

IMAGE RESTORATION: A WAVELET FRAME BASED MODEL FOR PIECEWISE SMOOTH FUNCTIONS AND BEYOND

JIAN-FENG CAI, BIN DONG, AND ZUOWEI SHEN

ABSTRACT. In this paper, we propose a new wavelet frame based image restoration model that explicitly treats images as piecewise smooth functions. It estimates both the image to be restored and its singularity set. It can well protect singularities, which are important image features, and provide enough regularization in smooth regions at the same time. This model penalizes the ℓ_2 -norm of the wavelet frame coefficients away from the singularity set, while penalizes the ℓ_1 -norm of the coefficients on the singularity set. This model explicitly models images as piecewise smooth functions with a general smoothness regularization and characterizes rather general singularity set, which includes both jump discontinuities and jumps after certain orders of differentiations. As we know, all types of singularities are important image features and need to be recovered. Furthermore, the singularity set can be robustly estimated by wavelet frame transform during the image recovery procedure, which makes our model easy to solve numerically; hence, the model is insensitive to the estimation of the singularity set.

The proposed model is in discrete setting and is a wavelet frame based approach. To further understand the piecewise smooth nature of the obtained solutions, we connect it to a variational model on the space of piecewise smooth functions and prove rigorously that the discrete model converges to the variational model as image resolution goes to infinity. Also, we show that the approximate solutions of the discrete model can be regarded as an approximation of those of the variational model. Through these theoretical analysis, we manage to connect the proposed discrete wavelet frame based model with the variational model. Such connection not only enables us to acquire deeper understandings of the discrete model, but also leads us to the discovery of a variational model new to the literature, which is more general and works better than the Mumford-Shah model [1] for image restoration problems. Although the focus of the paper is to propose the new model and provide theoretical studies, we still conduct numerical simulations to support our claims and theoretical findings. Our numerical studies show that the proposed model is the right one for image restorations, when the underlying solutions are piecewise smooth. Generally speaking, this model combines the merits of the PDE based approach [1–7] and the wavelet frame based approach [8–10].

1. INTRODUCTION

Image restoration, including image denoising, deblurring, inpainting, computed tomography, etc., is one of the most important areas in imaging science. Its major purpose is to enhance the quality of a given image that is corrupted in various ways during the process of imaging, acquisition and communication, and enable us to see crucial but subtle objects that reside in the image. Mathematics has become one of the main driving forces of the modern development of image restoration. There are several mathematics based approaches, the partial differential equation (PDE) based approach (e.g. variational methods and PDEs models), and wavelet frame based approach developed in the last few decades are successful examples among many.

Image restoration problems can be casted as solving a linear inverse problem

$$(1.1) \quad \mathbf{f} = \mathbf{A}\mathbf{u} + \boldsymbol{\eta}$$

2010 *Mathematics Subject Classification.* 35A15, 42C40, 45Q05, 65K10, 68U10.

Key words and phrases. Image restoration, piecewise smooth functions, (tight) wavelet frames, framelets, Mumford-Shah model, split Bregman, pointwise convergence, Γ -convergence.

Corresponding Author: Bin Dong (dongbin@math.pku.edu.cn).

Jian-Feng Cai is supported in part by NSF DMS-1418737. Bin Dong is supported in part by NSF DMS-1418772.

where the matrix \mathbf{A} is some linear operator (not invertible in general) and $\boldsymbol{\eta}$ denotes a perturbation caused by the additive noise in the observed image (or measurements), which is typically assumed to be white Gaussian noise. Different image restoration problem corresponds to different type of \mathbf{A} , e.g., the identity operator for image denoising, a restriction operator for inpainting, a convolution operator for image deconvolution, partial Radon transform for CT imaging, partial Fourier transform for MR Imaging, etc. The problem (1.1) is usually ill-posed, which makes solving (1.1) nontrivial.

In order to obtain a high quality recovery of the unknown image from the ill-posed linear inverse problem (1.1), we need a proper modeling of images. The modeling should be different for different classes of images, such as natural images, biological images, textures, etc. Here, just like many previous work in the literature, we shall focus on natural images, including photographs of architecture, landscaping, portrait, etc. A fundamental question one should always ask before going into the image restoration process is: what is an appropriate model/descriptor for images. In variational and nonlinear PDE based modeling, such as the well-known Rudin-Osher-Fatemi (ROF) model [2], images are described by functions of bounded variations (BV). For wavelet or wavelet frame based methods (see e.g. [9, 11, 12]) with the regularization term being the ℓ_1 -norm of the wavelet frame coefficients, images are essentially approximated by functions in the homogenous Besov space $B_{1,1}^1$. The BV space is a fairly large function space. It is known (see e.g. [6]) that a BV function can always be decomposed to the sum of an absolute continuous part with respect to the Lebesgue measure, a jump part (e.g. edges), and a Cantor measure. However, since images always have a limited resolution, we normally cannot observe the Cantor measure in images. On the other hand, the Besov space $B_{1,1}^1$ sometimes is not large enough to include some natural images.

In this paper, we model images as piecewise smooth functions, and we propose a new wavelet frame based image restoration model that seeks piecewise smooth solutions to the linear inverse problem (1.1). The proposed model estimates both the image to be restored and its singularity set (which shall be simply called jump set throughout the rest of this paper), so that it can well protect singularities, which are important image features, and provide enough regularization in smooth regions at the same time. The proposed model combines the merits of the PDE based approach [1–5, 7] and the wavelet frame based approach [8–10]. Here, we provide a first glance of the model while the detailed definition and analysis of it are available in Section 3–5:

$$(1.2) \quad \inf_{\mathbf{u}, \Gamma} \|\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}\|_{\Gamma^c}^2 + \|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1 + \frac{1}{2}\|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2,$$

where \mathbf{W} is the wavelet frame transformation, \mathbf{u} is the image to be recovered and Γ is the estimated jump sets. When $\Gamma = \emptyset$, this model reduces to the standard Tikhonov regularization method. When $\Gamma^c = \emptyset$, this model reduces to the wavelet frame based analysis model of [9]. Although the Tikhonov regularization method keeps regularity of smooth part of images, it smears out the edges in images. On the other hand, the sparsity based wavelet frame analysis model of [9] preserves the edges by pursuing a sparse approximation of the underlying solution through ℓ_1 -regularization in wavelet frame domain, while it can also introduce artifact, or unwanted singularities, in smooth regions of images.

The proposed model (1.2) is to take advantage of the fact that the ℓ_2 -regularization in wavelet frame domain to keep smooth components of images smooth, and the ℓ_1 -regularization to keep the edges sharp. The use of the ℓ_2 -norm in the term $\|\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}\|_{\Gamma^c}^2$ is because \mathbf{u} is supposed to be smooth away from its jump set. Due to the present of wavelet frames with multiple orders of vanishing moments, this term in fact corresponds to general regularization term for the smooth part of image. The term $\|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1$ regularizes both *jumps* (jump discontinuities) and *hidden jumps* (jump discontinuities after certain orders of differentiations) at Γ , so that it preserves the singularity set of the image, and, at the same time, promote regularity along the singularity set of the image.

The key to the success of the model (1.2) is that there is a fast and robust way to determine the jump set Γ , since the large wavelet frame coefficients indicate the locations of jump set. The jump discontinuities are naturally extracted by wavelet frames with vanishing moments of order 1, while the hidden jumps are extracted by wavelet frames with higher order of vanishing moments. Since the high pass filters associated to wavelet frame functions can automatically detect the jump set of various types (jumps and hidden jumps), the set Γ can be obtained iteratively in a numerically simple way, as will be shown in later parts of this paper. Furthermore, model (1.2) is not sensitive to the estimation of Γ , because the specific form of the second term $\|[\gamma \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1$. Indeed, when $\Gamma^c = \emptyset$, our model (1.2) reduces to the analysis based model [9, 13, 14] which is proven in the literature to be a good image restoration model. Intuitively, the reconstruction results of our model (1.2) cannot be worse than those of the analysis based model. In fact, our numerical experiments in Section 3.3 will show that if Γ is more informative than $\Gamma^c = \emptyset$, we can significantly improve the image restoration results for images that can be well modeled as piecewise smooth functions. Finally, the built-in multiscale structure of wavelet frame systems enables a reliable estimation of the jump set at the present of noise.

In the literature of mathematical modeling in image restoration, PDE based approach and wavelet frame based approach are successful examples among many. In this paper, we shall refer to both variational models and the (nonlinear) PDE based models as PDE based approach. Although both types of models view image restoration problems from different perspectives, they are closely related. Indeed, at the beginning of variational modeling for image restoration, variational models are solved via PDEs (e.g. gradient flows). In general, variational techniques are frequently used to aid the analysis of the associated PDEs. Therefore, we shall classify both variational modeling and PDE based modeling as PDE based approach. In general, when we mention the PDE based approach, we are including the variational models such as the refined Rudin-Osher-Fatemi (ROF) model [2] and the well-known Mumford-Shah model [1]; and nonlinear PDE models such as the anisotropic diffusion model by Perona and Malik [3]. In PDE based approach, images are essentially assumed as functions in the space of functions with bounded variation (BV). These ground-breaking models started the modern trend of mathematical imaging sciences and have inspired lots of interesting and exciting developments in the field ever since (see, e.g., [4–7, 15]). Almost all PDE based approaches for image restoration aim at preserving or enhancing image features, such as edges, while regularizing smooth image components at the same time. Wavelet frame based approach, on the other hand, achieves a similar objective as the PDE based approach by promoting the sparsity of the wavelet frame coefficients of the images via shrinkage operators so that singularities such as edges can be preserved/enhanced.

The wavelet and wavelet frame based image restoration models with the ℓ_1 -norm regularization implicitly approximates images by functions in a certain Besov space, such as $B_{1,1}^1$. The wavelet frame based image processing started from [8, 16] for high-resolution image reconstructions and was later generalized by [17, 18] which lead to an effective image restoration model known as the balanced model. The balanced model includes two other wavelet frame based models as special cases. One is known as the synthesis based model [11, 12, 19–21], and the other is known as the analysis based model, [9, 13, 14]. The three approaches are different from each other, unless the underlying wavelet frame systems is in fact orthonormal/biorthogonal. However, what they have in common, is the penalization of the sparsity of the wavelet frame coefficients of the image to be restored to enhance the features of images. In recent years, variational models are discretized first and then optimization algorithms are applied to solve the variational models in discrete settings. As shown by results in [22, 23] and this paper, implicitly, this approach essentially converts the PDE based approaches to various wavelet frame based approaches. This seems to make the PDE based models distant from variational methods. However, our recent work in [23] indicates that some such numerical optimization algorithms for variational models are closely related to (nonlinear) PDEs in the discrete setting on the other hand.

Regardless of the different spaces of images that the PDE based approach and wavelet frame based approach may assume, both approaches eventually want to model images as piecewise smooth functions. To properly recover a piecewise smooth function, we need to explicitly estimate its jump set, while regularize the function as much as possible elsewhere. However, most of the widely used image restoration models do not model jump set explicitly due to the computation challenge to simultaneously recover the jumps sets and the image. Instead, these models make use of the ℓ_1 -norm (or even the ℓ_0 -norm) and the sparsity of the image under certain transformation to maintain sharpness of edges while regularizing smooth image components. However, such treatment cannot always guarantee sharpness of edges and image smoothness at the same time since the ℓ_1 -norm treats each pixel location of an image equally which means that each pixel has a chance to be selected into the jump set. In the literature, a variational model that attempted to model images as piecewise smooth function explicitly is the Mumford-Shah model [1]. Mumford-Shah model treats images as the so-called special functions with bounded variation (SBV) [24], which are BV functions without the component of Cantor measure. The SBV space can be regarded as something in between of $B_{1,1}^1$ and BV space. However, SBV is still not a space large enough to include most natural images inside. For one thing, images may be much smoother than merely absolute continuous away from jump discontinuities. Furthermore, the Mumford-Shah model is difficult to solve numerically mainly due to the sensitivity of the image solutions to the estimation of the jump set.

To further understand the piecewise smooth nature of the underlying solution of the proposed model (1.2), we shall analyze its asymptotic property as image resolution goes to infinity. We will prove that there is a variational model (see (5.2)) for piecewise smooth functions, to which the discrete model (1.2) converges. Also, through the convergence analysis, the approximate solutions of (1.2) can be regarded as discrete approximations of the solutions to the variational model under suitable assumptions. Through these theoretical analysis, we are managed to connect the discrete wavelet frame based model with the continuum variational model. Such connection not only enables us to acquire deeper understandings of the discrete model, but also leads us to the discovery of a variational model new to the literature. As a consequence of our analysis, we are able to link this new wavelet frame based model to the Mumford-Shah model.

The analysis in this paper is motivated by our earlier work [22, 23]. In [22], a fundamental connection between a wavelet frame based approach, namely the analysis based approach, and a generic variation model (which takes the ROF model as a special case) were established. It was shown in [22] that the analysis based model using generic wavelet frame systems can be regarded as an approximation of the variational model in the discrete setting. More recently in [23], a fundamental connection between generic wavelet frame shrinkage and nonlinear evolution PDEs (which takes the PM (Perona-Malik) equation and many other classical PDE models as special cases) were established. In [23], interpretations and analytical studies of such connection were provided and new algorithms for image restoration were proposed based on the new understandings. From our work in [22, 23] and that of the current paper, we have a fairly complete picture of how wavelet frame based approach is connected with the PDE based approach. The connections we have established in the three papers automatically give wavelet frame based approach a geometric explanation through variational models and nonlinear PDEs; and at the same time, they enable us to equip the PDE based approach with a multi-scaled time frequency analysis. In particular, wavelet frames can be used as a new and useful tool in numerical analysis to discretize and solve various variational and PDE models. Furthermore, such connections lead to new and inspiring interpretations of both approaches which enables us to create novel and more effective PDE and wavelet frame based models/algorithms for image restoration problems.

The rest of the paper is organized as follows. In Section 2, we present a brief review of wavelet frame theory and introduce some basic concepts of wavelet frame that will be used in later sections. We propose the discrete wavelet frame based model and its associated algorithm in Section 3. At the end of this section, we conduct numerical experiments on the proposed model and algorithm

which will show great potential of our proposed approach. In Section 4, we present a variational model for piecewise smooth functions to which the proposed discrete model converges. In Section 5, we rigorously prove that, under suitable assumptions, the proposed discrete wavelet frame based model does converge to the variational model as the resolution goes to infinite. As a result of our analysis, we can see how the (approximate) solutions of the discrete model approximate those of the variational model. Through such analysis, we can also see that our discrete model does acquire piecewise smooth solutions of the associated linear inverse problem in the discrete setting.

2. PRELIMINARIES OF WAVELET FRAMES

In this subsection, we briefly introduce the concept of tight frames and tight wavelet frames. The interested readers should consult [25–27] for theories of frames and wavelet frames, [28] for a short survey on the theory and applications of frames, and [10] for a more detailed survey.

A countable set $X \subset L_2(\mathbb{R}^d)$, with $d \in \mathbb{Z}^+$, is called a tight frame of $L_2(\mathbb{R}^d)$ if

$$(2.1) \quad f = \sum_{g \in X} \langle f, g \rangle g \quad \forall f \in L_2(\mathbb{R}^d),$$

where $\langle \cdot, \cdot \rangle$ is the inner product of $L_2(\mathbb{R}^d)$.

For given $\Psi := \{\psi_1, \dots, \psi_L\} \subset L_2(\mathbb{R}^d)$, the corresponding quasi-affine system $X(\Psi)$ generated by Ψ is defined by the collection of the dilations and the shifts of Ψ as

$$(2.2) \quad X(\Psi) = \{\psi_{\ell, n, \mathbf{k}} : 1 \leq \ell \leq L; n \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^d\},$$

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

$$(2.3) \quad \psi_{\ell, n, \mathbf{k}} := \begin{cases} 2^{\frac{nd}{2}} \psi_{\ell}(2^n \cdot -\mathbf{k}), & n \geq 0; \\ 2^{nd} \psi_{\ell}(2^n \cdot -2^n \mathbf{k}), & n < 0. \end{cases}$$

When $X(\Psi)$ forms a (tight) frame of $L_2(\mathbb{R}^d)$, each function ψ_{ℓ} , $\ell = 1, \dots, L$, is called a (tight) framelet and the whole system $X(\Psi)$ is called a (tight) wavelet frame system. Note that in the literature, the affine (or wavelet) system is commonly used, which corresponds to the decimated wavelet (frame) transforms. The quasi-affine system, which corresponds to the undecimated wavelet (frame) transforms, was first introduced and analyzed by [25]. Here, we only discuss the quasi-affine system (2.3), since it works better in image restoration and its connection to PDEs is more natural than the affine system. The interested reader can find further details on the affine wavelet frame systems and its relation to the quasi-affine frames in [10, 25, 29].

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{a}_0 satisfying

$$(2.4) \quad \phi = 2^d \sum_{\mathbf{k} \in \mathbb{Z}^d} \mathbf{a}_0[\mathbf{k}] \phi(2 \cdot -\mathbf{k}).$$

The idea of an MRA-based construction of framelets $\Psi = \{\psi_1, \dots, \psi_q\} \subset L_2(\mathbb{R}^d)$ is to find masks \mathbf{a}_{ℓ} , which are finite sequences, such that

$$(2.5) \quad \psi_{\ell} = 2^d \sum_{\mathbf{k} \in \mathbb{Z}^d} \mathbf{a}_{\ell}[\mathbf{k}] \phi(2 \cdot -\mathbf{k}), \quad \ell = 1, 2, \dots, q.$$

The sequences $\mathbf{a}_1, \dots, \mathbf{a}_q$ are called wavelet frame masks, or the high pass filters of the system, and the refinement mask \mathbf{a}_0 is also known as the low pass filter.

The unitary extension principle (UEP) of [25] provides a general theory of the construction of MRA-based tight wavelet frames. Roughly speaking, as long as $\{\mathbf{a}_1, \dots, \mathbf{a}_q\}$ are finitely supported

and their Fourier series satisfy

$$(2.6) \quad \sum_{\ell=0}^q |\widehat{\mathbf{a}}_{\ell}(\xi)|^2 = 1 \quad \text{and} \quad \sum_{\ell=0}^q \widehat{\mathbf{a}}_{\ell}(\xi) \overline{\widehat{\mathbf{a}}_{\ell}(\xi + \nu)} = 0,$$

for all $\nu \in \{0, \pi\}^d \setminus \{\mathbf{0}\}$ and $\xi \in [-\pi, \pi]^d$, the quasi-affine system $X(\Psi)$ (as well as the traditional wavelet system) with $\Psi = \{\psi_1, \dots, \psi_q\}$ defined by (2.5) forms a tight frame in $L_2(\mathbb{R}^d)$.

We now show two simple but useful examples of univariate framelets. The framelet given in Example 2.1 is known as the Haar wavelet. When one uses a wavelet (affine) system, it generates an orthonormal basis of $L_2(\mathbb{R})$. The quasi-affine system that the Haar wavelet generates, however, is not an orthonormal basis, but a tight frame of $L_2(\mathbb{R})$ instead. We shall refer to ψ_1 in Example 2.1 as the ‘‘Haar framelet’’. The framelets given by Example 2.2 is constructed from piecewise linear B-spline which was first given in [25]. The masks of all the B-splines framelets constructed by [25] are exactly discrete difference operators up to a scaling. These framelets, especially the ones in Example 2.2, are widely used in frame based image restoration problems because they provide sparse approximations to piecewise smooth functions such as images (see, e.g., [8, 9, 16–18, 29–33]). We shall refer to ψ_1 and ψ_2 in Example 2.2 as ‘‘piecewise linear framelets’’. For a comprehensive introduction to B-splines, we refer the interested readers to [34].

Example 2.1. Let $\mathbf{a}_0 = \frac{1}{2}[1, 1]$ be the refinement mask of the piecewise constant B-spline $B_1(x) = 1$ for $x \in [0, 1]$ and 0 otherwise. Define $\mathbf{a}_1 = \frac{1}{2}[1, -1]$. Then \mathbf{a}_0 and \mathbf{a}_1 satisfy (2.6). Hence, the system $X(\psi_1)$ defined in (2.2) is a tight frame of $L_2(\mathbb{R})$. The mask \mathbf{a}_1 corresponds to a first order difference operator up to a scaling.

Example 2.2. Let $\mathbf{a}_0 = \frac{1}{4}[1, 2, 1]$ be the refinement mask of the piecewise linear B-spline $B_2(x) = \max(1 - |x|, 0)$. Define $\mathbf{a}_1 = \frac{\sqrt{2}}{4}[1, 0, -1]$ and $\mathbf{a}_2 = \frac{1}{4}[-1, 2, -1]$. Then \mathbf{a}_0 , \mathbf{a}_1 and \mathbf{a}_2 satisfy (2.6). Hence, the system $X(\Psi)$ where $\Psi = \{\psi_1, \psi_2\}$ defined in (2.2) is a tight frame of $L_2(\mathbb{R})$. The masks \mathbf{a}_1 and \mathbf{a}_2 correspond to the first order and second order difference operators respectively up to a scaling.

For practical concerns, we need to consider tight frames of $L_2(\mathbb{R}^d)$ with $d = 2$ or 3, since a typical image is a discrete function with its domain in 2 or 3 dimensional space. In this paper, we shall consider the case $d = 2$. One way to construct tight frames for $L_2(\mathbb{R}^2)$ (or $L_2(\mathbb{R}^d)$ in general) is by taking *tensor products* of univariate tight frames. Given a set of univariate masks $\{\mathbf{a}_{\ell} : \ell = 0, 1, \dots, r\}$, define the 2D masks $\mathbf{a}_{\mathbf{i}}[\mathbf{k}]$, with $\mathbf{i} := (i_1, i_2)$ and $\mathbf{k} := (k_1, k_2)$, as

$$(2.7) \quad \mathbf{a}_{\mathbf{i}}[\mathbf{k}] := \mathbf{a}_{i_1}[k_1] \mathbf{a}_{i_2}[k_2], \quad 0 \leq i_1, i_2 \leq r; (k_1, k_2) \in \mathbb{Z}^2.$$

Then the corresponding 2D refinable function and framelets are defined by

$$\psi_{\mathbf{i}}(x, y) = \psi_{i_1}(x) \psi_{i_2}(y), \quad 0 \leq i_1, i_2 \leq r; (x, y) \in \mathbb{R}^2,$$

where we have let $\psi_0 := \phi$ for convenience. We denote

$$\Psi := \{\psi_{\mathbf{i}}; 0 \leq i_1, i_2 \leq r; \mathbf{i} \neq (0, 0)\}.$$

If the univariate masks $\{\mathbf{a}_{\ell}\}$ are constructed from UEP, then it is easy to verify that $\{\mathbf{a}_{\mathbf{i}}\}$ satisfies (2.6) and thus $X(\Psi)$ is a tight frame for $L_2(\mathbb{R}^2)$.

Recall that there are m framelets with vanishing moment from 1, \dots , m constructed from the B-spline of order m by the UEP in [25]. Recall that the vanishing moment of a function is the order of the zero of its Fourier transform at the origin. We shall order the indices of the framelets according to their orders of vanishing moments. Note that the corresponding index for any B-spline is 0, since it has no vanishing moment. Then, for the tensor product framelet $\psi_{\mathbf{i}}$ with $\mathbf{i} = (i_1, i_2)$, if there is a differential operator associate with it, the differential operator should be $D_{\mathbf{i}}$, i.e. applying the i_1 derivative of the first variable and the i_2 derivative of the second variable. The order of vanishing moment of the framelet is $|\mathbf{i}|$ which is equal to the order of the differential operator $D_{\mathbf{i}}$. Such

association of framelets with differential operators was first discovered in [22] and was extensively used in [23]. The reason behind it will become more evident in later sections.

We now provide the masks of 2-dimensional Haar and piecewise linear framelets constructed by tensor product in the following example.

Example 2.3.

- (1) The tensor-product 2-dimensional Haar tight frame system has filters

$$\begin{aligned}\mathbf{a}_{0,0} &= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{a}_{0,1} = \frac{1}{4} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \\ \mathbf{a}_{1,0} &= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \mathbf{a}_{1,1} = \frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.\end{aligned}$$

- (2) The tensor-product 2-dimensional piecewise linear B-spline tight frame system has filters

$$\begin{aligned}\mathbf{a}_{0,0} &= \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}, \mathbf{a}_{0,1} = \frac{\sqrt{2}}{16} \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}, \mathbf{a}_{0,2} = \frac{1}{16} \begin{pmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{pmatrix}, \\ \mathbf{a}_{1,0} &= \frac{\sqrt{2}}{16} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}, \mathbf{a}_{1,1} = \frac{1}{8} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \mathbf{a}_{1,2} = \frac{\sqrt{2}}{16} \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{pmatrix}, \\ \mathbf{a}_{2,0} &= \frac{1}{16} \begin{pmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{pmatrix}, \mathbf{a}_{2,1} = \frac{\sqrt{2}}{16} \begin{pmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{pmatrix}, \mathbf{a}_{2,2} = \frac{1}{16} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}.\end{aligned}$$

In the discrete setting, let an image \mathbf{f} be a 2-dimensional array. We denote by

$$\mathcal{I}_2 := \mathbb{R}^{N_1 \times N_2}$$

the set of all 2-dimensional images. We will further assume that all images are square images, i.e. $N_1 = N_2 = N$ and they all have supports in the open unit d -dimensional cube $\Omega = (0, 1)^d$. Note that these assumptions are not essential, and all arguments and results in this paper can be easily extended to more general cases. We denote the 2-dimensional fast (discrete) framelet transform (see, e.g., [10]) with levels of decomposition L as

$$(2.8) \quad \mathbf{W}\mathbf{u} = \{\mathbf{W}_{l,i}\mathbf{u} : 0 \leq l \leq L-1, 0 \leq i_1, i_2 \leq r\}, \quad \mathbf{u} \in \mathcal{I}_2.$$

We denote the wavelet frame bands (high frequency bands) as $\mathbb{B} = \{\mathbf{i} : 0 \leq i_1, i_2 \leq r\} \setminus \{\mathbf{0}\}$. The fast framelet transform \mathbf{W} is a linear operator with $\mathbf{W}_{l,i}\mathbf{u} \in \mathcal{I}_2$ denoting the frame coefficients of \mathbf{u} at level l and band \mathbf{i} . Furthermore, we have

$$\mathbf{W}_{l,i}\mathbf{u} := \mathbf{a}_{l,i}[-\cdot] \circledast \mathbf{u},$$

where \circledast denotes the convolution operator with a certain boundary condition, e.g., periodic boundary condition, and $\mathbf{a}_{l,i}$ is defined as

$$(2.9) \quad \mathbf{a}_{l,i} = \tilde{\mathbf{a}}_{l,i} \circledast \tilde{\mathbf{a}}_{l-1,0} \circledast \dots \circledast \tilde{\mathbf{a}}_{0,0} \quad \text{with} \quad \tilde{\mathbf{a}}_{l,i}[\mathbf{k}] = \begin{cases} \mathbf{a}_i[2^{-l}\mathbf{k}], & \mathbf{k} \in 2^l\mathbb{Z}^2; \\ 0, & \mathbf{k} \notin 2^l\mathbb{Z}^2. \end{cases}$$

Notice that $\mathbf{a}_{0,i} = \mathbf{a}_i$.

We denote the inverse framelet transform as \mathbf{W}^\top , which is the adjoint operator of \mathbf{W} , and we will have the perfect reconstruction formula

$$\mathbf{u} = \mathbf{W}^\top \mathbf{W}\mathbf{u}, \quad \text{for all } \mathbf{u} \in \mathcal{I}_2.$$

We use \mathbf{H} and \mathbf{H}^\top to denote the decomposition and reconstruction using Haar framelets. We will also denote the general fast framelet decomposition and reconstruction as \mathbf{W}_n and \mathbf{W}_n^\top , whenever the image resolution level n becomes relevant.

3. WAVELET FRAME BASED MODEL FOR PIECEWISE SMOOTH FUNCTIONS

In this section, we present our wavelet frame based image restoration model for piecewise smooth functions. In this paper, we say a function is a piecewise smooth function if the function itself and its derivatives (especially the first order derivatives) belong to certain Sobolev spaces on sub-domains and have a relatively regular jump discontinuity set. In this section, the proposed model and algorithm are all in discrete setting, where all data are discrete arrays that are sampled from the underlying functions. The solutions we seek should be understood as sampled data from piecewise smooth functions. We will provide the precise definition of piecewise smooth functions and the sampling associated to the wavelet frame systems in Section 5, where we study the asymptotic property of the proposed discrete model as image resolution goes to infinity.

3.1. Image Restoration Model. The basic linear image restoration model is usually given as

$$(3.1) \quad \mathbf{f} = \mathbf{A}\mathbf{u} + \boldsymbol{\eta},$$

where \mathbf{A} is some linear operator (not invertible in general) mapping \mathcal{I}_2 into itself, e.g., the identity operator for image denoising, a convolution operator for image deconvolution, or partial Radon/Fourier transform for CT/MR imaging; and $\boldsymbol{\eta}$ denotes a perturbation caused by the additive noise in the observed image (or measurements), which is typically assumed to be a white Gaussian noise. Since the linear inverse problem (3.1) for image restoration is often ill-posed, which means \mathbf{A} usually has a large or even infinite condition number. Therefore, at the presence of noise $\boldsymbol{\eta}$, solving (3.1) without any further restrictions on the solution \mathbf{u} usually leads to very undesirable restored images. Hence, all modern image restoration models and algorithms enforce additional regularity requirements to the restored images.

The basic idea of wavelet frame based image restoration is to approximate a solution of (3.1) while maintaining desirable regularity of the solution via proper penalization of the wavelet frame coefficients. The regularity that is normally used in the literature is the minimization of the ℓ_1 -norm [8,9,11,12,16–21,29–33] or the ℓ_0 -norm [35–37] of the wavelet frame coefficients. Since wavelet frames can sparsely approximate piecewise smooth functions such as images, the penalization of the ℓ_1 - or ℓ_0 -norm has the effect of smoothing the restored image while maintaining key features such as edges. However, since neither of the ℓ_1 - nor ℓ_0 -norm penalization explicitly identifies the locations of singularity of images, neither of them is able to protect sharp image features (such as edges) and maintain smoothness away from these sharp features at the same time. As a result, simply applying ℓ_1 - or ℓ_0 -norm penalization may not always produce ideal image reconstruction results. For instance, the penalization of the ℓ_1 - or ℓ_0 -norm may introduce artifact, or in other words, unwanted singularities, in smooth parts of images. Tuning the regularization parameter(s) in the model may reduce these artifacts, while it may smear out edges (wanted singularities) at the same time.

Since images are better modeled as piecewise smooth functions, instead of passively using a sparsity promoting norm (such as the ℓ_1 - or ℓ_0 -norm) hoping the paradox between smoothness and sharpness can be resolved automatically, an ideal model should actively split images into smooth and rough (singularities) regions and use different norms that are most suitable to each of these regions. If we can estimate fairly accurately the locations of the singularities (sharp edges, ridges, etc.) in the image, we can easily avoid smoothing across singularities while maintaining smoothness in smooth regions, which will further improve image restoration quality. Here, we propose a wavelet frame based image restoration model that explicitly model images as piecewise smooth functions by simultaneously recovery of the corrupted image and estimation of the locations of the singularities. In other words, our proposed model is to obtain a piecewise smooth solution of the linear inverse problem (3.1).

We denote by $\mathbb{O}^2 := \{0, 1, \dots, N-1\}^2$ the set of indices of the $N \times N$ Cartesian grid that discretize the domain $\Omega = (0, 1)^2$. Recall that the space of all 2-dimensional array on the grid \mathbb{O}^2 is denoted as \mathcal{I}_2 . Let $\Gamma \subset \mathbb{O}^2$ be the set of singularities that needs to be estimated. We propose

our wavelet frame based image restoration model as

$$(3.2) \quad \inf_{\mathbf{u} \in \mathcal{I}_2, \Gamma \subset \mathbb{O}^2} \left\| [\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}]_{\Gamma^c} \right\|_2^2 + \|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1 + \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2,$$

where

$$\|[\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}]_{\Gamma^c}\|_2^2 := \sum_{\mathbf{k} \in \mathbb{O}^2 \setminus \Gamma} \sum_{l=0}^{L-1} \sum_{\mathbf{i} \in \mathbb{B}} \lambda_{l,\mathbf{i}}[\mathbf{k}] \left| (\mathbf{W}_{l,\mathbf{i}}\mathbf{u})[\mathbf{k}] \right|^2$$

and

$$\|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1 := \sum_{\mathbf{k} \in \Gamma} \left[\sum_{l=0}^{L-1} \left(\sum_{\mathbf{i} \in \mathbb{B}} \gamma_{l,\mathbf{i}}[\mathbf{k}] \left| (\mathbf{W}_{l,\mathbf{i}}\mathbf{u})[\mathbf{k}] \right|^2 \right)^{\frac{1}{2}} \right].$$

Note that the type of singularity in Γ that we shall specifically focus on is not only jump discontinuities (or simply jumps), but also hidden jumps which means jump discontinuities after certain orders of differentiations. Both jumps and hidden jumps are important image features. As will be rigorously analyzed in Section 5 that to extract jumps and hidden jumps, we only need to properly choose the parameter $\boldsymbol{\gamma}$ in the second term of (3.2). More specifically, the terms in $\|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1$ that correspond to the wavelet frame band $\{|\mathbf{i}| = 1 : \mathbf{i} \in \mathbb{B}\}$ can extract jumps, while the terms correspond to $\{|\mathbf{i}| = 2 : \mathbf{i} \in \mathbb{B}\}$ can extract first order hidden jumps (jumps after first order differentiations). We can easily extract higher order hidden jumps (jumps after higher order differentiation) by using a higher order B-spline wavelet frame system and properly choosing $\boldsymbol{\gamma}$ for $\{|\mathbf{i}| \geq 2 : \mathbf{i} \in \mathbb{B}\}$. However, we shall only focus on jumps and first order hidden jumps in this paper, since they are more important image features.

The proposed model (3.2) is particularly effective for images that can be well approximated by piecewise smooth functions. Many types of real images satisfy such assumption. Examples of these images will be shown in Section 3.3, where numerical studies are presented. We further note that textures cannot be modeled as piecewise smooth functions [15]. However, since textures are sparse under systems of oscillating patterns, such as local cosine basis, we can use a two-system model to handle textures gracefully (see e.g. [9, 13, 14, 17, 38]). We will not discuss details of a two-system version of the proposed model, since it is outside the scope of this paper. We shall focus on the single-system model (3.2) for images.

3.2. Algorithm for Model (3.2). We present an algorithm that approximates a solution of (3.2). Since this model is nonconvex, good initialization of \mathbf{u} and Γ may be preferable to get a satisfactory result. We delay the detailed choice of initialization to Section 3.3.

We propose the following alternative minimization algorithm for (3.2).

Alternative Minimization of (3.2). Let \mathbf{u}^0 and Γ^0 be some initial data. For $k = 1, 2, \dots$,

(1) Given Γ^{k-1} , compute \mathbf{u}^k by

$$\mathbf{u}^k = \arg \min_{\mathbf{u} \in \mathcal{I}_2} \left\| [\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}]_{(\Gamma^{k-1})^c} \right\|_2^2 + \|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma^{k-1}}\|_1 + \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2.$$

Note that \mathbf{u}^k can be solve by the split Bregman algorithm [9, 39–41], which is recently widely used in solving optimization problems in variational and wavelet frame based image restoration. For a given exterior iteration k , we see an approximation of \mathbf{u}^k by iterating the following split Bregman algorithm: let $\mathbf{d}^0 = \mathbf{b}^0 = \mathbf{0}$, for $j = 1, 2, \dots$

$$(3.3) \quad \begin{cases} \mathbf{u}^{k,j} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2 + \frac{\mu}{2} \|\mathbf{W}\mathbf{u} - \mathbf{d}^{j-1} + \mathbf{b}^{j-1}\|_2^2, \\ \mathbf{d}^j = \arg \min_{\mathbf{d}} \left\| [\boldsymbol{\lambda} \cdot \mathbf{d}]_{(\Gamma^{k-1})^c} \right\|_2^2 + \|[\boldsymbol{\gamma} \cdot \mathbf{d}]_{\Gamma^{k-1}}\|_1 + \frac{\mu}{2} \|\mathbf{d} - \mathbf{W}\mathbf{u}^{k,j} - \mathbf{b}^{j-1}\|_2^2, \\ \mathbf{b}^j = \mathbf{b}^{j-1} + (\mathbf{W}\mathbf{u}^{k,j} - \mathbf{d}^j). \end{cases}$$

(2) Given \mathbf{u}^k , estimate Γ^k by

$$\Gamma^k = \arg \min_{\Gamma \subset \mathbb{O}^2} \left\| \left[\boldsymbol{\lambda} \cdot \mathbf{W} \mathbf{u}^k \right]_{\Gamma^c} \right\|_2^2 + \left\| \left[\boldsymbol{\gamma} \cdot \mathbf{W} \mathbf{u}^k \right]_{\Gamma} \right\|_1.$$

It is easy to see that Γ^k has a closed form solution given as

$$\Gamma^k = \left\{ \mathbf{p} \in \mathbb{O}^2 : \sum_{l=0}^{L-1} \left(\sum_{\mathbf{i} \in \mathbb{B}} \gamma_{l,\mathbf{i}} \left| (\mathbf{W}_{l,\mathbf{i}} \mathbf{u}^k)[\mathbf{p}] \right|^2 \right)^{\frac{1}{2}} \leq \sum_{l=0}^{L-1} \sum_{\mathbf{i} \in \mathbb{B}} \lambda_{l,\mathbf{i}} \left| (\mathbf{W}_{l,\mathbf{i}} \mathbf{u}^k)[\mathbf{p}] \right|^2 \right\}.$$

3.3. Numerical Simulations. In this subsection, we conduct some numerical simulations using our proposed approach described in Section 3, and compare the results with the analysis based approach that has been widely used recently in image restoration. We note, however, that numerical simulation is not the emphasis of this paper.

We compare our proposed model (3.2) and the associated algorithm described in Section 3 with the following analysis based model solved by split Bregman algorithm [9, 39]:

$$\inf_{\mathbf{u} \in \mathcal{I}_2} \|\boldsymbol{\gamma} \cdot \mathbf{W} \mathbf{u}\|_1 + \frac{1}{2} \|\mathbf{A} \mathbf{u} - \mathbf{f}\|_2^2.$$

The analysis based model is recently widely used in image restoration with success. Comparing to the total variation based model, which is the other model that is widely used in image restoration, the analysis based model can generate higher quality images for various image restoration problems (see e.g. [9, 42, 43]). Furthermore, the analysis based model can be regarded as a special case of the proposed model (3.2). Therefore, we shall focus on comparisons of the proposed model with the analysis based model.

The specific image restoration problem we shall consider here is image deblurring. The reason we choose deblurring is because it is a fundamental while challenging image restoration problem that also has wide applications. To be more precise, the operator \mathbf{A} is taken to be the convolution operator with the kernel generated in MATLAB by “fspecial(‘gaussian’,15,1.5)”. Additive Gaussian noise (with standard deviation= 4 and images taking integer values in [0, 255]) is also added. To measure quality of the restored image, we use the PSNR value defined by

$$\text{PSNR} := -20 \log_{10} \frac{\|\mathbf{u} - \tilde{\mathbf{u}}\|_2}{N},$$

where \mathbf{u} and $\tilde{\mathbf{u}}$ are the original and restored images respectively, and N is the total number of pixels.

We initialize our algorithm by choosing $\mathbf{u}^0 = \mathbf{0}$. For Γ^0 , we first compute an edge function \mathbf{g} from the observed image \mathbf{f} by

$$\mathbf{g}_0[\mathbf{k}] := \sum_{0 \leq l \leq L-1, \mathbf{i} \in \mathbb{B}} |(\mathbf{W}_{l,\mathbf{i}} \mathbf{f})[\mathbf{k}]|^2 \quad \text{and} \quad \mathbf{g} = \mathbf{g}_0 / \|\mathbf{g}_0\|_{\infty}.$$

Then, we compute Γ^0 by

$$\Gamma^0 = \{ \mathbf{k} \in \mathbb{O}^2 : \mathbf{g}[\mathbf{k}] \geq \tau \}.$$

The wavelet frame system used to compute Γ^0 is the piecewise linear B-spline tight frame system given in (2) of Example 2.3 and the level of decomposition is taken to be one, i.e. $L = 1$. Throughout our numerical experiments, the threshold τ is fixed to be 0.1 (numerically, the reconstruction results of our approach is relatively stable w.r.t. the choice of τ).

Note that we use piecewise linear B-spline framelet system for both the analysis based model and our model (3.2). In order to achieve best image restoration quality (numerically), the level of decomposition of wavelet frame transform for the analysis based model is chosen to be 4, while the level of decomposition for our proposed models chosen to be 2. The parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ of (3.2) is chosen as $\lambda_{l,\mathbf{i}} = \lambda$ and $\gamma_{l,\mathbf{i}} = \gamma$ for all $0 \leq l \leq L - 1$ and $\mathbf{i} \in \mathbb{B}$, where λ and γ are scalars manually

chosen for optimal restoration quality. Note that the ratio λ/γ determines the estimation of the jump set Γ . The bigger is the ratio, the larger is the set of Γ .

3.3.1. *Synthetic Images.* We start with two synthetic images which are presented in Figure 1. The image on the left has one set of jump discontinuities that forms a centaur-shaped boundary; and the image on the right has multiple jump sets. The comparisons of our approach with the analysis based model solved by split Bregman algorithm is shown in Figure 2 and Figure 3. The estimated jump set Γ for both synthetic images are shown in Figure 4.

Our results shown in Figure 2 and Figure 3 indicate that the proposed approach is significantly better than the analysis based model which indicates the great potential of our proposed model (3.2). In fact, our proposed model is particularly good for images that have relatively sparse jump set with the rest of the regions smooth (such as the synthetic images we used here). In order to further support this claim, in the next subsection, we will choose some non-synthetic images as test images which has relatively sparse jump set. We note that for images that have relatively dense jump set, such as textures, the propose model has a comparable performance with the analysis based model. In fact, for textures, neither model is suitable since textures cannot be modeled as piecewise smooth functions.

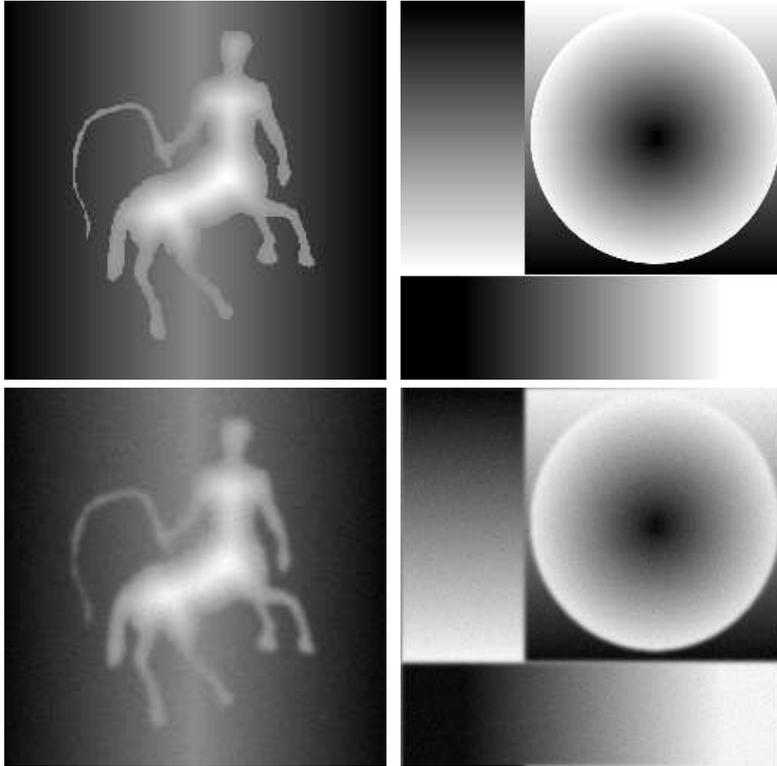


FIGURE 1. Synthetic images: original (first row) and observed (second row). The PSNR values of the observed images are **31.5046** and **24.6239** respectively.

3.3.2. *Non-Synthetic Images.* We downloaded several non-synthetic images from “google/image”, which are shown in Figure 5. We shall refer to these images (from left to right in Figure 5) as: “Car”, “Goldengate”, “Interior”, “Pitt” and “Samantha”. The comparison of our approach with the analysis based model solved by split Bregman algorithm is shown in Figure 6. The estimated jump set Γ is also given in Figure 6. The PSNR values of the reconstructed images for both methods are summarized in Table 1. As one can easily see from the table that our approach

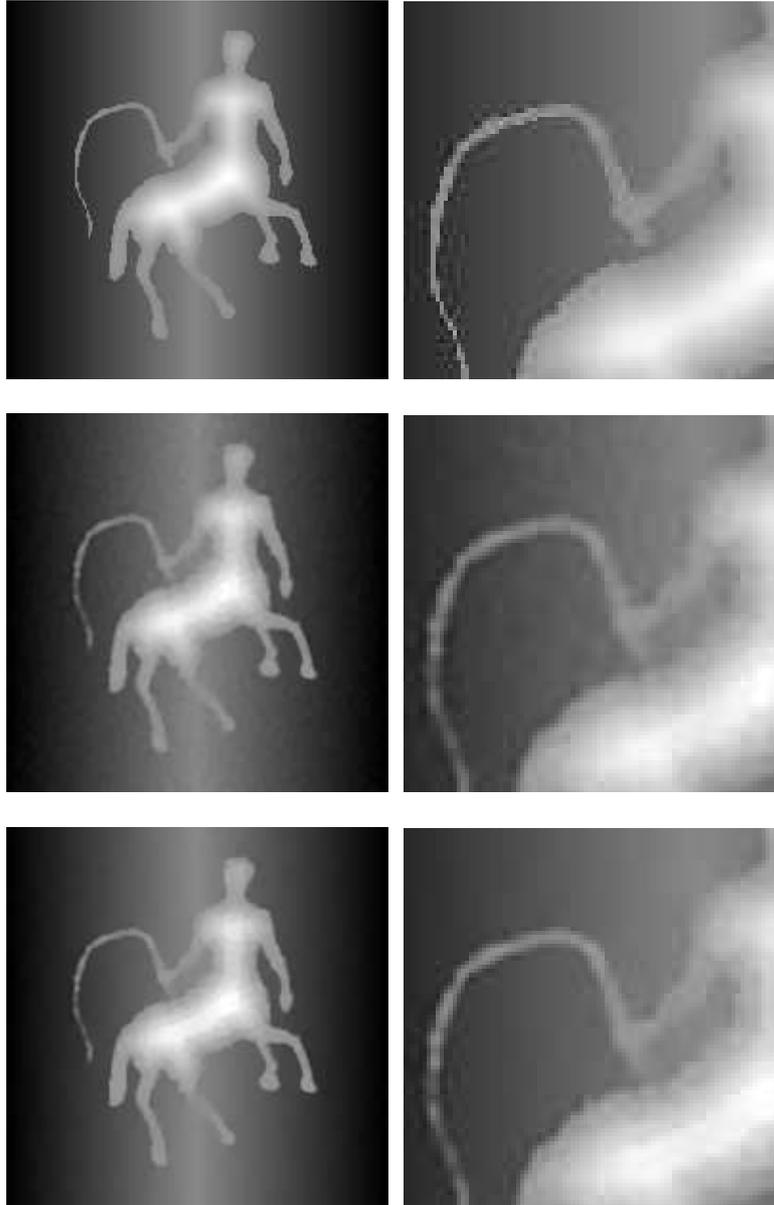


FIGURE 2. The first row of images are the observed image and one of its zoom-in view. The second row are the results of the analysis based model with PSNR=**35.2572**. The third row are the results of our proposed approach with PSNR=**35.9223**.

outperform the analysis based approach. The improvements of image restoration quality are also visually observable from Figure 6 and Figure 7.

To see the stability of the proposed model (3.2) in recovering \mathbf{u} and estimating $\mathbf{\Gamma}$, we tested our algorithm on deblurring of a relatively complex image data with three different noise levels with standard deviation = 2, 4, 6. Same blur kernel as before is used. Original and observed images are given in Figure 8. Results are presented in Figure 9. As we can see that, the restored image \mathbf{u} is gradually degrading as noise level increases. The estimated jump set $\mathbf{\Gamma}$ also degrades as noise level increases, but the degradation is rather slow. This shows the stability of the proposed model to increased noise level.

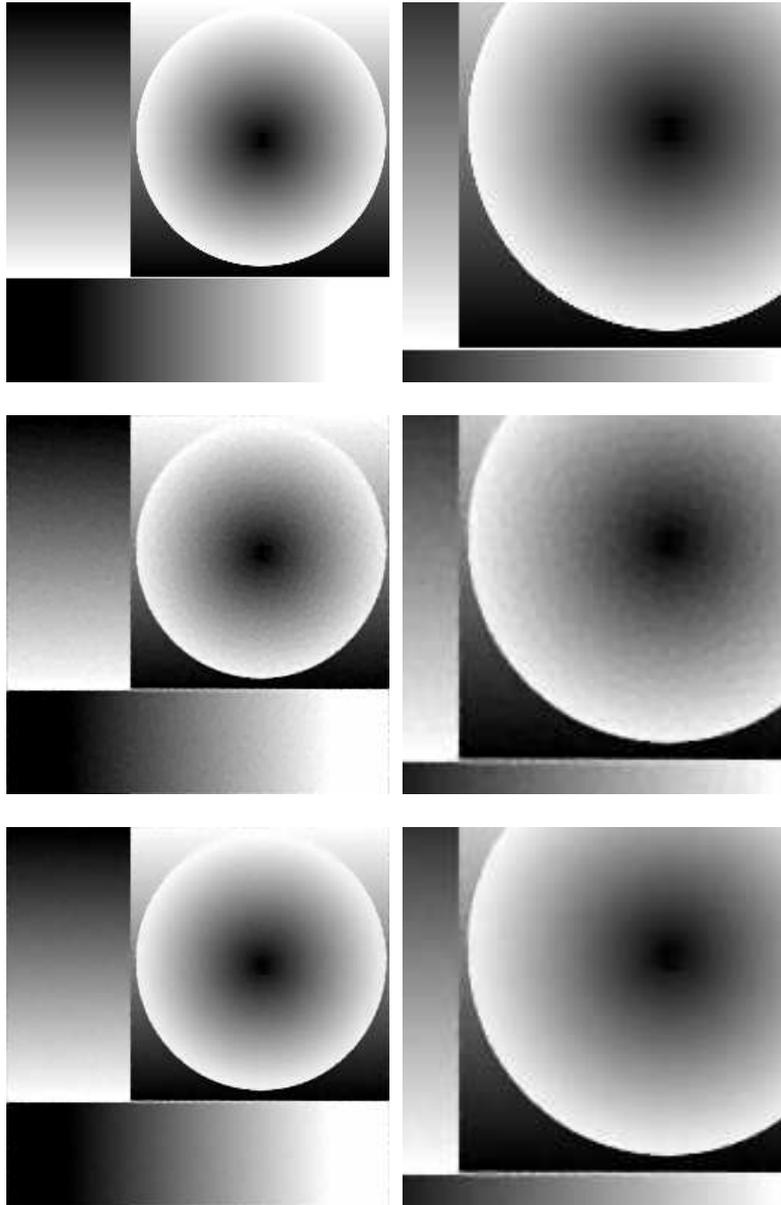


FIGURE 3. The first row of images are the observed image and one of its zoom-in view. The second row are the results of the analysis based model with PSNR=**31.7154**. The third row are the results of our proposed approach with PSNR=**34.2676**.

TABLE 1. Comparisons for image deconvolution using the images given in Figure 5.

Image Name	Analysis Based Model	Our Approach
Car	27.3194	27.5443
Goldgate	27.5312	27.8618
Interior	29.6087	30.0355
Pitt	29.4654	29.6716
Samantha	30.9207	31.0085

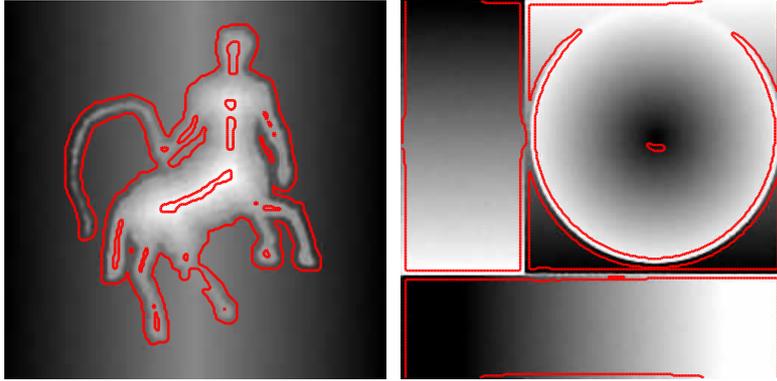


FIGURE 4. The estimated jump set Γ is enclosed in the red curves.

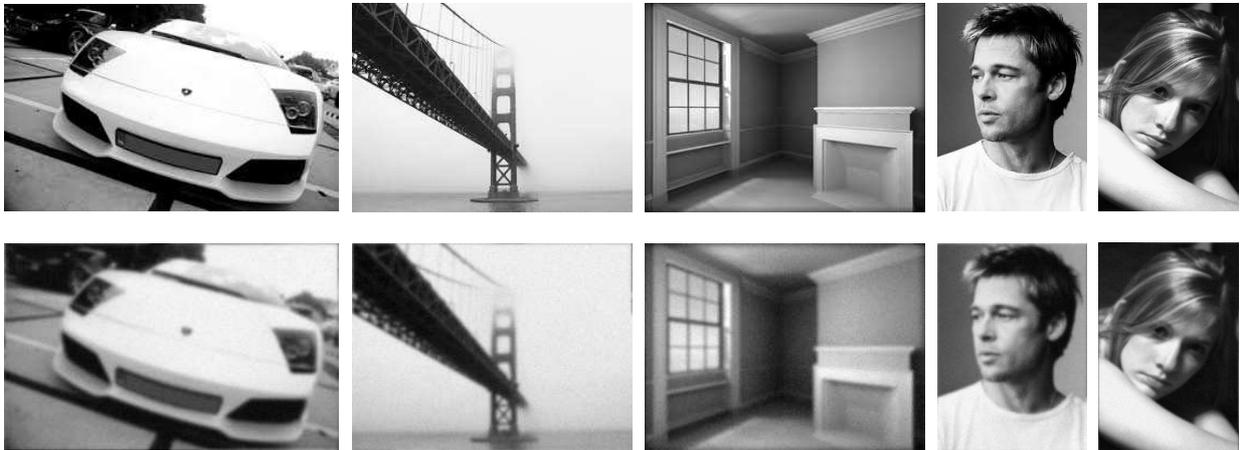


FIGURE 5. Non-synthetic images: original (first row) and observed (second row). The PSNR values of the observed images are **22.5976**, **24.1189**, **26.1993**, **25.9306** and **27.0862** respectively.

4. VARIATIONAL MODEL FOR PIECEWISE SMOOTH FUNCTIONS

Model (3.2) works well because, in wavelet domain, the large wavelet frame coefficients reflect the positions of jumps of different orders, while small wavelet frame coefficients normally reflect the “smooth components” of the images. Here, 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 given image data. Since the wavelet frame filters are designed to be standard difference operators with various orders, the locations of big wavelet frame coefficients indicate jumps of give images and their discrete differentiations. The locations of small wavelet frame coefficients indicate the region where image is smooth. It is natural to ask that, in which sense, this discrete model (3.2) does describe piecewise smooth functions, since in discrete setting, the concept of piecewise smoothness is unclear. Therefore, we need (and will) establish an asymptotic property of the model (3.2) as image resolution goes to infinity. In particular, we shall establish a variational model to which the discrete model (3.2) converges in a proper sense. Furthermore, this variational model clearly shows that it selects an optimal solution in a piecewise smooth function space. Finally, the detailed analysis shows that model (3.2) can be viewed as an approximation of this variational model. Consequently, it implies that model (3.2) is designed for piecewise smooth functions.



FIGURE 6. The first row of images are the ground truth images. The second row of images are the observed noisy and blurry images. The third row of images are the results of the analysis based model. The fourth row of images are the results of our proposed approach. The last row shows the computed jump set Γ .

This section focus on the introduction of the variational model, a precise definition of piecewise smooth function space, and the relation between the discrete model (3.2) and the variational model. The rigorous analysis of the convergence and approximation of the variational model by the discrete model (3.2) will be given in Section 5.

4.1. Space of Piecewise Smooth Functions and Properties. In this section, we give the precise definition of the space of piecewise smooth functions and introduce some of the properties of such functions that will be needed in later sections.

Piecewise smooth functions are functions that have different regularities in different sub-domains of the domain $\Omega = (0, 1)^2 \subset \mathbb{R}^2$. In this paper, we shall focus on piecewise smooth functions satisfy the following two conditions:

- (1) First order weak derivatives are integrable in sub-domains $\{\Omega_j\}$ of Ω with jump discontinuities at the boundaries of these sub-domains, which is the *jump set*.
- (2) Second order or higher weak derivatives are integrable in the sub-domains $\{\Omega_{j,\bar{j}}\} \subset \Omega_j$ and the first order derivatives have jump discontinuities at the boundaries of $\Omega_{j,\bar{j}}$, which shall be called the *(first order) hidden jump set*.



FIGURE 7. Zoom-in views of the images in Figure 6.

Note that we can define similarly the piecewise smooth functions with higher order hidden jumps sets, i.e. jump discontinuities of their second or higher order derivatives. However, we shall focus on the aforementioned space since jumps and first order hidden jumps are most important image features.

Now, we give the precise definitions of these sets and the space of piecewise smooth functions. Let $H^s(\Omega)$, with $s = 1, 2, \dots$, be the Sobolev space equipped with the norm $\|f\|_{H^s(\Omega)} = \sum_{0 \leq |i| \leq s} \|D_i f\|_{L_2(\Omega)}$, where $D_i(f(x, y)) = \frac{\partial^{|i|} f}{\partial x^{i_1} \partial y^{i_2}}$. Let $\{\Omega_j : j = 1, 2, \dots, m\}$ be a collection of open subsets of Ω satisfying

$$(4.1) \quad \bigcup_j \overline{\Omega_j} = \overline{\Omega} \quad \text{and} \quad \mathfrak{L}(\overline{\Omega_{j_1}} \cap \overline{\Omega_{j_2}}) = 0, \text{ for } j_1 \neq j_2,$$

with $\mathfrak{L}(\cdot)$ being the Lebesgue measure. We assume that Ω_j are Lipschitz domains [44] and $\partial\Omega_j$ are piecewise C^1 . Let the set of curves $\{\Gamma_j : j = 1, 2, \dots, \tilde{m}\}$ (see Figure 10), with smallest possible \tilde{m} , be such that

$$(4.2) \quad \overline{\bigcup_j \Gamma_j} = \bigcup_j \partial\Omega_j \setminus \partial\Omega,$$

and for each j , Γ_j has exactly one domain of $\{\Omega_j\}$ at each side (see Figure 10 for an illustration of Γ_j and Ω_j). We denote the two domains on each side of Γ_j as Ω_j^+ and Ω_j^- . Obviously, $\Omega_j^\pm \in$



FIGURE 8. First row shows the original and the observed noisy and blurry images with noise's standard deviation = 2, 4, 6 respectively. The PSNRs of the observed images are **23.9284**, **23.7344** and **23.4290** respectively. Second row shows the zoom-in views.

$\{\Omega_i : i = 1, 2, \dots, m\}$ for each j . Similarly, let the set of Lipschitz domains $\{\Omega_{j,\tilde{j}} : \tilde{j} = 1, 2, \dots, m_j\}$ be a partition of Ω_j for each $j = 1, \dots, \tilde{m}$ and satisfies a similar condition as (4.1). Let $\{\Gamma_{j,\tilde{j}} : \tilde{j} = 1, 2, \dots, \tilde{m}_j\}$ be the piecewise C^1 boundary curves associated to $\{\Omega_{j,\tilde{j}}\}$ that satisfy a similar condition as (4.2). We denote the two domains on each side of $\Gamma_{j,\tilde{j}}$ as $\Omega_{j,\tilde{j}}^+$ and $\Omega_{j,\tilde{j}}^-$.

Now, we define the space of piecewise smooth functions (with jumps and first order hidden jumps) as

$$\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) := \{f \in L_2(\Omega) : \|f\|_{\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})} < \infty\},$$

where

$$(4.3) \quad \|f\|_{\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})} := \sum_{j=1}^m \left[\|f\|_{H^1(\Omega_j)} + \sum_{\tilde{j}=1}^{m_j} \|f\|_{H^{s_{j,\tilde{j}}}(\Omega_{j,\tilde{j}})} \right],$$

where $s = \min\{s_{j,\tilde{j}}\}$ and $s_{j,\tilde{j}} \geq 2$. Now, we have the following property of the space of piecewise smooth functions.

Proposition 4.1. *The space of piecewise smooth functions $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) \subset L_2(\Omega)$ is a Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle_\Omega = \sum_{j=1}^m \langle \cdot, \cdot \rangle_{\Omega_j}$, with $\langle u, v \rangle_{\Omega_j} = \int_{\Omega_j} uv$.*

Note that we shall drop the subscript in the inner products and simple use $\langle \cdot, \cdot \rangle$ whenever the domain of integration is clear in context.

Recall that, the trace operator $\mathfrak{T} : H^1(B) \mapsto H^{\frac{1}{2}}(\partial B)$, for a general Lipschitz domain $B \subset \mathbb{R}^2$, is a continuous linear operator (see e.g. [45, 46]). The trace operator is defined on the dense subspace $C^\infty(\overline{B}) \subset H^s(B)$ as: $\mathfrak{T}(u) = u|_{\partial B}$ for $u \in C^\infty(\overline{B})$. With the trace operator, we can define the generalized integration by parts formula for functions in the Sobolev space $H^s(B)$. Given $\langle u, D_{\mathbf{i}}v \rangle$ for $|\mathbf{i}| \geq 2$, there exists multiple paths for integration by parts that can lead us to $(-1)^{|\mathbf{i}|} \langle D_{\mathbf{i}}u, v \rangle$ plus boundary terms which depend on the specific path we take. Therefore, to be more precise in our analysis, we introduce a notation, i.e. the index set $\mathbb{D}_{\mathbf{i}}$, for the path we take when doing

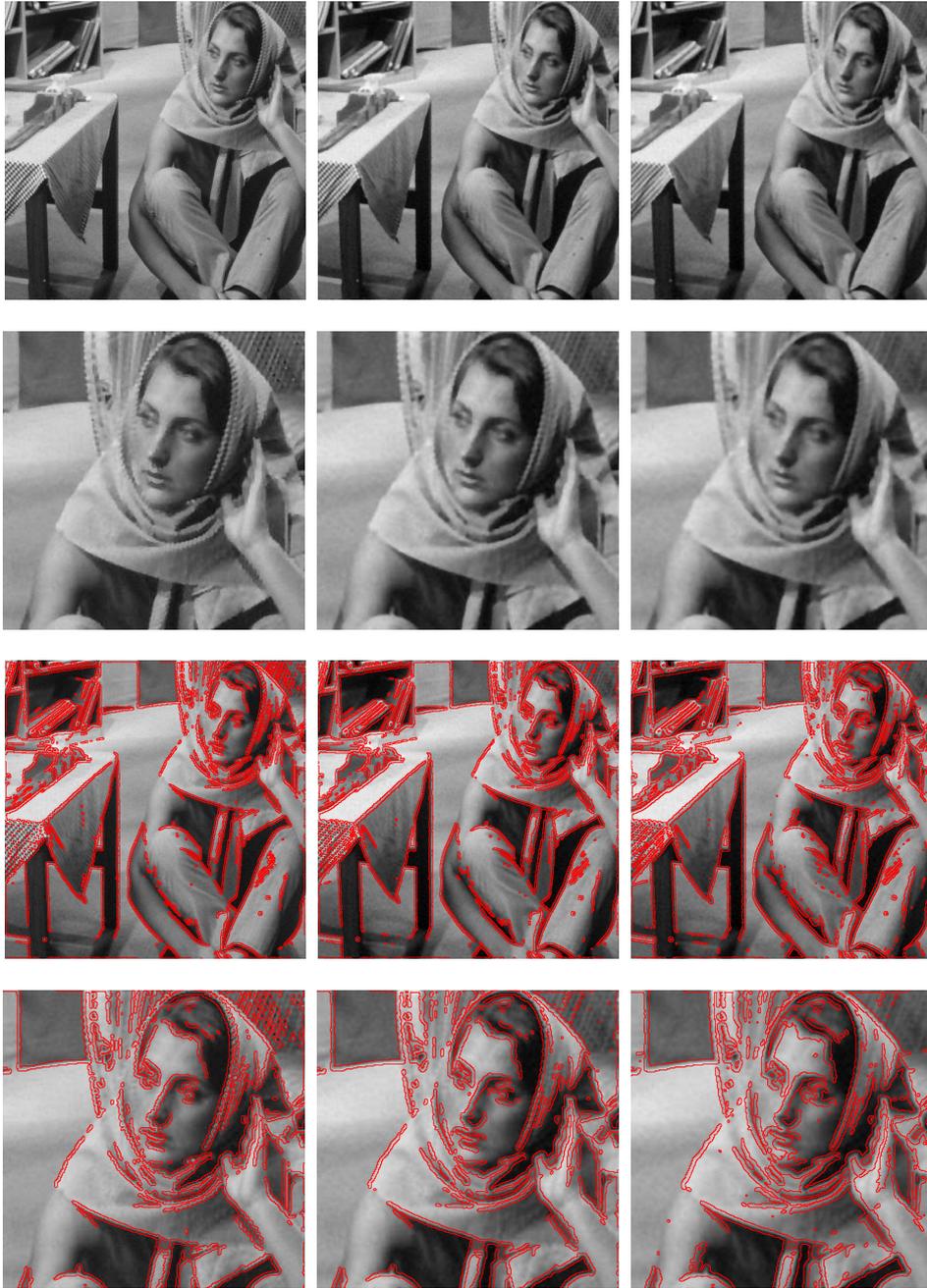
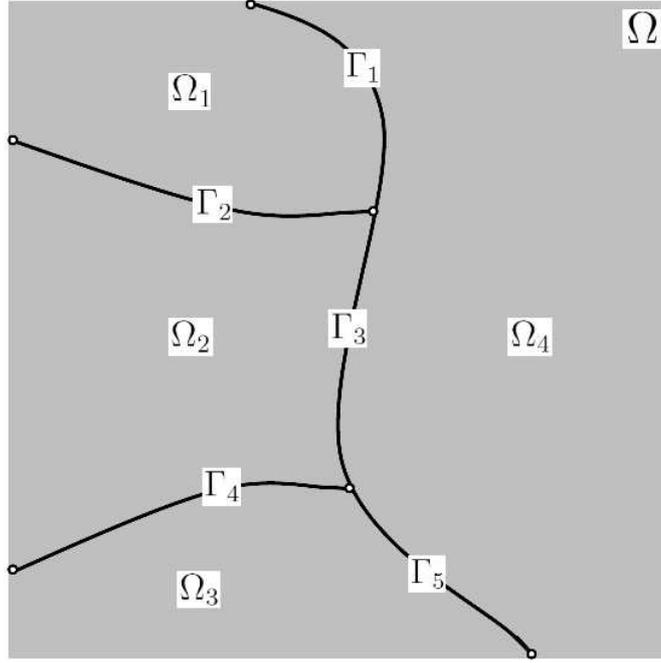


FIGURE 9. The three columns of images from left to right are the results of restored \mathbf{u} and estimated $\mathbf{\Gamma}$ from the observed images with noise level 2, 4, and 6 respectively. Row one and two: restored images and their zoom-in views. The PSNRs of the restored images are **24.5950**, **24.3374** and **24.0700** respectively. Row three and four: estimated jump sets and the their zoom-in views.

integration by parts of $\langle u, D_i v \rangle$. We define the set \mathbb{D}_i as

$$(4.4) \quad \mathbb{D}_i := \{j_l < i : |j_l| = l - 1; j_l < j_{l+1}; l = 1, 2, \dots, |i|\}.$$

FIGURE 10. Illustrations of domains Ω_j and curves Γ_j .

Note that, for a given \mathbf{i} , the set $\mathbb{D}_{\mathbf{i}}$ may not be uniquely defined. However, the integration by parts formula given in Proposition 4.2 holds for any instance of $\mathbb{D}_{\mathbf{i}}$. The vector $\mathbf{n}_j = n_1$ if $\mathbf{j} = (1, 0)^\top$ and $\mathbf{n}_j = n_2$ if $\mathbf{j} = (0, 1)^\top$ with $\mathbf{n} = (n_1, n_2)^\top$ being the outward normal of ∂B .

Example 4.1. The set $\mathbb{D}_{\mathbf{i}}$ indicates the type of differential operators that appears on u at the boundary after the operation of integration by parts on $\langle u, D_{\mathbf{i}}\varphi \rangle$. Here we provide a few examples.

- (1) For $\mathbf{i} = (1, 0)^\top$ or $(0, 1)^\top$: $\mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top\}$.
- (2) For $\mathbf{i} = (1, 1)^\top$:

$$\mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top, (1, 0)^\top\} \quad \text{or} \quad \mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top, (0, 1)^\top\}.$$

- (3) For $\mathbf{i} = (2, 1)^\top$:

$$\mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top, (1, 0)^\top, (1, 1)^\top\}, \quad \mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top, (0, 1)^\top, (1, 1)^\top\}, \quad \text{or} \quad \mathbb{D}_{\mathbf{i}} = \{(0, 0)^\top, (1, 0)^\top, (2, 0)^\top\}.$$

Proposition 4.2. Let $u \in H^s(B)$ and $\varphi \in C^s(\overline{B})$ with $B \subset \Omega$ a Lipschitz domain with piecewise C^1 boundary ∂B . Then, for any $1 \leq |\mathbf{i}| \leq s$, we have the following formula of integration by parts

$$\langle u, D_{\mathbf{i}}\varphi \rangle = \sum_{\mathbf{j}_l \in \mathbb{D}_{\mathbf{i}}, 1 \leq l \leq |\mathbf{i}|} (-1)^{l-1} \int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u) D_{\mathbf{i}-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds + (-1)^{|\mathbf{i}|} \langle D_{\mathbf{i}} u, \varphi \rangle,$$

where $\mathfrak{T}(\cdot)$ is the trace operator defined on $H^s(B)$.

Proof. Given any $u \in H^s(B)$ and the sequence $u_m \in C^\infty(\overline{B})$ such that $u_m \rightarrow u$ in $H^s(B)$, we have

$$\langle u_m, D_{\mathbf{i}}\varphi \rangle = \sum_{\mathbf{j}_l \in \mathbb{D}_{\mathbf{i}}, 1 \leq l \leq |\mathbf{i}|} (-1)^{l-1} \int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u_m) D_{\mathbf{i}-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds + (-1)^{|\mathbf{i}|} \langle D_{\mathbf{i}} u_m, \varphi \rangle,$$

where $\mathfrak{T}(D_{\mathbf{j}_l} u_m) = (D_{\mathbf{j}_l} u_m)|_{\partial B}$. First of all, by Cauchy-Schwartz inequality, we have

$$|\langle u_m, D_{\mathbf{i}}\varphi \rangle - \langle u, D_{\mathbf{i}}\varphi \rangle| \leq \|u_m - u\|_{H^s(B)} \|D_{\mathbf{i}}\varphi\|_{L_2(B)},$$

which implies that $\langle u_m, D_i \varphi \rangle \rightarrow \langle u, D_i \varphi \rangle$ as $m \rightarrow \infty$. Similarly, we can prove that $\langle D_i u_m, \varphi \rangle \rightarrow \langle D_i u, \varphi \rangle$. It remains to show that for each l and $\mathbf{j}_l \in \mathbb{D}_i$, we have

$$\int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u_m) D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds \rightarrow \int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u) D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds.$$

Indeed, since $\mathfrak{T} : H^1(B) \mapsto H^{\frac{1}{2}}(\partial B)$ is linear and bounded, and ∂B is of finite length, we have

$$\begin{aligned} & \left| \int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u_m) D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds - \int_{\partial B} \mathfrak{T}(D_{\mathbf{j}_l} u) D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} ds \right| \\ & \leq \| \mathfrak{T}(D_{\mathbf{j}_l} u_m - D_{\mathbf{j}_l} u) \|_{L_2(\partial B)} \| D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} \|_{L_2(\partial B)} \\ & \leq \| \mathfrak{T}(D_{\mathbf{j}_l} u_m - D_{\mathbf{j}_l} u) \|_{H^{\frac{1}{2}}(\partial B)} \| D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} \|_{L_2(\partial B)} \\ & \leq C \| D_{\mathbf{j}_l} (u_m - u) \|_{H^1(B)} \| D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} \|_{L_2(\partial B)} \\ & \leq C \| u_m - u \|_{H^s(B)} \| D_{i-\mathbf{j}_{l+1}} \varphi \mathbf{n}_{\mathbf{j}_{l+1}-\mathbf{j}_l} \|_{L_2(\partial B)}. \end{aligned}$$

This concludes the proof of the proposition. \square

4.2. Variational Model for Image Restoration. In this subsection, we present a new variational model, which can be approximated by (3.2) (for any fixed jump set) under suitable assumptions. From the variational model, one can clearly see that the discrete model (3.2) does approximate piecewise smooth solutions of the underlying linear inverse problem. This subsection focuses on the introduction of the variational model, while a detailed convergence analysis between (3.2) and the variational model will be given in Section 5.

The continuum counterpart of the linear inverse problem (3.1) can be written as

$$(4.5) \quad f = Au + \eta,$$

where A is some linear bounded operator mapping $L_2(\Omega)$ into itself. Here, and throughout the rest of this paper, we assume that the solution u that we seek for is a piecewise smooth function in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) \subset L_2(\Omega)$. To obtain a piecewise smooth function as an approximated solution of the linear inverse problem (4.5), we consider the following variational model

$$(4.6) \quad \begin{aligned} & \inf_{u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}), \{\Gamma_j\}, \{\Gamma_{j,\tilde{j}}\}} \| \boldsymbol{\nu} \cdot \mathbf{D}u \|_2^2 + \sum_{j=1}^{\tilde{m}} \left[\mu_1 \int_{\Gamma_j} | \mathfrak{T}_j^+(u) - \mathfrak{T}_j^-(u) | ds \right. \\ & \quad \left. + \mu_2 \sum_{\tilde{j}=1}^{\tilde{m}_j} \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} | \mathfrak{T}_{j,\tilde{j}}^+(D_{\mathbf{i}}u) - \mathfrak{T}_{j,\tilde{j}}^-(D_{\mathbf{i}}u) |^2 \right)^{\frac{1}{2}} ds \right] + \frac{1}{2} \| Au - f \|_{L_2(\Omega)}^2, \end{aligned}$$

where

$$\| \boldsymbol{\nu} \cdot \mathbf{D}u \|_2^2 := \sum_{j=1}^m \left[\nu_1 \sum_{|\mathbf{j}|=1} \| D_{\mathbf{j}} u \|_{L_2(\Omega_j)}^2 + \nu_2 \sum_{\tilde{j}=1}^{m_j} \left(\sum_{1 \leq |\mathbf{i}_{j,\tilde{j}}| \leq s_{j,\tilde{j}}} \| D_{\mathbf{i}_{j,\tilde{j}}} u \|_{L_2(\Omega_{j,\tilde{j}})}^2 \right) \right],$$

and $\mathfrak{T}_{j,\tilde{j}}^{\pm}$ is the trace operator defined for $H^{s_{j,\tilde{j}}}(\Omega_{j,\tilde{j}}^{\pm})$. Note that by definition of the curves $\{\Gamma_j\}$ and $\{\Gamma_{j,\tilde{j}}\}$, they are (part of the) boundaries of $\{\Omega_j\}$ and $\{\Omega_{j,\tilde{j}}\}$ respectively. Therefore, the sets $\{\Omega_j\}$ and $\{\Gamma_j\}$ (resp. $\{\Omega_{j,\tilde{j}}\}$ and $\{\Gamma_{j,\tilde{j}}\}$) are in direct correspondence to each other.

Remark 4.1. The variational model (4.6) is related, but more general than the Mumford-Shah model [1] in a certain sense. Recall the Mumford-Shah functional

$$E(u, \Gamma) = \nu \int_{\Omega \setminus \Gamma} |\nabla u|^2 + \mu |\Gamma| + \frac{1}{2} \| u - f \|_{L_2(\Omega)}^2,$$

where $|\Gamma|$ denotes the length of Γ . Then, we can see that:

- (1) The variational model (4.6) has a more general regularization term (i.e. $\|\mathbf{D}u\|_2^2$) than that used by Mumford-Shah model.
- (2) The term $\int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|i|=1} \left| \mathfrak{F}_{j,\tilde{j}}^+(D_i u) - \mathfrak{F}_{j,\tilde{j}}^-(D_i u) \right|^2 \right)^{\frac{1}{2}} ds$ that measures the size of the first order hidden jumps of u at Γ is not considered in the Mumford-Shah model. Hidden jumps are also important image features and a proper recovery of them will help with the quality of image restoration.
- (3) We have a general image restoration problem embedded in the term $\frac{1}{2}\|Au - f\|_{L_2(\Omega)}^2$, while Mumford-Shah model only has $\frac{1}{2}\|u - f\|_{L_2(\Omega)}^2$ (denoising).
- (4) When we take $\mathbf{D} = \nabla$, $\nu = \nu$, $\mu_1 = \mu$, $\mu_2 = 0$ and $A = I$, the variational model (4.6) reduces to

$$(4.7) \quad \inf_{u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}), \{\Gamma_j\}} \nu \|\nabla u\|_2^2 + \mu \sum_{j=1}^{\tilde{m}} \|v_j(u)\|_{L_1(\Gamma_j)} + \frac{1}{2}\|u - f\|_{L_2(\Omega)}^2,$$

where $v_j(u) = \mathfrak{F}_j^+(u) - \mathfrak{F}_j^-(u)$ which represents the jump function of u on Γ_j . Let $\|v\|_{L_0(\Gamma)}$ denote the measure of the length of support of v on Γ . Then we can see that, if we replace the term $\|v_j(u)\|_{L_1(\Gamma_j)}$ by $\|v_j(u)\|_{L_0(\Gamma_j)}$, the energy function of (4.7) becomes the Mumford-Shah functional. In other words, our model penalizes the jump values of the image on the jump set instead of the length of the jump set, which is more suitable for image restoration. In addition, the L_1 -minimization is easier to compute than the L_0 -minimization. When $\|\cdot\|_{L_0}$ is chosen instead of $\|\cdot\|_{L_1}$ in (4.7), the computation result will be highly sensitive to the estimation of the support set of $v_j(u)$, and normal numerical algorithms such as Algorithm 3.3 may easily stuck at an unfavorable local minimum. This explains why our proposed model has significant computational advantage over the Mumford-Shah model.

- (5) When finite difference method is used to discretize the Mumford-Shah model, the relation between the discrete and the continuum functionals was established in [24, 47] through Gamma-convergence. Their work seems related to ours. However, the analysis here is entirely different from that of [24, 47]. Furthermore, the model given here is more general and complex than Mumford-Shah model; the discrete model is much more sophisticated and has shown in [22, 23] to be superior than some standard finite difference discretization.

In Section 5, relations between the discrete model (3.2) and the variational model (4.6) will be established through an asymptotic analysis as the image resolution goes to infinity. This leads to the conclusion that images are indeed modeled as piecewise smooth functions when model (3.2) is used. Furthermore, this also implies that the discrete model (3.2) can be used to obtain approximate solution of the variational model (4.6).

5. ASYMPTOTIC ANALYSIS

This section is devoted to establish the assertion that the image restoration model (3.2) assumes that the underlying solutions are piecewise smooth functions. This is established by an asymptotic analysis as the image resolution goes to infinity. In particular, we will show that, under suitable assumptions and as the image resolution going to infinity, the energy functional of (3.2) converges to the variational model (4.6) for some fixed jump set $\{\Gamma_j\}$ and $\{\Gamma_{j,\tilde{j}}\}$ if the parameters of (3.2) are properly chosen. Consequently, through such convergence, we can draw connections between the discrete (approximated) solutions of (3.2) with those of the variational model for some fixed jump set.

The reason that we only consider the case with fixed jump set, is because it makes the analysis possible, which is already very technical. This is also justified by the facts that the jump set can be robustly estimated by wavelet frame transform numerically due to multiscale structure of wavelet frames and their multiple orders of vanishing moments. Recall that the major purpose of

the introduction of (4.6) is to support our claim that the discrete model (3.2) is to find piecewise smooth solutions of the linear inverse problem in discrete setting, and thus, analyzing (4.6) with the jump set fixed is sufficient for such purpose. Therefore, this section is focused on studying the relation between the following discrete and continuum models, for fixed Γ :

$$(5.1) \quad \inf_{\mathbf{u} \in \mathcal{L}_2} \|\boldsymbol{\lambda} \cdot \mathbf{W}\mathbf{u}\|_{\Gamma^c}^2 + \|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1 + \frac{1}{2}\|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2, \quad (\text{discrete})$$

(5.2)

$$\begin{aligned} \inf_{u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})} \|\boldsymbol{\nu} \cdot \mathbf{D}u\|_2^2 + \sum_{j=1}^{\tilde{m}} \left[\mu_1 \int_{\Gamma_j} |\mathfrak{T}_j^+(u) - \mathfrak{T}_j^-(u)| ds \right. \\ \left. + \mu_2 \sum_{\tilde{j}=1}^{\tilde{m}_j} \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} |\mathfrak{T}_{j,\tilde{j}}^+(D_{\mathbf{i}}u) - \mathfrak{T}_{j,\tilde{j}}^-(D_{\mathbf{i}}u)|^2 \right)^{\frac{1}{2}} ds \right] + \frac{1}{2}\|Au - f\|_{L_2(\Omega)}^2, \end{aligned} \quad (\text{continuum})$$

5.1. Wavelet Frame Based Model: Revisited. For convenience, we introduce some symbols and notation that will be used throughout the rest of the paper. In particular, it will be helpful in precisely stating the assumptions we make to the wavelet frame based model (5.1) so that it can be proved to converge to the variational model (5.2).

Notation 5.1. We focus our analysis on \mathbb{R}^2 , i.e., the 2-dimensional cases. All the 2-dimensional refinable functions and framelets are assumed to be constructed by tensor products of univariate B-splines and the associated framelets obtained from the UEP [25].

- (1) We assume all functions we consider are defined on the open unit square $\Omega := (0, 1)^2 \subset \mathbb{R}^2$, and that their discrete versions, i.e. digital images, are defined on an $N \times N$ Cartesian grid on $\bar{\Omega}$ with $N = 2^n + 1$ for $n \geq 0$. We denote by $h = 2^{-n}$ the meshsize of the $N \times N$ grid.
- (2) We use bold-face letters \mathbf{i} , \mathbf{j} , and \mathbf{k} to denote double indices in \mathbb{Z}^2 . We denote by $\mathbb{O}^2 := \{0, 1, \dots, N-1\}^2$ as the set of indices of the $N \times N$ Cartesian grid.
- (3) For 2-dimensional cases, $\phi_{n,\mathbf{k}}$ (also $\varphi_{n,\mathbf{k}}$ and $\psi_{n,\mathbf{k}}$, etc.) takes the form

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

Since we focus on the quasi-affine system, we have

$$\phi_{n-1,\mathbf{k}} = 2^{n-2} \phi(2^{n-1} \cdot -\mathbf{k}/2).$$

- (4) For simplicity, we assume that the level of wavelet frame decomposition is 1, i.e. $L = 1$. In that case, we have

$$\mathbf{W}\mathbf{u} = \{\mathbf{W}_{\mathbf{i}}\mathbf{u} : 0 \leq i_1, i_2 \leq r\}, \quad \mathbf{W}_{\mathbf{i}}\mathbf{u} := \mathbf{a}_{\mathbf{i}}[-\cdot] \otimes \mathbf{u}, \quad \text{with } \mathbf{u} \in \mathbb{R}^{M^2}.$$

Recall that the wavelet frame bands (high frequency bands) is denoted as $\mathbb{B} = \{\mathbf{i} : 0 \leq i_1, i_2 \leq r\} \setminus \{\mathbf{0}\}$. Let

$$\mathbb{B}_1 := \{(1, 0), (0, 1)\}, \quad \mathbb{B}_2 := \{(1, 1), (2, 0)(0, 2)\}, \quad \text{etc},$$

where \mathbb{B}_l denotes the wavelet frame bands corresponding to the framelets of vanishing moment l .

- (5) We shall divide the index set \mathbb{O}^2 into different subsets, where each of them plays a specific role of characterizing the locations of jumps and hidden jumps, or smooth regions. The supports of ϕ and the framelets $\psi_{\mathbf{i}}$ are important in properly defining these index sets. We denote $\Lambda_{n,\mathbf{k}}$ the intersection of the supports of $\phi_{n,\mathbf{k}}$ and $\psi_{\mathbf{i},n,\mathbf{k}}$ for all $\mathbf{i} \in \mathbb{B}$. Note that for the tensor-product B-spline wavelet frame systems constructed in [25], we have $\text{supp}\phi_{n,\mathbf{k}} = \text{supp}\psi_{\mathbf{i},n,\mathbf{k}}$ for each $\mathbf{i} \in \mathbb{B}$. Therefore, the support $\Lambda_{n,\mathbf{k}}$ is the common support of both the refinable function and the framelets. Although this property will be used in the proofs, it is not essential and can be easily removed.

- (6) Given a wavelet frame system and its corresponding refinable function ϕ , we denote \mathbb{M}^2 the set of indices $\mathbf{k} \in \mathbb{O}^2$ such that $\Lambda_{n,\mathbf{k}}$ is completely supported in $\overline{\Omega}$. Then obviously, $\mathbb{M}^2 \subset \mathbb{O}^2$. We denote the set of sequences defined on the grids \mathbb{M}^2 as \mathbb{R}^{M^2} with M^2 the cardinality of \mathbb{M}^2 .
- (7) We denote $\mathbb{K}^2 \subset \mathbb{M}^2$ as the index set when the boundary condition of $\mathbf{a}_i[-\cdot] \otimes \mathbf{u}$ is inactive for all \mathbf{i} , or in other words, $\mathbf{a}_i[-\cdot] * \mathbf{u}$ is well defined for all \mathbf{i} , where “ $*$ ” is the standard discrete convolution operator. Let K^2 be the cardinality of \mathbb{K}^2 . Then $\mathbf{W}_i : \mathbb{R}^{M^2} \mapsto \mathbb{R}^{K^2}$ for each $(0,0) \leq \mathbf{i} \leq (r,r)$. Note that the index sets \mathbb{O}^2 , \mathbb{M}^2 and \mathbb{K}^2 all depend on the image resolution n .
- (8) Given a partition $\{\Omega_j\}$ of Ω satisfying (4.1) and its associated boundary curves $\{\Gamma_j\}$, we define $\mathbb{M}_j^2 \subset \mathbb{M}^2$ to be the index set such that $\Lambda_{n,\mathbf{k}} \subset \overline{\Omega}_j$. Define $\mathbb{G}_j^2 \subset \mathbb{K}^2$ to be the index set such that the interior of $\Lambda_{n-1,\mathbf{k}}$ has a nonempty intersect with Γ_j and has an empty intersection with all $\Gamma_{j'}$ for $j' \neq j$. The latter requirement of \mathbb{G}_j^2 is to exclude index \mathbf{k} such that $\Lambda_{n-1,\mathbf{k}}$ include a multi-junction in its interior. For simplicity, we use $\Gamma_j \cap \Lambda_{n-1,\mathbf{k}} \neq \emptyset$ to denote that the interior of $\Lambda_{n-1,\mathbf{k}}$ has a nonempty intersect with Γ_j . Similar to the definition of \mathbb{K}^2 , we can define the index sets $\mathbb{K}_j^2 \subset \mathbb{M}_j^2$ be such that the boundary condition of $\mathbf{a}_i[-\cdot] \otimes \mathbf{u}$, for $\mathbf{u} \in \mathbb{R}^{M_j^2}$, is inactive for all \mathbf{i} .
- (9) Similarly, given a partition $\{\Omega_{j,\tilde{j}}\}$ of Ω_j satisfying (4.1) and its associated boundary curves $\{\Gamma_{j,\tilde{j}}\}$ for each $j = 1, \dots, \tilde{m}$, we define $\mathbb{M}_{j,\tilde{j}}^2 \subset \mathbb{M}_j^2$ to be the index set such that $\Lambda_{n-1,\mathbf{k}} \subset \overline{\Omega}_{j,\tilde{j}}$. Define $\mathbb{G}_{j,\tilde{j}}^2 \subset \mathbb{K}_j^2$ to be the index set such that the interior of $\Lambda_{n-1,\mathbf{k}}$ has a nonempty intersect with $\Gamma_{j,\tilde{j}}$ and has an empty intersection with all $\Gamma_{j,\tilde{j}'}$ for $\tilde{j}' \neq \tilde{j}$. We define the index sets $\mathbb{K}_{j,\tilde{j}}^2 \subset \mathbb{M}_{j,\tilde{j}}^2$ be such that the boundary condition of $\mathbf{a}_i[-\cdot] \otimes \mathbf{u}$, for $\mathbf{u} \in \mathbb{R}^{M_{j,\tilde{j}}^2}$, is inactive for all \mathbf{i} .
- (10) In order to link the continuous and discrete settings, we need to take resolution into account. Therefore, for any array $\mathbf{v} \in \mathbb{R}^{M^2}$, the discrete ℓ_p -norm we are using now is defined as

$$(5.3) \quad \|\mathbf{v}\|_p^p := \sum_{\mathbf{i} \in \mathbb{M}^2} |\mathbf{v}[\mathbf{i}]|^p h^2.$$

Given a partition $\{\Omega_{j,\tilde{j}}\}$ of the domain Ω and the associated boundary curves $\{\Gamma_{j,\tilde{j}}\}$, if we assume that both $\{\Omega_{j,\tilde{j}}\}$ and $\{\Gamma_{j,\tilde{j}}\}$ are known and fixed, and if we only consider the second order band \mathbb{B}_2 for the term $\|[\boldsymbol{\gamma} \cdot \mathbf{W}\mathbf{u}]_{\Gamma}\|_1$, then (5.1) can be explicitly written as:

$$(5.4) \quad \inf_{\mathbf{u} \in \mathbb{R}^{M^2}} \sum_{j=1}^m \left[\left\| [\boldsymbol{\lambda}_j \cdot \mathbf{W}\mathbf{u}]_{\Omega_j} \right\|_2^2 + \sum_{\tilde{j}=1}^{m'} \left\| [\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}} \cdot \mathbf{W}\mathbf{u}]_{\Omega_{j,\tilde{j}}} \right\|_2^2 \right] \\ + \sum_{j=1}^{\tilde{m}} \left[\left\| [\boldsymbol{\gamma}_j \cdot \mathbf{W}\mathbf{u}]_{\Gamma_j} \right\|_1 + \sum_{\tilde{j}=1}^{\tilde{m}_j} \left\| [\tilde{\boldsymbol{\gamma}}_{j,\tilde{j}} \cdot \mathbf{W}\mathbf{u}]_{\Gamma_{j,\tilde{j}}} \right\|_1 \right] + \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_2^2,$$

where

$$\left\| [\boldsymbol{\lambda}_j \cdot \mathbf{W}\mathbf{u}]_{\Omega_j} \right\|_2^2 := h^2 \sum_{\mathbf{k} \in \mathbb{K}_j^2} \sum_{\mathbf{i} \in \mathbb{B}} \lambda_{i,j}[\mathbf{k}] \left| (\mathbf{W}_{n,i}\mathbf{u})[\mathbf{k}] \right|^2, \\ \left\| [\boldsymbol{\gamma}_j \cdot \mathbf{W}\mathbf{u}]_{\Gamma_j} \right\|_1 := h \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left(\sum_{\mathbf{i} \in \mathbb{B}_1} \gamma_{i,j}[\mathbf{k}] \left| (\mathbf{W}_{n,i}\mathbf{u})[\mathbf{k}] \right|^2 \right)^{\frac{1}{2}},$$

$$\left\| \left[\tilde{\lambda}_{j,\tilde{j}} \cdot \mathbf{W}\mathbf{u} \right]_{\Omega_{j,\tilde{j}}} \right\|_2^2 := h^2 \sum_{\mathbf{k} \in \mathbb{K}_{j,\tilde{j}}^2} \sum_{i \in \mathbb{B}} \tilde{\lambda}_{i,j,\tilde{j}}[\mathbf{k}] \left| (\mathbf{W}_{n,i}\mathbf{u})[\mathbf{k}] \right|^2$$

and

$$\left\| \left[\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}\mathbf{u} \right]_{\Gamma_{j,\tilde{j}}} \right\|_1 := h \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2} \left(\sum_{i \in \mathbb{B}_2} \tilde{\gamma}_{i,j,\tilde{j}}[\mathbf{k}] \left| (\mathbf{W}_{n,i}\mathbf{u})[\mathbf{k}] \right|^2 \right)^{\frac{1}{2}}.$$

Here, we recall that

$$\mathbb{B}_1 = \{(1, 0), (0, 1)\} \quad \text{and} \quad \mathbb{B}_2 = \{(1, 1), (2, 0), (0, 2)\}.$$

In the next subsection, we will show that the wavelet frame based model (5.4) is related to the variational model (5.2) in the sense that the energy function of (5.4) convergence to the energy functional of (5.2). The specific type of convergence we establish implies Gamma-convergence. More importantly, based on such convergence, the approximated solutions of (5.4) can be shown to approximate those of (5.2). Therefore, we can regard (5.4) as a certain discretization of the variational model (5.2). We finally note that the use of wavelet frame band \mathbb{B}_1 in the term $\left\| [\gamma_j \cdot \mathbf{W}\mathbf{u}]_{\Gamma_j} \right\|_1$ is to approximate the jumps of function values across Γ_j , and the use of wavelet frame band \mathbb{B}_2 in the term $\left\| [\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}\mathbf{u}]_{\Gamma_{j,\tilde{j}}} \right\|_1$ is to approximate the jumps of the values of the first order derivatives across $\Gamma_{j,\tilde{j}}$. The model (5.1) is in fact more general than (5.4) since it contains bands other than \mathbb{B}_1 and \mathbb{B}_2 . We can apply a similar analysis to show that (5.4) converges to a variational model similar to (5.2) with higher order hidden jumps. However, we shall focus on jumps and first order hidden jumps for our theoretical analysis, i.e. we focus on the relation between (5.4) and (5.2).

5.2. Connection between Model (5.4) and Model (5.2). To find a connection between (5.4) and (5.2), we need to convert (5.4) to another equivalent optimization problem on the function space $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$. Define the operator \mathbf{T}_n on the space $L_2(\Omega)$ as

$$(5.5) \quad \mathbf{T}_n u = \{2^n \langle u, \phi_{n,\mathbf{k}} \rangle : \mathbf{k} \in \mathbb{M}^2\} \in \mathbb{R}^{M^2}, \quad u \in L_2(\Omega).$$

Then the optimization problem (5.4) is equivalent to the following one

$$(5.6) \quad \inf_{u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})} E_n(u),$$

where

$$\begin{aligned} E_n(u) = & \sum_{\tilde{j}=1}^m \left[\left\| [\lambda_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_j} \right\|_2^2 + \sum_{\tilde{j}=1}^{m'} \left\| [\tilde{\lambda}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_{j,\tilde{j}}} \right\|_2^2 \right] \\ & + \sum_{j=1}^{\tilde{m}} \left[\left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 + \sum_{\tilde{j}=1}^{\tilde{m}_j} \left\| [\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_{j,\tilde{j}}} \right\|_1 \right] + \frac{1}{2} \|\mathbf{A}_n \mathbf{T}_n u - \mathbf{T}_n f\|_2^2. \end{aligned}$$

where the subscript “ n ” of \mathbf{W}_n and \mathbf{A}_n is to emphasize the dependence of the operators on image resolution level n . The equivalence between (5.4) and (5.6) can be shown similarly as [22, Proposition 3.1]. The equivalence is in the sense that (5.4) and (5.6) have the same infimum value, and the minimizers, should they exist, can be constructed from each other (see [22, Proposition 3.1] for details).

Therefore, we can focus on analyzing the relation between the energy functional of (5.2) and that of (5.6). We denote the energy functional of (5.2) as

$$\begin{aligned} E(u) &:= \|\boldsymbol{\nu} \cdot \mathbf{D}u\|_2^2 + \mu_1 \sum_{j=1}^{\tilde{m}} \int_{\Gamma_j} \left| \mathfrak{T}_j^+(u) - \mathfrak{T}_j^-(u) \right| ds \\ &\quad + \mu_2 \sum_{j=1}^{\tilde{m}} \sum_{\tilde{j}=1}^{\tilde{m}_j} \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} \left| \mathfrak{T}_{j,\tilde{j}}^+(D_{\mathbf{i}}u) - \mathfrak{T}_{j,\tilde{j}}^-(D_{\mathbf{i}}u) \right|^2 \right)^{\frac{1}{2}} ds + \frac{1}{2} \|Au - f\|_{L_2(\Omega)}^2 \\ &=: E^{(1)}(u) + E^{(2)}(u) + E^{(3)}(u) + E^{(4)}(u) \end{aligned}$$

and

$$\begin{aligned} E_n(u) &= \sum_{j=1}^m \left[\left\| [\boldsymbol{\lambda}_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_j} \right\|_2^2 + \sum_{\tilde{j}=1}^{m'} \left\| [\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_{j,\tilde{j}}} \right\|_2^2 \right] \\ &\quad + \sum_{j=1}^{\tilde{m}} \left\| [\boldsymbol{\gamma}_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 + \sum_{j=1}^{\tilde{m}} \sum_{\tilde{j}=1}^{\tilde{m}_j} \left\| [\tilde{\boldsymbol{\gamma}}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_{j,\tilde{j}}} \right\|_1 + \frac{1}{2} \|\mathbf{A}_n \mathbf{T}_n u - \mathbf{T}_n f\|_2^2. \\ &=: E_n^{(1)}(u) + E_n^{(2)}(u) + E_n^{(3)}(u) + E_n^{(4)}(u). \end{aligned}$$

Without loss of generality, we take $\mu_1 = \mu_2 = 1$ and $\boldsymbol{\nu} = (1, 1)$ for $E(u)$. To draw an asymptotic relation between E_n and E , we need some proper assumption on \mathbf{A}_n , which needs to be a proper discretization of the operator A . The assumption we need is given as follows

$$(5.7) \quad \lim_{n \rightarrow \infty} \|\mathbf{T}_n A u - \mathbf{A}_n \mathbf{T}_n u\| = 0 \quad \text{for all } u \in L_2(\Omega),$$

where $A : L_2(\Omega) \mapsto L_2(\Omega)$ is a continuous linear operator. Note that operator A that corresponds to image denoising, deblurring and inpainting indeed satisfies the above assumption [22].

The first relation between the wavelet frame based model (5.6) (equivalently (5.4)) and the variational model (5.2) is given by the following Theorem 5.1, which shows the pointwise convergence of $E_n(u)$ to $E(u)$ for each u . The proof of this theorem is technical and will be presented in a later part of this section.

Theorem 5.1. (*Pointwise Convergence*) *Given any tensor-product B-spline wavelet frame system and its associated energy functional $E_n(u)$ and assuming that (5.7) is satisfied, then, with proper choices of the parameters $\{\boldsymbol{\lambda}_j\}$, $\{\boldsymbol{\gamma}_j\}$, $\{\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}}\}$ and $\{\tilde{\boldsymbol{\gamma}}_{j,\tilde{j}}\}$,*

$$\lim_{n \rightarrow \infty} E_n(u) = E(u) \quad \text{for every } u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}).$$

If we assume that Theorem 5.1 is true, we can further show that the sequence E_n is equicontinuous.

Proposition 5.1. *Given any tensor-product B-spline wavelet frame system and its associated energy functional $E_n(u)$ with proper choices of the parameters $\{\boldsymbol{\lambda}_j\}$, $\{\boldsymbol{\gamma}_j\}$, $\{\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}}\}$ and $\{\tilde{\boldsymbol{\gamma}}_{j,\tilde{j}}\}$, and assuming that (5.7) is satisfied, then for an arbitrary $u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$ and any $\epsilon > 0$, there exist an integer \mathcal{N} and $\delta > 0$ (independent of n) such that for all $v \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$ satisfying $\|u - v\|_{\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})} < \delta$ and $n > \mathcal{N}$, we have $|E_n(v) - E_n(u)| < \epsilon$.*

Proof. Note that the property that needs to be shown for $E_n^{(1)}(u)$ and $E_n^{(4)}(u)$ follows from [22, Proposition 3.2] if the parameters $\{\boldsymbol{\lambda}_j\}$ and $\{\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}}\}$ are properly chosen. In addition, the proof of $E_n^{(3)}$ is entirely analogous to that of $E_n^{(2)}$. Therefore, we shall skip the repeated proof and focus on the proof of $E_n^{(2)}$.

Define the space $\ell_1^*(\mathbb{Z}) := \{\mathbf{b} : \|\mathbf{b}\|_1^* < +\infty\}$ with

$$\|\mathbf{b}\|_1^* = \sum_{k \in \mathbb{Z}} \left(\sum_{i \in \mathbb{B}_1} |\mathbf{b}_i[k]|^2 \right)^{\frac{1}{2}},$$

which can be regarded as a finite tensor of the space of all absolute summable sequences on \mathbb{Z} . By the definition of ℓ_1 -norm for $E_n^{(2)}$ (5.6) (except now the sequence is on \mathbb{Z} instead of \mathbb{Z}^2), we have

$$\|\mathbf{b}\|_1 = \sum_{k \in \mathbb{Z}} \left(\sum_{i \in \mathbb{B}_1} |\mathbf{b}_i[k]|^2 \right)^{\frac{1}{2}} 2^{-n}.$$

Since for any given n and $v \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, and for each $j = 1, 2, \dots, \tilde{m}$, we have $[\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j} \in \ell_1^*(\mathbb{Z})$ and

$$\|2^{-n} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j}\|_1^* = \|[\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j}\|_1.$$

Since \mathbf{T}_n is a bounded linear operator on $L_2(\Omega)$ and \mathbf{W}_n is a linear operator on a finite dimensional space, then we have

$$\|2^{-n} \lambda_n \cdot \mathbf{W}_n \mathbf{T}_n v\|_1^* \leq C_n \|v\|_{L_2(\Omega)} \leq \tilde{C}_n \|v\|_{\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})}.$$

This shows that

$$2^{-n} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n (\cdot)]_{\Gamma_j} : \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) \mapsto \ell_1^*(\mathbb{Z})$$

is a bounded linear operator. In addition, for any fixed $v \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, Theorem 5.1 gives us

$$\left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j} \right\|_1 \rightarrow \int_{\Gamma_j} \left| \mathfrak{T}_j^+(v) - \mathfrak{T}_j^-(v) \right| ds,$$

if the parameters $\{\gamma_j\}$ are properly chosen. Therefore,

$$\sup_n \|2^{-n} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j}\|_1^* = \sup_n \|[\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j}\|_1 < +\infty$$

for every $v \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$. By applying the uniform boundedness principle, we get

$$\sup_n \|2^{-n} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j}\|_{op} \leq C_j,$$

where $\|\cdot\|_{op}$ stands for the operator norm and C_j is a constant independent of n . Then,

$$\begin{aligned} |E_n^{(2)}(u) - E_n^{(2)}(v)| &= \left\| \sum_{j=1}^{\tilde{m}} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 - \left\| \sum_{j=1}^{\tilde{m}} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n v]_{\Gamma_j} \right\|_1 \\ &\leq \sum_{j=1}^{\tilde{m}} \left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n (u - v)]_{\Gamma_j} \right\|_1 = \sum_{j=1}^{\tilde{m}} \left\| 2^{-n} [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n (u - v)]_{\Gamma_j} \right\|_1^* \\ &\leq C \|u - v\|_{\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})}, \end{aligned}$$

with $C = \sum_{j=1}^{\tilde{m}} C_j$. This implies that $E_n^{(2)}$ is Lipschitz continuous with Lipschitz constant independent of n . Therefore, by choosing $N = 1$ and $\delta = \epsilon/C$, we shall have the desired property for $E_n^{(2)}$. \square

With Theorem 5.1 and Proposition 5.1, we can show that the convergence of E_n to E is stronger than merely pointwise. A direct consequence of such stronger convergence is the Gamma-convergence of E_n to E in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$. These results are summarized by Theorem 5.2. We shall first recall the definition of Gamma-convergence, followed by the statement of Theorem 5.2.

Definition 5.1. Given $E_n(u) : \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) \mapsto \bar{\mathbb{R}}$ and $E(u) : \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\}) \mapsto \bar{\mathbb{R}}$, we say that E_n Gamma-converges to E if:

- (i) for every sequence $u_n \rightarrow u$ in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, $E(u) \leq \liminf_{n \rightarrow \infty} E_n(u_n)$;
- (ii) for every $u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, there is a sequence $u_n \rightarrow u$ in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, such that $E(u) \geq \limsup_{n \rightarrow \infty} E_n(u_n)$.

Theorem 5.2. *Given any tensor-product B-spline wavelet frame system and its associated energy functional $E_n(u)$ (5.6) and assuming that (5.7) is satisfied, then, with proper choices of the parameters $\{\lambda_j\}$, $\{\gamma_j\}$, $\{\tilde{\lambda}_{j,\tilde{j}}\}$ and $\{\tilde{\gamma}_{j,\tilde{j}}\}$, we have $\lim_{n \rightarrow \infty} E_n(u_n) = E(u)$ for every sequence $u_n \rightarrow u$ in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$. Consequently, E_n Gamma-converges to E in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$.*

The proof of Theorem 5.2 follows similarly as in [22, Theorem 3.2] provided that the following Theorem 5.1 and Proposition 5.1 are established. We postpone the more technical proof of Theorem 5.1 to the latter part of this section. Also, when we say “with properly chosen parameters $\{\lambda_j\}$, $\{\gamma_j\}$, $\{\tilde{\lambda}_{j,\tilde{j}}\}$ and $\{\tilde{\gamma}_{j,\tilde{j}}\}$ ”, the exact meaning of it will be revealed by the proof of Theorem 5.1. In particular, the choices of $\{\gamma_j\}$ and $\{\tilde{\gamma}_{j,\tilde{j}}\}$ are given in the proof of Lemma 5.2.

In numerical computations, the task is to find an approximate minimizer, i.e., the one on which the value of the corresponding objective functional is close to its infimum. We say that u^* is an ϵ -optimal solution to a given objective functional E if

$$(5.8) \quad E(u^*) \leq \inf_u E(u) + \epsilon, \quad \text{for some } \epsilon > 0.$$

We say that u^* is a minimizer of E if $E(u^*) = \inf_u E(u)$. It is clear that ϵ -optimal solutions of E and E_n always exist because each of them has an infimum. In this paper, whenever we say that u^* is an approximate minimizer of E we mean that u^* is an ϵ -optimal solution to E with some sufficiently small ϵ . In other words, $E(u^*)$ is very close to $\inf E$.

The main task for both the variational and wavelet frame based approaches for image restoration is to find sparse approximate solutions in transform domains. These solutions are usually those at which the underlying objective functionals assume smaller values. The actual minimizers for either variational or wavelet frame based models are not the focus here. This is because in numerical computations, exact minimizers are usually difficult and, in fact, unnecessary to find. Indeed, most of the numerical algorithms solving these models are iterative in nature and the minimizer is only attained when the iteration goes to infinity. Therefore, in practice, we have to stop at a finite iteration, which inevitably leads to an approximate solution.

Since E_n pointwise converges to E by Theorem 5.1, it is natural to use E_n as a discrete approximation of E . Then it is natural to ask whether the approximate solutions to E can be approximated by those to E_n in some proper sense. The following corollary answers this question, which follows from Theorem 5.2.

Corollary 5.1. *Suppose the assumptions of Theorem 5.2 are satisfied. Let u_n^* be an ϵ -optimal solution of E_n for a given $\epsilon > 0$ and for all n .*

- (1) *We have*

$$(5.9) \quad \limsup_{n \rightarrow \infty} E_n(u_n^*) \leq \inf_u E(u) + \epsilon.$$

In particular, when u_n^ is a minimizer of E_n , we have*

$$\limsup_{n \rightarrow \infty} E_n(u_n^*) \leq \inf_u E(u).$$

- (2) *If, in addition, the set $\{u_n^*\}$ has a cluster point u^* , then u^* is an ϵ -optimal solution to E . In particular, when u_n^* is a minimizer of E_n and u^* a cluster point of the set $\{u_n^*\}$, then*

$$E(u^*) = \limsup_{n \rightarrow \infty} (E_n(u_n^*)) = \inf_u E(u)$$

and u^ is a minimizer of E .*

Proof. Part (1): For any $u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, let $\{u_n\}$ be the sequence as given in item (ii) of the definition of Γ -convergence. Then, we have

$$E(u) \geq \limsup_{n \rightarrow \infty} E_n(u_n) \geq \limsup_{n \rightarrow \infty} \left(\inf_u E_n(u) \right) \geq \limsup_{n \rightarrow \infty} E_n(u_n^*) - \epsilon,$$

which implies (5.9).

Part (2): If u^* is a cluster point of $\{u_n^*\}$, let $\{u_{n_k}^*\}$ be a subsequence of $\{u_n^*\}$ such that $u_{n_k}^* \rightarrow u^*$ in $\mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$ as $k \rightarrow \infty$. Then by item (i) of the definition of Γ -convergence, we have

$$E(u^*) \leq \liminf_{k \rightarrow \infty} E_{n_k}(u_{n_k}^*) \leq \limsup_{k \rightarrow \infty} E_{n_k}(u_{n_k}^*) \leq \limsup_{n \rightarrow \infty} E_n(u_n^*) \leq \inf_u E(u) + \epsilon,$$

where the last inequality follows from (5.9). This shows that u^* is an ϵ -optimal solution to E . \square

5.3. Proof of Theorem 5.1. To prove Theorem 5.1, we show that $E_n^{(q)}(u) \rightarrow E^{(q)}(u)$ for each $u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$ and $q = 1, 2, 3, 4$. Note that the convergence $E_n^{(4)}(u) \rightarrow E^{(4)}(u)$ for every $u \in L_2(\Omega)$ is guaranteed by [22, Lemma 3.2] under assumption (5.7). Therefore, we focus on the convergence of $E_n^{(q)}(u) \rightarrow E^{(q)}(u)$ for each $u \in$ for $q = 1, 2, 3$.

5.3.1. Convergence of $E_n^{(1)} \rightarrow E^{(1)}$. We first recall the following approximation lemma from [22].

Lemma 5.1. (*[22, Lemma 4.1]*) *Let φ and $\tilde{\varphi}$ be two compactly supported bounded functions satisfying $\int_{\mathbb{R}^2} \varphi d\mathbf{x} = 1$ and $\int_{\mathbb{R}^2} \tilde{\varphi} d\mathbf{x} = 1$. In addition, we assume that φ satisfies the partition of unity, i.e. $\sum_{\mathbf{j} \in \mathbb{Z}^2} \varphi(\cdot + \mathbf{j}) = 1$. Let \mathbb{S}^2 be the index set for \mathbf{k} such that $\varphi_{n,\mathbf{k}}$ is completely supported in \overline{B} with $B \subset \Omega$ an open Lipschitz domain whose boundary ∂B is piecewise C^1 . In addition, we assume that the support of φ is contained within the support of $\tilde{\varphi}$. Then, for any $u \in L_p(B)$, for $1 \leq p \leq \infty$, we have*

$$(5.10) \quad \lim_{n \rightarrow \infty} \left\| u - \sum_{\mathbf{k} \in \mathbb{S}^2} \langle u, \tilde{\varphi}_{n,\mathbf{k}} \rangle \varphi_{n,\mathbf{k}} \right\|_{L_p(B)} = 0.$$

Remark 5.1. In [22], the domain B is taken to be $(0,1)^2$. However, the proof of [22, Lemma 4.1] stays the same when B satisfies the assumptions in Lemma 5.1.

Recall that

$$E^{(1)}(u) = \sum_{j=1}^m \left[\sum_{|\mathbf{i}_j|=1} \|D_{\mathbf{i}_j} u\|_{L_2(\Omega_j)}^2 + \sum_{\tilde{j}=1}^{m_j} \left(\sum_{1 \leq |\mathbf{i}_{j,\tilde{j}}| \leq s_{j,\tilde{j}}} \|D_{\mathbf{i}_{j,\tilde{j}}} u\|_{L_2(\Omega_{j,\tilde{j}})}^2 \right) \right],$$

and

$$E_n^{(1)}(u) = \sum_{j=1}^m \left[\left\| [\boldsymbol{\lambda}_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_j} \right\|_2^2 + \sum_{\tilde{j}=1}^{m_j} \left\| [\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_{j,\tilde{j}}} \right\|_2^2 \right].$$

To show that for each $u \in \mathcal{H}^{1,s}(\{\Omega_{j,\tilde{j}}\})$, $E_n^{(1)}(u) \rightarrow E^{(1)}(u)$ as $n \rightarrow \infty$, it suffices to show

$$(5.11) \quad \left\| [\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Omega_{j,\tilde{j}}} \right\|_2^2 \rightarrow \sum_{1 \leq |\mathbf{i}_{j,\tilde{j}}| \leq s_{j,\tilde{j}}} \|D_{\mathbf{i}_{j,\tilde{j}}} u\|_{L_2(\Omega_{j,\tilde{j}})}^2, \quad \text{for each } j \text{ and } \tilde{j},$$

with a properly chosen $\tilde{\boldsymbol{\lambda}}_{j,\tilde{j}}$. Since the proof is identical for every j and \tilde{j} , for simplicity of notation, we drop the subscripts “ j, \tilde{j} ” in (5.11) and switch $\Omega_{j,\tilde{j}}$ to some generic domain $B \subset (0,1)^2$ satisfying the assumption in Lemma 5.1. Then, we consider the convergence

$$(5.12) \quad \left\| [\boldsymbol{\lambda} \cdot \mathbf{W}_n \mathbf{T}_n u]_B \right\|_2^2 \rightarrow \sum_{1 \leq |\mathbf{i}| \leq s} \|D_{\mathbf{i}} u\|_{L_2(B)}^2, \quad \text{for each } u \in H^s(B)$$

with a properly chosen $\boldsymbol{\lambda} = \{\lambda_i : i \in \mathbb{B}\}$. The proof of (5.12) is similar to that of [22, Lemma 3.3], except that: (1) the ℓ_1 -norm of the wavelet frame coefficients and the L_1 -norm of $D_i u$ was used in [22]; (2) the domain B was taken as $B = (0, 1)^2$ in [22]. For completeness of this paper, we present a detailed proof here.

Consider

$$\begin{aligned} \|[\boldsymbol{\lambda} \cdot \mathbf{W} \mathbf{u}]_B\|_2^2 &= h^2 \sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{B}} \lambda_i |\mathbf{k}| \left| (\mathbf{W}_{n,i} \mathbf{u})[\mathbf{k}] \right|^2 \\ &= h^2 \sum_{\mathbf{k} \in \mathbb{K}^2} \left(\sum_{i \in \mathbb{I}} \lambda_i |(\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}]|^2 + \sum_{j \in \mathbb{J}} \lambda_j |(\mathbf{a}_j[-\cdot] * \mathbf{T}_n u)[\mathbf{k}]|^2 \right), \end{aligned}$$

where we have dropped the subscript “ j, \tilde{j} ” of $\mathbb{K}_{j, \tilde{j}}^2$ as well. From now on, the index set $\mathbb{M}^2 \subset \mathbb{O}^2$ is such that $\Lambda_{n,\mathbf{k}}$ is completely supported in \overline{B} , and $\mathbb{K}^2 \subset \mathbb{M}^2$ is such that the boundary condition of $\mathbf{a}_i[-\cdot] \otimes \mathbf{u}$ is inactive for all i . We define the index sets \mathbb{I} and \mathbb{J} as

$$\mathbb{I} := \{i : D_i \text{ is in } \mathbf{D}\} \quad \text{and} \quad \mathbb{J} := \mathbb{B} \setminus \mathbb{I}.$$

Note that we may choose $\lambda_j = 0$ for $j \in \mathbb{J}$ and the proof will be simpler. However, as shown in [22,23] that having inactive wavelet frame bands \mathbb{J} in the system is beneficial in image restoration, or in other words, it makes a better discretization of $\mathbf{D}u$.

Given a system of tensor-product B-spline wavelet frame system $\boldsymbol{\Psi} = \{\psi_i : i \in \mathbb{B}\}$, there exists an s_i -differentiable function φ_i for each $i \in \mathbb{B}$ with $c_i = \int_{\mathbb{R}^2} \varphi_i d\mathbf{x} \neq 0$ such that $D_i \varphi_i = \psi_i$ and φ_i has the same support as ψ_i (see [22] for more details). Same as in [22], the parameter $\boldsymbol{\lambda}$ needs to be properly chosen. We choose $\lambda_{0,0} = 0$. For every $i \in \mathbb{I}$, we set $\lambda_i = \frac{1}{c_i} 2^{2s_i(n-1)}$. For every $j \in \mathbb{J}$, we set $0 \leq \lambda_j \leq O(2^{2s_{j'}(n-1)})$ for some $j' \in \mathbb{B} \cup \{\mathbf{0}\}$ such that $\mathbf{0} \leq j' < j$ and $s_{j'} \leq s$.

Let us first show (5.12) when $\mathbb{J} = \emptyset$. Note that $D_i \varphi_i = \psi_i$ implies

$$D_i \varphi_{i,n-1,\mathbf{k}} = 2^{(s_i+1)(n-1)-1} \psi_i(2^{n-1} \cdot -\mathbf{k}/2) = 2^{s_i(n-1)} \psi_{i,n-1,\mathbf{k}}.$$

Since φ_i is smooth and compactly supported and both $\psi_{i,n-1,\mathbf{k}}$ and $\varphi_{i,n-1,\mathbf{k}}$ are supported on B for $\mathbf{k} \in \mathbb{K}^2$, according to the definition of weak derivatives, we have

$$\langle D_i u, \varphi_{i,n-1,\mathbf{k}} \rangle = (-1)^{s_i} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle = (-1)^{s_i} 2^{s_i(n-1)} \langle u, \psi_{i,n-1,\mathbf{k}} \rangle \quad \text{for } \mathbf{k} \in \mathbb{K}^2.$$

Notice that

$$\begin{aligned} \lambda_i (\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}] &= \lambda_i \sum_{j \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}] (\mathbf{T}_n u)[\mathbf{j}] = 2^n \lambda_i \sum_{j \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}] \langle u, \phi_{n,j} \rangle \\ &= 2^n \lambda_i \langle u, \sum_{j \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}] \phi_{n,j} \rangle = 2^n \lambda_i \langle u, \psi_{i,n-1,\mathbf{k}} \rangle, \end{aligned}$$

where S_i is the support of the filter \mathbf{a}_i . Thus,

$$\lambda_i |(\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}]|^2 = 2^{2n} \frac{1}{c_i} 2^{2s_i(n-1)} |\langle u, \psi_{i,n-1,\mathbf{k}} \rangle|^2,$$

and hence

$$\begin{aligned} \|\boldsymbol{\lambda}_n \cdot \mathbf{W}_n \mathbf{T}_n u\|_2^2 &= h^2 \sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \lambda_i |(\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}]|^2 \\ &= \sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} |\langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle|^2. \end{aligned}$$

Let $\square_{\mathbf{k}}$ be 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 $\mathbf{k} = (k_1, k_2)$. Since $\square_{\mathbf{k}} \subset \Lambda_{n,\mathbf{k}}$, we have $\square_{\mathbf{k}} \subset B$ for all $\mathbf{k} \in \mathbb{K}^2$. Note that for any $a, b, c \geq 0$, we have the inequality $|\sqrt{a+b} - \sqrt{c}| \leq$

$|\sqrt{a} - \sqrt{c}| + \sqrt{b}$. Then, we have

$$\begin{aligned}
(5.13) \quad & \left| \|Du\|_2 - \|\lambda_n \cdot W_n T_n u\|_2 \right| \\
&= \left| \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \int_{\square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} + \int_{B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} - \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} |\langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} \right| \\
&\leq \left| \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \int_{\square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} - \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} |\langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} \right| + \left(\int_{B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} \\
&= \left| \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} - \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} |2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle|^2 d\mathbf{x} \right)^{\frac{1}{2}} \right| + \left(\int_{B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} \left| D_i u - 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \right|^2 \right)^{\frac{1}{2}} + \left(\int_{B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}}.
\end{aligned}$$

It is easy to show that $\lim_{n \rightarrow \infty} \mathfrak{L}(B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}) = 0$. Therefore,

$$\left(\int_{B \setminus \cup_{\mathbf{k} \in \mathbb{K}^2} \square_{\mathbf{k}}} \sum_{i \in \mathbb{I}} |D_i u|^2 d\mathbf{x} \right)^{\frac{1}{2}} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

It remains to show that

$$\left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} \left| D_i u - 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \right|^2 \right)^{\frac{1}{2}} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Let $\mathbb{S}^2 \subset \mathbb{O}^2$ be the index set of \mathbf{k} such that $\square_{\mathbf{k}}$ is completely supported in \overline{B} . Since $\mathbb{K}^2 \subset \mathbb{S}^2$, we have

$$\begin{aligned}
(5.14) \quad & \left(\sum_{\mathbf{k} \in \mathbb{K}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} \left| D_i u - 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \right|^2 \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{\mathbf{k} \in \mathbb{S}^2} \sum_{i \in \mathbb{I}} \int_{\square_{\mathbf{k}}} \left| D_i u - 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \right|^2 \right)^{\frac{1}{2}} \\
&= \left(\sum_{i \in \mathbb{I}} \int_B \left| D_i u \chi_{\cup_{\mathbf{k} \in \mathbb{S}^2} \square_{\mathbf{k}}} - \sum_{\mathbf{k} \in \mathbb{S}^2} 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \chi_{\square_{\mathbf{k}}} \right|^2 d\mathbf{x} \right)^{\frac{1}{2}} \\
&\leq \sum_{i \in \mathbb{I}} \left\| D_i u \chi_{\cup_{\mathbf{k} \in \mathbb{S}^2} \square_{\mathbf{k}}} - \sum_{\mathbf{k} \in \mathbb{S}^2} 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \chi_{\square_{\mathbf{k}}} \right\|_{L_2(B)} \\
&\leq \sum_{i \in \mathbb{I}} \left(\left\| D_i u \chi_{\cup_{\mathbf{k} \in \mathbb{S}^2} \square_{\mathbf{k}}} - D_i u \right\|_{L_2(B)} + \left\| D_i u - \sum_{\mathbf{k} \in \mathbb{S}^2} 2^n \langle D_i u, \frac{1}{c_i} \varphi_{i,n-1,\mathbf{k}} \rangle \chi_{\square_{\mathbf{k}}} \right\|_{L_2(B)} \right).
\end{aligned}$$

The identity above follows from the fact that $\mathfrak{L}(\square_{\mathbf{k}} \cap \square_{\mathbf{j}}) = 0$ for $\mathbf{k} \neq \mathbf{j}$. It remains to show that

$$(5.15) \quad \lim_{n \rightarrow \infty} \left\| D_{\mathbf{i}} u \chi_{\cup_{\mathbf{k} \in \mathbb{S}^2} \square_{\mathbf{k}}} - D_{\mathbf{i}} u \right\|_{L_2(B)} = 0$$

and

$$(5.16) \quad \lim_{n \rightarrow \infty} \left\| D_{\mathbf{i}} u - \sum_{\mathbf{k} \in \mathbb{S}^2} 2^n \langle D_{\mathbf{i}} u, \frac{1}{c_{\mathbf{i}}} \varphi_{\mathbf{i}, n-1, \mathbf{k}} \rangle \chi_{\square_{\mathbf{k}}} \right\|_{L_2(B)} = 0, \quad \text{for } \mathbf{i} \in \mathbb{I}.$$

For (5.15), it is easy to show that $\lim_{n \rightarrow \infty} \chi_{\cup_{\mathbf{k} \in \mathbb{S}^2} \square_{\mathbf{k}}} = 1$ on B . Therefore, by the Lebesgue dominated convergence theorem, we get (5.15). For (5.16), notice that $2^n \chi_{\square_{\mathbf{k}}} = \phi_{n, \mathbf{k}}^{(H)}$, where $\phi^{(H)}$ is the characteristic function on the unit square, i.e., the tensor product of the piecewise constant B-spline which satisfies the partition of unity property. Furthermore, by the definition of quasi-affine systems given in (3) of Notation 5.1, we have

$$\frac{1}{c_{\mathbf{i}}} \varphi_{\mathbf{i}, n-1, \mathbf{k}} = \frac{1}{4c_{\mathbf{i}}} 2^n \varphi_{\mathbf{i}}(2^{n-1}x - \mathbf{k}/2) = \left(\frac{\varphi_{\mathbf{i}}(\cdot/2)}{4c_{\mathbf{i}}} \right)_{n, \mathbf{k}}$$

and $\int_B \frac{\varphi_{\mathbf{i}}(\cdot/2)}{4c_{\mathbf{i}}} d\mathbf{x} = 1$. We also note that the support of $\varphi_{\mathbf{i}}(\cdot/2)$ contains the support of $\phi^{(H)}$. Together with $D_{\mathbf{i}} u \in L_2(B)$, we establish (5.16) by Lemma 5.1.

In the case of $\mathbb{J} \neq \emptyset$, if we can show that, for all $\mathbf{j} \in \mathbb{J}$,

$$(5.17) \quad \lim_{n \rightarrow \infty} \lambda_{\mathbf{j}} \| \mathbf{a}_{\mathbf{j}}[-\cdot] * \mathbf{T}_n u \|_2^2 = 0,$$

then we get (5.12). Note that there exist $\varphi_{\mathbf{j}}$ and $\varphi_{\mathbf{j}'}$ such that $D_{\mathbf{j}} \varphi_{\mathbf{j}} = \psi_{\mathbf{j}}$ and $D_{\mathbf{j}'} \varphi_{\mathbf{j}'} = \psi_{\mathbf{j}'}$. Choose \mathbf{j}' satisfying $\mathbf{0} \leq \mathbf{j}' < \mathbf{j}$ and $s_{\mathbf{j}'} \leq s$. Note that such a \mathbf{j}' always exists, since, for example, one may pick $\mathbf{j}' = \mathbf{0}$. Let $\bar{\psi}_{\mathbf{j}} = D_{\mathbf{j}-\mathbf{j}'} \varphi_{\mathbf{j}}$. Then obviously we have $D_{\mathbf{j}'} \bar{\psi}_{\mathbf{j}} = \psi_{\mathbf{j}}$, due to the tensor product structure of $\varphi_{\mathbf{j}}$ that ensures $D_{\mathbf{j}'} D_{\mathbf{j}-\mathbf{j}'} \varphi_{\mathbf{j}} = D_{\mathbf{j}} \varphi_{\mathbf{j}}$. For any real number $t \geq 0$, the function

$$\tilde{\varphi}_t := \frac{1}{c_{\mathbf{j}'}} \varphi_{\mathbf{j}'} + t \bar{\psi}_{\mathbf{j}},$$

with $c_{\mathbf{j}'} = \int_{\mathbb{R}^2} \varphi_{\mathbf{j}'} d\mathbf{x}$, is smooth, compactly supported, and of integral 1 (since obviously $\int_{\mathbb{R}^2} \bar{\psi}_{\mathbf{j}} d\mathbf{x} = 0$). This together with $D_{\mathbf{j}'} \tilde{\varphi}_t = \frac{1}{c_{\mathbf{j}'}} \psi_{\mathbf{j}'} + t \psi_{\mathbf{j}}$ leads to, for $u \in H^s(B)$,

$$\langle D_{\mathbf{j}'} u, \tilde{\varphi}_{t, n-1, \mathbf{k}} \rangle = (-1)^{s_{\mathbf{j}'}} 2^{s_{\mathbf{j}'(n-1)}} \langle u, \frac{1}{c_{\mathbf{j}'}} \psi_{\mathbf{j}', n-1, \mathbf{k}} + t \psi_{\mathbf{j}, n-1, \mathbf{k}} \rangle.$$

Therefore,

$$\| 2^{s_{\mathbf{j}'(n-1)}} \left(\frac{1}{c_{\mathbf{j}'}} \mathbf{a}_{\mathbf{j}'}[-\cdot] + t \mathbf{a}_{\mathbf{j}}[-\cdot] \right) * \mathbf{T}_n u \|_2^2 = \sum_{\mathbf{k} \in \mathbb{K}^2} |\langle D_{\mathbf{j}'} u, \tilde{\varphi}_{t, n-1, \mathbf{k}} \rangle|^2.$$

Following the exact same steps as in (5.13) and (5.14), by removing the summation with respect to $\mathbf{i} \in \mathbb{I}$, replacing φ by $\tilde{\varphi}_t$, $D_{\mathbf{i}}$ by $D_{\mathbf{j}'}$ and $c_{\mathbf{i}}$ by $c_{\mathbf{j}'}$, we have

$$\| D_{\mathbf{j}'} u \|_2 = \lim_{n \rightarrow \infty} \| 2^{s_{\mathbf{j}'(n-1)}} \left(\frac{1}{c_{\mathbf{j}'}} \mathbf{a}_{\mathbf{j}'}[-\cdot] + t \mathbf{a}_{\mathbf{j}}[-\cdot] \right) * \mathbf{T}_n u \|_2.$$

In particular, when $t = 0$, we have

$$\| D_{\mathbf{j}'} u \|_2 = \lim_{n \rightarrow \infty} \| 2^{s_{\mathbf{j}'(n-1)}} \frac{1}{c_{\mathbf{j}'}} \mathbf{a}_{\mathbf{j}'}[-\cdot] * \mathbf{T}_n u \|_2.$$

These two equations imply that

$$t \limsup_{n \rightarrow \infty} \| 2^{s_{\mathbf{j}'(n-1)}} \mathbf{a}_{\mathbf{j}}[-\cdot] * \mathbf{T}_n u \|_2 \leq 2 \| D_{\mathbf{j}'} u \|_2.$$

Since t is arbitrary, we must have

$$\lim_{n \rightarrow \infty} \|2^{s_{j'}(n-1)} \mathbf{a}_j[-\cdot] * \mathbf{T}_n u\|_2 = 0.$$

In view of $0 \leq \lambda_j \leq O(2^{2s_{j'}(n-1)})$, we get (5.17). This concludes the proof of the convergence $E_n^{(1)} \rightarrow E^{(1)}$.

5.3.2. *Convergence of $E_n^{(2)} \rightarrow E^{(2)}$.* It suffices to show that

$$(5.18) \quad \left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 \rightarrow \int_{\Gamma_j} \left| \mathfrak{T}_j^+(u) - \mathfrak{T}_j^-(u) \right| ds, \quad \text{for each } j.$$

Notice that

$$\begin{aligned} \gamma_{i,j}(\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}] &= \gamma_{i,j} \sum_{\mathbf{j} \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}](\mathbf{T}_n u)[\mathbf{j}] = 2^n \gamma_{i,j} \sum_{\mathbf{j} \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}] \langle u, \phi_{n,j} \rangle \\ &= 2^n \gamma_{i,j} \langle u, \sum_{\mathbf{j} \in S_i + \mathbf{k}} \mathbf{a}_i[\mathbf{j} - \mathbf{k}] \phi_{n,j} \rangle = 2^n \gamma_{i,j} \langle u, \psi_{i,n-1,\mathbf{k}} \rangle, \end{aligned}$$

where S_i is the support of the filter \mathbf{a}_i . Let φ_i be such that $D_i \varphi_i = \psi_i$ and $\text{supp} \varphi_i = \text{supp} \psi_i$. Then

$$D_i \varphi_{i,n-1,\mathbf{k}} = 2^{s_i(n-1)} \psi_{i,n-1,\mathbf{k}}.$$

and thus, for each $\mathbf{k} \in \mathbb{G}_j^2$, we have

$$\gamma_{i,j}[\mathbf{k}] |(\mathbf{a}_i[-\cdot] * \mathbf{T}_n u)[\mathbf{k}]|^2 = 2^{2n} \gamma_{i,j}[\mathbf{k}] |\langle u, \psi_{i,n-1,\mathbf{k}} \rangle|^2 = 2^{2s_i(1-n)} 2^{2n} \gamma_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2.$$

Thus

$$\begin{aligned} \left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 &= h \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left(\sum_{\mathbf{i} \in \mathbb{B}_1} \gamma_{i,j}[\mathbf{k}] \left| (\mathbf{W}_{n,i} \mathbf{T}_n u)[\mathbf{k}] \right|^2 \right)^{\frac{1}{2}} \\ &= \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left(\sum_{\mathbf{i} \in \mathbb{B}_1} 2^{2s_i(1-n)} \gamma_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

To apply Proposition 4.2, we split $\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle$ into two parts as follows

$$\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle = \int_{\Omega_j^+ \cap \Lambda_{n-1,\mathbf{k}}} u(\mathbf{x}) (D_i \varphi_{i,n-1,\mathbf{k}})(\mathbf{x}) d\mathbf{x} + \int_{\Omega_j^- \cap \Lambda_{n-1,\mathbf{k}}} u(\mathbf{x}) (D_i \varphi_{i,n-1,\mathbf{k}})(\mathbf{x}) d\mathbf{x},$$

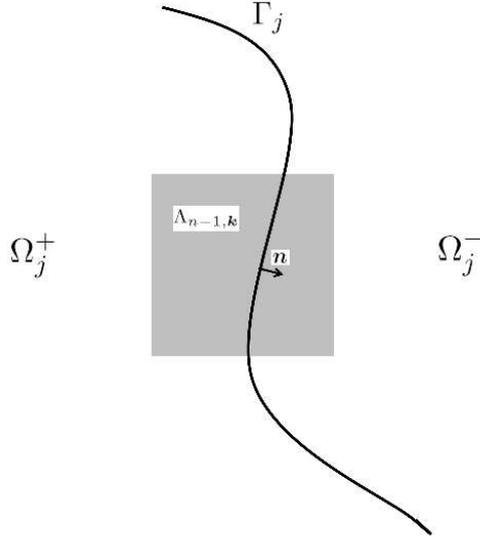
where Ω_j^\pm , $\Lambda_{n-1,\mathbf{k}}$ and Γ_j are illustrated in Figure 11. By the definition of the set \mathbb{G}_j^2 , for any $\mathbf{k} \in \mathbb{G}_j^2$, the above splitting of the integration domain is always valid. Now, we apply Proposition 4.2 for the case $|\mathbf{i}| = 1$ to each of the two integrals and obtain

$$(5.19) \quad \begin{aligned} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle &= \int_{\Gamma_j \cap \Lambda_{n-1,\mathbf{k}}} v_j \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_i ds \\ &\quad - \int_{\Omega_j^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}} - \int_{\Omega_j^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}}. \end{aligned}$$

where

$$v_j = \mathfrak{T}_j^+(u) - \mathfrak{T}_j^-(u) \in L_2(\Gamma_j) \subset L_1(\Gamma_j)$$

and \mathfrak{T}_j^\pm is the trace operator defined on $H^{s_j}(\Omega_j^\pm)$.

FIGURE 11. Illustrations of domains Ω_j^\pm , $\Gamma_{n-1,\mathbf{k}}$ and curve Γ_j .

Consider,

$$\begin{aligned}
& \left| \left\| [\boldsymbol{\gamma}_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 - \int_{\Gamma_j} |v_j| ds \right| \\
&= \left| \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left(\sum_{i \in \mathbb{B}_1} 2^{2s_i(1-n)} \gamma_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \int_{\Gamma_j} |v_j| ds \right| \\
&= \left| \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left[\left(\sum_{i \in \mathbb{B}_1} 2^{2s_i(1-n)} \gamma_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \int_{\Gamma_j \cap \square_{n,\mathbf{k}}} |v_j| ds \right] \right| + o(1) \\
&= \left| \sum_{\mathbf{k} \in \mathbb{G}_j^2} \left[\left(\sum_{i \in \mathbb{B}_1} 2^{2s_i(1-n)} \gamma_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \int_{\Gamma_j \cap \square_{n,\mathbf{k}}} \left(\sum_{i \in \mathbb{B}_1} |v_j \mathbf{n}_i|^2 \right)^{\frac{1}{2}} ds \right] \right| + o(1),
\end{aligned}$$

where $\square_{n,\mathbf{k}}$ is the rectangular domain $(\frac{k_1-1/2}{2^n}, \frac{k_1+1/2}{2^n}] \times (\frac{k_2-1/2}{2^n}, \frac{k_2+1/2}{2^n}]$ with $\mathbf{k} = (k_1, k_2) \in \mathbb{M}^2$. The second identity above follows from the fact that $\square_{n,\mathbf{k}} \subset \Lambda_{n-1,\mathbf{k}}^\circ$. Indeed, whenever $\square_{n,\mathbf{k}}$ has a nonempty intersection with Γ_j , so does $\Lambda_{n-1,\mathbf{k}}^\circ$ except only when $\mathbf{k} \in \mathbb{G}_j^2$ is near the boundary of Ω or a junction of multiple $\Gamma_{j'}$. Therefore, we have $\bigcup_{\mathbf{k} \in \mathbb{G}_j^2} \Gamma_j \cap \square_{n,\mathbf{k}} \rightarrow \Gamma_j$ as $n \rightarrow \infty$, in the sense that

$$\sum_{\mathbf{k} \in \mathbb{G}_j^2} \chi_{\Gamma_j \cap \square_{n,\mathbf{k}}}(s) \rightarrow 1 \quad \text{for almost all } s \in \Gamma_j.$$

In fact, the above convergence holds everywhere on Γ_j except at the end-points of Γ_j .

Notice that Γ_j may have an empty intersection with $\square_{n,\mathbf{k}}$ for $\mathbf{k} \in \mathbb{G}_j^2$. Define the index set $\mathbb{H}_j^2 \subset \mathbb{G}_j^2$ be such that

$$\mathbb{H}_j^2 := \{ \mathbf{k} \in \mathbb{G}_j^2 : \mathcal{L}(\Gamma_j \cap \square_{n,\mathbf{k}}) = 0 \},$$

where \mathcal{L} is the 1-D Lebesgue measure. Note that \mathbb{H}_j^2 is not generally an empty set (see Figure 12). For the rest of the proof we choose

$$\gamma_{i,j}[\mathbf{k}] = 0, \quad \text{for } \mathbf{k} \in \mathbb{H}_j^2, \quad i \in \mathbb{B}_1.$$

Letting

$$(5.20) \quad w_{i,j}[\mathbf{k}] = \frac{2^{2s_i(1-n)}}{\mathcal{L}^2(\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}})} \gamma_{i,j}[\mathbf{k}], \quad \text{for } \mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2,$$

we have

$$\begin{aligned} & \left| \left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 - \int_{\Gamma_j} |v_j| ds \right| \\ &= \left| \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \left[\left(\sum_{i \in \mathbb{B}_1} w_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \left(\sum_{i \in \mathbb{B}_1} |v_j \mathbf{n}_i|^2 \right)^{\frac{1}{2}} \right] ds \right| + o(1) \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \left| \left(\sum_{i \in \mathbb{B}_1} w_{i,j}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \left(\sum_{i \in \mathbb{B}_1} |v_j \mathbf{n}_i|^2 \right)^{\frac{1}{2}} \right| ds + o(1) \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \sum_{i \in \mathbb{B}_1} \left| \sqrt{w_{i,j}[\mathbf{k}]} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle - v_j \mathbf{n}_i \right| ds + o(1). \end{aligned}$$

Using the integration by parts formula (5.19), we have

$$\begin{aligned} & \left| \left\| [\gamma_j \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_j} \right\|_1 - \int_{\Gamma_j} |v_j| ds \right| \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \left[\sum_{i \in \mathbb{B}_1} \left| \sqrt{w_{i,j}[\mathbf{k}]} \int_{\Gamma_j \cap \Lambda_{n-1,\mathbf{k}}} v_j \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_i d\tilde{s} - v_j \mathbf{n}_i \right| \right] ds \\ &+ \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \left[\sum_{i \in \mathbb{B}_1} \sqrt{w_{i,j}[\mathbf{k}]} \left| \int_{\Omega_j^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u, \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_j^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u, \varphi_{i,n-1,\mathbf{k}} \right| \right] ds + o(1) \\ &=: G_n^{(1)}(u) + G_n^{(2)}(u) + o(1). \end{aligned}$$

We first show that there exists $w_{i,j}[\mathbf{k}]$, i.e. $\gamma_{i,j}[\mathbf{k}]$ for $\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2$, such that $G_n^{(1)}(u) \rightarrow 0$ as $n \rightarrow \infty$. Consider

$$\begin{aligned} G_n^{(1)}(u) &= \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \sum_{i \in \mathbb{B}_1} \left| \sqrt{w_{i,j}[\mathbf{k}]} \int_{\Gamma_j \cap \Lambda_{n-1,\mathbf{k}}} v_j \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_i d\tilde{s} - v_j \mathbf{n}_i \right| ds \\ &= \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \int_{\Gamma_j} \sum_{i \in \mathbb{B}_1} \left| \sqrt{w_{i,j}[\mathbf{k}]} \left(\int_{\Gamma_j \cap \Lambda_{n-1,\mathbf{k}}} v_j \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_i d\tilde{s} \right) \chi_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} - v_j \mathbf{n}_i \chi_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} \right| ds \\ &= \int_{\Gamma_j} \sum_{i \in \mathbb{B}_1} \left| \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \sqrt{w_{i,j}[\mathbf{k}]} \left(\int_{\Gamma_j \cap \Lambda_{n-1,\mathbf{k}}} v_j \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_i d\tilde{s} \right) \chi_{\Gamma_j \cap \tilde{\square}_{n,\mathbf{k}}} - v_j \mathbf{n}_i \right| ds + o(1), \end{aligned}$$

where the last identity follows from the fact that $\{\Gamma_j \cap \square_{n,\mathbf{k}} : \mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2\}$ are disjoint for different \mathbf{k} , the curve Γ_j is of finite length, and

$$(5.21) \quad \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \chi_{\Gamma_j \cap \square_{n,\mathbf{k}}}(s) \rightarrow 1 \quad \text{for almost all } s \in \Gamma_j.$$

Notice that for $\mathbf{i} \in \mathbb{B}_1$, $s_{\mathbf{i}} = |\mathbf{i}| = 1$. Letting

$$(5.22) \quad K_n(s, \tilde{s}) := \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \sqrt{w_{\mathbf{i},j}[\mathbf{k}]} \varphi_{\mathbf{i},n-1,\mathbf{k}}(\tilde{s}) \chi_{\Gamma_j \cap \square_{n,\mathbf{k}}}(s) = \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \frac{2^{(1-n)} \sqrt{\gamma_{\mathbf{i},j}[\mathbf{k}]}}{\mathfrak{L}(\Gamma_j \cap \square_{n,\mathbf{k}})} \varphi_{\mathbf{i},n-1,\mathbf{k}}(\tilde{s}) \chi_{\Gamma_j \cap \square_{n,\mathbf{k}}}(s),$$

then we have

$$G_n^{(1)}(u) = \sum_{\mathbf{i} \in \mathbb{B}_1} \int_{\Gamma_j} \left| \int_{\Gamma_j} v_j(\tilde{s}) \mathbf{n}_{\mathbf{i}}(\tilde{s}) K_n(s, \tilde{s}) d\tilde{s} - v_j(s) \mathbf{n}_{\mathbf{i}}(s) \right| ds + o(1).$$

To show that $G_n^{(1)}(u) \rightarrow 0$ as $n \rightarrow \infty$, we need the following lemma.

Lemma 5.2. *Let $K_n(s, \tilde{s})$ be defined in (5.22) with $\mathbf{i} \in \mathbb{B}_1 = \{(1, 0)^\top, (0, 1)^\top\}$ and $\varphi_{\mathbf{i}}$ being nonnegative and continuous.*

(1) *There exists $0 \leq \gamma_{\mathbf{i},j} \leq O(1)$ such that, as $n \rightarrow \infty$,*

$$\int_{\Gamma_j} K_n(s, \tilde{s}) d\tilde{s} \rightarrow 1 \quad \text{for almost all } s \in \Gamma_j.$$

(2) *For any $\delta > 0$ and an n large enough, we have*

$$K_n(s, \tilde{s}) = 0$$

for $(s, \tilde{s}) \in \{(s, \tilde{s}) \in \Gamma_j \times \Gamma_j, |s - \tilde{s}| \geq \delta\}$.

(3) *We have $\int_{\Gamma_j} |K_n(s, \tilde{s})| d\tilde{s} \leq M$ uniformly for all $s \in \Gamma_j$*

Consequently, with the choices of $\{\gamma_{\mathbf{i},j}\}$ in (1),

$$(5.23) \quad \lim_{n \rightarrow \infty} \left\| \int_{\Gamma_j} K_n(\cdot, \tilde{s}) f(\tilde{s}) d\tilde{s} - f(\cdot) \right\|_{L_1(\Gamma_j)} = 0,$$

for any $f \in L_1(\Gamma_j)$.

Proof. We first note that, for each $\mathbf{i} \in \mathbb{B}_1$ and $\mathbf{k} \in \mathbb{G}_j^2$, $\varphi_{\mathbf{i},n-1,\mathbf{k}}$ is nonnegative and continuous on the interior of its support $\Lambda_{n-1,\mathbf{k}}$. Consider (1). Let $I_{j,n,\mathbf{k}} := \Gamma_j \cap \square_{n,\mathbf{k}}$. We have

$$\int_{\Gamma_j} K_n(s, \tilde{s}) d\tilde{s} = \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \left[\frac{2^{(1-n)} \sqrt{\gamma_{\mathbf{i},j}[\mathbf{k}]}}{\mathfrak{L}(I_{j,n,\mathbf{k}})} \int_{\Gamma_j} \varphi_{\mathbf{i},n-1,\mathbf{k}}(\tilde{s}) d\tilde{s} \right] \chi_{I_{j,n,\mathbf{k}}}(s).$$

Then, if we choose

$$(5.24) \quad \sqrt{\gamma_{\mathbf{i},j}[\mathbf{k}]} = \begin{cases} 2^{n-1} \mathfrak{L}(I_{j,n,\mathbf{k}}) \left(\int_{\Gamma_j} \varphi_{\mathbf{i},n-1,\mathbf{k}}(\tilde{s}) d\tilde{s} \right)^{-1} & \text{for } \mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2 \\ 0 & \text{otherwise,} \end{cases}$$

equation (5.21) leads to $\int_{\Gamma_j} K_n(s, \tilde{s}) d\tilde{s} \rightarrow 1$ for almost all $s \in \Gamma_j$. It now remains to show that $\gamma_{j,\mathbf{k}} = O(1)$ for $\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2$, which means $\gamma_{j,\mathbf{k}}$ is well defined and is a reasonable parameter to choose in practice.

Let $\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2$ be fixed. Since $\mathfrak{L}(I_{j,n,\mathbf{k}}) = O(2^{-n})$, it suffices to show that $\int_{\Gamma_j} \varphi_{\mathbf{i},n-1,\mathbf{k}}(\tilde{s}) d\tilde{s} \geq C$ for some $C > 0$. Let $\Delta_{n,\mathbf{k}}$ be the square box whose center is the common center of $\square_{n,\mathbf{k}}$ and $\Lambda_{n-1,\mathbf{k}}$ and side length their average (see Figure 12). Then, since φ is strictly positive in the interior of its support, $\varphi_{\mathbf{i},n-1,\mathbf{k}}(\mathbf{x}) \geq C_1 2^{n-2}$ for all $\mathbf{x} \in \Delta_{n,\mathbf{k}}$, where $C_1 > 0$ is a constant. Furthermore, since

$\Gamma_j \cap \square_{n,\mathbf{k}} \neq \emptyset$ and $\Gamma_j \cap \Lambda_{n-1,\mathbf{k}} \neq \emptyset$, for sufficiently large n , the continuous curve Γ_j passes through the boundary of $\Lambda_{n-1,\mathbf{k}}$ and an interior point of $\square_{n,\mathbf{k}}$. Therefore, $\mathfrak{L}(\Gamma_j \cap \Delta_{n,\mathbf{k}}) \geq C_2 2^{-n}$. Finally,

$$(5.25) \quad \int_{\Gamma_j} \varphi_{i,n-1,\mathbf{k}}(\tilde{s}) d\tilde{s} \geq C_1 2^{n-2} C_2 2^{-n} = C_1 C_2 / 4 > 0.$$

For (2), we observe that for any $s, \tilde{s} \in \Gamma_j$, when n is large enough, we should have either $\varphi_{i,n-1,\mathbf{k}}(\tilde{s}) = 0$ or $\chi_{I_{j,n,\mathbf{k}}}(s) = 0$, or both, on $\{|s - \tilde{s}| \geq \delta\}$ for $\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2$.

For (3), the uniform boundedness of $\int_{\Gamma_j} |K_n(s, \tilde{s})| d\tilde{s}$ is obvious from the proof of (1) and the fact that Γ_j is of finite length.

Finally, by (1) and (3) and the fact that Γ_j is of finite length, the Lebesgue dominated convergence theorem implies that

$$\left\| f(\cdot) - f(\cdot) \int_{\Gamma_j} K_n(\cdot, \tilde{s}) d\tilde{s} \right\|_{L_1(\Gamma_j)} = o(1).$$

Therefore, for some $\delta > 0$,

$$\begin{aligned} \left\| \int_{\Gamma_j} K_n(\cdot, \tilde{s}) f(\tilde{s}) d\tilde{s} - f(\cdot) \right\|_{L_1(\Gamma_j)} &\leq \int_{\Gamma_j} \left| \int_{\Gamma_j} K_n(s, \tilde{s}) f(\tilde{s}) d\tilde{s} - f(s) \int_{\Gamma_j} K_n(s, \tilde{s}) d\tilde{s} \right| ds + o(1) \\ &\leq \int_{\Gamma_j} \int_{\Gamma_j} K_n(s, \tilde{s}) |f(\tilde{s}) - f(s)| d\tilde{s} ds + o(1) \\ &= \int_{\Gamma_j} \int_{|s-\tilde{s}| \geq \delta} K_n(s, \tilde{s}) |f(\tilde{s}) - f(s)| d\tilde{s} ds \\ &\quad + \int_{\Gamma_j} \int_{|s-\tilde{s}| < \delta} K_n(s, \tilde{s}) |f(\tilde{s}) - f(s)| d\tilde{s} ds + o(1). \end{aligned}$$

For n large enough, part (2) implies that

$$\int_{\Gamma_j} \int_{|s-\tilde{s}| \geq \delta} K_n(s, \tilde{s}) |f(\tilde{s}) - f(s)| d\tilde{s} ds = 0.$$

Now, if we can show that $K_n(s, \tilde{s}) |f(\tilde{s}) - f(s)| \in L_1(\Gamma_j \times \Gamma_j)$, then by continuity of Lebesgue integral, we establish the desired limit (5.23).

Using (3) and Fubini's theorem, we have

$$\int_{\Gamma_j \times \Gamma_j} |K_n(s, \tilde{s}) f(s)| d(s, \tilde{s}) = \int_{\Gamma_j} |f(s)| \left(\int_{\Gamma_j} |K_n(s, \tilde{s})| d\tilde{s} \right) ds \leq M \|f\|_{L_1(\Gamma_j)},$$

which means $K_n(s, \tilde{s}) |f(s)| \in L_1(\Gamma_j \times \Gamma_j)$. It remains to show that $K_n(s, \tilde{s}) |f(\tilde{s})| \in L_1(\Gamma_j \times \Gamma_j)$. By Fubini's theorem again, we have

$$\int_{\Gamma_j \times \Gamma_j} |K_n(s, \tilde{s}) f(\tilde{s})| d(s, \tilde{s}) = \int_{\Gamma_j} \left(\int_{\Gamma_j} |K_n(s, \tilde{s})| ds \right) |f(\tilde{s})| d\tilde{s}.$$

Consider

$$\begin{aligned} \int_{\Gamma_j} |K_n(s, \tilde{s})| ds &= \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \frac{2^{(1-n)} \sqrt{\gamma_{i,j}[\mathbf{k}]}}{\mathfrak{L}(I_{j,n,\mathbf{k}})} \varphi_{i,n-1,\mathbf{k}}(\tilde{s}) \int_{\Gamma_j} \chi_{I_{j,n,\mathbf{k}}}(s) ds \\ &= \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} 2^{(1-n)} \sqrt{\gamma_{i,j}[\mathbf{k}]} \varphi_{i,n-1,\mathbf{k}}(\tilde{s}) \\ &= \frac{1}{2} \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \sqrt{\gamma_{i,j}[\mathbf{k}]} \varphi_i(2^{n-1} x(\tilde{s}) - \mathbf{k}/2). \end{aligned}$$

Since $\gamma_{i,j} = O(1)$ and the support $\Lambda_{n-1,\mathbf{k}}$ overlaps finitely many times (independent of n) for all $\mathbf{k} \in \mathbb{M}^2$, we have

$$\left\| \int_{\Gamma_j} K_n(s, \cdot) ds \right\|_{L_\infty(\Gamma_j)} \leq C \|\varphi_{\mathbf{i}}\|_{L_\infty(\Omega)} < \infty.$$

Then Hölder's inequality implies $K_n(s, \tilde{s})|f(s)| \in L_1(\Gamma_j \times \Gamma_j)$. \square

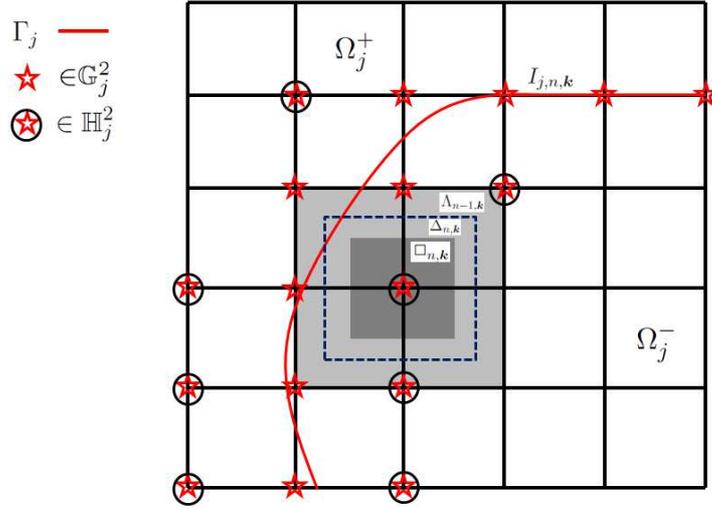


FIGURE 12. Illustrations for the proof of part (1) of Lemma 5.2 for the case of piecewise constant B-spline.

With Lemma 5.2, it is clear that $G_n^{(1)}(u) \rightarrow 0$. Now, we show that, with $\gamma_{i,j} = O(1)$, we have $G_n^{(2)}(u) \rightarrow 0$. Indeed,

$$\begin{aligned} G_n^{(2)}(u) &= \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \sum_{\mathbf{i} \in \mathbb{B}_1} 2^{1-n} \sqrt{\gamma_{i,j}[\mathbf{k}]} \left| \int_{\Omega_j^+ \cap \Lambda_{n-1,\mathbf{k}}} D_{\mathbf{i}} u \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_j^- \cap \Lambda_{n-1,\mathbf{k}}} D_{\mathbf{i}} u \varphi_{i,n-1,\mathbf{k}} \right| \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_j^2 \setminus \mathbb{H}_j^2} \sum_{\mathbf{i} \in \mathbb{B}_1} 2^{1-n} O(1) |o(1) + o(1)| = o(1), \end{aligned}$$

where we have used the fact that $|\mathbb{G}_j^2 \setminus \mathbb{H}_j^2| = O(2^n)$. Therefore, we have concluded the proof of $E_n^{(2)}(u) \rightarrow E^{(2)}(u)$.

5.3.3. Convergence of $E_n^{(3)} \rightarrow E^{(3)}$. Most part of the proof of $E_n^{(3)} \rightarrow E^{(3)}$ is analogous to that of $E_n^{(2)} \rightarrow E^{(2)}$. Note that it suffices to prove that, for each j and \tilde{j} and a properly chosen set of parameters $\{\tilde{\gamma}_{j,\tilde{j}}[\mathbf{k}] : \mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2\}$,

$$\left\| [\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_{j,\tilde{j}}} \right\|_1 \rightarrow \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|i|=1} |\mathfrak{T}_{j,\tilde{j}}^+(D_{\mathbf{i}} u) - \mathfrak{T}_{j,\tilde{j}}^-(D_{\mathbf{i}} u)|^2 \right)^{\frac{1}{2}} ds.$$

Letting

$$\tilde{v}_{i,j,\tilde{j}} := \mathfrak{T}_{j,\tilde{j}}^+(D_i u) - \mathfrak{T}_{j,\tilde{j}}^-(D_i u),$$

we have

$$\begin{aligned} & \left| \left\| \left[\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u \right]_{\Gamma_{j,\tilde{j}}} \right\|_1 - \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} |\tilde{v}_{i,j,\tilde{j}}|^2 \right)^{\frac{1}{2}} ds \right| \\ &= \left| \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2} \left(\sum_{\mathbf{i} \in \mathbb{B}_2} 2^{2s_{\mathbf{i}}(1-n)} \tilde{\gamma}_{i,j,\tilde{j}}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} |\tilde{v}_{i,j,\tilde{j}}|^2 \right)^{\frac{1}{2}} ds \right|. \end{aligned}$$

We now rewrite $\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle$ using Proposition 4.2. We first split $\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle$ into two parts as follows

$$\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle = \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} u(\mathbf{x}) (D_i \varphi_{i,n-1,\mathbf{k}})(\mathbf{x}) d\mathbf{x} + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} u(\mathbf{x}) (D_i \varphi_{i,n-1,\mathbf{k}})(\mathbf{x}) d\mathbf{x},$$

Now, we apply Proposition 4.2 and obtain

(5.26)

$$\begin{aligned} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle &= - \int_{\Gamma_{j,\tilde{j}} \cap \Lambda_{n-1,\mathbf{k}}} \tilde{v}_{(1,0)^\top, j,\tilde{j}} \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_{(0,1)^\top} ds && \text{(for } \mathbf{i} = (1,1)^\top \text{: version 1)} \\ &+ \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}}; \end{aligned}$$

(5.27)

$$\begin{aligned} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle &= - \int_{\Gamma_{j,\tilde{j}} \cap \Lambda_{n-1,\mathbf{k}}} \tilde{v}_{(0,1)^\top, j,\tilde{j}} \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_{(1,0)^\top} ds && \text{(for } \mathbf{i} = (1,1)^\top \text{: version 2)} \\ &+ \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}}; \end{aligned}$$

(5.28)

$$\begin{aligned} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle &= - \int_{\Gamma_{j,\tilde{j}} \cap \Lambda_{n-1,\mathbf{k}}} \tilde{v}_{(1,0)^\top, j,\tilde{j}} \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_{(1,0)^\top} ds && \text{(for } \mathbf{i} = (2,0)^\top) \\ &+ \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}}; \end{aligned}$$

and

(5.29)

$$\begin{aligned} \langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle &= - \int_{\Gamma_{j,\tilde{j}} \cap \Lambda_{n-1,\mathbf{k}}} \tilde{v}_{(0,1)^\top, j,\tilde{j}} \varphi_{i,n-1,\mathbf{k}} \mathbf{n}_{(0,1)^\top} ds && \text{(for } \mathbf{i} = (0,2)^\top) \\ &+ \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}} + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} D_i u \varphi_{i,n-1,\mathbf{k}}. \end{aligned}$$

Notice that

$$\int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} |\tilde{v}_{i,j,\tilde{j}}|^2 \right)^{\frac{1}{2}} ds = \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|\mathbf{i}|=1} \sum_{|\tilde{\mathbf{i}}|=1} |\tilde{v}_{i,j,\tilde{j}} \tilde{\mathbf{n}}_{\tilde{\mathbf{i}}}|^2 \right)^{\frac{1}{2}} ds.$$

Define the index set $\mathbb{H}_{j,\tilde{j}}^2 \subset \mathbb{G}_{j,\tilde{j}}^2$ as

$$\mathbb{H}_{j,\tilde{j}}^2 := \left\{ \mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 : \mathfrak{L}(\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}) > 0 \right\},$$

where $\square_{n,\mathbf{k}}$ is the rectangular domain $(\frac{k_1-1/2}{2^n}, \frac{k_1+1/2}{2^n}] \times (\frac{k_2-1/2}{2^n}, \frac{k_2+1/2}{2^n}]$. Now, letting $\gamma_{i,j,\tilde{j}}[\mathbf{k}] = 0$ for $\mathbf{k} \in \mathbb{H}_{j,\tilde{j}}^2$ and

$$w_{i,j,\tilde{j}}[\mathbf{k}] = \begin{cases} \frac{2^{2s_i(1-n)}}{\mathfrak{L}^2(\Gamma_{j,\tilde{j}} \cap \square_{n-1,\mathbf{k}})} \tilde{\gamma}_{i,j,\tilde{j}}[\mathbf{k}], & \mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2, \\ 0, & \mathbf{k} \in \mathbb{H}_{j,\tilde{j}}^2, \end{cases}$$

we have

$$\begin{aligned} & \left| \left\| [\tilde{\gamma}_{j,\tilde{j}} \cdot \mathbf{W}_n \mathbf{T}_n u]_{\Gamma_{j,\tilde{j}}} \right\|_1 - \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|i|=1} |\tilde{v}_{i,j,\tilde{j}}|^2 \right)^{\frac{1}{2}} ds \right| \\ &= \left| \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2} \left(\sum_{i \in \mathbb{B}_2} 2^{2s_i(1-n)} \tilde{\gamma}_{i,j,\tilde{j}}[\mathbf{k}] |\langle u, D_i \varphi_{i,n-1,\mathbf{k}} \rangle|^2 \right)^{\frac{1}{2}} - \int_{\Gamma_{j,\tilde{j}}} \left(\sum_{|i|=1} |\tilde{v}_{i,j,\tilde{j}}|^2 \right)^{\frac{1}{2}} ds \right| \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \int_{\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}} \left(\left| \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(1,1)^\top} \varphi_{(1,1)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(1,0)^\top, j,\tilde{j}} \mathbf{n}_{(0,1)^\top} \right| \right. \\ &\quad + \left| \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(1,1)^\top} \varphi_{(1,1)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(0,1)^\top, j,\tilde{j}} \mathbf{n}_{(1,0)^\top} \right| \\ &\quad + \left| \sqrt{w_{(2,0)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(2,0)^\top} \varphi_{(2,0)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(1,0)^\top, j,\tilde{j}} \mathbf{n}_{(1,0)^\top} \right| \\ &\quad \left. + \left| \sqrt{w_{(0,2)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(0,2)^\top} \varphi_{(0,2)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(0,1)^\top, j,\tilde{j}} \mathbf{n}_{(0,1)^\top} \right| \right) ds + o(1). \end{aligned}$$

We shall prove that each of the four terms above goes to 0 as $n \rightarrow \infty$ by using the integration formula (5.26)-(5.29) respectively for each term. Since the proof for each term is entirely analogous, we only focus on proving the convergence of the first term, i.e.

$$(5.30) \quad \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \int_{\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}} \left| \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(1,1)^\top} \varphi_{(1,1)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(1,0)^\top, j,\tilde{j}} \mathbf{n}_{(0,1)^\top} \right| ds \rightarrow 0$$

as $n \rightarrow \infty$. Using the integration by parts formula (5.26), we have

$$\begin{aligned} & \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \int_{\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}} \left| \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \langle u, D_{(1,1)^\top} \varphi_{(1,1)^\top, n-1,\mathbf{k}} \rangle - \tilde{v}_{(1,0)^\top, j,\tilde{j}} \mathbf{n}_{(0,1)^\top} \right| ds \\ &\leq \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \int_{\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}} \left| \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \int_{\Gamma_{j,\tilde{j}} \cap \Lambda_{n-1,\mathbf{k}}} \tilde{v}_{(1,0)^\top, j,\tilde{j}} \varphi_{(1,1)^\top, n-1,\mathbf{k}} \mathbf{n}_{(0,1)^\top} d\tilde{s} - \tilde{v}_{(1,0)^\top, j,\tilde{j}} \mathbf{n}_{(0,1)^\top} \right| ds \\ &\quad + \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \int_{\Gamma_{j,\tilde{j}} \cap \square_{n,\mathbf{k}}} \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j,\tilde{j}}[\mathbf{k}]} \left| \int_{\Omega_{j,\tilde{j}}^+ \cap \Lambda_{n-1,\mathbf{k}}} D_{(1,1)^\top} u \varphi_{(1,1)^\top, n-1,\mathbf{k}} \right. \\ &\quad \left. + \int_{\Omega_{j,\tilde{j}}^- \cap \Lambda_{n-1,\mathbf{k}}} D_{(1,1)^\top} u \varphi_{(1,1)^\top, n-1,\mathbf{k}} \right| ds \\ &=: \tilde{G}_n^{(1)}(u) + \tilde{G}_n^{(2)}(u). \end{aligned}$$

Same as in the proof of convergence $E_n^{(2)}(u) \rightarrow E^{(2)}(u)$, $\tilde{G}_n^{(1)}$ can be written as

$$\tilde{G}_n^{(1)}(u) = \int_{\Gamma_{j,\tilde{j}}} \left| \int_{\Gamma_{j,\tilde{j}}} v_{(1,0)^\top, j, \tilde{j}}(\tilde{s}) \mathbf{n}_{(0,1)^\top}(\tilde{s}) \tilde{K}_n(s, \tilde{s}) d\tilde{s} - v_{(1,0)^\top, j, \tilde{j}}(\tilde{s}) \mathbf{n}_{(0,1)^\top}(s) \right| ds + o(1),$$

where

$$\begin{aligned} \tilde{K}_n(s, \tilde{s}) &:= \frac{1}{\sqrt{2}} \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \sqrt{w_{(1,1)^\top, j, \tilde{j}}[\mathbf{k}]} \varphi_{(1,1)^\top, n-1, \mathbf{k}}(\tilde{s}) \chi_{\Gamma_{j,\tilde{j}} \cap \square_{n, \mathbf{k}}}(s) \\ (5.31) \quad &= \frac{1}{\sqrt{2}} \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \frac{2^{(2-2n)} \sqrt{\tilde{\gamma}_{(1,1)^\top, j, \tilde{j}}[\mathbf{k}]}}{\mathfrak{L}(\Gamma_{j,\tilde{j}} \cap \square_{n, \mathbf{k}})} \varphi_{(1,1)^\top, n-1, \mathbf{k}}(\tilde{s}) \chi_{\Gamma_{j,\tilde{j}} \cap \square_{n, \mathbf{k}}}(s). \end{aligned}$$

To show that $\tilde{G}_n^{(1)} \rightarrow 0$, we need the following lemma, which is a similar one as Lemma 5.2.

Lemma 5.3. *Let $\tilde{K}_n(s, \tilde{s})$ be defined in (5.31) with $\mathbf{i} = (1, 1)^\top$ and $\varphi_{\mathbf{i}}$ being nonnegative and continuous.*

(1) *There exists $0 \leq \tilde{\gamma}_{\mathbf{i}, j, \tilde{j}} \leq O(2^{2n})$ such that, as $n \rightarrow \infty$,*

$$\int_{\Gamma_{j,\tilde{j}}} \tilde{K}_n(s, \tilde{s}) d\tilde{s} \rightarrow 1 \quad \text{for almost all } s \in \Gamma_{j,\tilde{j}}.$$

(2) *For any $\delta > 0$ and an n large enough, we have*

$$\tilde{K}_n(s, \tilde{s}) = 0$$

for $(s, \tilde{s}) \in \{(s, \tilde{s}) \in \Gamma_{j,\tilde{j}} \times \Gamma_{j,\tilde{j}}, |s - \tilde{s}| \geq \delta\}$.

(3) *We have $\int_{\Gamma_{j,\tilde{j}}} |\tilde{K}_n(s, \tilde{s})| d\tilde{s} \leq M$ uniformly for all $s \in \Gamma_{j,\tilde{j}}$*

Consequently, for any $f \in L_1(\Gamma_{j,\tilde{j}})$, we have

$$(5.32) \quad \lim_{n \rightarrow \infty} \left\| \int_{\Gamma_{j,\tilde{j}}} \tilde{K}_n(\cdot, \tilde{s}) f(\tilde{s}) d\tilde{s} - f(\cdot) \right\|_{L_1(\Gamma_{j,\tilde{j}})} = 0.$$

Proof of Lemma 5.3. The proof of (2), (3) and (5.32) is the same as that of Lemma 5.2, whence we establish (1). In fact, the proof for (1) is also similar to that of Lemma 5.2. However, we shall still provide more detail on the proof of (1).

We first note that $\varphi_{\mathbf{i}, n-1, \mathbf{k}}$ is nonnegative and continuous on the interior of its support $\Lambda_{n-1, \mathbf{k}}$. Let $I_{j,\tilde{j}, n, \mathbf{k}} := \Gamma_{j,\tilde{j}} \cap \square_{n, \mathbf{k}}$. We have

$$\int_{\Gamma_{j,\tilde{j}}} \tilde{K}_n(s, \tilde{s}) d\tilde{s} = \sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \left[\frac{2^{(3/2-2n)} \sqrt{\tilde{\gamma}_{\mathbf{i}, j, \tilde{j}}[\mathbf{k}]}}{\mathfrak{L}(I_{j,\tilde{j}, n, \mathbf{k}})} \int_{\Gamma_{j,\tilde{j}}} \varphi_{\mathbf{i}, n-1, \mathbf{k}}(\tilde{s}) d\tilde{s} \right] \chi_{I_{j,\tilde{j}, n, \mathbf{k}}}(s).$$

Then, we choose

$$(5.33) \quad \sqrt{\tilde{\gamma}_{\mathbf{i}, j, \tilde{j}}[\mathbf{k}]} = \begin{cases} 2^{2n-3/2} \mathfrak{L}(I_{j,\tilde{j}, n-1, \mathbf{k}}) \left(\int_{\Gamma_{j,\tilde{j}}} \varphi_{\mathbf{i}, n-1, \mathbf{k}}(\tilde{s}) d\tilde{s} \right)^{-1} & \text{for } \mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2 \\ 0 & \text{otherwise,} \end{cases}$$

which leads to $\int_{\Gamma_{j,\tilde{j}}} \tilde{K}_n(s, \tilde{s}) d\tilde{s} \rightarrow 1$ a.e. because

$$\sum_{\mathbf{k} \in \mathbb{G}_{j,\tilde{j}}^2 \setminus \mathbb{H}_{j,\tilde{j}}^2} \chi_{I_{j,\tilde{j}, n, \mathbf{k}}}(s) \rightarrow 1 \quad \text{for almost all } s \in \Gamma_{j,\tilde{j}}.$$

Also, similar as in Lemma 5.2, we can show that $0 \leq \tilde{\gamma}_{\mathbf{i}, j, \tilde{j}} \leq O(2^{2n})$. This concludes the proof of Lemma 5.3. \square

With Lemma 5.3, it is clear that $\tilde{G}_n^{(1)}(u) \rightarrow 0$. Now, we show that, with $\tilde{\gamma}_{(1,1)^\top, j, \tilde{j}} = O(2^{2n})$, we have $\tilde{G}_n^{(2)}(u) \rightarrow 0$. Indeed,

$$\begin{aligned}
\tilde{G}_n^{(2)}(u) &= \sum_{\mathbf{k} \in \mathbb{G}_{j, \tilde{j}}^2 \setminus \mathbb{H}_{j, \tilde{j}}^2} \int_{\Gamma_{j, \tilde{j}} \cap \Lambda_{n-1, \mathbf{k}}} \frac{1}{\sqrt{2}} \sqrt{w_{(1,1)^\top, j, \tilde{j}}[\mathbf{k}]} \left| \int_{\Omega_{j, \tilde{j}}^+ \cap \Lambda_{n-1, \mathbf{k}}} D_{(1,1)^\top} u, \varphi_{(1,1)^\top, n-1, \mathbf{k}} \right. \\
&\quad \left. + \int_{\Omega_{j, \tilde{j}}^- \cap \Lambda_{n-1, \mathbf{k}}} D_{(1,1)^\top} u, \varphi_{(1,1)^\top, n-1, \mathbf{k}} \right| ds \\
&= \sum_{\mathbf{k} \in \mathbb{G}_{j, \tilde{j}}^2 \setminus \mathbb{H}_{j, \tilde{j}}^2} 2^{3/2-2n} \sqrt{\tilde{\gamma}_{(1,1)^\top, j, \tilde{j}}[\mathbf{k}]} \left| \int_{\Omega_{j, \tilde{j}}^+ \cap \Lambda_{n-1, \mathbf{k}}} D_{(1,1)^\top} u, \varphi_{(1,1)^\top, n-1, \mathbf{k}} \right. \\
&\quad \left. + \int_{\Omega_{j, \tilde{j}}^- \cap \Lambda_{n-1, \mathbf{k}}} D_{(1,1)^\top} u, \varphi_{(1,1)^\top, n-1, \mathbf{k}} \right| ds \\
&\leq \sum_{\mathbf{k} \in \mathbb{G}_{j, \tilde{j}}^2 \setminus \mathbb{H}_{j, \tilde{j}}^2} 2^{3/2-2n} O(2^n) |o(1) + o(1)| = o(1).
\end{aligned}$$

This concludes the proof of $E_n^{(3)}(u) \rightarrow E^{(3)}(u)$, hence the proof of Theorem 5.1.

REFERENCES

- [1] 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.
- [2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
- [3] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [4] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*. Springer, 2003.
- [5] G. Sapiro, *Geometric partial differential equations and image analysis*. Cambridge University Press, 2001.
- [6] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*. Springer, 2006.
- [7] T. Chan and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. Society for Industrial Mathematics, 2005.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] M. Fadili and J. Starck, "Sparse representations and Bayesian image inpainting," *Proc. SPARS*, vol. 5, 2005.
- [13] 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.
- [14] 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.
- [15] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*. American Mathematical Society, 2001.
- [16] R. Chan, S. Riemenschneider, L. Shen, and Z. Shen, "Tight frame: an efficient way for high-resolution image reconstruction," *Applied and Computational Harmonic Analysis*, vol. 17, no. 1, pp. 91–115, 2004.
- [17] 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.
- [18] J. Cai and Z. Shen, "Framelet based deconvolution," *J. Comp. Math*, vol. 28, no. 3, pp. 289–308, 2010.

- [19] M. Fadili, J. Starck, and F. Murtagh, “Inpainting and zooming using sparse representations,” *The Computer Journal*, vol. 52, no. 1, p. 64, 2009.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] B. Dong, Q. Jiang, and Z. Shen, “Image restoration: Wavelet frame shrinkage, nonlinear evolution pdes, and beyond,” *UCLA CAM Report*, vol. 13-78, 2013.
- [24] A. Chambolle, “Finite-differences discretizations of the mumford-shah functional,” *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 33, no. 02, pp. 261–288, 1999.
- [25] 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.
- [26] I. Daubechies, *Ten lectures on wavelets*, vol. CBMS-NSF Lecture Notes, SIAM, nr. 61. Society for Industrial and Applied Mathematics, 1992.
- [27] 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.
- [28] Z. Shen, “Wavelet frames and image restorations,” in *Proceedings of the International Congress of Mathematicians*, vol. 4, pp. 2834–2863, 2010.
- [29] A. Chai and Z. Shen, “Deconvolution: A wavelet frame approach,” *Numerische Mathematik*, vol. 106, no. 4, pp. 529–587, 2007.
- [30] J. Cai, R. Chan, L. Shen, and Z. Shen, “Restoration of chopped and noded images by framelets,” *SIAM J. Sci. Comput*, vol. 30, no. 3, pp. 1205–1227, 2008.
- [31] 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.
- [32] J. Cai, R. Chan, and Z. Shen, “A framelet-based image inpainting algorithm,” *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 131–149, 2008.
- [33] R. Chan, Z. Shen, and T. Xia, “A framelet algorithm for enhancing video stills,” *Applied and Computational Harmonic Analysis*, vol. 23, no. 2, pp. 153–170, 2007.
- [34] C. De Boor, *A Practical Guide to Splines*. Springer Verlag, 2001.
- [35] Y. Zhang, B. Dong, and Z. Lu, “ ℓ_0 minimization of wavelet frame based image restoration,” *Mathematics of Computation*, vol. 82, pp. 995–1015, 2013.
- [36] B. Dong and Y. Zhang, “An efficient algorithm for ℓ_0 minimization in wavelet frame based image restoration,” *Journal of Scientific Computing*, vol. 54 (2-3), pp. 350–368, 2013.
- [37] J.-F. Cai, H. Ji, Z. Shen, and G.-B. Ye, “Data-driven tight frame construction and image denoising,” *Applied and Computational Harmonic Analysis*, 2013, to appear.
- [38] 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.
- [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] 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.
- [41] J. Eckstein and D. Bertsekas, “On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.
- [42] X. Jia, B. Dong, Y. Lou, and S. Jiang, “GPU-based iterative cone-beam CT reconstruction using tight frame regularization,” *Physics in Medicine and Biology*, vol. 56, pp. 3787–3807, 2011.
- [43] B. Dong, J. Li, and Z. Shen, “X-ray CT image reconstruction via wavelet frame based regularization and Radon domain inpainting,” *Journal of Scientific Computing*, vol. 54 (2-3), pp. 333–349, 2013.
- [44] G. Verchota, “Layer potentials and regularity for the dirichlet problem for laplace’s equation in lipschitz domains,” *Journal of Functional Analysis*, vol. 59, no. 3, pp. 572–611, 1984.
- [45] E. Gagliardo, “Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili,” *Rendiconti del seminario matematico della università di Padova*, vol. 27, pp. 284–305, 1957.
- [46] Z. Ding, “A proof of the trace theorem of sobolev spaces on Lipschitz domains,” *Proceedings of the American Mathematical Society*, vol. 124, no. 2, pp. 591–600, 1996.
- [47] A. Chambolle, “Image segmentation by variational methods: Mumford and shah functional and the discrete approximations,” *SIAM Journal on Applied Mathematics*, vol. 55, no. 3, pp. 827–863, 1995.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF IOWA, IOWA CITY, IA, 52242-1419.

E-mail address: jianfeng-cai@uiowa.edu

BEIJING INTERNATIONAL CENTER FOR MATHEMATICAL RESEARCH, PEKING UNIVERSITY, BEIJING, CHINA, 100871.

E-mail address: dongbin@math.math.edu.cn

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, BLOCK S17, 10 LOWER KENT RIDGE ROAD, SINGAPORE, 119076.

E-mail address: matzuows@nus.edu.sg