Optimal Control for Switched Distributed Delay Systems with Refractory Period

Erik I. Verriest
School of ECE
Georgia Institute of Technology
Atlanta, GA 30332-250
erik.verriest@ee.gatech.edu

Abstract—The optimal switching of a multi-mode distributed time delay system is considered. The class of multi-mode systems consists of systems where the control variables are the switching times in a sequence of fixed vector fields. We assume that the systems considered all have a refractory period. Necessary conditions for a stationary solution are derived and shown to extend those reported for systems with a single or commensurate delays [5]. A method based on excision and regularization to determine the optimal mode sequence is presented.

I. INTRODUCTION

The switching control problem for a finite dimensional multi-mode system involves control actions at discrete instants. The control variables are the switching times between the fixed systems \( \dot{x} = f_i(x, t) \), where \( i \in \{1, \ldots, m\} = \Xi \). We assume that at the switching times, the state is carried over from one mode to the next in a continuous fashion.

In general, also the particular sequence of vector fields needs to be optimized. Such a global control is then parameterizable by the number of switches, \( N - 1 \), a “word” of length \( N \) with alphabet, \( \Xi \), and the sequence of switching times \( \{T_1, \ldots, T_{N-1}\} \). This global problem involves comparison of \( N(N-1)^{m-1} \) individual time optimized strategies, itself manifestly leading to an obvious combinatorial explosion. Optimal sequencing can also be obtained by the excision method [5].

The objective of this work is to derive the necessary conditions for optimality for modes modelled by functional differential equations. It is assumed that the systems considered have a refractory period, in the sense that once an action is taken, it takes a non-infinitesimal amount of time before a subsequent action can be taken. Refractory periods are ubiquitous, not only in physiological systems but also in many technological systems, (e.g., time required to recharge a capacitor). A refractory time provides a safeguard towards unwanted high frequency switching. The paper extends the results of [5] to systems with distributed delays. We believe the derivation of the optimality conditions via a classical variational approach [1] to be somewhat more straightforward than the one carried out in [2] for the finite dimensional case. However, the presence of delays adds a nontrivial twist to the original problem posed in [9]. It should also be noted that the optimal switching problem bears some relation to the optimal impulsive control problem in [6]. Necessary conditions for the optimal switching policy with fixed mode sequence are determined in Section 2. In Section 3, a regularization method is presented to determine the optimal mode sequence as well.

II. VARIATIONAL APPROACH TO OPTIMAL SWITCHING

Consider a distributed delay system with a maximal delay, \( \tau \). As usual, \( x_t \) denotes the data \( \{x(t + \theta) \mid -\tau \leq \theta \leq 0\} \) [3]. First we investigate a fixed sequence of vector fields: \( f_i(x_t); i = 1, \ldots, N \), satisfying the usual Lipschitz conditions to guarantee well-posedness of the problem. The state space for this multi-mode delay system is \( C([-\tau, 0], \mathbb{R}^n) \), and the instants of switching are the sole control variables. As in the delay free problem, we assume that the entire state \( x_t \) is carried over from one mode to the next at the switch, thus preserving the continuity.

The vector \( x(t) \in \mathbb{R}^n \) is called the partial state at \( t \). Obviously, continuity of \( x \) implies continuity of the state \( x_t \). The problem is to minimize the performance index

\[
J = \int_0^T L(x, \xi) \, dt + \Phi(x(T)) \tag{1}
\]

for a fixed terminal time \( T \) by an optimal choice of the switching times. Here, \( \xi(t) \) is a discrete state, taking values in the finite set, \( \Xi \), and denotes the operating mode at time \( t \). If \( \xi(t) = a \), then the dynamical system at \( t \) is the autonomous system \( \dot{x}(t) = f(a)(x_t) \). Denote the nominal switching times by \( T_i; i = 1, \ldots, N - 1 \), and define \( T_0 = 0 \) and \( T_N = T \) (assumed fixed). For simplicity, of notation, set \( L(x, \xi(t)) = L_i(x) \) and \( f(\xi(t)) = f_i \) in the interval \((T_{i-1}, T_i)\). The performance index (1) expands to

\[
\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} L_i(x) \, dt + \Phi(x(T)) \tag{2}
\]

for \( T_{i-1} \leq t \leq T_i \). Consider now arbitrary, independent perturbations of the nominal \( T_i \) with scale parameter, \( \epsilon \), which we will let eventually tend to zero, i.e., \( T_i \rightarrow T_i + \epsilon \theta_i \). Adjoining the dynamical constraints with different Lagrange multipliers, defined in each appropriate subinterval, will not alter the value of \( J \). Assume further that optimal values \( T_i \) exist, giving the nominal performance index, \( J_0 \), equal to

\[
\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} [L_i(x) + \lambda_i^f(f_i(x_t) - \dot{x})] \, dt + \Phi(x(T)) \tag{3}
\]
It is very important to remark that due to the requisite continuity of the state, a change in $T_j$ say will have an effect on all the modes $i > j$. This happens because the change from $T_j$ to $T_j + \epsilon \theta_j$ (keeping everything else the same) now changes the final (partial) state in the $j$-th mode from $x(T_j)$ to $(\theta_j > 0$ is assumed, the other case being similar)

$$x(T_j + \epsilon \theta_j) = x(T_j) + \dot{x}(T_j^-) \epsilon \theta_j.$$  

Note that the left derivative is taken with mode $j$, and this is

$$\dot{x}(T_j^-) = f_j(x_{T_j}).$$

If however, no perturbation were made to $T_j$, then the value of the partial state $x(T_j + \epsilon \theta_j)$ would have been

$$x(T_j + \epsilon \theta_j) = x(T_j) + \dot{x}(T_j^+) \epsilon \theta_j.$$ 

This gives a difference in the state at the beginning of the $j + 1$-st mode of

$$\Delta_{T_j} x = [f_j(x_{T_j}) - f_{j+1}(x_{T_j})] \epsilon \theta_j. \quad (4)$$

As each subsequent switch will add such a term, it is clear that the effects of all such perturbations will accumulate in subsequent modes. Keeping track of all these effects will complicate the derivation requiring the explicit computation of perturbations as done in [2]. In keeping with the philosophy of calculus of variations, we shall avoid having to keep track of these by introducing a sequence of induced variations, $\{\eta_j\}$, in the same way as we introduced independent Lagrange multipliers $\lambda_j$ for each mode, i.e., in each subinterval. Equivalently, we may model the induced partial state variation $\eta(t)$ as a possibly discontinuous function with discontinuities at the switching times. The same also holds for the costates $\lambda(t)$. These costates can then be chosen in a very convenient way in order to avoid computation of the induced variations.

Defining the Hamiltonian functionals,

$$H_i(x_t, \lambda) = L_i(x) + \lambda'_j f_j(x_t), \quad (5)$$

we find for a neighboring solution (with $\eta$ possibly discontinuous)

$$\mathcal{J}_i = \Phi(x(T) + \epsilon \eta(T)) +$$

$$+ \sum_{i=1}^{N} \int_{T_{i-1} + \epsilon \theta_i}^{T_i} [H_i(x_t + \epsilon \eta_t) - \lambda'_j \dot{x}] \, dt$$

$$+ \Phi(x(T)) + \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} [H_i(x) - \lambda'_j \dot{x}] \, dt +$$

$$+ \epsilon \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} [\mathbf{D}_x H_i(x_t, \lambda; \eta_t) - \lambda'_j \dot{\eta}] \, dt +$$

$$+ \epsilon \frac{\partial \Phi}{\partial x} [\eta(T)] + \sum_{i=1}^{N} \int_{T_{i-1} + \epsilon \theta_i}^{T_i} [H_i - \lambda'_j \dot{x}] \, dt +$$

$$- \sum_{i=1}^{N} \int_{T_{i-1} + \epsilon \theta_i}^{T_i} \{H_i - \lambda'_j \dot{x}\} \, dt. \quad (6)$$

In the above, $\mathbf{D}_x$ is the Fréchet derivative defined by

$$\mathbf{D}_x H_i(x_t, \lambda; \eta_t) = \lim_{\epsilon \to 0} \frac{H_i(x_t + \epsilon \eta_t, \lambda) - H_i(x_t, \lambda)}{\epsilon}. \quad (7)$$

Note that actually $\theta_0 = \theta_N = 0$, since initial and final time were considered fixed.

A. Progressively Distributed Delay Mode Systems

We shall leave the full generality of the problem behind and consider from now on only systems having progressive distributed delay dynamics, by which we mean system modes representable by the functional differential equations

$$\dot{x}(t) = \int_{0}^{\tau} G_i(x(t - \sigma), \sigma) \, dB_i(\sigma). \quad (8)$$

Here the $G_i$ are smooth differentiable maps $\mathbb{R}^n \times [0, \tau] \to \mathbb{R}^n$ and $B_i(\cdot)$ has bounded variation. The integration boundaries $0$ and $\tau$ indicate that possible crisp (or point-) delays for $\sigma = 0$ (the no-delay term) and $\sigma = \tau$ are included. However, the form (8) excludes correlations such as the product $x(t) x(t - \tau)$. This class includes the separable mode systems (with crisp point delays) [5]. At this point we remark that the above restriction is only made for notational purposes in what is to follow. Full generality can be achieved since as we shall see it suffices to apply the perturbation to the linear variational problem. Thus, the Hamiltonians, $H_i(x_t, \lambda_i)$ and final cost adjoined with the state continuity constraints are respectively defined by

$$L_i(x(t)) + \lambda'_i \left[ \int_{0}^{\tau} G_i(x(t - \sigma), \sigma) \, dB_i(\sigma) \right] \quad (9)$$

$$\Psi(x, \mu) = \Phi(x) + \sum_{i=1}^{N} \mu'_i [x(T_i^+) - x(T_i^-)]. \quad (10)$$

B. Delay effect of a single switch

We shall say that a function $y$ is $C_k$ at $t_0$, if the $k$-th derivative of $y$ is continuous at $t_0$, but the $(k + 1)$-st is not. Obviously, this implies that the derivatives of order $i$ are all continuous at $t_0$ for $i \leq k$.

Assume that a single controlled switch occurs at time $T$, switching from mode $i$ to $i + 1$. This makes $\dot{x}$ discontinuous at $T$. Consequently, $x$ has a ‘kink’ (i.e., is non-differentiable) at $T$. From the continuity assumption, $x$ is $C_0$ at $T$. But then $\int_{0}^{\tau} G(x(t - \sigma), \sigma) \, dB(\sigma)$ is $C_0$ at $t = T + \tau$ if $B(\theta) = b(\theta - \tau)$ (see [5]), and $C_1$ if $B$ is smooth. In turn this implies that $\dot{x}$ is at least $C_0$ at $T + \tau$, inducing again at least $C_1$ behavior in $x$ at $T + \tau$ and at least $C_2$ behavior in $\int G \, dB$ at time $T + \tau$, and so on. We summarize the chain of events (where $C_k \subseteq C_k$ means “at least $C_k$”):
Lemma 1: A controlled switch at time $T$ induces the following behavior in the delay system (8)

<table>
<thead>
<tr>
<th>$x(t)$</th>
<th>$T$</th>
<th>$T + \tau$</th>
<th>$T + 2\tau$</th>
<th>...</th>
<th>$T + k\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>jump</td>
<td>$C_0$</td>
<td>$\geq C_1$</td>
<td>$\geq C_2$</td>
<td>...</td>
<td>$\geq C_{k-1}$</td>
</tr>
<tr>
<td>$x(t)$</td>
<td>$C_0$</td>
<td>$\geq C_1$</td>
<td>$\geq C_2$</td>
<td>...</td>
<td>$\geq C_k$</td>
</tr>
<tr>
<td>$\int_0^T G(x_t) dB$</td>
<td>$\geq C_1 \quad \geq C_2 \quad \geq C_1 \quad \geq C_k$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. Effect of variation in a single switching time

Let us first recall a simple result:

Lemma 2: If $y$ is $C_k$ at $t_0$, then the variation of $y$ in an interval of length $\theta$ about $t_0$ is of order $k + 1$ in $\theta$.

Consider now for smooth functions $G$ a functional of the form

$$I(T) = \int_0^T y(t) \, dt + \sum_{i=1}^k \int_{T+(i-1)\tau}^{T+i\tau} y(t) \, dt.$$  

where $x$ satisfies the dynamics (8) with a single switch at time $T$. Observe (by Lemma 1) that this induces $C_0$ behavior at $T$, $C_1$ behavior at $T + \tau$, ..., $C_m$ behavior at $T + m\tau$ in $x(t)$.

Partition the integral as

$$I(T) = \int_0^T y(t) \, dt + \sum_{i=1}^k \int_{T+(i-1)\tau}^{T+i\tau} y(t) \, dt.$$  

where

$$y(t) = \int_0^T q(x(t-\sigma), \sigma) \, dB(\sigma).$$

If the switch occurred $\epsilon$ time units later than $T$, then $I(T)$ becomes $I(T + \epsilon)$, given by

$$\int_0^{T+\epsilon} y(t) \, dt + \sum_{i=1}^k \int_{T+(i-1)\tau+\epsilon}^{T+i\tau+\epsilon} [y(t) + \epsilon \eta_i(t)] \, dt.$$  

The induced variation in the integral is $I(T + \epsilon) - I(T)$ and (assuming first $\epsilon > 0$) after some regrouping of terms, reduces to

$$\int_{T}^{T+\epsilon} [y_-(t) - y_+(t)] \, dt + \sum_{i=1}^{k-1} \int_{T+i\tau}^{T+i\tau+\epsilon} [y_-(t) + \epsilon \eta_i - y_+(t) - \epsilon \eta_{i+1}] \, dt + \int_{T+k\tau}^{T+k\tau+\epsilon} y_-(t) + \epsilon \eta_k(t) \, dt + \sum_{i=1}^k \int_{T+(i-1)\tau}^{T+i\tau} \epsilon \mathbf{D}_{\eta_i} [y(t)] \, dt.$$  

(12)

The subscript $\pm$ reminds us that the integrand needs to be computed on the left (-) and right(+) hand side of the point. The analysis for $\epsilon < 0$ is analogous.

The first integral, $\int_T^{T+\epsilon} [y_-(t) - y_+(t)] \, dt$, is of order $2$ in $\epsilon$. Indeed, $y_-(T) = \int_T^{T+\epsilon} q(x(T-\sigma), \sigma) dB(\sigma)$, and involves the dynamics present before the switch. Likewise $y_+(T)$ must be computed with the dynamics reigning after the switch. The integral near $T + \tau$ contributes a perturbation proportional to at least second degree in $\epsilon$. Likewise, the induced perturbations at $T + i\tau$ are of order at least $i + 1$ in $\epsilon$. All these terms will be neglected.

This leaves the $\mathbf{D}_{\eta_i} [y(t)]$ integrals, where $\mathbf{D}_{\eta_i}$ denotes the Fréchet derivative. Hence

$$\mathbf{D}I(T) = \sum_{i=1}^k \int_{T+(i-1)\tau}^{T+i\tau+\epsilon} \mathbf{D}_{\eta_i} [y(t)] \, dt.$$  

This was the situation if there is only a single switch. If another switch occurs, say at $T_1$, then for some $k, T + k\tau < T_1 < T + (k + 1)\tau$, and the $G$-term would induce another $\epsilon$ perturbation. Therefore the bookkeeping of all perturbation terms will be quite complicated, especially in view of the fact that all possibilities (of relative positions of switching instants) need to be taken into account.

D. Refractory period

Let us now assume the existence of a refractory period. In particular, we consider the case where this refractory time exceeds the delay time, $\tau$. In this case, the problem greatly simplifies, and the aforementioned complexity disappears, as only two adjacent intervals need to be considered [5]. Indeed, the delayed effect of the $k$-th switch hits before the $k + 1$-st switch. This yields a situation which is akin to the case with a single switch in the previous section.

As in (11), with the Hamiltonian functionals defined in (9), we express the first variation in the performance index as the limit for $\epsilon \to 0$ of

$$\delta J = \lim_{\epsilon \to 0} \frac{J_\epsilon - J}{\epsilon}.$$  

We start with (3) for the progressively distributed mode systems mode form

$$\sum_{i=1}^N \int_{T_{i-1}}^{T_i} \left[ L_i(x) + \lambda_i' \left( \int_0^\tau G_i(x(t-\sigma), \sigma) \, dB_i(\sigma) - \dot{x} \right) \right] \, dt$$  

+ $\Phi(x(T)) + \sum_{i=1}^N \mu_i [x(T_i^+) - x(T_i^-)],$$  

(14)

in which we separate the delayed terms

$$\mathcal{J} = \sum_{i=1}^N \int_{T_{i-1}}^{T_i} [ L_i(x) - \lambda_i' \dot{x} ] \, dt$$  

+ $\sum_{i=1}^N \int_{T_{i-1}}^{T_i} \lambda_i' \left( \int_0^\tau G_i(x(\tau-\sigma), \sigma) \, dB_i(\sigma) \right) \, dt$  

+ $\Phi(x(T)) + \sum_{i=1}^N \mu_i [x(T_i^+) - x(T_i^-)]$  

(15)
The integrals of the delayed terms in (15) may be rearranged by permuting the order of integration as follows:

\[
\int_{T_{i-1}}^{T_i} \chi_i'(t) \left( \int_{0}^{\tau+} G_i(x(t - \sigma), \sigma) \, dB_i(\sigma) \right) \, dt \\
= \int_{T_{i-1}}^{T_i} \chi_{[T_{i-1} - \tau, T_{i-1}]} \left( \int_{T_{i-1} - \tau}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x(t)) \, dt + \\
+ \int_{T_{i-1}}^{T_i} \chi_{(T_{i-1}, T_{i-1} - \tau)} \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x(t)) \, dt + \\
+ \int_{T_{i-1}}^{T_i} \chi_{(T_{i-1} - \tau, T_{i-1})} \left( \int_{-T_{i-1} - \tau}^{T_{i-1} - \tau} (\lambda_i^T)' \, dB_i \right) G_i(x(t)) \, dt.
\]

In these expressions, \( \chi_i \) is the indicator function of the interval \( I_i \), and the advanced term \( \lambda(t + \sigma) \) is denoted by \( \lambda^T \).

Hence, by rearranging terms the expression (15) reduces to (for simplicity, we also denote \( \chi_{i+1}^{+} = \chi_{(T_i, T_{i+1} - \tau)} \) and \( \chi_{i+1}^{-} = \chi_{[T_{i-1}, T_{i-1} - \tau)} \))

\[
\mathcal{J}_i = \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[ L_i(x) - \Lambda_i^T \right] \, dt + \\
+ \sum_{i=1}^{N-1} \int_{T_{i-1}}^{T_i} \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \, dt + \\
+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \chi_{i+1}^{+} \left( \int_{0}^{\tau} (\lambda_{i+1}^T)' \, dB_{i+1} \right) G_{i+1}(x) \, dt + \\
+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \chi_{i+1}^{-} \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \, dt + \\
+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \chi_{i+1}^{-} \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \, dt + \\
+ \Phi(x(T)) + \sum_{i=1}^{N} \mu_i^{[x(T_i^{+}) - x(T_i^{-})]}).
\]

By adding \( \lambda_{N+1}(t) = 0 \) for \( t > T_N \), the sum from \( i = 0 \) to \( N - 1 \), may be changed to the sum from \( i = 0 \) to \( N \), thus,

\[
\mathcal{J}_i = \sum_{i=1}^{N} \int_{0}^{T_i} \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \, dt + \\
+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[ L_i(x) + M_i(x) - \Lambda_i^T \right] \, dt + \\
+ \Phi(x(T)) + \sum_{i=1}^{N} \mu_i^{[x(T_i^{+}) - x(T_i^{-})]},
\]

where

\[
M_i(x) = \left[ \chi_i^{+} T_i^{(1)} + \chi_{i+1}^{-} T_i^{(2)} + T_i^{(3)} \right],
\]

and

\[
\mathcal{I}_i^{(1)}(x) = \left( \int_{0}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \\
\mathcal{I}_i^{(2)}(x) = \left( \int_{0}^{T_i-\tau} (\lambda_i^T)' \, dB_i \right) G_i(x) \\
\mathcal{I}_i^{(3)}(x) = \left( \int_{T_{i-1} - \tau}^{\tau} (\lambda_i^T)' \, dB_i \right) G_i(x).
\]

The integrals in (16) involve expressions of the form

\[
\mathcal{K}_i = \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[ K_i(x) - \Lambda_i T_i \right] \, dt,
\]

for which we now consider independent perturbations of the switching times. To that effect introduce the variables \( \{ \tilde{\lambda}_i \} \) and a scale parameter \( \epsilon \) and the induced perturbations \( \eta_i \):

\[
\mathcal{K}_i = \sum_{i=1}^{N} \int_{T_{i-1} + \epsilon \theta_i}^{T_i} \left[ K_i(x + \epsilon \eta_i) - \Lambda_i T_i \right] \, dt + \\
+ \sum_{i=1}^{N} \left( \int_{T_{i-1} + \epsilon \theta_i}^{T_i} \left[ K_i(x) - \Lambda_i T_i \right] \, dt + \\
- \int_{T_{i-1}}^{T_{i-1} + \epsilon \theta_i} \left[ K_i(x) - \Lambda_i T_i \right] \, dt \right) + \\
+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left( \frac{\partial K_i}{\partial x} + \lambda_i^T \right) \, \epsilon \eta_i \, dt + \\
+ \epsilon \sum_{i=1}^{N} \left[ -\Lambda_i T_i \eta_i(T_i) + \lambda_i T_i \eta_i(T_i) \right],
\]

where as usual, we integrated by parts. Subtracting the nominal value, dividing by \( \epsilon \) and taking the limit gives the induced variation \( \delta \mathcal{K} \):

\[
\delta \mathcal{K} = \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left( \frac{\partial K_i}{\partial x} + \lambda_i^T \right) \eta_i(t) \, dt + \\
+ \sum_{i=1}^{N} \theta_i [K_i - \Lambda_i T_i] \mathcal{I}_i^{(2)} - \sum_{i=0}^{N-1} \theta_i [K_{i+1} - \Lambda_{i+1} T_i] \mathcal{I}_i^{(3)} + \\
- \sum_{i=1}^{N} \lambda_i(T_i) \eta_i(T_i) + \sum_{i=0}^{N-1} \lambda_{i+1}(T_i) \eta_{i+1}(T_i).
\]

Note that for \( i = 0, \theta_0 = 0 \) and for \( i = N, \theta_N = 0 \), since initial and final time were fixed.

The \( \{ \tilde{\theta}_i \} \)-induced perturbation of the non-integral term in (16), \( \Psi(x, \mu) = \Phi(x) + \sum \mu_i^{[\tilde{\lambda}(T_i) - \tilde{\lambda}(T_i^+) + \lambda_i T_i]} \), follows from (4),

\[
\delta \Psi = \frac{\partial \Phi}{\partial x} \eta_N(T) + \sum_{i=1}^{N-1} \mu_i^{[\tilde{\lambda}(T_i^+) - \tilde{\lambda}(T_i^-)]} \theta_i + \\
+ \eta_i(T_i^-) - \eta_{i+1}(T_i^+). \]
Combine the two perturbations $\delta K$ and $\delta \Psi$ and choose $\lambda_i$ in the intervals $[T_{i-1}, T_i]$ to solve

$$\dot{\lambda}_i = - \left( \frac{\partial K_i}{\partial x} \right)'.$$

(24)

This yields $\delta K + \delta \Psi$ in the form

$$\sum_{i=1}^{N-1} \left[ A_i \theta_i + B_i \eta(T_{i}^+) + C_i \eta(T_{i}^-) \right] + \lambda_1(0^+)\eta(0^+) + \left( \frac{\partial \Phi}{\partial x} - \lambda_N(T_N^-) \right) \eta(T_N^-),$$

where

$$A_i = L_i(x(T_{i}^-)) - L_{i+1}(x(T_{i}^+)) + \mu_i' [\dot{x}(T_{i}^-) - \dot{x}(T_{i}^-)]$$

$$B_i = -\mu_i + \lambda_{i+1}(T_{i}^+)$$

$$C_i = \mu_i - \lambda_i(T_{i}^-)$$

(25)

(26)

(27)

(28)

With the initial data given $\eta(0^+)$ must be zero. Computation of the requisite perturbations $\{\lambda_i\}$ is avoided if one chooses

$$\lambda_i(T_{i}^-) = \mu_i = \lambda_{i+1}(T_{i}^+),$$

(29)

thus specifying the boundary conditions for the differential equations (24). This implies that we can choose the costates $\lambda_i$ in $[T_{i-1}, T_i]$ to concatenate to a continuous functions in $[0, T]$.

It follows that the first order variation of $J$ reduces to

$$\delta J = \sum_{i=1}^{N-1} A_i \theta_i.$$

(31)

E. Main Result

The above derivations are now put together to obtain necessary conditions for the general progressively distributed delay system. Since the $\theta_i$ are independent, necessary conditions for optimality are the vanishing of the $A_i$ in (31). In view of the choice (29) of the multipliers $\mu_i$ and boundary conditions, it gives for $i = 1$ to $N - 1$

$$H_i(x_{T_i}, \lambda(T_i)) = H_{i+1}(x_{T_i}, \lambda(T_i)).$$

(32)

Simply stated, it means the continuity of the Hamiltonian functional $H$ at the switching times.

In formulating our main theorem below, we will assume that the vector fields $G_i(x(t - \sigma), \sigma)$ as well as the functions $L_i(x)$ are smooth, and we let $N - 1$ be the total number of switches, with $T_0 = 0$ and $T_N = t_f$ being fixed.

**Theorem 3:**

The separable mode switched system in equation (8), with fixed mode sequence, minimizes the performance index $J$ in (1) if the switching times $T_i$ are chosen as follows:

Euler-Lagrange Equations:

$$\dot{\lambda}_i = - \left( \frac{\partial L_i}{\partial x} \right)^T - \left( \frac{\partial G_i}{\partial x} \right)^T (\lambda_{i+1}^T T_i^+ + \lambda_{i+1}^T T_i^+) + \left( \frac{\partial \Phi}{\partial x} \right)^T \lambda_{i+1}^T T_i^+$$

(33)

where the $I_1$ and $I_2$, defined in (20) depend on $\lambda_{i}^T$ and $I_3$ on $\lambda_{i+1}^T$. Also, $T_{i-1} < t < T_i$, $i = 1, \ldots, N - 1$, with $\lambda_i^+(t) = 1$ if $t \in [T_{i-1}, T_i - \tau]$ and 0 otherwise, $\lambda_{i+1}^+(t) = 1$ if $t \in [T_i - \tau, T_i]$ and 0 otherwise, and $\lambda_i^T = \lambda_i(t + \tau)$. Moreover,

$$\dot{\lambda}_N = - \left( \frac{\partial L_N}{\partial x} \right)^T - \left( \frac{\partial G_N}{\partial x} \right)^T (\lambda_{i+1}^T T_i^+ + \lambda_{i+1}^T T_i^+)$$

(34)

Boundary Conditions:

$$\lambda_N(T_N) = \left( \frac{\partial \Phi}{\partial x} \right)^T$$

(35)

$$\lambda_i(T_i^-) = \lambda_{i+1}(T_i^+)$$

(36)

Optimality Conditions:

$$H_i(x_{T_i}, \lambda_i(T_i)) = H_{i+1}(x_{T_i}, \lambda_{i+1}(T_i)).$$

(37)

**Proof:** All one has to do is to recall what the function $K$ was and realize that both indicator functions in its definition evaluate to 1 at the switching point $T_i$. 

F. Point delay systems

The separable mode case discussed in [5] is retrieved by taking the $B_i(\theta)$ to be step functions. For systems with a single delay, let

$$\int_{0}^{\tau^+} G_i(x(t - \sigma), \sigma) dB_i(\sigma) = f_i(x(t) + g_i(x(t - \tau)).$$

(38)

Systems with multiple delays are treated by multi-level step functions $B_i$, whereas a system with commensurate delays can be recast as a system with a simple delay using state augmentation (see [5]).

III. Optimal Sequencing

So far, it was assumed that the sequence of modes was fixed. If $m$ modes are available, then with $N - 1$ switches, $N(N - 1)^{m-1}$ possible mode sequences exist, and their optimized performance needs to be evaluated and compared in order to find the global optimal switched control. Instead we refine an alternative method, first presented in [5].
A. Excision

In the absence of a refractory period, consider the fixed mode sequence by cycling the modes: 1 → 2 → ... → m → 1 → 2 → ... → m → 1 → 2 → ..., and optimize their switching times. If it is found that the optimal switching time sequence has $T_{i-1} = T_i$ for some $i$, it means that mode $i$ only occurs for duration 0, and therefore should be excised. Performing all such excisions will leave the optimal mode sequence associated with at most $N-1$ switches. Thus by formulating first the optimal control problem for a fixed mode sequence, no generality is lost in the global optimal control problem if an upper bound on the number of switches is imposed.

B. Regularization of the refractory period

With a refractory period, the time between switches is bounded below and the excision method is not applicable. There is however a way around this. First the problem in Section 2 can be generalized by adding a switching cost term to the performance index. Thus let (1) be replaced by

$$J = \int_0^T L(x, \xi) \, dt + \sum_{i=1}^N \Phi_i(x(T_i), \{T_j\}_{j=i+1}^N)$$

which leads to a more general set of boundary and optimality conditions in Theorem 3 (see [7] for the case of impulsive control.) Consider a switching cost function of the form

$$\Phi_i((x(T_i), \{T_j\}_{j=i+1}^N) = \phi(T_i - T_{i-1})$$

The existence of a true refractory period, $\tau$, is then equivalent to $\phi = \phi_r$ with $\phi_r(\theta) = 0$ for $\theta > \tau$, and $\phi_r(\theta) = \infty$ for $\theta < \tau$. It follows that one can regularize the refractory period phenomenon by considering instead of the above $\phi$ a more general smooth cost function: Examples are (with $\{\phi_i\}$ control.) Consider a switching cost function of the form

$$J(\theta) = \int_0^T L(x, \xi) \, dt + \sum_{i=1}^N \Phi_i(x(T_i), \{T_j\}_{j=i+1}^N)$$

of a multi-mode system, where also the optimal sequence of the modes needs to be found. To avoid a combinatorial search, we proposed an adaptation of the excision method, first presented in [5] for systems without refractory period constraints, via regularization of an equivalent cost function. Alternatively, a regularization method as for instance presented in [4] could be invoked to obtain a first approximation.

IV. CONCLUSIONS

We derived necessary conditions for stationarity of the performance index of a multi-mode distributed delay system controlled by switchings between a prespecified mode sequence. This is a first step in the complete optimal control of a multi-mode system, where also the optimal sequence of the modes needs to be found. To avoid a combinatorial search, we proposed an adaptation of the excision method, first presented in [5] for systems without refractory period constraints, via regularization of an equivalent cost function. Alternatively, a regularization method as for instance presented in [4] could be invoked to obtain a first approximation.

V. REFERENCES