Infimal convolution of oscillation total generalized variation for the recovery of images with structured texture

Yiming Gao[*] Kristian Bredies[†]

April 17, 2018

Abstract: We propose a new type of regularization functional for images called oscillation total generalized variation (TGV) which can represent structured textures with oscillatory character in a specified direction and scale. The infimal convolution of oscillation TGV with respect to several directions and scales is then used to model images with structured oscillatory texture. Such functionals constitute a regularizer with good texture preservation properties and can flexibly be incorporated into many imaging problems. We give a detailed theoretical analysis of the infimal-convolution-type model with oscillation TGV in function spaces. Furthermore, we consider appropriate discretizations of these functionals and introduce a first-order primal-dual algorithm for solving general variational imaging problems associated with this regularizer. Finally, numerical experiments are presented which show that our proposed models can recover textures well and are competitive in comparison to existing state-of-the-art methods.

Mathematics subject classification: 94A08, 68U10, 26A45, 90C90.

Key words: Oscillation total generalized variation, infimal-convolution-type regularization, texture modeling, cartoon/texture decomposition, image denoising, image inpainting, undersampled magnetic resonance imaging.

1   Introduction

After the introduction of total variation (TV) to image processing [46], variational models were developed for a wide range of applications in imaging during the last decades. In that context, variational models that base on image decomposition have also become increasingly popular. The concept of these models is to decompose an image into two or more components which inherit distinguishing characteristics. The classical TV (ROF) model for image denoising can, for instance, be regarded as a decomposition model:

$$\min_{\substack{u\in\mathrm{BV}(\Omega),\ v\in L^2(\Omega)\\u+v=f}} \frac{\lambda}{2}\|v\|^2 + J(u), \tag{1}$$

where f is the observed noisy image, $J(u) = \int_\Omega |Du|$ is the total variation, i.e., the Radon norm of the distributional derivative, and $\|\cdot\|$ represents the L^2 norm, both associated with a bounded domain $\Omega \subset \mathbb{R}^d$. In such an approach, we decompose the image to a noise-free image

[*]Department of Mathematics, Nanjing University of Science and Technology, 210094 Nanjing, Jiangsu, China. Email: gao_yiming@hotmail.com.

[†]Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria. Email: kristian.bredies@uni-graz.at. The Institute for Mathematics and Scientific Computing is a member of NAWI Graz (http://www.nawigraz.at/).

component u belonging to BV(Ω) and a noise or small-scale texture component in $L^2(\Omega)$. In [42], Y. Meyer pointed out some limitations of the above model and introduced a new space $G(\Omega) = W^{-1,\infty}(\Omega)$, which is larger than $L^2(\Omega)$, to model oscillating patterns that are typical for structured texture:

$$\min_{\substack{u\in\text{BV}(\Omega),\ v\in G(\Omega) \\ u+v=f}} \lambda\|v\|_G + J(u), \tag{2}$$

where $G(\Omega)$ denotes the space $G(\Omega) = \{v = \operatorname{div} g : g = (g_1,\ldots,g_d) \in L^\infty(\Omega)^d\}$ with the norm $\|v\|_G = \inf \{\|g\|_\infty : v = \operatorname{div} g, g = (g_1,\ldots,g_d) \in L^\infty(\Omega)^d\}$ and again, $J(u) = \int_\Omega |Du|$. In this model, u represents a piecewise constant function, consisting of homogeneous regions with sharp boundaries and is usually called cartoon component. The other part v in $G(\Omega)$ contains oscillating patterns, like textures and noise. This model can be considered as the original cartoon-texture decomposition model. However, the G-space may be difficult to handle in implementations. Vese and Osher [52] approximated the G-space by a Sobolev space of negative differentiability order $W^{-1,p}(\Omega)$, which can, in practice, more easily be implemented by partial differential equations. Later, Aujol et al. [4] made a modification to the G-norm term in (2) which is replaced by constraining v to the set $G_\mu(\Omega) = \{v \in G(\Omega) : \|v\|_G \le \mu\}$. As a consequence, the problem can be solved alternatingly by Chambolle's projection method [23] with respect to the two variables u and v.

Certainly, the aim of above models is to find an appropriate norm to describe textures. However, in general, the above norms can represent all kinds of oscillatory parts. As a consequence, since noise can be regarded as small-scale oscillating texture, these models do not deal well with noise. In order to tackle this disadvantage, in [47], Schaeffer and Osher incorporated robust principal component analysis (PCA) [22] into cartoon-texture decomposition models based on the patch method. For the texture part, the authors suggest to decompose the image into small, non-overlapping patches, written as vectors. The collection of the patch vectors is assumed to be (highly) linearly dependent and thus to have low rank. This suggests to minimize the nuclear norm of the patch-vector matrix. However, the choice of the block-size of the patches for this model is challenging and also, the numerical optimization procedures for the associated functionals have to rely on computationally expensive matrix-decomposition methods such as singular-value thresholding, for instance. Alternatively, rather than considering the cartoon and texture components separately, another efficient texture preservation approach is using non-local methods which is also based on patches. For more details, we refer readers to [15, 16, 30, 53]. Furthermore, in order to enhance and reconstruct line structures well, Holler and Kunisch [33] proposed an infimal convolution of TGV-type (ICTGV) functionals which can capture some directional line textures. Disadvantages are, of course, the bias of this regularization approach towards certain directions and the fact that we cannot expect to recover texture that is not composed of different line structures. In [38], Kongskov and Dong proposed a new directional total generalized variation (DTGV) functional to capture the directional information of an image, which can be considered as a special case of ICTGV.

Another approach to decompose an image is to compute sparse representations based on dictionaries. The basic idea is to find two suitable dictionaries, one adapted to represent textures, and the other to represent the cartoon part. In the papers [27, 49], Elad et al. combined the basis pursuit denoising (BPDN) algorithm and the TV-regularization scheme to the following decomposition model:

$$\min_{\alpha_t,\alpha_n} \|\alpha_t\|_1 + \|\alpha_n\|_1 + \frac{\lambda}{2}\|f - T_t\alpha_t - T_n\alpha_n\|^2 + \gamma J(T_n\alpha_n), \tag{3}$$

where T_t, T_n are dictionary synthesis operators for the texture and cartoon parts α_t, α_n, respectively, and J is again the total variation. Penalizing with TV forces the cartoon content $T_n\alpha_n$ to have a sparser gradient and to be closer to a piecewise constant image. Here, the choice of the dictionary synthesis operators T_t, T_n is the main challenge for this approach. The best

choice for one image depends on experience and may be not suitable for another image with different structures. The authors recommended that for the texture content, one may use local discrete cosine transform (DCT) or the Gabor transform, and for the cartoon content one may use curvelet, ridgelets, contourlets and so on. This approach is reported to perform well for the restoration of texture images [27, 49]. Inspired by this, Cai, Osher and Shen [21] constructed a cartoon-texture model with tight framelets [20] and local DCT which can easily be realized algorithmically by split Bregman iterations.

Recently, filtering approaches have been widely used in image decomposition. Gilboa et al. proposed a generalization of the structure-texture decomposition method based on TV scale space which is also known as TV flow in [28, 29, 34]. The aim is to construct a TV spectral framework to analyze dominant features in the transform domain, filter in that domain, and then perform an inverse transform back to the spatial domain to obtain the filtered response. In [19], Burger et al. extended this notion to general one-homogeneous functionals, not only TV regularization. Furthermore, Buades et al. proposed a non-linear filter that decomposes the image into a geometric and an oscillatory part in [17, 18], in which a non-linear low-high frequency decomposition is computed by using isotropic filters. However, the cartoon parts produced by these filters are not well-structured and textures also contain other oscillatory components. Furthermore, these methods are not variational models and can not easily be transferred to other image processing problems such as solving inverse problems, for instance.

In this paper, we propose a new regularization which can represent structured texture that consists of different directions and frequencies. The regularization is called oscillation total generalized variation (oscillation TGV) and defined as follows:

$$\text{TGV}^{\text{osci}}_{\alpha,\beta,\mathbf{c}}(u) = \min_{w\in\text{BD}(\Omega)} \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}}, \quad \text{for } u \in L^1(\Omega), \tag{4}$$

where $\alpha, \beta > 0$, $\text{BD}(\Omega)$ denotes the space of vector fields of bounded deformation, the operator ∇ denotes the weak derivative, the operator $\mathcal{E}$ is the weak symmetrized derivative $\mathcal{E}w = \frac{1}{2}(\nabla w + \nabla w^T)$ and $\|\cdot\|_{\mathcal{M}}$ is the Radon norm. Moreover, $\mathbf{c} \in \mathbb{R}^{d\times d}$ and $\mathbf{c}u$ corresponds to the matrix field given by $(\mathbf{c}u)_{ij} = c_{ij}u$ for $i,j = 1,\ldots,d$. In (4), we set $\mathbf{c} = \omega \otimes \omega$ for some $\omega \in \mathbb{R}^d$, i.e., $c_{ij} = \omega_i\omega_j$. For a precise description of these notions, we ask for the reader's patience until Section 2. This regularization functional reduces to the original second-order TGV functional when $\mathbf{c} = 0$, i.e.,

$$\text{TGV}^2_{\alpha,\beta}(u) = \min_{w\in\text{BD}(\Omega)} \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}}, \tag{5}$$

which has been widely-used in imaging applications, such as, for instance, denoising, (dynamic) MRI and QSM reconstruction, diffusion tensor image restoration, JPEG and MPEG decompression, zooming, optical flow estimation [9, 10, 8, 11, 12, 14, 33, 35, 37, 39, 44, 48, 51], and many more. Furthermore, we consider m-fold infimal convolution of oscillation TGV:

$$\text{ICTGV}^{\text{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}(u) = (\text{TGV}^{\text{osci}}_{\alpha_1,\beta_1,\mathbf{c}_1} \square \ldots \square \text{TGV}^{\text{osci}}_{\alpha_m,\beta_m,\mathbf{c}_m})(u)$$

$$= \inf_{u_1+\ldots+u_m=u} \sum_{i=1}^{m} \text{TGV}^{\text{osci}}_{\alpha_i,\beta_i,\mathbf{c}_i}(u_i) \tag{6}$$

where $\vec{\alpha} = (\alpha_1\ldots,\alpha_m)$, $\vec{\beta} = (\beta_1,\ldots,\beta_m)$ and $\vec{\mathbf{c}} = (\mathbf{c}_1,\ldots,\mathbf{c}_m)$. In applications, $\mathbf{c}_1 = 0$ such that u_1 represents the cartoon part interpreted as piecewise smooth function, and $u_2,\ldots,u_m$ are the texture components with the directions depending on the values of $\mathbf{c}_2,\ldots,\mathbf{c}_m$. We give a detailed theoretical analysis in an infinite-dimensional function space setting and apply regularization with $\text{ICTGV}^{\text{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}$ to several image processing problems.

The paper is organized as follows. In Section 2, we start with the introduction of the oscillation TGV functional which, on the one hand, measures smoothness but possesses, on the other hand, a kernel consisting of structured oscillations. Then, we give a detailed analysis of it and

3

show some main properties. In Section 3, we focus on $\text{ICTGV}^{\text{osci}}_{\vec{\alpha},\vec{\beta},\vec{c}}$. First, we establish a theoretic analysis for the general functionals of infimal-convolution type and apply it to infimal convolutions of oscillation TGV. As a result, we obtain well-posedness for Tikhonov regularization of linear inverse problems with $\text{ICTGV}^{\text{osci}}$-penalty. The discretization and a first-order primal-dual optimization algorithm for solving Tikhonov-regularized linear inverse problems with infimal-convolution of oscillation TGV regularization are presented in Section 4. In particular, we show how the differential operators in TGV^{osci} have to be discretized such that the discrete kernel coincides with discrete structured oscillations. In Section 5, we illustrate our approach for several imaging problems such as denoising, inpainting and MRI reconstruction, and demonstrate the efficiency of our new regularization compared to a variety of methods, including state-of-the-art methods. Finally, conclusions are drawn in Section 6.

2   Oscillation TGV and basic properties

This section is devoted to the introduction of oscillation total generalized variation (TGV^{osci}) and the derivation of some immediate properties. Throughout this paper, we will focus on second-order TGV and its generalizations although further generalization to higher orders might also be possible.

Let us first introduce the required spaces and notions. For simplicity, we assume for the rest of the paper that $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain where $d \geq 1$ is the image dimension. If μ is an X-valued distribution on Ω where $X \in \{\mathbb{R}, \mathbb{R}^d, S^{d\times d}\}$ and $S^{d\times d}$ denotes the set of symmetric $d \times d$ matrices equipped with the Frobenius matrix inner product and norm, then the Radon norm of μ is given as

$$\|\mu\|_{\mathcal{M}} = \sup \{\langle \mu, \phi \rangle : \phi \in C_c^\infty(\Omega, X), \ \|\phi\|_\infty \leq 1\}$$

where $C_c^\infty(\Omega, X)$ denotes the space of X-valued test functions on Ω, i.e., the set of arbitrarily smooth $\phi : \Omega \to X$ with compact support, and $\langle \cdot, \cdot \rangle$ denotes the duality pairing of distributions and test functions. Here, the supremum is ∞ whenever the set is unbounded from above. In case $\|\mu\|_{\mathcal{M}} < \infty$, the distribution μ can continuously be extended, by density, to a Radon measure in $\mathcal{M}(\Omega, X) = C_0(\Omega, X)^*$ in which case the Radon norm is the dual norm and we can identify μ with its corresponding measure. Note that in particular, the density of $\{\phi \in C_c^\infty(\Omega, X) : \|\phi\|_\infty \leq 1\}$ in the closed unit ball of $C_0(\Omega, X)$ can easily be established using smooth cut-off functions and the fact that for non-negative mollifiers ρ we have $\|\rho * \phi\|_\infty \leq \|\phi\|_\infty$.

A $\mathbb{R}$-valued distribution u always admits a $\mathbb{R}^d$-valued distribution as weak derivative which we denote by ∇u and which is given by $\langle \nabla u, \phi \rangle = \langle u, -\operatorname{div} \phi \rangle$ for each $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$. Likewise, for an $\mathbb{R}^d$-valued distribution w, the weak symmetrized derivative is a $S^{d\times d}$-valued distribution which is denoted by $\mathcal{E}w$ and defined by $\langle \mathcal{E}w, \phi \rangle = \langle w, -\operatorname{div} \phi \rangle$ for each $\phi \in C_c^\infty(\Omega, S^{d\times d})$. With the usual (tacit) identification of functions in $u \in L^1_{\text{loc}}(\Omega, X)$ with X-valued distributions via $T_u(\phi) = \int_\Omega u \cdot \phi \, dx$ for $\phi \in C_c^\infty(\Omega, X)$, the spaces of bounded variation $\text{BV}(\Omega)$ and bounded deformation $\text{BD}(\Omega)$ are given by

$$\text{BV}(\Omega) = \{u \in L^1(\Omega) : \|\nabla u\|_{\mathcal{M}} < \infty\}, \qquad \|u\|_{\text{BV}} = \|u\|_1 + \|\nabla u\|_{\mathcal{M}},$$
$$\text{BD}(\Omega) = \{w \in L^1(\Omega, \mathbb{R}^d) : \|\mathcal{E}w\|_{\mathcal{M}} < \infty\}, \quad \|w\|_{\text{BD}} = \|u\|_1 + \|\mathcal{E}w\|_{\mathcal{M}}.$$

These spaces are well-studied Banach spaces and we refer to [1, 2, 31, 50] for an overview and more details.

2.1   Definition of oscillation TGV

We start with a motivation for the definition as in (4). Observe that functions u representing sinusoidal oscillations along a fixed direction are given as follows:

$$u(x) = C_1 \cos(\omega \cdot x) + C_2 \sin(\omega \cdot x) \quad \text{for} \quad x \in \Omega. \tag{7}$$
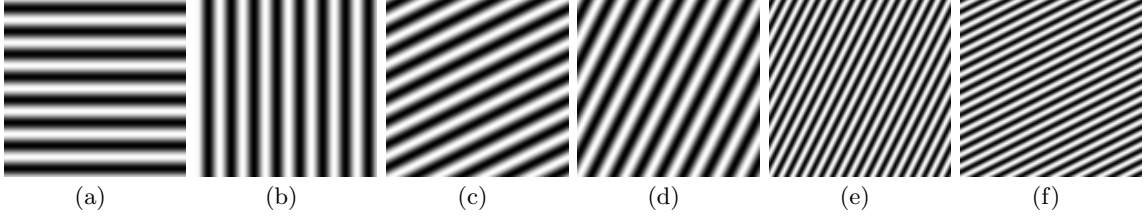
Figure 1: Visualization of sinusoidal oscillations on $\Omega \subset \mathbb{R}^2$ according to (7). (a) $\omega_1 = 0, \omega_2 = 0.5$; (b) $\omega_1 = 0.5, \omega_2 = 0$; (c) $\omega_1 = 0.25, \omega_2 = 0.5$; (d) $\omega_1 = 0.5, \omega_2 = 0.25$; (e) $\omega_1 = 1, \omega_2 = 0.5$; (f) $\omega_1 = 0.5, \omega_2 = 1$.

Here, $C_1, C_2 \in \mathbb{R}$ are constants, $\omega \in \mathbb{R}^d$, $\omega \neq 0$ is a direction vector and $\cdot$ represents the standard inner product of vectors in $\mathbb{R}^d$. Figure 1 shows a visualization of these oscillatory functions (7) in $\Omega \subset \mathbb{R}^2$ with different values of ω_1 and ω_2. Our goal is to find a differential equation that is, in some sense, compatible with the second-order TGV functional according to [12] and such that the functions u according (7) are exactly the solutions for all $C_1, C_2 \in \mathbb{R}$. For this purpose, consider the equation

$$\mathcal{E}\nabla u + \mathbf{c}u = 0, \tag{8}$$

where $\mathbf{c} = \omega \otimes \omega$ is the tensor (or outer) product, i.e., $c_{ij} = \omega_i \omega_j$, $i, j = 1, \ldots, d$. Indeed, this equation already possesses the desired properties.

Lemma 1. *Let $\mathbf{c} = \omega \otimes \omega$ with $\omega \in \mathbb{R}^d$, $\omega \neq 0$. Then, a function $u \in C^2(\Omega)$ solves (8) if and only if u has a representation (7) for some $C_1, C_2 \in \mathbb{R}$.*

Proof. Note that it suffices to prove the statement locally, i.e., without loss of generality, Ω is an open hypercube in $\mathbb{R}^d$ of the form $\Omega =]a_1, b_1[\times \cdots \times]a_d, b_d[$.

Clearly, $(\mathcal{E}\nabla)_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}$ and $\frac{\partial^2}{\partial x_i \partial x_j} \sin(\omega \cdot x) = -\omega_i \omega_j \sin(\omega \cdot x)$ as well as $\frac{\partial^2}{\partial x_i \partial x_j} \cos(\omega \cdot x) = -\omega_i \omega_j \cos(\omega \cdot x)$. Thus, if $u \in C^2(\Omega)$ has the representation (7), then (8) follows.

For the converse direction, we proceed inductively with respect to the dimension d. Starting with $d = 1$, equation (8) reads as $\frac{\partial^2 u}{\partial x_1^2} u + \omega_1^2 u = 0$ where $\omega_1 \neq 0$ which is an ordinary differential equation with solutions of the form (7). Next, fix $d \geq 1$ and suppose that (8) implies (7) for some $C_1, C_2 \in \mathbb{R}$ whenever $u \in C^2(\Omega)$ for $\Omega \subset \mathbb{R}^d$ an open hypercube. Let $\omega \in \mathbb{R}^{d+1}$ such that $\omega \neq 0$. After a possible interchange of axes, we can assume, without loss of generality, that $(\omega_1, \ldots, \omega_d) \neq 0$. Let $\Omega \subset \mathbb{R}^{d+1}$ be an open hypercube and $u \in C^2(\Omega)$ satisfy (8) (in dimension $d + 1$). Then, in particular, $\frac{\partial^2 u}{\partial x_i \partial x_j} + \omega_i \omega_j u = 0$ for $i, j = 1, \ldots, d$. Denoting by x' the projection of a vector $x \in \mathbb{R}^{d+1}$ onto the first d components, i.e., $x' = (x_1, \ldots, x_d)$, the induction hypothesis gives

$$u(x) = C_1(x_{d+1}) \sin(\omega' \cdot x') + C_2(x_{d+1}) \cos(\omega' \cdot x') \quad \text{for} \quad x \in \Omega.$$

As $x \mapsto \omega' \cdot x'$ is non-constant, we are able to choose $x, \bar{x} \in \Omega$ such that the matrix

$$\begin{bmatrix} \sin(\omega' \cdot x') & \cos(\omega' \cdot x') \\ \sin(\omega' \cdot \bar{x}') & \cos(\omega' \cdot \bar{x}') \end{bmatrix}$$

has full rank. Using that $\frac{\partial^2 u}{\partial x_{d+1}^2} + \omega_{d+1}^2 u = 0$ in neighborhoods of x and $\bar{x}$ consequently yields that there are constants $C_{1,1}, C_{1,2}, C_{2,1}, C_{2,2}$ such that

$$C_k(x_{d+1}) = \begin{cases} C_{k,1} x_{d+1} + C_{k,2} & \text{if } \omega_{d+1} = 0, \\ C_{k,1} \sin(\omega_{d+1} x_{d+1}) + C_{k,2} \cos(\omega_{d+1} x_{d+1}) & \text{otherwise,} \end{cases}$$

for $k = 1, 2$. If $\omega_{d+1} = 0$, then $\frac{\partial^2 u}{\partial x_i \partial x_{d+1}} + \omega_i \omega_{d+1} u = 0$ for $i \in \{1, \ldots, d\}$ such that $\omega_i \neq 0$ gives that $C_{1,1} \cos(\omega' \cdot x') - C_{2,1} \sin(\omega' \cdot x') = 0$ for $x \in \Omega$. Consequently, $C_{1,1} = C_{2,1} = 0$ and the induction step is complete.

5

In the case $\omega_{d+1} \neq 0$, $\frac{\partial^2 u}{\partial x_i \partial x_{d+1}} + \omega_i \omega_{d+1} u = 0$ for $i \in \{1, \ldots, d\}$ such that $\omega_i \neq 0$ implies

$$u(x) = \big(C_{1,2} \sin(\omega_{d+1} x_{d+1}) - C_{1,1} \cos(\omega_{d+1} x_{d+1})\big) \cos(\omega' \cdot x')$$
$$+ \big(C_{2,1} \cos(\omega_{d+1} x_{d+1}) - C_{2,2} \sin(\omega_{d+1} x_{d+1})\big) \sin(\omega' \cdot x')$$

for $x \in \Omega$. Comparing coefficients yields $C_{2,2} = -C_{1,1}$ as well as $C_{2,1} = C_{1,2}$. Plugging this into the representation of u gives, by virtue of the angle sum identities for sine and cosine,

$$u(x) = C_{2,1} \sin(\omega' \cdot x' + \omega_{d+1} x_{d+1}) + C_{2,2} \cos(\omega' \cdot x' + \omega_{d+1} x_{d+1}) = C_{2,1} \sin(\omega \cdot x) + C_{2,2} \cos(\omega \cdot x)$$

for $x \in \Omega$, proving the induction step for $\omega_{d+1} \neq 0$. $\qquad\square$

Based on this observation, we aim at designing a functional that is suitable to measure piecewise oscillatory functions u with direction $\omega \neq 0$, i.e.,

$$u = \sum_{i=1}^{n} \chi_{\Omega_i}(x) u_i(x), \quad \text{with} \quad \chi_{\Omega_i}(x) = \begin{cases} 1, & \text{if } x \in \Omega_i, \\ 0, & \text{otherwise,} \end{cases}$$

where $\Omega_1 \cup \ldots \cup \Omega_n = \Omega$, $\Omega_i \cap \Omega_j = \emptyset$, $\forall i \neq j$ and $u_i(x) = C_{1,i} \cos(\omega \cdot x) + C_{2,i} \sin(\omega \cdot x)$, for some $C_{1,i}, C_{2,i} \in \mathbb{R}$, $i = 1, \ldots, n$.

Let us now motivate the definition of oscillation TGV. It is easily seen that the kernel of second-order TGV according to (5) can be obtained by solving $\nabla u - w = 0$ and $\mathcal{E}w = 0$ for $u \in \mathrm{BV}(\Omega), w \in \mathrm{BD}(\Omega)$. This leads to all u for which $\mathcal{E}\nabla u = 0$ holds in the distributional sense which can be shown to be exactly the affine functions. Based on this and the fact that oscillatory functions are solving (8), we see in turn that this class can be represented by $\nabla u - w = 0$ and $\mathcal{E}w + \mathbf{c}u = 0$ for $u \in \mathrm{BV}(\Omega), w \in \mathrm{BD}(\Omega)$. This leads to the following generalization of second-order TGV we call oscillation TGV.

Definition 1. *Let $\alpha, \beta > 0$ and $\mathbf{c} = \omega \otimes \omega$ where $\omega \in \mathbb{R}^d$. For $u \in L^1(\Omega)$, the oscillation TGV functional is defined as follows:*

$$\mathrm{TGV}_{\alpha,\beta,\mathbf{c}}^{\mathrm{osci}}(u) = \min_{w \in \mathrm{BD}(\Omega)} \alpha \|\nabla u - w\|_{\mathcal{M}} + \beta \|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}} \tag{9}$$

if $u \in \mathrm{BV}(\Omega)$ and $\mathrm{TGV}_{\alpha,\beta,\mathbf{c}}^{\mathrm{osci}}(u) = \infty$ if $u \in L^1(\Omega) \setminus \mathrm{BV}(\Omega)$.

The usage of the minimum will be justified later. Note that for $\omega = 0$, the definition reduces to the well-known second-order TGV with weights α, β, i.e., we have $\mathrm{TGV}_{\alpha,\beta,\mathbf{c}}^{\mathrm{osci}}(u) = \mathrm{TGV}_{\alpha,\beta}^2(u)$ if $\omega = 0$. One can easily see that oscillation TGV is proper and convex; we focus on showing its lower semi-continuity in L^p-spaces along with some basic properties in the following.

2.2 Basic properties

For the purpose of establishing lower semi-continuity, we need some tools and notions from convex analysis, see, for instance, [45, 5], for an overview. For a functional $F : X \to \,]{-}\infty, \infty]$, the effective domain is defined as $\mathrm{dom}\, F = \{x \in X : F(x) < \infty\}$, and F is called proper if $\mathrm{dom}\, F \neq \emptyset$. The indicator functional is defined as

$$\mathcal{I}_K(x) = \begin{cases} 0 & \text{if } x \in K, \\ +\infty & \text{otherwise.} \end{cases}$$

As usual, we also denote by $\langle \cdot, \cdot \rangle$ the duality product of X^* and X between the Banach spaces X^* and X. Moreover, recall that the convex conjugate functional for a proper $F : X \to \,]{-}\infty, \infty]$ is given by $F^* : X^* \to \,]{-}\infty, \infty]$ with $F^*(x^*) = \sup_{x \in X} \langle x^*, x \rangle - F(x)$. With these prerequisites, we give a predual formulation of oscillation TGV which will then imply lower semi-continuity and in particular justify the minimum in the definition (9).

Proposition 1. *Fix $\alpha, \beta > 0$ and $\mathbf{c} = \omega \otimes \omega$, $\omega \in \mathbb{R}^d$. For each $u \in L^1(\Omega)$, we have*

$$\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) = \sup \left\{ \int_\Omega u(\mathrm{div}^2 \phi + \mathbf{c} \cdot \phi) \, \mathrm{d}x : \phi \in C_c^\infty(\Omega, S^{d\times d}), \|\phi\|_\infty \leq \beta, \|\mathrm{div}\, \phi\|_\infty \leq \alpha \right\},$$
(10)

where $\mathbf{c} \cdot \phi$ denotes the pointwise Frobenius inner product of $\mathbf{c}$ and the symmetric matrix field ϕ. Moreover, the minimum in the definition of $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u)$ according to (9) is obtained.

Proof. Firstly observing that $C_c^\infty(\Omega, S^{d\times d})$ is dense in the Banach space $C_0^2(\Omega, S^{d\times d})$ with respect to the C^2 norm given by $\|\phi\|_{C^2} = \max(\|\phi\|_\infty, \|\nabla \phi\|_\infty, \|\nabla^2 \phi\|_\infty)$, we can replace $C_c^\infty(\Omega, S^{d\times d})$ in (10) by $C_0^2(\Omega, S^{d\times d})$: Indeed, for $\phi \in C_0^2(\Omega, S^{d\times d})$ such that $\|\phi\|_\infty \leq \beta$ and $\|\mathrm{div}\, \phi\|_\infty \leq \alpha$ one may choose a sequence $\{\phi^n\}$ in $C_c^2(\Omega, S^{d\times d})$ such that $\lim_{n\to\infty} \phi^n = \phi$ in $C_0^2(\Omega, S^{d\times d})$. Convolving with smooth mollifiers, we can moreover ensure that, without loss of generality, each $\phi^n \in C_c^\infty(\Omega, S^{d\times d})$. Now, letting $\lambda_n = \max(\|\phi^n\|_\infty/\beta, \|\mathrm{div}\, \phi^n\|_\infty/\alpha, 1)$ we see by continuous embedding $C_0^2(\Omega, S^{d\times d}) \hookrightarrow C_0(\Omega, S^{d\times d})$ as well as continuity of $\mathrm{div} : C_0^2(\Omega, S^{d\times d}) \to C_0(\Omega, \mathbb{R}^d)$ that $\lim_{n\to\infty} \lambda_n = 1$. Thus, $\lim_{n\to\infty} \lambda_n^{-1} \phi^n = \phi$ and for each n, we have $\|\lambda_n^{-1} \phi^n\|_\infty \leq \beta$ as well as $\|\mathrm{div}\, \lambda_n^{-1} \phi^n\|_\infty \leq \alpha$. This proves the density.

Choosing $X = C_0^2(\Omega, S^{d\times d})$, $Y = C_0^1(\Omega, \mathbb{R}^d)$ with norm $\|\eta\|_{C^1} = \max(\|\eta\|_\infty, \|\nabla \eta\|_\infty)$, $\Lambda = \mathrm{div} \in \mathcal{L}(X, Y)$, and defining

$$F_1 : X \to\,]-\infty, \infty], \quad F_1(\phi) = \mathcal{I}_{\{\|\cdot\|_\infty \leq \beta\}}(\phi) - \langle \mathbf{c}u, \phi \rangle,$$
$$F_2 : Y \to\,]-\infty, \infty], \quad F_2(\eta) = \mathcal{I}_{\{\|\cdot\|_\infty \leq \alpha\}}(\eta) - \langle u, \mathrm{div}\, \eta \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the L^1–L^∞ duality pairing, we can equivalently write the supremum in (10) as

$$-\inf_{\phi \in X} F_1(\phi) + F_2(\Lambda \phi).$$

According to [3], the Fenchel–Rockafellar duality formula is valid if

$$Y = \bigcup_{\lambda \geq 0} \lambda(\mathrm{dom}\, F_2 - \Lambda \, \mathrm{dom}\, F_1).$$

This is indeed true since for each $\eta \in Y$, we can write $\eta = \lambda(\lambda^{-1}\eta - \mathrm{div}\, 0)$ with $\|\lambda^{-1}\eta\|_\infty \leq \alpha$ and $0 \in \mathrm{dom}\, F_1$. Thus, there is no duality gap between the primal and the dual problems and the dual problem admits a solution, i.e.,

$$\left(\inf_{\phi \in X} F_1(\phi) + F_2(\Lambda \phi) \right) + \left(\min_{w \in Y^*} F_1^*(-\Lambda^* w) + F_2^*(w) \right) = 0.$$

(In the case the primal functional is unbounded from below, this equation has to be understood in the sense that the dual functional is constant ∞.) Now, we calculate the convex conjugates on the right-hand side. For, a distribution $w \in Y^* = C_0^1(\Omega, \mathbb{R}^d)^*$ we have that $-\Lambda^* w = -\mathrm{div}^* w = \mathcal{E}w$ in the distributional sense: Indeed, testing $-\mathrm{div}^* w$ with $\phi \in C_c^\infty(\Omega, S^{d\times d})$ gives $\langle -\mathrm{div}^* w, \phi \rangle = \langle w, -\mathrm{div}\, \phi \rangle = \langle \mathcal{E}w, \phi \rangle$ for the duality pairing $\langle \cdot, \cdot \rangle$ of distributions and test functions. Thus,

$$F_1^*(-\Lambda^* w) = \sup_{\substack{\phi \in X, \\ \|\phi\|_\infty \leq \beta}} \langle w, -\mathrm{div}\, \phi \rangle + \langle \mathbf{c}u, \phi \rangle = \sup_{\substack{\phi \in C_c^\infty(\Omega, S^{d\times d}), \\ \|\phi\|_\infty \leq \beta}} \langle \mathcal{E}w, \phi \rangle + \langle \mathbf{c}u, \phi \rangle = \beta \|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}}$$

where, by definition of the Radon norm for distributions, $\|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}} = \infty$ if $\mathcal{E}w + \mathbf{c}u$ is not in $\mathcal{M}(\Omega, S^{d\times d})$. Note that for the above identity, we used that $\{\phi \in C_c^\infty(\Omega, S^{d\times d}) : \|\phi\|_\infty \leq \beta\}$ is dense in $\{\phi \in C_0^2(\Omega, S^{d\times d}) : \|\phi\|_\infty \leq \beta\}$ which can be seen analogously to the density argument at the beginning of the proof. Now, since $u \in L^1(\Omega)$, we have $\mathbf{c}u \in L^1(\Omega, S^{d\times d})$ and $\|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}}$ is finite if and only if $\mathcal{E}w \in \mathcal{M}(\Omega, S^{d\times d})$. But this is the case if and only if $w \in \mathrm{BD}(\Omega)$.

Furthermore, we see that $\mathrm{div}^* = -\nabla$ in the distributional sense, so by arguments analogous to the above, it follows that

$$F_2^*(w) = \sup_{\substack{\eta \in Y, \\ \|\eta\|_\infty \le \alpha}} \langle w, \eta \rangle + \langle u, \mathrm{div}\,\eta \rangle = \alpha \|\nabla u - w\|_{\mathcal{M}}$$

where, again, $\|\nabla u - w\|_{\mathcal{M}} = \infty$ if $\nabla u - w \notin \mathcal{M}(\Omega, \mathbb{R}^d)$. Consequently, for $w \in \mathrm{BD}(\Omega)$, the norm is finite if and only if $u \in \mathrm{BV}(\Omega)$.

These considerations yield that indeed, minimizing $F_1^*(-\Lambda^* w) + F_2^*(w)$ is equivalent to minimizing the functional in (9), that it suffices to choose w from $\mathrm{BD}(\Omega)$ instead of Y^*, and that the minimum is ∞ if and only if $u \notin \mathrm{BV}(\Omega)$. This gives the desired identity (10).

In particular, as the optimal value for the dual problem is attained, there always exists a $w \in \mathrm{BD}(\Omega)$ which admits the minimum in (9). $\qquad \square$

Since the dual formulation of oscillation TGV is a pointwise supremum of continuous functions, we can easily derive that it is lower semi-continuous with respect to the L^p-topology for each $p \in [1, \infty]$. In summary, it is a semi-norm that shares the following properties:

Proposition 2. *The functional* $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ *is a proper, and lower semi-continuous one-homogeneous functional on* $L^1(\Omega)$ *that satisfies the triangle inequality. Further,* $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) = 0$ *if and only if u satisfies (7) for some $C_1, C_2 \in \mathbb{R}$ in case of $\omega \neq 0$ and u affine in case of $\omega = 0$.*

Proof. We only discuss the case $\omega \neq 0$ as, by the above remark, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ reduces to second-order TGV in case $\omega = 0$ for which the claimed properties were established in [12].

By Proposition 1, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ is the convex conjugate of a convex indicator functional associated with a nonempty, convex and balanced set. Thus, it is proper, convex and positive one-homogeneous; we refer to the proof of statement 1 of Proposition 3.3 in [12]. In particular, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ satisfies the triangle inequality. Further, for each $\phi \in C_c^\infty(\Omega, S^{d \times d})$, the mapping $u \mapsto \int_\Omega u(\mathrm{div}^2\,\phi + \mathbf{c} \cdot \phi)\,\mathrm{d}x$ is linear and continuous on $L^1(\Omega)$, so $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ is lower semi-continuous on $L^1(\Omega)$ as the pointwise supremum of continuous functionals, see (10).

Finally, for $u \in L^1(\Omega)$, we have $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) = 0$ if and only if $\int_\Omega u(\mathrm{div}^2\,\phi + \mathbf{c} \cdot \phi)\,\mathrm{d}x = 0$ for all $\phi \in C_c^\infty(\Omega, S^{d \times d})$ with $\|\phi\|_\infty \le \beta$, $\|\mathrm{div}\,\phi\|_\infty \le \alpha$. The latter, however, is equivalent to $\int_\Omega u(\mathrm{div}^2\,\phi + \mathbf{c} \cdot \phi)\,\mathrm{d}x = 0$ for all $\phi \in C_c^2(\Omega, S^{d \times d})$ which, in turn, means $\mathcal{E}\nabla u + \mathbf{c}u = 0$ in Ω in the weak sense.

We show that this equation is also satisfied in the strong sense. For this purpose, observe that it is in particular true for $u * \rho_\epsilon$ with a mollifier $\rho_\epsilon \in C_c^\infty(B_\epsilon(0))$ and each subdomain $\Omega' \subset \Omega$ with $\mathrm{dist}(\Omega', \partial\Omega) > \epsilon$. Indeed, this follows from

$$\int_{\Omega'} (u * \rho_\epsilon)(\mathrm{div}^2\,\phi + \mathbf{c} \cdot \phi)\,\mathrm{d}x = \int_\Omega u\big(\mathrm{div}^2(\phi * \bar\rho_\epsilon) + \mathbf{c} \cdot (\phi * \bar\rho_\epsilon)\big)\,\mathrm{d}x = 0$$

where $\bar\rho_\epsilon(x) = \rho_\epsilon(-x)$, since $\phi * \bar\rho_\epsilon \in C_c^\infty(\Omega, S^{d \times d})$ whenever $\phi \in C_c^\infty(\Omega', S^{d \times d})$. But $u * \rho_\epsilon \in C_c^\infty(\Omega')$, so we have $\mathcal{E}\nabla(u * \rho_\epsilon) + \mathbf{c}(u * \rho_\epsilon) = 0$ in Ω' also in the strong sense. Thus, $u * \rho_\epsilon$ satisfies (7) in Ω'. Fixing Ω' and letting $\epsilon \to 0$ now gives $u * \rho_\epsilon \to u$ in $L^1(\Omega')$. In particular, $u * \rho_\epsilon$ converges to u in the finite-dimensional space of functions in $L^1(\Omega')$ satisfying (7), which implies by closedness of finite-dimensional spaces that u also satisfies (7) for some $C_1, C_2 \in \mathbb{R}$ and, consequently, $u \in C^2(\Omega')$. As Ω' can be chosen arbitrarily such that $\overline{\Omega'}$ is compact in Ω, the latter also holds true in Ω. $\qquad \square$

Furthermore, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ is translation invariant as well as rotationally invariant with respect to a rotated $\mathbf{c}$:

Proposition 3. *For each $y \in \mathbb{R}^d$ and $u \in L^1(\Omega)$, if $\tilde{u}(x) = u(x - y)$, then $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(\tilde{u}) = \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u)$, i.e., $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ is translation-invariant.*

Furthermore, for each orthonormal matrix $O \in \mathbb{R}^{d\times d}$ and $u \in L^1(\Omega)$, if $\tilde{u}(x) = u(Ox)$, then $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(\tilde{u}) = \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\tilde{\mathbf{c}}}(u)$ with $\tilde{\mathbf{c}} = O\omega \otimes O\omega$. In other words, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ is rotationally invariant.

Proof. First, let us define the following set

$$K_{\alpha,\beta}(\Omega) = \left\{ \phi \in C_c^\infty(\Omega, S^{d\times d}) : \|\phi\|_\infty \leq \beta, \ \|\operatorname{div}\phi\|_\infty \leq \alpha \right\}.$$

For $\phi \in K_{\alpha,\beta}(\Omega + y)$ we set $\tilde{\phi}(x) = \phi(x + y)$. Then, $\|\tilde{\phi}\|_\infty \leq \beta$, $\operatorname{div}\tilde{\phi}(x) = \operatorname{div}\phi(x + y)$ and $\operatorname{div}^2 \tilde{\phi}(x) = \operatorname{div}^2 \phi(x+y)$ such that $\|\operatorname{div}\tilde{\phi}\|_\infty \leq \alpha$ and, consequently, $\tilde{\phi} \in K_{\alpha,\beta}(\Omega)$. Interchanging the roles of ϕ and $\tilde{\phi}$, it follows that $\phi \in K_{\alpha,\beta}(\Omega + y)$ if and only if $\tilde{\phi} \in K_{\alpha,\beta}(\Omega)$. For $\phi \in K_{\alpha,\beta}(\Omega + y)$ it holds that

$$\int_{\Omega+y} \tilde{u}(x)\big(\operatorname{div}^2 \phi(x) + \mathbf{c} \cdot \phi(x)\big)\, \mathrm{d}x = \int_\Omega u(x)\big(\operatorname{div}^2 \tilde{\phi}(x) + \mathbf{c} \cdot \tilde{\phi}(x)\big)\, \mathrm{d}x,$$

hence taking the supremum yields with the above that $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(\tilde{u}) = \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u)$.

For $\phi \in K_{\alpha,\beta}(O^T\Omega)$ we set $\tilde{\phi}(x) = O\phi(O^T x)O^T$ which satisfies $\tilde{\phi} \in C_c^2(\Omega, S^{d\times d})$ as well as $\|\tilde{\phi}\|_\infty \leq \beta$, the latter due to the invariance of the Frobenius norm under orthonormal similarity transforms. Moreover, we have

$$\operatorname{div}\tilde{\phi}(x) = \operatorname{div}\phi(O^T x)O^T, \qquad \operatorname{div}^2 \tilde{\phi}(x) = \operatorname{div}^2 \phi(O^T x),$$

see, for instance, (A.4) in [12]. Therefore, $\|\operatorname{div}\tilde{\phi}\|_\infty \leq \alpha$, so $\tilde{\phi} \in K_{\alpha,\beta}(\Omega)$. Interchanging the roles of ϕ and $\tilde{\phi}$ it follows that $\phi \in K_{\alpha,\beta}(O^T\Omega)$ if and only if $\tilde{\phi} \in K_{\alpha,\beta}(\Omega)$.

Eventually, for each $\phi \in K_{\alpha,\beta}(O^T\Omega)$, we have

$$\begin{aligned}
\int_{O^T\Omega} \tilde{u}(x)\big(\operatorname{div}^2 \phi(x) + \mathbf{c} \cdot \phi(x)\big)\, \mathrm{d}x &= \int_{O^T\Omega} u(Ox)\big(\operatorname{div}^2 \phi(x) + \mathbf{c} \cdot \phi(x)\big)\, \mathrm{d}x \\
&= \int_\Omega u(x)\big(\operatorname{div}^2 \phi(O^T x) + \mathbf{c} \cdot \phi(O^T x)\big)\, \mathrm{d}x \\
&= \int_\Omega u(x)\big(\operatorname{div}^2 \tilde{\phi}(x) + \mathbf{c} \cdot (O^T \tilde{\phi}(x)O)\big)\, \mathrm{d}x \\
&= \int_\Omega u(x)\big(\operatorname{div}^2 \tilde{\phi}(x) + (O\mathbf{c}O^T) \cdot \tilde{\phi}(x)\big)\, \mathrm{d}x.
\end{aligned}$$

Taking the supremum and noting that $O\mathbf{c}O^T = O\omega \otimes O\omega = \tilde{\mathbf{c}}$ then leads to $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(\tilde{u}) = \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\tilde{\mathbf{c}}}(u)$. $\qquad\square$

We would like to examine Banach spaces that are associated with the semi-norm $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$. To construct a norm, we add, as it is also done in case of TGV, the L^1-norm. The resulting Banach space

$$\mathrm{BGV}^{\mathrm{osci}}(\Omega) = \{u \in L^1(\Omega) : \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) < \infty\}, \qquad \|u\|_{\mathrm{BGV}^{\mathrm{osci}}} = \|u\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u)$$

then coincides with $\mathrm{BV}(\Omega)$. Indeed, this is a consequence of the equivalence of norms which is proven in the following lemma.

Lemma 2. *For each $\omega \in \mathbb{R}^d$, there exist constants $0 < c < C < \infty$ such that for each $u \in L^1(\Omega)$, we have*

$$c\|u\|_{\mathrm{BV}} \leq \|u\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) \leq C\|u\|_{\mathrm{BV}}. \tag{11}$$

Proof. Setting $w = 0$ in the definition (9) implies immediately that

$$\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) \leq \alpha\|\nabla u\|_{\mathcal{M}} + \beta\|\mathbf{c}u\|_1 \leq \alpha\,\mathrm{TV}(u) + \beta|\omega|^2\|u\|_1.$$

Adding the L^1-norm on both sides give the second inequality of (11) for $C = \max\{\alpha, 1 + \beta|\omega|^2\}$.

For the converse inequality, we may assume $u \in \mathrm{BV}(\Omega)$ since otherwise, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) = \infty$ by definition. Due to the norm equivalence in $\mathrm{BV}(\Omega)$ and $\mathrm{BGV}^2_{\alpha,\beta}(\Omega)$ shown, e.g., in [51, 14, 11], there exists a constant $C_1 > 0$ such that for each $w \in \mathrm{BD}(\Omega)$,

$$
\begin{aligned}
C_1\|u\|_{\mathrm{BV}} &\leq \|u\|_1 + \mathrm{TGV}^2_{\alpha,\beta}(u) \\
&\leq \|u\|_1 + \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}} \\
&= \|u\|_1 + \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w + \mathbf{c}u - \mathbf{c}u\|_{\mathcal{M}} \\
&\leq \|u\|_1 + \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}} + \beta\|\mathbf{c}u\|_1 \\
&\leq C_2(\|u\|_1 + \alpha\|\nabla u - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w + \mathbf{c}u\|_{\mathcal{M}})
\end{aligned}
$$

where $C_2 = \max\{1, 1 + \beta|\omega|^2\}$. Choosing $c = \frac{C_1}{C_2}$ and minimizing over $w \in \mathrm{BD}(\Omega)$ finally yields the first inequality of (11). $\qquad\square$

Depending on the results of Lemma 2, we can further deduce following inequality of Poincaré–Wirtinger type which implies the coercivity of oscillation TGV in the BV-space.

Proposition 4. *Let* $P : L^1(\Omega) \to \ker(\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}})$ *a linear and continuous projection onto the kernel* $\ker(\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}})$ *(which exists), then there is a* $C > 0$ *such that*

$$\|u - Pu\|_{\mathrm{BV}} \leq C\,\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) \quad \text{for all} \quad u \in \mathrm{BV}(\Omega). \tag{12}$$

Proof. If the statement is not true, then there exists a sequence $\{u^n\}$ in $\mathrm{BV}(\Omega)$ such that

$$\|u^n - Pu^n\|_{\mathrm{BV}} = 1, \quad \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^n) \leq \frac{1}{n},$$

for all $n \in \mathbb{N}$. By the compact embedding $\mathrm{BV}(\Omega) \hookrightarrow L^1(\Omega)$, it follows that $\lim_{n\to\infty} u^n - Pu^n = u^*$ in $L^1(\Omega)$ for some subsequence (not relabeled). We have $P(u^n - Pu^n) = 0$ for all n, consequently, the limit has to satisfy $Pu^* = 0$ by continuity of P in $L^1(\Omega)$.

As P projects onto the kernel of $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$, we have $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(Pu^n) = 0$ for all n and can deduce that $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^n - Pu^n) = \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^n)$. On the other hand, the lower semi-continuity of $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ implies

$$\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^*) \leq \liminf_{n\to\infty} \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^n) = 0,$$

hence $u^* \in \ker(\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}) \cap \ker(P)$. As P is a projection with $\mathrm{ran}(P) = \ker(\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}})$, this implies $u^* = Pu^* = 0$, so $\lim_{n\to\infty} u^n - Pu^n = 0$ in $L^1(\Omega)$.

Now, according to Lemma 2, there is a $C_1 > 0$ such that

$$\|u^n - Pu^n\|_{\mathrm{BV}} \leq C_1\big(\|u^n - Pu^n\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u^n)\big), \tag{13}$$

so $\lim_{n\to\infty} u^n - Pu^n = 0$ in $\mathrm{BV}(\Omega)$ which is a contradiction to $\|u^n - Pu^n\|_{\mathrm{BV}} = 1$ for all n. $\quad\square$

Remark 1. *By continuous embedding* $\mathrm{BV}(\Omega) \hookrightarrow L^p(\Omega)$ *for* $1 \leq p \leq d/(d-1)$ *we have in particular coercivity of* $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ *in* $L^p(\Omega)$ *according to*

$$\|u - Pu\|_p \leq C\,\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}(u) \quad \text{for each} \quad u \in L^p(\Omega).$$

3   Infimal convolution of oscillation TGV functionals

So far, we examined oscillation TGV functionals associated with a single direction $\omega \in \mathbb{R}^d$. Next, we aim at incorporating multiple directions $\omega_1, \ldots, \omega_m \in \mathbb{R}^d$ into one functional in order to represent structured oscillations in an image with respect to multiple orientations and frequencies. This will be done in terms of the infimal convolution. Recall that, given two functionals J_1 and J_2, their infimal convolution is defined by

$$(J_1 \square J_2)(u) = \inf_v J_1(u-v) + J_2(v) = \inf_{u = u_1 + u_2} J_1(u_1) + J_2(u_2). \tag{14}$$

For m functionals $J_1, \ldots, J_m$ the definition is iterated leading to

$$(J_1 \square \ldots \square J_m)(u) = \inf_{u = u_1 + \ldots + u_m} \sum_{i=1}^m J_i(u_i).$$

We aim at studying the m-fold infimal convolution of oscillation TGV given as follows:

$$\text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{\mathbf{c}}}(u) = (\text{TGV}^{\text{osci}}_{\alpha_1, \beta_1, \mathbf{c}_1} \square \ldots \square \text{TGV}^{\text{osci}}_{\alpha_m, \beta_m, \mathbf{c}_m})(u) = \inf_{u = u_1 + \ldots + u_m} \sum_{i=1}^m \text{TGV}^{\text{osci}}_{\alpha_i, \beta_i, \mathbf{c}_i}(u_i). \tag{15}$$

Here, $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$, $\vec{\beta} = (\beta_1, \ldots, \beta_m)$ are vectors of positive weights and $\vec{\mathbf{c}} = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ is given by $\mathbf{c}_i = \omega_i \otimes \omega_i$ for $\omega_1, \ldots, \omega_m \in \mathbb{R}^d$.

It is well-known that the infimal convolution of the convex functionals is also convex. Also, since $\text{TGV}^{\text{osci}}_{\alpha_i, \beta_i, \mathbf{c}_i}(0) = 0$ for each i, one can easily obtain that $\text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{\mathbf{c}}}$ is proper. Therefore, in the following section, we concentrate on the coercivity and lower semi-continuity of $\text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{\mathbf{c}}}$ in L^p-spaces.

As it is typical for infimal convolutions, the lower semi-continuity and coercivity of $\text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{\mathbf{c}}}$ is not immediate. We study these properties in a more abstract framework that also allows for generalization to a larger class of regularizers.

Definition 2. *Let X be a real Banach space and $\Phi : X \to [0, \infty]$. Then, Φ is called a* lower semi-continuous semi-norm *on X if Φ satisfies the following properties:*

1. *Φ is proper,*
2. *Φ is positive one-homogeneous, i.e., $\Phi(\lambda u) = |\lambda| \Phi(u)$ for all $\lambda \in \mathbb{R}$, $u \in X$,*
3. *Φ satisfies the triangle inequality, i.e., $\Phi(u + v) \leq \Phi(u) + \Phi(v)$ for all $u, v \in X$,*
4. *Φ is lower semi-continuous, i.e., $\Phi(u) \leq \liminf_{n \to \infty} \Phi(u^n)$ whenever $u^n \to u$ in X as $n \to \infty$,*
5. *there is a linear, continuous and onto projection operator $P : X \to \ker(\Phi)$.*

Further, Φ is called coercive *in X if there exists a $C > 0$ such that*

$$\|u - Pu\|_X \leq C\Phi(u) \quad \text{for all} \quad u \in X.$$

Clearly, $\text{TGV}^{\text{osci}}_{\alpha, \beta, \mathbf{c}}$ are lower semi-continuous semi-norms on $\text{BV}(\Omega)$ and $L^p(\Omega)$ for all $1 \leq p \leq \infty$ (see Proposition 2 and recall that one can always find continuous projections onto finite-dimensional subspaces). They are coercive in $\text{BV}(\Omega)$ and in $L^p(\Omega)$ for $1 \leq p \leq d/(d-1)$, see Proposition 4 and Remark 1.

In the following, we fix a Banach space X and let Φ_1, Φ_2 be lower semi-continuous semi-norms on X with associated projections

$$P_1 : X \to \ker(\Phi_1), \qquad P_2 : X \to \ker(\Phi_2).$$

Further, assume that there is a linear, continuous and onto projection operator

$$Q : X \to \ker(\Phi_1) \cap \ker(\Phi_2).$$

In this situation, we have the following basic estimate.

Lemma 3. *Let Φ_1, Φ_2 be lower semi-continuous semi-norms on X and $Q : X \to \ker(\Phi_1) \cap \ker(\Phi_2)$ be a linear, continuous and onto projection. Then, there is a $C > 0$ such that*

$$\|v\|_X \leq C(\|v - P_1 v\|_X + \|v - P_2 v\|_X) \quad \text{for all} \quad v \in \operatorname{ran}(I - Q). \tag{16}$$

Proof. Consider the mapping $T : \operatorname{ran}(I - Q) \to \operatorname{ran}(T)$ according to

$$Tv = \begin{pmatrix} v - P_1 v \\ v - P_2 v \end{pmatrix},$$

which defines a mapping between Banach spaces since the fact that P_1, P_2, Q are continuous projection operators implies that $\operatorname{ran}(I - Q)$ and $\operatorname{ran}(T)$ are closed. Obviously, T is linear, continuous and surjective. To see that T is also injective, let $Tv = 0$ for some $v \in X$, i.e., $v - P_1 v = 0$ and $v - P_2 v = 0$. Then, $v \in \ker(\Phi_1) \cap \ker(\Phi_2)$ which implies $Qv = v$ such that $(I - Q)v = 0$ follows. Hence, $v \in \operatorname{ran}(I - Q) \cap \ker(I - Q) = \{0\}$, i.e., $v = 0$, as Q is, once again, a continuous projection operator. By the open mapping theorem, we have that T^{-1} is bounded, which leads to the desired result. $\square$

First, we show the exactness and lower semi-continuity of the infimal convolution of two lower semi-continuous semi-norms under suitable assumptions.

Lemma 4. *Let Φ_1, Φ_2 be lower semi-continuous and coercive semi-norms on the reflexive Banach space X with finite-dimensional kernels. Then $\Phi_1 \square \Phi_2$ is exact, i.e., for each $u \in X$ there is a $v^* \in X$ such that*

$$v^* \in \operatorname*{arg\,min}_{v \in X} \ \Phi_1(u - v) + \Phi_2(v). \tag{17}$$

Moreover, $\Phi_1 \square \Phi_2$ constitutes a lower semi-continuous semi-norm on X with finite-dimensional kernel.

For general X, the statements remain true if, for $i = 1$ or $i = 2$, the sublevel sets $\{u \in X : \|u\|_X + \Phi_i(u) \leq C\}$ are weakly relatively compact for each $C > 0$.

Proof. Clearly, $\Phi_1 \square \Phi_2$ is a non-negative functional. Choose a convergent sequence $\{u^n\}$ in X with limit u such that $\liminf_{n \to \infty}(\Phi_1 \square \Phi_2)(u^n) < \infty$. Going to subsequences (without relabeling), we may assume that $\lim_{n \to \infty}(\Phi_1 \square \Phi_2)(u^n)$ exists and is finite. Thus, we can find a sequence $\{v^n\}$ in X such that $\Phi_1(u^n - v^n) + \Phi_2(v^n) \leq (\Phi_1 \square \Phi_2)(u^n) + \frac{1}{n}$ for all n, meaning in particular that $\liminf_{n \to \infty} \Phi_1(u^n - v^n) + \Phi_2(v^n) \leq \lim_{n \to \infty}(\Phi_1 \square \Phi_2)(u^n)$.

Further, as $\ker(\Phi_1), \ker(\Phi_2), \ker(\Phi_1) \cap \ker(\Phi_2)$ are finite-dimensional subspaces of X, we can find linear, continuous and onto projections P_1, P_2, Q onto the respective spaces as previously stated.

Now, choosing $\tilde{v}^n = v^n - Qv^n$ instead of v^n does not change the functional values since

$$\Phi_1(u^n - \tilde{v}^n) + \Phi_2(\tilde{v}^n) = \Phi_1(u^n - v^n + Qv^n) + \Phi_2(v^n - Qv^n)$$
$$= \Phi_1(u^n - v^n) + \Phi_2(v^n),$$

as a consequence of each $Qv^n \in \ker(\Phi_1) \cap \ker(\Phi_2)$. In particular, we have that $\Phi_1(u^n - \tilde{v}^n)$ and $\Phi_2(\tilde{v}^n)$ are bounded. Since Φ_1 and Φ_2 are coercive in X, there is a constant $C > 0$ such that $\|u^n - \tilde{v}^n - P_1(u^n - \tilde{v}^n)\|_X \leq C\Phi_1(u^n - \tilde{v}^n)$ and $\|\tilde{v}^n - P_2\tilde{v}^n\|_X \leq C\Phi_2(\tilde{v}^n)$ for each n. Since $\|u^n - P_1 u^n\|_X$ is bounded due to the convergence $u^n - P_1 u^n \to u - P_1 u$, we have that $\|\tilde{v}^n - P_1\tilde{v}^n\|_X$ and $\|\tilde{v}^n - P_2\tilde{v}^n\|_X$ are bounded, so Lemma 3 implies that $\{\tilde{v}^n\}$ is bounded in X. By reflexivity, there exists a weakly convergent subsequence (not relabeled) such that $\lim_{n \to \infty} \tilde{v}^n = v^*$ weakly in X. Lower semi-continuity and convexity of Φ_1 and Φ_2 implies $\Phi_1(u - v^*) \leq \liminf_{n \to \infty} \Phi_1(u^n - \tilde{v}^n)$ and $\Phi_2(v^*) \leq \liminf_{n \to \infty} \Phi_2(\tilde{v}^n)$. Then,

$$(\Phi_1 \square \Phi_2)(u) \leq \Phi_1(u - v^*) + \Phi_2(v^*) \leq \liminf_{n \to \infty} \Phi_1(u^n - v^n) + \Phi_2(v^n) \leq \lim_{n \to \infty} (\Phi_1 \square \Phi_2)(u^n).$$

12

This shows that $\Phi_1 \square \Phi_2$ is lower semi-continuous. Plugging in constant sequences, i.e., $u^n = u$ for all n, yields in particular that $(\Phi_1 \square \Phi_2)(u) = \Phi_1(u - v^*) + \Phi_2(v^*)$, meaning that $\Phi_1 \square \Phi_2$ is exact.

For $\Phi_1 \square \Phi_2$ being a lower semi-continuous semi-norm in X, it remains to show that the functional is proper, positive one-homogeneous and satisfies the triangle inequality. But these properties are immediate from the definition and Φ_1, Φ_2 being lower semi-continuous semi-norms. Finally, note that $(\Phi_1 \square \Phi_2)(u) = 0$ if and only if $u \in \overline{\ker(\Phi_1) + \ker(\Phi_2)}$. But since, again, $\ker(\Phi_1), \ker(\Phi_2)$ are finite-dimensional we have that $\overline{\ker(\Phi_1) + \ker(\Phi_2)} = \ker(\Phi_1) + \ker(\Phi_2)$ is finite-dimensional and therefore, there exists a linear, continuous and onto projection $P : X \to \ker(\Phi_1) + \ker(\Phi_2)$. This establishes the remaining property.

To deduce the result without the reflexivity assumption on X, we assume, without loss of generality, that the sets $\{u \in X : \|u\|_X + \Phi_2(u) \leq C\}$ are weakly relatively compact for each $C > 0$. Proceeding along the lines of the above proof and observing that $\{\tilde{v}^n\}$ as well as $\{\Phi_2(\tilde{v}^n)\}$ are still bounded then allows to extract the weakly convergent subsequence of $\{\tilde{v}^n\}$. The rest of the proof is completely analogous. $\qquad\square$

In order to use $\Phi_1 \square \Phi_2$ as a regularizer, we still need criteria for coercivity of this infimal convolution functional. We will obtain this property from a norm equivalence result that is a generalization of the norm equivalence in Lemma 2.

Proposition 5. *Let X, Y be Banach spaces with X continuously embedded in Y. Assume that Φ_1, Φ_2 are lower semi-continuous semi-norms on Y, are coercive on X, possess a finite-dimensional kernel and satisfy*

$$c_i \|u\|_X \leq \|u\|_Y + \Phi_i(u) \leq C_i \|u\|_X \quad \text{for all} \quad u \in X$$

for constants $0 < c_i < C_i < \infty$, $i = 1, 2$ independent of u. Then, $\Phi_1 \square \Phi_2$ satisfies

$$c\|u\|_X \leq \|u\|_Y + (\Phi_1 \square \Phi_2)(u) \leq C\|u\|_X \quad \text{for all} \quad u \in X \tag{18}$$

for $0 < c < C < \infty$ independent of u.

Moreover, if X is compactly embedded in Y, then $\Phi_1 \square \Phi_2$ is a lower semi-continuous semi-norm on Y and coercive on X, i.e., there exists a $C' > 0$ such that

$$\|u - Pu\|_X \leq C'(\Phi_1 \square \Phi_2)(u) \quad \text{for all} \quad u \in X, \tag{19}$$

where $P : Y \to \ker(\Phi_1) + \ker(\Phi_2)$ is a linear, continuous and onto projector on the kernel of $\Phi_1 \square \Phi_2$.

Proof. We will first establish the norm equivalence of $u \mapsto \|u\|_Y + (\Phi_1 \square \Phi_2)(u)$ with the norm in X. For that purpose, let $u \in X$ and estimate, employing the definition of the infimal convolution and the fact that Φ_2 is semi-norm,

$$\|u\|_Y + (\Phi_1 \square \Phi_2)(u) \leq \|u\|_Y + \Phi_1(u) \leq C_1 \|u\|_X.$$

Next, suppose that the converse estimate is not true. Then, there exists a sequence $\{u^n\}$ in X with

$$\|u^n\|_X = 1, \qquad \|u^n\|_Y + (\Phi_1 \square \Phi_2)(u^n) \leq \frac{1}{n}$$

for all n. Hence, $\lim_{n \to \infty} u^n = 0$ in Y and one can find a sequence $\{v^n\}$ in X such that $\Phi_1(u^n - v^n) \leq \frac{2}{n}$ and $\Phi_2(v^n) \leq \frac{2}{n}$ for all n. By the coercivity of Φ_1 and Φ_2 in X, we have

$$\|(u^n - v^n) - P_1(u^n - v^n)\|_X \leq C_1' \Phi_1(u^n - v^n), \qquad \|v^n - P_2 v^n\|_X \leq C_2' \Phi_2(v^n)$$

where $P_i : Y \to \ker(\Phi_i)$ and C_i' are the projections and coercivity constants associated with the Φ_i, respectively, for $i = 1, 2$. Consequently,

$$\|u^n - P_1 u^n - (P_2 - P_1)v^n\|_X \leq \|u^n - v^n - P_1(u^n - v^n)\|_X + \|v^n - P_2 v^n\|_X \to 0$$

as $n \to \infty$. Since $\|u^n\|_Y \to 0$ as $n \to \infty$, it follows by continuous embedding $X \hookrightarrow Y$ that $P_1 u^n + (P_2 - P_1)v^n \to 0$ as $n \to \infty$ in Y. However, this sequence is contained in the finite-dimensional space $\ker(\Phi_1) + \ker(\Phi_2)$, so we also have $\lim_{n \to \infty} P_1 u^n + (P_2 - P_1)v^n = 0$ in X. Hence, $\|u^n\|_X \to 0$ as $n \to \infty$ which is a contradiction. So, there has to be a constant $c > 0$ such that

$$c\|u\|_X \leq \|u\|_Y + (\Phi_1 \square \Phi_2)(u) \quad \text{for all} \quad u \in X.$$

Now, assume that X is compactly embedded in Y. Then, $\Phi_1 \square \Phi_2$ is a lower semi-continuous semi-norm on Y. This follows from applying Lemma 4 setting $X = Y$ and observing that the sublevel sets $\{u \in Y : \|u\|_Y + \Phi_i(u) \leq C\}$ are relatively compact in Y for each $C > 0$ and $i = 1, 2$. Indeed, the latter follows from the norm estimate $\|u\|_X \leq c_i^{-1}(\|u\|_Y + \Phi_i(u))$ as well as the compact embedding of X in Y.

Finally, the coercivity of $\Phi_1 \square \Phi_2$ in X follows in the lines of proof of Proposition 4. Choose a projector $P : Y \to \ker(\Phi_1) + \ker(\Phi_2)$ according to the statement which has to exist as the kernels are finite-dimensional. Assuming that (19) does not hold leads to the existence of a sequence $\{u^n\}$ for which

$$\|u^n - Pu^n\|_X = 1, \qquad (\Phi_1 \square \Phi_2)(u^n) \leq \frac{1}{n}$$

for each n. By compact embedding, for a subsequence of $\{u^n\}$ (not relabeled), we have $u^n - Pu^n \to u^*$ as $n \to \infty$ in Y for some $u^* \in Y$. Since P is a projection, one obtains $P(u^n - Pu^n) = 0$ for each n such that by continuity of P in Y, it follows that $Pu^* = 0$, i.e., $u^* \in \ker(P)$. Moreover, $(\Phi_1 \square \Phi_2)(u^n - Pu^n) = (\Phi_1 \square \Phi_2)(u^n)$ for each n, such that $(\Phi_1 \square \Phi_2)(u^n - Pu^n) \to 0$ as $n \to \infty$. By lower semi-continuity of $\Phi_1 \square \Phi_2$ in Y, $(\Phi_1 \square \Phi_2)(u^*) = 0$, meaning that $u^* \in \text{ran}(P)$. Together with $u^* \in \ker(P)$ and P being a projection, this implies $u^* = 0$.

Thus, the norm equivalence (18) implies

$$\|u^n - Pu^n\|_X \leq c^{-1}(\|u^n - Pu^n\|_Y + (\Phi_1 \square \Phi_2)(u^n)) \to 0 \quad \text{as} \quad n \to \infty,$$

which is a contradiction to $\|u^n - Pu^n\|_X = 1$ for all n. Hence, $\Phi_1 \square \Phi_2$ has to be coercive as stated. $\qquad \square$

Having a lower semi-continuous and coercive semi-norm Φ allows to solve linear inverse problems $Ku = f$ with Φ as regularization functional. The following result is easily obtained using techniques from calculus of variations (see, for instance, [14, 11, 7]). However, we provide the argumentation for the sake of completeness.

Proposition 6. *Let X, Y be Banach spaces, X be reflexive, $K : X \to Y$ linear and continuous, $F : Y \to]-\infty, \infty]$ convex, lower semi-continuous, coercive and bounded from below, and Φ a lower semi-continuous and coercive semi-norm on X with finite-dimensional kernel. Then, there exists a solution to the minimization problem*

$$\min_{u \in X} \ F(Ku) + \Phi(u). \tag{20}$$

Proof. There is only something to show if the objective functional is proper. In that case, it is bounded from below, so one can choose an infimizing sequence $\{u^n\}$ for which the objective functional converges to a finite value. Denoting by $Z = \ker(\Phi) \cap \ker(K)$, we see that Z has finitely many dimensions and therefore, there exists a linear, continuous and onto projection operator

14

$Q : X \to Z$. Denoting by $P : X \to \ker(\Phi)$ the projection associated with the semi-norm Φ, we can moreover achieve that $PQ = Q$.

As Z is a subspace of both $\ker(\Phi)$ and $\ker(K)$, we can replace u^n by $u^n - Qu^n$ and get

$$F\big(K(u^n - Qu^n)\big) + \Phi(u^n - Qu^n) = F(Ku^n) + \Phi(u^n),$$

implying that $\{u^n - Qu^n\}$ is an infimizing sequence. Thus, we can assume without loss of generality that $\{u^n\}$ is an infimizing sequence in the reflexive Banach space $\operatorname{ran}(I - Q)$. Now, $\{u^n\}$ has to be bounded: From the coercivity of Φ in X we get that $\|u^n - Pu^n\|_X$ is bounded. Furthermore, as F is coercive, $\|Ku^n\|_Y$ is bounded, so we have

$$\|KPu^n\|_Y \leq \|Ku^n\|_Y + \|K\|\|u^n - Pu^n\|_X \leq C$$

for some $C > 0$. However, K is injective on $\operatorname{ran}\big(P(I - Q)\big) = \operatorname{ran}(P - Q)$: For $Ku = 0$ and $u \in \operatorname{ran}(P - Q) \subset \operatorname{ran}(P)$ it follows that $u \in \ker(K) \cap \ker(\Phi) = Z$, so $Qu = u$ as well as $Pu = u$ since $Z \subset \ker(\Phi)$. Hence, $(P - Q)u = 0$ which implies that $u = 0$ as $u \in \operatorname{ran}(P - Q)$ and $P - Q$ is a projection. Now, $\operatorname{ran}(P - Q)$ is finite-dimensional, so there exists a constant $c > 0$ such that $c\|Pu\|_X \leq \|KPu\|_Y$ for each $u \in \operatorname{ran}(I - Q)$. This implies, since $\|KPu^n\|_Y$ is bounded, that $\|Pu^n\|_X$ is bounded and hence, that $\{u^n\}$ is bounded in X.

By reflexivity of X, we can extract a weakly convergent subsequence with limit u^*. One easily sees that the objective functional is convex and lower semi-continuous, so u^* has to be a minimizer. $\square$

It is then immediate that Tikhonov functionals for the regularized solution of $Ku = f$ admit minimizers.

Corollary 1. *In the situation of Proposition 6, let $f \in Y$ and $1 \leq q < \infty$. Then, there exists a minimizer of*

$$\min_{u \in X} \frac{\|Ku - f\|_Y^q}{q} + \Phi(u). \tag{21}$$

and the minimum is finite.

The general results presented here allow us to obtain properties of the general infimal convolution of oscillation TGV model with an arbitrary number of functionals.

Theorem 1. *The functional $\operatorname{ICTGV}_{\vec{\alpha},\vec{\beta},\vec{c}}^{\mathrm{osci}}$ according to (15) is a lower semi-continuous semi-norm on $L^1(\Omega)$ with*

$$\ker(\operatorname{ICTGV}_{\vec{\alpha},\vec{\beta},\vec{c}}^{\mathrm{osci}}) = \operatorname{span}\{x \mapsto \sin(\omega_i \cdot x), x \mapsto \cos(\omega_i \cdot x) : i = 1, \ldots, m \text{ with } \omega_i \neq 0\}$$

$$+ \begin{cases} \{x \mapsto a \cdot x + b : a \in \mathbb{R}^d, b \in \mathbb{R}\} & \text{if } \omega_i = 0 \text{ for some } i, \\ \{0\} & \text{else.} \end{cases}$$

It is coercive on $\operatorname{BV}(\Omega)$ and there exist constants $0 < c < C < \infty$ such that for each $u \in \operatorname{BV}(\Omega)$, it holds that

$$c\|u\|_{\mathrm{BV}} \leq \|u\|_1 + \operatorname{ICTGV}_{\vec{\alpha},\vec{\beta},\vec{c}}^{\mathrm{osci}}(u) \leq C\|u\|_{\mathrm{BV}}.$$

Proof. We proceed inductively with respect to m. For $m = 1$, the result is clear from Propositions 2 and 4 as well as Lemma 2, as already mentioned in the beginning of this section. The induction step is then a consequence of Proposition 5 with $X = \operatorname{BV}(\Omega)$ and $Y = L^1(\Omega)$, where X is compactly embedded in Y. $\square$

Remark 2. *In imaging applications, it is meaningful to choose $\omega_i = 0$ for some i, i.e., $\operatorname{TGV}_{\alpha_i,\beta_i,\mathbf{c}_i}^{\mathrm{osci}} = \operatorname{TGV}_{\alpha_i,\beta_i}^2$, such that piecewise smooth parts are captured the usual TGV functional of second order.*

Moreover, one might require that texture parts are sparse in a certain sense. The latter can, for instance, be incorporated in the infimal convolution model by adding a L^1-term such that, for $\omega_i \neq 0$, the infimal convolution is performed with respect to $\gamma_i \|\cdot\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha_i,\beta_i,\mathbf{c}_i}$ for some $\gamma_i > 0$. This result is the following variant of (15):

$$\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}},\vec{\gamma}} = (\gamma_1 \|\cdot\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha_1,\beta_1,\mathbf{c}_1}) \square \ldots \square (\gamma_m \|\cdot\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha_m,\beta_m,\mathbf{c}_m}) \tag{22}$$

where $\vec{\gamma} = (\gamma_1, \ldots, \gamma_m)$ satisfies $\gamma_1, \ldots, \gamma_m \geq 0$.

One can easily see that each $\gamma_i \|\cdot\|_1 + \mathrm{TGV}^{\mathrm{osci}}_{\alpha_i,\beta_i,\mathbf{c}_i}$ is a lower semi-continuous semi-norm on $L^1(\Omega)$ that is coercive on $\mathrm{BV}(\Omega)$. Its kernel, however, is trivial if $\gamma_i > 0$. Thus, Theorem 1 can be extended to this variant of $\mathrm{ICTGV}^{\mathrm{osci}}$ with the modification that only i for which $\gamma_i = 0$ contribute to the kernel.

Finally, the infimal convolution of oscillation TGV can be used as regularization term for Tikhonov functionals associated with linear inverse problems.

Theorem 2. *Let $1 \leq p < \infty$, $p \leq d/(d-1)$ and $K : L^p(\Omega) \to Y$ a linear and continuous operator mapping to a Banach space Y, $f \in Y$ and $1 \leq q < \infty$. Then, the problem*

$$\min_{u \in L^p(\Omega)} \frac{1}{q} \|Ku - f\|_Y^q + \mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}(u) \tag{23}$$

admits a solution.

Proof. Without loss of generality, one can assume that $p > 1$ since $d/(d-1) > 1$ for all d and $K|_{L^p(\Omega)} : L^p(\Omega) \to Y$ is linear and continuous if $K : L^1(\Omega) \to Y$ is linear and continuous. Then, with $X = L^p(\Omega)$, $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}$ is a lower semi-continuous semi-norm on X as it is a lower semi-continuous semi-norm on $L^1(\Omega)$ and $X \hookrightarrow L^1(\Omega)$. Moreover, it is coercive on X as it is coercive on $\mathrm{BV}(\Omega)$ and $\mathrm{BV}(\Omega) \hookrightarrow X$ (see also Remark 1). Hence, application of Corollary 1 yields the result. $\square$

Remark 3. *It is clear that the theorem is still true if one replaces $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}$ by the variant $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}},\vec{\gamma}}$ according to (22).*

4 Discretization and a numerical optimization algorithm

4.1 A finite-difference discretization

We start with a finite-difference discretization of $\mathrm{TGV}^{\mathrm{osci}}$. Regarding the applications in this paper we just consider the two-dimensional case and rectangular domains. The framework can, however, easily be adapted to higher dimensions and more general domains. Following essentially the presentation in [12, 7, 13], we first replace the domain $\Omega \subset \mathbb{R}^2$ by a discrete rectangular unit-length grid

$$\Omega = \{(i,j) : i,j \in \mathbb{N}, 1 \leq i \leq N_1, 1 \leq j \leq N_2\}$$

where (N_1, N_2), $N_1, N_2 \geq 2$ are the image dimensions. Let us further introduce the vector spaces of functions

$$U = \{u : \Omega \to \mathbb{R}\}$$

and $V = U \times U$, $W = U \times U \times U$. The space U will be endowed with the scalar product $\langle u, u' \rangle_U = \sum_{(i,j) \in \Omega} u_{i,j} u'_{i,j}$ for $u, u' \in U$ and the norm $\|u\|_U = \sqrt{\langle u, u \rangle_U}$. Likewise, the spaces V and W are equipped with the scalar products $\langle v, v' \rangle_V = \langle v_1, v'_1 \rangle_U + \langle v_2, v'_2 \rangle_U$ and $\langle w, w' \rangle_W = \langle w_1, w'_1 \rangle_U + \langle w_2, w'_2 \rangle_U + 2\langle w_3, w'_3 \rangle_U$, respectively.

The oscillation TGV functional will be discretized by finite differences. The forward and backward partial differentiation operators are:

$$(\partial_x^+ u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } 1 \le i < N_1, \\ 0, & \text{if } i = N_1, \end{cases} \qquad (\partial_x^- u)_{i,j} = \begin{cases} u_{i,j} - u_{i-1,j}, & \text{if } 1 < i \le N_1, \\ 0, & \text{if } i = 1, \end{cases}$$

$$(\partial_y^+ u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } 1 \le j < N_2, \\ 0, & \text{if } j = N_2, \end{cases} \qquad (\partial_y^- u)_{i,j} = \begin{cases} u_{i,j} - u_{i,j-1}, & \text{if } 1 < j \le N_2, \\ 0, & \text{if } j = 1, \end{cases}$$

With these, the gradient ∇ and symmetric gradient $\mathcal{E}$ are given as follows:

$$\nabla : U \to V, \quad \nabla u = \begin{pmatrix} \partial_x^+ u \\ \partial_y^+ u \end{pmatrix}, \qquad \mathcal{E} : V \to W, \quad \mathcal{E} v = \begin{pmatrix} \partial_x^- v_1 \\ \partial_y^- v_2 \\ \frac{1}{2}(\partial_y^- v_1 + \partial_x^- v_2) \end{pmatrix}.$$

The divergences of vector fields and symmetric matrix fields are defined as the negative adjoint of the gradient and symmetric gradient, respectively. These can be expressed in terms of partial differentiation operators with suitable boundary conditions:

$$(\partial_x^{*+} u)_{i,j} = \begin{cases} u_{2,j}, & \text{if } i = 1, \\ u_{i+1,j} - u_{i,j}, & \text{if } 1 < i < N_1, \\ -u_{N_1,j}, & \text{if } i = N_1, \end{cases} \qquad (\partial_x^{*-} u)_{i,j} = \begin{cases} u_{1,j}, & \text{if } i = 1, \\ u_{i,j} - u_{i-1,j}, & \text{if } 1 < i < N_1, \\ -u_{N_1-1,j}, & \text{if } i = N_1, \end{cases}$$

$$(\partial_y^{*+} u)_{i,j} = \begin{cases} u_{i,2}, & \text{if } j = 1, \\ u_{i,j+1} - u_{i,j}, & \text{if } 1 < j < N_2, \\ -u_{i,N_2}, & \text{if } j = N_2. \end{cases} \qquad (\partial_y^{*-} u)_{i,j} = \begin{cases} u_{i,1}, & \text{if } j = 1, \\ u_{i,j} - u_{i,j-1}, & \text{if } 1 < j < N_2, \\ -u_{i,N_2-1}, & \text{if } j = N_2, \end{cases}$$

such that

$$\text{div}_1 : V \to U, \quad \text{div}\, v = \partial_x^{*-} v_1 + \partial_y^{*-} v_2, \qquad \text{div}_2 : W \to V, \quad \text{div}_2 w = \begin{pmatrix} \partial_x^{*+} w_1 + \partial_y^{*+} w_3 \\ \partial_x^{*+} w_3 + \partial_y^{*+} w_2 \end{pmatrix}.$$

By direct computation, one sees that indeed, $\nabla^* = -\text{div}_1$ as well as $\mathcal{E}^* = -\text{div}_2$. A discrete definition of oscillation TGV can now either be derived from the primal representation (4) or the dual representation (10). In any case, we need the notion of the 1 and ∞ norm in the discrete spaces:

$$u \in U: \quad \|u\|_1 = \sum_{(i,j) \in \Omega} |u_{i,j}|, \qquad\qquad\qquad \|u\|_\infty = \max_{(i,j) \in \Omega} |u_{i,j}|,$$

$$v \in V: \quad \|v\|_1 = \sum_{(i,j) \in \Omega} \left((v_1)_{i,j}^2 + (v_2)_{i,j}^2\right)^{1/2}, \qquad \|v\|_\infty = \max_{(i,j) \in \Omega} \left((v_1)_{i,j}^2 + (v_2)_{i,j}^2\right)^{1/2},$$

$$w \in W: \quad \|w\|_1 = \sum_{(i,j) \in \Omega} \left((w_1)_{i,j}^2 + (w_2)_{i,j}^2 + 2(w_3)_{i,j}^2\right)^{1/2},$$

$$\|w\|_\infty = \max_{(i,j) \in \Omega} \left((w_1)_{i,j}^2 + (w_2)_{i,j}^2 + 2(w_3)_{i,j}^2\right)^{1/2}.$$

With these prerequisites, the discrete functional associated with $\alpha > 0$, $\beta > 0$ and $\mathbf{c} = \begin{bmatrix} c_1 & c_3 \\ c_3 & c_2 \end{bmatrix}$ reads as follows

$$\text{TGV}_{\alpha,\beta,\mathbf{c}}^{\text{osci}}(u) = \min_{v \in V} \ \alpha \|\nabla u - v\|_1 + \beta \|\mathcal{E} v + \mathbf{c} u\|_1$$

$$= \max_{p \in V, \ q \in W} \ \langle u, \text{div}_1 p + \mathbf{c} \cdot q \rangle_U \quad \text{subject to} \quad \begin{cases} \|p\|_\infty \le \alpha, \|q\|_\infty \le \beta, \\ p = \text{div}_2 q. \end{cases} \tag{24}$$

Indeed, one can easily verify that Fenchel–Rockafellar duality holds, for instance, by continuity of $v \mapsto \alpha \|\nabla u - v\|_1 + \beta \|\mathcal{E} v + \mathbf{c} u\|_1$ in the finite-dimensional space V, and that the minimization and maximization problems in (24) are dual to each other.

4.2 Discrete kernel representation

We did not detail yet on the choice of $\mathbf{c}$. As our motivation for oscillation TGV was to capture oscillations of the type (7) by adapting the kernel of the regularizer exactly to these functions, our aim is to transfer this property to the discrete setting. However, a straightforward adaptation of the choice of $\mathbf{c}$ in the continuous setting, i.e., $\mathbf{c} = \omega \otimes \omega$ for $\omega \in \mathbb{R}^2$, $\omega \neq 0$ will lead to a non-trivial kernel only in exceptional situations. The reason for this is that in the continuous setting, elements in the kernel satisfy certain differential equations with entries of $\mathbf{c}$ as coefficients (see the proof of Lemma 1). These become difference equations in the discrete setting whose solutions generally do not coincide with discrete versions of the continuous solutions if one does not adapt the coefficients in $\mathbf{c}$. This induces a mismatch that typically leads to a trivial kernel; see the derivation below for more details.

Hence, we have to choose the coefficients of $\mathbf{c}$ in order to fit the above discretization of the differential operators. For this purpose, we examine the kernel of $\mathrm{TGV}^{\mathrm{osci}}_{\alpha,\beta,\mathbf{c}}$ according to (24) for a $\mathbf{c} = \begin{bmatrix} c_1 & c_3 \\ c_3 & c_2 \end{bmatrix}$ with $c_1, c_2, c_3 \in \mathbb{R}$ such that $\mathbf{c} \neq 0$. Then, one immediately sees that the kernel is given by all $u \in U$ that satisfy $\mathcal{E}\nabla u + \mathbf{c}u = 0$ in Ω. This corresponds to

$$\begin{cases} \partial_x^- \partial_x^+ u + c_1 u = 0, \\ \partial_y^- \partial_y^+ u + c_2 u = 0, \\ \dfrac{\partial_y^- \partial_x^+ + \partial_x^- \partial_y^+}{2} u + c_3 u = 0. \end{cases} \tag{25}$$

For the sake of simplicity we focus, in the following, on the "interior points" of Ω and neglect boundary effects. As a consequence, Equation (25) might not be satisfied at discrete boundary points. This leads to

$$\begin{cases} u_{i+1,j} - 2u_{i,j} + u_{i-1,j} + c_1 u_{i,j} = 0, \\ u_{i,j+1} - 2u_{i,j} + u_{i,j-1} + c_2 u_{i,j} = 0, \\ u_{i+1,j} - u_{i,j} - u_{i+1,j-1} + u_{i,j-1} + u_{i-1,j} - u_{i,j} - u_{i-1,j+1} + u_{i,j+1} + 2c_3 u_{i,j} = 0, \end{cases} \tag{26}$$

for all $(i,j) \in \mathbb{Z}^2$ where u is defined on all discrete points, i.e., $u : \mathbb{Z}^2 \to \mathbb{R}$.

We would like to achieve that for $\omega \in \mathbb{R}^2$, u satisfies (26) if and only if there exist constants $C_1, C_2 \in \mathbb{R}$ such that for each (i,j) it holds that

$$u_{i,j} = C_1 \sin(\omega_1 i + \omega_2 j) + C_2 \cos(\omega_1 i + \omega_2 j). \tag{27}$$

For that purpose let us, in analogy to the presentation in Section 2, derive the solution space for (26). Taking into account the first equation of (26), which can be regarded as one-dimensional if we fix the subscript j, and rearranging leads to the recurrence relation

$$u_{i+1,j} = (2 - c_1)u_{i,j} - u_{i-1,j} \tag{28}$$

whose solutions can be described as follows. Denoting by $\lambda \in \mathbb{C}$ a root of $t \mapsto t^2 - (2 - c_1)t + 1$, it holds that

$$u_{i,j} = \begin{cases} C_{1,j}\lambda^i + C_{2,j}\overline{\lambda}^i & \text{if } \lambda \text{ is a simple root,} \\ C_{1,j}\lambda^i + C_{2,j}i\lambda^i & \text{else,} \end{cases} \tag{29}$$

for suitable constants $C_{1,j}, C_{2,j} \in \mathbb{C}$. Without loss of generality, λ is given by

$$\lambda = \frac{2 - c_1 + \sqrt{c_1^2 - 4c_1}}{2}$$

where $\sqrt{\cdot}$ is a complex square root that is defined on the real axis. Let us first assume that λ is simple such that $\overline{\lambda}$ is the other root. In view of the Euler formulas $\sin(t) = \frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}t} - \mathrm{e}^{-\mathrm{i}t})$,

$\cos(t) = \frac{1}{2}(e^{it} + e^{-it})$, the representation (27) can only be achieved if $|\lambda| = 1$. This in turn, is the case if and only if $c_1^2 - 4c_1 < 0$, which is equivalent to $0 < c_1 < 4$. Thus, in order to satisfy (27), the equation

$$\lambda = \frac{2 - c_1 + i\sqrt{4c_1 - c_1^2}}{2} = e^{i\omega_1} = \cos(\omega_1) + i\sin(\omega_1) \tag{30}$$

has to hold for some complex root, which means that

$$c_1 = 2 - 2\cos(\omega_1). \tag{31}$$

In particular, this can only be true if $\omega_1 \notin \pi\mathbb{Z}$, in which case we can express u by

$$u_{i,j} = D_{1,j}\cos(\omega_1 i) + D_{2,j}\sin(\omega_1 i) \quad \text{for all} \quad i, j \tag{32}$$

where $D_{1,j}, D_{2,j} \in \mathbb{R}$ for each j. Fixing, for instance, $i = 1, 2$ and noting that the matrix $\begin{bmatrix} \cos(\omega_1) & \sin(\omega_1) \\ \cos(2\omega_1) & \sin(2\omega_1) \end{bmatrix}$ has full rank, the second equation in (26) leads to the recurrence relation $D_{i,j+1} = (2 - c_2)D_{i,j} - D_{i,j-1}$ for $i = 1, 2$ and all j. Analogous to the above we get, in case of $0 < c_2 < 4$ that $c_2 = 2 - 2\cos(\omega_2)$, $\omega_2 \notin \pi\mathbb{Z}$ and

$$D_{i,j} = A_i\cos(\omega_2 j) + B_i\sin(\omega_2 j) \quad \text{for} \quad i = 1, 2 \tag{33}$$

and all j. Plugging this into (32) yields, by virtue of the angle sum and difference identities,

$$u_{i,j} = A_1\cos(\omega_1 i)\cos(\omega_2 j) + B_1\cos(\omega_1 i)\sin(\omega_2 j) + A_2\sin(\omega_1 i)\cos(\omega_2 j) + B_2\sin(\omega_1 i)\sin(\omega_2 j)$$
$$= T_1\cos(\omega_1 i + \omega_2 j) + T_2\cos(\omega_1 i - \omega_2 j) + T_3\sin(\omega_1 i + \omega_2 j) + T_4\sin(\omega_1 i - \omega_2 j)$$

with $T_1 = \frac{A_1 - B_2}{2}$, $T_2 = \frac{A_1 + B_2}{2}$, $T_3 = \frac{A_2 + B_1}{2}$ and $T_4 = \frac{A_2 - B_1}{2}$. If one plugs the first two equations of (26) into the third equation, one gets

$$u_{i+1,j-1} + u_{i-1,j+1} = (2 - c_1 - c_2 + 2c_3)u_{ij} \tag{34}$$

which corresponds to, using the above as well as $c_1 = 2 - 2\cos(\omega_1)$, $c_2 = 2 - 2\cos(\omega_2)$,

$$2\cos(\omega_1 - \omega_2)\big(T_1\cos(\omega_1 i + \omega_2 j) + T_3\sin(\omega_1 i + \omega_2 j)\big)$$
$$+ 2\cos(\omega_1 + \omega_2)\big(T_2\cos(\omega_1 i - \omega_2 j) + T_4\sin(\omega_1 i - \omega_2 j)\big)$$
$$= 2(\cos(\omega_1) + \cos(\omega_2) + c_3 - 1)\big(T_1\cos(\omega_1 i + \omega_2 j) + T_2\cos(\omega_1 i - \omega_2 j)$$
$$+ T_3\sin(\omega_1 i + \omega_2 j) + T_4\sin(\omega_1 i - \omega_2 j)\big).$$

As $\omega_1, \omega_2 \notin \pi\mathbb{Z}$, we have $\cos(\omega_1 - \omega_2) \neq \cos(\omega_1 + \omega_2)$. Thus, setting $c_3 = 1 + \cos(\omega_1 - \omega_2) - \cos(\omega_1) - \cos(\omega_2)$ implies that the equation can only be true for all i, j if $T_2 = T_4 = 0$. But this means that u has the desired representation (27). Note that here, the above choice of c_3 is crucial as any other choice (expect $c_3 = 1 + \cos(\omega_1 + \omega_2) - \cos(\omega_1) - \cos(\omega_2)$) implies a trivial kernel. As mentioned above, this is what typically happens without an adaptation of $\mathbf{c}$ to the discrete setting.

We still have to discuss the remaining cases. Starting with $c_2 \in \{0, 4\}$ which corresponds, in view of $c_2 = 2 - 2\cos(\omega_2)$, to $\omega_2 \in \pi\mathbb{Z}$, such that the root in (29) is not simple and the representation (33) becomes

$$D_{i,j} = (A_i + B_i j)\cos(\omega_2 j) \quad \text{for} \quad i = 1, 2. \tag{35}$$

Now, since $\cos(\omega_2) \in \{-1, 1\}$, it follows that $\cos(\omega_1 i)\cos(\omega_2 j) = \cos(\omega_1 i + \omega_2 j)$ as well as $\sin(\omega_1 i)\cos(\omega_2 j) = \sin(\omega_1 i + \omega_2 j)$ such that

$$u_{i,j} = (A_1 + B_1 j)\cos(\omega_1 i + \omega_2 j) + (A_2 + B_2 j)\sin(\omega_1 i + \omega_2 j).$$

Employing (34) then leads to

$$2\cos(\omega_1-\omega_2)\big((A_1+B_1j)\cos(\omega_1i+\omega_2j)+(A_2+B_2j)\sin(\omega_1i+\omega_2j)\big)$$
$$+2\sin(\omega_1-\omega_2)\big(B_1\sin(\omega_1i+\omega_2j)-B_2\cos(\omega_1i+\omega_2j)\big)$$
$$=2(\cos(\omega_1)+\cos(\omega_2)+c_3-1)\big((A_1+B_1j)\cos(\omega_1i+\omega_2j)+(A_2+B_2j)\sin(\omega_1i+\omega_2j)\big),$$

such that setting again $c_3=1+\cos(\omega_1-\omega_2)-\cos(\omega_1)-\cos(\omega_2)$, this can only be true for all i,j if $B_1=B_2=0$ since $\sin(\omega_1-\omega_2)\neq 0$ as a consequence of $\omega_1\notin\pi\mathbb{Z}$, $\omega_2\in\pi\mathbb{Z}$. But this means that u has the desired representation (27). If we interchange the i- and j-axis, the same reasoning applies in case of $c_1\in\{0,4\}$ and $0<c_2<4$, or, equivalently, $\omega_1\in\pi\mathbb{Z}$ and $\omega_2\notin\pi\mathbb{Z}$.

Finally, if $c_1,c_2\in\{0,4\}$ which corresponds to $\omega_1,\omega_2\in\pi\mathbb{Z}$, the identity (32) becomes

$$u_{i,j}=(D_{1,j}+D_{2,j}i)\cos(\omega_1i)\quad\text{for all}\quad i,j.$$

Proceeding as above leads to (35) and, consequently,

$$u_{i,j}=(A_1+B_1j+A_2i+B_2ij)\cos(\omega_1i+\omega_2j).$$

Employing (34) again yields

$$2\cos(\omega_1-\omega_2)\big(A_1+B_1j+A_2i+B_2(ij-1)\big)\cos(\omega_1i+\omega_2j)$$
$$=2(\cos(\omega_1)+\cos(\omega_2)+c_3-1)\big(A_1+B_1j+A_2i+B_2ij\big)\cos(\omega_1i+\omega_2j).$$

With $c_3=1+\cos(\omega_1-\omega_2)-\cos(\omega_1)-\cos(\omega_2)$ and considering all i,j, we can, however, only deduce that $B_2=0$, such that the representation (27) is not obtainable in this situation. Thus, we have to assume that $(\omega_1,\omega_2)\notin\pi\mathbb{Z}^2$.

In order to be complete let us eventually note that if $c_1\notin[0,4]$ or $c_2\notin[0,4]$, the root λ in (29) is simple and satisfies $|\lambda|\neq 1$ such that (27) can also not be obtained. In summary, we have proved the following proposition:

Proposition 7. *Let $\omega=(\omega_1,\omega_2)\in\mathbb{R}^2\setminus\pi\mathbb{Z}^2$ and*

$$\mathbf{c}=\begin{bmatrix}2-2\cos(\omega_1) & 1+\cos(\omega_1-\omega_2)-\cos(\omega_1)-\cos(\omega_2)\\ 1+\cos(\omega_1-\omega_2)-\cos(\omega_1)-\cos(\omega_2) & 2-2\cos(\omega_2)\end{bmatrix}. \tag{36}$$

Then, $u:\mathbb{Z}^2\to\mathbb{R}$ satisfies $\mathcal{E}\nabla u+\mathbf{c}u=0$ in the sense that (26) holds for all $(i,j)\in\mathbb{Z}^2$ if and only if there are constants $C_1,C_2\in\mathbb{R}$ such that

$$u_{i,j}=C_1\cos(\omega_1i+\omega_2j)+C_2\sin(\omega_1i+\omega_2j)\qquad\text{for all}\qquad(i,j)\in\mathbb{Z}^2.$$

In particular, given a pair $\omega=(\omega_1,\omega_2)\in\mathbb{R}^2\setminus\pi\mathbb{Z}^2$, we can represent texture with oscillations in the corresponding direction and frequency by the discrete functional (24) and $\mathbf{c}$ according to (36). For $(\omega_1,\omega_2)=0$, Equation (36) yields $\mathbf{c}=0$, such that (24) corresponds to a discrete $\mathrm{TGV}^2_{\alpha,\beta}$. Thus, a discrete infimal convolution of oscillation TGV reads as follows:

$$\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha},\vec{\beta},\vec{\mathbf{c}}}(u)=\min_{\substack{u_1,\dots,u_m\in U,\\ w_1,\dots,w_m\in V,\\ u=u_1+\dots+u_m}}\sum_{i=1}^{m}\alpha_i\|\nabla u_i-w_i\|_1+\beta_i\|\mathcal{E}w_i+\mathbf{c}_iu_i\|_1$$

$$=\max_{\substack{r\in U,\ p_1,\dots,p_m\in V,\\ q_1,\dots,q_m\in W}}\langle u,r\rangle\quad\text{subject to}\quad\begin{cases}\|p_i\|_\infty\leq\alpha_i,\|q_i\|_\infty\leq\beta_i,\\ \quad p_i=\mathrm{div}_2\,q_i,\\ \quad r=\mathrm{div}_1\,p_i+\mathbf{c}_i\cdot q_i,\end{cases} \tag{37}$$

where $\vec{\alpha}=(\alpha_1,\dots,\alpha_m)$, $\vec{\beta}=(\beta_1,\dots,\beta_m)$ are vectors of positive parameters and $\vec{\mathbf{c}}=(\mathbf{c}_1,\dots,\mathbf{c}_m)$ are given by (36) for pairs $(\omega_{1,1},\omega_{1,2}),\dots,(\omega_{m,1},\omega_{m,2})\in\mathbb{R}^2\setminus\pi\mathbb{Z}^2\cup\{0\}$.

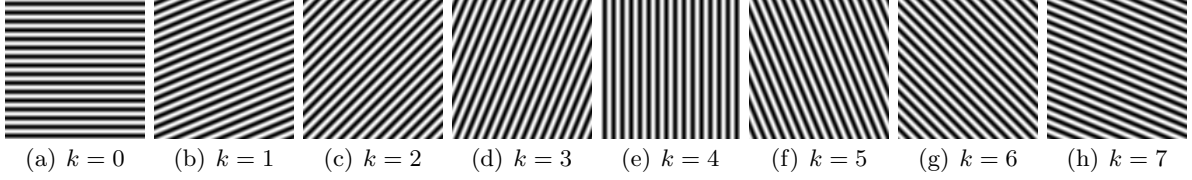| (a) $k=0$ | (b) $k=1$ | (c) $k=2$ | (d) $k=3$ | (e) $k=4$ | (f) $k=5$ | (g) $k=6$ | (h) $k=7$ |

Figure 2: Elements from the kernel of discrete oscillation TGV according to (24) for the choice $\omega_1 = \sin(\frac{k\pi}{8})$, $\omega_2 = \cos(\frac{k\pi}{8})$ and $\mathbf{c}$ according to (38) and constants $C_1 = C_2 = 1$.

Remark 4. *The direction and frequency parameters in* (36) *can, for instance, be chosen as follows. Suppose, we would like to detect eight different texture directions with unit frequency as well as a cartoon part, i.e., $m = 9$. The first part in the infimal-convolution model is the cartoon component and therefore, $\omega_{1,1} = \omega_{1,2} = 0$, leading to $\mathbf{c}_1 = 0$. Further, we choose eight pairs (ω_1, ω_2) to uniformly cover the (unsigned) unit directions:*

$$\begin{cases} \omega_{i,1} = \sin(\frac{(i-2)\pi}{8}), \\ \omega_{i,2} = \cos(\frac{(i-2)\pi}{8}), \end{cases} \quad for \quad i = 2, 3, \ldots, 9. \tag{38}$$

If one aims at obtaining textures of higher frequency, one can, for instance, either replace $(\omega_{i,1}, \omega_{i,2})$ by $(2\omega_{i,1}, 2\omega_{i,2})$ or add the latter pairs to the model, resulting in $m = 17$, i.e., one cartoon component, eight directions of low frequency and eight directions of high frequency. In our experiments, we utilized either the simple model ($m = 9$) or the extended model ($m = 17$).

See Figure 2 for a visualization of the kernels u associated with $\mathrm{TGV}^{\mathrm{osci}}_{\alpha_i, \beta_i, \mathbf{c}_i}$ for the above choice.

4.3 A numerical optimization algorithm

Now, we focus on numerical algorithms for solving the Tikhonov minimization problem (23) for $q = 2$ which is associated with $\mathrm{ICTGV}^{\mathrm{osci}}$-regularization of the inverse problem $Ku = f$. We assume that the linear operator $K : U \to Y$, where Y is a Hilbert space is already given in a discretized form. With the representation (37), problem (23) reads as

$$\min_{\substack{u_1, \ldots, u_m \in U, \\ w_1, \ldots, w_m \in V}} \frac{1}{2} \|K \sum_{i=1}^{m} u_i - f\|^2 + \sum_{i=1}^{m} \alpha_i \|\nabla u_i - w_i\|_1 + \beta_i \|\mathcal{E} w_i + \mathbf{c}_i u_i\|_1.$$

For the image processing problems we are considering, we also use a discrete version of the $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{\mathbf{c}}, \vec{\gamma}}$-functionals according to (22) with additional parameters $\vec{\gamma} = (\gamma_1, \ldots, \gamma_m)$ where $\gamma_i \geq 0$ for each $i = 1, \ldots, m$. This is might in particular be beneficial for texture detection in the presence of noise as it enforces sparsity of the respective texture component and prevents noise being recognized as texture. The modification then leads to the following minimization problem:

$$\min_{\substack{u_1, \ldots, u_m \in U, \\ w_1, \ldots, w_m \in V}} \frac{1}{2} \|K \sum_{i=1}^{m} u_i - f\|^2 + \sum_{i=1}^{m} \alpha_i \|\nabla u_i - w_i\|_1 + \beta_i \|\mathcal{E} w_i + \mathbf{c}_i u_i\|_1 + \gamma_i \|u_i\|_1, \tag{39}$$

where we set $\gamma_i = 0$ if $\mathbf{c} = 0$ in order to prevent a sparsification of the cartoon component (recall that then, $\mathrm{TGV}^{\mathrm{osci}}_{\alpha, \beta, \mathbf{c}} = \mathrm{TGV}^2_{\alpha, \beta}$). Our aim is to employ primal-dual algorithms which operate on equivalent saddle-point formulations of the minimization problem. For (39), one can easily see that Fenchel–Rockafellar duality is applicable (see, e.g., [26]), such that primal-dual solutions

are equivalent to solutions of the saddle-point problem

$$
\min_{\substack{u_1,\ldots,u_m\in U,\\ w_1,\ldots,w_m\in V}} \max_{\substack{p_1,\ldots,p_m\in V,\\ q_1,\ldots,q_m\in W,\\ \lambda\in Y}} \left(\langle K\sum_{i=1}^m u_i, \lambda\rangle + \sum_{i=1}^m \langle \nabla u_i - w_i, p_i\rangle + \langle \mathcal{E} w_i + \mathbf{c}_i u_i, q_i\rangle \right)
$$

$$
+ \left(\sum_{i=1}^m \gamma_i \|u_i\|_1 \right) - \left(\langle f, \lambda\rangle + \frac{\|\lambda\|^2}{2} + \sum_{i=1}^m \mathcal{I}_{\{\|\cdot\|_\infty \le \alpha_i\}}(p_i) + \mathcal{I}_{\{\|\cdot\|_\infty \le \beta_i\}}(q_i) \right). \tag{40}
$$

This reformulation allows to use, e.g., the primal-dual method in [24], which we will shortly describe in the following. It solves abstract convex-concave saddle-point problems of the form

$$
\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \langle \mathcal{K}x, y\rangle + G(x) - F^*(y), \tag{41}
$$

where $\mathcal{X}, \mathcal{Y}$ are Hilbert spaces, $\mathcal{K} : \mathcal{X} \to \mathcal{Y}$ is a continuous linear mapping, and the functionals $G : \mathcal{X} \to \,]-\infty, \infty]$ and $F^* : \mathcal{Y} \to \,]-\infty, \infty]$ are proper, convex and lower semi-continuous. The problem (41) is associated to the Fenchel–Rockafellar primal-dual problems

$$
\min_{x\in\mathcal{X}} F(\mathcal{K}x) + G(x), \qquad \max_{y\in\mathcal{Y}} -F^*(y) - G^*(-\mathcal{K}^*y). \tag{42}
$$

In order to state the algorithm clearly, we have to give the notion of resolvent operators $(I + \tau\partial G)^{-1}$ and $(I + \sigma\partial F^*)^{-1}$, respectively, which correspond to the solution operators of certain minimization problems, the so-called proximal operators:

$$
x^* = (I + \tau\partial G)^{-1}(\bar{x}) = \arg\min_{x\in\mathcal{X}} \frac{\|x - \bar{x}\|_\mathcal{X}^2}{2} + \tau G(x),
$$

$$
y^* = (I + \sigma\partial F^*)^{-1}(\bar{y}) = \arg\min_{y\in\mathcal{Y}} \frac{\|y - \bar{y}\|_\mathcal{Y}^2}{2} + \sigma F^*(y)
$$

where $\tau, \sigma > 0$ are step-size parameters we need to choose suitably. Given the initial point $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$ and setting $\bar{x}^0 = x^0$, the iterative procedure in [24] can be written as follows:

$$
\begin{cases}
y^{n+1} = (I + \sigma\partial F^*)^{-1}(y^n + \sigma\mathcal{K}\bar{x}^n), \\
x^{n+1} = (I + \tau\partial G)^{-1}(x^n - \tau\mathcal{K}^*y^{n+1}), \\
\bar{x}^{n+1} = 2x^{n+1} - x^n.
\end{cases} \tag{43}
$$

It converges if the condition $\tau\sigma\|\mathcal{K}\|^2 < 1$ is satisfied which corresponds to choosing appropriate values for $\tau, \sigma > 0$.

Next, we will delineate the iteration (43) adapted to our problem (40) which can be reformulated into above structure by redefining its variables and operators as:

$$
x = (u_1, w_1, \ldots, u_m, w_m) \in \mathcal{X} = \underbrace{(U \times V) \times \cdots \times (U \times V)}_{m\text{-times}},
$$

$$
y = (p_1, q_1, \ldots, p_m, q_m, \lambda) \in \mathcal{Y} = \underbrace{(V \times W) \times \cdots \times (V \times W)}_{m\text{-times}} \times Y,
$$

$$
\mathcal{K} = \begin{pmatrix}
\begin{pmatrix} \nabla & -I \\ \mathbf{c}_1 & \mathcal{E} \end{pmatrix} & 0 & \cdots & 0 \\
0 & \begin{pmatrix} \nabla & -I \\ \mathbf{c}_2 & \mathcal{E} \end{pmatrix} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \begin{pmatrix} \nabla & -I \\ \mathbf{c}_m & \mathcal{E} \end{pmatrix} \\
\begin{pmatrix} K & 0 \end{pmatrix} & \begin{pmatrix} K & 0 \end{pmatrix} & \cdots & \begin{pmatrix} K & 0 \end{pmatrix}
\end{pmatrix},
$$

Algorithm 1 Primal-Dual Method for infimal convolution oscillation TGV

Initialization: Choose τ, σ and x^0, y^0, set $\bar{x}^0 = x^0$;

Iteration: Update according to the following steps

$$\lambda^{n+1} = \left(\lambda^n + \sigma(K \sum_{i=1}^{m} \bar{u}_i^n - f)\right)/(1+\sigma);$$

for $i = 1, \ldots, m$ **do**

$\quad p_i^{n+1} = \mathcal{P}_{\alpha_i}\left(p_i^n + \sigma(\nabla \bar{u}_i^n - \bar{w}_i^n)\right);$

$\quad q_i^{n+1} = \mathcal{P}_{\beta_i}\left(q_i^n + \sigma(\mathcal{E}\bar{w}_i^n + \mathbf{c}_i\bar{u}_i^n)\right);$

$\quad \tilde{u}_i^{n+1} = u_i^n - \tau(K^*\lambda^{n+1} - \mathrm{div}_1\, p_i^{n+1} + \mathbf{c}_i q_i^{n+1});$

$\quad u_i^{n+1} = \mathrm{Shrink}_{\tau\gamma_i}(\tilde{u}_i^{n+1});$

$\quad w_i^{n+1} = w_i^n + \tau(p_i^{n+1} + \mathrm{div}_2\, q_i^{n+1});$

$\quad \bar{u}_i^{n+1} = 2u_i^{n+1} - u_i^n;$

$\quad \bar{w}_i^{n+1} = 2w_i^{n+1} - w_i^n;$

end for

If converged: Return $u = \sum_{i=1}^{m} u_i^n;$

---

as well as

$$G(x) = \sum_{i=1}^{m} \gamma_i \|u_i\|_1,$$

$$F^*(y) = \langle f, \lambda \rangle + \frac{\|\lambda\|^2}{2} + \sum_{i=1}^{m} \mathcal{I}_{\{\|\cdot\|_\infty \le \alpha_i\}}(p_i) + \mathcal{I}_{\{\|\cdot\|_\infty \le \beta_i\}}(q_i).$$

In order to guarantee the convergence condition $\tau\sigma\|\mathcal{K}\|^2 < 1$, we have to estimate the norm of the operator $\mathcal{K}$. At first, if $\mathbf{c} = \begin{bmatrix} c_1 & c_3 \\ c_3 & c_2 \end{bmatrix}$, then

$$\|\mathbf{c}\|^2 = \frac{c_1^2}{2} + \frac{c_2^2}{2} + c_3^2 + \frac{c_1 + c_2}{2}\sqrt{(c_1 - c_2)^2 + 4c_3^2}.$$

If we denote by $c^2 = \max\{\|\mathbf{c}_1\|^2, \ldots, \|\mathbf{c}_m\|^2\}$ and observe that $\|\nabla\|^2 < 8$ and $\|\mathcal{E}\|^2 < 8$, then, after some computations, one obtains the estimate

$$\|\mathcal{K}\|^2 < \frac{c^2 + (c+1)\sqrt{(c-1)^2 + 32} + 17}{2} + m\|K\|^2 = C_0(c, m, K),$$

thus, in order to ensure convergence, it suffices to choose $\tau, \sigma > 0$ such that $\sigma\tau = \frac{1}{C_0(c,m,K)}$.

By the above analysis, the primal-dual method for solving the saddle-point formulation (40) of the imaging problem (39) reads as Algorithm 1. Note that in the algorithm, the corresponding proximal operators are already written in an explicit form (see, for instance, [7] for more details). In particular, the expressions in the updates of p_i^{n+1}, q_i^{n+1} and u_i^{n+1} read as follows:

$$\mathcal{P}_\eta(t) = \min\left(1, \frac{\eta}{|t|}\right)t, \qquad \mathrm{Shrink}_\eta(t) = \max\left(0, 1 - \frac{\eta}{|t|}\right)t.$$

5 Applications and numerical experiments

In this section, we discuss applications of infimal convolution of oscillation TGV regularization to several image processing problems as well as associated numerical experiments including comparisons to some existing state-of-the-art methods for the respective problems. The applications cover cartoon/texture decomposition, image denoising, image inpainting and undersampled magnetic resonance imaging (MRI).

In particular, we assess the reconstruction results in terms of the peak signal to noise ratio (PSNR) and structural similarity index (SSIM). All parameters in the proposed and reference methods were manually tuned to give the best results in terms of these measures; we report them for the proposed methods in the figure captions of each case. The manual tuning of the parameters was performed, for each method, on a reasonably exhaustive set of parameters and was based on our experience as well as the recommendations of the authors of the respective method; we refer to the supplementary material Section D for additional details on the implementations and parameter choice. All experiments were performed in MATLAB R2016a running on a computer with Intel Core i5 CPU at 3.30 GHz and 8 GB of memory. For some implementations, we utilized the "gpuArray" feature of MATLAB's parallelization toolbox to achieve an embarrasing parallel implementation of the code for which the numerical experiments were also run on a NVIDIA GTX 1070 GPU. In all of the examples, the range of image values is $[0, 1]$. Note also that in all experiments, the iteration number was set to a fixed value of 2000, a number that was experimentally determined to be sufficient for convergence of the algorithm. Moreover, unless stated otherwise, direction and frequency parameters are chosen according to Remark 4, i.e., for $m = 9$, there are eight equispaced texture directions and for $m = 17$, there are eight equispaced texture directions and for each direction, there is a low and high frequency component. Also note that we did not perform an estimation of the directions and frequencies from the data.

5.1 Cartoon/texture decomposition and image denoising

The proposed regularization can be applied to both cartoon/texture decomposition as well as image denoising by setting $K = I$ in the Tikhonov minimization problem (23) as well as $Y = L^2(\Omega)$ and $p = q = 2$, i.e.,

$$\min_{u \in L^2(\Omega)} \quad \frac{1}{2}\|u - f\|_2^2 + \mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}}(u). \tag{44}$$

In the case of cartoon/texture decomposition, we can use $\mathrm{ICTGV}^{\mathrm{osci}}$ according to (15), i.e., setting $\gamma_i = 0$ for all i in the algorithm. In contrast, in the case of noise, which can be regarded as small-scale textures, the noise will influence the texture detection and lead to each texture component having undesired ghost stripes. In order to avoid this effect, we use the variant with added L^1-norm according to (22) in order to sparsify the texture which leads to replacing $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}}$ by $\mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}, \vec{\gamma}}$ in (44).

We performed several numerical experiments in order to study the effectiveness as well as benefits and potential weaknesses of this approach. In our first experiment, we show the performance of the new regularization for a synthetic image with one piecewise affine region and four directional textures corresponding to the angles 0, $\frac{\pi}{4}$, $\frac{\pi}{2}$ and $\frac{3\pi}{4}$. Choosing a $\mathrm{ICTGV}^{\mathrm{osci}}$ functional as in Remark 4 but with these four directions instead of eight (leading to $m = 5$) gives a decomposition as depicted in Figure 3. One can see that the infimal convolution of oscillation TGV models are almost perfect in decomposing the image into the respective cartoon and texture parts. Moreover, Figure 3 shows denoising results for the same synthetic image which are also almost perfect. Of course, the reason for this behavior is that the synthetic ground truth corresponds, in the respective regions, exactly to the kernels of the $\mathrm{TGV}^{\mathrm{osci}}$-functionals utilized in the model; leading to a low value of the $\mathrm{ICTGV}^{\mathrm{osci}}$-functional and a stable minimum of the objective functional.

In Figure 4 we show the decomposition/denoising performance for a synthetic image with cartoon part and some textured regions in comparison to the TV-G-norm model [4], TV-H^{-1} model [43] and the framelet/local discrete cosine transform (Fra+LDCT) model [21]. We performed experiments both for a noise-free and noisy image (with Gaussian noise of variance $\sigma = 0.05$). In the results, one can see that some cartoon components (such as edges) and undesired information (such as noise), appear in the texture parts of the TV-G-norm and TV-H^{-1} model which
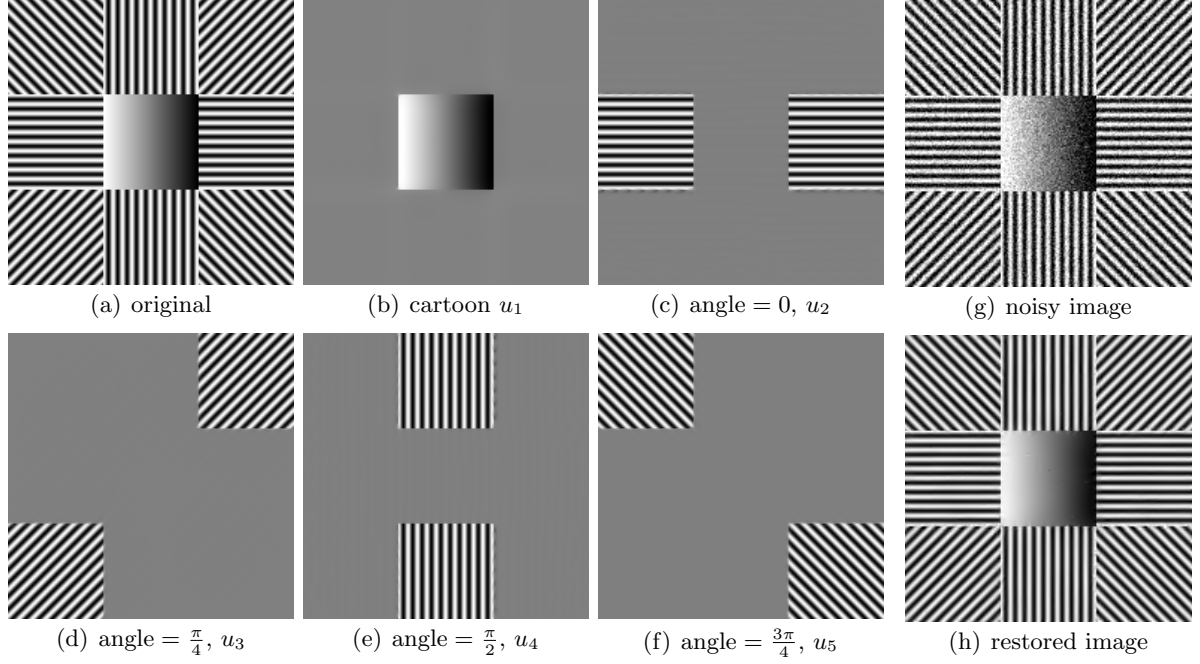
(a) original   (b) cartoon u_1 (c) angle $= 0$, u_2 (g) noisy image

(d) angle $= \frac{\pi}{4}$, u_3 (e) angle $= \frac{\pi}{2}$, u_4 (f) angle $= \frac{3\pi}{4}$, u_5 (h) restored image

Figure 3: Decomposition/denoising of a synthetic image. (a) Original image of dimensions 192×192; (b) Cartoon component of the image; (c)–(f) The four directional texture components. Parameter choice: $\alpha_1 = 0.06, \beta_1 = 2\alpha_1, \gamma_1 = 0, \alpha_i = 0.5\alpha_1, \beta_i = 2\alpha_i, \gamma_i = 0, i = 2, \ldots, 5$. (g) Noisy image corrupted by Gaussian noise with zero mean and standard deviation $\sigma = 0.1$; (h) Image restored by model (44). Parameter choice: $\alpha_1 = 0.12, \beta_1 = 2\alpha_1, \gamma_1 = 0, \alpha_i = 0.5\alpha_1, \beta_i = 2\alpha_i, \gamma_i = 0.4\alpha_i, i = 2, \ldots, 5$.

do not appear in the Fra+LDCT and the proposed model. Comparing the cartoon components of all models, one can observe staircasing effects in the TV-G-norm, TV-H^{-1} and Fra+LDCT model as well as some texture artifacts in the TV-H^{-1} model. Due to the utilization of TGV for the cartoon component, these artifacts do not appear in the ICTGV$^{\text{osci}}$ model. Focusing on the texture component, both Fra+LDCT and the proposed model yield high-quality results. In the Fra+LDCT results, however, one can, for instance, see some low-frequency components in the trousers that are attributable to a shading and hence, should belong to the cartoon part. In contrast, the proposed ICTGV$^{\text{osci}}$ model does correctly assign these shadings to the cartoon part. Moreover, the boundaries of the texture seem to be better delineated in the proposed model compared to Fra+LDCT decomposition which shows some spill-over effects in the texture part. In summary, one can observe that the ICTGV$^{\text{osci}}$ decomposition/denoising model is effective in separating cartoon, texture and noise.

We also performed numerical experiments for the denoising of three natural images ("barbara", "zebra" and "parrots") and two different noise levels (Gaussian noise with zero mean and standard deviation $\sigma = 0.05$ and 0.1, respectively). The "barbara" and "zebra" test images have plentiful textures, some of which are large scale and some are small. They can be interpreted as different frequency components, so we used the ICTGV$^{\text{osci}}$-model with sixteen texture components according to Remark 4, i.e., $m = 17$ for these images. For the "parrots" image, we performed the experiments with eight texture components, i.e., $m = 9$.

The results were compared to a selection of popular and state-of-the-art variational image models: TGV of second order, non-local TV (NLTV) [30, 53], ICTGV [33] as well as framelet/local DCT (Fra+LDCT) [21]. Additionally, we compared to BM3D [25] which is a dedicated denoising method. For all models, we manually tuned the parameters to give the best peak signal-to-noise ratio (PSNR). However, as some methods tend to produce ghost artifacts for PSNR-optimal parameters, we also tuned parameters in order to yield visually pleasing results;
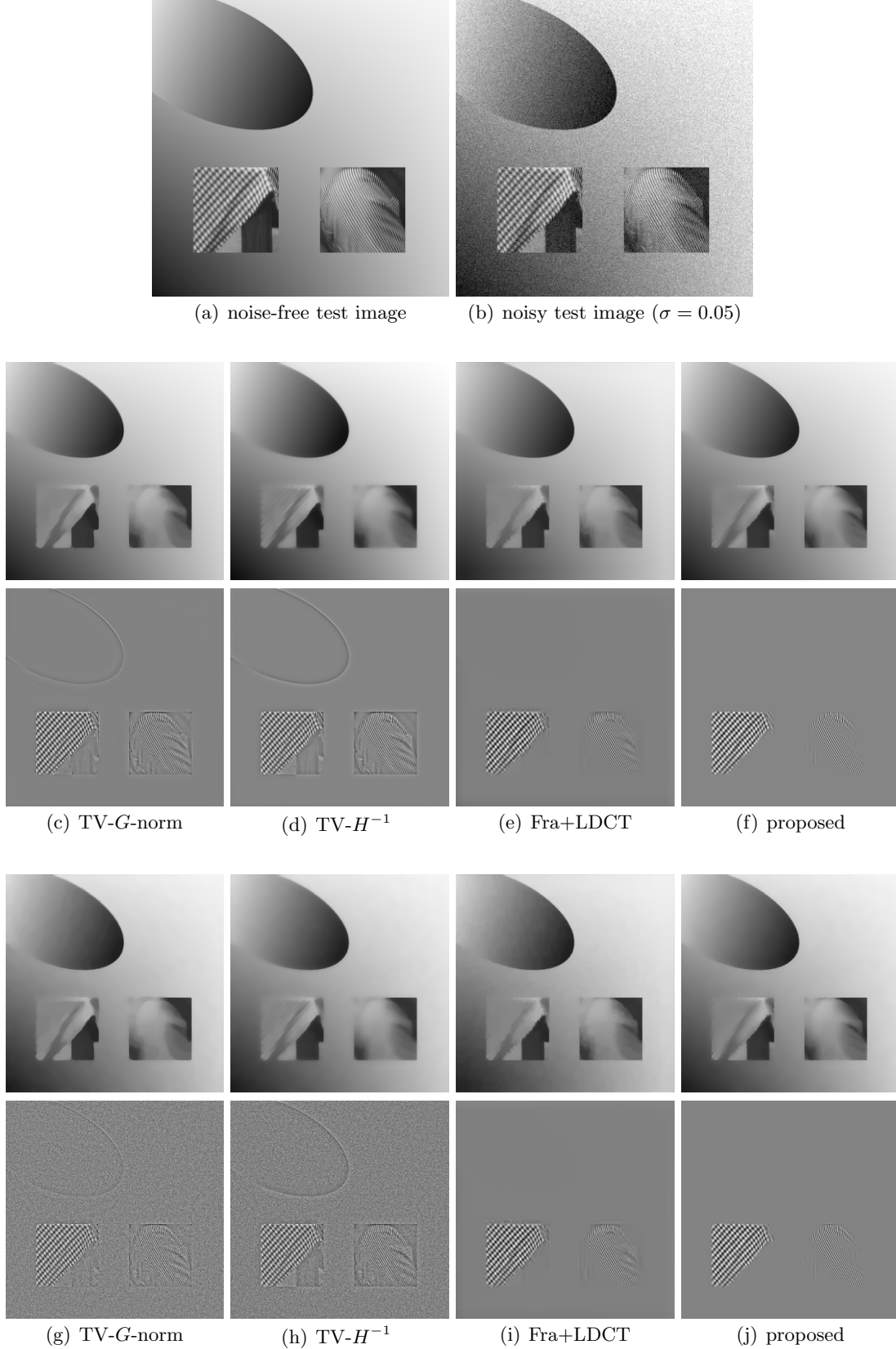
Figure 4: Comparison of cartoon/texture decomposition methods for a synthetic image. (a) A noise-free test image consisting of a cartoon part and textured regions (352×352 pixels); (b) A noisy version of (a) with Gaussian noise of zero mean and variance $\sigma = 0.05$; (c)–(f) Decomposition results for the noise-free image (top: cartoon part, bottom: texture part); (g)–(j) Decomposition results for the noisy image. Parameter choice for the proposed model: $\alpha_1 = 0.07, \beta_1 = 2\alpha_1, \gamma_1 = 0, \alpha_i = 0.9\alpha_1, \beta_i = \alpha_i, \gamma_i = 0.1\alpha_i, i = 2, \ldots, 17$.

noise level	image	TGV	NLTV	ICTGV	Fra+LDCT	BM3D	proposed model
	barbara	27.41	32.05 (31.28)	31.15 (30.33)	31.70 (31.19)	**34.43**	32.21 (31.44)
$\sigma = 0.05$	zebra	28.74	31.16 (30.07)	31.08 (30.43)	30.81 (30.06)	**32.25**	31.33 (30.46)
	parrots	34.77	35.10 (34.08)	36.49 (35.63)	36.10 (35.49)	**37.40**	36.61 (36.01)
	barbara	25.34	27.58 (27.15)	27.35 (26.71)	28.05 (27.66)	**31.06**	28.45 (27.89)
$\sigma = 0.1$	zebra	25.24	27.35 (26.93)	27.35 (26.83)	27.06 (26.44)	**28.80**	27.65 (27.00)
	parrots	32.51	31.39 (30.13)	33.28 (32.35)	32.91 (32.65)	**34.08**	33.32 (32.70)

Table 1: Comparison of denoising performance for noise levels $\sigma = 0.05$ and $\sigma = 0.1$ in terms of PSNR. (The values in parentheses correspond to the visually-optimized results.)

noise level	image	TGV	NLTV	ICTGV	Fra+LDCT	BM3D	proposed model
	barbara	0.8043	0.9005 (0.8917)	0.8859 (0.8833)	0.8857 (0.8988)	**0.9365**	0.9004 (0.8988)
$\sigma = 0.05$	zebra	0.8386	0.8647 (0.8419)	0.8761 (0.8735)	0.8653 (0.8671)	**0.8998**	0.8859 (0.8777)
	parrots	0.9157	0.8956 (0.9062)	0.9324 (0.9272)	0.9259 (0.9257)	**0.9416**	0.9358 (0.9303)
	barbara	0.7297	0.7553 (0.7814)	0.7835 (0.7840)	0.8214 (0.8201)	**0.8920**	0.8235 (0.8120)
$\sigma = 0.1$	zebra	0.7472	0.7271 (0.7540)	0.7796 (0.7897)	0.7868 (0.7795)	**0.8237**	0.7965 (0.7948)
	parrots	0.8887	0.8312 (0.8530)	0.8975 (0.8889)	0.8903 (0.8928)	**0.9025**	0.8979 (0.8936)

Table 2: Comparison of denoising performance for noise levels $\sigma = 0.05$ and $\sigma = 0.1$ in terms of SSIM. (The values in parentheses correspond to the visually-optimized results.)

at least from the subjective viewpoint of the authors (namely, the NLTV, ICTGV, Fra+LDCT and the proposed model). The PSNR and SSIM values of these experiments are reported in Tables 1 and 2, respectively.

Figure 5 shows an example of the denoising performance of the methods on the "barbara" test image with noise level $\sigma = 0.05$ in detail. Additionally, the outcome of the ICTGV$^{\mathrm{osci}}$-denoising for all tests is presented in Figure 6. For the remaining results, we refer to the supplementary material (Section A). One can see in Figure 5 that the proposed model is able to preserve more textures than the TGV model alone which is designed for piecewise smooth functions but does not specifically take textures into account. In contrast, NLTV is, due to its non-local nature, able to reconstruct parts of the texture. It suffers, however, from the staircasing behaviour that is typical for TV-like approaches. The Fra+LDCT and BM3D method yield smoother reconstructions in pure cartoon regions while also capturing the texture very well. They suffer, however, from ghosting artifacts in the smooth regions which leads, for instance, to a unnatural appearance of the face in the "barbara" image. The ICTGV model does not show these artifacts, but the texture denoising is not very well, see the trousers of the "barbara" image, where the line-like textures still look irregular. In comparison, the proposed ICTGV$^{\mathrm{osci}}$ model is able to capture both the texture and the smooth regions without producing ghosting artifacts. Nevertheless, for very fine textures, such as on the left trouser of the "barbara" image, the Fra+LDCT and BM3D approaches still reconstruct more of the details. This might be the reason for BM3D slightly outperforming the proposed methods in terms of quantitative measures (see Tables 1 and 2). The BM3D method is, however, specifically optimized for denoising and not variational, such that a generalization to inverse problems is not immediate. Among the variational approaches, the ICTGV$^{\mathrm{osci}}$ model performs best in terms of quantitative measures.

Convergence graphs that show the evolution of the root mean square error (RMSE) for TGV, ICTGV and ICTGV$^{\mathrm{osci}}$ ($m = 9, 17$) in terms of iteration number and computation time (with respect to GPU-supported MATLAB implementations) can be found in Figure 7. The experiments were performed for denoising the "barbara" image (dimensions 720×576) and with primal-dual methods for TGV and ICTGV that are similar to Algorithm 1. One can see that the essential improvement of the RMSE is achieved, for all methods, in a few hundreds of iterations and within 30 seconds, respectively. From a practical point of view this means that the proposed model and algorithm quickly improves image quality. However, these results stem from an interplay between model and algorithm, and there might be more efficient algorithms for

(a) ground truth   (b) noisy image ($\sigma =$ 0.05)   (c) TGV   (d) NLTV

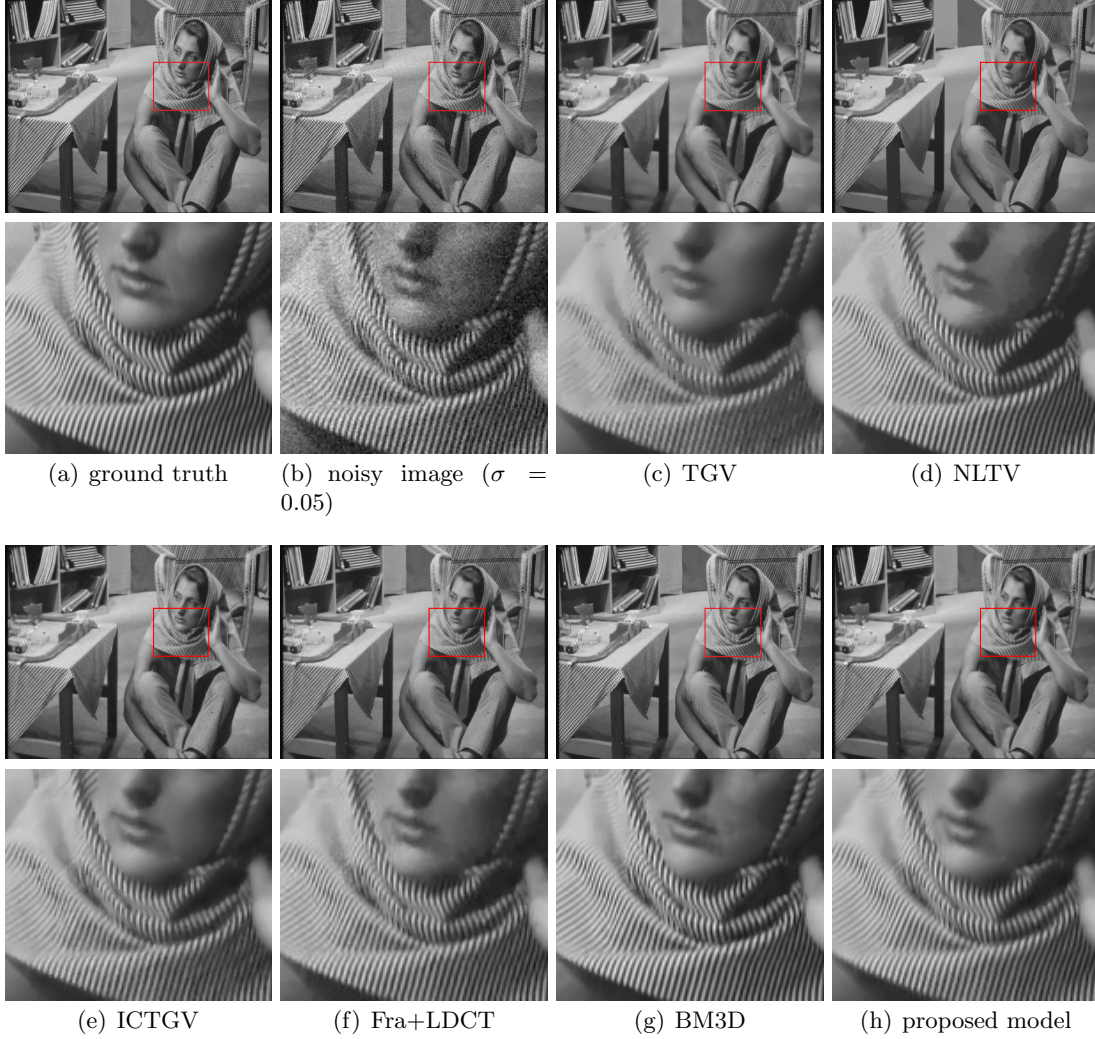(e) ICTGV   (f) Fra+LDCT   (g) BM3D   (h) proposed model

Figure 5: Denoising results for barbara image with noise level $\sigma = 0.05$. All parameters are tuned for visually-optimal results. Parameter choice for the proposed model: $\alpha_1 = 0.05, \beta_1 = \alpha_1, \gamma_1 = 0, \alpha_i = 0.9\alpha_1, \beta_i = \alpha_i, \gamma_i = 0.1\alpha_i, i = 2, \ldots, 17$.

the same model for which the behaviour is different in terms of RMSE improvement. Finally, to give an overview of the computational complexity, we also ran each algorithm with 2000 iterations for three times and took the average of the computation time, see, again Figure 7.

Finally, the supplementary material also contains a comparison for the "barbara" image with a varying number of oscillation directions in the extended model, i.e., 4, 6, 8, 10 and 12 directions (Section A.3). There, one can see that the restored images with 4 and 6 directions lose some textures, whereas the majority of textures is recovered using 8, 10 and 12 directions. Hence, the choice of 8 directions in (38) indeed offers a good balance between performance and complexity of the model.

5.2 Image inpainting

Inpainting, i.e., the restoration of missing regions of an image, is a well-studied standard image processing problem. It can, for example, be used to repair corruptions in digital photos or in ancient drawings, and to fill in the missing pixels of images transmitted through a noisy channel. In this subsection, we will apply the new regularizer to a restoration problem where pixels are missing randomly with a certain rate. We model this situation as follows. For Ω the image domain, let $\Omega' \subset \Omega$ the non-empty subdomain on which data is not missing. Given

(a) barbara ($\sigma = 0.05$)   (b) zebra ($\sigma = 0.05$)   (c) parrots ($\sigma = 0.05$)

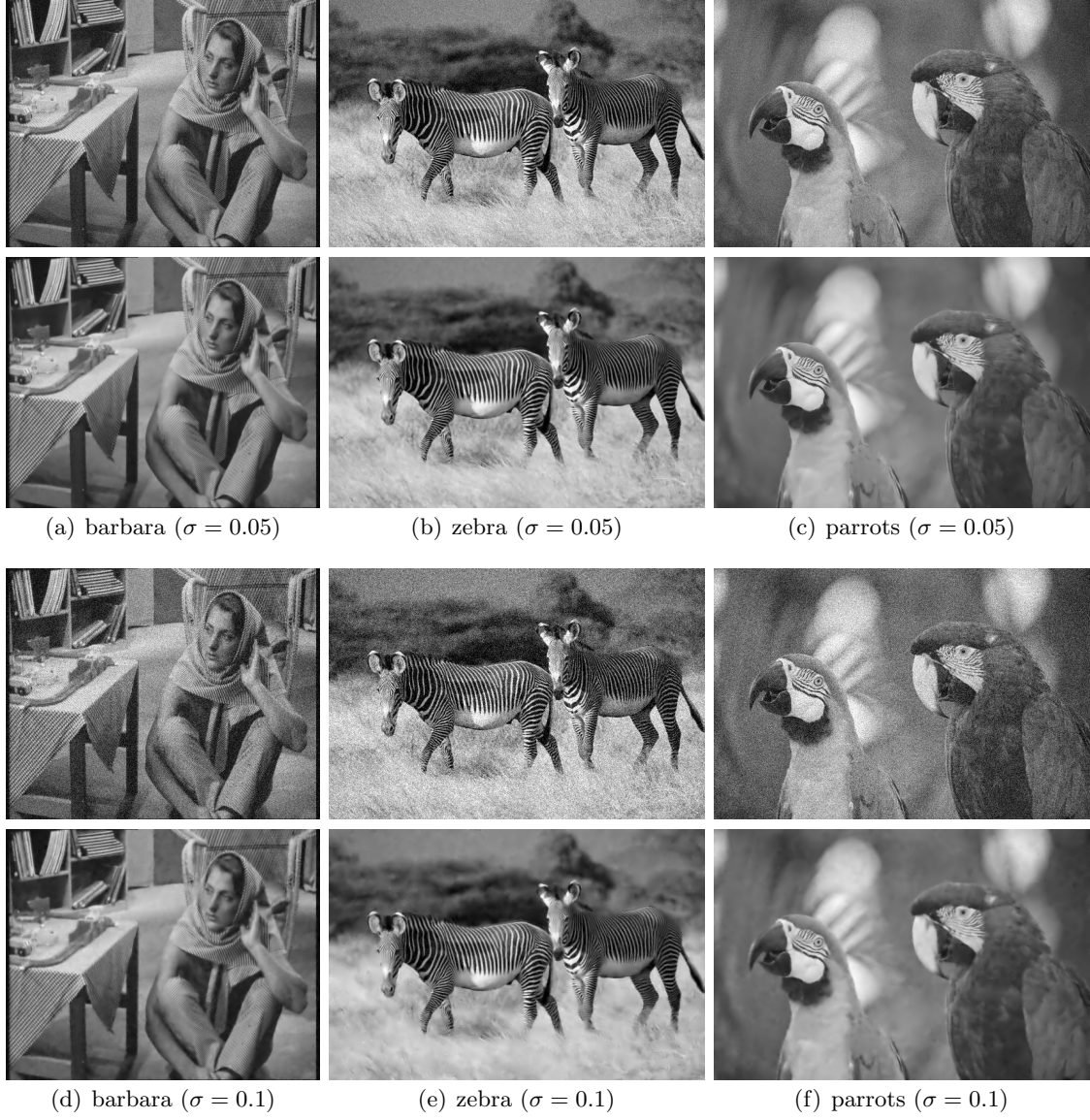(d) barbara ($\sigma = 0.1$)   (e) zebra ($\sigma = 0.1$)   (f) parrots ($\sigma = 0.1$)

Figure 6: Results of ICTGV$^{\text{osci}}$-denoising for natural images at different noise levels. For more details and comparisons, see the supplemental material.
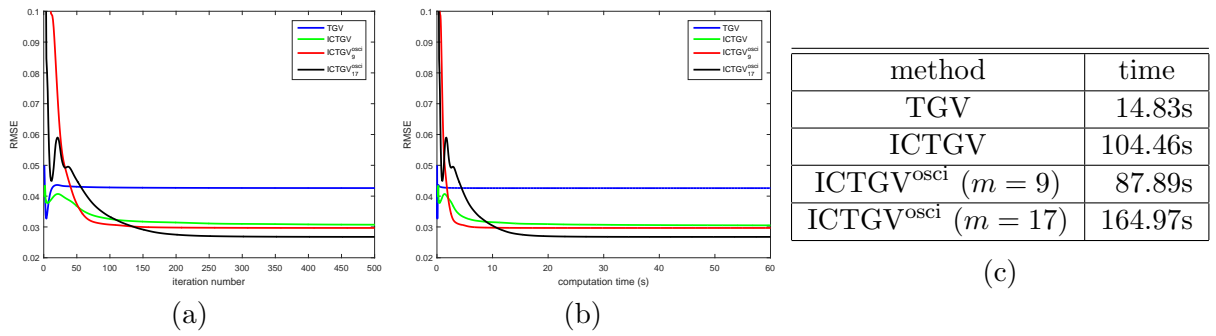

method	time
TGV	14.83s
ICTGV	104.46s
ICTGV$^{\text{osci}}$ ($m = 9$)	87.89s
ICTGV$^{\text{osci}}$ ($m = 17$)	164.97s

(c)

(a)   (b)

Figure 7: Convergence graph and computational complexity of TGV, ICTGV and ICTGV$^{\text{osci}}$ for "barbara" image with noise level $\sigma = 0.05$. (a) the RMSE values versus iterations; (b) the RMSE values versus GPU time; (c) the computation time (GPU) with 2000 iterations.
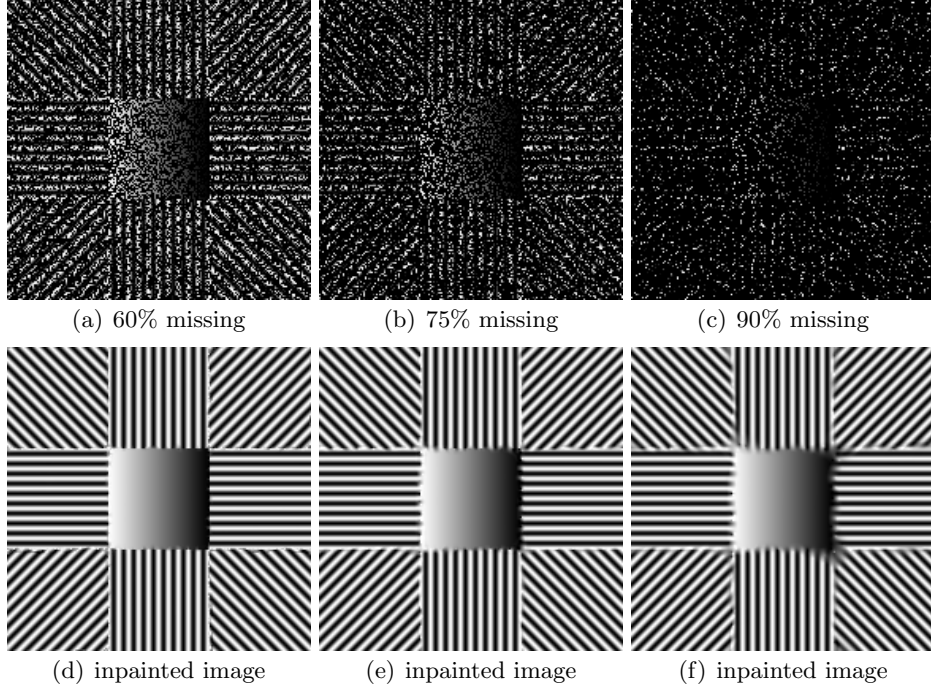
29

Figure 8: Inpainting results for a synthetic image (192×192 pixels). (a)–(c) Corrupted image with 60%, 75% and 90% pixels missing, respectively; (d)–(f) Images inpainted by model (45). Parameter choice: $\alpha_1 = 0.04, \beta_1 = \alpha_1, \gamma_1 = 0, \alpha_i = \alpha_1, \beta_i = 0.5\alpha_i, \gamma_i = 0.02\alpha_i, i = 2, \ldots, 5$.

$f \in L^2(\Omega')$, we are seeking a $u^* \in L^2(\Omega)$ that minimizes ICTGV$^{\text{osci}}$ in the variant that promotes sparse textures. In terms of the operator $K : L^2(\Omega) \to L^2(\Omega')$, $Ku = u|_{\Omega'}$ projecting u to the non-corrupted region, this corresponds to solving

$$\min_{u \in L^2(\Omega)} \text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}, \vec{\gamma}}(u) \quad \text{subject to} \quad Ku = f. \tag{45}$$

Proposition 6 applied to $F = \mathcal{I}_{\{f\}}$ and $\Phi = \text{ICTGV}^{\text{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}, \vec{\gamma}}$ yields existence in this situation. In order to solve (45), one can use a slightly modified version of Algorithm 1: Note that the only difference between the discrete version of this model and the version (40) is the fidelity term, where the term $\langle K \sum_{i=1}^m u_i - f, \lambda \rangle - \frac{\|\lambda\|^2}{2}$ is replaced by $\langle K \sum_{i=1}^m u_i - f, \lambda \rangle$. Thus, λ in Algorithm 1 is updated according to

$$\lambda^{n+1} = \lambda^n + \sigma\Big(K \sum_{i=1}^m \bar{u}_i^n - f\Big).$$

We performed numerical experiments using the inpainting model (45) and the described modification of Algorithm 1. Our first experiment aimed at recovering the synthetic image already utilized in the decomposition/denoising experiments, see Figure 3. In this experiment, 60%, 75% and 90% of the pixels were randomly removed from the image and inpainted by ICTGV$^{\text{osci}}$-minimization. The results are shown in Figure 8. There, one can see that the recovery from 60% and 75% of the missing pixels is almost perfect; whereas for 90% of missing pixels, the texture is also perfectly recovered while the boundaries of the regions deviate from the ground truth. The latter may be explained by the fact that in this situation, not enough data is available to unambiguously reconstruct the boundaries. In contrast, there is still enough information to determine the respective parameters of the affine and oscillatory parts. Overall, the experiment shows that ICTGV$^{\text{osci}}$-regularization is able to effectively reconstruct oscillatory texture in inpainting problems.

image	TGV			Fra+LDCT			proposed model		
	50%	60%	70%	50%	60%	70%	50%	60%	70%
barbara	27.49	26.26	25.09	32.75	30.76	28.53	**34.03**	**31.86**	**29.49**
zebra	26.00	24.23	22.65	28.50	26.67	25.08	**30.09**	**28.11**	**26.42**
fish	22.70	20.81	19.01	24.54	23.01	21.22	**25.10**	**23.50**	**21.54**

Table 3: Comparison of inpainting performance for different rates of missing pixels in terms of PSNR.

image	TGV			Fra+LDCT			proposed model		
	50%	60%	70%	50%	60%	70%	50%	60%	70%
barbara	0.8942	0.8562	0.8081	0.9549	0.9320	0.8953	**0.9591**	**0.9390**	**0.9078**
zebra	0.9016	0.8597	0.8051	0.9213	0.8863	0.8438	**0.9463**	**0.9194**	**0.8839**
fish	0.8676	0.8059	0.7255	0.8923	0.8462	0.7902	**0.9059**	**0.8596**	**0.7977**

Table 4: Comparison of inpainting performance for different rates of missing pixels in terms of SSIM.

Further, experiments with natural images ("barbara", "zebra" and "fish") were carried out. For these images, we also performed a comparison to TGV of second order and the framelet/local DCT model. Figure 9 shows a comparison of the inpainting performances for selected examples; we refer to the supplementary material Section B for more detailed results. Again, for the "barbara" and "zebra" image, sixteen texture components were used, i.e., $m = 17$, and for the "fish" image, experiments were performed with eight texture directions, i.e., $m = 9$. One can clearly see in Figure 9 that our proposed model is able to restore the corrupted images in high quality with the majority of the textures being recovered, such as, for instance, the table cloth and the scarf in the "barbara" image as well as the stripes on the bodies of the zebra and the fish. The same holds for the Fra+LDCT-based model while the TGV-model does not recover well the texture and leaves many unconnected "speckles". The latter is typical for derivative- and measure-based models and does also apply, e.g., to TV-inpainting. Furthermore, Tables 3 and 4 show the PSNR and SSIM values of the three models at different rates of missing pixels. As we see, the proposed model performs better than other two models by these two evaluation standards. However, despite the favorable quantitative comparison to Fra+LDCT model, the proposed model has only limited ability of recovering textures of very high frequency. Looking at the left trouser in the "barbara" image near the arm, there are some textures that are not inpainted correctly. In contrast, in the result of Fra+LDCT-model, this region is recovered well.

5.3 Undersampled magnetic resonance imaging

Since compressed sensing techniques have been shown to be effective for image reconstruction in magnetic resonance imaging (MRI) with undersampling [40, 41], i.e., where much less data is available than usually required, variational methods were becoming increasingly popular for undersampling MRI [35, 6, 36] in particular and medical image processing in general. As MRI samples data in the Fourier domain, it is particularly well-suited for these methods. In this subsection, we discuss ICTGV$^{\text{osci}}$-regularization for recovering an image given incomplete Fourier data. We model the situation as follows. Let $\Omega \subset \mathbb{R}^2$ be a bounded image domain and (Ω', μ) a finite Borel measure space on a bounded $\Omega' \subset \mathbb{R}^2$, where Ω' represents the set of measured Fourier data. The restriction of the Fourier transform $\mathcal{P}\mathcal{F}$ then maps $L^2(\Omega) \to L^2(\Omega', \mu)$ since both Ω and Ω' are bounded and the Fourier transform $\mathcal{F}$ on the whole space maps $L^1(\mathbb{R}^2)$ functions to continuous functions on $\mathbb{R}^2$. Given $f \in L^2(\Omega', \mu)$, the undersampled MRI reconstruction corresponds to solving $Ku = f$ for $K = \mathcal{P}\mathcal{F}$, such that the ICTGV$^{\text{osci}}$-regularized solution (in
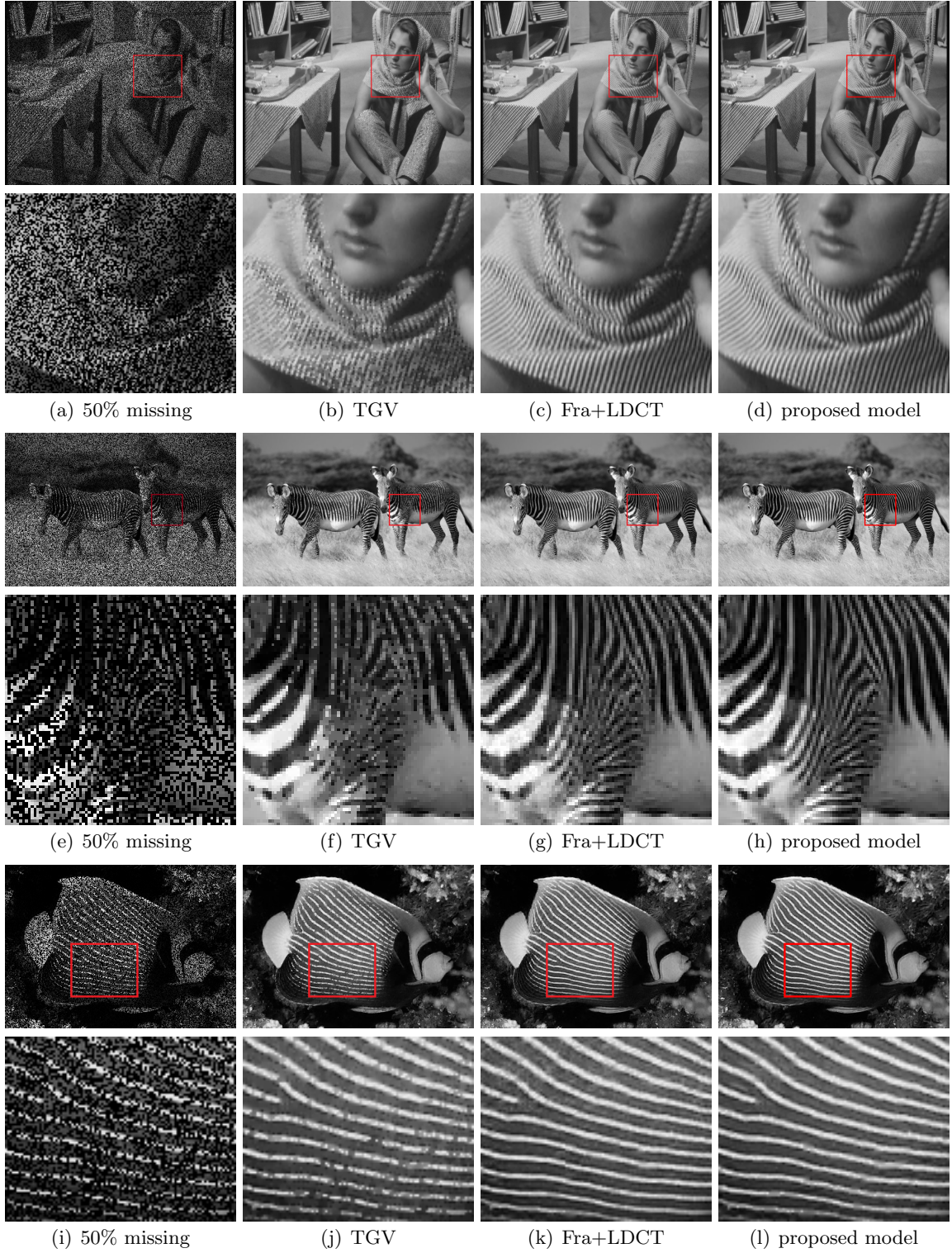
Figure 9: Inpainting results for "barbara", "zebra" and "fish" image. (a)–(d), (e)–(h), (i)–(l): Reconstruction from 50% missing pixels for proposed and reference methods for the respective test images. For more details and results, see the supplementary material.
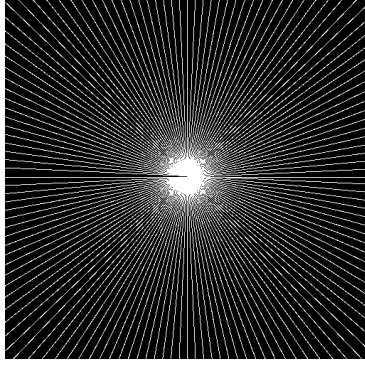
Figure 10: Example of a sampling pattern associated with radial sampling (70 equispaced radial lines, image resolution: 512×512 pixels).

sampling lines		40	50	60	70	80	90	100
TGV-shearlet	foot	**30.81**	**32.59**	33.75	34.99	36.51	37.34	38.59
	knee	28.01	28.77	29.59	30.10	30.84	31.50	32.07
	brain	34.05	35.71	37.18	38.44	39.66	40.51	41.43
proposed model	foot	29.76	32.41	**33.89**	**35.27**	**37.07**	**38.25**	**39.66**
	knee	**28.40**	**29.16**	**29.98**	**30.56**	**31.23**	**31.82**	**32.36**
	brain	**34.64**	**36.97**	**38.99**	**40.33**	**41.65**	**42.71**	**43.67**

Table 5: Comparison of MRI reconstruction performance for different sampling rates in terms of PSNR.

its sparse-texture variant) reads as

$$\min_{u \in L^2(\Omega)} \frac{1}{2} \int_{\Omega'} |Ku - f|^2 \, \mathrm{d}\mu + \mathrm{ICTGV}^{\mathrm{osci}}_{\vec{\alpha}, \vec{\beta}, \vec{c}, \vec{\gamma}}(u). \tag{46}$$

Clearly, the problem (46) admits a solution, see Theorem 2. In applications, the Borel measure space (Ω', μ) could, for instance, be $\Omega' \subset \mathbb{R}^2$ a bounded domain and $\mu = \mathcal{L}^2$ the Lebesgue measure, Ω' a finite collection of line segments and $\mu = \mathcal{H}^1 \llcorner \Omega'$ the one-dimensional Hausdorff measure restricted to Ω' (which models so-called *radial sampling* in case of radially centered line segments), or Ω' is a finite collection of points and $\mu = \mathcal{H}^0 \llcorner \Omega'$ a collection of delta-measures (modelling compressed-sensing-inspired sampling).

For the discretization, we chose $\mathcal{F}$ to represent the discrete Fourier transform on the discrete grid Ω which is usually realized by a fast algorithm (FFT). Further, the discrete domain Ω' is chosen as $\Omega' \subset \Omega$ and $\mathcal{P}$ is the projection operator which restricts Fourier coefficients to Ω'. In total, $K = \mathcal{P}\mathcal{F}$ as in the continuous setting and Algorithm 1 can be used to compute a solution.

We performed numerical experiments for undersampled MRI with radial sampling patterns associated with a varying number of sampling lines; see Figure 10 for the illustration of a selection mask with 70 equispaced radial sampling lines which corresponds to a sampling rate of approximately 14.74% for an image of 512×512 pixels. The experimental results for MR image reconstruction are shown in comparison to TGV-Shearlet MRI reconstruction in [32] which is based on simultaneous TGV and shearlet-transform-based regularization.

For the MRI reconstruction experiment, we chose some fully sampled MR test images from an online database (http://www.mr-tip.com/): a T_1-weighted foot image (512×512 pixels) and a T_1-weighted knee image (350×350 pixels); and a T_2-weighted brain image (490×490 pixels) (image courtesy of A. Farrall, University of Edinburgh). Undersampled measurements were then synthetically created by restricting the Fourier coefficients to radial sampling lines where the number of lines covers the range $40, 50, \ldots, 100$. Reconstructions were obtained from

sampling lines		40	50	60	70	80	90	100
TGV-shearlet	foot	0.8451	0.8669	0.8823	0.8956	0.9098	0.9175	0.9262
	knee	0.7356	0.7656	0.7950	0.8133	0.8348	0.8508	0.8661
	brain	0.8867	0.9059	0.9209	0.9318	0.9408	0.9466	0.9523
proposed model	foot	**0.8653**	**0.9280**	**0.9465**	**0.9567**	**0.9702**	**0.9774**	**0.9822**
	knee	**0.7581**	**0.7861**	**0.8148**	**0.8344**	**0.8570**	**0.8720**	**0.8878**
	brain	**0.9055**	**0.9421**	**0.9600**	**0.9681**	**0.9743**	**0.9784**	**0.9818**

Table 6: Comparison of MRI reconstruction performance for different sampling rates in terms of SSIM.

these measurements using the TGV+Shearlet model and by minimizing (46) where for both methods, parameters were tuned to yield optimal PSNR. Figure 11 shows exemplary results; see the supplementary material (Section C) for more detailed results. One can see in these results that both the TGV+Shearlet and the proposed model are able to preserve the textures well. Further, one can observe that the proposed model yields more naturally appearing bones and junctions that are also closer to the ground truth. The muscular tissues in the foot and knee image as well as the brain tissue are also reconstructed well. For the latter, the ICTGV$^{\text{osci}}$-model also has a desired smoothing effect that reduces the ringing artifacts originating from the MR measurements. Furthermore, note that there are some ghost stripes in the results of TGV+Shearlet model which do not appear in the ICTGV$^{\text{osci}}$ approach. The quantitative comparison in terms of PSNR and SSIM, see Tables 5 and 6, support these findings. Except for the PSNR for the foot image at low sampling rates, the proposed method consistently outperforms the TGV+Shearlet approach with respect to these quantitative measures. The only weakness one might observe in both models may the found in the foot image: There, the ground truth admits several fine directional structures in the bone (see the closeup) for which the reconstructions do not correctly recover the orientations and artificially introduce vertical structures. This effect can, however, be mitigated by switching to the extended model ($m = 17$) that also includes oscillations of higher frequencies, see the supplementary material.

6 Summary and conclusions

In this paper, we proposed the oscillation TGV regularizer which arises from modifying the second-order TGV functional such that structured oscillatory functions are in its kernel. The infimal convolution of these novel regularization functionals can be used to represent structured texture consisting of oscillations of different directions and frequencies. We have shown that the natural space for oscillation TGV is the space of functions of bounded variation in the sense that norm equivalence, lower semi-continuity and coercivity can be established in that space. Furthermore, we considered the m-fold infimal convolution of oscillation TGV which can be utilized for solving many variational image processing problems; in particular those for which TV-regularization is applicable. The analysis of the infimal-convolution model based on general properties of infimal convolution of lower semi-continuous semi-norms. For such functionals, we obtained general criteria for exactness, lower semi-continuity and coercivity of the infimal convolution. This allowed us to obtain existence results for variational problems regularized with this type of functional. We discussed a possible discretization of the oscillation TGV functional. It turned out that one has to choose the $\mathbf{c}_i$ specifically in order to preserve the kernel in the discretized version. Numerical experiments for denoising, inpainting and MRI reconstruction show the major properties of the proposed model in terms of texture preservation and reconstruction. The general framework obtained in this article allows in particular the extension to other imaging problems. A possible line of future research may focus on reducing the complexity of the proposed model by studying variants of oscillation TGV that allow for reconstruction of
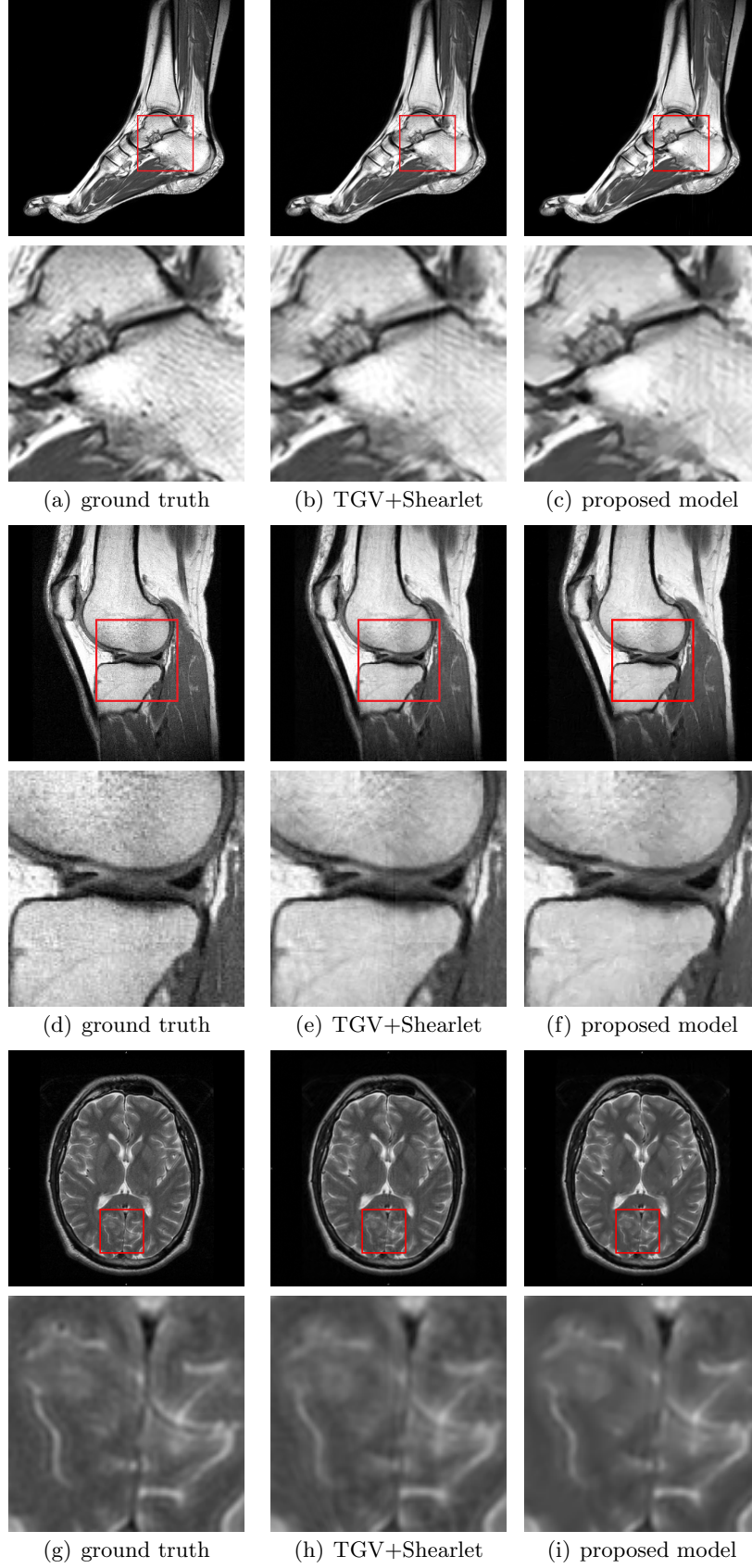
Figure 11: Comparison of MRI reconstruction results. (a) Foot image; (b)–(c) Recovery from 14.74% radially sampled coefficients; (d) Knee image; (e)–(f) Recovery from 21.12% radially sampled coefficients; (g) Brain image; (h)–(i) Recovery from 13.47% radially sampled coefficients. See the supplementary material for more detailed results.

oscillatory textures for multiple directions and frequencies with a simple, differentiation-based functional.

Acknowledgments

References

[1] L. Ambrosio, A. Coscia and G. Maso, *Fine properties of functions with bounded deformation*, Arch. Ration. Mech. Anal., 139 (1997), 201–238.

[2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, Oxford, 2000.

[3] H. Attouch and H. Brezis, *Duality for the sum of convex functions in general Banach spaces*, In Aspects of Mathematics and its Applications, 34 (1986), 125–133.

[4] J. F. Aujol, G. Aubert, L. Blanc-Féraud and A. Chambolle, *Image decomposition into a bounded variation component and an oscillating component and an oscillating component*, J. Math. Imaging Vis., 22 (2005), 71–88.

[5] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*, New York: Springer, 2011.

[6] K. T. Block, M. Uecker and J. Frahm, *Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint*, Magnetic resonance in medicine, 57(6) (2007), 1086–1098.

[7] K. Bredies, *Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty*, Lecture Notes in Computer Science, 8293 (2014), 44–77.

[8] K. Bredies and M. Holler, *Artifact-Free Variational MPEG Decompression*, Lecture Notes in Computer Science, 9087 (2015), 216–228.

[9] K. Bredies and M. Holler, *A TGV-based framework for variational image decompression, zooming and reconstruction. Part I: Analytics*, SIAM J. Imag. Sci., 8(4) (2015), 2814–2850.

[10] K. Bredies and M. Holler, *A TGV-based framework for variational image decompression, zooming and reconstruction. Part II: Numerics*, SIAM J. Imag. Sci., 8(4) (2015), 2851–2886.

[11] K. Bredies and M. Holler, *Regularization of linear inverse problems with total generalized variation*, Journal of Inverse and Ill-posed Problems, 22(6) (2014), 871–913.

[12] K. Bredies, K. Kunisch and T. Pock, *Total generalized variation*, SIAM J. Imag. Sci., 3 (2010), 492–526.

[13] K. Bredies and H. P. Sun, *Preconditioned Douglas-Rachford algorithms for TV and TGV regularized variational imaging problems*, J. Math. Imaging Vis., 52(3) (2015), 317–344.

[14] K. Bredies and T. Valkonen, *Inverse problems with second-order total generalized variation constraints*, In 9th International Conference on Sampling Theory and Applications, Singapore, 2011.

[15] A. Buades, B. Coll, and J. M. Morel, *A review of image denoising algorithms, with a new one*, Multiscale Model. Simul., 4(2) (2005), 490–530.

[16] A. Buades, B. Coll, and J. M. Morel, *Neighborhood filters and PDE's*, Numer. Math., 105(1) (2006), 1–34.

[17] A. Buades, T. Le, J. M. Morel and L. Vese, *Fast cartoon+texture image filters*, IEEE Trans. Image Process., 19(8) (2010), 1978–1986.

[18] A. Buades and J. L. Lisani, *Directional filters for color cartoon+texture image and video decomposition*, J. Math. Imaging Vis., 55 (2016), 125–135.

[19] M. Burger, G. Gilboa, M. Moeller, L. Eckardt and D. Cremers, *Spectral decompositions using one-homogeneous functionals*, SIAM J. Imag. Sci., 9(3) (2016), 1374–1408.

[20] J. F. Cai, R. H. Chan, and Z. W. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), 131–149.

[21] J. F. Cai, S. Osher, and Z. W. Shen, *Split Bregman methods and frame based image restoration*, Multiscale Model. Simul., 8(2) (2009), 337–369.

[22] E. J. Candès, X. Li, Y. Ma, and J. Wright, *Robust principal component analysis?*, Journal of the ACM, 58(3) (2011), 11.

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

[24] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vis., 40 (2011), 120–145.

[25] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, *Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering*, IEEE Trans. Image Processing, 16(8) (2007), 2080–2095.

[26] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.

[27] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho, *Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)*, Appl. Comput. Harmon. Anal., 19 (2005), 340–358.

[28] G. Gilboa, *A spectral approach to total variation*, Scale Space and Variational Methods, SSVM 2013, LNCS 7893(2013), 36–47.

[29] G. Gilboa, *A total variation spectral framework for scale and texture analysis*, SIAM J. Imag. Sci., 7(4) (2014), 1937–1961.

[30] G. Gilboa and S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul., 7(3) (2008), 1005–1028.

[31] E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser, Boston, 1984.

[32] W. H. Guo, J. Qin and W. T. Yin, *A new detail-preserving regularization scheme*, SIAM J. Imag. Sci., 7(2) (2014), 1309–1334.

[33] M. Holler and K. Kunisch, *On infimal convolution of TV-type functionals and applications to video and image reconstruction*, SIAM J. Imag. Sci., 7(4) (2014), 2258–2300.

[34] D. Horesh and G. Gilboa, *Separation surfaces in the spectral TV domain for texture decomposition*, IEEE Trans. Image Processing, 25(9) (2016), 4260–4270.

[35] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, *Second order total generalized variation (TGV) for MRI*, Magnetic Resonance in Medicine, 65(2) (2011), 480–491.

[36] F. Knoll, C. Clason, K. Bredies, M. Uecker and R. Stollberger, *Parallel Imaging with Non-linear Reconstruction using Variational Penalties*, Magnetic resonance in medicine, 67(1) (2012), 34–41.

[37] F. Knoll, M. Holler, T. Koesters, R. Otazo, K. Bredies and D. K. Sodickson, *Joint MR-PET reconstruction using a multi-channel image regularizer*, IEEE Transactions on Medical Imaging, 36(1) (2017), 1–16.

[38] R. D. Kongskov and Y. Q. Dong, *Directional Total Generalized Variation Regularization for Impulse Noise Removal*, Scale Space and Variational Methods, SSVM 2017, 221–231.

[39] C. Langkammer, K. Bredies and B. Poser et al., *Fast quantitative susceptibility mapping using 3D EPI and total generalized variation*, NeuroImage, 111 (2015), 622–630.

[40] M. Lustig, D. Donoho and J. M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magnetic resonance in medicine, 58(6) (2007), 1182–1195.

[41] M. Lustig, D. Donoho, J. M. Santos and J. M. Pauly, *Compressed sensing MRI*, IEEE Signal Processing Magazine, 25(2) (2008), 72–82.

[42] Y. Meyer, *Oscillating patterns in image processing and in some nonlinear evolution equations*, The 15th Dean Jacquelines B. Lewis Memorial Lectures, 2001.

[43] S. Osher, A. Sole, and L.A. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul., 1(2) (2003), 349–370.

[44] R. Ranftl, K. Bredies and T. Pock, *Non-local total generalized variation for optical flow estimation*, In: Computer Vision–ECCV 2014, Lecture Notes in Computer Science, 8689:439–454.

[45] R. T. Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[46] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), 61–79.

[47] H. Schaeffer and S. Osher, *A low patch-rank interpretation of texture*, SIAM J. Imag. Sci., 6(1) (2013), 226–262.

[48] M. Schloegl, M. Holler, A. Schwarzl, K. Bredies and R. Stollberger, *Infimal Convolution of Total Generalized Variation Functionals for dynamic MRI*, Magnetic Resonance in Medicine, 78(1) (2017), 142–155.

[49] J. L. Starck, M. Elad, and D. L. Donoho, *Image decomposition via the combination of sparse representations and a variational approach*, IEEE Trans. Image Process., 14 (2005), 1570–1582.

[50] R. Temam and G. Strang, *Functions of bounded deformation*, Arch. Ration. Mech. Anal., 75(1) (1980), 7–21.

[51] T. Valkonen, K. Bredies and F. Knoll, *Total generalized variation in diffusion tensor imaging*, SIAM J. Imag. Sci., 6 (2013), 487–525.

[52] L. Vese and S. Osher, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2003), 553–577.

[53] X. Q. Zhang, M. Burger, X. Bresson and S. Osher, *Bregmanized nonlocal regularization for deconvolution and sparse reconstruction*, SIAM J. Imag. Sci., 3(3) (2010), 253–276.