Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

CONTROL PARAMETERIZATION FOR OPTIMAL CONTROL PROBLEMS WITH CONTINUOUS INEQUALITY CONSTRAINTS: NEW CONVERGENCE RESULTS

Ryan Loxton, Qun Lin, Volker Rehbock, and Kok Lay Teo

Department of Mathematics and Statistics
Curtin University
GPO Box U1987 Perth, Western Australia 6845

Abstract. Control parameterization is a powerful numerical technique for solving optimal control problems with general nonlinear constraints. The main idea of control parameterization is to discretize the control space by approximating the control by a piecewise-constant or piecewise-linear function, thereby yielding an approximate nonlinear programming problem. This approximate problem can then be solved using standard gradient-based optimization techniques. In this paper, we consider the control parameterization method for a class of optimal control problems in which the admissible controls are functions of bounded variation and the state and control are subject to continuous inequality constraints. We show that control parameterization generates a sequence of suboptimal controls whose costs converge to the true optimal cost. This result has previously only been proved for the case when the admissible controls are restricted to piecewise continuous functions.

1. Introduction.

Real-world optimal control problems often involve *continuous inequality constraints* that restrict the state and/or control variables at every point in the time horizon. Such constraints are also called *path constraints*, *all-time constraints*, or *semi-infinite constraints* in the literature. They arise in many practical applications, such as chemistry [22], robotics [4], spacecraft control [1], underwater vehicles [3], zinc sulphate purification [20], and DC-DC power converters [11].

The control parameterization method (see [5, 12, 17]) is a popular numerical method for solving optimal control problems with continuous inequality constraints. This method involves partitioning the time horizon into a set of subintervals, and then approximating the control by a constant value on each subinterval. The optimal control problem is subsequently reduced to an approximate *semi-infinite programming problem*, which can be solved using existing techniques such as the constraint transcription methods in [6, 19], or the recently-developed exact penalty methods in [8, 21]. After solving the approximate problem, a suboptimal control for the original optimal control problem is easily obtained.

Convergence is an important issue for any numerical technique, and control parameterization is no exception. The central question is: how close is the suboptimal

2000 *Mathematics Subject Classification.* Primary: 65P99, 93C15; Secondary: 49M37.

Key words and phrases. Optimal control, Continuous inequality constraints, Control parameterization.

R. Loxton, V. Rehbock, and K. L. Teo are supported by grant from the Australian Research Council (Discovery Project DP110100083).

control generated by control parameterization to the true optimal control? In [17], it is shown that the cost of the suboptimal control converges to the true optimal cost as the number of subintervals approaches infinity. However, the proof of this result is only valid when the continuous inequality constraints are *pure state constraints*—i.e. constraints that only involve the state variables. In [12], improved convergence results are derived for the more difficult case in which the continuous inequality constraints restrict *both* the state and the control. However, these improved results come at a price: they require that the class of admissible controls consist only of piecewise continuous functions, whereas in [17] general measurable functions are allowed.

In this paper, we consider a class of optimal control problems in which the admissible controls are functions of bounded variation, the state and control are subject to continuous inequality constraints, and the cost function includes a term that penalizes changes in the control action. Our aim is to show that for this class of problems, control parameterization generates a suboptimal control whose cost converges to the true optimal cost as the discretization of the time horizon is refined. This new result supersedes the main convergence result in Chapter 10 of [17], which is only applicable to problems with pure state constraints.

2. **Problem Formulation.**

Consider the following dynamic system:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{u}(t)), \quad t \in [0, T], \tag{1}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}^0, \tag{2}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the *state* at time t, $\boldsymbol{u}(t) \in \mathbb{R}^r$ is the *control* at time t, $\boldsymbol{x}^0 \in \mathbb{R}^n$ is a given initial state, T is a given *terminal time*, and $\boldsymbol{f} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ is a given continuously differentiable function.

Let $u_i : [0, T] \to \mathbb{R}$ denote the ith component of $\boldsymbol{u} : [0, T] \to \mathbb{R}^r$. Then the *total variation* of u_i is defined by

$$\bigvee_0^T u_i := \sup \sum_{j=1}^{m} \left| u_i(t_j) - u_i(t_{j-1}) \right|,$$

where the supremum is taken over all finite partitions $\{t_j\}_{j=0}^m \subset [0, T]$ satisfying

$$0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = T.$$

The total variation of the vector-valued function $\boldsymbol{u} : [0, T] \to \mathbb{R}^r$ is defined by

$$\bigvee_0^T \boldsymbol{u} := \sum_{i=1}^{r} \bigvee_0^T u_i.$$

If the total variation of $\boldsymbol{u} : [0, T] \to \mathbb{R}^r$ is finite, then we say that $\boldsymbol{u}$ is of *bounded variation*. Let $\mathcal{U}$ denote the class of all such functions of bounded variation mapping $[0, T]$ into $\mathbb{R}^r$. Any $\boldsymbol{u} \in \mathcal{U}$ is called an *admissible control* for system (1)-(2).

Clearly, for each $\boldsymbol{u} \in \mathcal{U}$, there exists a corresponding real number $M > 0$ such that

$$\|\boldsymbol{u}(t)\| \leq M, \quad t \in [0, T],$$

where $\| \cdot \|$ denotes the Euclidean norm. Thus, each admissible control in $\mathcal{U}$ is bounded.

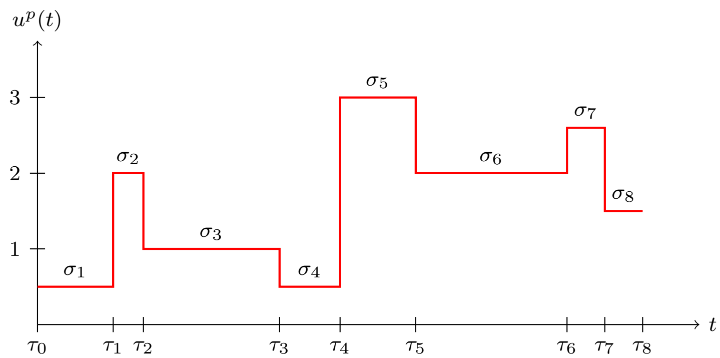

FIGURE 1. A piecewise-constant control approximation with $p = 8$.

As is customary (see [9, 12, 17]), we assume that there exists a constant $L > 0$ such that

$$\|\boldsymbol{f}(t, \boldsymbol{\xi}, \boldsymbol{\theta})\| \le L(1 + \|\boldsymbol{\xi}\| + \|\boldsymbol{\theta}\|), \quad (t, \boldsymbol{\xi}, \boldsymbol{\theta}) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^r. \tag{3}$$

This ensures that system (1)-(2) admits a unique Carathéodory solution corresponding to each admissible control $\boldsymbol{u} \in \mathcal{U}$ (see Theorem 3.3.3 in [2]). We denote this solution by $\boldsymbol{x}(\cdot|\boldsymbol{u})$.

Now, consider the following set of *continuous inequality constraints* involving both the state and the control:

$$h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) \ge 0, \quad t \in [0, T], \quad j = 1, \dots, q, \tag{4}$$

where each $h_j : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ is a given continuously differentiable function. Note that control bounds can be easily incorporated into (4).

Let $\mathcal{F}$ denote the set of all $\boldsymbol{u} \in \mathcal{U}$ satisfying (4). Controls in $\mathcal{F}$ are called *feasible controls*. Our optimal control problem is defined as follows.

Problem P. *Choose a feasible control $\boldsymbol{u} \in \mathcal{F}$ to minimize the cost functional*

$$J(\boldsymbol{u}) := \Phi(\boldsymbol{x}(T|\boldsymbol{u})) + \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t))dt + \gamma \bigvee_0^T \boldsymbol{u}, \tag{5}$$

where $\gamma \ge 0$ is a given weight and $\Phi : \mathbb{R}^n \to \mathbb{R}$ and $\mathcal{L} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ are given continuously differentiable functions.

The first term in (5) measures the system's *terminal cost* (as a function of the final state reached by the system), while the second term measures the system's *running cost* (as a function of the state and control at each time point). The last term in (5) is designed to penalize changes in the control input, and thereby discourage volatile control strategies that would be difficult to implement in practice.

3. **Control Parameterization.** To solve Problem P using the control parameterization method, we approximate $\boldsymbol{u}$ as follows:

$$\boldsymbol{u}(t) \approx \boldsymbol{u}^p(t) = \boldsymbol{\sigma}^k, \quad t \in [\tau_{k-1}, \tau_k), \quad k = 1, \dots, p,$$

where $p \ge 1$ is a given integer, τ_k, $k = 0, \dots, p$ are knot points, and $\boldsymbol{\sigma}^k \in \mathbb{R}^r$, $k = 1, \dots, p$ are vectors containing the approximate control values. This approximation scheme is illustrated in Figure 1.

The knot points satisfy

$$0 = \tau_0 \le \tau_1 \le \tau_2 \le \cdots \le \tau_{p-1} \le \tau_p = T. \tag{6}$$

The approximate control $\boldsymbol{u}^p$ can be written as

$$\boldsymbol{u}^p(t) = \sum_{k=1}^{p-1} \boldsymbol{\sigma}^k \chi_{[\tau_{k-1},\tau_k)}(t) + \boldsymbol{\sigma}^p \chi_{[\tau_{p-1},\tau_p]}(t), \tag{7}$$

where, for a given subinterval $\mathcal{I} \subset [0,T]$, the characteristic function $\chi_{\mathcal{I}} : \mathbb{R} \to \mathbb{R}$ is defined by

$$\chi_{\mathcal{I}}(t) := \begin{cases} 1, & \text{if } t \in \mathcal{I}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\boldsymbol{u}^p$ is a piecewise-constant function with potential discontinuities at the points $t = \tau_k$, $k = 1, \ldots, p-1$. These points are called *switching times*. Throughout this paper, we use the convention that $[\tau_{k-1}, \tau_k) = \emptyset$ if $\tau_{k-1} = \tau_k$.

Let σ_i^k denote the ith component of $\boldsymbol{\sigma}^k$. The following result shows that $\boldsymbol{u}^p$ is an admissible control for Problem P.

Theorem 3.1. *The piecewise-constant control* $\boldsymbol{u}^p$ *is of bounded variation with*

$$\bigvee_0^T \boldsymbol{u}^p \le \sum_{i=1}^{r} \sum_{k=1}^{p-1} \left|\sigma_i^{k+1} - \sigma_i^k\right|. \tag{8}$$

Proof. Let $\{t_j\}_{j=0}^m$ be an arbitrary partition of $[0,T]$ satisfying

$$0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = T.$$

Consider another partition of $[0,T]$ consisting of the control subintervals $[\tau_{k-1}, \tau_k)$, $k = 1, \ldots, p-1$ and $[\tau_{p-1}, \tau_p]$. Let $\kappa(j)$ denote the index of the unique control subinterval containing t_j. Then for each $j = 0, \ldots, m-1$, $\kappa(j)$ is the unique index in $\{1, \ldots, p\}$ such that

$$\tau_{\kappa(j)-1} \le t_j < \tau_{\kappa(j)}.$$

Furthermore, for $j = m$, we have $\kappa(m) = p$. Clearly, $\kappa(j)$ is non-decreasing in j.

For each $j = 1, \ldots, m$, let $\mathcal{E}_j$ denote the set of integers between $\kappa(j-1)$ and $\kappa(j) - 1$ inclusive. That is,

$$\mathcal{E}_j := \{\kappa(j-1), \ldots, \kappa(j) - 1\}, \quad j = 1, \ldots, m,$$

where $\mathcal{E}_j = \emptyset$ if $\kappa(j-1) = \kappa(j)$. Clearly,

$$\mathcal{E}_j \subset \{1, \ldots, p-1\}, \quad j = 1, \ldots, m. \tag{9}$$

We now show that $\{\mathcal{E}_j\}_{j=1}^m$ is a disjoint collection of subsets of $\{1, \ldots, p-1\}$. First, suppose that $\varsigma \in \mathcal{E}_{j'}$ and $\varsigma \in \mathcal{E}_{j''}$ for distinct integers j' and j'', where we assume without loss of generality that $j' < j''$. Then since $j' \le j'' - 1$, we must have $\kappa(j') < \kappa(j'')$ (otherwise $\mathcal{E}_{j''} = \emptyset$). Thus,

$$\varsigma \le \kappa(j') - 1 < \kappa(j') \le \kappa(j'' - 1) \le \varsigma.$$

But this is a contradiction. Hence,

$$\mathcal{E}_{j'} \cap \mathcal{E}_{j''} = \emptyset, \quad j' \ne j''. \tag{10}$$

Now,

$$\begin{aligned}\sum_{j=1}^{m}\left|u_i^p(t_j)-u_i^p(t_{j-1})\right| &= \sum_{j=1}^{m}\left|\sigma_i^{\kappa(j)}-\sigma_i^{\kappa(j-1)}\right| \\ &\leq \sum_{j=1}^{m}\sum_{l=\kappa(j-1)}^{\kappa(j)-1}\left|\sigma_i^{l+1}-\sigma_i^{l}\right| \\ &= \sum_{j=1}^{m}\sum_{l\in\mathcal{E}_j}\left|\sigma_i^{l+1}-\sigma_i^{l}\right|.\end{aligned}$$

Thus, in view of (9) and (10),

$$\sum_{j=1}^{m}\left|u_i^p(t_j)-u_i^p(t_{j-1})\right| \leq \sum_{j=1}^{m}\sum_{l\in\mathcal{E}_j}\left|\sigma_i^{l+1}-\sigma_i^{l}\right| \leq \sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|. \tag{11}$$

Since the right-hand side of (11) is independent of the partition $\{t_j\}_{j=0}^m$, we have

$$\bigvee_0^T u_i^p = \sup\sum_{j=1}^{m}\left|u_i^p(t_j)-u_i^p(t_{j-1})\right| \leq \sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|.$$

Consequently,

$$\bigvee_0^T \boldsymbol{u}^p = \sum_{i=1}^{r}\bigvee_0^T u_i^p \leq \sum_{i=1}^{r}\sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|,$$

as required. □

If the control knot points τ_k, $k=0,\dots,p$ are distinct, then $\{\tau_k\}_{k=0}^p$ is a valid partition of $[0,T]$ satisfying

$$0=\tau_0<\tau_1<\cdots<\tau_{p-1}<\tau_p=T.$$

Thus, by the definition of total variation,

$$\bigvee_0^T u_i^p \geq \sum_{k=1}^{p}\left|u_i^p(\tau_k)-u_i^p(\tau_{k-1})\right| = \sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|$$

and

$$\bigvee_0^T \boldsymbol{u}^p = \sum_{i=1}^{r}\bigvee_0^T u_i^p \geq \sum_{i=1}^{r}\sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|.$$

Combining this inequality with (8) yields

$$\bigvee_0^T \boldsymbol{u}^p = \sum_{i=1}^{r}\sum_{k=1}^{p-1}\left|\sigma_i^{k+1}-\sigma_i^{k}\right|.$$

Thus, if the control knot points are distinct, then inequality (8) in Theorem 3.1 holds with equality. This is the case in Chapter 10 of [17], where the knot points are assumed to be pre-fixed constants. Here, we have used a more flexible discretization scheme in which the knot points are decision variables to be chosen optimally.

Now, if the control knot points are not distinct—i.e. if two or more knot points coincide—then inequality (8) in Theorem 3.1 could be strict. For example, let $p=3$ and $r=1$, and define the knot points and control values as follows:

$$\tau_0=0,\quad \tau_1=3,\quad \tau_2=3,\quad \tau_3=8,\quad \sigma^1=3,\quad \sigma^2=0,\quad \sigma^3=1.$$

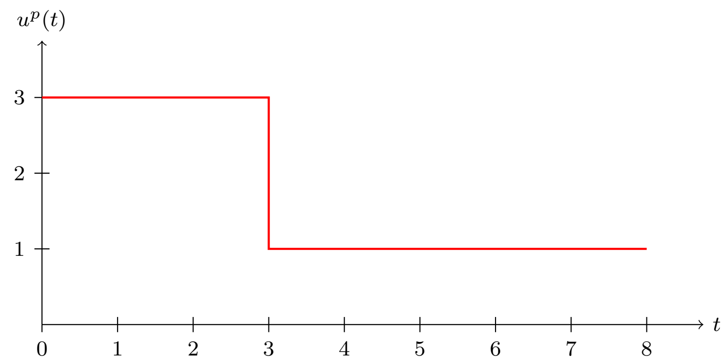

FIGURE 2. A piecewise-constant control with one switch at $t = 3$.

Note that τ_1 and τ_2 coincide at $t = 3$. The corresponding piecewise-constant control defined by (7) is shown in Figure 2. The total variation of this control is obviously equal to 2. However,

$$\sum_{k=1}^{p-1} \left|\sigma^{k+1} - \sigma^k\right| = \left|\sigma^2 - \sigma^1\right| + \left|\sigma^3 - \sigma^2\right| = |0 - 3| + |1 - 0| = 4 > 2.$$

Thus, in this case, inequality (8) in Theorem 3.1 is strict.

Now, let $\mathcal{U}^p$ denote the class of all piecewise-constant functions defined by (7) with switching times satisfying (6). Then clearly $\mathcal{U}^p \subset \mathcal{U}$.

Substituting (7) into the dynamic system (1)-(2) yields

$$\dot{\boldsymbol{x}}(t) = \sum_{k=1}^{p-1} \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{\sigma}^k)\chi_{[\tau_{k-1},\tau_k)}(t) + \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{\sigma}^p)\chi_{[\tau_{p-1},\tau_p]}(t), \quad t \in [0, T], \tag{12}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}^0. \tag{13}$$

Let

$$\boldsymbol{\tau} = [\tau_1, \ldots, \tau_{p-1}]^\top \in \mathbb{R}^{p-1}$$

and

$$\boldsymbol{\sigma} = \left[(\boldsymbol{\sigma}^1)^\top, \ldots, (\boldsymbol{\sigma}^p)^\top\right]^\top \in \mathbb{R}^{pr}.$$

Furthermore, let $\boldsymbol{x}^p(\cdot|\boldsymbol{\tau}, \boldsymbol{\sigma})$ denote the solution of (12)-(13) corresponding to the switching time vector $\boldsymbol{\tau} \in \mathbb{R}^{p-1}$ and the control value vector $\boldsymbol{\sigma} \in \mathbb{R}^{pr}$. Then clearly,

$$\boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}) = \boldsymbol{x}(t|\boldsymbol{u}^p), \quad t \in [0, T].$$

Substituting (7) into the continuous inequality constraints (4) yields

$$\sum_{k=1}^{p-1} h_j(t, \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}^k)\chi_{[\tau_{k-1},\tau_k)}(t) + h_j(t, \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}^p)\chi_{[\tau_{p-1},\tau_p]}(t) \geq 0, \quad t \in [0, T], \quad j = 1, \ldots, q. \tag{14}$$

Let Γ^p denote the set of all pairs $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \mathbb{R}^{p-1} \times \mathbb{R}^{pr}$ satisfying (6) and (14). Furthermore, let $\mathcal{F}^p$ denote the set of all $\boldsymbol{u}^p$ defined by (7) corresponding to pairs in Γ^p. Then

$$(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma^p \quad \Longleftrightarrow \quad \boldsymbol{u}^p \in \mathcal{F}^p.$$

Note that $\mathcal{F}^p \subset \mathcal{F}$.

Now, let $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma^p$ be a given pair, and let $\boldsymbol{u}^p$ be the corresponding piecewise-constant control defined by (7). Then

$$\begin{aligned} J(\boldsymbol{u}^p) &= \Phi(\boldsymbol{x}(T|\boldsymbol{u}^p)) + \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}^p), \boldsymbol{u}^p(t))dt + \gamma \bigvee_0^T \boldsymbol{u}^p \\ &= \Phi(\boldsymbol{x}^p(T|\boldsymbol{\tau}, \boldsymbol{\sigma})) + \sum_{k=1}^{p} \int_{\tau_{k-1}}^{\tau_k} \mathcal{L}(t, \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}^k)dt + \gamma \bigvee_0^T \boldsymbol{u}^p. \end{aligned}$$

By using Theorem 3.1, we obtain an upper bound for $J(\boldsymbol{u}^p)$ in terms of $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}$:

$$J(\boldsymbol{u}^p) \leq \Phi(\boldsymbol{x}^p(T|\boldsymbol{\tau}, \boldsymbol{\sigma})) + \sum_{k=1}^{p} \int_{\tau_{k-1}}^{\tau_k} \mathcal{L}(t, \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}^k)dt + \gamma \sum_{i=1}^{r} \sum_{k=1}^{p-1} \left|\sigma_i^{k+1} - \sigma_i^k\right|.$$

We will show later that if $\boldsymbol{u}^p$ is an optimal piecewise-constant control (i.e. a minimizer of J over $\mathcal{F}^p$), then this upper bound is tight. This suggests that Problem P can be approximated by the following finite-dimensional optimization problem.

Problem $\mathbf{P}_p$. *Choose a pair $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma^p$ to minimize the cost function*

$$\begin{aligned} J^p(\boldsymbol{\tau}, \boldsymbol{\sigma}) := \Phi(\boldsymbol{x}^p(T|\boldsymbol{\tau}, \boldsymbol{\sigma})) &+ \sum_{k=1}^{p} \int_{\tau_{k-1}}^{\tau_k} \mathcal{L}(t, \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \boldsymbol{\sigma}^k)dt \\ &+ \gamma \sum_{i=1}^{r} \sum_{k=1}^{p-1} \left|\sigma_i^{k+1} - \sigma_i^k\right|. \end{aligned} \tag{15}$$

Let $(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*) \in \Gamma^p$ be a solution of Problem P_p, where

$$\boldsymbol{\tau}^* = [\tau_1^*, \ldots, \tau_{p-1}^*]^\top$$

and

$$\boldsymbol{\sigma}^* = \left[(\boldsymbol{\sigma}^{1,*})^\top, \ldots, (\boldsymbol{\sigma}^{p,*})^\top\right]^\top.$$

Then the corresponding piecewise-constant control in $\mathcal{F}^p$ is defined as follows:

$$\boldsymbol{u}^{p,*}(t) = \sum_{k=1}^{p-1} \boldsymbol{\sigma}^{k,*} \chi_{[\tau_{k-1}^*, \tau_k^*)}(t) + \boldsymbol{\sigma}^{p,*} \chi_{[\tau_{p-1}^*, \tau_p^*]}(t), \tag{16}$$

where $\tau_0^* = 0$ and $\tau_p^* = T$. We now show that $\boldsymbol{u}^{p,*}$ is an optimal piecewise-constant control for Problem P. In other words, $\boldsymbol{u}^{p,*}$ minimizes J over $\mathcal{F}^p$.

Theorem 3.2. *Let $(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*) \in \Gamma^p$ be a solution of Problem P_p, and let $\boldsymbol{u}^{p,*} \in \mathcal{F}^p$ denote the corresponding piecewise-constant control defined by equation (16). Then $\boldsymbol{u}^{p,*}$ is a minimizer of the cost functional J over $\mathcal{F}^p$.*

Proof. The proof is by contradiction. Suppose that $\boldsymbol{u}^{p,*}$ does not minimize J over $\mathcal{F}^p$. Then there exists another piecewise-constant control $\boldsymbol{u}^p \in \mathcal{F}^p$ such that

$$J(\boldsymbol{u}^p) < J(\boldsymbol{u}^{p,*}) \leq J^p(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*). \tag{17}$$

Let $(\boldsymbol{\tau}, \boldsymbol{\sigma}) \in \Gamma^p$ denote the pair generating $\boldsymbol{u}^p$ through equation (7). Furthermore, let m denote the number of discontinuities of $\boldsymbol{u}^p$ on the open interval $(0, T)$, where $m = 0$ if $\boldsymbol{u}^p$ is continuous on $(0, T)$. Note that $m \leq p - 1$.

Define a set of points $\{\nu_j\}_{j=0}^{m+1} \subset [0, T]$ as follows:

(1) $\nu_0 = 0$.
(2) ν_j (for $j = 1, \ldots, m$) is the jth discontinuity of $\boldsymbol{u}^p$ on the open interval $(0, T)$.
(3) $\nu_{m+1} = T$.

Then clearly the points in $\{\nu_j\}_{j=0}^{m+1}$ are increasing.

For each $j = 1, \ldots, m+1$, there exists an integer k_j such that

$$\boldsymbol{u}^p(t) = \boldsymbol{\sigma}^{k_j}, \quad t \in [\nu_{j-1}, \nu_j),$$

where $\nu_j = \tau_{k_j}$ and $\nu_{j-1} = \tau_{k_j - 1}$. Define

$$\bar{\boldsymbol{\tau}} = [\bar{\tau}_1, \ldots, \bar{\tau}_{p-1}]^\top \in \mathbb{R}^{p-1},$$

where

$$\bar{\tau}_j = \begin{cases} \nu_j, & \text{if } j = 1, \ldots, m, \\ T, & \text{if } j = m+1, \ldots, p-1. \end{cases}$$

Furthermore, define

$$\bar{\boldsymbol{\sigma}} = \left[(\bar{\boldsymbol{\sigma}}^1)^\top, \ldots, (\bar{\boldsymbol{\sigma}}^p)^\top \right] \in \mathbb{R}^{pr},$$

where

$$\bar{\boldsymbol{\sigma}}^j = \begin{cases} \boldsymbol{\sigma}^{k_j}, & \text{if } j = 1, \ldots, m+1, \\ \boldsymbol{\sigma}^p, & \text{if } j = m+2, \ldots, p. \end{cases}$$

Let $\bar{\boldsymbol{u}}^p$ denote the piecewise-constant control in $\mathcal{U}^p$ corresponding to $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}})$. Then clearly,

$$\bar{\boldsymbol{u}}^p(t) = \boldsymbol{u}^p(t), \quad t \in [0, T],$$

and

$$\boldsymbol{x}^p(t|\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}}) = \boldsymbol{x}^p(t|\boldsymbol{\tau}, \boldsymbol{\sigma}), \quad t \in [0, T].$$

Thus, $\bar{\boldsymbol{u}}^p \in \mathcal{F}^p$ and $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}}) \in \Gamma^p$. Moreover, by virtue of (17),

$$J(\bar{\boldsymbol{u}}^p) = J(\boldsymbol{u}^p) < J(\boldsymbol{u}^{p,*}) \leq J^p(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*). \tag{18}$$

Now, $\{\nu_j\}_{j=0}^{m+1}$ is a valid partition of $[0, T]$ satisfying

$$0 = \nu_0 < \nu_1 < \cdots < \nu_m < \nu_{m+1} = T.$$

Thus, for each $i = 1, \ldots, r$,

$$\bigvee_0^T \bar{u}_i^p \geq \sum_{j=1}^{m+1} \left| \bar{u}_i^p(\nu_j) - \bar{u}_i^p(\nu_{j-1}) \right| = \sum_{k=1}^{p-1} \left| \bar{\sigma}_i^{k+1} - \bar{\sigma}_i^k \right|,$$

where $\bar{u}_i^p$ is the ith component of $\bar{\boldsymbol{u}}^p$. Therefore,

$$\bigvee_0^T \bar{\boldsymbol{u}}^p = \sum_{i=1}^r \bigvee_0^T \bar{u}_i^p \geq \sum_{i=1}^r \sum_{k=1}^{p-1} \left| \bar{\sigma}_i^{k+1} - \bar{\sigma}_i^k \right|. \tag{19}$$

But Theorem 3.1 implies

$$\bigvee_0^T \bar{\boldsymbol{u}}^p \leq \sum_{i=1}^r \sum_{k=1}^{p-1} \left| \bar{\sigma}_i^{k+1} - \bar{\sigma}_i^k \right|. \tag{20}$$

Combining inequalities (19) and (20) yields

$$\bigvee_0^T \bar{\boldsymbol{u}}^p = \sum_{i=1}^r \sum_{k=1}^{p-1} \left| \bar{\sigma}_i^{k+1} - \bar{\sigma}_i^k \right|.$$

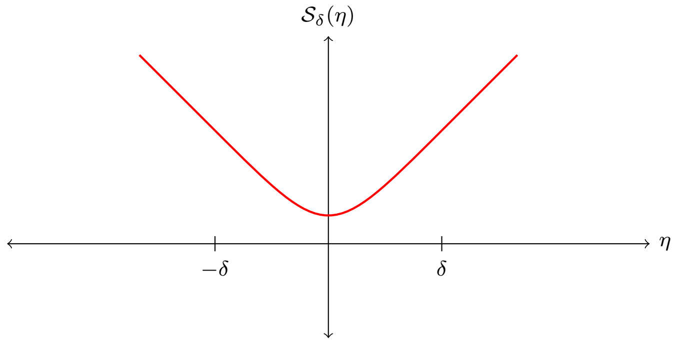

FIGURE 3. The smoothing function $\mathcal{S}_\delta$.

That is, inequality (8) in Theorem 3.1 holds with equality for $\bar{\boldsymbol{u}}^p$. Thus,

$$\begin{aligned} J(\bar{\boldsymbol{u}}^p) &= \Phi(\boldsymbol{x}(T|\bar{\boldsymbol{u}}^p)) + \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\bar{\boldsymbol{u}}^p), \bar{\boldsymbol{u}}^p(t))dt + \gamma \bigvee_0^T \bar{\boldsymbol{u}}^p \\ &= \Phi(\boldsymbol{x}^p(T|\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}})) + \sum_{k=1}^{p} \int_{\bar{\tau}_{k-1}}^{\bar{\tau}_k} \mathcal{L}(t, \boldsymbol{x}^p(t|\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}}), \bar{\boldsymbol{\sigma}}^k)dt + \gamma \sum_{i=1}^{r} \sum_{k=1}^{p-1} \left|\bar{\sigma}_i^{k+1} - \bar{\sigma}_i^k\right| \\ &= J^p(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}}). \end{aligned}$$

Combining this equation with (18) gives

$$J^p(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\sigma}}) = J(\bar{\boldsymbol{u}}^p) < J(\boldsymbol{u}^{p,*}) \leq J^p(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*).$$

But this contradicts the optimality of $(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*)$. Hence, the piecewise-constant control $\boldsymbol{u}^{p,*}$, which is generated by $(\boldsymbol{\tau}^*, \boldsymbol{\sigma}^*)$ through equation (16), must minimize J over $\mathcal{F}^p$. This completes the proof. □

Theorem 3.2 shows that a suboptimal control for Problem P can be generated by solving Problem P_p. Note that Problem P_p is a nonlinear optimization problem in which $\boldsymbol{\tau} \in \mathbb{R}^{p-1}$ and $\boldsymbol{\sigma} \in \mathbb{R}^{pr}$ need to be chosen to minimize the objective function (15) subject to the continuous inequality constraints (14). These constraints must be satisfied at *every* point in $[0, T]$ (an uncountable number of points). Hence, Problem P_p can be viewed as a semi-infinite optimization problem.

An algorithm for solving such problems is discussed in [17, 18]. This algorithm works by approximating the non-smooth absolute value term in (15) as follows:

$$\sum_{i=1}^{r} \sum_{k=1}^{p-1} \left|\sigma_i^{k+1} - \sigma_i^k\right| \approx \sum_{i=1}^{r} \sum_{k=1}^{p-1} \mathcal{S}_\delta(\sigma_i^{k+1} - \sigma_i^k), \tag{21}$$

where $\delta > 0$ is a fixed parameter and

$$\mathcal{S}_\delta(\eta) := \begin{cases} |\eta|, & \text{if } |\eta| > \delta, \\ (\eta^2 + \delta^2)/2\delta, & \text{if } |\eta| \leq \delta. \end{cases}$$

Note that $\mathcal{S}_\delta : \mathbb{R} \to \mathbb{R}$ is a smooth approximation of the absolute value function. This smoothing function is illustrated in Figure 3.

Substituting (21) into the objective function (15) gives

$$\begin{aligned} J^p(\boldsymbol{\tau},\boldsymbol{\sigma}) \approx \Phi(\boldsymbol{x}^p(T|\boldsymbol{\tau},\boldsymbol{\sigma})) + \sum_{k=1}^{p}\int_{\tau_{k-1}}^{\tau_k} \mathcal{L}(t,\boldsymbol{x}^p(t|\boldsymbol{\tau},\boldsymbol{\sigma}),\boldsymbol{\sigma}^k)dt \\ + \gamma\sum_{i=1}^{r}\sum_{k=1}^{p-1}\mathcal{S}_\delta(\sigma_i^{k+1}-\sigma_i^k). \end{aligned} \tag{22}$$

The continuous inequality constraints (14) can be handled using the constraint transcription method discussed in [6, 19]. This method involves transforming (14) into the following set of equivalent equality constraints:

$$\int_0^T \min\{g_j(t|\boldsymbol{\tau},\boldsymbol{\sigma}),0\}dt = 0, \quad j=1,\dots,q, \tag{23}$$

where

$$g_j(t|\boldsymbol{\tau},\boldsymbol{\sigma}) := \sum_{k=1}^{p-1} h_j(t,\boldsymbol{x}^p(t|\boldsymbol{\tau},\boldsymbol{\sigma}),\boldsymbol{\sigma}^k)\chi_{[\tau_{k-1},\tau_k)}(t) + h_j(t,\boldsymbol{x}^p(t|\boldsymbol{\tau},\boldsymbol{\sigma}),\boldsymbol{\sigma}^p)\chi_{[\tau_{p-1},\tau_p]}(t).$$

There are only a finite number of constraints in (23), and thus at first glance (23) appears much easier to work with than the continuous inequality constraints (14). Unfortunately, the equality constraints in (23) are non-smooth, and thus standard numerical optimization algorithms will likely struggle with these constraints. In the constraint transcription method, we approximate (23) by the following set of smooth inequality constraints:

$$\rho + \int_0^T \varphi_\epsilon(g_j(t|\boldsymbol{\tau},\boldsymbol{\sigma}))dt \geq 0, \quad j=1,\dots,q, \tag{24}$$

where $\epsilon > 0$ and $\rho > 0$ are fixed parameters and $\varphi_\epsilon : \mathbb{R} \to \mathbb{R}$ is defined by

$$\varphi_\epsilon(\eta) := \begin{cases} \eta, & \text{if } \eta < -\epsilon, \\ -(\eta-\epsilon)^2/4\epsilon, & \text{if } -\epsilon \leq \eta \leq \epsilon, \\ 0, & \text{if } \eta > \epsilon. \end{cases}$$

Note that φ_ϵ is a smooth approximation of the function $\min\{\cdot,0\}$; see Figure 4.

Problem P_p can now be approximated as follows: Choose $\boldsymbol{\tau} \in \mathbb{R}^{p-1}$ and $\boldsymbol{\sigma} \in \mathbb{R}^{pr}$ to minimize (22) subject to (6) and (24). This approximate problem contains only a finite number of constraints. Therefore, it can be solved using standard nonlinear programming techniques (see [10, 13, 14, 17]). In Chapter 10 of [17], it is shown that by updating the parameters δ, ϵ, and ρ according to certain rules, the solution of the approximate problem can be made to converge to a solution of Problem P_p.

We refer the reader to [17, 18] for more details on the computational aspects of solving Problem P_p. Our focus in this paper is on the theoretical convergence properties of the sequence of suboptimal controls generated by solving Problem P_p for increasing values of p. Specifically, we will show that the cost of the suboptimal control converges to the optimal cost of Problem P as p approaches infinity. The original proof of this result in [17] is only applicable to problems with pure state constraints, not the mixed state-control constraints considered in this paper.

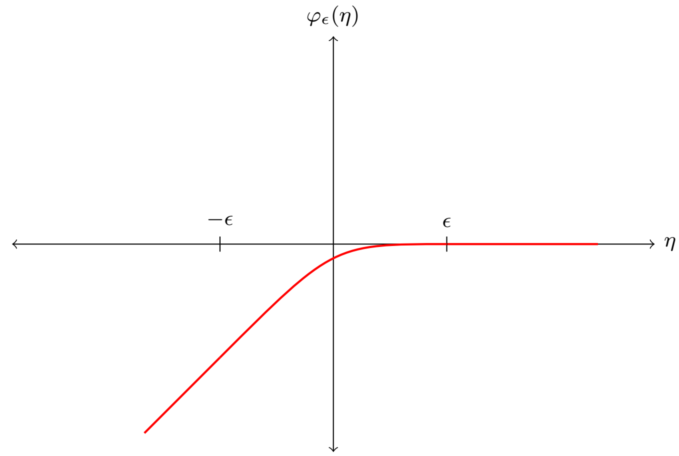

FIGURE 4. The smoothing function φ_ϵ.

4. **Preliminary Results.** The purpose of this section is to establish a series of preliminary results that will be needed later in Section 5.

Lemma 4.1. *Let $\varphi : [a,b] \to \mathbb{R}$ be a function of bounded variation. Furthermore, let $c \in [a,b]$ and $\eta \in \mathbb{R}$. Define a new function $\phi : [a,b] \to \mathbb{R}$ as follows:*

$$\phi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a,b] \setminus \{c\}, \\ \eta, & \text{if } t = c. \end{cases}$$

Then ϕ is also of bounded variation.

Proof. Since φ is of bounded variation, there exists a real number $M > 0$ such that

$$|\varphi(t)| \leq M, \quad t \in [a,b].$$

Let $\{t_j\}_{j=0}^m$ be an arbitrary partition of $[a,b]$ satisfying

$$a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b.$$

If $c \neq t_j$ for each $j = 0, \ldots, m$, then

$$\sum_{j=1}^m \left|\phi(t_j) - \phi(t_{j-1})\right| = \sum_{j=1}^m \left|\varphi(t_j) - \varphi(t_{j-1})\right| \leq \bigvee_0^T \varphi. \tag{25}$$

On the other hand, suppose that the point c coincides with one of the partition points. Then $c = t_l$ for some $l \in \{0, \ldots, m\}$. If $l \in \{1, \ldots, m-1\}$, then

$$\begin{aligned} |\phi(t_l) - \phi(t_{l-1})| &= |\eta - \varphi(t_{l-1})| \\ &\leq |\eta - \varphi(t_l)| + |\varphi(t_l) - \varphi(t_{l-1})| \\ &\leq |\varphi(t_l) - \varphi(t_{l-1})| + |\varphi(t_l)| + |\eta| \\ &\leq |\varphi(t_l) - \varphi(t_{l-1})| + M + |\eta|. \end{aligned} \tag{26}$$

Similarly,

$$|\phi(t_{l+1}) - \phi(t_l)| \leq |\varphi(t_{l+1}) - \varphi(t_l)| + M + |\eta|. \tag{27}$$

Using (26) and (27), we obtain

$$
\begin{aligned}
\sum_{j=1}^{m} \big|\phi(t_j) - \phi(t_{j-1})\big| &= \sum_{j=1}^{l-1} \big|\varphi(t_j) - \varphi(t_{j-1})\big| + \big|\phi(t_l) - \phi(t_{l-1})\big| + \big|\phi(t_{l+1}) - \phi(t_l)\big| \\
&\qquad\qquad + \sum_{j=l+2}^{m} \big|\varphi(t_l) - \varphi(t_{l-1})\big| \\
&\le \sum_{j=1}^{m} |\varphi(t_j) - \varphi(t_{j-1})| + 2M + 2|\eta| \\
&\le \bigvee_0^T \varphi + 2M + 2|\eta|. \qquad (28)
\end{aligned}
$$

This inequality is based on the assumption that $l \in \{1, \ldots, m-1\}$. If $l = 0$ or $l = m$, then similar arguments show that

$$\sum_{j=1}^{m} \big|\phi(t_j) - \phi(t_{j-1})\big| \le \bigvee_0^T \varphi + M + |\eta|. \qquad (29)$$

Recall that the choice of partition $\{t_j\}_{j=0}^m$ was arbitrary. Hence, in view of (25), (28), and (29), we have

$$\bigvee_0^T \phi \le \bigvee_0^T \varphi + 2M + 2|\eta|.$$

This shows that ϕ is of bounded variation, as required. □

Jordan's theorem states that a function of bounded variation can be written as the difference of two non-decreasing functions [7, 15]. Thus, since a non-decreasing function defined on $[a, b]$ has a left limit at every point in (a, b) (see [16]), a function of bounded variation defined on $[a, b]$ also has a left limit at every point in $(a, b]$. With this in mind, we present the following lemma.

Lemma 4.2. *Let $\varphi : [a, b] \to \mathbb{R}$ be a function of bounded variation. Furthermore, define a new function $\phi : [a, b] \to \mathbb{R}$ as follows:*

$$\phi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a, b), \\ \varphi(b^-), & \text{if } t = b, \end{cases}$$

where

$$\varphi(b^-) = \lim_{t \to b-} \varphi(t).$$

Then ϕ is of bounded variation and

$$\bigvee_a^b \phi = \bigvee_a^b \varphi - \big|\varphi(b) - \varphi(b^-)\big|. \qquad (30)$$

Proof. It follows from Lemma 4.1 that ϕ is of bounded variation. To prove (30), let $\{t_j\}_{j=0}^m$ be an arbitrary partition of $[a, b]$ such that

$$a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b.$$

Then

$$\begin{aligned}\sum_{j=1}^{m}\big|\varphi(t_j)-\varphi(t_{j-1})\big| &= \sum_{j=1}^{m-1}\big|\phi(t_j)-\phi(t_{j-1})\big| + \big|\varphi(b)-\phi(t_{m-1})\big| \\ &\leq \sum_{j=1}^{m-1}\big|\phi(t_j)-\phi(t_{j-1})\big| + \big|\varphi(b)-\varphi(b^-)\big| + \big|\varphi(b^-)-\phi(t_{m-1})\big| \\ &= \sum_{j=1}^{m}\big|\phi(t_j)-\phi(t_{j-1})\big| + \big|\varphi(b)-\varphi(b^-)\big| \\ &\leq \bigvee_a^b \phi + \big|\varphi(b)-\varphi(b^-)\big|.\end{aligned}$$

Thus, since the partition $\{t_j\}_{j=0}^m$ was chosen arbitrarily,

$$\bigvee_a^b \phi \geq \bigvee_a^b \varphi - \big|\varphi(b)-\varphi(b^-)\big|.$$

Suppose that this inequality is strict:

$$\bigvee_a^b \phi > \bigvee_a^b \varphi - \big|\varphi(b)-\varphi(b^-)\big|. \tag{31}$$

Then there exists a real number $\epsilon > 0$ such that

$$\bigvee_a^b \phi - \epsilon > \bigvee_a^b \varphi - \big|\varphi(b)-\varphi(b^-)\big|. \tag{32}$$

Since $\varphi(b^-)$ is the limit of φ as $t \to b-$, there exists a real number $\delta > 0$ such that

$$\big|\varphi(t)-\varphi(b^-)\big| < \tfrac{1}{4}\epsilon, \qquad t \in (b-\delta, b). \tag{33}$$

Let $\{t'_j\}_{j=0}^m$ be a partition of $[a,b]$ such that

$$\sum_{j=1}^{m}\big|\phi(t'_j)-\phi(t'_{j-1})\big| > \bigvee_a^b \phi - \tfrac{1}{4}\epsilon, \tag{34}$$

where

$$a = t'_0 < t'_1 < \cdots < t'_{m-1} < t'_m = b.$$

Choose a point $t^* \in (b-\delta, b)$ such that $t^* > t'_{m-1}$. Then we can define a new partition $\{t''_j\}_{j=0}^{m+1}$ as follows:

$$t''_j := \begin{cases} t'_j, & \text{if } j = 0,\dots,m-1, \\ t^*, & \text{if } j = m, \\ b, & \text{if } j = m+1. \end{cases}$$

Using (34) and the triangle inequality, we obtain

$$\begin{aligned}\sum_{j=1}^{m+1} \big|\phi(t_j'') - \phi(t_{j-1}'')\big| &= \sum_{j=1}^{m-1} \big|\phi(t_j'') - \phi(t_{j-1}'')\big| + \big|\phi(t_m'') - \phi(t_{m-1}'')\big| \\ &\qquad + \big|\phi(t_{m+1}'') - \phi(t_m'')\big| \\ &\geq \sum_{j=1}^{m-1} \big|\phi(t_j'') - \phi(t_{j-1}'')\big| + \big|\phi(t_{m+1}'') - \phi(t_{m-1}'')\big| \\ &= \sum_{j=1}^{m-1} \big|\phi(t_j') - \phi(t_{j-1}')\big| + \big|\phi(b) - \phi(t_{m-1}')\big| \\ &= \sum_{j=1}^{m} \big|\phi(t_j') - \phi(t_{j-1}')\big| \\ &> \bigvee_a^b \phi - \tfrac{1}{4}\epsilon. \end{aligned} \tag{35}$$

Now, recall from (32) that

$$\bigvee_a^b \varphi + \epsilon < \bigvee_a^b \phi + \big|\varphi(b) - \varphi(b^-)\big|.$$

Thus, using (35),

$$\begin{aligned}\bigvee_a^b \varphi + \epsilon &< \sum_{j=1}^{m+1} \big|\phi(t_j'') - \phi(t_{j-1}'')\big| + \tfrac{1}{4}\epsilon + \big|\varphi(b) - \varphi(b^-)\big| \\ &= \sum_{j=1}^{m} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \big|\phi(t_{m+1}'') - \phi(t_m'')\big| + \tfrac{1}{4}\epsilon + \big|\varphi(b) - \varphi(b^-)\big| \\ &= \sum_{j=1}^{m} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \big|\varphi(b^-) - \varphi(t^*)\big| + \tfrac{1}{4}\epsilon + \big|\varphi(b) - \varphi(b^-)\big|. \end{aligned} \tag{36}$$

Using (33) gives

$$\begin{aligned}\bigvee_a^b \varphi + \epsilon &< \sum_{j=1}^{m} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \tfrac{1}{2}\epsilon + \big|\varphi(b) - \varphi(b^-)\big| \\ &\leq \sum_{j=1}^{m} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \tfrac{1}{2}\epsilon + \big|\varphi(b) - \varphi(t^*)\big| + \big|\varphi(t^*) - \varphi(b^-)\big| \\ &< \sum_{j=1}^{m} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \big|\varphi(t_{m+1}'') - \varphi(t_m'')\big| + \tfrac{3}{4}\epsilon \\ &= \sum_{j=1}^{m+1} \big|\varphi(t_j'') - \varphi(t_{j-1}'')\big| + \tfrac{3}{4}\epsilon \\ &\leq \bigvee_a^b \varphi + \tfrac{3}{4}\epsilon. \end{aligned}$$

Hence,

$$\bigvee_a^b \varphi + \tfrac{1}{4}\epsilon < \bigvee_a^b \varphi.$$

But this is an obvious contradiction because $\epsilon > 0$. Hence, our initial assumption that (31) holds is false. This completes the proof. $\square$

The next lemma is similar to Lemma 4.2. For simplicity, we omit the proof.

Lemma 4.3. *Let $\varphi : [a,b] \to \mathbb{R}$ be a function of bounded variation. Furthermore, define a new function $\phi : [a,b] \to \mathbb{R}$ as follows:*

$$\phi(t) := \begin{cases} \varphi(t), & \text{if } t \in (a,b], \\ \varphi(a^+), & \text{if } t = a, \end{cases}$$

where

$$\varphi(a^+) = \lim_{t \to a+} \varphi(t).$$

Then ϕ is of bounded variation and

$$\bigvee_a^b \phi = \bigvee_a^b \varphi - \big|\varphi(a) - \varphi(a^+)\big|.$$

Lemmas 4.1-4.3 are needed to prove the next two results.

Lemma 4.4. *Let $\varphi : [a,b] \to \mathbb{R}$ be a function of bounded variation. Furthermore, let $\eta \in \mathbb{R}$ be a given constant. Define a new function $\phi : [a,b] \to \mathbb{R}$ as follows:*

$$\phi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a,b), \\ \eta, & \text{if } t = b. \end{cases}$$

Then ϕ is of bounded variation and

$$\bigvee_a^b \phi \le \bigvee_a^b \varphi + \big|\varphi(b) - \eta\big|. \tag{37}$$

Proof. Lemma 4.1 ensures that ϕ is of bounded variation. It remains to prove inequality (37).

Define

$$\psi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a,b), \\ \varphi(b^-), & \text{if } t = b. \end{cases}$$

Then it follows from Lemma 4.2 that ψ is of bounded variation and

$$\bigvee_a^b \psi = \bigvee_a^b \varphi - \big|\varphi(b) - \varphi(b^-)\big|. \tag{38}$$

Furthermore, since $\psi(t) = \phi(t) = \varphi(t)$ for all $t \neq b$, we have

$$\psi(b^-) = \phi(b^-) = \varphi(b^-).$$

Hence,

$$\psi(t) = \begin{cases} \phi(t), & \text{if } t \in [a,b), \\ \phi(b^-), & \text{if } t = b. \end{cases}$$

Thus, invoking Lemma 4.2 once again gives

$$\bigvee_a^b \psi = \bigvee_a^b \phi - \left|\phi(b) - \phi(b^-)\right| = \bigvee_a^b \phi - \left|\eta - \varphi(b^-)\right|. \tag{39}$$

Combining equations (38) and (39) and applying the reverse triangle inequality yields

$$\bigvee_a^b \phi = \bigvee_a^b \varphi + \left|\eta - \varphi(b^-)\right| - \left|\varphi(b) - \varphi(b^-)\right| \le \bigvee_a^b \varphi + \left|\varphi(b) - \eta\right|.$$

This completes the proof. □

Lemma 4.5. *Let $\varphi : [a, b] \to \mathbb{R}$ be a function of bounded variation. Furthermore, let $c \in [a, b)$ be a given constant. Define a new function $\phi : [a, b] \to \mathbb{R}$ as follows:*

$$\phi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a, b] \setminus \{c\}, \\ \varphi(c^+), & \text{if } t = c. \end{cases}$$

Then ϕ is of bounded variation and

$$\bigvee_a^b \phi \le \bigvee_a^b \varphi.$$

Proof. It follows from Lemma 4.1 that ϕ is of bounded variation. If $c = a$, then Lemma 4.3 implies that

$$\bigvee_a^b \phi = \bigvee_a^b \varphi - \left|\varphi(a) - \varphi(a^+)\right| \le \bigvee_a^b \varphi,$$

which completes the proof. Thus, we may assume that $c \in (a, b)$. Since ϕ and φ are of bounded variation on $[a, b]$, they are also of bounded variation on $[a, c]$ and $[c, b]$. Furthermore, by Lemma 4.4, the total variation of ϕ on $[a, c]$ is

$$\bigvee_a^c \phi \le \bigvee_a^c \varphi + \left|\varphi(c) - \varphi(c^+)\right|. \tag{40}$$

Meanwhile, it follows from Lemma 4.3 that the total variation of ϕ on $[c, b]$ is

$$\bigvee_c^b \phi = \bigvee_c^b \varphi - \left|\varphi(c) - \varphi(c^+)\right|. \tag{41}$$

Combining (40) and (41) gives

$$\bigvee_a^b \phi = \bigvee_a^c \phi + \bigvee_c^b \phi \le \bigvee_a^c \varphi + \bigvee_c^b \varphi = \bigvee_a^b \varphi.$$

This completes the proof. □

We are now ready to prove the three remaining lemmas in this section, each of which will be instrumental in proving the main convergence results in Section 5. The proofs of these remaining lemmas rely on Lemmas 4.1-4.5 above.

Lemma 4.6. *Let $\{\varphi_l\}_{l=1}^{\infty}$ be a sequence of functions defined on $[a,b]$, and suppose that there exists a constant $M > 0$ such that for all integers $l \geq 1$,*

$$\bigvee_a^b \varphi_l \leq M. \tag{42}$$

If $\{\varphi_l\}_{l=1}^{\infty}$ converges to a function $\varphi : [a,b] \to \mathbb{R}$ pointwise on $[a,b]$, then φ is of bounded variation with

$$\bigvee_a^b \varphi \leq \liminf_{l\to\infty} \bigvee_a^b \varphi_l.$$

Proof. Let $\{t_j\}_{j=0}^{m}$ be an arbitrary partition of $[a,b]$ satisfying

$$a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b.$$

Let $\kappa \geq 1$ be a fixed integer. Then since $\varphi_l \to \varphi$ pointwise on $[a,b]$ as $l \to \infty$, there exists an $l_1 > 0$ such that for each integer $l \geq l_1$,

$$\left|\varphi_l(t_j) - \varphi(t_j)\right| < \frac{1}{2m\kappa}, \quad j = 0, \ldots, m.$$

Thus, for each $l \geq l_1$,

$$\begin{aligned}\sum_{j=1}^{m} \left|\varphi(t_j) - \varphi(t_{j-1})\right| &\leq \sum_{j=1}^{m} \Big\{ \left|\varphi(t_j) - \varphi_l(t_j)\right| + \left|\varphi_l(t_j) - \varphi_l(t_{j-1})\right| \\ &\qquad\qquad + \left|\varphi_l(t_{j-1}) - \varphi(t_{j-1})\right| \Big\} \\ &< \frac{1}{\kappa} + \sum_{j=1}^{m} \left|\varphi_l(t_j) - \varphi_l(t_{j-1})\right| \\ &\leq \frac{1}{\kappa} + \bigvee_a^b \varphi_l.\end{aligned}$$

Hence,

$$\sum_{j=1}^{m} \left|\varphi(t_j) - \varphi(t_{j-1})\right| \leq \frac{1}{\kappa} + \inf_{l \geq l_1} \bigvee_a^b \varphi_l$$

and

$$\sum_{j=1}^{m} \left|\varphi(t_j) - \varphi(t_{j-1})\right| \leq \frac{1}{\kappa} + \liminf_{l\to\infty} \bigvee_a^b \varphi_l.$$

By (42), the right-hand side is finite and independent of the choice of partition. Thus,

$$\bigvee_a^b \varphi \leq \frac{1}{\kappa} + \liminf_{l\to\infty} \bigvee_a^b \varphi_l. \tag{43}$$

This shows that φ is of bounded variation.

Now, recall that κ was chosen arbitrarily. Taking $\kappa \to \infty$ in (43) gives

$$\bigvee_a^b \varphi \leq \liminf_{l\to\infty} \bigvee_a^b \varphi_l.$$

This completes the proof. $\square$

Our next result shows that any function of bounded variation can be made right-continuous by changing its value at a countable number of points. The new right-continuous function is equal to the original function *almost everywhere*. Furthermore, the new function has smaller total variation.

Lemma 4.7. *Let $\varphi : [a, b] \to \mathbb{R}$ be a function of bounded variation. Then there exists a function $\psi : [a, b] \to \mathbb{R}$, also of bounded variation, such that:*

(a) $\psi(t) = \varphi(t)$ *for almost all* $t \in [a, b]$.
(b) $\psi(t) = \psi(t^+)$ *for all* $t \in [a, b)$.
(c) $\psi(b) = \psi(b^-)$.
(d) $\bigvee_a^b \psi \leq \bigvee_a^b \varphi$.

Proof. Recall from Jordan's theorem that φ, a function of bounded variation, can be expressed as the difference of two monotonic functions [7, 15]. Thus, since monotonic functions have a countable number of discontinuities [15], φ also has a countable number of discontinuities.

Let $\mathcal{T} \subset [a, b)$ denote the set of discontinuities of φ in $[a, b)$. Define $\psi : [a, b] \to \mathbb{R}$ as follows:

$$\psi(t) := \begin{cases} \varphi(t), & \text{if } t \in [a, b) \setminus \mathcal{T}, \\ \varphi(t^+), & \text{if } t \in \mathcal{T}, \\ \varphi(b^-), & \text{if } t = b. \end{cases}$$

Since $\mathcal{T}$ is countable, it is clear that $\psi(t) = \varphi(t)$ for almost all $t \in [a, b]$, which proves part (a).

To prove part (b), let $t \in [a, b)$. Furthermore, let $\epsilon > 0$ be arbitrary but fixed. Then there exists a corresponding $\delta > 0$ such that

$$t < \tau < t + \delta \quad \Longrightarrow \quad \left|\varphi(\tau) - \varphi(t^+)\right| < \tfrac{1}{2}\epsilon.$$

Let $\tau \in (t, t + \delta)$. Note that $\psi(t) = \varphi(t) = \varphi(t^+)$ if $t \notin \mathcal{T}$, and $\psi(t) = \varphi(t^+)$ if $t \in \mathcal{T}$. Thus, if $\tau \notin \mathcal{T}$, then

$$\left|\psi(\tau) - \psi(t)\right| = \left|\varphi(\tau) - \varphi(t^+)\right| < \tfrac{1}{2}\epsilon. \tag{44}$$

On the other hand, if $\tau \in \mathcal{T}$, then there exists a $t' \in (t, t + \delta)$ such that

$$\left|\varphi(\tau^+) - \varphi(t')\right| < \tfrac{1}{2}\epsilon.$$

Thus,

$$\left|\psi(\tau) - \psi(t)\right| = \left|\varphi(\tau^+) - \varphi(t^+)\right| \leq \left|\varphi(\tau^+) - \varphi(t')\right| + \left|\varphi(t') - \varphi(t^+)\right| < \epsilon. \tag{45}$$

It follows from (44) and (45) that

$$t < \tau < t + \delta \quad \Longrightarrow \quad \left|\psi(\tau) - \psi(t)\right| < \epsilon.$$

Since ϵ was chosen arbitrarily, this shows that ψ is continuous from the right, thus proving part (b). Part (c) is proved in a similar manner.

It remains to prove part (d). Let $\mathcal{T} = \{\nu_l\}_{l \in \mathcal{S}}$, where $\mathcal{S} \subset \mathbb{N}$ is a countable index set. For each $l \in \mathcal{S}$, define $\psi_l : [a, b] \to \mathbb{R}$ as follows:

$$\psi_l(t) := \begin{cases} \psi_{l-1}(t), & \text{if } t \in [a, b] \setminus \{\nu_l\}, \\ \psi_{l-1}(t^+), & \text{if } t = \nu_l, \end{cases}$$

where

$$\psi_0(t) := \begin{cases} \varphi(t), & \text{if } t \in [a, b), \\ \varphi(b^-), & \text{if } t = b. \end{cases}$$

It follows from Lemmas 4.2 and 4.5 that for each $l \in \mathcal{S}$, ψ_l is of bounded variation and

$$\bigvee_a^b \psi_l \leq \bigvee_a^b \psi_{l-1}.$$

Therefore,

$$\bigvee_a^b \psi_l \leq \bigvee_a^b \psi_0 \leq \bigvee_a^b \varphi. \tag{46}$$

If $\mathcal{S}$ is a finite set with $|\mathcal{S}| = m$, then

$$\bigvee_a^b \psi = \bigvee_a^b \psi_m \leq \cdots \leq \bigvee_a^b \psi_1 \leq \bigvee_a^b \psi_0 \leq \bigvee_a^b \varphi,$$

as required. Thus, we may assume that $\mathcal{S}$ is an infinite set. We now show that

$$\lim_{l \to \infty} \psi_l(t) = \psi(t), \quad t \in [a, b]. \tag{47}$$

Clearly, for each $t \in [a, b) \setminus \mathcal{T}$, $\psi_l(t) = \varphi(t) = \psi(t)$ and hence (47) holds. Similarly, if $t = b$, then

$$\psi_l(b) = \varphi(b^-) = \psi(b),$$

and hence (47) also holds in this case. Finally, if $t = \nu_l$ for some $l \in \mathcal{S}$, then for each integer $k \geq l$,

$$\psi_k(\nu_l) = \varphi(\nu_l^+) = \psi(\nu_l).$$

Equation (47) then follows immediately. In view of (47), we see that ψ_l converges pointwise to ψ as $l \to \infty$. Moreover, (46) shows that the total variation of ψ_l is uniformly bounded with respect to l. Thus, by Lemma 4.6, ψ is of bounded variation. Furthermore, by (46),

$$\bigvee_a^b \psi \leq \liminf_{l \to \infty} \bigvee_a^b \psi_l \leq \bigvee_a^b \varphi.$$

This completes the proof. □

Our final preliminary result shows that, given a function of bounded variation with certain continuity properties, we can find a sequence of piecewise-constant functions converging to this function *uniformly*. This result is crucial to proving the main convergence result in the next section.

Lemma 4.8. *Let $\varphi : [a, b] \to \mathbb{R}$ be a function of bounded variation, and suppose that $\varphi(t) = \varphi(t^+)$ for each $t \in [a, b)$ and $\varphi(b) = \varphi(b^-)$. Then there exists a sequence of piecewise-constant functions $\{\varphi_l\}_{l=1}^{\infty}$ such that $\varphi_l \to \varphi$ uniformly on $[a, b]$ as $l \to \infty$. Furthermore, for each integer $l \geq 1$,*

$$\bigvee_a^b \varphi_l \leq \bigvee_a^b \varphi. \tag{48}$$

Proof. Recall from Jordan's theorem that any function of bounded variation can be expressed as the difference of two non-decreasing functions [7, 15]. In fact, we can express φ as follows:

$$\varphi(t) = \psi(t) - \phi(t), \quad t \in [a, b],$$

where $\psi : [a, b] \to \mathbb{R}$ and $\phi : [a, b] \to \mathbb{R}$ are defined by

$$\psi(t) = \bigvee_a^t \varphi,$$

$$\phi(t) = \bigvee_a^t \varphi - \varphi(t).$$

Both ψ and ϕ are bounded and non-decreasing (see [7]). In fact,

$$\psi(a) \le \psi(t) \le \psi(b), \quad t \in [a, b],$$

and

$$\phi(a) \le \phi(t) \le \phi(b), \quad t \in [a, b].$$

Hence, the range of ψ is a subset of the interval $[\psi(a), \psi(b)]$, and the range of ϕ is a subset of the interval $[\phi(a), \phi(b)]$.

Moreover, since φ is right-continuous on $[a, b)$ and left-continuous at $t = b$, both ψ and ϕ are also right-continuous on $[a, b)$ and left-continuous at $t = b$ (see [7]).

Now, let $l \ge 1$ be a given integer. For each $k = 0, \ldots, l$, define

$$\alpha_k := \psi(a) + \frac{k}{l}(\psi(b) - \psi(a)),$$

$$\beta_k := \phi(a) + \frac{k}{l}(\phi(b) - \phi(a)).$$

Then the intervals $[\alpha_{k-1}, \alpha_k)$, $k = 1, \ldots, l-1$ and $[\alpha_{l-1}, \alpha_l]$ constitute a partition of $[\psi(a), \psi(b)]$. Similarly, the intervals $[\beta_{k-1}, \beta_k)$, $k = 1, \ldots, l-1$ and $[\beta_{l-1}, \beta_l]$ constitute a partition of $[\phi(a), \phi(b)]$.

Now, for each $k = 0, \ldots, l$, define

$$\tilde{\tau}_k := \begin{cases} \inf\{\, t \in [a, b] : \ \psi(t) \ge \alpha_k \,\}, & \text{if } k = 0, \ldots, l-1, \\ b, & \text{if } k = l. \end{cases}$$

Note that $\psi(b) \ge \alpha_k$ for each $k = 0, \ldots, l$. Thus, each $\tilde{\tau}_k$ is well-defined. Furthermore, since ψ is right-continuous,

$$\psi(\tilde{\tau}_k) \ge \alpha_k, \quad k = 0, \ldots, l,$$

and

$$\psi(t) \le \alpha_k, \quad t \in [a, \tilde{\tau}_k), \quad k = 1, \ldots, l.$$

Clearly, $\tilde{\mathcal{P}} = \{\tilde{\tau}_0, \tilde{\tau}_1, \ldots, \tilde{\tau}_l\}$ is a partition of $[a, b]$ satisfying

$$a = \tilde{\tau}_0 \le \tilde{\tau}_1 \le \cdots \le \tilde{\tau}_l = b.$$

Now, for each $k = 0, \ldots, l$, define

$$\hat{\tau}_k := \begin{cases} \inf\{\, t \in [a, b] : \ \phi(t) \ge \beta_k \,\}, & \text{if } k = 0, \ldots, l-1, \\ b, & \text{if } k = l. \end{cases}$$

Then $\hat{\mathcal{P}} = \{\hat{\tau}_0, \hat{\tau}_1, \ldots, \hat{\tau}_l\}$ is another partition of $[a, b]$ satisfying

$$a = \hat{\tau}_0 \le \hat{\tau}_1 \le \cdots \le \hat{\tau}_l = b.$$

Furthermore, for each $k = 0, \ldots, l$, we have $\phi(\hat{\tau}_k) \ge \beta_k$ and $\phi(t) \le \beta_k$ for all $t < \hat{\tau}_k$.

Consider the union of the two partitions $\tilde{\mathcal{P}}$ and $\hat{\mathcal{P}}$:

$$\mathcal{P} = \tilde{\mathcal{P}} \cup \hat{\mathcal{P}} = \{\tau_0, \ldots, \tau_m\},$$

where $m+1 \leq 2l$ is the number of distinct points in $\tilde{\mathcal{P}}$ and $\hat{\mathcal{P}}$. Without loss of generality, we may assume that

$$a = \tau_0 < \tau_1 < \cdots < \tau_m = b.$$

Thus, for each $k = 0, \ldots, m$, either $\tau_k \in \tilde{\mathcal{P}}$ or $\tau_k \in \hat{\mathcal{P}}$ (or possibly $\tau_k \in \tilde{\mathcal{P}} \cap \hat{\mathcal{P}}$). Now, define a set of constants $\{\omega_k\}_{k=1}^m$ as follows:

$$\omega_k := \varphi(\tau_{k-1}) = \psi(\tau_{k-1}) - \phi(\tau_{k-1}), \quad k = 1, \ldots, m.$$

Furthermore, define a piecewise-constant function $\varphi_l : [a,b] \to \mathbb{R}$ as follows:

$$\varphi_l(t) := \sum_{k=1}^{m-1} \omega_k \chi_{[\tau_{k-1},\tau_k)}(t) + \omega_m \chi_{[\tau_{m-1},\tau_m]}(t), \tag{49}$$

where the indicator functions $\chi_{[\tau_{k-1},\tau_k)} : \mathbb{R} \to \mathbb{R}$ and $\chi_{[\tau_{m-1},\tau_m]} : \mathbb{R} \to \mathbb{R}$ are as defined in Section 3. Clearly, the values of φ_l and φ coincide at the partition points:

$$\varphi_l(\tau_k) = \omega_{k+1} = \varphi(\tau_k), \quad k = 0, \ldots, m-1.$$

Furthermore, notice that the definition of φ_l in equation (49) is consistent with the definition of the piecewise-constant controls in Section 3. Thus, by Theorem 3.1,

$$\bigvee_a^b \varphi_l \leq \sum_{k=1}^{m-1} \left|\omega_{k+1} - \omega_k\right| = \sum_{k=1}^{m-1} \left|\varphi(\tau_k) - \varphi(\tau_{k-1})\right| \leq \bigvee_a^b \varphi.$$

This proves inequality (48).

We now complete the proof by showing that φ_l defined by (49) converges to φ uniformly on $[a,b]$ as $l \to \infty$.

Let $t \in [a,b)$. Then there exists a unique integer $k \in \{1, \ldots, m\}$ such that $t \in [\tau_{k-1}, \tau_k)$. Furthermore, since $\tilde{\mathcal{P}} \subset \mathcal{P}$ and $\hat{\mathcal{P}} \subset \mathcal{P}$, there exists integers k' and k'' such that

$$\tilde{\tau}_{k'-1} \leq \tau_{k-1} \leq t < \tilde{\tau}_{k'}$$

and

$$\hat{\tau}_{k''-1} \leq \tau_{k-1} \leq t < \hat{\tau}_{k''}.$$

Thus, since ψ and ϕ are non-decreasing,

$$\alpha_{k'-1} \leq \psi(\tilde{\tau}_{k'-1}) \leq \psi(\tau_{k-1}) \leq \psi(t) \leq \alpha_{k'} \tag{50}$$

and

$$\beta_{k''-1} \leq \phi(\hat{\tau}_{k''-1}) \leq \phi(\tau_{k-1}) \leq \phi(t) \leq \beta_{k''}. \tag{51}$$

From (50), we obtain

$$-\frac{1}{l}(\psi(b) - \psi(a)) = \alpha_{k'-1} - \alpha_{k'} \leq \psi(t) - \psi(\tau_{k-1}) \leq \alpha_{k'} - \alpha_{k'-1} = \frac{1}{l}(\psi(b) - \psi(a)).$$

Similarly, from (51),

$$-\frac{1}{l}(\phi(b) - \phi(a)) \leq \phi(t) - \phi(\tau_{k-1}) \leq \frac{1}{l}(\phi(b) - \phi(a)).$$

Hence,

$$\left|\psi(t) - \psi(\tau_{k-1})\right| \leq \frac{1}{l}(\psi(b) - \psi(a)), \quad \left|\phi(t) - \phi(\tau_{k-1})\right| \leq \frac{1}{l}(\phi(b) - \phi(a)).$$

It follows that

$$\begin{aligned}\big|\varphi_l(t)-\varphi(t)\big| &= \big|\omega_k-(\psi(t)-\phi(t))\big| \\ &= \big|\psi(\tau_{k-1})-\phi(\tau_{k-1})-(\psi(t)-\phi(t))\big| \\ &\leq \big|\psi(t)-\psi(\tau_{k-1})\big|+\big|\phi(t)-\phi(\tau_{k-1})\big| \\ &\leq \frac{1}{l}(\psi(b)-\psi(a)+\phi(b)-\phi(a)), \end{aligned} \tag{52}$$

where $t \in [a, b)$. Recall that $\varphi(b) = \varphi(b^-)$. Furthermore, since $\tau_{m-1} < \tau_m$, we have $\varphi_l(b) = \varphi_l(b^-)$. In other words, both φ and φ_l are continuous from the left at $t = b$. It therefore follows that (52) also holds for $t = b$. Thus,

$$\big|\varphi_l(t)-\varphi(t)\big| \leq \frac{1}{l}(\psi(b)-\psi(a)+\phi(b)-\phi(a)), \quad t \in [a,b].$$

This shows that φ_l converges to φ uniformly, as required. □

It is important to note that the definition of φ_l in the proof of Lemma 4.8 (see equation (49)) is consistent with the definition of the piecewise-constant controls in Section 3. This observation will be exploited in the next section.

5. Main Convergence Results.

Our aim in this section is to show that the suboptimal control $\boldsymbol{u}^{p,*}$ generated from the solution of Problem P_p (see equation (16)) is such that $J(\boldsymbol{u}^{p,*}) \to J(\boldsymbol{u}^*)$ as $p \to \infty$, where $\boldsymbol{u}^*$ is an optimal control for Problem P. In other words, the cost of the suboptimal control converges to the true optimal cost as the number of subintervals approaches infinity.

Our first theorem follows readily from Lemma 4.7.

Theorem 5.1. *Let $\boldsymbol{u} \in \mathcal{U}$ be an arbitrary admissible control. Then there exists a corresponding $\boldsymbol{v} \in \mathcal{U}$ such that:*

(a) $\boldsymbol{v}(t) = \boldsymbol{u}(t)$ *for almost all* $t \in [0, T]$.
(b) $\boldsymbol{v}$ *is continuous from the right on* $[0, T)$.
(c) $\boldsymbol{v}(T) = \boldsymbol{v}(T^-)$.
(d) $\bigvee_0^T \boldsymbol{v} \leq \bigvee_0^T \boldsymbol{u}$.
(e) $J(\boldsymbol{v}) \leq J(\boldsymbol{u})$.

Proof. Lemma 4.7 implies that there exists an admissible control $\boldsymbol{v} : [0, T] \to \mathbb{R}^r$ satisfying properties (a)-(d). To prove that $\boldsymbol{v}$ also satisfies property (e), note that since $\boldsymbol{v}(t) = \boldsymbol{u}(t)$ almost everywhere,

$$\boldsymbol{x}(t|\boldsymbol{v}) = \boldsymbol{x}(t|\boldsymbol{u}), \quad t \in [0,T].$$

Hence,

$$\Phi(\boldsymbol{x}(T|\boldsymbol{v})) = \Phi(\boldsymbol{x}(T|\boldsymbol{u}))$$

and

$$\int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{v}), \boldsymbol{v}(t))dt = \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t))dt.$$

Combining these two equations with the inequality in part (d) yields

$$\begin{aligned} J(\boldsymbol{v}) &= \Phi(\boldsymbol{x}(T|\boldsymbol{v})) + \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{v}), \boldsymbol{v}(t))dt + \gamma \bigvee_0^T \boldsymbol{v} \\ &\leq \Phi(\boldsymbol{x}(T|\boldsymbol{u})) + \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t))dt + \gamma \bigvee_0^T \boldsymbol{u} \\ &= J(\boldsymbol{u}). \end{aligned}$$

This shows that $\boldsymbol{v}$ satisfies property (e), as required. ◻

The control $\boldsymbol{v}$ in Theorem 5.1 can be viewed as a "smoother" version of $\boldsymbol{u}$. Although both $\boldsymbol{u}$ and $\boldsymbol{v}$ produce the same state trajectory, $\boldsymbol{v}$ would normally be preferred in practice because a smoother control will be easier to implement. As in Chapter 10 of [17], we call $\boldsymbol{v}$ the *minimal bounded variation control* corresponding to $\boldsymbol{u} \in \mathcal{U}$. Let $\mathcal{V}$ denote the set of all such minimal bounded variation controls.

Our next theorem extends the results in Lemma 4.8 to control functions in $\mathcal{V}$.

Theorem 5.2. *Let $\boldsymbol{u} \in \mathcal{V}$. Then there exists a sequence of piecewise-constant controls $\{\boldsymbol{u}^{p_l}\}_{l=1}^{\infty}$, where each $\boldsymbol{u}^{p_l} \in \mathcal{U}^{p_l}$, such that $\boldsymbol{u}^{p_l} \to \boldsymbol{u}$ uniformly on $[0,T]$ as $l \to \infty$. Furthermore,*

$$\lim_{l\to\infty} J(\boldsymbol{u}^{p_l}) = J(\boldsymbol{u}). \tag{53}$$

Proof. Note that u_i, the ith component of $\boldsymbol{u} \in \mathcal{V}$, is of bounded variation, right-continuous on $[0,T)$, and left-continuous at $t = T$. Hence, according to Lemma 4.8, there exists a sequence of piecewise-constant functions $\{u_i^l\}_{l=1}^{\infty}$ such that $u_i^l \to u_i$ uniformly as $l \to \infty$, where

$$\bigvee_0^T u_i^l \leq \bigvee_0^T u_i, \quad l \geq 1. \tag{54}$$

Now, for each integer $l \geq 1$, define a function $\tilde{\boldsymbol{u}}^l : [0,T] \to \mathbb{R}^r$ as follows:

$$\tilde{\boldsymbol{u}}^l(t) := \left[u_1^l(t), \dots, u_r^l(t)\right]^\top, \quad t \in [0,T].$$

Recall from the proof of Lemma 4.8 that each u_i^l is right-continuous and piecewise-constant. Thus, $\tilde{\boldsymbol{u}}^l$ is also right-continuous and piecewise-constant. In fact, it is clear that $\tilde{\boldsymbol{u}}^l$ can be expressed in the form of equation (7) for some integer $p = p_l$. Thus, $\tilde{\boldsymbol{u}}^l \in \mathcal{U}^{p_l}$. Furthermore, since $u_i^l \to u_i$ uniformly on $[0,T]$ for each $i = 1, \dots, r$, $\tilde{\boldsymbol{u}}^l \to \boldsymbol{u}$ uniformly on $[0,T]$. Hence, $\{\tilde{\boldsymbol{u}}^l\}_{l=1}^{\infty}$ is the required sequence of piecewise-constant controls.

We now complete the proof by showing that $\{\tilde{\boldsymbol{u}}^l\}_{l=1}^{\infty}$ satisfies equation (53). First, since $\tilde{\boldsymbol{u}}^l \to \boldsymbol{u}$ uniformly and $\boldsymbol{u}$ is bounded, there exists a constant $M_1 > 0$ such that

$$\left\|\tilde{\boldsymbol{u}}^l(t)\right\| \leq M_1, \quad t \in [0,T], \quad l \geq 1. \tag{55}$$

Furthermore, recall from inequality (3) that

$$\|\boldsymbol{f}(t, \boldsymbol{\xi}, \boldsymbol{\theta})\| \leq L(1 + \|\boldsymbol{\xi}\| + \|\boldsymbol{\theta}\|), \quad (t, \boldsymbol{\xi}, \boldsymbol{\theta}) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^r, \tag{56}$$

where $L > 0$ is a constant and $\boldsymbol{f}$ is the function defining the system dynamics in Problem P. In view of (55) and (56), it follows from Lemma 6.4.2 of [17] that the sequence of state trajectories $\{\boldsymbol{x}(\cdot|\tilde{\boldsymbol{u}}^l)\}_{l=1}^{\infty}$ is uniformly bounded. Hence, there exists a constant $M_2 > 0$ such that

$$\left\|\boldsymbol{x}(t|\tilde{\boldsymbol{u}}^l)\right\| \leq M_2, \quad t \in [0,T], \quad l \geq 1.$$

Therefore, since $\mathcal{L} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ in the cost function (5) is continuous, there exists another constant $M_3 > 0$ such that

$$\left|\mathcal{L}(t, \boldsymbol{x}(t|\tilde{\boldsymbol{u}}^l), \tilde{\boldsymbol{u}}^l(t))\right| \leq M_3, \quad t \in [0,T], \quad l \geq 1. \tag{57}$$

Now, according to Lemma 6.4.3 of [17], if a sequence of admissible controls converges *almost everywhere*, then the corresponding sequence of state trajectories converges *uniformly*. We have already shown that $\tilde{\boldsymbol{u}}^l \to \boldsymbol{u}$ uniformly on $[0,T]$ as $l \to \infty$. Thus,

$$\lim_{l\to\infty} \boldsymbol{x}(t|\tilde{\boldsymbol{u}}^l) = \boldsymbol{x}(t|\boldsymbol{u}), \quad t \in [0,T].$$

Consequently, since Φ and $\mathcal{L}$ in the cost function (5) are continuous,

$$\lim_{l\to\infty} \Phi(\boldsymbol{x}(T|\tilde{\boldsymbol{u}}^l)) = \Phi(\boldsymbol{x}(T|\boldsymbol{u})) \tag{58}$$

and

$$\lim_{l\to\infty} \mathcal{L}(t, \boldsymbol{x}(t|\tilde{\boldsymbol{u}}^l), \tilde{\boldsymbol{u}}^l(t)) = \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)), \quad t \in [0,T]. \tag{59}$$

In view of (57) and (59), we may apply Lebesgue's dominated convergence theorem (see [15]) to obtain

$$\lim_{l\to\infty} \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\tilde{\boldsymbol{u}}^l), \tilde{\boldsymbol{u}}^l(t))dt = \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t))dt. \tag{60}$$

Now, from (54),

$$\bigvee_0^T \tilde{\boldsymbol{u}}^l = \sum_{i=1}^r \bigvee_0^T u_i^l \leq \sum_{i=1}^r \bigvee_0^T u_i = \bigvee_0^T \boldsymbol{u}, \quad l \geq 1. \tag{61}$$

This implies

$$\limsup_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l \leq \bigvee_0^T \boldsymbol{u}. \tag{62}$$

Note from (61) that the total variation of $\tilde{\boldsymbol{u}}^l$ is uniformly bounded with respect to l. Thus, by Lemma 4.6,

$$\bigvee_0^T \boldsymbol{u} = \sum_{i=1}^r \bigvee_0^T u_i \leq \sum_{i=1}^r \liminf_{l\to\infty} \bigvee_0^T u_i^l \leq \liminf_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l. \tag{63}$$

Combining (62) and (63) gives

$$\limsup_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l \leq \bigvee_0^T \boldsymbol{u} \leq \liminf_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l \leq \limsup_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l.$$

Thus,

$$\lim_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l = \liminf_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l = \limsup_{l\to\infty} \bigvee_0^T \tilde{\boldsymbol{u}}^l = \bigvee_0^T \boldsymbol{u}. \tag{64}$$

Equation (53) then follows from (58), (60), and (64). □

Theorem 5.2 asserts that for any $\boldsymbol{u} \in \mathcal{V}$, there exists a corresponding sequence of piecewise-constant controls converging to $\boldsymbol{u}$ *uniformly*. A similar result is proved in [17] with one important difference: the controls in [17] are assumed to be measurable, not necessarily of bounded variation, and thus the sequence of piecewise-constant controls is only guaranteed to converge *almost everywhere*. Here, we have exploited Jordan's theorem for functions of bounded variation to obtain uniform convergence in Theorem 5.2 (see the proof of Lemma 4.8). As we will see, uniform convergence is needed to prove the main result of this paper, as the continuous inequality constraints in Problem P depend on *both* the state and the control (in [17], only pure state constraints are considered).

Recall that $\mathcal{F}$, the feasible region for Problem P, is the set of all admissible controls $\boldsymbol{u} \in \mathcal{U}$ satisfying the following continuous inequality constraints (see (4)):

$$h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) \geq 0, \quad t \in [0, T], \quad j = 1, \ldots, q.$$

Let $\mathring{\mathcal{F}}$ denote the set of all admissible controls $\boldsymbol{u} \in \mathcal{U}$ such that

$$\inf_{t \in [0,T]} h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) > 0, \quad j = 1, \ldots, q. \tag{65}$$

We impose the following regularity condition.

Assumption 1. For each optimal control $\boldsymbol{u}^* \in \mathcal{F}$ of Problem P, there exists a corresponding $\bar{\boldsymbol{u}} \in \mathring{\mathcal{F}}$ such that

$$\lambda \bar{\boldsymbol{u}} + (1 - \lambda)\boldsymbol{u}^* \in \mathring{\mathcal{F}}, \quad \lambda \in (0, 1].$$

Similar assumptions are made in [6, 12, 17, 19].

Theorem 5.3. *Let $\boldsymbol{u} \in \mathcal{V}$ and $\{\boldsymbol{u}^{p_l}\}_{l=1}^{\infty}$ be as defined in Theorem 5.2, and suppose that $\boldsymbol{u} \in \mathring{\mathcal{F}}$. Then for all sufficiently large l, $\boldsymbol{u}^{p_l} \in \mathring{\mathcal{F}}$.*

Proof. We need to show that $\boldsymbol{u}^{p_l}$ satisfies inequality (65) for all sufficiently large l. Recall from the proof of Theorem 5.2 that:

1. $\boldsymbol{u}^{p_l} \in \mathcal{U}^{p_l} \subset \mathcal{U}$ for each $l \geq 1$.
2. $\boldsymbol{u}^{p_l} \to \boldsymbol{u}$ uniformly on $[0, T]$ as $l \to \infty$.
3. $\boldsymbol{x}(\cdot|\boldsymbol{u}^{p_l}) \to \boldsymbol{x}(\cdot|\boldsymbol{u})$ uniformly on $[0, T]$ as $l \to \infty$.
4. $\|\boldsymbol{u}^{p_l}(t)\| \leq M_1$ for all $t \in [0, T]$.
5. $\|\boldsymbol{x}(t|\boldsymbol{u}^{p_l})\| \leq M_2$ for all $t \in [0, T]$.

Let $M := \max\{M_1, M_2\}$. Furthermore, define

$$\mathcal{W}_1 := \{\, \boldsymbol{\xi} \in \mathbb{R}^n : \|\boldsymbol{\xi}\| \leq M \,\}$$

and

$$\mathcal{W}_2 := \{\, \boldsymbol{\theta} \in \mathbb{R}^r : \|\boldsymbol{\theta}\| \leq M \,\}.$$

Then the continuous functions h_j, $j = 1, \ldots, q$ are *uniformly continuous* on the compact set $[0, T] \times \mathcal{W}_1 \times \mathcal{W}_2$.

Since $\boldsymbol{u} \in \mathring{\mathcal{F}}$, there exists a constant $\delta > 0$ such that

$$h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) \geq \delta, \quad t \in [0, T], \quad j = 1, \ldots, q. \tag{66}$$

It follows from points 2 and 3 above, and from the uniform continuity of h_j, $j = 1, \ldots, q$, that there exists an $l_1 > 0$ such that for all integers $l \geq l_1$,

$$\left| h_j(t, \boldsymbol{x}(t|\boldsymbol{u}^{p_l}), \boldsymbol{u}^{p_l}(t)) - h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) \right| < \tfrac{1}{2}\delta, \quad t \in [0, T], \quad j = 1, \ldots, q.$$

Therefore,

$$h_j(t, \boldsymbol{x}(t|\boldsymbol{u}^{p_l}), \boldsymbol{u}^{p_l}(t)) > h_j(t, \boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{u}(t)) - \tfrac{1}{2}\delta, \quad t \in [0, T], \quad j = 1, \ldots, q. \tag{67}$$

Combining (66) and (67) gives

$$h_j(t, \boldsymbol{x}(t|\boldsymbol{u}^{p_l}), \boldsymbol{u}^{p_l}(t)) > \tfrac{1}{2}\delta, \quad t \in [0, T], \quad j = 1, \ldots, q,$$

which holds for all $l \geq l_1$. Thus, $\boldsymbol{u}^{p_l} \in \mathring{\mathcal{F}}$ for all integers $l \geq l_1$. ◻

We are now ready to prove the main convergence result of this paper.

Theorem 5.4. *Suppose that Problem P has an optimal control $\boldsymbol{u}^*$. For each $p \geq 1$, let $\boldsymbol{u}^{p,*}$ denote the suboptimal control constructed from the solution of Problem P_p according to equation (16). Then*

$$\lim_{p\to\infty} J(\boldsymbol{u}^{p,*}) = J(\boldsymbol{u}^*).$$

Proof. By Assumption 1, there exists a control $\bar{\boldsymbol{u}} \in \mathring{\mathcal{F}}$ such that for each $k \geq 1$,

$$\bar{\boldsymbol{u}}^k := \boldsymbol{u}^* + \frac{1}{k}(\bar{\boldsymbol{u}} - \boldsymbol{u}^*) \in \mathring{\mathcal{F}}. \tag{68}$$

Clearly, $\bar{\boldsymbol{u}}^k \to \boldsymbol{u}^*$ uniformly as $k \to \infty$. Thus, by using similar arguments to those used in the proof of Theorem 5.2, one can show that there exists a constant $M_1 > 0$ such that

$$\left|\mathcal{L}(t, \boldsymbol{x}(t|\bar{\boldsymbol{u}}^k), \bar{\boldsymbol{u}}^k(t))\right| \leq M_1, \quad t \in [0,T], \quad k \geq 1. \tag{69}$$

Furthermore, it follows from Lemma 6.4.3 in [17] that $\boldsymbol{x}(\cdot|\bar{\boldsymbol{u}}^k)$ converges uniformly to $\boldsymbol{x}(\cdot|\boldsymbol{u}^*)$ as $k \to \infty$. Thus, we have

$$\lim_{k\to\infty} \boldsymbol{x}(t|\bar{\boldsymbol{u}}^k) = \boldsymbol{x}(t|\boldsymbol{u}^*), \quad t \in [0,T].$$

Hence, since Φ and $\mathcal{L}$ are continuous functions,

$$\lim_{k\to\infty} \Phi(\boldsymbol{x}(T|\bar{\boldsymbol{u}}^k)) = \Phi(\boldsymbol{x}(T|\boldsymbol{u}^*)) \tag{70}$$

and

$$\lim_{k\to\infty} \mathcal{L}(t, \boldsymbol{x}(t|\bar{\boldsymbol{u}}^k), \bar{\boldsymbol{u}}^k(t)) = \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}^*), \boldsymbol{u}^*(t)), \quad t \in [0,T]. \tag{71}$$

In view of (69) and (71), we may apply the Lebesgue dominated convergence theorem (see [15]) to obtain

$$\lim_{k\to\infty} \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\bar{\boldsymbol{u}}^k), \bar{\boldsymbol{u}}^k(t))dt = \int_0^T \mathcal{L}(t, \boldsymbol{x}(t|\boldsymbol{u}^*), \boldsymbol{u}^*(t))dt. \tag{72}$$

Now, rearranging the definition of $\bar{\boldsymbol{u}}^k$ in (68) gives

$$\bar{\boldsymbol{u}}^k(t) = \frac{k-1}{k}\boldsymbol{u}^*(t) + \frac{1}{k}\bar{\boldsymbol{u}}(t), \quad t \in [0,T].$$

Thus,

$$\bigvee_0^T \bar{\boldsymbol{u}}^k \leq \frac{k-1}{k}\bigvee_0^T \boldsymbol{u}^* + \frac{1}{k}\bigvee_0^T \bar{\boldsymbol{u}} \leq \bigvee_0^T \boldsymbol{u}^* + \bigvee_0^T \bar{\boldsymbol{u}}. \tag{73}$$

This shows that the total variation of $\bar{\boldsymbol{u}}^k$ is uniformly bounded with respect to k. Therefore, by Lemma 4.6,

$$\bigvee_0^T \boldsymbol{u}^* = \sum_{i=1}^r \bigvee_0^T u_i^* \leq \sum_{i=1}^r \liminf_{k\to\infty} \bigvee_0^T \bar{u}_i^k \leq \liminf_{k\to\infty} \bigvee_0^T \bar{\boldsymbol{u}}^k, \tag{74}$$

where u_i^* is the ith component of $\boldsymbol{u}^*$ and $\bar{u}_i^k$ is the ith component of $\bar{\boldsymbol{u}}^k$. Suppose that the inequality in (74) is strict:

$$\bigvee_0^T \boldsymbol{u}^* < \liminf_{k\to\infty} \bigvee_0^T \bar{\boldsymbol{u}}^k.$$

Then there exists constants $k' > 0$ and $\bar{\epsilon} > 0$ such that

$$\bigvee_0^T \boldsymbol{u}^* + \bar{\epsilon} < \bigvee_0^T \bar{\boldsymbol{u}}^k, \quad k \geq k'. \tag{75}$$

Recall from (73) that

$$\bigvee_0^T \bar{\boldsymbol{u}}^k \leq \frac{k-1}{k} \bigvee_0^T \boldsymbol{u}^* + \frac{1}{k} \bigvee_0^T \bar{\boldsymbol{u}}. \tag{76}$$

Clearly,

$$\lim_{k\to\infty} \left\{ \frac{k-1}{k} \bigvee_0^T \boldsymbol{u}^* + \frac{1}{k} \bigvee_0^T \bar{\boldsymbol{u}} \right\} = \bigvee_0^T \boldsymbol{u}^*. \tag{77}$$

From (76) and (77), we see that there exists a constant $k'' > k'$ such that for each integer $k \geq k''$,

$$\bigvee_0^T \bar{\boldsymbol{u}}^k - \bigvee_0^T \boldsymbol{u}^* \leq \frac{k-1}{k} \bigvee_0^T \boldsymbol{u}^* + \frac{1}{k} \bigvee_0^T \bar{\boldsymbol{u}} - \bigvee_0^T \boldsymbol{u}^* < \bar{\epsilon}.$$

Hence,

$$\bigvee_0^T \bar{\boldsymbol{u}}^k < \bar{\epsilon} + \bigvee_0^T \boldsymbol{u}^*, \quad k \geq k''.$$

But this contradicts (75), and so our initial assumption that inequality (74) is strict is false. Therefore, we must have

$$\bigvee_0^T \boldsymbol{u}^* = \liminf_{k\to\infty} \bigvee_0^T \bar{\boldsymbol{u}}^k. \tag{78}$$

Combining (70), (72), and (78) yields

$$\liminf_{k\to\infty} J(\bar{\boldsymbol{u}}^k) = J(\boldsymbol{u}^*). \tag{79}$$

Now, let $\epsilon > 0$ be arbitrary but fixed. Then in view of (79), there exists an integer $\kappa \geq 1$ such that

$$\left| J(\bar{\boldsymbol{u}}^\kappa) - J(\boldsymbol{u}^*) \right| < \tfrac{1}{2}\epsilon. \tag{80}$$

Note that $\bar{\boldsymbol{u}}^\kappa \in \mathring{\mathcal{F}}$ (see (68)).

Let $\bar{\boldsymbol{v}}^\kappa$ denote the minimal bounded variation control in $\mathcal{V}$ corresponding to $\bar{\boldsymbol{u}}^\kappa$ (Theorem 5.1 ensures that $\bar{\boldsymbol{v}}^\kappa$ exists). We will show that $\bar{\boldsymbol{v}}^\kappa$ is feasible for Problem P. First, since $\bar{\boldsymbol{u}}^\kappa \in \mathring{\mathcal{F}}$, there exists a positive constant $\delta > 0$ such that

$$h_j(t, \boldsymbol{x}(t|\bar{\boldsymbol{u}}^\kappa), \bar{\boldsymbol{u}}^\kappa(t)) \geq \delta, \quad t \in [0,T], \quad j = 1, \ldots, q.$$

Thus, since $\bar{\boldsymbol{v}}^\kappa = \bar{\boldsymbol{u}}^\kappa$ almost everywhere on $[0,T]$, there exists a set $\mathcal{T} \subset [0,T]$ of measure zero such that

$$h_j(t, \boldsymbol{x}(t|\bar{\boldsymbol{v}}^\kappa), \bar{\boldsymbol{v}}^\kappa(t)) \geq \delta, \quad t \in [0,T] \setminus \mathcal{T}, \quad j = 1, \ldots, q. \tag{81}$$

Now, if $t \in \mathcal{T} \setminus \{T\}$, then there exists a sequence $\{t_i\}_{i=1}^\infty \subset [0,T] \setminus \mathcal{T}$ such that $t_i \to t+$. It follows from (81) that for each integer $i \geq 1$,

$$h_j(t_i, \boldsymbol{x}(t_i|\bar{\boldsymbol{v}}^\kappa), \bar{\boldsymbol{v}}^\kappa(t_i)) \geq \delta, \quad j = 1, \ldots, q.$$

Thus, since h_j and $\boldsymbol{x}(\cdot|\bar{\boldsymbol{v}}^\kappa)$ are continuous, and $\bar{\boldsymbol{v}}^\kappa$ is right-continuous,

$$h_j(t, \boldsymbol{x}(t|\bar{\boldsymbol{v}}^\kappa), \bar{\boldsymbol{v}}^\kappa(t)) = \lim_{i\to\infty} h_j(t_i, \boldsymbol{x}(t_i|\bar{\boldsymbol{v}}^\kappa), \bar{\boldsymbol{v}}^\kappa(t_i)) \geq \delta, \quad j = 1, \ldots, q, \tag{82}$$

where $t \in \mathcal{T} \setminus \{T\}$. A similar proof shows that (82) is also satisfied at $t = T$. Thus, inequalities (81) and (82) show that $\bar{\boldsymbol{v}}^\kappa \in \mathring{\mathcal{F}}$, so $\bar{\boldsymbol{v}}^\kappa$ is feasible for Problem P.

Now, using (80) we obtain

$$J(\bar{\boldsymbol{v}}^\kappa) \leq J(\bar{\boldsymbol{u}}^\kappa) < J(\boldsymbol{u}^*) + \tfrac{1}{2}\epsilon$$

and

$$\big|J(\bar{\boldsymbol{v}}^{\kappa}) - J(\boldsymbol{u}^*)\big| = J(\bar{\boldsymbol{v}}^{\kappa}) - J(\boldsymbol{u}^*) < \tfrac{1}{2}\epsilon. \tag{83}$$

Let $\{\bar{\boldsymbol{v}}^{\kappa,p_l}\}_{l=1}^{\infty}$ denote the sequence of piecewise-constant controls in Theorem 5.2 converging to $\bar{\boldsymbol{v}}^{\kappa}$ uniformly. Then by Theorems 5.2 and 5.3, there exists a constant $l_1 > 0$ such that for all $l \geq l_1$,

$$\bar{\boldsymbol{v}}^{\kappa,p_l} \in \mathring{\mathcal{F}}$$

and

$$\big|J(\bar{\boldsymbol{v}}^{\kappa,p_l}) - J(\bar{\boldsymbol{v}}^{\kappa})\big| < \tfrac{1}{2}\epsilon. \tag{84}$$

Now, choose a fixed $l \geq l_1$. For each integer $p \geq p_l$, we have $\mathcal{U}^{p_l} \subset \mathcal{U}^p$. Thus, since $\bar{\boldsymbol{v}}^{\kappa,p_l} \in \mathcal{U}^{p_l}$,

$$J(\boldsymbol{u}^{p,*}) \leq J(\bar{\boldsymbol{v}}^{\kappa,p_l}), \quad p \geq p_l.$$

This implies that for each integer $p \geq p_l$,

$$0 \leq J(\boldsymbol{u}^{p,*}) - J(\boldsymbol{u}^*) \leq J(\bar{\boldsymbol{v}}^{\kappa,p_l}) - J(\boldsymbol{u}^*) \leq \big|J(\bar{\boldsymbol{v}}^{\kappa,p_l}) - J(\bar{\boldsymbol{v}}^{\kappa})\big| + \big|J(\bar{\boldsymbol{v}}^{\kappa}) - J(\boldsymbol{u}^*)\big|.$$

Hence, by using (83) and (84),

$$0 \leq J(\boldsymbol{u}^{p,*}) - J(\boldsymbol{u}^*) \leq \epsilon, \quad p \geq p_l.$$

Since $\epsilon > 0$ was chosen arbitrarily, this shows that $J(\boldsymbol{u}^{p,*}) \to J(\boldsymbol{u}^*)$ as $p \to \infty$. □

6. **Conclusion.** In this paper, we have considered an optimal control problem in which the cost function contains a total variation term measuring changes in the control action, and the governing dynamic system is subject to continuous inequality constraints involving both the state and the control. Using the control parameterization technique, we showed that this optimal control problem can be approximated by a semi-infinite programming problem. Solving this semi-infinite programming problem yields a suboptimal control for the original problem. We showed that the control parameterization method is convergent in the sense that the cost of the suboptimal control converges to the true optimal cost as the discretization of the time horizon is refined. The proof of this result is based on Jordan's theorem, which states that a function of bounded variation can be expressed as the difference of two non-decreasing functions. It remains a challenge to prove Theorem 5.4 for the general case in which the admissible controls are general measurable functions. This is a topic for future research.

REFERENCES

[1] B. Açikmeşe and L. Blackmore, *Lossless convexification of a class of optimal control problems with non-convex control constraints*, Automatica, **47** (2011), 341-347.
[2] N. U. Ahmed, "Dynamic Systems and Control with Applications," World Scientific, Singapore, 2006.
[3] C. Büskens and H. Maurer, *SQP-methods for solving optimal control problems with control and state constraints: Adjoint variables, sensitivity analysis and real-time control*, Journal of Computational and Applied Mathematics, **120** (2000), 85-108.
[4] M. Gerdts and M. Kunkel, *A nonsmooth Newton's method for discretized optimal control problems with state and control constraints*, Journal of Industrial and Management Optimization, **4** (2008), 247-270.
[5] C. J. Goh and K. L. Teo, *Control parametrization: A unified approach to optimal control problems with general constraints*, Automatica, **24** (1988), 3-18.
[6] L. S. Jennings and K. L. Teo, *A computational algorithm for functional inequality constrained optimization problems*, Automatica, **26** (1990), 371-375.
[7] A. N. Kolmogorov and S. V. Fomin, "Introductory Real Analysis," Dover edition, Dover Publications, New York, 1975.

[8] B. Li, C. J. Yu, K. L. Teo, and G. R. Duan, *An exact penalty function method for continuous inequality constrained optimal control problem*, Journal of Optimization Theory and Applications, **151** (2011), 260-291.

[9] Q. Lin, R. Loxton, K. L. Teo, and Y. H. Wu, *A new computational method for a class of free terminal time optimal control problems*, Pacific Journal of Optimization, **7** (2011), 63-81.

[10] R. Loxton, K. L. Teo, and V. Rehbock, *Optimal control problems with multiple characteristic time points in the objective and constraints*, Automatica, **44** (2008), 2923-2929.

[11] R. Loxton, K. L. Teo, V. Rehbock, and W. K. Ling, *Optimal switching instants for a switched-capacitor DC/DC power converter*, Automatica, **45** (2009), 973-980.

[12] R. Loxton, K. L. Teo, V. Rehbock, and K. F. C. Yiu, *Optimal control problems with a continuous inequality constraint on the state and the control*, Automatica, **45** (2009), 2250-2257.

[13] D. G. Luenberger and Y. Ye, "Linear and Nonlinear Programming," 3rd edition, Springer, New York, 2008.

[14] J. Nocedal and S. J. Wright, "Numerical Optimization," 2nd edition, Springer, New York, 2006.

[15] H. L. Royden and P. M. Fitzpatrick, "Real Analysis," 4th edition, Prentice Hall, Boston, 2010.

[16] W. Rudin, "Principles of Mathematical Analysis," 3rd edition, McGraw-Hill, New York, 1976.

[17] K. L. Teo, C. J. Goh, and K. H. Wong, "A Unified Computational Approach to Optimal Control Problems," Longman Scientific and Technical, Essex, 1991.

[18] K. L. Teo and L. S. Jennings, *Optimal control with a cost on changing control*, Journal of Optimization Theory and Applications, **68** (1991), 335-357.

[19] K. L. Teo, V. Rehbock, and L. S. Jennings, *A new computational algorithm for functional inequality constrained optimization problems*, Automatica, **29** (1993), 789-792.

[20] L. Y. Wang, W. H. Gui, K. L. Teo, R. Loxton, and C. H. Yang, *Time delayed optimal control problems with multiple characteristic time points: Computation and industrial applications*, Journal of Industrial and Management Optimization, **5** (2009), 705-718.

[21] C. Yu, K. L. Teo, L. Zhang, and Y. Bai, *A new exact penalty method for semi-infinite programming problems*, Journal of Industrial and Management Optimization, **6** (2010), 895-910.

[22] Y. Zhao and M. A. Stadtherr, *Rigorous global optimization method for dynamic systems subject to inequality path constraints*, Industrial and Engineering Chemistry Research, **50** (2011), 12678-12693.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: r.loxton@curtin.edu.au
E-mail address: q.lin@curtin.edu.au
E-mail address: v.rehbock@curtin.edu.au
E-mail address: k.l.teo@curtin.edu.au