SIAM J. NUMER. ANAL.
Vol. 49, No. 2, pp. 845–868

ERROR BOUNDS FOR FINITE-DIFFERENCE METHODS FOR RUDIN–OSHER–FATEMI IMAGE SMOOTHING*

JINGYUE WANG† AND BRADLEY J. LUCIER‡

Abstract. We bound the difference between the solution to the continuous Rudin–Osher–Fatemi (ROF) image smoothing model and the solutions to various finite-difference approximations to this model. These bounds apply to "typical" images, i.e., images with edges or with fractal structure. These are the first bounds on the error in numerical methods for ROF smoothing.

Key words. total variation, variational methods, finite-difference methods, error bounds

AMS subject classifications. 65N06, 65N12, 94A08

DOI. 10.1137/090769594

1. Introduction.

Image noise removal based on total variation smoothing was introduced by Rudin, Osher, and Fatemi in [13]. Under this Rudin–Osher–Fatemi (ROF) model, one supposes a "true" image f defined on $\Omega = [0,1]^2$ and a "corrupted" image g derived from f (by adding noise, etc.) with $\|f-g\|^2_{L^2(\Omega)} = \sigma^2$. In an attempt to reconstruct f from g, one calculates a "smoothed" image u that minimizes

$$|v|_{\mathrm{BV}(\Omega)} = \int_\Omega |Dv| \text{ subject to the constraint } \|v-g\|^2_{L^2(\Omega)} \le \sigma^2. \tag{1.1}$$

(Precise definitions are given later.) We deal with the equivalent problem: If we calculate $\bar g$, the average of g on Ω, then for any σ with

$$\sigma^2 < \|g-\bar g\|^2_{L^2(\Omega)}$$

there exists a unique $\lambda > 0$ such that the minimizer of (1.1) is the minimizer u of the functional

$$E(v) = \frac{1}{2\lambda}\|v-g\|^2_{L^2(\Omega)} + |v|_{\mathrm{BV}(\Omega)}. \tag{1.2}$$

Here λ is a positive parameter that balances the relative importance of the smoothness of the minimizer (important when λ is large) and the $L^2(\Omega)$ distance between the minimizer and the initial data (important when λ is small). About the same time, Bouman and Sauer [1] proposed a discrete version of (1.2) in the context of tomography.

Practically, one discretizes $E(\cdot)$ to compute the minimizer of the discrete functional $E_h(\cdot)$. We assume the discrete corrupted image g^h of resolution $N \times N$ ($N = 1/h$) is simply the piecewise constant projection of the continuous corrupted image g, and define the discrete functional

$$E_h(v^h) = \frac{1}{2\lambda}\sum_i |v^h_i - g^h_i|^2\, h^2 + J_h(v^h), \tag{1.3}$$

*Received by the editors September 1, 2009; accepted for publication (in revised form) January 24, 2011; published electronically April 27, 2011. This work was partially supported by the Office of Naval Research, contract N00014-91-J-1152, and the Institute for Mathematics and its Applications, The University of Minnesota.
http://www.siam.org/journals/sinum/49-2/76959.html
†Department of Mathematics, University of Georgia, Athens, GA 30602 (jwang@math.uga.edu).
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907-2067 (lucier@math.purdue.edu).

where J_h is a discretized total variation. The most commonly used J_h is the discrete variation J_{++} used in [13],

$$J_{++}(v^h) = \sum_i \sqrt{\left(\frac{v^h_{i+(1,0)} - v^h_i}{h}\right)^2 + \left(\frac{v^h_{i+(0,1)} - v^h_i}{h}\right)^2}\, h^2. \tag{1.4}$$

Efficient algorithms have been developed to compute the discrete minimizer [2], [6], [3], [5].

In this paper, we study the relationship between the minimizer u of $E(\cdot)$ and the discrete minimizer u^h of $E_h(\cdot)$. It is well known that E_h Γ-converges to E in L^1. As a direct consequence u^h tends to u in L^1. Assuming the discrete variation J_h satisfies certain conditions that we explain later, we give a bound of the L^2 norm of the difference between u and u^h in Theorem 4.2.

Because the ROF model is often applied to images, an analysis of the error between the solutions of discrete approximations and the solution of the continuous model itself should apply to functions modeling images. "Typical" natural images have little smoothness, because of intensity discontinuities at the edges of objects and the fractal structure of many objects themselves (leaves in a tree, hair, etc.). Our results apply to functions in the Lipschitz spaces $\mathrm{Lip}(\alpha, L^2(\Omega))$, which contain functions with, roughly speaking, α "derivatives" in $L^2(\Omega)$. Here $0 < \alpha \leq 1/2$ for "images with edges": $f \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$ implies $f \in \mathrm{Lip}(1/2, L^2(\Omega))$, while functions with fractal structure usually have $\alpha < 1/2$; see [7].

Our convergence results in section 4 are proved for (1.3) with $J_h = J_*$, a discrete variation obtained by symmetrizing J_{++}. Nonetheless, our approach is quite general, and in section 5 we obtain the same results for J_U, an "upwind" discrete variation formulated in [12]. We remark that an iterative method for minimizing (1.3) with $J_h = J_U$ was given in [5].

While the ROF model has proved to be tremendously influential, and has been the basis of further algorithms in image processing, we know of no other results that bound the difference between the solution of the continuous problem and its finite-difference approximations. A finite element method applied to the time-dependent gradient descent problem associated with (1.2) was studied in [10]; we note that their Theorem 4 requires the initial data u_0 to have two continuous derivatives on $\overline{\Omega}$ so it does not apply to "typical" natural images with edges.

The rest of this section introduces notation and our main results. In section 2 we compare discrete and continuous variational functionals. In section 3 we note some properties of the minimizers of $E(\cdot)$ and $E_h(\cdot)$ that we use in section 4 to first bound the difference between the discrete and continuous functionals at their respective minimizers and then to bound the L^2 difference between the discrete and continuous minimizers themselves. In section 5, we prove a number of lemmas for the "upwind" discretization of the bounded variation (BV) semi-norm that allow us to prove similar error bounds for the discrete minimizer of the "upwind" scheme. Section 6 summarizes our results and points to variations that appear elsewhere.

Finally, we note that we present here a sequence of lemmas, the proofs of which are often omitted as "a tedious calculation," or "standard" given the previous lemmas. We found when developing these results, however, that dividing the argument into these smaller steps led to much greater clarity, and we preserve that structure in this paper.

1.1. Summary of main results. The major difficulty to overcome in our analysis is to compare $J_h(u^h)$, the discrete variation of the minimizer of the discrete functional (1.3), with $|u|_{\mathrm{BV}(\Omega)}$, the variation of the minimizer of the continuous functional (1.2). Indeed, if we consider $P_h u_i$, the discrete function that is computed as the average of u on subsquares $\Omega_i = h(i+\Omega)$, then, for general u, no matter which J_h we choose, we have that

$$\lim_{h\to 0} J_h(P_h u) \neq |u|_{\mathrm{BV}(\Omega)}.$$

So $E_h(P_h u)$ does not, in general, converge to $E(u)$ (since $\|P_h u - P_h g\|_{L^2(\Omega^h)} \to \|u-g\|_{L^2(\Omega)}$ as $h\to 0$), and the question of whether $E_h(u^h)$ *does* converge to $E(u)$ requires a more subtle analysis.

We note that J_{++} defined by (1.4) is a consistent approximation to $|u|_{\mathrm{BV}(\Omega)}$ for continuously differentiable u: $\lim_{h\to 0} J_{++}(P_h u) = |u|_{\mathrm{BV}(\Omega)}$ for smooth u. So, for general $u \in \mathrm{BV}(\Omega)$, we first mollify u, computing $\mathcal{S}_\epsilon u = \eta_\epsilon * u$, where η_ϵ is a mollifier and ϵ is a positive parameter tending to zero in a controlled way that depends on h, and compare $E_h(P_h\mathcal{S}_\epsilon u)$ to $E(u)$. Mollifying u introduces an error in the $L^2(\Omega)$ term of (1.2), but it reduces the J_h term, making it closer to $|u|_{\mathrm{BV}(\Omega)}$.

That is how we compare the continuous u to the discrete u^h; we also have to go the other way. To do that we require J_h to have a certain symmetry, so we consider first a symmetrized version J_* of J_{++} and then later an upwind discrete variation J_U. In more-or-less complete analogy with the continuous argument, we first compute $S_L u_i^h$, a discrete average of u_i^h on $(2L+1)\times(2L+1)$ squares (with the point i in the middle of the square), and then compute a piecewise linear interpolant $\operatorname{Int} S_L u^h$ (1.19) of $S_L u_i^h$, comparing $E(\operatorname{Int} S_L u^h)$ to $E_h(u^h)$. Now L is a positive parameter, depending on h, that tends to infinity as $h\to 0$.

We prove the following two theorems.

THEOREM 1.1 (functional difference). *Let $g \in \mathrm{Lip}(\alpha, L^2(\Omega))$ and assume u is the minimizer of $E(v)$ from* (3.3) *and u^h is the minimizer of $E_h(v^h)$ from* (3.1).

Then there is a constant C such that if $\epsilon = h^{1/(\alpha+1)}$, we have

$$E_h(P_h\mathcal{S}_\epsilon u) \le E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)},$$

and if L is set to the integer part of $h^{-\alpha/(\alpha+1)}$, then

$$E(\operatorname{Int} S_L u^h) \le E_h(u^h) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}.$$

Finally,

$$|E(u) - E_h(u^h)| \le \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}.$$

THEOREM 1.2 (minimizer difference). *Let $g \in \mathrm{Lip}(\alpha, L^2(\Omega))$ and assume that u is the minimizer of $E(v)$ from* (3.3) *and u^h is the minimizer of $E_h(v^h)$ from* (3.1). *Then there is a constant C such that*

$$\|I_h u^h - u\|^2_{L^2(\Omega)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)},$$

where $(I_h u^h)(x) = u_i^h$ for $x \in \Omega_i$.

1.2. Basic notations. We consider the usual $L^p(\Omega)$ spaces on $\Omega := [0,1]^2 \subset \mathbb{R}^2$, with $\|v\|_{L^p(\Omega)} := \left(\int_\Omega |v|^p\right)^{\frac{1}{p}}$ for $1 \le p < \infty$. (We assume the usual change for $p = \infty$.) We consider the discrete set Ω^h to be the set of all pairs $i = (i_1, i_2) \in \mathbb{Z}^2$, $\mathbb{Z}$ the integers, with $0 \le i_1, i_2 < N$, $h = 1/N$, and we refer to functions defined on Ω^h as *discrete functions*. So for discrete functions $v^h = v_i^h$, we define the discrete $L^p(\Omega^h)$ norms

$$\|v^h\|_{L^p(\Omega^h)} := \left(\sum_{i\in\Omega^h} |v_i^h|^p\, h^2\right)^{\frac{1}{p}} \quad \text{for } 1 \le p < \infty.$$

We define the translation operator for discrete functions by

$$(T_\ell(v^h))_i := v_{i+\ell}^h \quad \text{for any } \ell = (\ell_1, \ell_2) \in \mathbb{Z}^2.$$

To measure the size of a translation, we introduce $|\ell| = \max(|\ell_1|, |\ell_2|)$. Similarly,

$$(\mathcal{T}_\tau v)(x) = v(x+\tau) \quad \text{for any } \tau = (\tau_1, \tau_2) \in \mathbb{R}^2$$

and, for translations, we set $|\tau| = \max(|\tau_1|, |\tau_2|)$.

We often need to extend $v \in L^p(\Omega)$ and $v^h \in L^p(\Omega^h)$ to all of $\mathbb{R}^2$ and $\mathbb{Z}^2$, respectively; we denote the extensions by $\operatorname{Ext} v$ and $\operatorname{Ext}_h v^h$. For $v \in L^p(\Omega)$, we use the following procedure. First, $\operatorname{Ext} v(x) = v(x)$, $x \in \Omega$. We then reflect across the line $x_1 = 1$,

$$\operatorname{Ext} v(x_1, x_2) = \operatorname{Ext} v(2 - x_1, x_2), \quad 1 \le x_1 \le 2,\ 0 \le x_2 \le 1,$$

and reflect again across the line $x_2 = 1$,

$$\operatorname{Ext} v(x_1, x_2) = \operatorname{Ext} v(x_1, 2 - x_2), \quad 0 \le x_1 \le 2,\ 1 \le x_2 \le 2.$$

Having defined $\operatorname{Ext} v$ on 2Ω, we then extend $\operatorname{Ext} v$ periodically on all of $\mathbb{R}^2$.

We use the analogous construction of $\operatorname{Ext}_h v^h$ for discrete functions v^h. Note that the value of $\operatorname{Ext}_h v^h$ at any point immediately "outside" Ω^h is the same as the value of v^h at the closest point "inside" Ω^h.

For $v \in L^p(\Omega)$ we define the (first-order) $L^p(\Omega)$ modulus of smoothness by

$$\omega(v, t)_{L^p(\Omega)} = \sup_{\tau\in\mathbb{R}^2,\ |\tau|<t} \left(\int_{x, x+\tau\in\Omega} |v(x+\tau) - v(x)|^p\, dx\right)^{\frac{1}{p}}.$$

We also define

$$\omega(\operatorname{Ext} v, t)_{L^p(2\Omega)} := \sup_{\tau\in\mathbb{R}^2,\ |\tau|<t} \|\mathcal{T}_\tau \operatorname{Ext} v - \operatorname{Ext} v\|_{L^p(2\Omega)}.$$

The Lipschitz spaces $\operatorname{Lip}(\alpha, L^p(\Omega))$ consist of all functions v for which

$$|v|_{\operatorname{Lip}(\alpha, L^p(\Omega))} := \sup_{t>0} t^{-\alpha} \omega(v, t)_{L^p(\Omega)} < \infty;$$

we set $\|v\|_{\operatorname{Lip}(\alpha, L^p(\Omega))} := \|v\|_{L^p(\Omega)} + |v|_{\operatorname{Lip}(\alpha, L^p(\Omega))}$.

We also need a discrete modulus of smoothness. The discrete $L^p(\Omega^h)$ modulus of smoothness is

$$\omega(v^h, m)_{L^p(\Omega^h)} := \sup_{\ell\in\mathbb{Z}^2,\ |\ell|\le m} \left(\sum_{i, i+\ell\in\Omega^h} |v_{i+\ell}^h - v_i^h|^p\, h^2\right)^{\frac{1}{p}}.$$

For $\operatorname{Ext}_h v^h$ we similarly define

$$\omega(\operatorname{Ext}_h v^h, m)_{L^p(2\Omega^h)} = \sup_{\ell\in\mathbb{Z}^2,\ |\ell|\le m} \|T_\ell \operatorname{Ext}_h v^h - \operatorname{Ext}_h v^h\|_{L^p(2\Omega^h)}.$$

We have the following relationship between moduli of smoothness and our extension operators; the lemma can be proved as in [8, page 182].

LEMMA 1.1 (Whitney extension). *For all $1 \le p \le \infty$ there exists a constant C such that for all $v \in L^p(\Omega)$ and $v^h \in L^p(\Omega^h)$*

$$\|\mathcal{T}_\tau \operatorname{Ext} v - \operatorname{Ext} v\|_{L^p(2\Omega)} \le C\omega(v, |\tau|)_{L^p(\Omega)}, \quad \tau \in \mathbb{R}^2, \tag{1.5}$$

and

$$\|T_\ell \operatorname{Ext}_h v^h - \operatorname{Ext}_h v^h\|_{L^p(2\Omega^h)} \le C\omega(v^h, |\ell|)_{L^p(\Omega^h)}, \quad \ell \in \mathbb{Z}^2. \tag{1.6}$$

Moreover, for all positive $t \in \mathbb{R}, m \in \mathbb{Z}$ we have

$$\omega(\operatorname{Ext} v, t)_{L^p(2\Omega)} \le C\omega(v, t)_{L^p(\Omega)} \tag{1.7}$$

and

$$\omega(\operatorname{Ext}_h v^h, m)_{L^p(2\Omega^h)} \le C\omega(v^h, m)_{L^p(\Omega^h)}. \tag{1.8}$$

1.3. Variation functionals. The variation of a function $v \in L^1(\Omega)$ is defined as follows. We consider functions ϕ in the space of C^1 functions from Ω to $\mathbb{R}^2$ with compact support, i.e., $[C_0^1(\Omega)]^2$. The variation of a function $v \in L^1(\Omega)$ is then defined to be

$$|v|_{\mathrm{BV}(\Omega)} := \int_\Omega |Dv| := \sup_{\phi\in[C_0^1(\Omega)]^2,\ |\phi|\le 1 \text{ pointwise}} \int_\Omega v\nabla\cdot\phi.$$

We note that if v is in the Sobolev space $W^{1,1}(\Omega)$, so that its first distributional derivatives are in $L^1(\Omega)$, then

$$|v|_{\mathrm{BV}(\Omega)} = \int_\Omega |\nabla v|.$$

We need discrete analogues of the variation of a function. For $\oplus$ and $\ominus$ independently taking values in the set $\{+,-\}$ and any discrete function v^h, we define

$$J_{\oplus\ominus}(v^h) := \sum_{i\in\Omega^h} \sqrt{\left(\frac{\operatorname{Ext}_h v^h_{i\oplus(1,0)} - \operatorname{Ext}_h v^h_i}{h}\right)^2 + \left(\frac{\operatorname{Ext}_h v^h_{i\ominus(0,1)} - \operatorname{Ext}_h v^h_i}{h}\right)^2}\, h^2. \tag{1.9}$$

We note that the sum is over $i \in \Omega^h$, and $\operatorname{Ext}_h v^h_i = v^h_i$ for all $i \in \Omega^h$.

Having defined $J_{++}(v^h)$, $J_{+-}(v^h)$, $J_{-+}(v^h)$, and $J_{--}(v^h)$, for any nonnegative a, b, c, and d with $a+b+c+d=1$, we define

$$J_h(v^h) = a\,J_{++}(v^h) + b\,J_{+-}(v^h) + c\,J_{-+}(v^h) + d\,J_{--}(v^h), \tag{1.10}$$

and define the special "isotropic" discrete variation

$$J_*(v^h) := \frac{1}{4}\big(J_{++}(v^h) + J_{+-}(v^h) + J_{-+}(v^h) + J_{--}(v^h)\big);$$

J_* is invariant under rotations of Ω^h by 90 degrees, or under horizontal or vertical reflections.

At times we consider discrete variational functionals for discrete functions defined on $2\Omega^h$; for these purposes we denote by $J_{\oplus\ominus}^{\Omega^h}(v^h)$ the discrete variation defined in (1.9) and by $J_{\oplus\ominus}^{2\Omega^h}(\operatorname{Ext}_h v^h)$ the corresponding sum over $2\Omega^h$; similarly we write

$$J_h^{\Omega^h}(v^h) = J_h(v^h) \quad \text{and} \quad J_*^{\Omega^h}(v^h) = J_*(v^h),$$

and we use the notation $J_h^{2\Omega^h}(\operatorname{Ext}_h v^h)$ for

$$a\, J_{++}^{2\Omega^h}(\operatorname{Ext}_h v^h) + b\, J_{+-}^{2\Omega^h}(\operatorname{Ext}_h v^h) + c\, J_{-+}^{2\Omega^h}(\operatorname{Ext}_h v^h) + d\, J_{--}^{2\Omega^h}(\operatorname{Ext}_h v^h)$$

and $J_*^{2\Omega^h}(\operatorname{Ext}_h v^h)$ for the corresponding sum with $a = b = c = d = 1/4$.

We have the following relationships between continuous and discrete variations of functions and continuous and discrete extension operators; the lemma is proved simply by considering the symmetries of $J_*^{\Omega_h}$ and noticing that $\operatorname{Ext} v$ does not add any variation along the lines of reflection.

LEMMA 1.2 (TV symmetry). *For any discrete function v^h,*

$$J_{\oplus\ominus}^{2\Omega^h}(\operatorname{Ext}_h v^h) = 4 J_*^{\Omega^h}(v^h). \tag{1.11}$$

Thus, we have $J_*^{2\Omega^h}(\operatorname{Ext}_h v^h) = 4J_*^{\Omega^h}(v^h)$ and, for any $\ell \in \mathbb{Z}^2$, $J_*^{2\Omega^h}(T_\ell \operatorname{Ext}_h v^h) = 4J_*^{\Omega^h}(v^h)$.

Similarly, for any $v \in \mathrm{BV}(\Omega)$, we have $|\operatorname{Ext} v|_{\mathrm{BV}(2\Omega)} = 4|v|_{\mathrm{BV}(\Omega)}$.

We also define a discrete "anisotropic" variation that is analogous to the $W^{1,1}(\Omega)$ Sobolev semi-norm:

$$|v^h|_{W^{1,1}(\Omega^h)} = \sum_{i\in\Omega^h} \left\{ \left| \frac{\operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i}{h} \right| + \left| \frac{\operatorname{Ext}_h v^h_{i+(0,1)} - \operatorname{Ext}_h v^h_i}{h} \right| \right\} h^2. \tag{1.12}$$

Because $\sqrt{a^2+b^2} \le |a| + |b| \le \sqrt{2}\sqrt{a^2+b^2}$, there exist positive constants C_1 and C_2 such that for any discrete function v^h and any discrete functional J_h,

$$C_1 |v^h|_{W^{1,1}(\Omega^h)} \le J_h(v^h) \le C_2 |v^h|_{W^{1,1}(\Omega^h)}. \tag{1.13}$$

For some intermediate estimates we need second-order continuous and discrete semi-norms, so we define for v in the Sobolev space $W^{2,1}(2\Omega)$ with periodic boundary conditions (i.e., treating 2Ω as a torus)

$$|v|_{W^{2,1}(2\Omega)} = \int_{2\Omega} |D_1^2 v| + |D_2^2 v|,$$

and for periodic discrete functions v^h on $2\Omega^h$ we define $|v^h|_{W_h^{2,1}(2\Omega^h)}$ as

$$\sum_{i\in 2\Omega^h} \left\{ \left| \frac{v^h_{i+(1,0)} - 2v^h_i + v^h_{i-(1,0)}}{h^2} \right| + \left| \frac{v^h_{i+(0,1)} - 2v^h_i + v^h_{i-(0,1)}}{h^2} \right| \right\} h^2. \tag{1.14}$$

Note that these semi-norms do not include "cross" derivatives or differences, but we do not need these in our estimates.

LEMMA 1.3 (TV difference). *For any two discrete functionals $J_{\oplus\ominus}$ and $J_{\oplus'\ominus'}$, and any discrete function v^h, we have*

$$|J_{\oplus\ominus}(v^h) - J_{\oplus'\ominus'}(v^h)| \le h|\operatorname{Ext}_h v^h|_{W_h^{2,1}(2\Omega^h)}. \tag{1.15}$$

Proof. The quantities summed in (1.9) are the norms of two-vectors of divided differences, which we choose to write in the following way:

$$\begin{aligned}
&|J_{++}(v^h) - J_{+-}(v^h)| \\
&= \left| \sum_{i\in\Omega^h} \left| \frac{1}{h} \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ \operatorname{Ext}_h v^h_{i+(0,1)} - \operatorname{Ext}_h v^h_i \end{pmatrix} \right| h^2 \right. \\
&\qquad \left. - \sum_{i\in\Omega^h} \left| \frac{1}{h} \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ (-1)(\operatorname{Ext}_h v^h_{i-(0,1)} - \operatorname{Ext}_h v^h_i) \end{pmatrix} \right| h^2 \right| \\
&\le \frac{1}{h} \sum_{i\in\Omega^h} \left| \left| \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ \operatorname{Ext}_h v^h_{i+(0,1)} - \operatorname{Ext}_h v^h_i \end{pmatrix} \right| - \left| \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ (-1)(\operatorname{Ext}_h v^h_{i-(0,1)} - \operatorname{Ext}_h v^h_i) \end{pmatrix} \right| \right| h^2 \\
&\le \frac{1}{h} \sum_{i\in\Omega^h} \left| \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ \operatorname{Ext}_h v^h_{i+(0,1)} - \operatorname{Ext}_h v^h_i \end{pmatrix} - \begin{pmatrix} \operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i \\ (-1)(\operatorname{Ext}_h v^h_{i-(0,1)} - \operatorname{Ext}_h v^h_i) \end{pmatrix} \right| h^2 \\
&= \frac{1}{h} \sum_{i\in\Omega^h} \left| \begin{pmatrix} 0 \\ \operatorname{Ext}_h v^h_{i+(0,1)} - 2\operatorname{Ext}_h v^h_i + \operatorname{Ext}_h v^h_{i-(0,1)} \end{pmatrix} \right| h^2 \\
&\le h|\operatorname{Ext}_h v^h|_{W_h^{2,1}(2\Omega^h)}.
\end{aligned}$$

Analogous arguments apply to other differences. $\square$

1.4. Projectors, injectors, and smoothing operators. We define the piecewise constant injector of discrete functions v^h into $L^p(\Omega)$: $(I_h v^h)(x) = v^h_i$ for $x \in \Omega_i$, where $\Omega_i := h(\Omega + i)$. Later we define an injector from discrete functions into a space of continuous, piecewise linear functions.

We also consider the piecewise constant projector of $v \in L^1(\Omega)$ onto the space of discrete functions, defined by

$$(P_h v)_i = \frac{1}{|\Omega_i|} \int_{\Omega_i} v, \quad i \in \Omega^h,$$

where $|\Omega_i|$ is the measure of Ω_i.

LEMMA 1.4 (injector and projector). *There exists a constant C such that for all $v \in L^2(\Omega)$*

$$\|v - I_h P_h v\|_{L^2(\Omega)} \le C\omega(v,h)_{L^2(\Omega)}. \tag{1.16}$$

We also have for any periodic $v \in W^{2,1}(2\Omega)$

$$|P_h v|_{W_h^{2,1}(2\Omega^h)} \le |v|_{W^{2,1}(2\Omega)}\,. \tag{1.17}$$

Proof. Relationship (1.16) is a special case of a general bound for the error in spline approximation; see [8, Theorem 7.3, page 225].

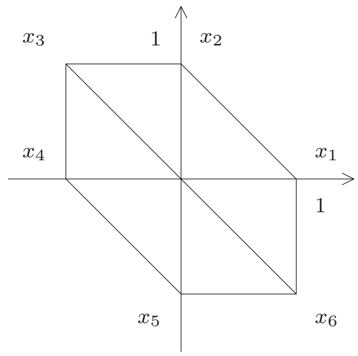

FIG. 1.1. *D, the support of* ϕ.

To prove (1.17), we deal with the differences in the horizontal direction:

$$\begin{aligned}
&\sum_{i\in 2\Omega^h}\left|\frac{(P_hv)_{i+(1,0)}-2(P_hv)_i+(P_h)v_{i-(1,0)}}{h^2}\right|h^2\\
&=\sum_{i\in 2\Omega^h}\frac{1}{h}\left|\frac{(P_hv)_{i+(1,0)}-(P_hv)_i}{h}-\frac{(P_hv)_i-(P_hv)_{i-(1,0)}}{h}\right|h^2\\
&=\sum_{i\in 2\Omega^h}\frac{1}{h^2}\left|\int_{\Omega_i}[(v(x+h,y)-v(x,y))-(v(x,y)-v(x-h,y))]\,dx\,dy\right|\\
&=\sum_{i\in 2\Omega^h}\frac{1}{h^2}\left|\int_{\Omega_i}\int_0^h[D_1v(x+t,y)-D_1v(x+t-h,y)]\,dt\,dx\,dy\right|\\
&=\sum_{i\in 2\Omega}\frac{1}{h^2}\left|\int_{\Omega_i}\int_0^h\int_{-h}^0 D_{11}v(x+t+s,y)\,ds\,dt\,dx\,dy\right|\\
&\le\int_{2\Omega}|D_{11}v|\,dx\,dy\quad\text{(exchange the order of integration and sum over } i\text{).}
\end{aligned}$$

Arguing similarly in the vertical direction, we see that (1.17) holds. ∎

We need another map taking $v^h\in L^2(\Omega^h)$ to $L^2(\Omega)$, in the form of a piecewise linear interpolant of the discrete values of v_i^h. To this end, let ϕ be the box spline function whose support is the hexagon D in Figure 1.1 with ϕ being linear on each triangle in Figure 1.1 and

$$\phi(i)=\begin{cases}1, & i=(0,0),\\ 0, & i\neq(0,0),\quad i\in\mathbb{Z}^2.\end{cases}$$

We dilate and translate ϕ to obtain the function

$$\phi_i^h(x):=\phi\Big(\frac{x}{h}-\Big(i+\Big(\frac{1}{2},\frac{1}{2}\Big)\Big)\Big). \tag{1.18}$$

We see that $\operatorname{supp}\phi_i^h$ is D dilated by h and translated by $\big(i+\big(\frac12,\frac12\big)\big)\,h$.

We define the interpolant $\operatorname{Int}v^h$ by

$$\operatorname{Int}v^h=\sum_{i\in\mathbb{Z}^2}\operatorname{Ext}_h v_i^h\phi_i. \tag{1.19}$$

We then have the following lemma.

LEMMA 1.5 (piecewise linear injector). *For any v^h in $L^2(\Omega^h)$ we have*

$$|\operatorname{Int} v^h|_{\mathrm{BV}(\Omega)} = \frac{1}{2}(J_{++}(v^h) + J_{--}(v^h)). \tag{1.20}$$

Additionally, there exists a constant C such that for all discrete functions v^h

$$\|I_h v^h - \operatorname{Int} v^h\|_{L^2(\Omega)} \le C\omega(v^h, 1)_{L^2(\Omega^h)}. \tag{1.21}$$

Proof. The proofs of (1.20) and (1.21) are just calculations, which can be found in [14]. For (1.20), the J_{++} terms come from triangles with the orientation of the triangle in the upper right quadrant of Figure 1.1, and the J_{--} terms come from triangles with the orientation of the triangle in the lower left quadrant. ∎

We need both continuous and discrete *smoothing operators*, which we define as follows. Assume that $\eta(x)$ is a fixed nonnegative, rotationally symmetric function with support in the unit disk; further, suppose that η is C^∞ and has integral 1. For $\epsilon > 0$ we define the scaled function $\eta_\epsilon(x) := \epsilon^{-2}\eta(x/\epsilon)$, $x \in \mathbb{R}^2$; we smooth a function $v \in L^p(\Omega)$, $1 \le p \le \infty$, by computing

$$(\mathcal{S}_\epsilon v)(x) := (\eta_\epsilon * \operatorname{Ext} v)(x) = \int_{\mathbb{R}^2} \eta_\epsilon(x - y) \operatorname{Ext} v(y)\, dy, \quad x \in \mathbb{R}^2.$$

Our discrete smoothing operator is defined simply as

$$S_L v^h := \frac{1}{(2L+1)^2} \sum_{|\ell| \le L} T_\ell \operatorname{Ext}_h v^h.$$

It is clear from these definitions that

$$T_\ell S_L \operatorname{Ext}_h v^h = S_L T_\ell \operatorname{Ext}_h v^h \quad \text{and} \quad \mathcal{T}_\tau \mathcal{S}_\epsilon \operatorname{Ext} v = \mathcal{S}_\epsilon \mathcal{T}_\tau \operatorname{Ext} v, \tag{1.22}$$

and that for any $1 \le p \le \infty$,

$$\|S_L v^h\|_{L^p(\Omega^h)} \le \|S_L v^h\|_{L^p(2\Omega^h)} \le \|\operatorname{Ext}_h v^h\|_{L^p(2\Omega^h)} \le 4\|v^h\|_{L^p(\Omega^h)} \tag{1.23}$$

and

$$\|\mathcal{S}_\epsilon v\|_{L^p(\Omega)} \le \|\mathcal{S}_\epsilon v\|_{L^p(2\Omega)} \le \|\operatorname{Ext} v\|_{L^p(2\Omega)} \le 4\|v\|_{L^p(\Omega)}. \tag{1.24}$$

For these continuous and discrete smoothing operators we have the following results.

LEMMA 1.6 (smoothing operators). *For all $v \in L^2(\Omega)$ and all discrete functions v^h, we have*

$$J_*(S_L v^h) \le J_*(v^h) \quad \textit{and} \quad |\mathcal{S}_\epsilon v|_{\mathrm{BV}(\Omega)} \le |v|_{\mathrm{BV}(\Omega)}. \tag{1.25}$$

There exists a constant $C > 0$ such that for all $M, t > 0$,

(1.26)
$$\omega(S_L v^h, M)_{L^2(\Omega^h)} \le C\omega(v^h, M)_{L^2(\Omega^h)} \quad \textit{and} \quad \omega(\mathcal{S}_\epsilon v, t)_{L^2(\Omega)} \le C\omega(v, t)_{L^2(\Omega)}.$$

Furthermore,

(1.27)
$$\|S_L v^h - v^h\|_{L^2(\Omega^h)} \le C\omega(v^h, L)_{L^2(\Omega^h)} \quad \textit{and} \quad \|\mathcal{S}_\epsilon v - v\|_{L^2(\Omega)} \le C\omega(v, \epsilon)_{L^2(\Omega)}.$$

We also have

$$|\mathcal{S}_\epsilon v|_{W^{2,1}(2\Omega)} \le \frac{C}{\epsilon}|v|_{\mathrm{BV}(\Omega)} \quad \textit{and} \quad |S_L v^h|_{W_h^{2,1}(2\Omega^h)} \le \frac{C}{Lh}|v^h|_{W^{1,1}(\Omega^h)}. \tag{1.28}$$

Proof. The first two inequalities follow simply because the BV semi-norm and J_* are convex and symmetric *on* $2\Omega^h$; see Lemma 1.2.

The two inequalities (1.26) follow from the definitions of S_L and $\mathcal{S}_\epsilon$ and Lemmas 1.1 and 1.2.

The next two inequalities follow from the definitions of S_L, $\mathcal{S}_\epsilon$, and the properties of the discrete and continuous moduli of smoothness; see also (1.5) of Lemma 1.1.

The bound on the discrete $W_h^{2,1}(2\Omega^h)$ semi-norm is a typical inverse inequality; to deal with the differences in the horizontal direction, we execute the following:

$$\begin{aligned}
&\sum_{i\in 2\Omega^h}\left|\frac{S_L v^h_{i+(1,0)} - 2S_L v^h_i + S_L v^h_{i-(1,0)}}{h^2}\right| h^2 \\
&= \sum_{i\in 2\Omega^h}\left|\frac{1}{(2L+1)^2}\sum_{|\ell|\le L}\frac{\operatorname{Ext}_h v^h_{i+(1,0)+\ell} - 2\operatorname{Ext}_h v^h_{i+\ell} + \operatorname{Ext}_h v^h_{i+\ell-(1,0)}}{h^2}\right| h^2 \\
&= \sum_{i\in 2\Omega^h}\frac{1}{(2L+1)^2}\left|\sum_{|\ell_2|\le L}\frac{\operatorname{Ext}_h v^h_{i+(L+1,\ell_2)} - \operatorname{Ext}_h v^h_{i+(L,\ell_2)}}{h^2}\right. \\
&\qquad\qquad \left. - \frac{\operatorname{Ext}_h v^h_{i-(L,\ell_2)} - \operatorname{Ext}_h v^h_{i-(L+1,\ell_2)}}{h^2}\right| h^2 \quad (\text{sum over } \ell_1) \\
&\le \sum_{i\in 2\Omega^h}\frac{1}{(2L+1)^2}\sum_{|\ell_2|\le L}\left\{\left|\frac{\operatorname{Ext}_h v^h_{i+(L+1,\ell_2)} - \operatorname{Ext}_h v^h_{i+(L,\ell_2)}}{h^2}\right|\right. \\
&\qquad\qquad \left. + \left|\frac{\operatorname{Ext}_h v^h_{i-(L,\ell_2)} - \operatorname{Ext}_h v^h_{i-(L+1,\ell_2)}}{h^2}\right|\right\} h^2 \\
&\le \frac{C}{(2L+1)h}\sum_{i\in 2\Omega^h}\left|\frac{\operatorname{Ext}_h v^h_{i+(1,0)} - \operatorname{Ext}_h v^h_i}{h}\right| h^2 \\
&\le \frac{C}{Lh}|v^h|_{W^{1,1}(\Omega^h)}.
\end{aligned}$$

For the bound on the $W^{2,1}(2\Omega)$ semi-norm, we again deal with derivatives in one direction only. We prove $\int_{2\Omega}|D_1^2\mathcal{S}_\epsilon v| \le \frac{C}{\epsilon}\int_\Omega |Dv|$.

In fact,

$$\begin{aligned}
\int_{2\Omega}|D_1^2\mathcal{S}_\epsilon v| &= \sup_{\phi\in C_0^1(2\Omega),\ |\phi|\le 1}\int_{\mathbb{R}^2}(D_1^2\mathcal{S}_\epsilon v)\,\phi \\
&= \sup_{\phi\in C_0^1(2\Omega),\ |\phi|\le 1}\int_{\mathbb{R}^2}(D_1\mathcal{S}_\epsilon v)\,D_1(-\phi) \\
&= \sup_{\phi\in C_0^1(2\Omega),\ |\phi|\le 1}\int_{\mathbb{R}^2} D_1(\eta_\epsilon * \operatorname{Ext} v)\,D_1(-\phi) \\
&= \sup_{\phi\in C_0^1(2\Omega),\ |\phi|\le 1}\int_{\mathbb{R}^2}(\operatorname{Ext} v)\,D_1(D_1\eta_\epsilon * \phi);
\end{aligned}$$

note that all but the first of these integrals are over $\mathbb{R}^2$. Notice

$$|D_1\eta_\epsilon * \phi| \le \|D_1\eta_\epsilon\|_{L^1}\|\phi\|_\infty \le \frac{C}{\epsilon},$$

and $D_1\eta_\epsilon * \phi \in C_0^\infty(2\Omega^\epsilon)$, where $2\Omega^\epsilon := \{x \mid \operatorname{dist}(x, \overline{2\Omega}) \le \epsilon\}$; therefore

$$\int_{2\Omega} |D_1^2 \mathcal{S}_\epsilon v| \le \frac{C}{\epsilon}\int_{2\Omega^\epsilon} |D \operatorname{Ext} v| \le \frac{C}{\epsilon}\int_{\Omega'} |D \operatorname{Ext} v| \le \frac{C}{\epsilon}\int_{\Omega} |Dv|\,,$$

where $\Omega' = \{(x,y) \mid |x|, |y| \le 3\}$. $\square$

2. Relationships between discrete and continuous variation and functionals. We need to compare continuous and discrete variation functionals, so we have the following technical lemma, which is proved in the appendix.

LEMMA 2.1 (TV bound). *There exists a $C > 0$ such that for any J_h and any $v \in L^1(\Omega)$,*

$$J_h(P_h v) \le |v|_{\mathrm{BV}(\Omega)} + Ch|\operatorname{Ext} v|_{W^{2,1}(2\Omega)}, \tag{2.1}$$

and for any v^h defined on Ω^h,

$$|\operatorname{Int} v^h|_{\mathrm{BV}(\Omega)} \le J_h(v^h) + Ch|\operatorname{Ext}_h v^h|_{W_h^{2,1}(2\Omega^h)}. \tag{2.2}$$

Our goal is to bound the difference between various continuous and discrete convex functionals defined on $L^2(\Omega)$ and $L^2(\Omega^h)$. We fix $\lambda > 0$. Given $g \in L^2(\Omega)$, we consider the (unique) minimizer u of the functional

$$E(v) = \frac{1}{2\lambda}\|v - g\|^2_{L^2(\Omega)} + |v|_{\mathrm{BV}(\Omega)}$$

and the (unique) minimizer u^h of the functional

$$E_h(v^h) = \frac{1}{2\lambda}\|v^h - P_h g\|^2_{L^2(\Omega^h)} + J_h(v^h),$$

where J_h is any of the discrete variational functionals defined above. Most of our analysis concerns the special case $J_h = J_*$.

It is difficult to compare u and u^h directly, because $J_*(u^h)$ and $|u|_{\mathrm{BV}(\Omega)}$ could be far apart, in general, even if $u^h \to u$ as $h \to 0$. However, there are smoothed versions of u and u^h, close to u and u^h, whose continuous and discrete variations are close, as the following lemma shows.

LEMMA 2.2 (TV consistency). *There exists a constant C such that for any discrete function $v^h \in L^2(\Omega^h)$ and any positive integer L, we have*

$$|\operatorname{Int} S_L v^h|_{\mathrm{BV}(\Omega)} \le J_*(v^h) + \frac{C}{L} J_*(v^h). \tag{2.3}$$

Furthermore, there is a constant C such that for any $v \in \mathrm{BV}(\Omega)$ and any positive ϵ and any discrete functional J_h, we have

$$J_h(P_h \mathcal{S}_\epsilon v) \le |v|_{\mathrm{BV}(\Omega)} + \frac{Ch}{\epsilon}|v|_{\mathrm{BV}(\Omega)}. \tag{2.4}$$

Proof. For the first inequality, we have from (2.2) in Lemma 2.1 (with $J_h = J_*$)

$$|\operatorname{Int} S_L v^h|_{\mathrm{BV}(\Omega)} \le J_*(S_L v^h) + Ch|S_L v^h|_{W_h^{2,1}(2\Omega^h)},$$

while from (1.25) in Lemma 1.6, $J_*(S_L v^h) \le J_*(v^h)$, and from (1.28) in the same lemma

$$|S_L v^h|_{W_h^{2,1}(2\Omega^h)} \le \frac{C}{Lh}|v^h|_{W^{1,1}(\Omega^h)} \le \frac{C}{Lh}J_*(v^h).$$

The second inequality follows from (1.13). Combining the previous inequalities gives (2.3).

For (2.4), we have from (2.1) in Lemma 2.1

$$J_h(P_h \mathcal{S}_\epsilon v) \le |\mathcal{S}_\epsilon v|_{\mathrm{BV}(\Omega)} + Ch|\mathcal{S}_\epsilon v|_{W^{2,1}(2\Omega)},$$

while (1.25) yields $|\mathcal{S}_\epsilon v|_{\mathrm{BV}(\Omega)} \le |v|_{\mathrm{BV}(\Omega)}$ and (1.28) gives $|\mathcal{S}_\epsilon v|_{W^{2,1}(2\Omega)} \le \frac{C}{\epsilon}|v|_{\mathrm{BV}(\Omega)}$. Combining these three inequalities yields (2.4). ∎

Now we compare discrete and continuous energy functionals.

LEMMA 2.3 (comparing discrete and continuous energies). *There exists a constant $C > 0$ such that for all J_h and for all $v \in \mathrm{BV}(\Omega)$*

$$\begin{aligned} E_h(P_h \mathcal{S}_\epsilon v) &\le E(v) + \frac{Ch}{\epsilon}|v|_{\mathrm{BV}(\Omega)} \\ &\quad + \frac{C}{\lambda}\|v-g\|_{L^2(\Omega)}\big(\omega(v,h)_{L^2(\Omega)} + \omega(v,\epsilon)_{L^2(\Omega)} + \omega(g,h)_{L^2(\Omega)}\big) \\ &\quad + \frac{C}{\lambda}\big(\omega(v,h)^2_{L^2(\Omega)} + \omega(v,\epsilon)^2_{L^2(\Omega)} + \omega(g,h)^2_{L^2(\Omega)}\big). \end{aligned} \tag{2.5}$$

Furthermore, for all discrete functions v^h

$$\begin{aligned} E(\operatorname{Int} S_L v^h) &\le E_h(v^h) + \frac{C}{L}J_*(v^h) \\ &\quad + \frac{C}{\lambda}\|v^h - P_h g\|_{L^2(\Omega^h)}\big(\omega(v^h,L)_{L^2(\Omega^h)} + \omega(g,h)_{L^2(\Omega)}\big) \\ &\quad + \frac{C}{\lambda}\big(\omega(v^h,L)^2_{L^2(\Omega)} + \omega(g,h)^2_{L^2(\Omega)}\big). \end{aligned} \tag{2.6}$$

Proof. We have

$$E_h(P_h\mathcal{S}_\epsilon v) = J_h(P_h\mathcal{S}_\epsilon v) + \frac{1}{2\lambda}\|P_h\mathcal{S}_\epsilon v - P_h g\|^2_{L^2(\Omega^h)}. \tag{2.7}$$

From (2.4), we see that the first term on the right is bounded by

$$|v|_{\mathrm{BV}(\Omega)} + \frac{Ch}{\epsilon}|v|_{\mathrm{BV}(\Omega)}. \tag{2.8}$$

Now

$$\|P_h\mathcal{S}_\epsilon v - P_h g\|^2_{L^2(\Omega^h)} = \|I_h P_h \mathcal{S}_\epsilon v - I_h P_h g\|^2_{L^2(\Omega)},$$

and the quantity on the right can be written as

$$\begin{aligned}
&\|(I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v)+(\mathcal{S}_\epsilon v-v)+(v-g)+(g-I_hP_hg)\|^2_{L^2(\Omega)}\\
&\le \|v-g\|^2_{L^2(\Omega)}+2\|v-g\|_{L^2(\Omega)}\\
&\quad\times\|(I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v)+(\mathcal{S}_\epsilon v-v)+(g-I_hP_hg)\|_{L^2(\Omega)}\\
&\quad+\|(I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v)+(\mathcal{S}_\epsilon v-v)+(g-I_hP_hg)\|^2_{L^2(\Omega)}\\
&\le \|v-g\|^2_{L^2(\Omega)}+2\|v-g\|_{L^2(\Omega)}\\
&\quad\times\big(\|I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v\|_{L^2(\Omega)}+\|\mathcal{S}_\epsilon v-v\|_{L^2(\Omega)}+\|g-I_hP_hg\|_{L^2(\Omega)}\big)\\
&\quad+C\big(\|I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v\|^2_{L^2(\Omega)}+\|\mathcal{S}_\epsilon v-v\|^2_{L^2(\Omega)}+\|g-I_hP_hg\|^2_{L^2(\Omega)}\big).
\end{aligned}$$

From (1.27) we can bound

$$\|\mathcal{S}_\epsilon v-v\|_{L^2(\Omega)}\le C\omega(v,\epsilon)_{L^2(\Omega)}$$

and from (1.16) we know that

$$\|I_hP_hg-g\|_{L^2(\Omega)}\le C\omega(g,h)_{L^2(\Omega)}.$$

We also have from (1.16) and (1.26)

$$\|I_hP_h\mathcal{S}_\epsilon v-\mathcal{S}_\epsilon v\|_{L^2(\Omega)}\le C\omega(\mathcal{S}_\epsilon v,h)_{L^2(\Omega)}\le C\omega(v,h)_{L^2(\Omega)}.$$

Thus,

$$\begin{aligned}
\|P_h\mathcal{S}_\epsilon v-P_hg\|^2_{L^2(\Omega^h)}\le{}&\|v-g\|^2_{L^2(\Omega)}+C\|v-g\|_{L^2(\Omega)}\\
&\times\big(\omega(v,h)_{L^2(\Omega)}+\omega(v,\epsilon)_{L^2(\Omega)}+\omega(g,h)_{L^2(\Omega)}\big)\\
&+C\big(\omega(v,h)^2_{L^2(\Omega)}+\omega(v,\epsilon)^2_{L^2(\Omega)}+\omega(g,h)^2_{L^2(\Omega)}\big).
\end{aligned}$$

Using this inequality as well as (2.8) in (2.7) yields (2.5).

Now let v^h be any discrete function. Then

$$E(\operatorname{Int}S_Lv^h)=|\operatorname{Int}S_Lv^h|_{\mathrm{BV}(\Omega)}+\frac{1}{2\lambda}\|\operatorname{Int}S_Lv^h-g\|^2_{L^2(\Omega)}. \tag{2.9}$$

By (2.3), the first term on the right is bounded by

$$J_*(v^h)+\frac{C}{L}J_*(v^h). \tag{2.10}$$

Now

$$\begin{aligned}
&\|\operatorname{Int}S_Lv^h-g\|^2_{L^2(\Omega)}\\
&=\|(\operatorname{Int}S_Lv^h-I_hS_Lv^h)+(I_hS_Lv^h-I_hv^h)+(I_hv^h-I_hP_hg)+(I_hP_hg-g)\|^2_{L^2(\Omega)}\\
&\le\|I_hv^h-I_hP_hg\|^2_{L^2(\Omega)}+2\|I_hv^h-I_hP_hg\|_{L^2(\Omega)}\\
&\qquad\times\|(\operatorname{Int}S_Lv^h-I_hS_Lv^h)+(I_hS_Lv^h-I_hv^h)+(I_hP_hg-g)\|_{L^2(\Omega)}\\
&\quad+\|(\operatorname{Int}S_Lv^h-I_hS_Lv^h)+(I_hS_Lv^h-I_hv^h)+(I_hP_hg-g)\|^2_{L^2(\Omega)}\\
&\le\|I_hv^h-I_hP_hg\|^2_{L^2(\Omega)}+2\|I_hv^h-I_hP_hg\|_{L^2(\Omega)}\\
&\qquad\times\big(\|\operatorname{Int}S_Lv^h-I_hS_Lv^h\|_{L^2(\Omega)}+\|I_hS_Lv^h-I_hv^h\|_{L^2(\Omega)}+\|I_hP_hg-g\|_{L^2(\Omega)}\big)\\
&\quad+C(\|\operatorname{Int}S_Lv^h-I_hS_Lv^h\|^2_{L^2(\Omega)}+\|I_hS_Lv^h-I_hv^h\|^2_{L^2(\Omega)}+\|I_hP_hg-g\|^2_{L^2(\Omega)}).
\end{aligned}$$

Since, for all discrete v^h, $\|I_h v^h\|_{L^2(\Omega)} = \|v^h\|_{L^2(\Omega^h)}$, the quantity above is bounded by

$$\begin{aligned}
&\|v^h - P_h g\|^2_{L^2(\Omega^h)} + 2\|v^h - P_h g\|_{L^2(\Omega^h)} \\
&\qquad \times \big(\|\operatorname{Int} S_L v^h - I_h S_L v^h\|_{L^2(\Omega)} + \|S_L v^h - v^h\|_{L^2(\Omega^h)} + \|I_h P_h g - g\|_{L^2(\Omega)}\big) \\
&\quad + C\big(\|\operatorname{Int} S_L v^h - I_h S_L v^h\|^2_{L^2(\Omega)} + \|S_L v^h - v^h\|^2_{L^2(\Omega^h)} + \|I_h P_h g - g\|^2_{L^2(\Omega)}\big).
\end{aligned}$$

From (1.27) we have

$$\|S_L v^h - v^h\|_{L^2(\Omega^h)} \le C\omega(v^h, L)_{L^2(\Omega^h)}.$$

By (1.21) and (1.26) we have

$$\begin{aligned}
\|\operatorname{Int} S_L v^h - I_h S_L v^h\|_{L^2(\Omega)} &\le C\omega(S_L v^h, 1)_{L^2(\Omega^h)} \\
&\le C\omega(v^h, 1)_{L^2(\Omega^h)} \le C\omega(v^h, L)_{L^2(\Omega^h)}.
\end{aligned}$$

Combining these inequalities, we have

$$\begin{aligned}
\|\operatorname{Int} S_L v^h - g\|^2_{L^2(\Omega)} \le \|v^h - P_h g\|^2_{L^2(\Omega^h)} &\\
+ C\|v^h - P_h g\|_{L^2(\Omega^h)}&\big(\omega(v^h, L)_{L^2(\Omega^h)} + \omega(g, h)_{L^2(\Omega)}\big) \\
+ C\big(\omega(v^h, L)^2_{L^2(\Omega)} &+ \omega(g, h)^2_{L^2(\Omega)}\big).
\end{aligned}$$

Combining this inequality with (2.9) and (2.10) yields (2.6). ∎

3. Properties of the continuous and discrete minimizers. We need to discuss some properties of minimizers of the discrete and continuous functionals. We begin by comparing functionals on Ω and Ω^h and the corresponding functionals on 2Ω and $2\Omega^h$. We remind the reader of the notations used in Lemma 1.2.

LEMMA 3.1 (extending minimizers). *If u^h is the minimizer of the functional*

$$E_h^{\Omega^h}(v^h) = E_h(v^h) = \frac{1}{2\lambda}\|v^h - g^h\|^2_{L^2(\Omega^h)} + J_*^{\Omega^h}(v^h), \tag{3.1}$$

then $\operatorname{Ext}_h u^h$ is the minimizer over all discrete functions v^h defined on $2\Omega^h$ of the functional

$$E_h^{2\Omega^h}(v^h) = \frac{1}{2\lambda}\|v^h - \operatorname{Ext}_h g^h\|^2_{L^2(2\Omega^h)} + J_*^{2\Omega^h}(v^h) \tag{3.2}$$

with periodic boundary conditions.

Similarly, if u is the minimizer of

$$E^{\Omega}(v) = E(v) = \frac{1}{2\lambda}\|v - g\|^2_{L^2(\Omega)} + |v|_{\mathrm{BV}(\Omega)}, \tag{3.3}$$

then $\operatorname{Ext} u$ is the minimizer of

$$E^{2\Omega}(v) = \frac{1}{2\lambda}\|v - \operatorname{Ext} g\|^2_{L^2(2\Omega)} + |v|_{\mathrm{BV}(2\Omega)}, \tag{3.4}$$

again with periodic boundary conditions.

Furthermore, if u and w are minimizers of (3.3) *with data g and h, respectively, then $\|u-w\|_{L^2(\Omega)} \le \|g-h\|_{L^2(\Omega)}$; similarly for the discrete and continuous minimizers of* (3.1)–(3.4). *Thus, for the two periodic problems* (3.2) *and* (3.4) *we have*

$$\|\operatorname{Ext}_h u^h - T_\ell \operatorname{Ext}_h u^h\|_{L^2(2\Omega^h)} \le \|\operatorname{Ext}_h g^h - T_\ell \operatorname{Ext}_h g^h\|_{L^2(2\Omega^h)} \tag{3.5}$$

and

$$\|\operatorname{Ext} u - \mathcal{T}_\tau \operatorname{Ext} u\|_{L^2(2\Omega)} \le \|\operatorname{Ext} g - \mathcal{T}_\tau \operatorname{Ext} g\|_{L^2(2\Omega)}. \tag{3.6}$$

Proof. Beginning with Lemma 1.2, the discrete extension result can be proved with a tedious calculation, which can be found in [14]. The rest of the lemma is standard. ∎

The results of the next lemma follow quickly from the previous one and Lemma 1.1 and we present them without proof.

LEMMA 3.2 (smoothness bounds). *Assume u is the minimizer of $E(v)$ from* (3.3) *and u^h is the minimizer of $E_h(v^h)$ from* (3.1). *Then*

$$\omega(u,\epsilon)_{L^2(\Omega)} \le C\omega(g,\epsilon)_{L^2(\Omega)} \tag{3.7}$$

and

$$\omega(u^h, L)_{L^2(\Omega^h)} \le C\omega(P_h g, L)_{L^2(\Omega^h)} \le C\omega(g, Lh)_{L^2(\Omega)}. \tag{3.8}$$

4. Proof of the main theorems. We now bound the difference between discrete and continuous functionals at their respective minimizers.

THEOREM 4.1 (functional difference). *Assume u is the minimizer of $E(v)$ from* (3.3) *and u^h is the minimizer of $E_h(v^h)$ from* (3.1).

Then if $\epsilon = h^{1/(\alpha+1)}$, we have

$$E_h(P_h\mathcal{S}_\epsilon u) \le E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}, \tag{4.1}$$

and if L is set to the integer part of $h^{-\alpha/(\alpha+1)}$, then

$$E(\operatorname{Int} S_L u^h) \le E_h(u^h) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}. \tag{4.2}$$

Finally,

$$|E(u) - E_h(u^h)| \le \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}. \tag{4.3}$$

Proof. We mainly use Lemmas 2.3 and 3.2. By (2.5) of Lemma 2.3,

$$\begin{aligned} E_h(P_h\mathcal{S}_\epsilon u) \le E(u) &+ \frac{Ch}{\epsilon}|u|_{\mathrm{BV}(\Omega)} \\ &+ \frac{C}{\lambda}\|u-g\|_{L^2(\Omega)}\big(\omega(u,h)_{L^2(\Omega)} + \omega(u,\epsilon)_{L^2(\Omega)} + \omega(g,h)_{L^2(\Omega)}\big) \\ &+ \frac{C}{\lambda}\big(\omega(u,h)^2_{L^2(\Omega)} + \omega(u,\epsilon)^2_{L^2(\Omega)} + \omega(g,h)^2_{L^2(\Omega)}\big). \end{aligned}$$

We then note $\|u-g\|_{L^2(\Omega)} \le \|g\|_{L^2(\Omega)}$ and $|u|_{\mathrm{BV}(\Omega)} \le \frac{1}{2\lambda}\|g\|^2_{L^2(\Omega)}$ and apply (3.7) to obtain

$$\begin{aligned} E_h(P_h\mathcal{S}_\epsilon u) \le{}& E(u) + \frac{Ch}{\epsilon\lambda}\|g\|^2_{L^2(\Omega)} \\ &+ \frac{C}{\lambda}\|g\|_{L^2(\Omega)}\big(\omega(g,\epsilon)_{L^2(\Omega)} + \omega(g,h)_{L^2(\Omega)}\big) \\ &+ \frac{C}{\lambda}\big(\omega(g,h)^2_{L^2(\Omega)} + \omega(g,\epsilon)^2_{L^2(\Omega)}\big). \end{aligned}$$

Now, $\omega(g,t)_{L^2(\Omega)} \le |g|_{\mathrm{Lip}(\alpha,L^2(\Omega))}t^\alpha$, $t>0$. Thus

$$\begin{aligned} E_h(P_h\mathcal{S}_\epsilon u) \le{}& E(u) + \frac{Ch}{\epsilon\lambda}\|g\|^2_{L^2(\Omega)} \\ &+ \frac{C}{\lambda}\|g\|_{L^2(\Omega)}|g|_{\mathrm{Lip}(\alpha,L^2(\Omega))}(\epsilon^\alpha + h^\alpha) \\ &+ \frac{C}{\lambda}|g|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}(\epsilon^{2\alpha} + h^{2\alpha}) \\ \le{}& E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}\Big(\frac{h}{\epsilon} + \epsilon^\alpha + h^\alpha + \epsilon^{2\alpha} + h^{2\alpha}\Big). \end{aligned}$$

We know at a minimum that $1 > \epsilon > h$, so setting the largest error terms h/ϵ and ϵ^α equal, i.e., setting $\epsilon = h^{1/(\alpha+1)}$, we have

$$E_h(P_h\mathcal{S}_\epsilon u) \le E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}\big(h^{\alpha/(\alpha+1)} + h^\alpha + h^{2\alpha} + h^{2\alpha/(\alpha+1)}\big).$$

Thus we obtain (4.1).

We point out that (4.1) holds for any discrete variation J_h defined in (1.10). More generally it holds for any discrete variation satisfying Lemma 2.2.

Similarly, if one begins with (2.6), notes that $\|u^h - P_hg\|_{L^2(\Omega^h)} \le \|P_hg\|_{L^2(\Omega^h)} \le \|g\|_{L^2(\Omega)}$ and $J_h(u^h) \le \frac{1}{2\lambda}\|P_hg\|^2_{L^2(\Omega^h)} \le \frac{1}{2\lambda}\|g\|^2_{L^2(\Omega)}$, and applies (3.8), one finds on setting L to the integer part of $h^{-\alpha/(\alpha+1)}$ that

$$E(\operatorname{Int} S_L u^h) \le E_h(u^h) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}h^{\alpha/(\alpha+1)},$$

which is (4.2).

Because u and u^h are minimizers of their respective functionals, we have

$$E_h(u^h) \le E_h(P_h\mathcal{S}_\epsilon u) \le E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}h^{\alpha/(\alpha+1)} \tag{4.4}$$

and

$$E(u) \le E(\operatorname{Int} S_L u^h) \le E_h(u^h) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))}h^{\alpha/(\alpha+1)}. \tag{4.5}$$

Then (4.3) is proved. ∎

To show the error bound for minimizers, we need the following result, which can be proved easily using classical arguments.

LEMMA 4.1. *Assume u is the minimizer of $E(v)$ from* (3.3) *and u^h is the minimizer of $E_h(v^h)$ from* (3.1). *Then for any $v \in \mathrm{BV}(\Omega)$,*

$$\|v-u\|^2_{L^2(\Omega)} \le 2\lambda(E(v) - E(u)). \tag{4.6}$$

Also, for any v^h *defined on* Ω^h,

$$\|v^h - u^h\|^2_{L^2(\Omega^h)} \le 2\lambda(E_h(v^h) - E_h(u^h)). \tag{4.7}$$

THEOREM 4.2 (minimizer difference). *Let* $g \in \mathrm{Lip}(\alpha, L^2(\Omega))$. *Assume that* u *is the minimizer of* $E(v)$ *from* (3.3) *and* u^h *is the minimizer of* $E_h(v^h)$ *from* (3.1). *Then*

$$\|I_h u^h - u\|^2_{L^2(\Omega)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}.$$

Proof. We apply (4.7) with $v^h = P_h\mathcal{S}_\epsilon u$ and $\epsilon = h^{1/(\alpha+1)}$:

$$\begin{aligned}\|P_h\mathcal{S}_\epsilon u - u^h\|^2_{L^2(\Omega^h)} &\le 2\lambda\big(E_h(P_h\mathcal{S}_\epsilon u^h) - E_h(u^h)\big)\\ &\le 2\lambda\Big[\Big(E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}\Big)\\ &\qquad + \Big(-E(u) + \frac{C}{\lambda}\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}\Big)\Big].\end{aligned}$$

The first substitution is by (4.1); the second is by (4.5). Thus we have

$$\|P_h\mathcal{S}_\epsilon u - u^h\|^2_{L^2(\Omega^h)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}. \tag{4.8}$$

Then

$$\begin{aligned}\|I_h u^h - u\|^2_{L^2(\Omega)} &= \|I_h u^h - I_hP_h\mathcal{S}_\epsilon u + I_hP_h\mathcal{S}_\epsilon u - \mathcal{S}_\epsilon u + \mathcal{S}_\epsilon u - u\|^2_{L^2(\Omega)}\\ &\le 3\big(\|I_h u^h - I_hP_h\mathcal{S}_\epsilon u\|^2_{L^2(\Omega)} + \|I_hP_h\mathcal{S}_\epsilon u - \mathcal{S}_\epsilon u\|^2_{L^2(\Omega)}\\ &\qquad + \|\mathcal{S}_\epsilon u - u\|^2_{L^2(\Omega)}\big).\end{aligned}$$

Because $\|I_h v^h\|_{L^2(\Omega)} = \|v^h\|_{L^2(\Omega^h)}$ for any v^h, it follows from (4.8) that

$$\|I_h u^h - I_hP_h\mathcal{S}_\epsilon u\|^2_{L^2(\Omega)} = \|P_h\mathcal{S}_\epsilon u - u^h\|^2_{L^2(\Omega^h)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}.$$

To bound $\|I_hP_h\mathcal{S}_\epsilon u - \mathcal{S}_\epsilon u\|_{L^2(\Omega)}$, by (1.16), (1.26), and (3.7), we have

$$\begin{aligned}\|I_hP_h\mathcal{S}_\epsilon u - \mathcal{S}_\epsilon u\|_{L^2(\Omega)} &\le C\omega(\mathcal{S}_\epsilon u, h)_{L^2(\Omega)}\\ &\le C\omega(u,h)_{L^2(\Omega)} \le C\omega(g,h)_{L^2(\Omega)}.\end{aligned} \tag{4.9}$$

Finally by (1.27) and (3.7),

$$\|\mathcal{S}_\epsilon u - u\|_{L^2(\Omega)} \le C\omega(u,\epsilon) \le C\omega(g,\epsilon). \tag{4.10}$$

Thus combining (4.8), (4.9), and (4.10), we have

$$\begin{aligned}\|I_h u^h - u\|^2_{L^2(\Omega)} &\le C\left(\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)} + \omega(g,h)^2_{L^2(\Omega)} + \omega(g,\epsilon)^2_{L^2(\Omega)}\right)\\ &\le C\left(\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)} + \|g\|^2_{L^2(\Omega)} h^{2\alpha} + \|g\|^2_{L^2(\Omega)} h^{2\alpha/(\alpha+1)}\right).\end{aligned}$$

Because the first term dominates the others, we have

$$\|I_h u^h - u\|^2_{L^2(\Omega)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}. \qquad \square$$

5. Error bounds for the upwind scheme. In this section, we provide error bounds for the "upwind" scheme. The upwind discrete gradient operator $-\nabla_h$ is defined by

$$(-\nabla_h)v_i^h = \frac{1}{h}\begin{pmatrix} \operatorname{Ext}_h v_i^h - \operatorname{Ext}_h v_{i+(1,0)}^h \\ \operatorname{Ext}_h v_i^h - \operatorname{Ext}_h v_{i-(1,0)}^h \\ \operatorname{Ext}_h v_i^h - \operatorname{Ext}_h v_{i+(0,1)}^h \\ \operatorname{Ext}_h v_i^h - \operatorname{Ext}_h v_{i-(0,1)}^h \end{pmatrix}. \tag{5.1}$$

The upwind discrete variation is then defined by

$$J_U(v^h) = \sum_{i\in\Omega^h} \left|(-\nabla_h)v_i^h \vee 0\right| h^2, \tag{5.2}$$

where 0 is the vector $(0,0,0,0)$, and $p\vee q$ and $p\wedge q$ are the componentwise maximum and minimum, respectively, of the vectors $p,q\in\mathbb{R}^4$.

In other words, we include a difference in the vector norm of the ith term in (5.2) only if v^h is increasing into v_i^h. Nothing changes in the following proofs (and one sees little change in the images themselves) if we change componentwise maximum ($\vee$) to componentwise minimum ($\wedge$) in (5.2). In their paper, Osher and Sethian [12] were solving Hamilton–Jacobi equations where this substitution could not be made: their problem, unlike ours, has a true notion of "wind."

To prove the result for the upwind scheme, we need to adapt to J_U the previous lemmas involving J_*.

First we shall prove the convexity of J_U.

LEMMA 5.1. *J_U is convex.*

Proof. First note that for two vectors $p,q\in\mathbb{R}^n$, it is easy to verify

$$0 \le (p+q)\vee 0 \le p\vee 0 + q\vee 0,$$

where inequality $p\le q$ means $p_i\le q_i$ for each index i. Thus,

$$|(p+q)\vee 0| \le |p\vee 0| + |q\vee 0|. \tag{5.3}$$

We apply (5.3) to each term in (5.2) of $J_U\left(\lambda f^h + (1-\lambda)g^h\right)$, where $1>\lambda>0$ and f^h and g^h are discrete functions, to find that

$$\begin{aligned} J_U\left(\lambda f^h + (1-\lambda)g^h\right) &= \sum_i \left|(-\nabla_h)(\lambda f^h + (1-\lambda)g^h)_i \vee 0\right| h^2 \\ &\le \sum_i \left\{\left|\lambda(-\nabla_h)f_i^h \vee 0\right| + \left|(1-\lambda)(-\nabla_h)g_i^h \vee 0\right|\right\} h^2 \\ &= \lambda J_U(f^h) + (1-\lambda)J_U(g^h). \qquad \square \end{aligned}$$

In the following we use the notation $\nabla_x^\oplus$ and $\nabla_y^\oplus$ defined for $\oplus\in\{+,-\}$ by

$$\nabla_x^\oplus v_i^h = \oplus\frac{\operatorname{Ext}_h v_{i\oplus(1,0)}^h - \operatorname{Ext}_h v_i^h}{h}, \quad \nabla_y^\oplus v_i^h = \oplus\frac{\operatorname{Ext}_h v_{i\oplus(0,1)}^h - \operatorname{Ext}_h v_i^h}{h}. \tag{5.4}$$

Note that the divided differences are applied to the *extended* discrete function, and that the difference is zero if $i\in\Omega^h$ and the other index is outside Ω^h.

Using these operators, we can write

$$J_U(v^h) = \sum_{i\in\Omega^h} \left| \begin{pmatrix} -\nabla_x^+ v_i^h \vee 0 \\ \nabla_x^- v_i^h \vee 0 \\ -\nabla_y^+ v_i^h \vee 0 \\ \nabla_y^- v_i^h \vee 0 \end{pmatrix} \right| h^2. \tag{5.5}$$

The following lemma corresponds to (1.13).

LEMMA 5.2. *J_U is equivalent to $|\cdot|_{W^{1,1}(\Omega^h)}$, where $|\cdot|_{W^{1,1}(\Omega^h)}$ is the discrete semi-norm defined in* (1.12).

Proof. Trivially,

$$\begin{aligned} &\frac{1}{2}\sum_i \Big\{ \left|-\nabla_x^+ v_i^h \vee 0\right| + \left|\nabla_x^- v_i^h \vee 0\right| + \left|-\nabla_y^+ v_i^h \vee 0\right| + \left|\nabla_y^- v_i^h \vee 0\right| \Big\} h^2 \\ &\le \sum_i \sqrt{\Big\{ \left|-\nabla_x^+ v_i^h \vee 0\right|^2 + \left|\nabla_x^- v_i^h \vee 0\right|^2 + \left|-\nabla_y^+ v_i^h \vee 0\right|^2 + \left|\nabla_y^- v_i^h \vee 0\right|^2 \Big\}}\, h^2 \\ &\le \sum_i \Big\{ \left|-\nabla_x^+ v_i^h \vee 0\right| + \left|\nabla_x^- v_i^h \vee 0\right| + \left|-\nabla_y^+ v_i^h \vee 0\right| + \left|\nabla_y^- v_i^h \vee 0\right| \Big\} h^2. \end{aligned}$$

The middle sum is $J_U(v^h)$, so we need to prove that the last sum equals $|v^h|_{W^{1,1}(\Omega^h)}$. Note that

$$\left|-\nabla_x^+ v_i^h \vee 0\right| + \left|\nabla_x^- v_{i+(1,0)}^h \vee 0\right| = \left|-\nabla_x^+ v_i^h \vee 0\right| + \left|\nabla_x^+ v_i^h \vee 0\right| = \left|\nabla_x^+ v_i^h\right|,$$

so the absolute value of each horizontal and vertical difference in v^h is included precisely once in the last sum, so it equals $|v^h|_{W^{1,1}(\Omega^h)}$. ∎

The following lemma corresponds to Lemma 1.3.

LEMMA 5.3.

$$\left|J_U(v^h) - J_{\oplus\ominus}(v^h)\right| \le h \left|\mathrm{Ext}_h\, v^h\right|_{W_h^{2,1}(2\Omega^h)}, \tag{5.6}$$

where $J_{\oplus\ominus}$ is any discrete variation defined in (1.9).

Proof. We prove only the case for $J_{\oplus\ominus} = J_{++}$. The other cases are the same.

Note that

$$|\nabla_x^+ v_i^h|^2 = |\nabla_x^+ v_i^h \vee 0|^2 + |(-\nabla_x^+) v_i^h \vee 0|^2,$$

so we can write $J_{++}(v^h)$ in a way similar to $J_U(v^h)$ as

$$J_{++}(v^h) = \sum_{i\in\Omega^h} \left| \begin{pmatrix} -\nabla_x^+ v_i^h \vee 0 \\ \nabla_x^+ v_i^h \vee 0 \\ -\nabla_y^+ v_i^h \vee 0 \\ \nabla_y^+ v_i^h \vee 0 \end{pmatrix} \right| h^2.$$

Thus,

$$\begin{aligned}\left|J_U(v^h)-J_{++}(v^h)\right| &= \left|\sum_{i\in\Omega^h}\left\{\left|\begin{pmatrix}-\nabla_x^+ v_i^h\vee 0\\ \nabla_x^- v_i^h\vee 0\\ -\nabla_y^+ v_i^h\vee 0\\ \nabla_y^- v_i^h\vee 0\end{pmatrix}\right| - \left|\begin{pmatrix}-\nabla_x^+ v_i^h\vee 0\\ \nabla_x^+ v_i^h\vee 0\\ -\nabla_y^+ v_i^h\vee 0\\ \nabla_y^+ v_i^h\vee 0\end{pmatrix}\right|\right\} h^2\right|\\
&\le \sum_{i\in\Omega^h}\left|\begin{pmatrix}0\\ \nabla_x^- v_i^h\vee 0-\nabla_x^+ v_i^h\vee 0\\ 0\\ \nabla_y^- v_i^h\vee 0-\nabla_y^+ v_i^h\vee 0\end{pmatrix}\right| h^2\\
&\le \sum_{i\in\Omega^h}\left(\left|\nabla_x^- v_i^h\vee 0-\nabla_x^+ v_i^h\vee 0\right|+\left|\nabla_y^- v^h\vee 0-\nabla_y^+ v^h\vee 0\right|\right)h^2.\end{aligned}$$

Because $|a\vee 0-b\vee 0|\le|a-b|$, we have

$$\begin{aligned}\left|J_U(v^h)-J_{++}(v^h)\right| &\le \sum_{i\in\Omega^h}\left(\left|\nabla_x^- v_i^h-\nabla_x^+ v_i^h\right|+\left|\nabla_y^- v_i^h-\nabla_y^+ v_i^h\right|\right)h^2\\
&\le h\left|\operatorname{Ext}_h v^h\right|_{W_h^{2,1}(2\Omega^h)}. \qquad \square\end{aligned}$$

We use Lemmas 5.2 and 5.3 to prove the following lemma that corresponds to Lemma 2.1.

LEMMA 5.4. *There exists a $C>0$ such for any $v\in L^1(\Omega)$,*

$$J_U(P_h v)\le |v|_{\mathrm{BV}(\Omega)}+Ch|\operatorname{Ext} v|_{W^{2,1}(2\Omega)}, \tag{5.7}$$

and for any v^h defined on Ω^h,

$$|\operatorname{Int} v^h|_{\mathrm{BV}(\Omega)}\le J_U(v^h)+Ch|\operatorname{Ext}_h v^h|_{W_h^{2,1}(2\Omega^h)}. \tag{5.8}$$

Proof. The second inequality can be proved by simply combining (2.2) and (5.6).

To prove the first inequality, again we assume that $\operatorname{Ext} v\in W^{2,1}(2\Omega)$; otherwise it is trivial. We apply Lemma 5.3 with $v^h=P_h v$; then

$$J_U(P_h v)\le J_{\oplus\ominus}(P_h v)+Ch|\operatorname{Ext}_h P_h v|_{W_h^{2,1}(2\Omega^h)}.$$

Then by (2.1) in Lemma 2.1,

$$\begin{aligned}J_U(P_h v) &\le |v|_{\mathrm{BV}(\Omega)}+Ch|\operatorname{Ext} v|_{W^{2,1}(2\Omega)}+Ch|\operatorname{Ext}_h P_h v|_{W_h^{2,1}(2\Omega^h)}\\
&= |v|_{\mathrm{BV}(\Omega)}+Ch|\operatorname{Ext} v|_{W^{2,1}(2\Omega)}+Ch|P_h\operatorname{Ext} v|_{W_h^{2,1}(2\Omega^h)}\\
&\le |v|_{\mathrm{BV}(\Omega)}+Ch|\operatorname{Ext} v|_{W^{2,1}(2\Omega)}.\end{aligned}$$

The last line follows from (1.17). $\square$

Lemma 5.5 is the counterpart of the first inequality (1.25) in Lemma 1.6.

LEMMA 5.5. *$J_U(S_L v^h)\le J_U(v^h)$.*

Proof. The result comes from the symmetry and convexity of J_U. The proof is exactly the same as the proof for J_* in Lemma 1.6. $\square$

We note that the proofs of Lemmas 3.1 and 4.1 carry over directly to J_U, and we obtain the following theorem for the upwind discrete variation.

TABLE 6.1

$L^2(\Omega)$ errors on grids of size 128, 256, and 512, and differences $\sigma = \|g-u\|_{L^2(\Omega)}$ of 16, 32, and 64, with initial data a multiple of the characteristic function of a disk (6.1). Columns 1–3 are the results with the anisotropic approximation $J_h = J_{++}$ to $|\cdot|_{\mathrm{BV}(\Omega)}$; columns 4–6 are the results with the upwind approximation $J_h = J_U$; α is the estimated order of convergence, $\|u-u^h\|_{L_2} \approx Ch^\alpha$.

	$L^2(\Omega)$ difference between continuous and discrete solutions					
	"Anisotropic" J_h			"Upwind" J_h		
Resolution	$\sigma = 16$	$\sigma = 32$	$\sigma = 64$	$\sigma = 16$	$\sigma = 32$	$\sigma = 64$
128×128	10.637	9.223	6.004	9.925	8.312	5.143
256×256	7.929	6.981	4.542	7.061	6.051	3.795
512×512	6.029	5.360	3.495	5.185	4.503	2.852
α	0.410	0.392	0.390	0.468	0.442	0.425

THEOREM 5.1 (error bounds for upwind scheme). *Assume u is the minimizer of $E(v)$ from (3.3) for $g \in \mathrm{Lip}(\alpha, L^2(\Omega))$ and u^h is the minimizer of the discrete functional*

$$E_h(v^h) = \frac{1}{2\lambda}\|v^h - P_h g\|^2_{L^2(\Omega^h)} + J_U(v^h).$$

Then

$$\|I_h u^h - u\|^2_{L^2(\Omega)} \le C\|g\|^2_{\mathrm{Lip}(\alpha,L^2(\Omega))} h^{\alpha/(\alpha+1)}.$$

The proof is the same as the proof for the symmetric discrete variation J_*.

6. Discussion and extensions. We remark that our error bounds are not optimal in an approximation-theory sense. In general, with suitably smooth piecewise polynomials, one can approximate a function in $\mathrm{Lip}(\alpha, L^p(\Omega))$ to order h^α in $L^p(\Omega)$ for $0 < p < \infty$. So one can approximate $\mathrm{Lip}(\alpha, L^2(\Omega))$ functions in $L^2(\Omega)$ to order h^α; in contrast, we derive an error bound of $h^{\alpha/(2\alpha+2)}$.

The characteristic function of a disk is in $\mathrm{Lip}(\alpha, L^p(\Omega))$ for $\alpha = 1/p$; if g is the characteristic function of a disk, then the minimizer u of $E(v)$ is again the characteristic function of a disk (for λ small enough). Thus one has $\alpha = 1/2$ and one can expect at most a convergence rate of $h^{1/2}$ in $L^2(\Omega)$. Our results bound the $L^2(\Omega)$ error by $Ch^{\alpha/(2(\alpha+1))} = Ch^{1/6}$.

In [5] some numerical experiments were conducted; with permission we reprint a table showing the results of these computations with initial data

$$g = 255\chi_{|x-(\frac12,\frac12)|\le\frac14}. \tag{6.1}$$

While the computations were iterative, the iterated approximation to the true discrete solution was provably within a distance in $L^2(\Omega)$ of $1/4$ to the true discrete solution.

We note that the upwind scheme has slightly smaller errors (because of less smoothing in the discrete solution at the edges of the disk) and a slightly higher estimated rate of convergence. In both cases, the estimated rate of convergence is strictly between our bound of $h^{1/6}$ and the optimal rate of $h^{1/2}$; we don't know whether this difference is real.

Estimating real rates of convergence for data in $\mathrm{Lip}(\alpha, L^2(\Omega))$ is difficult for many reasons. Even the optimal asymptotic rate of convergence, $O(h^\alpha)$, is quite slow, so one needs very small h to be convinced that the parameter h is in the asymptotic regime. Furthermore, a function in $\mathrm{Lip}(\alpha, L^2(\Omega))$ for $\alpha \le 1/2$ is, in general, not even bounded; if it is the characteristic function of a set, then this set need not have a bounded perimeter. For both these reasons, computing with "generic" $\mathrm{Lip}(\alpha, L^2(\Omega))$ data is quite difficult. We do not have an opinion on what the true rate of convergence might be.

Somewhat weaker results were proved by the first author in [14] for the functional

$$J_h(v^h) = \frac{1}{2}\big(J_{++}(v^h) + J_{--}(v^h)\big).$$

The arguments there exploit the fact that for this particular J_h,

$$|\operatorname{Int} v^h|_{\mathrm{BV}(\Omega)} = J_h(v^h);$$

they also require that $g \in \mathrm{Lip}(\beta, L^1(\Omega)) \cap L^\infty(\Omega)$, which implies that $g \in \mathrm{Lip}(\alpha, L^2(\Omega))$ for $\alpha = \beta/2$, and they achieve the same convergence rate of $h^{\alpha/(2\alpha+2)}$.

Finally, similar techniques have been applied to analyze a central difference approximation to $|v|_{\mathrm{BV}(\Omega)}$ in [11]; there the same convergence rate of approximation $O(h^{1/4})(\alpha = 1)$ was achieved, but for quite smooth functions: g is required to be in the Sobolev space $W^{1,2}(\Omega)$, a space that does not contain "images with edges."

Appendix. We include here the proof of a technical lemma.

Proof of Lemma 2.1. One proves the second inequality simply by combining (1.20) and (1.15).

As for the first inequality, the left-hand side is finite for $v \in L^1(\Omega)$, so if $\operatorname{Ext} v \notin W^{2,1}(2\Omega)$, we are done. So we assume that $\operatorname{Ext} v \in W^{2,1}(2\Omega)$ and we prove (2.1) for $J_h = J_{++}$, the other cases being the same.

We denote $P_h v$ by v^h and use the divided differences $\nabla_x^+ v_i^h$ and $\nabla_y^+ v_i^h$ from (5.4). In the argument that follows, we write v for $\operatorname{Ext} v$.

Then

$$\begin{aligned}
\nabla_x^+ v_i^h - \frac{1}{h^2}\int_{\Omega_i} D_1 v &= \frac{v_{i+(1,0)}^h - v_i^h}{h} - \frac{1}{h^2}\int_{\Omega_i} D_1 v \\
&= \frac{1}{h}\frac{1}{|\Omega_i|}\int_{\Omega_i} [v(x+h,y) - v(x,y)]\,dx\,dy - \frac{1}{h^2}\int_{\Omega_i} D_1 v\,.
\end{aligned}$$

The integrand of the first integral can be rewritten as an integral of $D_1 v$. Then combining these two integrals and once again rewriting the integrand as an integral of the second derivative of v, we have

$$\begin{aligned}
\nabla_x^+ v_i^h - \frac{1}{h^2}\int_{\Omega_i} D_1 v &= \frac{1}{h^3}\int_{\Omega_i}\int_0^h (D_1 v(x+t,y) - D_1 v(x,y))\,dt\,dx\,dy \\
&= \frac{1}{h^3}\int_{\Omega_i}\int_0^h\int_0^t D_1^2 v(x+s,y)\,ds\,dt\,dx\,dy\,.
\end{aligned}$$

Therefore

$$\nabla_x^+ v_i^h = \frac{1}{h^2}\int_{\Omega_i} D_1 v + \frac{1}{h^3}\int_{\Omega_i}\int_0^h\int_0^t D_1^2 v(x+s,y)\,ds\,dt\,dx\,dy\,.$$

Because we can write $\nabla_y^+ v_i^h$ in a similar way, we can bound the norm of $\nabla^+ v_i^h =$

$\begin{pmatrix} \nabla_x^+ v_i^h \\ \nabla_y^+ v_i^h \end{pmatrix}$ by

$$
\begin{aligned}
|\nabla^+ v_i^h| &\le \frac{1}{h^2} \left| \begin{pmatrix} \int_{\Omega_i} D_1 v \\ \int_{\Omega_i} D_2 v \end{pmatrix} \right| + \frac{1}{h^3} \left| \begin{pmatrix} \int_{\Omega_i} \int_0^h \int_0^t D_1^2 v(x+s,y)\, ds\, dt\, dx\, dy \\ \int_{\Omega_i} \int_0^h \int_0^t D_2^2 v(x,y+s)\, ds\, dt\, dx\, dy \end{pmatrix} \right| \\
&\le \frac{1}{h^2} \int_{\Omega_i} |Dv| \\
&\quad + \frac{1}{h^3} \int_{\Omega_i} \int_0^h \int_0^t |D_1^2 v(x+s,y)|\, ds\, dt\, dx\, dy \\
&\quad + \frac{1}{h^3} \int_{\Omega_i} \int_0^h \int_0^t |D_2^2 v(x,y+s)|\, ds\, dt\, dx\, dy\,. \qquad \text{(A.1)}
\end{aligned}
$$

The last line follows from the fact that $\left| \begin{pmatrix} \int f \\ \int g \end{pmatrix} \right| \le \int \sqrt{f^2+g^2}$ (by Jensen's inequality) and $\sqrt{a^2+b^2} \le |a| + |b|$.

To bound the discrete total variation $J_{++}(v^h)$, we sum (A.1) over all indices $i \in \Omega^h$ with weight h^2 at each index. We obtain

$$
J_{++}(v^h) \le \int_\Omega |Dv| + e_x + e_y,
$$

where

$$
\begin{aligned}
e_x &= \sum_i h^2 \frac{1}{h^3} \int_{\Omega_i} \int_0^h \int_0^t |D_1^2 v(x+s,y)|\, ds\, dt\, dx\, dy \\
&\le \frac{1}{h} \int_0^h \int_0^t \left\{ \int_\Omega |D_1^2 v(x+s,y)|\, dx\, dy \right\} ds\, dt \\
&\le \frac{C}{h} \int_0^h \int_0^t \left\{ \int_\Omega |D_1^2 v|\, dx\, dy \right\} ds\, dt \\
&\le Ch \int_\Omega |D_1^2 v|\,.
\end{aligned}
$$

Because we can bound e_y in a similar way, we have

$$
J_{++}(v^h) \le \int_\Omega |Dv| + Ch \int_\Omega (|D_1^2 v| + |D_2^2 v|)\,. \quad \square
$$

REFERENCES

[1] C. Bouman and K. Sauer, *Bayesian estimation of transmission tomograms using segmentation based optimization*, IEEE Trans. Nuclear Science, 39 (1992), pp. 1144–1152.

[2] J. L. Carter, *Dual Methods for Total Variation-Based Image Restoration*, Ph.D. Thesis, 2001, University of California, Los Angeles, Los Angeles, CA.

[3] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97.

[4] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, *Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Trans. Image Process., 7 (1998), pp. 319–335.

[5] A. Chambolle, S. E. Levine, and B. J. Lucier, *An upwind finite-difference method for total variation-based image smoothing*, SIAM J. Imaging Sci., 4 (2011), pp. 277–299.

[6] T. F. Chan, G. H. Golub, and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20 (1999), pp. 1964–1977.

[7] R. A. DeVore, B. Jawerth, and B. J. Lucier, *Image compression through wavelet transform coding*, IEEE Trans. Inform. Theory, 38 (1992), pp. 719–746.

[8] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, Berlin, 1993.

[9] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, FL, 1992.

[10] X. Feng, M. von Oehsen, and A. Prohl, *Rate of convergence of regularization procedures and finite element approximations for the total variation flow*, Numer. Math., 100 (2005), pp. 441–456.

[11] M.-J. Lai, B. Lucier, and J. Wang, *The convergence of a central-difference discretization of Rudin–Osher–Fatemi model for image denoising*, in Scale Space and Variational Methods in Computer Vision, X.-C. Tai et al., eds., Lecture Notes in Comput. Sci. 5567, Springer-Verlag, Berlin, Heidelberg, 2009, pp. 514–526.

[12] S. Osher and J. A. Sethian, *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations*, J. Comput. Phys., 79 (1988), pp. 12–49.

[13] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D., 60 (1992), pp. 259–268.

[14] J. Wang, *Error Bounds for Numerical Methods for the ROF Image Smoothing Model*, Ph.D. Thesis, 2008, Purdue University, West Lafayette, IN.