SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

A Total-Variation-based JPEG decompression model

M. Holler K. Bredies

SFB-Report No. 2011-010 May 2011

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

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

A Total-Variation-based JPEG decompression model*

K. Bredies and M. Holler†

Abstract. We propose a variational model for artifact-free JPEG decompression. It bases on the minimization of the total variation (TV) over the convex set U of all possible source images associated with given JPEG data. The general case where U represents a pointwise restriction with respect to a L^2-orthonormal basis is considered. Analysis of the infinite dimensional model is presented, including the derivation of optimality conditions. A discretized version is solved using a primal-dual algorithm supplemented by a primal-dual gap stopping criterion. Experiments illustrate the effect of the model. A good reconstruction quality is obtained even for highly compressed images, while a GPU implementation is shown to significantly reduce computation time, making the model suitable for real-time applications.

Key words. Total variation, artifact-free JPEG decompression, image reconstruction, optimality system, primal-dual gap.

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

1. Introduction.

This work is concerned with a constrained optimization problem motivated by the problem of artifact-free JPEG decompression. Being a lossy compression, a JPEG-compressed object yields a generally non-singleton convex set of possible source images containing the original one. A priori assuming an image to be approximately piecewise constant, our model chooses an image in accordance with the given data and minimal total variation, leading to a good approximation of the original image. The underlying infinite dimensional minimization problem reads

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

where U is the given image data set associated with the JPEG object.

The scope of the paper is the presentation and analysis of a general continuous model of this type and, in particular, its application to artifact-free JPEG decompression. This includes an existence result and the derivation of optimality conditions for a general setting, with the purpose of providing a basis for further investigations, for instance on qualitative properties of solutions and different solution strategies. To the best knowledge of the authors, an analysis for continuous JPEG decompression models does not appear in the literature. Moreover, in the concrete application to JPEG decompression, a new algorithm is proposed for solving the discrete optimization problem. This algorithm serves as a reliable, easy-to-implement solution strategy for numerical validation of the model. It further allows to estimate the difference of the current iterate to the optimal solution in terms of the TV functional and thus to ensure approximate optimality. Finally, an implementation for the GPU shows that

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

†Department of Mathematics, University of Graz, Heinrichstr. 36, A-8010 Graz, Austria (martin.holler@uni-graz.at, kristian.bredies@uni-graz.at).

this kind of mathematical framework, with the total variation functional as regularization, yields a practicable method to obtain a visually improved reconstruction from a given JPEG compressed image in real-time.

Since proposed in [23], the total variation enjoys a high popularity as a regularization functional for mathematical imaging problems as it is effective in preserving edges. Convex constraints appearing in these problems are often included directly, but also associated with a data fidelity term and realized by the introduction of a Lagrange multiplier leading to differentiable [16, 7, 8, 1] or non-differentiable [17, 12, 10] data fidelity terms. Our approach includes the convex constraint in the variational model and therefore leads to a non-smooth constrained minimization problem. The non-differentiability of both the TV-term as well as the constraint make the analysis of the problem, especially the derivation of optimality conditions, demanding. Nevertheless, the numerical algorithm we propose directly operates on the constraints via projections. It is therefore ensured that any approximate solution is contained in the given data set.

In the first part of this paper we study the infinite dimensional problem setting. There, the given image data set U is related to a general orthonormal basis of $L^2(\Omega)$ allowing to impose pointwise conditions on the coefficients for the basis representation of an optimal solution. The assumption of an arbitrary orthonormal basis together with weak assumptions on the coefficient intervals allow the potential application of the model to various image reconstruction problems such as DCT-based zooming. The results of the first part are as follows: Existence of a solution is proved and, ensuring additivity of the subdifferential for the general problem setting, optimality conditions are obtained.

In the second part, a discrete model formulation for artifact-free JPEG decompression is presented. We introduce and discuss a numerical algorithm which bases on the primal-dual algorithm introduced in [9]. We address additional topics such as the choice of a suitable stopping criterion and computation time of a GPU based implementation, with the aim of showing that the presented approach is already suitable for integration in user software.

Let us shortly discuss previous attempts to increase the reconstruction quality of JPEG images. Due to the high popularity of the JPEG standard, the development of such methods for JPEG encoded images is still an active research topic, even more than twenty years after the introduction of the standard. Consequently, there exist already many different numerical approaches with the aim to reduce typical artifacts and blocking effects. Most of them can be classified by – or are combinations of – one of the following basic principles: Approaches which are in no direct relation to the suggested model (and will thus not be discussed further) are post-processing methods based on filters or stochastic models. The filter based methods seem only effective if space varying filters combined with a pre-classification of image blocks are applied. Furthermore, there are segmentation based approaches, which are mostly used for document type images and thus neither are in the focus of our interest. A classical approach is to use algorithms based on projections onto convex sets (POCs), see for example [18, 28, 31]. Typically, one defines several convex sets according to data fidelity and regularization models, and then tries to find an image contained in the intersection of all those sets by iteratively projecting onto them. Difficulties of this method are the concrete implementation of these projections and its generally slow convergence.

The approach presented in this article can, among others, be classified as a constrained

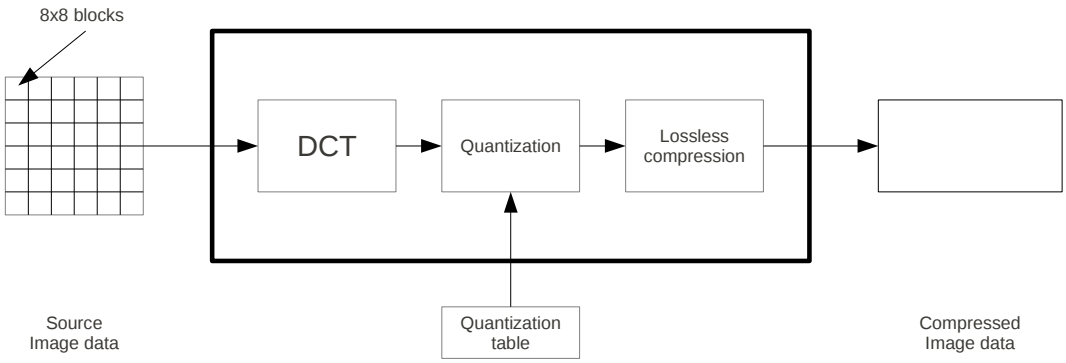

Figure 2.1. *Schematic overview of the JPEG compression procedure.*

optimization method, where we use the TV functional as regularization term. A first implementation of this idea has been given in [29] and a discrete model was derived in [2]. In both publications, a simple gradient descent algorithm together with a projection on the data set is used to solve the minimization problem. In contrast to that, we also focus on the continuous model which is, as the authors believe, fundamental for a well-developed discrete model. Usage of an efficient algorithm in the numerical part makes additional spatial weighting of the TV functional, as proposed in [2], unnecessary and the combination with a primal-dual gap based stopping criterion leads to an efficient, practicable solution strategy. There are of course other methods, unrelated to the previously presented principles. We refer to [21, 25, 24] for further reviews of current techniques.

The outline of the paper is as follows: In Section 2, using the JPEG compression as motivation, we present the general infinite dimensional problem formulation. In Section 3 preparatory results are established. Section 4 is the main section, where we define the infinite dimensional JPEG model as special case of the general problem setting, and then obtain existence and optimality results. In Section 5 we present and solve the discrete problem formulation and Section 6 concludes the paper by discussing the obtained results and giving a short outlook on further research.

2. Problem formulation.

The motivation for the infinite dimensional problem setting clearly comes from the application to JPEG decompression. Thus, to understand the abstract setting, we first give a short overview of the JPEG standard. For a more detailed explanation we refer to [27].

The JPEG compression is a lossy process, which means that most of the compression is obtained by loss of data, and the original image cannot be restored completely from the compressed object. Figure 2.1 illustrates the main steps of this process. At first, the image undergoes a blockwise discrete cosine transformation. As a result, the image is given as linear combination of different frequencies, making it easier to identify data with less importance to visual image quality such as high frequency variations. Then the image is quantized by pointwise division of each 8×8 block by a uniform quantization matrix. Next the quantized values are rounded to integer, which is where the loss of data takes place, and after that these

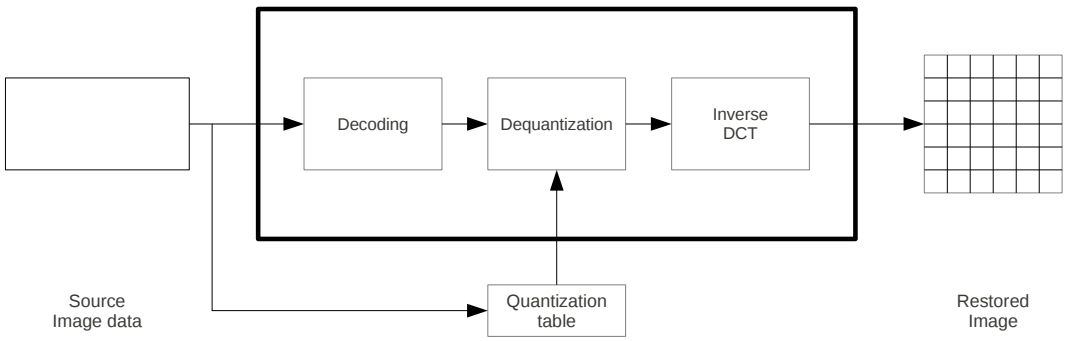

Figure 2.2. *Schematic overview of the standard JPEG decompression procedure.*

Figure 2.3. *JPEG image with typical blocking and ringing artifacts.*

integer values are further compressed by lossless compression. This data, together with the quantization matrix, is then stored in the JPEG object.

The standard JPEG decompression, as shown in Figure 2.2, simply reverses this process without taking into account that the given data is incomplete, i.e., that the compressed object delivers not a uniquely determined image, but a convex set of possible source images. Instead it just assumes the rounded integer value to be the true quantized DCT coefficient which leads to the well known JPEG artifacts as can be seen, for example, in Figure 2.3.

The principle idea of the model is now to use the given JPEG data to reconstruct one image within the set of all possible source images with minimal TV semi-norm. Hoping to reconstruct the original image this way, we implicitly assume that the given image is approximately piecewise constant. As we will see, this is a reasonable assumption and also suitable for a great range of realistic images.

First we will phrase the resulting minimization problem in an abstract setting. We denote by U a given set of possible source data which is the preimage of a set of coefficients I under a general linear basis transformation. We intend to find an element contained in U which has minimal TV semi-norm. The general problem formulation is to solve

$$\min_{u\in L^2(\Omega)} \mathrm{TV}(u) + \mathcal{I}_U(u), \tag{2.1}$$

where

$$\text{(A)}\quad \begin{cases} \Omega \subset \mathbb{R}^2 \text{ is a bounded Lipschitz domain,} \\ U = \{u \in L^2(\Omega) \,|\, Au \in I\} \text{ with} \\ A : L^2(\Omega) \to \ell^2 \text{ where } (Au)_n := (u, a_n)_{L^2}, \\ (a_n)_{n\in\mathbb{N}} \subset \mathrm{BV}(\Omega) \text{ is a complete orthonormal system of } L^2(\Omega) \\ I = \left\{z \in \ell^2 \,|\, z_n \in J_n \ \forall\, n \in \mathbb{N}\right\}, \\ (J_n)_{n\in\mathbb{N}} = ([l_n, r_n])_{n\in\mathbb{N}} \text{ is a sequence of non-empty, closed intervals,} \\ \overset{\circ}{U}_{\mathrm{int}} \neq \emptyset \text{ with } U_{\mathrm{int}} := \{u \in L^2(\Omega) \,|\, (Au)_n \in J_n \,\forall\, n \in \mathbb{N} \setminus W\}, \\ \text{where } W \subset \mathbb{N} \text{ is a finite index set,} \\ \text{and for at least one } n_0 \in \mathbb{N}, \text{ such that } \int_\Omega a_{n_0} \neq 0, J_{n_0} \text{ is bounded.} \end{cases}$$

For a definition of the functionals TV and $\mathcal{I}_U$ and further notation we refer to Section 3. Note that we allow free choice of the orthonormal system. Also we denote by ℓ^2 any space of square-integrable functions on a countable index set. Furthermore, only real spaces are considered. Existence of an element a_{n_0} such that $\int_\Omega a_{n_0} \neq 0$, as used in the last line of (A), is always satisfied for bounded Ω and will be addressed in Lemma 1. We will later see in Subsection 4.1 that these assumptions clearly include the infinite dimensional JPEG decompression model as special case.

After providing some analytical tools and results, we show existence of a solution to (2.1) under the assumption (A). Then we derive an optimality condition which allows to characterize all possible solutions.

3. Preliminaries.

This section is devoted to introduce some notation and present known results that are needed later on. Throughout this section, let $\Omega \subset \mathbb{R}^d$ always be a bounded Lipschitz domain. Further we often denote $\int_\Omega \phi$ or $\int_\Omega \phi \,\mathrm{d}x$ instead of $\int_\Omega \phi(x) \,\mathrm{d}x$ for the Lebesgue integral of an integrable function ϕ, when the usage of the Lebesgue measure and the integration variable are clear from the context.

Let us at first recall some basic definitions and results related to functions of bounded variation. For a detailed introduction including proofs we refer to [3, 30, 14].

Definition 1 (Finite Radon measure). *Let $\mathcal{B}(\Omega)$ be the Borel σ-algebra generated by the open subsets of Ω. We say that a function $\mu : \mathcal{B}(\Omega) \to \mathbb{R}^m$, for $m \in \mathbb{N}$, is a finite $\mathbb{R}^m$-valued Radon measure if $\mu(\emptyset) = 0$ and μ is σ-additive. We denote by $\mathcal{M}(\Omega)$ the space of all finite Radon*

measures on Ω. Further we denote by $|\mu|$ the variation of $\mu \in \mathcal{M}(\Omega)$, defined by

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

for $E \in \mathcal{B}(\Omega)$.

Definition 2 (Functions of bounded variation). *We say that a function $u \in L^1(\Omega)$ is of bounded variation, if there exists a finite $\mathbb{R}^d$-valued Radon measure, denoted by $Du = (D_1 u, ..., D_d u)$, such that for all $i \in \{1, ..., d\}$, $D_i u$ represents the distributional derivative of u with respect to the ith coordinate, i.e., we have*

$$\int_\Omega u \partial_i \phi = -\int_\Omega \phi \, \mathrm{d}D_i u \quad \text{for all } \phi \in C_c^\infty(\Omega)$$

By $\mathrm{BV}(\Omega)$ we denote the space of all functions $u \in L^1(\Omega)$ of bounded variation.

Definition 3 (Total variation). *For $u \in L^1(\Omega)$, we define the functional $\mathrm{TV} : L^1(\Omega) \to \overline{\mathbb{R}}$ as*

$$\mathrm{TV}(u) = \sup\left\{\int_\Omega u \operatorname{div} \phi \,\middle|\, \phi \in C_c^\infty(\Omega, \mathbb{R}^d),\, \|\phi\|_\infty \leq 1\right\}$$

where we set $\mathrm{TV}(u) = \infty$ if the set is unbounded from above. We call $\mathrm{TV}(u)$ the total variation of u.

Proposition 1. *The functional $\mathrm{TV} : L^1(\Omega) \to \overline{\mathbb{R}}$ is convex and lower semi-continuous with respect to L^1-convergence. For $u \in L^1(\Omega)$ we have that*

$$u \in \mathrm{BV}(\Omega) \text{ if and only if } \mathrm{TV}(u) < \infty.$$

In addition, the total variation of u coincides with the variation of the measure Du, i.e., $\mathrm{TV}(u) = |Du|(\Omega)$. Further,

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

defines a norm on $\mathrm{BV}(\Omega)$ and endowed with this norm, $\mathrm{BV}(\Omega)$ is a Banach space.

Definition 4 (Strict Convergence). *For $(u_n)_{n\in\mathbb{N}}$ with $u_n \in \mathrm{BV}(\Omega)$, $n \in \mathbb{N}$, and $u \in \mathrm{BV}(\Omega)$ we say that $(u_n)_{n\in\mathbb{N}}$ strictly converges to u if*

$$\|u_n - u\|_{L^1} \to 0 \text{ and } \mathrm{TV}(u_n) \to \mathrm{TV}(u)$$

as $n \to \infty$. Next we recall some standard notations and facts from convex analysis. For proofs and further introduction we refer to [13].

Definition 5 (Convex conjugate and subdifferential). *For a normed vector space V and a function $F : V \to \overline{\mathbb{R}}$ we define its convex conjugate, or Legendre-Fenchel transform, denoted by $F^* : V^* \to \overline{\mathbb{R}}$, as*

$$F^*(u^*) = \sup_{v\in V} \langle v, u^* \rangle_{V,V^*} - F(v).$$

Further F is said to be subdifferentiable at $u \in V$ if $F(u)$ is finite and there exists $u^* \in V^*$ such that

$$\langle v - u, u^* \rangle_{V,V^*} + F(u) \leq F(v)$$

for all $v \in V$. The element $u^* \in V^*$ is then called a subgradient of F at u and the set of all subgradients at u is denoted by $\partial F(u)$.

Definition 6 (Convex indicator functional). *For a normed vector space V and $U \subset V$ a convex set, we denote by $\mathcal{I}_U : V \to \overline{\mathbb{R}}$ the convex indicator function of U, defined by*

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

The space $H(\mathrm{div};\Omega)$, as defined on the following, will be needed to describe the subdifferential of the TV functional and to present an optimality condition for the infinite dimensional minimization problem. For further details we refer to [15, Chapter 1], where classical results like density of $C^\infty(\overline{\Omega},\mathbb{R}^d)$ and existence of a normal trace on $\partial\Omega$ for $H(\mathrm{div};\Omega)$ functions are derived.

Definition 7 (The space $H(\mathrm{div};\Omega)$). *Let $g \in L^2(\Omega,\mathbb{R}^d)$. We say that $\operatorname{div} g \in L^2(\Omega)$ if there exists $w \in L^2(\Omega)$ such that for all $\phi \in C_c^\infty(\Omega)$*

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

Furthermore we define

$$H(\mathrm{div};\Omega) = \left\{ g \in L^2(\Omega,\mathbb{R}^d) \mid \operatorname{div} g \in L^2(\Omega) \right\}$$

with the norm $\|g\|^2_{H(\mathrm{div})} := \|g\|^2_{L^2} + \|\operatorname{div} g\|^2_{L^2}$.

Remark 1. *Density of $C_c^\infty(\Omega)$ in $L^2(\Omega)$ implies that, if there exists $w \in L^2(\Omega)$ as above, it is unique. Hence it makes sense to write $\operatorname{div} g = w$. By completeness of $L^2(\Omega)$ and $L^2(\Omega,\mathbb{R}^d)$ it follows that $H(\mathrm{div};\Omega)$ is a Banach space when equipped with $\|\cdot\|_{H(\mathrm{div})}$.*

Definition 8. *We define*

$$H_0(\mathrm{div};\Omega) = \overline{C_c^\infty(\Omega,\mathbb{R}^d)}^{\|\cdot\|_{H(\mathrm{div})}}.$$

We further need a notion of normal trace for $H(\mathrm{div};\Omega)$ functions in the space $L^\infty(\Omega;|\mathrm{D}u|)$, $u \in \mathrm{BV}(\Omega)$, as it has been introduced in [4]. For the sake of self-containedness we provide here a basic definition and results necessary for our work. For more details and proofs however, we refer to [4]. Note that on the following we restrict ourselves to the special case $d = 2$. The reason for this restriction is, on the one hand, to maintain a simple notation as we have the continuous embedding of $\mathrm{BV}(\Omega)$ in $L^2(\Omega)$ [3, Corollary 3.49]. On the other hand, since assumption (A) contains $d = 2$, the results remain applicable.

Proposition 2. *Let $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitz domain, $u \in \mathrm{BV}(\Omega)$ and $g \in H(\mathrm{div};\Omega) \cap L^\infty(\Omega,\mathbb{R}^2)$. Then the linear functional $(g,\mathrm{D}u) : C_c^\infty(\Omega) \to \mathbb{R}$, defined as*

$$(g,\mathrm{D}u)(\phi) = \int_\Omega u \operatorname{div}(g\phi)\,\mathrm{d}x,$$

can be extended uniquely to a functional in $(C_0(\Omega))^*$ *and can be identified with a Radon measure in* Ω*. Further, this measure is absolutely continuous with respect to the measure* $|\mathrm{D}u|$ *and there exists a* $|\mathrm{D}u|$*-measurable function*

$$\theta(g, \mathrm{D}u) : \Omega \to \mathbb{R}$$

such that

$$\int_B (g, \mathrm{D}u) = \int_B \theta(g, \mathrm{D}u)|\mathrm{D}u| \quad \textit{for all Borel sets } B \subset \Omega$$

and

$$\|\theta(g, \mathrm{D}u)\|_{L^\infty(\Omega, |\mathrm{D}u|)} \leq \|g\|_\infty.$$

Definition 9 (Normal trace). *For* $\Omega \subset \mathbb{R}^2$ *a bounded Lipschitz domain,* $u \in \mathrm{BV}(\Omega)$ *and* $g \in H(\mathrm{div}; \Omega) \cap L^\infty(\Omega, \mathbb{R}^2)$ *we say that the function* $\theta(g, \mathrm{D}u) \in L^\infty(\Omega; |\mathrm{D}u|)$ *as in Proposition 2 is the normal trace of* g *in* $L^\infty(\Omega; |\mathrm{D}u|)$*.* The following proposition motivates the term normal trace and provides a generalized Gauss-Green theorem. For the latter, we need to recall (see [4, Proposition 1.3]) that for any $g \in H(\mathrm{div}; \Omega) \cap L^\infty(\Omega, \mathbb{R}^2)$, with $\Omega \subset \mathbb{R}^2$ being a bounded Lipschitz domain, we can define a boundary-normal trace on $\partial\Omega$, denoted by $(g, \nu) \in L^\infty(\partial\Omega; \mathcal{H}^1)$, satisfying

$$\int_\Omega \phi \,\mathrm{div}\, g \,\mathrm{d}x + \int_\Omega g \cdot \nabla\phi \,\mathrm{d}x = \int_{\partial\Omega} (g, \nu)\phi \,\mathrm{d}\mathcal{H}^1$$

for any $\phi \in C^1(\overline{\Omega})$.

Proposition 3. *With the assumptions of Proposition 2,* σ_u *denoting the density function of* $\mathrm{D}u$ *with respect to* $|\mathrm{D}u|$ *and* $|\mathrm{D}^a u|$ *the absolute continuous part of* $|\mathrm{D}u|$ *with respect to* $\mathcal{L}^2$*, we have*

$$\theta(g, \mathrm{D}u) = g \cdot \sigma_u$$

$|\mathrm{D}^a u|$*-almost everywhere in general and* $|\mathrm{D}u|$*-almost everywhere whenever* $g \in C_0(\Omega, \mathbb{R}^2)$ *or* $\mathrm{D}u \in L^1(\Omega, \mathbb{R}^2)$*. Further we have that*

$$\int_\Omega u \,\mathrm{div}\, g \,\mathrm{d}x + \int_\Omega \theta(g, \mathrm{D}u) \,\mathrm{d}|\mathrm{D}u| = \int_{\partial\Omega} (g, \nu)u \,\mathrm{d}\mathcal{H}^1.$$

Note that in the case $g \in H_0(\mathrm{div}; \Omega) \cap L^\infty(\Omega, \mathbb{R}^2)$ it follows that $(g, \nu) = 0$ on $\partial\Omega$ and thus

$$\int_\Omega u \,\mathrm{div}\, g \,\mathrm{d}x = -\int_\Omega \theta(g, \mathrm{D}u) \,\mathrm{d}|\mathrm{D}u|.$$

This notion of normal trace now allows us to provide a known characterization of the subdifferential of the TV functional. Since this is not a standard result, we present a proof for the sake of completeness.

Theorem 1 (Normal trace characterization). *For $\Omega \subset \mathbb{R}^2$ a bounded Lipschitz domain and $u \in L^2(\Omega)$, we have that $u^* \in \partial\,\mathrm{TV}(u)$ if and only if*

$$\begin{cases} u \in \mathrm{BV}(\Omega) \text{ and there exists } g \in H_0(\mathrm{div};\Omega) \\ \text{with } \|g\|_\infty \leq 1 \text{ such that } u^* = -\operatorname{div} g \text{ and} \\ \quad \theta(g, \mathrm{D}u) = 1,\ |\mathrm{D}u|\text{-almost everywhere.} \end{cases}$$

Proof. At first note that, for $g \in H_0(\mathrm{div};\Omega)$, $\|g\|_\infty \leq 1$, the equation $\theta(g, \mathrm{D}u) = 1\,|\mathrm{D}u|$-almost everywhere, is equivalent to $\int_\Omega \mathbf{1}\,\mathrm{d}|\mathrm{D}u| = -\int_\Omega u \operatorname{div} g$. Indeed, this is true since

$$\|\theta(g, \mathrm{D}u)\|_{L^\infty(\Omega;|\mathrm{D}u|)} \leq \|g\|_\infty \leq 1$$

and

$$-\int_\Omega u \operatorname{div} g = \int_\Omega \theta(g, \mathrm{D}u)\,\mathrm{d}|\mathrm{D}u|,$$

where we used that $(g, \nu) = 0$ on $\partial\Omega$ together with the generalized Gauss-Green theorem of Proposition 3. Thus, in order to show the desired characterization of the subdifferential, it suffices to show the assertion with $\theta(g, \mathrm{D}u) = 1$ replaced by $\int_\Omega \mathbf{1}\,\mathrm{d}|\mathrm{D}u| = -\int_\Omega u \operatorname{div} g$.

Denoting by $C = \left\{\operatorname{div}\phi \,|\, \phi \in C_c^\infty(\Omega, \mathbb{R}^2),\ \|\phi\|_\infty \leq 1\right\}$, we have

$$\mathrm{TV}(u) = \mathcal{I}_C^*(u),$$

where $\mathcal{I}_C^*$ denotes the convex conjugate of $\mathcal{I}_C$ (see Definition 5), and, consequently [13, Example I.4.3],

$$\mathrm{TV}^*(u^*) = \mathcal{I}_C^{**}(u^*) = \mathcal{I}_{\overline{C}}(u^*)$$

where the closure of C is taken with respect to the L^2 norm. Using the equivalence [13, Proposition I.5.1]

$$u^* \in \partial\,\mathrm{TV}(u) \Leftrightarrow \mathrm{TV}(u) + \mathrm{TV}^*(u^*) = (u, u^*)_{L^2},$$

we see that it suffices to show that

$$\overline{C} = \{\operatorname{div} g \,|\, g \in H_0(\mathrm{div}, \Omega),\ \|g\|_\infty \leq 1\} =: K$$

to obtain the desired assertion. Since clearly $C \subset K$, showing the closedness of K implies $\overline{C} \subset K$. For this purpose, take $(g_n)_{n\geq 0} \subset H_0(\mathrm{div};\Omega)$ with $\|g_n\|_\infty \leq 1$ such that

$$\operatorname{div} g_n \to h \text{ in } L^2(\Omega) \text{ as } n \to \infty.$$

By boundedness of $(g_n)_{n\geq 0}$ there exists a subsequence $(g_{n_i})_{i\geq 0}$ weakly converging to some $g \in L^2(\Omega, \mathbb{R}^d)$. Now for any $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega g \cdot \nabla\phi = \lim_{i\to\infty} \int_\Omega g_{n_i} \cdot \nabla\phi = \lim_{i\to\infty} -\int_\Omega \operatorname{div}(g_{n_i})\phi = -\int_\Omega h\phi,$$

from which follows that $g \in H(\mathrm{div};\Omega)$ and $\operatorname{div} g = h$. To show that $\|g\|_\infty \leq 1$ and $g \in H_0(\mathrm{div};\Omega)$ note that the set

$$\{(f, \operatorname{div} f) \,|\, f \in H_0(\mathrm{div};\Omega),\ \|f\|_\infty \leq 1\} \subset L^2(\Omega, \mathbb{R}^3)$$

forms a convex and closed – and therefore weakly closed – subset of $L^2(\Omega, \mathbb{R}^3)$ [13, Section I.1.2]. Since the sequence $((g_{n_i}, \operatorname{div} g_{n_i}))_{i\geq 0}$ is contained in this set and converges weakly in $L^2(\Omega, \mathbb{R}^3)$ to $(g, \operatorname{div} g)$, we have $g \in H_0(\operatorname{div}; \Omega)$ and $\|g\|_\infty \leq 1$. It remains to show that $K \subset \overline{C}$, for which it suffices to show that, for any $g \in H_0(\operatorname{div}; \Omega)$ with $\|g\|_\infty \leq 1$ fixed, we have for all $v \in L^2(\Omega)$ that

$$\int_\Omega v \operatorname{div} g \leq \mathrm{TV}(v)$$

since this implies $\mathrm{TV}^*(\operatorname{div} g) = \mathcal{I}_{\overline{C}}(\operatorname{div} g) = 0$. Now for such a $v \in L^2(\Omega)$ we can assume that $v \in \mathrm{BV}(\Omega)$ since otherwise the inequality is trivially satisfied. Thus we can take a sequence $(v_n)_{n\geq 0} \subset C^\infty(\overline{\Omega})$ strictly converging to v [3, Theorem 3.9], for which we can also assume that $v_n \to v$ with respect to $\|\cdot\|_{L^2}$. It eventually follows

$$\begin{aligned}\int_\Omega v \operatorname{div} g &= \lim_{n\to\infty} \int_\Omega v_n \operatorname{div} g = \lim_{n\to\infty} -\int_\Omega \nabla v_n \cdot g \\ &\leq \lim_{n\to\infty} \int_\Omega |\nabla v_n||g| \leq \lim_{n\to\infty} \mathrm{TV}(v_n) = \mathrm{TV}(v)\end{aligned}$$

which concludes the proof. ■

We now draw our attention to results related to the data fidelity term of our model. Important for the analytical development as well as for the practical implementation will be projections onto convex sets:

Proposition 4. *Let H be a Hilbert space, $U \subset H$ a nonempty, convex and closed subset and $u_0 \in H$. Then there exists exactly one $u \in U$ such that*

$$\|u - u_0\| = \inf_{v\in U} \|v - u_0\|$$

and this is equivalent to

$$(u_0 - u, v - u)_H \leq 0$$

for all $v \in U$.

Definition 10 (Projection). *In the situation of Proposition 4, we denote by $P_U : H \to U$ the function mapping each $u_0 \in H$ to its projection onto U, characterized by*

$$P_U(u) \in U : \quad (u_0 - P_U(u), v - u)_H \leq 0$$

for all $v \in U$.

We also need some general properties of orthonormal systems in Hilbert spaces (see for instance [26]):

Proposition 5 (Transform Operator). *Let H be a Hilbert space and $(a_n)_{n\in\mathbb{N}} \subset H$ a complete orthonormal system. Then the mapping*

$$\begin{aligned} A : H &\to \ell^2 \\ u &\mapsto ((u, a_n)_H)_{n\in\mathbb{N}}\end{aligned}$$

is an isometric isomorphism, i.e., A is linear, bijective and $\|Au\|_{l^2} = \|u\|_H$ for all $u \in H$. Further its adjoint $A^* : \ell^2 \to H$ is given by

$$A^* v = A^{-1} v = \sum_{n \in \mathbb{N}} a_n v_n$$

for $v = (v_n)_{n\in\mathbb{N}} \in \ell^2$.

The following lemma for complete orthonormal systems in $L^2(\Omega)$ will be needed to show existence of a solution to the minimization problem:

Lemma 1. *Let $(a_n)_{n\in\mathbb{N}}$ be a complete orthonormal system of $L^2(\Omega)$. If the constant functions are contained in $L^2(\Omega)$, i.e., if $|\Omega| < \infty$, then there exists a $n_0 \in \mathbb{N}$ such that*

$$\int_\Omega a_{n_0} \neq 0.$$

Proof. By boundedness of Ω, the constant function equal to one is contained in $L^2(\Omega)$ and thus by completeness of the orthonormal system, the assertion follows. ■

4. Analysis of the continuous problem. This section is devoted to the infinite dimensional JPEG model and the derivation of existence and optimality results for the general setting of assumption (A) which is shown to include the JPEG model. We therefore always assume (A) to be satisfied.

4.1. The JPEG model. In this subsection, we define the infinite dimensional JPEG decompression model. For this model, we assume the domain Ω to be a rectangle with width and height being multiples of 8, i.e. $\Omega = (0, 8k) \times (0, 8l)$ for some $k, l \in \mathbb{N}$, which reflects the fact that any grayscale JPEG image is processed with its horizontal and vertical number of pixels being multiples of 8.

At first, we give a definition of the orthonormal system related to the blockwise cosine transformation operator, which is simply a combination of standard cosine orthonormal systems on each 8×8 block of the domain. Figure 4.1 illustrates its finite dimensional equivalent used for JPEG decompression.

Definition 11 (Block-cosine system). *For $k, l \in \mathbb{N}$, let $\Omega = (0, 8k) \times (0, 8l) \subset \mathbb{R}^2$. For $i, j \in \mathbb{N}_0$, $0 \leq i < k$, $0 \leq j < l$ we define the squares*

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

and

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

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

$$b_{n,m}(x, y) = c_n c_m \cos(nx\pi) \cos(my\pi), \tag{4.1}$$

for $(x, y) \in \mathbb{R}^2$, where

$$c_s = \begin{cases} 1 & \text{if } s = 0 \\ \sqrt{2} & \text{if } s \neq 0. \end{cases}$$

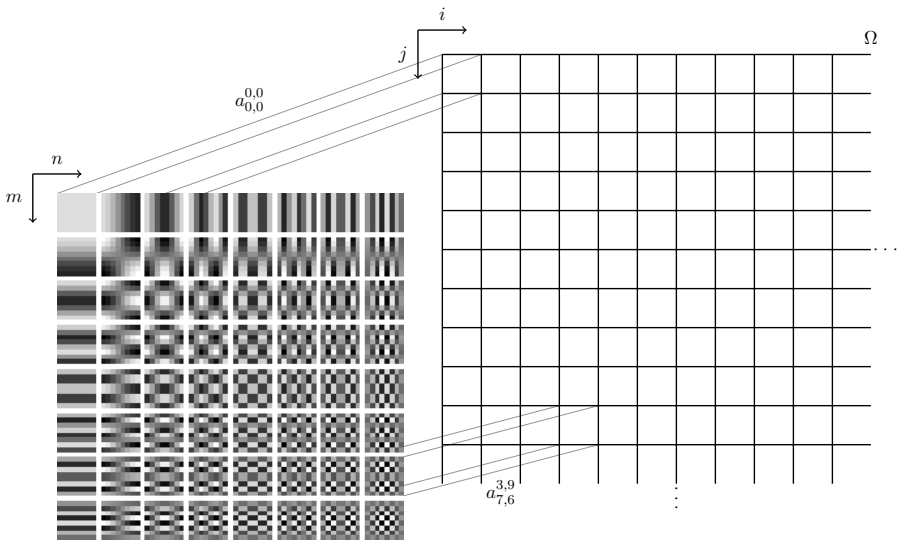

Figure 4.1. *Illustration of the finite dimensional block-cosine orthonormal basis used for JPEG decompression. As indicated, the basis function $a^{0,0}_{0,0}$ corresponds to the $(0,0)$ (upper left) 8×8-pixel block of the image and the $(0,0)$ frequency, while the basis function $a^{3,9}_{7,6}$ corresponds to the $(3,9)$ pixel block of the image and the $(7,6)$ frequency.*

With that, we define $a^{i,j}_{n,m} \in L^2(\Omega)$ as

$$a^{i,j}_{n,m}(x,y) = \frac{1}{8} b_{n,m}\left(\frac{x-8i}{8}, \frac{y-8j}{8}\right) \chi_{i,j}(x,y) \tag{4.2}$$

for $(x,y) \in \Omega$.

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

With given quantized integer coefficients $(z^{i,j}_{n,m})_{n,m\geq 0}$, for $0 \leq i < k$, $0 \leq j < l$, and $(q_{n,m})_{n,m\geq 0}$ given quantization values, which are both obtained from a compressed JPEG object, the data intervals describing the set of possible source images can be defined as

$$J^{i,j}_{n,m} = \left[q_{n,m} z^{i,j}_{n,m} - \frac{q_{n,m}}{2}, q_{n,m} z^{i,j}_{n,m} + \frac{q_{n,m}}{2}\right]. \tag{4.3}$$

The natural restriction $1 \leq q_{n,m} < \infty$ implies that $\overset{\circ}{J}{}^{i,j}_{n,m} \neq \emptyset$ and that all intervals $J^{i,j}_{n,m}$ are bounded. Consequently, the JPEG related minimization problem, with the orthonormal system as in (4.2) and the data intervals as in (4.3), is a special case of the general assumption (A) and results that are obtained for the general setting, as existence of a solution and optimality conditions, apply.

4.2. Existence. It is important for the following existence result that there is a finite index set W such that, with $U_{\mathrm{int}} := \{u \in L^2(\Omega) \,|\, (Au)_n \in J_n \,\forall\, n \in \mathbb{N} \setminus W\}$, we have $\overset{\circ}{U}_{\mathrm{int}} \neq \emptyset$.

Theorem 2 (Existence). *The minimization problem*

$$\min_{u\in L^2(\Omega)} \mathrm{TV}(u) + \mathcal{I}_U(u) \tag{4.4}$$

has a solution $\hat{u} \in \mathrm{BV}(\Omega)$.

Proof. First we show $F(u) := \mathrm{TV}(u)+\mathcal{I}_U(u)$ is proper, i.e., F takes nowhere the value $-\infty$ and is finite in at least one point. It is clear that $F(u) \geq 0$ for all $u \in L^2(\Omega)$, so it remains to find an element $u \in L^2(\Omega)$ such that $F(u) < \infty$. But this can be constructed as follows: Using that $\overset{\circ}{U}_{\mathrm{int}} \neq \emptyset$, by density of $C_c^\infty(\Omega) \subset \mathrm{BV}(\Omega)$ in $L^2(\Omega)$ there exists $u_0 \in \mathrm{BV}(\Omega)\cap U_{\mathrm{int}}$. Now we want to modify this u_0 such that we obtain $u \in \mathrm{BV}(\Omega)$, for which additionally $(u, a_n)_{L^2} \in J_n$ for $n \in W$.

But since $(a_n)_{n\geq 0} \subset \mathrm{BV}(\Omega)$ and W is finite, there exist constants $s_n \in \mathbb{R}$, $n \in W$, such that, with

$$u(x, y) := u_0(x, y) + \sum_{n\in W} s_n a_n(x, y),$$

for $(x, y) \in \Omega$, we have $u \in \mathrm{BV}(\Omega) \cap U$.

In order to find a minimizing element $\hat{u} \in L^2(\Omega)$, we take a minimizing sequence $(u_n)_{n\geq 0}$, i.e., $\lim_{n\to\infty} F(u_n) = \inf_{u\in X} F(u)$, for which, without loss of generality, we can assume that $(u_n)_{n\geq 0} \subset \mathrm{BV}(\Omega) \cap U$. To show boundedness of $(u_n)_{n\geq 0}$ with respect to $\|\cdot\|_{L^2}$, we define $\tilde{u}_n$ to be the constant function equal to $\frac{1}{|\Omega|}\int_\Omega u_n$. From the Poincaré inequality for BV [3, Remark 3.50], recall that Ω is a bounded Lipschitz domain, we get that

$$\|u_n - \tilde{u}_n\|_{L^2} \leq C\,\mathrm{TV}(u_n) \tag{4.5}$$

with $C \in \mathbb{R}$. Using that

$$\|\tilde{u}_n\|_{L^2} = |\Omega|^{\frac{1}{2}}|\tilde{u}_n| = |\Omega|^{\frac{1}{2}}|(A\tilde{u}_n)_{n_0}| \left|\int_\Omega a_{n_0}\right|^{-1}, \tag{4.6}$$

where n_0 is chosen as in assumption (A), it follows from the boundedness of J_{n_0} that $\|\tilde{u}_n\|_{L^2}$ and, consequently, $\|u_n\|_{L^2}$ is bounded. Hence, there exists a $\hat{u} \in L^2(\Omega)$ and a subsequence of $(u_n)_{n\geq 0}$, denoted by $(u_{n_k})_{k\geq 0}$, weakly converging to $\hat{u}$ in $L^2(\Omega)$. At last, convexity and closedness of U in $L^2(\Omega)$ imply weak closedness and thus, from lower semi-continuity of TV with respect to weak convergence in L^2, it follows that

$$F(\hat{u}) = \mathrm{TV}(\hat{u}) \leq \liminf_{k\to\infty} \mathrm{TV}(u_{n_k}) = \liminf_{k\to\infty} F(u_{n_k}) = \inf_{u\in L^2(\Omega)} F(u)$$

which implies that $\hat{u}$ is a solution to (4.4). ∎

4.3. Optimality conditions. The formulation of optimality conditions for solutions to the minimization problem we are giving in the following bases on a characterization of the subdifferential of the total variation functional as well as the indicator functional and the additivity of the subdifferential.

The characterization of the subdifferential of the total variation functional has been given in Section 3 and thus we now focus on the convex indicator functional. At first, recall that for a linear operator which is an isometric isomorphism, the following chain rule for the subdifferential holds:

Lemma 2. *Let X, Y be Hilbert spaces, $f : Y \to \overline{\mathbb{R}}$ proper, convex and lower semi-continuous, $T \in \mathcal{L}(X, Y)$ be an isometric isomorphism. Then we have*

$$\partial(f \circ T)(u) = T^* \partial f(Tu) \quad \text{for all } u \in X.$$

This enables us to provide a characterization of the subdifferential of the convex indicator functional $\mathcal{I}_U$. Recall for the following that according to assumption (A), $J_n = [l_n, r_n]$ for $n \in \mathbb{N}$.

Theorem 3. *We have that*

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

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

$$\begin{aligned} \lambda_n &\geq 0 \text{ if } (Au)_n = r_n \neq l_n, \\ \lambda_n &\leq 0 \text{ if } (Au)_n = l_n \neq r_n, \\ \lambda_n &= 0 \text{ if } (Au)_n \in \overset{\circ}{J}_n, \\ \lambda_n &\in \mathbb{R} \text{ if } (Au)_n = l_n = r_n. \end{aligned}$$

Proof.

With I corresponding to the given data set as in (A), using that A is an isometric isomorphism, Lemma 2 implies that

$$\partial \mathcal{I}_U(u) = \partial(\mathcal{I}_I \circ A)(u) = A^* \partial \mathcal{I}_I(Au)$$

and as a standard result in convex analysis we have

$$\lambda \in \partial \mathcal{I}_I(Au) \quad \Leftrightarrow \quad Au = P_I(Au + \lambda).$$

Finally, we can reduce P_I to a component-wise projection:

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

from which the assertion follows by a case study. ■

In order to show additivity of the subdifferential operator, we decompose $\mathcal{I}_U = \mathcal{I}_{U_{\text{int}}} + \mathcal{I}_{U_{\text{point}}}$ where, as before,

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

and

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

for a finite index set $W \subset \mathbb{N}$ such that $\overset{\circ}{U}_{\text{int}} \neq \emptyset$ and $J_n = \{j_n\}$ for all $n \in W$.

Theorem 4. *For all* $u \in L^2(\Omega)$,

$$\partial\left(\mathrm{TV} + \mathcal{I}_U\right)(u) = \partial\,\mathrm{TV}(u) + \partial\mathcal{I}_U(u).$$

Proof. Let $u \in L^2(\Omega)$. It is sufficient to show $\partial\left(\mathrm{TV} + \mathcal{I}_U\right)(u) \subset \partial\,\mathrm{TV}(u) + \partial\mathcal{I}_U(u)$, since the converse inclusion is always satisfied. Continuity of $\mathcal{I}_{U_{\mathrm{int}}}$ in at least one point $u \in \mathrm{BV}(\Omega) \cap U$ allows to apply [13, Proposition I.5.6] and assure that

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

In [5, Corollary 2.1], it was obtained, as a consequence of a more general result, that for two convex, lower semi-continuous and proper functions φ and ψ one has

$$\partial\left(\varphi + \psi\right) = \partial\varphi + \partial\psi$$

provided that

$$\bigcup_{\lambda \geq 0} \lambda\left(\mathrm{dom}(\varphi) - \mathrm{dom}(\psi)\right)$$

is a closed vector space. Thus, in order to establish

$$\partial(\mathrm{TV} + \mathcal{I}_{U_{\mathrm{point}}})(u) \subset \partial\,\mathrm{TV}(u) + \partial\mathcal{I}_{U_{\mathrm{point}}}(u)$$

it is sufficient to show that $\mathrm{dom}(\mathrm{TV}) - \mathrm{dom}(\mathcal{I}_{U_{\mathrm{point}}}) = L^2(\Omega)$. But this is true since, for any $w \in L^2(\Omega)$, denoting $J_n = \{j_n\}$ for $n \in W$, we can write

$$w = w_1 - w_2$$

where, using that $a_n \in \mathrm{BV}(\Omega)$ for $n \in \mathbb{N}$ and that W is finite,

$$w_1 = \sum_{n \in W} \left((a_n, w)_{L^2} + j_n\right) a_n \in \mathrm{dom}(\mathrm{TV})$$

and

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

Since

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

the assertion is proved. ■

Collecting the results, we can now write down the optimality conditions:

Theorem 5. *There exists a solution to*

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

and the following are equivalent:

1. $\hat{u} \in \arg\min_{u \in L^2(\Omega)} \left(\mathrm{TV}(u) + \mathcal{I}_U(u)\right) = \arg\min_{u \in U} \mathrm{TV}(u)$,

2. $\hat{u} \in \mathrm{BV}(\Omega) \cap U$ *and there exist* $g \in H_0(\mathrm{div}; \Omega)$ *and* $\lambda = (\lambda_n)_{n\in\mathbb{N}} \in \ell^2$ *satisfying*
   (a) $\|g\|_\infty \leq 1$,
   (b) $\theta(g, \mathrm{D}u) = 1$, $|\mathrm{D}u|$*-almost everywhere,*
   (c) $\mathrm{div}\, g = \sum_{n\in\mathbb{N}} \lambda_n a_n$,
   (d)
   $$\begin{cases} \lambda_n \geq 0 \text{ if } (A\hat{u})_n = r_n \neq l_n, \\ \lambda_n \leq 0 \text{ if } (A\hat{u})_n = l_n \neq r_n, \\ \lambda_n = 0 \text{ if } (A\hat{u})_n \in \overset{\circ}{J}_n, \end{cases} \qquad \forall n \in \mathbb{N},$$
   (note that, if $J_n = \{j_n\}$*, there is no additional condition on* λ_n*),*
3. $\hat{u} \in \mathrm{BV}(\Omega) \cap U$ *and there exists* $g \in H_0(\mathrm{div}; \Omega)$ *satisfying*
   (a) $\|g\|_\infty \leq 1$,
   (b) $\theta(g, \mathrm{D}u) = 1$, $|\mathrm{D}u|$*-almost everywhere,*
   (c)
   $$\begin{cases} (\mathrm{div}\, g, a_n)_{L^2} \geq 0 \ \ if (A\hat{u})_n = r_n \neq l_n, \\ (\mathrm{div}\, g, a_n)_{L^2} \leq 0 \ \ if (A\hat{u})_n = l_n \neq r_n, \\ (\mathrm{div}\, g, a_n)_{L^2} = 0 \ \ if (A\hat{u})_n \in \overset{\circ}{J}_n, \end{cases} \qquad \forall n \in \mathbb{N}.$$

Proof. Existence of a solution follows from Theorem 2. Equivalence of (2) and (3) is clear, so it is left to show the equivalence of (1) and (2):

(1) $\Rightarrow$ (2) : Let

$$\hat{u} \in \operatorname*{arg\,min}_{u\in L^2(\Omega)} \left(\mathrm{TV}(u) + \mathcal{I}_U(u)\right).$$

Thus $0 \in \partial\left(\mathrm{TV} + \mathcal{I}_U\right)(\hat{u})$ and by the additivity of the subdifferential, as shown in Theorem 4, we have $0 \in \partial\,\mathrm{TV}(\hat{u}) + \partial\mathcal{I}_U(\hat{u})$. Hence, there exist elements $z_1 \in \partial\,\mathrm{TV}(\hat{u})$ and $z_2 \in \partial\mathcal{I}_U(\hat{u})$ such that $0 = z_1 + z_2$. Now by Theorem 1, $\hat{u} \in \mathrm{BV}(\Omega)$ and there exists a $g \in H_0(\mathrm{div}; \Omega)$ such that $-\mathrm{div}\, g = z_1$, $\|g\|_\infty \leq 1$ and $\theta(g, \mathrm{D}u) = 1$, $|\mathrm{D}u|$-almost everywhere. Hence, we have

$$\mathrm{div}\, g = z_2.$$

Now since $\mathrm{div}\, g \in \partial\mathcal{I}_U(\hat{u})$, there exists, according to Theorem 3, $\lambda = (\lambda_n)_{n\in\mathbb{N}} \in \ell^2$ satisfying the element-wise conditions in 2.(*d*), as well as $A^*\lambda = \mathrm{div}\, g$. Finally the characterization of A^*, as given in Proposition 5, implies $\mathrm{div}\, g = \sum_{n\in\mathbb{N}} \lambda_n a_n$.

(2) $\Rightarrow$ (1) : Conditions (*a*) and (*b*) together with $\hat{u} \in \mathrm{BV}(\Omega)$ imply by Theorem 1 that $-\mathrm{div}\, g \in \partial\,\mathrm{TV}(\hat{u})$, while (*c*) and (*d*) together with $\hat{u} \in U$ imply by Theorem 3 that $\mathrm{div}\, g \in \partial\mathcal{I}_U(\hat{u})$. Hence $0 \in \partial\left(\mathrm{TV} + \mathcal{I}_U\right)(\hat{u})$ and $\hat{u}$ is a minimizer. ■

5. A numerical algorithm and examples. In this section we define the discrete minimization problem related to artifact-free JPEG decompression and present a solution strategy at use of a primal-dual algorithm. Further we show and discuss the outcome of numerical experiments with this framework.

To keep the illustration of the practical implementation simple, we only consider quadratic images, knowing that the generalization to rectangular images is straightforward. We define the space of discrete images as $X = \mathbb{R}^{8k\times 8k}$ together with the inner product

$$(u, v)_X = \sum_{0\leq i,j<8k} u_{i,j} v_{i,j}.$$

We further define the space $Y = X \times X$ together with the inner product

$$(u, v)_Y = \sum_{0 \le i,j < 8k} u^1_{i,j} v^1_{i,j} + u^2_{i,j} v^2_{i,j}.$$

The discrete minimization problem can then be written as

$$\min_{x \in X} F(Kx) + G(x) \tag{5.1}$$

where, with $|\cdot|$ being the Euclidean norm on $\mathbb{R}^2$, we define

$$\begin{aligned} F : Y &\to \mathbb{R} \\ y &\mapsto F(y) = \sum_{0 \le i,j < 8k} |(y^1_{i,j}, y^2_{i,j})| \end{aligned} \tag{5.2}$$

to be a discrete L^1 norm and

$$\begin{aligned} K : X &\to Y \\ x &\mapsto \nabla x, \end{aligned} \tag{5.3}$$

where

$$(\nabla x)^1_{i,j} = \begin{cases} (x_{i+1,j} - x_{i,j}) & \text{if } i < 8k, \\ 0 & \text{if } i = 8k, \end{cases}$$

$$(\nabla x)^2_{i,j} = \begin{cases} (x_{i,j+1} - x_{i,j}) & \text{if } j < 8k, \\ 0 & \text{if } j = 8k, \end{cases}$$

to be a discrete gradient operator using forward differences and a replication of image data as boundary condition. As there are only 8×8 local DCT basis elements per 8×8 block, we use, in contrast to Subsection 4.1, only two indices to identify a global basis element. The indices related to the local basis elements as in (4.1) for given data values $(z_{i,j})_{0 \le i,j < 8k}$ can then be obtained by taking the corresponding global indices modulo 8. Moreover, the 8×8 given quantization values are simply repeated up to $8k \times 8k$. Thus, with given image data $(z_{i,j})_{0 \le i,j < 8k}$ and given quantization values $(q_{n,m})_{0 \le n,m < 8k}$, both provided by the compressed JPEG object, we can define the data intervals as

$$J_{i,j} = \left[q_{i,j} z_{i,j} - \frac{q_{i,j}}{2}, q_{i,j} z_{i,j} + \frac{q_{i,j}}{2} \right], \quad 0 \le i, j < 8k. \tag{5.4}$$

The DCT-data set can then be given as

$$I = \left\{ v \in \mathbb{R}^{8k \times 8k} \,|\, v_{i,j} \in J_{i,j},\, 0 \le i, j < 8k \right\}. \tag{5.5}$$

Letting $A : X \to X$ be the blockwise discrete cosine transform operator, defined by

$$(Ax)_{8i+p,8j+q} = C_p C_q \sum_{n,m=0}^{7} x_{8i+n,8j+m} \cos\left(\frac{\pi(2n+1)p}{16} \right) \cos\left(\frac{\pi(2m+1)q}{16} \right), \tag{5.6}$$

for $0 \leq i, j < k$, $0 \leq p, q < 8$, where

$$C_s = \begin{cases} \frac{1}{\sqrt{8}} & \text{if } s = 0, \\ \frac{1}{2} & \text{if } 1 \leq s \leq 7, \end{cases}$$

the set of all possible source images is defined as

$$U = \{x \in X \,|\, Au \in I\} \tag{5.7}$$

and, consequently,

$$\begin{aligned} G : X &\to \overline{\mathbb{R}} \\ x &\mapsto G(x) = \mathcal{I}_U(x) = \begin{cases} 0 & \text{if } x \in U, \\ \infty & \text{else.} \end{cases} \end{aligned} \tag{5.8}$$

Remark 3. *Note that the blockwise DCT operator according to (5.6) is a basis transformation operator related to an orthonormal discrete block cosine basis, as illustrated in Figure 4.1, and thus an isometric isomorphism.*

5.1. A primal-dual algorithm. In order to obtain an optimal solution to (5.1) we apply the primal-dual algorithm presented in [9]. This algorithm is well-suited for the problem since we can ensure convergence and the iteration steps reduce to projections which are simple and fast to compute numerically. Another general advantage of the algorithm is its global region of attraction. As a drawback one has to mention that in this special case there is no immediate result about convergence rates.

As first step, we use Fenchel duality to formulate an equivalent saddle-point problem. This is justified in the following proposition, where also existence of a solution to (5.1) is established. Again, F^* and G^* denote the convex conjugates of F and G, respectively, as defined in Definition 5.

Proposition 6.

There exist $\hat{x} \in X$ and $\hat{y} \in Y$ such that

$$\hat{x} \in \operatorname*{arg\,min}_{x \in X} F(Kx) + G(x), \tag{5.9}$$

$$\hat{y} \in \operatorname*{arg\,max}_{y \in Y} -\left(G^*(-K^*y) + F^*(y)\right). \tag{5.10}$$

Further

$$\min_{x \in X} F(Kx) + G(x) = \max_{y \in Y} -\left(G^*(-K^*y) + F^*(y)\right) \tag{5.11}$$

and $\tilde{x}, \tilde{y}$ solving (5.9) and (5.10), respectively, is equivalent to $(\tilde{x}, \tilde{y})$ solving

$$\min_{x \in X} \max_{y \in Y} \left((y, Kx)_{Y,Y} - F^*(y) + G(x)\right). \tag{5.12}$$

Proof. Existence of a solution to (5.9) can be derived from boundedness of U by standard arguments. Having this, existence of a solution to the dual problem (5.10), equivalence to the

Algorithm 1 Abstract primal-dual algorithm

- Initialization: Choose $\tau, \sigma > 0$ such that $\|K\|^2\tau\sigma < 1$, $(x^0, y^0) \in X \times Y$ and set $\overline{x}^0 = x^0$
- Iterations ($n \geq 0$): Update $x^n, y^n, \overline{x}^n$ as follows:

$$\begin{cases} y^{n+1} = & (I + \sigma\partial F^*)^{-1}(y^n + \sigma K\overline{x}^n) \\ x^{n+1} = & (I + \tau\partial G)^{-1}(x^n - \tau K^* y^{n+1}) \\ \overline{x}^{n+1} = & 2x^{n+1} - x^n \end{cases}$$

saddle-point problem (5.12) as well as equality (5.11) follow from standard duality results as provided for example in [13, Chapter III], using that $F \circ K$ is continuous. ■

The idea of the primal-dual algorithm is now to solve the saddle-point problem (5.12) by alternately performing a gradient descent step for the primal variable and a gradient ascent for the dual one. Doing so, one has to evaluate resolvents of the subdifferential operators ∂F^* and ∂G, denoted $(I + \sigma\partial F^*)^{-1}$ and $(I + \tau\partial G)^{-1}$, respectively, in each iteration, as can be seen in the abstract formulation of the iteration steps given in Algorithm 1. The choice of the parameters σ, τ such that $\|K\|^2\tau\sigma < 1$, with K as in (5.3), ensures, according to [9, Theorem 1], convergence of the algorithm to a saddle-point.

Now, as already mentioned, for the setting we are considering, the resolvent operators in Algorithm 1 reduce to simple projections. Using that F^* is given by

$$F^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1, \\ \infty & \text{else,} \end{cases}$$

they can be described as follows [11]:

$$\left((I + \sigma\partial F^*)^{-1}(y)\right)_{i,j} = (\mathrm{proj}_1(y))_{i,j} = \begin{cases} y_{i,j} & \text{if } |y_{i,j}| \leq 1 \\ \frac{y_{i,j}}{|y_{i,j}|} & \text{if } |y_{i,j}| > 1 \end{cases} \tag{5.13}$$

and, for given data intervals $J_{i,j} = [l_{i,j}, r_{i,j}]$ and the related data set $I = \left\{v \in \mathbb{R}^{8k\times 8k} \,|\, v_{i,j} \in J_{i,j}\right\}$,

$$(I + \tau\partial G)^{-1}(x) = A^{-1}\,\mathrm{proj}_I(Ax),$$

where

$$(\mathrm{proj}_I(Ax))_{i,j} = \begin{cases} (Ax)_{i,j} & \text{if } (Ax)_{i,j} \in J_{i,j}, \\ r_{i,j} & \text{if } (Ax)_{i,j} > r_{i,j}, \\ l_{i,j} & \text{if } (Ax)_{i,j} < l_{i,j}. \end{cases} \tag{5.14}$$

Therefore, using the projection operator proj_1 and proj_I according to (5.13) and (5.14), respectively, a blockwise DCT operator BDCT according to (5.6) and its inverse denoted by IBDCT, the concrete implementation of the primal-dual algorithm can be given as in

Algorithm 2 Concrete primal-dual algorithm

function TV-JPEG(J_{comp})
  $(z, q) \leftarrow$ Entropy-decoding of JPEG-object J_{comp}
  $I \leftarrow$ Data set as in (5.5)
  $x \leftarrow$ Any initial image from U according to (5.7)
  $\overline{x} \leftarrow x, y \leftarrow 0$, choose $\sigma, \tau > 0$ such that $\sigma\tau \leq 1/8$
  repeat
    $y \leftarrow \text{proj}_1\left(y + \sigma(\nabla \overline{x})\right)$
    $x_{\text{new}} \leftarrow x + \tau(\operatorname{div} y)$
    $x_{\text{new}} \leftarrow \text{BDCT}(x_{\text{new}})$
    $x_{\text{new}} \leftarrow \text{proj}_I(x_{\text{new}})$
    $x_{\text{new}} \leftarrow \text{IBDCT}(x_{\text{new}})$
    $\overline{x} \leftarrow (2x_{\text{new}} - x)$
    $x \leftarrow x_{\text{new}}$
  until Stopping criterion satisfied
  return x_{new}
end function

Algorithm 2. Note that there we already replaced K by ∇ according to (5.3) and $K^* = \nabla^*$ by $-\operatorname{div}$, where

$$\operatorname{div}(y^1_{i,j}, y^2_{i,j}) = \begin{cases} y^1_{i,j} - y^1_{i-1,j} & \text{if } 0 < i < 8k-1, \\ y^1_{i,j} & \text{if } i = 0, \\ -y^1_{i-1,j} & \text{if } i = 8k, \end{cases} + \begin{cases} y^2_{i,j} - y^2_{i,j-1} & \text{if } 0 < j < 8k-1, \\ y^2_{i,j} & \text{if } j = 0, \\ -y^2_{i,j-1} & \text{if } j = 8k. \end{cases}$$

The restriction $\tau\sigma \leq \frac{1}{8}$ as set in Algorithm 2 results from the well-known estimate $\|K\| < \sqrt{8}$.

5.2. Extensions. In order to validate the optimal solutions obtained with the primal-dual algorithm mathematically and also to allow a choice of the level of TV smoothing, we seek for a suitable stopping criterion. Motivated by the results of Proposition 6, especially equation (5.11), we introduce the primal-dual gap $\mathcal{G} : X \times Y \to \overline{\mathbb{R}}$ defined by

$$\mathcal{G}(x, y) = F(Kx) + G(x) + G^*(-K^*y) + F^*(y). \tag{5.15}$$

As already mentioned, F^* is the convex indicator function of the set

$$\{y \in Y \,|\, \|y\|_\infty \leq 1\},$$

while G^* takes the form

$$G^*(x) = \sup_{u \in U} (x, u)_X. \tag{5.16}$$

The following proposition summarizes the basic properties of the primal-dual gap which make it suitable for a stopping criterion.

Proposition 7. *Let (x_n, y_n) be determined by Algorithm 2 and $(\hat{x}, \hat{y})$ be its limit as $n \to \infty$. Then, the primal-dual gap for (x_n, y_n) as defined in (5.15) converges to zero as $n \to \infty$. Further, for any $(x, y) \in X \times Y$, we have*

$$\mathcal{G}(x, y) \geq F(Kx) - F(K\hat{x}). \tag{5.17}$$

Proof. Due to the projections performed during the iterations of the primal-dual algorithm, it is ensured that, for each $n \in \mathbb{N}$, the iterates x_n and y_n are contained in U and $\{y \in Y \,|\, \|y\|_\infty \leq 1\}$, respectively. Thus the primal-dual gap of the iterates reduces to

$$\mathcal{G}(x_n, y_n) = F(Kx_n) + G^*(-K^* y_n).$$

By convergence of the iterates to a saddle-point $(\hat{x}, \hat{y})$, and since by (5.11) we have $\mathcal{G}(\hat{x}, \hat{y}) = 0$, to prove convergence of $\mathcal{G}(x_n, y_n)$ to zero it suffices to show continuity of $F \circ K + G^* \circ (-K^*)$. But, taking into account the representation of G^* in (5.16), this follows from the boundedness of U. Inequality (5.17) is trivially satisfied for $(x, y) \in X \times Y$ in the case that $x \notin U$ or $\|y\|_\infty > 1$, thus to show (5.17) we can suppose $x \in U$ and $\|y\|_\infty \leq 1$. Using this and equation (5.10) it follows that

$$-G^*(-K^* \hat{y}) \geq -G^*(-K^* y)$$

and, consequently,

$$\begin{aligned} \mathcal{G}(x, y) &= \mathcal{G}(x, y) - \mathcal{G}(\hat{x}, \hat{y}) \\ &= F(Kx) - F(K\hat{x}) + G^*(-K^* y) - G^*(-K^* \hat{y}) \\ &\geq F(Kx) - F(K\hat{x}). \end{aligned}$$

■

This result allows, for given $\epsilon > 0$, to use $\mathcal{G}(x_n, y_n) < \epsilon$ as stopping criterion. For determining an optimal solution, we use a normalized primal-dual gap defined by $\overline{\mathcal{G}}(x, y) = \frac{\mathcal{G}(x,y)}{64k^2}$. This is motivated by the inequality $\mathcal{G}(x, y) \geq F(Kx) - F(K\hat{x})$, since the right-hand side is a sum over all pixels which can be made independent of the image size by dividing through the number of pixels. As we will see in Subsection 5.3, this is indeed suitable.

5.3. Numerical experiments. This subsection is devoted to validate our mathematical model and the application of the primal-dual algorithm supplemented by the primal-dual gap based stopping criterion numerically. We have compared the standard JPEG decompression with an optimal solution to the discrete minimization problem for three images possessing different characteristics. The outcome can be seen in Figure 5.1, where also the image size, the iteration number and the compression rate with respect to the file size in bits per pixel (bpp) is given. The optimal solutions were obtained by reducing the normalized primal-dual gap below $10^{-2.5}$. As one can see, especially for the abstract test image, the TV-based reconstruction process removes almost all artifacts and gives a nice and clean reconstruction. Also for the partly cartoon-like cameraman image it leads to significantly improved image quality. Still, especially for the Barbara image, staircasing effects occur which are typical for

TV regularized images. This is a model related problem and to overcome it, one approach is to consider a different regularization functional (see Section 6 for a discussion on that).

However, in order to reduce the staircasing effects and also to save computation time, we suggest a relative decrease of the primal-dual gap to less than one third of the primal-dual gap corresponding to the standard decompression for the primal- and a zero initialization for the dual variable as alternative stopping criterion. This is still an heuristic choice depending on a subjective estimation of image quality and one may perform empirical studies to choose a more appropriate decrease factor. Still, as we see in Figure 5.2, this application of the primal-dual gap, which we will call the relative primal-dual gap stopping criterion, leads to improved image quality, even in comparison to the optimal solution obtained with the primal-dual algorithm, and this already at a low number of iterations.

As already mentioned in the introduction, the idea of applying TV minimization for artifact-free JPEG decompression is already appearing in the literature. In [29] and [2], a projected gradient descent algorithm with fixed and variable step-size, respectively, is proposed to obtain the optimal solution. Figure 5.3 allows to compare a TV based JPEG reconstruction of the cameraman image after 5000 iterations for different algorithms. As one can see, usage of the primal-dual algorithm as proposed in this paper results in the best approximation of the optimal solution. This is shown in the right column of the figure, where the development of the TV semi-norm for the three different implementations is plotted. The proposed primal-dual algorithm reaches a significantly lower TV level than the other ones after 5000 iterations, which is also visible in terms of visual image quality in the left column. Applying the gradient descent algorithm with a fixed step-size of 1 reduces the typical JPEG artifacts, but leads to new artifacts in the solution due to non-convergence. Using a variable step-size does not introduce these new artifacts, since convergence can be assured, but as a result of an iteratively decreasing step-size, the optimal solution is not reached after 5000 iterations. However, one has to mention that since we are interested in optimal solutions of the TV-based model without further modifications, no additional weighting of the TV functional, as proposed in [2], was used for the projected gradient descent algorithm with variable step-size.

To test scalability, we performed the primal-dual algorithm for varying image sizes. Table 5.1 shows the number of necessary iterations for obtaining an optimal solution. As one can see, the number of necessary iterations is quite stable in this respect. The test images are depicted in Figure 5.4. They were originally uncompressed with size of 4096×4096 pixels. We have successively reduced the resolution of the uncompressed image and then applied JPEG compression with a fixed quantization table. Note that again with optimal solution we mean an image having normalized primal-dual gap value below the bound $10^{-2.5}$.

We have also implemented the primal-dual algorithm for the CPU and GPU, using C++ and NVIDIA's Cuda [22], respectively. It can be seen in Table 5.2, that especially the GPU performs very fast, making the proposed method of JPEG decompression suitable for "real life" application. Note however, that the CPU implementation is not parallelized and clearly not optimized while we have simplified the GPU implementation by using the extragradient defined as $\overline{x} = 2x_{\text{new}} - x$ instead of x_{new} as input argument for the evaluation of the primal-dual gap. This allows less storage space demand, while by continuity of TV in finite dimensions, the qualitative behavior of the primal-dual gap remains the same.

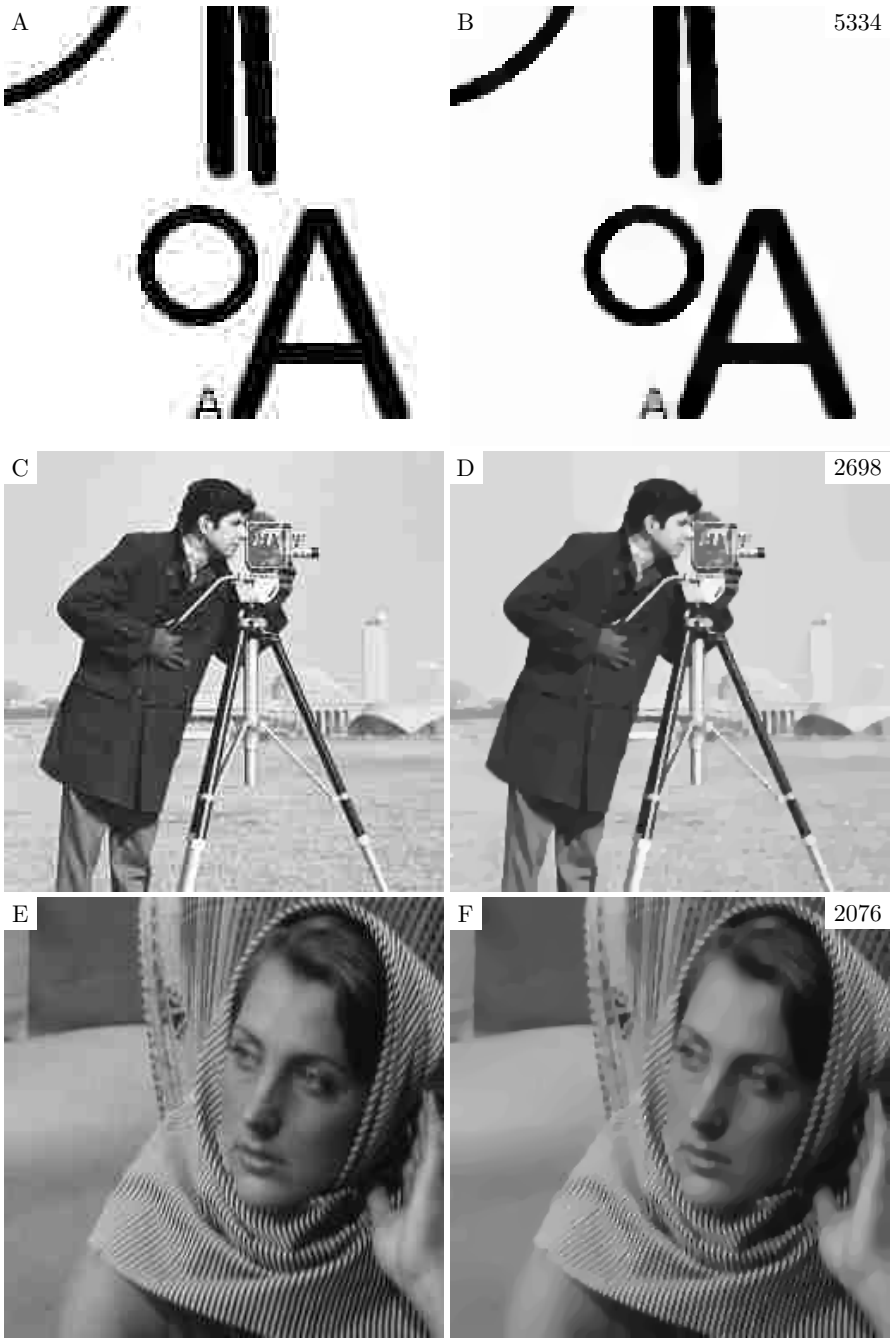

Figure 5.1. *On the left: Standard decompression, on the right: Optimal solution obtained with the primal-dual algorithm with iteration number on top right. A-B: Abstract image at 0.49 bpp (*128×128 *pixels). C-D: Cameraman image at 0.42 bpp (*256×256 *pixels). E-F: Section of the Barbara image at 0.62 bpp (*256×256 *pixels).*

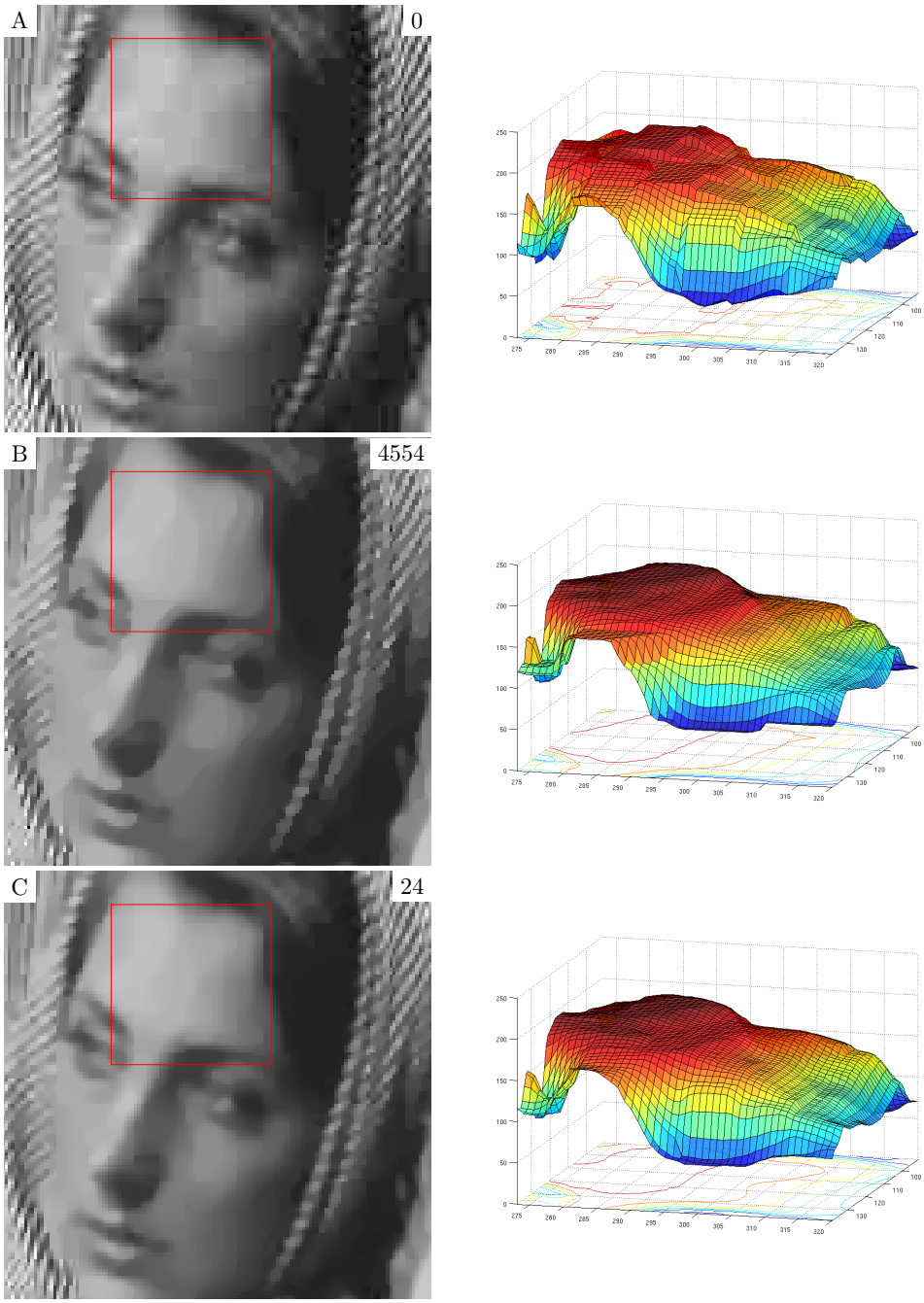

Figure 5.2. *Close-up of the Barbara image at 0.4 bpp with iteration numbers on top right. The marked region on the left is plotted as surface on the right. A: Standard Decompression. B: Optimal solution with normalized primal-dual gap below* $10^{-2.5}$. *C: Solution obtained by with the relative primal-dual gap stopping criterion.*

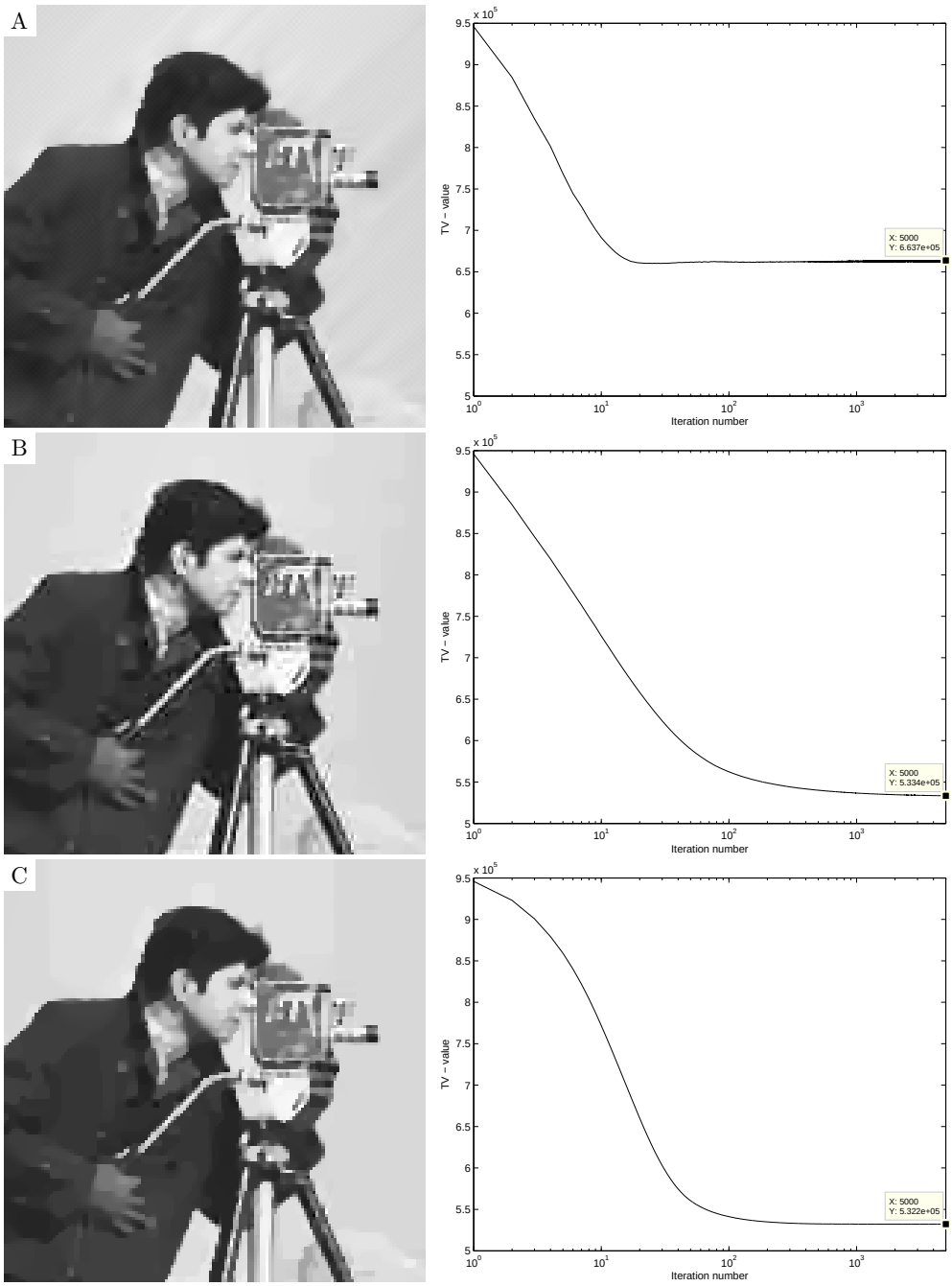

Figure 5.3. *Left: Numerical solutions for different algorithms after 5000 iterations for cameraman image. Right: The development of the* TV *functional during the iterations (logarithmic scale). Top: Projected gradient descent algorithm with fixed step-size of* 1*. Middle: Projected gradient descent algorithm with iteration dependent step-size of* $(n+1)^{-0.5001}$*. Bottom: Primal-dual algorithm with step-size* $\sigma = \tau = \sqrt{1/8}$*.*

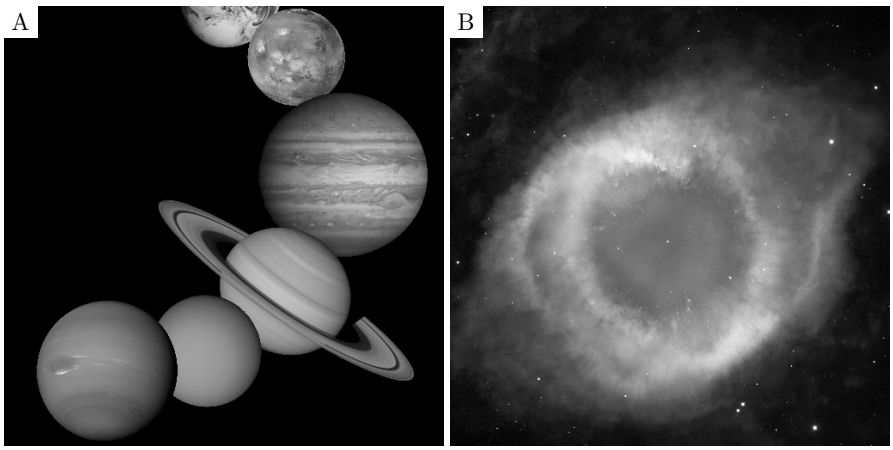

Figure 5.4. *A: Solar System image. B: Helix Nebula image. The images are gratefully taken from the NASA homepage [20, 19].*

	4096	**2048**	**1024**	**512**	**256**	**128**	**64**	**32**
Solar System	3897	3365	3087	2699	2460	2161	2823	2221
Helix Nebula	2252	2225	2282	2735	3219	3725	6161	2897

Table 5.1

Iteration numbers for different image sizes for Solar System *and* Helix Nebula *images. The optimal solutions were obtained as having a normalized primal-dual gap value below* $10^{-2.5}$. *The images are quadratic with width in pixels given in the first row.*

6. Discussion and conclusions. We analyzed a class of abstract imaging problems expressed as minimizing the sum of a regularization term and the indicator function of a convex set describing data compression modalities. This framework is applicable to different problems in digital image processing, such as artifact-free JPEG decompression or DCT-based zooming. The analytical treatment of the infinite dimensional problem formulation provides a basis for further research on qualitative properties of solutions and usage of more sophisticated numerical algorithms, possibly for applications in a different context.

The number of existing publications related to artifact-free JPEG decompression indicates that the resolution of this issue is of high interest to the digital imaging community. The numerical solutions obtained with the TV-based model and algorithm confirm effectivity of TV regularization in reducing noise without over-smoothing sharp boundaries. We see, however, that being approximately piecewise constant is not a completely accurate assumption for realistic images. The relative primal-dual gap stopping criterion is observed to reduce staircasing effects while allowing to perform only a few iterates to reconstruct an improved image. An additional weighting of the TV functional to reduce blocking effects, as proposed in [2], seems to be unnecessary, while we can confirm the benefit of early stopping for visual image quality.

As already mentioned, the observation that optimal solutions to the minimization problem often expose typical staircasing artifacts indicates that there is still need for a better

Device	512×512	1600×1200	4272×2848
CPU (AMD Phenom 9950)	0.5889	4.4592	28.257
GPU (Nvidia GTX 280)	0.0109	0.047	0.2651

Table 5.2

CPU and GPU computation time in seconds for decompressing three different image sizes. The reconstructions were obtained, using the relative primal-dual gap based stopping criterion, at iteration 15 for the first two, smaller images and at iteration 14 for the largest image.

regularization functional. The *Total Generalized Variation* (TGV), as recently proposed in [6], seems to resolve this issue. It can be seen as a generalization of the TV functional and allows a qualitative choice of smoothing, while still maintaining the advantages of TV.

We expect that similar implementations, using basically the same primal-dual framework but TGV as regularization, will result in far better image quality, without producing staircase effects. Motivated by this, the establishment of similar analytical and improved practical results as provided here, but for the TGV functional, will be the focus of further research.

REFERENCES

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

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

[3] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, 2000.

[4] G. Anzellotti, *Pairings between measures and bounded functions and compensated compactness*, Mat. Pura Appl., IV, 135 (1983), pp. 293–318.

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

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

[7] E. Casas, K. Kunisch, and C. Pola, *Regularization by functions of bounded variation and application to image enhancement*, Appl. Math. Optim., 40 (1999), pp. 229–257.

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

[9] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120–145.

[10] T. F. Chan and S. Esedoḡlu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math, 65 (2005), pp. 1817–1837.

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

[12] Y. Dong, M. Hintermüller, and Marrick Neri, *An efficient primal-dual method for L^1 TV image restoration*, SIAM J. Imaging Science, 2 (2009), pp. 1168–1189.

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

[14] L. C. Evans, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992.

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

[16] M. Hintermüller and K. Kunisch, *Total bounded variation as bilaterally constrained optimization problem*, SIAM J. Appl. Math., 64 (2004), pp. 1311–1333.

[17] M. Hintermüller and M. M. Rincon-Camacho, *Expected absolute value estimators for a spatially*

adapted regularization parameter choice rule in L^1-TV-based image restoration, Inverse Problems, 26 (2010), pp. 085005, 30.

[18] T. Kartalov, Z. A. Ivanovski, L. Panovski, and L. J. Karam, *An adaptive POCS algorithm for compression artifacts removal*, in 9th International Symposium on Signal Processing and Its Applications, 2007, pp. 1–4.

[19] NASA, *http://hubblesite.org/gallery/album/nebula/hires/true/*.

[20] ———, *http://photojournal.jpl.nasa.gov/favorites/PIA03153*.

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

[22] NVIDIA, *NVIDIA CUDA programming guide 2.0*. NVIDIA Cooperation, 2008.

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

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

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

[26] A. Vretblad, *Fourier Analysis and Its Applications*, Springer, 2003.

[27] Gregory K. Wallace, *The JPEG still picture compression standard*, Commun. ACM, 34 (1991), pp. 30–44.

[28] C. Weerasinghe, A. W.-C. Liew, and H. Yan, *Artifact reduction in compressed images based on region homogeneity constraints using the projection onto convex sets algorithm*, IEEE Transactions on Circuits and Systems for Video Technology, 12 (2002), pp. 891–897.

[29] S. Zhong, *Image coding with optimal reconstruction*, in International Conference on Image Processing, vol. 1, 1997, pp. 161–164.

[30] William P. Ziemer, *Weakly Differentiable Functions*, Springer, 1989.

[31] J. J. Zou and H. Yan, *A deblocking method for BDCT compressed images based on adaptive projections*, IEEE Transactions on Circuits and Systems for Video Technology, 15 (2005), pp. 430–435.