Zeitschrift für Analysis und ihre Anwendungen
Journal of Analysis and its Applications
Volume 34 (2015), 85–108
DOI: 10.4171/ZAA/1530

Optimal Control in Matrix-Valued Coefficients for Nonlinear Monotone Problems: Optimality Conditions I

Peter I. Kogut, Ol'ga P. Kupenko, and Günter Leugering

Abstract. In this article we study an optimal control problem for a nonlinear monotone Dirichlet problem where the controls are taken as matrix-valued coefficients in $L^\infty(\Omega; \mathbb{R}^{N\times N})$. For the exemplary case of a tracking cost functional, we derive first order optimality conditions. This first part is concerned with the general case of matrix-valued coefficients under some hypothesis, while the second part focuses on the special class of diagonal matrices.

Keywords. Nonlinear monotone Dirichlet problem, control in coefficients, adjoint equation

Mathematics Subject Classification (2010). 35J70, 49J20, 49J45, 93C73.

1. Introduction

The aim of this article is to derive a first order optimality system for a Dirichlet optimal control problem where the controls are taken as the matrix-valued coefficients in a nonlinear state equation. The controls are supposed to satisfy rather weak hypotheses. The optimal control problem amounts to minimizing the discrepancy between a given distribution $y_d \in L^p(\Omega)$, where Ω is an open bounded Lipschitz domain in $\mathbb{R}^N$, and the solution of a nonlinear Dirichlet problem by choosing an appropriate matrix of coefficients $\mathcal{U} \in L^\infty(D; \mathbb{R}^{N\times N})$. Namely, we consider the following minimization problem:

$$\text{Minimize} \quad \left\{ I_\Omega(\mathcal{U}, y) = \int_\Omega |y(x) - y_d(x)|^p \, dx \right\} \tag{1}$$

P. I. Kogut: Department of Differential Equations, Dnipropetrovsk National University, Gagarin av., 72, 49010 Dnipropetrovsk, Ukraine; p.kogut@i.ua

O. P. Kupenko: Institute for Applied System Analysis, National Academy of Sciences and Ministry of Education and Science of Ukraine, Peremogy av., 37/35, IPSA, 03056 Kyiv, Ukraine; kogut_olga@bk.ru

G. Leugering: Lehrstuhl für Angewandte Mathematik II, Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany; guenter.leugering@fau.de

subject to the constraints

$$\mathcal{U} \in U_{ad} \subset L^\infty(\Omega; \mathbb{R}^{N\times N}), \quad y \in W_0^{1,p}(\Omega), \tag{2}$$

$$-\text{div}\,\big(\mathcal{U}[(\nabla y)^{p-2}]\nabla y\big) + |y|^{p-2}y = f \quad \text{in } \Omega, \tag{3}$$

$$y = 0 \quad \text{on } \partial\Omega, \tag{4}$$

where U_{ad} is a class of admissible controls and

$$[(\nabla y)^{p-2}] = \text{diag}\left\{ \left|\frac{\partial y}{\partial x_1}\right|^{p-2}, \left|\frac{\partial y}{\partial x_2}\right|^{p-2}, \dots, \left|\frac{\partial y}{\partial x_N}\right|^{p-2} \right\}.$$

Clearly, the choice of the cost function in (1) is exemplary. A typical regularization of (1) would be of Tikhonov type. Optimal control for partial differential equations by the way of controlling the coefficients is a classical subject initiated by Lurie [19], Fleming [6] and Lions [16]. Zolezzi [33] picked up the theme and Tartar [28] showed examples of non-existence (see also [22,23]), which, in turn, initiated the theory of homogenization (see e.g. [29] also for the historical development). In particular, the notion of H-convergence was developed by Murat and Tartar ([24]) aiming at matrix-valued coefficients. Now, taking the coefficients of the leading differential operator as optimization variables amounts to a problem of material design, as those coefficients describe, via constitutive equations, the material behavior; e.g. conductivity in scalar equations or elasticity in vectorial problems. The possibility to optimize the material properties has triggered an enormous interest in material sciences in recent years. The subtle point is the choice of the topology in which minimizing sequences converge. Moreover, the limiting optimal coefficients have to be interpreted in the context of the application. Therefore, structural assumptions have to be considered during the optimization process in terms of constraints. One way of doing so is via proper parametrization of the material, respectively the coefficients, using mixtures, represented by characteristic functions. This has been pursued by Allaire [1] and many other authors in recent years. Other restriction can be realized via regularity of the coefficients and hard constraints. This procedure has been pursued first by Casas [2] for a scalar problem, as one of the first papers in that direction, and later by Haslinger et al. [10] in the context of what has come to be known as *Free Material Optimization* (FMO); see also [15], where slope-constraints are used for regularization. This direction of research is quite active in recent years. See e.g. the work of Dekelnick and Hinze [5] where tracking type optimization for scalar problems with a Tikhonov-type regularization of the controls are considered in the context of inverse problems. However, most of the results and methods rely on linear PDEs, while only very few articles deal with nonlinear problems, see O. Kogut [11] and P. Kogut and Leugering [12]. Another point of interest is degeneration in the coefficients which is typically avoided by assuming lower bounds on the coefficients. However, degeneration occurs genuinely in topology optimization, damage and crack problems. In Kogut and Leugering [13, 14] this problem has been considered in the context of linear problems. In this article, we extend our results to scalar nonlinear problems, where degeneration occurs already with respect to the states. We will continue the discussion for non-scalar ones in a forthcoming paper.

We restrict the set of admissible controls to the problem (1)–(4) by introducing so-called solenoidal matrices U_{ad} that are a uniformly bounded in $L^\infty(\Omega;\mathbb{R}^{N\times N})$. However, in contrast to the typical assumptions (see, for instance, [3,9,17,18,26,30]), U_{ad} belongs neither to the Sobolev space $W^{1,\infty}(\Omega;\mathbb{R}^{N\times N})$ nor to the space of matrices with bounded variation $BV(\Omega;\mathbb{R}^{N\times N})$. Thus, in some sense we try to avoid a situation of over-regularization for optimal solutions to the problem (1)–(4). We give the precise definition of such controls in Section 3 and show that in this case the original optimal control problem admits at least one solution.

In Section 4 we discuss the differentiability properties of the Lagrange functional associated to problem (1)–(4)

$$\Lambda(\mathcal{U}, y, \lambda) = I(\mathcal{U}, y) + a_{\mathcal{U}}(y, \lambda) - \langle f, \lambda\rangle_{W_0^{1,p}(\Omega)},$$

where

$$a_{\mathcal{U}}(y, \lambda) = \int_\Omega (\mathcal{U}(x)[(\nabla y)^{p-2}]\,\nabla y, \nabla\lambda)_{\mathbb{R}^N}\,dx + \int_\Omega |y|^{p-2} y\lambda\,dx$$

and show that it admits a one-sided directional derivative with respect to the variable $y \in \mathcal{D}_y^+\Lambda(\mathcal{U}, y, \lambda, h)$ for so-called non-degenerate directions $h \in W_0^{1,p}(\Omega)$ at the point y. Moreover, this derivative can be recovered in the form of the Gâteaux differential $\langle \mathcal{D}_y\,\Lambda(\mathcal{U}, y, \lambda), h\rangle_{W_0^{1,p}(\Omega)}$ if the given point y possesses some extra regularity properties.

In Section 5 we derive first-order optimality conditions for optimal control problem (1)–(4) and carry out their realization under additional assumptions. With that in mind, we introduce the notion of a quasi-adjoint state ψ_ε to an optimal solution $y_0 \in W_0^{1,p}(\Omega)$ that was proposed for linear problems by Serovajskiy [27]) and show that an optimality system for the original problem can be recovered in an explicit form if the mapping $U_{ad} \ni \mathcal{U} \mapsto \psi_\varepsilon(\mathcal{U})$ possesses the so-called $\mathcal{H}$-property with respect to the pair of spaces $\big(L^\infty(\Omega;\mathbb{R}^{N\times N}), W_0^{1,p}(\Omega)\big)$. However, it should be stressed that the fulfilment of this property is not proved for the case $p > 2$ and, thus, should be considered as some extra hypothesis. Moreover, the verification of the $\mathcal{H}$-property for quasi-adjoint states is not straightforward, in general. In order to relax the hypothesis, we focus on diagonal matrices in the second part of this article, published separately.

2. Notation and preliminaries

Throughout the paper Ω is a bounded open subset of $\mathbb{R}^N$, $N \geq 1$. The space $\mathcal{D}'(\Omega)$ of distributions in Ω is the dual of the space $C_0^\infty(\Omega)$. For real numbers $2 \leq p < +\infty$, and $1 < q < +\infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$, the space $W_0^{1,p}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in the Sobolev space $W^{1,p}(\Omega)$, while $W^{-1,q}(\Omega)$ is the space of distributions of the form $f = f_0 + \sum_j D_j f_j$, with $f_0, f_1, \dots, f_n \in L^q(\Omega)$ (i.e. $W^{-1,q}(\Omega)$ is the dual space of $W_0^{1,p}(\Omega)$). Let χ_E be the characteristic function of a set $E \subset \mathbb{R}^N$ and let $\mathcal{L}^N(E)$ be its the N-dimensional Lebesgue measure.

For any vector field $v \in L^q(\Omega; \mathbb{R}^N)$, the divergence is an element of the space $W^{-1,q}(\Omega)$ defined by the formula

$$\langle \operatorname{div} v, \varphi \rangle_{W_0^{1,p}(\Omega)} = -\int_\Omega (v, \nabla \varphi)_{\mathbb{R}^N}\, dx, \quad \forall\, \varphi \in W_0^{1,p}(\Omega), \tag{5}$$

where $\langle \cdot, \cdot \rangle_{W_0^{1,p}(\Omega)}$ denotes the duality pairing between $W^{-1,q}(\Omega)$ and $W_0^{1,p}(\Omega)$, and $(\cdot, \cdot)_{\mathbb{R}^N}$ denotes the scalar product of two vectors in $\mathbb{R}^N$. A vector field $\mathbf{v}$ is said to be solenoidal, if $\operatorname{div} \mathbf{v} = 0$.

Weak Compactness Criterion in $L^1(\Omega; \mathbb{R}^N)$. Let $\{f_k\}_{k\in\mathbb{N}}$ be a bounded sequence of vector-valued functions in $L^1(\Omega; \mathbb{R}^N)$. We recall that $\{f_k\}_{k\in\mathbb{N}}$ is called equi-integrable on Ω if for any $\delta > 0$, there is a $\tau = \tau(\delta)$ such that $\int_S \|f_k\|_{\mathbb{R}^N}\, dx < \delta$ for every measurable subset $S \subset \Omega$ of Lebesgue measure $\mathcal{L}^N(S) < \tau$. Then, the following assertions are equivalent for $L^1(\Omega; \mathbb{R}^N)$-bounded sequences:

(i) the sequence $\{f_k\}_{k\in\mathbb{N}}$ is weakly compact in $L^1(\Omega; \mathbb{R}^N)$,

(ii) the sequence $\{f_k\}_{k\in\mathbb{N}}$ is equi-integrable.

Lemma 2.1 (Lebesgue's Theorem)**.** *If a sequence* $\{f_k\}_{k\in\mathbb{N}} \subset L^1(\Omega; \mathbb{R}^N)$ *is equi-integrable and* $f_k \to f$ *almost everywhere in* Ω *then* $f_k \to f$ *in* $L^1(\Omega; \mathbb{R}^N)$.

Monotone operators. Let α and β be constants such that $0 < \alpha \leq \beta < +\infty$. We define $M_p^{\alpha,\beta}(\Omega)$ as a set of all square symmetric matrices $\mathcal{U}(x) = [a_{ij}(x)]_{1 \leq i,j \leq N}$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$ such that the following conditions of growth, monotonicity, and strong coercivity are fulfilled:

$$|a_{ij}(x)| \leq \beta \quad \text{a.e. in } \Omega,\ \forall\, i, j \in \{1, \dots, N\}, \tag{6}$$

$$\big(\mathcal{U}(x)([\zeta^{p-2}]\zeta - [\eta^{p-2}]\eta), \zeta - \eta\big)_{\mathbb{R}^N} \geq 0 \quad \text{a.e. in } \Omega,\ \forall\, \zeta, \eta \in \mathbb{R}^N, \tag{7}$$

$$\big(\mathcal{U}(x)[\zeta^{p-2}]\zeta, \zeta\big)_{\mathbb{R}^N} = \sum_{i,j=1}^N a_{ij}(x)|\zeta_j|^{p-2}\, \zeta_j\, \zeta_i \geq \alpha\, |\zeta|_p^p \quad \text{a.e in } \Omega, \tag{8}$$

where $|\eta|_p = \left(\sum_{k=1}^N |\eta_k|^p\right)^{\frac{1}{p}}$ is the Hölder norm of $\eta \in \mathbb{R}^N$ and

$$[\eta^{p-2}] = \operatorname{diag}\{|\eta_1|^{p-2}, |\eta_2|^{p-2}, \dots, |\eta_N|^{p-2}\} \quad \forall \eta \in \mathbb{R}^N. \tag{9}$$

Remark 2.2. It is easy to see that $M_p^{\alpha,\beta}(\Omega)$ is a nonempty subset of $L^\infty(\Omega;\mathbb{R}^{N\times N})$. Particular representatives are diagonal matrices of the form

$$\mathcal{U}(x) = \operatorname{diag}\{\delta_1(x), \delta_2(x), \dots, \delta_N(x)\},$$

where $\alpha \le \delta_i(x) \le \beta$ a.e. in Ω, $\forall\, i \in \{1,\dots,N\}$ (see [4]).

Let us consider the nonlinear operator $A : M_p^{\alpha,\beta}(\Omega) \times W_0^{1,p}(\Omega) \to W^{-1,q}(\Omega)$ defined as

$$A(\mathcal{U}, y) = -\operatorname{div}\big(\mathcal{U}(x)[(\nabla y)^{p-2}]\nabla y\big) + |y|^{p-2}y,$$

or via the paring

$$\langle A(\mathcal{U},y), v\rangle_{W_0^{1,p}(\Omega)} = \sum_{i,j=1}^{N} \int_\Omega \left(a_{ij}(x) \left|\frac{\partial y}{\partial x_j}\right|^{p-2} \frac{\partial y}{\partial x_j}\right) \frac{\partial v}{\partial x_i}\, dx + \int_\Omega |y|^{p-2} y\, v\, dx,$$

for all $v \in W_0^{1,p}(\Omega)$. In view of properties (6)–(8), for every fixed matrix $\mathcal{U} \in M_p^{\alpha,\beta}(\Omega)$, the operator $A(\mathcal{U},\cdot)$ turns out to be coercive, strongly monotone and demi-continuous in the following sense: $y_k \to y_0$ strongly in $W_0^{1,p}(\Omega)$ implies that $A(\mathcal{U}, y_k) \rightharpoonup A(\mathcal{U}, y_0)$ weakly in $W^{-1,q}(\Omega)$ (see [8, p. 79, Definition 3.1.1; p. 84, Lemma 3.1.3], [16, p. 173, Theorem 2.1.2]). Then by well-known existence results for nonlinear elliptic equations with strictly monotone semi-continuous coercive operators (see [8, p. 95, Theorem 3.2.1], [31]), one can easily see that for every $f \in W^{-1,q}(\Omega)$ the nonlinear Dirichlet boundary value problem

$$A(\mathcal{U}, y) = f \quad \text{in } \Omega, \quad y \in W_0^{1,p}(\Omega), \tag{10}$$

admits a unique weak solution in $W_0^{1,p}(\Omega)$ for every fixed matrix $\mathcal{U} \in M_p^{\alpha,\beta}(\Omega)$. Let us recall that a function y is the weak solution of (10) if

$$y \in W_0^{1,p}(\Omega),$$
$$\int_\Omega \big(\mathcal{U}(x)[(\nabla y)^{p-2}]\nabla y, \nabla v\big)_{\mathbb{R}^N}\, dx + \int_\Omega |y|^{p-2} y v\, dx = \int_\Omega f v\, dx \quad \forall\, v \in W_0^{1,p}(\Omega).$$

3. Setting of the optimal control problem

Let ξ_1, ξ_2 be given functions of $L^\infty(\Omega)$ such that $0 \le \xi_1(x) \le \xi_2(x)$ a.e. in Ω. Let $\{Q_1, \dots, Q_N\}$ be a collection of nonempty compact convex subsets of $W^{-1,q}(\Omega)$. To define the class of admissible controls, we introduce two sets

$$U_b = \big\{ \mathcal{U} = [a_{ij}] \in M_p^{\alpha,\beta}(\Omega) \big| \; \xi_1(x) \le a_{ij}(x) \le \xi_2(x) \text{ a.e. in } \Omega,\ \forall\, i,j = 1,\dots,N \big\}, \tag{11}$$

$$U_{sol} = \big\{ \mathcal{U} = [u_1,\dots,u_N] \in M_p^{\alpha,\beta}(\Omega) \big| \operatorname{div} u_i \in Q_i,\ \forall\, i = 1,\dots,N \big\}, \tag{12}$$

assuming that the intersection $U_b \cap U_{sol} \subset L^\infty(\Omega;\mathbb{R}^{N\times N})$ is nonempty.

Definition 3.1. We say that a matrix $\mathcal{U} = [a_{i\,j}]$ is an admissible control of solenoidal type to the nonlinear Dirichlet problem (10) if $\mathcal{U} \in U_{ad} := U_b \cap U_{sol}$.

Let us consider the optimal control problem

$$\text{Minimize} \quad \Big\{ I(\mathcal{U}, y) = \int_\Omega |y(x) - y_d(x)|^p \, dx \Big\}, \tag{13}$$

subject to the constraints

$$\int_\Omega \big(\mathcal{U}(x)[(\nabla y)^{p-2}]\nabla y, \nabla v\big)_{\mathbb{R}^N} \; dx + \int_\Omega |y|^{p-2} y v \, dx = \langle f, v \rangle_{W_0^{1,p}(\Omega)} \qquad \forall \, v \in W_0^{1,p}(\Omega), \tag{14}$$

$$\mathcal{U} \in U_{ad}, \quad y \in W_0^{1,p}(\Omega), \tag{15}$$

where $f \in W^{-1,q}(\Omega)$ and $y_d \in W_0^{1,p}(\Omega)$ are given distributions.

Hereinafter, $\Xi_{sol} \subset L^\infty(\Omega; \mathbb{R}^{N\times N}) \times W_0^{1,p}(\Omega)$ denotes the set of all admissible pairs to optimal control problem (13)–(15). Let τ be the topology on the set $L^\infty(\Omega; \mathbb{R}^{N\times N}) \times W_0^{1,p}(\Omega)$ which we define as a product of the weak-$*$ topology of $L^\infty(\Omega; \mathbb{R}^{N\times N})$ and the weak topology of $W_0^{1,p}(\Omega)$. Further we make use of the following results, which play a key role for the solvability of the problem (see e.g. [4, p. 698, Proposition 2.8], [12, p. 597, Theorem 16.14]).

Proposition 3.2. *For each $\mathcal{U} \in M_p^{\alpha,\beta}(\Omega)$ and every $f \in W^{-1,q}(\Omega)$, a weak solution y to variational problem* (14), (15) *satisfies the estimate*

$$\|y\|^p_{W_0^{1,p}(\Omega)} := \int_\Omega |\nabla y|^p_{\mathbb{R}^N} \, dx \le C \|f\|^q_{W^{-1,\,q}(\Omega)}, \tag{16}$$

where C is a constant depending only on p and α.

Theorem 3.3. *For every $f \in W^{-1,q}(\Omega)$ the set Ξ_{sol} is sequentially τ-closed, i.e. if $\{(\mathcal{U}_k, y_k) \in \Xi_{sol}\}_{k\in\mathbb{N}}$ is such that $\mathcal{U}_k \to \mathcal{U}_0$ weakly-$*$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$ and $y_k = y(\mathcal{U}_k) \to y_0$ weakly in $W_0^{1,p}(\Omega)$ then $(\mathcal{U}_0, y_0) \in \Xi_{sol}$, and, hence, $y_0 = y(\mathcal{U}_0)$.*

As was shown in [4, p. 705, Theorem 3.6] (see also [12, p. 600, Theorem 16.15]), we have the following existence result.

Theorem 3.4. *If $U_{ad} = U_b \cap U_{sol} \neq \emptyset$, then the optimal control problem* (13)–(15) *admits at least one solution*

$$(\mathcal{U}^{opt}, y^{opt}) \in \Xi_{sol} \subset L^\infty(\Omega; \mathbb{R}^{N\times N}) \times W_0^{1,p}(\Omega),$$
$$I(\mathcal{U}^{opt}, y^{opt}) = \inf_{(\mathcal{U},y)\in\Xi_{sol}} I(\mathcal{U}, y).$$

4. Some auxiliary results

The main goal of this paper is to derive the optimality conditions for optimal control problem (13)–(15). However, we deal with the case when we cannot apply the well-known classical approach (see, for instance, [7], [31]), since for a given distribution $f \in W^{-1,q}(\Omega)$ the mapping $\mathcal{U} \mapsto y(\mathcal{U})$ is not Fréchet differentiable on the class of solenoidal controls, in general. With that in mind, we consider the Lagrange functional associated to problem (13)–(15) and discuss its differentiable properties. We define this functional as follows

$$\Lambda(\mathcal{U}, y, \lambda) = I(\mathcal{U}, y) + a_{\mathcal{U}}(y, \lambda) - \langle f, \lambda \rangle_{W_0^{1,p}(\Omega)}, \quad \lambda \in W_0^{1,p}(\Omega), \tag{17}$$

where

$$a_{\mathcal{U}}(y, \lambda) = \int_\Omega (\mathcal{U}(x)[(\nabla y)^{p-2}]\, \nabla y, \nabla \lambda)_{\mathbb{R}^N}\, dx + \int_\Omega |y|^{p-2} y \lambda\, dx.$$

It is easy to see that the Lagrangian $\Lambda(\mathcal{U}, y, \lambda)$ is not Gâteaux differentiable, in general. Let us show, however, that for any $\mathcal{U} \in U_{ad}$ and $\lambda \in W_0^{1,p}(\Omega)$ this functional has a one-sided directional derivative with respect to the variable y [25]. Indeed, for given $h \in W_0^{1,p}(\Omega)$ and $\theta \in [0, 1]$, we introduce the following sets

$$\begin{gathered}
\Omega_0 = \{x \in \Omega\colon |y - y_d| > 0\}, \quad \Omega_{0,\theta} = \{x \in \Omega\colon |y - y_d + \theta h| > 0\}, \\
\Omega_i = \left\{x \in \Omega\colon \left|\frac{\partial y}{\partial x_i}\right| > 0\right\}, \ \Omega_{i,\theta} = \left\{x \in \Omega\colon \left|\frac{\partial y}{\partial x_i} + \theta \frac{\partial h}{\partial x_i}\right| > 0\right\}, \quad \forall\, i = 1, \dots, N, \\
\Omega_{N+1} = \{x \in \Omega\colon |y| > 0\}, \quad \Omega_{N+1,\theta} = \{x \in \Omega\colon |y + \theta h| > 0\}.
\end{gathered}$$

Clearly, we cannot claim that $\chi_{\Omega_{0,\theta}} \to \chi_{\Omega_0}$ in $L^r(\Omega)$ for some $1 \le r < \infty$, because convergence of the sequence $\{y - y_d + \theta h\}_{\theta \to +0}$ to $y - y_d$ does not imply, in general, the χ−convergence of subsets $\{\Omega_{0,\theta}\}_{\theta \to +0}$ to Ω_0 as $\theta \to +0$ [12, p. 218]. Indeed, let $y_d, h \in W_0^{1,p}(\Omega)$ be such that $|h(x)| > 0$ almost everywhere in Ω and $y = y_d$. Then $\Omega_0 = \emptyset$ whereas $\Omega_{0,\theta} = \Omega$ for all positive θ small enough. Hence, in this case the convergence $\chi_{\Omega_{0,\theta}} \to \chi_{\Omega_0}$ fails. The same remark is valid for the sequences $\{\Omega_{i,\theta}\}_{\theta \to +0}$. In view of this, it is reasonable to introduce use of the following notion.

Definition 4.1. We say that an element $h \in W_0^{1,p}(\Omega)$ is a non-degenerate direction at the point $y \in W_0^{1,p}(\Omega)$ if

$$\chi_{\Omega_{i,\theta}} \to \chi_{\Omega_i} \text{ in } L^1(\Omega) \text{ as } \theta \to +0, \quad \forall\, i = 0, 1, \dots, N+1. \tag{18}$$

Remark 4.2. It is easy to see that an element $h \in W_0^{1,p}(\Omega)$ is a non-degenerate direction at the point $y \in W_0^{1,p}(\Omega)$ provided that $\Omega_i = \Omega_{i,\theta}$ for all $i = 0, 1, \ldots, N+1$. In particular, this is so if

$$\{x \in \Omega : \ |h(x)| > 0\} \subseteq \Omega_0 \cap \Omega_{N+1} \quad \text{and}$$
$$\left\{x \in \Omega : \ \left|\frac{\partial h(x)}{\partial x_i}\right| > 0\right\} \subseteq \Omega_i, \quad \forall\, i = 1, \ldots, N.$$

In order to make the condition (18) more transparent, we concentrate on the case $i = 0$.

Proposition 4.3. *Let y, y_d, and h be given elements of $W_0^{1,p}(\Omega)$. If the closed set*

$$S = \mathrm{cl}\, \{x \in \Omega : \ |y(x) - y_d(x)| = 0\} \tag{19}$$

has zero Lebesgue measure, then $\chi_{\Omega_{0,\theta}} \to \chi_{\Omega_0}$ a.e. in Ω, and, hence, $\chi_{\Omega_{0,\theta}} \to \chi_{\Omega_0}$ strongly in $L^r(\Omega)$ for all $1 \le r < +\infty$.

Proof. If $x \in \Omega_0$, then, by definition, $|y(x) - y_d(x)| > 0$. Thus, there is a value $\theta_0 \in (0, 1]$ such that $|y(x) - y_d(x) + \theta h| > 0$ for all $\theta \in [0, \theta_0]$ (here we use the fact that each element of $W_0^{1,p}(\Omega)$ can be interpreted as a quasi-continuous function [32]). Hence, $\chi_{\Omega_{0,\theta}}(x) = \chi_{\Omega_0}(x) = 1$ for all $\theta \in [0, \theta_0]$. Since the set S has zero Lebesgue measure, we get $\chi_{\Omega_{0,\theta}} \to \chi_{\Omega_0}$ almost everywhere in Ω. To conclude the proof, it remains to note that $\|\chi_{\Omega_{0,\theta}} - \chi_{\Omega_0}\|^r_{L^r(\Omega)} = \int_\Omega |\chi_{\Omega_{0,\theta}}(x) - \chi_{\Omega_0}(x)|^r\, dx = \int_\Omega |\chi_{\Omega_{0,\theta}}(x) - \chi_{\Omega_0}(x)|\, dx = \|\chi_{\Omega_{0,\theta}} - \chi_{\Omega_0}\|_{L^1(\Omega)}$. □

Corollary 4.4. *If the elements $y, y_d \in W_0^{1,p}(\Omega)$ are such that the set S, given by* (19)*, has zero Lebesgue measure, then property* (18) *with $i = 0$ are fulfilled for any $h \in W_0^{1,p}(\Omega)$.*

Remark 4.5. It is easy to see that the conclusions similar to Proposition 4.3 and Corollary 4.4, can be stated for the sets Ω_i and $\Omega_{i,\theta}$ with $i = 1, 2, \ldots, N+1$. Indeed, in spite of the fact that the functions $\frac{\partial y(x)}{\partial x_i}$, $\frac{\partial h(x)}{\partial x_i}$ $(i = 1, \ldots, N)$ are not quasi-continuous, in general, but rather are elements of $L^p(\Omega)$, the pointwise inequality $\left|\frac{\partial y(x)}{\partial x_i} + \theta \frac{\partial h(x)}{\partial x_i}\right| > 0$ still makes sense if $x \in \Omega$ is a Lebesgue point of both y and h. In other words, the Lebesgue points of y and h are thus points where these functions do not oscillate too much, in an average sense. Moreover, the Lebesgue Differentiation Theorem states that, given any $f \in L^1(\Omega)$, almost every $x \in \Omega$ is a Lebesgue point. Hence, almost all Lebesgue points of $\frac{\partial y(x)}{\partial x_i}$ are the Lebesgue points of $\frac{\partial y(x)}{\partial x_i} + \theta \frac{\partial h(x)}{\partial x_i}$ for θ small enough.

Definition 4.6. We say that an element $y \in W_0^{1,p}(\Omega)$ is a regular point for the Lagrangian (17) if for each $v \in W_0^{1,p}(\Omega)$ the direction $h = v - y$ is non-degenerate in the sense of Definition 4.1.

As we will see later, in regular points for the Lagrangian (17) differentiation properties are guaranteed (see Lemma 4.9 and its Corollary). In view of this, it is important to have conditions which ensure that a given point $y \in W_0^{1,p}(\Omega)$ is regular for $\Lambda(\mathcal{U}, y, \lambda)$.

Proposition 4.7. *Let $y_d \in W_0^{1,p}(\Omega)$ be a given function. Then an element $y \in W_0^{1,p}(\Omega)$ is a regular point of the Lagrangian $\Lambda(\mathcal{U}, y, \lambda)$ if the set*

$$\Phi = \left\{ x \in \Omega : \left| y(x)\Big(y(x) - y_d(x)\Big) \prod_{i=1}^{N} \frac{\partial y(x)}{\partial x_i} \right| = 0 \right\}$$

has zero Lebesgue measure.

Proof. The assertion immediately follows from Proposition 4.3, Corollary 4.4, and the fact that the sets S and

$$\{x \in \Omega : \ |y(x)| = 0\} \quad \text{and} \quad \left\{ x \in \Omega : \ \left| \frac{\partial y(x)}{\partial x_i} \right| = 0 \right\}, \ \forall\, i = 1, \dots, N$$

are proper subsets of Φ. □

Remark 4.8. It is easy to observe that if y and v in $W_0^{1,p}(\Omega)$ are two regular points of the functional $\Lambda(\mathcal{U}, y, \lambda)$, then there exists a positive number $\alpha \in \mathbb{R}$ ($\alpha \neq 0$) such that each point of the segment $[y, \alpha v] = \{y + t(\alpha v - y) \colon \forall\, t \in [0,1]\} \subset W_0^{1,p}(\Omega)$ is also regular for $\Lambda(\mathcal{U}, y, \lambda)$.

We now study the differentiability properties of the Lagrangian $\Lambda(\mathcal{U}, y, \lambda)$.

Lemma 4.9. *If $\mathcal{U} \in U_{ad}$, $\lambda \in W_0^{1,p}(\Omega)$, then, for each non-degenerate direction $h \in W_0^{1,p}(\Omega)$ at the point y, the one-sided directional derivative with respect to the variable y*

$$\mathcal{D}_y^+ \Lambda(\mathcal{U}, y, \lambda, h) = \lim_{\theta \to +0} \frac{\Lambda(\mathcal{U}, y + \theta h, \lambda) - \Lambda(\mathcal{U}, y, \lambda)}{\theta}$$

exists and takes the form

$$\begin{aligned} \mathcal{D}_y^+ \Lambda(\mathcal{U}, y, \lambda, h) = (p-1) \int_\Omega \Big([(\nabla y)^{p-2}] \mathcal{U} \nabla \lambda, \nabla h \Big)_{\mathbb{R}^N} dx \\ + (p-1) \int_\Omega |y|^{p-2} \lambda\, h \, dx + p \int_\Omega |y - y_d|^{p-1} h \, dx. \end{aligned}$$

Proof. Let h be a non-degenerate direction at the point $y \in W_0^{1,p}(\Omega)$. Following the definition of directional derivative, we have

$$\mathcal{D}_y^+ \Lambda(\mathcal{U}, y, \lambda, h) = I_1 + I_2 + I_3, \tag{20}$$

where

$$I_1 = \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_\Omega |y - y_d + \theta h|^p \, dx - \int_\Omega |y - y_d|^p \, dx\Big],$$

$$I_2 = \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_\Omega \big(\mathcal{U}(x)\left[(\nabla(y+\theta h))^{p-2}\right] \nabla(y+\theta h), \nabla\lambda\big)_{\mathbb{R}^N} \, dx - \int_\Omega \big(\mathcal{U}(x)\left[(\nabla y)^{p-2}\right] \nabla y, \nabla\lambda\big)_{\mathbb{R}^N} \, dx\Big],$$

$$I_3 = \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_\Omega |y + \theta h|^{p-2}(y+\theta h)\lambda \, dx - \int_\Omega |y|^{p-2} y\lambda \, dx\Big].$$

To identify the term I_1, we make use of the following transformations

$$\begin{aligned} I_1 &= \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_\Omega |y-y_d+\theta h|^p \, dx - \int_\Omega |y-y_d|^p \, dx\Big] \\ &= \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} (y-y_d+\theta h)^p \, dx - \int_{\Omega_0} |y-y_d|^p \, dx\Big] \\ &= \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} \sum_{i=0}^{A} \theta^i \, \frac{p(p-1)\cdots(p-i+1)}{i!} h^i (y-y_d)^{p-i} \, dx \\ &\quad - \int_{\Omega_0} |y-y_d|^p \, dx\Big] \\ &= J_1 + J_2 + J_3, \end{aligned} \tag{21}$$

where $A = p$, if p is a natural number and $A = +\infty$, otherwise. Here,

$$J_1 = \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} (y-y_d)^p \, dx - \int_{\Omega_0} |y-y_d|^p \, dx\Big],$$

$$J_2 = \lim_{\theta\to+0} \Big[p \int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} (y-y_d)^{p-1} h \, dx\Big],$$

$$J_3 = \lim_{\theta\to+0} \frac{1}{\theta}\Big[\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} \sum_{i=2}^{A} \theta^i \, \frac{p(p-1)\cdots(p-i+1)}{i!} h^i (y-y_d)^{p-i} \, dx\Big].$$

Since $\operatorname{sign}(y-y_d+\theta h) = \operatorname{sign}(y-y_d)$ almost everywhere in Ω for θ small enough, it follows that $\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} (y-y_d)^p \, dx = \int_{\Omega_{0,\theta}} |y-y_d|^p \, dx = \int_{\Omega_0 \cap \Omega_{0,\theta}} |y-y_d|^p \, dx = \int_{\Omega_0} |y-y_d|^p \, dx - \int_{\Omega_0 \setminus \Omega_{0,\theta}} |y-y_d|^p \, dx$ and hence

$$\int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} (y-y_d)^p \, dx = \int_{\Omega_0} |y-y_d|^p \, dx - \int_\Omega \chi_{\Omega_0}\big(\chi_{\Omega_0} - \chi_{\Omega_{0,\theta}}\big) |y-y_d|^p \, dx. \tag{22}$$

Consequently, in view of the property (18), we obtain

$$J_1 \equiv 0 \quad \text{and} \quad J_2 = p \int_{\Omega_0} |y-y_d|^{p-1} h \, dx = p \int_\Omega |y-y_d|^{p-1} h \, dx.$$

It remains to evaluate the last term J_3. To this end, we use Hölder's inequality and uniform convergence of the power series. It leads us to the following estimate

$$\begin{aligned}
&\Big| \int_{\Omega_{0,\theta}} \frac{y-y_d+\theta h}{|y-y_d+\theta h|} \sum_{i=2}^{A} \theta^i \, \frac{p\,(p-1)\cdots(p-i+1)}{i!} \, h^i \, (y-y_d)^{p-i} \; dx \Big| \\
&\le \sum_{i=2}^{A} \theta^i \, \frac{p\,(p-1)\cdots(p-i+1)}{i!} \int_\Omega |h|^i \, |y-y_d|^{p-i} \; dx \\
&\le \sum_{i=2}^{A} \left[\theta^i \, \frac{p\,(p-1)\cdots(p-i+1)}{i!} \|h\|^i_{L^p(\Omega)} \|y-y_d\|^{p-i}_{L^p(\Omega)} \right] = o(\theta).
\end{aligned}$$

Therefore, $J_3 = 0$ and, hence,

$$I_1 = J_1 + J_2 + J_3 = p \int_\Omega |y-y_d|^{p-1} h \, dx.$$

As for the second term in (20), we can apply the similar arguments. Namely,

$$\begin{aligned}
I_2 = \Bigg(&\int_\Omega (\mathcal{U}(x) \left[(\nabla y + \theta \nabla h)^{p-2} \right] \nabla y, \nabla \lambda)_{\mathbb{R}^N} \, dx \\
&+ \theta \int_\Omega \big(\mathcal{U}(x) \left[(\nabla y + \theta \nabla h)^{p-2} \right] \nabla h, \nabla \lambda \big)_{\mathbb{R}^N} \; dx \\
&- \int_\Omega \big(\mathcal{U}(x) \left[(\nabla y)^{p-2} \right] \nabla y, \nabla \lambda \big)_{\mathbb{R}^N} \; dx \Bigg). \qquad (23)
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
&\left[(\nabla y + \theta \nabla h)^{p-2} \right] \nabla y \\
&= \operatorname{diag} \left\{ \left| \frac{\partial y}{\partial x_1} + \theta \frac{\partial h}{\partial x_1} \right|^{p-2}, \left| \frac{\partial y}{\partial x_2} + \theta \frac{\partial h}{\partial x_2} \right|^{p-2}, \dots, \left| \frac{\partial y}{\partial x_N} + \theta \frac{\partial h}{\partial x_N} \right|^{p-2} \right\} \nabla y \\
&= \begin{bmatrix}
\left| \frac{\partial y}{\partial x_1} + \theta \frac{\partial h}{\partial x_1} \right|^{p-2} \frac{\partial y}{\partial x_1} \\
\left| \frac{\partial y}{\partial x_2} + \theta \frac{\partial h}{\partial x_2} \right|^{p-2} \frac{\partial y}{\partial x_2} \\
\cdots \\
\left| \frac{\partial y}{\partial x_N} + \theta \frac{\partial h}{\partial x_N} \right|^{p-2} \frac{\partial y}{\partial x_N}
\end{bmatrix}.
\end{aligned}$$

Further, for an arbitrary $\lambda_i \in L^p(\Omega)$, we make use of the following transformations

$$\int_\Omega \left|\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right|^{p-2}\frac{\partial y}{\partial x_i}\lambda_i\,dx$$
$$=\int_{\Omega_{i,\theta}}\frac{\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}}{\left|\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right|}\left(\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right)^{p-2}\frac{\partial y}{\partial x_i}\lambda_i\,dx$$
$$=\int_{\Omega_{i,\theta}}\frac{\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}}{\left|\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right|}\left(\frac{\partial y}{\partial x_i}\right)^{p-1}\lambda_i\,dx+(p-2)\theta\int_{\Omega_{i,\theta}}\frac{\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}}{\left|\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right|}\left(\frac{\partial y}{\partial x_i}\right)^{p-2}\frac{\partial h}{\partial x_i}\lambda_i\,dx$$
$$+\theta^2\sum_{k=2}^{B}\theta^{k-2}\frac{(p-2)\cdots(p-1-k)}{k!}\int_{\Omega_{i,\theta}}\underbrace{\frac{\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}}{\left|\frac{\partial y}{\partial x_i}+\theta\frac{\partial h}{\partial x_i}\right|}\left(\frac{\partial y}{\partial x_i}\right)^{p-1-k}\left(\frac{\partial h}{\partial x_i}\right)^{k}}_{\zeta_{ki}(x,\theta)}\lambda_i\,dx,$$

where $B=p-2$, if p is a natural number and $B=+\infty$, otherwise. Let us show that $\zeta_{ki}\in L^q(\Omega)$, where $q=\frac{p}{p-1}$. Indeed, for each $k\in\{2,3,\dots,p-2\}$ and $1\le i\le N$, using Hölder's inequality, we get

$$\begin{aligned}\|\zeta_{ki}(\cdot,\theta)\|^{\frac{p}{p-1}}_{L^{\frac{p}{p-1}}(\Omega)}&=\int_\Omega|\zeta_{ki}|^{\frac{p}{p-1}}\,dx\\&\le\int_\Omega\left|\frac{\partial y}{\partial x_i}\right|^{\frac{p(p-1-k)}{p-1}}\left|\frac{\partial h}{\partial x_i}\right|^{\frac{pk}{p-1}}\,dx\\&\le\left(\int_\Omega\left|\frac{\partial y}{\partial x_i}\right|^{\frac{p(p-1-k)}{p-1}r}\,dx\right)^{\frac{1}{r}}\left(\int_\Omega\left|\frac{\partial h}{\partial x_i}\right|^{\frac{pk}{p-1}\widehat{r}}\,dx\right)^{\frac{1}{\widehat{r}}}\\&\qquad\left\{\text{for } r=\frac{p-1}{p-k-1},\ \widehat{r}=\frac{p-1}{k},\ r^{-1}+(\widehat{r})^{-1}=1\right\}\\&=\left\|\frac{\partial y}{\partial x_i}\right\|^{\frac{p(p-k-1)}{p-1}}_{L^p(\Omega)}\left\|\frac{\partial h}{\partial x_i}\right\|^{\frac{kp}{p-1}}_{L^p(\Omega)}=O(1).\end{aligned}$$

It remains to apply the arguments similar to given above (see (21)–(22)) and pass to the limit in (23) as $\theta\to+0$. As a result, we arrive at the following representations

$$\begin{aligned}I_2&=(p-2)\int_\Omega\big(\mathcal{U}(x)\left[(\nabla y)^{p-2}\right]\nabla h,\nabla\lambda\big)_{\mathbb{R}^N}\,dx+\int_\Omega\big(\mathcal{U}(x)\left[(\nabla y)^{p-2}\right]\nabla h,\nabla\lambda\big)_{\mathbb{R}^N}\,dx\\&=(p-1)\int_\Omega\big(\mathcal{U}(x)\left[(\nabla y)^{p-2}\right]\nabla h,\nabla\lambda\big)_{\mathbb{R}^N}\,dx\\&\qquad\left\{\text{in view of symmetric property of matrices } \mathcal{U} \text{ and } [(\nabla y)^{p-2}]\right\}\\&=(p-1)\int_\Omega\big(\left[(\nabla y)^{p-2}\right]\mathcal{U}(x)\nabla\lambda,\nabla h\big)_{\mathbb{R}^N}\,dx.\end{aligned}$$

By analogy with the previous line, it can be shown show that

$$I_3 = \lim_{\theta\to 0}\frac{1}{\theta}\left[\int_\Omega |y+\theta h|^{p-1}(y+\theta h)\lambda\,dx - \int_\Omega |y|^{p-1}y\lambda\,dx\right] = (p-1)\int_\Omega |y|^{p-2}\lambda h\,dx.$$

Therefore, the "right-hand" side directional derivative of the Lagrangian functional $\Lambda(\mathcal{U},y,\lambda)$ with respect to the variable y takes the form:

$$\begin{aligned}\mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,h) &= I_1 + I_2 + I_3\\ &= (p-1)\int_\Omega \left([(\nabla y)^{p-2}]\,\mathcal{U}\nabla\lambda, \nabla h\right)_{\mathbb{R}^N} dx\\ &\quad + (p-1)\int_\Omega |y|^{p-2}\lambda h\,dx + p\int_\Omega |y-y_d|^{p-1}h\,dx. \qquad \square \end{aligned} \tag{24}$$

Lemma 4.10. *Let* $\mathcal{U}\in U_{ad}$ *and* $\lambda, y_d \in W_0^{1,p}(\Omega)$ *be given distributions. If an element* $y\in W_0^{1,p}(\Omega)$ *is a regular point of the Lagrangian* $\Lambda(\mathcal{U},y,\lambda)$*, then the mapping* $W_0^{1,p}(\Omega)\ni v\mapsto \Lambda(\mathcal{U},v,\lambda)\in\mathbb{R}$ *is Gâteaux differentiable at* y *and its Gâteaux differential* $\langle \mathcal{D}_y\Lambda(\mathcal{U},y,\lambda),h\rangle_{W_0^{1,p}(\Omega)}$ *takes the form*

$$\langle \mathcal{D}_y\Lambda(\mathcal{U},y,\lambda),h\rangle_{W_0^{1,p}(\Omega)} = \mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,h), \quad \forall\, h\in W_0^{1,p}(\Omega).$$

Proof. Let $(\mathcal{U},y,\lambda)\in U_{ad}\times\left[W_0^{1,p}(\Omega)\right]^2$ be a given triplet. Since y is a regular point of Λ, it follows that Lemma 4.9 remains valid for all $h\in W_0^{1,p}(\Omega)$. Hence, taking into account the representation (24), it is easy to see that

$$\begin{aligned}\mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,h) &= \lim_{\theta\to+0}\frac{\Lambda(\mathcal{U},y+\theta h,\lambda)-\Lambda(\mathcal{U},y,\lambda)}{\theta}\\ &= \lim_{\theta\to 0}\frac{\Lambda(\mathcal{U},y+\theta h,\lambda)-\Lambda(\mathcal{U},y,\lambda)}{\theta}\\ &= \left\langle \mathcal{D}_y\left(\int_\Omega |y-y_d|^p\lambda\right), h\right\rangle_{W_0^{1,p}(\Omega)} + \left\langle \mathcal{D}_y\left(\int_\Omega |y|^{p-2}y\lambda\right), h\right\rangle_{W_0^{1,p}(\Omega)}\\ &\quad + \left\langle \mathcal{D}_y\left(\int_\Omega (\mathcal{U}[(\nabla y)^{p-2}]\nabla y,\nabla\lambda)_{\mathbb{R}^N}dx\right), h\right\rangle_{W_0^{1,p}(\Omega)}\\ &= p\int_\Omega |y-y_d|^{p-1}h\,dx + (p-1)\int_\Omega |y|^{p-2}\lambda h\,dx\\ &\quad + (p-1)\int_\Omega \left([(\nabla y)^{p-2}]\,\mathcal{U}\,\nabla\lambda,\nabla h\right)_{\mathbb{R}^N} dx, \end{aligned} \tag{25}$$

that is,

$$\mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,h) = -\mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,-h), \quad \forall\, h\in W_0^{1,p}(\Omega),$$

and, therefore, the mapping $h\mapsto \mathcal{D}_y^+\Lambda(\mathcal{U},y,\lambda,h)$ is linear on $W_0^{1,p}(\Omega)$. Thus, following the definition of the Gâteaux derivative, we arrive at the required conclusion. $\square$

Before deriving the optimality conditions, we need the following auxiliary result.

Lemma 4.11. *Let $\mathcal{U} \in U_{ad}$, $y \in W_0^{1,p}(\Omega)$, and $v \in W_0^{1,p}(\Omega)$ be given distributions. Assume that each point of the segment $[y,v]=\{y+\alpha(v-y) : \forall\, \alpha\in[0,1]\} \subset W_0^{1,p}(\Omega)$ is regular for the mapping $v \to \Lambda(\mathcal{U},v,\lambda)$. Then there exists a positive value $\varepsilon \in [0,1]$ such that*

$$\begin{aligned}
&\Lambda(\mathcal{U},v,\lambda) - \Lambda(\mathcal{U},y,\lambda)\\
&= \langle \mathcal{D}_y\Lambda(\mathcal{U},y+\varepsilon(v-y),\lambda), v-y\rangle_{W_0^{1,p}(\Omega)}\\
&= p\int_\Omega |y+\varepsilon(v-y)-y_d|^{p-1}(v-y)\,dx + (p-1)\int_\Omega |y+\varepsilon(v-y)|^{p-2}\lambda(v-y)\,dx\\
&\quad + (p-1)\int_\Omega \big([(\nabla y+\varepsilon(\nabla v-\nabla y))^{p-2}]\,\mathcal{U}\,\nabla\lambda, \nabla(v-y)\big)_{\mathbb{R}^N}\,dx. \qquad (26)
\end{aligned}$$

Proof. For given $\mathcal{U},\lambda,y_d,y$, and v, let us consider the scalar function $\varphi(t) = \Lambda(\mathcal{U},y+t(v-y),\lambda)$. Since by Lemma 4.10, the functional $\Lambda(\mathcal{U},\cdot,\lambda)$ is Gâteaux differentiable at each point of the segment $[y,v]$, it follows that the function $\varphi=\varphi(t)$ is differentiable on $[0,1]$ and

$$\varphi'(t) = \langle \mathcal{D}_y\,\Lambda(\mathcal{U},y+t(v-y),\lambda), v-y\rangle_{W_0^{1,p}(\Omega)}, \quad \forall\, t\in[0,1].$$

To conclude the proof, it remains to take into account (25) and apply the Mean Value Theorem (or, in other words, the generalization of Rolle's Theorem): $\varphi(1)-\varphi(0)=\varphi'(\varepsilon)$ for some $\varepsilon\in[0,1]$. □

5. Optimality conditions

In this section, we assume the fulfilment of the following hypothesis:

(H1) The distributions $f \in W^{-1,q}(\Omega)$ and $y_d \in W_0^{1,p}(\Omega)$ are such that, for each admissible control $\mathcal{U} \in U_{ad} := U_b \cap U_{sol}$, the corresponding weak solution $y = y(\mathcal{U})$ of the nonlinear Dirichlet boundary value problem (10) is a regular point of the Lagrangian $\Lambda(\mathcal{U},y,\lambda)$.

It is worth noting that due to the results of Manfredi (see [21]) this hypothesis appears natural and it is not restrictive supposition in practice. Indeed, following [21], we can ensure that the set $\{x\in\Omega :\ \nabla y=0\}$ for non-constant solutions of the p-Laplace equation (a p-harmonic function, i.e. with $f=0$), has zero Lebesgue measure.

Let $(\mathcal{U}_0,y_0)\in\Xi_{sol}$ be an optimal pair for problem (13)–(15). Then

$$\Delta\Lambda = \Lambda(\mathcal{U},y,\lambda) - \Lambda(\mathcal{U}_0,y_0,\lambda) \ge 0, \forall\,(\mathcal{U},y)\in\Xi_{sol}, \quad \forall\,\lambda\in W_0^{1,p}(\Omega). \qquad (27)$$

Hence,

$$\begin{aligned}\Lambda(\mathcal{U},y,\lambda)-\Lambda(\mathcal{U}_0,y_0,\lambda) &= \Lambda(\mathcal{U},y,\lambda)-\Lambda(\mathcal{U},y_0,\lambda)+\Lambda(\mathcal{U},y_0,\lambda)-\Lambda(\mathcal{U}_0,y_0,\lambda)\\ &= \Delta_y\Lambda(\mathcal{U},y_0,\lambda)+\Lambda(\mathcal{U}-\mathcal{U}_0,y_0,\lambda)\ge 0, \qquad (28)\end{aligned}$$

for all $\lambda \in W_0^{1,p}(\Omega)$ and $\mathcal{U} \in U_{ad}$ such that $(\mathcal{U}-\mathcal{U}_0) \in U_{ad}$. Due to Hypothesis (H1) and Remark 4.8, we can suppose that each point of the segment $[y_0, y] \subset W_0^{1,p}(\Omega)$ is regular for the mapping $v \to \Lambda(\mathcal{U}, v, \lambda)$. Then, by Lemma 4.11, there exists a positive value $\varepsilon \in [0,1]$ such that

$$\begin{aligned}\Delta_y\Lambda(\mathcal{U},y_0,\lambda) &= \Lambda\left(\mathcal{U},y,\lambda\right)-\Lambda\left(\mathcal{U},y_0,\lambda\right)\\ &= \left\langle \mathcal{D}_y\Lambda(\mathcal{U},y_0+\varepsilon(y-y_0),\lambda), y-y_0\right\rangle_{W_0^{1,p}(\Omega)}. \qquad (29)\end{aligned}$$

Now we introduce the concept of quasi-adjoint states that was first considered for linear problems by Serovajskiy [27]).

Definition 5.1. We say that, for a given $\mathcal{U} \in U_{sol}$, a distribution ψ_ε is the quasi-adjoint state to $y_0 \in W_0^{1,p}(\Omega)$ if ψ_ε satisfies the following integral identity:

$$\begin{aligned}&(p-1)\int_\Omega \left([(\nabla y_\varepsilon)^{p-2}]\mathcal{U}\nabla\psi_\varepsilon, \nabla\varphi\right)_{\mathbb{R}^N} dx\\ &+(p-1)\int_\Omega |y_\varepsilon|^{p-2}\,\psi_\varepsilon\,\varphi\,dx + p\int_\Omega |y_\varepsilon-y_d|^{p-1}\varphi\,dx = 0, \quad \forall\varphi\in W_0^{1,p}(\Omega). \quad (30)\end{aligned}$$

Here, $y_\varepsilon = y_0 - \varepsilon(y-y_0)$, $y = y(\mathcal{U})$ is the solution of problem (14), (15), and $\varepsilon = \varepsilon(\mathcal{U}) \in [0,1]$ is a constant coming from equality (29).

Remark 5.2. As follows from Lemma 4.11, the constant ε essentially depends on the choice of matrix $\mathcal{U} \in U_{ad}$, i.e. $\varepsilon = \varepsilon(\mathcal{U})$. Hence, the quasi-adjoint state ψ_ε should also be considered as a composite function $\psi_\varepsilon = \psi_\varepsilon(\mathcal{U})$ (to be more specific, we have to write $\psi_\varepsilon = \psi_\varepsilon(\mathcal{U}, y(\mathcal{U}))$).

Since our main intention in this section is to derive optimality conditions for optimal control problem (13)–(15) and carry out their thorough substantiation, we begin with the following concept.

Definition 5.3. We say that the mapping $U_{ad} \ni \mathcal{U} \mapsto \psi_\varepsilon(\mathcal{U})$ possesses the $\mathcal{H}$-property at the point $\widetilde{\mathcal{U}}$ with respect to the pair of spaces $\left(L^\infty(\Omega;\mathbb{R}^{N\times N}), W_0^{1,p}(\Omega)\right)$ if for each $\mathcal{U} \in U_{ad}$ we have: $\psi_{\varepsilon,\theta} := \psi_\varepsilon(\widetilde{\mathcal{U}} + \theta(\mathcal{U}-\widetilde{\mathcal{U}})) \in W_0^{1,p}(\Omega)$ for all $\theta \in [0,1]$ and the sequence $\{\psi_{\varepsilon,\theta}\}_\theta$ is uniformly bounded in $W_0^{1,p}(\Omega)$ with respect to $\theta \in [0,1]$.

Remark 5.4. In a much stronger form this concept was introduced by Serovajskiy with respect to optimal L^∞-control problems in coefficients for linear elliptic equations [27]. In fact, in [27] it has been proved that the so-called weakened continuity of the mapping $\mathcal{U} \mapsto \psi_\varepsilon(\mathcal{U})$ (i.e. it is a property when the strong convergence of controls $\mathcal{U}_k \to \mathcal{U}$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$ implies the weak convergence $\psi_\varepsilon(\mathcal{U}_k) \rightharpoonup \psi_\varepsilon(\mathcal{U})$ in $W_0^{1,p}(\Omega)$ for all $\mathcal{U}$ and $\varepsilon \in [0,1]$) is a characteristic property for quasi-adjoint states in the linear case. However, as we will see later, this property is not attributable to the quasi-adjoint state functions provided $p > 2$.

Here we focus on the case when for a given distribution $f \in W^{-1,q}(\Omega)$, the $\mathcal{H}$-property holds true for the mapping $U_{ad} \ni \mathcal{U} \mapsto \psi_\varepsilon(\mathcal{U})$ at some point. It allows us to derive optimality conditions in a correct way. Indeed, the characteristic feature of solenoidal controls U_{sol} is the fact that the weak-$*$ convergence of controls $\mathcal{U}_k \to \mathcal{U}$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$ leads to the weak convergence in $W_0^{1,p}(\Omega)$ of the corresponding solutions $y(\mathcal{U}_k) \to y(\mathcal{U})$ as $k \to \infty$ (see Theorem 3.3). At the same time, the following result shows that the mapping $\mathcal{U} \mapsto y(\mathcal{U})$, actually, possesses a little bit stronger property.

Lemma 5.5. *Assume that $\mathcal{U}_k \to \mathcal{U}$ strongly in $L^\infty(\Omega; \mathbb{R}^{N\times N})$. Then, for the corresponding solutions of boundary value problem* (14), (15), *we have strong convergence $y_k = y(\mathcal{U}_k) \to y = y(\mathcal{U})$ in $W_0^{1,p}(\Omega)$.*

Proof. Due to the properties of the class of admissible controls U_{ad}, we can take as an equivalent norm in $W_0^{1,p}(\Omega)$ the following:

$$|||y|||_{W_0^{1,p}(\Omega)} = \left(\int_\Omega |y|^p \, dx + \int_\Omega \big(\mathcal{U}(x)[(\nabla y)^{p-2}]\nabla y, \nabla y\big)_{\mathbb{R}^N} \, dx \right)^{\frac{1}{p}}.$$

Then, by definition of weak solutions to Dirichlet problem (14), (15), we have

$$\int_\Omega \big(\mathcal{U}_k(x)[(\nabla \, y_k)^{p-2}]\nabla y_k, \nabla \phi\big)_{\mathbb{R}^N} \, dx + \int_\Omega |y_k|^{p-2} y_k \, \phi \, dx = \langle f, \phi \rangle_{W_0^{1,p}(\Omega)}, \quad (31)$$

$$\int_\Omega \big(\mathcal{U}(x)[(\nabla \, y)^{p-2}]\nabla y, \nabla \phi\big)_{\mathbb{R}^N} \, dx + \int_\Omega |y|^{p-2} y \, \phi \, dx = \langle f, \phi \rangle_{W_0^{1,p}(\Omega)}, \quad (32)$$

where ϕ is an arbitrary element of $W_0^{1,p}(\Omega)$.

Having substituted in (31), (32) $\phi = y_k$, we observe that the right-hand sides of these relations coincide. Hence, the left-hand sides must coincide as well. Thus,

$$\begin{aligned} &\int_\Omega \big(\mathcal{U}_k(x)[(\nabla \, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N} \, dx + \int_\Omega |y_k|^p \, dx \\ &= \int_\Omega \big(\mathcal{U}(x)[(\nabla \, y)^{p-2}]\nabla y, \nabla y_k\big)_{\mathbb{R}^N} \, dx + \int_\Omega |y|^{p-2} y \, y_k \, dx. \end{aligned} \quad (33)$$

Taking into account estimate (16) and Theorem 3.3, we have the implication

$$\Big[\mathcal{U}_k \to \mathcal{U} \text{ strongly in } L^\infty(\Omega; \mathbb{R}^{N\times N})\Big] \Rightarrow \Big[y_k = y(\mathcal{U}_k) \rightharpoonup y = y(\mathcal{U}) \text{ in } W_0^{1,p}(\Omega)\Big]. \quad (34)$$

Consequently,

$$\lim_{k\to\infty}\left(\int_\Omega \big(\mathcal{U}(x)[(\nabla\, y)^{p-2}]\nabla y, \nabla y_k\big)_{\mathbb{R}^N}\ dx + \int_\Omega |y|^{p-2} y\, y_k\, dx\right)$$

$$= \underbrace{\int_\Omega \big(\mathcal{U}(x)[(\nabla\, y)^{p-2}]\nabla y, \nabla y\big)_{\mathbb{R}^N}\ dx + \int_\Omega |y|^p\, dx}_{M}.$$

On the other hand, using the lower semi-continuity of the norm in $L^p(\Omega)$ with respect to weak convergence, we have

$$\liminf_{k\to\infty}\left(\int_\Omega |y_k|^p\, dx + \int_\Omega \big(\mathcal{U}_k(x)[(\nabla\, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N}\ dx\right)$$

$$= \liminf_{k\to\infty}\left(\int_\Omega |y_k|^p\, dx + \int_\Omega \big(\mathcal{U}(x)[(\nabla\, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N}\ dx\right)$$

$$+ \liminf_{k\to\infty}\int_\Omega \big((\mathcal{U}_k(x) - \mathcal{U}(x))\, [(\nabla\, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N}\ dx$$

{by (34)}

$$= \liminf_{k\to\infty}\left(\int_\Omega |y_k|^p\, dx + \int_\Omega \big(\mathcal{U}(x)[(\nabla\, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N}\ dx\right)$$

$$\geq \int_\Omega |y|^p\, dx + \int_\Omega \big(\mathcal{U}(x)[(\nabla\, y)^{p-2}]\nabla y, \nabla y\big)_{\mathbb{R}^N}\ dx = M.$$

As a result, passing to the limit in (33) as $k\to\infty$, we obtain

$$\mathrm{M} \leq \liminf_{k\to\infty}\int_\Omega \Big(\big(\mathcal{U}_k(x)[(\nabla\, y_k)^{p-2}]\nabla y_k, \nabla y_k\big)_{\mathbb{R}^N} + |y_k|^p\Big)\ dx$$

$$= \liminf_{k\to\infty}\int_\Omega \Big(\big(\mathcal{U}(x)[(\nabla\, y)^{p-2}]\nabla y, \nabla y_k\big)_{\mathbb{R}^N} + |y|^{p-2} y\, y_k\Big)\ dx = \mathrm{M}.$$

Thus, we have $|||y_k|||_{W_0^{1,p}(\Omega)} \to |||y|||_{W_0^{1,p}(\Omega)}$ and $y_k \rightharpoonup y$ in $W_0^{1,p}(\Omega)$. Taking into account the equivalence of the norms $\|\cdot\|$ and $|||\cdot|||$ in $W_0^{1,p}(\Omega)$, this ensures strong convergence $y_k \to y$ in $W_0^{1,p}(\Omega)$ and the proof is complete. □

We are now in the position to derive the first order optimality conditions for optimal control problem (13)–(15).

Theorem 5.6. *Let us suppose that* $f \in W^{-1,q}(\Omega)$, $y_d \in W_0^{1,p}(\Omega)$, *and* $U_{ad} \neq \emptyset$ *are given with* $p \geq 2$. *Let* $(\mathcal{U}_0, y_0) \in L^\infty(\Omega; \mathbb{R}^{N\times N}) \times W_0^{1,p}(\Omega)$ *be an optimal pair to the problem* (13)–(15). *Assume that the quasi-adjoint state* $\psi_\varepsilon(\mathcal{U})$ *to* $y_0 \in W_0^{1,p}(\Omega)$, *defined by* (30), *possesses the* $\mathcal{H}$*-property at* $\mathcal{U}_0$ *in the sense of Definition* 5.3. *Then* (H1) *implies the existence of an element* $\overline{\psi} \in W_0^{1,p}(\Omega)$ *such that*

$$\int_\Omega \big((\mathcal{U} - \mathcal{U}_0)[(\nabla y_0)^{p-2}]\,\nabla y_0, \nabla\overline{\psi}\big)_{\mathbb{R}^N}\ dx \geq 0, \quad \forall \mathcal{U} \in U_{ad}, \tag{35}$$

$$\int_\Omega \left(\mathcal{U}_0[(\nabla y_0)^{p-2}]\nabla y_0, \nabla\varphi\right)_{\mathbb{R}^N}\, dx + \int_\Omega |y_0|^{p-2} y_0 \varphi\, dx$$
$$= \langle f, \varphi\rangle_{W_0^{1,p}(\Omega)}, \quad \forall \varphi \in W_0^{1,p}(\Omega), \tag{36}$$
$$(p-1)\int_\Omega \left([(\nabla y_0)^{p-2}]\,\mathcal{U}_0 \nabla\overline{\psi}, \nabla\varphi\right)_{\mathbb{R}^N} dx + (p-1)\int_\Omega |y_0|^{p-2}\,\overline{\psi}\,\varphi\, dx$$
$$= p\int_\Omega |y_0 - y_d|\varphi\, dx, \quad \forall\varphi \in W_0^{1,p}(\Omega). \tag{37}$$

Proof. Let $(\widehat{\mathcal{U}}, \widehat{y}) \in \Xi_{sol}$ be an admissible pair. We substitute $\mathcal{U}_\theta := \mathcal{U}_0 + \theta(\widehat{\mathcal{U}} - \mathcal{U}_0)$, where $\theta \in [0,1]$, in relations (27) and (28). Then, by Lemma 4.11, there exists a value $\varepsilon_\theta \in [0,1]$ such that the condition (27) can be represented as follows (see (29))

$$\begin{aligned}\Delta\Lambda &= \Lambda(\mathcal{U}_\theta, y_\theta, \lambda) - \Lambda(\mathcal{U}_0, y_0, \lambda)\\ &= \langle \mathcal{D}_y\Lambda(\mathcal{U}_\theta, y_0 + \varepsilon_\theta(y_\theta - y_0), \lambda), y_\theta - y_0\rangle_{W_0^{1,p}(\Omega)} + \Lambda(\mathcal{U}_\theta - \mathcal{U}_0, y_0, \lambda)\\ &= \langle \mathcal{D}_y\Lambda(\mathcal{U}_\theta, y_0 + \varepsilon_\theta(y_\theta - y_0), \lambda), y_\theta - y_0\rangle_{W_0^{1,p}(\Omega)} + \Lambda(\theta(\widehat{\mathcal{U}} - \mathcal{U}_0), y_0, \lambda)\\ &\geq 0,\end{aligned}$$

where $y_\theta := y\left(\mathcal{U}_\theta\right) = y\left(\mathcal{U}_0 + \theta(\widehat{\mathcal{U}} - \mathcal{U}_0)\right)$ is the corresponding solution of the boundary value problem problem (14), (15). Using (26) and (17), we obtain

$$\begin{aligned}\Delta\Lambda = p&\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0) - y_d|^{p-1}(y_\theta - y_0)\, dx\\ &+ (p-1)\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0)|^{p-2}\,\lambda\,(y_\theta - y_0)\, dx\\ &+ (p-1)\int_\Omega \left([(\nabla y_0 + \varepsilon_\theta(\nabla y_\theta - \nabla y_0))^{p-2}]\,\mathcal{U}_\theta\nabla\lambda, \nabla(y_\theta - y_0)\right)_{\mathbb{R}^N} dx\\ &+ \theta\int_\Omega \left((\widehat{\mathcal{U}} - \mathcal{U}_0)[(\nabla y_0)^{p-2}]\,\nabla y_0, \nabla\lambda\right)_{\mathbb{R}^N} dx\\ \geq 0,&\quad \forall \widehat{\mathcal{U}} \in U_{ad}.\end{aligned} \tag{38}$$

In view of the $\mathcal{H}$-property, let us define the element λ in (38) as the quasi-adjoint state to $y_0 \in W_0^{1,p}(\Omega)$, that is, we set $\lambda = \psi_{\varepsilon_\theta,\theta}$, where $\psi_{\varepsilon_\theta,\theta} := \psi_{\varepsilon_\theta}(\mathcal{U}_\theta)$ satisfies the following integral identity:

$$\begin{aligned}&(p-1)\int_\Omega \left([(\nabla y_0 + \varepsilon_\theta(\nabla y_\theta - \nabla y_0))^{p-2}]\,\mathcal{U}_\theta\nabla\psi_{\varepsilon_\theta,\theta}, \nabla\varphi\right)_{\mathbb{R}^N} dx\\ &+ (p-1)\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0)|^{p-2}\psi_{\varepsilon_\theta,\theta}\,\varphi\, dx + p\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0) - y_d|^{p-1}\varphi\, dx\\ &= 0, \quad \forall\,\varphi \in W_0^{1,p}(\Omega).\end{aligned} \tag{39}$$

As a result, dividing relation (38) by θ, we can simplify it to the form

$$\int_\Omega \Big((\widehat{\mathcal{U}} - \mathcal{U}_0)[(\nabla y_0)^{p-2}]\, \nabla y_0, \nabla \psi_{\varepsilon_\theta,\theta} \Big)_{\mathbb{R}^N}\, dx \ge 0, \quad \forall \widehat{\mathcal{U}} \in U_{ad}. \tag{40}$$

It remains to pass to the limit in (39), (40) as $\theta \to +0$. To this end, we note that

(A$_1$) by the initial suppositions, $\mathcal{U}_\theta \to \mathcal{U}_0$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$ as $\theta \to 0$;
(A$_2$) by Lemma 5.5, $y_\theta \to y_0$ in $W_0^{1,p}(\Omega)$ as $\theta \to 0$;
(A$_3$) by the $\mathcal{H}$-property, there exists an element $\overline{\psi} \in W_0^{1,p}(\Omega)$ such that (within a subsequence) $\psi_{\varepsilon_\theta,\theta} \rightharpoonup \overline{\psi}$ in $W_0^{1,p}(\Omega)$ as $\theta \to 0$.
(A$_4$) if $p > 3$, then

$$\big| |a|^{p-2} - |b|^{p-2} \big| \le (p-2)\left(|a| + |b|\right)^{p-3} |a-b|\,, \quad \forall\, a, b \in \mathbb{R}; \tag{41}$$

(A$_5$) if $2 \le p \le 3$, then

$$\big| |a|^{p-2} - |b|^{p-2} \big| \le |a-b|^{p-2}, \quad \forall\, a, b \in \mathbb{R}; \tag{42}$$

Then, passing to the limit in (40) immediately leads us to (35). Therefore, in order to end the proof, it remains to establish the validity of integral identity (37). With that in mind, we rewrite (39) as follows $(p-1)I_1^\theta + (p-1)I_2^\theta + pI_3^\theta = 0$. Since,

$$\begin{aligned} I_1^\theta &= \int_\Omega \left([(\nabla \widetilde{y}_\theta)^{p-2} - (\nabla y_0)^{p-2}]\, \mathcal{U}_\theta \nabla \psi_{\varepsilon_\theta,\theta}, \nabla\varphi \right)_{\mathbb{R}^N} dx \\ &\quad + \int_\Omega \left([(\nabla y_0)^{p-2}]\, \left(\mathcal{U}_\theta - \mathcal{U}_0\right) \nabla \psi_{\varepsilon_\theta,\theta}, \nabla\varphi \right)_{\mathbb{R}^N} dx \\ &\quad + \int_\Omega \left([(\nabla y_0)^{p-2}]\, \mathcal{U}_0 \left(\nabla \psi_{\varepsilon_\theta,\theta} - \nabla\overline{\psi} \right), \nabla\varphi \right)_{\mathbb{R}^N} dx \\ &\quad + \int_\Omega \left([(\nabla y_0)^{p-2}]\, \mathcal{U}_0 \nabla\overline{\psi}, \nabla\varphi \right)_{\mathbb{R}^N} dx \\ &= J_{1,1}^\theta + J_{1,2}^\theta + J_{1,3}^\theta + J_{1,4}, \end{aligned}$$

where $\widetilde{y}_\theta := y_0 + \varepsilon_\theta (y_\theta - y_0)$, let us show that $\lim_{\theta\to 0} J_{1,j}^\theta = 0$ (j=1,2,3), and, hence, $I_1^\theta \to J_{1,4}$ as $\theta \to +0$. Using the Cauchy-Schwarz inequality, condition (6), and equivalence of the Euclidean norm $\|\cdot\|_{\mathbb{R}^N}$ and Hölder's norm $|\cdot|_p$, we have

$$\begin{aligned} |J_{1,1}^\theta| &\le \int_\Omega \left\| [(\nabla \widetilde{y}_\theta)^{p-2} - (\nabla y_0)^{p-2}] \right\|_{\mathbb{R}^{N\times N}} \|\mathcal{U}_\theta\|_{\mathbb{R}^{N\times N}} \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{\mathbb{R}^N} \|\nabla\varphi\|_{\mathbb{R}^N}\, dx \\ &\le \beta c_1 \sum_{i=1}^N \int_\Omega \left| \left| \frac{\partial \widetilde{y}_\theta}{\partial x_i} \right|^{p-2} - \left| \frac{\partial y_0}{\partial x_i} \right|^{p-2} \right| |\nabla \psi_{\varepsilon_\theta,\theta}|_p \, |\nabla\varphi|_p \, dx. \end{aligned} \tag{43}$$

Therefore, if $p>3$, then, by (A_4) and Hölder's inequality with Hölder conjugates $r=\frac{p}{p-2}>1$ and $s=\frac{p}{2}$, we can estimate (43) as follows

$$\begin{aligned}|J_{1,1}^\theta| \le & \beta c_1 (p-2)\sum_{i=1}^N \int_\Omega \left(\left|\frac{\partial \widetilde{y}_\theta}{\partial x_i}\right| + \left|\frac{\partial y_0}{\partial x_i}\right|\right)^{p-3} \left|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right| |\nabla \psi_{\varepsilon_\theta,\theta}|_p \, |\nabla\varphi|_p \, dx \\ \le & c_2 \sum_{i=1}^N \left(\int_\Omega \left(\left|\frac{\partial \widetilde{y}_\theta}{\partial x_i}\right| + \left|\frac{\partial y_0}{\partial x_i}\right|\right)^{\frac{p(p-3)}{p-2}} \left|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right|^{\frac{p}{p-2}} dx \right)^{\frac{p-2}{p}} \\ & \times \left(\int_\Omega |\nabla \psi_{\varepsilon_\theta,\theta}|_p^{\frac{p}{2}} \, |\nabla\varphi|_p^{\frac{p}{2}} \, dx \right)^{\frac{2}{p}} \\ & \left\{ \text{by Hölder's inequality with } r=\tfrac{p-2}{p-3},\ s=(p-2) \right\} \\ \le & c_2 \sum_{i=1}^N \left(\int_\Omega \left(\left|\frac{\partial \widetilde{y}_\theta}{\partial x_i}\right| + \left|\frac{\partial y_0}{\partial x_i}\right|\right)^{p} dx \right)^{\frac{p-3}{p}} \left(\int_\Omega \left|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right|^{p} dx \right)^{\frac{1}{p}} \\ & \times \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{L^p(\Omega;\mathbb{R}^N)} \|\nabla\varphi\|_{L^p(\Omega;\mathbb{R}^N)} . \end{aligned} \tag{44}$$

Since $\sup_{\theta\in[0,1]} \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{L^p(\Omega;\mathbb{R}^N)} < +\infty$ by the $\mathcal{H}$-property of $\psi_{\varepsilon_\theta,\theta}$, and $\{\varepsilon_\theta\}\subset[0,1]$, the condition ($A_2$) implies that $\widetilde{y}_\theta \to y_0$ in $W_0^{1,p}(\Omega)$ and, therefore,

$$\max_{1\le i\le N} \sup_{\theta\in[0,1]} \int_\Omega \left(\left|\frac{\partial \widetilde{y}_\theta}{\partial x_i}\right| + \left|\frac{\partial y_0}{\partial x_i}\right|\right)^{p} dx < +\infty, \quad \left\|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right\|_{L^p(\Omega)} \overset{\theta\to 0}{\to} 0.$$

Thus, by estimate (44), we conclude: $\lim_{\theta\to 0} J_{1,1}^\theta = 0$.

As for the case $2\le p\le 3$, the inequality (43) and condition (A_5) lead us to the estimate $|J_{1,1}^\theta| \le \beta c_1 \sum_{i=1}^N \int_\Omega \left|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right|^{p-2} |\nabla \psi_{\varepsilon_\theta,\theta}|_p \, |\nabla\varphi|_p \, dx$. Further it remains to repeat the trick like in (44). As a result, we obtain

$$|J_{1,1}^\theta| \le \beta c_1 \sum_{i=1}^N \left(\int_\Omega \left|\frac{\partial \widetilde{y}_\theta}{\partial x_i} - \frac{\partial y_0}{\partial x_i}\right|^{p} dx \right)^{\frac{p-2}{p}} \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{L^p(\Omega;\mathbb{R}^N)} \|\nabla\varphi\|_{L^p(\Omega;\mathbb{R}^N)} .$$

Therefore, having applied the arguments given before, we can conclude: If $2\le p\le 3$, then $\lim_{\theta\to 0} J_{1,1}^\theta = 0$.

As for the term $J_{1,2}^\theta$, we have

$$\begin{aligned}|J_{1,2}^\theta| \le & \int_\Omega \left\|[(\nabla \widetilde{y}_\theta)^{p-2}]\right\|_{\mathbb{R}^{N\times N}} \|\mathcal{U}_\theta - \mathcal{U}_0\|_{\mathbb{R}^{N\times N}} \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{\mathbb{R}^N} \|\nabla\varphi\|_{\mathbb{R}^N} \, dx \\ & \le c_3 \|\mathcal{U}_\theta - \mathcal{U}_0\|_{L^\infty(\Omega;\mathbb{R}^{N\times N})} \|\nabla \widetilde{y}_\theta\|^{p-2}_{L^p(\Omega;\mathbb{R}^N)} \|\nabla \psi_{\varepsilon_\theta,\theta}\|_{L^p(\Omega;\mathbb{R}^N)} \|\nabla\varphi\|_{L^p(\Omega;\mathbb{R}^N)} \\ & \le c_3 \sup_{\theta\in[0,1]} \|\psi_{\varepsilon_\theta,\theta}\|_{W_0^{1,p}(\Omega)} \|\widetilde{y}_\theta\|^{p-2}_{W_0^{1,p}(\Omega)} \|\varphi\|_{W_0^{1,p}(\Omega)} \|\mathcal{U}_\theta - \mathcal{U}_0\|_{L^\infty(\Omega;\mathbb{R}^{N\times N})} \\ & \le c_4 \|\mathcal{U}_\theta - \mathcal{U}_0\|_{L^\infty(\Omega;\mathbb{R}^{N\times N})} \overset{\text{by } (A_1)}{\to} 0 \quad \text{as } \theta\to 0. \end{aligned}$$

To clarify the asymptotic behavior of the term $J_{1,3}^\theta$ as θ tends to zero, we note that

$$J_{1,3}^\theta := \int_\Omega \left([(\nabla y_0)^{p-2}]\mathcal{U}_0\left(\nabla\psi_{\varepsilon_\theta,\theta} - \nabla\overline{\psi}\right), \nabla\varphi\right)_{\mathbb{R}^N}\, dx$$
$$= \int_\Omega \left(\nabla\psi_{\varepsilon_\theta,\theta} - \nabla\overline{\psi}, \mathcal{U}_0\,[(\nabla y_0)^{p-2}]\,\nabla\varphi\right)_{\mathbb{R}^N}\, dx.$$

Since $\nabla\varphi \in L^p(\Omega;\mathbb{R}^N)$, $\mathcal{U}_0 \in L^\infty(\Omega;\mathbb{R}^{N\times N})$, and $\nabla y_0 \in L^p(\Omega;\mathbb{R}^N)$, the following inclusion $\mathcal{U}_0\,[(\nabla y_0)^{p-2}]\,\nabla\varphi \in L^q(\Omega;\mathbb{R}^N)$ holds true with $q = \frac{p}{p-1}$. Hence, the condition $\lim_{\theta\to 0} J_{1,3}^\theta = 0$ is ensured by the $\mathcal{H}$-property of $\psi_{\varepsilon_\theta,\theta}$. Indeed, in this case we can suppose that within a subsequence, we have the weak convergence $\nabla\psi_{\varepsilon_\theta,\theta} \rightharpoonup \nabla\overline{\psi}$ in $L^p(\Omega;\mathbb{R}^N)$.

Thus, summing up the results given above, we finally obtain

$$\lim_{\theta\to 0} I_1^\theta = \lim_{\theta\to 0}\left(\sum_{j=1}^{3} J_{1,j}^\theta + J_{1,4}\right) = \int_\Omega \left([(\nabla y_0)^{p-2}]\,\mathcal{U}_0\nabla\overline{\psi}, \nabla\varphi\right)_{\mathbb{R}^N}\, dx. \qquad (45)$$

In analogy with the previous case, it is easy to show that

$$\lim_{\theta\to 0} I_2^\theta = \lim_{\theta\to 0}\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0)|^{p-2}\,\psi_{\varepsilon_\theta,\theta}\,\varphi\, dx = \int_\Omega |y_0|^{p-2}\,\overline{\psi}\,\varphi\, dx.$$

As for the last term in (39),

$$I_3^\theta := \int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0) - y_d|^{p-1}\varphi\, dx = 0,$$

we see that $|\widetilde{y}_\theta - y_d|^{p-1} := |y_0 + \varepsilon_\theta(y_\theta - y_0) - y_d|^{p-1} \in L^q(\Omega)$ with $q = \frac{p}{p-1}$, for all $\theta \in [0,1]$. Moreover, $\left\||\widetilde{y}_\theta|^{p-1}\right\|_{L^q(\Omega)} = \|\widetilde{y}_\theta\|^{p-1}_{L^p(\Omega)}$. Hence, strong convergence $\widetilde{y}_\theta \to y_0$ in $W_0^{1,p}(\Omega)$ implies strong convergence

$$|\widetilde{y}_\theta - y_d|^{p-1} \to |y_0 - y_d|^{p-1} \quad \text{in } L^q(\Omega).$$

As a result,

$$\lim_{\theta\to 0} I_3^\theta = \lim_{\theta\to 0}\int_\Omega |y_0 + \varepsilon_\theta(y_\theta - y_0) - y_d|^{p-1}\varphi\, dx = \int_\Omega |y_0 - y_d|^{p-1}\varphi\, dx, \qquad (46)$$

for all $\varphi \in W_0^{1,p}(\Omega)$.

Thus, combining relations (45)–(46), it is easy to see that passing to the limit in (39) leads to variational problem (37). Moreover, as immediately follows from (37), the weak limit $\overline{\psi}$ in $W_0^{1,p}(\Omega)$ of the quasi-adjoint states $\{\psi_{\varepsilon_\theta,\theta}\}_\theta$ can also be interpreted as a quasi-adjoint state to $y_0 \in W_0^{1,p}(\Omega)$ with $\varepsilon = 0$, namely, $\overline{\psi} = \psi_0(\mathcal{U}_0, y_0)$. In this sense, $\overline{\psi}$ corresponds to the "classical" notion of adjoint state. This concludes the proof. □

Remark 5.7. We remark that the optimality conditions (35)–(37) can be recovered from [2] for scalar controls without state constraints. Similarly, in [5] these optimality conditions are derived in the context of linear problems with Tikhonov regularization.

Remark 5.8. In practice, the verification of the $\mathcal{H}$-property for quasi-adjoint states is not easy. Indeed, the set $\{x \in \Omega : \ |\nabla y(x)| = 0\}$, even with zero Lebesgue measure, prohibits the existence a positive constant $\delta > 0$ satisfying inequality

$$\left([(\nabla \widetilde{y}_\theta)^{p-2}]\mathcal{U}(x)\zeta, \zeta\right)_{\mathbb{R}^N} \geq \delta\, |\zeta|_2^2 \quad \text{a.e. in } \Omega$$

for all $\mathcal{U} \in U_{ad}$, $\theta \in [0,1]$, and $\zeta \in \mathbb{R}^N$, where as usual, $\widetilde{y}_\theta = y_0 + \varepsilon_\theta(y_\theta - y_0)$. So, we can not guarantee the boundedness of the sequence $\{\psi_{\varepsilon_\theta,\theta}\}_\theta$ with respect to the norm of Sobolev space $W_0^{1,p}(\Omega)$. That's why it is reasonable to consider the solvability of variational problem (37) in appropriate weighted spaces.

Remark 5.9. We give a few comments on inequalities (41), (42). As for relation (42), its validity comes from the fact that the functions $y_1(x) = \frac{1-x^\alpha}{(1-x)^\alpha}$ and $y_2(x) = \frac{1-x^\alpha}{(1+x)^\alpha}$ are monotonically decreasing on the interval $x \in [0,1]$ provided that $0 \leq \alpha \leq 1$. Hence, $y_i(x) \leq y_i(0) = 1$, and, therefore, for each $|a| \geq |b| > 0$, we get

$$\frac{|a|^\alpha - |b|^\alpha}{|a-b|^\alpha} = \frac{|a|^\alpha - |b|^\alpha}{(|a| \pm |b|)^\alpha} = \left(1 - \left(\frac{|b|}{|a|}\right)^\alpha\right)\left(1 \pm \left(\frac{|b|}{|a|}\right)\right)^{-\alpha} \equiv y_i\left(\frac{|b|}{|a|}\right) \leq 1.$$

As a result, we arrive at inequality (42).

If $\alpha > 1$, then we can apply the following reasoning for the substantiation of inequality (41). To begin with, we note that the function $f(x) = \frac{1-x^\alpha}{1-x}$ is monotonically increasing on the interval $x \in (0,1]$. Therefore,

$$f(x) \leq f(1) = \lim_{x\to 1} \frac{1-x^\alpha}{1-x} = \left\{\frac{0}{0}\right\} = \lim_{x \to 1} \frac{(1-x^\alpha)'}{(1-x)'} = \alpha.$$

As a result, we have $\frac{1-x^\alpha}{1-x} \leq \alpha$ for all $x \in (0,1]$, and hence,

$$\begin{aligned} ||a|^\alpha - |b|^\alpha| &= (\max\{|a|,|b|\})^\alpha \left|\left(\min\{|a|,|b|\}\,(\max\{|a|,|b|\})^{-1}\right)^\alpha - 1\right| \\ &\leq \alpha \left|\min\{|a|,|b|\}\,(\max\{|a|,|b|\})^{-1} - 1\right| (\max\{|a|,|b|\})^\alpha \\ &\leq \alpha\,|a-b|\,(\max\{|a|,|b|\})^{\alpha-1} \\ &\leq \alpha\,|a-b|\,(|a|+|b|)^{\alpha-1}. \end{aligned}$$

Acknowledgement. The last author gratefully acknowledges the support of the DFG CRC 814 *Additive Fertigung.* The authors also acknowledge support by the DFG Center of Excellence *Engineering of Advanced Materials* which covered a research stay of the first author at the university of Erlangen-Nuremberg in 2013.

References

[1] Allaire, G., *Shape Optimization by the Homogenization Method.* New York: Springer 2002.

[2] Casas, E., Optimal control in the coefficients of elliptic equations with state constraints. *Appl. Math. Optim.* 26 (1992), 21 – 37.

[3] Ciuperca, I., Talibi, M. A. and Jai, M., On the optimal control in coefficients in elliptic problems. Application to the optimization of the head slider. *ESAIM Control Optim. Calc. Var.* 11 (2005), 102 – 121.

[4] D'Apice, C., De Maio, U., Kogut, O. P., On shape stability of Dirichlet optimal control problems in coefficients for nonlinear elliptic equations. *Adv. Diff. Equ.* 15 (2010)(7–8), 689 – 720.

[5] Deckelnick, K. and Hinze, M., Convergence and error analysis of a numerical method for the identification of matrix parameters in elliptic PDEs. *Inverse Problems* 28 (2012)(11), 15 pp.

[6] Fleming, W., Duality and a Priori Estimates in Markovian Optimization Problems. *J. Math. Anal. Appl.* 16 (1966), 254 – 279.

[7] Fursikov, A. V., *Optimal Control of Distributed Systems. Theory and Applications.* Providence (RI): AMS 2000.

[8] Gajewski, H., Gröger, K. and Zacharias, K., *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen* (in German). Berlin: Akademie-Verlag 1974.

[9] Haslinger, J. and Neittaanmäki, P., *Finite Element Approximation of Optimal Shape. Material and Topology Design.* Chichester: John Wiley 1996.

[10] Haslinger, J., Kocvara, M., Leugering, G. and Stingl, M., Multidisciplinary free material optimization. *SIAM J. Appl. Math.* 70 (2010)(7), 2709 – 2728.

[11] Kogut, O. P., *Qualitative analysis of one class of optimization problems for nonlinear elliptic operators.* PhD thesis, Gluskov Institute of Cyberentics NAS Kiev (2010).

[12] Kogut, P. I. and Leugering, G., *Optimal Control Problems for Partial Differential Equations on Reticulated Domains.* Boston: Birkhäuser 2011.

[13] Kogut, P. I. and Leugering, G., Optimal L^1-control in coefficients for Dirichlet elliptic problems: W-optimal solutions. *J. Optim. Theory Appl.* 150 (2011)(2), 205 – 232.

[14] Kogut, P. I. and Leugering, G., Optimal L^1-control in coefficients for Dirichlet elliptic problems: H-optimal solutions. *Z. Anal. Anwend.* 31 (2012)(1), 31 – 53.

[15] Leugering, G. and Stingl, M., PDE-constrained optimization for advanced materials. *GAMM-Mitt.* 33 (2010)(2), 209 – 229.

[16] Lions, J.-L., *Some Methods of Solving Non-Linear Boundary Value Problems.* Paris: Dunod-Gauthier-Villars 1969.

[17] Lions, J.-L., *Optimal Control of Systems Governed by Partial Differential Equations.* New York: Springer 1971.

[18] Litvinov, V. G., *Optimization in Elliptic Boundary Value Problems with Applications to Mechanics* (in Russian). Moskow: Mir 1987.

[19] Lurie, K. A., Optimal control of conductivity of a fluid moving in a channel in a magnetic field. *J. Appl. Math. Mech.* 28 (1964), 316 – 327.

[20] Lurie, K. A., *Applied Optimal Control Theory of Distributed Systems.* New York: Plenum Press 1993.

[21] Manfredi, J. J., p-harmonic functions in the plane. *Proc. Amer. Math. Soc.* 103 (1988), 473 – 479.

[22] Murat, F., Un contre-exemple pour le probleme du controle dans les coefficients (in French). *C. R. Acad. Sci. Paris. Sér. A-B* 273 (1971), A708 – A711.

[23] Murat, F., Contre-examples pour divers problemes ou le controle intervient dans les coefficients (in French). *Ann. Math. Pura Appl.* 112 (1977), 49 – 68.

[24] Murat, F. and Tartar, L., H-convergence. In: *Topics in the Mathematical Modelling of Composite Materials* (eds.: A. Cherkaev et al.). Progr. Nonlinear Diff. Equ. Appl. 31. Boston: Birkhäuser 1997, pp. 21 – 43.

[25] Phelps, R. R., *Convex Functions, Monotone Operators and Differentiability.* Berlin: Springer 1993.

[26] Pironneau, O., *Optimal Shape Design for Elliptic Systems.* Berlin: Springer 1984.

[27] Serovajskij, S. Ya., Variational inequalities in non-linear optimal control problems. *Methods Means Math. Modelling* 1 (1977), 156 – 169.

[28] Tartar, L., Problèmes de controle des coefficients dans deséquations aux derivées partielles (in French). In: *Control Theory, Numerical Methods and Computer Systems Modelling* (eds.: A. Bensoussan et al.). Lect. Notes Economics Math. Systems 107. Berlin: Springer 1975.

[29] Tartar, L., *The General Theory of Homogenization. A Personalized Introduction.* Berlin: Springer 2009.

[30] Tiba, D., *Lectures on the Control of Elliptic Systems.* Lect. Notes 32. Jyväskylä (Finland): Department of Mathematics, University of Jyväskylä 1995.

[31] Zgurovsky, M. Z. and Mel'nik, V. S., *Nonlinear Analysis and Control of Physical Processes and Fields.* Berlin: Springer 2004.

[32] Ziemer, W. P., *Weakly Differentiable Functions.* Berlin: Springer 1989.

[33] Zolezzi, T., Teoremi d'esistenza per problemi di controllo ottimo retti da equazioni ellittiche o paraboliche (in Italian). *Rend. Semin. Mat. Univ. Padova.* 44 (1970), 155 – 173.

Received August 14, 2013; revised July 30, 2014