IEEE Control Systems Letters paper presented at
2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Discrete-Time Finite-Horizon Optimization of Singularly Perturbed Nonlinear Control Systems with State-Action Constraints

Jianzong Pi   Abhishek Gupta

***Abstract*— An algorithm is introduced for the computation of an approximate optimal control policy for discrete-time finite-horizon nonlinear singularly perturbed systems. This is achieved through timescale separation and by utilizing ideas from parametric optimization and dynamic programming. We demonstrate that our proposed method produces a control policy that is both theoretically robust and nearly optimal.**

***Index Terms*— Optimization algorithms; Constrained control; Optimal control.**

I. INTRODUCTION

Singularly perturbed systems play an important role in engineering, chemistry, biology, physics, and so on. In a singularly perturbed system, there may be two or more underlying coupled dynamical systems that are operating at different timescales.

In practice, the fast and the slow dynamics are typically decoupled in a way in which dynamic programming based solution is derived for the slow dynamics and a tracking control is applied to control the fast dynamics. Thus, by applying the singular perturbation theory, one can separately solve the control policy of the two timescales, and get the overall policy by combining them (this is made precise later in the paper). This approach greatly reduces the computational burden of applying dynamic programming over the entire dynamics with a higher dimensional state space. In turn, some optimality loss is incurred due to this approximate approach. Despite being extensively applied in practice, there are limited theoretical results regarding the timescale separation method for discrete time nonlinear systems with state-action constraints. Accordingly, we aim to bridge this gap in this paper: We derive an upper bound on the suboptimality gap for discrete-time finite-horizon nonlinear singularly-perturbed systems with state-action constraints.

A. Prior Work

Dynamic programming introduced by Bellman [1] has been successfully utilized to solve various types of sequential decision making problems, both finite and infinite in horizon [2] and with state-action constraints. Admittedly, singularly perturbed system can be optimized using dynamic programming. However, due to curse of dimensionality, the computational runtime is significantly high. Singular perturbation theory exploits the timescale separation structure in the problem to simplify the computation and arrive at a suboptimal policy.

The applications of singular perturbation theory are diverse, including aerospace [3], circuits and systems [4], [5], robotics [6] and so on. Comprehensive overviews of singular perturbation systems are presented in [7], [8], among several others. Robust control for singularly perturbed systems was studied in [9]–[13]. Much of the theoretical development has been done in continuous time [14]; there is scant literature on discrete time systems. Bidani et al. [15] and Kim et al. [16] studied discrete-time LQR and LQG optimal control problems. Thus, the study of optimization of discrete-time singularly perturbed system with nonlinear state transition function, cost function and state-action constraint is an important open problem in the field. This paper is an attempt in this direction.

The timescale separation in singularly perturbed system involves a parameter ε that is generally assumed to be a small positive constant. Thus, the optimization of singularly perturbed system can be viewed within the framework of parametric optimization. Continuity of the optimal solution in the parameter has been studied in [17]–[22]. Dutta et al. [23] investigated sufficient conditions under which the solutions of parametric dynamic programs are continuous in the parameters. More recently, [24] studied a warm start approach for perturbed dynamic program. This paper exploits the recent work to derive the suboptimality bound.

In this paper, we study the approximate algorithm for computing a suboptimal solution to discrete-time nonlinear singularly perturbed systems with state-action constraints. The algorithm we propose is a slight variation of the algorithms widely used in practice. In practice, the fast time scale problem is solved using an LQR controller with carefully tuned Q and R matrices. On the other hand, we use the value function derived from solving the slow dynamics optimization to derive a constrained quadratic program to solve the fast dynamics optimization. Consequently, our algorithm has certain desirable theoretical benefits in comparison to the ones used in practice. We utilize the theory of parametric dynamic programs and first-order perturbation methods for constrained optimization to derive the suboptimality gap. The suboptimality gap here refers to the difference between the total cost obtained by following our proposed control policy and the optimal cost achieved using the dynamic programming algorithm. To the best of our knowledge, this is the first paper to derive the suboptimality gap for

The authors are with the Department of Electrical and Computer Engineering at The Ohio State University, Columbus, OH, USA. Email: pi.35, gupta.706@osu.edu. The authors would like to thank Prof. Marcello Canova and Prof. Stephanie Stockar for insightful comments and discussions.

constrained singularly perturbed systems using the control policy obtained through the timescale separation approach.

B. Paper Outline

In Section II, we introduce the problem formulation of a two-timescale singularly perturbed system. Then, we study the decoupled algorithm to efficiently compute the approximately optimal policy. In Section III, we state the theoretical suboptimality gap and robustness of our approximated policy. Supplementary proofs are provided in Section IV. Finally, the conclusion is drawn in Section V.

C. Preliminaries

Throughout the paper, we endow the finite dimensional Euclidean space $\mathbb{R}^n$ with the usual ℓ_2 norm. Let $\mathcal{X} \subset \mathbb{R}^n$. By Heine–Borel theorem [25, 3.30], $\mathcal{X}$ is compact if and only if $\mathcal{X}$ is closed and bounded. The interior of $\mathcal{X}$, denoted by $\text{int}(\mathcal{X})$, is the largest open set that is contained in $\mathcal{X}$. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function. Then, the gradient $\nabla_x f(x) \in \mathbb{R}^{n\times m}$ is represented in a matrix form:

$$\nabla_x f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(x) & \cdots & \frac{\partial}{\partial x_1} f_m(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \end{bmatrix} \in \mathbb{R}^{n\times m}.$$

Fact 1 (Chain Rule): Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^p$ be differentiable functions. Then,

$$\nabla_x g \circ f|_{x_0} = \nabla_x f|_{x_0} \nabla_y g|_{f(x_0)}.$$

Fact 2 (First-order Taylor Expansion): Let $f : \mathcal{X} \to \mathbb{R}$ be a C^2 function. Let $x, x' \in \mathcal{X}$. Then, we write $f(x')$ as:

$$f(x') = f(x) + \nabla f(x)^T (x' - x) + O(\|x' - x\|^2).$$

We call $C : \mathcal{X} \rightrightarrows \mathcal{Y}$ a compact-valued correspondence (set-valued mapping) when $C(x)$ is a compact subset of $\mathcal{Y}$ for each $x \in \mathcal{X}$. A compact-valued correspondence C is upper-hemicontinuous at $x \in \mathcal{X}$ if and only if for all sequences $(x_n)_{n\in\mathbb{N}} \subset \mathcal{X}$, for all $y \in \mathcal{Y}$, and all sequences $(y_n)_{n\in\mathbb{N}}$ satisfying $y_n \in C(x_n)$,

$$x_n \to x \text{ and } y_n \to y \implies y \in C(x).$$

More information on correspondences can be found in [25].

II. Problem Formulation

Consider a two-timescale singularly perturbed control system with the following form:

$$\begin{aligned} x_{t+1} &= x_t + \varepsilon f_1(x_t, y_t, u_t), \\ y_{t+1} &= f_2(x_t, y_t, u_t), \end{aligned}$$

where $t = 1, \ldots, T$ is the time index, $x_t \in \mathcal{X} \subset \mathbb{R}^n$ is the slow state, $y_t \in \mathcal{Y} \subset \mathbb{R}^m$ is the fast state, $u_t \in \mathcal{U} \subset \mathbb{R}^k$ is the control input. f_1, f_2 are functions that characterize how the slow/fast states are updated, and $\varepsilon > 0$ is the perturbation parameter of the system. The model is a nonlinear generalization of the discrete time model considered in [15], [16] for linear systems. At time t, let $g_t : \mathcal{X} \times \mathcal{Y} \times \mathcal{U} \to \mathbb{R}^r$ be the state-action constraint function. In this paper, we assume all the constraints are of the form:

$$g_{t,i}(x_t, y_t, u_t) \le 0, \quad i \in \{1, \ldots r\},$$

let $c_t : \mathcal{X} \times \mathcal{Y} \times \mathcal{U} \to \mathbb{R}$ be the cost incurred at time t. Let $\pi_t : \mathcal{X} \times \mathcal{Y} \to \mathcal{U}$ be an admissible control policy such that $g_t(x, y, \pi_t(x, y)) \le 0$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$. The total cost is

$$J(\pi) := \sum_{t=1}^{T} c_t(x_t, y_t, \pi_t(x_t, y_t)).$$

The goal of an optimal control policy π^* is to minimize the total cost subject to the state-action constraints. For fixed $\varepsilon > 0$, we construct the following optimization problem:

$$\begin{aligned} P(\varepsilon) : \quad & \min_{\pi} J(\pi) \\ \text{subject to} \quad & g_t(x_t, y_t, u_t) \le 0, \\ & x_{t+1} = x_t + \varepsilon f_1(x_t, y_t, u_t), \\ & y_{t+1} = f_2(x_t, y_t, u_t), \quad u_t = \pi_t(x_t, y_t). \end{aligned}$$

We assume the cost function c_t and the constraint function g_t do not depend on ε for all $t = 1, \ldots, T$. Let π^*_ε be the solution to $P(\varepsilon)$. The computation of π^*_ε is done by solving $P(\varepsilon)$ recursively with dynamic programming. At the terminal state T, we first initialize the value functions to be

$$\begin{aligned} V_T(x_T, y_T) = \min_{u_T \in \mathcal{U}} c_T(x_T, y_T, u_T) \\ \text{subject to } g_T(x_T, y_T, u_T) \le 0 \end{aligned}$$

for each $x_T \in \mathcal{X}, y_T \in \mathcal{Y}$. Then, for $t = (T-1), \cdots, 1$, we define the state-action value function as:

$$\begin{aligned} Q_t(x_t, y_t, u_t) =& c_t(x_t, y_t, u_t) + \\ & V_{t+1}(x_t + \varepsilon f_1(x_t, y_t, u_t), f_2(x_t, y_t, u_t)). \end{aligned}$$

At stage t, we define the value function of the slow and fast states (x_t, y_t) as:

$$V_t(x_t, y_t) = \min_{u_t \in \mathcal{U}} Q_t(x_t, y_t, u_t) \text{ s.t. } g_t(x_t, y_t, u_t) \le 0. \quad (1)$$

With the assumption that the admissible action set has nonempty interior for each state, i.e. $\text{int}(\{u \in \mathcal{U} : g_t(x, y, u) \le 0\}) \neq \emptyset$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$ and $t = 1, \cdots, T$, dynamic programming is guaranteed to give us the optimal policy that minimizes the cost and satisfies the state-action constraints. However, solving the above dynamic programming $P(\varepsilon)$ can be time and memory-consuming when the dimension of the state space is as large as $n+m$ – this phenomenon is known as the *"curse of dimensionality"*.

Traditionally, to solve the problem without the state-action constraints, the slow and fast dynamics problems are decoupled, as described in [15], [16]. The two sub-problems are solved sequentially, and one obtains an approximately optimal policy with a small performance loss. In this paper, we extend that technique to nonlinear systems with state-action constraints formulated above. We are interested in computing a decoupled optimal policy $\hat{\pi}^*_{\varepsilon,t}$ of the form:

$$\hat{\pi}^*_{\varepsilon,t}(x_t, y_t) = \gamma^*_{\varepsilon,t}(x_t) + \mu^*_{\varepsilon,t}(x_t, y_t), \quad t = 1, \cdots, T$$

where we call $\gamma^*_{\varepsilon,t} : \mathcal{X} \to \mathcal{U}$ the optimal slow-state policy, and $\mu^*_{\varepsilon,t} : \mathcal{X} \times \mathcal{Y} \to \mathcal{U}$ the optimal fast-state policy. In this paper, we derive the decoupled approach to efficiently compute the approximate optimal control policy $\hat{\pi}^*_\varepsilon$ to the singularly perturbed optimal control problem $P(\varepsilon)$, so that $J(\hat{\pi}^*_\varepsilon) \approx J(\pi^*_\varepsilon)$. We analyze the performance loss for the approximated policy $\hat{\pi}^*_\varepsilon$. Moreover, we show the robustness of $\hat{\pi}^*_\varepsilon$ to variations in the singular perturbation parameter ε. We assume the system satisfies the following conditions.

Assumption 1: $\mathcal{X}, \mathcal{Y}, \mathcal{U}$ are compact, and f_1, f_2, c_t, g_t, V_t are C^2 mappings for every $t = 1, \dots, T$.

A. Slow Dynamics Optimization

In the sequel, we ignore the time index of the state and actions unless necessary. For the coupled slow and fast dynamics, we assume that the fast dynamics has a steady state that is related to the slow state x and the action u.

Assumption 2: There is a twice differentiable map $h : \mathcal{X} \times \mathcal{U} \to \mathcal{Y}$ such that

1) $f_2(x, h(x,u), u) = h(x,u)$,
2) for every $t = 1, \dots, T$ and $x \in \mathcal{X}$, the set $\{u \in \mathcal{U} : g_t(x, h(x,u), u) \le 0\}$ has nonempty interior.

With h defined in Assumption 2, we define the state update functions, cost function and the constraint function without the fast state for the slow dynamics as:

$$f_1^h(x,u) := f_1(x, h(x,u), u),\; c_t^h(x,u) := c_t(x, h(x,u), u),$$
$$g_t^h(x,u) := g_t(x, h(x,u), u), \tag{2}$$

assuming $y_t = h(x_t, u_t)$ for each $t = 1, \dots, T$. Then, we construct a dynamic programming problem which has a lower dimensional state space $\mathcal{X}$ instead of $\mathcal{X} \times \mathcal{Y}$:

$$\begin{aligned} P_1(\varepsilon) : \quad & \min_{\gamma_\varepsilon} \sum_{t=1}^{T} c_t^h(x_t, u_t) \\ \text{subject to} \quad & g_t^h(x_t, u_t) \le 0, \\ & x_{t+1} = x_t + \varepsilon f_1^h(x_t, u_t), \quad u_t = \gamma_{\varepsilon,t}(x_t). \end{aligned}$$

Given that the perturbation parameter ε is fixed, we solve $P_1(\varepsilon)$ using the dynamic programming algorithm. The optimal slow dynamics policy is derived by a sequence of constrained optimization problems as shown in (1). We denote the optimal slow policy of $P_1(\varepsilon)$ by $\gamma^*_\varepsilon = (\gamma^*_{\varepsilon,t})_{t=1}^T$, and let $\lambda^*_{\varepsilon,t}(x_t)$ be the Lagrange multiplier of action $\gamma^*_{\varepsilon,t}(x_t)$ associated with the constraint $g_t^h(x_t, \gamma^*_{\varepsilon,t}(x_t)) \le 0$. With the optimal slow dynamics policy γ^*_ε, we define the *steady state* of the fast dynamics: $\bar{y}_{\varepsilon,t} := h(x_t, \gamma^*_{\varepsilon,t}(x_t))$.

B. Fast Dynamics Optimization

In many cases, tracking control methods are used to determine the fast dynamics control policy after the computation of the slow dynamics policy. However, in this paper, we present a modified approach. Specifically, we leverage the optimal slow dynamics policy and the associated Lagrange multipliers to generate an approximate dynamic program, which we then solve using perturbed optimization techniques to efficiently obtain the optimal fast dynamics policy. With computed optimal slow-dynamic policy γ^*_ε, we define a new fast dynamics state $z_{\varepsilon,t} \in \mathcal{Y}$ capturing the difference between the actual fast state y_t and the steady state $\bar{y}_{\varepsilon,t}$: $z_{\varepsilon,t} := y_t - \bar{y}_{\varepsilon,t}$. At each backward induction stage $t = T, \dots, 1$, we aim to solve the following optimization:

$$\min_{u_t \in \mathcal{U}} \quad Q_t(x_t, y_t, u_t), \text{ subject to } g_t(x_t, y_t, u_t) \le 0.$$

Let $Q_t^h(x,u) := Q_t(x, h(x,u), u)$. Then, the state-action value function Q_t and the constraint g_t is approximated by their Taylor expansions:

$$Q_t(x,y,u) = Q_t^h(x,u) + \nabla_y Q_t(x, h(x,u), u)^T z + O(\|z\|^2),$$

$$g_t(x,y,u) = g_t^h(x,u) + \nabla_y g_t(x, h(x,u), u)^T z + O(\|z\|^2),$$

where $z = y - h(x,u)$. Note that the slow dynamics policy γ^*_ε is optimal when $z_{\varepsilon,t} = 0$ for $t = 1, \dots T$. At stage t, we denote the fast-dynamics action as $v_{\varepsilon,t}$: $v_{\varepsilon,t} := u_t - \gamma^*_{\varepsilon,t}(x)$. We derive an approximated fast policy by solving an approximated problem by linearizing the cost-to-go and constraint functions. The approximated problem is defined as $\hat{P}_2(\varepsilon)$:

$$\begin{aligned} \hat{P}_2(\varepsilon) : \quad & \min_{v \in \mathcal{V}_t(x)} \quad Q_t^h(x,u) + \nabla_y Q_t(x, h(x,u), u)^T z, \\ & \text{subject to } g_t^h(x,u) + \nabla_y g_t(x, h(x,u), u)^T z \le 0, \\ & \qquad u = \gamma^*_{\varepsilon,t}(x) + v, \end{aligned}$$

where $\mathcal{V}_t(x)$ is the set of feasible actions at state x:

$$\begin{aligned} \mathcal{V}_t(x) := \{ v \in \mathcal{U} : \gamma^*_{\varepsilon,t}(x) + v \in \mathcal{U} \\ \text{and } g_t^h(x, \gamma^*_{\varepsilon,t}(x) + v) \le 0 \}. \end{aligned}$$

In addition, we define $\mathcal{A}_t(x,u)$ as the index set of active constraints $\mathcal{A}_t(x,u) := \{ i \in \{1, \dots, r\} : g_{t,i}^h(x,u) = 0 \}$, where $g_{t,i}^h(x,u)$ denotes the i-th entry of vector $g_t^h(x,u)$. We observe that the objective and constraint functions in $\hat{P}_2(\varepsilon)$ are perturbed by $\nabla_y Q_t(x, h(x,u), u)^T z + O(\|z\|^2)$ and $\nabla_y g_t(x, h(x,u), u)^T z + O(\|z\|^2)$, respectively. Hence, we can leverage the ideas from Gupta et al. [24] for fast computation of perturbed optimization problems. At stage t and state x, we define the following:

$$\begin{aligned} H_{\varepsilon,t}(x) :=& \nabla_{uu} Q_t^h(x, \gamma^*_{\varepsilon,t}(x)), \\ A_{\varepsilon,t}(x) :=& [\nabla_u g_{t,i}^h(x, \gamma^*_{\varepsilon,t}(x))]_{i \in \mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))}, \\ \tilde{e}_{\varepsilon,t}(x) :=& \nabla^2_{uy} g_t^h(x, \gamma^*_{\varepsilon,t}(x)), \\ \tilde{b}_{\varepsilon,t}(x) :=& -[\nabla_y g_{t,i}(x, h(x, \gamma^*_{\varepsilon,t}(x)), \gamma^*_{\varepsilon,t}(x))]_{i \in \mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))}, \\ M_{\varepsilon,t}(x) :=& [A_{\varepsilon,t}(x) H^{-1}_{\varepsilon,t}(x) A^T_{\varepsilon,t}(x)]^{-1}, \\ w_{\varepsilon,t}(x) :=& [\tilde{e}^T_{\varepsilon,t}(x), \tilde{b}^T_{\varepsilon,t}(x)]^T, \\ B_{\varepsilon,t}(x) :=& H^{-1}_{\varepsilon,t}(x)[A^T_{\varepsilon,t}(x) M_{\varepsilon,t}(x) A_{\varepsilon,t}(x) H^{-1}_{\varepsilon,t}(x) - I | \\ & A^T_{\varepsilon,t}(x) M_{\varepsilon,t}(x)]. \end{aligned}$$

Assume the set of active constraints for the fast dynamics stays the same as $\mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))$, then we have a closed-form approximated optimal fast-dynamics action:

$$\hat{\mu}^*_{\varepsilon,t}(x,z) := B_{\varepsilon,t}(x) w_{\varepsilon,t}(x) z.$$

We show that $\hat{\mu}^*_{\varepsilon,t}$ approximates the true fast-state policy with a bounded loss in performance. In the case when $\mathcal{A}_t$ changes, we need to solve the original constrained optimization problem $\hat{P}_2(\varepsilon)$. The joint approximated policy $\hat{\pi}^*_{\varepsilon,t}$ is

$$\hat{\pi}^*_{\varepsilon,t}(x,y) := \gamma^*_{\varepsilon,t}(x) + \hat{\mu}^*_{\varepsilon,t}(x, y - h(x, \gamma^*_{\varepsilon,t}(x))). \quad (3)$$

The algorithm for deriving $\hat{\pi}^*_\varepsilon$ is shown in Algorithm 1. In the sequel, we provide our main results on robustness and performance characteristics of $\hat{\pi}^*_\varepsilon$.

Algorithm 1 Decoupled Algorithm.

Require: f_1, f_2, h and c_t, g_t for $t = 1, \ldots, T$
  Perform slow dynamics DP. For each $t = 1, \ldots, T$, compute the optimal policy $\gamma^*_{\varepsilon,t}(x_t)$ with the associated Lagrange multiplier $\lambda^*_{\varepsilon,t}(x_t)$ and the value functions V_t.
  for $t = T, \ldots, 1$ **do**
    for each $x_t \in \mathcal{X}, y_t \in \mathcal{Y}$ **do**
      $z_t \leftarrow y_t - h(x_t, \gamma^*_{\varepsilon,t}(x_t))$.
      Compute $A_{\varepsilon,t}(x_t), \tilde{b}_{\varepsilon,t}(x_t)$.
      Compute $B_{\varepsilon,t}(x_t), w_{\varepsilon,t}(x_t)$.
      if $\mathcal{A}_t$ not changing **then**
        $\hat{\mu}^*_{\varepsilon,t}(z_t) \leftarrow B_{\varepsilon,t}(x_t) w_{\varepsilon,t}(x_t) z_{\varepsilon,t}$.
      else
        Compute $\hat{\mu}^*_{\varepsilon,t}$ by constrained optimization.
      end if
    end for
  end for

III. Main Results

This section has two main parts. In the first part, we show that $\hat{\pi}^*_\varepsilon$ we derived is robust to the singular perturbation parameter ε. We study the differentiability of the optimal policy as a function of the parameter ε. In the second part, we derive the suboptimality gap.

A. Robustness Results

In this section, we present the sufficient conditions that ensures continuity or differentiablity of mapping $\varepsilon \mapsto \hat{\pi}^*_\varepsilon$. We now present the first theorem that outlines the sufficient conditions under which the slow dynamics optimal policy is continuous over the slow state and the perturbation parameter. The proof is provided in Section IV-A.

Theorem 1: Assume Assumption 1 holds. In addition, assume that for each $t = 1, \ldots, T$:

1) $\{u : g^h_{t,i}(x,u) \leq 0\} \subset \mathcal{U}$ is convex for every $x \in \mathcal{X}$,
2) $x \mapsto \{u : g^h_t(x,u) \leq 0\}$ is lower hemicontinuous,

then $\gamma^*_\varepsilon(\cdot)$ is lower-hemicontinuous on $\mathcal{X} \times \mathcal{E}$.
In the sequel, we present another robustness theorem that provides sufficient conditions that imply the slow dynamics policy to be continuous and differentiable over the perturbation parameter ε. For each $t = 1, \ldots, T$, the Lagrangian of the constrained cost-to-go problem is defined as

$$L_{\varepsilon,t}(x,u,\lambda) := c^h_t(x,u) + V_{t+1}(x + \varepsilon f^h_1(x,u)) + \sum_{i \in \mathcal{A}_t(x,u)} \lambda^*_{t,i}(x) g^h_{t,i}(x,u)$$

We now state the following assumption to establish the continuity properties of the optimal policy. The implication of the assumption is widely discussed in chapters about Lagrange Multiplier theory for constrained optimization problems; see, for example, [17, Chapter 4], and [26, Chapter 3].

Assumption 3: 1) $\nabla_u L_{\varepsilon,t}(x, \gamma^*_{\varepsilon,t}(x), \lambda^*_{\varepsilon,t}(x)) = 0$,
2) $\{\nabla_u g^h_{t,i}(x, \gamma^*_{\varepsilon,t}(x))\}_{i \in \mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))}$ is a set of linearly independent vectors,
3) $\lambda^*_{\varepsilon,t,i}(x) > 0$ for each $i \in \mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))$,
4) For any d such that $\nabla_u g^h_{t,i}(x, \gamma^*_{\varepsilon,t}(x))^T d = 0, i \in \mathcal{A}_t(x, \gamma^*_{\varepsilon,t}(x))$,

$$d^T \nabla^2_{uu} L_{\varepsilon,t}(x, \gamma^*_{\varepsilon,t}(x), \lambda^*_{\varepsilon,t}) d > 0.$$

The second theorem provides conditions that imply the continuity and differentiability of the slow dynamics policy. The proof is provided in [17, Theorem 4.4]

Theorem 2: If Assumptions 1 - 3 hold, then there exists $\bar{\varepsilon}$ such that $\gamma^*_{\varepsilon,t}(x)$ is differentiable over $\mathcal{X} \times [0, \bar{\varepsilon})$ for all t.
Note that $\hat{\mu}^*_{\varepsilon,t}$ is continuous over ε when γ^*_ε is continuous over ε if Assumptions 1 - 3 hold. This is because when γ^*_ε is continuous over ε, the matrices $B_{\varepsilon,t}, w_{\varepsilon,t}$ and the fast state $z_{\varepsilon,t}$ are all continuous over ε. Hence, we show the robustness of the joint policy:

Corollary 3: If Assumptions 1 - 3 hold, then the joint approximated policy $\hat{\pi}^*_\varepsilon$ is continuous over ε.

B. Approximation Results

We now present our last main theorem on the suboptimality gap for the approximated policy $\hat{\pi}^*_\varepsilon$. We first need the following assumption that ensures the fast state $\|z_{\varepsilon,t}\| = O(\|z_{\varepsilon,1}\|)$. Specifically, if $z_{\varepsilon,1}$ is insignificant, it is necessary that $z_{\varepsilon,2}$, $z_{\varepsilon,3}$, and so on up to $z_{\varepsilon,T}$ to also be insignificant. This ensures that for every backward induction step, the terms involving $z_{\varepsilon,t}$ is treated as a minor perturbation of the constrained optimization problem we aim to solve. This is made precise in Lemma 7. We now present the last assumption before establishing the approximation results. We assume that the distance between the steady state and the actual fast state is non-expanding with time.

Assumption 4: For each $t = 1, \ldots, T-1$, there exists $0 < \alpha < 1$ such that for every $x \in \mathcal{X}$, $u \in \mathcal{U}$, the fast state $y_{t+1} = f_2(x, y_t, u)$ satisfies

$$\|y_{t+1} - h(x,u)\| \leq \alpha \|y_t - h(x,u)\|.$$

Moreover, we assume that $\|y_1 - h(x_1, u_1)\| << 1$.
The following theorem shows the theoretical performance guarantee for $\hat{\pi}^*_\varepsilon$. The proof is shown in Section IV-B.

Theorem 4: Assuming Assumptions 1 - 4 hold. Let $\hat{\pi}^*_\varepsilon$ be defined in (3), and π^*_ε be the true optimal policy. Assuming γ^*_ε is continuous over the slow states $x \in \mathcal{X}$, $c_t \equiv c$ and $g_t \equiv g$ for $t = 1, \ldots, T$. Then

$$J(\hat{\pi}^*_\varepsilon) - J(\pi^*_\varepsilon) = \sum_{t=1}^T \nabla_y Q_t(x_t, h(x_t, \gamma^*_{\varepsilon,t}(x_t)), \gamma^*_{\varepsilon,t}(x_t))^T z_{\varepsilon,t} - \sum_{t=1}^T \tilde{e}^T_{\varepsilon,t}(x_t) H^{-1}_{\varepsilon,t}(x_t) \tilde{e}_{\varepsilon,t}(x_t) + O(\|z_{\varepsilon,1}\|, \varepsilon).$$

$$
\begin{aligned}
\gamma^*_{\varepsilon,2}(x) &= \underset{u\in\{u:g_2^h(x,u)\le 0\}}{\operatorname{argmin}} \; c(x,u) + V_3^h(x'(u)),\\
\gamma^*_{\varepsilon,1}(x) &= \underset{u\in\{u:g_1^h(x,u)\le 0\}}{\operatorname{argmin}} \; c(x,u) + V_2^h(x'(u))\\
&= \underset{u\in\{u:g_1^h(x,u)\le 0\}}{\operatorname{argmin}} \; c(x,u) + \min_v c(x'(u), \gamma^*_{\varepsilon,2}(x'(u)+v) + V_3^h(x'(u) + \varepsilon f_1^h(x + \varepsilon f_1^h(x'(u), \gamma^*_{\varepsilon,2}(x'(u)) + v))\\
&= \underset{u\in\{u:g_1^h(x,u)\le 0\}}{\operatorname{argmin}} \; c(x,u) + \min_v c(x, \gamma^*_{\varepsilon,2}(x)) + \nabla_u c(x,\gamma^*_{\varepsilon,2}(x))^T(\varepsilon \nabla \gamma^*_{\varepsilon,2}(x)^T f_1^h(x,u) + v)\\
&\quad + \varepsilon \nabla_x c(x, \gamma^*_{\varepsilon,2}(x))^T f_1^h(x,u) + V_3^h[x'(u) + \varepsilon \nabla V_3^h(x'(u))^T f_1^h(x'(u), \gamma^*_{\varepsilon,2}(x'(u)) + v)] + O(\varepsilon, \|v\|)\\
&= \underset{u\in\{u:g_1^h(x,u)\le 0\}}{\operatorname{argmin}} \; c(x,u) + V_3^h(x'(u)) + O(\varepsilon, \|v\|),
\end{aligned}
$$

Fig. 1. These equations are used in the proof of Lemma 7.

IV. Proof of Main Results

A. Proof of Theorem 1

Lemma 5: Let $F : \Theta \times \mathcal{U} \to \mathbb{R}^r$ be a function. Let the feasible action correspondence $C : \Theta \rightrightarrows \mathcal{U}$ be defined as $C(\theta) := \{u : F(\theta, u) \le 0\}$. If F_i is continuous on $\Theta \times \mathcal{U}$ and $\{u : F_i(\theta, u) \le 0\}$ is a convex subset of $\mathcal{U}$ for $i = 1, \dots, r$, then C is upper-hemicontinuous.

Proof: Let $\theta \in \Theta$, $u \in \mathcal{U}$ and sequences $(\theta_n)_n \subset \Theta$, $(u_n)_n \subset \mathcal{U}$ such that

$$\theta_n \to \theta, u_n \to u, \text{ and } u_n \in C(\theta_n) \text{ for all } n \in \mathbb{N}.$$

For every $n \in \mathbb{N}$, $u_n \in C(\theta_n)$ implies $(F(\theta_n, u_n))_n$ is a non-positive sequence. As F_i is continuous, we get $F_i(\theta, u) = \lim_{n\to\infty} F_i(\theta_n, u_n) \le 0$, which implies the mapping $\theta \mapsto \{u : F_i(\theta, u) \le 0\}$ is upper-hemicontinuous. As $C(\theta) = \{u : F(\theta, u) \le 0\} = \cap_i \{u : F_i(\theta, u) \le 0\}$, and $\{u : F_i(\theta, u) \le 0\}$ is convex for each $i = 1, \dots, r$, [27, Theorem B] implies C is upper-hemicontinuous. ∎

Proof of Theorem 1: First, as c, h, g are continuous on compact sets, so they are uniformly continuous functions. The constraint correspondence C as defined in Lemma 5 is continuous and compact-valued. Moreover, the state update constraints are indeed continuous and compact-valued as they are single-valued. As the objective function $J = \sum_{t=1}^T c_t(x_t, u_t)$ is continuous over $\mathcal{E}$, Berge's maximum theorem [25, 17.31] implies the result.

B. Proof of Theorem 4

We first show if Assumptions 1 - 4 hold, then $z_{\varepsilon,t}$ is in order of $\|z_{\varepsilon,0}\|$ for every $t = 1, \dots, T$.

Lemma 6: Consider the following parametric optimization problem:

$$\min_{u\in\mathcal{U}} F(u) + \varepsilon \tilde{F}(u) \text{ subject to } G(u) + \varepsilon \tilde{G}(u) \le 0.$$

Let u^*_ε be the solution to the above constrained optimization problem, and λ^*_ϵ be the corresponding Lagrange multiplier. Let $\mathcal{A}(u) := \{i : G_i(u) = 0\}$ be the index set of active constraints. Suppose the following hold:

1) $\nabla_u \left(F(u^*_\varepsilon) + \sum_{i\in\mathcal{A}(u^*_\varepsilon)} \lambda^*_{\varepsilon,i} G_i(u^*_\varepsilon)\right) = 0$,
2) $\{\nabla_u G_i(u^*_\varepsilon)\}_{i\in\mathcal{A}(u^*_\varepsilon)}$ is a set of linearly independent vectors,
3) $\lambda^*_{\varepsilon,i} > 0$ for $i \in \mathcal{A}(u^*_\varepsilon)$,
4) For any d satisfying $\nabla_u G_i(u^*_\varepsilon)^T d = 0$, $i \in \mathcal{A}(u^*_\varepsilon)$,

$$d^T \nabla^2_{uu} \left(F(u^*_\varepsilon) + \sum_{i\in\mathcal{A}(u^*_\varepsilon)} \lambda^*_{\varepsilon,i} G_i(u^*_\varepsilon) \right) d > 0.$$

Then, there exists $\bar{\varepsilon} > 0$ such that $\varepsilon \mapsto u^*_\varepsilon$ is differentiable for all $\varepsilon \in [0, \bar{\varepsilon})$.

Proof: The proof is completed by directly applying [17, Theorem 4.4]. ∎

Lemma 7: Suppose Assumptions 1 - 4 hold and $c_t \equiv c$ for all $t = 1, \dots, T$. Then, there exists $\bar{\varepsilon} > 0$ such that $\|z_{\varepsilon,t}\| = O(\|z_{\varepsilon,1}\|)$ whenever $\varepsilon \in [0, \bar{\varepsilon})$.

Proof: For time $t = 2$, we get

$$
\begin{aligned}
\|z_{\varepsilon,2}\| =& \|y_2 - h(x_2, \gamma^*_{\varepsilon,2}(x_2))\|\\
\le& \alpha \|y_1 - h(x_2, \gamma^*_{\varepsilon,2}(x_2))\|\\
\le& \alpha \|z_{\varepsilon,1}\| + \alpha \|h(x_1, \gamma^*_{\varepsilon,1}(x_1)) - h(x_2, \gamma^*_{\varepsilon,2}(x_2))\|,
\end{aligned}
$$

where we used Assumption 4. As $\mathcal{X}$ is compact and h is continuous, we conclude that h is uniformly continuous. Moreover, by Theorem 2 and its corollary, $\gamma^*_{\varepsilon,t}$ is also continuous for all t. Consequently, it suffices to bound $\|\gamma^*_{\varepsilon,2}(x) - \gamma^*_{\varepsilon,1}(x)\|$ pointwise. We can achieve so by invoking uniform continuity of h, $\gamma^*_{\varepsilon,1}$, and $\gamma^*_{\varepsilon,2}$. Let $x'(u) = f_1^h(x, u)$ denote the next state given the current state is x and the action is u. From the derivation in Fig. 1 and Lemma 6, there exists $\bar{\varepsilon}_1 > 0$ such that $\|\gamma^*_{\varepsilon,2}(x) - \gamma^*_{\varepsilon,1}(x)\|$ is sufficiently small whenever $\varepsilon \in [0, \bar{\varepsilon}_1)$. Hence in this case, $\|z_{\varepsilon,2}\| = O(\|z_{\varepsilon,1}\|)$. Following the similar proof steps as above for $t = 2, 3, \dots, T-1$, we can determine $\bar{\varepsilon}_2, \dots, \bar{\varepsilon}_{T-1} > 0$ such that $\|z_{\varepsilon,t+1}\| = O(\|z_{\varepsilon,t}\|) = O(\|z_{\varepsilon,1}\|)$ whenever $\varepsilon \in [0, \bar{\varepsilon}_t)$. Lastly, we conclude our proof by letting $\bar{\varepsilon} = \min(\bar{\varepsilon}_1, \dots, \bar{\varepsilon}_{T-1})$. ∎

Lemma 7 ensures that z_t is a small perturbation parameter for all $t = 1, \dots, T$. We now derive the suboptimality gap. We begin our presentation of suboptimality bound with a single timestep problem. The proof of the following lemma comes from [24, Theorem 3].

Lemma 8: Consider the constrained dynamic programming problem at time t. The suboptimality gap is

$$\nabla_y Q_t(x,h(x,\gamma^*_{\varepsilon,t}(x)),\gamma^*_{\varepsilon,t}(x))^T z_{\varepsilon,t} \\ -\tilde{e}^T_{\varepsilon,t}(x)H^{-1}_{\varepsilon,t}(x)\tilde{e}_{\varepsilon,t}(x) + O(\|z_{\varepsilon,t}\|).$$

Theorem 4 is proved by applying Lemma 8 recursively and using Lemma 7 to bound the last term. A few remarks are in order. In the algorithm proposed here, we need derivatives of the value function to solve the constrained tracking problem in the fast timescale optimization problem. One can conduct numerical differentiation of the value function to construct the tracking metric. This was the approach taken in [24]. Currently, the tracking metric is tuned through trial and error to get the desired performance. Our approach obviates the trial and error method at the expense of a higher computational burden due to numerical differentiation.

Our derivation of the suboptimality bound does not hold for infinite horizon problem, since the errors accumulated over time would add up to infinity. In this case, one can consider discounted cost problem, so that the errors over time are also discounted. A rigorous derivation for this case will be conducted in our future work. To ensure stability of the resulting control policy in the infinite horizon setup, one can exploit the Lyapunov stability theorem. In addition, if the cost function and the constraint are also functions of the parameter ε, one can use the approach of [24] to derive the algorithm for computing a suboptimal policy. Due to space constraints, we did not derive those expressions in this paper.

V. Conclusion

In this paper, we devised an approximate algorithm for optimal control of singularly perturbed nonlinear systems with state-action constraints. Under appropriate assumptions, we demonstrate that the control policy is continuous in the perturbation parameter ε. To derive the approximately optimal control policy using timescale separation, we first compute the optimal policy and the associated Lagrange multipliers for the slow dynamics optimization. Thereafter, we use them to derive an approximately optimal policy for the fast dynamics optimization using the WASP algorithm [24]. Under further assumptions, we derive the suboptimality gap due to timescale separation of the optimization problem. The benefit of the approach developed here is that the matrices for tracking problem in the fast dynamics optimization is automatically computed using the solution from the slow dynamics optimization.

In the future, one can study the case when the cost function and the constraints are also a function of the parameter ε and extend our analysis to infinite horizon settings.

References

[1] R. Bellman, "On the theory of dynamic programming," *Proceedings of the national Academy of Sciences*, vol. 38, no. 8, pp. 716–719, 1952.

[2] M. L. Littman, *Algorithms for sequential decision-making*. Brown University, 1996.

[3] K.-H. Shim and M. Sawan, "Approximate controller design for singularly perturbed aircraft systems," *Aircraft Engineering and Aerospace Technology*, vol. 77, no. 4, pp. 311–316, 2005.

[4] A. Soto-Cota, L. M. Fridman, A. G. Loukianov, and J. M. Cañedo, "Power system singularly perturbed discontinuous control," in *Proceedings of the 2004 American Control Conference*, vol. 3, pp. 2580–2585, IEEE, 2004.

[5] H. M. Nguyen and D. S. Naidu, "Time scale analysis and control of wind energy conversion systems," in *2012 5th International Symposium on Resilient Control Systems*, pp. 149–154, IEEE, 2012.

[6] H. Salmasi, R. Fotouhi, and P. N. Nikiforuk, "On the stability of a friction compensation strategy for flexible-joint manipulators," *Advanced Robotics*, vol. 24, no. 15, pp. 2059–2086, 2010.

[7] P. V. Kokotovic, R. E. O'Malley Jr, and P. Sannuti, "Singular perturbations and order reduction in control theory—an overview," *Automatica*, vol. 12, no. 2, pp. 123–132, 1976.

[8] Y. Zhang, D. S. Naidu, C. Cai, and Y. Zou, "Singular perturbations and time scales in control theories and applications: An overview 2002–2012," *Int. J. Inf. Syst. Sci*, vol. 9, no. 1, pp. 1–36, 2014.

[9] H. Yu, G. Lu, and Y. Zheng, "On the model-based networked control for singularly perturbed systems with nonlinear uncertainties," *Systems & control letters*, vol. 60, no. 9, pp. 739–746, 2011.

[10] J. Wu and S. Wo, "Non-fragile robust control for singular systems with nonlinear perturbation and time-delay," in *2010 Chinese Control and Decision Conference*, pp. 3443–3447, IEEE, 2010.

[11] Z. Wang, W. Liu, H. Dai, and D. S. Naidu, "Robust stabilization of model-based uncertain singularly perturbed systems with networked time-delay," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 7917–7922, IEEE, 2009.

[12] G. Grammel, "Robustness of exponential stability to singular perturbations and delays," *systems & control letters*, vol. 57, no. 6, pp. 505–510, 2008.

[13] A. Chakrabortty and M. Arcak, "A three-time-scale redesign for robust stabilization and performance recovery of nonlinear systems with input uncertainties," in *2007 46th IEEE Conference on Decision and Control*, pp. 3484–3489, IEEE, 2007.

[14] H. K. Khalil, *Nonlinear control*, vol. 406. Pearson New York, 2015.

[15] M. Bidani, N. Radhy, and B. Bensassi, "Optimal control of discrete-time singularly perturbed systems," *International journal of control*, vol. 75, no. 13, pp. 955–966, 2002.

[16] B.-S. Kim, Y.-J. Kim, and M.-T. Lim, "Lqg control for nonstandard singularly perturbed discrete-time systems," *J. Dyn. Sys., Meas., Control*, vol. 126, no. 4, pp. 860–864, 2004.

[17] G. Still, "Lectures on parametric optimization: An introduction," *Optimization Online*, 2018.

[18] A. V. Fiacco, *Mathematical programming with data perturbations*. CRC Press, 2020.

[19] B. Bereanu, "The continuity of the optimum in parametric programming and applications to stochastic programming," *Journal of Optimization Theory and Applications*, vol. 18, pp. 319–333, 1976.

[20] O. L. Mangasarian and T.-H. Shiau, "Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems," *SIAM Journal on Control and Optimization*, vol. 25, no. 3, pp. 583–595, 1987.

[21] N. N. Tam, "On continuity properties of the solution map in quadratic programming," *Acta Mathematica Vietnamica*, vol. 24, pp. 47–61, 1999.

[22] X. Li and S. Li, "Hölder continuity of perturbed solution set for convex optimization problems," *Applied Mathematics and Computation*, vol. 232, pp. 908–918, 2014.

[23] P. K. Dutta, M. K. Majumdar, and R. K. Sundaram, "Parametric continuity in dynamic programming problems," *Journal of Economic Dynamics and Control*, vol. 18, no. 6, pp. 1069–1092, 1994.

[24] A. Gupta, S. R. Deshpande, and M. Canova, "An algorithm to warm start perturbed (WASP) constrained dynamic programs," *IEEE Open Journal of Control Systems*, vol. 1, pp. 1–14, 2022.

[25] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin and Heidelberg GmbH & Company KG, 2006.

[26] D. P. Bertsekas, "Nonlinear programming," *Journal of the Operational Research Society*, vol. 48, no. 3, pp. 334–334, 1997.

[27] A. Lechicki and A. Spakowski, "A note on intersection of lower semicontinuous multifunctions," *Proceedings of the American Mathematical Society*, vol. 95, no. 1, pp. 119–122, 1985.