Unifying Synchronous and Asynchronous Message-Passing Models

Maurice Herlihy
Computer Science Department
Brown University
Providence, RI 02912
herlihy@cs.brown.edu

Sergio Rajsbaum
Instituto de Matemáticas
U.N.A.M., D.F. 04510, México
rajsbaum@servidor.unam.mx

Mark R. Tuttle
Digital Equipment Corporation
Cambridge Research Lab
One Kendall Square, Building 700
Cambridge, MA 02139
tuttle@crl.dec.com

March 25, 1998

Abstract

We take a significant step toward unifying the synchronous, semi-synchronous, and asynchronous message-passing models of distributed computation. The key idea is the concept of a *pseudosphere*, a new combinatorial structure in which each process from a set of processes is independently assigned a value from a set of values. Pseudospheres have a number of nice combinatorial properties, but their principal interest lies in the observation that the behavior of protocols in the three models can be characterized as simple unions of pseudospheres, where the exact structure of these unions is determined by the timing properties of the model. We use this pseudosphere construction to derive new and remarkably succinct proofs of bounds on consensus and k-set agreement in the asynchronous and synchronous models, as well as the first lower bound on wait-free k-set agreement in the semi-synchronous model.

To appear in the 16th Annual ACM Symposium on Principles of Distributed Computing (PODC98), Puerto Vallarta, Mexico, June 1998.

1 Introduction

The field of distributed computing embraces a bewildering variety of models [LL90, GY97]. A fundamental dimension along which these models differ is the degree to which process activity is synchronized. At one end of the spectrum is the *synchronous model* in which computation proceeds in a sequence of rounds. In each round, a process sends messages to the other processes, receives the messages sent to it by the other processes in that round, and changes state. All processes take steps at exactly the same rate, and all messages are delivered with exactly the same message delivery time. At the other end is the *asynchronous model* in which there is no bound on the amount of time that can elapse between process steps, and there is no bound on the time it can take for a message to be delivered. Between these extremes is the *semi-synchronous model* in which process step times and message delivery times can vary, but are bounded between constant upper and lower bounds. Proving a lower bound in any of these models requires a deep understanding of the global states that can arise in the course of a protocol's execution, and of how these global states are related.

The notion of *indistinguishability* or *similarity* [FLP85, HM90] has played a fundamental role in nearly every lower bound in distributed computation. Two global states are considered indistinguishable if one process has the same local state in both, and therefore cannot distinguish between them. The graph-theoretic representation of similarity, in which two global states are joined by an edge labeled with a process P if the global states are indistinguishable to P, has proven to be immensely powerful.

While the classical notion of similarity captures the notion of two states being indistinguishable to a single process, higher degrees of similarity have proved essential for understanding problems such as k-set agreement [Cha91] and renaming [ABND$^+$87, ABND$^+$90]. For example, it may matter that a pair of global states are indistinguishable to two processes, to three processes, and so on. To capture these higher degrees of similarity it is convenient to represent the global state of a system of $n+1$ processes with the n-dimensional analog of a triangle, called a *simplex*, where each vertex of the simplex representing a global state is labeled with the local states of processes in this global state. The set of all global states represented in this way forms a *simplicial complex*, sometimes called the *protocol complex*. The degree of similarity between two global states is represented geometrically by the number of vertices the corresponding simplexes have in common: two global states similar to one process share one vertex (namely that vertex labeled with the local state of this process having the same local state in both global states), states similar to two processes share two vertices, states similar to three processes share three vertices, and so on.

In this paper, we take a significant step toward unifying the synchronous, semi-synchronous, and asynchronous models of computation. The key, unifying idea in this paper is the concept of a *pseudosphere*, a simplicial complex in which each process from a set of processes is independently assigned a value from a set of values. Pseudospheres have a number of nice combinatorial properties,

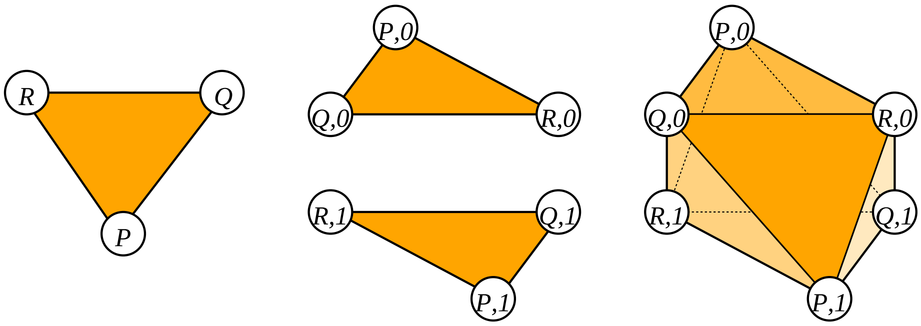

Figure 1: Construction of a three-process binary pseudosphere.

but their principal interest lies in the observation that protocol complexes in the three models can be characterized as simple unions of pseudospheres, where the exact structure of these unions is determined by the timing properties of the model. Because of the simple combinatorial properties of pseudospheres, reasoning about these unions can be accomplished by straightforward combinatorial arguments. We use this pseudosphere construction to derive new and remarkably succinct proofs of bounds on consensus [PSL80, Fis83] and k-set agreement [Cha91] in the asynchronous and synchronous models, as well as the first lower bound on wait-free k-set agreement in the semi-synchronous model.

A pseudosphere can be defined very simply. Start with an n-dimensional simplex where each vertex is labeled with a process id, and choose a finite set of values taken from an arbitrary domain. The pseudosphere is the complex constructed by taking multiple copies of this simplex and independently labeling each vertex with a value from the domain. For example, Figure 1 shows how to construct a pseudosphere by independently assigning binary values to a set of three processes. The left-hand figure shows a triangle labeled with process ids P, Q, and R. The central figure shows an intermediate stage where two copies of the triangle are each labeled with zeros and ones. The right-hand figure shows the complete construction, where copies of the triangle are labeled with all combinations of zeros and ones. We can just as easily assign values from a larger set than $\{0, 1\}$, although the result is harder to illustrate. We call this construct a pseudosphere because it is easily shown that the result of assigning binary values to $n + 1$ processes is topologically equivalent to an n-dimensional sphere.

The collection of initial global states for consensus or k-set agreement clearly forms a pseudosphere whose vertices are labeled with input values. For example, the right-hand figure in Figure 1 is the input complex for three-process consensus. The basic insight underlying the work presented in this paper is that protocol complexes in the models of interest have natural representations as unions of pseudospheres, except that the vertices are labeled with timing and failure information instead of input values. Reasoning about these protocol complexes reduces to the purely combinatorial problem of reasoning about unions of pseudospheres, and indeed the formal manipulations needed to derive our results are remarkably similar in all three models. In each model, we define a "round" structure appropriate for that model, and express the one-round executions as the union of pseudospheres. An r-round execution is constructed by inductively replacing each simplex in the single-round execution with the union of pseudospheres produced by the $(r-1)$-round protocol. The protocol complex produced by this iterative construction represents only a subset of the global states reachable in the model, but this set is large enough to prove the desired results for consensus, k-set agreement, renaming, and so on.

2 Related Work

Of course, we are not the first to propose iterative constructions representing the global states at the end of a protocol. As one example, Borowsky and Gafni [BG93] state a construction for the asynchronous model, and while the intuition behind the construction is compelling, it is not easy to write down a formal description of that construction. In a later paper [BG97], they define an iterated immediate snapshot model that has a recursive structure and has proven to be useful [HS97]. As another example, Chaudhuri, Herlihy, Lynch, and Tuttle [CHLT93] give an inductive construction for the synchronous model, and while the resulting "Bermuda Triangle" is visually appealing and an elegant combination of proof techniques from the literature, there is a fair amount of machinery needed in the formal description of the construction. In this sense, the formal presentation of our construction is substantially more succinct than these constructions.

We are also not the first to attempt to unify synchronous and asynchronous models of computation. All such attempts, including ours, restrict attention to a subset of well-behaved, round-based executions. There are a number of ways one could consider unifying models. One approach is to translate algorithms written for a synchronous model into an asynchronous model. This was the approach used by Awerbuch [Awe85] in the message-passing model when he constructed his *synchronizer* and showed how (in the absence of faults) synchronous protocols can be run in asynchronous systems in the presence of a synchronizer. This was also the approach used by Gafni and Yang [GY97] in the shared-memory model when they described a *round-by-round failure detector* and how these detectors can be used to run in an asynchronous model an algorithm written for a synchronous model. Another approach, which is basically our approach, is to identify a set of concepts that can be used to describe or reason about multiple models. This was the approach used by Moses and Rajsbaum [MR98] when they showed how the concepts of communication layering and mobile faults can be used to reason in a uniform way about the synchronous and asynchronous models. The translation approach assumes that the synchronous model is an easier model in which to work. This assumption seems justified for algorithm design, but lower bounds are typically easier to derive in the asynchronous model. Our pseudosphere construction illustrates this distinction in a striking way: in the asynchronous model, a single round of computation from a global state yields a single pseudosphere.

We are the first, however, to unify the synchronous, semi-synchronous, and asynchronous models of message-passing computation with a single concept, namely the pseudosphere. Gafni and Yang [GY97] do their most formal work in a synchronous and asynchronous shared-memory model, and while they sketch how their ideas might be extended to a semi-synchronous message-passing model, this extension requires changing the nature of the failure detector. Moses and Rajsbaum [MR98] focus on synchronous and asynchronous models. Their stated results apply to consensus whereas our results apply to both consensus and k-set agreement.

One indication that our construction is fundamental is that the pseudosphere constructions originally developed to unify the synchronous and asynchronous models extended cleanly to the semi-synchronous model. We consider this significant. Although variants of the semi-synchronous model have been around for a long time, we are aware of only one substantial lower bound in this model: the consensus bound of Attiya, Dwork, Lynch, and Stockmeyer [ADLS94]. The absence of other results suggests that it is very difficult to prove significant lower bounds in this model, and that results and proof techniques from other models do not translate into the semi-synchronous model as easily as one might hope. With our pseudosphere construction, however, we can prove the first lower bound for wait-free k-set agreement in this model. Another indication is that our construction can be used to simplify the proof of known results. For example, our protocol complex construction is significantly more succinct than the construction used by Herlihy and Shavit [HS93] in their asynchronous computability theorem, and our construction could be used to simplify the proof of one direction of that theorem. And, as mentioned, the formal analysis underlying our construction can be presented considerably more succinctly than the constructions used by Borowsky and Gafni [BG93] and by Chaudhuri, Herlihy, Lynch, and Tuttle [CHLT93].

Our constructions are guided by concepts and theorems taken from elementary combinatorial topology. As described above, we believe our results are interesting in their own right, even to readers unfamiliar with or uninterested in topological techniques. For readers interested in applications of topology to distributed computing, however, our constructions should be even more interesting. Our approach here replaces the existential arguments used by Herlihy and Shavit [HS93] to analyze protocol complexes in the asynchronous model with a constructive, inductive analysis. Although the existential analysis encompasses the entire complex, and ours is restricted to a well-structured subcomplex, we feel that the concise and constructive nature of our treatment makes a contribution, both in terms of simplicity and brevity, and in terms of intuitive appeal. Like us, Chaudhuri, Herlihy, Lynch, and Tuttle [CHLT93] explicitly construct a subset of the protocol complex in the synchronous model, although a smaller subset than ours. It is not clear, however, how to translate that construction into the asynchronous model. Moreover, since that construction, it has been a personal quest goal to find an inductive, round-by-round construction of the synchronous protocol complex.

3 Basic Topology

A *vertex* $\vec{v}$ is a point in a high-dimensional Euclidian space. Vertexes $\vec{v}_0, \ldots, \vec{v}_n$ are *affinely independent* if $\vec{v}_1 - \vec{v}_0, \ldots, \vec{v}_n - \vec{v}_0$ are linearly independent. An n*-dimensional simplex* (or n*-simplex*) $S^n = (\vec{s}_0, \ldots, \vec{s}_n)$ is the convex hull of a set of $n+1$ affinely-independent vertexes. For example, a 0-simplex is a vertex, a 1-simplex a line segment, a 2-simplex a solid triangle, and a 3-simplex a solid tetrahedron. Where convenient, we use superscripts to indicate dimensions of simplexes. We say that the $\vec{s}_0, \ldots, \vec{s}_n$ *span* S^n. By convention, a simplex of dimension $d < 0$ is an empty simplex. Simplex S^m is a (proper) *face* of T^n if the vertexes of S^m are a (proper) subset of the vertexes of T.

A *simplicial complex* (or complex) is a set of simplexes closed under containment and intersection. The *dimension* of a complex is the highest dimension of any of its simplexes. In this paper all the complexes of dimension n are *full* in the sense that every simplex is contained in some n-simplex. $\mathcal{L}$ is a *subcomplex* of $\mathcal{K}$ if every simplex of $\mathcal{L}$ is a simplex of $\mathcal{K}$. The m*-skeleton* of $\mathcal{K}$, denoted $skel^m(\mathcal{K})$, is the subcomplex consisting of all simplexes of $\mathcal{K}$ of dimension at most m. A map $\mu : \mathcal{K} \to \mathcal{L}$ carrying vertexes to vertexes is *simplicial* if it also carries simplexes to simplexes. Two complexes $\mathcal{K}$ and $\mathcal{L}$ are *isomorphic*, written $\mathcal{K} \cong \mathcal{L}$, if there is a surjective and one-to-one simplicial map $\iota : \mathcal{K} \to \mathcal{L}$.

Informally, a complex is k-connected if it has no holes in dimensions 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 [Spa66, p. 51]. (By convention, a complex is (-1)*-connected* if it is nonempty, and every complex is k*-connected* for $k < -1$.)

This definition says that a complex is 0-connected if it is connected in the graph-theoretic sense. The following theorem is an elementary consequence of the Mayer-Vietoris sequence [Mun84, p. 142]. It allows us to reason about a complex's connectivity in terms of the connectivity of its components.

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

4 Model

A set of $n+1$ sequential *processes* communicate by sending messages to one another. At any point, a process may *crash*: it stops and sends no more messages. There is a bound f on the number of processes that can fail. In this paper, we consider three distinct message-passing models. In the *asynchronous* model, there is no bound on process speed or message delivery time. In the *synchronous* model, processes take steps at the same rate, and messages take the same amount of time to be delivered. In the *semi-synchronous* model, the time between two consecutive steps of a process is at least c_1 and at most c_2, and the time to deliver a message is at most d, where c_1, c_2, and d are known constants. (The synchronous and asynchronous models are limiting cases of the semi-synchronous model.) In all three models, message delivery is reliable and FIFO.

Each process starts with an *input value* taken from a set V, and then executes a deterministic *protocol* in which it repeatedly receives one or more messages, changes its local state, and sends one or more messages. After a finite number of steps, each process chooses a *decision value* and halts. At any instant, a process's local state is given by its *view*: the input value and the the sequence of messages received so far. A protocol is uniquely determined by its *message function* and its *decision function*. The message function determines which messages a process should send in a given state, and the decision function determines which output value a process should choose in a given state (if any). A protocol is a *full-information protocol* [Had83, FL82, PSL80] if the message function causes each process to send its entire local state when it sends a message. We can assume without loss of generality that all protocols $\mathcal{P}$ we consider are *full-information* protocols [Had83, FL82, PSL80, DM90].

In the k-set agreement task [Cha91], processes are required to (1) choose a decision value after a finite number of steps, (2) choose as decision value some process's input value, and (3) collectively choose no more than k distinct decision values. When $k = 1$, this problem is usually called *consensus*.

We now show how to apply concepts from combinatorial topology to this model. An initial local state of process P is modeled as a vertex $\vec{v} = \langle P, v \rangle$ labeled with P's process id and initial value v. An initial global state is modeled as an n-simplex $S^n = (\langle P_0, v_0 \rangle, \ldots, \langle P_n, v_n \rangle)$, where the P_i are distinct. We use $\mathit{ids}(S^n)$ to denote the set of process ids associated with S^n, and $\mathit{vals}(S^n)$ the set of values. The set of all possible initial global states forms a complex, called the *input complex*.

Any protocol has an associated *protocol complex* $\mathcal{P}$, defined as follows. Each vertex is labeled with a process id and a possible view for that process. A set of vertexes $\langle P_{i_0}, v_{i_0} \rangle, \ldots, \langle P_{i_d}, v_{i_d} \rangle$ spans a simplex of $\mathcal{P}$ if and only if there is some protocol execution in which $P_{i_0}, \ldots, P_{i_d}$ finish the protocol with respective views $v_{i_0}, \ldots, v_{i_d}$. Each simplex thus corresponds to an equivalence class of executions that "look the same" to the processes at its vertexes. The protocol complex $\mathcal{P}$ depends both on the protocol and on the timing and failure characteristics of the model.

We use $\mathcal{P}(S^m)$ to denote the subcomplex of $\mathcal{P}$ corresponding to executions in which only the processes in $\mathit{ids}(S^m)$ participate (the rest fail before sending any messages). If $m < n - f$, then there are no such executions, and $\mathcal{P}(S^m)$ is empty. More generally, if $\mathcal{I}$ is a subcomplex of the input complex, then we define $\mathcal{P}(\mathcal{I})$ to be the union of $\mathcal{P}(S^m)$ for all S^m in $\mathcal{I}$. A protocol *solves* k-set agreement if the protocol's decision map δ carries vertexes of $\mathcal{P}$ to values in V such that if $\vec{p} \in \mathcal{P}(S^n)$, then $\delta(\vec{p}) \in \mathit{vals}(S^n)$.

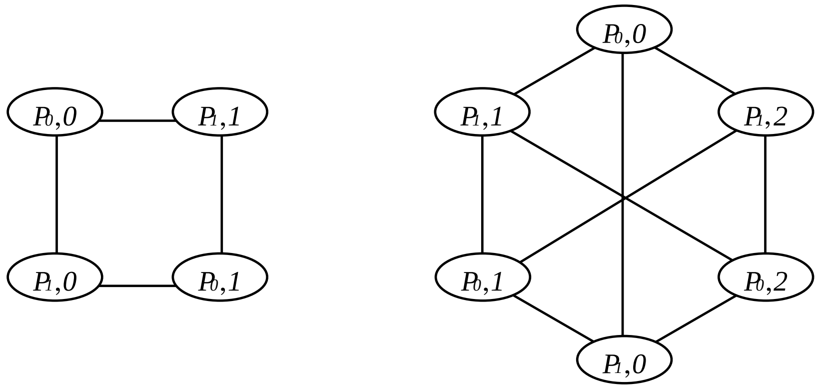

Figure 2: Pseudospheres $\psi(\{P_0, P_1\}; \{0,1\})$ and $\psi(\{P_0, P_1\}; \{0,1,2\})$.

5 Pseudospheres

Informally, a pseudosphere is a combinatorial structure in which each process from a set of processes is independently assigned a value from a set of values.

Definition 3: Let $S^m = (\vec{s}_0, \ldots, \vec{s}_m)$ be a simplex and $U_0, \ldots, U_m$ be a sequence of finite sets. The *pseudosphere* $\psi(S^m; U_0, \ldots, U_m)$ is the following complex. Each vertex is a pair $\langle \vec{s}_i, u_i \rangle$, where $\vec{s}_i$ is a vertex of S^m and $u_i \in U_i$. Vertexes $\langle \vec{s}_{i_0}, u_{i_0} \rangle, \ldots, \langle \vec{s}_{i_\ell}, u_{i_\ell} \rangle$ span a simplex of $\psi(S^m; U_0, \ldots, U_m)$ if and only if the $\vec{s}_i$ are distinct. A pseudosphere in which all U_i equal U is simply written $\psi(S^m; U)$.

We call this construct a pseudosphere because if S^n is an n-dimensional simplex, then $\psi(S^n; \{0,1\})$ is homeomorphic to an n-dimensional sphere. Pseudospheres are important because every complex considered here is either a pseudosphere or the union of pseudospheres. Because any process can start with any input from V, the input complex to k-set agreement is the pseudosphere $\psi(P^n; V)$, where P^n is a simplex whose vertexes are labeled with the $n+1$ distinct process ids.

Lemma 4: Pseudospheres satisfy the following simple combinatorial properties.

1. If U is a singleton set, then $\psi(S^m, U) \cong S^m$.

2. Let $S^m = (\vec{s}_0, \ldots, \vec{s}_m)$, and $S^{m-1} = (\vec{s}_0, \ldots, \widehat{\vec{s}_i}, \ldots \vec{s}_m)$, where circumflex denotes omission. If $U_i = \emptyset$, then

$$\psi(S^m; U_0, \ldots, U_m) \cong \psi(S^{m-1}; U_0, \ldots, \widehat{U}_i, \ldots, U_m).$$

3. $\psi(S_0; U_0, \ldots, U_m) \cap \psi(S_1; V_0, \ldots, V_m) \cong \psi(S_0 \cap S_1; U_0 \cap V_0, \ldots, U_m \cap V_m)$.

The next theorem shows how to exploit the nice combinatorial properties of pseudospheres. It states that if applying a protocol to a single simplex preserves connectivity below some dimension, then applying that protocol to any input pseudosphere also preserves that degree of connectivity. It is actually a theorem in topology, and so it applies to any model of computation.

Theorem 5: Let $\mathcal{P}$ be a protocol, S^m be a simplex, and c be a constant. If for every face S^ℓ of S^m and for every sequence $V_0, \ldots, V_\ell$ of singleton sets the protocol complex $\mathcal{P}(\psi(S^\ell; V_0, \ldots, V_\ell))$ is $(\ell - c - 1)$-connected, then for every sequence $U_0, \ldots, U_m$ of nonempty sets the protocol complex $\mathcal{P}(\psi(S^m; U_0, \ldots, U_m))$ is $(m - c - 1)$-connected.

Proof: Define the following partial order on sequences of sets: $(U_0, \ldots, U_m) \prec (V_0, \ldots, V_m)$ if each $U_i \subseteq V_i$, and for at least one set, the inclusion is strict.

We proceed by induction on m. For the base case, let $m = c$. Since $\mathcal{P}(\psi(S^m; U_0, \ldots, U_m))$ is the union of the $\mathcal{P}(\psi(S^m; V_0, \ldots, V_m))$ as the singletons V_i range over the values in U_i, and since each $\mathcal{P}(\psi(S^m; V_0, \ldots, V_m))$ is (-1)-connected (meaning nonempty), so is the union $\mathcal{P}(\psi(S^m; U_0, \ldots, U_m))$.

For the induction step for m, assume the hypothesis holds for simplexes of dimension less than m. We proceed by induction on the partially-ordered sequences $U_0, \ldots, U_m$. For the base case, when all the U_i are singletons, the claim is just the hypothesis.

For the induction step for the set sequences, assume the claim for every sequence less than $U_0, \ldots, U_m$. There must be some index i such that $U_i = V_i \cup \{v\}$, where V_i is nonempty. The pseudosphere is the union of

$$\mathcal{K} = \mathcal{P}(\psi(S^m; U_0, \ldots, V_i, \ldots, U_m))$$

and

$$\mathcal{L} = \mathcal{P}(\psi(S^m; U_0, \ldots, \{v\}, \ldots, U_m)).$$

By the induction hypothesis for the sets, both $\mathcal{K}$ and $\mathcal{L}$ are $(m-c-1)$-connected. Moreover,

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

This last expression is $\mathcal{P}$ applied to an $(m-1)$-dimensional pseudosphere, which is $(m - c - 2)$-connected by the induction hypothesis for m. Because the U_j are nonempty for $i \neq j$, this pseudosphere is nonempty. We have shown that $\mathcal{K}$ and $\mathcal{L}$ are $(m - c - 1)$-connected, and their intersection is nonempty and $(m - c - 2)$-connected, so their union is $(m - c - 1)$-connected by Theorem 2. $\square$

A consequence of this theorem is that any n-dimensional pseudosphere is $(n-1)$-connected (just let $\mathcal{P}$ be the trivial protocol in which each process halts immediately):

Corollary 6: If $U_0, \ldots, U_m$ are all nonempty, then $\psi(S^m; U_0, \ldots, U_m)$ is $(m-1)$-connected.

Naively, one might think that S^m is always m-connected, but note that although the empty simplex has dimension -1, it is not (-1)-connected. We can generalize Theorem 5 to multiple pseudospheres.

Theorem 7: Let $\mathcal{P}$ be a protocol satisfying the precondition of Theorem 5, and let $A_0, \ldots, A_\ell$ be a sequence of finite sets. If $\cap_{i=0}^{\ell} A_i \neq \emptyset$ then

$$\mathcal{P}\left(\bigcup_{i=0}^{\ell} \psi(S^m; A_i)\right) \text{ is } (m-c-1)\text{-connected.}$$

Proof: We argue by induction on ℓ. The base case, when $\ell = 0$, follows directory from Theorem 5.

For the induction step, assume that

$$\mathcal{K} = \mathcal{P}\left(\bigcup_{i=0}^{\ell-1} \psi(S^m; A_i)\right) \text{ is } (m-c-1)\text{-connected.}$$

By Theorem 5,

$$\mathcal{L} = \mathcal{P}(\psi(S^m; A_\ell)) \text{ is } (m-c-1)\text{-connected.}$$

The intersection is

$$\mathcal{K} \cap \mathcal{L} = \mathcal{P}\left(\left(\bigcup_{i=0}^{\ell-1} \psi(S^m; A_i)\right) \cap \psi(S^m; A_\ell)\right) = \mathcal{P}\left(\bigcup_{i=0}^{\ell-1} \psi(S^m; A_i \cap A_\ell)\right).$$

Define

$$\begin{aligned} B_i &= A_i \cap A_\ell \text{ for } 0 \leq i \leq \ell - 1 \\ \mathcal{B} &= \bigcup_{i=0}^{\ell-1} \psi(S^m; B_i) \end{aligned}$$

$\mathcal{B}$ is the union of $\ell - 1$ pseudospheres, and $\cap_{i=0}^{\ell-1} B_i \neq \emptyset$, so by the induction hypothesis,

$$\mathcal{K} \cap \mathcal{L} = \mathcal{P}(\mathcal{B}) = \mathcal{P}\left(\bigcup_{i=0}^{\ell-1} \psi(S^m; B_i)\right) \text{ is } (m-c-1)\text{-connected.}$$

The claim now follows from Theorem 2. $\square$

Letting $\mathcal{P}$ be the trivial protocol in which each process decides its input:

Corollary 8: $\mathcal{A}$ is $(m-1)$-connected.

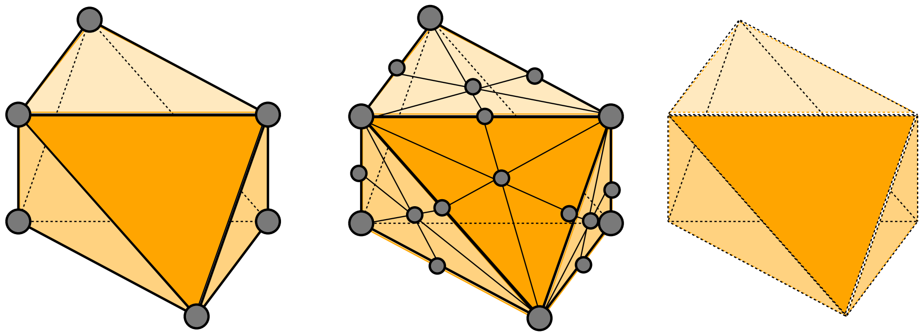

Figure 3: Simplicial complex, subdivision, and polyhedron

6 Connectivity vs k-Set Agreement

The notion of k-connectivity lies at the heart of all known lower bounds for k-set agreement. In this section, we prove a general theorem linking $(k-1)$-connectivity with impossibility of k-set agreement. This theorem is model independent in the sense that it depends on the connectivity properties of protocol complexes, not on explicit timing or failure properties of the model. This result was originally stated elsewhere [HR94], but for the sake of making this paper self-contained, we present the full proof here.

The point-set occupied by a complex $\mathcal{C}$ is called its *polyhedron*, and is denoted by $|\mathcal{C}|$. Any simplicial map $\phi : \mathcal{A} \to \mathcal{B}$ induces a piece-wise linear map $|\phi| : |\mathcal{A}| \to |\mathcal{B}|$ that agrees with ϕ on vertexes of $\mathcal{A}$.

A *subdivision* of a complex $\mathcal{A}$ is a complex $\mathcal{B}$ such that (1) each simplex of $\mathcal{B}$ is contained in a simplex of $\mathcal{A}$, and (2) each simplex of $\mathcal{A}$ is the union of finitely many simplexes of $\mathcal{B}$ [Mun84, p. 83]. This definition implies that $|\mathcal{A}| = |\mathcal{B}|$. If $\vec{b}$ is a vertex of $\mathcal{B}$, the *carrier* of $\vec{b}$ in $\mathcal{A}$, denoted $carrier(\vec{b}, \mathcal{A})$, is the smallest simplex of $\mathcal{A}$ that contains $\vec{b}$. Figure 3 illustrates a complex, a subdivision of that complex, and their underlying polyhedron.

We will need a step-by-step method for constructing subdivisions. Let $\mathcal{C}$ be a complex, and $\vec{w}$ a point with the property that any ray emanating from $\vec{w}$ intersects $|\mathcal{C}|$ in at most one point. Define the *cone* $\vec{w} \cdot \mathcal{C}$ to be the collection of all simplexes of the form $(\vec{w}, \vec{s}_0, \ldots, \vec{s}_k)$, where $(\vec{s}_0, \ldots, \vec{s}_k)$ is a simplex of $\mathcal{C}$, together with all faces of such simplexes. This cone is itself a complex, having $\mathcal{C}$ as a subcomplex [Mun84, p. 44]. Let σ be a subdivision of $skel^{\ell-1}(\mathcal{C})$, and $S_0^\ell, \ldots, S_L^\ell$ the ℓ-simplexes of $skel^\ell(\mathcal{C})$. For $0 \le i \le L$, let $\vec{w}_i$ be an interior point of $|S_i^\ell|$. Each cone $\vec{w}_i \cdot \sigma(S_i^\ell)$ is a subdivision of S_i^ℓ, and the union of these cones as i ranges from 0 to L is a subdivision of $skel^\ell(\mathcal{C})$ that agrees with σ on the $(\ell-1)$ skeleton [Mun84, p. 85]. The result is called the subdivision of $skel^\ell(\mathcal{C})$ obtained by *starring* σ. The subdivision shown in Figure 3 is the result of successive starring.

We use the following variant of Sperner's Lemma [Lef49, Lemma 5.5]:

Lemma 9 (Sperner's Lemma): Let $\sigma(S^n)$ be a subdivision of simplex S^n. If $F : \sigma(S^n) \rightarrow S^n$ is a map sending each vertex of $\sigma(S^n)$ to a vertex in its carrier, then there is at least one n-simplex $T^n = (\vec{t}_0, \ldots, \vec{t}_n)$ in $\sigma(S^n)$ such that the $F(\vec{t}_i)$ are all distinct.

We also exploit the following extension lemma, which appears in Glaser [Gla70, Theorem IV.2].

Lemma 10: Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be complexes such that $\mathcal{A} \subset \mathcal{B}$, and $f : |\mathcal{B}| \rightarrow |\mathcal{C}|$ is a continuous map such that f restricted to $|\mathcal{A}|$ is simplicial. There exists a subdivision τ of $\mathcal{B}$ such that $\tau(\mathcal{A}) = \mathcal{A}$, and a simplicial map $\phi : \tau(\mathcal{B}) \rightarrow \mathcal{C}$ extending the restriction of f to $|\mathcal{A}|$.

Theorem 11: Let $V = \{v_0, \ldots, v_k\}$ be a set of $k+1$ possible input values, and $\mathcal{P}$ a protocol with input complex $\psi(P_0, \ldots, P_n; V)$. If $\mathcal{P}$ has the property that for every n-dimensional pseudosphere $\psi(P_0, \ldots, P_n; U)$, where U is a nonempty subset of V, $\mathcal{P}(\psi(P_0, \ldots, P_n; U)$ is $(k-1)$-connected, then $\mathcal{P}$ cannot solve k-set agreement.

Proof: Let $S^k = (\vec{s}_0, \ldots, \vec{s}_k)$ be a simplex in which each vertex $\vec{s}_i$ is labeled with input value v_i. For any face S^j of S^k, let $\mathit{vals}(S^j)$ denote the set of values labeling S^j. We will construct a subdivision σ of S^k and a simplicial map

$$\phi : \sigma(S^k) \rightarrow \mathcal{P}(\psi(P_0, \ldots, P_n; V))$$

with the property that

$$\phi(\vec{s}) \in \mathcal{P}(\psi(P_0, \ldots, P_n; \mathit{vals}(\mathit{carrier}(\vec{s}, S^k)))). \tag{1}$$

Before undertaking this construction, we observe why it suffices to prove the theorem. We argue by contradiction. Recall that the protocol's decision map $\delta : D\,\mathcal{P} \rightarrow V$ carries a process's final local state to the value that process would choose in that local state. Each n-simplex of $\mathcal{P}(\psi(P_0, \ldots, P_n; \mathit{vals}(\mathit{carrier}(\vec{s}, S^k))))$ represents a protocol execution in which processes started with input values from $\mathit{vals}(\mathit{carrier}(\vec{s}, S^k))$. Because k-set agreement requires each process to choose some process's input, $\delta(\phi(\vec{s})) \in \mathit{vals}(\mathit{carrier}(\vec{s}, S^k))$ for every $\vec{s} \in \sigma(S^k)$. Define the map γ to send each $\vec{s} \in \sigma(S^k)$ to the unique vertex of S^k labeled with $\delta(\phi(\vec{s}))$. The map γ sends each vertex of $\sigma(S^k)$ to a vertex in its carrier, so by Sperner's Lemma there is a simplex $T^k = (\vec{t}_0, \ldots, \vec{t}_k)$ in $\sigma(S^k)$ such that $\gamma(\vec{t}_0), \ldots, \gamma(\vec{t}_k)$ are distinct. It follows that $\phi(\vec{t}_0), \ldots, \phi(\vec{t}_k)$ is a k-simplex of of $\mathcal{P}$ whose vertexes are mapped by δ to distinct values. This simplex is the face of an n-simplex representing an execution in which $k+1$ processes choose distinct values, a contradiction.

We now turn our attention to the construction of σ and ϕ. For $0 \leq \ell \leq k$, we will inductively construct a subdivision σ_ℓ and simplicial map

$$\phi_\ell : \sigma_\ell(\mathit{skel}^\ell(S^k)) \rightarrow \mathcal{P}$$

such that for every vertex $\vec{s} \in \sigma_\ell(skel^\ell(S^k))$,

$$\phi_\ell(\vec{s}) \in \mathcal{P}(\psi(P_0, \ldots, P_n; vals(carrier(\vec{s}, S^k)))). \qquad (2)$$

For the base step, let ϕ_0 carry each vertex $\vec{s}_i$ of $skel^0(S^k)$ to any vertex of $\mathcal{P}(\psi(P_0, \ldots, P_n; \{v_i\}))$, the protocol complex of executions in which every process starts with the same input v_i. It is immediate that ϕ_0 satisfies Equation 2.

For the induction step, we exploit connectivity to extend $\sigma_{\ell-1}$ and $\phi_{\ell-1}$ to dimension ℓ. Let τ be the subdivision of $skel^\ell(S^k)$ obtained by starring $\sigma_{\ell-1}(skel^{\ell-1}(S^k))$. Let $S_0^\ell, \ldots, S_L^\ell$ be all the ℓ-simplexes in $skel^\ell(S^k)$. For each S_i^ℓ, $0 \leq i \leq L$, let $\mathcal{A}_i$ be the complex $\tau(skel^{\ell-1}(S_i^\ell))$, and $\mathcal{B}_i$ the complex $\tau(S_i^\ell)$ Equation 2 implies that ϕ_ℓ sends $\mathcal{A}$ into $\mathcal{P}(\psi(P_0, \ldots, P_n; vals(S_i^\ell)))$. Let

$$g_i : |\mathcal{A}| \to |\mathcal{P}(\psi(P_0, \ldots, P_n; vals(S_i^\ell)))|$$

be the piece-wise linear map induced on the polyhedra. $\mathcal{A}$ is homeomorphic to an $(\ell - 1)$-sphere, and g_i is continuous (being piece-wise linear). Because $\mathcal{P}(\psi(P_0, \ldots, P_n; vals(S_i^\ell)))$ is $(k-1)$-connected, where $k \geq \ell$, g_i can be extended to a continuous

$$f_i : |\mathcal{B}_i| \to |\mathcal{P}(\psi(P_0, \ldots, P_n; vals(S_i^\ell)))|$$

By Lemma 10, there exists a subdivision ρ_i of $\mathcal{B}_i$ and a simplicial map

$$\mu_i : \rho_i(\mathcal{B}_i) \;\to\; \mathcal{P}(\psi(P_0, \ldots, P_n; vals(S_i^\ell)))$$

extending the restriction of $\phi_{\ell-1}$ to $|\mathcal{A}|$.

By construction, the ρ_i subdivisions agree with $\sigma_{\ell-1}$ on $skel^{\ell-1}(S^k)$, so together they define a subdivision ρ of $\tau(skel^\ell(S^k))$. Let $\sigma_\ell(skel^\ell(S^k)) = \rho(\tau(skel^\ell(S^k)))$. The maps μ_i also agree on $\sigma_{\ell-1}(skel^{\ell-1}(S^k))$, and together they define a simplicial map

$$\phi_\ell : \sigma_\ell(skel^\ell(S^k)) \to \mathcal{P}$$

satisfying Equation 2.

Finally, note that $skel^k(S^k) = S^k$, so the claim follows by letting $\sigma = \sigma_k$ and $\phi = \phi_k$. □

Theorems 5 and 11 imply

Corollary 12: If $\mathcal{P}(\mathcal{S}^m)$ is $(m - (n - k) - 1)$-connected for all m where with $n - f \leq m \leq n$, then $\mathcal{P}$ cannot solve k-set agreement in the presence of f failures.

7 Asynchronous Computation

In this section, we define the r-round asynchronous protocol complex $\mathcal{A}^r(S^m)$, and we prove that it is highly-connected, implying that it impossible to solve problems like k-set agreement in the asynchronous model.

$active_r$ is the set of active processors at the end of round r.
$delivered_{p,r}$ is the set of processors whose round r messages to p are delivered in round r.

choose $active_0 \subseteq P$ with $|active_0| \geq n - f + 1$
for each round $r = 1, \ldots, k$
  for each processor $p \in active_{r-1}$
    choose $delivered_{p,r} \subseteq active_{r-1} - \{p\}$ with $|delivered_{p,r}| \geq n - f$
    deliver all previously undelivered messages sent from processors in $delivered_{p,r} \cup \{p\}$ to p in rounds 1 through r
  choose $active_r \subseteq active_{r-1}$ with $|active_r| \geq n - f + 1$

Figure 4: Constructing the iterated executions.

The global states used in the construction of this complex are states that arise during a set of well-behaved, round-based executions of the full-information protocol defined informally as follows. In each round, each processor sends its state to every other processor, receives the messages delivered to it during that round, and undergoes a state transition. Because the model is asynchronous, a message m sent from p to q in round r may not be delivered in that round. When m is delivered, however, all previously undelivered messages sent from p to q in rounds 1 through r are delivered at the same time. This means that messages are delivered in FIFO order. In each round, each processor receives at least $n - f + 1$ messages sent during that round, including its message to itself. This is the greatest number of messages a processor can count on receiving when up to f processors can fail. At this level of detail, this set of executions looks something like a message-passing analog of the executions arising in the iterated immediate snapshot model defined by Borowsky and Gafni [BG97] for shared memory.

The formal definition of the states we use to construct $\mathcal{A}^r(S^m)$ is most easily given as the set of states that occur during iterated executions. An iterated execution is not necessarily an asynchronous execution, since a message sent during an iterated execution need not be delivered. Nonetheless, since every prefix of an iterated execution is a prefix of an asynchronous execution (one in which all undelivered messages are delivered immediately after the prefix), it follows that all global states appearing in iterated executions are global states that appear in asynchronous executions, and this is all that we need.

The class of *iterated executions* of the full-information protocol is defined formally as follows (see Figure 4). No processor fails during an iterated execution, so every processor broadcasts its local state to every processor in every round. A processor can become inactive, however, after which its messages fail to be delivered. For each round r, we choose a set $active_r$ of processors that are still active at the end of round r. These sets satisfy the conditions $active_r \subseteq active_{r-1}$ and $|active_r| \geq n - f + 1$ stating that inactive processors remain inactive that at most f processors become inactive. The initial condition $|active_0| \geq n - f + 1$ allows some processors to become inactive before sending a single message. For each round r and active processor $p \in active_{r-1}$, we choose a set $delivered_{p,r} \subseteq active_{r-1}$ of processors q whose round r messages are delivered to p during round r, along with p's message to itself, and along with any other previously undelivered messages from q to p. These sets satisfy the condition $|delivered_{p,r} \cup \{p\}| \geq n - f + 1$ stating that at most f round r messages fail to be delivered to p. An iterated execution is uniquely determined by the choice of these sets $active_r$ and $delivered_{p,r}$. The next lemma states that the states appearing in iterated executions are states appearing in asynchronous executions, so we can restrict our attention to iterated executions when constructing the protocol complex $\mathcal{A}^r(S^m)$.

Lemma 13: Let e be an iterated execution. For every round r, there is an asynchronous execution e_r in which no processors fail and in which the global states at time r in e and e_r is the same.

Proof: Let e_r be the execution obtained from e by causing all currently undelivered messages to be delivered at the end of round r' for all rounds $r' > r$, so that there are no undelivered messages at the end of any round r' in e_r beginning with round $r + 1$. $\square$

Let $\mathcal{A}^1(S^m)$ be the protocol complex of all one-round, iterated executions of the full information protocol with input simplex S^m. In these executions, the set of initially active processors is taken to be the set $active_0 = ids(S^m)$ of processor ids labeling S^m. Since $m \leq n$, the $n - m$ missing processors are inactive from the start of the execution, leaving room for only $f - (n - m)$ additional inactive processors during the remainder of the execution. Let 2^U_k denote the set of subsets of U of size at least k. Our first result says that the one-round protocol complex is a single pseudosphere:

Lemma 14: If $m \geq n - f$ and $Q = \{q_0, \ldots, q_m\}$ is the set of processor ids labeling the input simplex S^m, then

$$\mathcal{A}^1(S^m) \cong \psi(S^m; 2^{Q-\{q_0\}}_{n-f}, \ldots, 2^{Q-\{q_m\}}_{n-f}).$$

Proof: $\mathcal{A}^1(S^m)$ is the protocol complex of all one-round, iterated executions of the full-information protocol with $active_0 = ids(S^m) = Q$. Each vertex v in this complex is uniquely determined by a processor id $q \in Q$ and a set $delivered_{q,1}$ of processors (not including q) whose round 1 messages are delivered to q in round 1. These sets $delivered_{q,1}$ are subsets of $Q - \{q\}$ of size at least $n - f$ (since every processor receives a message from itself and from at least $n - f$ other processors). These sets $delivered_{q,1}$ are chosen independently for each processor (since each processor can hear from an independently chosen set of at least $n-f$ other processors). These sets therefore induce a pseudosphere on S^m. Define the vertex map

$$\iota : \mathcal{A}^1(S^m) \to \psi(S^m; 2^{Q-\{q_0\}}_{n-f}, \ldots, 2^{Q-\{q_m\}}_{n-f})$$

by $\iota(q_i, delivered_{q_i,1}) = \langle \vec{s}_i, delivered_{q_i,1} \rangle$. It is easy to see that ι is simplicial (sending simplexes to simplexes) and one-to-one. $\square$

The complex $\mathcal{A}^1(S^m)$ is highly-connected:

Lemma 15: $\mathcal{A}^1(S^m)$ is $(m-(n-f)-1)$-connected for all $m \geq 0$.

Proof: Suppose $0 \leq m < n-f$. In this case, $m-(n-f)-1 < -1$, and by convention every complex is k-connected when $k < -1$.

Suppose $n-f \leq m \leq n$. Then $\mathcal{A}^1(S^m)$ is an m-dimensional pseudosphere by Lemma 14 which is $(m-1)$-connected by Corollary 6. The claim follows since $m-1 \geq m-(n-f)-1$. $\square$

Let $\mathcal{A}^r(S^m)$ be the r-round protocol complex defined by induction to be the result of taking the union of $\mathcal{A}^{r-1}(T)$ for every simplex T in $A^1(S^m)$ with dimension at least $n-f$. The complex $\mathcal{A}^r(S^m)$ is also highly-connected:

Lemma 16: $\mathcal{A}^r(S^m)$ is $(m-(n-f)-1)$-connected for all $r \geq 1$ and all $m \geq 0$.

Proof: We proceed by induction on $r \geq 1$ to prove that $\mathcal{A}^r(S^m)$ is $(m-(n-f)-1)$-connected for all $m \geq 0$. If $r=1$, then Lemma 15 says that $\mathcal{A}^1(S^m)$ is $(m-(n-f)-1)$-connected, so suppose $r > 1$ and the induction hypothesis holds for $r-1$.

We proceed by induction on $m \geq 0$ to prove that $\mathcal{A}^r(S^m)$ is $(m-(n-f)-1)$-connected.

Suppose $0 \leq m < n-f$. In this case, $m-(n-f)-1 < -1$, and by convention every complex is k-connected when $k < -1$.

Suppose $n-f \leq m \leq n$ and the induction hypothesis holds for $m-1$. The induction hypothesis for $r-1$ implies that $\mathcal{A}^{r-1}(S^\ell)$ is $(\ell-(n-f)-1)$-connected for every proper face S^ℓ of S^m. It follows by Theorem 5 and Lemma 4 that $\mathcal{A}^{r-1}(\psi(S^m; U_0, \ldots, U_\ell))$ is $(m-(n-f)-1)$-connected for every sequence $U_0, \ldots, U_m$ of nonempty sets. Lemma 14 implies that $\mathcal{A}^1(S^m)$ is a pseudosphere of the form $\psi(S^m; U_0, \ldots, U_m)$ for some sequence $U_0, \ldots, U_m$ of sets. It follows that $\mathcal{A}^r(S^m) = \mathcal{A}^{r-1}(\mathcal{A}^1(S^m))$ is $(m-(n-f)-1)$-connected. $\square$

It follows that the asynchronous protocol complex is $(f-1)$-connected (when $m=n$), and thus we can prove the impossibility of asynchronous k-set agreement [BG93, HS93, SZ93]:

Theorem 17: If I the set of input values, and $\mathcal{I} = \psi(S^n, I)$ the input complex, then $\mathcal{A}^r(\mathcal{I})$ is $(f-1)$-connected for any $r > 0$.

Proof: For each input simplex S^m, $\mathcal{A}^r(S^m)$ is $(m-(n-f)-1)$-connected (Lemma 16), so $\mathcal{A}^r(\psi(S^m; I))$ is also $(m-(n-f)-1)$-connected by Theorem 5. $\square$

Corollary 18: There is no asynchronous f-resilient k-set agreement protocol for $k \leq f$.

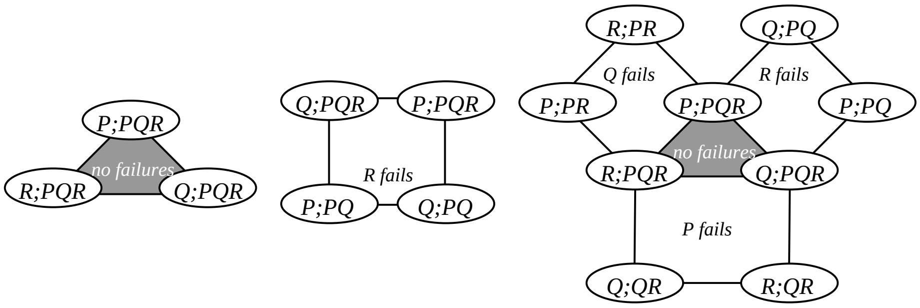

Figure 5: Construction of a one-round three-process protocol complex.

8 Synchronous Computation

We now define the r-round synchronous protocol complex $\mathcal{S}^r(S^m)$. Here too, we consider only a subset of all possible executions: executions in which no more than k processes fail in any round. We are interested in executions where no more than k processes fail in any round. Informally, we will show that the one-round protocol complex is the union of pseudospheres, where each pseudosphere corresponds to the set of executions in which a fixed set of processes fail. For example, Figure 8 illustrates the possible executions of a one-round protocol for three processes, P, Q, and R, starting from a fixed input simplex, in which no more than one process fails. Here, each vertex is labeled with a process, followed by the processes from which it has received messages. The figure on the left represents the execution in which in which no processes fail: this is a (degenerate) pseudosphere in which each process receives the same set of messages. The figure in the middle represents the executions in which R alone fails. This complex is a pseudosphere: P and Q independently do or do not receive a message from R. The figure on the right represents the entire one-faulty protocol complex. It is the union of the failure-free pseudosphere with the three single-failure pseudospheres.

8.1 Single-Round Protocols

Let $\mathcal{S}^1(S^n)$ be the complex of one-round executions of an $(n+1)$-process protocol with input simplex S^n in which at most k processes fail. It is the union of complexes $\mathcal{S}^1_K(S^n)$ of one-round executions starting from S^n in which *exactly* the processes in K fail. Given a set K of process ids, let $S^n \backslash K$ be the face of S^n labeled with the process ids *not* in K. Our next result says that $\mathcal{S}^1_K(S^n)$ is a pseudosphere, which means that $\mathcal{S}^1(S^n)$ is a union of pseudospheres:

Lemma 19: If $m \geq n - f$ and K is a subset of $ids(S^m)$ of size at most $f - (n-m)$, then

$$\mathcal{S}^1_K(S^m) \cong \psi(S^m \backslash K; 2^K).$$

Proof: Each vertex in $\mathcal{S}^1_K(S^m)$ has the form $\langle p_i, M\rangle$, where $p_i \in \mathit{ids}(S^m \backslash K)$ and M is a set of messages received from other processes in the round. Every nonfaulty processor receives a message from every other nonfaulty processor,, and also from some subset of the faulty processors in K. Define the vertex map

$$\iota : \mathcal{S}^1_K(S^m) \to \psi(S^m \backslash K; 2^K)$$

by $\iota\,\langle p_i, M\rangle = \langle \vec{s}_i, \mathit{ids}(M) \cap K\rangle$. The vertex map ι is simplicial and one-to-one. $\square$

Note the similarity between the two models: the one-round complex is a single pseudosphere in the asynchronous model (Lemma 14) and a union of pseudospheres in the synchronous model (Lemma 19). To compute the connectivity of this union using Theorem 2, we need to understand the intersections. The next lemma shows that these intersections have a simple structure: they are themselves the union of pseudospheres. Order the process sets lexicographically: the empty set first, followed by singleton sets, followed by two-element sets, and so on. Let $K_0, \ldots, K_\ell$ be the sequence of sets of process ids less than or equal to K_ℓ, listed in lexicographic order.

Lemma 20: Let $m \geq n - f$ and let $K_0, \ldots, K_k$ be the subsets of $\mathit{ids}(S^m)$ of size at most $f - (n - m)$ arranged in lexicographical order. If $K_0, \ldots, K_\ell$ is a prefix of this sequence, then

$$\bigcup_{i=0}^{\ell-1} \mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m) = \bigcup_{p \in K_\ell} \psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}}).$$

Proof: First consider the $\subseteq$ inclusion. Lemmas 4 and 19 imply that

$$\begin{aligned}\mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m) &= \psi(S^m \backslash K_i; 2^{K_i}) \cap \psi(S \backslash K_\ell; 2^{K_\ell}) \\ &= \psi(S^m \backslash K_i \cap S^m \backslash K_\ell; 2^{K_i \cap K_\ell})\end{aligned}$$

Notice that $S^m \backslash K_i \cap S^m \backslash K_\ell \subset S^m \backslash K_\ell$. Notice also that $K_i \cap K_\ell \subset K_\ell - \{p\}$ for some p, since the lexicographical ordering of K_i before K_ℓ guarantees that K_ℓ is not a subset of K_i, so $p \in K_\ell - K_i$ for some p. It follows that $\psi(S^m \backslash K_i \cap S^m \backslash K_\ell; 2^{K_i \cap K_\ell})$ is a subcomplex of $\psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}})$. Consequently, each $\mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m)$ is contained in some $\psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}})$.

Now consider the $\supseteq$ inclusion. Since $K_\ell - \{p\}$ is a proper subset of K_ℓ, it must precede K_ℓ in the lexicographical ordering, so $K_\ell - \{p\} = K_i$ for some $i < \ell$. Lemmas 4 and 19 imply that

$$\begin{aligned}\mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m) &= \mathcal{S}^1_{K_\ell - \{p\}}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m) \\ &= \psi(S^m \backslash K_\ell - \{p\}; 2^{K_\ell - \{p\}}) \cap \psi(S^m \backslash K_\ell; 2^{K_\ell}) \\ &= \psi(S^m \backslash (K_\ell - \{p\}) \cap S^m \backslash K_\ell; 2^{K_\ell - \{p\} \cap K_\ell}) \\ &= \psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}}).\end{aligned}$$

The last step follows because $(K_\ell - \{p\}) \cup K_\ell = K_\ell$ implies $S^m \backslash (K_\ell - \{p\}) \cap S^m \backslash K_\ell = S^m \backslash K_\ell$, and because $(K_\ell - \{p\}) \cap K_\ell = K_\ell - \{p\}$. Consequently, each $\psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}})$ is contained in some $\mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m)$. $\square$

Let $\mathcal{S}^1(S^n)$ denote the protocol complex for a one-round synchronous $(n+1)$-process protocol with input simplex S^n where no more than than k processes fail.

Lemma 21: $\mathcal{S}^1(S^m)$ is $(m-(n-k)-1)$-connected if $m \geq (n-f)+k$ and $n \geq 2k$.

Proof: Suppose $0 \leq m < n-k$. In this case, $m-(n-k)-1 < -1$, and by convention every complex is k-connected when $k < -1$.

Suppose $n-k \leq m \leq n$. Let $K_0, \ldots, K_k$ be subsets of $ids(S^m)$ of size at most k listed in lexicographical order so that

$$\mathcal{S}^1(S^m) = \bigcup_{i=0}^{\ell} \mathcal{S}^1_{K_i}(S^m).$$

We proceed by induction on ℓ to how that

$$\bigcup_{i=0}^{\ell} \mathcal{S}^1_{K_i}(S^m)$$

is $(m-(n-k)-1)$-connected, and the lemma will follow from the final case $\ell = k$.

Suppose $\ell = 0$. In this case, $K_0 = \emptyset$, so $\mathcal{S}^1_{K_0}(S^m) = \psi(S^m; 2^{\emptyset}) \cong S^m$. This m-dimensional simplex is $(m-1)$-connected, and the hypothesis holds since $m-1 \geq m-(n-k)-1$.

Suppose $\ell > 0$ and the induction hypothesis holds for $\ell - 1$. We must prove that $\mathcal{K} \cup \mathcal{L}$ is $(m-(n-k)-1)$-connected, where

$$\begin{aligned} \mathcal{K} &= \bigcup_{i=0}^{\ell-1} \mathcal{S}^1_{K_i}(S^m) \\ \mathcal{L} &= \mathcal{S}^1_{K_\ell}(S^m) \end{aligned}$$

By the induction hypothesis for $\ell - 1$, the complex $\mathcal{K}$ is $(m-(n-k)-1)$-connected. By Lemma 19, $\mathcal{L} = \mathcal{S}^1_{K_\ell}(S^m) = \psi(S^m \backslash K_\ell; 2^{K_\ell})$. By Corollary 6, the complex $\mathcal{L}$ is $(m-k-1)$-connected since the dimension of $S^m \backslash K_\ell$ is at least $m-k$, and it follows that $\mathcal{L}$ is $(m-(n-k)-1)$-connected since $n \geq 2k$ implies $m-k-1 \geq m-(n-k)-1$.

To finish the proof that $\mathcal{K} \cup \mathcal{L}$ is $(m-(n-k)-1)$-connected, the Mayer-Vietoris Theorem 2 says that it is enough to prove that $\mathcal{K} \cap \mathcal{L}$ is $(m-(n-k)-2)$-connected. Lemma 20 implies that

$$\mathcal{K} \cap \mathcal{L} = \bigcup_{i=0}^{\ell-1} \mathcal{S}^1_{K_i}(S^m) \cap \mathcal{S}^1_{K_\ell}(S^m) = \bigcup_{p \in K_\ell} \psi(S^m \backslash K_\ell; 2^{K_\ell - \{p\}}).$$

Since the dimension of $S^m \backslash K_\ell$ is at least $m-k$ and since the power sets $2^{K_\ell - \{p\}}$ share the empty set as an element, Corollary 8 implies that $\mathcal{K} \cap \mathcal{L}$ is $(m-k-1)$-connected. Since $n \geq 2k$ implies that $m-k-1 \geq m-(n-k)-1 \geq m-(n-k)-2$, it follows that $\mathcal{K} \cap \mathcal{L}$ is $(m-(n-k)-2)$-connected. $\square$

The reader is invited to compare both the statements and proofs of Lemmas 15 and 21.

8.2 Multi-Round Protocols

Let $\mathcal{S}^r(S^n)$ be the protocol complex for an r-round synchronous $(n+1)$-process protocol with input simplex S^n where no more than than k processes fail in each round. We can decompose this complex as follows. Let $K_0, \ldots, K_\ell$ be a sequence of sets of k or fewer process ids in lexicographic order. Recall that $\mathcal{S}^1_{K_i}(S^n) = \psi(S^n \backslash K_i; 2^{K_i})$ is the complex of one-round executions in which exactly the processes in K_i fail. The set of r-round executions in which exactly the processes in K_i fail in the first round can be written as $\mathcal{S}^{r-1}_i(\mathcal{S}^1_{K_i}(S^n))$, where $\mathcal{S}^{r-1}_i$ is the complex for an $(r-1)$-round, $(f-|K_i|)$-faulty, $(n-|K_i|+1)$-process full-information protocol. The $\mathcal{S}^{r-1}_i$ are considered distinct protocols because the $\mathcal{S}^1_{K_i}(S^n)$ have varying dimensions. Taking the union over all the K_i, we have

$$\mathcal{S}^r(S^n) = \bigcup_{i=0}^{\ell} \mathcal{S}^{r-1}_i(\mathcal{S}^1_{K_i}(S^n)).$$

The connectivity of a protocol $\mathcal{P}$ depends on the *degree* of the protocol. Consider the multi-round executions of $\mathcal{P}$ in which f_i is the maximum number of processes that fail at round i. The *degree* of $\mathcal{P}$ is the minimum f_i for any round. Define $\widetilde{\mathcal{S}}^{r-1}_\ell$ to be the protocol identical to $\mathcal{S}^{r-1}_\ell$ except that it fails at most $k-1$ processes in its first round. While $\mathcal{S}^{r-1}_\ell$ has degree k, $\widetilde{\mathcal{S}}^{r-1}_\ell$ has degree $k-1$. Our next result implies that intersections of the complexes comprising $\mathcal{S}^r$ are equivalent to $\widetilde{\mathcal{S}}^{r-1}_\ell$ applied to a union of pseudospheres, which makes it possible to use Theorem 2 to analyze the connectivity of $\mathcal{S}^r$.

Lemma 22:

$$\bigcup_{i=0}^{\ell-1} \mathcal{S}^{r-1}_i(\mathcal{S}^1_{K_i}(S^n)) \cap \mathcal{S}^{r-1}_\ell(\mathcal{S}^1_{K_\ell}(S^n)) = \widetilde{\mathcal{S}}^{r-1}_\ell \left(\bigcup_{j \in K_\ell} \psi(S^n \backslash K_\ell; 2^{K_\ell - \{j\}}) \right).$$

Proof: For the $\subseteq$ inclusion, we have seen in the proof of Lemma 20 that for each i, there is some $j \in K_\ell$ such that

$$\psi(S^n \backslash K_i \cap S \backslash K_\ell; 2^{K_i \cap K_\ell}) \subseteq \psi(S^n \backslash K_\ell; 2^{K_\ell - \{j\}}).$$

Under what circumstances can $\mathcal{S}^{r-1}_i$ and $\mathcal{S}^{r-1}_\ell$ agree on this complex? Because P_j has already failed in $\mathcal{S}^{r-1}_\ell$, the only executions of $\mathcal{S}^{r-1}_i$ that agree with $\mathcal{S}^{r-1}_\ell$ are ones in which P_j fails without sending any messages to nonfaulty processes. But then $\mathcal{S}^{r-1}_i$ (and therefore $\mathcal{S}^{r-1}_\ell$) can fail at most $k-1$ processes that do send messages to nonfaulty processes. Any such execution is also an execution of $\widetilde{\mathcal{S}}_\ell^{r-1}$.

For the reverse inclusion, we have seen that for each $j \in K_\ell$,

$$\mathcal{S}^1_{K_\ell - \{j\}}(S^n) \cap \mathcal{S}^1_{K_\ell}(S^n) = \psi(S^n \backslash K_\ell; 2^{K_\ell - \{j\}}).$$

The set of executions in which the two protocols agree are exactly those in which $\mathcal{S}_i^{r-1}$ immediately fails P_j, and in which $\mathcal{S}_\ell^{r-1}$ fails no more than $k-1$ processes. These executions comprise $\widetilde{\mathcal{S}}_\ell^{r-1}$. □

What follows is scratch work.

Let $\mathcal{S}_{P,K}(S^m)$ be the complex of one-round executions of the full-information protocol in which only the processors in K fail, and starting with the processors in P and the input simplex S^m, if $ids(S^m) \subseteq P$ and $P - ids(S^m) \subseteq K$; and define $\mathcal{S}_{P,K}(S^m)$ to be the empty complex otherwise. The condition $ids(S^m) \subseteq P$ says that initial inputs are provided for some of the processors, and the condition $P - ids(S^m) \subseteq K$ says that processors for which no input is provided can be considered to have failed immediately before having sent a single message. Given a sequence $\Sigma = K_1, \ldots, K_r$ of sets of processors, this definition of the one-round complex $\mathcal{S}_{P,K}(S^m)$ can be extended in the obvious way to an r-round complex by

$$\mathcal{S}_{P,\Sigma}(S^m) = \mathcal{S}_{P_r,K_r} \cdots \mathcal{S}_{P_2,K_2}\mathcal{S}_{P_1,K_1}(S^m)$$

where each $P_i = P - \bigcup_{j=1}^{i-1} K_j$. A consequence of these definitions is that $\mathcal{S}_{P,\Sigma}(S^m)$ is the empty set unless $ids(S^m) \subseteq P$ and $P - ids(S^m) \subseteq K_1$.

Define a *failure pattern* to be a sequence $\sigma = \sigma_1, \ldots, \sigma_r$ of integers representing upper bounds on the number of processes allowed to fail in each of the first r rounds of the full-information protocol. The failure pattern of most interest to us will be the failure pattern in which k processors fail in each round. We say that a failure pattern has *degree* k if it is a nondecreasing sequence of integers beginning with k. We say that a sequence $\Sigma = K_1, \ldots, K_r$ of processor sets satisfies σ if $|K_i| \leq \sigma_i$ for each i. The lexicographical ordering on sets of processors can be extended to sequences $K_1, \ldots, K_r$ of processor sets satisfying σ in the obvious way, by ordering with respect to K_1, then K_2, etc.

Let $\mathcal{S}^r_{P,\sigma}(S^m)$ be the complex of r-round executions of the full-information protocol with failure pattern σ, and starting with the processors in P and the input simplex S^m. This complex is defined by the union

$$\mathcal{S}^r_{P,\sigma}(S^m) = \bigcup_{\Sigma \text{ satisfying } \sigma} \mathcal{S}_{P,\Sigma}(S^m)$$

We say that $\mathcal{S}^r_{P,\sigma}(S^m)$ has degree k if σ has degree k.

Lemma 23: Let σ be a failure pattern of degree k and let $\Sigma_1, \ldots, \Sigma_k$ be the lexicographical ordering of the sequences of process sets satisfying σ.

$$\bigcup_{i=0}^{\ell-1} \mathcal{S}_i^{r-1}(\mathcal{S}_{K_i}^1(S^n)) \cap \mathcal{S}_\ell^{r-1}(\mathcal{S}_{K_\ell}^1(S^n)) = \widetilde{\mathcal{S}}_\ell^{r-1}\left(\bigcup_{j \in K_\ell} \psi(S^n \backslash K_\ell; 2^{K_\ell - \{j\}})\right).$$

Lemma 24: If $n \geq rk + k$, and $\mathcal{S}^r$ is an r-round, $(n+1)$-process protocol with degree k, then $\mathcal{S}^r(S^m)$ is $(m-(n-k)-1)$-connected.

Proof: We argue by induction on r. When $r = 1$, the claim follows from Lemma 21. As induction hypothesis, assume the claim for any $(r-1)$-round protocol.

First, consider the case where $m = n$. Let $K_0, \ldots, K_\ell$ be the sequence of all sets of process ids less than or equal to K_ℓ, listed in lexicographic order. We will show inductively that if $|K_\ell| \leq k$,

$$\bigcup_{i=0}^{\ell} \mathcal{S}_i^{r-1}(\mathcal{S}_{K_i}^1(S^n)) \text{ is } (k-1)\text{-connected.}$$

The claim then follows when K_ℓ is the maximal set of size k. For the base case, when $\ell = 0$, $K_0 = \emptyset$, $\mathcal{S}_\emptyset^1(S^n)$ is $\psi(S^n; 2^\emptyset) \cong S^n$, and $\mathcal{S}_0^{r-1}(S^n)$ is $(k-1)$-connected by the induction hypothesis for r.

For the induction step for ℓ, assume

$$\mathcal{K} = \bigcup_{i=0}^{\ell-1} \mathcal{S}_{K_i}^1(S^n) \text{ is } (k-1)\text{-connected.}$$

By Lemma 19,

$$\mathcal{L} = \mathcal{S}_\ell^{r-1}(\mathcal{S}_{K_\ell}^1(S^n)) = \mathcal{S}_\ell^{r-1}(\psi(S^n \backslash K_\ell; 2^{K_\ell})).$$

Recall that $\mathcal{S}_\ell^{r-1}$ is a k-faulty, $(n-|K_\ell|+1)$-process, $(r-1)$-round protocol, where $n - |K_\ell| \geq (r-1)k$, so by the induction hypothesis, for each simplex $S^d \in \psi(S^n \backslash K_\ell; 2^{K_\ell})$, $\mathcal{S}_\ell^{r-1}(S^d)$ is $(d-(n-|K_\ell|-k)-1)$-connected. By Theorem 5, $\mathcal{S}_\ell^{r-1}(\psi(S^n \backslash K_\ell; 2^{K_\ell}))$ is $(k-1)$-connected. By Lemma 22,

$$\mathcal{K} \cap \mathcal{L} = \widetilde{\mathcal{S}}_\ell^{r-1}\left(\bigcup_{i \in K_\ell} \psi(S^n \backslash K_\ell; 2^{K_\ell - \{i\}})\right).$$

This protocol has $(r-1)$-rounds and degree $(k-1)$. By the induction hypothesis, for any simplex S^{n-k}, $\widetilde{\mathcal{S}}_\ell^{r-1}(S^{n-k})$ is $(k-2)$-connected. Let $A_i = 2^{K_\ell - \{i\}}$, for $i \in K_\ell$. $\cap_{i \in K_\ell} A_i \neq$, so $\mathcal{K} \cap \mathcal{L}$ is also $(k-2)$-connected by Theorem 7. The claim now follows from Theorem 2.

If $n > m \geq n-k$ $\mathcal{S}^r(S^m)$ is equivalent to an m-process protocol in which at most $k-(n-m)$ processes fail in the first round, and k thereafter. This protocol has degree $k-(n-m)$, so $\mathcal{S}^r(S^m)$ is $(k-(n-m)-1)$-connected, which is $(m-(n-k)-1)$.

When $m < n-k$, $m-(n-k)-1 < -1$, and $\mathcal{S}^r(S^m)$ is empty, so the condition holds vacuously. □

The connectivity of this protocol complex implies the lower bound for synchronous k-set agreement [CHLT93]:

Theorem 25: If $n \geq f+k$, then any synchronous f-resilient k-set agreement protocol requires $\lfloor f/k \rfloor + 1$ rounds. If $n < f+k$, then any synchronous f-resilient k-set agreement protocol requires $\lfloor f/k \rfloor$ rounds.

Proof: If $n-k \geq f$, then $\mathcal{S}^{\lfloor f/k \rfloor}(\mathcal{I})$ is $(k-1)$-connected. If $n-k < f$, then $\mathcal{S}^{\lfloor f/k \rfloor - 1}(\mathcal{I})$ is $(k-1)$-connected. Either way, 11 states that the protocol cannot solve k-set agreement. □

9 Semi-Synchronous Computation

Finally, we define the r-round semi-synchronous protocol complex $\mathcal{M}^r(S^m)$. In this model, the time between two consecutive steps of a process is at least c_1 and at most c_2, and the time to deliver a message is at most d. The values c_1, c_2, and d are known constants, and we define $C = c_2/c_1$.

As in the previous two sections, we will show that any full-information protocol complex in the semi-synchronous model has a well-behaved subcomplex that can be expressed as a well-structured union of pseudospheres. The structure of this union is more complex than in the synchronous case, but the formally analysis is nearly the same.

9.1 Failure Patterns

Once again, we restrict attention to *round-structured* executions defined as follows. Each *round* takes exactly time d. All messages sent during a round are delivered at the very end of that round (at multiples of time d). All processes take steps in lockstep as quickly as possible (at multiples of time c_1). The interval between process steps is called a *microround*, and there are $\mu = \lfloor d/c_1 \rfloor$ microrounds per round. Each message is labeled with the microround in which it was sent. The *view* of a process P_i at the end of a round is a sequence $\langle \mu_0, \ldots, \mu_n \rangle$, where μ_j is the microround of the last message received from P_j (or 0 if no message was received from P_j).

Consider a set K of processes, and consider all single-round executions in which K is the set of failing processes. A *failure pattern* for K is a function F mapping each process $P_j \in K$ to the microround μ_j in which it fails. At the end of the round, there are a number of views consistent with F, since the last message received by P_i from a process P_j failing in microround μ_j will be sent either in microround μ_j or $\mu_j - 1$. We define $[F]$ to be the set of possible views $\langle \mu_0, \ldots, \mu_n \rangle$ produced by F:

$$\mu_i = \begin{cases} F(P_i) - 1 \text{ or } F(P_i) & \text{if } P_i \in K \\ \mu & \text{otherwise} \end{cases}$$

We define $[F \uparrow j]$ to be the subset of $[F]$ in which P_j's last message is delivered in microround μ_j (and not $\mu_j - 1$). If F is a failure pattern for K, and $j \in K$, then $[F \uparrow j]$ is defined to be the set of views $\langle \mu_0, \ldots, \mu_n \rangle$, where

$$\mu_i = \begin{cases} F(P_i) - 1 \text{ or } F(P_i) & \text{if } P_i \in K - \{j\} \\ F(P_i) & \text{if } i = j \\ \mu & \text{otherwise.} \end{cases}$$

We order failure patterns for a fixed set K in reverse lexicographical order: the first pattern fails all processs in K at micro-round μ, and the last at 0.

Let $\mathcal{M}^1(S^n)$ be the complex of one-round executions of an $(n+1)$-process protocol with input simplex S^n in which at most k processes fail. It is the union of complexes $\mathcal{M}^1_{K,F}(S^n)$ denoting protocol complex of one-round executions starting from S^n in which the processes in K fail with pattern F. The next result says that $\mathcal{M}^1_{K,F}(S^n)$ is a pseudosphere, so the one-round protocol complex is a union of pseudospheres:

Lemma 26: $\mathcal{M}^1_{K,F}(S^n) \cong \psi(S^n \backslash K; [F])$.

Proof: Each vertex in $\mathcal{M}^1_{K,F}(S^n)$ has the form $\langle P_i, M \rangle$, where $P_i \in \mathit{ids}(S^n \backslash K)$ and M is a set of messages received from other processes in the round. This set contains a message from every nonfaulty process in every microround, and a message from every faulty process in some (possibly empty) prefix of microrounds. Each message is labeled with the microround in which it was sent. Let ι map each vertex v of $\mathcal{M}^1_{K,F}(S^n)$ to the vertex v' of $\psi(S^n \backslash K; [F])$ where the j-th component of the view labeling v' is the microround of the last message from P_j in the set of messages labeling v. This is an isomorphism. □

Once again, notice the similarity among the models (Lemmas 14, 19, and 26). In each case, the one-round complex is the union of pseudospheres, where the structure of the union reflects the timing and failure properties of the model. The pseudospheres $\psi(S^n \backslash K; [F])$ forming the one-round complex $\mathcal{M}^1(S^n)$ are lexicographically ordered by the ordering on process sets K and the ordering on failure patterns F, ordering first by K and then by F. Let $\psi(S^n \backslash K_0; [F_0]), \ldots, \psi(S^n \backslash K_\ell; [F_\ell])$ be the sequence of pseudospheres less than or equal to $\psi(S^n \backslash K_\ell; [F_\ell])$, listed in this order. Let

$$\mathcal{K} = \bigcup_{i=0}^{\ell-1} \psi(S^n \backslash K_i, [F_i]) \text{ and } \mathcal{L} = \psi(S^n \backslash K_\ell; [F_\ell]).$$

Lemma 27:

$$\mathcal{K} \cap \mathcal{L} = \bigcup_{j \in K_\ell} \psi(S^n \backslash K_\ell; [F_\ell \uparrow j]).$$

Proof: For the $\subseteq$ inclusion, Lemma 4 implies that

$$\psi(S^n \backslash K_i; [F_i]) \cap \psi(S^n \backslash K_\ell; [F_\ell]) = \psi(S^n \backslash K_i \cap S^n \backslash K_\ell; [F_i] \cap [F_\ell]).$$

$S^n \backslash K_i \cap S^n \backslash K_\ell \subseteq S^n \backslash K_\ell$. If K_i is distinct from K_ℓ, then it precedes it lexicographically, and there is some $j \in K_\ell - K_i$. If $[F_i] \cap [F_\ell]$ is nonempty, then $F_\ell(j) = \mu$, because F_i does not fail P_j. Otherwise, if $K_i = K_\ell$, then F_i precedes F_ℓ, and there is some j such that $F_i(j) > F_\ell(j)$. In either case, $[F_i] \cap [F_\ell] \subseteq [F_\ell \uparrow j]$, proving the inclusion.

Consider the reverse inclusion. Suppose $F_\ell(j) = \mu$. Let F be the failure pattern that agrees with F_ℓ on $K_\ell - \{j\}$, but does not fail P_j. $[F] \cap [F_\ell] = [F_\ell \uparrow j]$ because every view has μ in position j, and either $F(i) - 1$ or $F(i)$ in position $i \in K_\ell - \{j\}$. $\psi(S^n \backslash (K_\ell - \{j\}); F)$ precedes $\psi(S^n \backslash K_\ell; F_\ell)$ in the sequence, and hence

$$\psi(S^n \backslash (K_\ell - \{j\}); [F]) \cap \psi(S^n \backslash K_\ell; [F_\ell]) = \psi(S^n \backslash K_\ell; [F_\ell \uparrow j]) \subset \mathcal{K} \cap \mathcal{L}.$$

Suppose instead that $F_\ell(j) < \mu$. Let F be the failure pattern that agrees with F_ℓ on $K_\ell - \{j\}$, but $F(j) = F_\ell(j) + 1$. $[F] \cap [F_\ell] = [F_\ell \uparrow j]$. $\psi(S^n \backslash K_\ell; F)$ precedes $\psi(S^n \backslash K_\ell; F_\ell)$ in the sequence, and hence

$$\psi(S^n \backslash K_\ell; [F]) \cap \psi(S^n \backslash K_\ell; [F_\ell]) = \psi(S^n \backslash K_\ell; [F_\ell \uparrow j]) \subset \mathcal{K} \cap \mathcal{L}.$$

$\square$

Let $\mathcal{M}^1(S^n)$ denote the protocol complex for a one-round semi-synchronous $(n+1)$-process protocol with input simplex S^n where no more than than k processes fail. This result says that the intersections of the pseudospheres making up $\mathcal{M}^1(S^n)$ are just the unions of other pseudospheres. This makes is possible to use Theorem 2 to compute the connectivity of their union $\mathcal{M}^1(S^n)$. For the case of one-round protocols, we can prove that one particular union $\mathcal{M}^1(S^m)$ of pseudospheres is $(m-(n-k)-1)$-connected for all $m \leq n$ when $n \geq 2k$. Iterating this construction r times to define the r-round protocol complex $\mathcal{M}^r(S^m)$, using techniques analogous to those used in the synchronous model, we can prove that $\mathcal{M}^r(S^m)$ is also highly-connected.

Lemma 28: If $n \geq 2k$, then $\mathcal{M}^1(S^m)$ is $m - (n-k) - 1)$-connected.

Proof: There are three cases: $m = n$, $n - t \leq m < n$, and $m < n - t$. First consider $m = n$. Let $\psi(S^n \backslash K_0, [F_0]), \ldots, \psi(S^n \backslash K_\ell, [F_\ell])$ be the sequence of pseudospheres less than or equal to $\psi(S^n \backslash K_\ell, [F_\ell])$, listed in this order. If K_ℓ is the maximal set of k process ids, and F_ℓ its maximal failure pattern, then

$$\mathcal{M}^1(S^n) = \bigcup_{i=0}^{\ell} \mathcal{M}^1_{K_i,F_i}(S^n).$$

We will show inductively that if $|K_\ell| \leq k$,

$$\bigcup_{i=0}^{\ell} \mathcal{M}^1_{K_i,F_i}(S^n) \text{ is } (k-1)\text{-connected.}$$

The claim then follows when $\psi(S^n\ K_\ell; [F_\ell])$ is the maximal pseudosphere where $|K_i| = k$.

For the base case, when $\ell = 0$, $K_0 = \emptyset$, F_0 fails no processe, $\mathcal{M}^1_{K_0,F_0}(S^n) = \psi(S^n; [F_0]) \cong S^n$, which is $(n-1)$-connected, and $n - 1 \geq k - 1$.

For the induction step, assume

$$\mathcal{K} = \bigcup_{i=0}^{\ell-1} \mathcal{M}^1_{K_i,F_i}(S^n) \text{ is } (k-1)\text{-connected.}$$

By Lemma 26,

$$\mathcal{L} = \mathcal{M}^1_{K_\ell,F_\ell}(S^n) = \psi(S^n \backslash K_\ell; [F_\ell]).$$

Since $\dim(S^n \backslash K_\ell) \geq n - k$, $\mathcal{L}$ is $(n-k-1)$-connected by Corollary 6, and because $n \geq 2k$, $n - k - 1 \geq k - 1$.

By Lemma 27,

$$\mathcal{K} \cap \mathcal{L} = \bigcup_{i=0}^{\ell-1} \mathcal{M}^1_{K_i,F_i}(S^n) \cap \mathcal{F}^1_{K_\ell,F_\ell}(S^n) = \bigcup_{i \in K_\ell} \psi(S^n \backslash K_\ell; [F_\ell \uparrow i])$$

Let $A_i = [F \uparrow i]$. $\cap_{i \in K_\ell} A_i \neq \emptyset$, and $S^n \backslash K_\ell$ has dimension at least $n - k$, so Corollary 8 implies that $\mathcal{K} \cap \mathcal{L}$ is $(n-k-1)$-connected. Since $n \geq 2k$, $n - k - 1 \geq k - 1$.

When $n - t \leq m < n$, recall that $\mathcal{M}^1(S^m)$ is the set of executions in which only the processes in S^m take steps. This protocol is equivalent to the complex of $(m+1)$-process, $(k-n+m)$-faulty protocol executions. Substituting m for n and $k - n + m$ for k, $\mathcal{M}^1(S^m)$ is $(m - (n-k) - 1)$-connected.

When $m < n - k$, $m - (n-k) - 1 < -1$, and $\mathcal{M}^1(S^m)$ is empty, so the condition holds vacuously. $\square$

The reader is invited to note the similarities between the statements and proofs of Lemmas 15, 21, and 28.

9.2 Multi-Round Protocols

We now consider multi-round protocol executions in which at most f_i processes fail at each round. As before, the *degree* of a protocol is the smallest value of f_i for any round.

Let $\mathcal{M}^r$ be the protocol complex for r-round, $(n+1)$-process protocol executions where no more than than k processes fail in each round. We can decompose this complex as follows. Let $\psi(S^n \backslash K_0, [F_0]), \ldots, \psi(S^n \backslash K_\ell, [F_\ell])$ be the sequence of pseudospheres less than or equal to $\psi(S^n \backslash K_\ell, [F_\ell])$, where K_ℓ is the maximal set of k process ids, and F_ℓ the maximal failure pattern for K, Recall that $\mathcal{M}_{K_i,F_i}(S^n) = \psi(S^n \backslash K_i; [F_i])$ is the complex of one-round executions in which exactly the processes in K_i fail with pattern F_i. The set of r-round executions in which exactly the processes in K_i fail in the first round with pattern F_i can be written as $\mathcal{M}^{r-1}_i(\mathcal{M}_{K_i,F_i}(S^n))$, where $\mathcal{M}^{r-1}_i$ is the complex for an $(r-1)$-round, $(n - |K_i| + 1)$-process full-information protocol. The $\mathcal{M}^{r-1}_i$ are considered distinct protocols because the $\mathcal{M}_{K_i,F_i}(S^n)$ have varying dimensions. Taking the union over all the K_i, we have

$$\mathcal{M}^r(S^n) = \bigcup_{i=0}^{\ell} \mathcal{M}_i^{r-1}(\mathcal{M}_{K_i,F_i}(S^n)).$$

Define $\widetilde{\mathcal{M}}_\ell^{r-1}$ to be the protocol identical to $\mathcal{M}_\ell^{r-1}$ except that it fails at most $k-1$ processes in its first round. While $\mathcal{M}_\ell^{r-1}$ has degree k, $\widetilde{\mathcal{M}}$ has degree $k-1$.

Lemma 29: If $F_\ell(j) < \mu$ for some $j \in K_\ell$,

$$\bigcup_{i=0}^{\ell-1} \mathcal{M}_i^{r-1}(\mathcal{M}^1_{K_i}(S^n)) \cap \mathcal{M}_\ell^{r-1}(\mathcal{M}^1_{K_\ell}(S^n)) = \mathcal{M}_\ell^{r-1}\left(\bigcup_{j \in K_\ell} \psi(S^n \backslash K_\ell; [F\ell \uparrow j])\right).$$

Proof: For the $\subseteq$ inclusion, we have seen in the proof of Lemma 27 that for each i there is a j such that

$$\mathcal{M}^1_{K_i}(S^n) \cap \mathcal{M}^1_{K_\ell}(S^n) \subseteq \psi(S^n \backslash K_\ell; [F\ell \uparrow j]).$$

It follows that

$$\mathcal{M}_i^{r-1}(\mathcal{M}^1_{K_i}(S^n)) \cap \mathcal{M}_\ell^{r-1}(\mathcal{M}^1_{K_\ell}(S^n)) \subseteq \mathcal{M}_\ell^{r-1}(\psi(S^n \backslash K_\ell; [F\ell \uparrow j])).$$

For the reverse inclusion, let F be the failure pattern that agrees with F_ℓ on $K_\ell - \{j\}$, but $F(j) = F_\ell(j) + 1$.

$$\mathcal{M}^1_{K_\ell,F}(S^n) \cap c M^1_{K_\ell,F_\ell}(S^n) = \psi(S^n \backslash K_\ell; [F]) \cap \psi(S^n \backslash K_\ell; [F_\ell]) = \psi(S^n \backslash K_\ell; [F_\ell \uparrow j])$$

Both $\mathcal{M}_i^{r-1}$ and $\mathcal{M}_\ell^{r-1}$ agree on this intersection. □

Lemma 30: If $F_\ell(j) = \mu$ for all $j \in K_\ell$,

$$\bigcup_{i=0}^{\ell-1} \mathcal{M}_i^{r-1}(\mathcal{M}^1_{K_i}(S^n)) \cap \mathcal{M}_\ell^{r-1}(\mathcal{M}^1_{K_\ell}(S^n)) = \widetilde{\mathcal{M}}_\ell^{r-1}\left(\bigcup_{j \in K_\ell} \psi(S^n \backslash K_\ell; [F\ell \uparrow j])\right).$$

Proof: Consider the $\subseteq$ inclusion. Since F_ℓ is the least failure pattern for K_ℓ, any K_i lexicographically precedes K_ℓ, and there exists $j \in K_\ell - K_i$. We have seen in the proof of Lemma 27 that

$$\psi(S^n \backslash K_i \cap S \backslash K_\ell; [F_i] \cap [F_\ell]) \subseteq \psi(S^n \backslash K_\ell; [F_\ell \uparrow j]).$$

Under what circumstances can $\mathcal{M}_i^{r-1}$ and $\mathcal{M}_\ell^{r-1}$ agree on this complex? Because P_j has already failed in $\mathcal{M}_\ell^{r-1}$, the only executions of $\mathcal{M}_i^{r-1}$ that agree with $\mathcal{M}_\ell^{r-1}$ are ones in which P_j fails without sending any messages to non-faulty processes. But then $\mathcal{M}_i^{r-1}$ (and therefore $\mathcal{M}_\ell^{r-1}$) can fail at most $k-1$

processes that do send messages to nonfaulty processes. Any such execution is also an execution of $\widetilde{\mathcal{M}}_\ell^{r-1}$.

For the reverse inclusion, let F be the failure pattern that agrees with F_ℓ on $K_\ell - \{j\}$, but does not fail P_j. $[F] \cap [F_\ell] = [F_\ell \uparrow j]$. The set of executions in which the two protocols agree are exactly those in which $\mathcal{M}_i^{r-1}$ immediately fails P_j, and in which $\mathcal{M}_\ell^{r-1}$ fails no more than $k-1$ processes. These executions comprise $\widetilde{\mathcal{M}}_\ell^{r-1}$. □

Lemma 31: If $n \geq (r+1)k$, then $\mathcal{M}^r(S^m)$ is $(m-(n-k)-1)$-connected.

Proof: We argue by induction on r. When $r=1$, the claim follows from Lemma 28.

As induction hypothesis, assume the claim for any $(r-1)$-round protocol.

First, consider the case where $m=n$. Let $\psi(S^n \backslash K_0, [F_0]), \ldots, \psi(S^n \backslash K_\ell, [F_\ell])$ be the sequence of pseudospheres less than or equal to $\psi(S^n \backslash K_\ell, [F_\ell])$, listed in this order. We will show inductively that if $|K_\ell| \leq k$,

$$\bigcup_{i=0}^{\ell} \mathcal{M}_i^{r-1}(\mathcal{M}_{K_i}^1(S^n)) \text{ is } (k-1)\text{-connected.}$$

The claim then follows when $\psi(S^n \backslash K_\ell; [F_\ell])$ is the maximal pseudosphere for $|K_\ell = k$.

For the base case, when $\ell = 0$, $K_0 = \emptyset$, $\mathcal{M}_{K_0}^1(S^n)$ is $\psi(S^n; [F_0]) \cong S^n$, and $\mathcal{M}_0^{r-1}(S^n)$ is $(k-1)$-connected by the induction hypothesis for r.

For the induction step for ℓ, assume

$$\mathcal{K} = \bigcup_{i=0}^{\ell-1} \mathcal{M}_{K_i}^1(S^n) \text{ is } (k-1)\text{-connected.}$$

By Lemma 26,

$$\mathcal{L} = \mathcal{M}_\ell^{r-1}(\mathcal{M}_{K_\ell}^1(S^n)) = \mathcal{M}_\ell^{r-1}(\psi(S^n \backslash K_\ell; [F_\ell])).$$

Recall that $\mathcal{M}_\ell^{r-1}$ is a k-faulty, $(n-|K_\ell|+1)$-process, $(r-1)$-round protocol, where $n - |K_\ell| \geq (r-1)k$, so by the induction hypothesis, for each simplex $S^d \in \psi(S^n \backslash K_\ell; 2^{K_\ell})$, $\mathcal{M}_\ell^{r-1}(S^d)$ is $(d-(n-|K_\ell|-k)-1)$-connected. By Theorem 5, $\mathcal{M}_\ell^{r-1}(\psi(S^n \backslash K_\ell; 2^{K_\ell}))$ is $(k-1)$-connected. By Lemmas 29 and 30, either

$$\mathcal{K} \cap \mathcal{L} = \mathcal{M}_\ell^{r-1}\left(\bigcup_{i \in K_\ell} \psi(S^n \backslash K_\ell; 2^{K_\ell - \{i\}})\right).$$

or

$$\mathcal{K} \cap \mathcal{L} = \widetilde{\mathcal{M}}_\ell^{r-1}\left(\bigcup_{i \in K_\ell} \psi(S^n \backslash K_\ell; 2^{K_\ell - \{i\}})\right).$$

This protocol has $(r-1)$-rounds and degree either k or $(k-1)$. By the induction hypothesis, for any simplex S^{n-k}, $\widetilde{\mathcal{M}}_\ell^{r-1}(S^{n-k})$ is at least $(k-2)$-connected.

Let $A_i = 2^{K_\ell - \{i\}}$, for $i \in K_\ell$. $\cap_{i \in K_\ell} A_i \neq$, so $\mathcal{K} \cap \mathcal{L}$ is also $(k-2)$-connected by Theorem 7. The claim now follows from Theorem 2.

If $n > m \geq n - k$ $\mathcal{M}^r(S^m)$ is equivalent to an m-process protocol in which at most $k - (n - m)$ processes fail in the first round, and k thereafter. This protocol has degree $k - (n - m)$, so $\mathcal{M}^r(S^m)$ is $(k - (n - m) - 1)$-connected, which is $(m - (n - k) - 1)$.

When $m < n - k$, $m - (n - k) - 1 < -1$, and $\mathcal{M}^r(S^m)$ is empty, so the condition holds vacuously. $\square$

9.3 Retiming

So far, we have shown that if $n \geq (r+1)k$, the r-round protocol executions in which at at most k processes fail comprise a $(k-1)$-connected complex, which cannot solve k-set agreement. These executions take time rd, which is short. We now show how to "stretch" the protocol.

First, consider the following wait-free case. There are exactly $n = (r+1)k+1$ processes, and we have a "failure budget" of $f = (r+1)k$. We have established that k-set agreement has no decision map on $\mathcal{M}^r(\mathcal{I})$.

Let $\mathcal{M}^{r+1}_\epsilon(\mathcal{I})$ denote the protocol complex at time $(r+1)d - \epsilon$ for $d > \epsilon > 0$, at time ϵ before round $r+1$. We claim that k-set agreement has no decision map on $\mathcal{M}^{r+1}_\epsilon(\mathcal{I})$. No process has received a message since time rd, so any decision it could make after waiting $d - \epsilon$ without a message could have been made at time rd, implying that $\mathcal{M}^r(\mathcal{I})$ has a decision map.

Let $\mathcal{M}^{r+1}_\infty(\mathcal{I})$ denote the protocol complex corresponding to the following executions: for each vertex $\vec{v}$ in $\mathcal{M}^r(\mathcal{I})$, where $P = id(\vec{v})$, fail all processes but P, and run P as slowly as possible, (taking steps at multiples of time c_2). At time $rd + Cd$, P will time out, but at time $rd + Cd - \epsilon$, this execution is indistinguishable to P from the corresponding execution in $\mathcal{M}^{r+1}_\epsilon(\mathcal{I})$, and therefore$\mathcal{M}^{r+1}_\infty(\mathcal{I})$ has no decision map. We have just shown a worst-case lower bound of time

$$\left(\left\lfloor \frac{f}{k} \right\rfloor - 1 \right) d + Cd$$

to solve k-set agreement wait-free with $f = n$ failures.

For $f < n$, we use a group G of $n - f + 1$ processes to "simulate" a single P_i as follows. Each process in G starts with the same input value, receives exactly the same messages, and every other process receives all or none of the messages sent by G. Messages between a member of G and a nonmember Each member of G receives messages from outside the group at intervals of d, as before, but it receives messages sent from other members of G to at multiples of d/c_2.

The resulting r-round, $(n+1)$-process protocol complex $\mathcal{G}^r_i(\mathcal{I})$ is not the same as the r-round, $(f+1)$-process protocol complex $\mathcal{M}^r(\mathcal{I})$, but the two complexes have the same homotopy type, (just contract each simplex spanned by processes in G to a point), and are therefore both $(k-1)$-connected.

is has the same homotopy type as but it has the same homotopy type as an f-dimensional pseudosphere (just contract each simplex spanned by processes in G to a point).

In every round, every process in G receives exactly the same messages. The resulting protocol complex is not a pseudosphere, but it has the same homotopy type as an f-dimensional pseudosphere (just contract each simplex spanned by processes in G to a point).

Using this result and the execution-stretching techniques from [ADLS94], we can prove that

Corollary 32: Any wait-free protocol for k-set agreement and $n+1$ processes in the semi-synchronous model requires time $\lfloor \frac{n}{k} \rfloor d + Cd$.

As noted above, this corollary is the first substantial new lower bound for the semi-synchronous model since the Attiya, Dwork, Lynch, and Stockmeyer consensus bound of 1993 [ADLS94]. We believe that this result can be extended to the f-resilient case, but this will require further work.

References

[ABND+87] Hagit Attiya, Amotz Bar-Noy, Danny Dolev, Daphne Koller, David Peleg, and Rudiger Reischuk. Achievable cases in an asynchronous environment. In *Proceedings of the 28th IEEE Symposium on Foundations of Computer Science*, pages 337–346, October 1987.

[ABND+90] Hagit Attiya, Amotz Bar-Noy, Danny Dolev, David Peleg, and Rudiger Reischuk. Renaming in an asynchronous environment. *Journal of the ACM*, July 1990.

[ADLS94] Hagit Attiya, Cynthia Dwork, Nancy Lynch, and Larry Stockmeyer. Bounds on the time to reach agreement in the presence of timing uncertainty. *Journal of the ACM*, 41(1):122–152, January 1994.

[Awe85] Baruch Awerbuch. Complexity of network synchronization. *Journal of the ACM*, 32:801–823, October 1985.

[BG93] Elizabeth Borowsky and Eli Gafni. Generalized FLP impossibility result for t-resilient asynchronous computations. In *Proceedings of the 25th ACM Symposium on Theory of Computing*, May 1993.

[BG97] Elizabeth Borowsky and Eli Gafni. A simple algorithmically reasoned characterization of wait-free computations. In *Proceedings of the 16th Annual ACM Symposium on Principles of Distributed Computing*, pages 189–198, August 1997.

[Cha91] Soma Chaudhuri. Towards a complexity hierarchy of wait-free concurrent objects. In *Proceedings of the 3rd IEEE Symposium on Parallel and Distributed Processing.* IEEE, December 1991. Also appeared as Technical Report No. 91-024, Iowa State University, 1991.

[CHLT93] Soma Chaudhuri, Maurice Herlihy, Nancy Lynch, and Mark R. Tuttle. A tight lower bound for k-set agreement. In *Proceedings of the 34th IEEE Symposium on Foundations of Computer Science*. IEEE, October 1993. To appear.

[DM90] Cynthia Dwork and Yoram Moses. Knowledge and common knowledge in a Byzantine environment: Crash failures. *Information and Computation*, 88(2):156–186, October 1990.

[Fis83] Michael J. Fischer. The consensus problem in unreliable distributed systems (a brief survey). In Marek Karpinsky, editor, *Proceedings of the 10th International Colloquium on Automata, Languages, and Programming*, pages 127–140. Springer-Verlag, 1983. A preliminary version appeared as Yale Technical Report YALEU/DCS/RR-273.

[FL82] Michael J. Fischer and Nancy A. Lynch. A lower bound for the time to assure interactive consistency. *Information Processing Letters*, 14(4):183–186, June 1982.

[FLP85] Michael J. Fischer, Nancy A. Lynch, and Michael S. Paterson. Impossibility of distributed consensus with one faulty processor. *Journal of the ACM*, 32(2):374–382, April 1985.

[Gla70] L. C. Glaser. *Geometrical Combinatorial Topology*, volume 1. Van Nostrand Reinhold, New York, 1970.

[GY97] Eli Gafni and Jiong Yang. A round-by-round failure detector — unifying synchrony and asynchrony. Unpublished manuscript, April 1997.

[Had83] Vassos Hadzilacos. A lower bound for Byzantine agreement with fail-stop processors. Technical Report TR–21–83, Harvard University, 1983.

[HM90] Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, July 1990.

[HR94] Maurice Herlihy and Sergio Rajsbaum. Set consensus using arbitrary objects. In *Proceedings of the 13th Annual ACM Symposium on Principles of Distributed Computing*, pages 324–333, August 1994.

[HS93] Maurice P. Herlihy and Nir Shavit. The asynchronous computability theorem for t-resilient tasks. In *Proceedings of the 25th ACM Symposium on Theory of Computing*, pages 111–120. ACM, May 1993.

[HS97] Gunnar Hoest and Nir Shavit. Towards a topological characterization of asynchronous complexity. In *Proceedings of the 16th Annual ACM Symposium on Principles of Distributed Computing*, pages 199–208, August 1997.

[Lef49] S. Lefschetz. *Introduction to Topology.* Princeton University Press, Princeton, New Jersey, 1949.

[LL90] Leslie Lamport and Nancy Lynch. Distributed computing: Models and methods. In J. van Leewen, editor, *Formal Models and Semantics*, volume B of *Handbook of Theoretical Computer Science*, chapter 19, pages 1157–1199. MITpress, 1990.

[MR98] Yoram Moses and Sergio Rajsbaum. The unified structure of consensus: a layered analysis approach. Unpublished manuscript, January 1998.

[Mun84] J. R. Munkres. *Elements Of Algebraic Topology.* Addison Wesley, Reading MA, 1984.

[PSL80] Marshall Pease, Robert Shostak, and Leslie Lamport. Reaching agreement in the presence of faults. *Journal of the ACM*, 27(2):228–234, 1980.

[Spa66] Edwin H. Spanier. *Algebraic Topology.* Springer-Verlag, New York, 1966.

[SZ93] Michael Saks and Fotis Zaharoglou. Wait-free k-set agreement is impossible: The topology of public knowledge. In *Proceedings of the 25th ACM Symposium on Theory of Computing*, May 1993.