



**HAL**  
open science

## An Anisotropic Inf-Convolution BV type model for dynamic reconstruction.

Maitine Bergounioux, E Papoutsellis

► **To cite this version:**

Maitine Bergounioux, E Papoutsellis. An Anisotropic Inf-Convolution BV type model for dynamic reconstruction. . SIAM Journal on Imaging Sciences, 2018. hal-01401408v3

**HAL Id: hal-01401408**

**<https://hal.science/hal-01401408v3>**

Submitted on 6 Sep 2017

**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.

# An Anisotropic Inf-Convolution BV type model for dynamic reconstruction.\*

Maitine Bergounioux<sup>†</sup> and E. Papoutsellis<sup>†</sup>

**Abstract.** We are interested in a spatial-temporal variational model for image sequences. The model involves a fitting data term adapted to different modalities such as denoising, deblurring or emission tomography. The regularizing term acts as an infimal-convolution type operator that takes into account the respective influence of space and time variables in a separate mode. We give existence and uniqueness results and provide optimality conditions via duality analysis.

**Key words.** Spatio-temporal Variational Regularization, Infimal Convolution Total Variation, Anisotropic Total variation, Optimality Conditions

**AMS subject classifications.** 65D18, 68U10,65K10

**1. Introduction.** In this paper, we examine variational inverse problems for dynamic image reconstruction. As in the context of image restoration, the goal regarding a video restoration is to recover a *clean* image sequence given a degraded dynamic datum. Certainly, one of the main differences between image and video restoration is the additional temporal domain where a collection of images-frames evolves over the time. Besides the spatial structures which are a significant factor on the output quality of the reconstruction, the time direction has an important role on the temporal consistency among the frames. Furthermore, in terms of video applications, one may consider applications inherited from the imaging context and extend them to the dynamical framework. To name a few, we have dynamic denoising, deblurring, inpainting, decompression and emission tomography such as Positron Emission Tomography and Magnetic Resonance Imaging.

The aim of this paper is to study variational regularization models in an infinite dimensional setting defined on a spatial-temporal domain. In particular, given a corrupted image sequence  $g$ , we look for a solution  $u$ , in a Banach space  $\mathcal{X}$ , to the following generic minimization problem

$$(1.1) \quad \inf_{u \in \mathcal{X}} \mathcal{H}(\mathcal{A}u, g) + \mathcal{N}(u).$$

The first and second terms represent the well known data fitting term (*fidelity*) and the *regularizer* respectively. The former is determined by the nature of degradation, e.g., a transformation through a continuous and linear operator  $\mathcal{A}$  with the presence of random noise, as well as the modality of the problem. The latter imposes a certain prior structure (regularity) on the solution  $u$ . Regarding image restoration, the minimization problem (1.1) has been extensively used and examined from both theoretical and numerical point of view for different applications. For instance, we refer the reader to the famous ROF variational model [36], where the use of functions of bounded variation (BV) and the total variation regularization (TV) was established

---

\*Submitted to the editors DATE.

**Funding:** This work was funded by by the french Région Centre (Project OTITE4D)

<sup>†</sup>Université d'Orléans , Fédération Poisson, Math., Labo. MAPMO, UMR 7349, Route de Chartres, BP 6759, 45067 Orléans cedex 2 ([maitine.bergounioux@univ-orleans.fr](mailto:maitine.bergounioux@univ-orleans.fr),[evangelos.papoutsellis@univ-orleans.fr](mailto:evangelos.papoutsellis@univ-orleans.fr))

35 in image processing. Moreover, it was analyzed in [1], [44] and several extensions have been  
 36 proposed in [15, 8, 12, 16, 24]. Now, concerning variational problems on a spatial-temporal  
 37 domain, one can witness significantly less work from a theoretical perspective compared to  
 38 a numerical one. Indeed, there is a plethora of numerical algorithms in the literature for  
 39 variational video processing. We refer the reader to some of them as [17, 38, 32].

40 A quite natural approach towards image sequence reconstruction is to apply the minimiza-  
 41 tion problem (1.1), acting on every image-frame of the sequence individually. For example, we  
 42 use the above problem in order to denoise each frame from a sequence corrupted by Gaussian  
 43 noise. We choose a non-smooth regularizer as the *total variation measure* over the spatial  
 44 domain  $\Omega \subset \mathbb{R}^2$ . It is known for the piecewise constant structures imposed to the solution  $u$   
 45 that can eliminate efficiently the noise while preserving the edges of the images. It is defined  
 46 as

$$47 \quad (1.2) \quad \mathcal{N}(u) = \alpha \text{TV}_x(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, dx : \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2), \|\varphi\|_{\infty} \leq \alpha \right\},$$

48 weighted by a positive parameter  $\alpha$  and

$$49 \quad (1.3) \quad \|\varphi\|_{\infty} = \operatorname{ess\,sup}_{x \in \Omega} |\varphi(x)|_r, \quad |\varphi(x)|_r = \begin{cases} \sqrt{\varphi_1^2(x) + \varphi_2^2(x)}, & r = 2, \text{ (isotropic)} \\ \max\{|\varphi_1(x)|, |\varphi_2(x)|\}, & r = \infty, \text{ (anisotropic)}. \end{cases}$$

50 This parameter is responsible for a proper balancing between the regularizer and the fidelity  
 51 term which is fixed as  $\mathcal{H}(u, g) = \frac{1}{2} \|u - g\|_{L^2(\Omega)}^2$  in this case. Although, this solution produces  
 52 a satisfying result on the spatial domain, it does not take into account the interaction between  
 53 time and space and some time artifacts, e.g. *flickering*, will be introduced. Note that one  
 54 can use the anisotropic norm instead of an isotropic one in (1.3). Although these norms are  
 55 equivalent in a finite dimensional setting, they have different effects on the corresponding  
 56 computed minimizers. In the isotropic case, sharp corners will not be allowed in the edge set  
 57 and smooth corners prevail. On the other hand, corners in the direction of the unit vectors  
 58 are favored in the anisotropic variant. For more details, we refer the reader to [29, 21, 34] on  
 59 the properties and differences between these two corresponding minimizers.

60 A more sophisticated path, referred as 3D denoising, is to extend the domain taking into  
 61 account the time activity and treat an image sequence as a 3D volume where the time plays  
 62 the role of the third variable. In this case, we write

$$63 \quad (1.4) \quad \mathcal{N}(u) = \text{TV}_{(t,x)}^{\alpha}(u) = \sup \left\{ \int_Q u \operatorname{div}_{\alpha} \varphi \, dx \, dt : \varphi \in \mathcal{C}_c^1(Q, \mathbb{R}^3), \|\varphi\|_{\infty} \leq 1 \right\}$$

where  $Q = \mathcal{T} \times \Omega \subset \mathbb{R}^3$  is the three-dimensional spatial-temporal domain with  $\mathcal{T} = (0, T)$ .  
 Here, we have a positive vector  $\alpha = (\alpha_1, \alpha_2)$  acting on the space and time respectively with

$$\operatorname{div}_{\alpha} = \alpha_1 \left( \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} \right) + \alpha_2 \frac{\partial}{\partial t} = \alpha_1 \operatorname{div}_x + \alpha_2 \operatorname{div}_t$$

64 and the TV smoothness is applied along both the spatial and the temporal directions. An  
 65 obvious question that rises on this particularly setting is the correlation between the space

66 and time. Video regularization approaches as in [17, 25, 30] combine spatial and temporal  
 67 domains under the corresponding dynamic isotropic norm  $\|\varphi\|_\infty = \operatorname{ess\,sup}_{(t,x) \in Q} |\varphi(t,x)|_2$ . Hence,  
 68 space and time are interacting with each other and contribute under some weight parameters  
 69 to the TV regularizer.

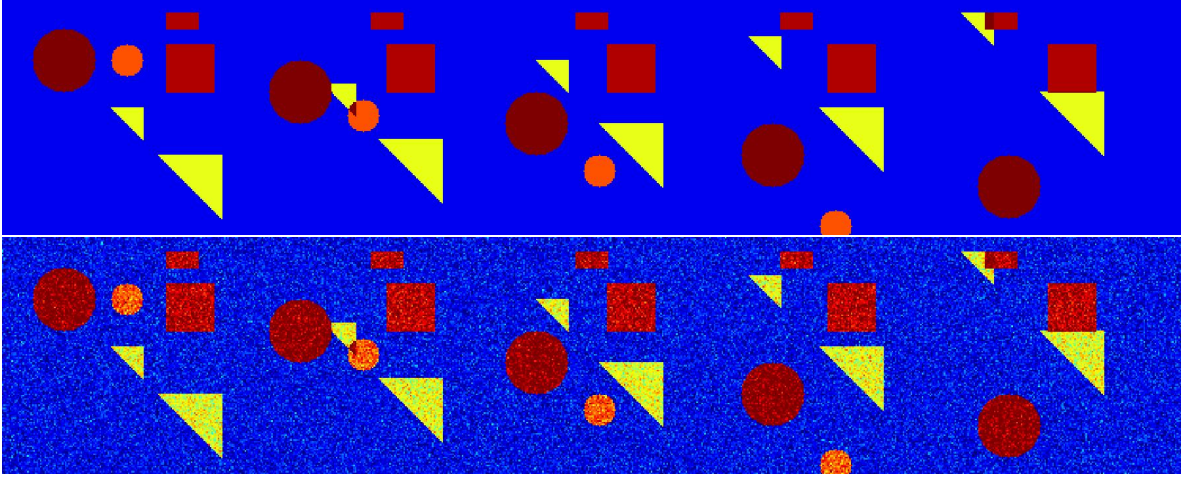


Figure 1.1: Image sequence of 5 frames and its noisy version corrupted with Gaussian noise. Geometrical shapes are moving in different directions with different moving speed.

70 However, this choice of norm is not very accurate concerning the preservation of spatial  
 71 and temporal discontinuities. Using the anisotropic norm,  $\|\varphi\|_\infty = \operatorname{ess\,sup}_{(t,x) \in Q} |\varphi(t,x)|_\infty$ , where  
 72 space and time are not correlated, has the advantage to focus on the discontinuities of  $\Omega$  and  
 73  $\mathcal{T}$  in separate modes respectively and preserves spatial and temporal details more accurately.  
 74 In particular, we can decompose (1.4) into a spatial and a temporal total variation, see [2],  
 75 and write

$$\begin{aligned}
 \operatorname{TV}_{(t,x)}^\alpha(u) &= \mathcal{TV}_x^{\alpha_1}(u) + \mathcal{TV}_t^{\alpha_2}(u), \text{ with} \\
 \mathcal{TV}_x^{\alpha_1}(u) &= \sup \left\{ \int_Q u \alpha_1 \left( \frac{\partial \varphi_1}{\partial x_1} + \frac{\partial \varphi_2}{\partial x_2} \right) dx dt : \varphi \in \mathcal{C}_c^1(Q, \mathbb{R}^3), \right. \\
 76 \quad (1.5) \quad &\quad \left. \max\{\sqrt{\varphi_1^2(t,x) + \varphi_2^2(t,x)}\} \leq 1 \right\}, \\
 \mathcal{TV}_t^{\alpha_2}(u) &= \sup \left\{ \int_Q u \alpha_2 \frac{\partial \varphi_3}{\partial t} dx dt : \varphi \in \mathcal{C}_c^1(Q, \mathbb{R}^3), \max\{|\varphi_3(t,x)|\} \leq 1 \right\}.
 \end{aligned}$$

77 This type of decomposition has already been proposed for several applications such as dy-  
 78 namic denoising, segmentation, video decompression and the reader is referred to [43, 38, 26,  
 79 17]. Although, this paper is rather theoretical we would like to intrigue the reader with a  
 80 simple numerical example. In Figure 1.1, we have an image sequence of 5 frames of sev-  
 81 eral geometrical objects moving in different directions and speed under a constant back-  
 82 ground. This is corrupted by Gaussian noise. In order to compare between isotropic (1.4)

83 and anisotropic (1.5) total variation spatial-temporal regularization we select the parameters  
 84 of the isotropic/anisotropic TV such that in both cases they will have the same distance for  
 85 the ground truth (as in PSNR ) namely  $\|\text{solution}_{ISTV} - \text{truth}\|_2 = \|\text{solution}_{ANTV} - \text{truth}\|_2$ .

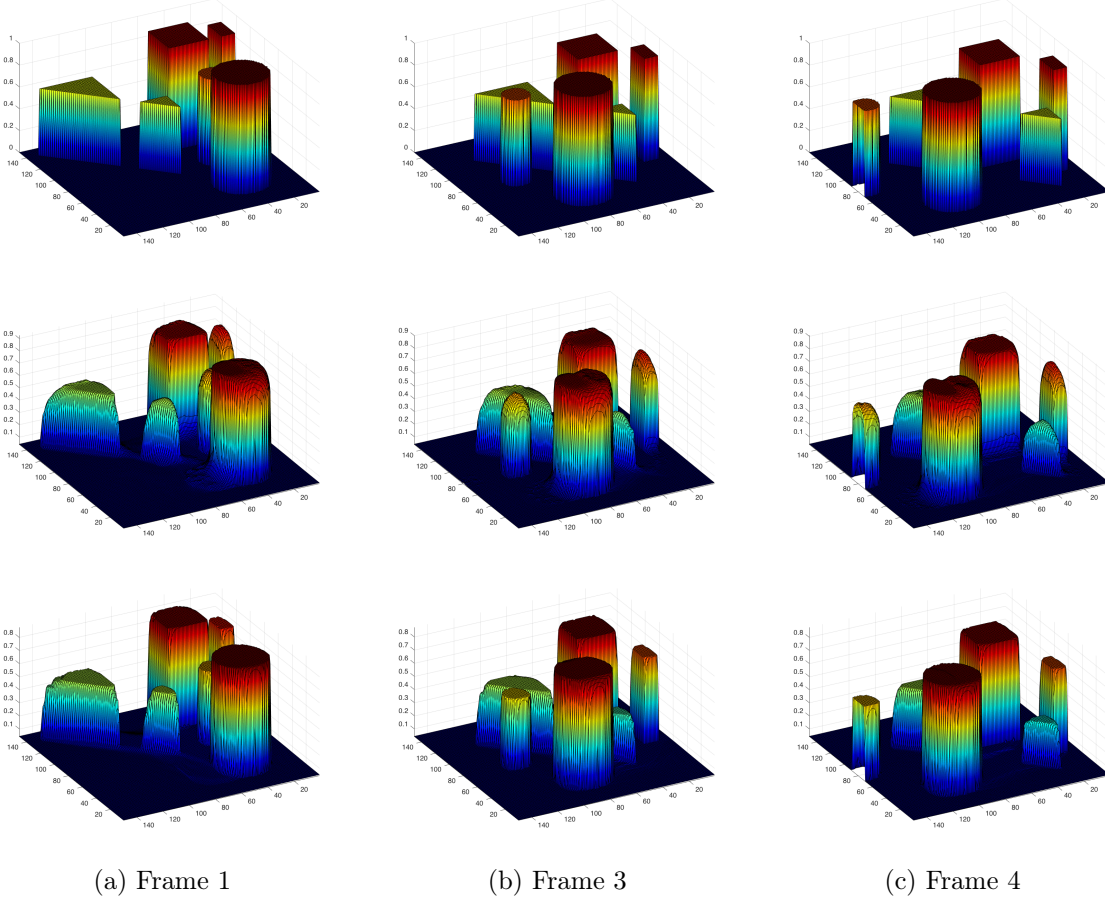


Figure 1.2: First row: True sequence, second row: Isotropic TV, third row: Anisotropic TV. We present frames 1, 3 and 5. The parameters are optimized such that  $\|\text{solution}_{ISTV} - \text{truth}\|_2 = \|\text{solution}_{ANTV} - \text{truth}\|_2 = 25.9559$  with  $\alpha_1^{ISTV} = \alpha_1^{ANTV} = 0.5$ ,  $\alpha_2^{ISTV} = 0.05$ ,  $\alpha_2^{ANTV} = 0.0501$ .

86 In Figure 1.2, we present the surface plots of three of the five frames of the correspond-  
 87 ing regularized solutions of (1.1) with the squared  $L^2$  norm fidelity term. We observe that  
 88 anisotropic regularization is able to preserve the geometry of these objects.

89 Motivated by (1.5), we proceed with a further decomposition in terms of the test function  
 90  $\varphi$  and define the following *decoupled* spatial-temporal total variation regularization,

$$91 \quad (1.6) \quad \mathcal{N}(u) = \alpha_1 \int_0^T \text{TV}_x(u(t, \cdot)) dt + \alpha_2 \int_{\Omega} \text{TV}_t(u(\cdot, x)) dx,$$

92 where  $\text{TV}_x$  is given by (1.2) and  $\text{TV}_t(u)$  is defined similarly (see (2.2)). They denote  
 93 the spatial total variation for every  $t \in \mathcal{T}$  and the temporal total variation for every  $x \in \Omega$   
 94 respectively. Note that in the above formulations the test functions are defined in  $\Omega$  and  $\mathcal{T}$   
 95 respectively.

96 Non-smooth regularization methods introduce different kind of modelling artifacts. As we  
 97 discussed above, a total variation regularizer tends to approximate non-constant noisy regions  
 98 with piecewise constant structures leading to the *staircasing* effect. This aspect is certainly  
 99 inherited in the dynamic framework and produces the *flickering* effect due to the staircasing  
 100 along the temporal dimension. In addition, one may observe some *ghost* artifacts on moving  
 101 objects, i.e., where certain features are overlapping between two consecutive frames. This is  
 102 due to the strong temporal regularization, namely when the ratio  $\frac{\alpha_1}{\alpha_2}$  is relatively small. In  
 103 order to overcome this kind of modelling artifacts, a combination of non-smooth regularizers  
 104 is used via the concept of the infimal convolution,

$$105 \quad (1.7) \quad \mathcal{N}(u) = F_1 \# F_2(u) = \inf_{v \in \mathcal{X}} F_1(u - v) + F_2(v).$$

106 This regularization functional is able to favor reconstructions with a relatively small  $F_1$  or  
 107  $F_2$  contribution. In the imaging context, this is introduced in [15], where a first and second  
 108 order TV-based regularizers are combined in order to reduce the staircasing phenomenon.  
 109 Under this regularizer, the corresponding solution  $u$  of (1.1) promotes both piecewise constant  
 110 and smooth structures due to the presence of higher order derivatives and in fact provides a  
 111 certain decomposition between piecewise constant and smooth regions. On the other hand,  
 112 Holler and Kunisch in [25], extend the notion of infimal convolution in the context of dynamic  
 113 processing. In such a setting, they propose the use of total variation functionals as in (1.4) with  
 114 an isotropic relation on the spatial and temporal regularities. As in the imaging framework,  
 115 one can decompose an image sequence into a sequence that captures spatial information and  
 116 a sequence that encodes temporal activity. This type of spatial-temporal regularizer will be  
 117 discussed in Section 3 under the anisotropic formulation (1.6) of separate action in space and  
 118 time. Specifically, we propose the following infimal convolution total variation regularization  
 119 for an image sequence  $u$ . Given two positive vectors  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$  and  $\boldsymbol{\mu} = (\mu_1, \mu_2)$ ,

$$120 \quad (1.8) \quad \mathcal{N}(u) = F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u) = \inf_{v \in \mathcal{X}} \int_0^T \lambda_1 \text{TV}_x(u - v)(t) dt + \int_{\Omega} \lambda_2 \text{TV}_t(u - v)(x) dx \\ + \int_0^T \mu_1 \text{TV}_x(v)(t) dt + \int_{\Omega} \mu_2 \text{TV}_t(v)(x) dx.$$

121 Depending on the choice of  $\boldsymbol{\lambda}$ ,  $\boldsymbol{\mu}$  one can enforce a certain regularity and either focus on space  
 122 or on time for the image sequences  $u - v$  and  $v$ . For example, if one selects that  $\lambda_1 = \mu_2 = \kappa$   
 123 and  $\lambda_2 = \mu_1 = 1$  with  $\kappa > 1$  then the first two terms impose a TV smoothness more on the  
 124 space direction that in time for the  $u - v$  term. For the other two terms, the TV smoothness  
 125 acts conversely for the  $v$  component. Therefore, it is a matter of proper balancing which is  
 126 tuned automatically via the infimal convolution and highlights the cost either on space or time.  
 127 The choice of parameters will be discussed in Section 3. We would like to mention that the  
 128 functionals in (1.7) are not necessarily total variational functionals and other combinations or  
 129 high order functionals may be used, see for instance [39, 7].

130 Finally, we would like to emphasize on the nature of the positive parameters defined above.  
 131 In the definitions (1.4), (1.5) and (1.8), we use parameters that are constant over the time do-  
 132 main. Equivalently, every *frame* is penalized with the same constant. This is a fair assumption  
 133 when the level of noise is assumed to be constant over time. However, in real world appli-  
 134 cations this is not always the case. There are situations when the noise is signal-dependent  
 135 e.g., Poisson noise and the noise-level variates over time. In the dynamic PET imaging and in  
 136 particular in *list-mode* PET, see [42], data can be binned into sinograms allowing frame dura-  
 137 tions to be determined after the acquisition. Under this approach, one has to choose between  
 138 longer scans with good counting statistics and shorter scans that are noisy but preserving  
 139 temporal resolution. A usual and fair choice is to select shorter scans in the beginning where  
 140 there is a high activity of the radioactive tracer and longer scans at the end. For example,  
 141 a 50 minutes acquisition in list mode rat-brain scans is rebinned into 27 frames under the  
 142 following scheme: 4x10s, 4x20s, 4x60s, 14x180s, 1x120s, see [40]. Hence, our goal is to allow  
 143 time dependent parameters on the above regularizers that can handle not only different levels  
 144 of noise *per frame* (1st term) but also balance the temporal activity in terms of a non-uniform  
 145 time domain discretization (2nd term), i.e.,

$$146 \quad (1.9) \quad \mathcal{N}(u) = \int_0^T \alpha_1(t) \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(\alpha_2(t)u)(x) dx.$$

147 Outline of the paper: The paper is organized as follows: we first recall some general properties  
 148 of functions of bounded variation and fix the notations in terms of the dynamic framework. We  
 149 continue with the definition of the regularizers used in this paper such as a *weighted* version  
 150 of the spatial-temporal total variation as well as its extension to the infimal convolution. In  
 151 addition, we define also the data fitting terms that are suitable for different applications. In  
 152 Section 4, we examine the well-posedness (existence, uniqueness and stability) of the associated  
 153 variation problem specifically for the infimal convolution regularizer and conclude in Section  
 154 5, with the corresponding optimality conditions. Finally, we would like to mention that the  
 155 nature of this paper is rather theoretical and we do not address any numerical issues. This  
 156 will be done in a forthcoming paper.

157 **2. Preliminaries.** Let us denote  $u : \mathcal{T} \times \Omega \rightarrow \mathbb{R}$ , an image sequence defined on an open  
 158 bounded set  $\Omega \subset \mathbb{R}^d$  with smooth boundary representing the space domain with  $d \geq 1$  and  
 159  $\mathcal{T} = (0, T)$ ,  $T > 0$  which represents the temporal domain. In this section, we recall some basic  
 160 notations related to functions of bounded variation (BV) extended to the spatial-temporal  
 161 context. In order to distinguish between spatial and temporal domains, we define the following  
 162 spaces

$$163 \quad (2.1) \quad \begin{aligned} L^1(\mathcal{T}; \text{BV}(\Omega)) &= \{u : \mathcal{T} \times \Omega \rightarrow \mathbb{R} \mid u(t, \cdot) \in \text{BV}(\Omega) \text{ a.e. } t \in \mathcal{T} \\ &\quad \text{and } t \mapsto \text{TV}_x(u)(t) \in L^1(\mathcal{T})\}, \\ L^1(\Omega; \text{BV}(\mathcal{T})) &= \{u : \mathcal{T} \times \Omega \rightarrow \mathbb{R} \mid u(\cdot, x) \in \text{BV}(\mathcal{T}) \text{ a.e. } x \in \Omega \\ &\quad \text{and } x \mapsto \text{TV}_t(u)(x) \in L^1(\Omega) \}. \end{aligned}$$

164 Here,  $\text{TV}_x$  and  $\text{TV}_t$  stand for the spatial and temporal total variation for every  $t \in \mathcal{T}$  and  
 165  $x \in \Omega$  respectively. In particular, we have

$$166 \quad (2.2) \quad \begin{aligned} \text{TV}_x(u)(t) &= \sup \left\{ \int_{\Omega} \xi(x)u(t, x) \, dx \mid \xi \in K_x \right\}, \\ \text{TV}_t(u)(x) &= \sup \left\{ \int_0^T \xi(t)u(t, x) \, dt \mid \xi \in K_t \right\}, \end{aligned}$$

167 with the corresponding sets

$$168 \quad (2.3) \quad \begin{aligned} K_x &:= \left\{ \xi = \text{div}_x(\varphi) \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^d), \|\varphi\|_{\infty, x} \leq 1 \right\}, \|\varphi\|_{\infty, x} = \text{ess sup}_{x \in \Omega} |\varphi(x)|_2 \\ K_t &:= \left\{ \xi = \frac{d\varphi}{dt} \mid \varphi \in \mathcal{C}_c^1(\mathcal{T}, \mathbb{R}), \|\varphi\|_{\infty, t} \leq 1 \right\}, \|\varphi\|_{\infty, t} = \text{ess sup}_{t \in \mathcal{T}} |\varphi(t)| \end{aligned}$$

169 where  $\text{div}_x$  is the divergence operator on the spatial domain and  $|\cdot|_2$  is the isotropic-euclidean  
 170 norm in space. Finally, we define the space of functions of bounded variation on the spatial-  
 171 temporal domain  $Q$ , acting isotropically in these two directions i.e.,

$$172 \quad (2.4) \quad \begin{aligned} \text{BV}(Q) &= \{u \in L^1(Q) \mid \text{TV}(u) < \infty\}, \text{ where} \\ \text{TV}_{(t,x)}(u) &= \sup \left\{ \int_Q \xi(t, x)u(t, x) \, dx \, dt \mid \xi \in K \right\}, \text{ and} \\ K &:= \left\{ \xi = \text{div}_{(t,x)}(\varphi) \mid \varphi \in \mathcal{C}_c^1(Q, \mathbb{R} \times \mathbb{R}^d), \|\varphi\|_{\infty} \leq 1 \right\}, \\ &\quad \|\varphi\|_{\infty} = \text{ess sup}_{(t,x) \in Q} |\varphi(t, x)|_2. \end{aligned}$$

173 In sequel we drop the index  $(t, x)$  in the total variation on  $Q$  notation so that  $\text{TV}$  stands  
 174 for  $\text{TV}_{(t,x)}$ . Note that  $\text{div}_{(t,x)} = \frac{\partial}{\partial t} + \text{div}_x$ . As we pointed out in the introduction, one may  
 175 consider an equivalent anisotropic norm using for any  $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_d) \in \mathcal{C}_c^1(Q, \mathbb{R} \times \mathbb{R}^d)$  :  
 176  $|\varphi(t, x)|_{\infty} = \max \left\{ \sqrt{\sum_{i=1}^d \varphi_i^2(t, x)}, |\varphi_0(t, x)| \right\} \leq 1$  and all the following results are still true.  
 177 In the following theorem, see [3, 5], we recall some useful properties on the  $\text{BV}(O)$  space, where  
 178  $O$  is a bounded, open set of  $\mathbb{R}^N$  (practically  $O = \Omega$  with  $N = d$  or  $O = Q$  with  $N = d + 1$ .)

**Theorem 2.1.** *Let  $O \subset \mathbb{R}^N$ ,  $N \geq 1$ . The space  $\text{BV}(O)$  endowed with the norm*

$$\|v\|_{\text{BV}(O)} := \|v\|_{L^1(O)} + \text{TV}(v)$$

*is a Banach space.*

(a) *For any  $u \in \text{BV}(O)$  there exists a sequence  $u_n \in \mathcal{C}^{\infty}(\bar{O})$  such that*

$$u_n \rightarrow u \text{ in } L^1(O) \text{ and } \text{TV}(u_n) \rightarrow \text{TV}(u).$$

(b) *The mapping  $u \mapsto \text{TV}(u)$  is lower semicontinuous from  $\text{BV}(O)$  endowed with the  $L^1(O)$  topology to  $\mathbb{R}^+$ .*



(c)  $BV(O) \subset L^p(O)$  with continuous embedding, for  $1 \leq p \leq \frac{N}{N-1}$  and we have the Poincaré-Wirtinger inequality (Remark 3.50 of [3]): there exists a constant  $C_O$  only depending on  $O$  such that for  $1 \leq p \leq \frac{N}{N-1}$

$$\forall u \in BV(O), \quad \|u - \bar{u}\|_{L^p(O)} \leq C_O \text{TV}(u),$$

179 where  $\bar{u}$  is the mean value of  $u$  on  $O$ .

180 (d)  $BV(O) \subset L^p(O)$  with compact embedding for  $1 \leq p < \frac{N}{N-1}$ .

181 The lemma below is essential for the forthcoming analysis and relates the spaces defined by  
182 (2.1) and (2.4). It is based on the definitions above as well as of some tools in the proof of  
183 [22, Theorem 2, Section 5.10.2]. A similar result (but in a different context) can be found in  
184 [9, Lemma 3].

185 **Lemma 2.1.** We have  $L^1(\mathcal{T}; BV(\Omega)) \cap L^1(\Omega; BV(\mathcal{T})) = BV(Q)$ . Moreover, for every  $u \in$   
186  $BV(Q)$

$$187 \quad (2.5) \quad \text{TV}(u) \leq \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \leq \sqrt{2} \text{TV}(u).$$

*Proof.* We start with the first inclusion,

$$L^1(\mathcal{T}; BV(\Omega)) \cap L^1(\Omega; BV(\mathcal{T})) \subset BV(Q).$$

Let be  $u \in L^1(\mathcal{T}; BV(\Omega)) \cap L^1(\Omega; BV(\mathcal{T}))$ . For any  $\xi \in K$  there exists  $\varphi = (\varphi_1, \varphi_2) \in \mathcal{C}_c^1(Q, \mathbb{R} \times \mathbb{R}^d)$  such that  $\|\varphi\|_{\infty} \leq 1$  and

$$\xi = \frac{\partial \varphi_1}{\partial t} + \text{div}_x \varphi_2 := \xi_1 + \xi_2$$

For every  $t \in \mathcal{T}$ ,  $\xi_2(t, \cdot) : x \mapsto \xi_2(t, x)$  belongs to  $K_x$  so that

$$\int_{\Omega} \xi_2(t, x) u(t, x) dx \leq \text{TV}_x(u)(t), \quad \text{a.e. } t \in \mathcal{T},$$

and

$$\int_0^T \int_{\Omega} \xi_2(t, x) u(t, x) dx dt \leq \int_0^T \text{TV}_x(u)(t) dt.$$

Similarly,

$$\int_{\Omega} \int_0^T \xi_1(t, x) u(t, x) dt dx \leq \int_{\Omega} \text{TV}_t(u)(x) dx.$$

188 Then, for every  $\xi \in K$ ,

$$189 \quad \begin{aligned} \int_Q \xi(t, x) u(t, x) dt dx &= \int_0^T \int_{\Omega} \xi_2(t, x) u(t, x) dx dt + \int_{\Omega} \int_0^T \xi_1(t, x) u(t, x) dt dx \\ &\leq \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx. \end{aligned}$$

The right hand side is finite independently of  $\xi$  since  $u \in L^1(\mathcal{T}; \text{BV}(\Omega)) \cap L^1(\Omega; \text{BV}(\mathcal{T}))$ . Therefore,  $u \in \text{BV}(Q)$  and

$$\text{TV}(u) \leq \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx .$$

Let us prove the converse inclusion. We first assume that  $u \in W^{1,1}(Q)$ . Then, using Fubini's theorem we get  $t \mapsto \int_{\Omega} |\nabla_{t,x} u|(t, x) dx \in L^1(\mathcal{T})$  and  $x \mapsto \int_0^T |\nabla_{t,x} u|(t, x) dt \in L^1(\Omega)$ . Here, we write  $|\nabla_{t,x} u|_2 = \sqrt{\left(\frac{\partial u}{\partial t}\right)^2 + \sum_{i=1}^d \left(\frac{\partial u}{\partial x_i}\right)^2}$  and

$$|\nabla_{t,x} u(t, x)|_2 \leq |\nabla_x u(t, x)|_2 + |\nabla_t u(t, x)| \leq \sqrt{2} |\nabla_{t,x} u(t, x)|_2.$$

190 Therefore,  $t \mapsto \int_{\Omega} |\nabla_x u(t, x)|_2 dx \in L^1(\mathcal{T})$ ,  $x \mapsto \int_0^T |\nabla_t u(t, x)| dt \in L^1(\Omega)$  and  
191  $u \in L^1(\mathcal{T}; \text{BV}(\Omega)) \cap L^1(\Omega; \text{BV}(\mathcal{T}))$  with

192 (2.6) 
$$\text{TV}(u) \leq \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \leq \sqrt{2} \text{TV}(u).$$
  
193

194 We now consider  $u \in \text{BV}(Q)$  and show that  $u \in L^1(\mathcal{T}; \text{BV}(\Omega))$ . As  $W^{1,1}(Q)$  is dense in  
195  $\text{BV}(Q)$  in the sense of the intermediate convergence [5], there exists a sequence of functions  
196  $u_k \in W^{1,1}(Q)$  such that  $u_k$  converges to  $u$  in  $L^1(Q)$  and  $\text{TV}(u_k) \rightarrow \text{TV}(u)$ . From Fubini's  
197 theorem, we infer that  $u_k(t, \cdot)$  converges to  $u(t, \cdot)$  in  $L^1(\Omega)$ , for almost every  $t \in \mathcal{T}$  and  $u_k(\cdot, x)$   
198 converges to  $u(\cdot, x)$  in  $L^1(\mathcal{T})$ , for almost every  $x \in \Omega$ . Moreover,  $\text{TV}(u_k) \rightarrow \text{TV}(u)$  is bounded.  
199 Using (2.6) and Fatou's Lemma we have that

200 (2.7) 
$$\begin{aligned} & \int_0^T \liminf_{k \rightarrow \infty} \text{TV}_x(u_k)(t) dt + \int_{\Omega} \liminf_{k \rightarrow \infty} \text{TV}_t(u_k)(x) dx \\ & \leq \liminf_{k \rightarrow \infty} \left( \int_0^T \text{TV}_x(u_k)(t) dt + \int_{\Omega} \text{TV}_t(u_k)(x) dx \right) \leq \sqrt{2} \text{TV}(u). \end{aligned}$$

Then,  $\liminf_{k \rightarrow \infty} \text{TV}_x(u_k)(t) < \infty$ , a.e  $t \in \mathcal{T}$  and  $\liminf_{k \rightarrow \infty} \text{TV}_t(u_k)(x) < \infty$ , a.e  $x \in \Omega$ . Now, for a.e.  $t \in \mathcal{T}$ , we have that

$$\forall \xi \in K_x, \quad \int_{\Omega} u_k(t, x) \xi(x) dx \leq \text{TV}_x(u_k)(t).$$

Hence,

$$\int_{\Omega} u(t, x) \xi(x) dx = \lim_{k \rightarrow +\infty} \int_{\Omega} u_k(t, x) \xi(x) dx \leq \liminf_{k \rightarrow \infty} \text{TV}_x(u_k)(t) < \infty,$$

and

$$\text{TV}_x(u)(t) = \sup_{\xi \in K_x} \int_{\Omega} u(t, x) \xi(x) dx \leq \liminf_{k \rightarrow \infty} \text{TV}_x(u_k)(t) < \infty.$$

This means  $u(t, \cdot) \in \text{BV}(\Omega)$  a.e.  $t \in \mathcal{T}$ . In a similar way, we have that  $u(\cdot, x) \in \text{BV}(\mathcal{T})$  a.e.  $x \in \Omega$ , since

$$\text{TV}_t(u)(x) = \sup_{\xi \in K_t} \int_0^T u(t, x) \xi(t) dt \leq \liminf_{k \rightarrow \infty} \text{TV}_t(u_k)(x) < \infty.$$

201 Finally, using (2.7), we get

$$\begin{aligned} & \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \\ 202 & \leq \int_0^T \liminf_{k \rightarrow \infty} \text{TV}_x(u_k)(t) dt + \int_{\Omega} \liminf_{k \rightarrow \infty} \text{TV}_t(u_k)(x) dx \leq \sqrt{2} \text{TV}(u). \end{aligned}$$

203 This ends the proof, and the inequality (2.6) is also valid for every  $u \in \text{BV}(Q)$ .  $\square$

204 **Remark 2.1.** Note that equation (2.6) depends on the choice of the  $\mathbb{R}^2$ -norm that appears in  
205 the definition of the total variation. If we choose another (equivalent)  $\mathbb{R}^2$ -norm, (2.6) remains  
206 valid with a different constant (instead of  $\sqrt{2}$ ). This does not change the theoretical analysis.  
207 However, the choice of the norm is an important numerical issue as we have pointed it out in  
208 the introduction.

209 **Remark 2.2.** The second inclusion of the previous lemma can be seen as a generalization  
210 of a function of bounded variation “in the sense of Tonelli” denoted by *TBV*, see [18, 4]. For  
211 instance, a function of two variables  $h(x, y)$  is *TBV* on a rectangle  $[a, b] \times [c, d]$  if and only if  
212  $\text{TV}_x h(\cdot, y) < \infty$  for a.e.  $y \in [c, d]$ ,  $\text{TV}_y h(x, \cdot) < \infty$  for a.e.  $x \in [a, b]$  and  $\text{TV}_x h(\cdot, y) \in L^1([a, b])$ ,  
213  $\text{TV}_y h(x, \cdot) \in L^1([c, d])$ .

214 **3. The variational model.** As already mentioned in the introduction we are interested in  
215 the following variational problem

$$216 \quad (3.1) \quad \inf_{u \in \mathcal{X}} \mathcal{H}(g, \mathcal{A}u) + \mathcal{N}(u),$$

217 where  $\mathcal{X} = \text{BV}(Q)$ . In this section, we describe the choice of the regularizer term  $\mathcal{N}(u)$  as  
218 well as the data fitting term  $\mathcal{H}(g, \mathcal{A}u)$ . Recall that  $\Omega \subset \mathbb{R}^d$  with  $d \geq 1$ ,  $\mathcal{T} = (0, T)$  with  $T > 0$   
219 and  $Q = \mathcal{T} \times \Omega \subset \mathbb{R}^{d+1}$ .

220 **3.1. Spatial-temporal regularizer.** In this section, we define the spatial-temporal total  
221 variation and infimal convolution total variation regularizers weighted by time dependent pa-  
222 rameters. Let  $\alpha$  be a positive time-dependent weight function  $\alpha \in W^{1,\infty}(\mathcal{T})$ . For the spatial  
223 and temporal variations, we write  $\Phi_{\alpha_1}(u)$  (in space) as the  $L^1(\mathcal{T})$  norm of  $t \mapsto \alpha_1(t) \text{TV}_x(u)(t)$ ,  
224 i.e.,

$$225 \quad (3.2) \quad \forall u \in L^1(\mathcal{T}; \text{BV}(\Omega)), \quad \Phi_{\alpha_1}(u) = \int_0^T \text{TV}_x[\alpha_1 u](t) dt = \int_0^T \alpha_1(t) \text{TV}_x[u](t) dt,$$

226 and for temporal penalization,  $\Psi_{\alpha_2}$  as

$$227 \quad (3.3) \quad \forall u \in L^1(\Omega; \text{BV}(\mathcal{T})), \quad \Psi_{\alpha_2}(v) = \int_{\Omega} \text{TV}_t[\alpha_2 u](x) dx.$$

228 Note that  $\Phi_{\alpha_1}, \Psi_{\alpha_2}$  are convex functionals and that the time dependent parameters  $\alpha_1, \alpha_2$   
 229 will satisfy

$$230 \quad (3.4) \quad \begin{cases} \alpha_1, \alpha_2 \in W^{1,\infty}(\mathcal{T}) \text{ and there exists} \\ \alpha_{min} > 0 \text{ s.t. } 0 < \alpha_{min} \leq \alpha_i(t) \text{ a.e. } t \in \mathcal{T}, i = 1, 2. \end{cases}$$

231 Therefore, using Lemma 2.1 and equations (3.2),(3.3) we have the following:

232 **Definition 3.1.** Let be  $\mathcal{X} = \text{BV}(Q)$  and  $\alpha = (\alpha_1, \alpha_2)$  that satisfies (3.4). We define the  
 233 spatial-temporal total variation regularizer  $F_\alpha$  on  $\mathcal{X}$  as

$$234 \quad (3.5) \quad F_\alpha(u) = \Phi_{\alpha_1}(u) + \Psi_{\alpha_2}(u),$$

that is

$$F_\alpha(u) = \int_0^T \text{TV}_x[\alpha_1 u](t) dt + \int_\Omega \text{TV}_t[\alpha_2 u](x) dx.$$

Moreover, for the spatial-temporal infimal convolution total variation regularization we fix  
 $\lambda = (\lambda_1, \lambda_2)$  and  $\mu = (\mu_1, \mu_2)$  that satisfy (3.4) and write

$$\forall u \in \mathcal{X}, \quad F_\lambda \# F_\mu(u) = \inf_{v \in \mathcal{X}} F_\lambda(u - v) + F_\mu(v).$$

235 **Proposition 3.1 (Lower semicontinuity of  $F_\alpha$ ).** For every  $\alpha = (\alpha_1, \alpha_2)$  that satisfies (3.4),  
 236 the functionals  $\Phi_{\alpha_1}$  and  $\Psi_{\alpha_2}$  are lower semicontinuous on  $L^1(\mathcal{T}; \text{BV}(\Omega))$  and  $L^1(\Omega; \text{BV}(\mathcal{T}))$   
 237 respectively, with respect to the  $L^1(Q)$  topology. In particular, the functional  $F_\alpha$  is lower semi-  
 238 continuous on  $\text{BV}(Q)$  with respect to the  $L^1(Q)$  topology. As a consequence, these functionals  
 239 are lower semicontinuous on  $\text{BV}(Q)$  for any  $L^p(Q)$  topology with  $p \geq 1$ .

240 *Proof.* We start with the lower semicontinuity of  $\Phi_{\alpha_1}$ . The proof is similar for the lower  
 241 semicontinuity of  $\Psi_{\alpha_2}$ . Let  $u_n \in L^1(\mathcal{T}; \text{BV}(\Omega))$  such that  $u_n \rightarrow u$  in  $L^1(Q)$ .

242 If  $\liminf_{n \rightarrow +\infty} \Phi_{\alpha_1}(u_n) = +\infty$  then the lower semicontinuity inequality is obviously satisfied.

243 Otherwise, one can extract a subsequence (still denoted  $u_n$ ) such that  
 244  $\sup_n \Phi_{\alpha_1}(u_n) = \sup_n \int_0^T \text{TV}_x[\alpha_1 u_n](t) dt < +\infty$ . Fatou's Lemma applied to the sequence  
 245  $\text{TV}_x(\alpha_1 u_n)$  gives

$$246 \quad \int_0^T \liminf_{n \rightarrow +\infty} \text{TV}_x[\alpha_1 u_n](t) dt \leq \liminf_{n \rightarrow +\infty} \int_0^T \text{TV}_x[\alpha_1 u_n](t) dt = \liminf_{n \rightarrow +\infty} \Phi_{\alpha_1}(u_n) < +\infty.$$

Moreover, for a.e.  $t \in \mathcal{T}$  we have

$$\forall \xi \in K_x, \quad \text{TV}_x[\alpha_1 u_n](t) \geq \int_\Omega \alpha_1(t) \xi(x) u_n(t, x) dx.$$

As  $u_n$  strongly converges to  $u$  in  $L^1(Q)$  then  $u_n(t, x) \rightarrow u(t, x)$  in  $L^1(\Omega)$  a.e.  $t \in \mathcal{T}$  up to a  
 subsequence. Therefore,

$$\forall \xi \in K_x, \text{ a.e. } t \in (0, T), \quad \liminf_{n \rightarrow +\infty} \text{TV}_x[\alpha_1 u_n](t) \geq \int_\Omega \alpha_1(t) \xi(x) u(t, x) dx,$$

and for almost every  $t \in \mathcal{T}$

$$\liminf_{n \rightarrow +\infty} \text{TV}_x[\alpha_1 u_n](t) \geq \sup_{\xi \in K_x} \int_{\Omega} \alpha_1(t) \xi(x) u(t, x) dx = \text{TV}_x[\alpha_1 u](t).$$

Finally,

$$\Phi_{\alpha_1}(u) = \int_0^T \text{TV}_x[\alpha_1 u](t) dt \leq \int_0^T \liminf_{n \rightarrow +\infty} \text{TV}_x[\alpha_1 u_n](t) dt \leq \liminf_{n \rightarrow +\infty} \Phi_{\alpha_1}(u_n).$$

248 Eventually, the functional  $F_{\alpha}$  is lower semicontinuous on  $\text{BV}(Q)$  as the sum of two lower  
249 semicontinuous functionals.  $\square$

250 Next result provides a relation between the total variation regularization which correlates  
251 space and time and the functional  $F_{\alpha}$  where these directions are treated separately. It is a key  
252 result to prove well-posedness results in the forthcoming analysis.

253 **Theorem 3.1.** *Assume that  $\alpha = (\alpha_1, \alpha_2)$  satisfies (3.4). Then, there exists positive con-*  
254 *stants  $C_{\alpha}^{-}$ ,  $C_{\alpha}^{+}$  depending on  $\alpha$ , such that for every  $u \in \text{BV}(Q)$*

$$255 \quad (3.6) \quad C_{\alpha}^{-} \text{TV}(\alpha_2 u) \leq F_{\alpha}(u) \leq C_{\alpha}^{+} \text{TV}(\alpha_2 u).$$

*Proof.* Let  $\alpha_{max} = \max\{\|\alpha_1\|_{L^{\infty}(\mathcal{T})}, \|\alpha_2\|_{L^{\infty}(\mathcal{T})}\}$  and note that  $\Phi_{\alpha_1}(u) = \Phi_1(\alpha_1 u)$ , for every  $u \in \text{BV}(Q)$ . Then, we have that

$$\frac{\alpha_{min}}{\alpha_{max}} \Phi_1(\alpha_2 u) \leq \Phi_{\alpha_1}(u) \leq \frac{\alpha_{max}}{\alpha_{min}} \Phi_1(\alpha_2 u), \quad \forall u \in \text{BV}(Q).$$

256 Since  $F_{\alpha}(u) = \Phi_{\alpha_1}(u) + \Psi_{\alpha_2}(u) = \Phi_1(\frac{\alpha_1}{\alpha_2} \alpha_2 u) + \Psi_1(\alpha_2 u)$  we conclude to

$$257 \quad \frac{\alpha_{min}}{\alpha_{max}} \Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u) \leq F_{\alpha}(u) \leq \frac{\alpha_{max}}{\alpha_{min}} \Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u) \Rightarrow$$

$$258 \quad \frac{\alpha_{min}}{\alpha_{max}} (\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u)) \leq F_{\alpha}(u) \leq \frac{\alpha_{max}}{\alpha_{min}} (\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u)),$$

259 since  $\frac{\alpha_{min}}{\alpha_{max}} \leq 1$  and  $\frac{\alpha_{max}}{\alpha_{min}} \geq 1$ . Using (2.5) in Lemma 2.1, we obtain

$$260 \quad (3.7) \quad \frac{\alpha_{min}}{\alpha_{max}} \text{TV}(\alpha_2 u) \leq F_{\alpha}(u) \leq \sqrt{2} \frac{\alpha_{max}}{\alpha_{min}} \text{TV}(\alpha_2 u).$$

262 Here  $C_{\alpha}^{-} = \frac{\alpha_{min}}{\alpha_{max}}$  and  $C_{\alpha}^{+} = \sqrt{2} \frac{\alpha_{max}}{\alpha_{min}}$ .  $\square$

263 In (3.6), we observe that the time dependent parameter  $\alpha_1$  that acts on the spatial domain  
264 of  $F_{\alpha}$  does not contribute to the correlated spatial-temporal total variation. In terms of the  
265 infimal convolution regularizer, a similar result is true when a certain assumption on the time  
266 dependent parameters is imposed.

267 **Proposition 3.2.** *Let  $\lambda = (\lambda_1, \lambda_2)$  and  $\mu = (\mu_1, \mu_2)$  be time dependent positive parameters*  
268 *that satisfy (3.4). Additionally, let  $\kappa > 0$  such that  $\mu_2 = \kappa \lambda_2$ . Then, there exists constants*  
269  *$C_1, C_2 > 0$  depending on  $\lambda, \mu$  and  $\kappa$  such that*

$$270 \quad (3.8) \quad \forall u \in \text{BV}(Q), \quad C_1 \text{TV}(\lambda_2 u) \leq F_{\lambda} \# F_{\mu}(u) \leq C_2 \text{TV}(\lambda_2 u).$$

271 *Proof.* Let be  $u \in \text{BV}(Q)$ , then for any  $v \in \text{BV}(Q)$  using Theorem 3.1, we have that

$$\begin{aligned}
272 \quad F_{\lambda}(u-v) + F_{\mu}(v) &\geq C_{\lambda}^{-} \text{TV}(\lambda_2(u-v)) + C_{\mu}^{-} \text{TV}(\mu_2 v) = C_{\lambda}^{-} \text{TV}(\lambda_2(u-v)) + \kappa C_{\mu}^{-} \text{TV}(\lambda_2 v) \\
273 \quad &\geq \min \{C_{\lambda}^{-}, \kappa C_{\mu}^{-}\} \left( \text{TV}(\lambda_2(u-v)) + \text{TV}(\lambda_2 v) \right) \geq C_1 \text{TV}(\lambda_2 u) \\
274
\end{aligned}$$

Passing to the infimum over  $v \in \text{BV}(Q)$  and obtain the left-hand side of (3.8). On the other hand, we have that

$$\inf_{v \in \text{BV}(Q)} F_{\lambda}(u-v) + F_{\mu}(v) \leq F_{\lambda}(u) \leq C_{\lambda}^{+} \text{TV}(\lambda_2 u) = C_2 \text{TV}(\lambda_2 u).$$

275

□

276 **Remark 3.1 (Choice of parameters).** *The assumption that there exists  $\kappa > 0$  such that  $\mu_2 =$*   
277  *$\kappa \lambda_2$  is a technical assumption and crucial for our analysis that follows. However, it is not*  
278 *too restrictive. Under this setting, one has to tune four parameters in total. Yet, we need to*  
279 *take into account the spatial and temporal regularization for each term. For instance, if one*  
280 *considers  $\lambda_i, \mu_i, i = 1, 2$  which satisfy (3.4) and  $\lambda_1 > \lambda_2, \mu_1 > \mu_2$  it is immediate that only*  
281 *a spatial regularization is enforced and vice versa. In order to employ an infimal convolution*  
282 *approach a certain relation between  $\lambda, \mu$  has to be imposed. For instance, one choice could be*  
283  *$\lambda_1 = \mu_2 = \lambda(t), \lambda_2 = \mu_1 = 1 - \lambda(t)$  with  $0 < \lambda(t) < 1$  for every  $t \in \mathcal{T}$ , see for instance [7].*  
284 *However, the assumption  $\mu_2 = \kappa \lambda_2$  may be not satisfied in that case except if we choose constant*  
285 *parameters. One could choose instead,  $\lambda_1(t), \lambda_2(t) \in (\lambda_{\min}, 1), \lambda_2(t) < \lambda_1(t), \mu_1(t) = 1 - \lambda_1(t)$*   
286 *and  $\mu_2(t) = \kappa \lambda_2(t)$  with  $\kappa > \frac{1 - \lambda_{\min}}{\lambda_{\min}}$  for example. In that case, we have  $\lambda_1 > \lambda_2$  and  $\mu_1 < \mu_2$ . In*  
287 *general the choice of parameters should follow a specific rule in order to avoid only spatial and*  
288 *only temporal regularization.*

289 The following is an immediate result when we consider constant parameters with respect  
290 to time.

291 **Corollary 3.1.** *Assume  $\alpha, \lambda$  and  $\mu$  are positive constant parameters. Then, we have the*  
292 *following relations for every  $u \in \text{BV}(Q)$ ,*

$$\begin{aligned}
293 \quad \alpha_{\max} C_{\alpha}^{-} \text{TV}(u) &\leq F_{\alpha}(u) \leq \alpha_{\min} C_{\alpha}^{+} \text{TV}(u) \\
294 \quad \lambda_{\min} C_1 \text{TV}(u) &\leq F_{\lambda} \# F_{\mu}(u) \leq \lambda_{\max} C_2 \text{TV}(u),
\end{aligned}$$

296 where  $\alpha_{\min} = \min \{\alpha_1, \alpha_2\}$  and  $\alpha_{\max} = \max \{\alpha_1, \alpha_2\}$  and respectively for  $\lambda_{\min}$  and  $\lambda_{\max}$ .

*Proof.* Recall that relation (2.5) gives

$$\text{TV}(u) \leq F_1(u) = \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \leq \sqrt{2} \text{TV}(u).$$

297 Next, we get

$$\begin{aligned}
298 \quad \alpha_{max} C_{\alpha}^{-} \text{TV}(u) &= \frac{\alpha_{max} \alpha_{min}}{\alpha_{max}} \text{TV}(u) \\
299 \quad &= \alpha_{min} \text{TV}(u) \\
300 \quad &\leq \alpha_{min} \left( \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \right) \\
301 \quad &\leq \alpha_1 \int_0^T \text{TV}_x(u)(t) dt + \alpha_2 \int_{\Omega} \text{TV}_t(u)(x) dx = F_{\alpha}(u) . \\
302
\end{aligned}$$

303 Similarly,

$$\begin{aligned}
304 \quad \alpha_{min} C_{\alpha}^{+} \text{TV}(u) &= \sqrt{2} \frac{\alpha_{min} \alpha_{max}}{\alpha_{min}} \text{TV}(u) \\
305 \quad &= \sqrt{2} \alpha_{max} \text{TV}(u) \\
306 \quad &\geq \alpha_{max} \left( \int_0^T \text{TV}_x(u)(t) dt + \int_{\Omega} \text{TV}_t(u)(x) dx \right) \\
307 \quad &\geq \alpha_1 \int_0^T \text{TV}_x(u)(t) dt + \alpha_2 \int_{\Omega} \text{TV}_t(u)(x) dx = F_{\alpha}(u) . \\
308
\end{aligned}$$

309 The second inequality is a direct consequence of Proposition 3.2.  $\square$

310 **3.2. Fitting data term.** In this section, we describe the possible choices of the data fitting  
311 term depending on the degradation of the input dynamic datum  $g$  as well as the linear operator  
312  $\mathcal{A}$ . Our setting is quite general and can be applied to any video denoising and deblurring  
313 application for instance, or even dynamic emission tomography (ET) such as Positron Emission  
314 Tomography (PET). We begin with two separate cases in terms of the linear operator  $\mathcal{A}$ .

315

316

### Case (1) : $\mathcal{A} = A$

317

318

We consider a linear and continuous operator with the following assumptions:

$$319 \quad (i) A \in \mathcal{L}(L^p(Q), L^q(Q)) \text{ with } 1 < p \leq \frac{d+1}{d}, \quad 1 \leq q < \infty,$$

$$320 \quad (3.9) \quad (ii) A \chi_Q \neq 0,$$

$$321 \quad (iii) A(\alpha(t)u) = \alpha(t)A(u), \text{ a.e. } t \in \mathcal{T}, \text{ for any positive time dependent parameter } \alpha.$$

322

Condition (ii) yields that  $A$  does not annihilate constant functions which is an important tool to derive existence results. Condition (iii) is obviously satisfied if  $\alpha$  is a positive constant. However, we require more: we need that an one-homogenous property holds for any positive time dependent function  $t \mapsto \alpha(t)$ . This may appear restrictive but it still allows to consider an identity operator for  $A$ : this is the case when we deal with denoising. This includes also spatial deblurring processes. Indeed, in that case we define  $A$  as a *spatial* convolution operator. Precisely, we may consider  $Au := h * u$ , where  $h$  is a spatially blurring kernel that remains constant over the time domain. Consequently, we get

$$A(\alpha(t)u(t, x)) = \alpha(t)A(u(t, x)) = \alpha(t)(h(x) * u(t, x)).$$

323 Next we may define,

$$324 \quad (3.10) \quad \mathcal{H}(g, \mathcal{A}u) = \frac{1}{q} \| \mathcal{A}u - g \|_{L^q(Q)}^q \quad \text{with } g \in L^q(Q).$$

325 as our data fitting term. This is suitable for dynamic data corrupted by noise that follows  
326 Gaussian distribution ( $q = 2$ ) or impulse noise ( $q = 1$ ) for example, see also [10].

327

328

### Case (2) : $\mathcal{A} = \mathcal{R}$

329

330 Here, we consider a linear operator related to emission imaging. The dynamic data that  
331 we obtained during a PET scan for instance, are connected through an integral (projection)  
332 operator known as the Radon transform  $\mathcal{R}$ . For every  $t \in \mathcal{T}$ , we write

$$333 \quad (3.11) \quad (\mathcal{R}u(\theta, s))(t) = \int_{x \cdot \theta = s} u(t, x) dx,$$

334 where  $\{x \in \mathbb{R}^d : x \cdot \theta = s\}$  is the hyperplane perpendicular to  $\theta \in \mathcal{S}^{d-1}$  with distance  $s \in \mathbb{R}$   
335 from the origin. For  $t \in \mathcal{T}$ ,  $(\mathcal{R}u(\theta, s))(t)$  lies on  $\{(\theta, s) : \theta \in \mathcal{S}^{d-1}, s \in \mathbb{R}\}$ , a cylinder of  
336 dimension  $d$  and is often referred as *projection space* or *sinogram space*. In the dynamic  
337 framework, we set  $\Sigma = \mathcal{T} \times \{(\theta, s) : \theta \in \mathcal{S}^{d-1}, s \in \mathbb{R}\}$  and the Radon transform is a continuous  
338 linear operator with

$$339 \quad (3.12) \quad \mathcal{R} : L^1(Q) \rightarrow L^1(\Sigma), \quad \|\mathcal{R}u\|_{L^1(\Sigma)} \leq C \|u\|_{L^1(Q)}.$$

340 We refer the reader to [31] for general continuity results of the Radon transform in  $L^p$  spaces.  
341 Furthermore, if  $p \geq \frac{d+1}{d}$ , the Radon transform is  $L^p$  discontinuous, since the function  
342  $u(x) = |x|^{-\frac{d+1}{p}} \frac{1}{\log(|x|)}$  belongs to  $L^p(Q)$ , for  $x \in Q$  but is not integrable over any hyperplane,  
343 see [28, Th. 3.32].

344 During the PET acquisition process, a certain amount of events e.g., photon-emissions are  
345 collected by the scanner (detectors) and organized into the so-called temporal bins  $g(\theta, s, t)$   
346 for every  $t \in \mathcal{T}$ . The associated noise in this data is called *photon* noise due to the ran-  
347 domness in the photo counting process and in fact, obeys the well-known Poisson probabil-  
348 ity distribution. For this kind of noise we use the *Kullback-Leibler* divergence, see [10],[27],  
349  $D_{KL} : L^1(\Sigma) \times L^1(\Sigma) \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ , defined as

$$350 \quad (3.13) \quad D_{KL}(w_1, w_2) = \begin{cases} \int_{\Sigma} \left( w_1 \log \left( \frac{w_1}{w_2} \right) - w_1 + w_2 \right) dx dt, & \forall w_1 \geq 0, w_2 > 0 \text{ a.e.} \\ +\infty & \text{otherwise} \end{cases}$$

351 This is in fact the *Bregman* distance of the *Boltzmann-Shannon* entropy, see [33]. We briefly  
352 recall some of the basic properties of the KL-functional which can be found in [11],[33] and  
353 will be used later.

354 **Lemma 3.1.** *The following properties hold true:*

355 (a)  $D_{KL}(w_1, w_2)$  is nonnegative and equal to 0 if and only if  $w_1 = w_2$ .

356 (b) The function  $(w_1, w_2) \mapsto D_{KL}(w_1, w_2)$  is convex.



357 (c) For fixed  $w_1 \in L^1_+(\Sigma)$  (resp.  $w_2 \in L^1_+(\Sigma)$ ), the function  $D_{KL}(w_1, \cdot)$  (resp.  $D_{KL}(\cdot, w_2)$ ) is  
 358 weakly lower semicontinuous with respect to  $L^1(\Sigma)$  topology.

359 (d) For every  $w_1, w_2 \in L^1_+(\Sigma)$

$$360 \quad (3.14) \quad \|w_1 - w_2\|_{L^1(\Sigma)}^2 \leq \left( \frac{2}{3} \|w_1\|_{L^1(\Sigma)} + \frac{4}{3} \|w_2\|_{L^1(\Sigma)} \right) D_{KL}(w_1, w_2).$$

361 In what follows, we fix  $w_1 = g$  as the dynamic datum. Assume that

$$362 \quad (3.15) \quad g \in L^\infty(\Sigma),$$

363 and set

$$364 \quad (3.16) \quad \forall w \in L^1(\Sigma), \quad \mathcal{H}(g, w) = \begin{cases} \int_{\Sigma} w - g \log w \, d\theta \, ds \, dt & \text{if } w > 0 \text{ and } \log w \in L^1(\Sigma) \\ +\infty & \text{else.} \end{cases}$$

365 With the above definition we have

$$366 \quad (3.17) \quad D_{KL}(g, w) = \mathcal{H}(g, w) - \mathcal{H}(g, g).$$

367 As we deal with the minimization problem (3.1), we can neglect the terms that are independent  
 368 of  $w$ . Indeed, the  $\mathcal{H}(g, g)$  term do not count on the minimization problem (3.1). Let us mention  
 369 that the domain of above expression is the cone of positive functions whose log belongs to  $L^1(\Sigma)$   
 370 and that  $\mathcal{H}(g, w) = +\infty$ , if  $w$  vanishes on a subset of  $\Sigma$  of non null measure or if  $\log w \notin L^1(\Sigma)$ .  
 371 The boundedness assumption (3.15) is true from the practical point of view since we deal with  
 372 a finite acquisition time.

373 **Lemma 3.2.** *The Radon transform  $\mathcal{R}$  satisfies (3.9) (ii) and (iii).*

*Proof.* Due to the definition of the Radon transform (3.11), we clearly have

$$\mathcal{R}(\alpha(t)u) = \alpha(t)\mathcal{R}(u).$$

374 Moreover, the Radon transform is injective ([28, Theorem 2.57]) so that it does not annihilate  
 375 constant functions and relation (3.9) (ii) is ensured.  $\square$

376 To conclude, we define

$$377 \quad (3.18) \quad \mathcal{H}(g, \mathcal{R}u) = \int_{\Sigma} (\mathcal{R}u - g \log \mathcal{R}u) \, d\theta \, ds \, dt,$$

378 whose domain is

$$379 \quad (3.19) \quad \mathcal{D} := \{u \in L^1(Q) \mid \mathcal{R}u > 0 \text{ and } \log \mathcal{R}u \in L^1(\Sigma)\}$$

as our data fitting term. Note that  $\mathcal{D} \subset L^1_+(Q)$  since  $u \geq 0$  a.e. implies that  $\mathcal{R}u \geq 0$  a.e. As  
 a direct consequence of Lemma 3.1 and the definitions above we get a lower semicontinuity  
 result for  $\mathcal{H}$ . Precisely, for every sequence  $(u_n) \in \mathcal{D}$  that strongly converges to  $u$  for the  $L^1(Q)$   
 topology we have

$$\mathcal{H}(g, \mathcal{R}u) \leq \liminf_{n \rightarrow +\infty} \mathcal{H}(g, \mathcal{R}u_n).$$

380 **Remark 3.2.** *Though we are mainly interested in the Radon transform case, one could re-*  
 381 *place  $\mathcal{R}$  with any operator that satisfies (3.9) as in Case 1. This may be suitable for Poisson*  
 382 *denoising and deblurring.*

383 **4. Well-posedness results.** In this section, we are interested in the well-posedness of the  
 384 minimization problem (3.1) for the regularizers described in Section 3.1 and the different choices  
 385 of the data fitting term in (3.10) and (3.18). We focus on the infimal convolution total variation  
 386 regularizer case i.e.,  $\mathcal{N}(u) := F_\lambda \# F_\mu(u)$ . In the case of the total variation regularizer, the  
 387 forthcoming analysis is similar and most of the proofs are the same with minor adaptations.  
 388 We prove well-posedness (existence, uniqueness and stability) via the direct method of calculus  
 389 of variations for

$$390 \quad (\mathcal{P}) \quad \inf_{u \in \text{BV}(Q)} \mathcal{E}(u),$$

391 where

$$392 \quad (4.1) \quad \mathcal{E}(u) := \mathcal{H}(g, \mathcal{A}u) + F_\lambda \# F_\mu(u).$$

393 In particular, we need the lower semicontinuity condition to be true for both the regularizing  
 394 and the fidelity term, together with some compactness properties. Note that the balancing  
 395 parameters between the fidelity term and the regularization term, namely  $\lambda_i, \mu_i, i = 1, 2$  are  
 396 involved in the definition of this regularization term. Precisely the cost functional of problem  
 397  $(\mathcal{P})$ , writes

$$398 \quad \mathcal{H}(g, \mathcal{A}u) + \inf_{v \in \text{BV}(Q)} \int_0^T (\text{TV}_x[\lambda_1(u-v)] + \text{TV}_x[\mu_1 v]) (t) dt +$$

$$399 \quad \int_\Omega (\text{TV}_t[\lambda_2(u-v)] + \text{TV}_t[\mu_2 v]) (x) dx.$$

400

401 **4.1. Lower semicontinuity of the inf-convolution operator.** Note that the lower semi-  
 402 continuity of the inf-convolution operator is not true in general, even if  $F_\lambda$  is, see [6, Example  
 403 12.13]. Additional assumptions have to be imposed such as coercivity on the underlying space  
 404 as well as exactness of the infimal convolution in order to get the lower semicontinuity. We  
 405 first need the following technical Lemma which provides an estimate on  $u \in \text{BV}(Q)$  when (3.4)  
 406 is satisfied. Precisely

407 **Lemma 4.1.** *Assume that  $\alpha \in W^{1,\infty}(\mathcal{T})$  and that there exists  $\alpha_{\min} > 0$  such that  $0 <$   
 408  $\alpha_{\min} \leq \alpha(t)$  a.e.  $t \in \mathcal{T}$ ; then  $1/\alpha \in W^{1,\infty}(\mathcal{T})$ . Moreover, if  $\alpha u \in \text{BV}(Q)$  then  $u \in \text{BV}(Q)$  as  
 409 well.*

*Proof.* Let  $\alpha$  be in  $W^{1,\infty}(\mathcal{T})$  such that  $0 < \alpha_{\min} \leq \alpha(t)$  a.e.  $t \in \mathcal{T}$ . We use Proposition  
 8.4 of [13] : a function  $f \in L^\infty(\mathcal{T})$  belongs to  $W^{1,\infty}(\mathcal{T})$  if and only if there exists a constant  
 $C$  such that

$$|f(x) - f(y)| \leq C|x - y| \text{ for a.e. } x, y \in \mathcal{T}.$$

Here, we assume that  $\alpha \in W^{1,\infty}(\mathcal{T})$  so that there exists  $C$  such that

$$|\alpha(x) - \alpha(y)| \leq C|x - y| \text{ for a.e. } x, y \in \mathcal{T}.$$

As  $0 < \frac{1}{\alpha} \leq \frac{1}{\alpha_{\min}}$  then the function  $\frac{1}{\alpha}$  belongs to  $L^\infty(\mathcal{T})$ . Moreover, for a.e.  $x, y \in \mathcal{T}$

$$\left| \left( \frac{1}{\alpha} \right) (x) - \left( \frac{1}{\alpha} \right) (y) \right| = \frac{|\alpha(x) - \alpha(y)|}{|\alpha(x)\alpha(y)|} \leq \frac{1}{\alpha_{\min}^2} |\alpha(x) - \alpha(y)| \leq \frac{C}{\alpha_{\min}^2} |x - y|.$$

410 Using again Proposition 8.4 of [13] this proves that  $1/\alpha \in W^{1,\infty}(\mathcal{T})$ . Moreover, if  $\alpha u \in \text{BV}(Q)$   
 411 then  $u \in \text{BV}(Q)$ . Indeed,  $u = \frac{1}{\alpha}(\alpha u)$  and

$$\begin{aligned} 412 \quad \|u\|_{\text{BV}(Q)} &= \left\| \frac{1}{\alpha}(\alpha u) \right\|_{\text{BV}(Q)} = \left\| \frac{1}{\alpha}(\alpha u) \right\|_{L^1(Q)} + \text{TV}\left(\frac{1}{\alpha}(\alpha u)\right) \\ 413 \quad &\leq \frac{1}{\alpha_{\min}} \|\alpha u\|_{L^1(Q)} + \text{TV}\left(\frac{1}{\alpha}(\alpha u)\right). \\ 414 \end{aligned}$$

415 Now, if  $\beta \in W^{1,\infty}(\mathcal{T})$  and  $v \in \text{BV}(Q)$  we get

$$416 \quad (4.2) \quad \text{TV}(\beta v) \leq \|\beta\|_{L^\infty(\mathcal{T})} \text{TV}(v) + \|\beta'\|_{L^\infty(\mathcal{T})} \|v\|_{L^1(Q)},$$

where  $\beta'$  is the (distributional) derivative of  $\beta$ . We set with  $\beta = \frac{1}{\alpha}$  and  $v = \alpha u$ :

$$\text{TV}\left(\frac{1}{\alpha}(\alpha u)\right) \leq \left\| \frac{1}{\alpha} \right\|_{L^\infty(\mathcal{T})} \text{TV}(\alpha u) + \left\| \left(\frac{1}{\alpha}\right)' \right\|_{L^\infty(\mathcal{T})} \|\alpha u\|_{L^1(Q)}.$$

417 Finally,

$$418 \quad (4.3) \quad \|u\|_{\text{BV}(Q)} \leq \left( \frac{1}{\alpha_{\min}} + \left\| \left(\frac{1}{\alpha}\right)' \right\|_{L^\infty(\mathcal{T})} \right) \|\alpha u\|_{L^1(Q)} + \frac{\text{TV}(\alpha u)}{\alpha_{\min}} \leq C_\alpha \|\alpha u\|_{\text{BV}(Q)} < +\infty$$

$$419 \quad \text{with } C_\alpha = \frac{1}{\alpha_{\min}} + \left\| \left(\frac{1}{\alpha}\right)' \right\|_{L^\infty(\mathcal{T})}. \quad \square$$

420 Next, we show that the inf-convolution operator is *exact* in our case.

**Lemma 4.2 (Exactness of  $F_\lambda \# F_\mu$ ).** *Assume that  $\lambda$  and  $\mu$  verify (3.4) and there exists  $\kappa > 0$  such that  $\mu_2 = \kappa \lambda_2$ . Then, for every  $u \in \text{BV}(Q)$ , there exists  $v_u \in \text{BV}(Q)$  such that*

$$v_u \in \underset{v \in \text{BV}(Q)}{\text{argmin}} F_\lambda(u - v) + F_\mu(v) \text{ and } \int_Q \mu_2(t) v_u(t, x) dt dx = 0.$$

*Proof.* Fix  $u \in \text{BV}(Q)$ . Let  $v_n$  be a minimizing sequence of

$$\inf_{v \in \text{BV}(Q)} F_\lambda(u - v) + F_\mu(v).$$

Then  $v_n \in \text{BV}(Q)$  and without loss of generality we may assume that the mean value of  $\mu_2 v_n$  is

$$\overline{\mu_2 v_n} := \frac{1}{|Q|} \int_Q \mu_2(t) v_n(t, x) dx dt = 0.$$

421 Indeed, since  $\mu_2 = \kappa \lambda_2$ , it is easy to see that

$$422 \quad F_\lambda \left( u - \left( v_n - \frac{1}{\mu_2} \overline{\mu_2 v_n} \right) \right) + F_\mu \left( v_n - \frac{1}{\mu_2} \overline{\mu_2 v_n} \right) = F_\lambda(u - v_n) + F_\mu(v_n), \\ 423$$

424 so that  $w_n := v_n - \frac{1}{\mu_2} \overline{\mu_2 v_n}$  is also a minimizing sequence that satisfies  $\int_Q \mu_2 w_n dx dt = 0$ .

425 As  $F_\lambda(u - v_n) + F_\mu(v_n)$  is bounded then Theorem 3.1 yields that  $\text{TV}(\mu_2 v_n)$  is bounded as  
 426 well. Moreover, we have  $\|\mu_2 v_n\|_{L^1(Q)} \leq C_Q \text{TV}(\mu_2 v_n)$  from the Poincaré-Wirtinger inequality  
 427 (see Theorem 2.1). Hence,  $(\mu_2 v_n)$  is BV-bounded. This implies that  $v_n$  is BV-bounded as well  
 428 (see Lemma 4.1 and (4.3)). Therefore, there exists  $v_u \in \text{BV}(Q)$  such that, up to subsequence,  
 429  $v_n \xrightarrow{w^*} v_u$  in  $\text{BV}(Q)$  which implies that  $v_n \rightarrow v_u$  for the  $L^1(Q)$  topology. We end the proof  
 430 with the lower semicontinuity of the functional with respect to the the  $L^1(Q)$  topology ( with  
 431 Proposition 3.1). In addition, since  $\int_Q \mu_2(t) v_n(t, x) dx dt = 0$ , we get from the  $L^1$  convergence  
 432 that  $\int_Q \mu_2(t) v_u(t, x) dx dt = 0$  as well.  $\square$

433 Now we prove a lower semicontinuity result of  $F_\lambda \# F_\mu$ . Here, we use the exactness of  $F_\lambda \# F_\mu$   
 434 and the BV coercivity of one of its terms. For more details on the lower semicontinuity of the  
 435 infimal convolution we refer to [41].

436 **Theorem 4.1.** *Assume that  $\lambda$  and  $\mu$  verify (3.4) and there exists  $\kappa > 0$  such that  $\mu_2 = \kappa \lambda_2$ .  
 437 Then, the infimal-convolution  $F_\lambda \# F_\mu$  operator is lower semicontinuous on  $\text{BV}(Q)$  with respect  
 438 to the  $L^1(Q)$  topology. Precisely, if  $u_n$  is a sequence in  $\text{BV}(Q)$  that converges to some  $u$  with  
 439 respect to the strong  $L^1(Q)$  topology then*

$$440 \quad (4.4) \quad F_\lambda \# F_\mu(u) \leq \liminf_{n \rightarrow +\infty} F_\lambda \# F_\mu(u_n).$$

*Proof.* Let  $u_n \in \text{BV}(Q)$  such that  $u_n \rightarrow u$  in  $L^1(Q)$ . If  $\liminf_{n \rightarrow +\infty} F_\lambda \# F_\mu(u_n) = +\infty$  then  
 relation (4.4) is satisfied. Otherwise, there exists a subsequence (denoted similarly) and a  
 constant  $C$  such that for every  $n \in \mathbb{N}$ ,  $F_\lambda \# F_\mu(u_n) \leq C$ . Since  $F_\lambda \# F_\mu$  is exact, there exists  
 $v_n \in \text{BV}(Q)$  such that

$$\forall n \in \mathbb{N} \quad F_\lambda(u_n - v_n) + F_\mu(v_n) = F_\lambda \# F_\mu(u_n) \text{ and } \int_Q \mu_2 v_n = 0.$$

We claim that  $(\mu_2 v_n)$  is BV-bounded (that is  $\|\mu_2 v_n\|_{\text{BV}(Q)}$  is uniformly bounded with respect  
 to  $n$ ). Indeed, Theorem (3.1) yields

$$\forall n \in \mathbb{N} \quad \text{TV}(\mu_2 v_n) \leq \frac{1}{C_\mu^-} F_\mu(v_n) \leq \frac{C}{C_\mu^-}.$$

Using Poincaré-Wirtinger inequality, we have that

$$\forall n \in \mathbb{N} \quad \|\mu_2 v_n\|_{L^1(Q)} \leq C_Q \text{TV}(\mu_2 v_n) \leq \frac{C C_Q}{C_\mu^-}.$$

Following similar steps as before, there exists a subsequence  $v_n \xrightarrow{w^*} \tilde{v}$  in  $\text{BV}(Q)$ . Due to the  
 lower semicontinuity  $F_\lambda$  and  $F_\mu$  with respect to the  $L^1(Q)$  topology and its exactness, we have

$$F_\lambda(u - \tilde{v}) + F_\mu(\tilde{v}) \leq \liminf_{n \rightarrow +\infty} F_\lambda(u_n - v_n) + F_\mu(v_n) = \liminf_{n \rightarrow +\infty} F_\lambda \# F_\mu(u_n)$$

and since  $F_\lambda \# F_\mu(u) \leq F_\lambda(u - \tilde{v}) + F_\mu(\tilde{v})$ , we conclude that

$$F_\lambda \# F_\mu(u) \leq \liminf_{n \rightarrow +\infty} F_\lambda \# F_\mu(u_n).$$

441

 $\square$

442 **4.2. Well-posedness.** Now we focus on the existence of a solution for  $(\mathcal{P})$ . The proof is  
 443 based on the corresponding results in [1, 44, 33] adapted to a spatial-temporal framework.

444 **Theorem 4.2 (Existence).** *Assume that*

445 • *Case (1): the data  $g \in L^q(Q)$  and  $A$  satisfies (3.9), or*

446 • *Case (2): the data  $g \in L^\infty(\Sigma)$ .*

447 *Let  $\lambda, \mu$  be parameters that satisfy (3.4) and that there exists a real number  $\kappa > 0$  such that*  
 448  *$\mu_2 = \kappa\lambda_2$ . Then, there exists a solution to problem  $(\mathcal{P})$ .*

449 *Proof.* We first observe that  $\mathcal{E}(u)$  is bounded from below and there exists  $u_0 \in \text{BV}(Q)$   
 450 such that  $\mathcal{E}(u_0) < +\infty$ . Let  $u_n \in \text{BV}(Q)$  be a minimizing sequence of problem  $(\mathcal{P})$ . Then  
 451 there exists  $M_0[g] > 0$  such that

$$452 \quad (4.5) \quad \forall n \in \mathbb{N}, \quad F_\lambda \# F_\mu(u_n) + \mathcal{H}(g, Au_n) \leq M_0[g] < +\infty.$$

This implies in particular that  $u_n \in \text{BV}(Q) \cap \mathcal{D}$  in case (2). In the sequel, we indicate the  
 dependence of the different bounding constants  $M_i$  with respect to  $g$  because we need a precise  
 estimate to prove Theorem 4.4.

Using Proposition 3.2, we deduce that  $\text{TV}(w_n)$  is bounded where we have set  $w_n = \lambda_2 u_n$ .  
 Therefore, with the Poincaré-Wirtinger inequality, then we have  $\|w_n - \bar{w}_n\|_{L^p(Q)} \leq M_1[g]$   
 with  $1 \leq p \leq \frac{d+1}{d}$  and  $M_1[g] = \frac{C_Q}{C_1} M_0[g]$  Moreover, we have

$$\|w_n\|_{L^p(Q)} \leq \|w_n - \bar{w}_n\|_{L^p(Q)} + \|\bar{w}_n\|_{L^p(Q)} \leq M_1[g] + |Q|^{\frac{1}{p}-1} \left| \int_Q w_n \, dx \, dt \right|.$$

453 The goal is to prove that the sequence  $(u_n)$  is bounded in  $\text{BV}(Q)$ . This is equivalent to find an  
 454 estimate on the last term of the above inequality. To achieve this, we consider the two cases  
 455 with respect to the choice of the fidelity term presented in Section 3.2.

456

$$457 \quad \textbf{Case (1)} : \mathcal{H}(g, Au) = \frac{1}{q} \|Au - g\|_{L^q(Q)}^q$$

458

459 Recall that  $g \in L^q(Q)$ ,  $A \in \mathcal{L}(L^p(Q), L^q(Q))$  with  $1 \leq p \leq \frac{d+1}{d}$ ,  $1 \leq q < \infty$ , and satisfy  
 460 (3.9). Then, one has that

$$\begin{aligned} 461 \quad \left| \int_Q w_n \, dx \, dt \right| \frac{\|A\chi_Q\|_{L^q(Q)}}{|Q|} &= \|A\bar{w}_n\|_{L^q(Q)} = \|A\bar{w}_n - Aw_n + Aw_n - \lambda_2 g + \lambda_2 g\|_{L^q(Q)} \\ 462 \quad &\leq \|A\| \|w_n - \bar{w}_n\|_{L^p(Q)} + \|A(\lambda_2 u_n) - \lambda_2 g\|_{L^q(Q)} + \|\lambda_2 g\|_{L^q(Q)} \\ 463 \quad &\leq \|A\| \|w_n - \bar{w}_n\|_{L^p(Q)} + \|\lambda_2\|_{L^\infty(\mathcal{T})} \left( \|Au_n - g\|_{L^q(Q)} + \|g\|_{L^q(Q)} \right) \\ 464 \quad &\leq \|A\| M_1 + \|\lambda_2\|_{L^\infty(\mathcal{T})} \left( (qM_0)^{1/q} + \|g\|_{L^q(Q)} \right) \leq M_2, \\ 465 \end{aligned}$$

466 where

$$\begin{aligned}
467 \quad M_2[g] &= \|A\| M_1[g] + \|\lambda_2\|_{L^\infty(\mathcal{T})} \left( (qM_0[g])^{1/q} + \|g\|_{L^q(Q)} \right) \\
468 \quad (4.6) \quad &= \|A\| \frac{C_Q}{C_1} M_0[g] + q^{1/q} \|\lambda_2\|_{L^\infty(\mathcal{T})} M_0[g]^{1/q} + \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^q(Q)}. \\
469
\end{aligned}$$

$$470 \quad \underline{\text{Case (2)}} : \mathcal{H}(g, \mathcal{A}u) = D_{KL}(g, \mathcal{R}u) + \mathcal{H}(g, g)$$

471

472 Recall that  $g \in L^\infty(\Sigma)$  and that we require an additional positivity constraint  $u_n \geq 0$ .  
473 Therefore, it suffices to bound  $\int_Q w_n dx dt$ . We employ (3.14) and using (3.12) we have

$$\begin{aligned}
474 \quad \|\mathcal{R}w_n - \lambda_2 g\|_{L^1(\Sigma)}^2 &\leq \left( \frac{2}{3} \|\lambda_2 g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}w_n\|_{L^1(\Sigma)} \right) D_{KL}(\lambda_2 g, \lambda_2 \mathcal{R}u_n) \\
475 &\leq \left( \frac{2}{3} \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}(w_n - \bar{w}_n) + \mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \right) \|\lambda_2\|_{L^\infty(\mathcal{T})} D_{KL}(g, \mathcal{R}u_n) \\
476 &\leq \left( \frac{2}{3} \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}\| \|w_n - \bar{w}_n\|_{L^1(Q)} + \frac{4}{3} \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \right) \|\lambda_2\|_{L^\infty(\mathcal{T})} M_0[g] \\
477 &\leq \left( \frac{2}{3} \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}\| |Q|^{1/p'} M_1[g] + \frac{4}{3} \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \right) \|\lambda_2\|_{L^\infty(\mathcal{T})} M_0[g]. \\
478
\end{aligned}$$

479 Hence,

$$480 \quad (4.7) \quad \|\mathcal{R}w_n - \lambda_2 g\|_{L^1(\Sigma)}^2 \leq \left( M_3[g] + \frac{4}{3} \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \right) M_4[g],$$

481 with

$$482 \quad (4.8) \quad M_3[g] = \frac{2}{3} \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}\| |Q|^{1/p'} M_1[g]$$

$$483 \quad (4.9) \quad = \frac{2}{3} \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} + \frac{4}{3} \|\mathcal{R}\| |Q|^{1/p'} \frac{C_Q}{C_1} M_0[g],$$

484

485 and

$$486 \quad (4.10) \quad M_4[g] = \|\lambda_2\|_{L^\infty(\mathcal{T})} M_0[g],$$

487 On the other hand,

$$\begin{aligned}
488 \quad \|\mathcal{R}w_n - \lambda_2 g\|_{L^1(\Sigma)}^2 &\geq \left( \|\mathcal{R}(w_n - \bar{w}_n) - \lambda_2 g\|_{L^1(\Sigma)} - \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \right)^2 \\
489 &\geq \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \left( \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} - 2 \|\mathcal{R}(w_n - \bar{w}_n) - \lambda_2 g\|_{L^1(\Sigma)} \right) \\
490 &\geq \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \left( \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} - 2 \left( \|\mathcal{R}\| |Q|^{1/p'} M_1[g] + \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)} \right) \right) \\
491 \quad (4.11) \quad &= \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} \left( \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} - M_5[g] \right), \\
492
\end{aligned}$$

493 with

$$494 \quad (4.12) \quad M_5[g] = 2 \|\mathcal{R}\| |Q|^{1/p'} M_1[g] + \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)}$$

$$495 \quad (4.13) \quad = 2 \|\mathcal{R}\| |Q|^{1/p'} \frac{C_Q}{C_1} M_0[g] + \|\lambda_2\|_{L^\infty(\mathcal{T})} \|g\|_{L^1(\Sigma)}$$

496  
497 Also, we have that  $\|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} = \frac{\int_Q w_n dx dt}{|Q|} \|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}$  that is

$$498 \quad (4.14) \quad \|\mathcal{R}\bar{w}_n\|_{L^1(\Sigma)} = \frac{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}{|Q|} \|\bar{w}_n\|_{L^1(Q)}.$$

499 Combining (4.7),(4.11) and (4.14), we derive that

$$500 \quad (4.15) \quad \frac{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}{|Q|} \|\bar{w}_n\|_{L^1(Q)} \left( \frac{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}{|Q|} \|\bar{w}_n\|_{L^1(Q)} - M_5[g] - \frac{4}{3}M_4[g] \right) \leq M_3[g]M_4[g].$$

501

Let  $B_n[g] = \frac{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}{|Q|} \|\bar{w}_n\|_{L^1(Q)} - M_5[g] - \frac{4}{3}M_4[g]$ . If  $n$  is such that  $B_n \geq 1$ , it is immediate from (4.15) and  $\mathcal{R}\chi_Q \neq 0$ , see Lemma 3.2, that

$$\|\bar{w}_n\|_{L^1(Q)} \leq \frac{M_3[g]M_4[g]|Q|}{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}$$

Otherwise, we have that

$$\|\bar{w}_n\|_{L^1(Q)} \leq \left( 1 + M_5[g] + \frac{4}{3}M_4[g] \right) \frac{|Q|}{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}}.$$

we finally obtain for every  $n \in \mathbb{N}$

$$\|\bar{w}_n\|_{L^1(Q)} \leq M_6[g],$$

502 where

$$503 \quad (4.16) \quad M_6[g] = \frac{|Q|}{\|\mathcal{R}\chi_Q\|_{L^1(\Sigma)}} \max \left\{ M_3[g]M_4[g], \left( 1 + M_5[g] + \frac{4}{3}M_4[g] \right) \right\},$$

To conclude, we have proved that in both cases  $w_n = \lambda_2 u_n$  is bounded in  $L^p(Q)$  and hence is bounded in  $BV(Q)$ . Using Lemma 4.1,  $u_n$  is bounded both in  $BV(Q)$  and  $L^p(Q)$ . Then, there exists subsequence still denoted by  $u_n$  such that  $u_n \xrightarrow{w^*} u$  in  $BV(Q)$  i.e.,  $u_n \rightarrow u$  in  $L^1(Q)$  and  $u_n \xrightarrow{w} u$  in  $L^p(Q)$ ,  $1 < p \leq \frac{d+1}{d}$ . Theorem 4.1 yields that

$$F_\lambda \# F_\mu(u) \leq \liminf_{n \rightarrow \infty} F_\lambda \# F_\mu(u_n).$$

Moreover, due to the lower semicontinuity of the fidelity terms as well as the continuity of  $A$  and  $\mathcal{R}$ , we conclude that

$$\mathcal{H}(g, \mathcal{A}u) \leq \liminf_{n \rightarrow \infty} \mathcal{H}(g, \mathcal{A}u_n).$$

504 This means that  $u$  is a solution to  $(\mathcal{P})$ . □

505 **Remark 4.1.** *To be consistent with the cases where either  $\mathcal{A}$  is the identity operator, let us*  
 506 *mention that the BV-boundedness is immediate since*

$$507 \quad \|u_n\|_{L^q(Q)} \leq \|u_n - g\|_{L^q(Q)} + \|g\|_{L^q(Q)}$$

$$508 \quad \|u_n\|_{L^1(Q)} - \|g\|_{L^\infty(Q)} \log \|u_n\|_{L^1(Q)} \leq \int_Q u_n - g \log u_n$$
 509

510 *We refer to [27] for the second case.*

511 **Theorem 4.3 (Uniqueness).** *Assume that the hypothesis of Theorem (4.2) are fulfilled and,*  
 512 *in addition that*

- 513 •  $A$  is injective and  $q \neq 1$  in Case (1),
- 514 •  $\inf_\Sigma g > 0$  in Case (2).

515 *Then the solution to  $(\mathcal{P})$  is unique.*

516 *Proof.* Note that  $F_\lambda \# F_\mu$  is convex since  $F_\lambda$  and  $F_\mu$  are convex. We first consider Case  
 517 (1) : since  $1 < q < \infty$  and  $A$  is injective then  $u \mapsto \frac{1}{q} \|Au - g\|_{L^q(Q)}^q$  is strictly convex.  
 518 In case (2), since  $\inf_\Sigma g > 0$  and  $\mathcal{R}$  is injective, see for instance [28, Theorem 2.57], then  
 519  $u \mapsto D_{KL}(g, \mathcal{R}u)$  is strictly convex. In both cases, we have that the energy  $\mathcal{E}$  is strictly convex  
 520 as a sum of a convex and a strictly convex terms. This gives uniqueness.  $\square$

521 **Remark 4.2.** *The assumption that  $\inf_\Sigma g > 0$  is a usual approximation for the continuous*  
 522 *setting which implies a positive systematic bias on the sinogram domain, see [33, 37]. This*  
 523 *is not far from the reality since for a reasonably long counting process, where some million*  
 524 *of photons are detected, all the PET detectors will record a certain amount of photons, even*  
 525 *if it is relatively small in practice. Note that one has to consider not only the recorded true*  
 526 *coincidence events but also the random coincidence events which occur when separate positron*  
 527 *emissions are detected within a time window and recorded as having originated from the same*  
 528 *emission. This results in an additional background noise on the sinogram domain.*

529 To conclude this section, we discuss the stability of minimizers of  $(\mathcal{P})$ , see [1, 33, 37] for  
 530 instance, with respect to a *small* perturbation on the data  $g$ . Let  $(g_n)$  be a perturbed dynamic  
 531 data sequence such that

$$532 \quad (4.17) \quad \begin{cases} \|g_n - g\|_{L^q(Q)} \rightarrow 0, & g_n \in L^q(Q) & \text{Case (1)} \\ \|g_n - g\|_{L^\infty(\Sigma)} \rightarrow 0, & g_n \in L^\infty(\Sigma) & \text{Case (2)} \end{cases}$$

533 and the corresponding perturbed minimization problem

$$534 \quad (4.18) \quad \inf_{u \in \text{BV}(Q)} \mathcal{H}(g_n, \mathcal{A}u) + (F_\lambda \# F_\mu)(u).$$
 535

536 **Theorem 4.4 (Stability).** *Assume the assumptions of Theorem 4.3 are fulfilled for param-*  
 537 *eters  $\lambda$  and  $\mu$  and every datum  $g_n$ . Then problem  $(\mathcal{P})$  is stable with respect to perturbations on*  
 538  *$g$ . Precisely, let be  $(g_n)$  as in (4.17) and  $u, u_n$  be the solutions to  $(\mathcal{P})$  and (4.18) respectively.*  
 539 *Then, there exists a subsequence of  $(u_n)$  that converges to  $u$  in  $\text{BV}(Q)$ - $w^*$ .*



540 *Proof.* Since  $u_n$  minimizes (4.18), then for every  $v \in \text{BV}(Q)$

$$541 \quad (4.19) \quad (F_\lambda \# F_\mu)(u_n) + \mathcal{H}(g_n, \mathcal{A}u_n) \leq (F_\lambda \# F_\mu)(v) + \mathcal{H}(g_n, \mathcal{A}v).$$

542 As in the previous proofs, we consider each case separately.

543

$$544 \quad \text{Case (1) : } \mathcal{H}(g, \mathcal{A}u) = \frac{1}{q} \|Au - g\|_{L^q(Q)}^q$$

545

546 Since  $g_n \rightarrow g$  in  $L^q(Q)$ , then there exists  $n_0 \in \mathbb{N}$  such that  $\|g - g_n\|_q^q \leq \frac{q}{2^{q-1}}$  for every  
547  $n \geq n_0$ . Then, for every  $n \geq n_0$

$$\begin{aligned} 548 \quad (F_\lambda \# F_\mu)(u_n) + \frac{1}{q} \|Au_n - g\|_{L^q(Q)}^q &\leq 2^{q-1} \left( (F_\lambda \# F_\mu)(u_n) + \frac{1}{q} \|Au_n - g_n\|_{L^q(Q)}^q + \frac{1}{q} \|g_n - g\|_{L^q(Q)}^q \right) \\ 549 &\leq 2^{q-1} \left( (F_\lambda \# F_\mu)(u) + \frac{1}{q} \|Au - g_n\|_{L^q(Q)}^q + \frac{1}{q} \|g_n - g\|_{L^q(Q)}^q \right) \\ 550 &\leq (M[g_n] + 1). \end{aligned}$$

Here, we used the convexity of the  $L^q$  norm ( $q > 1$ ) and relation (4.19) with  $v = u$ . Moreover

$$\|Au - g_n\|_{L^q(Q)} \leq \|Au - g\|_{L^q(Q)} + \|g_n - g\|_{L^q(Q)} \leq \|Au - g\|_{L^q(Q)} + q^{1/q} 2^{1/q-1}.$$

So  $M[g_n] + 1$  is bounded from above by a constant  $M_0[g]$  that does not depend on  $g_n$ . Following the same proof of Theorem 4.2, we can prove that  $(u_n)$  is uniformly bounded with respect to  $n$ , in  $\text{BV}(Q)$  and in  $L^p$ -bounded with  $1 < p \leq \frac{d+1}{d}$ . Therefore, we have that  $u_n \rightarrow \tilde{u}$  in  $L^1(Q)$ ,  $u_n \xrightarrow{w} \tilde{u}$  in  $L^p(Q)$ , with  $1 < p \leq \frac{d+1}{d}$ . It remains to show that  $\tilde{u}$  is a minimizer of  $(\mathcal{P})$ . Theorem 4.1 yields that

$$(F_\lambda \# F_\mu)(\tilde{u}) \leq \liminf_{n \rightarrow \infty} (F_\lambda \# F_\mu)(u_n).$$

Moreover  $Au_n - g_n \rightharpoonup A\tilde{u} - g$  in  $L^q(Q)$ . Since,

$$\forall v \in \text{BV}(Q), \quad (F_\lambda \# F_\mu)(u_n) + \frac{1}{q} \|Au_n - g_n\|_{L^q(Q)}^q \leq (F_\lambda \# F_\mu)(v) + \frac{1}{q} \|Av - g_n\|_{L^q(Q)}^q$$

552 we get for every  $v \in \text{BV}(Q)$  that

$$\begin{aligned} 553 \quad (F_\lambda \# F_\mu)(\tilde{u}) + \frac{1}{q} \|A\tilde{u} - g\|_{L^q(Q)}^q &\leq \liminf_{n \rightarrow \infty} \left[ (F_\lambda \# F_\mu)(u_n) + \frac{1}{q} \|Au_n - g_n\|_{L^q(Q)}^q \right] \\ 554 &\leq \lim_{n \rightarrow \infty} (F_\lambda \# F_\mu)(v) + \frac{1}{q} \|Av - g_n\|_{L^q(Q)}^q \\ 555 &\leq (F_\lambda \# F_\mu)(v) + \frac{1}{q} \|Av - g\|_{L^q(Q)}^q. \end{aligned}$$

557 So  $\tilde{u}$  is a minimizer and we conclude with uniqueness that  $u = \tilde{u}$ .

558

$$559 \quad \text{Case (2) : } \mathcal{H}(g, \mathcal{A}u) = D_{KL}(g, \mathcal{R}u) + \mathcal{H}(g, g) = \int_\Sigma \mathcal{R}u - g \log \mathcal{R}u \, d\theta \, ds \, dt,$$

560

561 Recall that we assumed that  $g, g_n \in L^\infty(\Sigma)$ ,  $\inf_\Sigma g, \inf_\Sigma g_n > 0$ . Using (4.19), we get

$$562 \quad (4.20) \quad (F_\lambda \# F_\mu)(u_n) + \mathcal{H}(g_n, \mathcal{R}u_n) \leq (F_\lambda \# F_\mu)(u) + \mathcal{H}(g_n, \mathcal{R}u).$$

As

$$\mathcal{H}(g_n, \mathcal{R}u) = \int_\Sigma \mathcal{R}u - g_n \log \mathcal{R}u \, d\theta \, ds \, dt,$$

$g_n \rightarrow g$  in  $L^\infty(\Sigma)$  and  $\log(\mathcal{R}u) \in L^1(\Sigma)$  then

$$\lim_{n \rightarrow \infty} \mathcal{H}(g_n, \mathcal{R}u) = \mathcal{H}(g, \mathcal{R}u).$$

In particular, there exists a constant  $C$  only dependent on  $g$  and  $u$  such that

$$\forall n \in \mathbb{N} \quad \mathcal{H}(g_n, \mathcal{R}u) \leq C.$$

563 Using (4.20), we get

$$564 \quad (F_\lambda \# F_\mu)(u_n) + \mathcal{H}(g_n, \mathcal{R}u_n) \leq (F_\lambda \# F_\mu)(u) + \mathcal{H}(g_n, \mathcal{R}u) \leq (F_\lambda \# F_\mu)(u) + C.$$

Again, we can use estimates as in Theorem 4.2 Case (2), with  $M_0 = (F_\lambda \# F_\mu)(u) + C$  that does not depend on  $n$ . Therefore,  $u_n$  is bounded in  $L^p(Q)$  with  $1 < p \leq \frac{d+1}{d}$  by a constant depending on  $\|g_n\|_{L^1(\Sigma)}$ . This bound is uniform with respect to  $n$  since  $\|g_n\|_{L^\infty(\Sigma)}$  (and thus  $\|g_n\|_{L^1(\Sigma)}$ ) is bounded. As before,  $u_n$  is bounded in  $BV(Q)$  and there exists  $\tilde{u} \in BV(Q)$  such that  $u_n \rightarrow \tilde{u}$  in  $L^1(Q)$ . Hence,  $\mathcal{R}u_n \rightarrow \mathcal{R}\tilde{u}$  in  $L^1(\Sigma)$  as well as pointwise convergent almost everywhere in  $\Sigma$ . By Fatou's Lemma applied to the sequence  $(\mathcal{R}u_n - g_n \log \mathcal{R}u_n)_n$ , we obtain

$$\mathcal{H}(g, \mathcal{R}\tilde{u}) \leq \liminf_{n \rightarrow \infty} \mathcal{H}(g_n, \mathcal{R}u_n).$$

566 Similarly to the previous case, we get for every  $v \in BV(Q)$ ,  $v \geq 0$  that

$$\begin{aligned} 567 \quad (F_\lambda \# F_\mu)(\tilde{u}) + \mathcal{H}(g, \mathcal{R}\tilde{u}) &\leq \liminf_{n \rightarrow \infty} (F_\lambda \# F_\mu)(u_n) + \mathcal{H}(g_n, \mathcal{R}u_n) \\ 568 &\leq \lim_{n \rightarrow \infty} (F_\lambda \# F_\mu)(v) + \mathcal{H}(g_n, \mathcal{R}v) \\ 569 \quad &\leq (F_\lambda \# F_\mu)(v) + \mathcal{H}(g, \mathcal{R}v). \end{aligned}$$

571 By uniqueness, we conclude that  $\tilde{u} = u$  is the minimizer of  $(\mathcal{P})$ . □

572 **4.3. An equivalent formulation.** We end this section by providing an equivalent formu-  
573 lation for  $(\mathcal{P})$  that may be useful for numerical computations. The key tool is the exactness  
574 of the inf-convolution operator. The original problem  $(\mathcal{P})$  also reads

$$575 \quad (\mathcal{P}') \quad \inf_{(u,v) \in BV(Q) \times BV(Q)} \mathcal{H}(g, \mathcal{A}u) + F_\lambda(u - v) + F_\mu(v).$$

576

577 **Theorem 4.5 (Equivalence).** Assume that  $\lambda$  and  $\mu$  verify (3.4) and there exists  $\kappa > 0$  such  
 578 that  $\mu_2 = \kappa\lambda_2$ .

579 1. If  $(\mathbf{u}, \mathbf{v})$  is a solution of  $(\mathcal{P}')$ , then  $\mathbf{u}$  is a solution of  $(\mathcal{P})$  and

$$580 \quad (4.21) \quad F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) = F_\lambda \# F_\mu(\mathbf{u}) = \inf_{v \in \text{BV}(Q)} \{F_\lambda(\mathbf{u} - v) + F_\mu(v)\}.$$

581 2. If  $\mathbf{u}$  is a solution of  $(\mathcal{P})$  and equation (4.21) is verified for some  $\mathbf{v} \in \text{BV}(Q)$ , then  $(\mathbf{u}, \mathbf{v})$   
 582 is a solution of  $(\mathcal{P}')$

583 *Proof.* Assume that  $(\mathbf{u}, \mathbf{v})$  is a solution to  $(\mathcal{P}')$ . Then, for every  $(u, v) \in \text{BV}(Q) \times \text{BV}(Q)$   
 584 we have

$$585 \quad (4.22) \quad F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) + \mathcal{H}(g, \mathcal{A}\mathbf{u}) \leq F_\lambda(u - v) + F_\mu(v) + \mathcal{H}(g, \mathcal{A}u) .$$

Taking  $u = \mathbf{u}$  gives

$$\forall v \in \text{BV}(Q), \quad F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) \leq F_\lambda(\mathbf{u} - v) + F_\mu(v) ,$$

that is  $F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) = F_\lambda \# F_\mu(\mathbf{u})$ . Let us fix  $u \in \text{BV}(Q)$ . Using (4.22), we obtain

$$\forall v \in \text{BV}(Q), \quad F_\lambda \# F_\mu(\mathbf{u}) + \mathcal{H}(g, \mathcal{A}\mathbf{u}) \leq F_\lambda(u - v) + F_\mu(v) + \mathcal{H}(g, \mathcal{A}u),$$

which results to

$$F_\lambda \# F_\mu(\mathbf{u}) + \mathcal{H}(g, \mathcal{A}\mathbf{u}) \leq \left( \inf_{v \in \text{BV}(Q)} F_\lambda(u - v) + F_\mu(v) \right) + \mathcal{H}(g, \mathcal{A}u) = F_\lambda \# F_\mu(u) + \mathcal{H}(g, \mathcal{A}u).$$

586 Therefore,  $\mathbf{u}$  is a solution to  $(\mathcal{P})$ .

587 Conversely, assume  $\mathbf{u}$  is a solution to  $(\mathcal{P})$ . As  $F_\lambda \# F_\mu$  is exact at  $\mathbf{u}$ , there exists  $\mathbf{v} \in \text{BV}(Q)$   
 588 such that  $F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) = F_\lambda \# F_\mu(\mathbf{u})$ . Then, for every  $(u, v) \in \text{BV}(Q) \times \text{BV}(Q)$

$$589 \quad F_\lambda(\mathbf{u} - \mathbf{v}) + F_\mu(\mathbf{v}) + \mathcal{H}(g, \mathcal{A}\mathbf{u}) = F_\lambda \# F_\mu(\mathbf{u}) + \mathcal{H}(g, \mathcal{A}\mathbf{u}) \leq F_\lambda \# F_\mu(u) + \mathcal{H}(g, \mathcal{A}u) \\ 590 \quad \leq F_\lambda(u - v) + F_\mu(v) + \mathcal{H}(g, \mathcal{A}u) .$$

591 This proves that  $(\mathbf{u}, \mathbf{v})$  is a solution to  $(\mathcal{P}')$ . □

593 **5. Optimality conditions.** In the final section of this paper, we deal with the optimality  
 594 conditions of  $(\mathcal{P})$ . Optimality conditions are useful since they provide qualitative information  
 595 on the solution of the minimization problem. In many cases, they are a useful tool to prove  
 596 convergence of the algorithms and get error estimates independent on the discretization grid.  
 597 Here, we use standard duality techniques based on the convex conjugate and the subdifferential  
 598 of a functional in order to characterize the solutions. However, as we often deal with the dual  
 599 of the underlying space, we prefer to use a reflexive framework since the dual of  $\text{BV}(Q)$  is  
 600 not easy to handle. Therefore we choose  $p$  with  $1 \leq p < \frac{d+1}{d}$ , so that  $\text{BV}(Q)$  is compactly  
 601 embedded in  $L^p(Q)$ .

We denote  $\langle \cdot, \cdot \rangle_{p', p}$  the duality product between  $L^p(Q)$  and its dual  $L^{p'}(Q)$  with  $\frac{1}{p} + \frac{1}{p'} = 1$   
 and

$$\forall u \in L^p(Q), \forall v \in L^{p'}(Q), \quad \langle v, u \rangle_{p', p} = \int_Q u(t, x) v(t, x) dt dx .$$

602 We start by extending  $\Phi_{\alpha_1}$ ,  $\Psi_{\alpha_2}$  and  $F_{\alpha}$  from their respective domains to  $L^p(Q)$  as follows:

$$603 \quad \tilde{\Phi}_{\alpha_1}(u) = \begin{cases} \Phi_{\alpha_1}(u) & \text{if } u \in L^1(\mathcal{T}; \text{BV}(\Omega)), \\ +\infty & \text{else,} \end{cases}, \quad \tilde{\Psi}_{\alpha_2}(u) = \begin{cases} \Psi_{\alpha_2}(u) & \text{if } u \in L^1(\Omega; \text{BV}(\mathcal{T})), \\ +\infty & \text{else,} \end{cases},$$

$$604 \quad \tilde{F}_{\alpha}(u) = \begin{cases} F_{\alpha}(u) & \text{if } u \in \text{BV}(Q), \\ +\infty & \text{if } u \in L^p(Q) \setminus \text{BV}(Q). \end{cases}$$

604 We define the *extended* problem as

$$605 \quad (\mathcal{P}^*) \quad \inf_{u \in L^p(Q)} \mathcal{H}(g, \mathcal{A}u) + (\tilde{F}_{\lambda} \# \tilde{F}_{\mu})(u).$$

606 With the definition of  $\tilde{F}_{\alpha}$ , it is clear that problems  $(\mathcal{P})$  and  $(\mathcal{P}^*)$  have the same solution set.  
607 So, we look for optimality conditions for  $(\mathcal{P}^*)$ . It is obvious that the lower semicontinuity for  
608 the extended regularizing terms as in Proposition 3.1 is still valid. Moreover,  $\tilde{\Phi}_{\alpha_1}$ ,  $\tilde{\Psi}_{\alpha_2}$  and  
609  $\tilde{F}_{\alpha}$  are convex as extensions of convex functions by  $+\infty$ . This may be summarized in the  
610 following corollary:

611 **Corollary 5.1.** *Let  $\alpha = (\alpha_1, \alpha_2)$  that satisfies (3.4). The functionals  $\tilde{\Phi}_{\alpha_1}$ ,  $\tilde{\Psi}_{\alpha_2}$  and  $\tilde{F}_{\alpha}$  are*  
612 *convex and lower semicontinuous on  $L^p(Q)$ .*

613 We next investigate the Fenchel conjugates of the corresponding regularizing terms and  
614 focus on the characterization of the subdifferential of  $\tilde{F}_{\lambda} \# \tilde{F}_{\mu} + \mathcal{H}(g, \mathcal{A}\cdot)$ .

615 **5.1. Fenchel conjugate of  $\tilde{F}_{\lambda} \# \tilde{F}_{\mu}$ .** One way to derive the optimality conditions of  $(\mathcal{P}^*)$ ,  
616 is by computing the subdifferentials of each term. A useful tool to achieve this goal is to  
617 compute the conjugate functionals. We start with the following theorem (see [5, Theorem  
618 9.5.1.]).

**Theorem 5.1.** *If  $V$  is a normed space with dual space  $V'$ , and  $f : V \rightarrow \mathbb{R} \cup \{+\infty\}$  is a lower semicontinuous convex and proper function, then*

$$\forall (u, u^*) \in V \times V' \quad u^* \in \partial f(u) \iff u \in \partial f^*(u^*),$$

where  $f^*$  is the Fenchel conjugate of  $f$  and the subdifferential of  $f$  at  $u$  is

$$\partial f(u) = \{u^* \in V^* \mid \forall v \in V, f(v) - f(u) \geq \langle u^*, v - u \rangle_{V', V}\}.$$

The first step is to compute the Fenchel conjugate of the regularizing term  $\tilde{F}_{\lambda} \# \tilde{F}_{\mu}$  starting by  $\tilde{F}_{\lambda}$ . Let us focus on the computation of the Fenchel-conjugate of  $\tilde{\Phi}_{\lambda}$ . We consider the set

$$\mathcal{K}_x := \left\{ \xi = \text{div}_x \varphi \mid \varphi \in L^{\infty}(\mathcal{T}; \mathcal{C}_c^1(\Omega, \mathbb{R}^d)), \|\varphi\|_{\infty} \leq 1 \right\} \subset L^{\infty}(Q).$$

We have the following lemma that provides a relation with the sets defined in (2.3). Let us define the injection  $\Upsilon$  from the space of functions defined almost everywhere on  $\Omega$  to the space of functions defined almost everywhere on  $\mathcal{T} \times \Omega$  as following: for every function  $\phi$  defined a.e. on  $\Omega$ ,  $\Upsilon(\phi) = \psi$  is defined a.e.; on  $\mathcal{T} \times \Omega$  with

$$\psi(t, x) = \phi(x), \text{ a.e. on } \mathcal{T} \times \Omega.$$

619

620 **Lemma 5.1.** *We have  $\Upsilon(K_x) \subset \mathcal{K}_x$ , where  $K_x$  is given by (2.3). Conversely, any  $\xi \in \mathcal{K}_x$*   
 621 *verifies  $\xi(t, \cdot) \in K_x$ , for almost every  $t \in \mathcal{T}$ .*

622 *Proof.* Let be  $\xi \in \mathcal{K}_x$ . There exists  $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^d)$  such that  $\xi = \operatorname{div}_x \varphi$  and  $\|\varphi\|_{\infty, x} \leq 1$ .  
 623 Let  $\psi = \Upsilon(\phi) \in L^\infty(\mathcal{T}; \mathcal{C}_c^1(\Omega, \mathbb{R}^d))$ . Then  $\|\psi\|_\infty \leq 1$  and  $\Upsilon(\xi) \in \mathcal{K}_x$ .  $\square$   
 624

**Theorem 5.2 ( $\tilde{\Phi}_\alpha$  Conjugate).** *For every function  $\alpha$  that satisfies (3.4), we have*

$$\tilde{\Phi}_\alpha^* = \mathbb{1}_{\alpha\overline{\mathcal{K}_x}}$$

625 *where,  $\mathbb{1}_C$  is the indicator function of the set  $C$  and  $\overline{\mathcal{K}_x}$  is the  $L^{p'}(Q)$ -closure of  $\mathcal{K}_x$ .*

626 *Proof.* Note that for every  $u^* \in L^{p'}(Q)$ ,

$$627 \quad (5.1) \quad \tilde{\Phi}_\alpha^*(u^*) = \sup_{v \in L^p(Q)} \langle u^*, v \rangle_{p', p} - \tilde{\Phi}_\alpha(v) = \sup_{v \in \operatorname{BV}(Q)} \langle u^*, v \rangle_{p', p} - \Phi_\alpha(v).$$

Let  $\xi \in \mathcal{K}_x$ , then  $\xi(t, \cdot) \in K_x$  for almost every  $t \in \mathcal{T}$  and (2.2) gives

$$\int_{\Omega} \xi(t, x) u(t, x) \, dx \leq \sup_{\zeta \in K_x} \int_{\Omega} \zeta(x) u(t, x) \, dx = \operatorname{TV}_x(u)(t)$$

629 using (3.2), we obtain that

$$630 \quad (5.2) \quad \sup_{\xi \in \alpha\mathcal{K}_x} \langle \xi, u \rangle_{p', p} = \sup_{\xi \in \alpha\mathcal{K}_x} \int_0^T \int_{\Omega} \xi(t, x) u(t, x) \, dx \, dt \leq \Phi_\alpha(u).$$

631 As  $\tilde{\Phi}_\alpha$  is positively homogeneous, then  $\tilde{\Phi}_\alpha^*$  is the indicator of some closed subset  $\tilde{\mathcal{K}}$  of  $L^{p'}(Q)$   
 632 (Corollary 13.2.1 of [35]).

633 • We first prove that  $\alpha\overline{\mathcal{K}_x} \subset \tilde{\mathcal{K}}$ . Let  $u^*$  be in  $\alpha\mathcal{K}_x$ . Using (5.1), (5.2) we have that for any  
 634  $v \in \operatorname{BV}(Q)$ ,  $\Phi_\alpha(v) \geq \langle u^*, v \rangle_{p', p}$  and so  $\tilde{\Phi}_\alpha^*(u^*) \leq 0$ . As  $\tilde{\Phi}_\alpha^*$  is an indicator function this means  
 635 that  $\tilde{\Phi}_\alpha^*(u^*) = 0$ . So  $u^* \in \tilde{\mathcal{K}}$  and  $\alpha\mathcal{K}_x \subset \tilde{\mathcal{K}}$ . As  $\tilde{\mathcal{K}}$  is  $L^{p'}(Q)$ -closed this gives  $\alpha\overline{\mathcal{K}_x} \subset \tilde{\mathcal{K}}$ .

636 • Let us prove the converse inclusion. Assume there exists  $u^* \in \tilde{\mathcal{K}}$  such that  $u^* \notin \alpha\overline{\mathcal{K}_x}$ . One  
 637 can separate  $u^*$  and  $\alpha\overline{\mathcal{K}_x}$ , see [13]: there exists  $\omega \in \mathbb{R}$  and  $u_0 \in L^p(Q)$  such that

$$638 \quad \langle u_0, u^* \rangle_{p, p'} = \langle u^*, u_0 \rangle_{p', p} > \omega \geq \sup_{v^* \in \alpha\overline{\mathcal{K}_x}} \langle v^*, u_0 \rangle_{p', p}$$

$$639 \quad (5.3) \quad \Rightarrow \sup_{v^* \in \alpha\overline{\mathcal{K}_x}} \langle v^* - u^*, u_0 \rangle_{p', p} < 0.$$

On the other hand, since  $\tilde{\Phi}_\alpha$  is convex and lower semicontinuous with respect to the  $L^p$ -  
 topology, then by Fenchel-Moreau theorem we have that  $\tilde{\Phi}_\alpha^{**} = \tilde{\Phi}_\alpha$ . So, for all  $u \in \operatorname{BV}(Q)$ ,

$$\Phi_\alpha(u) = \sup_{v^* \in L^{p'}(Q)} \langle v^*, u \rangle_{p', p} - \tilde{\Phi}_\alpha^*(v^*) = \sup_{v^* \in \tilde{\mathcal{K}}} \langle v^*, u \rangle_{p', p},$$

641 since  $\tilde{\Phi}_\alpha^*$  is the indicator of  $\tilde{\mathcal{K}}$ . In particular, as  $u^* \in \tilde{\mathcal{K}}$

$$642 \quad (5.4) \quad \Phi_\alpha(u) \geq \langle u^*, u \rangle_{p', p}.$$

Let us fix  $t \in \mathcal{T}$ , then

$$\forall \xi \in K_x, \quad \alpha(t)\xi(x)u(t, x) \leq \sup_{\zeta \in K_x} \alpha(t)\zeta(x)u(t, x) \quad \text{a.e. } x \in \Omega,$$

643 and taking the supremum we have that

$$\begin{aligned} 644 \quad \sup_{\xi \in K_x} \int_{\Omega} \alpha(t)\xi(x)u(t, x) \, dx &\leq \int_{\Omega} \sup_{\zeta \in K_x} \alpha(t)\zeta(x)u(t, x) \, dx, \\ 645 \quad \text{TV}_x(\alpha u)(t) &\leq \int_{\Omega} \sup_{\zeta \in K_x} \alpha(t)\zeta(x)u(t, x) \, dx. \\ 646 \end{aligned}$$

647 We integrate over the time domain  $\mathcal{T}$  and subtract both sides by  $\langle u^*, u \rangle_{p', p}$  to recover

$$\begin{aligned} 648 \quad \int_0^T \text{TV}_x(\alpha u)(t) \, dt - \int_0^T \int_{\Omega} u^*(t, x)u(t, x) \, dx \, dt &\leq \\ 649 \quad \int_0^T \int_{\Omega} \left[ \sup_{\zeta \in K_x} \alpha(t)\zeta(x) - u^*(t, x) \right] u(t, x) \, dx \, dt. \\ 650 \end{aligned}$$

651 Then, using (5.4) and Lemma 5.1, we have that for all  $u \in \text{BV}(Q)$

$$\begin{aligned} 652 \quad 0 \leq \Phi_{\alpha}(u) - \langle u^*, u \rangle_{p', p} &\leq \int_0^T \int_{\Omega} \left[ \sup_{\zeta \in \alpha K_x} \zeta(x) - u^*(t, x) \right] u(t, x) \, dx \, dt \\ 653 \quad &\leq \int_0^T \int_{\Omega} \left[ \sup_{\xi \in \alpha K_x} \xi(t, x) - u^*(t, x) \right] u(t, x) \, dx \, dt \\ 654 \quad &\leq \int_0^T \int_{\Omega} \left[ \sup_{\xi \in \overline{\alpha K_x}} \xi(t, x) - u^*(t, x) \right] u(t, x) \, dx \, dt. \\ 655 \quad & \\ 656 \end{aligned}$$

Hence, this implies

$$\forall u \in \text{BV}(Q), \quad \int_0^T \int_{\Omega} \left( \sup_{\xi \in \overline{\alpha K_x}} \xi(t, x) - u^*(t, x) \right) u(t, x) \, dx \, dt \geq 0.$$

Next, choosing  $-u$  instead of  $u$  we get

$$\forall u \in \text{BV}(Q), \quad \int_0^T \int_{\Omega} \left( \sup_{\xi \in \overline{\alpha K_x}} \xi(t, x) - u^*(t, x) \right) u(t, x) \, dx \, dt = 0.$$

Therefore  $\sup_{\xi \in \overline{\alpha K_x}} \xi - u^* = 0 \in \text{BV}'(Q)$ . Next, for every  $u \in \text{L}^p(Q)$  and for every  $\xi \in \overline{\alpha K_x}$  we have

$$\langle \xi - u^*, u \rangle_{p', p} \leq \left\langle \sup_{\xi \in \overline{\alpha K_x}} \xi - u^*, u \right\rangle_{p', p} = 0,$$

since  $\overline{\alpha\mathcal{K}_x} \subset L^{p'}(Q)$ . Once again, using  $-u$  we obtain for every  $u \in L^p(Q)$

$$\forall \xi \in \overline{\alpha\mathcal{K}_x}, \quad \langle \xi - u^*, u \rangle_{p',p} = 0,$$

that is

$$\sup_{\xi \in \overline{\alpha\mathcal{K}_x}} \langle \xi - u^*, u \rangle_{p',p} = 0,$$

since  $\overline{\alpha\mathcal{K}_x}$  is a closed subset of  $L^{p'}(Q)$ . As a consequence, we get

$$\sup_{\xi \in \overline{\alpha\mathcal{K}_x}} \langle \xi - u^*, u_0 \rangle_{p',p} = 0.$$

657 which is a contradiction by (5.3). □

658 The following is the analogous result of the previous theorem for the  $\tilde{\Psi}_\alpha$  functional and can  
659 be proved similarly.

**Theorem 5.3 ( $\tilde{\Psi}_\alpha$  Conjugate).** *For every function  $\alpha$  that satisfies (3.4), we have*

$$\tilde{\Psi}_\alpha^* = \mathbb{1}_{\overline{\alpha\mathcal{K}_t}}, \quad \text{where } \mathcal{K}_t := \left\{ \xi = \frac{d\psi}{dt} \mid \psi \in L^\infty(\Omega, \mathcal{C}_c^1(\mathcal{T}, \mathbb{R})), \|\psi\|_\infty \leq 1 \right\}.$$

Using the above theorems, we are able to compute the convex conjugate of the extended spatial-temporal total variation defined in (3.5). We use the following results for the convex conjugate of the infimal convolution and the convex conjugate of the sum, see [5, Chapter 9.4], i.e., for two proper, closed, convex functionals  $\phi, \psi$  we have

$$(\phi \# \psi)^* = \phi^* + \psi^* \quad \text{and} \quad (\phi + \psi)^* = (\phi^* \# \psi^*)^{**}.$$

**Corollary 5.2.** *For every  $\alpha$  that satisfies (3.4), we have that*

$$\tilde{F}_\alpha^* = \mathbb{1}_{\overline{\mathcal{K}_\alpha}} \quad \text{with} \quad \mathcal{K}_\alpha = \alpha_1 \overline{\mathcal{K}_x} + \alpha_2 \overline{\mathcal{K}_t}.$$

*Proof.* As  $\tilde{F}_\alpha = \tilde{\Phi}_{\alpha_1} + \tilde{\Psi}_{\alpha_2}$  and  $\tilde{\Phi}_{\alpha_1}, \tilde{\Psi}_{\alpha_2}$  are convex, lower semicontinuous, we have

$$\tilde{F}_\alpha^* = (\tilde{\Phi}_{\alpha_1} + \tilde{\Psi}_{\alpha_2})^* = (\tilde{\Phi}_{\alpha_1}^* \# \tilde{\Psi}_{\alpha_2}^*)^{**} = (\mathbb{1}_{\overline{\alpha_1\mathcal{K}_x}} \# \mathbb{1}_{\overline{\alpha_2\mathcal{K}_t}})^{**} = (\mathbb{1}_{\overline{\alpha_1\mathcal{K}_x + \alpha_2\mathcal{K}_t}})^{**} = (\mathbb{1}_{\mathcal{K}_\alpha})^{**},$$

660 where  $\mathcal{K}_\alpha = \alpha_1 \overline{\mathcal{K}_x} + \alpha_2 \overline{\mathcal{K}_t}$ . Moreover, one has that  $(\mathbb{1}_{\mathcal{K}_\alpha})^{**} = \mathbb{1}_{\overline{\mathcal{K}_\alpha}}$ , since the  $(L^{p'})$  closure  
661  $\overline{\mathcal{K}_\alpha}$  of  $\mathcal{K}_\alpha$  is convex, see [35, Chapter 13]. □

662

**Corollary 5.3 ( $\tilde{F}_\lambda \# \tilde{F}_\mu$  Conjugate).** *For every  $\lambda, \mu$  that satisfy (3.4), we have*

$$(\tilde{F}_\lambda \# \tilde{F}_\mu)^* = \mathbb{1}_{\overline{\mathcal{K}_\lambda \cap \mathcal{K}_\mu}},$$

663 where  $\mathcal{K}_\lambda, \mathcal{K}_\mu$  are the corresponding sets defined in Corollary 5.2.

664 We have computed the convex conjugate of our proposed regularizer and we proceed now with  
665 the optimality conditions of (P).

**5.2. Optimality conditions for  $(\mathcal{P})$ .** Since the problem  $(\mathcal{P}^*)$  is convex we have that  $\mathbf{u}$  is the solution if and only if  $0 \in \partial\mathcal{E}(\mathbf{u})$  where

$$\mathcal{E}(u) := (\tilde{F}_\lambda \# \tilde{F}_\mu)(u) + \mathcal{H}(g, Au).$$

666 We use the following result that allows to estimate the subdifferential of the sum of two  
667 functionals, see [5, Theorem 9.5.4].

668 **Theorem 5.4.** *Let  $(V, \|\cdot\|)$  be a normed space and let  $f, h : V \rightarrow \mathbb{R} \cup \{+\infty\}$  be two lower  
669 semicontinuous, convex and proper functions.*

670 (a) *The following inclusion is always true:  $\partial f + \partial h \subset \partial(f + h)$ .*

671 (b) *If  $f$  is finite and continuous at a point of  $\text{dom } h$ , then we have:  $\partial f + \partial h = \partial(f + h)$ .*

**5.2.1. Case (1).** In this subsection we focus on the first case where the  $L^q$  fidelity term is  $\mathcal{H}(g, Au) = \frac{1}{q} \|Au - g\|_{L^q(Q)}^q$  with  $1 \leq q < +\infty$  and  $A$  satisfies assumption (3.9). Clearly,  $\text{dom } \tilde{F}_\lambda \# \tilde{F}_\mu = \text{BV}(Q)$ ,  $\text{dom } \mathcal{H}(g, Au) = L^p(Q)$  and  $u \rightarrow \mathcal{H}(g, Au)$  is  $L^p$  continuous at  $0 \in \text{BV}(Q)$ . Therefore,

$$\partial\mathcal{E}(u) = \partial\tilde{F}_\lambda \# \tilde{F}_\mu(u) + \partial\mathcal{H}(g, Au).$$

672 Any  $u^*$  of  $\partial\mathcal{E}(u)$  writes  $u^* = u_1^* + u_2^*$  where  $u_1^* \in \partial\tilde{F}_\lambda \# \tilde{F}_\mu(u)$  and  $u_2^* \in \partial\mathcal{H}(g, Au)$ . In the  
673 sequel, we characterize the elements  $u_1^*, u_2^*$ . Starting with the subdifferential of  $\tilde{F}_\lambda \# \tilde{F}_\mu$ , it is  
674 easy to check that for every  $u \in \text{BV}(Q) \hookrightarrow L^p(Q)$ , we get

$$675 \quad (5.5) \quad u_1^* \in \partial\tilde{F}_\lambda \# \tilde{F}_\mu(u) \iff u_1^* \in \overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu} \text{ and } \forall v^* \in \overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu}, \langle u, v^* - u_1^* \rangle_{p,p'} \leq 0,$$

where  $\overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu}$  is a closed convex subset of  $L^{p'}(Q)$ . Indeed, we use Theorem 5.1, Corollary 5.3 and that  $\tilde{F}_\lambda \# \tilde{F}_\mu$  is convex and lower semicontinuous, to get

$$u \in \partial(\tilde{F}_\lambda \# \tilde{F}_\mu)^*(u_1^*) = \partial\mathbb{1}_{\overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu}}(u_1^*).$$

676 The subdifferential of the data fitting term using [20, Proposition 5.7] is

$$677 \quad (5.6) \quad \partial\mathcal{H}(g, Au) = \begin{cases} A^*(Au - g)^{q-1}, & \text{if } 1 < q < \infty \\ \left\{ A^*z, \|z\|_{L^\infty(Q)} \leq 1, z \in \text{sign}(Au - g) \right\}, & \text{if } q = 1. \end{cases}$$

Note that in the latter case one has

$$\partial(\|\cdot - g\|_{L^1(Q)})(v) = \partial(\|\cdot\|_{L^1(Q)})(v - g) = \{z \in L^\infty(Q) \mid \|z\|_{L^\infty(Q)} \leq 1, z \in \text{sign}(v - g)\}.$$

Overall, we have that

$$0 \in \partial\mathcal{E}(u) \iff \exists u^* \in \partial\mathcal{H}(g, Au) \text{ such that } -u^* \in \partial\tilde{F}_\lambda \# \tilde{F}_\mu(u)$$

678 and one concludes to the following result:

679 **Theorem 5.5.** *A function  $\mathbf{u} \in \text{BV}(Q)$  is a solution to  $(\mathcal{P})$  if and only if*

1.  $\forall v \in \overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu}, \langle \mathbf{u}, A^*(Au - g)^{q-1} - v \rangle_{p,p'} \leq 0, \text{ if } 1 < q < +\infty,$
2.  $\forall v \in \overline{\mathcal{K}_\lambda} \cap \overline{\mathcal{K}_\mu}, \langle \mathbf{u}, A^*z - v \rangle_{p,p'} \leq 0, \text{ if } q = 1 \text{ with}$   
680  $z \in L^\infty(Q), \|z\|_{L^\infty(Q)} \leq 1, z \in \text{sign}(Au - g).$



681 **5.2.2. Optimality conditions for  $(\mathcal{P})$ : case (2).** In this subsection we focus on the  
 682 Kullback-Leibler divergence see (3.18), i.e.,  $\mathcal{H}(g, \mathcal{R}u) = D_{KL}(g, \mathcal{R}u) + \mathcal{H}(g, g)$  where  $u \in \mathcal{D}$ ,  
 683 the domain of the fidelity term. We cannot follow the same strategy as before due to the  
 684 limitations of this fidelity in terms of continuity. It is known that a proper, convex, lower  
 685 semicontinuous function is continuous if and only if the interior of its domain is not empty,  
 686 i.e.,  $\text{int}(\text{dom}f) \neq \emptyset$ , see [20]. In our case the effective domain is in fact nowhere dense and  
 687  $D_{KL}(g, \mathcal{R}\cdot)$  is nowhere continuous in  $L^1(\Omega)$ , let alone in  $L^p(\Omega)$ , see [19, Remark 2.12]. More-  
 688 over,  $F_\lambda \# F_\mu$  is not continuous with respect to the  $L^p$  norm.

689 Therefore we use  $\text{BV}(Q)$  as the underlying functional space. In the sequel  $\langle \cdot, \cdot \rangle$  de-  
 690 notes the duality product between  $\text{BV}(Q)'$  and  $\text{BV}(Q)$ . We use Theorem 5.4 again with  
 691  $V = \text{BV}(Q)$ ,  $f = F_\lambda \# F_\mu$  and  $h = \mathcal{H}(g, \mathcal{R}(\cdot))$ . Indeed,  $f$  is lower semicontinuous due to  
 692 Theorem 4.1 and  $h$  due to the continuity properties of both the Radon transform and the  
 693 Kullback-Leibler divergence.

694 **Proposition 5.1.** *Assume that  $\lambda$  and  $\mu$  satisfy (3.4) and that there exists a real number*  
 695  *$\kappa > 0$  such that  $\mu_2 = \kappa\lambda_2$ . Then  $F_\lambda \# F_\mu$  is continuous on  $\text{BV}(Q)$  (and of course at any*  
 696 *element of  $\text{dom}f \cap \text{dom}h \subset \text{BV}_+(Q)$  the set of positive BV functions).*

697 *Proof.* Let  $u_1, u_2$  be in  $\text{BV}(Q)$ . As  $F_\lambda \# F_\mu$  is exact, there exists  $v_1 \in \text{BV}(Q)$  such that  
 698  $F_\lambda \# F_\mu(u_1) = F_\lambda(u_1 - v_1) + F_\mu(v_1)$ . We get

$$\begin{aligned} 699 \quad F_\lambda \# F_\mu(u_2) &= \inf_{v \in \text{BV}(Q)} F_\lambda(u_2 - v) + F_\mu(v) \leq F_\lambda(u_2 - v_1) + F_\mu(v_1) \\ 700 &\leq F_\lambda(u_2 - u_1) + F_\lambda(u_1 - v_1) + F_\mu(v_1) \\ 701 &= F_\lambda(u_2 - u_1) + F_\lambda \# F_\mu(u_1). \end{aligned}$$

Similarly

$$F_\lambda \# F_\mu(u_1) \leq F_\lambda(u_1 - u_2) + F_\lambda \# F_\mu(u_2),$$

and using Theorem 3.1

$$|F_\lambda \# F_\mu(u_1) - F_\lambda \# F_\mu(u_2)| \leq F_\lambda(u_1 - u_2) \leq C_\lambda^+ \text{TV}(\lambda_2(u_1 - u_2)).$$

703 Moreover

$$\begin{aligned} 704 \quad \text{TV}(\lambda_2(u_1 - u_2)) &\leq \|\lambda_2\|_{L^\infty(\mathcal{T})} \text{TV}(u_1 - u_2) + \left\| \lambda_2' \right\|_{L^\infty(\mathcal{T})} \|u_1 - u_2\|_{L^1(Q)} \\ 705 &\leq \|\lambda_2\|_{W^{1,\infty}(\mathcal{T})} \|u_1 - u_2\|_{\text{BV}}. \end{aligned}$$

This prove the continuity of  $F_\lambda \# F_\mu$  on  $\text{BV}(Q)$ .  $\square$

Recall that  $\mathcal{D}$ , given in (3.19), is the domain of the fidelity term. So  $\mathbf{u} \in \text{BV} \cap \mathcal{D}$  is a solution to  $(\mathcal{P})$  if and only if

$$0 \in \partial(F_\lambda \# F_\mu)(\mathbf{u}) + \partial\mathcal{H}(g, \mathcal{R}\mathbf{u}).$$

707 Equivalently, there exists  $\mathbf{u}^* \in \partial(F_\lambda \# F_\mu)(\mathbf{u})$  such that  $-\mathbf{u}^* \in \partial\mathcal{H}(g, \mathcal{R}(\cdot))(\mathbf{u})$ . As usual,  
 708 we have  $\mathbf{u}^* \in \partial(F_\lambda \# F_\mu)(\mathbf{u}) \iff \mathbf{u} \in \partial(F_\lambda \# F_\mu)^*(\mathbf{u}^*)$ . However, in this setting we are  
 709 in different topology. Though we have computed  $F_\lambda^*$  for previous case, the computation of  
 710  $F_\lambda^*$  is still challenging. Indeed, we cannot use the arguments used in Theorem 5.2 since the

711 underlying topology is now the BV one and not the  $L^p(Q)$  one any longer. In particular, we  
 712 loose reflexivity as well as an integral representation on the duality product, see [23].

Since  $F_\lambda$  is positively homogeneous functional, we know there exists a closed convex subset of  $BV'$  that we call  $\mathbf{K}_\lambda$  such that  $F_\lambda^* = \mathbb{1}_{\mathbf{K}_\lambda}(\mathbf{u}^*)$  is the indicator function of  $\mathbf{K}_\lambda$ . Unfortunately, we are not able to give an explicit description of  $\mathbf{K}_\lambda$ : we only know that  $\mathcal{K}_\lambda \subset \mathbf{K}_\lambda$ . We obtain

$$(F_\lambda \# F_\mu)^* = F_\lambda^* + F_\mu^* = \mathbb{1}_{\mathbf{K}_\lambda} + \mathbb{1}_{\mathbf{K}_\mu} = \mathbb{1}_{\mathbf{K}_\lambda \cap \mathbf{K}_\mu}.$$

Therefore,

$$\mathbf{u}^* \in \partial(F_\lambda \# F_\mu)(\mathbf{u}) \iff \mathbf{u}^* \in \mathbf{K}_\lambda \cap \mathbf{K}_\mu \text{ and } \forall w^* \in \mathbf{K}_\lambda \cap \mathbf{K}_\mu \quad \langle \mathbf{u}, w^* - \mathbf{u}^* \rangle \leq 0.$$

Next, we compute  $\partial\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u})$ . Let be  $w \in BV(Q) \cap \mathcal{D}$ :

$$-\mathbf{u}^* \in \partial\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) \implies \forall s > 0 \quad \frac{\mathcal{H}(g, \mathcal{R}(\mathbf{u} + s\mathbf{w})) - \mathcal{H}(g, \mathcal{R}\mathbf{u})}{s} \geq -\langle \mathbf{u}^*, \mathbf{w} \rangle.$$

713 Passing to the limit as  $s \rightarrow 0$  gives  $\langle \nabla\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) + \mathbf{u}^*, \mathbf{w} \rangle \geq 0$ .

Conversely, let us assume that  $\langle \nabla\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) + \mathbf{u}^*, \mathbf{w} \rangle \geq 0$  for every  $w \in BV \cap \mathcal{D}$  and prove that  $-\mathbf{u}^* \in \partial\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u})$  that is

$$\forall w \in BV(Q), \quad \mathcal{H}(g, \mathcal{R}(\mathbf{u} + w)) - \mathcal{H}(g, \mathcal{R}\mathbf{u}) \geq \langle (-\mathbf{u}^*), w \rangle.$$

Let be  $w \in BV(Q)$ : if  $\mathbf{u} + w \notin \mathcal{D}$  then

$$+\infty = \mathcal{H}(g, \mathcal{R}(\mathbf{u} + w)) - \mathcal{H}(g, \mathcal{R}\mathbf{u}) \geq \langle (-\mathbf{u}^*), w \rangle.$$

Otherwise, by convexity

$$\mathcal{H}(g, \mathcal{R}(\mathbf{u} + w)) - \mathcal{H}(g, \mathcal{R}\mathbf{u}) \geq \langle \nabla\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}), w \rangle \geq \langle (-\mathbf{u}^*), w \rangle.$$

Therefore

$$-\mathbf{u}^* \in \partial\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) \iff \forall w \in BV(Q) \cap \mathcal{D}, \quad \langle \nabla\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) + \mathbf{u}^*, w \rangle \geq 0.$$

A short computation gives

$$\nabla\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) = \mathcal{R}^* \left( \mathbf{1}_\Sigma - \frac{g}{\mathcal{R}\mathbf{u}} \right).$$

Finally ,

$$-\mathbf{u}^* \in \partial\mathcal{H}(g, \mathcal{R}\cdot)(\mathbf{u}) \iff \forall w \in BV(Q) \cap \mathcal{D}, \quad \left\langle \mathcal{R}^* \left( \mathbf{1}_\Sigma - \frac{g}{\mathcal{R}\mathbf{u}} \right) + \mathbf{u}^*, w \right\rangle \geq 0.$$

714 For this case, we conclude with the following optimality conditions

715 **Theorem 5.6.** *Let  $\mathbf{u} \in BV(Q) \cap \mathcal{D}$ . Then  $\mathbf{u}$  is a solution to  $(\mathcal{P})$  if and only if there exists*  
 716  *$\mathbf{u}^* \in \mathbf{K}_\lambda \cap \mathbf{K}_\mu \subset BV(Q)'$  such that*

$$717 \quad (5.7) \quad \forall w^* \in \mathbf{K}_\lambda \cap \mathbf{K}_\mu, \quad \langle \mathbf{u}, w^* - \mathbf{u}^* \rangle \leq 0,$$

$$718 \quad (5.8) \quad \forall w \in BV(Q) \cap \mathcal{D}, \quad \left\langle \mathcal{R}^* \left( \mathbf{1}_\Sigma - \frac{g}{\mathcal{R}\mathbf{u}} \right) + \mathbf{u}^*, w \right\rangle \geq 0.$$

719

720 **Remark 5.1.** *The difficulties met in order to establish the optimality conditions are closely*  
 721 *related to the so-called two-norm discrepancy in control theory (see [14] for example). We have*  
 722 *to deal with both the BV- norm and the  $L^p$ -norm. The qualification condition that we need to*  
 723 *describe the subdifferentials is easy to satisfy with the BV-norm. However, the computation of*  
 724 *the conjugate functions cannot be explicit within a non reflexive framework. On the contrary,*  
 725 *the use of  $L^p$ -norm leads to a nice description of conjugate functions while the splitting of the*  
 726 *differential cannot be done. In a discrete setting, these difficulties disappear of course.*

727 **6. Conclusion.** We perform a thorough analysis on the proposed spatial-temporal infimal-  
 728 convolution regularizer under time dependent weight parameters. It acts in a separate mode  
 729 on the spatial and temporal domains and it can be applied to a wide range of problems such as  
 730 denoising, deblurring and emission tomography with different kind of noise (impulse, gaussian  
 731 or Poisson). We focus on the well-posedness of the proposed minimization problem and provide  
 732 existence, uniqueness and stability results into a very general framework. We further derive  
 733 the optimality conditions using standard tools from duality theory. However, we have still  
 734 to focus in depth on the characterization of the sets  $K_\lambda$  to have a clear insight of the dual  
 735 variables. This implies that we have to deal with the dual of the BV space and use some  
 736 integral representations as in [23]. Another issue is to describe carefully the discretization  
 737 process and the dual problem in an appropriate way, especially with respect to isotropic or  
 738 anisotropic spatial-temporal discrete norms. Finally, in a forthcoming paper, we shall perform  
 739 numerics, especially for PET reconstruction, and compare this model to those that can be  
 740 found in the literature such as [25].

741 **Acknowledgments.** We would like to thank the two anonymous referees who have helped  
 742 us to improve this paper with their remarks and suggestions.

743

## REFERENCES

- 744 [1] R. ACAR AND C. R. VOGEL, *Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems,*  
 745 *Inverse Problems*, 10 (1994), pp. 1217–1229. <http://dx.doi.org/10.1088/0266-5611/10/6/003>.  
 746 [2] M. AMAR AND G. BELLETTINI, *A notion of total variation depending on a metric with discontinuous*  
 747 *coefficients*, *Annales de l'I.H.P. Analyse non linéaire*, 11 (1994), pp. 91–133. [http://eudml.org/doc/](http://eudml.org/doc/78325)  
 748 [78325](http://eudml.org/doc/78325).  
 749 [3] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity*  
 750 *problems*, *Oxford Mathematical Monographs*, The Clarendon Press Oxford University Press, New  
 751 *York*, 2000.  
 752 [4] J. APPELL, J. BANAS, AND N. MERENTES DÍAZ, *Bounded Variation and Around*, *De Gruyter series in*  
 753 *nonlinear analysis and applications*, The Clarendon Press Oxford University Press, Berlin, 2013.  
 754 [5] H. ATTOUCH, G. BUTTAZZO, AND G. MICHAILLE, *Variational analysis in Sobolev and BV spaces*, vol. 6  
 755 *of MPS/SIAM Series on Optimization*, Society for Industrial and Applied Mathematics (SIAM),  
 756 *Philadelphia, PA*, 2006. Applications to PDEs and optimization.  
 757 [6] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert*  
 758 *Spaces*, Springer Publishing Company, Incorporated, 1st ed., 2011.  
 759 [7] M. BENNING, C.-B. SCHÖNLIEB, T. VALKONEN, AND V. VLACIC, *Explorations on anisotropic regulari-*  
 760 *zation of dynamic inverse problems by bilevel optimisation*, arXiv preprint, (2016). [arXiv:1602.01278](https://arxiv.org/abs/1602.01278).  
 761 [8] M. BERGOUNIOUX AND L. PIFFET, *A Second-Order Model for Image Denoising*, *Set-Valued and Varia-*  
 762 *tional Analysis*, 18 (2010), pp. 277–306. <http://dx.doi.org/10.1007/s11228-010-0156-6>.  
 763 [9] M. BERGOUNIOUX AND E. TRÉLAT, *A variational method using fractional order Hilbert spaces for to-*  
 764 *mographic reconstruction of blurred and noised binary images*, *Journal of Functional Analysis*, 259

- 765 (2010), pp. 2296 – 2332. <http://dx.doi.org/10.1016/j.jfa.2010.05.016>.
- 766 [10] M. BERTERO, H. LANTÉRI, AND L. ZANNI, *Iterative image reconstruction: a point of view*, Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT), 7 (2008),  
767 pp. 37–63.
- 769 [11] J. M. BORWEIN AND A. S. LEWIS, *Convergence of best entropy estimates*, SIAM Journal on Optimization,  
770 1 (1991), pp. 191–205, doi:10.1137/0801014. <http://dx.doi.org/10.1137/0801014>.
- 771 [12] K. BREDIES, K. KUNISCH, AND T. POCK, *Total Generalized Variation*, SIAM Journal on Imaging  
772 Sciences, 3 (2010), pp. 492–526. <http://dx.doi.org/10.1137/090769521>.
- 773 [13] H. BREZIS, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer-Verlag New  
774 York, 2011. <http://dx.doi.org/10.1007/978-0-387-70914-7>.
- 775 [14] E. CASAS AND F. TRÖLTZSCH, *Second order optimality conditions and their role in PDE control*, Jahres-  
776 bericht der Deutschen Mathematiker-Vereinigung, 117 (2015), pp. 3–44.
- 777 [15] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*,  
778 Numer. Math., 76 (1997), pp. 167–188.
- 779 [16] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications*  
780 *to imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), pp. 120–145. [http://dx.doi.org/](http://dx.doi.org/10.1007/s10851-010-0251-1)  
781 [10.1007/s10851-010-0251-1](http://dx.doi.org/10.1007/s10851-010-0251-1).
- 782 [17] S. H. CHAN, R. KHOSHABEH, K. B. GIBSON, P. E. GILL, AND T. Q. NGUYEN, *An augmented lagrangian*  
783 *method for total variation video restoration*, IEEE Transactions on Image Processing, 20 (2011),  
784 pp. 3097–3111.
- 785 [18] J. A. CLARKSON AND R. C. ADAMS, *On definitions of bounded variation for functions of two variables*,  
786 Transactions of the American Mathematical Society, 35 (1933), pp. 824–854. [http://www.jstor.org/](http://www.jstor.org/stable/1989593)  
787 [stable/1989593](http://www.jstor.org/stable/1989593).
- 788 [19] B. DAN AND R. ELENA, *Bregman distances, totally convex functions, and a method for solving operator*  
789 *equations in banach spaces.*, Abstract and Applied Analysis, (2006), pp. Article ID 84919, 39 p.–Article  
790 ID 84919, 39 p., <http://eudml.org/doc/53750>.
- 791 [20] I. EKELAND AND R. TEMAM, *Convex analysis and variational problems*, SIAM, 1976.
- 792 [21] S. ESEDOGLU AND S. J. OSHER, *Decomposition of images by the anisotropic rudin-osher-fatemi model*,  
793 Communications on Pure and Applied Mathematics, 57 (2004), pp. 1609–1626. [http://dx.doi.org/10.](http://dx.doi.org/10.1002/cpa.20045)  
794 [1002/cpa.20045](http://dx.doi.org/10.1002/cpa.20045).
- 795 [22] L. C. EVANS AND R. F. GARIEPY, *Measure theory and fine properties of functions*, CRC press, 2015.
- 796 [23] N. FUSCO AND D. SPECTOR, *A remark on an integral characterization of the dual of BV*, Journal of  
797 Mathematical Analysis and Applications, (2017).
- 798 [24] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for  $L^1$  regularized problems*, SIAM Journal on  
799 Imaging Sciences, 2 (2009), pp. 323–343. <http://dx.doi.org/10.1137/080725891>.
- 800 [25] M. HOLLER AND K. KUNISCH, *On infimal convolution of TV-type functionals and applications to video*  
801 *and image reconstruction*, SIAM Journal on Imaging Sciences, 7 (2014), pp. 2258–2300. [http://dx.](http://dx.doi.org/10.1137/130948793)  
802 [doi.org/10.1137/130948793](http://dx.doi.org/10.1137/130948793).
- 803 [26] M. S. HOSSEINI AND K. N. PLATANIOTIS, *High-accuracy total variation with application to compressed*  
804 *video sensing*, IEEE Transactions on Image Processing, 23 (2014), pp. 3869–3884. [10.1109/TIP.2014.](https://doi.org/10.1109/TIP.2014.2332755)  
805 [2332755](https://doi.org/10.1109/TIP.2014.2332755).
- 806 [27] T. LE, R. CHARTRAND, AND T. J. ASAKI, *A variational approach to reconstructing images corrupted by*  
807 *poisson noise*, Journal of Mathematical Imaging and Vision, 27 (2007), pp. 257–263. [http://dx.doi.](http://dx.doi.org/10.1007/s10851-007-0652-y)  
808 [org/10.1007/s10851-007-0652-y](http://dx.doi.org/10.1007/s10851-007-0652-y).
- 809 [28] A. MARKOE, *Analytic Tomography*, Cambridge University Press, 2014. [http://dx.doi.org/10.1017/](http://dx.doi.org/10.1017/CBO9780511530012)  
810 [CBO9780511530012](http://dx.doi.org/10.1017/CBO9780511530012).
- 811 [29] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth*  
812 *Dean Jacqueline B. Lewis Memorial Lectures*, vol. 22, American Mathematical Society, 2001.
- 813 [30] Y. L. MONTAGNER, E. ANGELINI, AND J. C. OLIVO-MARIN, *Video reconstruction using compressed*  
814 *sensing measurements and 3D total variation regularization for bio-imaging applications*, in 2012 19th  
815 IEEE International Conference on Image Processing, 2012, pp. 917–920. [10.1109/ICIP.2012.6467010](https://doi.org/10.1109/ICIP.2012.6467010).
- 816 [31] D. OBERLIN AND E. STEIN, *Mapping properties of the radon transform*, Indiana Univ. Math. J., 31  
817 (1982), pp. 641–650.
- 818 [32] N. PUSTELNIK, C. CHAUX, J.-C. PESQUET, AND C. COMTAT, *Parallel algorithm and hybrid regular-*

- 819            *ization for dynamic PET reconstruction*, IEEE Medical Imaging Conference, (2010), pp. 2423–2427.  
820            <https://hal-upec-upem.archives-ouvertes.fr/hal-00733493>.
- 821 [33] E. RESMERITA AND R. S. ANDERSSON, *Joint additive Kullback-Leibler residual minimization and reg-*  
822 *ularization for linear inverse problems*, Mathematical Methods in the Applied Sciences, 30 (2007),  
823 pp. 1527–1544.
- 824 [34] W. RING, *Structural Properties of Solutions to Total Variation Regularisation Problems*, ESAIM: Mathe-  
825 *matical Modelling and Numerical Analysis*, 34 (2000), pp. 799–810. [http://dx.doi.org/10.1051/m2an:](http://dx.doi.org/10.1051/m2an:2000104)  
826 [2000104](http://dx.doi.org/10.1051/m2an:2000104).
- 827 [35] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, 1970.
- 828 [36] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*,  
829 *Physica D*, 60 (1992), pp. 259–268.
- 830 [37] A. SAWATZKY, C. BRUNE, T. KÖSTERS, F. WÜBBELING, AND M. BURGER, *EM-TV methods for inverse*  
831 *problems with poisson noise*, in *Level set and PDE based reconstruction methods in imaging*, Springer,  
832 2013, pp. 71–142.
- 833 [38] H. SCHAEFFER, Y. YANG, AND S. OSHER, *Space-time regularization for video decompression*, SIAM  
834 *Journal on Imaging Sciences*, 8 (2015), pp. 373–402.
- 835 [39] M. SCHLOEGL, M. HOLLER, A. SCHWARZL, K. BREDIES, AND R. STOLLBERGER, *Infimal convolution*  
836 *of total generalized variation functionals for dynamic MRI*, Magnetic Resonance in Medicine, (2016).  
837 <http://dx.doi.org/10.1002/mrm.26352>.
- 838 [40] S. SÉRIERE, C. TAUBER, J. VERCOULLIE, D. GUILLOTEAU, J.-B. DELOYE, L. GAR-  
839 *REAU, L. GALINEAU, AND S. CHALON, In vivo PET quantification of the dopamine trans-*  
840 *porter in rat brain with [18f]lbt-999*, Nuclear Medicine and Biology, 41 (2014), pp. 106 –  
841 113, doi:<http://dx.doi.org/10.1016/j.nucmedbio.2013.09.007>. [http://www.sciencedirect.com/science/](http://www.sciencedirect.com/science/article/pii/S0969805113002060)  
842 [article/pii/S0969805113002060](http://www.sciencedirect.com/science/article/pii/S0969805113002060).
- 843 [41] T. STRÖMBERG, *The operation of infimal convolution*, PhD Thesis, (1994).
- 844 [42] S. TONG, A. M. ALESSIO, AND P. E. KINAHAN, *Image reconstruction for PET/CT scanners: past*  
845 *achievements and future challenges*, Imaging in medicine, 2 (2010), pp. 529–545.
- 846 [43] M. UNGER, T. MAUTHNER, T. POCK, AND H. BISCHOF, *Tracking as Segmentation of Spatial-Temporal*  
847 *Volumes by Anisotropic Weighted TV*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 193–  
848 206. "[http://dx.doi.org/10.1007/978-3-642-03641-5\\_15](http://dx.doi.org/10.1007/978-3-642-03641-5_15)".
- 849 [44] L. VESE, *A Study in the BV Space of a Denoising-Deblurring Variational Problem*, Applied Mathematics  
850 and Optimization, 44 (2001), pp. 131–161.