Computational Optimization and Applications (2025) 91:759–785
https://doi.org/10.1007/s10589-025-00659-y

A Scaled Gradient Projection method for the realization of the balancing principle in TGV-based image restoration

Germana Landi[1] · **Marco Viola**[2] · **Fabiana Zama**[1]

Received: 1 April 2024 / Accepted: 23 January 2025 / Published online: 11 February 2025

Abstract

In the last few years, Total Generalized Variation (TGV) regularization has proved to be a valuable tool to remove blur and noise from an image while avoiding the staircase effect typical of the Total Variation (TV) and preserving the sharp edges. The TGV-regularized model depends on two regularization parameters whose values must be appropriately selected to obtain good-quality restored images. In this work, we propose the use of the Balancing Principle (BP) to formulate the TGV-based image restoration problem as a constrained minimization problem whose objective is an implicit function of the two regularization parameters depending on the image to be restored. The values of the regularization parameters, and the corresponding restored image, satisfying the optimality condition of the formulated problem guarantee that the data fidelity and regularization terms are balanced. We introduce a Scaled Gradient Projection (SGP) method specifically tailored to the BP-based optimization problem and test its effectiveness against the fixed-point iteration schemes proposed in the literature. The numerical results performed on real-life images, affected by both Gaussian and Poisson noise, show that the proposed approach can effectively restore input images corrupted by several kinds of noise and outperform the fixed-point strategies for the realization of the Balancing Principle.

Keywords Image restoration · Balancing principle · TGV regularization · Automatic regularization parameter estimation

Germana Landi, Marco Viola, and Fabiana Zama contributed equally to this work.

✉ Germana Landi
germana.landi@unibo.it

Marco Viola
marco.viola@dcu.ie

Fabiana Zama
fabiana.zama@unibo.it

[1] Department of Mathematics, University of Bologna, P.zza Porta S. Donato, 40127 Bologna, Italy

[2] School of Mathematical Sciences, Dublin City University, Collins Avenue Extension, Dublin D09V209, Ireland

Mathematics Subject Classification 90C25 · 65K05 · 94A08

1 Introduction

Image restoration plays a crucial role across various fields of applied sciences, as digital images often suffer from blur and noise during their acquisition process. Mathematically, image formation is usually formulated through a linear operator, i.e.,

$$A\mathbf{u} = \mathbf{b} \tag{1}$$

where $\mathbf{u} \in \mathbb{R}^n$ is the unknown image we aim to restore from the observed data $\mathbf{b} \in \mathbb{R}^n$, with $n = n_1 n_2$ representing the total number of pixels in an image of dimensions $n_1 \times n_2$. Here, the matrix A describes the discretized Point Spread Function (PSF) of the image acquisition system, which introduces blur artifacts and is known to be ill-conditioned. In image restoration, the measured data $\mathbf{b}$ are further compromised by noise, which relates to the blurred image according to the specific noise model employed. In a variational framework, the image restoration problem is formulated as an inverse linear problem of the form

$$\min_{\mathbf{u}\in\Omega} \left\{ \frac{1}{\lambda}\phi(\mathbf{u}) + \mathcal{R}(\mathbf{u}) \right\},$$

where $\phi : \Omega \to \mathbb{R}^+$ denotes the data fidelity function dependent on the discrete PSF A and the noise type on the data; $\lambda \in \mathbb{R}^+$ is the regularization parameter, whose value significantly influences the quality of the restored image, and Ω is a compact subset of $\mathbb{R}^n$ used to impose some physical constraints. $\mathcal{R}$ represents the regularization functional, accounting for desired properties in the solution (see [1] and references therein).

This work focuses on a specific regularization functional: Total Generalized Variation (TGV), introduced by Bredies et al. [2, 3]. TGV is an effective tool for image restoration, capable of preserving edges while effectively removing noise. However, its main challenge lies in its dependence on two positive weights α_0, α_1, crucial for balancing data fidelity and regularization. The automatic determination of these parameters remains a difficult task, and a common approach in the literature consists of setting them to some previously fixed values and adapting the regularization parameter λ to obtain satisfactory reconstructions.

A small amount of literature on TGV for image restoration has explored methodologies for estimating the TGV weight parameters α_0 and α_1. The work by Hintermüller et al. [4] introduces a bilevel optimization framework for automatically computing spatially dependent regularization parameters. The strategy proposed by di Serafino et al. [5] simplifies parameter selection to a convex combination of two parameters, effectively reducing the problem to a single-parameter setting, and proposes an automatic strategy for the selection of λ based on the Balancing Principle (BP) developed by Ito et al. [6]. Moreover, di Serafino and Pragliola proposed a hierarchical Bayesian method tailored for image restoration under Poisson noise [7] based on the work in

[8]. A somehow related work is [9], where the authors propose to fix the regularization parameters in a TGV-regularized restoration problem and adaptively choose the penalty parameters in the ADMM method they set to solve the problem.

In this work, we extend the approach developed in [5] for TGV-based restoration of Poissonian images. While in [5], the authors fix an a-priori value for the parameters α_0 and α_1 and focus on the determination of λ, we investigate the selection of the parameters α_0 and α_1 when λ has been fixed (e.g., to 1). This will allow for greater flexibility in determining the two separate terms of the TGV functional.

In detail, this paper considers a perspective of TGV as a multi-penalty regularization functional. It introduces an iterative method that computes a sequence of restored images along with their regularization parameters, exploiting the Balancing Principle [10–12] for multi-parameter setting. This approach has shown efficacy in algorithms based on L_1-L_2 and pixel-wise Tikhonov regularization [12, 13]. Following the framework in [10], this paper characterizes the TGV regularization parameters as the minimizers of the cost functional related to the TGV-penalized model and the product of the regularization parameters. We propose to solve the optimization problem through Scaled Gradient Projection iteration [14] identifying an appropriate scaling matrix. The key contributions of this paper are outlined as follows:

- development of an automatic strategy for computing regularization parameters based on the Balancing Principle;
- proposal of a Scaled Gradient Projection algorithm with a new scaling matrix for efficient computation of these parameters.

Numerical results validate the method's effectiveness and efficiency for image restoration problems with Gaussian and Poisson noise. Compared to the fixed-point methods proposed in the literature, the proposed approach produces better-restored images without requiring an increased computational effort.

The paper is organized as follows. Section 2 presents the balancing principle, adapting it to the structure of TGV and analyzing its theoretical properties. Then, Sect. 3 addresses the numerical implementation of the Balancing Principle, the fixed-point methods used for comparisons, and the novel Scaled Gradient Projection method proposed in this paper. Section 4 validates the proposed method by comparing it with fixed-point approaches through numerical tests on images corrupted by Gaussian and Poisson noise. Finally, we draw some conclusions in Sect. 5.

2 The balancing principle for TGV-based image restoration

Given an image $\mathbf{u} \in \mathbb{R}^n$, the discrete TGV function of order two with weights α_0, α_1 is defined as the function

$$\mathrm{TGV}^2(\mathbf{u}) = \min_{\mathbf{w}\in\mathbb{R}^{2n}} \alpha_0 \|\nabla\mathbf{u} - \mathbf{w}\|_{2,1|\mathbb{R}^{2n}} + \alpha_1 \|\mathcal{E}\mathbf{w}\|_{2,1|\mathbb{R}^{4n}}, \tag{2}$$

where $\mathbf{w} \in \mathbb{R}^{2n}$ is an auxiliary variable $\alpha_0, \alpha_1 \in (0,\ 1)$ are positive parameters balancing the two components of the regularization. The linear operators $\nabla \in \mathbb{R}^{2n\times n}$ and $\mathcal{E} \in \mathbb{R}^{4n\times 2n}$ represent the *discrete gradient operator* and the *symmetrized derivative*

operator, respectively, which are described below. Note that for any $\mathbf{v} \in \mathbb{R}^{2n}$ and $\mathbf{y} \in \mathbb{R}^{4n}$ we define

$$\|\mathbf{v}\|_{2,1|\mathbb{R}^{2n}} = \sum_{j=1}^{n} \sqrt{v_j^2 + v_{n+j}^2}, \tag{3}$$

and

$$\|\mathbf{y}\|_{2,1|\mathbb{R}^{4n}} = \sum_{j=1}^{n} \sqrt{y_j^2 + y_{n+j}^2 + y_{2n+j}^2 + y_{3n+j}^2}. \tag{4}$$

Let $D_H, D_V \in \mathbb{R}^{n\times n}$ represent the forward finite-difference operators along the horizontal and the vertical direction, respectively. The discrete gradient operator and the discrete symmetrized derivative operator are defined as

$$\nabla = \begin{bmatrix} D_H \\ D_V \end{bmatrix}, \qquad \mathcal{E} = \begin{bmatrix} D_H & 0 \\ \frac{1}{2}D_V & \frac{1}{2}D_H \\ \frac{1}{2}D_V & \frac{1}{2}D_H \\ 0 & D_V \end{bmatrix}.$$

In a variational framework, the TGV-based image restoration problem can be formulated as the nonsmooth optimization problem

$$\begin{aligned} &\min_{\mathbf{u},\mathbf{w}} \frac{1}{\lambda}\phi(\mathbf{u}) + \alpha_0 \|\nabla\mathbf{u} - \mathbf{w}\|_{2,1|\mathbb{R}^{2n}} + \alpha_1 \|\mathcal{E}\mathbf{w}\|_{2,1|\mathbb{R}^{4n}} \\ &\text{s.t.} \quad \mathbf{u} \geq 0, \end{aligned} \tag{5}$$

where $\phi(\mathbf{u})$ is the data fidelity term depending on the noise type and the parameter $\lambda \in (0, +\infty)$ regulates the relative weight between $\phi(\mathbf{u})$ and the two terms composing the TGV regularization. By following the maximum-likelihood approach, the term $\phi(\mathbf{u})$ is expressed as the L_2-norm

$$\phi(\mathbf{u}) = \|A\mathbf{u} - \mathbf{b}\|_2^2, \tag{6}$$

when the measured data is affected by Gaussian noise or by the Kullback-Leibler (KL) divergence

$$\phi(\mathbf{u}) = \sum_{i=1}^{n} \left(b_i \ln \frac{b_i}{[A\mathbf{u} + \boldsymbol{\nu}]_i} + [A\mathbf{u} + \boldsymbol{\nu}]_i - b_i \right), \tag{7}$$

where $\boldsymbol{\nu}$ represents the background radiation, for Poisson noise.

By introducing the auxiliary variable $\mathbf{x} = (\mathbf{u}, \mathbf{w}) \in \mathbb{R}^{3n}$, with the small abuse of notation $\phi(\mathbf{x}) \equiv \phi(\mathbf{u})$, we can reformulate problem (5) as

$$\begin{aligned} &\min_{\mathbf{x}} \phi(\mathbf{x}) + \eta_1 \psi_1(\mathbf{x}) + \eta_2 \psi_2(\mathbf{x}), \\ &\text{s.t.} \quad \mathbf{x} \in \Omega, \end{aligned} \tag{8}$$

where we set $\eta_1 = \lambda\alpha_0$, $\eta_2 = \lambda\alpha_1$ and we define the regularization functions as

$$\psi_1(\mathbf{x}) = \left\| \left[\nabla \; -I_{2n} \right] \mathbf{x} \right\|_{2,1|\mathbb{R}^{2n}}, \qquad \psi_2(\mathbf{x}) = \left\| \left[0_{4n \times n} \; \mathcal{E} \right] \mathbf{x} \right\|_{2,1|\mathbb{R}^{4n}} \tag{9}$$

and the feasible set as

$$\Omega \; = \; \{\mathbf{x} \mid \left[I_n \; 0_{n \times 2n} \right] \mathbf{x} \geq 0\}$$

with $I_m \in \mathbb{R}^{m \times m}$ identity matrix of size m and $0_{m \times l} \in \mathbb{R}^{m \times l}$ null matrix of size $m \times l$.

Thus, the TGV-based image restoration problem is reformulated as a regularization problem with multiple penalties. Here, we propose using the balancing principle as a criterion for choosing the regularization parameters η_1 and η_2.

This principle determines the regularization parameters η_1 and η_2 and the corresponding regularized solution $\mathbf{x}_{\boldsymbol{\eta}}$ such that

$$\begin{cases} \mathbf{x}_{\boldsymbol{\eta}} = \arg\min\limits_{\mathbf{x} \in \Omega} \{\phi(\mathbf{x}) + \eta_1 \psi_1(\mathbf{x}) + \eta_2 \psi_2(\mathbf{x})\}, \\ \eta_i = \dfrac{\phi(\mathbf{x}_{\boldsymbol{\eta}})}{\gamma \psi_i(\mathbf{x}_{\boldsymbol{\eta}})}, \quad i = 1, 2, \end{cases} \tag{10}$$

where γ is a positive parameter and $\boldsymbol{\eta} = (\eta_1, \eta_2)^T$. It is evident that, according to the BP, the optimal regularization parameters η_1 and η_2 are obtained by balancing, up to the multiplicative parameter γ, the data-fidelity term with the regularization functions ψ_i, $i = 1, 2$. The balancing principle has been formulated in [10, 11] for multi-parameter regularization under the hypothesis that the data fidelity term ϕ and the penalties ψ_i, $i = 1, 2$, are nonnegative, convex and weak lower semicontinuous; we remark that these hypotheses are satisfied by the TGV regularization and, e.g., by the standard data fidelities used for Gaussian, Poisson, or Laplace noise.

In [10–12], a variational characterization of the BP is given by introducing the function $\Phi_\gamma(\boldsymbol{\eta})$ defined as

$$\Phi_\gamma(\boldsymbol{\eta}) = \frac{F^{\gamma+2}(\boldsymbol{\eta})}{\eta_1 \eta_2}, \tag{11}$$

where $F(\boldsymbol{\eta})$ is the value function

$$F(\boldsymbol{\eta}) = \min_{\mathbf{x} \in \Omega} \{\phi(\mathbf{x}) + \eta_1 \psi_1(\mathbf{x}) + \eta_2 \psi_2(\mathbf{x})\}. \tag{12}$$

Remark 1 It is important to observe that the function $F(\boldsymbol{\eta})$ can only be zero when, fixed a pair of positive regularization parameters, the solution to the TGV-regularized restoration problem lies in the intersection between the kernel of the TGV operator and the kernel of the data fidelity term. We observe that the kernel of the TGV operator

consists of second order polynomials [2]. Hence, since most real-life images are not second order polynomials, we will assume that second order polynomials do not nullify the data fidelity term, which yields that $F(\boldsymbol{\eta}) > 0$ for every $\boldsymbol{\eta} > 0$.

The following proposition, proved in [11, 12], shows that the solutions of (10) coincide with the critical points of the function $\Phi_\gamma(\boldsymbol{\eta})$.

Proposition 2.1 *If the value function $F(\boldsymbol{\eta})$ is differentiable, all the critical points of the function Φ_γ are solutions to the BP system* (10).

The function $F(\boldsymbol{\eta})$ is a continuous function whose partial derivatives $\partial_i F$, $i = 1, 2$, exist and are continuous if, for any choice of $\boldsymbol{\eta} > 0$, it holds

$$\psi_i(\hat{\mathbf{x}}_{\boldsymbol{\eta}}) = \psi_i(\tilde{\mathbf{x}}_{\boldsymbol{\eta}}), \quad \text{for all } \hat{\mathbf{x}}_{\boldsymbol{\eta}}, \tilde{\mathbf{x}}_{\boldsymbol{\eta}} \in \mathcal{M}_{\boldsymbol{\eta}}, \tag{13}$$

where $\mathcal{M}_{\boldsymbol{\eta}}$ denotes the set of all the minimizers $\mathbf{x}_{\boldsymbol{\eta}}$ to $\phi + \eta_1\psi_1 + \eta_2\psi_2$. In particular, the latter holds whenever the minimizer of the function $\phi + \eta_1\psi_1 + \eta_2\psi_2$ is unique. It is worth mentioning that in the case of TGV regularized problems, the latter is satisfied if the linear operator A in (1) is injective [3]. Alternatively, this can be guaranteed by fostering the strong convexity of the data fidelity function ϕ with the addition of a term of the form $\sigma\|\mathbf{u}\|^2$.

3 Numerical realization of the balancing principle

In [10, 11], two fixed-point methods have been proposed for the numerical realization of the BP. We revise such iterative schemes in this section and propose our SGP method.

3.1 Fixed-point methods

The first method proposed in [10, 11] is a fixed-point iteration for computing a critical point of $\Phi_\gamma(\boldsymbol{\eta})$. The first-order derivatives of function $\Phi_\gamma(\boldsymbol{\eta})$ are given by

$$\nabla\Phi_\gamma(\boldsymbol{\eta}) = \frac{F^{1+\gamma}(\boldsymbol{\eta})}{\eta_1\eta_2} \begin{bmatrix} (2+\gamma)\psi_1(\mathbf{x}_{\boldsymbol{\eta}}) - \dfrac{F(\boldsymbol{\eta})}{\eta_1} \\ (2+\gamma)\psi_2(\mathbf{x}_{\boldsymbol{\eta}}) - \dfrac{F(\boldsymbol{\eta})}{\eta_2} \end{bmatrix}. \tag{14}$$

Since $F(\boldsymbol{\eta}) > 0$, critical points for $\Phi_\gamma(\boldsymbol{\eta})$ must satisfy

$$(2+\gamma)\psi_i(\mathbf{x}_{\boldsymbol{\eta}}) - \frac{\phi(\mathbf{x}_{\boldsymbol{\eta}}) + \eta_i\psi_i(\mathbf{x}_{\boldsymbol{\eta}}) + \eta_{-i}\psi_{-i}(\mathbf{x}_{\boldsymbol{\eta}})}{\eta_i} = 0, \quad i = 1, 2, \tag{15}$$

where $-i$ denotes the index that is not i. From (15) we get

$$\eta_i = \frac{\phi(\mathbf{x}_{\boldsymbol{\eta}}) + \eta_{-i}\psi_{-i}(\mathbf{x}_{\boldsymbol{\eta}})}{(1+\gamma)\psi_i(\mathbf{x}_{\boldsymbol{\eta}})}, \quad i = 1, 2, \tag{16}$$

which gives the fixed-point method outlined in Algorithm 1 and referred to as FP1. Convergence of FP1 was proved in [11, Proposition 3.1] under assumptions on the value function F and on the functions ψ_i that, however, are difficult to verify in practice.

Algorithm 1 Fixed-point method 1 (FP1)

1: Choose $\gamma > 0$, $\eta_1^{(0)}, \eta_2^{(0)} > 0$, and set $k = 0$;
2: **repeat**
3: $\quad \mathbf{x}_{\boldsymbol{\eta}}^{(k)} = \arg\min_{\mathbf{x}\in\Omega} \{\phi(\mathbf{x}) + \eta_1^{(k)}\psi_1(\mathbf{x}) + \eta_2^{(k)}\psi_2(\mathbf{x})\}$;
4: $\quad \eta_i^{(k+1)} = \dfrac{\phi(\mathbf{x}_{\boldsymbol{\eta}}^{(k)}) + \eta_{-i}\psi_{-i}(\mathbf{x}_{\boldsymbol{\eta}}^{(k)})}{(1+\gamma)\psi_i(\mathbf{x}_{\boldsymbol{\eta}}^{(k)})}$, $i = 1, 2$;
5: $\quad k = k + 1$;
6: **until** a stopping criterion is satisfied

The second algorithm proposed in [10], which we will refer to as FP2, is sketched in Algorithm 2 and consists of applying fixed-point iterations to the nonlinear equations (10) characterizing the BP. Convergence of FP2 is experimentally observed in [10] and proved in [12] where FP2 is shown to be equivalent to a Majorization–Minimization method [15] for minimizing $\Phi_\gamma(\boldsymbol{\eta})$, for a suitable choice of the surrogate function.

Algorithm 2 Fixed-point method 2 (FP2)

1: Choose $\gamma > 0$, $\eta_1^{(0)}, \eta_2^{(0)} > 0$, and set $k = 0$.
2: **repeat**
3: $\quad \mathbf{x}_{\boldsymbol{\eta}}^{(k)} = \arg\min_{\mathbf{x}\in\Omega} \{\phi(\mathbf{x}) + \eta_1^{(k)}\psi_1(\mathbf{x}) + \eta_2^{(k)}\psi_2(\mathbf{x})\}$;
4: $\quad \eta_i^{(k+1)} = \dfrac{\phi(\mathbf{x}_{\boldsymbol{\eta}}^{(k)})}{\gamma\,\psi_i(\mathbf{x}_{\boldsymbol{\eta}}^{(k)})}$, $i = 1, 2$;
5: $\quad k = k + 1$;
6: **until** a stopping criterion is satisfied

Observe that the iterations of FP1 and FP2 methods guarantee that the computed regularization parameters remain always positive.

3.2 Scaled Gradient Projection method

In this work, we address the computation of the regularization parameters η_1 and η_2, together with the corresponding regularized solution $\mathbf{x}_{\boldsymbol{\eta}}$, as the research for a local minimizer of $\Phi_\gamma(\boldsymbol{\eta})$.

Before introducing the proposed method, we observe that the problem of minimizing the nonconvex functional $\Phi_\gamma(\boldsymbol{\eta})$ in the positive orthant, i.e.,

$$\min_{\boldsymbol{\eta}>0} \quad \Phi_\gamma(\boldsymbol{\eta}), \tag{17}$$

is well defined. Furthermore, since $\Phi_\gamma(\boldsymbol{\eta}) \to +\infty$ whenever $\eta_1 \to 0$ or $\eta_2 \to 0$, by choosing a sufficiently small $\varepsilon > 0$, the constrained optimization problem

$$\min_{\boldsymbol{\eta} \geq \varepsilon} \quad \Phi_\gamma(\boldsymbol{\eta}), \tag{18}$$

is well defined and its stationary points are unconstrained stationary points for $\Phi_\gamma(\boldsymbol{\eta})$, thus realizing the Balancing Principle.

We here propose to solve problem (18) by means of a Scaled Gradient Projection (SGP) method, described by the general iteration

$$\boldsymbol{\eta}^{(k+1)} \;=\; \mathcal{P}_\varepsilon \left(\boldsymbol{\eta}^{(k)} - t^{(k)} D_k^{-1} \nabla \Phi_\gamma(\boldsymbol{\eta}^{(k)}) \right), \tag{19}$$

where D_k is a diagonal positive definite scaling matrix, $\mathcal{P}_\varepsilon(\cdot)$ denotes the projection operator onto the polytope $\{\boldsymbol{\eta} \geq \varepsilon\}$, and $t^{(k)}$ is a positive step-length parameter satisfying the condition

$$\Phi_\gamma(\boldsymbol{\eta}^{(k+1)}) \;\leq\; \Phi_\gamma(\boldsymbol{\eta}^{(k)}) + \mu \nabla \Phi(\boldsymbol{\eta}^{(k)})^\top \left(\boldsymbol{\eta}^{(k+1)} - \boldsymbol{\eta}^{(k)} \right), \tag{20}$$

with $0 < \mu < 1$. In the following, we will omit the superscripts/subscripts associated with the iteration number to lighten up the notation.

Remark 2 Since the objective function is well defined for strictly positive entries and it tends to infinity at the boundaries of the nonnegative orthant, ε only serves as a safeguard in case a line-search direction points outside the orthant. Indeed in this case one can guarantee that the point resulting from the line-search is strictly positive. In practice, ε can be arbitrarily small and our experiments suggest that for $\varepsilon < 10^{-3}$ the behavior of the proposed SGP method does not change.

Various choices have been proposed in literature for the SGP scaling matrix in the case of bound constrained problems in image restoration. Though the most common ones are based on the inverse of the current iterate norm, the most effective ones are based on particular splittings of the objective function [16]. However, in the specific case of the balancing principle the latter strategies cannot be applied. In this work we propose a tailored strategy for the selection of D, which is taken as

$$D^{-1} = \frac{\eta_1 \eta_2}{F^{1+\gamma}(\boldsymbol{\eta})} \begin{bmatrix} \dfrac{\eta_1}{(1+\gamma)\psi_1(\mathbf{x}_{\boldsymbol{\eta}})} & 0 \\ 0 & \dfrac{\eta_2}{(1+\gamma)\psi_2(\mathbf{x}_{\boldsymbol{\eta}})} \end{bmatrix} \tag{21}$$

which, by (14), gives

$$D^{-1} \nabla \Phi_\gamma(\boldsymbol{\eta}) = \begin{bmatrix} \dfrac{\eta_1}{(1+\gamma)\psi_1(\mathbf{x}_{\boldsymbol{\eta}})} & 0 \\ 0 & \dfrac{\eta_2}{(1+\gamma)\psi_2(\mathbf{x}_{\boldsymbol{\eta}})} \end{bmatrix} \begin{bmatrix} (2+\gamma)\psi_1(\mathbf{x}_{\boldsymbol{\eta}}) - \dfrac{F(\boldsymbol{\eta})}{\eta_1} \\ (2+\gamma)\psi_2(\mathbf{x}_{\boldsymbol{\eta}}) - \dfrac{F(\boldsymbol{\eta})}{\eta_2} \end{bmatrix}.$$

From

$$\begin{aligned}
&\frac{\eta_i}{(1+\gamma)\psi_i(\boldsymbol{\eta})}\left(\frac{(2+\gamma)\eta_i\psi_i(\boldsymbol{\eta})-F(\boldsymbol{\eta})}{\eta_i}\right)\\
&=\frac{\eta_i}{(1+\gamma)\psi_i(\boldsymbol{\eta})}\left(\frac{(1+\gamma)\eta_i\psi_i(\boldsymbol{\eta})-\phi(\boldsymbol{\eta})-\eta_{-i}\psi_{-i}(\boldsymbol{\eta})}{\eta_i}\right)\\
&=\eta_i-\frac{\phi(\boldsymbol{\eta})+\eta_{-i}\psi_{-i}(\boldsymbol{\eta})}{(1+\gamma)\psi_i(\boldsymbol{\eta})},\quad i=1,2,
\end{aligned}\tag{22}$$

we get that

$$D^{-1}\nabla\Phi_\gamma(\boldsymbol{\eta})=\boldsymbol{\eta}-\mathbf{d}$$

where

$$\mathbf{d}=\begin{bmatrix}\dfrac{\phi(\boldsymbol{\eta})+\eta_2\psi_2(\boldsymbol{\eta})}{(1+\gamma)\psi_1(\boldsymbol{\eta})}\\[2ex] \dfrac{\phi(\boldsymbol{\eta})+\eta_1\psi_1(\boldsymbol{\eta})}{(1+\gamma)\psi_2(\boldsymbol{\eta})}\end{bmatrix}.\tag{23}$$

Thus, the proposed scaling produces a search direction $\mathbf{d}$ which is equal to the fixed-point operator for FP1 (Algorithm 1).

With this choice, the SGP iteration (19) becomes:

$$\boldsymbol{\eta}^{(k+1)}\;=\;\mathcal{P}_\varepsilon\left((1-t^{(k)})\boldsymbol{\eta}^{(k)}+t^{(k)}\mathbf{d}^k\right).\tag{24}$$

Regarding the choice, for each k, of the step-length $t^{(k)}$ satisfying the Armijo condition (20), we resort to a backtracking strategy, i.e., we set

$$t^{(k)}=\rho^{n_k}\,t_0^{(k)},\tag{25}$$

where $0<\rho<1$, $t_0^{(k)}>0$, and n_k is the minimum integer such that $t^{(k)}$ satisfies (20). In detail, we set $t_0^{(k)}=1$ for the first two iterations and then use the short Barzilai-Borwein steplength (referred to as BB2) [17], i.e., we set

$$t_0^{(k)}=\begin{cases}1, & k<2,\\ \min\left\{t_{\min},\ \frac{(\mathbf{s}^{(k-1)})^\top\mathbf{y}^{(k-1)}}{\|\mathbf{y}^{(k-1)}\|^2},\ t_{\max}\right\}, & k\geq 2,\end{cases}\tag{26}$$

where $\mathbf{s}^{(k-1)}=\boldsymbol{\eta}^k-\boldsymbol{\eta}^{(k-1)}$, $\mathbf{y}^{(k-1)}=\nabla\Phi_\gamma(\boldsymbol{\eta}^k)-\nabla\Phi_\gamma(\boldsymbol{\eta}^{(k-1)})$, and $0<t_{\min}<t_{\max}$.

The resulting line-search SGP method is outlined in Algorithm 3.

Algorithm 3 Scaled Gradient Projection method (SGP)

1: Choose $\gamma > 0$, $\eta_1^{(0)}, \eta_2^{(0)} > 0$, $\mu \in (0,1)$, $0 < t_{\min} < t_{\max}$, and set $k = 0$;
2: **repeat**
3: $\mathbf{x}_{\boldsymbol{\eta}}^{(k)} = \arg\min_{\mathbf{x}\in\Omega} \{\phi(\mathbf{x}) + \eta_1^{(k)}\psi_1(\mathbf{x}) + \eta_2^{(k)}\psi_2(\mathbf{x})\}$;
4: compute $\mathbf{d}^k$ as in (23) and $t_0^{(k)}$ as in (26);
5: Set $\boldsymbol{\eta}^{(k+1)} = \mathcal{P}_\varepsilon\left((1-t^{(k)})\boldsymbol{\eta}^{(k)} + t^{(k)}\mathbf{d}^k\right)$, with $t^{(k)}$ as in (25) satisfying the Armijo condition (20);
6: $k = k + 1$;
7: **until** a stopping criterion is satisfied

3.3 Solution of the TGV-based subproblem

At each iteration of the FP1, FP2 and SGP methods, the constrained minimization problem

$$\begin{aligned} &\min_{\mathbf{u},\mathbf{w}} \phi(\mathbf{u}) + \eta_1 \|\nabla\mathbf{u} - \mathbf{w}\|_{2,1|\mathbb{R}^{2n}} + \eta_2 \|\mathcal{E}\mathbf{w}\|_{2,1|\mathbb{R}^{4n}} \\ &\text{s.t.} \quad \mathbf{u} \geq 0, \end{aligned} \tag{27}$$

has to be solved (we omitted the index k for easier notation). To this end, we use the Alternating Direction Method of Multipliers (ADMM) developed and analyzed in [18] to restore images using Directional TGV regularization. This is a two-block ADMM method where each subproblem can be solved exactly, at a low computational cost, when the linear operator A is a Block Circulant with Circulant Blocks (BCCB) matrix. The method, called "ResPoND",[1] was originally designed for images corrupted by Poisson noise (i.e., for ϕ equal to the Kullback-Leibler divergence in (7)). For the experiments presented in Sect. 4, we implemented an adaptation of the method to the case of Gaussian noise (i.e., with ϕ as in (6)).

3.4 Selection of the parameter γ

The BP depends on the parameter γ, whose value influences the quality of the restored images. In order to select a reliable value for γ, we use the uniform penalty principle [12] which suggests to choose γ so that

1. the regularization terms $\eta_1\psi_1(\mathbf{x})$ and $\eta_2\psi_2(\mathbf{x})$ are uniform and equal to a constant value $c > 0$, i.e:

$$\eta_1\psi_1(\mathbf{x}) = \eta_2\psi_2(\mathbf{x}) = c; \tag{28}$$

2. the data fidelity term is equal to the regularization term, i.e.:

$$\phi(\mathbf{x}) = \eta_1\psi_1(\mathbf{x}) + \eta_2\psi_2(\mathbf{x}). \tag{29}$$

[1] MATLAB implementation available on GitHub at https://github.com/diserafi/respond.

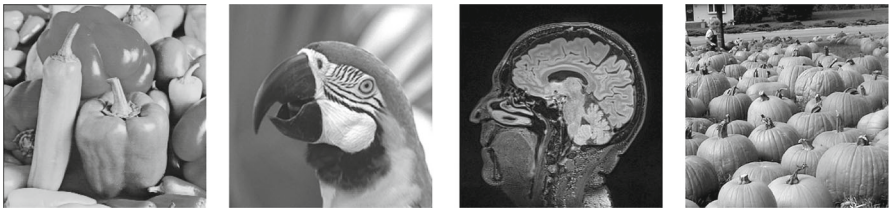

Fig. 1 The four reference images, from left to right: `peppers`, `parrot`, `mri_head` and `pumpkins`

Conditions (28) and (29) give

$$c = \frac{1}{2}\phi(\mathbf{x}) \quad \text{and} \quad \eta_i = \frac{\phi(\mathbf{x})}{2\psi_i(\mathbf{x})}, \ i = 1, 2,$$

from which we obtain $\gamma = 2$. In the more general case of multi penalty regularization with p penalty terms, the uniform penalty principle gives $\gamma = p$ [12].

4 Numerical experiments

In this section, we present the results of several tests to assess the performance of the balancing principle applied to the automatic computation of the TGV parameters. Initially, we demonstrate the advantage of adaptively choosing the two regularization parameters, η_1 and η_2, compared to setting a single parameter, λ, with fixed TGV weights η_1 and η_2. Then, we analyse the characteristics of the objective function $\Phi_\gamma(\boldsymbol{\eta})$ and the behavior of SGP for its minimization. Subsequently, we evaluate the effectiveness of our proposed SGP algorithm by comparing it with the fixed-point approaches FP1 and FP2, reported in Algorithms 1 and 2, respectively. A comprehensive set of results from image restoration tests involving Gaussian noise is extensively discussed in paragraph 4.3. Additionally, in paragraph 4.4, we provide a concise set of tests for Poisson noise including a comparison with the method presented in [7]. Finally, we present the results obtained on a computed tomography (CT) reconstruction problem to illustrate the potentiality of SGP in image reconstruction problems other than image deblurring. All the codes were implemented in MATLAB and the tests were run using MATLAB R2022b (Update 8) on a Windows 11 laptop equipped with an AMD Ryzen 5 5625U processor with 16 GB RAM.

Four reference grey-level images, with pixel values between 0 and 1, were used in the experiments: `peppers`, `parrot`, `mri_head` and `pumpkins`, shown in Figs. 1. The first three images have size 256×256, while the `pumpkins` image is 512×512.

To define the test problems, each reference image was convolved with a Gaussian PSF with variance 2, applying the `psfGauss` MATLAB function from [19]. The resulting image was then corrupted either by Gaussian or Poisson noise. In the case of Gaussian noise, noise levels $\delta = 2.5 \cdot 10^{-2}, 5 \cdot 10^{-3}$ were used (the noise level is defined as the ratio between the noise norm and the norm of the noise-free blurred image).

In our implementation we considered $\varepsilon = 10^{-5}$ for problem (18); furthermore, in Algorithm 3 we set $\mu = 10^{-4}$, $t_{\min} = 10^{-3}$, and $t_{\max} = 5$.

The stopping criteria for the outer iterations of all the algorithms are based on the relative distance between the regularization parameters at two consecutive iterations. This is mathematically represented as:

$$\|\boldsymbol{\eta}^{(k)} - \boldsymbol{\eta}^{(k-1)}\| \leq Tol_{\boldsymbol{\eta}} \|\boldsymbol{\eta}^{(k)}\|,$$

where $Tol_{\boldsymbol{\eta}}$ is set to 10^{-4}. Furthermore, a maximum number of 20 outer iterations is allowed for all the algorithms.

In our setting, we have assumed periodic boundary conditions for both the linear operator A and the finite-difference operators D_H and D_V, implying that they have a BCCB structure. Therefore, the computational cost of the ADMM method used for the solution of the regularized problems of the form (27) is reduced by resorting to 2D Fast Fourier Transforms (FFTs) as detailed in [18]. Each call to ADMM terminates when the approximate relative distance between two successive iterates falls below 10^{-5} or after a maximum value of 2000 iterations.

The quality of the restored images is quantitatively evaluated by means of the Root Mean Square Error (RMSE) value, and the efficiency of the algorithms is measured in terms of the total number of ADMM iterations.

Starting guesses We propose to automatically set the starting guess of the parameters η_1, η_2 in Algorithms 1, 2, 3 by using the noisy images as follows

$$\eta_1^0 = \frac{\phi(\mathbf{b})}{\left\|\nabla \hat{\mathbf{u}} - \hat{\mathbf{w}}\right\|_{2,1|\mathbb{R}^{2n}}}, \qquad \eta_2^0 = \frac{\phi(\mathbf{b})}{\left\|\mathcal{E}\hat{\mathbf{w}}\right\|_{2,1|\mathbb{R}^{4n}}},$$

where the vectors $\hat{\mathbf{u}}, \hat{\mathbf{w}}$ are rough approximations to the true image obtained by early stopping (five iterations for $\hat{\mathbf{u}}$ and one iteration for $\hat{\mathbf{w}}$) of the Projected Gradient method for the solution of

$$\min_{\mathbf{x} \geq 0} \quad \phi(\mathbf{x}).$$

4.1 Comparison between one and two parameters TGV settings

In this section, we compare the proposed approach for setting the values of the regularization parameters η_1 and η_2 with the approach described in [5] for computing the value of the regularization parameter λ in (5). In detail, we set the weights of the TGV regularizer as $\alpha_0 = \beta$ and $\alpha_1 = (1-\beta)$ with $\beta = \frac{1}{3}$, and the regularization parameter λ was chosen according to the one-parameter BP [6]:

$$\begin{cases} \mathbf{x}_\lambda = \arg\min\limits_{\mathbf{x} \in \Omega} \left\{ \phi(\mathbf{x}) + \lambda \big(\beta \psi_1(\mathbf{x}) + (1-\beta)\psi_2(\mathbf{x}) \big) \right\}, \\ \lambda = \dfrac{\phi(\mathbf{x}_\lambda)}{\gamma(\beta \psi_1(\mathbf{x}_\lambda) + (1-\beta)\psi_2(\mathbf{x}_\lambda))}. \end{cases} \tag{30}$$

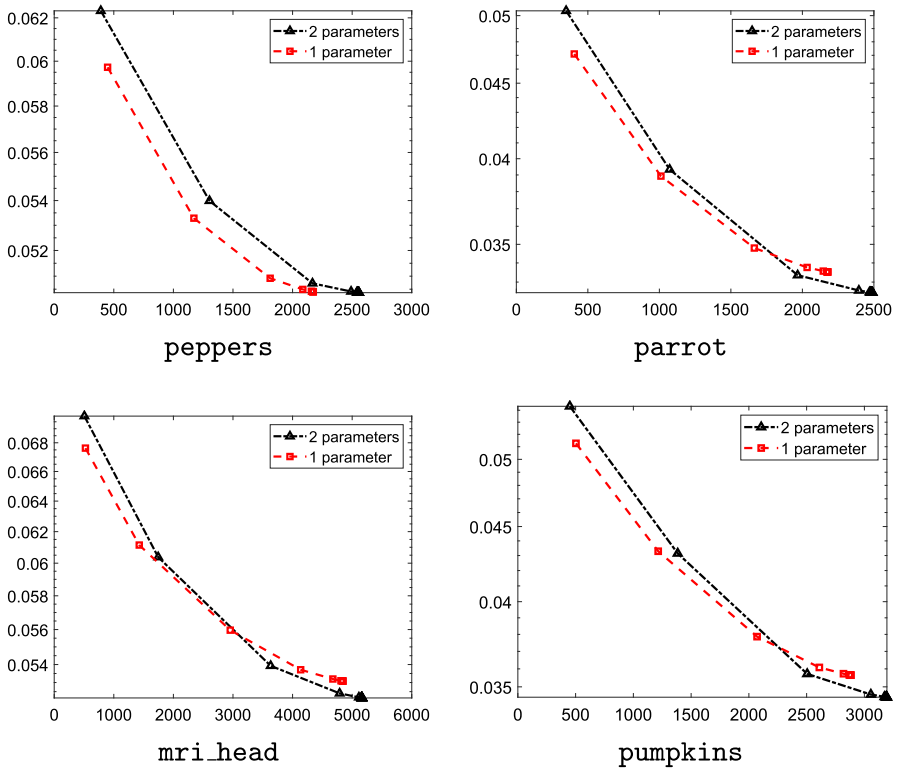

Fig. 2 RMSE history versus the number of ADMM iterations for one-parameter (red dashed line) and two-parameters (black dash-dotted line) FP methods with Gaussian noise $\delta = 5 \cdot 10^{-3}$

The value for γ was selected according to the uniform penalty principle as $\gamma = 1$. A solution to the BP nonlinear system (30) is obtained by employing fixed-point iterations as in FP2. The initial guess is chosen as

$$\lambda^{(0)} = \frac{\phi(\mathbf{b})}{\beta \left\| \nabla \hat{\mathbf{u}} - \hat{\mathbf{w}} \right\|_{2,1|\mathbb{R}^{2n}} + (1-\beta) \left\| \mathcal{E}\hat{\mathbf{w}} \right\|_{2,1|\mathbb{R}^{4n}}}. \tag{31}$$

We explore the impact of utilizing a one-parameter versus a two-parameter setting for TGV and we aim to demonstrate the enhanced accuracy attainable when the balancing principle is applied within a two-parameter framework. For a fair comparison with the approach of [5], we use Algorithm 2 for the two-parameter framework. For the sake of space, we limit the experiments in this section to the case of images affected by Gaussian noise.

The results, depicted in Figs. 2 and 3, provide a comparative analysis of the RMSE history plotted against the number of ADMM iterations, illustrating the performance of the one-parameter (red dashed line) and two-parameter (black dash-dotted line) Fixed-Point (FP) methods. Figure 2 shows low-noise tests ($\delta = 5 \cdot 10^{-3}$), while Fig. 3 corresponds to high-noise tests ($\delta = 2.5 \cdot 10^{-2}$). In all the cases but for the test with

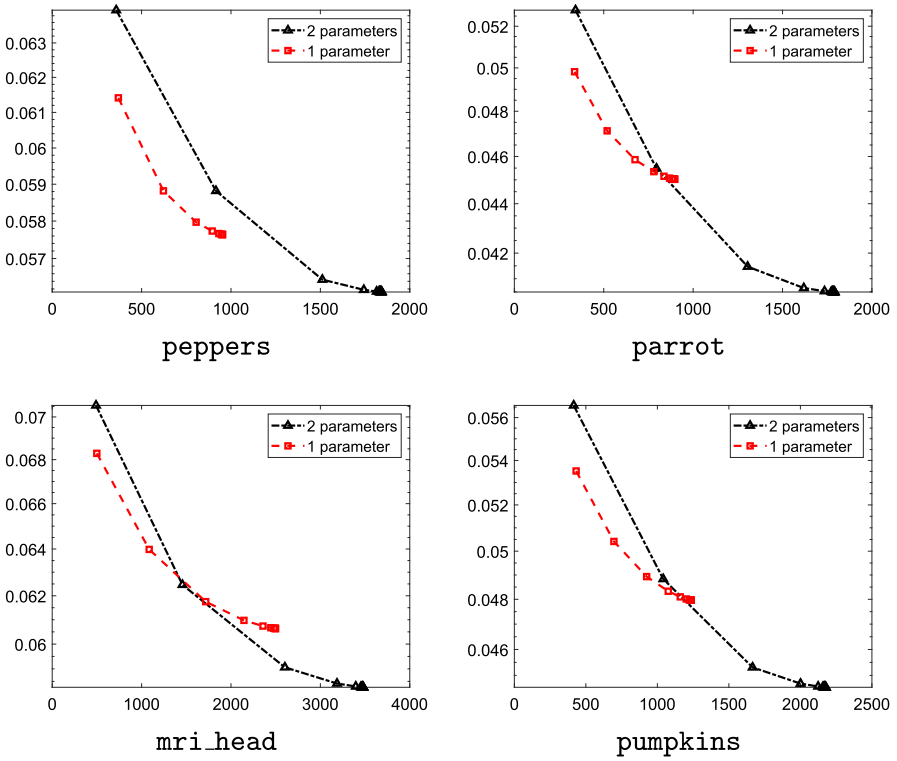

Fig. 3 RMSE history versus the number of ADMM iterations for one-parameter (red dashed line) and two-parameters (black dash-dotted line) FP methods with Gaussian noise $\delta = 2.5 \cdot 10^{-2}$

peppers image and $\delta = 5 \cdot 10^{-3}$, the two-parameter method reaches better solutions in terms of RMSE, especially in the noisier cases ($\delta = 5 \cdot 10^{-3}$). The results suggest that the additional degree of freedom in the choice of the regularizer can bring to improved performances without leading to increased computational cost.

To further assess the importance of the additional degree of freedom, in Fig. 4, we show, for the parrot test problem, the comparison between the two-parameter fixed-point scheme and different versions of the one-parameter scheme obtained by varying the value of β, i.e., the relative weight of the two components of the TGV regularization. From the plots, one can indeed appreciate the sensitivity of TGV with respect to the choice of β.

4.2 Analysis of the method's behavior

The aim of this section is twofold. First, we intend to evaluate the performance of the proposed SGP method for the minimization of $\Phi_\gamma(\boldsymbol{\eta})$. Then, we want to compare the parameters η_1 and η_2 computed by SGP with those heuristically obtained by trial and error. We focus on the test problem peppers with Gaussian noise and solve the TGV-based problem (27) for several (η_1, η_2) values in a 25×25 grid of

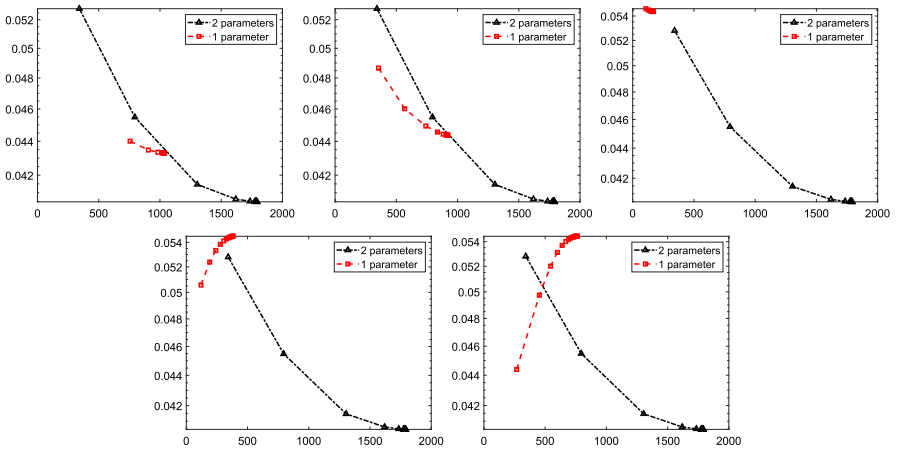

Fig. 4 `parrot` test problem: RMSE history versus the number of ADMM iterations for one-parameter (red dashed line) and two-parameters (black dash-dotted line) FP methods with Gaussian noise $\delta = 2.5 \cdot 10^{-2}$ and varying values of β. From left to right and top to bottom: $\beta = 0.1, 0.3, 0.5, 0.7, 0.9$

Table 1 Table presenting performance metrics for `peppers` test image under different noise levels (δ).

δ	Method	RMSE	η_1	η_2
$5 \cdot 10^{-3}$	SGP	$5.039 \cdot 10^{-2}$	$1.716 \cdot 10^{-4}$	2.815
	Best	$4.901 \cdot 10^{-2}$	$7.499 \cdot 10^{-5}$	$2.873 \cdot 10^{-4}$
$2.5 \cdot 10^{-2}$	SGP	$5.567 \cdot 10^{-2}$	$2.306 \cdot 10^{-3}$	5.014
	Best	$5.393 \cdot 10^{-2}$	$1.778 \cdot 10^{-3}$	$3.162 \cdot 10^{-3}$

Each row specifies the method for the TGV parameters computation (Method) and provides the values of RMSE and TGV parameters (η_1 and η_2)

logarithmically equally spaced points. Figures 5 and 6 represent the contour plot and surface of the objective function $\Phi_\gamma(\boldsymbol{\eta})$ (in $\log_{10}$ scale) and RMSE for the low-noise and high-noise tests, respectively. The red points in the contour plots indicate the best parameter pair that minimizes the RMSE over the grid. The blue points represent the SGP iterates $(\eta_1^{(k)}, \eta_2^{(k)})$. From Figs. 5 and 6, it is evident that $\Phi_\gamma(\boldsymbol{\eta})$ is a non-convex function that can have multiple local minima and SGP has a good performance in finding an approximation of a local minimum. Moreover, the RMSE function attains its minimum value inside an L-shaped valley that becomes narrower as the noise increases. Therefore, many parameter pairs (η_1, η_2) can give very close RMSE values corresponding to almost visually indistinguishable images. Comparing the contour plots in Figs. 5 and 6, it is evident that the local minima of $\Phi_\gamma(\boldsymbol{\eta})$ lie in the valley of the RMSE function. This shows that the Balancing Principle, based on minimizing $\Phi_\gamma(\boldsymbol{\eta})$, is a suitable criterion to select regularization parameters providing good restorations. Table 1 reports the best TGV parameters and those computed by SGP together with the corresponding RMSE values. These values confirm the previous considerations: since the RMSE valley is L-shaped, the best and SGP values of RMSE are very close as well as the values of the parameter η_1. However, the values of η_2 differ by three or

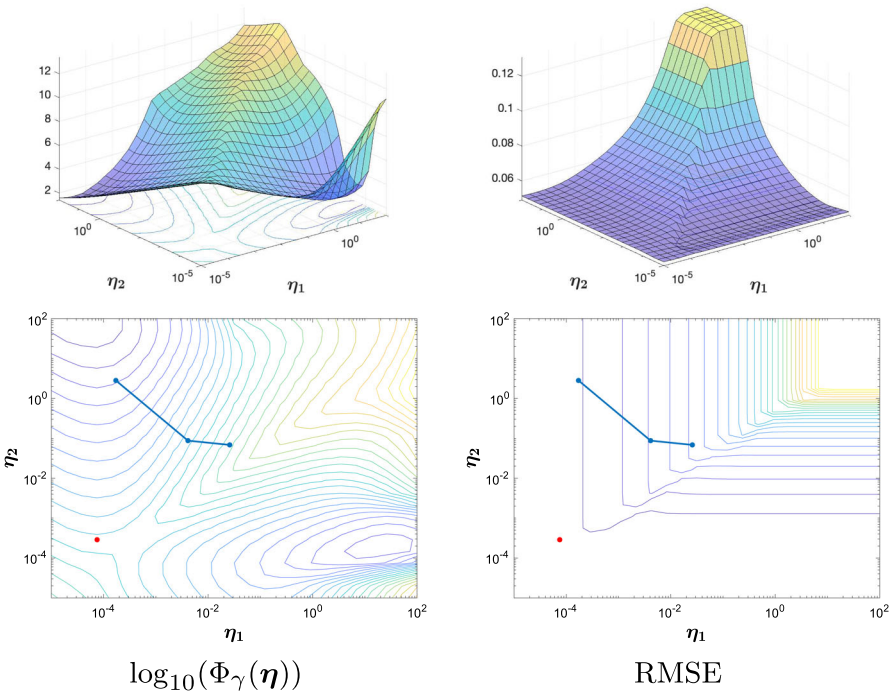

Fig. 5 Contour plot and surface of the objective function $\Phi_\gamma(\boldsymbol{\eta})$ (left) and RMSE (right) over the (η_1, η_2) grid for the test image `peppers`, Gaussian noise $\delta = 5 \cdot 10^{-3}$. The red point indicates where the RMSE is minimized over the grid; the blue points represent the SGP iterates $(\eta_1^{(k)}, \eta_2^{(k)})$

four orders of magnitude. The best and SGP restored images are depicted in Fig. 7. Their visual quality is comparable, and the SGP images tend to be smoother. This is somehow expected. Indeed, ψ_1 detects and highlights steep gradients at edges, while ψ_2 identifies and emphasizes smooth intensity transitions [2]; hence, a larger value for η_2, weighting ψ_2, will result in emphasized smoothness.

4.3 Gaussian noise

Here we report the results of the comparison between the proposed SGP method described in Algorithm 3 and the two fixed-point strategies described in Sect. 3.1. It is worth noting that, because of the peculiar choice we made for the scaling matrix D_k in (21) and t_k in (25)–(26), we expect the SGP method and FP1 (Algorithm 1) to behave identically in the first two iterations if no line-search iteration is performed.

Figures 8 and 9 report the comparison between the three algorithms in terms of RMSE against ADMM iterations for problems with $\delta = 5 \cdot 10^{-3}$ and $\delta = 2.5 \cdot 10^{-2}$, respectively. We observe that, as discussed in the previous paragraph, SGP and FP1 behave almost identically for the first two iterations for most of the cases, indicating that the line-search is not performed in the initial phases. However, starting from

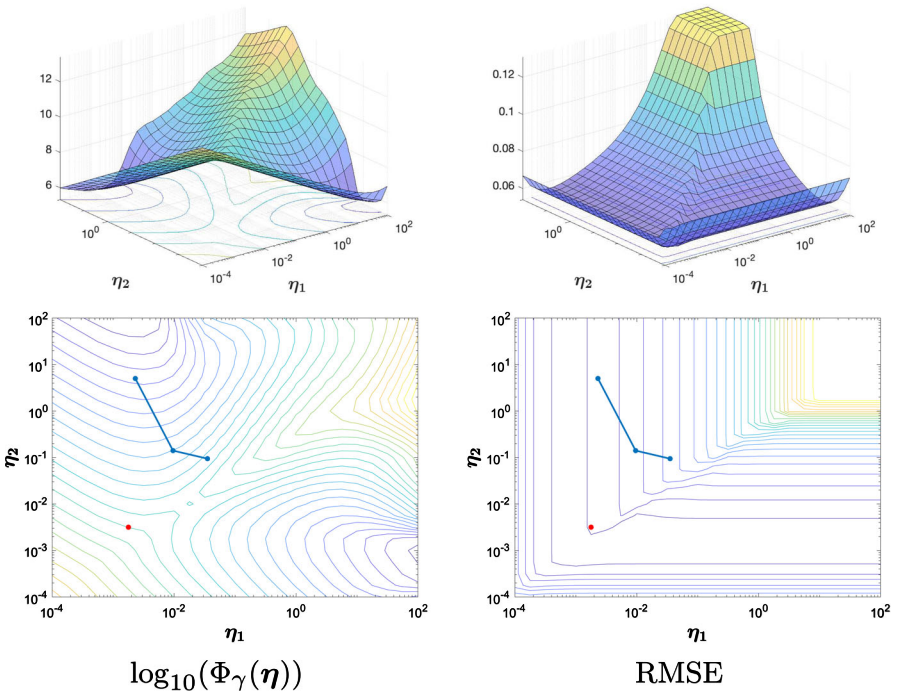

Fig. 6 Contour plot and surface of the objective function $\Phi_\gamma(\boldsymbol{\eta})$ (left) and RMSE (right) over the (η_1, η_2) grid for the test image `peppers`, Gaussian noise $\delta = 2.5 \cdot 10^{-2}$. The red point indicates where the RMSE is minimized over the grid; the blue points represent the SGP iterates $(\eta_1^{(k)}, \eta_2^{(k)})$

the third iteration, the combination of spectral steplengths and Armijo backtracking appears to be very effective to guide the method towards convergence and avoid the "semi-convergent" behaviour of the fixed-point iteration. This holds true especially for the case of higher noise level shown in Fig. 9. It is worth noting that both SGP and FP1 perform better than FP2 which appears to stagnate after just 3–4 iterations. In Table 2, we report the numerical values for accuracy (RMSE) and efficiency (ADMM iterations) shown in Figs. 8 and 9 and, in the last two columns, we show the values of the TGV parameters (η_1, η_2) computed by each method. We observe that the SGP method requires fewer ADMM iterations than FP1 and FP2, demonstrating faster convergence across various test images and noise levels.

To visualize the parameter values $\eta_1^{(k)}$ and $\eta_2^{(k)}$ computed during iterations k of the three methods, we plot them on the contour lines of both the objective function Φ_γ and the RMSE. For the image `peppers`, Fig. 10 displays the computed values from SGP (blue line with dots), FP1 (black line with dots), and FP2 (magenta line with dots), along with the optimal parameters (η_1, η_2). The left panels show the Φ_γ contours (in $\log_{10}$ scale), while the right panels show the RMSE contours. The top row corresponds to low-noise tests, and the bottom row to high-noise tests. By comparing the contour plots of Φ_γ and RMSE in Fig. 10, we observe that an oscillatory behaviour of the

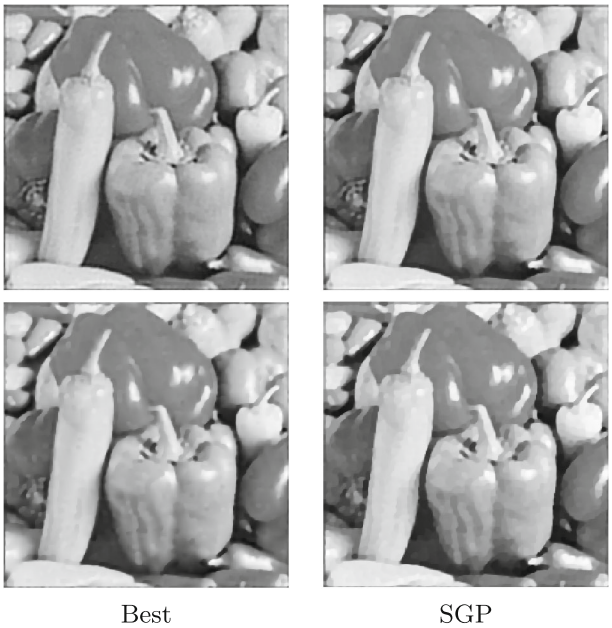

Fig. 7 Restored images corresponding to the minimum RMSE (left) and obtained with SGP (right) for noise levels $\delta = 5 \cdot 10^{-3}$ (top row) and $\delta = 2.5 \cdot 10^{-2}$ (bottom row)

objective function corresponds to the same oscillatory behaviour of the RMSE. This can explain why SGP, imposing a monotonic decrease of the objective function, tends to avoid a semiconvergence phenomenon.

To show the effectiveness of the proposed method we report in Figs. 11 and 12 the noisy images and the restorations obtained with SGP for all the test cases.

4.4 Poisson noise

To show the flexibility of the proposed approach with respect to the data fidelity term, in this section, we report the results of the experiments we performed on a pair of blurry images affected by Poisson noise. We observe that in the case of Poisson noise, to take into account the existence of some background emission, a constant term $\boldsymbol{\nu}$ equal to 10^{-10} was added to all pixels of the blurry image. The resulting image was then corrupted using the MATLAB function `imnoise`. The intensities of the original images were pre-scaled to get noisy and blurry images with Signal to Noise Ratio (SNR) equal to 40dB.

Figure 13 report the noisy images, the restored images resulting from the application of the proposed SGP method, and the comparison between SGP and the two fixed-point methods in terms of RMSE against total number of ADMM iterations. We observe that also in this case FP2 shows stagnation after a few iterates and is outperformed by SGP and FP1 in terms of RMSE at the solution. Furthermore, as for the case of Gaussian noise, SGP outperforms the related fixed-point method (FP1), avoiding the

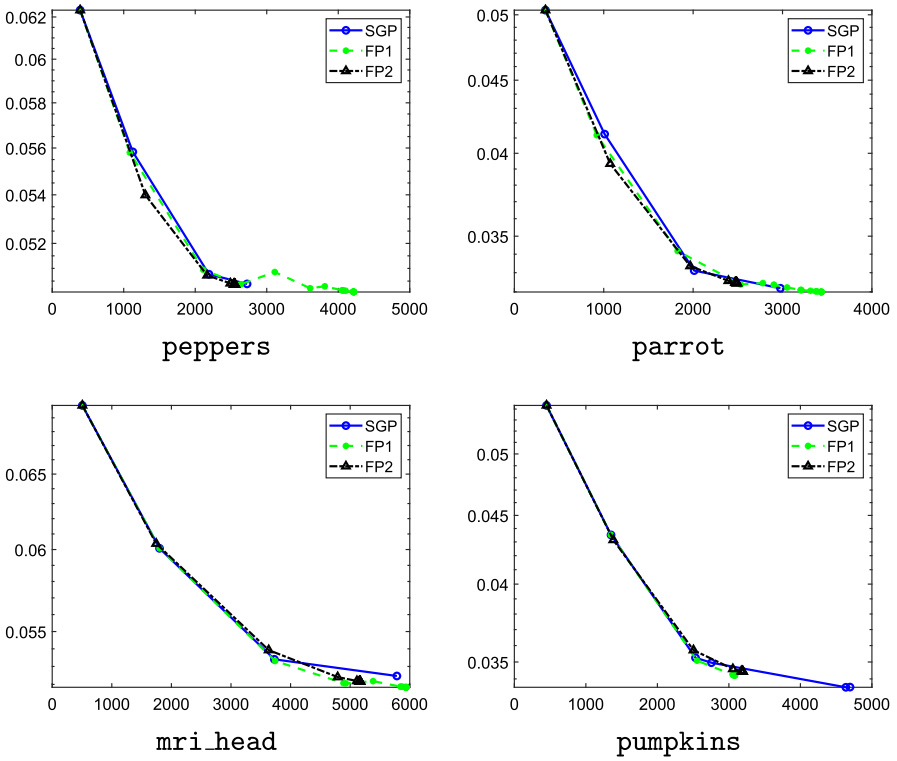

Fig. 8 Test problems with Gaussian noise level $\delta = 5 \cdot 10^{-3}$: RMSE history versus the number of ADMM iterations for all the methods

"semi-convergent" behavior. It is worth mentioning that in the case of the `parrot` image the convergence paths of FP1 and SGP split after the first iteration, suggesting that the full FP1 direction leads to an increased value of Φ_γ. The values of the TGV parameters (η_1, η_2), computed by each method, are reported in the last two columns of Table 3 together with the RMSE, the number of SGP iterations (for SGP) or fixed-point iterations (for FP1 and FP2) denoted by (k), the number of ADMM iterations (ADMM it), for the tests represented in Fig. 13.

4.4.1 Comparison with state of the art

In this section we compare the performance of the proposed algorithm with a state-of-the-art method for automatic estimation of the TGV regularization parameters pair. In detail, we consider the method proposed in [7, 8], in which the authors consider a hierarchical Bayesian formulation of the image restoration problem and couple it with a Maximum A Posteriori (MAP) estimation approach.

For this comparison, we considered images `peppers` and `parrot` in Fig. 1, and we generated problem instances in the same way used in [7]. We scaled them to have a maximum pixel intensity (i.e. maximum number of photons detected) equal,

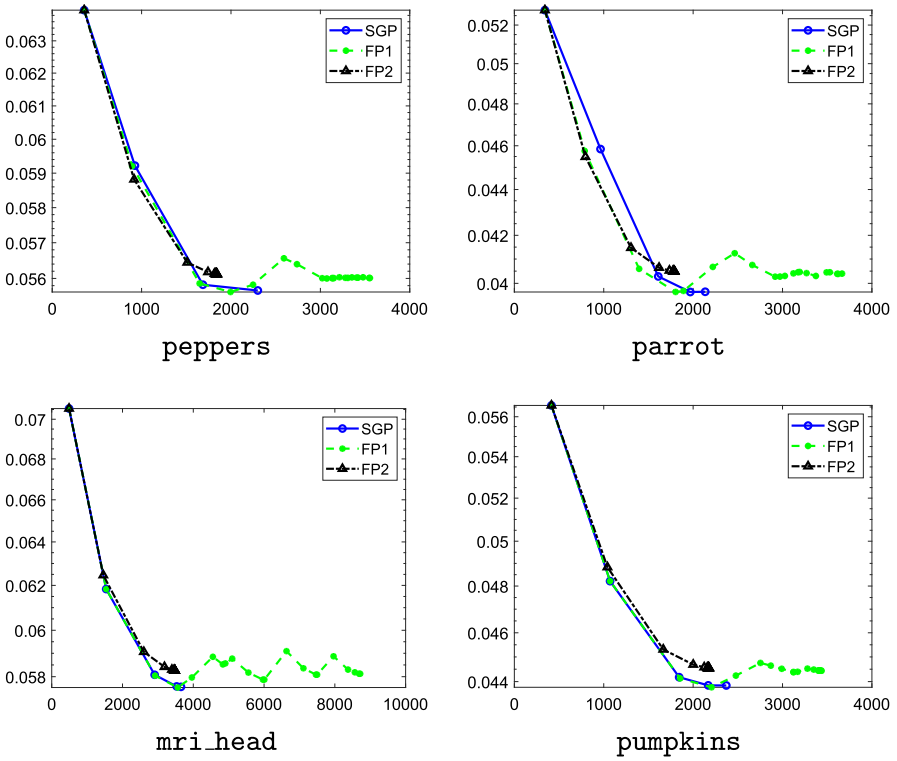

Fig. 9 Test problems with Gaussian noise level $\delta = 2.5 \cdot 10^{-2}$: RMSE history versus the number of ADMM iterations for all the methods

respectively, to $2.95 \cdot 10^3$ and $3.25 \cdot 10^3$. We applied a Gaussian blur with side length 5 pixels and standard deviation 1. The images have then been corrupted by Poisson noise using the MATLAB `poissrnd` function. Table 4 reports the RMSE and PSNR metrics for the restored images while Fig. 14 reports the noisy images (left) and the restored images resulting from the application of the proposed SGP method (center) and the method in [7] (right). From the table one might establish that the two methods perform comparably, with both methods are able to produce a good restorations of the image. However, from the images one can observe that the two restored images present quite distinct features: SGP seems to be better than the method in [7] in removing blur but tends to perform worse in terms of denoising (one can observe the "mosaic" effect in the images in the central column).

4.5 CT reconstruction

This last section aims to show that SGP has a promising potential when applied to image reconstruction problems other than image deblurring. We consider a CT image reconstruction problem where the test image, shown in Fig. 15, is a 256×256 slice of the 3D extended cardiac-torso (`XCAT`) phantom, a whole-body computer model of the

Table 2 Table presenting performance metrics for various test images and reconstruction methods under different noise levels (δ)

δ	Test Image	Method	Iterations		RMSE	TGV parameters	
			(k)	ADMM it		η_1	η_2
$5 \cdot 10^{-3}$	Peppers	SGP	4	6435	$5.039 \cdot 10^{-2}$	$1.716 \cdot 10^{-4}$	2.815
		FP1	21	49699	$5.037 \cdot 10^{-2}$	$8.931 \cdot 10^{-5}$	1.676
		FP2	21	75194	$5.007 \cdot 10^{-2}$	$8.567 \cdot 10^{-5}$	0.4352
	Parrot	SGP	4	6347	$3.218 \cdot 10^{-2}$	$2.631 \cdot 10^{-4}$	1.842
		FP1	21	47987	$3.246 \cdot 10^{-2}$	$1.026 \cdot 10^{-4}$	1.800
		FP2	21	60998	$3.199 \cdot 10^{-2}$	$1.016 \cdot 10^{-4}$	0.4225
	mri_head	SGP	4	11819	$5.247 \cdot 10^{-2}$	$1.087 \cdot 10^{-4}$	9.199
		FP1	21	98384	$5.216 \cdot 10^{-2}$	$2.498 \cdot 10^{-5}$	6.301
		FP2	21	110072	$5.184 \cdot 10^{-2}$	$2.468 \cdot 10^{-5}$	0.6626
	Pumpkins	SGP	6	16421	$3.353 \cdot 10^{-2}$	$1.026 \cdot 10^{-4}$	1.346
		FP1	13	36100	$3.444 \cdot 10^{-2}$	$6.923 \cdot 10^{-5}$	1.382
		FP2	21	59699	$3.422 \cdot 10^{-2}$	$6.813 \cdot 10^{-5}$	0.6727
$2.5 \cdot 10^{-2}$	peppers	SGP	4	5269	$5.567 \cdot 10^{-2}$	$2.306 \cdot 10^{-3}$	5.014
		FP1	21	35725	$5.612 \cdot 10^{-2}$	$2.734 \cdot 10^{-3}$	17.22
		FP2	21	58541	$5.601 \cdot 10^{-2}$	$2.678 \cdot 10^{-3}$	10.56
	Parrot	SGP	5	7019	$3.966 \cdot 10^{-2}$	$2.256 \cdot 10^{-3}$	20.38
		FP1	21	34344	$4.049 \cdot 10^{-2}$	$3.211 \cdot 10^{-3}$	9.148
		FP2	21	56536	$4.040 \cdot 10^{-2}$	$3.195 \cdot 10^{-3}$	9.427
	mri_head	SGP	5	12121	$5.757 \cdot 10^{-2}$	$5.619 \cdot 10^{-4}$	109.6
		FP1	21	66677	$5.827 \cdot 10^{-2}$	$8.044 \cdot 10^{-4}$	32.49
		FP2	21	120410	$5.812 \cdot 10^{-2}$	$7.578 \cdot 10^{-4}$	28.22
	Pumpkins	SGP	5	7870	$4.385 \cdot 10^{-2}$	$1.678 \cdot 10^{-3}$	22.00
		FP1	21	41969	$4.455 \cdot 10^{-2}$	$2.305 \cdot 10^{-3}$	15.08
		FP2	21	60012	$4.444 \cdot 10^{-2}$	$2.289 \cdot 10^{-3}$	9.960

Each row specifies a test image (Test Image), and the reconstruction method (Method), and provides detailed metrics: the number of SGP iterations (for SGP) or fixed-point iterations (for FP1 and FP2) denoted by (k), the number of ADMM iterations (ADMM it), the root mean square error (RMSE), and TGV parameters (η_1 and η_2). The table is organized by noise levels, with δ values of $5 \cdot 10^{-3}$ and $2.5 \cdot 10^{-2}$

human anatomy. Two scan geometries have been considered: a fan-beam geometry and a parallel-bean geometry [20]; the corresponding system matrices A have been obtained with the AIR Tools Matlab package [21]. For the fan-beam geometry, we have considered 360 equally spaced projection angles from 0 to 2π and 181 rays for each angle, while for the parallel-beam geometry, we have considered 180 equally spaced angles from 0 to π with 181 rays for each angle. The resulting sinograms have been corrupted by Gaussian noise with level $\delta = 5 \cdot 10^{-3}$. Figure 15 shows the noisy sinograms. Since the system matrix is not a BCCB matrix, we use the primal-dual ascent-descent (PD) method [22] to solve the TGV-based subproblems (27) at

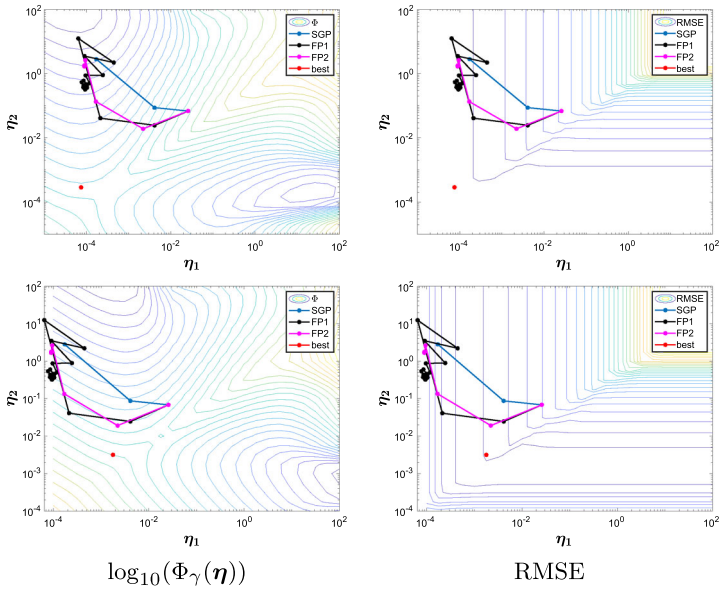

Fig. 10 Contour plot of the RMSE over the (η_1, η_2) grid for the test image `peppers` for noise levels $\delta = 5 \cdot 10^{-3}$ (top row) and $\delta = 2.5 \cdot 10^{-2}$ (bottom row). The red point indicates where the RMSE is minimized over the grid. The blue points represent the SGP iterates, the black points represent the FP1 iterates, and the magenta points represent the FP2 iterates

Fig. 11 Test problems with Gaussian noise level $\delta = 5 \cdot 10^{-3}$: noisy images (top row) and restored images obtained with SGP (bottom row)

each iteration of the SGP, FP1 and FP2 methods. Each call to PD terminates when the approximate relative distance between two successive iterates falls below 10^{-3} or after a maximum of 2000 iterations. The SGP restored images are depicted in Fig. 16. One can observe that in both cases the SGP method was able to produce a good reconstruction of the original CT scan, and this can also be seen by looking at the

Fig. 12 Test problems with Gaussian noise level $\delta = 2.5 \cdot 10^{-2}$: noisy images (top row) and restored images obtained with SGP (bottom row)

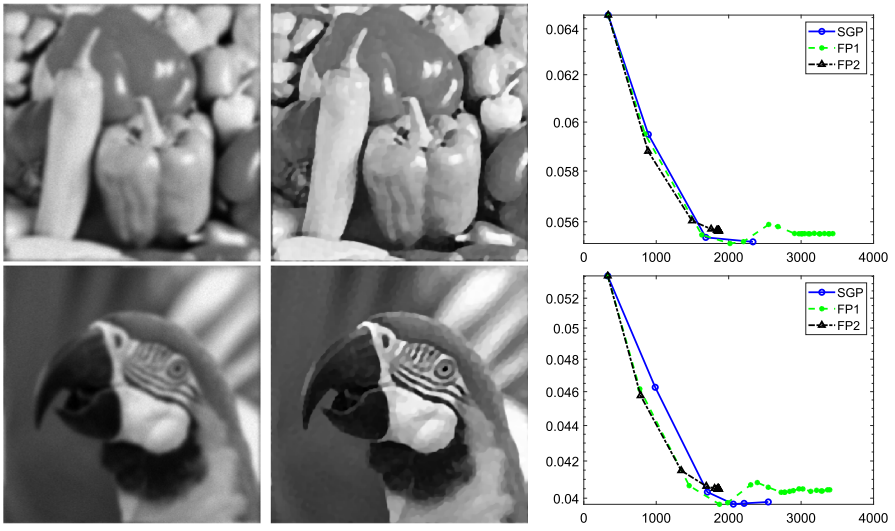

Fig. 13 Test problems with Poisson noise: noisy images (left column), restored images obtained with SGP (middle column), and RMSE history versus number of ADMM iterations

RMSE values reported in Table 5. In our opinion this result emphasizes the robustness of the balancing principle with respect to the change of the linear operator A. Indeed, no particular parameter tuning had to be performed to obtain such results, but we only had to plug in the new solver for the inner problems.

Table 3 Table presenting performance metrics for various test images and reconstruction methods for the Poisson noise cases represented in Fig. 13

Test image	Method	Iterations		RMSE	TGV parameters	
		(k)	ADMM it		η_1	η_2
Peppers	SGP	4	5247	$5.523 \cdot 10^{-2}$	$2.274 \cdot 10^{-3}$	4.638
	FP1	21	36045	$5.566 \cdot 10^{-2}$	$2.519 \cdot 10^{-3}$	$1.512 \cdot 10^{1}$
	FP2	21	56660	$5.555 \cdot 10^{-2}$	$2.619 \cdot 10^{-3}$	8.662
Parrot	SGP	6	9848	$3.980 \cdot 10^{-2}$	$2.186 \cdot 10^{-3}$	$1.864 \cdot 10^{1}$
	FP1	21	35670	$4.049 \cdot 10^{-2}$	$2.970 \cdot 10^{-3}$	$1.043 \cdot 10^{1}$
	FP2	21	53336	$4.044 \cdot 10^{-2}$	$3.009 \cdot 10^{-3}$	8.432

Each row specifies a test image (Test Image), and the reconstruction method (Method), and provides detailed metrics: the number of SGP iterations (for SGP) or fixed-point iterations (for FP1 and FP2) denoted by (k), the number of ADMM iterations (ADMM it), the root mean square error (RMSE), and TGV parameters (η_1 and η_2)

Table 4 Table presenting performance metrics for the test described in Sect. 4.4.1. Each row specifies a test image (Test Image), and the reconstruction method (Method), and provides RMSE and PSNR

Test image	Method	RMSE	PSNR
Peppers	SGP	$5.59 \cdot 10^{-2}$	25.053
	[7]	$5.58 \cdot 10^{-2}$	25.065
Parrot	SGP	$4.05 \cdot 10^{-2}$	27.86
	[7]	$4.22 \cdot 10^{-2}$	27.49

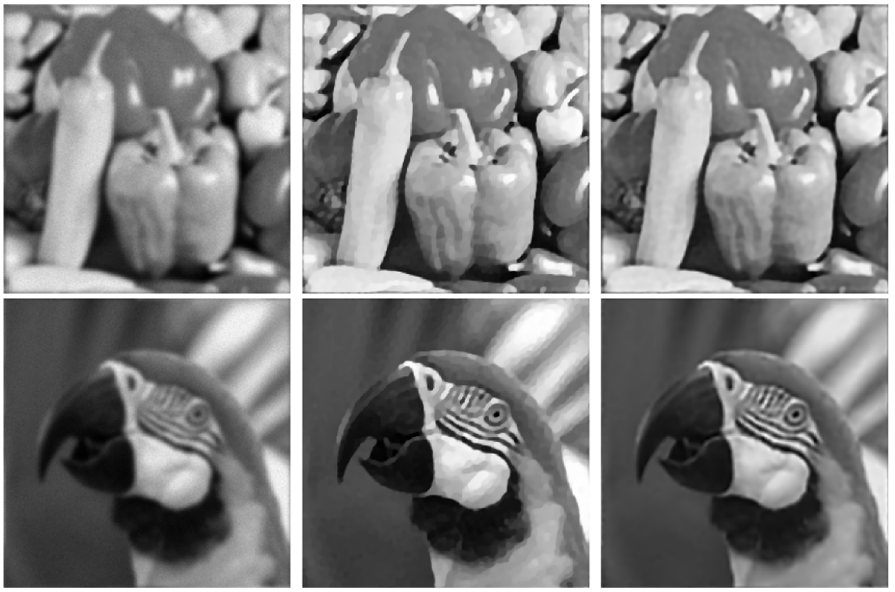

Fig. 14 Results of the test described in Sect. 4.4.1: noisy images (left column), restored images obtained with SGP (center column, and restored images with the method in [7])

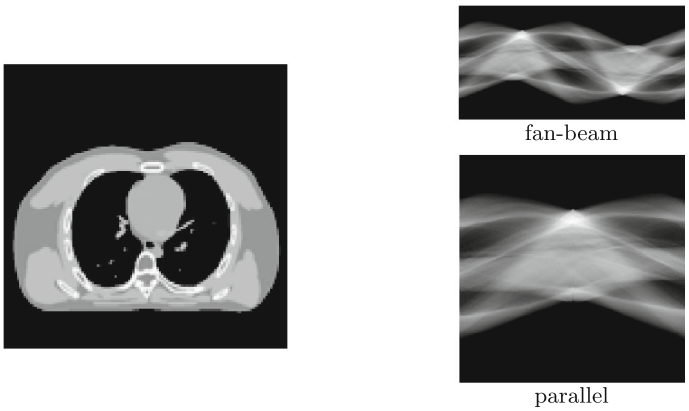

Fig. 15 XCAT test problem: exact image (left) and noisy sinograms (right)

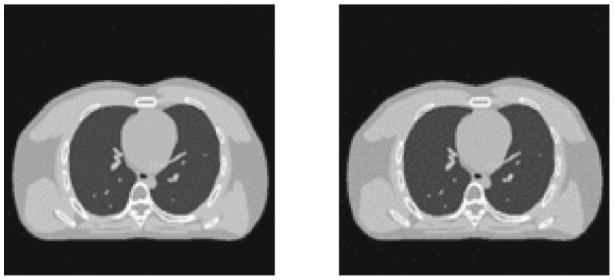

Fig. 16 XCAT test problem: restored image obtained with SGP for fan-beam (left) and parallel-beam (right) geometry

Table 5 Table presenting performance metrics for the CT image reconstruction problem

Geometry	Method	Iterations		RMSE	TGV parameters	
		(k)	PD it		η_1	η_2
Fan-beam	SGP	21	1471	$8.168 \cdot 10^{-3}$	$8.498 \cdot 10^{-1}$	$7.444 \cdot 10^{-1}$
Parallel-beam	SGP	16	862	$1.136 \cdot 10^{-2}$	$3.888 \cdot 10^{-1}$	$3.412 \cdot 10^{-1}$

Each row specifies the scan geometry and provides detailed metrics: the number of SGP iterations denoted by (k), the number of PD iterations (PD it), the root mean square error (RMSE), and TGV parameters (η_1 and η_2)

5 Conclusions

In this work we introduced an adaptation of the balancing principle for the multi-parameter setting to the case of TGV-regularized image restoration problems. We analyze two possible fixed-point schemes (here named FP1 and FP2) for automatic parameter selection, based on the work in [10, 11] and we introduce a tailored Scaled Gradient Projection method for the solution of the hyperparameter selection problem. The comparison of SGP with the related fixed-point methods on image restoration

problems affected by both Gaussian and Poisson noise reveals that the proposed strategy yields better restorations without leading to an increased computational cost. In contrast to FP1, to which it is strongly related, the SGP method does not exhibit the oscillatory behaviour of the relative error. When compared with other state-of-the-art strategies, the SGP showed reliable restoration performances. Additionally, we also tested the potential of the proposed method in a different scenario, namely the reconstruction of CT scans. Our experiments showed that the method is robust to the change of the linear operator, obtaining good reconstructions without the need of a parameter tuning.

Future works will deal with the analysis of the convergence properties of the proposed SGP method and the theoretical implications of inexactness in the solution of the TGV-regularized restoration problem (27) at each iteration.

In memoriam

We dedicate this paper to the memory of our friend and colleague Prof. Daniela di Serafino, who inspired our interest in regularization through TGV in image restoration. This work is a continuation of a research path we started with her a few years ago [5]. We will always miss her and we will be always thankful for her friendship and her guidance.

Acknowledgements The authors would like to thank Dr Monica Pragliola (University of Naples Federico II) for her assistance with the comparison presented in Sect. 4.4.1. This research was partially supported by the Istituto Nazionale di Alta Matematica, Gruppo Nazionale per il Calcolo Scientifico (INdAM-GNCS) "INdAM - GNCS Project", project code: CUP_E53C23001670001. GL is supported by the PRIN 2022 project "STILE: Sustainable Tomographic Imaging with Learning and rEgularization", project code: 20225STXSB, funded by the European Commission under the NextGeneration EU programme.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

References

1. Hansen, P.C., Nagy, J.G., O'Leary, D.P.: Deblurring Images: Matrices, Spectra, and Filtering. Society for Industrial and Applied Mathematics, USA (2006)
2. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM J. Imaging Sci. **3**, 492–526 (2010). https://doi.org/10.1137/09076952
3. Bredies, K., Holler, M.: Regularization of linear inverse problems with total generalized variation. J. Inverse Ill-Posed Probl. **22**, 871–913 (2014). https://doi.org/10.1515/jip-2013-0068

4. Hintermüller, M., Papafitsoros, K., Rautenberg, C.N., Sun, H.: Dualization and automatic distributed parameter selection of total generalized variation via bilevel optimization. Numer. Funct. Anal. Optim. **43**(8), 887–932 (2022)
5. di Serafino, D., Landi, G., Viola, M.: TGV-based restoration of poissonian images with automatic estimation of the regularization parameter. In: Proceedings of the 21th International Conference on Computational Science and Its Applications (ICCSA), Cagliari, Italy (2021). https://doi.org/10.1109/ICCSA54496.2021.00028
6. Ito, K., Jin, B., Takeuchi, T.: A regularization parameter for nonsmooth Tikhonov regularization. SIAM J. Sci. Comput. **33**(3), 1415–1438 (2011). https://doi.org/10.1137/100790756
7. di Serafino, D., Pragliola, M.: Automatic parameter selection for the TGV regularizer in image restoration under Poisson noise (2022). https://doi.org/10.48550/arXiv.2205.13439
8. Lanza, A., Pragliola, M., Sgallari, F.: Automatic fidelity and regularization terms selection in variational image restoration. BIT Numer. Math. **62**(3), 931–964 (2022). https://doi.org/10.1007/s10543-021-00901-z
9. Zietlow, C., Lindner, J.K.N.: ADMM-TGV image restoration for scientific applications with unbiased parameter choice. Numer. Algorithms (2024). https://doi.org/10.1007/s11075-024-01759-2
10. Ito, K., Jin, B., Takeuchi, T.: Multi-parameter Tikhonov regularization. Methods Appl. Anal. (2011). https://doi.org/10.4310/MAA.2011.v18.n1.a2
11. Ito, K., Jin, B., Takeuchi, T.: Multi-parameter Tikhonov regularization-an augmented approach. Chin. Ann. Math. Ser. B **35**(3), 383–398 (2014). https://doi.org/10.1007/s11401-014-0835-y
12. Bortolotti, V., Landi, G., Zama, F.: Uniform multi-penalty regularization for linear ill-posed inverse problems. SIAM J. Sci. Comput. (2025). https://doi.org/10.48550/arXiv.2309.14163
13. Bortolotti, V., Landi, G., Zama, F.: An automatic pixel-wise multi-penalty approach to image restoration. J. Imaging (2023). https://doi.org/10.3390/jimaging9110249
14. Bonettini, S., Zanella, R., Zanni, L.: A scaled gradient projection method for constrained image deblurring. Inverse Prob. **25**(1), 015002 (2008). https://doi.org/10.1088/0266-5611/25/1/015002
15. Hunter, D.R., Lange, K.: A tutorial on MM algorithms. Am. Stat. **58**(1), 30–37 (2004). https://doi.org/10.1198/0003130042836
16. Zanella, R., Boccacci, P., Zanni, L., Bertero, M.: Efficient gradient projection methods for edge-preserving removal of Poisson noise. Inverse Prob. **25**(4), 045010 (2009)
17. Barzilai, J., Borwein, J.M.: Two-point step size gradient methods. IMA J. Numer. Anal. **8**(1), 141–148 (1988)
18. di Serafino, D., Landi, G., Viola, M.: Directional TGV-based image restoration under Poisson noise. J. Imaging (2021). https://doi.org/10.3390/jimaging7060099
19. Nagy, J.G., Palmer, K., Perrone, L.: Iterative methods for image deblurring: a Matlab object-oriented approach. Numer. Algorithms **36**, 73–93 (2004). https://doi.org/10.1023/B:NUMA.0000027762.08431.64
20. Hansen, P.C., Jørgensen, J., Lionheart, W.R.B.: Computed Tomography: Algorithms, Insight, and Just Enough Theory. Society for Industrial and Applied Mathematics, Philadelphia, PA (2021). https://doi.org/10.1137/1.9781611976670
21. Hansen, P.C., Jørgensen, J.S.: AIR Tools II: algebraic iterative reconstruction methods, improved implementation. Numer. Algorithms **79**(1), 107–137 (2018)
22. Bredies, K.: Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. In: Bruhn, A., Pock, T., Tai, X.-C. (eds.) Efficient Algorithms for Global Optimization Methods in Computer Vision, pp. 44–77. Springer, Berlin, Heidelberg (2014)