# On a Variational Problem with a Nonstandard Growth Functional and Its Applications to Image Processing

Ciro D'Apice[1] · Peter I. Kogut[2,3] · Olha P. Kupenko[4,5] · Rosanna Manzo[6]

## Abstract

We propose a new variational model in Sobolev–Orlicz spaces with non-standard growth conditions of the objective functional and discuss its applications to image processing. The characteristic feature of the proposed model is that the variable exponent, which is associated with non-standard growth, is unknown a priori and it depends on a particular function that belongs to the domain of objective functional. So, we deal with a constrained minimization problem that lives in variable Sobolev–Orlicz spaces. In view of this, we discuss the consistency of the proposed model, give the scheme for its regularization, derive the corresponding optimality system, and propose an iterative algorithm for practical implementations.

**Keywords** Inverse problem · Nonconvex programming · Image reconstruction · Constrained minimization problems · Approximation methods · Sobolev–Orlicz space

**Mathematics Subject Classification** 90C90 · 94A08

## 1 Introduction

Over the past few decades, remote sensing has been increasingly contributing to many agricultural monitoring services by providing worldwide systematic observations of the so-called optical vegetation indexes. These indexes (like Normalized Difference Vegetation Index (NDVI), Infrared Percentage Vegetation Index (NPVI), and others) are well-established proxies for crop conditions and give early insights into how well the crops are doing and if they are in need of water or nutrients?

From physical point of view, the remote sensing sensors record earth's reflectance in different wavelength and these received reflectance values are processed to create separate image for each wavelength. The reflectance values stored for different wavelength in different layers, which are also called spectral channels, are present in such satellite images. In particular, the satellites Sentinel-2 delivers 13 spectral channels ranging from 10 to 60-m in pixel size. So, a hyperspectral image simply define when more than one spectrum and at least three wavelength data like RGB are recorded by sensor to create composite of satellite images.

Peter I. Kogut and Olha P. Kupenko have contributed equally to this work.

✉ Rosanna Manzo
rmanzo@unisa.it

Ciro D'Apice
cadapice@unisa.it

Peter I. Kogut
p.kogut@i.ua

Olha P. Kupenko
kupenko.olga@gmail.com

[1] Dipartimento di Scienze Aziendali - Management and Innovation Systems, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, Salerno, Italy

[2] Department of Differential Equations, Oles Honchar Dnipro National University, Gagarin av. 72, Dnipro 49005, Ukraine

[3] EOS Data Analytics Ukraine, Gagarin av. 103a, Dnipro 49005, Ukraine

[4] Department of System Analysis and Control, Dnipro University of Technology, Yavornitskii av. 19, Dnipro 49005, Ukraine

[5] Institute for Applied System Analysis of National Academy of Sciences and Ministry of Education and Science of Ukraine, Peremogy av. 37/35, IASA, Kyiv 03056, Ukraine

[6] Department of Information Engineering, Electrical Engineering and Applied Mathematics, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, Salerno, Italy

Typically, to provide useful information of crop status, it is important to monitor the selected agricultural fields in different spectrum on a regular basis in order to retrieve the complete time series of a vegetation index especially during those periods when conditions on the fields change drastically (It can be essential growth stages such as plant emergence, maturation period, and others.). However, in spite of the fact that optical images have a high resolution and are easily captured by low-cost cameras, the real-life satellite images frequently suffer from different types of noise, blur, and other atmosphere artifacts that can affect the radiation recovered by the sensors in different spectral channels and in a different manner. As a result, such satellite images lose their efficiency for the crop field monitoring problems and their utilization can lead to erroneous results and inferences.

In such situation, the problem we are going to deal with, can be shortly described as follows. We have a hyperspectral image $\mathbf{u}_0 = [u_{1,0}, u_{2,0}, \ldots, u_{M,0}]^T \in L^1(\Omega; \mathbb{R}^M)$ which is corrupted by noise or blur, or such that the features of interest in this image are hidden. Here, $\Omega \subset \mathbb{R}^2$ is a bounded open domain with sufficiently smooth boundary $\partial\Omega$. So, the challenge is to recover the original image $\mathbf{u} = [u_1, u_2, \ldots, u_M]^T$, in fact, the hidden image features, from the observed datum $\mathbf{u}_0$. In mathematical terms, this means that for each spectral channel $i = 1, \ldots, M$, one has to solve an inverse problem

$$T_i u_i + w = u_{i,0},$$

where the linear blur operator $T_i$ models the process through which the $i$-th spectral channel of $\mathbf{u}$ went before observation, and $w$ is the unknown noise affecting the measurements. Due to the presence of the noise and the fact that the blurring operator $T_i$ is usually ill-conditioned or even non-invertible, and $w$ is the unknown noise affecting the measurements, the recovery of $\mathbf{u}$ from the measurements $\mathbf{u}_0$ is an ill-posed problem [4,28]. In general, the ill-posed nature of this problem implies that there are too many ways one can obtain an approximate solution. A reasonable solution is to reformulate the image-deblurring problem by taking into account the image formation and acquisition process as well as any available prior information about the properties of the image to be restored.

The most popular way for that is to represent the denoising problem (in each separate spectral channel) in the form of an appropriate variational problem

$$u \in \underset{v \in B_2}{\text{Argmin}} \left\{ \mathcal{J}(v) = R(v) + \frac{1}{2\lambda} \|Tv - u_0\|_{B_1}^2 \right\}, \quad (1)$$

where $B_1$, $B_2$ are two Banach spaces over $\Omega$, $u_0 \in B_1$ is a given image, $\lambda > 0$ is the tuning parameter, $T \in \mathcal{L}(B_1, B_1)$ is a bounded linear operator, and $R : B_2 \to \mathbb{R}$ is the regularizing term which smooths the image $u$ and represents some

kind of a priori information about the minimizer $u$. Here, the term $\|Tu - u_0\|_{B_1}^2$ is the so-called fidelity term of the approach which forces the minimizer $u$ to stay close to the given image $u_0$ (how close depends on the size of $\lambda$). As for the choice of operator $T$, they usually set $T = Id$ (the identity in $B_1$) for the image denoising problems, and $T$ is a symmetric kernel with a smaller support than the image $u$ for the deblurring problems. The choice of a proper regularization $R(v)$, Banach spaces $B_1$, $B_2$, and a convenient fidelity term $\|Tv - u_0\|_{B_1}^2$ are always challenging problems since they affect the quality of the desired image and also the consistency with the given data.

There is a wide choice of the regularization term proposed in the literature to ensure the well-posedness of the denoising problem (1). In the most cases, this term can be represented as follows:

$$R(v) = \int_\Omega |Dv|^p \, dx \quad \text{with} \quad 1 \le p \le 2, \quad (2)$$

where $Dv$ stands for the generalized gradient.

In fact, it means that we can exploit the benefits of isotropic diffusion (when $p = 2$) arising from the minimization problem (1), total variation-based diffusion ($p = 1$), and more general anisotropic diffusion ($1 < p < 2$).

### 1.1 The Case $p = 1$ (Total Variation Minimization)

One of the classical regularization terms was the total variation (TV), introduced by Rudin et al. [48]. Basically, they introduced in [48] the following variational model (shortly, the ROF model)

$$\min_{u \in BV(\Omega)} \left[ \int_\Omega |Du| + \frac{\zeta}{2} \|u - f\|_{L^2(\Omega)}^2 \right], \quad (3)$$

where $BV(\Omega)$ stands for the space of functions with bounded variation [2,3], $\Omega \subset \mathbb{R}^2$ is an open bounded Lipschitz domain, $f \in L^2(\Omega)$ is a given image, $f = u + \xi$, $\xi$ is the 'noise', and $\int_\Omega |Du|$ denotes the total variation seminorm of $u$ on $\Omega$.

This model produces rather good results for removing noise and preserving edges in structured images, meaning images without texture-like components, i.e., without edges. Moreover, the choice $B_2 = BV(\Omega)$ is reasonable from mathematical point of view, since the space of functions of bounded variation is allowed for discontinuities which are necessary for edge reconstruction. At the same time, it fails in the presence of the latter. Namely, it cannot separate pure noise, i.e., well oscillatory components, from texture, but removes both equally. Also the ROF model suffers from the so-called blocky effects and it can also develop 'false edges' which can mislead a human or computer into identifying

erroneous features not present in the true image. These phenomena were a subject of long intensive discussions in the literature (see, for instance, [4,5,14,15,27,28,44,47]).

## 1.2 The Case $1 < p \leq 2$

As follows from (1)–(2), different types of diffusion can be explored. In particular, setting $p = 2$ in (1), we arrive at the isotropic diffusion which, in principle, solves the staircasing problem. Here, the "staircase effect" means the creation in the image of flat regions separated by artifact boundaries. However, this model is not good for image reconstruction since it has no mechanism for preserving edges. If we fix some value $p \in (1, 2)$, then we deal with an anisotropic diffusion which is somewhere between TV-based and isotropic smoothing. In spite of the fact that this type of diffusion can be effective in reconstructing piecewise smooth regions, a fixed value of $1 < p < 2$ may not allow for discontinuities, and thus obliterating edges.

## 1.3 High-Order Total Variation and Combination of TV-Based with Isotropic Diffusion

To overcome the drawbacks mentioned above, various modifications of the model (1) have been introduced. We can briefly indicate here the Adaptive Total Variation model [49] and a variety of high-order total variation regularizations (see [17,50] and the references therein). Recently, in order to reduce the staircasing effect, in a series of papers [6,11,12,50,53], authors proposed the models based on the so-called total generalized variation (TGV) involving a linear combination of higher-order derivatives and the total variation, on the total fractional-order derivatives, and on a combination of the total variation seminorm and $L^p$ norms.

One of the successful high-order regularization was the combination of TV and $TV^2$ terms proposed by Chambolle and Lions [15,16] (see also [46]). Namely, they proposed the following specification of (1):

$$
\begin{aligned}
&B_2 = BV(\Omega), \\
&R(v) = \frac{1}{2\beta} \int_{|Dv| \leq \beta} |Dv|^2 \, dx + \int_{|Dv| > \beta} |Dv| \, dx - \frac{\beta}{2}.
\end{aligned} \quad (4)
$$

As follows from (4), in this model, the diffusion is strictly perpendicular to the gradient, where $|Dv| > \beta$; that is, where edges are most likely present, and it is isotropic where $|Dv| \leq \beta$. So, this model is successful in restoring images where homogeneous regions are separated by distinct edges. However, if the image intensities representing objects are non-uniform or if an image is highly degraded, this model may become sensitive to the threshold $\beta$.

Not so long ago, Brinkmann et al. [8] proposed a new model that generalize most of the known regularizations,

including TV and TGV, and also the combination between them. This model is based on the introduction of a vector field whose nonzero entries are concentrated at the edges of the image.

## 1.4 Regularization Terms with Variable Exponent $(1 \leq p(x) \leq 2)$

This case is rather new in statement of the image denoising problem, albeit the conjecture to utilize this idea was firstly pushed forward by Blomgren et al. [5] at the end of 90th. In particular, they proposed the following variant of regularization term

$$
B_2 = BV(\Omega), \quad R(v) = \int_\Omega |Dv|^{p(|Dv|)} \, dx, \quad (5)
$$

where $\lim_{s \to 0} p(s) = 2$, $\lim_{s \to \infty} p(s) = 1$, and $p$ is a monotonically decreasing function. In spite of the natural expectations that this model will reap the benefits of both isotropic and TV-based diffusion, as well as a combination of the two, its real exploitation was rather questionable because of some principle difficulties coming from the absence of its rigorous mathematically substantiation.

Further specifications of this approach can be found in [18,19], where the authors discuss the following model:

$$
\min_{u \in BV(\Omega) \cap L^2(\Omega)} \left[ R(u) + \frac{\zeta}{2} \|u - f\|^2_{L^2(\Omega)} \right], \quad (6)
$$

with

$$
R(u) = \int_\Omega \left\{ \begin{array}{ll} \frac{1}{p(x)} |Du|^{p(x)}, & |Du| \leq \beta, \\ |Du| - \frac{\beta p(x) - \beta^{p(x)}}{p(x)}, & |Du| > \beta. \end{array} \right\} dx. \quad (7)
$$

Here, $\beta > 0$ is a given threshold,

$$
p(x) = 1 + \frac{1}{1 + k|\nabla G_\sigma * f(x)|^2}, \quad (8)
$$

$(G_\sigma * v)(x)$ determines the convolution of function $v$ with the $N$-dimensional Gaussian kernel $G_\sigma$,

$$
G_\sigma(x) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^N} \exp\left(-\frac{|x|^2}{2\sigma^2}\right), \quad x \in \mathbb{R}^N, \quad (9)
$$

and $k > 0$ and $\sigma > 0$ are fixed parameters.

A similar model to (6)–(9) has been later proposed in [42], where, instead of (7), the authors considered the variant $R(u) = \frac{1}{p(x)} |Du|^{p(x)}$ with the same choice of variable exponent $p(x)$.

The main benefit of the model (6)–(9) is the manner in which it accommodates the local image information. Where the gradient of the noisy image $f$ is sufficiently large (i.e.,

likely edges), only TV-based diffusion will be used. Where the gradient is close to zero (i.e., homogeneous regions), the model is isotropic. At all other locations, the filtering is somewhere between Gaussian and TV-based. However, as follows from (8), the type of anisotropy in (6) is completely predefined by the structure of the observed, noisy image $f$, whereas denoised image $u$ may have other structure with other location of edges and other shape of homogeneous regions.

In spite of the fact that there are many other different variants for the choice of regularization term using the so-called directional total variation [9,10] and flexible space-variants anisotropic regularization [12,39], to the best of our knowledge, the development of effective choice of Banach spaces $B_1$, $B_2$ and a proper regularization term $R(v)$ for general image denoising problem with different noise distributions remains an open problem for nowadays.

## 1.5 The Choice of Fidelity Term

As for the choice of the fidelity term $\frac{1}{2\lambda}\|Tv - u_0\|_{B_1}^2$ in (1), it mainly depends on the noise distribution. In most cases, when the noise is assumed to be additive with Gaussian distribution, the $L^2$-norm is considered (see, for example, [14]). At the same time, if the image is corrupted by noise with different statistical properties, other fidelity terms are investigated [7, 45,51]. For instance, when the image is corrupted by impulse noise, a typical choice for the space $B_1$ is $L^1(\Omega)$. In fact, the use of $L^1$-norm is more robust to recover the exact solution compared to the $L^2$-norm, which suffers from the contrast loss and other artifacts.

There exist many works in the literature regarding the choice of the fidelity term, where the authors propose different variants including $L^1 + L^2$ data fidelity [29] and the so-called infimal convolution combination of data fidelities [13,26]. Recently, the non-convex functions have been performed as a fidelity term and have achieved a considerable success in removing complex non-smooth noise [25,53]. In particular, the following concave variant of fidelity term has been proposed in [1]:

$$P\left(\|v - u_0\|_{L^1(\Omega)}\right) \quad \text{with} \quad P(t) = \frac{t}{t + \gamma},$$

where $\gamma$ is a strictly positive constant used to control the concavity rate. However, when the Gaussian noise level is very high compared to the impulse one, the exploitation of the concave variant of fidelity term can lead to the appearance of some artifacts in denoised images. So, the correct choice of the $\gamma$ parameter, which controls the concavity rate, is a nontrivial and challenging problem.

## 1.6 Motivations

A common feature of the models used in denoising and deblurring problems is the fact that a true image $u$ together with the rest feasible solutions belongs to the same functional space $B_2$ (with $B_2 = BV(\Omega)$ or $B_2 = W^{1,s}(\Omega)$ in a particular cases), whereas it is natural to suppose that the regularity of the restored image $u$ should be low in places in $\Omega$ where edges or discontinuities are present in $u$, and that it should be high in places where $u$ is smooth or contains homogeneous features.

Because of the specifics of the satellite images with crop fields, such images typically contain a significant portion of edges (boundaries of the crop locations). That is why the important question is to separate pure noise from edges. Moreover, since a typical satellite image $\mathbf{u}_0$ can be viewed as a number of separate scalar images (spectral channels) $\left[u_{1,0}, u_{2,0}, \ldots, u_{M,0}\right]^T$ which have been captured by different sensors in different zones of electromagnetic spectrum with different space resolution, the level of noise and its distribution can essentially variate from channel to channel. Figure 1 provides a visual comparison of red and near-infrared channels for a satellite image from Sentinel-2. So, it is plausible to suppose that different channels $u_{i,0}$, $u_{j,0}$, $i \neq j$, can 'live' in different functional spaces, for instance $u_{i,0} \in W^{1,s_i}(\Omega)$ and $u_{j,0} \in W^{1,s_j}(\Omega)$ with $s_i \neq s_j$. However, we cannot achieve such regularity for the recovered images in the framework of the standard setting of denoising problem. Even if we make use of the model (6)–(9), we can easily deduce that any its solution $u \in BV(\Omega) \cap L^2(\Omega)$ has the following Sobolev regularity: $u \in W^{1,p(x)}(\Omega)$, where the exponent $p(x)$ is completely determined by the original noisy image $f$, whereas regularity of images $f$ and $u$ and their texture can be drastically different. It says that the flexibility of the model (6)–(9) is not enough, especially for the hyperspectral images. In view of this the idea to involve into consideration the variational model (1) with an appropriate choice of variable character of exponent $p(s)$ in the regularization term (5), sounds more attractive.

## 1.7 Contribution

The main purpose of this paper is to present a robust approach for the denoising and deblurring of non-smooth hyperspectral satellite images $\mathbf{u}_0 \in L^1(\Omega; \mathbb{R}^M)$ using the energy functional with nonstandard growth. In fact, we use the $L^1$-norm of the noise in the minimization function and a special form of anisotropic diffusion tensor for the regularization term. Namely, we deal with the following optimization problem:

**Fig. 1** Example. Left panel: the original noisy satellite image. Middle panel: its red channel. Right panel: near-infrared channel



$$J(v) = \int_\Omega \frac{1}{\mathfrak{F}(v(x))} |\nabla v(x)|^{\mathfrak{F}(v(x))} \, dx$$
$$+ \mu \int_\Omega |Tv(x) - u_0(x)| \, dx \longrightarrow \inf \tag{10}$$

subject to the constraints

$$v \in W^{1,\mathfrak{F}(v(\cdot))}(\Omega), \quad 1 \le \gamma_0 \le v(x) \le \gamma_1 \quad \text{a.e. in } \Omega, \tag{11}$$

where $u_0 \in L^1(\Omega)$ is a particular spectral channel of the original noisy image $\mathbf{u}_0 = \left[u_{1,0}, u_{2,0}, \ldots, u_{M,0}\right]^T \in L^1(\Omega; \mathbb{R}^M)$, $\mu > 0$ is the tuning parameter of the problem, $W^{1,\mathfrak{F}(v(\cdot))}(\Omega)$ stands for the so-called Sobolev–Orlicz space associated with a feasible solution $v$, $T \in \mathcal{L}(L^1(\Omega))$ is a bounded linear operator with unbounded inverse,

$$\mathfrak{F}(v(x)) = 1 + g\left(|(\nabla G_\sigma * v)(x)|\right) \quad \text{in } \Omega, \tag{12}$$

$(G_\sigma * v)(x)$ determines the convolution of function $v$ with the $N$-dimensional Gaussian filter kernel $G_\sigma$ (see (9)),

$$(\nabla G_\sigma * v)(x) = \int_{\mathbb{R}^N} \nabla G_\sigma(x - y)\widetilde{v}(y) \, dy, \quad \forall x \in \Omega, \tag{13}$$

$\widetilde{v}$ is zero-extension of $v$ outside $\Omega$, $|\xi|$ stands for the Euclidean norm of $\xi \in \mathbb{R}^N$ given by the rule $|\xi| = \sqrt{(\xi,\xi)_{\mathbb{R}^N}}$, and $g : [0,\infty) \to (0,\infty)$ is a continuous monotone decreasing function such that $g(0) = 1$ and $g(t) > 0$ for all $t > 0$ with $\lim_{t \to +\infty} g(t) = 0$.

In particular, the definition of variable exponent $p(x) := \mathfrak{F}(u(x))$ in the form $1 + g\left(|(\nabla G_\sigma * u)(x)|\right)$, where the edge-stopping function $g(s)$ is taken in the form of the Cauchy law

$$g(t) = \frac{1}{1 + (t/a)^2} \quad \text{with an appropriate } a > 0, \tag{14}$$

implies that $p(x) \approx 1$ in places in $\Omega$ where edges or discontinuities are present in the spectral channels of an image $u(x)$, and $p(x) \approx 2$ in places where $u(x)$ is smooth or contains homogeneous features (see Lemma 1).

As for the data fidelity term, in contrast to the quadratic data-fitting term in the ROF model (3) (see also (1)), where it is specialized for zero-mean Gaussian noise, we take this term in $L^1$-norm just in order to apply the model to other types of noise, especially when the noise is impulsive or contains gross outliers. Following this way, we are going to increase the noise robustness of the model (10) albeit it makes such variational problem completely non-smooth and, hence, significantly more difficult from a minimization point of view.

So, the main characteristic feature of this model is that we involve into consideration the energy functional with the nonstandard growth where the edge information for restoration of $\mathbf{u}_{rec} = \left[u_1^{rec}, \ldots, u_M^{rec}\right]$ is accumulated in the variable exponents $\left[\mathfrak{F}(u_1^{rec}(x)), \ldots, \mathfrak{F}(u_M^{rec}(x))\right]$. Basically, by analogy with the model (6)–(9), here we utilize a similar manner of accommodating the local information. However, there is a principle difference between the models (6)–(9) and (10). In contrast to [18,19,42], we do not predefine the variable exponent $p(x)$ a priori using for that the original noisy image, but instead we associate this characteristic with each feasible solution. As a result, we admit that each feasible solution to this problem lives in the corresponding 'personal' functional space $W^{1,\mathfrak{F}(u_i(\cdot))}(\Omega)$. Formally it means that we look for the true image $\mathbf{u}_{rec}$ such that

$$\mathbf{u}_{rec} \in W^{1,\mathfrak{F}(u_1^{rec}(\cdot))}(\Omega) \times W^{1,\mathfrak{F}(u_2^{rec}(\cdot))}(\Omega) \times \ldots$$
$$\times W^{1,\mathfrak{F}(u_M^{rec}(\cdot))}(\Omega).$$

As follows from the definition of Sobolev–Orlicz space $W^{1,\mathfrak{F}(u_i^{rec}(\cdot))}(\Omega)$ (see Appendix B), its regularity is completely determined by the exponent $\mathfrak{F}(u_i^{rec}(\cdot))$ which depends on $i$-th spectral channel of the true image $\mathbf{u}_{rec}$ and, hence, is unknown a priori. Moreover, as it was shown in Fig. 1, the exponents $\left\{\mathfrak{F}(u_1^{rec}(\cdot)), \mathfrak{F}(u_2^{rec}(\cdot)), \ldots, \mathfrak{F}(u_M^{rec}(\cdot))\right\}$ may significantly differ from channel to channel. In particular, some pixels, which are the local minimum points in red channel, become local maximum points in near-infrared channel and vice versa. Moreover, the different feasible solutions $v_1 \ne v_2$ to the above problem live in different functional spaces: We have $v_1 \in W^{1,\mathfrak{F}(v_1(\cdot))}(\Omega)$, whereas $v_2 \in W^{1,\mathfrak{F}(v_2(\cdot))}(\Omega)$. As

a consequence, any minimizing sequence to this problem is, in fact, a sequence living in the scale of variable spaces. As a result, such notions as a convergence concept, compactness, density and others should be specified for the case of variable Sobolev–Orlicz spaces.

Thus, in spite of the fact that in the literature there are many approaches to the study of variational problems in abstract functional spaces, the above-mentioned circumstances make the problem (10) challenging (see [9,18,19,21,22,30] for the recent studies in this field).

In summary, the main contributions of our paper can be enumerated as follows:

- The variational statement for the hyperspectral image denoising and deblurring in the form of minimization problem in Sobolev–Orlicz spaces with non-standard growth conditions of the objective functional;
- Rigorous substantiation of the well posedness of the variational problem with non-standard growth functional;
- The proof of existence result to the proposed variational problem;
- The iterative algorithm for the numerical implementations;
- Rigorous derivation of the first-order necessary conditions for the approximating problem;
- The approximation scheme with fictitious control and rigorous substantiation of the attainability properties of its solutions.

The remainder of the paper is organized as follows: Sect. 2 focuses on the well-posedness of the proposed model for the image denoising and deblurring problem, and on the solvability issues. In Sect. 3, we propose an iterative algorithm for the approximate solution of the denoising problem and discuss its convergence properties. The deriving of optimality conditions for approximating problem and their rigorous substantiation, we provide in Sect. 4. Since the proposed iterative algorithm leads to the so-called weak solutions, the possible ways for the relaxation of the constrained minimization problem (10) are discussed in Sect. 5. With that in mind we introduce a family of minimization problems with fictitious control and show that each of these problem is solvable and their solutions converge to the solution of the original problem in an appropriate topology. For illustration of the proposed approach, we give in Sect. 6 some results of numerical experiments with real satellite images. Finally, we give in "Appendix" section the main auxiliary results concerning the Orlicz spaces and Sobolev spaces with variable exponent.

## 2 Existence Result

Our main intention in this section is to show that constrained minimization problem (10) is consistent and admits at least

one solution. Because of the specific form of the energy functional $J(v)$, we cannot assert that minimization problem (10) is well defined on the entire set $W^{1,\alpha}(\Omega)$ with some appropriate exponent $\alpha \geq 1$. Moreover, the structure and main topological properties of the set of feasible solution to minimization problem (10) are challenging issues. So, the study of these issues for minimization problem (10) is the main subject of this section. (We can refer to [5,9,18,23] for some specific details that can appear in this case.)

We begin with the following key assumption: The true intensities $u_i$ of all spectral channels for the retrieved image $\mathbf{u} = [u_1, \ldots, u_M]$ are subjected to the constraints $\gamma_{0,i} \leq u_i(x) \leq \gamma_{1,i}$ a.e. in $\Omega$, where

$$\gamma_{0,i} = \inf_{x \in \Omega} u_{i,0}, \qquad \gamma_{1,i} = \sup_{x \in \Omega} u_{i,0}. \tag{15}$$

We say that a function $\mathbf{u}^{rec} = \left[ u_1^{rec}, u_2^{rec}, \ldots, u_M^{rec} \right]^T : \Omega \to \mathbb{R}^M$ is the result of restoration of a noise contaminated image $\mathbf{u}_0 : \Omega \to \mathbb{R}^M$ if for a given regularization parameters $\mu > 0$ and given linear blur operators $T_i \in \mathcal{L}(L^1(\Omega))$ with $i = 1, \ldots, M$, each spectral component $u_i^{rec}$ is the solution of the following constrained minimization problem with the nonstandard growth energy functional

$$J_i(v) := \int_\Omega \frac{1}{\mathfrak{F}(v(x))} |\nabla v(x)|^{\mathfrak{F}(v(x))} \, dx$$
$$+ \mu \int_\Omega |T_i(v(x)) - u_{0,i}(x)| \, dx \longrightarrow \inf_{v \in \Xi_i} \tag{16}$$

i.e., for each $i = 1, \ldots, M$,

$$u_i^{rec} \in \Xi_i \text{ and } J_i \left( u_i^{rec} \right) = \inf_{v \in \Xi_i} J_i \left( v \right).$$

Here, $\Xi_i$ is the set of feasible solutions to the problem (16) which we define as follows:

$$\Xi_i = \Big\{ v \, | v \in W^{1,\mathfrak{F}(v(\cdot))}(\Omega),$$
$$0 \leq \gamma_{0,i} \leq v(x) \leq \gamma_{1,i} < +\infty \text{ a.e. in } \Omega \Big\},$$

where $\mathfrak{F}(v)$ is given by (2), and $W^{1,p(\cdot)}(\Omega)$ is the Sobolev–Orlicz space (for the details we refer to Appendix A).

Hereinafter, we associate with each spectral channel $u_i^{rec}$ of the restored optical image $\mathbf{u}^{rec} = \left[ u_1^{rec}, u_2^{rec}, \ldots, u_M^{rec} \right]^T : \Omega \to \mathbb{R}^M$ the so-called texture index $p_i : \Omega \to \mathbb{R}$ following the rule:

$$p_i(x) := \mathfrak{F}(u_i(x)) = 1 + g \left( |(\nabla G_\sigma * u_i)(x)| \right),$$
$$\forall x \in \Omega, \; \forall i = 1, \ldots, M, \tag{17}$$

where $g : [0, \infty) \to (0, \infty)$ is the edge-stopping function which we take in the form of the Cauchy law $g(t) = \frac{1}{1+(t/a)^2}$.

As follows from representation (12) and smoothness of the Gaussian filter kernel $G_\sigma$, we have the following estimates

$$|(\nabla G_\sigma * v)(x)| \leq \int_\Omega |\nabla G_\sigma(x - y)||v(y)\,\mathrm{d}y$$
$$\leq \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\|v\|_{L^1(\Omega)}$$
$$\leq \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}|\Omega|\gamma_{1,i},$$
$$\mathfrak{F}(v(x)) = 1 + \frac{a^2}{a^2 + (|(\nabla G_\sigma * v)(x)|)^2}$$
$$\geq 1 + \frac{a^2}{a^2 + \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}^2 \|v\|_{L^1(\Omega)}^2}$$
$$\geq 1 + \delta, \forall x \in \Omega,$$
$$\mathfrak{F}(v(x)) \leq 2 \quad \text{in } \Omega,$$

where

$$\delta = \frac{a^2}{a^2 + \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}^2 |\Omega|^2 \gamma_{1,i}^2} \ll 1,$$
$$\|G_\sigma\|_{C^1(\overline{\Omega - \Omega})} = \max_{\substack{z = x - y \\ x \in \overline{\Omega}, y \in \overline{\Omega}}} [|G_\sigma(z)| + |\nabla G_\sigma(z)|]$$
$$= \frac{1}{2\pi\sigma^2}e^{-1}\left[1 + \frac{1}{\sigma^2}\mathrm{diam}\,\Omega\right].$$

The following result plays a crucial role in the sequel.

**Lemma 1** *Let* $\{v_k\}_{k\in\mathbb{N}} \subset L^\infty(\Omega)$ *be a sequence of measurable non-negative functions such that* $v_k(x) \to v(x)$ *weakly-* $*$ *in* $L^\infty(\Omega)$ *for some* $v \in L^\infty(\Omega)$, *and each element of this sequence is extended by zero outside of* $\Omega$. *Let* $\{p_k = 1 + g(|(\nabla G_\sigma * v_k)|)\}_{k\in\mathbb{N}}$ *be the corresponding sequence of texture indices. Then,*

$$p_k(\cdot) \to p(\cdot) = 1 + g(|(\nabla G_\sigma * v)(\cdot)|)$$
*uniformly in* $\overline{\Omega}$ *as* $k \to \infty$,
$$\alpha := 1 + \delta \leq p_k(x) \leq \beta := 2,$$
$$\forall x \in \Omega, \ \forall k \in \mathbb{N}, \tag{18}$$
*with* $\delta = a^2\left[a^2 + \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}^2 \sup_{k\in\mathbb{N}}\|v_k\|_{L^1(\Omega)}^2\right]^{-1}$. $\tag{19}$

**Proof** In view of the initial assumptions, the sequence $\{v_k\}_{k\in\mathbb{N}}$ is uniformly bounded in $L^1(\Omega)$. Hence, by smoothness of the Gaussian filter kernel $G_\sigma$, it follows that

$$|(\nabla G_\sigma * v_k)(x)| \leq \int_\Omega |\nabla G_\sigma(x - y)|v_k(y)\,\mathrm{d}y$$
$$\leq \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\|v_k\|_{L^1(\Omega)},$$
$$p_k(x) = 1 + \frac{a^2}{a^2 + (|(\nabla G_\sigma * v_k)(x)|)^2}$$

$$\geq 1 + \frac{a^2}{a^2 + \|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}^2 \|v_k\|_{L^1(\Omega)}^2},$$
$$\forall x \in \Omega.$$

Then, $L^1$-boundedness of $\{v_k\}_{k\in\mathbb{N}}$ guarantees the existence of a positive value $\delta \in (0, 1)$ (see (18)) such that estimate (18) holds true for all $k \in \mathbb{N}$. Moreover, as follows from the relations

$$|p_k(x) - p_k(y)|$$
$$\leq a^2\left|\frac{|(\nabla G_\sigma * v_k)(x)|^2 - |(\nabla G_\sigma * v_k)(y)|^2}{(a^2 + |(\nabla G_\sigma * v_k)(x)|^2)(a^2 + |(\nabla G_\sigma * v_k)(y)|^2)}\right|$$
$$\leq \frac{2\|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\|v_k\|_{L^1(\Omega)}}{a^2}$$
$$\left||(\nabla G_\sigma * v_k)(x)| - |(\nabla G_\sigma * v_k)(y)|\right|$$
$$\leq \frac{2\|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\gamma_1^2|\Omega|}{a^2}$$
$$\int_\Omega |\nabla G_\sigma(x - z) - \nabla G_\sigma(y - z)|\,\mathrm{d}z,$$
$$\forall x, y \in \Omega \text{ with } \gamma_1 = \sup_{k\in\mathbb{N}}\|v_k\|_{L^\infty(\Omega)},$$

and smoothness of the function $\nabla G_\sigma(\cdot)$, there exists a positive constant $C_G > 0$ independent of $k$ such that

$$|p_k(x) - p_k(y)| \leq \frac{2\|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\gamma_{1,i}^2|\Omega|C_G}{a^2}$$
$$|x - y|, \ \forall x, y \in \Omega.$$

Setting

$$C := \frac{2\|G_\sigma\|_{C^1(\overline{\Omega - \Omega})}\gamma_{1,i}^2|\Omega|C_G}{a^2}, \tag{20}$$

we see that

$$\{p_k(\cdot)\} \subset \mathfrak{S}$$
$$= \left\{h \in C^{0,1}(\Omega) \left| \begin{array}{l} |h(x) - h(y)| \leq C|x - y|, \quad \forall x, y, \in \Omega, \\ 1 < \alpha \leq h(\cdot) \leq \beta \text{ in } \overline{\Omega}. \end{array}\right.\right\}$$

Since $\max_{x\in\overline{\Omega}}|p_k(x)| \leq \beta$ and each element of the sequence $\{p_k\}_{k\in\mathbb{N}}$ has the same modulus of continuity, it follows that this sequence is uniformly bounded and equi-continuous. Hence, by Arzelà–Ascoli theorem, the sequence $\{p_k\}_{k\in\mathbb{N}}$ is relatively compact with respect to the strong topology of $C(\overline{\Omega})$. Taking into account the estimate (20) and the fact that the set $\mathfrak{S}$ is closed with respect to the uniform convergence and

$$(\nabla G_\sigma * v_k)(x) \to (\nabla G_\sigma * v)(x) \text{ as } k \to \infty, \ \forall x \in \Omega$$

by definition of the weak-$*$ convergence in $L^\infty(\Omega)$, we deduce: $p_k(\cdot) \to p(\cdot)$ uniformly in $\overline{\Omega}$ as $k \to \infty$, where $p(x) = 1 + g\left(|(\nabla G_\sigma * v)(x)|\right)$ in $\Omega$. The proof is complete. $\qquad\square$

For our further analysis, we make use of the following result (we refer to [55, Lemma 13.3] for comparison) concerning the lower semicontinuity property of the variable $L^{p_k(\cdot)}$-norm with respect to the weak convergence in $L^{p_k(\cdot)}(\Omega)$.

**Proposition 1** *If a bounded sequence $\left\{f_k \in L^{p_k(\cdot)}(\Omega)\right\}_{k \in \mathbb{N}}$ converges weakly in $L^\alpha(\Omega)$ to $f$ for some $\alpha > 1$, then $f \in L^{p(\cdot)}(\Omega)$, $f_k \rightharpoonup f$ in variable $L^{p_k(\cdot)}(\Omega)$, and*

$$\liminf_{k \to \infty} \int_\Omega |f_k(x)|^{p_k(x)}\, dx \geq \int_\Omega |f(x)|^{p(x)}\, dx. \tag{21}$$

**Proof** In view of the fact that

$$\left| \int_\Omega |f_k(x)|^{p_k(x)}\, dx - \int_\Omega \frac{p(x)}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx \right|$$
$$\leq \|p_k - p\|_{C(\overline{\Omega})} \int_\Omega \frac{1}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx$$
$$\leq \frac{\|p_k - p\|_{C(\overline{\Omega})}}{\alpha} \int_\Omega |f_k(x)|^{p_k(x)}\, dx \overset{by\ (A9)}{\to} 0 \ \text{as}\ k \to \infty,$$

we see that

$$\liminf_{k \to \infty} \int_\Omega |\ f_k(x)|^{p_k(x)}\, dx$$
$$= \liminf_{k \to \infty} \int_\Omega \frac{p(x)}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx. \tag{22}$$

Using the Young's inequality $ab \leq |a|^p/p + |b|^{p'}/p'$, we have

$$\int_\Omega \frac{p(x)}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx \geq \int_\Omega p(x) f_k(x) \varphi(x)\, dx$$
$$- \int_\Omega \frac{p(x)}{p_k'(x)} |\varphi(x)|^{p_k'(x)}\, dx, \tag{23}$$

for $p_k'(x) = p_k(x)/(p_k(x) - 1)$ and any $\varphi \in C_0^\infty(\mathbb{R}^N)$.

Then, passing to the limit in (23) as $k \to \infty$ and utilizing property (A8) and the fact that

$$\lim_{k \to \infty} \int_\Omega f_k(x) \varphi(x)\, dx = \int_\Omega f(x) \varphi(x)\, dx$$
$$\text{for all}\ \varphi \in L^{\alpha'}(\Omega), \tag{24}$$

we obtain

$$\liminf_{k \to \infty} \int_\Omega |f_k(x)|^{p_k(x)}\, dx \geq \int_\Omega p(x) f(x) \varphi(x)\, dx$$

$$- \int_\Omega \frac{p(x)}{p'(x)} |\varphi(x)|^{p'(x)}\, dx. \tag{25}$$

Since the last inequality is valid for all $\varphi \in C_0^\infty(\mathbb{R}^N)$ and $C_0^\infty(\mathbb{R}^N)$ is a dense subset of $L^{p'(\cdot)}(\Omega)$, it follows that this relation also holds true for $\varphi \in L^{p'(\cdot)}(\Omega)$. So, taking $\varphi = |f(x)|^{p(x)-2} f(x)$, we arrive at the announced inequality (21). As a consequence of (21) and estimate (A4), we get: $f \in L^{p(\cdot)}(\Omega)$.

To end the proof, it remains to observe that relation (24) holds true for $\varphi \in C_0^\infty(\mathbb{R}^N)$ as well. From this the weak convergence $f_k \rightharpoonup f$ in the variable Orlicz space $L^{p_k(\cdot)}(\Omega)$ follows. $\qquad\square$

**Remark 1** Arguing in a similar manner and using, instead of (23), the estimate

$$\liminf_{k \to \infty} \int_\Omega \frac{1}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx \geq \int_\Omega f(x) \varphi(x)\, dx$$
$$- \int_\Omega \frac{1}{p_k'(x)} |\varphi(x)|^{p'(x)}\, dx,$$

it can be shown that the lower semicontinuity property (21) can be generalized as follows:

$$\liminf_{k \to \infty} \int_\Omega \frac{1}{p_k(x)} |f_k(x)|^{p_k(x)}\, dx \geq \int_\Omega \frac{1}{p(x)} |f(x)|^{p(x)}\, dx. \tag{26}$$

Following in some technical aspects the recent studies [21,22,30], we give the following existence result.

**Theorem 2** *For each $i = 1, \ldots, M$ and given $\mu > 0$, $\mathbf{u}_0 \in L^1(\Omega; \mathbb{R}^M)$, and $T_i \in \mathcal{L}(L^1(\Omega))$, the minimization problem (16) admits at least one solution $u_i^{rec} \in \Xi_i$.*

**Proof** Since $\Xi_i \neq \emptyset$ and $0 \leq J_i(v) < +\infty$ for all $v \in \Xi_i$, it follows that there exists a non-negative value $\zeta \geq 0$ such that $\zeta = \inf_{v \in \Xi_i} J_i(v)$. Let $\{v_k\}_{k \in \mathbb{N}} \subset \Xi_i$ be a minimizing sequence to the problem (16), i.e.,

$$v_k \in \Xi_i, \ \forall k \in \mathbb{N}, \ \text{and}\ \lim_{k \to \infty} J_i(v_k) = \zeta.$$

So, without loss of generality, we can suppose that $J_i(v_k) \leq \zeta + 1$ for all $k \in \mathbb{N}$. From this and the initial assumptions, we deduce

$$\int_\Omega |v_k(x)|^\alpha\, dx$$
$$\leq \int_\Omega \gamma_{1,i}^\alpha\, dx \leq \gamma_{1,i}^\alpha |\Omega|, \quad \forall k \in \mathbb{N}, \int_\Omega |\nabla v_k(x)|^{p_k(x)}\, dx$$
$$\leq 2 \int_\Omega \frac{1}{p_k(x)} |\nabla v_k(x)|^{p_k(x)}\, dx < 2(\zeta + 1), \forall k \in \mathbb{N}, \tag{27}$$

where $p_k(x) = 1 + g\left(|(\nabla G_\sigma * v_k)(x)|\right)$ in $\Omega$, and, therefore,

$$\sup_{k \in \mathbb{N}} \left[ \sup_{x \in \Omega} p_k(x) \right] \leq 2 \ (\text{see Lemma 1}).$$

Utilizing the fact that $v_k(x) \le \gamma_{1,i}$ for almost all $x \in \Omega$, we infer the following estimate

$$\|v_k\|_{L^1(\Omega)} \le \gamma_{1,i} |\Omega|, \quad \forall k \in \mathbb{N}.$$

Then, arguing as in Lemma 1, it can be shown that $p_k \in C^{0,1}(\overline{\Omega})$ and

$$\alpha := 1 + \delta \le p_k(x) \le \beta := 2, \quad \forall x \in \Omega, \quad \forall k \in \mathbb{N}, \quad (28)$$

$$\text{with} \quad \delta = \frac{a^2}{a^2 + \|G_\sigma\|^2_{C^1(\overline{\Omega-\Omega})} \gamma^2_{1,i} |\Omega|^2}. \quad (29)$$

Taking this fact into account, we deduce from (27), (28), and (A4) that

$$\|v_k\|_{W^{1,\alpha}(\Omega)} = \left( \int_\Omega \left[ |v_k(x)|^\alpha + |\nabla v_k(x)|^\alpha \right] \mathrm{d}x \right)^{1/\alpha}$$
$$\le (1 + |\Omega|)^{1/\alpha}$$
$$\left( \int_\Omega \left[ |v_k(x)|^{p_k(x)} + |\nabla v_k(x)|^{p_k(x)} \right] \mathrm{d}x + 2 \right)^{1/\alpha}$$
$$\overset{by\ (27)}{\le} (1 + |\Omega|)^{1/\alpha} \left( \gamma^2_{1,i} |\Omega| + 2\zeta + 4 \right)^{1/\alpha}$$

uniformly with respect to $k \in \mathbb{N}$. Therefore, there exist a subsequence of $\{v_k\}_{k \in \mathbb{N}}$, still denoted by the same index, and a function $u_i^{rec} \in W^{1,\alpha}(\Omega)$ such that

$$v_k \to u_i^{rec} \text{ strongly in } L^q(\Omega) \text{ for all } q \in [1, \alpha^*),$$
$$v_k \rightharpoonup u_i^{rec} \text{ weakly in } W^{1,\alpha}(\Omega) \text{ as } k \to \infty, \quad (30)$$

where, by Sobolev embedding theorem, $\alpha^* = \frac{N\alpha}{N-\alpha} = \frac{N+N\delta}{N-1-\delta} > 1 + \delta$.

Moreover, passing to a subsequence if necessary, we have (see Proposition 1 and Lemma 1):

$$v_k(x) \to u_i^{rec}(x) \text{ a.e. in } \Omega$$
$$v_k \rightharpoonup u_i^{rec} \text{ weakly in } L^{p_k(\cdot)}(\Omega),$$
$$\nabla v_k \rightharpoonup \nabla u_i^{rec} \text{ weakly in } L^{p_k(\cdot)}(\Omega; \mathbb{R}^N),$$
$$p_k(\cdot) \to p_i^{rec}(\cdot) = 1 + g\left( \left| \left( \nabla G_\sigma * u_i^{rec} \right)(\cdot) \right| \right)$$
$$\text{uniformly in } \overline{\Omega} \text{ as } k \to \infty, \quad (31)$$

where $u_i^{rec} \in W^{1,p^{rec}(\cdot)}(\Omega)$.

Since $\gamma_{0,i} \le v_k(x) \le \gamma_{1,i}$ a.a. in $\Omega$ for all $k \in \mathbb{N}$, it follows from (31) that the limit function $u_i^{rec}$ is also subjected to the same restriction. Thus, $u_i^{rec}$ is a feasible solution to minimization problem (16).

Let us show that $u_i^{rec}$ is a minimizer of this problem. With that in mind, we note that in view of the obvious inequality

$$|T_i(v_k(x)) - u_{0,i}(x)| \le \left( \|T_i\|_{\mathcal{L}(L^1(\Omega))} \gamma_{1,i} + |u_{0,i}(x)| \right),$$

we have: The sequence $\left\{ T_i(v_k(x)) - u_{0,i}(x) \right\}_{k \in \mathbb{N}}$ is bounded in $L^1(\Omega)$, equi-integrable in $\Omega$, and in view of (31), it strongly converges in $L^1(\Omega)$ to $T_i(u_i^{rec}) - u_{0,i}$. Hence,

$$\liminf_{k \to \infty} \int_\Omega |T_i(v_k(x)) - u_{0,i}(x)| \, \mathrm{d}x$$
$$= \int_\Omega \left| T_i(u_i^{rec}(x)) - u_{0,i}(x) \right| \mathrm{d}x. \quad (32)$$

It remains to notice that due to the properties (27), (30), the sequence $\left\{ |\nabla v_k| \in L^{p_k(\cdot)}(\Omega) \right\}_{k \in \mathbb{N}}$ is bounded and weakly convergent to $|\nabla u_i^{rec}|$ in $L^\alpha(\Omega)$. Hence, by Proposition 1, the following lower semicontinuous property

$$\liminf_{k \to \infty} \int_\Omega \frac{1}{p_k(x)} |\nabla v_k(x)|^{p_k(x)} \, \mathrm{d}x$$
$$\ge \int_\Omega \frac{1}{p_i^{rec}(x)} |\nabla u_i^{rec}(x)|^{p_i^{rec}(x)} \, \mathrm{d}x \quad (33)$$

holds true.

As a result, utilizing relations (32) and (33), we finally obtain

$$\zeta = \inf_{v \in \Xi_i} J_i(v) = \lim_{k \to \infty} J_i(v_k) = \liminf_{k \to \infty} J_i(v_k) \ge J_i(u_i^{rec}).$$

Thus, $u_i^{rec}$ is a minimizer to the problem (16), whereas its uniqueness remains an open question. $\qquad \square$

## 3 Iterative Algorithm Based on the Variational Convergence of Extremal Problems

Our prime interest in this section is to propose an approximation approach to the solution of the minimization problem (16) that would be effective from numerical simulations point of view. After discussion of the main steps of the proposed iterative algorithm, we supply it by a rigorous substantiation. As an alternative approach to the study of minimization problems (16), there can be recommended the modular-proximal gradient algorithms in variable Lebesgue spaces that were recently developed in [41]. However, the efficiency of such algorithms in the context of a particular form of the objective functional (65) seems to remain an open question nowadays.

Let $i \in \{1, \dots, M\}$ be a given spectral channel. Mainly basing on the concept of relaxation of extremal problems and their variational convergence [31,33,38], we propose the following algorithm. At the first step, we set up

$$p_0(x) = 1 + g\left( \left| \left( \nabla G_\sigma * u_{0,i} \right)(x) \right| \right), \quad \forall x \in \Omega, \quad (34)$$

$$u^0 = \underset{v \in \mathcal{B}_{i,p_0(\cdot)}}{\operatorname{Argmin}} \mathcal{J}_i(v, p_0(\cdot)), \quad (35)$$

where

$$\mathcal{B}_{i,p(\cdot)} = \left\{ v \in W^{1,p(\cdot)}(\Omega) : 1 \leq \gamma_{0,i} \leq v(x) \leq \gamma_{1,i} \ a.e. \ in \ \Omega \right\},$$

$$\mathcal{J}_i(v,p) = \int_\Omega \frac{1}{p(x)} |\nabla v(x)|^{p(x)} \, dx + \mu \int_\Omega |T_i(v(x)) - u_{0,i}(x)| \, dx.$$

It is clear that $\mathcal{J}_i(v, \mathfrak{F}(v(\cdot))) = J_i(v)$ for all $v \in \Xi_i$.

Then, for each $k \geq 1$, we set

$$p_k(x) = 1 + g\left(\left|\left(\nabla G_\sigma * u^{k-1}\right)(x)\right|\right), \ \forall x \in \Omega,$$

$$u^k = \operatorname*{Argmin}_{v \in \mathcal{B}_{i,p_k(\cdot)}} \mathcal{J}_i(v, p_k(\cdot)). \tag{36}$$

It is worth no notice that such definition of the elements $u^k$ is correct. Indeed, arguing as in the proof of Theorem 2 and using convexity arguments, it can be shown that for each $p(\cdot) \in \mathfrak{S}$, there exists a unique element $u_i^{0,p(\cdot)} \in \mathcal{B}_{i,p(\cdot)}$ such that $u_i^{0,p(\cdot)} = \operatorname{Argmin}_{v \in \mathcal{B}_{i,p(\cdot)}} \mathcal{J}_i(v,p(\cdot))$. This fact reflects the principle difference between optimization problems (37) and (16), where the problem (16) can be viewed as a minimization of the growth energy functional (16) with 'the frozen exponent' $p_k(x)$.

Before proceeding further, we introduce the following set:

$$\mathfrak{S} = \left\{ h \in C(\overline{\Omega}) \, \middle| \, \begin{array}{l} |h(x) - h(y)| \leq C|x-y|, \ \forall x, y \in \Omega, \\ \alpha := 1 + \delta \leq h(x) \leq \beta := 2, \quad \forall x \in \Omega, \end{array} \right\} \tag{37}$$

where $C > 0$ and $\delta > 0$ are defined by (20) and (29), respectively.

Utilizing the arguments of the proof of Theorem 2, it can be shown that for given $i = 1, \ldots, M$, $\mu > 0$, and $\mathbf{u}_0 \in L^1(\Omega, \mathbb{R}^M)$, the sequence $\{u^k \in W^{1,p_k(\cdot)}(\Omega)\}_{k \in \mathbb{N}}$ is compact with respect to the weak topology of $W^{1,\alpha}(\Omega)$, whereas the exponents $\{p_k\}_{k \in \mathbb{N}}$ are compact with respect to the strong topology of $C(\overline{\Omega})$.

We say that a function $\widehat{u}_i$ is a weak solution to the original problem (16) if

$$\widehat{u}_i = \operatorname*{Argmin}_{v \in \mathcal{B}_{i,\widehat{p}_i(\cdot)}} \mathcal{J}_i(v, \widehat{p}_i(\cdot)), \ \widehat{u}_i \in \mathcal{B}_{i,\widehat{p}_i(\cdot)},$$

$$\widehat{p}_i(x) = 1 + g\left(\left|(\nabla G_\sigma * \widehat{u}_i)(x)\right|\right), \ \forall x \in \Omega. \tag{38}$$

**Remark 2** The relation between a weak solution and a solution to the problem (16) is very intricate. Since the uniqueness of solutions to (16) is a rather questionable option, it follows that, in principle, these definitions can describe different functions in $\Xi_i$. As immediately follows from (38), a weak solution is a merely feasible one to the original problem. However, if the problem (16) admits a unique solution $(u_i^0, p_i^0) \in \Xi_i$, then (38) implies that this function is considered as a weak solution. On the other hand, if $u_i^* \in \Xi_i$ is some solution to (16), then setting in (34) $p_0(x) = \mathfrak{F}(u_i^*(x))$

for all $x \in \Omega$, we immediately arrive at the conclusion: $(u^k, p_k) = (u_i^*, \mathfrak{F}(u_i^*(x)))$ for all $k \in \mathbb{N}$ and, therefore, $u_i^*$ is a weak solution to the above problem. So, in general, it would be provocative to assert that all weak solutions to the problem (16) are local minimizers of (16). In view of arguments given above, these solutions are merely stationary points of the original problem.

Our main goal in this section is to establish the existence of a weak solution to the original problem (16) and show that it can be attained by some iterative algorithm.

We begin with some technical results.

**Lemma 2** *For given values $i \in \{1, \ldots, M\}$, $\mu > 0$, and $\mathbf{u}_0 \in L^1(\Omega, \mathbb{R}^M)$, the sequence of minimizers $\{u^k \in W^{1,p_k(\cdot)}(\Omega)\}_{k \in \mathbb{N}}$ of (36) is compact with respect to the weak topology of $W^{1,\alpha}(\Omega)$.*

**Proof** Let us show that the sequence of minimizers $\{u^k\}_{k \in \mathbb{N}}$ of (38) is bounded in the sense of condition (A9). Let $\widehat{u} \in C^1(\overline{\Omega})$ be an arbitrary function such that $\gamma_{0,i} \leq \widehat{u}(x) \leq \gamma_{1,i}$ in $\Omega$. Since

$$\mathcal{J}_i(u^k, p_k(\cdot)) \leq \mathcal{J}_i(\widehat{u}, p_k(\cdot)), \quad \forall k = 0, 1, 2, \ldots \tag{39}$$

and

$$\int_\Omega \frac{1}{p_k(x)} |\nabla \widehat{u}(x)|^{p_k(x)} \, dx \leq \int_\Omega \frac{1}{p_k(x)} \left(1 + \|\widehat{u}\|_{C^1(\overline{\Omega})}\right)^{p_k(x)} \, dx$$

$$\leq \frac{|\Omega|}{\alpha} \left(1 + \|\widehat{u}\|_{C^1(\overline{\Omega})}\right)^2,$$

$$\int_\Omega |T_i(\widehat{u}(x)) - u_{0,i}(x)| \, dx \leq \int_\Omega \left[\|T_i\| \, \|\widehat{u}\|_{C^1(\overline{\Omega})} + |u_{0,i}(x)|\right] \, dx$$

$$\leq |\Omega| \|T_i\| \, \|\widehat{u}\|_{C^1(\overline{\Omega})} + \|u_{0,i}\|_{L^1(\Omega)},$$

it follows that

$$\sup_{k \in \mathbb{N}} \mathcal{J}_i(u^k, p_k(\cdot)) \leq \sup_{k \in \mathbb{N}} \mathcal{J}_i(\widehat{u}, p_k(\cdot)) \leq \widehat{C}$$

with some constant $\widehat{C} > 0$.

From this and definition of the set $\mathcal{B}_{i,p_k(\cdot)}$, we deduce

$$\int_\Omega |u^k(x)|^\alpha \, dx \leq \gamma_{1,i}^\alpha |\Omega|, \quad \forall k \in \mathbb{N}, \tag{40}$$

$$\int_\Omega |\nabla u^k(x)|^{p_k(x)} \, dx$$

$$\leq 2 \int_\Omega \frac{1}{p_k(x)} |\nabla u^k(x)|^{p_k(x)} \, dx < 2\widehat{C}, \quad \forall k \in \mathbb{N}. \tag{41}$$

Then, estimate (A4) implies that the sequence $\{u^k\}_{k \in \mathbb{N}}$ is bounded in $W^{1,\alpha}(\Omega)$. So, its weak compactness is a direct consequence of the reflexivity of $W^{1,\alpha}(\Omega)$. □

We notice that boundedness of the sequence $\{u^k\}_{k\in\mathbb{N}}$ in $W^{1,\alpha}(\Omega)$ and compactness of the embedding $W^{1,\alpha}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in \left[1, \frac{N\alpha}{N-\alpha}\right)$ imply the existence of an element $u^* \in W^{1,\alpha}(\Omega)$ such that, up to a subsequence,

$$u^k(x) \to u^*(x) \text{ a.e. in } \Omega, \tag{42}$$

$$u^k \to u^* \text{ in } L^q(\Omega), \text{ and } \nabla u^k \rightharpoonup \nabla u^* \text{ in } L^\alpha(\Omega; \mathbb{R}^N). \tag{43}$$

Then, using (42) and passing to the limit in two-side inequality $\gamma_{0,i} \le u^k(x) \le \gamma_{1,i}$, we obtain

$$\gamma_{0,i} \le u^*(x) \le \gamma_{1,i} \quad \text{for } a.a. \ x \in \Omega. \tag{44}$$

Utilizing this fact together with the pointwise convergence (42), by the Lebesgue dominated convergence theorem, we have

$$\lim_{k\to\infty} \mathfrak{F}(u^k(x)) = 1 + \frac{a^2}{a^2 + \left(\lim_{k\to\infty} \left|\left(\nabla G_\sigma * u^k\right)(x)\right|\right)^2}$$

$$= 1 + \frac{a^2}{a^2 + \left(\left|\left(\nabla G_\sigma * \lim_{k\to\infty} u^k\right)(x)\right|\right)^2}$$

$$= \mathfrak{F}(u^*(x)), \quad \forall x \in \Omega. \tag{45}$$

Since, by Arzelà–Ascoli theorem, the set $\{p_k = 1 + g\left(\left|\left(\nabla G_\sigma * u^{k-1}\right)(x)\right|\right)\}_{k\in\mathbb{N}}$ is compact with respect to the strong topology of $C(\overline{\Omega})$, it follows from (45) (see also the proof of Lemma 1) that

$$\begin{aligned} p_k \to p^* &= \mathfrak{F}(u^*(x)) \text{ strongly in } C(\overline{\Omega}) \\ &\text{as } k \to \infty, \text{ and } p^* \in \mathfrak{S}. \end{aligned} \tag{46}$$

Then, properties (42)–(46) and Proposition 1 imply:

$$\begin{aligned} u^* &\in \mathcal{B}_{i,p^*(\cdot)} \\ &= \left\{u \in W^{1,p^*(\cdot)}(\Omega) \mid 1 \le \gamma_{0,i} \le u(x) \le \gamma_{1,i} \text{ a.e. in } \Omega\right\}. \end{aligned} \tag{47}$$

Thus, the iterative procedure (36) has a cluster point $(u^*, p^*) \in \mathcal{B}_{i,p^*(\cdot)} \times \mathfrak{S}$ with respect to the convergence (42)–(43), (46).

The main result of this section can be stated as follows:

**Theorem 3** *Let $\mu > 0$, and $\mathbf{u}_0 \in L^1(\Omega, \mathbb{R}^M)$ be given data. Then, for each $i \in \{1, \ldots, M\}$, the sequence of approximated solutions $\{(u^k, p_k)\}_{k\in\mathbb{N}}$ possesses the asymptotic properties:*

$$u^k(x) \to \widetilde{u}_i(x) \text{ a.e. in } \Omega, \tag{48}$$

$$u^k \to \widetilde{u}_i \text{ in } L^q(\Omega), \ \forall q \in \left[1, \frac{N\alpha}{N-\alpha}\right),$$

$$\nabla u^k \rightharpoonup \nabla \widetilde{u}_i \text{ in } L^\alpha(\Omega; \mathbb{R}^N), \tag{49}$$

$$p_k \to \widetilde{p}_i = \mathfrak{F}(\widetilde{u}_i) \text{ strongly in } C(\overline{\Omega}) \text{ as } k \to \infty, \tag{50}$$

*where $\widetilde{u}_i$ is a weak solution to the problem (16), that is,*

$$\widetilde{u}_i \in \mathcal{B}_{i,\widetilde{p}_i(\cdot)}, \quad \widetilde{u}_i = \operatorname*{Argmin}_{v\in\mathcal{B}_{i,\widetilde{p}_i(\cdot)}} \mathcal{J}_i(v, \widetilde{p}_i(\cdot)),$$

*and, in addition, the following variational property holds true*

$$\begin{aligned} \lim_{k\to\infty} \mathcal{J}_i(u^k, p_k(\cdot)) &= \lim_{k\to\infty}\left[\inf_{v\in\mathcal{B}_{i,p_k(\cdot)}} \mathcal{J}_i(v, p_k(\cdot))\right] \\ &= \inf_{v\in\mathcal{B}_{i,\widetilde{p}_i(\cdot)}} \mathcal{J}_i(v, \widetilde{p}_i(\cdot)) = J_i(\widetilde{u}_i). \end{aligned} \tag{51}$$

**Proof** As immediately follows from Lemma 2 and reasoning given above, the sequence $\{(u^k, p_k)\}_{k\in\mathbb{N}}$ is compact with respect to the convergence (48)–(50). Let $(\widetilde{u}_i, \widetilde{p}_i)$ be its cluster point. In order to show that the function $\widetilde{u}_i$ is a weak solution to the problem (16), we assume the converse— namely, there is another function $z \in \mathcal{B}_{i,\widetilde{p}_i(\cdot)}$ such that

$$\mathcal{J}_i(z, \widetilde{p}_i(\cdot)) = \inf_{v\in\mathcal{B}_{i,\widetilde{p}_i(\cdot)}} \mathcal{J}_i(v, \widetilde{p}_i(\cdot)) < \mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot)) \equiv J_i(\widetilde{u}_i). \tag{52}$$

Using the procedure of the standard direct smoothing, we set

$$u_\varepsilon(x) = \frac{1}{\varepsilon^N} \int_{\mathbb{R}^N} K\left(\frac{x-s}{\varepsilon}\right) \widetilde{z}(s)\,ds,$$

where $\varepsilon > 0$ is a small parameter, $K$ is a positive compactly supported smooth function with properties

$$K \in C_0^\infty(\mathbb{R}^N), \quad \int_{\mathbb{R}^N} K(x)\,dx = 1, \quad \text{and} \quad K(x) = K(-x),$$

and $\widetilde{z}$ is zero extension of $z$ outside of $\Omega$.

Since $z \in W^{1,\widetilde{p}(\cdot)}(\Omega)$ and $\widetilde{p}(x) \ge \alpha = 1+\delta$ in $\Omega$, it follows that $z \in W^{1,\alpha}(\Omega)$. Then,

$$\begin{aligned} &u_\varepsilon \in C_0^\infty(\mathbb{R}^N) \text{ for each } \varepsilon > 0, \\ &u_\varepsilon \to z \text{ in } L^\alpha(\Omega), \quad \nabla u_\varepsilon \to \nabla z \text{ in } L^\alpha(\Omega; \mathbb{R}^N) \end{aligned} \tag{53}$$

by the classical properties of smoothing operators. From this we deduce that

$$u_\varepsilon(x) \to z(x) \text{ a.e. in } \Omega. \tag{54}$$

Moreover, taking into account the estimates

$$u_\varepsilon(x) = \int_{\mathbb{R}^N} K(y)\widetilde{z}(x-\varepsilon y)\,dy \le \gamma_{1,i} \int_{\mathbb{R}^N} K(y)\,dy = \gamma_{1,i},$$

$$u_\varepsilon(x) \geq \int_{y \in \varepsilon^{-1}(x-\Omega)} K(y) \widetilde{z}(x - \varepsilon y) \, \mathrm{d}y$$

$$\geq \gamma_{0,i} \int_{y \in \varepsilon^{-1}(x-\Omega)} K(y) \, \mathrm{d}y \geq \gamma_{0,i},$$

we see that each element $u_\varepsilon$ is subjected to the pointwise constraints

$$\gamma_{0,i} \leq u_\varepsilon(x) \leq \gamma_{1,i} \text{ a.e. in } \Omega, \quad \forall \varepsilon > 0.$$

Since, for each $\varepsilon > 0$, $u_\varepsilon \in W^{1,p_k(\cdot)}(\Omega)$ for all $k \in \mathbb{N}$, it follows that $u_\varepsilon \in \mathcal{B}_{i,p_k(\cdot)}$, i.e., each element of the sequence $\{u_\varepsilon\}_{\varepsilon > 0}$ is a feasible solution to all approximating problems $\left\langle \inf_{v \in \mathcal{B}_{i,p_k(\cdot)}} \mathcal{J}_i(v, p_k(\cdot)) \right\rangle$. Hence,

$$\mathcal{J}_i(u^k, p_k(\cdot)) \leq \mathcal{J}_i(u_\varepsilon, p_k(\cdot)), \quad \forall \varepsilon > 0, \ \forall k = 0, 1, \dots \quad (55)$$

Further, we notice that

$$\liminf_{k \to \infty} \mathcal{J}_i(u^k, p_k(\cdot)) \geq \mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot)) \quad (56)$$

by Proposition 1 and Fatou's lemma, and

$$\lim_{k \to \infty} \mathcal{J}_i(u_\varepsilon, p_k(\cdot)) = \lim_{k \to \infty} \int_\Omega \frac{1}{p_k(x)} |\nabla u_\varepsilon(x)|^{p_k(x)} \, \mathrm{d}x$$

$$+ \mu \int_\Omega |T_i(u_\varepsilon(x)) - u_{0,i}(x)| \, \mathrm{d}x. \quad (57)$$

Since

$$\frac{1}{p_k(x)} |\nabla u_\varepsilon(x)|^{p_k(x)} \to \frac{1}{\widetilde{p}_i(x)} |\nabla u_\varepsilon(x)|^{\widetilde{p}_i(x)}$$
uniformly in $\Omega$ as $k \to \infty$,

it follows from the Lebesgue dominated convergence theorem and (57) that

$$\lim_{k \to \infty} \mathcal{J}_i(u_\varepsilon, p_k(\cdot)) = \mathcal{J}_i(u_\varepsilon, \widetilde{p}_i(\cdot)), \quad \forall \varepsilon > 0. \quad (58)$$

As a result, passing to the limit in (55) and utilizing properties (56)–(58), we obtain

$$\mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot)) \leq \mathcal{J}_i(u_\varepsilon, \widetilde{p}_i(\cdot))$$

$$= \int_\Omega \frac{1}{\widetilde{p}_i(x)} |\nabla u_\varepsilon(x)|^{\widetilde{p}_i(x)} \, \mathrm{d}x$$

$$+ \mu \int_\Omega |T_i(u_\varepsilon(x)) - u_{0,i}(x)| \, \mathrm{d}x, \quad (59)$$

for all $\varepsilon > 0$. Taking into account the pointwise convergence (see (54) and property (53))

$$|\nabla u_\varepsilon(x)|^{\widetilde{p}_i(x)} \to |\nabla z(x)|^{\widetilde{p}_i(x)},$$
$$|T_i(u_\varepsilon(x)) - u_{0,i}(x)| \to |T_i(z(x)) - u_{0,i}(x)|$$

as $\varepsilon \to 0$, and the fact that, for $\varepsilon$ small enough,

$$|\nabla u_\varepsilon|^{\widetilde{p}_i(\cdot)} \leq (1 + |\nabla z|)^{\widetilde{p}_i(\cdot)} \in L^1(\Omega),$$
$$|T_i(u_\varepsilon(\cdot)) - u_{0,i}(\cdot)| \leq \left[ \|T_i\| (1 + |z(\cdot)|) + |u_{0,i}(\cdot)| \right] \in L^1(\Omega),$$

we can pass to the limit in (59) as $\varepsilon \to 0$ by the Lebesgue's dominated convergence theorem. This yields

$$\mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot)) \leq \lim_{\varepsilon \to 0} \mathcal{J}_i(u_\varepsilon, \widetilde{p}_i(\cdot)) = \mathcal{J}_i(z, \widetilde{p}_i(\cdot)).$$

Combining this inequality with (59) and (52), we finally get

$$\mathcal{J}_i(z, \widetilde{p}_i(\cdot)) = \inf_{v \in \mathcal{B}_{i,\widetilde{p}_i(\cdot)}} \mathcal{J}_i(v, \widetilde{p}_i(\cdot)) < \mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot))$$

$$\leq \mathcal{J}_i(z, \widetilde{p}_i(\cdot)),$$

that leads us to conflict with the initial assumption. Thus,

$$J_i(\widetilde{u}_i) = \mathcal{J}_i(\widetilde{u}_i, \widetilde{p}_i(\cdot)) = \inf_{v \in \mathcal{B}_{i,\widetilde{p}_i(\cdot)}} \mathcal{J}_i(v, \widetilde{p}_i(\cdot)) \quad (60)$$

and, therefore, $\widetilde{u}_i$ is a weak solution to the original problem (16). As for the variational property (51), it is a direct consequence of (60) and (58). □

## 4 Optimality Conditions

To characterize the solution $u^{0,p(\cdot)} \in \mathcal{B}_{i,p(\cdot)}$ of the approximating optimization problem $\left\langle \inf_{v \in \mathcal{B}_{i,p(\cdot)}} \mathcal{J}_i(v, p(\cdot)) \right\rangle$, we check that the functional $\mathcal{J}_i(v, p(\cdot))$ is Gâteaux differentiable with respect to $v$, that is,

$$\lim_{t \to 0} \frac{\mathcal{J}_i(u^{0,p(\cdot)} + tv, p(\cdot)) - \mathcal{J}_i(u^{0,p(\cdot)}, p(\cdot))}{t}$$

$$= \int_\Omega \left( |\nabla u^{0,p(\cdot)}(x)|^{p(x)-2} \nabla u^{0,p(\cdot)}(x), \nabla v(x) \right) \mathrm{d}x$$

$$+ \mu \int_\Omega \frac{u^{0,p(\cdot)}}{|T_i(u^{0,p(\cdot)}) - u_{0,i}|} T_i(v) \, \mathrm{d}x, \quad \forall v \in W^{1,p(\cdot)}(\Omega).$$

$$(61)$$

To this end, we note that

$$\frac{|\nabla u^{0,p(\cdot)}(x) + t\nabla v(x)|^{p(x)} - |\nabla u^{0,p(\cdot)}(x)|^{p(x)}}{p(x)t}$$

$$\to \left( |\nabla u^{0,p(\cdot)}(x)|^{p(x)-2} \nabla u^{0,p(\cdot)}(x), \nabla v(x) \right) \text{ as } t \to 0$$

almost everywhere in $\Omega$. Since, by convexity,

$$|\xi|^p - |\eta|^p \leq 2p \left( |\xi|p - 1 + |\eta|^{p-1} \right) |\xi - \eta|,$$

it follows that

$$
\left| \frac{1}{p(x)t} \left( |\nabla u^{0,p(\cdot)}(x) + t\nabla v(x)|^{p(x)} - |\nabla u^{0,p(\cdot)}(x)|^{p(x)} \right) \right|
$$
$$
\leq 2 \left( \|\nabla u^{0,p(\cdot)}(x) + t\nabla v(x)\|^{p(x)-1} \right.
$$
$$
\left. + \|\nabla u^{0,p(\cdot)}(x)\|^{p(x)-1} \right) \|\nabla v(x)\|
$$
$$
\leq \text{const} \left( |\nabla u^{0,p(\cdot)}(x)|^{p(x)-1} + |\nabla v(x)|^{p(x)-1} \right) |\nabla v(x)|.
$$
(62)

Taking into account that

$$
\int_\Omega |\nabla u^{0,p(\cdot)}(x)|^{p(x)-1} |\nabla v(x)| \, dx
$$
$$
\leq 2\| |\nabla u^{0,p(\cdot)}(x)|^{p(x)-1} \|_{L^{p'(\cdot)}(\Omega)} \| |\nabla v(x)| \|_{L^{p(\cdot)}(\Omega)}
$$
$$
\leq 2\| |\nabla u^{0,p(\cdot)}(x)|^{p(x)-1} \|_{L^{p'(\cdot)}(\Omega,\mathbb{R}^N)} \|\nabla v(x)\|_{L^{p(\cdot)}(\Omega,\mathbb{R}^N)},
$$

and $\int_\Omega |\nabla v(x)|^{p(x)} \, dx \overset{by\,(A4)}{\leq} \|\nabla v\|^2_{L^{p(\cdot)}(\Omega,\mathbb{R}^N)} + 1$, we see that the right-hand side of inequality (62) is an $L^1(\Omega)$-function. Therefore,

$$
\int_\Omega \frac{|\nabla u^{0,p(\cdot)}(x) + t\nabla v(x)|^{p(x)} - |\nabla u^{0,p(\cdot)}(x)|^{p(x)}}{p(x)t} \, dx
$$
$$
\rightarrow \int_\Omega \left( |\nabla u^{0,p(\cdot)}(x)|^{p(x)-2} \nabla u^{0,p(\cdot)}(x), \nabla v(x) \right) \, dx
$$
$$
\text{as } t \rightarrow 0
$$

by the Lebesgue's dominated convergence theorem.

Utilizing similar arguments to the rest of terms in (10), we deduce that the representation (61) for the Gâteaux differential of $\mathcal{J}_i(\cdot, p(\cdot))$ at the point $u^{0,p(\cdot)} \in \mathcal{B}_{i,p(\cdot)}$ is valid.

Thus, in order to derive some optimality conditions for the minimizing element $u^{0,p(\cdot)} \in \mathcal{B}_{i,p(\cdot)}$ to the problem $\inf_{v \in \mathcal{B}_{i,p(\cdot)}} \mathcal{J}_i(v, p(\cdot))$, we note that $\mathcal{B}_{i,p(\cdot)}$ is a nonempty convex subset of $W^{1,p(\cdot)}(\Omega)$ and the objective functional $\mathcal{J}_i(\cdot, p(\cdot)) : \mathcal{B}_{i,p(\cdot)} \rightarrow \mathbb{R}$ is strictly convex. Hence, the well-known classical result (see [43, Theorem 1.1.3]) and representation (61) lead us to the following conclusion.

**Theorem 4** *Let $p_k(\cdot) \in \mathfrak{S}$ be an exponent given by the iterative rule (36). Then, the unique minimizer $u^k \in \mathcal{B}_{i,p_k(\cdot)}$ to the approximating problem $\inf_{v \in \mathcal{B}_{i,p_k(\cdot)}} \mathcal{J}_i(v, p_k(\cdot))$ is characterized by*

$$
\int_\Omega \left( |\nabla u^k(x)|^{p_k(x)-2} \nabla u^k(x), \nabla v(x) - \nabla u^k(x) \right) \, dx
$$
$$
+ \mu \int_\Omega \frac{u^k}{|T_i(u^k) - u_{0,i}|} \left( T_i \left( v(x) - u^k(x) \right) \right) \, dx \geq 0,
$$
$$
\forall \, v \in \mathcal{B}_{i,p_k(\cdot)}.
$$
(63)

# 5 On Approximation of the Reconstruction Problem

It is clear that because of the nonstandard growth energy functional and its non-convexity, constrained minimization problem (16) is not trivial in its practical implementation. The main difficulty in its study comes from the state constraints

$$
0 \leq \gamma_{0,i} \leq v(x) \leq \gamma_{1,i} \text{ a.e. in } \Omega, \; \mathfrak{F}(v(x))
$$
$$
= 1 + g\left(|(\nabla G_\sigma * v)(x)|\right)
$$
(64)

that we impose on the variable exponent and the set of feasible solutions $\Xi_i$. The iterative algorithm that was proposed in the previous sections allows to attain in the limit the so-called weak solution to the minimization problem (16). So, the question whether some optimal solutions can be reachable in such way remains open. This motivates us to pass to another relaxation scheme of the original variational problem.

As the main step of this procedure, we propose to consider the function $p(\cdot) := \mathfrak{F}(v(\cdot))$ as a fictitious control subjected to some special constraints and interpret the fulfillment of equality $\mathfrak{F}(v(x)) = 1 + g\left(|(\nabla G_\sigma * v)(x)|\right)$ with some accuracy in $\Omega$. To do so, we notice that if $v \in \Xi_i$ is a feasible solution to (16), then $\mathfrak{F}(v(\cdot))$ is subjected to the two-side inequality (28) with $\delta \in (0, 1)$ given by (29). Keeping this in mind and following in some aspects the standard penalty method (see, [52, Chapter 2], see also [34–37,40]), we consider the following family of approximating problems:

$$
\text{Minimize } J_{i,\varepsilon}(v, p) = \int_\Omega \frac{1}{p(x)} |\nabla v(x)|^{p(x)} \, dx
$$
$$
+ \mu \int_\Omega |T_i(v(x)) - u_{0,i}(x)| \, dx
$$
$$
+ \frac{1}{\varepsilon} \int_\Omega |p(x) - 1 - g\left(|(\nabla G_\sigma * v)(x)|\right)|^2 \, dx
$$
(65)

subject to the constraints $(v, p) \in \Xi_{i,\varepsilon}$, where

$$
\Xi_{i,\varepsilon} = \left\{ (v, p) \, \middle| \, \begin{array}{l} v \in W^{1,\alpha}(\Omega), \; p \in \mathfrak{S}_{ad}, \; J_{i,\varepsilon}(v, p) < +\infty, \\ 0 \leq \gamma_{0,i} \leq v(x) \leq \gamma_{1,i} \text{ a.e. in } \Omega, \end{array} \right\}
$$
(66)

$$
\mathfrak{S}_{ad} = \left\{ h \in C(\Omega) \, \middle| \, \begin{array}{l} |h(x) - h(y)| \leq C|x - y|, \; \forall \, x, y \in \Omega, \\ 1 < \alpha \leq h(\cdot) \leq \beta \text{ in } \overline{\Omega}. \end{array} \right\}
$$
(67)

with $C > 0$ given by (20), $\alpha = 1 + \delta$, and $\delta > 0$ is given by (29).

We assume that the parameter $\varepsilon$ varies within a strictly decreasing sequence of positive real numbers which converge to 0. So, when we write $\varepsilon > 0$, we consider only the elements of this sequence.

**Definition 1** We say that a pair $(v, p)$ is quasi-feasible to minimization problem (16) if $(v, p) \in \Xi_{i,\varepsilon}$. We also say that $(u^0_{i,\varepsilon}, p^0_{i,\varepsilon}) \in W^{1, p^0_\varepsilon(\cdot)}(\Omega) \times C^{0,1}(\overline{\Omega})$ is a quasi-optimal solution to the problem (16) if $(u^0_{i,\varepsilon}, p^0_{i,\varepsilon}) \in \Xi_{i,\varepsilon}$ and $J_{i,\varepsilon}(u^0_{i,\varepsilon}, p^0_{i,\varepsilon}) = \inf\limits_{(v,p)\in\Xi_{i,\varepsilon}} J_{i,\varepsilon}(v, p)$.

**Remark 3** It is clear that the condition $p \in \mathfrak{S}_{ad}$ together with the fact that $\mathfrak{S}_{ad}$ is a compact subset in $C(\overline{\Omega})$ imply: Every cluster point of a sequence $\{p_k\}_{k\in\mathbb{N}} \subset \mathfrak{S}_{ad}$ with respect to the uniform topology is a regular exponent, i.e., it is an exponent satisfying the log-Hölder continuity condition [55]. In this case, the set $C_0^\infty(\mathbb{R}^N)$ is dense in $W^{1,p(\cdot)}(\Omega)$ [20] and this fact plays a crucial role in approximation of minimization problem (16).

The principle point in the statement of approximated problem (65) is the fact that we pass from the state constrained optimization problem (16) with the variable exponent $p(x) = \mathfrak{F}(v(x))$ strongly depending on the function of interest $v$ to its approximation where we eliminate the equality constraint $p(x) = \mathfrak{F}(v(x))$ between the state $v(x)$ and exponent $p(x)$ and allow such pairs run freely in their respective sets of feasibility.

We begin with the following existence result.

**Theorem 5** *For each $i = 1, \dots, M$, every positive value $\varepsilon > 0$, and given $\mu > 0$, $u_{i,0} \in L^1(\Omega)$, and $T_i \in \mathcal{L}(L^1(\Omega))$, the minimization problem (65) has at least one solution.*

**Proof** The proof follows the steps of that of Theorem 2. Since the set $\Xi_{i,\varepsilon}$ is nonempty, we can take a minimizing sequence $\{(u_k, p_k)\}_{k\in\mathbb{N}} \subset \Xi_{i,\varepsilon}$. Then, arguing as in the proof of Theorem 2, we deduce the boundedness of $\{(u_k, p_k)\}_{k\in\mathbb{N}}$ in $W^{1,p_k(\cdot)}(\Omega) \times C^{0,1}(\Omega)$ and, hence, the existence of a subsequence, still denoted in the same way, and a pair $(u^0_\varepsilon, p^0_\varepsilon) \in \Xi_{i,\varepsilon}$ such that $p_k \to p^0_\varepsilon$ in $C(\overline{\Omega})$, $u_k \rightharpoonup u^0_\varepsilon$ in $W^{1,\alpha}(\Omega)$ and in variable $W^{1,p_k(\cdot)}(\Omega)$. Then, by the Sobolev embedding theorem, we deduce that $u_k \to u^0_\varepsilon$ strongly in $L^q(\Omega)$ for all $q \in [1, \frac{\alpha N}{N-\alpha})$, and, therefore, we can suppose that $u_k(x) \to u^0_\varepsilon(x)$ almost everywhere in $\Omega$ as $k \to \infty$. As a result, we have

$$\gamma_{0,i} \le u^0_\varepsilon(x) \le \gamma_{1,i} \quad \text{and} \quad \alpha \le p^0_\varepsilon(x) \le \beta \text{ a.a. in } \Omega,$$

$$\lim_{k\to\infty} \int_\Omega |T_i(u_k(x)) - u_{0,i}(x)| \, dx$$
$$= \int_\Omega |T_i(u^0_\varepsilon(x)) - u_{0,i}(x)| \, dx,$$

$$\liminf_{k\to\infty} \int_\Omega \frac{1}{p_k(x)} |\nabla u_k(x)|^{p_k(x)} \, dx$$
$$\overset{by\ (26)}{\ge} \int_\Omega \frac{1}{p^0_\varepsilon(x)} |\nabla u^0_\varepsilon(x)|^{p^0_\varepsilon(x)} \, dx$$

Thus, $(u^0_\varepsilon, p^0_\varepsilon) \in \Xi_{i,\varepsilon}$. It remains to notice that

$$\left| p_k - 1 - g\left(|(\nabla G_\sigma * u_k)(x)|\right) \right|^2$$
$$\to \left| p^0_\varepsilon - 1 - g\left(\left|\left(\nabla G_\sigma * u^0_\varepsilon\right)(x)\right|\right) \right|^2 \text{ in } C(\overline{\Omega}),$$

and the Lebesgue's dominated convergence theorem implies

$$\lim_{k\to\infty} \int_\Omega \left| p_k - 1 - g\left(|(\nabla G_\sigma * u_k)(x)|\right) \right|^2 dx$$
$$= \int_\Omega \left| p^0_\varepsilon - 1 - g\left(\left|\left(\nabla G_\sigma * u^0_\varepsilon\right)(x)\right|\right) \right|^2 dx.$$

Utilizing these properties, we finally obtain

$$J_{i,\varepsilon}(u^0_\varepsilon, p^0_\varepsilon) \le \liminf_{k\to\infty} J_{i,\varepsilon}(u_k, p_k) = \inf_{(v,p)\in\Xi_i} J_{i,\varepsilon}(v, p).$$

So, $(u^0_\varepsilon, p^0_\varepsilon) \in \Xi_{i,\varepsilon}$ is an optimal pair to the problem (65). □

Taking this existence result into account, we pass to the study of approximation properties of the problems (65). Namely, we establish the convergence of minima of (65) to minima of (16) as $\varepsilon$ tends to zero. In other words, we show that some optimal solutions to (16) can be approximated by the quasi-optimal solutions of this problem.

**Theorem 6** *Let $\left\{(u^0_\varepsilon, p^0_\varepsilon) \in \Xi_{i,\varepsilon}\right\}_{\varepsilon>0}$ be a sequence of minimizers to the problem (65). Then, there exists a subsequence of $\left\{(u^0_\varepsilon, p^0_\varepsilon)\right\}_{\varepsilon>0}$, still denoted by the same index $\varepsilon$, such that*

$$p^0_\varepsilon \to p^0 \text{ in } C(\overline{\Omega}) \text{ as } \varepsilon \to 0, \tag{68}$$
$$u^0_\varepsilon \rightharpoonup u^0 \text{ in } W^{1,\alpha}(\Omega) \text{ as } \varepsilon \to 0, \tag{69}$$
$$u^0_\varepsilon \rightharpoonup u^0 \text{ in } W^{1,p^0_\varepsilon(\cdot)}(\Omega), \; u^0 \in W^{1,p^0(\cdot)}(\Omega), \tag{70}$$
$$p^0(x) = 1 + g\left(\left|\left(\nabla G_\sigma * u^0\right)(x)\right|\right) \quad in \; \Omega, \tag{71}$$
$$J_i(u^0) = \inf_{(v)\in\Xi_i} J_i(v)$$
$$= \lim_{\varepsilon\to 0} \inf_{(u,p)\in\Xi_{i,\varepsilon}} J_{i,\varepsilon}(v, p) = \lim_{\varepsilon\to 0} J_{i,\varepsilon}(u^0_\varepsilon, p^0_\varepsilon), \tag{72}$$

*where $u^0 \in \Xi_i$.*

**Proof** Let $u^* \in \Xi_i$ be an arbitrary feasible solution to the original problem (16). We set $p^* = \mathfrak{F}(u^*(\cdot))$ in $\Omega$. Then, $u^* \in W^{1,\alpha}(\Omega)$, $p^* \in \mathfrak{S}_{ad}$, $J_{i,\varepsilon}(u^*, p^*) = J_i(u^*) < +\infty$, and, as a consequence, $(u^*, p^*) \in \Xi_{i,\varepsilon}$ for each $\varepsilon > 0$.

Since $J_{i,\varepsilon}(u^0_\varepsilon, p^0_\varepsilon) \le J_{i,\varepsilon}(u^*, p^*) = J_i(u^*) =: C^*$, it follows from (65) that

$$\sup_{\varepsilon>0} \int_\Omega \frac{1}{p^0_\varepsilon(x)} |\nabla u^0_\varepsilon(x)|^{p^0_\varepsilon(x)} \, dx \le C^*, \tag{73}$$
$$\int_\Omega \left| p^0_\varepsilon(x) - 1 - g\left(\left|\left(\nabla G_\sigma * u^0_\varepsilon\right)(x)\right|\right) \right|^2 dx$$
$$\le \varepsilon C^*, \quad \forall \varepsilon > 0. \tag{74}$$

Since $\left\{ p_\varepsilon^0 \in C^{0,1}(\overline{\Omega}) \right\}$ is a bounded sequence in $C(\overline{\Omega})$ with the same modulus of continuity, it follows, by Arzelà–Ascoli theorem, that this sequence is relatively compact with respect to the strong topology of $C(\overline{\Omega})$. Without loss of generality, we can suppose that there exists a function $p^0 \in C(\overline{\Omega})$ such that assertion (68) is valid. Moreover, as follows from definition of the set $\mathfrak{S}_{ad}$, the limit function $p^0$ is subjected to the pointwise constraints

$$\alpha := 1 + \delta \leq p^0(x) \leq \beta := 2, \quad \forall\, x \in \Omega. \tag{75}$$

Arguing in a similar manner, we can infer from (73) and the two-side inequality

$$0 \leq \gamma_{0,i} \leq u_\varepsilon^0(x) \leq \gamma_{1,i} \quad \text{a.a. in } \Omega, \ \forall\, \varepsilon > 0 \tag{76}$$

that the sequence $\left\{ u_\varepsilon^0 \right\}$ is relatively compact with respect to the weak topology of $W^{1,\alpha}(\Omega)$. Indeed, taking into account (76) and observing that

$$\sup_{\varepsilon > 0} \int_\Omega |u_\varepsilon^0(x)|^{p_\varepsilon^0(x)}\, dx \overset{by\ (76)}{\leq} +\infty,$$

$$\sup_{\varepsilon > 0} \int_\Omega |\nabla u_\varepsilon^0(x)|^{p_\varepsilon^0(x)}\, dx$$

$$\leq \int_\Omega \frac{\sup_{\varepsilon > 0} \|p_\varepsilon^0\|_{C(\overline{\Omega})}}{p_\varepsilon^0(x)} |\nabla u_\varepsilon^0(x)|^{p_\varepsilon^0(x)}\, dx$$

$$\overset{by\ (73)}{\leq} \sup_{\varepsilon > 0} \|p_\varepsilon^0\|_{C(\overline{\Omega})} C^* < +\infty,$$

we see that $u_\varepsilon^0 \in W^{1,p_\varepsilon^0(\cdot)}(\Omega)$ for all $\varepsilon > 0$ and the sequence $\left\{ u_\varepsilon^0 \right\}$ is bounded in variable space $W^{1,p_\varepsilon^0(\cdot)}(\Omega)$. Hence, this sequence is bounded in $W^{1,\alpha}(\Omega)$. Therefore, in view of completeness of $W^{1,\alpha}(\Omega)$, there exists a function $u^0 \in W^{1,\alpha}(\Omega)$ such that, up to a subsequence, property (69) holds true. As a result, Proposition 1 and Sobolev embedding theorem lead us to the conclusion:

$$u_\varepsilon^0 \rightharpoonup u^0 \text{ in } W^{1,p_\varepsilon^0(\cdot)}(\Omega), \ u^0 \in W^{1,p^0(\cdot)}(\Omega),$$
$$u_\varepsilon^0 \to u^0 \text{ strongly in } L^q(\Omega) \text{ for all } q \in [1, \alpha^*), \tag{77}$$

where $\alpha^* = \frac{N\alpha}{N-\alpha}$. So, we can suppose that $u_\varepsilon^0(x) \to u^0(x)$ a.e. in $\Omega$. Then, passing to the limit in (76) as $\varepsilon \to 0$, we see that the limit function $u^0$ is also subjected by the pointwise constraints

$$0 \leq \gamma_{0,i} \leq u^0(x) \leq \gamma_{1,i} \quad \text{a.a. in } \Omega. \tag{78}$$

Moreover, utilizing the estimate (74) and properties (68)–(69), we get

$$\lim_{\varepsilon \to 0} \int_\Omega \left| p_\varepsilon^0(x) - 1 - g\left( \left| \left( \nabla G_\sigma * u_\varepsilon^0 \right)(x) \right| \right) \right|^2 dx$$

$$= \int_\Omega \left| p^0(x) - 1 - g\left( \left| \left( \nabla G_\sigma * u^0 \right)(x) \right| \right) \right|^2 dx = 0.$$

Hence, $p^0(x) = 1 + g\left( \left| \left( \nabla G_\sigma * u^0 \right)(x) \right| \right)$ in $\Omega$. Thus, $u^0 \in W^{1, \mathfrak{F}(u^0(\cdot))}(\Omega)$. Combining this fact with (78), we see that the limit function $u^0$ is a feasible solution to minimization problem (16).

Let us show that this function is optimal to the problem (16). Since

$$\lim_{\varepsilon \to 0} \int_\Omega |T_i(u_\varepsilon^0(x)) - u_{0,i}(x)|\, dx$$
$$\overset{by\ (77)}{=} \int_\Omega |T_i(u^0(x)) - u_{0,i}(x)|\, dx$$

it follows from Proposition 1 and Remark 1 that

$$\liminf_{\varepsilon \to 0} J_{i,\varepsilon}(u_\varepsilon^0, p_\varepsilon^0) \geq J_i(u^0). \tag{79}$$

Then, assuming the converse—namely, there is a function $\widehat{u} \in \Xi_i$ such that $J_i(\widehat{u}) < J_i(u^0)$, we get: $(\widehat{u}, \widehat{p}) \in \Xi_{i,\varepsilon}$ for all $\varepsilon > 0$ with $\widehat{p} := \mathfrak{F}(\widehat{u}(\cdot))$, and

$$J_i(\widehat{u}) \equiv J_{i,\varepsilon}(\widehat{u}, \widehat{p}) \geq \inf_{(v,p) \in \Xi_{i,\varepsilon}} J_{i,\varepsilon}(v, p) = J_{i,\varepsilon}(u_\varepsilon^0, p_\varepsilon^0),$$
$$J_i(\widehat{u}) \geq \limsup_{\varepsilon \to 0} J_{i,\varepsilon}(u_\varepsilon^0, p_\varepsilon^0) \geq \liminf_{\varepsilon \to 0} J_{i,\varepsilon}(u_\varepsilon^0, p_\varepsilon^0)$$
$$\overset{by\ (79)}{\geq} J_i(u^0). \tag{80}$$

So, we come into contradiction with the initial assumption. Thus, $u^0$ is a solution of the original problem (16). In order to establish the equality (72), it is enough, instead of $(\widehat{u}, \widehat{p})$, to take $(u^0, p^0)$ in (80). □

Since Theorem 6 does not give an answer whether the entire set of solutions to problem (16) can be attained in such a way, the following result sheds some light on this matter.

**Corollary 1** *Let $u^0 \in \Xi$ be a minimizer to optimization problem* (16) *such that there is a closed neighborhood $\mathcal{U}(u^0)$ of $u^0$ in the norm topology of $L^\alpha(\Omega)$ satisfying*

$$J_i(u^0) < J_i(v) \quad \forall\, v \in \Xi_i \cap \mathcal{U}(u^0). \tag{81}$$

*Then, there exists a sequence of local minima $\left\{ (u_\varepsilon^0, p_\varepsilon^0) \right\}_{\varepsilon > 0}$ of problems* (65) *such that*

$$(u_\varepsilon^0, p_\varepsilon^0) \to (u^0, \mathfrak{F}(u^0(\cdot))) \quad \text{in the sense of Theorem 6.}$$

**Proof** By the strict local optimality of $u^0$, we have that it is the unique solution of the problem

$$\min_{v \in \Xi_i, v \in \mathcal{U}(u^0)} I_i(v). \tag{82}$$

**Fig. 2** Example. Left panel: noisy satellite image. Middle panel: reconstruction using total variation (TV) approach. Right panel: reconstruction using the approach introduced in this paper
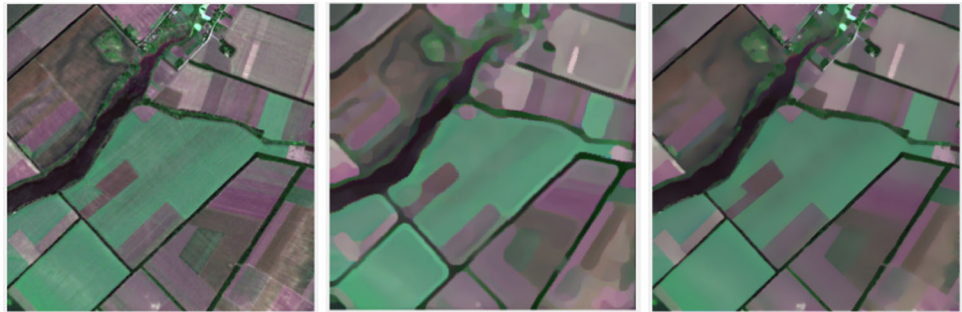
For every $\varepsilon > 0$, let us consider the following optimization problems:

$$\min_{(v,p) \in \Xi_{i,\varepsilon}, v \in \mathcal{U}(u^0)} J_{i,\varepsilon}(v, p). \tag{83}$$

Since the set $\{(v, p) \in \Xi_{i,\varepsilon}, v \in \mathcal{U}(u^0)\}$ is nonempty, it follows that the problem (83) has at least one solution $(u_\varepsilon^0, p_\varepsilon^0)$ for every $\varepsilon > 0$. Now, arguing as in the proof of Theorem 6, we deduce that $(u_\varepsilon^0, p_\varepsilon^0) \to (\tilde{u}^0, \tilde{p}^0)$ in the sense of convergences (68)–(72), and $\tilde{u}^0$ is a solution of (82). Since $u^0$ is the unique solution of (82), we infer that $u^0 = \tilde{u}^0$ and, therefore, $(u_\varepsilon^0, p_\varepsilon^0) \to (u^0, \mathfrak{F}(u^0(\cdot)))$ in the sense of Theorem 6. This implies the existence of $\varepsilon^0 > 0$ such that $u_\varepsilon^0$ belongs to the interior of $\mathcal{U}(u^0)$ for every $\varepsilon \le \varepsilon^0$. Consequently, $(u_\varepsilon^0, p_\varepsilon^0)$ is a local minimum of (65) for every $\varepsilon \le \varepsilon^0$. This concludes the proof. $\qquad\square$

# 6 Numerical Experiments

In order to illustrate the proposed algorithm for the denoising of satellite hyperspectral images, we have provided some numerical experiments with $T_i = Id$ for three channels $i = R, G, B$. As input data, we have used some Sentinel-2 images over the Dnipro area, Ukraine (see Fig. 2). This region represents a typical agricultural area with medium sides fields of various shapes.

To conduct the numerical simulations, we have dropped the two-side constraints $1 \le \gamma_{0,i} \le v(x) \le \gamma_{1,i}$ from the sets $\mathcal{B}_{i,p(\cdot)}$, and instead, we controlled the fulfilment of this condition at each time step of the numerical approximations. The solution procedure uses a parabolic equation with time as an evolution parameter, or equivalently, the gradient descent method. It means that we utilize the iterative algorithm, described in Sect. 3, and Theorem 4. As a result, at $k$-th step of the iterative algorithm we solve the following Cauchy–Neumann problem:

$$\frac{\partial u^k}{\partial t} = \operatorname{div}\left(|\nabla u^k|^{p_k(x)-2} \nabla u^k\right)$$

$$-\mu \frac{T_i^* u^k}{|T_i(u^k) - u_{0,i}|}, \quad (t, x) \in (0, \infty) \times \Omega, \tag{84}$$

$$u^k(0, x) = u_{0,i} \quad \forall x \in \Omega, \tag{85}$$

$$\frac{\partial u^k(t, \cdot)}{\partial n} = 0 \quad \text{on } \partial\Omega \tag{86}$$

with $p_k(\cdot)$ defined in (36).

As $t$ increases, we approach a denoised version of our image and the $k$-th iteration. The numerical scheme in two spatial dimensions is as follows. We let:

$$x_l = lh, \quad y_j = jh, \quad l, j = 0, 1, \ldots, N \quad \text{with } Nh = 1,$$
$$t_n = n\Delta t, \quad n = 0, 1, \ldots, \quad u_{lj}^n = u^k(t_n, x_l, y_j),$$
$$u_{lj}^0 = u_{0,i}(lh, jh).$$

Then, the numerical approximation to the problem (84)–(86) takes the form:

$$u_{lj}^{n+1} = u_{lj}^n$$
$$+ \frac{\Delta t}{h} \Delta_-^x \left( \frac{\Delta_+^x u_{lj}^n}{\sqrt{\left(\Delta_+^x u_{lj}^n\right)^2 + \left(m\left(\Delta_+^y u_{lj}^n, \Delta_-^y u_{lj}^n\right)\right)^2 + \varepsilon}} \right)$$
$$+ \frac{\Delta t}{h} \Delta_-^x \left( \frac{\Delta_+^y u_{lj}^n}{\sqrt{\left(\Delta_+^y u_{lj}^n\right)^2 + \left(m\left(\Delta_+^x u_{lj}^n, \Delta_-^x u_{lj}^n\right)\right)^2 + \varepsilon}} \right)$$
$$- \Delta t \mu \frac{T_i^* u_{lj}^n}{|T_i u_{lj}^n - u_{lj}^0 + \varepsilon|}, \quad \forall l, j = 1, \ldots, N \tag{87}$$

with boundary conditions

$$u_{0j}^n = u_{1j}^n, \quad u_{Nj}^n = u_{N-1,j}^n, \quad u_{l0}^n = u_{l1}^n, \quad u_{lN}^n = u_{l,N-1}^n.$$

Here, $\Delta_\pm^x u_{lj} = \pm\left(u_{l\pm1,j} - u_{lj}\right), \Delta_\pm^y u_{lj} = \pm\left(u_{l,j\pm1} - u_{lj}\right)$, and

$$m(a, b) = \operatorname{minmod}(a, b) = \frac{\operatorname{sgn} a + \operatorname{sgn} b}{2} \min(|a|, |b|).$$

A step size restriction is imposed for stability by condition: $\Delta/h^2 \leq C$.

For numerical simulations in this section, we set: $\varepsilon = 0.01, \sigma = 0.5, a = 5, \beta = 2, \gamma_{0,i} = 1, \gamma_{1,i} = 255, T_i = Id$, $i = 1, 2, 3$. For the three-channel satellite image depicted on the left panel in Fig. 2, we have conducted $k = 5$ iterations with the time step $\Delta t = 0.01$ at each iteration to obtain result seen on the right panel in Fig. 2. For comparison, Fig. 2b shows the result of a denoising using the ROF model (3).

## Declarations

## Appendix A: On Orlicz Spaces

Let $p(\cdot)$ be a measurable exponent function on $\Omega$ such that $1 < \alpha \leq p(x) \leq \beta < \infty$ a.e. in $\Omega$, where $\alpha$ and $\beta$ are given constants. Let $p'(\cdot) = \frac{p(\cdot)}{p(\cdot)-1}$ be the corresponding conjugate exponent. It is clear that

$$1 \leq \underbrace{\frac{\beta}{\beta-1}}_{\beta'} \leq p'(x) \leq \underbrace{\frac{\alpha}{\alpha-1}}_{\alpha'} \quad \text{a.e. in } \Omega,$$

where $\beta'$ and $\alpha'$ stand for the conjugates of constant exponents. Denote by $L^{p(\cdot)}(\Omega)$ the set of all measurable functions $f(x)$ on $\Omega$ such that $\int_\Omega |f(x)|^{p(x)}\, dx < \infty$. Then, $L^{p(\cdot)}(\Omega)$ is a reflexive separable Banach space with respect to the Luxemburg norm (see [20,24] for the details)

$$\|f\|_{L^{p(\cdot)}(\Omega)} = \inf\left\{\lambda > 0 \mid \rho_p(\lambda^{-1}f) \leq 1\right\}, \tag{A1}$$

where $\rho_p(f) := \int_\Omega |f(x)|^{p(x)}\, dx$.

It is well known that $L^{p(\cdot)}(\Omega)$ is reflexive provided $\alpha > 1$, and its dual is $L^{p'(\cdot)}(\Omega)$, that is, any continuous functional $F = F(f)$ on $L^{p(\cdot)}(\Omega)$ has the form (see [55, Lemma 13.2]) $F(f) = \int_\Omega fg\, dx$, with $g \in L^{p'(\cdot)}(\Omega)$.

**Proposition 7** *The infimum in* (A1) *is attained if* $\rho_p(f) > 0$. *Moreover,*

$$if \ \lambda_* := \|f\|_{L^{p(\cdot)}(\Omega)} > 0, \ then \ \rho_p(\lambda_*^{-1}f) = 1. \tag{A2}$$

Taking this result and condition $1 \leq \alpha \leq p(x) \leq \beta$ into account, we see that

$$\frac{1}{\lambda_*^\beta} \int_\Omega |f(x)|^{p(x)}\, dx \leq \int_\Omega \left|\frac{f(x)}{\lambda_*}\right|^{p(x)}\, dx$$
$$\leq \frac{1}{\lambda_*^\alpha} \int_\Omega |f(x)|^{p(x)}\, dx,$$
$$\frac{1}{\lambda_*^\beta} \int_\Omega |f(x)|^{p(x)}\, dx \leq 1 \leq \frac{1}{\lambda_*^\alpha} \int_\Omega |f(x)|^{p(x)}\, dx.$$

Hence, (see [20,24,54] for the details)

$$\|f\|_{L^{p(\cdot)}(\Omega)}^\alpha \leq \int_\Omega |f(x)|^{p(x)}\, dx \leq \|f\|_{L^{p(\cdot)}(\Omega)}^\beta,$$
$$\text{if } \|f\|_{L^{p(\cdot)}(\Omega)} > 1,$$
$$\|f\|_{L^{p(\cdot)}(\Omega)}^\beta \leq \int_\Omega |f(x)|^{p(x)}\, dx \leq \|f\|_{L^{p(\cdot)}(\Omega)}^\alpha,$$
$$\text{if } \|f\|_{L^{p(\cdot)}(\Omega)} < 1, \tag{A3}$$

and, therefore,

$$\|f\|_{L^{p(\cdot)}(\Omega)}^\alpha - 1 \leq \int_\Omega |f(x)|^{p(x)}\, dx \leq \|f\|_{L^{p(\cdot)}(\Omega)}^\beta + 1,$$
$$\forall f \in L^{p(\cdot)}(\Omega), \tag{A4}$$
$$\|f\|_{L^{p(\cdot)}(\Omega)} = \int_\Omega |f(x)|^{p(x)}\, dx,$$
$$\text{if } \|f\|_{L^{p(\cdot)}(\Omega)} = 1. \tag{A5}$$

The following estimates are well known (see, for instance, [20,24,54]): If $f \in L^{p(\cdot)}(\Omega)$, then

$$\|f\|_{L^\alpha(\Omega)} \leq (1 + |\Omega|)^{1/\alpha} \|f\|_{L^{p(\cdot)}(\Omega)}, \tag{A6}$$
$$\|f\|_{L^{p(\cdot)}(\Omega)} \leq (1 + |\Omega|)^{1/\beta'} \|f\|_{L^\beta(\Omega)},$$
$$\beta' = \frac{\beta}{\beta-1}, \quad \forall f \in L^\beta(\Omega). \tag{A7}$$

Let $\{p_k\}_{k \in \mathbb{N}} \subset C^{0,\delta}(\overline{\Omega})$, with some $\delta \in (0, 1]$, be a given sequence of exponents. Hereinafter, in this subsection, we

assume that

$$p, p_k \in C^{0,\delta}(\overline{\Omega}) \text{ for } k = 1, 2, \ldots, \text{ and}$$

$$p_k(\cdot) \to p(\cdot) \text{ uniformly in } \overline{\Omega} \text{ as } k \to \infty. \tag{A8}$$

We associate with this sequence the following collection $\left\{ f_k \in L^{p_k(\cdot)}(\Omega) \right\}_{k \in \mathbb{N}}$. The characteristic feature of this set of functions is that each element $f_k$ lives in the corresponding Orlicz space $L^{p_k(\cdot)}(\Omega)$. We say that the sequence $\left\{ f_k \in L^{p_k(\cdot)}(\Omega) \right\}_{k \in \mathbb{N}}$ is bounded if (see [32, Section 6.2])

$$\limsup_{k \to \infty} \int_\Omega |f_k(x)|^{p_k(x)} \, dx < +\infty. \tag{A9}$$

**Definition 2** A bounded sequence $\left\{ f_k \in L^{p_k(\cdot)}(\Omega) \right\}_{k \in \mathbb{N}}$ is weakly convergent in the variable Orlicz space $L^{p_k(\cdot)}(\Omega)$ to a function $f \in L^{p(\cdot)}(\Omega)$, where $p \in C^{0,\delta}(\overline{\Omega})$ is the limit of $\{p_k\}_{k \in \mathbb{N}} \subset C^{0,\delta}(\overline{\Omega})$ in the uniform topology of $C(\overline{\Omega})$, if

$$\lim_{k \to \infty} \int_\Omega f_k \varphi \, dx = \int_\Omega f \varphi \, dx, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^N). \tag{A10}$$

## Appendix B: Sobolev Spaces with Variable Exponent

We recall here well-known facts concerning the Sobolev spaces with variable exponent. Let $p(\cdot)$ be a measurable exponent function on $\Omega$ such that $1 < \alpha \le p(x) \le \beta < \infty$ a.e. in $\Omega$, where $\alpha$ and $\beta$ are given constants. We associate with it the so-called Sobolev–Orlicz space

$$W^{1,p(\cdot)}(\Omega)$$
$$:= \left\{ u \in W^{1,1}(\Omega) : \int_\Omega \left[ |u(x)|^{p(x)} + |\nabla u(x)|^{p(x)} \right] dx < \infty \right\} \tag{B11}$$

and equip it with the norm $\|u\|_{W_0^{1,p(\cdot)}(\Omega)} = \|u\|_{L^{p(\cdot)}(\Omega)} + \|\nabla u\|_{L^{p(\cdot)}(\Omega;\mathbb{R}^N)}$.

It is well known that in general, unlike classical Sobolev spaces, smooth functions are not necessarily dense in $W = W_0^{1,p(\cdot)}(\Omega)$. Hence, with variable exponent $p = p(x)$ ($1 < \alpha \le p \le \beta$), we can associate another Sobolev space,

$H = H^{1,p(\cdot)}(\Omega)$ as the closure of the set
$C^\infty(\overline{\Omega})$ in $W^{1,p(\cdot)}(\Omega)$-norm.

Since the identity $W = H$ is not always valid, it makes sense to say that an exponent $p(x)$ is regular if $C^\infty(\overline{\Omega})$ is dense in $W^{1,p(\cdot)}(\Omega)$.

The following result reveals the important property that guarantees the regularity of exponent $p(x)$ [55].

**Proposition 8** *Assume that there exists* $\delta \in (0, 1]$ *such that* $p \in C^{0,\delta}(\overline{\Omega})$. *Then, the set* $C^\infty(\overline{\Omega})$ *is dense in* $W^{1,p(\cdot)}(\Omega)$, *and, therefore,* $W = H$.

## References

1. Afraites, L., Hadri, A., Laghrib, A., Nachaoui, M.: A non-convex denoising model for impulse and Gaussian noise mixture removing using bi-level parameter identification. Inverse Probl. Imaging (2022). https://doi.org/10.3934/ipi.2022001
2. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford University Press, New York (2000)
3. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and $BV$ Spaces: Applications to PDEs and Optimization. SIAM, Philadelphia (2006)
4. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations, Series: Applied Mathematical Sciences, Springer, New York, p. 147 (2006)
5. Blomgren, P., Chan, T.F., Mulet, P., Wong, C.: Total variation image restoration: numerical methods and extensions. In: Proceedings of the 1997 IEEE International Conference on Image Processing, III, pp. 384–387 (1997)
6. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM J. Imaging Sci. **3**(3), 492–526 (2010)
7. Benning, M., Burger, M.: Error estimates for general fidelities. Electron. Trans. Numer. Anal. **38**, 44–68 (2011)
8. Brinkmann, E.-M., Burger, M., Grah, J.S.: Unified models for second-order tv-type regularisation in imaging: a new perspective based on vector operators. J. Math. Imaging Vis. **61**(5), 571–601 (2019)
9. Bungert, L., Coomes, D.A., Ehrhardt, M.J., Rasch, J., Reisenhofer, R., Schönlieb, C.D.: Blind image fusion for hyperspectral imaging with the directional total variation. Inverse Prob. **34**(4), 044003 (2018)
10. Bungert, L., Ehrhardt, M.J.: Robust image reconstruction with misaligned structural information. IEEE Access **8**, 222944–222955 (2020)
11. Burger, M., Papafitsoros, K., Papoutsellis, E., Schönlieb, C.-B.: Infimal convolution regularisation functionals of $BV$ and $L^p$ spaces. J. Math. Imaging Vis. **55**(3), 343–369 (2016)
12. Calatroni, L., Lanza, A., Pragliola, M., Sgallari, F.: A flexible space-variant anisotropic regularization for image restoration with automated parameter selection. SIAM J. Imaging Sci. (2019). https://doi.org/10.1137/18M1227937
13. Calatroni, L., De Los Reyes, J.C., Schönlieb, C.-B.: Infimal convolution of data discrepancies for mixed noise removal. SIAM J. Imaging Sci. **10**(3), 1196–1233 (2017)
14. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**(1–2), 89–97 (2004)
15. Chambolle, A., Lions, P.-L.: Image recovery via total variation minimization and related problems. Numer. Math. **76**, 167–188 (1997)
16. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numer. **25**, 161–319 (2016)
17. Chan, T.F., Marquina, A., Mulet, P.: High-order total variation-based image restoration. SIAM J. Sci. Comput. **22**(2), 503–516 (2000)
18. Chen, Y., Levine, S., Rao, M.: Variable exponent, linear growth functionals in image restoration. SIAM J. Appl. Math. **66**(4), 1383–1406 (2006)

19. Chen, Y., Levine, S., Stanich, J.: Image Restoration via Nonstandard Diffusion, Duquesne University, Technical Report, p. 13 (2004)

20. Cruz-Uribe, D.V., Fiorenza, A.: Variable Lebesgue Spaces: Foundations and Harmonic Analysis. Birkhäuser, New York (2013)

21. D'Apice, C., De Maio, U., Kogut, P.I.: Thermistor problem: multi-dimensional modelling, optimization, and approximation. In: Proceedings of the 32nd European Conference on Modelling and Simulation, May 22nd–25th, Wilhelmshaven Germany (2018)

22. D'Apice, C., De Maio, U., Kogut, P.I.: An indirect approach to the existence of quasi-optimal controls in coefficients for multi-dimensional thermistor problem. In: Sadovnichiy, V.A., Zgurovsky, M. (eds.), Contemporary Approaches and Methods in Fundamental Mathematics and Mechanics, Chapter 24, Springer, pp. 489–522 (2020)

23. D'Apice, C., Kogut, P.I., Manzo, R., Uvarov, M.: Variational model with nonstandard growth conditions for restoration of satellite optical images using synthetic aperture radar. Eur. J. Appl. Math. (2022). https://doi.org/10.1017/S0956792522000031

24. Diening, L., Harjulehto, P., Hästö, M., Růžiĉk, P.: Lebesgue and Sobolev Spaces with Variable Exponents, Springer, New York (2011)

25. Fan, Y.-R., Buccini, A., Donatelli, M., Huang, T.-Z.: A non-convex regularization approach for compressive sensing. Adv. Comput. Math. **45**, 563–588 (2019)

26. Gao, Y., Bredies, K.: Infimal convolution of oscillation total generalized variation for the recovery of images with structured texture. SIAM J. Image Sci. (2018). https://doi.org/10.1137/17M1153960

27. Gilles, J., Osher, S.: Bregman implementation of Meyer's G-norm for cartoon+textures decomposition. In: UCLA Cam, pp. 11–73 (2011)

28. Goyal, B., Dogra, A., Agrawal, S., Sohi, B.S., Sharma, A.: Image denoising review: from classical to state-of-the-art approaches. Inf. Fusion **55**, 220–244 (2020)

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

30. Kogut, P.I.: On optimal and quasi-optimal controls in coefficients for multi-dimensional thermistor problem with mixed Dirichlet–Neumann boundary conditions. Control. Cybern. **48**(1), 31–68 (2019)

31. Kogut, P.I.: On approximation of an optimal boundary control problem for linear elliptic equation with unbounded coefficients. Discrete Continu. Dyn. Syst. Ser. A **34**(5), 2105–2133 (2014)

32. Kogut, P.I., Leugering, L.: Optimal Control Problems for Partial Differential Equations on Reticulated Domains. Approximation and Asymptotic Analysis, Series: Systems and Control, Birkhäuser, Boston (2011)

33. Kogut, P.I., Leugering, L.: On S-homogenization of an optimal control problem with control and state constraints. Zeitschrift fur Analysis und ihre Anwendung **20**(2), 395–429 (2001)

34. Kogut, P.I., Kupenko, O.P.: Approximation Methods in Optimization of Nonlinear Systems, De Gruyter Series in Nonlinear Analysis and Applications, vol. 32. Walter de Gruyter GmbH, Berlin (2019)

35. Kogut, P.I., Kupenko, O.P., Manzo, R.: On regularization of an optimal control problem for ill-posed nonlinear elliptic equations. In: Abstract and Applied Analysis, Article ID 7418707, 15 p (2020)

36. Kogut, P.I., Manzo, R.: On vector-valued approximation of state constrained optimal control problems for nonlinear hyperbolic conservation laws. J. Dyn. Control Syst. **19**(3), 381–404 (2013)

37. Kogut, P., Kohut, Y., Manzo, R.: Fictitious controls and approximation of an optimal control problem for Perona–Malik equation. J. Optim. Differ. Equ. Appl. (JODEA) **30**(1), 42–70 (2022)

38. Kogut, P.I., Manzo, R.: Some results on optimal control problem for ill-posed nonlinear elliptic equations. In: IP Conference Proceedings of ICNAAM 2021 (the 19th International Conference of Numerical Analysis and Applied Mathematics, 20th–26th September Rodi Grecia (2021)

39. Kohr, H.: Total variation regularization with variable Lebesgue prior. arXiv:1702.08807 [math.NA] (2017)

40. Kupenko, O.P., Manzo, R.: Approximation of an optimal control problem in the coefficient for variational inequality with anisotropic p-Laplacian. Nonlinear Differ. Equ. Appl. **23**(3), 18 (2016)

41. Lazzaretti, M., Calatroni, L., Estatico, C.: Modular-proximal gradient algorithms in variable exponent Lebesgue spaces. Preprint HAL Id: hal-03476299 (2021)

42. Li, F., Li, Z., Pi, L.: Variable exponent functionals in image restoration. Appl. Math. Comput. **216**, 870–882 (2010)

43. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)

44. Meyer, Y.: Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. American Mathematical Society, Lewis Memorial Lectures (2001)

45. Nikolova, M.: A variational approach to remove outliers and impulse noise. J. Math. Imaging and Vis. **20**, 99–120 (2004)

46. Papafitsoros, K., Schönlieb, C.-B.: A combined first and second order variational approach for image reconstruction. J. Math. Imaging Vis. **48**(2), 308–338 (2014)

47. Pragliola, M., Calatroni, L., Lanza, A., Sgallari, F.: On and beyond total variation regularization in imaging: the role of space variance. arXiv:2104.03650 [math.NA] (2021)

48. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica **60**, 259–268 (1992)

49. Strong, D.M., Chan, T.F.: Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. Technical Report, Univeristy of California, Los Angeles, CA, CAM96-46 (1996)

50. Valkonen, T., Bredies, K., Knoll, F.: Total generalized variation in diffusion tensor imaging. SIAM J. Optim. **13**(3), 487–525 (2018)

51. Werner, F., Hohage, T.: Convergence rates in expectation for Tikhonov-type regularization of inverse problems with Poisson data. Inverse Probl. **28**(10), 1–14 (2012)

52. Zaslavski, A.J.: Optimization on Metric and Normed Spaces. Series: Springer Optimization and Its Applications, Springer, New York, p. 44 (2010)

53. Zhang, X., Bai, M., Ng, M.K.: Nonconvex-TV based image restoration with impulse noise removal. SIAM J. Imaging Sci. **10**, 1627–1667 (2017)

54. Zhikov, V.V.: Solvability of the three-dimensional thermistor problem. Proc. Steklov Inst. Math. **281**, 98–111 (2008)

55. Zhikov, V.V.: On variational problems and nonlinear elliptic equations with nonstandard growth conditions. J. Math. Sci. **173**(5), 463–570 (2011)

**Ciro D'Apice** obtained the degree cum laude in Mathematics in 1991 and the Ph.D in Mathematics in 1997 at University of Naples "Federico II." He is Full Professor in Mathematical Analysis at Dipartimento di Scienze Aziendali-Management e Innovation Systems (DISA-MIS). His research interests include: variational calculus, homogenization and optimal control; complex networks modeling: conservation laws and applications to traffic networks, telecommunication ones, and supply chains, queueing systems and networks; analytical aspects for the temporal and spatial behavior of solutions of dynamic problems.

**Peter Kogut** received his Ph.D in 1989 and his habilitation in Mathematics (Calculus of Variations and Optimal Control Theory) in 1998 from Glushkov Institute of Cybernetics (the National Academy of Science of Ukraine). He is currently the Full Professor and Head of Department of Differential Equations of Oles Honchar Dnipro National University, Ukraine. Before he held C4-professor positions at the Dnipro Technical State University and Vinnitsa Technical State University. During the last ten years, he was invited as visiting professor by Friedrich-Alexander-Universität Erlangen-Nürnberg (Germany), Basque Center of Applied Mathematics., Bilbao (Spain), Conservatoire National des Arts et Metiers (CNAM, Paris), and Salerno University (Italy). He was partially employed as a team leader of Tokyo Institute of Super Compression Technology (2000–2004), as Principle Data Scientist (Level C4) of Oracle Corporation (2010–2017), and Data Scientist of EOS Data Analytics (from 2018 up to now). His research interests include different areas of optimization theory, optimal control problems for partial differential equations, asymptotic analysis, optimization in partially ordered spaces, homogenization theory, inverse problems, mathematical image processing, and applications in various fields.

**Olha P. Kupenko** received her Ph.D. (System Analysis and Optimal decisions Theory) in 2011 from Space Research Institute and her habilitation in Mathematics (Theoretical Basics of Cybernetics) from Oles Honchar Dnipro National University in 2017. From 2010 to 2018, she was Associate Professor of System Analysis and Control department at Dnipro University of Technology "Dnipro Polytechnics" and since 2016 she is a C4-professor of Mathematics at the same department. She is partially employed as a senior researcher of Institute for Applied System Analysis of National Technical University of Ukraine "Igor Siko-rsky Kyiv Polytechnic Institute" (from 2018 up to now). Her research interests include different areas of optimization theory, optimal control problems for partial differential equations, asymptotic analysis, and image processing.

**Rosanna Manzo** received degree cum laude in Mathematics in 1996 and Ph.D in Information Engineering in 2007 at University of Salerno. She is Associate Professor in Mathematical Analysis at the Department of Information and Electrical Engineering and Applied Mathematics (DIEM), University of Salerno. The research activity focuses on the following topics: fluid dynamic models for road networks, telecommunication networks, supply chains and blood flows; optimal control for hybrid systems; variational analysis and optimization in normed spaces; queueing systems and networks.