

# A TOTAL VARIATION DIMINISHING INTERPOLATION OPERATOR AND APPLICATIONS

SÖREN BARTELS, RICARDO H. NOCHETTO, AND ABNER J. SALGADO

ABSTRACT. We construct an interpolation operator that does not increase the total variation and is defined on continuous first degree finite elements over Cartesian meshes for any dimension  $d$  and right triangular meshes for  $d = 2$ . The operator is stable and exhibits second order approximation properties in any  $L^p$ ,  $1 \leq p \leq \infty$ . With the help of it we provide improved error estimates for discrete minimizers of the total variation denoising problem and for total variation flows. We also explore computationally the limitations of the total variation diminishing property over non-Cartesian meshes.

## 1. INTRODUCTION

The approximation of weakly differentiable functions by polynomials is a very useful tool when trying to understand the behavior of such functions. In fact, S.L. Sobolev himself [30] used a kind of averaged Taylor polynomial to discuss equivalent norms in the celebrated spaces that now bear his name. Polynomial approximation also plays a crucial role in numerics, as it is the basis of the analysis of finite element methods and discretization techniques of partial differential equations. It is of no surprise then, that there exists several different constructions and that the approximation properties of such polynomials are well studied. We refer the reader to [11, 12, 13, 14, 15, 16, 26, 29] for a by no means exhaustive list that refers, mostly, to such results in the finite element method context.

The quasitotality of the aforementioned works construct an operator that is well defined in  $L^1(\Omega)$  (see Section 2 for notation) and is stable in  $L^p(\Omega)$  for  $p \geq 1$ . In addition, their approximation properties are studied when the functions belong to certain Sobolev spaces. However, in problems dealing with free discontinuities or connected with geometric features such as minimization of coarea; mean curvature flows and others, it is most natural to work in the space  $BV(\Omega)$  and so the need for an approximation theory by piecewise polynomials in this space arises.

In prior work [7] we dealt with such issue by first regularizing the function and then taking the Lagrange interpolant of the regularization. By properly choosing the regularization parameter as a function of the mesh size, we were able to obtain the properties that served our purposes. However, much simpler arguments and better results would have been possible if a *TV diminishing interpolant* was available. In other words, we wonder about an operator

$$(TVD) \quad \Pi_h : L^1(\Omega) \rightarrow X_h : \quad \int_{\Omega} |\nabla \Pi_h w| \leq \int_{\Omega} |\nabla w|, \quad \forall w \in BV(\Omega),$$

---

2010 *Mathematics Subject Classification.* 65D05; 49M25; 65K15; 65M60; 65N15; 49J40 .

*Key words and phrases.* Total Variation, Interpolation, Approximation, Finite Elements.

This work is partially supported by NSF grants DMS-0807811 and DMS-1109325. AJS is also partially supported by an AMS-Simons grant.

where we denote by  $X_h$  the finite element space; note that the stability constant in (TVD) is 1. The purpose of this work is to construct such an operator under the, rather stringent, geometric assumption that the underlying mesh is Cartesian and the finite element space is made of continuous piecewise linear or multilinear elements. Moreover, by means of computational counterexamples we show that, without this geometric assumption, it is not possible to construct an operator that satisfies (TVD).

We organize our presentation as follows. Notation, functional and discrete spaces are introduced in Section 2. The core of our discussion is Section 3, where we construct our TV diminishing interpolant for periodic functions on quadrilateral elements and prove its main properties. The extension to non-periodic functions is in Section 4 and to simplicial meshes is provided in Section 5. Some applications are discussed in Section 6; namely we provide error estimates for total variation minimization and total variation flows. Finally, in Section 7, we discuss the existence of TV diminishing operators for more general meshes and alternative definitions of the total variation.

## 2. NOTATION AND PRELIMINARIES

In what follows we consider either  $\Omega = S^1 \times S^1$  (the two dimensional torus) or  $\Omega = (0, 1)^2$  (the unit square). Our arguments can be extended without difficulty to more space dimensions. If  $z \in \mathbb{R}^2$  we denote its coordinates by  $z = (z^1, z^2)$ .

We denote by  $L^p(\Omega)$  with  $p \in [1, \infty]$  the space of Lebesgue integrable functions with exponent  $p$ . Sobolev spaces will be denoted by  $W_p^s(\Omega)$ , where  $s \in \mathbb{R}$  is the differentiability order. Whenever  $X$  is a normed space, we denote by  $\|\cdot\|_X$  its norm.

We say that a function  $w \in BV(\Omega)$  (is of bounded variation), if  $w \in L^1(\Omega)$  and its derivative in the sense of distributions is a Radon measure. To simplify notation, we will use  $\nabla w$  to denote such measure. We define the total variation (TV) by

$$(2.1) \quad \text{TV}(w) = \int_{\Omega} |\nabla w|,$$

with  $|\cdot|$  being the  $\ell^1$  norm in  $\mathbb{R}^2$ , and we endow the space  $BV(\Omega)$  with the norm

$$\|w\|_{BV} = \|w\|_{L^1} + \text{TV}(w).$$

We refer the reader to [1, 33] for the most relevant facts about such functions.

**Remark 2.2** (Alternative definitions of the total variation). One could define the total variation using any  $\ell^p$ -norm of the gradient. In other words, for  $p \in [1, \infty]$  we can define

$$(2.3) \quad \text{TV}_p(w) = \int_{\Omega} |\nabla w|_{\ell^p} = \sup \left\{ \int_{\Omega} w \operatorname{div} \mathbf{z} : \mathbf{z} \in \mathcal{C}_0^1(\Omega) : |\mathbf{z}|_{\ell^q} \leq 1, \frac{1}{p} + \frac{1}{q} = 1 \right\},$$

where, for  $\mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d$ ,  $|\mathbf{x}|_{\ell^t}^t = \sum_{k=1}^d |x_k|^t$ ,  $t \in [1, \infty)$  with the usual modification for  $t = \infty$ . Notice that with this notation  $\text{TV} = \text{TV}_1$  which, for instance, is the usual choice in the study of multidimensional nonlinear conservation laws [23, equation (A.19)]. Evidently, any choice of  $p \in [1, \infty]$  in the definition of  $\text{TV}_p$  leads to the same space  $BV(\Omega)$  algebraically and topologically. However, for a given function, the actual value of its total variation might change. The implications of choosing a particular definition of variation will be discussed in more detail in Section 7.

To carry out the finite element discretization, we introduce the number of points in each direction  $N_i \in \mathbb{N}$ ,  $i = 1, 2$ . The mesh size in each direction is

$$h_i = \frac{1}{N_i}, \quad i = 1, 2.$$

The nodes of the mesh are

$$z_{k,l} = (kh_1, lh_2), \quad k = \overline{0, N_1}, \quad l = \overline{0, N_2}.$$

The triangulation

$$\mathcal{T} = \{T_{k,l}\}_{k=0, l=0}^{N_1-1, N_2-1}$$

is made of rectangles

$$T_{k,l} = \{z \in \mathbb{R}^2 : z_{k,l}^1 < z^1 < z_{k+1,l}^1, \quad z_{k,l}^2 < z^2 < z_{k,l+1}^2\}.$$

Clearly  $\bar{\Omega} = \cup_{T \in \mathcal{T}} \bar{T}$ . On the basis of such triangulation we define

$$(2.4) \quad X_h = \{w_h \in C^0(\bar{\Omega}) : w_h|_T \in \mathbb{Q}_1, \quad \forall T \in \mathcal{T}\}.$$

Notice that, given the special structure of the mesh, if we denote by  $\Lambda_{k,l}$  the Lagrange basis function associated with node  $z_{k,l}$ , then

$$\Lambda_{k,l}(z) = \lambda_k^1(z^1) \lambda_l^2(z^2),$$

where  $\lambda_k^i$ ,  $i = 1, 2$ , denotes the one-dimensional Lagrange basis function associated with node  $z_k^i$  on a mesh of size  $h_i$ . We discuss right triangles in Section 5.

### 3. TV DIMINISHING INTERPOLANT: PERIODIC DOMAINS

In this section we construct the TV diminishing interpolation operator in the torus,  $\Omega = S^1 \times S^1$ , and prove some of its approximation properties. It is based on a weighted averaged Taylor polynomial as in [11, 12, 16] but, to better exploit the symmetries related to our mesh, we use a very specific weighting function.

To begin with, let  $\chi_I$  be the characteristic function of the interval  $I$  and set

$$\psi^i = \frac{1}{h_i} \chi_{[-h_i/2, h_i/2]}, \quad i = 1, 2.$$

Notice that  $\int \psi^i = 1$  and  $\|\psi^i\|_{L^\infty} = 1/h_i$ . The weighting function is then

$$(3.1) \quad \psi(z) = \psi^1(z^1) \psi^2(z^2).$$

Notice that

$$(3.2) \quad Q = \text{supp } \psi = \left[-\frac{h_1}{2}, \frac{h_1}{2}\right] \times \left[-\frac{h_2}{2}, \frac{h_2}{2}\right].$$

Given a function  $w \in L^1(\Omega)$ , similarly to [11, 12, 16], we define

$$(3.3) \quad W_{k,l} = \int_Q w(z_{k,l} + \zeta) \psi(\zeta) \, d\zeta \in \mathbb{R}.$$

The interpolant  $\Pi_h w \in X_h$  is the unique function that satisfies

$$(3.4) \quad \Pi_h w(z_{k,l}) = W_{k,l}.$$

The stability of this operator is as follows.

**Lemma 3.5** ( $L^p$ -stability). *The operator  $\Pi_h : L^1(\Omega) \rightarrow X_h$  is well defined and stable in any  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ , namely*

$$(3.6) \quad \|\Pi_h w\|_{L^p} \leq \|w\|_{L^p}, \quad \forall w \in L^p(\Omega).$$

*Proof.* It suffices to show (3.6) for  $p = 1$  and  $p = \infty$ . The intermediate powers will follow by function space interpolation. Estimate (3.6) for  $p = \infty$  is evident. To obtain the  $L^1$ -estimate we consider  $T_{\kappa,\ell} \in \mathcal{T}$  and

$$\|\Pi_h w\|_{L^1(T_{\kappa,\ell})} = \int_{T_{\kappa,\ell}} \left| \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} W_{k,l} \Lambda_{k,l}(z) \right| dz \leq \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} |W_{k,l}| \int_{T_{\kappa,\ell}} \Lambda_{k,l}(z) dz,$$

Notice that, if  $k = \kappa, \kappa + 1$  and  $l = \ell, \ell + 1$ ,

$$\int_{T_{\kappa,\ell}} \Lambda_{k,l}(z) dz = \int_{z_{\kappa,\ell}^1}^{z_{\kappa+1,\ell}^1} \lambda_k^1(z^1) dz^1 \int_{z_{\kappa,\ell}^2}^{z_{\kappa,\ell+1}^2} \lambda_l^2(z^2) dz^2 = \frac{1}{4} h_1 h_2.$$

By definition (3.3),

$$|W_{k,l}| \leq \frac{1}{h_1 h_2} \int_{z_{k,l} + Q} |w(\zeta)| d\zeta.$$

From the considerations given above it follows that

$$\|\Pi_h w\|_{L^1} = \sum_{\kappa=0, \ell=0}^{N_1-1, N_2-1} \|\Pi_h w\|_{L^1(T_{\kappa,\ell})} \leq \frac{1}{4} \sum_{\kappa=0, \ell=0}^{N_1-1, N_2-1} \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} \int_{z_{k,l} + Q} |w(\zeta)| d\zeta,$$

To conclude it suffices to notice that every vertex belongs to exactly four cells.  $\square$

The main result of this contribution is the following.

**Theorem 3.7** (TV diminishing: periodic case). *The operator  $\Pi_h : L^1(\Omega) \rightarrow X_h$  defined in (3.4) satisfies (TVD).*

*Proof.* Notice that, thanks to the particular structure of the Cartesian mesh  $\mathcal{T}$ , the basis functions satisfy

$$\partial_1 \Lambda_{k,l} = (\lambda_k^1)' \lambda_l^2 = \pm \frac{1}{h_1} \lambda_l^2.$$

If  $w \in \mathcal{C}^1(\Omega)$  and  $T = T_{k,l} \in \mathcal{T}$ , then

$$\begin{aligned} \partial_1 \Pi_h w|_T &= W_{k,l} \left( \frac{-1}{h_1} \right) \lambda_l^2 + W_{k+1,l} \left( \frac{1}{h_1} \right) \lambda_l^2 \\ (3.8) \quad &+ W_{k,l+1} \left( \frac{-1}{h_1} \right) (1 - \lambda_l^2) + W_{k+1,l+1} \left( \frac{1}{h_1} \right) (1 - \lambda_l^2) \\ &= \frac{1}{h_1} (W_{k+1,l} - W_{k,l}) \lambda_l^2 + \frac{1}{h_1} (W_{k+1,l+1} - W_{k,l+1}) (1 - \lambda_l^2), \end{aligned}$$

where we have used that, for every  $l$ , we have  $\lambda_l^2 + \lambda_{l+1}^2 \equiv 1$  on  $[z_{k,l}^2, z_{k,l+1}^2]$ . In view of the definition of  $W_{k,l}$ , given in (3.3), we can write

$$\begin{aligned} W_{k+1,l} - W_{k,l} &= \int_Q (w(z_{k+1,l} + \zeta) - w(z_{k,l} + \zeta)) \psi(\zeta) d\zeta \\ &= h_1 \int_Q \int_0^1 \partial_1 w(s z_{k+1,l} + (1-s) z_{k,l} + \zeta) \psi(\zeta) ds d\zeta. \end{aligned}$$

Notice that, at this point, it was essential to have a  $\mathcal{C}^1$  function. Notice also that

$$\int_T \lambda_l^2(z^2) dz = \int_T (1 - \lambda_{l+1}^2(z^2)) dz = \frac{1}{2} |T| = \frac{1}{2} h_1 h_2.$$

With these observations we estimate the local variation in the first coordinate direction of the interpolant as

$$\begin{aligned} \int_T |\partial_1 \Pi_h w(z)| dz &\leq \frac{1}{2} h_1 h_2 \int_Q \int_0^1 |\partial_1 w(s z_{k+1,l} + (1-s) z_{k,l} + \zeta)| \psi(\zeta) ds d\zeta \\ &\quad + \frac{1}{2} h_1 h_2 \int_Q \int_0^1 |\partial_1 w(s z_{k+1,l+1} + (1-s) z_{k,l+1} + \zeta)| \psi(\zeta) ds d\zeta. \end{aligned}$$

The total variation in the first coordinate direction is then

$$\begin{aligned} \int_\Omega |\partial_1 \Pi_h w| &= \sum_{l=0}^{N_2-1} \sum_{k=0}^{N_1-1} \int_{T_{k,l}} |\partial_1 \Pi_h w(z)| dz \\ &\leq \frac{1}{2} h_1 h_2 \int_Q \sum_{l=0}^{N_2-1} \sum_{k=0}^{N_1-1} \int_0^1 (|\partial_1 w(s z_{k+1,l} + (1-s) z_{k,l} + \zeta)| \\ &\quad + |\partial_1 w(s z_{k+1,l+1} + (1-s) z_{k,l+1} + \zeta)|) \psi(\zeta) ds d\zeta. \end{aligned}$$

Introduce the change of variables  $(0, 1) \ni s \mapsto \sigma_k \in (z_{k,l}^1 + \zeta^1, z_{k+1,l}^1 + \zeta^1)$  given by

$$\sigma_k = s z_{k+1,l}^1 + (1-s) z_{k,l}^1 + \zeta^1,$$

and notice that with it we can add over  $k$  to obtain

$$\begin{aligned} \int_\Omega |\partial_1 \Pi_h w| &\leq \frac{1}{2} h_2 \int_Q \sum_{l=0}^{N_2-1} \left( \int_0^1 |\partial_1 w(\sigma, \zeta^2 + l h_2)| d\sigma \right. \\ &\quad \left. + \int_0^1 |\partial_1 w(\sigma, \zeta^2 + (l+1) h_2)| d\sigma \right) \psi(\zeta) d\zeta \\ &\leq h_2 \int_{-h_1/2}^{h_1/2} \psi^1(\zeta^1) d\zeta^1 \sum_{l=0}^{N_2-1} \int_{-h_2/2}^{h_2/2} \int_0^1 |\partial_1 w(\sigma, \zeta^2 + l h_2)| \psi^2(\zeta^2) d\sigma d\zeta^2, \end{aligned}$$

where, to arrive at the last inequality, we have used the definition of  $\psi$  given in (3.1). Finally, we recall that  $\|\psi^2\|_{L^\infty} = \frac{1}{h_2}$ ,  $\int \psi^1 = 1$ , and

$$\sum_{l=0}^{N_2-1} \int_{-h_2/2}^{h_2/2} \int_0^1 |\partial_1 w(\sigma, \zeta^2 + l h_2)| d\sigma d\zeta^2 = \int_\Omega |\partial_1 w|,$$

to conclude the proof for a  $\mathcal{C}^1$  function  $w$  and definition (2.1) of TV.

In general, i.e.,  $w \in BV(\Omega) \setminus \mathcal{C}^1(\Omega)$ , we use an approximation argument. We recall [33, Theorem 5.3.3] that smooth functions are dense in  $BV(\Omega)$  under strict convergence. In other words, there is a sequence  $\{w_n\}_{n \in \mathbb{N}} \subset \mathcal{C}^\infty(\Omega)$  such that

$$\lim_{n \rightarrow \infty} \|w_n - w\|_{L^1} = 0, \quad \limsup_{n \rightarrow \infty} \text{TV}(w_n) \leq \text{TV}(w).$$

Since  $\text{TV}(\Pi_h w_n) \leq \text{TV}(w_n)$  for all  $n \in \mathbb{N}$ , we deduce

$$\text{TV}(\Pi_h w) = \lim_{n \rightarrow \infty} \text{TV}(\Pi_h w_n) \leq \liminf_{n \rightarrow \infty} \text{TV}(w_n) \leq \text{TV}(w),$$

and thus conclude the proof.  $\square$

The approximation properties of this operator are summarized in the following.

**Theorem 3.9** (Approximation). *Let  $h = \max\{h_i : i = 1, 2\}$ . If  $w \in BV(\Omega)$ , then*

$$(3.10) \quad \|w - \Pi_h w\|_{L^1} \leq ch \, \mathbf{TV}(w),$$

*with  $c > 0$  a geometric constant. If  $w \in BV(\Omega) \cap L^p(\Omega)$  with  $p > 1$ , then for all  $q \in [1, p]$ ,*

$$(3.11) \quad \|w - \Pi_h w\|_{L^q} \leq ch^{1-s} \mathbf{TV}(w)^{1-s} \|w\|_{L^p}^s, \quad \frac{1}{q} = \frac{1-s}{1} + \frac{s}{p}.$$

*If  $w \in W_p^2(\Omega)$  with  $1 \leq p \leq \infty$ , then*

$$(3.12) \quad \|w - \Pi_h w\|_{L^p} \leq ch^2 |w|_{W_p^2}.$$

*Proof.* For (3.12) we use the symmetries of the mesh and the averaging procedure. In other words, if  $w \in \mathbb{P}_1$ , then  $w(z_{k,l}) = W_{k,l}$  and  $\Pi_h w = w$ . This, together with an argument *à la* Bramble-Hilbert, implies the result.

Estimate (3.11) follows from (3.6) and (3.10) using the well-known function space interpolation inequality [31]

$$\|w\|_{L^q} \leq \|w\|_{L^1}^{1-s} \|w\|_{L^p}^s, \quad \frac{1}{q} = \frac{1-s}{1} + \frac{s}{p}.$$

We only sketch the proof of the first statement, as the arguments are similar to those used to obtain Theorem 3.7. We write

$$\|w - \Pi_h w\|_{L^1} = \sum_{\kappa, \ell=0}^{N_1-1, N_2-1} \int_{T_{\kappa, \ell}} |w - \Pi_h w|,$$

and estimate the local differences as

$$|w - \Pi_h w| \leq |w - W_{k_0, l_0}| + |W_{k_0, l_0} - \Pi_h w|,$$

where  $(k_0, l_0)$  is such that  $z_{k_0, l_0}$  is a vertex of  $T_{\kappa, \ell}$ . Provided that  $w \in \mathcal{C}^1(\Omega)$  the bound on the first term is immediate for if  $z \in T_{\kappa, \ell}$ ,

$$\begin{aligned} |w(z) - W_{k_0, l_0}| &= \left| \int (w(z) - w(z_{k_0, l_0} + \zeta)) \psi(\zeta) \, d\zeta \right| \\ &\leq ch \int_Q \int_0^1 |\nabla w(sz + (1-s)(z_{k_0, l_0} + \zeta))| \psi(\zeta) \, ds \, d\zeta, \end{aligned}$$

so that

$$\sum_{\kappa, \ell=0}^{N_1-1, N_2-1} \int_{T_{\kappa, \ell}} |w - W_{k_0, l_0}| \leq ch \int_{\Omega} |\nabla w|,$$

because  $\|\psi\|_{L^\infty} = (h_1 h_2)^{-1} = |T_{\kappa, \ell}|^{-1}$ . For the second term we recall that the basis functions form a partition of unity, whence

$$\begin{aligned} |W_{k_0, l_0} - \Pi_h w(z)| &= \left| W_{k_0, l_0} \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} \Lambda_{k, l}(z) - \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} W_{k, l} \Lambda_{k, l}(z) \right| \\ &\leq \sum_{k=\kappa, l=\ell}^{\kappa+1, \ell+1} \Lambda_{k, l}(z) |W_{k, l} - W_{k_0, l_0}|, \end{aligned}$$

Using, again, the differentiability properties of  $w$ ,

$$|W_{k, l} - W_{k_0, l_0}| \leq h \int_Q \int_0^1 |\nabla w(sz_{k, l} + (1-s)(z_{k_0, l_0} + \zeta))| \, ds \, \psi(\zeta) \, d\zeta,$$

which, employing the facts that  $|\Lambda_{k,l}(z)| \leq 1$  and  $\|\psi\|_{L^\infty} = (h_1 h_2)^{-1}$ , implies

$$\sum_{\kappa, \ell=0}^{N_1-1, N_2-1} \int_{T_{\kappa, \ell}} |\Pi_h w - W_{k_0, l_0}| \leq ch \int_{\Omega} |\nabla w|.$$

The proof concludes via a density argument with  $\mathcal{C}^1$  functions.  $\square$

#### 4. TV DIMINISHING INTERPOLANT: NON-PERIODIC DOMAINS

Let us describe two possible ways of modifying our construction in the case  $\Omega = (0, 1)^2$ . Both constructions will satisfy (TVD) and possess optimal approximation properties in the space  $BV(\Omega)$ , so that all the results presented in Section 6 still hold in this case. Property (3.12) however, will no longer be valid in general.

**4.1. Interpolation based on homotetic transformations of the domain.** Set  $\epsilon = \frac{1}{2} \max\{h_1, h_2\}$ , we define

$$\Omega_\epsilon = (-\epsilon, 1 + \epsilon)^2$$

and notice that  $\Omega \Subset \Omega_\epsilon$ . For a function  $w \in BV(\Omega)$  we define

$$w_\epsilon(y) = \frac{1}{1 + 2\epsilon} w\left(\frac{y + \epsilon}{1 + 2\epsilon}\right), \quad y \in \Omega_\epsilon.$$

Notice that, employing the change of variables  $x = (y + \epsilon)/(1 + 2\epsilon)$  valid for functions of bounded variation [1, Theorem 3.16], we obtain

$$\begin{aligned} \text{TV}_{\Omega_\epsilon}(w_\epsilon) &= \frac{1}{(1 + 2\epsilon)^2} \int_{\Omega_\epsilon} \left| \nabla w\left(\frac{y + \epsilon}{1 + 2\epsilon}\right) \right| dy = \frac{(1 + 2\epsilon)^2}{(1 + 2\epsilon)^2} \int_{\Omega} |\nabla w(x)| dx \\ &= \text{TV}_{\Omega}(w), \end{aligned}$$

where by  $\text{TV}_A(w)$  we have denoted the total variation of the function  $w$  over the set  $A$ . Clearly  $w_\epsilon|_{\Omega} \in BV(\Omega)$ .

We define  $\mathcal{I}_h w = \Pi_h w_\epsilon$ , where the operator  $\Pi_h$  is defined as in Section 3. Notice that the averaging procedure necessary for the definition of  $\Pi_h w_\epsilon$  makes sense, since the function  $w_\epsilon$  is defined in  $\Omega_\epsilon$ . The results of Section 3 then imply

$$\text{TV}_{\Omega}(\mathcal{I}_h w) = \text{TV}_{\Omega}(\Pi_h w_\epsilon) \leq \text{TV}_{\Omega_\epsilon}(w_\epsilon) = \text{TV}_{\Omega}(w).$$

The approximation properties (3.10) and (3.11) of  $\mathcal{I}_h$  can be obtained similarly.

**Remark 4.1** (Second order approximation). The approximation property (3.12) cannot hold for  $\mathcal{I}_h$ . In fact if  $w \equiv c$ , then  $\mathcal{I}_h w = c/(1 + 2\epsilon)$ . This also shows that the operator does not preserve constants, which might be an undesirable feature.

**4.2. Interpolation based on rescaled local averages.** The construction of §4.1, although simple, does not preserve constants. For this reason we consider a slightly more complicated procedure. It is a modification of the construction of Section 3 to take into account the effects of the boundary.

The operator  $\mathcal{C}_h : L^1(\Omega) \ni w \mapsto \mathcal{C}_h w \in X_h$  is then such that if  $z_{k,l}$  is an internal node, the value of the interpolant is as in (3.4). On the other hand, if  $z_{k,l}$  is a boundary node, we define

$$Q_{\Omega}(k, l) = (z_{k,l} + Q) \cap \Omega,$$

and

$$(4.2) \quad W_{k,l} = \frac{1}{|Q_\Omega(k,l)|} \int_{Q_\Omega(k,l)} w(y) dy.$$

Observe that,  $|Q|/|Q_\Omega(k,l)| = 4$  for vertices of  $\Omega$  and  $|Q|/|Q_\Omega(k,l)| = 2$  for the rest of boundary nodes. This rescaling allows us to take into account the lack of symmetry on the interaction with neighboring nodes.

**Theorem 4.3** (TV diminishing; non-periodic case). *The operator  $\mathcal{C}_h : L^1(\Omega) \rightarrow X_h$  defined by  $\mathcal{C}_h w(z_{k,l}) = W_{k,l}$  with  $W_{k,l}$  as in (3.3) for  $k = \overline{1, N_1 - 1}$ ,  $l = \overline{1, N_2 - 1}$  and by (4.2) otherwise is nonexpansive in  $L^1(\Omega)$ , i.e.,*

$$(4.4) \quad \|\mathcal{C}_h w\|_{L^1} \leq \|w\|_{L^1},$$

and, moreover, satisfies (TVD).

*Proof.* We split  $Q$  into four congruent rectangles

$$\begin{aligned} Q^{++} &= \{z \in Q : z^1 > 0, z^2 > 0\}, & Q^{+-} &= \{z \in Q : z^1 > 0, z^2 < 0\}, \\ Q^{-+} &= \{z \in Q : z^1 < 0, z^2 > 0\}, & Q^{--} &= \{z \in Q : z^1 < 0, z^2 < 0\}, \end{aligned}$$

and introduce

$$R^{-+} = Q^{-+} \cup Q^{++}, \quad R^{+-} = Q^{+-} \cup Q^{--}.$$

It is necessary to compute the contribution of the cells that intersect the boundary. To illustrate the procedure we consider  $T_{0,0}$  and recall (3.8) to write:

$$\begin{aligned} \frac{2}{h_2} \int_{T_{0,0}} |\partial_1 \mathcal{C}_h w| &\leq \left| \frac{1}{|R^{-+}|} \int_{R^{-+}} w(z_{1,0} + \zeta) d\zeta - \frac{1}{|Q^{++}|} \int_{Q^{++}} w(z_{0,0} + \zeta) d\zeta \right| \\ &\quad + \left| \frac{1}{|Q|} \int_Q w(z_{1,1} + \zeta) d\zeta - \frac{1}{|R^{+-}|} \int_{R^{+-}} w(z_{0,1} + \zeta) d\zeta \right|. \end{aligned}$$

The first two terms on the right hand side can be estimated as

$$\begin{aligned} &\left| \frac{2}{h_1 h_2} \int_{R^{-+}} w(z_{1,0} + \zeta) d\zeta - \frac{4}{h_1 h_2} \int_{Q^{++}} w(z_{0,0} + \zeta) d\zeta \right| \\ &\leq \frac{2}{h_2} \int_{Q^{++}} \int_0^1 |\partial_1 w(s z_{0,1} + (1-s) z_{0,0} + \zeta)| d\zeta ds \\ &\quad + \frac{2}{h_2} \int_{Q^{-+}} \int_0^1 |\partial_1 w(s z_{1,0} + (1-s)(z_{0,0} + \frac{h_1}{2} \mathbf{e}_1 + \zeta))| d\zeta ds. \end{aligned}$$

The summation procedure is now as before. For this it suffices to notice that  $|Q^{++}| = |Q^{+-}| = |Q^{-+}| = |Q^{--}| = \frac{1}{4}|Q|$ .  $\square$

The operator  $\mathcal{C}_h$  is  $L^p$ -stable as in (3.6). Its approximation properties are as in (3.10) and (3.11). For brevity we skip the proofs, as they repeat the arguments we have already presented. Property (3.12) however, will not be valid in general.

**Remark 4.5** (Second order approximation). Denote

$$\mathcal{N}_\varepsilon = \{x \in \Omega : \text{dist}(x, \partial\Omega) < \varepsilon\}.$$

The proof of Theorem 4.3 shows that if  $w \in W_\infty^1(\mathcal{N}_\varepsilon)$  and  $\frac{3}{2}h \leq \varepsilon$ , then

$$\|w - \mathcal{C}_h w\|_{L^\infty(\mathcal{N}_h)} \leq ch|w|_{W_\infty^1(\mathcal{N}_{3h/2})}.$$

Therefore, since  $|\mathcal{N}_h| \leq ch$ , we deduce

$$\|w - \mathcal{C}_h w\|_{L^1} \leq ch^2.$$

**Remark 4.6** (Positivity preserving interpolation). In light of the results of [26], the construction of a TVD interpolant with second order approximation properties on the whole domain seems impossible. Indeed, if one wants to construct a *linear* interpolation operator based on weighted means that satisfies (TVD), then this operator is *positivity preserving*. Indeed, if  $w \geq 0$  we obtain (see (3.3) and (4.2)) that  $W_{k,l} \geq 0$  and this implies that

$$\mathcal{C}_h w = \sum_{k=0, l=0}^{N_1, N_2} W_{k,l} \Lambda_{k,l} \geq 0,$$

since  $\Lambda_{k,l} \geq 0$ . Then, as [26, Corollaries 4–6] show, this operator can at most be first order accurate, unless we restrict the domain of  $\mathcal{C}_h$  to be  $BV_0(\Omega)$  (i.e., functions of bounded variation with vanishing trace), in which case we define  $\mathcal{C}_h w(x) = 0$  for all  $x \in \partial\Omega$ , or periodic functions in  $BV(\Omega)$ .

## 5. RIGHT TRIANGULAR MESHES

Let us now extend our construction to a particular two dimensional case of triangular meshes. We will work in the two dimensional torus  $\Omega = S^1 \times S^1$ . We start with the triangulation  $\mathcal{T} = \{T_{k,l}\}_{k=0, l=0}^{N_1-1, N_2-1}$  and subdivide each rectangle into two by means of the line that connects the nodes  $z_{k,l}$  and  $z_{k+1, l+1}$ . With this procedure we obtain two triangles

$$T_{k,l}^- = \text{conv}\{z_{k,l}, z_{k, l+1}, z_{k+1, l+1}\}, \quad T_{k,l}^+ = \text{conv}\{z_{k,l}, z_{k+1, l}, z_{k+1, l+1}\},$$

where  $\text{conv}$  denotes the convex hull. We collect the so obtained triangles in

$$\tilde{\mathcal{T}} = \{T_{k,l}^r : 0 \leq k \leq N_1 - 1, 0 \leq l \leq N_2 - 1, r \in \{+, -\}\}.$$

On the basis of this triangulation we define the finite element space

$$X_h = \left\{ w_h \in C^0(\bar{\Omega}) : w_h|_{\tilde{T}} \in \mathbb{P}_1, \forall \tilde{T} \in \tilde{\mathcal{T}} \right\}.$$

Finally, the operator  $\tilde{\Pi}_h : L^1(\Omega) \rightarrow X_h$  maps a given function  $w \in L^1(\Omega)$  to the unique piecewise  $\mathbb{P}_1$  function that takes as nodal values the local averages, namely

$$(5.1) \quad \tilde{\Pi}_h w(z_{k,l}) = W_{k,l}.$$

The main properties of this operator are summarized in the following result.

**Theorem 5.2** (TV diminishing: triangular meshes). *The operator  $\tilde{\Pi}_h$ , defined by (5.1) is well defined and stable in any  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ , namely*

$$(5.3) \quad \|\tilde{\Pi}_h w\|_{L^p} \leq \|w\|_{L^p}, \quad \forall w \in L^p(\Omega).$$

*Moreover, the operator  $\tilde{\Pi}_h$  satisfies (TVD) and possesses the following approximation properties: If  $w \in BV(\Omega)$ , then*

$$(5.4) \quad \|w - \tilde{\Pi}_h w\|_{L^1} \leq ch \text{TV}(w).$$

*For any  $p \in [1, \infty]$ , if  $w \in BV(\Omega) \cap L^p(\Omega)$ , then*

$$(5.5) \quad \|w - \tilde{\Pi}_h w\|_{L^q} \leq ch^{1-s} \text{TV}(w)^{1-s} \|w\|_{L^p}^s, \quad \frac{1}{q} = \frac{1-s}{1} + \frac{s}{p}.$$

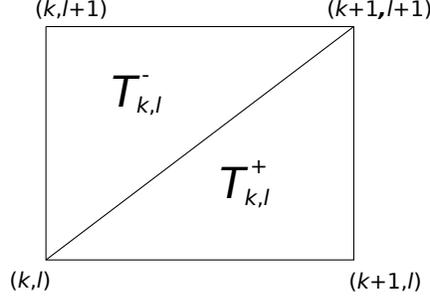


FIGURE 5.1. Two triangles of a mesh consisting of right simplices and their corresponding labels.

For any  $p \in [1, \infty]$ , if  $w \in W_p^2(\Omega)$ , then

$$(5.6) \quad \|w - \tilde{\Pi}_h w\|_{L^p} \leq ch^2 |w|_{W_p^2}.$$

*Proof.* We will just sketch the proof of (TVD), the other results follow *mutatis mutandis* the corresponding statements in Lemma 3.5 and Theorem 3.9.

Notice, first of all, that all considerations can be reduced to  $T_{k,l}^+ \cup T_{k,l}^-$  and that, in this patch of elements, we have

$$\begin{aligned} \nabla \Lambda_{k,l} &= \begin{cases} \left(-\frac{1}{h_1}, 0\right)^\top & \text{in } T_{k,l}^+, \\ \left(0, -\frac{1}{h_2}\right)^\top & \text{in } T_{k,l}^-, \end{cases} & \nabla \Lambda_{k,l+1} &= \begin{cases} 0 & \text{in } T_{k,l}^+, \\ \left(-\frac{1}{h_1}, -\frac{1}{h_2}\right)^\top & \text{in } T_{k,l}^-, \end{cases} \\ \nabla \Lambda_{k+1,l} &= \begin{cases} \left(\frac{1}{h_1}, -\frac{1}{h_2}\right)^\top & \text{in } T_{k,l}^+, \\ 0 & \text{in } T_{k,l}^-, \end{cases} & \nabla \Lambda_{k+1,l+1} &= \begin{cases} \left(0, \frac{1}{h_2}\right)^\top & \text{in } T_{k,l}^+, \\ \left(\frac{1}{h_1}, 0\right)^\top & \text{in } T_{k,l}^-, \end{cases} \end{aligned}$$

see Figure 5.1. This yields

$$\partial_1 \tilde{\Pi}_h w = \begin{cases} \frac{1}{h_1} (W_{k+1,l} - W_{k,l}) & \text{in } T_{k,l}^+, \\ \frac{1}{h_1} (W_{k+1,l+1} - W_{k,l+1}) & \text{in } T_{k,l}^-. \end{cases}$$

It remains then to notice that, again,

$$\begin{aligned} W_{k+1,l} - W_{k,l} &= \int_Q (w(z_{k+1,l} + \zeta) - w(z_{k,l} + \zeta)) \psi(\zeta) \, d\zeta \\ &= h_1 \int_Q \int_0^1 \partial_1 w(s z_{k+1,l} + (1-s) z_{k,l} + \zeta) \psi(\zeta) \, ds \, d\zeta, \end{aligned}$$

with a similar expression for  $W_{k+1,l+1} - W_{k,l+1}$ . This yields

$$\begin{aligned} \int_{T_{k,l}^+} |\partial_1 \tilde{\Pi}_h w| &\leq \frac{h_1 h_2}{2} \int_Q \int_0^1 |\partial_1 w(s z_{k+1,l} + (1-s) z_{k,l} + \zeta)| \psi(\zeta) \, d\zeta, \\ \int_{T_{k,l}^-} |\partial_1 \tilde{\Pi}_h w| &\leq \frac{h_1 h_2}{2} \int_Q \int_0^1 |\partial_1 w(s z_{k+1,l+1} + (1-s) z_{k,l+1} + \zeta)| \psi(\zeta) \, d\zeta. \end{aligned}$$

The rest of the proof proceeds as in Theorem 3.7.  $\square$

## 6. APPLICATIONS

In this section we present two applications of the operators constructed in Sections 3, 4 and 5. The fact that the operators satisfy (TVD) and (3.10) allows us to not only improve on existing results but also to have simpler arguments.

**6.1. Total variation minimization.** Let  $\alpha > 0$  and  $f \in L^\infty(\Omega)$ . Set

$$(6.1) \quad E(w) = \text{TV}(w) + \frac{\alpha}{2} \|w - f\|_{L^2}^2,$$

and consider the problem of finding  $\mathbf{v} \in BV(\Omega)$  that minimizes this functional. This problem arises, for instance, in connection with image processing [27]. Paper [6] presents a first order finite element method for the approximation of minimizers of (6.1) as well as an iterative scheme for their computation, thereby extending prior work of A. Chambolle and collaborators (see e.g. [9, 10]) and B. Lucier and J. Wang [32]. The main convergence result of [6] reads as follows: Let

$$(6.2) \quad \mathbf{v} = \underset{w \in BV(\Omega)}{\text{argmin}} E(w),$$

and

$$(6.3) \quad v_h = \underset{w_h \in X_h}{\text{argmin}} E(w_h).$$

If  $\mathbf{v} \in B_\infty^\beta(L^2(\Omega))$ , with  $\beta \in (0, 1]$ , then

$$(6.4) \quad \|\mathbf{v} - v_h\|_{L^2} \leq ch^{\frac{\beta}{2(1+\beta)}}.$$

Let us, with the help our TV diminishing operator, improve on this result, which is at best  $\mathcal{O}(h^{1/4})$ .

**Theorem 6.5** (Total variation minimization). *Let  $\mathbf{v}$  be as in (6.2) and  $v_h$  as in (6.3). Assume that  $\mathbf{v} \in BV(\Omega) \cap L^\infty(\Omega)$ . If there exists an operator  $\Pi_h : L^1(\Omega) \rightarrow X_h$  that is stable in any  $L^p(\Omega)$ , has first order approximation properties in  $L^1(\Omega)$  and satisfies (TVD), then*

$$\|\mathbf{v} - v_h\|_{L^2} \leq \sqrt{2}h^{1/2} (\|\mathbf{v}\|_{L^\infty} + \|f\|_{L^\infty})^{1/2} \text{TV}(\mathbf{v})^{1/2}.$$

*Proof.* Owing to the strict convexity of  $E$  and the fact that  $v_h$  is a minimizer over  $X_h$ , we have

$$\begin{aligned} \frac{\alpha}{2} \|\mathbf{v} - v_h\|_{L^2}^2 &\leq E(v_h) - E(\mathbf{v}) \leq E(\Pi_h \mathbf{v}) - E(\mathbf{v}) \\ &= \text{TV}(\Pi_h \mathbf{v}) - \text{TV}(\mathbf{v}) + \frac{\alpha}{2} (\|\Pi_h \mathbf{v} - f\|_{L^2}^2 - \|\mathbf{v} - f\|_{L^2}^2). \end{aligned}$$

Since  $\Pi_h$  satisfies (TVD) and is stable in  $L^\infty(\Omega)$ ,

$$\begin{aligned} \|\mathbf{v} - v_h\|_{L^2}^2 &\leq \langle \Pi_h \mathbf{v} - \mathbf{v}, \Pi_h \mathbf{v} + \mathbf{v} - 2f \rangle \leq 2\|\Pi_h \mathbf{v} - \mathbf{v}\|_{L^1} (\|\mathbf{v}\|_{L^\infty} + \|f\|_{L^\infty}) \\ &\leq 2h (\|\mathbf{v}\|_{L^\infty} + \|f\|_{L^\infty}) \text{TV}(\mathbf{v}), \end{aligned}$$

where, to obtain the last inequality, we used the first order approximation properties of  $\Pi_h$  in  $L^1(\Omega)$ .  $\square$

**Remark 6.6** (Reduced regularity). The conclusion of Theorem 6.5 is indeed an improvement over (6.4). It is not difficult to show that  $BV(\Omega) \cap L^\infty(\Omega) \hookrightarrow B_\infty^{1/2}(L^2(\Omega))$ ; see [31, Lemma 38.1] for a proof and, in some sense, the converse statement. In this case (6.4) yields a  $\mathcal{O}(h^{1/6})$  error estimate. Our results yields a rate which, given the smoothness of  $\mathbf{v}$ , is optimal in  $L^2(\Omega)$ .

**Remark 6.7** (First order accuracy). Let  $W$  be the space of functions in  $L^\infty(\Omega)$  whose gradients are in  $BV(\Omega)$ . A direct calculation leads to  $W \hookrightarrow B_\infty^1(L^2(\Omega))$ . Therefore, if  $\mathbf{v} \in W$  and  $\Pi_h$  is second order accurate in  $L^1(\Omega)$  (see (3.12) and (5.6)), then the last step in the proof of Theorem 6.5 yields the optimal rate

$$\|\mathbf{v} - v_h\|_{L^2} \leq ch (\|\mathbf{v}\|_{L^\infty} + \|f\|_{L^\infty})^{1/2} \text{TV}(\nabla \mathbf{v})^{1/2}.$$

**6.2. TV flows.** The equation

$$(6.8) \quad \mathbf{u}_t = \text{div} \left( \frac{\nabla \mathbf{u}}{|\nabla \mathbf{u}|} \right),$$

supplemented with suitable initial and boundary conditions is known as the *TV flow*, since it can be interpreted as the  $L^2$ -subgradient flow of the functional  $\text{TV}$ . We refer the reader to [2, 8, 21, 22] for the analysis of this problem and to [7, 17, 18] for its discretization. Concerning applications this problem, supplemented with a smooth strictly convex  $p$ -Laplacian term, arises, e.g., in the modeling of grain boundary motion [24]; facet formation and evolution [19]; and electromigration [20].

The precise mathematical meaning of (6.8) is as a subgradient flow. In other words, given  $\mathbf{u}_0 \in L^2(\Omega)$ , we seek a function  $\mathbf{u} : [0, T] \rightarrow L^2(\Omega) \cap BV(\Omega)$  such that

$$(6.9) \quad \begin{cases} \langle \mathbf{u}_t, \mathbf{u} - w \rangle + \text{TV}(\mathbf{u}) - \text{TV}(w) \leq 0, & \forall w \in L^2(\Omega) \cap BV(\Omega), \\ \mathbf{u}|_{t=0} = \mathbf{u}_0. \end{cases}$$

In prior work [7], we provided an error analysis for discretizations of (6.9) as well as a convergent solution scheme for it. We considered the fully discrete scheme: Find  $\{u_h^k\}_{k=0}^K \subset X_h$  such that

$$(6.10) \quad \begin{cases} \left\langle \frac{u_h^{k+1} - u_h^k}{\Delta t}, u_h^{k+1} - w_h \right\rangle + \text{TV}(u_h^{k+1}) - \text{TV}(w_h) \leq 0, & \forall w_h \in X_h, \\ u_h^0 = \mathcal{I}_h \mathbf{u}_0. \end{cases}$$

where  $\Delta t > 0$  is a time-step,  $K = [T/\Delta t]$  and  $\mathcal{I}_h$  is some interpolation operator. Our main convergence result stated that if  $\mathbf{u}_0 \in L^\infty(\Omega) \cap BV(\Omega)$  and  $\text{TV}(\mathcal{I}_h \mathbf{u}_0) \leq c < \infty$  uniformly in  $h$ , then

$$\|\mathbf{u} - \widehat{u}_h\|_{L^\infty(L^2)} \leq \|\mathbf{u}_0 - \mathcal{I}_h \mathbf{u}_0\|_{L^2} + c \left( \Delta t^{1/2} + h^{1/6} \right),$$

where  $\widehat{u}_h$  denotes the piecewise linear (in time) interpolant of the sequence  $\{u_h^k\}$ . This result, as [25, 28] show, is optimal with respect to time, given the regularity of the initial data.

Let us, with the help of the TV diminishing operator of Sections 3, 4 and 5 improve the convergence rate in space.

**Theorem 6.11** (Convergence of TV flows). *Let  $\mathbf{u}$  solve (6.9). Assume that the initial data  $\mathbf{u}_0 \in L^\infty(\Omega) \cap BV(\Omega)$ . Assume there exists an operator  $\Pi_h : L^1(\Omega) \rightarrow$*

$X_h$  that is stable in any  $L^p(\Omega)$ , has first order approximation properties in  $L^1(\Omega)$  and satisfies (TVD). If  $\{u_h^k\}_{k=0}^K \subset X_h$  solves (6.10), with  $u_h^0 = \Pi_h u_0$ , then

$$(6.12) \quad \|u - \widehat{u}_h\|_{L^\infty(L^2)} \leq c \left( \Delta t^{1/2} + h^{1/4} \right).$$

*Proof.* The error of the time discretization is obtained with standard techniques, cf. [25, 28]. For this reason we concentrate on the error of space discretization and compare the solution of (6.10) with the sequence  $\{u^k\}_{k=0}^K \subset L^2(\Omega) \cap BV(\Omega)$ , solution of the semidiscrete problem

$$\begin{cases} \left\langle \frac{u^{k+1} - u^k}{\Delta t}, u^{k+1} - w \right\rangle + \text{TV}(u^{k+1}) - \text{TV}(w) \leq 0, & \forall w \in L^2(\Omega) \cap BV(\Omega), \\ u^0 = u_0. \end{cases}$$

Set  $w = u_h^{k+1}$  in this inequality and  $w_h = \Pi_h u^{k+1}$  in (6.10) and add them. Denoting  $e^k = u^k - u_h^k$ , we obtain

$$(6.13) \quad \|e^{k+1}\|_{L^2}^2 + \|e^{k+1} - e^k\|_{L^2}^2 \leq \|e^k\|_{L^2}^2 + 2\Delta t \left\| \frac{u^{k+1} - u_h^k}{\Delta t} \right\|_{L^2} \|u^{k+1} - \Pi_h u^{k+1}\|_{L^2}$$

where we used (TVD) to cancel the TV terms.

Setting  $w_h = u_h^k$  in (6.10) is not difficult to show that

$$\frac{1}{\Delta t} \sum_{k=0}^K \|u_h^{k+1} - u_h^k\|_{L^2}^2 \leq c \text{TV}(u_0),$$

uniformly in  $h$ .

The results of [7] show that if  $u_0 \in L^\infty(\Omega) \cap BV(\Omega)$ , then the solution to the semidiscrete flow  $\{u^k\}_{k=0}^K \subset L^\infty(\Omega) \cap BV(\Omega)$ . Since the operator  $\Pi_h$  has first order approximation properties in  $L^1(\Omega)$  and is stable in  $L^\infty(\Omega)$  we can invoke an inequality similar to (3.11) with  $q = 2$  to conclude the proof.  $\square$

**Remark 6.14** (Error in space). At this stage one might ask whether an optimal rate of convergence  $\mathcal{O}(h^{1/2})$  can be obtained for (6.10). Our methods show that this is possible under two conditions: higher regularity of the solution or a uniform bound for the discrete time derivative. To realize that a more regular solution allows for a  $\mathcal{O}(h^{1/2})$  bound the reader can repeat our analysis for the heat equation (where the arguments are simpler) and convince himself that optimality is a consequence of the higher regularity that is expected from the solution. On the other hand, if we can show that

$$\sup \left\{ \left\| \frac{u_h^{k+1} - u_h^k}{\Delta t} \right\|_{L^\infty} : k = \overline{0, K-1} \right\} \leq c,$$

uniformly with respect to  $h$  and  $\Delta t$ , an  $\mathcal{O}(h^{1/2})$  error estimate easily follows from (6.13). This bound is observed, for instance, if the initial data is the characteristic of a ball, i.e.,  $u_0 = \chi_B$ , see [7].

**Remark 6.15** (Error in time). On the basis of the results presented in [25, 28] we can conclude that the convergence rate in time given in (6.12) is optimal, since we only assume  $\text{TV}(u_0) < \infty$ . As these references also show, if

$$(6.16) \quad \partial \text{TV}(u_0) \cap L^2(\Omega) \neq \emptyset,$$

the rate can be improved to  $\mathcal{O}(\Delta t)$ . However, a characterization of the subdifferential of  $\text{TV}(u_0)$  is not an easy task. Notice that, at least formally, (6.16) means

$$\text{div} \left( \frac{\nabla u_0}{|\nabla u_0|} \right) \in L^2(\Omega),$$

which cannot be related with a smoothness assumption for  $u_0$ . More details on the characterization of the subdifferential  $\partial \text{TV}$  are provided in Lemma 7.5.

## 7. LIMITATIONS OF TVD INTERPOLATION

At this point one might be inclined to believe that our constructions and results are too specialized to be anything but a mere mathematical curiosity, since they heavily rely on the Cartesian structure of the mesh and our particular definition of the total variation employing the  $\ell^1$ -norm in  $\mathbb{R}^d$ . It is the purpose of this section to provide evidence showing that these two assumptions seem necessary. We begin by showing that Lagrange interpolation might not satisfy (TVD).

**7.1. Failure of TVD for Lagrange interpolation.** If one knows *a priori* that the function to be interpolated is also continuous, then the most natural interpolant to use is the so-called nodal or Lagrange interpolant, which takes the point values at the nodes and uses them to define a discrete function. This choice satisfies (TVD) for dimension  $d = 1$  but the situation is strikingly different for  $d > 1$ . The following two examples illustrate this fact for  $d = 2$ .

**7.1.1. Failure due to mesh geometry.** Let  $s \in (0, 1]$  and set  $\Omega = (-1, 1) \times (-s, s)$ . Consider the mesh shown in Figure 7.1 and  $w(x) = 1 - |x^1| \in \mathcal{C}^0(\Omega)$ . The nodal interpolant,  $\mathcal{L}_h w$ , of  $w$  coincides with the nodal basis function associated with the origin. A straightforward calculation shows that, for any  $p \in [1, \infty]$ ,

$$\text{TV}_p(\mathcal{L}_h w) = 2s + 2.$$

Since  $\text{TV}_p(w) = 4s$  this shows that, whenever  $s < 1$ , the Lagrange interpolation operator  $\mathcal{L}_h$  does not satisfy (TVD).

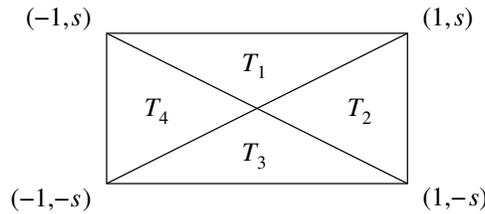


FIGURE 7.1. Triangulation on which the nodal interpolation operator does not satisfy (TVD).

**7.1.2. Failure due to function structure.** Let each quadrant of  $\Omega = (-1, 1)^2$  be divided into two right triangles as depicted in Figure 5.1, and consider the function

$$w(x) = \frac{1}{\varepsilon} (\varepsilon - |x|_{\ell^\infty})^+,$$

where by  $z^+$  we mean the positive part of  $z$  and  $0 < \varepsilon < 1$ . A simple calculation shows that, for any  $p \in [1, \infty]$ , we have  $\text{TV}_p(w) = 4\varepsilon$  whereas  $\int_T |\nabla \mathcal{L}_h w|_p = 1/2$

for 6 triangles  $T$  and zero for the remaining 2. Consequently,  $\text{TV}_p(\mathcal{L}_h w) = 3$  and the Lagrange interpolation operator  $\mathcal{L}_h$  does not satisfy (TVD) provided  $\varepsilon < 3/4$ .

**7.2. Consequences of TVD interpolation.** As described in Remark 2.2 there is more than one possible way to define the total variation which, for instance, implies that for a given  $f \in L^\infty(\Omega)$  the solution to (6.2) changes depending on the chosen definition. It is then natural to ask the question whether, given a  $p \neq 1$ , it is possible to construct an operator that satisfies (TVD). To contradict this fact we consider a family of TV minimization problems. Given  $p \in [1, \infty]$  we introduce the functional

$$(7.1) \quad E_p(w) = \text{TV}_p(w) + \frac{\alpha}{2} \|w - f\|_{L^2}^2,$$

so that the functional  $E$ , defined in (6.1) is actually  $E = E_1$ . Given  $f \in BV(\Omega) \cap L^\infty(\Omega)$  we define  $\mathbf{v}_p \in BV(\Omega) \cap L^2(\Omega)$  as

$$(7.2) \quad \mathbf{v}_p = \underset{w \in BV(\Omega)}{\text{argmin}} E_p(w),$$

and  $v_{p,h} \in X_h$  as

$$(7.3) \quad v_{p,h} = \underset{w_h \in X_h}{\text{argmin}} E_p(w_h).$$

With this notation the functions  $\mathbf{v}$  and  $v_h$ , defined in (6.2) and (6.3), respectively, become  $\mathbf{v} = \mathbf{v}_1$  and  $v_h = v_{1,h}$ .

The following result is nothing but a restatement of Theorem 6.5.

**Proposition 7.4.** *If the space  $X_h$  is such that there exists an operator  $\Pi_h : L^1(\Omega) \rightarrow X_h$  that is first order accurate in  $L^1(\Omega)$  and satisfies*

$$\text{TV}_p(\Pi_h w) \leq \text{TV}_p(w), \quad \forall w \in BV(\Omega),$$

then

$$\|\mathbf{v}_p - v_{p,h}\|_{L^2} \leq c\sqrt{h},$$

where  $c > 0$  depends on  $\|\mathbf{v}_p\|_{L^\infty}$ ,  $\text{TV}_p(\mathbf{v})$  and  $\|f\|_{L^\infty}$ .

**7.3. Explicit solutions of TV minimization.** To be able to proceed with the discussion we must first provide explicit solutions to problem (7.2). In this subsection, if  $p \in [1, \infty]$  the conjugate exponent is denoted by  $q$ , i.e.,  $q = p/(p-1)$ , with the usual modification if  $p = 1$ .

**Lemma 7.5** (Optimality conditions). *A function  $\mathbf{v}_p \in BV(\Omega) \cap L^2(\Omega)$  is a minimizer for  $E_p$  if and only if there exists a vector field  $\mathbf{z} \in \mathbf{L}^\infty(\Omega)$  with  $\text{div } \mathbf{z} \in L^2(\Omega)$  and  $|\mathbf{z}|_{\ell^q} \leq 1$  a.e. in  $\Omega$  such that*

$$\text{div } \mathbf{z} = \alpha(\mathbf{v}_p - f), \quad \mathbf{z} \in \partial \text{TV}_p(\mathbf{v}_p),$$

where the subdifferential inclusion is interpreted as

$$-\langle \mathbf{v}_p, \text{div}(\mathbf{q} - \mathbf{z}) \rangle \leq 0,$$

for all  $\mathbf{q} \in \mathbf{L}^\infty(\Omega)$  with  $\text{div } \mathbf{q} \in L^2(\Omega)$  and  $|\mathbf{q}|_{\ell^q} \leq 1$  a.e. in  $\Omega$ .

*Proof.* It follows from the characterization of  $\partial \text{TV}_p(w)$  provided in [2]; see also [3] and [6].  $\square$

With this result at hand we can find explicit minimizers of (7.1). In what follows we denote by  $B_r^t$  the ball centered at zero and with radius  $r$  in the  $\ell^t$  norm, i.e., the set

$$B_r^t = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}|_{\ell^t} \leq r\}.$$

**Proposition 7.6** (Explicit minimizers). *Let  $p \in \{1, 2\}$ ,  $r > 0$  be such that  $B_r^q \subset \Omega$ . If  $f = \chi_{B_r^q}$  is the characteristic of  $B_r^q$ , then*

$$v_p = \max \left\{ 0, 1 - \frac{d}{\alpha r} \right\} \chi_{B_r^q},$$

is the solution to (7.1).

*Proof.* Assume  $d \leq \alpha r$  and define

$$\mathbf{z}(\mathbf{x}) = \begin{cases} -\frac{1}{r}\mathbf{x}, & |\mathbf{x}|_{\ell^q} \leq r, \\ -\frac{r}{|\mathbf{x}|_{\ell^q}^d}\mathbf{x}, & |\mathbf{x}|_{\ell^q} > r. \end{cases}$$

Clearly  $\mathbf{z} \in \mathbf{L}^\infty(\Omega)$  with  $|\mathbf{z}|_{\ell^q} \leq 1$  and, moreover,  $\operatorname{div} \mathbf{z} = -\frac{d}{r}\chi_{B_r^q} \in L^2(\Omega)$ . Set  $v_p = \frac{1}{\alpha} \operatorname{div} \mathbf{z} + f$ . To verify that this is indeed the solution we must prove the subdifferential inclusion

$$(7.7) \quad -\langle v_p, \operatorname{div}(\mathbf{q} - \mathbf{z}) \rangle = -\left(1 - \frac{d}{\alpha r}\right) \int_{\partial B_r^q} (\mathbf{q} - \mathbf{z}) \cdot \boldsymbol{\nu} \leq 0,$$

for all suitable  $\mathbf{q}$ , where  $\boldsymbol{\nu}$  denotes the unit outer normal. If  $p = q = 2$  then it suffices to notice that  $\mathbf{z} = -\boldsymbol{\nu}$  on  $\partial B_r^2$  and  $\mathbf{q} \cdot \boldsymbol{\nu} \leq 1$ . If  $p = 1$  and  $q = \infty$  we have  $\mathbf{z} \cdot \boldsymbol{\nu} = -1$  on  $\partial B_r^\infty$  and  $\mathbf{q} \cdot \boldsymbol{\nu} \leq 1$ , since  $|\mathbf{q}|_{\ell^\infty} \leq 1$  and  $\boldsymbol{\nu}$  coincides with one of the canonical basis vectors of  $\mathbb{R}^d$  almost everywhere on  $\partial B_r^\infty$ .

If  $d \geq \alpha r$  define

$$\mathbf{z}(\mathbf{x}) = \begin{cases} -\frac{\alpha}{d}\mathbf{x}, & |\mathbf{x}|_{\ell^q} \leq r, \\ -\frac{\alpha r^{d-1}}{d|\mathbf{x}|_{\ell^q}^d}\mathbf{x}, & |\mathbf{x}|_{\ell^q} > r, \end{cases}$$

which clearly verifies  $\operatorname{div} \mathbf{z} = -\alpha \chi_{B_r^q} = -\alpha f$  and  $|\mathbf{z}|_{\ell^q} \leq \frac{\alpha r}{d} \leq 1$ . Since  $v_p = 0$  the variational inclusion  $-\langle v_p, \operatorname{div}(\mathbf{q} - \mathbf{z}) \rangle \leq 0$  is trivially satisfied.  $\square$

**7.4. Computational counterexamples.** With the explicit minimizers of (7.1) constructed in §7.3 at hand, let us now show that if either  $p \neq 1$  or if the mesh is non-Cartesian, then there is no operator that satisfies (TVD). To do so, we will show that the conclusion of Proposition 7.4 is false by means of a series of computational examples, thus yielding the non-existence of such operator  $\Pi_h$ .

We implemented, with the help of the `deal.II` library [4, 5], the TV minimization algorithm to solve (7.1) that was proposed and analyzed in [6]. In all the examples presented below we set  $\Omega = (-1, 1)^2$ ,  $\alpha = 10$  and  $f = \chi_{B_{1/2}^q}$ . The exit criterion for the iterative scheme was

$$\|\tau^{-1}(v_h^{k+1} - v_h^k)\|_{L^2} \leq h, \quad \|\tau^{-1}(\mathbf{z}_h^{k+1} - \mathbf{z}_h^k)\|_{\mathbf{L}^2} \leq h,$$

where  $h$  is the mesh size and  $\tau$  is the pseudo-time used in the iterative scheme (see [6] for details). Since the `deal.II` library is based on quadrilaterals, we ran the same experiments, with an in house code, on triangulations consisting of right

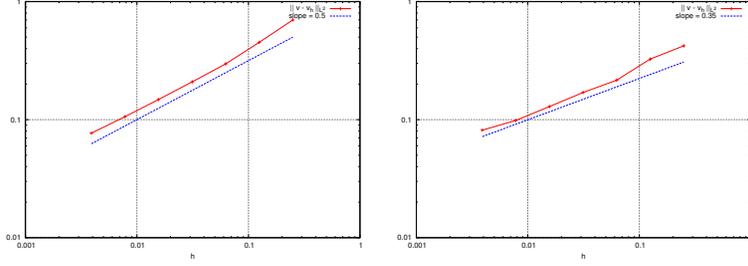


FIGURE 7.2. Convergence rate for total variation minimization with  $\text{TV}_1$  (see (7.1)) on Cartesian meshes (left) and non-Cartesian meshes (right). The experimental rate of convergence is  $\mathcal{O}(h^{1/2})$  on Cartesian meshes, i.e., in accordance with Proposition 7.4, whereas it is  $\mathcal{O}(h^{0.35})$  for non-Cartesian meshes, thus contradicting the conclusion of Proposition 7.4.

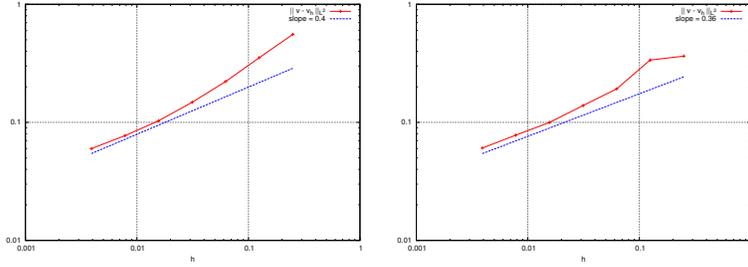


FIGURE 7.3. Convergence rate for total variation minimization with  $\text{TV}_2$  (see (7.1)) on Cartesian meshes (left) and non-Cartesian meshes (right). The experimental rate of convergence is  $\mathcal{O}(h^{0.4})$  and  $\mathcal{O}(h^{0.36})$ , respectively, thus contradicting the conclusion of Proposition 7.4 in both cases.

triangles and perturbed right triangles. The results we obtained are similar to those presented below for quadrilaterals and, thus, we do not show them.

Figure 7.2 shows the convergence rates for the approximation of the minimizers of  $E_1$  on Cartesian and non-Cartesian meshes. In accordance to Proposition 7.4 the convergence rate is  $\mathcal{O}(\sqrt{h})$  for Cartesian meshes. In contrast the convergence rate reduces to  $\mathcal{O}(h^{0.35})$  by just taking a non-Cartesian mesh of quadrilaterals obtained by random perturbation of the nodes. This shows that, in this case, the conclusion of Proposition 7.4 is not true and, consequently, a TVD operator cannot exist. The same conclusion applies to the minimizer of  $E_2$  either with Cartesian or non-Cartesian meshes, as Figure 7.3 shows.

The examples presented above show that, although our construction might seem restrictive it is indeed the only one possible to enforce (TVD). Let us finally point out that the definition of total variation involving the  $\ell_1$  norm in  $\mathbb{R}^d$ , i.e.,  $\text{TV}_1$  is common in the study of monotone schemes for nonlinear conservation laws [23, equation (A.19)].

**Acknowledgement.** We would like to thank Noel Walkington for pointing out a slight inaccuracy in an earlier version of this work, and Antonin Chambolle for suggesting the exact minimizers described in §7.3.

## REFERENCES

- [1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.
- [2] F. Andreu-Vailló, V. Caselles, and J.M. Mazón. *Parabolic quasilinear equations minimizing linear growth functionals*, volume 223 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2004.
- [3] G. Anzellotti. Pairings between measures and bounded functions and compensated compactness. *Ann. Mat. Pura Appl. (4)*, 135:293–318 (1984), 1983.
- [4] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II *Differential Equations Analysis Library, Technical Reference*. <http://www.dealii.org>.
- [5] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II — a general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.*, 33(4), 2007.
- [6] S. Bartels. Total variation minimization with finite elements: Convergence and iterative solution. *SIAM Journal on Numerical Analysis*, 50(3):1162–1180, 2012.
- [7] S. Bartels, R.H. Nochetto, and A.J. Salgado. Discrete TV flows without regularization. arXiv:1212.1137, 2012.
- [8] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
- [9] A. Chambolle. Total variation minimization and a class of binary MRF models. In A. Rangarajan, B. Vemuri, and A.L. Yuille, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 3757 of *Lecture Notes in Computer Science*, pages 136–152. Springer Berlin Heidelberg, 2005.
- [10] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40(1):120–145, 2011.
- [11] Z. Chen and R.H. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84(4):527–548, 2000.
- [12] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, 9(R-2):77–84, 1975.
- [13] W. Dahmen, B. Faermann, I.G. Graham, W. Hackbusch, and S.A. Sauter. Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method. *Math. Comp.*, 73(247):1107–1138, 2004.
- [14] T. Dupont and L.R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34(150):441–463, 1980.
- [15] R.G. Durán. On polynomial approximation in Sobolev spaces. *SIAM J. Numer. Anal.*, 20(5):985–988, 1983.
- [16] R.G. Durán and A.L. Lombardi. Error estimates on anisotropic  $Q_1$  elements for functions in weighted Sobolev spaces. *Math. Comp.*, 74(252):1679–1706 (electronic), 2005.
- [17] X. Feng and A. Prohl. Analysis of total variation flow and its finite element approximations. *M2AN Math. Model. Numer. Anal.*, 37(3):533–556, 2003.
- [18] 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(3):441–456, 2005.
- [19] P.-W. Fok, R.R. Rosales, and D. Margetis. Facet evolution on supported nanostructures: Effect of finite height. *Phys. Rev. B*, 78:235401, Dec 2008.
- [20] E.S. Fu, D.-J. Liu, M.D. Johnson, J.D. Weeks, and E.D. Williams. The effective charge in surface electromigration. *Surface Science*, 385(2–3):259 – 269, 1997.
- [21] M.-H. Giga and Y. Giga. Very singular diffusion equations: second and fourth order problems. *Jpn. J. Ind. Appl. Math.*, 27(3):323–345, 2010.

- [22] M.-H. Giga, Y. Giga, and R. Kobayashi. Very singular diffusion equations. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, pages 93–125. Math. Soc. Japan, Tokyo, 2001.
- [23] H. Holden and N.H. Risebro. *Front tracking for hyperbolic conservation laws*, volume 152 of *Applied Mathematical Sciences*. Springer, New York, 2011.
- [24] R. Kobayashi, J.A. Warren, and W.C. Carter. A continuum model of grain boundaries. *Physica D: Nonlinear Phenomena*, 140(12):141 – 150, 2000.
- [25] R.H. Nochetto, G. Savaré, and C. Verdi. A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations. *Comm. Pure Appl. Math.*, 53(5):525–589, 2000.
- [26] R.H. Nochetto and L.B. Wahlbin. Positivity preserving finite element approximation. *Math. Comp.*, 71(240):1405–1419 (electronic), 2002.
- [27] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1–4):259–268, 1992.
- [28] J. Rulla. Error analysis for implicit approximations to solutions to Cauchy problems. *SIAM J. Numer. Anal.*, 33(1):68–87, 1996.
- [29] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.
- [30] S.L. Sobolev. *Nekotorye primeneniya funktsional' nogo analiza v matematicheskoj fizike*. Izdat. Leningrad. Gos. Univ., Leningrad, 1950.
- [31] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin, 2007.
- [32] J. Wang and B.J. Lucier. Error bounds for finite-difference methods for Rudin-Osher-Fatemi image smoothing. *SIAM J. Numer. Anal.*, 49(2):845–868, 2011.
- [33] W.P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

(S. Bartels) ABTEILUNG FÜR ANGEWANDTE MATHEMATIK, ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG  
HERMANN-HERDER-STR. 1079104 FREIBURG I.BR., GERMANY.  
*E-mail address:* bartels@mathematik.uni-freiburg.de

(R.H. Nochetto) DEPARTMENT OF MATHEMATICS AND INSTITUTE FOR PHYSICAL SCIENCE AND  
TECHNOLOGY, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA.  
*E-mail address:* rhn@math.umd.edu

(A.J. Salgado) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK,  
MD 20742, USA.  
*E-mail address:* abnersg@math.umd.edu