Theoretical Computer Science 680 (2017) 25–35

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Selfish cops and passive robber: Qualitative games

Ath. Kehagias *, G. Konstantinidis

ARTICLE INFO

Article history:
Received 21 November 2016
Received in revised form 7 April 2017
Accepted 10 April 2017
Available online 24 April 2017
Communicated by J. van den Herik

Keywords:
Cops and robbers
Pursuit evasion
Game theory

ABSTRACT

Several variants of the *cops and robbers* (CR) game have been studied in the literature. In this paper we examine a novel variant, which is played between two cops, each one *independently* trying to catch a "passive robber". We call this the *Selfish Cops and Passive Robber* (*SCPR*) game. In short, SCPR is a stochastic two-player, zero-sum game where *the opponents are the two cop players*. We study sequential and concurrent versions of the SCPR game. For both cases we prove the existence of value and optimal strategies and present algorithms for the computation of these.

1. Introduction

Several variants of the *cops and robbers* (CR) game have been studied in the past. In this paper we examine a novel variant, which is played between two cops, each one *independently* trying to catch a "passive robber". We call this the *Selfish Cops and Passive Robber* (*SCPR*) game. Here is a brief and informal description of the game (a more detailed description will be provided in Section 2).

1. The game is played on an undirected, finite, simple and connected graph.
2. The game is played by two cop players C_1 and C_2, each controlling a *cop token* (the tokens will also be referred to as C_1 and C_2).
3. A *robber token* R is also used, which is moved according to a (deterministic or random) law known to both cop players.
4. At every turn of the game the tokens are moved from vertex to vertex, along the edges of the graph.
5. The winner is the first player whose token lies at the same vertex as the robber token (that is, the player who "captures the robber").

We emphasize that SCPR is a game played between *two cop players*; the robber is not associated to a player and *does not attempt to evade capture*. As far as we know this CR variant has not been previously studied.

As an example, consider two cops located at opposite ends of a path with N vertices and one robber located at some intermediate vertex. A cop wins if he reaches the robber before the other cop. We repeat that the robber is not actively trying to evade; his every move is governed by a probability law conditioned on the current positions of himself and the cops. This law is known to the cops. For each cop, what is the optimal strategy and what is the corresponding win probability?

* Corresponding author.
E-mail address: kehagiat@gmail.com (A. Kehagias).

http://dx.doi.org/10.1016/j.tcs.2017.04.004
0304-3975/

1. The simplest case is when the robber is stationary. In this case, the cop closest to him wins with probability one; each cop's optimal strategy is to always move towards the robber.
2. A slightly harder case is that of the "drunk robber" [16–18,20] who performs a random walk on the graph. The optimal cop strategies are the same as in the stationary robber case, but the computation of the win probabilities is not trivial (in Section 4 we provide an algorithm to compute them).
3. Examples can also be constructed in which the optimal cop strategy is *not* to always move towards the robber (one such example is given in Section 4.2).

The above examples should give a clearer idea of the problem which we study in the current paper. Of course our results are not limited to paths, but hold for general graphs.

The study of Cops and Robbers was initiated by Quilliot [27] and Nowakowski and Winkler [24] and an extensive graph theoretic literature exists on the problem (we will provide some references a little later). However, we study SCPR from a somewhat different angle, using the theory of *stochastic games* as presented in the book by Filar and Vrieze [13] (a stochastic game consists of a sequence of one-shot games where the game played at any time depends probabilistically on the previous game played and the actions of the agents in that game). In the current paper we study *qualitative* SCPR games, in which the *payoff* is the *winning probability*; in a forthcoming paper we will discuss *quantitative* SCPR games, in which the payoff is the *expected capture time*. The concepts of qualitative and quantitative *pursuit* games (a special category of which are CR games) have been introduced in [15] under the names "game of kind" and "game of degree". Informally, in a qualitative game a late capture is as good as a fast one; in a quantitative game the pursuer wants to capture the evader as soon as possible.

For the graph theoretic point of the view the reader can consult the recent book [25] which contains a good overview of the extensive literature. As already mentioned, this literature is mainly oriented to graph theoretic and combinatorial considerations. Indeed CR can be seen as a *combinatorial game*, as pointed out in [7,8]. On the topic of combinatorial games the reader can consult the introductory text [1] as well as the classic book (in four volumes) [5] by Berlekamp and Conway.

We believe that game theory offers a natural (but not often used in the "mainstream" CR literature) framework for the analysis of CR games. In particular, as already mentioned, we consider SCPR as a stochastic game. Stochastic games were introduced by Shapley [28]. A classic book on the subject is [13], which also contains a rich bibliography; see also [23].

Several game theoretic models can be applied to the study of CR games. For instance, as will be seen in Section 2, SCPR is a *recursive game* [12]: as soon as a non-zero-payoff is received the play moves to an absorbing state. This point of view can be applied to classical CR games as well.

Let us also mention a construction which has been used in several "classic" CR papers [7,8,14]. Suppose that a "classic" CR game is played between one cop and one robber on the undirected graph $G=(V,E)$. We now construct the *game digraph* $D=(S,A)$, where the vertex set is $S=V\times V\times\{1,2\}$ (with $i\in\{1,2\}$ denoting the player who has the next move) and the arc set A encodes possible vertex-to-vertex transitions. Then a play of the CR game can be understood as a walk on D; the cop wins if he can force the walk to pass through a vertex of the form (x,x,i). Hence CR can be seen as a game in which the two players push a token along the arcs of the digraph. As pointed out in [7,8] many CR variants and several other pursuit games on graphs (including their concurrent versions) can be formulated in a similar manner. It turns out that such "digraph games" have been studied by several researchers and the related literature is spread among many communities. The earliest such works of which we are aware are [4,22]. Related examples appear in [3,11,29]. But probably the most widespread application of this point of view appears in the literature of *reachability games* [6] and, more generally, *ω-regular games* [21]. In a reachability game two players take turns moving a token along the arcs of a digraph; player 1 wants to place the token on one of the nodes of a subset of the digraph vertices while player 2 wants to avoid this event. In addition to "classic" sequential reachability games, many other variants have been studied, for example, stochastic [9], concurrent [2], n-player [10] etc.

All of the above approaches find immediate application to both classical CR games and selfish cops variants, such as the one presented in the current paper.

The paper is organized as follows. In Section 2 we present definitions and notation which will be used in the rest of the paper. In Section 3 we study the *sequential* version of the SCPR game and in Section 4 the *concurrent* version. In both sections we obtain analogous results; namely both the sequential and concurrent SCPR game have a value, Cop 1 has an ε-optimal deterministic stationary Markovian strategy and Cop 2 has an optimal deterministic stationary Markovian strategy; furthermore we give algorithms which compute values and strategies efficiently; the algorithm for concurrent SCPR is somewhat more complicated but it can be simplified in case robber movement is governed by an "oblivious deterministic" law. Finally, in Section 5 we present concluding remarks and discuss future research directions.

2. Preliminaries

The SCPR game is played on an undirected, finite, simple and connected graph $G=(V,E)$, where V is the vertex set and E is the edge set. Unless otherwise stated, we will assume that the *cop number* $c(G)$ of the graph equals one (recall that $c(G)$ is the minimum number of cops required to guarantee capture of the robber on G).

The game proceeds in turns numbered by $t\in\mathbb{N}_0=\{0,1,2,\ldots\}$ and, as already mentioned, involves three tokens: C_1, C_2 and R. These will also be referred to as the first, second and third token, respectively, and their locations *at the end* of the t-th turn are indicated by X_t^1, X_t^2, X_t^3. The *starting position* at the 0-th turn is *given*: for $i \in \{1,2,3\}$, $X_0^i = x_0^i \in V$. In subsequent turns, the positions are changed according to the rules of the particular variant (sequential or concurrent) and subject to the constraint that movement always follows the graph edges: $X_{t+1}^i \in N\left[X_t^i\right]$ (the *closed* neighborhood of X_t^i). As will be seen in the sequel, in the general case token moves are governed by *probabilistic strategies*; hence X_t^1, X_t^2, X_t^3 are random variables.

2.1. Sequential SCPR

In the *sequential* version of SCPR players take turns in moving their tokens. More specifically, on odd-numbered turns C_1 is moved by the first cop player; on even-numbered turns first C_2 is moved by the second cop player and then R is moved according to a (deterministic or random) law known to both cop players. Consequently, for $t = 2l+1$ we have $X_t^2 = X_{t-1}^2$ and $X_t^3 = X_{t-1}^3$; for $t = 2l$ we have $X_t^1 = X_{t-1}^1$. An additional sequence of variables U_0, U_1, U_2, U_3, ... indicates the player to move in the next turn; in other words, $U_0 = U_2 = ... = 1$, $U_1 = U_3 = ... = 2$. We also define the vector $S_t = \left(X_t^1, X_t^2, X_t^3, U_t\right)$.

A *game position* or *state* is a vector $s = \left(x^1, x^2, x^3, u\right)$ where x^1, x^2, x^3 are the positions of the three tokens and u indicates which cop is about to play. For instance, $s = (2,3,5,1)$ denotes the situation in which C_1, C_2, R are located at vertices 2, 3 and 5, respectively, and C_1 will move in the next turn. We define the following sets of states

$$\text{for all } i \in \{1,2\} : \mathbf{S}_i = V \times V \times V \times \{i\} .$$

In other words, $\mathbf{S}_1$ (resp. $\mathbf{S}_2$) is the set of states “belonging” to the first (resp. second) player. A C_1-*capture state* is an $s = \left(x^1, x^2, x^3, u\right)$ such that $x^1 = x^3$. A C_2-*capture state* is an $s = \left(x^1, x^2, x^3, u\right)$ such that $x^2 = x^3$ and $x^1 \neq x^3$. We see that C_1 is slightly favored, since an (x,x,x,u) state is considered a C_1 capture; however, because of symmetry, reversing the definitions of C_i-captures would yield essentially the same results. We will also use a *terminal* state, denoted by τ; the behavior of the terminal state will be described in detail a little later. At any rate, the full *state space* of the sequential SCPR game is

$$\mathbf{S} = \mathbf{S}_1 \cup \mathbf{S}_2 \cup \{\tau\} .$$

The random variable A_t^i denotes the *move* (or *action*) of the i-th token at time t. When the game state is s, the set of moves available to the i-th token is denoted by $\mathbf{A}_i(s)$. For instance, when $s = \left(x^1, x^2, x^3, 1\right)$ we have $\mathbf{A}_1(s) = N\left[x^1\right]$ (the *closed* neighborhood of x^1) and $\mathbf{A}_2(s) = \left\{x^2\right\}$. Similar things hold for states $s = \left(x^1, x^2, x^3, 2\right)$. For $s = \tau$ we have $\mathbf{A}_i(s) = \{\lambda\}$, where λ is the *null move*. Legal moves result to “normal” state transitions; for example, suppose the current state is $s = (2,3,5,1)$ and the next moves are $a^1 = 3$, $a^2 = 3$, $a^3 = 5$; then, assuming $3 \in N[2]$, the next state is $s' = (3,3,5,2)$. However, the terminal state τ raises the following exceptions.

1. If the current state s is a C_i-capture state ($i \in \{1,2\}$), then the next state is $s' = \tau$, irrespective of the token moves. In other words, a capture state always transits to the terminal state.
2. If the current state s is the terminal (that is, $s = \tau$), then the next state is $s' = \tau$ irrespective of the token moves. In other words, the terminal always transits to itself.

A *play* or *infinite history* of the SCPR game is an infinite sequence $s_0 s_1 s_2 ... s_n ...$ of game states. The set of all infinite histories is denoted by

$$H^\infty = \{s_0 s_1 s_2 ... s_t ... : s_t \in S \text{ for } t \in \mathbb{N}_0\} .$$

A *finite history* is a sequence $s_0 s_1 s_2 ... s_n$ of game states; the set of all histories of length n is denoted by

$$H_n = \{s_0 s_1 s_2 ... s_{n-1} : s_t \in S \text{ for } t \in \{0, 1, ..., n-1\}\} ;$$

the set of all finite histories is $H = \bigcup_{n=0}^{\infty} H_n$.

We have already mentioned that each cop player moves his respective token. Rather than specifying each move separately, we assume (as is usual in Game Theory) that before the game starts, each cop player selects a *strategy* which controls all subsequent moves. Despite the fact that there is no robber player, we will assume that robber movement is also controlled by a “strategy”, which has been fixed before the game starts and *is known to the cop players*. Hence the i-th token ($i \in \{1,2,3\}$) is controlled by the strategy (conditional probability function):

$$\sigma_i(a|s_0 s_1 ... s_t) = \Pr\left(A_{t+1}^i = a | \left(X_0^1, X_0^2, X_0^3, U_0\right) = s_0, ..., \left(X_t^1, X_t^2, X_t^3, U_t\right) = s_t\right) .$$

The above definition is sufficiently general to describe every possible manner of move selection. We will only consider strategies which assign zero probability to illegal moves. The following classes of strategies are of particular interest.

1. A strategy σ_i is called *stationary Markovian* (or *positional*) if and only if $\sigma_i(a|s_0s_1...s_t) = \sigma_i(a|s_t)$; that is, the probability of the next move depends only on the current *state* of the game.
2. A strategy σ_i is called *oblivious* if and only if it is stationary Markovian and $\sigma_i\left(a|\left(y^1, y^2, y^3, u\right)\right) = \sigma_i\left(a|y^i, u\right)$; that is, the probability of the next move of the token depends only on (i) the current *location* of the token and (ii) the active player.
3. A strategy σ_i is called *deterministic* if and only if, for every $s_0s_1...s_t \in H$, $\sigma_i(x|s_0s_1...s_t) \in \{0, 1\}$; hence for every history, the i-th token moves to its next location deterministically.

To simplify presentation, we will often use the following notation for deterministic strategies. We define the *deterministic strategy* to be a function $\overline{\sigma}_i : H \to V$, defined as follows: for every finite history $s_0s_1...s_t$, $\overline{\sigma}_i(s_0s_1...s_t) = a$, where a is the unique vertex such that $\sigma_i(a|s_0s_1...s_t) = 1$. If σ_i is stationary Markovian then we write $\overline{\sigma}_i(s_t) = a$.

In all subsequent notation, the dependence on the fixed and known σ_3 is suppressed.

Suppose the game is in state s. Now C_1 plays a^1, C_2 plays a^2 and R's move a^3 is selected according to the (fixed) strategy σ_3; hence the game will move into some new state s' with a certain probability depending on a_1, a_2 and σ_3. We denote this probability by $\Pr\left(s'|s, a_1, a_2\right)$. Note that, when a cop reaches the vertex occupied by the robber we have a capture with probability one, irrespective of the robber's move.

Payoff is defined as follows. In each turn of the game, C_1 receives an *immediate payoff* equal to

$$q(s) = \begin{cases} 1 & \text{if and only if } s \text{ is a } C_1\text{-capture state,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

C_2 receives $-q(s)$. Hence, a play $s_0s_1....$ results in (total) *payoff*

$$Q(s_0s_1....) = \sum_{t=0}^{\infty} q(s_t) \tag{2}$$

for C_1 and $-Q(s_0s_1....)$ for C_2. Note that both players have an incentive to capture R.

1. If C_1 captures the robber, he receives a total payoff of one (comprising of immediate payoff of one for the capture turn and zero for all other turns); otherwise his total payoff is zero.
2. C_2 *never* receives positive payoff (even if he captures the robber). However, we have assumed $c(G) = 1$ and this implies that a single cop can always catch the robber. Hence, if C_2 does not capture R, then C_1 will and thus C_2 will receive a negative payoff; this provides the incentive for C_2 to capture R.

Sequential SCPR is a *stochastic zero sum game* [13]. Each player will try to maximize his *expected* payoff. Suppose the game starts at position s_0, C_i moves according to strategy σ_i (for $i \in \{1, 2\}$) and R moves according to a fixed and known strategy σ_3. Every triple $(\sigma_1, \sigma_2, \sigma_3)$ induces a probability measure on H^∞, the set of all infinite game histories. Hence the *expected* payoff to C_1 is

$$J(\sigma_1, \sigma_2|s_0) = \mathbb{E}\left(\sum_{t=0}^{\infty} q(s_t) \,|\, \left(X_0^1, X_0^2, X_0^3, U_0\right) = s_0\right) \tag{3}$$

and is well defined; $-J(\sigma_1, \sigma_2|s_0)$ is the expected payoff to C_2. It is easily seen that

$$J(\sigma_1, \sigma_2|s_0) = \Pr\left(\text{“}C_1 \text{ wins”}|\text{“the game starts at } s_0 \text{ and, for } i \in \{1, 2\}, C_i \text{ uses } \sigma_i\text{”}\right).$$

We always have

$$\sup_{\sigma_1} \inf_{\sigma_2} J(\sigma_1, \sigma_2|s_0) \le \inf_{\sigma_2} \sup_{\sigma_1} J(\sigma_1, \sigma_2|s_0); \tag{4}$$

if the two sides of (4) are equal, then we define the *value* of the game (when started at s_0) to be

$$v(s_0) = \sup_{\sigma_1} \inf_{\sigma_2} J(\sigma_1, \sigma_2|s_0) = \inf_{\sigma_2} \sup_{\sigma_1} J(\sigma_1, \sigma_2|s_0). \tag{5}$$

We will denote the vector of values for all starting states by $\mathbf{v}$; in other words, $\mathbf{v} = (v(s))_{s \in \mathbf{S}}$. Given some $\varepsilon \ge 0$, we say that:

1. a strategy $\sigma_1^\#$ is ε-*optimal* (for C_1) if and only if $\forall s_0 : v(s_0) - \inf_{\sigma_2} J\left(\sigma_1^\#, \sigma_2|s_0\right) \le \varepsilon$;
2. a strategy $\sigma_2^\#$ is ε-*optimal* (for C_2) if and only if $\forall s_0 : v(s_0) - \sup_{\sigma_1} J\left(\sigma_1, \sigma_2^\#|s_0\right) \ge -\varepsilon$.

A 0-optimal strategy is also simply called *optimal*.

Finally, let Γ be a *matrix game*: a (one-shot) two-player, zero-sum game with finite action set $\mathbf{A}_i$ for the i-th player and the payoff to the first player being $\Gamma\left(a^1, a^2\right)$ when i-th player plays $a^i \in \mathbf{A}_i$ (with $i \in \{1, 2\}$). As is well known [26], such a game always has a *value*, which we will denote by $\mathbf{Val}\left[\Gamma\left(a^1, a^2\right)\right]$.

2.2. Concurrent SCPR

Most of the CR literature studies sequential versions of the CR game. However, we have recently introduced a *concurrent* version of the classic CR [19]. Now we extend concurrency to the SCPR game.

The concurrent SCPR game differs from the sequential game in a basic aspect: in every turn the C_1, C_2, R tokens are moved *simultaneously* (hence, when making his move, each player does not know the other player's move; note that both of them know the probability of each possible R next move, since σ_3 is known in advance). Once again we will assume, unless otherwise indicated, that $\widehat{c}(G)=1$ (note that a graph G has *concurrent cop number* $\widehat{c}(G)=k$ if and only if it has sequential (in other words "classic") cop number $c(G)=k$ [19]).

In addition, in concurrent SCPR we can have "*en-passant capture*", in which a cop and the robber start at opposite ends of the same edge and move in opposite directions; in this case the robber is "swept" by the cop and moved into the cop's destination; in such a case the capture is credited to the sweeping cop (there is one exception: if the sweeping move results in a C_1-capture position (x,x,x), then the capture is credited to C_1 even when the sweeping cop is C_2).

With concurrent movement, game states are vectors $\left(x^1,x^2,x^3\right)$ where $x^i \in V$ indicates (as previously) the position of the i-th token; the u variable is no longer necessary, since all tokens are moved in every turn. Capture states now have the form $\left(x^1,x^2,x^3\right)$ with either $x^1=x^3$ or $x^2=x^3$ (or both) and the definition and behavior of the terminal state τ are the same as previously. For the state space, we define

$$\widehat{\mathbf{S}}_a = V \times V \times V, \quad \widehat{\mathbf{S}} = \widehat{\mathbf{S}}_a \cup \{\tau\}$$

and $\widehat{\mathbf{S}}$ is the full *state space* of the of concurrent SCPR game.

Regarding $\mathbf{A}_i(s)$ (the actions available to the i-th player when in game state s) we always have $\mathbf{A}_i\left(\left(x^1,x^2,x^3\right)\right) \in N\left[x^i\right]$. The definitions of (finite and infinite) histories and strategies are the same as in the sequential case, except that we now use the state space $\widehat{\mathbf{S}}$. The meaning of the sets $\widehat{H}_n$, $\widehat{H}$, $\widehat{H}^\infty$ is analogous to that of H_n, H, H^∞. The strategies σ_i $(i \in \{1,2,3\})$ are defined in the same manner as in the sequential case (again, for deterministic moves we introduce the deterministic strategy functions $\overline{\sigma}_i$).

Payoff of the concurrent SCPR game is defined in exactly the same manner as in the sequential case. Again, concurrent SCPR is a stochastic zero sum game and each player will try to maximize his expected payoff.

3. Results for the sequential SCPR

In this section we establish that sequential SCPR has a *value* which can be computed by *value iteration*.

Theorem 3.1. *Given some graph $G=(V,E)$. For every $s \in \mathbf{S}_1 \cup \mathbf{S}_2$, the sequential SCPR game starting at s has a value $v(s)$. The vector of values $\mathbf{v}=(v(s))_{s\in\mathbf{S}}$ is the smallest (componentwise) solution of the following* optimality equations*:*

$$v(\tau)=0; \tag{6}$$

$$\text{for all } s=\left(x^1,x^2,x^3,1\right)\in\mathbf{S}_1: \quad v(s)=\max_{a^1}\left[q(s)+\sum_{s'\in\mathbf{S}}\Pr\left(s'|s,a^1,x^2\right)v\left(s'\right)\right]; \tag{7}$$

$$\text{for all } s=\left(x^1,x^2,x^3,2\right)\in\mathbf{S}_2: \quad v(s)=\min_{a^2}\left[q(s)+\sum_{s'\in\mathbf{S}}\Pr\left(s'|s,x^1,a^2\right)v\left(s'\right)\right]. \tag{8}$$

Furthermore C_2 has a deterministic stationary Markovian optimal strategy and, for every $\varepsilon>0$, C_1 has a deterministic stationary Markovian ε-optimal strategy.

Proof. It is easily checked that, for every graph G and every starting position s, the sequential SCPR game is a *positive* zero sum stochastic game. Hence (by [13, Theorem 4.4.1]) it possesses a value which (by [13, Theorem 4.4.3]) is the smallest componentwise solution to the following system of *optimality equations*:

$$v(\tau)=0; \quad \text{for all } s\in\mathbf{S}_1\cup\mathbf{S}_2: v(s)=\mathbf{Val}\left[q(s)+\sum_{s'\in S}\Pr\left(s'|s,a^1,a^2\right)v\left(s'\right)\right]. \tag{9}$$

However, in each turn of the sequential SCPR game, one of the players has a single available action. For instance, when the state is $s=\left(x^1,x^2,x^3,1\right)$, C_2's action set is $\left\{x^2\right\}$. Hence in (9) we are taking the value of an one-shot game with the game matrix consisting of a single column. It follows that

$$\text{for all } s=\left(x^1,x^2,x^3,1\right): \mathbf{Val}\left[q(s)+\sum_{s'\in\mathbf{S}}\Pr\left(s'|s,a^1,a^2\right)v\left(s'\right)\right]=\max_{a^1}\left[q(s)+\sum_{s'\in\mathbf{S}}\Pr\left(s'|s,a^1,x^2\right)v\left(s'\right)\right]$$

which proves (7); (8) can be proved similarly.

The existence of stationary Markovian optimal strategy for C_2 follows from [13, Corollary 4.4.2]. It is a deterministic strategy because for each state $s \in \mathbf{S}_2$ the corresponding optimal C_2 move is the one minimizing (8). Similarly, the existence of a stationary Markovian ε-optimal strategy for C_1 follows from [13, Problem 4.16]; the strategy is deterministic, because for each state $s \in \mathbf{S}_1$ the corresponding optimal C_1 move is the one maximizing (7). $\square$

For the computation of the solution to (7)–(8) we have the following.

Proposition 3.2. *Given some graph $G = (V, E)$. Define $\mathbf{v}^{(0)}$ by*

$$v^{(0)}(\tau) = 0; \quad \textit{for all } s \in \mathbf{S}_1 \cup \mathbf{S}_2 : v^{(0)}(s) = q(s)$$

and $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ by the following recursion:

$$v^{(i)}(\tau) = 0; \tag{10}$$

$$\textit{for all } s = \left(x^1, x^2, x^3, 1\right) \in \mathbf{S}_1 : \quad v^{(i)}(s) = \max_{a^1} \left[q(s) + \sum_{s' \in \mathbf{S}} \Pr\left(s' | s, a^1, x^2\right) v^{(i-1)}\left(s'\right) \right]; \tag{11}$$

$$\textit{for all } s = \left(x^1, x^2, x^3, 2\right) \in \mathbf{S}_2 : \quad v^{(i)}(s) = \min_{a^2} \left[q(s) + \sum_{s' \in \mathbf{S}} \Pr\left(s' | s, x^1, a^2\right) v^{(i-1)}\left(s'\right) \right]. \tag{12}$$

Then, for every $s \in \mathbf{S}_1 \cup \mathbf{S}_2$, $\lim_{i \to \infty} v^{(i)}(s)$ exists and equals $v(s)$, the value of the sequential SCPR game played on G, starting from s.

Proof. Obviously, for all $s \in \mathbf{S}$, $v(s) \in [0, 1]$. Hence $\mathbf{v}$ is a (componentwise) finite vector. Then from [13, Theorem 4.4.4] we know that, defining $\mathbf{v}^{(0)}$ by

$$v^{(0)}(\tau) = 0; \quad \text{for all } s \in \mathbf{S}_1 \cup \mathbf{S}_2 : v^{(0)}(s) = q(s)$$

and $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ by the recursion

$$v^{(i)}(\tau) = 0, \quad \text{for all } s \in \mathbf{S}_1 \cup \mathbf{S}_2 : \quad v^{(i)}(s) = \mathbf{Val} \left[q(s) + \sum_{s' \in S} \Pr\left(s' | s, a_1, a_2\right) v^{(i-1)}\left(s'\right) \right], \tag{13}$$

we get $\lim_{i \to \infty} \mathbf{v}^{(i)} = \mathbf{v}$ (the value vector of Theorem 3.1). The equivalence of (13) to (10)–(12) is established by the argument used in the proof of Theorem 3.1. $\square$

Remark 3.3. The significance of Theorem 3.1 is the following. Since SCPR is a positive zero sum stochastic game, it will certainly have a value, which satisfies the optimality equations (9); each equation of the system (9) involves the value of a one-shot game. However, the optimality equations can be expressed in the simpler form (7)–(8) which shows that the values of the one-shot games can be computed by simple max and min operations.

Remark 3.4. Similar remarks can be made about Proposition 3.2, where the iteration (10)–(12) is computationally simpler (involves only max and min operations) than (13). Note the similarity of (10)–(12) to the algorithm of [14] for determining the winner of a classic CR game. The similarity becomes stronger in the case of deterministic σ_3. In this case, $\Pr\left(s' | s, a^1, x^2\right)$ equals 1 for a single $s' = \mathbf{T}\left(s, a^1, x^2\right)$ and $\Pr\left(s' | s, x^1, a^2\right)$ equals 1 for a single $s' = \mathbf{T}\left(s, x^1, a^2\right)$; where $\mathbf{T}\left(s, a^1, a^2\right)$ is the *transition function* which yields the next state when, from s, C_1 plays a^1 and C_2 plays a^2; there is also a suppressed dependence on the move of R, which is $\overline{\sigma}_3(s)$. Using this notation, (10)–(12) simplify to

$$\text{for all } s = \left(x^1, x^2, x^3, 1\right) \in \mathbf{S}_1 : v^{(i)}(s) = \max_{a^1} \left[q(s) + v^{(i-1)}\left(\mathbf{T}\left(s, a^1, x^2\right)\right) \right], \tag{14}$$

$$\text{for all } s = \left(x^1, x^2, x^3, 2\right) \in \mathbf{S}_2 : v^{(i)}(s) = \min_{a^2} \left[q(s) + v^{(i-1)}\left(\mathbf{T}\left(s, x^1, a^2\right)\right) \right]; \tag{15}$$

these parallel closely the algorithm of [14, p. 2494].

Remark 3.5. Finally, note that Theorem 3.1 and Proposition 3.2 hold even when $c(G) > 1$; the reason for which we have previously required $c(G) = 1$ has to do with the appropriateness of the payoff function introduced in Section 2. In particular, when $c(G) > 1$ our argument about C_2's incentive to capture R does not hold necessarily (hence, depending on σ_3, C_2 may ensure payoff of 0 without ever capturing R).

4. Results for the concurrent SCPR

In this section we establish that concurrent SCPR has a value which can be computed by value iteration. We first consider the case in which R is controlled by a general probability function σ_3 ("random robber") and then examine in greater detail the case in which σ_3 is oblivious deterministic ("oblivious deterministic robber").

4.1. Random robber

The two main results on concurrent SCPR are immediate consequences of the more general results of [13].

Theorem 4.1. *Given some graph $G = (V, E)$. For every $s = \left(x^1, x^2, x^3\right) \in \widehat{\mathbf{S}}_a$, the concurrent SCPR game starting at s has a value $v(s)$. The vector of values $\mathbf{v} = (v(s))_{s\in\mathbf{S}}$ is the smallest (componentwise) solution of the following* optimality *equations*

$$v(\tau) = 0; \quad \text{for all } s \in \widehat{\mathbf{S}}_a : v(s) = \mathbf{Val}\left[q(s) + \sum_{s' \in S} \Pr\left(s'|s, a_1, a_2\right) v\left(s'\right)\right]. \tag{16}$$

Furthermore, C_2 has a stationary Markovian optimal strategy and, for every $\varepsilon > 0$, C_1 has a stationary Markovian ε-optimal strategy.

Proof. For every graph G (and every starting position s) SCPR is a positive stochastic game. Hence (by [13, Theorem 4.4.1]) it possesses a value which (by [13, Theorem 4.4.3]) satisfies the optimality equation (16). Furthermore C_2 has a stationary Markovian optimal strategy by [13, Corollary 4.4.2] and, for every $\varepsilon > 0$, C_1 has a stationary Markovian ε-optimal strategy by [13, Problem 4.16]. □

Proposition 4.2. *Given some graph $G = (V, E)$, let $s = \left(x^1, x^2, x^3\right) \in \widehat{\mathbf{S}}_a$. Define $\mathbf{v}^{(0)}$ by*

$$v^{(0)}(\tau) = 0; \quad \text{for all } s \in \widehat{\mathbf{S}}_a : v^{(0)}(s) = q(s)$$

and $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ by the following recursion

$$v^{(i)}(\tau) = 0; \quad \text{for all } s \in \widehat{\mathbf{S}}_a : v^{(i)}(s) = \mathbf{Val}\left[q(s) + \sum_{s' \in \mathbf{S}} \Pr\left(s'|s, a_1, a_2\right) v^{(i-1)}\left(s'\right)\right]. \tag{17}$$

Then, for every $s \in \widehat{\mathbf{S}}_a$, $\lim_{i\to\infty} v^{(i)}(s)$ exists and equals $v(s)$, the value of the concurrent SCPR game played on G, starting from s.

Proof. This follows immediately from [13, Theorem 4.4.4]. □

4.2. Oblivious deterministic robber

Theorem 3.1 and Proposition 3.2 are "simpler" than Theorem 4.1 and Proposition 4.2, in the sense that the former do not involve the computation of matrix game values. We will now show that, when σ_3 is oblivious deterministic, we can obtain a similar simplification of Theorem 4.1. Before presenting these results in rigorous form, let us describe them informally.

1. Suppose first that a game is played between a single cop and an oblivious deterministic robber. We will prove that there exists a stationary Markovian deterministic cop strategy $\overline{\sigma}^*$ by which the cop can capture the robber in minimum time.
2. Next consider two cops and an oblivious deterministic robber. We will prove that the extension of $\overline{\sigma}^*$ to SCPR is *optimal* for both cops. More specifically, neither cop loses anything by using it; and one of the two will capture the robber with probability one.

Let us now formalize the above ideas. We pick any graph $G = (V, E)$ and any oblivious deterministic robber strategy $\overline{\sigma}_3$ and keep these fixed for the remainder of the discussion. Further, let $\mathcal{S}$ denote the set of all functions $\overline{\sigma} : V \times V \to V$ with the restriction that for all $\left(x^1, x^3\right) \in V \times V : \overline{\sigma}\left(x^1, x^3\right) \in N\left[x^1\right]$. In other words, $\mathcal{S}$ is the set of legal stationary Markovian deterministic cop strategies for the "classic" CR game of one cop and one robber.

Now pick some $\overline{\sigma} \in \mathcal{S}$ and play the game with starting positions $X_0^1 = x_0^1 \in V$ (for the cop) and $X_0^3 = x_0^3 \in V$ (for the robber). The following sequence (dependent on $\overline{\sigma}$, x_0^1, x_0^3) of cop and robber positions will be produced:

$$X_0^1 = x_0^1,\, X_0^3 = x_0^3,\, X_1^1 = \overline{\sigma}\left(x_0^1, x_0^3\right),\, X_1^3 = \overline{\sigma}_3\left(x_0^3\right), \ldots;$$

let $T_{\overline{\sigma}}\left(x_0^1, x_0^3\right)$ be the *capture time*, in other words, the smallest t such that $X_t^1 = X_t^3$, for the sequence produced by $\overline{\sigma}$, x_0^1, x_0^3 (and $\overline{\sigma}_3$). Also define

$$\overline{V \times V} = \left\{ \left(x^1, x^3\right) : x^1 \in V, x^3 \in V, x^1 \neq x^3 \right\}.$$

Then we have the following.

Lemma 4.3. *Given a graph $G = (V, E)$ and an oblivious deterministic robber strategy $\overline{\sigma}_3$. Let*

$$\textit{for all } x^1 \in V : T^{(0)}\left(x^1, x^1\right) = 0, \quad \textit{for all } \left(x^1, x^3\right) \in \overline{V \times V} : T^{(0)}\left(x^1, x^3\right) = \infty.$$

Now perform the following iteration for $i = 1, 2, \ldots$:

$$\textit{for all } x^1 \in V : T^{(i)}\left(x^1, x^1\right) = 0; \quad \textit{for all } \left(x^1, x^3\right) \in \overline{V \times V} : T^{(i)}\left(x^1, x^3\right) = \min_{x' \in N\left[x^1\right]} \left[1 + T^{(i-1)}\left(x', \sigma_3\left(x^3\right)\right)\right]; \tag{18}$$

$$\textit{for all } x^1 \in V : T^{(i)}\left(x^1, x^1\right) = 0; \quad \textit{for all } \left(x^1, x^3\right) \in \overline{V \times V} : \overline{\sigma}^{(i)}\left(x^1, x^3\right) = \arg\min_{x' \in N\left[x^1\right]} \left[1 + T^{(i-1)}\left(x', \sigma_3\left(x^3\right)\right)\right]. \tag{19}$$

Then the limits

$$\lim_{i \to \infty} \overline{\sigma}^{(i)}\left(x^1, x^3\right), \qquad \lim_{i \to \infty} T^{(i)}\left(x^1, x^3\right)$$

exist for all $(x^1, x^3) \in V \times V$. Furthermore, letting $\overline{\sigma}^*(x^1, x^3) = \lim_{i \to \infty} \overline{\sigma}^{(i)}(x^1, x^3)$ and $T^*(x^1, x^3) = \min_{\overline{\sigma} \in \mathcal{S}} T_{\overline{\sigma}}(x^1, x^3)$, we have

$$\textit{for all } \left(x^1, x^3\right) \in V \times V : \lim_{i \to \infty} T^{(i)}\left(x^1, x^3\right) = T_{\overline{\sigma}^*}\left(x^1, x^3\right) = T^*\left(x^1, x^3\right). \tag{20}$$

Proof. The proof is based on a standard dynamic programming argument. First note that, for every $(x^1, x^3) \in V \times V$, $T^*(x^1, x^3) < |V|$. This is true because C_1 can reach any vertex of V in at most $|V| - 1$ moves; so C_1 can simply go to $X^3_{|V|}$ (the *known* location of R at time $t = |V|$) and wait for the robber there.

Next we prove by induction that

$$T^*\left(x^1, x^3\right) = n \Rightarrow \left(\forall i \geq n : T^*\left(x^1, x^3\right) = T^{(i)}\left(x^1, x^3\right)\right). \tag{21}$$

For $n = 0$, $T^*(x^1, x^3) = 0$ implies $x^1 = x^3$ and, from the algorithm, $T^*(x^1, x^1) = 0 = T^{(i)}(x^1, x^1)$ for all $i \in \mathbb{N}_0$. Now suppose that (21) holds for $n = 1, 2, \ldots, k$ and consider the case $n = k + 1$, in which $T^*(x^1, x^3) = k + 1$ is the smallest number of steps in which C_1 can reach R. This also means that (i) there exists some $x' \in N\left[x^1\right]$ from which C_1 can reach R (who now starts at $\overline{\sigma}_3(x^3)$) in k steps and (ii) there does not exist any $x'' \in N\left[x^1\right]$ from which C_1 can reach R in $m < k$ steps (because then C_1 starting at x^1 could reach R in $m + 1 < k + 1$ steps). In other words

$$T^*\left(x^1, x^3\right) = k + 1 \Rightarrow T^*\left(x^1, x^3\right) = \min_{x' \in N\left[x^1\right]} \left[1 + T^{(k)}\left(x', \overline{\sigma}_3\left(x^3\right)\right)\right] = T^{(k+1)}\left(x^1, x^3\right).$$

It is also easy to check that:

$$T^{(k+1)}\left(x^1, x^3\right) = k + 1 \Rightarrow \left(\forall i > k + 1 : T^{(i)}\left(x^1, x^3\right) = k + 1\right).$$

Hence the induction has been completed.

Given (21), we see immediately that

$$\text{for all } \left(x^1, x^3\right) \in V \times V, i \geq |V| : T^{(i)}\left(x^1, x^3\right) = T^*\left(x^1, x^3\right)$$

which implies that both $\lim_{i \to \infty} T^{(i)}(x^1, x^3) = T^*(x^1, x^3)$ and $\lim_{i \to \infty} \sigma^{(i)}(x^1, x^3)$ exist. Taking the limit (as i tends to ∞) in (18)–(19) we get the *optimality equations*

$$T^*\left(x^1, x^3\right) = \min_{x' \in N\left[x^1\right]} \left[1 + T^*\left(x', \overline{\sigma}_3\left(x^3\right)\right)\right],$$

$$\overline{\sigma}^*\left(x^1, x^3\right) = \arg\min_{x' \in N\left[x^1\right]} \left[1 + T^*\left(x', \overline{\sigma}_3\left(x^3\right)\right)\right].$$

Hence, it is clear from the iteration (18)–(19) that $T_{\overline{\sigma}^*}(x^1, x^3) = T^*(x^1, x^3)$, for all $(x^1, x^3) \in V \times V$. $\square$

Now let us use $\overline{\sigma}^*$ of Lemma 4.3 to define strategies $\overline{\sigma}_i^*$ for C_i $(i \in \{1,2\})$ as follows:

$$\text{for all } \left(x^1,x^2,x^3\right) \in \widehat{\mathbf{S}}_a : \overline{\sigma}_1^*\left(x^1,x^2,x^3\right) = \overline{\sigma}^*\left(x^1,x^3\right);$$
$$\text{for all } \left(x^1,x^2,x^3\right) \in \widehat{\mathbf{S}}_a : \overline{\sigma}_2^*\left(x^1,x^2,x^3\right) = \overline{\sigma}^*\left(x^2,x^3\right).$$

Then the following holds.

Theorem 4.4. *Given some graph $G=(V,E)$, suppose SCPR is played on G and the robber is controlled by an oblivious deterministic strategy $\overline{\sigma}_3$. Then $\overline{\sigma}_i^*$ is an optimal strategy for C_i ($i \in \{1,2\}$), for every starting position $s=\left(x^1,x^2,x^3\right) \in \widehat{\mathbf{S}}_a$. Furthermore*

$$\text{for all } s = \left(x^1,x^2,x^3\right) \in \widehat{\mathbf{S}}_a : \begin{array}{lll} T_{\overline{\sigma}_1^*}\left(x^1,x^3\right) \le T_{\overline{\sigma}_2^*}\left(x^2,x^3\right) & \Rightarrow & v(s)=1, \\ T_{\overline{\sigma}_1^*}\left(x^1,x^3\right) > T_{\overline{\sigma}_2^*}\left(x^2,x^3\right) & \Rightarrow & v(s)=0. \end{array}$$

Proof. The key fact is this: when $\overline{\sigma}_3$ is oblivious deterministic, the players C_1 and C_2 interact only at the last phase of the game, when R is captured. In effect each cop plays a "decoupled" classic CR game, in which $\overline{\sigma}^*$ of Lemma 4.3 guarantees capture in minimum time. Of course in the full SCPR game there is always the possibility that the other cop can capture R at an earlier time. Hence the best C_i can do is to attempt to capture R at the earliest possible time and an optimal strategy to this end is $\overline{\sigma}_i^*$; he has no incentive to deviate from $\overline{\sigma}_i^*$ (by using another deterministic or probabilistic strategy) because this can never reduce his projected capture time. Hence $\overline{\sigma}_i^*$ is optimal for C_i. Since $\overline{\sigma}_1^*$, $\overline{\sigma}_2^*$ and $\overline{\sigma}_3$ are deterministic, the outcome of the game is also deterministic. In particular, when $T_{\overline{\sigma}_1^*}\left(x^1,x^3\right) \le T_{\overline{\sigma}_2^*}\left(x^2,x^3\right)$, with probability 1 C_1 reaches R before or at the same time as C_2; hence $v(s)=1$; when $T_{\overline{\sigma}_1^*}\left(x^1,x^3\right) > T_{\overline{\sigma}_2^*}\left(x^2,x^3\right)$, C_2 reaches R before C_1 with probability 1; hence $v(s)=0$. □

The next theorem gives an additional characterization of the value $v(s)$. In the statement of the theorem we will use the following notation: suppose the game is in the state s, C_1 plays a^1, C_2 plays a^2 and R plays the (predetermined) move $\overline{\sigma}_3(s)$; then we denote the next game state by $\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)$. We have the following.

Theorem 4.5. *Given some graph $G=(V,E)$, suppose SCPR is played on G and the robber is controlled by an oblivious deterministic strategy $\overline{\sigma}_3$. Then, for all $s \in \widehat{\mathbf{S}}_a$, we have*

$$v(s) = \max_{a^1}\min_{a^2}\left[q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)\right)\right] = \min_{a^2}\max_{a^1}\left[q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)\right)\right]. \tag{22}$$

Proof. Since $\overline{\sigma}_3$ is deterministic, $\Pr\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)|s,a^1,a^2\right)=1$. Hence, by [13, Theorem 4.4.3]:

$$v(s) = \mathbf{Val}\left[q(s) + \sum_{s' \in S}\Pr\left(s'|s,a_1,a_2\right)v\left(s'\right)\right] = \mathbf{Val}\left[q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)\right)\right].$$

Since $\overline{\sigma}_1^*$ and $\overline{\sigma}_2^*$ are also deterministic, at every turn of the game they produce an action with probability one. Hence there exist actions $\overline{a}^1 = \overline{\sigma}_1^*(s)$, $\overline{a}^2 = \overline{\sigma}_2^*(s)$ such that

$$v(s) = q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\overline{a}^2,\overline{\sigma}_3(s)\right)\right)\right).$$

From Theorem 4.4, $v(s) \in \{0,1\}$, hence we consider two cases.

1. Suppose $v(s)=1$. This means, that starting at s, C_1 will certainly capture R.
   (a) If s is a C_1-capture state, then $q(s)=1$ and, for any actions $\overline{a}^1,\overline{a}^2$, $\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\overline{a}^2,\overline{\sigma}_3(s)\right)\right)=\tau$, in which case

   $$v\left(\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\overline{a}^2,\overline{\sigma}_3(s)\right)\right)\right) = v(\tau) = 0.$$

   Hence $v(s) = \max_{a^1}\min_{a^2}\left[q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)\right)\right]=1$.

   (b) If s is not a C_1-capture state, then $q(s)=0$ and $v\left(\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\overline{a}^2,\overline{\sigma}_3(s)\right)\right)\right)=1$. Suppose there existed some $\widehat{a}^2$ such that $v\left(\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\widehat{a}^2,\overline{\sigma}_3(s)\right)\right)\right)=0$. This would mean that, starting at $\widehat{\mathbf{T}}\left(s,\left(\overline{a}^1,\widehat{a}^2,\overline{\sigma}_3(s)\right)\right)$, C_2 would certainly capture R before C_1 and, since $\overline{a}^1$ is the optimal (fastest capturing) move for C_1, we would also have

   $$\text{for all } a^1 \in \mathbf{A}_1(s) : q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,\widehat{a}^2,\overline{\sigma}_3(s)\right)\right)\right) = 0.$$

   But then $v(s) = \mathbf{Val}\left[q(s) + v\left(\widehat{\mathbf{T}}\left(s,\left(a^1,a^2,\overline{\sigma}_3(s)\right)\right)\right)\right]=0$, contrary to the assumption. So we must instead have

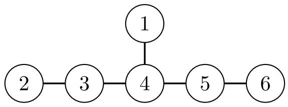

Fig. 1. An example where deterministic robber strategy results in probabilistic optimal cop strategies.

Table 1
A part of the robber strategy σ_3.

$\left(x_t^1, x_t^2, x_t^3\right)$	$x_{t+1}^3 = \overline{\sigma}_3\left(x_t^1, x_t^2, x_t^3\right)$
$(2, 6, 1)$	4
$(2, 6, 4)$	3
$(2, 5, 4)$	5
$(3, 6, 4)$	5
$(3, 5, 4)$	3

Table 2
Possible states at the end of the first turn.

$s_0 = (2, 6, 1)$	$a_1^1 = 2$	$a_1^2 = 6$	$a_1^3 = \overline{\sigma}_3(2, 6, 1) = 4$	$s_1 = (2, 6, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 2$	$a_1^2 = 5$	$a_1^3 = \overline{\sigma}_3(2, 6, 1) = 4$	$s_1 = (2, 5, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 3$	$a_1^2 = 6$	$a_1^3 = \overline{\sigma}_3(2, 6, 1) = 4$	$s_1 = (3, 6, 4)$
$s_0 = (2, 6, 1)$	$a_1^1 = 3$	$a_1^2 = 5$	$a_1^3 = \overline{\sigma}_3(2, 6, 1) = 4$	$s_1 = (3, 5, 4)$

$$\text{for all } a^2 \in \mathbf{A}_2(s) : q(s) + v\left(\widehat{\mathbf{T}}\left(s, \left(\overline{a}^1, a^2, \overline{\sigma}_3(s)\right)\right)\right) = 1$$

which implies $v(s) = \max_{a^1} \min_{a^2} \left[q(s) + v\left(\widehat{\mathbf{T}}\left(s, \left(a^1, a^2, \overline{\sigma}_3(s)\right)\right)\right)\right] = 1.$

2. Now suppose $v(s) = 0$. Then s is not a C_1-capture state, hence $q(s) = 0$. Now, we will show that

$$\text{for all } a^1 \in \mathbf{A}_1(s): \text{ there exists } a^2 \in \mathbf{A}_2(s) : v\left(\widehat{\mathbf{T}}\left(s, \left(a^1, a^2, \overline{\sigma}_3(s)\right)\right)\right) = 0. \tag{23}$$

If this is not the case, then

$$\text{there exists } \widetilde{a}^1 \in \mathbf{A}_1(s) : \text{for all } a^2 \in \mathbf{A}_2(s) : v\left(\widehat{\mathbf{T}}\left(s, \left(\widetilde{a}^1, a^2, \overline{\sigma}_3(s)\right)\right)\right) = 1.$$

Then C_1 will certainly capture R (before C_2) starting from the game position $\mathbf{T}\left(s, \left(\widetilde{a}^1, a^2, \overline{\sigma}_3(s)\right)\right)$ and this will be true for any $a^2 \in \mathbf{A}_2(s)$. But this means that C_1, starting from game position s and playing $\widetilde{a}^1$, will certainly capture R before C_2; which in turn means $v(s) = 1$, contrary to the hypothesis. Hence (23) holds and this implies that

$$\text{for all } a^1 \in \mathbf{A}_1(s) : \min_{a^2} v\left(\widehat{\mathbf{T}}\left(s, \left(a^1, a^2, \overline{\sigma}_3(s)\right)\right)\right) = 0$$
$$\Rightarrow \max_{a^1} \min_{a^2} v\left(\widehat{\mathbf{T}}\left(s, \left(a^1, a^2, \overline{\sigma}_3(s)\right)\right)\right) = 0.$$

Hence we have proved the first part of (22). The proof of the second part is similar and omitted. □

Remark 4.6. It must be emphasized that Theorems 4.4 and 4.5 do *not* hold for deterministic *non*-oblivious strategies $\overline{\sigma}_3$. This can be seen by the following counterexample. Suppose that concurrent SCPR is played on the graph of Fig. 1, starting from the state $(2, 6, 1)$.

Furthermore, the robber is controlled by the $\overline{\sigma}_3$ which is (partially) described in Table 1.

For every game state not listed above the robber stays in place, in other words, $x_{t+1}^3 = \overline{\sigma}_3\left(x_t^1, x_t^2, x_t^3\right) = x_t^3$. Now consider what the first moves of C_1 and C_2 should be. They know that R will move into vertex 4; C_1 can either stay at 2 or move into 3; C_2 can either stay at 6 or move into 5. After the first move is completed, the possible game states are in Table 2.

It is easy to check (from the respective $\overline{\sigma}_3$ values) that for $s_1 = (2, 6, 4)$ and $s_1 = (3, 5, 4)$ the capturing cop is C_1, while for $s_1 = (2, 5, 4)$ and $s_1 = (3, 6, 4)$ the capturing cop is C_2. Hence the game can be written out as the following (one-shot) matrix game (Table 3).

It is easy to compute, using standard methods, the optimal strategies for this game. C_1 must use $\Pr\left(a^1 = 2\right) = \Pr\left(a^1 = 3\right) = \frac{1}{2}$ and C_2 must use $\Pr\left(a^2 = 6\right) = \Pr\left(a^2 = 5\right) = \frac{1}{2}$. This implies that the optimal strategies $\overline{\sigma}_1^*$ and $\overline{\sigma}_2^*$ are probabilistic, despite the fact that $\overline{\sigma}_3$ is deterministic (but not oblivious). We can also see that in this case the optimal cop strategy is *not* to always move towards the robber. Many similar examples can be constructed. The important point is this: when $\overline{\sigma}_3$ is not oblivious, C_1 (resp. C_2) moves can influence future R moves and (since moves are performed simultaneously) this influence cannot be predicted by C_2 (resp. C_1).

Table 3
The one-shot matrix game equivalent to the original stochastic game.

	$a^2 = 6$	$a^2 = 5$
$a^1 = 2$	1	0
$a^1 = 3$	0	1

5. Conclusion

We have introduced the game of selfish cops and passive robber (SCPR game) and established its basic properties, namely the existence of value and optimal strategies for both the sequential and concurrent variants. We have also provided algorithms for the computation of the aforementioned quantities. In the current paper we have examined *qualitative* variants of the game, in which the goal of each cop is to maximize his probability of capturing the robber. In a forthcoming paper we will examine *quantitative* variants, in which the goal is to capture the robber *in the shortest possible time*.

Several additional issues merit further study and will be the subject of our future research. We have formulated SCPR as a zero-sum game; but reasonable formulations as a *non-zero-sum* game are also possible and we conjecture that these may lead to qualitatively different results. In addition, if we remove the assumption that the robber is passive and deal instead with the situation of two selfish cops and a robber *actively trying to avoid capture*, then we are left with a *three*-player game, which we intend to study in the future.

Settling the above mentioned questions (and additional ones which may arise in the process) will hopefully result in a comprehensive game theoretic framework encompassing the numerous CR variants which have appeared in the literature (and so far have been studied mainly from a combinatorial point of view).

References

[1] M.H. Albert, R. Nowakowski, D. Wolfe, Lessons in Play: An Introduction to Combinatorial Game Theory, 2007.
[2] L. de Alfaro, T.A. Henzinger, O. Kupferman, Concurrent reachability games, Theoret. Comput. Sci. 386 (2007) 188–217.
[3] V.J. Baston, F.A. Bostock, Infinite deterministic graphical games, SIAM J. Control Optim. 31 (1993) 1623–1629.
[4] A. Berarducci, B. Intrigila, On the cop number of a graph, Adv. in Appl. Math. 14 (1993) 389–403.
[5] E. Berlekamp, J.H. Conway, R. Guy, Winning Ways for Your Mathematical Plays, 1982.
[6] D. Berwanger, Graph games with perfect information, preprint, 2012.
[7] A. Bonato, G. MacGillivray, A general framework for discrete-time pursuit games, preprint, 2015.
[8] A. Bonato, G. MacGillivray, Characterizations and algorithms for generalized Cops and Robbers games, Contrib. Discrete Math. (2017), submitted for publication.
[9] K. Chatterjee, T.A. Henzinger, A survey of stochastic ω-regular games, J. Comput. System Sci. 78 (2012) 394–413.
[10] K. Chatterjee, M. Jurdziński, On Nash Equilibria in Stochastic Games, International Workshop on Computer Science Logic, Springer, Berlin–Heidelberg, 2004.
[11] A. Ehrenfeucht, J. Mycielski, Positional strategies for mean payoff games, Internat. J. Game Theory 8 (1979) 109–113.
[12] H. Everett, Recursive games, in: Contributions to the Theory of Games, vol. 3, 1957, pp. 47–78.
[13] J. Filar, K. Vrieze, Competitive Markov Decision Processes, Springer Science & Business Media, 1997.
[14] G. Hahn, G. MacGillivray, A note on k-cop, l-robber games on graphs, Discrete Math. 306 (2006) 2492–2497.
[15] R. Isaacs, Differential Games, John Wiley and Sons, 1965.
[16] Ath. Kehagias, P. Prałat, Some remarks on cops and drunk robbers, Theoret. Comput. Sci. 463 (2012) 133–147.
[17] Ath. Kehagias, D. Mitsche, P. Prałat, Cops and invisible robbers: the cost of drunkenness, Theoret. Comput. Sci. 481 (2013) 100–120.
[18] Ath. Kehagias, D. Mitsche, P. Prałat, The role of visibility in pursuit/evasion games, Robotics 3 (2014) 371–399.
[19] Ath. Kehagias, G. Konstantinidis, Simultaneously moving cops and robbers, Theoret. Comput. Sci. 645 (2016) 48–59.
[20] N. Komarov, P. Winkler, Capturing the drunk robber on a graph, arXiv preprint, arXiv:1305.4559, 2013.
[21] R. Mazala, Infinite games, in: Automata Logics, and Infinite Games, 2002, pp. 23–38.
[22] R. McNaughton, Infinite games played on finite graphs, Ann. Pure Appl. Logic 65 (1993) 149–184.
[23] J.F. Mertens, Stochastic games, in: Handbook of Game Theory with Economic Applications, vol. 3, 2002, pp. 1809–1832.
[24] R. Nowakowski, P. Winkler, Vertex to vertex pursuit in a graph, Discrete Math. 43 (1983) 230–239.
[25] R. Nowakowski, A. Bonato, The Game of Cops and Robbers on Graphs, AMS, 2011.
[26] M.J. Osborne, A. Rubinstein, A Course in Game Theory, MIT Press, 1994.
[27] A. Quilliot, Jeux et pointes fixes sur les graphes, Ph.D. Dissertation, Universite de Paris VI, 1978.
[28] L.S. Shapley, Stochastic games, Proc. Natl. Acad. Sci. USA 39 (1953) 1095–1100.
[29] A. Washburn, Deterministic graphical games, J. Math. Anal. Appl. 153 (1990) 84–96.