Verification of general Markov decision processes by approximate similarity relations and policy refinement

S. Haesaert[1], A. Abate[2], and P.M.J. Van den Hof[1]

[1] Department of Electrical Engineering, Eindhoven University of Technology
[2] Department of Computer Science, University of Oxford

Abstract. In this work we introduce new approximate similarity relations that are shown to be key for policy (or control) synthesis over general Markov decision processes. The models of interest are discrete-time Markov decision processes, endowed with uncountably-infinite state spaces and metric output (or observation) spaces. The new relations, underpinned by the use of metrics, allow in particular for a useful trade-off between deviations over probability distributions on states, and distances between model outputs. We show that the new probabilistic similarity relationscan be effectively employed over general Markov decision processes for verification purposes, and specifically for control refinement from abstract models.

1 Introduction

The formal verification of computer systems allows for the quantification of their properties and for their correct functioning. Whilst verification has classically focused on finite-state models, with the ever more ubiquitous embedding of digital components into physical systems richer models are needed, and correct functioning can only be expressed over the combined behaviour of both a digital computer and its surrounding physical system. It is in particular of interest to synthesise the part of the computer software that controls or interacts with the physical system automatically, with low likelihood of malfunctioning. Quite importantly, when computers interact with physical systems such as biological processes, power networks, and smart-grids, stochastic models are key.

Systems with uncertainty and non-determinism can be naturally modelled as Markov decision processes (MDP). In this work, we focus on general Markov decision processes (gMDP) with uncountable state spaces as well as metric output spaces. The characterisation of properties or the synthesis of policies over such processes can in general not be attained analytically [3], so an alternative is the approximation of the original (concrete) models by simpler (abstract) models that are prone to be analysed or algorithmically verified [11], such as finite-state MDP [12]. Clearly, it is then key to provide formal guarantees on this approximation step.

In this work we develop a new notion of approximate similarity relation to assist in the computationally efficient controller synthesis of gMDP. The use

of similarity relations on *finite-state* probabilistic models has been broadly investigated, either via exact notions of probabilistic simulation and bisimulation relations [16,20], or via approximate notions [9,10]. On the other hand, similar notions over general, uncountable-state spaces have been only recently studied: available relations either hinge on stability requirements on model outputs [15,23] (established via martingale theory or contractivity analysis), or alternatively enforce structural abstractions of a model [8] by exploiting continuity conditions on its probability laws [1,2].

In this work, we want to quantify properties with a certified precision *both* in the deviation of the probability laws for finite-time events (as in the classical notion of probabilistic bisimulation) and of the output trajectories (as studied for dynamical models). To this end, we generalise the exact probabilistic simulation and bisimulation relations to allow for errors in the probability laws *and* deviations over the output space (Sec. 4). A case study on smart buildings (Sec. 5) is used to evaluate this new approximate similarity relations, which are specifically tailored to perform control synthesis. The new approximate similarity relation generalises notions of probabilistic simulation relations [16,20], and their approximate versions [9,10]. Key to this work, we further show that a control strategy for a gMDP can be obtained as a refinement of a strategy synthesised for an abstract model, at the expense of bounded deviations in transition probabilities and outputs as defined by their similarity relation.

In view of space, details on measurability properties and precise derivations of proofs of the statements are relegated to an extended version [ARXIV], which also contains a more detailed comparison with literature.

2 Verification of general Markov decision processes

2.1 Preliminaries and notations

For two sets A and B a relation $\mathcal{R} \subset A \times B$ is a subset of their Cartesian product that relates elements $x \in A$ with elements $y \in B$, denoted as $x\mathcal{R}y$. We use the following notation for the mappings $\mathcal{R}(\tilde{A}) := \{y : x\mathcal{R}y,\ x \in \tilde{A}\}$ and $\mathcal{R}^{-1}(\tilde{B}) := \{x : x\mathcal{R}y,\ y \in \tilde{B}\}$ for $\tilde{A} \subseteq A$ and $\tilde{B} \subseteq B$. A relation over a set defines a preorder if it is reflexive, $\forall x \in A : x\mathcal{R}x$; and transitive, $\forall x, y, z \in A$: if $x\mathcal{R}y$ and $y\mathcal{R}z$ then $x\mathcal{R}z$. A relation $\mathcal{R} \subseteq A \times A$ is an equivalence relation if it is reflexive, transitive and symmetric, $\forall x, y \in A$: if $x\mathcal{R}y$ then $y\mathcal{R}x$.

A measurable space is a pair $(\mathbb{X}, \mathcal{F})$ with sample space $\mathbb{X}$ and σ-algebra $\mathcal{F}$ defined over $\mathbb{X}$, which is equipped with a topology. As a specific instance of $\mathcal{F}$ consider the Borel measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. In this work, we restrict our attention to Polish spaces and generally consider the Borel σ-field [6]. Recall that a Polish space is a separable and completely metrisable topological space. A simple example of such a space is the real line.

A probability measure $\mathbb{P}(\cdot)$ for $(\mathbb{X}, \mathcal{F})$ is a non-negative map, $\mathbb{P}(\cdot) : \mathcal{F} \to [0, 1]$ such that $\mathbb{P}(\mathbb{X}) = 1$ and such that for all countable collections $\{A_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{F}$, it holds that $\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$. Together with

the measurable space, such a probability measure $\mathbb{P}$ defines the probability space, which is denoted as $(\mathbb{X}, \mathcal{F}, \mathbb{P})$ and has realisations $x \sim \mathbb{P}$. Let us further denote the set of all probability measures for a given measurable pair $(\mathbb{X}, \mathcal{F})$ as $\mathcal{P}(\mathbb{X}, \mathcal{F})$. For a probability space[i] $(\mathbb{X}, \mathcal{F}_{\mathbb{X}}, \mathbb{P})$ and a measurable space $(\mathbb{Y}, \mathcal{F}_{\mathbb{Y}})$, a $(\mathbb{Y}, \mathcal{F}_{\mathbb{Y}})$-valued *random variable* is a function $y : \mathbb{X} \to \mathbb{Y}$ that is $(\mathcal{F}_{\mathbb{X}}, \mathcal{F}_{\mathbb{Y}})$-measurable, and which induces the probability measure $y_*\mathbb{P}$ in $\mathcal{P}(\mathbb{Y}, \mathcal{F}_{\mathbb{Y}})$. For a given set $\mathbb{X}$ a metric or distance function $\mathbf{d}_{\mathbb{X}}$ is a function $\mathbf{d}_{\mathbb{X}} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_0^+$.

2.2 gMDP models - syntax and semantics

General Markov decision processes are related to control Markov processes [1] and Markov decision processes [5,19], and are formalised as follows.

Definition 1 (Markov decision process (MDP)) *A discrete-time MDP* $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U})$ *is defined over an uncountable state space* $\mathbb{X}$*, and characterised by* $\mathbb{T}$*, a conditional stochastic kernel that assigns to each point* $x \in \mathbb{X}$ *and control* $u \in \mathbb{U}$ *a probability measure* $\mathbb{T}(\cdot \mid x, u)$ *over* $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$*. For any set* $A \in \mathcal{B}(\mathbb{X})$*,* $\mathbb{P}_{x,u}(x(t+1) \in A) = \mathbb{T}(A \,|\, x(t) = x, u)$*, where* $\mathbb{P}_{x,u}$ *denotes the conditional probability* $\mathbb{P}(\cdot \mid x, u)$*. The initial probability distribution is* $\pi : \mathcal{B}(\mathbb{X}) \to [0, 1]$*.*

At every state the state transition depends non-deterministically on the choice of $u \in \mathbb{U}$. When chosen according to a distribution $\mu_u : \mathcal{B}(\mathbb{U}) \to [0, 1]$, we refer to the stochastic control input as μ_u. Moreover the transition kernel is denoted 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 (possibly randomised) $u(0), u(1), \ldots, u(N)$, over a finite time horizon $\{0, 1, \ldots, N\}$, and an initial condition x_0 (sampled from distribution 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 MDP.

Definition 2 (General Markov decision process (gMDP)) *A discrete-time gMDP* $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ *is an MDP combined with a measurable output mapping* $h : \mathbb{X} \to \mathbb{Y}$*.*

The gMDP semantics are directly inherited from those of the MDP. Further, output traces of gMDP are obtained as mappings of MDP paths, namely $\{y(t)\}_{0:N} := y(0), y(1), \ldots, y(N)$, where $y(t) = h\big(x(t)\big)$. Denote the class of all gMDP with the metric output space $\mathbb{Y}$ as $\mathcal{M}_{\mathbb{Y}}$. Note that gMDP can be regarded as a super-class of the known labelled Markov processes (LMP) [8] as elucidated in [2].

Example 1. Consider a stochastic process defined as the solution of the stochastic difference equation

$$\mathbf{M} : x(t+1) = f(x(t), u(t)) + e(t), \qquad y(t) = h(x(t)) \in \mathbb{Y},$$

[i] The index $\mathbb{X}$ in $\mathcal{F}_{\mathbb{X}}$ distinguishes the given σ-algebra on $\mathbb{X}$ from that on $\mathbb{Y}$, which is denoted as $\mathcal{F}_{\mathbb{Y}}$. Whenever possible this index will be dropped.

with variables $x(t), u(t), e(t)$, taking values in $\mathbb{R}^n$, representing the state, control input (external non-determinism), and noise terms respectively. The process is initialised as $x(0) \sim \pi$, and driven by $e(t)$, a white noise sequence with zero-mean normal distributions and variance Σ_e. This stochastic process, defined as a dynamical model with dynamics characterised by the stochastic difference equation above, is a gMDP characterised by a tuple $(\mathbb{R}^n, \pi, \mathbb{T}, \mathbb{R}^n, h, \mathbb{Y})$, where the conditional transition kernel is defined as $\mathbb{T}(\cdot \mid x, u) = \mathcal{N}(\cdot | f(x(t), u(t)), \Sigma_e)$, a normal probability distribution with mean $f(x(t), u(t))$ and variance Σ_e. □

A policy is a selection of control inputs based on the past history of states and actions. We allow controls to be selected via universally measurable maps [5] from the state to the control space, so that time-bounded properties such as safety can be maximised [11]. When the selected controls are only dependent on the current states and thus conditionally independent of history (or memoryless), the policy is referred to as Markov. A Markov policy μ for a gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ is a sequence $\mu = (\mu_1, \mu_2, \mu_3, \ldots)$ of universally measurable maps $\mu_t : \mathbb{X} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$ $t = 0, 1, 2, \ldots$, from the state space $\mathbb{X}$ to the set of controls. Recall that a function $f : \mathbb{Z}_1 \to \mathbb{Z}_2$ is universally measurable if the inverse image of every Borel set is measurable with respect to every complete probability measure on $\mathbb{Z}_1$ that measures all Borel subsets of $\mathbb{Z}_1$.

The execution $\{x(t), t \in [0, N]\}$ initialised by $x_0 \in \mathbb{X}$ and controlled with Markov policy μ is a stochastic process defined on the canonical sample space $\Omega := \mathbb{X}^{N+1}$ endowed with its product topology $\mathcal{B}(\Omega)$. This stochastic process has a probability measure $\mathbb{P}$ uniquely defined by the transition kernel $\mathbb{T}$, policy μ, and initial distribution π [5, Prop. 7.45].

Of interest to us are time-dependent properties such as those expressed as specifications in a temporal logic of choice. This leads to problems where one maximises the probability that a sequence of labelled sets is reached within a time limit and in the right order. One can intuitively understand that in general the optimal policy leading to the maximal probability is not a Markov (memoryless) policy. We introduce the notion of a control strategy, and define it as a broader, memory-dependent version of the Markov policy above. Such a strategy for controlling a gMDP is formulated next as a Markov process that takes the state of the gMDP as input.

Definition 3 (Control strategy) *A control strategy* $\mathbf{C} = (\mathbb{X}_\mathbf{C}, x_{\mathbf{C}0}, \mathbb{X}, \mathbb{T}^t_\mathbf{C}, h^t_\mathbf{C})$ *for a gMDP* $\mathbf{M}$ *with state space* $\mathbb{X}$ *and control space* $\mathbb{U}$ *over the time horizon* $t = 0, 1, 2, \ldots, N$ *is an* inhomogenous Markov process *with state space* $\mathbb{X}_\mathbf{C}$*; an initial state* $x_{\mathbf{C}0}$*; inputs* $x \in \mathbb{X}$*; time-dependent, universally measurable kernels* $\mathbb{T}^t_\mathbf{C}$*,* $t = 0, 1, \ldots, N$*; and with universally measurable output maps* $h^t_\mathbf{C} : \mathbb{X}_\mathbf{C} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$*,* $t = 1, \ldots, N$*, with elements* $\mu \in \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$*.* □

Unlike a Markov policy, the control strategy is in general dependent on the history, as it has an internal state that can be used to remember relevant past events. Note that the first control $u(0)$ is selected by drawing $x_\mathbf{C}(1)$ according to $\mathbb{T}^0_\mathbf{C}(\cdot\,|x_\mathbf{C}(0), x(0))$, where $x_\mathbf{C}(0) = x_{\mathbf{C}0}$, and selecting $u(0)$ from measure $\mu^0_\mathbf{C} = h^0_\mathbf{C}(x_\mathbf{C}(1))$. This is then repeated at every time step, when the controller

selects a control $u(t)$ by updating its internal state $\mathbb{T}^t_{\mathbf{C}}(\cdot\,|x_{\mathbf{C}}(t), x(t))$ and then selecting $u(t)$ according to $\mu^t_{\mathbf{C}} = h^t_{\mathbf{C}}(x_{\mathbf{C}}(t+1))$. The control strategy applied to $\mathbf{M}$ can be both stochastic (it is a realisation of $\mathbb{T}^t_{\mathbf{C}}(\cdot\,|\,x_{\mathbf{C}}(t), x(t))\,$), a function of the initial state $x(0)$, and of time.
The execution $\{(x(t), x_{\mathbf{C}}(t)), t \in [0, N]\}$ of a gMDP $\mathbf{M}$ controlled with strategy $\mathbf{C}$, is defined on the canonical sample space $\Omega := (\mathbb{X} \times \mathbb{X}_{\mathbf{C}})^{N+1}$ endowed with its product topology $\mathcal{B}(\Omega)$. This stochastic process is associated to a unique probability measure $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}$, since the stochastic kernels $\mathbb{T}^t_{\mathbf{C}}$ and $\mathbb{T}$ are Borel measurable and composed via universally measurable policies [5, Prop. 7.45].

2.3 gMDP verification and strategy refinement: The idea

We qualitatively anticipate the main result of this work. We intend to provide a general framework to synthesise control policies over a formal abstraction $\tilde{\mathbf{M}}$ of a concrete complex model $\mathbf{M}$, with the understanding that $\tilde{\mathbf{M}}$ is much simpler to be manipulated (analytically or computationally) than $\mathbf{M}$ is. We define a simulation relation under which a policy $\tilde{\mathbf{C}}$ for the abstract Markov process $\tilde{\mathbf{M}}$ implies the existence of a policy $\mathbf{C}$ for $\mathbf{M}$, so that we can quantify differences in the stochastic transition kernels and in the output trajectories for the two closed-loop models. This allows us to derive bounds on the probability of satisfaction of a specification for $\mathbf{M}\times\mathbf{C}$ from the satisfaction probability of modified specifications for $\tilde{\mathbf{M}} \times \tilde{\mathbf{C}}$. This setup allows dealing with finite-horizon temporal properties, including safety verification as a relevant instance.
The results in this paper are to be used in parallel with optimisation, both for selecting the control refinement and for synthesising a policy on the abstract model. It has been shown in [5] that stochastic optimal control, even for a system on a "basic" state space, can lead to measurability issues: in order to avoid these issues we follow [5,9] and the developed theory for Polish spaces and Borel (or universally) measurable notions. Throughout the paper we will give as clarifying examples Markov processes evolving, as in Example 1, over Euclidean spaces which are a special instances of Polish spaces.

3 Exact (bi-)simulation relations based on lifting

In this section we define probabilistic simulation and bisimulation relations that are, respectively, a preorder and an equivalence relation on $\mathcal{M}_{\mathbb{Y}}$. Before introducing these relations, we first extend Segala's notion [20] of *lifting* to uncountable state spaces, which allows us to equate the transition kernels of two given gMDPs. Thereafter, we leverage liftings to define (bi-)simulation relations over $\mathcal{M}_{\mathbb{Y}}$, which characterise the similarity in the controllable behaviours of the two gMDPs. Subsequently we show that these similarity relations also imply controller refinement, i.e., within the similarity relation a control strategy for a given gMDP can be refined to a controller for another gMDP. In the next section, we show that this exact notion of similarity allows a more general notion of approximate probabilistic simulation. The new notions of similarity relations extend the known exact notions in [16], and the approximate notions of [9,10].

3.1 Lifting for general Markov decision processes

Consider two gMDP $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{M}_{\mathbb{Y}}$ mapping to a common output space $\mathbb{Y}$ with metric $\mathbf{d}_{\mathbb{Y}}$. For $\mathbf{M}_1 = (\mathbb{X}_1, \pi_1, \mathbb{T}_1, \mathbb{U}_1, h_1, \mathbb{Y})$ and $\mathbf{M}_2 = (\mathbb{X}_2, \pi_2, \mathbb{T}_2, \mathbb{U}_2, h_2, \mathbb{Y})$ at given state-action pairs $x_1 \in \mathbb{X}_1, u_1 \in \mathbb{U}_1$ and $x_2 \in \mathbb{X}_2, u_2 \in \mathbb{U}_2$, respectively, 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 $\mathbb{T}_2(\cdot \mid x_2, u_2) \in \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$.

Similar to the coupling of measures in $\mathcal{P}(\mathbb{X}, \mathcal{F})$ [4,17], consider the *coupling* of two arbitrary probability spaces $(\mathbb{X}_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\mathbb{X}_2, \mathcal{F}_2, \mathbb{P}_2)$ (cf. [21]). A probability measure $\mathbb{P}_c$ defined on $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{F})$ *couples* the two spaces if the projections p_1, p_2, with $x_1 = p_1(x_1, x_2)$ and $x_2 = p_2(x_1, x_2)$, define respectively an $(\mathbb{X}_1, \mathcal{F}_1)$- and an $(\mathbb{X}_2, \mathcal{F}_2)$-valued random variable, such that $\mathbb{P}_1 = p_{1*}\mathbb{P}_c$ and $\mathbb{P}_2 = p_{2*}\mathbb{P}_c$. For *finite- or countably infinite-state* stochastic processes a closely-related concept has been introduced in [20] and referred to as *lifting*: the transition probabilities are coupled using a weight function in a way that respects a given relation over the combined state spaces. Rather than using weight functions over a countable or finite domain [20], we introduce lifting as a coupling of measures over Polish spaces.

Since we assume that the state spaces are Polish and have a corresponding Borel σ-field for the given probability spaces $(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1), \mathbb{P}_1)$ and $(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2), \mathbb{P}_2)$ with $\mathbb{P}_1 := \mathbb{T}_1(\cdot \mid x_1, u_1)$ and $\mathbb{P}_2 := \mathbb{T}_2(\cdot \mid x_2, u_2)$, the natural choice for the σ-algebra becomes $\mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2) = \mathcal{B}(\mathbb{X}_1) \otimes \mathcal{B}(\mathbb{X}_2)$[ii] and the question of finding a coupling can be reduced to finding a probability measure in $\mathcal{P}(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2))$.

Definition 4 (Lifting for general state spaces) *Let $\mathbb{X}_1, \mathbb{X}_2$ be two sets with associated measure spaces $(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1))$ and $(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$ and let the Borel measurable set $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$ be a relation. We denote by $\bar{\mathcal{R}} \subseteq \mathcal{P}(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1)) \times \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$ the corresponding lifted relation, so that $\Delta \bar{\mathcal{R}} \Theta$ 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*

1. *for all $X_1 \in \mathcal{B}(\mathbb{X}_1)$: $\mathbb{W}(X_1 \times \mathbb{X}_2) = \Delta(X_1)$;*
2. *for all $X_2 \in \mathcal{B}(\mathbb{X}_2)$: $\mathbb{W}(\mathbb{X}_1 \times X_2) = \Theta(X_2)$;*
3. *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 $s\mathcal{R}t$ with probability 1, or equivalently that $\mathbb{W}(\mathcal{R}) = 1$.*

Remark 1. We have implicitly required that the σ-algebra $\mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2)$ contains not only sets of the form $X_1 \times \mathbb{X}_2$ and $\mathbb{X}_1 \times X_2$, but also specifically the sets that characterise the relation $\mathcal{R}$. Since the spaces $\mathbb{X}_1$ and $\mathbb{X}_2$ have been assumed to be Polish, it holds that every open (closed) set in $\mathbb{X}_1 \times \mathbb{X}_2$ belongs to $\mathcal{B}(\mathbb{X}_1) \otimes \mathcal{B}(\mathbb{X}_2) = \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2)$ [6, Lemma 6.4.2]. As an example also consider the diagonal relation $\mathcal{R}_{diag} := \{(x, x) : x \in \mathbb{X}\}$ over $\mathbb{X} \times \mathbb{X}$, of importance for some examples introduced later. This is a Borel measurable set [6, Theorem 6.5.7]. □

3.2 Exact probabilistic (bi-)simulation relations via lifting

Similar to the alternating notions for probabilistic game structures in [24], we provide a simulation that relates any input chosen for the abstract process with

[ii] $\mathcal{B}(\mathbb{X}_1) \otimes \mathcal{B}(\mathbb{X}_2)$ denotes the product σ-algebra of $\mathcal{B}(\mathbb{X}_1)$ and $\mathcal{B}(\mathbb{X}_2)$.

one for the concrete process. We aim to compare the models behaviour with respect to how they can be controlled, and thus allow for more elaborate handling of the inputs than in the probabilistic simulation relations of [9,10,20], paving the way to controller refinement. We introduce the notion of *interface function* in order to connect the controllable behaviour of the two gMDP:

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

where we require that $\mathcal{U}_v$ is a Borel measurable function. This means that $\mathcal{U}_v$ induces a Borel measurable stochastic kernel, denoted by $\mathcal{U}_v$, over $\mathbb{U}_2$ given $\mathbb{U}_1 \times \mathbb{X}_1 \times \mathbb{X}_2$. The notion of interface function is known in the context of correct-by-design controller synthesis and of hierarchical controller refinement [13,22]. The lifting of the transition kernels for the chosen interface generates a stochastic kernel $\mathbb{W}_{\mathbb{T}}$ conditioned on the inputs $\mathbb{U}_1$ and $\mathbb{X}_1 \times \mathbb{X}_2$. Let us trivially extend the interface function to $\mathcal{U}_v(\mu_1, x_1, x_2) := \int_{\mathbb{U}_1} \mathcal{U}_v(u_1, x_1, x_2)\mu_1(du_1)$.

Definition 5 (Probabilistic simulation) *Consider two gMDP* $\mathbf{M}_i, i = 1, 2$, $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$. *The gMDP* $\mathbf{M}_1$ *is stochastically simulated by* $\mathbf{M}_2$ *if there exists an interface function* $\mathcal{U}_v$ *and relation* $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2 \in \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2)$, *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* $\forall (x_1, x_2) \in \mathcal{R}$:

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

The relationship between the two models is denoted as $\mathbf{M}_1 \preceq \mathbf{M}_2$.

Definition 6 (Probabilistic bisimulation) *Under the same conditions as above,* $\mathbf{M}_1$ *is a probabilistic bisimulation of* $\mathbf{M}_2$ *if there exists a relation* $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$ *such that* $\mathbf{M}_1 \preceq \mathbf{M}_2$ *w.r.t.* $\mathcal{R}$ *and* $\mathbf{M}_2 \preceq \mathbf{M}_1$ *w.r.t. the inverse relation* $\mathcal{R}^{-1} \subseteq \mathbb{X}_2 \times \mathbb{X}_1$. $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *are said to be probabilistically bisimilar, which is denoted* $\mathbf{M}_1 \approx \mathbf{M}_2$.

For every gMDP $\mathbf{M}$: $\mathbf{M} \preceq \mathbf{M}$ and $\mathbf{M} \approx \mathbf{M}$. This can be seen by considering the diagonal relation $\mathcal{R}_{diag} = \{(x_1, x_2) \in \mathbb{X} \times \mathbb{X} \mid x_1 = x_2\}$ and selecting equal inputs for the associated interfaces. The resulting equal transition kernels $\mathbb{T}(\cdot|x, u)\bar{\mathcal{R}}_{diag}\mathbb{T}(\cdot|x, u)$ are lifted by the measure $\mathbb{W}_{\mathbb{T}}(dx_1' \times dx_2'|u, x_1, x_2) = \delta_{x_1'}(dx_2')\mathbb{T}(dx_1'|x_1, u)$ where δ denotes the Dirac distribution.

Example 2 (Lifting for diagonal relations). Consider the specific case of the gMDP ($\mathbf{M}_1$) introduced in Ex. 1, and a slight variation of it ($\mathbf{M}_2$), both given as stochastic dynamic processes as

$$\begin{aligned} \mathbf{M}_1 &: x(t+1) = ax(t) + bu(t) + e(t) \in \mathbb{R}, && y(t) = h(x(t)) \in \mathbb{R}, \\ \mathbf{M}_2 &: x(t+1) = ax(t) + bu(t) + \tilde{e}(t) + \tilde{u}(t) \in \mathbb{R}, && y(t) = h(x(t)) \in \mathbb{R}, \end{aligned}$$

with variables $x(t), x(t+1), u(t), \tilde{u}(t), e(t), \tilde{e}(t)$ and constants a, b taking values in $\mathbb{R}$, and with dynamics initialised with the same probability distribution at $t = 0$ and driven by white noise sequences $e(t), \tilde{e}(t)$, both with zero-mean normal distributions and with variance equal to 1 and 1.25, respectively. $\mathbf{M}_1 \preceq \mathbf{M}_2$. For every action u_1 chosen for $\mathbf{M}_1$, select the control input pair $(u_2, \tilde{u}_2) \in \mathbb{U}_2 = \mathbb{R}^2$ as $u_2 = u_1$, and $\tilde{u}_2$ according to the zero-mean normal distribution with variance 0.25, then the associated *interface* is $\mathcal{U}_v(\,\cdot\,|u_1, x_1, x_2) = \delta_{u_1}(du_2)\mathcal{N}(d\tilde{u}_2|0, 0.25)$. For this interface the stochastic dynamics of the two processes are equal, and can be lifted with $\mathcal{R}_{diag}$, namely $\mathbb{T}_1(\cdot|x, u)\bar{\mathcal{R}}_{diag}\mathbb{T}_2(\cdot|x, \mathcal{U}_v)$. □

Remark 2. Over $\mathcal{M}_{\mathbb{Y}}$, the class of gMDP with a shared output space, the relation $\preceq$ is a preorder, as it is reflexive (see Ex. 2) and transitive (see Cor. 6). Moreover $\approx$ is an equivalence relation as it is also symmetric (Cor. 6). □

3.3 Controller refinement via probabilistic simulation relations

The ideas underlying the controller refinement are first discussed, after which it is shown that the refined controller induces a strategy as per Def. 3. Finally the equivalence of properties defined over the closed-loop gMDPs is shown.

Consider two gMDP $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}, h_i, \mathbb{Y})$ $i = 1, 2$ with $\mathbf{M}_1 \preceq \mathbf{M}_2$. Given the entities $\mathcal{U}_v$ and $\mathbb{W}_{\mathbb{T}}$ associated to $\mathbf{M}_1 \preceq \mathbf{M}_2$, the distribution of the next state x_2' of $\mathbf{M}_2$ is given as $\mathbb{T}_2(\cdot \mid x_2, \mathcal{U}_v(u_1, x_1, x_2))$, and is equivalently defined via the lifted measure as the marginal of $\mathbb{W}_{\mathbb{T}}(\cdot|u_1, x_1, x_2)$ on $\mathbb{X}_2$. Therefore, the distribution of the combined next state (x_1', x_2'), defined as $\mathbb{W}_{\mathbb{T}}(\,\cdot\,|u_1, x_1, x_2)$, can be expressed as

$$\mathbb{W}_{\mathbb{T}}(dx_1' \times dx_2'|u_1, x_1, x_2) = \mathbb{W}_{\mathbb{T}}(dx_1'|x_2', u_1, x_1, x_2)\mathbb{T}_2(dx_2'|x_2, \mathcal{U}_v(u_1, x_1, x_2)),$$

where $\mathbb{W}_{\mathbb{T}}(dx_1'|x_2', u_1, x_1, x_2)$ is referred to as the conditional probability given x_2' (c.f. [7, Corollary 3.1.2]). Similarly, the conditional measure for the initialisation $\mathbb{W}_\pi$ is denoted as $\mathbb{W}_\pi(dx_1(0) \times dx_2(0)) = \mathbb{W}_\pi(dx_1(0)|x_2(0))\pi_2(dx_2(0))$.

Now suppose that we have a control strategy for $\mathbf{M}_1$, referred to as $\mathbf{C}_1$, and we want to construct the refined control strategy $\mathbf{C}_2$ for $\mathbf{M}_2$, which is such that events defined over the output space have equal probability. This refinement procedure follows directly from the interface and the conditional probability distributions, and is described in Algorithm 1. The above execution algorithm is separated into the refined control strategy $\mathbf{C}_2$ and its gMDP $\mathbf{M}_2$. $\mathbf{C}_2$ is composed of $\mathbf{C}_1$, the stochastic kernel $\mathbb{W}_{\mathbb{T}}$, and the interface $\mathcal{U}_v$, and it remembers the previous state of $\mathbf{M}_2$.

Theorem 1 (Refined control strategy) *Let gMDP* $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *be related as* $\mathbf{M}_1 \preceq \mathbf{M}_2$*, and consider the control strategy* $\mathbf{C}_1 = (\mathbb{X}_{\mathbf{C}_1}, x_{\mathbf{C}_1 0}, \mathbb{X}_1, \mathbb{T}^t_{\mathbf{C}_1}, h^t_{\mathbf{C}_1})$ *for* $\mathbf{M}_1$ *as given. Then there exists at least one refined control strategy* $\mathbf{C}_2 = (\mathbb{X}_{\mathbf{C}_2}, x_{\mathbf{C}_2 0}, \mathbb{X}_2, \mathbb{T}^t_{\mathbf{C}_2}, h^t_{\mathbf{C}_2})$ *as defined in Def. 3, with*

- *state space* $\mathbb{X}_{\mathbf{C}_2} := \mathbb{X}_{\mathbf{C}_1} \times \mathbb{X}_1 \times \mathbb{X}_2$*, with elements* $x_{\mathbf{C}_2} = (x_{\mathbf{C}_1}, x_1, x_2)$*;*
- *initial state* $x_{\mathbf{C}_2 0} := (x_{\mathbf{C}_1 0}, 0, 0)$*;*

Algorithm 1: Refinement of control strategy $\mathbf{C}_1$ as $\mathbf{C}_2$

Given *the interface function $\mathcal{U}_v$, and the (conditional) stochastic kernels $\mathbb{W}_{\mathbb{T}}(dx_1'|x_2', u_1, x_1, x_2)$ and $\mathbb{W}_{\pi}(dx_1(0)|x_2(0))$.*

Initialise *by drawing*

- *the initial state $x_2(0)$ from π_2, and*
- *the initial state $x_1(0)$ from $\mathbb{W}_{\pi}(\cdot \mid x_2(0))$.*

Run *starting at $t = 0$,*

1. *given $x_1(t)$, select $u_1(t)$ according $\mathbf{C}_1$,*
2. *choose randomised input $\mu_{2t} = \mathcal{U}_v(u_1(t), x_1(t), x_2(t))$, draw $x_2(t+1)$ from $\mathbb{T}_2(\cdot \mid x_2(t), \mu_{2t})$,*
3. *draw $x_1(t+1)$ from $\mathbb{W}_{\mathbb{T}}(\cdot\,|x_2(t+1), u_1(t), x_1(t), x_2(t))$,*
4. *set $t := t+1$, return.*

- *input variable $x_2 \in \mathbb{X}_2$, namely the state variable of $\mathbf{M}_2$;*
- *time-dependent stochastic kernels $\mathbb{T}^t_{\mathbf{C}_2}$, defined as*

$$\begin{aligned}\mathbb{T}^0_{\mathbf{C}_2}(dx_{\mathbf{C}_2}|x_{\mathbf{C}_2 0}, x_2(0)) &:= \mathbb{T}^0_{\mathbf{C}_1}(dx_{\mathbf{C}_1}|x_{\mathbf{C}_1 0}, x_1)\mathbb{W}_{\pi}(dx_1|x_2)\delta_{x_2(0)}(dx_2) \text{ and}\\ \mathbb{T}^t_{\mathbf{C}_2}(dx'_{\mathbf{C}_2}|x_{\mathbf{C}_2}(t), x_2(t)) &:= \mathbb{T}^t_{\mathbf{C}_1}(dx'_{\mathbf{C}_1}|x_{\mathbf{C}_1}, x_1')\\ &\qquad \mathbb{W}_{\mathbb{T}}(dx_1'|x_2', h^t_{\mathbf{C}_1}(x_{\mathbf{C}_1}), x_2, x_1)\delta_{x_2(t)}(dx_2') \text{ for } t \in [1, N];\end{aligned}$$

- *measurable output maps $h^t_{\mathbf{C}_2}(x_{\mathbf{C}_1}, \tilde{x}_1, x_2) := \mathcal{U}_v(h^t_{\mathbf{C}_1}(x_{\mathbf{C}_1}), x_1, x_2)$.* □

Both the time-dependent stochastic kernels $\mathbb{T}^t_{\mathbf{C}_2}$ and the output maps $h^t_{\mathbf{C}_2}$, for $t \in [0, N]$, are universally measurable, since Borel measurable maps are universally measurable and the latter are closed under composition [5, Chapter 7].

Since, by the above construction of $\mathbf{C}_2$, traces in the output spaces of the closed loop systems $\mathbf{C}_1 \times \mathbf{M}_1$ and $\mathbf{C}_2 \times \mathbf{M}_2$ have equal distribution, it follows that measurable events have equal probability, as stated next.

Theorem 2. *If $\mathbf{M}_1 \preceq \mathbf{M}_2$, then for all control strategies $\mathbf{C}_1$ there exists a control strategy $\mathbf{C}_2$ such that, for all measurable events $A \in \mathcal{B}\left(\mathbb{Y}^{N+1}\right)$,*

$$\mathbb{P}_{\mathbf{C}_1 \times \mathbf{M}_1}\left(\{y_1(t)\}_{0:N} \in A\right) = \mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}\left(\{y_2(t)\}_{0:N} \in A\right),$$

with respective output traces $\{y_1(t)\}_{0:N}$ and $\{y_2(t)\}_{0:N}$ of $\mathbf{C}_1 \times \mathbf{M}_1$ and $\mathbf{C}_2 \times \mathbf{M}_2$.

4 New approximate (bi-)simulation relations via lifting

The requirement on an exact simulation relation between two models is evidently restrictive. This is also shown in the following example of gMDPs.

Example 3 (Models with a shared noise source).
Consider an output space $\mathbb{Y} := \mathbb{R}^d$, with a metric $\mathbf{d}_{\mathbb{Y}}(x, y) := \|x - y\|$ (the Euclidean norm), and two gMDP expressed as noisy dynamic processes:

$$\begin{aligned}\mathbf{M}_1 &: x_1(t+1) = f(x_1(t), u_1(t)) + e_1(t), \qquad y_1(t) = h(x_1(t)),\\ \mathbf{M}_2 &: x_2(t+1) = f(x_2(t), u_2(t)) + e_2(t), \qquad y_2(t) = h(x_2(t)),\end{aligned}$$

where f and h are both globally Lipschitz. Namely, there is an $0 < L < 1$ such that $\|f(x_1, u) - f(x_2, u)\| \leq L\|x_1 - x_2\|$ for all $x_1, x_2 \in \mathbb{R}^n$ and for all

u, and in addition an $0 < H$ such that $\|h(x_1) - h(x_2)\| \leq H\|x_1 - x_2\|$. Suppose the probability distributions of the random variable e_1 and of e_2 can be coupled with distribution $\mathbb{P}_{e_1 \times e_2}$,and that there exists a value $c \in \mathbb{R}$, such that $\mathbb{P}_{e_1 \times e_2}[\|e_1 - e_2\| < c] = 1$. Then for every pair of states $x_1(t)$ and $x_2(t)$ of $\mathbf{M}_1$ and $\mathbf{M}_2$ respectively, the difference between state transitions is bounded as $\|x_1(t+1) - x_2(t+1)\| \leq L\|x_1(t) - x_2(t)\| + c$ with probability 1. Therefore, we know that if $\|x_1(0) - x_2(0)\| \leq \frac{c}{1-L}$, then for all $t \geq 0$, $\|x_1(t) - x_2(t)\| \leq \frac{c}{1-L}$, and $\|y_1(t) - y_2(t)\| \leq \frac{cH}{1-L}$.
Even though the difference in the output of the two models is bounded with probability 1, it is impossible to provide an approximation error using either the method in [15] (hinging on stochastic stability assumptions), nor using (approximate) relations as in [9,10]: with the former approach, for the same input sequence $u(t)$ the output trajectories of $\mathbf{M}_1$ and $\mathbf{M}_2$ have bounded difference, but do not converge to each other; with the latter approach, the relation defined via a normed difference cannot satisfy the required notion of transitivity. □

As mentioned before and highlighted in the previous Ex. 3, we are interested in introducing a new approximate version of the notion of probabilistic simulation relation, which allows for both δ-differences in the stochastic transition kernels, and ϵ-differences in the output trajectories. For the former prerequisite, we relax the requirements on the lifting in Def. 4.

Definition 7 (δ-lifting for general state spaces) *Let $\mathbb{X}_1, \mathbb{X}_2$ be two sets with associated measure 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 (acting on $\Delta\bar{\mathcal{R}}_\delta\Theta$), if there exists a probability space $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2), \mathbb{W})$ satisfying*

1. *for all $X_1 \in \mathcal{B}(\mathbb{X}_1)$: $\mathbb{W}(X_1 \times \mathbb{X}_2) = \Delta(X_1)$;*
2. *for all $X_2 \in \mathcal{B}(\mathbb{X}_2)$: $\mathbb{W}(\mathbb{X}_1 \times X_2) = \Theta(X_2)$;*
3. *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 $s\mathcal{R}t$ with probability at least $1 - \delta$, or equivalently that $\mathbb{W}(\mathcal{R}) \geq 1 - \delta$.*

We leverage Definition 7 to introduce a new approximate similarity relation that encompasses both approximation requirements, obtaining the following ϵ, δ-approximate probabilistic simulation.

Definition 8 (ϵ, δ-approximate probabilistic simulation) *Consider two gMDP $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y}), i = 1, 2$, 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 $\forall (x_1, x_2) \in \mathcal{R}$:*

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

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

Definition 9 (ϵ, δ-approximate probabilistic bisimulation) *Under the same conditions as before* $\mathbf{M}_1$ *is an* ϵ, δ*-probabilistic bisimulation of* $\mathbf{M}_2$ *if there exists a relation* $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$ *such that* $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$ *w.r.t.* $\mathcal{R}$ *and* $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$ *w.r.t.* $\mathcal{R}^{-1} \subset \mathbb{X}_2 \times \mathbb{X}_1$. $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *are said to be* ϵ, δ*-probabilistically bisimilar, denoted as* $\mathbf{M}_1 \approx_\epsilon^\delta \mathbf{M}_2$.

In the next section we use the introduced similarity relations is to quantify the probability of events of a gMDP via its abstraction and to refine controllers.

4.1 Controller refinement via approximate simulation relations

Consider two gMDP $\mathbf{M}_1$ and $\mathbf{M}_2$ for which $\mathbf{M}_1$ is the abstraction of the concrete model $\mathbf{M}_2$. The following result is an approximate version of Theorem 2, and provides the main result of this paper, i.e., approximate equivalence of properties defined over the gMDP $\mathbf{M}_1$ and $\mathbf{M}_2$.

Theorem 3. *If* $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$*, then for all control strategies* $\mathbf{C}_1$ *there exists a control strategy* $\mathbf{C}_2$ *such that for the output traces* $\{y_1(t)\}_{0:N}$ *and* $\{y_2(t)\}_{0:N}$ *of* $\mathbf{C}_1 \times \mathbf{M}_1$ *and* $\mathbf{C}_2 \times \mathbf{M}_2$*, it holds that for all measurable events* $A \subset \mathbb{Y}^{N+1}$

$$\mathbb{P}_{\mathbf{C}_1 \times \mathbf{M}_1}\big(\{y_1(t)\}_{0:N} \in A_{-\epsilon}\big) - \gamma \leq \mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}(\{y_2(t)\}_{0:N} \in A) \leq \mathbb{P}_{\mathbf{C}_1 \times \mathbf{M}_1}(\{y_1(t)\}_{0:N} \in A_\epsilon) + \gamma,$$

with constant $1 - \gamma := (1-\delta)^{N+1}$*, and with the* ϵ*-expansion of* A *defined as*

$$A_\epsilon := \big\{\{y_\epsilon(t)\}_{0:N} | \exists \{y(t)\}_{0:N} \in A : \max_{t \in [0,N]} \mathbf{d}_{\mathbb{Y}}(y_\epsilon(t), y(t)) \leq \epsilon\big\}$$

and similarly the ϵ*-contraction defined as* $A_{-\epsilon} := \{\{y(t)\}_{0:N} | \{\{y(t)\}_{0:N}\}_\epsilon \subset A\}$ *where* $\{\{y(t)\}_{0:N}\}_\epsilon$ *is the point-wise* ϵ*-expansion of* $\{y(t)\}_{0:N}$.

Key to show this result is the existence of a refined control strategy $\mathbf{C}_2$, which we detail next. Given a control strategy $\mathbf{C}_1$ over the time horizon $t \in \{0, \ldots, N\}$, there is a control strategy $\mathbf{C}_2$ that refines $\mathbf{C}_1$ over $\mathbf{M}_2$. The control strategy is conceptually given in Algorithm 2. Whilst the state (x_1, x_2) of $\mathbf{C}_2$ is in $\mathcal{R}$, the control refinement from $\mathbf{C}_1$ follows in the same way as for the exact case of Sec. 3.3. Hence, similar to the control refinement for exact probabilistic simulations, the *basic ingredients* of $\mathbf{C}_2$ are the states x_1 and x_2, whose 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 $\mathbb{W}_{\mathbb{T}}(dx_1' \mid x_2', u_1, x_2, x_1)$, *conditioned* on x_2'.

On the other hand, whenever the state (x_1, x_2) leaves $\mathcal{R}$ the control chosen by strategy $\mathbf{C}_1$ cannot be refined to $\mathbf{M}_2$ and fails. A new control strategy $\mathbf{C}_{rec}$, referred to as *recovery*, can be used to control the residual trajectory of $\mathbf{M}_2$. The choice is of no importance to the result in Theorem 3, as it bounds errors on probabilistic events based on the event that the states stay in the relation.

Theorem 4 (Refined control strategy) *Let gMDP* $\mathbf{M}_1$ *and* $\mathbf{M}_2$*, with* $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$*, and control strategy* $\mathbf{C}_1 = (\mathbb{X}_{\mathbf{C}_1}, x_{\mathbf{C}_1 0}, \mathbb{X}_1, \mathbb{T}^t_{\mathbf{C}_1}, h^t_{\mathbf{C}_1})$ *for* $\mathbf{M}_1$ *be given. Then for every recovery control strategy* $\mathbf{C}_{rec}$*, a refined control strategy* $\mathbf{C}_2 = (\mathbb{X}_{\mathbf{C}_2}, x_{\mathbf{C}_2 0}, \mathbb{X}_2, \mathbb{T}^t_{\mathbf{C}_2}, h^t_{\mathbf{C}_2})$ *is obtained as an* inhomogenous Markov process *with two discrete modes of operation,* {refinement} *and* {recovery}*, based on Algorithm 2.*

By dividing the execution in Algorithm 2 into a control strategy and a gMDP $\mathbf{M}_2$, we again obtain a refined control strategy with tuple $(\mathbb{X}_{\mathbf{C}_2}, x_{\mathbf{C}_2 0}, \mathbb{X}_2, \mathbb{T}^t_{\mathbf{C}_2}, h^t_{\mathbf{C}_2})$.

Algorithm 2: Refinement of $\mathbf{C}_1$ as $\mathbf{C}_2$

Given the interface function $\mathcal{U}_v$, the (conditional) stochastic kernels $\mathbb{W}_{\mathbb{T}}(dx_1'|x_2', u_1, x_1, x_2)$ and $\mathbb{W}_{\pi}(dx_1(0)|x_2(0))$, and the chosen recovery strategy $\mathbf{C}_{rec}$.

Initialise by drawing

- the initial state $x_2(0)$ from π_2, and
- the initial state $x_1(0)$ from $\mathbb{W}_{\pi}(\cdot \mid x_2(0))$.

Run starting at $t = 0$, **while** $t \leq N$

1. **if** $(x_1(t), x_2(t)) \in \mathcal{R}$ **go to** 2. **else skip to** 6.
2. given $x_1(t)$, select $u_1(t)$ from $\mathbf{C}_1$, {refine}
3. choose randomised input $\mu_{2t} = \mathcal{U}_v(u_1(t), x_1(t), x_2(t))$, draw $x_2(t+1)$ from $\mathbb{T}_2(\cdot \mid x_2(t), \mu_{2t})$,
4. draw $x_1(t+1)$ from $\mathbb{W}_{\mathbb{T}}(\cdot \,|x_2(t+1), u_1(t), x_1(t), x_2(t))$,
5. set $t := t + 1$, go to 1.
6. given $x_2(t)$, compute μ_t (from $\mathbf{C}_{rec}$), {recover}
7. draw $x_2(t+1)$ from $\mathbb{T}_2(\cdot \mid x_2(t), \mu_t)$,
8. set $t := t + 1$, go to 6.

4.2 Examples and properties

Example 4 (Models with a shared noise source – continued from above).
Based on the relation $\mathcal{R} := \{(x_1, x_2) : \|x_1 - x_2\| \leq \frac{c}{1-L}\}$ it can be shown that $\mathbf{M}_1 \approx_{\epsilon}^{0} \mathbf{M}_2$ with $\epsilon = \frac{Hc}{1-L}$, since, firstly, it holds that $\mathbf{d}_{\mathbb{Y}}(h(x_1), h(x_2)) \leq \epsilon$ for all $(x_1, x_2) \in \mathcal{R}$, with $\mathbf{d}_{\mathbb{Y}} = \|h(x_1) - h(x_2)\|$. Additionally, for all $(x_1, x_2) \in \mathcal{R}$ and for any input u_1 the selection $u_2 = u_1$ is such that $\mathbb{T}_1(\cdot|x_1, u_1)\bar{\mathcal{R}}_0\mathbb{T}_2(\cdot|x_2, u_1)$, note that $\bar{\mathcal{R}}_0$ is equal to $\bar{\mathcal{R}}$ (the lifted relation from $\mathcal{R}$). The lifted stochastic kernel is $\mathbb{W}_{\mathbb{T}}(dx_1' \times dx_2'|u_1, x_1, x_2) := \int_{\omega} \delta_{f(x_1,u_1)+g_1(\omega)}(dx_1')\delta_{f(x_2,u)+g_2(\omega)}(dx_2')\mathbb{P}_{\omega}(d\omega)$, this stochastic kernel is Borel measurable if $f(x_1, u_1) + g_1(\omega)$ and $f(x_2, u) + g_2(\omega)$ are assumed Borel measurable mappings. Note that the employed identity interface is also Borel measurable. □

Example 5 (Relationship to model with truncated noise).
Consider the stochastic dynamical process $\mathbf{M}_1 : x(t+1) = f(x(t), u(t)) + e(t)$ with output mapping $y(t) = h(x(t))$, operating over the Euclidean state space $\mathbb{R}^n$, and driven by a white noise sequence $e(t) \in \mathbb{R}^n$ with distribution $\mathbb{P}_e$. The output space $y \in \mathbb{Y} \subseteq \mathbb{R}^d$ is endowed with the Euclidean norm $\mathbf{d}_{\mathbb{Y}} = \|\cdot\|$. Select a domain $D \subset \mathbb{R}^n$ so that, at any given time instant t, $e(t) \in D$ with probability $1 - \delta$. Then define a truncated white noise sequence $\tilde{e}(t)$, with distribution $\mathbb{P}_e(\cdot \mid D)$. The resulting model $\mathbf{M}_2$ driven by $\tilde{e}(t)$ is $\mathbf{M}_2 : x(t+1) = f(x(t), u(t)) + \tilde{e}(t)$, with the same output mapping $y(t) = h(x(t))$. We show that $\mathbf{M}_2$ is a $0, \delta$-approximate probabilistic bisimulation of $\mathbf{M}_1$, i.e. $\mathbf{M}_1 \approx_0^{\delta} \mathbf{M}_2$. Select $\mathcal{R} := \{(x_1, x_2) \text{ for } x_1, x_2 \in \mathbb{R}^n | x_1 = x_2\}$, and choose as interface the identity function, i.e., $\mathcal{U}_v(u_1, x_1, x_2) = u_1$. Denote $t_1(e) = f(x_1, u_1) + e$ and $t_2(\tilde{e}) = f(x_2, u_1) + \tilde{e}$, then a lifting measure depending on $x_1, x_2 \in \mathcal{R}$ and u_1, is

$$\mathbb{W}_{\mathbb{T}}(dx_1' \times dx_2'|u_1, x_1, x_2) := \int_{e \in D} \delta_{x_1'}(dx_2')\delta_{t_1(e)}(dx_1')\mathbb{P}_e(de) \tag{1}$$
$$+ \int_{e \in \mathbb{R}^n \setminus D} \delta_{t_1(e)}(dx_1')\mathbb{P}_e(de) \int_{\tilde{e}} \delta_{t_2(\tilde{e})}(dx_2')\mathbb{P}_e(d\tilde{e}|D). \qquad \square$$

Example 6 (Relationship between noiseless and truncated-noise models). Continuing with Ex. 5, consider the model with truncated noise $\mathbf{M}_2$ as defined before. In what sense is $\mathbf{M}_2$ approximated by its noiseless version $\mathbf{M}_3$, namely $\mathbf{M}_3 : x(t+1) = f(x(t), u(t))$, with $y(t) = h(x(t))$? Under requirements on the Lipschitz continuity $\|f(x_1, u) - f(x_2, y)\| \leq L\|x_1 - x_2\|$ $0 < L < 1$, $\|h(x_1) - h(x_2)\| \leq H\|x_1 - x_2\|$, and on the boundedness of D and of $c = \max_{d \in D} \|d\|$, Ex. 3 can be leveraged by concluding that $\mathbf{M}_2 \approx_{\epsilon}^{0} \mathbf{M}_3$, with $\epsilon = \frac{Hc}{1-L}$. [iii] □

In the Ex. 5 and 6 we have that $\mathbf{M}_1$ is approximated by $\mathbf{M}_2$, which is subsequently approximated by $\mathbf{M}_3$. The following theorem and corollary attains a quantitative answer on the question whether $\mathbf{M}_1$ is approximated by $\mathbf{M}_3$.

Theorem 5 (Transitivity of $\preceq_{\epsilon}^{\delta}$). *Consider three gMDP $\mathbf{M}_i$, $i = 1, 2, 3$, defined by tuples $(\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, with shared output space.*

$$\textit{If } \mathbf{M}_1 \preceq_{\epsilon_a}^{\delta_a} \mathbf{M}_2 \textit{ and } \mathbf{M}_2 \preceq_{\epsilon_b}^{\delta_b} \mathbf{M}_3, \textit{ then } \mathbf{M}_1 \preceq_{\epsilon_a+\epsilon_b}^{\delta_a+\delta_b} \mathbf{M}_3.$$

Next, as a corollary of this theorem, we discuss further transitivity properties for simulation and bisimulation relations.

Corollary 6 (Transitivity properties) *Following Theorem 5, it holds that*

- *if* $\mathbf{M}_1 \approx_{\epsilon_a}^{\delta_a} \mathbf{M}_2$ *and* $\mathbf{M}_2 \approx_{\epsilon_b}^{\delta_b} \mathbf{M}_3$, *then* $\mathbf{M}_1 \approx_{\epsilon_a+\epsilon_b}^{\delta_a+\delta_b} \mathbf{M}_3$, *and*
- *if* $\mathbf{M}_1 \preceq \mathbf{M}_2$ *and* $\mathbf{M}_2 \preceq \mathbf{M}_3$, *then* $\mathbf{M}_1 \preceq \mathbf{M}_3$, *and*
- *if* $\mathbf{M}_1 \approx \mathbf{M}_2$ *and* $\mathbf{M}_2 \approx \mathbf{M}_3$, *then* $\mathbf{M}_1 \approx \mathbf{M}_3$.

Example 7 (Combination of Examples 5 and 6 via Corollary 6). For the models in Examples 5 and 6 we can conclude that $\mathbf{M}_1 \approx_{\epsilon}^{\delta} \mathbf{M}_3$. This means that a stochastic system as in $\mathbf{M}_1$ in Ex. 5 can be approximated via its deterministic counterpart, and that the approximation error can be expressed via the probability (i.e. amount of truncation cf. Ex. 5) and the output error (i.e. Ex. 6). This allows for explicit trading off between output deviation and deviation in probability. □

5 Case study: Energy management in smart buildings

We are interested in developing advanced solutions for the energy management of smart buildings. We consider a simple building that is divided in two connected zones, each with a radiator affecting the heat exchange in that zone by controlling the water temperature in a boiler. A model of the temperature dynamics in an office building with two zones to heat [14] assumes that the temperature fluctuations in the two zones and the ambient temperature dynamics can be modelled via $\mathbf{M}$ as a Gaussian process:

$$\mathbf{M} : x(t+1) = Ax(t) + Bu(t) + Fe(t), \qquad y(t) = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right] x(t), \quad (2)$$

[iii] Alternatively, if $\mathbf{M}_2$ with non-deterministic input $\tilde{e} \in D$ is an ϵ_a- alternating bisimulation [22] of $\mathbf{M}_3$ then $\mathbf{M}_2 \approx_{\epsilon_a}^{0} \mathbf{M}_3$.

with stable dynamics characterised by matrices

$$A = \begin{bmatrix} 0.8725 & 0.0625 & 0.0375 \\ 0.0625 & 0.8775 & 0.0250 \\ 0 & 0 & 0.9900 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0650 & 0 \\ 0 & 0.60 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0.05 & -0.02 & 0 \\ -0.02 & 0.05 & 0 \\ 0 & 0 & 0.1 \end{bmatrix},$$

where $x_{1,2}(t)$ are the temperatures in zone 1 and 2, respectively; $x_3(t)$ is the deviation of the ambient temperature from its mean; and $u(t) \in \mathbb{R}^2$ is the control input. The state variables are initiated as $x(0) = [16\ 14\ \text{-}5]^T$. The disturbance $e(t)$ is a sequence of independent and identically distributed standard normal distributions, for all $t \in \mathbb{R}^+$. This stochastic process can be written as a gMDP as detailed in Example 1. For the model abstraction, we select the controllable dynamics of the mean of the state variables, and consequently omit the ambient temperature:

$$\tilde{\mathbf{M}} : \begin{cases} \tilde{x}(t+1) = \tilde{A}\tilde{x}(t) + \tilde{B}\tilde{u}(t) \in \mathbb{R}^2, \text{ with } \tilde{A} := \left[\begin{smallmatrix} 0.8725 & 0.0625 \\ 0.0625 & 0.8775 \end{smallmatrix}\right], \\ \tilde{y}(t) \quad\quad = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]\tilde{x}(t), \qquad\qquad\qquad\quad \tilde{B} := \left[\begin{smallmatrix} 0.0650 & 0 \\ 0 & 0.60 \end{smallmatrix}\right]. \end{cases} \tag{3}$$

We then obtain that, as intuitive, $\tilde{\mathbf{M}} \preceq_\epsilon^\delta \mathbf{M}$.

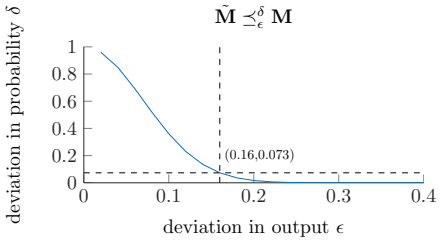

Fig. 1: Figure of trade-off between the output error ϵ and the probability error δ for the δ,ϵ-approximate probabilistic simulation $\tilde{\mathbf{M}} \preceq_\epsilon^\delta \mathbf{M}$. We have selected the pair $(\epsilon, \delta) = (0.16, 0.073)$ as an ideal trade-off.

In order to compute specific values of ϵ and δ, we select the relation $\mathcal{R} := \{(\tilde{x}, x) \in \mathbb{R}^2 \times \mathbb{R}^3 \mid \sqrt{(\tilde{x}_1 - x_1)^2 + (\tilde{x}_2 - x_2)^2} \leq \epsilon\}$ and the interface function $\mathcal{U}_v(\tilde{u}, \tilde{x}, x) = \tilde{u} + \tilde{B}^{-1}(\tilde{A}\tilde{x} - \bar{A}x)$, with $\bar{A} = \left[\begin{smallmatrix} 0.8725 & 0.0625 & 0.0375 \\ 0.0625 & 0.8775 & 0.0250 \end{smallmatrix}\right]$. A stochastic kernel $\mathbb{W}_\mathbb{T}$ for the lifting is $\mathbb{W}_\mathbb{T}(d\tilde{x}' \times dx' \mid \tilde{u}, \tilde{x}, x) = \int_e \delta_{\tilde{f}}(d\tilde{x}')\, \delta_{f(e)}(dx')\mathcal{N}(de \,|\, 0, I)$, with $\tilde{f} = \tilde{A}\tilde{x} + \tilde{B}\tilde{u}$ and $f(e) = Ax + B\mathcal{U}_v(\tilde{u}, \tilde{x}, x) + Fe$. The lower bound on $\mathbb{W}_\mathbb{T}(\mathcal{R} \mid \tilde{u}, \tilde{x}, x) \leq 1 - \delta$ has been computed and traded off against the output deviation in Fig. 1.

We are interested in the goal, expressed for the model $\mathbf{M}$, of increasing the likelihood of reaching the target set $T = [20.5,\ 21]^2$ and staying there thereafter. For the abstract model we have developed a strategy, as in[14], satisfying by construction the property expressed in LTL-like notation with the formula $\varphi = \Diamond\Box T$ and shrunken to $\varphi_{-\epsilon}$ (as per Theorem 3). This strategy is synthesised as a correct-by-construction controller using PESSOA [18], where the discrete-time dynamics are further discretised over state and action spaces: we have selected a state quantisation of 0.05 over the range $[15, 25]^2$, and an input quantisation of 0.05 over the set $[10, 30]^2$. It can be observed that the controller regulates the abstract model $\tilde{\mathbf{M}}$ to eventually remain within the target region, as shown in Fig. 2. We now want to verify that indeed, when refined to the concrete stochastic model, this strategy implies the reaching and staying in the safe set up to some probabilistic error. The refined strategy is obtained from this control strategy

as discussed in Section 4.1, and recovers from exits out of the relation $\mathcal{R}$ by resetting the abstract states in the relation. A simulation study is given in Fig.2: as predicted, the behaviour of the controlled concrete model $\mathbf{M}$ stays close to that of $\tilde{\mathbf{M}}$. Over a time horizon of 200 steps the output error exceeds the level $\epsilon = 0.16$ only a few (four) times. Indeed, the probability that the concrete state leaves the relation with the abstract model ($\leq \delta$, with $\delta = 0.073$) leads, over N time steps, to a bound on the probability that it does not satisfy the LTL property: Theorem 3 ensures that this probability is provably less than $1-(1-\delta)^N \approx N\delta$. In practice, whenever state exits the relation, then the controller recovers by resetting the state of the abstract model and re-applying the strategy again, and thanks to the ϵ-contraction $\varphi_{-\epsilon}$ of the concrete specification, $\mathbf{M}$ will abide by φ with a high confidence.

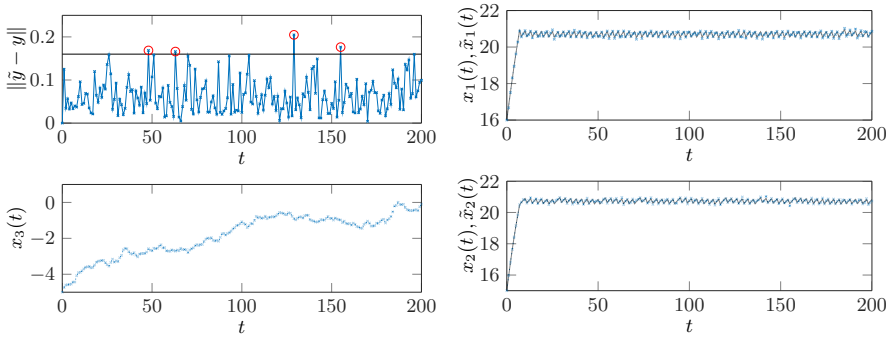

Fig. 2: Refined control for deterministic model applied to $\mathbf{M}$. The figure (above left) evaluates the accuracy of the approximation, and gives with red circles the instances in which the relation is left. The plot (below right) gives the ambient temperature. The plots on the right give the temperature inside the rooms. The (very small) blue crosses give the actual temperature in the rooms (x_1, x_2) and cover the deterministic simulation of $(\tilde{x}_1, \tilde{x}_2)$ drawn in black.

6 Conclusions

In this work we have discussed new approximate similarity relations for general control Markov processes, and shown that they can be effectively employed for abstraction-based verification and controller refinement. The new relations in particular allow for a useful trade-off over deviations between probability distributions on the states and distances between model outputs.

Alongside practical applications of the developed notions, current research efforts focus on further generalisation of Theorem 3 to specific quantitative properties expressed via temporal logics. We are moreover interested in expanding on the properties of the similarity relations.

References

1. A. Abate. Approximation metrics based on probabilistic bisimulations for general state-space Markov processes: a survey. *Electronic Notes in Theoretical Computer Science*, 297:3–25, 2013.

2. A. Abate, M. Kwiatkowska, G. Norman, and D. Parker. Probabilistic model checking of labelled Markov processes via finite approximate bisimulations. In *Horizons of the Mind – P. Panangaden Festschrift*, pages 40–58. Springer Verlag, 2014.
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. A. Abate, F. Redig, and I. Tkachev. On the effect of perturbation of conditional probabilities in total variation. *Statistics & Probability Letters*, 2014.
5. D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal control : The discrete time case*. Athena Scientific, 1996.
6. V. I. Bogachev. *Measure theory*. Springer Science & Business Media, 2007.
7. V. S. Borkar. *Probability theory: an advanced course*. Springer Science & Business Media, 2012.
8. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.
9. J. Desharnais, F. Laviolette, and M. Tracol. Approximate analysis of probabilistic processes: Logic, simulation and games. *QEST*, pages 264–273, Sept. 2008.
10. 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.
11. 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.
12. S. Esmaeil Zadeh Soudjani, C. Gevaerts, and A. Abate. FAUST ^{2}: Formal Abstractions of Uncountable-STate STochastic Processes. In *TACAS*, LNCS, pages 272–286. Springer Berlin Heidelberg, 2015.
13. A. Girard and G. J. Pappas. Hierarchical control system design using approximate simulation. *Automatica*, 45(2):566–571, 2009.
14. S. Haesaert, A. Abate, and P. M. J. Van den Hof. Correct-by-design output feedback of LTI systems. In *Conference on Decision and Control*, pages 6159–6164, 2015.
15. A. A. Julius and G. J. Pappas. Approximations of stochastic hybrid systems. *IEEE Trans. on Automatic Control*, 2009.
16. K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.
17. T. Lindvall. *Lectures on the coupling method*. Courier Corporation, 2002.
18. M. Mazo Jr, A. Davitian, and P. Tabuada. Pessoa: A tool for embedded controller synthesis. In *Computer Aided Verification*, pages 566–569. Springer, 2010.
19. S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., 1993.
20. R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, Massachusetts Institute of Technology, 1995.
21. H. J. Skala. The existence of probability measures with given marginals. *Ann. Probab.*, 21:136–142, 1993.
22. P. Tabuada. *Verification and control of hybrid systems*. Springer US, 2009.
23. M. Zamani, P. M. Esfahani, R. Majumdar, A. Abate, and J. Lygeros. Symbolic control of stochastic systems via approximately bisimilar finite abstractions. *IEEE Transactions on Automatic Control*, 59(12):3135–3150, 2014.
24. C. Zhang and J. Pang. On probabilistic alternating simulations. In *Theoretical Computer Science*, volume 323 of *IFIP Advances in Information and Communication Technology*, pages 71–85. Springer Berlin Heidelberg, 2010.