Finite Dimensional Approximation of Convex Regularization via Hexagonal Pixel Grids

Clemens Kirisits [*], Christiane Pöschl [†], Elena Resmerita [‡], and Otmar Scherzer [§]

May 5, 2021

Abstract

This work extends the existing convergence analysis for discrete approximations of minimizers of convex regularization functionals. In particular, some solution concepts are generalized, namely the standard minimum norm solutions for squared norm regularizers and the $\mathcal{R}$-minimizing solutions for general convex regularizers, respectively. A central part of the work addresses finite dimensional approximations of solutions of ill-posed operator equations with basis functions defined on *hexagonal* grids, which require the novel solution concept.

1 Introduction

This work is concerned with a follow-up to the paper [27], where we derived a convergence analysis for discrete approximations of minimizers of convex regularization functionals.

The general formulation is as follows: We are given an operator $F : \mathcal{U} \to \mathcal{V}$, where $\mathcal{U}$ and $\mathcal{V}$ denote infinite dimensional Banach spaces. Moreover, we

[*]Computational Science Center, University of Vienna, Oskar Morgenstern Platz 1, A-1090 Vienna *Clemens.Kirisits@univie.ac.at*

[†]Institute of Mathematics, Alpen Adria Universität Klagenfurt, Universitätsstraße 66-67, A-9020 Klagenfurt, Austria, *Christiane.Poeschl@aau.at*

[‡]Institute of Mathematics, Alpen Adria Universität Klagenfurt, Universitätsstraße 66-67, A-9020 Klagenfurt, Austria, *Elena.Resmerita@aau.at*

[§]Computational Science Center, University of Vienna, Oskar Morgenstern Platz 1, A-1090 Vienna, Austria and Radon Institute of Computational and Applied Mathematics, Altenbergerstraße 69, A-4040 Linz, Austria *Otmar.Scherzer@univie.ac.at*

are given noisy data v^δ (that are realizations of the ideal data v) and we wish to determine an approximation of a solution $u^\dagger$ of the equation

$$F(u) = v\,, \tag{1}$$

using the available data v^δ only. Based on the available prior information on the solution, we restrict our attention to $\mathcal{R}$-*approximate minimizing solutions* of the equation, that are solutions of (1) which approximately minimize a suitable convex and proper functional $\mathcal{R}$. To achieve this goal, our method of choice is *convex variational regularization* with not-necessarily quadratic fit-to-data term which, for given $\alpha > 0$, approximates an $\mathcal{R}$-minimizing solution of (1) by a minimizer u_α^δ of the functional

$$u \to \left\| F(u) - v^\delta \right\|^p + \alpha \mathcal{R}(u)\ , \quad p > 1.$$

For the numerical realization, we consider a sequence of finite dimensional subspaces (of increasing dimensionality)

$$\{\mathcal{U}_n \subseteq \mathcal{U}\}_{n\in\mathbb{N}}$$

and a sequence of operators

$$\{F_m\}_{m\in\mathbb{N}}$$

approximating F in an appropriate sense.

Thus, one obtains a sequence of elements $u_{m,n}^{\alpha,\delta} \in \mathcal{D}(F_m) \cap \mathcal{D}(\mathcal{R})$, which minimizes the functionals

$$u \to \left\| F_m(u) - v^\delta \right\|^p + \alpha \mathcal{R}(u)\,, \qquad p > 1 \tag{2}$$

over $\mathcal{U}_n$, respectively. The results in [27], which generalize the results of [25] from a Hilbert space setting to a Banach space setting (see also Sections 4.1.1-4.1.2 in [31] regarding the corresponding infinite dimensional framework), show that

1. $u_{m,n}^{\alpha,\delta} \to u^\dagger$ in a weak sense and
2. $\mathcal{R}(u_{m,n}^{\alpha,\delta}) \to \mathcal{R}(u^\dagger)$ for $m, n \to \infty$ and $\delta, \alpha(\delta, m, n) \to 0$ in an appropriate manner.

Earlier, in [25] we considered the particular case $\mathcal{R}(u) = \|u - u_0\|^2$ (for some given $u_0 \in \mathcal{U}$) and operators F_m between Hilbert spaces, in which case the two results on weak convergence and convergence of the regularization functional guarantee that

$$u_{m,n}^{\alpha,\delta} \to u^\dagger \text{ strongly.}$$

The regularization method with finite dimensional approximations considered in this paper cannot be analysed with previous settings ([25, 27]), and requires a weaker form of convergence for $n, m \to \infty$, which will be specified in Section 2.

Important applications are finite dimensional realizations of total variation minimization, which was introduced in [29]. Note that such realizations correspond to the most commonly used finite difference methods, for which our analysis now provides a general convergence framework. In contrast, implementation of regularization techniques with continuous finite element approximations can be analysed with the general results of [27]. The particular case of total variation minimization discretized by continuous finite elements has been studied before in [8]. The present paper deals also with piecewise constant approximations of bounded variation functions in the context of anisotropic total variation minimization. The drawback of the approach is that it is directionally sensitive. There are attempts in the literature to employ the isotropic version of the total variation seminorm; these use triangles instead of squares - see, e.g., [7]. Another idea for preserving rotational features in imaging is using hexagonal pixels, because hexagons better approximate circles than squares.

A central part of the current work concerns finite dimensional approximations of solutions of (1) with basis functions defined on *hexagonal* grids. For piecewise constant functions on regular hexagonal grids we find another application where the convergence analysis of [25, 27] is not reliable anymore, and the new approach of this paper becomes necessary.

Aside from theoretical consideration, this work has also practical relevance because hexagonal arrays are already implemented in photo-films, electronic paper [21], large-scale media displays and also large optical reflecting telescopes. In all these applications it is favourable to incorporate the hexagonal grid structure into the mathematical algorithm - see also [20] for references on hexagonal coordinate systems related topics. The idea of approximating bounded variation functions on hexagonal pixels has already been considered in a number of image or signal processing papers with remarkable implementation results - see, for instance [16] (more references to select from a long list from vision), and references in [24]. The aim here is to support this idea from a theoretical viewpoint.

This paper is organized as follows: In Section 2 we present a convergence analysis for generalized variational regularization methods. Section 3 is concerned with applications to generalized total variation minimization and in particular with applications to total variation minimization on hexagonal grids (Subsection 3.1). Section 4 presents practical realization for generalized total variation denoising and numerical results.

2 General Convergence Results

Before we state the main result on finite dimensional regularized approximations, we clarify the notation.

Let $\mathcal{D}(F)$ and $\mathcal{D}(\mathcal{R})$ stand for the domains of the operator F and of the proper convex functional $\mathcal{R}$, respectively. By $u^\dagger$ we denote an $\mathcal{R}$-minimizing solution of the inverse problem $F(u) = v$, that is $u^\dagger$ solves $\min \mathcal{R}(u)$ subject to $F(u) = y$. Here we are concerned with approaching *$\mathcal{R}$-approximate minimizing solutions.*

Definition 1. *An* $\mathcal{R}$-approximate minimizing solution $\tilde{u} \in \mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R})$ *is an element which satisfies*

$$\mathcal{R}(\tilde{u}) \leq c_{\mathcal{R}} \mathcal{R}(u) \text{ for all } u \in S \text{ and for some } c_{\mathcal{R}} \geq 1, \tag{3}$$

where S denotes the feasibility set, *which is defined by*

$$S := \{u \in \mathcal{D}, F(u) = v\} \ .$$

In the case $c_{\mathcal{R}} = 1$, this is just the concept of an $\mathcal{R}$-minimizing solution.

We will later give an example where $c_{\mathcal{R}}$ depends on the structure of the finite dimensional subsets that one uses for the approximation of the space $\mathcal{U}$.

Remark 1. *Observe that an $\mathcal{R}$-approximate minimizing solution $\tilde{u}$ is actually an ε-minimizer of the functional $\mathcal{R}$ over S for a certain $\varepsilon \geq 0$ (see, e.g. [3, Section 3.1.3]:*

$$\mathcal{R}(\tilde{u}) \leq c_{\mathcal{R}} \inf_{u \in S} \mathcal{R}(u) = \inf_{u \in S} \mathcal{R}(u) + \varepsilon,$$

with $\varepsilon = (c_{\mathcal{R}} - 1) \inf_{u \in S} \mathcal{R}(u)$.

We assume that the available data v^δ of v satisfy

$$\left\| v^\delta - v \right\| \leq \delta \ . \tag{4}$$

The method of choice is Tikhonov regularization with a norm power as a fit-to-data term, that is, it consists in minimizing the functional

$$u \to \mathcal{F}(u) := \left\| F(u) - v^\delta \right\|^p + \alpha \mathcal{R}(u)$$

over $\mathcal{D}$, with $p > 1$. Actually, we discretize this problem by working in finite dimensional subspaces $\mathcal{U}_n$, namely by considering the minimization of the functional $\mathcal{F}_m$ defined by

$$u \to \mathcal{F}_m(u) := \left\| F_m(u) - v^\delta \right\|^p + \alpha \mathcal{R}(u) \tag{5}$$

over $\mathcal{D}_{m,n}$, where F_m and $\mathcal{D}_{m,n}$ are approximations of F and $\mathcal{D}$, respectively.

We base our analysis on the following assumptions and conventions, which are summarized from [28, 23, 30, 27, 26].

Assumption 1. *1. The Banach space $\mathcal{U}$ is considered with two topologies: the strong topology induced by the norm and a weaker topology τ.*

2. *$\mathcal{V}$ is a reflexive Banach space. The according weak and strong convergence are denoted by $\rightharpoonup$, $\to$, respectively.*

3. *The power p satisfies $p > 1$.*

4. *The operator $F : \mathcal{D} \subseteq \mathcal{U} \to \mathcal{V}$ is sequentially τ-weakly closed. That is, if*

   $$\{u_k\} \subset \mathcal{D}\,, u_k \to_\tau u \text{ and } F(u_k) \rightharpoonup v$$

   then

   $$u \in \mathcal{D} \text{ and } v = F(u)\ .$$

   Moreover, F is continuous with respect to the norm topologies on $\mathcal{U}, \mathcal{V}$, respectively.

5. *The functional $\mathcal{R}$ is proper, convex, non-negative and sequentially τ-lower semi-continuous on $\mathcal{U}$.*

6. *$\{\mathcal{U}_n\}$ is a family of subspaces of $\mathcal{U}$.*

7. *The operators $F_m : \mathcal{D}(F_m) \subseteq \mathcal{U} \to \mathcal{V}$ satisfy the following properties:*

   - *For every pair of indices m, n,*

     $$\emptyset \neq \mathcal{D}_{m,n} := \mathcal{D}(F_m) \cap \mathcal{U}_n \cap \mathcal{D}(\mathcal{R}) \subseteq \mathcal{D}$$

     and $\mathcal{D}_{m,n}$ is τ-closed.

- *For every $m \in \mathbb{N}$, F_m is weakly continuous, i.e.,*

$$u_l \to_\tau u \Rightarrow F_m(u_l) \rightharpoonup F_m(u)\ .$$

- *Moreover, we assume that*

$$\|F(u) - F_m(u)\| \leq \rho_{m,n},\ \forall u \in \mathcal{D}_{m,n} \text{ and } \lim_{m,n\to\infty} \rho_{m,n} = 0\ . \quad (6)$$

8. *For every $M > 0$, $\alpha > 0$, $v \in \mathcal{V}$, $m \in \mathbb{N}$, the sets $\{u \in \mathcal{D}_{m,n} : \mathcal{F}_m(u) \leq M\}$ are τ-sequentially relatively compact.*

In this context, one can state well-posedness and stability of the proposed variational regularization (see, e.g., [27, Proposition 2.3]):

Proposition 1. *Let $m, n \in \mathbb{N}$ and $\alpha, \delta > 0$ be fixed. Then, for every $v^\delta \in \mathcal{V}$ there exists at least one minimizer $u_{m,n}^{\alpha,\delta} \in \mathcal{D}_{m,n}$ of the functional $\mathcal{F}_m$.*

Moreover, the minimizers of (5) are stable with respect to the data v^δ in the following sense: if $\{v_k\}_{k\in\mathbb{N}}$ converges strongly to v^δ, then every sequence $\{u_k\}_{k\in\mathbb{N}}$ of minimizers of (5) where v^δ is replaced by v_k has a subsequence $\{u_l\}_{l\in\mathbb{N}}$ which converges with respect to the topology τ to a minimizer $\tilde{u}$ of (5) and such that $\{\mathcal{R}(u_l)\}_{l\in\mathbb{N}}$ converges to $\mathcal{R}(\tilde{u})$, as $l \to \infty$.

For instance, [27] discusses the setting $\mathcal{U} = BV(\Omega)$ (the space of bounded variation functions) with τ being the weak-star topology in $BV(\Omega)$.

As regards convergence, we assumed in [27] that $\mathcal{U} = \overline{\bigcup_n \mathcal{U}_n}^d$ for some metric d that was connected to the regularization functional. We showed that the minimizers of the finite dimensional Tikhonov minimization problem converged to an $\mathcal{R}$ minimizing solution of the original problem. Now we are in the situation that we cannot find such a metric, hence we have to weaken our solution concept.

Theorem 1 below is technically standard aside that we weaken the assumptions on the finite dimensional subspaces $\mathcal{U}_n$. Thus, a proof of this theorem is added to highlight the new scenario.

Theorem 1. *Let Assumption 1 be satisfied.*

Moreover, assume that:

1. *There exist an $\mathcal{R}$-minimizing solution $u^\dagger$, a sequence $\phi_{m,n} \in \mathcal{D}_{m,n} := \mathcal{U}_n \cap \mathcal{D}(F_m) \cap \mathcal{D}(\mathcal{R})$ and a constant $c_\mathcal{R} \geq 1$, such that*

$$\phi_{m,n} \to_\tau u^\dagger \qquad \text{and} \qquad \limsup_{m,n\to\infty} \mathcal{R}(\phi_{m,n}) \leq c_\mathcal{R}\mathcal{R}(u^\dagger)\ . \qquad (7)$$

2. $\alpha := \alpha(m, n, \delta)$ is chosen such that

$$\alpha \to 0\,, \quad \frac{\delta^p}{\alpha} \to 0\,, \quad \frac{\rho_{m,n}^p}{\alpha} \to 0\,, \tag{8}$$

and

$$\frac{\|F(\phi_{m,n}) - v\|^p}{\alpha} \to 0 \text{ for } \delta \to 0\,,\ m, n \to \infty\,. \tag{9}$$

Let $\delta_k \to 0$, $n_k, m_k \to \infty$, and denote by $\alpha_k := \alpha(m_k, n_k, \delta_k)$, $\mathcal{D}_k := \mathcal{D}_{m_k,n_k}$ and

$$u_k := u_{m_k,n_k}^{\alpha_k,\delta_k} \in argmin_{u \in \mathcal{D}_k} \left\{ \left\| F_{m_k}(u) - v^{\delta_k} \right\|^p + \alpha_k \mathcal{R}(u) \right\}\,.$$

Then $\{u_k\}$ has a subsequence $\{u_l\}$ such that $u_l \to_\tau \tilde{u}$, where $\tilde{u}$ is an $\mathcal{R}$-approximate minimizing solution.

Proof. Let us denote $\phi_k := \phi_{m_k,n_k}$, then from the definition of u_k, (4) and (6), it follows that

$$\begin{aligned}
&\|F_{m_k}(u_k) - v^{\delta_k}\|^p + \alpha_k \mathcal{R}(u_k) \\
\le&\|F_{m_k}(\phi_k) - v^{\delta_k}\|^p + \alpha_k \mathcal{R}(\phi_k) \\
\le&(\|F_{m_k}(\phi_k) - F(\phi_k)\| + \|F(\phi_k) - F(u^\dagger)\| + \|F(u^\dagger) - v^{\delta_k}\|)^p + \alpha_k \mathcal{R}(\phi_k) \\
\le&(\rho_{m_k,n_k} + \|F(\phi_k) - v\| + \delta_k)^p + \alpha_k \mathcal{R}(\phi_k).
\end{aligned} \tag{10}$$

Therefore,

$$\mathcal{R}(u_k) \le \frac{(\rho_{m_k,n_k} + \|F(\phi_k) - v\| + \delta_k)^p}{\alpha_k} + \mathcal{R}(\phi_k)\,.$$

Using Assumptions (9) and (7) it follows that

$$\limsup_{k \to \infty} \mathcal{R}(u_k) \le c_{\mathcal{R}} \mathcal{R}(u^\dagger). \tag{11}$$

Since F is sequentially τ-weakly closed, it follows from (7)

$$\lim_{k \to \infty} \|F(\phi_k) - F(u^\dagger)\| = \lim_{k \to \infty} \|F(\phi_k) - v\| = 0\,.$$

Taking into account that the right hand side of (10) tends to 0 and $v^{\delta_k} \to v$ for $\alpha_k \to 0$, it follows that

$$F_{m_k}(u_k) \to v\,. \tag{12}$$

Now (6), (12), (11) and the compactness hypothesis in Assumption 1 yield existence of a subsequence $\{u_j\}_{j\in\mathbb{N}}$ which is τ-convergent to some solution $\tilde{u}$ of (1). Due to the lower semi-continuity of $\mathcal{R}$ and (11), we get

$$\mathcal{R}(\tilde{u}) \le \liminf_{j\to\infty} \mathcal{R}(u_j) \le \limsup_{j\to\infty} \mathcal{R}(u_j) \le c_{\mathcal{R}}\mathcal{R}(u^\dagger).$$

The fact that $u^\dagger$ is an $\mathcal{R}$-minimizing solution implies $\mathcal{R}(u^\dagger) \le \mathcal{R}(\tilde{u})$ and yields the conclusion. $\square$

The essential difference to our previous work [25, 27] is that we consider here $\mathcal{R}$-*approximate minimizing solutions*, which requires condition (7) in the analysis. We briefly sketch the different assumptions and results in the current work, as compared to [25, 27].

$\mathcal{U}$ Hilbert space	$u_k \to u^\dagger$ in norm	[25]
$\mathcal{U}$ Banach space, finite dimensional subspaces $\mathcal{U}_n$ are dense with respect to a metric connected to $\mathcal{R}$	$u_k \to_\tau u^\dagger$, $\mathcal{R}(u_k) \to \mathcal{R}(u^\dagger)$	[27]
$\mathcal{U}$ Banach space, there is no metric such that the finite dimensional subspaces $\mathcal{U}_n$ are dense with respect to this metric, (7) holds	$u_k \to_\tau \tilde{u}$, $\mathcal{R}(\tilde{u}) \le c_{\mathcal{R}}\mathcal{R}(u^\dagger)$	Here

3 Generalized Total Variation Minimization

In the literature, the anisotropy of numerical finite difference implementations of isotropic total variation regularization has been widely discussed [12, 32, 2].

In the following we investigate total variation (TV) regularization with generalized TV-seminorms, consisting in minimization of the functional

$$\mathcal{F}(u) = \left\|F(u) - v^\delta\right\|^2 + \alpha TV_g(u)\ .$$

Here TV_g denotes a generalized TV-seminorm, defined below, which is adapted to particular grid structure of the finite dimensional approximations.

Definition 2. ***Generalized total variation seminorm:*** *Let $g : \mathbb{R}^n \to [0, +\infty)$ be a norm on $\mathbb{R}^n$. Define*

$$TV_g(u) = \sup\left\{\int_\Omega u(x)\nabla\cdot\psi(x)\,dx : \psi \in C_0^\infty(\Omega;\mathbb{R}^n), g^*(\psi(\cdot)) \le 1\right\}$$

for $u \in BV(\Omega)$, where $g^* : \mathbb{R}^n \to [0, +\infty]$ is the dual of the norm g, that is

$$g^*(v) := \sup \{ \langle v, w \rangle \ : w \in \mathbb{R}^n, g(w) \leq 1 \} .$$

Remark 2. *i) We emphasize that for smooth functions $u \in C_0^1(\Omega)$ one has*

$$TV_g(u) = \int_\Omega g(\nabla u(x)) \, dx \ .$$

ii) Let $p \in [1, +\infty)$ and let p^* be its conjugate number. Then for $g(\cdot) = |\cdot|_{l^p}$ one has the standard TV-seminorm:

$$TV_{l^p}(u) = \sup \left\{ \int_\Omega u(x) \nabla \cdot \psi(x) \ dx : \psi \in C_0^\infty(\Omega; \mathbb{R}^n), |\psi(\cdot)|_{l^{p^*}} \leq 1 \right\} .$$

Finite dimensional approximations of TV-minimization for a fixed parameter α have been analyzed in detail in a functional analytical setting - see, e.g. [17], [8]. We refer for instance to [8], where $TV_g = TV_{l^p}$, with $p = 1, 2$ have been investigated in the case of linear operators F. The general setting of our work applies in principal to non-linear problems, however the core of this work are approximations of (1) by general TV-seminorms minimization.

We emphasize that the regularization results of this paper rely essentially on the penalty-approximation type assumption (7), which in this particular case means the approximation of functions of bounded variation on different grids.

We summarize below BV-function approximation results according to [8] and [7], where u_n denotes the finite dimensional approximation of a BV-function u:

grid	subdivision	Ansatz-function	
rectangles	regular	linear	$TV_{l_2}(u_n) \to TV_{l_2}(u)$ [8]
rectangles	regular	pcw. const.	$TV_{l_1}(u_n) \to TV_{l_1}(u)$ [8]
rectangles	flexible triangles	pcw. const.	$TV_{l_2}(u_n) \to TV_{l_2}(u)$ [7]

Under quite general conditions, similar convergence results hold for finite dimensional approximations of TV-minimization to a solution of the TV-minimization problem in the infinite dimensional setting involving a linear operator F, cf. [8].

3.1 Total Variation Minimization with Piecewise Constant Functions on Hexagonal Grids

What is missing in the table above is the convergence of piecewise constant approximations of BV-functions witrh respect to the isotropic TV_{l^2}-seminorm on regular grids.

In the beginning we specify the terminology and specifications which will be used throughout this section.

- We restrict attention to the 2-dimensional domain $\Omega = (0,1) \times \left(0, \frac{\sqrt{3}}{2}\right)$.
- Let $\{\Omega_m\}_{m \in \mathcal{I}_n}$ be a minimal, non-overlapping covering of Ω, consisting of hexagons whose centers are aligned on lines parallel to the x-axes, one center being $(0,0)$. For every $n \in \mathbb{N}$, let $h_n = \frac{1}{2^n}$ be the distance between centers of neighboring hexagons. The following situation is considered: If one divides Ω_m into two halves, then the left half is an open set and the right one a closed set, as shown in Figure 2 left. In this way one obtains a non-overlapping covering of Ω.

  Two hexagons with centers $\xi_m, \xi_p \in \Omega$ are neighbors if their closures have a common line segment.

  Other choices of h_n are possible but more complicated to handle: The choice $h_n = \frac{1}{n}$, for instance, leads to complicated situations of boundary hexagons, which we attempt to avoid. See Figure 1 top right.

- Let $\mathcal{I}_n^\circ$ be the index set of hexagons Ω_m that do not touch the boundary of Ω and denote $\Omega_n^\triangle = \{\Omega_m, m \in \mathcal{I}_n^\circ\}$.
- Let ξ_m be a center of a hexagon $\Omega_m \in \Omega_n^\triangle$. We denote the right, top and bottom neighboring hexagon centers (see Figure 2) by $\xi_{m_1}, \xi_{m_2}, \xi_{m_3}$ and define

  $$\triangle_m^+ := Triangle(\xi_m, \xi_{m_1}, \xi_{m_2}), \quad \triangle_m^- := Triangle(\xi_m, \xi_{m_1}, \xi_{m_3})\ .$$

  If either two of the $\xi_{m_1}, \xi_{m_2}, \xi_{m_3}$ are not defined, we set $\triangle_m^+ = \emptyset$ and $\triangle_m^- = \emptyset$, respectively.

- Moreover, we define the space of piecewise constant functions on the hexagonal grid as

  $$U_n := \left\{ u_n : u_n = \sum_{m \in \mathcal{I}_n^\circ} u^m \chi_{\Omega_m}, u^m \in \mathbb{R} \right\}\ . \tag{13}$$

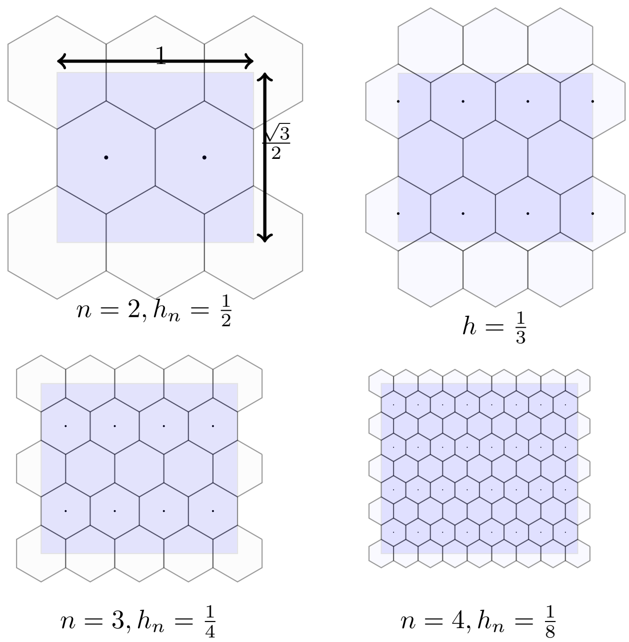


Figure 1: Refinement of Ω.

Note that we only sum over interior hexagons. In this way, when calculating the derivative, we do not have to take care about the boundary hexagons and thus, the sums in the proofs get less complicated. Since there are approximately $4\ \frac{1}{h_n}$ boundary hexagons, and the total area of those is given by $\mathcal{O}(h_n)$, we can neglect them in our asymptotic expansions, as the following calculations show.

- Denote by e_{x_1} and e_{x_2} the unitary vectors in x_1 and x_2 direction, and set

$$\begin{aligned} e_1 &= e_{x_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ e_2 &= \cos(60^\circ) e_{x_1} + \sin(60^\circ) e_{x_2} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \\ e_3 &= \cos(-60^\circ) e_{x_1} + \sin(-60^\circ) e_{x_2} = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}. \end{aligned} \tag{14}$$

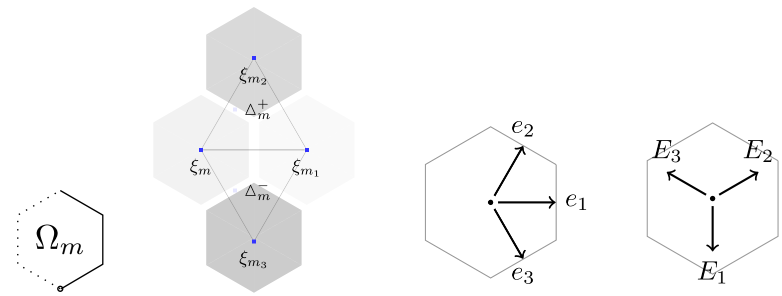


Figure 2: Notation: ξ_m, ξ_{m_i}. Hexagonal coordinate-system.

- Moreover, set

$$E_1 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} , \quad E_2 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} , \quad E_3 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} . \tag{15}$$

  The vectors are depicted in Figure 2.

In the following we introduce hexagonal norms on $\mathbb{R}^2$:

Definition 3. *Let $v = (v_1, v_2) \in \mathbb{R}^2$. Define the following* hexagonal norms*:*

$$|v|_{\hexagon} :== \frac{1}{\sqrt{3}} \sum_{k=1}^{3} |\langle E_k, v \rangle| \tag{16}$$

and

$$|v|_{\hexagon^T} :== \frac{1}{2} \sum_{k=1}^{3} |\langle e_k, v \rangle| \ . \tag{17}$$

Remark 3.
- *$|\cdot|_{\hexagon}$ and $|\cdot|_{\hexagon^T}$ are norms on $\mathbb{R}^2$. Their levelsets are hexagons - see Figure 3.*
- *One has $|e_k|_{\hexagon} = 1$ and $|E_k|_{\hexagon} = \frac{2}{\sqrt{3}}$ for all $k = 1, 2, 3$. Moreover one has*

$$\frac{3}{4} |v|_{\hexagon} \le |v|_{\hexagon^T} \le |v|_{\hexagon} \ , \qquad \forall v \in \mathbb{R}^2 \ .$$

Recall the strict convergence metric notion (see [1, Definition 3.14], [7])

$$d(u,v) = \|u - v\|_{L^1} + |TV_{l_p}(u) - TV_{l_p}(v)| \, , \qquad \forall u, v \in BV(\Omega) \ , \ p \in [1, +\infty).$$

According to [8], for $p = 1$, functions of bounded variation can be approximated by piecewise constant functions on a uniform rectangular grid with respect to the strict convergence metric.

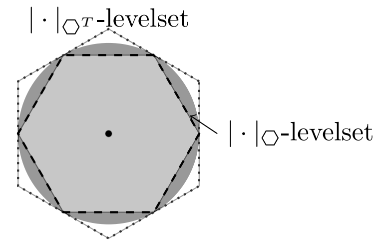


Figure 3: Levelsets of the hexagonal norms.

Now, we are interested in answering the question whether there exists a norm g, such that every function $u \in BV(\Omega)$ can be approximated by piecewise constant functions on hexagonal grids with respect to the metric

$$d(u,v) = \|u-v\|_{L^1} + |TV_g(u) - TV_g(v)| , \qquad \forall u,v \in BV(\Omega) . \tag{18}$$

The answer to this question is no, as shown in the following lemma.

Lemma 1. *There exists no norm g on $\mathbb{R}^2$ which satisfies the following two conditions:*

(i) $g(e_1) = g(e_2) = g(e_3)$*;*

(ii) The length of vertical and horizontal lines are preserved, after approximating these lines on a hexagonal grid.

Proof. Assume that there exists a norm g fulfilling (i)-(ii). Let L_h, L_v be a horizontal and a vertical line in the domain Ω, respectively, positioned as in Figure 5. The length of line L_h is equal to $2\frac{\sqrt{3}}{2}g(e_1)$ (see Figure 2).

When approximating the line L_h by boundaries of the hexagon, it is replaced by the two line segments a_1,b_1. These two lines have the same length as the vectors E_2, E_3, respectively. Because the length of the line segments a_1, b_1 equals the lengths of E_2, E_3, respectively, it follows that

$$g(\sqrt{3}e_1) = g(E_2) + g(E_3) . \tag{19}$$

L_v is approximated by the three line segments a_2,b_2,c_2 and d_2. These four lines have the same length as the vectors E_2, E_1, E_3, E_1, respectively. The second condition to hold then requires:

$$g(3 \cdot E_1) = g(2 \cdot E_1) + g(E_2) + g(E_3) ,$$

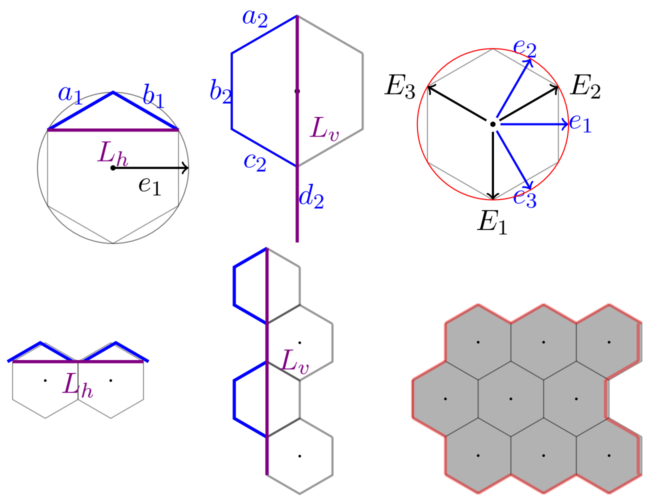


Figure 4: Illustration of the proof of Lemma 1.

which implies that

$$g(E_1) = g(E_2) + g(E_3) \ . \tag{20}$$

From (19), (20) it then follows that

$$g(E_1) = \sqrt{3} g(e_1) \ . \tag{21}$$

This, together with the fact that $E_1 = \frac{1}{\sqrt{3}}(e_2 - e_3)$ and the condition $g(e_1) = g(e_2) = g(e_3)$ shows that

$$\sqrt{3} g(e_1) = g(E_1) = \frac{1}{\sqrt{3}} g(e_2 - e_3) \leq \frac{1}{\sqrt{3}} \left(g(e_2) + g(e_3) \right) = \frac{2}{\sqrt{3}} g(e_1) \, ,$$

which is a contradiction. Next we can consider a grid-refinement as in Figure 5 (right). In this case, one obtains again (19) and (20) for L_h and L_v, respectively. These lead to a contradiction to the assumption that g is a norm. $\square$

Theorem 2. *Let g be a norm on $\mathbb{R}^2$ satisfying*

$$g(e_i) = 1, \, \forall i = 1, 2, 3. \tag{22}$$

Then U_n given by (13) *is a subset of $BV(\Omega)$, $\forall n \in \mathbb{N}$, and for each $u_n \in U_n$*

$$TV_g(u_n) = \sum_{m \in \mathcal{I}_n} \sum_{k=1,2,3} |u^m - u^{m_k}| \, \mathcal{H}^1 \left(\partial\Omega_m \cap \partial\Omega_{m_k} \right) . \tag{23}$$

Here the index m_k denotes the index of the hexagon neighbor to the one of index m in direction $+e_k$, and $\partial\Omega_m \cap \partial\Omega_{m_k} = \emptyset, u^{m^k} = 0$ if Ω_m has no neighbor in direction e_k. Moreover, the following convention is used: $u^m = 0$ for $m \notin \mathcal{I}_n^\circ$, so that

$$u_n = \sum_{m \in \mathcal{I}_n} u^m \chi_{\Omega_m} = \sum_{m \in \mathcal{I}_n^\circ} u^m \chi_{\Omega_m} \in U_n \,.$$

Proof. In this proof, $d(\partial\Omega_m)$ denotes the 1-dimensional measure on the boundary of Ω_m and $\mathbf{n}_{\Omega_m}$ stands for the normal vector field on the boundary of Ω_m. We consider on $BV(\Omega)$ the norm defined by means of the TV_g seminorm. In order to show that $U_n \subset BV(\Omega)$, it suffices to prove (23).

Note that $\sup \left\{ \left\langle e_k, \frac{v}{g^*(v)} \right\rangle : v \in \mathbb{R}^2 \right\}$ is taken for $v = e_k$ (Cauchy-Schwarz in the case for linear dependent vectors in a finite dimensional space). Hence, since $g(e_k) = 1$ we have

$$g^*(e_k) = \sup \left\{ \langle e_k, v \rangle, g(v) = 1 \right\} = \langle e_k, e_k \rangle = 1 \,.$$

With the same argument we obtain

$$\arg \sup_v \left\{ \langle e_k, v \rangle : g^*(v) = 1 \right\} = e_k \,, \quad \sup_v \left\{ \langle e_k, v \rangle : g^*(v) = 1 \right\} = 1 \,. \tag{24}$$

For m, k we can find a function $\phi_{m,k} \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^2)$ with $\phi_{m,k} = e_k$ on $(\partial\Omega_m \cap \partial\Omega_{m_k})^\circ$ (boundaries without the endpoints of the common segment) and $\phi_{m,k} = 0$ on all the other boundaries $(\partial\Omega_{\tilde{m}} \cap \partial\Omega_{\tilde{m}_k})^\circ, \tilde{m} \neq m$, with $\tilde{m} \in \mathcal{I}_n^\circ$. Then using (24), we see that

$$\begin{aligned} &\sup \left\{ a \int_{\partial\Omega_m \cap \partial\Omega_{m_k}} \langle e_k, \psi \rangle \; d(\partial\Omega_m) : \psi \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^2) : \| g^*(\psi) \|_{L^\infty} = 1 \right\} = \\ &\quad = |a| \int_{\partial\Omega_m \cap \partial\Omega_{m_k}} \langle e_k, \phi_{m,k} \rangle \; d(\partial\Omega_m) = |a| \int_{\partial\Omega_m \cap \partial\Omega_{m_k}} \langle e_k, e_k \rangle \; d(\partial\Omega_m) \\ &\quad = |a| \, \mathcal{H}^1 \left(\partial\Omega_m \cap \partial\Omega_{m_k} \right) \,, \end{aligned} \tag{25}$$

where $a \in \mathbb{R}$. Setting $\mathcal{B}_{0,g^*}^\infty := \left\{ \psi \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^2) : \| g^*(\psi) \|_{L^\infty} \leq 1 \right\}$ and plugging in the definition of u_n we have

$$\begin{aligned} TV_g(u_n) &= \sup\left\{\int_\Omega u_n(x)\nabla\cdot\psi(x)\,dx : \psi\in\mathcal{B}^\infty_{0,g^*}\right\} \\ &= \sup\left\{\sum_{m\in\mathcal{I}_n}\int_{\Omega_m} u^m\nabla\cdot\psi\,d\Omega_m : \psi\in\mathcal{B}^\infty_{0,g^*}\right\} \\ &= \sup\left\{\sum_{m\in\mathcal{I}_n}\int_{\partial\Omega_m} u^m\left\langle \mathbf{n}_{\Omega_m},\psi\right\rangle\, d(\partial\Omega_m) : \psi\in\mathcal{B}^\infty_{0,g^*}\right\}. \end{aligned}$$

Note that the vectors of the normal-field $\mathbf{n}_{\Omega_m}$ belong to the set $\{\pm e_k, k=1,2,3\}$. We can rearrange the sum in the supremum in the definition of the TV_g-functional and collect the integrals over $\partial\Omega_m\cap\partial\Omega_{m_k}$, setting $\partial\Omega_m\cap\partial\Omega_{m_k}=\emptyset$ if Ω_m has no neighbor in direction e_k. For any $\psi\in\mathcal{C}^\infty(\Omega,\mathbb{R}^2)$, one has

$$TV_g(u_n) = \sup\left\{\sum_{m\in\mathcal{I}_n}\sum_{k=1,2,3}\int_{\partial\Omega_m\cap\partial\Omega_{m_k}} (u^m-u^{m_k})\left\langle e_k,\psi\right\rangle\, d(\partial\Omega_m) : \psi\in\mathcal{B}^\infty_{0,g^*}\right\}.$$

Recalling (25), we see that the supremum above is taken for the function $\sum_{m\in\mathcal{I}_n}\sum_{k=1,2,3}(u^m-u^{m_k})\phi_{m,k}$, hence we obtain:

$$\begin{aligned} TV_g(u_n) &= \sum_{m\in\mathcal{I}_n}\sum_{k=1,2,3}\int_{\partial\Omega_m\cap\partial\Omega_{m_k}} (u^m-u^{m_k})\left\langle e_k,\phi_{m,k}\right\rangle\, d(\partial\Omega_m) \\ &= \sum_{m\in\mathcal{I}_n}\sum_{k=1,2,3}\int_{\partial\Omega_m\cap\partial\Omega_{m_k}} \left|u^m-u^{m_k}\right|\left\langle e_k,e_k\right\rangle\, d(\partial\Omega_m) \\ &= \sum_{m\in\mathcal{I}_n}\sum_{k=1,2,3} \left|u^m-u^{m_k}\right|\mathcal{H}^1\left(\partial\Omega_m\cap\partial\Omega_{m_k}\right). \end{aligned}$$

□

Theorem 3. *Let g be a norm on $\mathbb{R}^2$ satisfying (22) and $\|v\|_{\hexagon^T}\le g(v)$, $\forall v\in\mathbb{R}^2$. Then for every $u\in BV(\Omega)$, there exists a sequence $\{u_n\}$ of piecewise constant functions on hexagons in U_n satisfying*

$$\lim_{n\to\infty}\|u_n-u\|_{L^1}=0 \quad \textit{and} \tag{26}$$

$$TV_g(u)\le\liminf_{n\to\infty}TV_g(u_n)\le\limsup_{n\to\infty}TV_g(u_n)\le\frac{4}{3}TV_g(u).$$

Proof. Note first that every $\tilde{u} \in BV(\Omega)$ can be approximated by $u_k \in \mathcal{C}^\infty(\bar{\Omega})$ such that $\|u_k - u\|_{L^1} \to 0$ and $TV_g(u_k) \to TV_g(\tilde{u})$ (see [18, Theorem 1.17], which remains valid by replacing $\int |Du|$ by $TV_g(u)$).

Therefore, it is sufficient to prove that smooth functions can be approximated in the sense of (26) by piecewise constant functions. To this end, assume that $u \in \mathcal{C}^\infty(\bar{\Omega})$ and define

$$u_n(x) := \sum_{m \in \mathcal{I}_n^\circ} u(\xi_m) \chi_{\Omega_m}(x) \qquad \forall x \in \Omega \, .$$

For the sake of simplicity, we only took the sum over interior points, taking into account an $\mathcal{O}(h_n)$-term in the estimates below. We obtain the left inequality of (26) from the lower semicontinuity of the functional: For $k = 1, 2, 3$ we have for every $\psi \in \mathcal{B}^\infty_{0,g^*}$

$$\begin{aligned}
&\int_\Omega u_n(x) \nabla \cdot \psi(x) dx = \\
&\qquad - \sum_{m \in \mathcal{I}_n^\circ} \int_{\partial\Omega_m \cap \partial\Omega_{m_k}} (u(\xi_m) - u(\xi_{m_k})) \langle e_k, \psi \rangle \; d(\partial\Omega_m \cap \Omega) + \mathcal{O}(h_n)
\end{aligned}$$

where the index m_k denotes the index of the hexagon with a neighbor of index m in direction e_k. The $\mathcal{O}(h_n)$ term concerns those elements that are not taken into account in the sum over $\mathcal{I}_n^\circ$. Using (23), we obtain

$$\begin{aligned}
TV_g(u_n) =& \sum_{k=1}^3 \sum_{m \in \mathcal{I}_n^\circ} |u(\xi_m) - u(\xi_{m_k})| \, \mathcal{H}^1 \left(\partial\Omega_m \cap \partial\Omega_{m_k}\right) \\
&+ \mathcal{O}(h_n) \, \|\nabla u\|_\infty \\
\leq& \sum_{k=1}^3 \sum_{m \in \mathcal{I}_n^\circ} \left(|u(\xi_m) - u(\eta_{mm_k})| + |u(\eta_{mm_k}) - u(\xi_{m_k})|\right) \mathcal{H}^1 \left(\partial\Omega_m \cap \partial\Omega_{m_k}\right) \\
&\quad + \mathcal{O}(h_n) \, \|\nabla u\|_\infty \ ,
\end{aligned}$$

where $\eta_{mm_k} = \frac{\xi_m + \xi_{m_k}}{2}$. From the mean value theorem, it follows that there exist some points a^k_m and $b^k_{m_k}$ which belong to the segments $[\xi_m, \eta_{mm_k}]$ and

$[\eta_{mm_k}, \xi_{m_k}]$, respectively, such that the estimate above becomes

$$
\begin{aligned}
TV_g(u_n) \leq & \sum_{k=1}^{3} \sum_{m \in \mathcal{I}_n^\circ} \left(\left| \langle \nabla u(a_m^k), e_k \rangle \right| \frac{h_n}{2} + \left| \langle \nabla u(b_{m_k}^k), e_k \rangle \right| \frac{h_n}{2} \right) \underbrace{\mathcal{H}^1 \left(\partial\Omega_m \cap \partial\Omega_{m_k} \right)}_{= \frac{h_n}{\sqrt{3}}} \\
& + \mathcal{O}(h_n) \left\| \nabla u \right\|_\infty \\
\leq & \sum_{k=1}^{3} \sum_{m \in \mathcal{I}_n^\circ} \frac{1}{2} \left(\left| \langle \nabla u(a_m^k), e_k \rangle \right| + \left| \langle \nabla u(b_{m_k}^k), e_k \rangle \right| \right) \underbrace{h_n^2/\sqrt{3}}_{= \frac{2}{3}|\Omega_k|} \\
& + \mathcal{O}(h_n) \left\| \nabla u \right\|_\infty .
\end{aligned}
$$

The next step is to take the limit $n \to \infty$. Recall that in this case $h_n \to 0$ and that the sum over all intermediary points $a_m^k, b_{m_k}^k$ of the hexagons $\{\Omega_m\}_{m \in \mathcal{I}_n^\circ}$ approximates the integral over the domain Ω. Hence, in the limit, the estimate above becomes

$$
\begin{aligned}
\lim_{n \to} TV_g(u_n) \leq & \sum_{k=1}^{3} \frac{2}{3} \int_\Omega \left| \langle \nabla u, e_k \rangle \right| \, dx \underset{(17)}{=} \frac{4}{3} \int_\Omega \left| \nabla u \right|_{\hexagon^T} \, dx \\
\underset{(22)}{\leq} & \frac{4}{3} \int_\Omega g \left(\nabla u \right) = \frac{4}{3} TV_g(u) .
\end{aligned}
$$

Recall that u is a continuous function and therefore, the last equality holds.

$\square$

Remark 4. *We are interested in how much rectangular or hexagonal grids influence the change of length when approximating a line segment. Let $L(\alpha)$ be a line from $(0,0)$ to $(\cos(\alpha), \sin(\alpha))$.*

Moreover we denote by $L_{\hexagon,n}(\alpha), L_{\square,n}(\alpha)$ the approximations of a line $L(\alpha)$ by hexagonal or squared grid boundaries (see Figure 5).

Set

$$
l_{\hexagon}(\alpha) := \lim_{n \to 0} \mathcal{H}^1(L_{\hexagon,n}(\alpha)) \qquad l_{\square}(\alpha) := \lim_{n \to 0} \mathcal{H}^1(L_{\square,n}(\alpha)) .
$$

where $\mathcal{H}^1$ is the 1-dimensional Hausdorff measure.

Figure 6 illustrates the error of the approximations with respect to the angle α.

We observe the following:

- *There is no angle α such that hexagonal grids preserve the length of a line, where as rectangular grids preserve the length of lines in the directions $\alpha = k\frac{\pi}{2}, k \in \mathbb{N}$.*

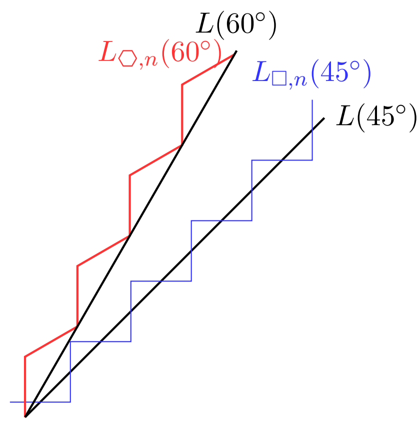


Figure 5: We approximate the lines $L(\alpha)$ on a hexagonal and a squared grid.

- *For rectangular grids, the worst approximation is for* $\alpha = \frac{\pi}{4} + k\frac{\pi}{2}$. *For hexagonal grids the worst approximation is for* $\alpha = \frac{\pi}{6} + k\frac{\pi}{3}$, *but the error in the length cause by the approximation is smaller than in the worst case with a rectangular grid.*

- *The difference of the length-error between best and worst case is much smaller for hexagonal grids compared to rectangular ones. Hence length-approximation on hexagonal grids is less anisotropic than on rectangular grids.*

To come back to our original problem, we want to find an estimate as in Theorem 1 (equation (7)) for the approximations on the hexagonal and the squared grid for the isotropic TV_{l^2} seminorm.

From the lower semi-continuity we obtain:

$$TV_{l^2}(u) \leq \liminf_{n\to\infty} TV_{l^2}(u_{n,\Box}) \qquad TV_{l^2}(u) \leq \liminf_{n\to\infty} TV_{l^2}(u_{n,\hexagon}) \; .$$

When approximating lines of length 1 on a hexagonal or a squared grid, the maximal error in length caused by the approximation on the regular grid is given by $\frac{4}{3}$ and $\sqrt{2}$ respectively, such that

$$\liminf_{n\to\infty} TV_{l^2}(u_{n,\Box}) \leq \sqrt{2} TV_{l^2}(u) \qquad \liminf_{n\to\infty} TV_{l^2}(u_{n,\hexagon}) \leq \frac{4}{3} TV_{l^2}(u) \; .$$

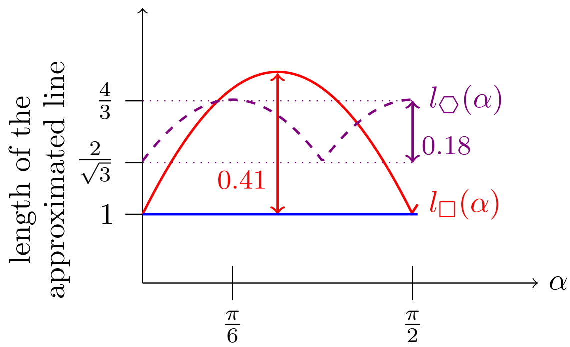


Figure 6: We see that the relative error of the TV_{l^2}-seminorm of an approximated line is smaller in the case where we use hexagonal grids. We never obtain an error-free approximation.

Hence we obtain

$$TV_{l_2}(u) \leq TV_{l_2}(u_{\square}) \leq \sqrt{2}\, TV_{l_2}(u)$$
$$TV_{l_2}(u) \leq TV_{l_2}(u_{\hexagon}) \leq \frac{4}{3} TV_{l_2}(u) \; .$$

This result states that the asymptotic error of the isotropic total variation caused by numerical approximation is much smaller for hexagonal grids.

4 Application to Image Denoising

This section is devoted to numerical experiments. We consider TV regularization for the problem of image denoising. In this case the operator F is the identity, and a given noisy image v^δ is denoised by finding a minimizer of

$$\mathcal{F}(u) = \left\| u - v^\delta \right\|^p + \alpha TV_{l^2}(u),$$

where $p \in \{1, 2\}$ is chosen according to the type of noise in v^δ.

We are mainly interested in comparing the performance of the model above for standard and hexagonally sampled images, respectively. Therefore, images are modelled as piecewise constant functions on a partition $\{\Omega_m\}_{m \in \mathcal{I}_n}$ of Ω into either squares or regular hexagons. In either case, $\mathcal{F}(u)$ can be written as

$$\mathcal{F}(u) = \sum_i \left|u^i - v^{\delta,i}\right|^p \mathcal{L}^2(\Omega_i) + \alpha \sum_{i<j} \left|u^i - u^j\right| \mathcal{H}^1(\Omega_i \cap \Omega_j),$$

where $u^i = u|_{\Omega_i}$, $v^{\delta,i} = v^\delta|_{\Omega_i}$ and $\mathcal{L}^2$ denotes the two-dimensional Lebesgue measure. In addition we only consider images quantized to the discrete set $\{0, 1, \ldots, 255\}$. To make the comparison fair, the two grids should have (approximately) the same amount of pixels, which means that the hexagons should have the same area as the squares. For a regular hexagon to have area h^2 it must have side length $h\sqrt{\frac{2}{3\sqrt{3}}} \approx 0.62h$.

Discrete energies like the one above can be minimized efficiently with so-called graph cut algorithms. Their main idea is to relate the minimization problem to a series of minimum cut problems on graphs, which in turn can be solved in low-order polynomial time. Good introductions to graph cuts can be found in [5, 6]. For their application to TV-based image restoration we refer to [22, 9, 15, 10, 19, 11]. Below we use the sequential algorithm from [15] adapted to hexagonal images together with the max-flow algorithm from [4].

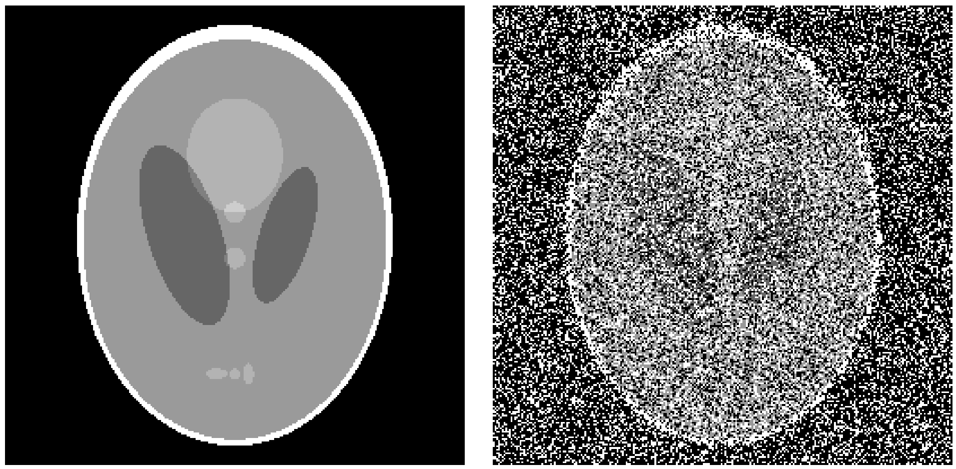
Apart from visually comparing our results, we use two quantitative performance measures: first, the L^1 distance between clean image $\bar{u}$ and restored image $u_n^{\alpha,\delta}$ divided by $|\mathcal{I}_n|$ (to make up for the fact that the square and hexagonal images do not have exactly the same amount of pixels)

$$\frac{1}{|\mathcal{I}_n|} \left\|\bar{u} - u_n^{\alpha,\delta}\right\|_{L^1}. \tag{27}$$

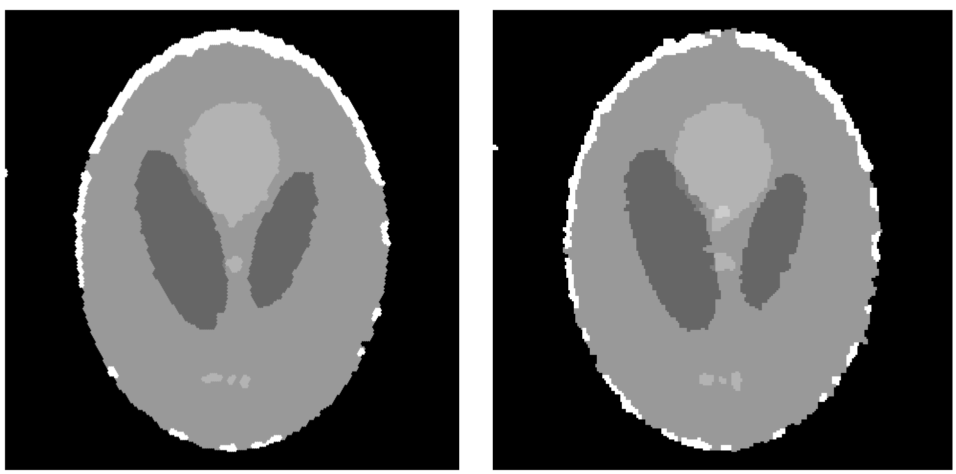
We also choose $\mathcal{L}^2(\Omega_i) = 1$ in the following experiments so that (27) can be interpreted as the mean absolute error per sampling point. As a second measure we use the ratio of correctly restored pixels

$$\frac{1}{|\mathcal{I}_n|} \left|\{i \in \mathcal{I}_n : \bar{u}^i = u_n^{\alpha,\delta,i}\}\right|. \tag{28}$$

Experiment no. 1 We chose the contrast enhanced Shepp-Logan phantom as a test image (Fig. 7a), and sampled it to a rectangular grid of size 256×256 and to a hexagonal one of approximately the same resolution. After adding 60% salt & pepper noise, the images were denoised with an L^1 fit-to-data term, which is known to be better suited than an L^2 fit-to-data term to remove this kind of noise. For different values of α, this procedure was repeated 50 times, in order to compensate for different realizations of noise. Denoised images are shown in Fig. 7, error plots are given in Fig. 8.



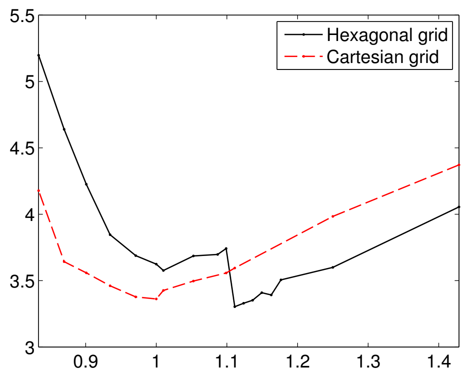
(a) Clean phantom. (b) Noisy phantom.



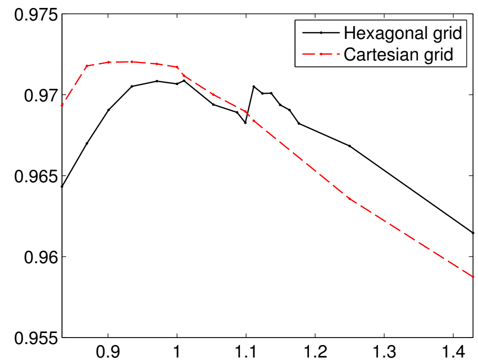
(c) Hexagonal grid, $\alpha = 1.1$. (d) Rectangular grid, $\alpha = 1$.

Figure 7: Results of Experiment no. 1. Clean and noisy image opposed to denoised images for different grids. The values of α have been chosen to match with optimal performance in terms of l^1 distance to ground truth (cf. Fig. 8a).

Experiment no. 2 Experiment no. 1 was repeated with a different image: a synthetic cosine with a resolution of 270×270 pixels, see Fig. 9a. The noise and L^1 data term are the same as before. Denoising results are presented in Figs. 9 and 10.



(a) Averaged l^1 distances plotted against α.



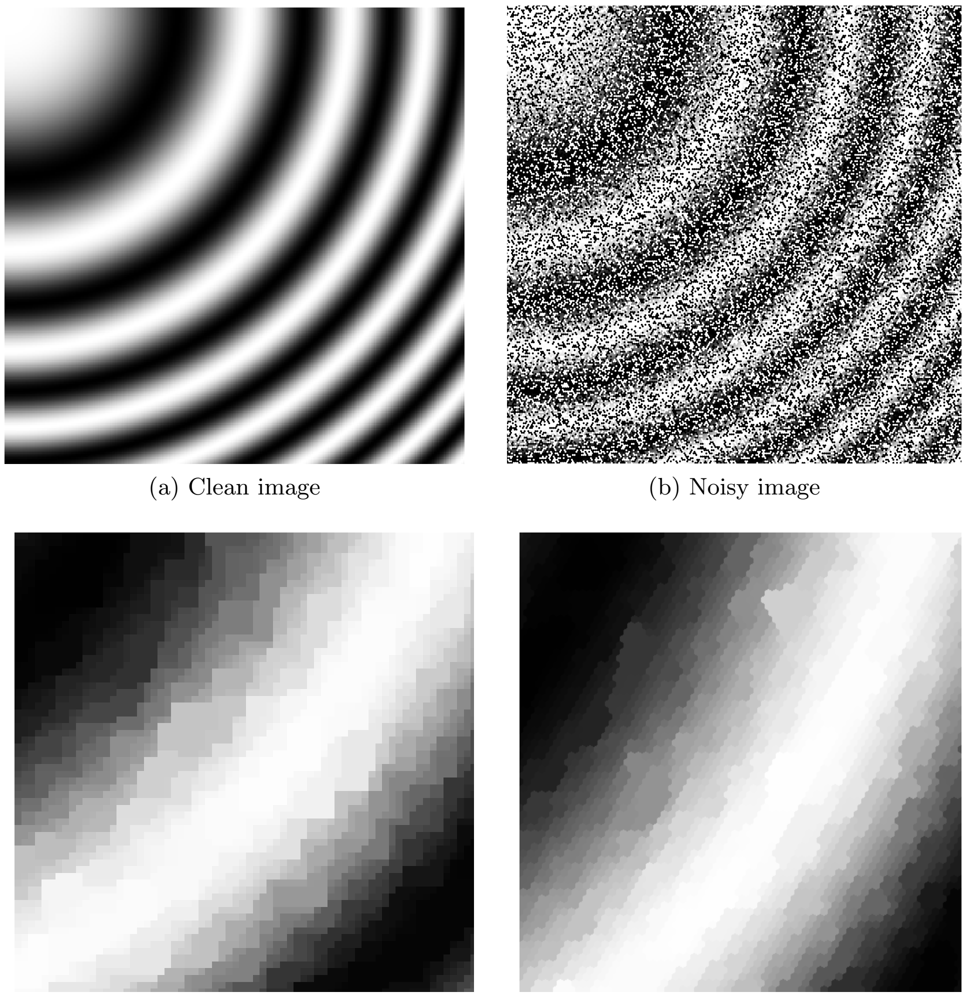
(b) Averaged ratios of correctly restored pixels plotted against α.

Figure 8: Results of Experiment no. 1. Performance measures (27) and (28) for different regularization parameters.

In contrast to the Shepp-Logan phantom the cosine test image consists entirely of smoothly varying intensity changes. TV-based denoising, however, is known to produce cartoon-like images, i.e. minimizers tend to be composed of subregions of more or less constant intensity separated by clear edges. This behaviour causes the so-called staircasing effect, which is clearly visible in Fig. 9c, where the restoration quality is additionally deteriorated by discretization errors, but also to some extent in 9d.

The error curves in Fig. 8 and even more so in Fig. 10 display a striking feature: they have significant discontinuities. This peculiarity of $L^1 - TV$ regularization has already been described in [13, 30]. The authors showed that, in general, the data fidelity of minimizers depends discontinuously on α, with at most countably many jumps. This behaviour, which is believed to be determined by the scales of image objects that rapidly merge at certain critical points, also manifests itself in measures (27) and (28).

Experiment no. 3 The first experiment was repeated with the phantom being contaminated by additive Gaussian noise of zero mean and a variance of 25.5, which corresponds to 10% of the range of grey values. Accordingly, an L^2 data term was employed to remove it. Since the $L^2 - TV$ model does not preserve the contrast, the number of correctly restored pixel intensities is close to zero for reasonable values of α. We therefore only plot measure (27) in Fig. 12. Denoising results are presented in Fig. 11.



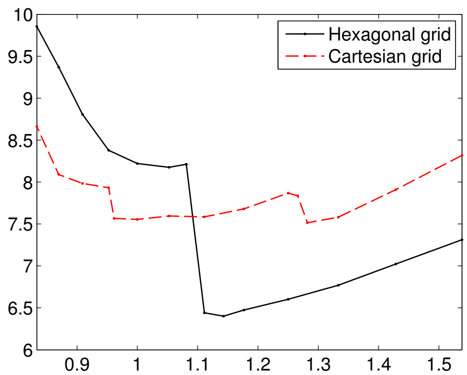
(a) Clean image (b) Noisy image

(c) Rectangular grid, $\alpha = 1.3$ (d) Hexagonal grid, $\alpha = 1.2$

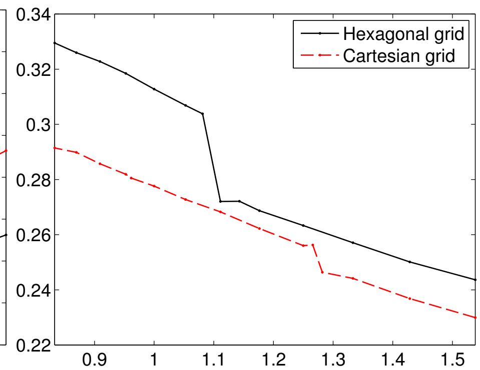
Figure 9: Results of Experiment no. 2. Clean and noisy cosine opposed to magnified portions of denoised image (cf. Fig. 9a) for the two grids. Values of α have again been chosen to roughly match with optimal performance according to Fig. 10a.

Experiment no. 4 Experiment no. 1 was repeated once more, but this time with a natural image, a 256×256 version of the camera man.

The clean image was resampled to the hexagonal grid by upsampling the original picture and downsampling it again to a hexagonal grid of approxi-



(a) Averaged l^1 distances plotted against α.



(b) Averaged ratios of correctly restored pixels plotted against α.

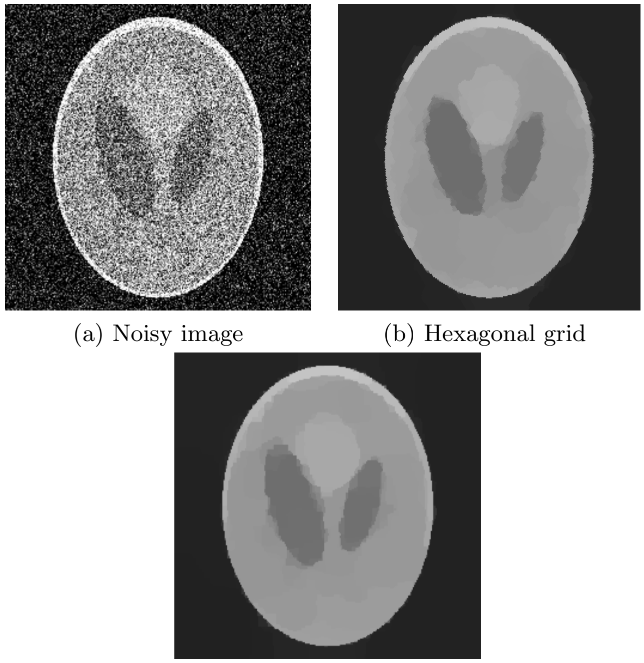
Figure 10: Results of Experiment no. 2. Performance measures (27) and (28) for different regularization parameters.

mately the same resolution, see Fig. 13. Downsampling was done by taking medians over hexagonal regions of the upsampled image.

After adding 60% salt & pepper noise, the images were denoised with an L^1 data term. Examples of denoised images are presented in Fig. 14, error plots can be found in Fig. 15. The latter figure seems to indicate that the hexagonal structure achieved significantly better results for all reasonable values of α. This has to be interpreted with care, as it is possible, that the better performance of the hexagonal grid is caused by the grid conversion process simplifying the image.

Experiment no. 5 Finally, we would like to compare the two different grids for one and the same noisy image. Recall that in the previous experiments the clean image was converted to the hexagonal grid first and only afterwards noise was added. The reason for this is that a generic conversion algorithm destroys the characteristics of the noise, so that an image corrupted with Gaussian noise, for example, would no longer be so after conversion.

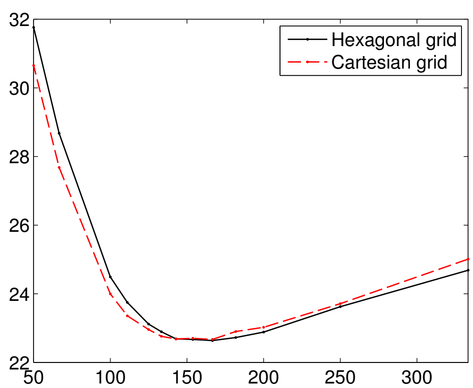
In a realistic application, however, we are usually given a noisy image on a rectangular grid. Then we would like to know if conversion to a hexagonal grid before denoising (and possible reconversion afterwards) leads to improved quality of the restored image. Therefore, we now fix a noisy standard image v^δ and convert it with an algorithm that transforms Gaussian



(a) Noisy image (b) Hexagonal grid

(c) Rectangular grid

Figure 11: Results of Experiment no. 3. Noisy image opposed to denoised images. The regularization parameter α was set to 250.



Averaged l^1 distances plotted against α.

Figure 12: Results of Experiment no. 3.

noise to Gaussian noise. For this purpose we implemented the conversion filter $h_\tau^{2,2}$ from [14]. Denoising results are depicted in Fig. 16. Apparently

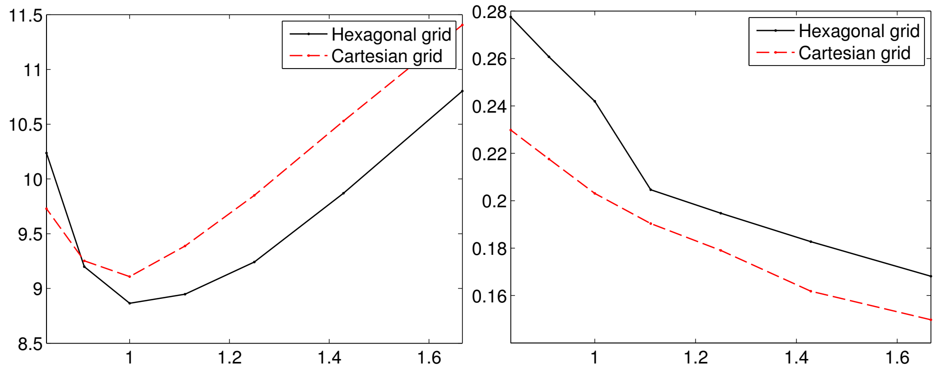
Figure 13: The two clean images used for Experiment no. 4. Original camera man image (right) and resampled to the hexagonal grid (left). It should be noted that the camera man features many edges that are perfectly aligned to Cartesian coordinates. Naturally those are better captured by a square pixel grid.

(a) Hexagonal grid

(b) Rectangular grid

Figure 14: Results of Experiment no. 4. Denoised images with $\alpha = 1.1$.

denoising on the hexagonal grid leads to less “blocky” images, that is, it suffers less from so-called metrication artefacts. On the downside, the reconversion to the square grid leads to overshooting artefacts near image



(a) Averaged l^1 distances plotted against α. (b) Averaged ratios of correctly restored pixels plotted against α.

Figure 15: Results of Experiment no. 4. Performance measures (27) and (28) for different regularization parameters.

edges.

Acknowledgements

The list of helpful references provided by Luminita Vese is gratefully acknowledged. C.P. and E.R. acknowledge the support by Austrian Science Fund, project J2970 (Schrödinger scholarship), project T644-N26 (Hertha Firnberg fellowship) and project FWF V82-N118 (Elise Richter fellowship), respectively. The work of E.R. has been partly done while affiliated with the Johannes Kepler University, Linz. The work of C.K. and O.S. is supported through the Austrian Science Fund (FWF) via projects S11704 and S10505.

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] S. Bartels. Total Variation Minimization with Finite Elements: Convergence and Iterative Solution. *SIAM J. Numerical Analysis*, 50(3):1162–1180, 2012.

(a) Noisy image.

(b) Denoised on Cartesian grid.

(c) Converted to hexagonal grid, then denoised.

(d) Converted, denoised, and converted back to Cartesian grid.

Figure 16: Results of Experiment no. 6. The regularization parameter α was set to 200.

[3] J.F. Bonans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, springer series in operations research and financial engineering edition, 2000.

[4] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE*

Trans. Pattern Anal. Mach. Intell., 26(9):1124–1137, September 2004.

[5] Y. Boykov and V. Kolmogorov. Basic Graph Cut Algorithms. In A. Blake, P. Kohli, and C. Rother, editors, *Markov Random Fields for Vision and Image Processing*, chapter 2, pages 31–49. MIT Press, 2011.

[6] Y. Boykov and O. Veksler. Graph Cuts in Vision and Graphics: Theories and Applications. In N. Paragios, Y. Chen, and O. Faugeras, editors, *Handbook of mathematical models in computer vision*, chapter 5, pages 79–96. Springer, New York, 2006.

[7] P. Bělík and M. Luskin. Approximation by piecewise constant functions in a BV metric. *Math. Models Methods Appl. Sci.*, 13(3):373–393, 2003. Dedicated to Jim Douglas, Jr. on the occasion of his 75th birthday.

[8] E. Casas, K. Kunisch, and C. Pola. Regularization by functions of bounded variation and applications to image enhancement. *Appl. Math. Optim.*, 40(2):229–257, 1999.

[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 Vision*, pages 136–152. Springer Berlin, Heidelberg, 2005.

[10] A. Chambolle and J. Darbon. On Total Variation Minimization and Surface Evolution Using Parametric Maximum Flows. *Int. J. Comput. Vision*, 84(3):288–307, April 2009.

[11] A. Chambolle and J. Darbon. A Parametric Maximum Flow Approach for Discrete Total Variation Regularization. In R. Lukac, editor, *Image Processing and Analysis with Graphs: Theory and Practice*, Digital Imaging and Computer Vision, chapter 4, pages 93–110. CRC Press / Taylor and Francis, 2012.

[12] A. Chambolle, S. Levine, and B. J. Lucier. An Upwind Finite-Difference Method for Total Variation-Based Image Smoothing. *SIAM J. Imaging Sciences*, 4(1):277–299, 2011.

[13] T. F. Chan and S. Esedoḡlu. Aspects of total variation regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65(5):1817–1837, 2005.

[14] L. Condat, D. Van De Ville, and B. Forster-Heinlein. Reversible, fast, and high-quality grid conversions. *IEEE Trans. Image Process.*, 17(5):679–693, May 2008.

[15] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation. Part I: Fast and exact optimization. *J. Math. Imaging Vision*, 26(3):261–276, 2006.

[16] F. Faille and M. Petrou. Bio-inspired, Invariant Image Reconstruction from Irregularly Placed Samples. In *BMVC'08*, pages –1–1, 2008.

[17] B. G. Fitzpatrick and S. L. Keeling. On approximation in total variation penalization for image reconstruction and inverse problems. *Numerical Functional Analysis and Optimization*, 18:941–958, 1997.

[18] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation.* Birkhäuser, Boston, 1984.

[19] D. Goldfarb and W. Yin. Parametric Maximum Flow Algorithms for Fast Total Variation Minimization. *SIAM J. Sci. Comput.*, 31(5):3712–3743, Oct 2009.

[20] X. He and W. Jia. Hexagonal structure for intelligent vision. In *Information and Communication Technologies, 2005. ICICT 2005. First International Conference on*, pages 52–64. IEEE, 2005.

[21] J. Heikenfeld, K. Zhou, E. Kreit, B. Raj, S. Yang, B. Sun, A. Milarcik, L. Clapp, and R. Schwartz. Electrofluidic displays using Young–Laplace transposition of brilliant pigment dispersions. *Nature Photonics*, 3(5):292–296, 2009.

[22] D. S. Hochbaum. An efficient algorithm for image segmentation, Markov random fields and related problems. *J. ACM*, 48(4):686–701 (electronic), 2001.

[23] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A Convergence Rates Result in Banach Spaces with Non-Smooth Operators. *Inverse Problems*, 23(3):987–1010, 2007.

[24] L. Middleton and J. Sivaswamy. Edge detection in a hexagonal-image processing framework. *Image and Vision Computing*, 19(14):1071–1081, 2001.

[25] A. Neubauer and O. Scherzer. Finite-dimensional approximation of Tikhonov regularized solutions of nonlinear ill-posed problems. *Numer. Funct. Anal. Optim.*, 11(1-2):85–99, 1990.

[26] C. Pöschl. *Tikhonov Regularization with General Residual Term.* PhD thesis, Leopold Franzens Universität Innsbruck, 2008.

[27] C. Pöschl, E. Resmerita, and O. Scherzer. Discretization of variational regularization in Banach spaces. *Inverse Problems*, 26(10):105017, 2010.

[28] E. Resmerita and O. Scherzer. Error estimates for non-quadratic regularization and the relation to enhancement. *Inverse Problems*, 22(3):801–814, 2006.

[29] 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.

[30] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*, volume 167 of *Applied Mathematical Sciences*. Springer Verlag, New York, 2009.

[31] T. Schuster, B. Kaltenbacher, B. Hofmann, and K.S. Kazimierski. *Regularization Methods in Banach Spaces.* Walter de Gruyter GmbH & Co. KG, radon series on computational and applied mathematics edition, 2012.

[32] J. Wang and B. J. Lucier. Error Bounds for Finite-Difference Methods for Rudin-Osher-Fatemi Image Smoothing. *SIAM J. Numerical Analysis*, 49(2):845–868, 2011.