East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.210117.060817a

Vol. **8**, No. 1, pp. 44-69
February 2018

Minimisation and Parameter Estimation in Image Restoration Variational Models with ℓ_1-Constraints

M. Tao*

Department of Mathematics, Nanjing University, Nanjing, 210093, China.

Received 21 January 2017; Accepted (in revised version) 6 August 2017.

Abstract. Minimisation of the total variation regularisation for linear operators under ℓ_1-constraints applied to image restoration is considered, and relationships between the Lagrange multiplier for a constrained model and the regularisation parameter in an unconstrained model are established. A constrained ℓ_1-problem reformulated as a separable convex problem is solved by the alternating direction method of multipliers that includes two sequences, converging to a restored image and the "optimal" regularisation parameter. This allows blurry images to be recovered, with a simultaneous estimation of the regularisation parameter. The noise level parameter is estimated, and numerical experiments illustrate the efficiency of the new approach.

AMS subject classifications: 68U10, 65J22, 65K10, 90C25

Key words: Parameter selection, ℓ_1-constraints, alternating direction method of multipliers, impulsive noise, image processing.

1. Introduction

The restoration of images is a challenging problem, and here images $\bar{x} \in \mathbb{R}^n$ corrupted by impulse noise and blurring effects are discussed. Blurring issues are almost unavoidable in contemporary imaging systems, and corruption by impulse noise emerges from bit errors in transmission, wrong pixels and faulty memory locations in hardware [1, 6, 12, 23]. In the model, an observed image f is represented by the equation

$$f = N(K\bar{x}), \tag{1.1}$$

where N and K denote the impulse noise and blurring effect, respectively. In applications, there are two main impulse noise sources — viz. salt-and-pepper and random-valued impulse noise [9]. Image restoration problems are usually ill conditioned and the direct solution of the system (1.1) rarely produces satisfactory results. To address this problem, one can use a regularisation procedure — e.g. the total variational (TV) regularisation [35], a

*Corresponding author. *Email address:* `taom@nju.edu.cn` (M. Tao)

wavelet transform [8], or a nonlocal regularisation [21]. The corresponding unconstrained model for the image deblurring can be written in the form

$$\min_{x\in\mathbb{R}^n}\{\|Dx\|+\lambda\Phi_{\text{fit}}(x,f)\}, \tag{1.2}$$

where D is a linear operator and $\Phi_{\text{fit}}(x,f)$ is the data-fitting term. In particular, if D is the discrete gradient operator ∇, then this model (1.2) represents the TV regularisation where the parameter λ balances the data fidelity with the regularity. The constrained counterpart of the model (1.2) is

$$\begin{array}{cl} \min_{x\in\mathbb{R}^n} & \|Dx\| \\ s.t. & \Phi_{\text{fit}}(x,f)\leq\tau \end{array} \tag{1.3}$$

with τ representing the noise level, and constrained models have been also used in image recovery [32, 41, 45]. For the impulse noise, one sets $\Phi_{\text{fit}}(x,f)=\|Kx-f\|_1$ in either of (1.2) or (1.3), which are respectively called unconstrained and constrained ℓ_1-models.

Numerical methods have been used to solve the unconstrained model (1.2) under the TV regularisation [17, 40, 45], a wavelet regulariser [14], and a nonlocal regularisation [15]. The associated minimiser $\hat{x}=\hat{x}(\lambda)$, and hence the recovered image, depends on the choice of λ. Usually, the parameter λ is determined manually by a trial-and-error method, but several techniques have been developed to detect the best parameter λ automatically — e.g. the L-curve criterion [24], the generalised cross-validation (GCV) [19,20], the normalised cumulative or residual periodogram approach [25,36], variational Bayes approaches [2, 3, 34], and Morozov's discrepancy principle (MDP) [31]. On the other hand, constrained ℓ_1-models have only recently been considered, and if *a priori* information about noise is available then the constrained model is more attractive [41, 47].

To solve unconstrained ℓ_1-models, various algorithms have been developed recently. In particular, a fast TV deblurring algorithm (FTVd) combines the variable splitting and quadratic penalised technique [40]. Each subproblem can be solved by either shrinkage or the fast Fourier transform (FFT), so the FTVd method performs much better than many other methods [17, 42]. An incremental version of the FTVd approach involving the alternating direction method of multipliers (ADMM) can be used to solve an unconstrained TV-ℓ_1-model. Wu *et al.* [44] developed the inexact augmented Lagrangian method equivalent to the ADMM. All of these approaches assume that a suitable regularisation parameter in the unconstrained model is known — and it can be determined manually by a trial-and-error method, but this procedure is often very slow, so one may prefer to restore images corrupted by impulse noise via the constrained TV-ℓ_1-model. Weiss *et al.* [45] used Nesterov's first-order scheme to do so, and this approach was improved by Ng *et al.* [33], but it requires inner iterations to describe projections onto an ℓ_1-ball.

The parameter λ in the unconstrained model (1.2) can also be determined by some classical methods — e.g. if the regularisation term has a quadratic form the GCV evaluation formula may be employed. However, this formula cannot be used directly in the seminorm-based model (1.2) due to the non-linearity of the seminorms $\|D\cdot\|_2$. The regularisation parameter can be estimated by quadratic approximation of the seminorm term [30], or by using an extra quadratic penalty term instead of the non-smooth term, but the resulting model involves an additional parameter, under-smoothes the solution and can lead to multiple minimisers for the GCV function [28]; and the L-curve method [24, 29] for finding the regularisation parameter is expensive and the location of the L-curves is difficult. Another popular method to evaluate the regularisation parameter is the MDP, which selects the optimal regularisation parameter λ by matching the norm of the residual (the violation of the data-fitting term) to an upper bound. This means that the solution $x(\hat{\lambda})$ satisfies the discrepancy equation

$$\Phi_{\text{fit}}(\hat{x}(\lambda), f) = \tau \, . \tag{1.4}$$

Recall that $\hat{x}(\lambda)$ is the minimiser of (1.2), and note that parameter-selection methods for Gaussian noise [47] and Poisson data [8, 41] based on the MDP allow the discrepancy equation to be solved iteratively. These approaches cannot be used for the evaluation of the regularisation parameter in unconstrained ℓ_1-models, however.

The goal of this work is to provide a fast scheme for simultaneous solution of constrained ℓ_1-models, and to propose a method for the evaluation of the regularisation parameter in the unconstrained model (1.2). The TV regulariser $\|\nabla x\|$ in the model (1.2) is isotropic if the l_2-norm is used, and anisotropic for the l_1-norm. Our approach is applicable in both isotropic and anisotropic TV regulariser. Here the TV regulariser for the models (1.2) and (1.3) in the isotropic case $\|\cdot\| := \|\cdot\|_2$ is considered, for the treatment of the anisotropic case is similar. Moreover, our approach also works for tight frames and nonlocal regularisers, and allows for more general data-fitting terms — e.g. for the ℓ_∞-norm, suitable for uniform noise [46], and the norm $\|SKx - f\|_1$ where S is a mask matrix used in image inpainting. Let us reformulate (1.3) as a separable convex problem, and apply the alternating direction method of multiplier method (ADMM) [18, 22]. This differs from reformulations in Refs. [33, 45], and provides a closed-form solution to each subproblem, simultaneously introducing a scalar sequence converging to the "optimal" regularisation parameter λ in (1.2). The recovered quality of the unconstrained model with such an "optimal" λ is comparable to the constrained model. Moreover, the ADMM can be replaced by other first-order algorithms. Theoretical analysis based on Lagrange dual theory shows the interconnection between the constrained and unconstrained models, and the high efficiency of our approach. A new method to evaluate the parameter τ in the model (1.3) is now proposed.

Throughout, the following notation is used. Let $x = (x_1, x_2, \cdots, x_n)^\top$ be extended by periodic boundary conditions. Only square images such that $n = m^2$, $m \in \mathbb{N}$ are considered. The operator $\nabla_i x \in \mathbb{R}^2$ represents the first-order finite difference of x at the pixel i in each of the horizontal and vertical directions — i.e.

$$\nabla_i x := ((\nabla^{(1)} x)_i, (\nabla^{(2)} x)_i)^\top \in \mathbb{R}^2, \quad i = 1, \cdots, n \, ,$$

where

$$(\nabla^{(1)}x)_i := \begin{cases} x_{i+m} - x_i & \text{if } 1 \le i \le m(m-1)\,, \\ x_{\mathrm{mod}(i,m)} - x_i & \text{otherwise}\;; \end{cases} \tag{1.5}$$

$$(\nabla^{(2)}x)_i := \begin{cases} x_{i+1} - x_i & \text{if } \mathrm{mod}(i,m) \neq 0\,, \\ x_{i-m+1} - x_i & \text{otherwise}\,. \end{cases} \tag{1.6}$$

The discrete gradient operators $\nabla^{(1)}$ and $\nabla^{(2)}$ are $n \times n$ matrices, where the i-th rows of $\nabla^{(1)}$ and $\nabla^{(2)}$ respectively correspond to the first and second rows of ∇_i. The discrete total variation seminorm is $\|\nabla x\| = \sum_{i=1}^{n} \|\nabla_i x\|_2$, where the quantity $\|\nabla_i x\|_2$ measures the total variation of x at the pixel i. The resulting total variation is called isotropic. (On the other hand, the corresponding total variation is referred to as anisotropic when the l_1-norm is used, but here the symbol $\|\cdot\|$ is reserved for the l_2-norm.) We also let $\nabla \equiv (\nabla^{(1)}; \nabla^{(2)}) \in \mathbb{R}^{2n\times n}$ be the global first-order finite-difference operator such that $\nabla^\top \nabla = \sum_{i=1}^{n} \nabla_i^\top \nabla_i$, and denote the range and null spaces of an operator A by $\mathscr{R}(A)$ and $\mathscr{N}(A)$, respectively. A diagonal matrix S is called a mask matrix if its diagonal entries are 1 or 0, where the values 1 and 0 correspond to sampled and missing pixels, respectively. Given a vector x and an index set $\Omega \subseteq \{1, \cdots, n\}$, the symbol Ω^C denotes the complement of Ω in $\{1, \cdots, n\}$, x_Ω denotes the entries of the vector x restricted on the set Ω, and $\mathring{D}$ the relative interior of the set D [4].

In Section 2, MDP-based approaches for parameter selection are discussed, but shown to be inapplicable in ℓ_1-models. Section 3 establishes relationships between the constrained and unconstrained models. In Section 4, an ADMM-based approach is used to solve the constrained ℓ_1-type model (1.3), and a sequence of the Lagrange multipliers is shown to converge to the “optimal" regularisation parameter. The numerical results presented in Section 6 confirm the validity of the theoretical analysis. Concluding remarks are in Section 7, and the Appendix contains the solution of an image inpainting problem.

2. Parameter Selection by MDP-based methods

In order to find the parameter λ by an MDP-based method, one has to find the solution of the discrepancy equation (1.4). Wen & Chan [47] developed a proximal point method using a representation of the TV term in the dual formulation, and similar MDP-based approaches for Poisson noise and multiplicative Gamma noise have been used [8,41] where first-order primal-dual algorithms to solve the constrained model have been considered. One of the least squares constrained problems studied is

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x-q\|^2 \quad \text{subject to} \quad \Phi_{\mathrm{fit}}(x,f) \le \tau\,,$$

where $q \in \mathbb{R}^n$ is a given vector. The corresponding unconstrained problem

$$\min_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-q\|^2 + \lambda \Phi_{\mathrm{fit}}(x,f) \right\} \tag{2.1}$$

has an analytical solution for the Gaussian noise data-fitting term $\Phi_{\text{fit}}(x,f) = \|Kx - f\|^2$, and for the Kullback-Leibler divergence distance

$$\Phi_{\text{fit}}(x,f) = \begin{cases} \langle 1, f\log(f) - f\log(Kx) + Kx - f \rangle & \text{if } (Kx) > 0\,, \\ \infty & \text{otherwise}\,, \end{cases}$$

in the case of Poisson or multiplicative Gamma noise. If $\hat{x}(q,\lambda)$ is the corresponding solution of (2.1), then the discrepancy equation

$$\Phi_{\text{fit}}\big(\hat{x}(q,\lambda), f\big) = \tau$$

has a unique solution since the function $\Phi_{\text{fit}}(\hat{x}(q,\lambda), f)$ is strictly monotone with respect to λ. Thus such methods produce two sequences $\hat{x}(q,\lambda)$ and λ, where the first converges to the solution of the constrained model and the second to the regularisation parameter of the unconstrained model. It is notable that the applicability of the above approach depends on both the exact analytical solution of the problem (2.1) and the strict monotonicity of the function $\Phi_{\text{fit}}(\hat{x}(q,\lambda), f)$ — and these two conditions are satisfied for Gaussian, Poisson and multiplicative Gamma noise. For unconstrained ℓ_1-models, the first condition can be satisfied by replacing the term Kx with a new variable such that the problem (2.1) still has a closed-form solution [8, 41]. On the other hand, the function $\Phi_{\text{fit}}(\hat{x}(q,\lambda), f)$ is not strict monotone with respect to λ and the discrepancy equation can have multiple solutions. Thus these approaches in Refs. [8, 41, 47] are not applicable to unconstrained ℓ_1-models, and one has to find another way to find a suitable parameter λ.

3. Unconstrained and Constrained Convex Problems

For a convex function $F : \mathbb{R}^n \to \mathbb{R}$ and a proper convex lower semi-continuous function $G : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, let us consider the minimisation problems

$$\min_{x\in\mathbb{R}^n} \big\{G(x) + \lambda F(x)\big\}, \quad \lambda > 0\,, \tag{3.1}$$

$$\min_{x\in\mathbb{R}^n} G(x) \ \text{ subject to } \ F(x) \le \tau\,. \tag{3.2}$$

The Lagrange function of the constrained problem (3.2) is defined as

$$L(x,p) = G(x) + p\big(F(x) - \tau\big)\,, \tag{3.3}$$

where $p \ge 0$ is the Lagrange multiplier. Recall $(\hat{x},\hat{p}) \in \mathbb{R}^n \times \mathbb{R}_+$ ($\mathbb{R}_+ := \{x \in \mathbb{R} : x \ge 0\}$) is a saddle point of the Lagrangian function (3.3) if and only if

$$L(\hat{x},p) \le L(\hat{x},\hat{p}) \le L(x,\hat{p})\,, \tag{3.4}$$

for all $p \ge 0$ and for all $x \in \mathbb{R}^n$. If the function (3.3) has a saddle point $(\hat{x},\hat{p})$, then $\hat{x}$ is a solution of the problem (3.2). To obtain the reverse conclusion, let us recall Slater's constraint qualification (SCQ):

$$\begin{cases} \textit{If } F(x) \textit{ is a nonlinear function,} \\ \qquad \textit{then there is } x_0 \in \operatorname{dom}(G) \bigcap \operatorname{dom}(F) \bigcap \overset{\circ}{D} \;\textit{ where } D = \{x \mid F(x) \le \tau\}; \\ \textit{If } F(x) \textit{ is a linear function, then there is } x_0 \in \operatorname{dom}(G) \bigcap \operatorname{dom}(F) \bigcap D\,. \end{cases} \tag{3.5}$$

Consequently, if the problem (3.2) has a solution $\hat{x}$ and the SCQ is fulfilled, then $(\hat{x},\hat{p})$ is a saddle point of (3.3) for some $\hat{p} \ge 0$.

The following theorem shows the relationship between constrained and unconstrained problems.

Theorem 3.1. *Assume that $F : \mathbb{R}^n \to \mathbb{R}$ is convex and $G : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper, convex, and lower semi-continuous. If the minima*

$$\tau_L := \min_{x \in \mathbb{R}^n} F(x) \quad \text{and} \quad \tau_U := \min_{x \in \arg\min\limits_{x \in \mathbb{R}^n} G(x)} F(x)$$

exist, then $\tau_L < \tau_U \iff \operatorname{argmin}_{x \in \mathbb{R}^n} F(x) \cap \operatorname{argmin}_{x \in \mathbb{R}^n} G(x) = \emptyset$.

Let
$$\tau_L < \tau < \tau_U\,. \tag{3.6}$$

i) If $\hat{x}$ is a solution of (3.2)*, then $F(\hat{x}) = \tau$.*

ii) Assume that the SCQ (3.5) *is satisfied. If $\hat{x}$ is a solution of* (3.2) *and $(\hat{x},\hat{p})$ is a saddle point of the Lagrangian* (3.4)*, then $\hat{p} > 0$. Moreover, $\hat{x}$ is a solution of* (3.1) *with $\lambda = \hat{p}$ — i.e.*

$$\hat{x} \in \operatorname*{argmin}_{x \in \mathbb{R}^n} \big\{ G(x) + \hat{p} F(x) \big\}.$$

Proof. It follows from the definition of τ_L and τ_U that $\tau_L \le \tau_U$. Moreover,

$$\tau_L = \tau_U \iff \arg\min_{x \in \mathbb{R}^n} F(x) \cap \arg\min_{x \in \mathbb{R}^n} G(x) \neq \emptyset, \quad \text{hence } \tau_L < \tau_U\,.$$

i) Let $\tau_L < \tau < \tau_U$, to proceed to show that any solution $\hat{x}$ of (3.2) satisfies the equation $F(x) = \tau$. First, it is notable that $\hat{x}$ is not a minimiser of $F(x)$ because $\tau > \tau_L$. Assume that $\hat{x}$ is a minimiser of $G(x)$. Then

$$F(\hat{x}) \le \tau < \tau_U = \min_{x \in \arg\min G(x)} F(x) \le F(\hat{x}),$$

which is a contradiction. Finally, the assumption $F(\hat{x}) < \tau$ implies the false conclusion that $\hat{x}$ is a minimiser of $G(x)$, hence our assumption is wrong and $F(\hat{x}) = \tau$.

ii) Let $\hat{x}$ be a solution of (3.2) and $(\hat{x},\hat{p})$ be a saddle point of the Lagrangian. Assuming that $\hat{p} = 0$, one obtains

$$\hat{x} \in \arg\min_{x \in \mathbb{R}^n} L(x,\hat{p}) = \arg\min_{x \in \mathbb{R}^n} G(x),$$

so that $\hat{x}$ is a minimiser of $G(x)$ and this contradicts the condition $\tau < \tau_U$. Hence, $\hat{p} > 0$. Finally, from the saddle point definition one has

$$\hat{x} \in \arg\min_{x\in\mathbb{R}^n} L(x,\hat{p}) = \arg\min_{x\in\mathbb{R}^n} \left\{G(x)+\hat{p}(F(x)-\tau)\right\} = \arg\min_{x\in\mathbb{R}^n} \left\{G(x)+\hat{p}F(x)\right\},$$

so $\hat{x}$ is a solution of (3.1), which completes the proof. $\square$

Let us now consider the models

$$\min_{x\in\mathbb{R}^n} \left\{\|Dx\| + \lambda\|Kx-f\|_1\right\}, \quad \lambda > 0\,, \tag{3.7}$$

$$\min_{x\in\mathbb{R}^n} \|Dx\| \quad \text{subject to} \quad \|Kx-f\|_1 \le \tau\,, \quad \tau > 0\,, \tag{3.8}$$

where $K \in \mathbb{R}^{s,n}$, $D \in \mathbb{R}^{t,n}$, $f \in \mathbb{R}^s$ and $\|\cdot\|$ is a norm on $\mathbb{R}^s$. Let us also assume that

$$\mathcal{N}(K)\cap\mathcal{N}(D) = \{0\}\,. \tag{3.9}$$

The function $\|Kx-f\|_1$ is coercive on any nontrivial subspace Y of $\mathbb{R}^n$ with $\mathcal{N}(K)\cap Y = \{0\}$, hence the minima

$$\tau_L := \min_{x\in\mathbb{R}^n} \|Kx-f\|_1, \quad \tau_U := \min_{x\in\mathcal{N}(D)} \|Kx-f\|_1 \tag{3.10}$$

exist. Recalling that the zero space of the operator $D := \nabla$ is $\mathcal{N}(\nabla) = \{\alpha 1_n : \alpha \in \mathbb{R}\}$ and the blur operator K usually possesses the property $K1_n = 1_n$, one has

$$\tau_U := \min_{\alpha\in\mathbb{R}} \|\alpha\, 1_n - f\|_1\,,$$

which is just the median of f — cf. Ref. [37]. Let us now assume that $\tau_L < \tau_U$. The following theorem provides solvability conditions for the models (3.7) and (3.8).

Theorem 3.2. *If $\mathcal{N}(K)\cap\mathcal{N}(D) = \{0\}$, then:*
i) the problem (3.7) *has a minimiser;*
ii) and if in addition $\tau \ge \tau_L$, the problem (3.8) *has a minimiser.*

Proof. Since all norms on $\mathbb{R}^n$ are equivalent, the theorem is valid for any norm used in the data and regularity terms. The proof follows standard arguments.

i) It suffices to show that the objective function $E(x) := \|Dx\| + \lambda\|Kx-f\|_1$ is coercive — i.e. that $E(x) \to +\infty$ as $\|x\| \to +\infty$. Assume this is not true. Then there is a sequence $\{x^k\} \in \mathbb{R}^n$ such that $\lim_{k\to\infty} \|x^k\| = \infty$ but

$$\|Dx^k\| + \lambda\|Kx^k - f\|_1 \le C\,,$$

and consequently

$$\|Dx^k\| \le C\,, \quad \|Kx^k - f\|_1 \le C\,, \tag{3.11}$$

for any $k \in \mathbb{N}$. Let $\mathcal{N}(D) \oplus \mathcal{R}(D^\top)$ be the orthogonal decomposition of $\mathbb{R}^n$ and let $x^k = x_0^k + x_1^k$ be the corresponding representation of an element x^k. The operator D considered on the space $\mathscr{R}(D^\top)$ is injective. Hence for the element $Dx^k = y^k$ one has $x_1^k = D^\dagger y^k$, where $D^\dagger$ is a Moore-Penrose inverse of D [5]; and using (3.11) one obtains

$$\|x_1^k\| = \|x^k - x_0^k\| = \|D^\dagger y^k\| \leq \|D^\dagger\|\|y^k\| = \|D^\dagger\|\|Dx^k\| \leq C_2\,. \tag{3.12}$$

Moreover, the triangle inequality and (3.11) imply that

$$\begin{aligned}\|Kx_0^k\|_1 &\leq \left\|K(x^k - x_0^k)\right\|_1 + \|f\|_1 + \|Kx^k - f\|_1 \\ &\leq C_3\|x^k - x_0^k\| + \|f\|_1 + \|Kx^k - f\|_1 \leq C_4\,.\end{aligned} \tag{3.13}$$

Let us assume there is a subsequence $x_0^{k_j}$ of x_0^k such that $\|x_0^{k_j}\| \to \infty$ as $j \to \infty$. Since $\mathscr{N}(K) \cap \mathscr{N}(D) = \{0\}$, the operator K is injective on $\mathscr{N}(D)$ such that the norms $\|Kx_0^{k_j}\|_1$ tend to ∞, which contradicts the inequality (3.13). Thus $\|x_0^k\| \leq C_5$ for all k — and from (3.12) the sequence $\{x^k\}$ is bounded, contrary to our assumption.

ii) Since $\tau \geq \tau_L$, the set

$$X := \left\{x \in \mathbb{R}^n : \|Kx - f\|_1 \leq \tau\right\}$$

is nonempty. Moreover, it is closed and convex. The directions of the recession of X and the objective function in (3.8) are respectively $\mathscr{N}(K)$ and $\mathscr{N}(D)$. Since $\mathscr{N}(K) \cap \mathscr{N}(D) = \{0\}$, it follows from Proposition 2.3.2 in Ref. [4] that the solution set of (3.8) is nonempty and compact. □

Corollary 3.1. *If $\mathscr{N}(K) \cap \mathscr{N}(D) = \{0\}$ and $\tau_L < \tau < \tau_U$, the problem* (3.8) *has a positive Lagrange multiplier.*

The proof of this result follows from the fact that the Slater condition for the problem (3.8) is fulfilled if $\tau \geq \tau_L$.

4. ADMM Approach to Constrained ℓ_1-Models

In this section, the constrained problem (3.8) is solved with the simultaneous evaluation of the regularisation parameter of the unconstrained model (3.7). More precisely, consider the model (3.8) with the TV regularisation — i.e. when $\|Dx\| = \sum_i \|\nabla_i x\|$ and the blur operator $K \in \mathbb{R}^{n\times n}$. The resulting models (3.7) and (3.8) are called unconstrained and constrained TV-ℓ_1-models, respectively. In passing, it is notable that our approach is also applicable to other regularisations.

The choice of the parameter τ in the constrained model (3.8) satisfies the condition (3.6). From Theorem 3.1(i), the constraint in (3.8) is tight. Assuming that the parameter τ satisfies the condition (3.6), one can rewrite the constrained TV-ℓ_1-model as

$$\begin{aligned}\min_{x,z_i} &\sum_{i=1}^n \|z_i\| \\ s.t.\quad &\nabla_i x = z_i \qquad\qquad : \mu_i \\ &\|Kx - f\|_1 = \tau \quad : \kappa\end{aligned} \tag{4.1}$$

where $z_i \in \mathbb{R}^2$ is an auxiliary vector, and $\mu_i \in \mathbb{R}^2$ and $\kappa \in \mathbb{R}$ are the Lagrange multipliers of the first and second equalities in the model (4.1).

Moreover, since the equation

$$\|Kx - f\|_1 = \tau$$

can be rewritten as

$$-y \leq Kx - f \leq y\,, \quad e^\top y = \tau\,,$$

where $e = (1,1,\cdots,1) \in \mathbb{R}^n$, the model (4.1) is equivalent to the minimisation problem

$$\begin{aligned}
&\min_{x,y,z_i,w_1,w_2} \sum_{i=1}^{n} \|z_i\| \\
s.t.\quad & \nabla_i x = z_i && : \mu_i \\
& Kx - f - y + w_1 = 0 && : \nu \\
& Kx - f + y - w_2 = 0 && : \delta \\
& e^\top y = \tau && : \kappa \\
& w_1, w_2 \geq 0
\end{aligned} \tag{4.2}$$

with the Lagrange multipliers $\mu_i \in \mathbb{R}^2$, $\nu \in \mathbb{R}^n$, $\delta \in \mathbb{R}^n$ and $\kappa \in \mathbb{R}$. Similar to the definitions of $(\nabla^{(1)}x)_i$ and $(\nabla^{(2)}x)_i$ in Eqs. (1.5) and (1.6), consider a vector

$$z := \begin{pmatrix} z^{(1)} \\ z^{(2)} \end{pmatrix} \in \mathbb{R}^{2n^2}, \quad z_i := \begin{pmatrix} (z^{(1)})_i \\ (z^{(2)})_i \end{pmatrix} \in \mathbb{R}^2, \quad i = 1,\cdots,n^2\,.$$

Then the augmented Lagrangian function of the problem (4.2) is

$$\begin{aligned}
&\mathscr{L}_{\mathscr{A}}(x,y,z,w_1,w_2;\mu,\nu,\delta,\kappa) \\
=&\sum_{i=1}^{n} \|z_i\| + \langle \nu, Kx - y - f + w_1\rangle + \frac{\beta_1}{2}\|Kx - y - f + w_1\|^2 \\
&+ \langle \delta, Kx + y - f - w_2\rangle + \frac{\beta_1}{2}\|Kx + y - f - w_2\|^2 + \langle \kappa, e^\top y - \tau\rangle \\
&+ \frac{\beta_2}{2}\|e^\top y - \tau\|^2 + \sum_{i=1}^{n}\langle \mu_i, \nabla_i x - z_i\rangle + \frac{\beta_3}{2}\sum_{i=1}^{n}\|\nabla_i x - z_i\|^2,
\end{aligned}$$

where $\beta_1,\ \beta_2,\ \beta_3 > 0$. Using the ADMM described in Refs. [18,22] starting at $x = x^k$,

$y = y^k$, $\mu = \mu^k$, $\nu = \nu^k$, $\delta = \delta^k$ and $\kappa = \kappa^k$, one arrives at the following iterative scheme:

$$\begin{pmatrix} w_1^{k+1} \\ w_2^{k+1} \\ z^{k+1} \end{pmatrix} \leftarrow \operatorname*{arg\,min}_{w_1 \in \mathbb{R}_+^n, w_2 \in \mathbb{R}_+^n, z} \mathscr{L}_{\mathscr{A}}(x^k, y^k, z, w_1, w_2; \mu^k, \nu^k, \delta^k, \kappa^k), \tag{4.3}$$

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \end{pmatrix} \leftarrow \operatorname*{arg\,min}_{x,y} \mathscr{L}_{\mathscr{A}}(x, y, z^{k+1}, w_1^{k+1}, w_2^{k+1}; \mu^k, \nu^k, \delta^k, \kappa^k), \tag{4.4}$$

$$\begin{pmatrix} \mu_i^{k+1} \\ \nu^{k+1} \\ \delta^{k+1} \\ \kappa^{k+1} \end{pmatrix} \leftarrow \begin{pmatrix} \mu_i^k + \beta_3(\nabla_i x^{k+1} - z_i^{k+1}) \\ \nu^k + \beta_1(Kx^{k+1} - f - y^{k+1} + w_1^{k+1}) \\ \delta^k + \beta_1(Kx^{k+1} - f + y^{k+1} - w_2^{k+1}) \\ \kappa^k + \beta_2(e^\top y^{k+1} - \tau) \end{pmatrix}. \tag{4.5}$$

This method converges for any positive numbers $\beta_1, \beta_2, \beta_3$ [7,27].

Now let us split the minimisation procedure (4.3) into three independent subproblems with respect to w_1, w_2, and z. The algorithms for w_1 and w_2 are obtained by projecting the corresponding terms on the space $\mathbb{R}_+^n$ — viz.

$$w_1^{k+1} = \mathscr{P}_{\mathbb{R}_+^n}\left[y^k + f - \frac{\nu^k}{\beta_1} - Kx^k \right], \tag{4.6}$$

$$w_2^{k+1} = \mathscr{P}_{\mathbb{R}_+^n}\left[y^k - f + \frac{\delta^k}{\beta_1} + Kx^k \right]. \tag{4.7}$$

The third subproblem is equivalent to n two-dimensional problems of the form

$$\min_{z_i \in \mathbb{R}^2} \left\{ \|z_i\| + \frac{\beta_3}{2} \left\| z_i - \left(\nabla_i x^k + \frac{1}{\beta_3}(\mu^k)_i \right) \right\|^2 \right\}, \quad i = 1, 2, \cdots, n,$$

and it has the analytical solution

$$z_i^{k+1} = \max\left\{ \left\| \nabla_i x^k + \frac{1}{\beta_3}(\mu^k)_i \right\| - \frac{1}{\beta_3}, 0 \right\} \frac{\nabla_i x^k + \frac{1}{\beta_3}(\mu^k)_i}{\|\nabla_i x^k + \frac{1}{\beta_3}(\mu^k)_i\|}, \quad i = 1, 2, \cdots, n, \tag{4.8}$$

presented in Refs. [39,43]. The computational cost in solving (4.8) grows linearly with respect to n.

The minimisation problem (4.4) can also be simplified by splitting it into two independent subproblems, one with respect to x and another with respect to y. The x subproblem is just a least squares problem with the normal equation

$$\left(\nabla^\top \nabla + 2\frac{\beta_1}{\beta_3} K^\top K \right) x^{k+1} = \nabla^\top \left(z^{k+1} - \frac{\mu^k}{\beta_3} \right) + \frac{\beta_1}{\beta_3} K^\top \left(2f - w_1^{k+1} + w_2^{k+1} - \frac{\nu^k + \delta^k}{\beta_1} \right). \tag{4.9}$$

If the condition (3.9) holds, the coefficient matrix in Eq. (4.9) is nonsingular. Moreover, for boundary conditions periodic in x, the matrices $\nabla^\top \nabla$ and $K^\top K$ are block circulant

[10, 23], so they are diagonalisable via the two-dimensional discrete Fourier transform (FFT) at $O(n\log n)$ cost. For a symmetric blur and Neumann boundary conditions, the coefficient matrix can be diagonalised by the discrete cosine transform (DCT) at the same cost — cf. Ref. [32]. The y subproblem is likewise a least squares problem with the normal equation

$$\left(2I+\frac{\beta_2}{\beta_1}ee^\top\right)y^{k+1}=\frac{\beta_2}{\beta_1}\tau e+\left(w_1^{k+1}+w_2^{k+1}\right)+\frac{\nu^k-\delta^k-\kappa^k e}{\beta_1}. \tag{4.10}$$

According to the Sherman-Morrison-Woodburg theorem,

$$\left(2I+\frac{\beta_2}{\beta_1}ee^\top\right)^{-1}=\frac{1}{2}I-\frac{\frac{\beta_2}{\beta_1}}{4+2\frac{\beta_2}{\beta_1}\cdot n}ee^\top,$$

and the cost of solving of the y-subproblem is $O(n)$. Finally, the update of (4.5) for μ_i, ν, δ, and κ is straightforward and requires $O(n)$ operations. Each subproblem preserves a closed-form solution, the corresponding regularisation parameter can be obtained as the limit of the sequence κ^k, and the per-iteration cost for the scheme (4.3)-(4.5) dominated by two FFT or two DCT operations is $O(n\log n)$.

Algorithm 4.1 ADMM approach for the constrained TV-ℓ_1-model (4.1).

Input f, K, $\tau>0$, β_1, β_2, $\beta_3>0$, **and** μ^0, ν^0, δ^0, κ^0. **Initialise** $x=f$, $y=f$ **and** $\mu=\mu^0$, $\nu=\nu^0$, $\delta=\delta^0$, **and** $\kappa=\kappa^0$.
Output $\hat{x}$, $\hat{\kappa}$.

While "a stopping criterion is not satisfied", **Do**
1) Compute w_1^{k+1} according to (4.6).
2) Compute w_2^{k+1} according to (4.7).
3) Compute z^{k+1} according to (4.8).
4) Compute x^{k+1} by solving (4.9).
5) Compute y^{k+1} by solving (4.10).
6) Update μ^{k+1}, ν^{k+1}, δ^{k+1}, and κ^{k+1} via (4.5).

End Do

This algorithm is just an application of the ADMM in nonsmooth settings with different scaling parameters. Mota *et al.* [27] established the global convergence of the method when the scaling parameters are equal to each other, but their proof can readily be extended to different scalings to produce the following theorem.

Theorem 4.1. *Assume that* $f\in\mathbb{R}^n$, *the matrices* $\nabla\in\mathbb{R}^{2n\times n}$, $K\in\mathbb{R}^{n\times n}$ *satisfy the condition* (3.9), *and* μ_L *and* μ_U *are defined by* (3.10). *If* $\mu_L<\tau<\mu_U$, *then the sequence* $\{(x^k,y^k,z^k,w_1^k,w_2^k;\mu^k,\nu^k,\delta^k,\kappa^k)\}$ *generated by Algorithm* 4.1 *converges to* $(\hat{x},\hat{y},\hat{z},\hat{w}_1,\hat{w}_2;\hat{\mu},\hat{\nu},\hat{\delta},\hat{\kappa})$, *where* $\hat{x}$ *is a solution of the problem* (3.8). *Moreover,* $\hat{x}$ *is also a solution of the problem* (3.7) *with* $\lambda=\hat{\kappa}$.

Remark 4.1. The equations in (4.2) can be written as

$$\begin{pmatrix} \nabla & 0 \\ K & -I \\ K & I \\ 0 & e^\top \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 & 0 & -I \\ I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ f \\ f \\ \tau \end{pmatrix}. \tag{4.11}$$

Under the assumption (3.9), the coefficient matrices in (4.11) have full column rank, and all the corresponding constraints are polyhedral sets, hence the constrained model (4.2) is satisfied under the conditions of Theorem 1 in Ref. [27].

5. Finding the Parameter τ

When using the constrained model (3.8) for the image recovery, one has to evaluate the parameter τ, which can be done as follows:

- apply a filter [16, 26] to locate the set Ω^C with pixel values corrupted by impulsive noise;
- obtain the estimate $\hat{u}$ for $K\bar{x}$ by solving the model

$$\begin{aligned} &\min_u \|\nabla u\| \\ &s.t. \quad u_\Omega = f_\Omega\,; \end{aligned} \tag{5.1}$$

- compute the estimate for τ from the formula

$$\tau := \|\hat{u} - f\|_1\,. \tag{5.2}$$

The main idea in this scheme is to use the solution of the model (5.1) as a replacement for $K\bar{x}$, in order to estimate τ.

6. Numerical Results

Several constrained TV-ℓ_1-problems were explored numerically, for two different data-fidelity terms — viz. $\Phi_{\text{fit}} = \|Kx - f\|_1$ and $\Phi_{\text{fit}} = \|S \cdot Kx - f\|_1$. The code is written in MATLAB 7.12 (R2011a), and all the numerical experiments were conducted on a ThinkPad notebook with Intel Core i5-2140M CPU, a 2.3-GHz processor, and 4 GB of memory. The quality of the restoration was evaluated by considering the signal-to-noise ratio (SNR), measured in decibels (dB) defined by

$$\text{SNR}(x) \triangleq 10 * \log_{10} \frac{\|\bar{x} - \tilde{x}\|^2}{\|\bar{x} - x\|^2}\,,$$

where $\bar{x}$ denotes the original image and $\tilde{x}$ the mean intensity value of $\bar{x}$. The recovered image quality was measured by the mean-square error (MSE). In all of the calculations, the algorithms were terminated if

$$Res < 5 \times 10^{-3}\,. \tag{6.1}$$

Figure 1: (a) Cameraman.tif, 256 × 256; (b) Lena.jpg, 512 × 512.

For the Algorithm 4.1, the residual *Res* is defined as

$$
\begin{aligned}
&Res^{(1)} := \max\{r_{x,z}, r_{x,y,w_1}, r_{x,y,w_2}, r_y\}\,,\\
&r_{x,z} := \nabla x - z\,,\\
&r_{x,y,w_1} := Kx - f - y + w_1\,,\\
&r_{x,y,w_2} := Kx - f + y - w_2\,,\\
&r_y := e^\top y - \tau\;.
\end{aligned}
$$

The FTVd solves the unconstrained TV-ℓ_1-model $\min_x \sum_i \|\nabla_i x\| + \mu\|Kx - f\|_1$ rewritten as

$$
\begin{aligned}
\min_{x,z_i,y} \; & \sum_i \|z_i\| + \mu\|y\|_1\\
s.t. \quad & \nabla_i x = z_i\\
& Kx - f = y
\end{aligned}
$$

by the ADMM, so the residual of the FTVd is defined by $Res^{(2)} := \max\{r_{x,z}, r_{x,y}\}$ and

$$
r_{x,z} := \nabla x - z\,, \quad r_{x,y} := Kx - f - y\;.
$$

In the simulations, both *salt-and-pepper* and *random-valued* impulse noise were considered, in seeking to establish the following:

- the scheme of Section 5 produces good estimates of τ in constrained models;
- Theorem 3.1(ii) is valid for TV-ℓ_1-models — i.e. the quality of the images restored by unconstrained models, with λ determined in Theorem 4.1, is comparable with constrained models; and
- the regularisation parameter is on par with one determined by the trial-and-error method.

The tested images are shown in Fig. 1.

6.1. Deblurring from impulsive noise

The first focus was on models with the data-fidelity term $\Phi_{\text{fit}} = \|Kx - f\|_1$. Three sets of tests were used to verify the above three considerations to: (1) show the efficiency of the method (5.1)-(5.2) in the selection of a suitable parameter τ; 2) compare the SNR of the images restored by the constrained and unconstrained models and (3) compare the "optimal" regularisation parameter of the unconstrained model determined by the trial-and-error method and Algorithm 4.1.

The vector f was generated in three steps: (1) all pixel values were scaled to the interval $[0,1]$; (2) the test images were blurred with the blurring (I)–`fspecial ('average',9)` and (II)–`fspecial ('gaussian',[9 9],3)`; and (3) the resulting images were corrupted by impulse noise at different levels. All algorithms used such blurred images initially. The first set of tests dealt with blurred and noisy images degraded under periodic boundary conditions, with the impulse noise at the level 30%, 40%, 50% or 60% added. For the *salt-and-pepper* noise, the noise candidate set Ω^C was detected by the AMF algorithm [26]. The maximum window size in the AMF algorithm was set to 19, as suggested in Ref. [26]. For the *random-value* impulsive noise, the set Ω^C was identified by the statistical method with a default setting [16]. The model (5.1) with the obtained Ω was then solved by applying the ADMM, reformulated as follows:

$$\begin{aligned} &\min_{u,z,y} \|y\| \\ &s.t. \quad \nabla u = y \\ &\qquad\quad u = z \\ &\qquad\quad z_\Omega = f_\Omega. \end{aligned} \tag{6.2}$$

The penalty parameters β_1 and β_2 in the ADMM method related to the first and second equations in the model (6.2) were set to 10 and 20, respectively. The solution $\hat{u}$ of (6.2) was used to calculate τ_2 by the formula (5.2). Table 1 shows $\tau_1 := \|K\bar{x} - f\|_1$, τ_2 and the relative error $|\tau_1 - \tau_2|/\tau_1$ for each case considered. The computed value τ_2 is comparable with the true value τ_1 used in the constrained models.

The Algorithm 4.1 and the FTVd of Ref. [40] were used to solve the constrained and unconstrained TV-ℓ_1-models, respectively. In the second set of tests, blurred and noisy images were degraded under periodic boundary conditions, and impulse noise at the level 30%, 40%, 50% and 60% was added. The parameters $\beta_1 = 200$, $\beta_2 = 10$, and $\beta_3 = 200$ in Algorithm 4.1 produce satisfactory results for all tests. In the FTVd, the same penalty parameters were chosen as in Algorithm 4.1 — i.e. the respective penalty parameters for the equations $\nabla x - z = 0$ and $Kx - f - y = 0$ were $\beta_1 = 200$ and $\beta_3 = 200$. Under Algorithm 4.1, the restored image and the last iteration for $\hat{\kappa}$ were obtained simultaneously, and then the FTVd† with the derived parameter $\hat{\kappa}$ and the stopping criterion (6.1), to find the solution of the TV-ℓ_1-model (3.7). The values of the parameter $\hat{\kappa}$ and SNRs for the constrained and unconstrained models are presented in Table 2. The recovered MSEs are

†http://www.caam.rice.edu/ optimization/L1/ftvd/v4.0/

Table 1: The values of τ_1 and τ_2.

Tested image	Blur	noise level	$\tau_1 := \|K\bar{x}-f\|_1$	τ_2 via (5.1)-(5.2)	$\|\tau_1-\tau_2\|/\tau_1$
salt-and-pepper impulse noise					
a	I	30%	9.8196e+003	9.8196e+003	2.3464e-006
		40%	1.3087e+004	1.3086e+004	7.3590e-005
		50%	1.6301e+004	1.6302e+004	9.1184e-005
		60%	1.9643e+004	1.9642e+004	4.9575e-005
	II	30%	9.8209e+003	9.8206e+003	2.5360e-005
		40%	1.3097e+004	1.3096e+004	3.9853e-005
		50%	1.6455e+004	1.6454e+004	3.9977e-005
		60%	1.9515e+004	1.9515e+004	5.1470e-006
random value impulse noise					
a	I	30%	9.9073e+003	9.9063e+003	1.0360e-004
		40%	1.3155e+004	1.3156e+004	9.6410e-005
		50%	1.6382e+004	1.6383e+004	7.2474e-005
		60%	1.9691e+004	1.9687e+004	1.9382e-004
	II	30%	1.3190e+004	1.3192e+004	1.5832e-004
		40%	1.3090e+004	1.3089e+004	7.3576e-005
		50%	1.6403e+004	1.6404e+004	2.9497e-005
		60%	1.9590e+004	1.9593e+004	1.6469e-004
salt-and-pepper impulse noise					
b	I	30%	9.8691e+003	9.8692e+003	1.0326e-005
		40%	1.3183e+004	1.3184e+004	3.0565e-005
		50%	1.6373e+004	1.6373e+004	2.4667e-005
		60%	1.9641e+004	1.9642e+004	1.2129e-005
	II	30%	9.8681e+003	9.8691e+003	1.0684e-004
		40%	1.3157e+004	1.3157e+004	1.9127e-005
		50%	1.6319e+004	1.6319e+004	2.2055e-006
		60%	1.9668e+004	1.9668e+004	1.1229e-005
random value impulse noise					
b	I	30%	9.7770e+003	9.7770e+003	7.8120e-007
		40%	1.3067e+004	1.3066e+004	9.2740e-005
		50%	1.6360e+004	1.6361e+004	5.5121e-005
		60%	1.9709e+004	1.9709e+004	1.9229e-005
	II	30%	9.8392e+003	9.8397e+003	5.1033e-005
		40%	1.3113e+004	1.3114e+004	2.2270e-005
		50%	1.6347e+004	1.6348e+004	6.2731e-005
		60%	1.9659e+004	1.9661e+004	8.3863e-005

the same for constrained and unconstrained models, so not reported here. Fig. 2 shows the images degraded by the `fspecial ('average',9)` blur along with the *salt-and-pepper* impulse noise with the images restored via the two models. The results are much the same as found in other cases.

In the third set of tests, the degraded images were generated similarly. By sampling a large range of the regularisation parameter λ that also contained $\hat{\kappa}$, the FTVd was run with different regularisation parameters to identify the particular "optimal" value λ^{opt} that

Table 2: Numerical results on the models (3.7) and (3.8).

Tested image	Blur	noise level	SNR (dB) via (3.8)	$\hat{\kappa}$	SNR (dB) via (3.7) with $\hat{\kappa}$	λ^{opt}	SNR (dB) via (3.7) with λ^{opt}
salt-and-pepper impulse noise							
a	I	30%	17.519	59.090	17.401	70	17.4636
		40%	14.720	22.473	14.658	46	15.5418
		50%	13.631	21.484	13.598	32	13.8932
		60%	12.404	18.623	12.375	21	12.3928
	II	30%	16.299	63.131	16.077	108	16.4733
		40%	13.741	22.013	13.678	69	14.7644
		50%	13.155	21.667	13.122	49	13.7138
		60%	12.381	19.881	12.359	32	12.5819
random value impulse noise							
a	I	30%	17.037	51.374	16.893	60	17.003
		40%	13.344	20.706	13.020	19	13.198
		50%	10.729	12.856	10.663	9	10.898
		60%	7.426	2.503	7.499	5	7.795
	II	30%	15.784	56.051	15.716	89	16.214
		40%	12.356	19.221	12.229	33	12.375
		50%	10.762	8.797	10.776	6	10.869
		60%	7.406	2.280	7.549	5	7.872
salt-and-pepper impulse noise							
b	I	30%	16.861	55.512	16.820	60	16.828
		40%	15.070	18.258	14.983	35	15.4508
		50%	14.239	17.263	14.189	25	14.3542
		60%	13.186	14.919	13.136	18	13.1740
	II	30%	16.254	57.615	15.958	82	16.0503
		40%	14.521	17.585	14.428	42	14.8931
		50%	14.040	17.548	13.963	34	14.2696
		60%	13.249	16.573	13.179	25	13.348
random value impulse noise							
b	I	30%	16.712	49.937	16.471	58	16.498
		40%	14.313	17.692	14.246	24	14.385
		50%	12.068	14.117	12.029	13	12.035
		60%	8.607	4.665	8.575	6	8.693
	II	30%	15.954	52.114	15.770	73	15.868
		40%	13.947	17.147	13.887	25	14.060
		50%	11.657	13.026	11.603	11	11.772
		60%	8.580	3.676	8.620	4	8.696

provided the highest SNR of the restored image. The results are also displayed in Table 2, where the optimal value λ^{opt} and SNR of the recovered image are listed. For the *salt-and-pepper* impulse noise, the discrepancy between the SNRs of the images recovered via $\hat{\kappa}$ and λ^{opt} is small. On the other hand, for the *random-valued* impulsive noise there is a clear difference between the SNRs for images recovered by constrained and "optimal" unconstrained models.

Figure 2: Left: Blurred and noisy image; Middle: Images restored via constrained model (3.8); Right: Images restored via unconstrained model (3.7). Noise level: Row 1 – 30%; Row 2 – 40%; Row 3 – 50%; Row 4 – 60%.

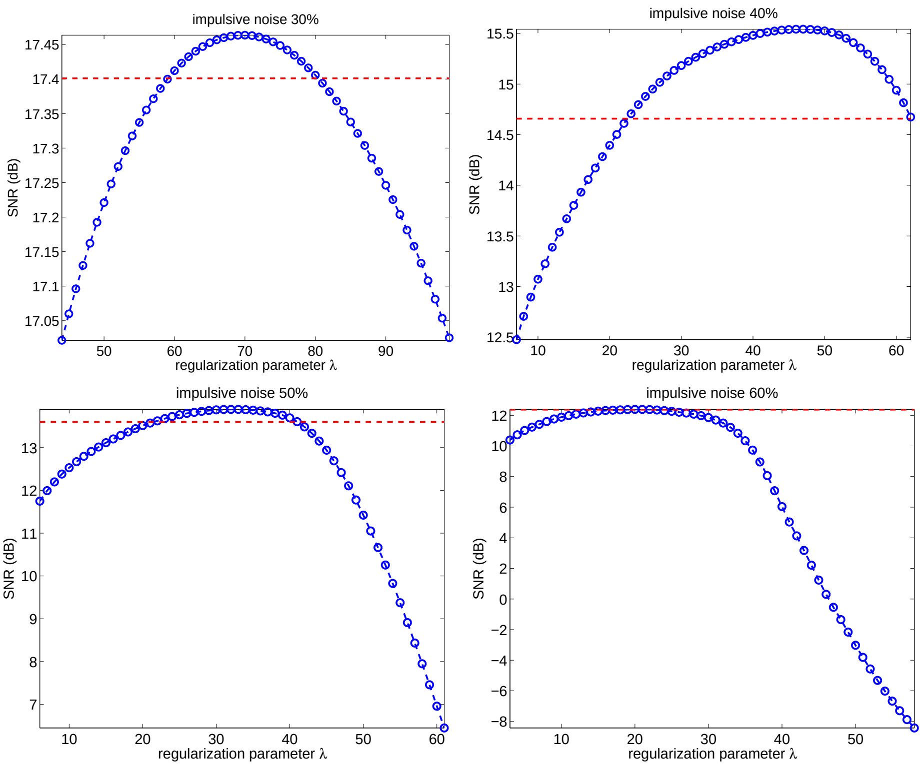

Figure 3: Restored SNR against λ, and the difference between the restored via $\hat{\kappa}$ and the trial-and-error method; Noise level: Row 1 – 30% and 40%; Row 2 – 50% and 60%. Noise: *salt-and-pepper*; Blur – I.

.

In Fig. 3, graphs of SNRs for the unconstrained TV-ℓ_1-model (3.7) with different regularisation parameters are shown. The red dashed line represents the SNR with $\hat{\kappa}$ as the regularisation parameter. The curve over the red dashed line shows the gap between the SNRs for images recovered via $\hat{\kappa}$ and the “optimal" value found. The graphs in Fig. 3 correspond to a different level of *salt-and-pepper* noise. For the “Cameraman.tif" image degraded by `fspecial ('average',9)` blur, for low and middle noise levels the discrepancy between the SNRs for $\hat{\kappa}$ and “optimal" values is about 0.2 dB, and diminishes for higher noise levels. The results of numerical experiments for the Neumann boundary conditions were similar, so not reported here.

6.2. Inpainting from impulsive noise

The efficiency of the approach was explored for the models (1.2) and (1.3) with the data-fidelity term $\Phi_{\text{fit}} = \|SKx - f\|_1$. The resulting models

$$\min_x \|\nabla x\| + \lambda \|SKx - f\|_1 \,, \tag{6.3}$$

$$\begin{aligned} &\min_x \|\nabla x\| \\ &s.t. \quad \|SKx - f\|_1 \le \tau \,, \end{aligned} \tag{6.4}$$

are suitable for image inpainting from impulsive noise. The parameter τ in the model (6.4) can be determined analogously to Section 5 as follows:

- apply a filter from Ref. [16] to locate the set Π^C where the pixel values are corrupted by impulsive noise;
- obtain the estimate $\hat{u}$ of $K\bar{x}$ by solving the model

$$\begin{aligned} &\min_u \|\nabla u\| \\ &s.t. \quad (Su)_\Pi = f_\Pi \,; \end{aligned} \tag{6.5}$$

- compute the estimate for τ from the formula

$$\tau := \|(S\hat{u}) - f\|_1 \,. \tag{6.6}$$

In the first set of tests, τ was obtained from the model (6.5) and Eq. (6.6). The images were degraded by blur and noise under periodic boundary conditions, and different sample ratios ranging from 30% to 70% with equal 10% distance were tested, followed by *salt-and-pepper* impulsive noise. The noise level was set to 5% and 15%, and the model (6.5) solved by the ADMM reformulated as

$$\begin{aligned} &\min_{u,y,w,z} \|y\| \\ &s.t. \quad \nabla u = y \\ &\qquad\; u = w \\ &\qquad\; Su = z \\ &\qquad\; z_\Pi = f_\Pi \,, \end{aligned} \tag{6.7}$$

where the chosen penalty parameter values in the ADMM were $\beta_1 = 10$, $\beta_2 = 20$, and $\beta_3 = 20$. Recall that β_1, β_2, and β_3 are related to the first, second, and third equations in (6.7), respectively. The solution $\hat{u}$ of (6.7) was obtained, and then computed τ_2 using the formula (6.6). The results are presented in Table 3, where the parameter τ_2 is seen to be a good approximation for the true value τ_1, so it was used in subsequent tests for the constrained model (6.4). The details of these tests are discussed in the Appendix, where the procedure from Section 6.1 was followed. Thus two types of impulsive noise and two types of the blur were considered, and vector f generated as follows. A convolution K was applied to $\bar{x}$ to obtain $K\bar{x}$ under the periodic boundary condition, and random samples were taken to get $SK\bar{x}$. Various sample ratios ranging from 30% to 70% with the equal distance 10% were tested, and then the impulsive noise added with the noise level set to 5% and 15%. The SNRs of the images restored by constrained and unconstrained models were compared, and the “optimal" regularisation parameter obtained was used in the corresponding unconstrained model.

Table 3: The values of τ_1 and τ_2.

Tested image	Blur	sample ratio	$\tau_1 := \|SK\bar{x} - f\|_1$	τ_2 via (6.5)-(6.6)	$\lvert\tau_1 - \tau_2\rvert/\tau_1$
salt-and-pepper (noise level - 5%)					
a	I	30%	8.4583e+002	8.4670e+002	1.0285e-003
		40%	1.1467e+003	1.1469e+003	1.8674e-004
		50%	1.3933e+003	1.3936e+003	2.4051e-004
		60%	1.6902e+003	1.6908e+003	3.6453e-004
		70%	1.9872e+003	1.9879e+003	3.0933e-004
	II	30%	8.6568e+002	8.6533e+002	4.1154e-004
		40%	1.1611e+003	1.1615e+003	3.3930e-004
		50%	1.4558e+003	1.4558e+003	3.1312e-005
		60%	1.7513e+003	1.7515e+003	1.0286e-004
		70%	2.0249e+003	2.0250e+003	7.6999e-005
salt-and-pepper (noise level - 15%)					
b	I	30%	2.5571e+003	2.5572e+003	4.4790e-005
		40%	3.4058e+003	3.4076e+003	5.1017e-004
		50%	4.2525e+003	4.2529e+003	9.7112e-005
		60%	5.1351e+003	5.1353e+003	3.6894e-005
		70%	5.9105e+003	5.9107e+003	3.5500e-005
	II	30%	2.5622e+003	2.5630e+003	3.1496e-004
		40%	3.4350e+003	3.4347e+003	6.6886e-005
		50%	4.3173e+003	4.3171e+003	5.2739e-005
		60%	5.1593e+003	5.1601e+003	1.6458e-004
		70%	6.0577e+003	6.0574e+003	4.3713e-005

The ADMM from Ref. [38] was used to solve the unconstrained model (6.3), and the results are presented in Table 4. It is notable that SNRs for the images recovered by these models are almost the same. Fig. 4 shows the restoration results for the constrained model and the penalised version for the sample ratios 30%, 40%, 50% and 60%.

Fig. 5 shows graphs of the SNR for the image restored by the penalised version (6.3) against λ. The red dashed lines represent the SNRs for $\lambda = \hat{\kappa}$ in the penalised model (6.3), and the curves over the red dashed lines indicate the gap between the SNRs of the images recovered via the parameter $\hat{\kappa}$ and via the optimal recovered value. It is again notable that the gap diminishes as the sample ratio increases.

7. Conclusions

Minimisation of the seminorms $\|D \cdot\|$ for ℓ_1-data-fidelity terms has been considered, where the solution of the discrepancy equation is not unique. Connexions between constrained and unconstrained models have been established, and the optimal regularisation parameter evaluated by solving constrained model. This approach preserves a closed-form solution of each subproblem, and recovers blurry images with simultaneous evaluation of the balance parameter of the unconstrained model. The numerical simulations illustrate the efficiency of the method for constrained ℓ_1-models.

Figure 4: Left: Blurred and noisy images. Middle: Images restored by the model (6.4). Right: Images restored by the model (6.3). Sample ratio: Row 1 – 30%; Row 2 – 40%; Row 3 – 50%; Row 4 – 60%. Noise level – 5%. Blur – I.

Table 4: Numerical results for the models (6.3) and (6.4).

Tested image	Blur	noise level	SNR (dB) via (6.4)	$\hat{\kappa}$	SNR (dB) via (6.3) with $\hat{\kappa}$	λ^{opt}	SNR (dB) via (6.3) with λ^{opt}
salt-and-pepper (noise level: 5%)							
a	I	30%	14.063	38.582636	14.028	67	14.192
		40%	14.885	34.529045	14.808	66	15.070
		50%	15.448	31.441490	15.358	66	15.528
		60%	15.806	28.988864	15.724	63	15.874
		70%	16.227	27.282183	16.150	61	16.291
	II	30%	13.163	32.655956	13.116	67	13.277
		40%	13.635	30.981656	13.595	65	13.759
		50%	13.986	29.048234	13.932	64	14.132
		60%	14.278	27.365302	14.213	62	14.515
		70%	14.569	26.609349	14.507	61	14.731
salt-and-pepper (noise level: 15%)							
a	I	30%	13.423	34.766208	13.377	60	13.595
		40%	14.125	31.727550	14.065	55	14.376
		50%	14.611	27.201878	14.529	55	14.761
		60%	15.060	27.201878	14.986	47	15.229
		70%	15.659	25.884016	15.554	47	15.790
	II	30%	12.700	27.544485	12.671	62	12.938
		40%	13.126	27.701021	13.095	60	13.381
		50%	13.595	26.514622	13.547	61	13.816
		60%	13.802	25.701829	13.748	60	14.016
		70%	14.136	24.938766	14.063	55	14.353

Acknowledgments

The author was supported by the Natural Science Foundation of China (Grant No. NSFC–1130128) and the sponsorship of Jiangsu overseas research and training program for university prominent young and middle-aged teachers and presidents. The author would like to thank the referees and the associate editor to improve the manuscript.

Appendix

Some basic information about image inpainting is provided here, involving the infill of missing or damaged regions in an image — cf. also Ref. [12]. Let $\bar{x}$ be an unknown image. The observed image f after image inpainting is given by

$$f = N\big(S(K\bar{x})\big),$$

where $S \in \mathbb{R}^{n\times n}$ is a mask operator and the corresponding constrained ℓ_1-model (6.4) is similar to that in Refs. [11,13]. If SK is considered the whole matrix K', under Algorithm 4.1 the x-related subproblem does not have a closed-form solution as the matrices on the left-hand side of equation (4.9) are not diagonalisable by a discrete FFT or discrete DCT.

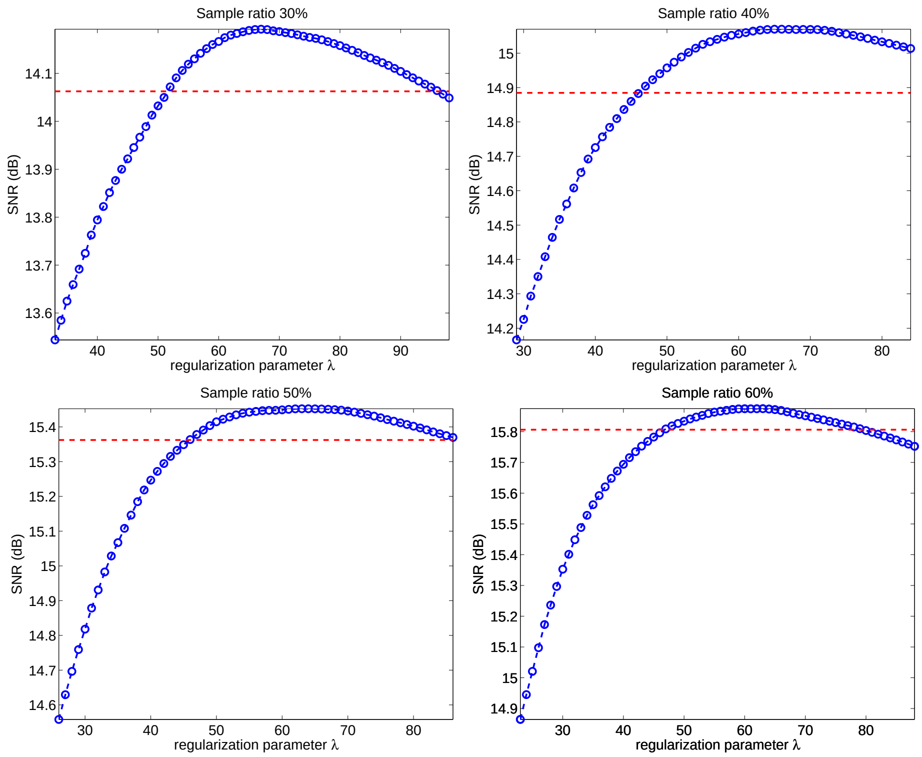

Figure 5: Restored SNR vs. λ, and difference between the restored SNR with $\hat{\kappa}$ and trial-and-error. Sample ratio: Row 1 – 30% and 40%; Row 2 – 50% and 60%. Noise level: 5%. Blur – I.

However, one can rewrite the model (6.4) as

$$
\begin{aligned}
&\min_{x,y,z_i,v,u_1,u_2,w} \sum_{i=1}^{n} \|z_i\| \\
s.t. \quad & \nabla_i x = z_i && : \mu_i \\
& Kx = v && : p_1 \\
& Sv - f = y && : p_2 \\
& y + u_1 - w = 0 && : \nu \\
& y - u_2 + w = 0 && : \delta \\
& e^\top w = \tau && : \kappa \\
& u_1, u_2 \geq 0.
\end{aligned}
$$

The augmented Lagrangian function of the above problem is

$$
\begin{aligned}
&\mathcal{L}_{\mathscr{A}}(x, y, z, v, u_1, u_2, w; \mu, p_1, p_2, \nu, \delta, \kappa) \\
=&\sum_{i=1}^{n} \|z_i\| + \sum_{i=1}^{n} \langle \mu_i, \nabla_i x - z_i \rangle + \frac{\beta_1}{2} \sum_{i=1}^{n} \|\nabla_i x - z_i\|^2
\end{aligned}
$$

Algorithm A.1 ADMM approach for the imaging inpainting model (6.4).

Input f, K, S, $\tau > 0$, β_1, β_2, β_3, β_4, $\beta_5 > 0$ **and** μ^0, p_1^0, p_2^0, ν^0, δ^0, κ^0. **Initialise** $x = f$, $y = f$, $w = f$ **and** $\mu = \mu^0$, $p_1 = p_1^0$, $p_2 = p_2^0$, $\nu = \nu^0$, $\delta = \delta^0$ **and** $\kappa = \kappa^0$.
Output $\hat{x}$, $\hat{\kappa}$.

While "a stopping criterion is not satisfied", **Do**

1) Compute z_i^{k+1} $(i = 1, \dots, n)$ according to
$$z_i^{k+1} = \max\left\{\left\|\nabla_i x^k + \frac{1}{\beta_1}(\mu^k)_i\right\| - \frac{1}{\beta_1}, 0\right\} \frac{\nabla_i x^k + \frac{1}{\beta_1}(\mu^k)_i}{\|\nabla_i x^k + \frac{1}{\beta_1}(\mu^k)_i\|}, \quad i = 1, 2, \cdots, n.$$
2) Compute ν^{k+1} according to
$$\nu^{k+1} = \left(I + \frac{\beta_3}{\beta_2} S^\top S\right)^{-1} \left[K x^k + \frac{1}{\beta_2} p_1^k + \frac{\beta_3}{\beta_2} S^\top (f + y^k) - \frac{1}{\beta_2} S^\top p_2^k\right].$$
3) Compute u_1^{k+1}, and u_2^{k+1} according to
$$u_1^{k+1} = \mathscr{P}_{\mathbb{R}_+^n}\left[w^k - y^k - \frac{\nu^k}{\beta_4}\right],$$
$$u_2^{k+1} = \mathscr{P}_{\mathbb{R}_+^n}\left[w^k + y^k + \frac{\delta^k}{\beta_4}\right].$$
4) Compute x^{k+1} by solving
$$x^{k+1} = (\nabla^\top \nabla + \frac{\beta_2}{\beta_1} K^\top K)^{-1}\left[\nabla^\top (z^{k+1} - \frac{1}{\beta_1}\mu^k) + \frac{1}{\beta_1} K^\top (\beta_2 \nu^{k+1} - p_1^k)\right].$$
5) Compute y^{k+1} by solving
$$y^{k+1} = \frac{1}{\beta_3 + 2\beta_4}\left[\beta_3(S\nu^{k+1} - f) + p_2^k - \nu^k - \delta^k + \beta_4(u_2^{k+1} - u_1^{k+1})\right].$$
6) Compute w^{k+1} by solving
$$w^{k+1} = (2I + \frac{\beta_5}{\beta_4} e e^\top)^{-1}\left[u_1^{k+1} + u_2^{k+1} + \frac{1}{\beta_4}(\nu^k - \delta^k + (\beta_5 \tau - \kappa^k) e)\right].$$
7) Update μ^{k+1}, p_1^{k+1}, p_2^{k+1}, ν^{k+1}, δ^{k+1} and κ^{k+1} via
$$\mu_i^{k+1} = \mu_i^k + \beta_1(\nabla_i x^{k+1} - z_i^{k+1}); \; i = 1, \cdots, n.$$
$$p_1^{k+1} = p_1^k + \beta_2(K x^{k+1} - \nu^{k+1}).$$
$$p_2^{k+1} = p_2^k + \beta_3(S \nu^{k+1} - f - y^{k+1}).$$
$$\nu^{k+1} = \nu^k + \beta_4(y^{k+1} + u_1^{k+1} - w^{k+1}).$$
$$\delta^{k+1} = \delta^k + \beta_4(y^{k+1} - u_2^{k+1} + w^{k+1}).$$
$$\kappa^{k+1} = \kappa^k + \beta_5(e^\top w^{k+1} - \tau).$$

End Do

$$+ \langle p_1, Kx - \nu\rangle + \frac{\beta_2}{2}\|Kx - \nu\|^2 + \langle p_2, S\nu - f - y\rangle + \frac{\beta_3}{2}\|S\nu - f - y\|^2$$
$$+ \langle \nu, y + u_1 - w\rangle + \frac{\beta_4}{2}\|y + u_1 - w\|^2 + \langle \delta, y - u_2 + w\rangle + \frac{\beta_4}{2}\|y - u_2 + w\|^2$$
$$+ \langle \kappa, e^\top w - \tau\rangle + \frac{\beta_5}{2}\|e^\top w - \tau\|^2,$$

where $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5 > 0$. With $\mu = (\mu^{(1)}; \mu^{(2)}) \in \mathbb{R}^{2n}$ and $\mu_i = (\mu_i^{(1)}, \mu_i^{(2)})$ $(i = 1, \cdots, n)$, the approach for the image inpainting model (6.4) is summarised in Algorithm A.1 above.

References

[1] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, Springer-Verlag (2006).

[2] S.D. Babacan, R. Molina and A.K. Katsaggelos, *Parameter estimation in TV image restoration using variational distribution approximation*, IEEE Trans. Image Process. **17**, 326-339 (2008).

[3] S.D. Babacan, R. Molina and A.K. Katsaggelos, *Variational Bayesian blind deconvolution using a total variation prior*, IEEE Trans. Image Process. **18**, 12-26 (2009).

[4] D.P. Bertsekas, *Convex Analysis and Optimization*, Athena Scientific (2003).

[5] A. Bjőrck, *Numerical Methods for Least Squares Problems*, SIAM (1996).

[6] A. Bovik, *Handbook of Image and Video Processing*, Academic Press (2000).

[7] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learning **3**, 1-122 (2010).

[8] M. Carlavan and L. Blanc-Feraud, *Sparse Poisson noisy image deblurring*, IEEE Trans. Image Process. **21**, 1834-1846 (2012).

[9] J.F. Cai, R.H. Chan and M. Nikolova, *Fast two-phase image deblurring under impulse noise*, J. Math. Imaging Vis. **36**, 46-53 (2010).

[10] R.H. Chan and M.K. Ng, *Conjugate gradient method for Toeplitz systems*, SIAM Rev. **38**, 427-482 (1996).

[11] T. Chan and J. Shen, *Mathematical models for local nontexture inpaintings*, SIAM J. Appl. Math. **62**, 1019-1043 (2002).

[12] T. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM (2005).

[13] T. Chan, J. Shen and H. Zhou, *Total variation wavelet inpainting*, J. Math. Imaging Vis. **25**, 107-125 (2006).

[14] R.H. Chan, J.F. Yang and X.M. Yuan, *Alternating direction method for image inpainting in wavelet domains*, SIAM J. Imaging Sci. **4**, 807-826 (2011).

[15] S.Y. Chun, Y.K. Dewaraja and J.A. Fessler, *Alternating direction method of multiplier for tomography with non-local regularisers*, IEEE Trans. Med. Imaging **33**, 1960-1968 (2014).

[16] Y.Q. Dong, R. Chan and S.F. Xu, *A detection statistic for random-valued impulse noise*, IEEE Trans. Image Process. **16**, 1112-1120 (2007).

[17] H. Fu, M. Ng, M. Nikolova and J. Barlow, *Efficient minimization methods of mixed ℓ_2-ℓ_1 and ℓ_1-ℓ_1 norms for image restoration*, SIAM J. Sci. Comput. **27**, 1881-1902 (2006).

[18] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite-element approximations*, Comput. Math. Appl. **2**, 17-40 (1976).

[19] N.P. Galatsanos and A.K. Katsaggelos, *Methods for choosing the regularisation parameter and estimating the noise variance in image restoration and their relation*, IEEE Trans. Image Process. **1**, 322-336 (1992).

[20] G.H. Gloub, M. Heath and G. Wahba, *Generalized cross-validation as method for choosing a good ridge parameter*, Technometrics **2**, 215-223 (1979).

[21] G. Gilboa and S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Sim. **7**, 1005-1028 (2008).

[22] R. Glowinski and A. Marrocco, *Sur l'approximation par éléments finis d'ordre un et la résolution par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires*, Revue Fr. Autom. Inform. Rech. Opér., Anal. Numér. **2**, 41-76 (1975).

[23] R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley (1992).

[24] P.C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev. **34**, 287-301 (1992).

[25] P.C. Hansen, M.E. Kilmer and R.H. Kjeldsen, *Exploiting residual information in the parameter choice for discrete ill-posed problems*, BIT Numer. Math. **46**, 41-59 (2006).

[26] H. Hwang, R.A. Haddad, *Adaptive median filters: New algorithms and results*, IEEE Trans. Image. Process. **4**, 499-502 (1995).

[27] J.F.C. Mota, J.M.F. Xavier, P.M.Q. Aguiar and M. Pűschel, *A proof of convergence for the alternating direction method of multipliers applied to polyhedral-constrained functions*, ArXiv **1112.2295** (2011).

[28] J.T. Kent and M. Mohammadzadeh, *Global optimization of the generalized cross-validation criterion*, Stat. Comput. **10**, 231-236 (2000).

[29] C. Lawson and R. Hanson, *Solving Least Squares Problems*, Classics Appl. Math. **15**, Prentice-Hall (1974).

[30] H. Liao, F. Li and M. Ng, *Selection of regularisation parameter in total variation image restoration*, J. Opt. Soc. Amer. A **26**, 2311-2320 (2009).

[31] V.A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer (1984).

[32] M.K. Ng, R.H. Chan and W.C. Tang, *A fast algorithm for deblurring models with Neumann boundary conditions*, SIAM J. Sci. Comput. **21**, 851-866 (1999).

[33] M.K. Ng, P. Weiss and X.M. Yuan, *Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods*, SIAM J. Sci. Comput. **32**, 2710-2736 (2010).

[34] J.P. Oliveira, J.M. Bioucas-Dias and M.A.T. Figueiredo, *Adaptive total variation image deblurring: A majorization-minimization approach*, Signal Process. **89**, 1683-1693 (2009).

[35] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D **60**, 259-268 (1992).

[36] B.W. Rust and D.P. OĹeary, *Residual periodograms for choosing regularisation parameters for ill-posed problems*, Inverse Probl. **24**, 034005 (2008).

[37] S. Setzer, G. Steidl and T. Teuber, *On vector and matrix median computation*, J. Comput. Appl. Math. **236**, 2200-2222 (2012).

[38] Y. Xiao, J. Yang and X. Yuan, *Alternating algorithms for total variation image reconstruction from random projections*, Inverse Probl. Imaging **6**, 547-563 (2012).

[39] J. Yang, W. Yin, Y. Zhang and Y. Wang, *A fast algorithm for edge-preserving variational multichannel image restoration*, SIAM J. Imaging Sci. **2**, 569-592 (2009).

[40] J. Yang, Y. Zhang and W. Yin, *An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise*, SIAM J. Sci. Comput. **31**, 2842-2865 (2009).

[41] T. Teuber, G. Steidl and R. Chan, *Minimization and parameter estimation for seminorm regularisation models with I-divergence constraints*, Inverse Probl. **29**, 035007 (2013).

[42] C.R. Vogel and M.E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput. **17**, 227-238 (1996).

[43] Y. Wang, J. Yang, W. Yin and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sci. **1**, 948-951 (2008).

[44] C. Wu, J. Zhang and X. Tai, *Augmented Lagrangian method for total variation restoration with non-quadratic fidelity*, Inverse Probl. Imaging **5**, 237-261 (2011).

[45] P. Weiss, L. Blance-Feraud and G. Aubert, *Efficient schemes for total variation minimization under constraints in image processing*, SIAM J. Sci. Comput. **31**, 2047-2080 (2009).

[46] P. Weiss, G. Aubert and L. Blanc-Féraud, *Some applications of ℓ^∞-constraints in image processing*, INRIA Research Report **6115** (2006).

[47] Y.W. Wen and R. Chan, *Parameter selection for total variation based restoration using discrepancy principle*, IEEE Trans. Image Process. **21**, 1770-1781 (2012).