Temporal logic control of general Markov decision processes by approximate policy refinement

Temporal logic control of general Markov decision processes by approximate policy refinement

Sofie Haesaert[1], Sadegh Soudjani[2], and Alessandro Abate[3]

[1] California Institute of Technology, United States
[2] School of Computing, Newcastle University, United Kingdom
[3] Computer Science Department, Oxford University, United Kingdom

Abstract. The formal verification and controller synthesis for Markov decision processes that evolve over uncountable state spaces are computationally hard and thus generally rely on the use of approximations. In this work, we consider the correct-by-design control of general Markov decision processes (gMDPs) with respect to temporal logic properties by leveraging approximate probabilistic relations between the original model and its abstraction. We newly work with a robust satisfaction for the construction and verification of control strategies, which allows for both deviations in the outputs of the gMDPs and in the probabilistic transitions. The computation is done over the reduced or abstracted models, such that when a property is robustly satisfied on the abstract model, it is also satisfied on the original model with respect to a refined control strategy.

1 Introduction

With the ever more ubiquitous embedding of digital components into physical systems new computational efficient verification and control synthesis methods for these cyber-physical systems are needed. The correct functioning of cyber-physical systems can only be expressed over the combined behaviour of both the digital component and its connected physical system. Quite importantly, stochastic models are key when computers interact with physical systems such as biological processes, power networks, and smart-grids. These dynamic systems with uncertainty and non-determinism can be modelled as Markov processes evolving over continuous spaces. Their potential safety critical impact on the environment they interact with makes it of particular interest to develop formal methods that assist in their verifiable design. In this work, we newly enable the verification and synthesis of these stochastic systems with respect to probabilistic linear temporal logic properties.

As our modelling framework, we consider the rich class of general Markov decisions processes (gMDPs), which are Markov decision processes evolving over continuous or uncountable state spaces and which have control-dependent stochastic transitions in combination with a metric output space. It is over this output that we define properties of interest. The characterisation of properties over such processes can in general not be attained analytically [3], so an alternative is to approximate these models by simpler processes, such as finite-state MDP [9] or continuous-space reduced order models [21] that are prone to be mathematically analysed or algorithmically verified [10]. In [13,12] we have proposed new approximate similarity relations to quantify the accuracy of the approximation utilising bounds on the output distance and on the transition probabilities. We have shown that for bounded safety properties these approximate similarity relations can be used to refine control strategies with bounded error on transition probability and deviations in the output space. The main goal of this paper is to study general temporal logic properties of gMDPs in combination with these approximate similarity relations. As the main contribution of the paper we show that the standard verification and control synthesis for gMDPs can be made robust a-priori to the introduced accuracy errors.

Related Work. Properties defined in PCTL, PLTL, and PCTL* for finite-state Markov (decision) processes can be verified using tools such as PRISM [17]. Moreover it is also well-known how to design policies, i.e., to control these Markov decision processes such that the satisfaction of these properties is maximised. The work in [2] has studied model-checking of automata specifications against autonomous (i.e. uncontrolled) discrete-time stochastic models over uncountable state spaces. It was shown that the computation of the probability of satisfying a specification expressed as deterministic finite automaton (DFA) can be restated in terms of a probabilistic reachability problem over the product between the original model and the DFA. This result has been extended in [22] to the case of controlled discrete-time Markov processes, which are a special subclass of gMDPs introduced in our work (obtained with identity output map, $h(x) = x$ for all $x \in \mathbb{X}$, and so with an output space $\mathbb{Y} = \mathbb{X}$). In this work we extend the model class discussed in [22] and build on their results by showing that an approximate model can be used to compute the solutions of Bellman equations associated to the stochastic reachability problem.

The paper is organised as follows. In the next section, we first define gMDPs and state the temporal logic control problem. In Section 3, we define approximate simulation relation on gMDPs and solve the robust probabilistic reachability problem. In Section 4, we extend the results to probabilistic temporal logic control utilising the product of gMDP and DFA. In Section 5, we detail the approximation procedure for linear stochastic dynamical systems and we apply it to a simple toy case. Throughout the paper proofs of the theorems have been relegated to the appendix of the extended version [14].

2 Problem statement: temporal logic control

2.1 Preliminaries and notations

In this work, we only consider Borel measurable spaces, i.e., $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, and we restrict our attention to Polish spaces [5]. Together with the measurable space

$(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, a probability measure $\mathbb{P}$ defines the probability space, denoted by $(\mathbb{X}, \mathcal{B}(\mathbb{X}), \mathbb{P})$ and has realisations $x \sim \mathbb{P}$. Let us further denote the set of all probability measures for a given measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ as $\mathcal{P}(\mathbb{X}, \mathcal{B}(\mathbb{X}))$.

For the sets A and B a relation $\mathcal{R} \subset A \times B$ is a subset of the Cartesian product $A \times B$. The relation $\mathcal{R}$ relates $x \in A$ with $y \in B$ if $(x, y) \in \mathcal{R}$, which is equivalently written as $x\mathcal{R}y$. For a given set $\mathbb{Y}$ a metric or distance function $\mathbf{d}_\mathbb{Y}$ is a function $\mathbf{d}_\mathbb{Y} : \mathbb{Y} \times \mathbb{Y} \to \mathbb{R}_{\geq 0}$ satisfying the following conditions: $\forall y_1, y_2, y_3 \in \mathbb{Y}$: $\mathbf{d}_\mathbb{Y}(y_1, y_2) = 0$ iff $y_1 = y_2$; $\mathbf{d}_\mathbb{Y}(y_1, y_2) = \mathbf{d}_\mathbb{Y}(y_2, y_1)$; and $\mathbf{d}_\mathbb{Y}(y_1, y_3) \leq \mathbf{d}_\mathbb{Y}(y_1, y_2) + \mathbf{d}_\mathbb{Y}(y_2, y_3)$.

2.2 General Markov decision processes and control strategies

General Markov decision processes are related to control Markov processes [1] and Markov decision processes [4,20,15], and formalised as follows.

Definition 1 (general Markov decision process (gMDP)). *A discrete-time gMDP is a tuple* $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ *with*

- $\mathbb{X}$, *an (uncountable) Polish state space with states* $x \in \mathbb{X}$ *as its elements;*
- $\mathbb{U}$, *the set of controls which is a Polish space;*
- π, *the initial probability measure* $\pi : \mathcal{B}(\mathbb{X}) \to [0, 1]$*;*
- $\mathbb{T} : \mathbb{X} \times \mathbb{U} \times \mathcal{B}(\mathbb{X}) \to [0, 1]$, *a conditional stochastic kernel assigning to each state* $x \in \mathbb{X}$ *and control* $u \in \mathbb{U}$ *probability measure* $\mathbb{T}(\cdot \mid x, u)$ *over* $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$*;*
- $\mathbb{Y}$, *the output space decorated with metric* $\mathbf{d}_\mathbb{Y}$*; and*
- $h : \mathbb{X} \to \mathbb{Y}$, *a measurable output map.*

For any set $A \in \mathcal{B}(\mathbb{X})$, $\mathbb{P}_{x,u}(x(t + 1) \in A) = \int_A \mathbb{T}(dx' \mid x(t) = x, u)$, where $\mathbb{P}_{x,u}$ denotes the conditional probability $\mathbb{P}(\cdot \mid x, u)$. At every state, the state transition depends non-deterministically on the choice of $u \in \mathbb{U}$. When chosen according to a probability measure $\mu_u : \mathcal{B}(\mathbb{U}) \to [0, 1]$, we refer to the stochastic control input as μ_u and denote the transition kernel as $\mathbb{T}(\cdot|x, \mu_u) = \int_\mathbb{U} \mathbb{T}(\cdot|x, u)\mu_u(du) \in \mathcal{P}(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. Given a string of inputs $\{u(t)\}_{t\leq N} := u(0), u(1), \ldots, u(N)$, over a finite time horizon $[0, N]$, and an initial condition x_0 (sampled from π), the state at the $(t + 1)$-st time instant, $x(t + 1)$, is obtained as a realisation of the controlled Borel-measurable stochastic kernel $\mathbb{T}(\cdot \mid x(t), u(t))$ – these semantics induce paths (or executions) of the gMDP. Further, output traces of gMDP are obtained by applying the output map $h(\cdot)$ to the paths of gMDP, namely $\{y(t)\}_{t\leq N} := y(0), y(1), \ldots, y(N)$ with $y(t) = h\big(x(t)\big)$ for all $t \in [0, N]$. Denote the class of all gMDP with the same metric output space $\mathbb{Y}$ as $\mathcal{M}_\mathbb{Y}$.

A policy is a selection of control inputs based on the past history of states and controls. When the selected controls are only dependent on the current states, the policy is referred to as Markov.

Definition 2 (Markov policy). *For a gMDP* $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$*, a Markov policy* μ *is a sequence* $\mu = (\mu_0, \mu_1, \mu_2, \ldots)$ *of universally measurable maps* $\mu_t = \mathbb{X} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$*,* $t \in \mathbb{N} := \{0, 1, 2, \ldots\}$*, from the state space* $\mathbb{X}$ *to the set of controls.*

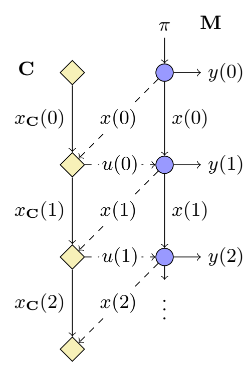
We allow controls to be selected via universally measurable maps [4] from the state to the space of stochastic control inputs, so that properties such as safety can be maximised [3]. We introduce the notion of a control strategy, and define it as a broader, memory-dependent version of the Markov policy above. This strategy is formulated again as a gMDP that takes as its input the state of the to-be-controlled gMDP.

Definition 3 (Control strategy). *A control strategy* $\mathbf{C} = (\mathbb{X}_{\mathbf{C}}, x_{\mathbf{C}0}, \mathbb{X}, \mathbb{T}_{\mathbf{C}}, h_{\mathbf{C}})$ *for a gMDP* $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ *is a gMDP with state space* $\mathbb{X}_{\mathbf{C}}$*; initial state* $x_{\mathbf{C}0}$*; input space* $\mathbb{X}$*; universally measurable kernels* $\mathbb{T}_{\mathbf{C}} : \mathbb{X}_{\mathbf{C}} \times \mathbb{X} \times \mathcal{B}(\mathbb{X}_{\mathbf{C}}) \to [0,1]$*; and with universally measurable output map* $h_{\mathbf{C}} : \mathbb{X}_{\mathbf{C}} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$.

Note that the stochastic transitions for the control strategy and the gMDP are selected in an alternating fashion. The output map of the strategy is indexed based on the time instant at which the resulting policy will be applied to the gMDP. This is further elucidated in Algorithm 1. The execution $\{(x(t), x_{\mathbf{C}}(t)), t \in [0, N]\}$

Algorithm 1 Execution semantics of the controlled model $\mathbf{C} \times \mathbf{M}$.

procedure EXECUTION($\mathbf{C}$, $\mathbf{M}$)
  set $t := 0$ and $x_{\mathbf{C}}(0) := x_{\mathbf{C}0}$
  draw $x(0) \sim \pi$
  for $t \leq N$ **do**
    draw $x_{\mathbf{C}}(t+1) \sim \mathbb{T}_{\mathbf{C}}(\,\cdot\,|x_{\mathbf{C}}(t), x(t))$
    draw $u(t)$ from $\mu_t := h_{\mathbf{C}}(x_{\mathbf{C}}(t+1))$
    draw $x(t+1) \sim \mathbb{T}(\,\cdot\,|x(t), u(t))$
    set $t := t + 1$
  end for
  return $\{x(t)\}_{t \leq N}$
end procedure



of a gMDP $\mathbf{M}$ controlled with strategy $\mathbf{C}$ (denoted by $\mathbf{C} \times \mathbf{M}$) is defined on the canonical sample space $\Omega := (\mathbb{X} \times \mathbb{X}_{\mathbf{C}})^{N+1}$ endowed with its product topology $\mathcal{B}(\Omega)$ and with a unique probability measure $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}$. A Markov policy is as a special case of control strategy which does not have an internal state that can be used to remember relevant past events.

Remark 1. Any Markov policy μ as defined in Def. 2 can be written as a control strategy $\mathbf{C}_\mu := (\mathbb{X}_\mu, x_{\mu 0}, \mathbb{X}, \mathbb{T}_\mu, h_\mu)$ with $\mathbb{X}_\mu := Q \times \mathbb{X}$, for which Q is the set of time indices $Q := \{-1, 0, 1, 2, 3, \ldots\}$, and with $x_{\mu 0} := (-1, x_0)$ the initial state for some $x_0 \in \mathbb{X}$. The probability measure on the next control state $x'_\mu = (q', \tilde{x}') \in \mathbb{X}_\mu$ given $x(t)$ and given the current state $x_\mu = (q, \tilde{x})$ is defined with a stochastic kernel as

$$\mathbb{T}_\mu(A|(q,\tilde{x}), x(t)) := \begin{cases} 1 \text{ if } (q+1, x(t)) \in A, \\ 0 \text{ else .} \end{cases}$$

The policies can then be embedded into the control strategy via the output map $h_\mu((q,\tilde{x})) := \mu_q(\tilde{x})$.[iv]

2.3 Probabilistic path properties of controlled gMDPs

Consider a measurable target set $K \subset \mathbb{Y}$. We say that an output trace $\{y(t)\}_{t\leq N}$ reaches a target set K, if there exists a time $t \in [0,N]$ such that $y(t) \in K$. This bounded reaching of K is denoted by $\Diamond^{\leq N}\{y \in K\}$ or briefly $\Diamond^{\leq N}K$. For $N \to \infty$, we denote the reachability property as $\Diamond K$, i.e., eventually K. For a given gMDP $\mathbf{M}$ with control strategy $\mathbf{C}$, a verification task consists of quantifying the probability that an output trace of $\mathbf{C} \times \mathbf{M}$ reaches K within the time horizon $[0,N]$, i.e., $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond^{\leq N}K)$, or that the target set K is eventually reached, i.e., $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond K)$, and verifying that it is within a given threshold.

More complex properties can be described using temporal logic. Consider a set of atomic propositions AP, the alphabet $\Sigma := 2^{AP}$, and infinite words that are string composed of elements from Σ, $\omega = \omega(0), \omega(1), \omega(2), \ldots \in \Sigma^{\mathbb{N}}$. Of interest are atomic propositions that are connected to the gMDP via a measurable labelling function $\mathsf{L} : \mathbb{Y} \to \Sigma$ from the output space to the alphabet Σ. Via its trivial extension, output traces $\{y(t)\}_{t\geq 0} \in \mathbb{Y}^{\mathbb{N}}$ can be mapped to the set of infinite words $\Sigma^{\mathbb{N}}$, as $\omega = \mathsf{L}(\{y(t)\}_{t\geq 0}) := \{\mathsf{L}(y(t))\}_{t\geq 0}$. Consider linear-time temporal logic properties with syntax

$$\psi ::= \text{true} \mid p \mid \neg\psi \mid \psi_1 \wedge \psi_2 \mid \bigcirc\psi \mid \psi_1 \mathbin{\mathsf{U}} \psi_2. \tag{1}$$

Let $\omega_t = \omega(t), \omega(t{+}1), \omega(t{+}2), \ldots$ be a postfix of the word ω, then the satisfaction relation between ω and a property ψ, expressed via LTL, is denoted by $\omega \vDash \psi$ (or equivalently $\omega_0 \vDash \psi$). The semantics of the satisfaction relation are defined recursively over ω_t and the syntax of the LTL formula ψ. An atomic proposition $p \in AP$ is satisfied by ω_t, i.e., $\omega_t \vDash p$, iff $p \in \omega(t)$. Furthermore, $\omega_t \vDash \neg\psi$ if $\omega_t \nvDash \psi$ and we say that $\omega_t \vDash \psi_1 \wedge \psi_2$ if $\omega_t \vDash \psi_1$ and if $\omega_t \vDash \psi_2$. The next operator $\omega_t \vDash \bigcirc\psi$ holds if the property holds at the next time instance $\omega_{t+1} \vDash \psi$. The temporal until operator $\omega_t \vDash \psi_1 \mathbin{\mathsf{U}} \psi_2$ holds if $\exists i \in \mathbb{N} : \omega_{t+i} \vDash \psi_2$, and $\forall j \in \mathbb{N} : 0 \leq j < i, \omega_{t+j} \vDash \psi_1$. Based on these semantics, operators such as disjunction ($\vee$) can also be defined through the negation and conjunction: $\omega_t \vDash \psi_1 \vee \psi_2 \;\Leftrightarrow\; \omega_t \vDash \psi_1$ or $\omega_t \vDash \psi_2$.

We are interested in a fragment of LTL property known as syntactically co-safe temporal logic (scLTL) [16]. Even though scLTL formulas are interpreted over infinite words, their satisfaction is guaranteed in finite time. This fragment is defined as follows.

Definition 4. *A scLTL over a set of atomic propositions AP has syntax*

$$\psi ::= \text{true} \mid p \mid \neg p \mid \psi_1 \wedge \psi_2 \mid \psi_1 \vee \psi_2 \mid \bigcirc\psi \mid \psi_1 \mathbin{\mathsf{U}} \psi_2 \mid \Diamond\psi_2 \tag{2}$$

with $p \in AP$.

[iv] Note that for this construction, the separability of the Polish space $\mathbb{X}$ is important as otherwise $\mathbb{T}_\mu$ would not be measurable in general.

In the remainder, we will mainly consider scLTL properties since their verification can be computed via a reachability property over a finite state automaton [16]. With respect to a scLTL property ψ, we say that a gMDP $\mathbf{M}$ satisfies ψ for a given control strategy $\mathbf{C}$ with probability at least p iff $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\mathsf{L}(\{y(t)\}_{t\geq 0}) \models \psi) \geq p$. Apart from this verification task, we are mostly interested in synthesising control strategies such that $\mathbf{C} \times \mathbf{M}$ satisfies the inequality.

2.4 Problem statement

In this paper, we tackle the control synthesis for the (bounded) probabilistic reachability problem and for the temporal logic control problem, defined next.

Problem 1 ((Bounded) probabilistic reachability) *Given a gMDP* $\mathbf{M}$ *and a set* $K \subset \mathbb{Y}$*, compute a control strategy* $\mathbf{C}$ *that maximises the bounded reachability probability* $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond^{\leq N} K)$ *or the reachability probability* $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond K)$*.*

Problem 2 (Temporal logic control) *Given a gMDP* $\mathbf{M}$*, a scLTL property* ψ *and a labelling function* L*, compute a control strategy* $\mathbf{C}$ *that maximises the probability of controlled Markov process* $\mathbf{C} \times \mathbf{M}$ *satisfying* ψ*, i.e.,*

$$\max_{\mathbf{C}} \ \mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\mathsf{L}(\{y(t)\}_{t\geq 0}) \models \psi). \tag{3}$$

Computation of (3) for a given controlled Markov process is generally impossible. In the next sections, we give a robust computations based on (ϵ, δ)-probabilistic simulation relations between a given model and its approximation.

3 Quantifying stochastic reachability using (ϵ, δ)-probabilistic simulation relations

Let a gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ and a target set $K \subset \mathbb{Y}$ be given. In the robust formulation of Problem 1, we compute a control strategy $\mathbf{C}$ and a quantified lower bound p, such that the probability of satisfying $\Diamond^{\leq N} K$, respectively $\Diamond K$, is lower bounded by p, i.e., $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond^{\leq N} K) \geq p$, respectively $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\Diamond K) \geq p$. We associate to the target set $K \subset \mathbb{Y}$, the corresponding set in the state space $K_{\mathbb{X}} := h^{-1}(K) \in \mathcal{B}(\mathbb{X})$. Let us denote by $r^{\mu}(K_{\mathbb{X}}, N)$, the probability that K is reached within N time steps when a Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$ is used. We iterate the computation of stochastic reachability of a Markov decision process, as explain in [3]. The value of $r^{\mu}(K_{\mathbb{X}}, N)$ can be computed by a backward recursion initialised with $V_N = 0$, and iterated for $k = N-1, \ldots, 0$ as

$$V_k(x) = \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(\bar{x}) + \mathbf{1}_{\mathbb{X}\setminus K_{\mathbb{X}}}(\bar{x}) V_{k+1}(\bar{x})\right] \mathbb{T}(d\bar{x}|x, \mu_k(x)). \tag{4}$$

Based on the final value function after N iterations, we have that

$$r^{\mu}(K_{\mathbb{X}}, N) = \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(x) + \mathbf{1}_{\mathbb{X}\setminus K_{\mathbb{X}}}(x) V_0(x)\right] \pi(dx). \tag{5}$$

Furthermore, the optimal value functions $V_k^*(x)$, $k \in [0, N]$, computed recursively with $V_N^*(x) = 0$ and for all $x \in \mathbb{X}$ and $k = N-1, \ldots, 1, 0$,

$$V_k^*(x) = \sup_{\mu_k} \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(\bar{x}) + \mathbf{1}_{\mathbb{X} \setminus K_{\mathbb{X}}}(\bar{x}) V_{k+1}^*(\bar{x}) \right] \mathbb{T}(d\bar{x}|x, \mu_k(x)), \tag{6}$$

give the optimal reachability probability

$$r^*(K_{\mathbb{X}}, N) = \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(x) + \mathbf{1}_{\mathbb{X} \setminus K_{\mathbb{X}}}(x) V_0^*(x) \right] \pi(dx). \tag{7}$$

Using $V_k^*(x)$ the control strategy $\mathbf{C}^*$ maximising $\mathbb{P}_{\mathbf{C} \times \mathbf{M}}(\Diamond K)$ is defined by a Markov policy μ^* with elements μ_k^*

$$\mu_k^*(x) \in \arg\sup_{\mu_k} \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(\bar{x}) + \mathbf{1}_{\mathbb{X} \setminus K_{\mathbb{X}}}(\bar{x}) V_{k+1}^*(\bar{x}) \right] \mathbb{T}(d\bar{x}|x, \mu_k(x)). \tag{8}$$

The unbounded optimal reachability probability $r^*(K_{\mathbb{X}})$ can be evaluated based on the fixed point of (6). More specifically, for $N \to \infty$ the value functions are strictly increasing and converging to the fixed point solution of

$$V^*(x) = \sup_{\mu} \int_{\mathbb{X}} \left[\mathbf{1}_{K_{\mathbb{X}}}(\bar{x}) + \mathbf{1}_{\mathbb{X} \setminus K_{\mathbb{X}}}(\bar{x}) V^*(\bar{x}) \right] \mathbb{T}(d\bar{x}|x, \mu(x)). \tag{9}$$

For a given policy μ, the unbounded reachability probability $r^{\mu}(K_{\mathbb{X}})$ and the associated value function $V^{\mu}(x)$ are formulated similarly. Thus we have that

$$\sup_{\mathbf{C}} \mathbb{P}_{\mathbf{C} \times \mathbf{M}}(\Diamond K) = r^*(K_{\mathbb{X}}), \text{ respectively, } \sup_{\mathbf{C}} \mathbb{P}_{\mathbf{C} \times \mathbf{M}}(\Diamond^{\leq N} K) = r^*(K_{\mathbb{X}}, N).$$

This together, with its universal measurability, has been discussed in [3]. Computation of backward recursions (4) and (6) is generally only tractable for finite state spaces. In the following, we define approximate probabilistic simulation relations over $\mathcal{M}_{\mathbb{Y}}$, as introduced in [13], and apply it to compute a lower bound on probabilistic reachability and the corresponding synthesis problem.

3.1 Approximate simulation relations for gMDPs

Consider two gMDPs $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, $i = 1, 2$, that have the same metric output space $(\mathbb{Y}, \mathbf{d}_{\mathbb{Y}})$. Given state-action pairs $x_1 \in \mathbb{X}_1, u_1 \in \mathbb{U}_1$ and $x_2 \in \mathbb{X}_2, u_2 \in \mathbb{U}_2$, we want to relate the corresponding transition kernels, namely the probability measures $\mathbb{T}_1(\cdot \mid x_1, u_1) \in \mathcal{P}(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1))$ and $\mathbb{T}_2(\cdot \mid x_2, u_2) \in \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$. As in [13], we introduce the concept of δ-lifting as follows.

Definition 5 (δ-lifting for general state spaces). *Let $\mathbb{X}_1, \mathbb{X}_2$ be two sets with associated measurable spaces $(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1)), (\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$, and let $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$ be a relation for which $\mathcal{R} \in \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2)$. We denote by $\bar{\mathcal{R}}_\delta \subseteq \mathcal{P}(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1)) \times \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$ the corresponding lifted relation so that $\Delta \bar{\mathcal{R}}_\delta \Theta$ holds if there exists a probability space $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2), \mathbb{W})$ (equivalently, a lifting $\mathbb{W}$) satisfying*

L1. *for all* $X_1 \in \mathcal{B}(\mathbb{X}_1)$*:* $\mathbb{W}(X_1 \times \mathbb{X}_2) = \Delta(X_1)$*;*
L2. *for all* $X_2 \in \mathcal{B}(\mathbb{X}_2)$*:* $\mathbb{W}(\mathbb{X}_1 \times X_2) = \Theta(X_2)$*;*
L3. *for the probability space* $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2), \mathbb{W})$ *it holds that* $x_1 \mathcal{R} x_2$ *with probability at least* $1 - \delta$*, or equivalently that* $\mathbb{W}(\mathcal{R}) \geq 1 - \delta$*.*

We introduce a notion of approximate probabilistic simulation relations which naturally leads to control refinement. For this, we build on the notion of *interface function* [11] to define probabilistic simulation relations that allow for the hierarchical control refinement of two gMDPs

$$\mathcal{U}_v : \mathbb{U}_1 \times \mathbb{X}_1 \times \mathbb{X}_2 \to \mathcal{P}(\mathbb{U}_2, \mathcal{B}(\mathbb{U}_2)).$$

This interface function $\mathcal{U}_v$ is required to be a Borel measurable function. Intuitively, an interface function implements (or refines) any control action synthesised over the abstract model to an action for the concrete model.

Definition 6 ((ϵ, δ)**-probabilistic simulation relation).** *Consider two gMDPs* $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$*,* $i = 1, 2$*, over a shared metric output space* $(\mathbb{Y}, \mathbf{d}_{\mathbb{Y}})$*.* $\mathbf{M}_1$ *is* (ϵ, δ)*-stochastically simulated by* $\mathbf{M}_2$ *if there exists an interface function* $\mathcal{U}_v$ *and a relation* $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$*, for which there exists a Borel measurable stochastic kernel* $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1, x_1, x_2)$ *on* $\mathbb{X}_1 \times \mathbb{X}_2$ *given* $\mathbb{U}_1 \times \mathbb{X}_1 \times \mathbb{X}_2$*, such that:*

APS1. $\forall (x_1, x_2) \in \mathcal{R}$*,* $\mathbf{d}_{\mathbb{Y}}(h_1(x_1), h_2(x_2)) \leq \epsilon$*;*
APS2. $\forall (x_1, x_2) \in \mathcal{R}$*,* $\forall u_1 \in \mathbb{U}_1$*:* $\mathbb{T}_1(\cdot|x_1, u_1)\ \bar{\mathcal{R}}_\delta\ \mathbb{T}_2(\cdot|x_2, \mathcal{U}_v(u_1, x_1, x_2))$*, with lifted probability measure* $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1, x_1, x_2)$*;*
APS3. $\pi_1 \bar{\mathcal{R}}_\delta \pi_2$*.*

The simulation relation is denoted as $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$*.*

This definition extends the known exact notions of probabilistic simulation in [18], and the approximate notions of [7,8] to gMDPs over Polish spaces as elaborated in [13]. The Borel measurability for both $\mathcal{U}_v$ (see above) and $\mathbb{W}_{\mathbb{T}}$ (as in this definition), which is used to prove the well-posedness of the controller refinement, can be relaxed to universal measurability [13].

3.2 δ-Robust probabilistic reachability

Definition 7 ((ϵ, δ)**-Robust satisfaction).** *Consider any* $\mathbf{M}_1 \in \mathcal{M}_{\mathbb{Y}}$*. We say that a Markov policy* μ *for* $\mathbf{M}_1$ (ϵ, δ)*-robustly satisfies* $\Diamond^{\leq N} K$ *with probability* p *if for every* $\mathbf{M}_2 \in \mathcal{M}_{\mathbb{Y}}$ *with* $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$ *there exists a controller* $\mathbf{C}_2$ *for* $\mathbf{M}_2$ *such that* $\mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}(\Diamond^{\leq N} K) \geq p$*.*

For a given universally measurable map $\nu : \mathbb{X}_1 \to \mathcal{P}(\mathbb{U}_1, \mathcal{B}(\mathbb{U}_1))$ and constant δ, define the operator $\mathbf{T}_\delta^\nu : \mathcal{F} \to \mathcal{F}$ acting on the set of functions $\mathcal{F} := \{f : \mathbb{X}_1 \to [0, 1]\}$ as

$$\mathbf{T}_\delta^\nu(V)(x) := \mathbf{L}\left(\int_{\mathbb{X}_1} [\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}) + \mathbf{1}_{\mathbb{X}_1 \setminus K_{\mathbb{X}_1}}(\bar{x}) V(\bar{x})] \mathbb{T}(d\bar{x}|x, \nu(x)) \ - \delta\right) \tag{10}$$

with $\mathbf{L} : \mathbb{R} \to [0, 1]$ being the truncation function $\mathbf{L}(\cdot) := \min(1, \max(0, \cdot))$. Define also the operator $\mathbf{T}_\delta^*(V)$ on $\mathcal{F}$ as $\mathbf{T}_\delta^*(V)(x) := \sup_\nu \mathbf{T}_\delta^\nu(V)(x)$.

Theorem 1. *A target set* $K \subset \mathbb{Y}$ *of gMDP* $\mathbf{M}_1$ *is reached* $(0,\delta)$*-robustly with Markov policy* μ *and with probability* $r^\mu(K_{\mathbb{X}_1}, N)$*, where*

$$r^\mu(K_{\mathbb{X}_1}, N) := \mathbf{L}\left(\int_{\mathbb{X}_1} [\mathbf{1}_{K_{\mathbb{X}_1}}(x) + \mathbf{1}_{\mathbb{X}_1 \setminus K_{\mathbb{X}_1}}(x) V_0^\delta(x)]\pi_1(dx) - \delta\right), \tag{11}$$

and $V_0^\delta(x)$ *is computed recursively according to* $V_k^\delta := \mathbf{T}_\delta^{\mu_k}(V_{k+1}^\delta)$*, for* $k = N-1, \ldots, 0$ *with initial value function* $V_N^\delta = 0$*. If* $V_0^{\delta,*}$ *is computed similarly as the solution of recursion* $V_k^{\delta,*} = \mathbf{T}_\delta^*(V_{k+1}^{\delta,*})$ *with initial value function* $V_N^{\delta,*} = 0$ *and* $\mu_k^* \in \arg\sup_{\mu_k} \mathbf{T}_\delta^{\mu_k}(V_{k+1}^{\delta,*})$ *then we call* $\mu^* = \{\mu_0, \mu_1, \ldots\}$ *the optimal* $(0,\delta)$*-robust policy.*

Notice that for $\delta = 0$ the computation of value functions V_k^δ in Theorem 1 is the same as (4). The proof of Theorem 1 builds on the construction of a refined control strategy as has been explained in [13]. For any $\mathbf{M}_2$ such that $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$ with the lifted probability measure $\mathbb{W}_\mathbb{T}$, the control policy $\mathbf{C}_2$ can be refined from $\mathbf{C}_1$ (cf. [13]). More precisely, a control strategy $\mathbf{C}_2$ that refines $\mathbf{C}_1$ over $\mathbf{M}_2$ is obtained by extending $\mathbf{C}_1$ with internal states (x_1, x_2). While the state (x_1, x_2) of $\mathbf{C}_2$ is in $\mathcal{R}$, the control refinement has as its *basic ingredients* the states x_1 and x_2, whose the stochastic transition to the pair (x_1', x_2') is governed firstly by a point distribution $\delta_{x_2(t)}(dx_2')$ based on the measured state $x_2(t)$ of $\mathbf{M}_2$; and subsequently, by the lifted probability measure $\mathbb{W}_\mathbb{T}(dx_1'|x_2', u_1, x_2, x_1)$, *conditioned* on x_2'.

Before tackling unbounded reachability properties, we first analyse the behaviour of $\mathbf{T}_\delta^\nu$ and $\mathbf{T}_\delta^*$. Suppose that $W_1(x) \geq W_2(x)$ for all $x \in \mathbb{X}_1$, then for a given map $\nu : \mathbb{X}_1 \to \mathcal{P}(\mathbb{U}_1, \mathcal{B}(\mathbb{U}_1))$ we have $\mathbf{T}_\delta^\nu(W_1)(x) \geq \mathbf{T}_\delta^\nu(W_2)(x)$, hence $\mathbf{T}_\delta^*(W_1)(x) \geq \mathbf{T}_\delta^*(W_2)(x)$. Then the series $V_N^\delta, V_{N-1}^\delta, \ldots, V_0^\delta$ constructed with $V_k^\delta = \mathbf{T}_\delta^{\mu_k} V_{k+1}^\delta$ and $V_N^\delta = 0$, is monotonically increasing. Therefore, for a given policy μ the functions $\{(\mathbf{T}_\delta^{\mu_k})^q(V)\}_{q\geq 0}$ initialised with $V = 0$ is point-wise converging since it is monotonically increasing and upper bounded. Additionally, the same holds for the function $\{(\mathbf{T}_\delta^*)^q(V)\}_{q\geq 0}$ initialised with $V = 0$. For unbounded reachability, this yields the following result.

Corollary 1 *A target set* $K \subset \mathbb{Y}$ *of gMDP* $\mathbf{M}_1$ *is reached* $(0,\delta)$*-robustly with time-homogeneous Markov policy* μ *and with probability* $r^\mu(K_{\mathbb{X}_1})$*, where*

$$r^\mu(K_{\mathbb{X}_1}) := \mathbf{L}\left(\int_{\mathbb{X}_1} [\mathbf{1}_{K_{\mathbb{X}_1}}(x) + \mathbf{1}_{\mathbb{X}_1 \setminus K_{\mathbb{X}_1}}(x) V^\delta(x)]\pi_1(dx) - \delta\right), \tag{12}$$

with $V^\delta : \mathbb{X}_1 \to [0,1]$ *the solution of* $V^\delta = \mathbf{T}_\delta^\mu(V^\delta)$ *computed as the limit of the sequence* $\{(\mathbf{T}_\delta^\mu)^q(V)\}_{q\geq 0}$ *that is initialised with* $V = 0$*. If* $V^{\delta,*}$ *is computed similarly as the solution of* $V^{\delta,*} = \mathbf{T}_\delta^*(V^{\delta,*})$ *and* $\mu^* \in \arg\sup \mathbf{T}_\delta^\mu(V^{\delta,*})$ *then we call* μ^* *the optimal* $(0,\delta)$*-robust policy.*

3.3 (ϵ,δ)-Robust probabilistic reachability

Consider a target set $K \subset \mathbb{Y}$, let K^ϵ be the largest Borel measurable set such that

$$K^\epsilon \subset \{y \mid \forall \bar{y} \in \mathbb{Y} \text{ with } \mathbf{d}_\mathbb{Y}(\bar{y}, y) \leq \epsilon : \bar{y} \in K\}. \tag{13}$$

We can now introduce an eroded version of the original target set $K_{\mathbb{X}_1}$ as $K^{\epsilon}_{\mathbb{X}_1} := h_1^{-1}(K^{\epsilon})$, such that for any pair x_1, x_2 if $x_1 \in K^{\epsilon}_{\mathbb{X}_1}$ and $x_1 \mathcal{R} x_2$ then $h_2(x_2) \in K$. As a consequence of Theorem 1 and of Corollary 1, we can evaluate the (ϵ, δ)-robust reachability with respect to $K^{\epsilon}_{\mathbb{X}_1}$.

Corollary 2 ((ϵ, δ)-robust probabilistic reachability) *A target set $K \subset \mathbb{Y}$ is reached (ϵ, δ)-robustly with Markov policy μ for $\mathbf{M}_1$ with respect to $r^{(\delta,\epsilon)}(K_{\mathbb{X}_1}, N)$ (or $r^{(\delta,\epsilon)}(K_{\mathbb{X}_1})$) if it reaches the target set K^{ϵ} as in (13), $(0, \delta)$-robustly with probability $r^{\mu}(K^{\epsilon}_{\mathbb{X}_1}, N)$ defined in (11) (or $r^{\mu}(K^{\epsilon}_{\mathbb{X}_1})$ defined in (12)).*

Hence for any model $\mathbf{M}_2$ that is in an (ϵ, δ)-probabilistic simulation relation with $\mathbf{M}_1$, the combination of Theorem 1 and Corollaries 1 and 2, mean that we can verify probabilistic reachability robustly over $\mathbf{M}_1$, and moreover, that we can synthesise a robust controller maximising the satisfaction probability robustly.

3.4 Upper-bounding the probabilistic reachability

We now question whether we can quantify an upper bound on a reachability probability using an approximate model $\mathbf{M}_1$. Consider the operator $\mathbf{T}^{\mu}_{-\delta}$, defined as

$$\mathbf{T}^{\mu}_{-\delta}(V)(x) := \mathbf{L}\left(\int_{\mathbb{X}_1} [\mathbf{1}_{K^{-\epsilon}_{\mathbb{X}_1}}(\bar{x}) + \mathbf{1}_{\mathbb{X}_1 \setminus K^{-\epsilon}_{\mathbb{X}_1}}(\bar{x}) V(\bar{x})] \mathbb{T}(d\bar{x}|x, \mu(x)) \ + \delta \right) \quad (14)$$

with $K^{-\epsilon}_{\mathbb{X}_1} := h_1^{-1}\left(K^{-\epsilon}\right)$ and $K^{-\epsilon} := \{y + y_{\epsilon} \in \mathbb{Y} | y \in K \text{ and } \mathbf{d}_{\mathbb{Y}}(0, y_{\epsilon}) \le \epsilon\}$.

Theorem 2. *Consider a target set $K \subset \mathbb{Y}$, then an upper bound on the maximal reachability $r^{\mu}(K_{\mathbb{X}_2}, N)$ of $\mathbf{M}_2$ for $\mathbf{M}_1 \succeq^{\delta}_{\epsilon} \mathbf{M}_2$ can be given as*

$$\forall \mu : \ r^{\mu}(K_{\mathbb{X}_2}, N) \le r^{(-\delta,-\epsilon)}(K_{\mathbb{X}_1}, N) \quad (15)$$

for which $r^{(-\delta,-\epsilon)}(K_{\mathbb{X}_1}, N) := r^{(-\delta)}(K^{-\epsilon}_{\mathbb{X}_1}, N)$ is computed with $\mathbf{M}_1$ as follows,

$$r^{(-\delta,-\epsilon)}(K_{\mathbb{X}_1}, N) := \mathbf{L}\left(\int_{\mathbb{X}_1} [\mathbf{1}_{K^{-\epsilon}_{\mathbb{X}_1}}(x) + \mathbf{1}_{\mathbb{X}_1 \setminus K^{-\epsilon}_{\mathbb{X}_1}}(x) V^{\delta}_0(x)] \pi_1(dx) + \delta \right)$$

with $V^{\delta}_k : \mathbb{X}_1 \to [0, 1]$ such that $V^{\delta}_k = \sup_{\mu} \mathbf{T}^{\mu}_{-\delta}(V^{\delta}_{k+1})$ and $V^{\delta}_N = 0$.

4 Temporal logic control leveraging (ϵ, δ)-probabilistic simulation relations

We extend the results on robust probabilistic reachability to the scLTL properties in Def. 4. For this purpose, we introduce a model known as Deterministic Finite-state Automaton (DFA).

Definition 8 (DFA). *A DFA is a tuple $\mathcal{A} = (Q, q_0, \Sigma, F, \mathtt{t})$, where Q is a finite set of locations, $q_0 \in Q$ is the initial location, Σ is a finite set, $F \subseteq Q$ is a set of accept locations, and $\mathtt{t} : Q \times \Sigma \to Q$ is a transition function.*

A finite word composed of letters of the alphabet defined by Σ, i.e., $\omega = (\omega(0), \ldots, \omega(n)) \in \Sigma^n$, is accepted by a DFA $\mathcal{A}$ if there exists a finite run $q = (q(0), \ldots, q(n+1)) \in Q^{n+2}$ such that $q(0) = q_0$, $q(i+1) = \mathsf{t}(q(i), \omega(i))$ for all $0 \le i \le n$ and $q(n+1) \in F$. Similarly, we say that an infinite word $w \in \mathfrak{S}$ is accepted by a DFA $\mathcal{A}$ if there exists a finite prefix of ω accepted by $\mathcal{A}$ as a finite word. More precisely, an infinite word $\omega \in \Sigma^{\mathbb{N}_0}$ is accepted by $\mathcal{A}$ if and only if there exists an infinite run $q \in Q^{\mathbb{N}_0}$ such that $q(0) = q_0$, $q(i+1) = \mathsf{t}(q(i), \omega(i))$ for all $i \in \mathbb{N}_0$ and there exists $j \in \mathbb{N}_0$ such that $q(j) \in F$. The accepted language of $\mathcal{A}$, denoted $\mathcal{L}(\mathcal{A})$, is the set of all words accepted by $\mathcal{A}$. For every scLTL property ψ, c.f. Definition 4, there exists a DFA $\mathcal{A}_\psi$ such that

$$\omega \vDash \psi \iff \omega \in \mathcal{L}(\mathcal{A}_\psi). \tag{16}$$

As a result, the satisfaction of the property ψ now becomes equivalent to the reaching of the accept locations in the DFA. We use the DFA $\mathcal{A}_\psi$ to specify properties of the gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ as follows. Remember that $\mathsf{L} : \mathbb{Y} \to \Sigma$ is a given measurable function. To each output $y \in \mathbb{Y}$ it assigns the letter $\mathsf{L}(y) \in \Sigma$. Given a control strategy $\mathbf{C}$, we can define the probability that a path of $\mathbf{M}$ satisfies a scLTL property ψ, i.e. $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi))$.

We can reduce the computation of $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi))$ over the traces ω of $\mathbf{M}$ to the reachability problem over another gMDP $\mathbf{M} \otimes \mathcal{A}_\psi$, which we refer to as a product of the gMDP $\mathbf{M}$ and the automaton $\mathcal{A}_\psi$. This was originally derived in [22] for MDPs. We give a similar definition of the product construction as.

Definition 9 (Product between automaton and gMDP). *Given a gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$, a finite alphabet Σ, a labelling function $\mathsf{L} : \mathbb{Y} \to \Sigma$ and a DFA $\mathcal{A}_\psi = (Q, q_0, \Sigma, F, \mathsf{t})$, we define the product between $\mathbf{M}$ and $\mathcal{A}_\psi$ to be another gMDP denoted as $\mathbf{M} \otimes \mathcal{A}_\psi = (\bar{\mathbb{X}}, \bar{\pi}, \bar{\mathbb{T}}, \mathbb{U}, \bar{h}, \mathbb{Y})$. Here $\bar{\mathbb{X}} = \mathbb{X} \times Q$, $\bar{h}(x, q) = h(x)$ for any $(x, q) \in \bar{\mathbb{X}}$, and*

$$\bar{\mathbb{T}}(A \times \{q'\} | x, q, u) = \int_{\tilde{x} \in A} \mathbf{1}(q' = \mathsf{t}(q, \mathsf{L}(h(\tilde{x})))) \cdot \mathbb{T}(d\tilde{x} | x, u).$$

initialised with $\bar{\pi}(dx, q) = \pi(dx)\mathsf{t}(q_0, \mathsf{L}(h(x))))$.

The quantity $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi))$ can be related to the reachability probability over the gMDP $\mathbf{M} \otimes \mathcal{A}$ with a goal state $G := \mathbb{X} \times F$, as it was shown to be the case for MDPs in [22].

Lemma 1. *Given gMDP $\mathbf{M}$, alphabet Σ, labelling function L, and scLTL specification ψ modelled with DFA $\mathcal{A}_\psi$, it holds that*

$$\mathbb{P}_{\mathbf{C}(\mu,\psi)\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi)) = \mathbb{P}_{\mu\times(\mathcal{A}_\psi\otimes\mathbf{M})}(\Diamond G))$$

For any Mrkov policy μ on the product space of $\mathcal{A}_\psi \otimes \mathbf{M}$ and any control strategy $\mathbf{C}(\mu, \psi)$ on $\mathbf{M}$ with properly defined mappings between μ and the control strategy.

4.1 δ-Robust satisfaction of scLTL properties

In the following we analyse the robust satisfaction of scLTL specifications, which are temporal specifications that go beyond reachability properties defined on $\mathbf{M}$. The probabily of satisfying such a temporal specification can be quantified as a reachability probability with respect to $\mathbf{M} \otimes \mathcal{A}_\psi$. For two gMDPs $\mathbf{M}_1$ and $\mathbf{M}_2$, subject to $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$, we show that ($\delta$-approximate) probabilistic simulation relations are preserved under a product with a DFA.

Theorem 3. *Let $\mathbf{M}_i, i = 1, 2$, $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, be two gMDPs such that $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$ and $\mathcal{A} = (Q, q_0, \Sigma, F, \mathsf{t})$ be an automaton. For any labeling function $\mathsf{L} : \mathbb{Y} \to \Sigma$ we have $\mathbf{M}_1 \otimes \mathcal{A} \preceq_0^\delta \mathbf{M}_2 \otimes \mathcal{A}$.*

This theorem enables us to quantify temporal logic properties for $\mathbf{M}_2$ with respect to $\mathbf{M}_1$. Consider a scLTL property ψ with a corresponding DFA $\mathcal{A}_\psi$ and two gMDPs $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}_\mathbb{Y}$ for which $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$. If there exists a Markov policy μ for $\mathbf{M}_1 \otimes \mathcal{A}_\psi$ such that the accepting states are reached with δ-robust probability p, then there exists a control strategy $\mathbf{C}_2$ for $\mathbf{M}_2$ such that the accepting states of $\mathcal{A}_\psi$ are reached with probability p under the evolution of $\mathbf{C}_2 \times \mathbf{M}_2$. More precisely, denote with $\bar{\mathbb{X}} := \mathbb{X}_1 \times Q$ the state space of $\mathbf{M}_1 \otimes \mathcal{A}_\psi$, then the mapping $\mathbf{T}_\delta^\nu$ becomes

$$\begin{aligned} \mathbf{T}_\delta^\nu(V)(x_1, q) = \mathbf{L}\Big(\int_{\mathbb{X}_1} &\big[\mathbf{1}_F(\mathsf{t}_x(q, x_1')) + \mathbf{1}_{Q\setminus F}(\mathsf{t}_x(q, x_1'))V(x_1', \mathsf{t}_x(q, x_1'))\big] \\ &\times \mathbb{T}(dx_1'|x_1, \nu(x_1, q)) - \delta\Big) \end{aligned} \tag{17}$$

with $\mathsf{t}_x(q, x_1') := \mathsf{t}(q, \mathsf{L}(h_1(x_1'))))$. For $\mathbf{T}_\delta^\mu(V)(x_1, q) = V(x_1, q)$, the δ-robust reachability probability is defined as

$$\begin{aligned} r^\mu(F \times \mathbb{X}_1) = \mathbf{L}\Big(\int_{\mathbb{X}_1} &[\mathbf{1}_F(\mathsf{t}_x(q_0, x_1)) + \mathbf{1}_{Q\setminus F}(\mathsf{t}_x(q_0, x_1)) \\ &\times V(x_1, \mathsf{t}_x(q_0, x_1))]\pi(dx_1) - \delta\Big). \end{aligned} \tag{18}$$

4.2 (ϵ, δ)-Robust satisfaction of scLTL properties

We now integrate the ϵ error in the output space into the robust synthesis problem via the effect it has on the labelling. Given $\mathsf{L} : \mathbb{Y} \to \Sigma$, we define $\mathsf{L}_\epsilon : \mathbb{Y} \to 2^\Sigma$ as

$$\mathsf{L}_\epsilon(y) := \{q \in \Sigma | \exists (y, y_\epsilon) : d_\mathbb{Y}(y, y_\epsilon) \le \epsilon \text{ and } q = \mathsf{L}(y_\epsilon)\}. \tag{19}$$

Consider $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$ with $\mathcal{R}_\epsilon$, then for all $(x_1, x_2) \in \mathcal{R}_\epsilon$, it holds that $\mathsf{L}(h_2(x_2)) \in \mathsf{L}_\epsilon(h_1(x_1))$. In Figure 1, an output space together with its labels is depicted. When taking into account an ϵ error in the output, the labelling becomes non-deterministic in some regions. This can be observed on the right of Figure 1.

Instead of integrating this relaxed labelling into the product construction of a given gMDP, we will immediately adapt the δ-robust reachability computations (17) to deal with this non-determinism. Consider the (ϵ, δ)-robust operator

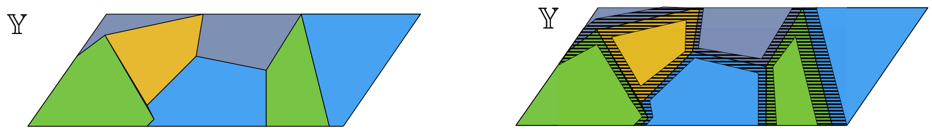


Fig. 1: A typical labelling over the output space. On the left a normal labelling is given, on the right the labelling is non-deterministic due to the output error.

$\mathbf{T}^{\nu}_{\epsilon,\delta}(V)(x_1, q)$ defined as

$$\mathbf{T}^{\nu}_{\epsilon,\delta}(V)(x_1, q) = \mathbf{L}\Big(\int_{\mathbb{X}_1} \min_{q' \in \bar{\mathsf{t}}_x(q, x_1')} [\mathbf{1}_F(q') + \mathbf{1}_{Q\setminus F}(q') V_{k+1}(x_1', q')] \times \\ \mathbb{T}(dx_1' | x_1, \nu(x_1, q)) - \delta\Big)$$

with $\bar{\mathsf{t}}_x(q, x_1) := \{\mathsf{t}(q, \alpha) \text{ with } \alpha \in \mathsf{L}_\epsilon(h_1(x_1))\}$. For a time-homogeneous Markov policy μ and $V(x_1, q)$ that satisfy $\mathbf{T}^{\mu}_{\epsilon,\delta}(V)(x_1, q) = V(x_1, q)$, the δ-robust reachability probability is defined as

$$r^{\epsilon,\delta}(F \times \mathbb{X}_1) = \mathbf{L}\Big(\int_{\mathbb{X}_1} \min_{q' \in \bar{\mathsf{t}}_x(q_0, x_1)} [1_F(q') + \mathbf{1}_{Q\setminus F}(q') V(x_1, q')] \pi(dx_1) - \delta\Big).$$

Consider a scLTL property ψ and the corresponding $\mathcal{A}_\psi$ with goal states F. If $F \times \mathbb{X}_1$ is δ-robust reachable with respect to $r^{\epsilon,\delta}(F \times \mathbb{X}_1)$, then we can refine μ to $\mathbf{C}_2(\mu, \psi)$ such that ψ is satisfied by $\mathbf{C}_2(\mu, \psi) \times \mathbf{M}_2$ with a probability $p \geq r^{\epsilon,\delta}(F \times \mathbb{X}_1)$. Of course the apparent non-determinism due to the relaxed labelling will be resolved in the refined control strategy by selecting the labels of the concrete model.

Building on Corollary 1, we can now also maximise the (ϵ, δ)-robust probability using $\mathbf{T}^{*}_{\epsilon,\delta}$, defined as $\mathbf{T}^{*}_{\epsilon,\delta}(V)(x_1, q) := \sup_\mu \mathbf{T}^{\mu}_{\epsilon,\delta}(V)(x_1, q)$, which yields an optimised robust Markov policy as

$$\mu^*(x_1, q) \in \arg\sup_{\mu} \mathbf{T}^{\mu}_{\epsilon,\delta}(V^*)(x_1, q) \text{ for } \mathbf{T}^{*}_{\epsilon,\delta}(V^*)(x_1, q) = V^*(x_1, q).$$

Hence, we have shown that we can leverage probabilistic simulation relations to use approximate models for the controller synthesis and verification of both probabilistic reachability (cf. Problem 1) and scLTL properties (cf. Problem 2).

5 Computational implementation and toy example

In this section, we apply our results to formal controller synthesis for stochastic linear dynamical systems. Existing formal controller synthesis results for this class of models either rely on model order reduction [19] or use abstraction techniques as finite-state MDPs [9,22]. Our new results combine these two approaches in one framework that benefits from both of them. More precisely, our abstract model used for synthesis is obtained by discretising the state space of a reduced-order version of the concrete model.

Concrete model. Consider the following linear dynamical model $\mathbf{M}_2$:

$$x_2(t+1) = A_2x_2(t) + B_2u_2(t) + B_{w2}w(t), \quad x_2(0) = x_{20} \in \mathbb{X}_2$$
$$y_2(t) = C_2x_2(t), \quad t = 0, 1, 2, \ldots,$$

where $x_2(\cdot) \in \mathbb{X}_2 \subset \mathbb{R}^n$, $u_2(\cdot) \in \mathbb{U}_2 \subset \mathbb{R}^m$, and $y_2(\cdot) \in \mathbb{Y} \subset \mathbb{R}^p$. Matrices A_2, B_2, B_{w2}, and C_2 have appropriate dimension and $w(\cdot)$ are iid random variables with standard multivariate Gaussian distributions.

Constructing the abstract model. Construction of the abstract model relies on partitioning a new space $\mathbb{X}_s \subset \mathbb{R}^{n_s}$, where $n_s < n$, as $\{\mathbb{A}_i \subset \mathbb{X}_s,\ i = 1, 2, \ldots, l\}$. Over this partition, we select representative points $\{z_i \in \mathbb{A}_i,\ i = 1, 2, \ldots, l\}$, which we call $\mathbb{X}_1$ and which becomes the state space of the abstract model $\mathbf{M}_1$. Introduce the operator $\Pi : \mathbb{X}_s \to \mathbb{X}_1$ that assigns to any $x_1 \in \mathbb{A}_i$, $i \in \{1, \ldots, l\}$ the representative point of $\mathbb{A}_i$, $z_i = \Pi(x_1)$.
Next we provide a dynamical characterisation of $\mathbf{M}_1$. The state evolution of $\mathbf{M}_1$ is written as

$$x_1(t+1) = \Pi\left(A_1x_1(t) + B_1u_1(t) + B_{w1}w(t)\right), \quad x_1(0) = x_{10} \in \mathbb{X}_1$$
$$y_1(t) = C_1x_1(t), \quad t = 0, 1, 2, \ldots,$$

with state $x_1(\cdot) \in \mathbb{X}_1$, input $u_1(\cdot) \in \mathbb{U}_1$, and output $y_1(\cdot) \in \mathbb{Y}$, and matrices A_1, B_1, B_{w1}, C_1 of appropriate dimensions. Note that the noise term $w(t)$ in $\mathbf{M}_1$ is the same as the one in $\mathbf{M}_2$, thereby allowing to define a lifting $\mathbb{W}_{\mathbb{T}}$.

Computing the (ϵ, δ)-simulation relation. Consider the linear interface function $u_2 = Ru_1 + Qx_1 + K(x_2 - Px_1)$, such that $PA_1 = A_2P + B_2Q$ for some matrix P. Define the relation $(x_1, x_2) \in \mathcal{R}_\delta^\epsilon$ iff $(x_2 - Px_1)^T M(x_2 - Px_1) \leq \epsilon^2$. Next we check conditions of Def. 6 under which $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$. It is guaranteed that $d_{\mathbb{Y}}(y_1, y_2) \leq \epsilon$ for any $(x_1, x_2) \in \mathcal{R}_\delta^\epsilon$ (cf. **APS1** in Def. 6) if $C_1 = C_2P$, and $C_2^TC_2 \leq M$. Condition **APS2** in Def. 6 holds if c_w is selected such that $\mathbb{P}(w^Tw \leq c_w) \geq 1 - \delta$ and the following inequality

$$(\bar{A}\bar{x} + \bar{B}u_1 + \bar{B}_w w + P\beta)^T M(\bar{A}\bar{x} + \bar{B}u_1 + \bar{B}_w w + P\beta) \leq \epsilon^2 \tag{20}$$

is satisfied for any $\bar{x}, u_1, w, \beta$ such that $w^Tw \leq c_w$, $u_1^2 \leq c_u$, $\bar{x}^TM\bar{x} \leq \epsilon^2$, $|\beta| \leq \boldsymbol{\delta}$. The matrices in (20) are defined as $\bar{A} := A_2 + B_2K$, $\bar{B} := B_2R - PB_1$, and $\bar{B}_{w2} := B_{w2} - PB_{w1}$. Vector $\boldsymbol{\delta}$ is the diameter of the partition $\{\mathbb{A}_i, i = 1, \ldots, l\}$, which satisfies $|x_s - x_s'| \leq \boldsymbol{\delta}$ component-wise for any $x_s, x_s' \in \mathbb{A}_i$ and any $i \in \{1, 2, \ldots, l\}$. Condition (20) can be checked using LMIs and the S-procedure [6].

For Condition **APS3** in Def. 6, we now question when there exists a deterministic initial state x_{10} for a given deterministic initial state x_{20} satisfying $(x_{20} - Px_{10})^TM(x_{20} - Px_{10}) \leq \epsilon^2$. We can choose $x_{10} := \hat{P}x_{20}$ with $\hat{P} := (P^TMP)^{-1}P^TM$ which is the minimum of left-hand side, and select it as the representative point of the associated partition set in $\mathbf{M}_1$. Alternatively, when the representative point cannot be freely chosen, we select $x_{10} = \Pi(\hat{P}x_{20})$.

In the former case, there exists an initial state x_{10} if $x_{20}^T M(\mathbb{I}_n - P\hat{P})x_{20} \leq \epsilon^2$, or, dually, ϵ is lower bounded for a given x_{20}. In the latter case, if $\|M^{\frac{1}{2}}(\mathbb{I}_n - P\hat{P})x_{20}\| + \|M^{\frac{1}{2}}P\delta\| \leq \epsilon$, then there exists an initial state x_{10} satisfying Condition **APS3**.

Toy example. We consider the specification $\psi = \Diamond\Box^{\leq n_2}\{y \in K\}$ which encodes reach and stay over bounded time intervals. The associated DFA is given in Figure 2, together with an illustration of a potential application of this toy example.

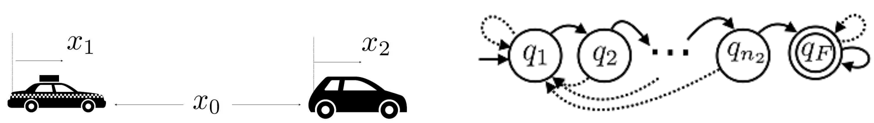


Fig. 2: A game of tag: $\Diamond\Box^{\leq n_2}\{x_a \in K\}$

Consider the original model $\mathbf{M}_2$, which is a 3 dimensional model with output $y_1(t) = x_a$ and

$$\begin{aligned} x_a(t+1) &= x_a(t) - a_1(x_b(t) - x_c(t)) - a_2 u(t) + a_3 w(t) \\ x_b(t+1) &= b x_b(t) + u(t) \\ x_c(t+1) &= c_1 x_c(t) + c_2 w(t) \end{aligned} \tag{21}$$

with $a_1 = .3$, $a_2 = 0.03$, $a_3 = 6e$-3, $b = c_1 = .8$ and $c_2 = .1$. For the game we consider the case with $n_2 = 3$. As in [13], we compute the lower dimensional

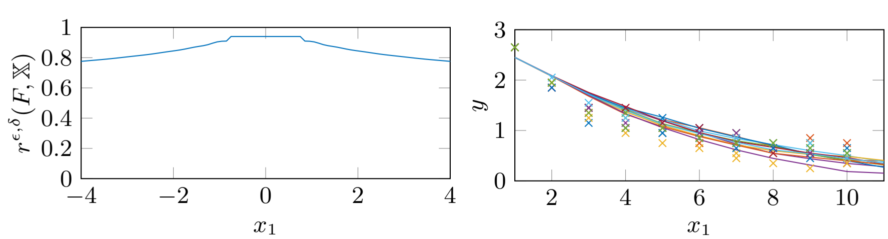


Fig. 3: On the left: (ϵ, δ)-robust satisfaction probability of $\Diamond\Box^{\leq n_2}\{y \in [-2, 2]\}$ with $\epsilon = 1.2266$ and $\delta = 0.03$. On the right: simulation runs for the original model and the abstract model with the composed robust controller.

model via the balance truncation of the original controlled model with a suitable feedback gain ($K = [-0.77380.9369 - 0.6829]$). In Figure 3, we give an example of such a robust temporal logic computation. On the right side of the figure, 10 simulation runs are given [v], the crosses are the outputs of $\mathbf{M}_1$ whereas the lines are the outputs of $\mathbf{M}_2$.

6 Conclusions and future work

In this paper, we have introduced a new robust way of synthesising control strategies and verifying probabilistic temporal logic properties. Beyond this theoretical contribution, future work will focus on the computational aspects of this approach to prepare for application on realistic sized problems.

References

1. A. Abate. Approximation metrics based on probabilistic bisimulations for general state-space Markov processes: a survey. *ENTCS*, 297:3–25, 2013.
2. A. Abate, J.-p. Katoen, and A. Mereacre. Quantitative Automata Model Checking of Autonomous Stochastic Hybrid Systems. In *Proc. 14th ACM Int. Conf. Hybrid Syst. Comput. Control*, pages 83–92, 2011.
3. A. Abate, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic Reachability and Safety for Controlled Discrete Time Stochastic Hybrid Systems. *Automatica*, 44(11):2724–2734, 2008.
4. D. Bertsekas and S. E. Shreve. *Stochastic Optimal control : The discrete time case.* Athena Scientific, 1996.
5. V. I. Bogachev. *Measure theory.* Springer Science & Business Media, 2007.
6. S. Boyd and L. Vandenberghe. *Convex Optimization.* CUP, Cambridge, 2004.
7. J. Desharnais, F. Laviolette, and M. Tracol. Approximate analysis of probabilistic processes: Logic, simulation and games. *Conf. on Quantitative Evaluation of Systems*, pages 264–273, Sept. 2008.
8. A. D'Innocenzo, A. Abate, and J.-P. Katoen. Robust PCTL model checking. In *Proceedings of the 15th ACM international conference on Hybrid Systems: computation and control*, pages 275–285, 2012.
9. S. Esmaeil Zadeh Soudjani and A. Abate. Adaptive and sequential gridding procedures for the abstraction and verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, 12(2):921–956, 2013.
10. S. Esmaeil Zadeh Soudjani, C. Gevaerts, and A. Abate. FAUST ^{2}: Formal Abstractions of Uncountable-STate STochastic Processes. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, Lecture Notes in Computer Science, pages 272–286. Springer Berlin Heidelberg, 2015.
11. A. Girard and G. J. Pappas. Hierarchical control system design using approximate simulation. *Automatica*, 45(2):566–571, 2009.
12. S. Haesaert, A. Abate, and P. M. J. Van den Hof. Verification of general markov decision processes by approximate similarity relations and policy refinement. In *13th International Conference on Quantitative Evaluation of Systems, QEST 2016, Quebec City, Canada, August 23-25*, pages 227–243, 2016.
13. S. Haesaert, S. Esmaeil Zadeh Soudjani, and A. Abate. Verification of general markov decision processes by approximate similarity relations and policy refinement. *SIAM Journal on Control and Optimization*, 55(4):2333–2367, 2017.

[v] initiated at $x_a = 2.45$, $x_b = 2.5$ and $x_c = 1.3$.

14. S. Haesaert, S. Soudjani, and A. Abate. Temporal logic control of general Markov decision processes by approximate policy refinement. Technical report, 10 2017.
15. O. Hernández-Lerma and J. B. Lasserre. *Discrete-time Markov control processes*, volume 30 of *Applications of Mathematics (New York)*. Springer Verlag, 1996.
16. O. Kupferman and M. Y. Vardi. Model checking of safety properties. *Formal Methods in System Design*, pages 291–314, 2001.
17. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of probabilistic real-time systems. In G. Gopalakrishnan and S. Qadeer, editors, *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.
18. K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.
19. A. Lavaei, S. Esmaeil Zadeh Soudjani, R. Majumdar, and M. Zamani. Compositional Abstractions of Interconnected Discrete-Time Stochastic Control Systems. *ArXiv e-prints*, Sept. 2017.
20. S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., 1993.
21. M. G. Safonov and R. Chiang. A Schur method for balanced-truncation model reduction. *IEEE Transactions on Automatic Control*, 34(7):729–733, 1989.
22. I. Tkachev, A. Mereacre, J. Katoen, and A. Abate. Quantitative automata-based controller synthesis for non-autonomous stochastic hybrid systems. In *HSCC*, pages 293–302, 2013.

A Additional proofs

*Proof (**Proof of Theorem 1**).* Consider any two models $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}_{\mathbb{Y}}$ with $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$. A Markov policy $\mu = (\mu_0, \ldots, \mu_N)$ for $\mathbf{M}_1$ can be refined to $\mathbf{M}_2$. Technically, this means that we first can write it as a control strategy for which we can do control refined as proven in [13]. Denote the control strategy that refines the Markov policy as $\mathbf{C}_2$. Then the composed system $\mathbf{C}_2 \times \mathbf{M}_2$ contains transitions over $\mathbb{X}_1 \times \mathbb{X}_2$ with stochastic transition kernels defined as

$$\mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_1), x_1, x_2) \text{ for } k \in \{0, \ldots, N-1\}$$

We now need to show that $r^\mu(K_{\mathbb{X}_1}, N)$ defined in (11) is a lower bound for the probability that target set K is reached by $\mathbf{C}_2 \times \mathbf{M}_2$ in N time steps, $\mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}(\Diamond^{\leq N} K)$. Onserve that the following holds

$$\mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}\left[((x_1, x_2) \in \mathcal{R}) \mathcal{U}^{\leq N} \left((x_1, x_2) \in \mathcal{R} \wedge (x_1 \in K_{\mathbb{X}_1})\right)\right] \leq \mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}\left[\Diamond^{\leq N}(y_2 \in K)\right], \tag{22}$$

since $((x_1, x_2) \in R \wedge (x_1 \in K_{\mathbb{X}_1}))$ implies $h_2(x_2) = h_1(x_1)$ and $h_1(x_1) \in K$. The left-hand side of (22) can be computed via backward recursion of reach-avoid property. For this the safe set (i.e., the complement of the avoid set) is $\mathcal{R}$ and the reach set is $G := (K_{\mathbb{X}_1} \times \mathbb{X}_2) \cap \mathcal{R}$. The value functions $\{V_k,\, k = 0, \ldots, N\}$ in the backward recursion are initialised with $V_N = 0$ and computed as

$$V_k(x_1, x_2) = \int_{\mathbb{X}_1 \times \mathbb{X}_2} \left[\mathbf{1}_G(\bar{x}_1) + \mathbf{1}_{\mathcal{R} \setminus G}(\bar{x}_1, \bar{x}_2) V_{k+1}(\bar{x}_1, \bar{x}_2)\right] \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_1), x_1, x_2).$$

This can be also written down as

$$V_k(x_1,x_2) = \int_{\mathcal{R}} \left[\mathbf{1}_{(K_{\mathbb{X}_1}\times\mathbb{X}_2)}(\bar{x}_1) + \mathbf{1}_{(\mathbb{X}_1\setminus K_{\mathbb{X}_1})\times\mathbb{X}_2}(\bar{x}_1,\bar{x}_2)V_{k+1}(\bar{x}_1,\bar{x}_2)\right] \\ \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1\times d\bar{x}_2|\mu_k(x_1),x_1,x_2).$$

We now want to compute a lower bound on $V_k(x_1,x_2)$ based on backwards computations over $\mathbf{M}_1$. The value functions $V_k^\delta : \mathbb{X}_1 \to [0,1]$ are defined inductively as $V_k^\delta = \mathbf{T}_\delta^{\mu_k}(V_{k+1}^\delta)$, which is

$$V_k^\delta(x_1) := \mathbf{L}\left(\int_{\mathbb{X}_1}\left[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1\setminus K_{\mathbb{X}_1}}(\bar{x}_1)V_{k+1}^\delta(\bar{x}_1)\right]\mathbb{T}(d\bar{x}_1|x_1,\mu_k(x_1)) - \delta\right)$$

and initialised with $V_k^\delta = 0$. With focus on the above two recursions we claim that if $V_{k+1}^\delta(x_1)$ is a lower bound for $V_{k+1}(x_1,x_2)$ for all $(x_1,x_2)\in\mathcal{R}$, then $V_k^\delta(x_1)$ is also a lower bound for $V_k(x_1,x_2)$. Once we prove this claim, by induction we get that $V_0^\delta(x_1)$ is a lower bound for $V_{k+1}(x_1,x_2)$. As a result, the δ-robust probability can be computed as

$$r^\mu(K_{\mathbb{X}_1},N) := \mathbf{L}\left(\int_{\mathbb{X}_1}[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1\setminus K_{\mathbb{X}_1}}(\bar{x}_1)V_0^\delta(\bar{x}_1)]\pi(d\bar{x}_1) - \delta\right).$$

To prove the claim we need to define $\tilde{V}_k^\delta : \mathbb{X}_1 \to [-\delta,1]$ as

$$\tilde{V}_k^\delta(x_1) := \int_{\mathbb{X}_1}\left[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1\setminus K_{\mathbb{X}_1}}(\bar{x}_1)V_{k+1}^\delta(\bar{x}_1)\right]\mathbb{T}(d\bar{x}_1|x_1,\mu_k(x_1)) - \delta,$$

such that $V_k^\delta(x_1) = \mathbf{L}\left(\tilde{V}_k^\delta(x_1)\right)$. For any $(x_1,x_2)\in\mathcal{R}$, we have

$$\begin{aligned}
\tilde{V}_k^\delta(x_1) + \delta &= \int\limits_{\mathbb{X}_1}\left[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1\setminus K_{\mathbb{X}_1}}(\bar{x}_1)V_{k+1}^\delta(\bar{x}_1)\right]\mathbb{T}(d\bar{x}_1|x_1,\mu_k(x_1)) \\
&= \int\limits_{\mathbb{X}_1\times\mathbb{X}_2}\left[\mathbf{1}_{(K_{\mathbb{X}_1}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{(\mathbb{X}_1\setminus K_{\mathbb{X}_1})\times\mathbb{X}_2}(\bar{x}_1,\bar{x}_2)V_{k+1}^\delta(\bar{x}_1)\right] \qquad (23) \\
&\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1\times d\bar{x}_2|\mu_k(x_1),x_1,x_2) \\
&\le \int\limits_{\mathcal{R}}\left[\mathbf{1}_{(K_{\mathbb{X}_1}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{(\mathbb{X}_1\setminus K_{\mathbb{X}_1})\times\mathbb{X}_2}(\bar{x}_1,\bar{x}_2)V_{k+1}(\bar{x}_1,\bar{x}_2)\right] \\
&\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1\times d\bar{x}_2|\mu_k(x_1),x_1,x_2) \\
&\qquad + \int\limits_{(\mathbb{X}_1\times\mathbb{X}_2)\setminus\mathcal{R}}\mathbb{W}_{\mathbb{T}}(d\bar{x}_1\times d\bar{x}_2|\mu_k(x_1),x_1,x_2) \qquad (24) \\
&\le \int\limits_{\mathcal{R}}\left[\mathbf{1}_{(K_{\mathbb{X}_1}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{(\mathbb{X}_1\setminus K_{\mathbb{X}_1})\times\mathbb{X}_2}(\bar{x}_1,\bar{x}_2)V_{k+1}(\bar{x}_1,\bar{x}_2)\right] \\
&\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1\times d\bar{x}_2|\mu_k(x_1),x_1,x_2) + \delta.
\end{aligned}$$

Note that (23) follows from the fact that $(x_1, x_2) \in \mathcal{R}$ and (24) holds since $V_{k+1}^{\delta}(\bar{x}_1)$ is upper bounded by 1 and is lower bounded by $V_{k+1}^{\delta}(\bar{x}_1, \bar{x}_2)$ over $\mathcal{R}$. Thus by the definition of V_k, we have also that $V_k(x_1, x_2) \leq \tilde{V}_k^{\delta}(x_1)$ for all $(x_1, x_2) \in \mathcal{R}$. Since V_k is also lower bounded by 0, we have that $V_k(x_1, x_2) \leq \mathbf{L}\left(\tilde{V}_k^{\delta}(x_1)\right) = V_k^{\delta}(x_1)$, which completes the proof.

*Proof (**Proof of Theorem 2**).* This proof upper-bounds the maximal reachability probability and follows an analogue path to the proof of Theorem 1. Consider any two models $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}_{\mathbb{Y}}$ with $\mathbf{M}_1 \succeq_{\epsilon}^{\delta} \mathbf{M}_2$. A Markov policy $\mu = (\mu_0, \ldots, \mu_N)$ for $\mathbf{M}_2$ can be refined to $\mathbf{M}_1$ such that the composed system $\mathbf{C}_1 \times \mathbf{M}_1$ contains transitions over $\mathbb{X}_1 \times \mathbb{X}_2$ with stochastic transition kernels defined as

$$\mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_2), x_1, x_2) \text{ for } k \in \{0, \ldots, N-1\}.$$

We now need to show that $\mathbb{P}_{\mu_2 \times \mathbf{M}_2}(\Diamond^{\leq N} K)$ has an upper bound computed as given in Theorem 2. Remark that $\Diamond^{\leq N}(y_2 \in K)$ implies the specification

$$\psi := [(x_1, x_2) \in \mathcal{R}]\, \mathcal{U}^{\leq N} \left[(x_1, x_2) \in \bar{\mathcal{R}} \vee (x_1 \in K_{\mathbb{X}_1}^{-\epsilon})\right]$$

with $\bar{\mathcal{R}} := \mathbb{X}_1 \times \mathbb{X}_2 \setminus \mathcal{R}$. Hence it follows that

$$\mathbb{P}_{\mu_2 \times \mathbf{M}_2}\left[\Diamond^{\leq N}(y_2 \in K)\right] \leq \mathbb{P}_{\mathbf{C}_1 \times \mathbf{M}_1}[\psi]\,. \tag{25}$$

The right-hand side of (25) can be computed via backward recursion of reach-avoid property. For this the safe set (i.e., the complement of the avoid set) is $\mathcal{R}$ and the reach set is $((K_{\mathbb{X}_1} \times \mathbb{X}_2) \cap \mathcal{R}) \cup \bar{\mathcal{R}}$. The value functions V_k in the backward recursion are initialised with $V_N = 0$ and computed as

$$\begin{aligned} V_k(x_1, x_2) = \int\limits_{\mathbb{X}_1 \times \mathbb{X}_2} & \Big[\mathbf{1}_{(K_{\mathbb{X}_1}^{-\epsilon} \times \mathbb{X}_2) \cap \mathcal{R}}(\bar{x}_1, \bar{x}_2) + \mathbf{1}_{\bar{\mathcal{R}}}(\bar{x}_1, \bar{x}_2) \\ & + \mathbf{1}_{\mathcal{R} \setminus (K_{\mathbb{X}_1}^{-\epsilon} \times \mathbb{X}_2)}(\bar{x}_1, \bar{x}_2) V_{k+1}(\bar{x}_1, \bar{x}_2)\Big]\, \mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_2), x_1, x_2). \end{aligned} \tag{26}$$

We now want to compute an upper bound on $V_k(x_1, x_2)$ based on backward recursion over $\mathbf{M}_1$. Let $V_k^{-\delta} : \mathbb{X}_1 \to [0, 1]$ be defined inductively as

$$V_k^{-\delta}(x_1) := \sup_{\bar{\mu}} \mathbf{L}\left(\int_{\mathbb{X}_1} \left[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1 \setminus K_{\mathbb{X}_1}}(\bar{x}_1) V_{k+1}^{-\delta}(\bar{x}_1)\right] \mathbb{T}(d\bar{x}_1 | \bar{\mu}_k(x_1), x_1) + \delta\right)$$

and initialised with $V_N^{-\delta} = 0$. We claim that if $V_{k+1}^{-\delta}(x_1) \geq V_{k+1}(x_1, x_2)$ for all $(x_1, x_2) \in \mathcal{R}$, then $V_k^{-\delta}(x_1) \geq V_k(x_1, x_2)$. Thus by induction we get $V_0^{-\delta}(x_1) \geq V_0(x_1, x_2)$. By repeating the same argument for initial measure π, we get Theorem 2.

In order to prove the above claim we need to define value functions $\tilde{V}_k^{\delta} : \mathbb{X}_1 \to [-\delta, 1]$ as

$$\tilde{V}_k^{-\delta}(x_1) := \sup_{\bar{\mu}} \left(\int_{\mathbb{X}_1} \left[\mathbf{1}_{K_{\mathbb{X}_1}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1 \setminus K_{\mathbb{X}_1}}(\bar{x}_1) V_{k+1}^{-\delta}(\bar{x}_1)\right] \mathbb{T}(d\bar{x}_1 | \bar{\mu}_k(x_1), x_1)\right) + \delta,$$

such that $V_k^{-\delta}(x_1) = \mathbf{L}\left(\tilde{V}_k^{-\delta}(x_1)\right)$ due to the fact that $\mathbf{L}$ and sup are interchangeable here. For any $(x_1, x_2) \in \mathcal{R}$ we have

$$\begin{aligned}
V_k(x_1,x_2) &\leq \int\limits_{\mathcal{R}} \Big[\mathbf{1}_{(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)\cap\mathcal{R}}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{\mathcal{R}\setminus(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2)V_{k+1}(\bar{x}_1,\bar{x}_2)\Big] \\
&\qquad\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_2), x_1, x_2) + \delta \\
&\leq \int\limits_{\mathcal{R}} \Big[\mathbf{1}_{(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{(\mathbb{X}_1\times\mathbb{X}_2)\setminus(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2)V_{k+1}^{-\delta}(\bar{x}_1)\Big] \\
&\qquad\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_2), x_1, x_2) + \delta \\
&\leq \int\limits_{\mathbb{X}_1\times\mathbb{X}_2} \Big[\mathbf{1}_{(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2) + \mathbf{1}_{(\mathbb{X}_1\times\mathbb{X}_2)\setminus(K_{\mathbb{X}_1}^{-\epsilon}\times\mathbb{X}_2)}(\bar{x}_1,\bar{x}_2)V_{k+1}^{-\delta}(\bar{x}_1)\Big] \\
&\qquad\qquad \times \mathbb{W}_{\mathbb{T}}(d\bar{x}_1 \times d\bar{x}_2 | \mu_k(x_2), x_1, x_2) + \delta \\
&\leq \sup_{\bar{\mu}} \int\limits_{\mathbb{X}_1} \Big[\mathbf{1}_{K_{\mathbb{X}_1}^{-\epsilon}}(\bar{x}_1) + \mathbf{1}_{\mathbb{X}_1\setminus K_{\mathbb{X}_1}^{-\epsilon}}(\bar{x}_1)V_{k+1}^{-\delta}(\bar{x}_1)\Big]\,\mathbb{T}(d\bar{x}_1|\bar{\mu}(x_1),x_1) + \delta \\
&= \tilde{V}_k^{-\delta}(x_1).
\end{aligned}$$

Since $V_k(x_1,x_2) : \mathbb{X}_1 \times \mathbb{X}_2 \to [0,1]$ it holds that if $V_k(x_1,x_2) \leq \tilde{V}_k^{-\delta}(x_1)$ then also $V_k(x_1,x_2) \leq \mathbf{L}\left(\tilde{V}_k^{-\delta}(x_1)\right) = V_k^{-\delta}(x_1)$. This completes the proof.

*Proof (**Proof of Theorem 3**).* Since $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$, according to Definition 6 there exists an interface function $\mathcal{U}_v$, a relation $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$, and a Borel measurable stochastic kernel $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1, x_1, x_2)$ such that

1. $\forall (x_1,x_2) \in \mathcal{R}$, $h_1(x_1) = h_2(x_2)$;
2. $\forall (x_1,x_2) \in \mathcal{R}$, $\forall u_1 \in \mathbb{U}_1$, $\mathbb{T}_1(\cdot|x_1,u_1)\ \bar{\mathcal{R}}_\delta\ \mathbb{T}_2(\cdot|x_2,\mathcal{U}_v(u_1,x_1,x_2))$, with lifted probability measure $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1,x_1,x_2)$;
3. $\pi_1 \bar{\mathcal{R}}_\delta \pi_2$.

Indicate the product gMDPs by $\mathbf{M}_i \otimes \mathcal{A} = (\bar{\mathbb{X}}_i, \bar{\pi}_i, \bar{\mathbb{T}}_i, \mathbb{U}, \bar{h}_i, \mathbb{Y})$. According to Definition 9 we have for any $i = 1,2$, $\bar{\mathbb{X}}_i = \mathbb{X}_i \times Q$, $\bar{\pi}_i(dx_i, q) = \pi_i(dx_i) \cdot 1(q = q_0)$, $\bar{h}_i(x_i, q) = h_i(x_i)$ for any $(x_i, q) \in \bar{\mathbb{X}}_i$, and

$$\bar{\mathbb{T}}_i(A_i \times \{q'\}|x_i,q,u_i) = \int_{\tilde{x}_i \in A_i} \mathbf{1}(q' = \mathsf{t}(q, \mathsf{L}(h_i(\tilde{x}_i)))) \cdot \mathbb{T}_i(d\tilde{x}_i|x_i,u_i).$$

In order to prove the theorem we construct the relation $\mathcal{R}^{\mathsf{p}}$, the interface function $\mathcal{U}_v^{\mathsf{p}}$, and the lifted measure $\mathbb{W}_{\mathbb{T}}^{\mathsf{p}}$ based on $\mathcal{R}, \mathcal{U}_v, \mathbb{W}_{\mathbb{T}}$.

1. Define $\mathcal{R}^{\mathsf{p}} \subseteq \bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2$ with $(x_1,q_1)\mathcal{R}^{\mathsf{p}}(x_2,q_2)$ iff $x_1 \mathcal{R} x_2$ and $q_1 = q_2$. Select any $\bar{x}_1 = (x_1, q_1) \in \bar{\mathbb{X}}_2$ and $\bar{x}_2 = (x_2, q_2) \in \bar{\mathbb{X}}_2$. Then
$$\bar{h}_1(\bar{x}_1) = \bar{h}_1(x_1,q_1) = h_1(x_1) \text{ and } \bar{h}_2(\bar{x}_2) = \bar{h}_2(x_2,q_2) = h_2(x_2).$$
Thus
$$(\bar{x}_1,\bar{x}_2) \in \mathcal{R}^{\mathsf{p}} \Rightarrow (x_1,x_2) \in \mathcal{R} \Rightarrow h_1(x_1) = h_2(x_2) \Rightarrow \bar{h}_1(\bar{x}_1) = \bar{h}_2(\bar{x}_2).$$

2. Since $\pi_1 \bar{\mathcal{R}}_\delta \pi_2$, there exists a lifted measure $\mathbb{W}_{\mathsf{Init}}$ such that $\mathbb{W}_{\mathsf{Init}}(\mathcal{R}) \geq 1-\delta$, $\mathbb{W}_{\mathsf{Init}}(A_1 \times \mathbb{X}_2) = \pi_1(A_1)$ for all $A_1 \in \mathcal{B}(\mathbb{X}_1)$, and $\mathbb{W}_{\mathsf{Init}}(\mathbb{X}_1 \times A_2) = \pi_2(A_2)$ for all $A_2 \in \mathcal{B}(\mathbb{X}_2)$. Define the probability space $\left(\bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2, \mathcal{B}(\bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2), \bar{\mathbb{W}}_{\mathsf{Init}}\right)$ with the property that for all $A_1 \in \mathcal{B}(\mathbb{X}_1)$, $A_2 \in \mathcal{B}(\mathbb{X}_2)$, and $q_1, q_2 \in Q$

$$\bar{\mathbb{W}}_{\mathsf{Init}}(\{q_1\} \times A_1 \times \{q_2\} \times A_2) := \mathbb{W}_{\mathsf{Init}}(A_1 \times A_2) 1(q_1 = q_0) 1(q_2 = q_0),$$

where q_0 is the initial state of the automaton $\mathcal{A}$. In words, $\bar{\mathbb{W}}_{\mathsf{Init}}$ assigns probabilities (which are the same as $\mathbb{W}_{\mathsf{Init}}$) to Borel measurable subsets of $\bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2$ if and only if the discrete modes are equal to q_0. For this particular lifted measure, we have

- $\bar{\mathbb{W}}_{\mathsf{Init}}(\{q_1\} \times A_1 \times \bar{\mathbb{X}}_2) = \mathbb{W}_{\mathsf{Init}}(A_1 \times \mathbb{X}_2) 1(q_1 = q_0) = \pi_1(A_1) 1(q_1 = q_0) = \bar{\pi}_1(\{q_1\} \times A_1)$.
- $\bar{\mathbb{W}}_{\mathsf{Init}}(\bar{\mathbb{X}}_1 \times \{q_2\} \times A_2) = \mathbb{W}_{\mathsf{Init}}(\mathbb{X}_1 \times A_2) 1(q_2 = q_0) = \pi_2(A_2) 1(q_2 = q_0) = \bar{\pi}_2(\{q_2\} \times A_2)$.
- $\bar{\mathbb{W}}_{\mathsf{Init}}(\mathcal{R}^{\mathsf{p}}) = \sum_{q_1, q_2} \bar{\mathbb{W}}_{\mathsf{Init}}((x_1, x_2) \in \mathcal{R} \wedge q_1 = q_2) = \mathbb{W}_{\mathsf{Init}}(\mathcal{R}) \geq 1 - \delta$.

Therefore $\bar{\pi}_1 \bar{\mathcal{R}}^{\mathsf{p}}_\delta \bar{\pi}_2$ with lifted measure $\bar{\mathbb{W}}_{\mathsf{Init}}$.

3. We know that for all $(x_1, x_2) \in \mathcal{R}$ and $u_1 \in \mathbb{U}_1$, $\mathbb{T}_1(\cdot|x_1, u_1)\, \bar{\mathcal{R}}\, \mathbb{T}_2(\cdot|x_2, \mathcal{U}_v(u_1, x_1, x_2))$, with lifted probability measure $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1, x_1, x_2)$. Define the new measure $\bar{\mathbb{W}}_{\mathbb{T}}(\,\cdot\,|u_1, \bar{x}_1, \bar{x}_2)$ on $\bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2$ given $\mathbb{U} \times \bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2$ with

$$\begin{aligned}&\bar{\mathbb{W}}_{\mathbb{T}}(\{q_1'\} \times A_1 \times \{q_2'\} \times A_2 \,|\, u_1, q_1, x_1, q_2, x_2) := \int_{\tilde{x}_1 \in A_1} \int_{\tilde{x}_2 \in A_2} \dots \\ &\mathbf{1}(q_1' = \mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1)))) \cdot \mathbf{1}(q_2' = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2)))) \cdot \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times d\tilde{x}_2 | u_1, x_1, x_2).\end{aligned} \tag{27}$$

In words, $\bar{\mathbb{W}}_{\mathbb{T}}$ assigns probabilities to Borel measurable subsets of $\bar{\mathbb{X}}_1 \times \bar{\mathbb{X}}_2$ which are the same probabilities as in $\mathbb{W}_{\mathbb{T}}$ and evolves the discrete mode of the two gMDP according to the automaton $\mathcal{A}$. For this particular lifted measure $\bar{\mathbb{W}}_{\mathbb{T}}$, we have for any $\bar{x}_1 = (x_1, q_1) \in \mathbb{X}_1$ and $\bar{x}_2 = (x_2, q_2) \in \mathbb{X}_2$ with $\bar{x}_1 \mathcal{R}^{\mathsf{p}} \bar{x}_2$ and any $u_1 \in \mathbb{U}_1$:

$$\begin{aligned}\bar{\mathbb{W}}_{\mathbb{T}}(\{q_1'\} \times A_1 \times \bar{\mathbb{X}}_2 \,|\, u_1, q_1, x_1, q_2, x_2) = \sum_{q_2' \in Q} \int_{\tilde{x}_2 \in \mathbb{X}_2} \int_{\tilde{x}_1 \in A_1} \dots \\ = \int_{\tilde{x}_1 \in A_1} \mathbf{1}(q_1' = \mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1)))) \int_{\tilde{x}_2 \in \mathbb{X}_2} \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times d\tilde{x}_2 | u_1, x_1, x_2) \\ = \int_{\tilde{x}_1 \in A_1} \mathbf{1}(q_1' = \mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1)))) \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times \mathbb{X}_2 | u_1, x_1, x_2) \\ = \int_{\tilde{x}_1 \in A_1} \mathbf{1}(q_1' = \mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1)))) \mathbb{T}_1(d\tilde{x}_1 | x_1, u_1) \\ = \bar{\mathbb{T}}_1(A_1 \times \{q_1'\} | x_1, q_1, u_1),\end{aligned}$$

and

$$\bar{\mathbb{W}}_{\mathbb{T}}(\bar{\mathbb{X}}_1 \times \{q_2'\} \times A_2 \,|\, u_1, q_1, x_1, q_2, x_2) = \sum_{q_1' \in Q} \int_{\tilde{x}_1 \in \mathbb{X}_1} \int_{\tilde{x}_2 \in A_2} \dots$$

$$= \int_{\tilde{x}_2 \in A_2} \mathbf{1}(q_2' = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2)))) \int_{\tilde{x}_1 \in \mathbb{X}_1} \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times d\tilde{x}_2 | u_1, x_1, x_2)$$

$$= \int_{\tilde{x}_2 \in A_2} \mathbf{1}(q_2' = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2)))) \mathbb{W}_{\mathbb{T}}(\mathbb{X}_1 \times d\tilde{x}_2 | u_1, x_1, x_2)$$

$$= \int_{\tilde{x}_2 \in A_2} \mathbf{1}(q_2' = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2)))) \mathbb{T}_2(d\tilde{x}_2 | x_2, \mathcal{U}_v(u_1, x_1, x_2))$$

$$= \bar{\mathbb{T}}_2(A_2 \times \{q_2'\} | x_2, q_2, \mathcal{U}_v(u_1, x_1, x_2)).$$

Take any $(x_1, q_1)\mathcal{R}^{\mathsf{p}}(x_2, q_2)$ which implies $q_1 = q_2$ and $x_1 \mathcal{R} x_2$, hence $h_1(x_1) = h_2(x_2)$. If we also assume $(\tilde{x}_1, q_1')\mathcal{R}^{\mathsf{p}}(\tilde{x}_2, q_2')$, then

$$\mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1))) = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2))),$$

and we get

$$\bar{\mathbb{W}}_{\mathbb{T}}(\mathcal{R}^{\mathsf{p}} \,|\, u_1, q_1, x_1, q_2, x_2) = \int_{(\tilde{x}_1, \tilde{x}_2) \in \mathcal{R}} \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times d\tilde{x}_2 | u_1, x_1, x_2)$$

$$\cdot \sum_{q_1', q_2' \in Q} \mathbf{1}(q_1' = \mathsf{t}(q_1, \mathsf{L}(h_1(\tilde{x}_1)))) \cdot \mathbf{1}(q_2' = \mathsf{t}(q_2, \mathsf{L}(h_2(\tilde{x}_2)))).$$

The above sum is equal to one due to $q_1 = q_2$, $(\tilde{x}_1, \tilde{x}_2) \in \mathcal{R}$, and the DFA being deterministic. Then

$$\bar{\mathbb{W}}_{\mathbb{T}}(\mathcal{R}^{\mathsf{p}} \,|\, u_1, q_1, x_1, q_2, x_2) = \int_{\mathcal{R}} \mathbb{W}_{\mathbb{T}}(d\tilde{x}_1 \times d\tilde{x}_2 | u_1, x_1, x_2) \geq 1 - \delta.$$

We have shown that

$$\bar{\mathbb{T}}_1(A_1 \times \{q_1'\} | x_1, q_1, u_1) \bar{\mathcal{R}}_{\delta}^{\mathsf{p}} \bar{\mathbb{T}}_2(A_2 \times \{q_2'\} | x_2, q_2, \mathcal{U}_v(u_1, x_1, x_2))$$

with lifted measure $\bar{\mathbb{W}}_{\mathbb{T}}$ defined in (27) and the same interface function $\mathcal{U}_v$.