



# Characterizing the maximum parameter of the total-variation denoising through the pseudo-inverse of the divergence

Charles-Alban Deledalle, Nicolas Papadakis, Joseph Salmon, Samuel Vaïter

## ► To cite this version:

Charles-Alban Deledalle, Nicolas Papadakis, Joseph Salmon, Samuel Vaïter. Characterizing the maximum parameter of the total-variation denoising through the pseudo-inverse of the divergence. 2016.

**HAL Id: hal-01412059**

**<https://hal.archives-ouvertes.fr/hal-01412059>**

Submitted on 7 Dec 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Characterizing the maximum parameter of the total-variation denoising through the pseudo-inverse of the divergence

Charles-Alban Deledalle    Nicolas Papadakis  
 IMB, CNRS, Bordeaux INP  
 Université Bordeaux, Talence, France  
 Email: firstname.lastname@math.u-bordeaux.fr

Joseph Salmon  
 LTCI, CNRS, Télécom ParisTech  
 Université Paris-Saclay, France  
 Email: joseph.salmon@telecom-paristech.fr

Samuel Vaiter  
 IMB, CNRS  
 Université de Bourgogne, Dijon, France  
 Email: samuel.vaiter@u-bourgogne.fr

**Abstract**—We focus on the maximum regularization parameter for anisotropic total-variation denoising. It corresponds to the minimum value of the regularization parameter above which the solution remains constant. While this value is well known for the Lasso, such a critical value has not been investigated in details for the total-variation. Though, it is of importance when tuning the regularization parameter as it allows fixing an upper-bound on the grid for which the optimal parameter is sought. We establish a closed form expression for the one-dimensional case, as well as an upper-bound for the two-dimensional case, that appears reasonably tight in practice. This problem is directly linked to the computation of the pseudo-inverse of the divergence, which can be quickly obtained by performing convolutions in the Fourier domain.

## I. INTRODUCTION

We consider the reconstruction of a  $d$ -dimensional signal (in this study  $d = 1$  or  $2$ ) from its noisy observation  $y = x + w \in \mathbb{R}^n$  with  $w \in \mathbb{R}^n$ . Anisotropic TV regularization writes, for  $\lambda > 0$ , as [1]

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \frac{1}{2} \|y - x\|_2^2 + \lambda \|\nabla x\|_1 \quad (1)$$

with  $\nabla x \in \mathbb{R}^{dn}$  being the concatenation of the  $d$  components of the discrete periodical gradient vector field of  $x$ , and  $\|\nabla x\|_1 = \sum_i |(\nabla x)_i|$  being a sparsity promoting term. The operator  $\nabla$  acts as a convolution which writes in the one dimensional case ( $d = 1$ )

$$\nabla = F^+ \operatorname{diag}(K_{\rightarrow})F \quad \text{and} \quad \operatorname{div} = F^+ \operatorname{diag}(K_{\leftarrow})F \quad (2)$$

where  $\operatorname{div} = -\nabla^\top$  (where  $\top$  denotes the adjoint),  $F : \mathbb{R}^n \mapsto \mathbb{C}^n$  is the Fourier transform,  $F^+ = \operatorname{Re}[F^{-1}]$  is its pseudo-inverse and  $K_{\rightarrow} \in \mathbb{C}^n$  and  $K_{\leftarrow} \in \mathbb{C}^n$  are the Fourier transforms of the kernel functions performing forward and backward finite differences respectively. Similarly, we define in the two dimensional case ( $d = 2$ )

$$\nabla = \begin{pmatrix} F^+ & 0 \\ 0 & F^+ \end{pmatrix} \begin{pmatrix} \operatorname{diag}(K_{\downarrow}) \\ \operatorname{diag}(K_{\rightarrow}) \end{pmatrix} F \quad (3)$$

$$\text{and} \quad \operatorname{div} = F^+ \begin{pmatrix} \operatorname{diag}(K_{\uparrow}) & \operatorname{diag}(K_{\leftarrow}) \\ 0 & F \end{pmatrix} \quad (4)$$

where  $K_{\rightarrow} \in \mathbb{C}^n$  and  $K_{\leftarrow} \in \mathbb{C}^n$  (resp.  $K_{\downarrow} \in \mathbb{C}^n$  and  $K_{\uparrow} \in \mathbb{C}^n$ ) perform forward and backward finite difference in the horizontal (resp. vertical) direction.

## II. GENERAL CASE

For the general case, the following proposition provides an expression of the maximum regularization parameter  $\lambda_{\max}$  as the solution of a convex but non-trivial optimization problem (direct consequence of the Karush-Khun-Tucker condition).

**Proposition 1.** Define for  $y \in \mathbb{R}^n$ ,

$$\lambda_{\max} = \min_{\zeta \in \operatorname{Ker}[\operatorname{div}]} \|\operatorname{div}^+ y + \zeta\|_{\infty} \quad (5)$$

where  $\operatorname{div}^+$  is the Moore-Penrose pseudo-inverse of  $\operatorname{div}$  and  $\operatorname{Ker}[\operatorname{div}]$  its null space. Then,  $x^* = \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\top y$  if and only if  $\lambda \geq \lambda_{\max}$ .

## III. ONE DIMENSIONAL CASE

In the 1d case,  $\operatorname{Ker}[\operatorname{div}] = \operatorname{Span}(\mathbb{1}_n)$  and thus the optimization problem can be solved by computing  $\operatorname{div}^+$  in the Fourier domain, in  $O(n \log n)$  operations, as shown in the next corollary.

**Corollary 1.** For  $d = 1$ ,  $\lambda_{\max} = \frac{1}{2} [\max(\operatorname{div}^+ y) - \min(\operatorname{div}^+ y)]$ ,

$$\text{where} \quad \operatorname{div}^+ = F^+ \operatorname{diag}(K_{\uparrow}^+)F \quad (6)$$

$$\text{and} \quad (K_{\uparrow}^+)_i = \begin{cases} \frac{(K_{\uparrow})_i^*}{|(K_{\uparrow})_i|^2} & \text{if } |(K_{\uparrow})_i|^2 > 0 \\ 0 & \text{otherwise} \end{cases},$$

and  $*$  denotes the complex conjugate.

Note that the condition  $|(K_{\uparrow})_i|^2 > 0$  is satisfied everywhere except for the zero frequency. In the non-periodical case,  $\operatorname{div}$  is the incidence matrix of a tree whose pseudo-inverse can be obtained following [2].

## IV. TWO DIMENSIONAL CASE

In the 2d case,  $\operatorname{Ker}[\operatorname{div}]$  is the orthogonal of the vector space of signals satisfying Kirchhoff's voltage law on all cycles of the periodical grid. Its dimension is  $n+1$ . It follows that the optimization problem becomes much harder. Since our motivation is only to provide an approximation of  $\lambda_{\max}$ , we propose to compute an upper-bound in  $O(n \log n)$  operations thanks to the following corollary.

**Corollary 2.** For  $d = 2$ ,  $\lambda_{\max} \leq \underbrace{\frac{1}{2} [\max(\operatorname{div}^+ y) - \min(\operatorname{div}^+ y)]}_{\lambda_{\text{bnd}}}$ ,

$$\text{where} \quad \operatorname{div}^+ = \begin{pmatrix} F^+ & 0 \\ 0 & F^+ \end{pmatrix} \begin{pmatrix} \operatorname{diag}(\tilde{K}_{\uparrow}^+) \\ \operatorname{diag}(\tilde{K}_{\leftarrow}^+) \end{pmatrix} F, \quad \text{and} \quad (7)$$

$$(\tilde{K}_{\uparrow}^+)_i = \begin{cases} \frac{(K_{\uparrow})_i^*}{|(K_{\uparrow})_i|^2 + |(K_{\leftarrow})_i|^2} & \text{if } |(K_{\uparrow})_i|^2 + |(K_{\leftarrow})_i|^2 > 0 \\ 0 & \text{otherwise} \end{cases},$$

$$(\tilde{K}_{\leftarrow}^+)_i = \begin{cases} \frac{(K_{\leftarrow})_i^*}{|(K_{\uparrow})_i|^2 + |(K_{\leftarrow})_i|^2} & \text{if } |(K_{\uparrow})_i|^2 + |(K_{\leftarrow})_i|^2 > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Note that the condition  $|(K_{\uparrow})_i|^2 + |(K_{\leftarrow})_i|^2 > 0$  is again satisfied everywhere except for the zero frequency. Remark also that this result can be straightforwardly extended to the case where  $d > 2$ .

## V. RESULTS AND DISCUSSION

Figure 1 and 2 provide illustrations of the computation of  $\lambda_{\max}$  and  $\lambda_{\text{bnd}}$  on a 1d signal and a 2d image respectively. The convolution kernel is a simple triangle wave in the 1d case but is more complex in the 2d case. The operator  $\operatorname{div} \operatorname{div}^+$  is in fact the projector onto the space of zero-mean signals, i.e.,  $\operatorname{Im}[\operatorname{div}]$ . Figure 3 illustrates the evolution of  $x^*$  with respect to  $\lambda$  (computed with the algorithm of [3]). Our upper-bound  $\lambda_{\text{bnd}}$  (computed in  $\sim 5$ ms) appears to be reasonably tight ( $\lambda_{\max}$  computed in  $\sim 25$ s with [3] on Problem (5)).

Future work will concern the generalization of these results to other  $\ell_1$  analysis regularization and to ill-posed inverse problems.

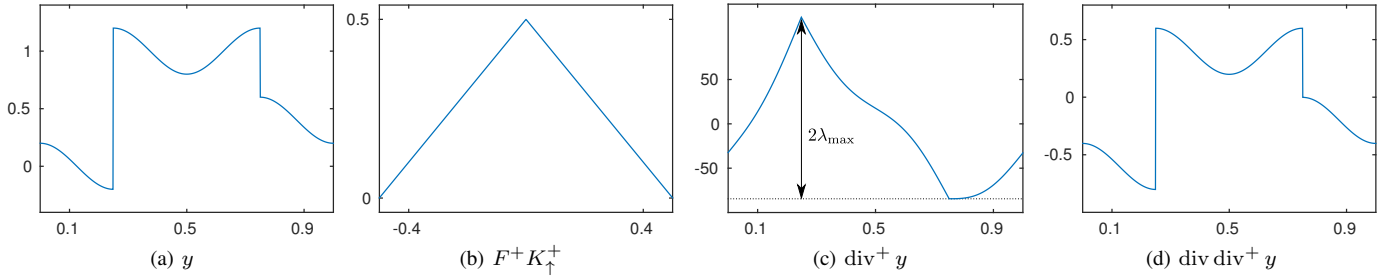


Fig. 1. (a) A 1d signal  $y$ . (b) The convolution kernel  $F^+K_\uparrow^+$  that realizes the pseudo inversion of the divergence. (c) The signal  $\text{div}^+ y$  on which we can read the value of  $\lambda_{\max}$ . (d) The signal  $\text{div div}^+ y$  showing that one can reconstruct  $y$  from  $\text{div}^+ y$  up to its mean component.

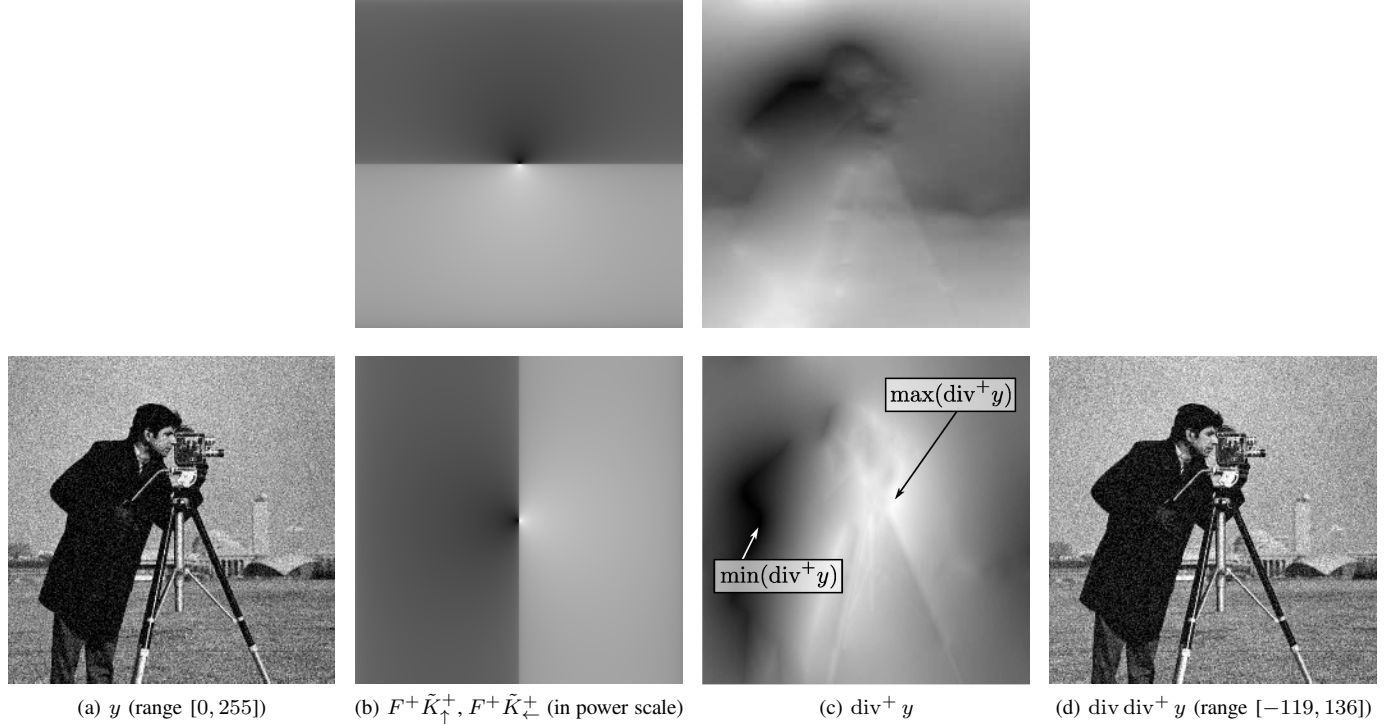


Fig. 2. (a) A 2d signal  $y$ . (b) The convolution kernels  $F^+K_\uparrow^+$  and  $F^+K_\pm^+$  that realizes the pseudo inversion of the divergence. (c) The absolute value of the two coordinates of the vector field  $\text{div}^+ y$  on which we can read the upper-bound  $\lambda_{\text{bnd}}$  of  $\lambda_{\max}$ . (d) The image  $\text{div div}^+ y$  showing again that one can reconstruct  $y$  from  $\text{div}^+ y$  up to its mean component.

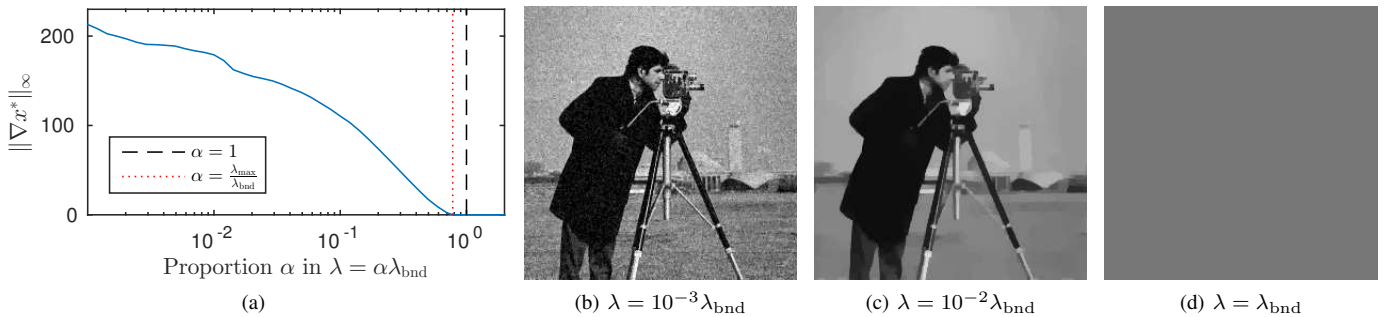


Fig. 3. (a) Evolution of  $\|\nabla x^*\|_\infty$  as a function of  $\lambda$ . (b), (c), (d) Results  $x^*$  of the periodical anisotropic total-variation for three different values of  $\lambda$ .

REFERENCES

[1] 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.

[2] R. Bapat, "Moore-penrose inverse of the incidence matrix of a tree," *Linear and Multilinear Algebra*, vol. 42, no. 2, pp. 159–167, 1997.

[3] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vis.*, vol. 40, pp. 120–145, 2011.