AUTOMATED PARAMETER SELECTION IN THE L^1-L^2-TV MODEL FOR REMOVING GAUSSIAN PLUS IMPULSE NOISE

ANDREAS LANGER*

Abstract. The minimization of a functional consisting of a combined L^1/L^2-data-fidelity term and a total variation term, named L^1-L^2-TV model, is considered to remove a mixture of Gaussian and impulse noise in images, which are possibly additionally deformed by some convolution operator. We investigate analytically the stability of this model with respect to its parameters and link it to a constrained minimization problem. Based on these investigations and a statistical characterization of the mixed Gaussian-impulse noise a fully automated parameter selection algorithm for the L^1-L^2-TV model is presented. It is shown by numerical experiments that the proposed method finds parameters with which noise is removed considerably while features are preserved in images.

1. Introduction. Total variation as regularization in image restoration was first introduced in [58] and has received considerable attention in image processing. This is in particular due to its ability to preserve edges in images [13, 20]. In this context, one typically minimizes a functional that consists of a data-fidelity term, which enforces the consistency between the recovered and the measured image, and the total variation as a regularization term. The choice of the data term typically depends on the type of noise affecting the measured image. Usually images are corrupted by different types of noise, such as Gaussian noise, Poisson noise, and impulse noise. This contamination usually happens during image acquisition, which describes the process of capturing an image by a camera and converting it into a measurable entity [51], and image transmission. If no data is lost, i.e., the image is not affected by impulse noise, then mixed Poisson-Gaussian noise can be efficiently transformed into additive white Gaussian noise [32]. This might be the reason why most of the literature is solely dedicated to Gaussian denoising. The task of removing this type of noise has been successfully performed by using a quadratic L^2-data-fidelity term in first order methods, see e.g. [12, 14, 15, 19, 22, 23, 24, 25, 28, 36, 52, 55, 62, 66], as well as in second order methods, see e.g. [41]. In this approach, which we refer to as the L^2-TV model, the original image $\hat{u}$ is recovered from the observed data g by solving

$$\min_{u\in BV(\Omega)} \alpha\|Tu-g\|^2_{L^2(\Omega)} + |Du|(\Omega), \tag{1.1}$$

where $\Omega \subset \mathbb{R}^2$ is an open bounded set with Lipschitz boundary, T is a bounded linear operator modeling the image-formation device (if the image is only corrupted by noise one sets $T = I$), and $\alpha > 0$ is a parameter. We recall, that for $u \in L^1(\Omega)$

$$V(u,\Omega) := \sup\left\{ \int_\Omega u \operatorname{div}\phi dx : \phi \in [C^1_c(\Omega)]^2, \|\phi\|_{L^\infty(\Omega)} \le 1 \right\}$$

is the variation of u in Ω. Here, $L^q(\Omega)$, with $q \in [1,\infty]$, denotes the usual Lebesgue space [2] and $C^l_c(\Omega)$, $l \in \mathbb{N}$, is the space of l-times continuously differentiable functions with compact support in Ω. In the event that $V(u,\Omega) < \infty$ we denote $|Du|(\Omega) = V(u,\Omega)$ and call it the total variation of u in Ω; see [4, 35] for more details. If $u \in W^{1,1}(\Omega)$, then $|Du|(\Omega) = \int_\Omega |\nabla u| dx$. Further, $BV(\Omega)$ denotes the space of functions with bounded variation, i.e., $u \in BV(\Omega)$ if and only if $V(u,\Omega) < \infty$. The space $BV(\Omega)$ endowed with the norm $\|u\|_{BV} = \|u\|_{L^1(\Omega)} + |Du|(\Omega)$ is a Banach space [35].

The L^2-TV model usually does not yield a satisfactory restoration in the presence of impulse noise. This type of noise is usually constituted due to malfunctioning pixels in camera sensors, faulty memory locations in hardware, or transmission over noisy digital links. There are two commonly used models of impulse noise considered in the literature. The first one, called *salt-and-pepper noise*, where the noisy image g is given by

$$g(x) = \begin{cases} 0 & \text{with probability } s_1 \in [0,1), \\ 1 & \text{with probability } s_2 \in [0,1), \\ T\hat{u}(x) & \text{with probability } 1 - s_1 - s_2, \end{cases}$$

* Institute for Applied Analysis and Numerical Simulation, University of Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany, `andreas.langer@mathematik.uni-stuttgart.de`

with $1 - s_1 - s_2 > 0$ [17]. Here and in the rest of the paper we assume that $T\hat{u}$ is in the dynamic range $[0, 1]$, i.e., $0 \leq T\hat{u} \leq 1$. The second model is called *random-valued impulse noise*, where g is described as

$$g(x) = \begin{cases} c & \text{with probability } s \in [0,1), \\ T\hat{u}(x) & \text{with probability } 1 - s, \end{cases}$$

with c being a uniformly distributed random variable in the image intensity range $[0, 1]$. For impulse noise contaminated images a more successful approach uses instead of a quadratic L^2-data-fidelity term a non-smooth L^1-data-fidelity term [3, 53, 54]. That is, instead of (1.1) one optimizes the following minimization problem

$$\min_{u \in BV(\Omega)} \alpha \|Tu - g\|_{L^1(\Omega)} + |Du|(\Omega), \tag{1.2}$$

which we call the L^1-TV model.

Instead of assuming that an image is only contaminated by one type of noise, in this paper we consider a mixture of Gaussian and impulse noise. Recently in [39] an L^1-L^2-data-fidelity term has been introduced and shown to be suited to the task of removing mixed Gaussian-impulse noise. In this approach, which we call L^1-L^2-TV model, an image is restored by solving

$$\min_{u \in BV(\Omega)} \alpha_1 \|T_1 u - g_1\|_{L^1(\Omega)} + \alpha_2 \|T_2 u - g_2\|^2_{L^2(\Omega)} + |Du|(\Omega), \tag{1.3}$$

where $T_i : L^2(\Omega) \to L^2(\Omega)$ is a bounded linear operator, $g_i \in L^2(\Omega)$ is a given datum, and $\alpha_i \geq 0$ for $i = 1, 2$ with $\alpha_1 + \alpha_2 > 0$. For the case of removing a mixture of Gaussian and impulse noise from an image g one typically sets $T_1 = T_2$ and $g_1 = g_2 = g$ in (1.3). In this setting it is easy to see that the L^1-L^2-TV model (1.3) is a generalization of (1.1) and (1.2). In particular, if we set $\alpha_2 = 0$ in (1.3) then we obtain the L^1-TV model while for $\alpha_1 = 0$ we obtain the L^2-TV model. Modifications of the L^1-L^2-TV model have been presented in [37], where the total variation is replaced by $\|Wu\|_1$ with W being a wavelet tight frame transform, and in [48], where the second order total generalized variation [9] has been used as regularization term and box-constraints, which assure that the reconstruction lies in the respective dynamic range, are incorporated. We also note, that for impulse noise-dominated contamination of image data the implementation of an impulse noise detector, such as the one in [16] and the references therein, enhance the model.

Other approaches for removing mixed Gaussian-impulse noise studied in the literature usually start by estimating or detecting outliers (impulse noise) in the image and then adapt or use a Gaussian noise removal; see for example [10, 34, 42, 46, 63, 65]. In general, algorithms for Gaussian plus impulse noise removal may be classified in the following way: filter approaches [33, 56, 65], regularization based approaches [10, 29, 34, 42, 46, 59, 63, 64], Bayesian-based approaches [49], and patch-based approaches [27, 45, 47].

The L^1-L^2-TV model and its aforementioned modifications clearly fall into the class of regularization based approaches and (the restoration quality of) its solution highly depends on the proper choice of α_i, $i = 1, 2$. In particular, small α_1 and α_2, which lead to an over-smoothed reconstruction, not only remove noise but also eliminate details in the image. On the contrary, large α_1 and α_2 lead to solutions that fit the given data properly but retain noise. Note, that α_1 and α_2 weight the importance of the L^1-term and L^2-term. In particular, we expect α_1 to be large if the noise in the image is impulse noise dominated, while for Gaussian noise dominated images α_2 should be sufficiently large. Hence a good reconstruction can be achieved by choosing α_1 and α_2 such that a good compromise of the aforementioned effects are made. In [48] it is suggested to select the parameters according to the variance σ^2 of the Gaussian noise and the energy of the impulse noise, i.e.,

$$\alpha_1 = \frac{E_I}{E_I + \sigma^2} \quad \text{and} \quad \alpha_2 = \frac{\sigma^2}{E_I + \sigma^2}, \tag{1.4}$$

where $E_I = \frac{s_1 + s_2}{2}$ for salt-and-pepper noise and $E_I = \frac{s}{3}$ for random-valued impulse noise. It is demonstrated in [48] that by setting the parameters according to (1.4) suitable restorations are obtained. Note, that this parameter selection depends on the noise-levels of the different contained noises without using the statistical behavior of mixed Gaussian-impulse noise.

For more general problems including (1.3) with $T_1 = T_2 = I$ in [26] based on a training set of pairs $(g_k, \hat{u}_k)$, $k = 1, 2, \dots, N \in \mathbb{N}$, where g_k is the noisy image and $\hat{u}_k$ represents the original image, a bilevel optimization approach is presented, which computes suitable parameters of the corresponding image model. Similar approaches are also discussed in [11, 43] and references therein. However, since in our setting we do not have a training set given, these approaches are not applicable here.

In this paper we are investigating the statistical characterization of mixed Gaussian-impulse noise and use it to formulate a fully automated parameter adjustment strategy based on the discrepancy principle to compute suitable α_1 and α_2 for (1.3). In order to construct such a method, we link the L^1-L^2-TV model with a constrained minimization problem consisting of two constraints, one related to the L^1-term and the other related to the L^2-term. In particular, there exist α_1 and α_2 such that a solution of the constrained minimization problem is also a minimizer of the L^1-L^2-TV model. Based on this result and on the constraints we suggest an iterative adjustment scheme, which either increases or decreases the parameter α_i, $i = 1, 2$. Since this update rule generates monotonic sequences of parameters, we are able to show that the proposed method indeed converges. Moreover, in each iteration the L^1-L^2-TV model has to be solved with the current parameters. An algorithm for solving such a minimization problem is presented in [39] without any theoretical justification of its convergence. Here we use the same algorithm and provide a convergence proof. In our numerical experiments we demonstrate that the proposed automated parameter selection method indeed finds parameters α_1 and α_2 such that the corresponding restoration is better with respect to some restoration quality measure than the one obtained with (1.4).

The rest of the paper is organized as follows: In Section 2 a statistical characterization of mixed Gaussian-impulse noise is given. The link between the L^1-L^2-TV model and a constrained minimization problem is investigated in Section 3. In this context, a collection of interesting properties of the L^1-L^2-TV model is given. For example we prove a stability result of its minimizers with respect to the parameters α_1 and α_2. Based on the constrained minimization problem together with the statistical characterization of the noise in Section 4 our proposed parameter selection algorithm is presented. This algorithm requires in each iteration the solution of the L^1-L^2-TV model for which a solution algorithm is stated in Section 5 together with its convergence properties. In Section 6 we show numerical experiments which demonstrate that the proposed algorithm indeed finds parameters α_1 and α_2 that provide a good compromise of the effects described above. Finally in Section 7 conclusions are drawn.

2. Statistical characterization of the noise. At a point $x \in \Omega$ the contaminated image, which might be written as $g(x) = T\hat{u}(x) + \eta_{\hat{u}}(x) + \rho_{\hat{u}}(x)$, is a stochastic observation where the random values $\eta_{\hat{u}}(x)$ and $\rho_{\hat{u}}(x)$ depend on the underlying noise. In particular, the random element $\eta_{\hat{u}}$ represents Gaussian noise with zero mean and variance σ^2, while $\rho_{\hat{u}}$ represents salt-and-pepper noise or random-valued impulse noise. The processes of contaminating an image by Gaussian noise and impulse noise are here assumed to be independent from each other, which seems natural, since usually Gaussian and impulse noise are constituted from different physical processes. For example, due to image registration Gaussian noise is added and later digital transmission adds impulse noise. Moreover, for any two points $x, y \in \Omega$ we assume that $\eta_{\hat{u}}(x)$ and $\eta_{\hat{u}}(y)$ as well as $\rho_{\hat{u}}(x)$ and $\rho_{\hat{u}}(y)$ are independent, cf. [40]. By analogous considerations as in [40] we obtain the following characterizations.

Gaussian noise. For $\eta_{\hat{u}}$ being normally distributed with zero mean and standard deviation σ the mean ($\mathbb{E}$), variance (Var), and expected absolute value (EAV) are

$$\mathbb{E}(\eta_{\hat{u}}) = 0, \quad \mathrm{Var}(\eta_{\hat{u}}) = \sigma^2, \quad \text{and} \quad \mathrm{EAV}(\eta_{\hat{u}}) = \sqrt{\frac{2}{\pi}}\sigma.$$

Salt-and-pepper noise. If $\rho_{\hat{u}}$ represents salt-and-pepper noise the mean, the variance, and the expected absolute value are depending on $\hat{u}$ and given by

$$\mathbb{E}(\rho_{\hat{u}}) = s_1(1 - T\hat{u}) - s_2 T\hat{u}, \quad \mathrm{Var}(\rho_{\hat{u}}) = s_1(1 - T\hat{u})^2 + s_2(T\hat{u})^2 - (s_1 - (s_1 + s_2)T\hat{u})^2$$

and

$$\mathrm{EAV}(\rho_{\hat{u}}) = s_1 - (s_1 - s_2)T\hat{u}.$$

Assuming that the range of $T\hat{u}$ belongs to the interval $[0, 1]$, we find

$$\mathbb{E}(\rho_{\hat{u}}) \in [-s_2, s_1], \quad \mathrm{Var}(\rho_{\hat{u}}) \in \left[\frac{s_1 s_2^2 + s_1^2 s_2}{(s_1 + s_2)^2}, \max\{s_1 - s_1^2, s_2 - s_2^2\}\right],$$

and

$$\mathrm{EAV}(\rho_{\hat{u}}) \in [\min\{s_1, s_2\}, \max\{s_1, s_2\}].$$

Random-valued impulse noise. For random-valued impulse noise the random variable $\rho_{\hat{u}}$ has the following mean, expected absolute value, and variance:

$$\mathbb{E}(\rho_{\hat{u}}) = s\left(\frac{1}{2} - T\hat{u}\right), \quad \mathrm{EAV}(\rho_{\hat{u}}) = s\left((T\hat{u})^2 - T\hat{u} + \frac{1}{2}\right),$$

and

$$\mathrm{Var}(\rho_{\hat{u}}) = s\left(\frac{1}{3} - T\hat{u} + (T\hat{u})^2\right) - s^2\left(\frac{1}{4} - T\hat{u} + (T\hat{u})^2\right).$$

Since $T\hat{u} \in [0, 1]$, we have

$$\mathbb{E}(\rho_{\hat{u}}) \in \left[-\frac{s}{2}, \frac{s}{2}\right], \quad \mathrm{EAV}(\rho_{\hat{u}}) \in \left[\frac{s}{4}, \frac{s}{2}\right], \quad \text{and} \quad \mathrm{Var}(\rho_{\hat{u}}) \in \left[\frac{s}{12}, \frac{s}{3} - \frac{s^2}{4}\right].$$

Mixed noise. Since $\eta_{\hat{u}}$ and $\rho_{\hat{u}}$ are independent random variables, we obtain for a combination of Gaussian and impulse noise that the variance is given by

$$\nu_2 := \mathrm{Var}(\eta_{\hat{u}} + \rho_{\hat{u}}) = \mathrm{Var}(\eta_{\hat{u}}) + \mathrm{Var}(\rho_{\hat{u}})$$

while the expected absolute value can be estimated from below and above by

$$|\mathbb{E}(\eta_{\hat{u}} + \rho_{\hat{u}})| \leq \nu_1 := \mathrm{EAV}(\eta_{\hat{u}} + \rho_{\hat{u}}) \leq \mathrm{EAV}(\eta_{\hat{u}}) + \mathrm{EAV}(\rho_{\hat{u}}).$$

Since $T\hat{u} \in [0, 1]$, we obtain by the above estimates the following bounds:

$$\begin{aligned}
\text{Gaussian + salt-and-pepper:} \quad & \nu_1 \in \left[0, \sqrt{\frac{2}{\pi}}\sigma + \max\{s_1, s_2\}\right] \\
& \nu_2 \in \sigma^2 + \left[\frac{s_1 s_2^2 + s_1^2 s_2}{(s_1 + s_2)^2}, \max\{s_1 - s_1^2, s_2 - s_2^2\}\right]. \\
\text{Gaussian + random-valued:} \quad & \nu_1 \in \left[0, \sqrt{\frac{2}{\pi}}\sigma + \frac{s}{2}\right] \\
& \nu_2 \in \sigma^2 + \left[\frac{s}{12}, \frac{s}{3} - \frac{s^2}{4}\right].
\end{aligned} \tag{2.1}$$

3. Constrained versus unconstrained problem. We define the functional in (1.3) as

$$\mathcal{J}_{\alpha_1,\alpha_2}(u) := \alpha_1\|T_1u - g_1\|_{L^1(\Omega)} + \alpha_2\|T_2u - g_2\|^2_{L^2(\Omega)} + |Du|(\Omega)$$

and link the optimization problem (1.3) to the constrained minimization problem

$$\min_{u\in BV(\Omega)} |Du|(\Omega) \quad \text{subject to (s.t.)} \quad \|T_1u - g_1\|_{L^1(\Omega)} \le \nu_1|\Omega| \quad \text{and} \quad \|T_2u - g_2\|^2_{L^2(\Omega)} \le \nu_2|\Omega|, \tag{3.1}$$

where $\nu_1, \nu_2 \ge 0$ denote the expected absolute value and the variance of the underlying noise, respectively. Here, we assume that ν_1 and ν_2 are fixed constants in the intervals as specified in (2.1). However, in our numerical experiments we report on results where ν_1 and ν_2 are chosen empirically based on some approximation of the true image. If $g_1 = g_2$, then we easily see from the previous section that ν_1 and ν_2 are correlated by the statistical values σ, s_1, s_2, and s of the noise. For example, if $\nu_1 = 0$, then also $\nu_2 = 0$ and hence no noise is present. Note, that in the general case where $g_1 \neq g_2$ is allowed, such a correlation might not be valid.

3.1. Existence of minimizers. For showing existence of a solution of (3.1) we start by adapting a result of [1]. For the sake of completeness we provide a proof which combines results of [30] and [40].

LEMMA 3.1. *Assume there exist $i \in \{1,2\}$ such that T_i does not annihilate constant functions, i.e., $T_i\chi_\Omega \neq 0$, where $\chi_\Omega(x) = 1$ if $x \in \Omega$. Then $\|u\|_{BV} \to \infty$ implies $\mathcal{J}_{1,1}(u) \to \infty$.*

Proof. Any $u \in BV(\Omega)$ can be written as $u = t + v$ with $t = \frac{\int_\Omega u dx}{|\Omega|}\chi_\Omega$ and $\int_\Omega v dx = 0$. Using the inequality $\|v\|_1 \le C_1|Dv|(\Omega)$ with $C_1 > 0$, which follows from the Sobolev inequality $\|v\|_{L^2(\Omega)} \le C_2|Dv|(\Omega)$ with $C_2 > 0$ [35, p. 24], we obtain

$$\|u\|_{BV} \le \|t\|_{BV} + \|v\|_{BV} = \|t\|_1 + |Dt|(\Omega) + \|v\|_1 + |Dv|(\Omega) \le \|t\|_1 + (C_1+1)|Dv|(\Omega).$$

Case 1: Assume T_1 and T_2 do not annihilate constant functions, i.e., there exist $C_3, C_4 > 0$ independent of t such that $\|T_1t\|_{L^1(\Omega)} \ge C_3\|t\|_{L^1(\Omega)}$ and $\|T_2t\|_{L^2(\Omega)} \ge C_4\|t\|_{L^1(\Omega)}$. Then we obtain,

$$\begin{aligned}
\mathcal{J}_{1,1}(u) &= |Du|(\Omega) + \|T_1u - g_1\|_{L^1(\Omega)} + \|T_2u - g_2\|^2_{L^2(\Omega)} \\
&\ge |Dv|(\Omega) + \|T_1t\|_{L^1(\Omega)} - \|T_1v - g_1\|_{L^1(\Omega)} + \|T_2t\|_{L^2(\Omega)}\left(\|T_2t\|_{L^2(\Omega)} - 2\|T_2v - g_2\|_{L^2(\Omega)}\right) \\
&\ge |Dv|(\Omega) + \|T_1t\|_{L^1(\Omega)} - \left(\|T_1\|\|v\|_{L^2(\Omega)} + \|g_1\|_{L^1(\Omega)}\right) \\
&\qquad + \|T_2t\|_{L^2(\Omega)}\left(\|T_2t\|_{L^2(\Omega)} - 2\left(\|T_2\|\|v\|_{L^2(\Omega)} + \|g_2\|_{L^2(\Omega)}\right)\right) \\
&\ge (1 - \|T_1\|C_2)|Dv|(\Omega) + C_3\|t\|_{L^1(\Omega)} - \|g_1\|_{L^1(\Omega)} \\
&\qquad + C_4\|t\|_{L^1(\Omega)}\left(C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right)\right).
\end{aligned} \tag{3.2}$$

If $C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right) \ge 0$, then

$$\mathcal{J}_{1,1}(u) \ge (1 - \|T_1\|C_2)|Dv|(\Omega) + C_3\|t\|_{L^1(\Omega)} - \|g_1\|_{L^1(\Omega)}$$

which yields

$$\begin{aligned}
\|t\|_{L^1(\Omega)} &\le \frac{\mathcal{J}_{1,1}(u) + (\|T_1\|C_2 - 1)|Dv|(\Omega) + \|g_1\|_{L^1(\Omega)}}{C_3} \\
&\le \frac{\mathcal{J}_{1,1}(u) + (\|T_1\|C_2)\mathcal{J}_{1,1}(u) + \|g_1\|_{L^1(\Omega)}}{C_3} \le C\mathcal{J}_{1,1}(u) + \frac{\|g_1\|_{L^1(\Omega)}}{C_3}
\end{aligned}$$

with $C = \max\{\frac{1}{C_3}, \frac{\|T_1\|C_2}{C_3}\}$, since $|Du|(\Omega) \le \mathcal{J}_{1,1}(u)$. Then we obtain

$$\|u\|_{BV} \le (C + C_1 + 1)\mathcal{J}_{1,1}(u) + \frac{\|g_1\|_{L^1(\Omega)}}{C_3}. \tag{3.3}$$

On the other hand, if $C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right) < 0$, then

$$\|t\|_{L^1(\Omega)} < \frac{2\|T_2\|C_2|Dv|(\Omega) + 2\|g_2\|_{L^2(\Omega)}}{C_4}$$

and consequently

$$\|u\|_{BV} \leq \frac{2\|T_2\|C_2\mathcal{J}_{1,1}(u) + 2\|g_2\|_{L^2(\Omega)}}{C_4} + (C_1+1)\mathcal{J}_{1,1}(u). \tag{3.4}$$

By (3.3) and (3.4) it follows that $\|u\|_{BV} \to \infty$ implies $\mathcal{J}_{1,1}(u) \to \infty$.

Case 2: Assume T_1 annihilates constant functions while T_2 does not. Then instead of (3.2) we get

$$\begin{aligned}\mathcal{J}_{1,1}(u) \geq{}& (1-\|T_1\|C_2)|Dv|(\Omega) - \|g_1\|_{L^1(\Omega)} \\ &+ C_4\|t\|_{L^1(\Omega)}\left(C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right)\right).\end{aligned}$$

If $C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right) \geq 1$, then

$$\|t\|_{L^1(\Omega)} \leq \tilde{C}\mathcal{J}_{1,1}(u) + \frac{\|g_1\|_{L^1(\Omega)}}{C_4},$$

where $\tilde{C} = \max\{\frac{1}{C_4}, \frac{\|T_1\|C_2}{C_4}\}$ and hence

$$\|u\|_{BV} \leq (\tilde{C} + C_1 + 1)\mathcal{J}_{1,1}(u) + \frac{\|g_1\|_{L^1(\Omega)}}{C_4}. \tag{3.5}$$

On the other hand, if $C_4\|t\|_{L^1(\Omega)} - 2\left(\|T_2\|C_2|Dv|(\Omega) + \|g_2\|_{L^2(\Omega)}\right) < 1$, then

$$\|t\|_{L^1(\Omega)} < \frac{1 + 2\|T_2\|C_2|Dv|(\Omega) + 2\|g_2\|_{L^2(\Omega)}}{C_4}$$

and

$$\|u\|_{BV} \leq \frac{2\|T_2\|C_2\mathcal{J}_{1,1}(u) + 2\|g_2\|_{L^2(\Omega)}}{C_4} + (C_1+1)\mathcal{J}_{1,1}(u),$$

which shows together with (3.5) that $\|u\|_{BV} \to \infty$ implies $\mathcal{J}_{1,1}(u) \to \infty$.

Case 3: Assume T_2 annihilates constant functions while T_1 does not. Then

$$\mathcal{J}_{1,1}(u) \geq (1-\|T_1\|C_2)|Dv|(\Omega) + C_3\|t\|_{L^1(\Omega)} - \|g_1\|_{L^1(\Omega)}$$

and hence

$$\|u\|_{BV} \leq (C + C_1 + 1)\mathcal{J}_{1,1}(u) + \frac{\|g_1\|_{L^1(\Omega)}}{C_3},$$

which concludes the proof. □

Next, we define the feasible set

$$U := \{u \in BV(\Omega) : \|T_1 u - g_1\|_{L^1(\Omega)} \leq \nu_1|\Omega| \text{ and } \|T_2 u - g_2\|^2_{L^2(\Omega)} \leq \nu_2|\Omega|\}.$$

Moreover, the convex problem (3.1) is *superconsistent* if there is a feasible point u of the problem such that $\|T_i u - g_i\|^i_{L^i(\Omega)} < \nu_i|\Omega|$ for $i = 1,2$ [57]. Note, that if $g_1 = g_2$ and $T_1 = T_2$, which is the most relevant case for removing a mixture of Gaussian-impulse noise, we have that $U \neq \emptyset$ and (3.1) is superconsistent, cf. for example [8, 15]. On the contrary, if $g_1 \neq g_2$, then the feasible set might be even empty. However, for example an assumption like $\nu_1|\Omega| > \|g_1\|_{L^1(\Omega)}$ and $\nu_2|\Omega| > \|g_2\|^2_{L^2(\Omega)}$ would guarantee the non-emptiness of U and the superconsistency of (3.1). For the sake of generality, in the sequel we will just assume that the set U is not empty or even that the problem (3.1) is superconsistent.

Now we are able to argue the existence of a minimizer of (3.1).

THEOREM 3.2. *Assume there exist $i \in \{1,2\}$ such that T_i does not annihilate constant functions and $U \neq \emptyset$. Then the problem in* (3.1) *has a solution $u \in BV(\Omega)$.*

Proof. Choose an infimal sequence $(u_n)_n \subset U$ of (3.1). Lemma 3.1 yields that $(u_n)_n$ is bounded in $BV(\Omega)$. Then there exists a subsequence $(u_{n_k})_k \subset U$ which converges weakly in $L^2(\Omega)$ to some $u^* \in L^2(\Omega)$. The lower semi-continuity of the total variation $|D\cdot|(\Omega)$ with respect to the $L^2(\Omega)$ topology [1, Theorem 2.3] implies $u^* \in BV(\Omega)$. The sequence $(Du_{n_k})_k$ converges weakly as a measure to Du^* [1, Lemma 2.5]. Since T_1 and T_2 are continuous linear operators, $(T_i u_{n_k})_k$ converges weakly to $T_i u^*$ in $L^2(\Omega)$. By the lower semi-continuity we have

$$\begin{aligned}\|T_1u^* - g_1\|_{L^1(\Omega)} &\leq \liminf_{k\to\infty} \|T_1 u_{n_k} - g_1\|_{L^1(\Omega)} \leq \nu_1|\Omega| \\ \|T_2u^* - g_2\|^2_{L^2(\Omega)} &\leq \liminf_{k\to\infty} \|T_2 u_{n_k} - g_2\|^2_{L^2(\Omega)} \leq \nu_2|\Omega|\end{aligned}$$

and hence $u^* \in BV(\Omega)$ is a solution of (3.1). □

The assumption that at least either T_1 or T_2 does not annihilating constant functions also ensures that $\mathcal{J}_{\alpha_1,\alpha_2}$ has a minimizer if $\alpha_1, \alpha_2 > 0$. In particular, we have the following result.

THEOREM 3.3. *Assume there exist $i \in \{1,2\}$ such that T_i does not annihilate constant functions and $\alpha_i > 0$. Then the problem in* (1.3) *has a solution $u \in BV(\Omega)$. If $\alpha_2 > 0$ and T_2 is injective, then the minimizer u is unique.*

Proof. Existence: The existence of a solution of (1.3) follows from the same arguments as the ones from the proof of Theorem 3.2 by noting that the lower semi-continuity yields

$$\mathcal{J}_{\alpha_1,\alpha_2}(u^*) \leq \liminf_{k\to\infty} \mathcal{J}_{\alpha_1,\alpha_2}(u_{n_k}),$$

where u_{n_k} is a subsequence of a minimizing sequence for $\mathcal{J}_{\alpha_1,\alpha_2}$ and u^* is its limit. Consequently $u^* \in BV(\Omega)$ is a minimizer of $\mathcal{J}_{\alpha_1,\alpha_2}$.

Uniqueness: If $\alpha_2 > 0$, then similar as in the proof of [60, Proposition 3.1], let $u, v \in BV(\Omega)$ be two minimizers of $\mathcal{J}_{\alpha_1,\alpha_2}$ and $T_2 u \neq T_2 v$. Then by the strict convexity of the L^2-term we get

$$\mathcal{J}_{\alpha_1,\alpha_2}\left(\tfrac{u+v}{2}\right) < \tfrac{1}{2}\mathcal{J}_{\alpha_1,\alpha_2}(u) + \tfrac{1}{2}\mathcal{J}_{\alpha_1,\alpha_2}(v) = \min_{w\in BV(\Omega)} \mathcal{J}_{\alpha_1,\alpha_2}(w).$$

Since u and v are minimizers, this inequality cannot be true, and hence $T_2 u = T_2 v$. If T_2 is injective, then we have $u = v$. □

The uniqueness of minimizers can be also obtained by the following stability result, cf. [7, Theorem 10.6] for the L^2-TV model with $T_2 = I$.

PROPOSITION 3.4. *For $g_1, g_2, f_1, f_2 \in L^2(\Omega)$ let the functions $u_g, u_f \in BV(\Omega)$ be minimizers of*

$$\min_{u\in BV(\Omega)} \alpha_1\|T_1u - g_1\|_{L^1(\Omega)} + \alpha_2\|T_2u - g_2\|^2_{L^2(\Omega)} + |Du|(\Omega)$$

and

$$\min_{u\in BV(\Omega)} \alpha_1\|T_1u - f_1\|_{L^1(\Omega)} + \alpha_2\|T_2u - f_2\|^2_{L^2(\Omega)} + |Du|(\Omega),$$

respectively. Then for $\alpha_2 > 0$ and $\alpha_1 \geq 0$ we have that

$$\|T_2(u_f - u_g)\|_{L^2(\Omega)} \leq \tfrac{1}{2}\|f_2 - g_2\|_{L^2(\Omega)} + \tfrac{1}{2\alpha_2}\sqrt{\alpha_2^2\|f_2 - g_2\|^2_{L^2(\Omega)} + 4\alpha_1\alpha_2\|f_1 - g_1\|_{L^1(\Omega)}}.$$

Proof. Define the convex functionals $G_g(u) := \alpha_2\|T_2u - g_2\|^2_{L^2(\Omega)}$, $G_f(u) := \alpha_2\|T_2u - f_2\|^2_{L^2(\Omega)}$, $F_g(u) := \alpha_1\|T_1u - g_1\|_{L^1(\Omega)} + |Du|(\Omega)$, $F_g(u) := \alpha_1\|T_1u - g_1\|_{L^1(\Omega)} + |Du|(\Omega)$ and set $\mathcal{J}_g(u) :=$

$G_g(u) + F_g(u)$ and $\mathcal{J}_f(u) := G_f(u) + F_f(u)$. We extend F_g and F_f to $L^2(\Omega)$ with the value $+\infty$. Moreover, we note that G_g and G_f are Fréchet differentiable.

For $x \in \partial F_g(u)$ and $y \in \partial F_f(v)$ we have by the definition of subdifferential, see for example [31], that

$$\begin{aligned} F_g(w) &\geq F_g(u) + \langle x, w-u\rangle \qquad \text{for all } w \in L^2(\Omega), \\ F_f(\tilde{w}) &\geq F_f(v) + \langle y, \tilde{w}-u\rangle \qquad \text{for all } \tilde{w} \in L^2(\Omega). \end{aligned}$$

Summing up these inequalities for $w = v$ and $\tilde{w} = u$ yields

$$\begin{aligned} \langle x-y, v-u\rangle &\leq \alpha_1 \left(\|T_1 v - g_1\|_{L^1(\Omega)} - \|T_1 v - f_1\|_{L^1(\Omega)} + \|T_1 u - f_1\|_{L^1(\Omega)} - \|T_1 u - g_1\|_{L^1(\Omega)}\right) \\ &\leq \alpha_1 \left(\|f_1 - g_1\|_{L^1(\Omega)} + \|f_1 - g_1\|_{L^1(\Omega)}\right). \end{aligned}$$

From this together with the optimality of u_g and u_f, i.e., $-\partial G_g(u_g) \in \partial F_g(u_g)$ and $-\partial G_f(u_f) \in \partial F_f(u_f)$, we obtain

$$2\alpha_2 \langle T_2(u_f - u_g) + g_2 - f_2, T_2(u_f - u_g)\rangle \leq 2\alpha_1 \|f_1 - g_1\|_{L^1(\Omega)}$$

which is equivalent to

$$2\alpha_2 \|T_2(u_f - u_g)\|^2_{L^2(\Omega)} + 2\alpha_2 \langle g_2 - f_2, T_2(u_f - u_g)\rangle \leq 2\alpha_1 \|f_1 - g_1\|_{L^1(\Omega)}.$$

Using Hölder's inequality implies then

$$\alpha_2 \|T_2(u_f - u_g)\|^2_{L^2(\Omega)} - \alpha_2 \|g_2 - f_2\|_{L^2(\Omega)} \|T_2(u_f - u_g)\|_{L^2(\Omega)} - \alpha_1 \|f_1 - g_1\|_{L^1(\Omega)} \leq 0.$$

This is a quadratic inequality in $\|T_2(u_f - u_g)\|_{L^2(\Omega)}$ and hence calculating the roots yields

$$\|T_2(u_f - u_g)\|_{L^2(\Omega)} \leq \tfrac{1}{2\alpha_2} \left(\alpha_2 \|f_2 - g_2\|_{L^2(\Omega)} + \sqrt{\alpha_2^2 \|f_2 - g_2\|^2_{L^2(\Omega)} + 4\alpha_1\alpha_2 \|f_1 - g_1\|_{L^1(\Omega)}}\right),$$

where we noted that $\sqrt{\alpha_2^2 \|f_2 - g_2\|^2_{L^2(\Omega)} + 4\alpha_1\alpha_2 \|f_1 - g_1\|_{L^1(\Omega)}} \geq \alpha_2 \|f_2 - g_2\|_{L^2(\Omega)}$. ∎

Motivated by results in [8] we link the constrained minimization problem (3.1) to the unconstrained minimization problem (1.3).

THEOREM 3.5. *Assume that T_i does not annihilate constant functions for $i = 1, 2$ and (3.1) is superconsistent. Then there exists an $\alpha = (\alpha_1, \alpha_2) \geq 0$ such that a solution of (1.3) satisfies the constraints in (3.1). Moreover, if $\alpha_i > 0$ then $\|T_i u - g_i\|^i_{L^i(\Omega)} = \nu_i|\Omega|$ for this value of i. In particular, there exist $i \in \{1, 2\}$ such that $\alpha_i > 0$, if at least one of the following conditions holds:*
(C1) $\inf_{c\in\mathbb{R}} \|g_1 - c\|_{L^1(\Omega)} > \nu_1|\Omega|$ *and* $T_1 \cdot 1 = 1$
(C2) $\inf_{c\in\mathbb{R}} \|g_2 - c\|^2_{L^2(\Omega)} > \nu_2|\Omega|$ *and* $T_2 \cdot 1 = 1$

Proof. We define the Lagrange function

$$L(u, \alpha) := |Du|(\Omega) + \alpha_1(\|T_1 u - g_1\|_{L^1(\Omega)} - \nu_1|\Omega|) + \alpha_2(\|T_2 u - g_2\|^2_{L^2(\Omega)} - \nu_2|\Omega|).$$

Let $u^* \in BV(\Omega)$ be a solution of (3.1), then, since the convex problem (3.1) is superconsistent, the Karush-Kuhn-Tucker Theorem [57, p. 182] yields that there exists an $\alpha^* = (\alpha_1^*, \alpha_2^*) \geq 0$ such that

$$L(u^*, \alpha) \leq L(u^*, \alpha^*) \leq L(u, \alpha^*) \tag{3.6}$$

for all $u \in BV(\Omega)$ and all $\alpha = (\alpha_1, \alpha_2) \geq 0$, and

$$\alpha_i^*(\|T_i u^* - g_i\|^i_{L^i(\Omega)} - \nu_i|\Omega|) = 0$$

for $i = 1, 2$. That is, $\|T_i u - g_i\|^i_{L^i(\Omega)} = \nu_i|\Omega|$ if $\alpha_i^* > 0$. By the second inequality in (3.6) we see that u^* is also a minimizer of (1.3).

Let us finally show that not all $\alpha_i = 0$, if condition (C1) and/or (C2) holds. If $\alpha_1 = \alpha_2 = 0$, then for the associated solution $\tilde{u}$ of (3.1) we would have that $|D\tilde{u}|(\Omega) \leq |Du|(\Omega)$ for all $u \in BV(\Omega)$ and hence $\tilde{u} = c \in \mathbb{R}$ is constant. Assume condition (C1) holds, then $T_1 \cdot 1 = 1$ and we obtain

$$\|g_1 - c\|_{L^1(\Omega)} = \|g_1 - T_1\tilde{u}\|_{L^1(\Omega)} \leq \nu_1|\Omega|$$

which is a contradiction to (C1). By the same arguments one shows the statement for (C2). ∎

3.2. Stability of the L^1-L^2-TV model with respect to its parameters. We define the minimum values of the energy $\mathcal{J}_{\alpha_1,\alpha_2}$ by

$$\mathcal{E}(\alpha_1,\alpha_2) := \min_{u\in BV(\Omega)} \mathcal{J}_{\alpha_1,\alpha_2}(u).$$

Following [18] we obtain the following result.

PROPOSITION 3.6. *For any given $g_i \in L^2(\Omega)$ for $i=1,2$ the function $\mathcal{E}$ has the following properties:*

1. $\mathcal{E}(0,0) = 0$.
2. $0 \le \mathcal{E}(\alpha_1,\alpha_2) \le \alpha_1\|g_1\|_{L^1(\Omega)} + \alpha_2\|g_2\|^2_{L^2(\Omega)}$ *for all* $\alpha_1,\alpha_2 \ge 0$.

Proof. Since $\mathcal{E}(0,0) = \min_{u\in BV(\Omega)} \mathcal{J}_{0,0}(u) = \min_{u\in BV(\Omega)} |Du|(\Omega) = 0$, the first statement follows. Further, we have $0 \le \mathcal{E}(\alpha_1,\alpha_2) \le \mathcal{J}_{\alpha_1,\alpha_2}(0)$, which shows the second statement. $\square$

Similar as for the L^1-TV model and the L^2-TV model, see [18, 15], we have a monotonicity property of the data-fidelity terms with respect to the parameters α_1 and α_2.

PROPOSITION 3.7. *Let $\beta_i > \alpha_i \ge 0$ and $\alpha_{\check{i}} \ge 0$ for $i=1,2$ and $\check{i} \in \{1,2\}\setminus\{i\}$. Assume u_{α_1,α_2}, u_{β_1,α_2}, u_{α_1,α_2}, and u_{α_1,β_2} are any four minimizers of $\mathcal{J}_{\alpha_1,\alpha_2}$, $\mathcal{J}_{\beta_1,\alpha_2}$, $\mathcal{J}_{\alpha_1,\alpha_2}$, and $\mathcal{J}_{\alpha_1,\beta_2}$, respectively. Then*

$$\begin{aligned}
&\|T_1u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)} \ge \|T_1u_{\beta_1,\alpha_2} - g_1\|_{L^1(\Omega)} \quad \text{and}\\
&\|T_2u_{\alpha_1,\alpha_2} - g_2\|_{L^2(\Omega)} \ge \|T_2u_{\alpha_1,\beta_2} - g_2\|_{L^2(\Omega)}
\end{aligned}$$

Proof. We start by showing the first inequality. Suppose it is not true, i.e., $\|T_1u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)} < \|T_1u_{\beta_1,\alpha_2} - g_1\|_{L^1(\Omega)}$. From the optimality of u_{α_1,α_2} we have that $\mathcal{J}_{\alpha_1,\alpha_2}(u_{\alpha_1,\alpha_2}) \le \mathcal{J}_{\alpha_1,\alpha_2}(u_{\beta_1,\alpha_2})$. Then we have

$$\begin{aligned}
\mathcal{J}_{\beta_1,\alpha_2}(u_{\alpha_1,\alpha_2}) &= \mathcal{J}_{\alpha_1,\alpha_2}(u_{\alpha_1,\alpha_2}) + (\beta_1-\alpha_1)\|T_1u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)}\\
&\le \mathcal{J}_{\alpha_1,\alpha_2}(u_{\beta_1,\alpha_2}) + (\beta_1-\alpha_1)\|T_1u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)}\\
&< \mathcal{J}_{\alpha_1,\alpha_2}(u_{\beta_1,\alpha_2}) + (\beta_1-\alpha_1)\|T_1u_{\beta_1,\alpha_2} - g_1\|_{L^1(\Omega)} = \mathcal{J}_{\beta_1,\alpha_2}(u_{\beta_1,\alpha_2}),
\end{aligned}$$

which is a contradiction, since u_{β_1,α_2} is a minimizer of $\mathcal{J}_{\beta_1,\alpha_2}$. This proves the first inequality.

By similar arguments one can show again by contradiction that the second inequality holds. $\square$

In order to show a stability result of the L^1-L^2-TV model we adapt [7, Lemma 10.2] to our more general setting.

LEMMA 3.8. *Let $u \in BV(\Omega)$ be a minimizer of $\mathcal{J}_{\alpha_1,\alpha_2}$. Then for every $v \in BV(\Omega)$ we have*

$$\alpha_2\|T_2(u-v)\|^2_{L^2(\Omega)} \le \mathcal{J}_{\alpha_1,\alpha_2}(v) - \mathcal{J}_{\alpha_1,\alpha_2}(u).$$

Proof. By setting $F(u) = \alpha_1\|T_1u - g_1\|_{L^1(\Omega)} + |Du|(\Omega)$ and $G(u) = \alpha_2\|T_2u - g_1\|^2_{L^2(\Omega)}$ the proof is analogue to the one of [7, Lemma 10.2]. $\square$

THEOREM 3.9. *Define $a_i(\alpha_1,\alpha_2) := \frac{1}{\alpha_i}(\alpha_1\|g_1\|_{L^1(\Omega)} + \alpha_2\|g_2\|^2_{L^2(\Omega)})$ for $i=1,2$ and let $\alpha_1,\overline{\alpha}_1,\alpha_2,\overline{\alpha}_2 > 0$. If u_{α_1,α_2} and $u_{\overline{\alpha}_1,\overline{\alpha}_2}$ are minimizers of $\mathcal{J}_{\alpha_1,\alpha_2}$ and $\mathcal{J}_{\overline{\alpha}_1,\overline{\alpha}_2}$ respectively, then we have*

$$\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L(\Omega)} \le \frac{|\alpha_1-\overline{\alpha}_1|}{\alpha_2+\overline{\alpha}_2}\max\{a_1(\overline{\alpha}_1,\overline{\alpha}_2), a_1(\alpha_1,\alpha_2)\} + \frac{|\alpha_2-\overline{\alpha}_2|}{\alpha_2+\overline{\alpha}_2}C =: B, \tag{3.7}$$

where $C = \min\{C_1,C_2\}$ with $C_1 = \max\{a_2(\overline{\alpha}_1,\overline{\alpha}_2), a_2(\alpha_1,\alpha_2)\}$,

$$C_2 = \frac{A_2^2|\alpha_2-\overline{\alpha}_2|}{2(\alpha_2+\overline{\alpha}_2)} + A_2\left((\alpha_2-\overline{\alpha}_2)^2A_2^2 + 4(\alpha_2+\overline{\alpha}_2)|\alpha_1-\overline{\alpha}_1|\max\{a_1(\overline{\alpha}_1,\overline{\alpha}_2), a_1(\alpha_1,\alpha_2)\}\right)^{1/2},$$

and $A_2 := a_2(\overline{\alpha}_1,\overline{\alpha}_2)^{1/2} + a_2(\alpha_1,\alpha_2)^{1/2}$.

If additionally $T_1 = T_2 =: T$ then

$$\|T(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|_{L^2(\Omega)} \le \min\{\sqrt{B}, \tilde{B}\} \tag{3.8}$$

where $\tilde{B} := \frac{|\alpha_1 - \overline{\alpha}_1|}{\alpha_2 + \overline{\alpha}_2}|\Omega|^{1/2} + \frac{|\alpha_2 - \overline{\alpha}_2|}{\alpha_2 + \overline{\alpha}_2} A_2$.

Proof. By Lemma 3.8 we have

$$\alpha_2 \|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le \mathcal{J}_{\alpha_1,\alpha_2}(u_{\overline{\alpha}_1,\overline{\alpha}_2}) - \mathcal{J}_{\alpha_1,\alpha_2}(u_{\alpha_1,\alpha_2})$$
$$\overline{\alpha}_2 \|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le \mathcal{J}_{\overline{\alpha}_1,\overline{\alpha}_2}(u_{\alpha_1,\alpha_2}) - \mathcal{J}_{\overline{\alpha}_1,\overline{\alpha}_2}(u_{\overline{\alpha}_1,\overline{\alpha}_2}).$$

Summing up these inequalities yields

$$\begin{aligned}(\alpha_2 + \overline{\alpha}_2)\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le (\alpha_1 - \overline{\alpha}_1)(\|T_1 u_{\overline{\alpha}_1,\overline{\alpha}_2} - g_1\|_{L^1(\Omega)} - \|T_1 u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)}) \\ +(\alpha_2 - \overline{\alpha}_2)(\|T_2 u_{\overline{\alpha}_1,\overline{\alpha}_2} - g_2\|^2_{L^2(\Omega)} - \|T_2 u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\Omega)}).\end{aligned} \tag{3.9}$$

By the monotonicity property, see Proposition 3.7, we obtain that both terms on the right-hand side of the latter inequality are non-negative. Moreover, note that

$$\|T_1 u_{\alpha_1,\alpha_2} - g_1\|^i_{L^i(\Omega)} \le a_i(\alpha_1, \alpha_2) \tag{3.10}$$

for $i = 1, 2$ and any $\alpha_1, \alpha_2 > 0$, see Proposition 3.6. These observations lead to

$$\begin{aligned}\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le & \frac{|\alpha_1 - \overline{\alpha}_1|}{\alpha_2 + \overline{\alpha}_2} \max\{a_1(\overline{\alpha}_1, \overline{\alpha}_2), a_1(\alpha_1, \alpha_2)\} \\ & + \frac{|\alpha_2 - \overline{\alpha}_2|}{\alpha_2 + \overline{\alpha}_2} \max\{a_2(\overline{\alpha}_1, \overline{\alpha}_2), a_2(\alpha_1, \alpha_2)\}.\end{aligned} \tag{3.11}$$

On the contrary, inequality (3.9) implies

$$\begin{aligned}(\alpha_2 + \overline{\alpha}_2)\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le |\alpha_1 - \overline{\alpha}_1| \max\{a_1(\overline{\alpha}_1, \overline{\alpha}_2), a_1(\alpha_1, \alpha_2)\} \\ + |\alpha_2 - \overline{\alpha}_2| \|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|_{L^2(\Omega)}(a_2(\overline{\alpha}_1, \overline{\alpha}_2)^{1/2} + a_2(\alpha_1, \alpha_2)^{1/2})\end{aligned}$$

where we used the binomial formula $a^2 - b^2 = (a + b)(a - b)$ for $a, b \in \mathbb{R}$, the triangle inequality, and (3.10). This is a quadratic inequality in $\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|_{L^2(\Omega)}$ yielding

$$\begin{aligned}&\|T_2(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|_{L^2(\Omega)} \\ &\quad \le \frac{|\alpha_2 - \overline{\alpha}_2|}{2(\alpha_2 + \overline{\alpha}_2)} A_2 + \frac{\sqrt{(\alpha_2 - \overline{\alpha}_2)^2 A_2^2 + 4(\alpha_2 + \overline{\alpha}_2)|\alpha_1 - \overline{\alpha}_1| \max\{a_1(\overline{\alpha}_1, \overline{\alpha}_2) + a_1(\alpha_1, \alpha_2)\}}}{2(\alpha_2 + \overline{\alpha}_2)}.\end{aligned} \tag{3.12}$$

Squaring (3.12) and combining it with (3.11) yields the assertion.

If $T_1 = T_2 = T$, then from (3.9) by using the triangle inequality and the above used binomial formula we get

$$\begin{aligned}(\alpha_2 + \overline{\alpha}_2)\|T(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|^2_{L^2(\Omega)} \le |\alpha_1 - \overline{\alpha}_1| \|T(u_{\overline{\alpha}_1,\overline{\alpha}_2} - u_{\alpha_1,\alpha_2})\|_{L^1(\Omega)} \\ + (\alpha_2 - \overline{\alpha}_2)\|T(u_{\overline{\alpha}_1,\overline{\alpha}_2} - u_{\alpha_1,\alpha_2})\|_{L^2(\Omega)}(\|T u_{\overline{\alpha}_1,\overline{\alpha}_2} - g_2\|_{L^2(\Omega)} + \|T u_{\alpha_1,\alpha_2} - g_2\|_{L^2(\Omega)}).\end{aligned}$$

By using Hölder inequality on the L^1-term and by using (3.10) we obtain

$$\|T(u_{\alpha_1,\alpha_2} - u_{\overline{\alpha}_1,\overline{\alpha}_2})\|_{L^2(\Omega)} \le \frac{|\alpha_1 - \overline{\alpha}_1|}{\alpha_2 + \overline{\alpha}_2}|\Omega|^{1/2} + \frac{|\alpha_2 - \overline{\alpha}_2|}{\alpha_2 + \overline{\alpha}_2} A_2.$$

Combining the latter inequality with (3.7) we get (3.8), which finishes the proof. □

Remark 3.10. *If $T_2 = I$, then the inequalities* (3.7) *and* (3.8) *provide us with an upper bound on the distance between two solutions obtained with different parameters. In particular, if the parameters in the L^1-L^2-TV model are slightly perturbed, only small changes are expected in the minimizer.*

3.3. Further properties of the L^1-L^2-TV model. In this section we essentially follow [18] to further investigate and prove properties of the L^1-L^2-TV model.

PROPOSITION 3.11. *Given $g_i \in L^2(\Omega)$, $i = 1, 2$. For each $\alpha_1, \alpha_2 > 0$ we denote by u_{α_1,α_2} the unique minimizer of $\mathcal{J}_{\alpha_1,\alpha_2}$ with $T_1 = T_2 = I$. Then the function $(\alpha_1, \alpha_2) \to \alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)} + \alpha_2\|u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\Omega)}$ is continuous.*

Proof. Fix $\alpha_1^*, \alpha_2^* > 0$ and let $u_{\alpha_1^*,\alpha_2^*}$ be the unique minimizer of $\mathcal{J}_{\alpha_1^*,\alpha_2^*}$. Let the sequence $(\alpha_1^j, \alpha_2^j)_j$ converge to (α_1^*, α_2^*). We consider the sequence $(u_{\alpha_1^j,\alpha_2^j})_j$ of corresponding minimizers. From the relation $\mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j}) \leq \mathcal{J}_{\alpha_1^j,\alpha_2^j}(0) = \alpha_1\|g_1\|_{L^1(\Omega)} + \alpha_2\|g_2\|^2_{L^2(\Omega)}$ follows that the sequence $(u_{\alpha_1^j,\alpha_2^j})_j$ has uniformly bounded total variation, L^1-norm, and L^2-norm. Moreover, it implies that

$$\alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Omega)} + \alpha_2\|u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\Omega)} \leq \alpha_1\|g_1\|_{L^1(\Omega)} + \alpha_2\|g_2\|^2_{L^2(\Omega)}. \tag{3.13}$$

The standard compactness property for functions with uniformly bounded total variation on compact sets implies that there exists a sequence, which we denote again by $(u_{\alpha_1^j,\alpha_2^j})_j$, such that $u_{\alpha_1^j,\alpha_2^j} \to v \in L^1_{loc}(\Omega)$ in L^1 on any bounded set. We may then pass to another subsequence to make sure that $u_{\alpha_1^j,\alpha_2^j}(x) \to v(x)$ pointwise almost everywhere as well. Fatou's lemma shows that

$$\begin{aligned}\|v - g_2\|_{L^2(\Omega)} &\leq \liminf_{j\to\infty} \|u_{\alpha_1^j,\alpha_2^j} - g_2\|_{L^2(\Omega)}\\ \|v - g_1\|_{L^1(\Omega)} &\leq \liminf_{j\to\infty} \|u_{\alpha_1^j,\alpha_2^j} - g_1\|_{L^1(\Omega)}\end{aligned} \tag{3.14}$$

and hence $v \in L^2(\Omega)$. By the lower semicontinuity of the total variation, i.e., $|Dv|(\Omega) \leq \liminf_{j\to\infty} |Du_{\alpha_1^j,\alpha_2^j}|(\Omega)$, we get $\mathcal{J}_{\alpha_1^*,\alpha_2^*}(v) \leq \liminf_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j})$.

Let us show that $\mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*}) \geq \limsup_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j})$. Assume it is not true. Then there exists an $\epsilon > 0$ and j arbitrary large such that $\mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*}) \leq \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j}) - \epsilon$. We also have $\lim_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^*,\alpha_2^*}) = \mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*})$ and hence $\mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^*,\alpha_2^*}) < \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j})$ for some large j, which is a contradiction, since $u_{\alpha_1^j,\alpha_2^j}$ is a minimizer of $\mathcal{J}_{\alpha_1^j,\alpha_2^j}$. Hence we can conclude that

$$\limsup_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j}) \leq \mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*}) \leq \mathcal{J}_{\alpha_1^*,\alpha_2^*}(v) \leq \liminf_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j}).$$

We thus see that v is a minimizer of $\mathcal{J}_{\alpha_1^*,\alpha_2^*}$ and by uniqueness we get that $v = u_{\alpha_1^*,\alpha_2^*}$.

We are left by showing that

$$\limsup_{j\to\infty} \alpha_1^j\|u_{\alpha_1^j,\alpha_2^j} - g_1\|_{L^1(\Omega)} + \alpha_2^j\|u_{\alpha_1^j,\alpha_2^j} - g_2\|^2_{L^2(\Omega)} \leq \alpha_1^*\|u_{\alpha_1^*,\alpha_2^*} - g_1\|_{L^1(\Omega)} + \alpha_2^*\|u_{\alpha_1^*,\alpha_2^*} - g_2\|^2_{L^2(\Omega)}.$$

Assume it is wrong. Then there exists an $\epsilon > 0$ and arbitrary j such that

$$\alpha_1^j\|u_{\alpha_1^j,\alpha_2^j} - g_1\|_{L^1(\Omega)} + \alpha_2^j\|u_{\alpha_1^j,\alpha_2^j} - g_2\|^2_{L^2(\Omega)} - \epsilon \geq \alpha_1^*\|u_{\alpha_1^*,\alpha_2^*} - g_1\|_{L^1(\Omega)} + \alpha_2^*\|u_{\alpha_1^*,\alpha_2^*} - g_2\|^2_{L^2(\Omega)}.$$

Then $\mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*}) \leq \liminf_{j\to\infty} \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j}) - \epsilon$ and $\mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^*,\alpha_2^*}) \to \mathcal{J}_{\alpha_1^*,\alpha_2^*}(u_{\alpha_1^*,\alpha_2^*})$ as $j \to \infty$. These last two statements lead as before to the contradiction that $\mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^*,\alpha_2^*}) \leq \mathcal{J}_{\alpha_1^j,\alpha_2^j}(u_{\alpha_1^j,\alpha_2^j})$. Hence we established continuity of the map for $\alpha_1, \alpha_2 > 0$. ∎

Now, let us deal with the behavior of the L^1-L^2-TV model if the parameters α_1 and α_2 are small.

PROPOSITION 3.12. *Let $T_1 = T_2 = I$ and $\alpha_1, \alpha_2 \geq 0$. There exists a threshold $\lambda^* = \lambda^*(\Omega)$ such that if $\alpha_1|\Omega|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Omega)} < \lambda^*$, then the minimizer u_{α_1,α_2} of $\mathcal{J}_{\alpha_1,\alpha_2}$ is constant.*

Proof. Let $u_\Omega := \frac{1}{|\Omega|}\int_\Omega u dx$, then there exists a constant $C > 0$ depending only on Ω such that

$$|Du|(\Omega) \geq C\|u - u_\Omega\|_{L^2(\Omega)} \quad \text{for all } u \in BV(\Omega);$$

see [4, Remark 3.50] or [35, p. 24]. From the optimality of $u := u_{\alpha_1,\alpha_2}$ we have $\mathcal{J}_{\alpha_1,\alpha_2}(u) \le \mathcal{J}_{\alpha_1,\alpha_2}(u_\Omega)$ and by the above inequality this yields

$$C\|u - u_\Omega\|_{L^2(\Omega)} + \alpha_1\|u - g_1\|_{L^1(\Omega)} + \alpha_2\|u - g_2\|^2_{L^2(\Omega)} \le \alpha_1\|u_\Omega - g_1\|_{L^1(\Omega)} + \alpha_2\|u_\Omega - g_2\|^2_{L^2(\Omega)}$$

which is equivalent to

$$C\|u - u_\Omega\|_{L^2(\Omega)} \le \alpha_1\|u_\Omega - g_1\|_{L^1(\Omega)} - \alpha_1\|u - g_1\|_{L^1(\Omega)} + \alpha_2\|u_\Omega - g_2\|^2_{L^2(\Omega)} - \alpha_2\|u - g_2\|^2_{L^2(\Omega)}.$$

By using the triangle inequality we get

$$C\|u - u_\Omega\|_{L^2(\Omega)} \le \alpha_1\|u_\Omega - u\|_{L^1(\Omega)} + \alpha_2\left(\|u_\Omega\|^2_{L^2(\Omega)} - 2\langle u_\Omega - u, g_2\rangle - \|u\|^2_{L^2(\Omega)}\right).$$

Note, that by the Cauchy-Schwarz inequality $\|u_\Omega\|^2_{L^2(\Omega)} \le \|u\|^2_{L^2(\Omega)}$ and hence we obtain

$$C\|u - u_\Omega\|_{L^2(\Omega)} \le \alpha_1\|u_\Omega - u\|_{L^1(\Omega)} + 2\alpha_2\|u_\Omega - u\|_{L^2(\Omega)}\|g_2\|_{L^2(\Omega)}.$$

By using Hölder's inequality on the L^1-term we get

$$C\|u - u_\Omega\|_{L^2(\Omega)} \le \left(\alpha_1|\Omega|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Omega)}\right)\|u - u_\Omega\|_{L^2(\Omega)}.$$

If $C > \alpha_1|\Omega|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Omega)}$, then $\|u - u_\Omega\|_{L^2(\Omega)} = 0$ and hence $u = u_\Omega$, which shows the assertion with $\lambda^* := C$. □

A similar result is obtained for images g_1, g_2 defined on $\mathbb{R}^d$, $d \in \mathbb{N}$, where the variational problem is written as

$$\min_{u\in BV(\mathbb{R}^d)}\{J_{\alpha_1,\alpha_2}(u) := \alpha_1\|u - g_1\|_{L^1(\mathbb{R}^d)} + \alpha_2\|u - g_2\|^2_{L^2(\mathbb{R}^d)} + |Du|(\mathbb{R}^d)\}. \tag{3.15}$$

PROPOSITION 3.13. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain, $g_i \in L^i(\mathbb{R}^d)$ given such that $\operatorname{supp}(g_i) \subset \Lambda$ for $i = 1, 2$, and $\alpha_1, \alpha_2 \ge 0$. Then there exists a threshold $\lambda^* = \lambda^*(\Lambda, d)$ such that if $\alpha_1|\Lambda|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Lambda)} < \lambda^*$, then a minimizer of J_{α_1,α_2} is given by $u_{\alpha_1,\alpha_2} \equiv 0$.*

Proof. By the Sobolev inequality, see e.g. [4, 35, 50], we have that there exists a constant $C(d) > 0$ such that

$$\int_{\mathbb{R}^d}|Du| \ge C(d)\|u\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} = C(d)\left(\|u\|^{\frac{d}{d-1}}_{L^{\frac{d}{d-1}}(\mathbb{R}^d\setminus\Lambda)} + \|u\|^{\frac{d}{d-1}}_{L^{\frac{d}{d-1}}(\Lambda)}\right)^{\frac{d-1}{d}}$$

for all $u \in BV(\mathbb{R}^d)$ with compact support. Then from the minimality of u_{α_1,α_2} we have $J_{\alpha_1,\alpha_2}(u_{\alpha_1,\alpha_2}) \le J_{\alpha_1,\alpha_2}(0)$ and hence by the isoperimetric inequality this means

$$\begin{aligned}C(d)\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} + \alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\mathbb{R}^d)} + \alpha_2\|u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\mathbb{R}^d)}&\\ \le \alpha_1\|g_1\|_{L^1(\mathbb{R}^d)} + \alpha_2\|g_2\|^2_{L^2(\mathbb{R}^d)}&.\end{aligned} \tag{3.16}$$

Since $\|\cdot\|^i_{L^i(\mathbb{R}^d)} = \|\cdot\|^i_{L^i(\mathbb{R}^d\setminus\Lambda)} + \|\cdot\|^i_{L^i(\Lambda)}$ and $\operatorname{supp}(g_i) \subset \Lambda$ for $i = 1, 2$ we also have

$$C(d)\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\Lambda)} + \alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Lambda)} + \alpha_2\|u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\Lambda)} \le \alpha_1\|g_1\|_{L^1(\Lambda)} + \alpha_2\|g_2\|^2_{L^2(\Lambda)},$$

which is equivalent to

$$\begin{aligned}C(d)\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\Lambda)} + \alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Lambda)} + \alpha_2\|u_{\alpha_1,\alpha_2}\|^2_{L^2(\Lambda)} + \alpha_2\|g_2\|^2_{L^2(\Lambda)}&\\ \le \alpha_1\|g_1\|_{L^1(\Lambda)} + \alpha_2\|g_2\|^2_{L^2(\Lambda)} + 2\alpha_2\langle u_{\alpha_1,\alpha_2}, g_2\rangle&.\end{aligned}$$

Now, we use the triangle inequality in the second term which yields

$$C(d)\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\Lambda)} \le \alpha_1\|u_{\alpha_1,\alpha_2}\|_{L^1(\Lambda)} + 2\alpha_2\langle u_{\alpha_1,\alpha_2}, g_2\rangle.$$

In the latter inequality we multiply the left side by $1 = \frac{1}{|\Lambda|^{\frac{2-d}{2d}}}\|1\|_{L^{\frac{2d}{2-d}}(\Lambda)}$ and use the generalized Hölder inequality, i.e., $\|uv\|_{L^r(\Lambda)} \le \|u\|_{L^p(\Lambda)}\|v\|_{L^q(\Lambda)}$ for $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}\le 1$ and $u \in L^p(\Lambda)$, $v \in L^p(\Lambda)$, to get

$$\frac{C(d)}{|\Lambda|^{\frac{2-d}{2d}}}\|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)} \le \alpha_1\|u_{\alpha_1,\alpha_2}\|_{L^1(\Lambda)} + 2\alpha_2\|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)}\|g_2\|_{L^2(\Lambda)},$$

where we used the Cauchy-Schwarz inequality on the right side. By using once more Hölder's inequality on the L^1-term we obtain

$$\frac{C(d)}{|\Lambda|^{\frac{2-d}{2d}}}\|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)} \le \left(\alpha_1|\Lambda|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Lambda)}\right)\|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)}.$$

If $\frac{C(d)}{|\Lambda|^{\frac{2-d}{2d}}} > \alpha_1|\Lambda|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Lambda)}$ then $\|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)} = 0$ and hence $u_{\alpha_1,\alpha_2} = 0$ in Λ.

We are left with showing that $u_{\alpha_1,\alpha_2} = 0$ in $\mathbb{R}^d \setminus \Lambda$ if $\|g_2\|_{L^2(\Lambda)} < \frac{1}{2\alpha_2}\left(\frac{C(d)}{|\Lambda|^{\frac{2-d}{2d}}} - \alpha_1|\Lambda|^{\frac{1}{2}}\right)$. By the inequality (3.16) we also have

$$C(d)\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} + \alpha_1\|u_{\alpha_1,\alpha_2} - g_1\|_{L^1(\Lambda)} + \alpha_2\|u_{\alpha_1,\alpha_2} - g_2\|^2_{L^2(\Lambda)} \le \alpha_1\|g_1\|_{L^1(\Lambda)} + \alpha_2\|g_2\|^2_{L^2(\Lambda)}.$$

Now we apply the triangle inequality and split the first term into integrations over $\mathbb{R}^d \setminus \Lambda$ and Λ, which gives

$$\begin{aligned} C(d)\left(\|u_{\alpha_1,\alpha_2}\|^{\frac{d}{d-1}}_{L^{\frac{d}{d-1}}(\mathbb{R}^d\setminus\Lambda)} + \|u_{\alpha_1,\alpha_2}\|^{\frac{d}{d-1}}_{L^{\frac{d}{d-1}}(\Lambda)}\right)^{\frac{d-1}{d}} + \alpha_1(\|g_1\|_{L^1(\Lambda)} - \|u_{\alpha_1,\alpha_2}\|_{L^1(\Lambda)}) \\ + \alpha_2(\|g_2\|_{L^2(\Lambda)} - \|u_{\alpha_1,\alpha_2}\|_{L^2(\Lambda)})^2 \le \alpha_1\|g_1\|_{L^1(\Lambda)} + \alpha_2\|g_2\|^2_{L^2(\Lambda)}. \end{aligned} \tag{3.17}$$

For $\alpha_1|\Lambda|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Lambda)} < \frac{C(d)}{|\Lambda|^{\frac{2-d}{2d}}} =: \lambda^*$ we have that $\|u_{\alpha_1,\alpha_2}\|_{L^p(\Lambda)} = 0$ for $p \in [1,\infty]$ and hence we obtain by (3.17) that $\|u_{\alpha_1,\alpha_2}\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d\setminus\Lambda)} = 0$, which concludes the proof. ☐

The assumptions $\alpha_1|\Omega|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Omega)} < \lambda^*$ and $\alpha_1|\Lambda|^{\frac{1}{2}} + 2\alpha_2\|g_2\|_{L^2(\Lambda)} < \lambda^*$ of the previous propositions clearly hold, if the parameters α_1 and α_2 are sufficiently small. These results somehow merge the behavior of the L^1-TV and L^2-TV model for small parameters, cf. [18, 50].

The last two statements dealt with the behavior of the L^1-L^2-TV model if α_1 and α_2 are small. Motivated by results for the L^1-TV model we state now properties of the L^1-L^2-TV model if α_1 is large. In particular, as for the L^1-TV model, see [18, Lemma 5.5], we have the following statement:

LEMMA 3.14. *Given $g_1 = g_2 =: g \in BV(\mathbb{R}^d)$, $d \in \mathbb{N}$. Assume there exists a vector field ϕ with the following properties:*

1. *$\phi(x) \in C^1_c(\mathbb{R}^d, \mathbb{R}^d)$,*
2. *$|\phi(x)| \le 1$ for all $x \in \mathbb{R}^d$,*
3. *$\int_{\mathbb{R}^d} g(x)\,\mathrm{div}\,\phi(x)dx = |Dg|(\mathbb{R}^d)$.*

Then there exists a threshold $\alpha_1^* \ge 0$ independent of α_2 such that the unique minimizer of J_{α_1,α_2} is given by $u_{\alpha_1,\alpha_2} = g$ for all $\alpha_1 \ge \alpha_1^*$ and $\alpha_2 \ge 0$.

Proof. For any $u \in BV(\mathbb{R}^d)$ we have

$$
\begin{aligned}
J_{\alpha_1,\alpha_2}(u) &= |Du|(\mathbb{R}^d) + \alpha_1 \int_{\mathbb{R}^d} |u-g|dx + \alpha_2 \int_{\mathbb{R}^d} |u-g|^2 dx \\
&\geq \int_{\mathbb{R}^d} u \operatorname{div} \phi dx + \alpha_1 \int_{\mathbb{R}^d} |u-g_1|dx + \alpha_2 \int_{\mathbb{R}^d} |u-g|^2 dx \\
&= \int_{\mathbb{R}^d} g \operatorname{div} \phi dx + \int_{\mathbb{R}^d} (u-g) \operatorname{div} \phi dx + \alpha_1 \int_{\mathbb{R}^d} |u-g|dx + \alpha_2 \int_{\mathbb{R}^d} |u-g|^2 dx \\
&= |Dg|(\mathbb{R}^d) + \int_{\mathbb{R}^d} (u-g) \operatorname{div} \phi dx + \alpha_1 \int_{\mathbb{R}^d} |u-g|dx + \alpha_2 \int_{\mathbb{R}^d} |u-g|^2 dx \\
&\geq J_{\alpha_1,\alpha_2}(g) + (\alpha_1 - \max_{x \in \mathbb{R}^d} |\operatorname{div} \phi|) \int_{\mathbb{R}^d} |u-g|dx + \alpha_2 \int_{\mathbb{R}^d} |u-g|^2 dx.
\end{aligned}
$$

For $\alpha_1 \geq \alpha_1^* := \max_{x \in \mathbb{R}^d} |\operatorname{div} \phi|$, the last inequality shows $J_{\alpha_1,\alpha_2}(u) \geq J_{\alpha_1,\alpha_2}(g)$. Assume u is a minimizer, which is unique, it follows that $u \equiv g$. ☐

When we apply Lemma 3.14 to binary images, we obtain the following theorem, cf. [18, Theorem 5.6].

THEOREM 3.15. *Let $\Lambda \subset \mathbb{R}^d$ be a bounded domain with C^2 boundary. Let $g(x) = g_1(x) = g_2(x) = 1_\Lambda(x)$ for all $x \in \mathbb{R}^d$ and $\alpha_2 \geq 0$. Then there exists a threshold $\alpha_1^* \geq 0$ such that whenever $\alpha_1 > \alpha_1^*$, the unique minimizer of J_{α_1,α_2} is $g = 1_\Lambda$ itself.*

4. Automated parameter selection. In order to motivate the L^1-L^2-TV model in [39] for the version (3.15) a simple and illustrative example is presented, in which its minimizer is compared with the one of the L^1-TV model, i.e., when $\alpha_2 = 0$ in (3.15), and with the one of the L^2-TV model, i.e., when $\alpha_1 = 0$ in (3.15). Note, that for $\alpha_2 > 0$ the functional in (3.15) is strictly convex and hence has a unique minimizer. Moreover, if $\alpha_1 = \alpha_2 = 0$, then any constant function is a minimizer of problem (3.15).

The amazing fact we observe from [39, Example 2.1] is that the L^1-L^2-TV model possesses the advantages of both other models, i.e., the L^1-TV model and L^2-TV model. That is, the L^1-L^2-TV model is able to recover the original image, has a unique solution $T_2 u_{\alpha_1,\alpha_2}$, since it is strictly convex with respect to $T_2 u$, and preserves even smaller details than the L^2-TV model.

We recall that for $g_1 = g_2 = 1_{B_r(0)}$ being the characteristic function of a disk $B_r(0)$ centered at the origin with radius $r > 0$ and $T_1 = T_2 = I$, the unique minimizer of (3.15) is given by

$$
u_{\alpha_1,\alpha_2} = \begin{cases} 0 & \text{if } 0 \leq r < \frac{2}{2\alpha_2+\alpha_1}, \\ \left(\frac{2\alpha_2+\alpha_1}{2\alpha_2} - \frac{1}{\alpha_2 r}\right) 1_{B_r(0)} & \text{if } \frac{2}{2\alpha_2+\alpha_1} \leq r \leq \frac{2}{\alpha_1}, \\ 1_{B_r(0)} & \text{if } r > \frac{2}{\alpha_1}. \end{cases} \tag{4.1}
$$

From this we clearly see that for the L^1-L^2-TV model there exist numerous different parameters α_1 and α_2 generating the same solution, even if $\frac{2}{2\alpha_2+\alpha_1} \leq r \leq \frac{2}{\alpha_1}$.

Our parameter selection approach is motivated by Theorem 3.5, from which we know that if $\alpha_i > 0$, then indeed $\|T_i u - g_i\|_{L^i(\Omega)}^i = \nu_i |\Omega|$ for this value of $i \in \{1,2\}$. In order to formulate an algorithm based on (3.1) we assume that the feasible set U is non-empty.

4.1. Uzawa's method. Assuming that ν_1 and ν_2 are at our disposal, we suggest to choose the parameters α_1 and α_2 depending on the constraints in (3.1). Hence the constrained minimization problem (3.1) might be solved by Uzawa's method [21]; see Algrotihm 1 below with $\nu_i(u^{(n)}) \equiv \nu_i$ constant. In general, as described in Section 2, ν_1 and ν_2 depend on the original (unknown) image. Nevertheless, instead of considering $\nu_i(u)$ in (3.1), which would result in a quite nonlinear problem, we choose a reference image and compute approximate values ν_1 and ν_2, leading to the following iterative scheme:

ALGORITHM 1 (Uzawa's method). *Initialize $\rho > 0$ (small enough), $\alpha_i^{(0)} > 0$ for $i = 1, 2$ and set $n = 0$;*

1) *Compute $u^{(n)} \in \arg\min_{u \in BV(\Omega)} \mathcal{J}_{\alpha_1^{(n)}, \alpha_2^{(n)}}(u)$*
2) *Update $\alpha_i^{(n+1)} = \max\{\alpha_i^{(n)} + \rho(H_i(u^{(n)}) - \nu_i(u^{(n)})|\Omega|), 0\}$ for $i = 1, 2$;*
3) *Stop or set $n = n + 1$ and continue with step 1).*

Here and below, $H_i(u) := \|T_i u - g_i\|_{L^i(\Omega)}^i$ for $i = 1, 2$. Observe, that if $H_i(u^{(n)}) < \nu_i(u^{(n)})|\Omega|$, then α_i is decreased, which relaxes the corresponding constraint, while for $H_i(u^{(n)}) > \nu_i(u^{(n)})|\Omega|$ the value α_i is increased and hence the associated constraint enforced. In step 2) it is ensured that the parameters α_i are always non-negative, whereby they are allowed to reach 0. The algorithm is stopped as soon as one of the following two conditions hold for the first time:

(S1) the distance between $H_i(u^{(n)})$ and $\nu_i(u^{(n)})|\Omega|$ is sufficiently small, i.e., $\frac{|H_i(u^{(n)}) - \nu_i(u^{(n)})|\Omega||}{\nu_i(u^{(n)})|\Omega|} < \varepsilon_1 = 10^{-4}$, or

(S2) the norm of the difference of two successive iterates $\alpha_i^{(n)}$ and $\alpha_i^{(n+1)}$ drops below a certain threshold, i.e., $\|\alpha_i^{(n)} - \alpha_i^{(n+1)}\| < \varepsilon_2 = 10^{-4}$.

In order to obtain convergence at all, the parameter $\rho > 0$ has to be chosen sufficiently small. If convergent, then clearly the magnitude of ρ has a significant influence on the convergence speed. In particular, a small ρ leads to a very slow convergence. Hence we would wish to choose ρ as large as possible but small enough such that the algorithm still converges. In all our numerical experiments we observed convergence of Algorithm 1 if we chose ρ at most 1. However, for this choice of ρ it turns out that the convergence speed is very slow, which makes this algorithm not really practical. Therefore, we present an alternative approach next.

4.2. The pAPS-algorithm. In [44] a fully automated parameter selection algorithm for the L^1-TV model, i.e., $\alpha_2 = 0$ in (1.3), and the L^2-TV model, i.e., $\alpha_1 = 0$ in (1.3), is proposed. We recall, that in contrast to Uzawa's method in the algorithm from [44] no additional parameter has to be chosen to find the regularization parameter α_1 or α_2 such that $u_{\alpha_1,0}$ solves

$$\min_{u \in BV(\Omega)} |Du|(\Omega) \quad \text{s.t.} \quad H_1(u) = \nu_1|\Omega|$$

and u_{0,α_2} is a minimizer of

$$\min_{u \in BV(\Omega)} |Du|(\Omega) \quad \text{s.t.} \quad H_2(u) = \nu_2|\Omega|.$$

The automated adjustment of the regularization parameter (α_1 or α_2) is performed iteratively depending on the constraint $H_1(u) = \nu_1|\Omega|$ in the case of the L^1-TV model or on the constraint $H_2(u) = \nu_2|\Omega|$ in the case of the L^2-TV model. For example, for the L^1-TV model the parameter α_1 is increased whenever $\frac{H_1(u_{\alpha_1,0})}{\nu_1|\Omega|} > 1$ and decreased if $\frac{H_1(u_{\alpha_1,0})}{\nu_1|\Omega|} < 1$. This leads to the following update scheme:

$$\alpha_1^{(n+1)} = \left(\frac{H_1(u_{\alpha_1^{(n)},0})}{\nu_1|\Omega|} \right)^p \alpha_1^{(n)},$$

where $p \geq 0$ such that $(H_1(u_{\alpha_1^{(n)},0}))_n$ is monotonically decreasing, if $H_1(u_{\alpha_1^{(0)},0}) > \nu_1|\Omega|$, and $(H_1(u_{\alpha_1^{(n)},0}))_n$ is monotonically increasing, if $H_1(u_{\alpha_1^{(0)},0}) \leq \nu_1|\Omega|$.

Motivated by this strategy, we suggest the following automated parameter selection algorithm for the L^1-L^2-TV model.

> ALGORITHM 2 (pAPS-Algorithm). *Initialize $p > 0$, $\alpha_i^{(0)} > 0$ for $i = 1, 2$ and set $n = 0$;*
> 1) *Compute $u^{(n)} \in \arg\min_{u \in BV(\Omega)} \mathcal{J}_{\alpha_1^{(n)}, \alpha_2^{(n)}}(u)$*
> 2) *Update $\alpha_i^{(n+1)} = \left(\frac{H_i(u^{(n)})}{\nu_i(u^{(n)})|\Omega|} \right)^p \alpha_i^{(n)}$ for $i = 1, 2$;*
> 3) *Solve $u^{(n+1)} \in \arg\min_{u \in BV(\Omega)} \mathcal{J}_{\alpha_1^{(n+1)}, \alpha_2^{(n+1)}}(u)$*
> 4) *For $i = 1, 2$ do*
>   (a) *if $H_i(u^{(0)}) \le \nu_i(u^{(0)})|\Omega|$*
>   (i) *if $H_i(u^{(n+1)}) > \nu_i(u^{(n+1)})|\Omega|$, decrease p, e.g., $p = p/2$, and go to step 2);*
>   (ii) *if $H_i(u^{(n+1)}) \le \nu_i(u^{(n+1)})|\Omega|$, continue;*
>   (b) *if $H_i(u^{(0)}) > \nu_i(u^{(0)})|\Omega|$*
>   (i) *if $H_i(u^{(n+1)}) < \nu_i(u^{(n+1)})|\Omega|$, decrease p, e.g., $p = p/2$, and go to step 2);*
>   (ii) *if $H_i(u^{(n+1)}) \ge \nu_i(u^{(n+1)})|\Omega|$, continue;*
> 5) *Stop or set $n := n + 1$ and return to step 2);*

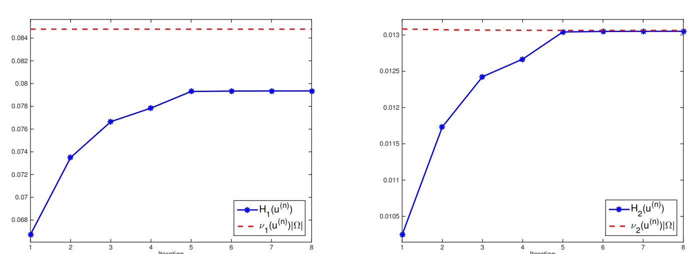
As a stopping criterion we use that either (S1), (S2), or
(S3) the power p is significant small, i.e., $p < \varepsilon_3 = 10^{-3}$;
holds for the first time.

Due to the adaptive choice of p in the pAPS-algorithm, we observe that the generated sequences $(H_i(u^{(n)}))_n$ and $(\alpha_i^{(n)})_n$ are monotonically decreasing or increasing, depending on the initial $\alpha_i^{(0)}$, for $i = 1, 2$, while for Algorithm 1 these monotonic behaviors are in general not guaranteed; see Figure 4.1, Figure 4.2 and Figure 4.3. In particular, we have the following result for the pAPS-algorithm.

LEMMA 4.1. *The pAPS-algorithm generates monotone sequences $(\alpha_i^{(n)})_n$, for $i = 1, 2$. In particular, we have*
(i) if $\alpha_i^{(0)}$ is such that $H_i(u^{(0)}) > \nu_i(u^{(0)})|\Omega|$, then $(\alpha_i^{(n)})_n$ is monotonically increasing, i.e., $\alpha_i^{(n)} \le \alpha_i^{(n+1)}$ for all $n \in \mathbb{N}$;
(ii) if $\alpha_i^{(0)}$ is such that $H_i(u^{(0)}) \le \nu_i(u^{(0)})|\Omega|$, then $(\alpha_i^{(n)})_n$ is monotonically decreasing, i.e., $\alpha_i^{(n)} \ge \alpha_i^{(n+1)}$ for all $n \in \mathbb{N}$.

Proof. For $H_i(u^{(0)}) > \nu_i(u^{(0)})|\Omega|$ we can show by induction that $\alpha_i^{(n+1)} > \alpha_i^{(n)}$ for all n and $i = 1, 2$. In particular, $H_i(u^{(n)}) > \nu_i(u^{(n)})|\Omega|$ implies $\alpha_i^{(n+1)} = \left(\frac{H_i(u^{(n)})}{\nu_i(u^{(n)})|\Omega|} \right)^p \alpha_i^{(n)} > \alpha_i^{(n)}$, where p is due to the pAPS-algorithm such that $H_i(u^{(n+1)}) \ge \nu_i(u^{(n+1)})|\Omega|$.

By similar arguments we obtain for $\alpha_i^{(0)}$ with $H_i(u^{(0)}) \le \nu_i(u^{(0)})|\Omega|$, that $(\alpha_i^{(n)})_n$ is monotonically decreasing. ∎



(a) (b)

FIG. 4.1. *Progress of $H_1(u^{(n)})$ and $H_2(u^{(n)})$ of the pAPS-algorithm with $\alpha_1^{(0)} = 1 = \alpha_2^{(0)}$ for restoring the image "cameraman" (see Figure 6.1(a)) corrupted by Gaussian white noise with $\sigma = 0.1$ and salt-and-pepper noise with $s_1 = s_2 = 0.005$.*

Due to the monotonicity property of the sequence $(\alpha_i^{(n)})_n$ for $i = 1, 2$ we have the following convergence property of the pAPS-algorithm.

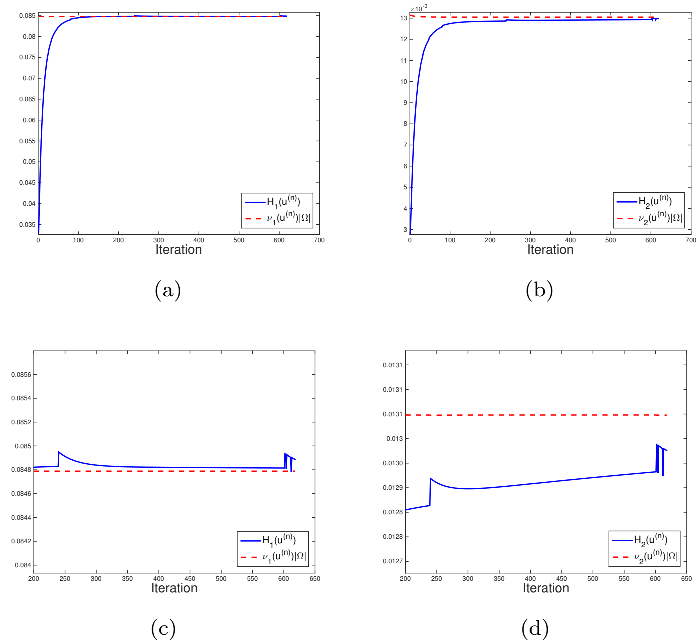


FIG. 4.2. *Progress of $H_1(u^{(n)})$ and $H_2(u^{(n)})$ of the Algorithm 1 with $\alpha_1^{(0)} = 1 = \alpha_2^{(0)}$ for restoring the image "cameraman" (see Figure 6.1(a)) corrupted by Gaussian white noise with $\sigma = 0.1$ and salt-and-pepper noise with $s_1 = s_2 = 0.005$. In (c) and (d) we zoomed in on the last few hundred iterations.*

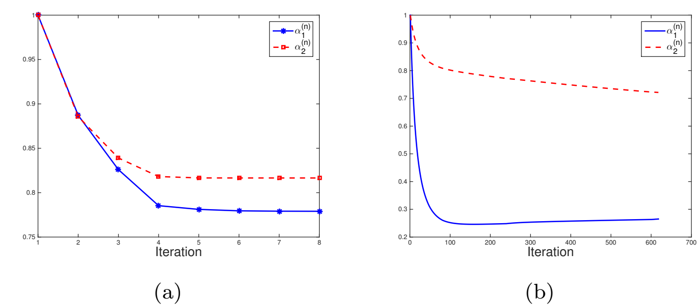


FIG. 4.3. *Progress of $\alpha_1^{(n)}$ and $\alpha_2^{(n)}$ of the pAPS-algorithm in (a) and Algorithm 1 in (b) with $\alpha_1^{(0)} = 1 = \alpha_2^{(0)}$ for restoring the image "cameraman" (see Figure 6.1(a)) corrupted by Gaussian white noise with $\sigma = 0.1$ and salt-and-pepper noise with $s_1 = s_2 = 0.005$.*

THEOREM 4.2. *For $i \in \{1, 2\}$ the pAPS-algorithm generates a convergent sequence $(\alpha_i^{(n)})_n$, i.e., $\lim_{n\to\infty} \alpha_i^{(n)} = \bar{\alpha}_i \in \mathbb{R}$, if one of the following conditions holds:*

(i) $\alpha_i^{(0)} > 0$ such that $H_i(u^{(0)}) \le \nu_i(u^{(0)})|\Omega|$;

(ii) there exist $\bar{\alpha}_1, \bar{\alpha}_2 > 0$ such that $H_i(u^{(n)}) > \nu_i(u^{(n)})|\Omega|$ for all $\alpha_i^{(n)} < \bar{\alpha}_i$ and $H_i(u_{\alpha_1,\alpha_2}) \le \nu_i(u_{\alpha_1,\alpha_2})|\Omega|$ for all $\alpha_1 \ge \bar{\alpha}_1$ and $\alpha_2 \ge \bar{\alpha}_2$, where u_{α_1,α_2} is a solution of (1.3).

Proof.

(i) Let $\alpha_i^{(0)} > 0$ such that $H_i(u^{(0)}) \le \nu_i(u^{(0)})|\Omega|$. Then by Lemma 4.1(ii) we have that $(\alpha_i^{(n)})_n$ is monotonically decreasing, i.e., $0 \le \alpha_i^{(n+1)} \le \alpha_i^{(n)} \le \alpha_i^{(0)}$ for all $n \in \mathbb{N}$, and hence it is bounded. The convergence follows by the monotone convergence theorem for sequences.

(ii) If there exist $\bar{\alpha}_1, \bar{\alpha}_2 > 0$ such that $H_i(u^{(n)}) > \nu_i(u^{(n)})|\Omega|$ for all $\alpha_i^{(n)} < \bar{\alpha}_i$ and $H_i(u_{\alpha_1,\alpha_2}) \le \nu_i(u_{\alpha_1,\alpha_2})|\Omega|$ for all $\alpha_1 \ge \bar{\alpha}_1$ and $\alpha_2 \ge \bar{\alpha}_2$, then $(\alpha_i^{(n)})_n$ is monotonically increasing, cf.

Lemma 4.1(i), and we deduce that $0 \leq \alpha_i^{(n)} \leq \alpha_i^{(n+1)} < \bar{\alpha}_i$ for all $n \in \mathbb{N}$. Hence $(\alpha_i^{(n)})_n$ is bounded and consequently convergent, which concludes the proof.

□

Note, that $\alpha_i^{(0)}$ has to be chosen positive for $i = 1, 2$, since if $\alpha_i^{(0)} = 0$ in the pAPS-algorithm, then $\alpha_i^{(n)} = 0$ for all $n \geq 0$, and we cannot expect a reasonable result in general.

5. A solution algorithm. For computing a minimizer of the problem (1.3) the authors suggested in [39] (without any convergence analysis) an algorithm, which is an adaptation of a method that was originally proposed for L^1-TV minimization problems in [6], based on replacing the functional $\mathcal{J}_{\alpha_1,\alpha_2}$ by

$$F(u, v) := \alpha_1 \|v\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1 u - g_1 - v\|^2_{L^2(\Omega)} + \alpha_2 \|T_2 u - g_2\|^2_{L^2(\Omega)} + |Du|(\Omega), \tag{5.1}$$

where $\gamma > 0$ is small, so that we have $g_1 \approx T_1 u - v$. Actually for $\gamma \to 0$ (5.1) approaches the objective functional in (1.3). Then (5.1) is minimized alternating with respect to u and v which results in the following algorithm:

ALGORITHM 3. *Initialize $u^{(0)} \in L^2(\Omega)$. For $n = 0, 1, \ldots$ do*

$$v^{(n+1)} = \arg \min_{v \in L^2(\Omega)} \alpha_1 \|v\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1 u^{(n)} - g_1 - v\|^2_{L^2(\Omega)} \tag{5.2}$$

$$u^{(n+1)} \in \arg \min_{u \in L^2(\Omega)} \frac{1}{2\gamma}\|T_1 u - g_1 - v^{(n+1)}\|^2_{L^2(\Omega)} + \alpha_2 \|T_2 u - g_2\|^2_{L^2(\Omega)} + |Du|(\Omega) \tag{5.3}$$

Now we are going to analyse the convergence of this algorithm.

THEOREM 5.1. *Weak accumulation points of the sequence $(u^{(n)}, v^{(n)})_n$ generated by Algorithm 3 are minimizers of F in $L^2(\Omega) \times L^2(\Omega)$ and $BV(\Omega) \times L^2(\Omega)$.*

The proof of this statement uses the same ideas as the ones of Proposition 5 in [6]. However, in contrary to [6] where the proof is done in a finite dimensional setting with the assumption of a continuous objective functional, we are working in an infinite dimensional space and our functional F is only lower semicontinuous, which requires additional arguments. Because of these reasons we state the complete proof here.

Proof. By Algorithm 3 we have

$$F(u^{(n)}, v^{(n)}) \geq F(u^{(n)}, v^{(n+1)}) \geq F(u^{(n+1)}, v^{(n+1)}). \tag{5.4}$$

Since F is bounded by 0 it follows that $(F(u^{(n)}, v^{(n)}))_n$ is convergent. Note that F is coercive in $L^2(\Omega) \times L^2(\Omega)$. From this and the convergence of $(F(u^{(n)}, v^{(n)}))_n$ we deduce that $(u^{(n)}, v^{(n)})_n$ is bounded in $L^2(\Omega) \times L^2(\Omega)$ and hence we can extract a weakly convergent subsequence. Moreover, due to the presence of the total variation $|Du|(\Omega)$ in F and $\alpha_1 + \alpha_2 > 0$ we obtain that $(u^{(n)}, v^{(n)})_n$ is bounded in $BV(\Omega) \times L^2(\Omega)$. The compact embedding $BV(\Omega) \hookrightarrow L^q(\Omega)$, $1 \leq q < \frac{d}{d-1}$ ($d = 2$ is the dimension of Ω), implies that a subsequence $(u^{(n_k)}, v^{(n_k)})_k$ converges in $L^q(\Omega) \times L^2(\Omega)$ to a limit $(u^*, v^*) \in L^2(\Omega) \times L^2(\Omega)$. By [5, Prop. 10.1.1] we even have that $(u^*, v^*) \in BV(\Omega) \times L^2(\Omega)$, $\liminf_{n_k \to \infty} |Du^{(n_k)}|(\Omega) \geq |Du^*|(\Omega)$, and $(u^{(n_k)}, v^{(n_k)})_k$ weakly converges to (u^*, v^*) in $BV(\Omega) \times L^2(\Omega)$ as $n_k \to +\infty$. Further, we have, for all $n_k \in \mathbb{N}$

$$F(u^{(n_k)}, v^{(n_k+1)}) \leq F(u^{(n_k)}, v)$$

for all $v \in L^2(\Omega)$ and

$$F(u^{(n_k)}, v^{(n_k)}) \leq F(u, v^{(n_k)}) \tag{5.5}$$

for all $u \in L^2(\Omega)$. Note that $(v^{(n_k+1)})_k$ is again bounded and let us denote by $\tilde{v}$ a corresponding cluster point.

Considering (5.4) we have that

$$F(u^{(n_k)}, v^{(n_k)}) - F(u^{(n_{k+1})}, v^{(n_{k+1})}) \geq F(u^{(n_k)}, v^{(n_k+1)}) - F(u^{(n_{k+1})}, v^{(n_{k+1})}).$$

Since F is bounded from below, we obtain $\lim_{n_k\to\infty}\left[F(u^{(n_k)}, v^{(n_k)}) - F(u^{(n_{k+1})}, v^{(n_{k+1})})\right] = 0$ and consequently

$$0 = \lim_{n_k\to\infty}\left[F(u^{(n_k)}, v^{(n_k+1)}) - F(u^{(n_{k+1})}, v^{(n_{k+1})})\right] = F(u^*, \tilde{v}) - F(u^*, v^*). \tag{5.6}$$

By passing (5.2) to the limit we get that $\tilde{v}$ is a solution of $\min_{v\in L^2(\Omega)} \alpha_1\|v\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1u^* - g_1 - v\|^2_{L^2(\Omega)}$. From (5.6) we know that $F(u^*, \tilde{v}) = F(u^*, v^*)$ and hence

$$\alpha_1\|\tilde{v}\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1u^* - g_1 - \tilde{v}\|^2_{L^2(\Omega)} = \alpha_1\|v^*\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1u^* - g_1 - v^*\|^2_{L^2(\Omega)}$$

By the uniqueness of the solution ($F(u^*, \cdot)$ is strictly convex) we conclude that $\tilde{v} = v^*$. Hence $v^{(n_k+1)} \to v^*$ for $n_k \to \infty$.

Moreover, $v^* = \arg\min_{v\in L^2(\Omega)} F(u^*, v)$, i.e.

$$F(u^*, v^*) \leq F(u^*, v) \quad \text{for all } v \in L^2(\Omega). \tag{5.7}$$

And by passing (5.5) to the limit we obtain

$$F(u^*, v^*) \leq \left(\liminf F(u^{(n_k)}, v^{(n_k)}) \leq \liminf F(u, v^{(n_k)}) =\right) F(u, v^*) \quad \text{for all } u \in L^2(\Omega). \tag{5.8}$$

From the definition of F the inequality in (5.7) is equivalent to

$$0 \in \frac{1}{\gamma}(v^* - T_1u^* + g_1) + \alpha_1\partial\|v^*\|_{L^1(\Omega)} \tag{5.9}$$

and (5.8) is equivalent to

$$0 \in \frac{1}{\gamma}T_1^*(T_1u^* - g_1 - v^*) + 2\alpha_2T_2^*(T_2u^* - g_2) + \partial|Du^*|(\Omega). \tag{5.10}$$

The subdifferential of F at (u^*, v^*) is given by

$$\partial F(u^*, v^*) = \begin{pmatrix} \frac{1}{\gamma}(v^* - T_1u^* + g_1) + \alpha_1\partial\|v^*\|_{L^1(\Omega)} \\ \frac{1}{\gamma}T_1^*(T_1u^* - g_1 - v^*) + 2\alpha_2T_2^*(T_2u^* - g_2) + \partial|Du^*|(\Omega) \end{pmatrix}.$$

According to (5.9) and (5.10) we have

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \partial F(u^*, v^*)$$

which is equivalent to $F(u^*, v^*) = \min_{(u,v)\in L^2(\Omega)\times L^2(\Omega)} F(u, v)$.
□

The minimizer $v^{(n+1)}$ of (5.2) can be easily computed via a soft thresholding, i.e., $v^{(n+1)} = \mathrm{ST}(T_1u^{(n)} - g_1, \gamma\alpha_1)$, where

$$\mathrm{ST}(g, \beta)(x) = \begin{cases} g(x) - \beta & \text{if } g(x) > \beta, \\ 0 & \text{if } |g(x)| \leq \beta, \\ g(x) + \beta & \text{if } g(x) < -\beta \end{cases}$$

for all $x \in \Omega$.

The solution of the minimization problem in (5.3) can be realized by replacing F by a family of *surrogate functionals*

$$S(u,a,v) := F(u,v) + \frac{1}{2\gamma}\left(\|u-a\|^2_{L^2(\Omega)} - \|T_1(u-a)\|^2_{L^2(\Omega)}\right) + \alpha_2\left(\|u-a\|^2_{L^2(\Omega)} - \|T_2(u-a)\|^2_{L^2(\Omega)}\right)$$

with $a,u,v \in L^2(\Omega)$. Here we assume that $\|T_i\| < 1$ for $i = 1,2$, which is not at all a restriction, since if a norm exceeds 1, then a proper rescaling of the problem reestablishes the desired setting. Note that

$$\min_{u\in L^2(\Omega)} S(u,a,v) \Leftrightarrow \min_{u\in L^2(\Omega)} \left\| u - \frac{\gamma}{1+2\alpha_2\gamma}\left(\frac{1}{\gamma}z_1 + 2\alpha_2 z_2\right)\right\|^2_{L^2(\Omega)} + \frac{2\gamma}{1+2\alpha_2\gamma}|Du|(\Omega). \tag{5.11}$$

where $z_1 = z_1(a) = a + T_1^*(g_1 + v - T_1 a)$ and $z_2 = z_2(a) = a + T_2^*(g_2 - T_2 a)$; cf. [39]. There exist several numerical methods for solving (5.11) efficiently; see for example [12, 19, 22, 25, 28, 36, 38, 41, 52, 55]. This leads to the following algorithm:

Algorithm 4. *Initialize:* $u^{(0,L)} \in L^2(\Omega)$. *For* $n = 0,1,\ldots$ *do*

$$\begin{aligned} v^{(n+1)} &= \arg\min_{v\in L^2(\Omega)} \alpha_1\|v\|_{L^1(\Omega)} + \frac{1}{2\gamma}\|T_1 u^{(n,L)} - g_1 - v\|^2_{L^2(\Omega)} \\ u^{(n+1,0)} &= u^{(n,L)} \\ u^{(n+1,\ell+1)} &= \arg\min_{u\in L^2(\Omega)} S(u, u^{(n+1,\ell)}, v^{(n+1)}), \quad \ell = 0,\ldots,L-1. \end{aligned} \tag{5.12}$$

Note that we do prescribe a finite number $L \in \mathbb{N}$ of inner iterations.

Theorem 5.2. *Let us assume* $\|T_i\| < 1$ *for* $i = 1,2$. *Then weak accumulation points of the sequence* $(u^{(n,L)}, v^{(n)})_n$ *generated by Algorithm 4 are minimizers of* F *in* $L^2(\Omega) \times L^2(\Omega)$ *and* $BV(\Omega) \times L^2(\Omega)$.

Proof. By Algorithm 4 we have

$$\begin{aligned} F(u^{(n,L)}, v^{(n)}) &\ge F(u^{(n,L)}, v^{(n+1)}) = S(u^{(n+1,0)}, u^{(n+1,0)}, v^{(n+1)}) \ge S(u^{(n+1,1)}, u^{(n+1,0)}, v^{(n+1)}) \\ &\ge S(u^{(n+1,1)}, u^{(n+1,1)}, v^{(n+1)}) \ge \ldots \ge S(u^{(n+1,L)}, u^{(n+1,L)}, v^{(n+1)}) = F(u^{(n+1,L)}, v^{(n+1)}). \end{aligned}$$

By the same arguments as in Theorem 5.1 we obtain

$$F(u^*, v^*) \le F(u^*, v) \quad \text{for all } v \in L^2(\Omega),$$

where $(u^*, v^*) \in BV(\Omega) \times L^2(\Omega)$ is a limit of the subsequence $(u^{(n_k,L)}, v^{(n_k)})_k$.

Next we want to show that $0 \in \partial F(u^*, v^*)$. Therefore we analyse the surrogate iteration (5.12) in more details. By the monotonic decrease of F and S we have

$$\begin{aligned} F(u^{(n,L)}, v^{(n)}) - F(u^{(n+1,1)}, v^{(n+1)}) &\ge F(u^{(n+1,0)}, v^{(n+1)}) - F(u^{(n+1,1)}, v^{(n+1)}) \\ &\ge S(u^{(n+1,1)}, u^{(n+1,0)}, v^{(n+1)}) - S(u^{(n+1,1)}, u^{(n+1,1)}, v^{(n+1)}) \\ &= \frac{1}{2\gamma}\left(\|u^{(n+1,1)} - u^{(n+1,0)}\|^2_{L^2(\Omega)} - \|T_1(u^{(n+1,1)} - u^{(n+1,0)})\|^2_{L^2(\Omega)}\right) \\ &\quad + \alpha_2\left(\|u^{(n+1,1)} - u^{(n+1,0)}\|^2_{L^2(\Omega)} - \|T_2(u^{(n+1,1)} - u^{(n+1,0)})\|^2_{L^2(\Omega)}\right) \\ &\ge \left(\frac{1}{2\gamma}C_1 + \alpha_2 C_2\right)\|u^{(n+1,1)} - u^{(n+1,0)}\|^2_{L^2(\Omega)}, \end{aligned}$$

where $C_i := (1 - \|T_i\|^2) > 0$. Moreover, we get

$$F(u^{(n+1,\ell)}, v^{(n+1)}) - F(u^{(n+1,\ell+1)}, v^{(n+1)}) \ge \left(\frac{1}{2\gamma}C_1 + \alpha_2 C_2\right)\|u^{(n+1,\ell+1)} - u^{(n+1,\ell)}\|^2_{L^2(\Omega)}.$$

Hence, after L steps we conclude

$$F(u^{(n,L)}, v^{(n)}) - F(u^{(n+1,L)}, v^{(n+1)}) \geq \left(\frac{1}{2\gamma}C_1 + \alpha_2 C_2\right) \sum_{\ell=0}^{L-1} \|u^{(n+1,\ell+1)} - u^{(n+1,\ell)}\|^2_{L^2(\Omega)}. \quad (5.13)$$

Since the sequence $(F(u^{(n,L)}, v^{(n)}))_n$ is convergent we deduce from (5.13) that

$$\sum_{\ell=0}^{L-1} \|u^{(n+1,\ell+1)} - u^{(n+1,\ell)}\|^2_{L^2(\Omega)} \to 0, \quad n \to \infty,$$

and hence

$$\lim_{n\to\infty} \|u^{(n+1,\ell+1)} - u^{(n+1,\ell)}\|^2_{L^2(\Omega)} = 0$$

for all $\ell \in \{0, \ldots, L-1\}$. Consequently, the sequences $(u^{(n_k,L)})$ and $(u^{(n_k,L-1)})$ have the same limit u^*.

By the optimality of $u^{(n_k,L)}$ we have

$$\begin{aligned}
0 &\in \partial S(\cdot, u^{(n_k,L-1)}, v^{(n_k)})(u^{(n_k,L)}) \\
&= \partial F(\cdot, v^{(n_k)})(u^{(n_k,L)}) + \frac{1}{\gamma}\left((u^{(n_k,L)} - u^{(n_k,L-1)}) - T_1^* T_1(u^{(n_k,L)} - u^{(n_k,L-1)})\right) \\
&\qquad + 2\alpha_2 \left((u^{(n_k,L)} - u^{(n_k,L-1)}) - T_2^* T_2(u^{(n_k,L)} - u^{(n_k,L-1)})\right)
\end{aligned}$$

Then, by letting $n_k \to \infty$ we obtain

$$0 \in \partial S(\cdot, u^*, v^*)(u^*) = \partial F(\cdot, v^*)(u^*).$$

The rest of the proof is analogous to the proof of Theorem 5.1. □

REMARK 5.3 (Denoising). *If $T_1 = T_2 = I$, then we do not need surrogate functionals and use Algorithm 3 directly, since the minimization problem in* (5.3) *is equivalent to*

$$\arg \min_{u \in L^2(\Omega)} \left\| u - \frac{\gamma}{1 + 2\alpha_2\gamma}\left(\frac{1}{\gamma}(g_1 + v) + 2\alpha_2 g_2\right)\right\|^2_{L^2(\Omega)} + \frac{2\gamma}{1 + 2\alpha_2\gamma}|Du|(\Omega)$$

and can be solved as (5.11) *by one of the methods mentioned above.*

6. Numerical Experiments. In this section we present several numerical experiments on image denoising and image deblurring to show the behavior of the proposed algorithm and their restoration potential. As a comparison for the different restoration qualities of the image we use the MSSIM [61] (mean structural similarity). In general, when comparing MSSIM, large values indicate better reconstruction than small values.

The minimization problem in the pAPS-algorithm as well as in Algorithm 1 is solved approximately by Algorithm 4, where $\gamma = 10^{-2}$. Moreover, the initial power p in the pAPS-algorithm is chosen to be 1 in all our experiments.

For our numerical studies we consider the images shown in Figure 6.1 of size 256×256 pixels. We recall, that the image intensity range of all examples considered in this paper is $[0, 1]$.

Since for a mixture of noise the expected absolute value is here not available and difficult to compute, in our numerics we set $\nu_1 = \mathrm{EAV}(\eta_{\hat{u}}) + \mathrm{EAV}(\rho_{\hat{u}})$, which is actually only an above approximation of the real expected absolute value. However, note that $\mathrm{EAV}(\eta_{\hat{u}}) + \mathrm{EAV}(\rho_{\hat{u}})$ is also an element of $\left[0, \sqrt{\frac{2}{\pi}}\sigma + \max\{s_1, s_2\}\right]$ or $\left[0, \sqrt{\frac{2}{\pi}}\sigma + \frac{s}{2}\right]$, respectively. Moreover, we recall, that ν_1 as well as ν_2 are computed based on some approximation of the true image.

(a) "cameraman" (b) "barbara"

Fig. 6.1. *Original images of size* 256×256 *pixels.*

$\alpha_1^{(0)}$	$\alpha_2^{(0)}$	MSSIM	α_1	α_2
1	1	0.7163	0.7790	0.8166
1	0.5	0.7051	0.8541	0.4562
1	0.1	0.6998	0.9230	0.0971
0.5	1	0.7498	0.5241	1.2271
0.5	0.5	0.7478	0.5943	0.8918
0.5	0.1	0.7405	0.7175	0.3330
0.1	1	0.7470	0.1644	3.0531
0.1	0.5	0.7483	0.2150	2.7935
0.1	0.1	0.7509	0.3710	1.9926

Table 6.1

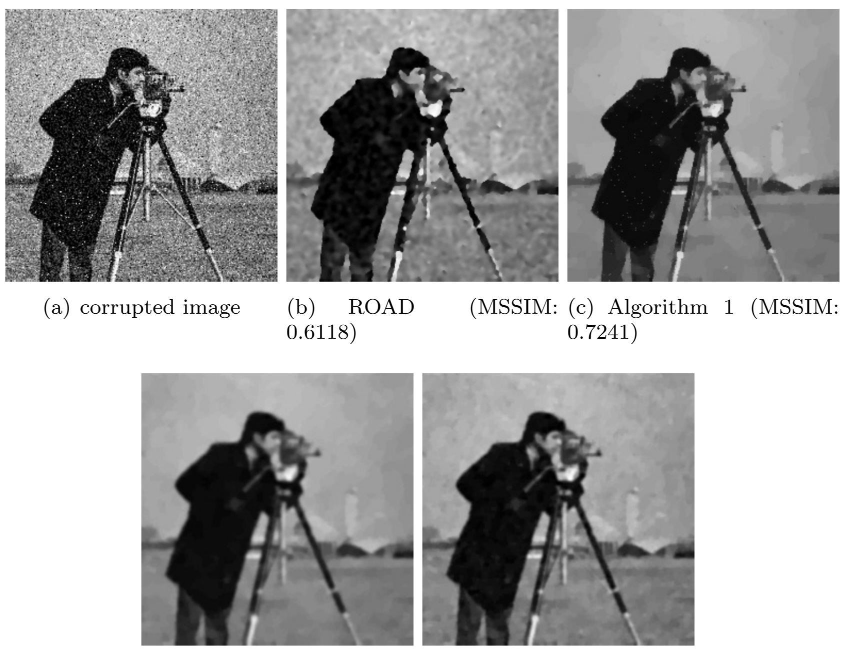
MSSIM results for the 256×256 *pixel image "cameraman" corrupted by Gaussian white noise with* $\sigma = 0.1$ *and salt-and-pepper noise with* $s_1 = s_2 = 0.005$ *obtained by the pAPS-algorithm.*

6.1. Initial value $\alpha_i^{(0)}$. We start by investigating the pAPS-algorithm concerning its stability with respect to the initial $\alpha_i^{(0)}$, $i = 1, 2$. For this purpose we consider the 256×256 pixel image "cameraman" corrupted by Gaussian white noise with $\sigma = 0.1$ and salt-and-pepper noise with $s_1 = s_2 = 0.005$ and test for $\alpha_1^{(0)}, \alpha_2^{(0)} \in \{0.1, 0.5, 1\}$. Our findings are summarized in Table 6.1, i.e., the obtained parameters α_1 and α_2 and the MSSIM of the corresponding received reconstructions. The obtained parameters α_1 and α_2 are always relatively close to the initial $\alpha_1^{(0)}$ and $\alpha_2^{(0)}$. Note, that even if problem (3.1) may have a unique minimizer there may exist pairs (α_1^1, α_2^1) and (α_1^2, α_2^2) with $(\alpha_1^1, \alpha_2^1) \neq (\alpha_1^2, \alpha_2^2)$ such that $u_{\alpha_1^1, \alpha_2^1} = u_{\alpha_1^2, \alpha_2^2}$, which can be, for example, easily seen from (4.1) and [39, Example 2.1]. We actually observe, that although the α_1's and α_2's differ significantly from each other, the MSSIM seems similar throughout the experiments.

In order to keep the number of iterations in the pAPS-algorithm small a good choice of the initial values is still desirable. Therefore in the sequel we choose $\alpha_i^{(0)}$, $i = 1, 2$, according to [48], i.e., we set $\alpha_1^{(0)}$ and $\alpha_2^{(0)}$ as in (1.4). In this case, for the example in Table 6.1 we actually get $\alpha_1^{(0)} = 0.5 = \alpha_2^{(0)}$. By incorporating (1.4) for the choice of the initial parameters in the pAPS-algorithm makes this method fully automatic for the user.

6.2. Gaussian plus impulse noise. For the simultaneous removal of Gaussian and impulse noise we compare the performance of Algorithm 1 and the pAPS-algorithm with the frequently used ROAD-trilateral filter [33], which is designed to remove a mixture of Gaussian noise (with zero mean and variance σ^2) and impulse noise. This filter is based on a simple statistic to detect outliers in an image. Moreover, we also report on the results obtained by the L^1-L^2-TV model with α_1 and α_2 chosen as suggest in [48], i.e., as in (1.4). In this case a minimizer is approximately computed by Algorithm 3 and in the case of deblurring by Algorithm 4. In the sequel we refer to them as the L^1-L^2-TV algorithm. For our comparison we restore the "cameraman" image (see Figure 6.1(a)) and the "barbara" image (see Figure 6.1(b)) for mixed Gaussian-impulse noise with different noise levels, i.e., $\sigma \in \{0.01, 0.1, \sqrt{0.02}\}$, $s_1 = s_2 \in \{0.005, 0.01, 0.05, 0.15\}$, and $s \in \{0.005, 0.01, 0.05, 0.15\}$.

For simultaneously removing Gaussian and salt-and-pepper noise in the "cameraman" image we summarize our findings in Table 6.2. There it is demonstrated that the L^1-L^2-TV algorithm with parameters chosen as in [48], i.e., as in equation (1.4), produces competitive results, which are actually always better than the ones generated by the ROAD-trilateral filter. Setting the initial parameters to (1.4) the pAPS-algorithm finds automatically new parameters (α_1, α_2) which improve the restoration quality of the L^1-L^2-TV algorithm. In particular, in Figure 6.2 we see that the numerical solution produced by the L^1-L^2-TV algorithm with the parameters as in [48] is over-smoothed, while the result generated by the pAPS-algorithm shows more details and has sharper edges. For this particular example in Figure 6.2 Algorithm 1 produces clearly the best result not only with respect to MSSIM but also visually. Figure 6.2(c) shows that edges are well preserved while noise is considerable removed. From Table 6.2 we further observe, that Algorithm 1 has the best performance with respect to MSSIM when $s_1 = s_2$ is sufficiently small. However, Algorithm 1 shows signs of weakness when $s_1 = s_2$ is large, i.e., $s_1 = s_2 = 0.15$ in our experiments. In contrast, the pAPS-algorithm does not suffer from this weakness and gives always better MSSIM than the ROAD-trilateral filter and the L^1-L^2-TV method with parameters as in (1.4); see Table 6.2 and Figure 6.2.



(a) corrupted image

(b) ROAD (MSSIM: 0.6118)

(c) Algorithm 1 (MSSIM: 0.7241)

(d) Parameters as in [48] (MSSIM: 0.7135)

(e) pAPS-algorithm (MSSIM: 0.7192)

FIG. 6.2. *Reconstruction of the images "cameraman" corrupted by mixed Gaussian - salt-and-pepper noise with* $\sigma = \sqrt{0.02}$, $s_1 = s_2 = 0.01$.

Similar observations as for denoising the "cameraman" image are also made for the "barbara" image; see Figure 6.3. There we again see that the pAPS-algorithm produces a result which improves the one from the L^1-L^2-TV algorithm with parameters as in [48].

For denoising the "cameraman" image corrupted by Gaussian and random-valued impulse noise we report in Table 6.3 on our findings. We observe as above that the pAPS-algorithm improves the restoration quality of the L^1-L^2-TV algorithm with parameters as in [48] or generates at least a result with the same MSSIM. This improvement is visible in Figure 6.4 for the "cameraman" image as well as for the "barbara" image. In particular, we see there for both examples that the results of the pAPS-algorithm preserve more details than the ones of the L^1-L^2-TV algo-

		ROAD	Algorithm 1			parameters as in [48]			pAPS-algorithm		
σ	$s_1 = s_2$	MSSIM	MSSIM	α_1	α_2	MSSIM	α_1	α_2	MSSIM	α_1	α_2
$\sqrt{0.02}$	0.005	0.6141	0.7383	0.1171	0.8031	0.7116	0.2000	0.8000	0.7245	0.2538	1.5717
	0.01	0.6134	0.7241	0.1115	0.7250	0.7135	0.3333	0.6667	0.7192	0.3720	1.0323
	0.05	0.5968	0.6639	0	0.8923	0.6619	0.7143	0.2857	0.6962	0.5114	0.2519
	0.15	0.4543	0.5671	0	0.6951	0.5160	0.8824	0.1176	0.6332	0.4392	0.0940
0.1	0.005	0.7011	0.7515	0.2650	0.7215	0.7290	0.3333	0.6667	0.7509	0.4552	1.5618
	0.01	0.7004	0.7446	0.2871	0.5769	0.7419	0.5000	0.5000	0.7429	0.5855	0.8821
	0.05	0.6859	0.6870	0.1094	0.5808	0.6956	0.8333	0.1667	0.7220	0.7013	0.1565
	0.15	0.5895	0.5569	0	0.8207	0.5954	0.9375	0.0625	0.6722	0.7002	0.0575
0.01	0.005	0.8293	0.8561	1.0360	0	0.8625	0.9804	0.0196	0.9461	2.6315	0.0373
	0.01	0.8283	0.8822	1.0535	0	0.8618	0.9901	0.0099	0.9375	2.1241	0.0143
	0.05	0.8163	0.8375	0.8304	0	0.8509	0.9980	0.0020	0.8844	1.3618	0.0021
	0.15	0.7562	0.7558	0.5866	0	0.8084	0.9993	0.0007	0.8206	1.2273	0.0008

TABLE 6.2

MSSIM results for the 256×256 *pixel image "cameraman" corrupted by Gaussian white noise and salt-and-pepper noise. The parameters of the ROAD-trilateral filter are* $\sigma_S = 1$, $\sigma_I = 40/255$, $\sigma_J = 30/255$, *and* σ_R *is optimized between* $10/255$ *and* $50/255$, *as suggested in [27].*

(a) corrupted image (b) ROAD (MSSIM: 0.5626) (c) Parameters as in [48] (MSSIM: 0.6030) (d) pAPS-algorithm (MSSIM: 0.6245)

FIG. 6.3. *Reconstruction of the images "barbara" corrupted by mixed Gaussian - salt-and-pepper noise with* $\sigma = \sqrt{0.02}$, $s_1 = s_2 = 0.01$.

(a) corrupted image (b) ROAD (MSSIM: 0.6145) (c) Parameters as in [48] (MSSIM: 0.7097) (d) pAPS-algorithm (MSSIM: 0.7311)

(e) corrupted image (f) ROAD (MSSIM: 0.5620 (g) Parameters as in [48] (MSSIM: 0.5851) (h) pAPS-algorithm (MSSIM: 0.6425)

FIG. 6.4. *Reconstruction of the images "cameraman" and "barbara" corrupted by mixed Gaussian - random-valued impulse noise with* $\sigma = \sqrt{0.02}$, $s_1 = s_2 = 0.01$.

σ	p	ROAD MSSIM	parameters as in [48] MSSIM	α_1	α_2	pAPS-algorithm MSSIM	α_1	α_2
$\sqrt{0.02}$	0.005	0.5629	0.7071	0.0769	0.9231	0.7298	0.1076	2.2475
	0.01	0.5620	0.7097	0.1429	0.8571	0.7311	0.1854	1.7897
	0.05	0.5577	0.7104	0.4545	0.5455	0.7104	0.4545	0.5455
	0.15	0.5439	0.6424	0.7143	0.2857	0.6721	0.5329	0.2592
0.1	0.005	0.6060	0.7161	0.1429	0.8571	0.7633	0.2278	2.9287
	0.01	0.6054	0.7232	0.2500	0.7500	0.7601	0.3602	2.0526
	0.05	0.6028	0.7365	0.6250	0.3750	0.7365	0.6250	0.3750
	0.15	0.5952	0.6823	0.8333	0.1667	0.6916	0.7666	0.1640
0.01	0.005	0.6542	0.8600	0.9434	0.0566	0.9516	3.4075	0.3228
	0.01	0.6540	0.8619	0.9709	0.0291	0.9480	2.6103	0.0790
	0.05	0.6535	0.8578	0.9940	0.0060	0.9209	1.7779	0.0078
	0.15	0.6518	0.8397	0.9980	0.0020	0.8734	1.5146	0.0024

Table 6.3

MSSIM results for the image "cameraman" corrupted by Gaussian white noise and random-valued impulse noise. The parameters of the ROAD-trilateral filter are $\sigma_S = 1$, $\sigma_I = 40/255$, $\sigma_J = 30/255$, and σ_R is optimized between $10/255$ and $50/255$, as suggested in [27].

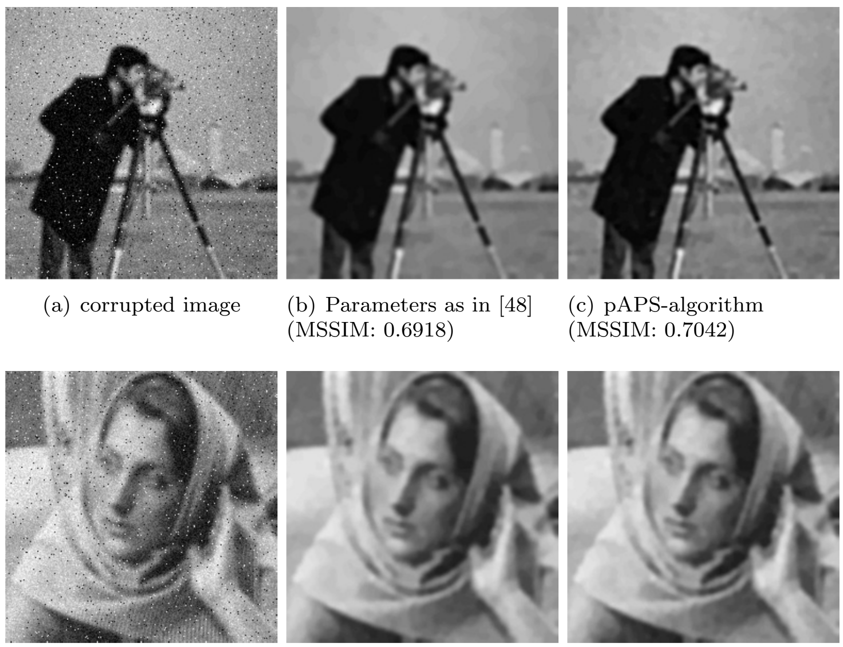
rithm with parameters as in [48]. Also here we observe that the ROAD-trilateral filter is clearly outperformed by the L^1-L^2-TV model.

Further we illustrate the successful application of our proposed algorithm when salt-and-pepper noise and Gaussian noise is disjoint present. More precisely, we consider the image in Figure 6.5 where the lower half g_1 is contaminated with salt-and-pepper noise and in the upper half g_2 only Gaussian noise is contained. This is an example where $g_1 \neq g_2$, although rather artificial, it is very interesting from a numerical point of view, since it is not possible to obtain a correct global solution by just cutting the image into an upper and a lower part due to the non-additivity of the total variation [39]. Note, that, since g_1 and g_2 are disjoint and T_1 and T_2 are restriction operators to the lower half and the upper half, respectively, problem (3.1) is superconsistent and consequently the feasible set U is non-empty. This justifies the use of the proposed pAPS-algorithm for this setting. In particular, we demonstrate with the help of this algorithm that with the correct choice of the parameters α_1 and α_2 the L^1-L^2-TV model is able to remove both type of noises considerably while preserving details at the same time from such images, see Figure 6.5(d). On the contrary, the parameters according to (1.4) obviously yield an over-smoothed restoration, see Figure 6.5(c), and thus this parameter choice rule is not suitable for such an application. Figure 6.5 also shows that the ROAD-trilateral filter does not work well for this task, since it does not remove the Gaussian noise sufficiently and over-smooths the salt-and-pepper contaminated part.

(a) corrupted image (b) ROAD (MSSIM: 0.7092) (c) Parameters as in [48] (MSSIM: 0.6707) (d) pAPS-algorithm (MSSIM: 0.7991)

Fig. 6.5. *Reconstruction of the image "cameraman" corrupted by Gaussian noise with $\sigma = 0.1$ (upper part) and salt-and-pepper noise with $s_1 = s_2 = 0.15$ (lower part).*

6.3. Reconstruction of blurred and noisy images. Now we are investigating the behavior of the proposed pAPS-algorithm for reconstructing blurred images which are additionally contaminated by mixed noise. In particular, we consider again the "cameraman" and "barbara" image and add Gaussian blur with kernel size 5×5 pixels and standard deviation 10 and corrupt it additionally by mixed Gaussian-impulse noise with $\sigma = 15/255$, $s_1 = s_2 = 0.01$ in the case of salt-and-pepper noise, and $s = 0.01$ in the case of random-valued impulse noise. For the sake of performance reference here we also compare the results of the pAPS-algorithm with the ones obtained by the L^1-L^2-TV algorithm with parameters as suggested in [48]. In Figure 6.6 and Figure 6.7 we show the respective results. We observe again that the parameters chosen by the pAPS-algorithm are more optimal than the ones suggested in [48], indicated by a larger MSSIM. This is also visible in Figure 6.6 and Figure 6.7, where the reconstructions of the pAPS-algorithm seem less blurred.



(a) corrupted image (b) Parameters as in [48] (MSSIM: 0.6918) (c) pAPS-algorithm (MSSIM: 0.7042)

(d) corrupted image (e) Parameters as in [48] (MSSIM: 0.5962) (f) pAPS-algorithm (MSSIM: 0.6037)

Fig. 6.6. *Reconstruction of the images "cameraman" and "barbara" image corrupted by Gaussian blur (kernel-size* 5×5*; standard deviation* 10*) and mixed Gaussian - salt-and-pepper noise with* $\sigma = 15/255$*,* $s_1 = s_2 = 0.01$*.*

7. Conclusion. To fully utilize the strength and advantages of the L^1-L^2-TV model the proper choice of the parameters α_1 and α_2 is essential, since they have a big influence on the restoration quality, as we see from our experiments. Therefore we present a fully automated algorithm, called pAPS-algorithm, for choosing appropriate parameters for the L^1-L^2-TV minimization problem. The automated adjustment of the parameters is based on the discrepancy principle and inspired by the work in [44]. As initial values $\alpha_i^{(0)}$ in the pAPS-algorithm we suggest to use the choice in (1.4). In this setting this method is fully automatic and generates parameters that give a satisfactory reconstruction, which are better than the ones obtained by the parameter-choice rule suggest in [48].

Due to the proposed automated parameter selection rule we are able to confirm and demonstrate once more that the L^1-L^2-TV model is suitable to reconstruct images corrupted by mixed Gaussian-impulse noise and possibly some blur; cf. [39].

Future improvements of the L^1-L^2-TV model may include spatially varying parameters, as considered in [30, 40, 44] for the L^1-TV model and L^2-TV model. In particular, large parameters α_1 and α_2 perform well in regions with small texture, while small parameters remove noise considerable in homogeneous regions. Also including an impulse noise detector in the model might be of future interest to enhance its performance of removing mixed Gaussian-impulse noise.

(a) corrupted image (b) Parameters as in [48] (MSSIM: 0.6922) (c) pAPS-algorithm (MSSIM: 0.7056)

(d) corrupted image (e) Parameters as in [48] (MSSIM: 0.5987) (f) pAPS-algorithm (MSSIM: 0.6123)

FIG. 6.7. *Reconstruction of the "cameraman" and "barbara" image corrupted by Gaussian blur (kernel-size* 5×5*; standard deviation* 10*) and mixed Gaussian - random-valued impulse noise with* $\sigma = 15/255$*,* $s = 0.01$.

Acknowledgment. The author would like to thank Andrea Barth (University of Stuttgart) for helpful discussions on stochastic processes and probability theory.

REFERENCES

[1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10(6):1217–1229, 1994.

[2] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[3] S. Alliney. A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Transactions on Signal Processing*, 45(4):913–917, 1997.

[4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[5] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, second edition, 2014. Applications to PDEs and optimization.

[6] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition – modeling, algorithms, and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.

[7] S. Bartels. *Numerical Methods for Nonlinear Partial Differential Equations*, volume 14. Springer, 2015.

[8] M. Bertalmío, V. Caselles, B. Rougé, and A. Solé. TV based image restoration with local constraints. *Journal of scientific computing*, 19(1-3):95–122, 2003.

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

[10] J.-F. Cai, R. H. Chan, and M. Nikolova. Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise. *Inverse Problems and Imaging*, 2(2):187–204, 2008.

[11] L. Calatroni, C. Chung, J. C. D. L. Reyes, C.-B. Schönlieb, and T. Valkonen. Bilevel approaches for learning of variational imaging models. *arXiv preprint arXiv:1505.02120*, 2015.

[12] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1-2):89–97, 2004. Special issue on mathematics and image analysis.

[13] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9:263–340, 2010.

[14] A. Chambolle and J. Darbon. On total variation minimization and surface evolution using parametric maximum flows. *International journal of computer vision*, 84(3):288–307, 2009.

[15] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.

[16] R. H. Chan, Y. Dong, and M. Hintermüller. An efficient two-phase-tv method for restoring blurred images with impulse noise. *IEEE Transactions on Image Processing*, 19(7):1731–1739, 2010.

[17] R. H. Chan, C.-W. Ho, and M. Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Transactions on Image Processing*, 14(10):1479–1485, 2005.

[18] T. F. Chan and S. Esedoḡlu. Aspects of total variation regularized L^1 function approximation. *SIAM J. Appl. Math.*, 65(5):1817–1837, 2005.

[19] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20(6):1964–1977, 1999.

[20] T. F. Chan and J. J. Shen. *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. Siam, 2005.

[21] P. G. Ciarlet. *Introduction to numerical linear algebra and optimisation*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1989. With the assistance of Bernadette Miara and Jean-Marie Thomas, Translated from the French by A. Buttigieg.

[22] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, 4(4):1168–1200 (electronic), 2005.

[23] J. Darbon and M. Sigelle. A fast and exact algorithm for total variation minimization. In *Pattern recognition and image analysis*, pages 351–359. Springer, 2005.

[24] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation. I. Fast and exact optimization. *J. Math. Imaging Vision*, 26(3):261–276, 2006.

[25] I. Daubechies, G. Teschke, and L. Vese. Iteratively solving linear inverse problems under general convex constraints. *Inverse Probl. Imaging*, 1(1):29–46, 2007.

[26] J. C. De los Reyes and C.-B. Schönlieb. Image denoising: Learning the noise model via nonsmooth pde-constrained optimization. *Inverse Problems & Imaging*, 7(4), 2013.

[27] J. Delon and A. Desolneux. A patch-based approach for removing impulse or mixed Gaussian-impulse noise. *SIAM J. Imaging Sci.*, 6(2):1140–1174, 2013.

[28] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34(5):1779–1791, 1997.

[29] B. Dong, H. Ji, J. Li, Z. Shen, and Y. Xu. Wavelet frame based blind image inpainting. *Applied and Computational Harmonic Analysis*, 32(2):268–279, 2012.

[30] Y. Dong, M. Hintermüller, and M. M. Rincon-Camacho. Automated regularization parameter selection in multi-scale total variation models for image restoration. *J. Math. Imaging Vision*, 40(1):82–104, 2011.

[31] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[32] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, 2008.

[33] R. Garnett, T. Huegerich, C. Chui, and W. He. A universal noise removal algorithm with an impulse detector. *IEEE Transactions on Image Processing*, 14(11):1747–1754, 2005.

[34] O. Ghita and P. F. Whelan. A new gvf-based image enhancement formulation for use in the presence of mixed noise. *Pattern Recognition*, 43(8):2646–2658, 2010.

[35] E. Giusti. *Minimal surfaces and functions of bounded variation*, volume 80 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1984.

[36] T. Goldstein and S. Osher. The split Bregman method for $L1$-regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–343, 2009.

[37] Z. Gong, Z. Shen, and K.-C. Toh. Image restoration with mixed or unknown noises. *Multiscale Modeling & Simulation*, 12(2):458–487, 2014.

[38] M. Hintermüller and K. Kunisch. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM Journal on Applied Mathematics*, 64(4):1311–1333, 2004.

[39] M. Hintermüller and A. Langer. Subspace correction methods for a class of nonsmooth and nonadditive convex variational problems with mixed L^1/L^2 data-fidelity in image processing. *SIAM J. Imaging Sci.*, 6(4):2134–2173, 2013.

[40] 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(8):085005, 30, 2010.

[41] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM J. Sci. Comput.*, 28(1):1–23, 2006.

[42] Y.-M. Huang, M. K. Ng, and Y.-W. Wen. Fast image restoration methods for impulse and Gaussian noises removal. *Signal Processing Letters, IEEE*, 16(6):457–460, 2009.

[43] K. Kunisch and T. Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM Journal on Imaging Sciences*, 6(2):938–983, 2013.

[44] A. Langer. Automated parameter selection for total variation minimization in image restoration. *arXiv preprint arXiv:1509.07442*, 2015.

[45] B. Li, Q. Liu, J. Xu, and X. Luo. A new method for removing mixed noises. *Science China Information Sciences*, 54(1):51–59, 2011.

[46] Y.-R. Li, L. Shen, D.-Q. Dai, and B. W. Suter. Framelet algorithms for de-blurring images corrupted by impulse plus Gaussian noise. *IEEE Transactions on Image Processing*, 20(7):1822–1837, 2011.

[47] J. Liu, X.-C. Tai, H. Huang, and Z. Huan. A weighted dictionary learning model for denoising images corrupted by mixed noise. *IEEE Transactions on Image Processing*, 22(3):1108–1120, 2013.

[48] R. W. Liu, L. Shi, S. C. H. Yu, and D. Wang. Box-constrained second-order total generalized variation minimization with a combined $L^{1,2}$ data-fidelity term for image reconstruction. *Journal of Electronic Imaging*, 24(3):033026–033026, 2015.

[49] E. López-Rubio. Restoration of images corrupted by Gaussian and uniform impulsive noise. *Pattern Recognition*, 43(5):1835–1846, 2010.

[50] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, volume 22. American Mathematical Soc., 2001.

[51] T. B. Moeslund. *Introduction to video and image processing: Building real systems and applications*. Springer Science & Business Media, 2012.

[52] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103(1, Ser. A):127–152, 2005.

[53] M. Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers. *SIAM J. Numer. Anal.*, 40(3):965–994 (electronic), 2002.

[54] M. Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20(1-2):99–120, 2004.

[55] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.*, 4(2):460–489 (electronic), 2005.

[56] S. Peng and L. Lucke. Fuzzy filtering for mixed noise removal during image processing. In *Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on*, pages 89–93. IEEE, 1994.

[57] A. L. Peressini, F. E. Sullivan, and J. J. Uhl. *The mathematics of nonlinear programming*, volume 1. Springer New York, 1988.

[58] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[59] Y. Shen, B. Han, and E. Braverman. Removal of mixed Gaussian and impulse noise using directional tensor product complex tight framelets. *Journal of Mathematical Imaging and Vision*, pages 1–14, 2015.

[60] L. Vese. A study in the BV space of a denoising-deblurring variational problem. *Applied Mathematics & Optimization*, 44(2):131–161, 2001.

[61] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[62] P. Weiss, L. Blanc-Féraud, and G. Aubert. Efficient schemes for total variation minimization under constraints in image processing. *SIAM J. Sci. Comput.*, 31(3):2047–2080, 2009.

[63] Y. Xiao, T. Zeng, J. Yu, and M. K. Ng. Restoration of images corrupted by mixed Gaussian-impulse noise via l_1 - l_0 minimization. *Pattern Recognition*, 44(8):1708–1720, 2011.

[64] M. Yan. Restoration of images corrupted by impulse noise and mixed Gaussian impulse noise using blind inpainting. *SIAM Journal on Imaging Sciences*, 6(3):1227–1245, 2013.

[65] J. X. Yang and H. R. Wu. Mixed Guassian and uniform impulse noise analysis using robust estimation for digital images. In *Digital Signal Processing, 2009 16th International Conference on*, pages 1–5. IEEE, 2009.

[66] M. Zhu and T. Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *UCLA CAM Report*, pages 08–34, 2008.