Research Article

The Natural Gas Cash-Out Problem: A Bilevel Optimal Control Approach

Vyacheslav V. Kalashnikov,^{1,2,3} Francisco Benita,¹ and Patrick Mehlitz⁴

¹*Departamento de Ingeniería Industrial y de Sistemas, Tecnológico de Monterrey (ITESM), Campus Monterrey, Avenida Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*

²*Department of Social Modeling, Central Economics and Mathematics Institute (CEMI), Russian Academy of Sciences (RAS), Nakhimovsky Prospekt 47, Moscow 117418, Russia*

³*Department of Electronics and Computing, Sumy State University, Rimsky-Korsakov Street 2, Sumy 40007, Ukraine*

⁴*Institut für Numerische Mathematik und Optimierung, Fakultät für Mathematik und Informatik, TU Bergakademie Freiberg, Prüferstrasse 9, 09596 Freiberg, Germany*

Correspondence should be addressed to Vyacheslav V. Kalashnikov; kalash@itesm.mx

Received 26 March 2015; Revised 10 June 2015; Accepted 18 June 2015

Academic Editor: Ching-Ter Chang

Copyright © 2015 Vyacheslav V. Kalashnikov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this paper is threefold: first, it formulates the natural gas cash-out problem as a bilevel optimal control problem (BOCP); second, it provides interesting theoretical results about Pontryagin-type optimality conditions for a general BOCP where the upper level boasts a Mayer-type cost function and pure state constraints, while the lower level is a finite-dimensional mixed-integer programming problem with exactly one binary variable; and third, it applies these theoretical results in order to find possible local minimizers of the natural gas cash-out problem.

1. Introduction

Bilevel programs are hierarchical optimization problems in the sense that their constraints are defined in part by a second parametric optimization problem. Hierarchical structures can be found in diverse scientific disciplines including environmental studies, classification theory, databases, network design, transportation, game theory, economics, and new applications, such as the gas cash-out problem. In its turn, this stimulates the development of both new theoretical results and efficient algorithms to solve bilevel programming problems.

A particular case of the bilevel programming problem is presented by the following optimal control problem arising in the natural gas industry. This problem has been formulated as a mixed-integer model [1–3] arising from the minimization of cash-out penalty costs of a natural gas shipping company. In this study, we formulate the gas cash-out problem as

a *bilevel optimal control problem* (BOCP) where the upper level is equipped with a Mayer-type cost function and pure state constraints, whereas the lower level is formulated as a finite-dimensional mixed-integer programming problem (MIP) with only one binary variable.

One of the main goals of this paper is to provide a maximum principle for this problem and to present a detailed discussion. The rest of the paper is organized as follows. Section 2 specifies the model and describes the system's behavior. Section 3 is the core of our whole approach: it contains the general formulation of the cash-out model, lists the assumptions we accept, reformulates the model as a single-level problem, provides Pontryagin-type optimality conditions, and illustrates the theory by means of a small example. Section 4 presents the results of application of the optimality conditions to the cash-out problem. Finally, Section 5 presents conclusions and possible items of the future research.

2. The Natural Gas Cash-Out Problem

Let us consider a system consisting of a single natural gas shipping company (NGSC) engaged in business with a Pipeline Operating Company (POC). The NGSC intends to transport the natural gas it buys from the wells to several consumption points (pool zones). The operative process is defined within a certain time frame $[0, T]$, for example, a month. In every $t \in [0, T]$, the NGSC buys gas at the wells, injects it into the pipeline at its string point, and extracts the amount of gas to sell it at the pool zones down the pipeline. The NGSC is required to book capacity in every pool zone before the operative period starts, so that the nominal extraction amounts are known by the POC beforehand, so that the pipeline's operations could be planned. However, the NGSC is allowed, within certain limits, to extract a different amount of fuel more than planned/booked originally, thus creating certain *imbalances*. An imbalance is formally defined as the difference between the booked capacity and the extracted amount. The gas extracted at every pool zone is then sold during each period of time, $t \in [0, T]$, to the Local Distribution Companies (LDCs) to fulfill prearranged contracts, with any unsold amounts marketed at different prices to external LDCs with no binding contracts to the NGSC.

To prevent the NGSC from generating large unjustified imbalances, the POC will limit the levels of imbalances, positive or negative, on each period of time in every pool zone, as well as the possible change, or “swing” of those imbalances from t to $t + \varepsilon$. At the end of the operative period, T , the POC will analyse the existing imbalances in its pipeline and redistribute them among the pool zones moving gas from zones with positive imbalances to zones with negative (and only negative!) imbalances, until only nonnegative or nonpositive imbalances exist at all pool zones (the same sign of the imbalances as the result of balancing procedure is also required). At this point, a cash-out will be arranged between the parties, with the pipeline charging high costs for the taken extra gas and paying low prices for the bought gas so that arbitrage on part of the NGSC be discouraged.

The natural gas cash-out problem has been modeled as an *optimistic bilevel optimization problem* in finite-dimensional spaces (cf. [1–3]). In these formulations the upper level corresponds to the decision of a single NGSC regarding its capacity booking, extraction levels, and imbalances whereas the lower level deals with the decision of the POC to come up with a balancing scheme that would please both parties. This optimization modeling is intended to be a tool for NGSCs to better manage the created imbalances in their continuous-time business. Therefore, we will interchangeably refer to the NGSC's (POC's) problem as the upper level (lower level) problem. Previous studies were focused on numerical solutions; however, in this work we formulate the problem as a BOCPP where the aim of the paper is to provide optimality conditions for such formulation.

2.1. Problem Statement. Following the studies of Dempe et al. [1] and Kalashnikov and Ríos-Mercado [3], the decision making process for the NGSC is to determine how to carry

out its daily imbalances such that the penalty that will be imposed by the POC is minimized. The following notation is used to describe the model.

Indices and Sets

i, j are zone pool indices; $i, j \in N = \{1, \dots, n\}$.

t is time index; $t \in [0, T]$.

Parameters

$x_i^L(t), x_i^U(t)$ are lower and upper bounds on imbalances at time $t \in [0, T]$ in zone $i \in N$.

$x^L(t), x^U(t)$ are lower and upper bounds for the sum of imbalances at time $t \in [0, T]$.

$u_i^L(t), u_i^U(t)$ are lower and upper bounds on balance swings at time $t \in [0, T]$ in zone $i \in N$.

a_{ij} is percentage of fuel retained for moving one unit of gas from zone i to zone j ; $i < j, i, j \in N$.

b_{ij} is transportation change for moving one unit of gas from zone i to zone j ; $i < j, i, j \in N$.

c_{ij} is backward haul credit for moving one unit of gas from zone j to zone i ; $i < j, i, j \in N$.

d_i is imbalance penalization parameter, $i \in N$.

$x_i(0)$ is initial imbalance at $t = 0$ in zone i ; $i \in N$.

Controls and States (Upper Level)

$u_i(t)$ is imbalance swing at time $t \in [0, T]$ in zone i ; $i \in N$.

$x_i(t)$ is imbalance at time $t \in [0, T]$ in zone i ; $i \in N$.

Decision Variables (Lower Level)

y_1 is total cash-out cost for the NGSC.

$y_{2,i}$ is final imbalance at zone i ; $i \in N$.

$y_{3,ij}$ is forward haul volume moved from zone i to zone j ; $i < j, i, j \in N$.

$y_{4,ij}$ is backward haul volume moved from zone j to zone i ; $i < j, i, j \in N$.

y_5 is auxiliary variable for linearization.

Auxiliary Variables

q is binary variable equal to 1 (0) if final imbalances are nonnegative (nonpositive). In the special case when all the imbalances are zeros, we accept $q = 1$.

Let us postulate a quite natural assumption that $x^L, x^U, \dots, x_n^L, x_n^U, x_1^U, \dots, x_n^U : [0, T] \rightarrow \mathbb{R}$ and the $u_1^L, \dots, u_n^L, u_1^U, \dots, u_n^U : [0, T] \rightarrow \mathbb{R}$ are measurable functions which satisfy $x_i^L(t) < x_i^U(t)$, $x^L(t) < x^U(t)$, $\sum_{i \in N} x_i^L(t) < x^L(t)$, $x^U(t) < \sum_{i \in N} x_i^U(t)$, and $u_i^L(t) < u_i^U(t)$ for any $t \in [0, T]$. Furthermore, we introduce

$$y_2 := (y_{2,1} \ \dots \ y_{2,n})^T,$$

$$y_l := (y_{l,12} \ \dots \ y_{l,1n} \ y_{l,23} \ \dots \ y_{l,2n} \ \dots \ y_{l,n-1n})^T \quad (1)$$

$$l = 3, 4.$$

The NGSC's problem is given by

$$y_1 \longrightarrow \min_{x,u,y,q} \quad (2a)$$

$$\dot{x}_i(t) = u_i(t) \quad \text{a.e. } t \in [0, T], \quad i \in N \quad (2b)$$

$$u_i^L(t) \leq u_i(t) \leq u_i^U(t) \quad \text{a.e. } t \in [0, T], \quad i \in N \quad (2c)$$

$$x_i^L(t) \leq x_i(t) \leq x_i^U(t) \quad \forall t \in [0, T], \quad i \in N \quad (2d)$$

$$x^L(t) \leq \sum_{i \in N} x_i(t) \leq x^U(t) \quad \forall t \in [0, T] \quad (2e)$$

$$x_i(0) = 0^n, \quad i \in N \quad (2f)$$

$$(y_1, y_2, y_3, y_4, y_5, q) \in \Psi(x(T)). \quad (2g)$$

The objective function of the NGSC (2a) is to minimize the total cash-out cost, given the control system (2b) and some initial conditions (2f). Constraints (2c)–(2e) reflect the bounds on the imbalances and imbalance swings. Note that constraints (2d)–(2e) involve only state variables. Moreover, the objective function variable y_1 depends on the POC's response to this upper level model; this is not a leader's decision variable *per se*. In the latter problem the NGSC's objective function is minimized with respect to x, u, y , and q ; thus this formulation is adopting the *optimistic approach* where the upper level decision maker is able to influence the POC so that the latter select his/her reaction in the leader's favor. Finally, $\Psi(x(T))$ is the solution set of the POC's problem for fixed $x(T)$. Hence, considering the vector of the imbalances in all zones at the final time moment $x(T) \in \mathbb{R}^n$ as given, the POC's problem can be formulated as follows:

$$y_5 \longrightarrow \min_{y_1, y_2, y_3, y_4, y_5, q} \quad (3a)$$

$$\sum_{j:i < j} y_{3,ij} + \sum_{k:k < i} y_{4,ki} \leq \max\{0, x_i(T)\}, \quad i \in N \quad (3b)$$

$$y_{3,ij} \leq \begin{cases} x_i(T) & \text{if } x_i(T) > 0, \quad x_j(T) < 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3c)$$

$$i, j \in N, \quad i < j$$

$$y_{4,ij} \leq \begin{cases} x_j(T) & \text{if } x_j(T) > 0, \quad x_i(T) < 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3d)$$

$$i, j \in N, \quad i < j$$

$$\min\{0, x_i(T)\} \leq y_{2,i} \leq \max\{0, x_i(T)\}, \quad i \in N \quad (3e)$$

$$y_{2,j} = x_j(T) + \sum_{i:i < j} [(1 - a_{ij}) y_{3,ij} - y_{4,ij}] \quad (3f)$$

$$+ \sum_{k:k > j} (y_{4,jk} - y_{3,jk}), \quad j \in N$$

$$y_1 = - \sum_{i \in N} d_i y_{2,i} - \sum_{(i,j):i < j} c_{ij} y_{4,ij} \quad (3g)$$

$$+ \sum_{(i,j):i < j} b_{ij} (1 - a_{ij}) y_{3,ij},$$

$$- y_5 \leq y_1 \leq y_5, \quad (3h)$$

$$- M_1 (1 - q) \leq y_{2,i} \leq M_1 q, \quad i \in N \quad (3i)$$

$$y_5 \geq 0, \quad (3j)$$

$$y_{3,ij}, y_{4,ij} \geq 0, \quad i, j \in N, \quad i < j$$

$$q \in \{0, 1\}, \quad (3k)$$

where $M_1 \in \mathbb{R}_+$ is a large enough scalar value, whose lower bound can be easily calculated, making use of duality theory, due to the compactness of the feasible set in states $x_i(T) \in \mathbb{R}$ (see, e.g., [2]). Constraint (3e) arises from the business rule saying that when being rearranged by the POC, the imbalances cannot change their signs. Constraints (3b)–(3e) imply that the pipeline can haul forward or backward an amount of gas from one zone to another that does not exceed (i) the positive imbalance in the source zone and (ii) the absolute value of the negative imbalance one has in the target pool zone. Constraint (3f) explains the relationship between the last moment imbalances $x(T) \in \mathbb{R}^n$, the final imbalances y_2 , and the forward/backward haul variables y_3, y_4 . Constraints (3g)–(3h) define the cash-out variable y_1 and show how the objective function variable y_5 is related to y_1 via linear inequalities (the latter linear inequalities help reduce the nonlinear original lower level objective function to a linear one). Restriction (3i) reflects the requirement that all the final imbalances must be of the same sign, and it is exactly where the binary variable q appears.

In [2], the mixed-integer nonlinear program (3a)–(3k) was reduced to an “almost” linear formulation. Indeed, only the presence of the binary variable q prevents the modified problem from being a linear program. Hence, for a given vector $x(T) \in \mathbb{R}^n$, we have the equivalent mixed-integer programming problem:

$$y_5 + M_2 \sum_{i \in N} (y_{6,i} + y_{7,i}) \longrightarrow \min_{y_1, y_2, y_3, y_4, y_5, y_6, y_7, q} \quad (4a)$$

$$(1 - a_{ij}) y_{3,ij} \leq y_{6,i}, \quad i, j \in N, \quad i < j \quad (4b)$$

$$(1 - a_{ij}) y_{3,ij} \leq y_{7,j}, \quad i, j \in N, \quad i < j \quad (4c)$$

$$y_{4,ij} \leq y_{6,j}, \quad i, j \in N, \quad i < j \quad (4d)$$

$$y_{4,ij} \leq y_{7,i}, \quad i, j \in N, \quad i < j \quad (4e)$$

$$-y_{7,i} \leq y_{2,i} \leq y_{6,i}, \quad i \in N \quad (4f)$$

$$y_{6,i} \geq x_i(T), \quad i \in N \quad (4g)$$

$$y_{7,i} \geq -x_i(T), \quad i \in N \quad (4h)$$

$$y_{6,i}, y_{7,i} \geq 0, \quad i \in N, \quad (4i)$$

$$(3f)-(3k), \quad (4j)$$

where $M_2 \in \mathbb{R}_+$ is a large enough scalar value, and from (4a), (4g)–(4i) we have $y_{6,i} = \max\{x_i(T), 0\}$ and $y_{7,i} = \max\{-x_i(T), 0\}$. Note that (4a)–(4j) is a linear programming problem involving a single binary variable. Moreover, (4a)–(4j) is a linear program for each value of q , and it was shown in [1] (for $M_2 = 0$) that this problem is solvable for at least one of the two possible values of $q = 0$ or $q = 1$. Then duality theory guarantees that there exists a finite penalty value of $M_2 > 0$ such that constraints (4g) and (4h) are active at any optimal solution of the program in question. From now on, without loss of generality, we will use $\Psi(x(T))$ in order to denote the solution set of (4a)–(4j) for a fixed parameter $x(T)$. Based upon that, we replace the last constraint (2g) with the inclusion $(y_1, y_2, y_3, y_4, y_5, y_6, y_7, q) \in \Psi(x(T))$, where

$$y_k := (y_{k,1} \ \dots \ y_{k,n})^\top, \quad k = 6, 7. \quad (5)$$

Due to the fact that the lower level problem is purely linear, it seems to be convenient to represent it in a matrix form. Therefore, we introduce some extra notation for vectors and matrices. Namely, by $I^k \in \mathbb{R}^{k \times k}$ we denote the identity matrix of dimension $k \times k$; $\mathbf{0}^{k \times l}$ denotes the zero matrix of dimension $k \times l$; and the zero-vector of dimension k is denoted by $0^k \in \mathbb{R}^k$. Let $e_l^k \in \mathbb{R}^k$ be the l th unity vector of dimension k and $s^k \in \mathbb{R}^k$ the all-ones vector of dimension k . We introduce the natural number $\bar{n} := (1/2)(n^2 - n)$ and use $\text{diag}(z)$ in order to define (for any $z \in \mathbb{R}^k$) a matrix in $\mathbb{R}^{k \times k}$ whose main diagonal is given by z , while all its other entries are zero. For two vectors $z_1, z_2 \in \mathbb{R}^k$, the product $z_1 \cdot z_2 \in \mathbb{R}^k$ will be understood as the componentwise product; that is, $(z_1 \cdot z_2)_j := z_{1j}z_{2j}$, $j = 1, \dots, k$.

Next, we define the lower level variable component vector y and some auxiliary vectors:

$$\begin{aligned} y &:= (y_1 \ y_2^\top \ y_3^\top \ y_4^\top \ y_5^\top \ y_6^\top \ y_7^\top)^\top, \\ a &:= (a_{1,2} \ \dots \ a_{1,n} \ a_{2,3} \ \dots \ a_{2,n} \ \dots \ a_{n-1,n})^\top, \\ b &:= (b_1 \ \dots \ b_n)^\top, \\ c &:= (c_{1,2} \ \dots \ c_{1,n} \ c_{2,3} \ \dots \ c_{2,n} \ \dots \ c_{n-1,n})^\top, \\ d &:= (d_1 \ \dots \ d_n)^\top. \end{aligned} \quad (6)$$

Let us introduce the matrices $H_1, H_2 \in \mathbb{R}^{\bar{n} \times n}$ and $H_3, H_4 \in \mathbb{R}^{n \times \bar{n}}$ as follows:

$$\begin{aligned} H_1 &:= \begin{pmatrix} -s^{n-1} & 0^{n-1} & \dots & 0^{n-1} & 0^{n-1} \\ 0^{n-2} & -s^{n-2} & & 0^{n-2} & 0^{n-2} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & -1 & 0 \end{pmatrix}; \\ H_2 &:= \begin{pmatrix} 0^{n-1} & -e_1^{n-1} & -e_2^{n-1} & \dots & -e_{n-1}^{n-1} \\ 0^{n-2} & 0^{n-2} & -e_1^{n-2} & \dots & -e_{n-2}^{n-2} \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & -1 \end{pmatrix}; \\ H_3 &:= \begin{pmatrix} -s^{n-1} & (1-a_{1,2})e_1^{n-1} & (1-a_{1,3})e_2^{n-1} & \dots & (1-a_{1,n})e_{n-1}^{n-1} \\ 0^{n-2} & -s^{n-2} & (1-a_{2,3})e_1^{n-2} & \dots & (1-a_{2,n})e_{n-2}^{n-2} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1-a_{n-1,n} \end{pmatrix}^\top; \\ H_4 &:= \begin{pmatrix} s^{n-1} & -e_1^{n-1} & -e_2^{n-1} & \dots & -e_{n-1}^{n-1} \\ 0^{n-2} & s^{n-2} & -e_1^{n-2} & \dots & -e_{n-2}^{n-2} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & -1 \end{pmatrix}^\top, \end{aligned} \quad (7)$$

and use the latter ones to compose the overall system's matrices. In order to do that, we define $A_1 \in \mathbb{R}^{(3n^2+5n+2) \times n}$, $A_2 \in \mathbb{R}^{(n+1) \times n}$, $B_1 \in \mathbb{R}^{(3n^2+5n+2) \times (n^2+2n+2)}$, $B_2 \in \mathbb{R}^{(n+1) \times (n^2+2n+2)}$, $C_1, D_1 \in \mathbb{R}^{3n^2+5n+2}$, $C_2, D_2 \in \mathbb{R}^{n+1}$, and $w \in \mathbb{R}^{n^2+2n+2}$ as follows:

$$A_1 := \begin{bmatrix} \mathbb{O}^{\bar{n} \times n} \\ \mathbb{O}^{\bar{n} \times n} \\ \mathbb{O}^{\bar{n} \times n} \\ \mathbb{O}^{\bar{n} \times n} \\ \mathbb{O}^{n \times n} \\ \mathbb{O}^{n \times n} \\ I^n \\ -I^n \\ \mathbb{O}^{n \times n} \\ \mathbb{O}^{n \times n} \\ 0 \\ 0 \\ \mathbb{O}^{n \times n} \\ \mathbb{O}^{n \times n} \\ \mathbb{O}^{\bar{n} \times n} \\ \mathbb{O}^{\bar{n} \times n} \end{bmatrix};$$

$$C_1 := \begin{bmatrix} 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^n \\ 0^n \\ 0^n \\ 0^n \\ 0^n \\ 0^n \\ 0 \\ 0 \\ M_1 s^n \\ -M_1 s^n \\ 0^{\bar{n}} \\ 0^{\bar{n}} \end{bmatrix};$$

$$D_1 := \begin{bmatrix} 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^{\bar{n}} \\ 0^n \\ 0^n \\ 0^n \\ 0^n \\ 0^n \\ 0 \\ 0 \\ M_1 s^n \\ 0^n \\ 0^{\bar{n}} \\ 0^{\bar{n}} \end{bmatrix};$$

$$A_2 := \begin{bmatrix} I^n \\ \mathbb{O}^{1 \times n} \end{bmatrix};$$

$$C_2 := \begin{bmatrix} 0^n \\ 0 \end{bmatrix};$$

$$D_2 := \begin{bmatrix} 0^n \\ 0 \end{bmatrix};$$

$$B_1$$

$$:= \begin{bmatrix} 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \text{diag}(s^{\bar{n}} - a) & \mathbb{O}^{\bar{n} \times \bar{n}} & 0^{\bar{n}} & H_1 & \mathbb{O}^{\bar{n} \times n} \\ 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \text{diag}(s^{\bar{n}} - a) & \mathbb{O}^{\bar{n} \times \bar{n}} & 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & H_2 \\ 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \mathbb{O}^{\bar{n} \times \bar{n}} & I^{\bar{n}} & 0^{\bar{n}} & H_2 & \mathbb{O}^{\bar{n} \times n} \\ 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \mathbb{O}^{\bar{n} \times \bar{n}} & I^{\bar{n}} & 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & H_1 \\ 0^n & -I^n & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & \mathbb{O}^{n \times n} & -I^n \\ 0^n & I^n & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & -I^n & \mathbb{O}^{n \times n} \\ 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & -I^n & \mathbb{O}^{n \times n} \\ 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & \mathbb{O}^{n \times n} & -I^n \\ 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & -I^n & \mathbb{O}^{n \times n} \\ 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & \mathbb{O}^{n \times n} & -I^n \\ -1 & \mathbb{O}^{1 \times n} & \mathbb{O}^{1 \times \bar{n}} & \mathbb{O}^{1 \times \bar{n}} & -1 & \mathbb{O}^{1 \times n} & \mathbb{O}^{1 \times n} \\ 1 & \mathbb{O}^{1 \times n} & \mathbb{O}^{1 \times \bar{n}} & \mathbb{O}^{1 \times \bar{n}} & -1 & \mathbb{O}^{1 \times n} & \mathbb{O}^{1 \times n} \\ 0^n & -I^n & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times n} \\ 0^n & I^n & \mathbb{O}^{n \times \bar{n}} & \mathbb{O}^{n \times \bar{n}} & 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times n} \\ 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & -I^{\bar{n}} & \mathbb{O}^{\bar{n} \times \bar{n}} & 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \mathbb{O}^{\bar{n} \times n} \\ 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \mathbb{O}^{\bar{n} \times \bar{n}} & -I^{\bar{n}} & 0^{\bar{n}} & \mathbb{O}^{\bar{n} \times n} & \mathbb{O}^{\bar{n} \times n} \end{bmatrix};$$

$$B_2 := \begin{bmatrix} 0^n & -I^n & H_3 & H_4 & 0^n & \mathbb{O}^{n \times n} & \mathbb{O}^{n \times n} \\ 1 & d^\top & [b \cdot (a - s^\top)]^\top & c^\top & 0 & \mathbb{O}^{1 \times n} & \mathbb{O}^{1 \times n} \end{bmatrix};$$

$$w := [0 \quad 0^{n^\top} \quad 0^{\bar{n}^\top} \quad 0^{\bar{n}^\top} \quad 1 \quad M_2 \cdot s^{n^\top} \quad M_2 \cdot s^{n^\top}]^\top.$$
(8)

Hence, we can rewrite (4a)–(4j) as

$$\begin{aligned} w^\top y &\longrightarrow \min_{y, q} \\ \text{subject to} \quad & A_1 x(T) + B_1 y + C_1 q \leq D_1, \\ & A_2 x(T) + B_2 y + C_2 q = D_2, \\ & q \in \{0, 1\}. \end{aligned}$$
(9)

3. General Setting and Optimality Conditions

This section deals with the theoretical results providing optimality conditions for the bilevel program in general, with an aim to apply them to our gas cash-out problem in Section 4.

3.1. Preliminaries and Further Notation. Throughout this paper, we denote by $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m \times n}$, $\mathbb{R}_+^n$, and $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ the set of real numbers, the space of all real vectors with n components, the space of all real matrices with m rows and n columns, the nonnegative orthant of $\mathbb{R}^n$, and the extended real line, respectively. For any matrix $X \in \mathbb{R}^{m \times n}$, let $X^\top \in \mathbb{R}^{n \times m}$ denote its transpose. Next, for an arbitrary subset $M \subseteq \mathbb{R}^n$, we use $\text{int}(M)$, $\text{ri}(M)$, and $\text{conv}(M)$ in order to denote the interior, the relative interior, and the convex hull of M , respectively. For $x^* \in M$, we define the *Fréchet normal cone* $\widehat{\mathcal{N}}_M(x^*)$ to M at x^* by means of

$$\widehat{\mathcal{N}}_M(x^*) := \left\{ d \in \mathbb{R}^n \mid \limsup_{x \rightarrow x^*, x \in M} \frac{(x - x^*)^\top d}{\|x - x^*\|} \leq 0 \right\}. \quad (10)$$

Here, $\|\cdot\|$ is an arbitrary norm in $\mathbb{R}^n$. Furthermore, the *basic normal cone* (sometimes referred to as *Mordukhovich normal cone*) $\mathcal{N}_M^B(x^*)$ to M at x^* is given by

$$\mathcal{N}_M^B(x^*) := \limsup_{x \rightarrow x^*, x \in M} \widehat{\mathcal{N}}_M(x), \quad (11)$$

where the upper limit has to be interpreted in the sense of Painlevé-Kuratowski (cf. [4]). If M is a convex set, then $\widehat{\mathcal{N}}_M(x^*)$ and $\mathcal{N}_M^B(x^*)$ both coincide with the normal cone of convex analysis, which we denote by $\mathcal{N}_M(x^*)$. Let us denote by $\mathbb{U}_{\mathbb{R}^n}$ and $\mathbb{B}_{\mathbb{R}^n}$ the open and closed unit balls in $\mathbb{R}^n$, and, for any $\varepsilon > 0$, $x^* \in \mathbb{R}^n$, we define $\mathbb{U}_{\mathbb{R}^n}^\varepsilon(x^*) := \{x^*\} + \varepsilon \mathbb{U}_{\mathbb{R}^n}$ and $\mathbb{B}_{\mathbb{R}^n}^\varepsilon(x^*) := \{x^*\} + \varepsilon \mathbb{B}_{\mathbb{R}^n}$, respectively. Given a differentiable function $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, its Jacobian at $x \in \mathbb{R}^n$ is represented by $\nabla \mathcal{F}(x) \in \mathbb{R}^{m \times n}$. Next, we denote by C and C^* the space of scalar continuous functions defined over the interval $[0, T]$ and the dual space of C , respectively. The set of elements in C^* taking nonnegative values at nonnegative functions in C is denoted by $C^\oplus$. Given a nonnegative measure

$\mu \in C^\oplus$, its norm $\|\mu\|_{C^\oplus}$ coincides with the total variation of μ defined as $\int_0^T |\mu|(ds)$ and the support of the measure μ ; written $\text{supp}(\mu)$ is the smallest closed set $A \subseteq [0, T]$ such that for any relatively open subset $B \subseteq [0, T] \setminus A$ one has $\mu(B) = 0$. The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted by $\|x\|$, the space of all Lebesgue-integrable functions w such that $w : [0, T] \rightarrow \mathbb{R}^n$ is described by L_n^1 , and by $W_n^{1,1}$ we represent the Sobolev space that contains all absolutely continuous $\mathbb{R}^n$ -valued functions on $[0, T]$, whose weak derivatives belong to L_n^1 . By L_n^∞ we denote the Banach space of all vector-functions with n essentially bounded components defined over $[0, T]$. Furthermore, by $\mathcal{L} \times \mathcal{B}^n$ we express the product σ -algebra generated by the Lebesgue set $\mathcal{L}$ and the Borel subsets of $\mathbb{R}^n$.

For a lower semicontinuous functional $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, we introduce its domain as $\text{dom}(f) := \{x \in \mathbb{R}^n \mid |f(x)| < +\infty\}$. Suppose that $x^* \in \text{dom}(f)$ is fixed. Then we define the *Fréchet subdifferential* $\widehat{\partial}f(x^*)$ of f at x^* as

$$\begin{aligned} \widehat{\partial}f(x^*) &:= \left\{ a \right. \\ &\left. \in \mathbb{R}^n \mid \liminf_{x \rightarrow x^*} \frac{f(x) - f(x^*) - (x - x^*)^\top a}{\|x - x^*\|} \geq 0 \right\}. \end{aligned} \quad (12)$$

Furthermore, we make use of the *singular subdifferential* $\partial^\infty f(x^*)$ of f at x^* , which is defined as stated below:

$$\partial^\infty f(x^*) := \limsup_{x \rightarrow x^*, \alpha \searrow 0} \alpha \cdot \widehat{\partial}f(x). \quad (13)$$

Note that $\partial^\infty f(x^*) = \{0^n\}$ holds if and only if f is locally Lipschitz continuous at x^* (cf. [5]). In the latter case $\partial^\circ f(x^*)$, the Clarke subdifferential of f at x^* is a nonempty, compact, and convex set (cf. [6]) and consequently we have $x^* \in \text{int}(\text{dom}(f))$. Suppose that f is a convex functional and $x^* \in \text{ri}(\text{dom}(f))$ is chosen arbitrarily. Then $\widehat{\partial}f(x^*)$ and $\partial^\circ f(x^*)$ both coincide with the classical subdifferential of f at x^* , the latter being denoted by $\partial f(x^*)$. For further information on these concepts of variational analysis, we refer the interested reader to [7].

Let $\Theta : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be an arbitrary set-valued mapping. The domain and graph of Θ are denoted, respectively, by $\text{dom}(\Theta)$ and $\text{graph}(\Theta)$. We say that Θ is a *polyhedral mapping* provided that its graph is the union of a finite collection of polyhedra. The mapping Θ is called *locally Lipschitz continuous* at $x^* \in \mathbb{R}^n$ whenever there exist scalars $L_\Theta, \varepsilon_\Theta > 0$ such that for all $x, x' \in \mathbb{U}_{\mathbb{R}^n}^\varepsilon(x^*)$ the following condition is satisfied:

$$\Theta(x) \subseteq \Theta(x') + L_\Theta \|x - x'\| \mathbb{B}_{\mathbb{R}^m}. \quad (14)$$

If the above inclusion holds only when $x' := x^*$ is fixed, then Θ is called *locally upper Lipschitz continuous* at x^* . We call Θ *upper (lower) semicontinuous* at x^* provided that for any open set $O \subseteq \mathbb{R}^m$ satisfying $\Theta(x^*) \subseteq O$ ($\Theta(x^*) \cap O \neq \emptyset$) we find $\varepsilon > 0$ such that $\Theta(x) \subseteq O$ ($\Theta(x) \cap O \neq \emptyset$) holds true for all $x \in \mathbb{U}_{\mathbb{R}^n}^\varepsilon(x^*)$. Now choose an arbitrary point $(x^*, y^*) \in \text{graph}(\Theta)$. We say that Θ is *inner semicontinuous*

at (x^*, y^*) if for any sequence $\{x^k\} \subseteq \mathbb{R}^n$ converging to x^* we find a sequence $\{y^k\} \subseteq \mathbb{R}^m$ converging to y^* and $k_0 \in \mathbb{N}$ such that $y^k \in \Theta(x^k)$ is true for all $k \in \mathbb{N}$ greater than k_0 . Obviously, if Θ is locally Lipschitz continuous at x^* , then it is inner semicontinuous at any point $(x^*, \bar{y})$ with $\bar{y} \in \Theta(x^*)$. The *normal coderivative* $D^*\Theta(x^*, y^*) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ of Θ at (x^*, y^*) is defined as stated below for any $y \in \mathbb{R}^m$:

$$D^*\Theta(x^*, y^*)(y) := \{x \in \mathbb{R}^n \mid (x, -y) \in \mathcal{N}_{\text{graph}(\Theta)}^B(x^*, y^*)\}. \quad (15)$$

Finally, we recall some helpful calculation rules concerning the Minkowski sum of sets. Due to its simplicity the corresponding proof is omitted here.

Lemma 1. *Let $A, B, C, D \subseteq \mathbb{R}^n$ be arbitrary sets. Then the following statements hold:*

- (i) $(A + B) \cup (C + D) \subseteq (A \cup C) + (B \cup D)$;
- (ii) $A \subseteq B + C \Rightarrow A \times D \subseteq (B \times D) + (C \times \{0^n\})$.

3.2. General Setting. In order to derive optimality conditions for the natural gas cash-out problem (2a)–(2g) where the lower level problem (3a)–(3k) is replaced with (4a)–(4j), we are going to develop a general formulation. The idea is to derive general optimality conditions, and by applying those theoretical results one can easily provide optimality conditions to the original problem of interest.

Let us rewrite problem (2a)–(2g) as an *optimistic* BOCBP such that

$$\begin{aligned} F(x(T), y, q) &\longrightarrow \min_{x, u, y, q} \\ \dot{x}(t) &= \phi(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T] \\ x(0) &= x_0 \\ u(t) &\in \mathcal{U}(t) \quad \text{a.e. } t \in [0, T] \\ G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \\ (y, q) &\in \Psi(x(T)), \end{aligned} \quad (16)$$

where

$$\Psi(x(T)) := \text{Argmin}_{y, q} \{w^\top y \mid (y, q) \in \Gamma(x(T))\}. \quad (17)$$

The problem's data comprise an interval $[0, T]$ for a fixed final time moment $T > 0$, given initial conditions $x_0 \in \mathbb{R}^{n_x}$, functions $F : \mathbb{R}^{n_x} \times \mathbb{R}^{m_y} \times \{0, 1\} \rightarrow \mathbb{R}$, $\phi : [0, T] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_x}$, $G : [0, T] \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_G}$, and a multivalued mapping $\mathcal{U} : [0, T] \rightrightarrows \mathbb{R}^{n_u}$ that is measurable on $[0, T]$. The feasible set of control functions is $U := \{u : [0, T] \rightarrow \mathbb{R}^{n_u} \mid u \in L_{n_u}^1, u(t) \in \mathcal{U}(t) \text{ a.e. } t \in [0, T]\}$. The state trajectory $x \in W_{n_x}^{1,1}$ is an absolutely continuous function that satisfies the differential equation and the initial conditions for some control function $u \in L_{n_u}^1$. Finally, $\Psi : \mathbb{R}^{n_x} \rightrightarrows \mathbb{R}^{m_y} \times \{0, 1\}$ denotes the solution set mapping of the lower level problem,

while $w \in \mathbb{R}^{m_y}$ is a fixed vector and $\Gamma : \mathbb{R}^{n_x} \rightrightarrows \mathbb{R}^{m_y} \times \{0, 1\}$ is a set-valued mapping that will be specified below.

In other words, problem (16) is an optimistic BOCBP with final-state-dependent finite-dimensional lower level. Moreover, the upper level problem boasts a Mayer-type cost function and state constraints. This formulation is notably interesting due to the parametric optimization problem at the lower level, whose parameter is nothing else but the final state of the upper level state variable. Hence, the natural gas cash-out problem specified by (2a)–(2g) fits into the formulation of problem (16).

Taking into account the POC's problem (4a)–(4j) one can define the set Γ as follows:

$$\begin{aligned} \Gamma(x(T)) &:= \{(y, q) \in \mathbb{R}^{m_y} \times \{0, 1\} \mid A_1 x(T) + B_1 y \\ &\quad + C_1 q \leq D_1, A_2 x(T) + B_2 y + C_2 q = D_2\}, \end{aligned} \quad (18)$$

where $A_1 \in \mathbb{R}^{m_g \times n_x}$, $A_2 \in \mathbb{R}^{m_h \times n_x}$, $B_1 \in \mathbb{R}^{m_g \times m_y}$, $B_2 \in \mathbb{R}^{m_h \times m_y}$, $C_1, D_1 \in \mathbb{R}^{m_g}$, and $C_2, D_2 \in \mathbb{R}^{m_h}$ are given matrices. When comparing the original problem (2a)–(2g) and the general formulation in (16), it is not difficult to see that we have $n_x = n_u = n$, $n_G = 2n + 2$, $m_y = n^2 + 2n + 2$, $m_g = 3n^2 + 5n + 2$, and $m_h = n + 1$.

Let us consider the space $W := W_{n_x}^{1,1} \times L_{n_u}^1 \times \mathbb{R}^{m_y} \times \{0, 1\}$. Then the process $(x, u, y, q) \in W$ is said to be *feasible* for (16) if it satisfies the dynamical system, the initial condition, and the pure state constraints, while (y, q) is a global optimal solution for the corresponding lower level problem. Now we introduce the definition of a local minimizer.

Definition 2. Take a feasible process $(x^*, u^*, y^*, q^*) \in W$. Then (x^*, u^*, y^*, q^*) is a local minimizer of problem (16) if there exists $\varepsilon > 0$ such that

$$F(x(T)^*, y^*, q^*) \leq F(x(T), y, q^*) \quad (19)$$

holds provided that $(x, y) \in \mathbb{U}_{W_{n_x}^{1,1} \times \mathbb{R}^{m_y}}^\varepsilon(x^*, y^*)$, and (x, u, y, q^*) is feasible for (16).

Clearly, the optimality conditions for local minimizers automatically hold for strong and hence for global minimizers. By following the approach in Vinter [8], the above definition respects our desire to focus on the state variables of (16) while ignoring somehow the control function's behavior. Additionally, fixing the binary variable q proposes a strategy about how to deal with (16). We compute possible minimizers for the problems that arise when setting $q = 0$ and $q = 1$ and then compare the resulting points. Note that assuming q to be free in the definition of local optimal solutions is another idea we are not going to examine here.

For an arbitrary local minimum point (x^*, u^*, y^*, q^*) of (16), we accept the following assumptions:

(A₁)

- (i) For a fixed $q := q^*$, the function F is continuously differentiable with respect to all their arguments.
- (ii) The function $(t, u) \mapsto \phi(t, x, u)$ is $\mathcal{L} \times \mathcal{B}^{n_u}$ -measurable for each $x \in \mathbb{R}^{n_x}$.

- (iii) There exist a $\mathcal{L} \times \mathcal{B}^{n_u}$ -measurable function $\zeta : \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ and a constant $\delta > 0$ such that $t \mapsto \zeta(u^*(t))$ is integrable and

$$\|\phi(t, x, u) - \phi(t, x', u)\| \leq \zeta \|x - x'\|. \quad (20)$$

For $x, x' \in \mathbb{U}_{\mathbb{R}^{n_x}}^\delta(x^*(t))$, $u \in \mathcal{U}(t)$ a.e. $t \in [0, T]$.

- (iv) The set $\text{graph}(\mathcal{U})$ is $\mathcal{L} \times \mathcal{B}^{n_u}$ measurable.
(v) The component functions $G_1, \dots, G_{n_G} : [0, T] \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ of G are continuously differentiable with respect to all their arguments. Furthermore, for any $t \in [0, T]$ the set $\{\nabla_x G_l(t, x^*(t)) \mid l \in I(t, x^*(t))\}$ is positive linearly independent; therein, for any $s' \in [0, T]$ and $x \in \mathbb{R}^{n_x}$ we define

$$\begin{aligned} I(s', x) &:= \{l \in \{1, \dots, n_G\} \mid G_l(s', x) \\ &= \max\{G_1(s', x); \dots; G_{n_G}(s', x)\}\}. \end{aligned} \quad (21)$$

(A₂)

- (i) If $G_l(0, x_0) = 0$ holds true for some $l \in \{1, \dots, n_G\}$, then there exist constants $\varepsilon_1, \varepsilon_2 > 0$ such that the set $\{\phi(s', x, v) \in \mathbb{R}^{n_x} \mid v \in \mathcal{U}(s')\}$ is convex for any $s' \in [0, \varepsilon_1)$ and any $x \in \mathbb{U}_{\mathbb{R}^{n_x}}^{\varepsilon_2}(x_0)$.
(ii) If $G_l(0, x_0) = 0$ holds true for some $l \in \{1, \dots, n_G\}$, then there exist constants $C, \bar{\delta}, \delta_1, \delta_2 > 0$ and a control $\tilde{u} \in U$ such that for almost every $s', t \in [0, \delta_1)$ and $x \in \mathbb{U}_{\mathbb{R}^{n_x}}^{\delta_2}(x_0)$ the following conditions hold:

$$\begin{aligned} \|\phi(t, x_0, u^*(t))\| &\leq C, \\ \|\phi(t, x_0, \tilde{u}(t))\| &\leq C, \\ \nabla_x G_i(s', x) [\phi(t, x_0, \tilde{u}(t)) - \phi(t, x_0, u^*(t))] &\leq -\bar{\delta}, \\ \forall i \in I(s', x). \end{aligned} \quad (22)$$

The assumptions stated in (A₁) are typical hypotheses in optimal control. Recall that, in assumption (A₁-(v)), $I(s', x)$ represents the set of active inequalities at $(s', x) \in [0, T] \times \mathbb{R}^{n_x}$. On the other hand, the assumptions in (A₂) are essential to ensure a nondegenerated maximum principle. The reader is referred to [9, 10] for a detailed discussion of these constraint qualifications.

It is not difficult to check that assumption (A₁) holds for the natural gas cash-out problem (2a)–(2g) at any feasible point. Finally, this general formulation allows us to deal with a smaller problem whose theoretical properties will be useful for further discussion on the gas cash-out problem. Our approach consists in reformulating problem (16) as a locally equivalent single-level problem. On the other hand, the presence of one binary variable at the lower level problem brings some additional difficulties. As mentioned earlier, we deal with that problem by considering the different cases $q = 0$ and $q = 1$.

3.3. Single-Level Reformulation. In this subsection, we replace problem (16) by an equivalent single-level optimal control problem. One typical approach in bilevel programming in finite-dimensional spaces consists in substituting the follower's problem by its Karush-Kuhn-Tucker conditions. They are necessary and sufficient for optimality as long as the program's functions are continuously differentiable and convex with respect to y for a fixed parameter, and some constraint qualification (such as Slater's condition or Mangasarian-Fromovitz constraint qualification (MFCQ)) is satisfied. Hence, if the lower level problem is differentiable, convex, and a regularity condition holds, the resulting problem would be equivalent to the original one in the sense of global optimal solutions. However, from [11] it is well-known that the surrogate problem may have local optimal solutions that do not correspond to local optimal solutions of the original bilevel programming problem. Similar difficulties may appear when using this approach to tackle problem (16).

In line with the above arguments, this study uses the lower level optimal value function, which allows us to derive a problem completely equivalent to (16). The resulting modified problem usually lacks differentiability and regularity [12]. Because of that, the optimal value approach has been discussed in many publications concerning bilevel programming both in the finite- and in infinite-dimensional case; compare, for instance, [13, 14]. The latter papers develop efficient realizations of the (lower level) optimal value method in optimistic bilevel programming by making use of formulas for the limiting/basic and singular subdifferential derived by Mordukhovich for the marginal/value functions (see Mordukhovich's book [7]).

In application to our case of the BOCP with the lower level represented by a conventional problem of parametric programming, the advanced techniques proposed in [13, 14] could be employed in order to find constructive (explicit) conditions to guarantee the required local Lipschitz continuity of the (lower level) objective function's optimal value, that is, of the functions φ and φ_r , $r = 1, 2$, by means of vanishing their singular subdifferentials. Moreover, necessary optimality conditions for our bilevel optimal control program could be developed through the basic subdifferential of the optimal value functions. However, these prospective targets will be postponed for our future research, since, in this paper, we make use of the concept of *partial calmness* introduced in [15].

Namely, we define the lower level optimal value function $\varphi : \mathbb{R}^{n_x} \rightarrow \overline{\mathbb{R}}$ by means of

$$\begin{aligned} \varphi(x(T)) &:= \begin{cases} +\infty & \text{if } \Gamma(x(T)) = \emptyset \\ \inf_{y,q} \{w^\top y \mid (y, q) \in \Gamma(x(T))\} & \text{if } \Gamma(x(T)) \neq \emptyset \end{cases} \end{aligned} \quad (23)$$

for any $x(T) \in \mathbb{R}^{n_x}$ and deduce

$$\begin{aligned} \Psi(x(T)) &= \{(y, q) \in \Gamma(x(T)) \mid w^\top y - \varphi(x(T)) \leq 0\} \end{aligned} \quad (24)$$

from the definition of φ . Hence, we can use the latter function to replace (16) by the following equivalent single-level optimal control problem:

$$\begin{aligned}
 F(x(T), y, q) &\longrightarrow \min_{x, u, y, q} \\
 \dot{x}(t) &= \phi(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T] \\
 x(0) &= x_0 \\
 u(t) &\in \mathcal{U}(t) \quad \text{a.e. } t \in [0, T] \\
 G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \\
 w^\top y - \varphi(x(T)) &\leq 0 \\
 (y, q) &\in \Gamma(x(T)).
 \end{aligned} \tag{25}$$

There are three issues that need to be especially addressed when using this approach. First, it is well-known that, in general, the value function φ may not be differentiable. Secondly, common constraint qualifications such as MFCQ, the linear independence constraint qualification, or at least their nonsmooth counterparts are too strong to hold at the boundary conditions of (25). Finally, the binary variable q could cause some problems in applying a proper constraint qualification. Hence, we have to expect that a nonsmooth maximum principle via (25) is likely to be degenerated (in certain sense).

In order to cope with the above-mentioned obstacles, we follow the ideas of Ye and Zhu [15] where the authors use a *partial penalization* to reach a proper constraint qualification. This partial penalization is often called *partial calmness condition*, which is frequently used for the *optimal value reformulation* of an optimistic bilevel programming problem (i.e., problem (25) in our case). This constraint qualification turns out to be closely related to the *calmness* of a nonlinear optimization problem at a certain point (cf. [6] for the definition). Calmness for multifunctions (for perturbations of the lower level solution set) might be considered as well. Therefore, one could introduce a multifunction $\Delta : \mathbb{R} \rightrightarrows \mathbb{R}^{n_x} \times \mathbb{R}^{m_y} \times \{0, 1\}$ such that we get

$$\begin{aligned}
 \Delta(\kappa) &:= \{(x(T), y, q) \in \mathbb{R}^{n_x} \times \mathbb{R}^{m_y} \times \{0, 1\} \mid (y, q) \\
 &\in \Gamma(x(T)), w^\top y - \varphi(x(T)) \leq \kappa\}
 \end{aligned} \tag{26}$$

for any point $(x, u, y, q) \in W$ and any $\kappa \in \mathbb{R}$. Observe that $\Delta(\kappa)$ is empty for any $\kappa < 0$ while $\Delta(0) = \text{graph}(\Psi)$ is valid. Results from Henrion and Surowiec [16] suggest that the calmness of Δ at $(0, x^*(T), y^*, q^*)$ for an arbitrary feasible point $(x^*, u^*, y^*, q^*) \in W$ of (16) (referred to as *value function constraint qualification* or VFCQ for short) is more restrictive than the partial calmness condition given in Definition 3. In [17], the authors generalize the aforementioned results from [16] to bilevel programs where the upper level is an optimal control problem. Note that Henrion and Surowiec criticize (cf. [16]) the fact that optimality conditions derived from partial calmness (or VFCQ) only contain first order information. However, this is justified enough since any

subdifferential of the optimal value function is a tool of *first-order* analysis. They recommend representing the lower level problem by means of a generalized equation and consider the calmness of some corresponding perturbation mapping in order to obtain optimality conditions containing second-order information of the lower level as well. Observe that the lower level problem in this paper is purely linear; that is, it makes use of no second-order information and hence it makes little sense to exploit this idea here. That is why we deal with a modification of the original concept of partial calmness, and it will turn out later that our problem of interest (16) possesses this partial calmness property at any of its local optimal solutions. We want to mention here that in [18] the author already made use of partial calmness to derive necessary optimality conditions for general optimal control problems but omitted a detailed discussion on when this condition holds in optimal control (apart from the situation where minimax problems are considered).

In order to define the partial calmness condition for (16), we need to consider a partially perturbed version of problem (25) given below for some real number κ :

$$\begin{aligned}
 F(x(T), y, q) &\longrightarrow \min_{x, u, y, q} \\
 \dot{x}(t) &= \phi(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T] \\
 x(0) &= x_0 \\
 u(t) &\in \mathcal{U}(t) \quad \text{a.e. } t \in [0, T] \\
 G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \\
 w^\top y - \varphi(x(T)) &\leq \kappa \\
 (y, q) &\in \Gamma(x(T)).
 \end{aligned} \tag{27}$$

Note that the feasible set of (27) is empty for any $\kappa < 0$. For the purpose of a space-saving notation we introduce the Banach space $X := W_{n_x}^{1,1} \times \mathbb{R}^{m_y} \times \mathbb{R}$.

Definition 3. Problem (25) is called *partially calm* at one of its local minimizers $(x^*, u^*, y^*, q^*) \in W$ if there exist scalars $\varepsilon > 0$ and $\eta > 0$ such that for any feasible point $(x, u, y, q^*, \kappa) \in W \times \mathbb{R}$ of (27) satisfying the condition $(x, y, \kappa) \in \mathbb{U}_X^\varepsilon(x^*, y^*, 0)$ the following inequality holds:

$$F(x(T), y, q^*) - F(x^*(T), y^*, q^*) + \eta\kappa \geq 0. \tag{28}$$

Adapting a result from [9] it becomes clear that partial calmness can be characterized equivalently via partial penalization provided the optimal value function φ is locally Lipschitz continuous in a certain neighborhood of the reference point.

Proposition 4. Let $(x^*, u^*, y^*, q^*) \in W$ be a local minimizer of (25) such that φ is locally Lipschitz continuous at $x^*(T)$. Then (25) is *partially calm* at (x^*, u^*, y^*, q^*) if and only if

there exists a scalar $\eta > 0$ such that (x^*, u^*, y^*, q^*) is a local minimizer of

$$\begin{aligned} \eta(w^\top y - \varphi(x(T))) + F(x(T), y, q) &\longrightarrow \min_{x, u, y, q} \\ \dot{x}(t) &= \phi(t, x(t), u(t)) \quad a.e. \ t \in [0, T] \\ x(0) &= x_0 \\ u(t) &\in \mathcal{U}(t) \quad a.e. \ t \in [0, T] \\ G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \\ (y, q) &\in \Gamma(x(T)). \end{aligned} \quad (29)$$

In order to apply the above result, we have to find a criterion which ensures the postulated Lipschitz-property of φ . Hence, we start a deeper analysis of the lower level problem by decomposing the lower level problem with respect to q . Let us introduce two additional set-valued mappings $\Gamma_r : \mathbb{R}^{n_x} \rightrightarrows \mathbb{R}^{m_y}$ and functions $\varphi_r : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$, $r = 0, 1$, such that, for all $x(T) \in \mathbb{R}^{n_x}$, we get

$$\begin{aligned} \Gamma_r(x(T)) &:= \{y \in \mathbb{R}^{m_y} \mid A_1 x(T) + B_1 y + C_1 r \\ &\leq D_1, A_2 x(T) + B_2 y + C_2 r = D_2\}, \\ \varphi_r(x(T)) &:= \begin{cases} +\infty & \text{if } \Gamma_r(x(T)) = \emptyset; \\ \inf_y \{w^\top y \mid y \in \Gamma_r(x(T))\} & \text{if } \Gamma_r(x(T)) \neq \emptyset. \end{cases} \end{aligned} \quad (30)$$

The following lemma summarizes our results on the continuity of Γ_0 , Γ_1 , and Γ .

Lemma 5. For any $x(T) \in \mathbb{R}^{n_x}$, the following statements hold true:

- (i) $\Gamma(x(T)) = (\Gamma_0(x(T)) \times \{0\}) \cup (\Gamma_1(x(T)) \times \{1\})$.
- (ii) The mapping Γ is polyhedral, closed, and locally upper Lipschitz continuous, as well as upper semicontinuous at any $x^*(T) \in \text{dom}(\Gamma)$. Furthermore, it is locally Lipschitz continuous and lower semicontinuous at any $x^*(T) \in \text{int}(\text{dom}(\Gamma_0) \cap \text{dom}(\Gamma_1))$.

Proof. (i) This relationship follows directly from the definitions of Γ , Γ_0 , and Γ_1 .

(ii) From (i) we deduce

$$\text{graph}(\Gamma) = (\text{graph}(\Gamma_0) \times \{0\}) \cup (\text{graph}(\Gamma_1) \times \{1\}). \quad (31)$$

Since $\text{graph}(\Gamma_0)$ and $\text{graph}(\Gamma_1)$ are polyhedra by the definition, $\text{graph}(\Gamma)$ is a polyhedral set; that is, Γ is a polyhedral mapping. Hence, it is locally upper Lipschitz continuous at any point $x^*(T) \in \text{dom}(\Gamma)$ (cf. [19]). It follows from the definition of upper Lipschitz continuity that any mapping boasting this property is upper semicontinuous. The closedness of Γ follows from its being polyhedral. Now fix $x^*(T) \in \text{int}(\text{dom}(\Gamma_0) \cap \text{dom}(\Gamma_1))$. Hence, Γ_0 and Γ_1 are locally Lipschitz

continuous at $x^*(T)$; that is, we can find scalars $L_r, \varepsilon_r > 0$ such that for all $x(T), x'(T) \in \mathbb{U}_{\mathbb{R}^{n_x}}^{\varepsilon_r}(x^*(T))$ one yields

$$\Gamma_r(x(T)) \subseteq \Gamma_r(x'(T)) + L_r \|x(T) - x'(T)\| \mathbb{B}_{\mathbb{R}^{m_y}}, \quad (32)$$

for $r = 0, 1$. Now we apply Lemma 1 twice. First, we note that for all $x(T), x'(T) \in \mathbb{U}_{\mathbb{R}^{n_x}}^{\varepsilon_r}(x^*(T))$

$$\begin{aligned} \Gamma_r(x(T)) \times \{r\} \\ \subseteq (\Gamma_r(x'(T)) \times \{r\}) \\ + L_r \|x(T) - x'(T)\| (\mathbb{B}_{\mathbb{R}^{m_y}} \times \{0\}), \end{aligned} \quad (33)$$

for $r = 0, 1$. Next, by defining $\varepsilon := \min\{\varepsilon_0; \varepsilon_1\}$ and $L := \max\{L_0; L_1\}$ we derive for all $x(T), x'(T) \in \mathbb{U}_{\mathbb{R}^{n_x}}^{\varepsilon}(x^*(T))$ the following inclusion:

$$\Gamma(x(T)) \subseteq \Gamma(x'(T)) + L \|x(T) - x'(T)\| \mathbb{B}_{\mathbb{R}^{m_y+1}}, \quad (34)$$

from statement (i) and Lemma 1. The above formula implies the local Lipschitz continuity of Γ at $x^*(T)$. Finally, it follows from the definition of local Lipschitz continuity that Γ is lower semicontinuous at $x^*(T)$ as well. Thus, the proof is complete. $\square$

Lemma 6. (i) For $r = 0, 1$, the function φ_r is convex, piecewise affine, and locally Lipschitz continuous on its domain.

(ii) The optimal value function can be represented by the formula $\varphi(x(T)) = \min\{\varphi_0(x(T)); \varphi_1(x(T))\}$ for any point $x(T) \in \mathbb{R}^{n_x}$.

(iii) The function φ is piecewise affine and lower semicontinuous on $\text{dom}(\varphi)$. If $x^*(T) \in \text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1))$ holds true, that φ is locally Lipschitz continuous at $x^*(T)$.

(iv) For $r = 0, 1$, the function φ_r is subdifferentiable in the classical sense (i.e., its subdifferential is nonempty) at any point $x^*(T) \in \text{dom}(\varphi_r)$, and the following formula holds:

$$\begin{aligned} \partial \varphi_r(x^*(T)) &= \{A_1^\top \lambda + A_2^\top \nu \in \mathbb{R}^{n_x} \mid \exists y^* \\ &\in \Gamma_r(x^*(T)) : w + B_1^\top \lambda + B_2^\top \nu = 0^{m_y}, (\lambda, \nu) \\ &\in \mathbb{R}_0^{m_{g^+}} \times \mathbb{R}^{m_h}, (A_1 x^*(T) + B_1 y^* + C_1 r - D_1)^\top \lambda \\ &= 0\}. \end{aligned} \quad (35)$$

(v) For $r = 0, 1$, the following criterion can be used to check whether $x^*(T) \in \text{int}(\text{dom}(\varphi_r))$. Suppose that $x^*(T) \in \text{dom}(\varphi_r)$ and $y^* \in \Gamma_r(x^*(T))$ satisfy $\varphi_r(x^*(T)) = w^\top y^*$. If the set

$$\begin{aligned} \{A_1^\top \lambda + A_2^\top \nu \in \mathbb{R}^{n_x} \mid B_1^\top \lambda + B_2^\top \nu = 0^{m_y}, (\lambda, \nu) \in \mathbb{R}_0^{m_{g^+}} \\ \times \mathbb{R}^{m_h}, (A_1 x^*(T) + B_1 y^* + C_1 r - D_1)^\top \lambda = 0\} \end{aligned} \quad (36)$$

only contains 0^n , then $x^*(T) \in \text{int}(\text{dom}(\varphi_r))$ holds true.

(vi) The function φ is Clarke subdifferentiable everywhere on its domain (i.e., its Clarke subdifferential is nonempty at the point $x^*(T) \in \text{dom}(\varphi)$). Introducing the set of active indices

of the optimal value function $I(x^*(T)) := \{r \in \{0, 1\} \mid \varphi(x^*(T)) = \varphi_r(x^*(T))\}$, the following formulas are correct:

(a) If $I(x^*(T)) = \{r\}$, then

$$\partial^c \varphi(x^*(T)) = \partial \varphi_r(x^*(T)). \quad (37)$$

(b) If $I(x^*(T)) = \{r, 1-r\}$, then

$$\partial^c \varphi(x^*(T)) \subseteq \text{conv}(\partial \varphi_r(x^*(T)) \cup \partial \varphi_{1-r}(x^*(T))). \quad (38)$$

Proof. (i) This statement is well-known from linear parametric optimization (cf. [20]).

(ii) We deduce this representation directly from statement (i) of Lemma 5:

$$\begin{aligned} \varphi(x(T)) &= \inf_{y,q} \{w^\top y \mid (y,q) \in \Gamma(x(T))\} \\ &= \inf_{y,q} \{w^\top y \mid (y,q) \in (\Gamma_0(x(T)) \times \{0\}) \cup (\Gamma_1(x(T)) \times \{1\})\} \\ &= \min \left\{ \inf_{y,q} \{w^\top y \mid (y,q) \in \Gamma_0(x(T)) \times \{0\}\}; \right. \\ &\quad \left. \inf_{y,q} \{w^\top y \mid (y,q) \in \Gamma_1(x(T)) \times \{1\}\} \right\} \\ &= \min \left\{ \inf_y \{w^\top y \mid y \in \Gamma_0(x(T))\}; \right. \\ &\quad \left. \inf_y \{w^\top y \mid y \in \Gamma_1(x(T))\} \right\} = \min \{\varphi_0(x(T)); \\ &\quad \varphi_1(x(T))\}. \end{aligned} \quad (39)$$

(iii) We already know from statement (i) that φ_0 and φ_1 are piecewise affine and lower semicontinuous functions, being also locally Lipschitz continuous on the interior of their domains. Due to the fact that φ is the pointwise minimum of these two functions (cf. statement (ii) of this lemma), it is piecewise affine as well.

Now, from assertion (ii) of Lemma 5 we know that Γ is upper semicontinuous, which yields that φ is a lower semicontinuous function (cf. [20]). On the other hand, at any point $x^*(T) \in \text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1))$ the functions φ_0 and φ_1 are locally Lipschitz continuous, so the same holds true for φ .

(iv) One may apply the results from [21] for fully convex parametric optimization problems to the given linear optimization problems using the duality of linear programming. Note that the classical subdifferential is *not bounded* at the boundary points of the domain of the optimal value function.

(v) For any fixed $r \in \{0, 1\}$, the parametric optimization problem

$$\min_y \{w^\top y \mid y \in \Gamma_r(x^*(T))\} \quad (40)$$

is completely linear and hence its solution set mapping is inner semicontinuous, while the singular subdifferential of its

objective function reduces to $\{0^{n+m}\}$ (cf. [22]). Note that y^* is a solution of (40). We apply Theorem 7 [5] to derive

$$\partial^\infty \varphi_r(x^*(T)) \subseteq D^* \Gamma_r(x^*(T), y^*)(0^m). \quad (41)$$

In order to evaluate the above coderivative, we have to compute the basic normal cone

$$\mathcal{N}_{\text{graph}(\Gamma_r)}^B(x^*(T), y^*). \quad (42)$$

Observe that $\text{graph}(\Gamma_r)$ is a polyhedron; that is, it is closed and convex. Hence, this cone coincides with the corresponding (classical) normal cone defined as

$$\begin{aligned} \mathcal{N}_{\text{graph}(\Gamma_r)}(x^*(T), y^*) &= \left\{ \begin{pmatrix} A_1^\top \lambda + A_2^\top \nu \\ B_1^\top \lambda + B_2^\top \nu \end{pmatrix} \right. \\ &\in \mathbb{R}^{n_x+m_y} \mid (\lambda, \nu) \in \mathbb{R}_0^{m_g,+} \\ &\times \mathbb{R}^{m_h}, (A_1 x^*(T) + B_1 y^* + C_1 r - D_1)^\top \lambda = 0 \left. \right\}, \end{aligned} \quad (43)$$

since at least Abadie Constraint Qualification holds at any feasible point of a purely linear optimization problem. It follows from (41) and the definition of the normal coderivative $D^* \Gamma_r(x^*(T), y)$ that the inclusion

$$\begin{aligned} \partial^\infty \varphi_r(x^*(T)) &\subseteq \{A_1^\top \lambda + A_2^\top \nu \in \mathbb{R}^{n_x} \mid B_1^\top \lambda + B_2^\top \nu \\ &= 0^{m_y}, (\lambda, \nu) \in \mathbb{R}_0^{m_g,+} \\ &\times \mathbb{R}^{m_h}, (A_1 x^*(T) + B_1 y^* + C_1 r - D_1)^\top \lambda = 0\} \end{aligned} \quad (44)$$

is true; that is, the criterion from statement (v) implies $\partial^\infty \varphi_r(x^*(T)) = \{0^n\}$. Consequently, φ_r is Lipschitz continuous in a neighborhood of $x^*(T)$, which is only possible if $x^*(T) \in \text{int}(\text{dom}(\varphi_r))$.

(vi) These formulas follow from the calculation rules of the Clarke subdifferential stated in [6]. $\square$

Note that the criterion from statement (v) of the above lemma always holds if, for a fixed $x^*(T)$, the linear problem (40) is solvable, while $\Gamma(x^*(T))$ possesses an interior point. In order to make this clear, recall that, in the latter case, the Slater condition holds true for problem (40) and the set of singular Lagrange multipliers

$$\begin{aligned} &\{(\lambda, \nu) \in \mathbb{R}_0^{m_g,+} \times \mathbb{R}^{m_h} \mid B_1^\top \lambda + B_2^\top \nu \\ &= 0^{m_y}, (A_1 x^*(T) + B_1 y^* + C_1 r - D_1)^\top \lambda = 0\} \end{aligned} \quad (45)$$

reduces to $\{(0^{m_g}, 0^{m_h})\}$. In general, one may use statement (v) as a criterion to check whether $x^*(T)$ comes from $\text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1)) = \text{int}(\text{dom}(\varphi_0)) \cap \text{int}(\text{dom}(\varphi_1))$, that is, whether φ is locally Lipschitz continuous in a neighborhood of $x^*(T)$. This condition will play an important role in our main results, that is, Theorems 7, 8, and 10.

In order to present our next result, we need to introduce two surrogate single-level optimal control problems that arise

from (25) and its partially perturbed version (27) by fixing the binary variable value q either to 0 or to 1. Hence, for any $r \in \{0, 1\}$ we define

$$\begin{aligned} F(x(T), y, r) &\longrightarrow \min_{x, u, y} \\ \dot{x}(t) &= \phi(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T] \\ x(0) &= x_0 \\ u(t) &\in \mathcal{U}(t) \quad \text{a.e. } t \in [0, T] \\ G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \end{aligned} \quad (46)$$

$$\begin{aligned} w^\top y - \varphi_r(x(T)) &\leq 0 \\ y &\in \Gamma_r(x(T)), \\ F(x(T), y, r) &\longrightarrow \min_{x, u, y} \\ \dot{x}(t) &= \phi(t, x(t), u(t)) \quad \text{a.e. } t \in [0, T] \\ x(0) &= x_0 \\ u(t) &\in \mathcal{U}(t) \quad \text{a.e. } t \in [0, T] \\ G(t, x(t)) &\leq 0^{n_G} \quad \forall t \in [0, T] \\ w^\top y - \varphi_r(x(T)) &\leq \kappa \\ y &\in \Gamma_r(x(T)). \end{aligned} \quad (47)$$

The following theorem reveals that (16) is *always* partially calm in the sense of Definition 3.

Theorem 7. *Let $(x^*, u^*, y^*, q^*) \in W$ be a feasible point of (16) satisfying $x^*(T) \in \text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1))$. Then (16) is partially calm at this point. In particular, if (x^*, u^*, y^*, q^*) is a local minimizer of (16), then there exists an $\eta > 0$ such that (x^*, u^*, y^*, q^*) is a local minimizer of (29), too.*

Proof. Denote by $W' := W_{n_x}^{1,1} \times L_{n_u}^1 \times \mathbb{R}^{m_y}$. Since (x^*, u^*, y^*, q^*) is a feasible point of (16), (y^*, q^*) is a global optimal solution for the lower level problem with the fixed parameter $x^*(T)$. Hence, we have $\varphi(x^*(T)) = \varphi_{q^*}(x^*(T)) = w^\top y^*$. That is why (x^*, u^*, y^*) is feasible for problem (46) with $r = q^*$. The function φ_{q^*} appearing therein is the optimal value function of a purely linear parametric optimization problem. Therefore, we may apply the results of Section 3 in [9] in order to find an $\varepsilon > 0$ and a scalar $\eta' > 0$ such that, for any feasible point $(x, u, y, \kappa) \in W' \times \mathbb{R}$ of (47) with $(x, y, \kappa) \in \mathbb{U}_X^\varepsilon(x^*, y^*, 0)$ (and $r = q^*$), we have

$$F(x(T), y, q^*) - F(x^*(T), y^*, q^*) + \eta' \kappa \geq 0. \quad (48)$$

Now choose an arbitrary feasible point $(x, u, y, q^*, \kappa) \in W \times \mathbb{R}$ of (27) satisfying $(x, y, \kappa) \in \mathbb{U}_X^\varepsilon(x^*, y^*, 0)$. Then (x, u, y, κ) is

feasible for problem (47) for $r = q^*$ since from Lemma 6 we have

$$\begin{aligned} \kappa &\geq w^\top y - \varphi(x(T)) \\ &= w^\top y - \min\{\varphi_0(x(T)); \varphi_1(x(T))\} \\ &\geq w^\top y - \varphi_{q^*}(x(T)). \end{aligned} \quad (49)$$

For any such point (48) holds true; that is, (16) is partially calm at (x^*, u^*, y^*, q^*) in the sense of Definition 3. The second part of the statement follows now easily taking Proposition 4 and Lemma 6 into account. $\square$

General classes of bilevel programming problems that are partially calm can be found in Ye and Zhu [15], Dempe and Zemkoho [23], and more recently Benita and Mehrlitz [17] and Benita et al. [9]. Furthermore, any minimax problem is partially calm at its local optimal solutions. Apart from these classes, there are two well-known conditions ensuring the partial calmness, namely, (i) the property of the lower level problem to have a so-called *uniformly weak sharp minimum* (cf. [15]) and (ii) the above-mentioned condition VFCQ. In [9, 17], it is asserted that a BOCIP with a finite-dimensional linear lower level problem is always partially calm since its lower level boasts a uniformly weak sharp minimum. Recall that we implicitly used this result when proving Theorem 7.

3.4. Optimality Conditions. In this subsection, we present an appropriate formulation of the maximum principle for problem (16). Since, in our notion of local optimality, we consider the binary variable to be fixed, it is possible to proceed as in [9]. We postulate assumptions that allow us to transform (16) into the single-level nonsmooth optimal control problem (29) by means of Theorem 7. Afterwards, we apply a maximum principle (cf. [8, 9]) to (29) and exploit the subdifferential formulas for φ we introduced in Lemma 6. We omit the proof here and refer the interested reader to the recent paper [9].

Theorem 8. *Let $(x^*, u^*, y^*, q^*) \in W$ be a local minimizer of (16), assume that (A_1) – (A_2) hold, and let $x^*(T) \in \text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1))$ be satisfied.*

(a) *Suppose that $I(x^*(T)) = \{q^*\}$ holds (recall Lemma 6 for the definition of this set). Then there exist a function $\theta \in W_{n_x}^{1,1}$, vectors $\lambda_1, \lambda_2 \in \mathbb{R}_0^{m_g, +}$, $\nu_1, \nu_2 \in \mathbb{R}^{m_h}$, measures $\mu_1, \dots, \mu_{n_G} \in C^\oplus$, and scalars $\tau \geq 0, \eta > 0$ such that we have (for almost every $t \in [0, T]$) the following:*

(i) *The nontriviality condition is*

$$\tau + \|p\|_{L_{n_u}^\infty} + \sum_{l=1}^{n_G} \mu_l([0, T]) > 0. \quad (50)$$

(ii) *The adjoint equation is*

$$-\dot{\theta}(t)^\top = p(t)^\top \nabla_x \phi(t, x^*(t), u^*(t)). \quad (51)$$

(iii) The Weierstrass-Pontryagin condition is

$$\begin{aligned} p(t)^\top \phi(t, x^*(t), u^*(t)) \\ = \max_{v \in \mathcal{U}(t)} p(t)^\top \phi(t, x^*(t), v). \end{aligned} \quad (52)$$

(iv) The transversality condition is

$$\begin{aligned} -p(T) = \tau \cdot \nabla_x F(x^*(T), y^*, q^*)^\top \\ + A_1^\top (\lambda_1 - \tau \eta \cdot \lambda_2) + A_2^\top (\nu_1 - \tau \eta \cdot \nu_2). \end{aligned} \quad (53)$$

(v) The support condition is

$$\begin{aligned} \text{supp}(\mu_l) \subseteq \{t \in [0, T] \mid G_l(t, x^*(t)) = 0\}, \\ l \in \{1, \dots, n_G\}. \end{aligned} \quad (54)$$

(vi) The lower level optimality condition is

$$\begin{aligned} w + B_1^\top \lambda_2 + B_2^\top \nu_2 = 0^{m_y}, \\ (A_1 x^*(T) + B_1 y^* + C_1 q^* - D_1)^\top \lambda_2 = 0. \end{aligned} \quad (55)$$

(vii) The complementarity condition is

$$(A_1 x^*(T) + B_1 y^* + C_1 q^* - D_1)^\top \lambda_1 = 0. \quad (56)$$

(viii) The multiplier condition is

$$\begin{aligned} \tau \cdot (\nabla_y F(x^*(T), y^*, q^*)^\top + \eta \cdot w) + B_1^\top \lambda_1 + B_2^\top \nu_1 \\ = 0^{m_y}. \end{aligned} \quad (57)$$

Here, the function p is defined as follows for almost every $t \in [0, T]$:

$$\begin{aligned} p(t) \\ := \begin{cases} \theta(t) + \int_{[0,t)} \sum_{l=1}^{n_G} \nabla_x G_l(\xi, x^*(\xi))^\top \mu_l(d\xi), & \text{if } t < T; \\ \theta(T) + \int_{[0,T]} \sum_{l=1}^{n_G} \nabla_x G_l(\xi, x^*(\xi))^\top \mu_l(d\xi), & \text{if } t = T. \end{cases} \end{aligned} \quad (58)$$

(b) Suppose that $I(x^*(T)) = \{q^*, 1 - q^*\}$ holds. Then we find a function $\theta \in W_{n_x}^{1,1}$, vectors $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_0^{m_g, +}$, $\nu_1, \nu_2, \nu_3 \in \mathbb{R}^{m_h}$, $\bar{y} \in \mathbb{R}^{m_y}$, measures $\mu_1, \dots, \mu_{n_G} \in C^\oplus$, constants $\tau \geq 0$, $\eta > 0$, and $\alpha \in [0, 1]$ such that conditions (i), (ii), (iii), (v), (vii), and (viii) hold true, together with relationships (ix) and (x) listed below:

(ix) The modified transversality condition

$$\begin{aligned} -p(T) = \tau \cdot \nabla_x F(x^*(T), y^*, q^*)^\top \\ + A_1^\top [\lambda_1 - \tau \eta \cdot (\alpha \lambda_2 + (1 - \alpha) \lambda_3)] \\ + A_2^\top [\nu_1 - \tau \eta \cdot (\alpha \nu_2 + (1 - \alpha) \nu_3)]. \end{aligned} \quad (59)$$

(x) The modified lower level optimality condition

$$\begin{aligned} w + B_1^\top \lambda_2 + B_2^\top \nu_2 = 0^{m_y}, \\ (A_1 x^*(T) + B_1 y^* + C_1 q^* - D_1)^\top \lambda_2 = 0, \\ w + B_1^\top \lambda_3 + B_2^\top \nu_3 = 0^{m_y}, \\ (A_1 x^*(T) + B_1 \bar{y} + C_1 (1 - q^*) - D_2)^\top \lambda_3 = 0, \\ A_1 x^*(T) + B_1 \bar{y} + C_1 (1 - q^*) \leq D_1, \\ A_2 x^*(T) + B_2 \bar{y} + C_2 (1 - q^*) = D_2. \end{aligned} \quad (60)$$

Here, the function p is defined as given in (58).

Now we illustrate the above theory by means of an example. Observe that it is an adaptation of the BOCP that the second and third authors already discussed in [9]. Here we modify the upper level objective function and the lower level problem in order to fit the settings of this paper.

Example 9. Consider the bilevel optimal control problem

$$\begin{aligned} x_2(3) + \frac{1}{2} (y_1 - 2)^2 + \frac{1}{2} (y_2 + 1)^2 \\ \rightarrow \min_{x, u, y, q} \\ \text{subject to } \dot{x}_1(t) = u(t) \quad \text{a.e. } t \in [0, 3], \\ \dot{x}_2(t) = x_1(t) \quad \text{a.e. } t \in [0, 3], \\ x_1(0) = 1, \\ x_2(0) = 0, \\ u(t) \in [-1, 1] \quad \text{a.e. } t \in [0, 3], \\ -x_1(t) \leq 0 \quad \forall t \in [0, 3], \\ (y, q) \in \Psi(x(3)), \end{aligned} \quad (61)$$

where

$$\begin{aligned} \Psi(x(3)) := \text{Argmin}_{y, q} \{-y_1 - 2y_2 \mid -x_1(3) + y_1 + y_2 \\ \leq 1, y_1 + 3q \leq 3, y_2 - q \leq 0, -y_1 \leq 0, -y_2 \leq 0, q \\ \in \{0, 1\}\}. \end{aligned} \quad (62)$$

A similar argumentation as presented in [9] reveals that its global minimizer (x^*, u^*, y^*, q^*) is given as stated below (for almost every $t \in [0, 3]$):

$$x_1^*(t) = \begin{cases} 1 - t, & \text{if } t \in [0, 1]; \\ 0, & \text{if } t \in [1, 2]; \\ t - 2, & \text{if } t \in [2, 3], \end{cases}$$

$$\begin{aligned}
x_2^*(t) &= \begin{cases} \frac{1}{2} - \frac{1}{2}(t-1)^2, & \text{if } t \in [0, 1]; \\ \frac{1}{2}, & \text{if } t \in [1, 2]; \\ \frac{1}{2} + \frac{1}{2}(t-2)^2, & \text{if } t \in [2, 3], \end{cases} \\
u^*(t) &= \begin{cases} -1, & \text{if } t \in [0, 1]; \\ 0, & \text{if } t \in [1, 2]; \\ 1, & \text{if } t \in [2, 3], \end{cases} \\
y_1^* &= 2, \\
y_2^* &= 0, \\
q^* &= 0.
\end{aligned} \tag{63}$$

It is easy to see that the assumptions of Theorem 8 hold for problem (61). Now we demonstrate that the corresponding necessary optimality conditions are satisfied.

Observe that the optimal value functions $\varphi_0, \varphi_1 : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ of the lower level problem (62) take the following values for an arbitrary (terminal) point $x(3) \in \mathbb{R}^2$:

$$\begin{aligned}
\varphi_0(x(3)) &= \begin{cases} +\infty, & \text{if } x_1(3) \in (-\infty, -1); \\ -x_1(3) - 1, & \text{if } x_1(3) \in [-1, 2]; \\ -3, & \text{if } x_1(3) \in (2, +\infty), \end{cases} \\
\varphi_1(x(3)) &= \begin{cases} +\infty, & \text{if } x_1(3) \in (-\infty, -1); \\ -2x_1(3) - 2, & \text{if } x_1(3) \in [-1, 0]; \\ -2, & \text{if } x_1(3) \in (0, +\infty). \end{cases}
\end{aligned} \tag{64}$$

Then $x^*(3)$ is an interior point of $\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1)$, and $I(x^*(3)) = \{0, 1\}$ holds true. Due to Theorem 8, there exist a function $\theta \in W_2^{1,1}$, vectors $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}_0^{5,+}$, $\bar{y} \in \mathbb{R}^2$, a measure $\mu \in C^\oplus$, constants $\tau \geq 0$, $\eta > 0$, and $\alpha \in [0, 1]$

which satisfy the following set of conditions (for almost every $t \in [0, 3]$):

(i) The nontriviality condition is

$$\tau + \|p\|_{L^\infty} + \mu((0, 3]) > 0. \tag{65}$$

(ii) The adjoint equation is

$$\begin{aligned}
-\dot{\theta}_1(t) &= p_2(t), \\
-\dot{\theta}_2(t) &= 0.
\end{aligned} \tag{66}$$

(iii) The Weierstrass-Pontryagin Condition is

$$\begin{aligned}
p_1(t) \cdot u^*(t) + p_2(t) \cdot x_1^*(t) \\
= \max_{v \in [-1, 1]} \{p_1(t) \cdot v + p_2(t) \cdot x_1^*(t)\}.
\end{aligned} \tag{67}$$

(v) The support condition is

$$\text{supp}(\mu) \subseteq [1, 2]. \tag{68}$$

(vii) The complementarity condition is

$$\begin{aligned}
\lambda_{1,2} &= 0, \\
\lambda_{1,4} &= 0.
\end{aligned} \tag{69}$$

(viii) The multiplier condition is

$$\begin{aligned}
-\tau\eta + \lambda_{1,1} + \lambda_{1,2} - \lambda_{1,4} &= 0, \\
\tau - 2\tau\eta + \lambda_{1,1} + \lambda_{1,3} - \lambda_{1,5} &= 0.
\end{aligned} \tag{70}$$

(x) The modified transversality condition is

$$\begin{aligned}
-p_1(3) &= -[\lambda_{1,1} - \tau\eta(\alpha\lambda_{2,1} + (1-\alpha)\lambda_{3,1})], \\
-p_2(3) &= \tau.
\end{aligned} \tag{71}$$

(xi) The modified lower level optimality condition is

$$\begin{aligned}
-1 + \lambda_{2,1} + \lambda_{2,2} - \lambda_{2,4} &= 0, & \lambda_{2,2} &= 0, & \lambda_{2,4} &= 0, \\
-2 + \lambda_{2,1} + \lambda_{2,3} - \lambda_{2,5} &= 0, \\
-1 + \lambda_{3,1} + \lambda_{3,2} - \lambda_{3,4} &= 0, & \lambda_{3,1} &= 0, & \lambda_{3,5} &= 0, \\
-2 + \lambda_{3,1} + \lambda_{3,3} - \lambda_{3,5} &= 0, & \bar{y}_1 &= 0, & \bar{y}_2 &= 1.
\end{aligned} \tag{72}$$

Here, the function $p \in L_2^\infty$ is defined as follows for almost every $t \in [0, 3]$:

$$\begin{aligned}
p_1(t) &= \begin{cases} \theta_1(t) - \int_{[0,t)} \mu(d\xi), & \text{if } t < 3; \\ \theta_1(3) - \int_{[0,3]} \mu(d\xi), & \text{if } t = 3, \end{cases} \\
p_2(t) &= \theta_2(t).
\end{aligned} \tag{73}$$

Note that we already computed the lower level solution $\bar{y}$ for the fixed $x^*(3)$ and $1 - q^*$ in (xi) to make the latter condition more convenient. As a consequence, (xi) yields $\lambda_{2,1} = 1$, $\lambda_{2,3} = \lambda_{2,5} + 1$, $\lambda_{3,3} = 2$, and $\lambda_{3,2} = \lambda_{3,4} + 1$. Taking (vi) and (vii) together we obtain $\lambda_{1,1} = \tau\eta$ and $\lambda_{1,3} = \lambda_{1,5} + \tau(\eta - 1)$. Substituting these representations into (x) leads to $p_1(3) = \tau\eta(1 - \alpha)$ and $p_2(3) = -\tau$.

Taking into account (ii) and the definition of p , we derive $\theta_2 \equiv -\tau$, $p_2 \equiv -\tau$, and $\theta_1(t) = \tau \cdot t + c$ for some constant $c \in \mathbb{R}$ and almost every $t \in [0, 3]$. We evaluate (iii) to obtain

$$\begin{aligned} t \in [0, 1] &\implies p_1(t) \leq 0, \\ t \in [1, 2] &\implies p_1(t) = 0, \\ t \in [2, 3] &\implies p_1(t) \geq 0. \end{aligned} \quad (74)$$

In order to satisfy the latter condition, as well as the definition of p , we choose $c := -\tau$ and $\mu := \tau \cdot (\mathbb{I} \circ \mathcal{J}_{[1,2]})$, where $\mathbb{I}$ denotes the Lebesgue measure and $\mathcal{J}_{[1,2]}$ represents the intersection mapping with the interval $[1, 2]$; that is, $\mathcal{J}_{[1,2]}(M) = M \cap [1, 2]$ for all $M \subseteq \mathbb{R}$. This leads to the following representation of θ_1 and p_1 for almost every $t \in [0, 3]$:

$$\begin{aligned} \theta_1(t) &= \tau \cdot t - \tau, \\ p_1(t) &= \begin{cases} \tau \cdot t - \tau, & \text{if } t \in [0, 1]; \\ 0, & \text{if } t \in [1, 2]; \\ \tau \cdot t - 2\tau, & \text{if } t \in [2, 3]. \end{cases} \end{aligned} \quad (75)$$

In order to satisfy the terminal condition $p_1(3) = \tau\eta(1 - \alpha)$ we derived earlier, we finally fix $\alpha := 1 - 1/\eta$ for an arbitrary $\eta \geq 1$ to guarantee $\alpha \in [0, 1]$.

Resuming the above calculations while taking into account (i), the following multipliers satisfy the necessary optimality conditions postulated in Theorem 8 (definitions hold for almost every $t \in [0, 3]$ when necessary) for any choice of $\eta \geq 1$:

$$\begin{aligned} \theta_1(t) &= \tau \cdot t - \tau, \\ \theta_2(t) &= -\tau, \\ \lambda_1 &= \begin{pmatrix} \tau\eta \\ 0 \\ \tau(\eta - 1) \\ 0 \\ 0 \end{pmatrix}, \\ \lambda_2 &= \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\ \lambda_3 &= \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \end{aligned}$$

$$\bar{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$\mu = \tau \cdot (\mathbb{I} \circ \mathcal{J}_{[1,2]}),$$

$$\tau > 0,$$

$$\alpha = 1 - \frac{1}{\eta}.$$

(76)

This result underlines that the partial calmness property (that we used to derive these optimality conditions) theoretically is closely related to partial penalization. One may interpret η as penalization parameter and observe that the necessary optimality conditions above hold for any η which is greater than or equal to some lower bound (which is 1 in this example).

4. Optimality Conditions of the Natural Gas Cash-Out Problem

In this section, we provide the optimality conditions for local minimizers of problem (2a)–(2g). We only present a result for the case $I(x^*(T)) = \{q^*\}$ using statement (a) of Theorem 8. The case $I(x^*(T)) = \{q^*, 1 - q^*\}$ can be handled similarly applying assertion (b) of the latter theorem. Recall that we already introduced the necessary notation for the matrix representation of the linear lower level problem (corresponding to the natural gas cash-out problem) in Section 2.1. Now these matrices play the role of the general matrices we considered in Section 3. For the brevity purpose, we introduce functions $g_1^L, \dots, g_n^L, g_1^U, \dots, g_n^U : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ by means of

$$\begin{aligned} g_i^L(t, x) &:= x_i^L(t) - x^\top e_i^n, \quad i \in N; \\ g^L(t, x) &:= x^L(t) - x^\top s^n; \\ g_i^U(t, x) &:= x^\top e_i^n - x_i^U(t), \quad i \in N; \\ g^U(t, x) &:= x^\top s^n - x^U(t), \end{aligned} \quad (77)$$

for any $t \in [0, T]$ and $x \in \mathbb{R}^n$.

Theorem 10. Let $(x^*, u^*, y^*, q^*) \in W_n^{1,1} \times L_n^1 \times \mathbb{R}^{n^2+2n+2} \times \{0, 1\}$ be a local minimizer of the natural gas cash-out problem (2a)–(2g) such that assumption (A_2) -(ii) is satisfied (note that (A_2) -(i) holds trivially) while $x^*(T) \in \text{int}(\text{dom}(\varphi_0) \cap \text{dom}(\varphi_1))$ and $I(x^*(T)) = \{q^*\}$. Then there exist a function $\theta \in W_n^{1,1}$, vectors $\lambda_1, \lambda_2 \in \mathbb{R}^{3n^2+5n+2,+}$, $v_1, v_2 \in \mathbb{R}^{n+1}$, measures $\mu_1^L, \dots, \mu_n^L, \mu_1^U, \dots, \mu_n^U \in C^\oplus$, and scalars $\tau \geq 0, \eta > 0$ such that one has (for almost every $t \in [0, T]$) the following:

(i) The nontriviality condition is

$$\begin{aligned} \tau + \|p\|_{L_n^\infty} + \sum_{i=1}^n \mu_i^L((0, T]) + \mu^L((0, T]) \\ + \sum_{i=1}^n \mu_i^U((0, T]) + \mu^U((0, T]) > 0. \end{aligned} \quad (78)$$

(ii) The adjoint equation is

$$\dot{\theta}(t) = 0^n. \quad (79)$$

(iii) The Weierstrass-Pontryagin Condition is

$$\begin{aligned} u_i^*(t) = u_i^L(t) &\implies p_i(t) \leq 0, \quad i \in N; \\ u_i^*(t) = u_i^U(t) &\implies p_i(t) \geq 0, \quad i \in N; \\ u_i^L(t) < u_i^*(t) < u_i^U(t) &\implies p_i(t) = 0, \quad i \in N. \end{aligned} \quad (80)$$

(iv) The transversality condition is

$$-p(T) = A_1^\top (\lambda_1 - \tau\eta \cdot \lambda_2) + A_2^\top (\nu_1 - \tau\eta \cdot \nu_2). \quad (81)$$

(v) The support condition is

$$\begin{aligned} \text{supp}(\mu_i^L) &\subseteq \{t \in [0, T] \mid g_i^L(t, x^*(t)) = 0\}, \quad i \in N; \\ \text{supp}(\mu^L) &\subseteq \{t \in [0, T] \mid g^L(t, x^*(t)) = 0\}; \end{aligned}$$

$$\begin{aligned} \text{supp}(\mu_i^U) &\subseteq \{t \in [0, T] \mid g_i^U(t, x^*(t)) = 0\}, \\ &i \in N; \end{aligned}$$

$$\text{supp}(\mu^U) \subseteq \{t \in [0, T] \mid g^U(t, x^*(t)) = 0\}. \quad (82)$$

(vi) The lower level optimality condition is

$$\begin{aligned} w + B_1^\top \lambda_2 + B_2^\top \nu_2 &= 0^{m_y}, \\ (A_1 x^*(T) + B_1 y^* + C_1 q^* - D_1)^\top \lambda_2 &= 0. \end{aligned} \quad (83)$$

(vii) The complementarity condition is

$$(A_1 x^*(T) + B_1 y^* + C_1 q^* - D_1)^\top \lambda_1 = 0. \quad (84)$$

(viii) The multiplier condition is

$$\tau \cdot (e_1^{n^2+2n+2} + \eta \cdot w) + B_1^\top \lambda_1 + B_2^\top \nu_1 = 0^{n^2+2n+2}. \quad (85)$$

Here, the function p is defined as follows for almost every $t \in [0, T]$:

$$p(t) := \begin{cases} \theta(t) + \sum_{i=1}^n [\mu_i^U([0, t]) - \mu_i^L([0, t])] \cdot e_i^n + [\mu^U([0, t]) - \mu^L([0, t])] \cdot s^n, & \text{if } t < T; \\ \theta(T) + \sum_{i=1}^n [\mu_i^U([0, T]) - \mu_i^L([0, T])] \cdot e_i^n + [\mu^U([0, T]) - \mu^L([0, T])] \cdot s^n, & \text{if } t = T. \end{cases} \quad (86)$$

5. Concluding Remarks

In this paper, we formulate the well-known natural gas cash-out model as a *bilevel optimal control problem*. The new formulation yields an interesting theoretical problem with three difficulties. First, its single-level reformulation implies that one has to guarantee that a local optimal solution of the single-level problem corresponds to a local optimal solution of the bilevel problem. Second, one has to find an approach to compute a general derivative of the lower level value function, and, third, the pure state constraints at the upper level and the binary variable at the lower level imply that common regularity conditions do not hold. An important contribution is that we develop a general maximum principle, which is likely to be nondegenerated in order to find the optimality conditions to the cash-out problem.

The natural targets for the future research include (i) deriving a theory where the binary variable need not be fixed, that is, where we do not have to evaluate the necessary optimality conditions twice, and (ii) implementing proper discretization techniques with the intention to find numerical solutions to the natural gas cash-out problem.

Another task of our future research is to apply to our case (of the BOCP with the lower level represented by a conventional problem of parametric programming) the advanced

techniques proposed in [13, 14]. The latter can be employed in order to find constructive (explicit) conditions to guarantee the required local Lipschitz continuity of the (lower level) objective function's optimal value, that is, of the functions φ and φ_r , $r = 1, 2$, by means of vanishing their singular subdifferentials. Moreover, necessary optimality conditions for our bilevel optimal control program will be developed through the basic subdifferential of the optimal value functions, similar to that made in the above-mentioned works by Dempe, Morduchovich, Dutta, and their coauthors.

Conflict of Interests

The authors declare that there is no conflict of interests for any of the authors of the paper.

Acknowledgments

The research activity of the first two authors was funded by the R&D Department of the Tecnológico de Monterrey (ITESM), Campus Monterrey, and by the SEP-CONACYT Project CB-2013-01-221676 (Mexico). The authors would also like to express their profound gratitude to the anonymous referee whose valuable comments and recommendations have helped them greatly improve the quality of the paper.

References

- [1] S. Dempe, V. V. Kalashnikov, and R. Z. Ríos-Mercado, "Discrete bilevel programming: application to a natural gas cash-out problem," *European Journal of Operational Research*, vol. 166, no. 2, pp. 469–488, 2005.
- [2] V. V. Kalashnikov, G. A. Pérez-Valdés, and N. I. Kalashnykova, "A linearization approach to solve the natural gas cash-out bilevel problem," *Annals of Operations Research*, vol. 181, no. 1, pp. 423–442, 2010.
- [3] V. V. Kalashnikov and R. Z. Ríos-Mercado, "A natural gas cash-out problem: a bilevel programming framework and a penalty function method," *Optimization and Engineering*, vol. 7, no. 4, pp. 403–420, 2006.
- [4] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer Science and Business Media, 1998.
- [5] B. S. Mordukhovich, N. M. Nam, and N. D. Yen, "Subgradients of marginal functions in parametric mathematical programming," *Mathematical Programming*, vol. 116, no. 1-2, pp. 369–396, 2009.
- [6] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, NY, USA, 1983.
- [7] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I and II*, Springer, Berlin, Germany, 2006.
- [8] R. Vinter, *Optimal Control*, Birkhäuser, Boston, Mass, USA, 2000.
- [9] F. Benita, S. Dempe, and P. Mehrlitz, *Bilevel Optimal Control Problems with Pure State Constraints and Finite-Dimensional Lower Level*, TU Freiberg, 2015.
- [10] S. O. Lopes, F. A. C. C. Fontes, and M. D. R. de Pinho, "An integral-type constraint qualification to guarantee nondegeneracy of the maximum principle for optimal control problems with state constraints," *Systems & Control Letters*, vol. 62, no. 8, pp. 686–692, 2013.
- [11] S. Dempe and J. Dutta, "Is bilevel programming a special case of a mathematical program with complementarity constraints?" *Mathematical Programming*, vol. 131, no. 1-2, pp. 37–48, 2012.
- [12] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, London, UK, 2002.
- [13] S. Dempe, J. Dutta, and B. S. Mordukhovich, "New necessary optimality conditions in optimistic bilevel programming," *Optimization*, vol. 56, no. 5-6, pp. 577–604, 2007.
- [14] B. S. Mordukhovich, N. M. Nam, and H. M. Phan, "Variational analysis of marginal functions with applications to bilevel programming," *Journal of Optimization Theory and Applications*, vol. 152, no. 3, pp. 557–586, 2012.
- [15] J. J. Ye and D. L. Zhu, "Optimality conditions for bilevel programming problems," *Optimization*, vol. 33, no. 1, pp. 9–27, 1995.
- [16] R. Henrion and T. Surowiec, "On calmness conditions in convex bilevel programming," *Applicable Analysis*, vol. 90, no. 6, pp. 951–970, 2011.
- [17] F. Benita and P. Mehrlitz, "Bilevel optimal control with final-statedependent finite-dimensional lower level," Tech. Rep., TU Freiberg, 2014.
- [18] J. J. Ye, "Optimal strategies for bilevel dynamic problems," *SIAM Journal on Control and Optimization*, vol. 35, no. 2, pp. 512–531, 1997.
- [19] S. M. Robinson, "Generalized equations and their solutions. I. Basic theory," *Mathematical Programming Studies*, vol. 10, pp. 128–141, 1979.
- [20] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer, *Non-Linear Parametric Optimization*, Birkhauser, Basel, Switzerland, 1983.
- [21] K. Shimizu, Y. Ishizuka, and J. F. Bard, *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.
- [22] F. Nožička, J. Guddat, H. Hollatz, and B. Bank, *Theorie der Linearen Parametrischen Optimierung*, Akademie, Berlin, Germany, 1974.
- [23] S. Dempe and A. B. Zemkoho, "The bilevel programming problem: reformulations, constraint qualifications and optimality conditions," *Mathematical Programming*, vol. 138, no. 1-2, pp. 447–473, 2013.