SIAM J. IMAGING SCIENCES
Vol. 8, No. 4, pp. 2814–2850
© 2015 Society for Industrial and Applied Mathematics

A TGV-Based Framework for Variational Image Decompression, Zooming, and Reconstruction. Part I: Analytics*

Kristian Bredies† and Martin Holler†

Abstract. A variational model for image reconstruction is introduced and analyzed in function space. Specific to the model is the data fidelity, which is realized via a basis transformation with respect to a Riesz basis followed by interval constraints. This setting in particular covers the task of reconstructing images constrained to data obtained from JPEG or JPEG 2000 compressed files. As image prior, the total generalized variation (TGV) functional of arbitrary order is employed. The present paper, the first of two works that deal with both analytical and numerical aspects of the model, provides a comprehensive analysis in function space and defines concrete instances for particular applications. A new, noncoercive existence result and optimality conditions, including a characterization of the subdifferential of the TGV functional, are obtained in the general setting.

Key words. image reconstruction, total generalized variation, JPEG decompression, JPEG 2000 decompression, variational zooming, optimality conditions

AMS subject classifications. 94A08, 49M29, 65F22, 49K30

DOI. 10.1137/15M1023865

1. Introduction.

The aim of this work is to provide a comprehensive analysis and concrete applications for a general, regularization-based model for image reconstruction. Specific to this model is the data fidelity, which is realized via interval constraints for the coefficients of some basis transformation of the L^2-space. The original motivation for this type of data constraint comes from JPEG decompression [50], where we aim at reconstructing an image subject to interval constraints of the blockwise cosine transform of the image. As image prior, the *total generalized variation* (TGV) [17] functional of arbitrary order is incorporated, and the model is formulated for multichannel images, particularly color images. A formal definition of the variational problem setting can be given as

$$\min_u \mathrm{TGV}_\alpha^k(u) + \mathcal{I}_{U_D}(u),$$

where TGV_α^k, the TGV functional of order k, generalizes the *total variation* (TV) functional by incorporating higher order smoothness information, and $\mathcal{I}_{U_D}$ is the convex indicator function of the set U_D, i.e., $\mathcal{I}_{U_D}(u) = 0$ if $u \in U_D$ and infinity else. The set U_D is formally given as

$$U_D = \{u \,|\, (Au)_i \in J_i\}$$

*Received by the editors June 1, 2015; accepted for publication (in revised form) October 2, 2015; published electronically December 3, 2015. This research was supported by the Austrian Science Fund (FWF) special research grant SFB-F32 *Mathematical Optimization and Applications in Biomedical Sciences*. The Institute of Mathematics and Scientific Computing of the University of Graz is a member of NAWI Graz (http://www.nawigraz.at/).
http://www.siam.org/journals/siims/8-4/M102386.html
†Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria (kristian.bredies@uni-graz.at, martin.holler@uni-graz.at).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with A a Riesz basis transformation operator and $(J_i)_i$ nonempty closed (not necessarily compact) intervals. In the application to JPEG decompression, one can obtain quantized coefficients of a blockwise cosine transform of the unknown, original image from the compressed file. Variational JPEG decompression with TGV^k_α regularization can then be realized in the proposed setting by defining A to be a blockwise cosine transform and the intervals $(J_i)_i$ to reflect the data loss due to quantization.

The present paper is the first of two papers that cover both a detailed analysis and the numerical realization of the problem setting of interest. This first part provides the analysis and defines concrete applications in function space, while the second part [16] deals with the numerical realization in a discrete setting.

Our work is motivated by previous papers on a TV-based JPEG decompression model [12] and on applications of TGV for regularized JPEG decompression [13] and for wavelet-based zooming [14]. The present paper provides, for the first time, a unified framework for all these settings and significantly extends the previous works both on the analysis and application sides. In terms of analysis, we deal with a general class of problems in function space that incorporate the TGV functional of any order for regularization and allow for arbitrary Riesz bases to describe the data constraints. A new, noncoercive existence result is obtained for this setting that allows a large class of interval constraints on the transform coefficients. Optimality conditions, including a characterization of the subdifferential of TGV, lay the basis for obtaining information about the structure of solutions. In terms of applications, we are able to obtain regularized reconstructions from both JPEG and JPEG 2000 compressed color images, extracting the information required for data fidelity from the encoded files. As a third application, a variational zooming method can also be derived from the general setting. These applications are realized in the second paper [16], where we use a unified algorithmic setup that yields globally convergent reconstruction algorithms in all cases. There, we define duality-based stopping criteria for the algorithm that allow us to estimate optimality in terms of the objective functional and an adaptive stepsize strategy that is needed to obtain a reasonably fast method in the case of JPEG 2000 decompression. For JPEG decompression, multicore CPU and GPU implementations are also presented.

As methods to improve standard decompression and zooming techniques are an active field of research, there exists a variety of works in this direction. In particular, the improvement upon standard JPEG decompression is an active research topic [45, 13, 12, 47, 1, 42, 46, 55]. While, in contrast to that, variational approaches designed explicitly for JPEG 2000 decompression are quite rare, the problem of wavelet coefficient inpainting is closely related and has been previously investigated in [52, 54, 22, 44, 23]. As for zooming techniques, we refer the reader to [41] for an overview and to [14, 2, 21, 39, 37, 24, 20] for methods that are related to our approach. For a more detailed discussion of existing techniques for each of the applications of interest, we ask for the reader's patience until the corresponding subsections 4.2, 4.3, and 4.4.

In contrast to application-oriented approaches, there are, besides the above-mentioned TV-based model of [12], to the best knowledge of the authors, no publications available that explicitly deal with a similar type of general problem setting in function space as it is done in the present work.

The present first part of our work is formulated in function space and deals with the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

analysis of the model. Its main section is section 3, where the model is defined and existence as well as optimality results are obtained. After that, concrete applications such as JPEG and JPEG 2000 decompression as well as variational image zooming are introduced, and appropriate frameworks are defined in function space in section 4. The second paper considers the discrete setting, deals with the algorithmic realization of the applications, and provides experimental results [16].

2. Functional-analytic background.

The aim of this section is to briefly introduce notation and mathematical concepts that are of particular relevance for this work, such as functions of bounded variation, the TGV functional, and the concept of Riesz basis. By $d \geq 2$ and m we always denote natural numbers, typically the dimension of the domain and the range of functions, respectively. By $\Omega \subset \mathbb{R}^d$ we denote a bounded Lipschitz domain. For the sake of brevity, we further denote the integral of a function ϕ over a set E with respect to a measure μ by $\int_E \phi \,\mathrm{d}\mu$ instead of $\int_E \phi(x) \,\mathrm{d}\mu(x)$, when the integration variable is clear from the context. Integration with respect to the Lebesgue measure is denoted by $\int_E \phi$.

2.1. Total variation and spaces of bounded variation.

Definition 2.1. *Let $\mathcal{B}(\Omega)$ be the Borel σ-algebra of sets in Ω. A mapping $\mathcal{B}(\Omega) \to \mathbb{R}^m$ is called an $\mathbb{R}^m$-valued finite Radon measure on Ω if it is σ-additive and $\mu(\emptyset) = 0$. We denote the variation of an $\mathbb{R}^m$-valued finite Radon measure μ by $|\mu| : \mathcal{B}(X) \to \mathbb{R}$, defined as*

$$|\mu|(E) = \sup \left\{ \sum_{i=0}^{\infty} |\mu(E_i)| \,\middle|\, (E_i)_{i\geq 0} \text{ in } \mathcal{B}(\Omega) \text{ pairwise disjoint, } E = \bigcup_{i=0}^{\infty} E_i \right\},$$

and by $\mathcal{M}(\Omega, \mathbb{R}^m)$ the space of all finite Radon measures on Ω.

The following classical result can be found in [3, Theorem 1.54].

Proposition 2.1. *The space $\mathcal{M}(\Omega, \mathbb{R}^m)$ equipped with $\|\mu\|_{\mathcal{M}} := |\mu|$ is a Banach space. Further, it can be identified with the dual of $\mathcal{C}_0(\Omega, \mathbb{R}^m)$ with the duality pairing*

$$\langle \mu, \phi \rangle = \int_\Omega \phi \,\mathrm{d}\mu := \sum_{i=1}^{m} \int_\Omega \phi_i \,\mathrm{d}\mu_i$$

for $\mu = (\mu_1, \ldots, \mu_m) \in \mathcal{M}(\Omega, \mathbb{R}^m)$, $\phi = (\phi_1, \ldots, \phi_m) \in \mathcal{C}_0(\Omega, \mathbb{R}^m)$, and the norm $\|\cdot\|_{\mathcal{M}}$ coincides with the dual norm.

Note that by $\langle \cdot, \cdot \rangle$ we always denote a duality pairing.

Definition 2.2. *We define the* total variation (TV) *functional* $\mathrm{TV} : L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m) \to \mathbb{R} \cup \{\infty\}$, *for $u = (u_1, \ldots, u_m) \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, as*

$$\mathrm{TV}(u) = \sup \left\{ \sum_{i=1}^{m} \int_\Omega u_i \,\mathrm{div}\, \phi_i \,\middle|\, \phi = (\phi_1, \ldots, \phi_m)^T \in \mathcal{C}^1_c(\Omega, \mathbb{R}^{m\times d}),\ \|\phi\|_\infty \leq 1 \right\},$$

where $\|\phi\|_\infty = \sup_{x\in\Omega} \sqrt{\sum_{j=1}^m |\phi_j(x)|^2}$ and $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$. We further define the space of functions of bounded variation

$$\mathrm{BV}(\Omega, \mathbb{R}^m) = \{u \in L^1(\Omega, \mathbb{R}^m) \mid \mathrm{TV}(u) < \infty\}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and

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

Functions of bounded variation are well known, particularly in the field of mathematical image processing, and have been extensively studied in the literature. We repeat some properties of functions of bounded variation that are most relevant to our work and refer the reader to [3, 31, 56] for proofs and further information.

Proposition 2.2. *A function* $u = (u_1, \dots, u_m) \in L^1(\Omega, \mathbb{R}^m)$ *belongs to* $\mathrm{BV}(\Omega, \mathbb{R}^m)$ *if and only if there exist finite Radon measures* $\mathrm{D}u_j = (\mathrm{D}_1 u_j, \dots, \mathrm{D}_d u_j) \in \mathcal{M}(\Omega, \mathbb{R}^d)$, $1 \le j \le m$, *such that*

$$\int_\Omega u_j \operatorname{div} \phi = -\int_\Omega \phi \, \mathrm{d}\mathrm{D}u_j \quad \forall \phi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^d).$$

Proposition 2.3. *The functional* TV *is proper, convex, and lower semicontinuous in* $L^1(\Omega, \mathbb{R}^m)$. *Further,* $\mathrm{TV}(u) = 0$ *if and only if there exist constants* $c_1, \dots c_m$ *such that* $u = (c_1, \dots, c_m)$.

Proposition 2.4. *The embedding*

$$i : \mathrm{BV}(\Omega, \mathbb{R}^m) \hookrightarrow L^p(\Omega, \mathbb{R}^m)$$

is continuous for $1 \le p \le \frac{d}{d-1}$ *and compact for* $1 \le p < \frac{d}{d-1}$.

Proposition 2.5. *Let* $u \in L^1(\Omega, \mathbb{R}^m)$. *Then* $u \in \mathrm{BV}(\Omega)$ *if and only if there exists a sequence* $(u_n)_n$ *in* $\mathcal{C}^\infty(\Omega, \mathbb{R}^m)$ *such that*

$$\|u_n - u\|_{L_1} \to 0 \quad \textit{and} \quad \mathrm{TV}(u_n) \to \mathrm{TV}(u).$$

2.2. The total generalized variation (TGV) functional. In this subsection we introduce the total generalized variation (TGV) functional for vector-valued functions. It will serve as the regularization term for the general image reconstruction problem settings in this work. The TGV functional can be considered as generalization of the TV functional that incorporates higher order smoothness. It still allows for jump discontinuities while, in contrast to TV, at the same time being able to employ higher order derivatives in smooth regions, hence avoiding the well-known staircasing effect.

The TGV functional was originally introduced in [17], and a generalization to the vector-valued case was presented in [11]. We refer the reader to [17] for a more detailed motivation and further properties and to [15] for its analysis in the context of inverse problems. Further topological properties relating the second order TGV functional to the first order TV functional were provided, for instance, in [19, 8].

Definition of the TGV functional requires the notion of spaces of symmetric tensors $\mathrm{Sym}^k(\mathbb{R}^d)$ and of tensor fields, i.e., functions mapping to spaces of symmetric tensors. We again refer the reader to [17] for an introduction of these spaces in the context of the TGV functional and provide, for the reader's convenience, a short summary in the appendix. In addition, the appendix also covers tuples of symmetric tensors, denoted by $\mathrm{Sym}^k(\mathbb{R}^d)^m$, which are required for the definition of the vectorial TGV functional as follows.

Definition 2.3. *We define the vectorial* TGV *functional of order* $k \in \mathbb{N}$ *and with parameters*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$\alpha = (\alpha_0, \dots, \alpha_{k-1}) \in (0,\infty)^k$, *for* $u \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, *as*

$$\mathrm{TGV}^k_\alpha(u) = \sup\left\{ \int_\Omega u \operatorname{div}^k \xi \;\middle|\; \xi \in \mathcal{C}^k_c(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m), \; \|\operatorname{div}^l \xi\|_\infty \le \alpha_l,\ l = 0, \dots, k-1 \right\}. \tag{2.1}$$

The norm $\|\cdot\|_\infty$ of the above definition takes the pointwise supremum with respect to a Frobenius-type tensor norm that results from an inner product, and we refer the reader to the appendix for further information. Note that we abuse notation by using the same notation as for the classical TGV functional. Also, when it is clear from the context, we will henceforth not mention the order k and the parameter vector α.

Remark 2.1. The above definition of the TGV functional is not the only possible generalization to vector-valued functions. In fact, the choice of norm on $\mathrm{Sym}^k(\mathbb{R}^d)^m$ influences this definition. The current choice has the advantage that we remain in a Hilbert space setting and hence can identify $\mathrm{Sym}^k(\mathbb{R}^d)^m$ with $\left(\mathrm{Sym}^k(\mathbb{R}^d)^m\right)^*$ with the same norm (see also [11]). In the context of color or hyperspectral image processing, different choices of norms on $\mathrm{Sym}^k(\mathbb{R}^d)^m$ might further enhance reconstruction quality, and we refer the reader to [33] and [18, section 6.3] for a discussion of suitable color norms.

Similar to the scalar case, we define the space $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$ as the set of all $L^1(\Omega, \mathbb{R}^m)$ functions such that the TGV functional is finite.

Definition 2.4. *We define*

$$\begin{aligned} \mathrm{BGV}^k(\Omega, \mathbb{R}^m) &= \left\{ u \in L^1(\Omega, \mathbb{R}^m) \mid \mathrm{TGV}^k_\alpha(u) < \infty \right\}, \\ \|u\|_{\mathrm{BGV}^k} &= \|u\|_1 + \mathrm{TGV}^k_\alpha(u). \end{aligned} \tag{2.2}$$

As one would hope, basic properties of the TGV^k_α functional and the space $\mathrm{BGV}^k(\Omega)$ can easily be transferred to the vectorial TGV^k_α functional and the space $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$. The basis for that is the following observation, which is provided in [11, Proposition 2].

Proposition 2.6. *There exist constants* $c, C > 0$ *such that, for any* $u = (u_1, \dots, u_m) \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$,

$$c \sum_{i=1}^m \mathrm{TGV}^k_\alpha(u_i) \le \mathrm{TGV}^k_\alpha(u) \le C \sum_{i=1}^m \mathrm{TGV}^k_\alpha(u_i).$$

We now summarize basic properties of the TGV^k_α functional for vector-valued functions. These assertions can either be shown to be similar to the scalar case or follow from the equivalence of Proposition 2.6 (see [17]).

Proposition 2.7. *The following statements hold:*

1. TGV^k_α *is a seminorm on the normed space* $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$.
2. TGV^k_α *and* $\mathrm{TGV}^k_{\tilde\alpha}$ *are equivalent for* $\tilde\alpha \in (0,\infty)^k$.
3. $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$ *is a Banach space.*
4. TGV^k_α *is proper, convex, and lower semicontinuous on each* $L^p(\Omega, \mathbb{R}^m)$, $1 \le p \le \infty$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

5. $\mathrm{TGV}^k_\alpha(u) = 0$ *for* $u \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ *if and only if each* u_i, $i \in \{1, \dots, m\}$, *is a polynomial of degree less than* k.

In particular, equivalence of TGV^k_α and $\mathrm{TGV}^k_{\tilde\alpha}$ for different $\alpha, \tilde\alpha \in (0,\infty)^k$ justifies the notion of $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$ independently of α.

Next we want to transfer two important results, shown in [15] for the scalar TGV functional, to the vector-valued case. These results are the minimum representation for TGV^k_α and the topological equivalence of BGV^k to BV. The proof of the minimum representation can be done almost exactly as in [15]; thus we provide only a short sketch.

Proposition 2.8. *For any* $u \in L^1(\Omega, \mathbb{R}^m)$ *we have*

$$\mathrm{TGV}^k_\alpha(u) = \min_{\substack{v_i \in \mathrm{BD}(\Omega, \mathrm{Sym}^i(\mathbb{R}^d)^m),\\ i=1,\dots,k,\\ v_0 = u,\, v_k = 0}} \sum_{i=1}^k \alpha_{k-i} \|\mathcal{E} v_{i-1} - v_i\|_{\mathcal{M}}, \tag{2.3}$$

where $\mathrm{BD}(\Omega, \mathrm{Sym}^i(\mathbb{R}^d)^m) = \mathrm{BD}(\Omega, \mathrm{Sym}^i(\mathbb{R}^d))^m$ *is the space of* m*-fold symmetric tensor fields of bounded deformation (see* [10]*), and* $\mathcal{E}v$ *denotes the symmetrized derivative of a tensor field* v*; see the appendix.*

Sketch of proof. Defining

$$X = \mathcal{C}^1_0(\Omega, \mathrm{Sym}^1(\mathbb{R}^d)^m) \times \cdots \times \mathcal{C}^k_0(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m),$$
$$Y = \mathcal{C}^1_0(\Omega, \mathrm{Sym}^1(\mathbb{R}^d)^m) \times \cdots \times \mathcal{C}^{k-1}_0(\Omega, \mathrm{Sym}^{k-1}(\mathbb{R}^d)^m),$$

the linear operator

$$\Lambda \in \mathcal{L}(X, Y), \qquad \Lambda v = \begin{pmatrix} -v_1 - \operatorname{div} v_2 \\ \cdots \\ -v_{k-1} - \operatorname{div} v_k \end{pmatrix},$$

and the proper, convex, and lower semicontinuous functionals

$$F : X \to]-\infty, \infty], \qquad F(v) = \sum_{l=1}^k I_{\{\|\cdot\|_\infty \le \alpha_{k-l}\}}(v_l) - \int_\Omega u \operatorname{div} v_1,$$
$$G : Y \to]-\infty, \infty], \qquad G(w) = \begin{cases} 0 & \text{if } w = 0, \\ \infty & \text{else,} \end{cases}$$

it follows that

$$\mathrm{TGV}^k_\alpha(u) = \sup_{v \in X} \; -F(v) - G(\Lambda v).$$

Applying [4, Corollary 2.3], we then obtain

$$\mathrm{TGV}^k_\alpha(u) = \min_{w^* \in Y^*} F^*(-\Lambda^* w^*) + G^*(w^*).$$

Rewriting the right-hand side to become (2.3) and using Proposition A.1 in the appendix, i.e.,

$$\mathcal{E} w_i^* \in \mathcal{M}(\Omega, \mathrm{Sym}^{i+1}(\mathbb{R}^d)^m) \quad \Rightarrow \quad w_i^* \in L^1(\Omega, \mathrm{Sym}^i(\mathbb{R}^d)^m), \quad 1 \le i < k,$$

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

the assertion follows. ∎

The basic result for topological equivalence follows immediately from the scalar case, as shown in [15], and Proposition 2.6.

Proposition 2.9. *Let $P_{k-1} : L^{d/(d-1)}(\Omega, \mathbb{R}^m) \to \ker \mathcal{E}^k$ be a linear, continuous, and onto projection. Then, there exists a constant $C > 0$, depending only on k, m, α, Ω, and P_{k-1}, such that*

$$\|\mathrm{D}u\|_{\mathcal{M}} \leq C\big(\|u\|_1 + \mathrm{TGV}^k_\alpha(u)\big) \qquad \text{as well as} \qquad \|u - P_{k-1}u\|_{d/(d-1)} \leq C\,\mathrm{TGV}^k_\alpha(u) \tag{2.4}$$

for all $u \in L^{d/(d-1)}(\Omega, \mathbb{R}^m)$.

Topological equivalence and an embedding result now follow trivially.

Corollary 2.1. *There exists a $\lambda > 0$ such that, for all $u \in \mathrm{BV}(\Omega, \mathbb{R}^m)$,*

$$\mathrm{TV}(u) \leq \lambda(\|u\|_1 + \mathrm{TGV}^k_\alpha(u)).$$

In particular, there exist $C > c > 0$ such that, for all $u \in \mathrm{BV}(\Omega, \mathbb{R}^m)$,

$$c(\|u\|_1 + \mathrm{TGV}^k_\alpha(u)) \leq \|u\|_1 + \mathrm{TV}(u) \leq C(\|u\|_1 + \mathrm{TGV}^k_\alpha(u)).$$

Corollary 2.2. *For $1 \leq p \leq \frac{d}{d-1}$ the space $\mathrm{BGV}^k(\Omega, \mathbb{R}^m)$ is continuously embedded into $L^p(\Omega, \mathbb{R}^m)$. If, moreover, $1 \leq p < \frac{d}{d-1}$, the embedding is compact.*

Thus, since by Corollary 2.1 $\mathrm{BV}(\Omega, \mathbb{R}^m) \simeq \mathrm{BGV}^k(\Omega, \mathbb{R}^m)$ as Banach spaces, we will in the following only use the notion $\mathrm{BV}(\Omega, \mathbb{R}^m)$.

We will require a smooth approximation of functions in $\mathrm{BV}(\Omega, \mathbb{R}^m)$ and $\mathrm{BD}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$ in a suitable topology. As these spaces are large, the topology has to be chosen sufficiently weak to achieve such approximations. The notions of strict convergence and weak-star convergence play an important role in this context, and we refer the reader to [15, 10] for results in these topologies. The approximation result needed in this work is the following.

Proposition 2.10. *For any $l \in \mathbb{N}$, $v \in \mathrm{BD}(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$, and $w \in \mathrm{BD}(\Omega, \mathrm{Sym}^{l+1}(\mathbb{R}^d)^m)$, there exists a sequence $(\phi_n)_n$ in $C^\infty(\overline{\Omega}, \mathrm{Sym}^l(\mathbb{R}^d)^m)$ such that*

$$\|\phi_n - v\|_{d/(d-1)} \to 0 \quad \text{and} \quad \|\mathcal{E}\phi_n - w\|_{\mathcal{M}} \to \|\mathcal{E}v - w\|_{\mathcal{M}} \quad \text{as} \quad n \to \infty.$$

Proof. A straightforward adaptation of [15, Lemma 5.4] to the vector-valued case gives a sequence $(\psi_n)_n$ in $C^\infty(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m) \cap \mathrm{BD}(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$ such that

$$\|\psi_n - v\|_1 \to 0 \quad \text{and} \quad \|\mathcal{E}\psi_n - w\|_{\mathcal{M}} \to \|\mathcal{E}v - w\|_{\mathcal{M}} \quad \text{as} \quad n \to \infty.$$

Now as $\mathrm{BD}(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$ is continuously embedded in $L^{d/(d-1)}(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$ (see [10, Theorem 4.16]), we can transfer L^p-mollification arguments (see [31, section 4.2], for instance) and replace the L^1 convergence of $(\psi_n)_n$ by $L^{d/(d-1)}$ convergence. Further, as the ϕ_n are in particular contained in $W^{1,1}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$ and Ω is a bounded Lipschitz domain, we can exploit standard Sobolev approximation techniques (see [31, Theorem 4.2.3]) to approximate each ϕ_n by $\psi_n \in C^\infty(\overline{\Omega}, \mathrm{Sym}^k(\mathbb{R}^d)^m)$ with respect to the $W^{1,1}$-topology, and the result follows.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.3. The notion of Riesz basis. The concept of a Riesz basis extends the classical notion of an orthonormal basis.

Definition 2.5. *Let H be a Hilbert space. We say that a sequence $(a_n)_n$ in H is a* Riesz basis *of H if* $\operatorname{span}(\{a_n | n \in \mathbb{N}\})$ *is dense in H and there exist $0 < A \leq B$ such that, for any $c = (c_i)_{i\in\mathbb{N}} \in \ell^2$, we have*

$$A \sum_{n\in\mathbb{N}} c_n^2 \leq \Big\| \sum_{n\in\mathbb{N}} c_n a_n \Big\|_H^2 \leq B \sum_{n\in\mathbb{N}} c_n^2. \tag{2.5}$$

If $(a_n)_n$ is an orthonormal basis, (2.5) holds with $A = B = 1$. Thus, orthonormal bases are indeed Riesz bases. An important property of any Riesz basis $(a_n)_n$ is the existence of a dual sequence $(\tilde{a}_n)_n$ that again is a Riesz basis.

Proposition 2.11. *Let $(a_n)_n$ be a Riesz basis of a Hilbert space H. Then, there exists a sequence $(\tilde{a}_n)_n$, the* dual Riesz basis*, such that $(\tilde{a}_n)_n$ is also a Riesz basis of H and*

$$(a_i, \tilde{a}_j)_H = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{else.} \end{cases}$$

Proof. See [53, Theorem 1.9]. ∎

As can be shown, the notion of the Riesz basis is the most general basis concept that ensures that a sequence is complete and that the resulting basis transformation is continuous and continuously invertible. Hence Riesz bases fit our data modeling very well in a general image reconstruction setting later on. There we mainly deal with componentwise bases, e.g., m possibly different Riesz bases $(a_n^i)_n$, $i = 1, \dots, m$, of $L^2(\Omega)$, and aim at reconstructing images contained in a set given as

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (a_n^i, u_i)_{L^2} \in J_n^i \text{ for all } n \in \mathbb{N},\, i = 1, \dots, m\}$$

with (J_n^i) being closed intervals.

The following remark emphasizes the connection of m componentwise Riesz bases to Riesz bases in $L^2(\Omega, \mathbb{R}^m)$. Also, in particular, interval restrictions such as those above naturally transfer to vector-valued bases.

Remark 2.2. For $i = 1, \dots, m$ let $(a_n^i)_n$ be Riesz bases of $L^2(\Omega)$. Then $(\overline{a}_n)_n = (\overline{a}_n^1, \dots, \overline{a}_n^m)_n$ defined by

$$\overline{a}_n^i = \begin{cases} a_k^i & \text{for } i = j + 1 \text{ with } n - 1 = km + j,\ k, j \in \mathbb{N}_0, \\ 0 & \text{else} \end{cases}$$

is a Riesz basis of $L^2(\Omega, \mathbb{R}^m)$. Further, given any $u = (u_1, \dots, u_m) \in L^2(\Omega, \mathbb{R}^m)$, and intervals $(J_n^i)_n$ for $1 \leq i \leq m$, $n \in \mathbb{N}$,

$$(a_n^i, u_i)_{L^2} \in J_n^i \quad \text{for all } 1 \leq i \leq m, n \in \mathbb{N},$$

is equivalent to

$$(\overline{a}_n, u)_{L^2} \in \overline{J}_n \quad \text{for all } n \in \mathbb{N},$$

where each $\overline{J}_n$ corresponds to one J_n^i.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3. The general reconstruction model. This is the main section of the work, where we study the general, TGV^k_α regularized reconstruction model. As already mentioned, the original motivation for this model was the application to artifact-free JPEG decompression. However, due to a general problem statement, it will be applicable to a broad class of problems in mathematical imaging. In particular, as will be discussed in subsection 4.1, the existence result that we obtain is applicable to *any* inverse problem with the forward map being a linear continuous operator between Hilbert spaces. We will start with a brief motivation followed by a definition and analysis of the model in a function space setting. Note that we will, without further comment, make use of the notation introduced in section 2.

3.1. Problem statement. Before we state the minimization problem and a set of generic assumptions, which make the problem setting precise, we would like to briefly sketch the original motivation for our considerations—a model for artifact-free JPEG decompression [1, 12]. For a detailed introduction to this topic we ask for the reader's patience until subsection 4.2.

The JPEG image compression standard is lossy; i.e., most of the compression is achieved by leaving out information that is believed to be less important for visual image quality. To identify such information, JPEG first applies a blockwise cosine transform to the given, uncompressed image. Following the assumption that high frequency brightness variations are less important for visual image quality, the associated cosine coefficients need only be stored with low accuracy. To this aim, the coefficients are divided by appropriate quantization values, according to a predefined quantization table, and subsequently rounded to integer. The resulting data undergoes a further lossless encoding and is, together with the quantization table, finally stored in the compressed file. Due to the quantization, this data does not provide enough information to determine a unique source image of the compression process. Standard JPEG decompression ignores this by simply using the rounded integer data for reconstruction. In contrast to that, our approach is to incorporate these data uncertainties into the model. Using the quantization matrix that is provided by the compressed file, it is possible to define a set of block-cosine coefficient data, whose quantization would coincide with the given, compressed data. Denoting this set of basis coefficient data D and the blockwise cosine transformation operator BDCT, we can formally define the (convex) set of possible source images for a given compressed JPEG file by

$$U_D = \{u \,|\, \mathrm{BDCT}(u) \in D\}.$$

Thus, reconstructing an image from the file corresponds to choosing one element of U_D. While the standard decompression method essentially picks an arbitrary element, our approach makes a choice according to a predefined image model, which is realized by employing the TGV functional as regularization term in a variational problem.

Expressed in a more abstract way, TGV regularized JPEG decompression hence means to minimize the TGV functional subject to interval constraints for the coefficients of some basis transformation of the unknown. By allowing a general class of linear basis transforms and interval constraints to realize data fidelity, not only the application to JPEG decompression, but also different decompression scenarios, such as JPEG 2000 [48] or DjVu [9] decompression, can be covered. Additionally, standard imaging problems such as zooming and in-painting also fit into this setting. This motivates an analytical investigation of this type of variational

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

approach in a general setting. To this aim, we consider the minimization problem

$$\min_{u\in L^2(\Omega,\mathbb{R}^m)} \mathrm{TGV}^k_\alpha(u) + \mathcal{I}_{U_D}(u), \tag{3.1}$$

where we use a set of generic assumptions, which are discussed below, and are given as

$$\text{(A)}\quad \begin{cases} \Omega \subset \mathbb{R}^2 \text{ is a bounded Lipschitz domain, } m \in \mathbb{N}, \\ (a_n)_n \text{ in } L^2(\Omega,\mathbb{R}^m) \text{ is a Riesz basis,} \\ (\tilde{a}_n)_n, \text{ the dual basis of } (a_n)_n, \text{ is contained in } \mathrm{BV}(\Omega,\mathbb{R}^m), \\ A : L^2(\Omega,\mathbb{R}^m) \to \ell^2, \quad (Au)_n := (u, a_n)_{L^2}, \\ (J_n)_n \text{ is a sequence of nonempty, closed intervals,} \\ D = \{z \in \ell^2 \,|\, z_n \in J_n \text{ for all } n \in \mathbb{N}\}, \\ U_D = \{u \in L^2(\Omega,\mathbb{R}^m) \,|\, Au \in D\}, \\ \text{there is a finite index set } W \subset \mathbb{N} \text{ such that} \\ U_{\mathrm{int}} := \{u \in L^2(\Omega,\mathbb{R}^m) \,|\, (Au)_n \in J_n \text{ for all } n \in \mathbb{N}\setminus W\} \text{ has nonempty interior,} \\ k \in \mathbb{N} \text{ and } \alpha = (\alpha_0,\dots,\alpha_{k-1}) \in (0,\infty)^k. \end{cases}$$

Also note that by $\mathcal{I}_{U_D}$ we denote the convex indicator function of U_D, i.e.,

$$\mathcal{I}_{U_D}(u) = \begin{cases} 0 & \text{if } u \in U_D, \\ \infty & \text{else.} \end{cases}$$

In other words, solving the minimization problem (3.1) amounts to finding a function in U_D minimizing TGV^k_α.

Let us now discuss the interpretation and motivation of (A) in more detail. The set Ω denotes the domain of our image representing functions, typically a rectangle, and Lipschitz regularity is required for embedding results. The number m denotes the number of image components, typically 3 for color images. Lines 2 to 4 of (A) define the operator A to be a basis transformation operator with respect to a Riesz basis. Thinking again of the application to JPEG decompression, A will be a blockwise cosine transform operator, and hence the basis $(a_n)_n$ will even be orthonormal. The more general assumption of a Riesz basis will, however, be needed for the application to JPEG 2000 decompression and zooming, where $(a_n)_n$ denotes a biorthogonal wavelet basis. The assumption of the dual basis being contained in BV is satisfied in all applications and is of a technical nature, as it will be needed to obtain optimality conditions. It seems, however, quite natural since the underlying assumption of all considered applications is that images can be well approximated by a finite linear combination of the dual basis.

Lines 5 to 7 of (A) define the data set via the basis transformation and interval constraints. In the application to JPEG decompression, the $(J_n)_n$ model data uncertainties due to quantization and will be bounded. We stress, however, that our model also allows unbounded intervals, which will be needed for the applications to JPEG 2000 decompression and zooming.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The assumption stated in lines 8 to 9 can be seen as a weaker version of the assumption that the data set U_D has nonempty interior. It is of a technical nature and is again only needed for the optimality conditions, in particular to ensure additivity of the subdifferential. Essentially it requires a positive minimal length of the data intervals, allowing finitely many exceptions. This is reasonable for applications where one aims to reconstruct images from compressed or low-resolution data, as in these situations the assumption corresponds to a finite precision of the given data. In particular, it is satisfied in all our target applications, while the assumption of U_D having nonempty interior would not hold true for the zooming case. The last line of (A) finally defines the order and the weight for the TGV functional as in Definition 2.3; in particular it is possible to use any order of TGV.

At last, let us again emphasize that the assumption that $(a_n)_n$ is a Riesz basis results in applicability of our framework beyond JPEG decompression, as will be discussed in more detail in subsection 4.1. Note also that assumption (A) allows us in particular to choose m different scalar-valued Riesz bases for $L^2(\Omega)$ and apply all results within assumption (A) to a corresponding Riesz basis of $L^2(\Omega, \mathbb{R}^m)$ as in Remark 2.2.

3.2. Existence of a solution. We show existence of a solution to the minimization problem (3.1) under assumption (A). First note that properness of the objective function follows easily from U_{int} having nonempty interior and $(\tilde{a}_n)_n$ in $\mathrm{BV}(\Omega, \mathbb{R}^m)$ but could have also been obtained by the more general assumption $U_D \cap \mathrm{BV}(\Omega, \mathbb{R}^m) \neq \emptyset$.

To obtain existence, we make one additional assumption which controls the number of half-bounded intervals, i.e., intervals of the form $[l, \infty)$ or $(-\infty, o]$ with $l, o \in \mathbb{R}$. For this purpose, we define $\mathcal{P}_{k-1}(\Omega, \mathbb{R}^m)$ to be the set of $\mathbb{R}^m$-valued polynomials of order less than k.

$$(\mathrm{EX}_k)\quad \left\{ \begin{array}{l} \text{With the definitions of (A), denote} \\ I = \{n \in \mathbb{N} \,|\, J_n \text{ is half-bounded and there exists } r \in \mathcal{P}_{k-1}(\Omega, \mathbb{R}^m) \text{ s.t. } (r, a_n) \neq 0\}, \\ \text{and assume that } I \text{ is a finite set.} \end{array} \right.$$

Remark 3.1. Note that in case only finitely many J_n are half-bounded, (EX_k) is trivially satisfied. In particular, (EX_k) allows arbitrarily many intervals to contain all of $\mathbb{R}$; thus it holds for any combination of bounded intervals and intervals containing all of $\mathbb{R}$ and is also satisfied in nontrivial settings where the objective functional is not coercive.

As a consequence of [6, Theorem 2.1], it suffices to show the following two assertions for $F := \mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D}$ to obtain existence of a solution to the minimization problem (3.1).

(H1) For each sequence $(x_n)_n$ in $L^2(\Omega, \mathbb{R}^m)$ satisfying

$$\|x_n\|_{L^2} \to \infty, \quad (F(x_n))_n \text{ bounded above, and } \frac{x_n}{\|x_n\|_{L^2}} \to x$$

we have

$$F(x_n - x) \leq F(x_n) \quad \text{for } n \text{ sufficiently large.}$$

(H2) For any real sequence $(t_n)_n$ with $t_n \to \infty$ and any bounded sequence $(x_n)_n$ with $x_n \rightharpoonup x$ weakly in $L^2(\Omega, \mathbb{R}^m)$ such that $F(t_n x_n)$ is bounded above, $(x_n)_n$ converges strongly to x.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In fact, given (H2) and that F is bounded below, it has been shown in [6] that even a slightly weaker version of (H1) is necessary and sufficient for existence of a solution. However, since (H1) is sufficient for our purposes, we stick with the modified version in order to avoid the introduction of additional notation.

Proposition 3.1. *Let* (A) *and* (EX_k) *be satisfied. Then there exists a solution to* (3.1).

Proof. We first verify (H2). Take the sequences $(t_n)_n$ and $(x_n)_n$ as in (H2). Since $\mathrm{TGV}^k_\alpha(t_n x_n)$ is bounded above, choosing $P_{k-1} : L^2(\Omega, \mathbb{R}^m) \to \mathcal{P}_{k-1}(\Omega, \mathbb{R}^m)$ to be a linear, continuous projection, we get by Proposition 2.9 that $t_n x_n - P_{k-1}(t_n x_n)$ is bounded. Hence, $x_n - P_{k-1}(x_n)$ converges strongly to zero. Now, weak convergence of $(x_n)_n$ to $x \in L^2(\Omega, \mathbb{R}^m)$ implies strong convergence of $(P_{k-1}(x_n))_n$ to $P_{k-1}(x)$. Hence we get from $x_n = x_n - P_{k-1}(x_n) + P_{k-1}(x_n)$ and uniqueness of the weak limit that x_n converges strongly to $x = P_{k-1}(x)$, and the claim follows.

Now assume there exists $(x_n)_n$ as in (H1). Again from boundedness of $\mathrm{TGV}^k_\alpha(x_n)$ it follows that $x_n - P_{k-1}(x_n)$ is bounded. Hence $x = \lim_{n\to\infty} \frac{x_n}{\|x_n\|_{L^2}} = \lim_{n\to\infty} \frac{P_{k-1}(x_n)}{\|x_n\|_{L^2}} \in \mathcal{P}_{k-1}(\Omega, \mathbb{R}^m)$.

Now fix $i \in \mathbb{N}$, and note first that, since $(x_n/\|x_n\|_{L^2}, a_i)_{L^2} \to (x, a_i)_{L^2}$, we can find $\epsilon_{n,i}$ such that $\epsilon_{n,i} \to 0$ as $n \to \infty$ and

$$\|x_n\|_{L^2}((x, a_i)_{L^2} + \epsilon_{n,i}) = (x_n, a_i)_{L^2} \in J_i. \tag{3.2}$$

Indeed, $(x_n, a_i)_{L^2} \in J_i$ for all i and n since $F(x_n)$ is bounded. We consider possible cases for i:

- If J_i is bounded, we can deduce from (3.2) and from $\|x_n\|_{L^2} \to \infty$ as $n \to \infty$ that $(x, a_i)_{L^2} = 0$.
- If $i \in I$ and $J_i = [l_i, \infty)$ with $l_i \in \mathbb{R}$, then necessarily $(x, a_i)_{L^2} \geq 0$, and, if the inequality is strict, we can find n_i such that for all $n \geq n_i$,
$$(x_n, a_i)_{L^2} = \|x_n\|_{L^2}((x, a_i)_{L^2} + \epsilon_{n,i}) \geq (x, a_i)_{L^2} + l_i.$$
- If $i \in I$ and $J_i = (-\infty, o_i]$ with $o_i \in \mathbb{R}$, then necessarily $(x, a_i)_{L^2} \leq 0$, and, if the inequality is strict, we can find n_i such that for all $n \geq n_i$,
$$(x_n, a_i)_{L^2} = \|x_n\|_{L^2}((x, a_i)_{L^2} + \epsilon_{n,i}) \leq (x, a_i)_{L^2} + o_i.$$
- In the remaining cases $i \notin I$ and either $(x, a_i)_{L^2} = 0$ (recall that $x \in \mathcal{P}_{k-1}(\Omega, \mathbb{R}^m)$ and the definition of I) or $J_i = \mathbb{R}$.

Since I is finite, we can define $n_0 = \max\{n_i \,|\, i \in I\}$ and get that

$$(x_n, a_i)_{L^2} - (x, a_i)_{L^2} \in J_i \quad \text{for all } i \in \mathbb{N},\ n \geq n_0,$$

and, consequently,

$$\mathrm{TGV}^k_\alpha(x_n - x) + \mathcal{I}_{U_D}(x_n - x) = \mathrm{TGV}^k_\alpha(x_n) + \mathcal{I}_{U_D}(x_n)$$

for all $n \geq n_0$, from which (H1) follows. ∎

Remark 3.2. By inspection of its proof, we note that the above existence result still holds for a weaker version of assumption (A). Indeed, instead of being a Riesz basis, $(a_n)_n$ can be any sequence in $L^2(\Omega, \mathbb{R}^m)$ (without dual basis), and the assumption that U_{int} has nonempty

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

interior was also not needed. However, in order to get existence of a nontrivial minimizer, one needs at least $U_D \cap \mathrm{BV}(\Omega, \mathbb{R}^m) \neq \emptyset$. The other assumptions of (A) will be necessary to obtain optimality conditions for (3.1).

Remark 3.3. It can also be seen from the proof of Proposition 3.1 that (EX_k) was necessary only to ensure (H1). Hence, the weaker version of (H1) presented in [6] is necessary and sufficient for existence of a solution to (3.1). However, the question whether (H1) is true or false without assuming (EX_k) remains open.

Remark 3.4. There is another possibility of obtaining existence of a solution to (3.1) for an arbitrary number of half-bounded intervals, which only requires $m\frac{k(k+1)}{2}$ (the dimension of the $\mathbb{R}^m$-valued polynomials of degree less than k) suitable intervals to be bounded. Both this assumption and (EX_k) hold for all applications considered in this work. However, since (EX_k) is easier to check, we do not discuss the alternative existence result here but rather refer the reader to [36].

Remark 3.5. Solutions to (3.1) are not unique in general. A simple case where infinitely many solutions exist can be obtained by choosing $(a_n)_n$ in (A) to be a cosine orthonormal basis and the data interval corresponding to the constant basis function to be nonsingleton. Then, one can always add a sufficiently small constant to an optimal solution without changing the objective functional. Thus, obtaining uniqueness would require additional assumptions and will not be considered further in the present work.

3.3. Optimality conditions.

Having obtained existence of a solution to (3.1) for reasonable assumptions, we now draw our attention to the derivation of optimality conditions. For this purpose, we will make use of the following obvious identity: Given F a function,

$$u^* = \operatorname*{arg\,min}_u F(u) \quad \Leftrightarrow \quad 0 \in \partial F(u^*).$$

The derivation of an optimality condition will thus be preceded by three main steps:

- Describe $\partial \mathrm{TGV}^k_\alpha$, the subdifferential of TGV^k_α.
- Describe $\partial \mathcal{I}_{U_D}$, the subdifferential of $\mathcal{I}_{U_D}$.
- Show additivity of the subdifferential operator under assumption (A).

3.3.1. Subdifferential of the TGV functional.

As the data fidelity term in our main minimization problem requires a Hilbert setting and $\mathrm{BV}(\Omega, \mathbb{R}^m)$ continuously embeds in $L^p(\Omega, \mathbb{R}^m)$ only for $p \leq d/(d-1)$, we are bound to the case $d = 2$ in the analysis within assumption (A). However, since the subdifferential of the TGV functional can be analyzed independently and such an analysis is of interest not only for our specific problem setting, we will for a moment leave the context of assumption (A) and, in this subsection, always use the following assumptions:

$$d \geq 2,\ p \in \mathbb{R} \text{ with } 1 < p \leq \tfrac{d}{d-1} \text{ and } m \in \mathbb{N}.$$

Further, we will denote the conjugate exponent of p by $p' := \frac{p}{p-1}$. Note that the restriction on p is to maintain a continuous embedding of $\mathrm{BV}(\Omega, \mathbb{R}^m)$ to $L^p(\Omega, \mathbb{R}^m)$ (see Proposition 2.4). Also, for this subsection, we always assume TGV^k_α to be a functional defined on $L^p(\Omega, \mathbb{R}^m)$.

A characterization of $\partial \mathrm{TGV}^k_\alpha$ requires a notion of tensor fields whose divergence up to a given order k can, in the weak sense, be identified with tensor fields in L^q. The space of such tensor fields, which we denote by $W^q(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$, is a generalization of the space

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$H(\mathrm{div};\Omega)$, as described, for example, in [32, Chapter 1], and also many properties can easily be generalized.

Definition 3.1. *Let* $1 \le q < \infty$, $g \in L^q(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$. *We say that* $w = \mathrm{div}\, g$ *in* $L^q(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ *if there exists* $w \in L^q(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ *such that for all* $\phi \in \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$

$$\int_\Omega (\nabla \otimes \phi) \cdot g = -\int_\Omega \phi \cdot w,$$

where $\nabla \otimes \phi$ *denotes the tensor field that is identified with the Fréchet derivative of* ϕ*; see the appendix. Furthermore, we define*

$$W^q(\mathrm{div}^k;\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m) = \Big\{ g \in L^q(\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m) \,|\, \mathrm{div}^l\, g \in L^q(\Omega,\mathrm{Sym}^l(\mathbb{R}^d)^m) \text{ for all } 1 \le l \le k \Big\}$$

with the norm $\|g\|^q_{W(\mathrm{div}^k)} := \sum_{l=0}^k \|\,\mathrm{div}^l\, g\|^q_{L^q}$.

Remark 3.6. Density of $\mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ in $L^q(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ implies that, if there exists $w \in L^q(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ as above, it is unique. By completeness of $L^q(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$, for $0 \le l \le k$ it follows that $W^q(\mathrm{div}^k;\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m)$ is a Banach space when equipped with $\|\cdot\|_{W^q(\mathrm{div}^k)}$.

Definition 3.2. *We define, again for* $1 \le q < \infty$,

$$W_0^q(\mathrm{div}^k;\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m) = \overline{\mathcal{C}_c^\infty(\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m)}^{\|\cdot\|_{W^q(\mathrm{div}^k)}},$$

i.e., the closure of $\mathcal{C}_c^\infty(\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m)$ *with respect to the norm* $\|\cdot\|_{W^q(\mathrm{div}^k)}$.

We now proceed toward a characterization of $\partial \mathrm{TGV}_\alpha^k$ by first describing the convex conjugate (or polar) of the TGV_α^k functional.

Proposition 3.2. *The convex conjugate of* TGV_α^k, *denoted by*

$$\mathrm{TGV}_\alpha^{k\,*} : L^{p'}(\Omega,\mathbb{R}^m) \to \overline{\mathbb{R}},$$

has the form

$$\mathrm{TGV}_\alpha^{k\,*}(v) = \mathcal{I}_{\overline{C_\alpha^k}}(v) = \begin{cases} 0 & \text{if } v \in \overline{C_\alpha^k}, \\ \infty & \text{if } v \notin \overline{C_\alpha^k}, \end{cases}$$

where

$$C_\alpha^k := \Big\{ \mathrm{div}^k \xi \,\Big|\, \xi \in \mathcal{C}_c^k(\Omega,\mathrm{Sym}^k(\mathbb{R}^d)^m),\ \|\,\mathrm{div}^l\, \xi\|_\infty \le \alpha_l,\ l = 0,\dots,k-1 \Big\}, \tag{3.3}$$

and the closure is taken with respect to the $L^{p'}$ *norm.*

Proof. This follows easily from convexity and lower semicontinuity of TGV_α^k and $\mathcal{I}_{\overline{C_\alpha^k}}$ since

$$\mathrm{TGV}_\alpha^k(u) = \mathcal{I}^*_{C_\alpha^k}(u),$$

and thus (see [30, Propositions 3.2 and 4.1]),

$$\mathrm{TGV}_\alpha^{k\,*}(v) = \mathcal{I}^{**}_{C_\alpha^k}(v) = \mathcal{I}_{\overline{C_\alpha^k}}(v). \qquad \blacksquare$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

A more detailed description of $\mathrm{TGV}_\alpha^{k\,*}$ follows from a study of $\overline{C_\alpha^k}$.

Proposition 3.3. *With $\overline{C_\alpha^k}$ as in Proposition 3.2, we have*

(3.4)
$$\overline{C_\alpha^k} = \left\{ \operatorname{div}^k g \,\middle|\, g \in W_0^{p'}(\operatorname{div}^k; \Omega, \operatorname{Sym}^k(\mathbb{R}^d)^m),\ \ \|\operatorname{div}^l g\|_\infty \le \alpha_l,\ l = 0, \dots, k-1 \right\} := K_\alpha^k.$$

Proof. In order to show that $\overline{C_\alpha^k} \subset K_\alpha^k$, it is sufficient to show that K_α^k is closed with respect to $\|\cdot\|_{L^{p'}}$. Define

$$W_0^{p',\alpha}(\operatorname{div}^k) := \{ g \in W_0^{p'}(\operatorname{div}^k; \Omega, \operatorname{Sym}^k(\mathbb{R}^d)^m) \,|\, \|\operatorname{div}^l g\|_\infty \le \alpha_l,\ l = 0, \dots, k-1 \}.$$

Now let $h \in \overline{K_\alpha^k}$. There exists a sequence $(g_n)_{n\ge 0}$ in $W_0^{p',\alpha}(\operatorname{div}^k)$ such that $\lim_{n\to\infty} \operatorname{div}^k g_n = h$. If we can show that there exists $g \in W_0^{p',\alpha}(\operatorname{div}^k)$ such that $\operatorname{div}^k g = h$, closedness of K_α^k follows. By boundedness of $\|\operatorname{div}^l g_n\|_\infty$, $0 \le l < k$, there exist $h^l \in L^{p'}(\Omega, \operatorname{Sym}^{k-l}(\mathbb{R}^d)^m)$ and a set of increasing indices $(n_i)_i$ in $\mathbb{N}$ such that

$$\operatorname{div}^l g_{n_i} \underset{L^{p'}}{\rightharpoonup} h^l \quad \text{as } i \to \infty \qquad \text{for all } 0 \le l < k.$$

Denoting $h^k = h$ it follows that, for $0 \le l \le k-1$ and $\phi \in \mathcal{C}_c^\infty(\Omega, \operatorname{Sym}^{k-1-l}(\Omega)^m)$,

$$\int_\Omega h^l \cdot \mathcal{E}\phi = \lim_{i\to\infty} \int_\Omega \operatorname{div}^l g_{n_i} \cdot \mathcal{E}\phi = \lim_{i\to\infty} (-1) \int_\Omega \operatorname{div}^{l+1} g_{n_i} \cdot \phi = (-1) \int_\Omega h^{l+1} \cdot \phi,$$

which implies $g := h^0 \in W^{p'}(\operatorname{div}^k; \Omega, \operatorname{Sym}^k(\mathbb{R}^d)^m)$ and $\operatorname{div}^l g = h^l$, $0 \le l \le k$.

In order to prove that $g \in W_0^{p',\alpha}(\operatorname{div}^k)$, we note that the set

$$\left\{ (z, \operatorname{div} z, \dots, \operatorname{div}^k z) | z \in W_0^{p',\alpha}(\operatorname{div}^k) \right\} \subset L^{p'}\left(\Omega, \bigtimes_{l=0}^{k} \operatorname{Sym}^{k-l}(\mathbb{R}^d)^m \right)$$

is convex and closed—and therefore weakly closed. Since the sequence $((g_{n_i}, \operatorname{div} g_{n_i}, \dots, \operatorname{div}^k g_{n_i}))_i$ is contained in this set and converges weakly to $(g, \operatorname{div} g, \dots, \operatorname{div}^k g)$, it follows that $g \in W_0^{p',\alpha}(\operatorname{div}^k)$.

Next, we prove $K_\alpha^k \subset \overline{C_\alpha^k}$. To this aim, it suffices to show that, for $g \in W_0^{p',\alpha}(\operatorname{div}^k)$ arbitrary, we have

$$\int_\Omega u \operatorname{div}^k g \le \mathrm{TGV}_\alpha^k(u) \quad \text{for all } u \in \mathrm{BV}(\Omega, \mathbb{R}^m),$$

since this implies that $\mathrm{TGV}_\alpha^{k\,*}(\operatorname{div}^k g) = 0$ and hence $\operatorname{div}^k g \in \overline{C_\alpha^k}$. In view of the equivalent characterization of TGV_α^k as given in Proposition 2.8, we prove the more general assertion that, for any $l = 1, \dots, k$, it holds that, for any $v \in \mathrm{BD}(\Omega, \operatorname{Sym}^{k-l}(\mathbb{R}^d)^m)$,

$$\left| \int_\Omega v \cdot \operatorname{div}^l g \right| \le \inf_{\substack{v_i \in \mathrm{BD}(\Omega, \operatorname{Sym}^{k-l+i}(\mathbb{R}^d)^m),\\ i=1,\dots,l,\\ v_0 = v,\, v_l = 0}} \sum_{i=1}^{l} \alpha_{l-i} \|\mathcal{E} v_{i-1} - v_i\|_{\mathcal{M}}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Setting $l = k$ then implies the result. We show this assertion by induction. For $l = 1$ we get, by a divergence theorem for tensor fields [17, Proposition 2.1], for $\phi \in C^\infty(\overline{\Omega}, \mathrm{Sym}^{k-1}(\mathbb{R}^d)^m)$ and $\psi \in C_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$,

$$\int_\Omega \phi \cdot \operatorname{div} \psi = -\int_\Omega \mathcal{E}\phi \cdot \psi.$$

Exploiting density, we can replace ψ by $g \in W_0^{p',\alpha}(\operatorname{div}^k)$ and estimate

$$\left| \int_\Omega \phi \cdot \operatorname{div} g \right| = \left| \int_\Omega \mathcal{E}\phi \cdot g \right| \leq \alpha_0 \|\mathcal{E}\phi\|_1.$$

Now approximating an arbitrary $v \in \mathrm{BD}(\Omega, \mathrm{Sym}^{k-1}(\mathbb{R}^d)^m)$ by a sequence $(\phi_n)_n$ contained in $C^\infty(\overline{\Omega}, \mathrm{Sym}^{k-1}(\mathbb{R}^d)^m)$ as in Proposition 2.10, the induction basis follows. For any $l \in \{2, \dots, k\}$ we take $\phi \in C^\infty(\overline{\Omega}, \mathrm{Sym}^{k-l}(\mathbb{R}^d)^m)$, apply the divergence theorem, and add and subtract $v_1 \in \mathrm{BD}(\Omega, \mathrm{Sym}^{k-l+1}(\mathbb{R}^d)^m)$ to get

$$\left| \int_\Omega \phi \cdot \operatorname{div}^l g \right| \leq \alpha_{l-1} \|\mathcal{E}\phi - v_1\|_1 + \left| \int_\Omega v_1 \cdot \operatorname{div}^{l-1} g \right|.$$

Again using a smooth approximation as in Proposition 2.10, the assertion follows from the induction hypothesis for $l - 1$. ∎

Having a sufficient description of $\mathrm{TGV}_\alpha^{k\,*}$, we can now characterize its subdifferential. The relation

$$u^* \in \partial \mathrm{TGV}_\alpha^k(u) \quad \Leftrightarrow \quad \mathrm{TGV}_\alpha^k(u) + \mathrm{TGV}_\alpha^{k\,*}(u^*) = \langle u, u^* \rangle$$

(see [30], Proposition I.5.1) together with the description of $\mathrm{TGV}_\alpha^{k\,*}$ immediately implies the following result.

Theorem 3.1. *Let $u \in L^p(\Omega, \mathbb{R}^m)$, $u^* \in L^{p'}(\Omega, \mathbb{R}^m)$. Then $u^* \in \partial \mathrm{TGV}_\alpha^k(u)$ if and only if* (3.5)

$$\begin{cases} u \in \mathrm{BV}(\Omega, \mathbb{R}^m) \text{ and there exists } g \in W_0^{p'}(\operatorname{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m) \text{ such that } \|\operatorname{div}^l g\|_\infty \leq \alpha_l, \\ l = 0, \dots, k-1,\ u^* = \operatorname{div}^k g, \text{ and} \\ \qquad \mathrm{TGV}_\alpha^k(u) = \displaystyle\int_\Omega u \operatorname{div}^k g. \end{cases}$$

3.3.2. Subdifferential of the data term. In order to describe $\partial \mathcal{I}_{U_D}$, first note that we can decompose $\mathcal{I}_{U_D} = \mathcal{I}_D \circ A$. Since $\partial \mathcal{I}_D$ can be described quite easily, we use a chain rule to deduce $\partial \mathcal{I}_{U_D} = \partial(\mathcal{I}_D \circ A) = A^* \partial \mathcal{I}_D \circ A$ and, consequently, to characterize $\partial \mathcal{I}_{U_D}$.

To this aim, we first summarize the relation between Riesz bases and transformation operators in the following proposition that can be shown by standard arguments.

Proposition 3.4. *Let $(a_n)_n$ and $(\tilde{a}_n)_n$ be two Riesz bases in duality in the Hilbert space H. Then, the operators*

$$\begin{aligned} A : H &\to \ell^2, & \tilde{A} : H &\to \ell^2 \\ u &\mapsto ((u, a_n)_H)_n & u &\mapsto ((u, \tilde{a}_n)_H)_n \end{aligned}$$

are both continuous and possess continuous inverses with

$$A^{-1} = \tilde{A}^*, \quad \tilde{A}^{-1} = A^*.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Their adjoints are given by

$$A^*\lambda = \sum_{n\in\mathbb{N}} \lambda_n a_n, \quad \tilde{A}^*\lambda = \sum_{n\in\mathbb{N}} \lambda_n \tilde{a}_n.$$

Using in particular bijectivity of A, the subdifferential of $\mathcal{I}_{U_D}$ can now be characterized as follows.

Proposition 3.5. *Let* (A) *be satisfied. Then*

$$u^* \in \partial\mathcal{I}_{U_D}(u) \quad \Leftrightarrow \quad u \in U_D \text{ and } u^* = A^*\lambda$$

with $\lambda = (\lambda_n)_n \in \ell^2$ *such that, for every* $n \in \mathbb{N}$,

$$\begin{cases} \lambda_n \geq 0 & \text{if} \quad (Au)_n = \sup(J_n) \neq \inf(J_n), \\ \lambda_n \leq 0 & \text{if} \quad (Au)_n = \inf(J_n) \neq \sup(J_n), \\ \lambda_n = 0 & \text{if} \quad (Au)_n \in \operatorname{int}(J_n), \\ \lambda_n \in \mathbb{R} & \text{if} \quad (Au)_n = \inf(J_n) = \sup(J_n). \end{cases} \tag{3.6}$$

Proof. At first, since $A : L^2(\Omega, \mathbb{R}^m) \to \ell^2$ is bijective and $\operatorname{dom}(\mathcal{I}_D) \neq \emptyset$, we can apply [7, Corollary 16.42] to obtain

$$u^* \in \partial\mathcal{I}_{U_D}(u) \quad \Leftrightarrow \quad u^* = A^*\lambda$$

for some $\lambda \in \partial\mathcal{I}_D(Au)$. By a standard result in convex analysis we have

$$\lambda \in \partial\mathcal{I}_D(Au) \quad \Leftrightarrow \quad Au = P_D(Au + \lambda),$$

where P_D denotes the projection onto the set D. A straightforward argument further allows us to reduce P_D to a componentwise projection;

$$Au = P_D(Au + \lambda) \quad \Leftrightarrow \quad (Au)_n = P_{J_n}((Au)_n + \lambda_n) \quad \text{for all } n \in \mathbb{N}.$$

From that, the assertion follows by an easy case study. ∎

3.3.3. Additivity of the subdifferential. Finally, we need to show that $\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D})(u) = \partial\mathrm{TGV}^k_\alpha(u) + \partial\mathcal{I}_{U_D}(u)$. For that, we first decompose $\mathcal{I}_{U_D} = \mathcal{I}_{U_{\text{int}}} + \mathcal{I}_{U_{\text{point}}}$, where, based on assumption (A),

$$U_{\text{int}} = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Au)_n \in J_n \text{ for all } n \in \mathbb{N} \setminus W\}$$

and

$$U_{\text{point}} = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Au)_n \in J_n \text{ for all } n \in W\}$$

for a finite index set $W \subset \mathbb{N}$ such that $\operatorname{int}(U_{\text{int}}) \neq \emptyset$.

Theorem 3.2. *Let* (A) *be satisfied. Then, for all* $u \in L^2(\Omega, \mathbb{R}^m)$,

$$\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D})(u) = \partial\mathrm{TGV}^k_\alpha(u) + \partial\mathcal{I}_{U_D}(u). \tag{3.7}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Let $u \in L^2(\Omega, \mathbb{R}^m)$. It is sufficient to show $\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D})(u) \subset \partial \mathrm{TGV}^k_\alpha(u) + \partial \mathcal{I}_{U_D}(u)$, since the other inclusion is always satisfied. Continuity of $\mathcal{I}_{U_{\mathrm{int}}}$ in at least one point $u \in \mathrm{BV}(\Omega, \mathbb{R}^m) \cap U_D$ allows us to apply [30, Proposition I.5.6], and to ensure that

$$\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_{\mathrm{point}}} + \mathcal{I}_{U_{\mathrm{int}}})(u) \subset \partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_{\mathrm{point}}})(u) + \partial \mathcal{I}_{U_{\mathrm{int}}}(u).$$

We now want to use [4, Corollary 2.1] to establish

$$\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_{\mathrm{point}}})(u) \subset \partial \mathrm{TGV}^k_\alpha(u) + \partial \mathcal{I}_{U_{\mathrm{point}}}(u),$$

for which it is sufficient to show that $\mathrm{dom}(\mathrm{TGV}^k_\alpha) - \mathrm{dom}(\mathcal{I}_{U_{\mathrm{point}}}) = L^2(\Omega, \mathbb{R}^m)$. But this is true since, for any $w \in L^2(\Omega, \mathbb{R}^m)$, by taking $j_n \in J_n$ for $n \in W$, we can write

$$w = w_1 - w_2,$$

where

$$w_1 = \sum_{n \in W} ((a_n, w)_{L^2} + j_n)\, \tilde{a}_n \in \mathrm{dom}(\mathrm{TGV}^k_\alpha)$$

as (A) assumes that each $\tilde{a}_n \in \mathrm{BV}(\Omega, \mathbb{R}^m)$, and

$$w_2 = -\sum_{n \in N \setminus W} (a_n, w)_{L^2} \tilde{a}_n + \sum_{n \in W} j_n \tilde{a}_n \in \mathrm{dom}(\mathcal{I}_{U_{\mathrm{point}}}).$$

Again, since

$$\partial \mathcal{I}_{U_{\mathrm{point}}}(u) + \partial \mathcal{I}_{U_{\mathrm{int}}}(u) \subset \partial(\mathcal{I}_{U_{\mathrm{point}}} + \mathcal{I}_{U_{\mathrm{int}}})(u) = \partial \mathcal{I}_{U_D}(u)$$

is always satisfied, the assertion is proved. ∎

3.3.4. Optimality system. The previous results now allow us to derive an optimality system.

Theorem 3.3. *Let* (A) *and* (EX_k) *be satisfied. Then there exists a solution of*

$$\min_{u \in L^2(\Omega, \mathbb{R}^m)} \left(\mathrm{TGV}^k_\alpha(u) + \mathcal{I}_{U_D}(u) \right),$$

and the following are equivalent:

1. $\hat{u} \in \arg\min_{u \in L^2(\Omega, \mathbb{R}^m)} \left(\mathrm{TGV}^k_\alpha(u) + \mathcal{I}_{U_D}(u) \right) = \arg\min_{u \in U_D} \mathrm{TGV}^k_\alpha(u)$.
2. $\hat{u} \in \mathrm{BV}(\Omega, \mathbb{R}^m) \cap U_D$, *and there exist* $g \in W^2_0(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^2)^m)$ *and* $\lambda = (\lambda_n)_n$ *in* ℓ^2 *satisfying*
   (a) $\|\mathrm{div}^l g\|_\infty \le \alpha_l$, $l = 0, \dots, k-1$,
   (b) $\mathrm{TGV}^k_\alpha(\hat{u}) = -\int_\Omega \hat{u}\, \mathrm{div}^k g$,
   (c) $\mathrm{div}^k g = \sum_{n \in \mathbb{N}} \lambda_n a_n$, *where, for all* $n \in \mathbb{N}$,
   $$\begin{cases} \lambda_n \ge 0 \text{ if } (A\hat{u})_n = \sup(J_n) \ne \inf(J_n), \\ \lambda_n \le 0 \text{ if } (A\hat{u})_n = \inf(J_n) \ne \sup(J_n), \\ \lambda_n = 0 \text{ if } (A\hat{u})_n \in \mathrm{int}(J_n) \end{cases}$$
   (note that, if $J_n = \{j_n\}$*, there is no additional condition on* λ_n*).*
3. $\hat{u} \in \mathrm{BV}(\Omega, \mathbb{R}^m) \cap U_D$, *and there exists* $g \in W^2_0(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^2)^m)$ *satisfying*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(a) $\|\operatorname{div}^l g\|_\infty \leq \alpha_l$, $l = 0, \dots, k-1$,
(b) $\mathrm{TGV}^k_\alpha(\hat{u}) = -\int_\Omega \hat{u} \operatorname{div}^k g$,
(c) *for all* $n \in \mathbb{N}$,

$$\begin{cases} (\operatorname{div}^k g, \tilde{a}_n)_{L^2} \geq 0 & \text{if } (A\hat{u})_n = \sup(J_n) \neq \inf(J_n), \\ (\operatorname{div}^k g, \tilde{a}_n)_{L^2} \leq 0 & \text{if } (A\hat{u})_n = \inf(J_n) \neq \sup(J_n), \\ (\operatorname{div}^k g, \tilde{a}_n)_{L^2} = 0 & \text{if } (A\hat{u})_n \in \operatorname{int}(J_n). \end{cases}$$

Proof. Existence of a solution follows from Proposition 3.1. Equivalence of 2 and 3 follows from biorthogonality of $(a_n)_n$ and $(\tilde{a}_n)_n$ (see Proposition 2.11), so it is left to show equivalence of 1 and 2. For this purpose, let

$$\hat{u} \in \operatorname*{arg\,min}_{u \in L^2(\Omega, \mathbb{R}^m)} \left(\mathrm{TGV}^k_\alpha(u) + \mathcal{I}_{U_D}(u) \right).$$

Thus $0 \in \partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D})(\hat{u})$, and by additivity of the subdifferential for this setting (see Theorem 3.2) we have $0 \in \partial \mathrm{TGV}^k_\alpha(\hat{u}) + \partial \mathcal{I}_{U_D}(\hat{u})$. Hence, there exist $z_1 \in \partial \mathrm{TGV}^k_\alpha(\hat{u})$ and $z_2 \in \partial \mathcal{I}_{U_D}(\hat{u})$ such that $0 = z_1 + z_2$. Now, by Theorem 3.1, $\hat{u} \in \mathrm{BV}(\Omega, \mathbb{R}^m)$, and there exists $g \in W^2_0(\operatorname{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^2)^m)$ satisfying 2(a) such that $z_1 = -\operatorname{div}^k g$ and $\mathrm{TGV}^k_\alpha(\hat{u}) = -\int_\Omega \hat{u} \operatorname{div}^k g$. Clearly, we have $\operatorname{div}^k g = z_2$. By Proposition 3.5 there exists $\lambda = (\lambda_n)_n \in \ell^2$ satisfying the elementwise conditions in 2(c) such that $\operatorname{div}^k g = A^*\lambda = \sum_{n \in \mathbb{N}} \lambda_n a_n$, the latter by Proposition 3.4. For the converse implication, observe that conditions 2(a) and 2(b) together with $\hat{u} \in \mathrm{BV}(\Omega, \mathbb{R}^m)$ imply that $-\operatorname{div}^k g \in \partial \mathrm{TGV}^k_\alpha(\hat{u})$ (Theorem 3.1), while 2(c) together with $\hat{u} \in U_D$ imply that $\operatorname{div}^k g \in \partial \mathcal{I}_{U_D}(\hat{u})$ (Proposition 3.5). Hence $0 \in \partial \mathrm{TGV}^k_\alpha(\hat{u}) + \partial \mathcal{I}_{U_D}(\hat{u}) = \partial(\mathrm{TGV}^k_\alpha(\hat{u}) + \mathcal{I}_{U_D})(\hat{u})$, and $\hat{u}$ is a minimizer. ∎

4. Application to data reconstruction.

The purpose of this section is to show how various models related to mathematical imaging problems are covered by the framework as derived in section 3. In the first subsection, we give some remarks about the general class of problem settings to which the theory of section 3 can be applied. Then, in the succeeding subsections, we will study the specific application to decompression and zooming problems in detail.

4.1. A general class of problem settings.

The aim of this subsection is to describe a class of inverse problems whose TGV^k_α regularization fits into the general framework of section 3. The basis for such a description is the following proposition, which replaces the Riesz basis transform in (A) with any bounded linear operator $B : L^2(\Omega, \mathbb{R}^m) \to \ell^2$ having closed range and allows for more general interval constraints.

Proposition 4.1. *Let $B : L^2(\Omega, \mathbb{R}^m) \to \ell^2$ be a bounded linear operator with closed range and, with $(e_i)_{i \in N}$, $N \subset \mathbb{N}$, a Riesz basis of its range and $(J_n)_n$ nonempty, closed intervals, define*

$$U_D := \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Bu, e_n)_{\ell^2} \in J_n \text{ for all } n \in N\}. \tag{4.1}$$

Then there exist nonempty, closed intervals $(\tilde{J}_n)_n$ and a Riesz basis $(a_n)_n$ of $L^2(\Omega, \mathbb{R}^m)$ such that, with $A : L^2(\Omega, \mathbb{R}^m) \to \ell^2$ the basis transformation operator corresponding to $(a_n)_n$, we have

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Au)_n \in \tilde{J}_n \text{ for all } n \in \mathbb{N}\}.$$

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Denote by $(\tilde{e}_i)_{i\in N}$ the dual Riesz basis to $(e_i)_{i\in N}$ and by $(z_i)_{i\in\mathbb{N}\setminus N}$ an orthonormal basis of $\ker(B)$. With these definitions, we choose sequences $(a_n)_n$, $(\tilde{a}_n)_n$ in $L^2(\Omega,\mathbb{R}^m)$ according to

$$a_i = \begin{cases} B^* e_i & \text{if } i \in N, \\ z_i & \text{if } i \in \mathbb{N}\setminus N \end{cases}$$

and

$$\tilde{a}_i = \begin{cases} B^{-1}\tilde{e}_i & \text{if } i \in N, \\ z_i & \text{if } i \in \mathbb{N}\setminus N, \end{cases}$$

where $B^{-1} : \mathrm{Rg}(B) \to \ker(B)^\perp$ denotes the inverse of $B : \ker(B)^\perp \to \mathrm{Rg}(B)$ which is linear and continuous. We would like to show that $(a_i)_i$ and $(\tilde{a}_i)_i$ are biorthogonal Riesz bases. For this purpose, according to [53, Theorem 1.9], it suffices to show that $(a_i)_i$ and $(\tilde{a}_i)_i$ both have dense linear span and are biorthogonal and that for any $f \in L^2(\Omega,\mathbb{R}^m)$ we have

$$\sum_{n\in\mathbb{N}} |(f,a_n)_{L^2}|^2 < \infty, \quad \sum_{n\in\mathbb{N}} |(f,\tilde{a}_n)_{L^2}|^2 < \infty.$$

Concerning density, suppose that, for arbitrary $w_1, w_2 \in L^2(\Omega,\mathbb{R}^m)$, $(a_n, w_1)_{L^2} = 0$ as well as $(\tilde{a}_n, w_2)_{L^2} = 0$ for all $n \in \mathbb{N}$. Given that $(z_i)_{i\in\mathbb{N}\setminus N}$ is a basis for $\ker(B)$, this implies $w_1, w_2 \in \ker(B)^\perp$. But $0 = (a_n, w_1)_{L^2} = (e_n, Bw_1)_{\ell^2}$ for all $n \in N$ implies that $Bw_1 = 0$; thus $w_1 \in \ker(B)$ and $w_1 = 0$. Similarly, $0 = (\tilde{a}_n, w_1)_{L^2} = (B^{-1}e_n, w_1)_{L^2}$ for all $n \in N$ implies, by surjectivity of $B^{-1} : \mathrm{Rg}(B) \to \ker(B)^\perp$, that also $w_2 = 0$. Thus both sequences have dense linear span. Now, for $i, j \in \mathbb{N}$, it follows that

$$(a_i, \tilde{a}_j)_{L^2} = \left.\begin{cases} (B^*e_i, B^{-1}\tilde{e}_j)_{L^2} & \text{if } i \in N, j \in N, \\ (B^*e_i, z_j)_{L^2} & \text{if } i \in N, j \in \mathbb{N}\setminus N, \\ (z_i, B^{-1}e_j)_{L^2} & \text{if } i \in \mathbb{N}\setminus N, j \in N, \\ (z_i, z_j)_{L^2} & \text{if } i \in \mathbb{N}\setminus N, j \in \mathbb{N}\setminus N \end{cases}\right\} = \delta_{i,j},$$

where we used that $BB^{-1}\tilde{e}_j = \tilde{e}_j$ and $B^{-1}e_j \in \ker(B)^\perp$ for $j \in N$, $Bz_j = 0$ for $j \in \mathbb{N}\setminus N$, and the fact that $(e_i)_{i\in N}$ and $(\tilde{e}_i)_{i\in N}$ are dual and $(z_i)_{i\in\mathbb{N}\setminus N}$ is an orthonormal basis. To show the remaining assertion, take any $f = f_1 + f_2 \in \ker(B)^\perp \oplus \ker(B) = L^2(\Omega,\mathbb{R}^m)$. Then, for constants $C_1, C_2 > 0$,

$$\sum_{n\in\mathbb{N}} |(f,a_n)_{L^2}|^2 = \sum_{n\in N} |(Bf_1, e_n)_{\ell^2}|^2 + \sum_{n\in\mathbb{N}\setminus N} |(f_2, z_n)_{L^2}|^2 \le C_1\|Bf_1\|_{\ell^2}^2 + \|f_2\|_{L^2}^2$$

and

$$\sum_{n\in\mathbb{N}} |(f,\tilde{a}_n)_{L^2}|^2 = \sum_{n\in N} |(B^{-*}f_1, \tilde{e}_n)_{\ell^2}|^2 + \sum_{n\in\mathbb{N}\setminus N} |(f_2, z_n)_{L^2}|^2 \le C_2\|B^{-*}f_1\|_{\ell^2} + \|f_2\|_{L^2}.$$

Consequently, $(a_n)_n, (\tilde{a}_n)_n$ are biorthogonal Riesz bases.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Finally, we define the intervals $(\tilde{J}_i)_{i\in\mathbb{N}}$ by $\tilde{J}_i = J_i$ for $i \in N$ and $\tilde{J}_i = \mathbb{R}$ for $i \in \mathbb{N} \setminus N$. Then, setting $\tilde{U}_D = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Au)_n \in \tilde{J}_n \text{ for all } n \in \mathbb{N}\}$, we have

$$u \in \tilde{U}_D \Leftrightarrow (u, B^* e_i)_{L^2} \in \tilde{J}_i \text{ for all } i \in N \Leftrightarrow (Bu, e_i)_{\ell^2} \in J_i \text{ for all } i \in N \Leftrightarrow u \in U_D. \qquad \blacksquare$$

We now consider the ill-posed operator equation

$$F(u) = d,$$

with $F : L^2(\Omega, \mathbb{R}^m) \to \ell^2$ a linear, bounded operator and $d \in \ell^2$ a given, degraded datum which is close to the true, unknown datum $d^\dagger$. We assume that information on the data acquisition process allows us to define an index set N, closed intervals $(J_n)_{n\in N}$, and a sequence $(e_n)_{n\in N}$ in ℓ^2 such that

$$d^\dagger \in D := \{v \in \ell^2 \,|\, (v, e_n)_{\ell^2} \in J_n \text{ for all } n \in N\}$$

and D is sufficiently "small." Motivated by the true signal being an image, our aim is to apply TGV^k_α regularization and reconstruct a signal $u^* \in L^2(\Omega, \mathbb{R}^m)$ which solves

$$\min_{F(u)\in D} \mathrm{TGV}^k_\alpha(u). \tag{4.2}$$

This setting is related to residual methods for inverse problems, and we refer the reader to [34] for a discussion of such methods in a more general context.

The theory of section 3 can now be applied to this situation as follows.

- If the forward mapping F is well behaved, i.e., has closed range, and the ill-posedness is hence only given in terms of nonuniqueness, and if further $(e_n)_{n\in N}$ constitutes a Riesz basis of $\mathrm{Rg}(F)$, then Proposition 4.1 can be applied, and the data constraints can equivalently be defined on the coefficients of the signal after a Riesz basis transform. If the data intervals are such that (EX_k) as well as the nonempty interior condition of (A) are satisfied and the dual Riesz basis is further contained in $\mathrm{BV}(\Omega, \mathbb{R}^m)$, all results of section 3 apply; in particular, existence of a solution is guaranteed, and the optimality conditions are valid. One example of this situation is given when F is a Riesz basis transform, but the $(e_n)_{n\in N}$ constitute a Riesz basis different from the standard basis in ℓ^2. Another example is the situation when F is a bounded surjective operator with nontrivial kernel for which interval restrictions on its coefficients can, for example, be realized by taking $(e_n)_{n\in N}$ as the standard basis of ℓ^2.
- If the data intervals $(J_n)_n$ only satisfy the assumption (EX_k), in particular if only finitely many of them are half-bounded, we can still guarantee existence of a solution to (4.2). Indeed, inspection of the proof of Proposition 4.1 shows that, without further assumptions on the sequence $(e_n)_n$ and the bounded linear operator $F : L^2(\Omega, \mathbb{R}^m) \to \ell^2$, we can set $(a_n)_{n\in N} = (F^* e_n)_{n\in N}$ and equivalently define U_D by using inner products with this sequence. In view of Remark 3.2, we can thus still guarantee existence of a solution.

Having discussed the general applicability of our model to inverse problems, we now turn to concrete applications in mathematical imaging.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

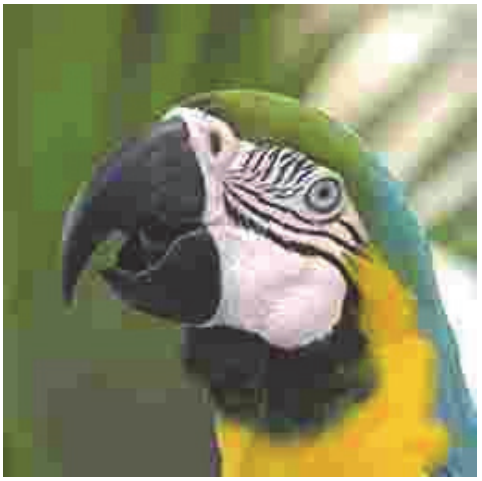

Figure 1. *JPEG image with typical blocking and ringing artifacts.*

4.2. Color JPEG decompression. As first application, we consider the problem of artifact-free decompression of JPEG compressed color images. This problem has already been addressed in various publications, of which the TV-based models of [12, 1, 55] are most related. Also, a discrete version of the problem using second order TGV regularization has already been published in [13] and in [45]. We further refer the reader to [12, 42, 47, 46] for a short overview of current standard techniques.

We start with a brief explanation of the basic steps of the JPEG compression standard. For further information about our modeling we refer the reader to [12, 13], and for a more detailed explanation of the JPEG compression procedure we refer the reader to [50].

The process of JPEG compression is lossy, which means that typically most of the compression is obtained by loss of data. As a consequence, the original image cannot be restored completely from the compressed object, which causes ringing and blocking artifacts in the reconstructed images, as can be seen, for example, in Figure 1. Figure 2 gives an overview of the basic steps of JPEG compression for color images that are important for our reconstruction framework. In particular, a further lossless coding of integer data is omitted here, since this procedure can be inverted without loss of data.

A color JPEG image is typically processed in the YCbCr color space, where the first (luminance) component essentially contains the brightness information and the second two (chroma) components the color information of the image. This color space is equivalent to the standard RGB color space, and images can be transformed from one to another without significant loss of data. The advantage of using the YCbCr color space is the following: Knowing that the human visual system is less sensitive to color than to brightness oscillations, as a first step of JPEG compression, data reduction can be achieved by subsampling the two chroma components.

Next, each component undergoes a discrete cosine transformation on each block of 8×8 pixels, resulting in a local representation of the components as a linear combination of different frequencies. Again, there is empirical evidence that the human visual system is less sensitive

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

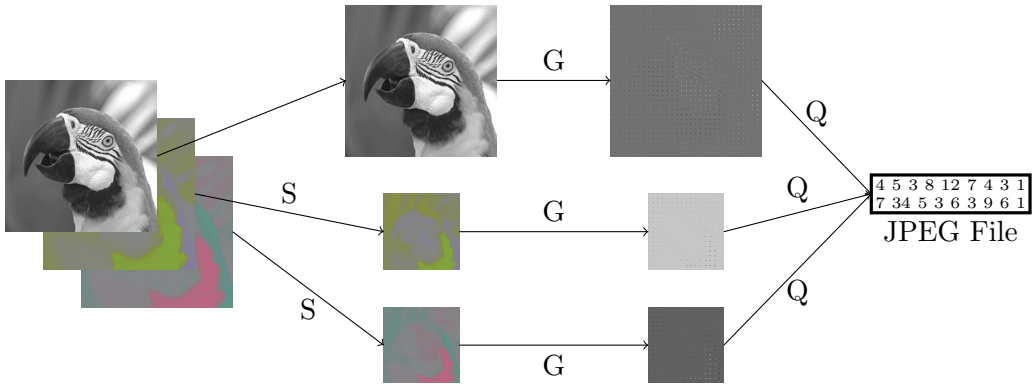

Figure 2. *Scheme of JPEG compression procedure. Here, S denotes a subsampling operation, G a blockwise discrete cosine transformation, and Q a quantization to integer, i.e., a blockwise division through a predefined quantization matrix followed by rounding to integer.*

to high frequency variations than to low frequency variations. Consequently, each 8×8 pixel block of the coefficients representing the cosine frequencies is divided pointwise by a predefined quantization matrix Q reflecting this empirical observation. The resulting data is then rounded to integer and, after further lossless compression, is stored in the compressed JPEG object.

In order to reconstruct an image from the compressed file, standard decompression algorithms now simply revert the compression process by pointwise multiplication with the quantization matrix, application of the inverse blockwise cosine transform, and color upsampling. It is thereby not taken into account that the data is incomplete, i.e., that it is a result of a rounding procedure and thus does not uniquely determine a single source image, but rather a set of possible source images. Indeed, since, in addition to the quantized integer coefficient data $d = (d^c_{i,j})$, the quantization matrix $Q = (Q^c_{i,j})$ can also be obtained from the compressed file, it is possible to define interval bounds

$$J^c_{i,j} = \left[Q^c_{i,j}\left(d^c_{i,j} - \frac{1}{2}\right), Q^c_{i,j}\left(d^c_{i,j} + \frac{1}{2}\right)\right] \tag{4.3}$$

for each quantized coefficient, and, consequently, a convex set of possible source data

$$D = \{(z^c_{i,j}) \,|\, z^c_{i,j} \in J^c_{i,j} \text{ for all } i, j, c\}. \tag{4.4}$$

Then D is the set of all coefficients that would, after quantization and rounding, result in the same data as given by the JPEG compressed file. Note that here, i and j denote the ith and jth coefficient of the blockwise cosine transform, respectively, while $c \in \{1, 2, 3\}$ denotes the color component.

Coupling the subsampling S and the cosine transformation operator C, as in Figure 2, we will see that with this the set of all possible source images of the compressed JPEG object can be described by D and an (even orthogonal) basis transformation operator. Thus it fits in our image reconstruction framework, where we aim at choosing one of all possible source images that minimizes the TGV^k_α functional.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

4.2.1. Continuous modeling. We consider color images in YCbCr color space as functions in $L^2(\Omega, \mathbb{R}^3)$, where $\Omega = (0, 8k) \times (0, 8l)$, $k, l \in \mathbb{N}$, is a rectangular domain, in particular a Lipschitz domain.

Subsampled image components are considered as functions in $L^2(\Omega_c)$, where $\Omega_c = (0, 8k_c) \times (0, 8l_c)$ are domains smaller than Ω, i.e., $k_c \leq k$, $l_c \leq l$. With these prerequisites, the subsampling process can be described color componentwise via the operators $S_c : L^2(\Omega) \to L^2(\Omega_c)$, $c \in \{1, 2, 3\}$, given by

$$S_c u(x, y) = u(s_c x, t_c y),$$

where $s_c = \frac{k}{k_c}, t_c = \frac{l}{l_c}$ are the subsampling factors.

Remark 4.1. Typically we have no subsampling for the luminance component, i.e., $s_1 = t_1 = 1$, while the chroma components are subsampled with factor 2, i.e., $s_2 = t_2 = s_3 = t_3 = 2$.

In order to define the blockwise cosine transform, we first need the following definition, which is taken from [12].

Definition 4.1 (blockwise cosine system). *For $t, r \in \mathbb{N}$, set $H = (0, 8t) \times (0, 8r) \subset \mathbb{R}^2$. For $i, j \in \mathbb{N}_0$, $0 \leq i < t$, $0 \leq j < r$, we define the squares*

$$E_{i,j} = \big([8i, 8i + 8) \times [8j, 8j + 8) \big) \cap H$$

and

$$\chi_{i,j} = \chi_{E_{i,j}},$$

their characteristic functions. Furthermore, let the standard cosine orthonormal system $(b_{n,m})_{n,m} \subset L^2((0, 1)^2)$ be defined as

$$b_{n,m}(x, y) = \lambda_n \lambda_m \cos(nx\pi) \cos(my\pi) \tag{4.5}$$

for $(x, y) \in \mathbb{R}^2$ and $n, m \in \mathbb{N}_0$, where

$$\lambda_l = \begin{cases} 1 & \text{if } l = 0, \\ \sqrt{2} & \text{if } l \neq 0. \end{cases}$$

We define the blockwise cosine system $g^{i,j}_{n,m} \in L^2(H)$ *as the collection of all $g^{i,j}_{n,m} \in L^2(H)$ according to*

$$g^{i,j}_{n,m}(x, y) = \frac{1}{8} b_{n,m} \left(\frac{x - 8i}{8}, \frac{y - 8j}{8} \right) \chi_{i,j}(x, y) \tag{4.6}$$

for $(x, y) \in H$. Here, $0 \leq i < t$, $0 \leq j < r$, and $n, m \in \mathbb{N}_0$.

Remark 4.2. It follows by reduction to the cosine orthonormal system $(b_{n,m})_{n,m}$ that $\{g^{i,j}_{n,m} \,|\, n, m \in \mathbb{N}_0, 0 \leq i < k, 0 \leq j < l\}$ is a complete orthonormal system in $L^2(H)$. Further, one can see that $\{g^{i,j}_{n,m} \,|\, n, m \in \mathbb{N}_0, 0 \leq i < k, 0 \leq j < l\} \subset \mathrm{BV}(H)$.

Denoting, for $c \in \{1, 2, 3\}$, by $(g^c_n)_n$ a blockwise cosine orthonormal system of $L^2(\Omega_c)$ as described in Definition 4.1 (note that we use a different index notation), the operators $G_c : L^2(\Omega_c) \to \ell^2$ are defined to be their corresponding basis transformation operators, i.e.,

$$(G_c v)_n = (g^c_n, v)_{L^2} \tag{4.7}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

for $v \in L^2(\Omega_c)$.

With these preliminaries, we define the operator modeling the JPEG compression procedure for each color component as $A_c : L^2(\Omega) \to \ell^2$ with $A_c = G_c S_c$. We further assume that we are given closed intervals $(J_n^c)_n$ such that

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^3) \,|\, (Au)_n^c \in J_n^c \text{ for } n \in \mathbb{N}, c \in \{1,2,3\}\} \tag{4.8}$$

defines the set of possible source image of a given, JPEG compressed file. Clearly, each A_c is bijective, and, following the proof of Proposition 4.1, A_c can be regarded as basis transformation operator, related to basis elements $(a_n^c)_n$, that can be given as

$$a_n^c(x,y) = S_c^* g_n^c(x,y) = \frac{1}{s_c t_c} g_n^c\left(\frac{x}{s_c}, \frac{y}{t_c}\right), \tag{4.9}$$

which are orthogonal and contained in $\mathrm{BV}(\Omega)$. Further, rewriting the componentwise operators A_c as basis transformation operator $A : L^2(\Omega, \mathbb{R}^3) \to \ell^2$, U_D can be rewritten into a form as in assumption (A), and the continuous minimization problem corresponding to color JPEG decompression is given by

$$\min_{u \in L^2(\Omega, \mathbb{R}^3)} \mathrm{TGV}_\alpha^k(u) + \mathcal{I}_{U_D}(u), \tag{4.10}$$

with U_D being equivalently defined in (4.8). In order to ensure that (A) and (EX_k) hold, we need to specify our continuous modeling of the given data intervals $(J_n^c)_n$. Remember that for each given integer coefficient d_n^c and the corresponding quantization value Q_n^c we can define an error interval J_n^c as

$$J_n^c = \left[Q_n^c\left(d_n^c - \frac{1}{2}\right), Q_n^c\left(d_n^c + \frac{1}{2}\right)\right]. \tag{4.11}$$

Now one can model the JPEG compression process by assuming that all coefficients $((a_n^c, u^\dagger))_n$ of the original image $u^\dagger \in L^2(\Omega, \mathbb{R}^m)$ are quantized, rounded, and stored. This means that all $(J_n^c)_n$ are given from data coefficients $(d_n^c)_n$, and quantization values $(Q_n^c)_n$, as in (4.11), in particular are bounded. Since the quantization values Q_n^c are typically nondecreasing for larger n, meaning that coefficients representing higher frequency are stored with less or the same precision, it is reasonable to assume that they are bounded below by some $\epsilon > 0$. Hence, as each coefficient d_n^c results from a rounding procedure of quantized data contained in ℓ^2, all but finitely many must be zero, and the resulting data intervals thus contain $[-\frac{\epsilon}{2}, \frac{\epsilon}{2}]$. Consequently, U_D has nonempty interior, and assumptions (A) and (EX_k) are clearly satisfied. This ensures existence of a solution and validity of the optimality condition as in Theorem 3.3.

Remark 4.3. As an alternative approach, one could regard the coefficient data d_n^c as a given, finite number of samples of the unknown image $u^\dagger \in L^2(\Omega, \mathbb{R}^m)$. This means that only finitely many intervals J_n^c can be constrained, as in (4.11), and all remaining intervals are set to be all of $\mathbb{R}$. Such a setting again satisfies our assumptions (A) and (EX_k), and thus all results of section 3 apply.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

4.3. Color JPEG 2000 decompression. As a second application, we employ the reconstruction model of section 3 for the improved reconstruction of JPEG 2000 color images, where the coding is essentially based on a biorthogonal wavelet transform. For the sake of self-containedness we will briefly explain basic features of JPEG 2000 compression that are necessary to understand the modeling. It will turn out that, again, the set of possible source images can be described by interval restrictions on the coefficients of the transformed image and thus fits our reconstruction model of section 3. However, due to the coding process, it will not be possible to restrict every coefficient by a bounded interval. After presenting an overview of JPEG 2000 compression, we will define and discuss the minimization problem for artifact-free JPEG 2000 decompression.

But at first, we discuss previous approaches to improving the reconstruction quality of the JPEG 2000 standard. To the best knowledge of the authors, in contrast to the JPEG decompression model, there does not exist any model or method designed particularly for improved JPEG 2000 decompression that is related to the present one. However, even if not designated to improve the JPEG 2000 compression/decompression procedure, some works on wavelet inpainting aim to solve a very similar task: Assuming that, due to transmission or storage error, some coefficients of the wavelet representation of an image are lost, the aim is to reconstruct an image that fits the known coefficients and minimizes the TV functional. In our terminology, given a suitable basis $(a_n)_n$ of $L^2(\Omega)$ and a source image u_0, this means solving

$$\min_{u\in L^2(\Omega)} \mathrm{TV}(u) + \mathcal{I}_V(u)$$

with

$$V = \{u \in L^2(\Omega) \,|\, (u, a_n)_{L^2} = (u_0, a_n)_{L^2} \text{ for all } n \in M\},$$

M being the index set of known coefficients. In [23], existence of a solution for this problem was established in the function space setting under the assumptions that $\Omega = \mathbb{R}^2$ and only finitely many coefficients are unknown. Numerical solution strategies for this, and a similar model with L^2 data fit, were presented in [23, 22, 52, 44]. In [54], the same model using nonlocal TV regularization was considered. In [29], the authors present the statement and numerical solution of a TV-wavelet denoising scheme whose formulation is also quite similar to those methods: Motivated by denoising with wavelet thresholding, the authors propose minimizing the TV functional subject to equality constraints on all wavelet coefficients with absolute value above a certain threshold.

However, even when considered solely as a method for wavelet-constrained optimization, our work differs significantly from those cited above. First, we use the TGV functional of arbitrary order as a regularization. Also, we are able to establish existence of a solution and optimality conditions in the case in which Ω is a *bounded Lipschitz domain* using *natural boundary extension* also in the function space setting. Additionally, we allow *infinitely many* wavelet coefficients to be unbounded and possible *interval constraints*. We also formulate the model for general biorthogonal wavelet bases from the very beginning, and our numerical solution scheme presented in [16] is different from those of previous works. Let us point out, however, that the assumptions of our work include the problem of wavelet inpainting; thus our method can also be seen as a generalization of the methods of [23, 29] using arbitrary order TGV regularization.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

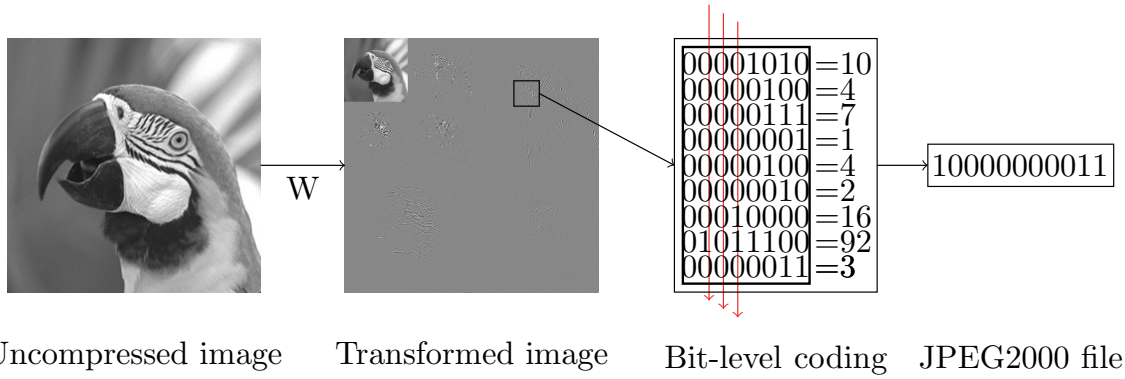

Figure 3. *Selected steps of JPEG* 2000 *compression including bit-level coding for one image component. See also* [40, *Figures* 10.16, 10.17].

Other methods that mainly focus on the concealment of wavelet data error due to transmission are featured in the works [38, 5, 25]. In [51], the aim is the reduction of artifacts due to tile separation of the image. We also refer the reader to [43] for a postprocessing method that attempts to improve reconstruction quality by reapplication of JPEG 2000 on shifted versions of the image.

The JPEG 2000 standard. We will now briefly discuss the JPEG 2000 compression procedure. For more information, we refer the reader to [40, 48, 49, 35] and the references therein.

Figure 3 gives a schematic overview of some main steps for JPEG 2000 compression that will be discussed in the following. As a first step, the image is split into color components and further into tiles, where each tile undergoes the same compression process. Next, a discrete wavelet transformation of arbitrary order is applied to each tile. Two types of wavelet transformation are possible within the standard—the *Cohen–Daubechies–Feauveau* (CDF) 9/7 and the *Le Gall* 5/3 wavelet transform (see [48, 35]). The numbers 9/7 and 5/3 indicate the support length of related filters. The resulting coefficients are then quantized depending on their importance for visual image quality. The values used for quantization are uniform on each subband, i.e., on each direction dependent part of each resolution level of each tile, and can in particular be obtained from the compressed code stream.

The quantized coefficients are then further split into different kinds of subunits, resulting finally in a set of code blocks. Each of these code blocks then undergoes a bit-level encoding consisting of three different passes. Starting from the highest nonzero bit-level, these three passes are repeated until the lowest bit-level has been encoded. This generates, for each code block, an independent bit-stream together with a set of valid truncation points (typically the last bit that has been encoded by each pass). Finally, the data from all code blocks is reorganized using mean-squared error estimations with respect to the original image. The result is a single bit-stream together with a set of possible truncation points which are, given a maximal number of bits to be saved, expected to be optimal in terms of peak signal-to-noise ratio (PSNR) (see [40, section 10.5.2], [35, section J.10]). When the compression rate is fixed by the user, this bit-stream is truncated to one of these points.

In the compressed JPEG 2000 file, the amount of information available in the bit-stream

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

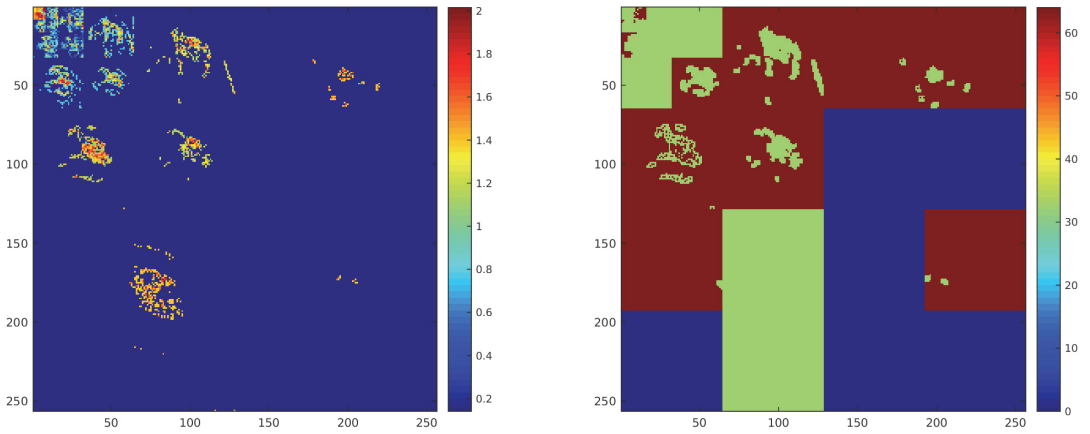

Figure 4. *Left: Wavelet coefficients of the brightness component of the bird image of Figure 3 (logarithmic scale). Right: Size of data intervals for the wavelet coefficients. (Note that 0, i.e., dark blue, indicates that the data is unbounded.)*

of one code block hence depends on the importance of the information in the code block for the PSNR rate. Thus, if, due to truncation, for one code block no bit-level information is left at all, the only information we can infer is that skipping information about its coefficients resulted in a better estimated PSNR rate than for other code blocks. However, since the original image is not known, we cannot use this information to obtain any estimate on its coefficients. Each individual coefficient could have been arbitrarily high as long as the overall information of the code block was less important for the PSNR value.

However, if at least one bit of coefficient information is left for a given code block, we can determine a bounded error interval for each of its coefficients as follows: As already explained, during compression, each code block is transformed into a bit-stream by repeating three passes, the *significance propagation pass*, the *magnitude refinement pass*, and the *cleanup pass* (see [35, Annex D]). Starting at the highest bit level, each pass follows predefined rules whether it encodes a particular bit or not. Thus, extracting the truncated, nonempty bit-stream and information about which pass has been performed last before truncation from the compressed file, we can determine, for each coefficient of the code block, up to which bit-level it has been encoded, i.e., its precision.

Using this knowledge, similar as for a JPEG compressed image, we can define a source value together with a (bounded) error interval for each coefficient of the code block. Thus, given any code block with nonzero information, we can define the set of its possible source coefficients again by bounded interval restrictions. As one can see in the numerical experiments in [16], this is possible for sufficiently many code blocks to keep the set of possible source images small and hence to achieve a good reconstruction quality. We refer the reader to Figure 4 for a visualization of the error bounds obtained from a JPEG 2000 compressed file.

Note that, in contrast to JPEG compression, the JPEG 2000 standard does not include explicit color subsampling. However, since due to the wavelet transformation the image is composed into a low-resolution part and a detail part, subsampling is still possible by skipping

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

the detail coefficients of the finest scale.

4.3.1. Continuous modeling. The definition of the wavelet transform for JPEG 2000 compression is based on either the Le Gall 5/3 or the CDF 9/7 wavelets. Let us detail the construction of a Riesz basis for $L^2(\Omega, \mathbb{R}^m)$ from these wavelets. We start with two finite length filter sequences $(h_n)_n$ and $(\tilde{h}_n)_n$ which yield either Le Gall 5/3 or CDF 9/7 wavelets and are defined in [26, Tables 6.1 and 6.2] for $N = \tilde{N} = 2$ and $N = \tilde{N} = 4$, respectively. As shown in [26], from both of these filter choices one can define scaling functions ϕ, $\tilde{\phi} \in L^2(\mathbb{R})$ and mother wavelets ψ, $\tilde{\psi} \in L^2(\mathbb{R})$ that allow, by translations and dilatations, the construction of biorthogonal Riesz bases. Indeed, defining for j, $k \in \mathbb{Z}$

$$\phi_{j,k}(x) = 2^{-j/2}\phi(2^{-j}x - k), \quad \psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k),$$

one obtains that, for any $R \in \mathbb{Z}$,

$$(\phi_{R,k})_k \cup (\psi_{j,k})_{j,k} \quad \text{for indices } k \in \mathbb{Z} \text{ and } j \leq R \tag{4.12}$$

is a Riesz basis of $L^2(\mathbb{R})$. Its dual basis is further denoted by

$$(\tilde{\phi}_{R,k})_k \cup (\tilde{\psi}_{j,k})_{j,k} \quad \text{for indices } k \in \mathbb{Z} \text{ and } j \leq R, \tag{4.13}$$

where $\tilde{\phi}_{j,k}$, $\tilde{\psi}_{j,k}$ are obtained from the dual scaling and wavelet functions $\tilde{\phi}$, $\tilde{\psi}$ again by translations and dilatations. We point out that, since we use only finite length filters for their construction, each of these basis elements has finite support length. Any signal $f \in L^2(\mathbb{R})$ can now be decomposed to the scale $R \in \mathbb{Z}$, using the bases (4.12) and (4.13), as

$$\begin{aligned} f = &\sum_{k\in\mathbb{Z}} (\tilde{\phi}_{k,R}, f)\phi_{k,R} + \sum_{k\in\mathbb{Z}, j\leq R} (\tilde{\psi}_{j,k}, f)\psi_{j,k} \\ = &\sum_{k\in\mathbb{Z}} (\phi_{k,R}, f)\tilde{\phi}_{k,R} + \sum_{k\in\mathbb{Z}, j\leq R} (\psi_{j,k}, f)\tilde{\psi}_{j,k}. \end{aligned} \tag{4.14}$$

The first sum in each of the terms can be interpreted as a low-resolution approximation of f, while the second contains detail information.

Next we want to obtain a Riesz basis of $L^2((0,1))$ from such a given Riesz basis of $L^2(\mathbb{R})$ which corresponds to symmetric boundary extension. For that purpose we apply a folding technique as in [27, section 2]. Given any compactly supported function $\eta \in L^2(\mathbb{R})$, we define its folded version $\eta^{\mathrm{f}} \in L^2((0,1))$ pointwise almost everywhere as

$$\eta^{\mathrm{f}}(x) = \sum_{n\in\mathbb{Z}} [\eta(x - 2n) + \eta(2n - x)].$$

Then, denoting by $u \in L^2((0,1))$ a function and by $\overline{u}$ its symmetric extension to all of $\mathbb{R}$, we get

$$\int_0^1 \eta^{\mathrm{f}} u = \int_{\mathbb{R}} \eta \overline{u}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Thus, testing $u \in L^2((0,1))$ with η^{f} corresponds to testing its symmetric extension with η. Using this technique and skipping redundant indices, it has been shown in [27, section 2] that for $R \in \mathbb{Z}$ the folded sequences

$$(\phi^{\mathrm{f}}_{R,k})_k \cup (\psi^{\mathrm{f}}_{j,k})_{j,k} \quad \text{and} \quad (\tilde{\phi}^{\mathrm{f}}_{R,k})_k \cup (\tilde{\psi}^{\mathrm{f}}_{j,k})_{j,k}$$

constitute Riesz bases of $L^2((0,1))$ in duality. Due to the support restriction, the folded bases contain only finitely many translations of the scaling functions.

From these two bases we then construct dual Riesz bases of $L^2((0,1)\times(0,1))$ by using tensor products of the basis elements, as done, for example, in [28, section 10.1] for orthogonal wavelet bases. Tensor products of two scaling functions give again a scaling function on $\mathbb{R}^2$, while tensor products of a wavelet with either a scaling function or another wavelet give a wavelet that resolves horizontal, vertical, or diagonal details, respectively. Grouping the scaling functions and the wavelets together and reindexing, we obtain, again for $R \in \mathbb{Z}$, two dual Riesz bases of $L^2((0,1)\times(0,1))$ written as

$$(\Phi_{R,k})_k \cup (\Psi_{j,k})_{j,k} \quad \text{and} \quad (\tilde{\Phi}_{R,k})_k \cup (\tilde{\Psi}_{j,k})_{j,k}. \tag{4.15}$$

Depending on our initial choice of filters $(h_n)_n$, $(\tilde{h}_n)_n$, these bases correspond to either Le Gall 5/3 or CDF 9/7 wavelets.

In order to apply the problem setting of section 3, we will need to ensure certain regularity assumptions on the Riesz basis; i.e., the dual basis must be contained in $\mathrm{BV}((0,1)\times(0,1))$. By construction (see [26]), all basis elements of the one dimensional dual basis in $L^2(\mathbb{R})$ can be expressed as finite linear combinations of translated, scaled versions of the scaling function $\tilde{\phi}$. Also, folded versions and tensor products of compactly supported BV functions are again in BV; hence it suffices to ensure regularity of $\tilde{\phi}$. In the case of Le Gall 5/3 filters the scaling function $\tilde{\phi}$ is just a piecewise linear spline (see [26, section 6.A], and note that there, ϕ and $\tilde{\phi}$ are interchanged) and thus is contained in $W^{1,1}(\mathbb{R})$. In the case of CDF 9/7 filters it has been shown, for example, in [49] that the scaling function corresponding to synthesis possesses a Sobolev regularity higher than 2 and in particular is also contained in $W^{1,1}(\mathbb{R})$.

We consider color images as functions in $L^2(\Omega,\mathbb{R}^3)$, where $\Omega = (0,k)\times(0,l)$ is a rectangular Lipschitz domain and k, l denote the number of tiles in which the images are split as part of compression. In contrast to JPEG compression, now also the definition of the basis used for reconstruction depends on the information obtained from a given compressed file. For each color component and tile encoded by a JPEG 2000 compressed file, we can now choose an appropriate resolution level and wavelet type and obtain a Riesz basis of $L^2((0,1)\times(0,1))$ as in (4.15). Using these bases, we can construct, for each color component, a block-wavelet basis of $L^2(\Omega)$. As in Remark 2.2, we obtain a Riesz basis of $L^2(\Omega,\mathbb{R}^3)$ together with a dual basis from these componentwise bases. We denote these bases by $(a_n)_n$ and $(\tilde{a}_n)_n$ and the corresponding basis transformation operators by A and $\tilde{A}$, respectively. Since each dual basis for each color component and tile is contained in $\mathrm{BV}(\Omega)$, so is $(\tilde{a}_n)_n$. As explained at the beginning of this section, we can further obtain data intervals $(J_n)_n$ such that all possible source images of the given JPEG 2000 compressed file must be contained in

$$U_D = \{u \in L^2(\Omega,\mathbb{R}^m) \,|\, (Au)_i \in J_i \text{ for all } i \in \mathbb{N}\}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In contrast to JPEG decompression, each of these intervals might also be unbounded. However, as the standard encodes the sign of a coefficient only when the first nonzero bit is encoded, no intervals are half-bounded and (EX_k) applies. During compression, the wavelet coefficients of the signal are quantized and rounded toward zero and a bit truncation is performed. The values used for quantization are uniform on each subband and are typically nondecreasing for higher subbands, meaning that coefficients corresponding to finer scales are saved with the same or less precision. Thus we again assume that the quantization values are bounded below by some $\epsilon > 0$, and, consequently, as the original sequence of coefficients is contained in ℓ^2, all but finitely many intervals must contain $[-\epsilon, \epsilon]$. This ensures that assumption (A) is valid, and thus again the existence result and the optimality conditions of Theorem 3.3 apply.

4.4. Variational image zooming. Apart from the data decompression models of subsections 4.2 and 4.3, we now consider the task of obtaining a high-resolution image from low-resolution data. A generic approach in this context is to perform a regularized inversion of a subsampling operation, i.e., given $\Omega = (0,1) \times (0,1)$, $L^2(\Omega, \mathbb{R}^3)$ the space of high-resolution color images and K a subsampling operator, one aims at solving

$$\min_{Ku = u_0} F(u),$$

where u_0 is the given low-resolution data and F is a regularization functional. Following this approach, we model subsampling as a linear operator mapping a function to a finite subset of its coefficients with respect to a Riesz basis. That is, given $(a_n)_n$ a Riesz basis of $L^2(\Omega, \mathbb{R}^3)$ and $N \subset \mathbb{N}$ a finite subset, we assume the subsampling operator K to be given as

$$Ku = ((a_n, u)_{L^2})_n \quad \text{for } n \in N.$$

As we will see, this can indeed be considered as a subsampling operation for images and covers also standard subsampling techniques such as averaging.

For regularization, we again use the TGV^k_α functional. With $A : L^2(\Omega, \mathbb{R}^3) \to \ell^2$ the basis transformation operator corresponding to $(a_n)_n$, the task of reconstructing a high-resolution image $u^\dagger \in L^2(\Omega, \mathbb{R}^3)$ from given low-resolution data $((a_n, u^\dagger)_n)_n$, for $n \in N$, then amounts to solving

$$\min_{u \in L^2(\Omega, \mathbb{R}^3)} \mathrm{TGV}^k_\alpha(u) + \mathcal{I}_{U_D}(u)$$

with

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^3) \,|\, (Au)_n = (a_n, u^\dagger)_{L^2} \text{ for all } n \in N\}.$$

In the notation of section 3 this means setting $J_n = \{(a_n, u^\dagger)_{L^2}\}$ for $n \in N$ and $J_n = \mathbb{R}$ else and corresponds to the case where the low resolution data is exactly given.

A particular case of this setting is given when $(a_n)_n$ results from a wavelet basis of $L^2(\Omega, \mathbb{R})$ and can be split into scaling functions $(\Phi_{j,k})_{j,k}$ and wavelet functions $(\Psi_{j,k})_{j,k}$. Having fixed a resolution level $R \in \mathbb{Z}$, U_D can then be defined as

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^3) \,|\, (\Phi_{R,k}, u)_{L^2} = (\Phi_{R,k}, u^\dagger)_{L^2} \text{ for all } k \in N\}$$

and indeed defines a low-resolution version of the original image $u^\dagger$; see Figure 5 for a visualization. This setting has already been discussed in [14] for the case of second order TGV

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

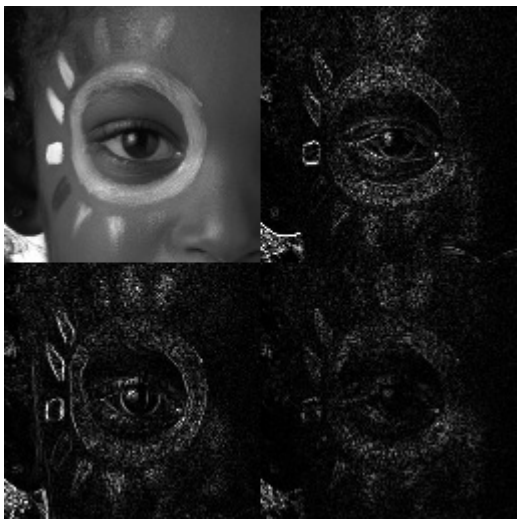

Figure 5. *Visualization of the low resolution image and the detail coefficients obtained from a high resolution image with one level of wavelet decomposition. Note that the wavelet coefficients have been rescaled for better visibility.*

regularization. As it allows any type of wavelet for the subsampling operation, it is still flexible in the choice of a subsampling operator. In particular, using the Haar wavelet corresponds to subsampling by averaging, and using the Le Gall 5/3 wavelet corresponds to the adjoint of bilinear interpolation as a subsampling operator (see [14]). We also refer the reader to [14] for a discussion of existing methods that are related to the present one.

Alternatively, one can also incorporate data from a given JPEG or JPEG 2000 compressed file in the zooming approach. Choosing the basis $(a_n)_n$ to be either a blockwise cosine basis or a tilewise wavelet basis with CDF 9/7 or Le Gall 5/3 wavelets, one can regard a compressed file as a finite number of truncated samples of the unknown image $u^\dagger \in L^2(\Omega, \mathbb{R}^3)$, define N to be the set of all coefficients for which data is available, and define the intervals $(J_n)_n$ to be the error bounds for these coefficients. Setting all remaining J_n to be all of $\mathbb{R}$, this yields a method for combined decompression and zooming of JPEG or JPEG 2000 compressed image files. We refer the reader to [13] for a discussion of this method in the case of JPEG files and second order TGV regularization.

In both of the above-discussed settings, all but finitely many intervals contain all of $\mathbb{R}$. Thus, for any choice of basis such that the dual basis is contained in $\mathrm{BV}(\Omega, \mathbb{R}^3)$, the assumptions (EX_k) and (A) are clearly satisfied, and all results of section 3 apply. Hence our general problem formulation is also directly applicable to a variational zooming as well as a combined decompression and zooming approach.

5. Conclusion.

Motivated by applications to image decompression, we have introduced a TGV regularized image reconstruction framework in a general function space setting. We have posed generic assumptions for which existence of a solution and optimality conditions for the resulting minimization problem were be obtained. These assumptions are quite general in the sense that an arbitrary Riesz basis together with a broad class of interval restrictions

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

can be used for data fidelity. This provides a common framework for a large class of problem settings in mathematical image processing, particular examples being the TGV regularized decompression of JPEG and JPEG 2000 images and a variational zooming model. After having established the theoretical foundations, the numerical realization and evaluation of these applications are the topic of a second paper [16], the content of which is strongly connected to that of the present paper.

Appendix A. Products of tensor spaces and related mappings. This appendix gives a short overview on functions mapping to products of tensor spaces, henceforth referred to as tensor fields. These spaces are needed to define the TGV functional for vector-valued data. The definitions and results stated in this subsection are straightforward generalizations of those presented in [17, 15] and are provided for the reader's convenience.

The space of symmetric tensors of order k is defined as

$$\mathrm{Sym}^k(\mathbb{R}^d) := \left\{ \xi : \left(\mathbb{R}^d\right)^k \to \mathbb{R} \,\middle|\, \xi : \ k - \text{linear and symmetric} \right\}, \tag{A.1}$$

respectively, with the scalar product

$$\xi \cdot \eta = \sum_{p \in \{1,\dots d\}^k} \xi(e_{p_1}, \dots, e_{p_k}) \eta(e_{p_1}, \dots, e_{p_k}), \tag{A.2}$$

for $\xi, \eta \in \mathrm{Sym}^k(\mathbb{R}^d)$, and the induced norm $|\xi| = \sqrt{\xi \cdot \xi}$. For a given, sufficiently smooth, tensor field ξ, its lth derivative can be identified with a (nonsymmetric) $(k+l)$ tensor field $\nabla^l \otimes \xi$, defined by

$$(\nabla^l \otimes \xi)(x)(a_1, \dots, a_{k+l}) = \left(\mathrm{D}^l \xi(x)(a_1, \dots, a_l) \right)(a_{l+1}, \dots, a_{k+l}),$$

where $\mathrm{D}^l \xi : \Omega \to \mathcal{L}^l\big(\mathbb{R}^d, \mathrm{Sym}^k(\mathbb{R}^d)\big)$ denotes the lth Fréchet derivative of ξ and $\mathcal{L}^l(X, Y)$ the space of l-linear and continuous mappings from X^l to Y. Further, we define a symmetrized derivative of a smooth tensor field $\xi : \Omega \to \mathrm{Sym}^k(\mathbb{R}^d)$ that can be identified with a symmetric tensor field:

$$\mathcal{E}^l \xi = |||(\nabla^l \otimes \xi). \tag{A.3}$$

Here $|||\eta$ denotes the symmetrization of a given tensor η defined by

$$(|||\eta)(a_1, \dots, a_k) = \frac{1}{k!} \sum_{\pi \in S_k} \eta(a_{\pi(1)}, \dots a_{\pi(k)}),$$

where S_k is the set of all permutations of $\{1, \dots, k\}$.

We also use the notion of l-divergence of a sufficiently smooth $(k+l)$ tensor field:

$$\mathrm{div}^l \eta = \mathrm{tr}^l(\nabla^l \otimes \eta)$$

for $\eta \in \mathrm{Sym}^{k+l}(\mathbb{R}^d)$, where, for $\xi \in \mathrm{Sym}^k(\mathbb{R}^d)$,

$$\mathrm{tr}(\xi) \in \mathrm{Sym}^{k-2}(\mathbb{R}^d), \quad \mathrm{tr}(\xi)(a_1, \dots, a_{k-2}) = \sum_{i=1}^d \xi(e_i, a_1, \dots, a_{k-2}, e_i).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that by definition of the trace operator, the divergence of η is symmetric.

Likewise, we equip the space $\mathrm{Sym}^k(\mathbb{R}^d)^m$ containing m-tuples of symmetric tensors, i.e.,

$$\mathrm{Sym}^k(\mathbb{R}^d)^m = \left\{ \xi = (\xi_1, \dots, \xi_m) \,\middle|\, \xi_i \in \mathrm{Sym}^k(\mathbb{R}^d),\ i \in \{1, \dots, m\} \right\}, \tag{A.4}$$

with the inner product and norm

$$\xi \cdot \eta = \sum_{i=1}^{m} \xi_i \cdot \eta_i \qquad \text{and} \qquad |\xi|^2 = \xi \cdot \xi. \tag{A.5}$$

For sufficiently smooth m-tuples of symmetric tensor fields, the differentiation operators $\nabla, \mathcal{E}, \mathrm{div}$ are defined componentwise. The spaces

$$L^p(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m), \quad \mathcal{C}_c^l(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m), \quad \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$$

are defined in the usual way, where we use tensor norm $|\cdot|$ as in (A.5) as vector norm. Spaces of measures and distributions Ω are defined by duality as

$$\mathcal{M}(\Omega, \mathrm{Sym}^{k+1}(\mathbb{R}^d)^m) = \left(\overline{\mathcal{C}_c\left(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m\right)}^{\|\cdot\|_\infty} \right)^*,$$
$$\mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m) = \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)^*.$$

For $\xi \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$, $\eta \in \mathcal{D}(\Omega, \mathrm{Sym}^{k+1}(\mathbb{R}^d)^m)$ is called the weak symmetrized derivative of ξ if

$$\langle \eta, \zeta \rangle = -\langle \xi, \mathrm{div}\, \zeta \rangle$$

for all $\zeta \in \mathcal{C}_c^1(\Omega, \mathrm{Sym}^{k+1}(\mathbb{R}^d)^m)$. In this case we denote $\mathcal{E}\xi = \eta$.

We will need that a distribution is represented by an L^1 function if its symmetrized gradient can be represented by a Radon measure. This has been shown in [15] for tensor spaces and can be generalized to products of tensor spaces as follows.

Proposition A.1. *If for $u \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$ we have $\mathcal{E}u \in \mathcal{M}(\Omega, \mathrm{Sym}^{k+1}(\mathbb{R}^d)^m)$, then $u \in L^1(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$.*

Proof. Given any distribution $u \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$, we apply the result of [15] to the distributions $u^i \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ defined by

$$\langle u^i, \phi \rangle = \langle u, (\ \underbrace{0, \dots, 0}_{(i-1)\text{ times}}\ , \phi, \ \underbrace{0 \dots, 0}_{(m-i)\text{ times}}\ ) \rangle$$

for $\phi \in \mathcal{C}_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$. ∎

REFERENCES

[1] F. Alter, S. Durand, and J. Froment, *Adapted total variation for artifact free decompression of JPEG images*, J. Math. Imaging Vision, 23 (2005), pp. 199–211.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[2] H. A. Aly and E. Dubois, *Image up-sampling using total-variation regularization with a new observation model*, IEEE Trans. Image Process., 14 (2005), pp. 1647–1659.
[3] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, Oxford, UK, 2000.
[4] H. Attouch and H. Brezis, *Duality for the sum of convex functions in general Banach spaces*, in Aspects of Mathematics and Its Applications, North–Holland Math. Library 34, North–Holland, Amsterdam, 1986, pp. 125–133.
[5] L. Atzori, *JPEG* 2000*-coded image error concealment exploiting convex sets projections*, IEEE Trans. Image Process., 14 (2005), pp. 487–498.
[6] A. Auslender, *Noncoercive optimization problems*, Math. Oper. Res., 21 (1996), pp. 769–782.
[7] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
[8] M. Benning, C. Brune, M. Burger, and J. Müller, *Higher-order TV methods—enhancement via Bregman iteration*, J. Sci. Comput., 54 (2013), pp. 269–310.
[9] L. Bottou, P. Haffner, P. G. Howard, P. Simard, Y. Bengio, and Y. L. Cun, *High quality document image compression with "DjVu,"* J. Electron. Imaging, 7 (1998), pp. 410–425.
[10] K. Bredies, *Symmetric tensor fields of bounded deformation*, Ann. Mat. Pura Appl. (4), 192 (2013), pp. 815–851.
[11] K. Bredies, *Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty*, in Efficient Algorithms for Global Optimization Methods in Computer Vision, Lecture Notes in Comput. Sci. 8293, Springer, Berlin, Heidelberg, 2014, pp. 44–77.
[12] K. Bredies and M. Holler, *A total variation–based JPEG decompression model*, SIAM J. Imaging Sci., 5 (2012), pp. 366–393.
[13] K. Bredies and M. Holler, *Artifact-free decompression and zooming of JPEG compressed images with total generalized variation*, in Computer Vision, Imaging and Computer Graphics. Theory and Application, Comm. Comput. Inform. Sci. 359, Springer, Berlin, 2013, pp. 242–258.
[14] K. Bredies and M. Holler, *A TGV regularized wavelet based zooming model*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Comput. Sci. 7893, Springer, Berlin, 2013, pp. 149–160.
[15] K. Bredies and M. Holler, *Regularization of linear inverse problems with total generalized variation*, J. Inverse Ill-Posed Probl., 22 (2014), pp. 871–913.
[16] K. Bredies and M. Holler, *A TGV-based framework for variational image decompression, zooming, and reconstruction. Part* II: *Numerics*, SIAM J. Imaging Sci., 8 (2015), pp. 2851–2886.
[17] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., 3 (2010), pp. 492–526.
[18] K. Bredies and D. Lorenz, *Mathematische Bildverarbeitung*, Vieweg+Teubner, Berlin, 2011.
[19] K. Bredies and T. Valkonen, *Inverse problems with second-order total generalized variation constraints*, in Proceedings of SampTA 2011: 9th International Conference on Sampling Theory and Applications, Singapore, 2011.
[20] W. K. Carey, D. B. Chuang, and S. S. Hemami, *Regularity-preserving image interpolation*, IEEE Trans. Image Process., 8 (1999), pp. 1293–1297.
[21] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 88–97.
[22] R. H. Chan, J. Yang, and X. Yuan, *Alternating direction method for image inpainting in wavelet domains*, SIAM J. Imaging Sci., 4 (2011), pp. 807–826.
[23] T. F. Chan, J. Shen, and H.-M. Zhou, *Total variation wavelet inpainting*, J. Math. Imaging Vision, 25 (2006), pp. 107–125.
[24] T. Chen, H. R. Wu, and B. Qiu, *Image interpolation using across-scale pixel correlation*, in Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01), Vol. 3, IEEE, Washington, DC, 2001, pp. 1857–1860.
[25] P. C. Chung, *A JPEG* 2000 *error resilience method using uneven block-sized information included markers*, IEEE Trans. Circuits Syst. Video Tech., 15 (2005), pp. 420–424.
[26] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[27] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harmon. Anal., 1 (1993), pp. 54–81.
[28] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61, SIAM, Philadelphia, 1992.
[29] S. Durand and J. Froment, *Reconstruction of wavelet coefficients using total variation minimization*, SIAM J. Sci. Comput., 24 (2003), pp. 1754–1767.
[30] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.
[31] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, FL, 1992.
[32] V. Girault and P.-A. Raviart, *Finite Element Method for Navier-Stokes Equation*, Springer, New York, 1986.
[33] B. Goldluecke, E. Strekalovskiy, and D. Cremers, *The natural total variation which arises from geometric measure theory*, SIAM J. Imaging Sci., 5 (2012), pp. 537–563.
[34] M. Grasmair, M. Haltmeier, and O. Scherzer, *The residual method for regularizing ill-posed problems*, Appl. Math. Comput., 218 (2011), pp. 2693–2710.
[35] Joint Bilevel Image Experts Group and Joint Photographic Experts Group, *JPEG* 2000 *image coding system*, ISO/IEC 15444-1, 2000.
[36] M. Holler, *Higher Order Regularization for Model Based Data Decompression*, Ph.D. thesis, University of Graz, Graz, Austria, 2013.
[37] N. Kaulgud and U. B. Desai, *Image zooming: Use of wavelets*, in Super-Resolution Imaging, The International Series in Engineering and Computer Science 632, Springer, New York, 2002, pp. 21–44.
[38] P. J. Lee, *Error concealment algorithm using interested direction for JPEG* 2000 *image transmission*, IEEE Trans. Consumer Electron., 49 (2003), pp. 1395–1401.
[39] F. Malgouyres and F. Guichard, *Edge direction preserving image zooming: A mathematical and numerical analysis*, SIAM J. Numer. Anal., 39 (2001), pp. 1–37.
[40] S. Mallat, *A Wavelet Tour of Signal Processing*, Elsevier, New York, 2009.
[41] S. Mallat and G. Yu, *Super-resolution with sparse mixing estimators*, IEEE Trans. Image Process., 19 (2010), pp. 2889–2900.
[42] A. Nosratinia, *Enhancement of JPEG-compressed images by re-application of JPEG*, J. VLSI Signal Process., 27 (2001), pp. 69–79.
[43] A. Nosratinia, *Postprocessing of JPEG*-2000 *images to remove compression artifacts*, IEEE Signal Process. Lett., 10 (2003), pp. 296–299.
[44] R. Oktem, *Regularization-based error concealment in JPEG* 2000 *coding scheme*, IEEE Signal Process. Lett., 14 (2001), pp. 956–959.
[45] S. Ono and I. Yamada, *Optimized JPEG image decompression with super-resolution interpolation using multi-order total variation*, in Proceedings of the 20th IEEE International Conference on Image Processing (ICIP), IEEE, Washington, DC, 2013, pp. 474–478.
[46] M.-Y. Shen and C.-C. J. Kuo, *Review of postprocessing techniques for compression artifact removal*, J. Visual Commun. Image Rep., 9 (1998), pp. 2–14.
[47] S. Singh, V. Kumar, and H. K. Verma, *Reduction of blocking artifacts in JPEG compressed images*, Digital Signal Process., 17 (2007), pp. 225–243.
[48] A. Skodras, C. Christopoulos, and T. Ebrahimi, *The JPEG* 2000 *still image compression standard*, IEEE Signal Process. Mag., 18 (2001), pp. 36–58.
[49] M. Unser and T. Blu, *Mathematical properties of the JPEG* 2000 *wavelet filters*, IEEE Trans. Image Process., 12 (2003), pp. 1080–1090.
[50] G. K. Wallace, *The JPEG still picture compression standard*, Commun. ACM, 34 (1991), pp. 30–44.
[51] J. Wei, M. Pickering, M. Frator, J. Arnold, J. Boman, and W. Zeng, *Boundary artefact reduction using odd tile length and low pass first convention (OTLPF)*, IEEE Trans. Image Process., 14 (2001), pp. 1033–1042.
[52] Y.-W. Wen, R. H. Chan, and A. M. Yip, *A primal-dual method for total variaton-based wavelet domain inpainting*, IEEE Trans. Image Process., 21 (2011), pp. 106–114.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[53] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 2001.
[54] X. Zhang and T. F. Chan, *Wavelet inpainting by nonlocal total variation*, Inverse Probl. Imaging, 4 (2010), pp. 191–210.
[55] S. Zhong, *Image coding with optimal reconstruction*, in Proceedings of the IEEE International Conference on Image Processing, Vol. 1, IEEE, Washington, DC, 1997, pp. 161–164.
[56] W. P. Ziemer, *Weakly Differentiable Functions*, Springer, New York, 1989.

Downloaded 05/03/19 to 143.50.47.147. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.