Abstract—This paper solves the problem of computing the maximum (and minimum) probability to reach a set of goal states within a given time bound in continuous-time Markov decision processes (CTMDPs). For the subclass of locally uniform CTMDPs, we define the class of late total time positional schedulers (\(TTPD_l\)) and prove that they suffice to resolve all non-deterministic choices in an optimal way. Our main contribution is a discretization technique which, for an a priori given error bound \(\varepsilon > 0\), induces a discrete-time MDP that approximates the maximum time-bounded reachability probability in the underlying locally uniform CTMDP up to \(\varepsilon\).

In a second part, we consider arbitrary CTMDPs. In this more general setting, \(TTPD_l\) schedulers are inapplicable and are replaced by the corresponding class of early \(TTPD\) schedulers. Using a measure preserving transformation from CTMDPs to interactive Markov chains (IMCs), we apply recent results on IMCs to compute the maximum time-bounded reachability probability under early scheduler in the CTMDP's induced IMC.

Index Terms—Markov processes; nondeterminism; continuous time; reachability probability; discretization; uniformization

I. INTRODUCTION

Continuous-time Markov decision processes (CTMDPs) [1], [2], [3], [4], [5] are a stochastic model which allows for non-determinism between transitions whose delay is governed by negative exponential distributions. As such, CTMDPs extend continuous-time Markov chains (CTMCs) [6], [7] with non-deterministic choices and discrete-time Markov decision processes (MDPs) [2] with exponentially distributed delays.

As CTMDPs in general exhibit non-determinism, their induced stochastic process is not uniquely determined. Therefore, we follow the MDP approach [2] and define schedulers that resolve the CTMDP's non-deterministic choices: Depending on the trajectory that led into the current state, a scheduler returns a probability distribution over the available actions and thereby resolves the action-nondeterminism in that state. Accordingly, the stochastic behavior of a CTMDP is described by upper and lower probability bounds induced by a given (usually uncountable) class of schedulers.

In general, the sojourn time distribution of the current state depends on the action that is chosen by the scheduler. This dependency requires the scheduler to decide early, that is, when entering the current state. Accordingly, we refer to such schedulers as early schedulers. However, locally uniform CTMDPs — which share the property that their states’ residence time distributions do not depend on the scheduler’s choice — allow for even more powerful schedulers: As shown in [3], local uniformity allows us to delay the scheduling decision until the current state is left; the resulting late schedulers, which are well-defined only for locally uniform CTMDPs, perform at least as good as any early scheduler and generally induce strictly better probability bounds [3].

By first restricting ourselves to locally uniform CTMDPs, we solve the time-bounded reachability problem in that we compute the maximum probability to reach a set \(G\) of goal states within a given time bound \(z\) under all late schedulers. More precisely, we characterize the maximum time-bounded reachability probability as the least fixed point of a higher-order operator which involves integration over the time domain. Exploiting this result, we prove that for time-bounded reachability, it suffices to consider late total time positional deterministic schedulers (\(TTPD_l\)) which base their decision only on the elapsed time and on the current state. This allows us to reduce the problem of computing time-bounded reachability probabilities in locally uniform CTMDPs to the probability of computing step-bounded reachability probabilities in discrete-time MDPs. Specifically, we show how to approximate the behavior of the locally uniform CTMDP up to an a priori specified error bound \(\varepsilon > 0\) by defining its discretized MDP such that its maximum step-bounded reachability probability coincides (up to \(\varepsilon\)) with the maximum time-bounded reachability probability of the underlying locally uniform CTMDP. Computing the maximum step-bounded reachability probability in MDPs is a well-studied problem [2] and can be done efficiently, e.g. by value iteration algorithms [8]. Furthermore, a small extension of the value iteration algorithm allows us to automatically synthesize the \(\varepsilon\)-optimal scheduler which induces the maximum time-bounded reachability probability.

In the second part of the paper we turn our attention to the problem of computing time-bounded reachability probabilities for general CTMDPs. In this setting, late schedulers are not applicable. Hence, we resort to early schedulers and introduce a measure preserving transformation from arbitrary CTMDPs to interactive Markov chains (IMCs). In [9], the time-bounded reachability problem has been solved for IMCs. Hence, the maximum (and minimum) time-bounded reachability probabilities for early schedulers and general CTMDPs can be computed by analyzing the CTMDPs’ induced IMCs.

In both cases, the complexity is in \(O(m \cdot (\lambda z)^2 / \varepsilon)\), where \(m\) denotes the size of the input model, \(\lambda\) is its maximal exit rate and \(z\) is the given time bound.
Uniformization has successfully been applied to compute transient and steady state probabilities in CTMCs. In the paper, we shortly discuss the subtle differences when applying uniformization in the setting of early and late schedulers for CTMDPs. More precisely, we prove that it is sound (w.r.t. maximum reachability probabilities) for locally uniform CTMDPs and late schedulers, while it is incorrect for early schedulers.

Although we present all results only for maximum time-bounded reachability probabilities, they equally hold for the dual problem of determining the minimum time-bounded reachability probability. The reachability analysis is the key ingredient to enable approximate model checking of CTMDPs with respect to logics like CSL [10]. Note however, that for CSL, it is necessary to consider time-interval (instead of time-bounded) reachability. The necessary steps to achieve this are quite technical and have been elaborated for IMCs in [9],[11].

Organization. Section II introduces CTMDPs. In Sec. III we provide a fixed point characterization of the maximum time-bounded reachability probability in locally uniform CTMDPs. In Sec. IV, we develop the discretization algorithm for locally uniform CTMDPs and late schedulers. Section V solves the problem to compute time-bounded reachability probabilities under early schedulers in arbitrary CTMDPs. In Sec. VI, we discuss uniformization for CTMDPs and time-dependent schedulers. To show the applicability of our approach, Sec. VII provides an example where the discretization from Sec. IV is applied to solve the stochastic job scheduling problem. Finally, Sec. VIII discusses related work and concludes the paper.

All proofs can be found in [11, Chapter 5].

II. CONTINUOUS-TIME MARKOV DECISION PROCESSES

Let \( X \) be a finite set. Probability distributions over \( X \) are functions \( \mu : X \rightarrow [0,1] \) with \( \sum_{x \in X} \mu(x) = 1 \). If \( \mu(x) = 1 \) for some \( x \in X \), \( \mu \) is degenerate, denoted \( \mu = \{ x \mapsto 1 \} \); in this case, we identify \( \mu \) and \( x \). The set of all probability distributions over \( X \) is denoted \( \text{Distr}(X) \).

Definition 1 (Continuous-time Markov decision process) A continuous-time Markov decision process (CTMDP) is a tuple \( \mathcal{C} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \nu) \) where \( \mathcal{S} \) and \( \mathcal{A} \) are finite sets of states and actions, \( \mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}_{\geq 0} \) is a three-dimensional rate matrix and \( \nu \in \text{Distr}(\mathcal{S}) \) is an initial distribution.

Intuitively, a CTMDP behaves as follows: If \( \mathcal{R}(s, \alpha, s') = \lambda \) and \( \lambda > 0 \), an \( \alpha \)-transition leads from state \( s \) to state \( s' \). Here, \( \lambda \) is the rate of an exponential distribution which governs the transition’s delay. If action \( \alpha \) is chosen, the transition executes in time interval \([a,b]\) with probability \(\eta_\alpha((a,b)) = \int_a^b \lambda e^{-\lambda t} \, dt \). If multiple \( \alpha \)-transitions emanate a state \( s \), a race condition occurs: The \( \alpha \)-transition whose delay expires first executes and thereby determines the sojourn time of state \( s \) under action \( \alpha \). As the minimum of independent exponentials is again exponentially distributed with the sum of their rates, the sojourn time of state \( s \) under action \( \alpha \) is distributed with exit rate \( E(s, \alpha) = \sum_{s' \in \mathcal{S}} \mathcal{R}(s, \alpha, s') \).

An action \( \alpha \) is enabled in state \( s \) iff \( E(s, \alpha) > 0 \); moreover, the set \( \mathcal{A}(s) = \{ \alpha \in \mathcal{A} \mid E(s, \alpha) > 0 \} \) is the set of enabled actions in state \( s \) and determines the available nondeterministic choices in state \( s \). A state \( s \) is absorbing if \( \mathcal{A}(s) = \emptyset \). Obviously, adding self-loops to absorbing states does not change the time-bounded reachability probabilities; thus we can safely restrict to CTMDPs without absorbing states. A CTMDP \( \mathcal{C} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \nu) \) induces the embedded discrete time MDP \( \mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \nu) \) where the (time-abstract) transition probability matrix \( \mathcal{P} \) satisfies \( \mathcal{P}(s, \alpha, s') = \frac{\mathcal{R}(s, \alpha, s')}{E(s, \alpha)} \) if \( E(s, \alpha) > 0 \) and 0, otherwise. In fact, \( \mathcal{P}(s, \alpha, s') \) is the time-abstract probability to move from state \( s \) to state \( s' \) if action \( \alpha \) is selected in state \( s \). For the initial distribution \( \nu \), we use \( \nu_s = \{ s \mapsto 1 \} \) to denote a single initial state.

Example 1 Consider the CTMDP in Fig. 1. If \( \alpha \) is chosen in state \( s_1 \), the \( \alpha \)-transitions to states \( s_4 \) and \( s_5 \) compete for execution. The sojourn time in state \( s_1 \) is exponentially distributed with rate \( E(s_1, \alpha) = 1 \). Moreover, the probability that the \( \alpha \)-transition to \( s_4 \) executes before the \( \alpha \)-transition to \( s_5 \) is \( \mathcal{P}(s_1, \alpha, s_4) = \frac{\mathcal{R}(s_1, \alpha, s_4)}{E(s_1, \alpha)} = \frac{1}{2} \).

A CTMDP only defines a stochastic process if the nondeterministic choices between actions are resolved by a scheduler. As shown in [3], locally uniform CTMDPs allow to define schedulers that cannot be defined for general CTMDPs and which perform strictly better, that is, they induce higher time-bounded reachability probabilities than general schedulers.

Definition 2 (Local uniformity) A CTMDP \( \mathcal{C} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \nu) \) is locally uniform if \( \forall s \in \mathcal{S}, \forall \alpha, \beta \in \mathcal{A}(s) \). \( E(s, \alpha) = E(s, \beta) \).

Local uniformity ensures that the sojourn time in any state does not depend on the action that is chosen in that state. Hence, we may use \( E(s) = E(s, \alpha) \) for some \( \alpha \in \mathcal{A}(s) \) to denote the exit rate of state \( s \). Note that the CTMDP in Fig. 1 is locally uniform.

Let us briefly recall some measure-theoretic concepts that we use in the remainder of the paper; for a more detailed discussion, we refer to [11],[12],[13].

A. The probability space

Finite paths in a CTMDP \( \mathcal{C} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \nu) \) are sequences \( \pi = s_0 \xrightarrow{t_0, \alpha_0} s_1 \xrightarrow{t_1, \alpha_1} \cdots \xrightarrow{t_{n-1}, \alpha_{n-1}} s_n \) where \( s_i \in \mathcal{S} \) is a state, \( t_i \in \mathbb{R}_{\geq 0} \) is the sojourn time in state \( s_i \) and \( \alpha_i \in \mathcal{A} \) is an action. The length of \( \pi \) (denoted \( |\pi| \)) is \( n \); accordingly, \( \pi_\downarrow = s_n \) is the last state of \( \pi \). Further, \( \Delta(\pi) = \sum_{k=0}^{n} t_k \) is the total time spent on \( \pi \). Note that as in [3], [13] we do not exclude invalid paths, i.e. we do not require \( \mathcal{P}(s_i, \alpha_i, s_{i+1}) > 0 \). As we shall see in Sec. II-C, this is justified by the fact that such paths are assigned probability zero.

Any path \( \pi \) is a concatenation of a state with a sequence of combined transitions from the set \( \Omega = \mathbb{R}_{\geq 0} \times \mathcal{A} \times \mathcal{S} \); hence,
Paths\(n\) is a set of paths of length \(n\) in \(C\); further, let Paths\(^*\)\((C)\), Paths\(^*\)\((C)\) and Paths\((C)\) denote the sets of finite, infinite, and all paths of \(C\). To simplify notation, we omit the reference to \(C\) wherever possible. Events in \(C\) are measurable sets of paths; as paths are Cartesian products of combined transitions, we first define the \(\sigma\)-field \(\mathfrak{F}_\text{Act} = \sigma(\mathfrak{B}_0 \times \mathfrak{B}_\text{Act} \times \mathfrak{B}_\text{Act})\) on subsets of \(\Omega\) where \(\mathfrak{B}_0\) denotes the Borel \(\sigma\)-field over \(\mathbb{R}_0\) and \(\mathfrak{B}_\text{Act} = 2^{\text{Act}}\). Based on \((\Omega, \mathfrak{F})\), we derive the product \(\sigma\)-field \(\mathfrak{F}_\text{Paths} = \sigma(\{S_0 \times M_0 \times \cdots \times M_{n-1} \mid S_0 \in \mathfrak{F}_\text{Act}, M_i \in \mathfrak{F}_\text{Act}\})\) of measurable subsets of Paths\(^*\). Finally, the cylinder-set construction [12] extends this uniquely to a \(\sigma\)-field over infinite paths: A set \(B \in \mathfrak{F}_\text{Paths}^n\) is called a base of an infinite cylinder \(C\), if it holds \(C = Cyl(B) = \{\pi \in \text{Paths}^* \mid \pi[0..n] \in B\}\), where \(\pi[0..n]\) denotes the prefix of length \(n\) of path \(\pi\). Now the desired \(\sigma\)-field \(\mathfrak{F}_\text{Paths}^*\) is generated by the set of cylinders, i.e. \(\mathfrak{F}_\text{Paths}^* = \sigma(\bigcup_{n=0}^\infty \{Cyl(B) \mid B \in \mathfrak{F}_\text{Paths}^n\})\).

### B. Schedulers

In order to reason about probabilities in CTMDPs, we introduce schedulers that resolve the nondeterminism that occurs in states with multiple enabled actions: If Act\((s) = \{\alpha_1, \ldots, \alpha_n\}\), a scheduler resolves the action-nondeterminism by providing a probability distribution over Act\((s)\). We recall the classical early schedulers which immediately decide on an action upon entering a state [11, 12, 13]. In locally uniform CTMDPs, any state’s sojourn time distribution is independent of the selected action. In this setting, we define a strictly more powerful class of schedulers [3] where the scheduler decision is delayed up to the point when the current state is left. Accordingly, the class of late schedulers additionally incorporates the current state’s sojourn time when making a decision:

**Definition 3 (Schedulers)** Let \(C = (S, \text{Act}, R, \nu)\) be a CTMDP.

- A mapping \(D: \text{Paths}^* \times \mathfrak{F}_\text{Act} \to [0,1]\) is a generic measurable scheduler (GM-scheduler) for \(C\) iff \(D(\pi, \cdot) \in \text{Distr}(\text{Act}(\pi_1))\) for all \(\pi \in \text{Paths}^*\) and if the functions \(D(\cdot, A) : \text{Paths}^* \to [0,1]\) are measurable for all \(A \in \mathfrak{F}_\text{Act}\).
- If \(C\) is locally uniform, a mapping \(D: \text{Paths}^* \times \mathbb{R}_0 \times \mathfrak{F}_\text{Act} \to [0,1]\) is a measurable late scheduler (ML-scheduler) for \(C\) iff \(D(\pi, t, \cdot) \in \text{Distr}(\text{Act}(\pi_1))\) for all \(t \in \mathbb{R}_0\) and \(\pi \in \text{Paths}^*\) and if the functions \(D(\cdot, \cdot, A) : \text{Paths}^* \times \mathbb{R}_0 \to [0,1]\) are measurable for all \(A \in \mathfrak{F}_\text{Act}\).

The measurability condition in Def. 3 excludes schedulers which resolve the nondeterminism in a way that induces non-measurable sets (like Vitali sets, see [12, p.34]), i.e., sets of paths that cannot be assigned a probability.

Let \(\pi\) be a finite path that ends in state \(s\) with \(|\text{Act}(s)| > 1\). Upon entering state \(s\), a GM-scheduler \(D_e\) yields a probability distribution \(D_e(\pi, \cdot)\) over the available actions in Act\((s)\) and thereby resolves the nondeterminism in state \(s\) for history \(\pi\). In contrast, an ML-scheduler \(D_l\) decides upon leaving state \(s\): Here, the probability distribution \(D_l(\pi, t, \cdot)\) additionally depends on the time \(t\) that has been spent in state \(s\). Hence, we refer to GM-schedulers as early schedulers whereas ML-schedulers are referred to as late schedulers.

A scheduler which only yields degenerate distributions is called deterministic; otherwise, it is called randomized. Further, \(D\) is time-abstract, if its decisions are independent of the timing information in \(\pi\) and (for ML-schedulers) the time spent in the last state. Where applicable, we use the term measurable schedulers to refer to the union \(GM \cup ML\). Moreover, we identify the degenerate distribution \(\{0 \mapsto 1\} \in \text{Distr}(\text{Act})\) with the action \(\alpha \in \text{Act}\).

### C. Probability measures

Given a CTMDP \(C\), any measurable scheduler \(D\) induces a unique stochastic process on \(C\). To define the induced probability measure on the measurable space (Paths\(^*\), \(\mathfrak{F}_\text{Paths}^*\)), we follow [13], [3] and start by deriving the probability of measurable sets of combined transitions:

**Definition 4 (Probability of combined transitions)** Let \(C = (S, \text{Act}, R, \nu)\) be a CTMDP. \(D\) a measurable scheduler and \(\pi \in \text{Paths}^*\). We define \(\mu_D(\pi, \cdot) : \mathfrak{F}_\pi \to [0,1]\) as follows:

\[
\text{If } D \in GM, \text{ then } \mu_D(\pi, M) = \int_{\text{Act}} D(\pi, \alpha) \int_{\mathbb{R}_0} \eta_{E(\pi, \alpha)}(dt) \int_{\mathfrak{F}_\text{Act}} I_M(t, \alpha, s') P(s, \alpha, ds').
\]

\[
\text{If } D \in ML, \text{ then } \mu_D(\pi, M) = \int_{\mathbb{R}_0} \eta_{E(\pi, 1)}(dt) \int_{\text{Act}} D(\pi, t, \alpha) \int_{\mathfrak{F}_\text{Act}} I_M(t, \alpha, s') P(s, \alpha, ds').
\]

Let us first explain the definition for ML-schedulers: Here, \(\eta_{E(\pi, 1)}\) is the exponential distribution of the sojourn time \(t\) of state \(\pi_1\) which has rate \(E(\pi_1)\) (recall that for ML-schedulers, the underlying CTMDP is locally uniform); further, \(I_M\) is the characteristic function of \(M \in \mathfrak{F}\). In fact, \(\mu_D(\pi, M)\) is the probability to continue with some combined transition in \(M\), given that we hit the current state \(\pi_1\) along the trajectory \(\pi\).

Note that for GM-schedulers, the first two integrals in Def. 4 must be swapped, as for general CTMDPs, the exit rate \(E(s, \cdot)\) depends on the chosen action [3].

Based on the probability measures \(\mu_D(\pi, \cdot)\), we can now define the probabilities of measurable sets of paths. Let \(C = (S, \text{Act}, R, \nu)\) be a CTMDP and \(D\) a measurable scheduler.

**Definition 5 (Probability measure)** For \(n \in \mathbb{N}\), we define the probability measures \(P_{\nu, D}^{n+1}\) on (Paths\(^n\), \(\mathfrak{F}_\text{Paths}^n\)) inductively:

\[
\text{For } n = 0 \text{ we set } P_{\nu, D}^{0} : \mathfrak{F}_\text{Paths}^n \to [0,1] : \Pi \mapsto \sum_{\alpha \in \text{Act}} \nu(s)\Pi \\
\text{and } P_{\nu, D}^{n+1} : \mathfrak{F}_\text{Paths}^{n+1} \to [0,1] : \\
\text{Pi} \mapsto \int_{\text{Paths}^n} P_{\nu, D}^{n}(d\pi) \int_{\mathfrak{F}_\text{Act}} I_{\Pi}(\pi \circ m) \mu_D(\pi, dm).
\]

For the base case, note that \(\Pi \in \mathfrak{F}_\text{Paths}^0\); hence, \(\Pi \subseteq S\) and we add the probabilities to start in one of the states in \(\Pi\). In the induction step, \(P_{\nu, D}^{n+1}\) measures a set of paths \(\Pi\) of length \(n+1\) by multiplying the probabilities \(P_{\nu, D}^{n}(dm)\) of path prefixes \(\pi\) (of length \(n\)) with the probability \(\mu_D(\pi, dm)\) of a combined transition \(m \in \mathcal{M}\) which extends \(\pi\) to a path in \(\Pi\).
Together, the measures $P^\nu_{\pi,D}$ extend to a unique measure on $\mathfrak{F}_\text{Paths}^\omega$: If $B \in \mathfrak{F}_\text{Paths}^\omega$ is a measurable base and $C = \text{Cyl}(B)$, we define $P^\nu_{\pi,D}(C) = P^\nu_{\pi,D}(B)$. Due to the inductive definition of $P^\nu_{\pi,D}$, the Ionescu–Tulcea extension theorem [12, Thm. 2.7.2] is applicable which yields a unique extension of $P^\nu_{\pi,D}$ from cylinders to sets in $\mathfrak{F}_\text{Paths}^\omega$.

Reasoning about a CTMDP’s behavior is relative to an a priori fixed class of schedulers. In this paper, we aim at computing the upper bounds on the probability to reach a set $G \subseteq \mathcal{S}$ of goal states within a given time bound $z$ (denoted $\bigtriangleup [0,z]G$) w.r.t. the classes of ML- and GM-schedulers.

**Definition 6 (Maximum reachability)** Let $\mathcal{C}$ be a CTMDP (locally uniform CTMDP) with state space $\mathcal{S}$ and let $\mathfrak{O} = \text{GM (or locally uniform CTMDP)}$ be a scheduler class. For $G \subseteq \mathcal{S}$, $s \in \mathcal{S}$ and $z \in \mathbb{R}_{\geq 0}$, $p^{C,\mathfrak{O},G}_{\text{max}} : \mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1] : (s, z) \mapsto \sup_{D \in \mathfrak{O}} P^\omega_{\nu,D}(\bigtriangleup [0,z]G)$ is the maximum time-bounded reachability for the scheduler class $\mathfrak{O}$, the set $G$ of goal states and time bound $z$.

The functions $p^{C,\mathfrak{O},G}_{\text{max}}$ share some mathematically nice properties that we need in the forthcoming sections:

**Lemma 1** The functions $p^{C,\mathfrak{O},G}_{\text{max}}(s, \cdot) : \mathbb{R}_{\geq 0} \to [0,1]$ are continuous and measurable for all $s \in \mathcal{S}$.

From now on, we omit the superscripts $C$, $\mathfrak{O}$ and $G$ of $p^{C,\mathfrak{O},G}_{\text{max}}$ whenever they are clear from the context.

### III. A FIXED POINT CHARACTERIZATION FOR TIME-BOUNDED REACHABILITY

In this section we consider locally uniform CTMDPs and the class of ML-schedulers and characterize the function $p_{\text{max}}$ as the least fixed point of a higher order operator $\Omega$.

**Theorem 1 (Fixed point characterization)** Let $\mathcal{C}$ be a locally uniform CTMDP with $\mathcal{S}$ and $\mathfrak{A}$ as before, and let $G \subseteq \mathcal{S}$ be a set of goal states. Then $p^{C,\mathfrak{ML},G}_{\text{max}}$ is the least fixed point of the higher-order operator $\Omega : (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1]) \to (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1])$ which is defined for $s \in \mathcal{S}$, $z \in \mathbb{R}_{\geq 0}$, and measurable function $F : \mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1]$ such that $\Omega(F)(s,z) = 1$ if $s \in G$ and for $s \notin G$, $\Omega(F)(s,z)$ equals:

$$\int_0^z E(s)e^{-E(s)t} \max_{\alpha \in \mathfrak{A}} \sum_{s' \in \mathcal{S}} P(s, \alpha, s') \cdot F(s', z-t) \, dt. \quad (1)$$

The term $E(s)e^{-E(s)t}$ in Eq. (1) is the density of the sojourn time in state $s$; accordingly, if state $s$ is left at time $t$, we multiply with the maximum (over $\alpha \in \mathfrak{A}$) probability to reach a goal state via some action $\alpha$ in the remaining $z-t$ time units. Briefly, the proof of Thm. 1 is split into two parts: First, we show that $p_{\text{max}}$ is a fixed point of $\Omega$. Second, we prove that $p_{\text{max}}$ is the least fixed point of $\Omega$ by decomposing the event $\bigtriangleup [0,z]G$ w.r.t. the number $n$ of transitions that are needed to reach a state in $G$. By induction on $n$, we then prove that $p_{\text{max}}(s,z) \leq F(s,z)$ for any other fixed point $F$ of $\Omega$ and all $s \in \mathcal{S}$ and $z \in \mathbb{R}_{\geq 0}$.

Theorem 1 implies that the maximum reachability probabilities are characterized by a set of Volterra integral equations [14], [15]. As for CTMCs [16], it can be solved by recursively solving the integrations, which is however reported to be time consuming and unstable. Therefore we resort to Thm. 1 and approximate the maximum reachability probability by discretization that we develop in Sec. IV. In the literature, similar fixed point characterizations have been used to derive the probability of CSL path formulas in CTMCs [16] and for timed automata specifications [17].

### A. Late total time positional schedulers

Whereas ML-schedulers take the complete history of the CTMDP into account, **late total time positional schedulers** [18], [3] ($\text{TPPR}_1$) base their decision only on the current state and the total elapsed time $t_{\text{past}}$. As we will see, $\text{TPPR}_1$ schedulers suffice to maximize time-bounded reachability objectives.

**Definition 7 (Late total-time positional scheduler)** Let $\mathcal{C}$ be a locally uniform CTMDP and $D \in \text{ML}$. The scheduler $D$ is a late-total-time positional randomized ($\text{TPPR}_1$) scheduler iff for all $\pi_1, \pi_2 \in \text{Paths}^\omega$ and for all $t_1, t_2 \in \mathbb{R}_{\geq 0}$ it holds that

$$\pi_1 \downarrow = \pi_2 \downarrow \land (\Delta(\pi_1) + t_1 = \Delta(\pi_2) + t_2) \implies D(\pi_1, t_1) = D(\pi_2, t_2).$$

A $\text{TPPR}_1$ scheduler is a deterministic $\text{TPPR}_1$ scheduler. Intuitively, a $\text{TPPR}_1$ scheduler yields the same distribution for trajectories $\pi_1$ and $\pi_2$ if $\pi_1$ and $\pi_2$ end in the same state (the current state) and if both states are left after the same amount of time ($\Delta(\pi_1) + t_1 = \Delta(\pi_2) + t_2$, resp.) has passed.

To maximize time-bounded reachability objectives, we define a $\text{TPPR}_1$ scheduler $D^z$ as follows:

**Definition 8 (The scheduler $D^z$)** Let $\mathcal{C} = (\mathcal{S}, \mathfrak{A}, \mathfrak{R}, \nu)$ be a locally uniform CTMDP, $G \subseteq \mathcal{S}$ a set of goal states and $z \in \mathbb{R}_{\geq 0}$ a time bound. Given an arbitrary (fixed) total order $<_{\mathfrak{A}}$ on $\mathfrak{A}$, we define the mapping $D^z$ such that for all $s \in \mathcal{S}$ and $t_{\text{past}} \leq z$:

$$D^z(s, t_{\text{past}}) = \min_{\beta \in \mathfrak{A} \setminus \{\alpha \in \mathfrak{A} \mid \nu(\beta) < \nu(\alpha)\}} \left\{ f(s, z - t_{\text{past}}, \beta) \leq f(s, z - t_{\text{past}}, \alpha) \right\},$$

where $f(s, z', \gamma) = \sum_{s' \in \mathcal{S}} P(s, \gamma, s') \cdot p_{\text{max}}(s', z')$. If $t_{\text{past}} > z$, set $D^z(s, t_{\text{past}}) = \min_{\alpha \in \mathfrak{A}} f(s, \alpha)$.

As the following lemma shows, the schedulers $D^z$ are measurable; hence, they are indeed $\text{TPPR}_1$ schedulers.

**Lemma 2** The schedulers $D^z$ are measurable for all $z \in \mathbb{R}_{\geq 0}$.

This follows directly from the measurability of the function $f$, which can be proved using Lemma 1.

Intuitively, in Def. 8, the function $f(s, z - t_{\text{past}}, \beta)$ denotes the maximum probability to reach a state in $G$ within the remaining $z-t_{\text{past}}$ time units via action $\beta$ for the case that $t_{\text{past}}$ time units have expired on the path that led to state $s$ and in state $s$ itself. However, multiple actions $\alpha$ may exist that maximize $f(s, z - t_{\text{past}}, \alpha)$. Hence, we fix some total order $<_{\mathfrak{A}}$ to ensure uniqueness of $D^z$. Note that Def. 8 implies that $D^z(s, t_{\text{past}} + t) = D^{z - t_{\text{past}}}(s, t)$ for all $s \in \mathcal{S}$, $t, z \in \mathbb{R}_{\geq 0}$ and $t_{\text{past}} \leq z$.
With these preliminaries, we prove that the schedulers $D^z$ maximize the probability of reaching $G$ within at most $z$ time units for any initial state $s$.

**Theorem 2 (Optimality)** Let $C$ be a locally uniform CTMDP with state space $S$, $G \subseteq S$ a set of goal states, $s \in S$ an initial state and $z \in \mathbb{R}_{\geq 0}$ a time bound. Then

$$P_{v_s^w, D^z} \left( \bigcirc^{[0,z]} G \right) = P_{\text{max}}^{C, ML, G} (s, z).$$

**B. Piecewise constant schedulers**

By Thm. 2, it suffices to consider $D^z$ to compute $P_{\text{max}}$. In particular, we do not need to consider the entire class of ML-schedulers. However, $D^z$ and TTPD$_1$ schedulers in general, are still continuous in their second argument. To obtain schedulers with a finite representation, we introduce piecewise constant TTPD$_1$ schedulers:

**Definition 9 (Piecewise constant TTPD$_1$ schedulers)** Let $C$ be a locally uniform CTMDP with state space $S$ and let $D : S \times \mathbb{R}_{\geq 0} \rightarrow \text{Act}$ be a TTPD$_1$ scheduler. $D$ is piecewise constant iff for all $s \in S$ and $\alpha \in \text{Act}(s)$ there exist disjoint intervals $A_{s, \alpha}^0, A_{s, \alpha}^1, A_{s, \alpha}^2, \ldots \subseteq \mathbb{R}_{\geq 0}$ such that for all $t_{past} \in \mathbb{R}_{\geq 0}$:

$$D(s, t_{past}) = \alpha \iff t_{past} \in \bigcup_{\varepsilon > 0} A_{s, \alpha}^\varepsilon.$$ 

A piecewise constant scheduler $D$ is non-Zeno if $\{ A_{s, \alpha}^\varepsilon \, | \, \inf A_{s, \alpha}^\varepsilon < z \} < \infty$ for all $s \in S$, $\alpha \in \text{Act}$ and $z \in \mathbb{R}_{\geq 0}$.

We use $P_C$ to denote the set of all piecewise constant and non-Zeno TTPD$_1$ schedulers. Intuitively, for state $s \in S$ and a given time-bound $z$, a $P_C$-scheduler changes its decision for an action only finitely many times: The intervals $A_{s, \alpha}$ in Def. 9 describe the time-periods, in which the scheduler chooses action $\alpha$ constantly if the current state is $s$. The non-Zeno assumption implies that only finitely many decision epochs occur up to time $z$.

**Theorem 3 (Optimality of PCD schedulers)** For locally uniform CTMDPs and ML-schedulers, it holds that

$$P_{\text{max}}^{C, ML, G} (s, z) = \sup_{D \in P_C} P_{v_s^w, D} \left( \bigcirc^{[0,z]} G \right).$$

For the discretization that we shall use later, we need to consider time intervals of equal length; hence, we further restrict to $\tau$-schedulers:

**Definition 10 ($\tau$-scheduler)** A scheduler $D \in P_C$ is a $\tau$-scheduler iff for all $s \in S$ and $k \in \mathbb{N}$: $\exists \alpha \in \text{Act}(s), \forall t_{past} \in [k\tau, (k+1)\tau), D(s, t_{past}) = \alpha$.

Any $P_C$-scheduler is a $\tau$-scheduler if its choices are constant on intervals of length $\tau$. As it turns out, the probabilities induced by $P_C$ and by $\tau$-schedulers converge for small $\tau$:

**Theorem 4 (Limiting $\tau$-scheduler)** Let $C$ be a locally uniform CTMDP, $G \subseteq S$ a set of goal states, $s \in S$ an initial state and $z \in \mathbb{R}_{\geq 0}$ a time bound. Then for any $D \in P_C$, there exist $\tau$-schedulers $D_{\tau}$ such that

$$\lim_{\tau \to 0} P_{v_s^w, D_{\tau}} \left( \bigcirc^{[0,z]} G \right) = P_{v_s^w, D} \left( \bigcirc^{[0,z]} G \right).$$

**IV. TIME-BOUNDED REACHABILITY PROBABILITIES IN LOCALLY UNIFORM CTMDPs**

In this section, we show how to compute the maximum time-bounded reachability probability for ML-schedulers and locally uniform CTMDPs. To this aim, we reduce this problem to the problem of maximizing the step-bounded reachability probability in (discrete-time) MDPs. The latter is a well-studied problem which can be solved efficiently, e.g. by value iteration algorithms [8]. The discretization we use for our reduction is defined such that it is exact up to an a priori given error bound $\varepsilon > 0$; hence, the results can be made arbitrarily precise. We study the complexity of our approach and show how to synthesize $\varepsilon$-optimal $\tau$-schedulers automatically.

**A. Reduction to step-bounded reachability in MDPs**

Let $C$ be a locally uniform CTMDP and $G \subseteq S$, $s \in S$ and $z \in \mathbb{R}_{\geq 0}$ as before. We aim at computing $P_{\text{max}}^{C, ML, G} (s, z)$ up to an a priori fixed error $\varepsilon > 0$. If $s \in G$, this is trivial as $P_{\text{max}} (s, z) = 1$ for all $z \in \mathbb{R}_{\geq 0}$. To compute $P_{\text{max}} (s, z)$ for $s \notin G$, we use the fixed point characterization of $P_{\text{max}}$ from Thm. 1. More precisely, we consider the first sub-interval $[0, \tau]$ of the integral in Eq. (1) separately, and split the integral accordingly:

$$P_{\text{max}} (s, z) = \Omega (P_{\text{max}} (s, z)) = \int_0^\tau E(s) e^{-E(s) t} \cdot \max_{\alpha \in \text{Act}} \left( s, s' \right) \cdot P_{\text{max}}^G (s', z - t) \, dt + \int_\tau^z E(s) e^{-E(s) t} \cdot \max_{\alpha \in \text{Act}} \left( s, s' \right) \cdot P_{\text{max}}^G (s', z - t) \, dt.$$

Now, let $A(s, z)$ and $B(s, z)$ denote the first, resp. second summand of $P_{\text{max}} (s, z)$. For small $\tau$, we approximate the term $P_{\text{max}} (s', z - \tau) - A(s, z)$, thereby making it independent of $t$. Accordingly, $A(s, z)$ is approximated by $A(s, z) \approx \max_{\alpha \in \text{Act}} \left( 1 - e^{-E(s) \tau} \right) \cdot \sum_{s' \in S} P_{\text{max}} (s', z - \tau).$

The summand $B(s, z)$ can be rewritten by shifting the range of integration by $-\tau$ such that $B(s, z) = e^{-E(s) \tau} \cdot P_{\text{max}} (s, z - \tau).$ Based on $B(s, z)$ and the approximation for $A(s, z)$, we propose a discretization for $P_{\text{max}}^{C, ML, G} (s, z)$ in a locally uniform CTMDP $C$ and a step duration $\tau$:

**Definition 11 (Discretization)** Let $C = (S, \text{Act}, R, \nu)$ be a locally uniform CTMDP, and let $\tau > 0$ be a step duration. The induced MDP $C_\tau = (S, \text{Act}, P_\tau, \nu)$ is defined such that for all $s, s' \in S$ and $\alpha \in \text{Act}(s)$:

$$P_\tau (s, \alpha, s') = \begin{cases} \left( 1 - e^{-E(s) \tau} \right) \cdot P(s, \alpha, s') & \text{if } s \neq s' \\ \left( 1 - e^{-E(s) \tau} \right) \cdot P(s, \alpha, s') + e^{-E(s) \tau} & \text{if } s = s'. \end{cases}$$

Further, for all $\alpha \notin \text{Act}(s)$, we define $P_\tau (s, \alpha, s') = 0$.
denotes the probability that no transition occurs within time \( \tau \) and thus \( s = s' \).

Let \( p_{c, \max}^z(s, k) \) be the maximum probability to reach \( G \) starting from state \( s \) in at most \( k \) steps in the (discrete-time) MDP \( C_r \). Therefore, \( p_{c, \max}^z(s, k) = 1 \) if \( s \in G \) and \( p_{c, \max}^z(s, 0) = 0 \) if \( s \notin G \). Further, for \( s \notin G \) and \( k > 0 \), \( p_{c, \max}^z(s, k) \) is defined recursively:

\[
p_{c, \max}^z(s, k) = \max_{\alpha \in Act(s)} \sum_{s' \in S} P_r(s, \alpha, s') \cdot p_{c, \max}^z(s', k-1).
\]

The next theorem proves that the probability to reach \( G \) from state \( s \) within at most \( k = \frac{\tau}{\bar{\tau}} \) steps in the discrete-time MDP \( C_r \) converges from below (for \( \tau \to 0 \)) to the corresponding time-bounded reachability probability \( p_{c, \max}^{C,\text{ML},G} \) in the CTMDP \( C \):

**Theorem 5** Let \( C \) be a locally uniform CTMDP; \( \lambda = \max_{s \in S} E(s) \), \( G \subseteq S \) a set of goal states, \( z \in \mathbb{R}_{\geq 0} \) a time bound and \( k \in \mathbb{N}, k > 0 \) the number of time steps, i.e. \( \tau = \frac{\tau}{\bar{\tau}} \). For all \( s \in S \), it holds:

\[
p_{c, \max}^z(s, k) \leq p_{c, \max}^{C,\text{ML},G}(s, z) \leq p_{c, \max}^z(s, k) + \frac{(\lambda z)^2}{2k}.
\]

**B. Algorithm and complexity**

Let \( C = (S, Act, R, \nu) \) be a locally uniform CTMDP, \( G \) a set of goal states and \( z \) a time bound. For some error bound \( \varepsilon > 0 \), let \( k \) be the number of steps needed to satisfy \( \varepsilon \geq \frac{(\lambda z)^2}{2k} \). Then \( \hat{\tau} = \frac{\varepsilon}{\lambda z} \) induces the discretized MDP \( C_\tau \) of \( C \) with step duration \( \tau \). By Thm. 5, the maximum probability to reach \( G \) within \( z \) time units in \( C \) can be approximated (up to \( \varepsilon \)) by maximizing the step-bounded reachability \( p_{c, \max}^z \) for \( G \) in \( C_\tau \) within \( k \) steps. The latter can be computed efficiently by the well-known value iteration approach [8]. Briefly, it starts with a probability vector \( v_0 \), with \( v_0(s) = 1 \) if \( s \in G \) and 0, otherwise. In each iteration, \( v_i \) is obtained from \( v_{i-1} \) according to Eq. (2). In each round, \( i \) is the number of steps in the MDP \( C_\tau \); hence, \( v_i(s) \) equals \( p_{c, \max}^z(s, i) \).

The value iteration approach on the discretized MDP \( C_\tau \) has the following complexity. For \( s \in S \) and \( \alpha \in Act(s) \), let \( post(s, \alpha) = \{ s' \in S \mid R(s, \alpha, s') > 0 \} \) be the set of \( \alpha \)-successors of state \( s \). The size of \( C \) is denoted by \( m = \sum_{s \in S} \sum_{\alpha \in Act(s)} |post(s, \alpha)| \). In the worst case, \( C_\tau \) is obtained by adding a self-loop for each state \( s \in S \) and action \( \alpha \in Act(s) \). Thus, the size of \( C_\tau \) is bounded by \( 2m \). For a given error bound \( \varepsilon \), it is easy to derive the number \( k \) of value iteration steps: By Thm. 5, \( |p_{c, \max}^z(s, z) - p_{c, \max}^{C,\text{ML},G}(s, k)| \leq \frac{(\lambda z)^2}{2k} \). Letting \( \frac{(\lambda z)^2}{2k} \leq \varepsilon \), we conclude that the smallest \( k \) to guarantee \( \varepsilon \) is \( \frac{(\lambda z)^2}{2k} \). In each value iteration step, the update of the vector \( v_i \) takes time \( 2m \). Thus, the worst-case time complexity of our approach is \( \Theta(m \cdot (\lambda z)^2/\varepsilon) \).

**C. Generation of \( \varepsilon \)-optimal schedulers**

Let \( C, G, \tau, C_r \) be as before. A byproduct of the value iteration on the discretized MDP \( C_\tau \) is an \( \varepsilon \)-optimal scheduler for the set of goal states \( G \) and time bound \( z \). More precisely, in any of the \( i \) value iteration steps, for each state \( s \in S \), an action \( \alpha_{s,i} \) is chosen according to Eq. (2). In this way, we obtain a history-dependent (or, to be more precise, a step-dependent) scheduler for the MDP \( C_r \). This scheduler induces a \( \tau \)-scheduler (denoted \( D_\tau \)) of the original CTMDP \( C \) as follows: \( D_\tau(s,t_{\text{past}}) = \alpha_{s,i} \) if \( t_{\text{past}} \in [(k-i)\tau, (k-i+1)\tau) \).

**Theorem 6** (\( \varepsilon \)-optimal scheduler) The scheduler \( D_\tau \) is an \( \varepsilon \)-optimal scheduler for \( C \) w.r.t. the maximum time-bounded reachability probability.

**V. TIME-BOUNDED REACHABILITY PROBABILITIES IN ARBITRARY CTMDPS**

In this section we consider general CTMDPs and the maximum time-bounded reachability probability with respect to GM-schedulers. First, we explain why our discretization technique from the previous section does not apply directly in this setting. As a consequence, we introduce a measure preserving transformation from CTMDPs to interactive Markov chains (IMCs). For the latter, a modified discretization that computes maximum and minimum time-bounded reachability probabilities has been studied in [9]. By the reduction that we describe here, these previous results enable us to compute maximum time-bounded reachability probabilities also in arbitrary CTMDPs under the class of GM-schedulers.

**A. A comparison of different scheduler classes**

Consider the CTMDP \( C \) which is depicted in Fig. 1. We compute the maximum time-bounded reachability probability for state \( s_4 \) with respect to initial state \( s_0 \), under different scheduler classes.

Note that the CTMDP in Fig. 1 is globally (and thus locally) uniform. This enables a comparison of all analysis methods with their underlying scheduler classes, that are currently available for CTMDPs. The resulting maximum probabilities are depicted in Fig. 2 for time bounds \( z \in [0, 8] \) and different scheduler classes.

- As \( C \) is locally uniform, the discretization from Sec. IV is applicable. As expected, the maximum probabilities for the underlying ML-schedulers outperform all other scheduler classes.

- For positional schedulers, the only relevant choice is between actions \( \alpha \) and \( \beta \) in state \( s_1 \); Fig. 2 depicts the results for both choices. Hence, the maximum reachability probability for the class of positional schedulers is the maximum of the two curves labeled \( \alpha \) and \( \beta \), resp.
• As $C$ is globally uniform, the algorithm in [1] is applicable, which computes the maximum time-bounded reachability probability for the class of time-abstract schedulers. Due to the restricted scheduler class, the obtained maxima are considerably smaller than those for $ML$- schedulers. In fact, in Fig. 2 they agree with the maximum that is achieved by positional schedulers. This is not surprising, as the only nondeterministic choice in $C$ occurs in state $s_1$, which is always entered along the (time-abstract) trajectory $\pi = s_0 \to s_1$.

• Finally, consider $C$ as a general CTMDP: The maximum time-bounded reachability probability under GM-schedulers is also depicted in Fig. 2. It is strictly higher than the one for time-abstract schedulers and strictly smaller than for late schedulers. A direct discretization technique as for late schedulers does not work here. For an example, see Sec. VI.

In the remainder of this section, we describe how to transform a general CTMDP into an equivalent IMC to compute maximum time-bounded reachability probabilities for GM-schedulers.

B. Interactive Markov chains

Opposed to CTMDPs, interactive Markov chains (IMCs) disentangle the relation between Markovian and nondeterministic behaviors. Therefore, IMCs strictly separate Markovian from interactive transitions. Below, we recall the notion of IMCs from [19]: For our purposes, it suffices to consider alternating IMCs in which all successor states of interactive states are Markovian states, and all successor states of Markovian states are interactive states:

**Definition 12 (Alternating IMCs)** An alternating interactive Markov chain is a tuple $M = (S, Act, IT, MT, \nu)$ where the set of states $S$ is partitioned into sets of interactive and Markovian states $S = MS \cup IS$. Act is the set of actions, $IT \subseteq IS \times Act \times IS$ is the set of interactive transitions and $MT \subseteq MS \times \mathbb{R}_{\geq 0} \times IS$ is the set of Markovian transitions. Further, $\nu \in \text{Distr}(S)$ is the initial distribution.

Paths in an IMC are sequences of the form $s_0 \overset{t_1, \alpha_1}{\rightarrow} s_1 \overset{t_2, \alpha_2}{\rightarrow} \cdots$ where $s_i \in S$, $\alpha_i \in Act \cup \{\perp\}$ and $t_i \in \mathbb{R}_{\geq 0}$. We use $\perp$ to denote Markovian transitions, i.e. if the IMC moves from state $s$ after a sojourn of $t$ time units along a Markovian transition to state $s'$, we write $s \overset{t}{\to} s'$. Otherwise, if an immediate transition with action $\alpha \in Act$ occurs, we write $s \overset{\alpha}{\to} s'$. We use $\text{Paths}^{\alpha}(M)$ and $\text{Paths}^\perp(M)$ to denote the sets of all finite paths and the set of all infinite paths. The corresponding $\sigma$-fields that constitute the measurable spaces are denoted $\mathcal{F}_{\text{Paths}^{\alpha}}$ and $\mathcal{F}_{\text{Paths}^\perp}$ and are obtained by the usual cylinder set construction (for details, we refer to [9] and [11]). For simplicity, we omit the reference to $M$ whenever possible. We define the set of enabled actions in an interactive state $s \in IS$ by $Act(s) = \{ \alpha \in Act \mid \exists s' \in S, (s, \alpha, s') \in IT\}$.

To resolve the corresponding nondeterministic choice, we use IMC schedulers:

**Definition 13 (IMC scheduler)** Let $(S, Act, IT, MT, \nu)$ be an IMC. An IMC scheduler is a mapping $D : \text{Paths}^\perp \to \text{Distr}(Act) \cup \{\perp\}$ such that for all $\pi \in \text{Paths}^\perp$, $D(\pi)(\alpha) \geq 0 \implies \alpha \in Act(\pi_{\downarrow})$ if $\pi_{\downarrow} \in IS$ and $D(\pi) = \perp$ if $\pi_{\downarrow} \in MS$. An IMC scheduler $D$ is measurable if $D^{-1}(\alpha) : \text{Paths}^\perp \to [0,1]$ is measurable w.r.t. $\mathcal{F}_{\text{Paths}^\perp}$ for all $\alpha \in Act$.

Note that no nondeterminism occurs if $\pi_{\downarrow}$ is a Markovian state, as the next state is determined probabilistically. Hence the scheduler yields the special action $\perp$ to indicate the absence of any decision. We use $GMI(M)$ to denote the class of measurable IMC schedulers for IMC $M$.

As for CTMDPs, the probability measure $\mu$ extends to entire sets of paths using standard measure theoretic constructions. Accordingly, we use $Pr_\mu$ to denote the probability measures induced by a GMI-scheduler $D$ on measurable sets of infinite paths.

C. Measure correspondence

In order to reduce the probabilistic reachability problem for general CTMDPs and GM-schedulers to the corresponding problem for IMCs, we first define the induced IMC $M(C)$ of a CTMDP $C$:

**Definition 14 (Induced IMC)** Let $C = (S, Act, R, \nu)$ be a CTMDP. Its induced IMC $M(C) = (S', Act, IT', MT', \nu')$ is defined by:

- $IS = S$ and $MS = \{ s^\alpha \mid s \in S \land \alpha \in Act(s) \}$
- $IT = \{(s, \alpha, s') \mid s \in S \land \alpha \in Act(s) \}$ and $MT = \{(s^\alpha, R(s, \alpha, s')) \mid s' \in S \land R(s, \alpha, s') > 0 \}$.

Further, $\nu'(s) = \nu(s)$ if $s \in IS$ and $\nu'(s) = 0$, otherwise.

Each state $s \in S$ in $C$ becomes an interactive state in $M(C)$. Intuitively, the induced IMC $M(C)$ mimics the CTMDP’s nondeterministic choices: For each action $\alpha \in Act(s)$, an interaction transition leads from state $s$ to a fresh Markovian state $s^\alpha$ which represents the scheduler’s choice of action $\alpha$. The Markovian transitions that leave state $s^\alpha$ represent the race between the exponential delays that ultimately lead to the corresponding successor state of $s$ in $C$.

To formally establish the relation between a CTMDP $C$ and its induced alternating IMC $M$, we first observe a correspondence between paths in $M$ and paths in $C$: Therefore, let $sep : \text{Paths}(C) \to \text{Paths}(M)$ be such that it separates the scheduler choices and the Markovian sojourn times on a path $\pi \in \text{Paths}(C)$. Formally, $sep(s_0 \overset{t_0, \alpha_0}{\rightarrow} s_1 \overset{t_1, \alpha_1}{\rightarrow} \cdots)$ equals $s_0 \overset{t_0, \alpha_0}{\rightarrow} s_1 \overset{t_1, \alpha_1}{\rightarrow} \cdots$. For infinite paths, we thus have a one-to-one correspondence between infinite paths in $C$ and infinite paths in $M$. For the following discussion, we extend the definitions of the functions $sep$ to sets of paths in the natural way.

The following theorem establishes a measure correspondence between arbitrary CTMDP and GM-schedulers and their induced IMC.

**Theorem 7 (Measure correspondence)** Given a CTMDP $C = (S, Act, R, \nu)$ and its induced IMC $M = (S', Act, IT, MT, \nu')$, it holds:
1) For $D^C \in GM(C)$ there exists $D^M \in GMI(M)$ such that for all measurable sets $\Pi \in S_{Paths^C(C)}$ in $C$ it holds $Pr^{\nu, D^C}_{\nu, D^C}(\Pi) = Pr^{\nu, D^M}_{\nu, D^M}(sep(\Pi))$.

2) For $D^M \in GMI(M)$ there exists $D^C \in GM(C)$ such that for all measurable sets $\Pi \in S_{Paths^C(C)}$ in $C$ it holds $Pr^{\nu, D^C}_{\nu, D^C}(\Pi) = Pr^{\nu, D^M}_{\nu, D^M}(sep(\Pi))$.

The above theorem was first established by Hermanns and John [20], [21]. A short proof that is tailored for alternating formation about the sojourn time in the current state and which there are only two early schedulers. Uniformization, however, the CTMDP depicted in Fig. 3. Upon starting in state $\alpha$ MDPs under late schedulers, namely $\{2\}$ is more general, as they consider a transformation from arbitrary — not necessarily alternating — IMCs to CTMDPs. For both CTMDPs and IMCs, time-bounded reachability objectives are measurable events. Applying the above theorem, one can exploit the algorithm [9], [11] for IMCs to solve the maximum time-bounded reachability problem for arbitrary CTMDPs under $GM$-schedulers. The overall complexity for computing the maximum reachability probability in alternating IMCs is the same as the one for locally uniformized CTMDPs under late schedulers, namely $O(m \cdot (\lambda z)^2 / \epsilon)$.

VI. UNIFORMIZATION

In the literature, uniformization has been applied to continuous-time Markov chains to compute transient and steady state probabilities. For CTMDPs, it is a well-known fact that uniformization increases the maximum reachability probability under time-abstract schedulers [1]. The same holds for $GM$-schedulers: To see this, consider the CTMDP depicted in Fig. 3. Upon starting in state $s_0$, an early scheduler directly chooses between $\alpha$ and $\beta$. Once the decision has been made, it cannot be changed anymore. Thus there are only two early schedulers. Uniformization, however, adds additional transitions which implicitly expose timing information about the sojourn time in the current state and which allow the early scheduler to change its decision (for a detailed discussion, see [3]). Hence, uniformization does not work for $GM$-schedulers.

However, in this section, we show that uniformization is sound for $ML$-schedulers and locally uniform CTMDPs. Intuitively, a late scheduler does not choose an action immediately upon entering a state, but rather at the time of leaving the state. Hence, no additional timing information can be revealed by uniformization.

Definition 15 Let $C = (S, Act, R, \nu)$ be a CTMDP and let $\lambda = \max_{s \in S} \max_{\alpha \in Act(s)} E(s, \alpha)$. We define the globally uniformized CTMDP $\mathcal{C}(\lambda')$ with uniformization rate $\lambda' \geq \lambda$. We set $\mathcal{C}(\lambda') = (S, Act, R, \nu)$, where $R, \nu(s, \alpha, s') = R(s, \alpha, s')$ if $s' \neq s$, and $\lambda' - E(s, \alpha)$ otherwise.

Theorem 8 Let $C = (S, Act, R, \nu)$ be a locally uniform CTMDP, $\lambda = \max_{s \in S} \max_{\alpha \in Act(s)} E(s, \alpha)$, $G \in S$ a set of goal states, $z \in \mathbb{R}_{\geq 0}$ a time bound. Further, assume that $\lambda, \lambda' \in \mathbb{Q}_{\geq 0}$.

For all $s \in S$ and all $\lambda' \geq \lambda$, it holds:

$$p_{ymax}^C(s, z) = p_{ymax}^{\lambda'}(s, z).$$

Proof: We exploit the discretization results established in Thm. 5. Let $\{k_i\}_{i \in N}$ be any increasing infinite sequence satisfying the condition $\tau_i = \frac{1}{k_i}$. Obviously $\lim_{i \to \infty} k_i = +\infty$ iff $\lim_{i \to \infty} \tau_i = 0$. Since $\lim_{i \to \infty} \frac{(\lambda z)^2}{k_i} = 0$, Thm. 5 implies that:

$$p_{ymax}(s, z) = \lim_{i \to \infty} p_{ymax}(s, k_i).$$

Now consider $C(\lambda')$: By similar arguments as above, we have:

$$p_{ymax}(s, z) = \lim_{i \to \infty} p_{ymax}(s, k_i) = \lim_{i \to \infty} C_{\tau_i}(s, k_i) = \lim_{i \to \infty} C_{\tau_i}(s, k_i) = \lim_{i \to \infty} C_{\tau_i}(s, k_i) = \lim_{i \to \infty} C_{\tau_i}(s, k_i).$$

which proves the theorem.

Similarly, it can be shown that for arbitrary IMCs, uniformization does not increase the power of IMC schedulers w.r.t. maximum time-bounded reachability probabilities.

It is the right place for an interesting open problem. Let us reconsider the plot in Fig. 2 and an infinite sequence of globally uniform CTMDPs with an increasing uniformization rate that converges to $+\infty$. We conjecture that the sequence of maximum reachability probabilities over time-abstract schedulers converges to the maximum reachable w.r.t. (time-dependent) late schedulers: Intuitively, with a large uniformization rate, a single step in the uniform CTMDP corresponds to a smaller time-interval, which then allows the time-abstract scheduler to approximate the optimal late scheduler more accurately.

VII. THE STOCHASTIC JOB SCHEDULING PROBLEM

We illustrate the applicability of our approach by considering the stochastic job scheduling problem (sJSP) from [22]. In their paper, the authors analyze the expected time to complete a set of stochastic jobs on a number of identical processors under a preemptive scheduling policy. An instance of the sJSP is a tuple $(m, n, \mu)$, where $m \geq 2$ is the number of processors, $J = \{1, \ldots, n\}$ is the set of stochastic jobs and $\mu: J \to \mathbb{R}_{\geq 0}$ specifies the jobs’ exponential service times, i.e. $\mu(i)$ is the rate of job $i$. The sJSP can be considered as a locally uniform CTMDP: A state of the sJSP is a tuple $(R, W)$, where $R, W \subseteq J$ (with $R \cap W = \emptyset$) are the sets of running and waiting jobs, respectively. When a job $j \in R$ completes, the decision which jobs to schedule next is nondeterministic. For $k$ remaining jobs, this gives rise to $\binom{k}{2}$ choices.

An action $\alpha \in Act((R, W))$ is a preemptive schedule: If state $(R, W)$ is left because job $j \in R$ finishes and if action $\alpha$ (where $\alpha: R \to 2^{R \cup W}$) is chosen, the set $\alpha(j)$ defines which jobs are executed next. In each state $(R, W)$, let $\text{Act}((R, W)) = \{\alpha: R \to 2^{R \cup W} \mid \forall j \in R. j \notin \alpha(j)\}$. This specifies the set of possible actions. For a fixed set of jobs, we can consider the set of all possible schedules.

For the stochastic job scheduling problem, the following property holds:

$$\text{sJSP}(R, W) = \bigcup_{\alpha \in \text{Act}((R, W))} \text{JSP}(\alpha(R, W))$$

where $\text{JSP}(\alpha(R, W))$ denotes the job scheduling problem with the schedule $\alpha(R, W)$.
Thus, given state $s_JSP$ represents a replacement strategy where jobs $\{1, \ldots, 4\}$ are executed next if job $1 \in R$ finishes first, and otherwise, the next jobs are $\{2, 4\}$. For action $\alpha_2$, the jobs $\{2, 4\}$ (or $\{1, 4\}$) are scheduled next if job 1 (resp. job 3) completes first.

Applying the results from Sec. IV, we are now able to algorithmically compute the maximum and minimum probabilities to finish all jobs within a given time bound $z$. In Fig. 4(b), we plot the maximum and minimum probabilities to finish jobs $\{1, \ldots, 4\}$ over a time bound $z \in [0, 15]$ for different values of $\mu$. The probabilities shown in Fig. 4(b) were obtained by implementing the discretization approach of Sec. IV for maximum and minimum time-bounded reachability. Clearly, for equally distributed job durations, i.e., if $\mu(i) = \mu(k)$ for all $i, k$, the maximum and minimum probabilities coincide. However, if $\mu(i) \neq \mu(k)$, the probabilities depend on the scheduling policy: In [22], the authors prove that the longest expected processing time first policy (LEPT) minimizes the expected makespan (i.e., the expected completion time) of the sJSP.

Although we consider a different quantitative measure (i.e., maximum time-bounded reachability instead of minimum expected makespan), we observe in our examples, that the $\varepsilon$-optimal $\tau$-scheduler which maximizes the reachability probabilities also adheres to the LEPT policy. Moreover, in our experiments, the $\tau$-scheduler that minimizes the time-bounded reachability probabilities obeys the shortest expected processing time first policy.

**VIII. Related Work**

Most of the existing results concerning CTMDPs focus on optimizing criteria such as the maximum expected total reward with a finite planning horizon [23], [18] or the maximum expected long-run average reward [24], [23], [25]. Uniformization and discretization have been applied to approximate the optimal rewards for both infinite and finite horizons [26], [18], [27], [28], [29].

So far, the analysis of CTMDPs with respect to maximum reachability probabilities has received scant attention. Directly related to our results is the work in [1], which provides an algorithm that computes time-bounded reachability probabilities in globally uniform CTMDPs. However, its applicability is severely restricted, as global uniformity — which requires the sojourn times in all states to be identically distributed — is hard to achieve. Moreover, the results in [1] only hold for time-abstract schedulers which are strictly less powerful than time-dependent ones [1], [3]. In [30], uniform IMCs have been transformed to (globally) uniform CTMDPs. By the transformation, maximum reachability under time-abstract schedulers for uniform IMCs can be computed using the algorithm in [1]. Recently, maximum reachability probabilities in CTMDPs have been studied in stochastic timed games [31], [4]: However, the authors of [4] also consider the strictly weaker classes of time-abstract schedulers only, while [31] addresses the decidability problem for qualitative reachability probabilities in stochastic timed games, i.e., compare probabilities to 1 or 0, respectively.

Therefore, this paper extends the existing results considerably: We provide an efficient algorithm that computes time-bounded reachability probabilities for the class of fully time- and history-dependent (early or late) schedulers up to an a priori given error bound $\varepsilon$. 

The approach in this paper is similar to the one in [9]. However, the results are complementary: In general, transforming IMCs to CTMDPs as done in [32], [30] does not yield locally (or globally) uniform CTMDPs. Hence, this transformation does not allow to analyze arbitrary IMCs (as it is done in [9]) by considering their induced CTMDP. Reversely, we show that time-bounded reachability probabilities in CTMDPs with respect to early schedulers can be computed by analyzing the CTMDP’s induced IMC.

Finally, let us mention two recent reports: In [33], Chen et al. consider maximum reachability probabilities for Markovian timed automata, an extension of timed automata [34] with exponential residence times. Similar to our approach, they also apply a discretization to approximate the maximum time-bounded reachability probability. In [35], Rabe and Schewe establish the existence of finite optimal timed schedulers for models which can be considered as extension of CTMDPs with instant states and two players games. Their extension with instant states has similar components as Markov automata [36], and thus generalizes both early and late schedulers.

ACKNOWLEDGMENTS

The authors thank Holger Hermanns and Joost-Pieter Katoen for their comments and for the many fruitful discussions. Moreover, we thank Christel Baier for pointing out errors in an early version of this paper.

This work was supported by the DFG as part of the Transregional Collaborative Research Centre SFB/TR 14 AVACS, by the European Community’s Seventh Framework Programme under grant agreement no. 214755 QUASIMODO, and by the NWO-DFG bilateral project ROCKS. Part of this work was done while Martin R. Neuhäußer was at RWTH Aachen, Germany and at the University of Twente, The Netherlands, and while Lijun Zhang was at Saarland University and at the Computing Laboratory, Oxford University. Lijun Zhang received partial support from MT-LAB, a VKR Centre of Excellence.

REFERENCES