"$\Gamma-$Accurate" Failure Detectors

Rachid Guerraoui and André Schiper

Département d'Informatique
Ecole Polytechnique Fédérale de Lausanne
1015 Lausanne, Switzerland

Abstract. The knowledge about failures needed to solve distributed agreement problems can be expressed in terms of completeness and accuracy properties of failure detectors introduced by Chandra and Toueg. The accuracy properties they have considered restrict the false suspicions that can be made *by all the processes in the system.* In this paper, we define "$\Gamma-$accurate" failure detectors, whose accuracy properties *(only)* restrict the false suspicions that can be made *by a subset Γ of the processes.* We discuss the relations between the classes of $\Gamma-$accurate failure detectors, and the classes of failure detectors defined by Chandra and Toueg. Then we point out the impact of these relations on the solvability of agreement problems.

1 Introduction

1.1 Restricting accuracy

Chandra and Toueg have expressed the knowledge about failures needed to solve distributed agreement problems in terms of *completeness* and *accuracy* properties of failure detectors [4]. Completeness properties require that every process that crashes is eventually permanently suspected, while accuracy properties restrict the mistakes (false suspicions) that can be made by *all the processes in the system.* We extend these accuracy properties, by considering properties that only restrict the false suspicions that can be made by a *subset Γ of the processes.* The new properties are called "Γ-accuracy" properties. Given a subset Γ of the processes in the system:

- *strong Γ-accuracy* is satisfied if no process p in Γ is suspected by any process in Γ before p crashes, and
- *weak Γ-accuracy* is satisfied if some correct process (not necessarily in Γ) is not suspected by any process in Γ.

1.2 Motivation

Our work is motivated by the observation that, because the accuracy properties defined in [4] span the whole system, the formalism does not apply to systems that can be subject to network partitions. Indeed, as failure suspicions are usually implemented with time-outs, the probability that any accuracy property holds (even weak accuracy) during a partition of the system, can be considered to be zero. For instance, in a system partitioned into Π_1 and Π_2, processes in Π_1 most probably suspect processes in Π_2, and processes in Π_2 most probably suspect processes in Π_1, i.e. there is no correct process not suspected by any process. Even the weak accuracy property does not hold while the system is partitioned. By restricting the accuracy properties to subsets Γ of the system, the Γ-accurate failure detectors address these concerns.

1.3 Reliable vs eventually reliable channels

Failure detectors of [4] have been considered in a system model with reliable channels. A reliable channel ensures that if a message m is sent by a process p to a process q, and q is correct, then m is eventually received by q. A reliable channel does not loose messages. The reliable channel assumption is however incompatible with network partitions, that are considered in the paper. Therefore we consider in the paper channels with a weaker reliability property that is called *eventually reliable channels*: if a message m is sent by a process p to a process q, and *both p and q are correct*, then m is eventually received by q. An eventually reliable channel can loose messages. The definition is close to the one considered in [1].

In the paper, we prove both possibility results (i.e. equivalence of failure detectors), and impossibility results (i.e non-equivalence results). In order for these results to be more general, we consider the eventually reliable channel assumption to prove possibility results, and the stronger reliable channel assumption to prove impossibility results.

1.4 Results

We use $\mathcal{P}(\Gamma)$ to denote the class of failure detectors that satisfy strong completeness and strong Γ-accuracy, $\mathcal{S}(\Gamma)$ to denote the class of failure detectors that satisfy strong completeness and weak Γ-accuracy, and $\mathcal{W}(\Gamma)$ to denote the class of failure detectors that satisfy weak completeness and weak Γ-accuracy. The failure detectors $\Diamond\mathcal{P}(\Gamma)$, $\Diamond\mathcal{S}(\Gamma)$ and $\Diamond\mathcal{W}(\Gamma)$ are similarly defined by requiring, roughly speaking, the corresponding Γ-accuracy property to eventually hold.

Consider a set Ω of processes. Among others, this paper establishes the following interesting relations between the Γ-accurate failure detector classes and the Chandra-Toueg classes:

1. Let f be the maximum number of processes of Ω that can crash. For any $\Gamma \subseteq \Omega$, and with eventually reliable channels, if $|\Gamma| > |\Omega|/2$ and $f < |\Omega|/2$, then any failure detector of $\Diamond\mathcal{S}(\Gamma)$ can be transformed into some failure detector of $\Diamond\mathcal{S}$ (which implies $\Diamond\mathcal{S}(\Gamma) \cong \Diamond\mathcal{S}$ [1]). Hence, given that $|\Gamma| > |\Omega|/2$ and $f < |\Omega|/2$, every problem that can be solved with $\Diamond\mathcal{S}$, (e.g consensus [4], uniform consensus [4], atomic broadcast [4], and non-blocking weak atomic

[1] This result has been informally stated in [3] for reliable channels.

commitment [6]) can also be solved with $\diamond\mathcal{S}(\Gamma)$ [2]. These problems can hence be solved whenever some correct process is, roughly speaking, eventually never suspected by a majority of correct processes.
2. For any $\Gamma \subset \Omega$, and with reliable channels, we cannot transform any failure detector of $\diamond\mathcal{W}(\Gamma)$ into some failure detector of $\diamond\mathcal{W}$ (which implies $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$). Hence, for any $\Gamma \subset \Omega$, problems that need $\diamond\mathcal{W}$ (e.g consensus, uniform consensus, atomic broadcast, and non-blocking weak atomic commitment [3]) cannot be solved with $\diamond\mathcal{W}(\Gamma)$.
3. For any $\Gamma \subset \Omega$, and with reliable channels, we cannot transform any failure detector of $\mathcal{P}(\Gamma)$ into some failure detector of $\mathcal{P}$ (which implies $\mathcal{P}(\Gamma) \prec \mathcal{P}$). Hence, for any $\Gamma \subset \Omega$, problems that need $\mathcal{P}$ (e.g election [10], genuine atomic multicast [7], and non-blocking atomic commitment [6]) cannot be solved with $\mathcal{P}(\Gamma)$.

The rest of the paper is organized as follows. Section 2 defines the system model. Section 3 defines "Γ-accurate" failure detectors. Section 4, where we consider eventually reliable channels, establishes the above result 1. In Sections 5, 6 and 7, we assume reliable channels. Section 5 establishes the result 2, and Section 6 establishes the result 3. Section 7 compares Γ-accurate failure detector classes. Finally, Section 8 uses the results established in the paper to compare the resilience of various atomic commitment protocols.

2 Model

Our model of asynchronous computation with failure detection is similar to the one described in [3].

2.1 Failures

A discrete global clock is assumed, and Φ, the range of the clock's ticks, is the set of natural numbers. Processes do not have access to the global clock. The distributed system consists of a set Ω of processes. Processes fail by *crashing*, and failures are permanent. A correct process is a process that does not crash. A *failure pattern* is a function F from Φ to 2^{Ω}, where $F(t)$ denotes the set of processes that have crashed through time t. We assume, as in [4], that in any failure pattern, there is at least one correct process. A *failure detector history* is a function from $\Omega \times \Phi$ to 2^{Ω}, where $H(p,t)$ denotes the set of processes suspected by process p at time t. A *failure detector* is a function $\mathcal{D}$ that maps each failure pattern F to a set of failure detector histories. The processes are connected through asynchronous, either (1) *reliable*, or (2) *eventually reliable* channels, represented by a *message buffer* (see Sect. 2.2):

[2] It has been shown that these problems are solvable with the specified failure detectors, and reliable channels. It can be shown that these problems are also solvable with eventually reliable channels, see [1].

– a *reliable channel* ensures that every message sent by a process p to a process q is eventually received by q, if q is correct.
– an *eventually reliable channel* ensures that every message sent by a process p to a process q is eventually received by q, if q *and* p *are both correct.*

The *eventually reliable* channel provides a weaker model than the *reliable channel*: an eventually reliable channel can loose messages. [3]

2.2 Algorithms

An *algorithm* is a collection A of n deterministic automata $A(p)$ (one per process p). Computation proceeds in steps of the algorithm. In each step of an algorithm A, a process p performs atomically the following phases: (1) p receives a message from q, or a "null" message λ; (2) p queries and receives a value d from its failure detector module (d is said to be *seen* by p); (3) p changes its state and sends a message (possibly null) to some process. This third phase is performed according to (a) the automaton $A(p)$, (b) the state of p at the beginning of the step, (c) the message received in phase 1, and (d) the value d seen by p in phase 2. The message received by a process is chosen non-deterministically among the messages in the message buffer destined to p, and the null message λ. A *configuration* is a pair (I, M) where I is a function mapping each process p to its local state, and M is a set of messages currently in the message buffer. A configuration (I, M) is an initial configuration if $M = \emptyset$. A step of an algorithm A is a tuple $e = (p, m, d, A)$, uniquely defined by the algorithm A, the identity of the process p that takes the step, the message m received by p, and the failure detector value d seen by p during the step. A step $e = (p, m, d, A)$ is *applicable to a configuration* (I, M) if and only if $m \in M \cup \{\lambda\}$. The *unique* configuration that results from applying e to $C = (I, M)$, is noted $e(C)$.

2.3 Schedules and runs

A *schedule* of an algorithm A is a (possibly infinite) sequence of steps of A, noted $S = S[1]; S[2]; \ldots S[k]; \ldots$. A schedule is applicable to a configuration C if (1) S is the empty schedule, or (2) $S[1]$ is applicable to C, $S[2]$ is applicable to $S[1](C)$, etc. Given any schedule S, we note $P(S)$ the set of the processes that have at least one step in S.

A *partial run* of A using a failure detector $\mathcal{D}$, is a tuple $R = < F, H, C, S, T >$ where, F is a failure pattern, H is a failure detector history and $H \in \mathcal{D}(F)$, C is an initial configuration of A, T is a finite sequence of increasing time values, and S is a finite schedule of A such that: (1) $|S| = |T|$, (2) S is applicable to

[3] The rational behind the definition is the following. To ensure eventual reception of the message m sent by p to q, the communication library linked to p will have to retransmit the message m, until m is eventually received by q. If p crashes, retransmission will stop, and q might never receive m. Therefore *eventually reliable* channels ensure reception only if the sender and the receiver are both correct.

C, and (3) for all $i \leq |S|$ where $S[i] = (p, m, d, A)$, we have $p \notin F(T[i])$ and $d = H(p, T[i])$.

A *run* of an algorithm A using a failure detector $\mathcal{D}$, is a tuple $R =< F, H, C, S, T >$ where F is a failure pattern, H is a failure detector history and $H \in \mathcal{D}(F)$, C is an initial configuration of A, S is an infinite schedule of A, T is an infinite sequence of increasing time values, and in addition to the conditions above of a partial run ((1), (2) and (3) above), the two following conditions are satisfied: (4) every correct process takes an infinite number of steps, (5) under the *reliable channel* assumption, every message sent by a process to a correct process is eventually received, and under the *eventually reliable channel* assumption, every message sent by a correct process to a correct process is eventually received.

Let $R =< F, H, C, S, T >$ be a partial run of some algorithm A. We say that $R' =< F', H', C', S', T' >$ is *an extension of* R, if R' is either a run or a partial run of A, and $F' = F$, $H' = H$, $C' = C$, $\forall i$ s.t. $T[1] \leq i \leq T[|T|]$: $S'[i] = S[i]$ and $T'[i] = T[i]$.

3 From accurate to "Γ-accurate" failure detectors

3.1 Accurate failure detectors

Failure detectors are abstractly characterized by completeness and accuracy properties. Completeness properties determine the degree to which crashed processes are suspected. Accuracy properties restrict the mistakes (false suspicions) that a process can make. Two completeness properties are defined in [4]: (1) *strong completeness*: eventually every process that crashes is permanently suspected by every correct process, and (2) *weak completeness*: eventually every process that crashes is permanently suspected by some correct process. The following accuracy properties are defined in [4]: (1) *strong accuracy*: no process is suspected before it crashes; (2) *weak accuracy*: some correct process is never suspected; (3) *eventual strong accuracy*: eventually correct processes are not suspected by any correct process, and (4) *eventual weak accuracy*: eventually some correct process is not suspected by any correct process.

A failure detector class is a set of failure detectors defined by some accuracy and some completeness property. Figure 1 shows the notations introduced in [4]. For example, the class $\diamond\mathcal{S}$ contains failure detectors that satisfy strong completeness and eventual weak accuracy.

3.2 Γ-accuracy properties

The accuracy properties defined by Chandra and Toueg restrict the mistakes made by all the processes in the system. We extend these properties, by considering properties that only restrict the mistakes of a subset Γ of the processes. Given $\Gamma \subseteq \Omega$, we define the Γ-accuracy properties as follows:

	Accuracy			
Completeness	Strong	Weak	$\diamond$Strong	$\diamond$Weak
Strong	P (*Perfect*)	S (*Strong*)	$\diamond P$	$\diamond S$
Weak	Q	W (*Weak*)	$\diamond Q$	$\diamond W$

Fig. 1. Accurate failure detector classes

- *Strong Γ-accuracy* is satisfied if no process p in Γ is suspected by any process in Γ, before p crashes.
- *Weak Γ-accuracy* is satisfied if some correct process (not necessarily in Γ) is never suspected by any process in Γ.
- *Eventual strong Γ-accuracy* is satisfied if eventually no correct process in Γ is suspected by any correct process in Γ.
- *Eventual weak Γ-accuracy* is satisfied if eventually some correct process (not necessarily in Γ) is never suspected by any correct process in Γ.

These accuracy properties can be viewed as generalizations of the accuracy properties defined in [4]. The latters correspond to the case where $\Gamma = \Omega$. The asymmetry of the definitions of Γ-accuracy properties ("p in Γ" in the definition of strong Γ-accuracy, "some correct process not necessarily in Γ" in the definition of weak Γ-accuracy) is because we want the Γ-accuracy properties to satisfy the *inclusion* property: given $\Gamma_1 \subset \Gamma_2$, if strong (respt. weak) Γ_2-accuracy holds, then strong (respt. weak) Γ_1-accuracy also holds. We comment on this in Section 3.4.

3.3 Γ-accurate failure detectors

A *Γ-accurate* failure detector is a failure detector defined by some completeness and some Γ-accuracy property. Figure 2 introduces the notations for Γ-accurate failure detector classes. For example, the class $\diamond\mathcal{S}(\Gamma)$ gathers failure detectors that satisfy strong completeness and eventual weak Γ-accuracy.

	Accuracy			
Completeness	Strong Γ-	Weak Γ-	$\diamond$ Strong Γ-	$\diamond$Weak Γ-
Strong	$P(\Gamma)$	$S(\Gamma)$	$\diamond P(\Gamma)$	$\diamond S(\Gamma)$
Weak	$Q(\Gamma)$	$W(\Gamma)$	$\diamond Q(\Gamma)$	$\diamond W(\Gamma)$

Fig. 2. Γ-accurate failure detector classes

3.4 Simple relations between failure detector classes

Given two failure detectors $\mathcal{D}_1$ and $\mathcal{D}_2$, if there is an algorithm that *transforms* $\mathcal{D}_1$ into $\mathcal{D}_2$, then $\mathcal{D}_2$ is said to be *reducible* to $\mathcal{D}_1$, written $\mathcal{D}_2 \preceq \mathcal{D}_1$ [4]. If every failure detector of a class $\mathcal{C}_2$ is reducible to some failure detector of a class $\mathcal{C}_1$, then $\mathcal{C}_2$ is said to be *weaker* than $\mathcal{C}_1$, written $\mathcal{C}_2 \preceq \mathcal{C}_1$. The relation $\preceq$ is an equivalence relation. If $\mathcal{C}_2 \preceq \mathcal{C}_1$ and $\mathcal{C}_1 \preceq \mathcal{C}_2$, then $\mathcal{C}_1$ and $\mathcal{C}_2$ are said to be *equivalent*, written $\mathcal{C}_1 \cong \mathcal{C}_2$. Finally, if $\mathcal{C}_2 \preceq \mathcal{C}_1$ and $\neg(\mathcal{C}_1 \preceq \mathcal{C}_2)$, then $\mathcal{C}_2$ is said to be *strictly weaker* than $\mathcal{C}_1$, written $\mathcal{C}_2 \prec \mathcal{C}_1$.

The inclusion property: The Γ-accuracy properties satisfy the *inclusion* property: given $\Gamma_1 \subset \Gamma_2$, if Γ_2-accuracy holds (strong, weak, eventually strong, or eventually weak), then Γ_1-accuracy also holds (strong, weak, eventually strong, or eventually weak). Roughly speaking, the inclusion property reflects the intuition that reducing the set Γ should not invalidate the accuracy property.

Note that the inclusion property would not hold if weak (respt. eventual weak) Γ-accuracy had been defined as follows: (eventually) some correct process " in Γ " is never suspected by any process (correct process) in Γ. Take this definition, and consider $\Gamma_1 \subset \Gamma_2$. If weak Γ_2-accuracy holds, then there is some process $p \in \Gamma_2$ that is never suspected by any process in Γ_2. However, $p \in \Gamma_2$ does not imply $p \in \Gamma_1$, i.e weak Γ_1-accuracy does not necessarily hold.

The following lemma state simple relations between the classes of accurate failure detectors, and the classes of $\Gamma-$accurate failure detectors.

Lemma 3.1 *Let $\mathcal{C}$ stand for $\mathcal{P}, \mathcal{Q}, \mathcal{S}, \mathcal{W}, \Diamond\mathcal{P}, \Diamond\mathcal{Q}, \Diamond\mathcal{S}$, or $\Diamond\mathcal{W}$. For any $\Gamma \subseteq \Omega$, and with both reliable and eventually reliable channels, we have $\mathcal{C}(\Gamma) \preceq \mathcal{C}$.*

Proof. Follows directly from the definition and the inclusion property. For any $\Gamma \subseteq \Omega$, any failure detector of class $\mathcal{C}$ satisfies the properties of class $\mathcal{C}(\Gamma)$. We thus trivially have $\mathcal{C} \subset \mathcal{C}(\Gamma)$, which implies $\mathcal{C}(\Gamma) \preceq \mathcal{C}$. □

From strong Γ-accuracy to weak Γ-accuracy: For any $\Gamma \subset \Omega$, eventual strong Γ-accuracy implies eventual weak Γ-accuracy. However, strong Γ-accuracy implies weak Γ-accuracy only if $f < |\Gamma|$ (i.e. if there is at least some correct process in Γ) (see Lemma 3.2). Indeed, assume the processes of Γ do not suspect each others, but suspect all the processes outside Γ and then crash. In this case, strong Γ-accuracy is satisfied whereas weak Γ-accuracy is not.

The following lemma states simple relation between the classes of $\Gamma-$accurate failure detectors.

Lemma 3.2 *For any $\Gamma \subseteq \Omega$, and with both reliable and eventually reliable channels, we have (1) $\mathcal{Q}(\Gamma) \preceq \mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma) \preceq \mathcal{S}(\Gamma)$, $\Diamond\mathcal{Q}(\Gamma) \preceq \Diamond\mathcal{P}(\Gamma)$, $\Diamond\mathcal{W}(\Gamma) \preceq \Diamond\mathcal{S}(\Gamma)$, (2) $\Diamond\mathcal{S}(\Gamma) \preceq \Diamond\mathcal{P}(\Gamma)$, $\Diamond\mathcal{W}(\Gamma) \preceq \Diamond\mathcal{Q}(\Gamma)$, and (3) if $f < |\Gamma|$, we also have $\mathcal{S}(\Gamma) \preceq \mathcal{P}(\Gamma)$ and $\mathcal{W}(\Gamma) \preceq \mathcal{Q}(\Gamma)$.*

Proof. As strong completeness implies weak completeness then we obviously have results (1). Consider now results (2). Let $\mathcal{D}$ be any failure detector of class $\diamond\mathcal{P}(\Gamma)$ (respt. of class $\diamond\mathcal{Q}(\Gamma)$). $\mathcal{D}$ satisfies strong completeness (respt. weak completeness) and eventual strong $\Gamma-$accuracy. Hence, for every failure pattern, eventually no correct process in Γ suspects any correct process in Γ. If there is some correct process in Γ, then $\mathcal{D}$ trivially satisfies eventual weak $\Gamma-$accuracy. Altogether, $\mathcal{D}$ is of class $\diamond\mathcal{S}(\Gamma)$ (respt. of class $\diamond\mathcal{W}(\Gamma)$), which implies $\diamond\mathcal{S}(\Gamma) \preceq \diamond\mathcal{P}(\Gamma)$ (respt. $\diamond\mathcal{W}(\Gamma) \preceq \diamond\mathcal{Q}(\Gamma)$).

Consider now results (3). Let $\mathcal{D}'$ be a failure detector of class $\mathcal{P}(\Gamma)$ (respt. of class $\mathcal{Q}(\Gamma)$). $\mathcal{D}'$ satisfies strong completeness (respt. weak completeness) and strong $\Gamma-$accuracy. As $f < |\Gamma|$, then for every failure pattern, there is some correct process in Γ, and this process is never suspected by any process in Γ. Hence, $\mathcal{D}$ satisfies weak $\Gamma-$accuracy. Altogether, $\mathcal{D}$ is of class $\mathcal{S}(\Gamma)$ (respt. of class $\mathcal{W}(\Gamma)$), which implies $\mathcal{S}(\Gamma) \preceq \mathcal{P}(\Gamma)$ (respt. $\mathcal{W}(\Gamma) \preceq \mathcal{Q}(\Gamma)$). □

4 About weak accuracy and strong completeness

In this section, we present an algorithm that transforms any failure detector of class $\diamond\mathcal{S}(\Gamma)$ (respt. $\mathcal{S}(\Gamma)$) into some failure detector of class $\diamond\mathcal{S}$ (respt. $\mathcal{S}$). Our transformation algorithm is correct under the assumptions (1) Γ is a majority of Ω ($|\Gamma| > |\Omega|/2$), (2) there is a majority of correct processes in Ω ($f < |\Omega|/2$), and (3) channels are eventually reliable. Under these assumptions, we have $\mathcal{S} \cong \mathcal{S}(\Gamma)$ and $\diamond\mathcal{S} \cong \diamond\mathcal{S}(\Gamma)$.

4.1 The transformation algorithm

Let Γ be a majority of Ω ($|\Gamma| > |\Omega|/2$), and assume a majority of correct processes ($f < |\Omega|/2$) and eventual reliable channels. With these assumptions, the algorithm in Figure 3 transforms any failure detector of $\mathcal{S}(\Gamma)$ (respt. of $\diamond\mathcal{S}(\Gamma)$) into some failure detector of $\mathcal{S}$ (respt. of $\diamond\mathcal{S}$). The algorithm uses any failure detector, say $\mathcal{D}_1$, of $\mathcal{S}(\Gamma)$ (respt. of $\diamond\mathcal{S}(\Gamma)$) to emulate the output of a failure detector $\mathcal{D}_2$ of $\mathcal{S}$ (respt. of $\diamond\mathcal{S}$). The emulation is done in a distributed variable, $output(\mathcal{D}_2)$. Each process p has a local copy of $output(\mathcal{D}_2)$, denoted $output(\mathcal{D}_2)_p$, which provides the information that should be given by the local failure detector module of $\mathcal{D}_2$ at process p (noted $\mathcal{D}_{2_p}$). The value of $output(\mathcal{D}_2)_p$ at time t is denoted $output(\mathcal{D}_2, t)_p$. Informally, the algorithm works as follows.

- Every process p periodically sends the message $(p, suspected_p)$ to all processes (line 3), where $suspected_p$ denotes the set of processes that p suspects according to its local failure detector module $\mathcal{D}_{1_p}$.
- When p receives a message of the form $(q, suspected_q)$ (line 4), it executes the following:
  (1) for each r in $suspected_q$, p adds q to $suspecting(r)_p$ (line 7), where $suspecting(r)_p$ denotes the set of processes p thinks are currently suspecting r. If $suspecting(r)_p$ contains a majority of processes, then p adds r to

$output(\mathcal{D}_2)_p$ (line 8);
(2) for each r not in $suspected_q$, p removes q from $suspecting(r)_p$ and removes r from $output(\mathcal{D}_2)_p$ (lines 10-11).[4]

```
   /* Every process p executes the following */

   /* Initialisation */
      suspected_p ← ∅;  /* The set of processes suspected by p */
      output(D_2)_p ← ∅;
            /* The local variable emulating the failure detector module D_2p */
      for each r in Ω: suspecting(r)_p ← ∅;
            /* The set of processes p thinks are currently suspecting r */

   cobegin  /* two concurrent tasks */
   ||  /* Task 1: */
1      repeat forever
2          suspected_p ← D_1p   /* p queries its failure detector module D_1p */
3          send (p, suspected_p) to all ;

   ||  /* Task 2: */
4      when (q, suspected_q) received from some q
5          for each r in Ω
6              if r in suspected_q then
7                  suspecting(r)_p ← (suspecting(r)_p ∪ {q}) ;
8                  if |suspecting(r)_p| > |Ω|/2
                       then output(D_2)_p ← (output(D_2)_p ∪ {r}) ;
9              else
10                 suspecting(r)_p ← (suspecting(r)_p − {q}) ;
11                 output(D_2)_p ← (output(D_2)_p − {r}) ;
   coend
```

Fig. 3. From $\mathcal{S}(\Gamma)$ (respt. $\diamond\mathcal{S}(\Gamma)$) to $\mathcal{S}$ (respt. $\diamond\mathcal{S}$)

4.2 Correctness of the transformation

By Lemma 4.1 below, if $|\Gamma| > |\Omega|/2$, the algorithm in Figure 3 transforms weak Γ-accuracy into weak accuracy. The proof is by contradiction. Similarly, by Lemma 4.2, if $|\Gamma| > |\Omega|/2$, the algorithm also transforms eventual weak Γ-accuracy into eventual weak accuracy (proof also by contradiction). Finally, by Lemma 4.3, if $f < |\Omega|/2$, the transformation of Figure 3 preserves strong

[4] The "correction phase" (lines 10-11) is needed to transform eventual weak Γ-accuracy into eventual weak accuracy, but is not needed to transform weak Γ-accuracy into weak accuracy.

completeness. Altogether, if $|\Gamma| > |\Omega|/2$ and $f < |\Omega|/2$, we get: $\mathcal{S} \cong \mathcal{S}(\Gamma)$ and $\diamond\mathcal{S} \cong \diamond\mathcal{S}(\Gamma)$ (Proposition 4.4).

If $f < |\Omega|/2$, the consensus [4], uniform consensus [4], atomic broadcast [4], and non-blocking weak atomic commitment [6] problems can be solved with any failure detector of the class $\diamond\mathcal{S}(\Gamma)$ and reliable channels. It can be shown that these problems are also solvable with any failure detector of the class $\diamond\mathcal{S}(\Gamma)$ and eventual reliable channels [1]. Thus from Proposition 4.4, $\forall\Gamma \subset \Omega$ such that $|\Gamma| > |\Omega|/2$, if $f < |\Omega|/2$, the consensus, uniform consensus, atomic broadcast and non-blocking weak atomic commitment problems can be solved with any failure detector of the class $\diamond\mathcal{S}(\Gamma)$.

Lemma 4.1 (from weak Γ-accuracy to weak accuracy) *Let $\mathcal{D}_1$ be a failure detector that satisfies weak Γ-accuracy. If $|\Gamma| > |\Omega|/2$ and with eventual reliable channels, the algorithm of Figure 3 transforms $\mathcal{D}_1$ into a failure detector $\mathcal{D}_2$ that satisfies weak accuracy.*

Proof. As $\mathcal{D}_1$ satisfies Γ-weak accuracy, there is a correct process r such that no process in Γ suspects r. Assume (by contradiction) that there is a process p such that r is in $output(D_2)_p$. This means that a majority of processes have suspected r (Fig. 2, line 8). As Γ contains a majority of processes, then some process in Γ must have suspected r: a contradiction. □

Lemma 4.2 (from $\diamond$ weak Γ-accuracy to $\diamond$ weak accuracy) *Let $\mathcal{D}_1$ be a failure detector that satisfies eventual weak Γ-accuracy. If $|\Gamma| > |\Omega|/2$, $f < |\Omega|/2$ and with eventual reliable channels, then the algorithm of Figure 3 transforms $\mathcal{D}_1$ into a failure detector $\mathcal{D}_2$ that satisfies eventual weak accuracy.*

Proof. (By contradiction). As $\mathcal{D}_1$ satisfies eventual weak Γ-accuracy, there is a correct process r, and a time t_1 after which no process in Γ suspects r. Hence, there is a time $t_2 > t_1$, from which no process receives a message $(q, suspected_q)$ from a process q in Γ, such that $suspected_q$ contains r.

As Γ contains at least one correct process (by the assumption $|\Gamma| > |\Omega|/2$ and $f < |\Omega|/2$), there is a process $p \in \Omega$ and a time $t_3 > t_2$, at which p receives a message $(q, suspected_q)$ from a correct process q in Γ, and $r \notin suspected_q$. Thus at t_3, we have $r \notin output(D_2)_p$ (Fig. 2, line 11). Assume (by contradiction) that there is a time $t_4 > t_3$ at which r is (again) in $output(D_2)_p$ (Fig. 2, line 8). This means that p has received, from a majority of processes, messages $(q, suspected_q)$ such that $r \in suspected_q$. As $|\Gamma| > |\Omega|/2$, some process in Γ must have suspected r after time t_1: a contradiction. □

Lemma 4.3 (preserving strong completeness) *Let $\mathcal{D}_1$ be a failure detector that satisfies strong completeness. If $f < |\Omega|/2$ and with eventual reliable channels, the algorithm of Figure 3 transforms $\mathcal{D}_1$ into a failure detector $\mathcal{D}_2$ that also satisfies strong completeness.*

PROOF. Consider time t_1 at which all processes that are not correct have crashed. After t_1, let r be a process that has crashed. As $\mathcal{D}_1$ satisfies strong completeness, there is a time $t_2 > t_1$ after which every correct process p suspects r forever, and sends its suspicion message $(p, suspected_p)$, where $r \in suspected_p$, to all (Fig. 2, line 3). These suspicions are thus sent by correct processes. As there is a majority of correct processes, and channels are eventually reliable, every correct process p eventually receives such a suspicion message from a majority of processes, and puts r into $output(D_2)_p$ forever (Fig. 2, line 8). □

Proposition 4.4 *Let $\Gamma \subset \Omega$ be such that $|\Gamma| > |\Omega|/2$, and consider $\mathcal{S}$, $\mathcal{S}(\Gamma)$, $\diamond\mathcal{S}$ and $\diamond\mathcal{S}(\Gamma)$. If $f < |\Omega|/2$ and with eventual reliable channels, we have $\mathcal{S} \cong \mathcal{S}(\Gamma)$ and $\diamond\mathcal{S} \cong \diamond\mathcal{S}(\Gamma)$.*

PROOF. By Lemma 3.1, we have $\mathcal{S}(\Gamma) \preceq \mathcal{S}$ and $\diamond\mathcal{S}(\Gamma) \preceq \diamond\mathcal{S}$. By Lemma 4.1 and Lemma 4.3, we have $\mathcal{S} \preceq \mathcal{S}(\Gamma)$. Thus $\mathcal{S} \cong \mathcal{S}(\Gamma)$. By Lemma 4.2 and Lemma 4.3, we have $\diamond\mathcal{S} \preceq \diamond\mathcal{S}(\Gamma)$. Thus $\diamond\mathcal{S} \cong \diamond\mathcal{S}(\Gamma)$. □

5 About weak completeness

This section compares the class $\diamond\mathcal{W}(\Gamma)$ (respt. $\mathcal{W}(\Gamma)$) of Γ-accurate failure detectors with the class $\diamond\mathcal{W}$ (respt. $\mathcal{W}$) of accurate failure detectors. We assume $|\Omega| > 2$ and we show that, for any subset $\Gamma \subset \Omega$ and even with reliable channels, we have: $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ and $\mathcal{W}(\Gamma) \prec \mathcal{W}$ [5].

It has been shown that $\diamond\mathcal{W}$ is the weakest failure detector class that enables to solve consensus, atomic broadcast, uniform consensus [3, 4], and non-blocking weak atomic commitment [6]. A consequence of $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ is that neither consensus, atomic broadcast, uniform consensus, nor non-blocking weak atomic commitment is solvable with $\diamond\mathcal{W}(\Gamma)$ (for any subset $\Gamma \subset \Omega$).

Failure detector $\mathcal{D}(\Gamma, r)$. The proofs of the above results ($\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ and $\mathcal{W}(\Gamma) \prec \mathcal{W}$) use a specific failure detector, noted $\mathcal{D}(\Gamma, r)$. The specification of $\mathcal{D}(\Gamma, r)$ is based on a failure pattern that we call *1-pattern*: we say that a failure pattern F is a *1-pattern* if at most one process crashes in F. Similarly, we say that a failure pattern F is a *0-pattern* if no process crashes in F. Consider a subset $\Gamma \subset \Omega$ and $r \in \Omega - \Gamma$. We define $\mathcal{D}(\Gamma, r)$ such that (a) in any *1-pattern* F, $\mathcal{D}(\Gamma, r)(F)$ is the set of histories such that: (a.1) as long as r does not crash, r permanently suspects every other process, and (a.2) $\forall r' \neq r$, as long as r' does not crash, r' permanently suspects r, but r' never suspects any other process, and (b) in any pattern F' that is not a 1-pattern, $\mathcal{D}(\Gamma, r)(F')$ is the set of histories that satisfy strong completeness and strong

[5] The assumption $|\Omega| > 2$ is only needed for $\mathcal{W}(\Gamma) \prec \mathcal{W}$ (for $|\Omega| = 2$, we have $\mathcal{W}(\Gamma) \cong \mathcal{W}$). $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ holds for $|\Omega| > 1$ (for $|\Omega| \leq 1$ all failure detector classes are equivalent). However, for presentation uniformity, we assume in this section that $|\Omega| > 2$.

accuracy (i.e. in any pattern that is not a 1-pattern, $\mathcal{D}(\Gamma, r)$ behaves like a failure detector of the class $\mathcal{P}$).

We show that $\mathcal{D}(\Gamma, r)$ is of class $\mathcal{W}(\Gamma)$ (and also of class $\mathcal{Q}(\Gamma)$, see Lemma 5.1), and no algorithm can transform $\mathcal{D}(\Gamma, r)$ into some failure detector of class $\diamond\mathcal{W}$. More precisely, we show that if there exists an algorithm $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ that transforms $\mathcal{D}(\Gamma, r)$ into some failure detector Δ that satisfies weak completeness, then Δ cannot satisfy eventual weak accuracy (Lemma 5.2 and Lemma 5.3).

Lemma 5.1 *$\mathcal{D}(\Gamma, r)$ is of the classes $\mathcal{W}(\Gamma)$ and $\mathcal{Q}(\Gamma)$.*

Proof. In any run of which failure pattern is not a 1-pattern, $\mathcal{D}(\Gamma, r)$ satisfies strong completeness and strong accuracy. Let $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ be any run with F a 1-pattern. If r is correct, then every process that crashes is permanently suspected by r. If r crashes, then all other processes are correct, and they all suspect r. Hence $\mathcal{D}(\Gamma, r)$ satisfies weak completeness in R. Consider now accuracy. As $|\Omega| > 2$, then there is at least some correct process in R that is never suspected by any process of Γ. Thus $\mathcal{D}(\Gamma, r)$ satisfies weak Γ-accuracy. As no process suspects any other process in Γ then $\mathcal{D}(\Gamma, r)$ satisfies strong Γ-accuracy. Hence, $\mathcal{D}(\Gamma, r)$ satisfies weak completeness, weak Γ-accuracy, and strong Γ-accuracy in R. □

Lemma 5.2 *Let $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ be any algorithm that transforms $\mathcal{D}(\Gamma, r)$ into some failure detector Δ. Let $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ be any partial run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ where F is a 0-pattern. If Δ satisfies weak completeness, then there is an extension $R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_\Omega, T_\Omega >$ of R, where for every correct process p, there is a correct process q, and a time t, $T[|T|] \leq t \leq T_\Omega[|T_\Omega|]$, such that $p \in output(\Delta, t)_q$ in R_Ω.*

Proof: Consider the partial run $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ where F is a $0-$pattern, and let p be any process in Ω. Let $R' =< F', H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ be a partial run such that F' is a $1-$pattern, similar to F, except that in F', p crashes at time $T[|T| + 1]$ (immediately after $T[|T|]$). A process such as p does exist as Ω contains at least two processes. As $\mathcal{D}(\Gamma, r)$ provides the same values both for F and F', and S is applicable to C, then R' is a partial run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$. By the weak completeness property of Δ, there is an extension $R'_p =< F', H_{\mathcal{D}(\Gamma,r)}, C, S_p, T_p >$ of R', and a correct process $q \in \Omega$, such that $p \in output(\Delta, T_p[|T_p|])_q$. Let $S_{Susp(p)}$ be the schedule of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ such that $S_p(C) = S_{Susp(p)}(S(C))$. The schedule $S_{Susp(p)}$ can be viewed as the schedule needed to put p into the $output(\Delta)_q$ of process q.

Consider now the run $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$. As $\mathcal{D}(\Gamma, r)$ provides the same values both for F and F', and S_p is applicable to C, then $R_p =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_p, T_p >$ is an extension of R, and there is a correct process q, such that $p \in output(\Delta, T_p[|T_p|])_q$. By iteratively applying the construction of the partial run R_p to every process $p \in \Omega$, the partial run R can be extended to a partial run $R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_\Omega, T_\Omega >$ where *every* process p is put in $output(\Delta)_q$ for some process q. $\square$

Lemma 5.3 *Let $A_{\mathcal{D}(\Gamma,r)\to\Delta}$ be any algorithm that transforms the failure detector $\mathcal{D}(\Gamma, r)$ into some failure detector Δ. If Δ satisfies weak completeness, then there is a run of $A_{\mathcal{D}(\Gamma,r)\to\Delta}$, where Δ does not satisfy eventual weak accuracy.*

Proof: Consider the partial run $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ with F a $0-$pattern. By Lemma 5.2, there is an extension of R, $R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_\Omega, T_\Omega >$, such that for every process $p \in \Omega$, there is a time t, $T[|T|] \leq t \leq T_\Omega[|T_\Omega|]$, and a correct process q, such that $p \in output(\Delta, t)_q$. Let (I, M) be the configuration $S_\Omega(C)$. Consider now a schedule S_{Mess}, of which steps are defined by: the reception by the processes of all messages in M not received in S_Ω, then the reception by every process of the null message λ. The schedule S_{Mess} is by construction applicable to $S(C)$, and we write $S_\Sigma(C) = S_{Mess}(S_\Omega(C))$. There is a sequence of increasing time values T_Σ, such that $R_\Sigma =< F, H_{\mathcal{D}_r}, C, S_\Sigma, T_\Sigma >$ is an extension of R. In R_Σ, all messages sent to p before time $T[|T|]$ are received by p before $T_\Sigma[|T_\Sigma|]$, and p takes at least one step between $T[|T|]$ and $T_\Sigma[|T_\Sigma|]$.

Therefore, given any partial run R of $A_{\mathcal{D}(\Gamma,r)\to\Delta}$, with F a $0-$pattern, we can extend R to a partial run R_Σ where every process is suspected at some process by Δ. We note $R^0_\Sigma = R_\Sigma$, R^1_Σ an extension of R obtained by applying the construction above to R^0_Σ, R^i_Σ an extension of R obtained by applying the construction above to R^{i-1}_Σ, etc., and $R^\infty_\Sigma = lim_{i\to\infty} R^i_\Sigma$.

In R^∞_Σ, the properties of a partial run are satisfied and, every process takes an infinite number of steps, and every message sent to a process is eventually received. Hence R^∞_Σ is a run of $A_{\mathcal{D}_r\to\Delta}$. Furthermore, for any time t and any process p, there is a time $t' \geq t$ and a process q, such that $p \in output(\Delta, t')_q$. Hence Δ does not satisfy eventual weak accuracy in R^∞_Σ. $\square$

Proposition 5.4 *Let $\Gamma \subset \Omega$, and consider $\diamond\mathcal{W}$ and $\diamond\mathcal{W}(\Gamma)$. We have $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ and $\mathcal{W}(\Gamma) \prec \mathcal{W}$.*

Proof. By Lemma 5.1 and Lemma 5.3, no algorithm can transform any failure detector of $\mathcal{W}(\Gamma)$, into some failure detector of $\diamond\mathcal{W}$. In other words, $\neg(\diamond\mathcal{W} \preceq \mathcal{W}(\Gamma))$. As $\diamond\mathcal{W} \preceq \mathcal{W}$ and $\diamond\mathcal{W}(\Gamma) \preceq \mathcal{W}(\Gamma)$, then $\neg(\diamond\mathcal{W} \preceq \diamond\mathcal{W}(\Gamma))$ and $\neg(\mathcal{W} \preceq \mathcal{W}(\Gamma))$. By Lemma 3.1, we have $\diamond\mathcal{W}(\Gamma) \preceq \diamond\mathcal{W}$ and $\mathcal{W}(\Gamma) \preceq \mathcal{W}$. Altogether, we have $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{W}$ and $\mathcal{W}(\Gamma) \prec \mathcal{W}$. $\square$

6 About Strong Accuracy

This section compares the classes of Γ-accurate failure detectors $\mathcal{P}(\Gamma)$, $\mathcal{Q}(\Gamma)$, $\diamond\mathcal{P}(\Gamma)$, and $\diamond\mathcal{Q}(\Gamma)$, with the classes of accurate failure detectors $\mathcal{P}$, $\mathcal{Q}$, $\diamond\mathcal{P}$, and $\diamond\mathcal{Q}$. We assume in this section that $|\Omega| > 1$, and we show in the following (Proposition 6.4) that, for any subset $\Gamma \subset \Omega$ and even with reliable channels, no algorithm can transform any failure detector of $\mathcal{P}(\Gamma)$ into some failure detector of $\diamond\mathcal{Q}$. Hence, we have: $\mathcal{P}(\Gamma) \prec \mathcal{P}$, $\mathcal{Q}(\Gamma) \prec \mathcal{Q}$, $\diamond\mathcal{P}(\Gamma) \prec \diamond\mathcal{P}$, and $\diamond\mathcal{Q}(\Gamma) \prec \diamond\mathcal{Q}$.

A consequence of $\mathcal{P}(\Gamma) \prec \mathcal{P}$ is that problems requiring $\mathcal{P}$ (e.g election [10], genuine atomic multicast [7], and non-blocking atomic commitment [6]) cannot be solved with $\mathcal{P}(\Gamma)$.

The proof of $\mathcal{P}(\Gamma) \prec \mathcal{P}$, $\mathcal{Q}(\Gamma) \prec \mathcal{Q}$, $\diamond\mathcal{P}(\Gamma) \prec \diamond\mathcal{P}$ and $\diamond\mathcal{Q}(\Gamma) \prec \diamond\mathcal{Q}$, is similar to the proof of the previous section. We introduce a specific failure detector $\mathcal{D}'(\Gamma, r)$ of class $\mathcal{P}(\Gamma)$, and we show that $\mathcal{D}'(\Gamma, r)$ cannot be transformed into some failure detector of $\diamond\mathcal{Q}$. More precisely, we show that if some algorithm $A_{\mathcal{D}'(\Gamma,r)\rightarrow\Delta}$ transforms $\mathcal{D}'(\Gamma, r)$ into some failure detector Δ that satisfies weak completeness, then Δ cannot satisfy eventual strong accuracy.

Failure detector $\mathcal{D}'(\Gamma, r)$. Let r be any process in Ω. We define the failure detector $\mathcal{D}'(\Gamma, r)$ using the notion of $\{r\}-$*pattern*. We say that a failure pattern F is a $\{r\}-$*pattern* if only r can crash in F. Using this notion, we define $\mathcal{D}'(\Gamma, r)$ as follows. Consider $\Gamma \subset \Omega$ and $r \in \Omega - \Gamma$. We define $\mathcal{D}'(\Gamma, r)$ such that (1) in any $\{r\}-$*pattern* F, $\mathcal{D}(\Gamma, r)(F)$ is the set of histories such that (a) r is permanently suspected by every process that has not crashed, and (b) $\forall r' \neq r$, r' is never suspected by any process, and (2) in any pattern F' that is not a $\{r\}-$*pattern*, $\mathcal{D}(\Gamma, r)(F')$ is the set of histories such that $\mathcal{D}'(\Gamma, r)$ satisfies strong completeness and strong accuracy.

Lemma 6.1 *$\mathcal{D}'_k$ is of class $\mathcal{P}(\Gamma)$.*

PROOF: In any $\{r\}-$*pattern*, $\mathcal{D}'(\Gamma, r)$ satisfies strong Γ-accuracy and strong completeness. In any run that is not a $\{r\}-$*pattern*, $\mathcal{D}'(\Gamma, r)$ satisfies strong accuracy and strong completeness. Altogether, $\mathcal{D}'(\Gamma, r)$ is thus of class $\mathcal{P}(\Gamma)$. $\Box$

Lemma 6.2 *Let $A_{\mathcal{D}'(\Gamma,r)\rightarrow\Delta}$ be any algorithm that transforms the failure detector $\mathcal{D}'(\Gamma, r)$ into some failure detector Δ. Let $R =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S, T >$ be any partial run of $A_{\mathcal{D}'(\Gamma,r)\rightarrow\Delta}$ where F is a $0-$pattern. If Δ satisfies weak completeness, then there is an extension $R_\Omega =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S_\Omega, T_\Omega >$ of R, a process q, and a time t, $T[|T|] \leq t \leq T_\Omega[|T_\Omega|]$, such that $r \in output_\Delta(q, t)$.*

PROOF: (similar to the proof of Lemma 5.2) Consider the partial run $R =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S, T >$ where F is a $0-$pattern. Let $R' =< F', H_{\mathcal{D}'(\Gamma,r)}, C, S, T >$ be a partial run where F' is a $1-$pattern, similar to F, except that in F', r crashes at time $T[|T| + 1]$ (immediately after $T[|T|]$). The process r can indeed crash as Ω contains at least two processes. As $\mathcal{D}'(\Gamma, r)$ provides the same values both for F and F', and S is applicable to C, then R' is a partial run of $A_{\mathcal{D}'(\Gamma,r)\rightarrow\Delta}$. By the weak completeness property of Δ, there is an extension $R'_q =< F', H_{\mathcal{D}'(\Gamma,r)}, C, S_q, T_q >$ of R', and a correct process q, such that $r \in output(\Delta, T_q[|T_q|])_q$. Let $S_{Susp(q)}$ be the schedule of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ such that

$S_p(C) = S_{Susp(q)}(S(C))$. Consider now the run $R =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S, T >$. As $\mathcal{D}'(\Gamma, r)$ provides the same values for both F' and F, and S_q is applicable to C, then
$R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_q, T_q >$ is an extension of R, and there is a correct process q, such that $r \in output(\Delta, T_q[|T_q|])_q$. □

Lemma 6.3 *Let $A_{\mathcal{D}'(\Gamma,r)\to\Delta}$ be any algorithm that transforms the failure detector $\mathcal{D}'(\Gamma, r)$ into some failure detector Δ. If Δ satisfies weak completeness, then there is a run of $A_{\mathcal{D}'(\Gamma,r)\to\Delta}$, where Δ does not satisfy eventual strong accuracy.*

Proof: (similar to the proof of Lemma 5.3) Consider the partial run $R =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S, T >$ with F a 0−pattern. By Lemma 6.2, there is an extension of R, $R_\Omega =< F, H_{\mathcal{D}'(\Gamma,r)}, C, S_\Omega, T_\Omega >$, a correct process q, and a time t, $T[|T|] \leq t \leq T_\Omega[|T_\Omega|]$, such that $r \in output(\Delta, t)_q$. As in proof of Lemma 5.3, we build a partial run R_Σ, that is an extension of R where all messages sent to every process p before time $T[|T|]$ are received by p in R_Σ, and p takes at least one step after $T[|T|]$ in R_Σ. We note $R^0_\Sigma = R_\Sigma$, R^1_Σ an extension of R obtained by applying the construction above to R^0_Σ, R^i_Σ an extension of R obtained by applying the construction above to R^{i-1}_Σ, etc., and $R^\infty_\Sigma = lim_{i\to\infty} R^i_\Sigma$.

In R^∞_Σ, the properties of a partial run are satisfied, every process takes an infinite number of steps, and every message sent to a process is eventually received. Hence R^∞_Σ is a run of $A_{\mathcal{D}',r\to\Delta}$. Furthermore, for any time t , there is a correct process q, and a time $t' \geq t$, such that $r \in output(\Delta, t')_q$. Hence Δ does not satisfy eventual strong accuracy in R^∞_Σ. □

Proposition 6.4 *Let $\Gamma \subset \Omega$, and consider $\mathcal{P}(\Gamma)$, $\mathcal{P}$, $\Diamond\mathcal{P}(\Gamma)$, $\Diamond\mathcal{P}$, $\mathcal{Q}(\Gamma)$, $\mathcal{Q}$, $\Diamond\mathcal{Q}(\Gamma)$, and $\Diamond\mathcal{Q}$. We have $\mathcal{P}(\Gamma) \prec \mathcal{P}$, $\Diamond\mathcal{P}(\Gamma) \prec \Diamond\mathcal{P}$, $\mathcal{Q}(\Gamma) \prec \mathcal{Q}$, and $\Diamond\mathcal{Q}(\Gamma) \prec \Diamond\mathcal{Q}$.*

Proof: By Lemma 6.1 and Lemma 6.3, no algorithm can transform any failure detector of $\mathcal{P}(\Gamma)$ into some failure detector of $\Diamond\mathcal{Q}$. In other words, $\neg(\Diamond\mathcal{Q} \preceq \mathcal{P}(\Gamma))$. As $\mathcal{P} \preceq \Diamond\mathcal{Q}$, $\Diamond\mathcal{P} \preceq \Diamond\mathcal{Q}$, $\Diamond\mathcal{Q} \preceq \mathcal{Q}$, $\Diamond\mathcal{Q}(\Gamma) \preceq \mathcal{P}(\Gamma)$, $\Diamond\mathcal{P}(\Gamma) \preceq \mathcal{P}(\Gamma)$, $\mathcal{Q}(\Gamma) \preceq \mathcal{P}(\Gamma)$, then $\neg(\mathcal{P} \preceq \mathcal{P}(\Gamma))$, $\neg(\Diamond\mathcal{P} \preceq \Diamond\mathcal{P}(\Gamma))$, $\neg(\mathcal{Q} \preceq \mathcal{Q}(\Gamma))$, $\neg(\Diamond\mathcal{Q} \preceq \Diamond\mathcal{Q}(\Gamma))$.

By Lemma 3.1, we have $\mathcal{P}(\Gamma) \preceq \mathcal{P}$, $\Diamond\mathcal{P}(\Gamma) \preceq \Diamond\mathcal{P}$, $\mathcal{Q}(\Gamma) \preceq \mathcal{Q}$, and $\Diamond\mathcal{Q}(\Gamma) \preceq \Diamond\mathcal{Q}$. Altogether, we have $\mathcal{P}(\Gamma) \prec \mathcal{P}$, $\Diamond\mathcal{P}(\Gamma) \prec \Diamond\mathcal{P}$, $\mathcal{Q}(\Gamma) \prec \mathcal{Q}$, and $\Diamond\mathcal{Q}(\Gamma) \prec \Diamond\mathcal{Q}$. □

7 Comparing "Γ-accurate" failure detectors

Chandra and Toueg have shown that, for accurate failure detectors, weak completeness can be transformed into strong completeness while preserving accuracy properties [4]. In other words, $\mathcal{Q} \cong \mathcal{P}$, $\Diamond\mathcal{Q} \cong \Diamond\mathcal{P}$, $\mathcal{W} \cong \mathcal{S}$, and $\Diamond\mathcal{W} \cong \Diamond\mathcal{S}$. This section shows that these results do not hold anymore for Γ-accurate failure detectors. More precisely, we show that given $|\Omega| > 2$, for any $\Gamma \subset \Omega$, and even with reliable channels, we have $\mathcal{Q}(\Gamma) \prec \mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma) \prec \diamond\mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma) \prec \mathcal{S}(\Gamma)$, and $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{S}(\Gamma)$ [6].

Our proof is based on the failure detector $\mathcal{D}(\Gamma, r)$, defined in Section 5, which was shown to be of the classes $\mathcal{Q}(\Gamma)$ and $\mathcal{W}(\Gamma)$. We show that $\mathcal{D}(\Gamma, r)$ cannot be transformed into some failure detector of class $\diamond\mathcal{S}(\Gamma)$. More precisely, we show that if an algorithm $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ transforms the failure detector $\mathcal{D}(\Gamma, r)$ into some failure detector Δ that satisfies strong completeness, then Δ cannot satisfy eventual weak $\Gamma-$accuracy. The proof is similar to those of Sections 5 and 6.

Lemma 7.1 *Let $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ be any algorithm that transforms the failure detector $\mathcal{D}(\Gamma, r)$ into some failure detector Δ. Let $R =< F, H_{\mathcal{D}(\Gamma,r)}, I, S, T >$ be any partial run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ where F is a $0-$pattern. If Δ satisfies strong completeness, then there is an extension $R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, I, S_\Omega, T_\Omega >$ of R, where for every process p and every process q in Ω, there is a time t, $T[|T|] \leq t \leq T_\Omega[|T_\Omega|]$, such that $p \in output_\Delta(q, t)$.*

Proof: (similar to the proof of Lemma 5.2). Let $R =< F, H_{\mathcal{D}(\Gamma,r)}, I, S, T >$ be any partial run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$, where F is a $0-$pattern. Let $R' =< F', H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ be a partial run where F' is a $1-$pattern, similar to F, except that in F', p crashes at time $T[|T| + 1]$ (immediately after $T[|T|]$). The process p can indeed crash as Ω contains at least two processes. As $\mathcal{D}(\Gamma, r)$ provides the same values both for F and F', and S is applicable to C, then R' is a partial run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$. By the strong completeness property of Δ, there is an extension $R'_p =< F', H_{\mathcal{D}(\Gamma,r)}, C, S_p, T_p >$ of R', such that for every process q, $p \in output(\Delta, T_q[|T_q|])_q$. As $\mathcal{D}(\Gamma, r)$ provides the same values for both F' and F, and S_p is applicable to C, then $R_p =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_p, T_p >$ is also an extension of R.

By iteratively applying the construction of the partial run R_p to every process $p \in \Omega$, the partial run R can be extended to a partial run R_Ω where every process p is put in $output(\Delta)_q$ for every process q. □

Lemma 7.2 Let $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$ be any algorithm that transforms the failure detector $\mathcal{D}(\Gamma, r)$ into some failure detector Δ. If Δ satisfies *strong completeness*, then there is a run of $A_{\mathcal{D}(\Gamma,r)\rightarrow\Delta}$, where Δ does not satisfy *eventual weak $\Gamma-$accuracy.*

Proof: (similar to the proof of Lemma 5.3) Consider the partial run $R =< F, H_{\mathcal{D}(\Gamma,r)}, C, S, T >$ with F a $0-$pattern. By Lemma 7.1, there is an extension of R, $R_\Omega =< F, H_{\mathcal{D}(\Gamma,r)}, C, S_\Omega, T_\Omega >$, where every process p is put in $output(\Delta)_q$ for every process q. As in proof of Lemma 5.3, we can thus build a run R^∞_Σ, that is an extension of R, where for any time t, for any pair of processes

[6] The assumption $|\Omega| > 2$ is only needed for $\mathcal{W}(\Gamma) \prec \mathcal{S}(\Gamma)$. The other results hold for $|\Omega| > 1$. However, for presentation uniformity we assume that $|\Omega| > 2$.

(p, q), there is a time $t' \geq t$, such that p is put in $output(\Delta)_q$ of q. Hence Δ does not satisfy strong completeness. $\square$

Proposition 7.3 Let $|\Omega| > 2$, $\Gamma \subset \Omega$, and consider $\mathcal{Q}(\Gamma)$, $\mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma)$, $\diamond\mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma)$, $\mathcal{S}(\Gamma)$, $\diamond\mathcal{W}(\Gamma)$, and $\diamond\mathcal{S}(\Gamma)$. We have: $\mathcal{Q}(\Gamma) \prec \mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma) \prec \diamond\mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma) \prec \mathcal{S}(\Gamma)$, and $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{S}(\Gamma)$.

Proof. By Lemma 7.1, and Lemma 7.2, no algorithm can transform any failure detector of $\mathcal{Q}(\Gamma)$, or $\mathcal{W}(\Gamma)$), into some failure detector of $\diamond\mathcal{S}$. In other words, $\neg(\diamond\mathcal{S} \preceq \mathcal{Q}(\Gamma))$ and $\neg(\diamond\mathcal{S} \preceq \mathcal{W}(\Gamma))$. As $\diamond\mathcal{S}(\Gamma) \preceq \mathcal{P}(\Gamma)$, $\diamond\mathcal{S}(\Gamma) \preceq \diamond\mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma) \preceq \mathcal{Q}(\Gamma)$, and $\diamond\mathcal{W}(\Gamma) \preceq \mathcal{W}(\Gamma)$, then $\neg(\mathcal{P}(\Gamma) \preceq \mathcal{Q}(\Gamma))$, $\neg(\diamond\mathcal{P}(\Gamma) \preceq \diamond\mathcal{Q}(\Gamma))$, $\neg(\mathcal{S}(\Gamma) \preceq \mathcal{W}(\Gamma))$, and $\neg(\diamond\mathcal{S}(\Gamma) \preceq \diamond\mathcal{W}(\Gamma))$.

By Lemma 3.2, $\mathcal{Q}(\Gamma) \preceq \mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma) \preceq \diamond\mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma) \preceq \mathcal{S}(\Gamma)$, and $\diamond\mathcal{W}(\Gamma) \preceq \diamond\mathcal{S}(\Gamma)$ Altogether, we have, $\mathcal{Q}(\Gamma) \prec \mathcal{P}(\Gamma)$, $\diamond\mathcal{Q}(\Gamma) \prec \diamond\mathcal{P}(\Gamma)$, $\mathcal{W}(\Gamma) \prec \mathcal{S}(\Gamma)$, and $\diamond\mathcal{W}(\Gamma) \prec \diamond\mathcal{S}(\Gamma)$. $\square$

8 Summary and Discussion

We have defined a formalism to express the knowledge about crash failures in a distributed system, in terms of $\Gamma-$accurate failure detectors. This formalism can be viewed as a generalization of the formalism of accurate failure detectors introduced in [4]. To reuse the results about the solvability of distributed agreement problems stated in [4], we have stated a set of relations between accurate and $\Gamma-$accurate failure detector classes. These relations are summarized in Figure 4. We assume in the figure that $\Gamma \subset \Omega$ and $|\Omega| > 2$. The notation $A \longleftrightarrow B$ means that failure detector classes A and B are equivalent. The notation $A \longrightarrow B$ means that A is strictly weaker than B. The notation $A \longmapsto B$ means that A is strictly weaker than B if $|\Gamma| > |\Omega|/2$ and $f < |\Omega|/2$.

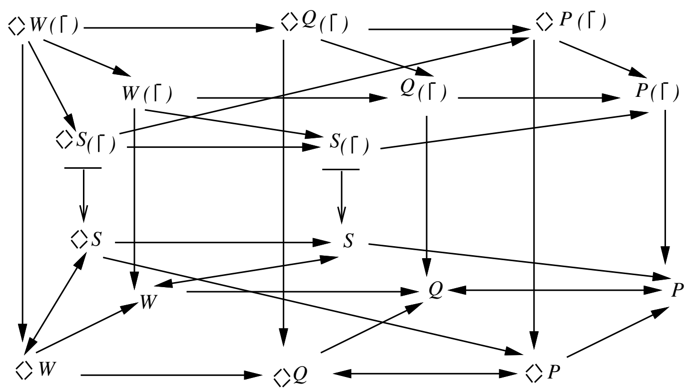

Fig. 4. Relations between failure detector classes

The formalism of Γ-accurate failure detectors enables to characterize and compare some well known distributed protocols designed with network partitions in mind. The initial 3PC protocol (*Three Phase Commit*) proposed by Skeen in 1981 [11] can be seen as requiring the failure detector class $\mathcal{P}$ and reliable channels. In 1982, Skeen proposed a variation of the 3PC protocol, *Quorum Three Phase Commit (Q3PC)* [12], which solves the non-blocking weak atomic commitment problem provided there is a "partition" Γ of correct processes that constitutes a quorum. Stated in our formalism, the Q3PC protocol requires eventual reliable channels, the failure detector class $\mathcal{P}(\Gamma)$, such that Γ is a majority (i.e $|\Gamma| > |\Omega|/2$), and all the processes in Γ correct. As $\mathcal{P}(\Gamma) \prec \mathcal{P}$ (Sect. 6), the Q3PC protocol is an improvement over 3PC.

Later, Keidar and Dolev have defined *Enhanced 3PC (E3PC)* which increases the resilience of Q3PC provided that the quorum exists "eventually" [9]. Stated in our formalism, E3PC assumes a failure detector of the class $\diamond\mathcal{P}(\Gamma)$, such that Γ is a majority (i.e. $|\Gamma| > |\Omega|/2$), and all the processes in Γ correct. As $\diamond\mathcal{P}(\Gamma) \prec \mathcal{P}(\Gamma)$, E3PC is an improvement over Q3PC.

Finally, Guerraoui has shown that the non-blocking weak atomic commitment problem can be solved with $\diamond\mathcal{S}$ [6]. By the result of Section 4, if $f < |\Omega|/2$ and $|\Gamma| > |\Omega|/2$, then $\diamond\mathcal{S} \cong \diamond\mathcal{S}(\Gamma)$. Because $\diamond\mathcal{S}(\Gamma) \prec \diamond\mathcal{P}(\Gamma)$ (Sect. 7), a protocol requiring only $\diamond\mathcal{S}(\Gamma)$ [7] can be seen as an improvement over E3PC. This comparison is somehow unfair as the E3PC protocol is based on a bounded buffer assumption, whereas our eventual reliable channel assumption implicitly requires unbounded buffers (used to store messages to be retansmitted).

The model underlying the E3PC protocol has been described in [5], and the results concerning atomic commitment have been generalized to other consensus-like problems. Babaoğlu et al. have adopted a complementary approach, by discussing the solvability of problems that are weaker than consensus [2], such as *weak-partial group* membership, in a weaker asynchronous system model where channels are not assumed to be eventually reliable. Both models (i.e. [5] and [2]) introduce new failure detector formalisms. Finding out a way to relate our Γ-accurate failure detectors to those introduced in [5] and [2] (and hence the associated results) is still an open issue.

Acknowledgments. We are grateful to Özalp Babaoğlu, Idith Keidar and Aleta Ricciardi, for their valuable comments on earlier drafts of this paper.

References

1. A. Basu, B. Charron-Bost and S. Toueg. Simulating Reliable Links with Unreliable Links in the Presence of Process Crashes. Proceedings of the *10th International Workshop on Distributed Algorithms*, LNCS, Springer Verlag, October 1996.
2. O. Babaoğlu, R. Davoli and A. Montresor. Failure Detectors, GroupMembership and View-Synchronous Communication in Partitionable Systems. *Technical Report, University of Bologna, Computer Science Department.* November 1995.

[7] Such a protocol can be obtained by combining the protocol of [6] which requires $\diamond\mathcal{S}$, with the protocol of Figure 2 (Sect. 4) which transforms $\diamond\mathcal{S}(\Gamma)$ into $\diamond\mathcal{S}$.

3. T. Chandra, V. Hadzilacos and S. Toueg. The weakest failure detector for solving consensus. Journal of the ACM, 43(4), July 1996. A preliminary version appeared in *Proceedings of the 11th ACM Symposium on Principles of Distributed Computing*, pp 147-159. ACM Press. August 1992.
4. T. Chandra and S. Toueg. Unreliable failure detectors for reliable distributed systems. Journal of the ACM, 34(1), pp 225-267, March 1996. A preliminary version appeared in *Proceedings of the 10th ACM Symposium on Principles of Distributed Computing*, pp 325-340. ACM Press. August 1991.
5. R. Friedman, I. Keidar, D. Malki, K. Birman and D. Dolev. Deciding in Partitionable Networks, *Technical Report, Cornell University, Computer Science Department.* November 1995.
6. R. Guerraoui. Revisiting the relationship between Non Blocking Atomic Commitment and Consensus problems. Proceedings of the *9th International Workshop on Distributed Algorithms*, pages 87-100, LNCS 972, Springer Verlag, September 1995.
7. R. Guerraoui and A. Schiper. Atomic Multicast harder than Atomic Broadcast. *Technical Report, Ecole Polytechnique Fédérale de Lausanne, Computer Science Department.* May 1996.
8. R. Guerraoui and A. Schiper. "$\Gamma-$"Accurate Failure Detectors. *Technical Report, Ecole Polytechnique Fédérale de Lausanne, Computer Science Department.* May 1996.
9. I. Keidar and D. Dolev. Increasing the Resilience of Atomic Commit, at No Additional Cost. Proceedings of the *ACM Symposium on Principles of Database Systems*, pages 245-254. ACM Press, May 1994.
10. L. Sabel and K. Marzullo. Election Vs. Consensus in Asynchronous Systems. *Technical Report TR95-1488*, Cornell Univ, 1995.
11. D. Skeen. NonBlocking Commit Protocols. Proceedings of the *ACM SIGMOD International Conference on Management of Data*, pages 133-142. ACM Press, 1981.
12. D. Skeen. A Quorum-Based Commit Protocol. Proceedings of the *Berkeley Workshop on Distributed Data Management and Computer Networks*, pages 69-80, Num 6, 1982.