Image Restoration: A General Wavelet Frame Based Model and Its Asymptotic Analysis

Bin Dong[1], Zuowei Shen[2] and Peichu Xie [2]

[1] *Beijing International Center for Mathematical Research (BICMR)*
Peking University, Beijing, China, 100871
dongbin@math.pku.edu.cn (corresponding author)

[2] *Department of Mathematics, National University of Singapore*
10 Lower Kent Ridge Road, Singapore, 119076
matzuows@nus.edu.sg (Zuowei Shen), xie@u.nus.edu (Peichu Xie)

November 4, 2016

Abstract

Image restoration is one of the most important areas in imaging science. Mathematical tools have been widely used in image restoration, where the wavelet frame based approach is one of the successful examples. In this paper, we introduce a generic wavelet frame based image restoration model, called the "general model", which includes most of the existing wavelet frame based models as special cases. Moreover, the general model also includes examples that are new to the literature. Motivated by our earlier studies [1–3], we provide an asymptotic analysis of the general model as image resolution goes to infinity, which establishes a connection between the general model in discrete setting and a new variatonal model in continuum setting. The variational model also includes some of the existing variational models as special cases, such as the total generalized variational model proposed by [4].

Keywords. Image restoration, variational method, (tight) wavelet frames, framelets, Γ-convergence.

1 Introduction

Image restoration, including image denoising, deblurring, inpainting, medical imaging, etc., is one of the most important areas in imaging science. Image restoration problems can be formulated as the following linear inverse problem

$$\mathbf{f} = \mathbf{A}\mathbf{u} + \varepsilon, \tag{1.1}$$

where the matrix $\mathbf{A}$ is some linear operator (not invertible in general) and ε denotes a perturbation caused by the additive noise in the observed image, which is typically assumed to be white Gaussian noise. As convention, we regard an image as a discrete function $\mathbf{u}$ defined on a regular grid $\mathbf{O} \subset (h\mathbb{Z})^2$ (where h indicates the size of each pixel): $\mathbf{u} : \mathbf{O} \to \mathbb{R}$.

Different image restoration problems correspond to a different type of $\mathbf{A}$ in (1.1). For example, $\mathbf{A}$ is the identity operator for image denoising, a restriction operator for image inpainting, a convolution operator for image deblurring, a partial collection of line integrations for CT imaging, a partial

[1]Bin Dong is supported in part by NSF DMS-1418772 and the Thousand Talents Plan of China.
[2]Zuowei Shen is supported by the Tan Chin Tuan Centennial Professorship at National University of Singapore.

Fourier transform for MR Imaging, etc. The problem (1.1) is usually ill-posed, which makes solving (1.1) non-trivial. A naive inversion of $\mathbf{A}$ may result in a recovered image with amplified noise and smeared-out edges. A good image restoration method should be capable of smoothing the image so that noise is suppressed to the greatest extent, while at the same time, restoring or preserving important image features such as edges, ridges, corners, etc.

Most of the existing image restoration methods are transformation based. A good transformation for image restoration should be capable of capturing both global patterns and local features of images. The global patterns are smooth image components that provide a global view of images, while the local features are sharp image components that characterize local singularities of images. One of the earliest transform used is the Fourier transform, which is effective on signals that are smooth and sinusoidal like. However, the Fourier transform can only well capture global patterns due to its poor spatial localization. This makes the Fourier transform much less effective on signals with multiple localized frequency components. Windowed Fourier transforms [5] were introduced to overcome the poor spatial localization of the Fourier transform. However, the high frequency coefficients in the transform domain are not ideally sparse for images due to the fixed time-frequency resolution of windowed Fourier transforms. This is why wavelets are much more effective for images than Fourier or windowed Fourier transforms, because of their varied time-frequency resolution which enables them to provide better sparse approximation to local image features. This leads to a very successful application of orthonormal or biorthogonal wavelets in image compression (see e.g. [6]).

Most transformation based image restoration algorithms adopt an iterative thresholding dynamics. The performance of a typical thresholding algorithm highly depends on the quality of the sparse approximation to local image features provided by the underlying wavelet frame system. Although orthogonal or biorthogonal wavelets provide sparse approximation to local image features, redundant systems such as wavelet frames can provide better sparse approximation which in turn leads to better image restoration results. One advantage of wavelet frames (e.g. the B-spline tight wavelet frame systems of [7]) over (bi)orthogonal wavelets, in terms of sparse approximation, is the presence of wavelet functions in the system with varied vanishing moments and generally shorter supports. The varied order of vanishing moments enables the wavelet frame system to possess subsystems specialized in the sparse approximation of different types of singularity, such as jump discontinuities and jumps after first order differentiation (ridges). The shorter supports of wavelet frames lead to more concentrated coefficients in the transform domain. Another advantage of wavelet frames over (bi)orthogonal wavelets is their robustness to errors, thanks to the redundancy of these systems. After thresholding in the transform domain, errors are inevitably introduced no matter how careful the thresholding operator is designed. However, if a redundant system is used, there is a good chance that these errors are canceled out after transforming back to the image domain. Therefore, wavelet frame systems have enough flexibility to better balance between smoothness and sparsity than (bi)orthogonal wavelets so that artifacts generated by the Gibbs phenomenon can be further reduced, which in turn leads to better image reconstruction. Other than wavelet frames, many other systems for efficiently representing images have been proposed in the literature, especially those that can well capture directional local features. This includes curvelets [8, 9], contourlets [10], bandlets [11], shearlets [12], complex tight framelets [13] and directional Gabor frames [14].

There are many different wavelet or wavelet frame based image restoration models proposed in the literature including the synthesis based approach [15–19], the analysis based approach [20–22], and the balanced approach [23–25]. For images that are better represented by a composition of two layers which can each be sparsely approximated by two different frame systems, two-system models were proposed in [1, 20–22, 26]. Although all these models are different in form from each other, they all share the same modeling philosophy, i.e. to penalize the ℓ_1-norm (or more generally, any sparsity promoting norms) of the sparse coefficients in the wavelet frame domain. This is because wavelet frame systems can sparsely approximate local features of piecewise smooth functions such as images.

In this paper, we study a generic wavelet frame based image restoration model which includes most of the aforementioned models as special cases. This model shall be referred to as the "general model" for image restoration. Moreover, the general model also includes some models that are new to the literature. Now, we present the general model for wavelet frame based image restoration as follows:

$$\inf_{\mathbf{u},\mathbf{v}} \left\{ a\|\mathbf{W}'\mathbf{u}-\mathbf{v}\|^p_{\ell_p(\mathbf{O})} + b\|\mathbf{W}''\mathbf{v}\|^q_{\ell_q(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\}, \tag{1.2}$$

where $\mathbf{W}'$ and $\mathbf{W}''$ are wavelet frame transforms associated to two wavelet frame systems, and $1 \le p, q \le 2$. Here, $\mathbf{u}$ is the image to be recovered, and $\mathbf{v}$ lives in the transform domain of $\mathbf{W}'$. The wavelet frame system corresponding to the transform $\mathbf{W}''$ consists of subsystems that are applied to each of the components of $\mathbf{v}$. For clarity of the presentation, details of the definition of (1.2) will be postponed to a later section.

Now, we observe that the general model (1.2) indeed takes many existing wavelet frame based models as special cases.

Case A: Let $\mathbf{W}' = \mathbf{W}$ be a certain wavelet frame transform, and $\mathbf{W}'' = \mathrm{Id}$; and choose $p = 2,\ q = 1$. By fixing $\mathbf{u} \equiv \mathbf{W}^T\mathbf{v}$, the general model (1.2) becomes the *Balanced Model* of [23–25]:

$$\inf_{\mathbf{v}} \left\{ a\|(Id-\mathbf{W}\mathbf{W}^T)\mathbf{v}\|^2_{\ell_2(\mathbf{O})} + b\|\mathbf{v}\|_{\ell_1(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{W}^T\mathbf{v}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\} \tag{1.3}$$

If we further enforce the condition $a = 0$ in model (1.3), then we obtain the *Synthesis Model* [15–19]:

$$\inf_{\mathbf{v}} \left\{ b\|\mathbf{v}\|_{\ell_1(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{W}^T\mathbf{v}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\} \tag{1.4}$$

If we formally set $a = \infty$ in (1.3), or more strictly, set $\mathbf{v} = \mathbf{0}$ directly in (1.2), we obtain the following *Analysis Model* [20–22]:

$$\inf_{\mathbf{u}} \left\{ a\|\mathbf{W}\mathbf{u}\|_{\ell_1(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\} \tag{1.5}$$

Case B: Let $\mathbf{W}' = \mathbf{W}'' = \mathbf{W}$, $\mathbf{v} = \mathbf{W}\mathbf{u}_2$, $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, and $p = q = 1$. The general model (1.2) becomes the two-layers *Wavelet-Packet Model* of [1]:

$$\inf_{\mathbf{u}_1,\mathbf{u}_2} \left\{ a\|\mathbf{W}\mathbf{u}_1\|_{\ell_1(\mathbf{O})} + b\|\mathbf{W}^2\mathbf{u}_2\|_{\ell_1(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}(\mathbf{u}_1+\mathbf{u}_2)-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\}. \tag{1.6}$$

Case C: Let $p = q = 1$. The general model (1.2) becomes the more general two-layers model proposed in [1]:

$$\inf_{\mathbf{u},\mathbf{v}} \left\{ a\|\mathbf{W}'\mathbf{u}-\mathbf{v}\|_{\ell_1(\mathbf{O})} + b\|\mathbf{W}''\mathbf{v}\|_{\ell_1(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\}. \tag{1.7}$$

The general model (1.2) also includes the following new model as its special case.

Case D (New): Let $p = 1$ and $q = 2$. The general model (1.2) becomes the following model:

$$\inf_{\mathbf{u},\mathbf{v}} \left\{ a\|\mathbf{W}'\mathbf{u}-\mathbf{v}\|_{\ell_1(\mathbf{O})} + b\|\mathbf{W}''\mathbf{v}\|^2_{\ell_2(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\}. \tag{1.8}$$

When model (1.8) is used, the image to be recovered is understood as having two layers: one layer contains sharp image features while the other layer consists of smooth image components. To see this, we first suppose that $\mathbf{v}$ is in the range of $\mathbf{W}'$. Then there exists $\mathbf{u}_2$ such that $\mathbf{v} = \mathbf{W}'\mathbf{u}_2$. Letting $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, model (1.8) can be rewritten as

$$\inf_{\mathbf{u}_1,\mathbf{u}_2} \left\{ a\|\mathbf{W}'\mathbf{u}_1\|_{\ell_1(\mathbf{O})} + b\|\mathbf{W}''\mathbf{W}'\mathbf{u}_2\|^2_{\ell_2(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}(\mathbf{u}_1+\mathbf{u}_2)-\mathbf{f}\|^2_{\ell_2(\mathbf{O})} \right\}. \tag{1.9}$$

The penalization of the ℓ_1-norm of $\mathbf{W}'\mathbf{u}_1$ ensures sharp image features are well captured by $\mathbf{u}_1$, while the penalization of the ℓ_2-norm of $\mathbf{W}''\mathbf{W}'\mathbf{u}_2$ ensures the smooth image components are well captured by $\mathbf{u}_2$. Note that we present model (1.9) to show that (1.8) implicitly assumes that the image to be recovered contains two layers. They are not equivalent in general since $\mathbf{v}$ in (1.8) does not have to be in the range of $\mathbf{W}'$.

Notably, the model (1.8) is related to the newly proposed *piecewise smooth image restoration model* of [3]. We first recall the piecewise smooth model of [3] as follows

$$\inf_{\mathbf{u},\,\boldsymbol{\Lambda}\subset\mathbf{O}} a\,\|[\mathbf{W}\mathbf{u}]_{\boldsymbol{\Lambda}}\|_{\ell_1(\mathbf{O})} + b\,\|[\mathbf{W}\mathbf{u}]_{\boldsymbol{\Lambda}^c}\|^2_{\ell_2(\mathbf{O})} + \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{f}\|^2_{\ell_2(\mathbf{O})}, \tag{1.10}$$

where $\boldsymbol{\Lambda}$ is a sub-index set of $\mathbf{O}$ that indicates the locations of the image singularities, and $[\mathbf{W}\mathbf{u}]_{\boldsymbol{\Lambda}}$ (resp. $[\mathbf{W}\mathbf{u}]_{\boldsymbol{\Lambda}^c}$) denotes the restricted coefficients on set $\boldsymbol{\Lambda}$ (resp. $\boldsymbol{\Lambda}^c$). The image recovered by the piecewise smooth model (1.10) can be written as $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ where

$$\mathbf{u}_1 = \begin{cases} [\mathbf{u}]_{\boldsymbol{\Lambda}} & \text{on } \boldsymbol{\Lambda} \\ 0 & \text{elsewhere} \end{cases} \quad \text{and} \quad \mathbf{u}_2 = \begin{cases} [\mathbf{u}]_{\boldsymbol{\Lambda}^c} & \text{on } \boldsymbol{\Lambda}^c \\ 0 & \text{elsewhere.} \end{cases}$$

Therefore, the piecewise smooth image restoration model (1.10) assumes the image to be recovered consists of two layers where one contains sharp image features and the other contains smooth image components.

Comparing model (1.9) with the piecewise smooth model (1.10), we can see that both models assume the images to be recovered can be decomposed into an addition of sharp image features $\mathbf{u}_1$ and smooth image components $\mathbf{u}_2$ via the penalization of the ℓ_1-norm of the wavelet frame coefficients of $\mathbf{u}_1$ and the ℓ_2-norm of the wavelet frame coefficients of $\mathbf{u}_2$. However, the difference between them is that $\mathbf{u}_1$ of the piecewise smooth model contains only sharp image features while $\mathbf{u}_1$ of the model (1.9) contains both sharp and some smooth image components. In other words, the decomposition $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ of the piecewise smooth model is non-overlapping, while that of the model (1.9) has some overlaps. It is not clear at this point whether such overlapping will lead to better image restoration results or not. However, model (1.9) (as well as model (1.8)) is convex while the piecewise smooth model (1.10) is nonconvex. Therefore, one may expect better behavior and theoretical support for the numerical algorithms solving (1.9) (and (1.8) as well). To properly compare the two models and their associated algorithms, we need to conduct comprehensive numerical studies. However, we shall omit these numerical studies in this paper since our focus is to provide a theoretical study of the general model (1.2). Nonetheless, numerical experiments in [3] on the piecewise smooth model (1.10) showed the advantage of the modeling philosophy that is adopted by both (1.10) and (1.8), i.e. modeling images as a summation of one image layer encoding sharp image features and another layer encoding smooth image components, and penalizing the ℓ_1-norm and ℓ_2-norm of them in transform domain respectively.

1.1 Analyzing Model (1.2)

The main objective, as well as contribution, of this paper is to provide an asymptotic analysis of the general model (1.2) as image resolution goes to infinity. This work is motivated by earlier studies of [1–3], where it was shown that wavelet frame transforms are discretization of differential operators in both variational and PDE frameworks, and such discretization is superior to some of the traditional finite difference schemes for image restoration. In particular, a fundamental connection of the wavelet frame based approach to the total variation model [27] was established in [1], to the Mumford-Shah model [28] was established in [3] and to the nonlinear evolution PDEs in [2]. This new understanding essentially merged the two seemingly unrelated areas: the wavelet frame base approach and the PDE based approach. It also gave birth to many innovative and more effective image restoration models and algorithms. Therefore, an asymptotic analysis of the general model (1.2) is important to the understanding of the model, as well as the corresponding variational model.

In [1], asymptotic analysis of the wavelet frame based analysis model (1.5) was provided. Asymptotic analysis of the piecewise smooth model (1.10) with a fixed $\mathbf{\Lambda}$ was given in [3]. However, asymptotic analysis of many other wavelet frame based models proposed in the literature, such as the examples in Case A-D which we presented in the previous subsection, is still missing. In this paper, we give a unified analysis of all these wavelet frame based models by providing an asymptotic analysis of the general model (1.2). Although the general idea of the proof presented in this paper is similar to that of [1, 3], the details are vastly different and more technical due to the more complicated structure of the energy functional to be analyzed.

In model (1.2), we view images as data samples of functions at a given resolution. The discrete wavelet frame coefficients are obtained by applying wavelet frame filters to the given image data. Since the operation of high-pass filtering in the wavelet frame transform can be regarded as applying a certain finite difference operator on the image, one can easily show using Taylor's expansion that when images are sampled from functions that are smooth enough, wavelet frame transforms indeed approximate differential operators if each of the wavelet frame band is properly weighted. This allows us to postulate that there is a certain variational model to which the general model (1.2) approximates. However, we need to justify such approximation in more general function spaces than smooth function spaces since images are by no means smooth. This requires more sophisticated analysis than simple Taylor's expansion.

The analysis used in this paper is based on what was used in [1, 3]. Let $\psi_{n,\mathbf{k}}$ be a wavelet frame function and $\phi_{n,\mathbf{k}}$ the corresponding refinable function at scale $n \in \mathbb{Z}$ and location $\mathbf{k} \in \mathbf{O}$. An image $\mathbf{u}$ is understood as a discrete sample of the associated function u via the inner product $\mathbf{u}[\mathbf{k}] = \delta_n \langle u, \phi_{n,\mathbf{k}} \rangle$ where δ_n is a constant depending on n. When discrete wavelet frame transform is applied on $\mathbf{u}$, the transform corresponding to the element $\psi_{n-1,\mathbf{k}}$ produces a coefficient proportional to $\langle u, \psi_{n-1,\mathbf{k}} \rangle$. One key observation that is crucial to our analysis is that there exists a function φ associated to ψ with non-zero integration and enough smoothness, such that $\langle u, \psi_{n-1,\mathbf{k}} \rangle$ is proportional to $\langle Du, \varphi_{n-1,\mathbf{k}} \rangle$, where D is a differential operator depending on the property of the wavelet frame function ψ. In other words, the wavelet frame coefficient associated to $\psi_{n-1,\mathbf{k}}$ can be understood as a sampling of Du via $\widetilde{\delta}_n \langle Du, \varphi_{n-1,\mathbf{k}} \rangle$ with $\widetilde{\delta}_n$ a constant depending on n.

Based on the aforementioned observations, we are able to find the variational model corresponding to the general model (1.2). We will show that the objective function of an equivalent form of the general model (1.2) converges to the energy functional of the variational model as image resolution goes to infinity. In particular, the energy functional Γ-converges (see e.g. [29]) to the corresponding variational model. Through such convergence, connections of the approximate minimizers of the general model to those of the variational model are also established. A summary of our main findings is given in the next subsection.

1.2 Main Results

We assume all functions/images we consider are defined on the open unit square $\Omega := (0,1)^2 \subset \mathbb{R}^2$. Let $\mathbf{O}_n \subset \Omega$ be a $2^n \times 2^n$ Cartesian grid on $\bar{\Omega}$ with $n \in \mathbb{N}$ indicating the resolution of the grid. Let $\mathbf{K}_n \subset \mathbf{O}_n$ be an appropriate index set on which wavelet transforms are well-defined. Let $\mathbf{u}_n$ be a real-valued array defined on $\mathbf{K}_n$, i.e. $\mathbf{u}_n \in \mathbb{R}^{|\mathbf{K}_n|}$, and $\mathbf{v}_n$ be a vector-valued array on $\mathbf{K}_n$ with J components, i.e. $\mathbf{v}_n \in \mathbb{R}^{J \cdot |\mathbf{K}_n|}$. Denote $\mathbf{W}'_n : \mathbb{R}^{|\mathbf{K}_n|} \mapsto \mathbb{R}^{J|\mathbf{K}_n|}$ and $\mathbf{W}''_n : \mathbb{R}^{J|\mathbf{K}_n|} \mapsto \mathbb{R}^{J^2|\mathbf{K}_n|}$ be wavelet frame transforms with each band weighted by a certain scalar depending on n. Details of these definitions can be found in Section 2.

We start with a more precise definition of the general model (1.2).

Definition 1.1. *At a given resolution $n \in \mathbb{N}$, rewrite the general model* (1.2) *as the following* *optimization problem:*

$$\inf_{\mathbf{u}_n,\mathbf{v}_n} F_n(\mathbf{u}_n,\mathbf{v}_n) \tag{1.11}$$

where

$$F_n(\mathbf{u}_n,\mathbf{v}_n) := \nu_1\|\mathbf{W}'_n\mathbf{u}_n-\mathbf{v}_n\|^p_{\ell_p(\mathbf{K}_n;\ell_2)}+\nu_2\|\mathbf{W}''_n\mathbf{v}_n\|^q_{\ell_q(\mathbf{K}_n;\ell_2)}+\frac{1}{2}\|\mathbf{A}_n\mathbf{u}_n-\mathbf{f}_n\|^2_{\ell_2(\mathbf{K}_n)}$$

and the norm of the m-vector-valued arrays are defined as follows

$$\|(\mathbf{f}_1,\cdots,\mathbf{f}_m)\|_{\ell_p(\mathbf{K}_n;\ell_q)} := \left(2^{-2n}\sum_{\mathbf{k}\in\mathbf{K}_n}\left(\sum_{i=1}^m|\mathbf{f}_i[\mathbf{k}]|^q\right)^{p/q}\right)^{1/p},$$

with $m=J$ for the first term of F_n and $m=J^2$ for the second term of F_n.

Let the operators $\mathbf{T}_n : L_2(\Omega)\to\mathbb{R}^{|\mathbf{K}_n|}$ and $\mathbf{S}_n : L_2(\Omega;\mathbb{R}^J)\to\mathbb{R}^{J\cdot|\mathbf{K}_n|}$ be sampling operators (see (2.13) and (2.17) for details). We define the functional $E_n(u,v)$ based on the objective function $F_n(\mathbf{u}_n,\mathbf{v}_n)$:

$$E_n(u,v) := F_n(\mathbf{T}_n u,\mathbf{S}_n v) \qquad \text{with } u\in L_2(\Omega) \text{ and } v\in L_2(\Omega;\mathbb{R}^J).$$

The relation between the problem $\inf_{\mathbf{u},\mathbf{v}} F_n(\mathbf{u}_n,\mathbf{v}_n)$ and $\inf_{u,v} E_n(u,v)$ for a fixed n will be given by Proposition 3.1 which states that $\inf_{\mathbf{u}_n,\mathbf{v}_n} F_n(\mathbf{u}_n,\mathbf{v}_n)=\inf_{u,v} E_n(u,v)$; and for any given minimizer $(\mathbf{u}_n^\star,\mathbf{v}_n^\star)$ of F_n, one can find $(u_n^\star,v_n^\star)$ that is a minimizer of E_n, and vice versa.

Definition 1.2. *We use $W^p_r(\Omega)$ to denote the Sobolev space in which the r-th weak derivative is in $L_p(\Omega)$, and equipped with norm $\|f\|_{W^r_p(\Omega)} := \sum_{|\boldsymbol{k}|\le r}\left\|\frac{\partial^{\boldsymbol{k}}f}{\partial\boldsymbol{x}^{\boldsymbol{k}}}\right\|_p$, where $\frac{\partial^{\boldsymbol{k}}}{\partial\boldsymbol{x}^{\boldsymbol{k}}}=\frac{\partial^{k_1+\cdots+k_n}}{\partial x_1^{k_1}\ldots\partial x_1^{k_n}}$ is the standard multi-index notation of (weak) derivative.*

Let $\mathbf{D}'=(D'_1,\ldots,D'_J)$ be a general differential operator with $\mathbf{D}' : W^p_s(\Omega)\to W^p_{s-|\mathbf{D}'|}(\Omega;\mathbb{R}^J)$ for $s>|\mathbf{D}'|:=\max_j|D'_j|>0$. Given $u\in W^p_s(\Omega)$, $\mathbf{D}'u:=(D'_1u,\ldots,D'_Ju)$. One example of $\mathbf{D}'$ is $\mathbf{D}'=\nabla=(\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2})$ with $\nabla : W^p_s(\Omega)\to W^p_{s-1}(\Omega;\mathbb{R}^2)$.

Given $\mathbf{D}'$, let $\mathbf{D}''=(\mathbf{D}',\ldots,\mathbf{D}')$ with $\mathbf{D}'' : W^p_s(\Omega;\mathbb{R}^J)\to W^p_{s-|\mathbf{D}''|}(\Omega;\mathbb{R}^{J^2})$ for $s>|\mathbf{D}''|=|\mathbf{D}'|>0$. Given $v\in W^p_s(\Omega;\mathbb{R}^J)$, $\mathbf{D}''v:=(D'v_1,\ldots,D'v_J)$. For example, when $\mathbf{D}'=\nabla$, we have $\mathbf{D}''=(\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2},\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2})$ with $\mathbf{D}'' : W^p_s(\Omega,\mathbb{R}^2)\to W^p_{s-1}(\Omega;\mathbb{R}^4)$.

Remark 1.1. *The differential operator $\mathbf{D}''$ in Definition 1.2 is formed by stacking J copies of $\mathbf{D}'$. Note that we can make $\mathbf{D}''$ entirely general, i.e. $\mathbf{D}''=(D''_{ij})_{1\le i,j\le J}$. The proof of our main theorem can be modified to facilitate such generalization. We only need to adjust the weights at each band of $\mathbf{W}''_n$ properly. However, for better readability and clarity, we shall focus on the choice of $\mathbf{D}''$ in Definition 1.2.*

We discovered that the corresponding variational model to the discrete model $E_n(u,v)$ takes the following form:

$$E(u,v) := \nu_1\left\|\mathbf{D}'u-v\right\|^p_{L_p(\Omega;\ell_2)}+\nu_2\left\|\mathbf{D}''v\right\|^q_{L_q(\Omega;\ell_2)}+\frac{1}{2}\left\|Au-f\right\|^2_{L_2(\Omega)}, \tag{1.12}$$

where the norm of the m-vector-valued functions are defined as follows

$$\|(f_1,\cdots,f_m)\|_{L_p(\Omega;\ell_q)} := \left(\int_\Omega\left(\sum_{i=1}^m|f_i[\mathbf{x}]|^q\right)^{p/q}\mathrm{d}\mathbf{x}\right)^{1/p},$$

with $m=J$ for the first term of E and $m=J^2$ for the second term of E.

Our main result reads as follows:

Theorem 3.1. For any given differential operators $\mathbf{D}'$ and $\mathbf{D}''$ given by Definition 1.2 with order $s > 0$, one can always select the wavelet frame transforms $\mathbf{W}_n'$ and $\mathbf{W}_n''$ with each wavelet frame bands properly weighted, such that E_n Γ-converges to E under the topology of $W_{2s}^p(\Omega) \times W_s^q(\Omega; \mathbb{R}^J)$.

Based on the Γ-convergence of Theorem 3.1, we have the following result that describes the relation between the (ϵ-optimal) solutions of E_n and those of E. Recall that $(u^\star, v^\star)$ is an ϵ-optimal solution of the problem $\inf_{u,v} E(u, v)$ if

$$E(u^\star, v^\star) \le \inf_{u,v} E(u, v) + \epsilon, \quad \text{for some } \epsilon > 0.$$

In particular, 0-optimal solutions of E_n or E will be called solutions or minimizers.

Corollary 3.1. If the sequence of the (ϵ-optimal) solutions of E_n has a cluster point (u^*, v^*), then this cluster point (u^*, v^*) is an (ϵ-optimal) solution of E.

We finally note that the variational model (1.12) is closely related to the (unsymmetrized) total generalized variational (TGV) model of [4] if the infinums in (1.12) are successively enforced on u and v. In particular, we have

$$\begin{aligned} \inf_v E(u, v) &= \inf_v \left\{ \nu_1 \|\nabla u - v\|_{L_1(\Omega;\ell_2)} + \nu_2 \|\nabla v\|_{L_1(\Omega;\ell_2)} \right\} + \frac{1}{2} \|Au - f\|_{L_2(\Omega)}^2 \\ &= TGV_{\nu_2,\nu_1}(u) + \frac{1}{2} \|Au - f\|_{L_2(\Omega)}^2, \end{aligned}$$

where

$$TGV_{\nu_2,\nu_1}(u) = \inf \left\{ \int u[\nabla^2 w] \middle| w \in C^\infty(\Omega; \mathbb{R}^{J^2}),\ \|w\|_{L_\infty(\Omega;\ell_2)} \le \nu_2,\ \|\nabla w\|_{L_\infty(\Omega;\ell_2)} \le \nu_1 \right\}.$$

If we restrict v in the range of the gradient operator ∇, i.e. there exists w such that $\nabla w = v$, then we can rewrite the problem $\inf_{u,v} E(u, v)$ as

$$\inf_{u,w} \nu_1 \|\nabla(u - w)\|_{L_1(\Omega;\ell_2)} + \nu_2 \|\nabla (\nabla w)\|_{L_1(\Omega;\ell_2)} + \frac{1}{2} \|Au - f\|_{L_2(\Omega)}^2.$$

Letting $u_1 = u - w$ and $u_2 = w$, the above problem can be written equivalently as

$$\inf_{u_1,u_2} \nu_1 \|\nabla u_1\|_{L_1(\Omega;\ell_2)} + \nu_2 \|\nabla (\nabla u_2)\|_{L_1(\Omega;\ell_2)} + \frac{1}{2} \|A(u_1 + u_2) - f\|_{L_2(\Omega)}^2,$$

which was known earlier as the inf-convolution model [30, 31]. Similar as the TGV model, the inf-convolution model has been proven successful in removing the staircase artifact of the well-known total variation model [27] by properly incorporating higher order differential operators into the functional. In [32], the authors studied a more general inf-convolution functional in a discrete setting by combining ℓ_1-type norms with linear operators fulfilling some general factorization properties, which is closely related to model (1.6) and (1.9). Inf-convolution model and its variants have been successful in solving various video and image reconstruction problems [33–35]

1.3 Numerical Simulations

We present some numerical simulations of the proposed model (1.2) for image deblurring. Note that the focus of this paper is to propose the general model and provide a unified asymptotic analysis of the model to draw connections of it with variational models. Therefore, numerical simulations of this section serve as simple proof of concept for the proposed general model (1.2).

We restrict our attention to two spacial cases of the general model (1.2), namely "Case C" and "Case D". We restate them as follows with simplified notation:

$$F(\mathbf{u}, \mathbf{v}) = \nu_1 \|\mathbf{W}'\mathbf{u} - \mathbf{v}\|_1 + \nu_2 \|\mathbf{W}''\mathbf{v}\|_q^q + \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2, \tag{1.13}$$

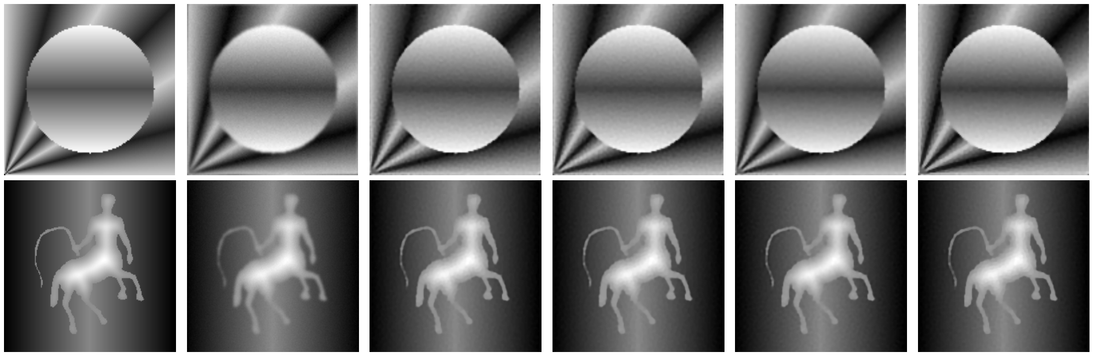

Figure 1: From left to right, the original image, the blurry and noisy, and the deblurring images using the total variation (TV) model [27], the analysis based model (1.5), the general model (1.13) with $q = 1$ and $q = 2$. The PSNR values of the recovered images for the first example (row one) are 34.32 (TV), 34.46 (Analysis), **35.12** ((1.13) with $q = 1$) and 34.87 ((1.13) with $q = 2$). The PSNR values for the second example (row two) are 36.28 (TV), 36.52 (Analysis), 36.74 ((1.13) with $q = 1$) and **36.77** ((1.13) with $q = 2$)

with $q = 1$ for "Case C" and $q = 2$ for "Case D". The above convex optimization problem can be solved using the idea of the alternating direction method of multipliers (ADMM) [36–38] which was later rediscovered as the split Bregman algorithm [20, 39]. We shall skip the derivation of the algorithms since they are quite standard. We present the algorithms minimizing $F(\mathbf{u}, \mathbf{v})$ for the case $q = 1$ and $q = 2$ as follows:

$$\begin{cases} \mathbf{u}_{k+1} = \left(\mathbf{A}^T\mathbf{A} + \mu\mathbf{I}\right)^{-1}\left(\mathbf{A}^T\mathbf{f} + \mu\mathbf{W}'^T(\mathbf{d}_k - \mathbf{p}_k)\right) \\ \mathbf{v}_{k+1} = \mathcal{T}_{\nu_1/\mu}\left(\mathbf{W}''^T(\mathbf{e}_k - \mathbf{q}_k) - \mathbf{d}_k\right) + \mathbf{d}_k \\ \mathbf{d}_{k+1} = \mathcal{T}_{\nu_1/\mu}\left(\mathbf{W}'\mathbf{u}_{k+1} + \mathbf{p}_k - \mathbf{v}_{k+1}\right) + \mathbf{v}_{k+1} \\ \mathbf{e}_{k+1} = \begin{cases} \mathcal{T}_{\nu_2/\mu}\left(\mathbf{W}''\mathbf{v}_{k+1} + \mathbf{q}_k\right) & (q = 1) \\ \frac{\mu}{2\nu_2+\mu}\left(\mathbf{W}''\mathbf{v}_{k+1} + \mathbf{q}_k\right) & (q = 2) \end{cases} \\ \mathbf{p}_{k+1} = \mathbf{p}_k + \delta\left(\mathbf{W}'\mathbf{u}_{k+1} - \mathbf{d}_{k+1}\right) \\ \mathbf{q}_{k+1} = \mathbf{q}_k + \delta\left(\mathbf{W}''\mathbf{v}_{k+1} - \mathbf{e}_{k+1}\right). \end{cases} \tag{1.14}$$

Figure 1 shows numerical results of image deblurring using algorithm (1.14), and comparisons with the total variation (TV) model [27] and the analysis based model (1.5) solved by the split Bregman algorithm/ADMM. The blur kernel is a known Gaussian filter of size 9×9. Mild Gaussian white noise is added to form the observed blurry and noisy image $\mathbf{f}$. For "Case C", i.e. (1.13) with $q = 1$, we choose framelets constructed by piecewise linear B-splines for $\mathbf{W}'$ and Haar framelets for $\mathbf{W}''$. For "Case D", i.e. (1.13) with $q = 2$, we choose Haar framelets for $\mathbf{W}'$ and framelets constructed by piecewise linear B-splines for $\mathbf{W}''$. All levels of decomposition are chosen to be 1. The specific choice of framelets and levels of decomposition is for best reconstruction results balanced with computation cost, and is empirical. Furthermore, all other parameters of the algorithms are manually chosen to obtain optimal reconstruction results as well. The first image example in Figure 1 contains jump discontinuities (edges) and also jump discontinuities after first order differentiation (ridges). For such images, model (1.13) with $q = 1$, i.e. "Case C", is more suitable as demonstrated by our numerical results. The second image example contains only jump discontinuities, while the image is smooth away from the discontinuities. For such images, model (1.13) with $q = 2$, i.e. "Case D", seems more suitable.

1.4 Organization of the Paper

In Section 2, we start with a review of wavelet frames followed by an introduction of basic notation and properties that will be needed in our analysis. Our main results are presented in Section 3, where the proof of the main theorem is given based on two technical lemmas that are proved later in Section 3.1 and Section 3.2 respectively.

2 Preliminaries

2.1 Wavelet Frames

In this section, we briefly introduce the concept of wavelet frames. The interested readers should consult [7, 40–43] for theories of frames and wavelet frames, [44, 45] for a short survey on the theory and applications of frames, and [46] for a more detailed survey.

A set $X = \{g_j : j \in \mathbb{Z}\} \subset L_2(\mathbb{R}^{\dim})$, with $\dim \in \mathbb{N}$, is called a frame of $L_2(\mathbb{R}^{\dim})$ if

$$A\|f\|^2_{L_2(\mathbb{R}^{\dim})} \le \sum_{j\in\mathbb{Z}} |\langle f, g_j\rangle|^2 \le B\|f\|^2_{L_2(\mathbb{R}^{\dim})}, \quad \forall f \in L_2(\mathbb{R}^{\dim}),$$

where $\langle\cdot,\cdot\rangle$ is the inner product of $L_2(\mathbb{R}^{\dim})$. We call X a tight frame if it is a frame with $A = B = 1$. For any given frame X of $L_2(\mathbb{R}^{\dim})$, there exists another frame $\widetilde{X} = \{\widetilde{g}_j : j \in \mathbb{Z}\}$ of $L_2(\mathbb{R}^{\dim})$ such that

$$f = \sum_{j\in\mathbb{Z}} \langle f, g_j\rangle \widetilde{g}_j \quad \forall f \in L_2(\mathbb{R}^{\dim}).$$

We call $\widetilde{X}$ a dual frame of X. We shall call the pair $(X, \widetilde{X})$ bi-frames. When X is a tight frame, we have

$$f = \sum_{j\in\mathbb{Z}} \langle f, g_j\rangle g_j \quad \forall f \in L_2(\mathbb{R}^{\dim}).$$

For given $\Psi = \{\psi_1, \ldots, \psi_J\} \subset L_2(\mathbb{R}^{\dim})$, the corresponding quasi-affine system $X^N(\Psi)$, $N \in \mathbb{Z}$ generated by Ψ is defined by the collection of the dilations and the shifts of Ψ as

$$X^N(\Psi) = \{\psi_{j,n,\mathbf{k}} :\ 1 \le j \le J; n \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^{\dim}\}, \tag{2.1}$$

where $\psi_{j,n,\mathbf{k}}$ is defined by

$$\psi_{j,n,\mathbf{k}} = \begin{cases} 2^{\frac{n\cdot\dim}{2}}\psi_j(2^n\cdot-\mathbf{k}), & n \ge N; \\ 2^{(n-\frac{N}{2})\cdot\dim}\psi_j(2^n\cdot-2^{n-N}\mathbf{k}), & n < N. \end{cases} \tag{2.2}$$

When $X^N(\Psi)$ forms a (tight) frame of $L_2(\mathbb{R}^{\dim})$, each function ψ_j, $j = 1, \ldots, J$, is called a (tight) framelet and the whole system $X^N(\Psi)$ is called a (tight) wavelet frame system. The quasi-affine system is obtained by over sampling the traditional affine (wavelet) system

$$\bar{X}(\Psi) := \{2^{\frac{n\cdot\dim}{2}}\psi_j(2^n\cdot-\mathbf{k}) : 1 \le j \le J, n \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^{\dim}\}$$

starting from level $n = N-1$ and downward. Hence, the whole quasi-affine system is a 2^{-N}-shift-invariant system. The quasi-affine system $X^0(\Psi)$ was first introduced in [7] to convert a non-shift invariant affine system to a shift invariant system. Further, it was shown in [7, Theorem 5.5] that a traditional affine system $\bar{X}(\Psi)$ is a tight frame of $L_2(\mathbb{R}^{\dim})$ if and only if the corresponding quasi-affine counterpart $X^N(\Psi)$ is a tight frame of $L_2(\mathbb{R}^{\dim})$.

Note that in the literature, the affine system $\bar{X}(\Psi)$ is commonly used, which corresponds to the decimated wavelet (frame) transforms. The quasi-affine system, which corresponds to the so-called undecimated wavelet (frame) transforms, was first introduced and analyzed by [7]. When a quasi-affine system $X^N(\Psi)$ is selected, the discrete image data is assumed to be sampled at level N and the associated undecimated wavelet (frame) decomposition starts at level N and goes downward. Here, we only discuss the quasi-affine system (2.2), since it works better in image restoration and its connection to variational models and PDEs is more natural than the affine system [1–3]. For simplicity, we denote $X(\Psi) := X^0(\Psi)$ and will focus on $X(\Psi)$ for the rest of this subsection. We will return to the generic quasi-affine system $X^N(\Psi)$ later when needed.

The constructions of framelets Ψ, which are desirably (anti)symmetric and compactly supported functions, are usually based on a multiresolution analysis (MRA) that is generated by some refinable function ϕ with refinement mask $\mathbf{p}$ and its dual MRA generated by $\widetilde{\phi}$ with refinement mask $\widetilde{\mathbf{p}}$ satisfying

$$\phi = 2^{\dim} \sum_{\mathbf{k}\in\mathbb{Z}^{\dim}} \mathbf{p}[\mathbf{k}]\phi(2\cdot - \mathbf{k}) \quad \text{and} \quad \widetilde{\phi} = 2^{\dim} \sum_{\mathbf{k}\in\mathbb{Z}^{\dim}} \widetilde{\mathbf{p}}[\mathbf{k}]\widetilde{\phi}(2\cdot - \mathbf{k}).$$

The idea of an MRA-based construction of bi-framelets $\Psi = \{\psi_1, \ldots, \psi_J\}$ and $\widetilde{\Psi} = \{\widetilde{\psi}_1, \ldots, \widetilde{\psi}_J\}$ is to find masks $\mathbf{q}^{(j)}$ and $\tilde{\mathbf{q}}^{(j)}$, which are finite sequences, such that, for $j = 1, 2, \ldots, J$,

$$\psi_j = 2^{\dim} \sum_{\mathbf{k}\in\mathbb{Z}^{\dim}} \mathbf{q}^{(j)}[\mathbf{k}]\widetilde{\phi}(2\cdot - \mathbf{k}) \quad \text{and} \quad \widetilde{\psi}_j = 2^{\dim} \sum_{\mathbf{k}\in\mathbb{Z}^{\dim}} \widetilde{\mathbf{q}}^{(j)}[\mathbf{k}]\phi(2\cdot - \mathbf{k}). \tag{2.3}$$

For a sequence $\{\mathbf{p}[\mathbf{k}]\}_{\mathbf{k}}$ of real numbers, we use $\widehat{\mathbf{p}}(\boldsymbol{\omega})$ to denote its Fourier series:

$$\widehat{\mathbf{p}}(\boldsymbol{\omega}) = \sum_{\mathbf{k}\in\mathbb{Z}^{\dim}} \mathbf{p}[\mathbf{k}]e^{-i\mathbf{k}\cdot\boldsymbol{\omega}}.$$

The mixed extension principle (MEP) of [40] provides a general theory for the construction of MRA-based wavelet bi-frames. Given two sets of finitely supported masks $\{\mathbf{p}, \mathbf{q}^{(1)}, \ldots, \mathbf{q}^{(J)}\}$ and $\{\tilde{\mathbf{p}}, \tilde{\mathbf{q}}_1, \ldots, \tilde{\mathbf{q}}_J\}$, the MEP says that as long as we have

$$\widehat{\mathbf{p}}(\boldsymbol{\xi})\overline{\widehat{\widetilde{\mathbf{p}}}(\boldsymbol{\xi})} + \sum_{j=1}^{J} \widehat{\mathbf{q}}^{(j)}(\boldsymbol{\xi})\overline{\widehat{\widetilde{\mathbf{q}}}^{(j)}(\boldsymbol{\xi})} = 1 \quad \text{and} \quad \widehat{\mathbf{p}}(\boldsymbol{\xi})\overline{\widehat{\widetilde{\mathbf{p}}}(\boldsymbol{\xi}+\boldsymbol{\nu})} + \sum_{j=1}^{J} \widehat{\mathbf{q}}^{(j)}(\boldsymbol{\xi})\overline{\widehat{\widetilde{\mathbf{q}}}^{(j)}(\boldsymbol{\xi}+\boldsymbol{\nu})} = 0, \tag{2.4}$$

for all $\boldsymbol{\nu} \in \{0, \pi\}^{\dim} \setminus \{\mathbf{0}\}$ and $\boldsymbol{\xi} \in [-\pi, \pi]^{\dim}$, the quasi-affine systems $X(\Psi)$ and $X(\widetilde{\Psi})$ with Ψ and $\widetilde{\Psi}$ given by (2.3) forms a pair of *bi-frames* in $L_2(\mathbb{R}^{\dim})$. In particular, when $\mathbf{p} = \widetilde{\mathbf{p}}$ and $\mathbf{q}^{(j)} = \widetilde{\mathbf{q}}^{(j)}$ for $j = 1, \ldots, J$, the MEP (2.4) become the following unitary extension principle (UEP) discovered in [7]:

$$|\widehat{\mathbf{p}}(\boldsymbol{\xi})|^2 + \sum_{j=1}^{J} |\widehat{\mathbf{q}}^{(j)}(\boldsymbol{\xi})|^2 = 1 \quad \text{and} \quad \widehat{\mathbf{p}}(\boldsymbol{\xi})\overline{\widehat{\mathbf{p}}(\boldsymbol{\xi}+\boldsymbol{\nu})} + \sum_{j=1}^{J} \widehat{\mathbf{q}}^{(j)}(\boldsymbol{\xi})\overline{\widehat{\mathbf{q}}^{(j)}(\boldsymbol{\xi}+\boldsymbol{\nu})} = 0, \tag{2.5}$$

and the system $X(\Psi)$ is a *tight frame* of $L_2(\mathbb{R}^{\dim})$. Here, $\mathbf{p}$ and $\widetilde{\mathbf{p}}$ are lowpass filters and $\mathbf{q}^{(j)}, \widetilde{\mathbf{q}}^{(j)}$ are highpass filters.

Now, we show two simple but useful examples of univariate tight framelets.

Example 2.1. ***(Haar)*** *Let* $\mathbf{p} = \frac{1}{2}[1, 1]$ *be the refinement mask of the piecewise constant B-spline* $B_1(x) = 1$ *for* $x \in [0, 1]$ *and* 0 *otherwise. Define* $\mathbf{q}^{(1)} = \frac{1}{2}[1, -1]$. *Then* $\mathbf{p}$ *and* $\mathbf{q}^{(1)}$ *satisfy both identities of* (2.5). *Hence, the system* $X(\psi_1)$ *defined in* (2.1) *is a tight frame of* $L_2(\mathbb{R})$.

Example 2.2. *[7]. Let* $\mathbf{p} = \frac{1}{4}[1, 2, 1]$ *be the refinement mask of the piecewise linear B-spline* $B_2(x) = \max(1 - |x|, 0)$. *Define* $\mathbf{q}^{(1)} = \frac{\sqrt{2}}{4}[1, 0, -1]$ *and* $\mathbf{q}^{(2)} = \frac{1}{4}[-1, 2, -1]$. *Then* $\mathbf{p}$, $\mathbf{q}^{(1)}$ *and* $\mathbf{q}^{(2)}$ *satisfy both identities of* (2.5). *Hence, the system* $X(\Psi)$ *where* $\Psi = \{\psi_1, \psi_2\}$ *defined in* (2.1) *is a tight frame of* $L_2(\mathbb{R})$.

In the discrete setting, let an image $\mathbf{u} \in \ell_2(\mathbb{Z}^{\dim})$ be a dim-dimensional array. We denote the fast $(L+1)$-level wavelet frame transform/decomposition with filters $\{\mathbf{q}^{(0)} = \mathbf{p}, \mathbf{q}^{(1)}, \cdots, \mathbf{q}^{(J)}\}$ (see, e.g., [46]) as

$$\mathbf{W}\mathbf{u} = \{\mathbf{W}_{j,l}\mathbf{u} : (j,l) \in \mathbb{B}\}, \tag{2.6}$$

where

$$\mathbb{B} = \{(j,l) : \ 1 \le j \le J, 0 \le l \le L\} \cup \{(0,L)\}.$$

The wavelet frame coefficients of $\mathbf{u}$ are computed by $\mathbf{W}_{j,l}\mathbf{u} = \mathbf{q}_{j,l}[-\cdot] \circledast \mathbf{u}$, where $\circledast$ denotes the convolution operator with a certain boundary condition, e.g., periodic boundary condition, and $\mathbf{q}_{j,l}$ is defined as

$$\mathbf{q}_{j,l} = \check{\mathbf{q}}_{j,l} \circledast \check{\mathbf{q}}_{0,l-1} \circledast \ldots \circledast \check{\mathbf{q}}_{0,0} \quad \text{with} \quad \check{\mathbf{q}}_{j,l}[\mathbf{k}] = \begin{cases} \mathbf{q}^{(j)}[2^{-l}\mathbf{k}], & \mathbf{k} \in 2^l\mathbb{Z}^{\dim}; \\ 0, & \mathbf{k} \notin 2^l\mathbb{Z}^{\dim}. \end{cases} \tag{2.7}$$

Similarly, we can define $\widetilde{\mathbf{W}}\mathbf{u}$ and $\widetilde{\mathbf{W}}_{j,l}\mathbf{u}$ given a set of dual filters $\{\tilde{\mathbf{q}}^{(0)} = \tilde{\mathbf{p}}, \tilde{\mathbf{q}}^{(1)}, \ldots, \tilde{\mathbf{q}}^{(J)}\}$. We denote the inverse wavelet frame transform (or wavelet frame reconstruction) as $\widetilde{\mathbf{W}}^\top$, which is the adjoint operator of $\widetilde{\mathbf{W}}$, and by the first condition of the MEP, we have the perfect reconstruction formula

$$\mathbf{u} = \widetilde{\mathbf{W}}^\top \mathbf{W}\mathbf{u}, \quad \text{for all } \mathbf{u} \in \ell_2(\mathbb{Z}^{\dim}). \tag{2.8}$$

In particular when $\mathbf{W}$ is the transform for a tight frame system, the first condition of the UEP gives us

$$\mathbf{u} = \mathbf{W}^\top \mathbf{W}\mathbf{u}, \quad \text{for all } \mathbf{u} \in \ell_2(\mathbb{Z}^{\dim}). \tag{2.9}$$

Identity (2.8) (resp. (2.9)) indicates that the associated filters generate a discrete bi-frame system (resp. a tight frame system) for the sequence space $\ell_2(\mathbb{Z}^{\dim})$.

In this paper, we will focus our analysis on the case dim $= 2$, i.e. for 2-dimensional images/functions. Also, we will only consider single-level wavelet frame transforms. For this case, we simply have

$$\mathbf{W}\mathbf{u} = \{\mathbf{q}^{(j)}[-\cdot] \circledast \mathbf{u} : 0 \le j \le J\}.$$

2.2 Notation, Assumptions and Simple Facts

Throughout the rest of this paper, we denote $\Psi = \{\psi_1, \ldots, \psi_J\}$ as the set of framelets, denote ϕ as the corresponding refinable function, and denote $\{\mathbf{q}^{(0)}, \mathbf{q}^{(1)}, \cdots, \mathbf{q}^{(J)}\}$ as the associated finitely supported filters. The following refinement equations are satisfied

$$\phi = 4\sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{q}^{(0)}[\mathbf{k}]\phi(2\cdot-\mathbf{k}) \quad \text{and} \quad \psi_j = 4\sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{k}]\phi(2\cdot-\mathbf{k}), \tag{2.10}$$

for $1 \le j \le J$. In this paper, we focus on the tensor-product B-spline tight wavelet frame systems constructed by [7], and ϕ is a tensor-product B-spline function. We shall refer to the elements in Ψ as B-spline framelets.

We start with the following basic definition:

Definition 2.1. *Let $\Omega = (0,1)^2 \subset \mathbb{R}^2$ and $n \in \mathbb{N}$. Define*

$$\begin{aligned} \mathbf{O}_n &:= \left\{\mathbf{k} \in \mathbb{Z}^2 : 2^{-n}\mathbb{Z}^2 \cap \overline{\Omega} \ne \emptyset\right\}, \\ \mathbf{M}_n &:= \left\{\mathbf{k} \in \mathbb{Z}^2 : \operatorname{supp}(\phi_{n,\mathbf{k}}) \subset \overline{\Omega}\right\}, \\ \mathbf{K}_n &:= \left\{\mathbf{k} \in \mathbf{M}_n : \mathbf{k} + C\cdot\left(|\operatorname{supp}(\mathbf{q}^{(j)})|\right) \in \mathbf{M}_n \text{ for all } 0 \le j \le J\right\}, \end{aligned} \tag{2.11}$$

where

$$\phi_{n,\mathbf{k}} = 2^n\phi(2^n\cdot-\mathbf{k}).$$

Note that the index set $\mathbf{K}_n$ (in particular, the constant C) is defined such that the boundary condition of $\left(\mathbf{q}^{(j)}[-\cdot] \circledast \mathbf{u}\right)[\mathbf{k}]$ and its high-order analogue(s) (see (2.14) *and* (2.19)*) are inactive for $\mathbf{k} \in \mathbf{K}_n$.*

Observe from the general model (1.2), the variable $\mathbf{v}$ has the same structure as the wavelet frame coefficients, which makes it a vector-valued array with J components where J is the total number of wavelet frame bands. We start with the following definition of vector-valued function/sequence spaces.

Definition 2.2. *Suppose f is a vector-valued function on a (continuum or discrete) domain Ω, i.e. for almost every $x \in \Omega$, $f(x)$ is specified as a vector in the Euclidean space $(\mathbb{R}^{J^r}, \|\cdot\|_{\ell_q})$ with $r = 0, 1, 2$. Let $W(\Omega)$ be a certain Banach space (such as an L_p, a Sobolev space or an ℓ_p space). Define*

$$\|f\|_{W(\Omega;B)} = \|f_B\|_{W(\Omega)}, \tag{2.12}$$

where f_B is the (almost everywhere defined) function such that

$$f_B(x) = \|f(x)\|_B,$$

*where $\|\cdot\|_B$ is the norm of the Banach space B. Note that we may only mention the norm (*e.g. *"$L_p(\Omega; \ell_2)$") or the space (*e.g. *"$W_s^p(\Omega; \mathbb{R}^J)$") whenever there is no confusion.*

Given $\{\phi_{n,\mathbf{k}} : \mathbf{k} \in \mathbb{Z}^2\}$, define the associated sampling operator as

$$(\mathbf{T}_n f)[\mathbf{k}] := 2^n \left\langle f, \phi_{n,\mathbf{k}} \right\rangle_{L_2(\Omega)} \qquad \text{for } f \in L_2(\Omega) \text{ and } \mathbf{k} \in \mathbf{K}_n. \tag{2.13}$$

In particular, we assume that image $\mathbf{u}_n \in \mathbb{R}^{|\mathbf{K}_n|}$ is sampled from its continuum counterpart $u \in L_2(\Omega)$ by $\mathbf{u}_n = \mathbf{T}_n u$. Therefore, when the undecimated wavelet frame transforms are applied to $\mathbf{u}_n$, the underlying quasi-affine system we use is $X^n(\Psi)$ (see (2.1) and (2.2) for the definition of $X^N(\Psi)$). Note that if $X^n(\Psi)$ is used, we have

$$\psi_{j,n-1,\mathbf{k}} = 2^{n-2}\psi_j(2^{n-1}\cdot - \mathbf{k}/2).$$

We write the standard single-level wavelet transform as

$$W_n f[j, \mathbf{k}] := 2^n \left\langle f, \psi_{j,n-1,\mathbf{k}} \right\rangle.$$

By the refinement equation (2.10), we have $\psi_{j,n-1,\mathbf{k}} = \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l} - \mathbf{k}]\phi_{n,\mathbf{l}}$, and hence

$$\begin{aligned} W_n f[j, \mathbf{k}] &= 2^n \left\langle f, \psi_{j,n-1,\mathbf{k}} \right\rangle = 2^n \left\langle f, \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l} - \mathbf{k}]\phi_{n,\mathbf{l}} \right\rangle \\ &= 2^n \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l} - \mathbf{k}] \left\langle f, \phi_{n,\mathbf{l}} \right\rangle \\ &= \left(\mathbf{q}^{(j)}[-\cdot] \circledast \mathbf{T}_n f\right)[\mathbf{k}], \qquad \text{for } \mathbf{k} \in \mathbf{K}_n. \end{aligned}$$

The key that links the wavelet frame transform $W_n f$ and differentiation of f is the observation made earlier in [1] that $\langle f, \psi_{j,n-1,\mathbf{k}}\rangle$ is a sampling of certain differentiation of f. This was justified in [1] for B-spline framelets [7]. However, the same result holds under a more general setting. What is essential in the analysis is only the vanishing moments of the framelet ψ_j. This will be proved in the following Lemma 2.1.

Let us first recall the definition of vanishing moments. We say $\mathbf{q}$ (and $\widehat{\mathbf{q}}(\boldsymbol{\xi})$) to have *vanishing moments of order* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, where $\boldsymbol{\alpha} \in \mathbb{Z}_+^2$, provided that $\sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{k}^{\boldsymbol{\alpha}}\mathbf{q}[\mathbf{k}] \neq 0$ and

$$\sum_{\mathbf{k}\in\mathbb{Z}^2} \mathbf{k}^{\boldsymbol{\beta}}\mathbf{q}[\mathbf{k}] = i^{|\boldsymbol{\beta}|}\frac{\partial^{\boldsymbol{\beta}}}{\partial\boldsymbol{\omega}^{\boldsymbol{\beta}}}\widehat{\mathbf{q}}(\boldsymbol{\xi})\Big|_{\boldsymbol{\xi}=\mathbf{0}} = 0$$

for all $\boldsymbol{\beta} \in \mathbb{Z}_+^2$ with $|\boldsymbol{\beta}| < |\boldsymbol{\alpha}|$, and for all $\boldsymbol{\beta} \in \mathbb{Z}_+^2$ with $|\boldsymbol{\beta}| = |\boldsymbol{\alpha}|$ but $\boldsymbol{\beta} \neq \boldsymbol{\alpha}$.

Lemma 2.1. *Let* $\mathbf{s} = (s_1, s_2)$ *be the vanishing moment of a finitely supported filter* $\mathbf{q}$*, and* ϕ *some compactly supported function in* $L_2(\mathbb{R}^2)$ *with support centered at* $\mathbf{0}$ *with radius* $|\mathrm{supp}(\phi)|$*, where* $|A| := ess\sup_{x\in A}|x|$ *for a set* A *in general. Let* $\psi = 4\sum_{\mathbf{k}\in\mathbb{Z}^2}\mathbf{q}[\mathbf{k}]\phi(2\cdot-\mathbf{k})$*. Then, there exists* φ *such that* $\int\varphi \neq 0$, $|supp(\varphi)| \le |supp(\psi)|$ *for some constant* $C > 0$*, and* $\frac{\partial^{\mathbf{s}}}{\partial \boldsymbol{x}^{\mathbf{s}}}\varphi = \psi$.

Proof. Observe that $\widehat{\psi}(\boldsymbol{\xi}) = \widehat{\mathbf{q}}(\boldsymbol{\xi})\cdot\widehat{\varphi}(\boldsymbol{\xi})$. Note that $\widehat{\mathbf{q}}(\boldsymbol{\xi}) \propto (i\boldsymbol{\xi})^{\mathbf{s}}$ near $\boldsymbol{\xi} = \mathbf{0}$, and $\widehat{\phi}(\mathbf{0}) = 1$. Define φ by

$$\widehat{\varphi}(\boldsymbol{\xi}) := (i\boldsymbol{\xi})^{-\mathbf{s}}\widehat{\mathbf{q}}(\boldsymbol{\xi})\widehat{\phi}(\boldsymbol{\xi}).$$

Obviously, we have $\widehat{\varphi}(\mathbf{0}) = \int\varphi \neq 0$ and $\frac{\partial^{\mathbf{s}}}{\partial \boldsymbol{x}^{\mathbf{s}}}\varphi = \psi$. By Schwartz's Paley-Wiener theorem,

$$|\widehat{\psi}(\boldsymbol{\xi})| \le C_1\cdot(1+|\boldsymbol{\xi}|)^{C_2}e^{|\mathrm{supp}(\psi)|\cdot|\mathrm{Im}\boldsymbol{\xi}|}$$

for $\boldsymbol{\xi}\in\mathbb{C}^2$. Denote $\Delta = \{\boldsymbol{\xi} : |\xi_1| \le 1,\ |\xi_2| \le 1\}$. It is immediate that

$$|\widehat{\varphi}(\boldsymbol{\xi})| = \left|(i\boldsymbol{\xi})^{-\mathbf{s}}\widehat{\psi}(\boldsymbol{\xi})\right| \le |\widehat{\psi}(\boldsymbol{\xi})|$$

for $\boldsymbol{\xi}\in\mathbb{C}^2\setminus\Delta$. Since ψ is compactly supported on $\mathbb{R}^2$, the function $\widehat{\psi}$ is analytic on $\mathbb{C}^2$. Therefore, $(i\boldsymbol{\xi})^{-\mathbf{s}}\widehat{\psi}(\boldsymbol{\xi})$ is analytic on $\mathbb{C}^2$, and

$$|\widehat{\varphi}(\boldsymbol{\xi})| = \left|(i\boldsymbol{\xi})^{-\mathbf{s}}\widehat{\psi}(\boldsymbol{\xi})\right| \le C_1'$$

for some constant C_1' for $\boldsymbol{\xi}\in\Delta$. Therefore,

$$|\widehat{\varphi}(\boldsymbol{\xi})| = \left|(i\boldsymbol{\xi})^{-\mathbf{s}}\widehat{\psi}(\boldsymbol{\xi})\right| \le \max\{C_1, C_1'\}\cdot(1+|\boldsymbol{\xi}|)^{C_2}e^{|\mathrm{supp}(\psi)|\cdot|\mathrm{Im}\boldsymbol{\xi}|},$$

for all $\boldsymbol{\xi}\in\mathbb{C}^2$. Consequently φ is compactly supported with

$$|\mathrm{supp}(\varphi)| \le |\mathrm{supp}(\psi)|\,. \tag{2.14}$$

This concludes the proof of the lemma. $\square$

Given B-spline framelets ψ_j, Lemma 2.1 implies that there exists a compactly supported function φ_j with $|supp(\varphi_j)| \le C\cdot|supp(\psi_j)|$ and $c_j := \int\varphi_j \neq 0$, such that $D_j\varphi_j = \psi_j$. The order of the differential operator D_j matches with the order of vanishing moments of ψ_j. In this case, we shall say that D_j is the differential operator associated to ψ_j. Then, letting $s_j = |D_j|$ be the total order of the differential operator D_j, we have

$$D_j\varphi_{j,n-1,\mathbf{k}} = 2^{s_j(n-1)}\psi_{j,n-1,\mathbf{k}}. \tag{2.15}$$

Therefore, the wavelet frame transform W_nf can be regarded as a sampling of the derivatives:

$$\begin{aligned} W_nf[j,\mathbf{k}] &= 2^n\langle f,\psi_{j,n-1,\mathbf{k}}\rangle = 2^{n-s_j(n-1)}\langle f, D_j\varphi_{j,n-1,\mathbf{k}}\rangle \\ &= (-1)^{s_j}2^{n-s_j(n-1)}\langle D_jf,\varphi_{j,n-1,\mathbf{k}}\rangle, \qquad \text{for } \mathbf{k}\in\mathbf{K}_n. \end{aligned}$$

Here, we have used integration by parts which is valid when f lives in a Sobolev space with sufficient regularity. Thus, we have, for $\mathbf{k}\in\mathbf{K}_n$,

$$\left(\mathbf{q}^{(j)}[-\cdot]\circledast\mathbf{T}_nf\right)[\mathbf{k}] = (-1)^{s_j}2^{n-s_j(n-1)}\langle D_jf,\varphi_{j,n-1,\mathbf{k}}\rangle\,. \tag{2.16}$$

For vector-valued function $v\in L_2\left(\Omega;\mathbb{R}^J\right)$, $v = (v_1,\ldots,v_J)$, we can define the sampling operator $\mathbf{S}_n : L_2\left(\Omega;\mathbb{R}^J\right)\to\mathbb{R}^{J|\mathbf{K}_n|}$ as follows

$$\left(\mathbf{S}_nv\right)[j;\mathbf{k}] = 2^n\left\langle v_j, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle, \quad \mathbf{k}\in\mathbf{K}_n, 1\le j\le J, \tag{2.17}$$

where $\varphi_{j,n-1,\mathbf{k}} = 2^{n-2}\varphi_j(2^{n-1}\cdot -\mathbf{k}/2)$. In particular, we assume that image $\mathbf{v}_n \in \mathbb{R}^{J|\mathbf{K}_n|}$ is sampled from its continuum counterpart $v \in L_2\left(\Omega;\mathbb{R}^J\right)$ by $\mathbf{v}_n = \mathbf{S}_n v$. One can verify that there exists a compactly supported function $\varphi_{ij,n-2,\mathbf{k}}$, with

$$\varphi_{ij,n-2,\mathbf{k}} = 2^{n-4}\varphi_{ij}(2^{n-2}\cdot -\mathbf{k}/4) \tag{2.18}$$

such that

$$\left(\mathbf{q}^{(i)}[-\cdot] \circledast c_j^{-1}\varphi_{j,n-1,\cdot}\right)[\mathbf{k}] = 2^{-(n-2)s_i} D_i\varphi_{ij,n-2,\mathbf{k}},$$

where s_i is the order of D_i. The existence of such $\varphi_{ij,n-2,\mathbf{k}}$ is guaranteed by Lemma 2.1. Note that

$$|\mathrm{supp}(\varphi_{ij,-1,\mathbf{0}})| \le C^{(i)} \cdot |\mathrm{supp}(\varphi_{j,0,\mathbf{0}})|\,. \tag{2.19}$$

Definition 2.3. *Define the weighted discrete wavelet transforms* $\mathbf{W}'_n : \mathbb{R}^{|\mathbf{K}_n|} \to \mathbb{R}^{J|\mathbf{K}_n|}$ *and* $\mathbf{W}''_n : \mathbb{R}^{J|\mathbf{K}_n|} \to \mathbb{R}^{J^2|\mathbf{K}_n|}$ *respectively as:*

$$\begin{aligned}
(\mathbf{W}'_n\mathbf{u}_n)[j;\mathbf{k}] &= \lambda'_j\left(\mathbf{q}^{(j)}[-\cdot] \circledast \mathbf{u}_n\right)[\mathbf{k}], \quad 1 \le j \le J;\\
(\mathbf{W}''_n\mathbf{t}_n)[i,j;\mathbf{k}] &= \lambda''_{ij}\left(\mathbf{q}^{(i)}[-\cdot] \circledast \mathbf{v}_n[j;\cdot]\right)[\mathbf{k}], \quad 1 \le i,j \le J.
\end{aligned} \tag{2.20}$$

In (2.20)*, the weights* λ'_j *and* $\lambda''_{i,j}$ *are chosen as*

$$\lambda'_j = c_j^{-1}(-1)^{s_j}2^{(n-1)s_j} \quad \textit{and} \quad \lambda''_{ij} = c_{ij}^{-1}(-1)^{s_i}2^{(n-2)s_i}.$$

As we will see from below that the values $\{s_j\}_{1\le j\le J} \subset \mathbb{N}$ *are the orders of the (partial) differential operators in* $\mathbf{D}'$ *and* $\mathbf{D}''$*.*

Let $\mathbf{D}'$ and $\mathbf{D}''$ be given by Definition 1.2 with $s_j := |D_j|$. By (2.16), we have

$$(\mathbf{W}'_n\mathbf{T}_n u)[j;\mathbf{k}] = 2^n\left\langle D_j u, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle, \quad \text{for } u \in L_2(\Omega). \tag{2.21}$$

Observe that

$$\begin{aligned}
\left(\mathbf{q}^{(i)}[-\cdot] \circledast 2^n\left\langle v_j, c_j^{-1}\varphi_{j,n-1,\cdot}\right\rangle\right)[\mathbf{k}] &= 2^n\left\langle v_j, \left(\mathbf{q}^{(i)}[-\cdot] \circledast c_j^{-1}\varphi_{j,n-1,\cdot}\right)[\mathbf{k}]\right\rangle\\
&= 2^n\left\langle v_j, 2^{-(n-1)s_i}D_i\varphi_{ij,n-2,\mathbf{k}}\right\rangle\\
&= (-1)^{s_i}2^{n-(n-1)s_i}\left\langle D_i v_j, \varphi_{ij,n-2,\mathbf{k}}\right\rangle
\end{aligned}$$

Here, we have used integration by parts which is valid when v_j lives in a Sobolev space with sufficient regularity. Therefore,

$$(\mathbf{W}''_n\mathbf{S}_n v)[i,j;\mathbf{k}] = 2^n\left\langle D_i v_j, c_{ij}^{-1}\varphi_{ij,n-2,\mathbf{k}}\right\rangle, \quad v \in L_2(\Omega;\mathbb{R}^J). \tag{2.22}$$

Furthermore,

$$\begin{aligned}
\left(\mathbf{S}_n\mathbf{D}'u\right)[j;\mathbf{k}] &= 2^n\left\langle D_j u, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle\\
&= 2^n\lambda'_j\left\langle u, \psi_{j,n-1,\mathbf{k}}\right\rangle\\
&= (\mathbf{W}'_n\mathbf{T}_n u)[j;\mathbf{k}].
\end{aligned} \tag{2.23}$$

Remark 2.1. *Note that identities* (2.21) *and* (2.22) *is the key that links wavelet frame transform and differential operator for the general wavelet frame based model. Identity* (2.21) *was derived earlier in [1], while* (2.22) *is new. As we mentioned earlier that we focus on single level wavelet frame transform for simplicity. However, similar relation between higher level wavelet frame transforms and* *differential operators holds. As a simple illustration, we observe that the wavelet frame coefficients in band* $\mathbf{k}$ *at level 2 satisfies*

$$\begin{aligned}
2^n \langle f, \psi_{j,n-2,\mathbf{k}} \rangle &= 2^n \left\langle f, 2^{n-2} \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l}] \phi(2^{n-1}\cdot - \mathbf{k}/2 - \mathbf{l}) \right\rangle \\
&= 2^n \left\langle f, 2^n \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l}] \sum_{\widetilde{\mathbf{l}}\in\mathbb{Z}^2} \mathbf{q}^{(0)}[\widetilde{\mathbf{l}}] \phi(2^n\cdot - \mathbf{k} - 2\mathbf{l} - \widetilde{\mathbf{l}}) \right\rangle \\
&= 2^n \left\langle f, \sum_{\mathbf{l}\in\mathbb{Z}^2} \mathbf{q}^{(j)}[\mathbf{l}] \left(\mathbf{q}^{(0)}[-\cdot] \circledast \phi_{n,\widetilde{\mathbf{l}}} \right) [\mathbf{k}+2\mathbf{l}] \right\rangle \\
&= \left(\mathbf{q}_{j,-1}[-\cdot] \circledast \mathbf{T}_n f \right)[\mathbf{k}],
\end{aligned}$$

where $\mathbf{q}_{j,-1}$ *is defined by* (2.7). *Since there exists differential operator* D_j *such that* $D_j\varphi_j = \psi_j$, *we have* $D_j\varphi_{j,n-2,\mathbf{k}} = 2^{s_j(n-2)}\psi_{j,n-2,\mathbf{k}}$. *Thus,*

$$\left(\mathbf{q}_{j,-1}[-\cdot] \circledast \mathbf{T}_n f \right)[\mathbf{k}] = (-1)^{s_j} 2^{n-s_j(n-2)} \langle D_j f, \varphi_{j,n-2,\mathbf{k}} \rangle .$$

This shows that the wavelet frame coefficients in band $\mathbf{k}$ *at level 2 are samplings of* $D_j f$. *In general, wavelet frame coefficients in the same band but at different levels are samplings of the same differential operator with different scaling constants. Therefore, by properly weighting each of the bands, we can establish a link between multi-level wavelet frame transforms and differential operators using similar arguments we use for single-level wavelet frame transforms.*

3 Asymptotic Analysis of Model (1.11)

Recall the definition of the objective function F_n of (1.11):

$$F_n(\mathbf{u}_n, \mathbf{v}_n) = \nu_1 \|\mathbf{W}'_n \mathbf{u}_n - \mathbf{v}_n\|^p_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))} + \nu_2 \|\mathbf{W}''_n \mathbf{v}_n\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} + \frac{1}{2}\|\mathbf{A}_n\mathbf{u}_n - \mathbf{f}_n\|_2^2, \tag{3.1}$$

with $1 \le p, q \le 2$. Here, the wavelet frame transforms $\mathbf{W}'_n$ and $\mathbf{W}''_n$ are given in Definition 2.3. Operator $\mathbf{A}_n$ is a certain discretization of its continuum counterpart $A : L_2(\Omega) \mapsto L_2(\Omega)$ satisfying the following condition:

$$\lim_{n\to\infty} \|\mathbf{T}_n A u - \mathbf{A}_n \mathbf{T}_n u\|_2 = 0 \quad \text{for all } u \in L_2(\Omega), \tag{3.2}$$

Note that operator A that corresponds to image denoising, deblurring and inpainting indeed satisfies the above assumption [1].

To study the asymptotic behaviour of the variation model (1.11) thoroughly, we first rewrite the objective function F_n to a new one that is defined on a function space instead of a finite dimensional Euclidean space. We regard $\mathbf{f}_n$, $\mathbf{u}_n$ and $\mathbf{v}_n$ as samples of their continuum counterparts $f \in L_2(\Omega)$ $u \in W^p_{2s}(\Omega)$ and $v \in W^q_s(\Omega;\mathbb{R}^J)$, i.e. $\mathbf{f}_n = \mathbf{T}_n f$, $\mathbf{u}_n = \mathbf{T}_n u$ and $\mathbf{v}_n = \mathbf{S}_n v$. Then, we define

$$\begin{aligned}
E_n(u,v) \quad &:= \quad \nu_1 \|\mathbf{W}'_n \mathbf{T}_n u - \mathbf{S}_n v\|^p_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))} + \nu_2 \|\mathbf{W}''_n \mathbf{S}_n v\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} + \frac{1}{2}\|\mathbf{A}_n\mathbf{T}_n u - \mathbf{T}_n f\|_2^2 \\
&=: \quad \tilde{V}_n(u,v) + \frac{1}{2}\|\mathbf{A}_n\mathbf{T}_n u - \mathbf{T}_n f\|_2^2
\end{aligned} \tag{3.3}$$

The following proposition ensures that the original problem $\inf_{\mathbf{u}_n,\mathbf{v}_n} F_n(\mathbf{u}_n, \mathbf{v}_n)$ is equivalent to the new problem $\inf_{u,v} E_n(u,v)$.

Proposition 3.1. *For any given* $\mathbf{u}_n \in \mathbb{R}^{|\mathbf{K}_n|}$, $\mathbf{v}_n \in \mathbb{R}^{J|\mathbf{K}_n|}$, *there exists* $u \in W^p_{2s}(\Omega)$ *and* $v \in W^q_s(\Omega; \mathbb{R}^J)$ *such that* $E_n(u, v) = F_n(\mathbf{u}_n, \mathbf{v}_n)$. *Conversely, for any* $u \in W^p_{2s}(\Omega)$ *and* $v \in W^q_s(\Omega; \mathbb{R}^J)$, *there exist* $\mathbf{u}_n$ *and* $\mathbf{v}_n$ *such that* $E_n(u, v) = F_n(\mathbf{u}_n, \mathbf{v}_n)$. *In particular, we have*

$$\inf_{\mathbf{u}_n \in \mathbb{R}^{|\mathbf{K}_n|}, \mathbf{v}_n \in \mathbb{R}^{J|\mathbf{K}_n|}} F_n(\mathbf{u}_n, \mathbf{v}_n) = \inf_{u \in W^p_{2s}(\Omega), v \in W^q_s(\Omega; \mathbb{R}^J)} E_n(u, v).$$

To prove Proposition 3.1, the following lemma is needed.

Lemma 3.1. *Given the sets of linearly independent compactly supported functions* $\{\phi_{n,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ *and* $\{\varphi_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ *for any* $1 \le j \le J$, *there exist dual functions* $\{\tilde{\phi}_{n,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ *and* $\{\tilde{\varphi}_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$, *with any prescribed smoothness, whose translations satisfy the relations*

$$\left\langle \tilde{\phi}_{n,\mathbf{k}'}, \phi_{n,\mathbf{k}} \right\rangle = \delta_{\mathbf{k}'\mathbf{k}} \quad \textit{and} \quad \langle \tilde{\varphi}_{j,n-1,\mathbf{k}'}, \varphi_{j,n-1,\mathbf{k}} \rangle = \delta_{\mathbf{k}'\mathbf{k}}.$$

Proof. The proof of existence of dual for $\{\phi_{n,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ was given by [1, Proposition 3.1]. Therefore, we focus on the proof of $\{\varphi_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ for any $1 \le j \le J$.

Since $\varphi_{j,n-1,0}$ has a compact support, $\hat{\varphi}_{j,n-1,0}$ is analytic, and thus has only isolated zeros. Consequently, for any $\hat{\mathbf{a}} \in L_2([-\pi, \pi]^2)$, $\hat{\mathbf{a}}(\boldsymbol{\xi})\hat{\varphi}_{j,n-1,0}(\boldsymbol{\xi}) = 0$ implies $\hat{\mathbf{a}}(\boldsymbol{\xi}) = 0$. In particular, given any sequence $\mathbf{a} \in \ell_0(\mathbb{Z}^2)$, $\sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{a}[\mathbf{k}]\varphi_{j,n-1,\mathbf{k}} = 0$ implies all coefficients $\mathbf{a}[\mathbf{k}] = 0$, i.e. the given system is finite linear independent. Let σ be some compactly supported function with a certain given smoothness, and $f^\sigma = f * \sigma$. Note that $\{\varphi^\sigma_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n\}$ is a linearly independent set since $\hat{\varphi^\sigma}_{j,n-1,\mathbf{k}}$ is analytic. Consequently, $\varphi^\sigma_{j,n-1,\mathbf{k}_0} \notin span\{\varphi^\sigma_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n - \{\mathbf{k}_0\}\}$, and there exists a function $f_0 \in (span\{\varphi^\sigma_{j,n-1,\mathbf{k}} : \mathbf{k} \in \mathbf{K}_n - \{\mathbf{k}_0\}\})^\perp \setminus (span\{\varphi^\sigma_{j,n-1,\mathbf{k}_0}\})^\perp \neq \emptyset$, where the total space is $L_2(\mathbb{R}^2)$. Define $\tilde{\varphi}_{j,n-1,\mathbf{k}_0} = \left(\left\langle f_0, \varphi^\sigma_{j,n-1,\mathbf{k}_0} \right\rangle^{-1} f_0 \right)^\sigma$, and apply the same process to the other $\mathbf{k} \in \mathbf{K}_n$. We obtain the desired result. □

Proof. (**of Proposition 3.1**) For one direction, given $\mathbf{u}_n$, $\mathbf{v}_n$, define $u = 2^{-n} \sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{u}_n[\mathbf{k}]\tilde{\phi}_{n,\mathbf{k}}$, $v_j = 2^{-n} \sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{v}_n[j; \mathbf{k}]\tilde{\varphi}_{j,n-1,\mathbf{k}}$, then

$$\begin{aligned} (\mathbf{T}_n u)[\mathbf{k}] &= \sum_{\mathbf{k}' \in \mathbb{Z}^2} \mathbf{u}_n[\mathbf{k}'] \left\langle \tilde{\phi}_{n,\mathbf{k}'}, \phi_{n,\mathbf{k}} \right\rangle = \mathbf{u}_n[\mathbf{k}] \\ (\mathbf{S}_n v)[j; \mathbf{k}] &= \sum_{\mathbf{k}' \in \mathbb{Z}^2} \mathbf{v}_n[j; \mathbf{k}'] \langle \tilde{\varphi}_{j,n-1,\mathbf{k}'}, \varphi_{j,n-1,\mathbf{k}} \rangle = \mathbf{v}_n[j; \mathbf{k}] \end{aligned}$$

consequently, $E_n(u, v) = F_n(\mathbf{u}_n, \mathbf{v}_n)$. Conversely, given u and v, define $\mathbf{u}_n = \mathbf{T}_n u$ and $\mathbf{v}_n = \mathbf{S}_n v$ then obviously, $F_n(\mathbf{u}_n, \mathbf{v}_n) = E_n(u, v)$. □

Consider the variational problem

$$\inf_{u \in W^p_{2s}(\Omega), v \in W^q_s(\Omega; \mathbb{R}^J)} E(u, v),$$

where

$$E(u, v) = \nu_1 \|\mathbf{D}' u - v\|^p_{L_p(\Omega; \ell_2(\mathbb{R}^J))} + \nu_2 \|\mathbf{D}'' v\|^q_{L_q(\Omega; \ell_2(\mathbb{R}^{J^2}))} + \frac{1}{2} \|Au - f\|^2_{L_2(\Omega)}. \tag{3.4}$$

Here, the differential operators $\mathbf{D}'$ and $\mathbf{D}''$ are defined in Definition 1.2 with $|\mathbf{D}'| = |\mathbf{D}''| = s$. Our main objective of this section is to show the relation between E_n and E, and how the solutions of $\inf E_n$ approximates that of $\inf E$. For this, we will use Γ-convergence [29] as the main tool.

Definition 3.1. *The sequence of functionals* $\{E_n\}$ *defined on a Banach space* B *equipped with a norm* $\|\cdot\|_B$ *is said to be* Γ*-convergent to the functional* E *if:*

1. $\liminf_{n\to\infty} E_n(u_n) \geq E(u)$ *for arbitrary* $u_n \xrightarrow{B} u$*;*

2. *For arbitrary* $u \in B$*, there exists* $u'_n \xrightarrow{B} u$ *such that:* $\limsup_{n\to\infty} E_n(u'_n) \leq E(u)$*.*

Here, $g_n \xrightarrow{B} g$ *means* $\|g_n - g\|_B \to 0$*.*

To show that E_n given by (3.3) indeed Γ-converges to the functional E given by (3.4), we will use the following two lemmas, which show that E_n converges to E pointwise and $\{E_n\}$ is equicontinuous. The proof of the two lemmas will be postponed to the later part of this section.

Lemma 3.2. *The energy functional* E_n *converges to* E *pointwise, that is, for* $(u,v) \in W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$*,*

$$\lim_{n\to\infty} E_n(u,v) = E(u,v)$$

Lemma 3.3. *The sequence of energy functionals* $\{E_n\}$ *forms an equicontinuous family in the sense that, for any given function* $(u,v) \in W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$*, and* $\varepsilon > 0$*, there exists an* $\eta > 0$ *independent from* n *such that* $|E_n(u',v') - E_n(u,v)| < \varepsilon$ *holds for any* (u',v') *satisfying* $\|u'-u\|_{W^p_{2s}(\Omega)} + \|v'-v\|_{W^q_s(\Omega;\mathbb{R}^J)} < \eta$*.*

Theorem 3.1. *Let* E_n *be given by* (3.3) *and* E *by* (3.4)*. Then, for every sequence* $(u_n, v_n) \to (u,v)$ *in* $W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$*, we have* $\lim_{n\to+\infty} E_n(u_n,v_n) = E(u,v)$*. Consequently,* E_n *Γ-converges to* E *in* $W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$*.*

Proof. The proof of Theorem 3.1 is essentially the same as [1, Theorem 3.2] once we have Lemma 3.2 and Lemma 3.3. However, for completeness, we include the proof here.

By Lemma 3.2 and Lemma 3.3, we have that, for an arbitrary given $(u,v) \in W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$, and $\epsilon > 0$,

(a) $\lim_{n\to+\infty} |E_n(u,v) - E(u,v)| = 0$;

(b) there exist an integer $\mathcal{N}$ and $\eta > 0$ satisfying $|E_n(u',v') - E_n(u,v)| < \epsilon$ whenever $\|u' - u\|_{W^p_{2s}(\Omega)} + \|v'-v\|_{W^q_s(\Omega;\mathbb{R}^J)} < \eta$ and $n > \mathcal{N}$.

Note that for arbitrary $(u_n, v_n) \in W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$, we have

$$|E_n(u_n,v_n) - E(u,v)| \leq |E_n(u,v) - E(u,v)| + |E_n(u,v) - E_n(u_n,v_n)|.$$

Let the sequence $(u_n,v_n) \to (u,v)$ in $W^p_{2s}(\Omega) \times W^q_s(\Omega;\mathbb{R}^J)$, and let $\epsilon > 0$ be a given arbitrary number. On one hand, by (a), there exists an $\mathcal{N}_1$ such that $|E_n(u,v) - E(u,v)| < \epsilon/2$ whenever $n > \mathcal{N}_1$. On the other hand, by (b), there exist $\mathcal{N}$ and η such that $|E_n(u',v') - E_n(u,v)| < \epsilon/2$ whenever $\|u'-u\|_{W^p_{2s}(\Omega)} + \|v'-v\|_{W^q_s(\Omega;\mathbb{R}^J)} < \eta$ and $n > \mathcal{N}$. Since $(u_n,v_n) \to (u,v)$, there exists $\mathcal{N}_2$ such that $\|u - u_n\|_{W^p_{2s}(\Omega)} + \|v - v_n\|_{W^q_s(\Omega;\mathbb{R}^J)} < \eta$ whenever $n > \mathcal{N}_2$. Letting $u' = u_n$ and $v' = v_n$ leads to $|E_n(u_n,v_n) - E_n(u,v)| < \epsilon/2$ whenever $n > \max\{\mathcal{N}, \mathcal{N}_2\}$. Therefore, we have

$$|E_n(u_n,v_n) - E(u,v)| \leq \epsilon$$

whenever $n > \max\{\mathcal{N}, \mathcal{N}_1, \mathcal{N}_2\}$. This shows that $\lim_{n\to+\infty} E_n(u_n,v_n) = E(u,v)$, and hence both conditions given in Definition 3.1 are satisfied. Therefore, E_n Γ-converges to E. □

By Theorem 3.1, we have the following result describing the relation between the ϵ-optimal solutions of E_n and that of E.

Corollary 3.1. *Let* $(u^\star_n, v^\star_n)$ *be an* ϵ*-optimal solution of* E_n *for a given* $\epsilon > 0$ *and for all* n*. If the set* $\{(u^\star_n, v^\star_n) : n\}$ *has a cluster point* $(u^\star, v^\star)$*, then* $(u^\star, v^\star)$ *is an* ϵ*-optimal solution to* E*. In particular, when* $(u^\star_n, v^\star_n)$ *is a minimizer of* E_n *and* $(u^\star, v^\star)$ *a cluster point of the set* $\{(u^\star_n, v^\star_n) : n\}$*, then* $(u^\star, v^\star)$ *is a minimizer of* E*.*

The rest of this section is dedicated to the proof of Lemma 3.2 and Lemma 3.3.

3.1 Proof of Lemma 3.2

The pointwise convergence of the third term of E_n to that of E has been shown by [1] under assumption (3.2). Therefore, we focus on the convergence of the first two terms. Let us first show that

$$\|\mathbf{W}_n''\mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} \to \|\mathbf{D}''v\|_{L_q(\Omega;\ell_2(\mathbb{R}^{J^2}))},$$

which will imply

$$\|\mathbf{W}_n''\mathbf{S}_n v\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} \to \|\mathbf{D}''v\|^q_{L_q(\Omega;\ell_2(\mathbb{R}^{J^2}))}.$$

Recall from Definition 1.2 that, given $\mathbf{D}'$, we have $\mathbf{D}'' = (\mathbf{D}', \dots, \mathbf{D}')$. Then,

$$\begin{aligned}
& \left| \|\mathbf{D}''v\|_{L_q(\Omega;\ell_2(\mathbb{R}^{J^2}))} - \|\mathbf{W}_n''\mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} \right| \\
= & \left| \left(\int_\Omega \left(\sum_{i,j=1}^J |D_i' v_j|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} - \left(2^{-2n} \sum_{\mathbf{k}\in\mathbf{K}_n} \left(\sum_{i,j=1}^J |(\mathbf{W}_n''\mathbf{S}_n v)[i,j;\mathbf{k}]|^2 \right)^{q/2} \right)^{1/q} \right| \\
\text{(By (2.22))} \quad = & \left| \left(\sum_{\mathbf{k}\in\mathbf{O}_n} \int_{I_\mathbf{k}} \left(\sum_{i,j=1}^J |D_i' v_j|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} \right. \\
& \qquad \left. - \left(2^{-2n} \sum_{\mathbf{k}\in\mathbf{K}_n} \left(\sum_{i,j=1}^J \left| \left\langle D_i' v_j, c_{ij}^{-1} \varphi_{ij,n-2,\mathbf{k}} \right\rangle \right|^2 \right)^{q/2} \right)^{1/q} \right| \\
= & \left| \left(\sum_{\mathbf{k}\in\mathbf{O}_n} \int_{I_\mathbf{k}} \left(\sum_{i,j=1}^J |D_i' v_j|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} \right. \\
& \qquad \left. - \left(\sum_{\mathbf{k}\in\mathbf{K}_n} \int_{I_\mathbf{k}} \left(\sum_{i,j=1}^J \left| \sum_{\mathbf{k}'\in\mathbf{K}_n} \left\langle D_i' v_j, c_{ij}^{-1} \varphi_{ij,n-2,\mathbf{k}'} \right\rangle \chi_{I_\mathbf{k}} \right|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} \right| \\
\le & \left(\sum_{\mathbf{k}\in\mathbf{K}_n} \int_{I_\mathbf{k}} \left(\sum_{i,j=1}^J \left| D_i' v_j - \sum_{\mathbf{k}'\in\mathbf{K}_n} \left\langle D_i' v_j, c_{ij}^{-1} \varphi_{ij,n-2,\mathbf{k}'} \right\rangle \chi_{I_\mathbf{k}} \right|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} \\
& + \left(\int_{\cup_{\mathbf{k}\in\mathbf{O}_n\setminus\mathbf{K}_n} I_\mathbf{k}} \left(\sum_{i,j=1}^J |D_i' v_j|^2 \right)^{q/2} \mathrm{d}\mathbf{x} \right)^{1/q} \\
\le & \sum_{i,j=1}^J \left(\left\| D_i' v_j - \sum_{\mathbf{k}\in\mathbf{K}_n} \left\langle D_i' v_j, c_{ij}^{-1} \varphi_{ij,n-2,\mathbf{k}} \right\rangle \chi_{I_\mathbf{k}} \right\|_{L_q(\Omega)} + \|D_i' v_j\|_{L_q(\cup_{\mathbf{k}\in\mathbf{O}_n\setminus\mathbf{K}_n} I_\mathbf{k})} \right).
\end{aligned}$$

Here, $I_\mathbf{k}$ is the rectangular domain $[\frac{k_1}{2^n}, \frac{k_1+1}{2^n}] \times [\frac{k_2}{2^n}, \frac{k_2+1}{2^n}]$ where $\mathbf{k} = (k_1, k_2)$, and the index sets $\mathbf{O}_n$ and $\mathbf{K}_n$ were defined in Definition 2.1. By the approximation lemma [1, Lemma 4.1], we have, for each i, j,

$$\lim_{n\to\infty} \left\| D_i' v_j - \sum_{\mathbf{k}\in\mathbf{O}_n} \left\langle D_i' v_j, c_{ij}^{-1} \varphi_{ij,n-2,\mathbf{k}} \right\rangle \chi_{I_\mathbf{k}} \right\|_{L_q(\Omega)} = 0.$$

Also, the Lebesgue measure $\mathfrak{L}$ of the set $\cup_{\mathbf{k}\in\mathbf{O}_n\setminus\mathbf{K}_n} I_{\mathbf{k}}$ satisfies

$$\mathfrak{L}(\cup_{\mathbf{k}\in\mathbf{O}_n\setminus\mathbf{K}_n} I_{\mathbf{k}}) \leq 4\cdot 2^n\left(\frac{\mathrm{diam}(\mathrm{supp}(\phi_{n,0}))}{2^{-n}}+1\right)\cdot(2^{-n})^2 = 4c\cdot 2^{-n}.$$

Thus, we have

$$\lim_{n\to\infty}\|D_i' v_j\|_{L_q(\cup_{\mathbf{k}\in\mathbf{O}_n-\mathbf{K}_n} I_{\mathbf{k}})} = 0,$$

since $v\in W_s^q(\Omega;\mathbb{R}^J)$ which implies that $D_i'v_j\in L_q(\Omega)$ (for each $(i,j)\in\{1,\cdots,J\}^2$). Altogether, we have

$$\|\mathbf{W}_n''\mathbf{S}_n v\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))}\to\|\mathbf{D}''v\|^q_{L_q(\Omega;\ell_2(\mathbb{R}^{J^2}))}.$$

Recall (2.23) that $\mathbf{W}_n'\mathbf{T}_n u = \mathbf{S}_n Du$. Then, following a similar proof as above by replacing the previous summing index i,j with merely j, we have:

$$\begin{aligned}
& \left|\|\mathbf{D}'u-v\|_{L_p(\Omega;\ell_2(\mathbb{R}^J))}-\|\mathbf{W}_n'\mathbf{T}_n u-\mathbf{S}_n v\|_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))}\right| \\
= & \left|\|\mathbf{D}'u-v\|_{L_p(\Omega;\ell_2(\mathbb{R}^J))}-\|\mathbf{S}_n\left(\mathbf{D}'u-v\right)\|_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))}\right| \\
\leq & \sum_{j=1}^J\left\|(Du-v)_j-\sum_{\mathbf{k}\in\mathbf{O}_n}\left\langle(Du-v)_j, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle\chi_{I_{\mathbf{k}}}\right\|_{L_p(\Omega)} \\
& +\sum_{j=1}^J\|D_j'u-v_j\|_{L_p(\cup_{\mathbf{k}\in\mathbf{O}_n\setminus\mathbf{K}_n} I_{\mathbf{k}})} \\
\to & \ 0.
\end{aligned}$$

Therefore, we have

$$\|\mathbf{W}_n'\mathbf{T}_n u-\mathbf{S}_n v\|^p_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))}\to\|\mathbf{D}'u-v\|^p_{L_p(\Omega;\ell_2(\mathbb{R}^J))},$$

which concludes the proof of the lemma.

3.2 Proof of Lemma 3.3

We shall focus on the equicontinuity of

$$\tilde{V}_n=\nu_1\|\mathbf{W}_n'\mathbf{T}_n u-\mathbf{S}_n v\|^p_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))}+\nu_2\|\mathbf{W}_n''\mathbf{S}_n v\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))},$$

since the equicontinuity of $\frac{1}{2}\|\mathbf{A}_n\mathbf{T}_n u-\mathbf{T}_n f\|_2^2$ has been established in [1, Proposition 3.2].

Let us begin with the bound of the linear operator $S_n : L_p(\Omega;\mathbb{R}^J)\to\mathbb{R}^{JK_n}$. We denote $\Lambda_{\mathbf{k}} := supp(\varphi_{j,n-1,\mathbf{k}})$. Note that when B-spline framelets are used, $\varphi_{j,n-1,\mathbf{k}}$ has the same support for different j. Now, consider

$$\begin{aligned}
\|\mathbf{S}_n v\|_{\ell_p(\mathbf{K}_n;\ell_2(\mathbb{R}^J))} &= \left(2^{-2n}\sum_{\mathbf{k}\in\mathbf{K}_n}\left(\sum_{j=1}^J|2^n\left\langle v_j, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle|^2\right)^{p/2}\right)^{1/p} \\
&\leq (2^{-n/p})^2\left(\sum_{\mathbf{k}\in\mathbf{K}_n}\left(\sum_{j=1}^J\left|2^n\left\langle v_j, c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\rangle\right|\right)^p\right)^{1/p} \\
&\leq (2^{-n/p})^2\cdot 2^n\left(\sum_{\mathbf{k}\in\mathbf{K}_n}\left(\sum_{j=1}^J\|v_j\|_{L_1(\Lambda_{\mathbf{k}})}\left\|c_j^{-1}\varphi_{j,n-1,\mathbf{k}}\right\|_{L_\infty(\Omega)}\right)^p\right)^{1/p}
\end{aligned}$$

$$
\begin{aligned}
&= (2^{-n/p})^2 \cdot 2^{2(n-1)} \left(\max_j \left\| c_j^{-1} \varphi_j \right\|_{L_\infty(\Omega)} \right) \left(\sum_{\mathbf{k} \in \mathbf{K}_n} \left(\sum_{j=1}^J \|v_j\|_{L_1(\Lambda_{\mathbf{k}})} \right)^p \right)^{1/p} \\
&\leq (2^{-n/p})^2 \cdot 2^{2(n-1)} \left(\max_j \left\| c_j^{-1} \varphi_j \right\|_{L_\infty(\Omega)} \right) \sum_{j=1}^J \left(\sum_{\mathbf{k} \in \mathbf{K}_n} \|v_j\|_{L_1(\Lambda_{\mathbf{k}})}^p \right)^{1/p} \\
&\leq C(n) \|v\|_{L_1(\Omega;\ell_p(\mathbb{R}^J))} \leq C'(n) \|v\|_{L_p(\Omega;\ell_2(\mathbb{R}^j))}, \qquad (3.5)
\end{aligned}
$$

where we applied after Hölder's inequality (in the third line) and the fact that

$$
\|\varphi_{j,n-1,\mathbf{k}}\|_{L_\infty(\Omega)} = 2^{n-2} \|\varphi_j\|_{L_\infty(\Omega)}
$$

which can be easily verified using $\varphi_{j,n-1,\mathbf{k}} = 2^{n-2} \varphi_j \left(2^{n-1} \cdot -\mathbf{k}/2 \right)$.

Consider the family of linear operators $2^{-2n/q} \mathbf{W}_n'' \mathbf{S}_n : W_s^q(\Omega; \mathbb{R}^J) \to \ell_{q,2}(\mathbb{Z}^2 \times J^2)$ ordered by n, where the norm of the latter is generally defined as

$$
\|\mathbf{a}[\mathbf{k}; i, j]\|_{\ell_{p,q}} = \left(\sum_{\mathbf{k} \in \mathbb{Z}^2} \left(\sum_{i,j=1}^J |\mathbf{a}[\mathbf{k}; i, j]|^q \right)^{p/q} \right)^{1/p}
$$

Based upon the above observation on $\mathbf{S}_n$, and the boundedness of the operator $\mathbf{W}_n''$ as a matrix, we have

$$
\|2^{-2n/q} \mathbf{W}_n'' \mathbf{S}_n v\|_{\ell_{q,2}} \leq C(n) \|v\|_{L_1(\Omega;\ell_q(\mathbb{R}^J))} \leq C'(n) \|v\|_{W_s^q(\Omega;\ell_2(\mathbb{R}^J))}
$$

Since we have proved that, for any given $v \in W_s^p(\Omega; \mathbb{R}^J)$,

$$
\lim_{n \to \infty} \|\mathbf{W}_n'' \mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n; \ell_2(\mathbb{R}^{J^2}))} = \|\mathbf{D}'' v\|_{L_q(\Omega; \ell_2(\mathbb{R}^{J^2}))}
$$

we have

$$
\sup_n \|2^{-2n/p} \mathbf{W}_n'' \mathbf{S}_n v\|_{\ell_{q,2}} = \sup_n \|\mathbf{W}_n'' \mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n; \ell_2(\mathbb{R}^{J^2}))} < \infty.
$$

Recall from the principle of uniform boundedness that if a sequence of bounded operators converges pointwise, then these pointwise limits define a bounded operator. Therefore, $\{2^{-2n/q} \mathbf{W}_n'' \mathbf{S}_n\}_{n=1}^\infty$ is uniformly bounded by some constant, i.e.

$$
\|\mathbf{W}_n'' \mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n; \ell_2(\mathbb{R}^{J^2}))} \leq C_1 \|v\|_{W_s^q(\Omega; \ell_2(\mathbb{R}^J))} \qquad (3.6)
$$

based on which the rest is justified by Sobolev's imbedding theorem (see, e.g., [47, Theorem 4.12]).

Applying the bound of $\mathbf{S}_n$ again, we have

$$
\begin{aligned}
\|\mathbf{W}_n' \mathbf{T}_n u - \mathbf{S}_n v\|_{\ell_p(\mathbf{K}_n; \ell_2(\mathbb{R}^J))} &= \|\mathbf{S}_n(\mathbf{D}' u - v)\|_{\ell_p(\mathbf{K}_n; \ell_2(\mathbb{R}^J))} \\
&\leq \|\mathbf{S}_n\|_{op} \|(\mathbf{D}' u - v)\|_{L_p(\Omega; \ell_2(\mathbb{R}^J))} \\
&\leq C_2(n) \left(\|u\|_{W_{2s}^p(\Omega)} + \|v\|_{W_s^q(\Omega; \ell_2(\mathbb{R}^J))} \right)
\end{aligned}
$$

where the Sobolev's imbedding theorem is applied in the last inequality. Since

$$
\lim_{n \to \infty} \|\mathbf{W}_n' \mathbf{T}_n u - \mathbf{S}_n v\|_{\ell_p(\mathbf{K}_n; \ell_2(\mathbb{R}^J))} = \|\mathbf{D}' u - v\|_{L_p(\Omega; \ell_2(\mathbb{R}^J))},
$$

following a similar argument using the resonance theorem, we have

$$
\|\mathbf{W}_n' \mathbf{T}_n u - \mathbf{S}_n v\|_{\ell_p(\mathbf{K}_n; \ell_2(\mathbb{R}^J))} \leq C \left(\|u\|_{W_{2s}^p(\Omega)} + \|v\|_{W_s^q(\Omega; \ell_2(\mathbb{R}^J))} \right). \qquad (3.7)
$$

Observe that $|x^q - y^q| \le q(\max\{x,y\})^{q-1}|x-y| \le q(y+|x-y|)^{q-1}|x-y|$ for $x, y \ge 0$ and $q \ge 1$. Using (3.6) and (3.7), for any (u^*, v^*) in the fixed unit neighborhood $B((u,v);\delta)$ (with $\delta \le 1$), we have

$$\begin{aligned}
&\left| \|\mathbf{W}''_n \mathbf{S}_n v^*\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} - \|\mathbf{W}''_n \mathbf{S}_n v\|^q_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} \right| \\
\le\ & q\left(\max\left\{\|\mathbf{W}''_n \mathbf{S}_n v^*\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))}, \|\mathbf{W}''_n \mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))}\right\}\right)^{q-1} \\
&\qquad\qquad \left| \|\mathbf{W}''_n \mathbf{S}_n v^*\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} - \|\mathbf{W}''_n \mathbf{S}_n v\|_{\ell_q(\mathbf{K}_n;\ell_2(\mathbb{R}^{J^2}))} \right| \\
\le\ & q\left(C_1\|v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))} + C_1\|v^*-v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))}\right)^{q-1} \cdot C_1\|v^*-v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))} \\
\le\ & C''(v)\,\|v^*-v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))}
\end{aligned} \tag{3.8}$$

where $C''(v) = q\left(\|v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))} + 1\right)^{q-1} C_1^q$. By the same argument, we obtain

$$\begin{aligned}
&\left| \|\mathbf{W}'_n \mathbf{T}_n u^* - \mathbf{S}_n v^*\|^p_{\ell_p(\mathbf{K}_n);\ell_2(\mathbb{R}^J)} - \|\mathbf{W}'_n \mathbf{T}_n u - \mathbf{S}_n v\|^p_{\ell_p(\mathbf{K}_n);\ell_2(\mathbb{R}^J)} \right| \\
\le\ & C'(u,v)\left(\|u^*-u\|_{W^p_{2s}(\Omega)} + \|v^*-v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))}\right).
\end{aligned} \tag{3.9}$$

for $C'(u,v) = p\left(\|\mathbf{D}'u - v\|_{L_p(\Omega;\ell_2(\mathbb{R}^J))} + 1\right)^{p-1} C_2^p \le p\left(\|u\|_{W^p_{2s}(\Omega)} + \|v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))} + 1\right)^{p-1} C_2^p$. Therefore,

$$\begin{aligned}
&\left| \tilde{V}_n(u^*, v^*) - \tilde{V}_n(u,v) \right| \\
\le\ & C(u,v)\left(\|u^*-u\|_{W^p_{2s}(\Omega)} + \|v^*-v\|_{W^q_s(\Omega;\ell_2(\mathbb{R}^J))}\right),
\end{aligned} \tag{3.10}$$

where $C(u,v) = \nu_1 C'(u,v) + \nu_2 C''(v)$. Since $C = C(u,v)$ does not depend on n, we can conclude that $\tilde{V}_n$ is equicontinuous.

References

[1] J. Cai, B. Dong, S. Osher, and Z. Shen, "Image restorations: total variation, wavelet frames and beyond," *Journal of American Mathematical Society*, vol. 25(4), pp. 1033–1089, 2012.

[2] B. Dong, Q. Jiang, and Z. Shen, "Image restoration: Wavelet frame shrinkage, nonlinear evolution pdes, and beyond," *UCLA CAM Report*, vol. 13-78, 2013.

[3] J. Cai, B. Dong, and Z. Shen, "Image restorations: a wavelet frame based model for piecewise smooth functions and beyond," *Applied and Computational Harmonic Analysis*, vol. 41(1), pp. 94–138, 2016. http://dx.doi.org/10.1016/j.acha.2015.06.009.

[4] K. Bredies, K. Kunisch, and T. Pock, "Total Generalized Variation," *SIAM Journal on Imaging Sciences*, vol. 3, p. 492, 2010.

[5] K. Gröchenig, *Foundations of time-frequency analysis*. Birkhauser, 2001.

[6] S. Mallat, *A wavelet tour of signal processing*, vol. 2nd ed. New York: Academic. Academic press, 1999.

[7] A. Ron and Z. Shen, "Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator," *Journal of Functional Analysis*, vol. 148, no. 2, pp. 408–447, 1997.

[8] E. Candes and D. Donoho, "Continuous curvelet transform:: II. Discretization and frames," *Applied and Computational Harmonic Analysis*, vol. 19, no. 2, pp. 198–222, 2005.

[9] E. Candes and D. Donoho, "New tight frames of curvelets and optimal representations of objects with C2 singularities," *Comm. Pure Appl. Math*, vol. 56, pp. 219–266, 2004.

[10] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on image processing*, vol. 14, no. 12, pp. 2091–2106, 2005.

[11] E. Le Pennec and S. Mallat, "Sparse geometric image representations with bandelets," *IEEE transactions on image processing*, vol. 14, no. 4, pp. 423–438, 2005.

[12] D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," in *Optics & Photonics 2005*, pp. 59140U–59140U, International Society for Optics and Photonics, 2005.

[13] B. Han and Z. Zhao, "Tensor product complex tight framelets with increasing directionality," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 997–1034, 2014.

[14] H. Ji, Z. Shen, and Y. Zhao, "Directional frames for image recovery: Multi-scale discrete gabor frames," *accepted by Journal of Fourier Analysis and Applications*, 2016.

[15] I. Daubechies, G. Teschke, and L. Vese, "Iteratively solving linear inverse problems under general convex constraints," *Inverse Problems and Imaging*, vol. 1, no. 1, p. 29, 2007.

[16] M. Fadili and J. Starck, "Sparse representations and Bayesian image inpainting," *Proc. SPARS*, vol. 5, 2005.

[17] M. Fadili, J. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *The Computer Journal*, vol. 52, no. 1, p. 64, 2009.

[18] M. Figueiredo and R. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, 2003.

[19] M. Figueiredo and R. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 2, pp. II–782, IEEE, 2005.

[20] J. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, vol. 8, no. 2, pp. 337–369, 2009.

[21] M. Elad, J. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, 2005.

[22] J. Starck, M. Elad, and D. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE transactions on image processing*, vol. 14, no. 10, pp. 1570–1582, 2005.

[23] R. Chan, T. Chan, L. Shen, and Z. Shen, "Wavelet algorithms for high-resolution image reconstruction," *SIAM Journal on Scientific Computing*, vol. 24, no. 4, pp. 1408–1432, 2003.

[24] J. Cai, R. Chan, L. Shen, and Z. Shen, "Convergence analysis of tight framelet approach for missing data recovery," *Advances in Computational Mathematics*, vol. 31, no. 1, pp. 87–113, 2009.

[25] J. Cai, R. Chan, and Z. Shen, "Simultaneous cartoon and texture inpainting," *Inverse Problems and Imaging (IPI)*, vol. 4, no. 3, pp. 379–395, 2010.

[26] B. Dong, H. Ji, J. Li, Z. Shen, and Y. Xu, “Wavelet frame based blind image inpainting,” *accepted by Applied and Computational Harmonic Analysis*, vol. 32, no. 2, pp. 268–279, 2011.

[27] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D*, vol. 60, pp. 259–268, 1992.

[28] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[29] G. Dal Maso, *Introduction to* Γ*-convergence*. Birkhauser, 1993.

[30] T. Chan, S. Esedoglu, and F. Park, “Image decomposition combining staircase reduction and texture extraction,” *Journal of Visual Communication and Image Representation*, vol. 18, no. 6, pp. 464–486, 2007.

[31] T. Chan, S. Esedoglu, and F. Park, “A fourth order dual method for staircase reduction in texture extraction and image restoration problems,” *UCLA CAM Report*, pp. 05–28, 2005.

[32] S. Setzer, G. Steidl, and T. Teuber, “Infimal convolution regularizations with discrete l1-type functionals,” *Comm. Math. Sci*, vol. 9, no. 3, pp. 797–872, 2011.

[33] S. Setzer and G. Steidl, “Variational methods with higher order derivatives in image processing,” *Approximation*, vol. 12, pp. 360–386, 2008.

[34] S. Setzer, G. Steidl, B. Popilka, and B. Burgeth, “Variational methods for denoising matrix fields,” pp. 341–360, 2009.

[35] M. Holler and K. Kunisch, “On infimal convolution of tv-type functionals and applications to video and image reconstruction,” *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2258–2300, 2014.

[36] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[37] D. Bertsekas and J. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., 1989.

[38] J. Eckstein and D. Bertsekas, “On the DouglasRachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.

[39] T. Goldstein and S. Osher, “The split Bregman algorithm for L1 regularized problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[40] A. Ron and Z. Shen, “Affine systems in $L_2(\mathbb{R}^d)$ ii: dual systems,” *Journal of Fourier Analysis and Applications*, vol. 3, no. 5, pp. 617–638, 1997.

[41] I. Daubechies, *Ten lectures on wavelets*, vol. CBMS-NSF Lecture Notes, SIAM, nr. 61. Society for Industrial and Applied Mathematics, 1992.

[42] I. Daubechies, B. Han, A. Ron, and Z. Shen, “Framelets: MRA-based constructions of wavelet frames,” *Applied and Computational Harmonic Analysis*, vol. 14, pp. 1–46, Jan 2003.

[43] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[44] Z. Shen, "Wavelet frames and image restorations," in *Proceedings of the International Congress of Mathematicians*, vol. 4, pp. 2834–2863, 2010.

[45] B. Dong and Z. Shen, "Image restoration: a data-driven perspective.," *Proceedings of the International Congress of Industrial and Applied Mathematics (ICIAM)*, pp. 65–108, 2015.

[46] B. Dong and Z. Shen, "MRA-Based Wavelet Frames and Applications," *IAS Lecture Notes Series, Summer Program on "The Mathematics of Image Processing", Park City Mathematics Institute*, 2010.

[47] R. Adams, *Sobolev Spaces.* Academic Press, New York, 1975.