# A CNC approach for Directional Total Variation

Gabriele Scrivanti, Émilie Chouzenoux, Jean-Christophe Pesquet

HAL Id: hal-03692840

https://hal.science/hal-03692840v2

Submitted on 29 Jun 2022

# A CNC approach for Directional Total Variation

Gabriele Scrivanti, Émilie Chouzenoux, Jean-Christophe Pesquet

*Université Paris-Saclay, Inria, CentraleSupélec, CVN, Gif-sur-Yvette, France.* (firstname.lastname@centralesupelec.fr)

*Abstract*—The core of many approaches for the resolution of variational inverse problems arising in signal and image processing consists of promoting the sought solution to have a sparse representation in a well-suited space. A crucial task in this context is the choice of a good sparsity prior that can ensure a good trade-off between the quality of the solution and the resulting computational cost. The recently introduced Convex-Non-Convex (CNC) strategy appears as a great compromise, as it combines the high qualitative performance of non-convex sparsity-promoting functions with the convenience of dealing with convex optimization problems. This work proposes a new variational formulation to implement CNC approach in the context of image denoising. By suitably exploiting duality properties, our formulation allows to encompass sophisticated directional total variation (DTV) priors. We additionally propose an efficient optimisation strategy for the resulting convex minimisation problem. We illustrate on numerical examples the good performance of the resulting CNC-DTV method, when compared to the standard convex total variation denoiser.

*Index Terms*—Directional Total Variation, Denoising, Non-convex and non-smooth regularisation, Primal-Dual Algorithm

## I. INTRODUCTION

A wide class of problems arising in image and signal processing can be described by sparsity-regularised variational models. The most natural sparsity inducing penalty is the $\ell_0$-pseudo-norm, but it induces the related problem to be NP hard and non-convex. A popular convex variational surrogate is the $\ell_1$-norm, though it has the drawback of under-estimating the high amplitude components of the considered signal. Non-convex variational regularisers manage to overcome this issue, at the cost of possibly introducing suboptimal local minima in the objective function. An efficient solution to keep only the best traits of these regularisers is represented by Convex-Non-Convex (CNC) strategies [1]–[5]. They consist of building convex objective functionals that include non-convex regularisation terms. Suitable CNC strategies have been designed for more and more general classes of problems in different fields of data processing (see [4] and the references therein). For the well-known Total Variation (TV) regularisation model [6], a CNC modification has shown to get around the notorious problems of boundary reduction and staircasing effect for image restoration [1], [4], [5] and image segmentation [7].

In this work we propose a new variational formulation extending the CNC strategy to a popular TV-based model, namely the Directional TV (DTV). DTV was firstly introduced in [8] for the restoration of images whose structures or textures

follow a single dominant direction. The idea is to provide the model with a suitable directional prior information in order to allow the regularisation to operate more efficiently. The model was then extended in [9] to handle multiple dominant directions by means of a spatially varying directional information that locally coincides with the edge directions of the image. In [10] and [11] the authors analysed the possibility to extend the directional approach to the Total Generalised Variation (TGV) model from [12], which takes into account higher-order derivatives of the image and allows to promote piecewise-affine reconstructions, in particular. A possibly non-convex version of DTV was proposed in [13] by considering a space-variant exponent for the DTV regulariser that adapts to the local smoothness of the image.

Our contributions in this work are the following. We infer a formulation for DTV that allows us to incorporate this regulariser into a CNC denoising framework. We then define a numerical procedure to efficiently address the resulting optimisation problem. We finally provide experimental results that support the use of the proposed regularisation method.

The paper is organised as follows. In Section II, we introduce a general TV-based image denoising problem. We describe the construction of a CNC sparsity-promoting function, and provide a sufficient condition for the convexity of the overall problem. Section III is dedicated to the derivation of our proposed CNC-DTV regularisation approach and Section IV provides the description of a dedicated optimisation procedure to tackle it. Section V presents our numerical results showing the interest of the proposed approach.

## II. IMAGE DENOISING VIA CNC TOTAL VARIATION

### A. Notation

Throughout this paper we denote by $\langle \cdot \mid \cdot \rangle$ the scalar product over $\mathbb{R}^n$ and $\| \cdot \|_2$ the associated Euclidean norm. $\mathsf{I}_n$ states for the $n \times n$ identity matrix.

The spectral norm is defined as $|||L||| = \sup\{\|Lz\|_2 \mid z \in \mathbb{R}^n, \|z\|_2 \leq 1\}$. $\Gamma_0(\mathbb{R}^n)$ indicates the class of functions $f : \mathbb{R}^n \to (-\infty, +\infty]$ that are proper (i.e., with a nonempty domain), lower semicontinuous and convex. For a function $f \in \Gamma_0(\mathbb{R}^n)$, function $f^* : \mathbb{R}^n \to [-\infty, +\infty]$ represents its *convex conjugate* that is defined as

$$(\forall u \in \mathbb{R}^n) \quad f^*(u) = \sup_{x \in \mathbb{R}^n} \{\langle x, u \rangle - f(x)\}.$$

### B. Image denoising

We focus on solving an image denoising problem, *i.e.* on restoring a source image $\bar{x} \in \mathbb{R}^n$ from an acquired

measurement $o \in \mathbb{R}^n$ that is related to the sought image through

$$o = \bar{x} + e, \tag{1}$$

where $e \in \mathbb{R}^n$ is an additive noise, here assumed to be i.i.d. zero-mean Gaussian. Then, a simple denoising strategy consists of defining an estimate $\hat{x} \in \mathbb{R}^n$ of $\bar{x}$ by solving the penalized least squares problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ \left\{ \mathcal{J}(x) = \psi(x) + \frac{\lambda}{2} \|x - o\|_2^2 \right\}. \tag{2}$$

Hereabove, function $\psi : \mathbb{R}^n \to (-\infty, +\infty]$ is the regularization term, associated to the regularization factor $\lambda^{-1}$, with $\lambda > 0$.

### C. Total-variation based regularizers

When dealing with images, a common approach consists of choosing $\psi$ so as to sparsify the sought image in a transformed space obtained via a linear transformation. This amounts to defining $\psi = \Psi \circ D$ where $D \in \mathbb{R}^{m \times n}$ is a linear operator and $\Psi : \mathbb{R}^m \to (-\infty, +\infty]$ is a sparsity-promoting term. Among possible choices for $\Psi$ and $D$, the total variation (TV) model [6] is probably the most celebrated one. TV regularizer promotes sparsity in the space of the vertical/horizontal gradients of the image, thus allowing piece-wise constant functions in the solution space. Matrix $D \in \mathbb{R}^{2n \times n}$ is set as the linear operator defined as $D = [D_{\mathrm{h}}^\top \ D_{\mathrm{v}}^\top]^\top$, where $(D_{\mathrm{h}}, D_{\mathrm{v}}) \in (\mathbb{R}^{n \times n})^2$ are the discrete horizontal and vertical 2D gradient operators obtained with a finite difference scheme. Then,

$$(\forall x \in \mathbb{R}^n) \quad \mathrm{TV}(x) = \|Dx\|_{1,2},$$
$$= \sum_{i=1}^n \|(Dx)_i\|_2. \tag{3}$$

Here, for every $u \in \mathbb{R}^{2n}$ and $i \in \{1, \ldots, n\}$, we use the compact notation: $\boldsymbol{u}_i = (u_i, u_{n+i}) \in \mathbb{R}^2$. We can also provide another definition of TV, based on duality [14]:

$$(\forall x \in \mathbb{R}^n) \quad \mathrm{TV}(x)$$
$$= \max_{u \in \mathbb{R}^{2n}} \{\langle Dx \mid u \rangle \mid \boldsymbol{u}_i \in B_2, \ i \in \{1, \ldots, n\}\}, \tag{4}$$

with $B_2 = \{v \in \mathbb{R}^2 \mid \|v\|_2 \leq 1\}$ the unit closed ball of $\mathbb{R}^2$.

One drawback of TV is that it gives an isotropic role to vertical and horizontal directions, that might not be well-adapted for natural images. In [8], a modified version of TV is proposed, that is more suitable for images containing objects with a (possibly non vertical/horizontal) dominant direction. The idea is to introduce an affine transformation in the dual space (i.e., the space of the image gradients). This transformation is parametrized by an expansion factor $\alpha \geq 1$ and a rotation angle $\theta \in [-\pi/2, \pi/2]$. We then define the transition matrices:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \Lambda_\alpha = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}.$$

These are used to build the elliptic set $E_{\alpha,\theta} \subset \mathbb{R}^2$ given by $E_{\alpha,\theta} = R_\theta \Lambda_\alpha B_2$, that is then substituted for $B_2$ in (4).

Actually, shapes present in images might have more than one directional orientation. In order to adapt the previous regularizer to each edge directions in the image, following [9], we can rely on a set of local ellipses $E_{\alpha_i,\theta_i} \subset \mathbb{R}^2$ for each pixel $i \in \{1, \ldots, n\}$. This leads to the so-called directional TV (DTV) regularizer, defined as

$$(\forall x \in \mathbb{R}^n) \ \mathrm{DTV}(x; \boldsymbol{\alpha}, \boldsymbol{\theta}) = \sum_{i=1}^n \max_{u \in E_{\alpha_i,\theta_i}} \langle (Dx)_i | u \rangle. \tag{5}$$

Here, $\boldsymbol{\alpha} = (\alpha_i)_{1 \leq i \leq n}$, $\boldsymbol{\theta} = (\theta_i)_{1 \leq i \leq n}$ are a predefined set of scaling factors/angles describing the edge direction information at each pixel of the image.

It is worth noticing that DTV is actually a generalized version of TV. Indeed, by choosing $\alpha_i = 1$ and $\theta_i = 0$ for every $i \in \{1, \ldots, n\}$, one retrieves the classic TV functional as expressed in (4), since $E_{1,0} = R_0 \Lambda_1 B_2 = B_2$.

DTV can be described in a more compact manner, by introducing the operator $\mathcal{A}_{\boldsymbol{\alpha},\boldsymbol{\theta}} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ that encodes the underlying local affine transformations. This linear operator is such that, for every $u \in \mathbb{R}^{2n}$, $\mathcal{A}_{\boldsymbol{\alpha},\boldsymbol{\theta}} u = (R_{\theta_i} \Lambda_{\alpha_i} \boldsymbol{u}_i)_{1 \leq i \leq n}$. We can also express its inverse, through $\mathcal{A}_{\boldsymbol{\alpha},\boldsymbol{\theta}}^{-1} u = (\Lambda_{1/\alpha_i} R_{-\theta_i} \boldsymbol{u}_i)_{1 \leq i \leq n}$ for every $u \in \mathbb{R}^{2n}$. Then, an equivalent definition of DTV is

$$(\forall x \in \mathbb{R}^n) \quad \mathrm{DTV}(x, \boldsymbol{\alpha}, \boldsymbol{\theta})$$
$$= \max_{u \in \mathbb{R}^{2n}} \{\langle Dx \mid u \rangle \mid (A_{\boldsymbol{\alpha},\boldsymbol{\theta}}^{-1} u)_i \in B_2, \ i \in \{1, \ldots, n\}\}. \tag{6}$$

The aforementioned (D)TV prior relies intrinsically on the convex sparsity measure $\ell_{1,2}$. However, as emphasized in [15], [16], non-convexity sparsity penalties might be key to obtain high quality results. The goal of this paper is to incorporate a CNC approach with the aim to obtain an enhanced image denoiser without complexifying the optimization procedure.

### III. PROPOSED CNC-DTV METHOD

### A. CNC approach

The idea of CNC, initially proposed in [4], [5], is to use the following construction for the regularization term $\psi = \Psi_M \circ D$ in (2), with $D$ the discrete gradient linear operator and

$$(\forall u \in \mathbb{R}^{2n}) \quad \Psi_M(u) = \varphi(u) - \underbrace{\inf_{t \in \mathbb{R}^{2n}} \left\{ \varphi(t) + \frac{1}{2} \|M(u-t)\|_2^2 \right\}}_{\varphi_M(u)}. \tag{7}$$

Function $\Psi_M$ is the non-convex modification of a convex sparsity promoting term $\varphi \in \Gamma_0(\mathbb{R}^{2n})$, parametrized by a matrix $M \in \mathbb{R}^{k \times 2n}$. It is obtained by subtracting from function $\varphi$ its so-called generalised Moreau envelope $\varphi_M : \mathbb{R}^{2n} \to \mathbb{R}$ depending on matrix $M$ [5, Definition 6]. This matrix plays a fundamental role, as it can be designed so as to guarantee both the efficiency of a non-convex regularisation approach and, at the same time, the overall convexity of the objective function, as we will discuss hereafter.

In [4], two main examples for the choice of function $\varphi$ are reported, which lead to different instances of CNC-TV

regularisations, namely $\varphi = \ell_1$ yields an anisotropic version of the TV while $\varphi$ equal to the Hessian Schatten norm [17] leads to a second-order extension of the TV regularizer. However, up to our knowledge, DTV (as defined in (6)) has not been explored in the context of CNC, and this is the aim of this work.

### B. Convexity condition

A key feature of CNC approach is that, despite the non-convexity of the introduced penalty $\Psi_M$, it is still possible to recast the minimization problem as a convex one. More precisely, in accordance with [18, Example 75], Problem (2) with $\psi = \Psi_M \circ D$ and $\Psi_M$ given in (7), can be reformulated as the following minimisation problem by introducing the dual variable $y \in \mathbb{R}^{2n}$:

$$\underset{(x,y)\in\mathbb{R}^n\times\mathbb{R}^{2n}}{\text{minimize}} \quad \tilde{\mathcal{J}}_M(x,y) = \frac{\lambda}{2}\|x-o\|_2^2 - \frac{1}{2}\|MDx\|_2^2 \quad (8)$$
$$+ \frac{1}{2}\|My\|_2^2 + \varphi(Dx) + \varphi^*(M^\top M(Dx-y)).$$

**Remark 1.** The equivalent problem in (8) is slightly different from the one proposed in [4, Eqs (88)-(89)] and [3, Eqs (52)-(53)], which relied on a saddle point formulation.

Since $\varphi$ is convex, so is its conjugate $\varphi^*$, as well as their compositions with linear mappings. It results that the only non-convex term in the objective function in (8) is the concave quadratic one $x \mapsto -\frac{1}{2}\|MDx\|_2^2$. In order to guarantee that $\tilde{\mathcal{J}}_M$ is convex, it is sufficient to ensure the convexity of the twice continuously differentiable function $x \mapsto \frac{\lambda}{2}\|x-o\|_2^2 - \frac{1}{2}\|MDx\|_2^2$ [1], [4], [5], that is equivalent to impose that the Hessian

$$H = \lambda \mathsf{I}_{2n} - D^\top M^\top M D \quad (9)$$

is a positive semidefinite matrix. In order to guarantee the existence of a unique minimizer, we further require $H$ to be positive definite and $M^\top M$ to be a full-rank matrix, thus making the objective function in (8) coercive.

### C. Proposed CNC Directional Total Variation

Let us now present our main contribution, that is a CNC formulation able to encompass the DTV regulariser (6) (itself being a generalisation of (4)).

First, by Fenchel-Rockafellar duality [19], from (6), we obtain an equivalent primal formulation for DTV, by noticing that the adjoint operator of $\mathcal{A}_{\alpha,\theta}^{-1}$ is given, for every $u \in \mathbb{R}^{2n}$, by $\mathcal{A}_{\alpha,\theta}^{-*}u = \left(R_{\theta_i}\Lambda_{1/\alpha_i}\boldsymbol{u}_i\right)_{1\leq i\leq n}$. This yields

$$(\forall x \in \mathbb{R}^n) \quad \text{DTV}(x,\boldsymbol{\alpha},\boldsymbol{\theta})$$
$$= \min_{v\in\mathbb{R}^{2n}} \{\|v\|_{1,2} \mid \mathcal{A}_{\alpha,\theta}^{-*}v = Dx\}$$
$$= (\mathcal{A}_{\alpha,\theta}^{-*} \rhd \|\cdot\|_{1,2})(Dx), \quad (10)$$

where $(\mathcal{A}_{\alpha,\theta}^{-*} \rhd g)$ denotes the *Exact Infimal Postcomposition* of a function $g$ by $\mathcal{A}_{\alpha,\theta}^{-*}$ [19, Chapter 12]. Then the proposed

CNC modification of DTV parametrised by some matrix $M \in \mathbb{R}^{k\times 2n}$ reads

$$(\forall x \in \mathbb{R}^n) \quad \text{CNC-DTV}(x)$$
$$= \varphi_{\mathcal{A}}(Dx) - \inf_{t\in\mathbb{R}^{2n}}\left\{\varphi_{\mathcal{A}}(t) + \frac{1}{2}\|M(Dx-t)\|_2^2\right\},$$

where $\varphi_{\mathcal{A}} = \mathcal{A}_{\alpha,\theta}^{-*} \rhd \|\cdot\|_{1,2}$. According to properties of the infimal postcomposition by a bounded operator [19, Proposition 13.24(iv)], the conjugate of $\varphi_{\mathcal{A}}$ is expressed as

$$(\forall y \in \mathbb{R}^{2n}) \quad \varphi_{\mathcal{A}}^*(y) = \iota_{B_{\infty,2}}(\mathcal{A}_{\alpha,\theta}^{-1}y), \quad (11)$$

where $B_{\infty,2}$ is the $\ell_{\infty,2}$ ball with center 0 and radius 1.

Hence, choosing $\varphi = \varphi_{\mathcal{A}}$ in (8) and introducing a constraint on the range of $x$ lead to the following optimisation problem where we set $\mathcal{H} = \mathbb{R}^n \times \mathbb{R}^{2n} \times \mathbb{R}^{2n}$:

$$\underset{(x,y,v)\in\mathcal{H}}{\text{minimize}} \quad F(x,y,v) + \|v\|_{1,2} + \iota_S(x,y,v)$$
$$+ \iota_E(Dx, \mathcal{A}_{\alpha,\theta}^{-*}v) + \iota_{B_{\infty,2}}(\mathcal{A}_{\alpha,\theta}^{-1}M^\top M(Dx-y)).$$
$$(12)$$

Here, $S = [0,1]^n \times \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ is the additional constraint set, $E$ is the vector space $\{(w,z)\in\mathbb{R}^{2n}\times\mathbb{R}^{2n} \mid w=z\}$ and, for every $(x,y,v) \in \mathcal{H}$,

$$F(x,y,v) = \frac{\lambda}{2}\|x-o\|_2^2 - \frac{1}{2}\|MDx\|_2^2 + \frac{1}{2}\|My\|_2^2. \quad (13)$$

Since $\varphi_{\mathcal{A}}$ is convex, the positive semidefiniteness condition for $H$ in (9) can also be applied as a convexity guaranty for our new objective function in (12). In the next section, we investigate how to address Problem (12) by means of an optimisation scheme that exploits the structure of this objective function.

### IV. OPTIMISATION ALGORITHM

### A. Primal-dual splitting

The Primal-Dual (PD) method in [20]–[22] allows to efficiently deal with problems involving several Lipschitzian, proximable, and linear composite terms. In (12) we identify the following functions and linear operators: for every $(y,v)\in\mathbb{R}^{2n}\times\mathbb{R}^{2n}$,

$$h_1(v) = \|v\|_{1,2} \qquad L_1 = [0\ 0\ \mathsf{I}_{2n}]$$
$$h_2(v) = \iota_{B_{\infty,2}}(v) \qquad L_2 = [\mathcal{A}_{\alpha,\theta}^{-1}M^\top MD - \mathcal{A}_{\alpha,\theta}^{-1}M^\top M, 0]$$
$$h_3(y,v) = \iota_E(y,v) \qquad L_3 = \begin{bmatrix} D & 0 & 0 \\ 0 & 0 & \mathcal{A}_{\alpha,\theta}^{-*} \end{bmatrix},$$

so that (12) amounts to minimizing $F + \sum_{i=1}^3 h_i \circ L_i + \iota_S$.

The iterations defining a sequence $(z_\ell)_{\ell\in\mathbb{N}} = (x_\ell, y_\ell, v_\ell)_{\ell\in\mathbb{N}}$ which, starting from point $z_0 \in \mathcal{H}$, converges to a solution $z_\infty \in \mathcal{H}$ to (12), are presented in Algorithm 1, where $\delta > 0$ is a Lipschitz constant of function $F$. For $f \in \Gamma_0(\mathbb{R}^n)$, $\text{prox}_f$ corresponds to the *proximity operator of $f$*. When $f = \iota_S$, $\text{prox}_f$ reduces to $\text{proj}_S$, the projection onto $S$.

**Algorithm 1** Primal-Dual Algorithm to solve (12)

---

**Initialize** $z_0 \in \mathcal{H}$, $w_{1,0} \in \mathbb{R}^n$, $w_{2,0} \in \mathbb{R}^{2n}$, $w_{3,0} \in \mathbb{R}^{2n}$
**Set** $\tau > 0$ and $\sigma > 0$ s.t. $\frac{1}{\tau} - \sigma ||| \sum_{i=1}^3 L_i^\top L_i ||| \geq \frac{\delta}{2}$
**for** $\ell = 0, 1, \ldots$ **do**
$\quad z_{\ell+1} = \text{proj}_S(z_\ell - \tau \nabla F(z_\ell) - \tau \sum_{i=1}^3 L_i^\top w_{i,\ell})$
$\quad$ **for** $i = 1, 2, 3$ **do**
$\quad\quad w_{i,\ell+1} = \text{prox}_{\sigma h_i^*}(w_{i,\ell} + \sigma L_i(2z_{\ell+1} - z_\ell))$
$\quad$ **end**
**end**

---

Since $|||D|||^2 = 8$, and $|||\mathcal{A}_{\alpha,\theta}^{-1}||| \leq 1$ (resp. $|||\mathcal{A}_{\alpha,\theta}^{-*}||| \leq 1$) as $\mathcal{A}_{\alpha,\theta}^{-1}$ (resp. $\mathcal{A}_{\alpha,\theta}^{-*}$) is the combination of rotations and contractions along one axis, the norm of the involved linear operator can be upper bounded as follows:

$$||| L_1^\top L_1 + L_2^\top L_2 + L_3^\top L_3 ||| \leq 1 + (1+8)(|||M|||^4) + 8. \quad (14)$$

It is worthy to note that Problem (12) and Algorithm 1 represent a unified framework for convex / non-convex, classical / directional TV, since proper choices for matrix $M$ and operator $\mathcal{A}_{\alpha,\theta}$ allow us to model the four different instances of the denoising problem at hand. For $M = 0_{2n} \in \mathbb{R}^{2n \times 2n}$, *i.e.* the null operator, we retrieve the convex TV formulation, whereas the classical TV formulation is obtained by setting, for every $i \in \{1, \ldots, n\}$, $\alpha_i = 1$ and $\theta_i = 0$.

In the proposed framework, all the proximal computations are exact since the PD method allows us to decouple the functions that are defined by composing a convex function with a linear operator. This represents an advantage with respect to the Forward-Backward strategies proposed in [5, Algorithm 3] and in [4, Proposition 10], which involve nested optimisation procedures, i.e. a subroutine has to be used to compute the proximity operator of the classical TV functional. We dedicate the next subsection to the description of the proximity operators of the involved terms.

*B. Practical implementation*

The proximity operator of function $h_2^* = h_1 = \| \cdot \|_{1,2}$ can easily be inferred from the one of the $\ell_2$ norm by applying the rule for a separable sum of terms:

$$(\forall u \in \mathbb{R}^{2n}) \quad \text{prox}_{\sigma h_2^*}(u) = \left( u_i - \frac{u_i}{\max(\frac{\|u_i\|_2}{\sigma}, 1)} \right)_{1 \leq i \leq n},$$

while the proximity operator of function $h_1^* = h_2$ is obtained by applying Moreau's identity: [19]

$$(\forall u \in \mathbb{R}^{2n}) \quad \text{prox}_{\sigma h_1^*}(u) = u - \sigma \text{prox}_{\frac{h_2}{\sigma}}\left(\frac{u}{\sigma}\right)$$
$$= \left( \frac{u_i}{\max(\|u_i\|_2, 1)} \right)_{1 \leq i \leq n}.$$



Fig. 1: Original images: `texture` (a), `barcode` (b), and `geometric` (c).

By applying again Moreau's formula, the proximity operator of $\sigma h_3^*$ is given by

$$(\forall (y,v) \in (\mathbb{R}^{2n})^2) \, \text{prox}_{\sigma h_3^*}(y,v) = (y,v) - \sigma \text{proj}_E \left( \frac{y}{\sigma}, \frac{v}{\sigma} \right)$$
$$= \frac{1}{2}\begin{bmatrix} y - v \\ v - y \end{bmatrix}.$$

## V. NUMERICAL RESULTS

We now evaluate the proposed approach for the restoration of noisy grayscale images. In accordance with [4], we choose matrix $M = \sqrt{\gamma}\mathsf{I}_{2n}$, so that $M^\top M = \gamma \mathsf{I}_{2n}$ and therefore the convexity condition reduces to ensuring the positive semidefiniteness of $H = \lambda \mathsf{I}_{2n} - \gamma D^\top D$. Since $|||D|||^2 = 8$, the convexity condition is satisfied as soon as $\gamma < \lambda/8$. We therefore set

$$\gamma = (\rho\lambda)/8,$$

with $\rho \in \{0, 0.99\}$. For $\rho = 0$, we retrieve the convex formulation of TV, whereas for $\rho = 0.99$ we get a high degree of non-convexity for $\psi_M$ while ensuring the convexity of the global problem.

We illustrate the performance of CNC-DTV on three synthetic images, `texture`, `barcode`, and `geometric`, shown in Figure 1. We added white zero-mean Gaussian noise with standard deviation $\sigma_e = 0.1$. In order to extract the directional information $\theta$ in DTV, we exploited the strategy proposed in [11, Section 5]. For the definition of parameter $\alpha$, we set $\alpha_i = \alpha$ for every $i$ as in [9]. We then chose the best combination of the two parameters $\alpha$ and $\lambda$ by means of a grid search to optimise the Peak Signal-to-Noise Ratio (PSNR) of the restored image.

In Table I, we report the best PSNR obtained when running our PD algorithm for the classic convex (i.e, $\rho = 0$) TV (C-TV), the non-convex (i.e, $\rho = 0.99$) TV (NC-TV), convex directional (C-DTV), and non-convex directional (NC-DTV) TV. This quantitative assessment shows that the combination of non-convex prior with a DTV-based space-variant regularisation yields an improvement w.r.t. the three other tested approaches.

| image | C-TV | CNC-TV | C-DTV ($1/\alpha$) | CNC-DTV ($1/\alpha$) |
|---|---|---|---|---|
| `texture` | 23.70 | 24.34 | 25.83 (0.1) | **26.76** (0.25) |
| `barcode` | 25.60 | 26.92 | 26.07 (0.4) | **27.69** (0.5) |
| `geometric` | 30.81 | 31.16 | 31.24 (0.45) | **32.30** (0.45) |

TABLE I: Best PSNR for $\sigma_e = 0.1$ optimised over a grid search for parameters $\lambda$ and $\alpha$.

We also provide in Figure 2 a visual illustration of the performance of the four approaches on `texture` image, showing the absolute error between the best estimated solution $\hat{x}$ and the original image $\bar{x}$. The reconstructions involving a non-convex regulariser or directional information show peculiar structures in the distribution of its absolute residual error, whereas the classic convex approach yields a rather dense error distribution. NC-DTV inherits the high coherence w.r.t the directions in the image from C-DTV and the high accuracy in noise removal and sharp transition reconstruction from NC-TV.
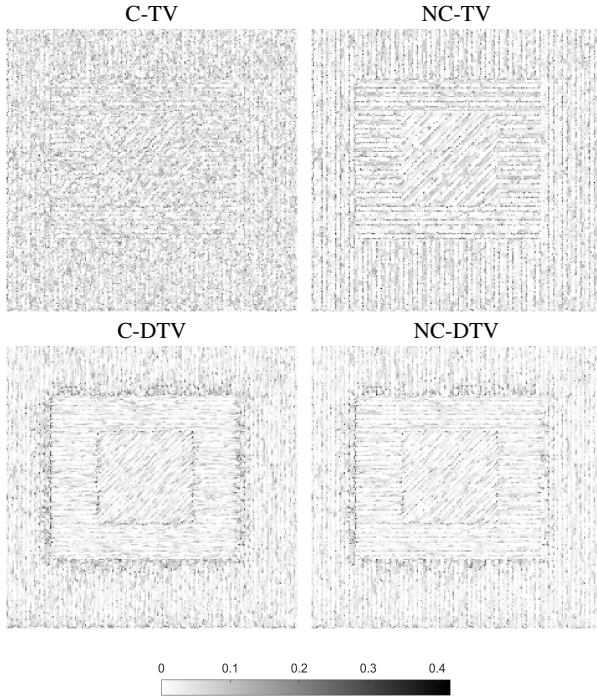


Fig. 2: Residual absolute error for `texture`.

Eventually, in Figure 3, we illustrate for image `geometric` the PSNR evolution along 3000 iterations (left) and the distance from the iterates $z_\ell$ to the solution $z_\infty$ in logarithmic scale (right) of the four approaches, which shows the fast and stable convergence behaviour of the proposed PD algorithm.
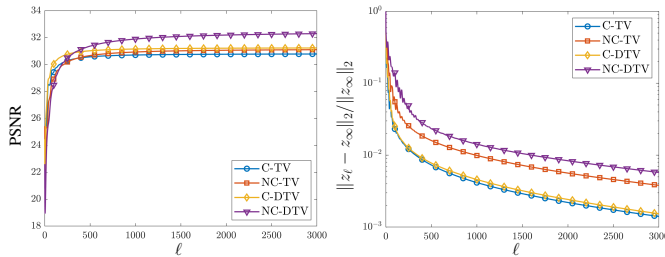


Fig. 3: PSNR versus iterations (left) and distance from the iterates to the solution versus iterations in logarithmic scale (right) for `geometric`.

## VI. CONCLUSIONS

In this work we investigated the extension of the CNC approach to a directional version of the TV regularisation model for image denoising. We proposed to address the resulting minimisation problem with a primal-dual procedure that efficiently exploits the structure of the objective function and we provided numerical results supporting the interest of the proposed approach.

## REFERENCES

[1] I. W. Selesnick, A. Parekh, and İ. Bayram, "Convex 1-d total variation denoising with non-convex regularization," *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 141–144, 2015.

[2] M. Nikolova, "Energy minimization methods," in *Handbook of Mathematical Methods in Imaging*. Springer New York, 2015, pp. 157–204.

[3] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4481–4494, 2017.

[4] A. Lanza, S. Morigi, I. W. Selesnick, and F. Sgallari, "Sparsity-inducing nonconvex nonseparable regularization for convex image processing," *SIAM Journal on Imaging Sciences*, vol. 12, no. 2, pp. 1099–1134, 2019.

[5] I. Selesnick, A. Lanza, S. Morigi, and F. Sgallari, "Non-convex total variation regularization for convex denoising of signals," *Journal of Mathematical Imaging and Vision*, vol. 62, 07 2020.

[6] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[7] R. H. Chan, A. Lanza, S. Morigi, and F. Sgallari, "Convex non-convex image segmentation," *Numerische Mathematik*, vol. 138, pp. 635–680, 2018.

[8] İ. Bayram and M. E. Kamasak, "Directional total variation," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 781–784, 2012.

[9] H. Zhang and Y. Wang, "Edge adaptive directional total variation," *The Journal of Engineering*, vol. 1, 11 2013.

[10] R. D. Kongskov, Y. Dong, and K. Knudsen, "Directional total generalized variation regularization," *BIT Numerical Mathematics*, pp. 1–26, 2019.

[11] S. Parisotto, S. Masnou, and C.-B. Schönlieb, "Higher-order total directional variation: Analysis," *SIAM Journal on Imaging Sciences*, vol. 13, pp. 474–496, 01 2020.

[12] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[13] L. Calatroni, A. Lanza, M. Pragliola, and F. Sgallari, "A flexible space-variant anisotropic regularization for image restoration with automated parameter selection," *SIAM Journal on Imaging Sciences*, vol. 12, no. 2, pp. 1001–1037, 2019.

[14] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 89–97, 2004.

[15] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Annals of Statistics*, vol. 38, no. 2, pp. 894–942, 04 2010.

[16] E. Soubies, L. Blanc-Féraud, and G. Aubert, "A continuous exact \ell_0 penalty (cel0) for least squares regularized problem," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1607–1639, 2015.

[17] S. Lefkimmiatis, J. P. Ward, and M. Unser, "Hessian schatten-norm regularization for linear inverse problems," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1873–1888, 2013.

[18] P. L. Combettes and J.-C. Pesquet, "Fixed point strategies in data science," *IEEE Transactions on Signal Processing*, vol. 69, p. 3878–3905, 2021.

[19] H. H. Bauschke and P. L. Combettes, "Convex analysis and monotone operator theory in hilbert spaces," in *CMS Books in Mathematics*, 2011.

[20] L. Condat, "A Primal–Dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of Optimization Theory and Applications*, vol. 158, 08 2013.

[21] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Advances in Computational Mathematics*, vol. 38, pp. 667–681, 2013.

[22] N. Komodakis and J.-C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp. 31–54, 2015.