Journal of Computational Mathematics
Vol.28, No.3, 2010, 289–308.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1001-m1002

FRAMELET BASED DECONVOLUTION*

Jian-Feng Cai
Department of Mathematics, University of California, Los Angeles, CA 90095
Email: cai@math.ucla.edu
Zuowei Shen
Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543
Email: matzuows@nus.edu.sg

Abstract

In this paper, two framelet based deconvolution algorithms are proposed. The basic idea of framelet based approach is to convert the deconvolution problem to the problem of inpainting in a frame domain by constructing a framelet system with one of the masks being the given (discrete) convolution kernel via the unitary extension principle of [26], as introduced in [6–9] . The first algorithm unifies our previous works in high resolution image reconstruction and infra-red chopped and nodded image restoration, and the second one is a combination of our previous frame-based deconvolution algorithm and the iterative thresholding algorithm given by [14, 16]. The strong convergence of the algorithms in infinite dimensional settings is given by employing proximal forward-backward splitting (PFBS) method. Consequently, it unifies iterative algorithms of infinite and finite dimensional setting and simplifies the proof of the convergence of the algorithms of [6].

Mathematics subject classification: 65T60, 90C90, 94A08
Key words: Framelet, deconvolution, wavelet, tight frame, soft-thresholding

1. Introduction

The deconvolution is to solve $\mathbf{v}$ given by the following convolution equation:

$$\mathbf{c} = \mathbf{h} * \mathbf{v} + \epsilon, \tag{1.1}$$

where $\mathbf{h}$, $\mathbf{c}$ and ϵ are all in $\ell_2(\mathbb{Z})$, and $*$ is the convolution operator. The sequence $\mathbf{h}$ is the blurring kernel, and $\mathbf{c}$ is the observed signal. The sequence ϵ is the error term satisfying $\|\epsilon\|_{\ell_2(\mathbb{Z})} \leq \varepsilon$. The basic idea of our framelet based approach is to convert the deconvolution problem to the problem of inpainting in a frame domain by constructing a framelet system with one of the masks being the given (discrete) convolution kernel $\mathbf{h}$ via the unitary extension principle of [26].

This framelet based approach for deconvolution was originally proposed in [7–9] for high-resolution image reconstruction, by using frames derived from bi-orthogonal wavelets or the unitary extension principle of [26]. It was then extended to video still enhancement [11] and to infrared image restoration [3]. Recently, this framelet based approach is further generalized to inpainting in the image domain by [4,10]. The numerical simulation results in those papers show clearly that this framelet based approach is numerically efficient and easy to implement. The framelet deconvolution algorithm was first analyzed in [6]. This paper is to unify the framelet based approaches in the literature and to give a complete analysis of the unified approach.

* Received December 31, 2008 / Revised version received May 20, 2009 / Accepted June 6, 2009 / Published online February 1, 2010 /

There are several papers on solving inverse problems, in particular deconvolution problems, by using wavelet methods. The wavelet-vaguelette decomposition methods by [18, 21], the deconvolution in mirror wavelet bases by [24, 25], Galerkin-type methods to inverse problems using an appropriate basis by [1, 12], and the orthonormal wavelet method by [14, 18] are examples of wavelet approaches. The connections and differences of the above approaches and the framelet approaches of [3,7–9,11] are detailed in [6], and interested readers should consult [6] for the details. Finally, we also refer the reader to the recent work of [5] that gives a different approach on framelet based deconvolution by using linearized Bregman iteration.

In this paper, we propose and analyze two algorithms. The first one unifies the previous works in high resolution image reconstruction [7–9] and infra-red chopped and nodded image restoration [3], and the second one is a combination of the frame-based deconvolution algorithm [6] and the iterative thresholding algorithm of [14, 16]. The strong convergence of the algorithms is given in infinite dimensional setting by employing proximal forward-backward splitting (PFBS) method [13]. Consequently, it unifies iterative algorithms of infinite and finite dimensional setting, simplifies the proof of the convergence of the algorithms, and improves the minimization results that the limits are satisfied, of [6].

Since the focus of this paper is to give a theoretical analysis of algorithms, the numerical simulation is not the focus of this paper. The interested readers should refer to [3, 7–9, 11] for the numerical simulations for various applications.

The paper is organized as follows. In Section 2, we give a review of framelets. In Section 3, we give algorithms for the framelet deconvolution approach. In Sections 4.1 and 4.2, we analyze the strong convergence of the algorithms by the theory of proximal forward-backward splitting. The corresponding results in finite dimensional setting are illustrated in Section 5.

2. Framelets

In this section, we review some of basics of framelet that are needed for the current paper. For those who are familiar with the notion of framelet may skip this section.

A countable set $X \subset L_2(\mathbb{R})$ is called a *tight frame* of $L_2(\mathbb{R})$ if

$$\|f\|^2_{L_2(\mathbb{R})} = \sum_{g\in X} |\langle f, g\rangle|^2,$$

or equivalently

$$f = \sum_{g\in X} \langle f, g\rangle\, g,$$

holds for all $f \in L_2(\mathbb{R})$, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|_{L_2(\mathbb{R})}$ are the inner product and the norm in $L_2(\mathbb{R})$ respectively. For given $\Psi := \{\psi_1, \dots, \psi_r\} \subset L_2(\mathbb{R})$, the *affine system* is defined by

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

When $X(\Psi)$ forms a tight frame of $L_2(\mathbb{R})$, it is called a *tight wavelet frame*, and ψ_ℓ, $\ell = 1, \dots, r$, are called the *tight framelets*.

The *quasi-affine system* from level J is defined as

$$X^q_J(\Psi) = \left\{\psi^q_{\ell,j,k} :\ 1 \le \ell \le r; j, k \in \mathbb{Z}\right\} \quad \text{with} \quad \psi^q_{\ell,j,k} := \begin{cases} 2^{j/2}\psi_\ell(2^j \cdot -k), & j \ge J; \\ 2^{j-\frac{J}{2}}\psi_\ell(2^j \cdot -2^{j-J}k), & j < J. \end{cases}$$

It is clear that $X_J^q(\Psi)$ is a 2^{-J}-shift invariant system, and it is obtained by over sampling the affine system starting from level $J-1$ and downward to a 2^{-J}-shift invariant system. It was shown in [26, Theorem 5.5] that $X_J^q(\Psi)$ is a tight frame of $L_2(\mathbb{R})$ if and only if its affine counterpart $X(\Psi)$ is a tight frame of $L_2(\mathbb{R})$.

To construct a set of tight framelets, one can start from a refinable function $\phi \in L_2(\mathbb{R})$ satisfying a refinement equation

$$\phi(x) = 2\sum_{k\in\mathbb{Z}} \mathbf{h}_0[k]\phi(2x-k), \tag{2.1}$$

where $\mathbf{h}_0 \in \ell_2(\mathbb{Z})$ is called the refinement mask. Under mild assumptions on $\mathbf{h}_0$, hence on ϕ, a multiresolution analysis (MRA) $\{V_j\}_{j\in\mathbb{Z}}$ can be formed (see [17,23] for details). Let $\mathbf{h}_1,\ldots,\mathbf{h}_r$ be framelet masks that are in $\ell_2(\mathbb{Z})$, and define

$$\psi_\ell(x) = 2\sum_{k\in\mathbb{Z}} \mathbf{h}_\ell[k]\phi(2x-k), \quad \ell = 1,\ldots,r. \tag{2.2}$$

The unitary extension principle (UEP) in [26] says that $X(\Psi)$ generated by $\Psi := \{\psi_1,\ldots,\psi_r\}$ is a tight frame of $L_2(\mathbb{R})$ provided that for all $p\in\mathbb{Z}$

$$\sum_{\ell=0}^{r}\sum_{k\in\mathbb{Z}} \overline{\mathbf{h}_\ell[k]}\mathbf{h}_\ell[k-p] = \delta_{0,p}, \quad \text{and} \quad \sum_{\ell=0}^{r}\sum_{k\in\mathbb{Z}}(-1)^{k-p}\overline{\mathbf{h}_\ell[k]}\mathbf{h}_\ell[k-p] = 0, \tag{2.3}$$

where $\delta_{0,p}$ is 1 when $p=0$, and 0 otherwise. Define the infinite dimensional Toeplitz matrix

$$H_\ell = (H_\ell[j,k]) = (\mathbf{h}_\ell[j-k]), \quad 0\le \ell\le r,$$

then the first condition in UEP (2.3) in this matrix form is equivalent to

$$H_0^*H_0 + H_1^*H_1 + \ldots + H_r^*H_r = I. \tag{2.4}$$

The UEP condition (2.3) can be written in the Fourier domain as

$$\sum_{\ell=0}^{r}|\widehat{\mathbf{h}_\ell}(\omega)|^2 = 1 \quad \text{and} \quad \sum_{\ell=0}^{r}\widehat{\mathbf{h}_\ell}(\omega)\overline{\widehat{\mathbf{h}_\ell}(\omega+\pi)} = 0, \quad \text{a.e.} \quad \omega\in[-\pi,\pi], \tag{2.5}$$

where

$$\widehat{\mathbf{h}_\ell}(\omega) := \sum_{k\in\mathbb{Z}}\mathbf{h}_\ell[k]e^{-ik\omega}.$$

It is well known that the refinement mask always satisfies $\widehat{\mathbf{h}_0}(0)=1$, i.e., $\mathbf{h}_0$ is a low pass filter. The first condition of (2.5) says that all the framelet masks satisfy $\widehat{\mathbf{h}_\ell}(0)=0$, $\ell=1,\ldots,r$. In other words, for the tight framelet system, the framelet masks are high pass filters.

In the following, we introduce the framelet decomposition in the quasi-affine system $X_0^q(\Psi)$, i.e., the normal framelet decomposition introduced in [15] without down sampling step. For simplicity, we define $\psi_0 := \phi$. For a given function $f\in L_2(\mathbb{R})$, we start from $\{\langle f,\psi_{0,0,k}^q\rangle\}$, the coefficient of f with respect to ϕ at level 0. Then, it was shown in [6] by applying (2.2), one has the decomposition algorithm (without downsampling)

$$\begin{aligned}\{\langle f,\psi_{s,-1,k}^q\rangle\} &= \left\{\langle f, \sum_{l\in\mathbb{Z}}\mathbf{h}_s[k-l]\psi_{0,0,l}^q\rangle\right\}\\ &= \left\{\sum_{l\in\mathbb{Z}}\mathbf{h}_s[k-l]\langle f,\psi_{0,0,l}^q\rangle\right\} = H_s\{\langle f,\psi_{0,0,k}^q\rangle\}.\end{aligned} \tag{2.6}$$

Therefore,

$$\mathcal{A}_{0\to-1} := [H_0; H_1; \ldots; H_r]^t$$

is a framelet decomposition operator from level 0 to level -1. Let $\mathcal{A}_{0\to-1}^*$ be the adjoint of $\mathcal{A}_{0\to-1}$. It is the framelet reconstruction operator from level -1 to level 0. By the UEP (2.4), we have $\mathcal{A}_{0\to-1}^*\mathcal{A}_{0\to-1} = I$. It implies that the decomposition and reconstruction is perfect, see [6] for details.

For the multilevel framelet decomposition in the quasi-affine system $X_0^q(\Psi)$, we define the filter $\mathbf{h}_{\ell,j}$ for the filter $\mathbf{h}_\ell$ in level $j < 0$ by

$$\mathbf{h}_{\ell,j}[k] = \begin{cases} \mathbf{h}_\ell[2^{j+1}k], & k \in 2^{-j-1}\mathbb{Z} \\ 0, & k \in \mathbb{Z} \setminus 2^{-j-1}\mathbb{Z}. \end{cases} \tag{2.7}$$

Let $H_{\ell,j} := (H_{\ell,j}[k,l]) = (\mathbf{h}_{\ell,j}[k-l])$. Then

$$\{\langle f, \psi_{\ell,j,k}^q\rangle\} = H_{\ell,j}\{\langle f, \psi_{0,j+1,k}^q\rangle\}, \text{ and } H_\ell = H_{\ell,-1}.$$

Therefore, one can define the decomposition operator from level $j+1$ to level j by

$$\mathcal{A}_{j+1\to j} := [H_{0,j}; H_{1,j}; \ldots; H_{r,j}]^t\,.$$

Again, by (2.4), one can verify that $\mathcal{A}_{j+1\to j}^*\mathcal{A}_{j+1\to j} = I$. One can define the multilevel framelet decompsition operator from level J_0 to J, $J < J_0 \le 0$, by

$$\begin{aligned} \mathcal{A}_{J_0\to J} := \Bigg[\left(\prod_{j=J}^{J_0-1} H_{0,j}\right); \left(H_{1,J} \prod_{j=J+1}^{J_0-1} H_{0,j}\right); \ldots; \\ \left(H_{r,J} \prod_{j=J+1}^{J_0-1} H_{0,j}\right); \ldots;\ H_{1,J_0-1}; \ldots; H_{r,J_0-1}\Bigg]^t. \end{aligned} \tag{2.8}$$

It can be easily proved that $\mathcal{A}_{J_0\to J}^*\mathcal{A}_{J_0\to J} = I$. For simplicity, we denote

$$\mathcal{A}_J := \mathcal{A}_{0\to J}, \quad \text{and} \quad \mathcal{A} := \lim_{J\to-\infty} \mathcal{A}_J.$$

Then $\mathcal{A}$ is a linear operator from $\ell_2(\mathbb{Z})$ to $\mathscr{H}$, where

$$\mathscr{H} := \bigotimes_{\ell=1,j=1}^{r,\infty} \ell_2^{\ell,-j}(\mathbb{Z}). \tag{2.9}$$

It was proven in [6] that $\mathcal{A}^*\mathcal{A} = I$, i.e., the decomposition and reconstruction is perfect. For an arbitrary given sequence $\mathbf{v}$, there are infinitely many $\tilde{\mathbf{v}}$ such that $\mathbf{v} = \mathcal{A}^*\tilde{\mathbf{v}}$, since the tight frame system is a redundant system. The sequence $\mathcal{A}\mathbf{v}$ is one of them that has the minimal ℓ_2 norm among all frame coefficient sequence $\tilde{\mathbf{v}}$ satisfying $\mathbf{v} = \mathcal{A}^*\tilde{\mathbf{v}}$. The sequence $\mathcal{A}\mathbf{v}$ is called the canonical frame coefficients of $\mathbf{v}$.

Now we can define the thresholding operator $\mathcal{T}^p$ from $\mathscr{H}$ to $\mathscr{H}$, which has been extensively used for denoising in the wavelet literature. For any given real numbers λ and p, $1 \le p < 2$, let the thresholding operator be

$$t_\lambda^p(x) := \arg\min_{y\in\mathbb{R}} \left\{(x-y)^2 + \lambda|y|^p\right\}. \tag{2.10}$$

When $p=1$, $t_\lambda(x) := t^1_\lambda(x) = \mathrm{sgn}(x)\max(|x|-\frac{\lambda}{2},0)$ is the soft-thresholding function [20]; when $1<p<2$, the thresholding function is defined by the inverse of the function

$$\mathrm{F}^p_\lambda(x) := x + \frac{p\lambda}{2}\mathrm{sgn}(x)|x|^{p-1}.$$

For a given sequence

$$\mathbf{w} := \{w_{\ell,j,k}\}^r_{\ell=1,j<0,k\in\mathbb{Z}} \in \mathscr{H},$$

the denoising operator $\mathcal{T}^p$ which applies the thresholding operator $t^p_{\lambda_{\ell,j,k}}$ to $w_{\ell,j,k}$ with the thresholding parameters $\lambda_{\ell,j,k}$ is defined as:

$$\mathcal{T}^p\mathbf{w} = \left\{t^p_{\lambda_{\ell,j,k}}(w_{\ell,j,k})\right\}^r_{\ell=1,j<0,k\in\mathbb{Z}}. \tag{2.11}$$

From the definitions (2.10) and (2.11), it can be easily proved that

$$\mathcal{T}^p\mathbf{w} = \arg\min_{\mathbf{u}\in\mathscr{H}}\left\{\|\mathbf{w}-\mathbf{u}\|^2_{\mathscr{H}} + \sum_{\ell=1}^r\sum_{j<0,k\in\mathbb{Z}}\lambda_{\ell,j,k}|u_{\ell,j,k}|^p\right\}.$$

To transfer the deconvolution problem into an inpainting problem in the framelet domain $\mathscr{H}$, the basic idea is to construct a set of framelets such that one of the masks (or filters) is the convolution kernel $\mathbf{h}$. For a given filter (high pass or low pass) there are many ways to construct a system of tight framelets, so that one of the masks is the given convolution kernel by applying the unitary extension principle of [26]. Here we provide one of them derived from the corresponding construction of [19].

Assume that $\mathbf{h}$ is finitely supported, and satisfies

$$|\widehat{\mathbf{h}}(\omega)|^2 + |\widehat{\mathbf{h}}(\omega+\pi)|^2 \le 1.$$

When $\mathbf{h}$ is a low pass filter, i.e., $\widehat{\mathbf{h}}(0)=1$, we set $\mathbf{h}_0=\mathbf{h}$. Then, it is well known that the associated refinement function ϕ exists in the sense of distribution, and the Fourier transform of ϕ is given by $\widehat{\phi}(\omega) = \prod_{j=1}^\infty \widehat{\mathbf{h}_0}(\frac{\omega}{2^j})$. Furthermore, by Proposition A.1 in [6], ϕ is in $L_2(\mathbb{R})$ and satisfies the refinement equation $\phi(x) = 2\sum_{k\in\mathbb{Z}}\mathbf{h}_0[k]\phi(2x-k)$.

Let

$$\xi(\omega) := 1 - |\mathbf{h}_0(\omega)|^2 - |\mathbf{h}_0(\omega+\pi)|^2, \text{ and } \varsigma(\omega) := \frac{\sqrt{\xi}}{2},$$

where $\sqrt{\xi}$ is obtained via the Fejér-Riesz lemma. Define

$$\widehat{\mathbf{h}_1}(\omega) := e^{-i\omega}\overline{\widehat{\mathbf{h}_0}(\omega+\pi)},\quad \widehat{\mathbf{h}_2}(\omega) := \varsigma(\omega) + e^{-i\omega}\varsigma(-\omega),\quad \widehat{\mathbf{h}_3}(\omega) := e^{-i\omega}\overline{\widehat{\mathbf{h}_2}(\omega+\pi)}. \tag{2.12}$$

It was proven in [19] that $\mathbf{h}_\ell$ for $\ell=0,1,2,3$ satisfies the UEP condition (2.3). Therefore, we have constructed a set of tight framelets by $\mathbf{h}_\ell$, with $\mathbf{h}=\mathbf{h}_0$, and ϕ via (2.2). Furthermore, if $\mathbf{h}_0$ is symmetric, the filters $\mathbf{h}_\ell$, $\ell=1,2,3$, are either symmetric or anti-symmetric; see [19, Construction 4.4].

For the case that $\mathbf{h}$ is a high pass filter, i.e., $\widehat{\mathbf{h}}(0)=0$, we assume that $\widehat{\mathbf{h}}(\pi)=1$. Then, define $\widehat{\mathbf{h}_0}(\omega) := e^{-i\omega}\overline{\widehat{\mathbf{h}}(\omega+\pi)}$, and define $\mathbf{h}_\ell$, $\ell=1,2,3$, by (2.12). Hence, this set of filters satisfies UEP. By direct calculation, we see that $\widehat{\mathbf{h}_1} = -\widehat{\mathbf{h}}$. Therefore, we have constructed a set of framelets such that one of the filters is $\mathbf{h}$. Again, if $\mathbf{h}_0$ is symmetric, the filters $\mathbf{h}_\ell$, $\ell=1,2,3$, are either symmetric or anti-symmetric.

3. Problem Formulation and Algorithms

The framelet deconvolution approach works in the quasi-affine tight frame system $X_J^q(\Psi)$. As we mentioned in [6], without loss of generality, we may assume that the data set is given on $\mathbb{Z}$ (i.e., $J=0$). In fact, when the data set is given on $2^{-J}\mathbb{Z}$, we consider the function $f(2^{-J}\cdot)$ instead of f. The approximation power of a function f in the space V_J is the same as that of the function $f(2^{-J}\cdot)$ in space V_0. Therefore, we only consider the framelet deconvolution approach in the quasi-affine tight frame system $X_0^q(\Psi)$.

Suppose that for the given convolution kernel $\mathbf{h}$, we have constructed via UEP a set of tight framelets $\Psi=\{\psi_1,\ldots,\psi_r\}$ with refinement function ϕ such that $\mathbf{h}\in\{\mathbf{h}_0,\mathbf{h}_1,\ldots,\mathbf{h}_r\}$. Then the model (1.1) can be rewritten as

$$\mathbf{c}=\mathbf{h}_s*\mathbf{v}+\epsilon, \tag{3.1}$$

where s is the index such that $\mathbf{h}_s=\mathbf{h}$. Here we have generalized the convolution model in [6], where the convolution kernel is assumed to be a low pass filter, i.e., $s=0$. Our generalization is motivated by the recent work of [3], where a tight framelet approach for deconvolution with convolution kernel being a high pass filter was proposed for applications arising from infrared imaging in astronomy. To simplify our notations, we use $\|\cdot\|:=\|\cdot\|_{\ell_2(\mathbb{Z})}$ or $\|\cdot\|:=\|\cdot\|_{\mathscr{H}}$.

3.1. Modeling and Algorithms

As it was done in [6], we start with the simplest case that the data set contains no error, i.e., $\mathbf{c}=\mathbf{h}_s*\mathbf{v}$, and $\mathbf{c}=\{\langle f,\psi_{s,-1,k}\rangle\}$ for some $f\in L_2(\mathbb{R})$. Then, $\mathbf{v}=\{\langle f,\psi^q_{0,0,k}\rangle\}$ is a solution. Indeed, by the decomposition algorithm (2.6), we have

$$\mathbf{c}=\mathbf{h}_s*\mathbf{v}=H_s\mathbf{v}=\{\langle f,\psi_{s,-1,k}\rangle\}. \tag{3.2}$$

In this case, the deconvolution problem $\mathbf{h}_s*\mathbf{v}=\mathbf{c}$ becomes exactly the inpainting problem in the framelet domain. Let $\Theta:=\{(\ell,j,k)\mid \ell=1,\ldots,r;\ j\in\mathbb{Z},j<0;\ k\in\mathbb{Z}\}$ be the set of indices of $\mathscr{H}$. To mark the given data part in the framelet domain, we consider the cases $s\neq 0$ and $s=0$ respectively.

- Let $s\neq 0$, i.e., $\mathbf{h}_s$ is a high pass filter. By (3.2), $\mathbf{c}$ is just the tight framelet coefficients $\mathcal{A}\mathbf{v}$ on the set of indices

$$\Gamma_s:=\Big\{(\ell,j,k)\mid \ell=s;\ j=-1;\ k\in\mathbb{Z}\Big\}\subset\Theta. \tag{3.3}$$

  Define $\mathbf{k}_s$ by

$$\mathbf{k}_s[\ell,j,k]=\begin{cases}c_k, & \text{if } (\ell,j,k)\in\Gamma_s,\\ 0, & \text{otherwise.}\end{cases}$$

- For the case that $\mathbf{h}_s$ is a low pass filter, i.e., $s=0$, $\mathbf{c}$ is the coarse level coefficients on level -1. If one further decomposes it from level -1 to get the framelet coefficients $\mathcal{A}_{-1\to-\infty}\mathbf{c}$, by applying the definition of $\mathcal{A}$, we have that $\mathcal{A}_{-1\to-\infty}\mathbf{c}$ is the coefficients $\mathcal{A}\mathbf{v}$ on the set of indices

$$\Gamma_0:=\Big\{(\ell,j,k)\mid \ell=1,\ldots,r;\ j\in\mathbb{Z},j\leq -2;\ k\in\mathbb{Z}\Big\}\subset\Theta. \tag{3.4}$$

Define $\mathbf{k}_0$ by

$$\mathbf{k}_0 = [\mathcal{A}_{-1\to-\infty}\mathbf{c}; \underbrace{\mathbf{0};\ldots;\mathbf{0}}_{r\ \mathbf{0}'s.}]^t.$$

Let $\mathcal{P}_{\Gamma_s}$ be the projection operator from $\mathscr{H}$ to $\mathscr{H}$ defined as

$$\mathbf{w} := \mathcal{P}_{\Gamma_s}\mathbf{u}, \quad \text{with} \quad w_{\ell,j,k} := \begin{cases} u_{\ell,j,k}, & \text{if } (\ell,j,k) \in \Gamma_s, \\ 0, & \text{otherwise.} \end{cases}$$

Then from (3.1) and the definition of $\mathbf{k}_s$, we obtain

$$\mathcal{P}_{\Gamma_s}\mathcal{A}\mathbf{v} = \mathbf{k}_s. \tag{3.5}$$

We solve (3.1) in the framelet domain, i.e., in the space $\mathscr{H}$. Note that for this special case $\mathbf{k}_s$ is the known part of the coefficient $\mathcal{A}\mathbf{v}$ on Γ_s. We recover $\mathbf{v}$ from $\mathbf{c}$ by approximating the missing part of the framelet coefficients. This formulation transfers deconvolution to inpainting in the framelet domain. The iteration is motivated by the identity

$$\mathbf{v} = \mathcal{A}^*[\mathcal{P}_{\Gamma_s}\mathcal{A}\mathbf{v} + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathbf{v}]. \tag{3.6}$$

The key of the algorithm is approximating the missing framlet coefficients by those of previous iteration, which leads to the following algorithm

$$\mathbf{v}_{n+1} = \mathcal{A}^*[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathbf{v}_n]. \tag{3.7}$$

This is essentially a Landweber iteration for (3.1)

$$\mathbf{v}_{n+1} = \mathbf{v}_n + H_s^*(\mathbf{c} - H_s\mathbf{v}_n),$$

which is equivalent to

$$\mathbf{v}_{n+1} = H_s^*\mathbf{c} + \sum_{\ell\neq s} H_\ell^* H_\ell \mathbf{v}_n. \tag{3.8}$$

The iteration (3.8) is Algorithm 2.1 in [6]. When $s = 0$, it was proven in [6] that $\{\mathbf{v}_n\}$ converges to $\mathbf{v}$ that satisfies $\mathbf{h}_0 * \mathbf{v} = \mathbf{c}$. The proof can be straightforwardly extended to the case when $s \neq 0$, so we will omit the details here.

However, the observed data $\mathbf{c} = \mathbf{h} * \mathbf{v}$ (i.e., we still assume there is no noise in the data) may not be in the form of $\{\langle f, \psi^q_{s,-1,k}\rangle\}$. To solve the deconvolution problem in the framelet domain $\mathscr{H}$, we should find a sequence $\tilde{\mathbf{v}} \in \mathscr{H}$ such that it coincides with the known data $\mathbf{c}$, i.e., $\mathbf{k}_s$ on Γ_s. Then, let the solution be $\mathbf{v} := \mathcal{A}^*\tilde{\mathbf{v}}$. To make $\mathbf{v}$ a meaningful approximation solution of the deconvolution equation (3.1), we need $\|\mathbf{h}_s * \mathbf{v} - \mathbf{c}\|^2 = \|\mathcal{P}_{\Gamma_s}\mathcal{A}\mathbf{v} - \mathbf{k}_s\|^2$ to be small. Since $\mathcal{P}_{\Gamma_s}\tilde{\mathbf{v}} = \mathbf{k}_s$, this implies that $\|\mathcal{P}_{\Gamma_s}(\mathcal{A}\mathbf{v} - \tilde{\mathbf{v}})\|^2$ should be small. This leads to that $\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2$ should be small. Furthermore, we need $\tilde{\mathbf{v}}$ to be sparse to keep the edge of $\mathbf{v}$ sharp. Altogether, we need $\tilde{\mathbf{v}}$ to minimize the functional

$$\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2 + \sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p \tag{3.9}$$

with the constraint $\mathcal{P}_{\Gamma_s}\tilde{\mathbf{v}} = \mathbf{k_s}$. The term $\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2$ also penalizes the distance between $\tilde{\mathbf{v}}$ and the range of $\mathcal{A}$, i.e., the distance to the canonical frame coefficients of $\mathbf{v}$. Since the canonical coefficients of a framelet system links to the regularity of the underlying function (see, e.g., [2, 22]), and since some weighted norm of canonical framelet coefficients can be equivalent to some norm of underlying function, the cost functional of (3.9) penalizes the regularity of the underlying function. Altogether, the cost functional of (3.9) balances the fidelity, regularity and sparsity of the solution which are desired.

Finally, we consider the case that the given data set is bounded to have errors, i.e., $\epsilon \neq \mathbf{0}$ in (1.1). For this case, we clean the data first by using the same framelet system, i.e., we apply a noise removing scheme to $\mathbf{k_s}$. Then we solve the same minimization problem with the constraint on the clean data set. That is to find $\tilde{\mathbf{v}}$ that satisfies

$$\min_{\tilde{\mathbf{v}}\in\mathscr{C}} \{\|(I-\mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2 + \sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p\}, \tag{3.10}$$

where $\mathscr{C} = \{\tilde{\mathbf{v}}|\tilde{\mathbf{v}} \in \mathscr{H};\ \mathcal{P}_{\Gamma_s}\tilde{\mathbf{v}} = \mathcal{T}^p\mathbf{k}_s\}$. Here the thresholding parameters $\lambda_{\ell,j,k}$ for $(\ell,j,k)\in\Gamma_s$ are set according to the noise level of given data. In particular, $\lambda_{\ell,j,k} = 0$ for $(\ell,j,k)\in\Gamma_s$, when there is no noise which coincides with the discussions of the case $\mathbf{c} = \mathbf{h} * \mathbf{v}$. Hence, the minimization problem of (3.10) unifies the all cases discussed above and the key is to find the minimizer of (3.10) .

As we will show in this paper, the minimizer of (3.10) can be obtained by the following simple algorithm

Algorithm 3.1

1. Choose an initial approximation $\mathbf{v}_0$ (e.g., $\mathbf{v}_0 = \mathbf{c}$);
2. Iterate on n until convergence

$$\mathbf{v}_{n+1} = \mathcal{A}^*\mathcal{T}^p[\mathbf{k}_s + (I-\mathcal{P}_{\Gamma_s})\mathcal{A}\mathbf{v}_n]. \tag{3.11}$$

It is noted that this algorithm is obtained by modifying (3.7) via applying the thresholding operator $\mathcal{T}^p$ to (3.7) before it is synthesized.

3.2. Algorithms in [6] and [16]

The idea of [6] for deconvolution is to apply the thresholding operator $\mathcal{T}^p$ to each iteration of (3.8). There are several different schemes suggested, the main one is essentially

$$\mathbf{v}_{n+1} = \mathcal{A}^*\mathcal{T}^p\mathcal{A}(H_s^*\mathbf{c} + \sum_{\ell\neq s} H_\ell^* H_\ell \mathbf{v}_n). \tag{3.12}$$

This can be written as

$$\mathbf{v}_{n+1} = \mathcal{A}^*\mathcal{T}^p\mathcal{A}(\mathbf{v}_n + H_s^*(\mathbf{c} - H_s\mathbf{v}_n)). \tag{3.13}$$

Let

$$\tilde{\mathbf{v}}_n := \mathcal{T}^p\mathcal{A}\left(H_s^*\mathbf{c} + \sum_{\ell\neq s} H_\ell^* H_\ell \mathbf{v}_n\right).$$

Then, iterations (3.12) and (3.13) can be rewritten into

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p\Big(\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}_n + \mathcal{A}H_s^*(\mathbf{c} - H_s\mathcal{A}^*\tilde{\mathbf{v}}_n)\Big). \tag{3.14}$$

Furthermore, the pair of limit $(\mathbf{v}, \tilde{\mathbf{v}})$ satisfies $\mathbf{v} = \mathcal{A}^*\tilde{\mathbf{v}}$. The convergence of this algorithm for the finite dimensional case (see Section 5 for the detailed description of the finite dimensional case) was proven by [6].

The convergence of the general case was proven in [6] by introducing an accelerate factor β that modifies the iteration (3.12) to

$$\mathbf{v}_{n+1} = \mathcal{A}^*\mathcal{T}^p\mathcal{A}\left(H_s^*\beta\mathbf{c} + \sum_{\ell\neq s} H_\ell^* H_\ell \beta\mathbf{v}_n\right), \quad 0 < \beta \leq 1, \tag{3.15}$$

and the final solution is $\mathbf{s}^\beta = \mathbf{v}^\beta/\beta$ with $\mathbf{v}^\beta = \lim_{n\to\infty}\mathbf{v}_n$. The convergence of (3.15) was established for $0 < \beta < 1$ by [6]. Furthermore, it was proven in [6] that the limit satisfies some minimization properties in the following sense. Let $(\mathbf{v}, \tilde{\mathbf{v}})$ satisfies $\mathbf{v} = \mathcal{A}^*\tilde{\mathbf{v}}$, and $\mathbf{v}$ be the limit of (3.15). Define $\mathbf{s}^\beta = \mathbf{v}^\beta/\beta$ and $\tilde{\mathbf{s}}^\beta = \tilde{\mathbf{v}}^\beta/\beta$. Then, for an arbitrary pair $(\eta, \tilde{\eta})$ satisfying $\tilde{\eta} = \mathcal{A}\eta \in \ell_p$, the following inequality holds:

$$\begin{aligned}
&\|H_s(\mathbf{s}^\beta + \eta) - \mathbf{c}\|^2 + \sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \lambda_j |\tilde{s}^\beta_{\ell,j,k} + \tilde{\eta}_{\ell,j,k}|^p \\
\geq &\|H_s\mathbf{s}^\beta - \mathbf{c}\|^2 + \sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k} |\tilde{s}^\beta_{\ell,j,k}|^p - \delta,
\end{aligned} \tag{3.16}$$

where δ can be arbitrary close to 0 when β is arbitrary close to 1. Furthermore, it was shown in [6] that the above minimization property holds with $\beta = 1$ and $\delta = 0$ for the limit of (3.12) in the finite dimensional case.

Next, we make a short review of the approach in [16], which discusses a more general inverse problem. When it is restricted to the deconvolution problem, the corresponding iterative algorithm is to improve the framelet coefficients $\tilde{\mathbf{v}}$ by

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p(\tilde{\mathbf{v}}_n + \mathcal{A}H_s^*(\mathbf{c} - H_s\mathcal{A}^*\tilde{\mathbf{v}}_n)). \tag{3.17}$$

The iteration is performed entirely in the frame domain. At each step, it corrects $\tilde{\mathbf{v}}_n$ by $\mathcal{A}H_s^*(\mathbf{c} - H_s\mathcal{A}^*\tilde{\mathbf{v}}_n)$. On the other hand, the iteration (3.12) is performed in the image domain and it corrects $\mathbf{v}_n$ by $H_s^*(\mathbf{c} - H_s\mathbf{v}_n)$ at each step. These two algorithms are "mirror symmetric" performed in the two different domains.

The frame system $\mathcal{A}$ in (3.17) can be chosen to be independent of the convolution kernel H_s. Hence, this approach does not model the deconvolution by an inpainting problem in a transform domain. As it is proven in [16], the sequence $\tilde{\mathbf{v}}_n$ converges. Let $\tilde{\mathbf{v}}$ be its limit, then the solution $\mathbf{v}$ is defined to be $\mathcal{A}^*\tilde{\mathbf{v}}$. The limit is a minimizer of the following cost functional

$$\min_{\tilde{\mathbf{v}}\in\mathscr{H}}\left\{\|\mathbf{c} - H_s\mathcal{A}^*\tilde{\mathbf{v}}\|^2 + \sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p\right\}. \tag{3.18}$$

The framelet deconvolution algorithm given below, in some sense, is a combination and unification of the above two mentioned algorithms of (3.14) and (3.17).

Algorithm 3.2

1. Choose an initial approximation $\mathbf{v}_0$ (e.g., $\mathbf{v}_0 = \mathbf{c}$);
2. Iterate on n until convergence

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p\big(\tilde{\mathbf{v}}_n - \mu\delta(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}_n + \delta\mathcal{A}H_s^*(\mathbf{c} - H_s\mathcal{A}^*\tilde{\mathbf{v}}_n)\big). \tag{3.19}$$

When $\mu = 0$ and $\delta = 1$, it is (3.17), and when $\mu = 1$ and $\delta = 1$, it is identical to (3.14). We will show the convergence of Algorithm 3.2 in infinite dimensional setting, and the limit is a solution of

$$\min_{\tilde{\mathbf{v}}\in\mathscr{H}}\left\{\|H_s\mathcal{A}^*\tilde{\mathbf{v}} - \mathbf{c}\|^2 + \mu\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2 + \frac{1}{\delta}\sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p\right\}. \tag{3.20}$$

Here the first term penalizes the fidelity, and the last term penalizes the sparsity of $\tilde{\mathbf{v}}$, especially, when p is close to 1. The second term $\mu\|(I-\mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2$ penalizes the distance between $\tilde{\mathbf{v}}$ and the range of $\mathcal{A}$, i.e., the distance to the canonical frame coefficients of $\mathbf{v}$. The larger μ makes the frame coefficients of $\mathbf{v}$ closer to the range of $\mathcal{A}$, i.e. the frame coefficients of $\mathbf{v}$ is closer to the canonical frame coefficients. Since the canonical coefficients of a framelet system link to the regularity of the underlying function (see, e.g., [2, 22]), and since some weighted norm of the canonical framelet coefficients can be equivalent to some norm of the underlying function, the second term together with the third term penalize the regularity of the underlying function. Here, we also notice that, since $\tilde{\mathbf{v}}$ does not interpolate the data, unlike the same term in (3.10), the term $\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2$ does not penalize the fidelity. However, the first term does. Altogether, we conclude that the cost functional of (3.20) balances the fidelity, regularity and sparsity of the solution which is exactly what we want.

In infinite dimensional setting, we are able to prove the convergence of (3.15) with $\beta = 1$ as a special case. This improves the results in [6]. Moreover, since, for an arbitrary pair $(\eta, \tilde{\eta})$ satisfying $\tilde{\eta} = \mathcal{A}\eta \in \ell_p$, $\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\eta}\|^2 = 0$, it is clear that (3.20) implies (3.16). Hence, we give here a more compact minimization form than those in [6] even for the finite dimensional case.

Comparing (3.18) with (3.20), the cost functional (3.20) has the additional term $\mu\|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2$ to balance the distance of $\tilde{\mathbf{v}}$ to the range of $\mathcal{A}$. Hence, Algorithm 3.2 balances the regularity and sparsity requirements of the solution, while iteration (3.17), which can be viewed as a special case of Algorithm 3.2, pursues the full sparsity of the redundance.

Finally, we point out that our analysis can be extended to the other algorithm given in [6]. For example, Algorithm 2.3 in [6] can be written as

$$\mathbf{v}_{n+1} = H_s^*\beta\mathbf{c} + \sum_{\ell\neq s}H_\ell^*(\mathcal{A}^*\mathcal{T}^p\mathcal{A})(\beta H_\ell\mathbf{v}_n), \quad 0 < \beta \le 1, \tag{3.21}$$

and the final solution is $\mathbf{s}^\beta = \mathbf{v}^\beta/\beta$ with $\mathbf{v}^\beta = \lim_{n\to\infty}\mathbf{v}_n$. This algorithm applies a different denoising scheme from that of (3.15). The convergence of this iteration is proven for $0 < \beta < 1$ and the corresponding minimization property is discussed in [6]. The analysis here can be extended to prove the convergence of (3.21) with $\beta = 1$, and the minimization property can also be discussed similarly. In particular, when $\beta = 1$, the sequence

$$\tilde{\mathbf{v}}_n := \Big(\mathcal{T}^p\mathcal{A}H_0\mathbf{v}_n, \mathcal{T}^p\mathcal{A}H_1\mathbf{v}_n, \ldots, \mathcal{T}^p\mathcal{A}H_{s-1}\mathbf{v}_n, \mathbf{c}, \mathcal{T}^p\mathcal{A}H_{s+1}\mathbf{v}_n, \ldots, \ldots, \mathcal{T}^p\mathcal{A}H_r\mathbf{v}_n\Big) \tag{3.22}$$

in the $r+1$-tuple converges to a minimizer of the following cost functional

$$\min_{\tilde{\mathbf{v}}\in\mathscr{C}_0}\left\{\|(I-\mathcal{B}\mathcal{B}^*)\tilde{\mathbf{v}}\|^2+\sum_{\ell\neq s}\sum_{\ell'=1}^{r}\sum_{j<0,k\in\mathbb{Z}}\lambda_j|\tilde{v}^{\ell}_{\ell',j,k}|^p\right\}, \tag{3.23}$$

where $\mathscr{C}_0=\{\tilde{\mathbf{v}}|\tilde{\mathbf{v}}^s=\mathbf{c}\}$ with $\tilde{\mathbf{v}}^s$ being the s-th component in tuple (3.22), and

$$\mathcal{B}=[\mathcal{A}H_0;\mathcal{A}H_1;\ldots;\mathcal{A}H_{s-1};H_s;\mathcal{A}H_{r+1};\ldots;\mathcal{A}H_r]^t.$$

We omit the detailed discussion here, since the extension is routine.

4. Analysis of Algorithms

In this section, we give a convergence analysis of Algorithms 3.1 and 3.2. In particular, we prove that Algorithm 3.1 converges to a minimizer of (3.10), and Algorithm 3.2 converges to a minimizer of (3.20). The main tool used in our analysis is the convergence theory for the proximal forward-backward splitting (PFBS) iteration proposed in [13]. The aim of the proximal forward-backward splitting is to solve the minimization problem

$$\min_{x\in\mathscr{X}}\left\{F_1(x)+F_2(x)\right\}, \tag{4.1}$$

where $\mathscr{X}$ is a Hilbert space, and $F_1:\mathscr{X}\mapsto(-\infty,\infty]$ and $F_2:\mathscr{X}\mapsto(-\infty,\infty)$ are two proper lower semi-continuous convex functionals such that F_2 is differentiable on $\mathscr{X}$. Moreover, the gradient of F_2 satisfies

$$\|\nabla F_2(x)-\nabla F_2(y)\|_{\mathscr{X}}\leq\frac{1}{\alpha}\|x-y\|_{\mathscr{X}},\quad\forall x,y\in\mathscr{X}. \tag{4.2}$$

The proposed (simplified) iteration is

$$x_{n+1}=\mathrm{prox}_{\gamma F_1}(x_n-\gamma\nabla F_2(x_n)), \tag{4.3}$$

where $\mathrm{prox}_{\gamma F_1}$ is the proximity operator defined by

$$\mathrm{prox}_{\gamma F_1}(y)=\arg\min_{x\in\mathscr{X}}\left\{\frac{1}{2}\|y-x\|^2_{\mathscr{X}}+\gamma F_1(x)\right\}. \tag{4.4}$$

The following convergence theorem is a special case of Theorem 3.4 of [13] for the sequence generated by (4.3).

Theorem 4.1. *Suppose that the set of minimizers for* (4.1)*, denoted by* $\mathscr{G}$*, is not empty. Fixing* $x_0\in\mathscr{X}$*, let the sequence* $\{x_n\}_{n=1}^{\infty}$ *be defined by* (4.3)*. Then, for* $0<\gamma<2\alpha$*, where* α *is defined by* (4.2)*, we have:*

(a) x_n *converges weakly to a point* $x\in\mathscr{G}$*;*

(b) $\sum_{n=1}^{\infty}\|\nabla F_2(x_n)-\nabla F_2(x)\|^2_{\mathscr{X}}<\infty$*;*

(c) $\sum_{n=1}^{\infty}\|x_{n+1}-x_n\|^2_{\mathscr{X}}<\infty$*.*

4.1. Convergence Analysis for Algorithm 3.1

In this subsection, we study the convergence of Algorithm 3.1. Let $\mathscr{C} := \{\tilde{\mathbf{v}} | \tilde{\mathbf{v}} \in \mathscr{H};\ \mathcal{P}_{\Gamma_s}\tilde{\mathbf{v}} = \mathcal{T}^p \mathbf{k}_s\}$ be a subset in $\mathscr{H}$. It is obviously a closed nonempty convex set. The indicator function of $\mathscr{C}$ is defined by

$$\iota_{\mathscr{C}}(\tilde{\mathbf{v}}) = \begin{cases} 0, & \tilde{\mathbf{v}} \in \mathscr{C}, \\ \infty, & \tilde{\mathbf{v}} \notin \mathscr{C}. \end{cases}$$

We will prove that Algorithm 3.1 converges to a minimizer of

$$\min_{\tilde{\mathbf{v}} \in \mathscr{C}} \left\{ \|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2 + \sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k} |\tilde{v}_{\ell,j,k}|^p \right\} \tag{4.5}$$

by splitting the cost functional as

$$F_1(\tilde{\mathbf{v}}) = \iota_{\mathscr{C}}(\tilde{\mathbf{v}}) + \sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k} |\tilde{v}_{\ell,j,k}|^p, \text{ and } F_2(\tilde{\mathbf{v}}) = \|(I - \mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2. \tag{4.6}$$

Define a new sequence in the space $\mathscr{H}$ as

$$\tilde{\mathbf{v}}_n = \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathbf{v}_n]. \tag{4.7}$$

Then we have

$$\mathbf{v}_{n+1} = \mathcal{A}^* \tilde{\mathbf{v}}_n,$$

and the iteration (3.11) can be rewritten as

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}_n]. \tag{4.8}$$

Lemma 4.1. *The sequence* $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$ *defined by* (4.8) *converges if and only if the sequence* $\{\mathbf{v}_n\}_{n=1}^{\infty}$ *defined by* (3.11) *does.*

Proof. By Proposition 1.2 in [6], the operator $\mathcal{A}$ satisfies $\mathcal{A}^*\mathcal{A} = I$. Therefore, $\mathcal{A}^*$ is a continuous linear operator, hence $\{\mathbf{v}_n\}_{n=1}^{\infty}$ converges if $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$ does.

On the other hand, by Proposition 3.2 in [6] the operator $\mathcal{T}^p$ is continuous. By the definition, the operator $\mathcal{P}_{\Gamma_s}$ is a continuous linear operator too. Therefore, the map from $\mathbf{v}_n$ to $\tilde{\mathbf{v}}_n$ in (4.7) is also continuous. Hence, $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$ converges if $\{\mathbf{v}_n\}_{n=1}^{\infty}$ does. □

Therefore, to show that $\{\mathbf{v}_n\}_{n=1}^{\infty}$ converges, we only need to show that $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$ converges, which will be done in the following. We first transform the iteration (3.11) into a proximal forward-backward splitting iteration for (4.5) by the splitting (4.6). Then, by applying Theorem 4.1, we obtain the weak convergence for the sequence $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$. Finally, we prove the strong convergence by a lemma in [14].

Lemma 4.2. *The iteration* (4.8) *(generated by* (3.11) *in Algorithm 3.1 is the same as the proximal forward-backward splitting iteration* (4.3) *with* F_1 *and* F_2 *being defined in* (4.6) *and* $\gamma = \frac{1}{2}$.

Proof. It is clear that

$$\frac{1}{2}\nabla F_2(\tilde{\mathbf{v}}) = \tilde{\mathbf{v}} - \mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}. \tag{4.9}$$

Comparing (4.8) with (4.9) and (4.3) where $\gamma = \frac{1}{2}$, we see that if we can prove that

$$\operatorname{prox}_{\frac{1}{2}F_1}(\mathbf{u}) = \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathbf{u}], \tag{4.10}$$

then we have done. We verify (4.10) by considering the definition of the proximity operator in the following

$$\operatorname{prox}_{\frac{1}{2}F_1}(\mathbf{u}) = \arg\min_{\mathbf{w}\in\mathscr{H}} \Big\{\frac{1}{2}\|\mathbf{w}-\mathbf{u}\|^2 + \frac{1}{2}\iota_{\mathscr{C}}(\mathbf{w}) + \frac{1}{2}\sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \lambda_{\ell,j,k}|w_{\ell,j,k}|^p\Big\}. \tag{4.11}$$

Note that

$$\iota_{\mathscr{C}}(\mathbf{w}) = \sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}} \iota_{\ell,j,k}(w_{\ell,j,k}),$$

where

$$\iota_{\ell,j,k}(w_{\ell,j,k}) = \begin{cases} 0, & \text{if } (\ell,j,k)\notin\Gamma_s,\\ 0, & \text{if } (\ell,j,k)\in\Gamma_s \text{ and } w_{\ell,j,k} = t^p_{\lambda_{\ell,j,k}}(\mathbf{k}_s[\ell,j,k]),\\ \infty, & \text{if } (\ell,j,k)\in\Gamma_s \text{ and } w_{\ell,j,k} \neq t^p_{\lambda_{\ell,j,k}}(\mathbf{k}_s[\ell,j,k]),\end{cases}$$

with $t^p_{\lambda_{\ell,j,k}}$ being defined in (2.10). By the definition (2.10), we have

$$\forall(\ell,j,k)\in\Gamma_s,\quad t^p_{\lambda_{\ell,j,k}}(\mathbf{k}_s[\ell,j,k]) = \arg\min_{w_{\ell,j,k}\in\mathbb{R}} \left\{\frac{1}{2}(w_{\ell,j,k}-u_{\ell,j,k})^2 + \frac{1}{2}\iota_{\ell,j,k}(w_{\ell,j,k}) + \frac{1}{2}|w_{\ell,j,k}|^p\right\} \tag{4.12}$$

and

$$\forall(\ell,j,k)\notin\Gamma_s,\quad t^p_{\lambda_{\ell,j,k}}(u_{\ell,j,k}) = \arg\min_{w_{\ell,j,k}\in\mathbb{R}} \left\{\frac{1}{2}(w_{\ell,j,k}-u_{\ell,j,k})^2 + \frac{1}{2}|w_{\ell,j,k}|^p\right\}. \tag{4.13}$$

Denote $\mathbf{x}$ by

$$x_{\ell,j,k} := \begin{cases} t^p_{\lambda_{\ell,j,k}}(\mathbf{k}_s[\ell,j,k]), & \text{for } (\ell,j,k)\in\Gamma_s,\\ t^p_{\lambda_{\ell,j,k}}(u_{\ell,j,k}), & \text{for } (\ell,j,k)\notin\Gamma_s.\end{cases}$$

Note that $\mathbf{k}_s\in\mathscr{H}$ and $\mathbf{u}\in\mathscr{H}$ imply $\mathbf{x}\in\mathscr{H}$. This, together with (4.12) and (4.13), leads to $\mathbf{x} = \operatorname{prox}_{\frac{1}{2}F_1}(\mathbf{u})$. On the other hand, by the definition of $\mathcal{T}^p$ in (2.11), we have

$$\mathbf{x} = \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathbf{u}].$$

Therefore, we have proved (4.10). □

By applying Theorem 4.1, we conclude that the sequence $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$ weakly converges to a minimizer of (4.5) if it not empty. Next lemma says that the minimizer of (4.5) is not empty.

Lemma 4.3. *Let $\tilde{\mathbf{v}}_n$ be the sequence defined by the iteration* (4.8). *Then, there exists a point $\tilde{\mathbf{v}}$ which is a minimizer of* (4.5) *such that*

(a) $\tilde{\mathbf{v}}_n \rightharpoonup \tilde{\mathbf{v}}$, *where* $\rightharpoonup$ *denotes weak convergence;*

(b) $\sum_{n=1}^{\infty}\|\nabla F_2(\tilde{\mathbf{v}}_n) - \nabla F_2(\tilde{\mathbf{v}})\|^2 < +\infty$;

(c) $\sum_{n=1}^{\infty}\|\tilde{\mathbf{v}}_{n+1} - \tilde{\mathbf{v}}_n\|^2 < +\infty$.

Proof. It is obvious that both functionals F_1 and F_2 are proper, semi-continuous and convex functionals, and F_2 is differentiable. By (4.9), for any vectors $\mathbf{f}, \mathbf{g} \in \mathscr{H}$, we have

$$\begin{aligned}\frac{1}{2}\|\nabla F_2(\mathbf{f}) - \nabla F_2(\mathbf{g})\| =&\|(I - \mathcal{A}\mathcal{A}^*)(\mathbf{f} - \mathbf{g})\| \\ \le&\|I - \mathcal{A}\mathcal{A}^*\|\|\mathbf{f} - \mathbf{g}\|.\end{aligned}$$

Since $\|I - \mathcal{A}\mathcal{A}^*\| = 1$, ∇F_2 is Lipschitz continuous with Lipshitz constant $1/\alpha = 2$. Therefore, $\alpha = \frac{1}{2}$. We conclude from Theorem 4.1 that the lemma follows if there exists a minimizer for (4.5).

The existence of the minimizers for (4.5) follows from the coercivity of F_1 hence $F_1 + F_2$. Since ℓ_p norm is always greater than ℓ_2 norm for $p \in [1, 2)$ and $\lambda = \inf_{(\ell,j,k)} \lambda_{\ell,j,k} > 0$, we have

$$F_1(\tilde{\mathbf{v}}) \ge \lambda \sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} |\tilde{v}_{\ell,j,k}|^p \ge \lambda (\sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} |\tilde{v}_{\ell,j,k}|^2)^{p/2} = \lambda \|\tilde{\mathbf{v}}\|^p.$$

Therefore, as $\|\tilde{\mathbf{v}}\| \to \infty$, $F_1(\tilde{\mathbf{v}}) \to \infty$. It means that F_1 is coercive. □

Next, we show that the convergence is in the strong topology, i.e., in the norm. For this, we need the following lemma, which follows immediately from Lemma 3.18 in [14].

Lemma 4.4. *Let* $\mathbf{a} \in \mathscr{H}$ *be a given vector and* $\{\mathbf{u}_n\}_{n=1}^{\infty}$ *be a sequence in* $\mathscr{H}$. *Assume that* $\mathbf{u}_n \rightharpoonup \mathbf{0}$ ($\rightharpoonup$ *denoted for the weak convergence), and* $\|\mathcal{T}^p(\mathbf{a} + \mathbf{u}_n) - \mathcal{T}^p(\mathbf{a}) - \mathbf{u}_n\| \to 0$, ($\to$ *denoted for the strong convergence). Then* $\|\mathbf{u}_n\| \to 0$, *i.e.,* $\mathbf{u}_n$ *converges to* $\mathbf{0}$ *strongly.*

With this lemma, we have the following desired result:

Theorem 4.2. *Iterations* (3.11) *and* (4.8) *converge in norm. Further, the limit pair* $(\mathbf{v}, \tilde{\mathbf{v}})$ *satisfies* $\mathbf{v} = \mathcal{A}^*\tilde{\mathbf{v}}$ *and* $\tilde{\mathbf{v}}$ *is a minimizer of the cost functional* (3.10).

Proof. It only remains to show the strong convergence of $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$. Let $\mathbf{u}_n = \tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}$, then $\mathbf{u}_n \rightharpoonup \mathbf{0}$ and $\tilde{\mathbf{v}}$ is a minimizer of (4.5) by Lemma 4.3. Set $\mathbf{a} = \mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}$. By Lemma 4.4, it is only left to show that

$$\|\mathcal{T}^p(\mathbf{a} + \mathbf{u}_n) - \mathcal{T}^p(\mathbf{a}) - \mathbf{u}_n\| \to 0. \tag{4.14}$$

Since $\tilde{\mathbf{v}}$ is a minimizer of (3.10), $\mathcal{T}^p(\mathbf{a}) = \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}] = \tilde{\mathbf{v}}$. Note that by Proposition 3.2 in [6] $\mathcal{T}^p$ is non-expansive. We get

$$\begin{aligned}&\|\mathcal{T}^p(\mathbf{a} + \mathbf{u}_n) - \mathcal{T}^p(\mathbf{a}) - \mathbf{u}_n\| \\ =&\|\mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}} + \tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}] - \tilde{\mathbf{v}}_n\| \\ =&\|\mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}} + \tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}] - \mathcal{T}^p[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}_{n-1}]\| \\ \le&\|[\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}} + \tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}] - [\mathbf{k}_s + (I - \mathcal{P}_{\Gamma_s})\mathcal{A}\mathcal{A}^*\tilde{\mathbf{v}}_{n-1}]\| \\ =&\|(I - \mathcal{P}_{\Gamma_s})(I - \mathcal{A}\mathcal{A}^*)(\tilde{\mathbf{v}}_{n-1} - \tilde{\mathbf{v}}) + (I - \mathcal{P}_{\Gamma_s})(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}_{n-1}) + \mathcal{P}_{\Gamma_s}(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}})\| \\ \le&\|(I - \mathcal{A}\mathcal{A}^*)(\tilde{\mathbf{v}}_{n-1} - \tilde{\mathbf{v}})\| + \|\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}_{n-1}\| + \|\mathcal{P}_{\Gamma_s}(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}})\| \\ =&\frac{1}{2}\|\nabla F_2(\tilde{\mathbf{v}}_{n-1}) - \nabla F_2(\tilde{\mathbf{v}})\| + \|\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}_{n-1}\| + \|\mathcal{P}_{\Gamma_s}(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}})\|.\end{aligned} \tag{4.15}$$

By (b) and (c) in Lemma 4.3, we get $\|\nabla F_2(\tilde{\mathbf{v}}) - \nabla F_2(\tilde{\mathbf{v}}_{n-1})\| \to 0$ and $\|\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}_{n-1}\| \to 0$. Since $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{v}}_n$ are all in $\mathscr{C}$, we have $\mathcal{P}_{\Gamma_s}(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}) = \mathbf{0}$, hence $\|\mathcal{P}_{\Gamma_s}(\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}})\| = 0$. Combining all together, (4.15) implies (4.14). □

4.2. Convergence Analysis for the Algorithm 3.2

In this subsection, we prove that the iteration (3.19) in Algorithm 3.2 converges to a minimizer of

$$\min_{\tilde{\mathbf{v}}\in\mathscr{H}}\left\{\|\mathbf{c}-H_s\mathcal{A}^*\tilde{\mathbf{v}}\|^2+\mu\|(I-\mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2+\frac{1}{\delta}\sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p\right\}. \tag{4.16}$$

Again, we first transform the iteration (3.12) into a proximal forward-backward splitting iteration for the cost functional (4.16). Then by applying Theorem 4.1, we obtain the weak convergence for the sequence $\{\tilde{\mathbf{v}}_n\}_{n=1}^\infty$. Finally, we prove the strong convergence by applying Lemma 4.4. For this, we split the cost functional of (4.16) into

$$F_1(\tilde{\mathbf{v}})=\frac{1}{\delta}\sum_{\ell=1}^{r}\sum_{j<0,k\in\mathbb{Z}}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p,\quad F_2(\tilde{\mathbf{v}})=\|\mathbf{c}-H_s\mathcal{A}^*\tilde{\mathbf{v}}\|^2+\mu\|(I-\mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}\|^2, \tag{4.17}$$

and prove the following:

Lemma 4.5. *Iteration* (3.19) *is exactly the same as the proximal forward-backward splitting iteration* (4.3) *with* F_1 *and* F_2 *being given in* (4.17) *and* $\gamma=\frac{\delta}{2}$.

Proof. By the definitions of $\mathcal{T}^p$ and F_1, we have $\mathcal{T}^p=\text{prox}_{\frac{\delta}{2}F_1}$. Comparing (3.19) with (4.3), we only need to show that

$$\frac{1}{2}\nabla F_2(\tilde{\mathbf{v}})=\mu(I-\mathcal{A}\mathcal{A}^*)\tilde{\mathbf{v}}_n-\mathcal{A}H_s^*(\mathbf{c}-H_s\mathcal{A}^*\tilde{\mathbf{v}}_n). \tag{4.18}$$

This follows directly by a straightforward computation. □

With this, we conclude that the sequence $\{\tilde{\mathbf{v}}_n\}_{n=1}^\infty$ weakly converges to a minimizer of (4.16) by applying Theorem 4.1 as stated in the next lemma:

Lemma 4.6. *Let* $\tilde{\mathbf{v}}_n$ *be a sequence defined by* (3.19). *Assume that* $0<\delta<\frac{2}{\max\{1,\mu\}}$. *Then, there exists a point* $\tilde{\mathbf{v}}$ *which is a minimizer of* (4.16) *such that*

(a) $\tilde{\mathbf{v}}_n\rightharpoonup\tilde{\mathbf{v}}$, *where* $\rightharpoonup$ *denotes weak convergence;*

(b) $\sum_{n=1}^\infty\|\nabla F_2(\tilde{\mathbf{v}}_n)-\nabla F_2(\tilde{\mathbf{v}})\|^2<+\infty$;

(c) $\sum_{n=1}^\infty\|\tilde{\mathbf{v}}_{n+1}-\tilde{\mathbf{v}}_n\|^2<+\infty$.

Proof. It is obvious that both functionals F_1 and F_2 are proper, semi-continuous and convex functionals, and F_2 is differentiable. By (4.18), for any vectors $\mathbf{f},\mathbf{g}\in\mathscr{H}$, we have

$$\begin{aligned}\frac{1}{2}\|\nabla F_2(\mathbf{f})-\nabla F_2(\mathbf{g})\|=&\|\big(\mu(I-\mathcal{A}\mathcal{A}^*)+\mathcal{A}H_s^*H_s\mathcal{A}^*\big)(\mathbf{f}-\mathbf{g})\|\\ \le&\|\mu(I-\mathcal{A}\mathcal{A}^*)+\mathcal{A}H_s^*H_s\mathcal{A}^*\|\|\mathbf{f}-\mathbf{g}\|.\end{aligned}$$

Hence, ∇F_2 is Lipschitz continuous with Lipshitz constant $1/\alpha=2\|\mu(I-\mathcal{A}\mathcal{A}^*)+\mathcal{A}H_s^*H_s\mathcal{A}^*\|$, which will be estimated in the following. For any vector $\mathbf{f}$,

$$\begin{aligned}&\|(\mu(I-\mathcal{A}\mathcal{A}^*)+\mathcal{A}H_s^*H_s\mathcal{A}^*)\mathbf{f}\|^2\\ =&\|\mu(I-\mathcal{A}\mathcal{A}^*)\mathbf{f}+\mathcal{A}H_s^*H_s\mathcal{A}^*\mathbf{f}\|^2=\|\mu(I-\mathcal{A}\mathcal{A}^*)\mathbf{f}\|^2+\|\mathcal{A}H_s^*H_s\mathcal{A}^*\mathbf{f}\|^2.\end{aligned}$$

The last equality follows from that the range of $I - \mathcal{A}\mathcal{A}^*$ is orthogonal to that of $\mathcal{A}$. This leads to

$$\begin{aligned}&\|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f}\|^2 + \|\mathcal{A}H_s^*H_s\mathcal{A}^*\mathbf{f}\|^2\\ \le&\|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f}\|^2 + \|\mathcal{A}\|^2\|H_s^*H_s\|^2\|\mathcal{A}^*\mathbf{f}\|^2 \le \|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f}\|^2 + \|\mathcal{A}\mathcal{A}^*\mathbf{f}\|^2.\end{aligned}$$

The last inequality follows from the facts that $\|\mathcal{A}\|^2\|H_s^*H_s\|^2 \le 1$ and $\|\mathcal{A}\mathbf{x}\|^2 = \|\mathbf{x}\|^2$ for any $\mathbf{x}$. Hence,

$$\|(\mu(I - \mathcal{A}\mathcal{A}^*) + \mathcal{A}H_s^*H_s\mathcal{A}^*)\mathbf{f}\|^2 \le \|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f} + \mathcal{A}\mathcal{A}^*\mathbf{f}\|^2,$$

because the range of $I - \mathcal{A}\mathcal{A}^*$ is orthogonal to that of $\mathcal{A}$ again. Furthermore, when $0 \le \mu \le 1$

$$\begin{aligned}\|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f} + \mathcal{A}\mathcal{A}^*\mathbf{f}\| =&\|\mu\mathbf{f} + (1-\mu)\mathcal{A}\mathcal{A}^*\mathbf{f}\|\\ \le&\mu\|\mathbf{f}\| + (1-\mu)\|\mathcal{A}\mathcal{A}^*\mathbf{f}\| \le \|\mathbf{f}\|,\end{aligned}$$

and when $\mu > 1$,

$$\begin{aligned}\|\mu(I - \mathcal{A}\mathcal{A}^*)\mathbf{f} + \mathcal{A}\mathcal{A}^*\mathbf{f}\| =& \|\mathbf{f} + (\mu-1)(I - \mathcal{A}\mathcal{A}^*)\mathbf{f}\|\\ \le& \|\mathbf{f}\| + (\mu-1)\|(I - \mathcal{A}\mathcal{A}^*)\mathbf{f}\| \le \mu\|\mathbf{f}\|.\end{aligned}$$

Therefore,

$$\|(\mu(I - \mathcal{A}\mathcal{A}^*) + \mathcal{A}H_s^*H_s\mathcal{A}^*)\mathbf{f}\|^2 \le (\max\{\mu, 1\})^2\|\mathbf{f}\|^2,$$

Consequently,

$$\|\mu(I - \mathcal{A}\mathcal{A}^*) + \mathcal{A}H_s^*H_s\mathcal{A}^*\| \le \max\{\mu, 1\}.$$

This implies that $\alpha = \frac{1}{2\max\{\mu,1\}}$. The conclusion follows once the existence of the minimizer for (4.16) is established which is achieved by showing the coercivity of F_1 hence $F_1 + F_2$. Since ℓ_p norm is always greater than ℓ_2 norm for $p \in [1, 2)$ and $\lambda = \inf_{(\ell,j,k)} \lambda_{\ell,j,k} > 0$, we have

$$F_1(\mathbf{f}) \ge \lambda \sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} |f_{\ell,j,k}|^p \ge \lambda(\sum_{\ell=1}^{r} \sum_{j<0,k\in\mathbb{Z}} |f_{\ell,j,k}|^2)^{p/2} = \lambda\|\mathbf{f}\|^p.$$

Therefore, as $\|\mathbf{f}\| \to \infty$, $F_1(\tilde{\mathbf{v}}) \to \infty$. It means that F_1 is coercive. □

Finally, by a similar proof of Theorem 4.2, which we omit, we get the strong convergence of the sequence $\{\tilde{\mathbf{v}}_n\}_{n=1}^{\infty}$.

Theorem 4.3. *Assume that* $0 < \delta < \frac{2}{\max\{1,\mu\}}$. *Then iteration* (3.19) *in Algorithm 3.2 converges in norm. Further, the limit* $\tilde{\mathbf{v}}$ *is a minimizer of* (4.16)*.*

5. Algorithms for Finite Dimensional Data

In applications, data sets are always given as finite dimensional vectors. In this section, we modify Algorithms 3.1 and 3.2 to suit this need. The key issue is to impose a proper boundary condition to convert the finite dimensional convolution $\mathbf{h} \circledast \mathbf{f}$ with kernel $\mathbf{h}$ to a matrix-vector multiplication $H\mathbf{f}$. There are several ways to generate convolution matrices from given convolution kernels. For example, when periodic boundary conditions are used as discussed in [6], the matrix H becomes a circulant matrix

$$H[l, k] = \mathbf{h}[(l-k) \mod N_0], \qquad 0 \le l, k < N_0.$$

Another example of boundary condition is the half-point symmetric extension condition. In this case, the matrix H is a Toeplitz-plus-Hankel matrix, and is given by

$$H[l,k] = \mathbf{h}_\ell(l-k) + \mathbf{h}(l+k-1) + \mathbf{h}(-1-(2N_0-l-k)), \qquad 0 \le l,k < N_0, \tag{5.1}$$

provided that the length of $\mathbf{h}$ is smaller than $2N_0$. Other possible boundary extensions include, for example, the whole-point symmetric extension discussed in [8].

Let the matrix H_ℓ be the finite convolution matrix with kernel $\mathbf{h}_\ell$, and $H_{\ell,j}$ be the finite convolution matrix with kernel $\mathbf{h}_{\ell,j}$ defined in (2.7). We do not specify any boundary condition here. We only assume that proper boundary conditions are incorporated such that the convolution matrices satisfy

$$\sum_{\ell=0}^{r} H_{\ell,j}^* H_{\ell,j} = I, \qquad \forall j < 0. \tag{5.2}$$

One can easily verify (5.2) for given boundary conditions with proper filter conditions, e.g., filters should be either symmetric or anti-symmetric for symmetric boundary conditions (see, e.g., [8]).

With the convolution matrices, we can define the tight framelet decomposition operator in $\mathbb{R}^{N_0}$. The decomposition operator (without down sampling) from level $j+1$ to level j is defined by

$$A_{j+1\to j} := [H_{0,j}; H_{1,j}; \ldots; H_{r,j}]^t. \tag{5.3}$$

By the assumption (5.2), we have $A_{j+1\to j}^* A_{j+1\to j} = I$. Similarly, the multilevel framelet decomposition operator from level J_0 to J, $J < J_0 \le 0$ is defined by

$$\begin{aligned} A_{J_0\to J} :=\Bigg[&\left(\prod_{j=J}^{J_0-1} H_{0,j}\right); \left(H_{1,J}\prod_{j=J+1}^{J_0-1} H_{0,j}\right); \ldots; \left(H_{r,J}\prod_{j=J+1}^{J_0-1} H_{0,j}\right); \\ &\ldots;\ H_{1,J_0-1}; \ldots; H_{r,J_0-1}\Bigg]^t. \end{aligned} \tag{5.4}$$

By (5.2), we obtain $A_{J_0\to J}^* A_{J_0\to J} = I$. Again, for simplicity, we denote $A_J := A_{0\to J}$. Therefore, $A_J^* A_J = I$, i.e., the decomposition and reconstruction is perfect. Hence, the rows of the matrix A_J form a tight frame in $\mathbb{R}^{N_0}$. For a given vector $\mathbf{w} \in \mathbb{R}^{(1+r|J|)N_0}$, we organize it according to the blocks of A_J and denote it as

$$\mathbf{w} := \left[\{w_{0,J,k}\}_{k=0}^{N_0-1}; \{w_{\ell,j,k}\}_{\ell=1,j=J,k=0}^{r,\ -1,\ N_0-1}\right]^t,$$

and the tresholding operator $\mathcal{T}^p$ applying to $\mathbf{w}$ is defined as:

$$\mathcal{T}^p\mathbf{w} = \left[\{t^p_{\lambda_{0,J,k}}(w_{0,J,k})\}_{k=0}^{N_0-1}; \{t^p_{\lambda_{\ell,j,k}}(w_{\ell,j,k})\}_{\ell=1,j=J,k=0}^{r,\ -1,\ N_0-1}\right]^t, \tag{5.5}$$

where $t^p_\lambda(x)$ is defined in (2.10). With this setting, (1.1) becomes

$$H_s\mathbf{v} = \mathbf{b} + \epsilon := \mathbf{c}. \tag{5.6}$$

As before, we consider the case $s = 0$ and $s \neq 0$ respectively.

- Let $s \neq 0$, i.e., $\mathbf{h}_s$ is a high pass filter. Define the set of indices on which $A_J\mathbf{v}$ is known by

$$\Gamma_s := \Big\{(\ell,j,k) \ \Big|\ \ell = s;\ j = -1;\ k = 0,1,\ldots,N_0-1\Big\}. \tag{5.7}$$

Define the sequence $\mathbf{k}_s$ by

$$\mathbf{k}_s[\ell,j,k] = \begin{cases} c_k, & \text{if } (\ell,j,k) \in \Gamma_s, \\ 0, & \text{otherwise.} \end{cases}$$

- For the case that $\mathbf{h}_s$ is a low pass filter, i.e., $s=0$, the set of indices of known coefficients $A_J\mathbf{v}$ is

$$\begin{aligned}\Gamma_0 :=& \Big\{(\ell,j,k) \;\big|\; \ell=1,\ldots,r;\; j=-2,-3,\ldots,J;\; k=0,1,\ldots,N_0-1\Big\} \\ &\cup \Big\{(\ell,j,k) \;\big|\; \ell=0;\; j=J;\; k=0,1,\ldots,N_0-1\Big\} \end{aligned} \tag{5.8}$$

Define the sequence $\mathbf{k}_0$ by

$$\mathbf{k}_0 = [A_{-1\to J}\mathbf{c}; \underbrace{\mathbf{0};\ldots;\mathbf{0}}_{r\ \mathbf{0}'s}]^t.$$

Let the matrix P_{Γ_s} be the diagonal matrix with diagonal entries 1 if the indices belong to Γ_s, and 0 otherwise. Then (3.11) in Algorithm 3.1 becomes

$$\mathbf{v}_{n+1} = A_J^*\mathcal{T}^p[\mathbf{k}_s + (I-P_{\Gamma_s})A_J\mathbf{v}_n], \tag{5.9}$$

and the iteration (4.8) becomes

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p[\mathbf{k}_s + (I-P_{\Gamma_s})A_JA_J^*\tilde{\mathbf{v}}_n]. \tag{5.10}$$

The following theorem can be shown similarly to what we have done in proving Theorem 4.2 by establishing lemmas similar to Lemmas 4.5 and 4.6. In fact, it is even simpler, since the weak convergence implies norm convergence in finite dimensional spaces. We omit repeating the same proof here.

Theorem 5.1. *Iterations* (5.9) *and* (5.10) *converge in norm. Further, the limit pair* $(\mathbf{v},\tilde{\mathbf{v}})$ *satisfies* $\mathbf{v} = A_J^*\tilde{\mathbf{v}}$ *and* $\tilde{\mathbf{v}}$ *is a minimizer of*

$$\min_{\tilde{\mathbf{v}}\in\mathscr{C}} \left\{ \|(I-A_JA_J^*)\tilde{\mathbf{v}}\|^2 + \sum_{\ell=1}^{r}\sum_{j=J}^{-1}\sum_{k=0}^{N_0-1} \lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p + \sum_{k=0}^{N_0-1}\lambda_{0,J,k}|\tilde{v}_{0,J,k}|^p \right\}, \tag{5.11}$$

where $\mathscr{C} = \{\tilde{\mathbf{v}} | \tilde{\mathbf{v}} \in \mathbb{R}^{(1+r|J|)N_0};\; P_{\Gamma_s}\tilde{\mathbf{v}} = \mathcal{T}^p\mathbf{k}_s\}$.

Analogously, (3.19) in Algorithm 3.2 becomes

$$\tilde{\mathbf{v}}_{n+1} = \mathcal{T}^p\big(\tilde{\mathbf{v}}_n - \mu\delta(I-A_JA_J^*)\tilde{\mathbf{v}}_n + \delta A_JH_s^*(\mathbf{c} - H_sA_J^*\tilde{\mathbf{v}}_n)\big). \tag{5.12}$$

For this iteration, we have the following:

Theorem 5.2. *Assume that* $0 < \delta < \frac{2}{\max\{1,\mu\}}$. *Then iteration* (5.12) *converge in norm. Further, the limit* $\tilde{\mathbf{v}}$ *is a minimizer of*

$$\begin{aligned}\min_{\tilde{\mathbf{v}}} \Big\{ &\|\mathbf{c} - H_sA_J^*\tilde{\mathbf{v}}\|^2 + \mu\|(I-A_JA_J^*)\tilde{\mathbf{v}}\|^2 \\ &+ \frac{1}{\delta}\sum_{\ell=1}^{r}\sum_{j=J}^{-1}\sum_{k=0}^{N_0-1}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p + \frac{1}{\delta}\sum_{k=0}^{N_0}\lambda_{0,J,k}|\tilde{v}_{0,J,k}|^p \Big\}. \end{aligned} \tag{5.13}$$

We remark that when $s=0$, i.e the convolution kernel is a low pass filter, the convergence of iterations (5.12) with $\mu=1$ and $\delta=1$ for periodic boundary condition was proved in [6] with a rate. Moreover, the limit pair $(\mathbf{v},\tilde{\mathbf{v}})$ satisfies a minimization condition in the following sense: for any pair $(\eta,\tilde{\eta})$ with $\tilde{\eta}=A_J\eta$,

$$\begin{aligned}&\|H_s(\mathbf{v}+\eta)-\mathbf{c}\|^2+\sum_{\ell=1}^{r}\sum_{j=J}^{-1}\sum_{k=0}^{N_0-1}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}+\tilde{\eta}_{\ell,j,k}|^p+\sum_{k=0}^{N_0-1}\lambda_{0,J,k}|\tilde{v}_{0,J,k}+\tilde{\eta}_{0,J,k}|^p\\ \geq&\|H_s\mathbf{v}-\mathbf{c}\|^2+\sum_{\ell=1}^{r}\sum_{j=J}^{-1}\sum_{k=0}^{N_0-1}\lambda_{\ell,j,k}|\tilde{v}_{\ell,j,k}|^p+\sum_{k=0}^{N_0-1}\lambda_{0,J,k}|\tilde{v}_{0,J,k}|^p. \qquad (5.14)\end{aligned}$$

The convergence rate depends on the smallest eigenvalue of the matrix H_0. It is clear that the solution of the minimization problem (5.13) implies (5.14), since $(I-A_JA_J^*)\tilde{\eta}=0$. Hence, our approach here not only generalize the results in [6] to more general setting (e.g., deconvolution with high pass filter kernel), and unifies the analysis of finite dimensional case and more general case discussed in Section 3, but also improves the results of this special case, which is the case of deconvolution with low pass filter kernel in [6].

References

[1] G. Beylkin, R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms. I, *Commun. Pur. Appl. Math.*, **44** (1991), 141-183.

[2] L. Borup, R. Gribonval and M. Nielsen, Bi-framelet systems with few vanishing moments characterize Besov spaces, *Appl. Comput. Harmon. A.*, **17** (2004), 3-28.

[3] J.-F. Cai, R.H. Chan, L. Shen, and Z. Shen, Restoration of chopped and nodded images by framelets, *SIAM J. Sci. Comput.*, **30** (2008), 1205-1227.

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

[5] J.-F. Cai, S. Osher, Z. Shen, Linearized Bregman iterations for frame-based image deblurring, *SIAM J. Imaging Sci.*, **2** (2009), 226-252.

[6] A. Chai and Z. Shen, Deconvolution: A wavelet frame approach, *Numer. Math.*, **106** (2007), 529-587.

[7] R.H. Chan, T.F. Chan, L. Shen, and Z. Shen, Wavelet algorithms for high-resolution image reconstruction, *SIAM J. Sci. Comput.*, **24** (2003), 1408-1432.

[8] R.H. Chan, S.D. Riemenschneider, L. Shen, and Z. Shen, Tight frame: an efficient way for high-resolution image reconstruction, *Appl. Comput. Harmon. A.*, **17** (2004), 91-115.

[9] R.H. Chan, S.D. Riemenschneider, L. Shen, and Z. Shen, High-resolution image reconstruction with displacement errors: A framelet approach, *Int. J. Imag. Syst. Tech.*, (2005), 91-104.

[10] R.H. Chan, L. Shen, and Z. Shen, A framelet-based approach for image inpainting, 2005. Research Report 2005-04(325), Department of Mathematics, The Chinese University of Hong Kong.

[11] R.H. Chan, Z. Shen, and T. Xia, A framelet algorithm for enhancing video stills, *Appl. Comput. Harmon. A.*, **23** (2007), 153-170.

[12] A. Cohen, M. Hoffmann, and M. Reiß, Adaptive wavelet Galerkin methods for linear inverse problems, *SIAM J. Numer. Anal.*, **42** (2004), 1479-1501.

[13] P.L. Combettes and V.R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Sim.*, **4** (2005), 1168-1200.

[14] I. Daubechies, M. Defrise, and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pur. Appl. Math.*, **57** (2004), 1413-1457.

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

[16] I. Daubechies, G. Teschke, and L. Vese, Iteratively solving linear inverse problems under general convex constraints, *Inverse Problems and Imaging*, **1** (2007), 29-46.

[17] C. de Boor, R.A. DeVore, and A. Ron, On the construction of multivariate (pre)wavelets, *Constr. Approx.*, **9** (1993), 123-166.

[18] C. De Mol and M. Defrise, A note on wavelet-based inversion algorithms, in Inverse problems, image analysis, and medical imaging (New Orleans, LA, 2001), vol. 313 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2002, 85-96.

[19] B. Dong and Z. Shen, Pseudo-splines, wavelets and framelets, *Appl. Comput. Harmon. A.*, **22** (2007), 78-104.

[20] D.L. Donoho, De-noising by soft-thresholding, *IEEE T. Inform. Theory*, **41** (1995), 613-627.

[21] D.L. Donoho, Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition, *Appl. Comput. Harmon. A.*, **2** (1995), 101-126.

[22] Y. Hur and A. Ron, CAPlets: wavelet representations without wavelets, preprint, (2005).

[23] R. Jia and Z. Shen, Multiresolution and Wavelets, *Proc. Edinburgh Math. Soc.*, **37** (1994), 271-300.

[24] J. Kalifa and S. Mallat, Thresholding estimators for linear inverse problems and deconvolutions, *Ann. Stat.*, **31** (2003), 58-109.

[25] J. Kalifa, S. Mallat, and B. Rougé, Deconvolution by thresholding in mirror wavelet bases, *IEEE T. Image Process.*, **12** (2003), 446-457.

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