Abstract—We present a stability criterion for switched systems which involves Lie brackets of the individual vector fields. We show that a switched system generated by a pair of globally asymptotically stable nonlinear vector fields which span a third-order nilpotent Lie algebra is globally asymptotically stable under arbitrary switching. This generalizes a known fact for switched linear systems and provides a partial solution to the open problem posed in [1].

To prove the result, we consider an optimal control problem which consists of finding the “most unstable” trajectory for an associated control system. We use the Agrachev-Gamkrelidze second-order maximum principle to show that there exists an optimal control that is piecewise continuous with no more than three switches over any interval of time. By construction, our criterion also holds for the more general case of differential inclusions.

Keywords: Switched nonlinear system, global asymptotic stability, Lie bracket, optimal control, maximum principle, differential inclusion.

I. INTRODUCTION

Consider the differential inclusion (DI)
\[ \dot{x} \in \text{co}\{f_0(x), f_1(x)\} \]  
(1)
where \( f_0, f_1 : \mathbb{R}^n \to \mathbb{R}^n \) are two analytic vector fields and \( \text{co} \) denotes the convex hull. A solution of (1) is an absolutely continuous function \( x(\cdot) : \mathbb{R} \to \mathbb{R}^n \) satisfying (1) for (almost) all \( t \). In particular, the set of solutions of (1) includes all the solutions of the \textit{switched system}
\[ \dot{x} = f_\sigma(x), \]  
(2)
where \( \sigma : [0, \infty) \to \{0, 1\} \) is a piecewise constant function of time, called a \textit{switching signal}. Switched systems have numerous applications and present a subject of extensive ongoing research (see, e.g., [2]).

An important special case, which we will occasionally use for illustration, is when the given vector fields are linear: \( f_1(x) = A_1x \), with \( A_i \in \mathbb{R}^{n \times n} \). Then (2) becomes the \textit{switched linear system}
\[ \dot{x} = A_\sigma x, \]  
(3)
The DI (1) is called \textit{globally asymptotically stable} (GAS) if there exists a class \( K \) function \( \beta \) such that for every initial condition \( x(0) \) every solution of (1) satisfies
\[ |x(t)| \leq \beta(|x(0)|, t) \quad \forall t \geq 0. \]  
(4)

This work was supported by the ISF under grant number 199/03.

Corresponding author: Dr. Michael Margaliot, School of Electrical Engineering-Systems, Tel Aviv University, Israel 69978; Homepage: www.eng.tau.ac.il/~michaelm; Email: michaelm@eng.tau.ac.il

1Recall that a function \( \alpha : [0, \infty) \to [0, \infty) \) is said to be of class \( K \) if it is continuous, strictly increasing, and \( \alpha(0) = 0 \). A function \( \beta : [0, \infty) \times [0, \infty) \to [0, \infty) \) is said to be of class \( K \) if \( \beta(\cdot, t) \) is of class \( K \) for each fixed \( t \geq 0 \) and \( \beta(s, t) \) decreases to 0 as \( t \to \infty \) for each fixed \( s \geq 0 \).

If the DI is GAS, then the switched system (2) is globally asymptotically stable, uniformly over the set of all switching signals, because solutions of (2) are contained in those of (1). We then say that (2) is \textit{globally uniformly asymptotically stable} (GUAS). Determining a necessary and sufficient condition for GUAS is a formidable challenge. In fact, for the special case (3) this is equivalent to solving one of the oldest open problems in the theory of control: the problem of absolute stability (see, e.g., [3]).

It is well known and easy to demonstrate that global asymptotic stability of the individual subsystems \( \dot{x} = f_i(x) \) is necessary but \textit{not} sufficient for GUAS of the switched system (2). In this paper, we are concerned with the problem of identifying conditions for the individual subsystems—apart from the obviously necessary requirement as to their global asymptotic stability—which guarantee GUAS of (2). This problem has received considerable attention in the literature; see [2, Chapter 2] for some available results.

The difficulty in analyzing the stability of (2) is that the switched system admits an infinite number of solutions for each initial condition. A natural idea is to try to characterize the “worst-case” (that is, the “most unstable”) switching law, and then analyze the behavior of the unique trajectory produced by this law. Pyatnitskiy and Rapoport [4], [5] developed a variational approach to describe the “worst-case” switching law, and the corresponding trajectory, for the switched linear system (3). Applying the Maximum Principle (MP), they derived an implicit characterization of this switching law in terms of a two-point boundary value problem. More recently, the same variational problem was addressed using a dynamic programming approach [6]. For the particular case of second-order switched linear systems, this approach yields an easily verifiable necessary and sufficient condition for GUAS [7] (see also [8] as well as the related work [9]).

Another particularly promising research avenue is to explore the role of commutation relations among the subsystems being switched. We now briefly recall available results, starting with the case of the switched linear system (3). The commutator, or \textit{Lie bracket}, is defined as \( [A_0, A_1] := A_0 A_1 - A_1 A_0 \). First, suppose that the matrices commute: \( [A_0, A_1] = 0 \). In this case,
\[ e^{A_1 t} e^{A_0 t} e^{A_0 t} e^{A_1 t} x_0 = e^{A_0 (t_3 + t_1)} e^{A_1 (t_4 + t_2)} x_0 = e^{A_0 (t_3 + t_1)} e^{A_1 (t_4 + t_2)} x_0 \]
so we can replace any solution of (3) with \( s = 3 \) switches with an equivalent solution that has a single switch. Obviously, this property actually holds for any \( s > 1 \). If \( A_0 \) and \( A_1 \) are Hurwitz matrices, then it is easy to see that this property implies GUAS of the system. It was shown in [10]
that the above statement remains true if the Lie bracket condition is relaxed to $[A_0, [A_0, A_1]] = [A_1, [A_0, A_1]] = 0$.

It is possible to generalize this idea using a Lie-algebraic approach. We say that the Lie algebra spanned by the pair $A_0, A_1$ is $k$th-order nilpotent if all iterated Lie brackets containing $k + 1$ terms vanish. Thus, the two results above can be restated as: if the Lie algebra is first-order or second-order nilpotent, then the switched system is GUAS. This result was extended in [11] to nilpotent (in fact, solvable) matrix Lie algebras for arbitrary $k$ by using simultaneous triangularization (Lie’s Theorem) (see also [12]).

The nonlinear switched system (2) is much less thoroughly understood. To tackle the global stability question, one can try to inspect commutation relations between the nonlinear vector fields $f_0, f_1$. The Lie bracket is now defined as

$$[f_1, f_0](x) := \frac{\partial f_1(x)}{\partial x} f_0(x) - \frac{\partial f_0(x)}{\partial x} f_1(x).$$

We say that the Lie algebra spanned by $f_0$ and $f_1$ is $k$th-order nilpotent if all iterated Lie brackets containing $k + 1$ terms vanish. It turns out that if the two vector fields commute (i.e., the Lie algebra is first-order nilpotent), then global asymptotic stability of the individual subsystems still implies GUAS of the switched system (2) [13].

Until very recently, all attempts to formulate global asymptotic stability criteria valid beyond the commuting nonlinear case have been unsuccessful. This is due to the fact that the methods employed to obtain the corresponding results for switched linear systems do not seem to apply. These issues are explained in [1], where this is presented as an open problem which seems to require a different approach altogether.

It is a well-known fact that Lie brackets play an essential role in the MP (see, e.g., [14],[15]). This suggests that the two approaches for stability analysis described above are actually related. Indeed, it was shown in [16] that it is possible to merge the variational approach with a Lie-algebraic analysis of the “worst-case” switching law. It follows from the MP that the “worst-case” switching law for the switched nonlinear system (2) is governed by the signs of a suitable function of time $\varphi : \mathbb{R} \to \mathbb{R}$. If $[f_0, [f_0, f_1]](x) = [f_1, [f_0, f_1]](x) = 0$, then $\varphi(t) = 0$, so $\varphi(t) = at + b$. If $a \neq 0$ or $b \neq 0$ then it is clear that the number of switches is bounded by one. The singular case $a = b = 0$ can be circumvented using a construction due to Sussmann [17] and it is possible to show that an optimal control, with no more than two switches, always exists. Then, GUAS of the switched system can be deduced from global asymptotic stability of the individual subsystems. This approach provided the first stability criterion for switched nonlinear systems which involves Lie brackets of the individual vector fields but does not require that these vector fields commute.

In this paper, we extend this approach to the case where the Lie algebra spanned by $f_0$ and $f_1$ is third-order nilpotent.

Note that there are two definitions of the Lie bracket in the literature which are identical up to a change of sign. The definition we use is consistent with the definition of the Lie bracket for matrices.

Our main result is that this property implies that the DI (1) is GAS. An analysis of the arguments in [16] shows that this result cannot be derived using the (classical) MP. Hence, we apply a new approach based on a powerful second-order MP developed by Agrachev and Gamkrelidze [18] (see also [19]).

The remainder of the paper is organized as follows. In the next section, we formulate the “worst-case” optimal control problem associated with (2). Our main result is stated in Section III. Section IV proves the main result for the particular case of bang-bang controls. Section V completes the proof by showing that it holds for controls containing singular arcs, as well. Section VI contains some concluding remarks.

II. OPTIMAL CONTROL APPROACH

Our starting point is to rewrite the differential inclusion (1) as the control system with drift

$$\dot{x} = f(x) + g(x)u, \quad u \in \mathcal{U}$$

(6)

where

$$f(x) := f_0(x), \quad g(x) := f_1(x) - f_0(x)$$

(7)

and $\mathcal{U}$ is the set of measurable functions $u(\cdot) : \mathbb{R} \to [0, 1]$.

It is obvious that every solution of (6) is also a solution of (1). It follows from Filippov’s Selection Lemma (see, e.g., [20, Section 14, Corollary 1] or [21, Theorem 2.3.13]) that every solution of (1) is also a solution of (6). Thus, the set of solutions of the control system (6) coincides with the set of solutions of the differential inclusion (1).

Note that trajectories of the switched system (2) correspond to piecewise constant controls taking values in \{0, 1\}. In particular, setting $u \equiv 0$ ($u \equiv 1$) in (6) yields $\dot{x} = f_0(x)$ [$\dot{x} = f_1(x)$]. We also remark that the switched linear system (3) is associated in this way with the bilinear control system $\dot{x} = A_0 x + (A_1 - A_0) x u$.

Fix an arbitrary point $p \in \mathbb{R}^n$, and let $x(\cdot; p, u)$ denote the solution of the system (6) with initial condition $x(0) = p$ corresponding to a control $u \in \mathcal{U}$. Since the right-hand side of (6) is bounded on every bounded ball in $\mathbb{R}^n$, there exists a largest time $T_{\text{max}} \in (0, \infty]$ (that depends on $|p|$) such that $x(\cdot; p, u)$ is well defined for all $u \in \mathcal{U}$ and all $t \in [0, T_{\text{max}})$. Picking a positive final time $t_f < T_{\text{max}}$, we define

$$J(t_f, p, u) := |x(t_f; p, u)|^2,$$

and pose the following optimal control problem.

Problem 1: Find a control $u \in \mathcal{U}$ that maximizes $J$ along the solutions of (6).

It follows from [22, §7, Theorem 3] that this problem is well posed, i.e., an optimal control $u^*$ does exist.

The intuitive interpretation of Problem 1 is clear: find a control that “pushes” the corresponding trajectory $x^*$ as far away from the origin as possible (from a given initial condition in a given amount of time). If we can show that $x^*$
satisfies the bound (4), then the same will be true for any other trajectory of the DI, and stability of the DI—as well as the switched system—will be established.

III. MAIN RESULT

We are now ready to state our main result. For $c \in [0,1]$, we denote

$$h_c(x) := f(x) + cg(x).$$

For $\tau, k > 0$ we let $PC(\tau, k) \subset U$ denote the set of piecewise continuous controls that have no more than $k$ switches over the interval $[0, \tau]$.

Theorem 1: Suppose that for any fixed $c \in [0,1]$ the vector field $h_c$ is analytic, and the Lie algebra spanned by $f_0$ and $f_1$ is third-order nilpotent. Then, there exists an optimal control $u^*$ that satisfies $u^* \in PC(t_f, 3)$ for any $t_f < T_{max}$.

The next result will allow us to apply Theorem 1 to the stability analysis of (1).

Proposition 1: Suppose that: (1) for any fixed $c \in [0,1]$ the vector field $h_c$ is analytic, and the system $\dot{x} = h_c(x)$ is GAS; and (2) there exists an absolute number $d$ and an optimal control $u^*$ that satisfies $u^* \in PC(t_f, d)$ for any $t_f < T_{max}$. Then, the DI (1) is GAS.

Proof: Consider Problem 1 with arbitrary $p$ and $t_f < T_{max}(p)$. Let $u^* : [0,t_f) \rightarrow [0,1]$ be the optimal control with no more than $d$ switches. The interval $[0,t_f]$ is thus divided into a maximum of $r := d + 1$ subintervals, on each of which the optimal trajectory $x^*$ satisfies the equation $\dot{x} = h_{c_i}(x)$, for some $c_i \in [0,1]$. This system is globally asymptotically stable, hence the solutions of $\dot{x} = h_{c_i}(x)$ satisfy $|x(t)| \leq \beta(|x(0)|, t)$ for some class $KL$ function $\beta$. Define $\beta := \max\{\beta_i\} \in KL$ and let $\alpha$ be the class $K$ function defined by $\alpha(r) := \beta(r, 0)$.

Considering the $r$ subintervals of $[0,t_f]$, we see that the length of at least one of these subintervals is no smaller than $t_f/r$. Considering all possible locations of this subinterval relative to the others, it is easy to verify that

$$|x^*(t_f)| \leq \beta(|p|, t_f) \tag{8}$$

where

$$\beta(r, t) := \max \{\alpha^k(\beta(\alpha^l(r), t/r)) : k, l \geq 0, k + l = d\}$$

is a class $KL$ function (here $\alpha^k$ denotes the composition of $\alpha$ with itself $k$ times).

Since $x^*$ is an optimal trajectory for Problem 1 it is clear that every other solution of (6) with $x(0) = p$ also satisfies $|x(t_f)| \leq \beta(|p|, t_f)$. In view of the bound $\beta(|p|, t_f) \leq \beta(|p|, 0)$ and the fact that $p$ and $t_f < T_{max}$ were arbitrary, we conclude that all solutions of (6) are bounded and exist globally in time. In other words, $t_f$ could be an arbitrary positive number, and all solutions of (6) satisfy the bound (8). The same is automatically true for the solutions of the DI (1), which in turn include all solutions of the switched system (2).

Loosely speaking, Proposition 1 states that to obtain instability in a DI, composed of GAS subsystems, we must never stop switching.

Combining Theorem 1 and Proposition 1 yields the following.

Corollary 1: Suppose that: (1) for any fixed $c \in [0,1]$ the vector field $h_c$ is analytic and the system $\dot{x} = h_c(x)$ is GAS; and (2) the Lie algebra spanned by $f_0$ and $f_1$ is third-order nilpotent. Then, the DI (1) is globally asymptotically stable, and in particular the switched system (2) is GUAS.

The remainder of this paper is devoted to the proof of Theorem 1. We divide the proof into two parts. First, we bound the number of switches for any optimal control that is bang-bang and then we consider the singular case.

IV. THE BANG-BANG CASE

In this section, we assume that $\varphi(t) = 0$ holds only on isolated points. This implies that any optimal control $u^*$ is bang-bang.

A. First-order analysis

The next result follows from applying the MP to Problem 1 (a simple proof of the MP for the problem considered here can be found in [23]). We use the notation $Df := \frac{\partial f}{\partial x}$.

Theorem 2: Let $u^*$ be an optimal control for Problem 1. Define the costate $\lambda : [0,t_f] \rightarrow \mathbb{R}^n$ as the (unique) solution of the equation

$$\dot{\lambda} = -(Df(x^*) + u^*Dg(x^*))^T \lambda, \quad \lambda(t_f) = x^*(t_f), \tag{9}$$

and let $\varphi(t) := \lambda^T(t)g(x^*(t))$. Then

$$u^*(t) = \begin{cases} 1 & \text{if } \varphi(t) > 0 \\ 0 & \text{if } \varphi(t) < 0. \end{cases} \tag{10}$$

The next result will allow us to relate $\varphi(t)$ to the Lie algebra spanned by $f$ and $g$. For easy reference, we state it formally.

Fact 1: If $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a smooth vector field, and

$$\psi(t) := \lambda^T(t)h(x(t)), \tag{11}$$

then

$$\dot{\psi}(t) = \lambda^T(t)(h, f) + (u(t)[h, g])(x(t)). \tag{12}$$

The proof follows immediately from differentiating the (absolutely continuous) function $\psi$ and using (6), (9) and (5).

Proposition 2: Suppose the Lie algebra spanned by the vector fields $f$ and $g$ is $k$th-order nilpotent for some integer $k > 0$. If $I \subset [0,t_f]$ is an interval such that $u^*(t) = c$ for all $t \in I$, then the restriction of $\varphi(t)$ to $I$ is a polynomial in $t$ with degree $< k$.

Proof: Fact 1 implies that the derivative of $\psi$, defined in (11), satisfies $\psi(t) = \lambda^T(t)(h, f) + (u(t)[h, g])(x(t)), \forall t \in I$. Hence, if $h$ is an iterated Lie bracket of $f$ and $g$ containing $s$ terms, then $\psi(t)$ contains iterated Lie brackets of $f$ and $g$ with $s + j$ terms. Noting that $\varphi = \lambda^T g(x^*)$, we get that $\varphi^k(0) \equiv 0$ on $I$, and using the absolute continuity of $\varphi$, and its derivatives, on $I$ completes the proof.

Proposition 2 implies that if $u^*$ is piecewise continuous, then the corresponding $\varphi(t)$ is piecewise polynomial in $t$, and every polynomial has a degree $< k$.

Proposition 3: If the Lie algebra spanned by $f$ and $g$ is third-order nilpotent, and $\varphi(t) = 0$ holds only on isolated
points, then one of the following two conditions holds: (1) \( u^* \) is periodic; or (2) \( u^* \in \mathcal{P}C(t_f, 2) \) for all \( t_f < T_{\text{max}} \).

Proof: It is sufficient to prove that any bang-bang control \( u^* \) with more than two switches is periodic. Thus, suppose that \( u^* \) has switches at times \( \tau_1 < \tau_2 < \tau_3 \). We assume, without loss of generality, that \( \tau_1 > 0, \tau_3 < t_f \), and that \( u^*(t) = 0 \) for \( t \in [0, \tau_1) \).

Differentiating \( \varphi \) and applying Fact 1 yields

\[
\dot{\varphi}(t) = a(t), \quad \ddot{\varphi}(t) = b + u(t)c,
\]

where

\[
\begin{align*}
a(t) &= \lambda^T(t)[g, f](x^*(t)) \\
b &= \lambda^T(t)[[g, f], f](x^*(t)) \\
c &= \lambda^T(t)[[g, f], g](x^*(t))
\end{align*}
\]

(the fact that \( b \) and \( c \) do not depend on \( t \) is an immediate consequence of Fact 1 and the absolute continuity of \( x^*(t) \) and \( \lambda(t) \)). Note that (14) implies that \( \varphi \) is absolutely continuous.

Let \( \varphi_i(t) \) denote the restriction of the absolutely continuous function \( \varphi \) on the interval \([\tau_i, \tau_{i+1})\). Combining Proposition 2 with the fact that \( \varphi \) vanishes on the switching points yields \( \varphi_i = p_i(t - \tau_i)(t - \tau_{i+1}) \), for some \( p_i \in \mathbb{R} \), so \( \dot{\varphi}_i(t) = p_i(2t - \tau_i - \tau_{i+1}) \). This implies that \( \dot{\varphi}_i(\tau_{i+1}) = -\dot{\varphi}_i(\tau_i) \), so \( a(\tau_1) = -a(\tau_2) = a(\tau_3) \).

It now follows from (13) that \( \varphi^{(j)}(\tau_i) = \varphi^{(j)}(\tau_3) \) for \( j = 0, 1, 2 \). Since \( \varphi \) is composed of second-order polynomials this implies that \( \varphi \) is periodic. Eq. (10) implies that \( u^* \) is periodic.

The analysis based on the classical, first-order, MP provides considerable information on \( u^* \). However, it cannot be used on its own to prove Theorem 1 because of the possibility that \( u^* \) is periodic.

B. Second-order analysis

In this section, we apply a second-order MP to prove that any bang-bang control with more than three switches is not optimal. To make this paper self-contained, we provide an independent proof for our particular case. Due to space limitations, and to makeke the proof more transparent, form here on we consider the special case of linear vector fields: \( f_0(x) = Ax \) and \( f_1(x) = Bx \). The generalization to the nonlinear case is not difficult, as all the tools that we will use hold for nonlinear vector fields as well.

Assume that there exists an optimal control \( u^* \) with exactly four switches \( 0 < \tau_1 < \tau_2 < \tau_3 < \tau_4 < t_f \). Without loss of generality, we assume that \( u^*(t) = 0 \) on \( t \in [0, \tau_1) \). For notational convenience, we define \( \tau_0 := 0 \) and \( \tau_5 := t_f \) (note, however, that these are not switching points). Then, \( x^*(t_f) = \exp(Aq_1) \exp(Bq_4) \exp(Aq_3) \exp(Bq_2) \exp(Aq_1) p \), where \( q_i := \tau_i - \tau_{i-1} \).

We define the set \( S^5 \subset \mathbb{R}^5 \) by

\[
S^5 := \{ \alpha = (\alpha_0, \ldots, \alpha_4)^T : \sum_{i=0}^4 \alpha_i = 0 \}.
\]

For \( \alpha \in S^5 \) and \( s > 0 \) we define a new control \( \tilde{u}(t) = \tilde{u}(t; \alpha, s) \) by perturbing the switching times of \( u^* \) to \( \tilde{\tau}_i := \tau_i + s \sum_{k=0}^{i-1} \alpha_k, i = 1, 2, \ldots, 5 \). In other words, \( \tilde{u}(t) = 0 \) for \( t \in [0, \tilde{\tau}_1) \), \( \tilde{u}(t) = 1 \) for \( t \in [\tilde{\tau}_1, \tilde{\tau}_2) \), and so on. Note that (15) implies that \( \tilde{\tau}_5 = T_f = t_f \), that is, the final time is not changed.

It is clear that for any \( \alpha \), there exists a sufficiently small \( s_0 \) such that \( \tilde{u}(s, \alpha) \in \mathcal{U} \) for all \( s \in [0, s_0] \). In other words, \( \tilde{u}(s, \alpha) \) is an admissible control for all sufficiently small \( s \). The corresponding trajectory satisfies

\[
\dot{x}(t_f; s, \alpha) = \exp(Aq_5) \exp(Bq_4) \exp(Aq_3) \exp(Bq_2) \exp(Aq_1) p,
\]

where \( \tilde{q}_i := \tilde{\tau}_i - \tilde{\tau}_{i-1} = q_i + s \alpha_{i-1} \).

It is easy to verify that

\[
\dot{x}(t_f; s, \alpha) - x^*(t_f) = \exp(Aq_5) \exp(Bq_4) \exp(Aq_3) \exp(Bq_2) V(s, \alpha) x^*(\tau_1).
\]

This yields the following necessary condition for the optimality of \( u^* \).

\begin{proposition}
If \( u^* \) is optimal then

\[
\lim_{s \to 0^+} \frac{1}{s} \lambda^T(\tau_1) V(s, \alpha) x^*(\tau_1) = 0, \quad \forall \alpha \in S^5.
\]

\end{proposition}

Proof: Seeking a contradiction, assume that (18) does not hold for some \( \alpha^0 \in S^5 \). This implies that there exists \( s_0 > 0 \) such that \( \tilde{u}(s_0, \alpha^0) \notin \mathcal{U} \) and \( \lambda^T(\tau_1) V(s_0, \alpha^0) x^*(\tau_1) > 0 \). Using (17) yields \( (x^*(t_f))^T \dot{x}(t_f; s_0, \alpha^0) - x^*(t_f)) > 0 \). However, this clearly implies that \( |x(t_f; s_0, \alpha^0)|^2 > |x^*(t_f)|^2 \) and this contradicts the optimality of \( u^* \).

To apply Proposition 4, we will expand \( V(s, \alpha) x^*(\tau_1) \) as a Taylor series in \( s \):

\[
V(s, \alpha) x^*(\tau_1) = s V_1(s, \alpha) x^*(\tau_1) + s^2 V_2(s, \alpha) x^*(\tau_1) + o(s^2),
\]

where \( o(s) \) denotes terms that satisfy \( \lim_{s \to 0} \frac{o(s)}{s} = 0 \).

Theorem 3:

\[
V_1(s, \alpha) x^*(\tau_1) = \sum_{i=0}^4 \alpha_i H_i x^*(\tau_1),
\]

where \( H_0 = A, H_1 := B, H_2 := \exp(-Bq_2) A \exp(Bq_2), H_3 := \exp(-Bq_2) \exp(-Aq_3) B \exp(Aq_3) \exp(Bq_2), \) and \( H_4 := \exp(-Bq_2) \exp(-Aq_3) \exp(-Bq_2) A \exp(Bq_4) \exp(Aq_3) \exp(Bq_2) \).
Furthermore, for any
\[ \alpha \in S_0^5 := \{ \alpha \in S^5 : \sum_{i=0}^{4} \alpha_i H_i \phi^*(\tau_1) = 0 \} \]
we have
\[ V_2(s, \alpha) \phi^*(\tau_1) = \sum_{i=0}^{4} \sum_{j=i+1}^{4} \alpha_i \alpha_j [H_j, H_i] \phi^*(\tau_1), \] (20)
and if \( \phi^* \) is optimal then
\[ \lambda^T(\tau_1) V_2(s, \alpha) \phi^*(\tau_1) \leq 0, \quad \forall \alpha \in S_0^5. \] (21)

Proof: Eqs. (19) and (20) follow directly from the definition of \( V \). The proof of (21) is similar to the proof of Proposition 4.

Theorem 3 is the Agrachev-Gamkrelidze MP specialized for our problem. It is important to note that the expression for \( V_2(s, \alpha) \phi^*(\tau_1) \) in (20) is true only for \( \alpha \in S_0^5 \), that is, when \( V_1(s, \alpha) \phi^*(\tau_1) \) vanishes. This makes (20) meaningful even in a coordinate-free setting (see [19] for more details).

Combining Proposition 4 and (19) yields
\[ \lambda^T(\tau_1) \sum_{i=0}^{4} \alpha_i H_i \phi^*(\tau_1) \leq 0, \quad \forall \alpha \in S^5, \] (22)
and noting that if \( \alpha \in S^5 \) then \( -\alpha \in S^5 \) immediately yields the following.

Proposition 5: If the bang-bang control \( \phi^* \) is optimal then
\[ \lambda^T(\tau_1) \sum_{i=0}^{4} \alpha_i H_i \phi^*(\tau_1) = 0, \quad \forall \alpha \in S^5. \] (23)

Note that we derived Proposition 5 using the first-order approximation of \( V \) for the perturbed control \( \tilde{u} \), and that \( \tilde{u} \) was obtained by applying several needle variations to \( \phi^* \). This suggests that Proposition 5 can be derived using the classical, first-order, MP. It is easy to verify that this is indeed so. Fortunately, the second-order approximation (20) will allow us to proceed further.

As a preliminary step, we use the nilpotency assumption, combined with the celebrated Baker-Campbell-Hausdorff (BCH) formula, to simplify the \( H_1 \)'s and their Lie brackets.

Theorem 4: (BCH)
\[ \exp(At)B \exp(-At) = B + [A, B]t + [A, [A, B]] \frac{t^2}{2} + \ldots \]
Note that this result can be considered as a particular case of the pullback formula (see, e.g., [24, Appendix I]), which provides a similar expansion for the case of nonlinear vector fields.

Denote \( C := [B, A], D := [B, [B, A]], \) and \( E := [A, [B, A]] \). Using Theorem 4, and the fact that Lie brackets involving four (or more) terms of \( A \) and \( B \) vanish, we get
\[ H_2 = A - q_2 C + \frac{1}{2} q_2^2 D \]
\[ H_3 = B + q_3 C - q_2 q_3 D - \frac{1}{2} q_3^2 E \]
\[ H_4 = H_2 - q_4 C + q_4 (q_2 + \frac{1}{2} q_4) D + q_3 q_4 E. \] (24)

Proposition 6: Let \( \alpha^0 := (q_3, 2q_2, 0, -2q_2, -q_3)^T \in S^5 \). Then, \( \alpha^0 \in S_0^5 \) and
\[ \lambda^T(\tau_1) V_2(s, \alpha^0) \phi^*(\tau_1) > 0. \] (25)

Proof: It follows from the proof of Proposition 3 that \( q_4 = q_2 \). Using this and (24), we get \( \sum_{i=0}^{4} \alpha_i^0 H_i = 0 \), so \( \alpha_0 \in S_0^5 \). Calculating, we find that
\[ \sum_{i=0}^{4} \sum_{j=i+1}^{4} \alpha_i^0 \alpha_j^0 [H_j, H_i] = 2q_2 q_3 (q_3 E + 2q_2 D). \] (26)

Let \( \alpha^1 := (q_3, q_2, -q_3, -q_2, 0)^T \in S^5 \), then \( \sum_{i=0}^{4} \alpha_i^1 H_i = \frac{1}{2} q_2 q_3 (2q_2 D + q_3 E) \). Proposition 5 yields
\[ \lambda^T(\tau_1) (q_2 D + q_3 E) \phi^*(\tau_1) = 0 \]
and substituting this in (26), we get
\[ \lambda^T(\tau_1) \sum_{i=0}^{4} \sum_{j=i+1}^{4} \alpha_i^0 \alpha_j^0 [H_j, H_i] \phi^*(\tau_1) \]
\[ = 2q_2^3 q_3 \lambda^T(\tau_1) D \phi^*(\tau_1). \] (27)

Arguing as in the proof of Proposition 3, we find that \( \phi^1(t) = p_1(t - \tau_1)(t - \tau_2) \). Using the fact that \( u^*(t) = 1 \) for \( t \in [\tau_1, \tau_2] \) yields \( \phi^1(t) = 2p_1 = -\lambda^T(\tau_1) D \phi^*(\tau_1) \), and also \( \phi(t) > 0 \) for \( t \in (\tau_1, \tau_2) \). Thus, \( p_1 < 0 \) and \( \lambda^T(\tau_1) D \phi^*(\tau_1) > 0 \). Combining this with (27) completes the proof.

Combining Proposition 6 and Theorem 3, we get that any bang-bang control with exactly four switches is not optimal. It is easy to see that this implies that if \( \phi^* \) is bang-bang then \( \phi^* \in PC(t_f, 3) \) for all \( t_f < T_{max} \).

V. THE SINGULAR CASE

We now consider the possibility that \( \phi^* \) contains an arc that is not bang-bang, i.e., there exists an interval \( I \subseteq [0, t_f] \) such that \( \phi(t) = 0 \), \( \forall t \in I \). It follows from (13) that \( a(t) = b + u(t)c = 0 \) on \( I \).

We now classify the singular control into one of two types.

Type 1: \( c \neq 0 \). In this case \( u(t) = -b/c \) so \( u \) is constant on \( I \). Now suppose that \( \phi^* \) contains a combination of singular and bang-bang arcs. Say, \( u^*(t) = -b/c \) for \( t \in [\tau_1, \tau_{i+1}] \) and \( u^* \) is bang-bang on \( t \in [\tau_{i+1}, \tau_{i+2}] \). In this case, \( \phi^i_{i+1} \) (that is, the restriction of \( \phi \) over the interval \( [\tau_{i+1}, \tau_{i+2}] \)) is a second-order polynomial and, since \( \phi(\tau_{i+1}) = \phi^i_{i+1} = 0 \), we get that \( \phi^i_{i+1}(t) = p(t - \tau_{i+1})^2 \). This implies that \( \phi(t) \neq 0 \) for all \( t > \tau_{i+1} \), so \( \tau_{i+2} = t_f \). Similarly, if there is a bang arc preceding the singular arc, say, \( u^*(t) = 0 \) for \( t \in [\tau_{i-1}, \tau_i] \) then we must have \( \tau_{i-1} = 0 \). Thus, the most general configuration possible is a bang arc, followed by a singular arc, and another bang arc. Hence,
\[ \phi^* \in PC(t_f, 2). \] (28)

Type 2: \( c = 0 \). In this case, (14) yields
\[ \lambda^T(t) P(x(t)) = 0, \quad \forall t \in I \] (29)
where \( P \) is the matrix defined by \( P(x) := (g, [g, f], [g, f], [g, f], [g, f]) (x) \).
Let \( q := x^*(t_f) \). It follows from the results in [17] that there exists a time \( s \in [0, t_f] \) and an admissible control
\[
v(t) = \begin{cases} 
w(t) & t \in [0,s) \\
1 & t \in [s, t_f)
\end{cases}
\]  
(30)

with the following properties: if we denote the trajectory of (6) corresponding to \( v \) by \( z \), with \( z(0) = p \), then (1) \( z(t_f) = q \) (so \( v \) is also an optimal control); and (2) there exists a corresponding costate \( \mu \) such that \( \mu^T(t)P(z(t)) \neq 0, \forall t \in [0, t_f) \). Comparing this with (29), we see that \( w(t), t \in [0,s) \), is either a bang-bang control or a type 1 singular control. Thus, \( v \) is either (1) a bang-bang control or (2) the concatenation of a type 1 singular control and a single bang arc. In the first case, it follows from the results in Section IV that \( v \in PC(t_f, 3) \). In the second case, it follows from (28) that \( v \in PC(t_f, 3) \). This concludes the proof of Theorem 1.

VI. CONCLUSIONS

We studied an optimal control problem for nonlinear differential inclusions and switched systems. We proved that if the vector fields that generate the DI span a third-order nilpotent Lie algebra, then an optimal control with no more than three switches over any interval of time always exists (Theorem 1). This results implies that if the vector fields are also GAS, then so is the DI (Corollary 1). This is a promising step toward a solution of the open problem described in [1].

ACKNOWLEDGEMENT

We thank Andrei Agrachev for helpful comments.

REFERENCES