

## ALGORITHMS FOR GAME METRICS\*

KRISHNENDU CHATTERJEE<sup>a</sup>, LUCA DE ALFARO<sup>b</sup>, RUPAK MAJUMDAR<sup>c</sup>,  
AND VISHWANATH RAMAN<sup>d</sup>

<sup>a</sup> IST Austria (Institute of Science and Technology Austria)  
*e-mail address:* krish.chat@ist.ac.at

<sup>b,d</sup> Computer Science Department, University of California, Santa Cruz  
*e-mail address:* {luca,vishwa}@soe.ucsc.edu

<sup>c</sup> Department of Computer Science, University of California, Los Angeles  
*e-mail address:* rupak@cs.ucla.edu

**ABSTRACT.** Simulation and bisimulation metrics for stochastic systems provide a quantitative generalization of the classical simulation and bisimulation relations. These metrics capture the similarity of states with respect to quantitative specifications written in the quantitative  $\mu$ -calculus and related probabilistic logics. We first show that the metrics provide a bound for the difference in *long-run average* and *discounted average* behavior across states, indicating that the metrics can be used both in system verification, and in performance evaluation. For turn-based games and MDPs, we provide a polynomial-time algorithm for the computation of the one-step metric distance between states. The algorithm is based on linear programming; it improves on the previous known exponential-time algorithm based on a reduction to the theory of reals. We then present PSPACE algorithms for both the decision problem and the problem of approximating the metric distance between two states, matching the best known algorithms for Markov chains. For the bisimulation kernel of the metric our algorithm works in time  $\mathcal{O}(n^4)$  for both turn-based games and MDPs; improving the previously best known  $\mathcal{O}(n^9 \cdot \log(n))$  time algorithm for MDPs.

For a concurrent game  $G$ , we show that computing the exact distance between states is at least as hard as computing the value of concurrent reachability games and the square-root-sum problem in computational geometry. We show that checking whether the metric distance is bounded by a rational  $r$ , can be done via a reduction to the theory of real closed fields, involving a formula with three quantifier alternations, yielding  $\mathcal{O}(|G|^{\mathcal{O}(|G|^5)})$  time complexity, improving the previously known reduction, which yielded  $\mathcal{O}(|G|^{\mathcal{O}(|G|^7)})$  time complexity. These algorithms can be iterated to approximate the metrics using binary search.

*1998 ACM Subject Classification:* F.4.1, F.1.1, F.1.1.

*Key words and phrases:* game semantics, minimax theorem, metrics,  $\omega$ -regular properties, quantitative  $\mu$ -calculus, probabilistic choice, equivalence of states, refinement of states.

\* A preliminary version of this paper titled “Algorithms for Game Metrics” appeared in the IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, December 2008. This is a version with full proofs and extensions.

## 1. INTRODUCTION

System metrics constitute a quantitative generalization of system relations. The bisimulation relation captures state *equivalence*: two states  $s$  and  $t$  are bisimilar if and only if they cannot be distinguished by any formula of the  $\mu$ -calculus [5]. The bisimulation *metric* captures the *degree of difference* between two states: the bisimulation distance between  $s$  and  $t$  is a real number that provides a tight bound for the difference in value of formulas of the *quantitative*  $\mu$ -calculus at  $s$  and  $t$  [12]. A similar connection holds between the simulation relation and the simulation metric.

The classical system relations are a basic tool in the study of *boolean* properties of systems, that is, the properties that yield a truth value. As an example, if a state  $s$  of a transition system can reach a set of target states  $R$ , written  $s \models \diamond R$  in temporal logic, and  $t$  can simulate  $s$ , then we can conclude  $t \models \diamond R$ . System metrics play a similarly fundamental role in the study of the quantitative behavior of systems. As an example, if a state  $s$  of a Markov chain can reach a set of target states  $R$  with probability 0.8, written  $s \models \mathbb{P}_{\geq 0.8} \diamond R$ , and if the metric simulation distance from  $t$  to  $s$  is 0.3, then we can conclude  $t \models \mathbb{P}_{\geq 0.5} \diamond R$ . The simulation relation is at the basis of the notions of system refinement and implementation, where qualitative properties are concerned. In analogous fashion, simulation metrics provide a notion of approximate refinement and implementation for quantitative properties.

We consider three classes of systems:

- *Markov decision processes*. In these systems there is one player. At each state, the player can choose a move; the current state and the move determine a probability distribution over the successor states.
- *Turn-based games*. In these systems there are two players. At each state, only one of the two players can choose a move; the current state and the move determine a probability distribution over the successor states.
- *Concurrent games*. In these systems there are two players. At each state, both players choose moves simultaneously and independently; the current state and the chosen moves determine a probability distribution over the successor states.

System metrics were first studied for Markov chains and Markov decision processes (MDPs) [12, 32, 33, 13, 14], and they have recently been extended to two-player turn-based and concurrent games [11]. The fundamental property of the metrics is that they provide a tight bound for the difference in value that formulas belonging to quantitative specification languages assume at the states of a system. More precisely, let  $q\mu$  indicate the *quantitative*  $\mu$ -calculus, a specification language in which many of the classical specification properties, including reachability and safety properties, can be written [10]. The metric bisimulation distance between two states  $s$  and  $t$ , denoted  $[s \simeq_g t]$ , has the property that  $[s \simeq_g t] = \sup_{\varphi \in q\mu} |\varphi(s) - \varphi(t)|$ , where  $\varphi(s)$  and  $\varphi(t)$  are the values  $\varphi$  assumes at  $s$  and  $t$ . To each metric is associated a *kernel*: the kernel of a metric  $d$  is the relation that relates the pairs of states that have distance 0; to each metric corresponds a metric kernel relation. The kernel of the simulation metric is *probabilistic simulation*; the kernel of the bisimulation metric is *probabilistic bisimulation* [27].

**Metric as bound for discounted and long-run average payoff.** Our first result is that the metrics developed in [11] provide a bound for the difference in *long-run average* and *discounted average* properties across states of a system. These average rewards play a central role in the theory of stochastic games, and in its applications to optimal control and

economics [4, 17]. Thus, the metrics of [11] are useful both for system verification, and for performance evaluation, supporting our belief that they constitute the canonical metrics for the study of the similarity of states in a game. We point out that it is possible to define a discounted version  $[\simeq_g]^\alpha$  of the game bisimulation metric; however, we show that this discounted metric does *not* provide a bound for the difference in discounted values.

**Algorithmic results.** Next, we investigate algorithms for the computation of the metrics. The metrics can be computed in iterative fashion, following the inductive way in which they are defined. A metric  $d$  can be computed as the limit of a monotonically increasing sequence of approximations  $d_0, d_1, d_2, \dots$ , where  $d_0(s, t)$  is the difference in value that variables can have at states  $s$  and  $t$ . For  $k \geq 0$ ,  $d_{k+1}$  is obtained from  $d_k$  via  $d_{k+1} = H(d_k)$ , where the operator  $H$  depends on the metric (bisimulation, or simulation), and on the type of system. Our main results are as follows:

- (1) *Metrics for turn-based games and MDPs.* We show that for turn-based games, and MDPs, the one-step metric operator  $H$  for both bisimulation and simulation can be computed in polynomial time, via a reduction to linear programming (LP). The only previously known algorithm, which can be inferred from [11], had EXPTIME complexity and relied on a reduction to the theory of real closed fields; the algorithm thus had more a complexity-theoretic, than a practical, value. The key step in obtaining our polynomial-time algorithm consists in transforming the original sup-inf *non-linear* optimization problem (which required the theory of reals) into a quadratic-size inf *linear* optimization problem that can be solved via LP. We then present PSPACE algorithms for both the decision problem of the metric distance between two states and for the problem of computing the approximate metric distance between two states for turn-based games and MDPs. Our algorithms match the complexity of the best known algorithms for the sub-class of Markov chains [31].
- (2) *Metrics for concurrent games.* For concurrent games, our algorithms for the  $H$  operator still rely on decision procedures for the theory of real closed fields, leading to an EXPTIME procedure. However, the algorithms that could be inferred from [11] had time-complexity  $\mathcal{O}(|G|^{\mathcal{O}(|G|^7)})$ , where  $|G|$  is the size of a game; we improve this result by presenting algorithms with  $\mathcal{O}(|G|^{\mathcal{O}(|G|^5)})$  time-complexity.
- (3) *Hardness of metric computation in concurrent games.* We show that computing the exact distance of states of concurrent games is at least as hard as computing the value of concurrent reachability games [15, 8], which is known to be at least as hard as solving the square-root-sum problem in computational geometry [18]. These two problems are known to lie in PSPACE, and have resisted many attempts to show that they are in NP.
- (4) *Kernel of the metrics.* We present polynomial time algorithms to compute the simulation and bisimulation kernel of the metrics for turn-based games and MDPs. Our algorithm for the bisimulation kernel of the metric runs in time  $\mathcal{O}(n^4)$  (assuming a constant number of moves) as compared to the previous known  $\mathcal{O}(n^9 \cdot \log(n))$  algorithm of [35] for MDPs, where  $n$  is the size of the state space. For concurrent games the simulation and the bisimulation kernel can be computed in time  $\mathcal{O}(|G|^{\mathcal{O}(|G|^3)})$ , where  $|G|$  is the size of a game.

Our formulation of probabilistic simulation and bisimulation differs from the one previously considered for MDPs in [2]: there, the names of moves (called “labels”) must be preserved by simulation and bisimulation, so that a move from a state has at most one

candidate simulator move at another state. Our problem for MDPs is closer to the one considered in [35], where labels must be preserved, but where a label can be associated with multiple probability distributions (moves).

For turn-based games and MDPs, the algorithms for probabilistic simulation and bisimulation can be obtained from the LP algorithms that yield the metrics. For probabilistic simulation, the algorithm we obtain coincides with the algorithm previously published in [35]. The algorithm requires the solution of feasibility-LP problems with a number of variables and inequalities that is quadratic in the size of the system. For probabilistic bisimulation, we are able to improve on this result by providing an algorithm that requires the solution of feasibility-LP problems that have linearly many variables and constraints. Precisely, as for ordinary bisimulation, the kernel is computed via iterative refinement of a partition of the state space [24]. Given two states that belong to the same partition, to decide whether the states need to be split in the next partition-refinement step, we present an algorithm that requires the solution of a feasibility-LP problem with a number of variables equal to the number of moves available at the states, and number of constraints linear in the number of equivalence classes. Overall, our algorithm for bisimulation runs in time  $\mathcal{O}(n^4)$  (assuming a constant number of moves), considerably improving the  $\mathcal{O}(n^9 \cdot \log(n))$  algorithm of [35] for MDPs, and providing for the first time a polynomial algorithm for turn-based games.

## 2. DEFINITIONS

**Valuations.** Let  $[\theta_1, \theta_2] \subseteq \mathbb{R}$  be a fixed, non-singleton real interval. Given a set of states  $S$ , a *valuation over  $S$*  is a function  $f : S \mapsto [\theta_1, \theta_2]$  associating with every state  $s \in S$  a value  $\theta_1 \leq f(s) \leq \theta_2$ ; we let  $\mathcal{F}$  be the set of all valuations. For  $c \in [\theta_1, \theta_2]$ , we denote by  $\mathbf{c}$  the constant valuation such that  $\mathbf{c}(s) = c$  at all  $s \in S$ . We order valuations pointwise: for  $f, g \in \mathcal{F}$ , we write  $f \leq g$  iff  $f(s) \leq g(s)$  at all  $s \in S$ ; we remark that  $\mathcal{F}$ , under  $\leq$ , forms a lattice. Given  $a, b \in \mathbb{R}$ , we write  $a \sqcup b = \max\{a, b\}$ , and  $a \sqcap b = \min\{a, b\}$ ; we also let  $a \oplus b = \min\{1, \max\{0, a + b\}\}$  and  $a \ominus b = \max\{0, \min\{1, a - b\}\}$ . We extend  $\sqcap, \sqcup, +, -, \oplus, \ominus$  to valuations by interpreting them in pointwise fashion.

**Game structures.** For a finite set  $A$ , let  $\text{Dist}(A)$  denote the set of probability distributions over  $A$ . We say that  $p \in \text{Dist}(A)$  is *deterministic* if there is  $a \in A$  such that  $p(a) = 1$ . We assume a fixed finite set  $\mathcal{V}$  of *observation variables*.

A (two-player, concurrent) *game structure*  $G = \langle S, [\cdot], \text{Moves}, \Gamma_1, \Gamma_2, \delta \rangle$  consists of the following components [1, 8]:

- A finite set  $S$  of states.
- A variable interpretation  $[\cdot] : \mathcal{V} \mapsto [\theta_1, \theta_2]^S$ , which associates with each variable  $v \in \mathcal{V}$  a valuation  $[v]$ .
- A finite set  $\text{Moves}$  of moves.
- Two move assignments  $\Gamma_1, \Gamma_2: S \mapsto 2^{\text{Moves}} \setminus \{\emptyset\}$ . For  $i \in \{1, 2\}$ , the assignment  $\Gamma_i$  associates with each state  $s \in S$  the nonempty set  $\Gamma_i(s) \subseteq \text{Moves}$  of moves available to player  $i$  at state  $s$ .
- A probabilistic transition function  $\delta: S \times \text{Moves} \times \text{Moves} \mapsto \text{Dist}(S)$ , that gives the probability  $\delta(s, a_1, a_2)(t)$  of a transition from  $s$  to  $t$  when player 1 plays move  $a_1$  and player 2 plays move  $a_2$ .

At every state  $s \in S$ , player 1 chooses a move  $a_1 \in \Gamma_1(s)$ , and simultaneously and independently player 2 chooses a move  $a_2 \in \Gamma_2(s)$ . The game then proceeds to the successor state  $t \in S$  with probability  $\delta(s, a_1, a_2)(t)$ . We let  $\text{Dest}(s, a_1, a_2) = \{t \in S \mid \delta(s, a_1, a_2)(t) > 0\}$ . The *propositional distance*  $p(s, t)$  between two states  $s, t \in S$  is the maximum difference in the valuation of any variable:

$$p(s, t) = \max_{v \in \mathcal{V}} |[v](s) - [v](t)| .$$

The kernel of the propositional distance induces an equivalence on states: for states  $s, t$ , we let  $s \equiv t$  if  $p(s, t) = 0$ . In the following, unless otherwise noted, the definitions refer to a game structure with components  $G = \langle S, [\cdot], \text{Moves}, \Gamma_1, \Gamma_2, \delta \rangle$ . We indicate the opponent of a player  $i \in \{1, 2\}$  by  $\sim i = 3 - i$ . We consider the following subclasses of game structures.

**Turn-based game structures.** A game structure  $G$  is *turn-based* if we can write  $S = S_1 \cup S_2$  with  $S_1 \cap S_2 = \emptyset$  where  $s \in S_1$  implies  $|\Gamma_2(s)| = 1$ , and  $s \in S_2$  implies  $|\Gamma_1(s)| = 1$ , and further, there exists a special variable  $\text{turn} \in \mathcal{V}$ , such that  $[\text{turn}]s = \theta_1$  iff  $s \in S_1$ , and  $[\text{turn}]s = \theta_2$  iff  $s \in S_2$ .

**Markov decision processes.** For  $i \in \{1, 2\}$ , we say that a structure is an  $i$ -MDP if  $\forall s \in S, |\Gamma_{\sim i}(s)| = 1$ . For MDPs, we omit the (single) move of the player without a choice of moves, and write  $\delta(s, a)$  for the transition function.

**Moves and strategies.** A *mixed move* is a probability distribution over the moves available to a player at a state. We denote by  $\mathcal{D}_i(s) \subseteq \text{Dist}(\text{Moves})$  the set of mixed moves available to player  $i \in \{1, 2\}$  at  $s \in S$ , where:

$$\mathcal{D}_i(s) = \{\mathcal{D} \in \text{Dist}(\text{Moves}) \mid \mathcal{D}(a) > 0 \text{ implies } a \in \Gamma_i(s)\} .$$

The moves in  $\text{Moves}$  are called *pure moves*. We extend the transition function to mixed moves by defining, for  $s \in S$  and  $x_1 \in \mathcal{D}_1(s)$ ,  $x_2 \in \mathcal{D}_2(s)$ ,

$$\delta(s, x_1, x_2)(t) = \sum_{a_1 \in \Gamma_1(s)} \sum_{a_2 \in \Gamma_2(s)} \delta(s, a_1, a_2)(t) \cdot x_1(a_1) \cdot x_2(a_2) .$$

A *path*  $\sigma$  of  $G$  is an infinite sequence  $s_0, s_1, s_2, \dots$  of states in  $S$ , such that for all  $k \geq 0$ , there exist moves  $a_1^k \in \Gamma_1(s_k)$  and  $a_2^k \in \Gamma_2(s_k)$  with  $\delta(s_k, a_1^k, a_2^k)(s_{k+1}) > 0$ . We write  $\Sigma$  for the set of all paths, and  $\Sigma_s$  for the set of all paths starting from state  $s$ .

A *strategy* for player  $i \in \{1, 2\}$  is a function  $\pi_i : S^+ \mapsto \text{Dist}(\text{Moves})$  that associates with every non-empty finite sequence  $\sigma \in S^+$  of states, representing the history of the game, a probability distribution  $\pi_i(\sigma)$ , which is used to select the next move of player  $i$ ; we require that for all  $\sigma \in S^*$  and states  $s \in S$ , if  $\pi_i(\sigma s)(a) > 0$ , then  $a \in \Gamma_i(s)$ . We write  $\Pi_i$  for the set of strategies for player  $i$ . Once the starting state  $s$  and the strategies  $\pi_1$  and  $\pi_2$  for the two players have been chosen, the game is reduced to an ordinary stochastic process, denoted  $G_s^{\pi_1, \pi_2}$ , which defines a probability distribution on the set  $\Sigma$  of paths. We denote by  $\text{Pr}_s^{\pi_1, \pi_2}(\cdot)$  the probability of a measurable event (sets of paths) with respect to this process, and denote by  $\mathbb{E}_s^{\pi_1, \pi_2}(\cdot)$  the associated expectation operator. For  $k \geq 0$ , we let  $X_k : \Sigma \rightarrow S$  be the random variable denoting the  $k$ -th state along a path.

**One-step expectations and predecessor operators.** Given a valuation  $f \in \mathcal{F}$ , a state  $s \in S$ , and two mixed moves  $x_1 \in \mathcal{D}_1(s)$  and  $x_2 \in \mathcal{D}_2(s)$ , we define the expectation of  $f$  from  $s$  under  $x_1, x_2$  by,

$$\mathbb{E}_s^{x_1, x_2}(f) = \sum_{t \in S} \delta(s, x_1, x_2)(t) f(t) .$$

For a game structure  $G$ , for  $i \in \{1, 2\}$  we define the *valuation transformer*  $\text{Pre}_i : \mathcal{F} \mapsto \mathcal{F}$  by, for all  $f \in \mathcal{F}$  and  $s \in S$  as,

$$\text{Pre}_i(f)(s) = \sup_{x_i \in \mathcal{D}_i(s)} \inf_{x_{\sim i} \in \mathcal{D}_{\sim i}(s)} \mathbb{E}_s^{x_i, x_{\sim i}}(f).$$

Intuitively,  $\text{Pre}_i(f)(s)$  is the maximal expectation player  $i$  can achieve of  $f$  after one step from  $s$ : this is the standard “one-day” or “next-stage” operator of the theory of repeated games [17].

**2.1. Quantitative  $\mu$ -calculus.** We consider the set of properties expressed by the *quantitative  $\mu$ -calculus* ( $q\mu$ ). As discussed in [20, 10, 22], a large set of properties can be encoded in  $q\mu$ , spanning from basic properties such as maximal reachability and safety probability, to the maximal probability of satisfying a general  $\omega$ -regular specification.

**Syntax.** The syntax of quantitative  $\mu$ -calculus is defined with respect to the set of observation variables  $\mathcal{V}$  as well as a set  $MVars$  of *calculus variables*, which are distinct from the observation variables in  $\mathcal{V}$ . The syntax is given as follows:

$$\varphi ::= c \mid v \mid V \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \oplus c \mid \varphi \ominus c \mid \text{pre}_1(\varphi) \mid \text{pre}_2(\varphi) \mid \mu V. \varphi \mid \nu V. \varphi$$

for constants  $c \in [\theta_1, \theta_2]$ , observation variables  $v \in \mathcal{V}$ , and calculus variables  $V \in MVars$ . In the formulas  $\mu V. \varphi$  and  $\nu V. \varphi$ , we furthermore require that all occurrences of the bound variable  $V$  in  $\varphi$  occur in the scope of an even number of occurrences of the complement operator  $\neg$ . A formula  $\varphi$  is *closed* if every calculus variable  $V$  in  $\varphi$  occurs in the scope of a quantifier  $\mu V$  or  $\nu V$ . From now on, with abuse of notation, we denote by  $q\mu$  the set of closed formulas of  $q\mu$ . A formula is a *player  $i$  formula*, for  $i \in \{1, 2\}$ , if  $\varphi$  does not contain the  $\text{pre}_{\sim i}$  operator; we denote with  $q\mu_i$  the syntactic subset of  $q\mu$  consisting only of closed player  $i$  formulas. A formula is in *positive form* if the negation appears only in front of constants and observation variables, i.e., in the context  $\neg c$  and  $\neg v$ ; we denote with  $q\mu^+$  and  $q\mu_i^+$  the subsets of  $q\mu$  and  $q\mu_i$  consisting only of positive formulas.

We remark that the fixpoint operators  $\mu$  and  $\nu$  will not be needed to achieve our results on the logical characterization of game relations. They have been included in the calculus because they allow the expression of many interesting properties, such as safety, reachability, and in general,  $\omega$ -regular properties. The operators  $\oplus$  and  $\ominus$ , on the other hand, are necessary for our results.

**Semantics.** A variable valuation  $\xi: MVars \mapsto \mathcal{F}$  is a function that maps every variable  $V \in MVars$  to a valuation in  $\mathcal{F}$ . We write  $\xi[V \mapsto f]$  for the valuation that agrees with  $\xi$  on all variables, except that  $V$  is mapped to  $f$ . Given a game structure  $G$  and a variable valuation  $\xi$ , every formula  $\varphi$  of the quantitative  $\mu$ -calculus defines a valuation  $\llbracket \varphi \rrbracket_\xi^G \in \mathcal{F}$  (the superscript  $G$  is omitted if the game structure is clear from the context):

$$\begin{aligned} \llbracket c \rrbracket_\xi &= \mathbf{c} & \llbracket v \rrbracket_\xi &= [v] \\ \llbracket V \rrbracket_\xi &= \xi(V) & \llbracket \neg\varphi \rrbracket_\xi &= \mathbf{1} - \llbracket \varphi \rrbracket_\xi \\ \llbracket \varphi \left\{ \begin{smallmatrix} \oplus \\ \ominus \end{smallmatrix} \right\} c \rrbracket_\xi &= \llbracket \varphi \rrbracket_\xi \left\{ \begin{smallmatrix} \oplus \\ \ominus \end{smallmatrix} \right\} \mathbf{c} & \llbracket \varphi_1 \left\{ \begin{smallmatrix} \vee \\ \wedge \end{smallmatrix} \right\} \varphi_2 \rrbracket_\xi &= \llbracket \varphi_1 \rrbracket_\xi \left\{ \begin{smallmatrix} \cup \\ \cap \end{smallmatrix} \right\} \llbracket \varphi_2 \rrbracket_\xi \\ \llbracket \text{pre}_i(\varphi) \rrbracket_\xi &= \text{Pre}_i(\llbracket \varphi \rrbracket_\xi) & \llbracket \left\{ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right\} V. \varphi \rrbracket_\xi &= \left\{ \begin{smallmatrix} \inf \\ \sup \end{smallmatrix} \right\} \{ f \in \mathcal{F} \mid f = \llbracket \varphi \rrbracket_{\xi[V \mapsto f]} \} \end{aligned}$$

where  $i \in \{1, 2\}$ . The existence of the fixpoints is guaranteed by the monotonicity and continuity of all operators and can be computed by Picard iteration [10]. If  $\varphi$  is closed,  $\llbracket \varphi \rrbracket_\xi$  is independent of  $\xi$ , and we write simply  $\llbracket \varphi \rrbracket$ .

**Discounted quantitative  $\mu$ -calculus.** A *discounted* version of the  $\mu$ -calculus was introduced in [9]; we call this  $d\mu$ . Let  $\Lambda$  be a finite set of discount parameters that take values in the interval  $[0, 1)$ . The discounted  $\mu$ -calculus extends  $q\mu$  by introducing discounted versions of the player pre modalities. The syntax replaces  $\text{pre}_i(\varphi)$  for player  $i \in \{1, 2\}$  with its discounted variant,  $\lambda \cdot \text{pre}_i(\varphi)$ , where  $\lambda \in \Lambda$  is a discount factor that discounts one-step valuations. Negation in the calculus is defined as  $\neg(\lambda \cdot \text{pre}_1(\varphi)) = (1 - \lambda) + \lambda \cdot \text{pre}_2(\neg\varphi)$ . This leads to two additional pre-modalities for the players,  $(1 - \lambda) + \lambda \cdot \text{pre}_i(\varphi)$ .

**2.2. Game bisimulation and simulation metrics.** A *directed metric* is a function  $d : S^2 \mapsto \mathbb{R}_{\geq 0}$  which satisfies  $d(s, s) = 0$  and the *triangle inequality*  $d(s, t) \leq d(s, u) + d(u, t)$  for all  $s, t, u \in S$ . We denote by  $\mathcal{M} \subseteq S^2 \mapsto \mathbb{R}$  the space of all directed metrics; this space, ordered pointwise, forms a lattice which we indicate with  $(\mathcal{M}, \leq)$ . Since  $d(s, t)$  may be zero for  $s \neq t$ , these functions are *pseudo-metrics* as per prevailing terminology [32]. In the following, we omit “directed” and simply say metric when the context is clear. For a metric  $d$ , we indicate with  $C(d)$  the set of valuations  $k \in \mathcal{F}$  where  $k(s) - k(t) \leq d(s, t)$  for every  $s, t \in S$ . A metric transformer  $H_{\preceq_1} : \mathcal{M} \mapsto \mathcal{M}$  is defined as follows, for all  $d \in \mathcal{M}$  and  $s, t \in S$ :

$$H_{\preceq_1}(d)(s, t) = p(s, t) \sqcup \sup_{k \in C(d)} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)) . \quad (2.1)$$

The *player 1 game simulation metric*  $[\preceq_1]$  is the least fixpoint of  $H_{\preceq_1}$ ; the *game bisimulation metric*  $[\simeq_1]$  is the least symmetrical fixpoint of  $H_{\preceq_1}$  and is defined as follows, for all  $d \in \mathcal{M}$  and  $s, t \in S$ :

$$H_{\simeq_1}(d)(s, t) = H_{\preceq_1}(d)(s, t) \sqcup H_{\preceq_1}(d)(t, s) . \quad (2.2)$$

The operator  $H_{\preceq_1}$  is monotonic, non-decreasing and continuous in the lattice  $(\mathcal{M}, \leq)$ . We can therefore compute  $H_{\preceq_1}$  using Picard iteration; we denote by  $[\preceq_1^n] = H_{\preceq_1}^n(\mathbf{0})$  the  $n$ -iterate of this. From the determinacy of concurrent games with respect to  $\omega$ -regular goals [21], we have that the game bisimulation metric is *reciprocal*, in that  $[\simeq_1] = [\simeq_2]$ ; we will thus simply write  $[\simeq_g]$ . Similarly, for all  $s, t \in S$  we have  $[s \preceq_1 t] = [t \preceq_2 s]$ .

The main result in [11] about these metrics is that they are logically characterized by the quantitative  $\mu$ -calculus of [10]. We omit the formal definition of the syntax and semantics of the quantitative  $\mu$ -calculus; we refer the reader to [10] for details. Given a game structure  $G$ , every closed formula  $\varphi$  of the quantitative  $\mu$ -calculus defines a valuation  $\llbracket \varphi \rrbracket \in \mathcal{F}$ . Let  $q\mu$  (respectively,  $q\mu_1^+$ ) consist of all quantitative  $\mu$ -calculus formulas (respectively, all quantitative  $\mu$ -calculus formulas with only the  $\text{Pre}_1$  operator and all negations before atomic propositions). The result of [11] shows that for all states  $s, t \in S$ ,

$$[s \preceq_1 t] = \sup_{\varphi \in q\mu_1^+} (\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)) \quad [s \simeq_g t] = \sup_{\varphi \in q\mu} |\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| . \quad (2.3)$$

**Metrics for the discounted quantitative  $\mu$ -calculus.** We call  $d\mu^\alpha$  the discounted  $\mu$ -calculus with all discount parameters  $\leq \alpha$ . We define the discounted metrics via an  $\alpha$ -discounted metric transformer  $H_{\preceq_1}^\alpha : \mathcal{M} \mapsto \mathcal{M}$ , defined for all  $d \in \mathcal{M}$  and all  $s, t \in S$  by:

$$H_{\preceq_1}^\alpha(d)(s, t) = p(s, t) \sqcup \alpha \cdot \sup_{k \in C(d)} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)) . \quad (2.4)$$

Again,  $H_{\preceq_1}^\alpha$  is continuous and monotonic in the lattice  $(\mathcal{M}, \leq)$ . The  $\alpha$ -discounted *simulation metric*  $[\preceq_1]^\alpha$  is the least fixpoint of  $H_{\preceq_1}^\alpha$ , and the  $\alpha$ -discounted *bisimulation metric*  $[\simeq_1]^\alpha$  is the least symmetrical fixpoint of  $H_{\preceq_1}^\alpha$ . The following result follows easily by induction on

the Picard iterations used to compute the distances [9]; for all states  $s, t \in S$  and a discount factor  $\alpha \in [0, 1)$ ,

$$[s \preceq_1 t]^\alpha \leq [s \preceq_1 t] \qquad [s \simeq_1 t]^\alpha \leq [s \simeq_1 t]. \quad (2.5)$$

Using techniques similar to the undiscounted case, we can prove that for every game structure  $G$  and discount factor  $\alpha \in [0, 1)$ , the fixpoint  $[\preceq_i]^\alpha$  is a directed metric and  $[\simeq_i]^\alpha$  is a metric, and that they are *reciprocal*, i.e.,  $[\preceq_1]^\alpha = [\succeq_2]^\alpha$  and  $[\simeq_1]^\alpha = [\simeq_2]^\alpha$ . Given the discounted bisimulation metric coincides for the two players, we write  $[\simeq_g]^\alpha$  instead of  $[\simeq_1]^\alpha$  and  $[\simeq_2]^\alpha$ . We now state without proof that the discounted  $\mu$ -calculus provides a logical characterization of the discounted metric. The proof is based on induction on the structure of formulas, and closely follows the result for the undiscounted case [11]. Let  $d\mu^\alpha$  (respectively,  $d\mu_1^{\alpha,+}$ ) consist of all discounted  $\mu$ -calculus formulas (respectively, all discounted  $\mu$ -calculus formulas with only the  $\text{Pre}_1$  operator and all negations before atomic propositions). It follows that for all game structures  $G$  and states  $s, t \in S$ ,

$$[s \preceq_1 t]^\alpha = \sup_{\varphi \in d\mu_1^{\alpha,+}} (\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)) \qquad [s \simeq_g t]^\alpha = \sup_{\varphi \in d\mu^\alpha} |\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)|. \quad (2.6)$$

**Metric kernels.** The kernel of the metric  $[\simeq_g]^\alpha$  ( $[\simeq_g]^\alpha$ ) defines an equivalence relation  $\simeq_g$  ( $\simeq_g^\alpha$ ) on the states of a game structure:  $s \simeq_g t$  ( $s \simeq_g^\alpha t$ ) iff  $[s \simeq_g t] = 0$  ( $[s \simeq_g^\alpha t] = 0$ ); the relation  $\simeq_g$  is called the *game bisimulation* relation [11] and the relation  $\simeq_g^\alpha$  is called the *discounted game bisimulation* relation. Similarly, we define the *game simulation* preorder  $s \preceq_1 t$  as the kernel of the directed metric  $[\preceq_1]^\alpha$ , that is,  $s \preceq_1 t$  iff  $[s \preceq_1 t] = 0$ . The *discounted game simulation* preorder is defined analogously.

### 3. BOUNDS FOR AVERAGE AND DISCOUNTED PAYOFF GAMES

From (2.3) it follows that the game bisimulation metric provides a tight bound for the difference in valuations of quantitative  $\mu$ -calculus formulas. In this section, we show that the game bisimulation metric also provides a bound for the difference in average and discounted value of games. This lends further support for the game bisimulation metric, and its kernel, the game bisimulation relation, being the canonical game metrics and relations.

**3.1. Discounted payoff games.** Let  $\pi_1$  and  $\pi_2$  be strategies of player 1 and player 2 respectively. Let  $\alpha \in [0, 1)$  be a discount factor. The  $\alpha$ -discounted payoff  $v_1^\alpha(s, \pi_1, \pi_2)$  for player 1 at a state  $s$  for a variable  $r \in \mathcal{V}$  and the strategies  $\pi_1$  and  $\pi_2$  is defined as:

$$v_1^\alpha(s, \pi_1, \pi_2) = (1 - \alpha) \cdot \sum_{n=0}^{\infty} \alpha^n \cdot \mathbb{E}_s^{\pi_1, \pi_2}([r](X_n)), \quad (3.1)$$

where  $X_n$  is a random variable representing the state of the game in step  $n$ . The discounted payoff for player 2 is defined as  $v_2^\alpha(s, \pi_1, \pi_2) = -v_1^\alpha(s, \pi_1, \pi_2)$ . Thus, player 1 wins (and player 2 loses) the “discounted sum” of the valuations of  $r$  along the path, where the discount factor weighs future rewards with the discount  $\alpha$ . Given a state  $s \in S$ , we are interested in finding the maximal payoff  $v_i^\alpha(s)$  that player  $i$  can ensure against all opponent strategies, when the game starts from state  $s \in S$ . This maximal payoff is given by:

$$w_i^\alpha(s) = \sup_{\pi_i \in \Pi_i} \inf_{\pi_{\sim i} \in \Pi_{\sim i}} v_i(s, \pi_i, \pi_{\sim i}).$$



These values can be computed as the limit of the sequence of  $\alpha$ -discounted,  $n$ -step rewards, for  $n \rightarrow \infty$ . For  $i \in \{1, 2\}$ , we define a sequence of valuations  $w_i^\alpha(0)(s)$ ,  $w_i^\alpha(1)(s)$ ,  $w_i^\alpha(2)(s)$ , ... as follows: for all  $s \in S$  and  $n \geq 0$ :

$$w_i^\alpha(n+1)(s) = (1-\alpha) \cdot [r](s) + \alpha \cdot \text{Pre}_i(w_i^\alpha(n))(s). \quad (3.2)$$

where the initial valuation  $w_i^\alpha(0)$  is arbitrary. Shapley proved that  $w_i^\alpha = \lim_{n \rightarrow \infty} w_i^\alpha(n)$  [28].

**3.2. Average payoff games.** Let  $\pi_1$  and  $\pi_2$  be strategies of player 1 and player 2 respectively. The *average payoff*  $v_1(s, \pi_1, \pi_2)$  for player 1 at a state  $s$  for a variable  $r \in \mathcal{V}$  and the strategies  $\pi_1$  and  $\pi_2$  is defined as

$$v_1(s, \pi_1, \pi_2) = \lim_{n \rightarrow \infty} \inf \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}_s^{\pi_1, \pi_2}([r](X_k)), \quad (3.3)$$

where  $X_k$  is a random variable representing the  $k$ -th state of the game. The reward for player 2 is  $v_2(s, \pi_1, \pi_2) = -v_1(s, \pi_1, \pi_2)$ . A game structure  $G$  with average payoff is called an average reward game. The *average value* of the game  $G$  at  $s$  for player  $i \in \{1, 2\}$  is defined as

$$w_i(s) = \sup_{\pi_i \in \Pi_i} \inf_{\pi_{\sim i} \in \Pi_{\sim i}} v_i(s, \pi_i, \pi_{\sim i}).$$

Mertens and Neyman established the determinacy of average reward games, and showed that the limit of the discounted value of a game as all the discount factors tend to 1 is the same as the average value of the game: for all  $s \in S$  and  $i \in \{1, 2\}$ , we have  $\lim_{\alpha \rightarrow 1} w_i^\alpha(s) = w_i(s)$  [23]. It is easy to show that the average value of a game is a valuation.

**3.3. Metrics for discounted and average payoffs.** We show that the game simulation metric  $[\preceq_1]$  provides a bound for discounted and long-run rewards. The discounted metric  $[\preceq_1]^\alpha$  on the other hand does not provide such a bound as the following example shows.



Figure 1: Example that shows that the discounted metric may not be an upper bound for the difference in the discounted value across states.

**Example 1.** Consider a game consisting of four states  $s, t, s', t'$ , and a variable  $r$ , with  $[r](s) = 2$ ,  $[r](s') = 2.1$ ,  $[r](t) = 5$ , and  $[r](t') = 8$  as shown in Figure 1. All players have only one move at each state, and the transition relation is deterministic. Consider a discount factor  $\alpha = 0.9$ . The 0.9-discounted metric distance between states  $s'$  and  $s$ , is  $[s' \simeq_g s]^{0.9} = 0.9 \cdot (8 - 5) = 2.7$ . For the difference in discounted values between the states we proceed as follows. Using formulation 3.2, taking  $w^\alpha(0)(t) = 5$ , since state  $t$  is absorbing, we get  $w^\alpha(1)(t) = (1 - 0.9) \cdot 5 + 0.9 \cdot 5 = 5$  which leads to  $w^\alpha(n)(t) = 5$  for all  $n \geq 0$ . Similarly  $w^\alpha(n)(t') = 8$  for all  $n \geq 0$ . Therefore, the difference in discounted values between  $s$  and  $s'$ , again using 3.2, is given by:  $w^\alpha(s') - w^\alpha(s) = (1 - 0.9) \cdot (2.1 - 2) + 0.9 \cdot (8 - 5) = 2.71$ . ■

In the following we consider player 1 rewards (the case for player 2 is identical).

**Theorem 1.** *The following assertions hold.*

- (1) For all game structures  $G$ ,  $\alpha$ -discounted rewards  $w_1^\alpha$ , for all states  $s, t \in S$ , we have,
  - (a)  $w_1^\alpha(s) - w_1^\alpha(t) \leq [s \preceq_1 t]$  and (b)  $|w_1^\alpha(s) - w_1^\alpha(t)| \leq [s \simeq_g t]$ .
- (2) There exists a game structure  $G$ , states  $s, t \in S$ , such that for all  $\alpha$ -discounted rewards  $w_1^\alpha$ ,  $w_1^\alpha(t) - w_1^\alpha(s) > [t \simeq_g s]^\alpha$ .

*Proof.* We first prove assertion (1)(a). As the metric can be computed via Picard iteration, we have for all  $n \geq 0$ :

$$[s \preceq_1^n t] = p(s, t) \sqcup \sup_{k \in C([\preceq_1^{n-1}])} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)). \quad (3.4)$$

We prove by induction on  $n \geq 0$  that  $w_1^\alpha(n)(s) - w_1^\alpha(n)(t) \leq [s \preceq_1^n t]$ . For all  $s \in S$ , taking  $w_1^\alpha(0)(s) = [r](s)$ , the base case follows. Assume the result holds for  $n - 1 \geq 0$ . We have:

$$\begin{aligned} w_1^\alpha(n)(s) - w_1^\alpha(n)(t) &= (1 - \alpha) \cdot [r](s) + \alpha \cdot \text{Pre}_1(w^\alpha(n - 1))(s) - \\ &\quad (1 - \alpha) \cdot [r](t) - \alpha \cdot \text{Pre}_1(w^\alpha(n - 1))(t) \\ &= (1 - \alpha) \cdot ([r](s) - [r](t)) + \\ &\quad \alpha \cdot (\text{Pre}_1(w^\alpha(n - 1))(s) - \text{Pre}_1(w^\alpha(n - 1))(t)) \\ &\leq (1 - \alpha) \cdot p(s, t) + \alpha \cdot [s \preceq_1^{n-1} t] \leq [s \preceq_1^n t], \end{aligned}$$

where the last step follows by (3.4), since by the induction hypothesis we have  $w_1^\alpha(n - 1) \in C([\preceq_1^{n-1}])$ . This proves assertion (1)(a). Given (1)(a), from the definition of  $[s \simeq_g t] = [s \preceq_1 t] \sqcup [t \preceq_1 s]$ , (1)(b) follows.

The example shown in Figure 1 proves the second assertion. ■

Using the fact that the limit of the discounted reward, for a discount factor that approaches 1, is equal to the average reward, we obtain that the metrics provide a bound for the difference in average values as well.

**Corollary 1.** *For all game structures  $G$  and states  $s$  and  $t$ , we have (a)  $w(s) - w(t) \leq [s \preceq_1 t]$  and (b)  $|w(s) - w(t)| \leq [s \simeq_g t]$ .*

**3.4. Metrics for total rewards.** The *total reward*  $v_1^T(s, \pi_1, \pi_2)$  for player 1 at a state  $s$  for a variable  $r \in \mathcal{V}$  and the strategies  $\pi_1 \in \Pi_1$  and  $\pi_2 \in \Pi_2$  is defined as [17]:

$$v_1^T(s, \pi_1, \pi_2) = \lim_{n \rightarrow \infty} \inf \frac{1}{n} \sum_{k=0}^{n-1} \sum_{j=0}^k \mathbb{E}_s^{\pi_1, \pi_2}([r](X_j)), \quad (3.5)$$

where  $X_j$  is a random variable representing the  $j$ -th state of the game. The payoff  $v_2^T(s, \pi_1, \pi_2)$  for player 2 is defined by replacing  $[r]$  with  $-[r]$  in (3.5). The *total-reward value* of the game  $G$  at  $s$  for player  $i \in \{1, 2\}$  is defined analogously to the average value, via,

$$w_i^T(s) = \sup_{\pi_i \in \Pi_i} \inf_{\pi_{\sim i} \in \Pi_{\sim i}} v_i^T(s, \pi_1, \pi_2).$$

While the game simulation metric  $[\simeq_g]$  provides an upper bound for the difference in discounted reward across states, as well as for the difference in average reward across states, it does not provide a bound for the difference in total reward. We now introduce a new

metric, the *total reward metric*,  $[\bowtie_g]$ , which provides such a bound. For a discount factor  $\alpha \in [0, 1)$ , we define a metric transformer  $H_{\sqsubseteq_1}^\alpha : \mathcal{M} \mapsto \mathcal{M}$  as follows. For all  $d \in \mathcal{M}$  and  $s, t \in S$ , we let:

$$H_{\sqsubseteq_1}^\alpha(d)(s, t) = p(s, t) + \alpha \cdot \sup_{k \in C(d)} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)) . \quad (3.6)$$

The metric  $[\sqsubseteq_1]^\alpha$  (resp.  $[\bowtie_1]^\alpha$ ) is obtained as the least (resp. least symmetrical) fixpoint of (3.6). We write  $[\sqsubseteq_1]$  for  $[\sqsubseteq_1]^1$ , and  $[\bowtie_1]$  for  $[\bowtie_1]^1$ . These metrics are *reciprocal*, i.e.,  $[\sqsubseteq_1]^\alpha = [\supseteq_2]^\alpha$  and  $[\bowtie_1]^\alpha = [\bowtie_2]^\alpha$ . If  $\alpha < 1$  we get the *discounted total reward metric* and if  $\alpha = 1$  we get the *undiscounted total reward metric*. While the discounted total reward metric is bounded, the undiscounted total reward metric may not be bounded. The total metrics provide bounds for the difference in discounted, average, and total reward between states.

**Theorem 2.** *The following assertions hold.*

- (1) For all game structures  $G$ , for all discount factors  $\alpha \in [0, 1)$ , for all states  $s, t \in S$ ,
- (a)  $[s \sqsubseteq_1 t]^\alpha \leq (\theta_2 - \theta_1)/(1 - \alpha)$ ,
  - (b)  $[s \sqsubseteq_1 t]^\alpha \leq [s \sqsubseteq_1 t]$ ,
  - (c)  $w_1^\alpha(s) - w_1^\alpha(t) \leq [s \sqsubseteq_1 t]^\alpha$ ,
  - (d)  $w_1(s) - w_1(t) \leq [s \sqsubseteq_1 t]$ ,
  - (e)  $w_1^T(s) - w_1^T(t) \leq [s \sqsubseteq_1 t]$ .
- (2) There exists a game structure  $G$  and states  $s, t \in S$  such that,  $[s \sqsubseteq_1 t] = \infty$ .

*Proof.* For assertion (1)(a), notice that  $p(s, t) \leq (\theta_2 - \theta_1)$ . Consider the  $n$ -step Picard iterate towards the metric distance. We have,

$$[s \sqsubseteq_1^n t]^\alpha \leq \sum_{i=0}^n \alpha^i \cdot (\theta_2 - \theta_1) .$$

In the limit this yields  $[s \sqsubseteq_1 t]^\alpha \leq (\theta_2 - \theta_1)/(1 - \alpha)$ . Assertion (1)(b) follows by induction on the Picard iterations that realize the metric distance. For all  $n \geq 0$ ,  $[s \sqsubseteq_1^n t]^\alpha \leq [s \sqsubseteq_1^n t]$ . Assertion (1)(c) follows by the definition of the discounted total reward metric where we have replaced the  $\sqcup$  with a  $+$ . By induction, for all  $n \geq 0$ , from the proof of Theorem 1 we have,

$$w_1^\alpha(n)(s) - w_1^\alpha(n)(t) \leq (1 - \alpha) \cdot p(s, t) + \alpha \cdot (\text{Pre}_1(w_1^\alpha(n-1))(s) - \text{Pre}_1(w_1^\alpha(n-1))(t)) \leq [s \sqsubseteq_1^n t]^\alpha .$$

For assertion (1)(d), towards an inductive argument on the Picard iterates that realize the metric, for all  $n \geq 0$ , we have  $[s \sqsubseteq_1^n t] \leq [s \sqsubseteq_1^n t]$ , which in the limit gives  $[s \sqsubseteq_1 t] \leq [s \sqsubseteq_1 t]$ . This leads to  $w_1(s) - w_1(t) \leq [s \sqsubseteq_1 t]$ , using Corollary 1. This proves assertion (1)(d). We now prove assertion (1)(e) by induction and show that for all  $n \geq 0$ ,  $w_1^T(n)(s) - w_1^T(n)(t) \leq [s \sqsubseteq_1^n t]$ . As the metric can be computed via Picard iteration, we have for all  $n \geq 0$ :

$$[s \sqsubseteq_1^n t] = p(s, t) + \sup_{k \in C([\sqsubseteq_1^{n-1}])} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)) . \quad (3.7)$$

We define a valuation transformer  $u : \mathcal{F} \mapsto \mathcal{F}$  as  $u(0) = [r]$  and for all  $n > 0$  and state  $s \in S$  as,

$$u(n)(s) = [r](s) + \text{Pre}_1(u(n-1))(s)$$

We take  $w_1^T(0) = u(0) = [r]$  and for  $n > 0$ , from the definition of total rewards (3.5), we get the  $n$ -step total reward value at a state  $s \in S$  in terms of  $u$  as,

$$w_1^T(n)(s) = \frac{1}{n} \cdot \sum_{i=1}^n u(i)(s).$$

Notice that  $w_1^T(n)(s) \leq u(n)$  for all  $n \geq 0$ . When  $n = 0$ , the result is immediate by the definition of  $w_1^T(0)$ , noticing that  $[s \preceq_1^0 t] = p(s, t)$ . Assume the result holds for  $n - 1 \geq 0$ . We have:

$$\begin{aligned} w_1^T(n)(s) - w_1^T(n)(t) &= \frac{1}{n} \cdot \sum_{i=1}^n u(i)(s) - \frac{1}{n} \cdot \sum_{i=1}^n u(i)(t) \\ &= \frac{1}{n} \cdot \sum_{i=1}^n (u(i)(s) - u(i)(t)) \\ &= \frac{1}{n} \cdot \sum_{i=1}^n (([r](s) - [r](t)) + \\ &\quad (\text{Pre}_1(u(i-1))(s) - \text{Pre}_1(u(i-1))(t))) \end{aligned} \quad (3.8)$$

$$\leq \frac{1}{n} \cdot \sum_{i=1}^n [s \preceq_1^i t] \quad (3.9)$$

$$\leq [s \preceq_1^n t], \quad (3.10)$$

where (3.9) follows from (3.8) by (3.7), since by our induction hypothesis we have  $w_1^T(i) \leq u(i) \in C([\preceq_1^i])$  for all  $0 \leq i < n$  and (3.10) follows from (3.9) from the monotonicity of the undiscounted total reward metric. To prove assertion (2), consider the game structure on the left hand side in Figure 1. The total reward at state  $s$  is unbounded;  $w_1^T(s) = 2+5+\dots = \infty$ . Now consider a modified version of the game, with identical structure and with states  $s'$  and  $t'$  corresponding to  $s$  and  $t$  of the original game. Let  $[r](t') = 0$ . In the modified game,  $w_1^T(s') = 2$ . From result (1)(e), since  $w_1^T(s) = \infty$  and  $w_1^T(s') = 2$ , we have  $[s \preceq_1 s'] = \infty$ . ■

It is a very simple observation that the quantitative  $\mu$ -calculus does not provide a logical characterization for  $[\preceq_1^\alpha]$  or  $[\preceq_1]$ . In fact, all formulas of the quantitative  $\mu$ -calculus have valuations in the interval  $[\theta_1, \theta_2]$ , while as stated in Theorem 2, the total reward can be unbounded. The difference is essentially due to the fact that our version of the quantitative  $\mu$ -calculus lacks a “+” operator. It is not clear how to introduce such a + operator in a context sufficiently restricted to provide a logical characterization for  $[\preceq_1^\alpha]$ ; above all, it is not clear whether a canonical calculus, with interesting formal properties, would be obtained.

**3.5. Metric kernels.** We now show that the kernels of all the metrics defined in the paper coincide: an algorithm developed for the game kernels  $\preceq_1$  and  $\simeq_g$ , compute the kernels of the corresponding discounted and total reward metrics as well.

**Theorem 3.** *For all game structures  $G$ , states  $s$  and  $t$ , all discount factors  $\alpha \in [0, 1)$ , the following statements are equivalent:*

$$(a) [s \preceq_1 t] = 0 \quad (b) [s \preceq_1 t]^\alpha = 0 \quad (c) [s \preceq_1 t]^\alpha = 0.$$

*Proof.* We prove  $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$ . We assume  $0 < \alpha < 1$ . Assertion (a) implies that  $p(s, t) = 0$  and  $\sup_{k \in C([\preceq_1])} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t)) \leq 0$ ; Since  $C([\preceq_1]^\alpha) \subseteq C([\preceq_1])$  from (2.5), (b) follows. We prove  $(b) \Rightarrow (c)$  by induction on the Picard iterations that compute  $[s \preceq_1 t]^\alpha$  and  $[s \triangleleft_1 t]^\alpha$ . The base case is immediate. Assume that for all states  $s$  and  $t$ ,  $[s \preceq_1^{n-1} t]^\alpha = 0$  implies  $[s \triangleleft_1^{n-1} t]^\alpha = 0$ . Towards a contradiction, assume  $[s \preceq_1^n t]^\alpha = 0$  but  $[s \triangleleft_1^n t]^\alpha > 0$ . Then there must be  $k \in C([\triangleleft_1^{n-1}]^\alpha)$  such that  $\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t) > 0$ . By our induction hypothesis, there exists a  $\delta > 0$  such that  $k' = \delta \cdot k \in C([\preceq_1^{n-1}]^\alpha)$ . Since Pre is multi-linear, the player optimal responses in  $\text{Pre}_1(k)(s)$  remain optimal for  $k'$ . But this means  $(\text{Pre}_1(k')(s) - \text{Pre}_1(k')(t)) > 0$  for  $k' \in C([\preceq_1^{n-1}]^\alpha)$ , leading to  $[s \preceq_1^n t]^\alpha > 0$ ; a contradiction. Therefore,  $(b) \Rightarrow (c)$ . In a similar fashion we can show that  $(c) \Rightarrow (a)$ . ■

#### 4. ALGORITHMS FOR TURN-BASED GAMES AND MDPs

In this section, we present algorithms for computing the metric and its kernel for turn-based games and MDPs. We first present a polynomial time algorithm to compute the operator  $H_{\preceq_i}(d)$  that gives the *exact* one-step distance between two states, for  $i \in \{1, 2\}$ . We then present a PSPACE algorithm to decide whether the limit distance between two states  $s$  and  $t$  (i.e.,  $[s \preceq_1 t]$ ) is at most a rational value  $r$ . Our algorithm matches the best known bound known for the special class of Markov chains [31]. Finally, we present improved algorithms for the important case of the kernel of the metrics. Since by Theorem 3 the kernels of the metrics introduced in this paper coincide, we present our algorithms for the kernel of the undiscounted metric. For the bisimulation kernel our algorithm is significantly more efficient compared to previous algorithms.

**4.1. Algorithms for the metrics.** For turn-based games and MDPs, only one player has a choice of moves at a given state. We consider two player 1 states. A similar analysis applies to player 2 states. We remark that the distance between states in  $S_i$  and  $S_{\sim i}$  is always  $\theta_2 - \theta_1$  due to the existence of the variable *turn*. For a metric  $d \in \mathcal{M}$ , and states  $s, t \in S_1$ , computing  $H_{\preceq_1}(d)(s, t)$ , given that  $p(s, t)$  is trivially computed by its definition, entails evaluating the expression,  $\sup_{k \in C(d)} (\text{Pre}_1(k)(s) - \text{Pre}_1(k)(t))$ , which is the same as,  $\sup_{k \in C(d)} \sup_{x \in \mathcal{D}_1(s)} \inf_{y \in \mathcal{D}_1(t)} (\mathbb{E}_s^x(k) - \mathbb{E}_t^y(k))$ , since  $\text{Pre}_1(k)(s) = \sup_{x \in \mathcal{D}_1(s)} (\mathbb{E}_s^x(k))$  and  $\text{Pre}_1(k)(t) = \sup_{y \in \mathcal{D}_1(t)} (\mathbb{E}_t^y(k))$  as player 1 is the only player with a choice of moves at state  $s$ . By expanding the expectations, we get the following form,

$$\sup_{k \in C(d)} \sup_{x \in \mathcal{D}_1(s)} \inf_{y \in \mathcal{D}_1(t)} \left( \sum_{u \in S} \sum_{a \in \Gamma_1(s)} \delta(s, a)(u) \cdot x(a) \cdot k(u) - \sum_{v \in S} \sum_{b \in \Gamma_1(t)} \delta(t, b)(v) \cdot y(b) \cdot k(v) \right). \quad (4.1)$$

We observe that the one-step distance as defined in (4.1) is a *sup-inf non-linear (quadratic)* optimization problem. We now present two lemmas by which we transform (4.1) to an *inf linear* optimization problem, which we solve by linear programming (LP). The first lemma reduces (4.1) to an equivalent formulation that considers only pure moves at state  $s$ . The second lemma further reduces (4.1), using duality, to a formulation that can be solved using LP.

**Lemma 1.** *For all turn-based game structures  $G$ , for all player  $i$  states  $s$  and  $t$ , given a metric  $d \in \mathcal{M}$ , the following equality holds,*

$$\sup_{k \in C(d)} \sup_{x \in \mathcal{D}_i(s)} \inf_{y \in \mathcal{D}_i(t)} (\mathbb{E}_s^x(k) - \mathbb{E}_t^y(k)) = \sup_{a \in \Gamma_i(s)} \inf_{y \in \mathcal{D}_i(t)} \sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)).$$

*Proof.* We prove the result for player 1 states  $s$  and  $t$ , with the proof being identical for player 2. Given a metric  $d \in \mathcal{M}$ , we have,

$$\begin{aligned} \sup_{k \in C(d)} \sup_{x \in \mathcal{D}_1(s)} \inf_{y \in \mathcal{D}_1(t)} (\mathbb{E}_s^x(k) - \mathbb{E}_t^y(k)) &= \sup_{k \in C(d)} \left( \sup_{x \in \mathcal{D}_1(s)} \mathbb{E}_s^x(k) - \sup_{y \in \mathcal{D}_1(t)} \mathbb{E}_t^y(k) \right) \\ &= \sup_{k \in C(d)} \left( \sup_{a \in \Gamma_1(s)} \mathbb{E}_s^a(k) - \sup_{y \in \mathcal{D}_1(t)} \mathbb{E}_t^y(k) \right) \end{aligned} \quad (4.2)$$

$$\begin{aligned} &= \sup_{k \in C(d)} \sup_{a \in \Gamma_1(s)} \inf_{y \in \mathcal{D}_1(t)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)) \\ &= \sup_{a \in \Gamma_1(s)} \sup_{k \in C(d)} \inf_{y \in \mathcal{D}_1(t)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)) \end{aligned} \quad (4.3)$$

$$= \sup_{a \in \Gamma_1(s)} \inf_{y \in \mathcal{D}_1(t)} \sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)) \quad (4.4)$$

For a fixed  $k \in C(d)$ , since pure optimal strategies exist at each state for turn-based games and MDPs, we replace the  $\sup_{x \in \mathcal{D}_1(s)}$  with  $\sup_{a \in \Gamma_1(s)}$  yielding (4.2). Since the difference in expectations is multi-linear,  $y \in \mathcal{D}_1(t)$  is a probability distribution and  $C(d)$  is a compact convex set, we can use the generalized minimax theorem [29], and interchange the innermost sup inf to get (4.4) from (4.3).  $\blacksquare$

The proof of Lemma 1 is illustrated using the following example.

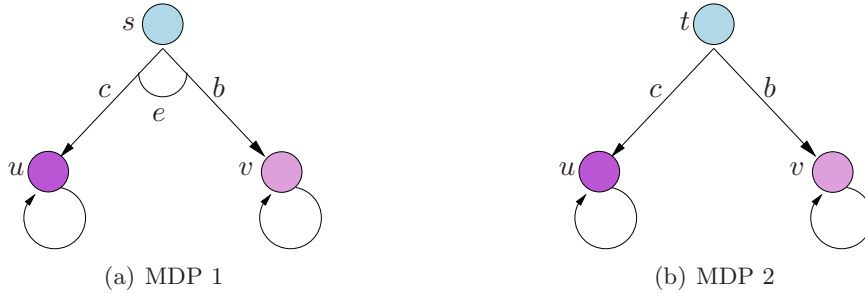


Figure 2: An example illustrating the proof of Lemma 1.

**Example 4.1.** Consider the example in Figure 2. In the MDPs shown in the figure, every move leads to a unique successor state, with the exception of move  $e \in \Gamma_1(s)$ , which leads to states  $u$  and  $v$  with equal probability. Assume the variable valuations are such that all states are at a propositional distance of 1. Without loss of generality, assume that the valuation  $k \in C(d)$  is such that  $k(u) > k(v)$ . By the linearity of expectations, for move  $c \in \Gamma_1(s)$ ,  $\mathbb{E}_s^c(k) \geq \mathbb{E}_s^x(k)$  for all  $x \in \mathcal{D}_1(s)$ . Similar arguments can be made for  $k(u) < k(v)$ . This gives an informal justification for step (4.2) in the proof; given a  $k \in C(d)$ , there exist pure optimal strategies for the single player with a choice of moves at each state. While we can use pure moves at states  $s$  and  $t$  if  $k \in C(d)$  is known, the principle difficulty in directly computing the left hand side of the equality arises from the uncountably many values for  $k$ ; the distance is the supremum over all possible values of  $k$ . In the final equality, step

(4.4), and hence by this theorem, we have avoided this difficulty, by showing an equivalent expression that picks a  $k \in C(d)$  to show the difference in distributions induced over states. As we shall see, this enables computing the one-step metric distance using a trans-shipping formulation. We remark that while we can use pure moves at state  $s$ , we cannot do so at state  $t$  in the right hand side of step (4.4) of the proof. Firstly, the proof of the theorem depends on  $y \in \mathcal{D}_1(t)$  being convex. Secondly, if we could restrict our attention to pure moves at state  $t$ , then we can replace  $\inf_{y \in \mathcal{D}_1(t)}$  with  $\inf_{f \in \Gamma_1(t)}$  on the right hand side. But this yields too fine a one-step distance. Consider move  $e$  at state  $s$ . We see that neither  $c$  nor  $b$  at state  $t$  yield distributions over states that match the distribution induced by  $e$ . We can then always pick  $k \in C(d)$  such that  $\mathbb{E}_s^e(k) - \mathbb{E}_t^f(k) > 0$ . If we choose  $y \in \mathcal{D}_1(t)$  such that  $y(b) = y(c) = \frac{1}{2}$ , we match the distribution induced by move  $e$  from state  $s$ , which implies that for any choice of  $k \in C(d)$ ,  $\mathbb{E}_s^e(k) - \mathbb{E}_t^{y(b)=y(c)=\frac{1}{2}}(k) = 0$ . Intuitively, the right hand side of the equality can be interpreted as a game between a protagonist and an antagonist, with the protagonist picking  $y \in \mathcal{D}_1(t)$ , for every pure move  $a \in \Gamma_1(s)$ , to match the induced distributions over states. The antagonist then picks a  $k \in C(d)$  to maximize the difference in induced distributions. If the distributions match, then no choice of  $k \in C(d)$  yields a difference in expectations bounded away from 0.

From Lemma 1, given  $d \in \mathcal{M}$ , we can write the player 1 one-step distance between states  $s$  and  $t$  as follows,

$$\text{OneStep}(s, t, d) = \sup_{a \in \Gamma_1(s)} \inf_{y \in \mathcal{D}_1(t)} \sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)). \quad (4.5)$$

Hence we compute for all  $a \in \Gamma_1(s)$ , the expression,

$$\text{OneStep}(s, t, d, a) = \inf_{y \in \mathcal{D}_1(t)} \sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)),$$

and then choose the maximum, i.e.,  $\max_{a \in \Gamma_1(s)} \text{OneStep}(s, t, d, a)$ . We now present a lemma that helps reduce the above inf – sup optimization problem to a linear program. We first introduce some notation. We denote by  $\lambda$  the set of variables  $\lambda_{u,v}$ , for  $u, v \in S$ . Given  $a \in \Gamma_1(s)$ , and a distribution  $y \in \mathcal{D}_1(t)$ , we write  $\lambda \in \Phi(s, t, a, y)$  if the following linear constraints are satisfied:

- (1) for all  $v \in S$  :  $\sum_{u \in S} \lambda_{u,v} = \delta(s, a)(v)$ ;    (2) for all  $u \in S$  :  $\sum_{v \in S} \lambda_{u,v} = \sum_{b \in \Gamma_1(t)} y(b) \cdot \delta(t, b)(u)$ ;  
 (3) for all  $u, v \in S$  :  $\lambda_{u,v} \geq 0$ .

**Lemma 2.** *For all turn-based game structures and MDPs  $G$ , for all  $d \in \mathcal{M}$ , and for all  $s, t \in S$ , the following assertion holds:*

$$\sup_{a \in \Gamma_1(s)} \inf_{y \in \mathcal{D}_1(t)} \sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)) = \sup_{a \in \Gamma_1(s)} \inf_{y \in \mathcal{D}_1(t)} \inf_{\lambda \in \Phi(s, t, a, y)} \left( \sum_{u, v \in S} d(u, v) \cdot \lambda_{u,v} \right).$$

*Proof.* Since duality always holds in LP, from the LP duality based results of [32], for all  $a \in \Gamma_1(s)$  and  $y \in \mathcal{D}_1(t)$ , the maximization over all  $k \in C(d)$  can be re-written as a minimization problem as follows:

$$\sup_{k \in C(d)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^y(k)) = \inf_{\lambda \in \Phi(s, t, a, y)} \left( \sum_{u, v \in S} d(u, v) \cdot \lambda_{u,v} \right).$$

The formula on the right hand side of the above equality is the *trans-shipping formulation*, which solves for the minimum cost of shipping the distribution  $\delta(s, a)$  into  $\delta(t, y)$ , with edge costs  $d$ . The result of the lemma follows. ■

Using the above result we obtain the following LP for  $\text{OneStep}(s, t, d, a)$  over the variables: (a)  $\{\lambda_{u,v}\}_{u,v \in S}$ , and (b)  $y_b$  for  $b \in \Gamma_1(t)$ :

$$\text{Minimize } \sum_{u,v \in S} d(u, v) \cdot \lambda_{u,v} \quad \text{subject to} \quad (4.6)$$

$$\begin{aligned} (1) \text{ for all } v \in S : \sum_{u \in S} \lambda_{u,v} &= \delta(s, a)(v); & (2) \text{ for all } u \in S : \sum_{v \in S} \lambda_{u,v} &= \sum_{b \in \Gamma_1(t)} y_b \cdot \delta(t, b)(u); \\ (3) \text{ for all } u, v \in S : \lambda_{u,v} &\geq 0; & (4) \text{ for all } b \in \Gamma_1(t) : y_b &\geq 0; & (5) \sum_{b \in \Gamma_1(t)} y_b &= 1. \end{aligned}$$

**Example 4.2.** We now use the MDPs in Figure 3(a) and 3(b) to compute the simulation distance between states using the results in Lemma 1 and Lemma 2. In the figure, states of the same color have a propositional distance of 0 and states of different colors have a propositional distance of 1;  $p(s, s') = p(t, t') = p(u, u') = p(v, v') = p(t', w') = 0$ . In MDP 1, shown in Figure 3(a),  $\delta(s, a)(t) = \delta(t, b)(v) = \delta(t, c)(u) = 1$  and  $\delta(t, f)(u) = \delta(t, f)(v) = \frac{1}{2}$ . In MDP 2, shown in Figure 3(b),  $\delta(s', a)(w') = \delta(s', b)(t') = 1$ ,  $\delta(t', c)(u') = \frac{1}{2} - \epsilon$ ,  $\delta(t', c)(v') = \frac{1}{2} + \epsilon$ ,  $\delta(w', e)(u') = \delta(w', f)(v') = 1 - \epsilon$  and  $\delta(w', e)(v') = \delta(w', f)(u') = \epsilon$ .

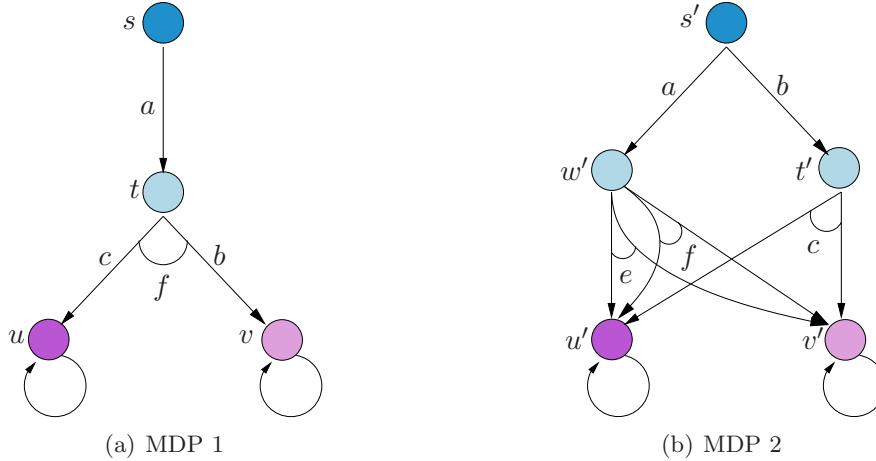


Figure 3: An example used to compute the simulation metric between states. States of the same color have a propositional distance of 0.

In Table 2, we show the simulation metric distance between states of the MDPs in Figure 3(a) and Figure 3(b). Consider states  $t$  and  $t'$ .  $c$  is the only move available to player 1 from state  $t'$  and it induces a transition probability of  $\frac{1}{2} + \epsilon$  to state  $v'$  and  $\frac{1}{2} - \epsilon$  to state  $u'$ . For the pure move  $c$  at state  $t$ , the induced transition probabilities and edge costs in the trans-shipping formulation are shown in Figure 4(a). It is easy to see that the trans-shipping cost in this case is  $\frac{1}{2} + \epsilon$ ; shown in Table 1 along the row corresponding to move  $c$  from state  $t$  and column corresponding to state  $t'$ . Similarly, the trans-shipping cost for the moves  $b$  and  $f$  from state  $t$  are  $\frac{1}{2} - \epsilon$  and  $\epsilon$  respectively. The metric distance



$t$	$w'$	$t'$		
$\Gamma_1(t)$	$x \in \mathcal{D}_1(w')$	$Cost$	$x \in \mathcal{D}_1(t')$	$Cost$
$b$	$x(f) = 1$	$\epsilon$	$x(c) = 1$	$\frac{1}{2} - \epsilon$
$c$	$x(e) = 1$	$\epsilon$	$x(c) = 1$	$\frac{1}{2} + \epsilon$
$f$	$x(f) = x(e) = \frac{1}{2}$	$0$	$x(c) = 1$	$\epsilon$

Table 1: The moves from states  $w'$  and  $t'$  that minimize the trans-shipping cost for each  $a \in \Gamma_1(t)$  and the corresponding costs.

$[\preceq]$	$s'$	$t'$	$w'$	$u'$	$v'$
$s$	$\epsilon$	$1$	$1$	$1$	$1$
$t$	$1$	$\frac{1}{2} + \epsilon$	$\epsilon$	$1$	$1$
$u$	$1$	$1$	$1$	$0$	$1$
$v$	$1$	$1$	$1$	$1$	$0$

Table 2: The simulation metric distance between states in MDP 1 and states in MDP 2.

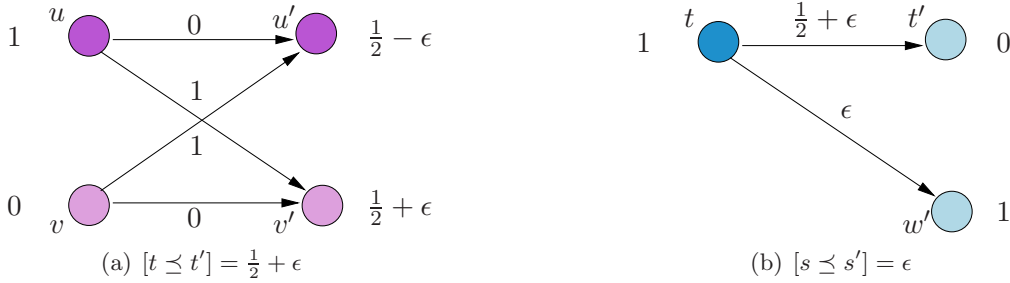


Figure 4: The trans-shipping formulation that gives the metric distances between states.

$[t \preceq t']$ , which is the maximum over these trans-shipping costs is then  $\frac{1}{2} + \epsilon$ . Now consider the states  $t$  and  $w'$ . In Table 1, we show for each pure move  $a \in \Gamma_1(t)$ , the move  $x \in \mathcal{D}_1(w')$  that minimizes the trans-shipping cost together with the minimum cost. In this case it is easy to see that  $[t \preceq w'] = \epsilon$ . Given  $[t \preceq t'] = \frac{1}{2} + \epsilon$  and  $[t \preceq w'] = \epsilon$ , we can calculate the distance  $[s \preceq s']$  from the trans-shipping formulation shown in Figure 4(b); the minimum cost is  $\epsilon$  that entails choosing move  $a$  from state  $s'$ , giving us  $[s \preceq s'] = \epsilon$ .

**Theorem 4.3.** *For all turn-based game structures and MDPs  $G$ , given  $d \in \mathcal{M}$ , for all states  $s, t \in S$ , we can compute  $H_{\preceq_1}(d)(s, t)$  in polynomial time by the Linear Program (4.6).*

For all states  $s, t \in S$ , iteration of  $\text{OneStep}(s, t, d)$  converges to the exact distance. However, in general, there are no known bounds for the rate of convergence. We now present a decision procedure to check whether the exact distance between two states is at most a rational value  $r$ . We first show how to express the predicate  $d(s, t) = \text{OneStep}(s, t, d)$ . We observe that since  $H_{\preceq_1}$  is non-decreasing, we have  $\text{OneStep}(s, t, d) \geq d(s, t)$ . It follows that the equality  $d(s, t) = \text{OneStep}(s, t, d)$  holds iff for every  $a \in \Gamma_1(s)$ , of which there are finitely many, all the linear inequalities of LP (4.6) are satisfied, and  $d(s, t) = \sum_{u, v \in S} d(u, v) \cdot \lambda_{u, v}$  holds. It then follows that  $d(s, t) = \text{OneStep}(s, t, d)$  can be written as a predicate in the theory of real closed fields. Given a rational  $r$ , two states  $s$  and  $t$ , we present an existential

theory of reals formula to decide whether  $[s \preceq_1 t] \leq r$ . Since  $[s \preceq_1 t]$  is the least fixed point of  $H_{\preceq_1}$ , we define a formula  $\Phi(r)$  that is true iff, in the fixpoint,  $[s \preceq_1 t] \leq r$ , as follows:

$$\exists d \in \mathcal{M}. \left[ \bigwedge_{u,v \in S} \text{OneStep}(u,v,d) = d(u,v) \wedge (d(s,t) \leq r) \right].$$

If the formula  $\Phi(r)$  is true, then there exists a fixpoint  $d$ , such that  $d(s,t)$  is bounded by  $r$ , which implies that in the least fixpoint  $d(s,t)$  is bounded by  $r$ . Conversely, if in the least fixpoint  $d(s,t)$  is bounded by  $r$ , then the least fixpoint is a witness  $d$  for  $\Phi(r)$  being true. Since the existential theory of reals is decidable in PSPACE [6], we have the following result.

**Theorem 4.4.** (*Decision complexity for exact distance*). *For all turn-based game structures and MDPs  $G$ , given a rational  $r$ , and two states  $s$  and  $t$ , whether  $[s \preceq_1 t] \leq r$  can be decided in PSPACE.*

**Approximation.** Given a rational  $\epsilon > 0$ , using binary search and  $\mathcal{O}(\log(\frac{\theta_2 - \theta_1}{\epsilon}))$  calls to check the formula  $\Phi(r)$ , we can obtain an interval  $[l, u]$  with  $u - l \leq \epsilon$  such that  $[s \preceq_1 t]$  lies in the interval  $[l, u]$ .

**Corollary 2. (Approximation for exact distance).** *For all turn-based game structures and MDPs  $G$ , given a rational  $\epsilon$ , and two states  $s$  and  $t$ , an interval  $[l, u]$  with  $u - l \leq \epsilon$  such that  $[s \preceq_1 t] \in [l, u]$  can be computed in PSPACE.*

**4.2. Algorithms for the kernel.** The kernel of the simulation metric  $\preceq_1$  can be computed as the limit of the series  $\preceq_1^0, \preceq_1^1, \preceq_1^2, \dots$ , of relations. For all  $s, t \in S$ , we have  $(s, t) \in \preceq_1^0$  iff  $s \equiv t$ . For all  $n \geq 0$ , we have  $(s, t) \in \preceq_1^{n+1}$  iff  $\text{OneStep}(s, t, 1_{\preceq_1^n}) = 0$ . Checking the condition  $\text{OneStep}(s, t, 1_{\preceq_1^n}) = 0$ , corresponds to solving an LP feasibility problem for every  $a \in \Gamma_1(s)$ , as it suffices to replace the minimization goal  $\gamma = \sum_{u,v \in S} 1_{\preceq_1^n}(u,v) \cdot \lambda_{u,v}$  with the constraint  $\gamma = 0$  in the LP (4.6). We note that this is the same LP feasibility problem that was introduced in [35] as part of an algorithm to decide simulation of probabilistic systems in which each label may lead to one or more distributions over states.

For the bisimulation kernel, we present a more efficient algorithm, which also improves on the algorithms presented in [35]. The idea is to proceed by partition refinement, as usual for bisimulation computations. The refinement step is as follows: given a partition, two states  $s$  and  $t$  belong to the same refined partition iff every pure move from  $s$  induces a probability distribution on equivalence classes that can be matched by mixed moves from  $t$ , and vice versa. Precisely, we compute a sequence  $\mathcal{Q}^0, \mathcal{Q}^1, \mathcal{Q}^2, \dots$ , of partitions. Two states  $s, t$  belong to the same class of  $\mathcal{Q}^0$  iff they have the same variable valuation (i.e., iff  $s \equiv t$ ). For  $n \geq 0$ , since by the definition of the bisimulation metric given in (2.2),  $[s \simeq_g t] = 0$  iff  $[s \preceq_1 t] = 0$  and  $[t \preceq_1 s] = 0$ , two states  $s, t$  in a given class of  $\mathcal{Q}^n$  remain in the same class in  $\mathcal{Q}^{n+1}$  iff both  $(s, t)$  and  $(t, s)$  satisfy the set of feasibility LP problems  $\text{OneStepBis}(s, t, \mathcal{Q}^n)$  as given below:

$\text{OneStepBis}(s, t, \mathcal{Q})$  consists of one feasibility LP problem for each  $a \in \Gamma(s)$ .

The problem for  $a \in \Gamma(s)$  has set of variables  $\{x_b \mid b \in \Gamma(t)\}$ , and set of

constraints:

$$(1) \text{ for all } b \in \Gamma(t) : x_b \geq 0, \quad (2) \sum_{b \in \Gamma(t)} x_b = 1,$$

$$(3) \text{ for all } V \in \mathcal{Q} : \sum_{b \in \Gamma(t)} \sum_{u \in V} x_b \cdot \delta(t, b)(u) \geq \sum_{u \in V} \delta(s, a)(u) .$$

In the following theorem we show that two states  $s, t \in S$  are  $n + 1$  step bisimilar iff  $\text{OneStepBis}(s, t, \mathcal{Q}^n)$  and  $\text{OneStepBis}(t, s, \mathcal{Q}^n)$  are feasible.

**Theorem 4.5.** *For all turn-based game structures and MDPs  $G$ , for all  $n \geq 0$ , given two states  $s, t \in S$  and an  $n$ -step bisimulation partition of states  $\mathcal{Q}^n$  such that  $\forall V \in \mathcal{Q}^n, \forall u, v \in V, [u \simeq_g v]^n = 0$ , the following holds,*

$$[s \simeq_g t]^{n+1} = 0 \text{ iff } \text{OneStepBis}(s, t, \mathcal{Q}^n) \text{ and } \text{OneStepBis}(t, s, \mathcal{Q}^n) \text{ are both feasible.}$$

*Proof.* We proceed by induction on  $n$ . Assume the result holds for all iteration steps up to  $n$  and consider the case for  $n + 1$ . In one direction, if  $[s \simeq_g t]^{n+1} = 0$ , then  $[s \preceq_1 t]^{n+1} = [t \preceq_1 s]^{n+1} = 0$  by the definition of the bisimulation metric. We need to show that given  $[s \preceq_1 t]^{n+1} = 0$ ,  $\text{OneStepBis}(s, t, \mathcal{Q}^n)$  is feasible. The proof is identical for  $[t \preceq_1 s]^{n+1} = 0$ . From the definition of the  $n + 1$  step simulation distance, given  $p(s, t) = 0$  by our induction hypothesis, we have,

$$\forall b \in \Gamma_1(s) \inf_{x \in \mathcal{D}_1(t)} \sup_{k \in C(d^n)} (\mathbb{E}_s^b(k) - \mathbb{E}_t^x(k)) \leq 0 . \quad (4.7)$$

Consider a player 1 move  $a \in \Gamma_1(s)$ . Since we can interchange the order of the inf and sup by the generalized minimax theorem in  $\inf_{x \in \mathcal{D}_1(t)} \sup_{k \in C(d^n)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^x(k))$ , the optimal values of  $x \in \mathcal{D}_1(t)$  and  $k \in C(d^n)$  exist and only depend on  $a$ . Let  $x_a$  and  $k_a$  be the optimal values of  $x$  and  $k$  that realize the inf and sup in  $\inf_{x \in \mathcal{D}_1(t)} \sup_{k \in C(d^n)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^x(k))$ . Using  $x_a$  and  $k_a$  in (4.7) we have:

$$\begin{aligned} \mathbb{E}_t^{x_a}(k_a) &\geq \mathbb{E}_s^a(k_a) \\ \sum_{u \in S} \delta(t, x_a)(u) \cdot k_a(u) &\geq \sum_{v \in S} \delta(s, a)(v) \cdot k_a(v) \\ \sum_{V \in \mathcal{Q}^n} \sum_{u \in V} \delta(t, x_a)(u) \cdot k_a(u) &\geq \sum_{V \in \mathcal{Q}^n} \sum_{v \in V} \delta(s, a)(v) \cdot k_a(v) \end{aligned} \quad (4.8)$$

$$\sum_{V \in \mathcal{Q}^n} \sum_{u \in V} \delta(t, x_a)(u) \geq \sum_{V \in \mathcal{Q}^n} \sum_{v \in V} \delta(s, a)(v) \quad (4.9)$$

$$\forall V \in \mathcal{Q}^n. \left( \sum_{u \in V} \delta(t, x_a)(u) \geq \sum_{u \in V} \delta(s, a)(u) \right), \quad (4.10)$$

where (4.9) follows from (4.8) by noting that for all  $V \in \mathcal{Q}^n$ , for all states  $u, v \in V$ ,  $d^n(u, v) = d^n(v, u) = 0$ , by our hypothesis, leading to  $k(u) - k(v) \leq d^n(u, v) = 0$  and  $k(v) - k(u) \leq d^n(v, u) = 0$ , which implies  $k(u) = k(v)$  for all  $k \in C(d^n)$ . To show (4.10) follows from (4.9), assume towards a contradiction that there exists a  $V' \in \mathcal{Q}^n$  such that  $\sum_{u \in V'} \delta(t, x_a)(u) < \sum_{u \in V'} \delta(s, a)(u)$ . Then there must be a  $V'' \in \mathcal{Q}^n$  such that  $\sum_{u \in V''} \delta(t, x_a)(u) > \sum_{u \in V''} \delta(s, a)(u)$  since  $\delta(t, x_a)$  is a probability distribution and the sum of the probability mass allocated to each equivalence class should be 1. Further, for all  $V \in \mathcal{Q}^n$ , for all  $u, v \in V$ , we have  $d^n(u, v) = d^n(v, u) = 0$  and for all  $u \in V$  and for all

$w \in S \setminus V$ , we have  $d^n(u, w) = d^n(w, u) = 1$ . Therefore, we can pick a feasible  $k' \in C(d^n)$  such that  $k'(v) > 0$  for all  $v \in V''$  and  $k'(v) = 0$  for all other states. Using  $k'$  we get  $\mathbb{E}_s^a(k') - \mathbb{E}_t^{x_a}(k') > 0$  which means  $k_a$  is not optimal, contradicting (4.7).

In the other direction, assume that  $\text{OneStepBis}(s, t, \mathcal{Q}^n)$  is feasible. We need to show that  $[s \preceq_1 t]^{n+1} = 0$ . Since  $\text{OneStepBis}(s, t, \mathcal{Q}^n)$  is feasible, there exists a distribution  $x_a \in \mathcal{D}_1(t)$  for all  $a \in \Gamma_1(s)$  such that,  $\forall V \in \mathcal{Q}^n. (\sum_{u \in V} \delta(t, x_a)(u) \geq \sum_{v \in V} \delta(s, a)(v))$ . By our induction hypothesis, this implies that for all  $k \in C(d^n)$ , we have  $(\mathbb{E}_s^a(k) - \mathbb{E}_t^{x_a}(k)) \leq 0$  and in particular  $\sup_{k \in C(d^n)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^{x_a}(k)) \leq 0$ . Since  $p(s, t) = 0$  by our hypothesis and we have shown,

$$\forall a \in \Gamma_1(s) \quad \inf_{x \in \mathcal{D}_1(t)} \sup_{k \in C(d^n)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^x(k)) \leq 0,$$

we have, from Lemma 1,

$$[s \preceq_1 t]^{n+1} = p(s, t) \sqcup \sup_{a \in \Gamma_1(s)} \inf_{x \in \mathcal{D}_1(t)} \sup_{k \in C(d^n)} (\mathbb{E}_s^a(k) - \mathbb{E}_t^x(k)) = 0.$$

In a similar fashion, if  $\text{OneStepBis}(t, s, \mathcal{Q}^n)$  is feasible then  $[t \preceq_1 s]^{n+1} = 0$ , which leads to  $[s \simeq_g t]^{n+1} = 0$  by the definition of the bisimulation metric, as required. ■

**Complexity.** The number of partition refinement steps required for the computation of both the simulation and the bisimulation kernel is bounded by  $O(|S|^2)$  for turn-based games and MDPs, where  $S$  is the set of states. At every refinement step, at most  $O(|S|^2)$  state pairs are considered, and for each state pair  $(s, t)$  at most  $|\Gamma(s)|$  LP feasibility problems needs to be solved. Let us denote by  $\text{LPF}(n, m)$  the complexity of solving the feasibility of  $m$  linear inequalities over  $n$  variables. We obtain the following result.

**Theorem 4.6.** *For all turn-based game structures and MDPs  $G$ , the following assertions hold:*

- (1) *the simulation kernel can be computed in  $\mathcal{O}(n^4 \cdot m \cdot \text{LPF}(n^2 + m, n^2 + 2n + m + 2))$  time;*
- (2) *the bisimulation kernel can be computed in  $\mathcal{O}(n^4 \cdot m \cdot \text{LPF}(m, n + m + 1))$  time;*

where  $n = |S|$  is the size of the state space, and  $m = \max_{s \in S} |\Gamma(s)|$ .

**Remark 1.** The best known algorithm for  $\text{LPF}(n, m)$  works in time  $\mathcal{O}(n^{2.5} \cdot \log(n))$  [34] (assuming each arithmetic operation takes unit time). The previous algorithm for the bisimulation kernel checked two way simulation and hence has the complexity  $\mathcal{O}(n^4 \cdot m \cdot (n^2 + m)^{2.5} \cdot \log(n^2 + m))$ , whereas our algorithm works in time  $\mathcal{O}(n^4 \cdot m \cdot m^{2.5} \cdot \log(m))$ . For most practical purposes, the number of moves at a state is constant (i.e.,  $m$  is constant). For the case when  $m$  is constant, the previous best known algorithm worked in  $\mathcal{O}(n^9 \cdot \log(n))$  time, whereas our algorithm works in time  $\mathcal{O}(n^4)$ .

## 5. ALGORITHMS FOR CONCURRENT GAMES

In this section we first show that the computation of the metric distance is at least as hard as the computation of optimal values in concurrent reachability games. The exact complexity of the latter is open, but it is known to be at least as hard as the square-root sum problem, which is in PSPACE but whose inclusion in NP is a long-standing open problem [16, 18]. Next, we present algorithms based on a decision procedure for the theory of real closed fields, for both checking the bounds of the exact distance and the kernel of the metrics. Our reduction to the theory of real closed fields removes one quantifier alternation

when compared to the previous known formula (inferred from [11]). This improves the complexity of the algorithm.

**5.1. Reduction of reachability games to metrics.** We will use the following terms in the result. A *proposition* is a boolean observation variable, and we say a state is labeled by a proposition  $q$  iff  $q$  is true at  $s$ . A state  $t$  is *absorbing* in a concurrent game, if both players have only one action available at  $t$ , and the next state of  $t$  is always  $t$  (it is a state with a self-loop). For a proposition  $q$ , let  $\diamond q$  denote the set of paths that visit a state labeled by  $q$  at least once. In concurrent reachability games, the objective is  $\diamond q$ , for a proposition  $q$ , and without loss of generality all states labeled by  $q$  are absorbing states.

**Theorem 4.** *Consider a concurrent game structure  $G$ , with a single proposition  $q$ , such that all states labeled by  $q$  are absorbing states. We can construct in linear-time a concurrent game structure  $G'$ , with one additional state  $t'$ , such that for all  $s \in S$ , we have*

$$[s \preceq_1 t'] = \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \Pr_s^{\pi_1, \pi_2}(\diamond q).$$

*Proof.* The concurrent game structure  $G'$  is obtained from  $G$  by adding an absorbing state  $t'$ . The states that are not labeled by  $q$ , and the additional state  $t'$ , are labeled by its complement  $\neg q$ . Observe there is only one proposition sequence from  $t'$ , and it is  $(\neg q)^\omega$ . To prove the desired claim we show that for all  $s \in S$  we have  $[s \preceq_1 t'] = \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \Pr_s^{\pi_1, \pi_2}(\diamond q)$ . From a state  $s$  in  $G$  the possible proposition sequences can be expressed as the following  $\omega$ -regular expression:  $(\neg q)^\omega \cup (\neg q)^* \cdot q^\omega$ . Since the proposition sequence from  $t'$  is  $(\neg q)^\omega$ , the supremum of the difference in values over  $q\mu$  formulas at  $s$  and  $t'$  is obtained by satisfying the set of paths formalized as  $(\neg q)^* \cdot q^\omega$  at  $s$ . The set of paths defined as  $(\neg q)^* \cdot q^\omega$  is the same as reaching  $q$  in any number of steps, since all states labeled by  $q$  are absorbing. Hence,

$$\sup_{\varphi \in q\mu^+} (\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t')) = \llbracket \mu X.(q \vee \text{Pre}_1(X)) \rrbracket(s).$$

It follows from the results of [10] that for all  $s \in S$  we have,

$$\llbracket \mu X.(q \vee \text{Pre}_1(X)) \rrbracket(s) = \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \Pr_s^{\pi_1, \pi_2}(\diamond q).$$

From the above equalities and the logical characterization result (2.3) we obtain the desired result.  $\blacksquare$

**5.2. Algorithms for the metrics.** We first prove a lemma that helps to obtain reduced-complexity algorithms for concurrent games. The lemma states that the distance  $[s \preceq_1 t]$  is attained by restricting player 2 to pure moves at state  $t$ , for all states  $s, t \in S$ .

**Lemma 3.** *For all concurrent game structures  $G$  and all metrics  $d \in \mathcal{M}$ , we have,*

$$\begin{aligned} & \sup_{k \in C(d)} \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{y_2 \in \mathcal{D}_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x_1, x_2}(k) - \mathbb{E}_t^{y_1, y_2}(k)) \\ &= \sup_{k \in C(d)} \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{b \in \Gamma_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x_1, x_2}(k) - \mathbb{E}_t^{y_1, b}(k)). \end{aligned} \quad (5.1)$$

*Proof.* To prove our claim we fix  $k \in C(d)$ , and player 1 mixed moves  $x \in \mathcal{D}_1(s)$ , and  $y \in \mathcal{D}_1(t)$ . We then have,

$$\sup_{y_2 \in \mathcal{D}_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x, x_2}(k) - \mathbb{E}_t^{y, y_2}(k)) = \inf_{x_2 \in \mathcal{D}_2(s)} \mathbb{E}_s^{x, x_2}(k) - \inf_{y_2 \in \mathcal{D}_2(t)} \mathbb{E}_t^{y, y_2}(k) \quad (5.2)$$

$$= \inf_{x_2 \in \mathcal{D}_2(s)} \mathbb{E}_s^{x, x_2}(k) - \inf_{b \in \Gamma_2(t)} \mathbb{E}_t^{y, b}(k) \quad (5.3)$$

$$= \sup_{b \in \Gamma_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x, x_2}(k) - \mathbb{E}_t^{y, b}(k)),$$

where (5.3) follows from (5.2) since the decomposition on the rhs of (5.2) yields two independent linear optimization problems; the optimal values are attained at a vertex of the convex hulls of the distributions induced by pure player 2 moves at the two states. This easily leads to the result.  $\blacksquare$

We now present algorithms for metrics in concurrent games. Due to the reduction from concurrent reachability games, shown in Theorem 4, it is unlikely that we have an algorithm in NP for the metric distance between states. We therefore construct statements in the theory of real closed fields, firstly to decide whether  $[s \preceq_1 t] \leq r$ , for a rational  $r$ , so that we can approximate the metric distance between states  $s$  and  $t$ , and secondly to decide if  $[s \preceq_1 t] = 0$  in order to compute the kernel of the game simulation and bisimulation metrics.

The statements improve on the complexity that can be achieved by a direct translation of the statements of [11] to the theory of real closed fields. The complexity reduction is based on the observation that using Lemma 3, we can replace a sup operator with finite conjunction, and therefore reduce the quantifier complexity of the resulting formula. Fix a game structure  $G$  and states  $s$  and  $t$  of  $G$ . We proceed to construct a statement in the theory of reals that can be used to decide if  $[s \preceq_1 t] \leq r$ , for a given rational  $r$ .

In the following, we use variables  $x_1, y_1$  and  $x_2$  to denote a set of variables  $\{x_1(a) \mid a \in \Gamma_1(s)\}$ ,  $\{y_1(a) \mid a \in \Gamma_1(t)\}$  and  $\{x_2(b) \mid b \in \Gamma_2(s)\}$  respectively. We use  $k$  to denote the set of variables  $\{k(u) \mid u \in S\}$ , and  $d$  for the set of variables  $\{d(u, v) \mid u, v \in S\}$ . The variables  $\alpha, \alpha', \beta, \beta'$  range over reals. For convenience, we assume  $\Gamma_2(t) = \{b_1, \dots, b_l\}$ .

First, notice that we can write formulas that state that a variable  $x$  is a mixed move for a player at state  $s$ , and  $k$  is a constructible predicate (i.e.,  $k \in C(d)$ ):

$$\text{IsDist}(x, \Gamma_1(s)) \equiv \bigwedge_{a \in \Gamma_1(s)} x(a) \geq 0 \wedge \bigwedge_{a \in \Gamma_1(s)} x(a) \leq 1 \wedge \sum_{a \in \Gamma_1(s)} x(a) = 1,$$

$$\text{kBounded}(k, d) \equiv \bigwedge_{u \in S} \left[ k(u) \geq \theta_1 \wedge k(u) \leq \theta_2 \right] \wedge \bigwedge_{u, v \in S} (k(u) - k(v) \leq d(u, v)).$$

In the following, we write bounded quantifiers of the form “ $\exists x_1 \in \mathcal{D}_1(s)$ ” or “ $\forall k \in C(d)$ ” which mean respectively  $\exists x_1. \text{IsDist}(x_1, \Gamma_1(s)) \wedge \dots$  and  $\forall k. \text{kBounded}(k, d) \rightarrow \dots$ .

Let  $\eta(k, x_1, x_2, y_1, b)$  be the polynomial  $\mathbb{E}_s^{x_1, x_2}(k) - \mathbb{E}_t^{y_1, b}(k)$ . Notice that  $\eta$  is a polynomial of degree 3. We write  $a = \max\{a_1, \dots, a_l\}$  for variables  $a, a_1, \dots, a_l$  for the formula

$$(a = a_1 \wedge \bigwedge_{i=1}^l a_1 \geq a_i) \vee \dots \vee (a = a_l \wedge \bigwedge_{i=1}^l a_l \geq a_i).$$

We construct the formula for game simulation in stages. First, we construct a formula  $\Phi_1(d, s, t, k, x, \alpha)$  with free variables  $d, k, x, \alpha$  such that  $\Phi_1(d, s, t, k, x_1, \alpha)$  holds for a valuation to the variables iff

$$\alpha = \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{b \in \Gamma_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x_1, x_2}(k) - \mathbb{E}_t^{y_1, b}(k)).$$

We use the following observation to move the innermost inf ahead of the sup over the finite set  $\Gamma_2(t)$  (for a function  $f$ ):

$$\sup_{b \in \Gamma_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} f(b, x_2, x) = \inf_{x_2^{b_1} \in \mathcal{D}_2(s)} \dots \inf_{x_2^{b_l} \in \mathcal{D}_2(s)} \max(f(b_1, x_2^{b_1}, x), \dots, f(b_l, x_2^{b_l}, x)).$$

The formula  $\Phi_1(d, s, t, k, x_1, \alpha)$  is given by:

$$\begin{aligned} & \forall y_1 \in \mathcal{D}_1(t). \forall x_2^{b_1} \in \mathcal{D}_2(s) \dots x_2^{b_l} \in \mathcal{D}_2(s). \forall w_1 \dots w_l. \forall a. \forall \alpha'. \\ & \exists \hat{y}_1 \in \mathcal{D}_1(t). \exists \hat{x}_2^{b_1} \in \mathcal{D}_2(s) \dots \hat{x}_2^{b_l} \in \mathcal{D}_2(s). \exists \hat{w}_1 \dots \hat{w}_l. \exists \hat{a}. \\ & \left[ \left[ \left\{ \begin{array}{l} (w_1 = \eta(k, x_1, x_2^{b_1}, y_1, b_1)) \\ \quad \wedge \dots \wedge \\ (w_l = \eta(k, x_1, x_2^{b_l}, y_1, b_l)) \end{array} \right\} \wedge (a = \max\{w_1, \dots, w_l\}) \right] \rightarrow (a \geq \alpha) \right] \wedge \\ & \left[ \left[ \left\{ \begin{array}{l} (\hat{w}_1 = \eta(k, x_1, \hat{x}_2^{b_1}, \hat{y}_1, b_1)) \\ \quad \wedge \dots \wedge \\ (\hat{w}_l = \eta(k, x_1, \hat{x}_2^{b_l}, \hat{y}_1, b_l)) \end{array} \right\} \wedge (\hat{a} = \max\{\hat{w}_1, \dots, \hat{w}_l\} \wedge \hat{a} \geq \alpha'(s, t)) \right] \rightarrow (\alpha \geq \alpha') \right]. \end{aligned}$$

Using  $\Phi_1$ , we construct a formula  $\Phi(d, s, t, \alpha)$  with free variables  $d \in \mathcal{M}$  and  $\alpha \in \mathcal{M}$  such that  $\Phi(d, s, t, \alpha)$  is true iff:

$$\alpha = \sup_{k \in C(d)} \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{b \in \Gamma_2(t)} \inf_{x_2 \in \mathcal{D}_2(s)} (\mathbb{E}_s^{x_1, x_2}(k) - \mathbb{E}_t^{y_1, b}(k)).$$

The formula  $\Phi$  is defined as follows:

$$\forall k \in C(d). \forall x_1 \in \mathcal{D}_1(s). \forall \beta. \forall \alpha'. \left[ \begin{array}{l} \Phi_1(d, s, t, k, x_1, \beta) \rightarrow (\beta(s, t) \leq \alpha) \wedge \\ (\forall k' \in C(d). \forall x_1' \in \mathcal{D}_1(s). \forall \beta'. \Phi_1(d, s, t, k', x_1', \beta') \wedge \beta'(s, t) \leq \alpha') \rightarrow \alpha \leq \alpha' \end{array} \right]. \quad (5.4)$$

Finally, given a rational  $r$ , we can check if  $[s \preceq_1 t] \leq r$  by checking if the following sentence is true:

$$\exists d \in \mathcal{M}. \exists a \in \mathcal{M}. \left[ \bigwedge_{u, v \in S} \Phi(d, u, v, a(u, v)) \wedge (d(u, v) = a(u, v)) \wedge (d(s, t) \leq r) \right]. \quad (5.5)$$

The above sentence is true iff in the least fixpoint,  $d(s, t)$  is bounded by  $r$ . Like in the case of turn-based games and MDPs, given a rational  $\epsilon > 0$ , using binary search and  $\mathcal{O}(\log(\frac{\theta_2 - \theta_1}{\epsilon}))$  calls to a decision procedure to check the sentence (5.5), we can compute an interval  $[l, u]$  with  $u - l \leq \epsilon$ , such that  $[s \preceq_1 t] \in [l, u]$ .

**Complexity.** Note that  $\Phi$  is of the form  $\forall \exists \forall$ , because  $\Phi_1$  is of the form  $\forall \exists$ , and appears in negative position in  $\Phi$ . The formula  $\Phi$  has  $(|S| + |\Gamma_1(s)| + 3)$  universally quantified variables, followed by  $(|S| + |\Gamma_1(s)| + 3 + 2(|\Gamma_1(t)| + |\Gamma_2(s)| \cdot |\Gamma_2(t)| + |\Gamma_2(t)| + 2))$  existentially quantified

variables, followed by  $2(|\Gamma_1(t)| + |\Gamma_2(s)| \cdot |\Gamma_2(t)| + |\Gamma_2(t)| + 1)$  universal variables. The sentence (5.5) introduces  $|S|^2 + |S|^2$  existentially quantified variables ahead of  $\Phi$ . The matrix of the formula is of length at most quadratic in the size of the game, and the maximum degree of any polynomial in the formula is 3. We define the size of a game  $G$  as:  $|G| = |S| + T$ , where  $T = \sum_{s,t \in S} \sum_{a,b \in \text{Moves}} |\delta(s, a, b)(t)|$ . Using the complexity of deciding a formula in the theory of real closed fields [3], which states that a formula with  $i$  quantifier blocks, where each block has  $l_i$  variables, of  $p$  polynomials, has a time complexity bound of  $\mathcal{O}(p^{\mathcal{O}(\prod(l_i+1))})$ , we get the following result.

**Theorem 5.1.** (*Decision complexity for exact distance*). *For all concurrent game structures  $G$ , given a rational  $r$ , and two states  $s$  and  $t$ , whether  $[s \preceq_1 t] \leq r$  can be decided in time  $\mathcal{O}(|G|^{\mathcal{O}(|G|^5)})$ .*

**Approximation.** Given a rational  $\epsilon > 0$ , using binary search and  $\mathcal{O}(\log(\frac{\theta_2 - \theta_1}{\epsilon}))$  calls to check the formula 5.5, we can obtain an interval  $[l, u]$  with  $u - l \leq \epsilon$  such that  $[s \preceq_1 t]$  lies in the interval  $[l, u]$ .

**Corollary 3. (Approximation for exact distance).** *For all concurrent game structures  $G$ , given a rational  $\epsilon$ , and two states  $s$  and  $t$ , an interval  $[l, u]$  with  $u - l \leq \epsilon$  such that  $[s \preceq_1 t] \in [l, u]$  can be computed in time  $\mathcal{O}(\log(\frac{\theta_2 - \theta_1}{\epsilon}) \cdot |G|^{\mathcal{O}(|G|^5)})$ .*

In contrast, the formula to check whether  $[s \preceq_1 t] \leq r$ , for a rational  $r$ , as implied by the definition of  $H_{\preceq_1}(d)(s, t)$ , that does not use Lemma 3, has five quantifier alternations due to the inner sup, which when combined with the  $2 \cdot |S|^2$  existentially quantified variables in the sentence (5.5), yields a decision complexity of  $\mathcal{O}(|G|^{\mathcal{O}(|G|^7)})$ .

**5.3. Computing the kernels.** Similar to the case of turn-based games and MDPs, the kernel of the simulation metric  $\preceq_1$  for concurrent games can be computed as the limit of the series  $\preceq_1^0, \preceq_1^1, \preceq_1^2, \dots$ , of relations. For all  $s, t \in S$ , we have  $(s, t) \in \preceq_1^0$  iff  $s \equiv t$ . For all  $n \geq 0$ , we have  $(s, t) \in \preceq_1^{n+1}$  iff the following sentence  $\Phi_s$  is true:

$$\forall a. \Phi(d^n, s, t, a) \rightarrow a = 0,$$

where  $\Phi$  is defined as in (5.4) and at step  $n$  in the iteration, the distance between any pair of states  $u, v \in S$  is defined as follows,

$$\forall u, v \in S. d^n(u, v) = \begin{cases} 0 & \text{if } (s, t) \in \preceq_1^n \\ 1 & \text{if } (s, t) \notin \preceq_1^n \end{cases}.$$

To compute the bisimulation kernel, we again proceed by partition refinement. For a set of partitions  $\mathcal{Q}^0, \mathcal{Q}^1, \dots$ , where  $(s, t) \in V$  for  $V \in \mathcal{Q}^n$  implies  $(s, t) \in \simeq_1^n$ ,  $(s, t) \in \simeq_1^{n+1}$  iff the following sentence  $\Phi_b$  is true for the state pairs  $(s, t)$  and  $(t, s)$ :

$$\forall a. \Phi(d^n, s, t, a) \rightarrow a = 0,$$

where  $\Phi$  is again as defined in (5.4) and at step  $n$  in the iteration, the distance between any pair of states  $u, v \in S$  is defined as follows,

$$\forall u, v \in S. d^n(u, v) = \begin{cases} 0 & \text{if } (s, t) \in \simeq_1^n \\ 1 & \text{if } (s, t) \notin \simeq_1^n \end{cases}.$$



**Complexity.** In the worst case we need  $\mathcal{O}(|S|^2)$  partition refinement steps for computing both the simulation and the bisimulation relation. At each partition refinement step the number of state pairs we consider is bounded by  $\mathcal{O}(|S|^2)$ . We can check if  $\Phi_s$  and  $\Phi_b$  are true using a decision procedure for the theory of real closed fields. Therefore, we need  $\mathcal{O}(|S|^4)$  decisions to compute the kernels. The partitioning of states based on the decisions can be done by any of the partition refinement algorithms, such as [25].

**Theorem 5.2.** *For all concurrent game structures  $G$ , states  $s$  and  $t$ , whether  $s \preceq_1 t$  can be decided in  $\mathcal{O}(|G|^{\mathcal{O}(|G|^3)})$  time, and whether  $s \preceq_g t$  can be decided in  $\mathcal{O}(|G|^{\mathcal{O}(|G|^3)})$  time.*

## 6. CONCLUSION: POSSIBLE APPLICATIONS AND OPEN PROBLEMS

We have shown theoretical applications of game metrics with respect to discounted and long-run average values of games. An interesting question regarding game metrics is related to their usefulness in real-world applications. We now discuss possible applications of game metrics.

- *State space reduction.* The kernels of the metrics are the simulation and bisimulation relations. These relations have been well studied in the context of transition systems with applications in program analysis and verification. For example, in [19] the authors show that bisimulation based state space reduction is practical and may result in an enormous reduction in model size, speeding up model checking of probabilistic systems.
- *Security.* Bisimulation plays a critical role in the formal analysis of security protocols. If two instances of a protocol, parameterized by a message  $m$ , are bisimilar for messages  $m$  and  $m'$ , then the messages remain secret [7]. The authors use bisimulation in probabilistic transition systems to analyze probabilistic anonymity in security protocols.
- *Computational Biology.* In the emerging area of computational systems biology, the authors of [30] use the metrics defined in the context of probabilistic systems [12, 32, 33] to compare reduced models of *Stochastic Reaction Networks*. These reaction networks are used to study intra-cellular behavior in computational systems biology. The reduced models are Continuous Time Markov Chains (CTMCs), and the comparison of different reduced models is via the metric distance between their initial states. A central question in the study of intra-cellular behavior is estimating the sizes of populations of various species that cohabit cells. The inter-cellular dynamics in this context is modeled as a stochastic process, representing the temporal evolution of the species' populations, represented by a family  $(X(t))_{t \geq 0}$  of random vectors. For  $0 \leq i < N$ ,  $N$  being the number of different species,  $X_i(t)$  is the population of species  $S_i$  at time  $t$ . In [26], the authors show how CTMCs that model system dynamics can be reduced to Discrete Time Markov Chains (DTMCs) using a technique called uniformization or discrete-time conversion. The DTMCs are stochastically identical to the CTMCs and enable more efficient estimation of species' populations. An assumption that is made in these studies is that systems are spatially homogeneous and thermally equilibrated; the molecules are well stirred in a fixed volume at a constant temperature. These assumptions enable the reduction of these systems to CTMCs and to DTMCs in some cases.

In the applications we have discussed, non-determinism is modeled probabilistically. In applications where non-determinism needs to be interpreted demonically, rather than probabilistically, MDPs or turn-based games would be the appropriate framework for analysis. If the interaction between various sources of non-determinism needs to be modeled simultaneously, then concurrent games would be the appropriate framework for analysis. For the analysis of these general models, our results and algorithms will be useful.

**Open Problems.** While we have shown polynomial time algorithms for the kernel of the simulation and bisimulation metrics for MDPs and turn-based games, the existence of a polynomial time algorithm for the kernel of both the simulation and bisimulation metrics for concurrent games is an open problem. The existence of a polynomial time algorithm to approximate the exact metric distance in the case of turn-based games and MDPs is an open problem. The existence of a PSPACE algorithm for the decision problem of the exact metric distance in concurrent games is an open problem.

**Acknowledgments.** The first, second and fourth author were supported in part by the National Science Foundation grants CNS-0720884 and CCR-0132780. The third author was supported in part by the National Science Foundation grants CCF-0427202, CCF-0546170. We would like to thank the reviewers for their detailed comments that helped us make the paper better.

## REFERENCES

- [1] R. Alur, T.A. Henzinger, and O. Kupferman. Alternating time temporal logic. *J. ACM*, 49:672–713, 2002.
- [2] C. Baier. Polynomial time algorithms for testing probabilistic bisimulation and simulation. In *CAV*, pages 50–61. Springer-Verlag, 1996.
- [3] S. Basu. New results on quantifier elimination over real closed fields and applications to constraint databases. *J. ACM*, 46(4):537–555, 1999.
- [4] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995. Volumes I and II.
- [5] M.C. Browne, E.M. Clarke, and O. Grumberg. Characterizing finite Kripke structures in propositional temporal logic. *Theoretical Computer Science*, 59:115–131, 1988.
- [6] J. F. Canny. Some algebraic and geometric computations in pspace. In *STOC*, pages 460–467. ACM Press, 1988.
- [7] K. Chatzikokolakis, G. Norman, and D. Parker. Bisimulation for demonic schedulers. In L. de Alfaro, editor, *12th International Conference on Foundations of Software Science and Computation Structures (FOSSACS'09)*, volume 5504 of *LNCS*, pages 318–332. Springer, 2009.
- [8] L. de Alfaro, T.A. Henzinger, and O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.
- [9] L. de Alfaro, T.A. Henzinger, and R. Majumdar. Discounting the future in systems theory. In *Proc. 30th Int. Colloq. Aut. Lang. Prog.*, volume 2719 of *Lect. Notes in Comp. Sci.*, pages 1022–1037. Springer-Verlag, 2003.
- [10] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004.
- [11] L. de Alfaro, R. Majumdar, V. Raman, and M. Stoelinga. Game relations and metrics. In *LICS*, pages 99–108. IEEE Computer Society Press, 2007.
- [12] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov systems. In *CONCUR*, pages 258–273. Springer-Verlag, 1999.
- [13] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labelled markov processes. *Information and Computation*, 2002.
- [14] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *LICS*, pages 413–422. ACM Press, 2002.

- [15] K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *ICALP (2)*, pages 324–335. Springer-Verlag, 2006.
- [16] K. Etessami and M. Yannakakis. On the complexity of Nash equilibria and other fixed points (extended abstract). In *FOCS*, pages 113–123. IEEE Computer Society Press, 2007.
- [17] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.
- [18] M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *STOC*, pages 10–22. ACM Press, 1976.
- [19] Joost-Pieter Katoen, Tim Kemna, Ivan S. Zapreev, and David N. Jansen. Bisimulation minimisation mostly speeds up probabilistic model checking. In *TACAS*, pages 87–101, 2007.
- [20] D. Kozen. A probabilistic PDL. In *Proc. 15th ACM Symp. Theory of Comp.*, pages 291–297, 1983.
- [21] D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
- [22] A. McIver and C. Morgan. *Abstraction, Refinement, and Proof for Probabilistic Systems*. Monographs in Computer Science. Springer-Verlag, 2004.
- [23] J.F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10:53–66, 1981.
- [24] R. Milner. Operational and algebraic semantics of concurrent processes. In *Handbook of Theoretical Computer Science*, volume B, pages 1202–1242. Elsevier Science Publishers, 1990.
- [25] R. Paige and R. E. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16(6):973–989, 1987.
- [26] Werner Sandmann and Verena Wolf. Computational probability for systems biology. In *FMSB*, pages 33–47, 2008.
- [27] R. Segala and N.A. Lynch. Probabilistic simulations for probabilistic processes. In *CONCUR*, pages 481–496. Springer-Verlag, 1994.
- [28] L.S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.
- [29] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics.*, 8:171–176, 1958.
- [30] D. Thorsley and E. Klavins. A theory of approximation for stochastic biochemical processes. to appear in B. Ingalls and P. Iglesias, eds., *Control-Theoretic Approaches to Systems Biology*. MIT Press, 2009.
- [31] F. van Breugel, B. Sharma, and J. Worrell. Approximating a behavioural pseudometric without discount for probabilistic systems. *CoRR*, abs/0803.3796, 2008.
- [32] F. van Breugel and J. Worrell. An algorithm for quantitative verification of probabilistic transition systems. In *CONCUR*, pages 336–350. Springer-Verlag, 2001.
- [33] F. van Breugel and J. Worrell. Towards quantitative verification of probabilistic transition systems. In *ICALP*, pages 421–432. Springer-Verlag, 2001.
- [34] Y. Ye. Improved complexity results on solving real-number linear feasibility problems. *Math. Program.*, 106(2):339–363, 2006.
- [35] L. Zhang and H. Hermanns. Deciding simulations on probabilistic automata. In *ATVA*, pages 207–222. Springer-Verlag, 2007.