Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. 94

On the Equivalence of Soft Wavelet Shrinkage, Total Variation Diffusion, Total Variation Regularization, and SIDEs

Gabriele Steidl, Joachim Weickert, Thomas Brox, Pavel Mrázek and Martin Welk

Saarbrücken 2003

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. 94
submitted: August 29, 2003

On the Equivalence of Soft Wavelet Shrinkage, Total Variation Diffusion, Total Variation Regularization, and SIDEs

Gabriele Steidl
Faculty of Mathematics and Computer Science
University of Mannheim
68131 Mannheim, Germany
steidl@kiwi.math.uni-mannheim.de

Joachim Weickert
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Building 27.1
66041 Saarbrücken, Germany
weickert@mia.uni-saarland.de

Thomas Brox
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Building 27.1
66041 Saarbrücken, Germany
brox@mia.uni-saarland.de

Pavel Mrázek
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Building 27.1
66041 Saarbrücken, Germany
mrazek@mia.uni-saarland.de

Martin Welk
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Building 27.1
66041 Saarbrücken, Germany
welk@mia.uni-saarland.de

This preprint is a revised version of Preprint 26 (February 2003) from Series DFG–SPP 1114, issued by DFG-Schwerpunktprogramm 1114 “Mathematical methods for time series analysis and image processing” at the University of Bremen.

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-Mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Abstract

Soft wavelet shrinkage, total variation (TV) diffusion, total variation regularization, and a dynamical system called SIDEs are four useful techniques for discontinuity preserving denoising of signals and images. In this paper we investigate under which circumstances these methods are equivalent in the 1-D case. First we prove that Haar wavelet shrinkage on a single scale is equivalent to a single step of space-discrete TV diffusion or regularization of two-pixel pairs. In the translationally invariant case we show that applying cycle spinning to Haar wavelet shrinkage on a single scale can be regarded as an absolutely stable explicit discretization of TV diffusion. We prove that space-discrete TV diffusion and TV regularization are identical, and that they are also equivalent to the SIDEs system when a specific force function is chosen. Afterwards we show that wavelet shrinkage on multiple scales can be regarded as a single step diffusion filtering or regularization of the Laplacian pyramid of the signal. We analyse possibilities to avoid Gibbs-like artifacts for multiscale Haar wavelet shrinkage by scaling the thesholds. Finally we present experiments where hybrid methods are designed that combine the advantages of wavelets and PDE / variational approaches. These methods are based on iterated shift-invariant wavelet shrinkage at multiple scales with scaled thresholds.

Contents

1 Introduction

Image denoising is a field where one is typically interested in removing noise without sacrificing important structures such as edges. This goal cannot be achieved with linear filters. Consequently, a large variety of nonlinear strategies has been proposed including rank-order filtering and mathematical morphology [28, 37], global stochastic methods [35, 40, 69], adaptive smoothing [7, 38, 54, 61], wavelet techniques [26, 27, 44], partial differential equations (PDEs) [2, 50, 55, 67] and variational methods [5, 6, 12, 49, 57, 63, 66].
Although these method classes serve the same purpose, relatively few publications exist where their similarities and differences are juxtaposed and their mutual relations are analysed. However, such an analysis is highly desirable, since it can help to transfer results from one of these classes to the others. Moreover, a deeper understanding of the differences between these classes might be helpful for designing novel hybrid methods that combine the advantages of the different classes.
The goal of the present paper is to address this problem by analysing relations between four important representatives of discontinuity preserving denoising methods:

- wavelet soft thresholding [26]
- space-discrete total variation (TV) diffusion [3, 4]

- discrete total variation (TV) regularization [53, 1]
- SIDEs, a dynamical system that has been inspired from space-discrete stabilized inverse diffusion equations [51]

Figure 1 gives an illustration of the denoising properties of soft wavelet shrinkage and TV regularization methods. We observe that the results do not differ very much. Indeed, we shall prove in our paper that all four before mentioned methods are very closely related.
In order to keep things as simple as possible we base our analysis on the 1-D case. Our basic strategy is to start with the simplest cases for which we can establish equivalence. Afterwards we extend these results to more general situations. The higher dimensional case is beyond the scope of the present paper, since it cannot be treated as a straightforward generalization of the 1-D ideas. For some preliminary results in 2-D, we refer to [47], where diffusion-inspired wavelet shrinkage with improved rotation invariance is introduced.

Our paper is organized as follows: in Section 2 we give a very brief description of the general ideas behind wavelet shrinkage, nonlinear diffusion filtering, variational image denoising and SIDEs. In Section 3 we specify these paradigms to the simplest cases where equivalence can be shown. In this section we restrict ourselves to 2-pixel signals, soft Haar wavelet shrinkage, the total variation diffusivity and its corresponding regularizer. Under these circumstances we prove equivalence between wavelet shrinkage, TV diffusion and TV regularization. These results are extended in Section 4 to the translationally invariant case with N-pixel signals. In the wavelet setting, we use a Haar wavelet based technique on a single scale with cycle spinning. We show that it can be regarded as a single iteration of a stabilized explicit scheme for TV diffusion, and we prove that this TV diffusion is equivalent to both TV regularization and SIDEs with an appropriate force function. In Section 5 we extend our wavelet results from a single scale to multiple scales. We show that multiple scale Haar wavelet soft shrinkage can be regarded as TV diffusion, TV regularization or SIDEs applied to a Laplacian pyramid decomposition of the signal. In Section 6 we present experiments where we compare iterated single scale filtering with noniterated and iterated multiscale filtering. The paper is concluded with a summary in Section 7.

Related work. Earlier applications of wavelets in the context of PDE-based denoising have been focusing on their use as basis functions in Galerkin schemes for nonlinear diffusion filtering [32, 34]. More recently a number of interesting connections between wavelet shrinkage of functions, regularization methods and PDEs has been established. A book by Meyer [46] presents a unified view on wavelets and nonlinear evolutions, and Shen and Strang [58, 59] have included wavelets into the solution of the linear heat equation. Chambolle et al. [16] showed that one may interpret wavelet shrinkage of functions as regularization processes in suitable Besov spaces. In particular Haar thresholding was considered in [21]. Furthermore, Cohen et al. [20] showed that the space of functions of bounded variation can be „almost" characterized by wavelet expansions. Chambolle and Lucier [18] considered iterated translationally invariant wavelet shrinkage and interpreted it as a nonlinear scale-space, that differs from other scale-spaces by the fact that it is not given in terms of PDEs. A stochastic interpretation of the connections between wavelets and edge-preserving regularization is given by Belge et al. [9].

There has also been a rapidly increasing interest in designing hybrid methods using both wavelet shrinkage and TV denoising methods. Durand and Froment [29] proposed to address the problem of pseudo-Gibbs artifacts in wavelet denoising by replacing the thresholded wavelet coefficients by coefficients that minimize the total variation. Their method is also close in spirit to approaches by Chan and Zhou [19] who postprocessed images obtained from wavelet shrinkage by a TV-like regularization technique. Coifman and Sowa [23] used functional minimization with wavelet constraints for postprocessing signals that have been degraded by wavelet thresholding or quantization. Candes and Guo [15] also presented related work, in which they combined ridgelets and curvelets with TV minimization strategies. Recently, Malgouyres [42, 43] proposed a hybrid method that uses both wavelet packets and TV approaches. His experiments showed that it may restore textured regions without introducing ringing artifacts.

Regarding the relations between wavelet shrinkage denoising of discrete signals and TV reduction, not much research has been done so far. One notable exception is a recent paper by Coifman and Sowa [24] where they propose TV diminishing flows that act along the direction of Haar wavelets. Bao and Krim [8] addressed the problem of texture loss in diffusion scale-spaces by incorporating ideas from wavelet analysis. An experimental evaluation of the denoising capabilities of 3-D wavelet shrinkage and nonlinear diffusion filters is presented in a paper by Frangakis et al. [33].

This discussion shows that our paper differs from preceding work by the fact that we investigate conditions under which we can prove equivalence between wavelet shrinkage of discrete signals, space-discrete TV diffusion or regularization, and SIDEs. Some preliminary results in this paper have been presented at conferences [62, 13].

2 The Basic Methods

The goal of this section is to give a brief introduction to the methods that are considered in this paper: soft Haar wavelet shrinkage, total variation diffusion, total variation denoising, and SIDEs.

2.1 Wavelet Shrinkage

During the last years wavelet methods have proved their use in various signal processing tasks. One of them is discontinuity-preserving denoising. The discrete wavelet transform represents a one-dimensional signal $f(x)$ in terms of shifted versions of a dilated lowpass scaling function $\varphi(x)$, and shifted and dilated versions of a bandpass wavelet function $\psi(x)$. In case of orthogonal wavelets, this gives

$$f(x) = \sum_{i \in \mathbb{Z}} \langle f, \varphi_i^{j_e} \rangle \, \varphi_i^{j_e}(x) + \sum_{j=-\infty}^{j_e} \sum_{i \in \mathbb{Z}} \langle f, \psi_i^j \rangle \, \psi_i^j(x),$$

where $\psi_i^j(x) := 2^{-j/2} \psi(2^{-j}x - i)$ and where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L_2(\mathbb{R})$. The wavelet representation employs scaling components only at one level j_e, and wavelet components at levels $j \leq j_e$ add higher resolution details to the signal.
If the measurements f are corrupted by white Gaussian noise, then this noise is contained to small amount in all wavelet coefficients $\langle f, \psi_i^j \rangle$, while the original signal is in general determined by few significant wavelet coefficients. Therefore wavelet shrinkage attempts to eliminate noise from the wavelet coefficients by the following three-step procedure:

- *Analysis*: transform the noisy data f to the wavelet coefficients $d_i^j = \langle f, \psi_i^j \rangle$ and scaling function coefficients $c_i^{j_e} = \langle f, \varphi_i^{j_e} \rangle$.
- *Shrinkage*: apply a shrinkage function S_τ with a threshold parameter τ related to the variance of the Gaussian noise to the wavelet coefficients, i.e., $S_\tau(d_i^j) = S_\tau(\langle f, \psi_i^j \rangle)$.
- *Synthesis*: reconstruct the denoised version u of f from the shrunken wavelet coefficients

$$u(x) := \sum_{i \in \mathbb{Z}} \langle f, \varphi_i^{j_e} \rangle \, \varphi_i^{j_e}(x) + \sum_{j=-\infty}^{j_e} \sum_{i \in \mathbb{Z}} S_\tau(\langle f, \psi_i^j \rangle) \, \psi_i^j(x).$$

In the literature a number of different shrinkage functions have been considered. In this paper we focus on one of the most popular strategies, namely Donoho's soft shrinkage [26]. It uses the soft thresholding with threshold parameter $\tau > 0$

$$S_\tau(x) = \begin{cases} x - \tau \operatorname{sgn}(x) & \text{if} \quad |x| > \tau, \\ 0 & \text{if} \quad |x| \leq \tau, \end{cases} \tag{1}$$

which shrinks all coefficients towards zero. Other shrinkage functions will be considered in a forthcoming paper.

Furthermore, in this paper we restrict our attention to Haar wavelets. They are well-suited for recovering piecewise constant signals with discontinuities. The Haar wavelet $\psi(x)$ and the corresponding scaling function $\varphi(x)$ are given by

$$\begin{aligned} \psi(x) &:= \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}, \\ \varphi(x) &:= \mathbf{1}_{[0,1)}, \end{aligned}$$

where $\mathbf{1}_{[a,b)}$ denotes the characteristic function of $[a, b)$:

$$\mathbf{1}_{[a,b)}(x) := \begin{cases} 1 & \text{if } x \in [a,b), \\ 0 & \text{else.} \end{cases}$$

Using the so-called "two-scale relation" of the wavelet and its scaling function, the coefficients c_i^j and d_i^j at higher level j can be computed from the coefficients c_i^{j-1} at lower level $j-1$ and conversely. This results in fast algorithms for the analysis step and synthesis step. For the Haar wavelets, we obtain

$$c_i^j = \frac{c_{2i}^{j-1} + c_{2i+1}^{j-1}}{\sqrt{2}}, \qquad d_i^j = \frac{c_{2i}^{j-1} - c_{2i+1}^{j-1}}{\sqrt{2}}, \tag{2}$$

and

$$c_{2i}^{j-1} = \frac{c_i^j + d_i^j}{\sqrt{2}}, \qquad c_{2i+1}^{j-1} = \frac{c_i^j - d_i^j}{\sqrt{2}}. \tag{3}$$

2.2 Diffusion Filtering

Let us now consider a function $f(x)$ on some interval $[a, b]$. The basic idea behind nonlinear diffusion filtering is to obtain a family $u(x, t)$ of filtered versions of the signal $f(x)$ as the solution of a suitable diffusion process with $f(x)$ as initial condition and homogeneous Neumann boundary conditions [50]:

$$\begin{aligned} u_t &= (g(u_x^2)\, u_x)_x && \text{on } (a,b) \times (0,\infty), \qquad (4) \\ u(x,0) &= f(x) && \text{for all } x \in [a,b], \\ u_x(a,t) &= u_x(b,t) = 0 && \text{for all } t \in (0,\infty), \end{aligned}$$

where subscripts denote partial derivatives, and the diffusion time t is a simplification parameter: larger values correspond to stronger filtering.
The diffusivity $g(u_x^2)$ is a nonnegative function that steers the amount of diffusion. Usually, it is decreasing in u_x^2. This ensures that strong edges are less blurred by the diffusion filter than noise and low-contrast details. In the present paper, we focus on the *total variation (TV) diffusivity*

$$g(u_x^2) \;:=\; \frac{1}{|u_x|}. \tag{5}$$

The resulting *TV diffusion filter* (also called *total variation flow*) has a number of interesting properties. It requires no additional parameters (besides t), it is well-posed [3, 10, 30], it preserves the shape of some objects [10], and it leads to constant signals in finite time [4].

2.3 Regularization Methods

Regularization methods constitute an alternative to diffusion filters when one is interested in a discontinuity-preserving denoising method for a continuous signal $f(x)$ with $x \in [a, b]$. Here the basic idea is to look for the minimizer u of the energy functional

$$E(u; \alpha, f) \;:=\; \int_a^b \Big((u-f)^2 \,+\, \alpha\,\Psi(u_x^2)\Big)\,dx. \tag{6}$$

The first term of this functional encourages similarity between the original signal $f(x)$ and its filtered version $u(x)$, while the second term penalizes deviations from smoothness. The increasing function Ψ is called *penalizer (regularizer)*, and the nonnegative *regularization parameter* α serves as smoothness weight: larger values correspond to a more pronounced filtering.
As is explained in detail in [56], there are strong relations between regularization methods and diffusion filters: A minimizer of (6) satisfies necessarily the Euler–Lagrange equation

$$\frac{u-f}{\alpha} \;=\; (\Psi'(u_x^2)\,u_x)_x,$$

with homogeneous Neumann boundary conditions. This equation may be regarded as a fully implicit time discretization of the diffusion equation (4) with diffusivity $g(u_x^2) = \Psi'(u_x^2)$, initial value $f(x)$, and stopping time α. Thus, one would expect that the minimizer of (6) *approximates* the diffusion filter (4), but is not identical to it.

In the present paper, we are interested in one of the most popular nonlinear regularization methods, namely *total variation (TV) regularization* [53, 1]. It uses the penalizer

$$\Psi(u_x^2) \;:=\; 2\,|u_x|$$

which corresponds to the TV diffusivity (5). This regularization strategy is well-known for its good denoising capabilities and its tendency to create blocky, segmentation-like results. Well-posedness results have been established in [17].

2.4 SIDEs

A SIDE is a dynamical system that has been inspired from a stabilized limiting case of a space-discrete nonlinear diffusion filter [51]. The name SIDE is an acronym for *stabilized inverse diffusion equation*.
Let us consider a discrete signal $f = (f_i)_{i=0}^{N-1}$. Then its SIDE evolution produces a sequence of filtered images $u(t) = (u_i(t))_{i=0}^{N-1}$ with $u(0) = f$. Increasing the time t leads to a consecutive merging of regions. The evolution between two merging events is governed by a dynamical system with discontinuous right hand side.
Assume that at some time t_j, a pixel with index i belongs to a constant region of size m_{i,t_j}, i.e. there exist $l \geq 1$ and $r \geq 0$ with $m_{i,t_j} = l + r$,

$$u_{i-l+1} = \ldots = u_i = u_{i+1} = \ldots = u_{i+r},$$

and

$$u_{i-l} \neq u_{i-l+1} \;\text{ if } i-l \geq 0, \quad u_{i+r} \neq u_{i+r+1} \;\text{ if } i+r \leq N-2.$$

Then the SIDEs algorithm proceeds as follows:

(i) *(Initialization)*
Start at time $t_0 = 0$ with the trivial segmentation, where each pixel i is regarded as a region of size $m_{i,0} = 1$:

$$u_i(0) = f_i.$$

(ii) *(Evolution)*
Given some segmentation at time t_j, the signal evolves according to

$$\dot{u}_i \;=\; \begin{cases} \frac{1}{m_{i,t_j}} F(u_{i+r+1} - u_{i+r}) & \text{if } i-l = -1, \\ \frac{-1}{m_{i,t_j}} F(u_{i-l+1} - u_{i-l}) & \text{if } i+r = N-1, \\ \frac{1}{m_{i,t_j}} \left(F(u_{i+r+1} - u_{i+r}) - F(u_{i-l+1} - u_{i-l})\right) & \text{else,} \end{cases} \tag{7}$$

where $\dot{u}_i$ denotes the derivative of u_i with respect to t and F is a so-called *force function* that satisfies a number of formal requirements [51]. The first case in (7) describes the evolution of the region at the left signal boundary, the second case applies for the right boundary region, and the third case specifies the evolution of all inner regions.

In [51], only the third case has been specified. We have supplemented the other two cases here in order to be able to treat the boundary regions in a proper way as well.

The evolution is stopped when two neighbouring regions attain equal grey values. This determines the new merging time t_{j+1}.

(iii) *(Merging)*
Merge the neighbouring regions with equal grey values.

(iv) *(Loop Control)*
Stop if all regions are merged to one, else go back to step (ii).

We see that the stabilization in SIDEs is achieved by an additional definition that results in merging neighbouring regions when they approach each other. This step is crucial for the performance of SIDEs, as it can be used for reducing the state variables of the dynamical system. The analytical solutions in the following sections will provide further theoretical justification for this region merging step.

In [51] several theoretical results for sides are proved, including a maximum principle, well-posedness properties and a finite extinction time.

The dynamic system suggests that for the specific case $m_i = 1$, one may regard a 1-D SIDE as a space discretization of the partial differential equation

$$u_t = (F(u_x))_x$$

with homogeneous Neumann boundary conditions. This is a diffusion equation with flux function F. Since we are specifically interested in the total variation case, we do not consider the specific choice in [51], but restrict ourselves to the *total variation (TV) force function*

$$F(v) \;\; := \;\; \begin{cases} \;\;\,1 & \text{if} \;\; v > 0, \\ -1 & \text{if} \;\; v < 0. \end{cases}$$

Then it is evident that, if $m_i = 1$ for all i, TV diffusion is approximated.

3 Two-Pixel Signals

In this section, we analyse relations between soft wavelet shrinkage, TV diffusion, TV regularization and SIDEs for the simplest signals, namely discrete signals with only two pixels. We will see that the restriction to two pixels allows us to find analytical solutions for these degenerated nonlinear processes.

3.1 Soft Haar Wavelet Shrinkage of Two-Pixel Signals

Let us now consider a discrete two-pixel signal $f = (f_0, f_1)$, and study its change under soft Haar wavelet shrinkage.
The analysis step produces the coefficients

$$c = \frac{f_0 + f_1}{\sqrt{2}}, \qquad d = \frac{f_0 - f_1}{\sqrt{2}}$$

of the scaling function and the wavelet. For simplicity, we have dropped the sub- and superscripts for c and d.
This step is followed by the shrinkage operation $S_\tau(d)$ with the soft shrinkage function (1) Then the synthesis step

$$u_0 \;=\; \frac{c + S_\tau(d)}{\sqrt{2}}, \qquad u_1 \;=\; \frac{c - S_\tau(d)}{\sqrt{2}}$$

gives the final result

$$u_0 \;=\; \begin{cases} f_0 + \frac{\tau}{\sqrt{2}}\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad \tau < |f_1 - f_0|/\sqrt{2}, \\ (f_0 + f_1)/2 & \text{else,} \end{cases} \tag{8}$$

$$u_1 \;=\; \begin{cases} f_1 - \frac{\tau}{\sqrt{2}}\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad \tau < |f_1 - f_0|/\sqrt{2}, \\ (f_0 + f_1)/2 & \text{else.} \end{cases} \tag{9}$$

This shows that by increasing the shrinkage threshold τ, the grey values of both pixels approach each other. For $\tau = |f_1 - f_0|/\sqrt{2}$, they merge, and for larger τ, they remain merged.

3.2 TV Diffusion of Two-Pixel Signals

Next we are interested in the space-discrete diffusion of two-pixel signals (f_0, f_1). The homogeneous Neumann boundary conditions are discretized by setting flows over the signal boundary to zero. In this case a space-discrete version of the TV diffusion equation

$$u_t = \left(\frac{u_x}{|u_x|}\right)_x$$

can be written as

$$\dot{u}_0 = \frac{u_1 - u_0}{|u_1 - u_0|}, \tag{10}$$

$$\dot{u}_1 = -\frac{u_1 - u_0}{|u_1 - u_0|}, \tag{11}$$

with initial conditions $u_0(0) = f_0$ and $u_1(0) = f_1$. Here, the dot denotes again temporal differentiation, and the pixel size is set to 1.
Setting $w(t) := u_1(t) - u_0(t)$ and $\eta := f_1 - f_0$, and subtracting (10) from (11), we obtain the following initial value problem:

$$\begin{aligned} \dot{w} &= -2\,\frac{w}{|w|}, \\ w(0) &= \eta. \end{aligned} \tag{12}$$

The right-hand side of this differential equation is discontinuous for $w = 0$ and thus requires a generalization of the concept of solution. We say that w is a solution of (12) if it is an absolutely continuous function which fulfils

$$\begin{aligned} \dot{w} &= -2\,\mathrm{sgn}\,(w), \\ w(0) &= \eta \end{aligned} \tag{13}$$

almost everywhere, where

(I) $\mathrm{sgn}\,(w) := 1$ if $w > 0$,
$\mathrm{sgn}\,(w) := -1$ if $w < 0$
and may take any value in $[-1, 1]$ if $w = 0$.

Note that this definition is in agreement with the frequently used concept of differential inclusions for differential equations with discontinuous right-hand sides [31], where absolutely continuous solutions of

$$-\frac{1}{2}\,\dot{w} \in \begin{cases} \{1\} & \text{if } w > 0 \\ \{-1\} & \text{if } w < 0 \\ [-1, 1] & \text{if } w = 0 \end{cases}$$

were considered. The solution of (13) can be obtained as follows: If $\eta \neq 0$, then we have by straightforward computation for $t < |\eta|/2$ that $w(t) = \eta - 2\,t\,\mathrm{sgn}\,(\eta)$, and in particular by continuity of w that $w(|\eta|/2) = 0$. Assume that $w(t) \neq 0$ for some $t > |\eta|/2$. Let without loss of generality $w(t) > 0$. The opposite assumption $w(t) < 0$ can be handled in the same way. Then $w(t) = -2t + C$, where we get by continuity of w, if t approaches $|\eta|/2$, that

$C = |\eta|$, and consequently $w(t) = 2(|\eta|/2 - t) < 0$. This contradicts our assumption. Thus $w(t) = 0$ for $t \geq |\eta|/2$. In summary, we obtain the solution

$$w(t) = \begin{cases} \eta - 2\,t\,\mathrm{sgn}\,(\eta) & \text{if } t < |\eta|/2, \\ 0 & \text{if } t \geq |\eta|/2. \end{cases} \tag{14}$$

This equation shows that the grey value difference $w(t) = u_1(t) - u_0(t)$ tends linearly to 0. Both pixels merge at time $t = |f_1 - f_0|/2$, and remain merged afterwards. Thus, already the simple two-pixel model indicates a finite extinction time for TV diffusion.
Since $\dot{u}_0 + \dot{u}_1 = 0$ and $u_0(0) + u_1(0) = f_0 + f_1$, we see further that the average grey value is preserved:

$$u_0(t) + u_1(t) = f_0 + f_1 \qquad \forall\, t \geq 0. \tag{15}$$

Using (14) and (15), we obtain the analytical solution

$$u_0(t) \;=\; \begin{cases} f_0 + t\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad t < |f_1 - f_0|/2, \\ (f_0 + f_1)/2 & \text{else,} \end{cases} \tag{16}$$

$$u_1(t) \;=\; \begin{cases} f_1 - t\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad t < |f_1 - f_0|/2, \\ (f_0 + f_1)/2 & \text{else.} \end{cases} \tag{17}$$

Interestingly, this result is identical to the results (8)–(9) for soft Haar wavelet shrinkage if one identifies the diffusion time t with the threshold parameter $\tau = \sqrt{2}t$.

3.3 TV Regularization of Two-Pixel Signals

Let us now turn our attention to the regularization framework.
Again we are only interested in the two-pixel model (f_0, f_1). We consider a space-discrete variant of (6) with a TV penalizer:

$$E(u_0, u_1; \alpha, f) \;=\; (f_0 - u_0)^2 + (f_1 - u_1)^2 + 2\alpha\,|u_1 - u_0|, \tag{18}$$

Straightforward computation results in the following minimizer of (18):

$$u_0 \;=\; \begin{cases} f_0 + \alpha\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad \alpha < |f_1 - f_0|/2, \\ (f_0 + f_1)/2 & \text{else.} \end{cases}$$

$$u_1 \;=\; \begin{cases} f_1 - \alpha\,\mathrm{sgn}\,(f_1 - f_0) & \text{if} \quad \alpha < |f_1 - f_0|/2, \\ (f_0 + f_1)/2 & \text{else.} \end{cases}$$

This result coincides with the outcome of a single Haar wavelet shrinkage step with shrinkage parameter $\tau = \sqrt{2}\,\alpha$. Moreover, it is identical to TV diffusion, if one replaces the diffusion time t by the regularization parameter α. Thus, all three methods are equivalent by setting

$$\tau = \sqrt{2}\,t = \sqrt{2}\,\alpha.$$

It is remarkable that TV diffusion and TV regularization give identical evolutions in the two-pixel case. From the considerations in Section 2.3 one would only expect that the processes approximate each other. In Section 4.3 we will investigate if this equivalence also holds in the general space-discrete case with N pixels.

3.4 SIDEs for Two-Pixel Signals

If we consider the SIDE evolution of a two-pixel signal (f_0, f_1), we obtain for the case of a TV force function the dynamical system

$$\begin{aligned} \dot{u}_0 &= \frac{u_1 - u_0}{|u_1 - u_0|}, \\ \dot{u}_1 &= -\frac{u_1 - u_0}{|u_1 - u_0|} \end{aligned}$$

with initial conditions $u_0(0) = f_0$ and $u_1(0) = f_1$.
This is the same evolution as in the TV diffusion case. Hence, its solution is given by (16)–(17), and there is a finite merging time $t = |f_1 - f_0|/2$.

4 N-Pixel Signals

So far we have focused on the two-pixel case. Let us now investigate which of the equivalences carry over to the general 1-D case with N pixels. To this end we will consider shift invariant wavelet shrinkage on a single scale, show that it performs a numerical approximation to TV diffusion, prove the equivalence of space-discrete TV diffusion and discrete TV regularization by deriving analytical solutions of both processes, and show that this solution coincides with SIDEs with TV force functions.

4.1 Shift Invariant Wavelet Shrinkage on a Single Scale

Let us first reconsider the soft Haar wavelet shrinkage on a single scale with N pixels, where N is even. Figure 2 illustrates this computation as two-channel filter bank. As usual we apply the z-transform notation $f(z) = \sum_{i=0}^{N-1} f_i z^{-i}$.

Then $\boxed{H_i(z)}$ $(i = 0, 1)$ denotes the convolution of f with the lowpass filter $(i = 0)$ and the highpass filter $(i = 1)$, i.e. $f(z)H_i(z)$, $\boxed{2\downarrow}$ and $\boxed{2\uparrow}$ downsampling and upsampling by 2, respectively, and the circle soft thresholding by S_τ. Finally, $\bullet$ signifies addition; see also [64].

The use of Haar wavelets creates a natural decomposition of the signal into two-pixel pairs of type (f_{2j}, f_{2j+1}) $(j = 0, \ldots, N/2 - 1)$. This two-pixel clustering, however, also causes a lack of translation invariance which may be responsible for visual artifacts.
One method to improve the quality of the denoised signal considerably is to "average out" the translation dependence. This method was termed *cycle spinning* by Coifman et al. [22]. For a single wavelet decomposition step, the basic idea of cycle spinning on a single scale reads as follows:

(a) perform wavelet shrinkage (8), (9) on successive pairs of the original signal,

(b) shift the signal one pixel to the right and perform wavelet shrinkage on successive pairs of the shifted signal, shift the resulting signal one pixel back to the left,

(c) average both results.

The shifting process requires the incorporation of boundary conditions for f. Again we mirror the signal f at its ends. The steps (a)–(c) are equivalent to denoising the signal using the nonsubsampled filter bank in Figure 3. More sophisticated material on oversampled filter banks, corresponding wavelet frames and undecimated wavelet transforms can be found in [25, 44].

4.2 Equivalence to a Numerical Scheme for TV Diffusion

We have seen that, in order to improve the performance of wavelet shrinkage and to make wavelet-based denoising translationally invariant, cycle spinning can be used. Since there is an equivalence between Haar wavelet shrinkage and TV diffusion in the two-pixel case, it would be natural to ask if there is a TV diffusion scheme equivalent to translationally invariant soft Haar wavelet shrinkage on a single level. This leads us to an interesting novel scheme for TV diffusion.

Derivation of the Scheme. We have been able to derive an analytical solution for TV diffusion in the two-pixel case. We can use this two-pixel solution to create a numerical scheme for N pixels. In order to derive such a scheme for some time step size $\triangle t$, we proceed in three steps that are inspired by the cycle spinning technique:

(a) perform TV diffusion with step size $2\triangle t$ on all pixel pairs (u_{2j}, u_{2j+1}),

(b) perform TV diffusion with step size $2\triangle t$ on all pixel pairs (u_{2j-1}, u_{2j}),

(c) average both results.

Obviously one step of this iterative scheme is equivalent to a translationally invariant Haar wavelet shrinkage with threshold $\tau = 2\sqrt{2}\triangle t$ on a single level. So let us investigate this scheme in more detail.

At iteration level k, we assume that our signal is given by $(u_i^k)_{i=0}^{N-1}$. We denote the resulting signal of step (a) by $(v_i^{k+1})_{i=0}^{N-1}$ and the spatial grid size by h. From our analysis of the two-pixel situation, it follows that v_i in some even pixel $i = 2j$ is given by

$$v_i^{k+1} \;=\; \frac{u_i^k + u_{i+1}^k}{2} \;-\; \begin{cases} \max\left(\frac{u_{i+1}^k - u_i^k}{2} - \frac{2\triangle t}{h}, 0\right) & \text{if } u_{i+1}^k \geq u_i^k, \\ \min\left(\frac{u_{i+1}^k - u_i^k}{2} + \frac{2\triangle t}{h}, 0\right) & \text{if } u_{i+1}^k < u_i^k. \end{cases} \tag{19}$$

To simplify the notation, we assume only in this subsection instead of the third agreement in (I) that $\operatorname{sgn}(0) := 0$. It is not difficult to see that (19) can be rewritten as

$$v_i^{k+1} \;=\; u_i^k + \frac{2\triangle t}{h} \operatorname{sgn}\left(u_{i+1}^k - u_i^k\right) \min\left(1,\; \frac{h}{4\triangle t}|u_{i+1}^k - u_i^k|\right). \tag{20}$$

Step (b) leads to a resulting signal $(w_i^{k+1})_{i=0}^{N-1}$. For $i = 2j$ it is given by

$$w_i^{k+1} \;=\; u_i^k - \frac{2\triangle t}{h} \operatorname{sgn}\left(u_i^k - u_{i-1}^k\right) \min\left(1,\; \frac{h}{4\triangle t}|u_i^k - u_{i-1}^k|\right). \tag{21}$$

Thus, the averaging step (c) gives the final scheme for TV diffusion:

$$\begin{aligned} u_i^{k+1} \;=\; u_i^k \;&+\; \frac{\triangle t}{h} \operatorname{sgn}\left(u_{i+1}^k - u_i^k\right) \min\left(1,\; \frac{h}{4\triangle t}|u_{i+1}^k - u_i^k|\right) \\ &-\; \frac{\triangle t}{h} \operatorname{sgn}\left(u_i^k - u_{i-1}^k\right) \min\left(1,\; \frac{h}{4\triangle t}|u_i^k - u_{i-1}^k|\right). \end{aligned} \tag{22}$$

The same scheme can also be derived if i is odd, since the construction (a)–(c) ensures that the result is translationally invariant. Hence it holds for every inner pixel $i \in \{1, \ldots, N-2\}$. It is even valid for the boundary pixels $i = 0$

and $i = N-1$, if we realize the homogeneous Neumann boundary conditions by introducing dummy values $u_{-1}^k := u_0^k$ and $u_N^k := u_{N-1}^k$.

Stability. Let us now investigate the stability properties of the explicit finite difference scheme (20). Since (20) satisfies

$$\min(u_i^k, u_{i+1}^k) \;\le\; v_i^{k+1} \;\le\; \max(u_i^k, u_{i+1}^k)$$

and (21) fulfils the estimate

$$\min(u_{i-1}^k, u_i^k) \;\le\; w_i^{k+1} \;\le\; \max(u_{i-1}^k, u_i^k)$$

we can conclude that

$$\min(u_{i-1}^k, u_i^k, u_{i+1}^k) \;\le\; u_i^{k+1} \;\le\; \max(u_{i-1}^k, u_i^k, u_{i+1}^k).$$

With the initial condition $u_j^0 = f_j$ for $j = 0, \ldots, N-1$, it follows that the two-pixel scheme (20) satisfies the discrete maximum-minimum principle

$$\min_j f_j \;\le\; u_i^{k+1} \;\le\; \max_j f_j$$

for all pixels $i \in \{0, \ldots, N-1\}$, all iteration levels $k = 0,\, 1,\, 2, \ldots$, and all time step sizes $\triangle t > 0$. In particular, this shows that the scheme is absolutely stable in the maximum norm.
We may regard (22) as a stabilization of the naive explicit scheme

$$u_i^{k+1} \;=\; u_i^k + \frac{\triangle t}{h}\,\mathrm{sgn}\,(u_{i+1}^k - u_i^k) - \frac{\triangle t}{h}\,\mathrm{sgn}\,(u_i^k - u_{i-1}^k) \tag{23}$$

which becomes unstable for arbitrary small time steps if neighbouring values become arbitrarily close.

Consistency. The absolute stability in scheme (22) is at the expense that its consistency is no longer unconditional. This effect is typical for absolutely stable explicit schemes; see for example the DuFort–Frankel scheme for linear diffusion [60]. In our case, equation (22) is an $O(\triangle t + h^2)$ approximation to the continuous TV diffusion for

$$\triangle t \;\le\; \frac{h}{4}\,\min\left(|u_{i+1}^k - u_i^k|,\, |u_i^k - u_{i-1}^k|\right),$$

since it coincides with scheme (23) then. For larger time step sizes, the scheme performs averaging within the neighbourhood of each pixel. By using small time step sizes, these averaging effects only appear in regions that are already almost flat such that the difference to real TV diffusion becomes invisible.

This two-pixel scheme may be regarded as an alternative to classical finite difference schemes that are based on the regularized TV flow

$$u_t = \left(\frac{u_x}{\sqrt{\epsilon^2 + u_x^2}} \right)_x . \tag{24}$$

The ϵ-regularization is necessary for making the diffusivity bounded. It has a similar effect as the deviation from consistency in the two-pixel scheme (22): For small $|u_x|$, a PDE is approximated that differs from TV diffusion and has better stability properties. Indeed, in Section 6 we shall see that both schemes give very similar results.

Related Schemes. The idea to split up a diffusion process into pairwise interactions has also proved to be fruitful in other fields. In the context of fluid dynamic problems, related schemes have been formulated by Richardson, Ferrell and Long [52]. These authors, however, use multiplicative splittings, i.e. they first compute the diffusion of the pairs of type (u_{2j}, u_{2j+1}), which is then used as initial state for the subsequent diffusion of the shifted pairs. In a general nonlinear setting, such a scheme is not translationally invariant. Our approach computes the diffusion of the pairs and the shifted pairs in parallel and averages afterwards. This additive splitting guarantees translation invariance. The splitting into two-pixel interactions distinguishes scheme (22) from other additive operator splittings [41, 68]. They use directional splittings along the coordinate axis.

4.3 Equivalence of Space-Discrete TV Diffusion and Discrete TV Regularization

The equivalence of TV diffusion and TV regularization in the two-pixel case gives rise to the question whether this equivalence also holds in the N-pixel situation. In order to clarify this, we now derive analytical solutions for both processes.

4.3.1 Space-Discrete TV-Diffusion

We consider the following dynamical system designed to describe space-discrete TV flow on a one-dimensional signal with N pixels:

$$\left.\begin{aligned} \dot{u}_0 &= \operatorname{sgn}(u_1 - u_0), \\ \dot{u}_i &= \operatorname{sgn}(u_{i+1} - u_i) - \operatorname{sgn}(u_i - u_{i-1}) \qquad (i = 1, \ldots, N-2), \\ \dot{u}_{N-1} &= -\operatorname{sgn}(u_{N-1} - u_{N-2}), \\ u(0) &= f. \end{aligned}\right\} \tag{25}$$

In the following, we further set $u_{-1} := u_0$ and $u_N := u_{N-1}$. Since the right-hand side of this system is discontinuous, we need again a more detailed specification of when a system of functions is said to satisfy these differential equations. A vector-valued function u is said to fulfil the system (25) over the time interval $[0, T]$ if the following holds true:

(II) u is an absolutely continuous vector-valued function which satisfies (25) almost everywhere, where sgn is defined by (I) on Page 11.

(III) If $\dot{u}_i(t)$ and $\dot{u}_{i+1}(t)$ exist for the same t, and $u_{i+1}(t) = u_i(t)$ holds, then the expression $\operatorname{sgn}(u_{i+1}(t) - u_i(t))$ occurring in both the right-hand sides for $\dot{u}_i(t)$ and $\dot{u}_{i+1}(t)$ must take the same value in both equations.

Proposition 4.1 (Properties of Space-Discrete TV Diffusion)
The system (25) has a unique solution $u(t)$ in the sense of (II) and (III). This solution has the following properties:

(i) *(Finite Extinction Time)*
There exists a finite time $T \geq 0$ such that for all $t \geq T$ the signal becomes constant:

$$u_i(t) = \frac{1}{N} \sum_{k=0}^{N-1} f_k \qquad \text{for all } i = 0, \ldots, N-1.$$

(ii) *(Finite Number of Merging Events)*
There exists a finite sequence $0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = T$ such that the interval $[0, T)$ splits into sub-intervals $[t_j, t_{j+1})$ with the property that for all $i = 0, \ldots, N-2$ either $u_i(t) = u_{i+1}(t)$ or $u_i(t) \neq u_{i+1}(t)$ throughout $[t_j, t_{j+1})$. The absolute difference between neighbouring pixels does not become larger for increasing $t \in [t_j, t_{j+1})$.

(iii) *(Analytical Solution)*
In each of the sub-intervals $[t_j, t_{j+1})$ constant regions of $u(t)$ evolve linearly:
For a fixed index i let us consider a constant region given by

$$u_{i-l+1} = \ldots = u_i = u_{i+1} = \ldots = u_{i+r} \qquad (l \geq 1, r \geq 0) \tag{26}$$

and

$$u_{i-l} \neq u_{i-l+1} \text{ if } i - l \geq 0, \quad u_{i+r} \neq u_{i+r+1} \text{ if } i + r \leq N - 1$$

for all $t \in [t_j, t_{j+1})$. We call (26) a region of size $m_{i,t_j} = l + r$. For $t \in [t_j, t_{j+1})$ let $\triangle t = t - t_j$. Then $u_i(t)$ is given by

$$u_i(t) = u_i(t_j) + \mu_{i,t_j} \frac{2\triangle t}{m_{i,t_j}} \,,$$

where μ_{i,t_j} reflects the relation between the region containing u_i and its neighbouring regions. It is given as follows:
For inner regions (i.e. $i - l \geq 0$ and $i + r \leq N - 1$) we have

$$\mu_{i,t_j} = \begin{cases} 0 & \text{if} \quad (u_{i-l}, u_i, u_{i+r+1}) \text{ is strictly monotonic,} \\ 1 & \text{if} \quad u_i \text{ is minimal in } (u_{i-l}, u_i, u_{i+r+1}), \\ -1 & \text{if} \quad u_i \text{ is maximal in } (u_{i-l}, u_i, u_{i+r+1}) \end{cases} \tag{27}$$

and in the boundary case ($i - l + 1 = 0$ or $i + r = N - 1$), the evolution is half as fast:

$$\mu_{i,t_j} = \begin{cases} 0 & \text{if} \quad m = N, \\ \frac{1}{2} & \text{if} \quad u_i \text{ is minimal in } (u_{i-l}, u_i, u_{i+r+1}), \\ -\frac{1}{2} & \text{if} \quad u_i \text{ is maximal in } (u_{i-l}, u_i, u_{i+r+1}). \end{cases} \tag{28}$$

Proof: Let u be a solution of (25). We show that u is uniquely determined and satisfies the rules (i)–(iii).
Our proof proceeds in four steps.

1. If $\dot{u}(t)$ exists at a fixed time t and $u_i(t)$ lies at this time in some region

$$u_{i-l+1}(t) = \ldots = u_i(t) = \ldots = u_{i+r}(t) \qquad (l \geq 1, r \geq 0),$$

$$u_{i-l}(t) \neq u_{i-l+1}(t) \;\text{ if } i - l \geq 0, \quad u_{i+r}(t) \neq u_{i+r+1}(t) \;\text{ if } i + r \leq N - 1$$

of size $m_{i,t}$, then it follows by (25) and (III) in the non-boundary case $i - l \geq 0$ and $i + r \leq N - 1$ that

$$u_i(t) \;=\; \frac{1}{m_{i,t}} \sum_{k=-l+1}^{r} u_{i+k}(t),$$

and therefore

$$\begin{aligned} \dot{u}_i(t) &= \frac{1}{m_{i,t}} \sum_{k=-l+1}^{r} \dot{u}_{i+k}(t) \\ &= \frac{1}{m_{i,t}} \left(\operatorname{sgn}\left(u_{i+r+1}(t) - u_i(t)\right) \,-\, \operatorname{sgn}\left(u_i(t) - u_{i-l}(t)\right) \right) \\ &= \mu_{i,t} \frac{2}{m_{i,t}} \,, \end{aligned} \tag{29}$$

where $\mu_{i,t}$ describes the relation between the region containing u_i and its neighbours at time t as in (27). In the boundary case $i - l + 1 = 0$ or $i + r = N - 1$ we follow the same lines and obtain (29) with $\mu_{i,t}$ defined by (28).

2. Let $\dot{u}(t)$ exist in some small interval (τ_0, τ_1) and assume that $u_i(t) \neq u_{i+1}(t)$ for some $i \in \{0, \ldots, N-2\}$ and all $t \in (\tau_0, \tau_1)$. By continuity of u we may assume that $u_i(t) < u_{i+1}(t)$ throughout (τ_0, τ_1). The opposite case $u_i(t) > u_{i+1}(t)$ can be handled in the same way. Then we obtain by (29) and definition of $\mu_{i,t}$ for all $t \in (\tau_0, \tau_1)$ that

$$\begin{aligned} \dot{u}_i(t) &\geq 0 \quad \text{if} \quad i - l \geq 0, && (30)\\ \dot{u}_i(t) &> 0 \quad \text{if} \quad i - l + 1 = 0, && (31)\\ \dot{u}_{i+1}(t) &\leq 0 \quad \text{if} \quad i + r \leq N - 2, && (32)\\ \dot{u}_{i+1}(t) &< 0 \quad \text{if} \quad i + r = N - 1. && (33) \end{aligned}$$

Set $w(t) := u_{i+1}(t) - u_i(t)$. Then the mean value theorem yields

$$w(\tau_1) - w(\tau_0) = (\tau_1 - \tau_0)\, \dot{w}(t^*)$$

for some $t^* \in (\tau_0, \tau_1)$ and we get by (30)–(33) that

$$w(\tau_1) - w(\tau_0) \leq 0$$

with strict inequality in the boundary case. Consequently, the difference between pixels cannot become larger in the considered interval. In particular, by continuity of u, pixels cannot be split. Once merged they stay merged.

3. Now we start at time $t_0 = 0$. Let t_1 be the largest time such that $\dot{u}(t)$ exists and no merging of regions appears in $(0, t_1)$. Then, for all $i \in \{0, \ldots, N-1\}$, a function u_i is in the same region with the same relations to its neighbouring regions throughout $[0, t_1)$. Thus, we conclude by (29) that

$$\dot{u}_i(t) = \mu_{i,0} \frac{2}{m_{i,0}} \qquad (t \in (0, t_1))$$

and consequently

$$\begin{aligned} u_i(t) &= \mu_{i,0} \frac{2t}{m_{i,0}} + C_{i,0}\\ &= f_i + \mu_{i,0} \frac{2t}{m_{i,0}} \qquad (t \in [0, t_1]), \end{aligned}$$

where the last equality follows by continuity of u_i if t approaches 0.

4. We are now in the position to analyse the entire chain of merging events successively.
   Next we consider the largest interval (t_1, t_2) without merging events in the same way, where we take the initial setting $u(t_1)$ instead of f into account. Then we obtain

$$u_i(t) = \mu_{i,t_1} \frac{2t}{m_{i,t_1}} + C_{i,t_1},$$

   where by continuity of u_i $u_i(t_1) = \mu_{i,t_1} \frac{2t_1}{m_{i,t_1}} + C_{i,t_1}$, and consequently

$$u_i(t) = u_i(t_1) + \mu_{i,t_1} \frac{2(t - t_1)}{m_{i,t_1}}.$$

   Now we can continue in the same way by considering $[t_2, t_3)$ and so on. Since we have only a finite number N of pixels and some of these pixels merge at the points t_j the process stops after a finite number of n steps with output

$$u_i(t_n) = \frac{1}{N} \sum_{k=0}^{N-1} f_k$$

   for all $i = 0, \ldots, N-1$.

Conversely, it is easy to check that a function u with (i)–(iii) is a solution of the system (25). This completes the proof of the proposition. □

4.3.2 Discrete TV Regularization

Next we will prove that discrete TV regularization satisfies the same rules as space-discrete TV diffusion. For given initial data $f = (f_0, \ldots, f_{N-1})$ discrete TV regularization consists in constructing the minimizer

$$u(\alpha) = \min_u E(u; \alpha, f) \tag{34}$$

of the functional

$$E(u; \alpha, f) = \sum_{i=0}^{N-1} \left((u_i - f_i)^2 + 2\alpha |u_{i+1} - u_i| \right), \tag{35}$$

where we suppose again Neumann boundary conditions $u_{-1} = u_0$ and $u_N = u_{N-1}$.

For a fixed regularization parameter $\alpha \geq 0$, the minimizer of (35) is uniquely determined since $E(u;\alpha,f)$ is strictly convex in $u_0,\ldots,u_{N-1}$. Further, $E(u,\alpha;f)$ is a continuous function in $u_0,\ldots,u_{N-1},\alpha$. Consequently, $u(\alpha)$ is a (componentwise) continuous function in α.
The following proposition implies together with Proposition 4.1 the equivalence of space-discrete TV diffusion and discrete TV regularization.

Proposition 4.2 (Properties of Discrete TV Regularization)
The function $u(\alpha)$ in (34) is uniquely determined by the following rules:

(i) *(Finite Extinction Parameter)*
There exists a finite $A \geq 0$ such that for all $\alpha \geq A$ the signal becomes constant:

$$u_i(\alpha) = \frac{1}{N}\sum_{k=0}^{N-1} f_k \qquad \text{for all } i = 0,\ldots,N-1.$$

(ii) *(Finite Number of Merging Events)*
There exists a finite sequence $0 = a_0 < a_1 < \ldots < a_{n-1} < a_n = A$ such that the interval $[0,A)$ splits into sub-intervals $[a_j,a_{j+1})$ with the property that for all $i = 0,\ldots,N-2$ either $u_i(\alpha) = u_{i+1}(\alpha)$ or $u_i(\alpha) \neq u_{i+1}(\alpha)$ throughout $[a_j,a_{j+1})$. The absolute difference between neighbouring pixels does not become larger for increasing $\alpha \in [a_j,a_{j+1})$.

(iii) *(Analytical Solution)*
In each of the sub-intervals $[a_j,a_{j+1})$ constant regions of $u(\alpha)$ evolve linearly:
For a fixed index i let us consider a constant region given by

$$u_{i-l+1} = \ldots = u_i = u_{i+1} = \ldots = u_{i+r} \qquad (l \geq 1, r \geq 0) \tag{36}$$

and

$$u_{i-l} \neq u_{i-l+1} \text{ if } i-l \geq 0, \quad u_{i+r} \neq u_{i+r+1} \text{ if } i+r \leq N-2 \tag{37}$$

for all $\alpha \in [a_j,a_{j+1})$. We call (36) a region of size $m_{i,a_j} = l + r$. For $\alpha \in [a_j,a_{j+1})$ let $\triangle\alpha = \alpha - a_j$.
Then $u_i(\alpha)$ is given by

$$u_i(\alpha) = u_i(a_j) + \mu_{i,a_j}\,\frac{2\triangle\alpha}{m_{i,a_j}}\,,$$

where μ_{i,a_j} reflects the relation between the region containing u_i and its neighbouring regions. It is given as follows:
For inner regions (i.e. $i-l \geq 0$ and $i+r \leq N-2$) we have

$$\mu_{i,a_j} = \begin{cases} 0 & \text{if} \quad (u_{i-l}, u_i, u_{i+r+1}) \text{ is strictly monotonic,} \\ 1 & \text{if} \quad u_i \text{ is minimal in } (u_{i-l}, u_i, u_{i+r+1}), \\ -1 & \text{if} \quad u_i \text{ is maximal in } (u_{i-l}, u_i, u_{i+r+1}) \end{cases} \tag{38}$$

and in the boundary case ($i-l+1=0$ or $i+r=N-1$), the evolution is half as fast:

$$\mu_{i,a_j} = \begin{cases} 0 & \text{if} \quad m = N, \\ \frac{1}{2} & \text{if} \quad u_i \text{ is minimal in } (u_{i-l}, u_i, u_{i+r+1}), \\ -\frac{1}{2} & \text{if} \quad u_i \text{ is maximal in } (u_{i-l}, u_i, u_{i+r+1}). \end{cases} \tag{39}$$

Proof: Our proof proceeds in four steps.

1. Let us first verify the solution $u(\alpha)$ of (34) for an arbitrary but fixed $\alpha > 0$.
If $u_i(\alpha)$ is contained in some region of size $m_{i,\alpha}$ with (36), (37), then, in case $i-l \geq 0$ and $i+r \leq N-2$, we have that $u(\alpha)$ can be obtained as minimizer of

$$\begin{aligned} E(u_0, \ldots, u_{i-l}, u_i, u_{i+r+1}, \ldots, u_{N-1}; \alpha, f) \; = \\ \sum_{k=-l+1}^{r} (u_i - f_{i+k})^2 \; + \; 2\alpha \left(|u_i - u_{i-l}| + |u_{i+r+1} - u_i| \right) \\ + F(u_0, \ldots, u_{i-l}, u_{i+r+1}, \ldots, u_{N-1}) \end{aligned}$$

with some function F independent of u_i. By (36), (37) the partial derivative of E with respect to u_i exists and is given by

$$\frac{\partial E}{\partial u_i} = 2 \sum_{k=-l+1}^{r} (u_i - f_{i+k}) \; - \; 4\alpha\mu_{i,\alpha}.$$

Here $\mu_{i,\alpha}$ describes the relation between the region containing u_i and its neighbours for the regularization parameter α as in (38). Setting the partial derivative to zero, we obtain that

$$u_i(\alpha) = \frac{1}{m_{i,\alpha}} \sum_{k=-l+1}^{r} f_{i+k} \; + \; \mu_{i,\alpha} \frac{2\alpha}{m_{i,\alpha}}\,. \tag{40}$$

In the boundary case $i-l+1=0$ or $i+r=N-1$ we follow the same lines and obtain (40) with $\mu_{i,\alpha}$ defined by (39).

2. Next we show that initially merged pixels will not be split for any α in a small interval $[0, a_1]$.

   For $\alpha = 0$ we have that $u(0) = f$. Let $f_i = u_i(0)$ be contained in some region of the form

$$f_{i-l_0+1} = \ldots = f_i = f_{i+1} = \ldots = f_{i+r_0} \qquad (l_0, r_0 \geq 1)$$

   and

$$f_{i-l_0} \neq f_{i-l_0+1} \text{ if } i - l_0 \geq 0, \quad f_{i+r_0} \neq f_{i+r_0+1} \text{ if } i + r_0 \leq N-2.$$

   By continuity of $u(\alpha)$ we can choose $\alpha_1 > 0$ so that $u_i(\alpha) \neq u_{i-l_0}(\alpha)$ and $u_{i+1}(\alpha) \neq u_{i+r_0}(\alpha)$ throughout $[0, \alpha_1)$. Assume that there exists $\alpha \in (0, \alpha_1)$ so that $u_i(\alpha) \neq u_{i+1}(\alpha)$, where we may assume that

$$u_i(\alpha) < u_{i+1}(\alpha). \tag{41}$$

   The opposite case $u_i(\alpha) > u_{i+1}(\alpha)$ can be handled in the same way. Note that at time α more pixels than u_i and u_{i+1} may be separated. However, we have by (40) with some $1 \leq l \leq l_0$ and some $1 \leq r \leq r_0$ that

$$\begin{aligned} u_i(\alpha) &= \frac{1}{l} \sum_{k=-l+1}^{0} f_{i+k} + \mu_{i,\alpha} \frac{2\alpha}{l} = f_i + \mu_{i,\alpha} \frac{2\alpha}{l}, \\ u_{i+1}(\alpha) &= \frac{1}{r} \sum_{k=1}^{r} f_{i+k} + \mu_{i+1,\alpha} \frac{2\alpha}{r} = f_i + \mu_{i+1,\alpha} \frac{2\alpha}{r}, \end{aligned}$$

   where we see by (41) and (38), (39) that $\mu_{i,\alpha} \geq 0$ and $\mu_{i+1,\alpha} \leq 0$. Thus, $u_i(\alpha) \geq u_{i+1}(\alpha)$ which contradicts (41). Consequently $u_i(\alpha) = u_{i+1}(\alpha)$ throughout $[0, \alpha_1)$, i.e., the pixels of our initial region stay merged.

   Let $a_1 > 0$ denote the largest number such that no merging of regions appears in $[0, a_1)$. Then we have for all $i = 0, \ldots, N-1$ and all $\alpha \in [0, a_1)$ that $\mu_{i,\alpha} = \mu_{i,0}$ and regarding that $u(\alpha)$ is continuous that

$$u_i(\alpha) = f_i + \mu_{i,0} \frac{2\alpha}{m_{i,0}} \qquad (\alpha \in [0, a_1]). \tag{42}$$

3. Now we show that the absolute difference between neighbouring regions cannot become larger with increasing $\alpha \in [0, a_1)$.
   Without loss of generality let for some fixed index i

$$u_{i-l+1} = \ldots = u_i < u_{i+1} = \ldots = u_{i+r} \qquad (l, r \geq 1)$$

and

$$u_{i-l} \neq u_{i-l+1} \text{ if } i-l \geq 0, \quad u_{i+r} \neq u_{i+r+1} \text{ if } i+r \leq N-2.$$

We consider the non-boundary case $i-l \geq 0$ and $i+r \leq N-2$ first. By (42) we obtain for $\alpha+\delta \in [0, a_1)$, $\delta > 0$ that

$$\begin{aligned} d_i(\alpha) &= u_{i+1}(\alpha) - u_i(\alpha) = f_{i+1} - f_i + 2\alpha\left(\frac{\mu_{i+1,0}}{r} - \frac{\mu_{i,0}}{l}\right), \\ d_i(\alpha+\delta) &= u_{i+1}(\alpha+\delta) - u_i(\alpha+\delta) \\ &= f_{i+1} - f_i + 2(\alpha+\delta)\left(\frac{\mu_{i+1,0}}{r} - \frac{\mu_{i,0}}{l}\right) \end{aligned}$$

and consequently

$$d_i(\alpha+\delta) - d_i(\alpha) = 2\delta\left(\frac{\mu_{i+1,0}}{r} - \frac{\mu_{i,0}}{l}\right).$$

By (38) it follows that

$$\frac{\mu_{i+1,0}}{r} - \frac{\mu_{i,0}}{l} = \begin{cases} 0 & \text{if } u_{i-l} < u_i < u_{i+1} < u_{i+r+1}, \\ -\frac{1}{r} & \text{if } u_{i-l} < u_i \text{ and } u_{i+1} > u_{i+r+1}, \\ -\frac{1}{l} & \text{if } u_{i-l} > u_i \text{ and } u_{i+1} < u_{i+r+1}, \\ -\frac{1}{r} - \frac{1}{l} & \text{if } u_{i-l} > u_i \text{ and } u_{i+1} > u_{i+r+1} \end{cases}$$

which yields the desired property $d_i(\alpha) \geq d_i(\alpha+\delta)$.

In case of boundary regions we follow the same lines but replace (38) by (39). Then we see that the absolute difference between neighboring regions becomes smaller with increasing $\alpha \in [0, a_1)$.

4. We are now in the position to analyse the entire chain of merging events successively.
   For $\alpha > a_1$ and $\triangle\alpha = \alpha - a_1$, we consider

   $$\tilde{u}_i(\triangle\alpha) = \min_u E(u; \triangle\alpha, u(a_1)).$$

   We can repeat the same considerations as in Part 2 of the proof but with initial setting $u(a_1)$ instead of f. It follows that there exists a_2 such that for all $i = 0, \ldots, N-2$ either $\tilde{u}_i(\triangle\alpha) = \tilde{u}_{i+1}(\triangle\alpha)$ or $\tilde{u}_i(\triangle\alpha) \neq \tilde{u}_{i+1}(\triangle\alpha)$ throughout $[a_1, a_2)$, where the absolute difference between neighbouring pixels does not become larger for increasing $\triangle\alpha$. Further, we obtain by (42) and (40) that

   $$\begin{aligned} \tilde{u}_i(\triangle\alpha) &= u_i(a_1) + \mu_{i,a_1}\frac{2\triangle\alpha}{m_{i,a_1}} \\ &= \frac{1}{m_{i,a_1}}\sum_{j \in R_{i,a_1}} f_j + \mu_{i,a_1}\frac{2a_1}{m_{i,a_1}} + \mu_{i,a_1}\frac{2\triangle\alpha}{m_{i,a_1}}, \end{aligned}$$

where $R_{i,\alpha} = \{j : u_j(\alpha) \text{ is in the region of } u_i(\alpha)\}$ while m_{i,a_1} denotes the size of the region containing $u_i(a_1)$ and μ_{i,a_1} reflects the relation between the region containing $u_i(a_1)$ and its neighbouring regions. Since the relations between regions do not change for $\triangle\alpha \in [0, a_2 - a_1)$ we can rewrite $\tilde{u}_i(\triangle\alpha)$ as

$$\begin{aligned} \tilde{u}_i(\triangle\alpha) &= \frac{1}{m_{i,a_1+\triangle\alpha}} \sum_{j \in R_{i,a_1+\triangle\alpha}} f_j + \mu_{i,a_1+\triangle\alpha} \frac{2(a_1 + \triangle\alpha)}{m_{i,a_1+\triangle\alpha}} \\ &= \frac{1}{m_{i,\alpha}} \sum_{j \in R_{i,\alpha}} f_j + \mu_{i,\alpha} \frac{2\alpha}{m_{i,\alpha}}. \end{aligned}$$

On the other hand, we have by (40) that

$$u_i(\alpha) = \frac{1}{m_{i,\alpha}} \sum_{j \in R_{i,\alpha}} f_j + \mu_{i,\alpha} \frac{2\alpha}{m_{i,\alpha}} .$$

Thus, $u_i(\alpha) = \tilde{u}_i(\triangle\alpha)$.

Now we can continue in the same way by considering $[a_2, a_3)$ and so on. Since we have only a finite number N of pixels and some of these pixels merge at the points a_j the process stops after a finite number of n steps with output $u(a_n)$ which by (40) reads as

$$u_i(a_n) = \frac{1}{N} \sum_{k=0}^{N-1} f_k$$

for all $i = 0, \ldots, N-1$. This completes the proof. □

Analytical results for some convex regularization problems applied to specific test signals have been presented by Li [39]. Strong [65] derived analytical results in the case of continuous TV regularization methods with step functions as initializations. Equivalent results have been obtained by Mammen and van de Geer [45] for the taut-string algorithm in statistics; see also [36]. Our results are in accordance with these findings, but our proof shows that they can be derived in a different way: The structure of our proof is in complete analogy with the proof for the TV diffusion case. After the submission of our manuscript, an alternative way of deriving explicit solutions for TV diffusion has been proposed by Bellettini et al. [11] by considering an eigenvalue problem.

4.4 Equivalence to SIDEs with TV Force Functions

In Section 2.4 we have seen that 1-D SIDEs with region size 1 and TV force function are identical to space-discrete TV diffusion. Moreover, in Section 4.3 we have derived analytical solutions of space-discrete TV diffusion and discrete TV regularization that show the same merging behaviour as SIDEs with TV force functions. Consequently, 1-D SIDEs can be interpreted as an exact solution of space-discrete TV diffusion or regularization in general.
This also confirms that the merging steps in the SIDE evolution are much more than a heuristic stabilization that speeds up the evolution: They are a natural consequence of the degenerated diffusivities that are unbounded in 0. Last but not least, our considerations can be regarded as a theoretical justification of region merging in terms of variational and PDE-based techniques.

5 Multiple Scales

So far we have only considered soft wavelet shrinkage on a single scale. In almost all practical applications, however, wavelet shrinkage is performed on multiple scales. In this section, we interpret *multiscale* soft shrinkage with Haar wavelets as application of nonlinear TV-based diffusion to two-pixel groups of *hierarchical* signals. First we consider the standard situation without shift invariance, then we discuss the shift-invariant case. Finally, we address a frequent problem that occurs with wavelet shrinkage on multiple scale: the presence of Gibbs-like artifacts. We analyse ways to circumvent this phenomenon by using scale-dependent thresholds.
Throughout this section we deal with signals of length $N = 2^n$ ($n \in \mathbb{N}$).

5.1 Standard Case without Shift Invariance

Haar wavelet shrinkage on two scales is described by the filter bank in Figure 4. To obtain more than two scales we further split up the upper branch of the inner filter bank and so on until we arrive at scale $n = \log_2 N$, where the successive downsampling by 2 results in a one-pixel signal.
Next we briefly recall the concept of *Gaussian* and *Laplacian pyramids* [14] with respect to the Haar filters. The Gaussian pyramid we are interested in is the sequence of H_0-smoothed and downsampled versions of an initial signal f given by

$$f = f^{(0)} \longrightarrow f^{(1)} = Rf \longrightarrow \ldots \longrightarrow f^{(n)} = R^n f,$$

where R denotes the operator for H_0-smoothing and subsequent downsampling by 2, i.e.,

$$f_i^{(j+1)} = (Rf^{(j)})_i = \frac{f_{2i}^{(j)} + f_{2i+1}^{(j)}}{\sqrt{2}} \quad (j = 0, \ldots, n-1;\ i = 0, \ldots, \tfrac{N}{2^{j+1}} - 1).$$

Let $Pf^{(j)}$ denote the prolongated version of $f^{(j)}$ given by

$$(Pf^{(j)})_{2i} = (Pf^{(j)})_{2i+1} = \frac{f_i^{(j)}}{\sqrt{2}} \quad (j = 1, \ldots, n;\ i = 0, \ldots, \tfrac{N}{2^j} - 1). \tag{43}$$

Then the corresponding Laplacian pyramid is the sequence

$$f - Pf^{(1)} \longrightarrow f^{(1)} - Pf^{(2)} \longrightarrow \ldots \longrightarrow f^{(n-1)} - Pf^{(n)} \longrightarrow f^{(n)}.$$

By

$$f^{(j)} = Pf^{(j+1)} + \left(f^{(j)} - Pf^{(j+1)}\right) \qquad (j = n-1, \ldots, 0)$$

we can reconstruct f from its Laplacian pyramid.
Let diff_t denote the operator of nonlinear diffusion with TV diffusivity and stopping time t, applied to the successive two-pixel parts of a signal. By Subsection 3.2 we know that diff_t performs like a single wavelet shrinkage step with soft threshold parameter $\tau = \sqrt{2}t$. In other words, the result of the filter bank in Figure 2 is $u = \mathrm{diff}_t(f)$. Further, we see that the upper branch of this filter bank produces $Pf^{(1)}$ so that the lower branch must produce $\mathrm{diff}_t(f) - Pf^{(1)}$. By (43) and (3.1) it is easy to check that the nonlinear operator diff_t fulfils

$$\mathrm{diff}_t(f) - Pf^{(1)} = \mathrm{diff}_t(f - Pf^{(1)}).$$

Thus, one wavelet shrinkage step is given by

$$u = Pf^{(1)} + \mathrm{diff}_t(f - Pf^{(1)}).$$

Now the multiscale Haar wavelet shrinkage up to scale n can be described by successive application of diff_t to the Laplacian pyramid:

$$u^{(n)} = f^{(n)}, \tag{44}$$
$$u^{(j)} = Pu^{(j+1)} + \mathrm{diff}_t(f^{(j)} - Pf^{(j+1)}) \qquad (j = n-1, \ldots, 0). \tag{45}$$

The result of the multiscale wavelet shrinkage is $u = u^{(0)}$.

5.2 Shift-Invariant Case

Now we consider translation invariant multiscale wavelet shrinkage. In the multiscale setting we apply cycle spinning over the range of all N shifts of f. This is equivalent to denoising using the undecimated or stationary wavelet transform [25, 27, 44]. The filter bank which corresponds to two scales of translation-invariant Haar wavelet shrinkage is shown in Figure 5. Note that the inner filter bank uses z^2 instead of z in H_i $(i = 0, 1)$. In general we have to replace z by $z^{2^{j-1}}$ at scale j. While ordinary wavelet shrinkage requires $O(N)$ arithmetic operations, its translation invariant version needs $O(N \log_2 N)$ arithmetic operations.

In Subsection 4.2 we have deduced a numerical scheme for TV diffusion. Each iteration is given by (22). This coincide with a single translation invariant Haar wavelet shrinkage step with threshold $\tau = 2\sqrt{2}t$. Using our operator diff. and the operator S which shifts a signal one pixel to the right, the result u of the single scale translation invariant filter bank in Figure 3 is given by

$$u = \frac{1}{2} \left(\mathrm{diff}_{2t}(f) + S^{-1} \mathrm{diff}_{2t}(Sf) \right).$$

Now the multiscale translation invariant Haar wavelet shrinkage can be interpreted as application of diff. to a multiple Laplacian pyramid. We define a multiple Gaussian pyramid by

$$\begin{aligned} f^{(0,0)} &\longrightarrow \left(f^{(1,0)}, f^{(1,1)} \right) \longrightarrow \left(f^{(2,0)}, f^{(2,1)}, f^{(2,2)}, f^{(2,3)} \right) \longrightarrow \ldots \\ &\longrightarrow \left(f^{(n,0)}, \ldots, f^{(n,2^n-1)} \right), \end{aligned}$$

where $f = f^{(0,0)}$. Here $f^{(j,k)}$ is obtained by successive application of the operators R and RS on f as follows: Let 0 denote the application of R and 1 the application of RS, then these operators are applied to f in the order of the binary representation $(k_{j-1}, \ldots, k_0)_2$ of k, where we start from the left. For example we get $f^{(2,1)} = f^{(2,(0,1)_2)} = RS\,Rf$ and $f^{(2,2)} = f^{(2,(1,0)_2)} = R\,RSf$. Then the multiple Laplacian pyramid is given by

$$\begin{aligned} &\left(f^{(0,0)} - Pf^{(1,0)}, Sf^{(0,0)} - Pf^{(1,1)} \right) \\ &\quad \longrightarrow \left(f^{(1,0)} - Pf^{(2,0)}, Sf^{(1,0)} - Pf^{(2,1)}, f^{(1,1)} - Pf^{(2,2)}, Sf^{(1,1)} - Pf^{(2,3)} \right) \\ &\quad \longrightarrow \ldots \longrightarrow \left(f^{(n,0)}, \ldots, f^{(n,2^n-1)} \right) \end{aligned}$$

and the translation invariant version of (44)–(45) can be obtained from this multiple Laplacian pyramid by

$$\begin{aligned}
u^{(n,k)} &= f^{(n,k)} \qquad (k = 0, \ldots, 2^n - 1), \\
u^{(j,k)} &= \frac{1}{2} \left(Pu^{(j+1,2k)} + \mathrm{diff}_{2t}(f^{(j,k)} - Pf^{(j+1,2k)}) \right. \\
&\quad \left. + S^{-1} \left(Pu^{(j+1,2k+1)} + \mathrm{diff}_{2t}(Sf^{(j,k)} - Pf^{(j+1,2k+1)}) \right) \right) \\
&\qquad\qquad\qquad (j = n-1, \ldots, 0; k = 0, \ldots, 2^j - 1).
\end{aligned}$$

The result of the multiscale translation invariant wavelet shrinkage is $u = u^{(0,0)}$.

5.3 Scale-Dependent Thresholds

Cycle spinning techniques can be used to make wavelet shrinkage not only translationally invariant, but also to reduce artifacts. However, it is still possible that oscillatory (Gibbs-like) artifacts appear if multiple scales are used. We want to demonstrate that the use of the scale-dependent thresholds

$$\tau_j = \frac{\tau}{\sqrt{2}^{j-1}} \qquad (j = 1, \ldots, n) \tag{46}$$

suppresses oscillations in the shrinkage process.

In this subsection, we consider signals $f = (f_0, \ldots, f_{N-1})$ with periodic boundary conditions. Note that mirror boundary conditions can easily be transferred into periodic ones by doubling the signal.

The decimated Haar wavelet shrinkage with full n-scale decomposition and thresholds (46) consists of three operations. It starts with the linear transform (2) of f yielding the wavelet coefficients $(c^n, d^n, d^{n-1}, \ldots, d^1)$, where $d^j := (d_0^j, \ldots, d_{N/2^j-1}^j)$. The wavelet coefficients then undergo the soft wavelet thresholding

$$s_i^j := S_{\tau_j}(d_i^j) \qquad (j = 1, \ldots, n; \quad i = 0, \ldots, N/2^j - 1)$$

followed by the inverse linear transform (3) of $(c^n, s^n, s^{n-1}, \ldots, s^1)$ which gives the denoised signal $u(\tau)$. In particular we have $u(0) = f$. Note that by the semigroup property

$$S_{\tau+\tilde{\tau}}(x) = S_{\tilde{\tau}}\left(S_\tau(x)\right)$$

of our shrinkage function (1), the signal $u(K\tau)$ obtained by one n-scale wavelet shrinkage cycle with threshold $K\tau$ coincides with the signal which results from K times repeating one n-scale wavelet shrinkage cycle with smaller

threshold τ. Of course this is no longer true for the translation-invariant wavelet transform. In our examples in the next section we will consider iterated translation-invariant Haar wavelet shrinkage with small thresholds τ. Since oscillatory (Gibbs-like) artifacts are characterized by the emergence of new local extrema, we study the behaviour of local extrema of the signal under the shrinkage process. We call u_i an *extremal pixel*, if either $u_{i-1} < u_i$, $u_i > u_{i+1}$ or $u_{i-1} > u_i$, $u_i < u_{i+1}$.
First we consider the dynamics of "infinitesimal translation-invariant soft Haar wavelet shrinkage", i.e. the speed at which pixels of the signal evolve with respect to the threshold $\tau \in [0, T]$ in the limit case $T \to 0$.

Proposition 5.1 (Suppression of Gibbs-like Artifacts by Scaled Thresholds)
Under infinitesimal translation-invariant soft Haar wavelet shrinkage, an extremal pixel f_i evolves as follows:

(i) The value of the extremal pixel decreases, i.e. $\dot{u}_i < 0$, if it is a maximum and increases, i.e. $\dot{u}_i > 0$, if it is a minimum. Here the dot denotes differentiation with respect to τ.

(ii) The absolute value of the difference of the extremal pixel to each of its two neighbours decreases, i.e. $\dot{u}_i - \dot{u}_{i\pm1} < 0$ for a maximum and $\dot{u}_i - \dot{u}_{i\pm1} > 0$ for a minimum.

Statement (i) holds also for the decimated Haar wavelet shrinkage while statement (ii) cannot be established in that setting.

Proof: For the decimated Haar wavelet shrinkage with full n-scale decomposition and thresholds (46) it is easy to check that the resulting signal $\tilde{u}_i$ is given by

$$\tilde{u}_i = \mu + \sum_{j=1}^{n} 2^{-j/2} \varepsilon_j(i) s^j_{\lfloor i/2^j \rfloor},$$

where $\lfloor x \rfloor$ denotes the largest integer $\leq x$. Moreover, $\mu := \frac{1}{N} \sum_{i=0}^{N-1} f_i$ is the average value, and

$$\varepsilon_j(i) := \begin{cases} 1 & \text{if } \lfloor i/2^{j-1} \rfloor \text{ is even,} \\ -1 & \text{if } \lfloor i/2^{j-1} \rfloor \text{ is odd.} \end{cases}$$

For the translation-invariant Haar wavelet shrinkage, the sum on the right-hand side of this equation is replaced by the average of N sums of the same kind containing the back-shifted shrunken wavelet coefficients of N forward-shifted initial signals, i.e.

$$u_i = \mu + \frac{1}{N} \sum_{\nu=0}^{N-1} \sum_{j=1}^{n} 2^{-j/2} \varepsilon_j(i+\nu) s^j_{\lfloor (i+\nu)/2^j \rfloor, \nu} \tag{47}$$

where $s^j_{i,\nu}$ denotes the i-th coefficient of the j-th level of the ν-shifted initial signal, and the coefficients are treated $N/2^j$-periodic with respect to i. Of course, some coefficients coincide for different ν, more precisely

$$s^j_{\lfloor (i+\nu)/2^j \rfloor, \nu} = s^j_{\lfloor (i+\nu+r2^j)/2^j \rfloor, \nu+r2^j} \qquad (\nu = 0, \ldots, 2^j - 1; \quad r = 0, \ldots, N/2^j - 1).$$

This equation allows us to rewrite (47) as

$$\begin{aligned} u_i &= \mu + \frac{1}{2\sqrt{2}} \left(\varepsilon_1(i) s^1_{\lfloor i/2 \rfloor, 0} + \varepsilon_1(i+1) s^1_{\lfloor (i+1)/2 \rfloor, 1} \right) \\ &\quad + \frac{1}{N} \sum_{j=2}^{n} 2^{-j/2} \sum_{\nu=0}^{2^j-1} \sum_{r=0}^{N/2^j-1} \varepsilon_j(i+\nu+r2^j) s^j_{\lfloor (i+\nu+r2^j)/2^j \rfloor, \nu+r2^j}, \\ &= \mu + \frac{s_{i,+} - s_{i,-}}{2\sqrt{2}} + \sum_{j=2}^{n} 2^{-3j/2} \sum_{\nu=0}^{2^j-1} \varepsilon_j(i+\nu) s^j_{\lfloor (i+\nu)/2^j \rfloor, \nu}, \end{aligned}$$

where $s_{i,+} := S_\tau(d_{i,+}) = S_\tau\big((f_i - f_{i+1})/\sqrt{2}\big)$ and $s_{i,-} := S_\tau(d_{i,-}) = S_\tau\big((f_{i-1} - f_i)/\sqrt{2}\big)$. Now the evolution of u_i under infinitesimal soft wavelet shrinkage is described by

$$\dot{u}_i = \frac{\dot{s}_{i,+} - \dot{s}_{i,-}}{2\sqrt{2}} + \sum_{j=2}^{n} 2^{-3j/2} \sum_{\nu=0}^{2^j-1} \varepsilon_j(i+\nu) \dot{s}^j_{\lfloor (i+\nu)/2^j \rfloor, \nu}, \tag{48}$$

where

$$\dot{s}^j_{\cdot} = \frac{\mathrm{d}S_{\tau_j}(d^j_{\cdot})}{\mathrm{d}\tau_j} \cdot \frac{\mathrm{d}\tau_j}{\mathrm{d}\tau} = \frac{-\operatorname{sgn}(d^j_{\cdot})}{\sqrt{2}^{j-1}}.$$

Inserting this into (48), we obtain

$$\dot{u}_i = \frac{-\operatorname{sgn}(d_{i,+}) + \operatorname{sgn}(d_{i,-})}{2\sqrt{2}} - A_i, \tag{49}$$

where

$$A_i := \sqrt{2} \sum_{j=2}^{n} 4^{-j} \sum_{\nu=0}^{2^j-1} \varepsilon_j(i+\nu) \operatorname{sgn}(d^j_{\lfloor (i+\nu))/2^j \rfloor, \nu}).$$

By the triangle inequality we can estimate

$$|A_i| \leq \sqrt{2} \sum_{j=2}^{n} 2^{-j} < \frac{1}{\sqrt{2}}. \tag{50}$$

If f_i is an extremal pixel, then we have that $\operatorname{sgn}(d_{i,+}) = -\operatorname{sgn}(d_{i,-}) = 1$ for a maximum and -1 for a minimum. This implies by (49) and (50) that

$$\operatorname{sgn}(\dot{u}_i) = -\operatorname{sgn}(d_{i,+}) \tag{51}$$

proving statement (i) of the proposition.
By subtracting from (49) its counterpart for pixel u_{i+1}, we obtain by $d_{i,+} = d_{i+1,-}$ that

$$\dot{u}_i - \dot{u}_{i+1} = \frac{\operatorname{sgn}(d_{i,-}) - 2\operatorname{sgn}(d_{i,+}) + \operatorname{sgn}(d_{i+1,+})}{2\sqrt{2}} - (A_i - A_{i+1}). \tag{52}$$

In

$$\begin{aligned} A_i - A_{i+1} = \sqrt{2} \sum_{j=2}^{n} 4^{-j} \sum_{\nu=0}^{2^j-1} \Big(& \varepsilon_j(i+\nu) \operatorname{sgn}(d^j_{\lfloor (i+\nu)/2^j \rfloor, \nu}) \\ & - \varepsilon_j(i+1+\nu) \operatorname{sgn}(d^j_{\lfloor (i+1+\nu)/2^j \rfloor, \nu}) \Big) \end{aligned}$$

the values in the inner brackets cancel except for the two indices $\nu = \nu_k^j \in \{0, \ldots, 2^j - 1\}$ $(k = 0, 1)$ with $\nu_k^j + 1 + i \equiv 0 \bmod 2^{j-1}$. For these indices the signs of $\varepsilon_j(i + \nu_k)$ and $\varepsilon_j(i + 1 + \nu_k)$ are opposite. Consequently, for each j, the inner sum contains only four summands, and we can estimate

$$|A_i - A_{i+1}| \leq \sqrt{2} \sum_{j=2}^{n} 4^{-j} \cdot 4 < \frac{\sqrt{2}}{3}. \tag{53}$$

By inserting this into (52), it becomes clear that for an extremal pixel f_i we get

$$\operatorname{sgn}(\dot{u}_i - \dot{u}_{i+1}) = -\operatorname{sgn}(d_{i,+}). \tag{54}$$

We have therefore proven that the difference of an extremal pixel to its right neighbour decreases under infinitesimal soft wavelet shrinkage. Analogous considerations apply to the left neighbour which completes the proof of (ii). $\square$

It follows particularly from Proposition 5.1 that under iterated infinitesimal soft wavelet shrinkage, no oscillatory (Gibbs-like) artifacts can appear. Any

artifact of this type would include at least one local extremum evolving from a flat region which would, for continuity, have to grow over a finite time interval in contradiction to Proposition 5.1.
It should be noted that a single step of infinitesimal shrinkage does not effectively change the signal any more since $T \to 0$. To investigate true changes of the signal by the shrinkage procedure, one has to consider iterated shrinkage. Summing up τ over all iteration steps, a "total evolution time" t is obtained; for fixed t, the number of iteration steps tends to infinity as τ goes to zero. Infinitesimal translation-invariant soft Haar wavelet shrinkage thus becomes a dynamic process parametrized by t and Proposition 5.1 describes its behaviour at a single point of time.
Of course, this analysis can be extended to a time interval. Then one has to take care of the discontinuity of sgn at 0. Similarly as in the proof of Proposition 4.1 this can be done by splitting the time axis into intervals in which no sign changes of wavelet coefficients occur. However, since once-merged pixels can (and will in general) split again in the process considered here, sgn(0) will in most cases occur only in discrete time points.

Now we turn to consider finite-size shrinkage steps τ. The ideas used in the proof of Proposition 5.1 can also be applied to analyse soft wavelet shrinkage with finite threshold τ by simply replacing the derivatives $\dot{u}_i$, $\dot{s}^j_{i,\nu}$ by differences $\Delta u_i := u_i(\tau) - u_i(0)$ and $\Delta s^j_{i,\nu} := s^j_{i,\nu} - d^j_{i,\nu}$, respectively. Then we obtain instead of (49) that

$$\Delta u_i = \frac{(s_{i,+} - d_{i,+}) - (s_{i,-} - d_{i,-})}{2\sqrt{2}} + A_i,$$

where

$$A_i = \sum_{j=2}^{n} 2^{-3j/2} \sum_{\nu=0}^{2^j-1} \varepsilon_j(i+\nu) \Delta s^j_{\lfloor (i+\nu)/2^j \rfloor, \nu}.$$

By (46) and (1) we obtain instead of (50) the estimate

$$|A_i| \le \tau\sqrt{2} \sum_{j=2}^{n} 2^{-j} < \frac{\tau}{\sqrt{2}}.$$

However, the implication from inequality (50) to equation (51) can be transferred only if $|d_{i,+}| \ge \tau$ and $|d_{i,-}| \ge \tau$.
Similarly, we conclude instead of (52) that

$$\Delta u_i - \Delta u_{i+1} = \frac{-\Delta s_{i,-} + 2\Delta s_{i,+} - \Delta s_{i+1,+}}{2\sqrt{2}} + (A_i - A_{i+1}) \qquad (55)$$

and estimate the later difference by

$$|A_i - A_{i+1}| \leq \tau\sqrt{2}\sum_{j=2}^{n} 4^{-j}\, 4 < \frac{\tau\sqrt{2}}{3}. \tag{56}$$

However, the conclusion from (53) to (54) can be transferred only if $|\Delta s_{i,-} - 2\Delta s_{i,+} + \Delta s_{i+1,+}| \geq 4\tau/3$. The latter holds true if (but not only if) $|d_{i,+}| \geq \tau$ and $|d_{i,-}| \geq \tau$, i.e., if $f_i - f_{i\pm1} \geq \sqrt{2}\tau$. In this case we obtain by (55), (56) and their counterparts for the left neighbours of f_i that

$$-\frac{\tau\sqrt{2}}{3} \leq u_i(\tau) - u_{i\pm1}(\tau) \leq f_i - f_{i\pm1} - \frac{\tau\sqrt{2}}{6}$$

if f_i is a maximum. Analogous inequalities hold true if f_i is a minimum. We can therefore state the following corollary.

Corollary 5.2 (Behaviour of Extrema under Haar Wavelet Shrinkage)
Under translation-invariant soft Haar wavelet shrinkage with thresholds (46) an extremal pixel f_i, which differs at least by $\sqrt{2}\tau$ from each of its neighbours, evolves as follows:

(i) The value of the extremal pixel decreases, i.e. $\Delta u_i < 0$, if it is a maximum and increases, i.e. $\Delta u_i > 0$, if it is a minimum.

(ii) The absolute value of the difference of the extremal pixel to each of its two neighbours decreases; in particular, one has $\Delta u_i - \Delta u_{i\pm1} < 0$ for a maximum and $\Delta u_i - \Delta u_{i\pm1} > 0$ for a minimum.

It can be shown by examples that each of the statements (i) and (ii) of the corollary can be violated if the extremal pixel f_i differs from its neighbours by not more than $\sqrt{2}\tau$. In summary, it follows that Gibbs-like artifacts can in principle still occur under finite-size steps of soft Haar wavelet shrinkage but are restricted in amplitude.

6 Experiments

In this section we illustrate the interplay of iterations and multiscale soft Haar wavelet shrinkage by two examples. As in the previous section we consider initial signals $f = (f_0, \ldots, f_{N-1})$, where $N = 2^n$ is a power of 2. Furthermore, we restrict our attention to reflecting (Neumann) boundary conditions. Then we can perform multiscale wavelet shrinkage up to some assigned scale $m \leq n$.

We start with a simple example which demonstrates the influence of the interplay between iterations and multiscale wavelet shrinkage on Gibbs-like artifacts and its relation to TV diffusion. We consider the initial signal in Fig. 6 and apply iterated translation-invariant single-scale and multi-scale soft Haar wavelet shrinkage with various threshold parameters. The resulting signals are presented in Fig. 7.
Consider the left column of Figure 7. In Subsection 4.2 we have shown that translation-invariant soft Haar wavelet shrinkage corresponds to a stable numerical scheme for TV diffusion which represents real TV diffusion if the shrinkage parameter τ is small enough. The first row demonstrates the local effect of the single-scale wavelet shrinkage with threshold $\tau = 1$. The K-times iterated processes with thresholds $\tau = 1/K$ in the second and third rows spreads the information globally over the signal. For $K = 1000$, the scheme is a very good approximation to TV diffusion.
The middle and the right columns of Fig. 7 deal with multi-scale wavelet shrinkage which does not fully correspond to TV diffusion. Already a single iteration results in global effects here. Iterating the multiple-scale wavelet shrinkage flattens homogeneous regions, as desired also in TV diffusion. In the middle column, we can observe Gibbs-like phenomena. In the right column, they are avoided by scaling the thresholds.

In our second example we are concerned with the initial signal in Fig. 8 obtained using the *WaveLab* package [44]. Figure 9 presents the denoised signal, where the parameter of each method (threshold value or number of iterations) was chosen to optimize the signal-to-noise ratio on the output. We have applied the following techniques:

A. 1 level, regularized TV scheme (24) with $\epsilon = \frac{0.04}{2\sqrt{2}}$, iterated with $\tau = \frac{0.01}{2\sqrt{2}}$, $K = 53707$ iterations

B. 1 level two-pixel scheme (22), iterated with $\tau = 0.01$, $K = 53707$ iterations

C. 13 levels, 1 iteration, uniform threshold $\tau = 37.4$

D. 13 levels, iterated, $\tau = 0.01$, $K = 3244$ iterations

E. 13 levels, 1 iteration, scaled thresholds, $\tau = 92.6$

F. 13 levels, iterated, $\tau = 0.01$, $K = 7800$ iterations

Table 1 shows the signal-to-noise ratio of the denoised signals. In Table 1, the best restoration results in terms of the signal-to-noise ratio are obtained using the regularized TV diffusion scheme (A), iterated single-scale wavelet shrinkage (B) or the iterated n-scale wavelet shrinkage with adapted thresholds (E). Although these methods are not exactly equivalent, they reveal a

method	SNR	$K \cdot \tau$
A. regularized TV flow	24.6	537.1
B. single-level two-pixel scheme, iterated	24.5	537.1
C. multi-level, single threshold, single step	18.3	37.4
D. multi-level, single threshold, iterated	21.3	32.4
E. multi-level, scaled thresholds, single step	21.9	92.6
F. multi-level, scaled thresholds, iterated	24.3	78.0

Table 1: Numerical evaluation of the filtering performance of several methods on the data of Fig. 8. The filtered data are shown in Fig. 9.

high level of visual similarity and provide a good piecewise constant approximation to the original signal. The single step multi-scale wavelet shrinkage with scale-adapted threshold (E) performs slightly worse. The single step and iterated multi-scale wavelet shrinkage techniques with a uniform threshold on all scales (C, D) are less satisfactory, also visually.
These experiments show that TV denoising (which approximates iterated shift-invariant soft Haar wavelet shrinkage on a single scale) outperforms many soft wavelet shrinkage strategies. On the other hand, this is at the expense of a relatively high numerical effort. In order to make wavelets competitive, the shrinkage should be shift invariant, iterative, and use multiple scales with scaled thresholds. In those cases where it is possible to reduce the number of iterations without severe quality degradations, one obtains a hybrid method that combines the speed of multiscale wavelet techniques with the quality of variational or PDE-based denoising methods. For more experiments on multiscale ideas versus iterations and for an evaluation of the computational complexity of these techniques, we refer to [48].

7 Summary

The goal of the present paper was to investigate under which conditions one can prove equivalence between four discontinuity preserving denoising techniques in the 1-D case: soft wavelet thresholding, TV diffusion, TV regularization, and SIDEs. Starting from a simple two-pixel case we were able to derive analytical solutions. These two-pixel solutions have been used

- to establish equivalence between soft Haar wavelet shrinkage with threshold parameter τ and TV diffusion of two-pixel signal pairs with diffusion time $t = \tau/\sqrt{2}$.

- to prove also equivalence to TV regularization of two-pixel pairs with regularization parameter $\alpha = \tau/\sqrt{2}$.
- to conjecture equivalence of space-discrete TV-diffusion and discrete TV regularization for general N-pixel signals. This conjecture has been proven in a later section.
- to prove that space-discrete TV diffusion and discrete TV regularization are also equivalent to a SIDE evolution when a TV-based force function is used. This gives a sound theoretical justification for the heuristically introduced evolution rules for SIDEs.
- to design a novel numerical scheme for TV diffusion of N-pixel signals. It is based on an additive operator splitting (AOS) into two-pixel interactions where analytical solutions exist for arbitrary large time step sizes. Thus, the numerical scheme is explicit and absolutely stable.

We showed that wavelet shrinkage on multiple scales can also be regarded as two-pixel TV diffusion or regularization on the Laplacian pyramid of the signal. On the wavelet side, our experiments show that one can improve the denoising performance by rescaling the thresholds for each wavelet level, and by iterating the translation-invariant wavelet shrinkage. On the PDE side, it is possible to achieve a speed-up without significant quality deterioration by using iterated multiple scales instead of iterated single scale denoising. Thus, the resulting hybrid methods combine the advantages of wavelet and PDE based denoising.
In our future work we intend to consider more advanced wavelet methods (other shrinkage functions, different wavelets) and to analyse the multidimensional case. In 2-D, first results on diffusion-inspired wavelet shrinkage with improved rotation invariance are presented in [47]. We will also consider extensions of the numerical two-pixel schemes for TV diffusion.

Acknowledgements

Our research activities in this paper are partly funded by the DFG projects We 2602/2-1 and We 2602/1-1 / So 363/9-1. This is gratefully acknowledged. Joachim Weickert also thanks Stephen Keeling (Graz, Austria) for interesting discussions on two-pixel signals.

References

[1] R. ACAR AND C. R. VOGEL, *Analysis of bounded variation penalty methods for ill–posed problems*, Inverse Problems, 10 (1994), pp. 1217–1229.

[2] L. ALVAREZ, P.-L. LIONS, AND J.-M. MOREL, *Image selective smoothing and edge detection by nonlinear diffusion. II*, SIAM Journal on Numerical Analysis, 29 (1992), pp. 845–866.

[3] F. ANDREU, C. BALLESTER, V. CASELLES, AND J. M. MAZÓN, *Minimizing total variation flow*, Differential and Integral Equations, 14 (2001), pp. 321–360.

[4] F. ANDREU, V. CASELLES, J. I. DIAZ, AND J. M. MAZÓN, *Qualitative properties of the total variation flow*, Journal of Functional Analysis, 188 (2002), pp. 516–547.

[5] G. AUBERT AND P. KORNPROBST, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of Applied Mathematical Sciences, Springer, New York, 2002.

[6] G. AUBERT AND L. VESE, *A variational method in image recovery*, SIAM Journal on Numerical Analysis, 34 (1997), pp. 1948–1979.

[7] V. AURICH AND J. WEULE, *Non-linear Gaussian filters performing edge preserving diffusion*, in Mustererkennung 1995, G. Sagerer, S. Posch, and F. Kummert, eds., Springer, Berlin, 1995, pp. 538–545.

[8] Y. BAO AND H. KRIM, *Towards bridging scale-space and multiscale frame analyses*, in Wavelets in Signal and Image Analysis, A. A. Petrosian and F. G. Meyer, eds., vol. 19 of Computational Imaging and Vision, Kluwer, Dordrecht, 2001, ch. 6.

[9] M. BELGE, M. E. KILMER, AND E. L. MILLER, *Wavelet domain image restoration with adaptive edge-preserving regularization*, IEEE Transactions on Image Processing, 9 (2000), pp. 597–608.

[10] G. BELLETTINI, V. CASELLES, AND M. NOVAGA, *The total variation flow in* R^N, Journal of Differential Equations, 184 (2002), pp. 475–525.

[11] ——, *Explicit solutions of the eigenvalue problem* $-div(Du/|Du|) = u$, tech. rep., Dept. of Mathematics, SNS di Pisa, Italy, June 2003.

[12] A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Cambridge, MA, 1987.

[13] T. Brox, M. Welk, G. Steidl, and J. Weickert, *Equivalence results for TV diffusion and TV regularisation*, in Scale-Space Methods in Computer Vision, L. D. Griffin and M. Lillholm, eds., vol. 2695 of Lecture Notes in Computer Science, Berlin, 2003, Springer, pp. 86–100.

[14] P. J. Burt and E. H. Adelson, *The Laplacian pyramid as a compact image code*, IEEE Transactions on Communications, 31 (1983), pp. 532–540.

[15] E. J. Candés and F. Guo, *New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction*, Signal Processing, 82 (2002), pp. 1519–1543.

[16] A. Chambolle, R. A. DeVore, N. Lee, and B. L. Lucier, *Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Transactions on Image Processing, 7 (1998), pp. 319–335.

[17] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik, 76 (1997), pp. 167–188.

[18] A. Chambolle and B. L. Lucier, *Interpreting translationally-invariant wavelet shrinkage as a new image smoothing scale space*, IEEE Transactions on Image Processing, 10 (2001), pp. 993–1000.

[19] T. F. Chan and H. M. Zhou, *Total variation improved wavelet thresholding in image compression*, in Proc. Seventh International Conference on Image Processing, Vancouver, Canada, Sept. 2000.

[20] A. Cohen, W. Dahmen, I. Daubechies, and R.DeVore, *Harmonic analysis in the space BV*, Tech. Rep. 195, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Germany, Sept. 2000.

[21] A. Cohen, R. DeVore, P. Petrushev, and H. Xu, *Nonlinear approximation and the space $BV(R^2)$*, American Journal of Mathematics, 121 (1999), pp. 587–628.

[22] R. R. Coifman and D. Donoho, *Translation invariant denoising*, in Wavelets in Statistics, A. Antoine and G. Oppenheim, eds., Springer, New York, 1995, pp. 125–150.

[23] R. R. Coifman and A. Sowa, *Combining the calculus of variations and wavelets for image enhancement*, Applied and Computational Harmonic Analysis, 9 (2000), pp. 1–18.

[24] ——, *New methods of controlled total variation reduction for digital functions*, SIAM Journal on Numerical Analysis, 39 (2001), pp. 480–498.

[25] Z. Cvetković and M. Vetterli, *Oversampled filter banks*, IEEE Transactions on Signal Processing, 46 (1998), pp. 1245–1255.

[26] D. L. Donoho, *De-noising by soft thresholding*, IEEE Transactions on Information Theory, 41 (1995), pp. 613–627.

[27] D. L. Donoho and I. M. Johnstone, *Minimax estimation via wavelet shrinkage*, Annals of Statistics, 26 (1998), pp. 879–921.

[28] E. R. Dougherty and J. Astola, eds., *Nonlinear Filters for Image Processing*, SPIE Press, Bellingham, 1999.

[29] S. Durand and J. Froment, *Reconstruction of wavelet coefficients using total-variation minimization*, Tech. Rep. 2001–18, Centre de Mathématiques et de Leurs Applications, ENS de Cachan, France, 2001.

[30] X. Feng and A. Prohl, *Analysis of total variation flow and its finite element approximations*, Tech. Rep. 1864, Institute of Mathematics and its Applications, University of Minnesota, Minneapolis, MN, July 2002. Submitted to Communications on Pure and Applied Mathematics.

[31] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, Kluwer, Dordrecht, 1988.

[32] F. L. Fontaine and S. Basu, *Wavelet-based solution to anisotropic diffusion equation for edge detection*, International Journal of Imaging Systems and Technology, 9 (1998), pp. 356–368.

[33] A. S. Frangakis, A. Stoschek, and R. Hegerl, *Wavelet transform filtering and nonlinear anisotropic diffusion assessed for signal reconstruction performance on multidimensional biomedical data*, IEEE Transactions on Biomedical Engineering, 48 (2001), pp. 213–222.

[34] J. Fröhlich and J. Weickert, *Image processing using a wavelet algorithm for nonlinear diffusion*, Tech. Rep. 104, Laboratory of Technomathematics, University of Kaiserslautern, Germany, Mar. 1994.

[35] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6 (1984), pp. 721–741.

[36] W. Hinterberger, M. Hintermüller, K. Kunisch, M. von Oehsen, and O. Scherzer, *Tube methods for BV regularization*, Tech. Rep. 6, Department of Computer Science, University of Innsbruck, Austria, Dec. 2002.

[37] R. Klette and P. Zamperoni, *Handbook of Image Processing Operators*, Wiley, New York, 1996.

[38] J. S. Lee, *Digital image smoothing and the sigma filter*, Computer Vision, Graphics and Image Processing, 24 (1983), pp. 255–269.

[39] S. Z. Li, *Close-form solution and parameter selection for convex minimization-based edge-preserving smoothing*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20 (1998), pp. 916–932.

[40] S. Z. Li and A. K. Jain, *Markov Random Field Modeling in Image Analysis*, Springer, Tokyo, second ed., 2001.

[41] T. Lu, P. Neittaanmäki, and X.-C. Tai, *A parallel splitting up method and its application to Navier–Stokes equations*, Applied Mathematics Letters, 4 (1991), pp. 25–29.

[42] F. Malgouyres, *Combining total variation and wavelet packet approaches for image deblurring*, in Proc. First IEEE Workshop on Variational and Level Set Methods in Computer Vision, Vancouver, Canada, July 2001, IEEE Computer Society Press, pp. 57–64.

[43] ——, *Mathematical analysis of a model which combines total variation and wavelet for image restoration*, Inverse Problems, 2 (2002), pp. 1–10.

[44] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, second ed., 1999.

[45] E. Mammen and S. van de Geer, *Locally adaptive regression splines*, Annals of Statistics, 25 (1997), pp. 387–413.

[46] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, vol. 22 of University Lecture Series, AMS, Providence, 2001.

[47] P. Mrázek and J. Weickert, *Rotationally invariant wavelet shrinkage*, in Pattern Recognition, B. Michaelis, ed., Lecture Notes in Computer Science, Berlin, 2003, Springer. In press.

[48] P. Mrázek, J. Weickert, G. Steidl, and M. Welk, *On iterations and scales of nonlinear filters*, in Proc. Eighth Computer Vision Winter Workshop, O. Drbohlav, ed., Valtice, Czech Republic, Feb. 2003, Czech Pattern Recognition Society, pp. 61–66.

[49] N. Nordström, *Biased anisotropic diffusion – a unified regularization and diffusion approach to edge detection*, Image and Vision Computing, 8 (1990), pp. 318–327.

[50] P. Perona and J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (1990), pp. 629–639.

[51] I. Pollak, A. S. Willsky, and H. Krim, *Image segmentation and edge enhancement with stabilized inverse diffusion equations*, IEEE Transactions on Image Processing, 9 (2000), pp. 256–266.

[52] J. L. Richardson, R. C. Ferrell, and L. N. Long, *Unconditionally stable explicit algorithms for nonlinear fluid dynamics problems*, Journal of Computational Physics, 104 (1993), pp. 69–74.

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

[54] P. Saint-Marc, J. S. Chen, and G. Medioni, *Adaptive smoothing: a general tool for early vision*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13 (1991), pp. 514–529.

[55] G. Sapiro, *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, Cambridge, UK, 2001.

[56] O. Scherzer and J. Weickert, *Relations between regularization and diffusion filtering*, Journal of Mathematical Imaging and Vision, 12 (2000), pp. 43–63.

[57] C. Schnörr, *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, Journal of Mathematical Imaging and Vision, 4 (1994), pp. 189–198.

[58] J. Shen, *A note on wavelets and diffusion*, Journal of Computational Analysis and Applications, 5 (2003), pp. 147–159.

[59] J. Shen and G. Strang, *On wavelet fundamental solutions to the heat equation – heatlets*, Journal of Differential Equations, 161 (2000), pp. 403–421.

[60] G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Clarendon Press, Oxford, third ed., 1985.

[61] S. M. Smith and J. M. Brady, *SUSAN: A new approach to low-level image processing*, International Journal of Computer Vision, 23 (1997), pp. 45–78.

[62] G. Steidl and J. Weickert, *Relations between soft wavelet shrinkage and total variation denoising*, in Pattern Recognition, L. Van Gool, ed., vol. 2449 of Lecture Notes in Computer Science, Springer, Berlin, 2002, pp. 198–205.

[63] R. L. Stevenson, B. E. Schmitz, and E. J. Delp, *Discontinuity preserving regularization of inverse visual problems*, IEEE Transactions on Systems, Man and Cybernetics, 24 (1994), pp. 455–469.

[64] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley–Cambridge Press, Wellesley, 1996.

[65] D. M. Strong, *Adaptive Total Variation Minimizing Image Restoration*, PhD thesis, Department of Mathematics, University of California, Los Angeles, CA, 1997.

[66] L. Vese, *A study in the BV space of a denoising–deblurring variational problem*, Applied Mathematics and Optimization, 44 (2001), pp. 131–161.

[67] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.

[68] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, *Efficient and reliable schemes for nonlinear diffusion filtering*, IEEE Transactions on Image Processing, 7 (1998), pp. 398–410.

[69] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, vol. 27 of Applications of Mathematics, Springer, Berlin, 1995.

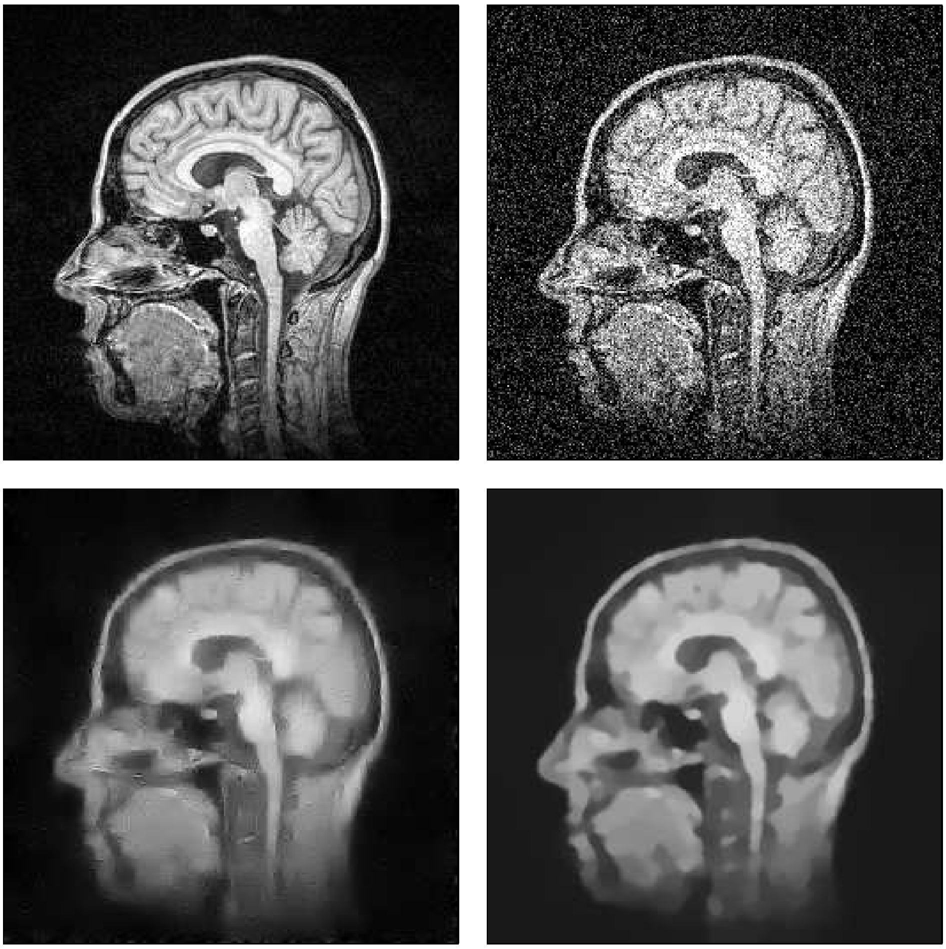

Figure 1: (a) TOP LEFT: Original MR image. (b) TOP RIGHT: MR image degraded with Gaussian noise with standard deviation 50. (c) BOTTOM LEFT: Wavelet denoising of (b) using translation invariant soft shrinkage with Haar wavelets. (d) BOTTOM RIGHT: Total variation diffusion of (b).

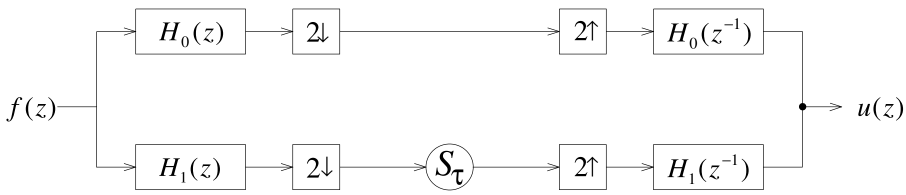

Figure 2: Two-channel filter bank with $H_0(z) = \frac{1+z}{\sqrt{2}}$ and $H_1(z) = \frac{1-z}{\sqrt{2}}$.

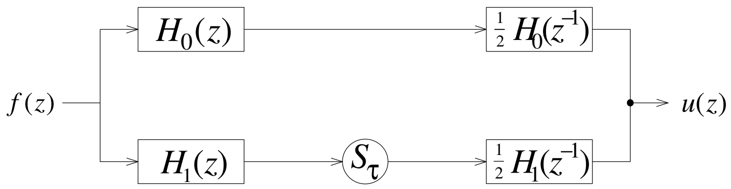

Figure 3: Nonsubsampled two-channel filter bank with $H_0(z) = \frac{1+z}{\sqrt{2}}$ and $H_1(z) = \frac{1-z}{\sqrt{2}}$.

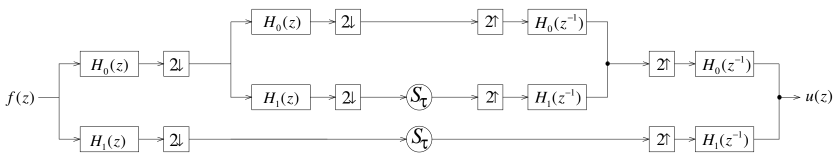

Figure 4: Two scales of Haar wavelet shrinkage with $H_0(z) = \frac{1+z}{\sqrt{2}}$ and $H_1(z) = \frac{1-z}{\sqrt{2}}$.

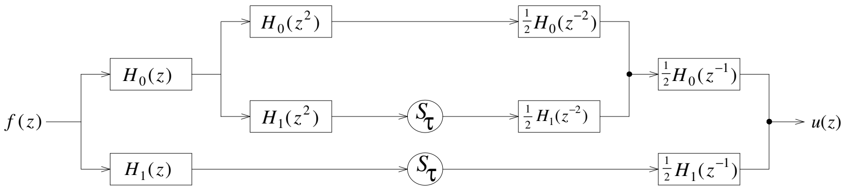

Figure 5: Two scales of translation-invariant Haar wavelet shrinkage with $H_0(z) = \frac{1+z}{\sqrt{2}}$ and $H_1(z) = \frac{1-z}{\sqrt{2}}$.

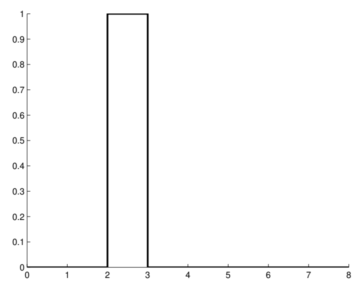

Figure 6: Test signal with $N = 8$ pixels.

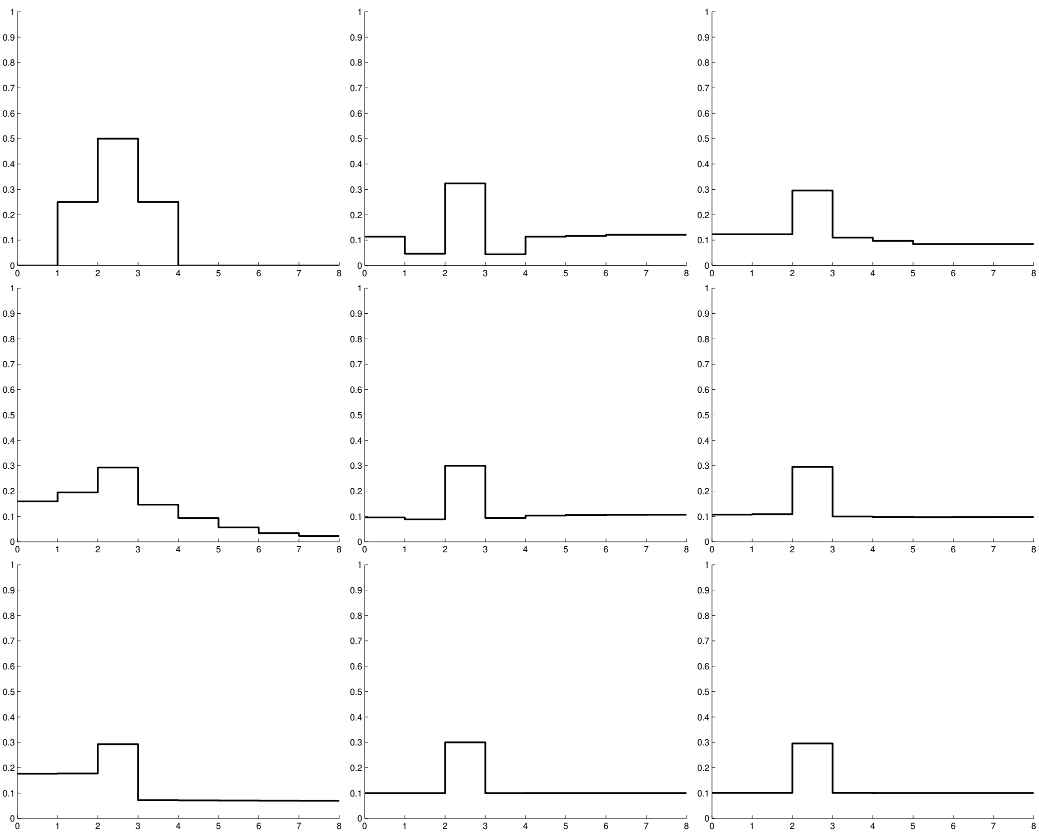

Figure 7: $K = 1, 20, 1000$ iterations (top to bottom) of translation-invariant soft Haar wavelet shrinkage with thresholds τ/K applied to the signal in Fig. 6. Left column: single-scale wavelet shrinkage with $\tau = 1$. Center column: multi-scale wavelet shrinkage ($m = 4$) with uniform threshold $\tau = 0.48$ on all scales. Right column: multi-scale wavelet shrinkage ($m = 4$) with $\tau = 0.585$ and scale-adapted thresholds according to (46).

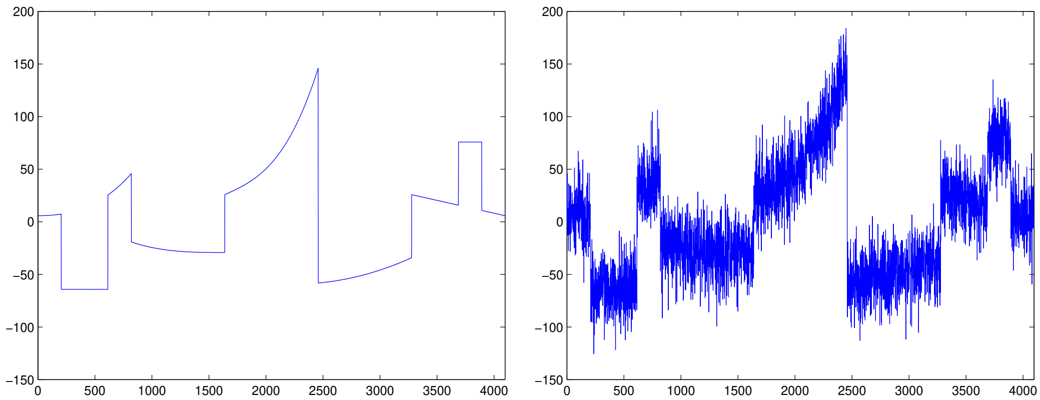

Figure 8: Piecewise polynomial signal. Left: original. Right: with additive Gaussian white noise (SNR = 8 dB) as input for the filtering procedures.

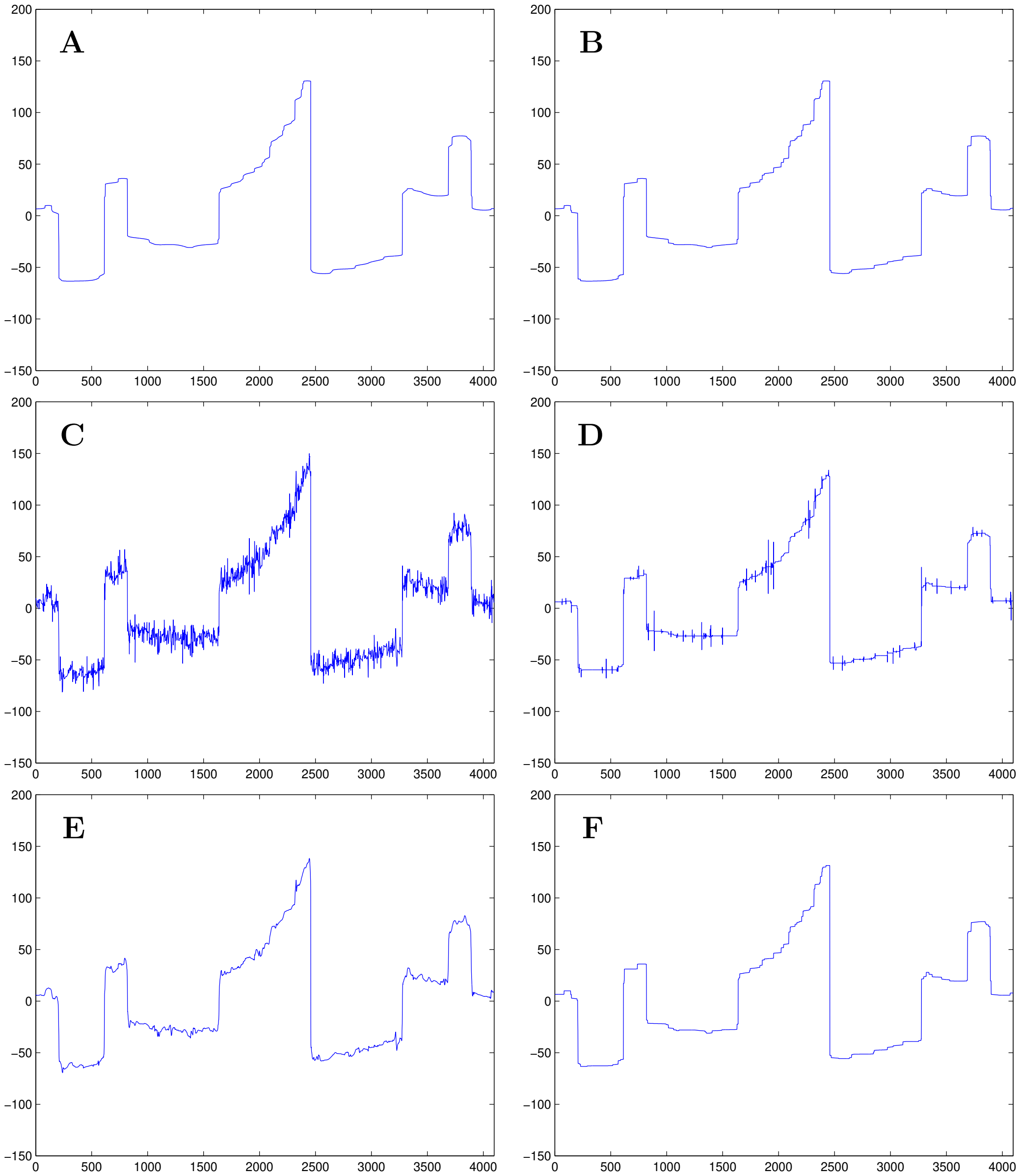

Figure 9: Optimal filtering results of several variants of procedures based on TV or wavelet filtering when run on the noisy data of Fig. 8.
A. Iterated classical scheme for the regularized TV flow (24).
B. Iterated single-level shrinkage (equivalent to the scheme (22) for TV flow).
C. Multiple levels with a single threshold, single step (i.e. noniterated).
D. Iterated multiple level with a single threshold at each of the levels.
E. Multiple levels with thresholds scaled according to (46), single step.
F. Iterated multiple level with scaled thresholds.
See text and Table 1 for explanation and numerical evaluation of the results.