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

Maïtine Bergounioux, E Papoutsellis

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

Maïtine Bergounioux† and E. Papoutsellis †

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

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

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)

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

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

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

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

weighted by a positive parameter α and

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

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

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

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

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 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ acting on the space and time respectively with

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

and the TV smoothness is applied along both the spatial and the temporal directions. An obvious question that rises on this particularly setting is the correlation between the space and time. Video regularization approaches as in [17, 25, 30] combine spatial and temporal domains under the corresponding dynamic isotropic norm $\|\varphi\|_\infty = \operatorname*{ess\,sup}_{(t,x)\in Q} |\varphi(t,x)|_2$. Hence, space and time are interacting with each other and contribute under some weight parameters to the TV regularizer.

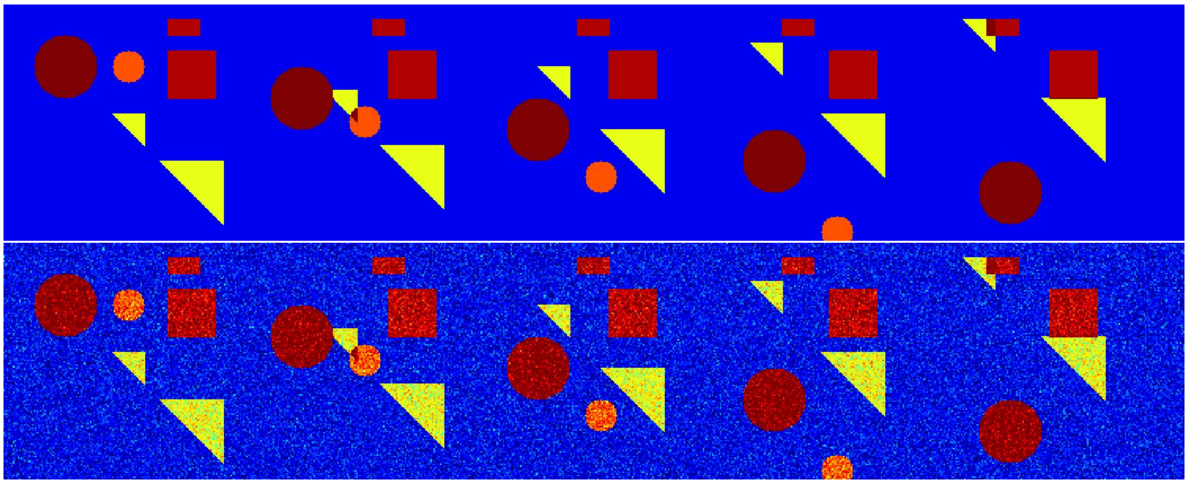
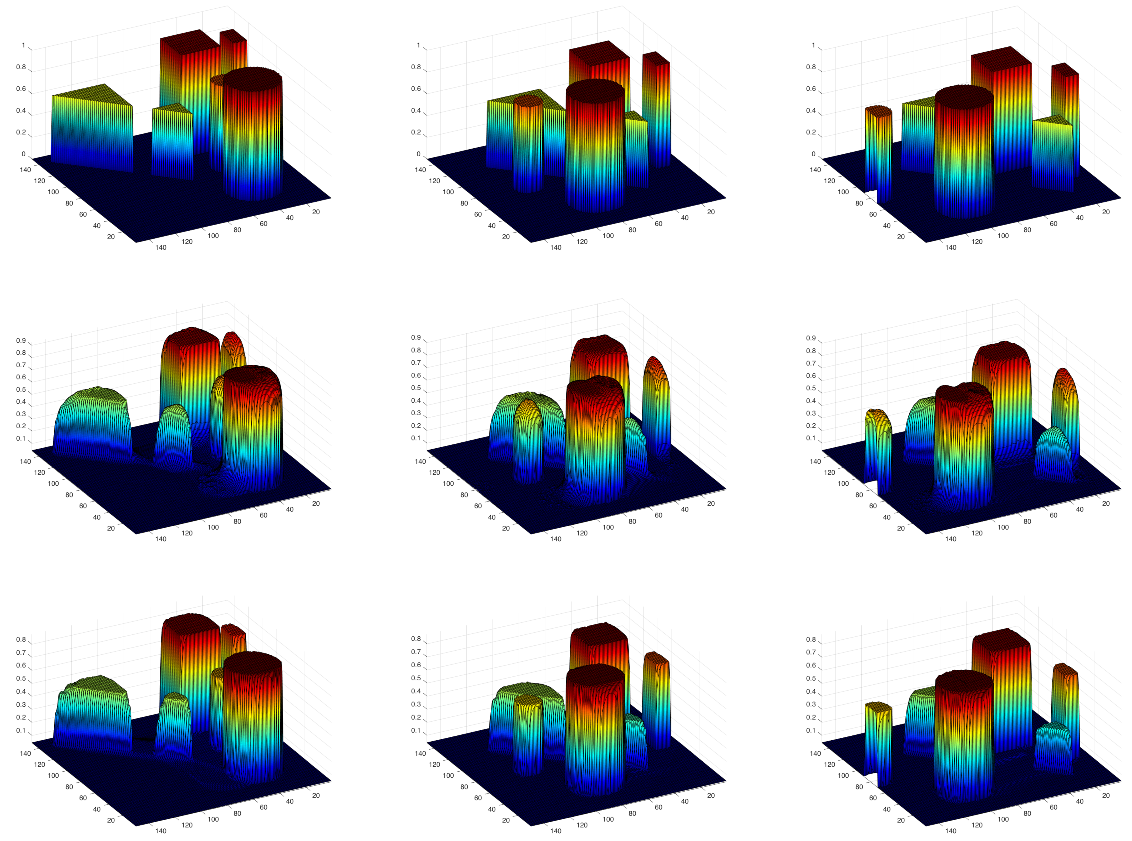


Figure 1.1: Image sequence of 5 frames and its noisy version corrupted with Gaussian noise. Geometrical shapes are moving in different directions with different moving speed.

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

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

This type of decomposition has already been proposed for several applications such as dynamic denoising, segmentation, video decompression and the reader is referred to [43, 38, 26, 17]. Although, this paper is rather theoretical we would like to intrigue the reader with a simple numerical example. In Figure 1.1, we have an image sequence of 5 frames of several geometrical objects moving in different directions and speed under a constant background. This is corrupted by Gaussian noise. In order to compare between isotropic (1.4) and anisotropic (1.5) total variation spatial-temporal regularization we select the parameters of the isotropic/anisotropic TV such that in both cases they will have the same distance for the ground truth (as in PSNR) namely $||\text{solution}_{ISTV} - \text{truth}||_2 = ||\text{solution}_{ANTV} - \text{truth}||_2$.



(a) Frame 1 (b) Frame 3 (c) Frame 4

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 $\alpha_1^{ISTV} = \alpha_1^{ANTV} = 0.5$, $\alpha_2^{ISTV} = 0.05$, $\alpha_2^{ANTV} = 0.0501$.

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

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

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

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

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

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

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

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

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

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

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

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

2. Preliminaries.

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

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

Here, TV_x and TV_t stand for the spatial and temporal total variation for every $t \in \mathcal{T}$ and $x \in \Omega$ respectively. In particular, we have

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

with the corresponding sets

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

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

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

In sequel we drop the index (t,x) in the total variation on Q notation so that TV stands for $\mathrm{TV}_{(t,x)}$. Note that $\mathrm{div}_{(t,x)} = \frac{\partial}{\partial t} + \mathrm{div}_x$. As we pointed out in the introduction, one may consider an equivalent anisotropic norm using for any $\varphi = (\varphi_0, \varphi_1, \cdots, \varphi_d) \in \mathcal{C}_c^1(Q, \mathbb{R}\times\mathbb{R}^d)$: $|\varphi(t,x)|_\infty = \max\left\{\sqrt{\sum_{i=1}^d \varphi_i^2(t,x)}, |\varphi_0(t,x)|\right\} \le 1$ and all the following results are still true.

In the following theorem, see [3, 5], we recall some useful properties on the $\mathrm{BV}(O)$ space, where 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 $O \subset \mathbb{R}^N$, $N \ge 1$. The space $\mathrm{BV}(O)$ endowed with the norm*

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

is a Banach space.
(a) For any $u \in \mathrm{BV}(O)$ there exists a sequence $u_n \in \mathcal{C}^\infty(\bar{O})$ such that

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

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

(c) $\mathrm{BV}(O) \subset \mathrm{L}^p(O)$ *with continuous embedding, for* $1 \le p \le \dfrac{N}{N-1}$ *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* $1 \le p \le \dfrac{N}{N-1}$

$$\forall u \in \mathrm{BV}(O), \qquad \|u - \bar{u}\|_{\mathrm{L}^p(O)} \le C_O \, \mathrm{TV}(u) \ ,$$

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

(d) $\mathrm{BV}(O) \subset \mathrm{L}^p(O)$ *with compact embedding for* $1 \le p < \dfrac{N}{N-1}$.

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

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

$$\mathrm{TV}(u) \le \int_0^T \mathrm{TV}_x(u)(t)dt + \int_\Omega \mathrm{TV}_t(u)(x)dx \le \sqrt{2}\,\mathrm{TV}(u). \tag{2.5}$$

Proof. We start with the first inclusion,

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

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

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

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

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

and

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

Similarly,

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

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

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

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

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

Let us prove the converse inclusion. We first assume that $u \in \mathrm{W}^{1,1}(Q)$. Then, using Fubini's theorem we get $t \mapsto \int_\Omega |\nabla_{t,x}u|(t,x)\ dx \in \mathrm{L}^1(\mathcal{T})$ 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

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

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

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

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

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

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

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

Hence,

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

and

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

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

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

Finally, using (2.7), we get

$$\int_0^T \mathrm{TV}_x(u)(t)\ dt + \int_\Omega \mathrm{TV}_t(u)(x)\ dx \\ \leq \int_0^T \liminf_{k\to\infty} \mathrm{TV}_x(u_k)(t)\ dt + \int_\Omega \liminf_{k\to\infty} \mathrm{TV}_t(u_k)(x)\ dt \leq \sqrt{2}\,\mathrm{TV}(u)\ .$$

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

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

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

3. The variational model.

As already mentioned in the introduction we are interested in the following variational problem

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

where $\mathcal{X} = \mathrm{BV}(Q)$. In this section, we describe the choice of the regularizer term $\mathcal{N}(u)$ as 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 $T > 0$ and $Q = \mathcal{T} \times \Omega \subset \mathbb{R}^{d+1}$.

3.1. Spatial-temporal regularizer.

In this section, we define the spatial-temporal total variation and infimal convolution total variation regularizers weighted by time dependent parameters. Let α be a positive time-dependent weight function $\alpha \in \mathrm{W}^{1,\infty}(\mathcal{T})$. For the spatial and temporal variations, we write $\Phi_{\alpha_1}(u)$ (in space) as the $\mathrm{L}^1(\mathcal{T})$ norm of $t \mapsto \alpha_1(t)\mathrm{TV}_x(u)(t)$, i.e.,

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

and for temporal penalization, Ψ_{α_2} as

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

Note that Φ_{α_1}, Ψ_{α_2} are convex functionals and that the time dependent parameters α_1, α_2 will satisfy

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

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

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

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

that is

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

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

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

Proposition 3.1 (Lower semicontinuity of $F_{\boldsymbol{\alpha}}$). *For every $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ that satisfies (3.4), the functionals Φ_{α_1} and Ψ_{α_2} are lower semicontinuous on $\mathrm{L}^1(\mathcal{T}; \mathrm{BV}(\Omega))$ and $\mathrm{L}^1(\Omega; \mathrm{BV}(\mathcal{T}))$ respectively, with respect to the $\mathrm{L}^1(Q)$ topology. In particular, the functional $F_{\boldsymbol{\alpha}}$ is lower semicontinuous on $\mathrm{BV}(Q)$ with respect to the $\mathrm{L}^1(Q)$ topology. As a consequence, these functionals are lower semicontinuous on $\mathrm{BV}(Q)$ for any $\mathrm{L}^p(Q)$ topology with $p \geq 1$.*

Proof. We start with the lower semicontinuity of Φ_{α_1}. The proof is similar for the lower semicontinuity of Ψ_{α_2}. Let $u_n \in \mathrm{L}^1(\mathcal{T}; \mathrm{BV}(\Omega))$ such that $u_n \to u$ in $\mathrm{L}^1(Q)$.
If $\liminf\limits_{n \to +\infty} \Phi_{\alpha_1}(u_n) = +\infty$ then the lower semicontinuity inequality is obviously satisfied.
Otherwise, one can extract a subsequence (still denoted u_n) such that $\sup_n \Phi_{\alpha_1}(u_n) = \sup_n \int_0^T \mathrm{TV}_x[\alpha_1 u_n](t)\, dt < +\infty$. Fatou's Lemma applied to the sequence $\mathrm{TV}_x(\alpha_1 u_n)$ gives

$$\int_0^T \liminf_{n \to +\infty} \mathrm{TV}_x[\alpha_1 u_n](t)\, dt \leq \liminf_{n \to +\infty} \int_0^T \mathrm{TV}_x[\alpha_1 u_n](t)\, dt = \liminf_{n \to +\infty} \Phi_{\alpha_1}(u_n) < +\infty.$$

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

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

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

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

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

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

Finally,

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

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

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

Theorem 3.1. *Assume that $\boldsymbol\alpha = (\alpha_1, \alpha_2)$ satisfies (3.4). Then, there exists positive constants $C^-_{\boldsymbol\alpha}$, $C^+_{\boldsymbol\alpha}$ depending on $\boldsymbol\alpha$, such that for every $u \in \mathrm{BV}(Q)$*

$$C^-_{\boldsymbol\alpha}\mathrm{TV}(\alpha_2 u) \leq F_{\boldsymbol\alpha}(u) \leq C^+_{\boldsymbol\alpha}\mathrm{TV}(\alpha_2 u)\ . \tag{3.6}$$

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

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

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

$$\begin{aligned} &\frac{\alpha_{min}}{\alpha_{max}}\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u) \leq F_{\boldsymbol\alpha}(u) \leq \frac{\alpha_{max}}{\alpha_{min}}\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u) \Rightarrow \\ &\frac{\alpha_{min}}{\alpha_{max}}\left(\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u)\right) \leq F_{\boldsymbol\alpha}(u) \leq \frac{\alpha_{max}}{\alpha_{min}}\left(\Phi_1(\alpha_2 u) + \Psi_1(\alpha_2 u)\right), \end{aligned}$$

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

$$\frac{\alpha_{min}}{\alpha_{max}}\mathrm{TV}(\alpha_2 u) \leq F_{\boldsymbol\alpha}(u) \leq \sqrt{2}\,\frac{\alpha_{max}}{\alpha_{min}}\mathrm{TV}(\alpha_2 u). \tag{3.7}$$

Here $C^-_{\boldsymbol\alpha} = \frac{\alpha_{min}}{\alpha_{max}}$ and $C^+_{\boldsymbol\alpha} = \sqrt{2}\,\frac{\alpha_{max}}{\alpha_{min}}$. $\square$

In (3.6), we observe that the time dependent parameter α_1 that acts on the spatial domain of $F_{\boldsymbol\alpha}$ does not contribute to the correlated spatial-temporal total variation. In terms of the infimal convolution regularizer, a similar result is true when a certain assumption on the time dependent parameters is imposed.

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

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

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

$$
\begin{aligned}
F_{\boldsymbol{\lambda}}(u-v) + F_{\boldsymbol{\mu}}(v) &\geq C_{\boldsymbol{\lambda}}^{-}\mathrm{TV}(\lambda_2(u-v)) + C_{\boldsymbol{\mu}}^{-}\mathrm{TV}(\mu_2 v) = C_{\boldsymbol{\lambda}}^{-}\mathrm{TV}(\lambda_2(u-v)) + \kappa C_{\boldsymbol{\mu}}^{-}\mathrm{TV}(\lambda_2 v) \\
&\geq \min\left\{C_{\boldsymbol{\lambda}}^{-}, \kappa C_{\boldsymbol{\mu}}^{-}\right\}\Big(\mathrm{TV}(\lambda_2(u-v)) + \mathrm{TV}(\lambda_2 v)\Big) \geq C_1 \mathrm{TV}(\lambda_2 u)
\end{aligned}
$$

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

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

□

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

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

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

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

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

Proof. Recall that relation (2.5) gives

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

Next, we get

$$\begin{aligned}
\alpha_{max} C_{\boldsymbol{\alpha}}^{-} \mathrm{TV}(u) &= \frac{\alpha_{max}\alpha_{min}}{\alpha_{max}} \mathrm{TV}(u) \\
&= \alpha_{min} \mathrm{TV}(u) \\
&\leq \alpha_{min} \left(\int_0^T \mathrm{TV}_x(u)(t)\, dt + \int_\Omega \mathrm{TV}_t(u)(x)\, dx \right) \\
&\leq \alpha_1 \int_0^T \mathrm{TV}_x(u)(t)\, dt + \alpha_2 \int_\Omega \mathrm{TV}_t(u)(x)\, dx = F_{\boldsymbol{\alpha}}(u) \ .
\end{aligned}$$

Similarly,

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

The second inequality is a direct consequence of Proposition 3.2. □

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

<u>**Case**</u> **(1) :** $\mathcal{A} = A$

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

$$\begin{aligned}
&(i)\ A \in \mathcal{L}(\mathrm{L}^p(Q), \mathrm{L}^q(Q)) \text{ with } 1 < p \leq \frac{d+1}{d}, \quad 1 \leq q < \infty, \\
&(ii)\ A\chi_Q \neq 0, \\
&(iii)\ A(\alpha(t)u) = \alpha(t)A(u), \text{ a.e. } t \in \mathcal{T}, \quad \text{for any positive time dependent parameter } \alpha.
\end{aligned} \tag{3.9}$$

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 α 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

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

Next me may define,

$$\mathcal{H}(g, \mathcal{A}u) = \frac{1}{q}\|Au - g\|^q_{\mathrm{L}^q(Q)} \text{ with } g \in \mathrm{L}^q(Q). \tag{3.10}$$

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

<u>Case</u> (2) : $\mathcal{A} = \mathcal{R}$

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

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

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

$$\mathcal{R} : \mathrm{L}^1(Q) \to \mathrm{L}^1(\Sigma), \quad \|\mathcal{R}u\|_{\mathrm{L}^1(\Sigma)} \le C\,\|u\|_{\mathrm{L}^1(Q)}\,. \tag{3.12}$$

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

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

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

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

Lemma 3.1. *The following properties hold true:*
(a) $D_{KL}(w_1, w_2)$ is nonnegative and equal to 0 if and only if $w_1 = w_2$.
(b) The function $(w_1, w_2) \mapsto D_{KL}(w_1, w_2)$ is convex.

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

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

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

$$g \in \mathrm{L}^\infty(\Sigma), \tag{3.15}$$

and set

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

With the above definition we have

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

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

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

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

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

To conclude, we define

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

whose domain is

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

as our data fitting term. Note that $\mathcal{D} \subset \mathrm{L}^1_+(Q)$ 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 $\mathrm{L}^1(Q)$ topology we have

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

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

4. Well-posedness results.

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

$$(\mathcal{P}) \qquad \inf_{u \in \mathrm{BV}(Q)} \mathcal{E}(u),$$

where

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

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

$$\mathcal{H}(g, \mathcal{A}u) + \inf_{v \in \mathrm{BV}(Q)} \int_0^T (\mathrm{TV}_x[\lambda_1(u - v)] + \mathrm{TV}_x[\mu_1 v])\,(t)\,dt + \int_\Omega (\mathrm{TV}_t[\lambda_2(u - v)] + \mathrm{TV}_t[\mu_2 v])\,(x)\,dx.$$

4.1. Lower semicontinuity of the inf-convolution operator.

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

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

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

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

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

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

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

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

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

$$\begin{aligned}\|u\|_{\mathrm{BV}(Q)} &= \|\frac{1}{\alpha}(\alpha u)\|_{\mathrm{BV}(Q)} = \|\frac{1}{\alpha}(\alpha u)\|_{\mathrm{L}^1(Q)} + \mathrm{TV}(\frac{1}{\alpha}(\alpha u)) \\ &\leq \frac{1}{\alpha_{min}}\|\alpha u\|_{\mathrm{L}^1(Q)} + \mathrm{TV}(\frac{1}{\alpha}(\alpha u)) \ .\end{aligned}$$

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

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

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

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

Finally,

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

with $C_\alpha = \dfrac{1}{\alpha_{min}} + \left\|\left(\dfrac{1}{\alpha}\right)'\right\|_{\mathrm{L}^\infty(\mathcal{T})}$. □

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

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

$$v_u \in \operatorname*{argmin}_{v\in\mathrm{BV}(Q)}\ F_{\boldsymbol{\lambda}}(u-v) + F_{\boldsymbol{\mu}}(v)\ \textit{and}\ \int_Q \mu_2(t)\ v_u(t,x)\,dt\ dx = 0.$$

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

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

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

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

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

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

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

As $F_{\boldsymbol{\lambda}}(u - v_n) + F_\mu(v_n)$ is bounded then Theorem 3.1 yields that $\mathrm{TV}(\mu_2 v_n)$ is bounded as well. Moreover, we have $\|\mu_2 v_n\|_{\mathrm{L}^1(Q)} \le C_Q \mathrm{TV}(\mu_2 v_n)$ from the Poincaré-Wirtinger inequality (see Theorem 2.1). Hence, $(\mu_2 v_n)$ is BV-bounded. This implies that v_n is BV-bounded as well (see Lemma 4.1 and (4.3)). Therefore, there exists $v_u \in \mathrm{BV}(Q)$ such that, up to subsequence, $v_n \overset{w^*}{\rightharpoonup} v_u$ in $\mathrm{BV}(Q)$ which implies that $v_n \to v_u$ for the $\mathrm{L}^1(Q)$ topology. We end the proof with the lower semicontinuity of the functional with respect to the the $\mathrm{L}^1(Q)$ topology (with 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 that $\int_Q \mu_2(t) v_u(t,x)\, dx\, dt = 0$ as well. □

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

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

$$F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u) \le \liminf_{n \to +\infty} F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_n). \tag{4.4}$$

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

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

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

$$\forall n \in \mathbb{N} \qquad \mathrm{TV}(\mu_2 v_n) \le \frac{1}{C_\mu^-} F_{\boldsymbol{\mu}}(v_n) \le \frac{C}{C_\mu^-}.$$

Using Poincaré-Wirtinger inequality, we have that

$$\forall n \in \mathbb{N} \qquad \|\mu_2 v_n\|_{\mathrm{L}^1(Q)} \le C_Q \mathrm{TV}(\mu_2 v_n) \le \frac{C\, C_Q}{C_\mu^-}.$$

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

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

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

$$F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u) \le \liminf_{n \to +\infty} F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_n).$$

□

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

Theorem 4.2 (Existence). *Assume that*

- *Case (1): the data $g \in \mathrm{L}^q(Q)$ and A satisfies (3.9), or*
- *Case (2): the data $g \in \mathrm{L}^\infty(\Sigma)$.*

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

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

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

This implies in particular that $u_n \in \mathrm{BV}(Q) \cap \mathcal{D}$ 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 $\mathrm{TV}(w_n)$ is bounded where we have set $w_n = \lambda_2 u_n$. Therefore, with the Poincaré-Wirtinger inequality, then we have $\|w_n - \overline{w_n}\|_{\mathrm{L}^p(Q)} \leq M_1[g]$ with $1 \leq p \leq \frac{d+1}{d}$ and $M_1[g] = \dfrac{C_Q}{C_1} M_0[g]$ Moreover, we have

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

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

$$\underline{\textbf{Case}}\ \textbf{(1)}\ :\ \mathcal{H}(g, \mathcal{A}u) = \tfrac{1}{q}\|Au - g\|^q_{\mathrm{L}^q(Q)}$$

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

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

where

$$M_2[g] = \|A\|\, M_1[g] + \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})} \left((qM_0[g])^{1/q} + \|g\|_{\mathrm{L}^q(Q)} \right)$$

$$= \|A\|\, \frac{C_Q}{C_1}\, M_0[g] + q^{1/q}\, \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, M_0[g]^{1/q} + \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, \|g\|_{\mathrm{L}^q(Q)}\,. \tag{4.6}$$

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

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

$$\|\mathcal{R}w_n - \lambda_2 g\|^2_{\mathrm{L}^1(\Sigma)} \leq \left(\frac{2}{3}\, \|\lambda_2 g\|_{\mathrm{L}^1(\Sigma)} + \frac{4}{3}\, \|\mathcal{R}w_n\|_{\mathrm{L}^1(\Sigma)} \right) D_{KL}(\lambda_2 g, \lambda_2 \mathcal{R}u_n)$$

$$\leq \left(\frac{2}{3}\, \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, \|g\|_{\mathrm{L}^1(\Sigma)} + \frac{4}{3}\, \|\mathcal{R}(w_n - \overline{w_n}) + \mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \right) \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, D_{KL}(g, \mathcal{R}u_n)$$

$$\leq \left(\frac{2}{3}\, \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, \|g\|_{\mathrm{L}^1(\Sigma)} + \frac{4}{3}\, \|\mathcal{R}\|\, \|w_n - \overline{w_n}\|_{L^1(Q)} + \frac{4}{3}\, \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \right) \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, M_0[g]$$

$$\leq \left(\frac{2}{3}\, \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, \|g\|_{\mathrm{L}^1(\Sigma)} + \frac{4}{3}\, \|\mathcal{R}\|\, |Q|^{1/p'} M_1[g] + \frac{4}{3}\, \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \right) \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, M_0[g].$$

Hence,

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

with

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

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

and

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

On the other hand,

$$\|\mathcal{R}w_n - \lambda_2 g\|^2_{\mathrm{L}^1(\Sigma)} \geq \left(\|\mathcal{R}(w_n - \overline{w_n}) - \lambda_2 g\|_{\mathrm{L}^1(\Sigma)} - \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \right)^2$$

$$\geq \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \left(\|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} - 2\, \|\mathcal{R}(w_n - \overline{w_n}) - \lambda_2 g\|_{\mathrm{L}^1(\Sigma)} \right)$$

$$\geq \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \left(\|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} - 2\left(\|\mathcal{R}\|\, |Q|^{1/p'} M_1[g] + \|\lambda_2\|_{\mathrm{L}^\infty(\mathcal{T})}\, \|g\|_{\mathrm{L}^1(\Sigma)} \right) \right)$$

$$= \|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} \left(\|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} - M_5[g] \right), \tag{4.11}$$

with

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

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

Also, we have that $\|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} = \dfrac{\int_Q w_n\,dx\,dt}{|Q|}\,\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}$ that is

$$\|\mathcal{R}\overline{w_n}\|_{\mathrm{L}^1(\Sigma)} = \frac{\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}}{|Q|}\,\|\overline{w_n}\|_{\mathrm{L}^1(Q)}\,. \tag{4.14}$$

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

$$\frac{\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}}{|Q|}\,\|\overline{w_n}\|_{\mathrm{L}^1(Q)} \left(\frac{\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}}{|Q|}\,\|\overline{w_n}\|_{\mathrm{L}^1(Q)} - M_5[g] - \frac{4}{3}M_4[g] \right) \le M_3[g]M_4[g]. \tag{4.15}$$

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

$$\|\overline{w_n}\|_{\mathrm{L}^1(Q)} \le \frac{M_3[g]M_4[g]|Q|}{\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}}$$

Otherwise, we have that

$$\|\overline{w_n}\|_{\mathrm{L}^1(Q)} \le \left(1 + M_5[g] + \frac{4}{3}M_4[g]\right) \frac{|Q|}{\|\mathcal{R}\chi_Q\|_{\mathrm{L}^1(\Sigma)}}.$$

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

$$\|\overline{w_n}\|_{\mathrm{L}^1(Q)} \le M_6[g]\ ,$$

where

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

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

$$F_{\boldsymbol{\lambda}}\#F_{\boldsymbol{\mu}}(u) \le \liminf_{n\to\infty} F_{\boldsymbol{\lambda}}\#F_{\boldsymbol{\mu}}(u_n).$$

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

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

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

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

$$\|u_n\|_{\mathrm{L}^q(Q)} \leq \|u_n - g\|_{\mathrm{L}^q(Q)} + \|g\|_{\mathrm{L}^q(Q)}$$
$$\|u_n\|_{\mathrm{L}^1(Q)} - \|g\|_{\mathrm{L}^\infty(Q)} \log \|u_n\|_{\mathrm{L}^1(Q)} \leq \int_Q u_n - g \log u_n$$

We refer to [27] for the second case.

Theorem 4.3 (Uniqueness). *Assume that the hypothesis of Theorem (4.2) are fulfilled and, in addition that*
- *A is injective and $q \neq 1$ in Case (1),*
- *$\inf_\Sigma g > 0$ in Case (2).*

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

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

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

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

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

and the corresponding perturbed minimization problem

$$\inf_{u \in \mathrm{BV}(Q)} \mathcal{H}(g_n, \mathcal{A}u) + (F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(u). \tag{4.18}$$

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

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

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

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

<u>**Case**</u> **(1)** : $\mathcal{H}(g, \mathcal{A}u) = \frac{1}{q}\|Au - g\|^q_{\mathrm{L}^q(Q)}$

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

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

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

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

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 $\mathrm{BV}(Q)$ and in L^p-bounded with $1 < p \leq \frac{d+1}{d}$. Therefore, we have that $u_n \to \tilde{u}$ in $\mathrm{L}^1(Q)$, $u_n \overset{w}{\rightharpoonup} \tilde{u}$ in $\mathrm{L}^p(Q)$, 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

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

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

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

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

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

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

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

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

$$(F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(u_n) + \mathcal{H}(g_n, \mathcal{R}u_n) \le (F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(u) + \mathcal{H}(g_n, \mathcal{R}u). \tag{4.20}$$

As

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

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

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

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

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

Using (4.20), we get

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

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

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

Similarly to the previous case, we get for every $v \in \mathrm{BV}(Q)$, $v \ge 0$ that

$$\begin{aligned}
(F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(\tilde{u}) + \mathcal{H}(g, \mathcal{R}\tilde{u}) &\le \liminf_{n\to\infty} (F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(u_n) + \mathcal{H}(g_n, \mathcal{R}u_n) \\
&\le \lim_{n\to\infty} (F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(v) + \mathcal{H}(g_n, \mathcal{R}v) \\
&\le (F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(v) + \mathcal{H}(g, \mathcal{R}v).
\end{aligned}$$

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

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

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

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

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

$$F_{\boldsymbol{\lambda}}(\boldsymbol{u} - \boldsymbol{v}) + F_{\boldsymbol{\mu}}(\boldsymbol{v}) = F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(\boldsymbol{u}) = \inf_{v \in \mathrm{BV}(Q)} \{F_{\boldsymbol{\lambda}}(\boldsymbol{u} - v) + F_{\boldsymbol{\mu}}(v)\}. \tag{4.21}$$

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

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

$$F_{\boldsymbol{\lambda}}(\boldsymbol{u} - \boldsymbol{v}) + F_{\boldsymbol{\mu}}(\boldsymbol{v}) + \mathcal{H}(g, \mathcal{A}\boldsymbol{u}) \le F_{\boldsymbol{\lambda}}(u - v) + F_{\boldsymbol{\mu}}(v) + \mathcal{H}(g, \mathcal{A}u) \ . \tag{4.22}$$

Taking $u = \boldsymbol{u}$ gives

$$\forall v \in \mathrm{BV}(Q), \qquad F_{\boldsymbol{\lambda}}(\boldsymbol{u} - \boldsymbol{v}) + F_{\boldsymbol{\mu}}(\boldsymbol{v}) \le F_{\boldsymbol{\lambda}}(\boldsymbol{u} - v) + F_{\boldsymbol{\mu}}(v) \ ,$$

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

$$\forall v \in \mathrm{BV}(Q), \qquad F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(\boldsymbol{u}) + \mathcal{H}(g, \mathcal{A}\boldsymbol{u}) \le F_{\boldsymbol{\lambda}}(u - v) + F_{\boldsymbol{\mu}}(v) + \mathcal{H}(g, \mathcal{A}u),$$

which results to

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

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

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

$$\begin{aligned} F_{\boldsymbol{\lambda}}(\boldsymbol{u} - \boldsymbol{v}) + F_{\boldsymbol{\mu}}(\boldsymbol{v}) + \mathcal{H}(g, \mathcal{A}\boldsymbol{u}) = F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(\boldsymbol{u}) + \mathcal{H}(g, \mathcal{A}\boldsymbol{u}) &\le F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u) + \mathcal{H}(g, \mathcal{A}u) \\ &\le F_{\boldsymbol{\lambda}}(u - v) + F_{\boldsymbol{\mu}}(v) + \mathcal{H}(g, \mathcal{A}u) \ . \end{aligned}$$

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

5. Optimality conditions.

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

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

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

We start by extending Φ_{α_1}, Ψ_{α_2} and $F_{\boldsymbol{\alpha}}$ from their respective domains to $\mathrm{L}^p(Q)$ as follows:

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

$$\tilde F_{\boldsymbol{\alpha}}(u) = \begin{cases} F_{\boldsymbol{\alpha}}(u) & \text{if } u \in \mathrm{BV}(Q), \\ +\infty & \text{if } u \in \mathrm{L}^p(Q)\backslash \mathrm{BV}(Q). \end{cases}$$

We define the *extended* problem as

$$(\mathcal{P}*) \qquad \inf_{u\in \mathrm{L}^p(Q)} \ \mathcal{H}(g,\mathcal{A}u) + (\tilde F_{\boldsymbol{\lambda}} \# \tilde F_{\boldsymbol{\mu}})(u)\ .$$

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

Corollary 5.1. *Let $\boldsymbol{\alpha} = (\alpha_1,\alpha_2)$ that satisfies* (3.4)*. The functionals $\tilde\Phi_{\alpha_1}$, $\tilde\Psi_{\alpha_2}$ and $\tilde F_{\boldsymbol{\alpha}}$ are convex and lower semicontinuous on $\mathrm{L}^p(Q)$.*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

using (3.2), we obtain that

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

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

- 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 $v \in \mathrm{BV}(Q)$, $\Phi_\alpha(v) \ge \langle u^*, v\rangle_{p',p}$ and so $\tilde{\Phi}^*_\alpha(u^*) \le 0$. As $\tilde{\Phi}^*_\alpha$ is an indicator function this means 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 $\mathrm{L}^{p'}(Q)$-closed this gives $\alpha\overline{\mathcal{K}_x} \subset \tilde{\mathcal{K}}$.
- Let us prove the converse inclusion. Assume there exists $u^* \in \tilde{\mathcal{K}}$ such that $u^* \notin \alpha\overline{\mathcal{K}_x}$. One can separate u^* and $\alpha\overline{\mathcal{K}_x}$, see [13]: there exists $\omega \in \mathbb{R}$ and $u_0 \in \mathrm{L}^p(Q)$ such that

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

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

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

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

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

$$\Phi_\alpha(u) \ge \langle u^*, u\rangle_{p',p}\,. \tag{5.4}$$

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

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

and taking the supremum we have that

$$\sup_{\xi \in K_x} \int_\Omega \alpha(t)\xi(x)u(t,x)\ dx \leq \int_\Omega \sup_{\zeta \in K_x} \alpha(t)\zeta(x)u(t,x)\ dx\ ,$$

$$\mathrm{TV}_x(\alpha u)(t) \leq \int_\Omega \sup_{\zeta \in K_x} \alpha(t)\zeta(x)u(t,x)\ dx.$$

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

$$\int_0^T \mathrm{TV}_x(\alpha u)(t)\, dt - \int_0^T \int_\Omega u^*(t,x)u(t,x)\ dx\ dt \leq$$
$$\int_0^T \int_\Omega \left[\sup_{\zeta \in K_x} \alpha(t)\zeta(x) - u^*(t,x)\right] u(t,x)\ dx\ dt.$$

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

$$0 \leq \Phi_\alpha(u) - \langle u^*, u\rangle_{p',p} \leq \int_0^T \int_\Omega \left[\sup_{\zeta \in \alpha K_x} \zeta(x) - u^*(t,x)\right] u(t,x)\ dx\ dt$$
$$\leq \int_0^T \int_\Omega \left[\sup_{\xi \in \alpha \mathcal{K}_x} \xi(t,x) - u^*(t,x)\right] u(t,x)\ dx\ dt$$

$$\leq \int_0^T \int_\Omega \left[\sup_{\xi \in \alpha \overline{\mathcal{K}_x}} \xi(t,x) - u^*(t,x)\right] u(t,x)\ dx\ dt.$$

Hence, this implies

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

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

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

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

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

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

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

that is

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

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

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

which is a contradiction by (5.3). □

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

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

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

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 ϕ, ψ we have

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

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

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

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

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

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

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

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

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

We have computed the convex conjugate of our proposed regularizer and we proceed now with the optimality conditions of ($\mathcal{P}$).

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

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

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

Theorem 5.4. *Let $(V, \|\cdot\|)$ be a normed space and let $f, h : V \to \mathbb{R} \cup \{+\infty\}$ be two lower semicontinuous, convex and proper functions.*
(a) The following inclusion is always true: $\partial f + \partial h \subset \partial(f+h)$.
(b) If f is finite and continuous at a point of 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 $\mathcal{H}(g, \mathcal{A}u) = \frac{1}{q}\|Au - g\|^q_{L^q(Q)}$ with $1 \leq q < +\infty$ and A satisfies assumption (3.9). Clearly, $\operatorname{dom} \tilde{F}_{\boldsymbol{\lambda}} \# \tilde{F}_{\boldsymbol{\mu}} = \mathrm{BV}(Q)$, $\operatorname{dom} \mathcal{H}(g, \mathcal{A}u) = \mathrm{L}^p(Q)$ and $u \to \mathcal{H}(g, \mathcal{A}u)$ is L^p continuous at $0 \in \mathrm{BV}(Q)$. Therefore,

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

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

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

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

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

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

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

Note that in the latter case one has

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

Overall, we have that

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

and one concludes to the following result:

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

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

5.2.2. Optimality conditions for $(\mathcal{P})$: case (2). In this subsection we focus on the 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 $u \in \mathcal{D}$, the domain of the fidelity term. We cannot follow the same strategy as before due to the limitations of this fidelity in terms of continuity. It is known that a proper, convex, lower semicontinuous function is continuous if and only if the interior of its domain is not empty, i.e., $int(dom f) \neq \emptyset$, see [20]. In our case the effective domain is in fact nowhere dense and $D_{KL}(g, \mathcal{R}\cdot)$ is nowhere continuous in $\mathrm{L}^1(\Omega)$, let alone in $\mathrm{L}^p(\Omega)$, see [19, Remark 2.12]. Moreover, $\tilde{F}_{\boldsymbol{\lambda}} \# \tilde{F}_{\boldsymbol{\mu}}$ is not continuous with respect to the L^p norm.

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

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

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

$$
\begin{aligned}
F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_2) = \inf_{v \in \mathrm{BV}(Q)} F_{\boldsymbol{\lambda}}(u_2 - v) + F_{\boldsymbol{\mu}}(v) &\leq F_{\boldsymbol{\lambda}}(u_2 - v_1) + F_{\boldsymbol{\mu}}(v_1) \\
&\leq F_{\boldsymbol{\lambda}}(u_2 - u_1) + F_{\boldsymbol{\lambda}}(u_1 - v_1) + F_{\boldsymbol{\mu}}(v_1) \\
&= F_{\boldsymbol{\lambda}}(u_2 - u_1) + F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_1).
\end{aligned}
$$

Similarly

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

and using Theorem 3.1

$$
|F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_1) - F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}(u_2)| \leq F_{\boldsymbol{\lambda}}(u_1 - u_2) \leq C^+_{\boldsymbol{\lambda}} \mathrm{TV}(\lambda_2(u_1 - u_2)).
$$

Moreover

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

This prove the continuity of $F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}}$ on $\mathrm{BV}(Q)$. □

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

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

Equivalently, there exists $\boldsymbol{u}^* \in \partial(F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(\boldsymbol{u})$ such that $-\boldsymbol{u}^* \in \partial \mathcal{H}(g, \mathcal{R}(\cdot))(\boldsymbol{u})$. As usual, we have $\boldsymbol{u}^* \in \partial(F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})(\boldsymbol{u}) \iff \boldsymbol{u} \in \partial(F_{\boldsymbol{\lambda}} \# F_{\boldsymbol{\mu}})^*(\boldsymbol{u}^*)$. However, in this setting we are in different topology. Though we have computed $\tilde{F}^*_{\boldsymbol{\lambda}}$ for previous case, the computation of $F^*_{\boldsymbol{\lambda}}$ is still challenging. Indeed, we cannot use the arguments used in Theorem 5.2 since the underlying topology is now the BV one and not the $L^p(Q)$ one any longer. In particular, we loose reflexivity as well as an integral representation on the duality product, see [23].

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

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

Therefore,

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

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

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

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

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

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

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

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

Otherwise, by convexity

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

Therefore

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

A short computation gives

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

Finally ,

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

For this case, we conclude with the following optimality conditions

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

$$\forall w^* \in \boldsymbol{K}_{\boldsymbol{\lambda}} \cap \boldsymbol{K}_{\boldsymbol{\mu}}, \quad \langle \boldsymbol{u}, w^* - \boldsymbol{u}^* \rangle \leq 0, \tag{5.7}$$

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

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

6. Conclusion.

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

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

REFERENCES

[1] R. Acar and C. R. Vogel, *Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems*, Inverse Problems, 10 (1994), pp. 1217–1229. http://dx.doi.org/10.1088/0266-5611/10/6/003.

[2] M. Amar and G. Bellettini, *A notion of total variation depending on a metric with discontinuous coefficients*, Annales de l'I.H.P. Analyse non linéaire, 11 (1994), pp. 91–133. http://eudml.org/doc/78325.

[3] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[4] J. Appell, J. Banas, and N. Merentes Díaz, *Bounded Variation and Around*, De Gruyter series in nonlinear analysis and applications, The Clarendon Press Oxford University Press, Berlin, 2013.

[5] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, vol. 6 of MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer Publishing Company, Incorporated, 1st ed., 2011.

[7] M. Benning, C.-B. Schönlieb, T. Valkonen, and V. Vlacic, *Explorations on anisotropic regularisation of dynamic inverse problems by bilevel optimisation*, arXiv preprint, (2016). arXiv:1602.01278.

[8] M. Bergounioux and L. Piffet, *A Second-Order Model for Image Denoising*, Set-Valued and Variational Analysis, 18 (2010), pp. 277–306. http://dx.doi.org/10.1007/s11228-010-0156-6.

[9] M. Bergounioux and E. Trélat, *A variational method using fractional order Hilbert spaces for tomographic reconstruction of blurred and noised binary images*, Journal of Functional Analysis, 259

(2010), pp. 2296 – 2332. http://dx.doi.org/10.1016/j.jfa.2010.05.016.

[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), pp. 37–63.

[11] J. M. Borwein and A. S. Lewis, *Convergence of best entropy estimates*, SIAM Journal on Optimization, 1 (1991), pp. 191–205, doi:10.1137/0801014. http://dx.doi.org/10.1137/0801014.

[12] K. Bredies, K. Kunisch, and T. Pock, *Total Generalized Variation*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526. http://dx.doi.org/10.1137/090769521.

[13] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer-Verlag New York, 2011. http://dx.doi.org/10.1007/978-0-387-70914-7.

[14] E. Casas and F. Tröltzsch, *Second order optimality conditions and their role in PDE control*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 117 (2015), pp. 3–44.

[15] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.

[16] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), pp. 120–145. http://dx.doi.org/10.1007/s10851-010-0251-1.

[17] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, *An augmented lagrangian method for total variation video restoration*, IEEE Transactions on Image Processing, 20 (2011), pp. 3097–3111.

[18] J. A. Clarkson and R. C. Adams, *On definitions of bounded variation for functions of two variables*, Transactions of the American Mathematical Society, 35 (1933), pp. 824–854. http://www.jstor.org/stable/1989593.

[19] B. Dan and R. Elena, *Bregman distances, totally convex functions, and a method for solving operator equations in banach spaces.*, Abstract and Applied Analysis, (2006), pp. Article ID 84919, 39 p.–Article ID 84919, 39 p., http://eudml.org/doc/53750.

[20] I. Ekeland and R. Temam, *Convex analysis and variational problems*, SIAM, 1976.

[21] S. Esedoglu and S. J. Osher, *Decomposition of images by the anisotropic rudin-osher-fatemi model*, Communications on Pure and Applied Mathematics, 57 (2004), pp. 1609–1626. http://dx.doi.org/10.1002/cpa.20045.

[22] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC press, 2015.

[23] N. Fusco and D. Spector, *A remark on an integral characterization of the dual of BV*, Journal of Mathematical Analysis and Applications, (2017).

[24] T. Goldstein and S. Osher, *The split Bregman method for L^1 regularized problems*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 323–343. http://dx.doi.org/10.1137/080725891.

[25] M. Holler and K. Kunisch, *On infimal convolution of TV-type functionals and applications to video and image reconstruction*, SIAM Journal on Imaging Sciences, 7 (2014), pp. 2258–2300. http://dx.doi.org/10.1137/130948793.

[26] M. S. Hosseini and K. N. Plataniotis, *High-accuracy total variation with application to compressed video sensing*, IEEE Transactions on Image Processing, 23 (2014), pp. 3869–3884. 10.1109/TIP.2014.2332755.

[27] T. Le, R. Chartrand, and T. J. Asaki, *A variational approach to reconstructing images corrupted by poisson noise*, Journal of Mathematical Imaging and Vision, 27 (2007), pp. 257–263. http://dx.doi.org/10.1007/s10851-007-0652-y.

[28] A. Markoe, *Analytic Tomography*, Cambridge University Press, 2014. http://dx.doi.org/10.1017/CBO9780511530012.

[29] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, vol. 22, American Mathematical Society, 2001.

[30] Y. L. Montagner, E. Angelini, and J. C. Olivo-Marin, *Video reconstruction using compressed sensing measurements and 3D total variation regularization for bio-imaging applications*, in 2012 19th IEEE International Conference on Image Processing, 2012, pp. 917–920. 10.1109/ICIP.2012.6467010.

[31] D. Oberlin and E. Stein, *Mapping properties of the radon transform*, Indiana Univ. Math. J., 31 (1982), pp. 641–650.

[32] N. Pustelnik, C. Chaux, J.-C. Pesquet, and C. Comtat, *Parallel algorithm and hybrid regular-*

ization for dynamic PET reconstruction, IEEE Medical Imaging Conference, (2010), pp. 2423–2427. https://hal-upec-upem.archives-ouvertes.fr/hal-00733493.

[33] E. Resmerita and R. S. Anderssen, *Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems*, Mathematical Methods in the Applied Sciences, 30 (2007), pp. 1527–1544.

[34] W. Ring, *Structural Properties of Solutions to Total Variation Regularisation Problems*, ESAIM: Mathematical Modelling and Numerical Analysis, 34 (2000), pp. 799–810. http://dx.doi.org/10.1051/m2an:2000104.

[35] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[36] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.

[37] A. Sawatzky, C. Brune, T. Kösters, F. Wübbeling, and M. Burger, *EM-TV methods for inverse problems with poisson noise*, in Level set and PDE based reconstruction methods in imaging, Springer, 2013, pp. 71–142.

[38] H. Schaeffer, Y. Yang, and S. Osher, *Space-time regularization for video decompression*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 373–402.

[39] M. Schloegl, M. Holler, A. Schwarzl, K. Bredies, and R. Stollberger, *Infimal convolution of total generalized variation functionals for dynamic MRI*, Magnetic Resonance in Medicine, (2016). http://dx.doi.org/10.1002/mrm.26352.

[40] S. Sérriere, C. Tauber, J. Vercouillie, D. Guilloteau, J.-B. Deloye, L. Garreau, L. Galineau, and S. Chalon, *In vivo PET quantification of the dopamine transporter in rat brain with [18f]lbt-999*, Nuclear Medicine and Biology, 41 (2014), pp. 106 – 113, doi:http://dx.doi.org/10.1016/j.nucmedbio.2013.09.007. http://www.sciencedirect.com/science/article/pii/S0969805113002060.

[41] T. Strömberg, *The operation of infimal convolution*, PhD Thesis, (1994).

[42] S. Tong, A. M. Alessio, and P. E. Kinahan, *Image reconstruction for PET/CT scanners: past achievements and future challenges*, Imaging in medicine, 2 (2010), pp. 529–545.

[43] M. Unger, T. Mauthner, T. Pock, and H. Bischof, *Tracking as Segmentation of Spatial-Temporal Volumes by Anisotropic Weighted TV*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 193–206. "http://dx.doi.org/10.1007/978-3-642-03641-5_15".

[44] L. Vese, *A Study in the BV Space of a Denoising-Deblurring Variational Problem*, Applied Mathematics and Optimization, 44 (2001), pp. 131–161.