Population Protocols with Faulty Interactions: the Impact of a Leader

Giuseppe A. Di Luna[a,*], Paola Flocchini[a], Taisuke Izumi[b], Tomoko Izumi[c], Nicola Santoro[d], Giovanni Viglietta[a]

[a]*Department of Electrical and Computer Engineering, University of Ottawa, 800 King Edward Avenue, Ottawa.*
[b]*Department of Computer Science and Engineering, Nagoya Institute of Technology, Gokisocho, Showa Ward, Nagoya 466-8555.*
[c]*College of Information Science and Engineering, Ritsumeikan University, 56-1 Tojiin Kitamachi, Kita-ku, Kyoto 603-8577.*
[d]*School of Computer Science, Carleton University, 125 Colonel By Drive, Ottawa.*

Abstract

We consider the problem of simulating traditional population protocols under weaker models of communication, which include one-way interactions (as opposed to two-way interactions) and omission faults (i.e., failure by an agent to read its partner's state during an interaction), which in turn may be detectable or undetectable. We focus on the impact of a leader, and we give a complete characterization of the models in which the presence of a unique leader in the system allows the construction of simulators: when simulations are possible, we give explicit protocols; when they are not, we give proofs of impossibility. Specifically, if each agent has only a finite amount of memory, the simulation is possible only if there are no omission faults. If agents have an unbounded amount of memory, the simulation is possible as long as interactions are detectable. If an upper bound on the number of omissions involving the leader is known, the simulation is possible in all omissive models.

---

[*]Corresponding author

Email addresses: `gdiluna@uottawa.ca` (Giuseppe A. Di Luna), `paola.flocchini@uottawa.ca` (Paola Flocchini), `t-izumi@nitech.ac.jp` (Taisuke Izumi), `izumi-t@fc.ritsumei.ac.jp` (Tomoko Izumi), `santoro@scs.carleton.ca` (Nicola Santoro), `gvigliett@uottawa.ca` (Giovanni Viglietta)

A preliminary version of the results contained in this paper have been published in [25].

1. Introduction

1.1. Framework

Consider a system of *interacting computational entities*, called agents, whose interaction is however not under their control but decided by an external scheduler. Such are for example systems of wireless mobile entities where two entities can interact (i.e., exchange information) when they are within the communication range of each other, but their movements, and thus their interactions, are unpredictable. Systems satisfying this condition, sometimes called *opportunistic mobility* or *passive mobility*, have been extensively examined under a variety of assumptions, especially within the context of distributed computing in highly dynamic networks and time-varying graphs (for recent surveys see [14, 30]).

In particular, in the *population protocol* model (PP), introduced in the seminal paper [3], the entities are assumed to be finite-state and anonymous (i.e., identical), execute the same protocol, and interactions are always between pairs of agents. The roles of the two agents involved in an interaction are asymmetric: one agent is considered the *starter* and the other is the *reactor*. Still, the communication is *two-way*: each agent receives the state of the other and executes the protocol to update its own state based on the received information and its own state. Furthermore, in the selection of the occurrences of the interactions, the scheduler is constrained to satisfy some fairness assumption.

The restricted computational universe defined by the basic assumptions of PP has been subsequently expanded in an attempt to overcome the inherent computability limitations and to examine the computational impact of factors such as non-constant memory (e.g., [1, 2, 17]), presence of a leader (e.g., [6]), storage of information on edges (e.g., [16, 18, 21, 29]), etc.

In all these models, including the original one, the interaction is assumed to be fault-free. An immediate important question is what happens if interactions are subject to *failures*. Very little is known in this regard, and most works have focused on faults pertaining agents, i.e., Byzantine failures or crash-stop failures [7, 22, 23]. To the best of our knowledge, the investigation of failures

on interactions is new. An insight comes from the study of the so-called *one-way interaction models* [4], where the starter of an interaction is not able to see the state of the reactor (*immediate transmission*), or it is not even able to detect that the interaction has taken place (*immediate observation*). This study showed that, under one-way interactions, the computational power of the agents is strictly weaker than it is with the usual bidirectional interactions. In particular, if the interactions are not detectable by the starter (i.e., immediate observation), the agents can compute only the threshold predicates, i.e., if a specific symbol in the input is above a certain threshold: $\#(a) > t$ [4]. In the case of two-way interactions, the class of computable predicates coincides with the semilinear predicates, i.e., those corresponding to first-order formulas in Presburger arithmetic. An example of a predicate that is semilinear but is not a threshold predicate is $\#(a) \equiv 3 \pmod{5}$.

In this paper we focus on *omission failures*: during an interaction, the state of an agent is not transmitted. This implies that one, or both agents, may fail to see the state of the other or to detect that an interaction occurred.

The one-way interaction models capture a class of *permanent omission failures* involving the starter's side, those where the starter never sees the state of the reactor agent. Clearly, there are many more types of omission failures, such as those occurring at the reactor's side and, more insidious, those whose occurrence is dynamic and unpredictable. And of course, for each of these types there are different variations, depending on the kind of fault detection assumed.

The complete range of *dynamic omission failures* has been classified in [24], where the following general question was posed: *Under what additional system capabilities is it possible to correctly execute every traditional two-way population protocol in spite of dynamic omission failures?* More specifically, *under what conditions (if any) is it possible to* simulate *the execution of every two-way population protocol for a given class of omission failures?* A population protocol $\mathcal{P}$, working in the presence of some kind of omission failures, is said to simulate a population protocol $\mathcal{P}'$ if each execution of $\mathcal{P}$ can be "reinterpreted" as an execution of $\mathcal{P}'$, in the sense that the sequence of state changes in the agents

3

executing $\mathcal{P}$ is the same as the one that could be observed if the agents executed $\mathcal{P}'$ instead. In other words, a simulator provides an interface between the simulated protocol $\mathcal{P}$ and the physical communication layer, giving the system the illusion of being in a fault-free two-way environment.

The existence of simulators is important in scenarios in which we do not only concern ourselves with the final output of a population protocol, but also with the *execution* that leads to the result. We may want, for instance, to guarantee that our simulating agents enter some critical states exactly as many times as they would if they were actually executing the protocol that is being simulated. A critical state could be a state that determines an outcome on a physical asset; for example, entering a certain state could trigger a money transaction between two bank accounts.

The existence of fault-tolerant one-way simulators of two-way protocols has been investigated in [24] in terms of the amount of *memory* required by the agents to perform such simulations, and a variety of models and results were established. It is shown that, with no *a-priori* knowledge, the simulation of two-way protocols in the presence of omissions is impossible even if the agents have infinite memory. In the weakest models investigated, this impossibility holds even if the number of omission failures in each execution is limited to one.

On the other hand, it is also shown that simulation is possible if agents have unique IDs or the total number of agents is known. Moreover, in some restricted models, simulation is possible when an upper bound on the number of omission faults is known. Notice that the focus of the investigation in [24] and of our paper is on simulation. That is, given a particular two-way protocol $\mathcal{P}$ that solves some problem, we do not investigate in a general sense the possibility of solving the same problem in a different model: we only focus on the possibility of simulating $\mathcal{P}$ in such a model, incidentally solving the same problem as $\mathcal{P}$.

In this paper we continue this general line of research and investigate how the presence in the system of a distinguished agent, a *leader*, can impact the capability of the system to tolerate dynamic omission failures. More precisely, we study the possibility and impossibility of simulation of two-way protocols

4

with the aid of a leader, with respect to the different classes of omission failures and one-way interactions.

1.2. Main Contributions

As in [24], we consider all the computationally distinct models that arise from the introduction of omission faults and/or one-way interactions in two-way protocols: TW, IT, IO, T_i $(i = 1, 2, 3)$, and I_j $(j = 1, 2, 3, 4)$; see Figure 3, where the type of transition function δ, detailed in Section 2, uniquely identifies each model. In particular, TW refers to two-way protocols without omissions; IT and IO refer to the one-way models with *immediate transmission* and *immediate observation*, introduced in [4]; the T_i's and I_i's refer to the distinct two-way and one-way models with omissions, respectively.

We consider two types of omission adversaries: informally, a "malignant" one (the unfair omissive, **UO**), which is able to arbitrarily insert omission faults into "globally fair" sequences of interactions, and a "benign" one (the eventually non omissive, $\diamond$**NO**), which inserts some omission faults, but eventually stops. To make our results stronger, we always assume the benign adversary in the impossibility proofs and the malignant one in the possibility proofs.

We study the negative impact that omissions have on computability, and we show that the simulation of two-way protocols is impossible even with the aid of a leader (Theorem 1), assuming that the amount of memory available to the simulator is bounded.

On the other hand, we show that the presence of both a leader and infinite memory on each agent makes the simulation possible in the weak intermediate one-way models I_1 and I_2 (Theorem 4), and thus in all the upper models of Figure 3, Section 2.5. The fact that this possibility does not apply to IO and T_1 is not accidental: indeed we prove that, for these two models, the simulation is impossible even with both a leader and infinite memory, even against the benign omission adversary (Theorem 2).

Finally, we study what happens when a bound on the omission failures involving the leader is known, and essentially we show that simulators exists for

models I_1 and I_2 (Theorem 5) and model T_1 (Theorem 6), and these imply the possibility of simulations in all other omissive models.

For non-omissive models, we show that two-way simulation is possible in the IT model (Theorem 7). In light of the fact that with constant memory, in absence of additional capabilities, IT protocols are strictly less powerful than TW (see [4]), our results show that this computational gap can be overcome by using a leader.

Our main results are summarized in Figure 1, where white blobs represent possibilities, and gray blobs impossibilities. As a consequence of these results, we have a complete characterization of the feasibility of simulations with respect to the presence of a leader. In Figure 2, there is a summary of the results contained in [24] for the unbounded-memory case and the case of knowledge on omissions: as we can see, the assumption of a leader drastically changes the possibility of designing a simulator.

1.3. Related Work

Since their introduction, there have been extensive investigations on Population Protocols (e.g., see [5, 8, 11, 13, 15, 19, 20, 30, 27]), and the basic assumptions of the original model have been expanded in several directions, typically to overcome inherent computability restrictions. For example, allowing each agent to have non-constant memory [1, 2, 17]; assuming the presence of a leader [6]; allowing a certain amount of information to be stored on the edges [16, 18, 29] of the interaction graph.

The issue of dependable computations in population protocols, first raised in [22], has been considered and studied only with respect to processors' faults, and the basic model has necessarily been expanded. In [23] it has been shown how to compute functions tolerating $\mathcal{O}(1)$ crash-stops and transient failures, assuming that the number of failures is bounded and known. In [7] the specific majority problem under $\mathcal{O}(\sqrt{n})$ Byzantine failures, assuming a fair probabilistic scheduler, has been studied. In [28] unique IDs are assumed, and it is shown how to compute functions tolerating a bounded number of Byzantine faults, under

the assumption that Byzantine agents cannot forge IDs. Interestingly, in [28] it is shown that, in the population protocol model without IDs, the presence of a single Byzantine agent prevents the computation of any non-trivial function. This setting is different from ours since a single Byzantine agent can send any number of bogus states to the other agents.

Self-stabilizing solutions have been devised for specific problems such as leader election (assuming knowledge of the system's size and a non-constant number of states [12], or assuming a leader detection oracle [26]) and counting (assuming the presence of a leader [9]). Moreover, in [10] a self-stabilizing transformer for general protocols has been studied in a slightly different model and under the assumption of unbounded memory and a leader.

Finally, to the best of our knowledge, the one-way model without omissions, has been studied only in [4], where it is shown that IT and IO, when equipped with constant memory, can compute a set of functions that is strictly included in that of TW. Combined with our results in Figure 1, this implies that, without using extra resources (e.g., infinite memory, leader, etc.), simulations are impossible in all the one-way and omissive models.
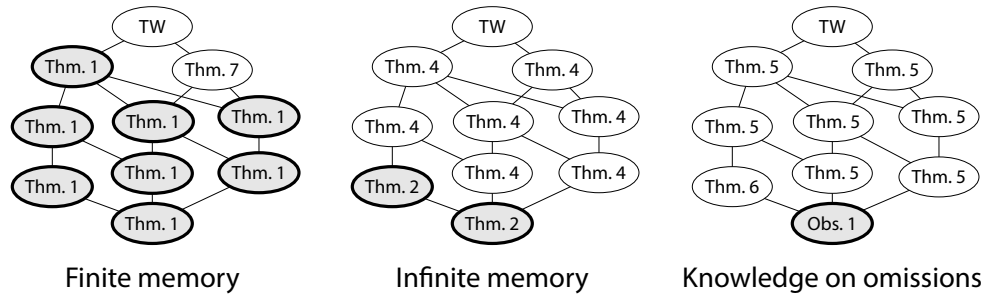

Figure 1: Map of results (cf. Figure 3 page 12). White blobs denote the existence of simulators; grey blobs indicate that simulations are not possible.
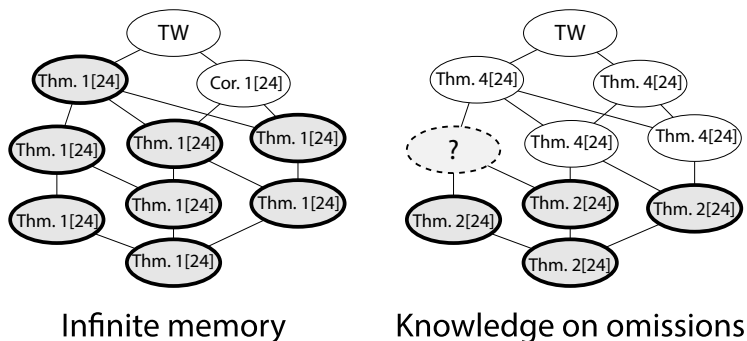
7

Figure 2: Map of results when a leader is not present, see [24] (cf. Figure 1). White blobs denote the existence of simulators; grey blobs indicate that simulations are not possible.

2. Model and Terminology

In this section we define the computation model, the notion of omission, and the notion of simulator.

2.1. Interacting Entities

We consider a system consisting of a set $A = \{a_1, \ldots, a_n\}$ of interacting computational entities, called *agents*. Each *interaction* involves only two agents with asymmetric roles: one agent is the *starter*, a_s, and the other is the *reactor* a_r, that is, an interaction between a_s and a_r is indicated by the ordered pair $i = (a_s, a_r)$.

Interactions occur at discrete times, and at every "time unit" exactly one interaction occurs.

When two agents interact, they exchange information and perform a local computation according to the same protocol $\mathcal{P}$. A protocol is a triple $\mathcal{P} = (Q_{\mathcal{P}}, Q'_{\mathcal{P}}, \delta_{\mathcal{P}})$, where $Q_{\mathcal{P}}$ is a set of local states, $Q'_{\mathcal{P}} \subseteq Q_{\mathcal{P}}$ is a set of initial states, potentially infinite, and $\delta_{\mathcal{P}} \colon Q_{\mathcal{P}} \times Q_{\mathcal{P}} \to Q_{\mathcal{P}} \times Q_{\mathcal{P}}$ is the transition function defining the states of the two interacting agents at the end of their local computation. With a small abuse of notation, and when no ambiguity arises, we will use the same literal (e.g., a_i) to indicate both an agent and its local state.

8

The protocol $\mathcal{P}$ and the number of agents n are enough to uniquely determine the static structure of the system, therefore we will refer to the system as the tuple $(\mathcal{P}, n)$. A *configuration* C of a system $(\mathcal{P}, n)$ is an n-tuple of local states, that is $C \in Q_{\mathcal{P}}^n$. Initially, the systems starts in an *initial configuration* $C_0 \in Q_{\mathcal{P}}'^n$. Given a tuple $t = (x_1, x_2, \ldots, x_m)$, we indicate with $t[j]$, with $j \in [1, m]$, the element x_j. When clear from the context we omit the subscript $\mathcal{P}$.

We model the presence of a *leader* in the system by stipulating that, in every initial configuration, there is exactly one agent in a distinguished state (or set of states).

2.2. Executions

The occurrence of an interaction $i = (a_j, a_k)$ turns a configuration of the form $C = (a_1, \ldots, a_j, \ldots, a_k, \ldots, a_n)$ into one of the form

$$C' = (a_1, \ldots, \delta(a_j, a_k)[0], \ldots, \delta(a_j, a_k)[1], \ldots, a_n)$$

This is indicated by the notation $C \xrightarrow{i} C'$. We say that an infinite sequence of interactions $I = (i_0, i_1, \ldots)$ is a *run* of $\mathcal{P}$. Starting from an initial configuration C_0, a run I induces an infinite sequence of configurations, $\Gamma_I(C_0) = (C_0, C_1, \ldots)$ such that $C_j \xrightarrow{i_j} C_{j+1}$ for every $j \geq 0$. This sequence of configurations is an *execution* of $\mathcal{P}$.

2.3. Non Omissive Interaction Models

Depending on the conditions imposed on the transition function, three main models of interactions have been identified: the standard *two-way* model and the one-way models, *immediate transmission* and *immediate observation*, presented in [4].

Two-Way Interaction Model (TW). The state transition function consists of two functions $f_s \colon Q_{\mathcal{P}} \times Q_{\mathcal{P}} \to Q_{\mathcal{P}}$ and $f_r \colon Q_{\mathcal{P}} \times Q_{\mathcal{P}} \to Q_{\mathcal{P}}$, one for the starter and the other for the receiver respectively, with $\delta_{\mathcal{P}}(a_s, a_r) = (f_s(a_s, a_r), f_r(a_s, a_r))$.

Immediate Transmission Model (IT). The state transition function consists of two functions $g\colon Q_{\mathcal{P}} \to Q_{\mathcal{P}}$ and $f\colon Q_{\mathcal{P}} \times Q_{\mathcal{P}} \to Q_{\mathcal{P}}$, with $\delta_{\mathcal{P}}(a_s, a_r) = (g(a_s), f(a_s, a_r))$. Note that, in the IT interaction model, the starter does not read the state of the reactor but it explicitly detects the interaction, as it applies function g to its own state.

Immediate Observation Model (IO). The state transition function has the form $\delta_{\mathcal{P}}(a_s, a_r) = (a_s, f(a_s, a_r))$. Note that, in the IO model, there is no detection of the interaction (or proximity) by the starter.

2.4. Omissions

An *omission* is a fault involving a single interaction. In an omissive interaction, an agent does not receive any information about the state of the other.[2] If omissions can occur in the system, then transition functions become more general relations. We consider the omissive models introduced in [24].

Two-Way Omissive Models. In the most general omissive model, T_3, the transition relation has the form

$$\delta_{\mathcal{P}}(a_s, a_r) = \{(f_s(a_s, a_r), f_r(a_s, a_r)),\ (o(a_s), f_r(a_s, a_r)),\ (f_s(a_s, a_r), h(a_r)),$$
$$(o(a_s), h(a_r))\}.$$

The first pair is the outcome of an interaction when no omission is present; the other three pairs represent all possible outcomes when there is an omission: respectively, an omission on the starter's side, on the reactor's side, and on both sides. The functions o and h represent the detection capabilities of each agent: omissions are *undetectable* on a certain side if the corresponding function is equal to the identity. From this we obtain the weaker models T_2 and T_1 depicted in Figure 3 (see [24] for more details).

One-Way Omissive Models. These models are defined by the transition

---

[2]This happens also in the IT and IO models, but it is due to the nature of one-way interactions, and not because of failures.

relation

$$\delta_{\mathcal{P}}(a_s, a_r) = \{(g(a_s), f(a_s, a_r)), \ (o(a_s), h(a_r))\}.$$

The first pair is the outcome of an interaction when no omission is present, and the second pair when there is an omission. Notice that, when no omissions are present, we obtain the IT model. When no omission are present and g is the identity function, we obtain the IO model. As for the previous case, omissions are undetectable starter-side if o is the identity function or if $o = g$. Moreover, if $h = g$, the reactor has detected the *proximity* of another agent, but it cannot read its state or even determine who is the starter and who is the reactor. Collectively, these variations give rise to models I_1 to I_4 in Figure 3. Other combinations of omissions and detections are possible, but they are provably equivalent to some of the aforementioned ones (see [24] for more details). When omissions are considered, the interaction between a_s and a_r is denoted by the ordered pair (a_s, a_r) and additional information specifying if the interaction is omissive and on which side. Omissions are introduced by an adversarial entity. We consider two types of adversaries:

Definition 1 ([24] Unfair Omissive (UO) Adversary). *The* **UO** *adversary takes a run I and outputs a new sequence I', which is obtained by inserting a (possibly empty) finite sequence of omissive interactions between each pair of consecutive interactions of I.*

Definition 2 ([24] Eventually Non-Omissive ($\diamond$**NO**/$\diamond$**NO$_1$**) Adversary). *The* $\diamond$**NO** *adversary takes a run I and outputs a new sequence I', which is obtained by inserting any finite sequence of omissive interactions between finitely many pairs of consecutive interactions of I. The $\diamond$**NO$_1$** adversary is even weaker, and can only output interaction sequences with at most one omission.*

2.4.1. Fairness

We say that a set of configurations $\mathcal{C} \subseteq Q_{\mathcal{P}}^n$ is *closed* if, for every $C \in \mathcal{C}$, and for every configuration $\widehat{C}$, which is a permutation of C, also $\widehat{C} \in \mathcal{C}$.
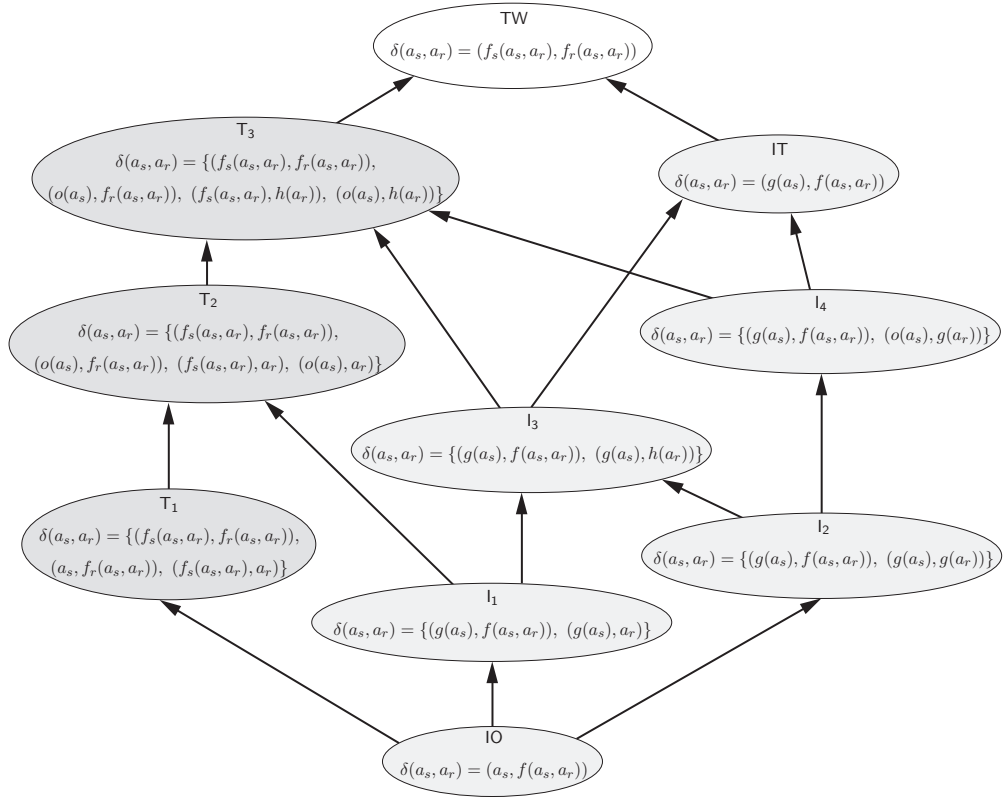
Figure 3: Interaction models (up to equivalence) and their computational relationships. An arrow between two blobs indicates that the class of solvable problems in the source blob is included in that of the destination blob. The models on the left, T_1, T_2, T_3, are the two-way models with omissions. The models on the right, I_1, I_2, I_3, I_4, are the one-way models with omissions.

We say that an execution Γ is *globally fair* (GF) when: for every two (possibly infinite) closed sets of configurations $\mathcal{C}, \mathcal{C}' \subseteq Q_{\mathcal{P}}^n$ such that for every $C \in \mathcal{C}$ there exists a non-omissive interaction i and some $C' \in \mathcal{C}'$ such that $C \xrightarrow{i} C'$ if infinitely many configurations of Γ belong to $\mathcal{C}$, then infinitely many configurations of Γ belong to $\mathcal{C}'$ (notice that this does not necessarily mean that a configuration in $\mathcal{C}'$ is an immediate successor of configurations of $\mathcal{C}$ in the execution Γ).

This definition of global fairness, introduced in [24], extends the standard one, which only deals with single configurations, as opposed to sets (see [8]). When we consider protocols that use only finitely many states the two definitions are equivalent, but when we consider protocols that work with infinitely many states only the extended definition is effective.

2.5. Simulation of Two-Way Protocols

In this section we define the *two-way protocol simulator*, we first give an intuitive explanation of this concept, and then we formally define it.

Informal Definition. Let $\mathcal{P}$ be a two-way protocol, and let $\mathcal{S}(\mathcal{P})$ be any protocol (which could be one-way, omissive, or both). Next we are going to informally define what it means for $\mathcal{S}(\mathcal{P})$ to *simulate* $\mathcal{P}$; the formal definition can be found later in the text.

We want the set of local states of $\mathcal{S}(\mathcal{P})$ to be of the form $Q_{\mathcal{P}} \times Q_{\mathcal{S}}$, where $Q_{\mathcal{P}}$ is the set of local states of $\mathcal{P}$ (the "simulated states"), and $Q_{\mathcal{S}}$ is some additional memory space used in the simulation. Suppose to start an execution of $\mathcal{S}(\mathcal{P})$ on a system of $n \geq 2$ agents from a given initial configuration. An initial configuration for $\mathcal{S}(\mathcal{P})$ is such that, when restricted to the simulated states, it is an initial configuration of $\mathcal{P}$. Agents are allowed to freely change the $Q_{\mathcal{S}}$ component of their local states; but when they change their $Q_{\mathcal{P}}$ component, we want the change to reflect the transition function of $\mathcal{P}$. That is, if $\delta_{\mathcal{P}}(a_s, a_r) = (f_s(a_s, a_r), f_r(a_s, a_r))$, then for every agent whose simulated state changes from a_s to $f_s(a_s, a_r)$, there must be some other agent (at some point in time) whose simulated state changes from a_r to $f_r(a_s, a_r)$. Moreover, there must be a perfect matching between such

13

transitions, in such a way that each starter of a simulated transition can be implicitly mapped to an appropriate reactor. Also, such a perfect matching must be "consistent", i.e., there must be an ordering of the simulated two-way interactions that is an execution of the simulated protocol.

We additionally require that, if the execution of $\mathcal{S}(\mathcal{P})$ is globally fair, then also the resulting simulated execution of $\mathcal{P}$ is globally fair.

Formal Definition. Given a two-way protocol $\mathcal{P}$, consider a protocol $\mathcal{S}(\mathcal{P})$, whose set of local states is $Q_\mathcal{P} \times Q_\mathcal{S}$, where $Q_\mathcal{P}$ is the set of "simulated states", that are local states of $\mathcal{P}$, and the set $Q_\mathcal{S}$ is additional memory space used in the simulation. The function $\pi_\mathcal{P} \colon Q_\mathcal{P} \times Q_\mathcal{S} \to Q_\mathcal{P}$ is the projection function onto the set of local states of $\mathcal{P}$. Given a configuration C of $\mathcal{S}(\mathcal{P})$, we write $\pi_\mathcal{P}(C)$ to indicate the configuration of $\mathcal{P}$ consisting of the projections of the states of the agents of C.

Given an execution $\Gamma_I(C_0)$ of $\mathcal{S}(\mathcal{P})$, where $I = (i_0, i_1, \dots)$, we say that $E(\Gamma_I(C_0)) = (e_0, e_1, \dots)$ is a sequence of *events* for $\Gamma_I(C_0)$ if it is a weakly increasing sequence of indices of interactions of I, with the following two constraints: no three indices are the same, it contains at least the indices of the interactions that determine the update of the simulated state of some agent in the execution $\Gamma_I(C_0)$ (if an interaction updates the simulated states of two agents, then its index must appear twice in $E(\Gamma_I(C_0))$). Therefore, at each event e_j in $E(\Gamma_I(C_0))$, we can associate a unique agent involved in the interaction i_{e_j}; preferably, this agent is one that changes simulated state as a result of i_{e_j}. We allow in $E(\Gamma_I(C_0))$ also events associated with agents that do not change simulated state, this is necessary to simulate two-way protocols that could leave the state of an agent unmodified.

If $\Gamma_I(C_0) = (C_0, C_1, \dots)$, we let $C_j^- = C_{e_j}$ and $C_j^+ = C_{e_j+1}$; the configurations C_j^- and C_j^+ are the ones before and after the j-th update of the simulated state, respectively.

Definition 3 ([24] Perfect matching of events). *Given an execution $\Gamma_I(C_0)$ of a run I and a sequence of events $E = E(\Gamma_I(C_0))$, a perfect matching $M(E)$ is*

a partition of $\mathbb{N}$ *into ordered pairs (viewed as indices of events of* $E(\Gamma_I(C_0))$*) such that, if* $(e_j, e_k) \in M(E)$*, where* e_j *is associated with agent* a_x *and* e_k *with agent* a_y*, then* $x \neq y$ *and*

$$\delta_{\mathcal{P}}(\pi_{\mathcal{P}}(C_j^-[x]), \pi_{\mathcal{P}}(C_k^-[y])) = (\pi_{\mathcal{P}}(C_j^+[x]), \pi_{\mathcal{P}}(C_k^+[y])).$$

Intuitively, a pair (e_j, e_k) in the perfect matching represents a simulated two-way interaction of the simulated protocol $\mathcal{P}$: where the event e_j corresponds to the update of the simulated state in the starter, and, similarly, e_k corresponds to the update of the simulated state in the reactor.

Given a matching $M(E)$, we can obtained a *derived run* D of the simulated protocol $\mathcal{P}$. This run is obtained as follows: sort the pairs (e_j, e_k) of $M(E)$ by increasing $\min\{e_j, e_k\}$, obtaining the sorted sequence M'. Now, we consider the m-th element of M', let it be (e'_j, e'_k), where the event e'_j is associated with agent a_x and e'_k with agent a_y, and we insert in D as m-th element the interaction (x, y). It is clear, that the run D of $\mathcal{P}$ induces a *derived execution* $\Gamma_D(\pi_{\mathcal{P}}(C_0))$.

Definition 4 ([24] Simulation)**.** *A protocol* $\mathcal{S}(\mathcal{P})$ *simulates* $\mathcal{P}$ *if, for any initial configuration* C_0 *of n agents of* $\mathcal{S}(\mathcal{P})$*, and any run* I *whose execution* $\Gamma_I(C_0)$ *satisfies the* GF *condition, there exists a sequence of events* $E(\Gamma_I(C_0))$ *with a perfect matching* $M(E)$ *whose derived execution is an execution of n agents of* $\mathcal{P}$ *starting from the initial configuration* $\pi_{\mathcal{P}}(C_0)$ *and satisfying the* GF *condition. We further require that, for each initial configuration* C_0*, every finite initial sequence of interactions of* $\mathcal{S}(\mathcal{P})$ *(possibly with omissions) can be extended to an infinite one* I*, having no additional omissions, whose execution* $\Gamma_I(C_0)$ *satisfies the* GF *condition.*

3. Simulation with a Leader in Omissive Models: Impossibility

In this section we prove that the presence of a leader, alone, might not be sufficient to overcome dynamic omissions. Indeed, we prove that there are two-way protocols that cannot be simulated with omissive interactions even if a leader is present.

Next we consider the *Pairing Problem* introduced in [24]: a set of agents A is given, partitioned into *consumer* agents A_c, starting in state c, and *producer* agents A_p, starting in state p. Informally, we want a consumer to enter a specific state cs if and only if it is paired with a producer, and we want to pair each producer with at most one consumer.

We say that a protocol $\mathcal{P}$ solves the *Pairing Problem* if it enforces the following properties:

(i) *Irrevocability:* $\mathcal{P}$ has a state cs that only agents in state c can reach; once an agent has state cs, its state cannot change any more;

(ii) *Safety:* At any time, the number of agents in state cs is at most $|A_p|$;

(iii) *Liveness:* In all globally fair executions of $\mathcal{P}$, eventually the number of agents in state cs is equal to $\min\{|A_c|, |A_p|\}$.

This problem can be solved in the standard two-way model by the simple protocol below:

Pairing Protocol $\mathcal{P}_{PP}$. $Q_{\mathcal{P}_{PP}} = \{cs, c, p, \bot\}$. The only non-trivial transition rules are $(c, p) \mapsto (cs, \bot)$ and $(p, c) \mapsto (\bot, cs)$.

However, as we will show, this protocol cannot be simulated, in spite of the presence of a leader, even in the strongest of the omissive models.

3.1. Impossibility with Finite Memory

We investigate what happens when we introduce a distinguished leader node, but we restrict the memory of agents to be bounded. We show our impossibility results directly for the T_3 omissive model. The results clearly carry over to all the less powerful omissive models.

Definition 5. (Omission-Recurrent Configuration) *Let $\{\ell, a\}$ be a system of two agents and the communication model be T_3, where ℓ is the leader, and let $C = (q_\ell, q_a)$ be a configuration. Suppose that there exists a finite non-empty*

sequence of interactions $I = (i_1, i_2, \ldots, i_t)$, where i_1 is omissive on both sides, such that, if I is executed according to the transition rules of a population protocol $\mathcal{P}$ starting from configuration C, eventually the state of ℓ is again q_ℓ. Then, if ℓ is the starter (respectively, the reactor) of i_1, we say that C is a starter-omission-recurrent configuration (respectively, an reactor-omission-recurrent configuration) for protocol $\mathcal{P}$.

Note that, in the above definition, since i_1 is omissive on both sides, the system transitions into configuration $(o(q_\ell), h(q_a))$ or $(h(q_\ell), o(q_a))$ after executing i_1. Intuitively, if a configuration $C = (q_\ell, q_a)$ is omission-recurrent, then it is possible to make the leader agent ℓ, in state q_ℓ, "loop back" to state q_ℓ after an omission. This property will be used in our impossibility proof to make ℓ interact in the same state with multiple agents without noticing it.

To do so, we first prove a structural property of simulators of a system of two agents. In Lemma 1 we show that, every time the two agents are able to exchange their states, they are in a configuration that is omission recurrent, and indeed the simulation of a two-way interaction is based exclusively on the information contained in omission recurrent configurations.

Lemma 1. Let $\mathcal{S}(\mathcal{P})$ be a simulator having a finite number k of states in total and working under the $\diamond$**NO** adversary. Let a system of two agents $\{\ell, a\}$ be given, where ℓ is the leader. Let C_0 be an initial configuration for ℓ and a, and let $I = (i_1, i_2, \ldots)$ be an infinite sequence of interactions with no omissions between ℓ and a such that, if $\mathcal{S}(\mathcal{P})$ is executed according to I starting from C_0, then the execution is globally fair. Then there exists a finite sequence of interactions $I' = (i'_1, i'_2, \ldots i'_t)$ with the following properties.

(1) I' is obtained by introducing at most k omissive interactions into an initial finite sub-sequence of I. All the omissive interactions of I' are omissive on both sides.

(2) If ℓ and a execute $\mathcal{S}(\mathcal{P})$ according to I' starting from C_0, they both do a simulated state transition (according to $\delta_\mathcal{P}$).

(3) *Suppose that ℓ and a execute $\mathcal{S}(\mathcal{P})$ according to I' starting from C_0, and let C_j be the configuration of ℓ and a immediately before executing interaction i'_{j+1}. Then, if i'_{j+1} is not omissive and has ℓ as the starter (respectively, reactor), C_j is starter-omission-recurrent (respectively, reactor-omission-recurrent) for $\mathcal{S}(\mathcal{P})$.*

Proof. First we will insert at most k omissive interactions into I (thus building an infinite sequence that satisfies property (1)) in such a way as to satisfy property (3). Then we will choose t so as to satisfy property (2).

We will construct I' incrementally by an inductive procedure. Suppose we have constructed I' up to i'_j, and let C_j be the configuration of ℓ and a after executing the first j interactions of I' starting from C_0 (the base case is with $j = 0$). Suppose also that the partial sequence $(i'_1, \ldots, i'_j)$ has been obtained by adding some omissive interactions to some initial sub-sequence of I, say $(i_1, \ldots, i_{v(j)})$ (with $v(0) = 0$, i.e., in the base case the sub-sequence is empty). Let $i_{v(j)+1}$ have ℓ as its starter (respectively, reactor). Then, if C_j is starter-omission-recurrent (respectively, reactor-omission-recurrent), we set $i'_{j+1} = i_{v(j)+1}$ and $v(j+1) = v(j) + 1$. Otherwise, we let i'_{j+1} be omissive on both sides with ℓ as the starter (respectively, reactor), and we set $v(j+1) = v(j)$. Notice that, by constructing the sequence in the aforementioned way we are satisfying property (3): for any interaction i'_{j+1} that is not omissive between ℓ and a, we have that C_j is either reactor-omission-recurrent or starter-omission-recurrent. We claim that, if we continue this process indefinitely, we put at most k omissive interactions in I'. Indeed, suppose that an omissive interaction i'_j with ℓ as the starter (respectively, reactor) has been inserted in I', which means that C_{j-1} is not a starter-omission-recurrent (respectively, reactor-omission-recurrent) configuration. By definition of omission-recurrent configuration, ℓ will never get the simulated state it had in C_{j-1} after executing i'_j. It follows that the same state will contribute to the addition of at most one omissive interaction to I'. Since the possible states for ℓ are at most k, there can be at most k omissive interactions in I'.

18

We now have to decide when to stop the incremental construction of I'. Recall that the execution of I starting from C_0 is globally fair, and observe that adding finitely many omissive interactions to it preserves its global fairness. So, by definition of simulator, the simulated states of ℓ and a also change in a globally fair way as they execute I'. Therefore, at some point, they will conclude a simulated interaction, changing their local states according to $\delta_{\mathcal{P}}$. At this point we stop the construction of I', obtaining a sequence of finite length t that satisfies all three properties. $\qquad\square$

In the next theorem (Theorem 1) we prove our impossibility results. In the proof, we show that there exists a system of several agents where it is possible to trick the leader ℓ and two other agents to believe that they are in a system of only two agents. Each of these three agents sees a sequence of interactions I that respects the property dictated by Lemma 1. From this we show that it is impossible to simulate the aforementioned Pairing Protocol.

Theorem 1. *A system of agents, each of which has a finite amount of memory, cannot simulate every two-way protocol in the T_3 model (hence in all the omissive models), even with the presence of a leader and under the $\diamond\mathbf{NO}$ adversary.*

Proof. We will show that the Pairing Protocol $\mathcal{P}_{PP}$ cannot be simulated in T_3. Suppose by contradiction that there is a simulator $\mathcal{S}(\mathcal{P}_{PP})$ for it, and let us consider a system of two agents ℓ and a, where ℓ is the leader. Let C_0 be an initial configuration in which ℓ has simulated state p and a has simulated state c, and let I be an omission-less infinite sequence of interactions for the two agents whose execution starting from C_0 is globally fair. According to $\mathcal{P}_{PP}$, eventually I will make both ℓ and a change simulated state to $\perp$ and cs, respectively.

Let us apply Lemma 1 to $\mathcal{S}(\mathcal{P}_{PP})$, C_0, and I, which yields a sequence of interactions $I' = (i'_1, \ldots, i'_t)$ of length t, some of which are omissive on both sides. The sequence I' guarantees that both agents will change simulated state to cs and $\perp$, as per property (2). Let C_j be the configuration of the two agents after executing the first j interactions of I starting from configuration C_0.

19

Now we construct a larger system of agents: $\{\ell, a, b_1, b_2, \ldots, b_m, d\}$, where $m = 2^t$. Let C_0' be the configuration of this system in which ℓ and a have the same state as in C_0 and all other agents have the same state as a. We will show how to modify I' by inserting some extra interactions in it, obtaining an expanded sequence I'' that involves also the other members of the system (as opposed to only ℓ and a). We will then show that executing I'' makes the simulator behave in a way that is not compatible with the Pairing Problem.

We will construct I'' inductively by inserting a sequence of interactions before each interaction $i_j' \in I'$. Say we have already done so up to i_j', with $0 \le j < t$ (the base case being with $j = 0$). Executing I'' up to this point from configuration C_0' makes the system reach configuration C_j'. Let $w(j) = m/2^j$, and suppose that agents $a, b_1, b_2, \ldots, b_{w(j)}$ have the same state in C_j' (this is certainly true for $j = 0$). Also suppose that, in C_j', ℓ and a have the same state as in C_j (again, this is true for $j = 0$). As we construct I'' and j increases, we will also prove that these properties are preserved.

Now consider the next interaction i_{j+1}', which could be either non-omissive or omissive on both sides. Next we are going to explain what interactions we add to I'' between i_j' and i_{j+1}'.

Suppose that i_{j+1}' is omissive on both sides, and let ℓ be the starter and a the reactor. Then we introduce in I'' (right after i_j') the sequence of interactions $(d, b_1), (d, b_2), \ldots, (d, b_{w(j)/2})$, omissive on both sides. Finally we introduce i_{j+1}' into I''. If a is the starter and ℓ the reactor, we insert the same interactions, but with starter and reactor exchanged. It is immediate to see that, after executing these interactions from configuration C_j', agents $a, b_1, b_2, \ldots, b_{w(j+1)}$ have equal states. Indeed, they have the same state in C_J', and then they all execute one omissive interaction as starters or as reactors. Moreover, in C_{j+1}', ℓ and a have the same state as in C_{j+1}.

Suppose that i_{j+1}' is not omissive, and let ℓ be the starter and a the reactor. By property (3) of I', configuration C_j is starter-omission-recurrent. Let q_ℓ be the state of ℓ in C_j (and therefore in C_j'). By definition of starter-omission-recurrent configuration, there exists a sequence of interactions $I^* = (i_1^*, i_2^*, \ldots, i_z^*)$, with i_1^*

omissive on both sides and having ℓ as the starter and a as the reactor, such that the state of ℓ becomes q_ℓ again if I^* is executed from configuration C_j (and hence from C_j'). Note that the same happens if the partner of ℓ in the interactions of I^* is not a but any of the b_x's, with $1 \leq x \leq w(j)$, since all these agents have the same state in C_j' by inductive hypothesis. Given these premises, we introduce in I'' (right after i_j') the following interactions.

- For all $1 \leq x \leq w(j)/2$, we insert:

    - the interaction (ℓ, b_x), omissive on ℓ's side;
    - the interaction $(d, b_{x+w(j)/2})$, omissive on both sides;
    - the sequence of interactions $(i_2^*, i_3^*, \ldots, i_z^*)$, with $b_{x+w(j)/2}$ as ℓ's partner instead of a.

- Finally, we insert i_{j+1}'.

The case in which ℓ is the reactor of i_{j+1}' and a the starter is handled in a similar fashion, but we exchange starter and reactor in the interactions that we add to I''. Suppose now that the system executes the above sequence of interactions starting from configuration C_j', and let us focus on the sub-system consisting of ℓ and $b_{x+w(j)/2}$, where $x \in [1, w(j)/2]$. The agent ℓ starts in state q_ℓ and then executes an omissive interaction, while $b_{x+w(j)/2}$ executes another omissive interaction. Together, these two interactions have the same effect on ℓ and $b_{x+w(j)/2}$ as i_1^* (note that the partners of ℓ and $b_{x+w(j)/2}$ were irrelevant, because the interactions were omissive on their sides). Then ℓ and $b_{x+w(j)/2}$ execute all the interactions of I^* except i_1^*. Since $b_{x+w(j)/2}$ started in the same state as a, by definition of I^* it follows that the state of ℓ is reset again to q_ℓ after this sequence. By induction, this is true for all $1 \leq x \leq w(j)/2$. Finally, the state of ℓ correctly changes according to i_{j+1}' as it interacts with a. On the other hand, a and all the b_x's see ℓ exactly once when it is in state q_ℓ and, since they have the same state in C_j', they also have the same state in C_{j+1}'.

We have shown that agents a, b_1, b_2, ..., $b_{w(j)}$ have the same state in C_j' for all $1 \leq j \leq t$. In particular, for $j = t$, we have that $w(j) = m/2^t = 1$, which

21

means that a and b_1 have the same state in C'_t. In turn, the simulated state of a in C'_t is the same as in C_t, i.e., cs. Since at the beginning there was only one agent with simulated state p (i.e., ℓ), and now we have two agents with simulated state cs, we have violated the safety property of the Pairing Problem, meaning that $\mathcal{S}(\mathcal{P}_{PP})$ cannot be a simulator.

As there is only a finite number of omissions in I'', this sequence of interactions can be extended to an infinite one with the addition of non-omissive interactions, which is compatible with the $\diamond\textbf{NO}$ adversary. $\qquad\square$

3.2. Impossibility with Infinite Memory

For this case we can show that simulation is impossible in the omissive two-way model without detection, and thus in $\textsf{IO}$.

Theorem 2. *A system of agents, each of which has an infinite amount of memory, cannot simulate every two-way protocol in the $\textsf{T}_1$ model (hence in $\textsf{IO}$), even with the presence of a leader and under the $\diamond\textbf{NO}$ adversary.*

Proof. We will show that the Pairing Protocol $\mathcal{P}_{PP}$ cannot be simulated in $\textsf{T}_1$. Let $\mathcal{S}(\mathcal{P}_{PP})$, ℓ, a, C_0, and $I = (i_1, i_2, \ldots)$ be defined as in the first paragraph of the proof of Theorem 1. By definition of simulator and by the Pairing Problem, if we execute I from C_0 according to $\mathcal{S}(\mathcal{P}_{PP})$, at some point we reach a configuration in which ℓ has simulated state $\bot$ and a has simulated state cs. Say that this happens after executing i_j, and let $I_j = (i_1, i_2, \ldots, i_j)$.

Let us now extend the system with a third agent b, initially having the same state as a in C_0, and let us show how to insert interactions into I_j involving b as well, in order to obtain a contradictory finite sequence of interactions I'. Recall that I is omission-less, and consider the interaction i_x, with $1 \leq x \leq j$. If $i_x = (\ell, a)$, we insert the interaction (ℓ, b) right before it, with omission on ℓ's side. If $i_x = (a, \ell)$, we insert the interaction (b, ℓ) right before it, again with omission on ℓ's side.

Since omissions are undetectable, it is easy to see that the extended sequence I' will make ℓ undergo the same state transitions as I_j (but at half the "speed").

On the other hand, a and b will always see ℓ in the same state and will never see each other, so they will both have the same state throughout the execution of I'. It follows that a and b will eventually have simulated state cs, which violates the safety property of the Pairing Problem.

Note that the sequence I' contains finitely many omissions, and therefore it can be extended to an infinite sequence that is compatible with the $\diamond\textbf{NO}$ adversary. $\qquad\square$

Observation 1. *Since in* IO *there are no omissions, the statement of Theorem 2 for the* IO *model trivially extends to the scenario in which the number of omissions in the sequence of interactions is known in advance by the agents.*

4. Simulation in Omissive Models

In this section we are going to make use of a result that appears in [24] as Theorem 4.5. This theorem assumes each agent to have a unique ID, which is a non-negative integer, as part of its local state.

Theorem 3. *Assuming* IO*, unique IDs, and* $\mathcal{O}(\log(\max\text{ID}))$ *bits of memory on each agent (where* $\max$ ID *is the maximum ID in the system), there exists a simulator for every two-way protocol, even under the* $\textbf{UO}$ *adversary.* $\qquad\square$

What this theorem says is that, if the agents initially have unique IDs, they can perform a simulation of any two-way protocol, even if the simulation runs in the weakest model, IO, and against the strongest adversary, **UO**. The technique used in [24] to prove Theorem 3 is to use IDs to implement a locking mechanism between agents: if an agent a is seen by agent b in a state that indicates that a is looking for a matching agent, then agent b uses the ID of a and its own ID to put a in a state in which it is forced to pair with b. In this section we assume the presence of a leader and we show that, in certain models, we can implement a *naming algorithm*, i.e., an algorithm that assign unique IDs to all agents. Once an ID has been assigned to an agent, it cannot change. Therefore, the naming algorithm and the simulator of Theorem 3 can be combined into a single protocol

and where they run in parallel: if an agent has no ID yet, the simulator simply ignores every interaction involving this agent. By global fairness, eventually all agents will have unique IDs, and the simulation will finally involve the entire system, producing a globally fair simulated execution.

The protocols will be presented using an algorithmic style: for each interaction of the form (a_s, a_r), the starter agent a_s executes function Upon Event Starter sends() and the reactor agent a_r executes Upon Event Reactor receives (var^s), where var^s is the variable var in the local state of agent a_s.

4.1. Naming Algorithm with Infinite Memory

If the leader has infinite memory, it can implement a simple naming algorithm under certain models. Since Theorem 2 already states the impossibility of simulation under models T_1 and IO, we will assume model I_1 or model I_2. Constructing a simulator for these models will imply the existence of a simulator for all other models except T_1 and IO (refer to Figure 1).

Theorem 4. *Assuming* I_1 *or* I_2, *the presence of a leader, and an infinite amount of memory on each agent, there exists a simulator for every two-way protocol, even under the* **UO** *adversary.*

Proof. By the above discussion, it suffices to implement a naming algorithm. Each agent has a local variable my_ID and the leader also has a second variable $next\_ID$. The leader has $my\_ID = 0$ and $next\_ID = 1$, while all other agents initially have $my\_ID = \bot$. Every time the leader detects the proximity of another agent (i.e., it applies function g to its own state), it increments the variable $next\_ID$. Every time an agent, who does not have yet an ID, sees the state of the leader (i.e., it applies function f), it sets its own my_ID to the value found in the leader's $next\_ID$ variable.

Since the leader increments $next\_ID$ every time it is involved in an interaction (even if the interaction is omissive on the other side), no two agents can get the same ID. Moreover, by global fairness, all agents will eventually see the leader in a non-omissive interaction, and will therefore acquire an ID. □

4.2. Naming Algorithms with Knowledge on Omissions

Now we assume that agents have only a finite amount of memory, but they know in advance a finite upper bound L on the number of omission faults that the adversary is going to insert *in interactions that involve the leader*. Note that the adversary can still be **UO** even if only finitely many omissive interactions involve the leader.

4.2.1. Naming Algorithm for I_1 and I_2

We refine the naming algorithm of Theorem 4 to cope with the fact that now memory is bounded by a function of L and the size of the system, n. It is worth mentioning that the precise value of n is not known to the agents, and L is only an upper bound on the number of omissions involving the leader, not necessarily the exact number.

Theorem 5. *Assuming I_1 or I_2, the presence of a leader, knowledge of an upper bound L on the number of omission failures in interactions that involve the leader, and $\Theta(L \log nL)$ bits of memory on each agent (where n is the number of agents), there exists a simulator for every two-way protocol, even under the* **UO** *adversary.*

Proof. We implement the naming algorithm presented in Figure 4. Compared to the algorithm of Theorem 4, here the leader has an array of $L + 1$ $next\_ID$ variables, as opposed to only one. This array is initialized to $[1, 2, \ldots, L+1]$ and, when an ID is assigned, the corresponding entry of the array will be incremented by $L + 1$, so that no two equal IDs can be generated.

All entries of $next\_ID$ are initially *unlocked*: this information is stored in the leader's Boolean array *locked*. The *active ID* is defined as the unlocked entry of $next\_ID$ having minimum index, if there is any (line 10). This is the ID that will tentatively be assigned next. Whenever the leader detects the proximity of another agent (i.e., it executes function g on its own state, or function Upon Event Starter sends in the algorithm of Figure 4), it locks the active entry of $next\_ID$ (line 12). The purpose of locking an entry of $next\_ID$ (as opposed to

just incrementing it as in Theorem 4) is that the leader cannot allow its value to grow indefinitely, because now memory is limited. Instead, the leader will make the entry temporarily inactive, and will keep it on hold until it gathers more information in the following interactions.

On the other hand, if an agent a sees the leader (i.e., it executes function f or function Upon Event Reactor receives in Figure 4), and a does not have an ID yet, then it assigns itself the active ID from the leader's $next\_ID$ variable (line 30). So, the next time the leader sees a, it will read its new ID and it will know that the corresponding entry of $next\_ID$ can be unlocked (line 23) and its value can be incremented by $L + 1$ (line 24).

It may happen that the leader is involved in an omissive interaction, and therefore the entry of $next\_ID$ that it locks will never be unlocked again. However, this can happen at most L times, while the array has $L + 1$ entries.

This is not sufficient yet, because the same agent a may see the leader multiple times in a row and cause all entries of $next\_ID$ to become locked. If a only stores one ID, it will have no way to tell the leader that more than one entry of $next\_ID$ has to be unlocked. This is why a also has a variable called $redundant$, which is a Boolean array that will store information on all the active entries of $next\_ID$ that a sees after receiving an ID. So, if the agent a already has an ID and it sees the leader again, it sets to $true$ the entry of $redundant$ corresponding to the active ID of the leader (line 32).

Now, suppose that the leader sees that a has an entry of $redundant$ set to $true$. This implies that the corresponding entry of $next\_ID$ is currently locked and should be unlocked. However, this cannot be done right away: the leader wants to give a an "acknowledgment", so that a will set the entry of $redundant$ to $false$ first. This is to prevent the scenario in which the entry of $next\_ID$ gets unlocked, becomes active, another agent b sees it, and takes it as its own ID. If then the leader sees a again (still with $redundant$ on $true$), it will unlock the entry of $next\_ID$. Then perhaps yet another agent c will see the leader, getting the same ID as b.

To prevent such an incorrect behavior, the leader has another variable array

called *waiting*, in which it stores the IDs of the agents that should reset their *redundant* variables. The size of such an array is $L + 1$, since at most $L + 1$ different entries of *next_ID* can be locked by agents with IDs. So, when the leader sees that a has some entry of *redundant* set to *true*, it stores the (unique) ID of a in the corresponding entry of *waiting* (line 18). Then, when a sees the leader again and reads its own ID in the *waiting* array, it knows that it has to set to *false* the corresponding entries of *redundant* (line 34). Finally, when the leader sees a again and notices that the entry of *redundant* has been set to *false*, it can reset the corresponding entry of *waiting* (line 20) and unlock the entry of *new_ID* (line 21).

The fact that the algorithm does not give the same ID to two different agents follows from the observation that at most one agent can keep an entry of *new_ID* locked at any given time, which in turn follows from the way the two variables *redundant* and *waiting* function together. If no omission occurs and the leader is observed by some agent a, then a will store information about the currently active ID. If a takes this ID for itself, that entry of *next_ID* will be incremented before any other agent can get the same ID. If a has already an ID, the entry of *next_ID* will remain locked until a has reset its own *redundant* variable. Moreover, the fact that the algorithm will eventually assign every agent an ID immediately follows from the global fairness of the adversarial scheduler.

Since the IDs in the *next_ID* array increase by $L + 1$ every time one is assigned, and since there are n agents in total, the value of every ID is $\mathcal{O}(nL)$. Hence, $\mathcal{O}(L \log nL)$ bits of memory are required to store each agent's arrays, and $\mathcal{O}(\log nL)$ more bits are required to run the simulator of Theorem 3. The total amount of memory needed per agent is therefore $\mathcal{O}(L \log nL)$ bits. □

4.2.2. Naming Algorithm for T_1

Observe that the naming protocol of Figure 4 does not work for model T_1, because it is based on the fundamental assumption that the leader can always detect an interaction. This is not possible in model T_1: if an omission occurs, is it possible for the leader to be the starter of an interaction without realizing

```
 1: Variables
 2: my_ID                                    ▷ the leader has this variable initialized to 0, non-leaders to ⊥
 3: next_ID[] := [1, 2, . . . , L + 1]                                              ▷ leader variable
 4: locked[] := [false, false, . . . , false]              ▷ leader variable, array of size L + 1
 5: waiting[] := [⊥, ⊥, . . . , ⊥]                          ▷ leader variable, array of size L + 1
 6: redundant[] := [false, false, . . . , false]           ▷ leader variable, array of size L + 1
 7:
 8: Upon Event Starter sends()
 9: if my_ID = 0 then                                       ▷ I am the leader
10:     j := min{j | locked[j] = false, L + 2}
11:     if j < L + 2 then
12:         locked[j] := true
13:
14: Upon Event Reactor receives (my_ID^s, next_ID^s[], locked^s[], waiting^s[], redundant^s[])
15: if my_ID = 0 then                                       ▷ I am the leader
16:     for all j ∈ {1, 2, . . . , L + 1} do
17:         if redundant^s[j] = true then
18:             waiting[j] := my_ID^s
19:         else if waiting[j] = my_ID^s then
20:             waiting[j] := ⊥
21:             locked[j] := false
22:     if ∃j, next_ID[j] = my_ID^s then
23:         locked[j] := false
24:         next_ID[j] := next_ID[j] + L + 1
25: else                                                    ▷ I am not the leader
26:     if my_ID^s = 0 then                                 ▷ my partner is the leader
27:         j = min{j | locked^s[j] = false, L + 2}
28:         if j < L + 2 then
29:             if my_ID = ⊥ then
30:                 my_ID := next_ID^s[j]
31:             else
32:                 redundant[j] := true
33:         if my_ID ≠ ⊥ ∧ ∃j, waiting^s[j] = my_ID then
34:             redundant[j] := false
```

Figure 4: Naming algorithm for l_1 and l_2 with knowledge on omissions, used in Theorem 5

```
 1: Variables
 2: role     ▷ the leader has this variable initialized to leader, non-leaders to
     available
 3: state_𝒫 := initial_state_𝒫
 4: token := ⊥
 5:
 6: Upon Event Starter sends()
 7: token := ⊥
 8: if role = leader then
 9:     role := available
10: else if role = moving then
11:     role := available
12: else if role = starter then
13:     role := pending
14:
15: Upon Event Reactor receives (role^s, state_𝒫^s, token^s)
16: token := token^s
17: if role^s = leader then
18:     role := moving
19: else if role^s = moving then
20:     role := starter
21: else if role^s = starter then
22:     token := state_𝒫
23:     state_𝒫 := f_r(state_𝒫^s, state_𝒫)
24: else if role = pending ∧ token ≠ ⊥ then
25:     state_𝒫 := f_s(state_𝒫, token)
26:     role := leader
27:     token := ⊥
```

Figure 5: Simulation protocol for IT with finite memory, used in Theorem 7

it. Moreover, Theorem 2 does not hold when some kind of upper bound on omissions is known.

Theorem 6. *Assuming* T_1*, the presence of a leader, knowledge of an upper bound L on the number of omission failures in interactions that involve the leader, and $\Theta(L \log nL)$ bits of memory on each agent (where n is the number of agents), there exists a simulator for every two-way protocol, even under the* **UO** *adversary.*

Proof. It is sufficient to give a naming algorithm. We modify the one used in Theorem 4 to work in T_1 with $\mathcal{O}(L \log nL)$ memory. The leader has the same two local variables, but each other agent has an array my_ID of $L + 1$ local variables, each of which is initially set to $\bot$. If an agent a sees the leader, and the local array my_ID of a has some entries still set to $\bot$, then a changes one of them from $\bot$ to the value of the leader's $next\_ID$ variable. On the other hand, if the leader sees an agent whose local array my_ID has some entries still set to $\bot$, it increments $next\_ID$. When all entries of an agent's array my_ID have been set, the entire array is taken as the agent's ID.

Since the execution is globally fair, eventually all agents will have their my_ID array completely set. Observe that, whenever the leader increments $next\_ID$, there is an agent a that removes one occurrence of $\bot$ from its local array my_ID, unless the interaction is omissive on a's side. But there can only be L such omissive interactions, which means that the maximum value of $next\_ID$ will be $\mathcal{O}(nL)$. So, $\Theta(L \log nL)$ bits of memory are enough for an agent to store its local array my_ID. By Theorem 3, combining this naming algorithm with the simulator does not require more than $\Theta(L \log nL)$ bits of memory on each agent.

Suppose for a contradiction that two agents receive equal IDs. Therefore, both agents have observed the leader $L+1$ times, and at the jth observation both agents must have read the same value in variable $next\_ID$, for all $1 \leq j \leq L+1$. So, the leader has failed to increment $next\_ID$ for at least $L + 1$ times, implying that there have been $L + 1$ omissive interactions involving the leader, which

contradicts the theorem's assumptions. Thus all agents receive distinct IDs, and the naming algorithm is correct. □

5. Simulation for IT

Notice that IT is the only finite-memory model for which the impossibility result of Theorem 1 does not hold (see Figure 1). It turns out that in this model we can implement a simulator that sequentializes the simulated two-way interactions via a token-passing technique.

Theorem 7. *Assuming* IT, *the presence of a leader, a constant amount of memory on each agent and at least $n > 2$ agents, there exists a simulator for every two-way protocol.*

Proof. The simulation algorithm is reported in Figure 5. Suppose we are given a two-way protocol $\mathcal{P}$ whose transition function is $\delta_{\mathcal{P}}(a_s, a_r) = (f_s(a_s, a_r), f_r(a_s, a_r))$. In our simulator, each agent has a local variable called $state_{\mathcal{P}}$, which is the state of $\mathcal{P}$ that the agent is simulating, plus an auxiliary variable $role$, which is used to coordinate the simulation. Initially, the role of one agent is $leader$, while all others are $available$. When the leader meets another agent a, the leadership is "transferred" to a: the role of the leader becomes $available$ (line 9), and the role of a becomes $moving$ (line 18). Note that the leader does not have to see the state of a to perform this operation.

The next agent b that sees a becomes the starter of a new simulated interaction: the role of b becomes $starter$ (line 20) and the role of a becomes $available$ again (line 11). Now, the first agent c that observes b becomes the reactor of the simulated interaction: it executes function f_r on its own simulated state using b's simulated state as part of the input (line 23), while b's role becomes $pending$ (line 13).

Now, in order for b to perform its side of the simulated transition, it has to retrieve the simulated state that c had before transitioning. To deliver this information to b, the agent c stores its own simulated state in a variable called $token$ before performing the transition (line 22). Now, as soon as an agent sees

31

c, it copies the token (line 16), while c erases its own copy (line 7). This token circulation protocol is executed until the token reaches b.

When b finally obtains the token, it uses it as part of the input to function f_s and changes its simulated state accordingly (line 25). Now both sides of the simulated transition have been performed correctly, and b resets its role to *leader* (line 26) and destroys the token (line 27). At this point we have exactly one agent whose role is *leader*, while all other agents have role *available*, as we had at the beginning. The next steps of the simulation are thus performed in the same fashion.

Note that, if the two-way protocol $\mathcal{P}$ has a constant number k of states, then our simulator has $\mathcal{O}(k^2)$ states, independently of the size of the system.

The correctness of the simulator can be proven by observing that, due to the uniqueness of the leader, there is at most one pending transition at all times. Moreover, any agent in the system can become the starter (including the leader itself), thanks to the extra step that creates an agent with role *moving*: any non-leader agent can become *moving*, and then any agent (including the original leader) can become *starter*. Note that this is not true if the system consists of $n = 2$ agents (in this case the leader will necessarily become the starter of every simulated interaction), but by hypothesis we have $n > 2$ agents.

We have to prove that the system will perform infinitely many simulated interactions, in such a way that the simulated execution is globally fair. By the global fairness of the simulator, a token will certainly be created and will be passed around by the agents; again by global fairness, the token will eventually reach the agent with $role = starter$, and the simulated interaction will be concluded. The global fairness of the entire simulated execution also follows immediately from the global fairness of the simulator itself. $\qquad\square$

6. Conclusions

In this paper we investigated the impact of the presence of a leader on the problem of simulating traditional population protocols under weaker models of

communication, which include omission faults. The obtained results show that the presence of a leader effectively allows the simulation when there is an upper bound on the number of omissions or memory is not constrained. This is in contrast with what happens where a leader is not present, see [24]. However, when the memory is constrained and there is no upper bound on omissions, despite the presence of the leader, simulations are impossible in presence of omission faults. We remark that, since the impossibility holds for any amount of finite memory, it also holds for uniform message passing algorithms where the communication is dynamic and point to point, showing that implementing an atomic exchange of messages between two parties s not trivial under omission faults. An interesting open problem is to determine which problems become unsolvable when omissions are present, so to gain an insight into the computational challenges introduced by omissions.

References

[1] Alistarh, D., & Gelashvili, R. (2015). Polylogarithmic-time leader election in population protocols. In *42th International Colloquium on Automata, Languages and Programming, ICALP* (pp. 479–491).

[2] Alistarh, D., Gelashvili, R., & Vojnovic, M. (2015). Fast and exact majority in population protocols. In *4th Annual ACM Symposium on Principles of Distributed Computing, PODC* (pp. 47–56).

[3] Angluin, D., Aspnes, J., Diamadi, Z., Fischer, M., & Peralta, R. (2006). Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, *18*, 235–253.

[4] Angluin, D., Aspnes, J., & Eisenstat, D. (2005). On the power of anonymous one-way communication. In *9th International Conference on Principles of Distributed Systems, OPODIS* (pp. 396–411).

[5] Angluin, D., Aspnes, J., & Eisenstat, D. (2006). Stably computable predicates are semilinear. In *25th Annual ACM Symposium on Principles of Distributed Computing, PODC* (pp. 292–299).

[6] Angluin, D., Aspnes, J., & Eisenstat, D. (2008). Fast computation by population protocols with a leader. *Distributed Computing*, *21*, 61–75.

[7] Angluin, D., Aspnes, J., & Eisenstat, D. (2008). A simple population protocol for fast robust approximate majority. *Distributed Computing*, *2*, 87–102.

[8] Aspnes, J., & Ruppert, E. (2007). An introduction to population protocols. In *Bulletin of the European Association for Theoretical Computer Science* (pp. 98–117). volume 93.

[9] Beauquier, J., Burman, J., Clavière, S., & Sohier, D. (2015). Space-optimal counting in population protocols. In *29th International Symposium on Distributed Computing, DISC* (pp. 631–649).

[10] Beauquier, J., Burman, J., & Kutten, S. (2011). A self-stabilizing transformer for population protocols with covering. *Theoretical Computer Science*, *412*, 4247–4259.

[11] Bournez, O., Chassaing, P., Cohen, J., Gerin, J., & Koegler, X. (2009). On the convergence of population protocols when population goes to infinity. *Applied Mathematics and Computation*, *215*, 1340–1350.

[12] Cai, S., Izumi, T., & Wada, K. (2012). How to prove impossibility under global fairness: on space complexity of self-stabilizing leader election on a population protocol model. *Theory of Computing Systems*, *50*, 433–445.

[13] Canepa, D., & Gradinariu Potop-Butucaru, M. (2010). Self-stabilizing tiny interaction protocols. In *3rd International Workshop on Reliability, Availability, and Security, WRAS* (pp. 1–6).

[14] Casteigts, A., Flocchini, P., Quattrociocchi, W., & Santoro., N. (2012). Time-varying graphs and dynamic networks. *Int. Journal of Parallel, Emergent and Distributed Systems*, *27*, 387–408.

[15] Chatzigiannakis, I., Dolev, S., Fekete, S., Michail, O., & Spirakis, P. (2009). Not all fair probabilistic schedulers are equivalent. In *13th International Conference on Principles of Distributed Systems, OPODIS* (pp. 33–47).

[16] Chatzigiannakis, I., Michail, O., Nikolaou, S., Pavlogiannis, A., & Spirakis, P. (2010). All symmetric predicates in NSPACE(n^2) are stably computable by the mediated population protocol model. In *35th International Symposium on Mathematical Foundations of Computer Science, MFCS* (pp. 270–281).

[17] Chatzigiannakis, I., Michail, O., Nikolaou, S., Pavlogiannis, A., & Spirakis, P. (2011). Passively mobile communicating machines that use restricted space. *Theoretical Computer Science*, *412*, 6469–6483.

[18] Chatzigiannakis, I., Michail, O., & Spirakis, P. (2010). Stably decidable graph languages by mediated population protocols. In *12th International Symposium on Stabilization, Safety, and Security of Distributed Systems, SSS* (pp. 252–266).

[19] Chatzigiannakis, I., & Spirakis, P. (2008). The dynamics of probabilistic population protocols. In *22th International symposium on Distributed Computing, DISC* (pp. 498–499).

[20] Chen, H.-L., Cummings, R., Doty, D., & Soloveichik, D. (2014). Speed faults in computation by chemical reaction networks. In *28th International Symposium on Distributed Computing, DISC* (pp. 16–30).

[21] Das, S., Di Luna, G., Flocchini, P., Santoro, N., & Viglietta, G. (2017). Mediated population protocols: Leader election and applications. In *14th*

Annual Conference on Theory and Applications of Models of Computation, TAMC (pp. 172–186).

[22] Delporte-Gallet, C., Fauconnier, H., & Guerraoui, R. (2005). What dependability for networks of mobile sensors? In *1st Workshop on Hot Topics in System Dependability, HotDep* (p. 8).

[23] Delporte-Gallet, C., Fauconnier, H., Guerraoui, R., & Ruppert, E. (2006). When birds die: making population protocols fault-tolerant. In *2nd IEEE International Conference on Distributed Computing in Sensor Systems, DCOSS* (pp. 51–66).

[24] Di Luna, G., Flocchini, P., Izumi, T., Izumi, T., Santoro, N., & Viglietta, G. (2017). On the power of weaker pairwise interaction: Fault-tolerant simulation of population protocols. In *37th IEEE International Conference on Distributed Computing Systems, ICDCS* (pp. 2472–2477).

[25] Di Luna, G., Flocchini, P., Izumi, T., Izumi, T., Santoro, N., & Viglietta, G. (2017). Population protocols with faulty interactions: the impact of a leader. In *10th International Conference on Algorithms and Complexity, CIAC* (pp. 454–466).

[26] Fischer, M., & Jiang, H. (2006). Self-stabilizing leader election in networks of finite-state anonymous agents. In *10th International Conference on Principles of Distributed Systems, OPODIS* (pp. 395–409).

[27] Guerraoui, R., & Ruppert, E. (2007). *Even small birds are unique: population protocols with identifiers*. Technical Report CSE-2007-04, York University.

[28] Guerraoui, R., & Ruppert, E. (2009). Names trump malice: tiny mobile agents can tolerate byzantine failures. In *36th International Colloquium on Automata, Languages and Programming, ICALP* (pp. 484–495). volume 2.

[29] Michail, O., Chatzigiannakis, I., & Spirakis, P. (2011). Mediated population protocols. *Theoretical Computer Science*, *412*, 2434–2450.

[30] Michail, O., Chatzigiannakis, I., & Spirakis, P. (2011). *New models for population protocols*. Morgan & Claypool.