Stochastic Best-Effort Strategies for Borel Goals

Benjamin Aminof*, Giuseppe De Giacomo*†, Sasha Rubin‡ and Florian Zuleger§
*University of Rome "La Sapienza", Italy
†University of Oxford, United Kingdom
†University of Sydney, Australia
‡TU Vienna, Austria
{benjamin.aminof, florian.zuleger}@tuwien.ac.at,
degiacomo@diag.uniroma1.it, giuseppe.degiacomo@cs.ox.ac.uk, sasha.rubin@sydney.edu.au

***Abstract*—We study reactive systems with Borel goals operating in a possibly non-Markovian stochastic environment. Moreover, the specific environment is not known, only its support is, i.e., at each step one knows which transitions are possible and which are impossible, but the probability distribution amongst the possible transitions is unknown. We consider system strategies that are maximal in the dominance order, i.e., no other strategy achieves the goal with at least the same probability in all environments, and with a higher probability in some environment. We call such strategies "stochastic best-effort". We prove the very general result that stochastic best-effort strategies exist for any Borel goal. We do this by providing local characterizations in terms of a three-valued abstraction of the probability of achieving the goal at a history. The correctness of the characterization is shown using a version of the Lebesgue Density Theorem from geometric measure theory. On the more practical side, we consider goals given in linear temporal logic. We establish the computational complexity of synthesizing a stochastic best-effort strategy, and show that it is not harder than synthesizing an optimal strategy in a domain with fixed known probabilities.**

I. Introduction

Synthesis is the problem of producing a reactive system, i.e., one that operates in a dynamic environment, from a temporal specification [1]. This is often modeled by having, at each time step, the system take an action and the environment respond by changing the state. In the stochastic setting, the environment's choice at each time step is governed by some probability distribution over states. When quality stochastic information can be obtained, and the probability distribution for a given state and action does not change over time, an effective approach is to model the environment as a Markov Decision Process and to use algorithms that find optimal strategies [2], [3]. However, what if very little is known about the exact probabilities? What if the probabilities change over time? What sort of system strategies are appropriate?

We study the fundamental general case that environments can be modeled by a set $\mathfrak{D}$ of finite-state non-Markovian stochastic domains that share the same support. Intuitively, the system is ignorant about which particular domain in $\mathfrak{D}$ it is operating in. By "non-Markovian" we mean that the probability distributions may depend on the entire history of the interaction of the system and its environment, not only on the last state and action. By domains sharing the "same support" we mean that for every history it is known which transitions are possible and which are not, but the probability distribution amongst the possible transitions is unknown.

To address the issue of which strategies should the system employ under these conditions, we draw inspiration from the non-stochastic setting. In this setting, one similarly asks "what should the system do when it has little information about the exact environment in which it operates?" The standard answer is to take a worst-case view and employ a strategy that achieves its goal against all environments [4], [5] (in the stochastic setting one possible analogue of the worst-case view is to achieve the goal with optimal probability in each stochastic domain in $\mathfrak{D}$). However, the more ignorant the system is about its environment, the more unreasonable it is to expect that such a strategy exists. A more refined answer draws on a basic principle from Decision Theory: the system should not use a strategy that is dominated by another [6]. In other words, it should not employ a strategy if there is another which does at least as well in all environments and better in some environments. This and related dominance notions have been studied for synthesis and game-solving in the non-stochastic setting, e.g., [7], [8], [9], [10], [11], [12], [13]. Motivated by this, we introduce a dominance order between strategies operating in stochastic domains which says that $\sigma_1 \succcurlyeq \sigma_2$ if in every domain in $\mathfrak{D}$, the probability that the goal is achieved using σ_1 is at least as large as when using σ_2. We call a strategy that is maximal in this order "stochastic best-effort".

We provide local characterizations of the dominance order and of maximality in terms of a three-valued abstraction of the probability of achieving the goal at a history. To reason about this abstraction we make use of properties of the Lebesgue measure, including a version of the Lebesgue Density Theorem for our probability spaces. We then use the characterization of maximality to prove the surprising result that stochastic best-effort strategies exist for every Borel goal. We conclude with addressing the algorithmic problem of finding a stochastic best-effort strategy for the important special case that goals are specified in linear temporal logic, and the domains are bounded and have finitely-generated support (but need not be Markovian). We completely address the computational complexity and show an upper bound (obtained using the characterization of maximality), and a matching lower bound.

We note that we have tried to provide proofs that assume as little as possible about the reader's knowledge. Nonetheless,

979-8-3503-3587-3/23/$31.00

the paper unavoidably assumes that the reader is familiar with basic concepts of topology, measure theory, and the Lebesgue measure.

II. Preliminaries

A. Notation

A set or sequence is *countable* if it is finite or countably infinite. Countable sequences may be written $(x_0, x_1, \cdots)$ or $x_0 x_1 \cdots$. The start of the sequence is x_0. If h is a finite sequence then $last(h)$ denotes its last element, and $|h|$ denotes its length. If h is a prefix of h' we say that h' *extends* h; if, in addition, $h \neq h'$, then we say that h is a *proper prefix* of h' and that h' is a *proper extension* of h. The empty sequence is denoted ε. For a finite set X, let $Dbn(X)$ denote the set of *distributions* over X, i.e., functions $d : X \to [0, 1]$ such that $\sum_{x \in X} d(x) = 1$. The *support* of d is the set $sup(d)$ of elements $x \in X$ such that $d(x) > 0$.

B. States, Actions, Plays, Histories

Let St be a finite set of *states* and Act a finite set of *actions*. Elements of $\Omega := (St \cdot Act)^\omega$ are called *infinite paths*. Elements of $FinPaths := (St \cdot Act)^* \cdot St$ are called *finite paths*. If $\pi = (s_0, a_0, s_1, a_1, \cdots, s_{n-1}, a_{n-1}, s_n)$ then we write $\pi_i = (s_0, a_0, s_1, a_1, \cdots, s_i)$ for its prefix that ends in s_i $(0 \leq i \leq n)$. Similarly, if $\pi = (s_0, a_0, s_1, a_1, \cdots)$ is an infinite path, we write π_i for its prefix that ends in s_i $(0 \leq i)$.

C. Stochastic domains

A *stochastic domain*, or simply *domain*, is a tuple $D = (St, Act, \iota, Pr)$ where

- $\iota \in St$ is the *initial* state, and
- $Pr : FinPaths \times Act \to Dbn(St)$ is the *transition function* that associates to each finite path x, each action $a \in Act$, and each state $s \in St$, the probability $Pr(x, a)(s)$ of a transition from x to s when action a is selected.

In other words, a domain is a transition system with finitely many states, every action available at every state, and transitions are probabilistic and may depend on the entire sequence of states and actions seen so far, and not just on the current state and action. In particular, stochastic domains are *not Markovian*, e.g., the probability of s when doing action a from a state s' may depend on the past states and actions.

The *support function* of D is the function

$$\Delta : \{\varepsilon\} \cup (FinPaths \times Act) \to 2^{St}$$

defined by (i) $\Delta(\epsilon) = \{\iota\}$, and (ii) for $x \in FinPaths$, $s \in \Delta(x, a)$ iff $s \in sup(Pr(x, a))$. An infinite path $\pi = (s_0, a_0, s_1, a_1, \cdots)$ (resp. finite path $\pi = (s_0, a_0, s_1, a_1, \cdots, s_n)$) is called a *play* (resp. *history*) if it satisfies that $\Delta(\epsilon) = \{s_0\}$ and $s_{i+1} \in \Delta(\pi_i, a_i)$ for every i (resp. $i < n$). Note that plays and histories only depend on the support Δ, and not on the particular values of the non-zero probabilities. Intuitively, plays and histories are paths that start at the initial state, and use non-zero probability transitions at every step. We will typically denote a history by the letter h. The set of histories is denoted $Hist$.

A domain D is *Markovian* if $Pr(h, a)$ only depends on $last(h)$ and a, i.e., $last(h) = last(h')$ implies that $Pr(h, a) = Pr(h', a)$. We say that D has *Markovian support* if $\Delta(h, a)$ only depends on $last(h)$ and a, i.e., $last(h) = last(h')$ implies that $\Delta(h, a) = \Delta(h', a)$. In this case, we may write the support function as $\Delta : \{\varepsilon\} \cup (St \times Act) \to 2^{St}$. Observe that every Markovian domain has a Markovian support, but that the converse may or may not be true. In the rest of the paper we assume a fixed support Δ, and will usually not mention it.

D. Strategies, split points, shifts

A *strategy* is a function $\sigma : Hist \to Act$ that assigns an action to every history.[1] A strategy is *finite-state* (aka *finite-memory*) if it can be represented as a finite-state input/output automaton that, on reading $h \in Hist$, outputs the action $\sigma(h)$.

An play $\pi = (s_0, a_0, s_1, a_1, \cdots)$ (resp. history $\pi = (s_0, a_0, s_1, a_1, \cdots, s_n)$) is *consistent* with a strategy σ if $\sigma(\pi_i) = a_i$ for every i (resp. for every $i < n$). In this case we will also say that π is a σ-play (resp. σ-history).

A history h is a *split point* of strategies σ_1, σ_2 if $\sigma_1(h) \neq \sigma_2(h)$ but $\sigma_1(h') = \sigma_2(h')$ for all histories h' that are proper prefixes of h.

For strategies σ_1, σ_2 and a history h, define the *shift* of σ_1 by σ_2 at h, denoted $\sigma \doteq \sigma_1[h \leftarrow \sigma_2]$, to be the following strategy: for a history h', let $\sigma(h') \doteq \sigma_2(h')$ if h' extends h, and let $\sigma(h') \doteq \sigma_1(h')$ otherwise. I.e., σ does what σ_2 does from h (including) onwards, and elsewhere does what σ_1 does.

E. Borel Space

Let x be a finite prefix of an infinite path. Define $C_x \subseteq \Omega$ to consist of all infinite paths $\omega \in \Omega$ such that x is a prefix of ω. The sets C_x are called *cones*. For technical convenience, the empty set $\emptyset$ is also considered to be a cone. Notice that if $C_x \cap C_y \neq \emptyset$ then either x is a prefix of y, or y is a prefix of x, and so $C_x \subseteq C_y$ or $C_y \subseteq C_x$, and so $C_x \cap C_y \in \{C_x, C_y\}$. In other words, the intersection of two non-disjoint cones is equal to one of them. The cones form a basis for a topological space (Ω, τ), i.e., $X \in \tau$ iff X is a union of cones. The sets in τ are called *open*. It is worth noting that this topological space is metrizable, i.e., its topology is induced by the following definition of distance d in Ω: $d(\omega_1, \omega_2) = 2^{-k}$ where k is the length of the longest common prefix of ω_1, ω_2.[2] Let $\mathfrak{B}$ be the smallest Sigma-algebra (i.e., family of subsets of Ω closed under countable union and complement) containing τ. This gives the Borel space $(\Omega, \mathfrak{B})$. Borel sets in Ω will also be called *goals*.

For a strategy σ, define $\Omega_\sigma \subseteq \Omega$ to consist of all plays consistent with σ. The *induced* topological space (resp. Borel space) is the pair $(\Omega_\sigma, \tau_\sigma)$ (resp. $(\Omega_\sigma, \mathfrak{B}_\sigma)$), where τ_σ (resp. $\mathfrak{B}_\sigma$) consists of sets of the form $A \cap \Omega_\sigma$ where $A \in \tau$ (resp. $A \in \mathfrak{B}$). For every finite prefix x of a play in Ω_σ, the cone

[1] Strategies are sometimes called "schedulers", e.g., [3], [2].

[2] In fact, this distance is an ultrametric, it satisfies the following strengthening of the triangle inequality, i.e., $d(\omega_1, \omega_2) \leq \max(d(\omega_1, \omega_3), d(\omega_3, \omega_2))$.

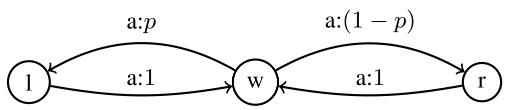

Fig. 1: Single-action stochastic domain for simulating a random walk on $\mathbb{Z}$ where $p \in (0, 1)$ is the probability of going left.

C_x in Ω_σ is thus the set of all plays in Ω_σ that extend x. Note that the notation is overloaded, i.e., C_x is in the induced topology Ω_σ, and thus also depends on σ. Nonetheless, this should not cause confusion since σ will always be clear from the context. As is conventional, we may simply write Ω_σ for this topological (resp. Borel) space. For $X \in \tau_\sigma$ (resp. $X \in \mathfrak{B}_\sigma$) we say that X is *open in* Ω_σ (resp. *Borel in* Ω_σ). If Ω_σ is understood from the context, we may write $\neg X$ as a shorthand for $\Omega_\sigma \setminus X$.

III. Probability spaces induced by strategies

To assign probabilities to Borel sets in a domain we need to fix a strategy (in other words, we need to pick a single action at every point). A stochastic domain $D = (St, Act, \iota, Pr)$ and a strategy σ induce a probability measure $\mu_{D,\sigma} : \mathfrak{B}_\sigma \to [0, 1]$, i.e., $\mu_{D,\sigma}$ is countably additive[3] and $\mu_{D,\sigma}(\Omega_\sigma) = 1$. This is done in the standard way, as follows. Define $\mu_{D,\sigma}$ as satisfying that $\mu_{D,\sigma}(C_\iota) = 1$ and $\mu_{D,\sigma}(C_{has}|C_{ha}) = Pr(h, a)(s)$ for $a = \sigma(h)$, and $\mu_{D,\sigma}(C_{ha}|C_h) = 1$ for $a = \sigma(h)$, and extend this uniquely to the Borel sets by Carathéodory's Extension Theorem. The tuple $(\Omega_\sigma, \mathfrak{B}_\sigma, \mu_{D,\sigma})$ is called a *probability space*, although we may simply refer to it as Ω_σ. We call $\mu_{D,\sigma}(X)$ the *probability* or *measure* of X (in Ω_σ). If $\mu_{D,\sigma}(X) = 1$ we may say that X *holds almost-surely*.

Remark 1. Note that given $h = (s_0, a_0, \cdots, s_n)$, the probability of a cone $C_h \subseteq \Omega_\sigma$ is equal to $\prod_{0 \le i < n} Pr(h_i, a_i)(s_{i+1})$, and is thus completely determined by the domain D (independently of σ). Consequently, $\mu_{D,\sigma}(C_h) = \mu_{D,\sigma'}(C_h)$ for every strategy σ' with which h is consistent.

Example 1 (Random Walk)**.** Let D_p be the domain pictured in Fig. 1 with $St = \{l, w, r\}, Act = \{a\}$, initial state w, and parameter $p \in (0, 1)$. Note that D_p and D_q have the same support for $p, q \in (0, 1)$. Let G consist of all traces that have some prefix with more occurrences of l than of r. The set G is Borel (in fact, it is an open set).[4] Since there is a single action, there is a single strategy σ. Observe that a trace is in G iff it represents a one-dimensional random walk on the integers that starts at 1 and visits 0. Thus, $\mu_{D_p,\sigma}(G) = 1$ if $p \ge 1/2$, and $0 < \mu_{D_p,\sigma}(G) < 1$ otherwise [14].

The following is a list of some basic properties of probability measures that we use implicitly. For Borel sets $A_i \subseteq \Omega_\sigma$:

- Monotonicity: $A_1 \subseteq A_2$ implies $\mu_{D,\sigma}(A_1) \le \mu_{D,\sigma}(A_2)$.
- Law of total probability: $\mu_{D,\sigma}(A_1) = \mu_{D,\sigma}(A_1 \cap A_2) + \mu_{D,\sigma}(A_1 \setminus A_2)$.
- Law of conditional probability: $\mu_{D,\sigma}(A_1 \cap A_2) = \mu_{D,\sigma}(A_1|A_2)\mu_{D,\sigma}(A_2)$.
- Subadditivity: $\mu_{D,\sigma}(\cup A_i) \le \sum_i \mu_{D,\sigma}(A_i)$ where $(A_i)_i$ is countable.

A countable sequence $(U_i)_i$ of cones in Ω_σ such that $A \subseteq \cup_i U_i$ is called a *cone-covering of* A. The *(Lebesgue) outer measure* is the function $\mu^*_{D,\sigma}$ defined for $A \subseteq \Omega_\sigma$ by $\mu^*_{D,\sigma} \doteq \inf\{\sum_i \mu_{D,\sigma}(U_i)\}$ where the infimum is taken over all cone-coverings $(U_i)_i$ of A. Note that this is well-defined since Ω_σ itself is a cone, and thus the inf ranges over a non-empty set.

The following lemma states the basic property that the outer measure of every Borel set is equal to its measure (i.e., probability). A proof in a more general setting can be found in [15].

Lemma 1. *If A is Borel in Ω_σ then $\mu_{D,\sigma}(A) = \mu^*_{D,\sigma}(A)$.*

We now state and prove two useful lemmas concerning our probability spaces.

The following lemma states that, in a given stochastic domain, the set of plays consistent with two strategies have the same measure in their respective probability spaces.

Lemma 2. *For a domain D, a Borel set G, and two strategies σ_1, σ_2, if $T \subseteq \Omega$ is the set of plays consistent with both σ_1 and σ_2, then T is Borel in Ω and $\mu_{D,\sigma_1}(T \cap G) = \mu_{D,\sigma_2}(T \cap G)$.*

Proof. First, if $T \cap G = \emptyset$ the lemma clearly holds, so we can assume that $T \cap G \neq \emptyset$. Note that Ω_σ is Borel in Ω: indeed, the set of plays consistent with a strategy σ is the countable intersection $\cap_n H_n$, where H_n is the union of all the cones C_h of histories h of length n consistent with σ. Hence, T is the intersection of two Borel sets, and thus also a Borel set.

To prove the equality of measures, we assume w.l.o.g. that the cone-coverings $(U_i)_i$ of a set A in Lemma 1 satisfy that $U_i \cap A \neq \emptyset$ for every i (indeed, discarding cones that cover nothing will not change the coverage, and only reduce the sum of measures). Take a cone-covering $(U_i)_i$ in Ω_{σ_1} of $T \cap G$ s.t. $U_i \cap (T \cap G) \neq \emptyset$ for every i. Since U_i is a cone in Ω_{σ_1}, there is a history h consistent with σ_1 such that U_i is the set of all plays in Ω_{σ_1} that extend h. Note that h is also consistent with σ_2 (since h is a prefix of some play in T), and define V_i to be the set of all plays in Ω_{σ_2} extending h. Recall that (c.f. Remark 1) $\mu_{D,\sigma_1}(U_i) = \mu_{D,\sigma_2}(V_i)$, and obtain that $\sum_i \mu_{D,\sigma_1}(U_i) = \sum_i \mu_{D,\sigma_2}(V_i)$. Observe that $(V_i)_i$ is a cone-covering of $T \cap G$ in Ω_{σ_2}. It follows that $\mu^*_{D,\sigma_1}(T \cap G) \ge \mu^*_{D,\sigma_2}(T \cap G)$. By symmetry, the converse is also true, and thus $\mu^*_{D,\sigma_1}(T \cap G) = \mu^*_{D,\sigma_2}(T \cap G)$. The lemma now follows from Lemma 1. ☐

The next lemma says that, in a given domain, probabilities conditioned on a cone only depend on distributions inside it:

Lemma 3. *Fix a domain $D = (St, Act, \iota, Pr)$, a strategy σ, a σ-history h, and a Borel set G. The measure $\mu_{D,\sigma}(G|C_h)$*

[3]This means that $\mu(\cup A_i) = \sum_i \mu(A_i)$ for every countable sequence $(A_i)_i$ of pairwise disjoint Borel sets.

[4]G is not omega-regular since, intuitively, one needs to remember the difference between the number of ls and rs, which requires infinite memory (formally, this can be proved using a standard pumping argument).

only depends on the distributions $Pr(h', -)$ where h' properly extends h.

Proof. Write μ for $\mu_{D,\sigma}$. Let $(U_i)_i$ be a cone-covering of $G \cap C_h$. Let $V_i = U_i \cap C_h$ and note the following: $V_i \subseteq C_h$, $G \cap C_h \subseteq \cup_i V_i$ and V_i is a cone (being the intersection of cones). Thus, $(V_i)_i$ is a cone-covering of $G \cap C_h$, and $\sum_i \mu(V_i) \leq \sum_i \mu(U_i)$ since $V_i \subseteq U_i$ for every i. Thus, by Lemma 1, $\mu(G \cap C_h) = \inf\{\sum_i \mu(V_i)\}$ where the infimum is taken over all cone-coverings $(V_i)_i$ of $G \cap C_h$ such that each $V_i \subseteq C_h$. Given $(V_i)_i$, let $\nu_i \in (0, 1]$ be such that $\mu(V_i) = \mu(C_h) \times \nu_i$. That is, if $h = (s_0, a_0, \cdots, s_m)$ and $V_i = C_x$ where $x = (s_0, a_0, \cdots, s_n)$ for $m \leq n$, then $\nu_1 = 1$ if $m = n$, and $\nu_i = \prod_{m \leq k < n} Pr(h_k, a_k)(s_{k+1})$ if $m < n$. So, $\mu(G|C_h) = \mu(G \cap C_h)/\mu(C_h) = \inf\{\sum_i \mu(V_i)\}/\mu(C_h) = \inf\{\sum_i \mu(C_h)\nu_i\}/\mu(C_h) = \inf\{\sum_i \nu_i\}$. Note that ν_i only depends on $Pr(h', -)$ for certain histories h' that properly extend h. □

We now state the Lebesgue Density Theorem (LDT). This theorem has its origins in the field of real analysis, and extensions of it to domains other than the real line or Euclidean spaces (as is our case) are usually found in the literature of geometric measure theory. A proof of it applying to our probability spaces can be found in the Appendix, or in [16] where it is given for the Cantor space.[5]

Fix a domain D, a strategy σ, and a goal G. For $\omega \in \Omega_\sigma$, let $C_{\omega[0,n)}$ be the cone C_h where h is the prefix of ω of length $n \geq 1$. The *density of G in D, σ at ω* is defined as

$$\lim_{n \to \infty} \mu_{D,\sigma}(G|C_{\omega[0,n)})$$

if this limit exists (if the limit does not exist, we say that the density is undefined). Let $\Lambda(G)$ be the set of plays $\omega \in G$ such that the density of G in D, σ at ω is equal to 1.

Theorem 1 (Lebesgue Density Theorem)**.** *Fix a domain D, a strategy σ, and a goal G. Then $\mu_{D,\sigma}(G \setminus \Lambda(G)) = 0$ (and thus $\mu_{D,\sigma}(G) = \mu_{D,\sigma}(\Lambda(G))$).*

We immediately get the following useful corollary.

Corollary 1. *Fix a domain D, a strategy σ, a σ-history h, and a goal G. If $\mu_{D,\sigma}(G|C_h) > 0$ then there exists a σ-play ω extending h s.t. the density of G in D, σ at ω is equal to* 1.

IV. Stochastic Best-Effort strategies

The ignorance that the system has about its stochastic environment is modeled as a set of "similar" domains.

Definition 1. *For $i = 0, 1$, consider domains $D_i = (St_i, Act_i, \iota_i, Pr_i)$, say with support functions Δ_i. Call the domains* similar *if they have the same set of states $St_0 = St_1$, the same set of actions $Act_0 = Act_1$, the same initial state $\iota_0 = \iota_1$, and the same support function $\Delta_0 = \Delta_1$.*

In particular, in similar domains only the values of the non-zero probabilities may be different. For instance, if D has Markovian support, then every domain E similar to D also has Markovian support (but E need not be Markovian).

[5]Not to be confused with the related, but different, Cantor set.

To define strategies that are Stochastic Best-Effort (SBE), we first define the "dominance" relation that says what it means for one strategy to do as least as well as another.

Definition 2 (Dominance)**.** *Let $\mathfrak{D}$ be a set of similar domains, and let G be a goal.*

(1) Define $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$ to mean that $\mu_{D,\sigma_1}(G) \geq \mu_{D,\sigma_2}(G)$ for every $D \in \mathfrak{D}$. In this case we say that σ_1 dominates *σ_2 wrt $\mathfrak{D}, G$.*

(2) Define $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$ to mean that $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$ and, in addition, $\mu_{D,\sigma_1}(G) > \mu_{D,\sigma_2}(G)$ for some $D \in \mathfrak{D}$. In this case we say that σ_1 strictly dominates *σ_2 wrt $\mathfrak{D}, G$.*

If $\mathfrak{D}, G$ are understood from the context, we may drop them from the subscripts. Intuitively, σ_1 dominates σ_2 means that σ_1 does at least as well as σ_2 in achieving the goal G in every domain in $\mathfrak{D}$, and it strictly dominates if it does strictly better in at least one domain $\mathfrak{D}$. The relation $\succcurlyeq$ is a pre-order (reflexive, transitive) and $\succ$ is a strict partial order (irreflexive, transitive) on the set of strategies. We say that σ_1, σ_2 are *incomparable* (wrt $\succcurlyeq$) if neither dominates the other.

Definition 3 (SBE)**.** *Fix a set $\mathfrak{D}$ of similar domains, and let G be a goal. A strategy σ is* maximal *(wrt $\succcurlyeq_{\mathfrak{D},G}$) if there is no σ' such that $\sigma' \succ_{\mathfrak{D},G} \sigma$. Maximal strategies are called* stochastic best-effort wrt $\mathfrak{D}, G$, *or* SBE *for short.*

Intuitively, if σ is not maximal, say σ' strictly dominates it, then a player that uses σ is not doing its "best" to achieve the goal since it could use σ' instead. Conversely, if σ is maximal, then there is no other strategy that would do "better" (although there may be other maximal strategies that are incomparable to σ).

V. Characterization of dominance and SBE

In this section we provide a characterization of dominance (Theorem 3), which in turn provides a useful characterization of SBE strategies (Theorem 4). The characterizations are in terms of a three-valued abstraction of the conditional probability $\mu_{D,\sigma}(G|C_h)$ for histories h consistent with the strategy σ. The characterizations are crucial for proving results in the rest of the paper, and we believe are of independent interest.

Definition 4. *[17] For a domain D, a goal G, a strategy σ, and a σ-history h, define*

$$val_{D,G}(\sigma, h) := \begin{cases} \textit{winning} & \text{if } \mu_{D,\sigma}(G|C_h) = 1 \\ \textit{pending} & \text{if } 0 < \mu_{D,\sigma}(G|C_h) < 1 \\ \textit{losing} & \text{if } \mu_{D,\sigma}(G|C_h) = 0 \end{cases}$$

We call $val_{D,G}(\sigma, h)$ the value of σ at h *or just the* value of h *or just the* value *when D, G, σ and h are clear from the context.*

Intuitively, *winning* means that σ almost-surely achieves the goal from h, *losing* that it almost-surely does not, and *pending* otherwise. We order the values as follows: *losing* < *pending* <

winning. When $\mathfrak{D}, G$ are given, we will use the following shorthand. Given a strategy σ and a σ-history h:

$$\min(h,\sigma) \doteq \min_{D\in\mathfrak{D}} val_{D,G}(h,\sigma)$$
$$\max(h,\sigma) \doteq \max_{D\in\mathfrak{D}} val_{D,G}(h,\sigma)$$

For example, $\max(h,\sigma) =$ *losing* means that for every domain $D \in \mathfrak{D}$ the value of σ at h is *losing*, and $\min(h,\sigma) =$ *winning* means that for every domain $D \in \mathfrak{D}$ the value of σ at h is *winning*. Also $\min(h,\sigma_1) \geq \max(h,\sigma_2)$ means that the best value that σ_2 achieves across all domains is no better than the worst value that σ_1 achieves across all domains.

A. Characterization of the dominance order

On the way to the desired characterization we will provide two intermediate, slightly weaker, characterizations.

To give the first characterization, we introduce the concept of an amplification of a domain along a history.

Definition 5 (Amplification)**.** *Given $\epsilon \in (0,1)$, a domain $D = (St, Act, \iota, Pr_D)$, and a history $h = s_0, a_0, \cdots, s_n$, we call $E = (St, Act, \iota, Pr_E)$ an* (h,ϵ)-amplification *of D if it is similar to D and satisfies:*

(1) $\prod_{i<n} Pr_E(s_i,a_i)(s_{i+1}) \geq \epsilon$;
(2) $Pr_E(x,a) = Pr_D(x,a)$ for every action a and every history x such that xa is not a proper prefix of h.

Intuitively, A (h,ϵ)-amplification is a domain E derived from D by modifying the transition function only for proper prefixes of h, and in a way that makes $\mu_{D,\sigma}(C_h) \geq \epsilon$ for every strategy σ which h is consistent with.

The following lemma is an easy consequence of the fact that a domain and an amplification of it have almost identical transition functions.

Lemma 4. *If E is an (h,ϵ)-amplification of D, then for every goal G and every strategy σ, we have that $\mu_{D,\sigma}(G|C_x) = \mu_{E,\sigma}(G|C_x)$ for every σ-history x such that $x\sigma(x)$ is not a proper prefix of h.*

Proof. Let Pr_D (resp. Pr_E) denote the transition function of domain D (resp. E). We consider two cases. The case that x is not a proper prefix of h is immediate from item 2 of the definition of amplification (Definition 5) and Lemma 3. For the case that x is a proper prefix of h, but $x\sigma(x)$ is not, let $s_1, \cdots, s_k$ be the states in the support of $Pr_E(x,\sigma(x))$. Then

$$\begin{aligned}\mu_{E,\sigma}(G|C_x) &= \sum_i \mu_{E,\sigma}(G|C_{x\sigma(x)s_i}) Pr_E(x,\sigma(x))(s_i)\\ &= \sum_i \mu_{D,\sigma}(G|C_{x\sigma(x)s_i}) Pr_E(x,\sigma(x))(s_i)\\ &= \sum_i \mu_{D,\sigma}(G|C_{x\sigma(x)s_i}) Pr_D(x,\sigma(x))(s_i)\\ &= \mu_{D,\sigma}(G|C_x)\end{aligned}$$

The first equality is by the definition of conditional probability. The second equality is since the σ-history $x\sigma(x)s_i$ is not a proper prefix of h and using the previous case. The third equality is by item 2 of Definition 5 and our assumption that $x\sigma(x)$ is not a proper prefix of h. The fourth equality is by definition of conditional probability and the fact that Pr_E and Pr_D have the same support since D and E are similar. □

We can now give one direction of the first characterization. In words, it says that if $\sigma_1 \succcurlyeq \sigma_2$ then for every split point h and every domain D the value of σ_1 is no worse than the value of σ_2, and both these values cannot be pending. Its proof uses the Lebesgue Density Theorem (Theorem 1).

Lemma 5. *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, let G be a goal, and let σ_1, σ_2 be such that $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$. Then for every split point h and every $D \in \mathfrak{D}$, the following two conditions hold:*

(1) $val_{D,G}(h,\sigma_1) \geq val_{D,G}(h,\sigma_2)$;
(2) both values above cannot be pending.

Proof. Suppose that for some domain $D \in \mathfrak{D}$ and split point h at least one of the two conditions fail. Then $val_{D,G}(h,\sigma_2)$ is not *losing* and $val_{D,G}(h,\sigma_1)$ is not *winning*. In order to conclude that it is not the case that $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$, we will find a domain $E \in \mathfrak{D}$ such that $\mu_{E,\sigma_2}(G) > \mu_{E,\sigma_1}(G)$. Intuitively, E should boost the probability that σ_2 satisfies G beyond that of σ_1. We will do this by choosing E to be an amplification of D along an extension h' of h so that $\mu_{E,\sigma_2}(C_{h'})$ is large (due to amplification), $\mu_{E,\sigma_2}(G|C_{h'})$ is large (due to the LDT), which will mean that $\mu_{E,\sigma_2}(G)$ is large. The tricky part is to ensure that while doing this the probability of σ_1 satisfying G, which also changes in this process, does not grow too much.

Let $\alpha \doteq \mu_{D,\sigma_1}(G|C_h)$, and note that $0 \leq \alpha < 1$ since $val_{D,G}(h,\sigma_1)$ is not *winning*. Since $val_{D,G}(h,\sigma_2)$ is not *losing*, apply Corollary 1 to get a σ_2-play ω extending h such that the density of G in D,σ_2 at ω is 1. Hence, we can pick a history h', that extends h and is a prefix of ω, such that $\beta \doteq \mu_{D,\sigma_2}(G|C_{h'}) \geq \sqrt{\alpha'}$ where $\alpha' \doteq (\alpha+1)/2$. Let E be an $(h',\sqrt{\alpha'})$-amplification of D. Note that $\mu_{E,\sigma_2}(G) \geq \mu_{E,\sigma_2}(G|C_{h'})\mu_{E,\sigma_2}(C_{h'}) \geq \beta\sqrt{\alpha'} \geq \alpha'$, where the second inequality is by Lemma 4. Since h is a split point of σ_1,σ_2, Lemma 4 also tells us that $\alpha = \mu_{E,\sigma_1}(G|C_h)$. Also, item 1 of the definition of amplification (Definition 5) means that $\mu_{E,\sigma_1}(C_h) \geq \sqrt{\alpha'}$. Thus:

$$\begin{aligned}&\mu_{E,\sigma_1}(G)\\ &= \mu_{E,\sigma_1}(G|C_h)\mu_{E,\sigma_1}(C_h) + \mu_{E,\sigma_1}(G|\neg C_h)\mu_{E,\sigma_1}(\neg C_h)\\ &\leq \alpha\mu_{E,\sigma_1}(C_h) + 1(1-\mu_{E,\sigma_1}(C_h))\\ &\leq \alpha + (1-\sqrt{\alpha'})\\ &< \alpha + (1-\alpha') = (\alpha+1) - (\alpha+1)/2 = \alpha'\end{aligned}$$

where the last inequality is because $\alpha' \in (0,1)$. Overall, $\mu_{E,\sigma_2}(G) > \mu_{E,\sigma_1}(G)$, as promised. □

We are now ready to give our first characterization of dominance:

Theorem 2. *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, let G be a goal, and let σ_1, σ_2 be two strategies. Then,*

$\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$ iff for every split point h and every $D \in \mathfrak{D}$, the following two conditions hold:
(1) $val_{D,G}(h, \sigma_1) \geq val_{D,G}(h, \sigma_2)$.
(2) both values cannot be pending.
Also, $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$ iff, in addition, there is a split point h and a domain $D \in \mathfrak{D}$ such that $val_{D,G}(h, \sigma_1) > val_{D,G}(h, \sigma_2)$.

Proof. The $\Rightarrow$ of direction for $\succcurlyeq$ is by Lemma 5. We now prove the $\Leftarrow$ direction for $\succcurlyeq$. Consider a domain $D \in \mathfrak{D}$. We will show that $\mu_{D,\sigma_1}(G) \geq \mu_{D,\sigma_2}(G)$. Let T be the set of plays consistent both with σ_1 and σ_2. By the fact that $\mu_{D,\sigma_i}(G) = \mu_{D,\sigma_i}(G \setminus T) + \mu_{D,\sigma_i}(G \cap T)$ and by Lemma 2 it is sufficient to show that $\mu_{D,\sigma_1}(G \setminus T) \geq \mu_{D,\sigma_2}(G \setminus T)$. Now, $\mu_{D,\sigma_i}(G \setminus T) = \sum_h \mu_{D,\sigma_i}(G \cap C_h)$ where the sum is over all split points h of σ_1, σ_2. But notice that the two assumed conditions of the Theorem give us that $\mu_{D,\sigma_1}(G|C_h) \geq \mu_{D,\sigma_2}(G|C_h)$, and so $\mu_{D,\sigma_1}(G \cap C_h) \geq \mu_{D,\sigma_2}(G \cap C_h)$.

The statement for $\succ$ follows from that for $\succcurlyeq$. Indeed, if $\sigma_1 \succcurlyeq \sigma_2$, then $\sigma_1 \succ \sigma_2$ iff $\sigma_2 \not\succcurlyeq \sigma_1$; in other words, iff there is some split point h and domain D such that either the first condition fails, i.e., $val_{D,G}(h, \sigma_2) < val_{D,G}(h, \sigma_1)$ or the second condition fails. However, the second condition cannot fail since we assumed that $\sigma_1 \succcurlyeq \sigma_2$. □

We now give a second characterization of dominance which does not compare values one domain at a time (as Theorem 2 does), but rather over all domains at once (for a given history). In words, it will say that $\sigma_1 \succcurlyeq \sigma_2$ means that the worst value that σ_1 achieves (across all domains) is no worse than the best value that σ_2 achieves (across all domains). In order to do that, we will combine different domains using the following definition.

Definition 6 (Shift). *Given two similar domains $E_i = (St, Act, \iota, Pr_i)$, a history h, and an action a, let $D = (St, Act, \iota, Pr)$ be the similar domain where $Pr(h', a')$ is defined as follows: for every action a' and history h',*

- *$Pr(h', a') = Pr_1(h', a')$ if ha is a prefix of h',*
- *$Pr(h', a') = Pr_2(h', a')$, otherwise.*

We call D the shift of E_2 by E_1 at history h and action a.

Intuitively, the shift domain D looks like E_2, except that following action a from h it looks like E_1. This means that if two strategies σ_1, σ_2 split at h, and if D is the shift of E_2 by E_1 at h and action $\sigma_1(h)$, then by Lemma 3, σ_i achieves the same value at h in D as it does in E_i ($i = 1, 2$).

Proposition 1. *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, let G be goal, and let σ_1, σ_2 be two strategies. Then, $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$ iff for every split point h, the following two conditions hold:*
(I) $\min(h, \sigma_1) \geq \max(h, \sigma_2)$
(II) this min and max are not both pending.
Also, $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$ iff, in addition, there is a split point h, such that $\max(h, \sigma_1) > \min(h, \sigma_2)$.

Proof. Fix a split point h. We will show that both conditions (I) and (II) hold iff both conditions (1) and (2) in Theorem 2 hold.

Suppose (I) and (II) hold. Consider a domain $E \in \mathfrak{D}$. Then:

$$val_{E,G}(h, \sigma_1) \geq \min(h, \sigma_1) \geq \max(h, \sigma_2) \geq val_{E,G}(h, \sigma_2),$$

which implies that both conditions (1) and (2) hold.

For the converse, suppose that either (I) or (II) fail. We will find a domain $D \in \mathfrak{D}$ such that either (1) or (2) fail. Let $E_1, E_2 \in \mathfrak{D}$ be domains that witness the min and max respectively, i.e., $val_{E_1,G}(h, \sigma_1) = \min(h, \sigma_1)$ and $val_{E_2,G}(h, \sigma_2) = \max(h, \sigma_2)$. Say $E_i = (St, Act, \iota, Pr_i)$. Let D be the shift of E_2 by E_1 at history h and action $\sigma_1(h)$. Note that $D \in \mathfrak{D}$ (it is similar to, e.g., E_1), and by Lemma 3, $val_{D,G}(h, \sigma_1) = val_{E_1,G}(h, \sigma_1)$ and $val_{D,G}(h, \sigma_2) = val_{E_2,G}(h, \sigma_2)$. If (I) fails then $val_{D,G}(h, \sigma_1) < val_{D,G}(h, \sigma_2)$ and so (1) fails. If (II) fails then $val_{D,G}(h, \sigma_1) = val_{D,G}(h, \sigma_2) = $ *pending*, and so (2) fails.

To see the statement for $\succ_{\mathfrak{D},G}$, suppose that $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$. Then $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$ iff it is not the case that $\sigma_2 \succcurlyeq_{\mathfrak{D},G} \sigma_1$ iff there is some split point h such that either (I) fails, i.e., it is not the case that $\min(h, \sigma_2) \geq \max(h, \sigma_1)$, which is the stated condition, or (II) fails.

We now rule out (II) failing, i.e., there is no split point h such that $\max(h, \sigma_1)$ and $\min(h, \sigma_2)$ are *pending*. Similar to above, let E_1 witness $\max(h, \sigma_1)$ and let E_2 witness $\min(h, \sigma_2)$, and let D be as before. Then $val_{D,G}(h, \sigma_1) = val_{D,G}(h, \sigma_2) = $ *pending*, which contradicts Lemma 5. □

Proposition 1 yields our final characterization of dominance.

Theorem 3 (Characterization of the dominance order)**.** *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, let G be a goal, and let σ_1, σ_2 be two strategies. Then $\sigma_1 \succcurlyeq \sigma_2$ iff for every split point h, either*
(1) $\min(h, \sigma_1) = $ winning,
(2) $\max(h, \sigma_2) = $ losing.
Also, $\sigma_1 \succ \sigma_2$ iff, in addition, for some split point h, either
(A) $\min(h, \sigma_1) = $ winning and $\min(h, \sigma_2) \neq $ winning;
(B) $\max(h, \sigma_2) = $ losing and $\max(h, \sigma_1) \neq $ losing.

Proof. Suppose condition (1) or condition (2) of this Theorem hold for h, then clearly (I) and (II) of Proposition 1 hold for h. Conversely, suppose both (1) and (2) fail for h. Then $\min(h, \sigma_1)$ is either *losing* or *pending*, and $\max(h, \sigma_2)$ is either *winning* or *pending*. If (I) holds, then both values must be *pending*, which means (II) doesn't hold.

To see the characterization of $\succ_{\mathfrak{D},G}$ holds, suppose $\sigma_1 \succcurlyeq_{\mathfrak{D},G} \sigma_2$. So we have, for every split point h, by (I) and (II), that $\max(h, \sigma_1) \geq \min(h, \sigma_1) \geq \max(h, \sigma_2) \geq \min(h, \sigma_2)$, and $\min(h, \sigma_1), \max(h, \sigma_2)$ are not both *pending*.

Suppose for some split point h, (A) holds or (B) holds. In the first case $\max(h, \sigma_1)$ is *winning* and $\min(h, \sigma_2)$ is either *pending* or *losing*, and in the second case $\max(h, \sigma_1)$ is either *pending* or *winning* and $\min(h, \sigma_2)$ is *losing*. In either case, $\max(h, \sigma_1) > \min(h, \sigma_2)$, thus $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$ by Proposition 1.

Conversely, suppose $\sigma_1 \succ_{\mathfrak{D},G} \sigma_2$. By Proposition 1, there is a split point h such that (*) $\max(h, \sigma_1) > \min(h, \sigma_2)$. We will show that either (A) holds or (B) holds. There are three cases depending on the value $\min(h, \sigma_1)$. First, if $\min(h, \sigma_1)$ is *winning*, then $\max(h, \sigma_1)$ is *winning*, and by (*) we have $\min(h, \sigma_2)$ is not *winning*, and so (A) holds. Second, if $\min(h, \sigma_1)$ is *losing*, then $\max(h, \sigma_2)$ is *losing*, and by (*) we have that $\max(h, \sigma_1)$ is not *losing*, and so (B) holds. Third, if $\min(h, \sigma_1)$ is *pending*, then since $\max(h, \sigma_2)$ is not *pending*, we must have that $\max(h, \sigma_2)$ is *losing*, and so $\min(h, \sigma_2)$ is losing, and so by (*), $\max(h, \sigma_1)$ is not *losing*, and so (B) holds. ☐

B. Characterization of SBE strategies

Theorem 3 allows us to characterize SBE as those strategies σ that, at every σ-history, win almost surely in every domain, if this is possible by any strategy at all, and otherwise do not lose almost surely in every domain, if this is possible by any strategy at all.

Theorem 4 (Characterization of SBE strategies)**.** *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, and let G be a goal. A strategy σ is SBE wrt $\mathfrak{D}, G$ iff for every σ-history h:*

(i) If there exists σ' such that $\min(h, \sigma') =$ winning then $\min(h, \sigma) =$ winning;

(ii) If there exists σ' such that $\max(h, \sigma') \neq$ losing then $\max(h, \sigma) \neq$ losing;

Proof. Suppose σ is not SBE. Then there is some $\sigma' \succ_{\mathfrak{D},G} \sigma$. By Theorem 3, there is a split point h such that (A) or (B) holds. If (A) holds then (i) fails, and if (B) holds then (ii) fails.

Conversely, suppose σ fails the characterization, i.e., there is some σ-history h such that either (i) fails or (ii) fails. We will show σ is not maximal.

Suppose (i) fails, and let σ' be the stated strategy such that $val_{D,G}(h, \sigma')$ is *winning* for every $D \in \mathfrak{D}$, and let $E \in \mathfrak{D}$ be the domain such that $val_{D,G}(h, \sigma)$ is not *winning*. In particular: $\mu_{D,\sigma'}(G|C_h) \geq \mu_{D,\sigma}(G|C_h)$ for every $D \in \mathfrak{D}$, and $\mu_{E,\sigma'}(G|C_h) > \mu_{E,\sigma}(G|C_h)$. Let $\sigma_2 = \sigma[h \leftarrow \sigma']$, i.e., the shift of σ by σ' at h (defined in Section II-D). We claim that $\sigma_2 \succ_{\mathfrak{D},G} \sigma$. Intuitively this is because σ_2 is the same as σ except after following h it achieves the goal with probability 1 (as σ' does) instead of with probability < 1 (as σ does). Formally, for every domain D:

$$
\begin{aligned}
&\mu_{D,\sigma_2}(G) \\
&= \mu_{D,\sigma_2}(G|C_h)\mu_{D,\sigma_2}(C_h) + \mu_{D,\sigma_2}(G|\neg C_h)\mu_{D,\sigma_2}(\neg C_h) \\
&= \mu_{D,\sigma'}(G|C_h)\mu_{D,\sigma}(C_h) + \mu_{D,\sigma}(G|\neg C_h)\mu_{D,\sigma}(\neg C_h) \\
&\geq \mu_{D,\sigma}(G|C_h)\mu_{D,\sigma}(C_h) + \mu_{D,\sigma}(G|\neg C_h)\mu_{D,\sigma}(\neg C_h) \\
&= \mu_{D,\sigma}(G),
\end{aligned}
$$

and the inequality is strict for $D = E$. The second equality follows from Lemma 3.

The case (ii) is similar. Suppose (ii) fails. Then $val_{D,G}(h, \sigma)$ is *losing* for every domain D, and let σ' be the strategy such that $val_{D,G}(h, \sigma')$ is not *losing* for some $E \in \mathfrak{D}$. In particular $\mu_{E,\sigma'}(G|C_h) > \mu_{E,\sigma}(G|C_h)$. Let $\sigma_2 = \sigma[h \leftarrow \sigma']$, i.e., the shift of σ by σ' at h. Then, by the same calculation as before, we have that $\sigma_2 \succ_{\mathfrak{D},G} \sigma$. ☐

VI. Existence

In this section we use the characterization of SBE (Theorem 4) to prove that SBE strategies exist in quite a general setting. We begin with the following observation that a *winning* or *losing* value of a history is inherited by the histories that extend it.

Lemma 6. *For a domain D, a goal G, a strategy σ, and a σ-history h: if $val_{D,G}(h, \sigma) \in \{winning, losing\}$ then for every σ-history h' extending h we have that $val_{D,G}(h, \sigma) = val_{D,G}(h', \sigma)$.*

Proof. We will prove the case $val_{D,G}(h, \sigma) =$ *losing*. The other case follows by observing that for every σ-history w we have that $val_{D,G}(w, \sigma) =$ *winning* iff $val_{D,\Omega\setminus G}(w, \sigma) =$ *losing*, and applying the first case to $\Omega \setminus G$. Assume then that $\mu_{D,\sigma}(G|C_h) = 0$, and by contradiction that $\mu_{D,\sigma}(G|C_{h'}) > 0$. Recall that $\mu_{D,\sigma}(G|C_{h'}) = \mu_{D,\sigma}(G \cap C_{h'})/\mu_{D,\sigma}(C_{h'})$, and thus $\mu_{D,\sigma}(G \cap C_{h'}) > 0$. Since h' extends h we have that $G \cap C_{h'} \subseteq G \cap C_h$, hence $\mu_{D,\sigma}(G \cap C_h) > 0$. Conclude that $\mu_{D,\sigma}(G \cap C_h)/\mu_{D,\sigma}(C_h) = \mu_{D,\sigma}(G|C_h) > 0$, which is a contradiction. ☐

The following theorem states that SBE strategies exist for every Borel goal.

Theorem 5. *Let $\mathfrak{D}$ be the set of all domains similar to a fixed domain, and let G be a Borel set. Then there exists a stochastic best-effort strategy wrt $\mathfrak{D}, G$.*

Proof. Fix $\mathfrak{D}, G$ as in the statement of the theorem, and write $\succ$ for $\succ_{\mathfrak{D},G}$.

To define σ, we define a sequence of strategies, $\sigma_0, \sigma_1, \sigma_2, \cdots$ that eventually stabilizes on each history, i.e., for every h there exists m such that $\sigma_m(h) = \sigma_{m+1}(h) = \sigma_{m+2}(h) = \cdots$. We then define σ as the point-wise limit of this sequence, i.e., $\sigma(h) := \lim_i \sigma_i(h)$. The sequence is constructed as follows (note that the construction makes possibly many arbitrary choices, so it is not unique).

Fix an ordering on $St \cup Act$, and note that this induces a length-lexicographic ordering on histories. Start with all histories *unmarked*, and with σ_0 being an arbitrary strategy. At the start of round $i \geq 0$, consider the smallest unmarked history h in the length-lexicographic order. Mark h as stabilized, and define σ_{i+1}, the strategy resulting from round i, as follows:

(L1) If $\min(h, \sigma_i) <$ *winning*, but $\min(h, \sigma') =$ *winning* for some σ', then let $\sigma_{i+1} = \sigma_i[h \leftarrow \sigma']$;

(L2) otherwise, if $\max(h, \sigma_i) =$ *losing*, but $\max(h, \sigma') >$ *losing* for some σ', then let $\sigma_{i+1} = \sigma_i[h \leftarrow \sigma']$;

(L3) otherwise, let $\sigma_{i+1} = \sigma_i$.

Finally, if $\min(h, \sigma_{i+1}) =$ *winning* or $\max(h, \sigma_{i+1}) =$ *losing* then: (L4) mark all the proper extensions of h as stabilized.

We now prove that σ is SBE. Assume, for a contradiction, that there is some strategy $\sigma' \succ \sigma$. By Theorem 3, we know that for some split point h of σ, σ', either:

(A) $\min(h, \sigma')$ is *winning* and $\min(h, \sigma)$ is not *winning*; or

(B) $\max(h, \sigma)$ is *losing* and $\max(h, \sigma')$ is not *losing*.

Let i be the round in the construction of σ where h was marked as stabilized. We consider two cases depending on whether or not (L4) was invoked at round i.

First, suppose that (L4) was invoked at round i, and let h' be the history considered at this round. Note that h' is a (not necessarily proper) prefix of h. Since all extensions of h' are marked as stabilized in round i, we know that σ_{i+1} and σ agree on h' and all its extensions, and thus have the same value at h' in all domains. Since (L4) was invoked, either $\min(h', \sigma) = \min(h', \sigma_{i+1}) = \textit{winning}$ or $\max(h', \sigma) = \max(h', \sigma_{i+1}) = \textit{losing}$. The option $\min(h', \sigma) = \textit{winning}$ implies, by Lemma 6, that $\min(h, \sigma) = \textit{winning}$, which contradicts both (A) and (B); whereas the option $\max(h', \sigma_{i+1}) = \textit{losing}$ also leads to a contradiction since our assumption that either (A) or (B) hold implies that $\max(h, \sigma') \neq \textit{losing}$, which would have ensured (by (L1) and (L2)) that $\max(h', \sigma_{i+1}) \neq \textit{losing}$.

Second, suppose that (L4) was not invoked at round i. Then, in particular, h was the history considered at this round. We claim that (A) does not hold. Indeed, if $\min(h, \sigma') = \textit{winning}$, then by (L1) also $\min(h, \sigma_{i+1}) = \textit{winning}$, which would have invoked (L4). We are left with the case that (B) holds, i.e., that: (†) $\max(h, \sigma) = \textit{losing}$ and $\max(h, \sigma') \neq \textit{losing}$. The last fact implies, by (L1) and (L2), that also $\max(h, \sigma_{i+1}) \neq \textit{losing}$. Intuitively, this means that the value of h deteriorated in the limit σ compared to when it was marked. We complete the proof by showing that this is not possible due to the greedy nature of the construction.

We first find a split point w of σ, σ_{i+1} such that: (‡) w extends h and $\min(w, \sigma_{i+1}) \neq \textit{losing}$. To do this, observe that since $\max(h, \sigma_{i+1}) \neq \textit{losing}$ and $\max(h, \sigma) = \textit{losing}$, we have that $\mu_{D,\sigma_{i+1}}(G|C_h) > 0$ for some $D \in \mathfrak{D}$, and $\mu_{D,\sigma}(G|C_h) = 0$ for all $D \in \mathfrak{D}$. Hence, together with Definition 2, we get that $\sigma_{i+1} \succ_{\mathfrak{D},F} \sigma$, where F is a new goal consisting of all traces in G that extend h. By Theorem 2, there is some split point w of σ_{i+1}, σ and some domain $D \in \mathfrak{D}$, where $val_{D,F}(w, \sigma_{i+1}) > val_{D,F}(w, \sigma)$. Note that this implies, in particular, that $\mu_{D,\sigma_{i+1}}(F) > 0$ and thus w extends h and $\mu_{D,\sigma_{i+1}}(G) > 0$, so w satisfies ‡.

Let $k > i$ be the smallest round in which a prefix (not necessarily proper) x of w was considered and (L3) not invoked (i.e., a shift in (L1) or (L2) was done). Such a k exists because, being a split point, $\sigma_{i+1}(w) \neq \sigma(w)$, so k is bounded from above by the number of the round in which w was marked as stabilized. The minimality of k implies that from round $i + 1$ up to the start of round k no shift was invoked on any prefix of w (and thus on any prefix of x), so σ_{i+1} and σ_k agree on x and all of its extensions. Hence, $\min(x, \sigma_{i+1}) = \min(x, \sigma_k)$. Observe that ‡ and Lemma 6 imply that $\min(x, \sigma_{i+1}) \neq \textit{losing}$, so it must be that the shift at round k was done in (L1). Hence, $\min(x, \sigma_{k+1}) = \textit{winning}$, and (L4) was invoked at the end of round k making σ and σ_{k+1} agree on x and all its extensions, and thus also $\min(x, \sigma) = \textit{winning}$. Observe that since x and h are both prefixes of w, and x was considered at a later round than h, it must be that x extends h. Alas, this is a contradiction since then † and Lemma 6 yield that $\max(x, \sigma) = \textit{losing}$. □

VII. Complexity of finding SBE strategies for linear-temporal goals

While in the previous section we showed existence of SBE strategies, in this section we study the computational complexity of the problem of finding such strategies. We thus need to consider finite representations of $\mathfrak{D}$ and G. We focus on the important special case where G is given by a formula of linear-temporal logic, and $\mathfrak{D}$ is a set of domains that share the same Markovian (resp. finitely generated) support function. Additionally we require $\mathfrak{D}$ to be a set of *bounded* domains, where a domain D is bounded if there is some $\epsilon > 0$ such that $Pr(h, a)(s) \geq \epsilon$ for all h, a and s such that $s \in sup(Pr(h, a))$; technically, boundedness guarantees that fair traces have probability 1 (cf [4]).

Before we get to the upper and lower bounds, we make the following important observation. Note that Theorem 5, which states that SBE strategies exist for all Borel goals, does not apply to the set of all bounded domains that are similar to some fixed domain. However, it turns out that the proof of that Theorem can easily be generalized to this setting. Indeed, the only place in the proof where we took advantage of the fact that we started with the set of all domains similar to some fixed domain is in Lemma 5 where we picked an amplification domain and in Proposition 1 where we picked a domain using a shift. We say that $\mathfrak{D}$ is *closed under amplification* if for every $D \in \mathfrak{D}$ and every h, ϵ there is some (h, ϵ)-amplification of D in $\mathfrak{D}$. We say that $\mathfrak{D}$ is *closed under shift* if for every $E_1, E_2 \in \mathfrak{D}$ and every history h and action a, the shift of E_2 by E_1 at h and a is in $\mathfrak{D}$. Thus, we get that if $\mathfrak{D}$ is a set of similar domains that is closed under amplification and shift, and G is a Borel set, then there exists a stochastic best-effort strategy wrt $\mathfrak{D}, G$. Note that the set of all bounded domains $\mathfrak{D}$ that are similar to some fixed D is closed under amplification and shift. This yields:

Corollary 2. *Let $\mathfrak{D}$ be the set of bounded domains that are similar to a fixed domain D, and let G be a goal. Then:*

1) *There exists a stochastic best-effort strategy wrt $\mathfrak{D}, G$.*
2) *A strategy σ is SBE wrt $\mathfrak{D}, G$ iff for every σ-history h:*
   (i) *If there exists σ' such that $\min(h, \sigma') = \textit{winning}$ then $\min(h, \sigma) = \textit{winning}$;*
   (ii) *If there exists σ' such that $\max(h, \sigma') \neq \textit{losing}$ then $\max(h, \sigma) \neq \textit{losing}$;*

Remark 2. In general, some closure properties on $\mathfrak{D}$ are necessary for SBE strategies to exist. To see this, consider a set of domains $\mathfrak{D}$ that consists of a single domain D (note that $\mathfrak{D}$ is obviously not closed under shifts and amplifications). Then, a strategy σ is SBE wrt $\mathfrak{D}, G$ iff σ is optimal for D. However, it is well-known that optimal strategies do not need to exist, e.g., [17] exhibits a domain D that has Markovian support (but is non-Markovian), and a reachability goal G, for which there is no optimal strategy.

We now recall the definition of linear-time temporal logic. The *formulas of* LTL over a finite set AP of *atoms* are defined by the following BNF (where $p \in AP$):

$$\varphi ::= p \,|\, \varphi \vee \varphi \,|\, \neg\varphi \,|\, \mathsf{X}\varphi \,|\, \varphi \,\mathsf{U}\, \varphi.$$

We use the usual abbreviations including the following: $\mathsf{F}\,\varphi \doteq \mathsf{true}\,\mathsf{U}\,\varphi$, $\mathsf{G}\,\varphi \doteq \neg\,\mathsf{F}\,\neg\varphi$. A *trace* τ is an infinite sequence of valuations of the atoms, i.e., $\tau \in (2^{AP})^\omega$. For $n \geq 0$, write τ_n for the valuation at position n. Given τ, n, and φ, the satisfaction relation $(\tau, n) \models \varphi$, stating that φ holds at step n of the sequence τ, is defined as follows:

- $(\tau, n) \models p$ iff $p \in \tau_n$;
- $(\tau, n) \models \varphi_1 \vee \varphi_2$ iff $(\tau, n) \models \varphi_1$ or $(\tau, n) \models \varphi_2$;
- $(\tau, n) \models \neg\varphi$ iff $(\tau, n) \models \varphi$ does not hold;
- $(\tau, n) \models \mathsf{X}\,\varphi$ iff $(\tau, n+1) \models \varphi$; and
- $(\tau, n) \models \varphi_1 \,\mathsf{U}\, \varphi_2$ iff there exists m with $n \leq m$ such that: $(\tau, m) \models \varphi_2$ and $(\tau, j) \models \varphi_1$ for all j with $n \leq j < m$.

If $\tau, 0 \models \psi$ we write $\tau \models \psi$ and say that τ *satisfies* ψ and that τ is a *model* of ψ.

We treat LTL formulas φ as goals. To do this, we introduce a *labeling function* $\lambda : St \cup Act \rightarrow 2^{AP}$. If $\pi = (s_0, a_0, s_1, a_1, \cdots)$ is an infinite path (or a finite path that ends in an action), define the trace $\lambda(\pi)$ as $(\lambda(s_0) \cup \lambda(a_0), \lambda(s_1) \cup \lambda(a_1), \cdots)$. If π is an infinite path, we say that π *satisfies* φ iff $\lambda(\pi) \models \varphi$. Let $[\varphi]_\lambda \subseteq \Omega$ denote the set of infinite paths that satisfy φ. We remark that $[\varphi]_\lambda$ is Borel (this follows from the more general fact that $[\varphi]_\lambda$ is omega-regular and thus Borel [18]). We may write $val_{D,\varphi}(-,-)$ instead of the more correct $val_{D,[\varphi]_\lambda}(-,-)$.

In the rest of this section we establish matching upper and lower bounds for the computational complexity of the following.

Definition 7. *The following problem is called* SBE synthesis for LTL goals and bounded domains with Markovian support*: Given a Markovian domain D, a labeling function λ, and an LTL formula φ, find a finite-state strategy that is stochastic best-effort wrt* $\mathfrak{D}, [\varphi]_\lambda$ *where* $\mathfrak{D}$ *is the set of all bounded domains that are similar to D.*

Although Corollary 2 only guarantees the existence of a SBE strategy, the algorithm in Section VII-A will show that there exists a finite-state one. Thus, the SBE synthesis problem is a search problem that always has a solution.

A. Upper bound

We establish a 2EXPTIME-upper bound for the problem in Definition 7.

Theorem 6. *The SBE synthesis problem for LTL goals and bounded domains with Markovian support can be solved in* 2*EXPTIME.*

Proof. We use the following from [17]. For a domain D, an LTL formula φ, and a labeling λ, let $G = [\varphi]_\lambda$ and consider the following property: $(\dagger_D)$ for every σ-history h,

$$val_{D,G}(\sigma, h) = \max_{h \text{ is a } \sigma'\text{-history}} val_{D,G}(\sigma', h)$$

In words, $(\dagger_D)$ means that for every σ-history h and every strategy σ', the value of σ at h is at least the value of σ' at h. That paper also proves that if $\mathfrak{D}$ is any set of similar domains, all of which are bounded and have Markovian support, then there exists a finite-state strategy σ that satisfies $(\dagger)_D$ for every $D \in \mathfrak{D}$, and one such strategy can be computed in 2EXPTIME (from the common Markovian support function $St \times Act \rightarrow 2^{St}$, the formula φ, and the labeling λ).[6]

We now prove our result. Take an instance D, λ, φ of the SBE synthesis problem. Let $\mathfrak{D}$ be the set of all bounded domains similar to the Markovian domain D. Observe that a strategy satisfying $(\dagger)_D$ for every $D \in \mathfrak{D}$ satisfies the conditions in our characterization of SBE (Corollary 2), and thus is SBE wrt $\mathfrak{D}, G$. ☐

Remark 3. Let $(\dagger)_\mathfrak{D}$ denote that $(\dagger)_D$ holds for all $D \in \mathfrak{D}$. The work in [17] calls a strategy SBE if it satisfies $(\dagger)_\mathfrak{D}$. In contrast, this work calls a strategy SBE if it is maximal in the dominance order (Definition 3). Unfortunately, these two definitions do not always coincide[7]: while for every Borel goal a strategy satisfying $(\dagger)_\mathfrak{D}$ also satisfies the conditions in Theorem 4, and is thus maximal in the dominance order (i.e., it is also SBE by Definition 3), the reverse is not true. In other words, $(\dagger)_\mathfrak{D}$ cannot, in general, be taken as a characterization of maximal strategies in the dominance order. Indeed, whereas Theorem 5 shows that maximal strategies always exist, strategies satisfying $(\dagger)_\mathfrak{D}$ may not exist for goals that are not omega-regular, as the following example demonstrates.

Example 2. Consider the domains in Figure 2 with $p \in (0, 1)$. There are two types of strategies: σ_0 does action a on the first move, and σ_1 does action b in the first move (after the first move it does not matter whether a or b is chosen). Define the goal G to consist of all traces that have some prefix with more occurrences of l_0 than of r_0 or more occurrences of l_1 than of r_1. Note that G is Borel (it is open), but not omega-regular.

Observe that a trace is in the goal iff it represents a one-dimensional random walk that starts at 1 and reaches 0. If in a random walk the probability of going left is $p \in (0, 1)$, then the probability the goal is satisfied is equal to 1 if $p \geq 1/2$ and is not equal to 1, but is positive, if $p < 1/2$ (as in Example 1).

Let D_0 (resp. D_1) be the domain obtained by taking $p = 3/4$ (resp. $p = 1/4$). Observe that, for $i \in \{0, 1\}$, we have that $\mu_{D_i,\sigma_i}(G) = 1$, and $0 < \mu_{D_i,\sigma_{1-i}}(G) < 1$. It follows

[6] Technically, the domains in [17] and our domains are not exactly the same. The former uses a specific representation of states and actions: the state set is of the form 2^F for some finite set F of Boolean variables, and the action set is of the form 2^A for some finite set A of Boolean variables, and the atoms of the formula are over $F \cup A$ (and thus there is no need for a labeling function). It is not hard to suitably encode our stochastic domains in this representation. Alternatively, the techniques in [17] do not rely on this special representation, and apply just as well to our representation of stochastic domains.

[7] The reason we use the term SBE in Definition 3 despite the potential confusion with [17] is two-fold. First, we see the work in [17] as a preliminary step in the road to this work; Second, [17] only considers omega-regular goals, and for these the two definitions of SBE coincide (if one further restricts to finite-state strategies). Thus, there is in fact only a small room for confusion.

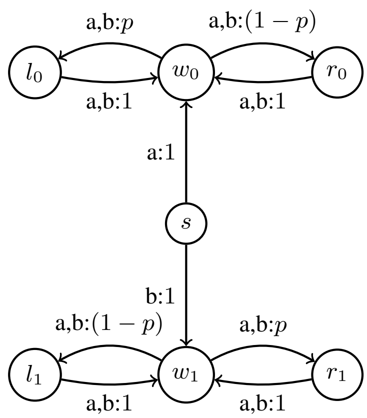

Fig. 2: Stochastic domain for simulating two random walks

that in D_i there is a strategy with value *winning* at the history $h = s$, but neither σ_0 nor σ_1 achieve value *winning* on both domains. So, taking $\mathfrak{D}$ to be the set of all domains (resp. bounded domains) similar to, say, D_1, we see that there is no strategy that satisfies $(\dagger)_{\mathfrak{D}}$.

B. Lower bound

We establish a matching 2EXPTIME lower-bound. We first recall the following decision problem:

Theorem 7. *[3] The following problem is 2EXPTIME-hard: Given a Markovian stochastic domain D, a labeling function λ, and an* LTL *formula φ, decide if there exists a strategy σ that wins almost surely, i.e., such that $\mu_{D,\sigma}([\varphi]_\lambda) = 1$.*[8]

We will use the fact that *deterministic* complexity classes are closed under complementation, i.e., if $\mathcal{C} \subseteq \mathbf{2}^{\{0,1\}^*}$ is a deterministic complexity class, and $L \in \mathcal{C}$, then also $L^c = \{0,1\}^* \setminus L \in \mathcal{C}$. Moreover, if H is hard for $\mathcal{C}$ then so is H^c. Indeed, for $L \in \mathcal{C}$, also $L^c \in \mathcal{C}$, and so there is a Karp reduction of L^c to H, i.e., a PTIME-function $f : \{0,1\}^* \to \{0,1\}^*$ such that $x \in L^c$ iff $f(x) \in H$. Note that f is also a Karp reduction of L to H^c.

We now state and prove our promised lower bound:

Theorem 8. *The SBE synthesis problem for LTL goals and bounded domains with Markovian support is 2EXPTIME-hard.*

Proof. Let L be a 2EXPTIME (decision) problem. We will show how to reduce L to the SBE synthesis problem.[9] Let H be the decision problem in Theorem 7. Then H^c is also 2EXPTIME-hard. So, there exists a Karp reduction f^+ from L to H, and a Karp reduction f^- from L to H^c. For every input x to L, and $b \in \{-,+\}$, let $f^b(x) = (D^b, \lambda^b, \varphi^b)$ and $D^b = (St^b, Act^b, \iota^b, Pr^b)$. Then (*): $x \in L$ iff there is a strategy σ such that $\mu_{D^+,\sigma}([\varphi^+]_{\lambda^+}) = 1$ iff there is no strategy σ such that $\mu_{D^-,\sigma}([\varphi^-]_{\lambda^-}) = 1$.

From this, we construct a Markovian stochastic domain $D = (St, Act, \iota, Pr)$, a labeling function λ, and an LTL formula φ. We may assume, without loss of generality, that St^+, St^- are disjoint, and Act^+, Act^- are disjoint. Let $\iota, sink$ be new states, and let $plus, minus$ be new actions. Define D, λ, φ as follows:

- $St = St^+ \cup St^- \cup \{\iota, sink\}$;
- $Act = Act^+ \cup Act^- \cup \{plus, minus\}$;
- The initial state is ι.
- The labeling function $\lambda : St \cup Act \to 2^{AP \cup \{plus, minus\}}$ is defined as follows: for $x \in St^b \cup Act^b$ define $\lambda(x) = \lambda^b(x)$, and for $x \in \{plus, minus\}$ define $\lambda(x) = \{x\}$, and for $x \in \{\iota, sink\}$ define $\lambda(x) = \emptyset$.
- Before giving the probability function, we give the support function $\Delta : St \times Act \to 2^{St}$. We set $\Delta(\iota, plus) = \{\iota^+\}$ and $\Delta(\iota, minus) = \{\iota^-\}$. For $s \in St^b, a \in Act^b$ define $\Delta(s,a) = \Delta^b(s,a)$. In all other cases, $\Delta(\cdot,\cdot) = \{sink\}$.
- Now, define the probability as the uniform distribution, i.e., for $s \in St, a \in Act$, if $|\Delta(s,a)| = N$ then define $Pr(s,a)(s') = 1/N$ for every $s' \in \Delta(s,a)$. Clearly D is Markovian.
- Finally, define the formula φ as

$$(\mathsf{G}\,\neg sink) \wedge (plus \to \mathsf{X}\,\varphi^+) \wedge (minus \to \mathsf{X}\,\varphi^-).$$

In words, the initial state of D transitions on action $plus$ (resp. $minus$) with probability 1 to the initial state of D^+ (resp. D^-). Since in our formalism every action must be available from every state, we introduce a sink state to catch all other transitions. The goal φ says that the sink is never reached, and if the first action is $plus$ then the formula φ^+ should hold from the second state onwards, and if the first action is $minus$ then the formula φ^- should hold from the second state onwards.

The reduction then calls the oracle to the SBE problem with this data (D, λ, φ) which returns a finite-state strategy, call it σ_x, that is SBE wrt $\mathfrak{D}, [\varphi]_\lambda$ where $\mathfrak{D}$ is the set of all bounded domains similar to D. The reduction then looks at the first action of σ_x. If it is $plus$ then it returns "$x \in L$", and if it is $minus$ then it returns "$x \notin L$".

We now argue that the reduction is correct. To do this, we will first argue that $\mu_{D',\sigma_x}([\varphi]_\lambda) = 1$ for all $D' \in \mathfrak{D}$.

Claim 1: There exists a strategy δ such that $\mu_{D,\delta}([\varphi]_\lambda) = 1$, i.e., $val_{D,[\varphi]_\lambda}(\delta, \iota) = $ *winning*. The claim follows from (*). Indeed, if $x \in L$ then there is a strategy σ that satisfies φ^+ in D^+ with probability 1, and thus the strategy δ that first does action $plus$ and then follows σ satisfies φ with probability 1. On the other hand, if $x \notin L$ then a symmetric argument holds with D^- replacing D^+ and $minus$ replacing $plus$.

[8] Actually, [3] states the dual problem, i.e., decide if every strategy wins almost surely. Their proof reduces from the membership problem for alternating EXPSPACE Turing machines. It transforms such a TM T and input word x into a Markovian domain D, a labeling function λ, and an LTL formula φ such that the T rejects x iff there exists a strategy σ such that $\mu_{D,\sigma}([\varphi]_\lambda) > 0$. However, in a footnote, they remark that D, λ, φ has the following additional property: there exists a strategy σ such that $\mu_{D,\sigma}([\varphi]_\lambda) > 0$ iff there exists a strategy σ' such that $\mu_{D,\sigma'}([\varphi]_\lambda) = 1$. Thus, their reduction also establishes the 2EXPTIME-hardness of the present problem.

[9] The reduction we use is a Cook reduction since we will be reducing from a decision problem to a search problem that always has a solution, which makes Karp and Levin reductions inappropriate. Our Cook reduction will have a Levin flavor as it calls the oracle only once.

Moreover, optimal strategies for LTL goals can be taken to be finite-state [4]; thus, we can assume that δ is a finite-state.

Now, [17] proved that given similar bounded domains with Markovian support D_1, D_2, an LTL formula ψ, a labeling function λ, and a finite-state strategy σ, the following holds for every σ-history h (and thus for the history $h = \iota$ in particular):

$$val_{D_1,[\psi_\lambda]}(\sigma, h) = val_{D_2,[\psi_\lambda]}(\sigma, h).$$

From this and Claim 1, we can immediately conclude that for $\mu_{D',\delta}([\varphi]_\lambda) = 1$ for every $D' \in \mathfrak{D}$. Thus δ is SBE wrt $\mathfrak{D}, [\varphi]_\lambda$, since it cannot be strictly dominated, and so every SBE strategy has this property, in particular σ_x does. Thus we have shown that $\mu_{D',\sigma_x}([\varphi]_\lambda) = 1$ for all $D' \in \mathfrak{D}$.

Finally, we show that if the first action of σ_x is $plus$ then $x \in L$, and if it is $minus$ then $x \notin L$ (note that it cannot be any other action since this would violate φ with probability 1). We show just the $plus$ case (the $minus$ case is symmetric). If the first action of σ_x is $plus$, then the strategy σ in D^+ that copies what σ_x does after it does $plus$ has the property that $\mu_{D^+,\sigma}([\varphi^+]_{\lambda^+}) = 1$. Thus by (*), $x \in L$. □

Putting the upper and lower bounds together, we get:

Theorem 9. *The SBE synthesis problem for LTL goals and bounded domains with Markovian support is* 2*EXPTIME-complete.*

C. Finitely-generated support

Theorem 9 is easily generalized to the setting of finitely-generated support, as follows. A stochastic domain $D = (St, Act, \iota, Pr)$ has *finitely-generated support* if there exists a finite-state machine $M = (Q, q_I, \tau, \Gamma)$ where Q is a finite set of states, $q_0 \in Q$ is the initial state, $\tau : Q \times St \to Q$ is the transition function, and $\Gamma : Q \times St \times Act \to 2^{St}$ is the output function, such that the following holds. If $h = s_0, a_0, \cdots, s_n$ is a history, and $q_0, q_1, \cdots, q_{n+1}$, which we call the *run of* M *on* h, is determined by $q_0 = q_I$ and $q_{i+1} = \tau(q_i, s_i)$, then $\Delta(h, a) = \Gamma(q_{n+1}, s_n, a)$. We say that the machine M *generates* the support Δ.

Theorem 10. *The following problem is* 2*EXPTIME-complete: given a domain D with finitely-generated support, a labeling function λ, and an LTL formula φ, find a finite-state strategy that is stochastic best-effort wrt $\mathfrak{D}, [\varphi]_\lambda$ where $\mathfrak{D}$ is the set of all bounded domains that are similar to D.*

Proof. For the lower-bound use Theorem 8 and the fact that a Markovian support is finitely generated (i.e., by a machine with $|Q| = 1$).

For the upper bound, we first establish some facts about a simple product construction. Let D be a domain whose support is generated by M, and let λ be a labeling function. Define the product domain (St', Act', ι', Pr') with support Δ', and labeling λ', as follows:

- $St' = Q \times St$;
- $Act' = Act$;
- $\iota' = (q_I, \iota)$;
- $\Delta'(h', a) = \Gamma(\tau(q, s), s, a)$ where $last(h') = (q, s)$ (note that Δ' is Markovian since it only depends on $last(h')$);
- define the probability $Pr'(h', a)$ for $h' = ((q_0, s_0), a_0, (q_1, s_1), a_1, \cdots, (q_n, s_n))$ to be $Pr(h, a)$ where $h = (s_0, a_0, \cdots, s_n)$;
- $\lambda'(q, s) = \lambda(s)$ and $\lambda'(a) = \lambda(a)$.

Call this domain $M \times D$. Every history h in D determines a history h' in $M \times D$ by "adding" the run of M on h, and conversely, every history h' in $M \times D$ determines a history h in D by "removing" the run of M on h. This correspondence extends to plays π, π', strategies σ, σ', and sample spaces $\Omega_\sigma, \Omega_{\sigma'}$. Also: (*) $\mu_{D,\sigma}([\varphi]_\lambda) = \mu_{M\times D,\sigma'}([\varphi]_{\lambda'})$ (because this fact is true of corresponding cones, i.e., $\mu_{D,\sigma}(C_h) = \mu_{M\times D,\sigma'}(C_{h'})$ by the definition of Pr'). Finally, σ' is finite state iff σ is finite-state.

Now, for the upper bound, we are given D, λ, φ where the support of D is finitely-generated, say by M. Let $\mathfrak{D}'$ be the set of all bounded domains that are similar to $M \times D$. Apply the 2EXPTIME algorithm in Theorem 6 to a Markovian domain from $\mathfrak{D}'$, with the labeling λ', and the formula φ, to produce a finite-state strategy σ' that is SBE wrt $\mathfrak{D}', [\varphi]_{\lambda'}$. Then the finite-state strategy σ is SBE wrt $\mathfrak{D}, [\varphi]_\lambda$. Indeed, if $\delta \succ_{\mathfrak{D},[\varphi]_\lambda} \sigma$ then by (*) also $\delta' \succ_{\mathfrak{D}',[\varphi]_\lambda} \sigma'$, contradicting that σ' is maximal. □

VIII. Related Work

In the non-stochastic setting maximal strategies in the dominance order have been studied in the synthesis community, where they are sometimes called "admissible" or "best effort" [7], [8], [9], [10], [11], [12], [13]. A common approach to transfer techniques from the non-stochastic setting to the stochastic setting is to capture the effects of stochasticity by some form of fairness, e.g., [19], [20], [21], [22]. This was the approach taken in [17]. Unfortunately, that work does not extend beyond omega-regular conditions (Remark 3), which are low in the Borel hierarchy, and it is unclear if it can handle strategies that are not finite-state. Thus, in order to fully handle omega-regular conditions in particular, and the entire Borel hierarchy in general, we took a different approach, i.e., the use of geometric measure theory and the Lebesgue Density Theorem.

While [17] proposes a preliminary notion of SBE, the main differences are: it defines its notion of SBE 'locally', in terms of the value of histories, and does not consider any dominance order or notion of maximality; it only considers omega-regular goals while we consider all Borel goals; and it only considers domains with Markovian support while we consider domains with arbitrary support. Finally, the definition and proof techniques there are tailored to the special case they consider. In particular, strategies satisfying their definition may not even exist for non omega-regular goals (see Remark 3 for more details).

Uncertainty in the stochastic setting — usually but not exclusively studied in the context of Markov chains — has been modeled by constraining the probability distributions, e.g., [23] assigns each transition a set of allowed probabilities, and [24] restricts the probabilities by arithmetic constraints. Many later works constrain the probabilities to lie in given intervals and studied algorithmic properties, e.g., [25], [26], [27], [28]: Goals are typically omega-regular, and strategies should provide worst-case guarantees and may not always exist (e.g., for an LTL formula φ and number $p \in [0, 1]$, that the probability is at least p of satisfying φ in every specified domain). In comparison, we consider the general case of not-necessarily Markovian domains where the system has complete ignorance about the values of non-zero probabilities. Our work introduces and studies the dominance order on strategies, as well as the maximal strategies in this order which we show always exist not only for omega-regular goals but for all Borel goals.

IX. Conclusion and Future Work

Ruling out dominated strategies is a classic rationality principle in Decision Theory [6]. We may formulate our problem as a type of decision problem as follows. Consider the tree-unwinding of a finite edge-labeled graph from a start vertex, where each edge is labeled by an action of the agent (this unwinding corresponds to the *support function* in this paper). The agent's goal is given as a set of infinite paths through the tree. The agent's strategies are deterministic, i.e., functions from states to actions. A possible environment is a function from state/action pairs to probability distributions over the immediate-successor states in the tree. The agent does not know which environment it will face. If the agent fixes a strategy then every environment becomes an (infinite-state) Markov chain, and one can take the agent's payoff to be the probability that G is satisfied in this Markov chain. Stochastic best-effort strategies are those that are not dominated in this order. This formulation suggests one might be able to exploit Decision Theory or even Game Theory (for a multi-player variant) and obtain a rich theory of synthesis under the sort of uncertainty/ignorance studied in this paper.

To prove existence of SBE we used properties of the Lebesgue measure, notably the Lebesgue Density Theorem adapted to our probability spaces. In more restricted settings one does not need such machinery. For instance, consider the setting that $\mathfrak{D}$ consists of all bounded domains that are similar to some Markovian domain D and the goal G is omega-regular. In this setting, our characterization theorem can be proven for finite-state strategies using standard automata-theoretic techniques, such as those used in [17]. However, even in this restricted setting we do not see how such techniques can handle all strategies and not just finite-state ones. One direction may be to try and handle other finitely-representable strategies, e.g., pushdown strategies, using automata-theoretic techniques.

Acknowledgments

Partially supported by the Austrian Science Fund (FWF) P 32021; ERC Advanced Grant WhiteMech (No. 834228); EU ICT-48 2020 project TAILOR (No. 952215); PRIN project RIPER (No. 20203FFYLK); PNRR MUR project FAIR (No. PE0000013).

Appendix

For convenience, we include a proof of the Lebesgue Density Theorem that is tailored for our spaces and is perhaps more accessible than most proofs in the literature.

Proof of Theorem 1 (Lebesgue Density Theorem). Write $\mu(-)$ as shorthand for $\mu_{D,\sigma}(-)$. First, observe that if $x_n \leq 1$ for all n, then $\liminf_{n\to\infty} x_n = 1$ iff $\lim_{n\to\infty} x_n$ exists and is equal to 1.[10] Thus, the density of G in D, σ at ω is 1 iff $\liminf_{i\to\infty} \mu(G|C_{\omega[0,i)}) = 1$. For $n \geq 1$, define

$$S_n \doteq \{\omega \in G : \liminf_{i\to\infty} \mu(G|C_{\omega[0,i)}) < (n-1)/n\}.$$

Note that $\cup_n S_n$ consists of the plays in G whose density is not equal to 1. We will show that $\mu(S_n) = 0$ for every n. The result then follows since $\mu(G \setminus X) = \mu(G \cap (\cup S_n)) \leq \mu(\cup S_n) \leq \sum_n \mu(S_n) = 0$ (recall from the statement of the Theorem being proved that X stands for the plays in G whose density is equal to 1).

We first show that each S_n is a Borel set. Note that if $(x_i)_i$ is a sequence of numbers, $\liminf x_i < B$ iff there exists $k \in \mathbb{N}$ and infinitely many i such that $x_i + 1/k < B$.[11] Thus, $\omega \in S_n$ iff $\omega \in G$ and there exists $k \in \mathbb{N}$ such that for every $M \in \mathbb{N}$ there exists $i \geq M$ such that $\mu(G|C_{\omega[0,i)}) + 1/k < (n-1)/n$. Thus, for every $k, i \in \mathbb{N}$ define the set $U_{k,i}$ to be the set of traces ω such that $\mu(G|C_{\omega[0,i)}) + 1/k < (n-1)/n$. Note that $U_{k,i}$ is a union of cones (i.e., C_h where $|h| = i$ and $\mu(G|C_h) + 1/k < (n-1)/n$), and

$$S_n = G \cap (\cup_{k\in\mathbb{N}} \cap_{M\in\mathbb{N}} \cup_{i\geq M} U_{k,i}),$$

and thus S_n is Borel (recall that by our assumption G is a goal, i.e., a Borel set).

Next, we show that $\mu(S_n) = 0$ for all n. We introduce some helpful terminology. A cone-covering $(U_i)_i$ of Y is called:

- *disjoint* if $U_i \cap U_j = \emptyset$ for every $i \neq j$.
- *n-tight* if $\mu(\cup_i U_i) < \mu(Y)n/(n-1)$.
- *n-sparse* if $\mu(Y|U_i) < (n-1)/n$ for every i.

Suppose, towards a contradiction, that $\mu(S_n) > 0$ for some n. We now find a disjoint n-tight n-sparse covering of S_n. By Lemma 1, there exists an n-tight cone-covering $(U_i)_i$ of S_n. By definition of cover, every $\omega \in S_n$ is in some U_i. By definition of S_n, we can take a cone $B_\omega \subseteq U_i$ containing ω such that $\mu(S_n|B_\omega) < (n-1)/n$. Then $(B_\omega)_{\omega\in S_n}$ is cone-covering of S_n which is n-sparse and n-tight since $\cup_{\omega\in S_n} B_\omega \subseteq \cup_i U_i$. However, it need not be disjoint. So, refine it by first removing repetitions, and then simultaneously

[10] Indeed: use the definition of limit, i.e., if the liminf is equal to the limsup then the limit exists and is this common value; conversely, use the fact that the liminf is at most the limsup, which itself is at most 1 by the assumption.

[11] Indeed: if the right-hand side holds then there is some k and infinitely many i such that $\inf_{m\geq i} x_m < B - 1/k$, and thus $\liminf x_i \leq B - 1/k < B$; if the right-hand side does not hold, then for every k there exists m such that for every $i \geq m$ we have $x_i \geq B - 1/k$, and thus for every k there is an m such that $\inf_{i\geq m} x_i \geq B - 1/k$, and thus for every k we have that $\liminf x_i \geq B - 1/k$, and thus $\liminf x_i \geq B$.

removing every cone that is a subset of some other cone in the sequence. This results in a disjoint n-sparse n-tight cone-covering $(V_i)_i$ of S_n (it is n-tight by monotonicity of μ). Disjointness means that $\mu(\cup_i V_i) = \sum_i \mu(V_i)$ (by countable additivity). Putting this together:

$$\begin{aligned}\mu(S_n) &= \mu(S_n \cap (\cup_i V_i)) = \mu(\cup_i (S_n \cap V_i)) \\ &\leq \sum_i \mu(S_n \cap V_i) = \sum_i \mu(S_n | V_i)\mu(V_i) \\ &< ((n-1)/n) \sum_i \mu(V_i) \\ &= ((n-1)/n)\mu(\cup_i V_i) < \mu(S_n)\end{aligned}$$

which is impossible since the inequality is strict. □

References

[1] A. Pnueli and R. Rosner, "On the synthesis of a reactive module," in *POPL*, 1989, pp. 179–190.

[2] A. Bianco and L. de Alfaro, "Model checking of probabilistic and nondeterministic systems," in *FSTTCS*, 1995, pp. 499–513.

[3] C. Courcoubetis and M. Yannakakis, "The complexity of probabilistic verification," *J. ACM*, vol. 42, no. 4, pp. 857–907, 1995.

[4] C. Baier, J.-P. Katoen, and K. Guldstrand Larsen, *Principles of Model Checking*. The MIT Press, 2008.

[5] B. Aminof, G. De Giacomo, A. Murano, and S. Rubin, "Planning under LTL environment specifications," in *ICAPS*, 2019.

[6] M. D. Resnik, *Choices: An introduction to decision theory*. U of Minnesota Press, 1987.

[7] D. Berwanger, "Admissibility in infinite games," in *STACS*, 2007, pp. 188–199.

[8] M. Faella, "Admissible strategies in infinite games over graphs," in *MFCS*, 2009, pp. 307–318.

[9] R. Brenguier, J. Raskin, and M. Sassolas, "The complexity of admissibility in omega-regular games," in *CSL-LICS*, 2014, pp. 23:1–23:10.

[10] R. Brenguier, J. Raskin, and O. Sankur, "Assume-admissible synthesis," *Acta Inf.*, vol. 54, no. 1, 2017.

[11] B. Aminof, G. De Giacomo, A. Lomuscio, A. Murano, and S. Rubin, "Synthesizing strategies under expected and exceptional environment behaviors," in *IJCAI*, 2020, pp. 1674–1680.

[12] ——, "Synthesizing best-effort strategies under multiple environment specifications," in *KR*, 2021, pp. 42–51.

[13] B. Aminof, G. De Giacomo, and S. Rubin, "Best-effort synthesis: Doing your best is not harder than giving up," in *IJCAI*, 2021, pp. 1766–1772.

[14] R. Bhattacharya and E. C. Waymire, *Random Walk, Brownian Motion, and Martingales*. Springer, 2021.

[15] A. E. Taylor, *General theory of functions and integration*. Courier Corporation, 1985.

[16] A. Andretta and R. Camerlo, "The descriptive set theory of the lebesgue density theorem," *Advances in Mathematics*, vol. 234, pp. 1–42, 2013.

[17] B. Aminof, G. D. Giacomo, S. Rubin, and F. Zuleger, "Beyond strong-cyclic: Doing your best in stochastic environments," in *IJCAI*, 2022, pp. 2525–2531.

[18] M. Y. Vardi, "Automatic verification of probabilistic concurrent finite-state programs," in *FOCS*, 1985, pp. 327–338.

[19] A. Pnueli, "On the extremely fair treatment of probabilistic algorithms," in *STOC*, 1983.

[20] C. Courcoubetis, M. Y. Vardi, and P. Wolper, "Reasoning about fair concurrent programs," in *STOC*. ACM, 1986, pp. 283–294.

[21] C. Baier and M. Z. Kwiatkowska, "On the verification of qualitative properties of probabilistic processes under fairness constraints," *Inf. Process. Lett.*, vol. 66, no. 2, pp. 71–79, 1998.

[22] B. Aminof, G. De Giacomo, and S. Rubin, "Stochastic fairness and language-theoretic fairness in planning on nondeterministic domains," in *ICAPS*, 2020.

[23] B. Jonsson and K. G. Larsen, "Specification and refinement of probabilistic processes," in *LICS*, 1991, pp. 266–277.

[24] A. Aziz, V. Singhal, and F. Balarin, "It usually works: The temporal logic of stochastic systems," in *CAV*, 1995, pp. 155–165.

[25] I. Kozine and L. V. Utkin, "Interval-valued finite markov chains," *Reliab. Comput.*, vol. 8, no. 2, pp. 97–113, 2002.

[26] K. Sen, M. Viswanathan, and G. Agha, "Model-checking markov chains in the presence of uncertainties," in *TACAS*, 2006, pp. 394–410.

[27] E. M. Wolff, U. Topcu, and R. M. Murray, "Robust control of uncertain markov decision processes with temporal logic specifications," in *CDC*. IEEE, 2012, pp. 3372–3379.

[28] T. Chen, T. Han, and M. Z. Kwiatkowska, "On the complexity of model checking interval-valued discrete time markov chains," *Inf. Process. Lett.*, vol. 113, no. 7, pp. 210–216, 2013.