

# Maximal Cost-Bounded Reachability Probability on Continuous-Time Markov Decision Processes\*

Hongfei Fu\*\*

Lehrstuhl für Informatik II, RWTH Aachen University, Germany

**Abstract.** In this paper, we consider multi-dimensional maximal cost-bounded reachability probability over continuous-time Markov decision processes (CTMDPs). Our major contributions are as follows. Firstly, we derive an integral characterization which states that the maximal cost-bounded reachability probability function is the least fixed-point of a system of integral equations. Secondly, we prove that the maximal cost-bounded reachability probability can be attained by a measurable deterministic cost-positional scheduler. Thirdly, we provide a numerical approximation algorithm for maximal cost-bounded reachability probability. We present these results under the setting of both early and late schedulers. Besides, we correct a fundamental proof error in the PhD Thesis by Martin Neuhäüßer on maximal time-bounded reachability probability by completely new proofs for the more general case of multi-dimensional maximal cost-bounded reachability probability.

## 1 Introduction

The class of continuous-time Markov decision processes (CTMDPs) (or controlled Markov chains) [13,12] is a stochastic model that incorporates both features from continuous-time Markov chains (CTMCs) [6] and discrete-time Markov decision processes (MDPs) [13]. A CTMDP extends a CTMC in the sense that it allows non-deterministic choices, and it extends an MDP in the sense that it incorporates negative exponential time-delays. Due to its modelling capability of real-time probabilistic behaviour and non-determinism, CTMDPs are widely used in dependability analysis and performance evaluation [2].

In a CTMDP, non-determinism is resolved by *schedulers* [16]. Informally, a scheduler determines the non-deterministic choices depending on the finite trajectory of the CTMDP so far and possibly the sojourn time of the current state. A scheduler is assumed to be *measurable* so that it induces a well-defined probability space over the infinite trajectories of the underlying CTMDP. Measurable schedulers are further divided into categories of *early schedulers* and *late schedulers* [10,16]. A scheduler that makes the choice solely by the trajectory so far is called an early scheduler, while a scheduler that utilizes both the trajectory and

---

\* Partially funded by the EU FP7 projects CARP and SENSATION. Full version available at [7].

\*\* Supported by a CSC scholarship.

the sojourn time (at the current state) is called a late scheduler. With schedulers, one can reason about quantitative information such as the maximal/minimal probability/expectation of certain property.

In this paper, we focus on the problem to compute *max/min resource-bounded reachability probability* on a CTMDP. Typical resource types considered here are time and cost, where a time bound can be deemed as a special cost bound with unit-cost 1. In general, the task is to compute or approximate the optimal (max/min) reachability probability to certain target states within a given resource bound (e.g., a time bound).

Optimal time-bounded reachability probability over CTMDPs has been widely studied in recent years. Neuhäuser *et al.* [11] proved that the maximal time-bounded reachability probability function is the least fixed point of a system of integral equations. Rabe and Schewe [14] showed that the max/min time-bounded reachability probability can be attained by a deterministic piecewise-constant time-positional scheduler. Efficient approximation algorithms are also developed by, e.g., Neuhäuser *et al.* [11], Brázdil *et al.* [3], Hatefi *et al.* [8] and Rabe *et al.* [5].

As to optimal cost-bounded reachability probability, much less is known. To the best of the author's knowledge, the only prominent result is by Baier *et al.* [1], which establishes a certain duality property between time and cost bound. Their result is restrictive in the sense that (i) it assumes that the CTMDP have everywhere positive unit-cost values, (ii) it only takes into account one-dimensional cost-bound aside the time-bound, and (iii) it does not really provide an approximation algorithm when both time- and cost-bounds are present.

Besides resource-bounded reachability probability, we would like to mention another research field on CTMDPs with costs (or dually, rewards), which is (discounted) accumulated reward over finite/infinite horizon (cf. [4,12], just to mention a little).

**Our Contribution.** We consider multi-dimensional maximal cost-bounded reachability probability (*abbr.* MMCRP) over CTMDPs, under the setting of both early and late schedulers, for which the unit-cost is constant. We first prove that the MMCRP function is the least fixed-point of a system of integral equations. Then we prove that deterministic cost-positional measurable schedulers suffice to achieve the MMCRP value. Finally, we describe a numerical algorithm which approximates the MMCRP value with an error bound. The approximation algorithm relies on a differential characterization which in turn is derived from the least fixed-point characterization. The complexity of the approximation algorithm is polynomial in the size of the CTMDP and the reciprocal of the error bound, and exponential in the dimension of cost vectors.

Besides, we point out a fundamental proof error in the treatment of maximal time-bounded reachability probability on continuous-time Markov decision processes [9,11]. We fix this error in the more general setting of maximal cost-bounded reachability probability by completely new proofs.

**Structure of the Paper.** Section 2 introduces some preliminaries of CTMDPs. Section 3 illustrates the definition of schedulers and the probability spaces they induce. In Section 4, we define the notion of maximal cost-bounded reachability probability and derive the least-fixed-point characterization, while we also point out the proof error in [9,11]. In Section 5, we prove that the maximal cost-bounded reachability probability can be reached by a measurable deterministic cost-positional scheduler. In Section 6, we derive a differential characterization which is crucial to our approximation algorithm. In Section 7, we present our approximation algorithm. Finally, Section 8 concludes the paper.

Due to page limit, we omit all the proofs and only present the results for late schedulers. The details can be found at [7].

## 2 Continuous-Time Markov Decision Processes

In the whole paper, we will use the following convention for notations. We will denote by  $\mathbb{R}_{\geq 0}$  the set of non-negative real numbers and by  $\mathbb{N}_0$  the set of non-negative integers. We use  $x, d, t, \tau$  to range over real numbers,  $m, n, i, j$  to range over  $\mathbb{N}_0$ , and bold-face letters  $\mathbf{x}, \mathbf{c}, \mathbf{d}$  to range over (column) real vectors. Given  $\mathbf{c} \in \mathbb{R}^k$  ( $k \in \mathbb{N}$ ), we denote by  $\mathbf{c}_i$  ( $1 \leq i \leq k$ ) the  $i$ -th coordinate of  $\mathbf{c}$ . We denote by  $\mathbf{0}$  the real vector whose coordinates are all equal to 0 (with the implicitly known dimension). We extend  $\{\leq, <, \geq, >\}$  to real vectors and functions in a pointwise fashion: for two real vectors  $\mathbf{c}, \mathbf{d}$ ,  $\mathbf{c} \leq \mathbf{d}$  iff  $\mathbf{c}_i \leq \mathbf{d}_i$  for all  $i$ ; for two real-valued functions  $g, h$ ,  $g \leq h$  iff  $g(y) \leq h(y)$  for all  $y$ . Given a set  $Y$ , we let  $\mathbf{1}_Y$  be the indicator function of  $Y$ , i.e.,  $\mathbf{1}_Y(y) = 1$  if  $y \in Y$  and  $\mathbf{1}_Y(y) = 0$  for  $y \in X - Y$ , where  $X \supseteq Y$  is an implicitly known set. Given a positive real number  $\lambda > 0$ , let  $f_\lambda(t) := \lambda \cdot e^{-\lambda \cdot t}$  ( $t \geq 0$ ) be the probability density function of the negative exponential distribution with rate  $\lambda$ . Besides, we will use  $g, h$  to range over general functions.

### 2.1 The Model

**Definition 1.** A Continuous-Time Markov Decision Process (CTMDP) is a tuple  $(L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  where

- $L$  is a finite set of states (or locations);
- $Act$  is a finite set of actions;
- $\mathbf{R} : L \times Act \times L \rightarrow \mathbb{R}_{\geq 0}$  is the rate matrix;
- $\{\mathbf{w}_i : L \times Act \rightarrow \mathbb{R}_{\geq 0}\}_{1 \leq i \leq k}$  is the family of  $k$  unit-cost functions ( $k \in \mathbb{N}$ );

An action  $a \in Act$  is enabled at state  $s \in L$  if  $\mathbf{E}(s, a) := \sum_{u \in L} \mathbf{R}(s, a, u)$  is non-zero. The set of enabled actions at  $s \in L$  is denoted by  $\text{En}(s)$ . We assume that for each state  $s \in L$ ,  $\text{En}(s) \neq \emptyset$ .

Let  $(L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  be a CTMDP. For each  $s, s' \in L$  and  $a \in \text{En}(s)$ , we define  $\mathbf{P}(s, a, s') := \frac{\mathbf{R}(s, a, s')}{\mathbf{E}(s, a)}$  to be the discrete transition probability from  $s$  to  $s'$  via  $a$ . We denote by  $\mathbf{w}(s, a)$  the real vector  $\{\mathbf{w}_i(s, a)\}_{1 \leq i \leq k}$  for each

$(s, a) \in L \times Act$ . Given  $s \in L$  and  $a \in Act$ , we denote by  $\mathcal{D}[s]$  the Dirac distribution (over  $L$ ) at  $s$  (i.e.,  $\mathcal{D}[s](s) = 1$  and  $\mathcal{D}[s](s') = 0$  for  $s' \in L - \{s\}$ ) and by  $\mathcal{D}[a]$  the Dirac distribution (over  $Act$ ) at  $a$ . Moreover, we define (with  $\min \emptyset := 1$ ):

- $\mathbf{w}_{\min} := \min\{\mathbf{w}_i(s, a) \mid 1 \leq i \leq k, s \in L, a \in \text{En}(s), \mathbf{w}_i(s, a) > 0\}$  ;
- $\mathbf{w}_{\max} := \max\{\mathbf{w}_i(s, a) \mid 1 \leq i \leq k, s \in L, a \in \text{En}(s)\}$  ;
- $\mathbf{E}_{\max} := \max\{\mathbf{E}(s, a) \mid s \in L, a \in \text{En}(s)\}$  ;

We will use  $s, s'$  (resp.  $a, b$ ) to range over states (resp. actions) of a CTMDP.

Often, a CTMDP is accompanied with an initial distribution which specifies the initial stochastic environment (for the CTMDP).

**Definition 2.** Let  $\mathcal{M} = (L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  be a CTMDP. An initial distribution (for  $\mathcal{M}$ ) is a function  $\alpha : L \rightarrow [0, 1]$  such that  $\sum_{s \in L} \alpha(s) = 1$ .

Intuitively, the execution of a CTMDP  $(L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  with a scheduler is as follows. At the beginning, an initial state  $s$  is chosen (as the current state) w.r.t the initial distribution  $\alpha$ . Then the scheduler chooses an action  $a$  enabled at  $s$  either before or after a time-delay occurs at the state  $s$ . After the time-delay, the current state is switched to an arbitrary state  $s' \in L$  with probability  $\mathbf{P}(s, a, s')$ , and so forth. Besides, each cost function  $\mathbf{w}_i$  assigns to each state-action pair  $(s, a)$  the  $i$ -th constant unit-cost  $\mathbf{w}_i(s, a)$  (per time unit) when the CTMDP dwells at state  $s$ . Basically, the scheduler makes the decision of the action to be chosen when entering a new state, and has two distinct objectives: either to maximize a certain property or (in contrast) to minimize a certain property. In this paper, we will focus on the objective to maximize a cost-bounded reachability probability for a certain target set of states.

In this paper, we focus on an important subclass of CTMDPs, called *locally-uniform* CTMDPs (cf. [10]).

**Definition 3.** A CTMDP  $(L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  is locally-uniform if  $\mathbf{E}(s, a) = \mathbf{E}(s, b)$  and  $\mathbf{w}_i(s, a) = \mathbf{w}_i(s, b)$  for all  $1 \leq i \leq k$ ,  $s \in L$  and  $a, b \in \text{En}(s)$ .

Intuitively, a locally uniform CTMDP has the property that the time-delay and the cost is independent of the action chosen at each state. For locally-uniform CTMDPs, we simply use  $\mathbf{E}(s)$  to denote  $\mathbf{E}(s, a)$  ( $a \in \text{En}(s)$  is arbitrary), and  $\mathbf{w}(s), \mathbf{w}_i(s)$  for  $\mathbf{w}(s, a), \mathbf{w}_i(s, a)$  likewise.

## 2.2 Paths and Histories

In this part, we introduce the notion of paths and histories. Intuitively, paths reflect infinite executions of a CTMDP, whereas histories reflect finite executions of a CTMDP. Below we fix a CTMDP  $\mathcal{M} = (L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$ .

**Definition 4.** A (*n infinite*) path  $\pi$  is an infinite sequence

$$\pi = \left\langle s_0 \xrightarrow{a_0, t_0} s_1 \xrightarrow{a_1, t_1} s_2 \dots \right\rangle$$

such that  $s_i \in L$ ,  $t_i \in \mathbb{R}_{\geq 0}$  and  $a_i \in \text{Act}$  for all  $i \geq 0$ ; We denote  $s_i, t_i$  and  $a_i$  by  $\pi[i], \pi\langle i \rangle$  and  $\pi(i)$ , respectively. A (finite) history  $\xi$  is a finite sequence

$$\xi = \left\langle s_0 \xrightarrow{a_0, t_0} s_1 \xrightarrow{a_1, t_1} s_2 \dots s_m \right\rangle \quad (m \geq 0)$$

such that  $s_i \in L$ ,  $t_i \in \mathbb{R}_{\geq 0}$  and  $a_i \in \text{Act}$  for all  $0 \leq i \leq m-1$ , and  $s_m \in L$ ; We denote  $s_i, t_i, a_i$  and  $m$  by  $\xi[i], \xi\langle i \rangle, \xi(i)$  and  $|\xi|$ , respectively. Moreover, we define  $\xi \downarrow := \xi[|\xi|]$  to be the last state of the history  $\xi$ .

Below we introduce more notations on paths and histories. We denote the set of paths and histories (of  $\mathcal{M}$ ) by  $\text{Paths}(\mathcal{M})$  and  $\text{Hists}(\mathcal{M})$ , respectively. We define  $\text{Hists}^n(\mathcal{M}) := \{\xi \in \text{Hists}(\mathcal{M}) \mid |\xi| = n\}$  to be the set of all histories with length  $n$  ( $n \geq 0$ ). For each  $n \in \mathbb{N}_0$  and  $\pi \in \text{Paths}(\mathcal{M})$ , we define the history  $\pi[0..n]$  to be the finite prefix of  $\pi$  up to  $n$ ; Formally,

$$\pi[0..n] := \left\langle \pi[0] \xrightarrow{\pi(0), \pi\langle 0 \rangle} \dots \pi[n] \right\rangle .$$

Given  $\pi \in \text{Paths}(\mathcal{M})$  and  $(s, a, t) \in L \times \text{Act} \times \mathbb{R}_{\geq 0}$ , we denote by  $s \xrightarrow{a, t} \pi$  the path obtained by “putting” the prefix “ $s \xrightarrow{a, t}$ ” before  $\pi$ ; Formally,

$$s \xrightarrow{a, t} \pi := \left\langle s \xrightarrow{a, t} \pi[0] \xrightarrow{\pi(0), \pi\langle 0 \rangle} \pi[1] \xrightarrow{\pi(1), \pi\langle 1 \rangle} \dots \right\rangle .$$

Analogously, we define  $s \xrightarrow{a, t} \xi$  (for  $\xi \in \text{Hists}(\mathcal{M})$ ) to be the history obtained by “putting” “ $s \xrightarrow{a, t}$ ” before the history  $\xi$ .

Intuitively, a path  $\pi$  reflects a whole execution (trajectory) of the CTMDP where  $\pi[i]$  is the current state at the  $i$ -th stage,  $\pi(i)$  is the action chosen at  $\pi[i]$  and  $\pi\langle i \rangle$  is the dwell-time (time-delay) on  $\pi[i]$ . On the other hand, a history  $\xi$  is a finite prefix of a path which reflects the execution up to  $|\xi|$  stages.

Below we extend sets of histories to sets of paths in a cylindrical fashion.

**Definition 5.** Suppose  $n \in \mathbb{N}_0$  and  $\Xi \subseteq \text{Hists}^n(\mathcal{M})$ . The cylinder extension of  $\Xi$ , denoted  $\text{Cyl}(\Xi)$ , is defined as follows:

$$\text{Cyl}(\Xi) := \{\pi \in \text{Paths}(\mathcal{M}) \mid \pi[0..n] \in \Xi\} .$$

In this paper, we concern costs on paths and histories. The cost is assigned linearly w.r.t the unit-cost and the time spent in a state. The following definition presents the details.

**Definition 6.** Given a path  $\pi \in \text{Paths}(\mathcal{M})$  and a set  $G \subseteq L$  of states, we denote by  $\mathbf{C}_j(\pi, G)$  ( $1 \leq j \leq k$ ) the  $j$ -th accumulated cost along  $\pi$  until  $G$  is reached; Formally, if  $\pi[m] \in G$  for some  $m \geq 0$  then

$$\mathbf{C}_j(\pi, G) := \sum_{i=0}^m \mathbf{w}_j(\pi[i], \pi(i)) \cdot \pi(i)$$

where  $n \in \mathbb{N}_0 \cup \{-1\}$  is the smallest integer such that  $\pi[n+1] \in G$ ; otherwise  $\mathbf{C}_j(\pi, G) := +\infty$ . Given a history  $\xi \in \text{Hists}(\mathcal{M})$ , we denote by  $\mathbf{C}_j(\xi)$  ( $1 \leq j \leq k$ ) the accumulated cost of  $\xi$  w.r.t the  $j$ -th unit-cost function; Formally,

$$\mathbf{C}_j(\xi) := \sum_{i=0}^{|\xi|-1} \mathbf{w}_j(\xi[i], \xi(i)) \cdot \xi(i).$$

We denote by  $\mathbf{C}(\pi, G)$  and  $\mathbf{C}(\xi)$  the vectors  $\{\mathbf{C}_j(\pi, G)\}_{1 \leq j \leq k}$  and  $\{\mathbf{C}_j(\xi)\}_{1 \leq j \leq k}$ .

### 2.3 Measurable Spaces on Paths and Histories

In the following, we define the measurable spaces for paths and histories, following the definitions of [16,10]. Below we fix a CTMDP  $\mathcal{M} = (L, \text{Act}, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$ . Firstly, we introduce the notion of combined actions and its measurable space.

**Definition 7.** A combined action is a tuple  $(a, t, s)$  where  $a \in \text{Act}$ ,  $t \in \mathbb{R}_{\geq 0}$  and  $s \in L$ . The measurable space  $(\Gamma_{\mathcal{M}}, \mathcal{U}_{\mathcal{M}})$  over combined actions is defined as follows:

- $\Gamma_{\mathcal{M}} := \text{Act} \times \mathbb{R}_{\geq 0} \times L$  is the set of combined actions;
- $\mathcal{U}_{\mathcal{M}} := 2^{\text{Act}} \otimes \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes 2^L$  is the product  $\sigma$ -algebra for which  $\mathcal{B}(\mathbb{R}_{\geq 0})$  is the Borel  $\sigma$ -field on  $\mathbb{R}_{\geq 0}$ .

The following definition introduces the notion of templates which will be used to define the measurable spaces.

**Definition 8.** A template  $\theta$  is a finite sequence  $\theta = \langle s, U_1, \dots, U_m \rangle$  ( $m \geq 0$ ) such that  $s \in L$  and  $U_i \in \mathcal{U}_{\mathcal{M}}$  for  $1 \leq i \leq m$ ; The length of  $\theta$ , denoted by  $|\theta|$ , is defined to be  $m$ . The set of histories  $\text{Hists}(\theta)$  spanned by a template  $\theta$  is defined by:

$$\text{Hists}(\langle s, U_1, \dots, U_m \rangle) := \left\{ \xi \in \text{Hists}^m(\mathcal{M}) \mid \xi[0] = s \text{ and } (\xi(i), \xi(i), \xi[i+1]) \in U_{i+1} \text{ for all } 0 \leq i < m \right\}.$$

Now we introduce the measurable spaces on paths and histories, as in the following definition.

**Definition 9.** The measurable space  $(\Omega_{\mathcal{M}}^n, \mathcal{S}_{\mathcal{M}}^n)$  over  $\text{Hists}^n(\mathcal{M})$  ( $n \in \mathbb{N}_0$ ) is defined as follows:  $\Omega_{\mathcal{M}}^n = \text{Hists}^n(\mathcal{M})$  and  $\mathcal{S}_{\mathcal{M}}^n$  is generated by the family

$$\{\text{Hists}(\theta) \mid \theta \text{ is a template and } |\theta| = n\}$$

of subsets of  $\text{Hists}^n(\mathcal{M})$ .

The measurable space  $(\Omega_{\mathcal{M}}, \mathcal{S}_{\mathcal{M}})$  over  $\text{Paths}(\mathcal{M})$  is defined as follows:  $\Omega_{\mathcal{M}} = \text{Paths}(\mathcal{M})$  and  $\mathcal{S}_{\mathcal{M}}$  is the smallest  $\sigma$ -algebra generated by the family

$$\{\text{Cyl}(\Xi) \mid \Xi \in \mathcal{S}_{\mathcal{M}}^n \text{ for some } n \geq 0\}$$

of subsets of  $\text{Paths}(\mathcal{M})$ .

### 3 Schedulers and Their Probability Spaces

The stochastic feature of a CTMDP is endowed by a (measurable) scheduler which resolves the action when a state is entered. In the following, we briefly introduce late schedulers for CTMDPs. Most notions in this part stem from [16,10]. Below we fix a locally-uniform CTMDP  $\mathcal{M} = (L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$ .

**Definition 10.** A late scheduler  $D$  is a function

$$D : Hists(\mathcal{M}) \times \mathbb{R}_{\geq 0} \times Act \rightarrow [0, 1]$$

such that for each  $\xi \in Hists(\mathcal{M})$  and  $t \in \mathbb{R}_{\geq 0}$ , the following conditions hold:

- $\sum_{a \in Act} D(\xi, t, a) = 1$  ;
- for all  $a \in Act$ ,  $D(\xi, t, a) > 0$  implies  $a \in \text{En}(\xi \downarrow)$  .

$D$  is measurable iff for all  $n \geq 0$  and  $a \in Act$ , the function  $D(\cdot, \cdot, a)$  is measurable w.r.t  $(\Omega_{\mathcal{M}}^n \times \mathbb{R}_{\geq 0}, \mathcal{S}_{\mathcal{M}}^n \otimes \mathcal{B}(\mathbb{R}_{\geq 0}))$ , provided that the domain of  $D(\cdot, \cdot, a)$  is restricted to  $Hists^n(\mathcal{M}) \times \mathbb{R}_{\geq 0}$  .

Intuitively, a late scheduler  $D$  chooses a distribution over actions immediately after the time-delay at the current state  $s$  (i.e., the last state of a history) is ended; the time-delay observes the negative exponential distribution with rate  $\mathbf{E}(s)$ . The decision  $D(\xi, t, \cdot)$  is based on the history  $\xi$  and the dwell time  $t$  at  $\xi \downarrow$ ; the next state is determined stochastically w.r.t  $\mathbf{P}(\xi \downarrow, a, \cdot)$ , where  $a$  is in turn determined w.r.t  $D(\xi, t, \cdot)$ . The local-uniformity allows a late scheduler to make such decision, without mathematical ambiguity on the accumulated cost and the probability density function for the time-delay. Moreover, the measurability condition will be needed to define a probability measure for the measurable space  $(\Omega_{\mathcal{M}}, \mathcal{S}_{\mathcal{M}})$ .

Each measurable late scheduler will induce a probability measure on combined actions, when applied to a specific history. Below we introduce the probability measure induced by measurable late schedulers.

**Definition 11.** Let  $\xi \in Hists(\mathcal{M})$  be a history and  $D$  a measurable late scheduler. The probability measure  $\mu_{\mathcal{M}}^D(\xi, \cdot)$  for the measurable space  $(\Gamma_{\mathcal{M}}, \mathcal{U}_{\mathcal{M}})$  is defined as follows:

$$\mu_{\mathcal{M}}^D(\xi, U) := \int_{\mathbb{R}_{\geq 0}} f_{\mathbf{E}(\xi \downarrow)}(t) \cdot \left\{ \sum_{a \in \text{En}(\xi \downarrow)} D(\xi, t, a) \cdot \left[ \sum_{s \in L} \mathbf{1}_U(a, t, s) \cdot \mathbf{P}(\xi \downarrow, a, s) \right] \right\} dt$$

for each  $U \in \mathcal{U}_{\mathcal{M}}$ .

Now we define the probability spaces on histories and paths. Firstly, we define the probability space on histories. To this end, we introduce the notion of concatenation as follows.

**Definition 12.** Let  $\xi \in Hists(\mathcal{M})$  be a history and  $(a, t, s) \in \Gamma_{\mathcal{M}}$  be a combined action. We define  $\xi \circ (a, t, s) \in Hists(\mathcal{M})$  to be the history obtained by concatenating  $(a, t, s)$  to  $\xi \downarrow$  (i.e.  $\xi \circ (a, t, s) = \xi[0] \dots \xi \downarrow \xrightarrow{a, t} s$ ) .

Then the probability space on histories of fixed length is given as follows.

**Definition 13.** Suppose  $D$  is a measurable late scheduler and  $\alpha$  is an initial distribution. The sequence  $\{\Pr_{\mathcal{M},D,\alpha}^n : \mathcal{S}_{\mathcal{M}}^n \rightarrow [0, 1]\}_{n \geq 0}$  of probability measures is inductively as follows:

$$\Pr_{\mathcal{M},D,\alpha}^0(\Xi) := \sum_{s \in \Xi} \alpha(s) ;$$

$$\Pr_{\mathcal{M},D,\alpha}^{n+1}(\Xi) := \int_{\Omega_{\mathcal{M}}^n} \left[ \int_{\Gamma_{\mathcal{M}}} \mathbf{1}_{\Xi}(\xi \circ \gamma) \mu_{\mathcal{M}}^D(\xi, d\gamma) \right] \Pr_{\mathcal{M},D,\alpha}^n(d\xi) ;$$

for each  $\Xi \in \mathcal{S}_{\mathcal{M}}^n$ .

Finally, the probability space on paths is given as follows.

**Definition 14.** Let  $D$  be a measurable late scheduler and  $\alpha$  be an initial distribution. The probability space  $(\Omega_{\mathcal{M}}, \mathcal{S}_{\mathcal{M}}, \Pr_{\mathcal{M},D,\alpha})$  is defined as follows:

- $\Omega_{\mathcal{M}}$  and  $\mathcal{S}_{\mathcal{M}}$  is defined as in Definition 9;
- $\Pr_{\mathcal{M},D,\alpha}$  is the unique probability measure such that

$$\Pr_{\mathcal{M},D,\alpha}(\text{Cyl}(\Xi)) = \Pr_{\mathcal{M},D,\alpha}^n(\Xi)$$

for all  $n \geq 0$  and  $\Xi \in \mathcal{S}_{\mathcal{M}}^n$ .

We end this section with a fundamental property asserting that the role of initial distribution  $\alpha$  can be decomposed into Dirac distributions on individual states.

**Proposition 1.** For each measurable late scheduler  $D$  and each initial distribution  $\alpha$ ,  $\Pr_{\mathcal{M},D,\alpha}(\Pi) = \sum_{s \in L} \alpha(s) \cdot \Pr_{\mathcal{M},D,\mathcal{D}[s]}(\Pi)$  for all  $\Pi \in \mathcal{S}_{\mathcal{M}}$ .

## 4 Maximal Cost-Bounded Reachability Probability

In this section, we consider maximal cost-bounded reachability probabilities. Below we fix a locally-uniform CTMDP  $\mathcal{M} = (L, \text{Act}, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  and a set  $G \subseteq L$ . For the sake of simplicity, we omit the ‘ $\mathcal{M}$ ’ which appear in the subscript of ‘Pr’.

**Definition 15.** Let  $D$  be a measurable late scheduler. Define the function  $\text{prob}_G^D : L \times \mathbb{R}^k \rightarrow [0, 1]$  by:  $\text{prob}_G^D(s, \mathbf{c}) := \Pr_{D,\mathcal{D}[s]}(\Pi_G^{\mathbf{c}})$  where

$$\Pi_G^{\mathbf{c}} := \{\pi \in \text{Paths}(\mathcal{M}) \mid \mathbf{C}(\pi, G) \leq \mathbf{c}\} .$$

Define  $\text{prob}_G^{\max} : L \times \mathbb{R}^k \rightarrow [0, 1]$  by:  $\text{prob}_G^{\max}(s, \mathbf{c}) := \sup_{D \in \mathcal{L}_{\mathcal{M}}} \text{prob}_G^D(s, \mathbf{c})$  for  $s \in L$  and  $\mathbf{c} \in \mathbb{R}^k$ , where  $\mathcal{L}_{\mathcal{M}}$  is the set of all measurable late schedulers.



From the definition, we can see that  $\Pi_G^c$  is the set of paths which can reach  $G$  within cost  $\mathbf{c}$ ,  $\text{prob}_G^{\max}(s, \mathbf{c})$  is the maximal probability of  $\Pi_G^c$  with initial distribution  $\mathcal{D}(s)$  (i.e., fixed initial state  $s$ ) ranging over all late schedulers. It is not hard to verify that  $\Pi_G^c$  is measurable, thus all functions in Definition 15 are well-defined. It is worth noting that if  $\mathbf{c} \not\geq \mathbf{0}$ , then both  $\text{prob}_G^D(s, \mathbf{c})$  and  $\text{prob}_G^{\max}(s, \mathbf{c})$  is zero.

The following theorem mainly presents the fixed-point characterization for  $\text{prob}_G^{\max}$ , while it also states that  $\text{prob}_G^{\max}$  is Lipschitz continuous.

**Theorem 1.** *The function  $\text{prob}_G^{\max}$  is the least fixed-point (w.r.t  $\leq$ ) of the high-order operator  $\mathcal{T}_G : [L \times \mathbb{R}^k \rightarrow [0, 1]] \rightarrow [L \times \mathbb{R}^k \rightarrow [0, 1]]$  defined as follows:*

- $\mathcal{T}_G(h)(s, \mathbf{c}) := \mathbf{1}_{\mathbb{R}_{\geq 0}^k}(\mathbf{c})$  if  $s \in G$ ;
- If  $s \notin G$  then

$$\mathcal{T}_G(h)(s, \mathbf{c}) := \int_0^\infty f_{\mathbf{E}(s)}(t) \cdot \max_{a \in \text{En}(s)} \left[ \sum_{s' \in L} \mathbf{P}(s, a, s') \cdot h(s', \mathbf{c} - t \cdot \mathbf{w}(s)) \right] dt ;$$

for each  $h : L \times \mathbb{R}^k \rightarrow [0, 1]$ . Moreover,

$$|\text{prob}_G^{\max}(s, \mathbf{c}) - \text{prob}_G^{\max}(s, \mathbf{c}')| \leq \frac{\mathbf{E}_{\max}}{\mathbf{w}_{\min}} \cdot \|\mathbf{c} - \mathbf{c}'\|_\infty$$

for all  $\mathbf{c}, \mathbf{c}' \geq \mathbf{0}$  and  $s \in L$ .

The Lipschitz constant  $\frac{\mathbf{E}_{\max}}{\mathbf{w}_{\min}}$  will be crucial to the error bound of our approximation algorithm.

Now we describe the proof error in [9,11]. The error lies in the proof of [9, Lemma 5.1 on Pages 119] which tries to prove that the time-bounded reachability probability functions are continuous. In detail, the error is at the proof for right-continuity of the functions. Let us take the sentence ‘‘This implies ... for some  $\xi \leq \frac{\epsilon}{2}$ .’’ from line -3 to line -2 on page 119 as (\*). (\*) is wrong in general, as one can treat  $D$ 's as natural numbers, and define

$$\text{Pr}_n(\text{‘‘reach } G \text{ within } z\text{’’}) := \begin{cases} n \cdot z & \text{if } z \in [0, \frac{1}{n}] \\ 1 & \text{if } z \in (\frac{1}{n}, \infty) \end{cases} .$$

Then  $\sup_n \text{Pr}_n(\text{‘‘reach } G \text{ within } z\text{’’})$  equals 1 for  $z > 0$  and 0 for  $z = 0$ . Thus  $\sup_D \text{Pr}_D(\text{‘‘reach } G \text{ within } z\text{’’})$  on  $z \geq 0$  is right-discontinuous at  $z = 0$ , which does not satisfy (\*) (treat  $D$  as a natural number). Note that a concrete counterexample does not exist as [9, Lemma 5.1] is correct due to this paper; it is the proof that is flawed. Also note that Lemma 5.1 is important as the least fixed-point characterization [9, Theorem 5.1 on Page 120] and the optimal scheduler [9, Theorem 5.2 on page 124] directly rely on it. We fix the error in the more general setting of cost-bounded reachability probability by providing new proofs.

## 5 Optimal Schedulers

In this section, we establish optimal late schedulers for maximal cost-bounded reachability probability. We show that there exists a deterministic cost-positional late scheduler that achieves the maximal cost-bounded reachability probability. Below we fix a locally-uniform CTMDP  $\mathcal{M} = (L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$ . We first introduce the notion of deterministic cost-positional schedulers.

**Definition 16.** *A measurable late scheduler  $D$  is deterministic cost-positional iff (i)  $D(\xi, t, \bullet) = D(\xi', t', \bullet)$  whenever  $\xi \downarrow = \xi' \downarrow$  and  $\mathbf{C}(\xi) + t \cdot \mathbf{w}(\xi \downarrow) = \mathbf{C}(\xi') + t' \cdot \mathbf{w}(\xi' \downarrow)$ , and (ii)  $D(\xi, t, \bullet)$  is Dirac for all histories  $\xi$  and  $t \geq 0$ .*

Intuitively, a deterministic cost-positional scheduler makes its decision solely on the current state and the cost accumulated so far, and its decision is always Dirac. The following theorem shows that such a scheduler suffices to achieve maximal cost-bounded reachability probability.

**Theorem 2.** *For all  $\mathbf{c} \in \mathbb{R}^k$  and  $G \subseteq L$ , there exists a deterministic cost-positional measurable late scheduler  $\mathsf{D}$  such that  $\text{prob}_G^{\max}(s, \mathbf{c}) = \text{Pr}_{\mathsf{D}, \mathcal{D}[s]}(\Pi_G^{\mathbf{c}})$  for all  $s \in L$ .*

## 6 Differential Characterization for Maximal Reachability Probabilities

In this section, we derive differential characterization for the function  $\text{prob}_G^{\max}$ . The differential characterization will be fundamental to our approximation algorithm. Below we fix a locally-uniform CTMDP  $(L, Act, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$  and a set  $G \subseteq L$ .

The differential characterization relies on a notion of directional derivative as follows.

**Definition 17.** *Let  $s \in L - G$  and  $\mathbf{c} \geq \mathbf{0}$ . Define*

$$\nabla^+ \text{prob}_G^{\max}(s, \mathbf{c}) := \lim_{t \rightarrow 0^+} \frac{\text{prob}_G^{\max}(s, \mathbf{c} + t \cdot \mathbf{w}(s)) - \text{prob}_G^{\max}(s, \mathbf{c})}{t} .$$

*If  $\mathbf{c}_i > 0$  whenever  $\mathbf{w}_i(s) > 0$  ( $1 \leq i \leq k$ ), define*

$$\nabla^- \text{prob}_G^{\max}(s, \mathbf{c}) := \lim_{t \rightarrow 0^-} \frac{\text{prob}_G^{\max}(s, \mathbf{c} + t \cdot \mathbf{w}(s)) - \text{prob}_G^{\max}(s, \mathbf{c})}{t} ;$$

*Otherwise, let  $\nabla^- \text{prob}_G^{\max}(s, \mathbf{c}) = \nabla^+ \text{prob}_G^{\max}(s, \mathbf{c})$ .*

The following theorem gives a characterization for the directional derivative.

**Theorem 3.** *For all  $s \in L - G$  and  $\mathbf{c} \geq \mathbf{0}$ ,  $\nabla^+ \text{prob}_G^{\max}(s, \mathbf{c}) = \nabla^- \text{prob}_G^{\max}(s, \mathbf{c})$ . Moreover,*

$$\nabla^+ \text{prob}_G^{\max}(s, \mathbf{c}) = \max_{a \in \text{En}(s)} \sum_{s' \in L} \mathbf{R}(s, a, s') \cdot (\text{prob}_G^{\max}(s', \mathbf{c}) - \text{prob}_G^{\max}(s, \mathbf{c})) .$$

As  $\nabla^+ \text{prob}_G^{\max}(s, \mathbf{c}) = \nabla^- \text{prob}_G^{\max}(s, \mathbf{c})$ , we will solely use  $\nabla \text{prob}_G^{\max}(s, \mathbf{c})$  to denote both of them.

Theorem 3 allows one to approximate  $\text{prob}_G^{\max}(s, \mathbf{c} + t \cdot \mathbf{w}(s))$  by  $\text{prob}_G^{\max}(s, \mathbf{c})$  and  $\nabla^+ \text{prob}_G^{\max}(s, \mathbf{c})$ . An exception is the case  $\mathbf{w}(s) = \mathbf{0}$ . Below we tackle this situation.

**Proposition 2.** *Let  $Y_G := \{s \in L \mid \mathbf{w}(s) = \mathbf{0}\}$ . For each  $\mathbf{c} \geq \mathbf{0}$ , the function  $s \mapsto \text{prob}_G^{\max}(s, \mathbf{c})$  is the least fixed-point (w.r.t  $\leq$ ) of the high-order operator  $\mathcal{Y}_{\mathbf{c}, G} : [Y_G \rightarrow [0, 1]] \rightarrow [Y_G \rightarrow [0, 1]]$  defined as follows:*

$$\mathcal{Y}_{\mathbf{c}, G}(h)(s) := \max_{a \in \text{En}(s)} \left[ \sum_{s' \in Y_G} \mathbf{P}(s, a, s') \cdot h(s') + \sum_{s' \in L - Y_G} \mathbf{P}(s, a, s') \cdot \text{prob}_G^{\max}(s', \mathbf{c}) \right].$$

## 7 Numerical Approximation Algorithms

In this section, we develop an approximation algorithm to compute the maximal cost-bounded reachability probability under late schedulers. In the following we fix a locally-uniform CTMDP  $\mathcal{M} = (L, \text{Act}, \mathbf{R}, \{\mathbf{w}_i\}_{1 \leq i \leq k})$ . Our numerical algorithm will achieve the following tasks:

**Input:** a set  $G \subseteq L$ , a state  $s \in L$ , a vector  $\mathbf{c} \in \mathbb{N}_0^k$  and an error bound  $\epsilon > 0$ ;

**Output:** a value  $x \in [0, 1]$  such that  $|\text{prob}_G^{\max}(s, \mathbf{c}) - x| \leq \epsilon$ .

For computational purposes, we assume that each  $\mathbf{w}_i(s)$  is an integer; rational numbers (and simultaneously the input cost bound vector) can adjusted to integers by multiplying a common multiplier of the denominators, without changing the maximal probability value to be approximated.

We base our approximation scheme on Theorem 3 and Proposition 2. In the following we fix a set  $G \subseteq L$ .

Below we illustrate the discretization and the approximation scheme for late schedulers. Note that  $\text{prob}_G^{\max}(s, \mathbf{c}) = 1$  whenever  $s \in G$  and  $\mathbf{c} \geq \mathbf{0}$ . Thus we do not need to incorporate those points into the discretization.

**Definition 18.** *Let  $\mathbf{c} \in \mathbb{N}_0^k$  and  $N \in \mathbb{N}$ . Define*

$$\text{Disc}(\mathbf{c}, N) := \{\mathbf{d} \in \mathbb{R}^k \mid \mathbf{0} \leq \mathbf{d} \leq \mathbf{c} \text{ and } N \cdot \mathbf{d}_i \in \mathbb{N}_0 \text{ for all } 1 \leq i \leq k\}.$$

The set  $\text{D}_N^{\mathbf{c}}$  of discretized grid points is defined as follows:

$$\text{D}_N^{\mathbf{c}} := (L - G) \times \text{Disc}(\mathbf{c}, N).$$

The following definition presents the approximation scheme on  $\text{D}_N^{\mathbf{c}}$ .

**Definition 19.** *Define  $X_G := (L - G) - Y_G$ . The approximation scheme  $\Upsilon_{\mathbf{c}, N}^G$  on  $\text{D}_N^{\mathbf{c}}$  consists of the following items:*

- exactly one rounding argument for each element of  $D_N^c$  ;
- a system of equations for elements in  $X_G \times \text{Disc}(\mathbf{c}, N)$  ;
- a linear program on  $Y_G$  for each  $\mathbf{d} \in \text{Disc}(\mathbf{c}, N)$  .

**Rounding Arguments:** For each element  $y \in D_N^c$ , the rounding argument for  $y$  is as follows:

$$\overline{\text{prob}}_G(y) := \frac{K}{N^2} \text{ if } \text{prob}_G(y) \in \left[ \frac{K}{N^2}, \frac{K+1}{N^2} \right) \text{ for some integer } 0 \leq K \leq N^2 .$$

**Equations:** The system of equations is described as follows. For all  $(s, \mathbf{d}) \in D_N^c$  with  $\mathbf{w}(s) \neq \mathbf{0}$  and  $\mathbf{d} - \frac{1}{N} \cdot \mathbf{w}(s) \geq \mathbf{0}$ , there is a linear equation

$$\frac{\text{prob}_G(s, \mathbf{d}) - \overline{\text{prob}}_G(s, \text{pre}(\mathbf{d}, s))}{\frac{1}{N}} = \max_{a \in \text{En}(s)} \sum_{s' \in L} \mathbf{R}(s, a, s') \cdot (\overline{\text{prob}}_G(s', \text{pre}(\mathbf{d}, s)) - \overline{\text{prob}}_G(s, \text{pre}(\mathbf{d}, s))) \quad (\text{E1})$$

where  $\text{pre}(\mathbf{d}, s) := \mathbf{d} - \frac{1}{N} \cdot \mathbf{w}(s)$ . For all  $(s, \mathbf{d}) \in D_N^c$  with  $\mathbf{w}(s) \neq \mathbf{0}$  and  $\mathbf{d} - \frac{1}{N} \cdot \mathbf{w}(s) \not\geq \mathbf{0}$ , there is a linear equation

$$\text{prob}_G(s, \mathbf{d}) = 0 . \quad (\text{E2})$$

**Linear Programs:** For each  $\mathbf{d} \in \text{Disc}(\mathbf{c}, N)$ , the collection  $\{\text{prob}_G(s, \mathbf{d})\}_{s \in Y_G}$  of values is the unique optimum solution of the linear program as follows:

- $\min \sum_{s \in Y_G} \text{prob}_G(s, \mathbf{d})$ , subject to:
- $\text{prob}_G(s, \mathbf{d}) \geq \sum_{s' \in L} \mathbf{P}(s, a, s') \cdot \text{prob}_G(s', \mathbf{d})$  for all  $s \in Y_G$  and  $a \in \text{En}(s)$ ;
  - $\text{prob}_G(s, \mathbf{d}) \in [0, 1]$  for all  $s \in Y_G$ ;

where the values  $\{\text{prob}_G(s, \mathbf{d})\}_{s \in X_G}$  are assumed to be known. All  $\text{prob}_G(s, \mathbf{d})$ 's and  $\overline{\text{prob}}_G(s, \mathbf{d})$  above with  $s \in G$  are predefined to be 1.

Generally,  $\text{prob}_G(s, \mathbf{d})$  approximates  $\text{prob}_G^{\max}(s, \mathbf{d})$  and  $\overline{\text{prob}}_G(s, \mathbf{d})$  approximates the same value with a rounding operation. A detailed computational sequence of the approximation scheme is described in Algorithm 1.

In principle, we compute the ‘‘higher’’ grid point  $\text{prob}_G(s, \mathbf{d} + \frac{1}{N} \cdot \mathbf{w}(s))$  by  $\text{prob}_G(s, \mathbf{d})$  and (E1), and then update other ‘‘higher’’ points through the linear program. The rounding argument is incorporated to avoid precision explosion caused by linear programming. The following proposition shows that Algorithm 1 indeed terminates after a finite number of steps.

**Proposition 3.** *Algorithm 1 terminates after a finite number of steps for all  $\mathbf{c} \in \mathbb{N}_0^k$  and  $N \in \mathbb{N}$ .*

The following theorem states that the approximation scheme really approximates  $\text{prob}_G^{\max}$ . To ease the notation, we shall use  $\text{prob}_G(s, \mathbf{d})$  or  $\overline{\text{prob}}_G(s, \mathbf{d})$  to denote both the variable at the grid point and the value it holds under the approximation scheme.

---

**Algorithm 1.** The Computation of  $\mathcal{Y}_{\mathbf{c},N}^G$  (for late schedulers)
 

---

- 1: Set all grid points in  $\{(s, \mathbf{d}) \in \mathbb{D}_N^{\mathbf{c}} \mid \mathbf{d} - \frac{1}{N} \cdot \mathbf{w}(s) \not\geq \mathbf{0}\}$  to zero by (E2);
  - 2: Compute all  $\text{prob}_G(s, \mathbf{d})$  that can be directly obtained through the linear program;
  - 3: Compute all  $\overline{\text{prob}}_G(s, \mathbf{d})$  that can be directly obtained by the rounding argument;
  - 4: Compute all  $\text{prob}_G(s, \mathbf{d})$  that can be directly obtained through (E1), and back to Step 2. until all grid points in  $\mathbb{D}_N^{\mathbf{c}}$  are computed;
- 

**Theorem 4.** Let  $\mathbf{c} \in \mathbb{N}_0^k$  and  $N \in \mathbb{N}$  with  $N \geq \mathbf{E}_{\max}$ . For each  $(s, \mathbf{d}) \in \mathbb{D}_N^{\mathbf{c}}$ ,

$$|\text{prob}_G(s, \mathbf{d}) - \text{prob}_G^{\max}(s, \mathbf{d})| \leq \left( \frac{2 \cdot \mathbf{E}_{\max}^2 \cdot \mathbf{w}_{\max}}{N \cdot \mathbf{w}_{\min}} + \frac{1}{N} \right) \cdot \left[ \sum_{i=1}^k \mathbf{d}_i \right] + \frac{\mathbf{E}_{\max}}{N}$$

and

$$\begin{aligned} |\overline{\text{prob}}_G(s, \mathbf{d}) - \text{prob}_G^{\max}(s, \mathbf{d})| \leq \\ \left( \frac{2 \cdot \mathbf{E}_{\max}^2 \cdot \mathbf{w}_{\max}}{N \cdot \mathbf{w}_{\min}} + \frac{1}{N} \right) \cdot \left[ \sum_{i=1}^k \mathbf{d}_i \right] + \frac{\mathbf{E}_{\max}}{N} + \frac{1}{N^2} \end{aligned}$$

From Theorem 4, we derive our approximation algorithm as follows.

**Corollary 1.** There exists an algorithm such that given any  $\epsilon > 0$ ,  $s \in L$ ,  $G \subseteq L$  and  $\mathbf{c} \in \mathbb{N}_0^k$ , the algorithm outputs a  $d \in [0, 1]$  which satisfies that  $|d - \text{prob}_G^{\max}(s, \mathbf{c})| \leq \epsilon$ . Moreover, the algorithm runs in

$$\mathcal{O}\left(\left(\max\left\{\mathbf{E}_{\max}, \frac{M}{\epsilon}\right\}\right)^k \cdot \left(\prod_{i=1}^k \mathbf{c}_i\right) \cdot \left(|\mathcal{M}| + \log \frac{M}{\epsilon}\right)^8\right)$$

time, where  $M := \left(2 \cdot \mathbf{E}_{\max}^2 \cdot \frac{\mathbf{w}_{\max}}{\mathbf{w}_{\min}} + 1\right) \cdot \left[\sum_{i=1}^k \mathbf{c}_i\right] + \mathbf{E}_{\max} + 1$  and  $|\mathcal{M}|$  is the size of  $\mathcal{M}$ .

*Proof.* The algorithm is an simple application of Theorem 4. If  $s \in G$ , the algorithm just returns 1. Otherwise, the algorithm just calls Algorithm 1 with  $N := \lfloor \max\{\mathbf{E}_{\max}, \frac{M}{\epsilon}\} \rfloor + 1$  and set  $d = \text{prob}_G(s, \mathbf{c})$ ; By Theorem 4, we directly obtain that  $|d - \text{prob}_G^{\max}(s, \mathbf{c})| \leq M \cdot \frac{1}{N}$ . For each  $\mathbf{d} \in \text{Disc}(\mathbf{c}, N)$ , the total computation of  $\{\text{prob}_G(s, \mathbf{d})\}_{s \in X_G \cup Y_G}$  and  $\{\overline{\text{prob}}_G(s, \mathbf{d})\}_{s \in X_G \cup Y_G}$  takes  $\mathcal{O}(|\mathcal{M}| + \log \frac{M}{\epsilon})^8$  time since the most time consuming part is the linear program which takes  $\mathcal{O}(|\mathcal{M}| + \log N)^8$  time (cf. [15]). Thus the total running time of the algorithm is  $\mathcal{O}\left(\left(\max\left\{\mathbf{E}_{\max}, \frac{M}{\epsilon}\right\}\right)^k \cdot \left(\prod_{i=1}^k \mathbf{c}_i\right) \cdot \left(|\mathcal{M}| + \log \frac{M}{\epsilon}\right)^8\right)$  since the size of  $\text{Disc}(\mathbf{c}, N)$  is  $\mathcal{O}(N^k \cdot \left(\prod_{i=1}^k \mathbf{c}_i\right))$ .

## 8 Conclusion

In this paper, we established an integral characterization for multi-dimensional maximal cost-bounded reachability probabilities in continuous-time Markov decision processes, the existence of deterministic cost-positional optimal scheduler

and an algorithm to approximate the cost-bounded reachability probability with an error bound, under the setting of both early and late schedulers. The approximation algorithm is based on a differential characterization of cost-bounded reachability probability which in turn is derived from the integral characterization. The error bound is obtained through the differential characterization and the Lipschitz property described. Moreover, the approximation algorithm runs in polynomial time in the size of the CTMDP and the reciprocal of the error bound, and exponential in the dimension of the unit-cost vector. An important missing part is the generation of an  $\epsilon$ -optimal scheduler. However, we conjecture that an  $\epsilon$ -optimal scheduler is not difficult to obtain given that the approximation scheme has been established.

A future direction is to determine an  $\epsilon$ -optimal scheduler, under both early and late schedulers. Besides, we believe that the paradigms developed in this paper can also be applied to minimum cost-bounded reachability probability and even stochastic games [12] with multi-dimensional cost-bounded reachability objective.

**Acknowledgement.** I thank Prof. Joost-Pieter Katoen for his valuable advices on the writing of the paper, especially for the Introduction part. I also thank anonymous referees for valuable comments.

## References

1. Baier, C., Haverkort, B.R., Hermanns, H., Katoen, J.P.: Reachability in continuous-time Markov reward decision processes. In: Flum, J., Grädel, E., Wilke, T. (eds.) *Logic and Automata. Texts in Logic and Games*, vol. 2, pp. 53–72. Amsterdam University Press (2008)
2. Baier, C., Haverkort, B.R., Hermanns, H., Katoen, J.P.: Performance evaluation and model checking join forces. *Commun. ACM* 53(9), 76–85 (2010)
3. Brázdil, T., Forejt, V., Kreál, J., Kretínský, J., Kučera, A.: Continuous-time stochastic games with time-bounded reachability. *Inf. Comput.* 224, 46–70 (2013)
4. Buchholz, P., Schulz, I.: Numerical analysis of continuous time Markov decision processes over finite horizons. *Computers & OR* 38(3), 651–659 (2011)
5. Fearnley, J., Rabe, M., Schewe, S., Zhang, L.: Efficient approximation of optimal control for continuous-time Markov games. In: Chakraborty, S., Kumar, A. (eds.) *FSTTCS. LIPIcs*, vol. 13, pp. 399–410. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2011)
6. Feller, W.: *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, New York (1966)
7. Fu, H.: Maximal cost-bounded reachability probability on continuous-time Markov decision processes. *CoRR* abs/1310.2514 (2013)
8. Hatefi, H., Hermanns, H.: Improving time bounded reachability computations in interactive Markov chains. In: Arbab, F., Sirjani, M. (eds.) *FSEN 2013. LNCS*, vol. 8161, pp. 250–266. Springer, Heidelberg (2013)
9. Neuhäuser, M.R.: *Model checking nondeterministic and randomly timed systems*. Ph.D. thesis, RWTH Aachen (2010)

10. Neuhäüßer, M.R., Stoelinga, M., Katoen, J.-P.: Delayed nondeterminism in continuous-time Markov decision processes. In: de Alfaro, L. (ed.) FOSSACS 2009. LNCS, vol. 5504, pp. 364–379. Springer, Heidelberg (2009)
11. Neuhäüßer, M.R., Zhang, L.: Time-bounded reachability probabilities in continuous-time Markov decision processes. In: QEST, pp. 209–218. IEEE Computer Society (2010)
12. Prieto-Rumeau, T., Hernández-Lerma, O.: Selected Topics on Continuous-Time Controlled Markov Chains and Markov Games. Imperial College Press, London (2012)
13. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming, 1st edn. John Wiley & Sons, Inc., New York (1994)
14. Rabe, M.N., Schewe, S.: Finite optimal control for time-bounded reachability in CTMDPs and continuous-time Markov games. *Acta Inf.* 48(5-6), 291–315 (2011)
15. Schrijver, A.: Theory of Linear and Integer Programming. John Wiley & Sons, Inc., New York (1986)
16. Wolovick, N., Johr, S.: A characterization of meaningful schedulers for continuous-time Markov decision processes. In: Asarin, E., Bouyer, P. (eds.) FORMATS 2006. LNCS, vol. 4202, pp. 352–367. Springer, Heidelberg (2006)