# The Second Order Shannon Total Generalized Variation for Image Restoration

Alireza Hosseini[a,1], Sohrab Bazm[b]

[a] School of Mathematics, Statistics and Computer science, University of Tehran, P.O. Box 14115-175, Tehran, Iran. email: hosseini.alireza@ut.ac.ir

[b] Department of Mathematics, Faculty of Science, University of Maragheh, 55136-553 Maragheh, Iran

**Abstract**

In this paper, a Shannon-based second order TGV (Shannon TGV) is proposed. Shannon first and second order operators are defined and their corresponding adjoints are investigated leading to design a variational model as an optimization problem. We obtain the dual form of the proposed model and utilize it to formulate imaging problems, i.e., denoising and deconvolution. Moreover, we examine numerically the effectiveness of the proposed scheme in imaging problems and compare the results to the classic total variation (TV), the second order total generalized variation (TGV) and Shannon total variation. The outcomes confirm that the proposed model retains the advantages of both TGV and Shannon TV in elimination of artifacts simultaneously and admit the greater capability of the new model to remove artifacts.

*Keywords:* Total variation, total generalized variation, Shannon TV, Shannon TGV, image denoising, image deconvolution, primal-dual algorithm, staircase effects.

## 1. Introduction

The variational approach to image processing constitutes the computation of a reconstructed image $u$ based on the observed image (or more generally data) $x$ as a minimizer of a functional. Assume $x \in L^2(\Omega)$, where $\Omega \subset \mathbb{R}^2$ is a rectangular image domain and $A : L^2(\Omega) \to L^2(\Omega)$ is a bounded linear operator. A generic minimization problem for recovering $u$ from $x$ reads

$$\min_{u \in L^2(\Omega)} \lambda \mathcal{R}(u) + \frac{1}{2} \int_\Omega |Au(t) - x(t)|^2 dt. \tag{1}$$

Here, the so-called regularizer $\mathcal{R}$ corresponds to the a-priori information that we have on the reconstructed image $u$ and $\lambda > 0$ is the so-called regularizing parameter and acts like a balance between data model and the regularizer. One of the most popular variational approaches for image reconstruction is the total variation (TV) model, which takes $\mathcal{R}(u) = TV(u)$. Total variation model for image processing tasks introduced for image denoising and reconstruction in a celebrated paper of 1992 by Rudin, Osher and Fatemi [1]. Inspired by this approach, various kinds of first and high order variational models have been proposed. For first order contributions, we can refer to Condat's discrete TV [2], a model which uses oblique directions to define derivative [3] and an adaptive weighted TV model [4]. For high order models, we can point to total Laplace (TL) [5], bounded Hessian (BH) [6]-[9], TV and Laplacian model (TVL) [11], a model, which combines TV and BH (TVBH) [12], infimal-convolution (INFCON) [13] and a nonconvex, nonsmooth variational model [14]. In the another pioneer contribution, Bredies, Kunisch, and Pock, proposed the total generalized variation (TGV) functional [15]. The $k-$th order TGV with regularization vector $\alpha = (\alpha_0, \cdots, \alpha_{k-1}) \in (\mathbb{R}^+)^k$ is denoted by $\mathrm{TGV}^k_\alpha$. TGV is a generalization of TV, i.e., the first order TGV with the corresponding regularization parameter $\alpha_0 = 1$ coincides with TV. Meanwhile, to apply variational models for the real world image restoration, suitable discretization strategies are required. The simplest discretization scheme is defining discrete difference operators for continuous derivatives (this can be described in different ways). The discretization formulas for TV, second order TGV and all above mentioned high order approaches are discussed in the literature as well. In the computational aspects, the second order TGV is the most popular TGV model because of its simplicity and its distinguishing performance to diminish noises and artifacts. Practically, for image reconstruction problems, (e.g., denoising and deblurring), the higher order TGV models ($k > 1$) outperform TV in the sense of attenuating artifacts,

specially staircase effects. Total variation (as well as other first order models), however, also has some shortcomings, most notably the staircasing phenomenon. To briefly explain this effect, we assume that A=I, so that (1) describes the image denoising problem. If the true image contains not only flat, but also slanted regions, then mathematically, it is piecewise affine like function, then the image reconstructed by total variation model tends to be piecewise constant (staircasing). This phenomenon is because of nature of TV, whereas only first order derivatives is used in its structure and TV can restore only constant functions exactly, that is, if $x$ is constant signal, then $TV(x) = 0$. For instance, the second order TGV is exact for piecewise affine functions, that is, if $x$ is an affine (linear) function then $\mathrm{TGV}^2_\alpha(x) = 0$, therefore flat and stanted regions can be restored almost exactly using the second order TGV and in contrast to the usual total variation denoising model it does not exhibit the staircasing effect. Staircasing effect for TGV model is studied in [15]. It is worth mentioning, because of TGV's essence, the capability of the second (or third) order TGV, for removing unwanted effects is more visible in the plain (smooth) parts of the restored image versus the parts containing the details.

Recently, some efforts have been made to reduce disadvantages of TV directly through modification in discretization strategies, without attempting to increase the order of the derivatives. One of these schemes is using grid domain converter operators and reformulating the discrete dual definition of the discrete TV, along with involving the operators in the corresponding optimization problem's constraints (see [2] and [3] for instance). Another scheme is interpolation (refer to [16] and the references therein), where, instead of employing discrete TV formulation for a discrete image $u$, a suitable continuous interpolation of $u$ (let denote this image by $u^{\mathrm{int}}$) is obtained and the continuous definition of TV is applied for $u^{\mathrm{int}}$ to find $\mathrm{TV}(u^{\mathrm{int}})$. Generally computing $\mathrm{TV}(u^{\mathrm{int}})$ is impossible, so we are inevitable to discretize it, using mesh grids with number of grids preferably more than size of the initial discrete image $u$ (commonly, numbers of mesh grids in each row and column is twice (or third) as row and column sizes of $u$). In the present work, we get idea from interpolation approaches to give a new discretization for the second order TGV. Specifically, we consider Abergel and Moisan interpolation scheme (called Shannon TV) [16], which is a reconciliation of total variation with Shannon interpolation and is estimated by Fourier transforms. Shannon TV, behaves much better in terms of grid invariance, isotropy and artifact removal compared to the classic discrete TV model. However, Shannon TV is not able to defeat discrete second order TGV [15], in removing artifacts, specially staircase effects. On the other hand, Shannon TV, affect on all parts of the image, almost equally by diminishing the deep artifacts (where, still narrow artifacts remain), whereas, the performance of the second order TGV is case dependent (but generally, even narrow artifacts can be attenuated as well).

In this paper, a Shannon-based second order TGV (Shannon TGV) is proposed. In other words, similar to Shannon TV, which is a modification of discrete classic TV by means of Shannon interpolation, we design a model which is obtained from the discrete second order TGV, using Shannon interpolation approach. We compare the new model numerically with the classic discrete TV [1], Shannon TV [16] and the discrete second order TGV [15] for denoising and deconvolution problems. It is illustrated that the new model preserves the benefits of both Shannon TV and the second order discrete TGV simultaneously in removing artifacts and noises, with the higher quality.

The rest of the paper is organized as follows: Section 2 is a short review on some popular variational models, i.e., classic TV, Shannon TV and the second order TGV. In Section 3, the new first and second order Shannon operators are defined, and their adjoint operators are demonstrated. In Section 4, based on the introduced Shannon operators, the second order Shannon TGV is defined as an optimization problem and its dual formulation is given. In Section 5, we mathematically formulate denoising and deconvolution problems through the proposed model and investigate their dual formulation and propose two algorithms for denoising and deconvolution problems. In addition, we study the computational complexity of the model and compare it with the other models. In Section 6, numerical results of the proposed model is presented and compared with the other variational approaches. Finally, we give a conclusion in Section 7.

## 2. Some variational models and their discretization

In this section, some preliminaries about some popular variational models are expressed. First, we give the definition of the basic variational model, i.e., TV. Then, Shannon TV, as an interpolation-based model is defined and finally we give a brief explanation of the second order TGV.

In the sequel of the paper, let $I_M = \{0, 1, \cdots, M-1\}$, $\widehat{I_M} = [-\frac{M}{2}, \frac{M}{2}) \cap \mathbb{Z}$, where $M$ is a nonnegative integer.

Moreover, we denote the set of operators $u : I_P \times I_Q \to \mathbb{R}^k$ by $(\mathbb{R}^{P\times Q})^k$ and the set of operators $u : I_P \times I_Q \to \left\{\begin{pmatrix} a & b \\ b & c \end{pmatrix} : a, b, c \in \mathbb{R}\right\}$ by $S(\mathbb{R}^{P\times Q})^{2\times 2}$.

*2.1. Total variation*

Assume $u : \Omega \to \mathbb{R}^d$ and $u \in L^1_{loc}(\Omega)$, $\Omega \subset \mathbb{R}^d$, then

$$\mathrm{TV}(u) = \sup\left\{-\int_\Omega u \cdot \mathrm{div}\phi \, dX : \phi \in C^1_c(\Omega, \mathbb{R}^d), |\phi(X)| \le 1 \ \forall X \in \Omega\right\}. \tag{2}$$

If $u \in C^1(\Omega)$ (or $W^{1,1}(\Omega)$), then it can be proved that

$$\mathrm{TV}(u) = \int_\Omega |\nabla u| \, dX. \tag{3}$$

Moreover, discrete TV (called isotropic TV), for a discrete image $u : I_M \times I_N \to \mathbb{R}$ is defined by

$$\begin{array}{ll} \mathrm{TV}_i(u) = \sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2}\sqrt{((Du)_1(n_1,n_2))^2 + ((Du)_2(n_1,n_2))^2}, & \text{where} \\ (Du)_1(n_1,n_2) = u(n_1+1,n_2) - u(n_1,n_2), & n_1 \in I_M, n_2 \in I_N \\ (Du)_2(n_1,n_2) = u(n_1,n_2+1) - u(n_1,n_2), & u(M,n_2) = u(M-1,n_2), u(n_1,N) = u(n_1,N-1). \end{array} \tag{4}$$

The discrete TV is a simple model for implementation. However, for imaging problems, this model is unable to clear artifacts and clean up noises as well. One of the modified TV models which overcomes these shortcomings is Shannon TV approach which is explained in the following.

*2.2. Shannon total variation*

The Shannon TV is a modification of the discrete TV, which uses Shannon interpolation for discrete signal $u$ and then applies the direct definition of the continuous TV (3) over it. Finally, for the implementations, the model is discretized through the standard Riemann sum. To define the Shannon total variation, we need the following definitions:

**Definition 2.1.** [16] Assume $u : I_M \times I_N \to \mathbb{R}$, Shannon interpolation of $u$ is defined as the following:

$$U(x,y) = \frac{1}{MN}\sum_{-\frac{M}{2}\le\alpha\le\frac{M}{2},-\frac{N}{2}\le\beta\le\frac{N}{2}} \varepsilon_M(\alpha)\varepsilon_N(\beta).\hat{u}(\alpha,\beta)e^{2\pi\left(\frac{\alpha x}{M}+\frac{\beta y}{N}\right)}, \tag{5}$$

where

$$\varepsilon_M(\alpha) = \begin{cases} \frac{1}{2}, & |\alpha| = \frac{M}{2}, \\ 1, & \text{otherwise.} \end{cases} \tag{6}$$

From definition 2.1, we get

$$\nabla U(x,y) = \begin{pmatrix} \nabla_x U(x,y) \\ \nabla_y U(x,y) \end{pmatrix}, \ \nabla_x U(x,y) = 2\pi i\frac{\alpha}{M}U(x,y), \nabla_y U(x,y) = 2\pi i\frac{\beta}{N}U(x,y), \tag{7}$$

**Definition 2.2.** [16] Assume $n \ge 1$, the $n$ - Shannon gradient operator $\nabla_n : \mathbb{R}^{M\times N} \to (\mathbb{R}^{nM\times nN})^2$, is defined by

$$\begin{array}{l} \nabla_n u(k,l) = \begin{pmatrix} (\nabla_n)_1 u(k,l) \\ (\nabla_n)_2 u(k,l) \end{pmatrix}, u \in \mathbb{R}^{M\times N}, \\ (\nabla_n)_1 u(k,l) = \nabla_x U(\frac{k}{n},\frac{l}{n}), (\nabla_n)_2 u(k,l) = \nabla_y U(\frac{k}{n},\frac{l}{n}), k,l \in \Omega_n = I_{nM}\times I_{nN}. \end{array} \tag{8}$$

That is

$$\nabla_n u(k,l) = \frac{1}{MN}\sum_{-\frac{M}{2}\le\alpha\le\frac{M}{2},-\frac{N}{2}\le\beta\le\frac{N}{2}} e^{2\pi\left(\frac{\alpha k}{nM}+\frac{\beta l}{nN}\right)}g_{\hat{u}}(\alpha,\beta), \quad g_{\hat{u}}(\alpha,\beta) = 2\pi i\varepsilon_M(\alpha)\varepsilon_N(\beta)\hat{u}(\alpha,\beta)\begin{pmatrix} \frac{\alpha}{M} \\ \frac{\beta}{N} \end{pmatrix}. \tag{9}$$

Now, we are ready to introduce Shannon TV and its discrete version of the order $n$ (we call it $n$ - Shannon TV):

**Definition 2.3.** [16] Let $|\cdot|$ denotes the $l^2$ norm (in discrete or continuous domains), $\Omega = I_M \times I_N$, and $u \in \mathbb{R}^{M\times N}$, the Shannon total variation of $u$ is defined by

$$\mathrm{TV}_{\mathrm{SH}}(u) = \int_{[0,M]\times[0,N]} |\nabla U(x,y)| dxdy. \tag{10}$$

Moreover, for $n \geq 2$, the $n$ - Shannon total variation of $u$ is defined by

$$\mathrm{TV}_{\mathrm{SH(n)}}(u) = \frac{1}{n^2} \sum_{(k,l)\in\Omega_n} |\nabla_n u(k,l)|, \tag{11}$$

where, $\nabla_n$ and $\Omega_n$ are defined in definition 2.2.

The following lemme helps us to compute the discrete Fourier transform of the $n$ - Shannon gradient operator as well as its adjoint efficiently, which leads to a simple way to compute the $n$ - Shannon total variation.

**Lemma 2.4.** *[16] For any $n \geq 2$ and $(\alpha,\beta) \in \widehat{I_{nN}} \times \widehat{I_{nM}}$, we have*

$$\widehat{\nabla_n} u(\alpha,\beta) = \begin{cases} n^2 g_{\hat{u}}(\alpha,\beta), & |\alpha| \leq \frac{M}{2}, |\beta| \leq \frac{N}{2}, \\ 0, & otherwise, \end{cases} \tag{12}$$

*Moreover, assume $(\nabla_n)^* = -div_n$, then, for any $p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in (\mathbb{R}^{M\times N})^2$, any $n \geq 2, (\alpha,\beta) \in \widehat{I_M} \times \widehat{I_N}$,*

$$\widehat{div_n(p)}(\alpha,\beta) = 2\pi i \Big( \frac{\alpha}{M} k_1(p_1)(\alpha,\beta) + \frac{\beta}{N} k_2(p_2)(\alpha,\beta) \Big), \tag{13}$$

*where*

$$k_1(p_1)(\alpha,\beta) = \begin{cases} \hat{p_1}(\alpha,\beta), & |\alpha| < \frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p_1}(\alpha,\beta) - \hat{p_1}(-\alpha,\beta)), & \alpha = -\frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p_1}(\alpha,\beta) + \hat{p_1}(\alpha,-\beta)), & |\alpha| < \frac{M}{2}, \beta = -\frac{N}{2}, \\ \frac{1}{4}\sum_{s_1=\pm1, s_2=\pm1} \hat{p_1}(s_1\alpha, s_2\beta), & (\alpha,\beta) = (-\frac{M}{2}, -\frac{N}{2}), \end{cases} \tag{14}$$

$$k_2(p_2)(\alpha,\beta) = \begin{cases} \hat{p_2}(\alpha,\beta), & |\alpha| < \frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p_2}(\alpha,\beta) - \hat{p_2}(-\alpha,\beta)), & \alpha = -\frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p_2}(\alpha,\beta) + \hat{p_2}(\alpha,-\beta)), & |\alpha| < \frac{M}{2}, \beta = -\frac{N}{2}, \\ \frac{1}{4}\sum_{s_1=\pm1, s_2=\pm1} \hat{p_2}(s_1\alpha, s_2\beta), & (\alpha,\beta) = (-\frac{M}{2}, -\frac{N}{2}), \end{cases} \tag{15}$$

### *2.3. The second order TGV*

TGV is a celebrated variational model, because of its capability to remove staircase artifacts. Moreover, in the discrete imaging problems, TGV performance is far better than TV and even Shannon TV. It is worth to mention that TGV and the Shannon TV remove artifacts in different ways. That is, in the second order TGV model, a trade off between first and second derivatives leads to track the undesirable effects, whereas, in Shannon TV model, a trigonometric like (Shannon) interpolation leads to reduce the deep staircase effects. Therefore, introducing a model that preserves the advantages of both models would be valuable. In this paper, we are going to fulfill this aim. Now, we give a brief explanation of the second order TGV.
Let $u \in L^1_{loc}(\Omega), \Omega \subseteq \mathbb{R}^N, \alpha_0 > 0, \alpha_1 > 0$, then, the second order TGV of $u$ is defined by

$$\mathrm{TGV}^2_\alpha(u) = \sup \left\{ \int_\Omega u \cdot \mathrm{div}^2 v \, dX : v \in C^2_c(\Omega, \mathrm{Sym}^2(\mathbb{R}^d)), \|\mathrm{div}^l v\|_\infty \leq \alpha_l, l = 0, 1 \right\}. \tag{16}$$

In this paper, we focus on the two dimensional setting, i.e., $N = 2$. If $u \in C^2(\Omega)$, this definition can be simplified by

$$\mathrm{TGV}^2_\alpha(u) = \min_w \left\{ \alpha_1 \int_\Omega |\nabla u - w|| \, dX + \alpha_0 \int_\Omega |\epsilon(w)| \, dX \right\}, \epsilon(w) = \begin{pmatrix} \frac{\partial w_1}{\partial x} & \frac{\frac{\partial w_1}{\partial y} + \frac{\partial w_2}{\partial x}}{2} \\ \frac{\frac{\partial w_1}{\partial y} + \frac{\partial w_2}{\partial x}}{2} & \frac{\partial w_2}{\partial y} \end{pmatrix}. \tag{17}$$

$C_c^2(\Omega, \mathrm{Sym}^2(\mathbb{R}^2))$ is the set of symmetric $\mathbb{R}^{2\times 2}$ matrices whose components belong to $C_c^2(\Omega)$. To study the higher order TGV models, the properties, and the standard discretization scheme of the second order TGV, we refer the reader to [15].

## 3. Shannon TGV operators and their properties

Inspired by Shannon TV model, which is introduced in Section 2, we are going to give a new discretization for the second order TGV. To define the new model, we need to define the suitable derivative operators and their adjoints. The $n$ - Shannon gradient operator in Definition 2.2 is not enough for our target, so we define the new Shannon operators (lets call them Shannon TGV operators) and explain their adjoints.

### *3.1. Shannon first and second derivative operators*

Assume $u : I_M \times I_N \to \mathbb{R}$ is a two dimensional discrete gray scaled image. Let $U$ is the Shannon interpolation of $u$, then from Definition 2.1, we get

$$\nabla^2 U(x,y) = \begin{pmatrix} \nabla^2_{xx} & \nabla^2_{xy} \\ \nabla^2_{xy} & \nabla^2_{yy} \end{pmatrix}, \quad \begin{array}{c} \nabla^2_{xx} = (2\pi i)^2(\frac{\alpha^2}{M^2})U(x,y), \nabla^2_{xy} = (2\pi i)^2(\frac{\alpha\beta}{MN})U(x,y) \\ \nabla^2_{yy} = (2\pi i)^2(\frac{\beta^2}{N^2})U(x,y). \end{array} \tag{18}$$

**Definition 3.1.** (*Shannon TGV operators*) Assume $n \geq 2$,
($a$) The $n$- Shannon Hessian operator is defined by $\nabla^2_n : \mathbb{R}^{M\times N} \to S(\mathbb{R}^{nM\times nN})^{2\times 2}$, with

$$\nabla^2_n u(k,l) = \nabla^2 U(\frac{k}{n}, \frac{l}{n}), k, l \in \Omega_n = I_{nM} \times I_{nN}, \tag{19}$$

that is

$$\nabla^2_n u(k,l) = \frac{1}{MN}\sum_{-\frac{M}{2}\leq\alpha\leq\frac{M}{2}, -\frac{N}{2}\leq\beta\leq\frac{N}{2}} e^{2\pi\left(\frac{\alpha k}{nM}+\frac{\beta l}{nN}\right)} S_{\hat{u}}(\alpha,\beta), \quad S_{\hat{u}}(\alpha,\beta) = (2\pi i)^2 \varepsilon_M(\alpha)\varepsilon_N(\beta)\hat{u}(\alpha,\beta)\begin{pmatrix} \frac{\alpha^2}{M^2} & \frac{\alpha\beta}{MN} \\ \frac{\alpha\beta}{MN} & \frac{\beta^2}{N^2} \end{pmatrix}. \tag{20}$$

($b$) The $n$ - Shannon gradient operator $\mathcal{E}_n : (\mathbb{R}^{nM\times nN})^2 \to S(\mathbb{R}^{nM\times nN})^{2\times 2}$, for $u, v : I_{nM} \times I_{nN} \to \mathbb{R}$, is defined by

$$\mathcal{E}_n\begin{pmatrix} u \\ v \end{pmatrix}(k,l) = \begin{pmatrix} n(\nabla_1)_1 u(k,l) & \frac{n((\nabla_1)_1 v(k,l) + (\nabla_1)_2 u(k,l))}{2} \\ \frac{n((\nabla_1)_1 v(k,l) + (\nabla_1)_2 u(k,l))}{2} & n(\nabla_1)_2 v(k,l) \end{pmatrix}, k, l \in \Omega_n = I_{nM} \times I_{nN}, \tag{21}$$

where, operators $(\nabla_1)_1$ and $(\nabla_1)_2$ are defined in (8), for $n = 1$.

We are interested in obtaining the efficient ways to compute the required objects for our simulation aspects. From the definition of the discrete Fourier transform, we get the following simple way to compute the discrete Fourier transform of the $n$ - Shannon Hessian, for a given signal:

**Lemma 3.2.** *For any $n \geq 2$ and $(\alpha,\beta) \in \widehat{I_{nN}} \times \widehat{I_{nM}}$, we have*

$$\widehat{\nabla^2_n u}(\alpha,\beta) = \begin{cases} n^2 S_{\hat{u}}(\alpha,\beta), & |\alpha| \leq \frac{M}{2}, |\beta| \leq \frac{N}{2}, \\ 0, & \textit{otherwise}. \end{cases} \tag{22}$$

*Remark* 3.3. (12) and (22), hold for $n = 1$, if $\hat{u}(-\frac{M}{2},\beta) = \hat{u}(\alpha,-\frac{N}{2}) = 0$, for any $(\alpha.\beta) \in \widehat{I_M} \times \widehat{I_N}$.

It is worth to note that, in the definition of the Shannon operators for defining corresponding Shannon TGV model, we tried to use the definition of continuous operators, defined in TGV models (16) and (17), for the Shannon interpolation of the discrete signal $u$ (which is denoted by $U$ (5)). More precisely, $\nabla_n u \approx \nabla U, \mathcal{E}_n u \approx \epsilon U$ and $\nabla^2_n u \approx \nabla^2 U$. It is natural to expect that, relations between continuous operators in TGV definitions would be preserved for the new discrete operators. One of these relations for the continuous operators is $\epsilon\nabla = \nabla^2$. In the following theorem, we show that $\mathcal{E}_n\nabla_n = \nabla^2_n$. This, justifies the existence of the factor "$n$" as a coefficient in the elements of the defined $2 \times 2$ matrix in the definition of $\mathcal{E}_n$ (21).

**Theorem 3.4.** *Assume $u \in \mathbb{R}^{M\times N}$, then, for any $k,l \in \Omega_n = I_{nM} \times I_{nN}$, $\mathcal{E}_n(\nabla_n)u(k,l) = \nabla_n^2 u(k,l)$.*

*Proof.* From definition of $\mathcal{E}_n$, we get

$$\begin{aligned}\mathcal{E}_n(\nabla_n u)(k,l) &= \mathcal{E}_n\begin{pmatrix}(\nabla_n)_1 u\\ (\nabla_n)_2 u\end{pmatrix}(k,l) = \\ &= \begin{pmatrix} n(\nabla_1)_1((\nabla_n)_1 u)(k,l) & \frac{n((\nabla_1)_1((\nabla_n)_2 u)(k,l) + (\nabla_1)_2((\nabla_n)_1 u)(k,l))}{2}\\ \frac{n((\nabla_1)_1((\nabla_n)_2 u)(k,l) + (\nabla_1)_2((\nabla_n)_1 u)(k,l))}{2} & n(\nabla_1)_2((\nabla_n)_2 u)(k,l)\end{pmatrix}.\end{aligned} \tag{23}$$

It is enough to show that for any $k,l \in \Omega_n = I_{nM} \times I_{nN}$, (23) equals to (20). We prove this for the element $(1,2)$ (which is the same as $(2,1)$). The remaining elements have the similar procedures.

$$\begin{aligned} n((\nabla_1)_1((\nabla_n)_2 u)(k,l) &= n\frac{1}{(nM)(nN)} \sum_{\substack{-\frac{nM}{2}\le\alpha\le\frac{nM}{2},\\ -\frac{nN}{2}\le\beta\le\frac{nN}{2}}} e^{2\pi\left(\frac{\alpha k}{nM}+\frac{\beta l}{nN}\right)}(g_{\widehat{(\nabla_n)_2 u}})_1(\alpha,\beta),\\ (g_{\widehat{(\nabla_n)_2 u}})_1(\alpha,\beta) &= 2\pi i\varepsilon_{nM}(\alpha)\varepsilon_{nN}(\beta)\widehat{(\nabla_n)_2 u}(\alpha,\beta)(\frac{\alpha}{nM}).\end{aligned} \tag{24}$$

$\widehat{(\nabla_n)_2 u}$ given in (22), that is

$$\widehat{(\nabla_n)_2 u}(\alpha,\beta) = \begin{cases} 2n^2\pi i\varepsilon_M(\alpha)\varepsilon_N(\beta)\hat{u}(\alpha,\beta)(\frac{\beta}{N}), & -\frac{M}{2}\le\alpha\le\frac{M}{2}, -\frac{N}{2}\le\beta\le\frac{N}{2}\\ 0, & otherwise\end{cases}$$

Substituting in (24) and the fact that in (24), sentences outside the indexes $-\frac{M}{2}\le\alpha\le\frac{M}{2}, -\frac{N}{2}\le\beta\le\frac{N}{2}$ will be zero, we get $\varepsilon_{nM}(\alpha) = 1, \varepsilon_{nN}(\beta) = 1$, for the nonzero sentences. Consequently

$$\begin{aligned} n((\nabla_1)_1((\nabla_n)_2 u)(k,l) &= n\frac{1}{(nM)(nN)} \sum_{\substack{-\frac{M}{2}\le\alpha\le\frac{M}{2}\\ -\frac{N}{2}\le\beta\le\frac{N}{2}}} e^{2\pi\left(\frac{\alpha k}{nM}+\frac{\beta l}{nN}\right)}(g_{\widehat{(\nabla_n)_2 u}})_1(\alpha,\beta),\\ (g_{\widehat{(\nabla_n)_2 u}})_1(\alpha,\beta) &= -n^2 4\pi^2 i\varepsilon_M(\alpha)\varepsilon_N(\beta)\hat{u}(\alpha,\beta)(\frac{\beta}{N})(\frac{\alpha}{nM}).\end{aligned} \tag{25}$$

Comparing (25) with (20), we get

$$n((\nabla_1)_1((\nabla_n)_2 u)(k,l) = ((\nabla_n^2)u(k,l))(1,2) = ((\nabla_n^2)u(k,l))(2,1),$$

similarly

$$n((\nabla_1)_2((\nabla_n)_1 u)(k,l) = ((\nabla_n^2)u(k,l))(1,2) = ((\nabla_n^2)u(k,l))(2,1).$$

This proves that elements $(2,1)$ (as well as $(1,2)$) of $\mathcal{E}_n(\nabla_n)u$ and $\nabla_n^2 u$ are equal. □

### *3.2. Adjoint Operators*

In the sequel of this section, we propose an applicable way to compute the adjoints of the derivative operators, by means of discrete Fourier transform. The results are expressed in the following theorem:

**Theorem 3.5.** *Assume $(\nabla_n^2)^* = div_n^2$ and $(\mathcal{E}_n)^* = -div_n^{'}$, then*

1. *Let $p = \begin{pmatrix} p_1 & p_2\\ p_2 & p_3\end{pmatrix} \in S(\mathbb{R}^{nM\times nN})^{2\times 2}$, then for any $n \ge 2, (\alpha,\beta) \in \widehat{I_M} \times \widehat{I_N}$,*

$$\widehat{div_n^2(p)}(\alpha,\beta) = -4\pi^2\left(\frac{\alpha^2}{M^2}h_1(p_1)(\alpha,\beta) + \frac{2\alpha\beta}{MN}h_2(p_2)(\alpha,\beta) + \frac{\beta^2}{N^2}h_1(p_3)(\alpha,\beta)\right), \tag{26}$$

*where, for* $p_i, i = 1, 3$,

$$h_1(p_i)(\alpha,\beta) = \begin{cases} \hat{p}_i(\alpha,\beta), & |\alpha| < \frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p}_i(\alpha,\beta) + \hat{p}_i(-\alpha,\beta)), & \alpha = -\frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p}_i(\alpha,\beta) + \hat{p}_i(\alpha,-\beta)), & |\alpha| < \frac{M}{2}, \beta = -\frac{N}{2}, \\ \frac{1}{4}\sum_{s_1=\pm1, s_2=\pm1} \hat{p}_i(s_1\alpha, s_2\beta), & (\alpha,\beta) = (-\frac{M}{2}, -\frac{N}{2}), \end{cases} \tag{27}$$

*and*

$$h_2(p_2)(\alpha,\beta) = \begin{cases} \hat{p}_2(\alpha,\beta), & |\alpha| < \frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p}_2(\alpha,\beta) - \hat{p}_2(-\alpha,\beta)), & \alpha = -\frac{M}{2}, |\beta| < \frac{N}{2}, \\ \frac{1}{2}(\hat{p}_2(\alpha,\beta) - \hat{p}_2(\alpha,-\beta)), & |\alpha| < \frac{M}{2}, \beta = -\frac{N}{2}, \\ \frac{1}{4}\sum_{s_1=\pm1, s_2=\pm1} \hat{p}_2(s_1\alpha, s_2\beta), & (\alpha,\beta) = (-\frac{M}{2}, -\frac{N}{2}). \end{cases} \tag{28}$$

2. *Assume* $p = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_3 \end{pmatrix} \in S(\mathbb{R}^{nM\times nN})^{2\times2}$, $n \geq 1, (\alpha,\beta) \in \widehat{I_{nM}} \times \widehat{I_{nN}}$, *moreover,* $\hat{p}_i(-\frac{nM}{2},\beta) = \hat{p}_i(\alpha,-\frac{nN}{2}) = 0, i = 1, 2, 3$, *then*

$$\widehat{div_n'(p)}(\alpha,\beta) = \begin{pmatrix} w_1(\alpha,\beta) \\ w_2(\alpha,\beta), \end{pmatrix} \tag{29}$$

$$w_1(\alpha,\beta) = 2\pi i\left(\frac{\alpha}{M}k_1(p_1)(\alpha,\beta) + \frac{\beta}{N}k_2(p_2)(\alpha,\beta)\right), \tag{30}$$

$$w_2(\alpha,\beta) = 2\pi i\left(\frac{\alpha}{M}k_1(p_2)(\alpha,\beta) + \frac{\beta}{N}k_2(p_3)(\alpha,\beta)\right), \tag{31}$$

*where* $k_1, k_2$ *are defined in (14) and (15).*

*Proof.* For the proof of part 1, we should find $div_n^2$ such that for each $u \in \mathbb{R}^{M\times N}$, and $S(\mathbb{R}^{nM\times nN})^{2\times2}$,

$$\langle \nabla_n^2 u, p\rangle = \langle u, div_n^2 p\rangle, \tag{32}$$

on the other hand we know that for two signal of the same size, $x, y \in \mathbb{R}^M$, $\langle x, y\rangle = \frac{1}{M}\langle \hat{x}, \hat{y}\rangle$, thus it is enough to find $div_n^2 p$ such that $\langle \widehat{\nabla_n^2 u}, \hat{p}\rangle = n^2\langle \hat{u}, \widehat{div_n^2 p}\rangle$. This is equivalent to the following equation

$$\begin{aligned} \sum_{\substack{|\alpha| \leq \frac{nM}{2} \\ |\beta| \leq \frac{nN}{2}}} & \widehat{\nabla_n^2 u}(\alpha,\beta)(1,1)(\hat{p}_1(\alpha,\beta))^* + 2\widehat{\nabla_n^2 u}(\alpha,\beta)(1,2)(\hat{p}_2(\alpha,\beta))^* + \widehat{\nabla_n^2 u}(\alpha,\beta)(2,2)(\hat{p}_3(\alpha,\beta))^* \\ & = n^2 \sum_{\substack{-\frac{M}{2} < \alpha \leq \frac{M}{2} \\ -\frac{N}{2} < \beta \leq \frac{N}{2}}} (\widehat{div_n^2 p}(\alpha,\beta))^* \hat{u}(\alpha,\beta). \end{aligned} \tag{33}$$

Substituting $\widehat{\nabla_n^2 u}(\alpha,\beta)(i,j), i, j = 1, 2$, from (22)), we obtain

$$\begin{aligned} (2\pi i)^2 n^2 \sum_{\substack{|\alpha| \leq \frac{M}{2} \\ |\beta| \leq \frac{N}{2}}} & \varepsilon_M(\alpha)\varepsilon_N(\beta)\hat{u}(\alpha,\beta)\left\{\tfrac{\alpha^2}{M^2}(\hat{p}_1(\alpha,\beta))^* + 2\tfrac{\alpha\beta}{MN}(\hat{p}_2(\alpha,\beta))^* + \tfrac{\beta^2}{N^2}(\hat{p}_3(\alpha,\beta))^*\right\} \\ & = n^2 \sum_{\substack{-\frac{M}{2} < \alpha \leq \frac{M}{2} \\ -\frac{N}{2} < \beta \leq \frac{N}{2}}} (\widehat{div_n^2 p}(\alpha,\beta))^* \hat{u}(\alpha,\beta). \end{aligned} \tag{34}$$

From the definition of fast Fourier transform, $\hat{u}(\alpha + M, \beta + N) = \hat{u}(\alpha,\beta)$. the indexes $(\gamma,\eta)$ such that $\hat{u}(\alpha,\beta) = \hat{u}(\gamma,\eta)$ in the left side of equation (34) are listed in the second column of Table 1 and the corresponding $\varepsilon_M, \varepsilon_N$ coefficients are listed in the third column.

As a result, it can be seen that for a given $\alpha \in [-\frac{M}{2}, \frac{M}{2}), \beta \in [-\frac{M}{2}, \frac{M}{2})$, the coefficient of $\hat{u}(\alpha,\beta)$ in the right hand side

| $(\alpha,\beta)$ | The indexes $(\gamma,\eta)$, appeared in the left side of (34), where $\hat{u}(\alpha,\beta) = \hat{u}(\gamma,\eta)$ | The corresponding $\varepsilon_M(\gamma), \varepsilon_N(\eta)$ coefficients |
|---|---|---|
| $\lvert\alpha\rvert < \frac{M}{2}, \lvert\beta\rvert < \frac{N}{2}$ | $(\gamma,\eta) = (\alpha,\beta)$ | $\varepsilon_M(\gamma) = \varepsilon_N(\eta) = 1$ |
| $\alpha = -\frac{M}{2}, \lvert\beta\rvert < \frac{N}{2}$ | $(\gamma,\eta) = (\pm\frac{M}{2},\beta)$ | $\varepsilon_M(\gamma) = \frac{1}{2}, \varepsilon_N(\eta) = 1$ |
| $\lvert\alpha\rvert < \frac{M}{2}, \beta = -\frac{M}{2}$ | $(\gamma,\eta) = (\alpha,\pm\frac{N}{2})$ | $\varepsilon_M(\gamma) = 1, \varepsilon_N(\eta) = \frac{1}{2}$ |
| $\alpha = -\frac{M}{2}, \beta = -\frac{M}{2}$ | $(\gamma,\eta) = (\pm\frac{M}{2},\pm\frac{N}{2})$ | $\varepsilon_M(\gamma) = \varepsilon_N(\eta) = \frac{1}{2}$ |

Table 1: List of all indexes in the ranges of the summation in the left side of (34) are shown in first column, second column illustrate $(\gamma,\eta)$ indexes with the property $\hat{u}(\gamma,\eta) = \hat{u}(\alpha,\beta)$, and third column shows the corresponding $\varepsilon_M, \varepsilon_N$ values

of (34) is equal to $(\widehat{div_n^2 p}(\alpha,\beta))^*$ which can be obtained from the left hand side of (34), by the following calculation:

$$(\widehat{div_n^2 p}(\alpha,\beta))^* = -4\pi^2 \sum_{\substack{|\gamma| \le \frac{M}{2}, |\eta| \le \frac{N}{2}: \\ \hat{u}(\alpha,\beta) = \hat{u}(\gamma,\eta)}} \varepsilon_M(\gamma)\varepsilon_N(\eta)\hat{u}(\gamma,\eta)\left\{\frac{\alpha^2}{M^2}(\hat{p_1}(\gamma,\eta))^* + 2\frac{\gamma\eta}{MN}(\hat{p_2}(\gamma,\eta))^* + \frac{\gamma^2}{N^2}(\hat{p_3}(\gamma,\eta))^*\right\}. \tag{35}$$

Observing Table 1 and taking complex adjoint from both sides, (26) will be obtained. By referring to Remark 3.3, part 2, can be proved in the similar way and the proof of the theorem is completed. □

Suppose $X$ and $Y$ are finite-dimensional real vector spaces, $A : X \to Y$ is a continuous linear operator and $F : X \to [-\infty, +\infty]$ is a convex, proper and lower semicontinuous function. Hereafter, we assume $A^* : Y^* \to X^*$ is adjoint of $A$, $F^*$, is Fenchel - Legendre conjugate of $F$ and $\text{prox}_{\delta F}$ is proximal operator associated to function $F$ (see more details about convex analysis in [17]).

## 4. Shannon TGV

Based on the defined first and second order Shannon TGV operators, we are ready to define the second order Shannon TGV and derive its dual formulation.

**Definition 4.1.** Assume $u \in \mathbb{R}^{M\times N}$, then, for a given $\alpha = (\alpha_0, \alpha_1) \in \mathbb{R}^2$, $n$ - Shannon second order TGV of $u$ with weight vector $\alpha$, is defined by

$$\text{TGV}^{2(\alpha)}_{\text{SH}(n)}(u) = \max_v \left\{\frac{1}{n^2}\langle u, div_n^2 v\rangle : v \in S(\mathbb{R}^{nM\times nN})^{2\times 2}, \|v\|_{2,\infty} \le \alpha_0, \|div_n' v\|_{2,\infty} \le \alpha_1\right\}, \tag{36}$$

where for $v = \begin{pmatrix} v_1 & v_2 \\ v_2 & v_3 \end{pmatrix}, v_i \in \mathbb{R}^{P\times Q}, i = 1,2,3,$

$$\|v\|_{2,\infty} = \max\left\{\sqrt{v_1^2(i,j) + 2v_2^2(i,j) + v_3^2(i,j)}, (i,j) \in I_P \times I_Q\right\},$$

and for $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, w_i \in \mathbb{R}^{P\times Q}, i = 1, 2,$

$$\|w\|_{2,\infty} = \max\left\{ \sqrt{w_1^2(i,j) + w_2^2(i,j)}, (i,j) \in I_P \times I_Q \right\}.$$

Let

$$K = \left\{(v, w, s) \in S(\mathbb{R}^{nM\times nN})^{2\times 2} \times (\mathbb{R}^2)^{nM\times nN} \times \mathbb{R}^{M\times N} : \|v\|_{2,\infty} \le \alpha_0, \|w\|_{2,\infty} \le \alpha_1, w = div_n^{'} v, s = div_n^2 v\right\}, \tag{37}$$

then for $u \in \mathbb{R}^{M\times N}$

$$\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}(u) = \max_{(v,w,s)} \left\{ \frac{1}{n^2}\langle u, s\rangle - I_k\{(v, w, s)\} \right\}, \tag{38}$$

where $I_K$ is indicator function which is defined by $I_K(p) = \begin{cases} 0, & p \in K, \\ +\infty, & p \notin K. \end{cases}$.
Define linear operator $L : S(\mathbb{R}^{nM\times nN})^{2\times 2} \times (\mathbb{R}^2)^{nM\times nN} \times \mathbb{R}^{M\times N} \to S(\mathbb{R}^{nM\times nN})^{2\times 2} \times (\mathbb{R}^2)^{nM\times nN} \times \mathbb{R}^{M\times N} \times (\mathbb{R}^2)^{nM\times nN}$ and its adjoint $L^*$ (which can be obtained using Theorem 3.5) by

$$L = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -div_n^2 & 0 & I \\ -div_n^{'} & I & 0 \end{pmatrix}, \qquad L^* = \begin{pmatrix} I & 0 & -\nabla_n^2 & \mathcal{E}_n \\ 0 & I & 0 & I \\ 0 & 0 & I & 0 \end{pmatrix}. \tag{39}$$

In order to formulate imaging problems via the proposed model, we need to find the dual form of it. In the following theorem, we give the dual formulation of the $n$ - Shannon TGV, through a minimization problem.

**Theorem 4.2.** *(Dual formulation of $TGV^{2(\alpha)}_{SH(n)}$)*

$$\begin{array}{ll} TGV^{2(\alpha)}_{SH(n)}(u) = & \min \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} \\ s.t. & \bar{v} - \mathcal{E}_n\bar{w} = \frac{1}{n^2}\nabla_n^2 u. \end{array} \tag{40}$$

*where*

$$v = \begin{pmatrix} v_1 & v_2 \\ v_2 & v_3 \end{pmatrix}, v_i \in \mathbb{R}^{P\times Q}, i = 1, 2, 3, \|v\|_{2,1} = \sum_{(i,j)\in I_P\times I_Q} \sqrt{v_1^2(i,j) + 2v_2^2(i,j) + v_3^2(i,j)},$$

$$w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, w_i \in \mathbb{R}^{P\times Q}, i = 1, 2, \|w\|_{2,1} = \sum_{(i,j)\in I_P\times I_Q} \sqrt{w_1^2(i,j) + w_2^2(i,j)}.$$

*Proof.* From (38) and definition of operators in (39) we have

$$\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}(u) = \max\left\{ \frac{1}{n^2}\langle u, s\rangle : L\begin{pmatrix} v \\ w \\ s \end{pmatrix} = \begin{pmatrix} \bar{v} \\ \bar{w} \\ 0 \\ 0 \end{pmatrix}, \|\bar{v}\|_{2,\infty} \le \alpha_0, \|\bar{w}\|_{2,\infty} \le \alpha_1 \right\}.$$

Assume $\mathbf{u} = (v, w, s)^T$, $\mathbf{v} = (\bar{v}, \bar{w}, \bar{z}, \bar{t})^T$, $F(\mathbf{u}) = \frac{1}{n^2}\langle u, w\rangle$, and $G(\mathbf{v}) = I_{\{\|\bar{v}\|_{2,\infty}\le\alpha_0, \|\bar{w}\|_{2,\infty}\le\alpha_1, \bar{z}=0, \bar{t}=0\}}(\mathbf{v})$.
Therefore $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}(u)$ can be reformulated by the following optimization problem:

$$\max_{\mathbf{u}} -F(\mathbf{u}) - G(L(\mathbf{u})). \tag{41}$$

Note that (41) is equivalent to the following problem

$$\max_{\mathbf{u}} F(\mathbf{u}) - G(L(\mathbf{u})),$$

in the other word, if $\mathbf{u} = (v, s, w)$ is a solution of this problem, then $-\mathbf{u} = (-v, -w, -s)$ is the corresponding solution of (41), with the same optimal value.
From [17], It is well known that the dual formulation of (41) is:

$$\min_{\mathbf{v}} G^*(\mathbf{v}) + F^*(-L^*(\mathbf{v})), \tag{42}$$

where $G^*(\bar{v}, \bar{w}, \bar{s}, \bar{t}) = \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1}$ and

$$F^*(\tilde{v}, \tilde{w}, \tilde{s}) = \begin{cases} 0, & \tilde{v} = 0, \tilde{w} = 0, \tilde{s} = \frac{1}{n^2}u, \\ +\infty, & \text{othewise.} \end{cases}$$

Therefore (42) can be reformulated by

$$\begin{aligned} \mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}(u) = \min\ & \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} \\ & L^*\begin{pmatrix} \bar{v} \\ \bar{w} \\ \bar{s} \\ \bar{t} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{n^2}u \end{pmatrix}. \end{aligned} \tag{43}$$

By substituting $L^*$ in (43) we obtain (40) and the proof is completed. □

## 5. Applications on image processing

In this section, we apply the proposed model to formulate image restoration problems; e.g., denoising and deconvolution. For the both problems, the equivalent Fenchel-Rochafellar dual forms are obtained and a primal-dual algorithm is employed to solve the problems numerically (see [17] for instance). The analysis of the primal-dual algorithm for image restoration by means of Shannon TV is reviewed in Appendix as well.

### *5.1. Denoising*

Assume $x \in \mathbb{R}^{M\times N}$, and $\mathcal{R}$ is a discrete regularization function. Consider, the following denoising problem:

$$\min_u \frac{1}{2}\|x - u\|_2^2 + \mathcal{R}(u), \tag{44}$$

where, a positive multiplication of any variational functional such as TV, $\mathrm{TGV}^2_\alpha$, $\mathrm{TV}_{\mathrm{SH}(n)}$ and $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}$, can play the role of $\mathcal{R}$. In the sequel, for the proposed Shannon TGV, the dual problem of (44) will be explained, and a primal-dual algorithm will be suggested to solve the problem. Consider the following problem:

$$n\text{- Shannon TGV Denoising Problem} \quad \min_u \tfrac{1}{2}\|x - u\|_2^2 + \mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}(u), \tag{45}$$

From Theorem 4.2, (45), has the following equivalent version:

$$\begin{aligned} \min_{(u,\bar{v},\bar{w})} \quad & \tfrac{1}{2}\|x - u\|_2^2 + \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} \\ \text{s.t.} \quad & \bar{v} - \mathcal{E}_n\bar{w} = \tfrac{1}{n^2}\nabla_n^2 u. \end{aligned} \tag{46}$$

Define linear operator $T$ by $T = [-\frac{1}{n^2}\nabla_n^2 \quad -\mathcal{E}_n \quad I]$, then we get another equivalent form of (45):

$$\begin{aligned} \min_{(u,\bar{v},\bar{w})} \quad & \tfrac{1}{2}\|x - u\|_2^2 + \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} + I_K\left(T\begin{pmatrix} u \\ \bar{w} \\ \bar{v} \end{pmatrix}\right) \\ \text{s.t.} \quad & K = \{0\}. \end{aligned} \tag{47}$$

*5.1.1. Dual form of the n - Shannon TGV denoising*

Assume

$$\mathbf{u} = (u, \bar{w}, \bar{v})^T, G(\mathbf{u}) = \frac{1}{2}\|x-u\|_2^2 + \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1}, \mathbf{p} = p, F(\mathbf{p}) = I_K(p). \tag{48}$$

Now, we rewrite (47) by

$$\min_{\mathbf{u}} F(T(\mathbf{u})) + G(\mathbf{u}). \tag{49}$$

It is well known that, the dual form of (49) is

$$\max_{p} -F^*(\mathbf{p}) - G^*(-T^*(\mathbf{p})), \tag{50}$$

where

$$F^*(p) = 0, \quad G^*(u_1, w_1, v_1) = (u_1, x) + \frac{1}{2}\|u_1\|_2^2 + I_{\{\|v_1\|_{2,\infty}\leq\alpha_0, \|w_1\|_{2,\infty}\leq\alpha_1\}}(w_1, v_1). \tag{51}$$

As $T^* = [-\frac{1}{n^2}div_n^2 \quad div_n^{'} \quad I]^T$, by substituting $T^*$, in (51), we get the following dual formulation of (45):

**Theorem 5.1.** *n-Shannon TGV denoising problem (45) is equivalent to the following dual formulation:*

$$\begin{aligned} &\max_p \quad -\tfrac{1}{2n^4}\|div_n^2 p\|_2^2 - \tfrac{1}{n^2}\langle div_n^2 p, x\rangle \\ &s.t. \qquad \|p\|_{2,\infty} \leq \alpha_0, \quad \|div_n^{'} p\|_{2,\infty} \leq \alpha_1. \end{aligned} \tag{52}$$

*5.1.2. Primal-dual algorithm*

To solve problem (49) and its dual (50), simultaneously, we apply primal-dual algorithm 1.

**Data**: $T, T^*, F^*, G$, and $\epsilon$ as a tolerance
**Result**: $\mathbf{u}^k$ as primal solution and $\mathbf{p}^k$ as dual solution
initialization: Choose parameters $0 < \tau, \sigma < 1/\|T\|$ and initial estimates $(\mathbf{u}^0, \mathbf{p}^0)$, $\tilde{\mathbf{u}}^0 = \mathbf{u}^0$.
**while** *for* $k = 0, 1, \cdots$, $F(T(\boldsymbol{u}^k)) + G(\boldsymbol{u}^k) + F^*(\boldsymbol{p}^k) + G^*(-T^*(\boldsymbol{p}^k)) > \epsilon$ **do**
  1-$\mathbf{p}^{k+1} = \text{prox}_{\sigma F^*}(\mathbf{p}^k + \sigma T(\tilde{\mathbf{u}}^k))$
  2- $\mathbf{u}^{k+1} = \text{prox}_{\tau G}(\mathbf{u}^k - \tau T^* \mathbf{p}^{k+1})$
  3-$\tilde{\mathbf{u}}^{k+1} = 2\mathbf{u}^{k+1} - \mathbf{u}^k$
**end**

**Algorithm 1:** The Chambolle-Pock algorithm for solving problems (49) and (50)

To employ Algorithm 1 for denoising problem (45) and its dual (52), we need to compute $\text{prox}_{\sigma F^*}(p)$ and $\text{prox}_{\tau G}(\mathbf{u})$, for given $p$ and $\mathbf{u}$, where $G$ and $F^*$ are defined in 48 and 51 respectively. Obviously $\text{prox}_{\sigma F^*}(p) = p$ and

$$\text{prox}_{\tau G}(\mathbf{u}) = \left(\text{prox}_{\tau\frac{1}{2}\|x-u\|_2^2}(u), \text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(\bar{w}), \text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(\bar{v})\right)^T, \tag{53}$$

where $\text{prox}_{\tau\frac{1}{2}\|x-u\|_2^2}(u) = \frac{u+\tau x}{1+\tau}$ and for $(i, j) \in I_M \times I_N$,

$$\text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(\bar{v})(i,j) = \left(1 - \frac{\tau\alpha_0}{\max\{\tau\alpha_0, \|\bar{v}(i,j)\|_2\}}\right)\bar{v}(i,j), \quad \text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(\bar{w})(i,j) = \left(1 - \frac{\tau\alpha_1}{\max\{\tau\alpha_1, \|\bar{w}(i,j)\|_2\}}\right)\bar{w}(i,j). \tag{54}$$

As a result we propose Algorithm 2 for solving (45) and (52).

*5.2. Deconvolution*

Let $A$ is a convolution operator and $x$ is a kind of blurred image, which is concluded from a clean image $u$, through $A$ ($Au = x$). For a given regularizer $\mathcal{R}$, the Tikhonov type deconvolution problem is formulated with the following optimization model:

$$\min_u \frac{1}{2}\|x - Au\|_2^2 + \mathcal{R}(u), \tag{55}$$

**Data**: $x$ as a noisy image, and $\epsilon$ as a tolerance
**Result**: $\mathbf{u}^k = (u^k, v^k, w^k)$ as primal solution and $p^k$ as dual solution
initialization: Choose parameters $0 < \tau, \sigma < 1/\|T\|$ and initial estimates $(\mathbf{u}^0, p^0)$, $\quad \tilde{\mathbf{u}}^0 = \mathbf{u}^0$.
**while** *for* $k = 0, 1, \cdots,$ $\quad \frac{1}{2}\|x - u^k\|_2^2 + \alpha_0\|v^k\|_{2,1} + \alpha_1\|w^k\|_{2,1} + \frac{1}{2n^4}\|div_n^2 p^k\|_2^2 + \frac{1}{n^2}\langle div_n^2 p^k, x\rangle > \epsilon$ **do**

$1\text{-}p^{k+1} = p^k + \sigma(-\frac{1}{n^2}\nabla_n^2\tilde{u}^k - \mathcal{E}_n\tilde{w}^k + \tilde{v}^k)$,

$2\text{-}\ u^{k+1} = \frac{u^k + \tau\left(\frac{1}{n^2} div_n^2(p^{k+1}) + x\right)}{1+\tau}$,

$3\text{-}\ v^{k+1} = \text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(v^k - \tau p^{k+1})$,

$4\text{-}\ w^{k+1} = \text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(w^k - \tau div_n'(p^{k+1}))$,

$5\text{-}\tilde{u}^{k+1} = 2u^{k+1} - u^k$,

$6\text{-}\ \tilde{v}^{k+1} = 2v^{k+1} - v^k$,

$7\text{-}\ \tilde{w}^{k+1} = 2w^{k+1} - w^k$.

**end**

**Algorithm 2:** The Chambolle-Pock algorithm for solving problems (45) and (52)

We are going to solve this problem via variational models studied in this paper. Shannon TGV deconvolution is formulated as the following:

$$\text{n-Shannon TGV Deconvolution Problem} \qquad \min_u \frac{1}{2}\|x - Au\|_2^2 + \text{TGV}^{2(\alpha)}_{\text{SH}(n)}(u), \tag{56}$$

From Theorem 4.2, (56), has the following equivalent version:

$$\begin{array}{ll} \min_{(u,\bar{v},\bar{w})} & \frac{1}{2}\|x - Au\|_2^2 + \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} \\ \text{s.t.} & \bar{v} - \mathcal{E}_n\bar{w} = \frac{1}{n^2}\nabla_n^2 u. \end{array} \tag{57}$$

Let

$$T = \begin{pmatrix} -\frac{1}{n^2}\nabla_n^2 & I & -\mathcal{E}_n \\ A & 0 & 0 \end{pmatrix}, \tag{58}$$

then (57) can be reformulated by

$$\begin{array}{ll} \min_{(u,\bar{v},\bar{w},p,s)} & \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1} + \frac{1}{2}\|x - s\|_2^2 + I_{\{0\}}(p) \\ \text{s.t.} & T\begin{pmatrix} u \\ \bar{v} \\ \bar{w} \end{pmatrix} = \begin{pmatrix} p \\ s \end{pmatrix}. \end{array} \tag{59}$$

*5.2.1. Dual form of n-Shannon TGV deconvolution*

Assume $\mathbf{u} = (u, \bar{v}, \bar{w}), \mathbf{p} = (p, s)$, and define $G(\mathbf{u}) = \alpha_0\|\bar{v}\|_{2,1} + \alpha_1\|\bar{w}\|_{2,1}$, $F(\mathbf{p}) = \frac{1}{2}\|x - s\|_2^2 + I_{\{0\}}(p)$. The primal Fenchel-Rochafellar form of (59) is "$\min_{\mathbf{u}} F(T\mathbf{u}) + G(\mathbf{u})$". Now, we need to obtain the dual form "$\max_{\mathbf{p}} -F^*(\mathbf{p}) - G^*(-T^*(\mathbf{p}))$". One can verify

$$\begin{aligned} &F^*(\mathbf{p}) = F^*(p, s) = \langle s, x\rangle + \tfrac{1}{2}\|s\|_2^2, \\ &G^*(\mathbf{u}) = G^*(u, \bar{v}, \bar{w}) = \begin{cases} 0, & u = 0, \|\bar{v}\|_{2,\infty} \le \alpha_0, \|\bar{w}\|_{2,\infty} \le \alpha_1, \\ +\infty, & \text{otherwise}, \end{cases} \quad , T^* = \begin{pmatrix} -\frac{1}{n^2}\text{div}_n^2 & A^* \\ I & 0 \\ \text{div}_n' & 0 \end{pmatrix}. \end{aligned} \tag{60}$$

Consequently we get the following result:

**Theorem 5.2.** *n-Shannon TGV deconvolution problem (56) is equivalent to the following dual formulation:*

$$\begin{array}{ll} \max_{p,s} & -\langle s, x\rangle - \frac{1}{2}\|s\|_2^2 \\ s.t. & -\frac{1}{n^2} div_n^2 p + A^* s = 0, \\ & \|p\|_{2,\infty} \le \alpha_0, \quad \|div_n' p\|_{2,\infty} \le \alpha_1. \end{array} \tag{61}$$

*5.2.2. Primal-dual algorithm*

We solve n-Shannon TGV deconvolution problem (56) and its dual form (61), by means of Algorithm 1. The required proximal operators are as the following:

$$\begin{aligned} &\text{prox}_{\tau G}(\mathbf{u}) = \Big(u, \text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(\bar{w}), \text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(\bar{v})\Big)^T, \\ &\text{prox}_{\tau F^*}(\mathbf{p}) = \Big(p, \tfrac{s-\tau x}{1+\tau}\Big), \end{aligned} \tag{62}$$

where $\text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(\bar{w})$ and $\text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(\bar{v})$ are defined in (54). As a result, Algorithm 3, is a primal-dual algorithm to solve deconvolution problems (56) and (61).

**Data**: $x$ as a blurry image, and $\epsilon$ as a tolerance
**Result**: $\mathbf{u}^k = (u^k, v^k, w^k)$ as primal solution and $\mathrm{p}^k = (p^k, s^k)$ as dual solution
initialization: Choose parameters $0 < \tau, \sigma < 1/\|T\|$ and initial estimates $(\mathbf{u}^0, \mathbf{p}^0)$, $\quad \tilde{\mathbf{u}}^0 = \mathbf{u}^0$.
**while** *for* $k = 0, 1, \cdots,$ $\quad \frac{1}{2}\|x - u^k\|_2^2 + \alpha_0\|v^k\|_{2,1} + \alpha_1\|w^k\|_{2,1} + \langle s^k, x\rangle + \frac{1}{2}\|s^k\|_2^2 > \epsilon$ **do**

1-$p^{k+1} = p^k + \sigma(-\frac{1}{n^2}\nabla_n^2\tilde{u}^k - \mathcal{E}_n\tilde{w}^k + \tilde{v}^k)$,
2- $s^{k+1} = \frac{s^k + \sigma(A\tilde{u}^k - x)}{1+\sigma}$
3- $u^{k+1} = u^k - \tau\Big(A^* s - \frac{1}{n^2} div_n^2(p^{k+1})\Big)$,
4- $v^{k+1} = \text{prox}_{\tau\alpha_0\|\bar{v}\|_{2,1}}(v^k - \tau p^{k+1})$,
5- $w^{k+1} = \text{prox}_{\tau\alpha_1\|\bar{w}\|_{2,1}}(w^k - \tau div_n'(p^{k+1}))$,
6-$\tilde{u}^{k+1} = 2u^{k+1} - u^k$,
7- $\tilde{v}^{k+1} = 2v^{k+1} - v^k$,
8- $\tilde{w}^{k+1} = 2w^{k+1} - w^k$.

**end**

**Algorithm 3:** The Chambolle-Pock algorithm for solving problems (56) and (61)

*5.3. Computational complexity*

In the following, we address the computational complexity of Shannon TV and Shannon TGV versus TV and TGV, using primal-dual algorithm for denoising an $M \times N$ gray image per each iteration:

*TV computational costs*: computing discrete gradient and divergence take $4MN$ flops, $14MN$ flops, $MN$ taking square roots and $MN$ comparison are required for computation of proximals. Overall we need about $19MN = O(MN)$ flops which is completely efficient.

*n - Shannon TV computational costs*: computing each of discrete Shannon gradient and Shannon divergence needs $2MN$ flops plus calculating FFT (or IFFT) of an $M\times N$ discrete signal and FFT (or IFFT) of two $nM\times nN$ signals which take about $2MN(\log(M) + \log(N)) + n^2MN(\log(nM) + \log(nN))$ flops. The main body of the primal-dual algorithm needs $14n^2MN$ flops (computation of proximals) as well as $n^2MN$ square roots and comparisons. Overall we need about $4MN(\log(M)+\log(N))+2n^2MN(\log(nM)+\log(nN))+4MN+14n^2MN = O(n^2MN(\log(nM)+\log(nN)))$ flops.

*The second order TGV computational costs*: $5MN$ flops for computing $\mathcal{E}$ of a given signal, where $\mathcal{E} : (\mathbb{R}^{M\times N})^2 \to S(\mathbb{R}^{M\times N})^{2\times 2}$ is the first order gradient operator. $4MN$ flops for $\text{div}' = \mathcal{E}^*$, $8MN$ flops for the second order gradient $\nabla^2$ and $8MN$ flops for $\text{div}^2 = (\nabla^2)^*$. $10MN$ flops for proximal operators as well as $2MN$ square roots and $2MN$ comparisons. Main primal-dual algorithm body requires $32MN$ flops. Overall number of floating point computations is $67MN = O(MN)$.

*n-Shannon TGV computational costs*: four Shannon operators; $\mathcal{E}_n, \text{div}_n' = (\mathcal{E}_n)^*, \nabla_n^2$, and $\text{div}_n^2 = (\nabla_n^2)^*$ would be computed per iteration (each operator corresponds to an operator in TGV). To compute $\mathcal{E}_n$, FFT of two $nM \times nN$ matrices and IFFT of three $nM\times nN$ matrices are required ($5n^2MN(\log(nM)+\log(nN))$ flops). For $\nabla_n^2$, FFT of one $M\times N$ matrix and IFFT of three $nM \times nN$ is needed ($MN(\log(M) + \log(N)) + 3n^2MN(\log(nM) + \log(nN))$ flops). Moreover,

we need $5n^2MN(\log(nM)+\log(nN))$ flops for $\text{div}_n^{'}$, and $MN(\log(M)+\log(N))+2n^2MN(\log(nM)+\log(nN))$ flops for $\text{div}_n^2$. Proximal operators and main body primal-dual algorithm cost $42n^2MN$ (similar to TGV) we should add $2n^2MN$ square roots and comparisons. Finally, we get:

$$\text{Number of flops} = 2MN(\log(M)+\log(N)) + 15n^2MN(\log(nM)+\log(nN)) + 42n^2NM.$$

One can verify that for $n = 2, M = N = 128$

$$\text{Number of Shannon TV flops} \approx 11.14 \times \text{Number of TV flops}$$

$$\text{Number of Shannon TGV flops} \approx 12.58 \times \text{Number of TGV flops}$$

The above discussion confirms that, if we apply Shannon interpolation for any other variational model and set $n = 2, M = N = 128$ , we would expect more computational complexity, about 12 times compared to the original model.

## 6. Numerical experiments

In this section, we present numerical result of the proposed model and compare it with some authentic first and second order variational models. In the following, models which are compared in the experiments, parameter setting, noise ratios, number of iterations and sample image properties are expressed:

- Variational models in the comparisons: we have compared our method with three first order models, i.e., TV, anisotrpic TV and Shannon TV and three second order models, i.e., the second order TGV, infimal convolution (INFCON) and total Laplace (TL). The infimal convolution model and TGV are the most reliable higher order variational models in the literature.

- Parameters of the primal-dual algorithms: in the algorithms we need two parameters $\tau$ and $\sigma$. We set $\sigma = \tau$, for the both denoising and deconvolution problems. For Shannon TV and Shannon TGV we set $\tau = 0.1$ and for the others we set $\tau = \frac{5}{37}$.

- Regularization parameters: the optimal parameters are determined for each method and for the restoration of a given image separately. For instance, we set optimal parameter values $(\alpha_0^*, \alpha_1^*)$, such that we get the best result in terms of psnr values. In other words, assume $x$ is noisy (or blurry) , and $x_{ref}$ is the clean image and $\alpha_0^* = 2\alpha_1^*$ (or $\alpha_0^* = \alpha_1^*$) . Define

$$\alpha_1^*(x) = \text{argmax}_{\alpha_1>0} = \Big\{\text{psnr}(u_\alpha^*, x_{ref}), \alpha = (\alpha_0, \alpha_1), \alpha_0 = 2\alpha_1 \text{ (or } \alpha_0 = \alpha_1), u_\alpha^* \text{ is the solution of (56)}\Big\}. \quad (63)$$

  the optimal parameter $\lambda^*$ for the first order models (TV, Shannon TV and anisotropic TV), and $\alpha^* = (\alpha_0^*, \alpha_1^*)$ for the second order models (TGV, infimal convolution and Shannon TGV), for each image and each noise ratio are reported in tables 2 and 3 Note that for the second order denoising models we assumed $\lambda^* = \alpha_1^* = \frac{1}{2}\alpha_0^*$ and for deconvolution $\lambda^* = \alpha_1^* = \alpha_0^*$.

- Noise ratios: artificial noises are added to the image for denoising problems. For the images with intensity values in the range $[0, 255]$, the added noises are Gaussian noises with 0 mean and the range of variances from 150 to 1000. Larger variance needs larger value of the regularization parameter and clearly the restored image is less similar to the reference image. This facts can be observed in tables ... .

- Iteration numbers: in our experiments for all models, the number of iteration for denoising is 2000 and for deconvolution is 5000.

- Sample images: we have solved denoising (with the range of variances from 150 to 1000) and deconvolution problems for sample "png" images; "Bird ($186 \times 154$)", "Cat & Dog ($261 \times 261$)", "Girl ($150 \times 150$)" and " Butterfly ($285 \times 193$)" (see Fig. 1).

(a) Bird

(b) Butterfly

(c) Girl

(d) Cat & Dog

Figure 1: Reference images.

### *6.1. Denoising*

Assume $u$ is a clean image, we add a Gaussian noise vector $v$ to $u$ and obtain $x = u + v$ as a noisy image. Algorithm 2 is employed for the proposed model. Figures 2-4 display the restored images for "Cat & Dog" image with noise ratio 400. The size of zoomed in images (details) is $90 \times 90$. As it is discussed in [16], by transition from TV to $\mathrm{TV}_{\mathrm{SH}}$, the staircase effects resulted from TV, are attenuated via $\mathrm{TV}_{\mathrm{SH}}$ regularizer. However, the artifacts still remain with the narrower depth. TGV is a well-known model in deceasing artifacts, especially staircases [15]. The results show that the outcome of the second order TGV model admits distinguished performance to diminish staircase effects in the flat areas. As we expect, transition from the TGV to 2-Shannon TGV, leads to attenuation and narrowing the remaining artifacts observed in the TGV regularizer (compare (e) and (f) in the Fig. 4). In other words, $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$ is an authentic generalization for $\mathrm{TV}_{\mathrm{SH}}$ in computational aspects, as well as the theoretical foundations.
For any added random noise with the noise ratio in the range [150, 1000], we have executed each algorithm for a given image, 10 times, and the mean psnr and mean ssim values are reported in Table 2. The results confirm the effectiveness of Shannon interpolation for both TV and TGV models. Specifically, we get higher accuracy of $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$ compared to $\mathrm{TGV}^2_\alpha$.

#### *6.1.1. High and low frequencies analysis of the model*

The proposed model performs better than other state of the art models in both high and low frequencies. TGV model is famous for reducing artifacts in smooth areas (low frequencies) and Shannon TV is capable to preserve details (high frequencies). The introduced model preserves the benefits of Shannon TV to reduce the artifacts by means of trigonometric interpolation. Moreover, it retains the distinguishing property of the second order discrete TGV, to remove staircase effects. In other words, for low frequencies the proposed model acts like the second order TGV and for high frequencies acts like Shannon TV. To illustrate this fact, we have applied ideal high pass filter for denoising results to address the differences of variational models in high frequencies. Ideal high pass filter in frequency domain is defined by

$$H(u,v) = \begin{cases} 0, & D(u,v) \leq D_0, \\ 1, & D(u,v) > D_0. \end{cases},$$

(*a*) reference image

(*b*) noisy image

(*c*) TV - restored

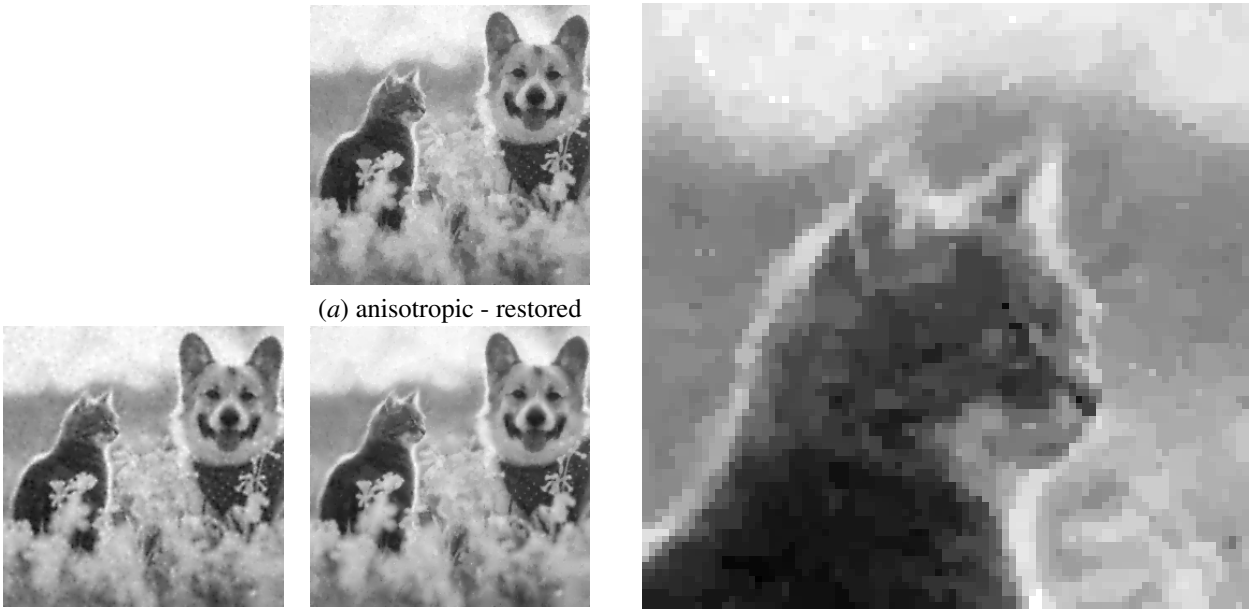
(*d*) reference image: details

(*e*) noisy image: details

(*f*) TV - restored: details

.

Figure 2: Denoising process and TV result: reference image, noisy image and restored image using TV model are shown in (a), (b) and (c) respectively. Some details of reference and noisy images are shown in (d) and (e). The details of restored image using TV variational model is shown in (f), where deep staircase artifacts around the borders are evident.

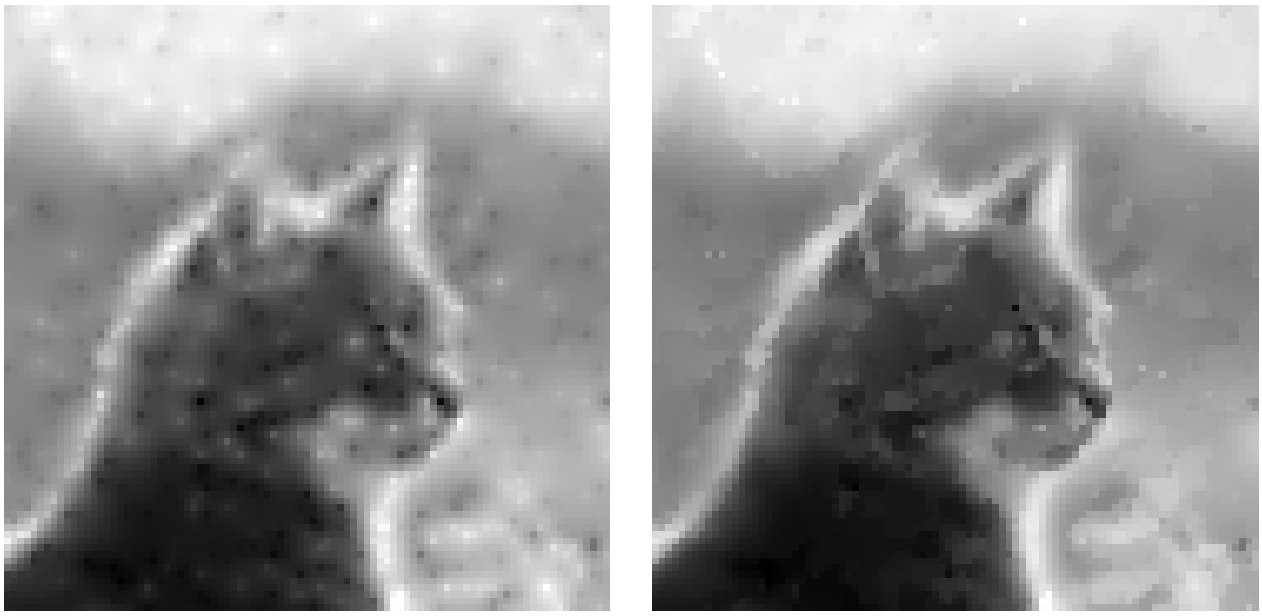


($a$) anisotropic - restored

($b$) TL - restored

($c$) INFCON - restored

($d$) anisotropic - restored : details



($e$) TL - restored : details

($f$) INFCON - restored : details

Figure 3: Comparing results of the variational denoising problems: restored images via anisotrpic TV, total Laplace and the infimal convolution variational models are shown in (a), (b) and (c) respectively. Consider details of restored images illustrated in (d), (e) and (f). Anisotropic TV artifacts is very similar to TV model (Fig. 2 (f)), whreas in both second order models; TL and infimal convolution, the staircase effects are diminished. However, some some sharp borders in smooth areas for infimal convolution and white noises in TL model are appeared.

($a$) $\mathrm{TV}_{\mathrm{SH}(2)}$ - restored

($b$) $\mathrm{TGV}^2_{\alpha}$ - restored

($c$) $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$ - restored

($d$) $\mathrm{TV}_{\mathrm{SH}(2)}$ - restored : details

($e$) $\mathrm{TGV}^2_{\alpha}$ - restored : details

($f$) $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$ - restored : details

Figure 4: Comparing results of the variational denoising problems: restored images via 2- Shannon TV, the second order TGV and the second order 2- Shannon TGV variational models are shown in (a), (b) and (c) respectively. Consider details of restored images illustrated in (d), (e) and (f). Obviously the deep artifacts observed in the restored image using TV model (Fig. 2 (f)) are attenuated in (d). However, some undesirable cloudy like pieces still remain, whereas for TGV model (e), these parts are completely faded. The restored image using proposed Shannon TGV (f), preserves the properties of both TGV and Shannon TV models simultaneously. Referring to Fig 3, obviously, TGV result is very similar to infimal convolution, but a little bit better in smooth areas, whereas Shannon TGV performs better than all variational models in the comparison in terms of eliminating staircase effects and noise removal.

where $D(u,v) = \sqrt{u^2+v^2}$. We have compared the denoising results after ideal high pass filter for $D_0 = 10$. See figures 5 and 6 for more details.

### 6.2. Deconvolution

Assume $\ker : \{-\frac{k-1}{2}, \cdots, -1, 0, 1, \cdots, \frac{k-1}{2}\} \times \{-\frac{k-1}{2}, \cdots, -1, 0, 1, \cdots, \frac{k-1}{2}\} \to \mathbb{R}, k$ odd, is a kernel convolution and $u : \{1, 2, \cdots, m\} \times \{1, 2, \cdots, n\} \to \mathbb{R}$ is a grayscale image. Let

$$\bar{u} : \{-\frac{k-1}{2}, \cdots, -1, 0, 1, \cdots, m + \frac{k-1}{2}\} \times \{-\frac{k-1}{2}, \cdots, -1, 0, 1, \cdots, n + \frac{k-1}{2}\} \to \mathbb{R}$$

is the extension of image $u$ with the property $\bar{u}(i,j) = u(i,j), (i,j) \in \{1, 2, \cdots, m\} \times \{1, 2, \cdots, n\}$ and at the other parts of the domain $\bar{u}(i,j)$ is determined by symmetric boundary conditions. In our simulation, the linear convolution operator $A : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ is defined by

$$Au(i,j) = \sum_{\alpha=-\frac{k-1}{2}}^{\frac{k-1}{2}} \ker(\alpha,\beta)\bar{u}(i-\alpha, j-\beta), (i,j) \in \{1, 2, \cdots, m\} \times \{1, 2, \cdots, n\} \tag{64}$$

One can verify that the adjoint operator $A^*$, is

$$A^*v(i,j) = \sum_{\alpha=-\frac{k-1}{2}}^{\frac{k-1}{2}} \ker(\alpha,\beta)\bar{v}(i+\alpha, j+\beta), v : \{1, 2, \cdots, m\} \times \{1, 2, \cdots, n\} \to \mathbb{R}, \tag{65}$$

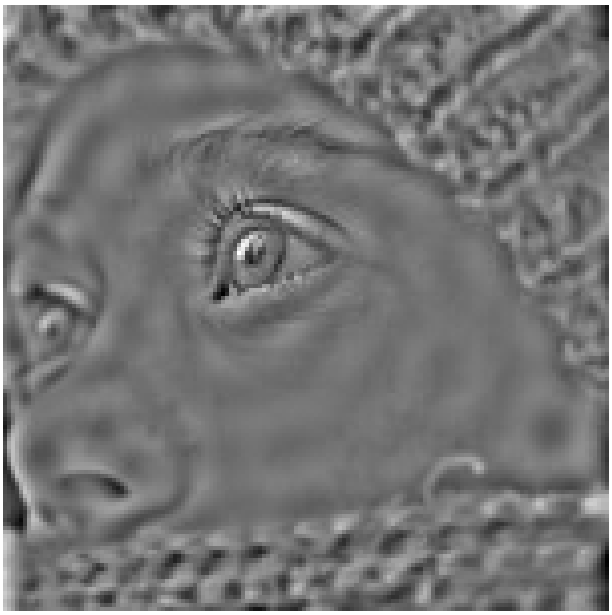
where $\bar{v}$ is the extended version of $v$, with the suitable boundary conditions.
Assume $u$ is the original image and $x = Au + v$ is the blurry image obtained from $u$ through a linear convolution operator $A$ with a kernel convolution and a Gaussian noise vector $v$ (mean=0 and variance = 7, for the images with intensity values in the range $[0, 255]$). Our purpose is restoring $x$ by means of variational models using primal-dual Algorithm 1. Specifically, to solve problem 56, we apply Algorithm 3. The reconstructed images for the sample image "Girl" for infimal convolution, Shannon TV, TGV and the proposed model are illustrated in figures 7 and 8. The size of zoomed in images (details) is $90 \times 90$. The outcome of $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(n)}$ show superiority in noise removal, preserving the details and diminishing staircase artifact. The quantitative details of the obtained results are expressed in Table 3.

## 7. Conclusion

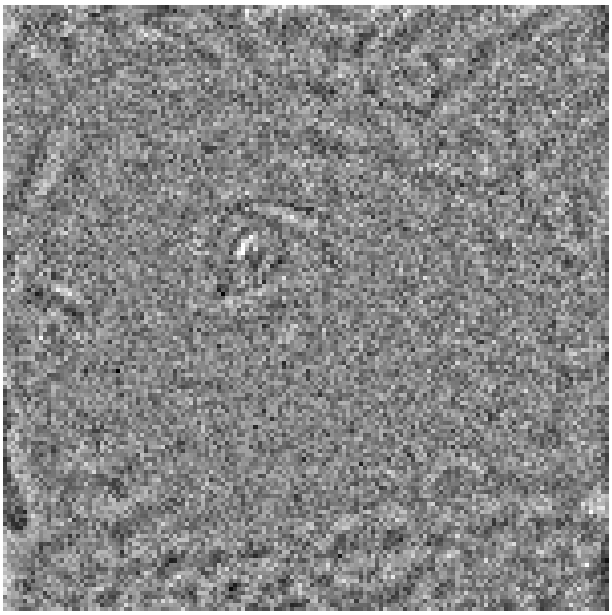
In this contribution, we get an idea from the Shannon TV model of Abergel and Moisan [16], and propose a new discrete second order TGV model based on Shannon interpolation. The new discrete derivative operators are defined and the divergence operators as the adjoints of the derivatives are expressed. Consequently, the new discrete TGV model is defined as an optimization problem. In the simulations results, it is verified that the introduced model, preserves the benefits of Shannon TV to reduce the artifacts by means of trigonometric interpolation. Moreover, it retains the distinguishing property of the second order discrete TGV, to remove staircase effects. It is possible to extend the present model for the higher order TGVs or other higher order variational models in the same way.

## 8. References

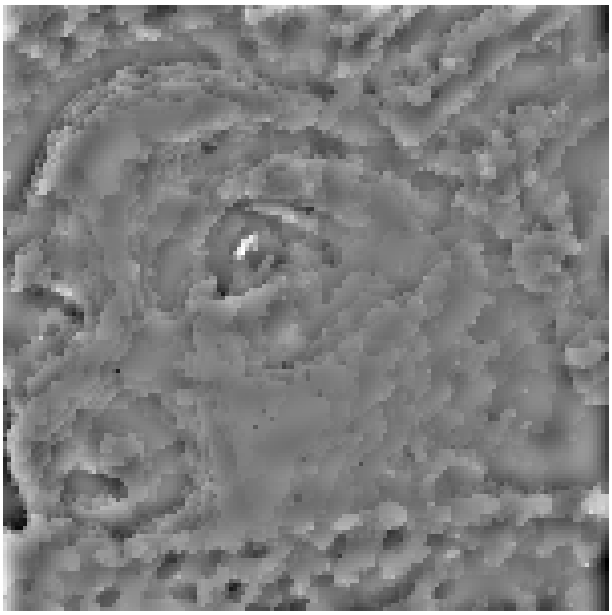

[1] Rudin, L. I., Osher, S. and Fatemi, E.: Nonlinear total variation based noise removal algorithms, Phys. D 60, 259-268 (1992).
[2] Condat, L.: Discrete total variation: new definition and minimization, SIAM Journal on Imaging Sciences 10, 1258-1290 (2017).
[3] Hosseini, A.: New discretization of total variation functional for image processing tasks, Signal Processing: Image Communication 78, 62-76 (2019).
[4] Pang, Z. F., Zhang, H. L., Luo, S. and Zeng, T.: Image denoising based on the adaptive weighted $TV^p$ regularization, Signal Processing 167, 107325 (2020).
[5] You, Y. L. and Kaveh, M.: Fourth-order partial differential equations for noise removal. IEEE Transactions on Image Processing 9, 1723-1730 (2000).



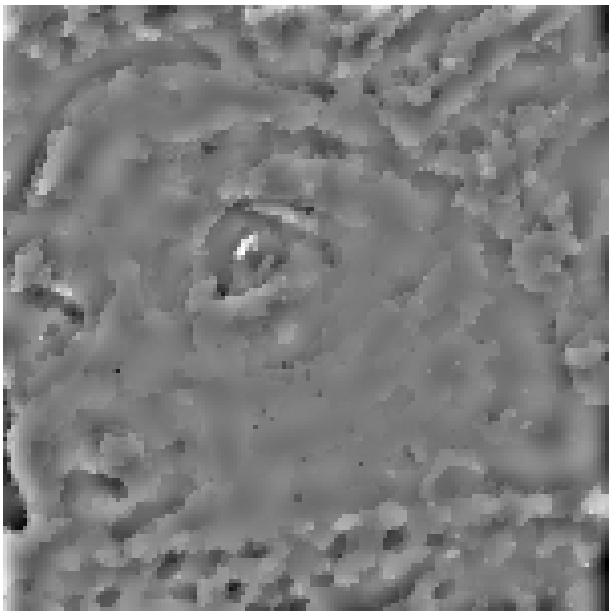
(a) reference



(b) noisy

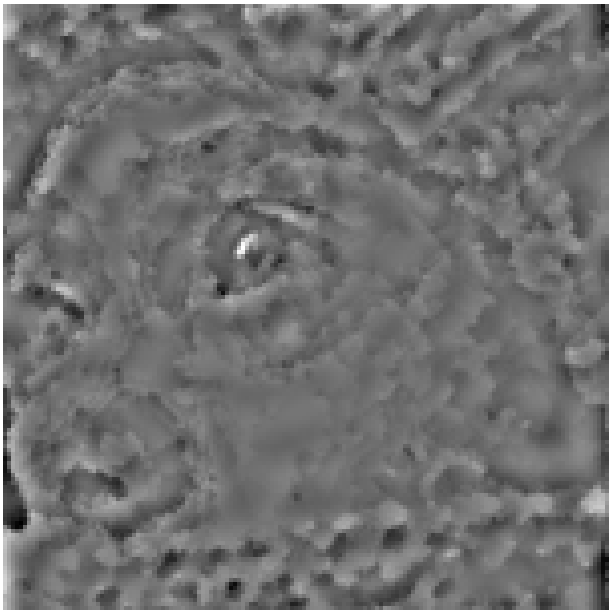


(c) TV

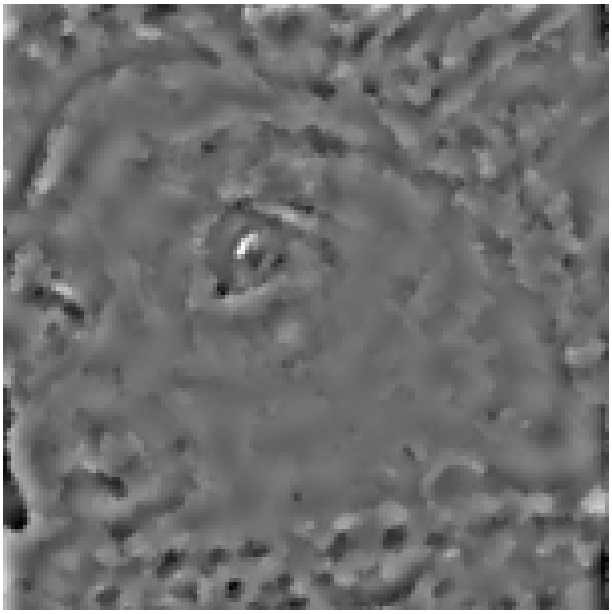


(d) $\mathrm{TGV}_{\alpha}^{2}$

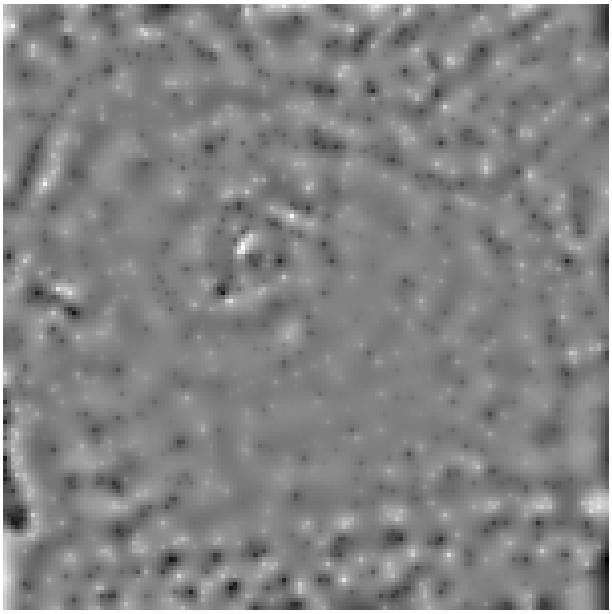
Figure 5: Denoising sample image "Girl" results after ideal high pass (a) reference, (b) noisy, (c) TV and (d) second order TGV. It is clear that TGV and TV are very similar in the areas containing details (see hat and scarf textures). TV as a first order model could not recognize the smooth areas (all over the cheek contains edges), whereas TGV as a second order mode is capable to restore smooth areas.



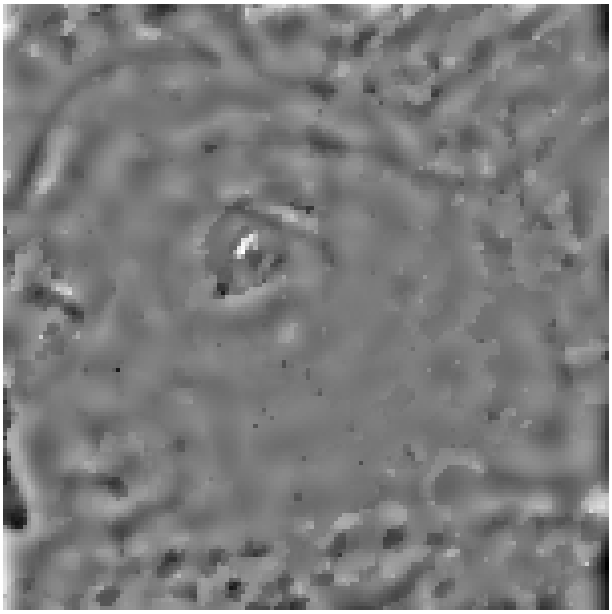
(a) $\mathrm{TV}_{\mathrm{SH}(2)}$



(b) $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$



(c) total Laplace



(d) infimal convolution

Figure 6: Denoising sample image "Girl" results after ideal high pass (a) Shannon TV, (b) Shannon TGV, (c) total Laplace and (d) infimal convolution. Shannon TV suffers from the same drawback of any first order models, similar to TV, the smooth parts contain edges, whereas, compared to TGV, the details (hat and scarf texture) are more like the reference (in high frequencies Shannon TV is better than TGV). The Shannon TGV model (proposed) performance in the areas containing details is very similar to Shannon TV, whereas in smooth areas performs even better than TGV. In total Laplace and infimal convolution which are second order models, results are blurry and the main edges are attenuated which is a drawback.

(*a*) reference image

(*b*) blurry & noisy image

(*c*) TV - restored

(*d*) reference image: details

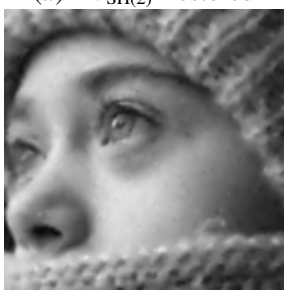
(*e*) blurry & noisy image: details

(*f*) TV - restored: details

Figure 7: Deconvolution problem and infimal convolution result: reference image, blurry & noisy image and restored image using infimal convolution model are shown in (a), (b) and (c) respectively. Some details of reference and blurry & noisy images are shown in (d) and (e). The details of restored image using infimal convolution is shown in (f).

($a$) $\mathrm{TV}_{\mathrm{SH}(2)}$- restored

($b$) $\mathrm{TGV}^2_\alpha$- restored



($c$) $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$- restored

($d$) $\mathrm{TV}_{\mathrm{SH}(2)}$- restored : details

($e$) $\mathrm{TGV}^2_\alpha$- restored : details

($f$) $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$- restored : details

Figure 8: Comparing results of the variational deconvolution problems: restored images via 2- Shannon TV, the second order TGV and the second order 2- Shannon TGV variational models are shown in (a), (b) and (c) respectively. Consider details of restored images illustrated in (d), (e) and (f). Obviously the staircase artifacts observed in the restored image using infimal convolution model (Fig. 7 (f)) are disappeared in (d). However, in the smooth areas, some unwanted edges are observed in TGV result which are eliminated with the proposed model.

| Noise(variance) | | 150 | | | 400 | | | 500 | | | 600 | | | 1000 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Metric | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM |
| Image | | | | | | | Cat & Dog | | | | | | | | |
| TV | 0.030 | 31.98 | 0.9008 | 0.050 | 29.81 | 0.8267 | 0.065 | 28.92 | 0.8047 | 0.065 | 28.21 | 0.8270 | 0.100 | 27.62 | 0.7877 |
| TL | 0.030 | 31.95 | 0.9079 | 0.050 | 29.89 | 0.8426 | 0.060 | 29.13 | 0.8297 | 0.060 | 28.54 | 0.8349 | 0.100 | 27.54 | 0.7820 |
| Anisotropic | 0.030 | 31.65 | 0.8983 | 0.050 | 29.80 | 0.8425 | 0.050 | 28.54 | 0.8231 | 0.055 | 28.06 | 08205 | 0.082 | 27.43 | 0.7791 |
| INFCON | 0.030 | 32.54 | **0.9144** | 0.055 | 30.53 | 0.8644 | 0.065 | 29.73 | 0.8545 | 0.067 | 28.80 | 0.8501 | 0.098 | 28.19 | **0.8157** |
| $TV_{SH(2)}$ | 0.028 | 32.31 | 0.9095 | 0.050 | 30.52 | 0.8627 | 0.060 | 29.66 | 0.8517 | 0.060 | 28.65 | 0.8454 | 0.083 | 28.12 | 0.8067 |
| $TGV^2_\alpha$ | 0.030 | 32.22 | 0.9081 | 0.060 | 30.24 | 0.8621 | 0.065 | 29.41 | 0.8436 | 0.066 | 28.44 | 0.8372 | 0.100 | 27.85 | 0.8040 |
| $TGV^{2(\alpha)}_{SH(2)}$ | 0.027 | **32.56** | **0.9143** | 0.050 | **30.65** | **0.8709** | 0.055 | **29.81** | **0.8560** | 0.058 | **28.81** | **0.8518** | 0.082 | **28.22** | 0.8144 |
| Image | | | | | | | Girl | | | | | | | | |
| TV | 0.032 | 31.87 | 0.9039 | 0.057 | 29.30 | 0.8469 | 0.065 | 28.71 | 0.8324 | 0.070 | 28.47 | 0.8189 | 0.096 | 27.30 | 0.7826 |
| TL | 0.024 | 32.60 | 0.9189 | 0.048 | 29.80 | 0.8649 | 0.057 | 29.39 | 0.8585 | 0.070 | 29.05 | 0.8482 | 0.098 | 27.76 | 0.8084 |
| Anisotropic | 0.026 | 31.65 | 0.8987 | 0.045 | 29.11 | 0.8384 | 0.055 | 29.00 | 0.8274 | 0.058 | 28.25 | 0.8102 | 0.078 | 27.09 | 0.7710 |
| INFCON | 0.030 | 32.87 | 0.9279 | 0.055 | 30.20 | 0.8798 | 0.067 | 29.55 | 0.8690 | 0.070 | 29.39 | 0.8592 | 0.096 | 28.13 | 0.8268 |
| $TV_{SH(2)}$ | 0.026 | 32.58 | 0.9176 | 0.047 | 29.89 | 0.8673 | 0.052 | 29.36 | 0.8547 | 0.060 | 29.10 | 0.8452 | 0.080 | 27.81 | 0.8086 |
| $TGV^2_\alpha$ | 0.032 | 32.40 | 0.9203 | 0.057 | 29.80 | 0.8708 | 0.065 | 29.19 | 0.8581 | 0.072 | 28.97 | 0.8479 | 0.098 | 27.79 | 0.8138 |
| $TGV^{2(\alpha)}_{SH(2)}$ | 0.030 | **32.91** | **0.9309** | 0.045 | **30.22** | **0.8814** | 0.050 | **29.72** | **0.8706** | 0.080 | **29.46** | **0.8640** | 0.082 | **28.15** | **0.8294** |
| Image | | | | | | | Bird | | | | | | | | |
| TV | 0.028 | 31.16 | 0.8693 | 0.052 | 28.71 | 0.8025 | 0.060 | 28.12 | 0.7831 | 0.066 | 27.81 | 0.7694 | 0.090 | 26.78 | 0.7363 |
| TL | 0.018 | 30.72 | 0.8563 | 0.038 | 28.12 | 0.7780 | 0.046 | 27.55 | 0.7552 | 0.054 | 27.27 | 0.7465 | 0.084 | 26.38 | 0.7244 |
| Anisotropic | 0.022 | 31.03 | 0.8639 | 0.042 | 28.58 | 0.7965 | 0.055 | 28.00 | 0.7743 | 0.054 | 27.69 | 0.7626 | 0.074 | 26.67 | 0.7279 |
| INFCON | 0.028 | 31.34 | 0.8805 | 0.052 | 28.88 | 0.8178 | 0.058 | 28.26 | 0.7946 | 0.066 | 27.97 | 0.7857 | 0.088 | 26.98 | 0.7521 |
| $TV_{SH(2)}$ | 0.022 | 31.36 | 0.8752 | 0.042 | 28.91 | 0.8132 | 0.048 | 28.30 | 0.7939 | 0.054 | 28.05 | 0.7848 | 0.072 | 26.99 | 0.7509 |
| $TGV^2_\alpha$ | 0.028 | 31.24 | 0.8735 | 0.052 | 28.79 | 0.8182 | 0.060 | 28.19 | 0.7890 | 0.068 | 27.89 | 0.7788 | 0.090 | 26.86 | 0.7420 |
| $TGV^{2(\alpha)}_{SH(2)}$ | 0.023 | **31.40** | **0.8810** | 0.044 | **28.92** | **0.8216** | 0.050 | **28.31** | **0.8016** | 0.054 | **28.08** | **0.7890** | 0.070 | **27.04** | **0.7534** |
| Image | | | | | | | Butterfly | | | | | | | | |
| TV | 0.028 | 31.51 | 0.8881 | 0.052 | 28.78 | 0.8376 | 0.060 | 28.12 | 0.8260 | 0.066 | 27.74 | 0.8109 | 0.086 | 26.38 | 0.7706 |
| TL | 0.016 | 30.68 | 0.8514 | 0.034 | 27.83 | 0.7798 | 0.040 | 27.20 | 0.7627 | 0.048 | 26.77 | 0.7564 | 0.068 | 25.48 | 0.7117 |
| Anisotropic | 0.022 | 31.37 | 0.8804 | 0.042 | 28.63 | 0.8282 | 0.048 | 28.01 | 0.8143 | 0.054 | 27.60 | 0.8023 | 0.072 | 26.26 | 0.7649 |
| INFCON | 0.028 | 31.77 | 0.9028 | 0.052 | 29.03 | 0.8600 | 0.058 | 28.37 | 0.8425 | 0.066 | 27.97 | 0.8356 | 0.086 | 26.59 | 0.7977 |
| $TV_{SH(2)}$ | 0.022 | 31.73 | 0.8926 | 0.042 | 29.05 | 0.8488 | 0.048 | 28.41 | 0.8370 | 0.054 | 28.03 | 0.8277 | 0.070 | 26.65 | 0.7869 |
| $TGV^2_\alpha$ | 0.030 | 31.65 | 0.9059 | 0.052 | 28.91 | 0.8517 | 0.060 | 28.26 | 0.8417 | 0.066 | 27.86 | 0.8264 | 0.088 | 26.50 | 0.7946 |
| $TGV^{2(\alpha)}_{SH(2)}$ | 0.023 | **31.85** | **0.9060** | 0.042 | **29.13** | **0.8605** | 0.050 | **28.47** | **0.8566** | 0.054 | **28.10** | **0.8411** | 0.070 | **26.70** | **0.8003** |

Table 2: Summary of experiments for denoising


[6] Lysaker, M., Lundervold, A. and Tai, X. C.: Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Transactions on Image Processing 12, 1579-1590 (2003).

[7] Scherzer, O.: Denoising with higher order derivatives of bounded variation and an application to parameter estimation, Computing 60, 1-27 (1998).

[8] Hinterberger, W. and Scherzer O.: Variational methods on the space of functions of bounded hessian for convexification and denoising, Computing, 76, 109-133 (2006).

[9] Lai, R. J., Tai, X. C. and Chan, T. F.: A ridge and corner preserving model for surface restoration, SIAM Journal on Scientific Computing, 35, A675-A695 (2013).

[10] Bergounioux, M. and Piffet, L.: A second-order model for image denoising, Set-Valued and Variational Analysis 18, 277-306 (2010).

[11] Chan, R. H., Liang, H. X., Wei, S. H., Nikolova, M. and Tai, X. C.: High-order total variation regularization approach for axially symmetric object tomography from a single radiograph, Inverse Problems and Imaging 9, 55-77 (2015).

[12] Papafitsoros, K. and Schönlieb, C. B.: A combined first and second order variational approach for image reconstruction, Journal of Mathematical Imaging and Vision 48, 308-338 (2014).

[13] Chambolle, A. and Lions, P. L.: Image recovery via total variation minimization and related problems, Numerische Mathematik 76, 167-188 (1997).

[14] Zhang, H., Tang, L., Fang, Z., Xiang, C. and Li, C.: Nonconvex and nonsmooth total generalized variation model for image restoration, Signal Processing 143, 69-85 (2018).

[15] Bredies, K., Kunisch, K. and Pock, T.: Total generalized variation, SIAM J. Imaging Sci. 3, 492-526 (2010).

[16] Abergel, R. and Moisan, L.: The Shannon total variation. J. Math Imaging Vis. 59, 341-370 (2017).

[17] Chambolle, A., Caselles, V., Cremers, D., Novaga, M. and Pock, T.: An Introduction to Total Variation for Image Analysis, in Theoretical Foundations and Numerical Methods for Sparse Recovery, Germany, Berlin: De Gruyter, vol. 9, 263-340, 2010.

| Image | Cat & Dog | | | Girl | | | Bird | | | Butterfly | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Metric | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM | $\lambda^*$ | M-PSNR | M-SSIM |
| TV | $1.20\times10^{-4}$ | 31.62 | 0.8998 | $1.18\times10^{-4}$ | 31.67 | 0.9048 | $8.60\times10^{-4}$ | 29.86 | 0.8556 | $7.80\times10^{-4}$ | 29.16 | 0.8742 |
| TL | $6.60\times10^{-4}$ | 31.28 | 0.8817 | $7.80\times10^{-4}$ | 31.86 | 0.9110 | $5.60\times10^{-4}$ | 29.14 | 0.8310 | $4.60\times10^{-4}$ | 28.05 | 0.8163 |
| INFCON | $9.80\times10^{-4}$ | 32.24 | 0.9107 | $9.40\times10^{-4}$ | 32.56 | 0.9232 | $7.00\times10^{-4}$ | 29.93 | 0.8540 | $6.20\times10^{-4}$ | 29.21 | 0.8543 |
| $\mathrm{TV}_{\mathrm{SH}(2)}$ | $1.02\times10^{-3}$ | 32.32 | 0.9114 | $9.00\times10^{-4}$ | 32.60 | 0.9212 | $7.00\times10^{-4}$ | 30.18 | 0.8665 | $6.40\times10^{-4}$ | **29.50** | 0.8852 |
| $\mathrm{TGV}^2_\alpha$ | $5.00\times10^{-4}$ | 31.89 | 0.8802 | $7.00\times10^{-4}$ | 32.91 | 0.9251 | $4.60\times10^{-4}$ | 29.70 | 0.8378 | $7.00\times10^{-4}$ | 29.00 | 0.8653 |
| $\mathrm{TGV}^{2(\alpha)}_{\mathrm{SH}(2)}$ | $7.50\times10^{-4}$ | **32.49** | **0.9131** | $7.50\times10^{-4}$ | **33.20** | **0.9330** | $6.30\times10^{-4}$ | **30.19** | **0.8683** | $6.00\times10^{-4}$ | **29.50** | **0.8871** |

Table 3: Summary of Experiments for deconvolution

## Appendix

*Shannon TV primal-dual algorithm for image restoration*

Assume $u \in \mathbb{R}^{M\times N}$. From 10, image restoration problem, using $n$ - Shannon TV model can be formulated by

$$\min_u \frac{1}{2}\|Au - x\|^2 + \frac{1}{n^2}\sum_{(k,l)\in\Omega_n} |\nabla_n u(k,l)|, \tag{66}$$

or equivalently

$$\min_u \frac{1}{2}\|Au - x\|^2 + \frac{1}{n^2}\|\nabla_n\|_{2,1}, \tag{67}$$

Assume $T = [A, \nabla_n]^T, \mathbf{p} = (p, q_1, q_2) \in (\mathbb{R}^{M\times N})^3, \mathbf{q} = (q_1, q_2) \in (\mathbb{R}^{M\times N})^2$. Now define

$$F(\mathbf{p}) = \frac{1}{2}\|p - x\|_2^2 + \frac{\lambda}{n^2}\|\mathbf{q}\|_{2,1}, \quad G : \mathbb{R}^{M\times N} \to \mathbb{R}, G = 0. \tag{68}$$

One can investigate that

$$F^*(\mathbf{p}) = \frac{1}{2}\|p\|_2^2 + \langle p, x\rangle + I_{\left\{\|\mathbf{q}\|_{2,\infty}\le\frac{\lambda}{n^2}\right\}}(\mathbf{q}), \quad G^*(u) = I_{\{u=0\}}(u), \quad T^* = [A^*, -\mathrm{div}_n]. \tag{69}$$

From (47) and (50), the dual formulation of (66) is

$$\begin{array}{ll} \max_{\mathbf{p}=(p,\mathbf{q})} & -\frac{1}{2}\|p\|_2^2 - \langle p, x\rangle \\ & A^*p - \mathrm{div}_n(\mathbf{q}) = 0, \\ & \|q\|_{2,\infty} \le \frac{\lambda}{n^2}. \end{array} \tag{70}$$

Moreover, required proximal operators for primal - dual algorithms are as follow

$$\mathrm{prox}_{\sigma F^*}(\mathbf{p}) = (r(p), s_1(\mathbf{q}), s_2(\mathbf{q})), r_1 = \frac{p}{1+\sigma}, \quad s_k(\mathbf{q}))(i,j) = \begin{cases} q_k(i,j), & \|\mathbf{q}(i,j)\|_2 \le \frac{\lambda}{n^2}, \\ \frac{q_k(i,j)}{\|\mathbf{q}(i,j)\|_2}(\frac{\lambda}{n^2}), & \text{else} \end{cases}, k = 1, 2. \tag{71}$$

$$\mathrm{prox}_{\tau G}(u) = u.$$

Consequently, we get Algorithm 4. Note that for denoising $A = I$ and for deconvolution, $A$ is defined as (64).

**Data**: $x$ as a noisy image, and $\epsilon$ as a tolerance
**Result**: $u^k$ as primal solution and $\mathbf{p}^k = (p^k, q_1^k, q_2^k), \mathbf{q}^k = (q_1^k, q_2^k)$ as dual solution
initialization: Choose parameters $0 < \tau, \sigma < 1/\|T\|$ and initial estimates $(u^0, \mathbf{p}^0, ), \quad \tilde{u}^0 = u^0$.
**while** *for* $k = 0, 1, \cdots, \quad \frac{1}{2}\|Au - x\|^2 + \frac{1}{n^2}\|\nabla_n\|_{2,1} + \frac{1}{2}\|p\|_2^2 + \langle p, x\rangle > \epsilon$ **do**
  1-$p^{k+1} = \frac{p_k + \sigma A\tilde{u}^k}{1+\sigma}$
  2- $\mathbf{q}^{k+1} = (q_1^{k+1}, q_2^{k+1}) = (s_1(\mathbf{q}^k + \sigma\nabla_n\tilde{u}^k), s_2(\mathbf{q}^k + \sigma\nabla_n\tilde{u}^k))$,
  3- $u^{k+1} = u^k + \tau(\mathrm{div}_n(\mathbf{q}^{k+1} - A^* p_1^{k+1})$,
  4-$\tilde{u}^{k+1} = 2u^{k+1} - u^k$,
**end**

**Algorithm 4:** The Chambolle-Pock algorithm for solving problems (67) and (70)