A Note on Set Agreement with Omission Failures*

Rachid Guerraoui, Petr Kouznetsov, Bastian Pochon
Distributed Programming Laboratory,
Swiss Federal Institute of Technology in Lausanne (EPFL)

Abstract

This paper considers the k-set agreement problem in a synchronous distributed system model with send-omission failures in which at most f processes can fail by send-omission. We show that, in a system of $n+1$ processes ($n+1 > f$), no algorithm can solve k-set agreement in $\lfloor \frac{f}{k} \rfloor$ rounds. Our lower bound proof uses topological techniques to characterize subsets of executions of our model. The characterization has a surprisingly regular structure which leads to a simple and succinct proof. We also show that the lower bound is tight by exhibiting a new algorithm that solves k-set agreement in $\lfloor \frac{f}{k} \rfloor + 1$ rounds.

Contact author Petr Kouznetsov
e-mail: petr.kouznetsov@epfl.ch

1 Introduction

Context. K-set agreement [4] is a generalization of the consensus [7] problem in which processes should decide on some final values based on their initial proposed values in such a way that: (1) the set of decided values contains at most k distinct values, (2) every decided value is a proposed value, and (3) every correct process eventually decides. The problem cannot be solved in a crash stop asynchronous model if the number f of processes that can crash is at least k [2]. This is a generalization of the FLP impossibility result [7] stating that consensus is not solvable if at least one process can crash: in this case, $k = 1$ and $f = 1$. It can be shown that in a synchronous model of $n+1$ processes, where up to f processes can crash, k-set agreement requires exactly $\lfloor f/k \rfloor + 1$ rounds if $\lfloor \frac{f}{k} \rfloor k \leq n-k$, and exactly $\lfloor f/k \rfloor$ rounds if $\lfloor \frac{f}{k} \rfloor k > n-k$: this is a simple generalization of [5], where only the case $f \leq n-k$ was considered.

Model. In this paper, we consider a synchronous message-passing model with send-omission failures (we will simply say an omission model). In this model, processes proceed in a round-by-round manner: in each round, every process sends a message to all, receives messages from other processes, and updates its

*This work is partially supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

local state. The only failures allowed in our model are send-omission failures (we will simply say omissions): a process can send a message that is never received.

Contributions. We show that in the omission model of $n+1$ processes, where at most f processes can fail by omission ($f \leq n$), k-set agreement cannot be solved in $\lfloor f/k \rfloor$ rounds. To prove this lower bound, we use the convenient notion of *pseudosphere* from [9] to describe the topological structure corresponding to a one-round execution of our model. In comparison with the proofs given in [5, 9] on k-set agreement lower bounds in a synchronous model with crash failures, our result is much easier to derive. This is due to the observation that the protocol complex corresponding to a bounded number of rounds of our omission model has a very regular structure: it is a complex homeomorphic to a union of n-*dimensional* pseudospheres. As a result, the connectivity of the complex giving the lower bound for k-set agreement can be easily computed. We also present a new algorithm that solves the problem in $\lfloor f/k \rfloor + 1$ rounds. Thus, for any $f < n+1$, k-set agreement requires exactly $\lfloor f/k \rfloor + 1$ rounds of a synchronous model of $n+1$ processes with at most f processes that can fail by omission.

Roadmap. Section 2 discusses the link between our result and known lower bounds on k-set agreement. Section 3 presents our model. Section 4 recalls some basic topological results used in this paper. Section 5 proves the lower bound. Section 6 proves that the lower bound is tight by giving an algorithm that matches it.

2 Related work

An execution of a synchronous model of $n+1$ processes with up to f crash failures can be viewed as an execution of our omission model: a crash failure is modeled in our case as a special case of an omission where, having failed by omission in a given round, a process fails by omitting all its messages in the subsequent rounds. In a synchronous model of $n+1$ processes, with up to f crash failures where $\lfloor \frac{f}{k} \rfloor k \leq n-k$, k-set agreement cannot be solved in $\lfloor f/k \rfloor$ rounds [5, 9]. Hence, no algorithm can solve k-set agreement in our omission model where up to f processes can fail by omission (we will call these processes *unreliable* in order to disambiguate with the notion of faulty process in the crash-prone model) and $\lfloor \frac{f}{k} \rfloor \leq n-k$, using less than $\lfloor \frac{f}{k} \rfloor + 1$ rounds: this would otherwise contradict the lower bound of [5].

However, in the case where $\lfloor f/k \rfloor k > n-k$, the lower bound of $\lfloor f/k \rfloor + 1$ rounds does not hold anymore for the synchronous model with crash failures: one can easily derive an algorithm that solves the problem in exactly $\lfloor f/k \rfloor$ rounds. We show in this paper that the lower bound for the omission model holds for all $f < n+1$ (not only for $\lfloor \frac{f}{k} \rfloor k \leq n-k$). Thus, for the case where $n-k < \lfloor f/k \rfloor k < n+1$, our result does not follow from [5].

With respect to k-set agreement, a round-based asynchronous model of $n+1$ processes with the strong failure detector $\mathcal{S}$ of [3] is equivalent to our omission model with $f = n$ in following sense: whenever k-set agreement is solvable in one model, it is also solvable in the other model. Thus, the lower bound of $n+1$ rounds for consensus holds for this model too Note that the lower bound for consensus in this model was obtained in [6] independently of our general proof for k-set agreement. On the other hand, in a synchronous model with at most n crash failures, consensus can be solved in n rounds. In this sense, our tight lower bound captures an interesting difference between the synchronous model with omission failures and the synchronous model with crash failures.

An alternative proof of our lower bound for the omission model was obtained in [8] by reduction of the first $\lfloor f/k \rfloor$ rounds of the model to the asynchronous *round-by-round failure detector* atomic snapshot shared memory model with at most k crash failures. The latter model is known to be too weak to solve k-set agreement [2] which implies that $\lfloor f/k \rfloor$ rounds of the omission model are not enough. The lower bound proof is based on two fundamental results in distributed computing: the impossibility of k-set agreement in the asynchronous model [2] and the atomic snapshot shared memory construction [1]. Neither of these is easy to derive. The proof we give in this paper is self-contained and is based on an interesting regularity of the omission model. Moreover, we show here that the lower bound is tight by presenting an optimal k-set agreement algorithm.

3 Model

The system we consider is a set of $n+1$ processes $\Pi = \{p_0, ..., p_n\} (n > 0)$. The processes evolve in synchronized rounds. In each round r, every process p_i executes the following steps: p_i sends a message to all other processes, receives a set of messages $M_{i,r}$ from other processes, and then updates its state. We assume that all protocols we consider are *full-information* protocols where, in each round, every process sends its local state to all processes. The only failures allowed are (send) *omission* failures: messages sent by a process to a subset of other processes can be lost. It is known that no deterministic algorithm can solve k-set agreement in a model with omissions where, in *every* round, *some* k processes can fail by omission [11, 13]. We assume here that at most f processes can fail by omission ($f < n+1$). As we pointed out in the introduction, we call these processes unreliable. By definition, in our model, every process is correct. Thus, when the k-set agreement problem is invoked in the omission model, every process, even an unreliable one, should eventually decide on some value according to the problem specification (recalled in the introduction).

4 Background

This section recalls some notions and results from basic algebraic topology (presented, for example in [12]) and some remarkable definitions and results from [9] that we use in this paper.

4.1 Simplexes and complexes

It is convenient to model a global state of a system of $n+1$ processes as an *n-dimensional simplex* $S^n = (s_0, ..., s_n)$, where $s_i = \langle p_i, v_i \rangle$ defines local state v_i of process p_i [10]. We say that the vertexes $s_0, ..., s_n$ *span* the simplex S^n. We say that a simplex T is a *face* of a simplex S if all vertexes of T are vertexes of S. A set of global states is modeled as a set of simplexes, closed under containment, called a *complex*.

4.2 Protocols

A *protocol* $\mathcal{P}$ is a subset of executions of our model. For any initial state represented as an n-simplex S, a *protocol complex* $\mathcal{P}(S)$ defines the set of final states reachable from them through the executions in $\mathcal{P}$. In other words, a set of vertexes $\langle p_{i_0}, v_{i_0}\rangle, ..., \langle p_{i_n}, v_{i_n}\rangle$ span a simplex in $\mathcal{P}(S)$ if and only if (1) S defines the initial state of $p_{i_0}, ..., p_{i_n}$, and (2) there is an execution in $\mathcal{P}$ in which $p_{i_0}, ..., p_{i_n}$ finish the protocol with states $v_{i_0}, ..., v_{i_n}$. For a set $\{S_i\}$ of possible initial states, $\mathcal{P}(\cup_i S_i)$ is defined as $\cup_i \mathcal{P}(S_i)$. If S^m is a face of S^n, then we define $\mathcal{P}(S^m)$ to be a subcomplex of $\mathcal{P}(S^m)$ corresponding to an execution in which only processes of S^m take steps and processes of $S^n \backslash S^m$ failed by omitting all their messages. For $m < n - f$, $\mathcal{P}(S^m) = \emptyset$, since in our model, there is no execution in which more than f processes fail by omissions.

For any two complexes $\mathcal{K}$ and $\mathcal{L}$, $\mathcal{P}(\mathcal{K} \cap \mathcal{L}) = \mathcal{P}(\mathcal{K}) \cap \mathcal{P}(\mathcal{L})$: any state of $\mathcal{P}(\mathcal{K} \cap \mathcal{L})$ belongs both to $\mathcal{P}(\mathcal{K})$ and $\mathcal{P}(\mathcal{L})$, any state from $\mathcal{P}(\mathcal{K}) \cap \mathcal{P}(\mathcal{L})$ defines the final states of processes originated from $\mathcal{K} \cap \mathcal{L}$ and, thus, belongs to $\mathcal{P}(\mathcal{K} \cap \mathcal{L})$.

We denote by $\mathcal{I}$ a complex corresponding to a set of possible initial configurations. Informally, a protocol $\mathcal{P}$ solves k-set agreement for $\mathcal{I}$ if there exists a map δ that carries each vertex of $\mathcal{P}(\mathcal{I})$ to a decision value in such a way that, for any $S^m = (\langle p_{i_0}, v_{i_0}\rangle, ..., \langle p_{i_m}, v_{i_m}\rangle) \in \mathcal{I}$ $(m \geq n - f)$, we have $\delta(\mathcal{P}(S^m)) \subseteq \{v_{i_0}, ..., v_{i_m}\}$ and $|\delta(\mathcal{P}(S^m))| \leq k$. (The formal definition of a *solvable task* is given in [10].)

Thus, in order to show that k-set agreement is not solvable in r rounds, it is sufficient to find an r-round protocol $\mathcal{P}$ that cannot solve the problem for some $\mathcal{I}$. Such a protocol can be interpreted as a set of worst-case executions in which no decision can be taken.

4.3 Connectivity

Informally, a complex is said to be k-connected if it has no holes in dimension k or less. More precisely:

Definition 1 *A complex $\mathcal{K}$ is* k-connected *if every continuous map of the k-sphere to $\mathcal{K}$ can be extended to a continuous map of the $(k+1)$-disk. By convention, a complex is (-1)-connected if it is non-empty, and every complex is k-connected for $k < -1$.*

We will also use the following corollary to the Mayer-Vietoris sequence [12] that helps define the connectivity of the result of $\mathcal{P}$ applied to a union of complexes:

Theorem 2 *If $\mathcal{K}$ and $\mathcal{L}$ are k-connected complexes, and $\mathcal{K} \cap \mathcal{L}$ is $(k-1)$-connected, then $\mathcal{K} \cup \mathcal{L}$ is k-connected.*

4.4 Pseudospheres

To prove our lower bound, we use the notion of *pseudosphere* introduced in [9] as a convenient abstraction to describe the topological structure corresponding to a bounded number of rounds of our model. To make the paper self-contained, we recall the definition of [9] here:

Definition 3 *Let $S^m = (s_0, ..., s_m)$ be a simplex and $U_0, ..., U_m$ be a sequence of finite sets. The* pseudosphere *$\psi(S^m; U_0, ..., U_m)$ is a complex defined as follows. Each vertex of $\psi(S^m; U_0, ..., U_m)$ is a pair $\langle s_i, u_i \rangle$, where s_i is a vertex of S^m and $u_i \in U_i$. Vertexes $\langle s_{i_0}, u_{i_0} \rangle, ..., \langle s_{i_l}, u_{i_l} \rangle$ define a simplex of $\psi(S^m; U_0, ..., U_m)$ if and only if all s_{i_j} $(0 \leq j \leq l)$ are distinct. If for all $0 \leq i \leq m$, $U_i = U$, the pseudosphere is written $\psi(S^m; U)$.*

The following properties of pseudospheres follow from their definition:

1. If $U_0, ..., U_m$ are singleton sets, then $\psi(S^m; U_0, ..., U_m) \cong S^m$.
2. $\psi(S^m; U_0, ..., U_m) \cap \psi(S^m; V_0, ..., V_m) \cong \psi(S^m; U_0 \cap V_0, ..., U_m \cap V_m)$.
3. If $U_i = \emptyset$, then $\psi(S^m; U_0, ..., U_m) \cong \psi(S^{m-1}; U_0, ..., \widehat{U_i}, ..., U_m)$, where circumflex means that U_i is omitted in the sequence $U_0, ..., U_m$.

4.5 Impossibility and connectivity

The following theorem, borrowed from [9], is based on Sperner's lemma [12]: it relates the connectivity of a protocol complex derived from a pseudosphere, with the impossibility of k-set agreement:

Theorem 4 *Let $\mathcal{P}$ be a protocol. If for every n-dimensional pseudosphere $\psi(p_0, ..., p_n; V)$, where V is non-empty, $\mathcal{P}(\psi(p_0, ..., p_n; V))$ is $(k-1)$-connected, and there are more than k possible input values, then $\mathcal{P}$ cannot solve k-set agreement.*

5 Lower bound

In this section we prove our lower bound by presenting a counter-example: a protocol $\mathcal{P}$, such that the corresponding complex satisfies the precondition of Theorem 4: for any pseudosphere $\psi(p_0, ..., p_n; V)$, where V is non-empty, $\mathcal{P}(\psi(p_0, ..., p_n; V))$ is $(k-1)$-connected. More precisely, we consider a set of executions in which, in every round, at most k processes are allowed to fail by omission. The corresponding protocol complex can be viewed as a union of n-dimensional pseudospheres which makes the reasoning about its connectivity very simple.

5.1 Connectivity theorem

The following generalization of Theorem 9 and Theorem 11 of [9] helps define the connectivity of a union of pseudospheres. The proof which basically reuses the arguments from [9] is given here to make the paper self-contained.

Theorem 5 *Let $\mathcal{P}$ be a protocol, S^m a simplex, and c a constant integer. Let for every face S^l of S^m, the protocol complex $\mathcal{P}(S^l)$ be $(l-c-1)$-connected. Then for every sequence of finite sets $\{A_{0_j}\}_{j=0}^m, ..., \{A_{l_j}\}_{j=0}^m$, such that for any*

$j \in [0, m]$, $\bigcap_{i=0}^{l} A_{i_j} \neq \emptyset$, *the protocol complex*

$$\mathcal{P}\left(\bigcup_{i=0}^{l} \psi(S^m; A_{i_0}, ..., A_{i_m})\right) \text{ is } (m-c-1)\text{-connected.} \qquad \text{(Eq. 1)}$$

Proof: Since for any sequence $V_0, ..., V_l$ of singleton sets, $\psi(S^l; V_0, ..., V_l) \cong S^l$, we notice that $\mathcal{P}(\psi(S^l; V_0, ..., V_l)) \cong \mathcal{P}(S^l)$ is $(l-c-1)$-connected.

(i) First, we prove that, for any m and any non-empty sets $U_0, ..., U_m$, the protocol complex $\mathcal{P}(\psi(S^m; U_0, ..., U_m))$ is $(m-c-1)$-connected. We introduce here the partial order on the sequences $U_0, ..., U_m$: $(V_0, ..., V_m) \prec (U_0, ..., U_m)$ if and only if each $V_i \subseteq U_i$ and for some j, $V_j \subset U_j$. We proceed by induction on m. For $m = c$ and any sequence $U_0, ..., U_m$, the protocol complex $\mathcal{P}(\psi(S^m; U_0, ..., U_m))$ is non-empty and, by definition, (-1)-connected.

Now assume that the claim holds for all simplexes of dimension less than m $(m > c)$. We proceed by induction on the partially-ordered sequences of sets $U_0, ..., U_m$. For the case where $(U_0, ..., U_m)$ are singletons, the claim follows from the theorem condition. Assume that the claim holds for all sequences smaller than $U_0, ..., U_m$ and there is an index i, such that $U_i = v \cup V_i$, where V_i is non-empty $(v \notin V_i)$. $\mathcal{P}(\psi(S^m; U_0, ..., U_m))$ is the union of $\mathcal{K} = \mathcal{P}(\psi(S^m; U_0, ..., V_i, ..., U_m))$ and $\mathcal{L} = \mathcal{P}(\psi(S^m; U_0, ..., \{v\}, ..., U_m))$ which are both $(m-c-1)$-connected by the induction hypothesis. The intersection is:

$$\begin{aligned} \mathcal{K} \cap \mathcal{L} &= \mathcal{P}(\psi(S^m; U_0, ..., V_i \cap \{v\}, ..., U_m)) = \\ &= \mathcal{P}(\psi(S^m; U_0, ..., \emptyset, ..., U_m)) \cong \\ &\cong \mathcal{P}(\psi(S^{m-1}; U_0, ..., \widehat{\emptyset}, ..., U_m)). \end{aligned}$$

The argument of $\mathcal{P}$ in the last expression represents an $(m-1)$-dimensional pseudosphere which is $(m-c-2)$-connected by the induction hypothesis. By Theorem 2, $\mathcal{K} \cup \mathcal{L} = \mathcal{P}(\psi(S^m; U_0, ..., U_m))$ is $(m-c-1)$-connected.

(ii) Now we prove our theorem by induction on l. We show that for any $l \geq 0$ and any sequence of sets $\{A_{i_j}\}$ satisfying the condition of the theorem, Equation 1 is guaranteed. The case $l = 0$ follows directly from (i). Now assume that, for some $l > 0$,

$$\mathcal{K} = \mathcal{P}\left(\bigcup_{i=0}^{l-1} \psi(S^m; A_{i_0}, ..., A_{i_m})\right) \text{ is } (m-c-1)\text{-connected.} \qquad \text{(Eq. 2)}$$

By (i), $\mathcal{L} = \mathcal{P}(\psi(S^m; A_{l_0}, ..., A_{l_m}))$ is $(m-c-1)$-connected. The intersection is

$$\begin{aligned} \mathcal{K} \cap \mathcal{L} &= \mathcal{P}\left((\bigcup_{i=0}^{l-1} \psi(S^m; A_{i_0}, ..., A_{i_m})) \cap \psi(S^m; A_{l_0}, ..., A_{l_m}) \right) = \\ &= \mathcal{P}\left(\bigcup_{i=0}^{l-1} \psi(S^m; A_{i_0} \cap A_{l_0}, ..., A_{i_m} \cap A_{l_m}) \right). \end{aligned}$$

By the initial assumption (Equation 2), for any j, $\bigcap_{i=0}^{l-1}(A_{i_j} \cap A_{l_j}) = \bigcap_{i=0}^{l} A_{i_j} \neq \emptyset$. Thus by the induction hypothesis,

$$\mathcal{K} \cap \mathcal{L} = \mathcal{P}\left(\bigcup_{i=0}^{l-1} \psi(S^m; A_{i_0} \cap A_{l_0}, ..., A_{i_m} \cap A_{l_m})\right) \text{ is } (m-c-1)\text{-connected.}$$

By Theorem 2, $\mathcal{K} \cup \mathcal{L}$ is $(m-c-1)$-connected.

□

Considering an identity protocol gives

Corollary 6 $\bigcup_{i=0}^{l} \psi(S^m; A_{i_0}, ..., A_{i_m})$ *is* $(m-1)$*-connected.*

5.2 One round

Now we define the protocol complex $\mathcal{R}^1(S^l)$ corresponding to one round of execution of our model, starting from an initial configuration S^l, in which up to k processes can fail by omission.[1]

Lemma 7 *Let* $S^l = (p_{i_0}, ..., p_{i_l})$ *be a simplex. If* $l \geq n-k$*, then*

$$\mathcal{R}^1(S^l) \cong \bigcup_{|K| \leq k} \psi(S^l; 2^{K-\{p_{i_0}\}}, ..., 2^{K-\{p_{i_l}\}}). \qquad \text{(Eq. 3)}$$

If $l < n-k$*, then* $\mathcal{R}^1(S^l)$ *is empty.*

Proof: Consider first the case $l \geq n-k$. Each vertex of $\mathcal{R}^1(S^l)$ has the form $\langle p_i, M_i \rangle$, where $p_i \in S^l$ and M_i is the set of messages received by p_i in the first round. Consider a particular set of executions in which exactly a subset $K \subset \Pi$ failed by omission in the first round. Each process p_i receives all messages from $\Pi \backslash K$ and a subset of messages from $K - \{p_i\}$ (p_i always knows its own message). Thus we can map in a one-to-one manner each vertex $\langle p_i, M_i \rangle$ of our protocol complex to a vertex labeled with a value from $2^K - \{p_i\}$. All combinations of the form $\langle p_i, u_i \rangle$, where $p_i \in S^l$ and $u_i \in 2^{K-\{p_i\}}$, give us a pseudosphere $\psi(S^l; 2^{K-\{p_{i_0}\}}, ..., 2^{K-\{p_{i_l}\}})$. The union over all sets K, such that $|K| \leq k$ gives the characterization of Equation 3.

The case $l < n-k$ is trivial: by the initial assumption, at most k processes can fail by omission. Thus no execution in which less then $n+1-k$ processes participate exists in the protocol complex. □

Example. Figure 1 depicts a protocol complex $\mathcal{R}^1(S^n)$, where $n = 2$, $f = 1$ and $k = 1$, corresponding to one round of the omission model of 3 processes of which at most one can fail by omission. Each vertex of the protocol complex corresponding to a reachable *local* state of a process is defined by the process id and the set of messages received by this process in the first round. Since at least two

[1]Naturally, we consider the case where $k \leq f$. Otherwise the protocol complex is trivial.

processes never fail by omission, each process receives at least two messages in each round. Moreover, in every simplex of the protocol complex corresponding to a reachable *global* state of the system, all sets of received messages include at least two common elements and every process is aware of its own message. Geometrically, the complex of Figure 1 consists of four pyramids starring from the vertexes $p : \{p, q, r\}$, $q : \{p, q, r\}$, $r : \{p, q, r\}$ with the base quadrangles corresponding to all possible executions where pairs of processes (q, r), (p, r) and (p, q) can miss the message of, respectively, p, q and r. These pyramids are homeomorphic to pseudospheres of the type $\psi(\{p, q, r\}; 2^{K-\{p\}}, 2^{K-\{q\}}, 2^{K-\{r\}})$, where K is, respectively, $\{p\}$, $\{q\}$ and $\{r\}$.

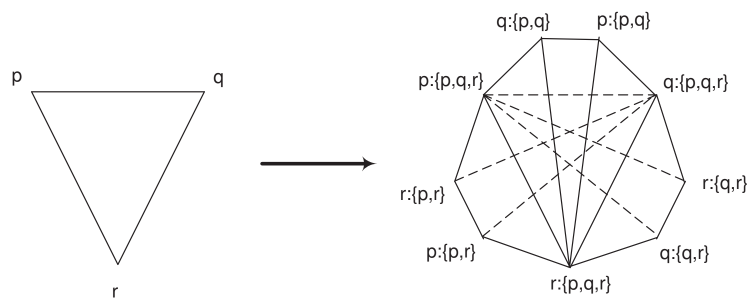

Figure 1: One-round protocol complex for three processes and one unreliable process.

By Lemma 7 and Corollary 6, $\mathcal{R}^1(S^l)$ is $(l-1)$-connected for all $l \geq n-k$. Since for all $l < n-k$, $\mathcal{R}^1(S^l)$ is (-2)-connected, we have:

Lemma 8 *For all l, $\mathcal{R}^1(S^l)$ is $(l-(n-k)-1)$-connected.*

5.3 Multiple rounds

Now we are ready to derive our main result.

Theorem 9 *If $rk \leq f$, then no protocol can solve k-set agreement in r rounds.*

Proof: We apply Theorem 4 by showing that, for any non-empty set V and $rk \leq f$, $\mathcal{R}^r(\psi(S^n; V))$ is $(k-1)$-connected. First, we prove that, for any m, $\mathcal{R}^r(S^m)$ is $(m-(n-k)-1)$-connected. Then we apply Theorem 5 showing that $\mathcal{R}^r(\psi(S^n; V))$ is $(k-1)$-connected.

We proceed by induction. The initial step ($r = 1$) trivially follows from Lemma 8. Now assume that, for all m, $\mathcal{R}^{r-1}(S^m)$ is $(m-(n-k)-1)$-connected under the condition $rk \leq f$. Thus,

$$\mathcal{R}^r(S^m) = \mathcal{R}^{r-1}(\mathcal{R}^1(S^m)) \cong \mathcal{R}^{r-1}\left(\bigcup_{|K|\leq k} \psi(S^m; 2^{K-\{p_{i_0}\}}, ..., 2^{K-\{p_{i_m}\}})\right).$$

Since, for any $j \in [0, n]$, $\bigcap_{|K|\leq k} 2^{K-\{p_{i_j}\}} = \{\emptyset\} \neq \emptyset$, by Theorem 5, $\mathcal{R}^r(S^m)$ is $(m-(n-k)-1)$-connected. $\square$

6 Algorithm

Figure 2 presents an algorithm that matches our lower bound of Theorem 9. The algorithm solves k-set agreement in our model and guarantees that, in every execution, every process decides in round $\lfloor f/k \rfloor + 1$. The algorithm can be viewed as a generalization of the consensus algorithm of [3] defined for the asynchronous model augmented with the strong failure detector $\mathcal{S}$.[2] The idea of our algorithm is the following:

1. Each process p_i sets its decision estimate est_i to its initial proposal v_i.
2. In each round r from 1 to $\lfloor \frac{f}{k} \rfloor + 1$, each process p_i sends its current decision estimate est_i to all and receives the set $M_{i,r}$ of estimates of other processes.
3. If a message (j, u) is received from some p_j, such that $(r-1)k \leq j \leq rk-1$, then p_i sets its estimate to u.
4. p_i decides its est_i after running $\lfloor \frac{f}{k} \rfloor + 1$ rounds.

$est_i := v_i$
for $r = 1..\lfloor \frac{f}{k} \rfloor + 1$ **do**
   send (i, est_i) to all processes
   receive $M_{i,r}$
   if $(\exists j)((r-1)k \leq j \leq rk-1) \wedge ((j, u) \in M_{i,r})$ **then**
      $est_i := u$
decide est_i

Figure 2: An algorithm for k-set agreement: process p_i.

Theorem 10 *The algorithm of Figure 2 solves k-set agreement in an omission model with $f < n+1$ unreliable processes.*

Proof: Every process decides after $\lfloor \frac{f}{k} \rfloor + 1$ rounds of computation. By the algorithm, the decided value is a proposed value of some process. Now we need to show that, in any execution, the set of decided values does not include more than k distinct values.

By the definition of the model, in every execution, there is a set E of at least $(n+1-f)$ processes, such that $\forall p_i, \forall r \geq 1, \forall p_j \in E : (p_j, u) \in M_{i,r}$. In every round r, process p_i expects to receive messages from a set of k processes p_j, such that $(r-1)k \leq j \leq rk-1$. Thus, in total per execution, p_i expects to receive messages from $\lfloor \frac{f}{k} \rfloor k + k$ distinct processes. Since $\lfloor \frac{f}{k} \rfloor k + k > f$ and there are at most f processes that can fail by omission, there exists a round r' in which every process receives a message from some $p_j \in E$, such that $(r'-1)k \leq j \leq r'k-1$. Thus, at round r', every process updates its estimate with some received value. Since no more than k estimates can be adopted in a round (only messages from k distinct processes are taken into account), k-set agreement is guaranteed. □

[2]The generalization is twofold: (1) we extend the algorithm from consensus to k-set agreement and (2) we make it f-resilient ($f \leq n$) instead of n-resilient (for a system of $n+1$ processes) as in [3].

7 Acknowledgements

Maurice Herlihy got us excited about the use of algebraic topology in distributed computing a year ago. Eli Gafni and Sergio Rajsbaum helped us control our excitement. Our colleagues from the Institute of Mathematics of EPFL, Sylvestre Blanc, Kathryn Hess, Paul-Eugene Parent and Orin Sauvageot, were instrumental in our understanding of simplexes and complexes.

References

[1] Y. Afek, H. Attiya, D. Dolev, E. Gafni, M. Merrit, and N. Shavit. Atomic snapshots of shared memory. *Journal of the ACM*, 40(4):873–890, September 1993.

[2] E. Borowsky and E. Gafni. Generalized FLP impossibility result for t-resilient asynchronous computations. In *Proceedings of the 25th ACM Symposium on Theory of Computing*, pages 91–100, May 1993.

[3] T. D. Chandra and S. Toueg. Unreliable failure detectors for reliable distributed systems. *Journal of the ACM*, 43(2):225–267, March 1996.

[4] S. Chaudhuri. Agreement is harder than consensus: set consensus in totally asynchronous systems. In *Proceedings of the 9th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 311–324, August 1990.

[5] S. Chaudhuri, M. Herlihy, N. A. Lynch, and M. R. Tuttle. Tight bounds for k-set agreement. *Journal of the ACM*, 47(5):912–943, 2000.

[6] C. Delporte-Gallet and H. Fauconnier. Latency measures and lower bounds for consensus with failure detectors. In *Proceedings of the 9th International Colloquium on Structural Information and Communication Complexity (SIROCCO 2002)*, Andros, 2002.

[7] M. J. Fischer, N. A. Lynch, and M. S. Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(3):374–382, April 1985.

[8] E. Gafni. A round-by-round failure detector - unifying synchrony and asynchrony. In *Proceedings of the 17th ACM Symposium on Principles of Distributed Computing (PODC)*, June 1998.

[9] M. Herlihy, S. Rajsbaum, and M. R. Tuttle. Unifying synchronous and asynchronous message-passing models. In *Proceedings of the 17th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 133–142, June 1998.

[10] M. Herlihy and N. Shavit. The topological structure of asynchronous computability. *Journal of the ACM*, 46(6):858–923, November 1999.

[11] I. Keidar and S. Rajsbaum. On the cost of fault-tolerant consensus when there are no faults - a tutorial. Technical report, MIT, May 2001.

[12] J. R. Munkres. *Elements of Algebraic Topology.* Addison-Wesley, Menlo Park, Calif., 1984.

[13] N. Santoro and P. Widmayer. Time is not a healer. In *Proceedings of the 6th Annual Symp. Theor. Aspects of Computer Science*, volume 349 of LNCS, pages 304–313, Padernorn, Germany, February 1989. Springer Verlag.