# Analysis of Optimal Control Problems of Semilinear Elliptic Equations by BV-Functions

**Eduardo Casas · Karl Kunisch**

**Abstract** Optimal control problems for semilinear elliptic equations with control costs in the space of bounded variations are analysed. BV-based optimal controls favor piecewise constant, and hence 'simple' controls, with few jumps. Existence of optimal controls, necessary and sufficient optimality conditions of first and second order are analysed. Special attention is paid on the effect of the choice of the vector norm in the definition of the BV-seminorm for the optimal primal and adjoined variables.

**Keywords** optimal control · bounded variation functions · sparsity · first and second order optimality conditions · semilinear elliptic equations

**Mathematics Subject Classification (2000)** 35J61 · 49J20 · 49J52 · 49K20

## 1 Introduction

This paper is dedicated to the study of the optimal control problem

$$\text{(P)} \quad \min_{u \in BV(\omega)} J(u) = \frac{1}{2} \int_{\Omega} |y - y_d|^2 \, dx + \alpha \int_{\omega} |\nabla u| + \frac{\beta}{2} \left( \int_{\omega} u(x) \, dx \right)^2 + \frac{\gamma}{2} \int_{\omega} u^2(x) \, dx,$$

where $y$ is the unique solution to the Dirichlet problem

$$\begin{cases} -\Delta y + f(x, y) = u\chi_{\omega} & \text{in } \Omega, \\ \qquad\qquad\quad y = 0 & \text{on } \Gamma. \end{cases} \tag{1.1}$$

E. Casas
Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39005 Santander, Spain
E-mail: eduardo.casas@unican.es

K. Kunisch
Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria
E-mail: karl.kunisch@uni-graz.at

The control domain $\omega$ is an open subset of $\Omega$ with a Lipschitz boundary. We assume that $\alpha > 0$, $\beta \geq 0$, $\gamma \geq 0$, $y_d \in L^2(\Omega)$, and $\Omega$ is a bounded domain in $\mathbb{R}^n$, $n = 2$ or 3, with Lipschitz boundary $\Gamma$. Additionally we make the following hypothesis:

$$\text{if } n = 3, \quad \text{then} \quad \gamma > 0 \text{ is assumed.} \tag{1.2}$$

Here, $BV(\omega)$ denotes the space of functions of bounded variation in $\omega$ and $\int_\omega |\nabla u|$ stands for the total variation of $u$. The assumptions on the nonlinear term $f(x, y)$ in the state equation will be formulated later. By introducing the penalty term involving the mean of $u$ when $\beta > 0$ we realize the fact that constants functions constitute the kernel of the BV-seminorm. If $\gamma = 0$, in dependence on the order of the nonlinearity $f$ it can be necessary to choose $\beta > 0$ to guarantee that (P) admits a solution.

The use of the BV-seminorm in (P) enhances that the optimal controls are piecewise constant in space. Thus the cost functional in (P) models the objective of simultaneously determining a control of simple structure and resulting in a state $y = y(u)$ which is as close to $y_d$ as possible. Comparing with the common formulation of using $L^2(\omega)$ or $L^p(\omega)$ control-cost functionals, with $p > 2$ to match the nonlinearity $f$, these later functionals will produce smooth optimal controls which may be more intricate to realize in practice than controls which result from the $BV-$formulation. Piecewise constant behavior of the optimal controls can also be obtained by introducing bilateral bounds $\underline{a} \leq u(x) \leq \overline{b}$ together with only the tracking term in (P). In this case we can expect optimal controls which exhibit bang-bang structure. If an $L^1(\omega)$ control cost term is added then the optimal control will be of the form bang-zero-bang. But this type of behavior is distinctly different from that which is allowed in (P), since the value of the piecewise constants plateaus is not prescribed. This is distinctly different from the bilaterally constraint case where the optimal control typically assumes one of the extreme values $\underline{a}$ or $\overline{b}$. This in turn can lead to unnecessarily high control costs.

Possibly one of the first papers where this was pointed out, but not systematically investigated is [15]. In [9] semilinear parabolic equations with temporally dependent BV-functions as controls were investigated. Thus we were focusing on controls which are optimally switching in time. The analysis for this case is simpler and exploits specific properties of BV-functions in dimension one. Numerically the simple structure of the controls which is obtained for BV-constrained control problems was already demonstrated in [5,9] and a recent master thesis [19]. BV-seminorm control costs are also employed in [8], where the control appears as coefficient in the $p$-Laplace equation. Beyond these papers the choice of the control costs related to BV-norms or BV-seminorms has not received much attention in the optimal control literature yet.

In mathematical image analysis, to the contrary, the BV-seminorm has received a tremendous amount of attention. The beginning of this activity is frequently dated to [22]. Let us also mention the recent paper [2] which gives interesting insight into the structure of the subdifferential of the BV-seminorm. Fine properties of BV-functions, in the context of image reconstruction problems, in particular the stair casing effect were, analyzed for the one-dimensional case in [21], and in higher dimensions in [20,14], for example. In [11] the authors provided a convergence analysis for BV-regularized mathematical imaging problems by finite elements, paying special attention to the choice of the vector norm in the definition of the BV-seminorm.

Let us also compare the use of the BV-term in (P) with the efforts that have been made for studying optimal control problems with sparsity constraints. These formulations involve either measure-valued norms of the control or $L^1$-functionals combined with pointwise constraints on the control. We cite [5,7] from among the many results which are now already available. The BV-seminorm therefore can also be understood as a sparsity constraint for the first derivative.

Let us briefly describe the structure of the paper. Section 2 contains an analysis of the state equation and the smooth part of the cost-functional. The non-smooth part of the cost-functional is investigated in Section 3. Special attention is given to the consequences which arise from the specific choice which is made for the vector norm in the variational definition of the BV-seminorm. In particular, we consider the Euclidean and the infinity norms, which at times are also referred to as the isotropic and anisotropic cases. Existence of optimal solutions and first order optimality conditions are obtained in Section 4. Second order sufficient optimality conditions are provided in Section 5. Finally in Section 6 we consider (P) with an additional $H^1(\omega)$ regularisation term and investigate the asymptotic behavior as the weight of the $H^1(\omega)$ regularisation tends to 0.

## 2 Analysis of the state equation and the cost functional

We recall that a function $u \in L^1(\omega)$ is a function of bounded variation if its distributional derivatives $\partial_{x_i} u$, $1 \leq i \leq n$, belong to the Banach space of real and regular Borel measures $\mathcal{M}(\omega)$. Given a measure $\mu \in \mathcal{M}(\omega)$, its norm is given by

$$\|\mu\|_{\mathcal{M}(\omega)} = \sup\{\int_\omega z \, d\mu : z \in C_0(\omega) \text{ and } \|z\|_{C_0(\omega)} \leq 1\} = |\mu|(\omega),$$

where $C_0(\omega)$ denotes the Banach space of continuous functions $z : \bar{\omega} \longrightarrow \mathbb{R}$ such that $z = 0$ on $\partial\omega$, and $|\mu|$ is the total variation measure associated with $\mu$. On the product space $\mathcal{M}(\omega)^n$ we define the norm

$$\|\mu\|_{\mathcal{M}(\omega)^n} = \sup\{\int_\omega z \, d\mu : z \in C_0(\omega)^n \text{ and } |z(x)| \leq 1 \ \forall x \in \omega\}, \qquad (2.1)$$

where $|\cdot|$ is a norm in $\mathbb{R}^n$.

On $BV(\omega)$ we consider the usual norm

$$\|u\|_{BV(\omega)} = \|u\|_{L^1(\omega)} + \|\nabla u\|_{\mathcal{M}(\omega)^n},$$

that makes $BV(\omega)$ a Banach space; see [1, Chapter 3] or [18, Chapter 1] for details. We recall that the total variation of $u$ is given by

$$\|\nabla u\|_{\mathcal{M}(\omega)^n} = \sup\{\int_\omega \text{div} z \, u \, dx : z \in C_0^\infty(\omega)^n \text{ and } |z(x)| \leq 1 \ \forall x \in \omega\}.$$

We also use the notation

$$\int_\omega |\nabla u| = \|\nabla u\|_{\mathcal{M}(\omega)^n},$$

as already employed in (P). For these topologies $\nabla : BV(\omega) \longrightarrow \mathcal{M}(\omega)^n$ is a linear continuous mapping.

In the sequel we will denote

$$a_u = \frac{1}{|\omega|} \int_\omega u(x)\, dx \ \text{ and } \ \hat{u} = u - a_u \ \text{ for every } u \in BV(\omega).$$

Using [1, Theorem 3.44], it is easy to deduce the existence of a constant $C_\omega$ such that

$$\|u\| := |a_u| + \|\nabla u\|_{\mathcal{M}(\omega)^n} \leq \max\left(1, \frac{1}{|\omega|}\right)\|u\|_{BV(\omega)} \leq C_\omega \|u\|. \tag{2.2}$$

In addition, we mention that $BV(\omega)$ is the dual space of a separable Banach space; see [1, Remark 3.12] for a description of this space. Therefore every bounded sequence $\{u_k\}_{k=1}^\infty$ in $BV(\omega)$ has a subsequence converging weakly* to some $u \in BV(\omega)$. The weak* convergence $u_k \overset{*}{\rightharpoonup} u$ implies that $u_k \to u$ strongly in $L^1(\omega)$ and $\nabla u_k \overset{*}{\rightharpoonup} \nabla u$ in $\mathcal{M}(\omega)^n$; see [1, pages 124-125]. Recall that $\omega$ has a Lipschitz boundary and, hence, it is an extension domain [1, pp. 130–131]. We will also use that $BV(\omega)$ is continuously embedded in $L^p(\omega)$ with $1 \leq p \leq \frac{n}{n-1}$, and compactly embedded in $L^p(\omega)$ for every $p < \frac{n}{n-1}$; see [1, Corollary 3.49]. From this property we deduce that the convergence $u_k \overset{*}{\rightharpoonup} u$ in $BV(\omega)$ implies that $u_k \to u$ strongly in every $L^p(\omega)$ for all $p < \frac{n}{n-1}$.

We make the following assumption on the nonlinear term of the state equation. We assume that $f : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Borel function, of class $C^2$ with respect to the last variable, and satisfies

$$f(\cdot, 0) \in L^{\hat{p}}(\Omega) \ \text{ with } \hat{p} > \frac{n}{2}, \tag{2.3}$$

$$\frac{\partial f}{\partial y}(x, y) \geq 0 \quad \forall y \in \mathbb{R}, \tag{2.4}$$

$$\forall M > 0 \ \exists C_M : \left|\frac{\partial f}{\partial y}(x, y)\right| + \left|\frac{\partial^2 f}{\partial y^2}(x, y)\right| \leq C_M \quad \forall |y| \leq M, \tag{2.5}$$

$$\begin{cases} \forall M > 0 \text{ and } \forall \rho > 0 \ \exists \varepsilon > 0 \text{ such that} \\ \left|\frac{\partial^2 f}{\partial y^2}(x, y_2) - \frac{\partial^2 f}{\partial y^2}(x, y_1)\right| \leq \rho \text{ if } |y_2 - y_1| < \varepsilon \text{ and } |y_1|, |y_2| \leq M, \end{cases} \tag{2.6}$$

for almost all $x \in \Omega$.

Let us observe that if $f$ is an affine function, $f(x, y) = c_0(x)y + d_0(x)$, then (2.3)-(2.6) hold if $c_0 \geq 0$ in $\Omega$, $c_0 \in L^\infty(\Omega)$, and $d_0 \in L^{\hat{p}}(\Omega)$.

By using these assumptions, the following theorem can be proved in a standard way; see, for instance, [26, §4.2.4]. For the Hölder continuity result, the reader is referred to [17, Theorem 8.29].

**Proposition 1** *For every $u \in L^{\hat{p}}(\omega)$ the state equation* (1.1) *has a unique solution $y_u \in C^\sigma(\bar{\Omega}) \cap H_0^1(\Omega)$ for some $\sigma \in (0, 1)$. In addition, for every $M > 0$ there exists a constant $K_M$ such that*

$$\|y_u\|_{C^\sigma(\bar{\Omega})} + \|y_u\|_{H_0^1(\Omega)} \leq K_M \ \ \forall u \in L^{\hat{p}}(\omega) : \|u\|_{L^{\hat{p}}(\omega)} \leq M. \tag{2.7}$$

In the sequel we will denote $Y = C(\bar{\Omega}) \cap H_0^1(\Omega)$ and $S : L^{\hat{p}}(\omega) \longrightarrow Y$ the mapping associating to each control $u$ the corresponding state $S(u) = y_u$. We have the following differentiability property of $S$.

**Proposition 2** *The mapping $S : L^{\hat{p}}(\omega) \longrightarrow Y$ is of class $C^2$. For all elements $u, v$ and $w$ of $L^{\hat{p}}(\omega)$, the functions $z_v = S'(u)v$ and $z_{vw} = S''(u)(v, w)$ are the solutions of the problems*

$$\begin{cases} -\Delta z + \dfrac{\partial f}{\partial y}(x, y_u)z = v\chi_\omega & in \ \Omega, \\ \qquad\qquad\qquad z = 0 & on \ \Gamma, \end{cases} \tag{2.8}$$

*and*

$$\begin{cases} -\Delta z + \dfrac{\partial f}{\partial y}(x, y_u)z + \dfrac{\partial^2 f}{\partial y^2}(x, y_u)z_v z_w = 0 & in \ \Omega, \\ \qquad\qquad\qquad\qquad\qquad\qquad z = 0 & on \ \Gamma, \end{cases} \tag{2.9}$$

*respectively.*

The proof is a consequence of the implicit function theorem. Let us give a sketch. We define the space

$$V = \{y \in Y : \Delta y \in L^{\hat{p}}(\Omega)\}$$

endowed with the norm

$$\|y\|_V = \|y\|_{C(\bar{\Omega}))} + \|y\|_{H_0^1(\Omega)} + \|\Delta y\|_{L^{\hat{p}}(\Omega)}.$$

Thus, $V$ is a Banach space. Now we introduce the mapping $\mathcal{F} : V \times L^{\hat{p}}(\Omega) \longrightarrow L^{\hat{p}}(\Omega)$ by

$$\mathcal{F}(y, u) = -\Delta y + f(x, y) - u.$$

From (2.5) we deduce that $\mathcal{F}$ is of class $C^2$ and

$$\frac{\partial \mathcal{F}}{\partial y}(y, u)z = -\Delta z + \frac{\partial f}{\partial y}(x, y)z.$$

From the monotonicity condition (2.4), we obtain that $\frac{\partial \mathcal{F}}{\partial y}(y, u) : V \longrightarrow L^{\hat{p}}(\Omega)$ is an isomorphism. Hence, the implicit function theorem and Proposition 1 with $\omega = \Omega$ imply the existence of a $C^2$ mapping $\hat{S} : L^{\hat{p}}(\Omega) \longrightarrow Y$ associating to every element $u$ its corresponding state $\hat{S}(u) = y_u$. When $\omega \subsetneq \Omega$, we use that $S = \hat{S} \circ S_\omega$, where $S_\omega : L^{\hat{p}}(\omega) \longrightarrow L^{\hat{p}}(\Omega)$ is defined by $S_\omega u = u\chi_\omega$. Hence the chain rule leads to the result.

Next, we separate the smooth and the non smooth parts in $J$: $J(u) = F(u) + \alpha G(u)$ with

$$F(u) = \frac{1}{2}\int_\Omega |y_u - y_d|^2 \, dx + \frac{\beta}{2}\Big(\int_\omega u(x) \, dx\Big)^2 + \frac{\gamma}{2}\int_\omega u^2(x) \, dx \ \text{ and } \ G(u) = g(\nabla u),$$

where $g : \mathcal{M}(\omega)^n \longrightarrow \mathbb{R}$ is given by $g(\mu) = \|\mu\|_{\mathcal{M}(\omega)^n}$. In the rest of this section we study the differentiability of $F$. From Proposition 2 and the chain rule the following proposition can be obtained.

**Proposition 3** *The functional $F : L^2(\omega) \longrightarrow \mathbb{R}$ is of class $C^2$. The derivatives of $F$ are given by*

$$F'(u)v = \int_\omega \Big[\varphi_u(x) + \gamma u(x) + \beta\Big(\int_\omega u(s) \, ds\Big)\Big]v(x) \, dx, \tag{2.10}$$

*and*

$$F''(u)(v,w) = \int_\Omega \big(1 - \varphi_u \frac{\partial^2 f}{\partial y^2}(x, y_u)\big) z_v z_w \, dx + \gamma \int_\omega vw \, dx + \beta \Big( \int_\omega v \, dx \Big) \Big( \int_\omega w \, dx \Big)$$
$$(2.11)$$

*with $z_v = S'(u)v$, $z_w = S'(u)w$, and $\varphi_u \in Y$ the adjoint state which satisfies*

$$\begin{cases} -\Delta \varphi_u + \dfrac{\partial f}{\partial y}(x, y_u)\varphi_u = y_u - y_d & in \ \Omega, \\[2mm] \qquad\qquad\quad \varphi_u = 0 & on \ \Gamma. \end{cases} \qquad (2.12)$$

The $C(\bar{\Omega})$ regularity of $\varphi_u$ follows from the assumptions on $y_d \in L^2(\Omega)$ and the fact that $y_u \in L^\infty(\Omega)$.

*Remark 1* If $n = 2$, since $BV(\omega)$ is embedded in $L^2(\omega)$, then the functional $F : BV(\omega) \longrightarrow \mathbb{R}$ is well defined and it is of class $C^2$ with derivatives given by (2.10) and (2.11). However, if $n = 3$, then $BV(\omega)$ is only embedded in $L^{3/2}(\omega)$. Hence, for elements $u \in BV(\omega)$ Proposition 1 is not applicable and, therefore, the functional $F$ is not defined in $BV(\omega)$. To deal with the case $n = 3$ we introduced the assumption (1.2), i.e. $\gamma > 0$. Hence, the functional $F : BV(\omega) \cap L^2(\omega) \longrightarrow \mathbb{R}$ is well defined and of class $C^2$.

The assumption (1.2) can be avoided if we suppose that the nonlinearity $f(x, y)$ has only polynomial growth of arbitrary order in $y$. In this case, Propositions 1 and 2 hold if we change $Y$ to $Y_q = L^q(\Omega) \cap H_0^1(\Omega)$ with $q < \infty$ arbitrarily big. We recall that for a right hand side of the state equation belonging to $L^{3/2}(\Omega)$ the solution of the state equation does not belong to $L^\infty(\Omega)$, in general, even for linear equations. However, since $L^{3/2}(\Omega) \subset W^{-1,3}(\Omega)$, we can use [25, Theorem 4.2] to deduce that $y_u \in L^q(\Omega) \ \forall q < \infty$. To analyze the semilinear case one can follow the classical approach of truncation of the nonlinear term, Schauder's fix point theorem, and $L^q$-estimates from the linear case combined with the monotonicity of the nonlinear term. Finally, since $\gamma = 0$, we have that the functional $F : BV(\omega) \longrightarrow \mathbb{R}$ is of class $C^2$.

*Remark 2* In the state equation, the Laplace operator $-\Delta$ can be replaced by a more general linear elliptic operator with bounded coefficients. All the results proved in this paper hold for these general operators.

## 3 Analysis of the functional $G$

Now, we analyze the functional $G$. We already expressed $G$ as the composition $G = g \circ \nabla$. Concerning the functional $g$, we note that it is Lipschitz continuous and convex. Hence, it has a subdifferential and a directional derivative, which are denoted by $\partial g(\mu)$ and $g'(\mu; \nu)$, respectively. Before giving an expression for $\partial g(\mu)$ and $g'(\mu; \nu)$, we have to specify the norm that we use in $\mathbb{R}^n$. Indeed, in the definition of the norm $\|\mu\|_{\mathcal{M}(\omega)^n}$ we have considered a generic norm $|\cdot|$ in $\mathbb{R}^n$. The choice of the specific norm strongly influences the structure of the optimal controls. In this paper, we focus on the Euclidean and the $|\cdot|_\infty$ norms, which lead to different properties for $g$, that we consider separately in the following two

subsections. To illustrate one aspect, let us observe that the use of the $|\cdot|_\infty$ norm on $\mathbb{R}^n$ in the definition of $\|\cdot\|_{\mathcal{M}(\omega)^n}$ implies that

$$\|\mu\|_{\mathcal{M}(\omega)^n} = \sum_{j=1}^n \|\mu_j\|_{\mathcal{M}(\omega)} \quad \forall \mu \in \mathcal{M}(\omega)^n. \tag{3.1}$$

In particular, it holds that

$$\int_\omega |\nabla u| = \sum_{j=1}^n \|\partial_{x_j} u\|_{\mathcal{M}(\omega)} \quad \forall u \in BV(\omega).$$

However, if in (2.1) we take the Euclidean norm for the constraint $|z(x)| \leq 1$, then, in general,

$$\|\mu\|_{\mathcal{M}(\omega)^n} \neq \Big( \sum_{j=1}^n \|\mu_j\|_{\mathcal{M}(\omega)}^2 \Big)^{1/2}. \tag{3.2}$$

Indeed, the identity (3.1) is an immediate consequence of the definitions of the norms $\|\cdot\|_{\mathcal{M}(\omega)}$ and $\|\cdot\|_{\mathcal{M}(\omega)^n}$. To verify (3.2) we give an example. Let us fix $n$ different points $\{\xi^i\}_{i=1}^n$ in $\omega$ and take $\varepsilon > 0$ small enough such that the balls $B_\varepsilon(\xi^i)$ are disjoint. Now, applying Urysohn's lemma, cf. [23, Lemma 2.12], we get functions $z_i \in C_0(\omega)$ such that $0 \leq z_i(x) \leq 1 \ \forall x \in \omega$, $z_i(\xi^i) = 1$ and $\mathrm{supp}(z_i) \subset B_\varepsilon(\xi^i)$. We set $z = (z_1, \ldots, z_n)$ and $\mu = (\delta_{\xi^1}, \ldots, \delta_{\xi^n})$. Then, since $|z(x)|_2 \leq 1$ $\forall x \in \omega$, we have

$$\|\mu\|_{\mathcal{M}(\omega)^n} = \sum_{i=1}^n \int_\omega z_i(x) \, d\mu_i(x) = \sum_{i=1}^n z_i(\xi^i) = n.$$

On the other hand, we get

$$\Big( \sum_{j=1}^n \|\mu_j\|_{\mathcal{M}(\omega)}^2 \Big)^{1/2} = \sqrt{n}.$$

3.1 The use of the Euclidean norm $|\cdot|_2$

In order to give an expression for $\partial g(\mu)$ and $g'(\mu; \nu)$, let us introduce some notation. We recall that if $\mu \in \mathcal{M}(\omega)^n$, its associated total variation measure is defined as a positive scalar measure as follows

$$|\mu|(A) = \sup \Big\{ \sum_{k=1}^\infty |\mu(E_k)|_2 : \{E_k\}_k \subset \mathcal{B} \text{ are pairwise disjoint and } A = \bigcup_{k=1}^\infty E_k \Big\},$$

where $\mathcal{B}$ is the $\sigma$-algebra of Borel sets in $\omega$, and $|\mu(E_k)|_2$ denotes the Euclidean norm in $\mathbb{R}^n$ of the vector $\mu(E_k)$. Let us denote by $h_\mu$ the Radon-Nikodym derivative of $\mu$ with respect to $|\mu|$. Thus we have

$$h_\mu \in L^1(\omega, |\mu|)^n, \ |h_\mu(x)|_2 = 1 \text{ for } |\mu|-\text{a.e. } x \in \omega \text{ and } \mu(A) = \int_A h_\mu(x) \, d|\mu|(x) \ \forall A \in \mathcal{B}.$$

Given a second vector measure $\nu \in \mathcal{M}(\omega)^n$, the following Lebesgue decomposition holds: $\nu = \nu_a + \nu_s$, $d\nu_a = h_\nu d|\mu|$, where $\nu_a$ and $\nu_s$ are the absolutely continuous

and singular parts of $\nu$ with respect to $|\mu|$, and $h_\nu$ is the Radon-Nikodym derivative of $\nu$ with respect to $|\mu|$. Then, the following identity is fulfilled

$$\|\nu\|_{\mathcal{M}(\omega)^n} = \|\nu_a\|_{\mathcal{M}(\omega)^n} + \|\nu_s\|_{\mathcal{M}(\omega)^n} = \int_\omega |h_\nu(x)|_2\, d|\mu|(x) + \|\nu_s\|_{\mathcal{M}(\omega)^n}.$$

The reader is referred to [1, Chapter 1].

Now, we analyze the subdifferential $\partial g(\mu)$. It is well known that an element $\lambda \in \partial g(\mu)$ if

$$\langle \lambda, \nu - \mu \rangle + \|\mu\|_{\mathcal{M}(\omega)^n} \le \|\nu\|_{\mathcal{M}(\omega)^n} \quad \forall \nu \in \mathcal{M}(\omega)^n. \tag{3.3}$$

This is equivalent to the next two relations

$$\langle \lambda, \mu \rangle = \|\mu\|_{\mathcal{M}(\omega)^n}, \tag{3.4}$$
$$\langle \lambda, \nu \rangle \le \|\nu\|_{\mathcal{M}(\omega)^n} \quad \forall \nu \in \mathcal{M}(\omega)^n. \tag{3.5}$$

Observe that $\lambda$ belongs to the dual of $\mathcal{M}(\omega)^n$, which is not a space of distributions. However, due to (3.5), the restriction of the functional $\lambda$ to $L^1(\omega)^n \subset \mathcal{M}(\omega)^n$ can be identified with a function of $L^\infty(\omega)^n$ with $\|\lambda\|_{L^\infty(\omega)^n} \le 1$. However, we observe that much information on $\lambda$ can be lost when it is restricted to $L^1(\omega)^n$. Indeed, let us consider the following example. Let $\{A_j\}_{j=1}^n$ a family of pairwise disjoint compact subsets of $\omega$ with zero Lebesgue measure. Set $\mu = (\mu_1, \ldots, \mu_n) \in \mathcal{M}(\omega)^n$ with $\mathrm{supp}(\mu_j) \subset A_j$ and $\mu_j$ being a positive measure for all $j = 1, \ldots, n$. Now, we define $\lambda \in [\mathcal{M}(\omega)^n]^*$ by

$$\langle \lambda, \nu \rangle = \sum_{j=1}^n \nu_j(A_j) \quad \forall \nu \in \mathcal{M}(\omega)^n.$$

It is easy to check that (3.4) and (3.5) hold, hence $\lambda \in \partial g(\mu)$. If we consider the restriction of $\lambda$ to $L^1(\omega)^n$, due to the fact that every $A_j$ has a zero Lebesgue measure, we have

$$\langle \lambda, v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} = \sum_{j=1}^n \int_{A_j} v_j(x)\, dx = 0 \quad \forall v \in L^1(\omega)^n.$$

Therefore, the restriction of $\lambda$ is zero.

In the special case where $\lambda$ is weakly$^*$ continuous, then $\lambda$ is identified with a function of $C_0(\omega)^n$. Thus, the action of the functional $\lambda$ is given by the integral

$$\langle \lambda, \nu \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} = \sum_{j=1}^n \int_\omega \lambda_j(x)\, d\nu_j(x) \quad \forall \nu \in \mathcal{M}(\omega)^n,$$

and we can establish some precise relations between $\lambda$ and $\mu$. Before proving these relations, let us mention that here we have

$$\|z\|_{C_0(\omega)^n} = \sup\{|z(x)|_2 : x \in \omega\} \quad \forall z \in C_0(\omega)^n.$$

**Proposition 4** *If $\lambda \in C_0(\omega)^n \cap \partial g(\mu)$, then $\|\lambda\|_{C_0(\omega)^n} \le 1$. Moreover, if $\mu \ne 0$, then the following properties hold*

1. *$\|\lambda\|_{C_0(\omega)^n} = 1$, and*
2. *$\mathrm{supp}(\mu) \subset \{x \in \omega : |\lambda(x)|_2 = 1\}$.*

*Proof* The inequality $\|\lambda\|_{C_0(\omega)^n} \leq 1$ follows from (3.5). Additionally, if $\mu \neq 0$, then (3.4) implies 1. To prove 2. we use (3.4) as follows

$$\int_\omega d|\mu|(x) = \|\mu\|_{\mathcal{M}(\omega)^n} = \langle \lambda, \mu \rangle = \int_\omega \lambda(x)\, d\mu(x) = \int_\omega \lambda(x) \cdot h_\mu(x)\, d|\mu|(x).$$

Then, using that $|\lambda(x)|_2 \leq 1 \ \forall x \in \omega$ and $|h_\mu(x)|_2 = 1 \ |\mu|$-a.e. in $\omega$ we deduce from the identity

$$\int_\omega d|\mu|(x) = \int_\omega \lambda(x) \cdot h_\mu(x)\, d|\mu|(x)$$

that $\lambda(x) \cdot h_\mu(x) = 1 \ |\mu|$-a.e. in $\omega$. Using again that $|h_\mu(x)|_2 = 1$, $|\mu|$-a.e., we conclude that $\lambda(x) = h_\mu(x)$, $|\mu|$-a.e. Therefore, we have that

$$|\mu|\big(\{x \in \omega : |\lambda(x)|_2 < 1\}\big) = 0,$$

which implies 2. $\qquad\blacksquare$

Next we study the directional derivatives of $g$.

**Proposition 5** *Let $\mu, \nu \in \mathcal{M}(\omega)^n$, then*

$$g'(\mu; \nu) = \int_\omega h_\nu\, d\mu + \|\nu_s\|_{\mathcal{M}(\omega)^n}, \tag{3.6}$$

*where $\nu = \nu_a + \nu_s = h_\nu d|\mu| + \nu_s$ is the Lebesgue decomposition of $\nu$ respect to $|\mu|$.*

*Proof* As above, let us write $d\mu = h_\mu d|\mu|$. Then we have

$$g'(\mu; \nu) = \lim_{\rho \searrow 0} \frac{\|\mu + \rho\nu\|_{\mathcal{M}(\omega)^n} - \|\mu\|_{\mathcal{M}(\omega)^n}}{\rho}$$

$$= \lim_{\rho \searrow 0} \frac{\|\mu + \rho\nu_a\|_{\mathcal{M}(\omega)^n} + \|\rho\nu_s\|_{\mathcal{M}(\omega)^n} - \|\mu\|_{\mathcal{M}(\omega)^n}}{\rho}$$

$$= \lim_{\rho \searrow 0} \frac{1}{\rho}\left(\int_\omega |h_\mu(x) + \rho h_\nu(x)|_2\, d|\mu|(x) - \int_\omega |h_\mu(x)|_2\, d|\mu|(x)\right) + \|\nu_s\|_{\mathcal{M}(\omega)^n}$$

$$= \int_\omega \lim_{\rho \searrow 0} \frac{|h_\mu(x) + \rho h_\nu(x)|_2 - |h_\mu(x)|_2}{\rho}\, d|\mu|(x) + \|\nu_s\|_{\mathcal{M}(\omega)^n}$$

$$= \int_\omega \frac{h_\mu(x) \cdot h_\nu(x)}{|h_\mu(x)|_2}\, d|\mu|(x) + \|\nu_s\|_{\mathcal{M}(\omega)^n} = \int_\omega h_\nu\, d\mu + \|\nu_s\|_{\mathcal{M}(\omega)^n}.$$

Since the quotients are dominated by $|h_\nu|_2$, we applied Lebesgue's dominated convergence theorem above. Moreover, we use that $|h_\mu(x)|_2 = 1 \ |\mu|$-a.e. in $\omega$ in the last equality and also to justify the differentiability of the norm $|\cdot|_2$ at every $h_\mu(x)$ with $x$ in the support of $|\mu|$. $\qquad\blacksquare$

Now, we come back to the mapping $G$. To this end, let us recall that the adjoint operator $\nabla^*$ is defined by

$$\nabla^* : [\mathcal{M}(\omega)^n]^* \longrightarrow BV(\omega)^*, \quad \langle \nabla^* \lambda, u \rangle_{BV(\omega)^*, BV(\omega)} = \langle \lambda, \nabla u \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n}.$$

**Proposition 6** *The following identities hold for all $u \in BV(\omega)$:*

$$\partial G(u) = \partial(g \circ \nabla)(u) = \nabla^* \partial g(\nabla u), \tag{3.7}$$

$$G'(u; v) = (g \circ \nabla)'(u; v) = \int_\omega h_v \, d(\nabla u) + \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n}, \tag{3.8}$$

*where $\nabla v = h_v d|\nabla u| + (\nabla v)_s$ is the Lebesgue decomposition of $\nabla v$ with respect to $|\nabla u|$.*

*Proof* Since $\nabla : BV(\omega) \longrightarrow \mathcal{M}(\omega)^n$ is a linear and continuous mapping and $g : \mathcal{M}(\omega)^n \longrightarrow \mathbb{R}$ is convex and continuous, we can apply the chain rule [16, Chapter I, Proposition 5.7] to deduce that $\partial(g \circ \nabla)(u) = \nabla^* \partial g(\nabla u)$, which immediately leads to (3.7).

To verify (3.8) it is enough to observe that

$$(g \circ \nabla)'(u; v) = g'(\nabla u; \nabla v)$$

and to apply (3.6). $\qquad \blacksquare$

3.2 The use of the $|\cdot|_\infty$ norm

The use of $|\cdot|_\infty$ norm implies that

$$\|z\|_{C_0(\omega)^n} = \sup\{|z(x)|_\infty : x \in \omega\} \quad \forall z \in C_0(\omega)^n.$$

We recall that every scalar real measure $\mu \in \mathcal{M}(\omega)$ admits a Jordan decomposition $\mu = \mu^+ - \mu^-$, where $\mu^+$ and $\mu^-$ are positive measures with disjoint supports. Further, if $h_\mu$ is the Radon-Nikodym derivative of $\mu$ with respect to $|\mu|$, then $\mu^+ = h^+ d|\mu|$ and $\mu^- = h^- d|\mu|$, where $h = h^+ - h^-$ is the decomposition of $h$ in positive and negative parts.

**Proposition 7** *If $\lambda \in C_0(\omega)^n \cap \partial g(\mu)$, then $\|\lambda_j\|_{C_0(\omega)} \leq 1$ for all $j = 1, \ldots, n$. Moreover, if $\mu_j \neq 0$, then the following properties hold*

1. $\|\lambda_j\|_{C_0(\omega)} = 1$, and
2. $\mathrm{supp}(\mu_j^+) \subset \{x \in \omega : \lambda_j(x) = +1\}$ and $\mathrm{supp}(\mu_j^-) \subset \{x \in \omega : \lambda_j(x) = -1\}$.

*Proof* Inserting (3.1) in (3.4) and (3.5) we get

$$\sum_{i=1}^n \langle \mu_i, \lambda_i \rangle = \sum_{i=1}^n \|\mu_i\|_{\mathcal{M}(\omega)}, \tag{3.9}$$

$$\sum_{i=1}^n \langle \nu_i, \lambda_i \rangle \leq \sum_{i=1}^n \|\nu_i\|_{\mathcal{M}(\omega)} \quad \forall \nu \in \mathcal{M}(\omega)^n. \tag{3.10}$$

Let us fix $1 \leq j \leq n$ and take in (3.10) $\nu_i = 0$ for every $i \neq j$ and $\nu_j = \pm \delta_x$ with $x \in \omega$ arbitrary. Then, we obtain

$$\pm \lambda_j(x) = \langle \nu_j, \lambda_j \rangle \leq \|\nu_j\|_{\mathcal{M}(\omega)} = 1.$$

This proves that $|\lambda_j(x)| \leq 1 \ \forall x \in \omega$ for every $j$. Then, from (3.9) we infer

$$\sum_{i=1}^{n} \|\mu_i\|_{\mathcal{M}(\omega)} = \sum_{i=1}^{n} \langle \mu_i, \lambda_i \rangle \leq \sum_{i=1}^{n} \|\mu_i\|_{\mathcal{M}(\omega)} \|\lambda_i\|_{C_0(\omega)} \leq \sum_{i=1}^{n} \|\mu_i\|_{\mathcal{M}(\omega)}.$$

This implies that $\|\lambda_i\|_{C_0(\omega)} = 1$ and $\langle \mu_i, \lambda_i \rangle = \|\mu_i\|_{\mathcal{M}(\omega)}$ for every $i$ such that $\mu_i \neq 0$. Hence, 1. holds. The second part was proved in [6, Lemma 3.4].

Now, we compute the directional derivatives of $g'(\mu; \nu)$. Then, we have the following expression which is similar but different from the one obtained in Proposition 5.

**Proposition 8** *Let $\mu, \nu \in \mathcal{M}(\omega)^n$, then*

$$g'(\mu; \nu) = \int_{\omega} h_\nu \, d\mu + \|\nu_s\|_{\mathcal{M}(\omega)^n} = \sum_{j=1}^{n} \left\{ \int_{\omega} h_{\nu_j} \, d\mu_j + \|(\nu_j)_s\|_{\mathcal{M}(\omega)} \right\}, \qquad (3.11)$$

*where $\nu_j = (\nu_j)_a + (\nu_j)_s = h_{\nu_j} d|\mu_j| + (\nu_j)_s$ is the Lebesgue decomposition of $\nu_j$ with respect to $|\mu_j|$ for $1 \leq j \leq n$.*

*Proof* For the proof it is enough use (3.1) to obtain

$$g'(\mu; \nu) = \lim_{\rho \searrow 0} \frac{\|\mu + \rho\nu\|_{\mathcal{M}(\omega)^n} - \|\mu\|_{\mathcal{M}(\omega)^n}}{\rho} = \sum_{i=1}^{n} \lim_{\rho \searrow 0} \frac{\|\mu_i + \rho\nu_i\|_{\mathcal{M}(\omega)^n} - \|\mu_i\|_{\mathcal{M}(\omega)^n}}{\rho}.$$

Then, we proceed as in the proof of [10, Proposition 3.3].

With the same proof we infer that Proposition 6 is also true for the $|\cdot|_\infty$ norm with (3.8) being interpreted as follows

$$G'(u; v) = (g \circ \nabla)'(u; v) = \int_{\omega} h_v \, d(\nabla u) + \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n}$$

$$= \sum_{j=1}^{n} \left\{ \int_{\omega} h_{v,j} \, d(\partial_{x_j} u) + \|(\partial_{x_j} v)_s\|_{\mathcal{M}(\omega)} \right\}, \qquad (3.12)$$

where $\partial_{x_j} v = h_{v,j} |\partial_{x_j} u| + (\partial_{x_j} v)_s$ is the Lebesgue decomposition of $\partial_{x_j} v$ with respect to $|\partial_{x_j} u|$.

*Remark 3* For the well-posedness analysis of (P) we could equally well use the norm $\left( \sum_{j=1}^{n} \|\partial_{x_j} u\|_{M(\omega)}^2 \right)^{1/2}$ instead of $\sum_{j=1}^{n} \|\partial_{x_j} u\|_{M(\omega)}$, which was mentioned in (3.2) above. For the sake of interest, we specify the changes which become necessary for this norm when compared to the assertions made in Proposition 7:

1. $\left( \sum_{j=1}^{n} \|\lambda_j\|_{C_0(\omega)}^2 \right)^{1/2} \leq 1$.
2. If $\mu \neq 0$, then $\left( \sum_{j=1}^{n} \|\lambda_j\|_{C_0(\omega)}^2 \right)^{1/2} = 1$.
3. If $\mu_j \neq 0$, then $\text{supp}(\mu_j^+) \subset \{x \in \omega : \lambda_j(x) = +\|\lambda_j\|_{C_0(\omega)}\}$ and $\text{supp}(\mu_j^-) \subset \{x \in \omega : \lambda_j(x) = -\|\lambda_j\|_{C_0(\omega)}\}$.

Since, for our purposes, we see no advantages of this norm over the norms defined by the duality in (2.1) we shall not follow it up.

## 4 Existence of an optimal control and first order optimality conditions

The proof of the existence of an optimal control follows the lines of [9, Theorem 3.1] with the obvious modifications.

**Theorem 1** *Let us assume that one of the following assumptions hold.*

1. $\beta + \gamma > 0$.
2. *There exist $q \in [1, 2)$ and $C > 0$ such that*

$$\frac{\partial f}{\partial y}(x, y) \le C(1 + |y|^q) \quad \text{for a.a. } x \in \Omega \text{ and } \forall y \in \mathbb{R}.$$

*Then, problem* (P) *has at least one solution. Moreover, if $f$ is affine with respect to $y$, the solution is unique.*

Now, we prove the first order optimality conditions satisfied by any local minimum of (P).

**Theorem 2** *Let $\bar{u}$ be a local solution of* (P)*. Then, there exists $\bar{\lambda} \in \partial g(\nabla \bar{u})$ such that*

$$\alpha \langle \bar{\lambda}, \nabla v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} + \int_\omega \left( \bar{\varphi} + \gamma \bar{u} + \beta \int_\omega \bar{u} \, dz \right) v \, dx = 0 \quad \forall v \in BV(\omega) \cap L^2(\omega), \tag{4.1}$$

*where $\bar{\varphi} \in H_0^1(\Omega) \cap C(\bar{\Omega})$ is the adjoint state corresponding to $\bar{u}$.*

*Proof* Let us denote by $\bar{\varphi} \in C(\bar{\Omega}) \cap H_0^1(\Omega)$ the adjoint state corresponding to the local solution $\bar{u}$. Given $v \in BV(\omega) \cap L^2(\omega)$, from the local optimality of $\bar{u}$ and the convexity of $G$ we deduce for every $0 < \rho < 1$ small enough

$$0 \le \frac{J(\bar{u} + \rho v) - J(\bar{u})}{\rho} = \frac{F(\bar{u} + \rho v) - F(\bar{u})}{\rho} + \alpha \frac{G(\bar{u} + \rho v) - G(\bar{u})}{\rho}$$
$$\le \frac{F(\bar{u} + \rho v) - F(\bar{u})}{\rho} + \alpha [G(\bar{u} + v) - G(\bar{u})].$$

Passing to the limit as $\rho \to 0$ in the above inequality and using (2.10) we obtain for every $v \in BV(\omega)$

$$0 \le \int_\omega \left( \bar{\varphi}(x) + \gamma \bar{u}(x) + \beta \int_\omega \bar{u} \, ds \right) v(x) \, dx + \alpha [G(\bar{u} + v) - G(\bar{u})].$$

Replacing $v$ by $u - \bar{u}$, this inequality can be written

$$-\frac{1}{\alpha} \int_\omega \left( \bar{\varphi} + \gamma \bar{u} + \beta \int_\omega \bar{u} \, ds \right)(u - \bar{u}) \, dx + G(\bar{u}) \le G(u) \quad \forall u \in BV(\omega) \cap L^2(\omega).$$

This along with (3.7) implies

$$-\frac{1}{\alpha} \left( \bar{\varphi} + \gamma \bar{u} + \beta \int_\omega \bar{u} \, ds \right) \in \partial G(\bar{u}) = \nabla^* \partial g(\nabla \bar{u}).$$

Hence, there exists $\bar{\lambda} \in \partial g(\nabla \bar{u}) \subset [\mathcal{M}(\omega)^n]^*$ such that

$$\langle \bar{\lambda}, \nabla v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} = \frac{-1}{\alpha} \int_\omega \left[ \bar{\varphi} + \beta \int_\omega \bar{u} \, ds \right] v \, dx \quad \forall v \in BV(\omega) \cap L^2(\omega),$$

which implies (4.1).

*Remark 4* As mentioned after inequality (3.5), the restriction of $\bar{\lambda}$ to $L^1(\omega)^n$, denoted by $\bar{\lambda}|_{L^1}$, can be identified with a function in $L^\infty(\omega)^n$. Then, if we take $v \in C_0^\infty(\omega)^n$ in (4.1), we deduce that the identity

$$\mathrm{div}\bar{\lambda}|_{L^1} = \bar{\varphi} + \gamma\bar{u} + \beta \int_\omega \bar{u}\,dx$$

holds in the distribution sense, and hence $\mathrm{div}\bar{\lambda}|_{L^1} \in BV(\omega) \subset L^p(\omega)$ for $p = n/(n-1)$. Then, it is well known that $\bar{\lambda}|_{L^1} \cdot \boldsymbol{\nu}$ is defined in a trace sense on $\partial\omega$, where $\boldsymbol{\nu}$ is the outward unit normal vector to $\partial\omega$. Moreover, taking $v \in C^\infty(\bar{\omega})^n$ in (4.1) and using Green's formula, we infer that $\bar{\lambda}|_{L^1} \cdot \boldsymbol{\nu} = 0$ on $\partial\omega$. Hence, we conclude that $\bar{\lambda}|_{L^1}$ belongs to $L^\infty(\omega) \cap W_0^p(\mathrm{div};\omega) \cap BV(\omega)$. Here we use the standard notation

$$W_0^p(\mathrm{div};\omega) = \{v \in L^p(\omega)^n : \mathrm{div}\,v \in L^p(\omega) \text{ and } v \cdot \boldsymbol{\nu} = 0 \text{ on } \partial\omega\}.$$

See Proposition 7 of [2] for a related result.

Since $\bar{\lambda} \in [\mathcal{M}(\omega)^n]^*$ and $[\mathcal{M}(\omega)^n]^*$ is not a space of distributions, sometimes it can be more convenient to handle a different optimality system involving distributional spaces, mainly if we think of the numerical analysis. To this end, we present the following equivalent optimality conditions.

**Theorem 3** *Let us assume that $n = 2$. Given $\bar{u} \in BV(\omega)$, let $\bar{y}$ and $\bar{\varphi}$ be the associated state and adjoint state. Then, there exists $\bar{\lambda} \in \partial g(\nabla\bar{u})$ satisfying (4.1) if and only if there exists $\bar{\Phi} \in C_0(\omega)^n$ such that*

$$\alpha\langle\nabla v, \bar{\Phi}\rangle_{\mathcal{M}(\omega)^n, C_0(\omega)^n} + \int_\omega \left[\bar{\varphi} + \gamma\bar{u} + \beta\int_\omega \bar{u}\,ds\right]v\,dx = 0 \quad \forall v \in BV(\omega), \qquad (4.2)$$

$$\langle\nabla v, \bar{\Phi}\rangle_{\mathcal{M}(\omega)^n, C_0(\omega)^n} \le \|\nabla v\|_{\mathcal{M}(\omega)^n} \quad \forall v \in BV(\omega), \qquad (4.3)$$

$$\langle\nabla\bar{u}, \bar{\Phi}\rangle_{\mathcal{M}(\omega)^n, C_0(\omega)^n} = \|\nabla\bar{u}\|_{\mathcal{M}(\omega)^n}. \qquad (4.4)$$

*Proof* Assume that $\bar{\lambda} \in \partial g(\nabla\bar{u})$ satisfies (4.1). We define a linear form $T_0$ in $\mathcal{M}(\omega)^n$ as follows

$$D(T_0) = \{\nabla v : v \in BV(\omega)\} \quad \text{and} \quad T_0(\mu) = \langle\bar{\lambda}, \nabla v\rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} \text{ if } \mu = \nabla v.$$

From (3.4) and (3.5) we have

$$T_0(\nabla\bar{u}) = \|\nabla\bar{u}\|_{\mathcal{M}(\omega)^n}, \qquad (4.5)$$

$$T_0(\mu) \le \|\mu\|_{\mathcal{M}(\omega)^n} \quad \forall\mu \in D(T_0). \qquad (4.6)$$

Let us prove that $T_0$ is weakly* continuous on its domain. First, we prove that $T_0^{-1}(0)$ is sequentially weakly* closed in $\mathcal{M}(\omega)^n$. Indeed, let $\{\mu_k\}_k \subset T_0^{-1}(0)$ and $\mu \in \mathcal{M}(\omega)^n$ be such that $\mu_k \overset{*}{\rightharpoonup} \mu$ in $\mathcal{M}(\omega)^n$. By definition of $D(T_0)$ there exist elements $\{v_k\}_k \subset BV(\omega)$ such that $\mu_k = \nabla v_k$. Without loss of generality we assume that the integrals of each $v_k$ in $\omega$ are zero. Then, using (2.2), we know that $\{v_k\}_k$ is bounded in $BV(\omega)$. Therefore, there exist a subsequence, denoted in the same form, and an element $v \in BV(\omega)$ such that $v_k \overset{*}{\rightharpoonup} v$ in $BV(\omega)$. This implies that $\mu_k = \nabla v_k \overset{*}{\rightharpoonup} \nabla v$ in $\mathcal{M}(\omega)^n$. Hence, the identity $\nabla v = \mu$ holds and,

consequently, $\mu \in D(T_0)$. We have to prove that $T_0(\mu) = 0$. From the continuity of the embedding $BV(\omega) \subset L^2(\omega)$ due to $n = 2$ and the convergence $v_k \xrightarrow{*} v$ in $BV(\omega)$, we obtain that $v_k \rightharpoonup v$ in $L^2(\omega)$. Therefore, we get with (4.1)

$$
\begin{aligned}
0 &= \lim_{k \to \infty} T_0(\mu_k) = \lim_{k \to \infty} \langle \bar{\lambda}, \nabla v_k \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} \\
&= \lim_{k \to \infty} \frac{-1}{\alpha} \int_\omega \Big[ \bar{\varphi} + \gamma \bar{u} + \beta \int_\omega \bar{u} \, ds \Big] v_k \, dx \qquad\qquad (4.7) \\
&= \frac{-1}{\alpha} \int_\omega \Big[ \bar{\varphi} + \gamma \bar{u} + \beta \int_\omega \bar{u} \, ds \Big] v \, dx = \langle \bar{\lambda}, \nabla v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} = T_0(\mu).
\end{aligned}
$$

Thus $\mu \in T_0^{-1}(0)$ holds and, hence, $T_0^{-1}(0)$ is sequentially weakly* closed in $\mathcal{M}(\omega)^n$. As a consequence, we also have that $T_0^{-1}(0) \cap \bar{B}_r(0)$ is sequentially weakly* closed in $\mathcal{M}(\omega)^n$ for every closed ball $\bar{B}_r(0)$ of $\mathcal{M}(\omega)^n$ centered at 0 and arbitrary radius $r > 0$. Due to the fact that $C_0(\omega)^n$ is a separable Banach space, we have that $\bar{B}_r(0)$ is metrizable in the weak* topology of $\mathcal{M}(\omega)^n$. Hence, $T_0^{-1}(0) \cap \bar{B}_r(0)$ is sequentially weakly* closed in $\mathcal{M}(\omega)^n$ if and only if it is weak* closed. Finally, from the Krein-Šmulian Theorem [3, Theorem 3.33], we infer that $T_0^{-1}(0)$ is weak* closed in $\mathcal{M}(\omega)^n$, which implies the weak* continuity of $T_0$.

Hence, there exists a weakly* continuous linear form $T : \mathcal{M}(\omega)^n \longrightarrow \mathbb{R}$ extending $T_0$; [24, Theorem 3.6]. In this case, we know that $T$ can be identified with an element $\bar{\Phi} \in C_0(\omega)^n$, i.e.

$$
T(\mu) = \langle \mu, \bar{\Phi} \rangle_{\mathcal{M}(\omega)^n, C_0(\omega)^n} = \int_\omega \bar{\Phi} \, d\mu \quad \forall \mu \in \mathcal{M}(\omega)^n;
$$

see [3, Proposition 3.14]. The function $\bar{\Phi}$ fulfills (4.2)–(4.4). Indeed, (4.2) follows from the definition of $T_0$ and (4.1), and (4.3)-(4.4) are the same as (4.5)-(4.6).

Reciprocally, assume that $\bar{\Phi} \in C_0(\omega)^n$ satisfies (4.2)–(4.4). This time we define the linear operator

$$
D(T_0) = \{ \nabla v : v \in BV(\omega) \} \quad \text{and} \quad T_0(\mu) = \langle \nabla v, \bar{\Phi} \rangle_{\mathcal{M}(\omega)^n, C_0(\omega)^n} \quad \text{if } \mu = \nabla v.
$$

From (4.3) we know that $T_0$ is a continuous operator in $D(T_0)$ for the strong topology of $\mathcal{M}(\omega)^n$, and $\|T_0\|_{[\mathcal{M}(\omega)^n]^*} \leq 1$. Hence, the Hahn-Banach theorem implies the existence of an operator $\bar{\lambda} \in [\mathcal{M}(\omega)^n]^*$ extending $T_0$ and such that $\|\bar{\lambda}\|_{[\mathcal{M}(\omega)^n]^*} \leq 1$. This along with (4.3) implies that

$$
\begin{aligned}
\langle \bar{\lambda}, \nabla \bar{u} \rangle &= \|\nabla \bar{u}\|_{\mathcal{M}(\omega)^n}, \\
\langle \bar{\lambda}, \nu \rangle &\leq \|\nu\|_{\mathcal{M}(\omega)^n} \quad \forall \nu \in \mathcal{M}(\omega)^n.
\end{aligned}
$$

Hence, we have $\bar{\lambda} \in \partial g(\nabla \bar{u})$; see (3.3)–(3.5). Finally, (4.1) follows from (4.2) and the definition of $T_0$. This concludes the proof.

*Remark 5* Theorem 3 is still valid in dimension $n = 3$ if we take $\gamma = 0$ and we assume that the nonlinearity of $f(x, y)$ has a polynomial growth of arbitrary order with respect to the variable $y$; see Remark 1. Indeed, let us observe that the limit (4.7) is still valid because $v_k \rightharpoonup v$ in $L^{3/2}(\Omega)$ and $\bar{\varphi} + \beta \int_\omega \bar{u} \, ds$ is a continuous function in $\bar{\Omega}$.

*Remark 6* It would be interesting to prove the existence of a function $\bar{\Phi} \in C_0(\omega)^n \cap \partial g(\nabla \bar{u})$ satisfying (4.3)–(4.5). Indeed, Theorem 3 does not guarantee that $\|\Phi\|_{C_0(\omega)^n} \leq 1$. In this hypothetic case, we could deduce from Propositions 4 and 7 the following sparsity structure of $\nabla \bar{u}$.

1. For the $|\cdot|_2$ norm, if $\nabla \bar{u} \neq 0$ we have $\|\bar{\Phi}\|_{C_0(\omega)^n} = 1$ and

$$\mathrm{supp}(\nabla \bar{u}) \subset \{x \in \omega : |\bar{\Phi}(x)|_2 = 1\}.$$

2. For the $|\cdot|_\infty$ norm, for any $1 \leq j \leq n$ such that if $\partial_{x_j} \bar{u} \neq 0$ we have $\|\bar{\Phi}_j\|_{C_0(\omega)} = 1$, and

$$\mathrm{supp}([\partial_{x_j} u]^+) \subset \{x \in \omega : \bar{\Phi}_j(x) = +1\},$$
$$\mathrm{supp}([\partial_{x_j} \bar{u}]^-) \subset \{x \in \omega : \bar{\Phi}_j(x) = -1\}.$$

## 5 Second order optimality conditions

The goal of this section is to prove necessary and sufficient second order optimality conditions for problem (P). In the whole section, $\bar{u}$ will denote a fixed element of $BV(\omega) \cap L^2(\omega)$ satisfying the optimality conditions given in Theorem 2. As in Section 3, we will distinguish the cases where the norms $|\cdot|_2$ and $|\cdot|_\infty$ in $\mathbb{R}^n$ are used in the definition of the measure $\|\nabla u\|_{\mathcal{M}(\omega)^n}$. First, we analyze the case corresponding to the norm $|\cdot|_\infty$. For this case, the gap between the necessary and sufficient second order optimality conditions is smaller than the ones that we provide for the case of the norm $|\cdot|_2$.

5.1 The use of the $|\cdot|_\infty$ norm

As pointed out in (3.1), the use of the $|\cdot|_\infty$ norm in $\mathbb{R}^n$ leads to the identity

$$\|\nabla v\|_{\mathcal{M}(\omega)^n} = \sum_{j=1}^{n} \|\partial_{x_j} v\|_{\mathcal{M}(\omega)} = \sum_{j=1}^{n} \left\{ \int_\omega |h_{v,j}| \, d|\partial_{x_j} \bar{u}| + \|(\partial_{x_j} v)_s\|_{\mathcal{M}(\omega)} \right\} \quad (5.1)$$

$\forall v \in BV(\omega)$, where $\partial_{x_j} v = h_{v,j} |\partial_{x_j} \bar{u}| + (\partial_{x_j} v)_s$ is the Lebesgue decomposition of $\partial_{x_j} v$ with respect to the measure $|\partial_{x_j} \bar{u}|$, $1 \leq j \leq n$. Moreover, for every $1 \leq j \leq n$ there exists a Borel function $\bar{h}_j$ such that

$$|\bar{h}_j(x)| = 1, \quad |\partial_{x_j} \bar{u}| - \text{a.e., and } \partial_{x_j} \bar{u} = \bar{h}_j |\partial_{x_j} \bar{u}|. \quad (5.2)$$

In the sequel, we will denote $h_v = (h_{v,1}, \ldots, h_{v,n})$ and $\bar{h} = (\bar{h}_1, \ldots, \bar{h}_n)$.

First, we state the second order necessary optimality conditions. To this end we define the cone of critical directions $C_{\bar{u}}$ as the closure in $L^2(\omega)$ of the cone

$$C_{\bar{u}} = \{v \in BV(\omega) \cap L^2(\omega) : F'(\bar{u})v + \alpha G'(\bar{u}; v) = 0$$
$$\text{and } h_{v,j} \in L^2(|\partial_{x_j} \bar{u}|), \ 1 \leq j \leq n\}. \quad (5.3)$$

Then, we have the following result.

**Theorem 4** *If $\bar{u}$ is a local minimum of* (P), *then $F''(\bar{u})v^2 \geq 0 \ \forall v \in C_{\bar{u}}$.*

*Proof* We will prove the result for every $v \in E_{\bar{u}}$. Then, the theorem follows by using the continuity of quadratic form $v \in L^2(\omega) \to F''(\bar{u})v^2 \in \mathbb{R}$. Given $v \in E_{\bar{u}}$ and $\rho > 0$ we set

$$\omega_{\rho,j} = \{x \in \omega : \rho|h_{v,j}(x)| \le \frac{1}{2}\} \quad 1 \le j \le n.$$

We have with Schwarz inequality

$$|\partial_{x_j}\bar{u}|(\omega \setminus \omega_{\rho,j}) \le 2\rho \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)| \, d|\partial_{x_j}\bar{u}|$$

$$\le 2\rho\sqrt{|\partial_{x_j}\bar{u}|(\omega \setminus \omega_{\rho,j})} \Big( \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)|^2 \, d|\partial_{x_j}\bar{u}| \Big)^{1/2},$$

which implies

$$\sqrt{|\partial_{x_j}\bar{u}|(\omega \setminus \omega_{\rho,j})} \le 2\rho\Big( \int_{\omega \setminus \omega_{\rho,j}} |h_{v.j}(x)|^2 \, d|\partial_{x_j}\bar{u}| \Big)^{1/2} \quad 1 \le j \le n. \tag{5.4}$$

Taking into account (5.2) we get for $1 \le j \le n$

$$\frac{|\bar{h}_j(x) + \rho h_{v,j}(x)| - |\bar{h}_j(x)|}{\rho} = h_{v,j}(x)\bar{h}_j(x) \quad [|\partial_{x_j}\bar{u}|]-\text{a.e. } x \in \omega_{\rho,j}.$$

Using this identity and (5.1) we get

$$\frac{G(\bar{u} + \rho v) - G(\bar{u})}{\rho}$$

$$= \sum_{j=1}^{n} \Big\{ \int_{\omega_{\rho,j}} \frac{|\bar{h}_j + \rho h_{v,j}| - |\bar{h}_j|}{\rho} d|\partial_{x_j}\bar{u}| + \|(\partial_{x_j}v)_s\|_{\mathcal{M}(\omega)^n} \Big\}$$

$$+ \sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} \frac{|\bar{h}_j + \rho h_{v,j}| - |\bar{h}|}{\rho} d|\partial_{x_j}\bar{u}|$$

$$= \sum_{j=1}^{n} \Big\{ \int_{\omega_{\rho,j}} (h_{v,j}\bar{h}_j) \, d|\partial_{x_j}\bar{u}| + \|(\partial_{x_j}v)_s\|_{\mathcal{M}(\omega)^n} \Big\}$$

$$+ \sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} \frac{|\bar{h}_j + \rho h_{v,j}| - |\bar{h}_j|}{\rho} d|\partial_{x_j}\bar{u}|$$

$$= \sum_{j=1}^{n} \Big\{ \int_{\omega} h_{v,j} \, d\partial_{x_j}\bar{u} + \|(\partial_{x_j}v)_s\|_{\mathcal{M}(\omega)^n} \Big\}$$

$$+ \sum_{j=1}^{n} \Big\{ \int_{\omega \setminus \omega_{\rho,j}} \frac{|\bar{h}_j + \rho h_{v,j}| - |\bar{h}_j|}{\rho} d|\partial_{x_j}\bar{u}| - \int_{\omega \setminus \omega_{\rho,j}} (h_{v,j}\bar{h}_j) \, d|\partial_{x_j}\bar{u}| \Big\}.$$

Now, from (3.12), (5.2), Schwarz inequality, and (5.4) we infer

$$\frac{G(\bar{u} + \rho v) - G(\bar{u})}{\rho} \leq G'(\bar{u}; v) + 2\sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}| d|\partial_{x_j}\bar{u}|$$

$$\leq G'(\bar{u}; v) + 2\sum_{j=1}^{n} \sqrt{|\partial_{x_j}\bar{u}|(\omega \setminus \omega_{\rho,j})} \Big( \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)|^2 \, d|\partial_{x_j}\bar{u}| \Big)^{1/2}$$

$$\leq G'(\bar{u}; v) + 4\rho \sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)|^2 \, d|\partial_{x_j}\bar{u}|.$$

Next we use the local optimality of $\bar{u}$. By a Taylor expansion of $F$ around $\bar{u}$ and using that $v \in E_{\bar{u}}$, we get for $\rho > 0$ small enough

$$0 \leq J(\bar{u} + \rho v) - J(\bar{u}) = \rho[F'(\bar{u})v + \alpha G'(\bar{u}; v)]$$

$$+ \frac{\rho^2}{2} \Big( F''(\bar{u} + \theta_\rho v)v^2 + 8\alpha \sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)|^2 \, d|\partial_{x_j}\bar{u}| \Big)$$

$$= \frac{\rho^2}{2} \Big( F''(\bar{u} + \theta_\rho v)v^2 + 8\alpha \sum_{j=1}^{n} \int_{\omega \setminus \omega_{\rho,j}} |h_{v,j}(x)|^2 \, d|\partial_{x_j}\bar{u}| \Big)$$

with $0 \leq \theta_\rho \leq 1$. Dividing the above expression by $\rho^2/2$, passing to the limit as $\rho \to 0$, and taking into account that $h_{v,j} \in L^2(|\partial_{x_j}\bar{u}|)$ and $|\partial_{x_j}\bar{u}|(\omega \setminus \omega_{\rho,j}) \to 0$, we conclude that $F''(\bar{u})v^2 \geq 0$.

For the sufficient second order conditions we introduce the critical cones

$$C_{\bar{u}}^{\tau} = \{v \in BV(\omega) \cap L^2(\omega) : F'(\bar{u})v + \alpha G'(\bar{u}; v) \leq \tau \|z_v\|_{L^2(\Omega)}\}, \qquad (5.5)$$

where $\tau > 0$ and $z_v = S'(\bar{u})v$. The reader is referred to [4] for some second order conditions based on these cones; see also [12] and [13]. Let us observe that (4.1) and the inequality $G'(\bar{u}; v) \geq \langle \bar{\lambda}, \nabla v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n}$ imply that $\forall v \in BV(\omega) \cap L^2(\omega)$

$$F'(\bar{u})v + \alpha G'(\bar{u}; v) \geq F'(\bar{u})v + \alpha \langle \bar{\lambda}, \nabla v \rangle_{[\mathcal{M}(\omega)^n]^*, \mathcal{M}(\omega)^n} = 0. \qquad (5.6)$$

**Theorem 5** *Let $\bar{u} \in BV(\omega) \cap L^2(\omega)$ satisfy the first order optimality conditions stated in Theorem 2 and the second order condition*

$$\exists \delta > 0 \ \text{and} \ \exists \tau > 0 : F''(\bar{u})v^2 \geq \delta \|z_v\|_{L^2(\Omega)}^2 \quad \forall v \in C_{\bar{u}}^{\tau}. \qquad (5.7)$$

*Then, there exist $\kappa > 0$ and $\varepsilon > 0$ such that*

$$J(\bar{u}) + \frac{\kappa}{2} \|y_u - \bar{y}\|_{L^2(\omega)}^2 \leq J(u) \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon, \qquad (5.8)$$

*where $y_u = S(u)$ and $\bar{y} = S(\bar{u})$.*

*Proof* We follow the proof of [4, Theorem 3.6] with some changes. First, from [4, Lemma 2.7] we deduce the existence of $\varepsilon_0 > 0$ such that

$$|[F''(u) - F''(\bar{u})]v^2| \leq \frac{\delta}{2} \|z_v\|_{L^2(\Omega)}^2 \ \forall v \in L^2(\omega) \ \text{and all} \ \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon_0. \qquad (5.9)$$

Moreover, from Proposition 2 we deduce the existence of a constant $C_1 > 0$ such that

$$\|z_v\|_{L^2(\Omega)} = \|S'(\bar{u})v\|_{L^2(\Omega)} \leq C_1 \|v\|_{L^2(\omega)} \quad \forall v \in L^2(\omega). \tag{5.10}$$

Now, from (2.7) we have that there exists a constant $K$ such that $\|y_u\|_{C(\bar{\Omega})} \leq K$ if $\|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon_0$. From the adjoint state equation (2.12) and (2.4) we deduce that $\|\varphi_u\|_{C(\bar{\Omega})} \leq K'$ for every $\|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon_0$ and some constant $K'$. Finally, using these estimates, (2.5) and the expression (2.11) we infer the existence of a constant $C_2 > 0$ such that

$$F''(u)v^2 \geq \gamma \|v\|_{L^2(\omega)}^2 - C_2 \|z_v\|_{L^2(\Omega)}^2 \quad \text{for all } \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon_0 \text{ and } \forall v \in L^2(\omega). \tag{5.11}$$

Now, we set

$$\varepsilon = \min\left\{\varepsilon_0, \frac{2\tau}{(\delta + C_2)C_1}\right\}$$

with $\tau$ and $\delta$ given in (5.7). Let $u \in BV(\omega) \cap L^2(\omega)$ such that $\|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon$. We distinguish two cases.

*Case I: $u - \bar{u} \in C_{\bar{u}}^\tau$.* Making a Taylor expansion of $F$ around $\bar{u}$, using the convexity of $G$ and (5.6), (5.7) and (5.9), we get for some $0 \leq \theta \leq 1$

$$J(u) - J(\bar{u}) \geq [F'(\bar{u})(u - \bar{u}) + \alpha G'(\bar{u}; u - \bar{u})] + \frac{1}{2}F''(\bar{u} + \theta(u - \bar{u}))(u - \bar{u})^2$$

$$\geq \frac{1}{2}F''(\bar{u})(u - \bar{u})^2 + \frac{1}{2}[F''(\bar{u} + \theta(u - \bar{u})) - F''(\bar{u})](u - \bar{u})^2$$

$$\geq \frac{\delta}{2}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2 - \frac{\delta}{4}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2 = \frac{\delta}{4}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2. \tag{5.12}$$

*Case II: $u - \bar{u} \notin C_{\bar{u}}^\tau$.* This implies that

$$F'(\bar{u})(u - \bar{u}) + \alpha G'(\bar{u}; u - \bar{u}) > \tau \|z_{u-\bar{u}}\|_{L^2(\Omega)}. \tag{5.13}$$

Moreover, from (5.10) and the definition of $\varepsilon$ we infer

$$\|z_{u-\bar{u}}\|_{L^2(\Omega)} \leq C_1 \|u - \bar{u}\|_{L^2(\Omega)} \leq \frac{2\tau}{\delta + C_2},$$

and therefore

$$\frac{\delta + C_2}{2\tau}\|z_{u-\bar{u}}\|_{L^2(\Omega)} \leq 1. \tag{5.14}$$

Using again the convexity of $G$, (5.11), (5.13) and (5.14) we infer

$$J(u) - J(\bar{u}) \geq [F'(\bar{u})(u - \bar{u}) + \alpha G'(\bar{u}; u - \bar{u})] + \frac{1}{2}F''(\bar{u} + \theta(u - \bar{u}))(u - \bar{u})^2$$

$$\geq \tau \|z_{u-\bar{u}}\|_{L^2(\Omega)} - C_2 \|z_{u-\bar{u}}\|_{L^2(\Omega)}^2$$

$$\geq \frac{\delta + C_2}{2}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2 - \frac{C_2}{2}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2 = \frac{\delta}{2}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2. \tag{5.15}$$

From (5.12) and (5.15) we deduce that [4, page 2364]

$$J(u) - J(\bar{u}) \geq \frac{\delta}{4}\|z_{u-\bar{u}}\|_{L^2(\Omega)}^2 \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon.$$

Finally, choosing $\varepsilon$ still smaller, if necessary, we have that [4, page 2364]

$$\frac{1}{2}\|y_u - \bar{y}\|_{L^2(\Omega)} \leq \|z_{u-\bar{u}}\|_{L^2(\Omega)} \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon.$$

The last two inequalities imply (5.8) with $\kappa = \frac{\delta}{8}$.

We observe that (5.7) is a sufficient second order condition for strict local optimality of $\bar{u}$ in the $L^2(\omega)$ sense. Moreover, by using (5.8) we can prove stability of the optimal states with respect to perturbations in the data of the control problem. However, it does not provide information on the optimal controls. If $\gamma > 0$ we can change (5.7) by a stronger assumption leading to a quadratic growth of the controls instead of the states; i.e. $\|y_u - \bar{y}\|^2_{L^2(\Omega)}$ can be replaced by $\|u - \bar{u}\|^2_{L^2(\omega)}$ in (5.8). However, if $\gamma = 0$, then this is not possible; see [4].

**Theorem 6** *Suppose that $\gamma > 0$ and let $\bar{u} \in BV(\omega) \cap L^2(\omega)$ satisfy the first order optimality conditions stated in Theorem 2 and the second order condition*

$$\exists \delta > 0 \text{ and } \exists \tau > 0 : F''(\bar{u})v^2 \geq \delta\|v\|^2_{L^2(\omega)} \quad \forall v \in C^\tau_{\bar{u}}. \tag{5.16}$$

*Then, there exist $\kappa > 0$ and $\varepsilon > 0$ such that*

$$J(\bar{u}) + \frac{\kappa}{2}\|u - \bar{u}\|^2_{L^2(\omega)} \leq J(u) \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon. \tag{5.17}$$

*Proof* Using again [4, Lemma 2.7] along with (5.10) we infer the existence of $\varepsilon > 0$ such that

$$|[F''(u) - F''(\bar{u})]v^2| \leq \frac{\delta}{2}\|v\|^2_{L^2(\Omega)} \ \forall v \in L^2(\omega) \text{ and all } \|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon. \tag{5.18}$$

Arguing similarly to (5.12), but using (5.16) and (5.18) we obtain for every $u \in BV(\omega) \cap L^2(\omega)$ such that $\|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon$ and $u - \bar{u} \in C^\tau_{\bar{u}}$

$$
\begin{aligned}
J(u) - J(\bar{u}) &\geq [F'(\bar{u})(u - \bar{u}) + \alpha G'(\bar{u}; u - \bar{u})] + \frac{1}{2}F''(\bar{u} + \theta(u - \bar{u}))(u - \bar{u})^2 \\
&\geq \frac{1}{2}F''(\bar{u})(u - \bar{u})^2 + \frac{1}{2}[F''(\bar{u} + \theta(u - \bar{u})) - F''(\bar{u})](u - \bar{u})^2 \\
&\geq \frac{\delta}{2}\|u - \bar{u}\|^2_{L^2(\omega)} - \frac{\delta}{4}\|u - \bar{u}\|^2_{L^2(\omega)} = \frac{\delta}{4}\|u - \bar{u}\|^2_{L^2(\omega)}. \tag{5.19}
\end{aligned}
$$

Thus, (5.17) holds with $\kappa = \frac{\delta}{2}$ assuming that $u - \bar{u} \in C^\tau_{\bar{u}}$. Now, we argue by contradiction, and we assume that there do not exist $\kappa > 0$ and $\varepsilon > 0$ such that (5.17) holds for all the elements $u \in BV(\omega) \cap L^2(\omega)$ with $\|u - \bar{u}\|_{L^2(\omega)} \leq \varepsilon$. This implies that for every integer $k > 0$, there exists an element $u_k \in BV(\omega) \cap L^2(\omega)$ with

$$\|u_k - \bar{u}\|_{L^2(\omega)} \leq \frac{1}{k} \quad \text{and} \quad J(\bar{u}) + \frac{1}{2k}\|u_k - \bar{u}\|^2_{L^2(\omega)} > J(u_k). \tag{5.20}$$

From (5.19) we know that $u_k - \bar{u} \notin C^\tau_{\bar{u}}$, hence with (5.14)

$$F'(\bar{u})(u_k - \bar{u}) + \alpha G'(\bar{u}; u_k - \bar{u}) > \tau\|z_{u_k-\bar{u}}\|_{L^2(\Omega)} \geq \frac{\delta + C_2}{2}\|z_{u_k-\bar{u}}\|^2_{L^2(\Omega)} \tag{5.21}$$

for every $k$ large enough. Using (5.11), (5.20) and (5.21) we obtain

$$
\begin{aligned}
\frac{1}{2k}\|u_k - \bar{u}\|^2_{L^2(\omega)} &> J(u_k) - J(\bar{u}) \\
&\geq [F'(\bar{u})(u_k - \bar{u}) + \alpha G'(\bar{u}; u_k - \bar{u})] + \frac{1}{2}F''(\bar{u} + \theta_k(u_k - \bar{u}))(u_k - \bar{u})^2 \\
&\geq \frac{\delta + C_2}{2}\|z_{u-\bar{u}}\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u_k - \bar{u}\|^2_{L^2(\omega)} - \frac{C_2}{2}\|z_{u_k-\bar{u}}\|^2_{L^2(\Omega)} \geq \frac{\gamma}{2}\|u_k - \bar{u}\|^2_{L^2(\omega)},
\end{aligned}
$$

which is a contradiction because $\gamma > 0$.

5.2 The use of the $|\cdot|_2$ norm

Given an element $v \in BV(\omega)$, we consider the Lebesgue decomposition of $\nabla v$ with respect to the positive measure $|\nabla \bar{u}|$: $\nabla v = h_v d|\nabla \bar{u}| + (\nabla v)_s$. Hence, we have

$$\|\nabla v\|_{\mathcal{M}(\omega)^n} = \int_\omega |h_v(x)|_2 \, d|\nabla \bar{u}| + \|(\nabla v)_s\|_{\mathcal{M}(\omega)}. \qquad (5.22)$$

We also set $\nabla \bar{u} = \bar{h}|\nabla \bar{u}|$, where $|\bar{h}(x)|_2 = 1$ $|\nabla \bar{u}|$-a.e. in $\omega$. Then, we have with (3.8)

$$G'(\bar{u}; v) = \int_\omega (\bar{h} \cdot h_v) \, d|\nabla \bar{u}| + \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n}. \qquad (5.23)$$

Now, we define the cone of critical directions

$$C_{\bar{u}} = \{v \in BV(\omega) \cap L^2(\omega) : F'(\bar{u})v + \alpha G'(\bar{u}; v) = 0 \text{ and } |h_v|_2 \in L^2(|\nabla \bar{u}|)\}. \qquad (5.24)$$

Then, we have the following second order necessary optimality conditions.

**Theorem 7** *If $\bar{u}$ is a local minimum of* (P), *then*

$$F''(\bar{u})v^2 + \alpha \int_\omega \left( |h_v(x)|_2^2 - (\bar{h}(x) \cdot h_v(x))^2 \right) d|\nabla \bar{u}| \geq 0 \quad \forall v \in C_{\bar{u}}. \qquad (5.25)$$

*Proof* For fixed $v \in C_{\bar{u}}$ and given $\rho > 0$, we define

$$\omega_\rho = \{x \in \omega : \rho|h_v(x)|_2 \leq \frac{1}{2}\}.$$

Arguing as in the proof of Theorem 4 we get the following inequality analogous to (5.4)

$$\sqrt{|\nabla \bar{u}|(\omega \setminus \omega_\rho)} \leq 2\rho \Big( \int_{\omega \setminus \omega_\rho} |h_v(x)|_2^2 \, d|\nabla \bar{u}| \Big)^{1/2}. \qquad (5.26)$$

Using the differentiability of the $|\cdot|_2$-norm $x \in \mathbb{R}^n \to |x|_2$ for every $x \neq 0$, the fact that $|\bar{h}(x)|_2 = 1$ $|\nabla \bar{u}|$-a.e., (5.23), the Schwarz inequality, and (5.26), we get for $0 \leq \theta_\rho(x) \leq 1$

$$\frac{G(\bar{u} + \rho v) - G(\bar{u})}{\rho}$$

$$= \int_{\omega_\rho} \frac{|\bar{h} + \rho h_v|_2 - |\bar{h}|_2}{\rho} \, d|\nabla \bar{u}| + \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n} + \int_{\omega \setminus \omega_\rho} \frac{|\bar{h} + \rho h_v|_2 - |\bar{h}|_2}{\rho} \, d|\nabla \bar{u}|$$

$$= \int_{\omega_\rho} \left[ \bar{h} \cdot h_v + \frac{\rho}{2} \Big( \frac{|h_v|_2^2}{|\bar{h} + \theta_\rho \rho h_v|_2} - \frac{(\bar{h} + \theta_\rho \rho h_v) \cdot h_v}{|\bar{h} + \theta_\rho \rho h_v|_2^3} \Big) \right] d|\nabla \bar{u}|$$

$$+ \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n} + \int_{\omega \setminus \omega_\rho} \frac{|\bar{h} + \rho h_v|_2 - |\bar{h}|_2}{\rho} \, d|\nabla \bar{u}|$$

$$\leq \int_\omega (\bar{h} \cdot h_v) \, d|\nabla \bar{u}| + \frac{\rho}{2} \int_{\omega_\rho} \left[ \frac{|h_v|_2^2}{|\bar{h} + \theta_\rho \rho h_v|_2} - \frac{(\bar{h} + \theta_\rho \rho h_v) \cdot h_v}{|\bar{h} + \theta_\rho \rho h_v|_2^3} \right] d|\nabla \bar{u}|$$

$$+ \|(\nabla v)_s\|_{\mathcal{M}(\omega)^n} + 2 \int_{\omega \setminus \omega_\rho} |h_v|_2 \, d|\nabla \bar{u}|$$

$$\leq G'(\bar{u}; v) + \frac{\rho}{2} \Big\{ \int_{\omega_\rho} \left[ \frac{|h_v|_2^2}{|\bar{h} + \theta_\rho \rho h_v|_2} - \frac{(\bar{h} + \theta_\rho \rho h_v) \cdot h_v}{|\bar{h} + \theta_\rho \rho h_v|_2^3} \right] d|\nabla \bar{u}| + 8 \int_{\omega \setminus \omega_\rho} |h_v|_2^2 \, d|\nabla \bar{u}| \Big\}.$$

Using this inequality and the local optimality of $\bar{u}$, we infer with $u_\rho = \bar{u} + \theta_\rho \rho v$, $0 \le \theta_\rho \le 1$,

$$0 \le J(\bar{u} + \rho v) - J(\bar{u}) = \rho\Big[ F'(\bar{u})v + \alpha \frac{G(\bar{u} + \rho v) - G(\bar{u})}{\rho} \Big] + \frac{\rho^2}{2} F''(u_\rho)v^2$$

$$\le \rho[F'(\bar{u})v + \alpha G'(\bar{u}; v)] + \frac{\rho^2}{2}\Big\{ F''(u_\rho)v^2$$

$$+ \alpha \int_{\omega_\rho} \Big[ \frac{|h_v|_2^2}{|\bar{h} + \theta_\rho \rho h_v|_2} - \frac{(\bar{h} + \theta_\rho \rho h_v) \cdot h_v}{|\bar{h} + \theta_\rho \rho h_v|_2^3} \Big] d|\nabla \bar{u}| + 8\alpha \int_{\omega \backslash \omega_\rho} |h_v|_2^2 \, d|\nabla \bar{u}| \Big\}.$$

Now, taking into account that $v \in C_{\bar{u}}$ and dividing the above inequality by $\rho^2/2$ we get

$$0 \le F''(u_\rho)v^2 + \alpha \int_{\omega_\rho} \Big[ \frac{|h_v|_2^2}{|\bar{h} + \theta_\rho \rho h_v|_2} - \frac{(\bar{h} + \theta_\rho \rho h_v) \cdot h_v}{|\bar{h} + \theta_\rho \rho h_v|_2^3} \Big] d|\nabla \bar{u}| + 8\alpha \int_{\omega \backslash \omega_\rho} |h_v|_2^2 \, d|\nabla \bar{u}|.$$

Finally, using that $|\nabla \bar{u}|(\omega \backslash \omega_\rho) \to 0$ as $\rho \to 0$, $|\bar{h}(x)|_2 = 1$, and

$$\frac{1}{2} \le 1 - \rho |h_v(x)|_2 \le |\bar{h} + \theta_\rho \rho h_v|_2 \le 1 + \rho |h_v(x)|_2 \le \frac{3}{2} \quad |\nabla \bar{u}|\text{-a.e. in } \omega_\rho,$$

we pass to the limit as $\rho \to 0$ in the above inequality with the aid of the Lebesgue dominated convergence theorem and we obtain (5.25).

The proofs of Theorems 5 and 6 can be used without changes to prove the following analogous results.

**Theorem 8** *Let $\bar{u} \in BV(\omega) \cap L^2(\omega)$ satisfy the first order optimality conditions stated in Theorem 2 and the second order condition*

$$\exists \delta > 0 \text{ and } \exists \tau > 0 : F''(\bar{u})v^2 \ge \delta \|z_v\|_{L^2(\Omega)}^2 \quad \forall v \in C_{\bar{u}}^\tau, \tag{5.27}$$

*where $C_{\bar{u}}^\tau$ is defined by (5.5). Then, there exist $\kappa > 0$ and $\varepsilon > 0$ such that*

$$J(\bar{u}) + \frac{\kappa}{2} \|y_u - \bar{y}\|_{L^2(\omega)}^2 \le J(u) \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \le \varepsilon, \tag{5.28}$$

*where $y_u = S(u)$ and $\bar{y} = S(\bar{u})$.*

**Theorem 9** *Suppose that $\gamma > 0$ and let $\bar{u} \in BV(\omega) \cap L^2(\omega)$ satisfy the first order optimality conditions stated in Theorem 2 and the second order condition*

$$\exists \delta > 0 \text{ and } \exists \tau > 0 : F''(\bar{u})v^2 \ge \delta \|v\|_{L^2(\omega)}^2 \quad \forall v \in C_{\bar{u}}^\tau. \tag{5.29}$$

*Then, there exist $\kappa > 0$ and $\varepsilon > 0$ such that*

$$J(\bar{u}) + \frac{\kappa}{2} \|u - \bar{u}\|_{L^2(\omega)}^2 \le J(u) \quad \forall u \in BV(\omega) \cap L^2(\omega) : \|u - \bar{u}\|_{L^2(\omega)} \le \varepsilon. \tag{5.30}$$

*Remark 7* To reduce the gap between the necessary and sufficient conditions for optimality, we should prove that the conditions

$$F''(\bar{u})v^2 + \alpha \int_\omega \Big( |h_v(x)|_2^2 - (\bar{h}(x) \cdot h_v(x))^2 \Big) d|\nabla \bar{u}| \ge \delta \|z_v\|_{L^2(\Omega)}^2 \quad \forall v \in C_{\bar{u}}^\tau$$

and

$$F''(\bar{u})v^2 + \alpha \int_\omega \Big( |h_v(x)|_2^2 - (\bar{h}(x) \cdot h_v(x))^2 \Big) d|\nabla \bar{u}| \ge \delta \|v\|_{L^2(\omega)}^2 \quad \forall v \in C_{\bar{u}}^\tau$$

imply (5.28) and (5.30), respectively. This, however, remains as a challenge.

## 6 A regularization of problem (P)

Here we briefly discuss the effect of introducing an $H^1(\omega)$-regularization term. Throughout this section we assume that assumptions 1. and 2. of Theorem 1 hold. We discuss the optimality conditions of the regularized problems and investigate their asymptotic behavior as $\varepsilon \to \infty$. We also compare the optimality conditions which arise from the asymptotics of the regularized optimality conditions with those obtained in Section 4. For $\varepsilon > 0$ we thus consider

$$(\mathrm{P}_\varepsilon) \quad \min_{u \in H^1(\omega)} J_\varepsilon(u) = J(u) + \frac{\varepsilon}{2} \int_\omega |\nabla u(x)|^2 \, dx,$$

subject to (1.1), and denote a solution by $u_\varepsilon$. Let us set

$$J_\varepsilon(u) = F_\varepsilon(u) + G(u),$$

where $F_\varepsilon(u) = F(u) + \frac{\varepsilon}{2} \int_\omega |\nabla u|^2 \, dx$ for $u \in H^1(\omega)$. We have

$$F_\varepsilon'(u)v = F'(u)v + \varepsilon \int_\omega \nabla u \cdot \nabla v \, dx, \quad \text{and} \quad \partial G(u) = \nabla^* \partial g(\nabla u) \text{ for } u \in H^1(\omega),$$

where now $G(\cdot) = (g \circ \nabla)(\cdot)$, $\nabla : H^1(\omega) \to L^2(\omega)^n$, and $g : L^2(\omega)^n \to \mathbb{R}$ is given by $g(v) = \|v\|_{L^1(\omega)^n}$. We have the analog of Theorem 2, i.e. for every local solution $u_\varepsilon$ of $(\mathrm{P}_\varepsilon)$ there exists $\lambda_\varepsilon \in \partial g(\nabla u_\varepsilon)$ such that

$$\alpha(\lambda_\varepsilon, \nabla v)_{L^2(\omega)^n} + F_\varepsilon'(u_\varepsilon)v = 0, \text{ for all } v \in H^1(\omega). \tag{6.1}$$

Let us focus on $\lambda_\varepsilon \in \partial g(\nabla u_\varepsilon)$ next. It is equivalent to

$$(\lambda_\varepsilon, \nabla u_\varepsilon) = \|\nabla u_\varepsilon\|_{L^1(\omega)^n}, \text{ and } (\lambda_\varepsilon, v) \le \|v\|_{L^1(\omega)^n} \text{ for all } v \in L^1(\omega)^n. \tag{6.2}$$

**The use of the Euclidean norm** $|\cdot|_2$: Here (6.2) results in

$$\sum_{i=1}^n (\lambda_{\varepsilon,i}, \partial_{x_i} u_\varepsilon) = \int_\omega \left( \sum_{i=1}^n |\partial_{x_i} u_\varepsilon|^2 \right)^{\frac{1}{2}} dx, \text{ and } \sum_{i=1}^n (\lambda_{\varepsilon,i}, v_i) \le \int_\omega \left( \sum_{i=1}^n |v_i|^2 \right)^{\frac{1}{2}} dx,$$
$$\tag{6.3}$$

for all $v \in L^1(\omega)^n$. The second expression in (6.3) implies that $\|\lambda_\varepsilon\|_{L^\infty(\omega, \mathbb{R}^n)} \le 1$. Moreover, if $\nabla u_\varepsilon \ne 0$,

$$\|\lambda_\varepsilon\|_{L^\infty(\omega, \mathbb{R}^n)} = 1 \text{ and } \operatorname{supp} \nabla u_\varepsilon \subset \{x \in \omega : |\lambda_\varepsilon(x)|_2 = 1\}. \tag{6.4}$$

The first claim follows from the equality in (6.3). This equality can also be expressed as $\int_\omega |\nabla u_\varepsilon|_2 \, dx = \int_\omega (\nabla u_\varepsilon \cdot \lambda_\varepsilon) \, dx$, which, together with $|\lambda_\varepsilon(x)|_2 \le 1$ implies the second assertion in (6.4).

**The use of the** $|\cdot|_\infty$**-norm**: In this case (6.2) results in

$$\sum_{i=1}^n (\lambda_{\varepsilon,i}, \partial_{x_i} u_\varepsilon) = \sum_{i=1}^n \|\partial_{x_i} u_\varepsilon\|_{L^1(\omega)} \text{ and } \sum_{i=1}^n (\lambda_{\varepsilon,i}, v_i) \le \sum_{i=1}^n \|v_i\|_{L^1(\omega)}, \tag{6.5}$$

for all $v \in L^1(\omega)^n$. This implies that $\|\lambda_{\varepsilon,j}\|_{L^\infty(\omega)} \le 1$ for all $j = 1, \ldots, n$ and if $\partial_{x_j} u_\varepsilon \ne 0$

$$\|\lambda_{\varepsilon,j}\|_{L^\infty(\omega)} = 1, \text{ and } \operatorname{supp}(\partial_{x_j} u_\varepsilon)^\pm \subset \{x \in \omega : \lambda_{\varepsilon,j} = \pm 1\}. \tag{6.6}$$

In fact, for any $1 \leq j \leq n$, let $\nu_i = 0$ for all $i \neq j$ and $\nu_j = \lambda_{\varepsilon,j}$ on $S_j^+ = \{x : \lambda_{\varepsilon,j} > 1\}$, and equal to 0 otherwise. Then $\int_{S_j^+} (\lambda_{\varepsilon,j}^2 - \lambda_{\varepsilon,j})(x)\, dx \leq 0$, while the integrand is strictly positive a.e. Hence $meas(S_j^+) = 0$. In an analogous form we exclude the case $\lambda_{\varepsilon,j} < -1$, and hence $\|\lambda_{\varepsilon,j}\|_{L^\infty(\omega)} \leq 1$, for all $j$. Using the first expression in (6.5) we have

$$\sum_{i=1}^n \|\partial_{x_i} u_\varepsilon\|_{L^1(\omega)} = \sum_{i=1}^n (\lambda_{\varepsilon,i}, \partial_{x_i} u_\varepsilon) \leq \sum_{i=1}^n \|\partial_{x_i} u_\varepsilon\|_{L^1(\omega)},$$

which implies (6.6).

**Asymptotic behavior:** Finally we consider the asymptotic behavior of (6.1), (6.2) as $\varepsilon \to 0^+$. From the inequality $J_\varepsilon(u_\varepsilon) \leq J(0)$ for all $\varepsilon > 0$, we deduce with (1.2) the boundedness of $\{u_\varepsilon\}_\varepsilon$ in $BV(\omega) \cap L^2(\omega)$. Moreover, (6.4) and (6.6) imply the boundedness of $\{\lambda_\varepsilon\}_\varepsilon$ in $L^\infty(\omega)^n$. Hence there exists $(\bar{u}, \hat{\lambda}) \in (BV(\omega) \cap L^2(\omega)) \times L^\infty(\omega)^n$ such that on a subsequence $(u_\varepsilon, \lambda_\varepsilon) \xrightarrow{*} (\bar{u}, \hat{\lambda})$ weakly* in $(BV(\omega) \cap L^2(\omega)) \times L^\infty(\omega)$. In particular $\lambda_\varepsilon \xrightarrow{*} \hat{\lambda}$ in $L^\infty(\omega)$. Moreover $y_{u_\varepsilon} \to y_{\bar{u}}$ in $L^2(\Omega)$.

Now, given an arbitrary element $u \in H^1(\omega)$, the optimality of $u_\varepsilon$ and the structure of $J$ implies

$$J(\bar{u}) \leq \liminf_{\varepsilon \to 0} J(u_\varepsilon) \leq \limsup_{\varepsilon \to 0} J(u_\varepsilon) \leq \limsup_{\varepsilon \to 0} J_\varepsilon(u_\varepsilon) \leq \limsup_{\varepsilon \to 0} J_\varepsilon(u) = J(u).$$

Since $H^1(\Omega)$ is dense in $BV(\omega) \cap L^2(\omega)$, the above inequality implies that $\bar{u}$ is a solution of (P) and

$$J(\bar{u}) = \lim_{\varepsilon \to 0} J(u_\varepsilon) = \limsup_{\varepsilon \to 0} J_\varepsilon(u_\varepsilon) = \inf(\mathrm{P}) = J(\bar{u}). \tag{6.7}$$

This implies that $J(u_\varepsilon) \to J(\bar{u})$ and $\frac{\varepsilon}{2} \int_\omega |\nabla u_\varepsilon|^2\, dx \to 0$. Moreover, from the convergence properties of $\{u_\varepsilon\}_\varepsilon$ and $\{y_\varepsilon\}_\varepsilon$ we deduce that

$$\lim_{\varepsilon \to 0} \left[ \frac{1}{2} \|y_{u_\varepsilon} - y_d\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \left( \int_\omega u_\varepsilon\, dx \right)^2 \right] = \frac{1}{2} \|y_{\bar{u}} - y_d\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \left( \int_\omega \bar{u}\, dx \right)^2, \tag{6.8}$$

$$\int_\omega |\nabla \bar{u}| \leq \liminf_{\varepsilon \to 0} \int_\omega |\nabla \bar{u}_\varepsilon|. \tag{6.9}$$

Combining (6.8) with the convergence $J(u_\varepsilon) \to J(\bar{u})$ we infer

$$\lim_{\varepsilon \to 0} \left( \frac{\gamma}{2} \|u_\varepsilon\|_{L^2(\omega)}^2 + \alpha \int_\omega |\nabla u_\varepsilon| \right) = \frac{\gamma}{2} \|\bar{u}\|_{L^2(\omega)}^2 + \alpha \int_\omega |\nabla \bar{u}|. \tag{6.10}$$

If $\gamma = 0$ then this identity is reduced to $\int_\omega |\nabla u_\varepsilon| \to \int_\omega |\nabla \bar{u}|$. Let us prove that this convergence property also holds for $\gamma > 0$. Using (6.10), the convergence $u_\varepsilon \rightharpoonup \bar{u}$ in $L^2(\omega)$, and (6.9) we obtain

$$\frac{\gamma}{2} \|\bar{u}\|_{L^2(\omega)}^2 \leq \liminf_{\varepsilon \to 0} \frac{\gamma}{2} \|u_\varepsilon\|_{L^2(\omega)}^2 \leq \limsup_{\varepsilon \to 0} \frac{\gamma}{2} \|u_\varepsilon\|_{L^2(\omega)}^2$$

$$\leq \limsup_{\varepsilon \to 0} \left( \frac{\gamma}{2} \|u_\varepsilon\|_{L^2(\omega)}^2 + \alpha \int_\omega |\nabla u_\varepsilon| \right) - \alpha \liminf_{\varepsilon \to 0} \int_\omega |\nabla \bar{u}_\varepsilon|$$

$$\leq \left( \frac{\gamma}{2} \|\bar{u}\|_{L^2(\omega)}^2 + \alpha \int_\omega |\nabla \bar{u}| \right) - \alpha \int_\omega |\nabla \bar{u}| = \frac{\gamma}{2} \|\bar{u}\|_{L^2(\omega)}^2.$$

Therefore, $\|u_\varepsilon\|_{L^2(\omega)} \to \|\bar{u}\|_{L^2(\omega)}$ holds. Combining this fact with the weak convergence we conclude that $u_\varepsilon \to \bar{u}$ strongly in $L^2(\omega)$. Inserting this in (6.10) it follows that $\int_\omega |\nabla u_\varepsilon| \to \int_\omega |\nabla \bar{u}|$.

From (6.1) we have that

$$\alpha(\lambda_\varepsilon, \nabla v) + \int_\omega \Big(\varphi(u_\varepsilon) + \gamma u_\varepsilon + \beta \int_\omega u_\varepsilon \, dz\Big) v \, dx - \varepsilon \int_\omega u_\varepsilon \Delta v \, dx = 0, \quad \forall v \in C_0^\infty(\omega).$$

Taking the limit $\varepsilon \to 0$ we obtain

$$\alpha(\hat{\lambda}, \nabla v) + \int_\omega \Big(\varphi(\bar{u}) + \gamma \bar{u} + \beta \int_\omega \bar{u} \, dz\Big) v \, dx = 0, \quad \forall v \in C_0^\infty(\omega),$$

which corresponds to (4.1). Moreover, since $C_0^\infty(\omega)$ is dense in $L^2(\omega)$ the above relation implies that $\hat{\lambda} \in H^{\mathrm{div}}(\omega)$, and

$$-\alpha \, \mathrm{div}\hat{\lambda} + \varphi(\bar{u}) + \gamma \bar{u} + \beta \int_\omega \bar{u} \, dz = 0 \quad \text{in } L^2(\omega). \tag{6.11}$$

From (6.1), the above identity, and the established convergence $\varepsilon \int_\omega |\nabla u_\varepsilon|^2 \, dx \to 0$ we find

$$\lim_{\varepsilon \to 0}(\lambda_\varepsilon, \nabla u_\varepsilon) = -\lim_{\varepsilon \to 0} \frac{1}{\alpha} F_\varepsilon'(u_\varepsilon) u_\varepsilon = -\lim_{\varepsilon \to 0} \frac{1}{\alpha}\Big(F'(u_\varepsilon)u_\varepsilon - \varepsilon \int_\omega |\nabla u_\varepsilon|^2 \, dx\Big)$$
$$= -\frac{1}{\alpha} F'(\bar{u})\bar{u} = -(\mathrm{div}\,\hat{\lambda}, \bar{u}).$$

Now, from (6.2) and the convergence $\int_\omega |\nabla u_\varepsilon| \to \int_\omega |\nabla \bar{u}|$ we infer

$$\lim_{\varepsilon \to 0}(\lambda_\varepsilon, \nabla u_\varepsilon) = \|\nabla \bar{u}\|_{\mathcal{M}(\omega)^n}.$$

From the last two identities, and using again (6.2) along with the convergence $\lambda_\varepsilon \overset{*}{\rightharpoonup} \hat{\lambda}$ in $L^\infty(\omega)$ we obtain

$$(-\mathrm{div}\,\hat{\lambda}, \bar{u})_{L^2(\omega)} = \|\nabla \bar{u}\|_{\mathcal{M}(\omega)^n}, \text{ and } \langle \hat{\lambda}, v\rangle_{L^\infty, L^1} \le |v|_{L^1(\omega)^n} \text{ for all } v \in L^1(\omega)^n. \tag{6.12}$$

We now summarize the important issues of the above developments as a theorem.

**Theorem 10** *Let assumptions 1. and 2. of Theorem 1 hold, and let $\{u_\varepsilon\}_{\varepsilon>0}$ denote a family of solutions to $(\mathrm{P}_\varepsilon)$. Then for each $\varepsilon$ there exists $\lambda_\varepsilon \in L^\infty(\omega)$ such that*

$$\alpha(\lambda_\varepsilon, \nabla v)_{L^2(\omega)^n} + F_\varepsilon'(u_\varepsilon)v = 0, \text{ for all } v \in H^1(\omega), \text{ and}$$
$$(\lambda_\varepsilon, \nabla u_\varepsilon) = \|\nabla u_\varepsilon\|_{L^1(\omega)^n}, \text{ and } (\lambda_\varepsilon, v) \le \|v\|_{L^1(\omega)^n} \text{ for all } v \in L^1(\omega)^n.$$

*For each weakly* $^*$ *convergent subsequence of $\{(u_\varepsilon, \lambda_\varepsilon)\}_\varepsilon \subset (BV(\omega) \cap L^2(\omega)) \times L^\infty(\omega)$ with limit $(\bar{u}, \hat{\lambda})$, of which there exists at least one, the following properties hold: $\hat{\lambda} \in H^{\mathrm{div}}(\omega)$, and $(\bar{u}, \hat{\lambda})$ satisfy the necessary optimality conditions (6.11), (6.12).*

*Remark 8* Here we compare $\bar{\lambda}$ from Section 4 to $\hat{\lambda}$ obtained by the regularisation approach. Since $L^1(\omega)^n \subset \mathcal{M}(\omega)^n$ we can conclude from $\bar{\lambda} \in \partial g(\nabla \bar{u})$ and (3.5) that $\bar{\lambda}$ restricted to $L^1(\omega)^n$ can be identified with an $L^\infty(\omega)^n$ function that we denote by $\bar{\lambda}|_{L^1}$. Restricting (4.1) to $v \in C_0^\infty(\omega)$ and using density of $C_0^\infty(\omega)$ in $L^2(\omega)$ we conclude that $\bar{\lambda}|_{L^1} \in H^{\mathrm{div}}(\omega)$, and we have

$$-\alpha \operatorname{div}\bar{\lambda}|_{L^1} + \varphi(\bar{u}) + \gamma \bar{u} + \beta \int_\omega \bar{u}\,dz = 0 \;\; \text{in } L^2(\omega).$$

Thus $\operatorname{div}\bar{\lambda}|_{L^1}$ from Section 4 coincides with $\operatorname{div}\hat{\lambda}$ obtained by regularization and it is uniquely defined by (6.11). We cannot assert, however, that $\hat{\lambda} = \bar{\lambda}|_{L^1}$.

But we can compare (6.12) and $\bar{\lambda} \in \partial g(\nabla \bar{u}) \subset [\mathcal{M}(\omega)^n]^*$ which was obtained in Theorem 2. By (3.4) and (3.5) the latter implies that $\langle \nabla^* \bar{\lambda}, \bar{u} \rangle_{BV(\omega)^*, BV(\omega)} = \|\nabla \bar{u}\|_{\mathcal{M}(\omega)^n}$ and $\langle \bar{\lambda}, \nu \rangle \leq \|\nu\|_{\mathcal{M}(\omega)^n}$ for all $\nu \in \mathcal{M}(\omega)^n$, which coincides with (6.12), when restricted to $\nu \in L^1(\omega)^n$. Summarizing, the regularisation approach recovers $\bar{\lambda}$ when restricted to $L^1(\omega)^n$.

## 7 Conclusions

An analysis for BV-regularised optimal control problems associated to semilinear elliptic equations was provided. Existence, first order necessary and second order sufficient optimality conditions were investigated. Special attention was given to the different cases which arise due to the choice of a particular vector norm in the definition of the BV-seminorm. If (P) is additionally regularised by an $H^1(\omega)$-seminorm, then the set where the gradient of the optimal solution vanishes, can be characterised conveniently by an adjoint variable, see (6.4) and (6.6). For the original problem (P) without $H^1(\omega)$-seminorm regularisation, such a transparent description of the set where the measure $|\nabla \bar{u}|$ vanishes is not available, rather it was replaced by the properties specified in Theorem 3.

## References

1. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, New York, 2000.
2. K. Bredies and M. Holler, *A pointwise characterization of the subdifferential of the total variation functional*, tech. report.
3. H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, New York, 2011.
4. E. Casas, *Second order analysis for bang-bang control problems of PDEs*, SIAM J. Control Optim., 50 (2012), pp. 2355–2372.
5. E. Casas, C. Clason, and K. Kunisch, *Approximation of elliptic control problems in measure spaces with sparse solutions*, SIAM Journal on Control and Optimization, 50 (2012), pp. 1735–1752.
6. E. Casas, C. Clason, and K. Kunisch, *Parabolic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim., 51 (2013), pp. 28–63.
7. E. Casas, R. Herzog, and G. Wachsmuth, *Optimality conditions and error analysis of semilinear elliptic control problems with $L^1$ cost functional*, SIAM J. Optim., 22 (2012), pp. 795–820.
8. E. Casas, P. Kogut, and G. Leugering, *Approximation of optimal control problems in the coefficient for the p-Laplace equation. I. Convergence result*, SIAM J. Control Optim., 54 (2016), pp. 1406–1422.

9. E. Casas, F. Kruse, and K. Kunisch, *Optimal control of semilinear parabolic equations by BV-functions*, SIAM J. Control Optim., 55 (2017), pp. 1752–1788.

10. E. Casas and K. Kunisch, *Optimal control of semilinear elliptic equations in measure spaces*, SIAM J. Control Optim., 52 (2014), pp. 339–364.

11. E. Casas, K. Kunisch, and C. Pola, *Regularization by functions of bounded variation and applications to image enhancement*, App. Math. Optim., 40 (1999), pp. 229–257.

12. E. Casas and F. Tröltzsch, *Second order optimality conditions and their role in pde control*, Jahresber Dtsch Math-Ver, 117 (2015), pp. 3–44.

13. ———, *Second order optimality conditions for weak and strong local solutions of parabolic optimal control problems*, Vietnam J. Math., 44 (2016), pp. 181–202.

14. A. Chambolle, V. Duval, G. Peyré, and C. Poon, *Geometric properties of solutions to the total variation denoising problem*, Inverse Problems, 33 (2017), pp. 015002, 44.

15. C. Clason and K. Kunisch, *A duality-based approach to elliptic control problems in non-reflexive Banach spaces*, ESAIM: Control, Optimisation and Calculus of Variations, 17 (2011), pp. 243–266.

16. I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland-Elsevier, New York, 1976.

17. D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin Heidelberg, 1983.

18. E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Boston, 1984.

19. D. Hafemayer, *Regularization and Discretization of a BV-controlled, Elliptic Problem: A Completely Adaptive Approach*, master's thesis, TU-Munich, Germany, 2017.

20. K. Jalalzai, *Some remarks on the staircasing phenomenon in total variation-based image denoising*, J. Math. Imaging Vision, 54 (2016), pp. 256–268.

21. W. Ring, *Structural properties of solutions to total variation regularization problems*, M2AN Math. Model. Numer. Anal., 34 (2000), pp. 799–810.

22. L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268. Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991).

23. W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book Co., London, 1970.

24. ———, *Functional Analysis*, McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973.

25. G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.

26. F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Philadelphia, 2010.