model we consider four parameters. In the latter, we selected two different objective functions leading to uni-modal and bi-modal stationary distributions.

The techniques presented in this technical note could also aid the design of novel gene regulatory circuits with desirable properties, or it could be used in determining how to best combine circuits—each matrix H_i representing a different one.

ACKNOWLEDGMENT

REFERENCES

[1] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*. Amsterdam, The Netherlands: North Holland, 2007.

[2] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Nashua, NH: Athena Scientific, 1997.

[3] M. Gallivan and R. Murray, "Model reduction and system identification for master equation control systems," in *Proc. Amer. Control Conf.*, 2003, pp. 3561–3566.

[4] D. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *J. Comp. Phys.*, vol. 22, no. 403, pp. 403–434, 1976.

[5] D. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Comp. Phys.*, vol. 81, no. 25, pp. 2340–2361, 1977.

[6] D. Gillespie, *Markov Processes: An Introduction for Physical Scientists*. San Diego, CA: Academic Press, 1992.

[7] J. Hespanha and A. Singh, "Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems," *Int. J. Robust Control, Special Iss. Control Small Scales*, vol. 1, no. 15, pp. 669–689, 2005.

[8] D. G. Luenberger, *Optimization by Vector Space Methods (Series in Decision and Control)*. New York: Wiley-Interscience, January 25, 1997.

[9] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *J. Chem. Phys.*, vol. 124, no. 4, 2006 [Online]. Available: http://ccdc.mee.ucsb.edu/ pdf/ ccdc-05-0505.pdf

[10] A. Singh and J. Hespanha, "LogNormal moment closures for biochemical reactions," in *Proc. Conf. Decision Control*, 2006, pp. 2063–2068.

Computational Method for a Class of Switched System Optimal Control Problems

Ryan Christopher Loxton, Kok Lay Teo, and Volker Rehbock

Abstract—We consider an optimal control problem with dynamics that switch between several subsystems of nonlinear differential equations. Each subsystem is assumed to satisfy a linear growth condition. Furthermore, each subsystem switch is accompanied by an instantaneous change in the state. These instantaneous changes—called "state jumps"—are influenced by a set of control parameters that, together with the subsystem switching times, are decision variables to be selected optimally. We show that an approximate solution for this optimal control problem can be computed by solving a sequence of conventional dynamic optimization problems. Existing optimization techniques can be used to solve each problem in this sequence. A convergence result is also given to justify this approach.

Index Terms—Nonlinear differential equations.

I. INTRODUCTION

A novel time-scaling transformation for switched system optimal control problems is discussed in [1]–[3]. This transformation converts the original problem—in which the subsystem switching times are decision variables—into a new optimal control problem that is easier to solve. The new problem is also governed by a switched system, but its switching times are fixed points, not decision variables. In fact, each subsystem in the new switched system is active for the same duration of time. This facilitates accurate numerical integration and ensures the new problem can be solved using a nonlinear programming algorithm.

The time-scaling transformation introduced in [1]–[3] is only applicable when the governing switched system does not contain state jumps. A similar transformation has been developed for systems in which the state, as well as the dynamics, change instantaneously at the switching times [4], [5]. Practical examples of such systems include switched-capacitor dc/dc power converters [6], [7] and bioconversion reactors [8], [9]. Note that the stability of switched systems has been investigated in [10]–[12].

To apply the transformation discussed in [4], [5], the optimal number of switches must be known *a priori*. This is because the time-scaling transformation assumes that the number of switches in the system is fixed and *every* available switch is applied. Even if two or more switching times coincide at a single point (and thus combine to form a single switch), they still correspond to *different* switching times in the new switched system obtained via the transformation. Consequently, whenever multiple switching times coincide, the transformation introduces additional state jumps that do not occur in the original switched system.

In practice, excessive switching between subsystems can adversely affect the overall system. For example, in a switched-capacitor dc/dc power converter, each subsystem represents a different circuit topology, and although changing the topology helps to regulate the output voltage, it also induces a voltage leak in the capacitors. Hence, it is usually not optimal to apply every available switch; the optimal solution may involve "deleting" switches by merging two or more

switching times to form a single switch. In this case, the optimal switching times coincide, and the transformation developed in [4], [5] cannot preserve the true effect of switching on the system. The aim of this technical note is to develop a new technique that is capable of handling this case.

II. PROBLEM FORMULATION

Let $T > 0$ be a given terminal time and define

$$\Gamma \triangleq \left\{ \boldsymbol{\tau} \in \mathbb{R}^n : \tau_{i-1} \leq \tau_i, \ i = 1, \ldots, n+1 \right\}$$

where $\tau_0 \triangleq 0$ and $\tau_{n+1} \triangleq T$. Additionally, define

$$\Xi \triangleq \left\{ \boldsymbol{\sigma} \in \mathbb{R}^r : \sigma_i^l \leq \sigma_i \leq \sigma_i^u, \ i = 1, \ldots, r \right\}$$

where, for each $i = 1, \ldots, r$, σ_i^l and σ_i^u are given constants such that $\sigma_i^l < \sigma_i^u$.

Now, let $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma \times \Xi$ and consider the following switched system:

$$\dot{\mathbf{x}}(t) = \mathbf{f}^i(\mathbf{x}(t), \boldsymbol{\sigma}), \ t \in (\tau_{i-1}, \tau_i), \ i = 1, \ldots, n+1 \qquad (1)$$

and

$$\mathbf{x}(\tau_i^+) = \begin{cases} \mathbf{x}^0, & \text{if } i = 0, & \text{(2a)} \\ \mathbf{x}(\tau_i^-) + \boldsymbol{\phi}^i(\mathbf{x}(\tau_i^-), \boldsymbol{\sigma}), & \text{if } i \in \{1, \ldots, n\}, \ \tau_{i-1} < \tau_i < T. \\ & \text{(2b)} \end{cases}$$

Here, $\mathbf{x}(t) \in \mathbb{R}^m$ is the state at time t; $\mathbf{x}^0 \in \mathbb{R}^m$ is a given initial state; and $\mathbf{f}^i : \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}^m$ and $\boldsymbol{\phi}^i : \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}^m$ are continuously differentiable functions. Furthermore, $\boldsymbol{\sigma}$ is a vector of control parameters and τ_i, $i = 1, \ldots, n$, are switching times.

The state starts from $\mathbf{x}^0$ at time $t = 0$ and evolves smoothly according to (1) until the first *positive* switching time is reached. It then jumps to a new point, which is given by (2b). Starting from this new point, the state again evolves smoothly until the next switching time is reached, and so on for the remainder of the time horizon. The solution of (1)–(2) obtained in this way is denoted by $\mathbf{x}(\cdot | \boldsymbol{\tau}, \boldsymbol{\sigma})$. We define an optimal control problem as follows.

Problem 1: Choose $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma \times \Xi$ to minimize the objective function

$$J(\boldsymbol{\tau}, \boldsymbol{\sigma}) \triangleq \sum_{i=1}^{n+1} \int_{\tau_{i-1}}^{\tau_i} \mathcal{L}_i(\mathbf{x}(t | \boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}) dt$$

where $\mathcal{L}_i : \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}$ are continuously differentiable functions.

We assume that the subsystems in (1) satisfy the following linear growth condition: there exists a positive constant K such that for each $i = 1, \ldots, n+1$:

$$\|\mathbf{f}^i(\mathbf{x}, \boldsymbol{\sigma})\| \leq K(1 + \|\mathbf{x}\|), \quad (\mathbf{x}, \boldsymbol{\sigma}) \in \mathbb{R}^m \times \Xi. \qquad (3)$$

This assumption[1] ensures that the system does not blow up in finite time, so that Problem 1 is well-defined.

III. EQUIVALENT PROBLEM

Since the state in Problem 1 depends on the switching times, it is difficult to integrate (1)–(2) numerically. In this section, we will derive a more tractable equivalent problem. To begin, let

$$\Theta \triangleq \left\{ \boldsymbol{\theta} \in \mathbb{R}^{n+1} : \theta_i \geq 0; \ \theta_1 + \cdots + \theta_{n+1} = T \right\}.$$

[1]In this technical note, $\|\cdot\|$ denotes the Euclidean norm.

For each $\boldsymbol{\theta} \in \Theta$, define a function $\mu(\cdot | \boldsymbol{\theta}) : [0, n+1] \to \mathbb{R}$ by

$$\mu(s | \boldsymbol{\theta}) \triangleq \begin{cases} \sum_{j=1}^{\lfloor s \rfloor} \theta_j + \theta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor), & \text{if } s \in [0, n+1), \\ T, & \text{if } s = n+1 \end{cases}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Clearly

$$\mu(i | \boldsymbol{\theta}) = \sum_{j=1}^i \theta_j, \quad i = 0, \ldots, n+1. \qquad (4)$$

It is tedious, but not difficult, to verify the following additional properties for $\mu(\cdot | \boldsymbol{\theta})$:
   a) $\mu(0 | \boldsymbol{\theta}) = 0$, $\mu(n+1 | \boldsymbol{\theta}) = T$, and $\mu(s | \boldsymbol{\theta}) \in [0, T]$ for all $s \in [0, n+1]$;
   b) For each $i = 1, \ldots, n+1$, $\mu(\cdot | \boldsymbol{\theta})$ is strictly increasing on $[i-1, i]$ if and only if $\theta_i > 0$;
   c) For each $i = 1, \ldots, n+1$, $\mu(\cdot | \boldsymbol{\theta})$ is constant on $[i-1, i]$ if and only if $\theta_i = 0$;
   d) $\mu(\cdot | \boldsymbol{\theta})$ is a continuous function.

Since $\mu(\cdot | \boldsymbol{\theta})$ is non-decreasing on each closed subinterval $[i-1, i]$, it is non-decreasing on the entire interval $[0, n+1]$. Hence, for every $\boldsymbol{\theta} \in \Theta$

$$\mu(i-1 | \boldsymbol{\theta}) \leq \mu(i | \boldsymbol{\theta}), \quad i = 1, \ldots, n+1$$

which implies that

$$\boldsymbol{\nu}(\boldsymbol{\theta}) \triangleq [\mu(1 | \boldsymbol{\theta}), \ldots, \mu(n | \boldsymbol{\theta})]^T \in \Gamma. \qquad (5)$$

In other words, the components of $\boldsymbol{\nu}(\boldsymbol{\theta})$ are valid switching times for Problem 1. In fact[2]

$$\Gamma = \{\boldsymbol{\nu}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}. \qquad (6)$$

This shows that valid switching times for Problem 1 are generated by a corresponding vector in Θ. By property (a) and (6), we can define a new state variable as follows:

$$\tilde{\mathbf{x}}(s | \boldsymbol{\theta}, \boldsymbol{\sigma}) \triangleq \mathbf{x}(\mu(s | \boldsymbol{\theta}) | \boldsymbol{\nu}(\boldsymbol{\theta}), \boldsymbol{\sigma}), \quad s \in [0, n+1]. \qquad (7)$$

To derive a new optimal control problem, we must determine the dynamic behavior of $\tilde{\mathbf{x}}(\cdot | \boldsymbol{\theta}, \boldsymbol{\sigma})$ on the new time horizon $[0, n+1]$. Clearly

$$\dot{\mu}(s | \boldsymbol{\theta}) = \theta_i, \quad s \in (i-1, i), \quad i = 1, \ldots, n+1. \qquad (8)$$

Properties (a)–(d) and (1)–(2), (7)–(8) can be used to show that

$$\dot{\tilde{\mathbf{x}}}(s) = \theta_i \mathbf{f}^i(\tilde{\mathbf{x}}(s), \boldsymbol{\sigma}), \ s \in (i-1, i), \ i = 1, \ldots, n+1 \qquad (9)$$

and, for each $i = 0, \ldots, n$

$$\tilde{\mathbf{x}}(i^+) = \begin{cases} \mathbf{x}^0, & \text{if } i = 0, & \text{(10a)} \\ \tilde{\mathbf{x}}(i^-) + \boldsymbol{\phi}^i(\tilde{\mathbf{x}}(i^-), \boldsymbol{\sigma}), & \text{if } \mu(i-1 | \boldsymbol{\theta}) < \mu(i | \boldsymbol{\theta}) < T & \text{(10b)} \\ \tilde{\mathbf{x}}(i^-), & \text{otherwise} & \text{(10c)} \end{cases}$$

where for simplicity we have written $\tilde{\boldsymbol{x}}(\cdot | \boldsymbol{\theta}, \boldsymbol{\sigma})$ as $\tilde{\boldsymbol{x}}(\cdot)$. We define a new optimal control problem as follows.

Problem 2: Choose $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$ to minimize the objective function

$$\tilde{J}(\boldsymbol{\theta}, \boldsymbol{\sigma}) \triangleq J(\boldsymbol{\nu}(\boldsymbol{\theta}), \boldsymbol{\sigma}) = \sum_{i=1}^{n+1} \int_{i-1}^i \theta_i \mathcal{L}_i(\tilde{\mathbf{x}}(s | \boldsymbol{\theta}, \boldsymbol{\sigma}), \boldsymbol{\sigma}) ds.$$

[2]It follows immediately from (5) that $\{\boldsymbol{\nu}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\} \subseteq \Gamma$. To prove the opposite inclusion, let $\boldsymbol{\tau} \in \Gamma$ and put $\theta_1 = \tau_1$ and $\theta_i = \tau_i - \tau_{i-1}$ for $i = 2, \ldots, n+1$. Then $\boldsymbol{\theta} \in \Theta$ and $\mu(i | \boldsymbol{\theta}) = \tau_i$, $i = 1, \ldots, n$. Consequently, we have $\boldsymbol{\tau} = \boldsymbol{\nu}(\boldsymbol{\theta})$.

Remark 1: Problems 1 and 2 are equivalent. To see this, suppose that $(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) \in \Theta \times \Xi$ is optimal for Problem 2 and let $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma \times \Xi$ be arbitrary. Then by (6), there exists a $\boldsymbol{\theta} \in \Theta$ such that $\boldsymbol{\tau} = \boldsymbol{\nu}(\boldsymbol{\theta})$. Hence, since $(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$ is optimal for Problem 2

$$J(\boldsymbol{\nu}(\boldsymbol{\theta}^*), \boldsymbol{\sigma}^*) = \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) \le \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\sigma})$$
$$= J(\boldsymbol{\nu}(\boldsymbol{\theta}), \boldsymbol{\sigma}) = J(\boldsymbol{\tau}, \boldsymbol{\sigma})$$

which shows that $(\boldsymbol{\nu}(\boldsymbol{\theta}^*), \boldsymbol{\sigma}^*) \in \Gamma \times \Xi$ is optimal for Problem 1. Similarly, if $(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*) \in \Gamma \times \Xi$ is optimal for Problem 1, then there exists a $\boldsymbol{\theta}^* \in \Theta$ such that $\boldsymbol{\tau}^* = \boldsymbol{\nu}(\boldsymbol{\theta}^*)$ and $(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$ is optimal for Problem 2.

Remark 2: Let $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$ be such that $\mu(q|\boldsymbol{\theta}) = T$ for some $q \in \{1, \dots, n\}$. Then we have $\mu(q|\boldsymbol{\theta}) = \mu(n+1|\boldsymbol{\theta})$, so (4) implies that $\theta_j = 0, j = q+1, \dots, n+1$. Hence

$$\tilde{J}(\boldsymbol{\theta}, \boldsymbol{\sigma}) = \sum_{i=1}^{q} \int_{i-1}^{i} \theta_i \mathcal{L}_i(\tilde{\mathbf{x}}(s|\boldsymbol{\theta}, \boldsymbol{\sigma}), \boldsymbol{\sigma}) ds.$$

In this case, the objective function value does not depend on $\tilde{\mathbf{x}}(s|\boldsymbol{\theta}, \boldsymbol{\sigma})$, $s \in [q, n+1]$, and therefore the condition $\mu(i|\boldsymbol{\theta}) < T$ in (10b) can be omitted. Using properties (b)-(c), the state jump conditions (10) can be expressed more concisely as follows:

$$\tilde{\mathbf{x}}(i^+) = \begin{cases} \mathbf{x}^0, & \text{if } i = 0, \quad (11a) \\ \tilde{\mathbf{x}}(i^-) + \chi(\theta_i) \boldsymbol{\phi}^i(\tilde{\mathbf{x}}(i^-), \boldsymbol{\sigma}), & \text{if } i \in \{1, \dots, n\} \\ & \hspace{4.5cm} (11b) \end{cases}$$

where $\chi(0) \triangleq 0$ and $\chi(\eta) \triangleq 1$ whenever $\eta > 0$.

Remark 3: The switched system for Problem 2 is (9), (11), which has switching times at the fixed points $s = i, i = 1, \dots, n$, and is therefore easier to integrate numerically than (1)–(2).

Remark 4: Let $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$, where $\theta_i = 0$ and $\theta_{i-1} > 0$ for some $i \in \{2, \dots, n+1\}$. Then in view of (4)–(5), the new switching times $s = i-1$ and $s = i$ correspond to the same point $\nu_{i-1}(\boldsymbol{\theta}) = \nu_i(\boldsymbol{\theta})$ in the original time horizon. For this reason, a jump is applied to $\tilde{\mathbf{x}}(\cdot|\boldsymbol{\theta}, \boldsymbol{\sigma})$ at $s = i-1$ but not at $s = i$; see (10b)–(10c) or (11b). In contrast, if the technique from [4], [5] is used to transform Problem 1, then separate jumps will be applied at both $s = i-1$ and $s = i$. In this case, $\mathbf{x}(\cdot|\boldsymbol{\nu}(\boldsymbol{\theta}), \boldsymbol{\theta})$ experiences a "double jump" at $t = \nu_{i-1}(\boldsymbol{\theta}) = \nu_i(\boldsymbol{\theta})$. Such "double jumps" do not reflect the original system described in Section II.

IV. APPROXIMATE PROBLEMS

Since the state jump conditions (11) contain a discontinuous function, Problem 2 cannot be solved using conventional optimization methods. To overcome this, we approximate (11) by

$$\tilde{\mathbf{x}}(i^+) = \begin{cases} \mathbf{x}^0, & \text{if } i = 0, \quad (12a) \\ \tilde{\mathbf{x}}(i^-) + \chi_\epsilon(\theta_i) \boldsymbol{\phi}^i(\tilde{\mathbf{x}}(i^-), \boldsymbol{\sigma}), & \text{if } i \in \{1, \dots, n\} (12b) \end{cases}$$

where $\epsilon > 0$ and

$$\chi_\epsilon(\eta) \triangleq \begin{cases} -\frac{2}{\epsilon^3}\eta^3 + \frac{3}{\epsilon^2}\eta^2, & \text{if } 0 \le \eta \le \epsilon, \\ 1, & \text{if } \eta > \epsilon. \end{cases}$$

Note that χ_ϵ is continuously differentiable and $\chi_\epsilon(\eta) = \chi(\eta)$ whenever $\eta \notin (0, \epsilon)$. Let $\tilde{\mathbf{x}}^\epsilon(\cdot|\boldsymbol{\theta}, \boldsymbol{\sigma})$ denote the solution of (9), (12) corresponding to $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$ and $\epsilon > 0$. Furthermore, define

$$\tilde{J}^\epsilon(\boldsymbol{\theta}, \boldsymbol{\sigma}) \triangleq \sum_{i=1}^{n+1} \int_{i-1}^{i} \theta_i \mathcal{L}_i(\tilde{\mathbf{x}}^\epsilon(s|\boldsymbol{\theta}, \boldsymbol{\sigma}), \boldsymbol{\sigma}) ds.$$

For given parameters $\epsilon > 0$ and $\gamma > 0$, consider the following optimization problem.

Problem 3: Choose $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$ to minimize the objective function

$$\tilde{G}^{\epsilon,\gamma}(\boldsymbol{\theta}, \boldsymbol{\sigma}) \triangleq \tilde{J}^\epsilon(\boldsymbol{\theta}, \boldsymbol{\sigma}) + \gamma \sum_{i=1}^{n} \chi_\epsilon(\theta_i)(1 - \chi_\epsilon(\theta_i)).$$

Unlike χ, the continuous function χ_ϵ can assume values in $(0,1)$. Hence, the new state jump conditions (12) are not always an accurate approximation of (11). The last term in $\tilde{G}^{\epsilon,\gamma}$ is used to penalize "fractional jumps", so that (12) is a good reflection of (11) at the optimal solution of Problem 3. It is also evident that $\chi_\epsilon \to \chi$ pointwise on $[0, \infty)$ as $\epsilon \to 0$. We thus expect that Problem 3 is a good approximation of Problem 2 when ϵ is small and γ is large. This observation is made rigorous in the next section.

Since (12) is constructed from smooth functions, the partial derivatives of $\tilde{G}^{\epsilon,\gamma}$ can be computed using gradient formulae given in [4], [5]. These formulae depend on the solution of a so-called costate system, whose value at the terminal time is known. Therefore, the costate system is integrated in the opposite direction to the state system (recall that the initial state is given). Hence, given $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$, the gradient of the objective function for Problem 3 can be computed via the following algorithm: i) Solve the state system (9), (12) forward in time; ii) Solve the costate system backwards in time; iii) Evaluate the gradient formulae. This procedure can be applied in conjunction with a gradient-based global optimization technique, such as the filled function method [13], to solve Problem 3. Alternatively, if n and r are not too large, then Problem 3 can be solved by repeatedly applying a local optimization technique with different starting points.

It is important to note that $\tilde{G}_{\epsilon,\gamma}$ is non-convex in general. In fact, since $\tilde{G}_{\epsilon,\gamma}$ contains a penalty term that approximates a discontinuous function, Problem 3 will usually have many local solutions. Accordingly, it is imperative that a global search strategy, such as those suggested above, is used to solve Problem 3.

V. CONVERGENCE RESULTS

In the previous section, we introduced a class of approximate problems for Problem 2 and discussed how each problem in this class can be solved using existing optimization methods. In this section, we will establish an important convergence result that links the solutions of the approximate problems with the solution of Problem 2.

Lemma 1: There exists a function $\vartheta : (0, \infty) \to (0, \infty)$ of order $O(1/\epsilon)$ such that whenever $\boldsymbol{\theta}^1, \boldsymbol{\theta}^2 \in \Theta$ and $\boldsymbol{\sigma} \in \Xi$

$$\left| \tilde{J}^\epsilon(\boldsymbol{\theta}^1, \boldsymbol{\sigma}) - \tilde{J}^\epsilon(\boldsymbol{\theta}^2, \boldsymbol{\sigma}) \right| \le \vartheta(\epsilon) \left\| \boldsymbol{\theta}^1 - \boldsymbol{\theta}^2 \right\|.$$

Proof: First, inequality (3) and Gronwall's Lemma can be used to show that the state and costate functions (see the gradient formulae in [4], [5]) are equibounded on the interval $[0, n+1]$ with respect to $\epsilon > 0$ and $(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Theta \times \Xi$. Furthermore, it is not difficult to see that

$$|\chi_\epsilon'(\eta)| = \chi_\epsilon'(\eta) \le \frac{3}{(2\epsilon)}, \quad \eta \ge 0.$$

The function ϑ is constructed by invoking Taylor's Theorem and applying the above results to the gradient formulae in [4], [5]. ∎

For the remainder of this section, let $(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) \in \Theta \times \Xi$, where $\epsilon > 0$ and $\gamma > 0$, denote an optimal solution of Problem 3. Additionally, define

$$\hat{\epsilon} \triangleq \min \left\{ 1, \frac{T}{(n^2 + 2n + 1)} \right\}.$$

Theorem 1: For each $\epsilon \in (0, \hat{\epsilon})$, there exists a corresponding $\gamma_\epsilon > 0$ such that if $\gamma \ge \gamma_\epsilon$, then

$$\theta_i^{\epsilon,\gamma,*} < \frac{\epsilon^{3/2}}{2} \quad or \quad \theta_i^{\epsilon,\gamma,*} > \epsilon - \frac{\epsilon^{3/2}}{2}, \quad i = 1, \dots, n. \quad (13)$$

Proof: Let $(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*}) \in \Theta \times \Xi$ denote the minimum of $\tilde{J}^{\epsilon}$ on $\Theta \times \Xi$. Then

$$\tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*}) \leq \tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}), \quad \gamma > 0.$$

Adding a penalty term to both sides of this inequality gives

$$\tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*}) + \gamma \sum_{i=1}^{n} \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})(1 - \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*}))$$
$$\leq \tilde{G}^{\epsilon,\gamma}(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}).$$

Hence, since $(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$ is optimal for Problem 3

$$\tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*}) + \gamma \sum_{i=1}^{n} \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})(1 - \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*}))$$
$$\leq \tilde{G}^{\epsilon,\gamma}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\sigma}}), \ \gamma > 0 \quad (14)$$

where $\bar{\boldsymbol{\sigma}} \in \Xi$ and $\bar{\theta}_i \triangleq T/(n+1)$, $i = 1, \ldots, n+1$. Since

$$\epsilon < \hat{\epsilon} \leq \frac{T}{(n+1)}$$

we have

$$\chi_{\epsilon}(\bar{\theta}_i) = 1, \quad i = 1, \ldots, n.$$

Thus, the penalty term in $\tilde{G}^{\epsilon,\gamma}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\sigma}})$ vanishes and inequality (14) becomes

$$\tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*}) + \gamma \sum_{i=1}^{n} \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})(1 - \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})) \leq \tilde{J}^{\epsilon}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\sigma}}).$$

This can be rearranged to give

$$\chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})(1 - \chi_{\epsilon}(\theta_i^{\epsilon,\gamma,*})) \leq \frac{\tilde{J}^{\epsilon}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\sigma}}) - \tilde{J}^{\epsilon}(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*})}{\gamma} \quad (15)$$

for each $i = 1, \ldots, n$. Since $(\boldsymbol{\theta}^{\epsilon,*}, \boldsymbol{\sigma}^{\epsilon,*})$ is the minimum of $\tilde{J}^{\epsilon}$, the right-hand side of (15) is non-negative. Hence, by choosing a sufficiently large value for γ, we can make the penalty term on the left-hand side arbitrarily small. This forces $\theta_i^{\epsilon,\gamma,*}$ to either zero or exceed ϵ. ∎

Theorem 2: If $\epsilon \in (0, \hat{\epsilon})$, $\gamma > 0$, and $\boldsymbol{\theta}^{\epsilon,\gamma,*}$ satisfies (13), then there exists a $\hat{\boldsymbol{\theta}}^{\epsilon,\gamma} \in \Theta$ such that

$$\left\| \hat{\boldsymbol{\theta}}^{\epsilon,\gamma} - \boldsymbol{\theta}^{\epsilon,\gamma,*} \right\| \leq n\epsilon^{3/2} \quad (16)$$

and

$$\chi_{\epsilon}(\hat{\theta}_i^{\epsilon,\gamma}) = \chi(\hat{\theta}_i^{\epsilon,\gamma}), \quad i = 1, \ldots, n. \quad (17)$$

Proof: Define

$$\mathcal{I}_1^{\epsilon,\gamma} \triangleq \left\{ i \in \{1, \ldots, n\} : \ 0 < \theta_i^{\epsilon,\gamma,*} < \frac{\epsilon^{3/2}}{2} \right\} \quad (18)$$

and

$$\mathcal{I}_2^{\epsilon,\gamma} \triangleq \left\{ i \in \{1, \ldots, n\} : \ \epsilon - \frac{\epsilon^{3/2}}{2} < \theta_i^{\epsilon,\gamma,*} < \epsilon \right\}. \quad (19)$$

Since $\epsilon < 1$, the index sets $\mathcal{I}_1^{\epsilon,\gamma}$ and $\mathcal{I}_2^{\epsilon,\gamma}$ are disjoint. Let $\kappa_{\epsilon,\gamma}$ be an integer in $\{1, \ldots, n+1\}$ such that

$$\theta_{\kappa_{\epsilon,\gamma}}^{\epsilon,\gamma,*} = \max_{1 \leq i \leq n+1} \theta_i^{\epsilon,\gamma,*} \geq \frac{T}{n+1} > \epsilon. \quad (20)$$

It is clear from (20) that $\kappa_{\epsilon,\gamma} \notin \mathcal{I}_1^{\epsilon,\gamma} \cup \mathcal{I}_2^{\epsilon,\gamma}$. For each $i = 1, \ldots, n+1$, define

$$\hat{\theta}_i^{\epsilon,\gamma} \triangleq \begin{cases} 0, & \text{if } i \in \mathcal{I}_1^{\epsilon,\gamma}, \\ \epsilon, & \text{if } i \in \mathcal{I}_2^{\epsilon,\gamma}, \\ \theta_i^{\epsilon,\gamma,*} - \zeta_{\epsilon,\gamma}, & \text{if } i = \kappa_{\epsilon,\gamma}, \\ \theta_i^{\epsilon,\gamma,*}, & \text{otherwise}, \end{cases}$$

$$\text{where}$$

$$\zeta_{\epsilon,\gamma} \triangleq \sum_{i \in \mathcal{I}_2^{\epsilon,\gamma}} (\epsilon - \theta_i^{\epsilon,\gamma,*}) - \sum_{i \in \mathcal{I}_1^{\epsilon,\gamma}} \theta_i^{\epsilon,\gamma,*}.$$

We will show that $\hat{\boldsymbol{\theta}}^{\epsilon,\gamma} \triangleq [\hat{\theta}_1^{\epsilon,\gamma}, \ldots, \hat{\theta}_{n+1}^{\epsilon,\gamma}]^T$ is an element of Θ satisfying (16)–(17). First, it follows from (18)–(19) that:

$$|\zeta_{\epsilon,\gamma}| < \frac{n\epsilon^{3/2}}{2}. \quad (21)$$

Using (18)–(19), (21), we obtain

$$\left\| \hat{\boldsymbol{\theta}}^{\epsilon,\gamma} - \boldsymbol{\theta}^{\epsilon,\gamma,*} \right\|^2 = \zeta_{\epsilon,\gamma}^2 + \sum_{i \in \mathcal{I}_1^{\epsilon,\gamma}} (\theta_i^{\epsilon,\gamma,*})^2$$
$$+ \sum_{i \in \mathcal{I}_2^{\epsilon,\gamma}} (\epsilon - \theta_i^{\epsilon,\gamma,*})^2$$
$$< n^2\epsilon^3$$

proving (16). Next, combining (20) and (21) gives

$$\hat{\theta}_{\kappa_{\epsilon,\gamma}}^{\epsilon,\gamma} = \theta_{\kappa_{\epsilon,\gamma}}^{\epsilon,\gamma,*} - \zeta_{\epsilon,\gamma}$$
$$> \frac{T}{n+1} - \frac{n\epsilon^{3/2}}{2} > \frac{T}{n+1} - n\epsilon^{3/2}. \quad (22)$$

Since $\epsilon < \hat{\epsilon}$, the inequality $(n+1)\epsilon < T/(n+1)$ holds. Substituting this into (22) gives

$$\hat{\theta}_{\kappa_{\epsilon,\gamma}}^{\epsilon,\gamma} > (n+1)\epsilon - n\epsilon^{3/2} = \epsilon + n(\epsilon - \epsilon^{3/2}) > \epsilon \quad (23)$$

where the last inequality follows from $\epsilon < 1$. Equation (17) is a consequence of (13), (23), and the definition of $\hat{\theta}_i^{\epsilon,\gamma}$. Finally, we have

$$\sum_{i=1}^{n+1} \hat{\theta}_i^{\epsilon,\gamma} = \sum_{i \in \mathcal{I}_2^{\epsilon,\gamma}} \epsilon + \theta_{\kappa_{\epsilon,\gamma}}^{\epsilon,\gamma,*} - \zeta_{\epsilon,\gamma} + \sum_{\substack{i \notin \mathcal{I}_1^{\epsilon,\gamma} \cup \mathcal{I}_2^{\epsilon,\gamma} \\ i \neq \kappa_{\epsilon,\gamma}}} \theta_i^{\epsilon,\gamma,*}$$
$$= \sum_{i=1}^{n+1} \theta_i^{\epsilon,\gamma,*} = T. \quad (24)$$

Equations (23) and (24), together with the definition of $\hat{\theta}_i^{\epsilon,\gamma}$ show that $\hat{\boldsymbol{\theta}}^{\epsilon,\gamma} \in \Theta$. ∎

The next theorem is the main result of this section.

Theorem 3: Suppose that $(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) \in \Theta \times \Xi$ is an optimal solution of Problem 2. Furthermore, for each $\epsilon \in (0, \hat{\epsilon})$, let γ be sufficiently large so that (13) is satisfied (recall from Theorem 1 that such a γ exists) and let $\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}$ be as defined in Theorem 2. Then

$$\tilde{J}(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) - \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) = O(\sqrt{\epsilon}).$$

Proof: Define

$$\bar{\epsilon} \triangleq \min_{1 \leq i \leq n} \{\theta_i^* : \ \theta_i^* > 0\}$$

and let $\epsilon < \min(\hat{\epsilon}, \bar{\epsilon})$. Then for each $i = 1, \ldots, n$, either $\theta_i^* > \epsilon$ or $\theta_i^* = 0$. Hence

$$\chi_{\epsilon}(\theta_i^*) = \chi(\theta_i^*), \quad i = 1, \ldots, n.$$

This means that (11) coincides with its approximation (12) when $(\boldsymbol{\theta}, \boldsymbol{\sigma}) = (\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$. Consequently

$$\mathbf{x}(s|\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) = \mathbf{x}^\epsilon(s|\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*), \qquad s \in [0, n+1]$$

and so[3]

$$\tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) = \tilde{J}^\epsilon(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) = \tilde{G}^{\epsilon,\gamma}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*).$$

Similarly, (17) shows that $\mathbf{x}(s|\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) = \mathbf{x}^\epsilon(s|\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$ for all $s \in [0, n+1]$, and

$$\tilde{J}(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) = \tilde{J}^\epsilon(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}). \tag{25}$$

Now, since the penalty term in $\tilde{G}^{\epsilon,\gamma}$ is non-negative and $(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$ is optimal for Problem 3

$$\tilde{J}^\epsilon(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) \leq \tilde{G}^{\epsilon,\gamma}(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$$
$$\leq \tilde{G}^{\epsilon,\gamma}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) = \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*).$$

Consequently, $-\tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) \leq -\tilde{J}^\epsilon(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$. By using this result and (25), we obtain

$$0 \leq \tilde{J}(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) - \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$$
$$\leq \tilde{J}^\epsilon(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) - \tilde{J}^\epsilon(\boldsymbol{\theta}^{\epsilon,\gamma,*}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}).$$

(The lower bound is zero here because $(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$ is optimal for Problem 2.) Therefore, applying Lemma 1 gives

$$0 \leq \tilde{J}(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) - \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*) \leq \vartheta(\epsilon) \left\| \hat{\boldsymbol{\theta}}^{\epsilon,\gamma} - \boldsymbol{\theta}^{\epsilon,\gamma,*} \right\|.$$

Since $\vartheta(\epsilon) = O(1/\epsilon)$, estimate (16) ensures that the right-hand side of the above inequality is a function of order $O(\sqrt{\epsilon})$. ∎

Remark 5: Theorem 3 implies that $\tilde{J}(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*}) \rightarrow \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\sigma}^*)$ as $\epsilon \rightarrow 0$, where γ is chosen so that (13) is satisfied. As a consequence, we can solve Problem 1 numerically as follows. First, transform Problem 1 into Problem 2 as shown in Section III. Then, choose an initial $\epsilon \in (0, \hat{\epsilon})$ and solve Problem 3 for increasing values of $\gamma > 0$ until (13) is satisfied (Theorem 1 ensures that Problem 3 only needs to be solved a finite number of times here). Next, construct $\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}$ from $\boldsymbol{\theta}^{\epsilon,\gamma,*}$ using the formula given in the proof of Theorem 2, and decrease ϵ. Repeat these steps until ϵ is sufficiently small; at this stage, $(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$ is a good approximation of the optimal solution for Problem 2. In view of Remark 1, we can use $(\hat{\boldsymbol{\theta}}^{\epsilon,\gamma}, \boldsymbol{\sigma}^{\epsilon,\gamma,*})$ to construct an approximate solution for Problem 1.

VI. ILLUSTRATIVE EXAMPLE

Consider the switched-capacitor dc/dc power converter discussed in [14]. This electrical circuit, which is designed to deliver half of the input voltage to an attached load, consists of three primary capacitors and four switches. For each $i = 1, 2, 3$, let $x_i(t)$ denote the voltage across the ith primary capacitor at time t. Let $y(t)$ denote the output voltage at time t. We assume that the dc input is 3.6 V and the load resistance is 75 Ω.

The function of each capacitor—in particular, whether it stores energy from the input or delivers energy to the load—is determined by the switch configuration. Consequently, modifying the switch configuration changes the circuit topology of the power converter. Three distinct topologies are possible. The ith topology is modeled by the dynamics

$$\dot{\mathbf{x}}(t) = A^i \mathbf{x}(t) + 3.6B^i$$
$$y(t) = C^i \mathbf{x}(t) + 3.6D^i$$

where $A^i \in \mathbb{R}^{3 \times 3}$, $B^i \in \mathbb{R}^{3 \times 1}$, $C^i \in \mathbb{R}^{1 \times 3}$, and $D^i \in \mathbb{R}$ are obtained using Kirchhoff's laws. The output voltage is regulated to the desired

---

1.8 V (half of the 3.6 V input) by switching between these topologies in an appropriate manner.

An *operating schedule* for the power converter specifies the order in which topologies are operated (the switching sequence) and the times at which the topologies are switched (the switching times). Since the ideal output is 1.8 V, the operating schedule should be chosen to minimize

$$\int_0^T (y(t) - 1.8)^2 dt \tag{26}$$

where here the terminal time $T = 1.0 \times 10^{-4}$ seconds. We use the method suggested in [2] and model the power converter by the following dynamics:

$$\dot{\mathbf{x}}(t) = A^{\alpha_j} \mathbf{x}(t) + 3.6B^{\alpha_j}$$
$$t \in (\tau_{j-1}, \tau_j), \quad j = 1, \ldots, 9 \tag{27}$$
$$y(t) = C^{\alpha_j} \mathbf{x}(t) + 3.6D^{\alpha_j}$$
$$t \in [\tau_{j-1}, \tau_j), \quad j = 1, \ldots, 9 \tag{28}$$

where $\tau_0 \triangleq 0$; $\tau_9 \triangleq 1.0 \times 10^{-4}$; $\tau_j, j = 1, \ldots, 8$, are switching times such that $\tau_{j-1} \leq \tau_j$; and

$$\alpha_j \triangleq \mathrm{mod}(j-1, 3) + 1, \quad j = 1, \ldots, 9.$$

Equations (27)–(28) can replicate any operating schedule. For example, if

$$0 = \tau_0 = \tau_1 < \tau_2 = \tau_3 < \tau_4 = \tau_5 < \tau_6 = \tau_7 = \tau_8 = \tau_9 = T$$

then the switching sequence is $\{2, 1, 3\}$[4] and the switching times are $\tau_2 = \tau_3$ and $\tau_4 = \tau_5$. Clearly, unnecessary subsystems in (27)–(28) are "deleted" by combining some of the switches.

In practice, topology switches are accompanied by a voltage leak from the capacitors in the circuit [6], [7]. We assume that this leak is 10% of the voltage immediately before the switch. Hence, the following state jump conditions are imposed:

$$\mathbf{x}(\tau_j^+) = \begin{cases} [0, 0, 0]^T, & \text{if } j = 0, \tag{29a} \\ 0.9\mathbf{x}(\tau_j^-), & \text{if } j \in \{1, \ldots, 8\}, \tau_{j-1} < \tau_j < T. \end{cases}$$
$$\tag{29b}$$

Our optimal control problem is as follows: Choose switching times $\tau_j, j = 1, \ldots, 8$, such that (26) is minimized subject to the switched system (27)–(29). We solved this problem using a Fortran 90 implementation of the algorithm discussed in Remark 5. In this implementation, Problem 3 was solved by applying the optimization routine NLPQLP [15] ten times with random starting points. Initially, $\epsilon = 0.9\hat{\epsilon}$ (here, $n = 8$ and $T = 1.0 \times 10^{-4}$) and $\gamma = 2.0$. The algorithm was terminated when $\epsilon < 9\hat{\epsilon}/10^4$.

The optimal switching times in (27)–(28) are

$$\tau_0^* = \tau_1^* = 0.0$$
$$\tau_2^* = \tau_3^* = \tau_4^* = 1.0122 \times 10^{-6}$$
$$\tau_5^* = \tau_6^* = 9.4824 \times 10^{-6}$$
$$\tau_7^* = 1.9633 \times 10^{-5}$$
$$\tau_8^* = \tau_9^* = 1.0 \times 10^{-4}.$$

This solution "deletes" subsystems $\{1, 3, 4, 6, 9\}$ in (27)–(28). In fact, the optimal switching sequence is $\{2, 1, 2\}$. Fig. 1 shows the voltage across the load and capacitors for this optimal operating schedule.

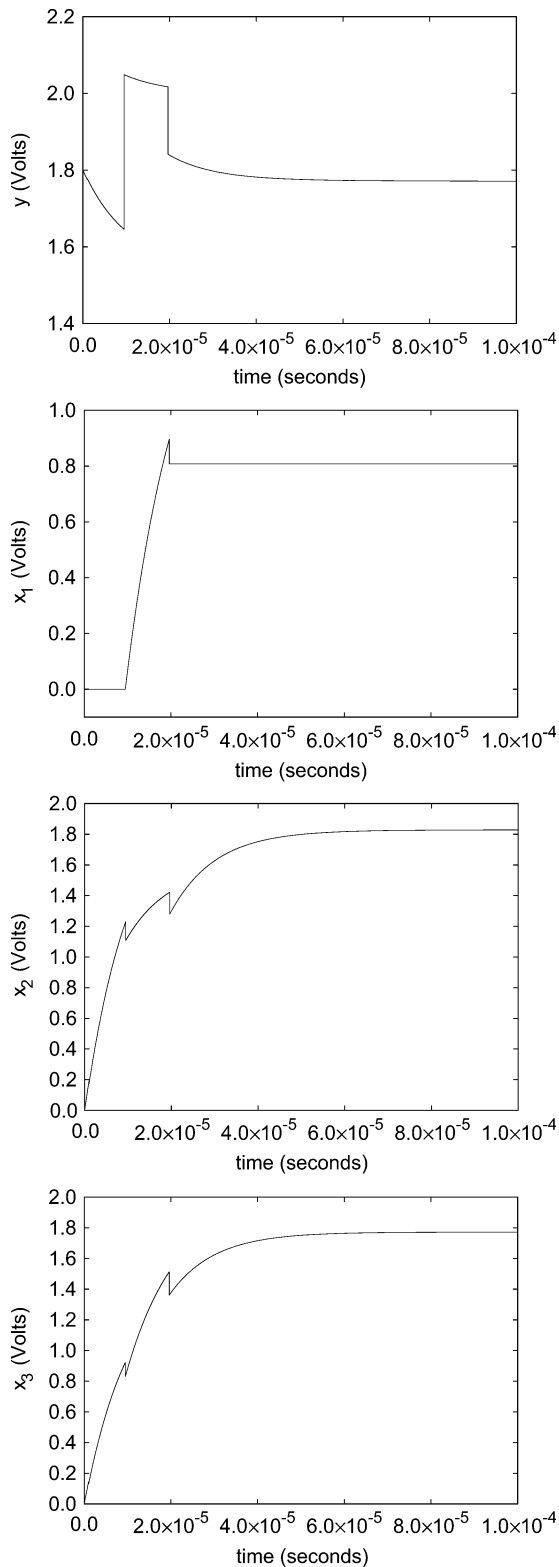[4]This means that the topologies are operated in the following order: Topology 2, Topology 1, Topology 3.

Fig. 1. Output and state voltages for the optimal operating schedule.

REFERENCES

[1] R. Li, K. L. Teo, K. H. Wong, and G. R. Duan, "Control parameterization enhancing transform for optimal control of switched systems," *Math. Comput. Modelling*, vol. 43, pp. 1393–1403, Jun. 2006.

[2] T. Ruby and V. Rehbock, "Numerical solutions of optimal switching control problems," in *Optimization and Control with Applications*, L. Qi, K. L. Teo, and X. Yang, Eds. New York: Springer, 2005, pp. 447–459.

[3] X. Xu and P. J. Antsaklis, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Trans. Autom. Control*, vol. 49, no. 1, pp. 2–16, Jan. 2004.

[4] Y. Liu, K. L. Teo, L. S. Jennings, and S. Wang, "On a class of optimal control problems with state jumps," *J. Optim. Theory Appl.*, vol. 98, pp. 65–82, Jul. 1998.

[5] C. Z. Wu and K. L. Teo, "Global impulsive optimal control computation," *J. Ind. Manag. Optim.*, vol. 2, pp. 435–450, Nov. 2006.

[6] C. Y. F. Ho, B. W. K. Ling, Y. Q. Liu, P. K. S. Tam, and K. L. Teo, "Optimal PWM control of switched-capacitor dc-dc power converters via model transformation and enhancing control techniques," *IEEE Trans. Circuits Syst. I*, vol. 55, no. 6, pp. 1382–1391, Jun. 2008.

[7] R. C. Loxton, K. L. Teo, V. Rehbock, and W. K. Ling, "Optimal switching instants for a switched-capacitor dc/dc power converter," *Automatica*, vol. 45, pp. 973–980, Apr. 2009.

[8] C. Gao, E. Feng, Z. Wang, and Z. Xiu, "Nonlinear dynamical systems of bio-dissimilation of glycerol to 1,3-propanediol and their optimal controls," *J. Ind. Manag. Optim.*, vol. 1, pp. 377–388, Aug. 2005.

[9] C. Gao, K. Li, E. Feng, and Z. Xiu, "Nonlinear impulsive system of fed-batch culture in fermentative production and its properties," *Chaos, Solitons Fractals*, vol. 28, pp. 271–277, Apr. 2006.

[10] J. P. Hespanha and S. Morse, "Switching between stabilizing controllers," *Automatica*, vol. 38, pp. 1905–1917, Nov. 2002.

[11] D. Chatterjee and D. Liberzon, "Stability analysis of deterministic and stochastic switched systems via a comparison principle and multiple lyapunov functions," *SIAM J. Control Optim.*, vol. 45, pp. 174–206, 2006.

[12] L. Vu and D. Liberzon, "Common Lyapunov functions for families of commuting nonlinear systems," *Syst. Control Lett.*, vol. 54, pp. 405–416, May 2005.

[13] Z. Y. Wu, F. S. Bai, H. W. J. Lee, and Y. J. Yang, "A filled function method for constrained global optimization," *J. Global Optim.*, vol. 39, pp. 495–507, Dec. 2007.

[14] T. Umeno, K. Takahashi, I. Oota, F. Ueno, and T. Inoue, "New switched-capacitor dc-dc converter with low input current ripple and its hybridization," in *Proc. 33rd Midwest Symp. Circuits Syst.*, 1990, pp. 1091–1094.

[15] K. Schittkowski, *NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm With Distributed and Non-Monotone Line Search*. Bayreuth, Germany: Dept. Comp. Sci., Univ. Bayreuth, 2007.