Journal of Mathematical Modeling
Vol. 13, No. 3, 2025, pp. 533-552. Research Article

JMM

A new bilevel optimization problem for the image restoration

Abdelmajid El Mourabit, Idriss El Mourabit*

MIMSC, EST d'Essaouira, Universit Cadi Ayyad, Morocco
Email(s): elmourabitabdelmajid@gmail.com, idriss.el.mourabit88@gmail.com

Abstract. This paper presents a novel bilevel optimization approach for a nonlinear partial differential equation. The approach aims to enhance the quality of image denoising by estimating certain parameters within this equation. Our work deals with both analytical and numerical results. Analytically, we establish the existence of a solution to the bilevel optimization problem and apply the Alternating Direction Method of Multipliers algorithm to approximate this solution. Furthermore, the method fine-tunes the restoration process, effectively reducing noise while preserving crucial image features. Finally, numerical results validate the performance of our method, surpassing traditional denoising approaches. This research makes an important contribution to image restoration, paving the way for high-quality practical applications.

Keywords: Bilevel optimization, PDE, image processing, denoising, ADMM.
AMS Subject Classification 2010: 35-XX,49Jxx, 65K10

1 Introduction and motivation

Image processing is a crucial field in applied sciences, touching on various sectors such as medical imaging, remote sensing, and computer vision. In these contexts, captured images are often degraded by noise, which can compromise the accuracy and reliability of analyses and interpretations. To address these issues, various techniques have been developed to enhance image quality. Noise reduction is an essential method, with techniques such as Gaussian filtering [17,18], median filtering [6], and anisotropic diffusion [2, 16, 19, 21] helping to reduce noise while preserving important image features.

In particular, non-linear Partial Differential Equations (PDEs) [1, 8, 9, 16, 19] have proven their effectiveness in image denoising by capturing the complex features of visual structures. However, the performance of these methods is strongly linked to the accuracy of the parameters involved in the PDEs, which paves the way for an intensive exploration of optimization and parameter identification techniques [3, 7, 12].

*Corresponding author
Received: 7 August 2024 / Revised: 5 November 2024 / Accepted: 11 January 2025
DOI: 10.22124/jmm.2025.28007.2482

A classic model of image degradation is given by $u_0 = u + n$, where u is the original image, u_0 is the noisy image, and n represents additive noise. This simplified model describes how noise can corrupt images, presenting a significant challenge for denoising techniques.

To address this challenge, various mathematical approaches have been developed, generally categorized into three main areas: statistical, algebraic, and variational and PDE. Each approach offers unique tools and methods for tackling the denoising problem, each with its own advantages and limitations. Among classical models, the heat equation is often used as an example of linear diffusion:

$$\partial_t u + \Delta u = 0, \quad u(0) = u_0. \tag{1}$$

While this model effectively removes noise; it has a major drawback: it also removes important image details, such as edges and fine textures, which can degrade the quality of the restored images.

To overcome this issue, Perona and Malik [16] proposed a modification to the diffusion equation by introducing a non-constant diffusivity:

$$\partial_t u + \operatorname{div}\left(\frac{1}{1+c|\nabla u|^2}\nabla u\right) = 0, \quad u(0) = u_0. \tag{2}$$

This approach reduces diffusion in regions of high gradient, such as edges, while increasing diffusion in uniform regions. Additionally, Catt et al. [5] suggested scaling the diffusivity using the gradient of a uniformly smoothed image:

$$\partial_t u + \operatorname{div}\left(\frac{1}{1+c|\nabla u_\sigma|^2}\nabla u\right) = 0, \quad u(0) = u_0, \tag{3}$$

where $u_\sigma = K_\sigma * u$ represents the convolution of u with a Gaussian kernel of variance σ^2. This approach helps to stop diffusion at edges, offering an improvement over previous models by better preserving important image details.

A similar model to the Perona-Malik diffusion was developed by Aboulaich et al. in [1] and is given by

$$\begin{cases} \dfrac{\partial u}{\partial t} - \operatorname{div}\left(\left(\frac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha(x)\right)\nabla u\right) + \lambda(u-u_0) = 0 \ \text{ in } \ \Omega\times(0,T), \\ \left\langle \left(\frac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha(x)\right)\nabla u, n\right\rangle = 0 \ \text{ on } \ \partial\Omega\times(0,T), \\ u(x,0) = u_0 \ \text{ in } \ \Omega. \end{cases} \tag{4}$$

The simplified equation resembles the Perona-Malik one. This model incorporates a diffusion term that depends on the image gradient, with $\frac{1}{\sqrt{1+|\nabla u(x)|^2}}\nabla u$ representing a velocity that modulates diffusion based on gradient magnitude.

Despite advancements in these models, parameter selection remains a significant challenge. Manual tuning of parameters such as the diffusivity in these models is often laborious and may not ensure optimal results. This highlights the need for more sophisticated approaches that automate the parameter selection process.

In this context, our work focuses on automating the selection of fundamental parameters within the nonlinear PDEs (4) to improve the quality of noisy image restoration.

We formulate a bilevel optimization problem to automate the selection of parameters α and λ as:

$$\inf_{\alpha,\lambda\in\mathscr{E}_{ad}} \mathbf{J}(\alpha,\lambda) = \frac{1}{2}\int_\Omega |u(x,T)-u_0^\sigma|^2 dx + \gamma_1 \int_\Omega |\alpha(x)|^2 dx + \gamma_2 \int_\Omega |\lambda(x)|^2 dx, \tag{5}$$

subject to:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \operatorname{div}\left(\left(\dfrac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha(x)\right)\nabla u\right) + \lambda(u-u_0) = 0 \ \text{ in } \ \Omega\times(0,T), \\ \left\langle \left(\dfrac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha(x)\right)\nabla u, n\right\rangle = 0 \ \text{ on } \ \partial\Omega\times(0,T), \\ u(x,0) = u_0 \ \text{ in } \ \Omega, \end{cases} \tag{6}$$

where γ_1 and γ_2 are positive constants, and $\mathscr{E}_{ad}$ is the set of admissible functions defined by:

$$\mathscr{E}_{ad} = \{(\alpha,\lambda) : \ \alpha\in \mathbf{U}_{ad}, \ \ \lambda\in\mathscr{U}_{ad} \ \text{ with } \ (\alpha,\lambda) \text{ satisfying } \ (6)\}, \tag{7}$$

with constraints:

$$\mathscr{U}_{ad} = \{\lambda\in L^2(\Omega) : \lambda_a \leqslant \lambda(x) \leqslant \lambda_b \ \ a.e \ in \ \Omega\}, \tag{8}$$

$$\mathbf{U}_{ad} = \{\alpha\in L^2(\Omega) : \alpha_c \leqslant \alpha(x) \leqslant \alpha_d \ \ a.e \ in \ \Omega\}, \tag{9}$$

where α_c, α_d, λ_a, $\lambda_b > 0$ in $\mathbb{R}$, $\alpha_c < \alpha_d$, and $0 < \lambda_a < \lambda_b$.

By automating the parameter selection process, our approach allows for more precise adjustment of denoising processes while preserving essential image details. We propose using the Alternating Direction Method of Multipliers (ADMM) algorithm to solve this problem and validate our approach with numerical results demonstrating significant improvements over traditional denoising methods. This work represents a significant advancement by introducing an automated method for parameter identification in nonlinear PDEs, offering promising prospects for practical applications requiring accurate image restoration.

2 Existence of a solution to the optimization problem

This section aims to establish the existence of a solution to the bilevel optimisation problem (5)-(6), taking into account the constraints imposed on the domain $(\alpha,\lambda)\in\mathscr{E}_{ad}$. To this end, we assume that the space $L^2(\Omega)\times L^2(\Omega)$ is equipped with the following norm:

$$||(\alpha,\lambda)||_{L^2(\Omega)\times L^2(\Omega)} = ||\alpha||_{L^2(\Omega)} + ||\lambda||_{L^2(\Omega)}.$$

Proposition 1. *Suppose that $(\alpha_k,\lambda_k)_{k\in\mathbb{N}}\subset\mathscr{E}_{ad}$ is a minimizing sequence of (5) which converges to the pair $(\alpha^*,\lambda^*)\subset\mathbf{U}_{ad}\times\mathscr{U}_{ad}$ in $L^2(\Omega)\times L^2(\Omega)$, and let u_k denote the solution of (6) for all $k\in\mathbb{N}$, where $u_k\equiv u(\alpha_k,\lambda_k)$, then*
1. there is a subsequence $(u_k)_k$ that converges almost everywhere to u^*, which is a solution of equation (6).
2. the functional cost $\mathbf{J}$ exhibits weak lower semi-continuity and fulfills the following property:

$$\mathbf{J}(\alpha^*,\lambda^*) \leqslant \liminf_{k\to\infty} \mathbf{J}(\alpha_k,\lambda_k).$$

Proof. 1. Since (α_k, λ_k) is a minimizing sequence of (5), we have

$$\lim_{k\to\infty} \mathbf{J}(\alpha_k, \lambda_k) = \inf_{(\alpha,\lambda)\in\mathscr{E}_{ad}} \mathbf{J}(\alpha, \lambda).$$

On the other hand, we have

$$\begin{aligned}\mathbf{J}(\alpha_k, \lambda_k) &> \gamma_1 \int_\Omega |\alpha_k(x)|^2 dx + \gamma_2 \int_\Omega |\lambda_k(x)|^2 dx \\ &> \min(\gamma_1, \gamma_2)(||\alpha_k||^2_{L^2(\Omega)} + ||\lambda_k||^2_{L^2(\Omega)}).\end{aligned}$$

Therefore, there is $\mu = \min(\gamma_1, \gamma_2) > 0$, such that

$$\mathbf{J}(\alpha_k, \lambda_k) > \mu \left(||\alpha_k||^2_{L^2(\Omega)} + ||\lambda_k||^2_{L^2(\Omega)} \right),$$

hence $\exists M > 0$, such that

$$||\alpha_k||^2_{L^2(\Omega)} + ||\lambda_k||^2_{L^2(\Omega)} \leqslant M.$$

As a result, the sequence $(\alpha_k, \lambda_k)_k$ is bounded in $L^2(\Omega) \times L^2(\Omega)$. Since $L^2(\Omega) \times L^2(\Omega)$ is reflexive, there exists a subsequence, again denoted as (α_k, λ_k), such that

$$\alpha_k \rightharpoonup \alpha^* \quad \text{in} \quad L^2(\Omega)$$

and

$$\lambda_k \rightharpoonup \lambda^* \ \text{ in } L^2(\Omega). \tag{10}$$

Since $(\lambda_k)_k$ is a sequence of $\mathscr{U}_{ad}$, we have

$$\lambda_a \leqslant \lambda_k \leqslant \lambda_b \ \text{ a.e } \ \text{in } \Omega.$$

For a measurable subset B of Ω, we have

$$\int_B \lambda_a dx \leqslant \int_B \lambda_k(x) dx \leqslant \int_B \lambda_b dx.$$

So

$$\int_B \lambda_a dx \leqslant \int_\Omega \lambda_k(x) 1_B dx \leqslant \int_B \lambda_b dx.$$

According to (10) and when $k \to \infty$

$$\int_B \lambda_a dx \leqslant \int_\Omega \lambda^*(x) 1_B dx \leqslant \int_B \lambda_b dx.$$

Then

$$\int_B \lambda_a dx \leqslant \int_B \lambda^*(x) dx \leqslant \int_B \lambda_b dx.$$

Thus

$$\lambda_a \leqslant \lambda^* \leqslant \lambda_b \ \text{ a.e } \ \text{in } \Omega. \tag{11}$$

Therefore $\lambda^* \in \mathscr{U}_{ad}$. Similary, we have $\alpha^* \in \mathbf{U}_{ad}$.

Consider the sequence $(\alpha_k,\lambda_k)_k \in \mathscr{E}_{ad}$, where for each (α,λ_k), the corresponding solution u_k satisfies equation (6). Consequently, for every $\psi \in H_0^1(\Omega)$:

$$\left\langle \frac{\partial u_k}{\partial t},\psi \right\rangle_{(H^1(\Omega))',H^1(\Omega)} + \int_Q \left(\frac{1}{\sqrt{1+|\nabla u_k(x)|^2}} + \alpha_k(x) \right) \nabla u_k \nabla \psi dxdt \tag{12}$$

$$+ \int_Q \lambda_k (u_k - u_0)\psi dxdt = 0. \tag{13}$$

Let $t \in (0,T)$ and choose $\psi = u_k(t)$ as test function in the variational formulation (13), we get

$$\frac{1}{2}\int_\Omega u_k^2(x,t)dx - \frac{1}{2}\int_\Omega u_0^2(x)dx + \int_0^t \int_\Omega \left(\frac{1}{\sqrt{1+|\nabla u_k(x)|^2}} + \alpha_k(x) \right) (\nabla u_k)^2 dxdt \tag{14}$$

$$+ \int_0^t \int_\Omega \lambda_k u_k^2(x,t)dxdt = \int_0^t \int_\Omega \lambda_k u_k(x,t)u_0(x)dxdt, \tag{15}$$

By using Young's inequality and the fact that $\alpha_k \in \mathbf{U}_{ad}$ and $\lambda_k \in \mathscr{U}_{ad}$, we can show that

$$\begin{aligned}
\frac{1}{2}\int_\Omega u_k^2(x,t)dx - \frac{1}{2}\int_\Omega u_0^2(x)dx + \int_0^t \int_\Omega \left(\frac{1}{\sqrt{1+|\nabla u_k(x)|^2}} + \alpha_k \right) (\nabla u_k)^2 dxdt & \\
+ \lambda_a \int_0^t \int_\Omega u_k^2(x,t)dxdt & \\
\leqslant \lambda_b \int_0^t \int_\Omega u_k(x,t)u_0(x)dxdt & \\
\leqslant \frac{\lambda_b}{2}\int_0^t \int_\Omega u_k^2(x,t)dxdt + \frac{\lambda_b}{2}\int_0^t \int_\Omega u_0^2(x,t)dxdt, &
\end{aligned}$$

then

$$\int_\Omega u_k^2(x,t)dx \leqslant (T\lambda_b + 1)\int_\Omega u_0^2(x,t)dx + \lambda_b \int_0^t \int_\Omega u_k^2(x,t)dxdt.$$

By Gronwall inequality

$$\int_\Omega u_k^2(x,t)dx \leqslant \left((T\lambda_b + 1)\int_\Omega u_0^2(x)dx \right) \exp(t\lambda_b) \ \ \forall t \in (0,T),$$

this implies that

$$\int_\Omega u_k^2(x,t)dx \leqslant C, \tag{16}$$

for some constant $C > 0$ depending only on T and λ_b. Then $||u_k||_{L^\infty(0,T;L^2(\Omega))} \leqslant C$. The function $\mu : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ defined as

$$\mu(|\nabla u|) = \frac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha_c,$$

is bounded (see [1]). Hence, we have $\int_0^t ||u_k(\tau)||^2 d\tau \leqslant C, \quad \forall t \in [0,T]$. The operator A defined by

$$(Au_k,\psi) = \int_\Omega \left(\frac{1}{\sqrt{1+|\nabla u_k(x)|^2}} + \alpha_c \right) \nabla u_k \nabla \psi dx,$$

is a monotone operator. We deduce that the approximate solution u_k of the problem (13) converges to a weak solution u of problem (6).

2- We prove that $\mathbf{J}$ obeys the inequality

$$\liminf_{k\to\infty} \mathbf{J}(\alpha_k,\lambda_k) \geqslant \mathbf{J}(\alpha^*,\lambda^*),$$

where α^* and λ^* are the optimal solutions of the optimization problem and α_k and λ_k are the solutions obtained by an iterative algorithm with step size k. We have

$$\lim_{k\longrightarrow+\infty} \frac{1}{2}\int_\Omega |u_k(x,T)-u_0^\sigma(x)|^2dx = \frac{1}{2}\int_\Omega |u^*(x,T)-u_0^\sigma(x)|^2dx,$$

$$\frac{\gamma_1}{2}\int_\Omega |\alpha^*(x)|^2dx \leqslant \liminf_{k\longrightarrow+\infty} \frac{\gamma_1}{2}\int_\Omega |\alpha_k(x)|^2dx,$$

and

$$\frac{\gamma_2}{2}\int_\Omega |\lambda^*(x)|^2dx \leqslant \liminf_{k\longrightarrow+\infty} \frac{\gamma_2}{2}\int_\Omega |\lambda_k(x)|^2dx,$$

then

$$\begin{aligned}
\mathbf{J}(\alpha^*,\lambda^*) &= \frac{\gamma_2}{2}\int_\Omega |\lambda^*(x)|^2dx + \frac{\gamma_1}{2}\int_\Omega |\alpha^*(x)|^2dx + \frac{1}{2}\int_\Omega |u^*(x,T)-u_0^\sigma(x)|^2dx \\
&\leqslant \liminf_{k\longrightarrow+\infty} \frac{1}{2}\int_\Omega |u_k(x,T)-u_0^\sigma(x)|^2dx + \frac{\gamma_2}{2}\liminf_{k\longrightarrow+\infty}\int_\Omega |\lambda_k(x)|^2dx \\
&+ \frac{\gamma_1}{2}\liminf_{k\longrightarrow+\infty}\int_\Omega |\alpha_k(x)|^2dx \\
&\leqslant \liminf_{k\longrightarrow+\infty}\left(\frac{1}{2}\int_\Omega |u_k(x,T)-u_0^\sigma(x)|^2dx + \frac{\gamma_2}{2}\int_\Omega |\lambda_k(x)|^2dx + \frac{\gamma_1}{2}\int_\Omega |\alpha_k(x)|^2dx\right) \\
&\leqslant \liminf_{k\longrightarrow+\infty} \mathbf{J}(\alpha_k,\lambda_k).
\end{aligned}$$

□

The previous proposition provides evidence to support the following theorem.

Theorem 1. *The bilevel optimization problem (5) has a solution* (α,λ) *in* $\mathscr{E}_{ad}$.

3 The proposed algorithm and numerical implementation.

3.1 Algorithm

In order to obtain an approximation of the solution to our problem (5), we develop an algorithm that aims to overcome the challenges inherent in this problem, and we will analyze in depth the key features of this algorithm. The algorithmic procedure outlined in this section is presented as a numerical optimization method specifically designed to solve a well-defined problem. In 1976, Gabay, Mercier, Glowinski, and Marrocco introduced the ADMM [10, 14]. This method combines the robustness of the method of multipliers with the ability to support decomposition. Additionally, the ADMM has proven to be a powerful tool for solving a wide range of constrained optimization problems [11, 14, 15, 20]. In the following

sections, we directly apply the ADMM to our parabolic optimal control problem with constraints (5)-(6). The ADMM can be viewed as a fractional version of the classical augmented Lagrangian method; for further insights, we refer to [20]. This approach allows us to partition the original problem into easily solvable sub-problems, resulting in greater efficiency compared to the augmented Lagrangian method, which utilizes Gauss-Seidel techniques at each iteration.

In many instances, pinpointing the exact parameters becomes challenging due to the presence of numerous parameters, the intricate nature of the image, and the increased computational time with higher iterations. Given that an image comprises pixels, any parameter chosen manually applies uniformly to all pixels. Therefore, methods need to either identify these parameters or determine suitable parameters for each pixel to achieve optimal image restoration. This process can be encapsulated by introducing the operator G, which acts as an affine solution operator linked with the state equation (6), as follows:

$$\begin{aligned} G : \mathbf{U}_{ad} \times \mathscr{U}_{ad} &\longrightarrow L^2(\Omega) \\ (\alpha,\lambda) &\mapsto G(\alpha,\lambda) := u(x,T). \end{aligned}$$

The operator G is bounded, continuous and compact. For $u = G(\alpha,\lambda)$, the problem (5)-(6) can be reformulated as the following optimal control problem:

$$\inf_{\alpha,\lambda \in \mathbf{U}_{ad}\times\mathscr{U}_{ad}} \mathscr{G}(G(\alpha,\lambda)) + \mathscr{R}(\alpha,\lambda),$$

with $\mathscr{G}$ is the fidelity term given by

$$\mathscr{G}(G(\alpha,\lambda)) = \frac{1}{2}\int_\Omega |G(\alpha,\lambda) - u_0^\sigma(x)|^2 dx,$$

$\mathscr{R}$ is the regularization term defined by

$$\mathscr{R}(\alpha,\lambda) = \frac{\gamma_1}{2}\int_\Omega |\alpha(x)|^2 dx + \frac{\gamma_2}{2}\int_\Omega |\lambda(x)|^2 dx,$$

and γ_1 and γ_2 are two regularization parameters. Now, we consider the auxiliary variables $y_1, y_2 \in L^2(\Omega)$ such that $(\alpha,\lambda) = (y_1,y_2)$, consequently the problem (5)-(6) becomes

$$\begin{cases} \min\limits_{\alpha,\lambda,y_1,y_2} \mathbf{J}(\alpha,\lambda) + \mathscr{I}_{\mathbf{U}_{ad}}(y_1) + \mathscr{I}_{\mathscr{U}_{ad}}(y_2) \\ \text{subjet to } \ (\alpha,\lambda) = (y_1,y_2), \end{cases} \tag{17}$$

where $\mathscr{I}_K(.)$ is the indicator functional of the set K defined by

$$\mathscr{I}_K(y) := \begin{cases} 0, & \text{if } y \in K \\ +\infty & \text{if } y \in L^2(\Omega)\backslash K, \end{cases}$$

and

$$\mathbf{J}(\alpha,\lambda) = \mathscr{G}(G(\alpha,\lambda)) + \mathscr{R}(\alpha,\lambda).$$

To start applying the ADMM to the problem (5)-(6), we need to define the augmented Lagrangian functional associated with the minimization problem (17) as

$$\begin{aligned} L(\alpha,\lambda,y_1,y_2,\beta_1,\beta_2,\mu_1,\mu_2) = \mathbf{J}(\alpha,\lambda) + \mathscr{I}_{\mathbf{U}_{ad}}(y_1) + \mathscr{I}_{\mathscr{U}_{ad}}(y_2) - (\beta_1,\alpha-y_1) \\ + \frac{\mu_1}{2}||\alpha-y_1||^2 - (\beta_2,\lambda-y_2) + \frac{\mu_2}{2}||\lambda-y_2||^2, \end{aligned} \tag{18}$$

with μ_1 and μ_2 are positive penalty parameters and β_1, $\beta_2 \in L^2(\Omega)$ are the Lagrange multipliers associated to the constraint $(\alpha,\lambda) = (y_1,y_2)$. Otherwise, finding the saddle point of the function (18) is equivalent to obtaining the minimum of the problem (5)-(6). Then leads to ADMM steps:

$$\begin{cases} \alpha^{k+1} = \arg\min_{\alpha} L(\alpha,\lambda^k,y_1^k,y_2^k,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) & (19a)\\ \lambda^{k+1} = \arg\min_{\lambda} L(\alpha^k,\lambda,y_1^k,y_2^k,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) & (19b)\\ y_1^{k+1} = \arg\min_{y_1} L(\alpha^{k+1},\lambda^{k+1},y_1,y_2^k,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) & (19c)\\ y_2^{k+1} = \arg\min_{y_2} L(\alpha^{k+1},\lambda^{k+1},y_1^k,y_2,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) & (19d)\\ \beta_1^{k+1} = \beta_1^k - \alpha_1(\alpha^{k+1}-y_1^{k+1}) & (19e)\\ \beta_2^{k+1} = \beta_2^k - \alpha_1(\lambda^{k+1}-y_2^{k+1}) & (19f) \end{cases}$$

In what follows, we are interested in solving the above minimization subproblems. The subproblems (19a) and (19b) are equivalent to the following unconstrained optimal control problems

$$\min_{\gamma\in L^2(\Omega)} \mathbf{J}(\alpha,\lambda^k) - (\beta_1^k,\alpha-y_1^k) + \frac{\mu_1}{2}||\alpha-y_1^k||^2, \tag{20}$$

$$\min_{\lambda\in L^2(\Omega)} \mathbf{J}(\alpha^k,\lambda) - (\beta_2^k,\lambda-y_2^k) + \frac{\mu_2}{2}||\lambda-y_2^k||^2. \tag{21}$$

To solve the minimization (20) and (21), we need to compute the gradients of

$$\Gamma(\alpha) = \mathbf{J}(\alpha,\lambda^k) - (\beta_1^k,\gamma-y_1^k) + \frac{\mu_1}{2}||\alpha-y_1^k||^2,$$

$$\Lambda(\lambda) = \mathbf{J}(\alpha^k,\lambda) - (\beta_2^k,\lambda-y_2^k) + \frac{\mu_2}{2}||\lambda-y_2^k||^2,$$

which are given as follows:

$$\begin{cases} D\Gamma(\alpha) = \gamma_1\alpha + \left[\frac{\partial G}{\partial\alpha}\right]^*\big(G(\alpha,\lambda^k)\big) - \beta_1^k + \mu_1(\alpha-y_1^k),\\ D\Lambda(\lambda) = \gamma_2\lambda + \left[\frac{\partial G}{\partial\lambda}\right]^*\big(G(\alpha^k,\lambda)\big) - \beta_2^k + \mu_2(\lambda-y_2^k), \end{cases} \tag{22}$$

where $[\mathscr{Z}]^*$ is the adjoint operator of $\mathscr{Z}$ and the adjoint state equation is given as following

$$(\mathscr{AE}) \begin{cases} -\frac{\partial w}{\partial t} + \operatorname{div}\left(\left(\frac{|\nabla u|^2}{(1+|\nabla u(x)|^2)^{\frac{3}{2}}} - \frac{1}{\sqrt{1+|\nabla u(x)|^2}} - \alpha(x)\right)\nabla w\right) - \lambda w = 0, \ \ on\ ]0,T[\times\Omega,\\ w(T,x) = u_0(x) - u(x,T), \ \ on\ \ ]0,T[\times\partial\Omega,\\ u(0,x) = u_0(x). \end{cases} \tag{23}$$

For the subproblems (19c) and (19d), we have

$$\begin{cases} L(\gamma^{k+1},\lambda^{k+1},y_1,y_2^k,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) = J(\gamma^{k+1},\lambda) + \mathscr{I}_{H^1(\Omega)}(y_1) - (\beta_1^k,\gamma^{k+1}-y_1)\\ \qquad\qquad + \frac{\mu_1}{2}||\gamma^{k+1}-y_1||^2,\\ L(\gamma^{k+1},\lambda^{k+1},y_1^k,y_2,\beta_1^k,\beta_2^k,\mu_1^k,\mu_2^k) = J(\gamma,\lambda^{k+1}) + \mathscr{I}_{\Lambda_{ad}}(y_2) - (\beta_2^k,\lambda^{k+1}-y_2)\\ \qquad\qquad + \frac{\mu_2}{2}||\lambda^{k+1}-y_2||^2, \end{cases} \tag{24}$$

and their solution can be computed by the projection on the admissible set $\mathbf{U}_{ad}$ and $\mathscr{U}_{ad}$, hence

$$\begin{cases} y_1^{k+1} = arg \min\limits_{y_1 \in L^2(\Omega)} \mathscr{I}_{\mathbf{U}_{ad}}(y_1) - (\beta_1^k, \alpha^{k+1} - y_1) + \dfrac{\mu_1}{2}||\alpha^{k+1} - y_1||^2, \\ y_2^{k+1} = arg \min\limits_{y_2 \in L^2(\Omega)} \mathscr{I}_{\mathscr{U}_{ad}}(y_2) - (\beta_2^k, \lambda^{k+1} - y) + \dfrac{\mu_2}{2}||\lambda^{k+1} - y_2||^2, \end{cases} \tag{25}$$

so,

$$\begin{cases} y_1^{k+1} = \mathbf{Proj}_{\mathbf{U}_{ad}}\left(\alpha^{k+1} - \frac{\beta_1^k}{\mu_1}\right) = \max\left(\alpha_c, \min\left(\alpha_d, \alpha^{k+1} - \frac{\beta_1^k}{\mu_1}\right)\right), \\ y_2^{k+1} = \mathbf{Proj}_{\mathscr{U}_{ad}}\left(\lambda^{k+1} - \frac{\beta_2^k}{\mu_2}\right) = \max\left(\lambda_b, \min\left(\lambda_a, \lambda^{k+1} - \frac{\beta_2^k}{\mu_2}\right)\right), \end{cases} \tag{26}$$

with $\mathbf{Proj}_P(.)$ is the projection operator in P. Hence, the ADMM algorithm for the problem (5)-(6) is summarized in Algorithm 1.

Algorithm 1 ADMM algorithm.

1: **Choose** u_0,α^0, λ^0, y_1^0, y_2^0, β_1^0, β_2^0, γ_1^0,γ_2^0, α_c, α_d, λ_a, λ_b, τ_1^0 and τ_2^0.

2: **For** $k = 0, ..., N$, **compute**

3: u^k solution of problem (6).

4: w^k solution of problem (23).

5: $D\Gamma(\alpha)$ and $D\Lambda(\lambda)$ from the equation (22):

6: **Update** α^{k+1} and λ^{k+1} by

$$\begin{cases} \alpha^{k+1} = \alpha^k - \tau_1^k D\Gamma(\alpha \\ \lambda^{k+1} = \lambda^k - \tau_2^k D\Lambda(\lambda), \end{cases} \tag{27}$$

with τ_1^k and τ_2^k being computed by the Armijo line search algorithm.

8: Update y_1^{k+1} and y_2^{k+1} by

$$\begin{cases} y_1^{k+1} = \max\left\{\underline{\alpha}, \min\left\{\alpha^{k+1} - \frac{\beta_1^k}{\mu_1}, \overline{\alpha}\right\}\right\} \\ y_2^{k+1} = \max\left\{\underline{\lambda}, \min\left\{\lambda^{k+1} - \frac{\beta_2^k}{\mu_2}, \overline{\lambda}\right\}\right\} \end{cases} \tag{28}$$

9: Update the Lagrange multipliers: β_1^{k+1} and β_2^{k+1} by

$$\begin{cases} \beta_1^{k+1} = \beta_1^k - \mu_1(\alpha^{k+1} - y_1^{k+1}) \\ \beta_2^{k+1} = \beta_2^k - \mu_2(\lambda^{k+1} - y_2^{k+1}) \end{cases} \tag{29}$$

10: end

3.2 Convergence

The convergence of the ADMM algorithm for a PDE-constrained optimization problem has been treated in the literature [11, 15], where the cost function was assumed to be convex. Here, we only focus on proving the convexity of the cost function **J**. In following, we establish the convexity of the function **J** with respect to the variable α, assuming that the parameter λ to be constant. The convexity with respect to λ can then be justified in a similar way. Firstly, we suppose that the solution u is differentiable, which can be proved in the same manner as in [13]. And the differentiable function $u^{'} = DG(\alpha)h$ with $h \in \mathbf{U}_{ad}$, is the unique solution to the following equation

$$\begin{cases} \dfrac{\partial u^{'}}{\partial t} - \operatorname{div}\left(g(u,\alpha+h)\nabla u^{'}\right) - div\left(\dfrac{\partial g(u,\alpha)}{\partial \alpha}h\nabla u\right) + \lambda u^{'} = 0 \quad \text{in} \quad \Omega\times(0,T), \\ \left\langle g(u,\alpha+h)\nabla u^{'} + \dfrac{\partial g(u,\alpha)}{\partial \alpha}h\nabla u, n\right\rangle = 0 \quad \text{on} \quad \partial\Omega\times(0,T), \\ u^{'}(x,0) = 0 \quad \text{in} \quad \Omega, \end{cases} \tag{30}$$

where

$$g(u,\alpha) = \left(\frac{1}{\sqrt{1+|\nabla u(x)|^2}} + \alpha(x)\right).$$

Now, let us denote $J(\alpha) = \mathbf{J}(\alpha,\lambda)$. For any $\alpha \in \mathbf{U}_{ad}$, and h perturbation, we can show that

$$J'(\alpha)h = \int_{\Omega_1}(u(x,T) - u_0^{\sigma})u' dxdt + \gamma_1\int_{\Omega}\alpha h dx. \tag{31}$$

In order to establish the convexity of J, we need the following two lemmas.

Lemma 1. Let $\alpha \in \mathbf{U}_{ad}$ and, $h \in L^2(\Omega\times(0,1))$, u_h and u be the solution of (5) associated to $\alpha+h$ and α, respectively. Then

$$\|u_h - u\|_{L^{\infty}(0,1;L^2(\Omega))} + \|\nabla u_h - \nabla u\|_{L^2(0,1;L^2(\Omega))} \leq CT\|h\|^2_{L^2(\Omega\times(0,1))}.$$

Lemma 2. Let $h \in L^2(\Omega\times(0,1))$ such that $\alpha+h \in \mathbf{U}_{ad}$ and, let u_h and u be the solutions of the problem (5) associated to $\alpha+h$ and α, respectively. And let $u^{'}$ be the unique solution of (30), then

$$\|u_h - u - u'\|^2_{L^{\infty}(0,1;L^2(\Omega))} \leq CT\|h\|^2_{L^2(\Omega\times(0,1))}. \tag{32}$$

The proofs of these lemmas can be found easily, following a similar approach to that in [13]. We justify the convexity of the functional cost **J** in the following theorem.

Theorem 2. *For any γ_1, the functional* **J** *is strictly convex.*

Proof. Let α and $\alpha+h$ be in $\mathbf{U}_{ad}$ and let u and u_h be the associated solutions, respectively. We have

$$\begin{aligned} J(\alpha+h) - J(\alpha) - J'(\alpha)h = &\frac{1}{2}\int_{\Omega}|u_h(x,T) - u_0^{\sigma}(x)|^2 dx + \frac{\gamma_1}{2}\int_{\Omega}|\alpha+h|^2 dx - \frac{1}{2}\int_{\Omega}|u(x,T) - u_0^{\sigma}(x)|^2 dx \\ &- \frac{\gamma_1}{2}\int_{\Omega}|\alpha|^2 dx - \int_{\Omega_1}(u(x,T) - u_0^{\sigma})u' dx - \gamma_1\int_{\Omega}\alpha h dx. \end{aligned}$$

After a simple calculation, we can show

$$J(\alpha+h)-J(\alpha)-J'(\alpha)h=\frac{1}{2}\int_\Omega \left(u_h(x,T)+u(x,T)-2u_0^\sigma(x)\right)\left(u_h(x,T)-u(x,T)\right)dx \\ -\int_\Omega (u(x,T)-u_0^\sigma)u'dx+\frac{\gamma_1}{2}\int_\Omega |h|^2dx.$$

It is easy to see that the last term is non-negative, it remains to show that the difference between the first terms is positive, then

$$J(\alpha+h)-J(\alpha)-J'(\alpha)h=J_1+J_2+J_3+\frac{\gamma_1}{2}\int_\Omega |h|^2dx, \tag{33}$$

where

$$J_1=\frac{1}{2}\int_\Omega \left(u_h(x,T)-u(x,T)-u'(x,T)\right)\left(u(x,T)-u_0^\sigma\right)dx, \tag{34}$$

$$J_2=\frac{1}{2}\int_\Omega \left(u_h(x,T)-u(x,T)-u'(x,T)\right)\left(u_h(x,T)-u_0^\sigma\right)dx \tag{35}$$

and

$$J_3=\frac{1}{2}\int_\Omega \left(u_h(x,T)-u(x,T)\right)u'(x,T)dx. \tag{36}$$

Using the Cauchy-Schwartz inequality and Lemmas 1 and 2, we obtain

$$|J_1|\leq C_1T\|h\|^2_{L^2(\Omega\times(0,1))}, \tag{37}$$

$$|J_2|\leq C_2T\|h\|^2_{L^2(\Omega\times(0,1))} \tag{38}$$

and

$$|J_3|\leq C_3T\|h\|^2_{L^2(\Omega\times(0,1))}. \tag{39}$$

By (37), (38) and (39) we conclude that

$$J(\alpha+h)-J(\alpha)-J'(\alpha)h>\left(\frac{\gamma_1}{2}-CT\right)\|h\|^2_{L^2(\Omega\times(0,1))}. \tag{40}$$

Thus, for $T<\dfrac{\gamma_1}{2C}$, the functional J is strictly convex. □

Concerning the convexity of the function F, where $F(\lambda)=\mathbf{J}(\alpha,\lambda)$, it can be proven similarly the convexity for the variable α. And we can prove that

$$F(\lambda+h)-F(\lambda)-F'(\lambda)h>\left(\frac{\gamma_2}{2}-CT\right)\|h\|^2_{L^2(\Omega\times(0,1))}.$$

For $T<\dfrac{\gamma_2}{2C}$, the functional F is strictly convex.

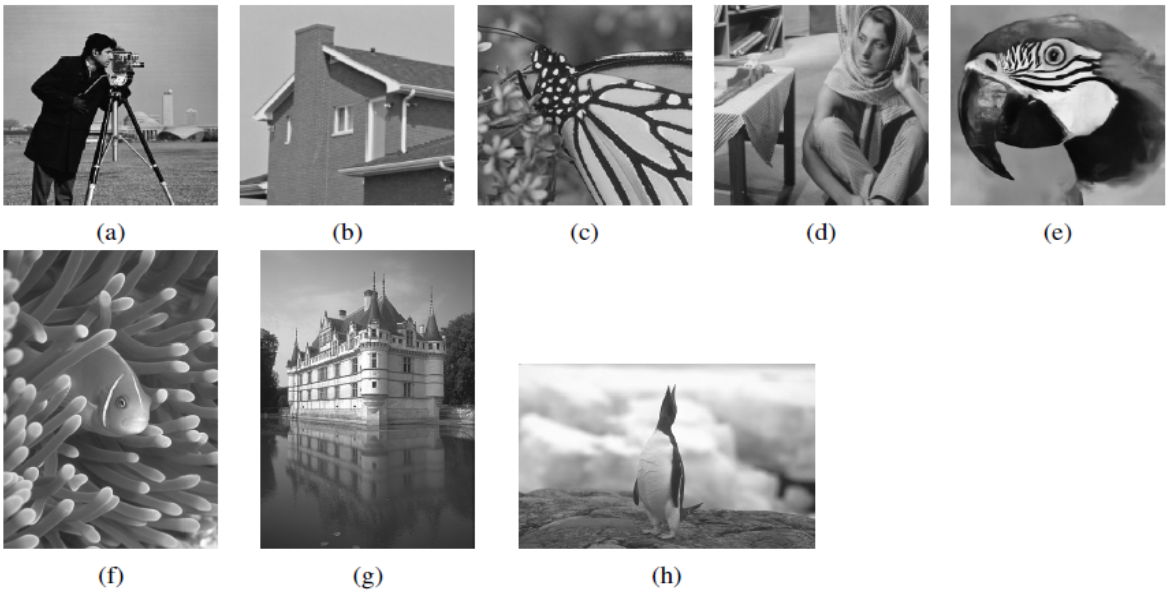

Figure 1: The test images were used for comparison experiments.

3.3 Experimental results

In the digital section, we present some experimental results to demonstrate the effectiveness of the proposed model. We compare it with several competing noise reduction methods to highlight its performance. Note that we tested our model using the classical datasets BSD68 and Set12, some images of which we show here in Figure 1.

We evaluate our results against various competing denoising techniques, including an optimal variable exponent bilevel model as described in the work [12], the TGV (Total Generalized Variation) model [4], the Perona Malik (PM) method proposed in [16], and the approach based on the total variation function TV.

The peak signal-noise ratio (PSNR) (the ratio between the maximum value of the image and the mean square error between the original image and the restored image) and (SSIM) (calculated on multiple windows of given image) are intended to measure the quality of the images obtained, which are defined as follows

$$PSNR = 10\log_{10}\left(\frac{255^2}{mse}\right), \tag{41}$$

with the *mse* is the mean squared error giving as follows

$$mse = \frac{1}{nm}\sum_{i=1}^{m}\sum_{j=1}^{n}(u(i,j)-v(i,j))^2,$$

where u be the noise-free image of size $m \times n$ and v the restored image. Also

$$SSIM(u,v) = \frac{(2\sigma_u\sigma_v + r_2)\,(2\mu_u\mu_v + r_1)\,(2\,\mathrm{cov}\,uv + r3)}{(\mu_u^2+\mu_v^2+r_1)\,(\sigma_u^2+\sigma_v^2+r_2)\,(\sigma_u\sigma_v+r_3)} \tag{42}$$

with cov_{uv} is covariance and $r_1 = (m_1 l)^2, r_2 = (m_2 l)^2$ are two stabilizing constants; and l the dynamics of the pixel values. μ_u and μ_v; σ_u^2 and σ_v^2 are mean; variance of u and v, respectively.

Using various regularization techniques, we generated the diagrams shown in Figures 2-3, providing an overview of the evolution of PSNR and SSIM for reconstructed outputs over 200 iterations, with a fixed noise level of $\sigma_{noise} = 25$. These measures enable us to assess the quality of the reconstructions as we apply the different denoising methods.

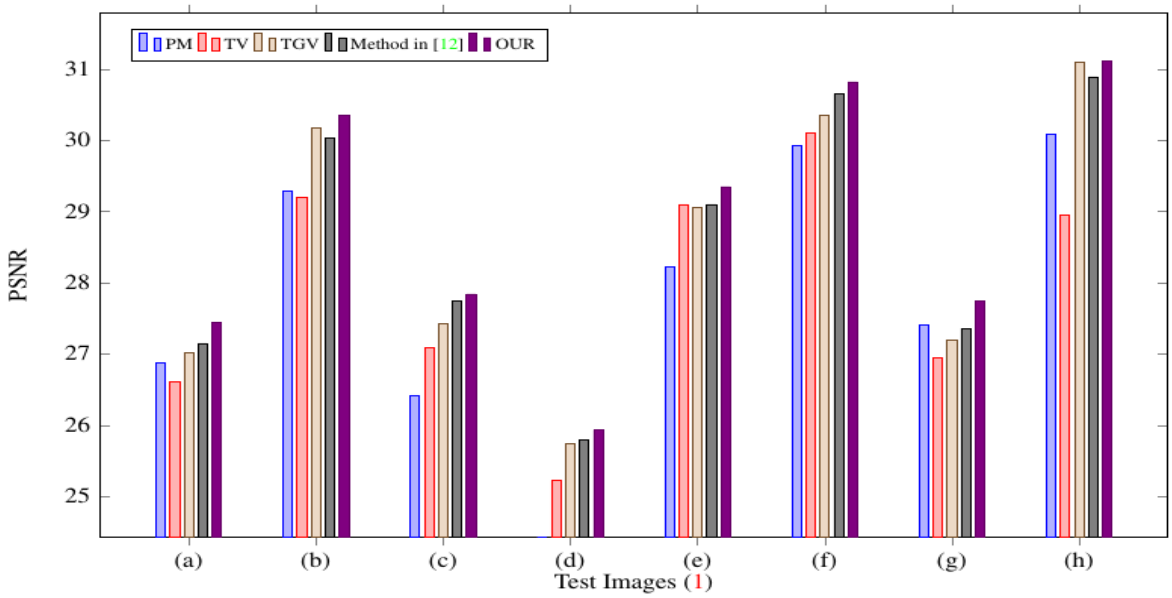

Figure 2: Histogram of PSNR values

Our proposed method has been rigorously tested using a set of reference images, as illustrated in Figure 1. To assess the performance of our method, we carried out a series of simulation experiments in which we introduced Gaussian noise with a degradation factor $\sigma_{noise} = 25$ into each of the original images. In Figures 4, 5 and 6, where $\sigma_{noise} = 25$, and in Figure 7 , where $\sigma_{noise} = 35$, we tested our denoising method alongside various other commonly employed techniques. The results of this evaluation concern the restoration of original images from their degraded versions.

The results indicate that our method is particularly effective in restoring images with high quality and efficiency. Compared with other image denoising approaches, our method stands out for its remarkable restoration quality. The differences are clearly visible to the naked eye, highlighting the power of our image restoration proposal.

To reinforce this qualitative assessment, we also collected digital data using the metrics PSNR and SSIM, as illustrated in Figures 2 and 3 respectively. PSNR measures restoration fidelity by quantifying the quality of the denoised image compared to the original image, while SSIM evaluates the structural similarity between the denoised image and the original image. The figures obtained in these figures convincingly confirm the superior performance of our method over the other denoising methods tested. In particular, our model is able to reduce the staircasing and cartoon-like affects produced by the TV and TGV models (see Figure 7).

In conclusion, our simulation experiments and quantitative evaluations convincingly demonstrate the

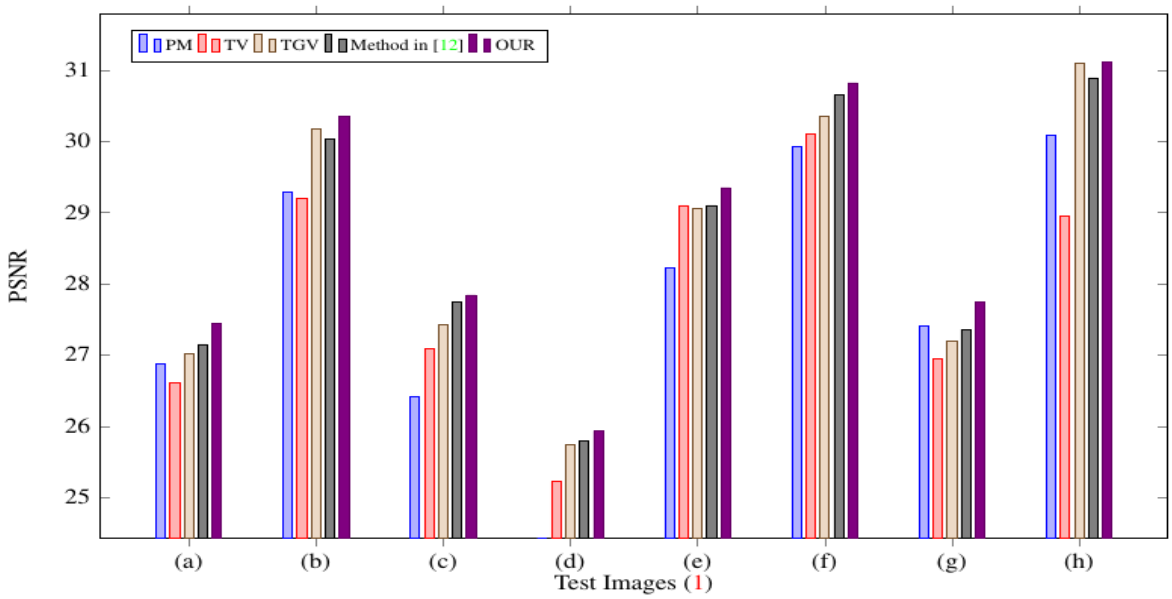

Figure 3: Histogram of SSIM values.

effectiveness and robustness of our image denoising method. Visual results and numerical measurements confirm its ability to restore images exceptionally well, making it a valuable tool for a variety of image processing applications.

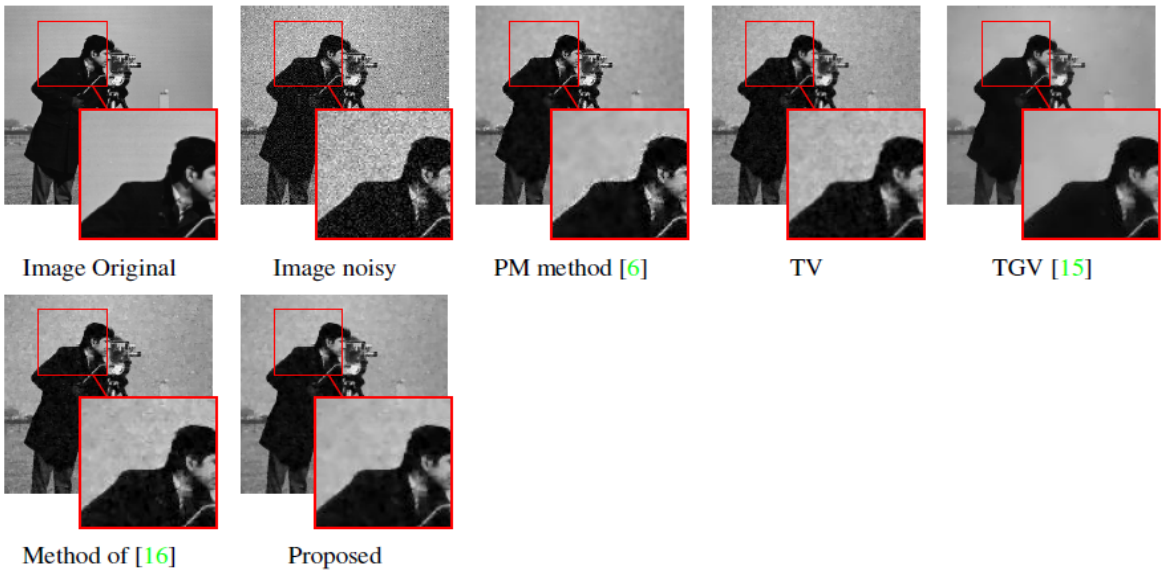

Figure 4: The outcomes achieved through the utilization of various techniques on the House image as contrasted with our methodology ($\sigma_{noisy} = 25$).

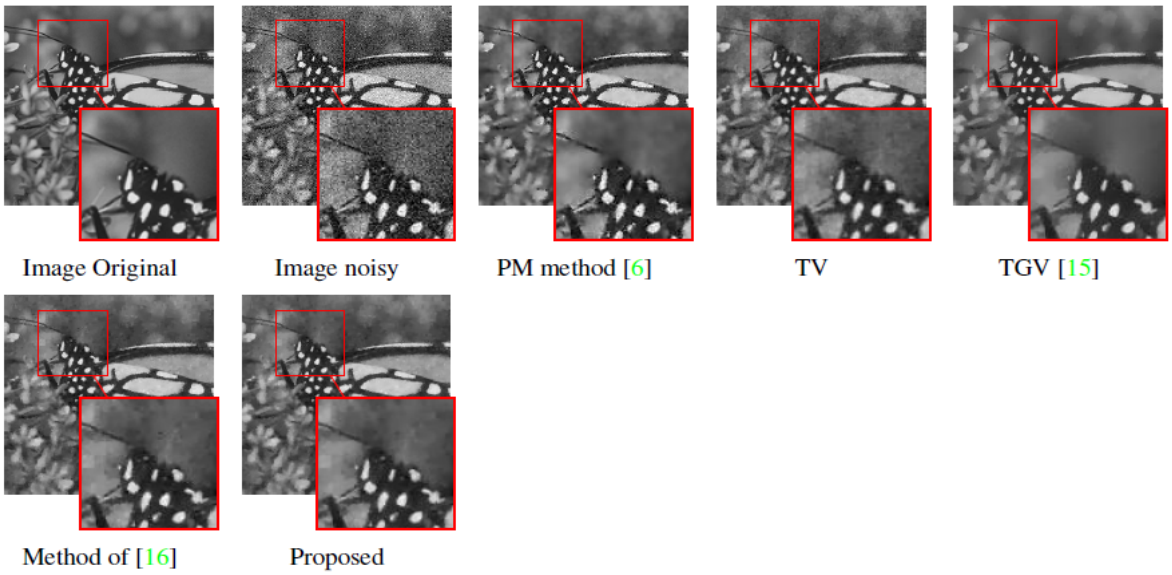

Figure 5: The outcomes achieved through the utilization of various techniques on the Fly image as contrasted with our methodology ($\sigma_{noisy} = 25$).

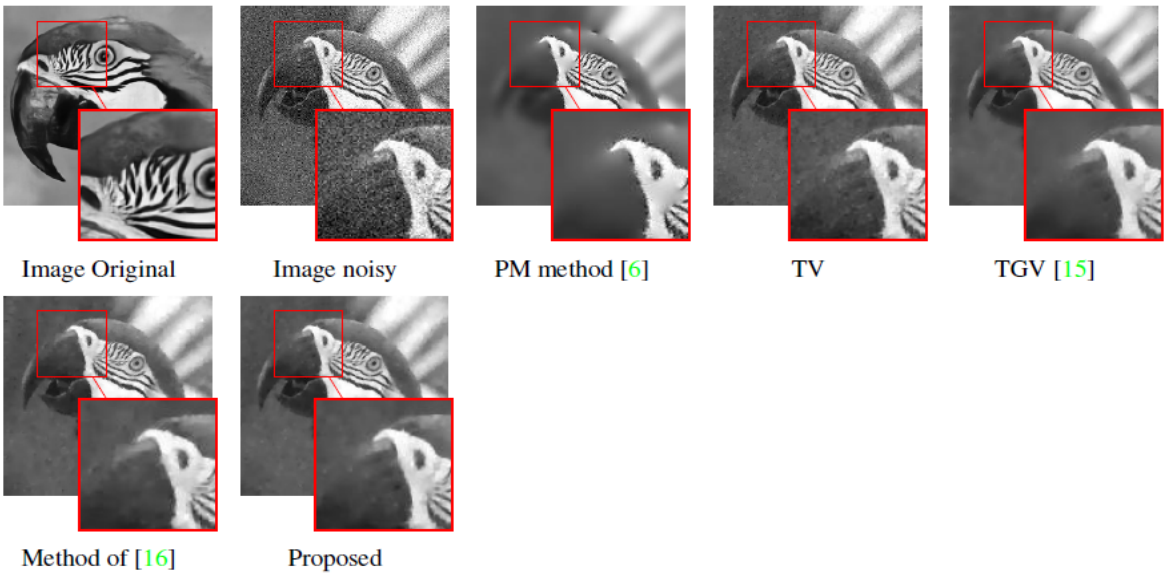

Figure 6: The outcomes achieved through the utilization of various techniques on the Parrots image as contrasted with our methodology ($\sigma_{noisy} = 25$).

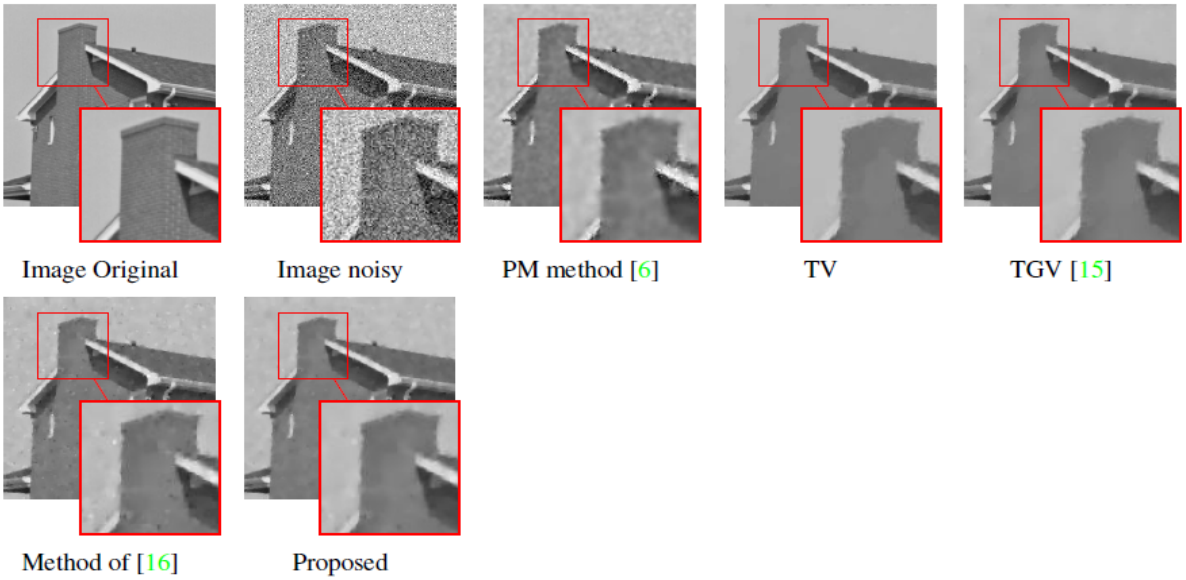

Figure 7: The outcomes achieved through the utilization of various techniques on the House image as contrasted with our methodology ($\sigma_{noisy} = 35$).

Image		PSNR					
	Noise	Noisy	PM [16]	TV	TGV [4]	Method in [12]	Our method
(a)	25	20.16	26.78	26.61	27.02	27.15	27.44
	35	17.30	24.98	25.8	25.83	26.01	26.05
(b)	25	20.13	29.29	29.20	30.17	30.03	30.36
	35	17.25	27.75	28.60	28.65	28.78	28.92

Table 1: The PSNR values.

Image		SSIM					
	Noise	Noisy	PM [16]	TV	TGV [4]	Method in [12]	Our method
(a)	25	0.33	0.78	0.67	0.80	0.72	0.76
	35	0.25	0.74	0.76	0.78	0.70	0.76
(b)	25	0.27	0.79	0.80	0.82	0.78	0.83
	35	0.19	0.76	0.78	0.80	0.77	0.80

Table 2: The SSIM values.

Tables 1 and 2 provide a detailed analysis of the performance of our image denoising model, showcasing numerical values for the PSNR and the SSIM. These evaluations were conducted under conditions

of high Gaussian noise with degradation factors of $\sigma_{noise} = 25$ and $\sigma_{noise} = 35$. These high noise parameters are often encountered in real-world scenarios, where image restoration is particularly challenging. One of the key features of our model is its ability to automatically select regularization parameters λ and α based on input data. This automatic selection is crucial as it allows our model to adapt to varying noise levels without requiring laborious manual calibration. The results presented in Tables 1 and 2 once again confirm the outstanding performance of our model. Our PSNR and SSIM values significantly outperform those achieved by other well-established image denoising approaches, as referenced in studies [16], [4], [12]. This indicates that our model is capable of restoring images even under extremely high noise conditions, thereby preserving the details and fidelity of the original image. These remarkable performances suggest that our model is particularly well-suited for demanding applications such as object detection, computer vision, and medical imaging.

To demonstrate the robustness of our algorithm when applied to images with different noise levels and features, we make comparisons with other competitive denoising PDEs. In fact, the results obtained are compared with the optimal variable exponent model for image denoising, which in turn has been compared with models involving either automatic parameter identification. We use three images " Cameraman ", " Barbara " and " Castle " to select the proposed PDE constrained by the choice of α and λ variables in space Both tests are performed using different levels of Gaussian noise, with $\sigma_{noisy} = 35$ in Figure 7 and $\sigma_{noisy} = 45$ in Figure 8. The robustness of the proposed PDE is evident, as it effectively avoids various artifacts compared to other methods. The parameters are adjusted based on the highest PSNR achieved. For the [12] approach, the parameters used for the " Cameraman " image are: $p = 0.98$ (initial), $\beta = 1$, $\alpha = 0.00001$, $\lambda = 1$. For our approach, the parameters used are: $\gamma_1 = 10^{-5}$, $\gamma_2 = 10^{-6}$, $\beta_1 = 1$, $\beta_2 = 1$, $\alpha = 10^{-6}$ (initial), $\lambda = 10^{-2}$ (initial), $\alpha_c = 10^{-3}$, $\alpha_d = 10^{-1}$, $\lambda_a = 10^{-4}$, $\lambda_b = 10^{-2}$. In order to more precisely assess the quality of the reconstructions, we present a table showing the elapsed CPU time (in seconds). Table 3 shows the elapsed CPU time (in seconds) for the different denoising comparison approaches. The suggested method stands out for having lower values, confirming its robustness.

Image			Method				
	σ_{noisy}	Metric	PM [16]	TV	TGV [4]	Method in [12]	Our method
Barbara	35	CPU	80	71	83	97	122
Cameraman	35	CPU	72	69	78	95	112
Castle	35	CPU	97	89	102	134	141

Table 3: The elapsed CPU time (in seconds) for the tests above. The best (lowest) score in each row is shown in bold.

In summary, our numerical results underscore the undeniable effectiveness of our image denoising model, especially when faced with high noise levels. The automatic selection of regularization parameters λ and α is a major asset of our approach, and our superior performance compared to other established methods enhances the credibility and relevance of our model across a wide range of applications.

(a) PSNR=15.09, SSIM=0.198 (b) PSNR=15.05, SSIM=0.168 (c) PSNR=15.07, SSIM=0.447

(d) PSNR=24.55, SSIM=0.741 (e) PSNR=25.22, SSIM=0.734 (f) PSNR=23.64, SSIM=0.750

(g) PSNR=25.03, SSIM=0.745 (h) PSNR=25.38, SSIM=0.739 (i) PSNR=23.75, SSIM=0.752

Figure 8: The obtained denoised image compared to other PDE approach [12] with respect to both quality measures *PSNR* and *SSIM*. First row: Noisy images with $\sigma_{noisy} = 45$. Second row: Approach of [12]. Third row: Our approach.

4 Conclusions

In conclusion, this study introduces a novel bilevel optimization approach for solving a nonlinear partial differential equation proposed in [1], aimed at improving image denoising quality by estimating certain parameters of the equation. Our work offers both analytical results, demonstrating the existence of minimizers and numerical approximations of the solution, and numerical results confirming the excellent performance of our method compared to traditional denoising methods. This research represents a significant contribution to image restoration, paving the way for high-quality practical applications in this constantly evolving field.

References

[1] R. Aboulaich, S. Boujena, E. El Guarmah, *A nonlinear parabolic model in processing of medical image*, Math. Model. Nat. Phenom. **6** (2008) 131–145.

[2] R. Aboulaich, D. Meskine, A. Souissi, *New diffusion models in image processing*, Comput. Math. Appl. **56** (2008) 874–882.

[3] H. Antil, Z.W. Di, R. Khatri, *Bilevel optimization, deep learning and fractional Laplacian regularization with applications in tomography*, Inverse Probl. **36** (2020) 064001.

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

[5] F. Catt, P.L. Lions, J.M. Morel, T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal. **29** (1992) 182–193.

[6] C.C. Chang, H. Ju-Yuan, H. Chih-Ping, *An adaptive median filter for image denoising*, 2008 Second International Symposium on Intelligent Information Technology Application, Shanghai, China, (2008) 346-350.

[7] A. El Mourabit, I. El Mourabit, *Adaptive nonlinear diffusion for image restoration via bilevel optimization*, Moroccan J. Pure Appl. Anal. **10** (2024) 220–235.

[8] I. El Mourabit, A. Laghrib, A. Hadri, A. Hakim, *A theoretical study of a bilateral term with a tensor-based fourth-order PDE for image super-resolution*, Adv. Comput. Math. **48** (2022) 83.

[9] I. El Mourabit, M. El Rhabi, A. Hakim, A. Laghrib, E. Moreau, *A new denoising model for multi-frame super-resolution image reconstruction*, Signal Process. **132** (2017) 51–65.

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

[11] R. Glowinski, Y. Song, X. Yuan, *An ADMM numerical approach to linear parabolic state constrained optimal control problems*, Numer. Math. **144** (2020) 931–966.

[12] A. Hadri, A. Laghrib, and H. Oummi, *An optimal variable exponent model for magnetic resonance images denoising*, Pattern Recognit. Lett. **151** (2021) 302–309.

[13] A. Hadri, A. Laghrib, I. El Mourabit, *A new learning space-variant anisotropic constrained-PDE for image denoising*, Appl. Math. Model. **125** (2024) 139–163.

[14] N. Komodakis, J.-C. Pesquet, *Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems*, IEEE Signal Process. Mag. **32** (2015) 31–54.

[15] A. Oulmelk, M. Srati, L. Afraites, *Implementation of the ADMM approach to constrained optimal control problem with a nonlinear time-fractional diffusion equation*, Discrete Contin. Dyn. Syst.-S **18** (2022) 2238–2253.

[16] P. Perona, J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell. **12** (1990) 629–639.

[17] B.K. Shreyamsha Kumar, *Image denoising based on Gaussian/bilateral filter and its method noise thresholding*, Signal Image Video Process. **7** (2013) 1159–1172.

[18] M. Wang, S. Zheng, X. Li, X. Qin, *A new image denoising method based on Gaussian filter*, 2014 International Conference on Information Science, Electronics and Electrical Engineering, Sapporo, Japan, (2014)163-167.

[19] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teuber Verlag, Stuttgart, 1998.

[20] J. Yang, X. Yuan, *Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization*, Math. Comput. **82** (2013) 301–329.

[21] J. Zhang, R. Lai, C.C. J. Kuo, *Adaptive directional total-variation model for latent fingerprint segmentation*, IEEE Trans. Inf. Forensics Security **8** (2013) 1261–1273.