5-1-2004

Optimal Control of Delay Systems with Differential and Algebraic Dynamic Constraints

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

Lianwen Wang
Wayne State University

OPTIMAL CONTROL OF DELAY SYSTEMS WITH DIFFERENTIAL AND ALGEBRAIC DYNAMIC CONSTRAINTS

BORIS S. MORDUKHOVICH and LIANWEN WANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2004 Series
#5

This research was partly supported by the National Science Foundation.

OPTIMAL CONTROL OF DELAY SYSTEMS WITH DIFFERENTIAL AND ALGEBRAIC DYNAMIC CONSTRAINTS

BORIS S. MORDUKHOVICH[1]
Department of Mathematics, Wayne State University, Detroit, MI 48202
boris@math.wayne.edu
LIANWEN WANG
Department of Mathematics and Computer Science, Central Missouri State University,
Warrensburg, MO 64093
lwang@cmsu1.cmsu.edu

Abstract. This paper concerns constrained dynamic optimization problems governed by delay control systems whose dynamic constraints are described by both delay-differential inclusions and linear algebraic equations. This is a new class of optimal control systems that, on one hand, may be treated as a specific type of variational problems for neutral functional-differential inclusions while, on the other hand, is related to a special class of differential-algebraic systems with a general delay-differential inclusion and a linear constraint link between "slow" and "fast" variables. We pursue a two-hold goal: to study variational stability for this class of control systems with respect to discrete approximations and to derive necessary optimality conditions for both delayed differential-algebraic systems under consideration and their finite-difference counterparts using modern tools of variational analysis and generalized differentiation. The authors are not familiar with any results in these directions for such systems even in the delay-free case. In the first part of the paper we establish the value convergence of discrete approximations as well as the strong convergence of optimal arcs in the classical Sobolev space $W^{1,2}$. Then using discrete approximations as a vehicle, we derive necessary optimality conditions for the initial continuous-time systems in both Euler-Lagrange and Hamiltonian forms via basic generalized differential constructions of variational analysis.

Keywords. Optimal control, variational analysis, functional-differential inclusions of neutral type, differential and algebraic dynamic constraints, discrete approximations, generalized differentiation, necessary optimality conditions.

AMS subject classification. 49K24, 49K25, 49J53, 49M25, 90C31, 93C30.

Running title: Optimal control of dynamic systems with differential and algebraic constraints

1 Introduction

This paper deals with a new class of dynamic optimization problems modeled as follows:

$$\text{minimize} \quad J[x,z] := \varphi(x(a),x(b)) + \int_a^b f(x(t),x(t-\Delta),z(t),\dot{z}(t),t)\,dt \tag{1.1}$$

[1]Research was partly supported by the National Science Foundation under grants DMS-0072179 and DMS-0304989.

subject to the constraints

$$\dot{z}(t) \in F(x(t), x(t-\Delta), z(t), t) \quad \text{a.e. } t \in [a,b], \tag{1.2}$$

$$z(t) = x(t) + Ax(t-\Delta), \quad t \in [a,b], \tag{1.3}$$

$$x(t) = c(t), \quad t \in [a-\Delta, a), \tag{1.4}$$

$$(x(a), x(b)) \in \Omega \subset \mathbb{R}^{2n}, \tag{1.5}$$

where $x : [a-\Delta, b] \to \mathbb{R}^n$ is continuous on $[a-\Delta, a)$ and $[a,b]$ (with a possible jump at $t=a$), and where $z(t)$ is absolutely continuous on $[a,b]$. We always assume that $F\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times [a,b] \rightrightarrows \mathbb{R}^n$ is a set-valued mapping of closed graph, that Ω is a closed set, that $\Delta > 0$ is a constant delay, and that A is a constant $n \times n$ matrix. Let us label this problem as (P) and note that the methods used in this paper allow us to include the cases of multiple delays $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_m > 0$ as well as variable delays $\Delta(t)$ with $|\dot{\Delta}(t)| < \alpha \in (0,1)$ a.e. $t \in [a,b]$, which are not consider here for simplicity. We do not also consider the limiting case of $\Delta \downarrow 0$ required additional assumptions.

Observe that the variational problem (P) involves two kinds of state variables: "slow" x and "fast" z, which satisfy interrelated *dynamic constraints* given by the *delay-differential inclusion* (1.2) and the *linear delay-algebraic equation* (1.3). Furthermore, the integral functional in (1.1) depends on on both slow and fast variables as well as on the *time-derivative of fast variables* (slow variables may not differentiable in time). All these specific features are highly essential for the methods developed and the results obtained in this papers.

On one hand, problem (P) containing both differential and algebraic constraints on slow and fast variables may be viewed as a special subclass of delayed *differential-algebraic* control systems providing, by definition, descriptions of control process via combinations of interconnected differential and algebraic dynamic relations. There are many applications of such dynamic models (called DAEs, i.e., differential-algebraic equations) especially in process systems engineering, robotics, mechanical systems with holonomic and nonholonomic constraints, etc.; see [1, 14, 15] and the references therein. Observe, however, that the dynamic relations (1.2) and (1.3) are generally different from those conventional in the control theory for DAEs. In particular, the only algebraic constraint (1.3) is *linear*, while it does *not* satisfy the *index one* assumption that is usually imposed in known results on necessary optimality conditions; see [2, 14] and their references. The most advanced results in this direction for index one DAEs are obtained in[14], where it was particularly discovered that optimal processes in such systems do *not* satisfy the (strong) Maximum Principle in the absence of a convexity hypothesis on the velocity sets.

On the other hand, the problem (P) under consideration is strongly related to functional-differential control systems of the so-called *neutral type*, which contain time *delays* in *velocity* variables. Indeed, the dynamic constraints (1.2) and (1.3) can be unified as

$$\frac{d}{dt}\big[x(t) + Ax(t-\Delta)\big] \in F\big(x(t), x(t-\Delta), x(t) + Ax(t-\Delta), t\big) \quad \text{a.e.},$$

that, provided the absolute continuity of $\bar{x}(t)$ (which is not the case under the assumptions made), may be written in the general form of neutral delay differential inclusions

$$\dot{x}(t) \in G(x(t), x(t-\Delta), \dot{x}(t-\Delta), t) \quad \text{a.e.} \tag{1.6}$$

Similarly, the cost functional (1.1) transfers under this substitution into the form

$$\varphi(x(a), x(b)) + \int_a^b g\big(x(t), x(t-\Delta), \dot{x}(t), \dot{x}(t-\Delta), t\big)\, dt \tag{1.7}$$

Thus we can treat problem (P) as a special case of Bolza-type variational problems for neutral delay-differential inclusions. However, in this way we loose the principal feature of the considered problem (P), which is crucial for the methods applied as well as for the results obtained below. This specific feature of problem (P) is as follows: both the dynamic constraint (1.6) and the cost functional (1.7) depend in fact not on $\dot{x}(t)$ and $\dot{x}(t-\Delta)$ but on the derivative of the *same linear combination* $x(t) + Ax(t-\Delta)$. That is why we treat this linear combination as a new state variable in (1.3) and consider problem (P) in the natural form (1.1)–(1.5), which emphasizes both *delay-differential* and *linear algebraic constraint* on the system dynamics. In our opinion, the recognition of this special class of dynamic optimization problems is a significant contribution of this paper.

Our approach is based on the *method of discrete approximations*, in the line developed in [8, 10, 11, 12] for nondelayed differential inclusions, delay-differential inclusions with $A = 0$, and for a special class of the neutral-type problems that corresponds to (P) with F independent of z and with f independent of $(z, \dot{z})$. Some results for delayed differential-algebraic problems of type (P) were announced in [13] in the case when both F and f are independent of z, while f depends on the velocity $\dot{z}$ described by (1.2).

The discrete approximation method is of undoubted interest from qualitative as well as numerical viewpoints, and the present paper contains results in both of these directions. Our main emphasis, however, is the *qualitative aspect*, which allows us to derive necessary optimality conditions for delayed differential-algebraic systems by passing to the limit from their discrete-time analogues. A crucial issue is to establish *variational stability* of discrete approximations that ensures an appropriate *strong convergence* of optimal solutions.

Once such a stability is established, discrete-time control problems for delayed *difference-algebraic inclusions* reduce to special *finite-dimensional* problems of *nonsmooth programming* with an increasing number of *geometric constraints* that may have empty interiors. To handle such problems, we use appropriate generalized differentiation tools of variational analysis introduced earlier by the first author. In this way we derive necessary optimality conditions for the discrete-time and then for continuous-time problems under consideration.

The rest of the paper is organized as follows. In Section 2 we show that any admissible pair to the delayed differential-algebraic system (1.2) and (1.3) can be *strongly approximated* by the corresponding admissible pairs to its finite-difference counterparts. This result important for its

own sake plays a crucial role in the construction of *well-posed* discrete approximations to the original problem (P) and in the subsequent justification of the strong convergence of their optimal solutions to the given optimal solution for (P).

Such a convergence analysis is conducted in Section 3 involving appropriate *perturbations* of the endpoint constraints (1.5) that is consistent with the step of discretization. The required strong convergence of optimal solutions is justified under an intrinsic property of the original problem (P) called *relaxation stability*. This property imposing the equality between the optimal values in (P) and its relaxation goes far beyond the convexity assumption on the velocity sets $F(x,y,z,t)$.

Section 4 contains basic constructions and required material on generalized differentiation that are appropriate for performing a variational analysis of discrete-time and continuous-time optimal control problems in the subsequent sections. These constructions and calculus rules of generalized differentiation are used in Section 5 for deriving general necessary optimality conditions for non-convex discrete-time inclusions arising in discrete approximations of the original control problem (P). The main necessary optimality conditions in the forms of Euler-Lagrange and Hamiltonian inclusions for (P) are derived in Section 6 via passing to the limit from discrete approximations.

Our notation is basically standard; cf. [8] and [17]. Recall that, given a set-valued mapping (or multifunction) $F\colon X \rightrightarrows Y$ between finite-dimensional spaces, the *Painlevé-Kuratowski upper/outer limit* of $F(x)$ as $x \to \bar{x}$ is defined by

$$\operatorname*{Limsup}_{x\to\bar{x}} F(x) := \big\{y \in Y |\ \exists\ x_k \to \bar{x},\ \exists\ y_k \to y \ \text{ with } \ y_k \in F(x_k) \quad \text{for all } \ k \in I\!N\big\},$$

where $I\!N$ stands for the collection of all natural numbers.

2 Discrete approximations of differential-algebraic inclusions

This section deals with discrete approximations of an arbitrary admissible pair to the delayed differential-algebraic system (1.2)–(1.4) *without* taking into account the endpoint constraints. We show that, under fairly general assumptions, any admissible pair to (1.2)–(1.4) can be *strongly approximated* in the sense indicated below by the corresponding admissible pairs to finite-difference inclusions obtained from (1.2)–(1.4) by the classical Euler scheme. This result is *constructive* providing *efficient estimates of the approximation rate*, and hence it is certainly of independent interest for numerical analysis of delayed differential-algebraic inclusions.

Let $(\bar{x},\bar{z})$ be an *admissible pair* to (1.2)–(1.4), i.e., $\bar{x}(\cdot)$ is continuous on $[a-\Delta,a)$ and $[a,b]$ (with a possible jump at $t=a$), $\bar{z}(\cdot)$ is absolutely continuous on $[a,b]$, and relations (1.2)–(1.4) are satisfied. Note that the endpoint constraints (1.5) may not hold for $(\bar{x},\bar{z})$; if they do hold, $(\bar{x},\bar{z})$ is called *feasible* to (P). The following *standing assumptions* are imposed throughout the paper:

(H1) There are two open sets $U \subset I\!R^n$, $V \subset I\!R^n$ and two positive numbers ℓ_F, m_F such that $\bar{x}(t) \in U$ for all $t \in [a-\Delta,b]$ and $\bar{z}(t) \in V$ for all $t \in [a,b]$, that the sets $F(x,y,z,t)$ are

closed, and that one has

$$F(x,y,z,t) \subset m_F I\!B \text{ for all } (x,y,z,t) \in U \times U \times V \times [a,b],$$
$$F(x_1,y_1,z_1,t) \subset F(x_2,y_2,z_2,t) + \ell_F(|x_1-x_2|+|y_1-y_2|+|z_1-z_2|)I\!B$$

if $(x_1,y_1,z_1),(x_2,y_2,z_2) \in U \times U \times V$ and $t \in [a,b]$, where $I\!B$ stands for the closed unit ball in $I\!R^n$.

(H2) $F(x,y,z,t)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $(x,y,z) \in U \times U \times V$.

(H3) The function $c(\cdot)$ is continuous on $[a-\Delta,a]$.

Following [3], we consider the so-called *averaged modulus of continuity* for the multifunction $F(x,y,z,t)$ with $(x,y,z) \in U \times U \times V$ and $t \in [a,b]$ defined by

$$\tau(F;h) := \int_a^b \sigma(F;t,h)\,dt,$$

where $\sigma(F;t,h) := \sup\big\{\vartheta(F;x,y,z,t,h)\big|\ (x,y,z) \in U \times U \times V\big\}$ with

$$\vartheta(F;x,y,z,t,h) := \sup\Big\{\text{haus}(F(x,y,z,t_1),F(x,y,z,t_2))\ \Big|\ (t_1,t_2) \in \big[t-\frac{h}{2},t+\frac{h}{2}\big] \cap [a,b]\Big\},$$

and where $\text{haus}(\cdot,\cdot)$ stands for the Hausdorff distance between two compact sets. It is proved in [3] that if $F(x,y,z,t)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $(x,y,z) \in U \times U \times V$, then $\tau(F;h) \to 0$ as $h \to 0$. This fact is essentially used in what follows.

Let us construct a sequence of discrete approximations of the given delayed differential-algebraic inclusion replacing the derivative in (1.2) by the classical *Euler finite difference*

$$\dot{z}(t) \approx \frac{z(t+h)-z(t)}{h}.$$

For any $N \in I\!N := \{1,2,\ldots\}$ we consider the *step of discretization* $h_N := \Delta/N$ and define the *discrete partition* $t_j := a + jh_N$ as $j = -N,\ldots,k$ and $t_{k+1} := b$, where k is a natural number determined from $a+kh_N \le b < a+(k+1)h_N$. Then the corresponding *delayed difference-algebraic inclusions* associated with (1.2)–(1.4) are described by

$$\begin{cases} z_N(t_{j+1}) \in z_N(t_j) + h_N F(x_N(t_j),x_N(t_j-\Delta),z_N(t_j),t_j) \quad \text{for} \quad j=0,\ldots,k, \\ z_N(t_j) = x_N(t_j) + Ax_N(t_j-\Delta) \quad \text{for} \quad j=0,\ldots,k+1, \\ x_N(t_j) = c(t_j) \quad \text{for} \quad j=-N,\ldots,-1. \end{cases} \tag{2.1}$$

Given discrete functions $x_N(t_j)$ and $z_N(t_j)$ satisfying (2.1), we consider the *extension* of $x_N(t_j)$ to the continuous-time intervals $[a-\Delta,b]$ such that $x_N(t)$ are defined piecewise-linearly on $[a,b]$ and

piecewise-constantly, continuously from the right on $[a-\Delta, a)$. We also define piecewise-constant *extension of discrete velocities* on $[a,b]$ by

$$v_N(t) := \frac{z_N(t_{j+1}) - z_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}),\ j = 0, \ldots, k.$$

It is easy to see that

$$z_N(t) = z_N(a) + \int_a^t v_N(s)\, ds \quad \text{for } t \in [a,b],$$

where $z_N(t) = x_N(t) + Ax_N(t-\Delta)$.

Let $W^{1,2}[a,b]$ be the classical Sobolev space with the norm

$$\|x(\cdot)\|_{W^{1,2}} := \max_{t\in[a,b]} |x(t)| + \Big(\int_a^b |\dot{x}(t)|^2 dt\Big)^{1/2}.$$

The following theorem, which plays an essential role in the subsequent constructions and results of the paper being also important for its own sake, establishes the *strong $W^{1,2}$-approximation* of any admissible pair to the given delayed differential-algebraic inclusion by corresponding solutions to its discrete-time counterparts.

Theorem 2.1 *Let $(\bar{x}, \bar{z})$ be an admissible pair to (1.2)–(1.4) under hypotheses (H1)–(H3). Then there is a sequence $(\widehat{x}_N(t_j), \widehat{z}_N(t_j))$ of solutions to discrete inclusions (2.1) with $\widehat{x}_N(t_0) = \bar{x}(a)$ for all $N \in \mathbb{N}$ such that their extensions $\widehat{x}_N(t)$, $a-\Delta \le t \le b$, converge uniformly to $\bar{x}(\cdot)$ on $[a-\Delta, b]$ while $\widehat{z}_N(t)$, $a \le t \le b$, converge to $\bar{z}(t)$ in the $W^{1,2}$-norm on $[a,b]$ as $N \to \infty$.*

Proof. Using the density of step-functions in $L^1[a,b]$, we first select a sequence $\{\omega_N(\cdot)\}$, $N \in \mathbb{N}$, such that each $\omega_N(t)$ is constant on the interval $[t_j, t_{j+1})$ for $j = 0, \ldots, k$ and that $\omega_N(\cdot)$ converge to $\dot{\bar{z}}(\cdot)$ as $N \to \infty$ in the norm topology of $L^1[a,b]$. It follows from (H1) that

$$|\omega_N(t)| \le \big|\omega_N(t) - \dot{\bar{z}}(t)\big| + \big|\dot{\bar{z}}(t)\big| \le 1 + m_F$$

for all $t \in [a,b]$ and $N \in \mathbb{N}$. In the estimates below we use the sequence

$$\xi_N := \int_a^b \big|\omega_N(t) - \dot{\bar{z}}(t)\big|\, dt \to 0 \ \text{ as } \ N \to \infty.$$

Denote $\omega_{N_j} := \omega_N(t_j)$ and define discrete pairs $(u_N(t_j), s_N(t_j))$ recurrently by

$$\begin{cases} u_N(t_j) := \bar{x}(t_j) \ \text{ for } \ j = -N, \ldots, 0, \\ s_N(t_j) := u_N(t_j) + Au_N(t_j - \Delta) \quad \text{for } j = 0, \ldots, k+1, \\ s_N(t_{j+1}) := s_N(t_j) + h_N \omega_{N_j} \quad \text{for } j = 0, \ldots, k. \end{cases}$$

Then the extended discrete functions satisfy

$$\begin{cases} u_N(t) = \bar{x}(t_j) \quad \text{for } t \in [t_j, t_{j+1}),\ j = -N, \ldots, -1, \\ s_N(t) = u_N(t) + Au_N(t-\Delta) \quad \text{for } t \in [a,b], \\ s_N(t) = \bar{z}(a) + \displaystyle\int_a^t \omega_N(s)\, ds \quad \text{for } t \in [a,b]. \end{cases}$$

Next we want to prove that $u_N(t)$ converge uniformly to $\bar{x}(t)$ on $[a,b]$. Denote $r_N(t) := u_N(t) - \bar{x}(t)$ and $y_N(t) := |r_N(t) + Ar_N(t-\Delta)|$. For any $t \in [a,b]$ one has

$$y_N(t) = \big|s_N(t) - \bar{z}(t)\big| \le \int_a^t |\omega_N(s) - \dot{\bar{z}}(s)|\, ds \le \ \xi_N, \tag{2.2}$$

which implies the estimates

$$\begin{aligned} |r_N(t)| &\le y_N(t) + |A||r_N(t-\Delta)| \le y_N(t) + |A|y_N(t-\Delta) + |A|^2|r_N(t-2\Delta)| \le \ldots \\ &\le y_N(t) + |A|y_N(t-\Delta) + \ldots + |A|^m y_N(t-m\Delta) + |A|^{m+1}|r_N(t-(m+1)\Delta)|. \end{aligned}$$

Observe that $c(\cdot)$ is *uniformly continuous* on $[a-\Delta, a]$ due to assumption (H3). Picking an arbitrary sequence $\beta_N \downarrow 0$ as $N \to \infty$, we therefore have

$$|c(t') - c(t'')| \le \beta_N \ \text{ whenever } \ t', t'' \in [t_j, t_{j+1}], \quad j = -N, \ldots, -1.$$

Choose an integer number m such that $a - \Delta \le b - (m+1)\Delta < a$. Then $t - (m+1)\Delta \in [t_j, t_{j+1})$ for some $j \in \{-N, \ldots, -1\}$, which implies that

$$|r_N(t-(m+1)\Delta)| \le |c(t_j) - c(t-(m+1)\Delta)| \le \beta_N.$$

Since $m \in I\!N$ does not depend on N, this gives

$$|r_N(t)| \le \xi_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}\beta_N := \varrho_N \to 0 \quad \text{as } N \to \infty \tag{2.3}$$

for all $t \in [a,b]$ due to the construction of $r_N(\cdot)$. Now consider a sequence $\{\zeta_N\}$ defined by

$$\zeta_N := \ h_N \sum_{j=0}^{k} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), s_N(t_j), t_j))$$

and show that $\zeta_N \downarrow 0$ as $N \to \infty$. By construction of ζ_N and the averaged modulus of continuity

$\tau(F;h)$ we get the following estimates:

$$\zeta_N = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t_j))\, dt$$

$$= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t))\, dt +$$

$$\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left[\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t_j)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t))\right] dt$$

$$\le \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t))\, dt + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \sigma(F; t, h_N)\, dt$$

$$\le \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t))\, dt + \tau(F; h_N).$$

Further, assumption (H1) implies that for any $t \in [t_j, t_{j+1})$ with $j = 0, \dots, k$ one has

$$\begin{aligned}
&\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), s_N(t), t)) \\
&\le \operatorname{dist}(F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t), F(u_N(t), u_N(t-\Delta), s_N(t), t)) \\
&\le \ell_F(|u_N(t_j) - u_N(t)| + |u_N(t_j-\Delta) - u_N(t-\Delta)| + |s_N(t_j) - s_N(t)|).
\end{aligned}$$

Taking into account that

$$\left|s_N(t_j) - s_N(t)\right| = \left| \int_{t_j}^{t} \omega_N(s)\, ds \right| \le (1 + m_F)(t_{j+1} - t_j) = (1 + m_F) h_N := a_N \downarrow 0,$$

we arrive at

$$\begin{aligned}
|u_N(t) - u_N(t_j)| &\le a_N + |A||u_N(t-\Delta) - u_N(t_j-\Delta)| \\
&\le a_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}|u_N(t-(m+1)\Delta) - u_N(t_j-(m+1)\Delta)| \\
&\le a_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}\beta_N := b_N \downarrow 0 \ \text{ as } \ N \to \infty
\end{aligned}$$

and hence ensure that

$$\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), s_N(t), t)) \le (a_N + 2b_N)\ell_F.$$

It follows from (H1), (2.2) and (2.3) that for any $t \in [t_j, t_{j+1})$ and $j = 0, \dots, k$ one has

$$\begin{aligned}
&\operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), s_N(t), t)) - \operatorname{dist}(\omega_N(t); F(\bar{x}(t), \bar{x}(t-\Delta), \bar{z}(t), t)) \\
&\le \operatorname{dist}(F(u_N(t), u_N(t-\Delta), s_N(t), t),\ F(\bar{x}(t), \bar{x}(t-\Delta), \bar{z}(t), t)) \\
&\le \ell_F(|u_N(t) - \bar{x}(t)| + |u_N(t-\Delta) - \bar{x}(t-\Delta)| + |s_N(t) - \bar{z}(t)|) \\
&\le (2\varrho_N + \xi_N)\ell_F.
\end{aligned}$$

Combining the above estimates and denoting $\mu_N := a_N + 2b_N + 2\varrho_N + \xi_N$, we arrive at

$$\begin{aligned}&\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t))\\ &\le \ell_F\mu_N + \operatorname{dist}(\omega_{N_j}; F(\bar{x}(t), \bar{x}(t-\Delta), \bar{z}(t), t)) \le \ell_F\mu_N + \left|\omega_{N_j} - \dot{\bar{z}}(t)\right|\end{aligned}$$

and finally conclude that

$$\begin{aligned}\zeta_N &\le \sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} \left(\left|\omega_{N_j} - \dot{\bar{z}}(t)\right| + \ell_F\mu_N\right) dt + \tau(F; h_N)\\ &= \xi_N + \ell_F\mu_N(b-a) + \tau(F; h_N) := \gamma_N \downarrow 0 \ \text{ as } \ N \to \infty.\end{aligned} \tag{2.4}$$

Note that the discrete functions $(u_N(t_j), s_N(t_j))$ may *not* be a admissible pair for (2.1) because the inclusions $\omega_{N_j} \in F(u_N(t_j), u_N(t_j-\Delta), s_N(t_j), t_j)$ may not be true for $j = 0, \ldots, k$. Let us construct the desired pair $(\widehat{x}_N(t_j), \widehat{z}_N(t_j))$ by the following *proximal algorithm*:

$$\begin{cases}\widehat{x}_N(t_j) = c(t_j) \ \text{ for } \ j = -N, \ldots, -1, \quad \widehat{x}_N(t_0) = \bar{x}(a),\\ \widehat{z}_N(t_{j+1}) = \widehat{z}_N(t_j) + h_N v_{N_j} \quad \text{for } \ j = 0, \ldots, k,\\ \widehat{z}_N(t_j) = \widehat{x}_N(t_j) + A\widehat{x}_N(t_j-\Delta) \quad \text{for } \ j = 0, \ldots, k+1,\\ v_{N_j} \in F(\widehat{x}_N(t_j), \widehat{x}_N(t_j-\Delta), \widehat{z}_N(t_j), t_j) \quad \text{for } \ j = 0, \ldots, k,\\ |v_{N_j} - \omega_{N_j}| = \operatorname{dist}(\omega_{N_j}; F(\widehat{x}_N(t_j), \widehat{x}_N(t_j-\Delta), \widehat{z}_N(t_j), t_j)) \quad \text{for } \ j = 0, \ldots, k.\end{cases} \tag{2.5}$$

It follows from the construction (2.5) that $(\widehat{x}_N(t_j), \widehat{z}_N(t_j))$ is a feasible pair to the discrete inclusion (2.1) for each $N \in \mathbb{N}$. Note that

$$|\widehat{x}_N(t) - \bar{x}(t)| = |\widehat{x}_N(t_j) - \bar{x}(t)| = |c(t_j) - c(t)| < \beta_N \quad \text{for } t \in [t_j, t_{j+1}), \ j = -N, \ldots, -1,$$

which implies that the extensions of $\widehat{x}_N(\cdot)$ converge to $\bar{x}(t)$ uniformly on $[a-\Delta, a)$. Let us analyze the situation on $[a, b]$.

First we claim that $\widehat{x}_N(t_j) \in U$ and $\widehat{z}_N(t_j) \in V$ for $j = 0, \ldots, k+1$. Arguing by induction, we obviously have $\widehat{x}_N(t_0) \in U$ and $\widehat{z}_N(t_0) \in V$. Assume that $\widehat{x}_N(t_j) \in U$ and $\widehat{z}_N(t_j) \in U$ for all $j = 1, \ldots, m$ with some fixed $m \in \{1, \ldots, k\}$. Then

$$\begin{aligned}&|\widehat{x}_N(t_{m+1}) - u_N(t_{m+1})| = |\widehat{z}_N(t_{m+1}) - A\widehat{x}_N(t_{m+1}-\Delta) - s_N(t_{m+1}) + Au_N(t_{m+1}-\Delta)|\\ &\le |A||\widehat{x}_N(t_{m+1}-\Delta) - u_N(t_{m+1}-\Delta)| + |\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})|\\ &\le |A||\widehat{x}_N(t_{m+1}-\Delta) - u_N(t_{m+1}-\Delta)| + |A||\widehat{x}_N(t_m-\Delta) - u_N(t_m-\Delta)|\\ &+ |\widehat{x}_N(t_m) - u_N(t_m)| + h_N\operatorname{dist}(\omega_{N_m}; F(\widehat{x}_N(t_m), \widehat{x}_N(t_m-\Delta), \widehat{z}_N(t_m), t_m)).\end{aligned}$$

Taking into account that

$$\begin{aligned}&|\widehat{x}_N(t_m) - u_N(t_m)| \le |A||\widehat{x}_N(t_{m-N}) - u_N(t_{m-N})| + |A||\widehat{x}_N(t_{m-1-N}) - u_N(t_{m-1-N})|\\ &+ |\widehat{x}_N(t_{m-1}) - u_N(t_{m-1})| + h_N\operatorname{dist}(\omega_{N_{m-1}}; F(\widehat{x}_N(t_{m-1}), \widehat{x}_N(t_{m-1-N}), \widehat{z}_N(t_{m-1}), t_{m-1})),\end{aligned}$$

that

$$\begin{aligned}
&\operatorname{dist}(\omega_{N_{m-1}}; F(\widehat{x}_N(t_{m-1}), \widehat{x}_N(t_{m-1-N}), \widehat{z}_N(t_{m-1}), t_{m-1})) \\
&\le \operatorname{dist}(\omega_{N_{m-1}}; F(u_N(t_{m-1}), u_N(t_{m-1-N}), s_N(t_{m-1}), t_{m-1})) \\
&+\ell_F(|\widehat{x}_N(t_{m-1}) - u_N(t_{m-1})| + |\widehat{z}_N(t_{m-1}) - s_N(t_{m-1})| \\
&+|\widehat{x}_N(t_{m-1-N}) - u_N(t_{m-1-N})|),
\end{aligned} \tag{2.6}$$

that

$$|\widehat{z}_N(t_m) - s_N(t_m)| \le |\widehat{x}_N(t_m) - u_N(t_m)| + |A||\widehat{x}_N(t_{m-N}) - u_N(t_{m-N})|, \tag{2.7}$$

and that $|\widehat{x}_N(t_j) - u_N(t_j)| = 0$ for $j \le 0$, one has

$$|\widehat{x}_N(t_{m+1}) - u_N(t_{m+1})| \le M_1 h_N \sum_{j=0}^{m} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), s_N(t_j), t_j)) \le M_1 \gamma_N \tag{2.8}$$

with some constant $M_1 > 0$. Now invoking (2.3) and increasing M_1 if necessary, we arrive at

$$|\widehat{x}_N(t_{m+1}) - \bar{x}(t_{m+1})| \le \xi_N + M_1 \gamma_N \to 0 \ \text{ as } \ N \to \infty,$$

which implies that $\widehat{x}_N(t_j) \in U$ for $j = 0, \ldots, k+1$.

Observing further that

$$\begin{aligned}
|\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})| &\le |\widehat{z}_N(t_m) - s_N(t_m)| + h_N |v_{N_m} - \omega_{N_m}| \\
&\le |\widehat{z}_N(t_m) - s_N(t_m)| + h_N \operatorname{dist}(\omega_{N_m}; F(\widehat{x}_N(t_m), \widehat{x}_N(t_m - \Delta), \widehat{z}_N(t_m), t_m)),
\end{aligned}$$

we derive from (2.6) and (2.7) the estimate

$$|\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})| \le M_2 h_N \sum_{j=0}^{m} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), s_N(t_j), t_j)) \le M_2 \gamma_N \tag{2.9}$$

with some constant $M_2 > 0$. Note that

$$|\widehat{z}_N(t_{m+1}) - \bar{z}_N(t_{m+1})| \le |\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})| + |s_N(t_{m+1}) - \bar{z}_N(t_{m+1})| \le M_2 \gamma_N + \xi_N,$$

which ensures the inclusions $\widehat{z}_N(t_j) \in V$ for $j = 0, \ldots, k+1$. It remains to prove that $\widehat{z}_N(t)$ converge to $\bar{z}(t)$ in the $W^{1,2}$-norm on $[a, b]$, which means that

$$\max_{t \in [a,b]} \big|\widehat{z}_N(t) - \bar{z}(t)\big| + \int_a^b \big|\dot{\widehat{z}}_N(t) - \dot{\bar{z}}(t)\big|^2 \, dt \to 0 \quad \text{as } \ N \to \infty. \tag{2.10}$$

To furnish this, we use (2.8) and (2.9) to get the estimates

$$\sum_{j=0}^{k+1}\left|\widehat{x}_N(t_j)-u_N(t_j)\right|\le\sum_{j=0}^{k+1}M_1\sum_{m=0}^{j-1}h_N\mathrm{dist}(\omega_{N_m};F(u_N(t_m),u_N(t_m-\Delta),s_N(t_m),t_m))$$
$$\le M_1(b-a)\sum_{j=0}^{k}\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j)),$$
$$\sum_{j=0}^{k+1}\left|\widehat{z}_N(t_j)-s_N(t_j)\right|\le\sum_{j=0}^{k+1}M_2\sum_{m=0}^{j-1}h_N\mathrm{dist}(\omega_{N_m};F(u_N(t_m),u_N(t_m-\Delta),s_N(t_m),t_m))$$
$$\le M_2(b-a)\sum_{j=0}^{k}\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j)),$$

which imply by (H1) that

$$\int_a^b\left|\dot{\widehat{z}}_N(t)-\omega_N(t)\right|dt=\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\left|\dot{\widehat{z}}_N(t)-\omega_N(t)\right|dt$$
$$=\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}|v_{N_j}-\omega_{N_j}|\,dt\;=\sum_{j=0}^{k}h_N\mathrm{dist}(\omega_{N_j};F(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),t_j))$$
$$=\sum_{j=0}^{k}h_N\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j))$$
$$+\sum_{j=0}^{k}h_N[\mathrm{dist}(\omega_{N_j};F(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),t_j))-\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j))]$$
$$\le\sum_{j=0}^{k}h_N\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j))$$
$$+\sum_{j=0}^{k}\ell_F h_N\left[|\widehat{x}_N(t_j)-u_N(t_j)|+|\widehat{x}_N(t_j-\Delta)-u_N(t_j-\Delta)|+|\widehat{z}_N(t_j)-s_N(t_j)|\,\right]$$
$$\le\gamma_N+\sum_{j=0}^{k}\ell_F h_N\left[|\widehat{x}_N(t_j)-u_N(t_j)|+|\widehat{x}_N(t_j-\Delta)-u_N(t_j-\Delta)|+|\widehat{z}_N(t_j)-s_N(t_j)|\,\right]$$
$$\le\gamma_N+2(M_1+M_2)(b-a)\ell_F\sum_{j=0}^{k}h_N\mathrm{dist}(\omega_{N_j};F(u_N(t_j),u_N(t_j-\Delta),s_N(t_j),t_j))$$
$$\le\gamma_N+2(M_1+M_2)\ell_F(b-a)\gamma_N.$$

The latter ensures the estimates

$$\int_a^b\left|\dot{\widehat{z}}_N(t)-\dot{\bar{z}}(t)\right|dt\le\int_a^b\left|\dot{\widehat{z}}_N(t)-\omega_N(t)\right|dt+\int_a^b\left|\omega_N(t)-\dot{\bar{z}}(t)\right|dt$$
$$\le\gamma_N(1+2(M_1+M_2)(b-a)\ell_F)+\xi_N.$$

Due to $\widehat{x}_N(t)\in U$ and $\widehat{z}_N(t)\in V$, we get from (H1) by (1.2) and (2.5) that $\left|\dot{\widehat{z}}_N(t)\right|\le m_F$,

$|\dot{\bar{z}}(t)| \le m_F$, and hence

$$\int_a^b \big|\dot{\widehat{z}}_N(t) - \dot{\bar{z}}(t)\big|^2\,dt = \int_a^b \big|\dot{\widehat{z}}_N(t) - \dot{\bar{z}}(t)\big|\big|\dot{\widehat{z}}_N(t) + \dot{\bar{z}}(t)\big|\,dt$$
$$\le\ 2m_F[\gamma_N(1+2(M_1+M_2)(b-a)\ell_F)+\xi_N]\downarrow 0 \ \text{ as } \ N\to\infty.$$

Observing finally that

$$\max_{t\in[a,b]}\big|\widehat{z}_N(t)-\bar{z}(t)\big|^2 \le (b-a)\int_a^b \big|\dot{\widehat{z}}_N(t)-\dot{\bar{z}}(t)\big|^2\,dt,$$

we arrive at (2.10) and complete the proof of the theorem. $\triangle$

3 Strong convergence of discrete approximations

The goal of this section is to construct a sequence of *well-posed* discrete approximations of the dynamic optimization problem (P) such that optimal solutions to discrete approximation problems *strongly converge*, in the sense described below, to a given optimal solution to the original optimization problem governed by delayed differential-algebraic inclusions. The following construction explicitly involves the optimal solution $(\bar{x},\bar{z})$ to the problem (P) under consideration for which we aim to derive necessary optimality conditions in the subsequent sections.

For any natural number N we consider the following *discrete-time* dynamic optimization problem (P_N):

$$\begin{aligned}
&\text{minimize}\quad J_N[x_N,z_N]:=\varphi(x_N(t_0),x_N(t_{k+1}))+|x_N(t_0)-\bar{x}(a)|^2\\
&+h_N\sum_{j=0}^{k} f(x_N(t_j),x_N(t_j-\Delta),z_N(t_j),\frac{z_N(t_{j+1})-z_N(t_j)}{h_N},t_j)\\
&+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\Big|\frac{z_N(t_{j+1})-z_N(t_j)}{h_N}-\dot{\bar{z}}(t)\Big|^2\,dt
\end{aligned}\tag{3.1}$$

subject to the *dynamic constraints* governed by delayed difference-algebraic inclusions (2.1), the *perturbed endpoint constraints*

$$(x_N(t_0),x_N(t_{k+1}))\in\Omega_N:=\Omega+\eta_N I\!B,\tag{3.2}$$

where $\eta_N:=|\widehat{x}_N(t_{k+1})-\bar{x}(b)|$ with the approximation $\widehat{x}_N(t)$ of $\bar{x}(t)$ from Theorem 2.1, and the *auxiliary constraints*

$$|x_N(t_j)-\bar{x}(t_j)|\le\varepsilon,\quad |z_N(t_j)-\bar{z}(t_j)|\le\varepsilon,\quad j=1,\ldots,k+1,\tag{3.3}$$

with some $\varepsilon>0$. The latter auxiliary constraints are needed to guarantee the *existence* of optimal solutions in (P_N) and can be *ignored* in the derivation of necessary optimality conditions; see below.

In what follows we select $\varepsilon > 0$ in (3.3) such that $\bar{x}(t) + \varepsilon I\!B \subset U$ for all $t \in [a - \Delta, b]$ and $\bar{z}(t) + \varepsilon I\!B \subset V$ for all $t \in [a, b]$. Take sufficiently large N ensuring that $\eta_N < \varepsilon$. Note that problems (P_N) have *feasible* solutions, since the pair $(\widehat{x}_N, \widehat{z}_N)$ from Theorem 2.1 satisfy all the constraints (2.1), (3.2), and (3.3). Therefore, by the classical Weierstrass theorem, each (P_N) admits an *optimal* pair $(\bar{x}_N, \bar{z}_N)$ under the following assumption imposed in addition to (H1)–(H3):

(H4) φ is continuous on $U \times U$, $f(x, y, z, v, \cdot)$ is continuous for a.e. $t \in [a, b]$ uniformly in $(x, y, z, v) \in U \times U \times V \times m_F I\!B$, $f(\cdot, \cdot, \cdot, \cdot, t)$ is continuous on $U \times U \times V \times m_F I\!B$ uniformly in $t \in [a, b]$, and Ω is locally closed around $(\bar{x}(a), \bar{x}(b))$.

We are going to justify the strong convergence of $(\bar{x}_N, \bar{z}_N)$ to $(\bar{x}, \bar{z})$ in the sense of Theorem 2.1. To proceed, we need to involve an important intrinsic property of the original problem (P) called *relaxation stability*. Let us consider, along with the original system (1.2), the *convexified* delayed differential-algebraic system

$$\begin{cases} \dot{z}(t) \in \operatorname{co} F(x(t), x(t - \Delta), z(t), t) \quad \text{a.e. } t \in [a, b], \\ z(t) = x(t) + Ax(t - \Delta), \quad t \in [a, b], \end{cases} \tag{3.4}$$

where "co" stands for the convex hull of a set. Further, given the integrand f in (1.1), we take its restriction

$$f_F(x, y, z, v, t) := f(x, y, z, v, t) + \delta(v; F(x, y, z, t))$$

to F in (1.2), where $\delta(\cdot; F)$ stands for the indicator function of a set. Denote by $\widehat{f}_F(x, y, z, v, t)$ the *convexification* of f_F in the v variable and define the *relaxed generalized Bolza problem* (R) for delayed differential-algebraic systems as follows:

$$\text{minimize } \widehat{J}\,[x, z] := \varphi(x(a), x(b)) + \int_a^b \widehat{f}_F(x(t), x(t - \Delta), z(t), \dot{z}(t), t)\, dt \tag{3.5}$$

over feasible pairs (x, z) with the same analytic properties as in (P) subject to the tail (1.4) and endpoint (1.5) constraints. Every feasible pair to (R) is called a *relaxed pair* to (P).

One clearly has $\inf(R) \le \inf(P)$ for the optimal values of the cost functionals in the relaxed and original problems. We say that the original problem (P) is *stable with respect to relaxation* if

$$\inf(P) = \inf(R).$$

This property, which obviously holds under the convexity assumptions on the sets $F(x, y, z, t)$ and the integrand f in v, goes far beyond the convexity. General sufficient conditions for the relaxation stability of (P) follow from [4]. We also refer the reader to [8, 10, 19, 20] for more detailed discussions on the validity of the relaxation stability property for various classes of differential and functional-differential control systems.

Now we are ready to establish the following strong convergence theorem for optimal solutions to discrete approximations, which makes a bridge between optimal control problems governed by delayed differential-algebraic and difference-algebraic systems.

Theorem 3.1 *Let $(\bar{x},\bar{z})$ be an optimal pair to problem (P), which is assumed to be stable with respect to relaxation. Suppose also that hypotheses (H1)-(H4) hold. Then any sequence $\{(\bar{x}_N,\bar{z}_N)\}$, $N\in\mathbb{N}$, of optimal pairs to (P_N) extended to the continuous interval $[a-\Delta,b]$ and $[a,b]$ respectively, strongly converges to $(\bar{x},\bar{z})$ as $N\to\infty$ in the sense that $\bar{x}_N$ converge to $\bar{x}$ uniformly on $[a-\Delta,b]$ and $\bar{z}_N$ converge to $\bar{z}$ in the $W^{1,2}$-norm on $[a,b]$.*

Proof. We know from the above discussion that (P_N) has an optimal pair $(\bar{x}_N,\bar{z}_N)$ for all N sufficiently large; suppose that it happens for all $N\in\mathbb{N}$ without loss of generality. We consider the sequence $(\widehat{x}_N,\widehat{z}_N)$ from Theorem 2.1. Since each $(\widehat{x}_N,\widehat{z}_N)$ is feasible to (P_N), one has

$$J_N[\bar{x}_N,\bar{z}_N]\le J_N[\widehat{x}_N,\widehat{z}_N]\ \text{ for all }\ N\in\mathbb{N}.$$

For convenience we represent $J_N[\widehat{x}_N,\widehat{z}_N]$ as the sum of three terms:

$$\begin{aligned}J_N[\widehat{x}_N,\widehat{z}_N]&=\varphi(\widehat{x}_N(t_0),\widehat{x}_N(t_{k+1}))+h_N\sum_{j=0}^{k}f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\frac{\widehat{z}_N(t_{j+1})-\widehat{z}_N(t_j)}{h_N},t_j)\\&\quad+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\Big|\frac{\widehat{z}_N(t_{j+1})-\widehat{z}_N(t_j)}{h_N}-\dot{\widehat{z}}(t)\Big|^2dt:=I_1+I_2+I_3.\end{aligned}$$

It follows from Theorem 2.1 and the assumption on φ in (H4) that

$$I_1\to\varphi(\bar{x}(a),\bar{x}(b))\ \text{ as }\ N\to\infty.$$

Moreover, using the sign "$\approx$" for expressions equivalent as $N\to\infty$ and the notation

$$\widehat{v}_N(t):=\frac{\widehat{z}_N(t_{j+1})-\widehat{z}_N(t_j)}{h_N},\quad t\in[t_j,t_{j+1}),\ j=0,\dots,k,$$

we have the relations

$$\begin{aligned}I_2&=h_N\sum_{j=0}^{k}f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t_j),t_j)\\&=\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t),t)\,dt\\&\quad+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\big[f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t),t_j)-f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t),t)\big]\,dt\\&=\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t),t)\,dt+\tau(f;h_N)\\&\approx\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}f(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\Delta),\widehat{z}_N(t_j),\widehat{v}_N(t),t)\,dt\\&\to\int_a^b f(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),\dot{\bar{z}}(t),t)\,dt\ \text{ as }\ N\to\infty,\end{aligned}$$

$$I_3 = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left|\widehat{v}_N(t) - \dot{\widetilde{z}}(t)\right|^2 dt = \int_a^b \left|\widehat{v}_N(t) - \dot{\widetilde{z}}(t)\right|^2 dt$$
$$= \int_a^b \left|\dot{\widehat{z}}_N(t) - \dot{\widetilde{z}}(t)\right|^2 dt \to 0 \ \text{ as } \ N \to \infty.$$

This implies that $J_N[\widehat{x}_N, \widehat{z}_N] \to J[\bar{x}, \bar{z}]$ as $N \to \infty$, and therefore

$$\limsup_{N\to\infty} J_N[\bar{x}_N, \bar{z}_N] \le J[\bar{x}, \bar{z}]. \tag{3.6}$$

It is easy to observe that the strong convergence claimed in the theorem follows from

$$\rho_N := \left|\bar{x}_N(a) - \bar{x}(a)\right|^2 + \int_a^b \left|\dot{\bar{z}}_N(t) - \dot{\bar{z}}(t)\right|^2 dt \to 0 \ \text{ as } \ N \to \infty.$$

On the contrary, suppose that the latter does not hold. Then there are a constant $\alpha > 0$ and a subsequence $\{N_m\} \subset \mathbb{N}$ for which $\rho_{N_m} \to \alpha$ as $m \to \infty$. Employing the standard compactness arguments based on (2.1) and the boundedness assumption in (H1), we find an absolutely continuous function $\widetilde{z}\colon [a,b] \to \mathbb{R}^n$ and a function $\widetilde{x}\colon [a-\Delta, b]$ continuous on $[a-\Delta, a)$ and $[a,b]$ such that

$$\dot{\bar{z}}_N(t) \to \dot{\widetilde{z}}(t) \quad \text{weakly in } L^2[a,b],$$

that $\bar{x}_N(t) \to \widetilde{x}(t)$ uniformly on $[a-\Delta, b]$ as $N \to \infty$ (without loss of generality), and that $\widetilde{z}(t) = \widetilde{x}(t) + A\widetilde{x}(t-\Delta)$ for $t \in [a,b]$. By the classical Mazur theorem there is a sequence of *convex combinations* of $\dot{\bar{z}}_N(t)$ that converges to $\dot{\widetilde{z}}(t)$ in the norm topology of $L^2[a,b]$ and hence *pointwisely* for a.e. $t \in [a,b]$ along some subsequence. Therefore

$$\begin{cases} \dot{\widetilde{z}}(t) \in \operatorname{co} F(\widetilde{x}(t), \widetilde{x}(t-\Delta), \widetilde{z}(t), t) \ \text{ a.e. } \ t \in [a,b], \\ \widetilde{z}(t) = \widetilde{x}(t) + A\widetilde{x}(t-\Delta), \ \ t \in [a,b]. \end{cases}$$

Since $\widetilde{x}(\cdot)$ obviously satisfies the initial tail condition (1.4) and the endpoint constraints (1.5), it is a feasible solution to the relaxed problem (R). Note that

$$h_N \sum_{j=0}^{k} f(\bar{x}_N(t_j), \bar{x}_N(t_j - \Delta), \bar{z}_N(t_j), \frac{\bar{z}_N(t_{j+1}) - \bar{z}_N(t_j)}{h_N}, t_j)$$
$$= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(\bar{x}_N(t_j), \bar{x}_N(t_j - \Delta), \bar{z}_N(t_j), \dot{\bar{z}}_N(t), t_j)\, dt \to \int_a^b f(\widetilde{x}(t), \widetilde{x}(t-\Delta), \widetilde{z}(t), \dot{\widetilde{z}}(t), t)\, dt$$

as $N \to \infty$ due to the assumptions made. Observe also that the integral functional

$$I[v] := \int_a^b \left|v(t) - \dot{\widetilde{z}}(t)\right|^2 dt$$

is lower semicontinuous in the weak topology of $L^2[a,b]$ by the *convexity* of the integrand in v. Since one has

$$\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left|\frac{\bar{z}_N(t_{j+1}) - \bar{z}_N(t_j)}{h_N} - \dot{\widetilde{z}}(t)\right|^2 dt = \int_a^b \left|\dot{\bar{z}}_N(t) - \dot{\widetilde{z}}(t)\right|^2 dt,$$

the latter implies that

$$\int_a^b \left|\dot{\widetilde{z}}(t)-\dot{\bar{z}}(t)\right|^2 dt \le \liminf_{N\to\infty}\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\left|\frac{\bar{z}_N(t_{j+1})-\bar{z}_N(t_j)}{h_N}-\dot{\bar{z}}(t)\right|^2 dt.$$

Using the above relationships and passing to the limit in the expression (3.1) for $J_N[\bar{x}_N,\bar{z}_N]$ as $N\to\infty$, we arrive at the inequality

$$J[\widetilde{x},\widetilde{z}]+\alpha \le \lim_{N\to\infty} J_N[\bar{x}_N,\bar{z}_N].$$

By (3.6) one therefore has

$$J[\widetilde{x},\widetilde{z}] \le J[\bar{x},\bar{z}]-\alpha < J[\bar{x},\bar{z}] \ \text{ if } \ \alpha>0.$$

This clearly contradicts the optimality of pair $(\bar{x},\bar{z})$ in the relaxed problem (R) due to the assumption on relaxation stability. Thus $\alpha=0$, which completes the proof of the theorem. $\triangle$

4 Tools of generalized differentiation

The convergence results of the previous section allow us to make a *bridge* between the original infinite-dimensional optimization problem (P) for delayed differential-algebraic inclusions and the family of finite-dimensional dynamic optimization problems (P_N) for delayed difference-algebraic inclusions. The further strategy is now clear: to obtain first necessary optimality conditions for the discrete approximation problems (P_N) and then to derive necessary optimality conditions for the original problem (P) by passing to the limit from the ones for (P_N) as $N\to\infty$.

Observe that problems (P_N) are essentially *nonsmooth*, even in the case of smooth functions φ and f in the cost functional and the absence of endpoint constraints. The main source of nonsmoothness comes from the (increasing number of) *geometric constraints* in (2.1), which reflect the discrete dynamics and may have empty interiors. To conduct a variational analysis of such problems, we use appropriate tools of generalized differentiation in *finite dimensions* introduced in [5] and then developed and applied in many publications; see, in particular, the books [6, 17, 19] for detailed treatments and further references.

Recall the the basic (general,limiting) *normal cone* to the set $\Omega\subset\mathbb{R}^n$ at the point $\bar{x}\in\Omega$ is

$$N(\bar{x};\Omega) := \operatorname*{Limsup}_{x\overset{\Omega}{\to}\bar{x}} \widehat{N}(x;\Omega), \tag{4.1}$$

where $x\overset{\Omega}{\to}\bar{x}$ means that $x\to\bar{x}$ with $x\in\Omega$, and where

$$\widehat{N}(\bar{x};\Omega) := \Big\{x^*\in\mathbb{R}^n\Big|\ \limsup_{x\overset{\Omega}{\to}\bar{x}}\frac{\langle x^*,x-\bar{x}\rangle}{|x-\bar{x}|}\le 0\Big\} \tag{4.2}$$

is the cone of Fréchet (regular, strict) normals to Ω at $\bar{x}$. For convex sets Ω both cones $N(\bar{x};\Omega)$ and $\widehat{N}(\bar{x};\Omega)$ reduce to the normal cone of convex analysis. Note that the basic normal cone (4.1) is often *nonconvex* while satisfying a comprehensive calculus, which is to the case for (4.2).

Given an extended-real-valued function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}} := [-\infty,\infty]$ finite at $\bar{x}$, the basic (general, limiting) *subdifferential* of φ at $\bar{x}$ is defined geometrically

$$\partial\varphi(\bar{x}) := \big\{x^* \in \mathbb{R}^n \big|\ (x^*,-1) \in N((\bar{x},\varphi(\bar{x})); \operatorname{epi}\varphi)\big\} \tag{4.3}$$

via basic normals to the epigraph $\operatorname{epi}\varphi := \{(x,\mu) \in \mathbb{R}^{n+1} |\ \mu \geq \varphi(x)\}$; equivalent analytic representations of (4.3) can be found in the books [6, 17, 18, 19]. One of the most convenient representations of (4.3) is via the Painlevé-Kuratowski upper limit (i.e., robust regularization) of Fréchet subgradients, which are the same as subgradients in the sense of viscosity solutions.

Given a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with the graph

$$\operatorname{gph} F := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m |\ y \in F(x)\},$$

the *coderivative* $D^*F(\bar{x},\bar{y})\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ of F at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by

$$D^*F(\bar{x},\bar{y})(y^*) := \big\{x^* \in \mathbb{R}^n \big|\ (x^*,-y^*) \in N((\bar{x},\bar{y}); \operatorname{gph} F)\big\}. \tag{4.4}$$

Note the useful relationships

$$\partial\varphi(\bar{x}) = D^*E_\varphi(\bar{x},\varphi(\bar{x}))(1) \quad \text{and} \quad D^*g(\bar{x})(y^*) = \partial\langle y^*,g\rangle(\bar{x}), \quad y^* \in \mathbb{R}^m,$$

between the subdifferential and coderivative introduced, where $E_\varphi(x) := \{\mu \in \mathbb{R} |\ \mu \geq \varphi(x)\}$ is the epigraphical multifunctions associated with $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, and where $\langle y^*,g\rangle(x) := \langle y^*,g(x)\rangle$ is the scalarized function associated with a locally Lipschitzian mapping $g\colon \mathbb{R}^n \to \mathbb{R}^m$. Observe that

$$D^*g(\bar{x})(y^*) = \big\{\nabla g(\bar{x})^* y^*\big\} \ \text{ for all } \ y^* \in \mathbb{R}^m$$

if g is strictly differentiable at $\bar{x}$. This means that the coderivative (4.4) is a natural extension of the *adjoint* derivative operator to nonsmooth and set-valued mappings.

The subdifferential/coderivative constructions (4.3) and (4.4) enjoy a variety of useful calculus rules that can be found in the books mentioned above and their references. Let us formulate two results crucial in the method of discrete approximations. The first one gives a complete coderivative characterization of the classical local Lipschitzian property of multifunctions imposed in our standing assumption (H1); cf. [7, Theorem 5.11] and [17, Theorem 9.40].

Theorem 4.1 *Let $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a closed-graph multifunction locally bounded around $\bar{x}$. Then the following conditions are equivalent:*

(i) *F is locally Lipschitzian around $\bar{x}$.*

(ii) *There exist a neighborhood U of $\bar{x}$ and a number $\ell > 0$ such that*

$$\sup\Big\{|x^*| \;\Big|\; x^* \in D^*F(x,y)(y^*)\Big\} \le \ell|y^*| \quad \textit{for all } x \in U,\ y \in F(x),\ y^* \in I\!R^m.$$

The next result (see, e.g., [6, Corollary 7.5] and [18, Theorem 3.17]) provides necessary optimality conditions for a general problem (MP) of *nonsmooth mathematical programming* with *many geometric constraints*:

$$\begin{cases} \text{minimize } \phi_0(z) \quad \text{subject to} \\ \phi_j(z) \le 0, \quad j = 1,\dots,r, \\ g_j(z) = 0, \quad j = 0,\dots,m, \\ z \in \Lambda_j, \quad j = 0,\dots,l, \end{cases}$$

where $\phi_j\colon I\!R^d \to I\!R$, $g_j\colon I\!R^d \to I\!R^n$, and $\Lambda_j \subset I\!R^d$.

Theorem 4.2 *Let $\bar{z}$ be an optimal solution to (MP). Assume that all ϕ_i are Lipschitz continuous, that g_j are continuously differentiable, and that Λ_j are locally closed near $\bar{z}$. Then there exist real numbers $\{\mu_j |\ j = 0,\dots,r\}$ as well as vectors $\{\psi_j \in I\!R^n |\ j = 0,\dots,m\}$ and $\{z_j^* \in I\!R^d |\ j = 0,\dots,l\}$, not all zero, such that*

$$\mu_j \ge 0 \quad \textit{for} \quad j = 0,\dots,r, \tag{4.5}$$

$$\mu_j \phi_j(\bar{z}) = 0 \quad \textit{for} \quad j = 1,\dots,r, \tag{4.6}$$

$$z_j^* \in N(\bar{z};\Lambda_j) \quad \textit{for} \quad j = 0,\dots,l, \tag{4.7}$$

$$-\sum_{j=0}^{l} z_j^* \in \partial\Big(\sum_{j=0}^{r} \mu_j\phi_j\Big)(\bar{z}) + \sum_{j=0}^{m} \nabla g_j(\bar{z})^*\psi_j. \tag{4.8}$$

For applications in this paper in the case of nonautonomous continuous-time systems we need the following modifications of the basic constructions (4.1), (4.3), and (4.4) for sets, functions, and set-valued mappings *depending on a parameter* t from a topological space T (in our case $T = [a,b]$).

Given $\Omega\colon T \rightrightarrows I\!R^n$ and $\bar{x} \in \Omega(\bar{t})$, we define the *extended normal cone* to $\Omega(\bar{t})$ at $\bar{x}$ by

$$\widetilde{N}(\bar{x};\Omega(\bar{t})) := \operatorname*{Limsup}_{(t,x) \overset{\operatorname{gph}\Omega}{\to} (\bar{t},\bar{x})} \widehat{N}(x;\Omega(t)). \tag{4.9}$$

For $\varphi\colon I\!R^n \times T \to \overline{I\!R}$ finite at $(\bar{x},\bar{t})$ and for $F\colon I\!R^n \times T \rightrightarrows I\!R^m$ with $\bar{y} \in F(\bar{x},\bar{t})$, the *extended subdifferential* of φ at $(\bar{x},\bar{t})$ and the *extended coderivative* of F at $(\bar{x},\bar{y},\bar{t})$ with respect to x are given, respectively, by

$$\widetilde{\partial}_x\varphi(\bar{x},\bar{t}) := \big\{x^* \in I\!R^n \big|\ (x^*,-1) \in \widetilde{N}((\bar{x},\varphi(\bar{x},\bar{t}));\operatorname{epi}\varphi(\bar{t}))\big\} \tag{4.10}$$

and, whenever $y^* \in I\!R^m$, by

$$\widetilde{D}^*_x F(\bar{x},\bar{y},\bar{t})(y^*) := \big\{x^* \in I\!R^n \big|\ (x^*,-y^*) \in \widetilde{N}((\bar{x},\bar{y});\operatorname{gph} F(\bar{t}))\big\}. \tag{4.11}$$

Note that the sets (4.9)–(4.11) may be bigger in some situations than the corresponding sets $N(\bar{x};\Omega(\bar{t}))$, $\partial_x\varphi(\bar{x},\bar{t})$, and $D^*_x F(\bar{x},\bar{y},\bar{t})(y^*)$, where the latter two sets stand for the subdifferential (4.3) of $\varphi(\cdot,\bar{t})$ at $\bar{x}$ and the coderivative (4.4) of $F(\cdot,\bar{t})$ at $(\bar{x},\bar{y},\bar{t})$, respectively. Efficient conditions ensuring equalities for these sets are discussed in [8, 9, 11].

It is not difficult to check that the extended constructions (4.9)–(4.11) are *robust* with respect to their variables, which is important for performing limiting procedures in what follows. In particular,

$$\widetilde{N}(\bar{x};\Omega(\bar{t})) = \operatorname*{Limsup}_{(t,x)\ \overset{\operatorname{gph}\Omega}{\to}\ (\bar{t},\bar{x})} \widetilde{N}(x;\Omega(t)). \tag{4.12}$$

Note also that the constructions (4.9)–(4.11) enjoy a full generalized differential calculus similar to one for (4.1), (4.3), and (4.4). We do not need this calculus in the present paper, however.

5 Necessary optimality conditions for difference-algebraic systems

In this section we derive necessary optimality conditions for the discrete approximation problems (P_N) by reducing them to those in Theorem 4.2 for nonsmooth mathematical programming problems with many geometric constraints.

Given $n \in I\!N$, consider problem (MP) with the decision vector

$$w := (x_0^N,\dots,x_{k+1}^N, z_0^N,\dots,z_{k+1}^N, v_0^N, v_1^N,\dots,v_k^N) \in I\!R^{n(3k+5)}$$

and the following data:

$$\begin{aligned}\phi_0(w) \ :=\ & \varphi(x_0^N,x_{k+1}^N) + |x_0^N - \bar{x}(a)|^2 + h_N \sum_{j=0}^{k} f\big(x_j^N, x_{j-N}^N, z_j^N, v_j^N, t_j\big)\\ & + \sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} \big|v_j^N - \dot{\bar{z}}(t)\big|^2\,dt,\end{aligned} \tag{5.1}$$

$$\begin{aligned}\phi_j(w) \ :=\ & \begin{cases} |x_j^N - \bar{x}(t_j)| - \varepsilon, & j = 1,\dots,k+1,\\ |z_{j-k-1}^N - \bar{z}(t_{j-k-1})| - \varepsilon, & j = k+2,\dots,2k+2,\end{cases}\\
\Lambda_j \ :=\ & \big\{(x_0^N,\dots,v_k^N)\ \big|\ v_j^N \in F(x_j^N, x_{j-N}^N, z_j^N, t_j)\big\}, \quad j = 0,\dots,k,\\
\Lambda_{k+1} \ :=\ & \big\{(x_0^N,\dots,v_k^N)\ \big|\ (x_0^N, x_{k+1}^N) \in \Omega_N\big\},\\
g_j(w) \ :=\ & z_{j+1}^N - z_j^N - h_N v_j^N, \quad j = 0,\dots,k,\\
h_j(w) \ :=\ & z_j^N - x_j^N - A x_{j-N}^N, \quad j = 0,\dots,k+1,\end{aligned}$$

where $x_j^N := c(t_j)$ for $j < 0$. Let $\bar{w}^N := (\bar{x}_0^N, \ldots, \bar{x}_{k+1}^N, \bar{z}_0^N, \ldots, \bar{z}_{k+1}^N, \bar{v}_0^N, \ldots, \bar{v}_k^N)$ be a given optimal solution to (MP). Applying Theorem 4.2 with the equality constraints given by (g_j, h_j), we find real numbers μ_j^N for $j = 0, \ldots, 2k+2$, vectors $w_j^* \in I\!R^{n(3k+5)}$ for $j = 0, \ldots, k+1$, vectors $\eta_j^N \in I\!R^n$ for $j = 0, \ldots, k$, as well as vectors $\psi_j^N \in I\!R^n$ for $j = 0, \ldots, k+1$, not all zero, such that conditions (4.5)–(4.8) are satisfied.

Taking $w_j^* := (x_{0,j}^*, \ldots, x_{k+1,j}^*, z_{0,j}^*, \ldots, z_{k+1,j}^*, v_{0,j}^*, \ldots, v_{k,j}^*) \in N(\bar{w}^N; \Lambda_j)$ for $j = 0, \ldots, k$, we observe that all but one components of w_j^* are zero and the remaining one satisfies

$$(x_{j,j}^*, x_{j-N,j}^*, z_{j,j}^*, v_{j,j}^*) \in N((\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N, \bar{v}_j^N); \operatorname{gph} F(\cdot,\cdot,\cdot,t_j)), \quad j = 0, \ldots, k.$$

Similarly notice that the condition $w_{k+1}^* \in N(\bar{z}^N; \Lambda_{k+1})$ is equivalent to

$$(x_{0,k+1}^*, x_{k+1,k+1}^*) \in N((\bar{x}_0^N, \bar{x}_{k+1}^N); \Omega_N)$$

with all the other components of w_{k+1}^* equal to zero. Employing Theorem 3.1 on the convergence of discrete approximations, we have $\phi_j(\bar{w}^N) < 0$ for $j = 1, \ldots, 2k+2$ whenever N is sufficiently large. Thus $\mu_j^N = 0$ for these indexes due to the complementary slackness conditions (4.6). Let $\lambda^N := \mu_0^N \geq 0$. Observe further that

$$\sum_{j=0}^{k} (\nabla g_j(\bar{w}^N))^* \eta_j^N = \big(0, \ldots, 0, \eta_0^N, \eta_0^N - \eta_1^N, \eta_{k-1}^N - \eta_k^N, \eta_k^N, -h_N\eta_0^N, \ldots, -h_N\eta_k^N\big),$$

and that one has

$$\begin{aligned}\sum_{j=0}^{k+1} (\nabla h_j(\bar{w}^N))^* \psi_j^N = \big(&- \psi_0^N + A^*\psi_N^N, -\psi_1^N + A^*\psi_{N+1}^N, \ldots, \\ &- \psi_{k-N+1}^N + A^*\psi_{k+1}^N, -\psi_{k-N+2}^N, \ldots, -\psi_{k+1}^N, \psi_0^N, \ldots, \psi_{k+1}^N, 0, \ldots, 0\big).\end{aligned}$$

From the subdifferential sum rule for ϕ_0 in (5.1) we deduce that

$$\begin{aligned}\partial\phi_0(\bar{w}^N) \quad \subset\, & \partial\varphi(\bar{x}_0^N, \bar{x}_{k+1}^N) + 2(\bar{x}_0^N - \bar{x}(a)) + h_N \sum_{j=0}^{k} \partial f(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N, \bar{v}_j^N, t_j) \\ & + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} 2\Big(\bar{v}_j^N - \dot{\bar{z}}(t)\Big)\, dt\end{aligned}$$

with ∂f standing here and in what follows for the basic subdifferential of f with respect to the first

four variables. Thus the inclusion (4.8) in Theorem 4.2 is equivalent to the relationships

$$\begin{cases} -x^*_{0,0} - x^*_{0,N} - x^*_{0,k+1} = \lambda^N u^N_0 + \lambda^N h_N \vartheta^N_0 + \lambda^N h_N \kappa^N_0 + \\ \qquad 2\lambda^N(\bar{x}^N_0 - \bar{x}(a)) - \psi^N_0 - A^*\psi^N_N, \\ -x^*_{j,j} - x^*_{j,j+N} = \lambda^N h_N \vartheta^N_j + \lambda^N h_N \kappa^N_j - \psi^N_j - A^*\psi^N_{j+N}, \quad j = 1,\dots,k-N+1, \\ -x^*_{j,j} = \lambda^N h_N \vartheta^N_j - \psi^N_j, \quad j = k-N+2,\dots,k, \\ -x^*_{k+1,k+1} = \lambda^N u^N_{k+1} - \psi^N_{k+1}, \\ -z^*_{j,j} = \lambda^N h_N \sigma^N_j + \psi^N_j + \eta^N_{j-1} - \eta^N_j, \quad j = 0,\dots,k, \\ -v^*_{j,j} = \lambda^N h_N \iota^N_j + \lambda^N \theta^N_j - h_N \eta^N_j, \quad j = 0,\dots,k, \end{cases} \tag{5.2}$$

with the notation

$$(u^N_0, u^N_{k+1}) \in \partial\varphi(\bar{x}^N_0, \bar{x}^N_{k+1}), \quad (\vartheta^N_j, \kappa^N_{j-N}, \sigma^N_j, \iota^N_j) \in \partial f(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{z}^N_j, \bar{v}^N_j, t_j),$$
$$\theta^N_j := 2\int_{t_j}^{t_{j+1}} \big(\bar{v}^N_j - \dot{\bar{z}}(t)\big)\, dt.$$

Based on the above relationships, we arrive at the following necessary optimality conditions for discrete-time problems (P_N), where $f_j(\cdot,\cdot,\cdot,\cdot) := f(\cdot,\cdot,\cdot,\cdot,t_j)$ and $F_j(\cdot,\cdot,\cdot) := F(\cdot,\cdot,\cdot,t_j)$.

Theorem 5.1 *Let $\bar{w}^N$ be an optimal solution to problem (P_N). Assume that the sets Ω and $\operatorname{gph} F_j$ are closed and that the functions φ and f_j are Lipschitz continuous around the points $(\bar{x}^N_0, \bar{x}^N_{k+1})$ and $(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{z}^N_j, \bar{v}^N_j)$, respectively, for all $j = 0,\dots,k$. Then there exist $\lambda^N \ge 0$, p^N_j $(j = 0,\dots,k+N+1)$, q^N_j $(j = -N,\dots,k+1)$, and r^N_j $(j = 0,\dots,k+1)$, not all zero, such that*

$$p^N_j = 0, \quad j = k+2,\dots,k+N+1, \tag{5.3}$$

$$q^N_j = 0, \quad j = k-N+1,\dots,k+1, \tag{5.4}$$

$$(p^N_0 + q^N_0, -p^N_{k+1}) \in \lambda^N \partial\varphi(\bar{x}^N_0, \bar{x}^N_{k+1}) + N((\bar{x}^N_0, \bar{x}^N_{k+1}); \Omega_N), \tag{5.5}$$

$$\begin{aligned} &\Big(\frac{P^N_{j+1} - P^N_j}{h_N}, \frac{Q^N_{j-N+1} - Q^N_{j-N}}{h_N}, \frac{r^N_{j+1} - r^N_j}{h_N}, -\frac{\lambda^N \theta^N_j}{h_N} + p^N_{j+1} + q^N_{j+1} + r^N_{j+1}\Big) \\ &\in \lambda^N \partial f_j(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{z}^N_j, \bar{v}^N_j) + N((\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{z}^N_j, \bar{v}^N_j); \operatorname{gph} F_j), \quad j = 1,\dots,k, \end{aligned} \tag{5.6}$$

with the notation

$$P^N_j := p^N_j + A^* p^N_{j+N}, \quad Q^N_j := q^N_j + A^* q^N_{j+N}.$$

Proof. Most of the proof has been actually done above, where we transformed the necessary optimality conditions for (MP) into the ones for (P_N) written in the form of nonsmooth mathematical

programming. What we need to do is to change the notation in the relationships of (5.2). Let us first denote

$$S_j^N := \begin{cases} \psi_j^N & \text{for } j = 1, \dots, k+1, \\ 0 & \text{for } j = k+2, \dots, k+N; \end{cases}$$

$$T_j^N := \begin{cases} \lambda^N \kappa_j^N + x^*_{j,j+N}/h_N & \text{for } j = 1, \dots, k-N+1, \\ 0 & \text{for } j = k-N+2, \dots, k; \end{cases}$$

$$\widetilde{r}_j^N := \eta_{j-1}^N \quad \text{for } j = 1, \dots, k+1.$$

It follows from (5.2) that

$$\begin{cases} S_j^N + A^* S_{j+N}^N - T_j^N = \lambda^N \vartheta_j^N + x^*_{j,j}/h_N, \\ T_{j-N}^N = \lambda^N \kappa_{j-N}^N + x^*_{j-N,j}/h_N, \\ \dfrac{\widetilde{r}_{j+1}^N - \widetilde{r}_j^N}{h_N} - S_j^N = \lambda^N \sigma_j^N + \dfrac{z^*_{j,j}}{h_N}, \\ -\lambda^N \theta_j^N/h_N + \widetilde{r}_{j+1}^N = \lambda^N \iota_j^N + v^*_{j,j}/h_N \end{cases} \tag{5.7}$$

for all $j = 1, \dots, k$. Define the sequences $\widetilde{p}_j^N$, and $\widetilde{q}_j^N$ by the recurrent formulae

$$\widetilde{p}_j^N := \widetilde{p}_{j+1}^N - S_j^N h_N \quad \text{with} \quad \widetilde{p}_j^N = 0 \quad \text{for} \quad j = k+2, \dots, k+N+1,$$
$$\widetilde{q}_j^N := \widetilde{q}_{j+1}^N - T_j^N h_N \quad \text{with} \quad \widetilde{q}_j^N = 0 \quad \text{for} \quad j = k-N+1, \dots, k+N+1.$$

Putting $q_j^N := \widetilde{q}_j^N + A^* \widetilde{q}_{j+N}^N$, we rewrite (5.7) as

$$\begin{cases} \dfrac{(\widetilde{p}_{j+1}^N - q_{j+1}^N) - (\widetilde{p}_j^N - q_j^N)}{h_N} + A^* \dfrac{(\widetilde{p}_{j+N+1}^N - q_{j+N+1}^N) - (\widetilde{p}_{j+N}^N - q_{j+N}^N)}{h_N} \\ = \lambda^N \vartheta_j^N + \dfrac{x^*_{j,j}}{h_N}, \quad j = 1, \dots, k, \\ \dfrac{(q_{j-N+1}^N + A^* q_{j+1}^N) - (q_{j-N}^N + A^* q_j^N)}{h_N} = \lambda^N \kappa_{j-N}^N + \dfrac{x^*_{j-N,j}}{h_N}, \quad j = 1, \dots, k, \\ \dfrac{\widetilde{r}_{j+1}^N - \widetilde{r}_j^N}{h_N} - \dfrac{\widetilde{p}_{j+1}^N - \widetilde{p}_j^N}{h_N} = \lambda^N \sigma_j^N + \dfrac{z^*_{j,j}}{h_N}, \quad j = 1, \dots, k, \\ -\lambda^N \theta_j^N/h_N + \widetilde{r}_{j+1}^N = \lambda^N \iota_j^N + \dfrac{v^*_{j,j}}{h_N}, \quad j = 1, \dots, k. \end{cases} \tag{5.8}$$

Letting finally

$$p_0^N := \lambda^N u_0^N + x^*_{0,k+1} - q_0^N,$$
$$p_j^N := \widetilde{p}_j^N - q_j^N \quad \text{for } j = 1, \dots, k+N+1,$$
$$r_j^N := \widetilde{r}_j^N - \widetilde{p}_j^N \quad \text{for } j = 1, \dots, k+1,$$

we arrive at all the relationships (5.3)–(5.6) and complete the proof of the theorem. $\triangle$

Corollary 5.2 *In addition to the assumptions in Theorem 5.1, suppose that the mapping F_j is bounded and Lipschitz continuous around $(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N)$ for each $j = 0, \ldots, k$. Then the conditions $\lambda^N \geq 0$ and (5.3)–(5.6) hold with $(\lambda^N, p_{k+1}^N, r_{k+1}^N) \neq 0$, i.e., one can let*

$$(\lambda^N)^2 + |p_{k+1}^N|^2 + |r_{k+1}^N|^2 = 1. \tag{5.9}$$

Proof. If $\lambda^N = 0$, then (5.6) together with (5.3) and (5.4) imply that

$$\Big(\frac{p_{k+1}^N - p_k^N}{h_N}, \frac{-q_{k-N}^N}{h_N}, \frac{r_{k+1}^N - r_k^N}{h_N}\Big) \in D^* F_k(\bar{x}_k^N, \bar{x}_{k-N}^N, \bar{z}_k^N, \bar{v}_k^N)(-p_{k+1}^N - r_{k+1}^N).$$

Assuming now that $p_{k+1}^N = 0$ and $r_{k+1}^N = 0$, we get

$$\Big(\frac{-p_k^N}{h_N}, \frac{-q_{k-N}^N}{h_N}, \frac{-r_k^N}{h_N}\Big) \in D^* F_k(\bar{x}_k^N, \bar{x}_{k-N}^N, \bar{z}_k^N, \bar{v}_k^N)(0),$$

which yields $p_k^N = 0$, $q_{k-N}^N = 0$, and $r_k^N = 0$ by Theorem 4.1. Repeating the above procedure, we arrive at the contradiction with the nontriviality assertion in Theorem 5.1. $\triangle$

6 Optimality conditions for differential-algebraic inclusions

In the concluding section of the paper we obtain the main results of this study that provide necessary optimality conditions for the original dynamic optimization problem (P) in both extended Euler-Lagrange and Hamiltonian forms involving generalized differential constructions of Section 4. Our major theorem establishes the following conditions of the *Euler-Lagrange type* derived by the limiting procedure from discrete approximations.

Theorem 6.1 *Let $(\bar{x}, \bar{z})$ be an optimal solution pair to problem (P) under hypotheses (H1)–(H4), where φ and $f(\cdot, \cdot, \cdot, \cdot, t)$ are assumed to be Lipschitz continuous instead of the plain continuity. Suppose also that (P) is stable with respect to relaxation. Then there exist a number $\lambda \geq 0$, piecewise continuous functions $p\colon [a, b+\Delta] \to \mathbb{R}^n$ and $q\colon [a-\Delta, b] \to \mathbb{R}^n$ (whose points of discontinuity are confined to multiples of the delay time Δ), and an absolutely continuous function $r\colon [a, b] \to \mathbb{R}^n$ such that $p(t) + A^* p(t+\Delta)$ and $q(t-\Delta) + A^* q(t)$ are absolutely continuous on $[a, b]$ and that the following conditions hold:*

$$\lambda + |p(b)| + |r(b)| = 1, \tag{6.1}$$

$$p(t) = 0 \ \text{ for } \ t \in (b, b+\Delta], \quad q(t) = 0 \ \text{ for } \ t \in (b-\Delta, b], \tag{6.2}$$

$$(p(a) + q(a), -p(b)) \in \lambda \partial \varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega), \tag{6.3}$$

$$
\begin{aligned}
&\Big(\frac{d}{dt}[p(t)+A^*p(t+\Delta)],\ \frac{d}{dt}[q(t-\Delta)+A^*q(t)],\dot{r}(t)\Big)\\
&\in \mathrm{co}\Big\{(u,v,w)\ \big|\ (u,v,w,p(t)+q(t)+r(t))\in\lambda\widetilde{\partial}f(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),\dot{\bar{z}}(t),t)\\
&+\widetilde{N}\big((\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),\dot{\bar{z}}(t));\mathrm{gph}\,F(t)\big)\Big\}\quad a.e.\ \ t\in[a,b].
\end{aligned}
\tag{6.4}
$$

Proof. To prove this theorem by the method of discrete approximations, we first construct a sequence of discrete-time problems (P_N) whose optimal solutions $(\bar{x}_N,\bar{z}_N)$ strongly approximate $(\bar{x},\bar{z})$ in the sense of Theorem 2.1. By necessary optimality conditions for $(\bar{x}_N,\bar{z}_N)$ from Corollary 5.2 we find $\lambda^N\ge 0$, p_j^N, q_j^N, and r_j^N satisfying relationships (5.3)–(5.9) for all $N\in\mathbb{N}$.

Without loss of generality we suppose that $\lambda^N\to\lambda$ as $N\to\infty$ for some $\lambda\ge 0$. Recall that the symbols $\bar{x}^N(t)$, $\bar{z}^N(t)$, $p^N(t)$, $q^N(t-\Delta)$, $r^N(t)$, $P^N(t)$, and $Q^N(t-\Delta)$ stand for the piecewise linear extensions of the corresponding discrete functions from Theorem 5.1 with their piecewise constant derivatives on the continuous-time interval $[a,b]$.

Considering θ_j from Theorem 5.1, we define $\theta^N(t):=\theta_j^N/h_N$ for $t\in[t_j,t_{j+1})$ as $j=0,\ldots,k$ and conclude by Theorem 2.1 that

$$
\begin{aligned}
\int_a^b|\theta^N(t)|\,dt&=\sum_{j=0}^k|\theta_j^N|\le 2\sum_{j=0}^k\int_{t_j}^{t_{j+1}}\Big|\dot{\bar{z}}(t)-\bar{v}_j^N\Big|\,dt\\
&=2\int_a^b\Big|\dot{\bar{z}}(t)-\dot{\bar{z}}^N(t)\Big|\,dt:=\nu_N\to 0\ \text{ as }\ N\to\infty.
\end{aligned}
$$

We may assume without loss of generality that

$$
\bar{v}^N(t):=\dot{\bar{z}}^N(t)\to\dot{\bar{z}}(t)\quad\text{and}\quad\theta^N(t)\to 0\ \text{ a.e. }\ t\in[a,b]\ \text{ as }N\to\infty.
$$

Let us estimate $(p^N(t),q^N(t-\Delta),r^N(t))$ for large N. Using (5.3) and (5.4), we derive from (5.6) the inclusions

$$
\begin{aligned}
&\Big(\frac{p_{j+1}^N-p_j^N}{h_N}-\lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N-q_{j-N}^N}{h_N}-\lambda^N\kappa_{j-N}^N,\ \frac{r_{j+1}^N-r_j^N}{h_N}-\lambda^N\sigma_j^N,\\
&-\frac{\lambda^N\theta_j^N}{h_N}+p_{j+1}^N+r_{j+1}^N-\lambda^N\iota_j^N\Big)\in N((\bar{x}_j^N,\bar{x}_{j-N}^N,\bar{z}_j^N,\bar{v}_j^N);\mathrm{gph}\,F_j)
\end{aligned}
$$

with some $(\vartheta_j^N,\kappa_{j-N}^N,\sigma_j^N,\iota_j^N)\in\partial f_j(\bar{x}_j^N,\bar{x}_{j-N}^N,\bar{z}_j^N,\bar{v}_j^N)$ for all $j=k-N+2,\ldots,k+1$. This means, by definition of the coderivative (4.4), that

$$
\begin{aligned}
&\Big(\frac{p_{j+1}^N-p_j^N}{h_N}-\lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N-q_{j-N}^N}{h_N}-\lambda^N\kappa_{j-N}^N,\ \frac{r_{j+1}^N-r_j^N}{h_N}-\lambda^N\sigma_j^N\Big)\\
&\in D^*F_j(\bar{x}_j^N,\bar{x}_{j-N}^N,\bar{z}_j^N,\bar{v}_j^N)\Big(\lambda^N\iota_j^N+\frac{\lambda^N\theta_j^N}{h_N}-p_{j+1}^N-r_{j+1}^N\Big)
\end{aligned}
$$

for such j. Thus it follows from Theorem 4.1 that

$$
\begin{aligned}
&\Big|\Big(\frac{p_{j+1}^N-p_j^N}{h_N}-\lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N-q_{j-N}^N}{h_N}-\lambda^N\kappa_{j-N}^N,\ \frac{r_{j+1}^N-r_j^N}{h_N}-\lambda^N\sigma_j^N\Big)\Big|\\
&\le\ell_F\Big|\lambda^N\iota_j^N+\frac{\lambda^N\theta_j^N}{h_N}-p_{j+1}^N-r_{j+1}^N\Big|
\end{aligned}
$$

for $j = k-N+2, \ldots, k+1$. Since $|(\vartheta_j^N, \kappa_{j-N}^N, \sigma_j^N, \iota_j^N)| \le \ell_f$ due to the Lipschitz continuity of f with modulus ℓ_f, we derive from the above that

$$\begin{aligned}
&|(p_j^N, q_{j-N}^N, r_j^N)| \le \ell_F|\theta_j^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N+1)|(p_{j+1}^N, q_{j-N+1}^N, r_{j+1}^N)| \\
&\le \ell_F|\theta_j^N| + (\ell_F h_N+1)\ell_F|\theta_{j+1}^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N+1)(\ell_F+1)h_N\ell_f \\
&+ (\ell_F h_N+1)^2|(p_{j+2}^N, q_{j-N+2}^N, r_{j+2}^N)| \le \cdots \\
&\le \exp[\ell_F(b-a)](1+\ell_f(\ell_F+1)/\ell_F + \ell_F\nu_N), \quad j = k-N+2, \ldots, k+1,
\end{aligned}$$

which implies the uniform boundedness of $\{(p_j^N, q_{j-N}^N, r_j^N) \,|\, j = k-N+2, \ldots, k+1\}$ and hence that of $(p^N(t), q^N(t-\Delta), r^N(t))$ on $[b-\Delta, b]$.

Next we consider $j = k-2N+2, \ldots, k-N+1$ and derive from (5.6) that

$$\begin{aligned}
&\Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N\vartheta_j^N, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N\kappa_{j-N}^N, \frac{r_{j+1}^N - r_j^N}{h_N} - \lambda^N\sigma_j^N\Big)\Big| \\
&\le \ell_F\Big|\lambda^N\iota_j^N + \frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N - r_{j+1}^N\Big| + \Big|\Big(\frac{A^*p_{j+N+1}^N - A^*p_{j+N}^N}{h_N}, \frac{A^*q_{j+1}^N - A^*q_j^N}{h_N}, 0\Big)\Big|.
\end{aligned}$$

This implies due to Theorem 4.1 and the uniform boundedness of the above vectors p_j^N and q_j^N by some constant $\alpha > 0$, and so

$$\begin{aligned}
&\Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N\vartheta_j^N, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N\kappa_{j-N}^N, \frac{r_{j+1}^N - r_j^N}{h_N} - \lambda^N\sigma_j^N\Big)\Big| \\
&\le \ell_F\Big|\lambda^N\iota_j^N + \frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N - r_{j+1}^N\Big| + \frac{\alpha}{h_N}
\end{aligned}$$

for $j = k-2N+2, \ldots, k-N+1$. Therefore

$$\begin{aligned}
&|(p_j^N, q_{j-N}^N, r_j^N)| \le \ell_F|\theta_j^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N+1)|(p_{j+1}^N, q_{j-N+1}^N, r_{j+1}^N)| + (\ell_F h_N+1)\alpha \\
&\le \ell_F|\theta_j^N| + (\ell_F h_N+1)\ell_F|\theta_{j+1}^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N+1)(\ell_F+1)h_N\ell_f \\
&+ (\ell_F h_N+1)(\ell_F+1)\alpha + (\ell_F h_N+1)^2|(p_{j+2}^N, q_{j-N+2}^N, r_{j+2}^N)| \le \cdots \\
&\le \exp[\ell_F(b-a)](1+(\ell_f+\alpha)(\ell_F+1)/\ell_F + \ell_F\nu_N), \quad j = k-2N+2, \ldots, k-N+1.
\end{aligned}$$

This shows that p_j^N, q_{j-N}^N, and r_j^N are uniformly bounded for $j = k-2N+2, \ldots, k-N+1$, and hence the sequence $\{p^N(t), q^N(t-\Delta), r^N(t)\}$ is uniformly bounded on $[b-2\Delta, b-\Delta]$. Repeating the above procedure, we conclude that both sequences $\{p^N(t), q^N(t-\Delta), r^N(t)\}$ and $\{P^N(t), Q^N(t-\Delta)\}$ are uniformly bounded on the whole interval $[a, b]$.

Next we estimate $\big(\dot{P}^N(t), \dot{Q}^N(t-\Delta), \dot{r}^N(t)\big)$ on $[a, b]$ using (5.6) and Theorem 4.1. This yields, for $t_j \le t < t_{j+1}$ with $j = 0, \ldots, k$, that

$$\begin{aligned}
&|(\dot{P}^N(t), \dot{Q}^N(t-\Delta), \dot{r}^N(t))| = \Big|\Big(\frac{P_{j+1}^N - P_j^N}{h_N}, \frac{Q_{j-N+1}^N - Q_{j-N}^N}{h_N}, \frac{r_{j+1}^N - r_j^N}{h_N}\Big)\Big| \\
&\le \ell_F\Big|\lambda^N\iota_j^N + \frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N - r_{j+1}^N\Big| + \ell_f \\
&\le \ell_F|\theta^N| + \ell_F|p_{j+1}^N| + \ell_F|q_{j+1}^N| + \ell_F|r_{j+1}^N| + (\ell_F+1)\ell_f.
\end{aligned}$$

Thus the sequence $\{\dot{P}^N(t), \dot{Q}^N(t-\Delta), \dot{r}^N(t)\}$ is weakly compact in $L^1[a,b]$. Taking the whole sequence of $N \in I\!N$ without loss of generality, we find three absolutely continuous functions $P(\cdot)$, $Q(\cdot-\Delta)$, and $r(\cdot)$ on $[a,b]$ such that

$$\dot{P}^N(t) \to \dot{P}(t), \quad \dot{Q}^N(t-\Delta) \to \dot{Q}(t-\Delta), \quad \dot{r}^N(t) \to \dot{r}(t) \ \text{ weakly in } \ L^1[a,b]$$

and $P^N(t) \to P(t)$, $Q^N(t-\Delta) \to Q(t-\Delta)$, $r^N(t) \to r(t)$ uniformly on $[a,b]$ as $N \to \infty$. Since $p^N(t)$ and $q^N(t-\Delta)$ are uniformly bounded on $[a, b+\Delta]$, they surely converge to some functions $p(t)$ and $q(t-\Delta)$ weakly in $L^1[a,b+\Delta]$. Taking into account the above convergence of $P^N(t)$ and $Q^N(t-\Delta)$, we get that $p(\cdot)$ and $q(\cdot)$ satisfy (6.2), that

$$P(t) = p(t) + A^*p(t+\Delta), \quad Q(t-\Delta) = q(t-\Delta) + A^*q(t), \quad t \in [a,b],$$

and that $p(t)$ and $q(t)$ are piecewise continuous on $[a, b+\Delta]$ and $[a-\Delta, b]$, respectively, with possible discontinuity (from the right) at the points $b - i\Delta$ at $i = 0, 1, \ldots$. Conditions (6.1) and (6.3) follow by passing to the limit from (5.9) and (5.5), respectively, by taking into account the robustness of the basic subdifferential (4.3) and the normal cone (4.1).

It remains to justify the Euler-Lagrange inclusion (6.4). To furnish this, we rewrite the discrete Euler-Lagrange inclusion (5.6) in the form

$$\begin{aligned}
&(\dot{P}^N(t), \dot{Q}^N(t-\Delta), \dot{r}^N(t)) \\
&\in \Big\{(u,v,w)\Big| \ \Big(u, v, w, p^N(t_{j+1}) + q^N(t_{j+1}) + r^N(t_{j+1}) - \frac{\lambda^N \theta_j^N}{h_N}\Big) \\
&\in \lambda^N \partial f(\bar{x}^N(t_j), \bar{x}^N(t_j-\Delta), \bar{z}^N(t_j), \bar{v}_j^N, t_j) + N\big((\bar{x}^N(t_j), \bar{x}^N(t_j-\Delta), \bar{z}^N(t_j), \bar{v}_j^N); \operatorname{gph} F_j\big)\Big\}
\end{aligned} \tag{6.5}$$

for $t \in [t_j, t_{j+1}]$ with $j = 0, \ldots, k$. By the classical Mazur theorem there is a sequence of convex combinations of the functions $(\dot{P}^N(t), \dot{Q}^N(t-\Delta), \dot{r}^N(t))$ that converges to $(\dot{P}(t), \dot{Q}(t-\Delta), \dot{r}(t))$ for a.e. $t \in [a,b]$. Passing the limit in (6.5) and taking into account the pointwise convergence of $\theta^N(t)$ and $\bar{v}^N(t)$ established above, as well as the constructions of the extended normal cone (4.9) and the extended subdifferential (4.10) and their robustness property (4.12) with respect to all variables and parameters, we arrive at (6.4) and complete the proof of the theorem. $\triangle$

Observe that for the *Mayer problem* (P_M), which is (1.1)–(1.5) with $f = 0$, the generalized Euler-Lagrange inclusion (6.4) is equivalently expressed in terms of the extended coderivative (4.11) with respect to the first three variables of $F = F(x,y,z,t)$, i.e., in the form

$$\begin{aligned}
&\Big(\frac{d}{dt}[p(t) + A^*p(t+\Delta)], \ \frac{d}{dt}[q(t-\Delta) + A^*q(t)], \dot{r}(t)\Big) \\
&\in \operatorname{co} \widetilde{D}^*_{x,y,z} F\big(\bar{x}(t), \bar{x}(t-\Delta), \bar{z}(t), \dot{\bar{z}}(t)\big)\big(-p(t) - q(t) - r(t)\big) \ \text{ a.e. } \ t \in [a,b].
\end{aligned} \tag{6.6}$$

It turns out that the extended Euler-Lagrange inclusion obtained above implies, under the *relaxation stability* of the original problems, two other principal optimality conditions expressed

in terms of the Hamiltonian function built upon the mapping F in (1.2). The first condition called the extended *Hamiltonian inclusion* is given below in terms of a *partial convexification* of the basic subdifferential (4.3) for the Hamiltonian function. The second one is an analogue of the classical *Weierstrass-Pontryagin maximum condition* for the differential-algebraic inclusions under consideration. Recall that an analogue of the Maximum Principle (centered around the maximum condition) does *not generally hold* for differential-algebraic systems, even in the case of optimal control problems governed by smooth functional-differential equations of neutral type that are a special case of (P).

The following relationships between the extended Euler-Lagrange and Hamiltonian inclusions are based on Rockafellar's dualization theorem [16] (see also [19, Section 7.6] for another, more simple proof) that concerns subgradients of abstract Lagrangian and Hamiltonian associated with set-valued mappings *regardless of the dynamics* in (1.2). For simplicity we consider the case of the Mayer problem (P_M) for autonomous differential-algebraic systems. Then the *Hamiltonian* function for F in (1.2) is defined by

$$H(x,y,z,p) := \sup\big\{\langle p,v\rangle \big|\, v \in F(x,y,z)\big\}. \tag{6.7}$$

Corollary 6.2 *Let $(\bar{x},\bar{z})$ be an optimal solution pair to the Mayer problem (P_M) for the autonomous delayed differential-algebraic inclusion* (1.2) *under the assumptions of Theorem 6.1. Then there exist a number $\lambda \geq 0$, piecewise continuous functions $p\colon [a,b+\Delta] \to I\!R^n$ and $q\colon [a-\Delta,b] \to I\!R^n$ (whose points of discontinuity are confined to multiples of the delay time Δ), and an absolutely continuous function $r\colon [a,b] \to I\!R^n$ such that $p(t)+A^*p(t+\Delta)$ and $q(t-\Delta)+A^*q(t)$ are absolutely continuous on $[a,b]$ and, besides* (6.1)-(6.4), *one has the extended Hamiltonian inclusion*

$$\begin{aligned}&\Big(\frac{d}{dt}[p(t)+A^*p(t+\Delta)],\ \frac{d}{dt}[q(t-\Delta)+A^*q(t)],\dot{r}(t)\Big)\\ &\in \operatorname{co}\Big\{(u,v,w)\ \Big|\ \Big(-u,-v,-w,\dot{\bar{z}}(t)\Big) \in \partial H\big(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),p(t)+q(t)+r(t)\big)\Big\}\end{aligned} \tag{6.8}$$

and the maximum condition

$$\big\langle p(t)+q(t)+r(t),\dot{\bar{z}}(t)\big\rangle = H\big(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),p(t)+q(t)+r(t)\big) \tag{6.9}$$

for a.e. $t \in [a,b]$. If moreover F is convex-valued around $(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t))$, then (6.8) *is equivalent to the Euler-Lagrange inclusion*

$$\begin{aligned}&\Big(\frac{d}{dt}[p(t)+A^*p(t+\Delta)],\ \frac{d}{dt}[q(t-\Delta)+A^*q(t)],\dot{r}(t)\Big)\\ &\in \operatorname{co} D^*F\big(\bar{x}(t),\bar{x}(t-\Delta),\bar{z}(t),\dot{\bar{z}}(t)\big)\big(-p(t)-q(t)-r(t)\big)\ \ \text{a.e. } t \in [a,b],\end{aligned} \tag{6.10}$$

which automatically implies the maximum condition (6.9) *in this case.*

Proof. Since (P_M) is stable with respect to relaxation, the pair $(\bar{x}, \bar{z})$ is an optimal solution to the relaxed problem (R_M) whose only difference from (P_M) is that the delayed differential-algebraic inclusion (1.2) is replaced by its convexification (3.4). By Theorem 6.1 the optimal solution $(\bar{x}, \bar{z})$ satisfies conditions (6.1)–(6.4) and the relaxed counterpart of (6.6), which is the same as (6.10) in this case with F replaced by the convex hull $\operatorname{co} F$. According to [16, Theorem 3.3] and [19, Theorem 7.6.5] one has

$$\begin{aligned} &\operatorname{co}\Big\{(u,v,w)\Big|\ (u,v,w,p) \in N((x,y,z,q);\operatorname{gph}(\operatorname{co} F)\Big\} \\ &\subset \operatorname{co}\Big\{(u,v,w)\Big|\ (-u,-v,-w,q) \in \partial H_R(x,y,z,p)\Big\}, \end{aligned} \tag{6.11}$$

where H_R stands for the Hamiltonian (6.7) of the relaxed system, i.e., with F replaced by $\operatorname{co} F$. It is easy to check that $H_R = H$. Thus the extended Euler-Lagrange inclusion for the relaxed system implies the extended Hamiltonian inclusion (6.8), which surely yields the maximum condition (6.9). When F is *convex-valued*, (6.8) and (6.10) are *equivalent* due to the equality in (6.11) proved in [16]. This completes the proof of the corollary. △

Acknowledgments. The authors are grateful to two anonymous referees for valuable suggestions and remarks that helped us to improve the original presentation.

References

[1] K.E. Brennan, S.L. Campbell and L.R. Pretzold, *Numerical Solution of Initial Value Problems in Edifferential-Algebraic Equations.* North-Holland, New York (1989).

[2] E.N. Devdariani and Yu.S. Ledyaev, Maximum principle for implicit control systems. *Appl. Math. Optim.* **40**(1999) 79–103.

[3] A.L. Dontchev and E.M. Farhi, Error estimates for discretized differential inclusions. *Computing* **41** (1989) 349–358.

[4] M. Kisielewicz, *Differential Inclusions and Optimal Control.* Kluwer, Dordrecht (1991).

[5] B.S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.* **40** (1976) 960–969.

[6] B.S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control,* Nauka, Moscow (1988).

[7] B.S. Mordukhovich, Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. *Trans. Amer. Math. Soc.* **340** (1993) 1–35.

[8] B.S. Mordukhovich, Discrete approximations and refined Euler–Lagrange conditions for nonconvex differential inclusions. *SIAM J. Control Optim.* **33** (1995) 882–915.

[9] B.S. Mordukhovich, J.S. Treiman and Q.J. Zhu, An extended extremal principle with applications to multiobjective optimization. *SIAM J. Optim.* **14** (2003) 359–379.

[10] B.S. Mordukhovich and R. Trubnik, Stability of discrete approximation and necessary optimality conditions for delay-differential inclusions. *Ann. Oper. Res.* **101** (2001) 149-170.

[11] B.S. Mordukhovich and L. Wang, Optimal control of constrained delay-differential inclusions with multivalued initial condition. *Control and Cybernetics* **32** (2003) 585–609.

[12] B.S. Mordukhovich and L. Wang, Optimal control of neutral functional-differential inclusions. *SIAM J. Control Optim.* (to appear).

[13] B.S. Mordukhovich and L. Wang, Optimal control of differential-algebraic inclusions, in *Optimal Control, Stabilization, and Nonsmooth Analysis*, M. de Queiroz et al., eds., Lectures Notes in Control and Information Sciences, Vol. 301, pp. 73–83, Springer-Verlag, Heidelberg (2004).

[14] M. d. R. de Pinho and R.B. Vinter, Necessary conditions for optimal control problems involving nonlinear differential algebraic equations. *J. Math. Anal. Appl.* **212** (1997) 493–516.

[15] C. Pantelides, D. Gritsis, K. P. Morison and R.W.H. Sargent, The mathematical modelling of transient systems using differential-algebraic equations. *Comput. Chem. Engrg.* **12** (1988) 449–454.

[16] R.T. Rockafellar, Equivalent subgradient versions of Hamiltonian and Euler–Lagrange conditions in variational analysis. *SIAM J. Control Optim.* **34** (1996) 1300–1314.

[17] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis.* Springer-Verlag, Berlin (1998).

[18] G.V. Smirnov, *Introduction to the Theory of Differential Inclusions.* American Mathematical Society, Providence, RI (2002).

[19] R.B. Vinter, *Optimal Control.* Birkhäuser, Boston (2000).

[20] J. Warga, *Optimal Control of Differential and Functional Equations.* Academic Press, New York (1972).