

# Ping Pong in Dangerous Graphs: Optimal Black Hole Search with Pure Tokens

Paola Flocchini<sup>1</sup>, David Ilcinkas<sup>2</sup>, and Nicola Santoro<sup>3</sup>

<sup>1</sup> SITE, University of Ottawa, Canada  
flocchin@site.uottawa.ca

<sup>2</sup> CNRS, Université de Bordeaux, France  
david.ilcinkas@labri.fr

<sup>3</sup> School of Computer Science, Carleton University, Canada  
santoro@scs.carleton.ca

**Abstract.** We prove that, for the black hole search problem, the pure token model is computationally as powerful as the whiteboard model; furthermore the complexity is exactly the same. More precisely, we prove that a team of *two* asynchronous agents, each endowed with a single identical pebble (that can be placed only on nodes, and with no more than one pebble per node) can locate the black hole in an arbitrary network of known topology; this can be done with  $\Theta(n \log n)$  moves, where  $n$  is the number of nodes, even when the links are not FIFO.

**Key words:** distributed computing, graph exploration, mobile agents, autonomous robots, dangerous graphs.

## 1 Introduction

### 1.1 The Framework

*Black Hole Search* (BHS) is the distributed problem in a networked system (modeled as a simple edge-labelled graph  $G$ ) of determining the location of a *black hole* (BH): a site where any incoming agent is destroyed without leaving any detectable trace. The problem has to be solved by a team of identical system agents injected into  $G$  from a safe site (the homebase). The team operates in presence of an adversary that chooses e.g., the edge labels, the location of the black hole, the delays, etc. The problem is solved if at least one agent survives and all surviving agents know the location of the black hole (e.g., see [15]).

The practical interest of BHS derives from the fact that a black hole can model several types of faults, both hardware and software, and security threats arising in networked systems supporting code mobility. For example, the crash failure of a site in an asynchronous network turns such a site into a black hole; similarly, the presence at a site of a malicious process (e.g., a virus) that thrashes any incoming message (e.g., by classifying it as spam) also renders that site a black hole. Clearly, in presence of such a harmful host, the first step must be to determine and report its location.

From a theoretical point of view, the natural interest in the computational and complexity aspects of this distributed problem is amplified by the fact that it opens a new dimension in the classical *graph exploration* problem. In fact, the black hole can be located only after all the nodes of the network but one have been visited and are found to be safe; in this exploration process some agents may disappear in the black hole. In other words, while the existing wide body of literature on *graph exploration* (e.g., see [1, 2, 8, 9, 16, 17]) assumes that the graph is *safe*, BHS opens the research problems of the exploration of *dangerous graphs*.

Indeed BHS has been studied in several settings, under a variety of assumptions on the power of the adversary and on the capabilities of the agents; e.g., on the level of synchronization of the agents; on whether or not the links are FIFO; on the type of mechanisms available for inter agent communication and coordination; on whether or not the agents have a map of the graph. In these investigations, the research concern has been to determine under what conditions and at what cost mobile agents can successfully accomplish this task. The main complexity measures are the size of the team (i.e., the number of agents employed) and the number of moves performed by the agents; sometimes also time complexity is considered.

In this paper we are interested in the weakest settings that still make the problem solvable. Thus we will make no assumptions on timing or delays, and focus on the *asynchronous* setting. Indeed, while the research has also focused on the *synchronous* case [5–7, 18, 19] where all agents are synchronized and delays are unitary, the main body of the investigations has concentrated on the asynchronous one (e.g., [4, 10–13]).

## 1.2 The Quest and its Difficulties

In the asynchronous setting, the majority of the investigations operate in the *whiteboard* model: every node provides a shared space for the arriving agents to read and write (in fair mutual exclusion). The whiteboard model is very powerful: it endows the agents not only with direct and explicit communication capabilities, but also with the means to overcome severe network limitations; in particular, it allows the software designer to assume FIFO links (even when not supported by the system). Additionally, whiteboards allow to break symmetry among identical agents. Indeed, whiteboards (and even stronger inter-agent coordination mechanisms) are supported by most existing mobile agent platforms [3]. The theoretical quest, on the contrary, has been for the weakest interaction mechanism allowing the problem to be solved.

A weaker and less demanding interaction mechanism is the one assumed by the *token* model, used in the early investigations on (safe) graph exploration; it is provided by identical *pebbles* (that can be placed on nodes, picked up and carried by the agents) without any other form of marking or communication (e.g., [2]).

The research quest is to determine if pebbles are computationally as powerful as whiteboards with regards to BHS. The importance of this quest goes beyond

the specific problem, as it would shed some light on the relative computational power of these two interaction mechanisms.

Two results have been established so far in this quest. In [10] it has been shown that  $\Delta + 1$  agents<sup>4</sup> without a map (the minimum team size under these conditions), each endowed with an identical pebble, can locate the black hole with a (very high but) polynomial number of moves. In [13] it has been shown that two agents with a map (the minimum team size under these conditions), each endowed with a constant number of pebbles, can locate the black hole in a ring network with  $\Theta(n \log n)$  moves, where  $n$  denotes the number of nodes in the network.

Although they indicate that BHS can be solved using pebbles instead of whiteboards, these results do not prove yet the computational equivalence for BHS of these two inter-agent coordination mechanisms. There are two main reasons for this. The first main reason is that both results assume *FIFO links*; note that the whiteboard model allows to work assuming FIFO links, but does not require them. Hence, the class of networks for which the results of [10, 13] apply is smaller than that covered with whiteboards; also such an assumption is a powerful computational help to any solution protocol. The second and equally important reason is that these results are *not* established within the “pure” token model used in the traditional exploration problem. In fact, in [10, 13] the agents are allowed to place pebbles not only on nodes but also on links (e.g., to indicate on which link it is departing); this gives immediately to a single token the computational power of  $O(\log \Delta)$  bits of information. In [13], where the network considered is only a ring, each agent has available several tokens, and multiple tokens can be placed at the exact same place (node or link) to store more than one bit of information.

### 1.3 Our Results

In this paper, we provide the first proof that indeed the pure token model is computationally as powerful as the whiteboard model for BHS.

The context we examine is the one of agents with a map in an arbitrary graph. For this context we prove that: A team of *two* asynchronous agents, each endowed with a single identical pebble (that can be placed only on nodes, and at no more than one pebble per node) and a map of the graph can locate the black hole with  $\Theta(n \log n)$  moves, even if the links are not FIFO.

In other words, for networks of known topology, using pure tokens it is possible to obtain exactly the same optimal bounds for team size and number of moves as using whiteboards.

Note that our result implies as a corollary an optimal solution for the whiteboard model using only a single bit of shared memory per node; the existing solution [11] requires a whiteboard of  $O(\log n)$  bits at each node.

Our results are obtained using a new and (surprisingly) simple technique called *ping pong*. In its bare form, this technique solves the problem but with

---

<sup>4</sup>  $\Delta$  denotes the maximum node degree in  $G$

$O(n^2)$  moves. To obtain the optimal bound, the technique is enhanced by integrating it with additional mechanisms, exploiting two ideas developed in previous investigations: “split work” [12], and “distance counting” [13]. The mechanisms that we have developed use a variety of novel not-trivial techniques, and are the first to overcome the severe limitation imposed by the lack of the FIFO assumption (available instead in all previous investigations with whiteboards or tokens).

The paper is organized as follows. We first present our techniques, prove their properties and analyze their complexity in the case of ring networks (Section 3). Then, in Section 4, we show how to modify and enhance those techniques so to obtain the same bounds also in the case of arbitrary graphs.

Due to space limitations, some proofs and the code of the algorithms are omitted; the interested reader can find them in [14].

## 2 Terminology and Definitions

Let  $G = (V, E)$  be a simple biconnected<sup>5</sup> graph with  $n = |V|$  nodes. At each node  $x$ , there is a distinct label from a totally ordered set associated to each of its incident links. We shall denote by  $(G, \lambda)$  the resulting edge-labelled graph.

Operating in  $(G, \lambda)$  is a team of identical autonomous mobile agents (or robots). All agents enter the system from the same node, called homebase. The agents have computing capabilities, computational storage (polynomially bounded by the size of the graph), and a map of  $(G, \lambda)$  with the indication of the homebase; they can move from node to neighbouring node, and obey the same set of behavioral rules (the *algorithm*). Every agent has a *pebble*; all pebbles are identical. A pebble can be carried, put down at a node if no other pebble is already there, and picked up from a node by an agent without pebbles.

When an agent enters a node, it can see if there is a pebble dropped there; it might be however unable to see other agents there or to determine whether they are carrying a pebble with them.

The system is *asynchronous* in the sense that (i) each agent can enter the system at an arbitrary time; (ii) traveling to a node other than the black hole takes a finite but otherwise unpredictable amount of time; and (iii) an agent might be idle at a node for a finite but unpredictable amount of time. The basic computational step of an agent (executed either when the agent arrives to a node, or upon wake-up) is to look for the presence of a pebble, drop or pick up the pebble if wanted, and leave the node through some chosen port (or terminate). The whole computational step is performed in local mutual exclusion as an atomic action, i.e. as if it took no time to execute it. Links are *not* FIFO: two agents moving on the same link in the same direction at the same time might arrive at destination in an arbitrary order.

To simplify the model, we can assume without loss of generality that the transition between two states of the agent at a node plus the corresponding

<sup>5</sup> Note that biconnectivity is necessary for BHS to be solvable [11].

move are instantaneous. In other words, the waiting due to asynchrony only occurs after the move of the agent. Furthermore we can assume that also the actions of agents at different nodes occur at different instants.

A *black hole* is a node that destroys any incoming agent; no observable trace of such a destruction will be evident to the other agents. The location of the black hole is unknown to the agents. The *Black Hole Search* problem is to find the location of the black hole. More precisely, the problem is solved if at least one agent survives, and all surviving agents know the location of the black hole.

The two measures of complexity of a solution protocol are the number of agents used to locate the black hole and the total number of moves performed by the agents.

### 3 Black hole search in rings

#### 3.1 Preliminaries

Without loss of generality, we can assume that the clockwise direction is the same for both agents: for example, the direction implied by the link with the smallest label at the homebase. In the following, going right (resp. left) means going in the clockwise (resp. counterclockwise) direction. An agent exploring to the right (resp. left) is said to be a *right* (resp. *left*) agent. Using this definition, an agent *changes role* if it was a left agent and becomes a right agent or vice versa. For  $i \geq 0$ , the node at distance  $i$  to the right, resp. to the left, of the home base will be called node  $i$ , resp. node  $-i$ . Hence node  $i$  and  $i - n$  represent the same node, for  $0 \leq i \leq n$ .

In the algorithm the agents obey the two following metarules:

1. An agent always ensures that a pebble is lying at  $u$  before traversing an unknown edge  $\{u, v\}$  from  $u$  to  $v$  (i.e. an edge that it does not know to be safe).
2. An agent never traverses an unknown edge  $\{u, v\}$  from  $u$  to  $v$  if a pebble lies at  $u$  and the pebble was not dropped there by this agent.

These metarules imply that the two agents never enter the black hole from the same edge. Moreover, each agent keeps track of its progress by storing the number of the most-right, resp. most-left, node in a variable **Last\_Right**, resp. **Last\_Left**, used to detect termination: when only one node remains unexplored, this node is the black hole and the agent can stop.

A (right) agent is said to traverse an edge  $\{u, v\}$  from  $u$  to  $v$  using *cautious walk* if it has one pebble, it drops it at  $u$ , traverses the edge (in state **Explore-Right**), comes back to  $u$  (in state **Pick-Up-Right**), retrieves the pebble and goes again to  $v$  (in state **Ping-Right**). A (left) agent is said to traverse an edge  $\{u, v\}$  from  $u$  to  $v$  using *double cautious walk* if it has one pebble and the other is at  $u$ , it goes to  $v$  (in state **Explore-Left**) carrying one pebble, the other pebble staying at  $u$ , drops the pebble at node  $v$ , comes back to  $u$  (in state **Pick-Up-Left**), retrieves the other pebble and goes again to  $v$  (in state **Ping-Left**). We will see later that double cautious walk is employed only by left agents. Note that these two cautious explorations obey the first metarule.

### 3.2 The Algorithm

Our algorithm is based on a novel coordination and interaction technique for agents using simple tokens, *Ping-Pong*. The idea at the basis of this technique is the following: one agent explores the “right” side and one the “left” side (the side assigned to an agent changes dynamically, due to the non-FIFO nature of the links). However, only one agent at a time is allowed to explore; the agent willing to do so must first “steal” the pebble of the other, and then can proceed to explore its allowed side. When an agent discovers that its pebble has been stolen, it goes to find it and steal the other pebble as well. This generate a “ping-pong” movements of the agents on the ring. The actual Ping-Pong technique based on this idea must however take into account the non-FIFO nature of the links, which creates a large variety of additional situations and scenarios (e.g., an agent moving to steal the pebble of the other, might “jump over” the other agent).

Algorithm **EnhancedPingPong** is divided in two phases, each one further divided into stages. The first phase is the Ping-Pong technique. The second phase, whose function is to ensure that the costs are kept low, in some cases may not be executed at all. Inside a phase, a *stage* is a maximal period during which no agent changes role.

In the first phase, exploration to the right is always done using cautious walk, while exploration to the left is always done using double cautious walk (i.e., after stealing a pebble). Note that, since an agent exploring to the right uses one pebble and an agent exploring to the left uses two pebbles, the agents cannot make progress simultaneously in two different directions because there are only two pebbles in total. This also implies that while an agent is exploring new nodes it knows all the nodes that have already been explored, as well as the position of the only unexplored node where the other agent possibly died. This prevents the agents from exploring the same node and thus from dying in the black hole from two different directions.

**Phase 1.** Initially both agents explore to the right. Since links are not FIFO, an agent may pass the other and take the lead without any of the two noticing it. Nevertheless, it eventually happens that one agent  $L$  finds the pebble of the other agent  $R$ , say at node  $p$  (at the latest it happens when one agent locates or dies in the black hole). When this happens  $L$  drops its pebble at node  $p - 1$  (if its pebble is not already there) and steals  $R$ 's pebble. Having control on the two pebbles,  $L$  starts to explore left using double cautious walk. The stage has now an even number. When/if  $R$  comes back to  $p$  to retrieve its pebble, it does not find it. It then goes left in state **Pong-Right** until it finds a pebble. Agent  $R$  does eventually find a pebble because at the beginning of the stage there is a pebble at its left (at node  $p - 1$ ), and Agent  $L$  never removes a pebble before putting the other pebble further to the left. At this point  $R$  retrieves the pebble and goes right again in state **Ping-Right** and explores to the right. When/if  $L$  realizes that one of its pebble has been stolen, it changes role (and the stage changes) and explores to the right using its remaining pebble. At this point, both agents explore to the right. Again, one agent will find and steal the pebble of the other. To ensure progress in exploration, a right agent puts down its pebble

only when it reaches the last visited node to the right it knows (using its variable `Last_Right`). Consequently the stealing at the end of an odd stage always occurs at least one node further to the right from two stages before. Hence the algorithm of Phase 1 is in fact correct by itself but the number of moves can be  $\Theta(n^2)$  in the worst case (one explored node every  $O(n)$  moves). To decrease the worst case number of moves to  $O(n \log n)$ , the agents switch to Phase 2 as soon as at least two nodes have been explored to the right.

**Phase 2.** Phase 2 uses the *halving* technique, based on an idea of [12], but highly complicated by the absence of whiteboards and by the lack of FIFO. The idea is to regularly divide the workload (the unexplored part) in two. One agent has the left half to explore (using variable `Goal_Left`), while the second agent explores the right half (using variable `Goal_Right`). These explorations are performed concurrently by using (simple) cautious walk (for a right agent, in states `Halving-Explore-Right`, `Halving-Pick-Up-Right` and `Halving-Ping-Right`). After finite time, exactly one agent finishes its part and joins the other in exploring the other part, changing role and thus changing the stage number. At some point, one agent  $A$  will see the other agent's pebble.  $A$  steals the pebble and moves it by one position to indicate a change of stage to the second agent  $B$ . It then computes the new workload, divide it into two parts (using the function `Update_Goal_Left` or `Update_Goal_Right`), and goes and explores its newly assigned part, changing role again by switching to state `Halving-From-Left-To-Right` or `Halving-From-Right-To-Left`. This can happen several times (if  $B$  remains blocked by the asynchronous adversary or if it is dead in the black hole). When/if agent  $B$  comes back to retrieve its pebble, it does not find it. It further goes back to retrieve its pebble in state `Halving-Pong-Right` (if it is a right agent). The number of moves it has to perform to find the pebble indicates how many halvings (pair of stages) it misses. Knowing that, it can compute what is the current unexplored part and what is its current workload. It then starts to explore its part. Since there are at most  $O(\log n)$  stages of  $O(n)$  moves each, this leads to a total number of moves of  $O(n \log n)$ .

The algorithm starts with a few stages of Phase 1 because Phase 2 needs some safe nodes to put the pebble that is used as a message to indicate the current partition of the workload.

Several other technical details and precautions have to be taken because of asynchrony and lack of FIFO. Due to space restrictions, the code describing all the details of the state transitions can be found in [14].

### 3.3 Correctness and complexity

As explained before the algorithm consists of up to two *phases*. The first one corresponds to the case where both agents are in one of the eight states `Ping-Right`, `Ping-Left`, `Explore-Right`, `Explore-Left`, `Pick-Up-Right`, `Pick-Up-Left`, `Pong-Right`, `Put-Pebble-Right`. If this is not the case, we say that the algorithm is in its second phase. (Note that this phase may not exist in all possible

executions.) An agent is said to be a *right*, resp. *left*, agent if its state ends with `-Right`, resp. `-Left`. Using this definition, an agent *changes role* if it was a left agent and becomes a right agent or vice versa. Finally, inside a phase, a *stage* is a maximal period during which no agent changes role.

For the purpose of the proofs of the main theorems, we will use the three following properties.

**Property  $\mathcal{P}(p)$** , with  $p \in \{0, 1\}$ : There is a left agent  $L$  and a right agent  $R$ . The agent  $L$  is waiting at node  $p-1$ , where one pebble is located. Agent  $L$  is carrying the other pebble and is in state `Ping-Left`. Moreover, its variable `Last_Right` has value  $p$ . Agent  $R$ , empty-handed, is in one of the following situations:

- it is dead in the black hole located at node  $p+1$ ;
- it is at node  $p+1$  in state `Explore-Right` and its variable `Last_Right` has value  $p$ ;
- it is already back from node  $p+1$  at node  $p$  in state `Pick-Up-Right` and its variable `Last_Right` has value  $p+1$ .

Moreover, the termination condition of agent  $L$  is not satisfied, and in the last two cases, the value `Last_Left` is the same for each agent.

**Property  $\mathcal{P}'_L(p, q)$** , with  $p \geq 2$ ,  $q \leq 0$  and  $p - q < n - 2$ : There is a left agent  $L$  and a right agent  $R$ . There exists some  $k \geq 0$ , with  $p - k - 1 > q$ , such that  $L$  is waiting at node  $p - k - 1$  where one pebble is located. Agent  $L$  is carrying the other pebble and is in state `Halving-From-Right-To-Left`. Moreover, its variable `Last_Right`, resp. `Last_Left`, has value  $p$ , resp.  $q$ . Its variable `Goal_Left` has value `Update.Goal_Left(p, q, 1)`. (Its variable `Goal_Right` has value `Goal_Left + n - 1`.) Agent  $R$ , empty-handed, is in one of the following situations:

- it is dead in the black hole located at node  $p+1$ ;
- it is waiting at  $p+1$  in state `Explore-Right` and its variable `Last_Right` has value  $p$ ;
- it is already back from node  $p+1$  at node  $p$  in state `Pick-Up-Right` and its variable `Last_Right` has value  $p+1$ ;
- it is waiting at node  $p+1$  in state `Halving-Explore-Right` and its variable `Last_Right` has value  $p$ ;
- it is already back from node  $p+1$  at node  $p$  in state `Halving-Pick-Up-Right` and its variable `Last_Right` has value  $p+1$ ;
- it is waiting at node  $p-i$ ,  $1 \leq i \leq k$ , in state `Halving-Pong-Right`, its variable `Last_Right` has value  $p+1$  and its variable `Counter` has value  $i-1$ .

Moreover, in the second and third cases, the value `Goal_Left` of Agent  $L$  is equal to `Update.Goal_Left(p, q', k+1)`, where  $q'$  is the value `Last_Left` of Agent  $R$ . In the last three cases, the value `Goal_Left` of Agent  $L$  is equal to `Update.Goal_Left(p, q', k+1-Counter)`, where  $q'$  equals `Goal_Left` of Agent  $R$ .

**Property  $\mathcal{P}'_R(p, q)$** , with  $p \geq 2$ ,  $q \leq 0$  and  $p - q < n - 2$ : There is a left agent  $L$  and a right agent  $R$ . There exists some  $k \geq 0$ , with  $q + k + 1 < p$ , such that  $R$  is waiting at node  $q + k + 1$  where one pebble is located. Agent  $R$  is carrying the other pebble and is in state `Halving-From-Left-To-Right`. Moreover, its variable `Last_Left`, resp. `Last_Right`, has value  $q$ , resp.  $p$ . Its variable `Goal_Right` has value `Update.Goal_Right(p, q, 1)`. (Its variable `Goal_Left` has

value  $\text{Goal\_Right} - n + 1$ .) Agent  $L$ , empty-handed, is in one of the following situations:

- it is dead in the black hole located at node  $q - 1$ ;
  - it is waiting at  $q - 1$  in state **Halving-Explore-Left** and its variable **Last\_Left** has value  $q$ ;
  - it is already back from node  $q - 1$  at node  $q$  in state **Halving-Pick-Up-Left** and its variable **Last\_Left** has value  $q - 1$ ;
  - it is waiting at node  $q + i$ , for some  $1 \leq i \leq k$  in state **Halving-Pong-Left**, its variable **Last\_Left** has value  $q - 1$  and its variable **Counter** has value  $i - 1$ .
- Moreover, in the last three cases, the value **Goal\_Right** of Agent  $R$  is equal to **Update\_Goal\_Right** ( $p', q, k + 1 - \text{Counter}$ ), where  $p'$  is **Goal\_Right** of Agent  $L$ .

**Lemma 1.** *Consider a  $n$ -node ring containing a homebase and a black hole, and two agents running Algorithm **EnhancedPingPong** from the homebase. After finite time, one of the following situations occurs:*

- Stage 2 of Phase 1 begins and Property  $\mathcal{P}(p)$  holds for some  $p \in \{0, 1\}$ ;
- Phase 2 begins and Property  $\mathcal{P}'_L(p, 0)$  holds for some integer  $p$  such that  $2 \leq p \leq n - 2$ ;
- all agents of the non-empty set of surviving agents have terminated and located the black hole.

*Moreover, at that time, each edge has been traversed at most a constant number of times since the beginning of the algorithm.*

**Lemma 2.** *Consider a  $n$ -node ring containing a homebase and a black hole, and two agents running Algorithm **EnhancedPingPong** from the homebase. Assume that at some time  $t$  a Phase-1 stage of even number  $i$  begins and that Property  $\mathcal{P}(p)$  holds for some  $p \in \{0, 1\}$ . Then at some time  $t' > t$  one of the following situations occurs:*

- Stage  $i + 2$  of Phase 1 begins and Property  $\mathcal{P}(p')$  holds for some integer  $p'$  such that  $p < p' \leq 1$  (thus  $p' = 1$ );
- Phase 2 begins and Property  $\mathcal{P}'_L(p', q)$  holds for some integers  $p'$  and  $q$  such that  $p' \geq 2$ ,  $q \leq 0$  and  $p' - q < n - 2$ ;
- all agents of the non-empty set of surviving agents have terminated and located the black hole.

*Moreover, each edge has been traversed at most a constant number of times between times  $t$  and  $t'$ .*

**Lemma 3.** *Consider a  $n$ -node ring containing a homebase and a black hole, and two agents running Algorithm **EnhancedPingPong** from the homebase. Assume that at some time  $t$  a Phase-2 stage of odd number  $i$  begins and that either Property  $\mathcal{P}'_L(p, q)$  or Property  $\mathcal{P}'_R(p, q)$  holds for some integers  $p$  and  $q$  such that  $p \geq 2$ ,  $q \leq 0$  and  $p - q < n - 2$ . Then at some time  $t' > t$  one of the following situations occurs:*

- Stage  $i + 2$  of Phase 2 begins and either Property  $\mathcal{P}'_L(p', q')$  or Property  $\mathcal{P}'_R(p', q')$  holds for some integers  $p'$  and  $q'$  such that  $p' \geq p$ ,  $q' \leq q$  and  $n - (p' - q' + 1) \leq \lceil \frac{n(p - q + 1)}{2} \rceil$ ;

- all agents of the non-empty set of surviving agents have terminated and located the black hole.

Moreover, each edge has been traversed at most a constant number of times between times  $t$  and  $t'$ .

**Theorem 1.** *Algorithm `EnhancedPingPong` is correct.*

More precisely, consider a  $n$ -node ring containing a homebase and a black hole, and two agents running Algorithm `EnhancedPingPong` from the home base. After finite time, there remains at least one surviving agent and all surviving agents have terminated and located the black hole.

*Proof.* From Lemmas 1 and 2, we know that the first phase contains at most five stages, each one ending after finite time. Furthermore we know that after finite time, either the algorithm terminates correctly, or Property  $\mathcal{P}'_L(p, q)$  or  $\mathcal{P}'_R(p, q)$  holds, for some integers  $p$  and  $q$  such that  $q \leq 0 < p$  and  $0 < p - q < n - 2$ . From Lemma 3, we know that a stage of Phase 2 ends after finite time. We also know that if the algorithm does not terminate after two stages  $i, i + 1$  in Phase 2, then Property  $\mathcal{P}'_L(p', q')$  or  $\mathcal{P}'_R(p', q')$  holds, for some integers  $p'$  and  $q'$  such that the positive value  $p' - q'$  is strictly less than  $p - q$ . Hence, after finite time, neither  $\mathcal{P}_L(p, q)$  nor  $\mathcal{P}'_R(p, q)$  can be satisfied and the algorithm terminates correctly.

**Theorem 2.** *The total number of moves performed by two agents running Algorithm*

*EnhancedPingPong in a  $n$ -node ring is at most  $O(n \log n)$ .*

*Proof.* From Lemmas 1 and 2, there are at most five stages in Phase 1 and for each of them the number of edge traversals performed by each agent is at most  $O(n)$ . From Lemma 3, there are at most  $O(\log n)$  stages in Phase 2 because the unexplored part is basically halved every two stages. From the same lemma, we have that for each Phase-2 stage the number of edge traversals performed by each agent is at most  $O(n)$ . Hence, overall, the total number of moves performed by two agents running Algorithm `EnhancedPingPong` in a  $n$ -node ring is at most  $O(n \log n)$ .

The optimality of the algorithm follows from the fact that, in a ring, the problem cannot be solved with less agents or (asymptotically) less moves [12], and clearly not with less pebbles.

## 4 Black hole search in arbitrary graphs

### 4.1 Preliminaries

In this section, both agents are provided with a map of the network containing all edge labels and a mark showing the position of the homebase in this network. Thus, each node of the map can be uniquely identified (for example by a list of edge labels leading to it from the homebase). Therefore, each agent is able to

know where it lies at any point of the execution of the algorithm. It also knows where each edge incident to its position leads.

The algorithm `GeneralizedEnhancedPingPong` we propose for arbitrary networks is an adaptation of the algorithm `EnhancedPingPong` that we described for rings. To be able to apply `EnhancedPingPong` in a general graph, each agent will maintain a partial mapping between the node numbers used in the algorithm and the actual nodes in the network (or its map), such that at any point in time an agent knows what means “go left” and “go right”.

During the execution of the algorithm, each agent maintains two walks  $W_R$  and  $W_L$ , defined as two sequences  $(r_0, r_1, \dots, r_P)$  and  $(l_0, l_1, \dots, l_Q)$  of nodes of the network. The nodes  $r_0$  and  $l_0$  correspond to the homebase. Since  $W_R$  and  $W_L$  are walks, we have that  $\{r_i, r_{i+1}\}$  and  $\{l_j, l_{j+1}\}$  are edges of the graph, for all  $0 \leq i < P$  and  $0 \leq j < Q$ .

From these two walks, we define recursively function  $\sigma$  as follows. First  $\sigma(0) = 0$ . Assume that  $\sigma$  is defined for all  $j$  such that  $0 \leq j \leq i$ , for some  $i \geq 0$ . Then if there exists an element  $r_K$  in  $W_R$  such that  $r_K \notin \{r_{\sigma(0)}, r_{\sigma(1)}, \dots, r_{\sigma(i)}\}$  but for all  $k < K$ ,  $r_k \in \{r_{\sigma(0)}, r_{\sigma(1)}, \dots, r_{\sigma(i)}\}$ , then  $\sigma(i+1) = K$ , otherwise  $\sigma(i+1)$  is not defined. Similarly, assume that  $\sigma$  is defined for all  $j$  such that  $i \leq j \leq 0$ , for some  $i \leq 0$ . Then if there exists an element  $l_K$  in  $W_L$  such that  $l_K \notin \{l_{\sigma(0)}, l_{\sigma(1)}, \dots, l_{\sigma(i)}\}$  but for all  $k < K$ ,  $l_k \in \{l_{\sigma(0)}, l_{\sigma(1)}, \dots, l_{\sigma(i)}\}$ , then  $\sigma(i-1) = K$ , otherwise  $\sigma(i-1)$  is not defined.

Let us assume that an agent lies at some node  $i$ . If  $i \geq 0$  (i.e., the agent is at the homebase or somewhere in the explored part to the right) going one step right from node  $i$  means following the sub-walk  $(r_{\sigma(i)}, r_{\sigma(i)+1}, \dots, r_{\sigma(i+1)})$  of  $W_R$ . Going left from node  $i+1$  to node  $i$  means following this sub-walk in reverse order. Similarly, if  $i \leq 0$  (i.e., the agent is at the homebase or in the explored part to the left) going one step left from node  $i$  means following the sub-walk  $(l_{\sigma(i)}, l_{\sigma(i)+1}, \dots, l_{\sigma(i-1)})$  of  $W_L$ . Going right from node  $i-1$  to node  $i$  means following this sub-walk in reverse order.

## 4.2 The Algorithm

We now describe the definitions of the walks  $W_R$  and  $W_L$  throughout the algorithm. First of all, node  $v_0$  denotes the homebase. Node  $v_1$  is the neighbor of node  $v_0$  reachable by the smallest edge label while node  $v_{-1}$  is the neighbor of node  $v_0$  reachable by the largest edge label.

At the beginning of the algorithm, let  $T_R$  be a tree spanning all nodes except for node  $v_{-1}$  and containing the edge  $\{v_0, v_1\}$ . Let  $W_R$  be a DFS traversal of  $T_R$  starting from node  $v_0$  by the edge  $\{v_0, v_1\}$ . Let  $W_L$  be  $(v_0, v_{-1})$ . Clearly, nodes  $v_{-1}, v_0$  and  $v_1$  are the nodes -1, 0 and 1.

Assume that the stage changes from an odd number to an even number in Phase 1. Let  $p$  and  $q$  be the values `Last_Right` and `Last_Left` of the left agent. Then the new walk  $W_R$  consists of the first  $\sigma(p+1) + 1$  elements of the old  $W_R$ , that is the sequence  $(r_0, r_1, \dots, r_{\sigma(p+1)})$ . In addition, let  $T_L$  be a tree spanning all nodes except for node  $p+1$ . Let  $S_L$  be a DFS traversal of  $T_L$  starting from node  $q-1$ . Then the new walk  $W_L$  is the concatenation of the old  $W_L$  and of

the sequence  $S_L$ . The left agent does these updates of the walks when changing role. The other agent does these updates when it finds out that its pebble has been stolen. More precisely, it updates its walks just before switching to state **Pong-Right**. Note that both agents agree on the new definition of the walks because they use the same values for  $p$  and for  $q$  (cf. Property  $\mathcal{P}(p)$ ).

Similarly assume that the stage changes from an even number to an odd number in Phase 1. Let  $p$  and  $q$  be the values **Last\_Right**  $- 1$  and **Last\_Left** of the right agent. Then the new walk  $W_L$  consists of the first  $\sigma(q-1)+1$  elements of the old  $W_L$ , that is the sequence  $(l_0, l_1, \dots, l_{\sigma(q-1)})$ . In addition, let  $T_R$  be a tree spanning all nodes except node  $q-1$ . Let  $S_R$  be a DFS traversal of  $T_R$  starting at node  $p+1$ . Then the new walk  $W_R$  is the concatenation of the old  $W_R$  and of the sequence  $S_R$ . The right agent does these updates of the walks when it retrieves a pebble, just before switching from state **Pong-Right** to state **Ping-Right**. The other agent does these updates when it finds out that its pebble has been stolen. More precisely, it updates its walks just before changing role. Note that again both agents agree on the new definition of the walks because they use the same values for  $p$  and for  $q$ .

The walks are also updated at the beginning of each Phase-2 stage of odd number  $i$ . More precisely this is done by an agent each time and just after it updates its knowledge of the unexplored part and its goals. Assume w.l.o.g. that the stage is now  $i$  because a right agent became a left agent. Let  $p$ ,  $q$  and  $g$  be the values, respectively, of **Last\_Right**, **Last\_Left** and **Goal\_Right** just after the update of the goals. Let  $\{V_{ex}, V_{uex}\}$  be a partition of the nodes of the graph such that  $V_{ex}$  is the set of nodes  $\{q, q+1, \dots, p-1, p\}$ . From Lemma 5.2 in [11],  $V_{uex}$  can be partitioned into  $V_R$  and  $V_L$  such that  $|V_R| = p-g$ , the node  $p+1$  is in  $V_R$ , and the graphs  $G_R$  and  $G_L$  induced by, respectively,  $V_{ex} \cup V_R$  and  $V_{ex} \cup V_L$  are connected. Let  $T_R$  and  $T_L$  be spanning trees of  $G_R$  and  $G_L$ . Let  $S_R$  be a DFS traversal of  $T_R$  starting at node  $p+1$ . Similarly let  $S_L$  be a DFS traversal of  $T_L$  starting at node  $q$ . Finally, let  $W'_R$ , resp.  $W'_L$ , consists of the first  $\sigma(p+1)$ , resp.  $\sigma(q)$ , elements of  $W_R$ , resp.  $W_L$ . Then the new walks  $W_R$  and  $W_L$  are respectively the concatenation of  $W'_R$  and  $S_R$  and the concatenation of  $W'_L$  and  $S_L$ . Note that both agents agree on the new definition of the walks because they use the same values for  $p$ ,  $q$  and  $g$  (cf. Properties  $\mathcal{P}'_R(p, q)$  and  $\mathcal{P}'_L(p, q)$ ).

In some cases, it is possible to use (safe) shortcuts to decrease the number of moves. Indeed, always following the walks  $W_R$  and  $W_L$  to go right and left may lead to a total of  $n \log^2 n$  moves. The algorithm is modified as follows. During Phase 2, each agent maintains an additional variable **Last\_Seen\_Pebble** that basically memorizes the last place where the agent has seen the other pebble. When an agent finishes a half and switches to state **Halving-Ping-Left**, resp. **Halving-Ping-Right**, it goes directly to node **Last\_Seen\_Pebble** and if there are no pebbles at this node, it then goes directly to node **Last\_Left**, resp. **Last\_Right**. This is done by traversing only nodes that are known to be safe.

### 4.3 Correctness and complexity

**Theorem 3.** *Algorithm `GeneralizedEnhancedPingPong` is correct.*

*More precisely, consider a  $n$ -node graph containing a homebase and a black hole, and two agents running Algorithm `GeneralizedEnhancedPingPong` from the homebase. After finite time, at least one agent survives and all surviving agents have terminated and located the black hole.*

*Proof.* As noticed in the description of the algorithm in the previous subsection, the two agents agree on the definition of the walks and thus of the node numbers. Moreover, one can easily check that the function  $\sigma$  defining the node numbers always gives the same number to the same node as soon as this node has been explored by at least one agent. Indeed, if a node  $i \geq 0$  is explored, then the initial part  $(r_0, r_1, \dots, r_{\sigma(i)})$  of  $W_R$  is kept unchanged forever. A similar property holds for  $i \leq 0$ . Finally note that a node of the graph has at most one pre-image by  $\sigma$ .

To summarize, Algorithm `GeneralizedEnhancedPingPong` behaves exactly the same as Algorithm `EnhancedPingPong`. The only difference is that traversing an edge in the ring may correspond to the traversals of (finitely) many edges in an arbitrary graph. Nevertheless, since Algorithm `EnhancedPingPong` is correct, Algorithm `GeneralizedEnhancedPingPong` is correct as well.

**Theorem 4.** *The total number of moves performed by two agents running Algorithm*

*`GeneralizedEnhancedPingPong` in a  $n$ -node graph is at most  $O(n \log n)$ .*

*Proof.* In this proof we call the number of edge traversals performed by an agent going from node  $i$  to node  $i + 1$  ( $-n < i < n - 1$ ) the length of the virtual edge  $\{i, i + 1\}$ . We now bound the total number of moves performed by each agent in each phase.

As in the case of the ring, the first phase consists of at most five stages. Moreover, each update of the walks increases their length by at most  $2n$  because the path appended to a walk is a DFS traversal of a tree. Hence, the sum of all the lengths of the virtual edges traversed in the first phase is at most  $10n$ . From Lemmas 1 and 2, each edge of the network has been traversed at most a constant number of times during Phase 1. Hence, the total number of moves performed by two agents running Algorithm `GeneralizedEnhancedPingPong` is at most  $O(n \log n)$  in the first phase.

From the lemmas 1, 2 and 3, either Property  $\mathcal{P}'_R(p, q)$  or Property  $\mathcal{P}'_L(p, q)$  holds, for some integers  $p$  and  $q$ , at the beginning of a Phase-2 stage of odd number. Let  $p_i$  and  $q_i$  be the two integers corresponding to the stage  $2i + 1$  of the second phase, for  $0 \leq i \leq s$ . Note that  $s$  is at most  $O(\log n)$ . Let  $p_{s+1}$ , resp.  $q_{s+1}$ , be the right, resp. left, neighbor of the black hole. By definition of the properties and from the lemmas, we have that  $q_{s+1} \leq \dots \leq q_0 \leq 0 \leq p_0 \leq \dots \leq p_{s+1}$ . Since the walks  $W_R$  and  $W_L$  are updated when and only when the goals are updated, and since a walk is always extended by a DFS traversal of a tree, we obtain that the sum of all the lengths of the virtual edges from node  $p_i$  to  $p_{i+1}$ , and from node  $q_i$  to  $q_{i+1}$ , is at most  $O(n)$ , for  $0 \leq i \leq s$ . Moreover, from the

previous paragraph, the sum of all the lengths of the virtual edges from node  $q_0$  to  $p_0$  is at most  $O(n)$ .

Consider a stage  $2i + 1$  of the second phase, for  $0 \leq i \leq s$ . Let  $A$  be the agent that started the stage by changing role and let  $B$  the other agent. Without loss of generality, assume that  $A$  is a left agent. The total number of moves performed by  $A$  in this stage is at most  $O(n)$  because  $A$  first goes directly (by a shortest safe path) to the beginning  $q_i$  of its workload, thus in at most  $n$  moves, and then stays in the final part of  $W_L$  that corresponds to the DFS traversal of a tree to explore its assigned workload, which incurs at most  $O(n)$  moves (from Lemma 3). If it succeeds to explore its half, then it goes directly to the node where it left the other pebble (thanks to the variable `Last.Seen.Pebble`). If the pebble is not there anymore, it further goes directly to node  $p_i$  and starts to explore the right half. Hence, in any case, for the same reasons as before,  $A$  performs at most  $O(n)$  moves in stage  $2i + 2$ . Concerning  $B$ , if it retrieves its pebble in stage  $2i + 1$  or  $2i + 2$ , it will perform at most  $O(n)$  moves in these two stages, without counting the moves done in state `Halving-Pong-Right`. Indeed, again, exploring a half or going directly to the beginning of it costs at most a linear number of moves.

It remains to bound the number of moves done while in one of the states `Halving-Pong-Right` or `Halving-Pong-Left`. This is done globally over the whole second phase. Each edge traversed in one of these two states may cost up to  $O(n)$  moves. However, there are at most  $O(\log n)$  such traversals because any of them corresponds to an update of the workloads, which happens only a logarithmic number of times in the entire algorithm.

One can now conclude that the total number of moves performed by two agents running Algorithm `GeneralizedEnhancedPingPong` in a  $n$ -node graph is at most  $O(n \log n)$ .

The optimality of the algorithm follows from the fact that, in an arbitrary graph, BHS cannot be solved with less agents or (asymptotically) less moves [11], and clearly not with less pebbles.

**Acknowledgment** This work was done during the stay of David Ilcinkas at the University of Ottawa, as a postdoctoral fellow. Paola Flocchini was partially supported by the University Research Chair of the University of Ottawa. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Discovery grants.

## References

1. S. Albers and M. Henzinger. Exploring unknown environments. *29th ACM Symposium on Theory of Computing (STOC)*, 416–425, 1997.
2. M. A. Bender, A. Fernández, D. Ron, A. Sahai, and S. P. Vadhan. The power of a pebble: Exploring and mapping directed graphs. *Information and Computation*, 176(1):1–21, 2002.

3. J. Cao and S. Das (Eds), *Mobile Agents in Networking and Distributed Computing*, John Wiley, 2008.
4. J. Chalopin, S. Das, N. Santoro. Rendezvous of mobile agents in unknown graphs with faulty links. *21st Conf. on Distributed Comp.* (DISC), 108-122, 2007.
5. C. Cooper, R. Klasing, and T. Radzik. Searching for black-hole faults in a network using multiple agents. *10th Int. Conf. on Principles of Distributed Systems* (OPODIS), 320-332, 2006.
6. J. Czyzowicz, D. Kowalski, E. Markou, and A. Pelc. Complexity of searching for a black hole. *Fundamenta Informaticae*, 71 (2-3): 229-242, 2006
7. J. Czyzowicz, D. Kowalski, E. Markou, and A. Pelc. Searching for a black hole in synchronous tree networks. *Combinatorics, Probability & Computing* 16: 595-619, 2007.
8. S. Das, P. Flocchini, S. Kutten, A. Nayak, and N. Santoro. Map construction of unknown graphs by multiple agents. *Theoretical Computer Science* 385(1-3): 34-48, 2007.
9. X. Deng and C. H. Papadimitriou. Exploring an unknown graph. *J. Graph Theory* 32 (3): 265-297, 1999.
10. S. Dobrev, P. Flocchini, R. Kralovic, and N. Santoro. Exploring a dangerous unknown graph using tokens. *5th IFIP Int. Conf. on Theoretical Computer Science* (TCS), 131-150, 2006.
11. S. Dobrev, P. Flocchini, G. Prencipe, and N. Santoro. Searching for a black hole in arbitrary networks: optimal mobile agents protocol. *Distributed Computing* 19 (1): 1-19, 2006.
12. S. Dobrev, P. Flocchini, G. Prencipe, and N. Santoro. Mobile search for a black hole in an anonymous ring. *Algorithmica* 48: 67-90, 2007.
13. S. Dobrev, R. Kralovic, N. Santoro, and W. Shi. Black hole search in asynchronous rings using tokens. *6th Conf. on Algorithms and Complexity* (CIAC), 139-150, 2006.
14. P. Flocchini, D. Ilcinkas, and N. Santoro. Optimal black hole search with pure tokens. <http://www.scs.carleton.ca/~santoro/Reports/PingPong.pdf>.
15. P. Flocchini and N. Santoro. *Distributed Security Algorithms For Mobile Agents*. Chapter 5 of [3], 2008.
16. P. Fraigniaud, L. Gasieniec, D. Kowalski, and A. Pelc. Collective tree exploration. *Networks*, 48 (3): 166-177, 2006.
17. P. Fraigniaud, D. Ilcinkas, G. Peer, A. Pelc, and D. Peleg. Graph exploration by a finite automaton. *Theoretical Computer Science*, 345 (2-3): 331-344, 2005.
18. R. Klasing, E. Markou, T. Radzik, and F. Sarracco. Approximation bounds for black hole search problems. *9th Int. Conf. on Principles of Distributed Systems* (OPODIS) 261-274, 2005.
19. R. Klasing, E. Markou, T. Radzik, and F. Sarracco. Hardness and approximation results for black hole search in arbitrary networks. *Theoretical Computer Science* 384 (2-3): 201-221, 2007.