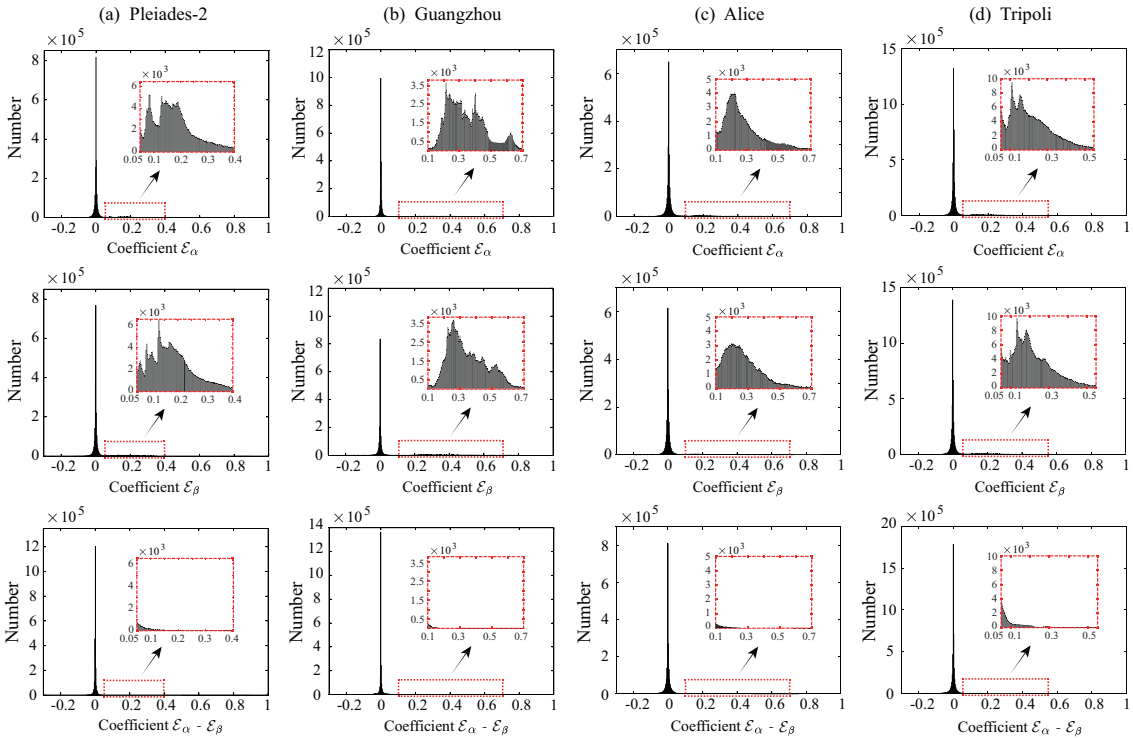

Fig. 1.1: Illustration of sparse coefficient characteristics on the reduced-resolution (a) Pléiades-2 dataset (sensor: Pléiades), (b) Guangzhou dataset (sensor: GF-2), (c) Alice dataset (sensor: WV-4), and (d) Tripoli dataset (sensor: WV-3). Variables $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ represent the framelet coefficients of the simulated HR-MS image and the histogram-matched PAN image, respectively.

where $\{\omega_i\}_{i=1}^S$ and $\{\kappa_i\}_{i=1}^S$ are the regularization parameters and scale coefficients, respectively, $\{\mathcal{H}_i\}_{i=1}^S$ are the projection operations, $\{\boldsymbol{X}_{[i]}\}_{i=1}^S$ are the frontal slices of $\mathcal{X}$, and $\{\boldsymbol{C}_{[i]}\}_{i=1}^S$ are the constant intercepts. Generally, $\{\omega_i\}_{i=1}^S$ are assumed to be ones, and $\{\boldsymbol{C}_{[i]}\}_{i=1}^S$ are supposed to be zeros since $\mathcal{H}_i(\boldsymbol{X}_{[i]})$ and $\mathcal{H}_i(\boldsymbol{P})$ have similar data characteristics. Moreover, $\{\mathcal{H}_i\}_{i=1}^S$ can be a common transformation, such as gradient [25,59] one and wavelet [24,40] one. Comparatively to (1.4), determining $\{\kappa_i\}_{i=1}^S$ is another crucial issue for (1.5) that yields a variety of contributing strategies. For instance, Aly et al. [7] utilize the spectral responsivity to compute $\{\kappa_i\}_{i=1}^S$ in advance; Fu et al. [25] adopt a patch-based strategy to update $\{\kappa_i\}_{i=1}^S$ iteratively; Wu et al. [58] design a pixel-based criterion to obtain $\{\kappa_i\}_{i=1}^S$ dynamically; and Xiao et al. [59] consider adaptively estimating $\{\kappa_i\}_{i=1}^S$ from nonlocal self-similarity perspective.

However, there are still three limitations to most paradigm (1.5) based methods. Firstly, more complex designs for computing $\{\kappa_i\}_{i=1}^S$ only achieve minor performance gains, but instead causing more regularization parameters and even optimization variables, e.g., [7], leading to higher model complexity and ill-posedness. Secondly, the previous methods only independently establish J_1 and J_2 of (1.2), while ignoring that appropriate $\{\mathcal{H}_i\}_{i=1}^S$ and $\{\boldsymbol{C}_{[i]}\}_{i=1}^S$ can also realize J_2 when modeling J_1, e.g., [24], resulting in an increased model redundancy. Thirdly, the high-pass operator is an irreversible transformation, which inevitably causes the underutilization of the PAN image, thus leading to sub-optimal solutions. Therefore, we propose the nonconvex framelet sparse reconstruction method, abbreviated NC-FSRM, aiming to overcome these drawbacks.

As is well known, matter distribution in nature is locally continuous, thus the underlying HR-MS images usually show the piece-wise smoothness property. To characterize it, total variation (TV) regularization has recently been introduced into the pansharpening field see, e.g., [30, 59]. Nevertheless, since the theoretical sparsity-based assumption significantly deviates from reality [43], TV regularizer unavoidably generates

staircase artifacts [21, 41], obscuring the intricacies and geometric features of the recovered images. In contrast, the framelet [13, 19, 45] is an orthogonal basis generation that relaxes the orthogonality and linear independence criteria. Because of the redundancy, it can well-preserve geometric characteristics and details [12, 14, 15, 17]. Given a framelet transform $\mathcal{H}$, the piece-wise smoothness can be reflected by the sparse coefficient $\mathcal{E}_\alpha$, see Fig. 1.1. Afterwards, the piece-wise prior can be formulated by a framelet-based sparse reconstruction, i.e.,

$$\min_{\mathcal{X},\mathcal{E}_\alpha} \|\mathcal{E}_\alpha\|_0, \quad \text{s.t. } \mathcal{H}(\mathcal{X}) = \mathcal{E}_\alpha, \tag{1.6}$$

where $\mathcal{E}_\alpha$ is the framelet coefficients of the underlying HR-MS image, $\|\cdot\|_0$ is the so-called ℓ_0 pseudonorm, which counts the number of nonzero elements of its argument. Certainly, model (1.6) is only formed as the regularizer J_2, and we will deduce the corresponding J_1 from J_2 detailedly. Assume that $\mathcal{H}(\mathcal{P}) = \mathcal{E}_\beta$, in which $\mathcal{P} \in \mathbb{R}^{H\times W\times S}$ implies the histogram-matched version of $\boldsymbol{P}$ with respect to the LR-MS image, then Fig. 1.1 reveals that $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ share a similar distribution over the same framelet transform. Also, $\mathcal{E}_\alpha - \mathcal{E}_\beta$ exhibits a more sparse structure. Let $\mathcal{E} = \mathcal{E}_\alpha - \mathcal{E}_\beta$, we then impose the sparsity condition of $\mathcal{E}$ into (1.6) as follows,

$$\min_{\mathcal{X},\mathcal{E}_\alpha,\mathcal{E}} \|\mathcal{E}_\alpha\|_0 + \lambda\|\mathcal{E}\|_0, \quad \text{s.t. } \mathcal{E}_\alpha - \mathcal{E}_\beta = \mathcal{E}, \mathcal{H}(\mathcal{X}) = \mathcal{E}_\alpha, \tag{1.7}$$

where λ is a regularization parameter. Since $\mathcal{E}_\alpha = \mathcal{E} + \mathcal{E}_\beta$ with $\mathcal{E}_\beta$ being a sparse constant, a sparser $\mathcal{E}$ forces a sparser $\mathcal{E}_\alpha$. That is, $\|\mathcal{E}_\alpha\|_0$ can be incorporated into $\|\mathcal{E}\|_0$. Thus, (1.7) can be simplified as

$$\min_{\mathcal{X},\mathcal{E}} \|\mathcal{E}\|_0, \quad \text{s.t. } \mathcal{H}(\mathcal{X}) - \mathcal{E}_\beta = \mathcal{E}. \tag{1.8}$$

Finally, based on the regularized representation of model (1.8), we can formulate the J_1 function of the proposed NC-FSRM as

$$\int_\Omega \Big(\mathcal{H}(\mathcal{X}) - \mathcal{H}(\mathcal{P}) - \mathcal{E}\Big)^2 d\boldsymbol{l} + \lambda\|\mathcal{E}\|_0. \tag{1.9}$$

Compared with the previous methods, (1.9) mainly includes three advantages. Specifically, (*i*) the scale coefficients $\{\kappa_i\}_{i=1}^S$ that need to be estimated are omitted and removed, reducing the number of parameters; (*ii*) By constraining the sparsity of variable $\mathcal{E}$, both the utilization of the PAN image, i.e., J_1, and the piece-wise smoothness prior of the underlying HR-MS image, i.e., J_2, can simultaneously be implemented without imposing additional regularizers, leading to a simpler form in modeling; (*iii*) The framelet transform is reversible, which more fully utilizes the information of the PAN image, not just its high frequencies.

The remainder of the paper is organized as follows. In Section 2, we briefly introduce the preliminaries. In Section 3, we formulate the final regularization model, i.e., NC-FSRM. In Section 4, we develop a proximal alternating minimization (PAM) based algorithm and prove its convergence theoretically. In Section 5, we perform the numerical experiments to evaluate the performance. Finally, conclusions are drawn in Section 6.

2 Preliminaries

2.1 Notation

In general, scalars, vectors, matrices, and tensors are denoted by lowercase letters, e.g., x, lowercase bold letters, e.g., $\boldsymbol{x}$, uppercase bold letters, e.g., $\boldsymbol{X}$, and calligraphic letters, e.g., $\mathcal{X}$, respectively. For an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, its $(i_1, i_2, \cdots, i_N)$-th element is expressed by $\mathcal{X}(i_1, i_2, \cdots, i_N)$ or, more compactly, $x_{i_1 i_2\cdots i_N}$. The inner product between tensors $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ and $\mathcal{Y} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ is defined as $\langle\mathcal{X}, \mathcal{Y}\rangle := \sum_{i_1,i_2,\cdots,i_N} x_{i_1 i_2\cdots i_N} y_{i_1 i_2\cdots i_N}$. The Frobenius norm of $\mathcal{X}$ is defined as $\|\mathcal{X}\|_F := \sqrt{\langle\mathcal{X}, \mathcal{X}\rangle}$. The ℓ_0 pseudonorm of $\mathcal{X}$ is defined as $\|\mathcal{X}\|_0 := \sum_{i_1,i_2,\cdots,i_N} 1_{x_{i_1 i_2\cdots i_N}\neq 0}$, where $1_{x_{i_1 i_2\cdots i_N}\neq 0}$ denotes a condition function, i.e., $1_{x_{i_1 i_2\cdots i_N}\neq 0} = 1$ if $x_{i_1 i_2\cdots i_N} \neq 0$, and $1_{x_{i_1 i_2\cdots i_N}\neq 0} = 0$ otherwise. The mode-k matricization of $\mathcal{X}$ is defined as $\boldsymbol{X}_{(k)} \in \mathbb{R}^{I_k\times\prod_{j\neq k} I_j}$, whose elements satisfy $\boldsymbol{X}_{(k)}(i_k, \overline{i_1\cdots i_{k-1}i_{k+1}\cdots i_N}) = \mathcal{X}(i_1, i_2, \cdots, i_N)$ with index $\overline{i_1\cdots i_{k-1}i_{k+1}\cdots i_N} = 1 + \sum_{n=1,n\neq k}^N (i_n - 1)\prod_{j=1,j\neq k}^{n-1} I_j$. Given a row

vector $\boldsymbol{v} \in \mathbb{R}^{1\times I_k}$, then the k-mode multiplication between $\mathcal{X}$ and the vector $\boldsymbol{v}$ yields a tensor $\mathcal{X} \times_k \boldsymbol{v} \in \mathbb{R}^{I_1\times\cdots\times I_{k-1}\times 1\times I_{k+1}\times\cdots\times I_N}$, which is computed by $\boldsymbol{v}\boldsymbol{X}_{(k)}$. Especially, when $N = 3$, we denote the i_3-th frontal slice of tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times I_3}$ by $\boldsymbol{X}_{[i_3]} \in \mathbb{R}^{I_1\times I_2}$, which requires $\boldsymbol{X}_{[i_3]}(i_1, i_2) = \mathcal{X}(i_1, i_2, i_3)$.

2.2 Framelet

A countable function subset $X \subset L^2(\mathbb{R})$ is called a tight frame of $L^2(\mathbb{R})$, if

$$f = \sum_{h\in X} \langle f, h\rangle h, \quad \forall f \in L^2(\mathbb{R}),$$

which is equivalent to

$$\|f\|^2 = \sum_{h\in X} |\langle f, h\rangle|^2, \quad \forall f \in L^2(\mathbb{R}),$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product for function variables of $L^2(\mathbb{R})$, and $\|\cdot\| = \langle\cdot,\cdot\rangle^{1/2}$. Given a finite set $\Psi := \{\psi_1, \psi_2, \cdots, \psi_r\} \subset L^2(\mathbb{R})$, the collection of its dilations and shifts is defined as an affine (or wavelet) system $X(\Psi)$, i.e.,

$$X(\Psi) := \{\psi_{l,t,k} : 1 \le l \le r; l, t, k \in \mathbb{Z}\} \quad \text{with} \quad \psi_{l,t,k} := 2^{t/2}\psi_l(2^t \cdot -k).$$

When $X(\Psi)$ forms a tight frame, it is called a tight wavelet frame, and $\{\psi_l\}_{l=1}^r$ are termed (tight) framelets.

Starting from a compactly supported refinable function $\phi \in L^2(\mathbb{R})$ with a refinement mask (low-pass filter) $\zeta_\phi \in L^2(\mathbb{Z})$ satisfying $\phi = \sum_j \zeta_\phi(j)\phi(2\cdot -j)$, a compactly supported framelet system can be constructed by designing an appropriate set of framelets $\Psi := \{\psi_1, \psi_2, \cdots, \psi_r\} \subset L^2(\mathbb{R})$ or, more essentially, framelet masks (high-pass filters) $\{\zeta_{\psi_1}, \zeta_{\psi_2}, \cdots, \zeta_{\psi_r}\} \subset L^2(\mathbb{Z})$, because $\psi_l := \sum_j \zeta_{\psi_l}(j)\phi(2\cdot -j),\ l = 1, 2, \cdots, r$. According to the unitary extension principle (UEP) [48], the wavelet system $X(\Psi)$ can form a tight frame in $L^2(\mathbb{R})$, if the masks ζ_ϕ and $\{\zeta_{\psi_1}, \zeta_{\psi_2}, \cdots, \zeta_{\psi_r}\}$ for almost all $\omega \in \mathbb{R}$ satisfy

$$\xi_\phi(\omega)\overline{\xi_\phi(\omega + \gamma\pi)} + \sum_{l=1}^{r} \xi_{\psi_l}(\omega)\overline{\xi_{\psi_l}(\omega + \gamma\pi)} = \delta(\gamma), \quad \gamma = 0, 1,$$

where $\xi_\phi(\omega) = \sum_j \zeta_\phi(j)e^{-ij\omega}$, $\xi_{\psi_l}(\omega) = \sum_j \zeta_{\psi_l}(j)e^{-ij\omega}$, $l = 1, 2, \cdots, r$, and $\delta(\gamma)$ is a delta function. In our numerical experiments, we adopt the piece-wise linear B-spline framelet [48], in which the piece-wise linear B-spline function is used as ϕ. Under this case, the refinement mask $\zeta_\phi(\omega) = \cos^2(\omega/2)$ and two deduced framelet masks $\zeta_{\psi_1}(\omega) = -(\sqrt{2}i\sin(\omega))/2$, $\zeta_{\psi_2} = \sin^2(\omega/2)$ correspond to the filters as follows,

$$\boldsymbol{\zeta}_\phi = \Big[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\Big], \quad \boldsymbol{\zeta}_{\psi_1} = \Big[\frac{\sqrt{2}}{4}, 0, -\frac{\sqrt{2}}{4}\Big], \quad \boldsymbol{\zeta}_{\psi_2} = \Big[-\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}\Big].$$

For a given discrete image $\boldsymbol{f} \in \mathbb{R}^N$, a framelet system H in $\mathbb{R}^N$ can be regarded as the transform (decomposition) operator to obtain the coefficients of the finite-dimensional vector. The numerical computation can directly be realized by an analysis matrix $\boldsymbol{H} \in \mathbb{R}^{K\times N}(K \ge N)$, whose row vectors form the system H. Remarkably, such a matrix has orthogonal columns, i.e., $\boldsymbol{H}^\top\boldsymbol{H} = \boldsymbol{I}$, while $\boldsymbol{H}\boldsymbol{H}^\top \ne \boldsymbol{I}$ in general.

In what follows, we construct $\boldsymbol{H}$ from the filters related to the multi-level tight frame system, which corresponds to the commonly-used tight framelet decomposition without down-sampling. Let $\zeta = \{\zeta(j)\}_{j=-J}^J$ be a filter, the convolution operator $\boldsymbol{T}(\zeta)$ with filter ζ under the Neumann (symmetric) boundary condition is a Toeplitz-plus-Hankel matrix [13, 48], i.e.,

$$\boldsymbol{T}(\zeta) = \begin{bmatrix} \zeta(0) & \cdots & \zeta(-J) & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \zeta(J) & \ddots & \ddots & \ddots & \zeta(-J) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \zeta(J) & \cdots & \zeta(0) \end{bmatrix} + \begin{bmatrix} \zeta(1) & \zeta(2) & \cdots & \zeta(J) & 0 \\ \zeta(2) & \ddots & \ddots & \ddots & \zeta(-J) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \zeta(J) & \ddots & \ddots & \ddots & \zeta(-2) \\ 0 & \zeta(-J) & \cdots & \zeta(-2) & \zeta(-1) \end{bmatrix}.$$

Particularly, the filters $\zeta^{(l)}$ at level l corresponding to the decomposition without down-sampling is

$$\zeta^{(l)} = \{\zeta(-J), \underbrace{0, \cdots, 0}_{2^{l-1}-1}, \zeta(-J+1), 0, \cdots, \cdots, 0, \zeta(0), 0, \cdots, \cdots, 0, \zeta(J-1), \underbrace{0, \cdots, 0}_{2^{l-1}-1}, \zeta(J)\}.$$

Recalling the low-pass filter ζ_ϕ and high-pass filters $\{\zeta_{\psi_1}, \zeta_{\psi_2}, \cdots, \zeta_{\psi_r}\}$, also denoting $\boldsymbol{Z}_\phi^{(l)} \equiv \boldsymbol{T}(\zeta_\phi^{(l)})$ and $\boldsymbol{Z}_{\psi_i}^{(l)} \equiv \boldsymbol{T}(\zeta_{\psi_i}^{(l)})$ for $i = 1, 2, \cdots, r$, then the multi-level decomposition matrix $\boldsymbol{H}$ with level L is given by

$$\boldsymbol{H} = \begin{bmatrix} \prod_{l=0}^{L-1} \boldsymbol{Z}_\phi^{(L-l)} \\ \boldsymbol{Z}_{\psi_1}^{(L)} \prod_{l=1}^{L-1} \boldsymbol{Z}_\phi^{(L-l)} \\ \vdots \\ \boldsymbol{Z}_{\psi_r}^{(L)} \prod_{l=1}^{L-1} \boldsymbol{Z}_\phi^{(L-l)} \\ \vdots \\ \boldsymbol{Z}_{\psi_1}^{(1)} \\ \vdots \\ \boldsymbol{Z}_{\psi_r}^{(1)} \end{bmatrix}.$$

Afterwards, the framelet decomposition for a third-order image $\mathcal{X} \in \mathbb{R}^{H\times W\times S}$ can concisely be achieved by $\boldsymbol{H}\boldsymbol{X}_{(3)}^\top$. Conversely, the framelet reconstruction $\boldsymbol{H}^\top$ allows $\boldsymbol{X}_{(3)}^\top = \boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{X}_{(3)}^\top$ because of $\boldsymbol{H}^\top\boldsymbol{H} = \boldsymbol{I}$. In our implementations, the decomposition level is set as 1, i.e., $L = 1$, which can experimentally achieve satisfactory numerical performance while preserving a reduced computational burden.

3 The proposed model

According to Section 2.2, we firstly modify model (1.9) as follows,

$$\int_\Omega \left(\boldsymbol{H}\boldsymbol{X}_{(3)}^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}\right)^2 d\boldsymbol{l} + \lambda\|\boldsymbol{E}\|_0. \tag{3.10}$$

Following (1.2) with (3.10), the final constrained model, i.e., NC-FSRM, can formally be formulated by

$$\min_{\boldsymbol{X}_{(3)},\boldsymbol{E}} \int_\Omega \left(\boldsymbol{H}\boldsymbol{X}_{(3)}^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}\right)^2 d\boldsymbol{l} + \lambda\|\boldsymbol{E}\|_0, \quad \text{s.t. } \boldsymbol{Y}_{(3)} = \boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} + \boldsymbol{N}. \tag{3.11}$$

To eliminate the Gaussian noise $\boldsymbol{N}$, (3.11) can also be transformed into an unconstrained model, i.e.,

$$\min_{\boldsymbol{X}_{(3)},\boldsymbol{E}} \frac{1}{2}\int_\Omega \left(\boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\right)^2 d\boldsymbol{l} + \lambda_1 \int_\Omega \left(\boldsymbol{H}\boldsymbol{X}_{(3)}^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}\right)^2 d\boldsymbol{l} + \lambda_2 \|\boldsymbol{E}\|_0, \tag{3.12}$$

where $\lambda_i > 0$, $i = 1, 2$, are two regularization parameters. Regarding the nonconvex ℓ_0 norm, there are many available optimization algorithms, e.g., the mean doubly augmented Lagrangian (MDAL) algorithm [22], the proximal iterative hard-thresholding (PIHT) algorithm [66], and the Moreau envelope augmented Lagrangian (MEAL) algorithm [64]. Interestingly, we develop a proximal alternating minimization (PAM) [11] based solving algorithm, then prove its convergence theoretically.

4 The proposed algorithm

This section consists in applying the PAM framework [11] to the proposed nonsmooth and nonconvex model. Before that, model (3.12) is firstly rewritten as follows,

$$\min_{\boldsymbol{X}_{(3)},\boldsymbol{E}} \Phi(\boldsymbol{X}_{(3)}, \boldsymbol{E}) := h(\boldsymbol{X}_{(3)}, \boldsymbol{E}) + f(\boldsymbol{X}_{(3)}) + g(\boldsymbol{E}), \tag{4.13}$$

where $h(\boldsymbol{X}_{(3)}, \boldsymbol{E}) = \lambda_1\|\boldsymbol{H}\boldsymbol{X}_{(3)}^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}\|_F^2$, $f(\boldsymbol{X}_{(3)}) = 1/2\|\boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\|_F^2$, and $g(\boldsymbol{E}) = \lambda_2\|\boldsymbol{E}\|_0$. Definitely, h, $f \in C^1$ are two continuous functions with locally Lipschitz continuous gradient in the feasible domain, and g can be verified to be proper lower semi-continuous, thus $\Phi : \mathbb{R}^{S\times HW} \times \mathbb{R}^{3H3W\times S} \to \mathbb{R}\cup\{+\infty\}$ in (4.13) is a proper lower semi-continuous function. Under the PAM framework, variables $\boldsymbol{X}_{(3)}$ and $\boldsymbol{E}$ can be alternately iterated as follows,

$$\begin{cases} \boldsymbol{X}_{(3)}^{k+1} \in \arg\min\limits_{\boldsymbol{X}_{(3)}} \left\{ h(\boldsymbol{X}_{(3)}, \boldsymbol{E}^k) + f(\boldsymbol{X}_{(3)}) + \dfrac{\rho}{2}\|\boldsymbol{X}_{(3)} - \boldsymbol{X}_{(3)}^k\|_F^2 \right\}, \\ \boldsymbol{E}^{k+1} \in \arg\min\limits_{\boldsymbol{E}} \left\{ h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}) + g(\boldsymbol{E}) + \dfrac{\sigma}{2}\|\boldsymbol{E} - \boldsymbol{E}^k\|_F^2 \right\}, \end{cases} \tag{4.14}$$

where ρ, $\sigma > 0$ are two proximal parameters.

4.1 The solution of $\boldsymbol{X}_{(3)}$-subproblem

Relying upon (4.13) and (4.14), we obtain concretely the minimization form of the $\boldsymbol{X}_{(3)}$-subproblem, i.e.,

$$\min_{\boldsymbol{X}_{(3)}} \frac{1}{2}\|\boldsymbol{X}_{(3)}\boldsymbol{B}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\|_F^2 + \lambda_1\|\boldsymbol{H}\boldsymbol{X}_{(3)}^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}^k\|_F^2 + \frac{\rho}{2}\|\boldsymbol{X}_{(3)} - \boldsymbol{X}_{(3)}^k\|_F^2. \tag{4.15}$$

Although the objection function in (4.15) is theoretically differentiable, forcing its derivative to be zero will cause an ensuing difficulty resulting from the coupling of blurring and down-sampling matrices. Here, we further introduce an ADMM-based algorithm. By making variable substitution, we have the equivalent constrained problem as follows,

$$\begin{aligned} \min_{\boldsymbol{X}_{(3)},\boldsymbol{U},\boldsymbol{V}} \ & \frac{1}{2}\|\boldsymbol{U}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\|_F^2 + \lambda_1\|\boldsymbol{H}\boldsymbol{V} - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}^k\|_F^2 + \frac{\rho}{2}\|\boldsymbol{X}_{(3)} - \boldsymbol{X}_{(3)}^k\|_F^2 \\ & \text{s.t. } \boldsymbol{U} = \boldsymbol{X}_{(3)}\boldsymbol{B},\ \boldsymbol{V} = \boldsymbol{X}_{(3)}^\top. \end{aligned} \tag{4.16}$$

The augmented Lagrangian function of the constrained model (4.16) is formulated by

$$\begin{aligned} \mathcal{L}(\boldsymbol{X}_{(3)}, \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{\Lambda}, \boldsymbol{\Theta}) = & \frac{1}{2}\|\boldsymbol{U}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\|_F^2 + \lambda_1\|\boldsymbol{H}\boldsymbol{V} - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}^k\|_F^2 + \frac{\rho}{2}\|\boldsymbol{X}_{(3)} - \boldsymbol{X}_{(3)}^k\|_F^2 \\ & + \langle \boldsymbol{\Lambda}, \boldsymbol{X}_{(3)}\boldsymbol{B} - \boldsymbol{U}\rangle + \frac{\eta_1}{2}\|\boldsymbol{X}_{(3)}\boldsymbol{B} - \boldsymbol{U}\|_F^2 + \langle \boldsymbol{\Theta}, \boldsymbol{X}_{(3)}^\top - \boldsymbol{V}\rangle + \frac{\eta_2}{2}\|\boldsymbol{X}_{(3)}^\top - \boldsymbol{V}\|_F^2, \end{aligned} \tag{4.17}$$

where $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$ are two Lagrange multipliers, and $\eta_i > 0$, $i = 1, 2$, are the penalty parameters. Afterwards, (4.17) can be solved alternatively and iteratively via the following procedure,

$$\begin{cases} \boldsymbol{X}_{(3)}^{k+1,p+1} \in \arg\min\limits_{\boldsymbol{X}_{(3)}} \mathcal{L}(\boldsymbol{X}_{(3)}, \boldsymbol{U}^p, \boldsymbol{V}^p, \boldsymbol{\Lambda}^p, \boldsymbol{\Theta}^p) \\ \qquad = \arg\min\limits_{\boldsymbol{X}_{(3)}} \left\{ \dfrac{\rho}{2}\|\boldsymbol{X}_{(3)} - \boldsymbol{X}_{(3)}^k\|_F^2 + \dfrac{\eta_1}{2}\|\boldsymbol{X}_{(3)}\boldsymbol{B} - \boldsymbol{U}^p + \dfrac{\boldsymbol{\Lambda}^p}{\eta_1}\|_F^2 + \dfrac{\eta_2}{2}\|\boldsymbol{X}_{(3)}^\top - \boldsymbol{V}^p + \dfrac{\boldsymbol{\Theta}^p}{\eta_2}\|_F^2 \right\}, \\ \boldsymbol{U}^{p+1} \in \arg\min\limits_{\boldsymbol{U}} \mathcal{L}(\boldsymbol{X}_{(3)}^{k+1,p+1}, \boldsymbol{U}, \boldsymbol{V}^p, \boldsymbol{\Lambda}^p, \boldsymbol{\Theta}^p) \\ \qquad = \arg\min\limits_{\boldsymbol{U}} \left\{ \dfrac{1}{2}\|\boldsymbol{U}\boldsymbol{S} - \boldsymbol{Y}_{(3)}\|_F^2 + \dfrac{\eta_1}{2}\|\boldsymbol{X}_{(3)}^{k+1,p+1}\boldsymbol{B} - \boldsymbol{U} + \dfrac{\boldsymbol{\Lambda}^p}{\eta_1}\|_F^2 \right\}, \\ \boldsymbol{V}^{p+1} \in \arg\min\limits_{\boldsymbol{V}} \mathcal{L}(\boldsymbol{X}_{(3)}^{k+1,p+1}, \boldsymbol{U}^{p+1}, \boldsymbol{V}, \boldsymbol{\Lambda}^p, \boldsymbol{\Theta}^p) \\ \qquad = \arg\min\limits_{\boldsymbol{V}} \left\{ \lambda_1\|\boldsymbol{H}\boldsymbol{V} - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}^k\|_F^2 + \dfrac{\eta_2}{2}\|(\boldsymbol{X}_{(3)}^{k+1,p+1})^\top - \boldsymbol{V} + \dfrac{\boldsymbol{\Theta}^p}{\eta_2}\|_F^2 \right\}, \\ \boldsymbol{\Lambda}^{p+1} = \boldsymbol{\Lambda}^p + \eta_1(\boldsymbol{X}_{(3)}^{k+1,p+1}\boldsymbol{B} - \boldsymbol{U}^{p+1}), \quad \boldsymbol{\Theta}^{p+1} = \boldsymbol{\Theta}^p + \eta_2\left((\boldsymbol{X}_{(3)}^{k+1,p+1})^\top - \boldsymbol{V}^{p+1}\right). \end{cases} \tag{4.18}$$

Algorithm 1 The inner ADMM-based algorithm.

Input: LR-MS image $\boldsymbol{Y}$, PAN image $\boldsymbol{P}$, $\boldsymbol{E}^k$, $\boldsymbol{X}_{(3)}^k$, λ_1, ρ, η_1, η_2, and p_{max}.

Initialization: $\boldsymbol{U}^0 \leftarrow \boldsymbol{U}^k$, $\boldsymbol{V}^0 \leftarrow \boldsymbol{V}^k$, $\boldsymbol{\Lambda}^0 \leftarrow \boldsymbol{\Lambda}^k$, and $\boldsymbol{\Theta}^0 \leftarrow \boldsymbol{\Theta}^k$.

1: **for** $p = 0$ to p_{max} **do**
2: Update $\boldsymbol{X}_{(3)}^{k+1,p+1}$ by (4.19);
3: Update $\boldsymbol{U}^{p+1}$ by (4.21);
4: Update $\boldsymbol{V}^{p+1}$ by (4.23);
5: $\boldsymbol{\Lambda}^{p+1} = \boldsymbol{\Lambda}^p + \eta_1(\boldsymbol{X}_{(3)}^{k+1,p+1}\boldsymbol{B} - \boldsymbol{U}^{p+1})$;
6: $\boldsymbol{\Theta}^{p+1} = \boldsymbol{\Theta}^p + \eta_2\left((\boldsymbol{X}_{(3)}^{k+1,p+1})^\top - \boldsymbol{V}^{p+1}\right)$.
7: **end for**

Output: The $(k+1)$-th values $\boldsymbol{X}_{(3)}^{k+1} \leftarrow \boldsymbol{X}_{(3)}^{k+1,p_{max}+1}$, $\boldsymbol{U}^{k+1} \leftarrow \boldsymbol{U}^{k+1,p_{max}+1}$, $\boldsymbol{V}^{k+1} \leftarrow \boldsymbol{V}^{k+1,p_{max}+1}$, $\boldsymbol{\Lambda}^{k+1} \leftarrow \boldsymbol{\Lambda}^{k+1,p_{max}+1}$, and $\boldsymbol{\Theta}^{k+1} \leftarrow \boldsymbol{\Theta}^{k+1,p_{max}+1}$.

Updating $\boldsymbol{X}_{(3)}$: This deconvolution problem has a closed-form solution, which can be given by the fast Fourier transform (FFT) based on the condition of a periodic boundary, as follows,

$$\boldsymbol{X}_{(3)}^{k+1,p+1} = \mathcal{F}^{-1}\left(\frac{\rho\mathcal{F}(\boldsymbol{X}_{(3)}^k) + (\eta_1\mathcal{F}(\boldsymbol{U}^p) - \mathcal{F}(\boldsymbol{\Lambda}^p)) \circ \mathcal{F}(\boldsymbol{B})^\ddagger + \eta_2\mathcal{F}((\boldsymbol{V}^p)^\top) - \mathcal{F}((\boldsymbol{\Theta}^p)^\top)}{\eta_1\mathcal{F}(\boldsymbol{B}) \circ \mathcal{F}(\boldsymbol{B})^\ddagger + (\rho + \eta_2)\boldsymbol{1}}\right), \tag{4.19}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ indicate the FFT and its inverse operators, respectively, $\ddagger$ denotes the complex conjugate, $\circ$ is the Hadamard product, and component-wise division is performed. According to [58], a unique blurring matrix can be generated for the i-th frontal slice $\mathbf{X}_{[i]}$, $i = 1, 2, \cdots, S$, by the Gaussian-shaped kernel $h_i \in \mathbb{R}^{41\times 41}$ with a corresponding standard deviation, whose estimated value is directly advisable, see [53, 54].

Updating $\boldsymbol{U}$: Forcing the derivative of $\mathcal{L}(\boldsymbol{X}_{(3)}^{k+1,p+1}, \boldsymbol{U}, \boldsymbol{V}^p, \boldsymbol{\Lambda}^p, \boldsymbol{\Theta}^p)$ with respect to $\boldsymbol{U}$ to be zero, we easily have

$$\boldsymbol{U}\boldsymbol{S}\boldsymbol{S}^\top + \eta_1\boldsymbol{U} = \boldsymbol{Y}_{(3)}\boldsymbol{S}^\top + \eta_1\boldsymbol{X}_{(3)}^{k+1,p+1}\boldsymbol{B} + \boldsymbol{\Lambda}^p, \tag{4.20}$$

where $\boldsymbol{S}\boldsymbol{S}^\top \in \mathbb{R}^{HW\times HW}$ is a diagonal matrix whose entries are ones at sampled positions. Consequently,

$$\boldsymbol{U}^{p+1} = \frac{\boldsymbol{Y}_{(3)}\boldsymbol{S}^\top + \eta_1\boldsymbol{X}_{(3)}^{k+1,p+1}\boldsymbol{B} + \boldsymbol{\Lambda}^p}{\boldsymbol{S}\boldsymbol{S}^\top + \eta_1\boldsymbol{I}}. \tag{4.21}$$

Updating $\boldsymbol{V}$: Similarly, we can obtain the sub-gradient equation as follows,

$$2\lambda_1\boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{V} + \eta_2\boldsymbol{V} = 2\lambda_1\boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{P}_{(3)}^\top + 2\lambda_1\boldsymbol{H}^\top\boldsymbol{E}^k + \eta_2(\boldsymbol{X}_{(3)}^{k+1,p+1})^\top + \boldsymbol{\Theta}^p, \tag{4.22}$$

whose solution is given by

$$\boldsymbol{V}^{p+1} = \frac{2\lambda_1(\boldsymbol{P}_{(3)}^\top + \boldsymbol{H}^\top\boldsymbol{E}^k) + \eta_2(\boldsymbol{X}_{(3)}^{k+1,p+1})^\top + \boldsymbol{\Theta}^p}{2\lambda_1 + \eta_2}. \tag{4.23}$$

The numerical procedure of the whole ADMM-based algorithm is detailed in Algorithm 1. Especially, a warm-start strategy is adopted, i.e., all the intermediate variables in the ADMM algorithm are reused as initial values in the next iteration, aiming to accelerate the algorithm convergence.

4.2 The solution of $\boldsymbol{E}$-subproblem

Following (4.13) and (4.14), the minimization problem of $\boldsymbol{E}$-subproblem can be presented as follows,

$$\min_{\boldsymbol{E}} \ \lambda_1\|\boldsymbol{H}(\boldsymbol{X}_{(3)}^{k+1})^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top - \boldsymbol{E}\|_F^2 + \lambda_2\|\boldsymbol{E}\|_0 + \frac{\sigma}{2}\|\boldsymbol{E} - \boldsymbol{E}^k\|_F^2. \tag{4.24}$$

Algorithm 2 The PAM-based solver for the proposed NC-FSRM.

Input: LR-MS image $\boldsymbol{Y}$, PAN image $\boldsymbol{P}$, λ_1, λ_2, η_1, η_2, ρ, σ, r, ε, k_{max}, p_{max}.
Initialization: $k = 0$, $\boldsymbol{X}_{(3)}^0 \leftarrow \mathbf{0}$, $\boldsymbol{E}^0 \leftarrow \mathbf{0}$, $\boldsymbol{U}^0 \leftarrow \mathbf{0}$, $\boldsymbol{V}^0 \leftarrow \mathbf{0}$, $\boldsymbol{\Lambda}^0 \leftarrow \mathbf{0}$, $\boldsymbol{\Theta}^0 \leftarrow \mathbf{0}$.
1: **while** *not converged* and $k < k_{max}$ **do**
2: Update $\boldsymbol{X}_{(3)}^{k+1}$ via Algorithm 1;
3: Update $\boldsymbol{E}^{k+1}$ via (4.26);
4: Check the convergence criterion: $\|\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k}\|_F / \|\boldsymbol{X}_{(3)}^{k+1}\|_F < \varepsilon$;
5: $k \leftarrow k + 1$.
6: **end while**
Output: The pansharpened HR-MS image $\boldsymbol{X}_{(3)}$.

The above problem is a sparse coding problem, which can be concisely rewritten as

$$\min_{\boldsymbol{E}} \left\| \frac{2\lambda_1 \boldsymbol{H}(\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{P}_{(3)})^\top + \sigma \boldsymbol{E}^k}{2\lambda_1 + \sigma} - \boldsymbol{E} \right\|_F^2 + \epsilon^2 \|\boldsymbol{E}\|_0 \quad \text{with} \quad \epsilon^2 = \frac{2\lambda_2}{2\lambda_1 + \sigma}. \tag{4.25}$$

The closed-form solution of (4.25) can be obtained by

$$\boldsymbol{E}^{k+1} = \text{Hard}\left(\frac{2\lambda_1 \boldsymbol{H}(\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{P}_{(3)})^\top + \sigma \boldsymbol{E}^k}{2\lambda_1 + \sigma}, \, \epsilon \right), \tag{4.26}$$

where $\text{Hard}(\cdot, \cdot)$ implies a hard thresholding operator defined by $(\text{Hard}(\boldsymbol{E}, \vartheta))(i_1, i_2) := e_{i_1 i_2} 1_{|e_{i_1 i_2}| > \vartheta}$.

The proposed PAM-based algorithm for solving model (3.12) is summarized in Algorithm 2.

4.3 The Convergence Analysis

Before performing the proof of theoretical convergence, we firstly provide some basic definitions of the semi-algebraic function and Kurdyka–Łojasiewicz (KŁ) property. For a proper and lower semicontinuous function $\Psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, the domain of Ψ is defined as

$$\text{dom}(\Psi) := \left\{ \boldsymbol{u} \in \mathbb{R}^n : \Psi(\boldsymbol{u}) < +\infty \right\}.$$

For any subset $S \subset \mathbb{R}^n$ and any point $\boldsymbol{u} \in \mathbb{R}^n$, the distance from $\boldsymbol{u}$ to S is given by

$$\text{dist}(\boldsymbol{u}, S) := \inf \{ \|\boldsymbol{v} - \boldsymbol{u}\| : \boldsymbol{v} \in S \}.$$

Definition 4.1 (Semi-algebraic sets and functions [10, 44]). *(a) A subset S of $\mathbb{R}^n$ is a real semi-algebraic set if there exists a finite number of real polynomial functions f_{ij}, $h_{ij} : \mathbb{R}^n \to \mathbb{R}$ such that*

$$S = \bigcup_{j=1}^{p} \bigcap_{i=1}^{q} \left\{ \boldsymbol{u} \in \mathbb{R}^n : f_{ij}(\boldsymbol{u}) = 0 \text{ and } h_{ij}(\boldsymbol{u}) < 0 \right\}.$$

(b) A function $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called semi-algebraic if its graph

$$\left\{ (\boldsymbol{u}, z) \in \mathbb{R}^{n+1} : g(\boldsymbol{u}) = z \right\}$$

is a semi-algebraic subset of $\mathbb{R}^{n+1}$.

Definition 4.2 (KŁ property [8]). *Let function $\Psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous. (a) This function is provided with the KŁ property at $\boldsymbol{u}^* \in dom(\partial \Psi)$ if there exists $\eta \in (0, +\infty]$, a neighborhood U of $\boldsymbol{u}^*$ and a continuous concave function $\varphi : [0, \eta) \to \mathbb{R}_+$ such that*

i) $\varphi(0) = 0$*;*
ii) φ *is* C^1 *on* $(0, \eta)$*;*
iii) for all $s \in (0, \eta)$*,* $\varphi'(s) > 0$*;*
iv) for all $\boldsymbol{u} \in U \cap \{\boldsymbol{u} : \Psi(\boldsymbol{u}^*) < \Psi(\boldsymbol{u}) < \Psi(\boldsymbol{u}^*) + \eta\}$*, the following KŁ inequality holds*

$$\varphi'\big(\Psi(\boldsymbol{u}) - \Psi(\boldsymbol{u}^*)\big)\mathrm{dist}\big(0,\ \partial\Psi(\boldsymbol{u})\big) \geq 1.$$

(b) When Ψ *satisfies the KŁ inequality at each point of* $\mathrm{dom}(\partial\Psi)$*, it is called KŁ function.*

Lemma 1. *The objective function* Φ *in* (4.13) *is a KŁ function.*

Proof. Following [8, 10], since h is a polynomial function of two coupling variations, it is a real analytic function. Also, f is obviously a real-analytic polynomial function. Moreover, ℓ_0 norm, i.e., $\|\cdot\|_0$, is semi-algebraic. Thus, Φ, as a finite sum of real-valued analytic and semi-algebraic functions, is a KŁ function. □

Lemma 2 (Sufficient decrease condition)**.** *Let* $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ *be the sequence generated by the proposed Algorithm 2, where* $\boldsymbol{Z}^k := (\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k)$*. Then, the sequence* $\{\Phi(\boldsymbol{Z}^k)\}_{k\in\mathbb{N}}$ *satisfies the nonincreasing inequality,*

$$\Phi(\boldsymbol{Z}^k) - \Phi(\boldsymbol{Z}^{k+1}) \geq \min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)\|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^k\|_F^2,$$

where ρ*,* $\sigma > 0$ *are two proximal parameters.*

Proof. The sequence $\{\boldsymbol{X}^k_{(3)}\}_{k\in\mathbb{N}}$ is generated by Algorithm 1, which provides an iteration-based solution of the minimization problem (4.15). When Algorithm 1 converges globally and sufficiently, then $\boldsymbol{X}^{k+1}_{(3)}$ can be a minimizer in the $(k+1)$-th iteration, thus leading to inequality as follows,

$$h(\boldsymbol{X}^{k+1}_{(3)}, \boldsymbol{E}^k) + f(\boldsymbol{X}^{k+1}_{(3)}) + \frac{\rho}{2}\|\boldsymbol{X}^{k+1}_{(3)} - \boldsymbol{X}^k_{(3)}\|_F^2 \leq h(\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k) + f(\boldsymbol{X}^k_{(3)}). \tag{4.27}$$

Likewisely, the sequence $\{\boldsymbol{E}^k\}_{k\in\mathbb{N}}$ generated by (4.26) is also formed by minimizers, thus $\boldsymbol{E}^{k+1}$ can satisfy

$$h(\boldsymbol{X}^{k+1}_{(3)}, \boldsymbol{E}^{k+1}) + g(\boldsymbol{E}^{k+1}) + \frac{\sigma}{2}\|\boldsymbol{E}^{k+1} - \boldsymbol{E}^k\|_F^2 \leq h(\boldsymbol{X}^{k+1}_{(3)}, \boldsymbol{E}^k) + g(\boldsymbol{E}^k). \tag{4.28}$$

By combining inequalities (4.27)-(4.28), we easily deduce

$$\begin{aligned}
\Phi(\boldsymbol{Z}^k) - \Phi(\boldsymbol{Z}^{k+1}) &= h(\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k) + f(\boldsymbol{X}^k_{(3)}) + g(\boldsymbol{E}^k) - h(\boldsymbol{X}^{k+1}_{(3)}, \boldsymbol{E}^{k+1}) - f(\boldsymbol{X}^{k+1}_{(3)}) - g(\boldsymbol{E}^{k+1}) \\
&\geq \frac{\rho}{2}\|\boldsymbol{X}^{k+1}_{(3)} - \boldsymbol{X}^k_{(3)}\|_F^2 + \frac{\sigma}{2}\|\boldsymbol{E}^{k+1} - \boldsymbol{E}^k\|_F^2 \\
&\geq \min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)\left(\|\boldsymbol{X}^{k+1}_{(3)} - \boldsymbol{X}^k_{(3)}\|_F^2 + \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^k\|_F^2\right) \\
&= \min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)\|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^k\|_F^2.
\end{aligned}$$

The proof is completed. □

Lemma 3 (Relative error condition)**.** *Let* $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ *be the sequence generated by the proposed Algorithm 2, where* $\boldsymbol{Z}^k := (\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k)$*. Then, there exists for each* $k \in \mathbb{N}$*,*

$$\|\partial\Phi(\boldsymbol{Z}^{k+1})\|_F \leq (L_h + \sigma + \rho)\|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^k\|_F,$$

where L_h *is the Lipschitz constant of the partial derivative* $\partial_{\boldsymbol{X}_{(3)}} h$*.*

Proof. According to the solving procedure in (4.14), when Algorithm 1 converges globally and sufficiently, variables $\boldsymbol{X}_{(3)}^{k+1}$ and $\boldsymbol{E}^{k+1}$ are the minimum solutions. Suppose that the sub-gradients of the objective functions exist, then two minimum solutions must satisfy the first-order optimal condition, i.e., for all $k \in \mathbb{N}$,

$$\begin{cases} 0 \in \partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k}) + \partial f(\boldsymbol{X}_{(3)}^{k+1}) + \rho(\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k}), \\ 0 \in \partial_{\boldsymbol{E}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) + \partial g(\boldsymbol{E}^{k+1}) + \sigma(\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}). \end{cases} \tag{4.29}$$

Based on the sub-differentiability property, i.e.,

$$\partial \Phi(\boldsymbol{Z}^{k+1}) = \Big(\partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) + \partial f(\boldsymbol{X}_{(3)}^{k+1}),\ \partial_{\boldsymbol{E}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) + \partial g(\boldsymbol{E}^{k+1}) \Big),$$

then we have the triangle inequality as follows,

$$\|\partial \Phi(\boldsymbol{Z}^{k+1})\|_F \leq \|\partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) + \partial f(\boldsymbol{X}_{(3)}^{k+1})\|_F + \|\partial_{\boldsymbol{E}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) + \partial g(\boldsymbol{E}^{k+1})\|_F.$$

Substituting into the first-order optimal condition (4.29), we have

$$\begin{aligned} \|\partial \Phi(\boldsymbol{Z}^{k+1})\|_F &\leq \|\partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) - \partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k}) - \rho(\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k})\|_F \\ &\quad + \sigma \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}\|_F \\ &\leq \|\partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) - \partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k})\|_F + \rho \|\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k}\|_F \\ &\quad + \sigma \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}\|_F. \end{aligned} \tag{4.30}$$

Following the Lipschitz continuity, $h \in C^1$ is a continuously differentiable function with $\partial_{\boldsymbol{X}_{(3)}} h$ assumed to be L_h-Lipschitz continuous. Then, we easily have

$$\|\partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k+1}) - \partial_{\boldsymbol{X}_{(3)}} h(\boldsymbol{X}_{(3)}^{k+1}, \boldsymbol{E}^{k})\|_F \leq L_h \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}\|_F. \tag{4.31}$$

Resultantly, a backsubstitution from (4.31) into (4.30) yields for all $k \geq 0$,

$$\begin{aligned} \|\partial \Phi(\boldsymbol{Z}^{k+1})\|_F &\leq L_h \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}\|_F + \rho \|\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k}\|_F + \sigma \|\boldsymbol{E}^{k+1} - \boldsymbol{E}^{k}\|_F \\ &\leq (L_h + \sigma) \|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^{k}\|_F + \rho \|\boldsymbol{X}_{(3)}^{k+1} - \boldsymbol{X}_{(3)}^{k}\|_F \\ &\leq (L_h + \sigma + \rho) \|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^{k}\|_F. \end{aligned}$$

The relative error condition is proved. □

Lemma 4. *Let $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ be the sequence generated by the proposed Algorithm 2, where $\boldsymbol{Z}^k := (\boldsymbol{X}_{(3)}^k, \boldsymbol{E}^k)$. Then, $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ is bounded.*

Proof. According to Algorithm 2, the initial variables $\boldsymbol{X}_{(3)}^0$ and $\boldsymbol{E}^0$ are bounded, hence $\Phi(\boldsymbol{Z}^0)$ is bounded. Relying upon Lemma 2, i.e., the sequence $\{\Phi(\boldsymbol{Z}^k)\}_{k\in\mathbb{N}}$ decreases sufficiently, we then have $0 \leq \Phi(\boldsymbol{Z}^k) \leq \Phi(\boldsymbol{Z}^0)$ for $\forall k \in \mathbb{N}$. That is, $0 \leq h(\boldsymbol{Z}^k) + f(\boldsymbol{X}_{(3)}^k) + g(\boldsymbol{E}^k) \leq \Phi(\boldsymbol{Z}^0)$ with $h, f, g \geq 0$ for all the discrete points. Thus, sequences $\{h(\boldsymbol{Z}^k)\}_{k\in\mathbb{N}}$, $\{f(\boldsymbol{X}_{(3)}^k)\}_{k\in\mathbb{N}}$, and $\{g(\boldsymbol{E}^k)\}_{k\in\mathbb{N}}$ are bounded.

Firstly, since f is proper and coercive, there exists $\|\boldsymbol{X}_{(3)}^k\|_F \to \infty$ if and only if $f(\boldsymbol{X}_{(3)}^k) \to \infty$. Obviously, $f(\boldsymbol{X}_{(3)}^k) \not\to \infty$ for $\forall k \in \mathbb{N}$, thus $\|\boldsymbol{X}_{(3)}^k\|_F \not\to \infty$, i.e., the sequence $\{\boldsymbol{X}_{(3)}^k\}_{t\in\mathbb{N}}$ is bounded. Additionally, the triangle inequality gives

$$\|\boldsymbol{E}^k\|_F - \|\boldsymbol{H}(\boldsymbol{X}_{(3)}^k)^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top\|_F \leq \left\| \boldsymbol{E}^k - \left(\boldsymbol{H}(\boldsymbol{X}_{(3)}^k)^\top - \boldsymbol{H}\boldsymbol{P}_{(3)}^\top \right) \right\|_F, \quad \forall k \in \mathbb{N},$$

which is equivalent to

$$\begin{aligned}\|\boldsymbol{E}^k\|_F &\leq \left\|\boldsymbol{E}^k - \left(\boldsymbol{H}(\boldsymbol{X}^k_{(3)})^\top - \boldsymbol{H}\boldsymbol{P}^\top_{(3)}\right)\right\|_F + \|\boldsymbol{H}(\boldsymbol{X}^k_{(3)})^\top - \boldsymbol{H}\boldsymbol{P}^\top_{(3)}\|_F \\ &\leq \|\boldsymbol{H}(\boldsymbol{X}^k_{(3)})^\top - \boldsymbol{H}\boldsymbol{P}^\top_{(3)} - \boldsymbol{E}^k\|_F + \|\boldsymbol{H}(\boldsymbol{X}^k_{(3)})^\top\|_F + \|\boldsymbol{H}\boldsymbol{P}^\top_{(3)}\|_F \\ &= \frac{\lambda_1\|\boldsymbol{H}(\boldsymbol{X}^k_{(3)})^\top - \boldsymbol{H}\boldsymbol{P}^\top_{(3)} - \boldsymbol{E}^k\|_F}{\lambda_1} + \sqrt{\text{Trace}\left(\boldsymbol{X}^k_{(3)}(\boldsymbol{X}^k_{(3)})^\top\right)} + \sqrt{\text{Trace}\left(\boldsymbol{P}_{(3)}\boldsymbol{P}^\top_{(3)}\right)} \\ &= \sqrt{\frac{h(\boldsymbol{Z}^k)}{\lambda_1}} + \|(\boldsymbol{X}^k_{(3)})^\top\|_F + \|\boldsymbol{P}^\top_{(3)}\|_F \not\to \infty, \quad \forall\lambda_1 > 0,\ \forall k \in \mathbb{N}.\end{aligned} \tag{4.32}$$

Namely, the sequence $\{\boldsymbol{E}^k\}_{k\in\mathbb{N}}$ is bounded. Consequently, the sequence $\{(\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k)\}_{k\in\mathbb{N}}$, i.e., $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$, is certainly bounded. This completes the proof. □

Equipped with Lemmas 1-4, Theorem 1 theoretically establishes the convergence proof of Algorithm 2.

Theorem 1. *Let $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ be a sequence generated by the proposed Algorithm 2, where $\boldsymbol{Z}^k := (\boldsymbol{X}^k_{(3)}, \boldsymbol{E}^k)$. Then, $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ converges to a critical point $\boldsymbol{Z}^* := (\boldsymbol{X}^*_{(3)}, \boldsymbol{E}^*)$ of Φ.*

Proof. According to Lemma 4, the sequence $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ is bounded. Hence, there exists a convergence subsequence $\{\boldsymbol{Z}^{k_d}\}_{d\in\mathbb{N}}$ such that $\lim\limits_{d\to\infty} \boldsymbol{Z}^{k_d} = \bar{\boldsymbol{Z}}$, where $\bar{\boldsymbol{Z}}$ is a limit point of $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$. Since h, $f \in C^1$ are continuous functions, and g is lower semi-continuous, i.e., $\liminf\limits_{k\to\infty} g(\boldsymbol{E}^k) \geq g(\bar{\boldsymbol{E}})$, we can obtain

$$\lim_{k\to\infty} \Phi(\boldsymbol{Z}^k) = \Phi(\bar{\boldsymbol{Z}}). \tag{4.33}$$

(*i*) If there exists an integer $\bar{k} \in \mathbb{N}$ for which $\Phi(\boldsymbol{Z}^{\bar{k}}) = \Phi(\bar{\boldsymbol{Z}})$, then Lemma 2 can guarantee that $\Phi(\boldsymbol{Z}^{\bar{k}}) = \Phi(\boldsymbol{Z}^{\bar{k}+1}) = \cdots = \Phi(\boldsymbol{Z}^{k\to\infty}) = \Phi(\bar{\boldsymbol{Z}})$, leading to $\boldsymbol{Z}^{\bar{k}} = \boldsymbol{Z}^{\bar{k}+1} = \cdots = \boldsymbol{Z}^{k\to\infty} = \bar{\boldsymbol{Z}}$. Thus, the sequence $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ is stationary, and Theorem 1 obviously holds with $\boldsymbol{Z}^{\bar{k}}$ being the critical point of Φ.
(*ii*) For another case, the integer $\bar{k}$ is nonexistent. Reiterating Lemma 2, the nonincreasing property of Φ implies

$$\Phi(\boldsymbol{Z}^k) > \Phi(\bar{\boldsymbol{Z}}), \quad \forall k > 0. \tag{4.34}$$

Relying upon (4.33), for any $\delta > 0$, there exists $k_1 \in \mathbb{N}_+$ such that

$$\Phi(\boldsymbol{Z}^k) < \Phi(\bar{\boldsymbol{Z}}) + \delta, \quad \forall k > k_1. \tag{4.35}$$

Denote $\vartheta(\boldsymbol{Z}^0)$ as the limit point set of the sequence $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$, then an elementary consequence is followed as $\lim\limits_{k\to\infty} \text{dist}\left(\boldsymbol{Z}^k, \vartheta(\boldsymbol{Z}^0)\right) = 0$. That is, for any $\gamma > 0$, there exists $k_2 \in \mathbb{N}_+$ such that

$$\text{dist}\left(\boldsymbol{Z}^k, \vartheta(\boldsymbol{Z}^0)\right) < \gamma, \quad \forall k > k_2. \tag{4.36}$$

Let $\tau = \max(k_1, k_2)$, summarizing inequalities (4.34)-(4.36) yields

$$\boldsymbol{Z}^k \in \left\{\boldsymbol{Z} : \text{dist}\left(\boldsymbol{Z}, \vartheta(\boldsymbol{Z}^0)\right) < \gamma\right\} \cap \left\{\boldsymbol{Z} : \Phi(\bar{\boldsymbol{Z}}) < \Phi(\boldsymbol{Z}) < \Phi(\bar{\boldsymbol{Z}}) + \delta\right\}, \quad \forall k > \tau.$$

From Lemma 1, the objective function Φ in (4.13) is a KŁ function. Thus, for any $k > \tau$, the KŁ inequality of Definition 4.2 holds, i.e.,

$$\varphi'\left(\Phi(\boldsymbol{Z}^k) - \Phi(\bar{\boldsymbol{Z}})\right) \text{dist}\left(0, \partial\Phi(\boldsymbol{Z}^k)\right) \geq 1. \tag{4.37}$$

Moreover, the concavity of function φ gives

$$\varphi\left(\Phi(\boldsymbol{Z}^{k+1}) - \Phi(\bar{\boldsymbol{Z}})\right) - \varphi\left(\Phi(\boldsymbol{Z}^k) - \Phi(\bar{\boldsymbol{Z}})\right) \leq \varphi'\left(\Phi(\boldsymbol{Z}^k) - \Phi(\bar{\boldsymbol{Z}})\right)\left(\Phi(\boldsymbol{Z}^{k+1}) - \Phi(\boldsymbol{Z}^k)\right). \tag{4.38}$$

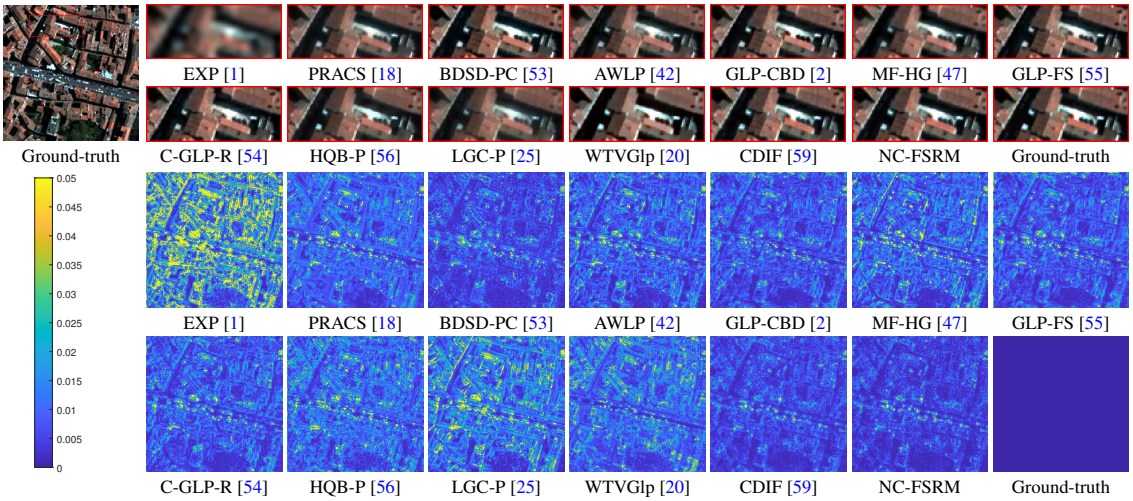

Fig. 4.2: The visual results on the reduced-resolution Pléiades2 dataset (sensor: Pléiades). The size of the ground-truth image is $256 \times 256 \times 4$. The top two rows: the close-ups of images recovered by all compared methods, also the proposed NC-FSRM. The bottom two rows: the corresponding residual maps.

For all b, $d \in \mathbb{N}$, and $\bar{\boldsymbol{Z}}$, we firstly define

$$\Gamma_{b,d} := \varphi\left(\Phi(\boldsymbol{Z}^b) - \Phi(\bar{\boldsymbol{Z}})\right) - \varphi\left(\Phi(\boldsymbol{Z}^d) - \Phi(\bar{\boldsymbol{Z}})\right), \quad \text{and} \quad C := \frac{\rho + \sigma + L_h}{\min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)} \in (0, +\infty). \tag{4.39}$$

Considering formulas (4.37)-(4.38) and Lemmas 2-3, we then obtain for any $k > \tau$,

$$\begin{aligned}\|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^k\|_F^2 \leq \frac{\Gamma_{k,k+1}}{\min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)} \frac{1}{\varphi'\left(\Phi(\boldsymbol{Z}^k) - \Phi(\bar{\boldsymbol{Z}})\right)} &\leq \frac{\Gamma_{k,k+1}}{\min\left(\frac{\rho}{2}, \frac{\sigma}{2}\right)} \operatorname{dist}\left(0, \partial\Phi(\boldsymbol{Z}^k)\right) \\ &\leq C\Gamma_{k,k+1}\|\boldsymbol{Z}^k - \boldsymbol{Z}^{k-1}\|_F.\end{aligned}$$

Based on the inequality of arithmetic and geometric means, i.e., $2\sqrt{\alpha\beta} \leq \alpha + \beta,\ \forall \alpha, \beta \geq 0$, we further have

$$2\|\boldsymbol{Z}^{k+1} - \boldsymbol{Z}^k\|_F \leq 2\sqrt{C\Gamma_{k,k+1}\|\boldsymbol{Z}^k - \boldsymbol{Z}^{k-1}\|_F} \leq C\Gamma_{k,k+1} + \|\boldsymbol{Z}^k - \boldsymbol{Z}^{k-1}\|_F, \quad \forall k > \tau. \tag{4.40}$$

Summing up (4.40) for all $i = \tau + 1, \cdots, k$ yields

$$\begin{aligned}2\sum_{i=\tau+1}^{k} \|\boldsymbol{Z}^{i+1} - \boldsymbol{Z}^i\|_F &\leq \sum_{i=\tau+1}^{k} \|\boldsymbol{Z}^i - \boldsymbol{Z}^{i-1}\|_F + C\sum_{i=\tau+1}^{k} \Gamma_{i,i+1} \\ &\leq \sum_{i=\tau+1}^{k} \|\boldsymbol{Z}^{i+1} - \boldsymbol{Z}^i\|_F + \|\boldsymbol{Z}^{\tau+1} - \boldsymbol{Z}^\tau\|_F + C\sum_{i=\tau+1}^{k} \Gamma_{i,i+1} \\ &= \sum_{i=\tau+1}^{k} \|\boldsymbol{Z}^{i+1} - \boldsymbol{Z}^i\|_F + \|\boldsymbol{Z}^{\tau+1} - \boldsymbol{Z}^\tau\|_F + C\Gamma_{\tau+1,k+1},\end{aligned}$$

which is equivalent to

$$\sum_{i=\tau+1}^{k} \|\boldsymbol{Z}^{i+1} - \boldsymbol{Z}^i\|_F \leq \|\boldsymbol{Z}^{\tau+1} - \boldsymbol{Z}^\tau\|_F + C\Gamma_{\tau+1,k+1}. \tag{4.41}$$

Following (4.33), (4.39), and (4.41), the sequence $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ easily has the finite length property as follows,

$$\begin{aligned}\sum_{i=0}^{\infty}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F &= \sum_{i=0}^{\tau}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F + \sum_{i=\tau+1}^{\infty}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F\\ &\leq \sum_{i=0}^{\tau}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F + \|\boldsymbol{Z}^{\tau+1}-\boldsymbol{Z}^{\tau}\|_F + C\Gamma_{\tau+1,\infty}\\ &= \sum_{i=0}^{\tau}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F + \|\boldsymbol{Z}^{\tau+1}-\boldsymbol{Z}^{\tau}\|_F + C\varphi\left(\Phi(\boldsymbol{Z}^{\tau+1})-\Phi(\bar{\boldsymbol{Z}})\right)\\ &< +\infty,\end{aligned}$$

which implies that for any $r>0$, there exists an integer $k_3>\tau$ such that

$$\|\boldsymbol{Z}^b-\boldsymbol{Z}^d\|_F = \left\|\sum_{i=d}^{b-1}(\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i)\right\|_F \leq \sum_{i=d}^{b-1}\|\boldsymbol{Z}^{i+1}-\boldsymbol{Z}^i\|_F < r, \quad \forall b>d>k_3.$$

This follows that $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. With the domain space of $\boldsymbol{Z}$ is completeness, the Cauchy sequence $\{\boldsymbol{Z}^k\}_{k\in\mathbb{N}}$ can certainly converge to a critical point $(\boldsymbol{X}^*_{(3)},\boldsymbol{E}^*)$ of Φ. The proof is completed. □

5 Experimental Results

This section is devoted to the comparison between the proposed NC-FSRM and other state-of-the-art methods, i.e., EXP [1], PRACS [18], BDSD-PC [53], AWLP [42], GLP-CBD [2], MF-HG [47], GLP-FS [55], C-GLP-R [54], HQB-P [56], LGC-P [25], WTVGlp [20], and CDIF [59], whose source codes are available at either of the website[1] or the authors' homepages. For fairness, all hyper-parameters involved in these methods are fine-tuned within a specific range suggested by their authors, aiming to achieve the best performance. To adequately verify the superiority of the proposed method, four commonly-used datasets[2] are employed, including the Pléiades2 dataset (sensor: Pléiades), the Guangzhou dataset (sensor: GF-2), the Alice dataset (sensor: WV-4), and the Tripoli dataset (sensor: WV-3). Notably, all experimental data are pre-normalized into [0, 1]. All numerical experiments are implemented in MATLAB (R2016a) on a computer of 16Gb RAM and an Intel(R) Core(TM) i7-5960X CPU: @3.00 GHz.

Regarding evaluation metrics, two types of experiments are distinguished. Specifically, when evaluated at reduced-resolution (i.e., simulated) images, six widely-used metrics are adopted, i.e., the peak signal-to-noise ratio (PSNR), the structural similarity index (SSIM) [57], the spectral angle mapper (SAM) [63], the spatial correlation coefficient (SCC) [42], the erreur relative globale adimensionnelle de synthèse (ERGAS) [6], and the $Q2^n$ [5]. Instead, when evaluated at full-resolution (i.e., real) images, the quality with no reference (QNR) [4] comprised of a spectral distortion index D_λ and a spatial distortion index D_s is specified.

5.1 Parameters Setting

As illustrated in Algorithm 2, ten parameters are involved, i.e., λ_1, λ_2, η_1, η_2, ρ, σ, r, ε, k_{max}, and p_{max}. For pansharpening, the hardware characteristics of both MS and PAN images fix the scale factor, i.e., $r=4$, in all experiments. Furthermore, $\varepsilon=2\times10^{-5}$ and $k_{max}=200$ can be determined empirically. Significantly, the warm-start strategy is utilized in Algorithm 1 to accelerate convergence, thus p_{max} can be set as a smaller value, e.g., $p_{max}=2$ in this paper. Moreover, we forcibly assign $\rho=\sigma$, aiming at reducing the parameter burden. Therefore, only five parameters, i.e., λ_1, λ_2, η_1, η_2, and ρ, need to be further adjusted manually. For reduced-resolution experiments, λ_1, λ_2, η_1, η_2, and ρ can be numerically selected from the candidate sets

[1] http://openremotesensing.net/kb/codes/pansharpening/
[2] http://www.digitalglobe.com/samples?search=Imagery

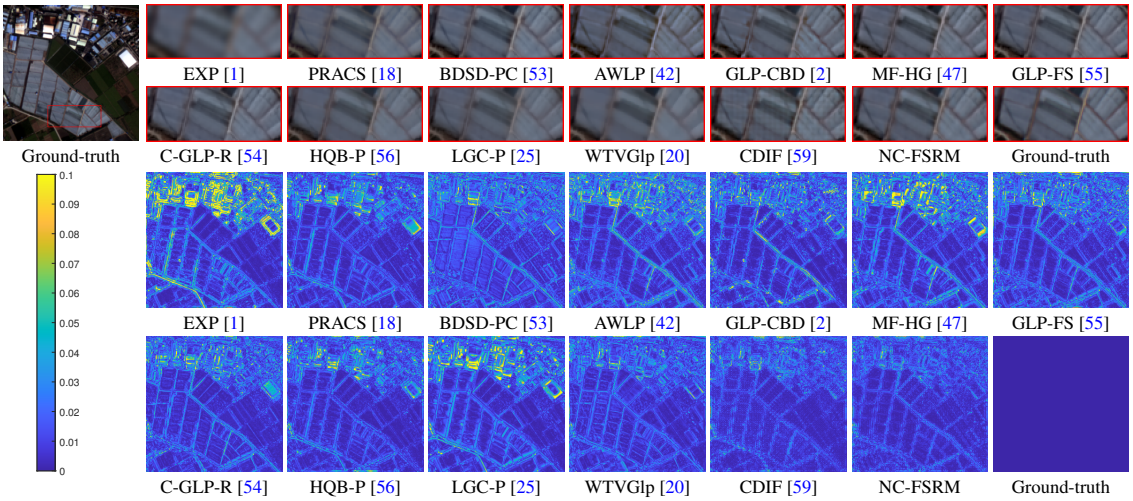

Fig. 5.3: The visual results on the reduced-resolution Guangzhou dataset (sensor: GF-2). The size of the ground-truth image is $256 \times 256 \times 4$. The top two rows: the close-ups of images recovered by all compared methods, also the proposed NC-FSRM. The bottom two rows: the corresponding residual maps.

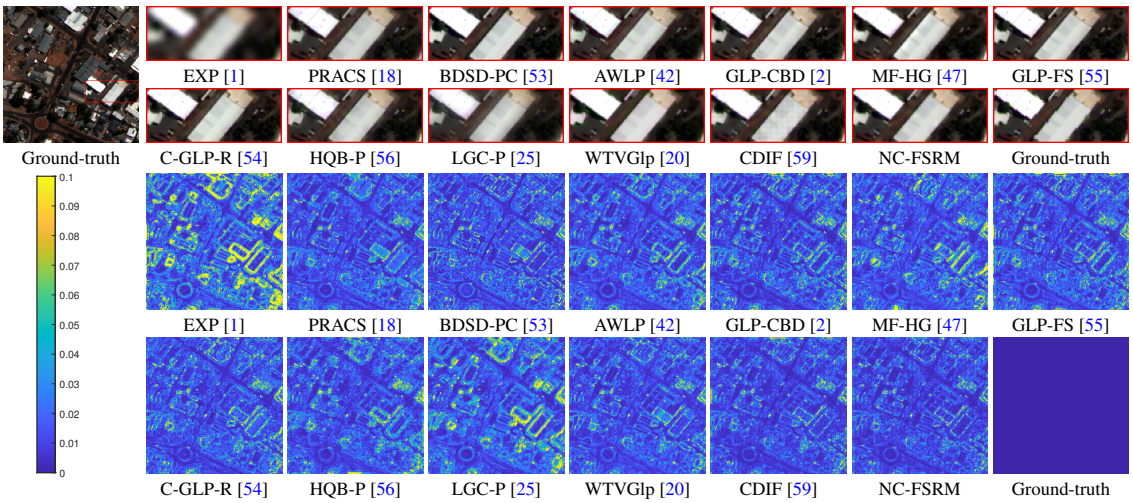

Fig. 5.4: The visual results on the reduced-resolution Alice dataset (sensor: WV-4). The size of the ground-truth image is $256 \times 256 \times 4$. The top two rows: the close-ups of images recovered by all compared methods, also the proposed NC-FSRM. The bottom two rows: the corresponding residual maps.

$\{4.5\times10^{-4}, 5.7\times10^{-4}\}$, $\{1.5\times10^{-7}, 3\times10^{-7}, 4.5\times10^{-7}, 6\times10^{-7}, 7.5\times10^{-7}\}$, $\{2\times10^{-2}, 4\times10^{-2}, 8\times10^{-2}, 1.6\times10^{-1}, 3.2\times10^{-1}\}$, $\{1.4\times10^{-5}, 2.8\times10^{-5}, 4.2\times10^{-5}\}$, $\{1.5\times10^{-3}, 1.5\times10^{-2}, 1.5\times10^{-1}, 1.5\times10^{0}\}$, respectively. For full-resolution experiments, λ_1, λ_2, η_1, η_2, and ρ are recommended to select from the candidate sets $\{3.5\times10^{-4}, 3.5\times10^{-3}, 3.5\times10^{-2}\}$, $\{1\times10^{-7}, 5\times10^{-7}, 1\times10^{-6}, 5\times10^{-6}, 1\times10^{-5}\}$, $\{3\times10^{-4}, 6\times10^{-4}, 1.2\times10^{-3}\}$, $\{3.5\times10^{-5}, 3.5\times10^{-4}, 3.5\times10^{-3}\}$, $\{1.5\times10^{-3}, 1.5\times10^{-2}, 1.5\times10^{-1}, 1.5\times10^{0}\}$, respectively. More specifically, Table 5.1 provides the hyper-parameter configurations for all the performed experiments.

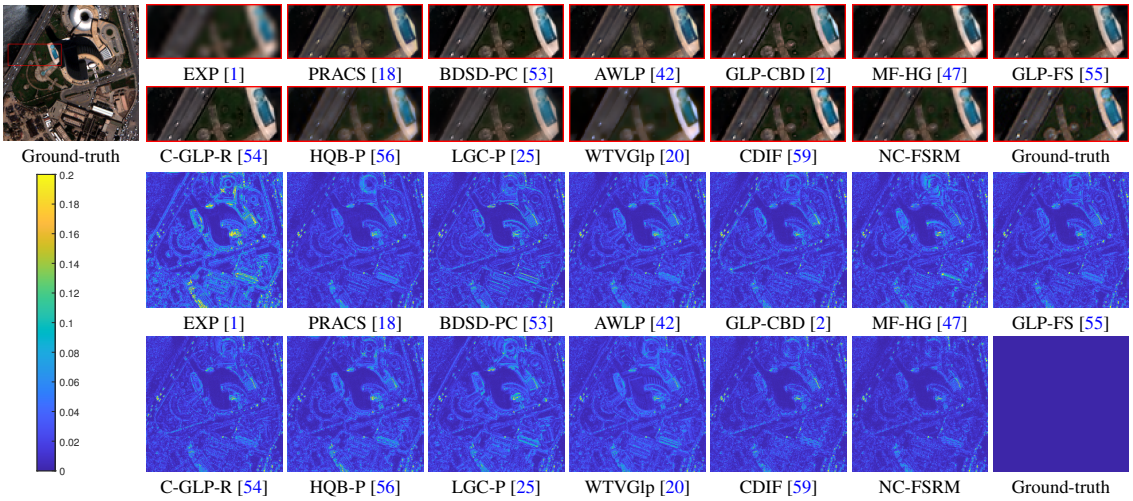

Fig. 5.5: The visual results on the reduced-resolution Tripoli dataset (sensor: WV-3). The size of the ground-truth image is $256 \times 256 \times 8$. The top two rows: the close-ups of images recovered by all compared methods, also the proposed NC-FSRM. The bottom two rows: the corresponding residual maps.

Table 5.1: The hyper-parameter configurations of the proposed NC-FSRM model for different experiments. (R: Reduced resolution; F: Full resolution)

Dataset	Sensor	Case	λ_1	λ_2	η_1	η_2	ρ
Pléiades2	Pléiades	R	5.7×10^{-4}	1.7×10^{-7}	3.0×10^{-1}	4.1×10^{-5}	5.8×10^{-2}
Guangzhou	GF-2	R	5.7×10^{-4}	7.3×10^{-7}	3.8×10^{-2}	4.0×10^{-5}	1.9×10^{-1}
		F	3.7×10^{-4}	4.3×10^{-7}	1.0×10^{-3}	4.0×10^{-5}	1.9×10^{-1}
Alice	WV-4	R	4.5×10^{-4}	1.0×10^{-5}	2.2×10^{-2}	2.2×10^{-5}	1.3×10^{-1}
Tripoli	WV-3	R	4.1×10^{-4}	9.0×10^{-7}	2.2×10^{-2}	1.4×10^{-5}	1.2×10^{-1}
		F	2.2×10^{-3}	1.3×10^{-6}	6.0×10^{-4}	3.0×10^{-5}	1.4×10^{-1}

5.2 The Reduced-Resolution Experiments

The simulated ground-truth, LR-MS, and PAN images are generated by the real LR-MS, blurred and then down-sampled LR-MS, and blurred and then down-sampled PAN images, respectively. Similar to the settings in (4.19), the Gaussian-shaped kernels $\{h_i \in \mathbb{R}^{41\times41}\}_{i=1}^{S}$ with the known standard deviations are utilized.

The qualitative comparison. As aforementioned, we here provide the comparative results on four reduced-resolution datasets, visualized in Figs. 4.2-5.5. Inspecting Fig. 4.2, only the results recovered by CDIF and NC-FSRM can inherently generate similar structures with the ground-truth image without significant spectral distortion, whereas other compared methods inevitably cause spectral and spatial distortion, leading to some artifacts and fading. From the corresponding residual maps, we further observe that the proposed NC-FSRM realizes the minimal error using the ground-truth image as a reference, confirming its superiority.

The performances in Figs. 5.3-5.5 are similar to that of Fig. 4.2. Among the different experiments, the proposed NC-FSRM consistently achieves promising results compared with all the other methods. Particularly, although CDIF obtains high performance, such a method creates rasterization in smooth image areas, especially for Figs. 5.3-5.4, resulting from its non-local coefficient estimating strategy. Different from CDIF, the proposed framelet-based sparse modeling simultaneously takes the piece-wise smoothness property into account, thus experimentally exhibiting better capability for prior characterization.

The quantitative comparison. To better illustrate the above analysis, we further provide quantitative

Table 5.2: Quality metrics of all the compared methods on the reduced-resolution (a) Pléiades2 dataset (sensor: Pléiades); (b) Guangzhou dataset (sensor: GF-2); (c) Alice dataset (sensor: WV-4); and (d) Tripoli dataset (sensor: WV-3). (Bold: best; Underline: second best)

Methods	(a) Pléiades2 dataset (sensor: Pléiades)						(b) Guangzhou dataset (sensor: GF-2)					
	PSNR↑	SSIM↑	SAM↓	SCC↑	ERGAS↓	Q4↑	PSNR↑	SSIM↑	SAM↓	SCC↑	ERGAS↓	Q4↑
EXP [1]	27.8770	0.7064	4.6437	0.8517	5.7290	0.7680	29.1360	0.8120	1.5222	0.9424	2.5521	0.8545
PRACS [18]	33.4027	0.9265	4.1063	0.9682	3.1673	0.9398	33.2858	0.9198	1.5457	0.9774	1.6267	0.9280
BDSD-PC [53]	36.1990	0.9606	3.2758	0.9749	2.4078	0.9710	34.0490	0.9274	1.5596	0.9835	1.4451	0.9399
AWLP [42]	34.2092	0.9469	3.4829	0.9653	2.8765	0.9543	32.7305	0.8945	2.1090	0.9730	1.7032	0.9154
GLP-CBD [2]	35.2919	0.9557	3.1223	0.9712	2.5706	0.9674	33.4882	0.9154	1.4558	0.9785	1.5120	0.9299
MF-HG [47]	33.4754	0.9404	3.3617	0.9580	3.0378	0.9495	32.2083	0.9141	1.5318	0.9733	1.8052	0.9250
GLP-FS [55]	35.0074	0.9544	3.3868	0.9706	2.5543	0.9647	33.6165	0.9177	1.5367	0.9806	1.4574	0.9328
C-GLP-R [54]	34.8654	0.9494	3.7962	0.9682	2.6971	0.9638	34.3090	0.9354	1.3721	0.9826	1.4004	0.9503
HQB-P [56]	33.9065	0.9246	3.7898	0.9692	2.9611	0.9502	34.5815	0.9426	1.1511	0.9858	1.3234	0.9555
LGC-P [25]	31.7598	0.9030	3.8345	0.9539	3.7047	0.9116	32.3290	0.9162	1.4591	0.9743	1.7842	0.9203
WTVGlp [20]	33.7182	0.9210	5.1323	0.9696	2.9433	0.9520	34.8210	0.9402	1.4189	0.9843	1.3146	0.9495
CDIF [59]	37.5372	0.9650	2.8498	0.9817	1.9872	0.9791	36.1829	0.9479	**1.1259**	0.9891	1.0810	**0.9670**
NC-FSRM	**38.2551**	**0.9695**	**2.6682**	**0.9837**	**1.8345**	**0.9815**	**36.7373**	**0.9485**	1.1992	**0.9902**	**1.0213**	0.9653
—	(c) Alice dataset (sensor: WV-4)						(d) Tripoli dataset (sensor: WV-3)					
	PSNR↑	SSIM↑	SAM↓	SCC↑	ERGAS↓	Q4↑	PSNR↑	SSIM↑	SAM↓	SCC↑	ERGAS↓	Q8↑
EXP [1]	24.6865	0.6526	3.9590	0.9169	5.1208	0.8050	27.3924	0.6992	4.2012	0.9159	4.9479	0.8097
PRACS [18]	29.1953	0.8725	4.4999	0.9693	3.1708	0.9302	30.9542	0.8785	4.3733	0.9634	3.2777	0.9209
BDSD-PC [53]	29.2411	0.8859	4.1863	0.9699	3.0916	0.9397	30.6140	0.8781	4.3452	0.9614	3.4689	0.9271
AWLP [42]	28.9846	0.8815	4.1295	0.9664	3.2228	0.9346	30.6673	0.8723	4.5644	0.9591	3.5129	0.9190
GLP-CBD [2]	28.6345	0.8780	4.0134	0.9637	3.3036	0.9327	30.1131	0.8629	4.4360	0.9516	3.7448	0.9057
MF-HG [47]	28.3126	0.8772	4.2105	0.9623	3.4386	0.9306	30.6589	0.8789	4.0756	0.9601	3.4463	0.9223
GLP-FS [55]	28.7045	0.8781	4.2601	0.9677	3.1359	0.9361	31.2984	0.8829	4.0283	0.9656	3.1557	0.9303
C-GLP-R [54]	28.9803	0.8795	4.1250	0.9667	3.2061	0.9360	31.0477	0.8809	4.2333	0.9635	3.2895	0.9307
HQB-P [56]	28.9437	0.8733	3.7650	0.9694	3.1313	0.9328	30.4370	0.8450	4.0849	0.9602	3.4685	0.9126
LGC-P [25]	27.5171	0.8268	4.1854	0.9570	3.8286	0.8919	30.4761	0.8592	4.1018	0.9608	3.5141	0.9053
WTVGlp [20]	29.3398	0.8791	4.3992	0.9705	3.0680	0.9379	30.0676	0.8540	4.6679	0.9586	3.5634	0.9065
CDIF [59]	30.2709	0.8967	**3.3572**	0.9760	2.7081	0.9532	32.0088	0.8955	3.7222	0.9715	2.9275	0.9457
NC-FSRM	**30.7283**	**0.9038**	3.4779	**0.9778**	**2.5998**	**0.9556**	**32.6974**	**0.9049**	**3.4775**	**0.9751**	**2.6747**	**0.9509**

comparisons of all methods on the four reduced-resolution experiments. The statistical values of six metrics, i.e., PSNR, SSIM, SAM, SCC, ERGAS, and $Q2^n$, are recorded in Table 5.2. From Table 5.2, we observe that the proposed NC-FSRM invariably achieves competitive results, indicating the effectiveness of NC-FSRM.

5.3 The Full-Resolution Experiments

The qualitative comparison. To avoid a potential bias privileging the methods applying a parallel hypothesis for synthesizing the LR-MS and PAN images, we further test the proposed NC-FSRM on full-resolution experiments, whose confirmation is more objective and convincing. For full-resolution experiments, the full-resolution Guangzhou dataset (sensor: GF-2) and Tripoli dataset (sensor: WV-3) are utilized. Under this case, the ground-truth images are unavailable, and thus, visual comparisons require the PAN images as spatial reference, while the upsampled/interpolated LR-MS images, e.g., EXP [1], as spectral reference.

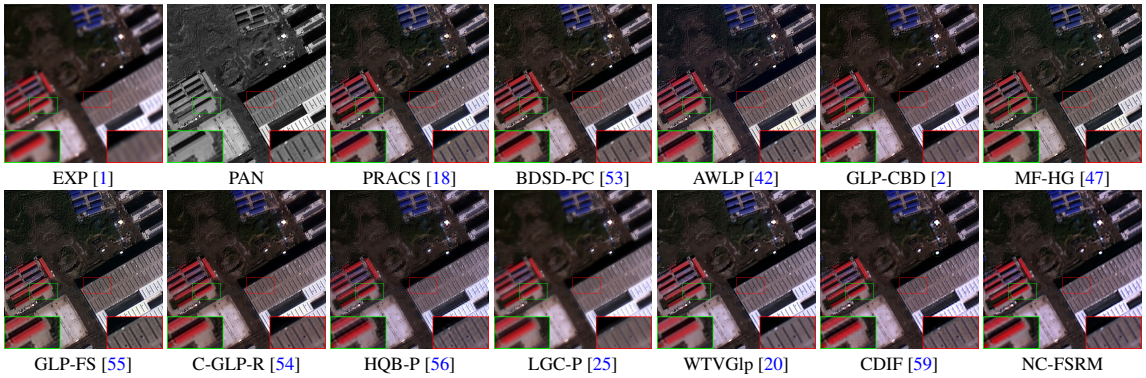

Fig. 5.6: The visual results on the full-resolution Guangzhou dataset (sensor: GF-2). The size of the PAN image is 400×400. Close-ups are depicted in corners, zooming in to see the details.

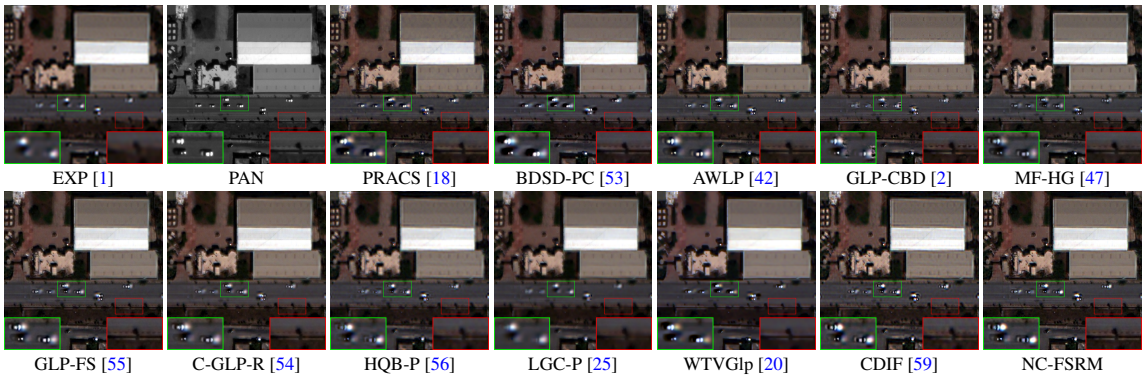

Fig. 5.7: The visual results on the full-resolution Tripoli dataset (sensor: WV-3). The size of the PAN image is 400×400. Close-ups are depicted in corners, zooming in to see the details.

Subsequently, all visual inspections are displayed in Figs. 5.6-5.7. Compared with the EXP method, all the other methods significantly contribute to a visual improvement in terms of spatial resolution, showing a relatively unambiguous and distinct sharpening. More specifically, the AWLP, MF-HG, GLP-FS, C-GLP-R, and WTVGlp algorithms exhibit varying levels of spectral distortion, which are reflected in the white striped area in the lower-right corner of Fig. 5.6. From the close-ups of Figs. 5.6-5.7, we can also observe that the HQB-P, LGC-P, and CDIF methods suffer from visible aliasing effects and blurring. Also, the PRACS, BDSD-PC, and especially GLP-CBD methods generate indistinguishable or nonexistent objects and structures. Conversely, the proposed NC-FSRM performs the details closest to the PAN image regarded as a spatial reference, yet prevents apparent spectral aberration, directly confirmed by the visual inspections.

The quantitative comparison. To numerically validate the superiority of the proposed NC-FSRM, the quantitative comparisons of all methods on the two full-resolution experiments are also conducted. The statistical values of three metrics, i.e., D_λ, D_s, and QNR, are provided in Table 5.2. From Table 5.2, we observe that the proposed NC-FSRM favorably outperforms the other methods regarding the QNR values, while also making an excellent trade-off between the mitigation of spectral and spatial distortions, showing the stability of NC-FSRM.

Table 5.3: Quantitative results for all the compared methods on the full-resolution (a) Guangzhou dataset (sensor: GF-2); and (b) Tripoli dataset (sensor: WV-3). (Bold: best; Underline: second best)

Methods	(a) Guangzhou dataset (sensor: GF-2)			(b) Tripoli dataset (sensor: WV-3)		
	$D_\lambda \downarrow$	$D_s \downarrow$	QNR↑	$D_\lambda \downarrow$	$D_s \downarrow$	QNR↑
EXP [1]	**0.0029**	0.1255	0.8720	**0.0024**	0.0514	0.9463
PRACS [18]	0.0478	0.0549	0.8999	0.0195	0.0503	0.9311
BDSD-PC [53]	0.0726	0.0731	0.8596	0.0123	0.0491	0.9392
AWLP [42]	0.0396	0.0696	0.8935	0.0289	0.0432	0.9291
GLP-CBD [2]	0.0588	0.0356	0.9077	0.0261	0.0311	0.9436
MF-HG [47]	0.0710	0.0661	0.8676	0.0433	0.0282	0.9297
GLP-FS [55]	0.0748	0.0618	0.8680	0.0312	0.0342	0.9357
C-GLP-R [54]	0.0629	0.0497	0.8905	0.0165	<u>0.0236</u>	0.9603
HQB-P [56]	0.0445	**0.0162**	0.9401	0.0111	0.0270	<u>0.9622</u>
LGC-P [25]	<u>0.0053</u>	0.0551	0.9398	<u>0.0054</u>	0.0339	0.9609
WTVGlp [20]	0.1018	0.0824	0.8242	0.0478	0.0682	0.8873
CDIF [59]	0.0294	<u>0.0220</u>	<u>0.9493</u>	0.0227	0.0258	0.9521
NC-FSRM	0.0132	0.0237	**0.9634**	0.0102	**0.0092**	**0.9807**

6 Conclusions

In this paper, we proposed a framelet sparse reconstruction method for pansharpening, i.e., NC-FSRM, which investigates the coefficient similarity among the underlying HR-MS image and the PAN image on the framelet domain, then characterizes the strong statistical sparsity of their error using ℓ_0 norm. The proposed NC-FSRM not only more precisely and concisely established the relationship between the underlying HR-MS image and the PAN image, but also simultaneously characterized the piece-wise smoothness prior of the former without adding additional regularizers, showing superior properties. Furthermore, we developed an efficient PAM-based solving algorithm for the proposed nonconvex regularization model. Also, we theoretically prove that the algorithm can globally converge to a critical point under some mild assumptions. Substantial experiments on reducted-resolution and full-resolution datasets confirmed that the proposed NC-FSRM markedly outperforms other state-of-the-art pansharpening methods, both qualitatively and quantitatively.

References

[1] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Trans. Geosci. Remote Sens.*, 40(10):2300–2312, 2002.

[2] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogramm. Eng. Remote Sens.*, 72(5):591–596, 2006.

[3] B. Aiazzi, S. Baronti, and M. Selva. Improving component substitution pansharpening through multivariate regression of MS + Pan data. *IEEE Trans. Geosci. Remote Sens.*, 45(10):3230–3239, 2007.

[4] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.*, 74(2):193–200, 2008.

[5] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.*, 1(4):313–317, 2004.

[6] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce. Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Trans. Geosci. Remote Sens.*, 45(10):3012–3021, 2007.

[7] H. A. Aly and G. Sharma. A regularized model-based optimization framework for pan-sharpening. *IEEE Trans. Image Process.*, 23(6):2596–2608, 2014.

[8] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Math. Program.*, 137(1):91–129, 2013.

[9] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rouge. A variational model for P+XS image fusion. *Int. J. Comput. Vis.*, 69(1):43–58, 2006.

[10] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Program.*, 146(1):459–494, 2014.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2011.

[12] J. F. Cai, R. H. Chan, L. Shen, and Z. Shen. Convergence analysis of tight framelet approach for missing data recovery. *Adv. Comput. Math.*, 31(1):87–113, 2009.

[13] J. F. Cai, R. H. Chan, and Z. Shen. A framelet-based image inpainting algorithm. *Appl. Comput. Harmon. Anal.*, 24(2):131–149, 2008.

[14] J. F. Cai, R. H. Chan, and Z. Shen. Simultaneous cartoon and texture inpainting. *Inverse Probl. Imaging*, 4(3):379, 2010.

[15] J. F. Cai, S. Osher, and Z. Shen. Split Bregman methods and frame based image restoration. *Multiscale Model. Simul.*, 8(2):337–369, 2010.

[16] W. Carper, T. Lillesand, and R. Kiefer. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.*, 56(4):459–467, 1990.

[17] D. Q. Chen and Y. Zhou. Wavelet frame based image restoration via combined sparsity and nonlocal prior of coefficients. *J. Sci. Comput.*, 66(1):196–224, 2016.

[18] J. Choi, K. Yu, and Y. Kim. A new adaptive component-substitution-based satellite image fusion by using partial replacement. *IEEE Trans. Geosci. Remote Sens.*, 49(1):295–309, 2010.

[19] I. Daubechies, B. Han, A. Ron, and Z. Shen. Framelets: MRA-based constructions of wavelet frames. *Appl. Comput. Harmon. Anal.*, 14(1):1–46, 2003.

[20] L. J. Deng, M. Y. Feng, and X. C. Tai. The fusion of panchromatic and multispectral remote sensing images via tensor-based sparse modeling and hyper-Laplacian prior. *Inf. Fusion*, 52:76–89, 2019.

[21] D. C. Dobson and F. Santosa. Recovery of blocky images from noisy and blurred data. *SIAM J. Appl. Math.*, 56(4):1181–1198, 1996.

[22] B. Dong and Y. Zhang. An efficient algorithm for ℓ_0 minimization in wavelet frame based image restoration. *J. Sci. Comput.*, 54(2):350–368, 2013.

[23] F. Fang, F. Li, C. Shen, and G. Zhang. A variational approach for pan-sharpening. *IEEE Trans. Image Process.*, 22(7):2822–2834, 2013.

[24] F. Fang, G. Zhang, F. Li, and C. Shen. Framelet based pan-sharpening via a variational method. *Neurocomputing*, 129:362–377, 2014.

[25] X. Y. Fu, Z. H. Lin, Y. Huang, and X. H. Ding. A variational pan-sharpening with local gradient constraints. In *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, pages 10265–10274, 2019.

[26] A. Garzelli. Pansharpening of multispectral images based on nonlocal parameter optimization. *IEEE Trans. Geosci. Remote Sens.*, 53(4):2096–2107, 2014.

[27] A. Garzelli, F. Nencini, and L. Capobianco. Optimal MMSE pan sharpening of very high resolution multispectral images. *IEEE Trans. Geosci. Remote Sens.*, 46(1):228–236, 2007.

[28] W. Guo, W. Wan, J. Liu, and H. Huang. Non-local blind hyperspectral image super-resolution via 4D sparse tensor factorization and low-rank. *Inverse Probl. Imaging*, 14(2), 2020.

[29] L. He, Y. Z. Rao, J. Li, J. Chanussot, A. Plaza, J. W. Zhu, and B. Li. Pansharpening via detail injection based convolutional neural networks. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 12(4):1188–1204, 2019.

[30] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia. A new pansharpening method based on spatial and spectral sparsity priors. *IEEE Trans. Image Process.*, 23(9):4160–4174, 2014.

[31] W. Huang, L. Xiao, Z. H. Wei, H. Y. Liu, and S. Z. Tang. A new pan-sharpening method with deep neural networks. *IEEE Geosci. Remote Sens. Lett.*, 12(5):1037–1041, 2015.

[32] Y. Y. Jiang, X. H. Ding, D. L. Zeng, Y. Huang, and J. Paisley. Pan-sharpening with a hyper-Laplacian penalty. In *Int. Conf. Comput. Vision (ICCV)*, pages 540–548, 2015.

[33] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, 2009.

[34] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. *Adv. in Neural Inf. Process. Syst. (NeurIPS)*, 22, 2009.

[35] P. Kwarteng and A. Chavez. Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens.*, 55(1):339–348, 1989.

[36] J. G. Liu. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. *Int. J. Remote Sens.*, 21(18):3461–3472, 2000.

[37] H. Lu, Y. Yang, S. Huang, W. Tu, and W. Wan. A unified pansharpening model based on band-adaptive gradient and detail correction. *IEEE Trans. Image Process.*, 31(pp):918–933, 2022.

[38] A. Marquina and S. J. Osher. Image super-resolution by TV-regularization and Bregman iteration. *J. Sci. Comput.*, 37(3):367–382, 2008.

[39] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa. Pansharpening by convolutional neural networks. *Remote Sens.*, 8(7):594, 2016.

[40] M. Moeller, T. Wittman, and A. L. Bertozzi. Variational wavelet pan-sharpening. *CAM Report*, pages 08–81, 2008.

[41] M. Nikolova. Local strong homogeneity of a regularized estimator. *SIAM J. Appl. Math.*, 61(2):633–658, 2000.

[42] X. Otazu, M. Gonzalezaudicana, O. Fors, and J. Nunez. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. *IEEE Trans. Geosci. Remote Sens.*, 43(10):2376–2385, 2005.

[43] J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng. Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing. *IEEE Trans. Image Process.*, 29:7889–7903, 2020.

[44] T. Pock and S. Sabach. Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems. *SIAM J. Imaging Sci.*, 9(4):1756–1787, 2016.

[45] Y. Quan, H. Ji, and Z. Shen. Data-driven multi-scale non-local wavelet frame construction and image recovery. *J. Sci. Comput.*, 63(2):307–329, 2015.

[46] R. Restaino, M. Dalla Mura, G. Vivone, and J. Chanussot. Context-adaptive pansharpening based on image segmentation. *IEEE Trans. Geosci. Remote Sens.*, 55(2):753–766, 2016.

[47] R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot. Fusion of multispectral and panchromatic images based on morphological operators. *IEEE Trans. Image Process.*, 25(6):2882–2895, 2016.

[48] A. Ron and Z. Shen. Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator. *J. Funct. Anal.*, 148(2):408–447, 1997.

[49] W. Z. Shao, H. S. Deng, and Z. H. Wei. Kullback–Leibler divergence based composite prior modeling for Bayesian super-resolution. *J. Sci. Comput.*, 60(1):60–78, 2014.

[50] M. J. Shensa. The discrete wavelet transform: wedding the a trous and Mallat algorithms. *IEEE Trans. Signal Process.*, 40(10):2464–2482, 1992.

[51] C. Song and Y. Wang. Nonlocal latent low rank sparse representation for single image super resolution via self-similarity learning. *Inverse Probl. Imaging*, 15(6):1347, 2021.

[52] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Trans. Geosci. Remote Sens.*, 46(5):1301–1312, 2008.

[53] G. Vivone. Robust band-dependent spatial-detail approaches for panchromatic sharpening. *IEEE Trans. Geosci. Remote Sens.*, 57(9):6421–6433, 2019.

[54] G. Vivone, S. Marano, and J. Chanussot. Pansharpening: Context-based generalized Laplacian pyramids by robust regression. *IEEE Trans. Geosci. Remote Sens.*, 58(9):6152–6167, 2020.

[55] G. Vivone, R. Restaino, and J. Chanussot. Full scale regression-based injection coefficients for panchromatic sharpening. *IEEE Trans. Image Process.*, 27(7):3418–3431, 2018.

[56] T. Wang, F. Fang, F. Li, and G. Zhang. High-quality Bayesian pansharpening. *IEEE Trans. Image Process.*, 28(1):227–239, 2018.

[57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, 2004.

[58] Z. C. Wu, T. Z. Huang, L. J. Deng, J. F. Hu, and G. Vivone. VO+Net: An adaptive approach using variational optimization and deep learning for panchromatic sharpening. *IEEE Trans. Geosci. Remote Sens.*, 60:1–16, 2022.

[59] J. L. Xiao, T. Z. Huang, L. J. Deng, Z. C. Wu, and G. Vivone. A new context-aware details injection fidelity with adaptive coefficients estimation for variational pansharpening. *IEEE Trans. Geosci. Remote Sens.*, 2022.

[60] J. F. Yang, X. Y. Fu, Y. W. Hu, Y. Huang, X. H. Ding, and J. Paisley. PanNet: A deep network architecture for pan-sharpening. In *Int. Conf. Comput. Vision (ICCV)*, pages 5449–5457, 2017.

[61] W. Yao, J. Shen, Z. Guo, J. Sun, and B. Wu. A total fractional-order variation model for image super-resolution and its SAV algorithm. *J. Sci. Comput.*, 82(3):1–18, 2020.

[62] H. Yin. PSCSC-Net: A deep coupled convolutional sparse coding network for pansharpening. *IEEE Trans. Geosci. Remote Sens.*, 60:1–16, 2022.

[63] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1:147–149, 1992.

[64] J. Zeng, W. Yin, and D. X. Zhou. Moreau envelope augmented Lagrangian method for nonconvex optimization with linear constraints. *J. Sci. Comput.*, 91(2):1–36, 2022.

[65] W. Zhang, J. Li, and Z. Hua. Attention-based multistage fusion network for remote sensing image pansharpening. *IEEE Trans. Geosci. Remote Sens.*, 60:1–16, 2022.

[66] X. Zhang and X. Zhang. A new proximal iterative hard thresholding method with extrapolation for ℓ_0 Minimization. *J. Sci. Comput.*, 79(2):809–826, 2019.