SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

A TGV-based framework for variational image decompression, zooming and reconstruction. Part I: Analytics

K. Bredies M. Holler

SFB-Report No. 2015-009 May 2015

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

A TGV-based framework for variational image decompression, zooming and reconstruction. Part I: Analytics*

Kristian Bredies and Martin Holler

May 22, 2015

Abstract

A variational model for image reconstruction is introduced and analyzed in function space. Specific about 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, being 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, non-coercive existence result and optimality conditions, including a characterization of the subdifferential of the TGV functional, are obtained in the general setting.

Mathematical subject classification: 94A08 49M29 65F22 49K30
Keywords: Image reconstruction, Total Generalized Variation, JPEG decompression, JPEG 2000 decompression, Variational zooming, Optimality conditions.

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 for 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 [47], where we aim at reconstructing an image subject to interval constraints of the block-cosine transform of the image. As image prior, the *Total Generalized Variation* (TGV) [13] functional of arbitrary order is incorporated and the model is formulated for multichannel images, in particular 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

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

*Support by the Austrian Science Fund (FWF) special research grant SFB-F32 *Mathematical Optimization and Applications in Biomedical Sciences* is gratefully acknowledged.

with A a Riesz-basis transformation operator and $(J_i)_i$ non-empty closed intervals. In the application to JPEG decompression, A is a block-wise cosine transform and, given a JPEG compressed file, U_D is the set of all images that, when compressed at the given rate, result in the same file.

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 [12] deals with the numerical realization in a discrete setting.

Our work is motivated by previous papers on a TV based JPEG decompression model [8] and on applications of TGV for regularized JPEG decompression [9] and for wavelet based zooming [10]. The present paper provides, for the first time, a unified framework for all these settings and significantly extents the previous works both on the analysis and application side. 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, non-coercive 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 third application, a variational zooming method can also be derived from the general setting. These applications are realized in the second paper [12], 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 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, multi-core 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 [42, 9, 8, 44, 1, 39, 43, 52]. 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 [49, 51, 18, 41, 19]. As for zooming techniques, we refer to [38] for an overview and to [10, 2, 17, 36, 34, 20, 16] 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, 4.4.

In contrast to application oriented approaches, there are, besides the above-mentioned TV based model of [8], 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 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 and deals with the algorithmic realization of the applications and provides experimental results [12].

2 Functional-analytic background

The aim of this section is to shortly introduce notation and mathematical concepts that are of particular relevance for this work, such as functions of bounded variation, the total generalized variation functional as well as 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.

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 a $\mathbb{R}^m$-valued finite Radon measure on Ω if it is σ-additive and $\mu(\emptyset) = 0$. We denote the variation of a $\mathbb{R}^m$-valued finite Radon measures 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*

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

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

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

and

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

Functions of bounded variation are well known, in particular in the field of mathematical image processing, and have been extensively studied in the literature. We repeat just some properties of functions of bounded variation that are most relevant to our work and refer to [3, 28, 53] 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*

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

Proposition 2.3. *The functional* TV *is proper, convex and lower semi continuous 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 functional (TGV)

In this subsection we introduce the total generalized variation functional (TGV) for vector-valued functions. It will serve as 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 has originally been introduced in [13] and a generalization to the vector-valued case has been presented in [15]. We refer to [13] for a more detailed motivation and further properties and to [11] for its analysis in the context of inverse problems.

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 to [13] for an introduction of these spaces in the context of the TGV functional and provide, for the readers 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 $\alpha = (\alpha_0, \dots, \alpha_{k-1}) \in (0, \infty)^k$, for $u \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, as*

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

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 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 [15]). 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 to [30] and [14, 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 total generalized variation functional is finite.

Definition 2.4. *We define*

$$\begin{aligned}\mathrm{BGV}^k(\Omega,\mathbb{R}^m) &= \left\{u\in L^1(\Omega,\mathbb{R}^m)\mid \mathrm{TGV}^k_\alpha(u)<\infty\right\},\\ \|u\|_{\mathrm{BGV}^k} &= \|u\|_1+\mathrm{TGV}^k_\alpha(u).\end{aligned} \tag{2}$$

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 [15, 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 similar to the scalar case or follow from the equivalence of Proposition 2.6 (see [13]).

Proposition 2.7. *The following statements hold:*

1. TGV^k_α *is a semi-norm 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, lower semi-continuous on each* $L^p(\Omega,\mathbb{R}^m)$, $1\le p\le\infty$,
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 that* 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 [11] 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 [11], thus we provide only a short sketch:

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

$$\mathrm{TGV}^k_\alpha(u) = \min_{\substack{v_i \in \mathrm{BD}(\Omega, \mathrm{Sym}^i(\mathbb{R}^d)^m), \\ i=1,\dots,k, \\ v_0 = u,\, v_k = 0}} \sum_{i=1}^{k} \alpha_{k-i} \|\mathcal{E} v_{i-1} - v_i\|_{\mathcal{M}}. \tag{3}$$

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 [7], and $\mathcal{E}v$ denotes the symmetrized derivative of a tensor field v, see the Appendix.

Sketch of proof. Defining

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

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 semi-continuous functionals

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

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 (3) and using Proposition A.1 in the Appendix, i.e.,

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

the assertion follows. □

The basic result for topological equivalence follows immediately from the scalar case, as shown in [11], 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 $C > 0$, only depending on k, m, α, Ω and P_{k-1} such that*

$$\|\mathrm{D}u\|_{\mathcal{M}} \le C\big(\|u\|_1 + \mathrm{TGV}^k_\alpha(u)\big) \qquad \text{as well as} \qquad \|u - P_{k-1}u\|_{d/(d-1)} \le C\,\mathrm{TGV}^k_\alpha(u) \tag{4}$$

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 notion of strict convergence and weak-star convergence play an important role in this context and we refer to [11, 7] 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*

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

Proof. A straight forward adaption of [11, 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 [7, Theorem 4.16]), we can transfer L^p–mollification arguments, see [28, 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 [28, 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.

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 $\mathrm{span}(\{a_n | n \in \mathbb{N}\})$ 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*

$$A \sum_{n \in \mathbb{N}} c_n^2 \leq \Big\| \sum_{n \in \mathbb{N}} c_n a_n \Big\|_H^2 \leq B \sum_{n \in \mathbb{N}} c_n^2. \tag{5}$$

If $(a_n)_n$ is an orthonormal basis, equation (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 also $(\tilde{a}_n)_n$ is a Riesz basis of H and*

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

Proof. See [50, Theorem 1.9]. □

As can be shown, the notion of Riesz basis is the most general basis concept that ensures a sequence to be complete and the resulting basis transformation to be continuous and continuously invertible. Hence Riesz bases fit very well to our data modeling in a general image reconstruction setting later on. There we mainly deal with component-wise 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

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

with (J_n^i) being closed intervals.

The following remark emphasizes the connection of m component-wise Riesz bases to Riesz bases in $L^2(\Omega, \mathbb{R}^m)$. In particular, also interval restrictions as 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 $1 \le i \le m$, $n \in \mathbb{N}$,

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

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.

3 The general reconstruction model

This is the main section of the work, where we study the general, TGV_α^k 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. We will start with a brief motivation followed by a definition and analysis of the model in 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, 8]. For a detailed introduction to this topic we ask for the readers patience until Subsection 4.2.

Given an image u, the main step of JPEG compression is a transformation by a blockwise cosine transformation operator followed by quantization to integer values. As a result, in the compressed JPEG file, the image u is described by quantized integers of its coefficients for a basis representation with respect to a blockwise cosine basis. Due to quantization, this data does not provide enough information to determine a unique source image of the compression process. But it is possible to define a set of basis coefficient data, whose quantization would coincide with the

compressed image 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 any given compressed JPEG file by

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

Thus, reconstructing an image from a given compressed JPEG file requires to choose one element of U_D as reconstruction. This choice can be motivated by a predefined image model, which is in our case realized by using the TGV functional as regularization term.

Keeping this application in mind, we now consider the following minimization problem

$$\min_{u \in L^2(\Omega, \mathbb{R}^m)} \mathrm{TGV}_\alpha^k(u) + \mathcal{I}_{U_D}(u), \tag{6}$$

where we use the assumptions

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

By $\mathcal{I}_{U_D}$ we denote the convex indicator function of U_D, i.e.,

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

Thus, solving the minimization problem (6) amounts to finding a function in U_D minimizing TGV_α^k.

Thinking again of JPEG decompression, the operator A in assumption (A) will be the BDCT operator, while $(a_n)_n$ will be an orthonormal blockwise cosine basis. The dimension of the image space, $m \in \mathbb{N}$, reflects the number of image components, typically $m = 3$ for color and $m = 1$ for grayscale images.

The assumption that U_{int} has non-empty interior can be seen as a generalization of the assumption that the data set U_D has non-empty interior, which is satisfied in the application to JPEG decompression, and will be needed for the analysis of the model. The more general assumption is motivated by the application of the model to zooming problems, where typically finitely many of the intervals $(J_n)_n$ will consist only of a single point.

At last let us emphasize that in particular the assumption of $(a_n)_n$ being 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 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 (6) under assumption (A). First note that properness of the objective function follows easily from U_{int} having non-empty 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.

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

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 arbitrary 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 non-trivial 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 (6).

(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.

In fact, given (H2) and that F is bounded below, it has been shown in [6] that a slightly weaker version of (H1) is even 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* (6).

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 onto 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

$$\|x_n\|_{L^2}((x, a_i)_{L^2} + \epsilon_{n,i}) = (x_n, a_i)_{L^2} \in J_i. \tag{7}$$

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 (7) 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 non-empty interior was also not needed. However, in order to get existence of a non-trivial 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* (6)*.*

Remark 3.3. *It can also be seen from the proof of Proposition 3.1 that (EX_k) was necessary only to assure (H1). Hence, the weaker version of (H1) presented in [6] is necessary and sufficient for existence of a solution to* (6)*. 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* (6) *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 that 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 to [33].*

3.3 Optimality Conditions

Having obtained existence of a solution to (6) 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,
$$u^* = \arg\min_u F(u) \quad \Leftrightarrow \quad 0 \in \partial F(u^*).$$
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 $H(\mathrm{div}; \Omega)$, as described for example in [29, Chapter 1], and also many properties can easily be generalized.

Definition 3.1. *Let $1 \leq q < \infty$, $g \in L^q(\Omega, \mathrm{Sym}^l(\mathbb{R}^d)^m)$. 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 $\phi \in \mathcal{C}^\infty_c(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$*

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

Furthermore, we define

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

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

Remark 3.5. *Density of $\mathcal{C}^\infty_c(\Omega, \mathrm{Sym}^{l-1}(\mathbb{R}^d)^m)$ 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 $0 \leq l \leq k$ 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 $1 \leq q < \infty$,*

$$W^q_0(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m) = \overline{\mathcal{C}^\infty_c(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)}^{\|\cdot\|_{W^q(\mathrm{div}^k)}}.$$

We now proceed towards a characterization of $\partial \mathrm{TGV}^k_\alpha$ by first describing the convex conjugate (or polar) of the TGV^k_α functional.

Proposition 3.2. *The convex conjugate of TGV^k_α, denoted by*

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

has the form

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

where

$$C_\alpha^k := \left\{ \operatorname{div}^k \xi \,\middle|\, \xi \in \mathcal{C}_c^k(\Omega, \operatorname{Sym}^k(\mathbb{R}^d)^m),\ \|\operatorname{div}^l \xi\|_\infty \le \alpha_l,\ l = 0, ..., k-1 \right\}, \tag{8}$$

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

Proof. This follows easily from convexity and lower semi-continuity 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 [27, Propositions 3.2 and 4.1]),

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

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*

$$\overline{C_\alpha^k} = \left\{ \operatorname{div}^k g \,\middle|\, g \in W_0^{p'}(\operatorname{div}^k; \Omega, \operatorname{Sym}^k(\mathbb{R}^d)^m),\ \ \|\operatorname{div}^l g\|_\infty \le \alpha_l,\ l = 0, ..., k-1 \right\} := K_\alpha^k. \tag{9}$$

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

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

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

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

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

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

is convex and closed – and therefore weakly closed. Since the sequence $((g_{n_i}, \operatorname{div} g_{n_i}, \ldots, \operatorname{div}^k g_{n_i}))_n$ is contained in this set and converges weakly to $(g, \operatorname{div} g, \ldots, \operatorname{div}^k g)$ 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 $\mathrm{TGV}_\alpha^{k\,*}(\mathrm{div}^k g) = 0$ and hence $\mathrm{div}^k g \in \overline{C_c^k}$. 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 $v \in \mathrm{BD}(\Omega, \mathrm{Sym}^{k-l}(\mathbb{R}^d)^m)$,

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

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 [13, Proposition 2.1], for $\phi \in C^\infty(\overline{\Omega}, \mathrm{Sym}^{k-1}(\mathbb{R}^d)^m)$ and $\psi \in C_c^\infty(\Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m)$,

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

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

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

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

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

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 [27], 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*

$$\begin{cases} u \in \mathrm{BV}(\Omega, \mathbb{R}^m) \text{ and there exists } g \in W_0^{p'}(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^d)^m) \text{ such that } \|\mathrm{div}^l g\|_\infty \leq \alpha_l, \\ l = 0, \dots, k-1,\ u^* = \mathrm{div}^k g \text{ and} \\ \qquad \mathrm{TGV}_\alpha^k(u) = \displaystyle\int_\Omega u\, \mathrm{div}^k g. \end{cases} \tag{10}$$

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*

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

are both continuous and possess continuous inverses with

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

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 $n \in \mathbb{N}$,

$$\begin{cases} \lambda_n \geq 0 & \text{if} \quad (Au)_n = \sup(J_n) \neq \inf(J_n), \\ \lambda_n \leq 0 & \text{if} \quad (Au)_n = \inf(J_n) \neq \sup(J_n), \\ \lambda_n = 0 & \text{if} \quad (Au)_n \in \operatorname{int}(J_n), \\ \lambda_n \in \mathbb{R} & \text{if} \quad (Au)_n = \inf(J_n) = \sup(J_n). \end{cases} \tag{11}$$

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 [24, 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, by a straightforward argument, P_D can be reduced to a component-wise projection;

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

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

3.3.3 Additivity of the subdifferential

At last, 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),

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

and

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

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

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

$$\partial(\mathrm{TGV}^k_\alpha + \mathcal{I}_{U_D})(u) = \partial\mathrm{TGV}^k_\alpha(u) + \partial\mathcal{I}_{U_D}(u). \tag{12}$$

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 $\mathcal{I}_{U_{\text{int}}}$ in at least one point $u \in \mathrm{BV}(\Omega, \mathbb{R}^m) \cap U_D$ allows to apply [27, Proposition I.5.6], and to assure that

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

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

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

for which it is sufficient to show that $\mathrm{dom}(\mathrm{TGV}^k_\alpha) - \mathrm{dom}(\mathcal{I}_{U_{\text{point}}}) = L^2(\Omega, \mathbb{R}^m)$. 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

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

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

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

Again, since

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

is always satisfied, the assertion is proved. □

3.3.4 Optimality system

The previous results now allow to derive an optimality system:

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

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

and the following are equivalent

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

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) $\|\mathrm{div}^l g\|_\infty \leq \alpha_l$, $l = 0, ..., k-1$,

   (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* $n \in \mathbb{N}$,

$$\begin{cases} \lambda_n \geq 0 \text{ if } (A\hat{u})_n = \sup(J_n) \neq \inf(J_n), \\ \lambda_n \leq 0 \text{ if } (A\hat{u})_n = \inf(J_n) \neq \sup(J_n), \\ \lambda_n = 0 \text{ if } (A\hat{u})_n \in \mathrm{int}(J_n), \end{cases}$$

   (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*

   (a) $\|\mathrm{div}^l g\|_\infty \leq \alpha_l$, $l = 0, ..., k-1$,

   (b) $\mathrm{TGV}^k_\alpha(\hat{u}) = -\int_\Omega \hat{u}\, \mathrm{div}^k g$,

   (c) *For all* $n \in \mathbb{N}$,

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

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 $g \in W^2_0(\mathrm{div}^k; \Omega, \mathrm{Sym}^k(\mathbb{R}^2)^m)$ satisfying 2.(a) such that $z_1 = -\mathrm{div}^k g$ and $\mathrm{TGV}^k_\alpha(\hat{u}) = -\int_\Omega \hat{u}\, \mathrm{div}^k g$. Clearly, we have $\mathrm{div}^k g = z_2$. By Proposition 3.5 there exists $\lambda = (\lambda_n)_n \in \ell^2$, satisfying the element-wise conditions in 2.(c), such that $\mathrm{div}^k g = A^*\lambda = \sum_{n\in\mathbb{N}} \lambda_n a_n$, 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 $-\mathrm{div}^k g \in \partial \mathrm{TGV}^k_\alpha(\hat{u})$ (Theorem 3.1), while 2.(c) together with $\hat{u} \in U_D$ imply that $\mathrm{div}^k g \in \partial \mathcal{I}_{U_D}(\hat{u})$ (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 $(J_n)_n$ non-empty, closed intervals, define*

$$U_D := \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Bu, e_n)_{\ell^2} \in J_n \text{ for all } n \in N\}. \tag{13}$$

Then there exist non-empty, 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

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

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 [50, Theorem 1.9], it suffices to show that $(a_i)_i$ and $(\tilde{a}_i)_i$ both have dense linear spam, 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 $w_1 = 0$. Similar, $0 = (\tilde{a}_n, w_1)_{L^2} = (B^{-1}e_n, w_1)_{L^2}$ 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 spam. 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 bases. 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 $C_1, C_2 > 0$,

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

and

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

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

At last 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 $\tilde{U}_D = \{u \in L^2(\Omega, \mathbb{R}^m) \,|\, (Au)_n \in \tilde{J}_n \,\forall n \in \mathbb{N}\}$ we have

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

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 data which is close to the true, unknown data $d^\dagger$. 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_α^k regularization and reconstruct a signal $u^* \in L^2(\Omega, \mathbb{R}^m)$ which solves

$$\min_{F(u)\in D} \mathrm{TGV}_\alpha^k(u). \tag{14}$$

This setting is related to residual methods for inverse problems and we refer to [31] 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 non-uniqueness, 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 non-empty 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 non-trivial 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.

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

- 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 (14). 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.

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 [8, 1, 52] are most related. Also, a discrete version of the problem using second order TGV regularization has already been published in [9] and in [42]. We further refer to [8, 39, 44, 43] 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 to [8, 9] and for a more detailed explanation of the JPEG compression procedure to [47].

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 hu-

man visual system is less sensitive to color than to brightness oscillations, as 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 linear combination of different frequencies. Again, there is empirical evidence that the human visual system is less sensitive to high frequency variations than to low frequency variations. Consequently, the coefficients representing the cosine frequencies are quantized by pointwise division of each 8×8 pixel block by a predefined quantization matrix reflecting this empirical observation. The resulting data is then rounded to integer and, after further lossless compression, 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 dequantization, 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 a set of possible source images. Indeed, since, besides the quantized integer coefficient data $d = (d^c_{i,j})$, also the quantization matrix $Q = (Q^c_{i,j})$ can be obtained from the compressed file, it is possible to define interval bounds

$$J^c_{i,j} = \left[Q^c_{i,j}(d^c_{i,j} - \frac{1}{2}), Q^c_{i,j}(d^c_{i,j} + \frac{1}{2})\right] \tag{15}$$

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

$$D = \{(z^c_{i,j}) \,|\, z^c_{i,j} \in J^c_{i,j} \text{ for all } i, j, c\}. \tag{16}$$

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 ith and jth coefficient of the block-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.

4.2.1 Continuous modeling

We consider color images 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.

Sub-sampled 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 component-wise 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 [8]:

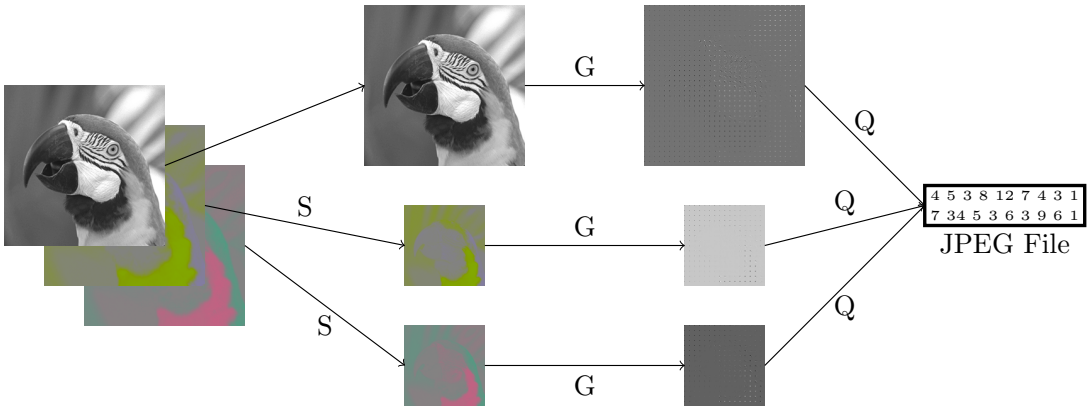

Figure 2: Scheme of JPEG compression procedure. Here, S denotes a subsampling operation, G a blockwise discrete cosine transformation and Q a quantization to integer, i.e., a blockwise division through a predefined quantization matrix followed by rounding to integer.

Definition 4.1 (Block-wise 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

$$b_{n,m}(x, y) = \lambda_n \lambda_m \cos(nx\pi) \cos(my\pi), \tag{17}$$

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*

$$g^{i,j}_{n,m}(x, y) = \frac{1}{8} b_{n,m} \left(\frac{x - 8i}{8}, \frac{y - 8j}{8} \right) \chi_{i,j}(x, y) \tag{18}$$

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 systems 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.

$$(G_c v)_n = (g^c_n, v)_{L^2}, \tag{19}$$

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

$$U_D = \{u \in L^2(\Omega, \mathbb{R}^3) \,|\, (Au)_n^c \in J_n^c \text{ for } n \in \mathbb{N}, c \in \{1,2,3\}\} \tag{20}$$

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

$$a_n^c(x,y) = S_c^* g_n^c(x,y) = \frac{1}{s_c t_c} g_n^c\left(\frac{x}{s_c}, \frac{y}{t_c}\right), \tag{21}$$

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

$$\min_{u \in L^2(\Omega, \mathbb{R}^3)} \mathrm{TGV}_\alpha^k(u) + \mathcal{I}_{U_D}(u), \tag{22}$$

with U_D being equivalently defined in equation (20). In order to assure (A) and (EX_k) to 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

$$J_n^c = [Q_n^c(d_n^c - \frac{1}{2}), Q_n^c(d_n^c + \frac{1}{2})]. \tag{23}$$

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 Equation (23), in particular are bounded. Since the quantization values Q_n^c are typically non-decreasing 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 non-empty 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 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 Equation* (23) *and all remaining 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.*

4.3 Color JPEG 2000 decompression

As 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 to 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 improve 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 in particular 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 to 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 to solve

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

with

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

and M being the index set of known coefficients. In [19] existence of a solution for this problem was established in function space setting under the assumption 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 [19, 18, 49, 41]. In [51] the same model using non-local TV regularization was considered. In [26] 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 to minimize the TV functional subject to equality constraints on all wavelet coefficients with absolute value above a certain threshold.

However, even when considered solely as method for wavelet-constrained optimization, our work differs significantly from the ones cited above. First of all we use the total generalized variation functional of arbitrary order as regularization. Also, we are able to establish existence of a solution and optimality conditions in the case Ω is a *bounded Lipschitz domain* and using *natural boundary extension* also in 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 [12] is different to the ones of previous works. Let us point out however, that the assumptions of our work include the problem of wavelet inpainting, thus it can also be seen as a generalization of methods of [19, 26] using arbitrary order TGV regularization.

Other methods mainly focused on the concealment of wavelet data error due to transmission are the works [35, 5, 21]. In [48] the aim is the reduction of artifacts due to tile separation of the image. We also refer to [40] for a post processing 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 to [37, 45, 46, 32] 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 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 [45, 32]). 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 non-zero 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 PSNR (see [37, Section 10.5.2],[32, 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 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 arbitrary 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 [32, 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, non-empty 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, we can define a source value together with a (bounded) error interval for each coefficient of the code block, exactly how we could do for a JPEG compressed image. Thus, given any code block with non-zero 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 [12], 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 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 and a detail-part, subsampling is still possible by skipping the detail coefficients of the finest scale.

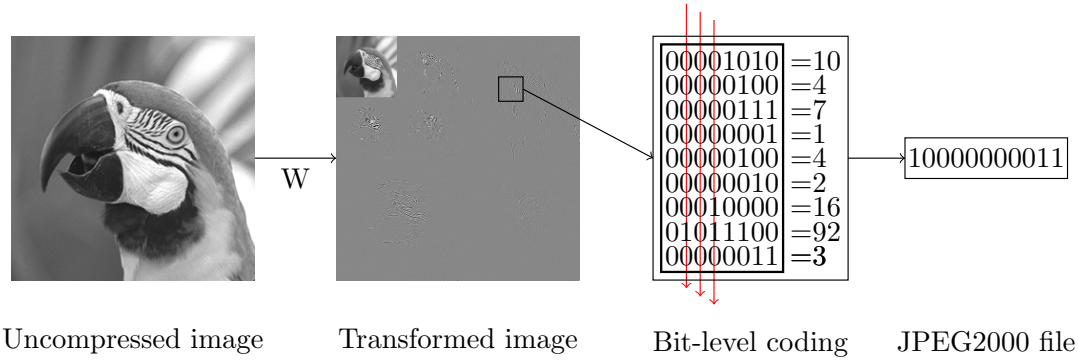

Figure 3: Selected steps of JPEG 2000 compression including bit-level coding for one image component. See also [37, Figures 10.16,10.17].

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 CDF 9/7 wavelets. Let us detail on 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 [22, Table 6.1 and Table 6.2] for $N = \tilde{N} = 2$ and $N = \tilde{N} = 4$, respectively. As shown in [22], 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 $R \in \mathbb{Z}$,

$$(\phi_{R,k})_k \cup (\psi_{j,k})_{j,k}, \qquad \text{for indices } k \in \mathbb{Z} \text{ and } j \leq R, \tag{24}$$

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

$$(\tilde{\phi}_{R,k})_k \cup (\tilde{\psi}_{j,k})_{j,k}, \qquad \text{for indices } k \in \mathbb{Z} \text{ and } j \leq R, \tag{25}$$

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 (24) and (25), as

$$\begin{aligned} f = & \sum_{k\in\mathbb{Z}} (\tilde{\phi}_{k,R}, f)\phi_{k,R} + \sum_{k\in\mathbb{Z}, j\leq R} (\tilde{\psi}_{j,k}, f)\psi_{j,k} \\ = & \sum_{k\in\mathbb{Z}} (\phi_{k,R}, f)\tilde{\phi}_{k,R} + \sum_{k\in\mathbb{Z}, j\leq R} (\psi_{j,k}, f)\tilde{\psi}_{j,k}. \end{aligned} \tag{26}$$

The first sum in each of the terms can be interpreted as a low resolution approximation of f while the second one 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

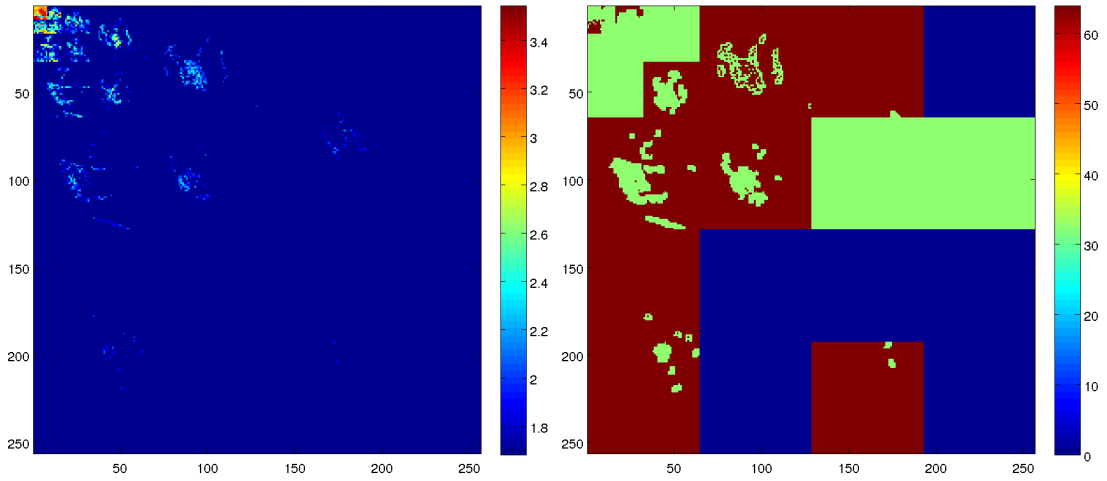

Figure 4: Left: Wavelet coefficients of the brightness component of the bird image of Figure 3 (logarithmic scale). Right: Size of data intervals for the wavelet coefficients (Note that 0, i.e., dark blue, indicates that the data is unbounded).

as in [23, 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

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

Thus, testing $u \in L^2((0,1))$ with η^{f} corresponds to testing its symmetric extension with η. Using this techniques and skipping redundant indices, it has been shown in [23, 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 [25, 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 gives a wavelet that resolves horizontal, vertical or diagonal details, respectively. Grouping the scaling functions and the wavelets together and re-indexing we obtain, again for $R \in \mathbb{Z}$, two dual Riesz bases of $L^2((0,1)\times(0,1))$ written as

$$(\Phi_{R,k})_k \cup (\Psi_{j,k})_{j,k} \quad \text{and} \quad (\tilde{\Phi}_{R,k})_k \cup (\tilde{\Psi}_{j,k})_{j,k}. \tag{27}$$

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

In order to apply the problem setting of Section 3, we will need to assure 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 [22]), 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 [22, Section 6.A] and note that there, ϕ and $\tilde{\phi}$ are interchanged), thus is contained in $W^{1,1}(\mathbb{R})$. In the case of CDF 9/7 filters it has been shown for example in [46] that the scaling function corresponding to synthesis possesses a Sobolev regularity higher than 2, 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 bases of $L^2((0,1)\times(0,1))$ as in (27). 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 bases of $L^2(\Omega,\mathbb{R}^3)$ together with a dual basis from these component-wise 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

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

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 non-zero bit is encoded, no intervals are half bounded and (EX_k) applies. During compression, the wavelet coefficients of the signal are quantized, rounded towards zero and a bit truncation is performed. The values used for quantization are uniform on each subband and are typically non-decreasing for higher subbands, meaning that coefficients corresponding to finer scales are saved 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 assumption (A) to be 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 a regularization functional. Following this approach, we model subsampling as a linear operator mapping a function to a finite subset of

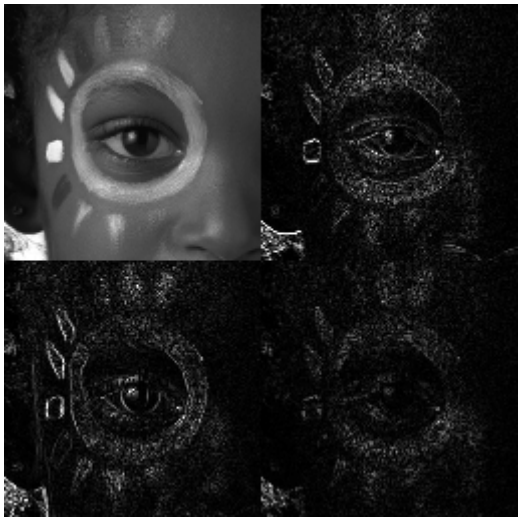

Figure 5: Visualization of of the low resolution image and the detail coefficients obtained from a high resolution image with one level of wavelet decomposition. Note that the wavelet coefficients have been rescaled for better visibility.

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 to set $J_n = \{(a_n, u^\dagger)_{L^2}\}$ 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 $(\Psi_{j,k})_{j,k}$. 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 [10] for the case of second order TGV 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 subsampling operator (see [10]). We also refer to [10] 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 block-wise cosine basis or a tile-wise 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 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 to [9] 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 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 is the topic of a second paper [12] whose contents are strongly connected to the present one.

A Products of tensor spaces and related mappings

The 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 total generalized variation functional for vector-valued data. The definitions and results stated in this subsection are straightforward generalizations of the ones presented in [13, 11] and are provided for the readers convenience.

The space of symmetric tensors of order k is defined as

$$\mathrm{Sym}^k(\mathbb{R}^d) := \left\{ \xi : \left(\mathbb{R}^d\right)^k \to \mathbb{R} \,\middle|\, \xi : \ k - \text{linear and symmetric} \right\}, \tag{28}$$

respectively, with the scalar product

$$\xi \cdot \eta = \sum_{p \in \{1,\dots d\}^k} \xi(e_{p_1}, \dots, e_{p_k}) \eta(e_{p_1}, \dots, e_{p_k}), \tag{29}$$

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 (non-symmetric) $(k+l)$ tensor field $\nabla^l \otimes \xi$,

defined by

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

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:

$$\mathcal{E}^l \xi = |||(\nabla^l \otimes \xi). \tag{30}$$

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 $\xi \in \mathrm{Sym}^k(\mathbb{R}^d)$,

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

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.,

$$\mathrm{Sym}^k(\mathbb{R}^d)^m = \Big\{ \xi = (\xi_1, \dots, \xi_m) \,\Big|\, \xi_i \in \mathrm{Sym}^k(\mathbb{R}^d),\, i \in \{1, \dots, m\} \Big\}, \tag{31}$$

with the inner product and norm

$$\xi \cdot \eta = \sum_{i=1}^{m} \xi_i \cdot \eta_i \qquad \text{and} \qquad |\xi|^2 = \xi \cdot \xi. \tag{32}$$

For sufficiently smooth m-tuples of symmetric tensor fields, the differentiation operators $\nabla, \mathcal{E}, \mathrm{div}$ are defined component-wise. 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 (32) 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 [11] 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 [11] to the distributions $u^i \in \mathcal{D}(\Omega, \mathrm{Sym}^k(\mathbb{R}^d))$ defined by

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

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. *Journal of Mathematical Imaging and Vision*, 23:199–211, 2005.

[2] H.A. Aly and E. Dubois. Image up-sampling using total-variation regularization with a new observation model. *IEEE Transactions on Image Processing*, 14(10):1647–1659, Oct 2005.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, 2000.

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

[5] L. Atzori. JPEG2000-coded image error concealment exploiting convex sets projections. *IEEE Transactions on Image Processing*, 14(4):487–498, 2005.

[6] A. Auslender. Noncoercive optimization problems. *Math. Oper. Res.*, 21(4):769–782, 1996.

[7] K. Bredies. Symmetric tensor fields of bounded deformation. *Annali di Matematica Pura ed Applicata*, 195(5):815–851, 2013.

[8] K. Bredies and M. Holler. A total variation–based JPEG decompression model. *SIAM Journal on Imaging Sciences*, 5(1):366–393, 2012.

[9] 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*, volume 359, pages 242–258. Springer, 2013.

[10] K. Bredies and M. Holler. A TGV regularized wavelet based zooming model. In *Scale Space and Variational Methods in Computer Vision*, volume 7893, pages 149–160. Springer, 2013.

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

[12] K. Bredies and M. Holler. A TGV-based framework for variational image decompression, zooming and reconstruction. part II: Numerics. *Submitted*, 2014.

[13] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[14] K. Bredies and D. Lorenz. *Mathematische Bildverarbeitung*. Vieweg+Teubner, 2011.

[15] Kristian 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*, volume 8293 of *Lecture Notes in Computer Science*, pages 44–77. Springer Berlin Heidelberg, 2014.

[16] W.K. Carey, D.B. Chuang, and S.S. Hemami. Regularity-preserving image interpolation. *Image Processing, IEEE Transactions on*, 8(9):1293–1297, Sep 1999.

[17] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20:88–97, 2004.

[18] R. H. Chan, J. Yang, and X. Yuan. Alternating direction method for image inpainting in wavelet domains. *SIAM Journal on Imaging Sciences*, 4(3):807–826, 2010.

[19] T. F. Chan, J. Shen, and H.-M. Zhou. Total variation wavelet inpainting. *Journal of Mathematical Imaging and Vision*, 25(1):107–125, 2006.

[20] Tao Chen, Hong Ren Wu, and Bin Qiu. Image interpolation using across-scale pixel correlation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, volume 3, pages 1857–1860 vol.3, 2001.

[21] P. C. Chung. A JPEG 2000 error resilience method using uneven block-sized information included markers. *IEEE Transactions on Circuits and Systems for Video Technology*, 15(3):420–424, 2005.

[22] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5):485–560, 1992.

[23] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, 1993.

[24] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4:1168–1200, 2005.

[25] I. Daubechies. *Ten Lectures on Wavelets*. Number 61 in CBMS-NSF Lecture Notes. SIAM, 1992.

[26] S. Durand and J. Froment. Reconstruction of wavelet coefficients using total variation minimization. *SIAM Journal on Scientific Computing*, 24(5):1754–1767, 2003.

[27] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. SIAM, 1999.

[28] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.

[29] V. Girault and P.-A. Raviart. *Finite Element Method for Navier-Stokes Equation*. Springer, 1986.

[30] B. Goldluecke, E. Strekalovskiy, and D. Cremers. The natural total variation which arises from geometric measure theory. *SIAM Journal on Imaging Sciences*, 5(2):537–563, 2012.

[31] Markus Grasmair, Markus Haltmeier, and Otmar Scherzer. The residual method for regularizing ill-posed problems. *Applied Mathematics and Computation*, 218(6):2693–2710, 2011.

[32] Joint Bilevel Image Experts Group and Joint Photographic Experts Group. *JPEG 2000 image coding system*, 2000. ISO/IEC 15444-1.

[33] Martin Holler. *Higher order regularization for model based data decompression.* PhD thesis, University of Graz, 2013.

[34] N. Kaulgud and U. B. Desai. Image zooming: Use of wavelets. In *The International Series in Engineering and Computer Science*, volume 632, pages 21–44. Springer, 2002.

[35] P. J. Lee. Error concealment algorithm using interested direction for JPEG 2000 image transmission. *IEEE Transactions on Consumer Electronics*, 49(4):1395–1401, 2003.

[36] F. Malgouyres and F. Guichard. Edge direction preserving image zooming: A mathematical and numerical analysis. *SIAM Journal on Numerical Analysis*, 39:1–37, 2001.

[37] S. Mallat. *A Wavelet Tour of Signal Processing.* Elsevier, 2009.

[38] S. Mallat and Guoshen Yu. Super-resolution with sparse mixing estimators. *IEEE Transactions on Image Processing*, 19(11):2889–2900, Nov 2010.

[39] A. Nosratinia. Enhancement of JPEG-compressed images by re-application of JPEG. *The Journal of VLSI Signal Processing*, 27:69–79, 2001.

[40] A. Nosratinia. Postprocessing of JPEG-2000 images to remove compression artifacts. *IEEE Signal Processing Letters*, 10(10):296–299, 2003.

[41] R. Oktem. Regularization-based error concealment in JPEG 2000 coding scheme. *IEEE Signal Processing Letters*, 14(12):956–959, 2001.

[42] S. Ono and I. Yamada. Optimized JPEG image decompression with super-resolution interpolation using multi-order total variation. In *20th IEEE International Conference on Image Processing (ICIP)*, pages 474–478, Sept 2013.

[43] M.-Y. Shen and C.-C. Jay Kuo. Review of postprocessing techniques for compression artifact removal. *Journal of Visual Communication and Image Representation*, 9(1):2–14, 1998.

[44] S. Singh, V. Kumar, and H. K. Verma. Reduction of blocking artifacts in JPEG compressed images. *Digital Signal Processing*, 17(1):225–243, 2007.

[45] A. Skodras, C. Christopoulos, and T. Ebrahimi. The JPEG 2000 still image compression standard. *IEEE Signal processing Magazine*, 18:36–58, 2001.

[46] M. Unser and T. Blu. Mathematical properties of the JPEG 2000 wavelet filters. *Image Processing, IEEE Transactions on*, 12:1080–1090, 2003.

[47] G. K. Wallace. The JPEG still picture compression standard. *Communications of the ACM*, 34(4):30–44, 1991.

[48] 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). *Image Processing, IEEE Transactions on*, 14(8):1033–1042, 2001.

[49] Y.-W. Wen, R. H. Chan, and A. M. Yip. A primal-dual method for total variaton-based wavelet domain inpainting. *Image Processing, IEEE Transactions on*, 21(1):106–114, 2011.

[50] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, 2001.

[51] X. Zhang and T. F. Chan. Wavelet inpainting by nonlocal total variation. *Inverse Problems and Imaging*, 4(1):191–210, 2010.

[52] S. Zhong. Image coding with optimal reconstruction. In *International Conference on Image Processing*, volume 1, pages 161–164, 1997.

[53] W. P. Ziemer. *Weakly Differentiable Functions.* Springer, 1989.