OPTIMAL STOPPING TIME ON SEMI-MARKOV PROCESSES WITH FINITE HORIZON

FANG CHEN,* *Sun Yat-Sen University*

XIANPING GUO,* *Sun Yat-Sen University*

ZHONG-WEI LIAO,** *Beijing Normal University at Zhuhai*

Abstract

In this paper, we consider the optimal stopping problem on semi-Markov processes (SMPs) with finite horizon, and aim to establish the existence and computation of optimal stopping times. To achieve the goal, we first develop the main results of finite horizon semi-Markov decision processes (SMDPs) to the case with *additional* terminal costs, introduce an explicit construction of SMDPs, and prove the equivalence between the optimal stopping problems on SMPs and SMDPs. Then, using the equivalence and the results on SMDPs developed here, we not only show the existence of optimal stopping time of SMPs, but also provide an algorithm for computing optimal stopping time on SMPs. Moreover, we show that the optimal and ε-optimal stopping time can be characterized by the hitting time of some special sets, respectively.

Keywords: optimal stopping time, semi-Markov processes, value function, semi-Markov decision processes, optimal policy, iterative algorithm

2020 Mathematics Subject Classification: Primary 90C40

Secondary 93E20, 60G40

1. Introduction

The optimal stopping problem is an important branch of the intersection of probability and control theory, which aims to find the optimal stopping time of stochastic systems under

---
* Postal address: School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, China.

** Postal address: Corresponding author. College of Education for the Future, Beijing Normal University at Zhuhai, Zhuhai 519087, China.

** Email address: zhwliao@hotmail.com & zhwliao@bnu.edu.cn

some certain criterion, and has been studied and widely applied in finance, such as the pricing of American options and the buying-selling problem, see the monographes [3, 6, 19] and the references therein. In most existing literature on optimal stopping problems, the case on discrete-time Markov processes (de Saporta et al. [8]; Dufour & Piunovskiy [9]; Huang & Zhou [13]) and the case on continuous-time Markov processes (Arkin & Slastnikov [1]; Bäuerle & Popp [2]; Christensen & Lindensjö [7]; Gapeev et al. [10]; Shao & Tian [18]) are commonly considered.

Note that between two jump epochs in discrete-time Markov processes and continuous-time Markov processes are constant and exponentially distributed, respectively. As is well known, semi-Markov processes (SMPs) are more general than these two processes since the time between two jump epochs in SMPs can not only follow more distribution but also depend on the present state and the next state. Thus, SMPs haven been widely studied and applied in many areas [12, 17, 14]. However, to the best of our knowledge, the optimal stopping problem on SMPs is addressed only in Boshuizen & Gouweleeuw [4] and Chen et al. [5]. More precisely, Boshuizen & Gouweleeuw [4] and Chen et al. [5] both studied the optimal stopping problems under the infinite-horizon discounted criterion, and show the existence of an optimal time for time-dependent costs in Boshuizen & Gouweleeuw [4] and state-dependent costs in Chen et al. [5], respectively. As in well-known in many practical applications, the underlying process will be forced to end at a certain time. Thus, it is natural and desirable to consider the optimal stopping problem on SMPs with finite horizon, which has not been studied yet, and which will be studied in this paper.

In practical applications, if one breaks a contract, one has to pay an additional penalty, which will be regarded as a terminal cost. Thus, in optimal stopping problems with finite horizon, the costs consist of both the running costs and additional terminal costs [3]. To deal with the optimal stopping problem on SMPs with finite horizon, we follow the idea of transforming the discrete-time optimal stopping problem into an equivalent discrete-time Markov decision process in [3], which has an obvious advantage that the existence and computation of optimal stopping times can be obtained together. Indeed, noting that "continue" or "stop" may be considered as an action, using the date of the optimal stopping problem on SMPs, we also construct a corresponding SMDP with the action space $\{0, 1\}$, where the action 0 and action 1 mean continuation and stop respectively, and prove the equivalence between the optimal stopping time problems on SMPs and the corresponding SMDPs, that is, given any

deterministic policy in the SMDPs and planning horizon, we can induce a stopping time with the same expected cost as that for the policy, and vice versa. To deal with optimal stopping problem on SMPs with a terminal cost by the equivalent SMDP, we need to extend the results in [12] without any terminal cost to the more natural case with an additional terminal cost, and establish the existence of an optimal policy and an approximation algorithm for the value function of the SMDP by a minimal non-negative solution method. Using this equivalence and the results about SMDPs developed here, we not only show the existence of optimal stopping time on SMPs, but also provide an algorithm for computing optimal stopping time on SMPs. Moreover, we show that the optimal and ε-optimal stopping time can be characterized by the hitting time of some special sets.

The rest of the paper is organized as follows. We describe optimal stopping problems on SMPs with finite horizon in Section 2. In section 3, we develop some results on SMDPs with additional terminal costs. Our main results on the existences and computation of optimal stopping times are given in Section 4 after giving the preliminaries in Section 3.

2. Optimal stopping problem on semi-Markov processes

Notation. If X is a Borel space, we denote by $\mathcal{B}(X)$ the Borel σ-algebra, by $\mathcal{P}(X)$ the set of all probability measures on $\mathcal{B}(X)$, by δ_x the Dirac measure at the point x, and by $\mathbb{1}_D$ the indicator function on the set $D \subset X$. Moreover, let $\mathbb{R} := (-\infty, +\infty)$, $\mathbb{R}_+ := [0, +\infty)$, $x^+ := \max\{x, 0\}$ and $x \wedge y := \min\{x, y\}$ (for all $x, y \in \mathbb{R}$). Finally, for any sequence $\{x_k\} \subset R$, we use the convention $\sum_{k=n}^{m} y_k = 0$ if $n > m$.

The model of SMPs is the two-tuples as below

$$\{E, Q(\cdot, \cdot|x)\} \tag{1}$$

where E is the state space, which is assumed to be a Borel space and the transition mechanism of the SMPs is defined by the semi-Markov kernel $Q(\cdot, \cdot|x)$ on $\mathbb{R}_+ \times E$ given E. It is assumed that:

  (i) given any $B \in \mathcal{B}(E)$ and $x \in E$, $Q(\cdot, B|x)$ is a non-decreasing right continuous real-valued function on $\mathbb{R}_+$, with $Q(0, B|x) = 0$;

  (ii) given any $t \in \mathbb{R}_+$, $Q(t, \cdot|\cdot)$ is a sub-stochastic kernel on E;

  (iii) $\lim_{t\to\infty} Q(t, \cdot|\cdot)$ is a stochastic kernel on E.

Then, we introduce the measurable space $(\Omega, \mathscr{F})$, which is based on the Kitaev construction (see [15, 16]),

$$\Omega = \{(x_0, t_1, x_1, \ldots, t_n, x_n, \ldots) : x_0 \in E, (t_n, x_n) \in \mathbb{R}_+ \times E, n \geqslant 1\},$$

and $\mathscr{F}$ is the corresponding product Borel σ-algebra. The history of SMPs up to the n-th jump epoch is

$$h_0 = x_0, \quad h_{n+1} = (x_0, t_1, x_1, \ldots, t_{n+1}, x_{n+1}), \quad n \geqslant 0.$$

Let H_n be the set of all histories h_n. For each $\omega = (x_0, t_1, \ldots, x_n, t_{n+1}, \ldots) \in \Omega$, define

$$X_n(\omega) = x_n, \quad T_0(\omega) = 0, \quad T_{n+1}(\omega) = t_{n+1}, \quad S_n(\omega) = \sum_{k=0}^{n} T_k(\omega), \quad \forall n \geqslant 0,$$

where S_n, T_{n+1} and X_n denote the n-th jump time, the sojourn time between the n-th and $(n+1)$-th jumps and the state at the n-th jump time, respectively. Further, we assume here that the decision may only depend on the observation of the marked point process $\{T_n, X_n, n \geqslant 0\}$. Thus we denote by $\mathscr{F}_n$ the filtration generated by $\{T_n, X_n, n \geqslant 0\}$, i.e.,

$$\mathscr{F}_n := \sigma(T_0, X_0, \ldots, T_n, X_n).$$

Hence, we can give the definition of stopping times as following.

Definition 2.1. A random variable $\tau : \Omega \to \mathbb{N} \cup \{+\infty\}$ is called $\mathscr{F}_n$-stopping time if for all $n \in \mathbb{N}$,

$$\{\tau = n\} \in \mathscr{F}_n.$$

This condition means that upon observing the marked point process $\{X_n, T_n, n \geqslant 0\}$ until n-th jump time we can decide whether or not τ has already occurred. Since the filtration will always be generated by $\{X_n, T_n, n \geqslant 0\}$ in this paper, we will not mention it explicitly. Denote by Γ the set of all stopping times.

Using the Tulcea theorem (see [11, Proposition C.10]), for each $x \in E$, there exists a unique probability measure $\mathbb{P}_x$ on $(\Omega, \mathscr{F})$ satisfying that $\mathbb{P}_x(T_0 = 0, X_0 = x) = 1$ and

$$\mathbb{P}_x(T_{n+1} \leqslant t, X_{n+1} \in B|Y_n) = Q(t, B|X_n),$$

where $Y_n = (X_0, \ldots, T_n, X_n)$. Denote by $\mathbb{E}_x$ the expectation with respect to $\mathbb{P}_x$. Moreover, we give the following assumption, which can ensure the regularity of SMPs, i.e., $\mathbb{P}_x(\lim_{n\to\infty} S_n = \infty) = 1$.

Assumption 1. *There exist $\delta > 0$ and $\epsilon > 0$, such that*

$$Q(\delta, E|x) \leqslant 1 - \epsilon, \quad \forall x \in E. \tag{2}$$

The Assumption 1 is a standard regular condition widely used in SMPs and SMDPs, see [5, 12, 17], for instance. According to [12], the Assumption 1 implies that

$$\mathbb{P}_x(\lim_{n \to \infty} S_n = \infty) = 1, \quad \forall x \in E.$$

Corresponding to $\{(T_n, X_n), n \geqslant 0\}$, we define an underlying continuous-time state process $\{X(t), t \in \mathbb{R}_+\}$ by

$$X(t) = X_n, \quad S_n \leqslant t < S_{n+1}.$$

Refer to Limnios and Oprisan [17] for more details about $\{X(t), t \in [0, \infty)\}$. Next step, we introduce the optimal stopping problem with finite horizon.

Let $c(x)$ and $g(x)$ be the nonnegative measurable real-valued functions on E, which represent the cost rates and the terminal costs, respectively. For a given planning horizon $T \in \mathbb{R}_+$, the optimal stopping time problem with finite horizon T implies that if we have not stopped before time T we must stop paying at time T. Thus, for a given planning horizon $T \in \mathbb{R}_+$, if we choose a stopping time $\tau \in \Gamma$, we pay the cost

$$R_\tau^T := \begin{cases} \int_0^{S_\tau} c(X(t))\text{t} + g(X(S_\tau)), & S_\tau < T; \\ \int_0^T c(X(t))\text{t}, & S_\tau \geqslant T. \end{cases} \tag{3}$$

Remark 2.1. From the definition R_τ^T, if we stop before time T, we need to pay the terminal cost at the time S_τ. This is very common in practical applications, such as house renting problems. Because if one breaks the contract, one has to pay an additional penalty.

The T-horizon expected cost of a stopping time τ is given by

$$V^\tau(T, x) := \mathbb{E}_x\left[R_\tau^T\right], \quad x \in E. \tag{4}$$

Then, the value function of optimal stopping problems with finite horizon T is defined by

$$V^*(T, x) := \inf_{\tau \in \Gamma} V^\tau(T, x). \tag{5}$$

Definition 2.2. Given any planning horizon $T \in \mathbb{R}_+$, a stopping time $\tau^* \in \Gamma$ is called T-optimal if it satisfies that

$$V^{\tau^*}(T, x) = V^*(T, x) = \inf_{\tau \in \Gamma} V^\tau(T, x), \quad \forall x \in E.$$

Here and what follows, we fix a planning horizon $T \in \mathbb{R}_+$. The main purpose of this paper is to find a T-optimal stopping time and give an algorithm for computing the value function V^*.

3. On semi-Markov decision processes

We want to solve the stopping time problem by formulating it as SMDPs, so we need to consider the model of SMDPs with a terminal cost and give some results about the model. Moreover, if the terminal cost is always equal to 0, these results are same as those about SMDPs without terminal cost in [12].

Here and in what follows, we always use " ˆ " to distinguish SMDPs from SMPs. The model of SMDPs is introduced by:

$$\{\hat{E}, A, (A(x), x \in \hat{E}), \hat{Q}(\cdot, \cdot | x, a), \hat{c}(x, a), \hat{g}(x, a)\}$$

where $\hat{E}$ is the state space and A is the action set, which are assumed to be a Borel space and a denumerable set, respectively; $A(x) \subset A$ denotes the set of admissible actions at $x \in \hat{E}$, which assume to be finite; $\hat{Q}(\cdot, \cdot | x, a)$ is the semi-Markov kernel on $\mathbb{R}_+ \times \hat{E}$ given K, where $K := \{(x, a) | x \in \hat{E}, a \in A(x)\}$ denotes the set of admissible state-action pairs. Assume that $K \in \mathcal{B}(\hat{E}) \times \mathcal{B}(A)$ and that there exists a measurable mapping $f : \mathbb{R} \times \hat{E} \to A$ such that $(x, f(t, x)) \in K$ for all $(t, x) \in \mathbb{R} \times \hat{E}$. Finally, the functions $\hat{c}(x, a)$ and $\hat{g}(x, a)$ on K represent the cost rates and terminal costs, which are assumed to be nonnegative and measurable.

Remark 3.1. If $\hat{E}$ is denumerable and $\hat{g} = 0$, the model is same as that in [12].

The evolution of the finite horizon SMDPs as follows. Initially, the system occupies some state $x_0 \in \hat{E}$ and the decision maker has a planning horizon $s \in \mathbb{R}$, then he/she chooses an action $a_0 \in A(x_0)$ according to the current state x_0 and the planning horizon s. As a consequence, the system jumps to state x_1 after a sojourn time t_1 in x_0, in which the transition law is subject to the semi-Markov kernel $\hat{Q}$. At time t_1, there is a remaining planning horizon $s - t_1$ for the decision maker. According to the current state x_1 and the current planning horizon $s - t_1$ as well

as the previous state and action (x_0, a_0) and the sojourn time t_1, the decision maker chooses an action $a_1 \in A(x_1)$ and the same sequence of events occur. The decision process evolves in this way and thus we obtain a remaining planning horizon $s - \sum_{k=1}^{n} t_k$ and an admissible history $\hat{h}_n$ and of the SMDPs up to the n-th decision epoch i.e.,

$$\hat{h}_n = (x_0, a_0, t_1, x_1, \ldots, a_{n-1}, t_n, x_n),$$

where $(x_m, a_m) \in K$, $t_{m+1} \in \mathbb{R}_+$ for all $m = 0, 1, \ldots, n-1$, $x_n \in \hat{E}$. Let $\hat{H}_n$ denote the set of all admissible histories $\hat{h}_n$ of the system up to the n-th decision epoch, which is endowed with the corresponding product σ-algebra.

Definition 3.1. A policy $\pi = \{\pi_n, n \geqslant 0\}$ is a sequence of stochastic kernels π_n on A given $\mathbb{R} \times \hat{H}_n$ satisfying

$$\pi_n(A(x_n)|s, \hat{h}_n) = 1, \quad \forall n \geqslant 0, \hat{h}_n = (x_0, a_0, t_1, x_1, \ldots, a_{n-1}, t_n, x_n) \in \hat{H}_n. \tag{6}$$

The set of all policies is denoted by Π.

Remark 3.2. The s in (6) means the remaining planning horizon up to the n-th decision epoch, and it is assumed can be negative just for convenience. Moreover, the definition of policies here is horizon-relevant, whereas that in infinite horizon case is not.

To distinguish the subclasses of Π, we introduce the following notations.

   Notation. Let Φ represent the set of stochastic kernels φ on A given $\mathbb{R} \times \hat{E}$ such that $\varphi(A(x)|s, x) = 1$ for all $(s, x) \in \mathbb{R} \times \hat{E}$, and $\mathbb{F}$ represent the set of measurable functions $f :$ $\mathbb{R} \times \hat{E} \to A$ such that $f(s, x) \in A(x)$ for all $(s, x) \in \mathbb{R} \times \hat{E}$.

Definition 3.2. (a) A policy $\pi = \{\pi_n\}$ is said to be randomized Markov if there is a sequence $\{\varphi_n\}$ of stochastic kernels $\varphi_n \in \Phi$ such that $\pi_n(\cdot|s, \hat{h}_n) = \varphi_n(\cdot|s, x_n)$ for every $(s, \hat{h}_n) \in \mathbb{R} \times \hat{H}_n$ and $n \geqslant 0$. We write such a policy as $\pi = \{\varphi_n\}$.

(b) A randomized Markov policy $\pi = \{\varphi_n\}$ is said to be randomized stationary if φ_n are independent of n. In this case, we write π as φ for simplicity.

(c) A policy $\pi = \{\pi_n\}$ is called deterministic if there exists a sequence $\{d_n\}$ of measurable functions $d_n : \mathbb{R} \times \hat{H}_n \to A$ such that for all $(s, \hat{h}_n) \in \mathbb{R} \times \hat{H}_n$ and $n \geqslant 0$, $d_n(s, \hat{h}_n) \in A(x_n)$ and $\pi_n(\cdot|s, \hat{h}_n)$ is the Dirac measure at $d_n(s, \hat{h}_n)$, i.e.,

$$\pi_n(a|s, \hat{h}_n) = \delta_{\{d_n(s, \hat{h}_n)\}}(a), \quad \forall a \in A.$$

We write such a policy as $\pi = \{d_n\}$.

(d) A randomized Markov policy $\pi = \{\varphi_n\}$ is said to be deterministic Markov if there is a

   sequence $\{f_n\}$ of functions $f_n \in \mathbb{F}$ such that $\varphi_n(\cdot|s, x)$ is concentrated at $f_n(s, x)$ for all

   $(s, x) \in \mathbb{R} \times \hat{E}$ and $n \geqslant 0$. We write such a policy as $\pi = \{f_n\}$.

(e) A deterministic Markov policy $\pi = \{f_n\}$ is said to be deterministic stationary if f_n are

   independent of n. In this case, we write π as f for simplicity.

For convenience, we denote by Π_{RM}, Π_{RS}, Π_{DH}, Π_{DM} and Π_{DS} the families of all randomized Markov, randomized stationary, deterministic, deterministic Markov and deterministic stationary policies, respectively. Obviously, $\mathbb{F} = \Pi_{DS} \subset \Pi_{DM} \subset \Pi_{DH} \subset \Pi$ and $\mathbb{F} \subset \Phi = \Pi_{RS} \subset \Pi_{RM} \subset \Pi$.

Let $\hat{\Omega} = (\hat{E} \times A \times \mathbb{R}_+)^\infty$ be a sample space and $\hat{\mathscr{F}}$ be the corresponding product σ-algebra. Similar to SMPs, for any $\hat{\omega} = (x_0, a_0, t_1, x_1, \ldots, a_n, t_{n+1}, x_{n+1}, \ldots) \in \hat{\Omega}$ and $n \geqslant 0$, we can define

$$\hat{T}_0(\hat{\omega}) = 0, \quad \hat{T}_{n+1}(\hat{\omega}) = t_{n+1}, \quad \hat{X}_n(\hat{\omega}) = x_n, \quad A_n(\hat{\omega}) = a_n.$$

Further, for all $n \geqslant 0$, let $\hat{S}_n = \sum_{m=0}^n \hat{T}_m$. And then, we define $\{\hat{X}(t), A(t), t \in \mathbb{R}_+\}$ by

$$\hat{X}(t) := \begin{cases} \hat{X}_n, & \hat{S}_n \leqslant t < \hat{S}_{n+1}; \\ \partial_S, & t \geqslant \lim_{n\to\infty} \hat{S}_n; \end{cases} \qquad A(t) := \begin{cases} A_n, \hat{S}_n \leqslant t < \hat{S}_{n+1}; \\ \partial_A, t \geqslant \lim_{n\to\infty} \hat{S}_n, \end{cases}$$

where ∂_S and ∂_A are the extra state and action jointed to $\hat{E}$ and A, respectively. Now, given $(s, x) \in \mathbb{R} \times \hat{E}$ and $\pi = \{\pi_n\} \in \Pi$, by the Ionescu Tulcea theorem (see [3, Proposition B.2.5]), there exists a unique probability measure $\hat{\mathbb{P}}^\pi_{(s,x)}$ on $(\hat{\Omega}, \hat{\mathscr{F}})$ such that

$$\hat{\mathbb{P}}^\pi_{(s,x)}(\hat{T}_0 = 0, \hat{X}_0 = x) = 1, \tag{7}$$

$$\hat{\mathbb{P}}^\pi_{(s,x)}(A_n = a|\hat{Y}_n) = \pi_n(a|(s - \hat{S}_n), \hat{Y}_n), \tag{8}$$

$$\hat{\mathbb{P}}^\pi_{(s,x)}(\hat{T}_{n+1} \leqslant t, \hat{X}_{n+1} \in B|\hat{Y}_n, A_n) = \hat{Q}(t, B|\hat{X}_n, A_n), \tag{9}$$

where $\hat{Y}_n := (\hat{X}_0, A_0, \hat{T}_1, \hat{X}_1, \ldots, A_{n-1}, \hat{T}_n, \hat{X}_n)$, $t \in \mathbb{R}_+$, $B \in \mathcal{B}(\hat{E})$, $a \in A$ and $n \geqslant 0$. The expectation operator with respect to $\hat{\mathbb{P}}^\pi_{(s,x)}$ is denoted by $\hat{\mathbb{E}}^\pi_{(s,x)}$. Recall that we fix a planning horizon T in Section 2. To treat T-horizon optimization problem, naturally, we assume a finite number of jumps until time T. Thus, we propose Assumption 2.

Assumption 2. *For all* $(s, x) \in \mathbb{R} \times \hat{E}$ *and* $\pi \in \Pi$, $\hat{\mathbb{P}}^\pi_{(s,x)}(\{\lim_{n\to\infty} \hat{S}_n = \infty\}) = 1$.

Given $(s, x) \in \mathbb{R}_+ \times \hat{E}$, we define the expected cost of a policy $\pi \in \Pi$ by

$$U^\pi(s, x) := \hat{\mathbb{E}}^\pi_{(s,x)} \left[\int_0^s \hat{c}(\hat{X}(t), A(t)) \mathfrak{t} + \hat{g}(\hat{X}(s), A(s)) \right]$$

and the value function (or minimum expected cost) by $U^*(s, x) := \inf_{\pi \in \Pi} U^\pi(s, x)$.

Definition 3.3. A policy $\pi^* \in \Pi$ is called T-horizon optimal if

$$U^{\pi^*}(s, x) = U^*(s, x), \quad \forall (s, x) \in [0, T] \times \hat{E}.$$

Remark 3.3. Noting that the stopping time τ^* is T-optimal if and only if it achieves V^* for the fixed planning horizon T, while if the policy $\pi^* \in \Pi$ is T-optimal, π^* needs to achieve the value function of SMDP for all $s \in [0, T]$.

The focus of this section is on finding an optimal policy in Π, which can deduce the optimal stopping time. The following result reveals that it suffices to seek for optimal policies in Π_{RM}.

Proposition 3.1. *Suppose that Assumption 2 holds. Then for each $\pi = \{\pi_n\} \in \Pi$ and $(s, x) \in \mathbb{R}_+ \times \hat{E}$, there exists a policy $\hat{\pi} = \{\varphi_n\} \in \Pi_{RM}$ such that $U^\pi(s, x) = U^{\hat{\pi}}(s, x)$.*

Proof. Under Assumption 2, the monotone convergence theorem gives that

$$U^\pi(s, x) = \sum_{m=0}^\infty \hat{\mathbb{E}}^\pi_{(s,x)} \left[((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \hat{c}(\hat{X}_m, A_m) + \mathbb{1}_{[\hat{S}_m, \hat{S}_m + \hat{T}_{m+1})}(s) \hat{g}(\hat{X}_m, A_m) \right]. \quad (10)$$

Hence, it suffices to show that there is a policy $\hat{\pi} = \{\varphi_n\} \in \Pi_{RM}$ such that

$$\hat{\mathbb{P}}^\pi_{(s,x)}(\hat{X}_n \in B, A_n = a, \hat{S}_n \leqslant t, \hat{T}_{n+1} \leqslant v) = \hat{\mathbb{P}}^{\hat{\pi}}_{(s,x)}(\hat{X}_n \in B, A_n = a, \hat{S}_n \leqslant t, \hat{T}_{n+1} \leqslant v),$$

for $n = 0, 1, \ldots, t, v \in \mathbb{R}_+$, $B \in \mathcal{B}(\hat{E})$, and $a \in A$. Moreover, noting that (9) implies

$$\hat{\mathbb{E}}^\pi_{(s,x)} \left[\mathbb{1}_{[0,v]}(\hat{T}_{n+1}) | \hat{X}_n, A_n, \hat{S}_n \right] = \hat{Q}(v, \hat{E} | \hat{X}_n, A_n),$$

we need only to prove that

$$\hat{\mathbb{P}}^\pi_{(s,x)}(\hat{X}_n \in B, A_n = a, \hat{S}_n \leqslant t) = \hat{\mathbb{P}}^{\hat{\pi}}_{(s,x)}(\hat{X}_n \in B, A_n = a, \hat{S}_n \leqslant t). \quad (11)$$

Indeed, fix $(s, x) \in \mathbb{R}_+ \times \hat{E}$, and define a randomized Markov policy $\hat{\pi} := \{\varphi_n\}$ (depending on (s, x)) by

$$\varphi_n(a|t, y) := \begin{cases} \hat{\mathbb{E}}^\pi_{(s,x)}[\mathbb{1}_{\{a\}}(A_n) | \hat{S}_n = s - t, \hat{X}_n = y], & t \leqslant s, y \in \hat{E}; \\ \frac{1}{|A(y)|}, & t > s, y \in \hat{E}; \end{cases} \quad (12)$$

where $|A(y)|$ is the cardinality of $A(y)$. We show by induction that (11) holds with $\hat{\pi}$ defined through (12). Clearly it holds with $n = 0$. Assume that (11) holds for some n ($n \geq 0$). Then,

$$
\begin{aligned}
&\hat{\mathbb{P}}^{\pi}_{(s,x)}(\hat{X}_{n+1} \in B, \hat{S}_{n+1} \leq t)\\
=&\hat{\mathbb{E}}^{\pi}_{(s,x)}[\hat{\mathbb{E}}^{\pi}_{(s,x)}[\mathbb{1}_B(\hat{X}_{n+1})\mathbb{1}_{[0,t]}(\hat{S}_{n+1})|\hat{X}_n, A_n, \hat{S}_n]]\\
=&\hat{\mathbb{E}}^{\pi}_{(s,x)}[\hat{Q}((t - \hat{S}_n)^+, B|\hat{X}_n, A_n)] \quad \text{(by (9))}\\
=&\hat{\mathbb{P}}^{\hat{\pi}}_{(s,x)}(\hat{X}_{n+1} \in B, \hat{S}_{n+1} \leq t). \quad \text{(by the induction hypothesis)} \qquad (13)
\end{aligned}
$$

Therefore, the definition of $\hat{\pi}$ and the above equality give that

$$
\begin{aligned}
&\hat{\mathbb{P}}^{\pi}_{(s,x)}(\hat{X}_{n+1} \in B, \hat{S}_{n+1} \leq t, A_{n+1} = a)\\
=&\hat{\mathbb{E}}^{\pi}_{(s,x)}[\mathbb{1}_B(\hat{X}_{n+1})\mathbb{1}_{[0,t]}(\hat{S}_{n+1})\hat{\mathbb{E}}^{\pi}_{(s,x)}[\mathbb{1}_{\{a\}}(A_{n+1})|\hat{S}_{n+1}, \hat{X}_{n+1}]]\\
=&\hat{\mathbb{E}}^{\pi}_{(s,x)}[\mathbb{1}_B(\hat{X}_{n+1})\mathbb{1}_{[0,t]}(\hat{S}_{n+1})\varphi_n(a|s - \hat{S}_{n+1}, \hat{X}_{n+1})] \quad \text{(by (12))}\\
=&\hat{\mathbb{E}}^{\pi}_{(s,x)}[\mathbb{1}_B(\hat{X}_{n+1})\mathbb{1}_{[0,t]}(\hat{S}_{n+1})\hat{\mathbb{E}}^{\hat{\pi}}_{(s,x)}[\mathbb{1}_{\{a\}}(A_{n+1})|\hat{S}_{n+1}, \hat{X}_{n+1}]] \quad \text{(by (8))}\\
=&\hat{\mathbb{P}}^{\hat{\pi}}_{(s,x)}(\hat{X}_{n+1} \in B, \hat{S}_{n+1} \leq t, A_{n+1} = a). \quad \text{(by(13))}
\end{aligned}
$$

Thus the induction hypothesis is satisfied and the proof is completed.   □

Due to Proposition 3.1, we limit our discussion to randomized Markov policies in the rest of this section. Next, we establish our main results about the SMDPs. That is, we prove that the value function is a minimum nonnegative solution to the optimality equation and that there exists an optimal deterministic stationary policy. Also, we derive an algorithm for computing optimal policies and the value function.

Let $\mathbb{M}$ be the set of Borel measurable functions $v : [0, T] \times \hat{E} \to \bar{\mathbb{R}}_+ := [0, \infty]$. Given any $a \in A$, define an operator $\mathbb{T}^a$ from $\mathbb{M}$ to itself as: for each $v \in \mathbb{M}$ and $x \in \hat{E}$ if $a \notin A(x)$, $\mathbb{T}^a v(s, x) := +\infty$, otherwise

$$
\begin{aligned}
\mathbb{T}^a v(s, x) :=&\hat{c}(x, a) \int_0^s (1 - \hat{Q}(t, \hat{E}|x, a)\mathrm{t} + \hat{g}(x, a)(1 - \hat{Q}(s, \hat{E}|x, a))\\
&+ \int_{[0,s]} \int_{\hat{E}} v(s - t, y)\hat{Q}(\mathrm{t}, \mathrm{y}|x, a).
\end{aligned}
$$

Moreover, any $v \in \mathbb{M}$, $\varphi \in \Phi$, and $(s, x) \in [0, T] \times \hat{E}$, let

$$
\mathbb{T}^{\varphi} v(s, x) := \sum_{a \in A(x)} \varphi(a|s, x)\mathbb{T}^a v(s, x), \quad \text{and} \quad \mathbb{T} v(s, x) := \min_{a \in A(x)} \mathbb{T}^a v(s, x).
$$

To establish the iteration algorithm for computing U^π and U^*, we define a function sequence $\{U_n^\pi\}$ as following:

$$U_{-1}^\pi(s, x) := 0,$$

$$U_n^\pi(s, x) := \sum_{m=0}^{n} \hat{\mathbb{E}}_{(s,x)}^\pi \left[((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \hat{c}(\hat{X}_m, A_m) + \mathbb{1}_{[\hat{S}_m, \hat{S}_{m+1})}(s) \hat{g}(\hat{X}_m, A_m) \right]$$

for every $(s, x) \in [0, T] \times \hat{E}$ and $n \geqslant 0$. Clearly, $U_n^\pi(s, x) \leqslant U_{n+1}^\pi(s, x)$ for every $n \geqslant -1$, and moreover, it follows from (10) that $\lim_{n \to \infty} U_n^\pi(s, x) = U^\pi(s, x)$.

The following lemma is basic to our results.

Lemma 3.1. *Suppose that Assumption 2 holds. Let $\pi = \{\varphi_n, n \geqslant 0\} \in \Pi_{RM}$ be arbitrary.*

(a) *For each $n \geqslant -1$, $U_{n+1}^\pi = \mathbb{T}^{\varphi_0} U_n^{(1)\pi}$ and $U^\pi = \mathbb{T}^{\varphi_0} U^{(1)\pi}$, where $^{(1)}\pi = \{\varphi_n, n \geqslant 1\}$.*

(b) *In particular, for each $\varphi \in \Phi$, $U_{n+1}^\varphi = \mathbb{T}^\varphi U_n^\varphi$ and $U^\varphi = \mathbb{T}^\varphi U^\varphi$.*

Proof. (a) First, using that π is Markovian and (7)-(9), we have

$$\sum_{m=1}^{n+1} \hat{\mathbb{E}}_{(s,x)}^\pi \left[((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \hat{c}(\hat{X}_m, A_m) + \mathbb{1}_{[\hat{S}_m, \hat{S}_{m+1})}(s) \hat{g}(\hat{X}_m, A_m) \right]$$

$$= \sum_{m=1}^{n+1} \sum_{a \in A(x)} \varphi_0(a|s, x) \int_{\mathbb{R}_+} \int_{\hat{E}} \hat{Q}(t, y|x, a) \left[\hat{\mathbb{E}}_{(s,x)}^\pi \left[((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \hat{c}(\hat{X}_m, A_m) \right. \right.$$

$$\left. \left. + \mathbb{1}_{[\hat{S}_m, \hat{S}_{m+1})}(s) \hat{g}(\hat{X}_m, A_m) \Big| \hat{X}_1 = y, \hat{T}_1 = t \right] \right]$$

$$= \sum_{a \in A(x)} \varphi_0(a|s, x) \int_{\mathbb{R}_+} \int_{\hat{E}} \hat{Q}(t, y|x, a) \sum_{m=1}^{n+1} \hat{\mathbb{E}}_{((s-t),y)}^{(1)\pi} \left[((s - t) - \hat{S}_{m-1})^+ \wedge \hat{T}_m) \right.$$

$$\left. \times \hat{c}(\hat{X}_{m-1}, A_{m-1}) + \mathbb{1}_{[\hat{S}_{m-1}, \hat{S}_m)}(s - t) \hat{g}(\hat{X}_{m-1}, A_{m-1}) \right]$$

$$= \sum_{a \in A(x)} \varphi_0(a|s, x) \int_{[0,s]} \int_{\hat{E}} U_n^{(1)\pi}(s - t, y) Q(t, y|x, a),$$

where the last equality is due to that if $t > s$, $\mathbb{1}_{[\hat{S}_{m-1}, \hat{S}_m)}(s - t) = 0$ and $((s - t) - \hat{S}_{m-1})^+ = 0$. Then, we have

$$U_{n+1}^\pi(s, x) = \sum_{a \in A(x)} \varphi_0(a|s, x) \left[\hat{c}(x, a) \int_{\mathbb{R}_+} (t \wedge s) \hat{Q}(t, \hat{E}|x, a) \right.$$

$$\left. + \hat{g}(x, a) \int_{\mathbb{R}_+} \mathbb{1}_{[0,t)}(s) \hat{Q}(t, \hat{E}|x, a) + \int_{[0,s]} \int_{\hat{E}} U_n^{(1)\pi}(s - t, y) \hat{Q}(t, y|x, a) \right]$$

$$= \sum_{a \in A(x)} \varphi_0(a|s, x) \left[\hat{c}(x, a) \int_0^s (1 - \hat{Q}(t, \hat{E}|x, a)) t + \hat{g}(x, a)(1 - \hat{Q}(s, \hat{E}|x, a)) \right.$$

$$+ \int_{[0,s]} \int_{\hat{E}} U_n^{(1)\pi}(s - t, y) \hat{Q}(\mathfrak{t}, \mathfrak{y}|x, a) \Big]$$

$$= \mathbb{T}^{\varphi_0} U_n^{(1)\pi}(s, x).$$

Further, noting that $A(x)$ is finite, under Assumption 2, the monotone convergence theorem implies $U^\pi = \mathbb{T}^{\varphi_0} U^{(1)\pi}$.

(b) It immediately follows from part (a).   □

At the end of this section, we state our main results about SMDPs with the terminal cost function. In detail, we provide an iterative algorithm for computing the value function U^*, and give the optimality equation and the existence of optimal policies.

Theorem 3.1. *Suppose that Assumption 2 holds. Then the following statements hold.*

(a) *(Value iteration) For every $n \geq 0$, let $U_0^* =: 0$ and $U_{n+1}^* := \mathbb{T} U_n^*$. Then, $U^* = \lim_{n\to\infty} U_n^* \in \mathbb{M}$.*

(b) *(Optimality equation) U^* is the minimum solution in $\mathbb{M}$ to the optimality equation $U^* = \mathbb{T} U^*$, that is, if $u \in \mathbb{M}$ satisfies that $u = \mathbb{T}u$, then $u \geq U^*$.*

(c) *(Optimal policy) There exists an $f^* \in \mathbb{F}$ such that $U^* = \mathbb{T}^{f^*} U^*$, and such a policy $f^* \in \mathbb{F}$ is T-horizon optimal.*

Proof. (a) Since $\hat{c}(x, a)$ and $\hat{g}(x, a)$ are nonnegative and $\mathbb{T}$ is a nondecreasing map from $\mathbb{M}$ to $\mathbb{M}$, we obtain $U_{n+1}^*(s, x) \geq U_n^*(s, x)$ and $U_n^* \in \mathbb{M}$ for all $n \geq 0$ by $U_0^* = 0$. Therefore, $u^* := \lim_{n\to\infty} U_n^* \in \mathbb{M}$. To prove part (a), it remains to establish that $u^* = U^*$. We show $u^* \leq U^*$ and $U^* \geq u^*$, respectively.

To show $u^* \leq U^*$, we prove that

$$U_{n+1}^* \leq U_n^\pi, \quad \forall n \geq -1, \pi \in \Pi_{RM}, \tag{14}$$

and do this by induction. It is obviously true for $n \geq -1$. Suppose that $U_{n+1}^* \leq U_n^\pi$ for some $n \geq -1$ and any $\pi \in \Pi_{RM}$. Then, fixed any $\pi = \{\varphi_n, n \geq 0\} \in \Pi_{RM}$, by Lemma 3.1 part (a), it holds that

$$U_{n+1}^\pi = \mathbb{T}^{\varphi_0} U_n^{(1)\pi} \geq \mathbb{T}^{\varphi_0} U_{n+1}^* \geq \mathbb{T} U_{n+1}^* = U_{n+2}^*,$$

where $^{(1)}\pi = \{\varphi_n, n \geq 1\} \in \Pi_{RM}$ and the second and third inequalities follow from inductive hypothesis and the definitions of $\mathbb{T}$ and $\mathbb{T}^{\varphi_0}$, respectively. Hence (14) holds, and thus $u^* \leq U^*$.

We now show that $u^* \geqslant U^*$. For every fixed $(s, x) \in [0, T] \times \hat{E}$, since $A(x)$ is finite, there exists an $a_n^*(s, x)$ satisfying that $\mathbb{T}^{a_n^*(s,x)} U_n^*(s, x) = \mathbb{T} U_n^*(s, x) = U_{n+1}^*(s, x)$. Using that $A(x)$ is finite again, there is a subsequence $\{n_k\}$ of $\{n\}$ and $a^*(s, x) \in A(x)$ such that $a_{n_k}^*(s, x) = a^*(s, x)$ for all n_k. Hence, $U_{n_k+1}^*(s, x) = \mathbb{T}^{a^*(s,x)} U_{n_k}^*(s, x)$. Letting $n_k \to \infty$, it holds that $u^*(s, x) = \mathbb{T}^{a^*(s,x)} u^*(s, x)$ by monotone convergence theorem, which implies that $u^*(s, x) \geqslant \mathbb{T} u^*(s, x)$. By the arbitrariness of (s, x), we have $u^* \geqslant \mathbb{T} u^*$. On the other hand, the finiteness of $A(x)$ and measurable selection theorem (see [11, Proposition D.5]) ensure that there is an $f^* \in \mathbb{F}$ such that

$$\mathbb{T}^{f^*} u^* = \mathbb{T} u^* \leqslant u^*.$$

Moreover, Since $u^* \geqslant 0 = U_{-1}^{f^*}$, by induction it holds that $u^* \geqslant \mathbb{T}^{f^*} U_n^{f^*} = U_{n+1}^{f^*}$ for all $n \geqslant -1$, which implies

$$u^* \geqslant \lim_{n \to \infty} U_n^{f^*} = U^{f^*} \geqslant U^* \geqslant u^* \tag{15}$$

Therefore, $u^* = U^*$.

(b) For every $\pi = \{\varphi_n, n \geqslant 0\} \in \Pi_{RM}$, by Lemma 3.1 (a), it holds that

$$U^\pi = \mathbb{T}^{\varphi_0} U^{(1)\pi} \geqslant \mathbb{T}^{\varphi_0} U^* \geqslant \mathbb{T} U^*.$$

Then, the arbitrariness of π implies that $U^* \geqslant \mathbb{T} U^*$. On the other hand,

$$U_{n+1}^*(s, x) = \mathbb{T} U_n^*(s, x) \leqslant \mathbb{T}^a U_n^*(s, x), \quad \forall (s, x) \in [0, T] \times \hat{E}, a \in A(x), n \geqslant 0.$$

Hence, by the monotone convergence theorem, we obtain that $U^*(s, x) \leqslant \mathbb{T}^a U^*(s, x)$, and so $U^*(s, x) \leqslant \mathbb{T} U^*(s, x)$. Therefore, $U^* = \mathbb{T} U^*$.

Let $u \in \mathbb{M}$ be an arbitrary solution to the equation $u = \mathbb{T} u$. Since $u \geqslant 0 = U_0^*$ and $u = \mathbb{T}^n u$, it follows from part (a) that

$$u = \lim_{n \to \infty} \mathbb{T}^{n+1} u \geqslant \lim_{n \to \infty} \mathbb{T}^{n+1} U_0^* = \lim_{n \to \infty} U_n^* = U^*,$$

where $T^1 v = Tv$ and $T^n v = T(T^{n-1} v)$ for all $n \geqslant 1$ and $v \in \mathbb{M}$. This means that U^* is the minimum solution in $\mathbb{M}$ to the optimality equation.

(c) By the proof of part (a), there is an $f^* \in \mathbb{F}$ such that $\mathbb{T} U^* = \mathbb{T}^{f^*} U^*$. Therefore, part (b) gives $U^* = \mathbb{T}^{f^*} U^*$. Hence, f^* is T-horizon optimal by part (a) and (15). □

Remark 3.4. In particular, if $\hat{E}$ is denumerable and $\hat{g} = 0$, the above results are same as [12, Theorem 3.1 and Theorem 3.2].

4. Existence and computation of optimal stopping times

In this section, we introduce the equivalent SMDPs corresponding to the original optimal stopping problem of SMPs in section 2. And then, we show that for every stopping time τ and $s \in [0, T]$, there is a policy π_τ such that the s-horizon expected cost of τ is equal to s-horizon expected cost of the policy π_τ. Hence, we can analyze the value function V^* and the optimal stopping time τ^* of SMPs through the conclusions of SMDPs given in section 3. Note that the regular condition (Assumption 1) is needed.

Intuitively, in the SMPs, "continue" or "stop" can be considered as a special action in the corresponding SMDPs. This intuition gives us an idea to construct the SMDPs. The details about the constructions of SMDPs are given as follows. Hence, the model of the corresponding SMDPs is

$$\left\{ \hat{E}, A, (A(x) \subset A, x \in \hat{E}), \hat{Q}_T(\cdot, \cdot | x, a), \hat{c}(x, a), \hat{g}(x, a) \right\}. \tag{16}$$

where the state space $\hat{E} := E \cup \{\Delta\}$ includes the state space E of SMPs and a virtual state Δ. $A(x)$, denoting the set of admissible actions at state $x \in \hat{E}$, is defined as

$$A(x) := \begin{cases} \{0, 1\}, & x \in E; \\ \{1\}, & x = \Delta, \end{cases}$$

where the action 0 means continuation and 1 means stop. The action space $A = \{0, 1\}$ is finite. Then the set of admissible state-action pairs $K = (E \times A) \cup \{(\Delta, 1)\}$ is a Borel subset of $\hat{E} \times A$. For each $t \geqslant 0$ and $B \in \mathcal{B}(\hat{E})$, the semi-Markov kernel $\hat{Q}_T(t, B|x, a)$ is given by

$$\hat{Q}_T(t, B|x, a) := \begin{cases} Q(t, B \setminus \{\Delta\}|x), & x \in E, a = 0; \\ \mathbb{1}_{[T+1, +\infty)}(t)\delta_\Delta(B), & x \in \hat{E}, a = 1, \end{cases} \tag{17}$$

where $Q(t, B|x)$ is the kernel of SMPs given in (1). Finally, the cost rate function and terminal cost function of SMDP are defined as

$$\hat{c}(x, a) := \begin{cases} c(x), & x \in E, a = 0; \\ 0, & \text{otherwise}, \end{cases} \tag{18}$$

$$\hat{g}(x, a) := \begin{cases} g(x), & x \in E, a = 1; \\ 0, & \text{otherwise}, \end{cases} \tag{19}$$

where c and g are the cost rate function and terminal cost function of SMPs, respectively. Since c and g are measurable on E, $\hat{c}(x, a)$ and $\hat{g}(x, a)$ are measurable on K. Firstly, we give a lemma to show that the above model satisfies Assumption 2.

Lemma 4.1. *Suppose that Assumption 1 holds. For the SMDPs as in (16), Assumption 2 is fulfilled.*

Proof. Under Assumption 1, there exist $\delta > 0$ and $\epsilon > 0$ such that (2) holds. Let $\hat{\delta} = \min\{\delta, \frac{1}{2}\}$, and thus

$$\hat{Q}_T(\hat{\delta}, \hat{E}|x, a) = \begin{cases} Q(\hat{\delta}, E|x) \leqslant Q(\delta, E|x) \leqslant 1 - \epsilon, & x \in E, a = 0; \\ \mathbb{1}_{[T+1,+\infty)}(\hat{\delta}) = 0 \leqslant 1 - \epsilon, & i \in \hat{E}, a = 1. \end{cases}$$

Hence, Assumption 2 is fulfilled by [12, Proposition 2.1]. $\square$

Next step, we will show the relationship between the stopping times $\tau \in \Gamma$ of SMPs and the policies $\pi \in \Pi_{DH}$ of SMDPs as in (16). To do so, for all $n \geqslant 0$ and history of SMPs up to the n-th jump epoch $h_n = (x_0, t_1, \ldots, x_{n-1}, t_n, x_n) \in H_n$, let

$$h_n^0 = (x_0, 0, t_1, x_1, \ldots, 0, t_n, x_n) \in \hat{H}_n. \tag{20}$$

The action 0 (means continuation) added to the equation (20) indicates that the system has been running incessantly before the n-th jump epoch. Obviously, by the definition of h_n^0, it holds that

$$C^0 := \{h_n^0 \,|\, h_n \in C\} \in \mathcal{B}(\hat{H}_n), \quad \forall C \in \mathcal{B}(H_n). \tag{21}$$

In particular, $H_n^0 = E \times (\{0\} \times \mathbb{R}_+ \times E)^n \in \mathcal{B}(\hat{H}_n)$.

Definition 4.1. Given any deterministic policy $\pi = \{d_n, n \geqslant 0\} \in \Pi_{DH}$ defined in Definition 3.2 (c) and $s \in \mathbb{R}$, define

$$\tau_\pi^s(\omega) := \inf \left\{ n \in \mathbb{N} \,\middle|\, d_n(s - S_n(\omega), Y_n^0(\omega)) = 1 \right\}, \quad \omega = (x_0, t_1, \ldots, x_n, t_{n+1}, \ldots) \in \Omega,$$

where $\inf\{\emptyset\} := +\infty$ and $Y_n^0 := (X_0, 0, T_1, X_1, \ldots, 0, T_n, X_n)$. Then τ_π^s is called the stopping time induced by the policy π and s.

Lemma 4.2. *For each deterministic policy $\pi = \{d_n, n \geqslant 0\}$ and $s \in \mathbb{R}$, the induced stopping time τ_π^s is a stopping time.*

Proof. Note that for each $n \geqslant 0$, the random variables Y_n^0 and S_n, and the function d_n are measurable in their corresponding spaces. Hence, we have

$$\{\tau_\pi^s = n\} = \left(\cap_{k=0}^{n-1} \{d_k((s - S_k), Y_k^0) = 0\} \right) \cap \{d_n((s - S_n), Y_n^0) = 1\} \in \sigma(Y_n) = \mathscr{F}_n,$$

which implies that τ_π^s is a stopping time. $\square$

We introduce a subclass Π_{DH}^0 of Π_{DH} by

$$\Pi_{DH}^0 := \{\pi = \{d_n, n \geqslant 0\} \in \Pi_{DH} \mid d_n(0, \hat{h}_n) = 0, \forall n \geqslant 0, \hat{h}_n \in H_n^0\}.$$

We now give a key theorem which establishes the relationship between the s-horizon expected cost of polices in Π_{DH}^0 and that of the stopping times induced by the polices.

Theorem 4.1. *Suppose that Assumption 1 holds. For any* $\pi = \{d_n, n \geqslant 0\} \in \Pi_{DH}^0$, *it holds that*

$$U^\pi(s, x) = V^{\tau_\pi^s}(s, x), \quad \forall x \in E, s \in [0, T],$$

where τ_π^s *is the stopping time induced by* π *and* s, *and* $V^{\tau_\pi^s}(s, x)$ *is the s-horizon expected cost of* τ_π^s.

Proof. For each $\hat{\omega} = (x_0, a_0, t_1, \ldots, x_n, a_n, t_{n+1}, \ldots) \in \hat{\Omega}$, recall that $\hat{S}_n(\omega) = \sum_{k=1}^n t_k$ and

$$\hat{Y}_n(\hat{\omega}) = (x_0, a_0, t_1, x_1 \ldots, a_{n-1}, t_n, x_n).$$

We define $C_n (n \geqslant 0)$ and C, the subsets of $\hat{\Omega}$, as

$$C_n := \left\{ \omega \in \hat{\Omega} : \inf \left\{ k \in \mathbb{N} : d_k(s - \hat{S}_k(\omega), \hat{Y}_k(\omega)) = 1 \right\} = n \right\}, \quad n \geqslant 0;$$

$$C := \left\{ \omega \in \hat{\Omega} : d_k(s - \hat{S}_k(\omega), \hat{Y}_k(\omega)) = 0, \forall k \geqslant 0 \right\}.$$

It is easy to know that $\{C, C_n, n \geqslant 0\}$ is a partition of $\hat{\Omega}$ and

$$\mathbb{1}_{C_n} = \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - \hat{S}_k, \hat{Y}_k)) \times \mathbb{1}_{\{1\}}(d_n(s - \hat{S}_n, \hat{Y}_n)).$$

Hence, the monotone convergence theorem implies that

$$U^\pi(s, x) = \sum_{n=0}^{\infty} \hat{\mathbb{E}}_{(s,x)}^\pi \left[\mathbb{1}_{C_n} \left(\int_0^s \hat{c}(\hat{X}(t), A(t)) t + \hat{g}(\hat{X}(s), A(s)) \right) \right]$$

$$+ \hat{\mathbb{E}}_{(s,x)}^\pi \left[\mathbb{1}_C \left(\int_0^s \hat{c}(\hat{X}(t), A(t)) t + \hat{g}(\hat{X}(s), A(s)) \right) \right]. \tag{22}$$

Noting that π is a deterministic, the definition of C_n and (8) give

$$\hat{\mathbb{P}}_{(s,x)}^\pi [A_k = 1 | C_n] = \hat{\mathbb{P}}_{(s,x)}^\pi \left[d_k(s - \hat{S}_k, \hat{Y}_k) = 1 \big| C_n \right] = \begin{cases} 0, & k < n; \\ 1, & k = n, \end{cases} \tag{23}$$

which, together with $\lim_{t \to \infty} \hat{Q}_T(t, \Delta | x, 1) = 1$ and $A(\Delta) = \{1\}$, implies that $\hat{\mathbb{P}}_{(s,x)}^\pi(\hat{X}_m = \Delta | C_n) = 1$ for all $m > n$. Thus, by Lemma 4.1, (10), (18) and (19) give

$$\hat{\mathbb{E}}_{(s,x)}^\pi \left[\mathbb{1}_{C_n} \left(\int_0^s \hat{c}(\hat{X}(t), A(t)) t + \hat{g}(\hat{X}(s), A(s)) \right) \right]$$

$$= \sum_{m=0}^{n-1} \hat{\mathbb{E}}_{(s,x)}^{\pi} \left[\hat{c}(\hat{X}_m, A_m)((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \mathbb{1}_{C_n} \right] + \hat{\mathbb{E}}_{(s,x)}^{\pi} \left[\mathbb{1}_{C_n} \mathbb{1}_{[\hat{S}_n, \hat{S}_{n+1})}(s) \hat{g}(\hat{X}_n, 1) \right]. \tag{24}$$

And then, for each $m < n$, using

$$\mathbb{1}_{\{\tau_\pi^s = n\}} = \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - S_k, Y_k^0)) \cdot \mathbb{1}_{\{1\}}(d_n(s - S_n, Y_n^0)) \tag{25}$$

and $\hat{Q}_T(\cdot, B|x, 0) = Q(\cdot, B \setminus \{\Delta\}|x)$ for all $x \in E$ and $B \in \mathscr{B}(\hat{E})$, it holds that

$$\hat{\mathbb{E}}_{(s,x)}^{\pi} \left[\hat{c}(\hat{X}_m, A_m)((s - \hat{S}_m)^+ \wedge \hat{T}_{m+1}) \mathbb{1}_{C_n} \right]$$

$$= \int_{\hat{E}} \delta_x(x_0) \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_1, x_1|x_0, d_0(s, \hat{h}_0)) \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_2, x_2|x_1, d_0(s - s_1, \hat{h}_1))$$

$$\cdots \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_n, x_n|x_{n-1}, d_{n-1}(s - s_{n-1}, \hat{h}_{n-1})) \hat{c}(x_m, d_m(s - s_m, \hat{h}_m))$$

$$((s - s_m)^+ \wedge t_{m+1}) \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - s_k, \hat{h}_k)) \cdot \mathbb{1}_{\{1\}}(d_n(s - s_n, \hat{h}_n))$$

$$= \int_E \delta_x(x_0) \int_E \int_{\mathbb{R}_+} Q(t_1, x_1|x_0) \cdots \int_E \int_{\mathbb{R}_+} Q(t_n, x_n|x_{n-1}) c(x_m)$$

$$((s - s_m)^+ \wedge t_{m+1}) \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - s_k, h_k^0)) \cdot \mathbb{1}_{\{1\}}(d_n(s - s_n, h_n^0))$$

$$= \mathbb{E}_x \left[((s - S_m)^+ \wedge T_{m+1}) c(X_m) \mathbb{1}_{\{\tau_\pi^s = n\}} \right], \tag{26}$$

where $s_k = \sum_{i=1}^k t_i$, $\hat{h}_0 = x_0$, $\hat{h}_{k+1} = (\hat{h}_k, d_k(s - s_k, \hat{h}_k), t_{k+1}, x_{k+1})$, $h_k = (x_0, t_1, x_1, \ldots, t_k, x_k)$, and h_k^0 defined in (20). For each $\hat{\omega} \in C_n$, $\pi \in \Pi_{HD}^0$ gives $\hat{Y}_n(\hat{\omega}) \in H_n^0$, and then $d_n(0, \hat{Y}_n(\hat{\omega})) = 0$. Further, by $d_n(s - \hat{S}_n(\hat{\omega}), \hat{Y}_n(\hat{\omega})) = 1$ for all $\hat{\omega} \in C_n$, $C_n \cap \{s = \hat{S}_n\} = \emptyset$. Therefore, we have $\mathbb{1}_{C_n} \mathbb{1}_{[\hat{S}_n, \hat{S}_{n+1})}(s) = \mathbb{1}_{C_n} \mathbb{1}_{(\hat{S}_n, \hat{S}_{n+1})}(s)$. Thus, (17) and (19) show

$$\hat{\mathbb{E}}_{(s,x)}^{\pi} \left[\mathbb{1}_{C_n} \mathbb{1}_{[\hat{S}_n, \hat{S}_{n+1})}(s) \hat{g}(\hat{X}_n, 1) \right]$$

$$= \int_{\hat{E}} \delta_x(x_0) \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_1, x_1|x_0, d_0(s, \hat{h}_0)) \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_2, x_2|x_1, d_0(s - s_1, \hat{h}_1))$$

$$\cdots \int_{\hat{E}} \int_{\mathbb{R}_+} \hat{Q}_T(t_{n+1}, x_{n+1}|x_n, d_n(s - s_n, \hat{h}_n)) \hat{g}(x_n, 1) \mathbb{1}_{(s_n, s_n + t_{n+1})}(s)$$

$$\times \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - s_k, \hat{h}_k)) \cdot \mathbb{1}_{\{1\}}(d_n(s - s_n, \hat{h}_n))$$

$$= \int_E \delta_x(x_0) \int_E \int_{\mathbb{R}_+} Q(t_1, x_1|x_0) \cdots \int_E \int_{\mathbb{R}_+} Q(t_n, x_n|x_{n-1}) g(x_n) \mathbb{1}_{(s_n, s_n + T + 1)}(s)$$

$$\times \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k(s - s_k, h_k^0)) \cdot \mathbb{1}_{\{1\}}(d_n(s - s_n, h_n^0)) \quad \text{(by } \hat{Q}_T(t, \hat{E}|x_n, 1) = \mathbb{1}_{[T+1, \infty)}(t))$$

$$=\mathbb{E}_x\left[\mathbb{1}_{(S_n,+\infty)}(s)g(X_n)\mathbb{1}_{\{\tau_\pi^s=n\}}\right]. \tag{27}$$

Moreover, by the definition of $R_{\tau_\pi^s}^s$ given in (3), (24), (26), and (27), we have

$$\mathbb{E}_x\left[\mathbb{1}_{\{\tau_\pi^s=n\}}R_{\tau_\pi^s}^s\right]$$

$$=\mathbb{E}_x\left[\mathbb{1}_{\{S_n<s\}}\mathbb{1}_{\{\tau_\pi^s=n\}}\left(\int_0^{S_n}c(X(t))\mathfrak{t}+g(X(S_n))\right)\right]+\mathbb{E}_x\left[\mathbb{1}_{\{S_n\geqslant s\}}\mathbb{1}_{\{\tau_\pi^s=n\}}\int_0^s c(X(t))\mathfrak{t}\right]$$

$$=\mathbb{E}_x\left[\mathbb{1}_{\{S_n<s\}}\mathbb{1}_{\{\tau_\pi^s=n\}}g(X_n)\right]+\mathbb{E}_x\left[\mathbb{1}_{\{S_n<s\}}\mathbb{1}_{\{\tau_\pi^s=n\}}\sum_{m=0}^{n-1}c(X_m)((s-S_m)^+\wedge T_{m+1})\right]$$

$$+\mathbb{E}_x\left[\mathbb{1}_{\{S_n\geqslant s\}}\mathbb{1}_{\{\tau_\pi^s=n\}}\sum_{m=0}^{n-1}c(X_m)((s-S_m)^+\wedge T_{m+1})\right]$$

$$=\hat{\mathbb{E}}_{(s,x)}^\pi\left[\mathbb{1}_{C_n}\left(\int_0^s\hat{c}(\hat{X}(t),A(t))\mathfrak{t}+\hat{g}(\hat{X}(s),A(s))\right)\right]. \tag{28}$$

Next, we calculate the second item of (22). Using that $\{C,C_n,n\geqslant 0\}$ is a partition of $\hat{\Omega}$ again, for all $k\geqslant 0$ we obtain that

$$\mathbb{1}_C=\mathbb{1}_C\times\prod_{m=0}^k(1-\mathbb{1}_{C_m})=(1-\sum_{n=0}^{+\infty}\mathbb{1}_{C_n})\times\prod_{m=0}^k(1-\mathbb{1}_{C_m})=\prod_{m=0}^k(1-\mathbb{1}_{C_m})-\sum_{n=k+1}^{+\infty}\mathbb{1}_{C_n},$$

which, combining with (10), implies that

$$\hat{\mathbb{E}}_{(s,x)}^\pi\left[\mathbb{1}_C\left(\int_0^s\hat{c}(\hat{X}(t),A(t))\mathfrak{t}+\hat{g}(\hat{X}(s),A(s))\right)\right]$$

$$=\sum_{k=0}^\infty\left\{\hat{\mathbb{E}}_{(s,x)}^\pi\left[\prod_{m=0}^k(1-\mathbb{1}_{C_m})\left(((s-\hat{S}_k)^+\wedge\hat{T}_{k+1})\hat{c}(\hat{X}_k,A_k)+\mathbb{1}_{[\hat{S}_k,\hat{S}_{k+1})}(s)\hat{g}(\hat{X}_k,A_k)\right)\right]\right.$$

$$\left.-\sum_{n=k+1}^{+\infty}\hat{\mathbb{E}}_{(s,x)}^\pi\left[\mathbb{1}_{C_n}\left(((s-\hat{S}_k)^+\wedge\hat{T}_{k+1})\hat{c}(\hat{X}_k,A_k)+\mathbb{1}_{[\hat{S}_k,\hat{S}_{k+1})}(s)\hat{g}(\hat{X}_k,A_k)\right)\right]\right\}.$$

Firstly, using that $\hat{g}(x,0)=0$ for all $x\in E$, (23) and (26), we obtain that for each $n>k$,

$$\hat{\mathbb{E}}_{(s,x)}^\pi\left[\mathbb{1}_{C_n}\left(((s-S_k)^+\wedge\hat{T}_{k+1})\hat{c}(\hat{X}_k,A_k)+\mathbb{1}_{[\hat{S}_k,\hat{S}_{k+1})}(s)\hat{g}(\hat{X}_k,A_k)\right)\right]$$

$$=\mathbb{E}_x\left[((s-S_m)^+\wedge T_{m+1})c(X_m)\mathbb{1}_{\{\tau_\pi^s=n\}}\right].$$

And for any $k\geqslant 0$, by the same methods of (26)

$$\hat{\mathbb{E}}_{(s,x)}^\pi\left[\prod_{m=0}^k(1-\mathbb{1}_{C_m})\left(((s-\hat{S}_k)^+\wedge\hat{T}_{k+1})\hat{c}(\hat{X}_k,A_k)+\mathbb{1}_{[\hat{S}_k,\hat{S}_{k+1})}(s)\hat{g}(\hat{X}_k,A_k)\right)\right]$$

$$=\mathbb{E}_x\left[\prod_{m=0}^k(1-\mathbb{1}_{\{\tau_\pi^s=m\}})((s-S_k)^+\wedge T_{k+1})c(X_k)\right].$$

Therefore, we obtain that

$$\mathbb{E}_x\left[\mathbb{1}_{\{\tau^s_\pi=+\infty\}}R^s_{\tau^s_\pi}\right] = \hat{\mathbb{E}}^\pi_{(s,x)}\left[\mathbb{1}_C\left(\int_0^s \hat{c}(\hat{X}(t),A(t))\text{t} + \hat{g}(X(s),A(s))\right)\right],$$

which, together with (22) and (28), shows that

$$U^\pi(s,x) = \sum_{n=0}^\infty \mathbb{E}_x\left[\mathbb{1}_{\{\tau^s_\pi=n\}}R^s_{\tau^s_\pi}\right] + \mathbb{E}_x\left[\mathbb{1}_{\{\tau^s_\pi=+\infty\}}R^s_{\tau^s_\pi}\right] = V^{\tau^s_\pi}(s,x), \forall x \in E, s \in [0,T].$$

The proof of Theorem 4.1 is completed. □

For any $s \in [0,T]$, the Definition 4.1, Lemma 4.2 and Theorem 4.1 say that for each policy $\pi \in \Pi^0_{DH}$, we can construct a stopping time τ^s_π such that their s-horizon expected costs are equal. On the other hand, for each stopping time we also can construct a policy which satisfies this condition, see Definition 4.2, Lemma 4.3 and Theorem 4.2.

Definition 4.2. Given any stopping time $\tau \in \Gamma$ and $n \geqslant 0$, let

$$B^\tau_n := \{Y_n(\omega) : \omega = (x_0,t_1,x_1,\ldots,t_k,x_k,\ldots) \in \{\tau = n\}\},$$

where $Y_n(\omega) = (x_0,t_1,x_1,\ldots,t_n,x_n)$. For each $\hat{h}_n = (x_0,a_0,t_1,x_1,\ldots,a_{n-1},t_n,x_n) \in \hat{H}_n$, $s \in \mathbb{R}$, define

$$d^\tau_n(s,\hat{h}_n) := \begin{cases} \mathbb{1}_{B^\tau_n}(x_0,t_1,x_1,\ldots,t_n,x_n)\mathbb{1}_{(0,\infty)}(s), & \hat{h}_n \in H^0_n; \\ 1, & \hat{h}_n \in \hat{H}_n \setminus H^0_n, \end{cases}$$

where $H^0_n = E \times (\{0\} \times \mathbb{R}_+ \times E)^n$. $\pi_\tau := \{d^\tau_n, n \geqslant 0\}$ is called the policy induced by τ.

Lemma 4.3. *For each stopping time $\tau \in \Gamma$, π_τ is in Π^0_{DH} on the corresponding SMDPs.*

Proof. By Definition 4.2, it can verify easily that $d^\tau_n(s,\hat{h}_n) \in A(x_n)$. Then, we just need to consider the measurability. Noting that $\{\tau = n\} \in \mathscr{F}_n = \sigma(Y_n)$, we obtain $B^\tau_n \in \mathcal{B}(H_n)$. Thus, we have

$$\{(s,\hat{h}_n) \in \mathbb{R} \times \hat{H}_n | d^\tau_n(s,\hat{h}_n) = 0\} = \left((0,+\infty) \times (H_n \setminus B^\tau_n)^0\right) \cup \left((-\infty,0] \times H^0_n\right) \in \mathcal{B}(\mathbb{R} \times \hat{H}_n),$$

where $(H_n \setminus B^\tau_n)^0$ is defined in (21). Hence, $\pi_\tau := \{d^\tau_n, n \geqslant 0\}$ is a deterministic policy of the corresponding SMDPs. Furthermore, for each $\hat{h}_n \in H^0_n$, it holds that $f^\tau_n(0,\hat{h}_n) = 0$, which implies that $\pi_\tau \in \Pi^0_{DH}$. □

Theorem 4.2. *Suppose that Assumption 1 holds. For each stopping time $\tau \in \Gamma$, let $\pi_\tau := \{d^\tau_n, n \geqslant 0\}$ be the policy induced by τ. Then,*

$$V^\tau(s,x) = U^{\pi_\tau}(s,x) = V^{\tau^s_{\pi_\tau}}(s,x) \quad \forall(s,x) \in [0,T] \times E, \tag{29}$$

where $\tau_{\pi_\tau}^s$ is the stopping time induced by π_τ and s.

Proof. To prove (29), by the Theorem 4.1, it suffices to show that $V^\tau(s, x) = V^{\tau_{\pi_\tau}^s}(s, x)$. By the (25) and the definition of d_n^τ, we have

$$\mathbb{1}_{\{\tau_{\pi_\tau}^s = n\}} = \prod_{k=0}^{n-1} \mathbb{1}_{\{0\}}(d_k^\tau(s - S_k, Y_k^0)) \cdot \mathbb{1}_{\{1\}}(d_n^\tau(s - S_n, Y_n^0)) = \mathbb{1}_{[0,s)}(S_n)\mathbb{1}_{\{\tau = n\}}, \qquad (30)$$

where $Y_k^0 = (X_0, 0, T_1, X_1, \ldots, 0, T_k, X_k)$. Moreover, using the definition of $\tau_{\pi_\tau}^s$ again,

$$\mathbb{1}_{\{\tau_{\pi_\tau}^s = \infty\}} = \prod_{n=0}^{\infty} \left(1 - \mathbb{1}_{B_n^\tau}(Y_n)\mathbb{1}_{(0,\infty)}(s - S_n) \right)$$

$$= \mathbb{1}_{\{\tau = \infty\}} \prod_{n=0}^{\infty} \left(1 - \mathbb{1}_{\{\tau = n\}}\mathbb{1}_{(0,\infty)}(s - S_n) \right) + \mathbb{1}_{\{\tau \neq \infty\}} \prod_{n=0}^{\infty} \left(1 - \mathbb{1}_{\{\tau = n\}}\mathbb{1}_{(0,\infty)}(s - S_n) \right)$$

$$= \mathbb{1}_{\{\tau = \infty\}} + \sum_{k=0}^{\infty} \mathbb{1}_{\{\tau = k\}}\mathbb{1}_{[s,\infty)}(S_k). \qquad (31)$$

According to the definition of R_τ^s given in (3), (30) and (31), we have that

$$V^{\tau_{\pi_\tau}^s}(s, x) = \sum_{n=0}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\{\tau_{\pi_\tau}^s = n\}} R_n^s \right] + \mathbb{E}_x \left[\mathbb{1}_{\{\tau_{\pi_\tau}^s = \infty\}} \int_0^s c(X(t))\mathrm{t} \right]$$

$$= \sum_{n=0}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{[0,s)}(S_n)\mathbb{1}_{\{\tau = n\}} R_n^s \right] + \mathbb{E}_x \left[\left(\mathbb{1}_{\{\tau = \infty\}} + \sum_{n=0}^{\infty} \mathbb{1}_{\{\tau = n\}}\mathbb{1}_{[s,\infty)}(S_n) \right) \int_0^s c(X(t))\mathrm{t} \right]$$

$$= \sum_{n=0}^{\infty} \left(\mathbb{E}_x \left[\mathbb{1}_{[0,s)}(S_n)\mathbb{1}_{\{\tau = n\}} R_n^s \right] + \mathbb{E}_x \left[\mathbb{1}_{\{\tau = n\}}\mathbb{1}_{[s,\infty)}(S_n) R_n^s \right] \right) + \mathbb{E}_x \left[\mathbb{1}_{\{\tau = \infty\}} R_\tau^s \right]$$

$$= V^\tau(s, x),$$

which is the desired result. □

Let $\mathcal{M}$ be the set of Borel measurable functions $u : [0, T] \times E \to [0, \infty]$. Next, we define the operator $\mathbb{G}$ from $\mathcal{M}$ into $\mathcal{M}$ as follows:

$$\mathbb{G}u(s, x) = \min \left\{ c(x) \int_0^s (1 - Q(t, E|x))\mathrm{t} + \int_E \int_{[0,s]} u(s - t, y)Q(\mathrm{t}, \mathrm{y}|x), g(x) \right\},$$

for each $u \in \mathcal{M}$ and $(s, x) \in [0, T] \times E$. Then, we can state our main results of optimal stopping problems, these are an algorithm for computing the value function V^* and a finite optimal stopping time.

Theorem 4.3. (Value iteration) *Suppose that Assumption 1 holds. For any $n \geqslant -1$, let $V_0^* \equiv 0$ and $V_{n+1}^* = \mathbb{G}V_n^*$. Then, we have $V^* = \lim_{n \to \infty} V_n^*$ and, moreover, $V^* = \mathbb{G}V^*$.*

Proof. We define the policy f^* for the SMDP as in (16) by

$$f^*(s, x) = \mathbb{1}_{(0,\infty) \times E}(s, x)\mathbb{1}_{\{U^*(s,x)=\hat{g}(x,1)\}}(s, x) + \mathbb{1}_{\{\Delta\}}(x). \tag{32}$$

Since $\{U^*(s, x) = g(x, 1)\}$ is a measurable subset of $\mathbb{R} \times \hat{E}$ and $f^*(\cdot, \Delta) = 1$, f^* is a deterministic stationary policy. Moreover, for any $x \in E$, we have that $f^*(0, x) = 0$, which implies $f^* \in \Pi^0_{DH}$. Next we will show that f^* is T-optimal.

(i) $x = \Delta$: Noting that $A(\Delta) = \{1\}$, by $U^*_0 \equiv 0$ and $\hat{c}(\Delta, 1) = \hat{g}(\Delta, 1) = 0$, it holds that

$$U^*(s, \Delta) = \mathbb{T}U^*(s, \Delta) = \mathbb{T}^{f^*(s,\Delta)}U^*(s, \Delta) = 0, \quad \forall s \in [0, T].$$

(ii) $s = 0$ and $x \in E$: By $f^*(0, x) = 0$, we have $U^*(0, x) = \min\{0, g(x)\} = \mathbb{T}^{f^*(0,x)}U^*(0, x)$.

(iii) $(s, x) \in ((0, T] \times E)$: Noting that $\mathbb{T}^1 U^*(s, x) = \hat{g}(x, 1)$, by $U^* = \mathbb{T}U^*$, we have

$$U^*(s, x) = \begin{cases} \mathbb{T}^1 U^*(s, x) = \mathbb{T}^{f^*(s,x)}U^*(s, x), & U^*(s, x) = \hat{g}(x, 1); \\ \mathbb{T}^0 U^*(s, x) = \mathbb{T}^{f^*(s,x)}U^*(s, x), & U^*(s, x) \neq \hat{g}(x, 1). \end{cases}$$

Hence, f^* is T-optimal by Theorem 3.1 part (c). Therefore, Theorem 4.1 implies that

$$U^*(s, x) = U^{f^*}(s, x) = V^{\tau^s_{f^*}}(s, x) \geqslant V^*(s, x), \quad \forall x \in E, s \in [0, T].$$

where $\tau^s_{f^*}$ is the stopping time induced by f^* and s. On the other hand, for each $\tau \in \Gamma$, Theorem 4.2 gives

$$V^\tau(s, x) = U^{\pi_\tau}(s, x) \geqslant U^*(s, x) \quad \forall x \in E, s \in [0, T].$$

By the arbitrariness of τ, $V^*(s, x) \geqslant U^*(s, x)$. Hence,

$$V^*(s, x) = U^*(s, x), \quad \forall x \in E, s \in [0, T]. \tag{33}$$

Next, we show that for all $n \geqslant 0$

$$U^*_n(s, x) = V^*_n(s, x), \quad \forall x \in E, s \in [0, T]. \tag{34}$$

Obviously, it holds that for $n = 0$. Assume that (34) holds for some n. Then

$$U^*_{n+1}(s, x) = \min\left\{\hat{g}(x, 1), \hat{c}(x, 0)\int_0^s (1 - \hat{Q}_T(t, \hat{E}|x, 0))\mathrm{t}\right.$$
$$\left. + \hat{g}(x, 0)(1 - \hat{Q}_T(s, \hat{E}|x, 0)) + \int_{[0,s]}\int_{\hat{E}} U^*_n(s - t, y)\hat{Q}_T(\mathrm{t}, \mathrm{y}|x, 0)\right\}$$
$$= \min\left\{g(x), c(x)\int_0^s (1 - Q(t, E|x))\mathrm{t} + \int_{[0,s]}\int_E V^*_n(s - t, y)Q(\mathrm{t}, \mathrm{y}|x)\right\}$$

$$= \mathbb{G}V_n^*(s, x) = V_{n+1}^*(s, x),$$

where the second equality is due to inductive hypothesis and (17)-(19). Hence, by Theorem 3.1 part (a), (33) and (34) give that $V^*(s, x) = \lim_{n\to\infty} V_n^*(s, x)$. Further, the monotone convergence theorem gives $V^* = \mathbb{G}V^*$. □

Theorem 4.4. (Optimal stopping time) *Suppose that Assumption 1 holds. Define a subset of* $(0, T] \times E$ *by*

$$D^* := \{(s, x) \in (0, T] \times E : V^*(s, x) = g(x)\}, \tag{35}$$

and for each $\omega = (x_0, t_1, x_1, \ldots, t_n, x_n, \ldots) \in \Omega$, *define*

$$\tau^*(\omega) := \inf\left\{n \middle| (T - \sum_{k=1}^{n} t_k, x_n) \in D^*\right\} \wedge \inf\left\{n \middle| T \leqslant \sum_{k=1}^{n} t_k\right\} \tag{36}$$

Then, τ^* *is a* T-*optimal stopping time and satisfies that* $\mathbb{P}_x(\tau^* = \infty) = 0$ *for all* $x \in E$.

Proof. The definition of f^* given in (32) and Definition 4.1 show that for each $\omega \in \Omega$

$$\tau_{f^*}^T(\omega) = \inf\{n \in \mathbb{N} | f^*(T - S(\omega), X_n(\omega)) = 1\}$$

$$= \inf\left\{n \in \mathbb{N} \middle| U^*(T - \sum_{k=1}^{n} t_k, x_n) = g(x_n), T > \sum_{k=1}^{n} t_k\right\}$$

$$= \inf\left\{n \in \mathbb{N} \middle| (T - \sum_{k=1}^{n} t_k, x_n) \in D^*\right\},$$

which, implies $\tau^*(\omega) = \tau_{f^*}^T(\omega) \wedge \inf\left\{n \middle| T \leqslant \sum_{k=1}^{n} t_k\right\}$. By $V^*(0, x) = 0 = V^{\tau^*}(0, x)$, we only consider the case $T > 0$. First, we have

$$\{\tau^* = 0\} = \{\tau_{f^*}^T = 0\} \cup \{\tau_{f^*}^T > 0, T \leqslant 0\} = \{\tau_{f^*}^T = 0\};$$

$$\{\tau^* = n + 1\} = \{\tau_{f^*}^T = n + 1\} \cup \{\tau_{f^*}^T > n + 1, S_n < T \leqslant S_{n+1}\}, \quad \forall n \geqslant 0;$$

$$\{\tau^* = \infty\} = \{\tau_{f^*}^T = \infty\} \cap \{T > \lim_{n\to\infty} S_n\}.$$

Hence, using that f^* is T-optimal, Theorem 4.1 and Theorem 4.3, we obtain that

$$V^{\tau^*}(T, x) = \sum_{n=1}^{\infty} \left\{\mathbb{E}_x\left[\mathbb{1}_{\{\tau_{f^*}^T = n\}} R_n^T\right] + \mathbb{E}_x\left[\mathbb{1}_{\{S_{n-1} < T \leqslant S_n\}} \mathbb{1}_{\{\tau_{f^*}^T > n\}} R_n^T\right]\right\} + \mathbb{E}_x\left[\mathbb{1}_{\{\tau_{f^*}^T = 0\}} R_n^T\right]$$

$$+ \mathbb{E}_x\left[\mathbb{1}_{\{\tau_{f^*}^T = \infty\}} \mathbb{1}_{\{T > \lim_{n\to\infty} S_n\}} \int_0^T c(X(t)) \mathrm{d} t\right]$$

$$= \sum_{n=0}^{\infty} \hat{\mathbb{E}}_x\left[\mathbb{1}_{\{\tau_{f^*}^T = n\}} R_n^T\right] + \sum_{n=1}^{\infty} \mathbb{E}_x\left[\mathbb{1}_{\{S_{n-1} < T \leqslant S_n\}} \mathbb{1}_{\{\tau_{f^*}^T = \infty\}} R_{\tau_{f^*}^T}^T\right]$$

$$=V^{\tau_{f^*}^T}(T,x) = V^*(T,x).$$

Therefore, τ^* is T-optimal stopping time. Moreover, using (36), we have

$$0 \leqslant \mathbb{P}_x(\tau^* = \infty) \leqslant \mathbb{P}_x(T > \lim_{n\to\infty} S_n) = 0, \forall x \in E,$$

which means τ^* is a finite stopping time, whereas $\tau_{f^*}^T$ does not necessarily. $\qquad\square$

The condition of Theorem 4.3 requires the value function V^*, but in practical applications, V^* is often unknown. Intuitively, we can replace V^* by the approximation function V_n^*, which is obtained by the iterative algorithm given in Theorem 4.3. Therefore, the concept of optimal stopping time will be replaced by ε-T-optimal, that is the following definition.

Definition 4.3. Given any $\varepsilon > 0$, a stopping time τ is called ε-T-optimal if it holds that $V^\tau(T,x) - V^*(T,x) \leqslant \varepsilon$ for all $x \in E$, where $V^\tau(T,x)$ is T-horizon expected cost of the stopping time τ given in (4).

The following theorem shows that for any $\varepsilon > 0$, we can iterate enough times and get an ε-T-optimal stopping time under some conditions. For the convenience of statement, we give two notations, i.e. $\|f\| := \sup_{x\in C} |f(x)|$ for any function f defined on the set C; $\lceil x \rceil := \min\{n \in \mathbb{N} : n \geqslant x\}$ for any $x \in \mathbb{R}_+$.

Theorem 4.5. *Suppose that c and g are bounded and that the semi-Markov kernel Q satisfies* $\sup_{x\in E} Q(T, E|x) =: \beta < 1$. *For any $\varepsilon > 0$, the number of iterations N_ε is given by*

$$N_\varepsilon := \left\lceil \frac{\log(\varepsilon(1-\beta)) - \log(M+1)}{\log\beta} \right\rceil,$$

where $M = T\|c\| + \|g\|$. Let $V_{N_\varepsilon}^*(s,x)$ be the N_ε-th step iterative function given in Theorem 4.3 and D^ε be the subset of $[0,T] \times E$ given by

$$D^\varepsilon := \left\{ (s,x) \in (0,T] \times E | g(x) = \mathbb{G}V_{N_\varepsilon}^*(s,x) \right\}.$$

Then, the following statements hold.

    (a) *Then, define the stopping time τ^ε by*

$$\tau^\varepsilon(\omega) = \inf\left\{ n \middle| ((T - \sum_{k=1}^n t_k, x_n) \in D^\varepsilon \right\} \wedge \inf\left\{ n \middle| T \leqslant \sum_{k=1}^n t_k \right\}.$$

    Then, τ^ε is an ε-T-optimal stopping time.

(b) *If it holds that*

$$\inf_{(s,x)\in((0,T]\times E)\backslash D^{\varepsilon}} \left(g(x) - \mathbb{G}V_{N_{\varepsilon}}^{*}(s,x)\right) > \varepsilon, \tag{37}$$

then, τ^{ε} is also the T-optimal stopping time.

Proof. By (17), we have that

$$\hat{Q}_{T}(T,\hat{E}|x,a) := \begin{cases} Q(T,E|x) \leqslant \beta, & x \in E, a = 0; \\ \mathbb{1}_{[T+1,+\infty)}(T)\delta_{\Delta}(\hat{E}) = 0 \leqslant \beta, & x \in \hat{E}, a = 1. \end{cases}$$

Then, using that c and g are bounded and $\mathbb{T}$ is a monotone operator, we have

$$\|U_{n}^{*}\| \leqslant \|U^{*}\| \leqslant T\|c\| + \|g\| = M, \quad \forall n \geqslant 0.$$

Moreover, by the definitions of $\mathbb{T}$, for each $(s,x) \in [0,T] \times E$, we have

$$\mathbb{T}U_{n+1}^{*}(s,x) \leqslant \mathbb{T}U_{n}^{*}(s,x) + \beta\|U_{n+1}^{*} - U_{n}^{*}\|,$$

which implies that $\|U_{n+1}^{*} - U_{n}^{*}\| = \beta\|U_{n}^{*} - U_{n-1}^{*}\| \leqslant \beta^{n}\|U_{1}^{*} - U_{0}^{*}\| \leqslant \beta^{n}M$. Then, we define a policy f^{ε} by

$$f^{\varepsilon}(s,x) = \mathbb{1}_{(0,\infty)\times E}(s,x)\mathbb{1}_{\{\mathbb{T}U_{N_{\varepsilon}}^{*}(s,x)=\hat{g}(x,1)\}}(s,x) + \mathbb{1}_{\{\Delta\}}(x),$$

and then, we have $f^{\varepsilon} \in \Pi_{HD}^{0}$. Moreover, similar to the proof of Theorem 4.3, we have that

$$U_{N_{\varepsilon}+1}^{*}(s,x) = \mathbb{T}U_{N_{\varepsilon}}^{*}(s,x) = \mathbb{T}^{f^{\varepsilon}(s,x)}U_{N_{\varepsilon}}^{*}(s,x), \quad \forall(s,x) \in [0,T] \times \hat{E}. \tag{38}$$

Then, for each $n \geqslant 0$, by induction, we can show that

$$U_{N_{\varepsilon}+1}^{*}(s,x) \geqslant U_{n}^{f^{\varepsilon}}(s,x) + \hat{\mathbb{E}}_{(s,x)}^{f^{\varepsilon}}\left[U_{N_{\varepsilon}+1}^{*}(s-\hat{S}_{n+1},\hat{X}_{n+1})\right] - \sum_{m=0}^{n}\beta^{m+1}\|U_{N_{\varepsilon}+1}^{*} - U_{N_{\varepsilon}}^{*}\|. \tag{39}$$

By the definition of $\mathbb{T}^{f^{\varepsilon}(s,x)}$, we obtain that

$$U_{N_{\varepsilon}+1}^{*}(s,x) = U_{0}^{f^{\varepsilon}}(s,x) + \int_{[0,s]}\int_{\hat{E}} U_{N_{\varepsilon}}^{*}(s-t,y)\hat{Q}_{T}(\mathrm{t},\mathrm{y}|x,f^{\varepsilon}(s,x))$$

$$\geqslant U_{0}^{f^{\varepsilon}}(s,x) + \hat{\mathbb{E}}_{(s,x)}^{f^{\varepsilon}}\left[U_{N_{\varepsilon}+1}^{*}(s-\hat{S}_{1},\hat{X}_{1})\right] - \beta\|U_{N_{\varepsilon}+1}^{*} - U_{N_{\varepsilon}}^{*}\|,$$

which means that (39) holds for $n = 0$. On the other hand, for any $n \geqslant 0$

$$\hat{\mathbb{E}}\left[U_{N_{\varepsilon}}^{*}(s-\hat{S}_{n+1},\hat{X}_{n+1})\right]$$

$$= \int_{\hat{E}}\delta_{x}(\mathrm{x}_{0})\int_{[0,s]}\int_{\hat{E}}\hat{Q}_{T}(\mathrm{t}_{1},\mathrm{x}_{1}|x_{0},f^{\varepsilon}(s,x_{0}))\int_{[0,s-s_{1}]}\int_{\hat{E}}\hat{Q}_{T}(\mathrm{t}_{2},\mathrm{x}_{2}|x_{1},f^{\varepsilon}(s-s_{1},x_{1}))$$

$$\cdots \int_{[0,s-s_n]} \int_{\hat{E}} \hat{Q}_T(\mathfrak{t}_{n+1}, \mathbf{x}_{n+1} | x_1, f^\varepsilon(s-s_n, x_n)) U^*_{N_\varepsilon}(s-s_{n+1}, x_{n+1})$$

$$\geqslant \int_{\hat{E}} \delta_x(\mathbf{x}_0) \int_0^s \int_{\hat{E}} \hat{Q}_T(\mathfrak{t}_1, \mathbf{x}_1 | x_0, f^\varepsilon(s, x_0)) \int_{[0,s-s_1]} \int_{\hat{E}} \hat{Q}_T(\mathfrak{t}_2, \mathbf{x}_2 | x_1, f^\varepsilon(s-s_1, x_1))$$

$$\cdots \int_{[0,s-s_n]} \int_{\hat{E}} \hat{Q}_T(\mathfrak{t}_{n+1}, \mathbf{x}_{n+1} | x_1, f^\varepsilon(s-s_n, x_n))(U^*_{N_\varepsilon+1}(s-s_{n+1}, x_{n+1}) - \|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\|)$$

$$\geqslant \hat{\mathbb{E}}\left[U^*_{N_\varepsilon+1}(s - \hat{S}_{n+1}, \hat{X}_{n+1})\right] - \beta^{n+1} \|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\|.$$

Thus, suppose that (39) holds for some n, it holds that by (8), (9) and (38)

$$U^*_{N_\varepsilon+1}(s,x) \geqslant U^{f^\varepsilon}_n(s,x) + \hat{\mathbb{E}}^{f^\varepsilon}_{(s,x)}\left[\mathbb{T}^{f^\varepsilon} U^*_{N_\varepsilon}(s - \hat{S}_{n+1}, \hat{X}_{n+1})\right] - \sum_{m=0}^n \beta^{m+1} \|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\|$$

$$= U^{f^\varepsilon}_{n+1}(s,x) + \hat{\mathbb{E}}^{f^\varepsilon}_{(s,x)}\left[U^*_{N_\varepsilon}(s - \hat{S}_{n+2}, \hat{X}_{n+2})\right] - \sum_{m=0}^n \beta^{m+1} \|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\|$$

$$\geqslant U^{f^\varepsilon}_{n+1}(s,x) + \hat{\mathbb{E}}^{f^\varepsilon}_{(s,x)}\left[U^*_{N_\varepsilon+1}(s - \hat{S}_{n+2}, \hat{X}_{n+2})\right] - \sum_{m=0}^{n+1} \beta^{m+1} \|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\|.$$

Hence, passing the limit $n \to \infty$ in (39), it holds that for all $(s,x) \in [0,T] \times E$

$$U^*(s,x) \geqslant U^*_{N_\varepsilon+1}(s,x) \geqslant U^{f^\varepsilon}(s,x) - \frac{\beta}{1-\beta}\|U^*_{N_\varepsilon+1} - U^*_{N_\varepsilon}\| \geqslant U^{f^\varepsilon}(s,x) - \varepsilon. \tag{40}$$

In the same method of Theorem 4.4, we can verify $\tau^\varepsilon = \tau^T_{f^\varepsilon} \wedge \inf\left\{n \big| T \leqslant S_n\right\}$ and

$$V^{\tau^\varepsilon}(T,x) = U^{f^\varepsilon}(T,x) \leqslant V^*(T,x) + \varepsilon,$$

i.e., τ^ε is an ε-T-optimal stopping time.

If $((0,T] \times E) \setminus D^\varepsilon = \varnothing$, then $(0,T] \times E = D^\varepsilon$ and the condition (37) holds naturally. Hence, for each $(s,x) \in (0,T] \times E$, we have $g(x) = V^*_{N_\varepsilon+1}(s,x) \leqslant V^*(s,x)$. That means $D^* = D^\varepsilon$, where D^* given in (35).

Next, we consider the case $((0,T] \times E) \setminus D^\varepsilon \neq \varnothing$. Again, using the monotonicity of $\mathbb{G}$, we have $D^\varepsilon \subset D^*$. Conversely, by the definition of $V^*_{N_\varepsilon}(s,x)$ and (40), it holds that

$$\mathbb{H}V^*_{N_\varepsilon}(s,x) = V^*_{N_\varepsilon+1}(s,x) \geqslant U^{f^\varepsilon}(s,x) - \varepsilon \geqslant U^*(s,x) - \varepsilon = V^*(s,x) - \varepsilon,$$

i.e. $V^*(s,x) \leqslant \mathbb{H}V^*_{N_\varepsilon}(s,x) + \varepsilon$ for each $(s,x) \in (0,T] \times E$. For each $(s,x) \in ((0,T] \times E) \setminus D^\varepsilon$, the condition (37) implies that

$$g(x) - V^*(s,x) \geqslant g(x) - \mathbb{G}V^*_{N_\varepsilon}(s,x) - \varepsilon > 0,$$

which means $(s,x) \in ((0,T] \times E) \setminus D^*$. Hence, we have $D^* \subset D^\varepsilon$ and then $D^* = D^\varepsilon$. Finally, we have $\tau^\varepsilon = \tau^*$, which is a T-optimal stopping time given in Theorem 4.3. $\qquad\square$

Hence, given any accuracy $\varepsilon > 0$ and planning horizon $T > 0$, we can devire an approach of computing ε-T-optimal stopping time.

An algorithm (for ε-T-optimal stopping time)

Step 1 (Initialization): Let $V_0^*(s, x) = 0$ for every $(s, x) \in [0, T] \times E$.

Step 2 (Iteration): Compute the function $V_n^*(s, x)$ for every $(s, x) \in [0, T] \times E$ by

$$V_{n+1}^*(s, x) = \min \left\{ c(x) \int_0^s (1 - Q(t, E|x)) \mathrm{t} + \int_E \int_{[0,s]} V_n^*(s - t, y) Q(\mathrm{t}, \mathrm{y}|x), g(x) \right\}.$$

Step 3 (Accuracy control): If $V_{n+1}^*(s, x) - V^*(s, x) \leqslant \varepsilon$ for every $(s, x) \in [0, T] \times E$, go to *Step 4*; otherwise, go to *Step 2* by replacing n with $n + 1$.

Step 4 (ε-T-optimal stopping time): Compute the set

$$D^\varepsilon = \left\{ (s, x) \in (0, T] \times E | g(x) = \mathbb{G} V_{N_\varepsilon}^*(s, x) \right\}$$

and the ε-T-optimal stopping time

$$\tau^\varepsilon(\omega) = \inf\{n \big| ((T - \sum_{k=1}^n t_k, x_n) \in D^\varepsilon\} \wedge \inf\{n \big| T \leqslant \sum_{k=1}^n t_k\}, \ \forall \omega = (x_0, \ldots, x_n, t_{n+1}, \ldots) \in \Omega.$$

Funding Information

References

[1] ARKIN, V. I. AND SLASTNIKOV, A. D. (2021). On optimal threshold stopping times for Ito diffusions. *Stochastics. An International Journal of Probability and Stochastic Processes.* **93**(5), 655-681.

[2] BÄUERLE, N. AND POPP, A. (2018). Risk-sensitive stopping problems for continuous-time Markov chains. *Stochastics. An International Journal of Probability and Stochastic Processes.* **90**(3), 411-431.

[3] BÄUERLE, N. AND RIEDER, U. (2011). *Markov Decision Processes with Applications to Finance.* Springer, Heidelberg.

[4] BOSHUIZEN, F. A. AND GOUWELEEUW, J. M. (1993). General optimal stopping theorems for semi-Markov processes. *Advances in Applied Probability.* **25**(4), 825-846.

[5] CHEN, F., GUO, X. P. AND LIAO, Z. W. (2021). Optimal stopping time on discounted semi-Markov processes. *Frontiers of Mathematics in China.* **16**(2), 303-324.

[6] CHOW, Y. S., ROBBINS, H. AND SIEGMUND, D. (1991). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin Company. Boston.

[7] CHRISTENSEN, S. AND LINDENSJÖ, K. (2020). On time-inconsistent stopping problems and mixed strategy stopping times.*Stochastic Processes and their Applications.* **130**(5), 2886-2917.

[8] DE SAPORTA, B., DUFOUR, F. AND NIVOT, C. (2017). Partially observed optimal stopping problem for discrete-time Markov processes. *4OR. A Quarterly Journal of Operations Research.* **15**(3), 277-302.

[9] DUFOUR, F. AND PIUNOVSKIY, A. B. (2010). Multiobjective stopping problem for discrete-time Markov processes: convex analytic approach. *Journal of Applied Probability.* **47**(4), 947-966.

[10] GAPEEV, P. V., KORT, P. M. AND LAVRUTICH, M. N. (2021). Multiobjective stopping problem for discrete-time Markov processes: convex analytic approach. *Advances in Applied Probability.* **53**(1), 189-219.

[11] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1996). *Discrete-time Markov control processes: Basic optimality criteria.* Springer-Verlag, New York.

[12] HUANG, Y. H. AND GUO, X. P. (2011). Finite horizon semi-Markov decision processes with application to maintenance systems. *European Journal of Operational Research.* **212**(1), 131-140.

[13] HUANG, Y. J. AND ZHOU, Z. (2019). The optimal equilibrium for time-inconsistent stopping problems: the discrete-time case. *SIAM Journal on Control and Optimization.***57**(1), 590-609.

[14] JAŚKIEWICZ, A. AND NOWAK, A. S. (2006). Optimality in Feller semi-Markov control processes. *Operations Research Letters.* **34**(6), 713-718.

[15] KITAEV, M. Y. (1986). Semi-Markov and jump Markov controlled models: average cost criterion. *Theory of Probability and Its Applications.* **30**(2), 272-288.

[16] KITAEV, M. Y. AND RYKOV, V. (1995). *Controlled Queueing Systems.* CRC Press, Boca Raton, FL.

[17] LIMNIOS, N. AND OPRISAN, G. (2001). *Semi-Markov processes and reliability.* Birkhäuser, Boston.

[18] SHAO, J. H. AND TIAN, T. R. (2021). Optimal stopping problem for jump-diffusion processes with regime-switching. *Nonlinear Analysis: Hybrid Systems.* **41**: 101029.

[19] SHIRYAEV, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag, New York-Heidelberg. Springer.