

Invariant φ -minimal sets and total variation denoising on graphs

C. Kirisits, O. Scherzer, E. Setterqvist

RICAM-Report 2018-24

Invariant φ -minimal sets and total variation denoising on graphs

Clemens Kirisits^{*,1}, Otmar Scherzer^{†,1,2}, and Eric Setterqvist^{‡,2}

¹Johann Radon Institute for Computational and Applied Mathematics (RICAM),
Austrian Academy of Sciences, Linz, Austria

²Computational Science Center, University of Vienna, Austria

July 27, 2018

Abstract

Total variation flow, total variation regularization and the taut string algorithm are known to be equivalent filters for one-dimensional discrete signals. In addition, the filtered signal simultaneously minimizes a large number of convex functionals in a certain neighbourhood of the data. In this article we study the question to what extent this situation remains true in a more general setting, namely for data given on the vertices of a finite oriented graph and the total variation being $J(f) = \sum_{i,j} |f(v_i) - f(v_j)|$. Relying on recent results on invariant φ -minimal sets we prove that the minimizer to the corresponding Rudin-Osher-Fatemi (ROF) model on the graph has the same universal minimality property as in the one-dimensional setting. Interestingly, this property is lost, if J is replaced by the discrete isotropic total variation. Next, we relate the ROF minimizer to the solution of the gradient flow for J . It turns out that, in contrast to the one-dimensional setting, these two problems are not equivalent in general, but necessary and sufficient conditions for equivalence are available.

1 Introduction

It is a well known fact that for one-dimensional discrete data total variation (TV) regularization and TV flow are equivalent. More specifically, denote by

$$J(u) = \sum_{i=1}^{n-1} |u_i - u_{i+1}|$$

the total variation of $u \in \mathbb{R}^n$, and let $f \in \mathbb{R}^n$ and $\alpha > 0$ be given. Then, as was shown in [23], the minimizer u_α of the functional

$$\frac{1}{2} \|f - u\|_2^2 + \alpha J(u)$$

coincides with the solution to the Cauchy problem

$$\begin{aligned} u'(t) &\in -\partial J(u(t)), \\ u(0) &= f, \end{aligned}$$

*clemens.kirisits@ricam.oeaw.ac.at

†otmar.scherzer@univie.ac.at

‡eric.setterqvist@univie.ac.at

at time $t = \alpha$. On the other hand, it is known since [17] that u_α can also be obtained by means of the taut string algorithm, which reads as follows.

1. Identify the vector $f \in \mathbb{R}^n$ with a piecewise constant function on the unit interval and integrate it to obtain the linear spline F .
2. Find the “taut string” U_α , that is, the element of minimal graph length in a tube of width 2α around F with fixed ends:

$$U_\alpha = \arg \min \left\{ \int_0^1 \sqrt{1 + U'(x)} dx : \|U - F\|_\infty \leq \alpha, U(0) = F(0), U(1) = F(1) \right\}$$

3. Differentiate U_α to obtain u_α .

Problems which essentially can be modelled and solved by the taut string algorithm appear in diverse applications. Examples include production planning, see e.g. [18], and energy and information transmission, see e.g. [20] and [24]. A generalization of the taut string algorithm for higher-dimensional data is proposed in [11]. Further suggestions of generalizations of the taut string algorithm, in both discrete and continuous settings, can be found in [22, Chap. 4.4].

It turns out that the taut string does not only have minimal graph length, but actually minimizes *every* integral of the form

$$\int_0^1 \varphi(U'(x)) dx,$$

where $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary convex function, in the 2α -tube around F . Very recently, this intriguing situation was studied in greater generality in [14, 15]. The authors coined the term *invariant φ -minimal* for sets which, like the 2α -tube, have an element that simultaneously minimizes a large class of distances. In addition they characterized these sets in the discrete setting.

In this article we study total variation regularization and total variation flow in a setting that contains the one outlined above as a special case. More specifically, we consider data f as given on the vertices of a finite oriented graph $G = (V, E)$ together with the total variation

$$J(f) = \sup_{H \in \mathcal{B}_1} \langle f, \operatorname{div} H \rangle = \sum_{v,w} |f(v) - f(w)|. \quad (1)$$

Here, the sum runs over all adjacent pairs of vertices v, w , the set \mathcal{B}_1 is the unit ball with respect to the ℓ^∞ -norm on the space of edge functions $H : E \rightarrow \mathbb{R}$, and div denotes the graph divergence.

Our first result concerns the subdifferential of J . In Theorem 4 we prove that $\partial J(f)$ is an invariant φ -minimal set for every $f : V \rightarrow \mathbb{R}$. It is noteworthy that, as is shown in Example 1, this property is not shared by the discrete isotropic total variation, which for $f \in \mathbb{R}^{m \times n}$ reads

$$\sum_{i,j} \sqrt{(f_{i+1,j} - f_{i,j})^2 + (f_{i,j+1} - f_{i,j})^2} \quad (2)$$

and has been used, for instance, in [5].

Next we consider the Rudin-Osher-Fatemi (ROF) model [19] on the graph

$$\min_{u: V \rightarrow \mathbb{R}} \frac{1}{2} \|f - u\|_2^2 + \alpha J(u), \quad \alpha \geq 0. \quad (3)$$

From its dual formulation it follows that problem (3) is equivalent to

$$\min_{u \in f + \alpha \partial J(0)} \|u\|_2 \quad (4)$$

and since, by Theorem 4, the set $f + \alpha\partial J(0)$ is invariant φ -minimal, the ℓ^2 -norm in (4) can be replaced by $\sum_{v \in V} \psi(u(v))$ for any strictly convex function $\psi : \mathbb{R} \rightarrow \mathbb{R}$. More generally, as in the one-dimensional setting, the ROF minimizer u_α simultaneously minimizes $\sum_{v \in V} \varphi(u(v))$ over $f + \alpha\partial J(0)$ for every convex φ . See Proposition 5 and Theorem 6. We stress again that the minimizer of the isotropic ROF model does not have this property. Despite its anisotropy model (3) has been used extensively for the problem of image denoising. This is in part due to the fact that, in contrast to (2), J as given by (1) is submodular and for the minimization of submodular functions many efficient algorithms are available, for instance, graph cut algorithms [6, 7, 8, 12].

Finally, we examine the gradient flow for J and how it relates to the ROF model. Such relations in higher dimensional settings are the subject of recent investigations. In [3], discrete higher-dimensional variational methods and gradient flows for convex one-homogeneous functionals, thereby covering total variation, are investigated and sufficient conditions for their equivalence are provided. A sufficient condition for the equivalence of TV regularization and TV flow in the plane with ℓ^1 anisotropy is given in [16]. Considering the continuous setting with isotropic TV, it is shown in [13] that TV regularization and TV flow coincide for radial data but in general are non-equivalent.

Our results in this direction are the following. First, it follows from the fact that the ROF minimizer u_α solves (4) that

$$\|u_\alpha\|_2 \leq \|u(\alpha)\|_2$$

holds for all $\alpha \geq 0$. Here $u(\alpha)$ denotes the solution to the total variation flow at time $t = \alpha$. Second, we adapt [16, Thm. 10] to the graph setting, which gives a sufficient condition for equality of ROF and TV flow solutions. In addition, it shows that — apart from being equal for one-dimensional graphs — u_α and $u(\alpha)$ always coincide at least until the time derivative of the latter changes for the first time. Note that both ROF and flow solutions are continuous piecewise affine functions of α and t , respectively, cf. [3, Thm. 4.6] or Propositions 8 and 11 below. In Propositions 15 and 16 we discuss conditions that are necessary and sufficient for equality of u_α and $u(\alpha)$. Lastly, we consider a 3×3 Cartesian graph and track both solutions for a specific datum f . Apart from showing that on the graph ROF and flow solutions are not equivalent in general, it also shows that the jump sets of both are not monotonically decreasing with respect to α or t , respectively.

To summarize, problem (4) considered on the graph may be seen as a generalization of the taut string algorithm to higher dimensions since

- it is a minimization problem over the set $f + \alpha\partial J(0)$ which in turn can be seen as a higher-dimensional analogue of the set of derivatives of the elements in the 2α -tube around F ,
- the solution u_α minimizes $\sum_{v \in V} \varphi(u(v))$ over $f + \alpha\partial J(0)$ for any convex function φ ,
- u_α minimizes the corresponding ROF model (3). Further, if α is either sufficiently small or sufficiently large, then u_α is also the TV flow solution.

This article is organized as follows. In Section 2 we introduce the graph setting and the total variation J . We also collect some of its properties and briefly discuss the concept of invariant φ -minimal sets. Sections 3 and 4 are dedicated to the two main problems considered in this paper, that is, total variation regularization and total variation flow, respectively. In the final part of the paper, Section 5 that is, we compare the flow and ROF solutions.

2 The graph setting

2.1 The divergence operator and total variation on graphs

Consider a finite oriented graph (V, E) . That is, both the vertex set V and the edge set E are finite and if E contains the edge $e = (v, w)$, then it cannot contain (w, v) . The edge (v, w) is interpreted as directed from v to w . Let \mathbb{R}^V and \mathbb{R}^E denote the space of real-valued functions defined on the vertices and edges, respectively. We consider ℓ^p -norms on \mathbb{R}^V according to

$$\|u\|_p = \left(\sum_{v \in V} |u(v)|^p \right)^{1/p}$$

for $1 \leq p < \infty$ and

$$\|u\|_\infty = \max_{v \in V} |u(v)|.$$

Analogous ℓ^p -norms will be considered on \mathbb{R}^E . For $\alpha \geq 0$, denote by \mathcal{B}_α the following set of functions in \mathbb{R}^E :

$$\mathcal{B}_\alpha = \{F \in \mathbb{R}^E : \|F\|_\infty \leq \alpha\}. \quad (5)$$

Given $F \in \mathbb{R}^E$, introduce the mapping $\text{div} : \mathbb{R}^E \rightarrow \mathbb{R}^V$ according to

$$(\text{div } F)(v) = \sum_{w \in V : (w, v) \in E} F((w, v)) - \sum_{w \in V : (v, w) \in E} F((v, w)), \quad (6)$$

i.e. the divergence at the vertex v can be thought about as the sum of the flows on the incoming edges minus the sum of the flows on the outgoing edges. Introduce further the natural scalar product on \mathbb{R}^V according to

$$\langle g, h \rangle_{\mathbb{R}^V} = \sum_{v \in V} g(v)h(v).$$

The total variation of a function $u \in \mathbb{R}^V$ is defined as:

$$J(u) = \sup_{H \in \mathcal{B}_1} \langle u, \text{div } H \rangle_{\mathbb{R}^V}. \quad (7)$$

We can compute (7) using the following formula.

Proposition 1.

$$J(u) = \sum_{e=(v_i, v_j) \in E} |u(v_j) - u(v_i)|. \quad (8)$$

Proof. Recalling (6), we rewrite $J(u)$ according to:

$$\begin{aligned} J(u) &= \sup_{H \in \mathcal{B}_1} \langle u, \text{div } H \rangle_{\mathbb{R}^V} = \sup_{H \in \mathcal{B}_1} \sum_{v \in V} u(v)(\text{div } H)(v) = \\ &= \sup_{H \in \mathcal{B}_1} \sum_{e=(v_i, v_j) \in E} (u(v_j) - u(v_i))H((v_i, v_j)). \end{aligned} \quad (9)$$

Taking into account the definition (5) of \mathcal{B}_α , we see that the supremum of (9) is attained for $H \in \mathcal{B}_1$ if and only if

$$H((v_i, v_j)) \in \begin{cases} \{1\}, & u(v_i) < u(v_j), \\ [-1, 1], & u(v_i) = u(v_j), \\ \{-1\}, & u(v_i) > u(v_j), \end{cases} \quad (10)$$

giving

$$J(u) = \sum_{e=(v_i, v_j) \in E} |u(v_j) - u(v_i)|.$$

□

Remark 1. Proposition 1 shows that J is independent of the orientation of edges, even though the divergence is not. All subsequent results remain true regardless of edge orientation, and also apply to simple undirected graphs once each edge has been oriented arbitrarily.

2.2 Properties of the subdifferential

Definition 1. For every $u \in \mathbb{R}^V$ the subdifferential $\partial J(u)$ is defined as the set of all elements $u^* \in \mathbb{R}^V$ such that

$$\langle g - u, u^* \rangle_{\mathbb{R}^V} + J(u) \leq J(g) \quad \text{for all } g \in \mathbb{R}^V. \quad (11)$$

Since $\partial J(u)$ is a closed, convex and non-empty subset of \mathbb{R}^V , we can highlight one particular subgradient.

Definition 2. The element of minimal ℓ^2 -norm in $\partial J(u)$ will be referred to as the minimal section of $\partial J(u)$. It is denoted by $\partial^\circ J(u)$, that is,

$$\partial^\circ J(u) = \arg \min_{u^* \in \partial J(u)} \|u^*\|_2.$$

Given a set $\Gamma \subset \mathbb{R}^E$, we denote by $\text{div } \Gamma$ the image of Γ under the action of the operator div , that is,

$$\text{div } \Gamma = \{u \in \mathbb{R}^V : u = \text{div } H, H \in \Gamma\}.$$

It turns out that the divergence of the following subset of \mathcal{B}_1

$$\mathcal{B}_{1,u} = \left\{ H \in \mathbb{R}^E : H((v_i, v_j)) \in \begin{cases} \{1\}, & u(v_i) < u(v_j), \\ [-1, 1], & u(v_i) = u(v_j), \\ \{-1\}, & u(v_i) > u(v_j) \end{cases} \right\}, \quad (12)$$

for a given $u \in \mathbb{R}^V$, is equal to $\partial J(u)$.

Lemma 2.

1. $\partial J(0) = \text{div } \mathcal{B}_1$.
2. $\partial J(u) = \{u^* \in \partial J(0) : \langle u, u^* \rangle_{\mathbb{R}^V} = J(u)\}$ for all $u \in \mathbb{R}^V$.
3. $\partial J(u) = \text{div } \mathcal{B}_{1,u}$ for all $u \in \mathbb{R}^V$.

Proof. The functional J is the support function of the closed and convex set $\text{div } \mathcal{B}_1 \subset \mathbb{R}^V$ and therefore $\partial J(0) = \text{div } \mathcal{B}_1$.

Turn next to item 2. Take $u^* \in \partial J(u)$, i.e. u^* satisfies (11). Inserting $g = 0$ and $g = 2u$ in (11) give together that $\langle u, u^* \rangle_{\mathbb{R}^V} = J(u)$. This reduces (11) to

$$\langle g, u^* \rangle_{\mathbb{R}^V} \leq J(g) \quad \text{for all } g \in \mathbb{R}^V,$$

i.e. $u^* \in \partial J(0)$ (take $u = 0$ in the definition (11) of the subdifferential $\partial J(u)$). Hence, $\partial J(u) \subset \{u^* \in \partial J(0) : \langle u, u^* \rangle_{\mathbb{R}^V} = J(u)\}$. For the reverse direction, take $h \in \{u^* \in \partial J(0) : \langle u, u^* \rangle_{\mathbb{R}^V} = J(u)\}$. We have

$$\langle g - u, h \rangle_{\mathbb{R}^V} + J(u) = \langle g, h \rangle_{\mathbb{R}^V} \leq J(g) \quad \text{for all } g \in \mathbb{R}^V.$$

So, $h \in \partial J(u)$ and we conclude that

$$\partial J(u) = \{u^* \in \partial J(0) : \langle u, u^* \rangle_{\mathbb{R}^V} = J(u)\} = \{u^* \in \operatorname{div} \mathcal{B}_1 : \langle u, u^* \rangle_{\mathbb{R}^V} = J(u)\}.$$

Regarding item 3, recall that $J(u) = \langle u, \operatorname{div} H \rangle_{\mathbb{R}^V}$ for $H \in \mathcal{B}_1$ if and only if H satisfies (10), i.e. $H \in \mathcal{B}_{1,u}$. In view of item 2, it then is clear that $\partial J(u) = \operatorname{div} \mathcal{B}_{1,u}$. \square

Remark 2. 1. *Since, according to item 3 in Lemma 2, the set $\mathcal{B}_{1,u}$ only depends on $\operatorname{sgn}(u(v_i) - u(v_j))$ for every edge $(v_i, v_j) \in E$, we have*

$$\partial J(u) = \partial J(w),$$

if and only if

$$\operatorname{sgn}(u(v_i) - u(v_j)) = \operatorname{sgn}(w(v_i) - w(v_j))$$

for each $(v_i, v_j) \in E$.

2. *It now follows immediately that, if the subdifferentials of J at u and w coincide, then they also coincide for every convex combination of u and w . That is, $\partial J(u) = \partial J(w)$ implies $\partial J(\lambda u + (1 - \lambda)w) = \partial J(u)$ for every $\lambda \in (0, 1)$.*
3. *Lemma 2 also implies that the number of different subdifferentials of J is finite. In particular,*

$$|\{\partial J(u) : u \in \mathbb{R}^V\}| \leq 3|E|.$$

This must not be confused with the fact that for any given $u \in \mathbb{R}^V$ the subdifferential $\partial J(u)$ might have infinitely many elements.

We end this subsection by recalling the notion of invariant φ -minimal sets and show that the subdifferential $\partial J(u)$ is an example of such set.

Definition 3. *A set $\Omega \subset \mathbb{R}^n$ is called invariant φ -minimal if for every $a \in \mathbb{R}^n$ there exists an element $x_a \in \Omega$ such that*

$$\sum_{i=1}^n \varphi(x_{a,i} - a_i) \leq \sum_{i=1}^n \varphi(x_i - a_i) \quad (13)$$

holds for all $x \in \Omega$ and all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$.

In an invariant φ -minimal set Ω there exist for every $a \in \mathbb{R}^n$ an element $x_a \in \Omega$ which is the element of best approximation with respect to all generalized distances $d_\varphi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$d_\varphi(x, y) = \sum_{i=1}^n \varphi(x_i - y_i).$$

One interesting property of invariant φ -minimal sets is the following. By considering the particular convex function $\varphi(x) = |x|^p$, $1 \leq p < \infty$, in (13) we obtain

$$\sum_{i=1}^n |x_{a,i} - a_i|^p \leq \sum_{i=1}^n |x_i - a_i|^p$$

for all $x \in \Omega$. Taking the p -th root and including the case $p = \infty$, which follows by limiting arguments, shows that the element x_a satisfies

$$\|x_a - a\|_p \leq \|x - a\|_p \quad (14)$$

for all $x \in \Omega$ and $1 \leq p \leq \infty$. That is, x_a is an element of best approximation of a in Ω with respect to all ℓ^p -norms, $1 \leq p \leq \infty$. The inequality (14) will be used later on when analysing invariance properties of the ROF-minimizer for isotropic TV in Subsection 3.2.

In order to show that the subdifferential $\partial J(u)$ is invariant φ -minimal we recall the following result in [14]:

Theorem 3. *A bounded, closed and convex set $\Omega \subset \mathbb{R}^n$ is invariant φ -minimal if and only if Ω is a convex polytope where the affine hull of any face of Ω is a shifted subspace of \mathbb{R}^n spanned by vectors of the type $e_i - e_j$, $i \neq j$, where $\{e_i\}_{i=1}^n$ denotes the standard basis of \mathbb{R}^n .*

We are now ready to show:

Theorem 4. *The subdifferential $\partial J(u)$ is an invariant φ -minimal set.*

Proof. Consider first $\partial J(0)$. In [15, Thm. 2.4, Rem. 2.5], it was shown that the bounded, closed and convex set $\text{div } \mathcal{B}_\alpha \subset \mathbb{R}^n$ is invariant φ -minimal for all $\alpha \geq 0$. So $\partial J(0) = \text{div } \mathcal{B}_1$ is an invariant φ -minimal set. Further, $\partial J(0)$ is a convex polytope with the structure of its faces specified by Theorem 3.

Take now a general $u \in \mathbb{R}^V$. We have $\partial J(u) = S \cap \partial J(0)$ where $S = \{u^* \in \mathbb{R}^n : \langle u^*, u \rangle_{\mathbb{R}^V} = J(u)\}$, recall Lemma 2. Consider the halfspace $\widehat{S} = \{u^* \in \mathbb{R}^n : \langle u^*, u \rangle_{\mathbb{R}^V} \leq J(u)\}$ bounded by S . Note that (i) $\partial J(0) \subset \widehat{S}$ and (ii) $S \cap \partial J(0) = \partial J(u) \neq \emptyset$ (the supremum of $J(u)$ is attained). So, S is a supporting hyperplane of $\partial J(0)$ and since $\partial J(0)$ is a convex polytope it follows that $\partial J(u) = S \cap \partial J(0)$ is a face of $\partial J(0)$ and itself a convex polytope. Recall further that if F_2 is a face of F_1 which is a face of a convex polytope P , then F_2 is a face of P , see e.g. [10, Chap. 3.1, Thm. 5]. Hence, every face of $\partial J(u)$ has the structure specified by Theorem 3 and is therefore an invariant φ -minimal set. \square

Remark 3. *As $\partial J(u)$ is an invariant φ -minimal set, it follows that the minimal section $\partial^\circ J(u)$ not only has minimal ℓ^2 -norm in $\partial J(u)$, but satisfies*

$$\sum_{v \in V} \varphi(\partial^\circ J(u)(v)) = \min_{u^* \in \partial J(u)} \sum_{v \in V} \varphi(u^*(v))$$

for every convex function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$.

3 The ROF model on the graph

With the graph setting introduced, we now turn to an analogue of the Rudin-Osher-Fatemi image denoising model on \mathbb{R}^V . Given $f \in \mathbb{R}^V$ and $\alpha \geq 0$ we consider the following minimization problem:

$$\min_{u \in \mathbb{R}^V} \left(\frac{1}{2} \|f - u\|_2^2 + \alpha J(u) \right). \quad (15)$$

Throughout this article the unique solution to (15) will be denoted by u_α .

3.1 Dual formulation and an invariance property of the ROF-minimizer

Proposition 5. *For every $f \in \mathbb{R}^V$ and $\alpha \geq 0$ problem (15) is equivalent to*

$$\min_{u \in f + \text{div } \mathcal{B}_\alpha} \|u\|_2. \quad (16)$$

Proof. The corresponding dual problem of (15) can be expressed as

$$\min_{w \in \mathbb{R}^V} \left(\frac{1}{2} \|f - w\|_2^2 + (\alpha J)^*(w) \right), \quad (17)$$

where $(\alpha J)^*$ denotes the convex conjugate of αJ . For general results underlying the derivation of (17) and the optimality conditions (18) and (19) below, see [9, Chap. III, Prop. 4.1, Rem. 4.2]. Let u_α and w_α denote solutions to the primal problem (15) and the dual problem (17) respectively. The optimality conditions are

$$w_\alpha \in \partial(\alpha J)(u_\alpha) = \alpha \partial J(u_\alpha) \quad (18)$$

and

$$u_\alpha = f - w_\alpha. \quad (19)$$

As

$$\alpha J(u) = \alpha \sup_{H \in \mathcal{B}_1} \langle u, \operatorname{div} H \rangle_{\mathbb{R}^V} = \sup_{h \in \operatorname{div} \mathcal{B}_\alpha} \langle u, h \rangle_{\mathbb{R}^V}$$

the conjugate function $(\alpha J)^*$ can be derived as

$$(\alpha J)^*(g) = \begin{cases} 0, & g \in \operatorname{div} \mathcal{B}_\alpha, \\ +\infty, & g \notin \operatorname{div} \mathcal{B}_\alpha. \end{cases}$$

Taking into account the characterization of $(\alpha J)^*$ in the dual formulation (17) gives that

$$w_\alpha = \arg \min_{w \in \operatorname{div} \mathcal{B}_\alpha} \|f - w\|_{L^2}.$$

That is, w_α is the orthogonal projection of f on the closed and convex set $\operatorname{div} \mathcal{B}_\alpha$.

For the solution u_α of the ROF minimization problem (15), we obtain due to symmetry of $\operatorname{div} \mathcal{B}_\alpha$:

$$\|u_\alpha\|_2 = \|f - w_\alpha\|_2 = \min_{w \in \operatorname{div} \mathcal{B}_\alpha} \|f - w\|_2 = \min_{w \in \operatorname{div} \mathcal{B}_\alpha} \|f + w\|_2 = \min_{u \in f + \operatorname{div} \mathcal{B}_\alpha} \|u\|_2.$$

□

Remark 4. Proposition 5 implies that the ROF-minimizer u_α can be written as

$$u_\alpha = f + \operatorname{div} F_\alpha, \quad F_\alpha \in \mathcal{B}_\alpha. \quad (20)$$

The representation (20) will be used from time to time in the sequel.

Theorem 6. The ROF-minimizer u_α satisfies

$$\sum_{v \in V} \varphi(u_\alpha(v)) = \min_{u \in f + \operatorname{div} \mathcal{B}_\alpha} \sum_{v \in V} \varphi(u(v)) \quad (21)$$

for every convex function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$.

Proof. In [15, Thm. 2.4, Rem. 2.5] it was shown that $\operatorname{div} \mathcal{B}_\alpha$ is invariant φ -minimal, recall Definition 3. From the above derivation of the dual formulation, we know that u_α is the ℓ^2 -minimizer in the set $f + \operatorname{div} \mathcal{B}_\alpha$. Taken together, this gives that

$$\sum_{v \in V} \varphi(u_\alpha(v)) = \min_{u \in f + \operatorname{div} \mathcal{B}_\alpha} \sum_{v \in V} \varphi(u(v))$$

holds for every convex function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. □

3.2 Isotropic total variation on a Cartesian graph

In this subsection we specialize to graphs on a Cartesian grid representing a two-dimensional discrete image of $N \times N$ uniformly sized pixels with pixel intensity $u = (u(v_{ij}))_{i,j=1}^N$. The vertices represent the pixels and the edges connect adjacent pixels, an example with $N = 3$ is given in Figure 1. In this setting, we will provide some insight regarding the validity of Theorem 6 with respect to the specific form of the total variation.

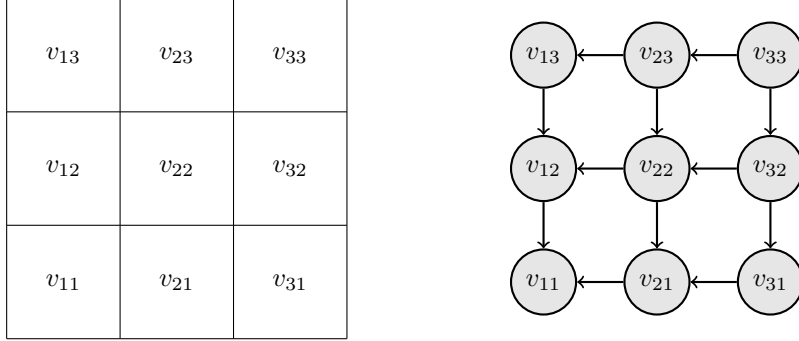


Figure 1: A digital image with uniformly sized pixels and its representation on a graph

For graphs on a Cartesian grid, the total variation (7) corresponds to discrete total variation with ℓ^1 anisotropy:

$$J(u) = \sum_{i=1}^{N-1} \sum_{j=1}^N |u(v_{(i+1)j}) - u(v_{ij})| + \sum_{i=1}^N \sum_{j=1}^{N-1} |u(v_{i(j+1)}) - u(v_{ij})|.$$

Define now the ball \mathcal{B}_α instead of (5) by

$$\mathcal{B}_\alpha = \left\{ F \in \mathbb{R}^E : \|F\| = \max_{1 \leq i,j \leq N} C_{ij}^F \leq \alpha \right\}, \quad (22)$$

where

$$C_{ij}^F = \sqrt{F((v_{(i+1)j}, v_{ij}))^2 + F((v_{i(j+1)}, v_{ij}))^2}, \quad i, j = 1, \dots, N-1.$$

For the right and upper boundaries we have

$$\begin{aligned} C_{iN}^F &= |F((v_{(i+1)N}, v_{iN}))|, \quad i = 1, \dots, N-1, \\ C_{Nj}^F &= |F((v_{N(j+1)}, v_{Nj}))|, \quad j = 1, \dots, N-1, \\ C_{NN}^F &= 0. \end{aligned}$$

The support function of $\text{div } \mathcal{B}_1$, with \mathcal{B}_1 given by (22), can be shown to be equal to

$$\begin{aligned} J(u) &= \sum_{i,j=1}^{N-1} \sqrt{|u(v_{(i+1)j}) - u(v_{ij})|^2 + |u(v_{i(j+1)}) - u(v_{ij})|^2} \\ &\quad + \sum_{i=1}^{N-1} |u(v_{(i+1)N}) - u(v_{iN})| + \sum_{j=1}^{N-1} |u(v_{N(j+1)}) - u(v_{Nj})|. \end{aligned} \quad (23)$$

This is the total variation used in e.g. [5] and can be viewed as one possible discretization of the isotropic total variation in the continuous setting. Note that it

is not obvious how to find an analogue of (23) and other isotropic total variation measures on general graphs without a specific structure such as the Cartesian grid.

An example is now provided which shows that Theorem 6 does not hold when considering the isotropic total variation (23).

Example 1. Consider a 2×2 Cartesian graph, analogous to the one shown in Figure 1, together with an image f with intensities $f(v_{11}) = 25$, $f(v_{21}) = 25 - \frac{3\sqrt{2}}{2} + 1$, $f(v_{12}) = 0$ and $f(v_{22}) = 10$. Recall that the inequality (14) for $a = 0$ must hold for u_α if (21) holds. Let $\alpha = 1$ and look for the ℓ^∞ -minimizer g in $f + \text{div } \mathcal{B}_1$ with minimal ℓ^2 -norm. Taking into account the structure of \mathcal{B}_1 , recall (22), it can be shown that $g = f + \text{div } G$ where $G \in \mathcal{B}_1$ is given by

$$\begin{aligned} G((v_{21}, v_{11})) &= G((v_{12}, v_{11})) = \frac{-1}{\sqrt{2}}, \\ G((v_{22}, v_{21})) &= -G((v_{22}, v_{12})) = -1, \end{aligned}$$

and therefore

$$g(v_{11}) = g(v_{21}) = 25 - \sqrt{2}, \quad g(v_{12}) = 1 + 1/\sqrt{2}, \quad g(v_{22}) = 10.$$

Consider next the edge function $H \in \mathcal{B}_1$ with

$$H((v_{21}, v_{11})) = 0, \quad H((v_{12}, v_{11})) = H((v_{22}, v_{21})) = -H((v_{22}, v_{12})) = -1,$$

and let $h = f + \text{div } H$. We then have

$$h(v_{11}) = 24, \quad h(v_{21}) = 25 - 3\sqrt{2}/2, \quad h(v_{12}) = 2, \quad h(v_{22}) = 10.$$

Clearly, $\|h\|_\infty > \|g\|_\infty$ and some further calculations show that $\|h\|_2 < \|g\|_2$. As the ROF-minimizer u_1 is the ℓ^2 -minimizer in $f + \text{div } \mathcal{B}_1$, it is clear that $u_1 \neq g$. Therefore, u_1 does not satisfy (14) for $a = 0$ and therefore does not satisfy the invariance property (21).

Remark 5. In the continuous setting it is known that an analogue of Theorem 6 holds for isotropic total variation, see [22, Thm. 4.46]. This can also be shown to be the case for total variation with ℓ^1 anisotropy in view of [15, Thm. 5.3, Rem. 5.4].

3.3 Further properties of the ROF-minimizer

In this subsection we study further properties of the ROF-minimizer u_α . We first give an auxiliary result.

Lemma 7. Let $0 \leq \alpha_1 < \alpha_2$. If $\partial J(u_{\alpha_1}) = \partial J(u_{\alpha_2})$, then for every $\alpha \in (\alpha_1, \alpha_2)$ the ROF-minimizer u_α is a convex combination of u_{α_1} and u_{α_2} . That is,

$$u_\alpha = \frac{\alpha_2 - \alpha}{\alpha_2 - \alpha_1} u_{\alpha_1} + \frac{\alpha - \alpha_1}{\alpha_2 - \alpha_1} u_{\alpha_2}, \quad \alpha_1 < \alpha < \alpha_2. \quad (24)$$

Proof. Denote the convex combination in (24) by $c(\alpha)$. It suffices to verify that $c(\alpha)$ satisfies the optimality conditions (18) and (19), that is, $f - c(\alpha) \in \alpha \partial J(c(\alpha))$. First, note that by item 2 in Remark 2 we have $\partial J(c(\alpha)) = \partial J(u_{\alpha_1})$. Next, let $w_{\alpha_i} = f - u_{\alpha_i}$, $i = 1, 2$. We now compute

$$\begin{aligned} \frac{f - c(\alpha)}{\alpha} &= \frac{1}{\alpha} \left[\frac{\alpha_2 - \alpha}{\alpha_2 - \alpha_1} w_{\alpha_1} + \frac{\alpha - \alpha_1}{\alpha_2 - \alpha_1} w_{\alpha_2} \right] \\ &= \frac{\alpha_1}{\alpha} \frac{\alpha_2 - \alpha}{\alpha_2 - \alpha_1} \frac{w_{\alpha_1}}{\alpha_1} + \frac{\alpha_2}{\alpha} \frac{\alpha - \alpha_1}{\alpha_2 - \alpha_1} \frac{w_{\alpha_2}}{\alpha_2}. \end{aligned}$$

It is straightforward to check that the last expression is a convex combination of w_{α_1}/α_1 and w_{α_2}/α_2 . By optimality of u_{α_i} and the assumption that $\partial J(u_{\alpha_1}) = \partial J(u_{\alpha_2})$, both w_{α_i}/α_i lie in the same convex set $\partial J(u_{\alpha_1})$. Therefore $(f - c(\alpha))/\alpha$ is in this set, too. We conclude that $c(\alpha)$ must be the ROF-minimizer u_α . \square

Using the above result we can show the following properties of the ROF-minimizer.

Proposition 8. *1. Problem (15) is mean-preserving, that is*

$$\sum_{v \in V} u_\alpha(v) = \sum_{v \in V} f(v) \quad \text{for all } \alpha \geq 0.$$

2. The function

$$\alpha \mapsto \|u_\alpha\|_2$$

is nonincreasing on $[0, \infty)$.

3. The solution u_α is a continuous piecewise affine function with respect to α . Its piecewise constant derivative $\partial_\alpha u_\alpha$ exists everywhere except for a finite number of values of $0 < \alpha_1 < \dots < \alpha_N < \infty$. In particular,

$$u_\alpha(v) = \frac{1}{|V|} \sum_{w \in V} f(w), \quad \text{for all } \alpha \geq \alpha_N \text{ and } v \in V.$$

Proof. 1. As $u_\alpha = f + \operatorname{div} F_\alpha$, $F_\alpha \in \mathcal{B}_\alpha$, it is clear that $\sum_{v \in V} u_\alpha(v) = \sum_{v \in V} f(v)$ for all $\alpha \geq 0$.

2. From the dual formulation of the ROF model, we know that u_α is the ℓ^2 -minimizer in the set $f + \operatorname{div} \mathcal{B}_\alpha$. Since $f + \operatorname{div} \mathcal{B}_{\alpha_1} \subset f + \operatorname{div} \mathcal{B}_{\alpha_2}$, $\alpha_1 \leq \alpha_2$, it then follows that $\alpha \mapsto \|u_\alpha\|_2$ is nonincreasing.

3. The continuity of u_α can be shown with arguments similar to [1, Thm. 3.2, Ex. 3.4]. The piecewise affine structure of u_α has been shown in [3, Thm. 4.6]. However, since our proof relies on different arguments, we choose to include it.

From Lemma 7 as well as Remark 2, items 2 and 3, we can derive two important facts. These two facts, combined with continuity of the map $\alpha \mapsto u_\alpha$, show that it must be piecewise affine on $[0, \infty)$. First, the subdifferential $\partial J(u_\alpha)$ can only change a finite number of times. Second, in intervals where it does not change, the minimizer u_α is an affine function of α .

Finally, consider u_α for $\alpha \geq \alpha_N$, where α_N is the last time $\partial J(u_\alpha)$ changes. Let \bar{f} denote the averaged initial image f , i.e.

$$\bar{f}(v) = \frac{1}{|V|} \sum_{w \in V} f(w), \quad \text{for all } v \in V. \quad (25)$$

For $\alpha \geq C$, where $C > 0$ is chosen large enough, it follows that $\bar{f} \in f + \operatorname{div} \mathcal{B}_\alpha$. Clearly, \bar{f} is the ℓ^2 -minimizer in $f + \operatorname{div} \mathcal{B}_\alpha$. Combined with the piecewise affine structure of u_α , we conclude that $u_\alpha = \bar{f}$ for $\alpha \geq \alpha_N$. \square

In general there might be many edge functions $F_\alpha \in \mathcal{B}_\alpha$ such that $u_\alpha = f + \operatorname{div} F_\alpha$, but a partial characterization can be provided. The following lemma will also be useful for the explicit computation of u_α in Sec. 5.2.

Lemma 9. *Suppose $e = (v_i, v_j) \in E$. If $u_\alpha(v_i) \neq u_\alpha(v_j)$, then*

$$F_\alpha((v_i, v_j)) = \alpha \operatorname{sgn}(u_\alpha(v_i) - u_\alpha(v_j)), \quad (26)$$

$$\partial_\alpha F_\alpha((v_i, v_j)) = \operatorname{sgn}(u_\alpha(v_i) - u_\alpha(v_j)). \quad (27)$$

Proof. The optimality condition (18) together with the equality $\partial J(u) = \operatorname{div} \mathcal{B}_{1,u}$ (recall Lemma 2, item 3) gives

$$F_\alpha((v_i, v_j)) = \begin{cases} -\alpha, & u_\alpha(v_i) < u_\alpha(v_j), \\ \alpha, & u_\alpha(v_i) > u_\alpha(v_j). \end{cases} \quad (28)$$

As the ROF-minimizer u_α depends continuously on α it follows that if $u_\alpha(v_i) < u_\alpha(v_j)$ then $u_{\alpha+t}(v_i) < u_{\alpha+t}(v_j)$, $t \in (-\varepsilon, \varepsilon)$, for $\varepsilon > 0$ small enough. Analogously, if $u_\alpha(v_i) > u_\alpha(v_j)$ then $u_{\alpha+t}(v_i) > u_{\alpha+t}(v_j)$, $t \in (-\varepsilon, \varepsilon)$, for $\varepsilon > 0$ small enough. In view of (28) we then have

$$\partial_\alpha F_\alpha((v_i, v_j)) = \begin{cases} -1, & u_\alpha(v_i) < u_\alpha(v_j), \\ 1, & u_\alpha(v_i) > u_\alpha(v_j). \end{cases}$$

□

4 The TV flow on the graph

In this section we consider the gradient flow associated to J . That is, given an initial datum $f : V \rightarrow \mathbb{R}$ we want to find a function $u : [0, \infty) \rightarrow \mathbb{R}^V$ that solves the Cauchy problem

$$\begin{aligned} u'(t) &\in -\partial J(u(t)) \quad \text{for a.e. } t > 0, \\ u(0) &= f. \end{aligned} \quad (29)$$

The statements in the next theorem follow from general results on nonlinear evolution equations and semigroup theory. See [2, Chap. 4] for a detailed treatment and [21, Sec. 2.1] for a brief introduction to the finite-dimensional setting.

Theorem 10. *Solutions to problem (29) have the following properties.*

1. *For every $f \in \mathbb{R}^V$ there is a unique solution and this solution depends continuously on f . In particular, if u_1 and u_2 are two solutions corresponding to initial conditions f_1 and f_2 , respectively, then*

$$\|u_1(t) - u_2(t)\|_2 \leq \|u_1(s) - u_2(s)\|_2 \quad \text{for all } t \geq s \geq 0.$$

2. *The solution u lies in $C([0, \infty), \mathbb{R}^V) \cap W^{1,\infty}([0, \infty), \mathbb{R}^V)$ and satisfies*

$$\|u'(t)\|_2 \leq \|\partial^\circ J(f)\|_2 \quad \text{for a.e. } t \geq 0.$$

3. *The solution is right differentiable everywhere. Its right derivative is right continuous and satisfies*

$$\frac{d^+}{dt} u(t) = -\partial^\circ J(u(t)) \quad \text{for all } t \geq 0.$$

4. *Define $S_t(f) = u(t)$. Then, for every $f \in \mathbb{R}^V$, we have*

$$S_t(S_s(f)) = S_{t+s}(f) \quad \text{for all } t, s \geq 0.$$

5. *The function $u(t) \in \mathbb{R}^V$ converges to a minimizer of J as $t \rightarrow \infty$.*

In fact, Theorem 10 holds true for any convex real-valued functional which admits a minimizer on \mathbb{R}^V . The solution to the TV flow in addition has the following properties.

Proposition 11. 1. *Problem (29) is mean-preserving, that is,*

$$\sum_{v \in V} u(t)(v) = \sum_{v \in V} f(v) \quad \text{for all } t \geq 0.$$

2. *The functions*

$$t \mapsto \|u(t)\|_2 \quad \text{and} \quad t \mapsto \left\| \frac{d^+}{dt} u(t) \right\|_2$$

are nonincreasing on $[0, \infty)$.

3. *The solution u is piecewise affine with respect to t . More specifically, the derivative $u'(t)$ does not exist for only a finite number of times $0 < t_1 < \dots < t_M$ and it is constant in between. It follows that a stationary solution is reached in finite time:*

$$u(t)(v) = \frac{1}{|V|} \sum_{w \in V} f(w) \quad \text{for all } t \geq t_M \text{ and } v \in V.$$

Proof. 1. Since the subdifferential of J consists entirely of divergences of edge functions, for a.e. $t \geq 0$ there is a $H(t) \in \mathbb{R}^E$ such that

$$u'(t) = -\operatorname{div} H(t). \quad (30)$$

Summing this equation over all $v \in V$ and using the fact that $\sum_{v \in V} \operatorname{div} H(v)$ vanishes for every $H \in \mathbb{R}^E$ gives

$$\frac{d}{dt} \sum_{v \in V} u(t)(v) = 0 \quad \text{for a.e. } t \geq 0$$

and the assertion follows.

2. Take the inner product of equation (30) with $u(t)$ and notice that, from the characterization of the subdifferential in Lemma 2, it follows that

$$\langle u(t), \operatorname{div} H(t) \rangle_{\mathbb{R}^V} = J(u(t)).$$

Therefore

$$-J(u(t)) = -\langle u(t), \operatorname{div} H(t) \rangle_{\mathbb{R}^V} = \langle u(t), u'(t) \rangle_{\mathbb{R}^V} = \frac{1}{2} \frac{d}{dt} \|u(t)\|_2^2$$

for a.e. $t > 0$, which shows that $t \mapsto \|u(t)\|_2$ is nonincreasing. That $t \mapsto \|d^+ u(t)/dt\|_2$ is nonincreasing as well, follows from items 2 and 4 in Theorem 10.

3. As for the ROF-minimizer the piecewise affine behaviour has been shown in [3, Thm. 4.6]. Our proof uses different arguments. According to item 3 in Remark 2 the number of different values the right derivative of u can take is finite. Since $d^+ u/dt$ is also right continuous, there must be an $\epsilon > 0$ for every $t_0 \geq 0$ such that

$$\frac{d^+}{dt} u(t) = -\partial^\circ J(u(t_0)) \quad \text{for all } t \in [t_0, t_0 + \epsilon)$$

with $d^+u/dt = u'$ on $(t_0, t_0 + \epsilon)$. This proves that $t \mapsto u(t)$ is piecewise affine on $[0, \infty)$.

That d^+u/dt only changes a finite number of times follows from the fact that, if it changes, then its norm becomes strictly smaller. To see this let $\hat{t} > 0$ and assume that $d^+u(t)/dt \equiv c$ is constant on $(\hat{t} - \epsilon, \hat{t})$ for some $\epsilon > 0$ and that $d^+u(\hat{t})/dt \neq c$. We now have

$$J(u(\hat{t})) = \lim_{t \rightarrow \hat{t}^-} J(u(t)) = \lim_{t \rightarrow \hat{t}^-} \langle u(t), -c \rangle = \langle u(\hat{t}), -c \rangle,$$

and therefore $-c \in \partial J(u(\hat{t}))$. However, since $-c = \partial^\circ J(u(t))$ for $t \in (\hat{t} - \epsilon, \hat{t})$ and the minimal section is the unique element of minimal norm in the subdifferential, we must have $\|d^+u(t)/dt\|_2 > \|d^+u(\hat{t})/dt\|_2$.

Thus $t \mapsto u(t)$ is a continuous piecewise affine function with a finite number of slope changes. Since, by item 5 in Theorem 10, $u(t)$ is convergent, it must reach its limit in finite time. Due to mean preservation, this limit has to be the averaged initial datum. □

Remark 6. *The solution to the TV flow can be represented in a way similar to the ROF-minimizer*

$$u(t) = f + \operatorname{div}(F(t)), \quad (31)$$

where $F(t) = -\int_0^t H(s)ds \in \mathcal{B}_t$ and $H(s) \in \mathcal{B}_{1,u(s)}$. Compare Remark 4. In this case it follows directly from integration of (29) and the characterization of the subdifferential in item 3 of Lemma 2. In addition F satisfies

$$F'(t)((v_i, v_j)) = \operatorname{sgn}(u(t)(v_i) - u(t)(v_j)), \quad (32)$$

whenever $u(t)(v_i) \neq u(t)(v_j)$ holds, which is the same as (27) in Lemma 9 for the ROF-minimizer. However, property (26) does not hold in general for the TV flow. Equations (31) and (32) will be used in Section 5.2 for computing $u(t)$ for a specific datum f .

5 Comparison of TV regularization and TV flow

In this section, various conditions for the equivalence of TV regularization and TV flow on general graphs are provided and analysed. Then a study is done on a particular graph where the solutions of the TV regularization and the TV flow are explicitly computed and compared.

5.1 Equivalence of TV regularization and TV flow

We first provide a result concerning the norms of the solutions of the TV regularization and the TV flow. Recall that \bar{f} denotes the averaged initial datum f , see (25).

Proposition 12. *For every $\alpha > 0$ let u_α and $u(\alpha)$ be the ROF and TV flow solutions, respectively, both corresponding to the same datum $f \in \mathbb{R}^V$. They satisfy*

$$\|\bar{f}\|_2 \leq \|u_\alpha\|_2 \leq \|u(\alpha)\|_2 \leq \|f\|_2, \quad \text{for all } \alpha > 0.$$

It follows that in general u_α reaches \bar{f} before $u(t)$, that is, $\alpha_N \leq t_M$, see Propositions 8 and 11.

Proof. Both $\|u_\alpha\|_2$ and $\|u(\alpha)\|_2$ are nonincreasing functions of α , recall property 2 in Propositions 8 and 11, and therefore bounded from above by $\|f\|_2$. On the other hand, due to mean preservation, recall property 1 in Propositions 8 and 11, they are bounded from below by $\|\bar{f}\|_2$. It remains to show that $\|u_\alpha\|_2 \leq \|u(\alpha)\|_2$. To see this, observe that both u_α and $u(\alpha)$ lie in $f + \text{div } \mathcal{B}_\alpha$ with u_α being the element of minimal norm in this set according to (16). \square

The next proposition, which is an adaptation of [16, Thm. 10] to the graph setting, gives a sufficient condition for equality of ROF and TV flow solutions.

Proposition 13. *Let $f \in \mathbb{R}^V$ and let u be the corresponding solution to the TV flow (29). If there is an $\alpha > 0$ such that*

$$-\langle u'(t), u(\alpha) \rangle_{\mathbb{R}^V} = J(u(\alpha)) \quad \text{for a.e. } t \in (0, \alpha), \quad (33)$$

then $u(\alpha) = u_\alpha$.

Proof. The proof is analogous to the one of [16, Thm. 10]. We therefore omit it. \square

Corollary 14. *The ROF and TV flow solutions always coincide in the interval $[0, t_1]$, where $t_1 > 0$ is, as discussed in Proposition 11, the first time the derivative of the latter changes.*

Proof. Recall that the flow solution satisfies

$$-u'(t) = \partial^\circ J(f) \in \partial J(u(t)), \quad t \in [0, t_1].$$

This implies by Lemma 2 that

$$\langle \partial^\circ J(f), u(t) \rangle = J(u(t)), \quad t \in [0, t_1],$$

and since u is continuous in t

$$\langle \partial^\circ J(f), u(t_1) \rangle = J(u(t_1)).$$

Therefore condition (33) is satisfied for every $\alpha \in [0, t_1]$. \square

Remark 7. *Combining Corollary 14 with Theorem 6 shows that the TV flow solution minimizes*

$$\sum_{v \in V} \varphi(u(t)(v))$$

over $f + \text{div } \mathcal{B}_t$ for every convex φ and at least for every $t \in [0, t_1]$. Thus, the situation encountered for one-dimensional signals, namely equivalence of TV flow and TV regularization together with φ -minimality of the solution (recall the first paragraph of Section 1), can be recovered in the more general setting of finite oriented graphs at least on the intervals $[0, t_1] \cup [t_M, \infty)$.

Remark 8. *Condition (33) holds true, given any $\alpha > 0$, for graphs of the type displayed in Figure 2 corresponding to one-dimensional space-discrete signals. This follows directly from the inclusion*

$$\partial J(u(\alpha_1)) \subset \partial J(u(\alpha_2)), \quad \alpha_1 \leq \alpha_2, \quad (34)$$

which applies in this setting. The derivation of (34) can be done with the following arguments. Consider a pair of adjacent pixels v_i and v_{i+1} . In [23, Prop. 4.1], it is shown that if $u(\alpha)(v_i) = u(\alpha)(v_{i+1})$ then $u(\beta)(v_i) = u(\beta)(v_{i+1})$ for any $\beta \geq \alpha$. This fact together with the continuity of $u(\alpha)$ implies that $\mathcal{B}_{1, u(\alpha_1)} \subset \mathcal{B}_{1, u(\alpha_2)}$ (recall (12) for the definition of $\mathcal{B}_{1, u}$). Taking into account the characterization of the subdifferential given by item 3 in Lemma 2, (34) then follows.

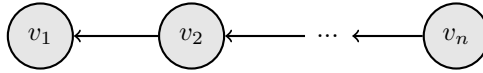


Figure 2: Graph corresponding to a one-dimensional space-discrete signal with n pixels.

Two necessary and sufficient conditions, in a general graph setting, concerning the equivalence of TV flow and TV regularization will now be provided. The first one is formulated in terms of the TV flow $u(\alpha)$ for a given $\alpha > 0$.

Proposition 15. *Given $\alpha > 0$,*

$$u(\alpha) = u_\alpha$$

if and only if

$$-\frac{1}{\alpha} \int_0^\alpha u'(t) dt \in \partial J(u(\alpha)). \quad (35)$$

Proof. We can express $u(\alpha) = f + \int_0^\alpha u'(t) dt$. Recalling the optimality conditions (18) and (19) for the ROF-minimizer u_α , it follows that $u(\alpha) = u_\alpha$ if and only if $-\frac{1}{\alpha} \int_0^\alpha u'(t) dt \in \partial J(u(\alpha))$. \square

Remark 9. *The above proposition gives that $u(\alpha) = u_\alpha$ if and only if the average time derivative $\frac{1}{\alpha} \int_0^\alpha u'(t) dt$ is in $-\partial J(u(\alpha))$. Compare with the pointwise inclusion $u'(t) \in -\partial J(u(t))$ which holds for a.e. $t > 0$. Note further that condition (35) is strictly weaker than (33).*

For the entire interval $\alpha \geq 0$, we have the following necessary and sufficient condition for equivalence. It is formulated in terms of the ROF solution operator defined by $T_\alpha(f) = u_\alpha$.

Proposition 16. *The identity*

$$u_\alpha = u(\alpha)$$

holds for all $\alpha \geq 0$, if and only if

$$T_t(T_s(f)) = T_{t+s}(f) \quad (36)$$

for all $t, s \geq 0$.

Proof. Let $u_\alpha = u(\alpha)$ for all $\alpha \geq 0$. It then follows from property 4 in Theorem 10 that $T_t(T_s(f)) = T_{t+s}(f)$ for all $t, s \geq 0$.

Start now with the assumption $T_t(T_s(f)) = T_{t+s}(f)$ for all $t, s \geq 0$. As the TV flow has an analogous property and the solutions to TV regularization and TV flow always coincide for the interval $[0, t_1]$, recall Corollary 14, it is then immediate that they coincide for all $\alpha \geq 0$. \square

5.2 TV denoising on a particular graph

In this subsection TV regularization and TV flow are studied for the particular graph $G = (V, E)$ and datum f given by Figure 3 on the interval $0 \leq \alpha, t \leq 4$.

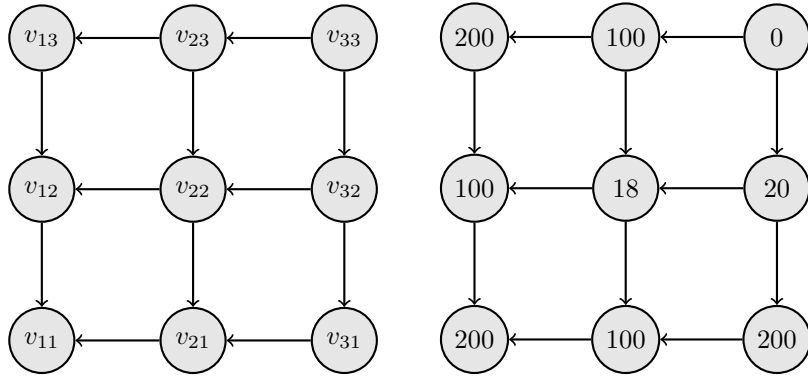


Figure 3: Graph structure (left) and pixel intensities (right) of the initial datum f .

5.2.1 TV regularization

For the ROF-minimizer $u_\alpha = f + \operatorname{div} F_\alpha$, where $F_\alpha \in \mathcal{B}_\alpha$, we have for any $v \in V$,

$$f(v) - \deg(v)\alpha \leq u_\alpha(v) \leq f(v) + \deg(v)\alpha, \quad (37)$$

where $\deg(v)$ denotes the degree of v , that is, the number of edges incident to v . Using (37) it is straightforward to show that

$$\operatorname{sgn}(u_\alpha(v_{ij}) - u_\alpha(v_{kl})) = \operatorname{sgn}(f(v_{ij}) - f(v_{kl})) \in \{\pm 1\} \quad (38)$$

for every edge except (v_{32}, v_{22}) and $0 \leq \alpha \leq 4$. From Lemma 9, it then follows that

$$F_\alpha((v_{ij}, v_{kl})) = \alpha \operatorname{sgn}(f(v_{ij}) - f(v_{kl})),$$

for all $(v_{ij}, v_{kl}) \in E \setminus \{(v_{32}, v_{22})\}$ and $0 \leq \alpha \leq 4$.

Consider now the special edge (v_{32}, v_{22}) on $0 \leq \alpha \leq 4$. Using the knowledge of F_α on the other edges, $u_\alpha(v_{22})$ and $u_\alpha(v_{32})$ are given by

$$\begin{aligned} u_\alpha(v_{22}) &= f(v_{22}) + F_\alpha((v_{32}, v_{22})) + F_\alpha((v_{23}, v_{22})) - F_\alpha((v_{22}, v_{12})) - F_\alpha((v_{22}, v_{21})) \\ &= 18 + F_\alpha((v_{32}, v_{22})) + 3\alpha, \end{aligned}$$

and

$$\begin{aligned} u_\alpha(v_{32}) &= f(v_{32}) - F_\alpha((v_{32}, v_{22})) + F_\alpha((v_{33}, v_{32})) - F_\alpha((v_{32}, v_{31})) \\ &= 20 - F_\alpha((v_{32}, v_{22})), \end{aligned}$$

on $0 \leq \alpha \leq 4$. Recall further that u_α is the ℓ^2 -minimizer in the set $f + \operatorname{div} \mathcal{B}_\alpha$, cf. Proposition 5, and that $F_\alpha((v_{32}, v_{22}))$ only appears in the terms $u_\alpha(v_{22})$ and $u_\alpha(v_{32})$. Minimizing $(u_\alpha(v_{22}))^2 + (u_\alpha(v_{32}))^2$ subject to the constraint $F_\alpha((v_{32}, v_{22})) \in [-\alpha, \alpha]$ then gives

$$F_\alpha((v_{32}, v_{22})) = \begin{cases} \alpha, & 0 \leq \alpha \leq 2/5, \\ (2 - 3\alpha)/2, & 2/5 \leq \alpha \leq 2, \\ -\alpha, & 2 \leq \alpha \leq 4. \end{cases}$$

The function F_α is now determined on all edges and u_α can then be computed on all vertices. The results can be seen in Figure 4.

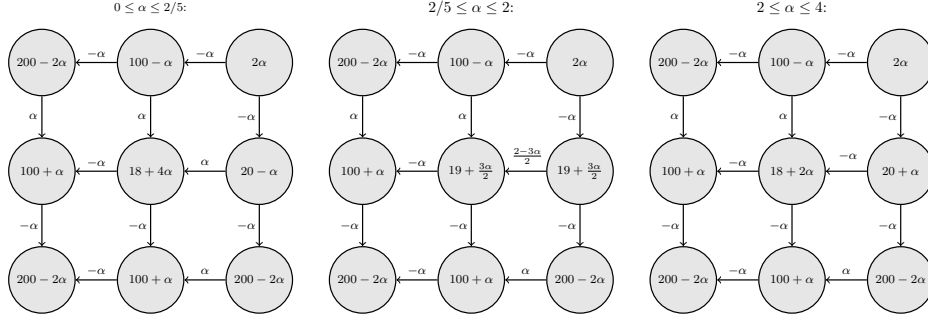


Figure 4: The evolution of the ROF-minimizer u_α (defined on the vertices) and the function F_α (defined on the edges) on the interval $0 \leq \alpha \leq 4$.

5.2.2 TV flow

For the TV flow $u(t) = f + \operatorname{div}(F(t))$, where $F(t) \in \mathcal{B}_t$ according to Remark 6, we have an analogous inequality to (37),

$$f(v) - \operatorname{deg}(v)t \leq u(t)(v) \leq f(v) + \operatorname{deg}(v)t \quad (39)$$

for all $v \in V$. Using (39), we obtain

$$\operatorname{sgn}(u(t)(v_{ij}) - u(t)(v_{kl})) = \operatorname{sgn}(f(v_{ij}) - f(v_{kl})) \in \{\pm 1\} \quad (40)$$

for any edge $(v_{ij}, v_{kl}) \in E \setminus \{(v_{32}, v_{22})\}$ and $0 \leq t \leq 4$. Recall Remark 6 and conclude from (40) that

$$F'(t)((v_{ij}, v_{kl})) = \operatorname{sgn}(f(v_{ij}) - f(v_{kl}))$$

for all $(v_{ij}, v_{kl}) \in E \setminus \{(v_{32}, v_{22})\}$ and $0 < t < 4$. Hence,

$$F(t)((v_{ij}, v_{kl})) = t \operatorname{sgn}(f(v_{ij}) - f(v_{kl})),$$

for all $(v_{ij}, v_{kl}) \in E \setminus \{(v_{32}, v_{22})\}$ and $0 \leq t \leq 4$.

Turn next to the computation of $F(t)((v_{32}, v_{22}))$ on $0 \leq t \leq 4$. Knowledge of $F(t)$ on the other edges gives

$$u(t)(v_{22}) = 18 + 3t + F(t)((v_{32}, v_{22})), \quad (41)$$

and

$$u(t)(v_{32}) = 20 - F(t)((v_{32}, v_{22})), \quad (42)$$

on $0 \leq t \leq 4$. From (41) and (42), together with $F(t) \in \mathcal{B}_t$, follow the inequalities

$$u(t)(v_{22}) \leq 18 + 4t < 20 - t \leq u(t)(v_{32}), \quad 0 \leq t < 2/5.$$

Recalling Remark 6, we then obtain

$$F'(t)((v_{32}, v_{22})) = \operatorname{sgn}(u(t)(v_{32}) - u(t)(v_{22})) = 1, \quad 0 < t < 2/5,$$

and therefore

$$F(t)((v_{32}, v_{22})) = t, \quad 0 \leq t \leq 2/5.$$

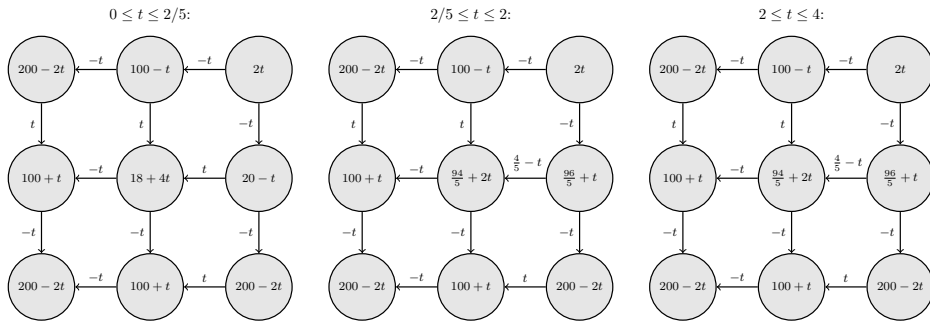


Figure 5: The evolution of the TV flow $u(t)$ (defined on the vertices) and the function $F(t)$ (defined on the edges) on the interval $0 \leq t \leq 4$.

Turn now to the interval $2/5 \leq t \leq 4$. Note that F can be written as $F(t) = -\int_0^t H(s)ds$, where $H(s) \in \mathcal{B}_{1,u(s)}$, recall Remark 6. We can now estimate

$$\begin{aligned} F(t)((v_{32}, v_{22})) &= F(2/5)((v_{32}, v_{22})) - \int_{2/5}^t H(s)((v_{32}, v_{22}))ds \\ &\geq 2/5 - (t - 2/5) \\ &= 4/5 - t \end{aligned}$$

for $t \geq 2/5$. This inequality together with (41) and (42) give

$$u(t)(v_{32}) \leq 96/5 + t < 94/5 + 2t \leq u(t)(v_{22}), \quad 2/5 < t \leq 4.$$

We conclude, recalling Remark 6 again, that

$$F'(t)((v_{32}, v_{22})) = \text{sgn}(u(t)(v_{32}) - u(t)(v_{22})) = -1, \quad 2/5 < t < 4.$$

From the continuity of F then follows

$$F(t)((v_{32}, v_{22})) = 4/5 - t, \quad 2/5 \leq t \leq 4.$$

The function $F(t)$ is now determined on all edges which enables the computation of $u(t)$ on all vertices. The results can be seen in Figure 5.

5.2.3 Comments

We now comment on the findings which are summarized in Figures 4 and 5. First, note that $u(\alpha) = u_\alpha$ for $0 \leq \alpha \leq 2/5$ but $u(\alpha) \neq u_\alpha$ for $2/5 < \alpha \leq 4$. So in higher-dimensional graph settings TV regularization and TV flow are not equivalent. In particular, as the provided example is a graph on a two-dimensional Cartesian grid, this implies that discrete TV regularization and discrete TV flow with ℓ^1 anisotropy are distinct image filters.

The necessary and sufficient condition for equivalence of TV regularization and TV flow in Proposition 16 gives that the TV regularization in general does not possess the semigroup property (36). This is in contrast to the situation for the TV flow, recall property 4 in Theorem 10.

Studying Figure 4, note that $\|\partial_\alpha u_{\alpha_1}\|_2 < \|\partial_\alpha u_{\alpha_2}\|_2$ for $\alpha_1 \in (2/5, 2)$ and $\alpha_2 \in (2, 4)$. Hence, $\|\partial_\alpha u_\alpha\|_2$ is in general not decreasing with respect to the parameter α . This is again in contrast to the situation for the TV flow, recall property 2 in Proposition 11.

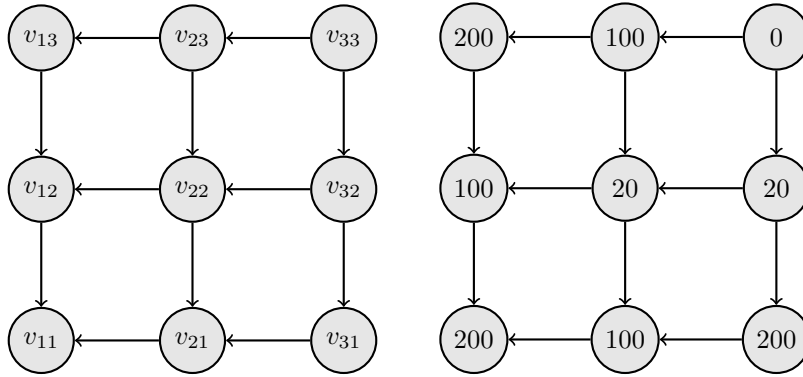


Figure 6: Graph structure (left) and pixel intensities (right) of the initial datum \tilde{f} .

For the TV regularization, note that

$$\text{sgn}(u_\alpha(v_{32}) - u_\alpha(v_{22})) = \begin{cases} 1, & 0 \leq \alpha < 2/5, \\ 0, & 2/5 \leq \alpha < 2, \\ -1, & 2 \leq \alpha \leq 4. \end{cases}$$

That is, the jump between $u_\alpha(v_{22})$ and $u_\alpha(v_{32})$ disappears for $2/5 \leq \alpha \leq 2$ but appears again, with reversed sign, for $2 < \alpha \leq 4$. For the TV flow, we have

$$\text{sgn}(u(t)(v_{32}) - u(t)(v_{22})) = \begin{cases} 1, & 0 \leq t < 2/5, \\ 0, & t = 2/5, \\ -1, & 2/5 < t \leq 4. \end{cases}$$

Here the jump between $u(t)(v_{22})$ and $u(t)(v_{32})$ disappears at $t = 2/5$ and then a jump with reversed sign appears for $2/5 < t \leq 4$. By the jump sets of u_α and $u(t)$, we denote the set of all edges $e = (v_{ij}, v_{kl}) \in E$ such that $\text{sgn}(u_\alpha(v_{ij}) - u_\alpha(v_{kl})) \in \{\pm 1\}$ respectively the set of all edges $e = (v_{ij}, v_{kl}) \in E$ such that $\text{sgn}(u(t)(v_{ij}) - u(t)(v_{kl})) \in \{\pm 1\}$. The provided example shows that the jump sets, in higher-dimensional graph settings, can grow with respect to α or t . Given one-dimensional graphs (as given by Figure 2), however, it can be shown that the jump sets are nonincreasing, see [23, Prop. 4.1 and 4.2].

Consider the slightly different initial datum \tilde{f} given by Figure 6. Note that $\tilde{f}(v_{ij}) = f(v_{ij})$ except for v_{22} where $\tilde{f}(v_{22}) = 20$. By doing analogous derivations as in the previous example for $0 \leq \alpha \leq 4$, results according to Figure 7 are obtained. A jump is created in the resulting images $u_\alpha = u(\alpha)$, $0 < \alpha \leq 4$, between the vertices v_{22} and v_{32} which is not present in the initial datum \tilde{f} . In higher-dimensional graph settings, the jump set of an image resulting from TV regularization or TV flow can therefore strictly contain the jump set of the initial datum. Note that this cannot happen for one-dimensional graphs as the jump sets then are nonincreasing. On the other hand, in the continuous anisotropic setting this phenomenon is known to occur, see [4, Rem. 4] and [16, Ex. 1].

6 Conclusion

In this article we have studied and compared TV regularization and TV flow for functions defined on the vertices of a finite oriented graph. Our motivation was the discrete one-dimensional setting, where the two problems are known to be equivalent and their solution minimizes a large class of convex functionals in a certain neighbourhood of the data [17, 23, 14].

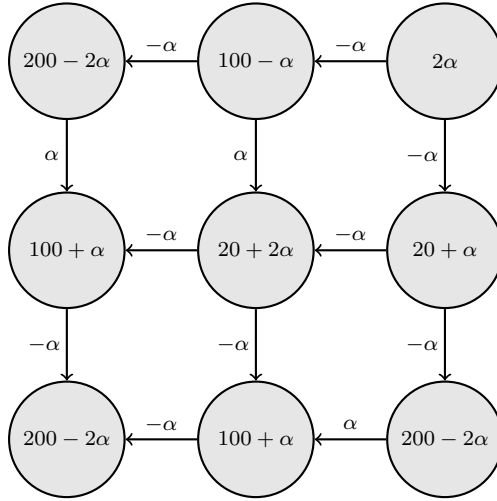


Figure 7: $u_\alpha = u(\alpha)$ (defined on the vertices) and $F_\alpha = F(\alpha)$ (defined on the edges) for $0 \leq \alpha \leq 4$.

It turns out that in the graph setting this situation can only be recovered for $\alpha, t \in [0, t_1] \cup [t_M, \infty)$, the reason being that on the complement (t_1, t_M) the ROF and flow solution are in general different. Here t_1 and t_M are the first and last times, respectively, the time derivative of the flow solution changes.

In addition we have shown that for every $\alpha \geq 0$ the ROF-minimizer u_α simultaneously minimizes all functionals of the form

$$u \mapsto \sum_{v \in V} \varphi(u(v))$$

over the set $f + \alpha \partial J(0)$, where $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is convex but otherwise arbitrary. In doing so we have relied on a recent characterization from [14, 15] of invariant φ -minimal sets, of which the subdifferential of J is one example. We have also constructed a simple example showing that the discrete isotropic total variation does not have this property.

Acknowledgements

We acknowledge support by the Austrian Science Fund (FWF) within the national research network ‘‘Geometry + Simulation,’’ S117.

References

- [1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.*, 10:1217–1229, 1994.
- [2] V. Barbu. *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer Monographs in Mathematics. Springer, New York, 2010.
- [3] M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers. Spectral decompositions using one-homogeneous functionals. *SIAM J. Imaging Sciences*, 9:1374–1408, 2016.
- [4] V. Caselles, A. Chambolle, and M. Novaga. The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Model. Simul.*, 6:879–894, 2007.

- [5] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20:89–97, 2004.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North-Holland, Amsterdam, 1976.
- [10] B. Grünbaum. *Convex Polytopes*. Pure and Applied Mathematics, XVI. Wiley: Interscience Publ., London, 1967.
- [11] W. Hinterberger, M. Hintermüller, K. Kunisch, M. von Oehsen, and O. Scherzer. Tube methods for BV regularization. *J. Math. Imaging Vision*, 19:219–235, 2003.
- [12] D. S. Hochbaum. An efficient algorithm for image segmentation, Markov random fields and related problems. *J. ACM*, 48(4):686–701, 2001.
- [13] K. Jalalzai. Some remarks on the staircasing phenomenon in total variation-based image denoising. *J. Math. Imaging Vision*, 54:256–268, 2016.
- [14] N. Kruglyak and E. Setterqvist. Discrete taut strings and real interpolation. *J. Funct. Anal.*, 270:671–704, 2016.
- [15] N. Kruglyak and E. Setterqvist. Invariant K -minimal sets in the discrete and continuous settings. *J. Fourier Anal. Appl.*, 23:672–711, 2017.
- [16] M. Lasica, S. Moll, and P. B. Mucha. Total variation denoising in ℓ^1 anisotropy. *SIAM J. Imaging Sciences*, 10:1691–1723, 2017.
- [17] E. Mammen and S. van de Geer. Locally adaptive regression splines. *Ann. Statist.*, 25:387–413, 1997.
- [18] F. Modigliani and F. E. Hohn. Production planning over time and the nature of the expectation and planning horizon. *Econometrica*, 23:46–66, 1955.
- [19] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [20] J. D. Salehi, Z.-L. Zhang, J. Kurose, and D. Towsley. Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing. *IEEE/ACM Trans. Networking*, 6:397–410, 1998.
- [21] F. Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bulletin of Mathematical Sciences*, 7(1):87–154, Apr 2017.
- [22] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Applied Mathematical Sciences, 167. Springer, New York, 2009.

- [23] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDs. *SIAM J. Numer. Anal.*, 42:686–713, 2004.
- [24] J. Yang and S. Ulukus. Optimal packet scheduling in an energy harvesting communication system. *IEEE Trans. Commun.*, 60:220–230, 2012.