Distributed Computing manuscript No.
(will be inserted by the editor)

How to Reconcile Fault-Tolerant Interval Intersection with the Lipschitz Condition*

Ulrich Schmid[1], **Klaus Schossmaier**[2]

[1] Technische Universität Wien, Department of Automation, Treitlstraße 1, A-1040 Vienna; e-mail: `s@auto.tuwien.ac.at`
[2] Technische Universität Wien, Department of Automation, Treitlstraße 1, A-1040 Vienna; e-mail: `Klaus.Schossmaier@cern.ch`

Received: 23 August 1999 / Revised version: 15 November 2000

Summary. We present a new fault-tolerant intersection function $\mathcal{F}$, which satisfies the Lipschitz condition for the uniform metric and is optimal among all functions with this property. $\mathcal{F}$ thus settles Lamport's question about such a function raised in [5]. Our comprehensive analysis reveals that $\mathcal{F}$ has exactly the same worst-case performance as the optimal Marzullo function $\boldsymbol{\mathcal{M}}$, which does not satisfy a Lipschitz condition. The utilized modelling approach in conjunction with a powerful hybrid fault model ensures compatibility of our results with any known application framework, including replicated sensors and clock synchronization.

Key words: Fault-tolerant interval intersection – Marzullo function – hybrid fault models – interval-based clock synchronization

1 Motivation

Consider some quantity like a point in real-time (for clock synchronization) or a temperature value (for replicated sensors) that is not known exactly but only within some range. Such a quantity t can be represented by a real interval $\boldsymbol{I} = [x, y]$ containing t, which makes the uncertainty explicit by its length $|\boldsymbol{I}| = y - x$. Now suppose that we are somehow provided with $n \geq 1$ different intervals $\boldsymbol{\mathcal{I}} = \{\boldsymbol{I}_1, \ldots, \boldsymbol{I}_n\}$ all representing the same t, and that we want to extract a single interval of minimum length that contains t. If all input intervals are accurate (i.e. non-faulty), in the sense that $t \in \boldsymbol{I}_i$, $1 \leq i \leq n$, it is obvious that $\boldsymbol{J} = \bigcap_{i=1}^{n} \boldsymbol{I}_i$ contains t and hence $\boldsymbol{J} \neq \emptyset$. In fact, $\boldsymbol{J}$ is the best (deterministic) information about t that can be deduced from $\boldsymbol{\mathcal{I}}$.

However, the question arises what to do if some of the input intervals are not accurate (i.e. faulty), that is, $t \notin \boldsymbol{I}_j$ for some (unknown) j's. Some sort of *fault-tolerant intersection* has to be employed here to compute an interval that is guaranteed to contain t.

It is well-known that, if at most f of the n input intervals may be faulty, the minimum length result containing t is provided by the Marzullo function $\boldsymbol{\mathcal{M}}_n^f(\boldsymbol{\mathcal{I}})$ introduced in [7]: It is the largest interval whose edges lie in the intersection of at least $n - f$ different $\boldsymbol{I}_j$'s. Therefore, to compute for example the left edge of $\boldsymbol{\mathcal{M}}_n^f(\boldsymbol{\mathcal{I}})$, one has to "sweep" over the set of intervals from left to right and stop when $n - f$ intervals intersect for the first time. Thus $\boldsymbol{\mathcal{M}}_n^f(\boldsymbol{\mathcal{I}})$ can be computed in $\mathcal{O}(n \log n)$ time by sorting the intervals' edges, cf. [8]. Fig. 1 shows an example with $n = 4$ and $f = 1$. Note that the unknown t cannot lie in the region between the right edge of $\boldsymbol{I}_3$ and the left edge of $\boldsymbol{I}_2$ in this example. However, since there is no way to decide whether t lies in the area left or right of this region, both must be covered to secure inclusion of t.

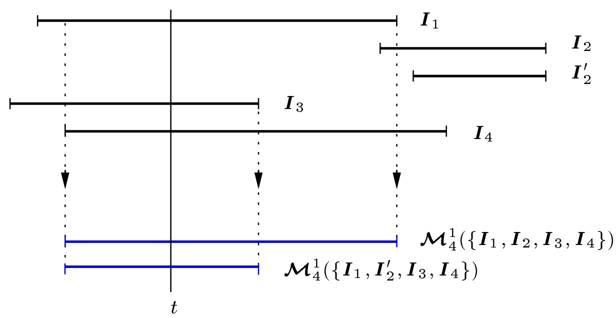

Fig. 1. *Example of the Marzullo function $\boldsymbol{\mathcal{M}}$ with $n = 4$ and $f = 1$. The edges of the result lie in $n - f = 3$ input intervals. Changing interval $\boldsymbol{I}_2$ to $\boldsymbol{I}'_2$ has a big impact on the result.*

It is easily seen, though, that $\boldsymbol{\mathcal{M}}$ exhibits a somewhat irregular behavior: If the left edge of $\boldsymbol{I}_2$ is slightly moved right, as given by $\boldsymbol{I}'_2$, then the right edge of the result suddenly jumps to the right edge of $\boldsymbol{I}_3$. Thus, moving $\boldsymbol{I}_2$ by a small amount ε, just large enough to prohibit intersection with $\boldsymbol{I}_1$, causes a variation by much more than ε. In [5], this behavior was formalized as violation of a *Lipschitz condition* w.r.t. a suitable metric defined on intervals. This is an undesirable

---

* This research has been conducted in our SynUTC-project `http://www.auto.tuwien.ac.at/Projects/SynUTC` supported by the Austrian START programme Y41-MAT.
Correspondence to: Ulrich Schmid (`s@auto.tuwien.ac.at`)

property, since it implies that $\mathcal{M}$ applied to two slightly different input sets may deliver quite different results.

For example, in the clock synchronization context, two nodes p and q usually obtain (slightly) different input sets $\mathcal{I}_p$ and $\mathcal{I}_q$ even if all senders are non-faulty, since intervals are time-dependent here. For that reason, Lamport did not use the Marzullo function for his *Synchronizing Time Servers* [5], but rather an averaging function A^f based on the Fault-Tolerant Average algorithm of [6]. However, Lamport wrote: *"While the averaging function A^f gives reasonable worst-case behavior, it does not make the best use of the available information because it ignores the widths of intervals. Very wide intervals are given the same weight as narrow ones, even though they provide less information. One can construct examples in which the function A^f does not provide the best possible approximation to UT. However, I know of no simple function F satisfying the Lipschitz condition that does better."*

The simple *Fault-Tolerant Interval* (FTI) intersection function $\mathcal{F}$ proposed and analyzed in this paper satisfies a Lipschitz condition and takes into account the widths of intervals. Since $\mathcal{F}$ is in fact optimal among all such functions, we can reasonably claim to have settled Lamport's question.

The remainder of our paper is organized as follows: FTI's definition and basic properties can be found in Section 2, along with the proof that it satisfies the Lipschitz condition and that it is optimal. Section 3 is devoted to the worst case analysis for local and distributed application of $\mathcal{F}$ in presence of faults. Some conclusions in Section 4 eventually round off the paper.

2 FTI Definition and Relations

We consider real intervals $\boldsymbol{I} = [x, y]$, $x \leq y$, where $|\boldsymbol{I}| = y - x$ denotes the interval's *length*, $x = \text{left}(\boldsymbol{I})$ its *left edge*, and $y = \text{right}(\boldsymbol{I})$ its *right edge*. The *intersection* of two intervals is the set-theoretic one, the *union* is defined as $[x, y] \cup [u, v] = [\min\{x, u\}, \max\{y, v\}]$, hence covers the closure of disjoint intervals as well. The *non-commutative union* (*nc-union*) is defined as $[x, y] \sqcup [u, v] = [x, v]$ if $x \leq v$ or $\emptyset$ otherwise. In what follows, we assume a single (unknown) value t, and a set of real intervals $\mathcal{I} = \{\boldsymbol{I}_1, \dots, \boldsymbol{I}_n\}$, $n \geq 1$, that all represent t; we will call such intervals *compatible*. An interval $\boldsymbol{I}$ that is meant to represent t is *accurate* (also termed *correct*) if $t \in \boldsymbol{I}$, otherwise it is *non-accurate* (also termed *faulty*).

The interval-based paradigm and the Marzullo function $\mathcal{M}$ was introduced in Marzullo's thesis [7], and several publications [5], [4], [10], [3], [11], [9], [2], [16], [13], [15], [14], etc. reveal that it is widely applied in practice. The properties of $\mathcal{M}$ have been studied thoroughly both in the context of replicated sensors [8] and clock synchronization [12]. Recall that the latter application differs fundamentally from the former due to the fact that two nodes usually perceive slightly different intervals even from a non-faulty sender.

Our novel Fault-Tolerant Interval intersection function $\mathcal{F}$ is similar to the Marzullo function and defined as follows:

Definition 1 (Function $\mathcal{F}$). *Let a set $\mathcal{I} = \{\boldsymbol{I}_1, \dots, \boldsymbol{I}_n\}$ of $n \geq 1$ non-empty compatible intervals $\boldsymbol{I}_i = [x_i, y_i]$ with at most $f < n$ of those being faulty be given. The* Fault-Tolerant Interval (FTI) *intersection function $\mathcal{F}^f_n(\mathcal{I})$ is defined as the interval*

$$[(f+1)\text{-}max\{x_1, \dots, x_n\}, (f+1)\text{-}min\{y_1, \dots, y_n\}],$$

where h-max$\{z_1, \dots, z_n\}$ resp. h-min$\{z_1, \dots, z_n\}$ gives the h-st largest resp. h-st smallest element of the set $\{z_1, \dots, z_n\}$.

Like $\mathcal{M}^f_n(\mathcal{I})$, our function is translation invariant in the sense that $\mathcal{F}^f_n(\boldsymbol{I}_1 + \Delta, \dots, \boldsymbol{I}_n + \Delta) = \mathcal{F}^f_n(\boldsymbol{I}_1, \dots, \boldsymbol{I}_n) + \Delta$ for any real Δ, and can be computed in $\mathcal{O}(n \log n)$ time by sorting the intervals' edges: In order to compute the left edge of $\mathcal{F}^f_n(\mathcal{I})$, one has to sweep over the set of intervals (by looking at their left edges) from right to left, discarding left edges until the $(f+1)$-st largest is encountered. Similarly, to find the right edge of $\mathcal{F}^f_n(\mathcal{I})$, one has to sweep over the set of intervals from left to right and stop when the $(f+1)$-st smallest right edge is encountered. Fig. 2 shows how $\mathcal{F}$ behaves in the scenario taken from Fig. 1.

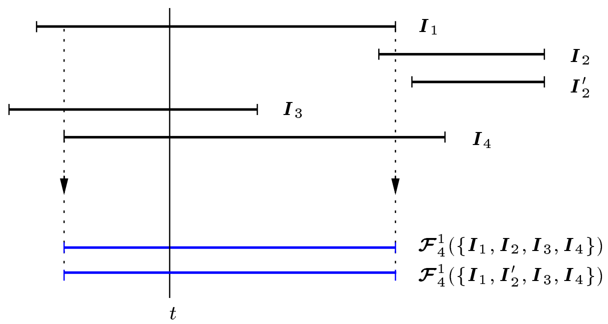

Fig. 2. *Example of the Fault-Tolerant Interval function $\mathcal{F}$ with $n = 4$ and $f = 1$. Changing interval $\boldsymbol{I}_2$ to $\boldsymbol{I}'_2$ has no impact on the result.*

Note carefully that the resulting interval is the same as computed by $\mathcal{M}$ when $\boldsymbol{I}_2$ is used as input interval. Nevertheless, it is apparent here that the right edge of $\mathcal{F}$'s result does not jump when the slightly moved interval $\boldsymbol{I}'_2$ is used instead. In Lemma 4 below, we will show that $\mathcal{F}$ indeed satisfies the desired Lipschitz condition of [5]; its optimality will be proved in Lemma 5.

Definition 1 reveals some similarity between $\mathcal{F}$ and the Fault-Tolerant Midpoint (FTM) algorithm of [6]: Given a set of point values $\mathcal{C} = \{c_1, \dots, c_n\}$ with at most f of those being faulty, $\text{FTM}^f_n(\mathcal{C})$ is defined as the midpoint (center) of the interval

$$[(f+1)\text{-min}\{c_1, \dots, c_n\}, (f+1)\text{-max}\{c_1, \dots, c_n\}].$$

$\mathcal{F}$ can hence be viewed as a generalization of FTM to the interval domain. In fact, $\mathcal{F}$ even emulates FTM when applied to intervals with identical length: If all $\boldsymbol{I}_j = [x_j, y_j]$, $1 \leq j \leq n$, satisfy $|\boldsymbol{I}_j| = l$, we may write $\boldsymbol{I}_j = [c_j - l/2, c_j + l/2]$ with midpoint $c_j = \text{center}(\boldsymbol{I}_j) = (x_j + y_j)/2$. From $\mathcal{F}$'s definition and the fact that h-max$\{x_1, \dots, x_n\} = h$-max$\{c_1, \dots, c_n\} - l/2 = h$-max$\{y_1, \dots, y_n\} - l$ for any h here, it is immediately apparent that $\text{FTM}^f_n(\{c_1, \dots, c_n\}) = \text{center}(\mathcal{F}^f_n(\mathcal{I}))$.

We start the detailed analysis of $\mathcal{F}$'s properties with the following technical Lemma 1, which is needful for proving Lemma 2 and Lemma 5.

Lemma 1 (Edges of $\mathcal{M}$ in presence of disjoint intervals). *Let a set $\mathcal{I} = \{\boldsymbol{I}_1, \ldots, \boldsymbol{I}_n\}$ of $n \geq 1$ non-empty compatible intervals $\boldsymbol{I}_i = [x_i, y_i]$ be given, which yields $\boldsymbol{M} = \mathcal{M}_n^f(\mathcal{I}) \neq \emptyset$. If there are exactly $f_l' + f_r' \leq f$ intervals $\boldsymbol{I}_x \in \mathcal{I}$ with $\boldsymbol{I}_x \cap \boldsymbol{M} = \emptyset$, where $f_l' \geq 0$ resp. $f_r' \geq 0$ of those lie strictly left resp. right of $\boldsymbol{M}$, then*

$$\textit{left}(\boldsymbol{M}) = (f+1-f_l')\text{-}\textit{max}\{x_1, \ldots, x_n\}$$
$$\textit{right}(\boldsymbol{M}) = (f+1-f_r')\text{-}\textit{min}\{y_1, \ldots, y_n\}.$$

Proof. By the definition of $\mathcal{M}$, there must be $n-f$ intervals containing left($\boldsymbol{M}$), and by assumption there are exactly f_l' other intervals strictly left of $\boldsymbol{M}$, which implies that the remaining $f - f_l'$ intervals have a left edge equal or right to left($\boldsymbol{M}$). Hence, the edge $(f - f_l' + 1)$-max$\{x_1, \ldots, x_n\}$ must determine left($\boldsymbol{M}$). The proof for right($\boldsymbol{M}$) is analogous. □

The following Lemma 2 establishes some relations between $\mathcal{F}$ and $\mathcal{M}$.

Lemma 2 (Relations between $\mathcal{F}$ and $\mathcal{M}$). *Given a set $\mathcal{I} = \{\boldsymbol{I}_1, \ldots, \boldsymbol{I}_n\}$ of $n \geq 1$ non-empty compatible intervals and any f such that both $\mathcal{M}_n^f(\mathcal{I}) \neq \emptyset$ and $\mathcal{F}_n^f(\mathcal{I}) \neq \emptyset$, the following relations hold true:*

$$\mathcal{F}_n^f(\mathcal{I}) \supseteq \mathcal{M}_n^f(\mathcal{I}) \tag{1}$$
$$\mathcal{F}_n^f(\mathcal{I}) = \mathcal{M}_n^f(\mathcal{I}) \quad \textit{if } \nexists \boldsymbol{I} \in \mathcal{I} \textit{ with } \boldsymbol{I} \cap \mathcal{M}_n^f(\mathcal{I}) = \emptyset \tag{2}$$
$$\mathcal{F}_n^{n-1}(\mathcal{I}) \equiv \mathcal{M}_n^{n-1}(\mathcal{I}) = \cup_{i=1}^n \boldsymbol{I}_i \tag{3}$$
$$\mathcal{F}_n^0(\mathcal{I}) \equiv \mathcal{M}_n^0(\mathcal{I}) = \cap_{i=1}^n \boldsymbol{I}_i \tag{4}$$

Proof. To show relation (1), let $\boldsymbol{F} = \mathcal{F}_n^f(\mathcal{I})$ and $\boldsymbol{M} = \mathcal{M}_n^f(\mathcal{I})$. Suppose that left($\boldsymbol{M}$) $<$ left($\boldsymbol{F}$), then left($\boldsymbol{M}$) has to be selected from the at most $n-(f+1)$ remaining left edges that are smaller than left($\boldsymbol{F}$). However, left($\boldsymbol{M}$) requires at least $n-f$ intersections, which is not possible here. A similar contradiction can be derived for right($\boldsymbol{F}$) $<$ right($\boldsymbol{M}$).

The equality relation (2) follows directly from Lemma 1 by setting $f_l' = f_r' = 0$ and recalling the definition of $\mathcal{F}$. Finally, the equivalences (3) and (4) are a direct consequence of the functions' definitions. □

Remarks:

1. Since $\mathcal{M}$ computes an accurate result, inclusion (1) guarantees that $\mathcal{F}$ is accurate as well.
2. The behavior of both functions "changes" from intersection to union as f increases, and is identical for the extreme settings.

Next we establish a few useful monotonicity relations of function $\mathcal{F}$ with respect to both parameters and input arguments.

Lemma 3 (Monotonicity). *Let a set $\mathcal{I} = \{\boldsymbol{I}_1, \ldots, \boldsymbol{I}_n\}$ of $n > f \geq 0$ non-empty compatible intervals with f', $0 \leq f' \leq f$, faulty ones among those be given. Then, $\mathcal{F}_n^f(\mathcal{I})$ satisfies the following monotonicity relations:*

(1) $\mathcal{F}_n^f(\mathcal{I}) \subseteq \mathcal{F}_n^{f+k}(\mathcal{I})$ for any integer k with $0 \leq k < n-f$,
(2) $\mathcal{F}_n^f(\mathcal{I}) \subseteq \mathcal{F}_n^f(\mathcal{J})$ for any $\mathcal{J} = \{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_n\}$ with $\boldsymbol{I}_l \subseteq \boldsymbol{J}_l$ for $1 \leq l \leq n$,
(3) For $f \geq f' \geq 1$, if $\mathcal{L} = \mathcal{I} \backslash \{\boldsymbol{I}_j\}$ is obtained by discarding some faulty interval $\boldsymbol{I}_j$ from $\mathcal{I}$, $\mathcal{F}_{n-1}^{f-1}(\mathcal{L})$ is accurate and satisfies

$$\mathcal{F}_{n-1}^{f-1}(\mathcal{L}) \subseteq \mathcal{F}_n^f(\mathcal{I}). \tag{5}$$

Proof. Item (1) of the lemma is trivial, since discarding $f+k$ edges of input intervals in $\mathcal{F}$'s Definition 1 provides a larger interval.

For proving item (2), it is sufficient to establish the following monotonicity properties of the h-smallest and h-largest element of a set:

$$h\text{-min}\{r_1, \ldots, r_n\} \leq h\text{-min}\{r_1 + \varepsilon_1, \ldots, r_n + \varepsilon_n\}$$
$$h\text{-max}\{l_1, \ldots, l_n\} \geq h\text{-max}\{l_1 - \varepsilon_1, \ldots, l_n - \varepsilon_n\}$$

for any $\varepsilon_i \geq 0$, $1 \leq i \leq n$, and any integer $1 \leq h \leq n$. Let us thus assume that, say, the property for the "h-min-part" was not true. However, this would imply that there are h different indices j with $r_j \leq r_j + \varepsilon_j \leq h\text{-min}\{r_1 + \varepsilon_1, \ldots, r_n + \varepsilon_n\} < h\text{-min}\{r_1, \ldots, r_n\}$, which provides the required contradiction.

Turning our attention to item (3), $n > f \geq 1$ implies that $n-1 > f-1 \geq 0$, hence $\boldsymbol{F}_{n-1}^{f-1}(\mathcal{L})$ is accurate. If the discarded interval $\boldsymbol{I}_j$ contributed a left resp. right edge to the f ones skipped by $\mathcal{F}_n^f(\mathcal{I})$, the same left resp. right edge is computed by $\mathcal{F}_{n-1}^{f-1}(\mathcal{L})$ as well. If $\boldsymbol{I}_j$ did not contribute, the same argument as used in the proof of item (1) applies and establishes relation (5). This eventually completes the proof of Lemma 3. □

Remarks:

1. Item (3) of Lemma 3 implies that one should always try to detect and discard faulty intervals before $\mathcal{F}$ is applied, since this can only improve the result. Note that this does not affect validity/applicability of the results of this paper.
2. Comparison[1] with [12, Lem. 3] reveals that $\mathcal{F}$ satisfies the same monotonicity properties as established for the Marzullo function $\mathcal{M}$.

As a prerequisite for defining the Lipschitz condition of an interval-valued function, a suitable metric ("distance function") on intervals needs to be chosen. The following ones were used in [5]:

- The *uniform metric* $\mu(\boldsymbol{U}, \boldsymbol{V})$ that equals the maximum of $|\text{left}(\boldsymbol{U}) - \text{left}(\boldsymbol{V})|$ and $|\text{right}(\boldsymbol{U}) - \text{right}(\boldsymbol{V})|$.
- The *midpoint pseudo-metric* $\overline{\mu}(\boldsymbol{U}, \boldsymbol{V})$ that equals the distance of the midpoints $|\text{center}(\boldsymbol{U}) - \text{center}(\boldsymbol{V})|$ of $\boldsymbol{U}$ and $\boldsymbol{V}$, where center($\boldsymbol{I}$) $=$ (left($\boldsymbol{I}$) $+$ right($\boldsymbol{I}$))$/2$. Note that $\overline{\mu}$ is not a metric, because $\overline{\mu}(\boldsymbol{U}, \boldsymbol{V}) = 0$ does not imply $\boldsymbol{U} = \boldsymbol{V}$.

Note carefully that $\mu(\boldsymbol{U}, \boldsymbol{V}) < \delta$ implies $\overline{\mu}(\boldsymbol{U}, \boldsymbol{V}) < \delta$, since writing left($\boldsymbol{V}$) $=$ left($\boldsymbol{U}$) $+ l$ and right($\boldsymbol{V}$) $=$ right($\boldsymbol{U}$) $+ r$

[1] Note carefully that we used the alternative notation $\mathcal{M}_n^{n-f}$ in [12], which is equal to $\mathcal{M}_n^f$ in this paper.

delivers $|l| < \delta$ and $|r| < \delta$, which leads to $\overline{\mu}(\boldsymbol{U}, \boldsymbol{V}) = |l + r|/2 < \delta$. The converse, however, is not true in general.

We will now show that $\mathcal{F}$ satisfies the Lipschitz condition for the uniform metric.[2]

Lemma 4 (Lipschitz condition for μ). *The FTI intersection function $\mathcal{F}$ satisfies the Lipschitz condition for the uniform metric μ, which means that for any $\delta > 0$ and any two sets $\mathcal{I} = \{\boldsymbol{I}_1, \dots, \boldsymbol{I}_n\}$, $\mathcal{I}' = \{\boldsymbol{I}'_1, \dots, \boldsymbol{I}'_n\}$ of non-empty compatible intervals with at most $f < n$ of those being faulty,*

$$\mu\Big(\mathcal{F}^f_n(\mathcal{I}), \mathcal{F}^f_n(\mathcal{I}')\Big) < \delta \tag{6}$$

provided that $\mu(\boldsymbol{I}_i, \boldsymbol{I}'_i) < \delta$, $1 \le i \le n$.

Proof. Let $\boldsymbol{F} = \mathcal{F}^f_n(\mathcal{I})$ and $\boldsymbol{F}' = \mathcal{F}^f_n(\mathcal{I}')$. Since

$$\text{left}(\boldsymbol{F}) = h\text{-max}\{\text{left}(\boldsymbol{I}_1), \dots, \text{left}(\boldsymbol{I}_n)\}$$
$$\text{right}(\boldsymbol{F}) = h\text{-min}\{\text{right}(\boldsymbol{I}_1), \dots, \text{right}(\boldsymbol{I}_n)\}$$

for $h = f + 1$, we can look separately at the involved left and right edges in our proof. Abbreviating $x_i = \text{left}(\boldsymbol{I}_i)$, $x'_i = \text{left}(\boldsymbol{I}'_i)$ and $y_i = \text{right}(\boldsymbol{I}_i)$, $y'_i = \text{right}(\boldsymbol{I}'_i)$, it boils down to show that

$$|h\text{-max}\{x_1, \dots, x_n\} - h\text{-max}\{x'_1, \dots, x'_n\}| < \delta,$$
$$|h\text{-min}\{y_1, \dots, y_n\} - h\text{-min}\{y'_1, \dots, y'_n\}| < \delta,$$

for any integer $1 \le h \le n$, where $x'_i = x_i + l_i$ and $y'_i = y_i + r_i$. Remember that we assumed $-\delta < l_i, r_i < \delta$, $1 \le i \le n$. To show the "h-max part", consider the following independent cases for $x_j = h\text{-max}\{x_1, \dots, x_n\}$ and $x'_{j'} = h\text{-max}\{x'_1, \dots, x'_n\}$:

1. Regarding x_j either $x'_j = x_j + l_j \le x'_{j'}$ or $\ge x'_{j'}$. The first case immediately provides $x_j - \delta < x_j + l_j \le x'_{j'}$. Otherwise there must be another $x_k \ge x_j$ with $x'_k = x_k + l_k \le x'_{j'}$, which again yields $x_j - \delta < x'_{j'}$. The existence of x_k is warranted, since $|\{x'_i : x'_i \ge x'_{j'}\} \setminus \{x_j + l_j\}| = h - 1 \ge 0$ but $|\{x_i : x_i \ge x_j\}| = h$.
2. Regarding $x'_{j'}$ either $x_{j'} = x'_{j'} - l_{j'} \le x_j$ or $\ge x_j$ The first case immediately provides $x'_{j'} \le x_j + l_{j'} < x_j + \delta$. Otherwise there must be another $x_k \le x_j$ with $x'_k = x_k + l_k \ge x'_{j'}$, which again yields $x'_{j'} < x_j + \delta$. The existence of x_k is warranted, since $|\{x_i : x_i \ge x_j\} \setminus \{x'_{j'} - l_{j'}\}| = h - 1 \ge 0$ but $|\{x'_i : x'_i \ge x'_{j'}\}| = h$.

Combining both cases, we arrive at $x_j - \delta < x'_{j'} < x_j + \delta$ as required. In order to show the "h-min part" we just note that the h-smallest element of a set with cardinality n is the $(n + 1 - h)$-largest one. □

The following Lemma 5 finally shows that $\mathcal{F}$ is optimal in the sense that its result is a lower bound for any proper intersection function $\mathcal{X}$ that satisfies the Lipschitz condition.

[2] Since we will primarily consider the uniform metric in our paper, the phrase "the Lipschitz condition" usually assumes this kind of metric.

Lemma 5 (Optimality). *Let a set $\mathcal{I} = \{\boldsymbol{I}_1, \dots, \boldsymbol{I}_n\}$ of $n > f \ge 0$ non-empty compatible intervals with at most f faulty ones among those be given. Any proper Fault-Tolerant Intersection function $\mathcal{X}^f_n(\mathcal{I})$ that satisfies the Lipschitz condition for the uniform metric fulfils $\mathcal{X}^f_n(\mathcal{I}) \supseteq \mathcal{F}^f_n(\mathcal{I})$.*

Proof. Abbreviating $\boldsymbol{X} = \mathcal{X}^f_n(\mathcal{I})$ and $\boldsymbol{F} = \mathcal{F}^f_n(\mathcal{I})$, let us assume that the statement of our lemma was not true, i.e., that w.l.o.g. $\text{left}(\boldsymbol{X}) > \text{left}(\boldsymbol{F})$. Then, there must be $f'_l > 0$ intervals $\boldsymbol{I}_x \in \mathcal{I}$ with $\boldsymbol{I}_x \cap \boldsymbol{M} = \emptyset$ lying strictly left of $\boldsymbol{M} = \mathcal{M}^f_n(\mathcal{I})$, since otherwise $\text{left}(\boldsymbol{F}) = \text{left}(\boldsymbol{M})$ according to Lemma 1 and Definition 1. This, however, would contradict our hypothesis due to $\mathcal{M}$'s optimality.

Let $\delta = \max_x\{\text{left}(\boldsymbol{M}) - \text{right}(\boldsymbol{I}_x)\}$, which is guaranteed to be non-negative by the above claim, and define $\mathcal{I}'$ to be the set of intervals obtained by extending any left and right edge of the intervals in $\mathcal{I}$ by $\delta/2$. Then, by the definition of $\mathcal{F}$, $\text{left}(\boldsymbol{F}) - \text{left}(\boldsymbol{F}') = \delta/2$ for $\boldsymbol{F}' = \mathcal{F}^f_n(\mathcal{I}')$. Moreover, since $\mu(\boldsymbol{I}_j, \boldsymbol{I}'_j) = \delta/2$, the Lipschitz property of $\mathcal{X}$ guarantees $\mu(\boldsymbol{X}, \boldsymbol{X}') \le \delta/2$ for $\boldsymbol{X}' = \mathcal{X}^f_n(\mathcal{I}')$. However, combining this with our hypothesis implies $\text{left}(\boldsymbol{X}') > \text{left}(\boldsymbol{F}')$ as well.

Now, since obviously $\text{left}(\boldsymbol{M}) - \text{left}(\boldsymbol{M}') \ge \delta/2$ for $\boldsymbol{M}' = \mathcal{M}^f_n(\mathcal{I}')$, we must have $\boldsymbol{M}' \cap \boldsymbol{I}' \ne \emptyset$ for any $\boldsymbol{I}' \in \mathcal{I}'$, which in turn implies $\boldsymbol{F}' = \boldsymbol{M}'$ according to eq. (2). Because $\boldsymbol{X}'$ must certainly include $\boldsymbol{M}'$, we have eventually constructed the required contradiction. An analogous argument can be used to disprove $\text{right}(\boldsymbol{X}) < \text{right}(\boldsymbol{F})$. □

Whereas this optimality result shows that no intersection function satisfying the Lipschitz property can do better, this does not mean that the midpoint of $\mathcal{F}$ always provides the best approximation of the unknown t, recall Section 1. In our deterministic worst-case setting, the optimal choice is the midpoint of the interval provided by the Marzullo function, as it guarantees the smallest maximum distance to t in the worst case. However, the midpoint of an interval that properly contains $\mathcal{M}$ is usually a sub-optimal approximation.

Even worse, it may well be the case that a definitely sub-optimal function like Lamport's A^f, which is just the Fault-Tolerant Average algorithm [6] (i.e., the average of the input values after discarding the f largest and f smallest ones) applied to the midpoints of the input intervals, provides a better approximation of t than $\mathcal{F}$ in some cases. We will show that this is true for any "good" intersection function $\mathcal{X}^1_4$ that satisfies the Lipschitz condition. An intersection function is "good" if it does not compute a sub-optimal result in the case where no input interval can be cast out as obviously faulty, i.e., when item (2) of Lemma 2 applies. Fig. 3 shows the scenario used in our argument, which is easily generalized to arbitrary n and f.

The input intervals have been chosen appropriately such that $\boldsymbol{X} = \mathcal{X}^1_4(\mathcal{I}) = \boldsymbol{M}$ due to $\mathcal{X}$'s "goodness". Replacing $\boldsymbol{I}_4$ by $\boldsymbol{I}'_4$ with $\mu(\boldsymbol{I}_4, \boldsymbol{I}'_4) \le \delta$ in $\boldsymbol{X}' = \mathcal{X}^1_4(\mathcal{I}')$, the Lipschitz property of $\mathcal{X}$ reveals $\mu(\boldsymbol{X}, \boldsymbol{X}') \le \delta$ and hence $\overline{\mu}(\boldsymbol{X}, \boldsymbol{X}') \le \delta$ as well, recall our remark on the introduction of the midpoint pseudo-metric. On the other hand, the left edge of $\mathcal{M}$'s result jumps by $> 2\delta$ upon the transition from $\boldsymbol{M}$ to $\boldsymbol{M}'$, which leads to $\overline{\mu}(\boldsymbol{M}, \boldsymbol{M}') > \delta$. Combining this

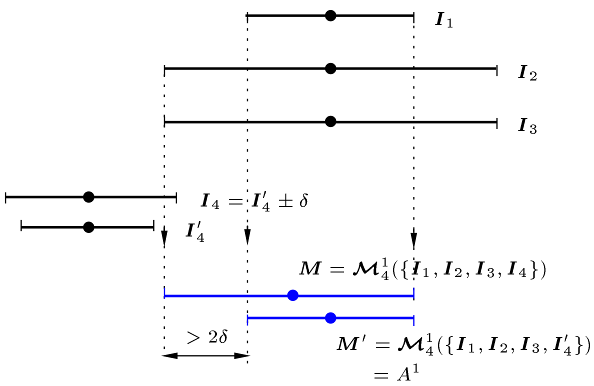

Fig. 3. *Scenario where the approximation of t provided by any "good" Fault-Tolerant Intersection function* $\boldsymbol{\mathcal{X}}_4^1(\{\boldsymbol{I}_1, \boldsymbol{I}_2, \boldsymbol{I}_3, \boldsymbol{I}_4'\})$ *that satisfies the Lipschitz condition is worse than the one provided by* A^1.

with the midpoint pseudo-metric's triangle inequality

$$\begin{aligned}\overline{\mu}(\boldsymbol{M}, \boldsymbol{M}') &\le \overline{\mu}(\boldsymbol{M}, \boldsymbol{X}') + \overline{\mu}(\boldsymbol{X}', \boldsymbol{M}') \\ &\le \overline{\mu}(\boldsymbol{X}, \boldsymbol{X}') + \overline{\mu}(\boldsymbol{X}', \boldsymbol{M}')\end{aligned}$$

eventually yields $0 < \overline{\mu}(\boldsymbol{X}', \boldsymbol{M}')$. Consequently, the midpoint of $\boldsymbol{X}'$ cannot be equal to the one of $\boldsymbol{M}'$. The midpoint provided by A^f, however, is equal to center$(\boldsymbol{M}')$ by construction, and the claimed sub-optimality of $\boldsymbol{\mathcal{X}}$ follows.

We hence conclude that, although we can expect the approximation of t provided by center$(\boldsymbol{\mathcal{F}})$ to surpass A^f in most cases, we cannot demand that it —like any other "good" intersection function satisfying the Lipschitz condition— always outperforms A^f.

3 Worst-Case Analysis in Presence of Faults

In this section, we analyze the worst-case performance of $\boldsymbol{\mathcal{F}}$ according to the framework introduced in [12]. Subsection 3.1 is devoted to the simple case of "local application", where $\boldsymbol{\mathcal{F}}$ is applied to a single input set, say, at a particular node p. In the following Subsection 3.2, we consider the more advanced "distributed application" scenario, where two instances of $\boldsymbol{\mathcal{F}}$ are applied to similar input sets at two different nodes p and q.

3.1 Local Application

In order to reason about the behavior of $\boldsymbol{\mathcal{F}}$ in presence of faults, a fault model is required. Any interval may be faulty due to the following reasons:

Definition 2 (Single Faults). *An interval* $\boldsymbol{I}$ *representing* t *can suffer from the following faults:*

- Omission*:* $\boldsymbol{I} = \emptyset$.
- Non-accurate interval*:* $t \notin \boldsymbol{I}$
- Unbounded accuracy*:* $t \in \boldsymbol{I}$ *but* $|\boldsymbol{I}|$ *too large according to some condition (that need not be known explicitly).*

Note that it is of course easy to recognize and discard an omissive faulty interval, but usually impossible to decide reliably whether an interval $\boldsymbol{I}$ is accurate or not. Masking or detecting —and thus ruling them out completely— unbounded accuracy faults is also difficult in most circumstances.

The following Lemma 6 reveals how $\boldsymbol{\mathcal{F}}$ behaves in presence of faults according to Definition 2. It answers the question of how many non-faulty intervals are required for tolerating at most f_n non-accurate intervals and f_u unbounded accuracy faults. The most important property shown is that $\boldsymbol{\mathcal{F}}$'s result lies within the intersection of $n - 2f_n - 3f_u \ge 1$ non-faulty input intervals.

Lemma 6 (Local Application). *Let* $\boldsymbol{\mathcal{J}} = \{\boldsymbol{J}_1, \dots, \boldsymbol{J}_n\}$ *be a set of* $n \ge 1$ *non-empty compatible accuracy intervals representing* t*, and define* w^h *to be the length of the largest intersection of* $h \ge 1$ *non-faulty intervals among those. If* $f_u' \ge 0$ *of the* $\boldsymbol{J}_j$ *suffer from unbounded accuracy faults and* $f_n' \ge 0$ *are non-accurate, where* $f_u' \le f_u$ *and* $f_n' \le f_n$ *with* $f_u' + f_n' = f' \le f_u + f_n = f < n$ *(so that* $n - f' \ge n - f > 0$ *of the* n *intervals are non-faulty), then:*

(1) $\boldsymbol{F} = \boldsymbol{\mathcal{F}}_n^f(\boldsymbol{\mathcal{J}})$ *is accurate and contains any intersection* $\boldsymbol{W}$ *of* $n - f \ge 1$ *different non-faulty input intervals* $\boldsymbol{J}_{m_1}, \dots, \boldsymbol{J}_{m_{n-f}}$*, i.e.,*

$$\boldsymbol{W} = \bigcap_{j=1}^{n-f} \boldsymbol{J}_{m_j} \subseteq \boldsymbol{F}, \tag{7}$$

so that $|\boldsymbol{F}| \ge w^{n-f}$ *(minimal intersection property).*

(2) There are at least $n - 2f - f_u' \ge n - 2f - f_u$ *different non-faulty input intervals* $\boldsymbol{J}_{b_1}, \dots, \boldsymbol{J}_{b_{n-2f-f_u'}} \in \boldsymbol{\mathcal{J}}$ *such that*

$$\boldsymbol{F} \subseteq \bigcap_{j=1}^{n-2f-f_u'} \boldsymbol{J}_{b_j} \subseteq \bigcap_{j=1}^{n-2f-f_u} \boldsymbol{J}_{b_j'}, \tag{8}$$

where the set of indices $\{b_j'\}_{1 \le j \le n-2f-f_u}$ *is obtained from* $\{b_j\}_{1 \le j \le n-2f-f_u'}$ *by discarding* $f_u - f_u'$ *arbitrary elements. Hence,* $|\boldsymbol{F}| \le w^{n-2f-f_u'} \le w^{n-2f-f_u}$.

(3) There are at least $f - f' + 1 \ge 1$ *non-faulty intervals* $\boldsymbol{J}_{\ell_k}$ *resp.* $\boldsymbol{J}_{r_k}$*,* $1 \le k \le f - f' + 1$*, in* $\boldsymbol{\mathcal{J}}$ *satisfying left*$(\boldsymbol{F}) \le$ *left*$(\boldsymbol{J}_{\ell_k})$ *resp. right*$(\boldsymbol{F}) \ge$ *right*$(\boldsymbol{J}_{r_k})$.

Proof. We first show that $\boldsymbol{F} = \boldsymbol{\mathcal{F}}_n^f(\boldsymbol{\mathcal{J}})$ contains any intersection of at least $n - f$ input intervals: By $\boldsymbol{\mathcal{F}}$'s definition, we have at most $n - (f + 1)$ intervals with left edge strictly smaller than left$(\boldsymbol{F})$. Therefore, assuming an intersection of $n - f$ intervals strictly left of left$(\boldsymbol{F})$ immediately leads to a contradiction. An analogous argument can be applied to the right edges. Finally, since inclusion of any intersection of at least $n - f$ intervals implies inclusion of any such intersection made up of non-faulty intervals only, it follows that $t \in \boldsymbol{F}$ and $|\boldsymbol{F}| \ge w^{n-f}$ as asserted in item (1) of the lemma.

Turning our attention to item (2), it is apparent that at least

$$g_l' + g_r' \ge 2n - 2f - 2f_u' - f_n' \tag{9}$$

non-faulty input intervals[3] must have a left edge left or equal to left($\boldsymbol{F}$) as well as a right edge right or equal to right($\boldsymbol{F}$). This is due to the fact that, apart from the $2f$ intervals contributing the f largest left edges and the f smallest right edges (which cannot have this property by Definition 1), there may be still up to f'_u intervals with unbounded accuracy faults that could have edges both left of left($\boldsymbol{F}$) and right of right($\boldsymbol{F}$). They must hence be subtracted twice in eq. (9). Similarly, there may also be up to f'_n non-accurate intervals, which must be subtracted only once since any such interval $\boldsymbol{J}_j$ could satisfy either left($\boldsymbol{J}_j$) $<$ left($\boldsymbol{F}$) or else right($\boldsymbol{J}_j$) $>$ right($\boldsymbol{F}$) but not both, due to to $t \notin \boldsymbol{J}_j$ but $t \in \boldsymbol{F}$.

However, since there are only $g' = n - f'$ different non-faulty intervals in the input set $\mathcal{J} = \{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_n\}$, the pigeonhole principle reveals that

$$\begin{aligned} g'_l + g'_r - g' &\geq 2n - 2f - 2f'_u - f'_n - n + f' \\ &\geq n - 2f - f'_u \end{aligned}$$

of the intervals counted in eq. (9), say $\boldsymbol{J}_{b_1}, \ldots, \boldsymbol{J}_{b_{n-2f-f'_u}}$, must be the same. Therefore, $\boldsymbol{F}$ must lie in the intersection of those intervals and $|\boldsymbol{F}| \leq w^{n-2f-f'_u}$ as asserted. The upper bound in eq. (8) follows immediately from $f'_u \leq f_u$.

Finally, item (3) of our lemma follows directly from $\mathcal{F}$'s Definition 1 in conjunction with the fact that at least $f - f'$ of the discarded left edges (and analogously for the right edges) must belong to non-faulty intervals. This eventually completes the proof of Lemma 6. □

Remarks:

1. We excluded omission faults in our lemma, since $\mathcal{F}$ as defined in Definition 1 cannot deal with empty intervals. However, intervals with omission faults can of course be discarded before $\mathcal{F}$ is applied. Therefore, if f'_o of presumed n intervals suffer from an omission fault, we just have to set $n := n - f'_o$ and $f := f - f'_o$ in Lemma 6 to obtain the results for this case as well. Note that it is feasible to let f depend on f'_o, see Lemma 7.
2. Interpreting item (2) of Lemma 6 and the previous remark in terms of the usual fault-tolerance degree notion, it follows that $n \geq f'_o + 2f + f'_u + 1$ nodes are required to guarantee that $\boldsymbol{F}$ remains bounded by the length of at least one non-faulty input interval. Hence, as many as

$f'_o + 1$	for f'_o omission faults,
$2f_n + 1$	for $f'_n \leq f_n$ non-accurate faults,
$2f_u + f'_u + 1$	for $f'_u \leq f_u$ unbounded accuracy faults

   nodes are required for tolerating faults of the given type. It is thus apparent that $\mathcal{F}$ can tolerate $\lfloor (n-1)/2 \rfloor$ non-accurate intervals but only $\lfloor (n-1)/3 \rfloor$ intervals that suffer from unbounded accuracy faults. Note carefully that the numbers above do not solely depend on the *actual* number of faults (e.g., f'_u), but also on their maximum number (e.g., f_u); this is due to the fact that the latter is compiled into the superscript argument of $\mathcal{F}$.
3. The lower bound on $|\boldsymbol{F}|$ in item (1) expresses the rather obvious fact that $\mathcal{F}$ cannot improve the accuracy beyond the one "hidden" in the input intervals; the term *minimal intersection property* was coined in [7]. Note that $\boldsymbol{F}$ contains any intersection of $n-f$ intervals, hence includes intersections involving unbounded accuracy faults as well.
4. Item (3) just says that $\boldsymbol{F}$ contains the left and right edge of at least one (not necessarily the same) non-faulty interval.
5. Comparison[2] of Lemma 6 and [12, Lem. 2] reveals that $\mathcal{F}$ has literally the same worst-case performance as the optimal Marzullo function $\mathcal{M}$. This means that both functions produce the same result for worst-case scenarios. Of course, for "average" input sets, $\mathcal{F}$ will usually provide a slightly larger interval.

[3] Note that we do not count different intervals here, but rather intervals according to the total number of edges.

3.2 Distributed Application

In this section, we will consider the case where $\mathcal{F}$ is applied to (similar) input sets $\mathcal{I}_p$, $\mathcal{I}_q$ at two different nodes. Those sets could be produced by a remote clock reading algorithm or replicated sensors, for example. It will turn out that the respective outcomes $\boldsymbol{F}_p = \mathcal{F}^f_n(\mathcal{I}_p)$ and $\boldsymbol{F}_q = \mathcal{F}^f_n(\mathcal{I}_q)$ cannot deviate too much from each other, even if faults lead to quite different input sets. Note carefully, however, that Lemma 4 does not help here, since exploiting the Lipschitz condition would require $\mu(\boldsymbol{I}^i_p, \boldsymbol{I}^i_q) < \delta$ for any $1 \leq i \leq n$. This requirement cannot be guaranteed when faults cause the input sets to differ at node p and q.

Of course, one might consider to employ a consensus protocol prior to $\mathcal{F}$'s application for alleviating such inconsistencies. This is expensive, though, since only complete agreement upon the set of faulty/non-faulty senders would render Lemma 4 applicable. Lemma 2 reveals that using a binary decision value v_j, meaning "$\boldsymbol{I}^s_j \cap \mathcal{M}(\mathcal{I}_j)$ empty/non-empty", as an input to the s-th instance, $1 \leq s \leq n$, of a consensus protocol would lead to consistent input sets $\mathcal{I}'_j$ that even guarantee $\mathcal{F}(\mathcal{I}'_j) = \mathcal{M}(\mathcal{I}'_j)$ for all non-faulty nodes j. Less costly (approximate) agreement protocols, however, are difficult to apply in our context for the reasons explained below.

In fact, any distributed application of $\mathcal{F}$ is considerably complicated by the fact that we cannot always assume that the information disseminated by a single sender s leads to the same interval at two receivers p and q, even if there is no fault at all. More specifically, in typical clock synchronization applications, it is not a constant-valued interval that is disseminated by s to p and q, but rather a time-dependent one. Any time-dependent quantity, however, is affected by transmission delays, clock granularities and related effects. As a consequence, p and q may not only receive slightly different information from non-faulty senders, but also perceive faults differently: An interval from sender s may be correct at p but faulty at q, both due to faults occurring at the sending and the receiving side. This implies that approximate agreement protocols are of limited use for alleviating inconsistencies (although part of our current research indicates some potential for improvement).

In order to be able to reason about faults in distributed applications, the single-interval faults of Definition 2 are complemented by faults of *pairs of intervals* $\boldsymbol{I}^s_p \in \mathcal{I}_p$ resp. $\boldsymbol{I}^s_q \in \mathcal{I}_q$ obtained at nodes p resp. q. This will lead to a *perception-based fault model* as introduced in [12], where the usual omniscient (= global) perception of faults is replaced by the local perceptions of any two non-faulty nodes in the system. This way, both node and link faults can be accurately modeled.

We therefore assume that the intervals in both input sets can be uniquely grouped as n pairs $\{\boldsymbol{I}_p^s \in \boldsymbol{\mathcal{I}}_p, \boldsymbol{I}_q^s \in \boldsymbol{\mathcal{I}}_q\}$ originating in the same source of information s, $1 \leq s \leq n$. We will use the term *ordered sets* for $\boldsymbol{\mathcal{I}}_p$ and $\boldsymbol{\mathcal{I}}_q$ to indicate this property. The corresponding intervals in two ordered sets need not be the same, although they should be reasonably similar. Definition 3 exhaustively specifies all possible faults of pairs of intervals:

Definition 3 (Pairwise Faults). *A pair of compatible accuracy intervals $\{\boldsymbol{I}_p^s, \boldsymbol{I}_q^s\}$ representing t suffers from*

- *a* crash fault *iff* $\boldsymbol{I}_p^s = \boldsymbol{I}_q^s = \emptyset$,
- *a* symmetric fault *iff either*
  (1) both $\boldsymbol{I}_p^s$ and $\boldsymbol{I}_q^s$ are not accurate in the sense of $t <$ left$(\boldsymbol{I}_p^s)$ and $t <$ left$(\boldsymbol{I}_q^s)$, or else $t >$ right$(\boldsymbol{I}_p^s)$ and $t >$ right$(\boldsymbol{I}_q^s)$,
  (2) without loss of generality, $\boldsymbol{I}_p^s = \emptyset$ and $\boldsymbol{I}_q^s \neq \emptyset$ does not suffer from an unbounded accuracy fault.
- *an* asymmetric fault *iff either*
  (1) both $\boldsymbol{I}_p^s$ and $\boldsymbol{I}_q^s$ are not accurate in the sense of $t >$ right$(\boldsymbol{I}_p^s)$ and $t <$ left$(\boldsymbol{I}_q^s)$ or else $t >$ right$(\boldsymbol{I}_q^s)$ and $t <$ left$(\boldsymbol{I}_p^s)$ (true Byzantine fault),
  (2) without loss of generality, $\boldsymbol{I}_p^s \neq \emptyset$ is faulty and $\boldsymbol{I}_q^s$ is arbitrary (and none of the other faults applies).

Remarks:

1. The "classical" asymmetric fault [17] is caused by disseminating information that is perceived differently at p and q. In our special context, it is characterized by the fact that node p arrives at the conclusion that the interval $\boldsymbol{I}_p^s$ from sender s is, say, strictly left of the sought value t, whereas q thinks that $\boldsymbol{I}_p^s$ is strictly right of t (or correct). This situation usually also occurs in presence of an unbounded accuracy fault.
2. The "classical" symmetric fault [17] is caused by disseminating information that is perceived identically at p and q. In our special context, both p and q must arrive at the same conclusion on whether the intervals from sender s are both left or right of t. Alternatively, one of the intervals may be missing due to a receive omission.
3. A crash fault causes an omission both at node p and q. Note carefully, though, that it is impossible for either node to decide locally (without further information) whether its omission is due to a crash fault or a more severe receive omission.
4. Note that Definition 3 does not cover the case where a more severe fault comes out as a less severe one. For example, it is reasonable to assume that an asymmetric fault could just be a symmetric or even a crash fault only. In this paper, we will typically use phrases like "asymmetric (or weaker) fault" to indicate such extensions.

Introducing different classes of faults as in Definition 3 is known as a *hybrid fault model* in literature, cf. [1], [17]. It allows us to exploit the fact that masking f symmetric faults requires only $n \geq 2f + 1$ nodes, whereas $n \geq 3f + 1$ are needed if all faults are asymmetric ones. Since a large number of asymmetric faults is very unlikely in practice, cf. [11], this effectively leads to a smaller n for tolerating a given number of faults.

We should explicitly mention, though, that our definition of symmetric and asymmetric faults extends and, in some cases, apparently contradicts the "classical" meaning of those terms. Still, we think that their usage is legitimate due to the fact that our extension preserves the essentials of their meaning: The meaning of symmetric / asymmetric fault is basically received identically / not identically at different nodes. In our context, however, we have to relax the meaning of "received identically" since we cannot assume identical information at different nodes even in the faultless case, as explained earlier. We also have to accept the fact that the interval-based paradigm introduces unbounded accuracy faults, which are not known in traditional settings but can create an asymmetric perception.

The following Lemma 7 gives the number of non-faulty pairs of intervals required by $\boldsymbol{\mathcal{F}}$ for tolerating a certain number of

- crash faults ($f_c' \leq f_c$),
- symmetric faults ($f_s' \leq f_s$),
- asymmetric faults ($f_a' \leq f_a$).

The most important result is an upper bound on the nc-union $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$, which must lie within at least $n - \min\{f_c' + f_s', 2f_c - f_c'\} - 2f_s - 3f_a \geq 1$ nc-unions $\boldsymbol{I}_p^s \sqcup \boldsymbol{I}_q^s$ of non-faulty input intervals. Note that using (nc-)unions in our lemma takes into account that two different nodes p and q may have slightly different input sets, even if there is no fault.

Lemma 7 (Distributed Application). *Let $\boldsymbol{\mathcal{I}}_p = \{\boldsymbol{I}_p^1, \dots, \boldsymbol{I}_p^n\}$ and $\boldsymbol{\mathcal{I}}_q = \{\boldsymbol{I}_q^1, \dots, \boldsymbol{I}_q^n\}$ be two ordered sets of $n > f_c + f_s + f_a$, $f_c, f_s, f_a \geq 0$, compatible (or empty) accuracy intervals representing t, where $f_a' \leq f_a$, $f_s' \leq f_s$, and $f_c' \leq f_c$ of the n pairs of intervals $\{\boldsymbol{I}_p^i, \boldsymbol{I}_q^i\}$ exhibit asymmetric, symmetric, and crash faults, respectively, and the remaining ones are non-faulty. Define u^h resp. v^h to be the length of the largest intersection of $h \geq 1$ nc-unions resp. intersections of pairs of non-faulty intervals, formally $u^h = \max\{|\boldsymbol{U}| : \boldsymbol{U} \in \boldsymbol{\mathcal{U}}_{pq}^h\}$ and $v^h = \max\{|\boldsymbol{V}| : \boldsymbol{V} \in \boldsymbol{\mathcal{V}}_{pq}^h\}$ for*

$$\boldsymbol{\mathcal{U}}_{pq}^h = \Big\{\boldsymbol{U} : \boldsymbol{U} = \bigcap_{i=1}^{h} \boldsymbol{I}_p^{u_i} \sqcup \boldsymbol{I}_q^{u_i} \text{ with } u_i \neq u_j,\, i \neq j, \text{ and } \boldsymbol{I}_p^{u_i} \in \boldsymbol{\mathcal{I}}_p,\, \boldsymbol{I}_q^{u_i} \in \boldsymbol{\mathcal{I}}_q \text{ being non-faulty}\Big\}$$

$$\boldsymbol{\mathcal{V}}_{pq}^h = \Big\{\boldsymbol{V} : \boldsymbol{V} = \bigcap_{i=1}^{h} \boldsymbol{I}_p^{v_i} \cap \boldsymbol{I}_q^{v_i} \text{ with } v_i \neq v_j,\, i \neq j, \text{ and } \boldsymbol{I}_p^{v_i} \in \boldsymbol{\mathcal{I}}_p,\, \boldsymbol{I}_q^{v_i} \in \boldsymbol{\mathcal{I}}_q \text{ being non-faulty}\Big\}.$$

Let d_p', $0 \leq d_p' \leq f_s'$, resp. e_p', $0 \leq e_p' \leq f_a'$, denote the (unknown) number of empty intervals caused by symmetric resp. asymmetric faults at node p, and $\boldsymbol{\mathcal{J}}_p = \{\boldsymbol{J}_1, \dots, \boldsymbol{J}_{n_p}\}$ be the set of $n_p = n - o_p$ non-empty intervals obtained from $\boldsymbol{\mathcal{I}}_p$ by discarding any of the (known) $o_p = f_c' + d_p' + e_p' \leq f_c + f_s + f_a$ empty intervals caused by crash and symmetric/asymmetric faults. Using the upper bound $f_p = f_s + f_a - \max\{0, o_p - f_c\}$ on the number of intervals in $\boldsymbol{J}_p$ that (still) *may be faulty in presence of o_p omissions, define*

$$\boldsymbol{F}_p = \boldsymbol{\mathcal{F}}_{n_p}^{f_p}(\boldsymbol{\mathcal{J}}_p)$$
$$\boldsymbol{F}_q = \boldsymbol{\mathcal{F}}_{n_q}^{f_q}(\boldsymbol{\mathcal{J}}_q).$$

Then,

(1) both $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ are accurate and

$$\boldsymbol{F}_p \cap \boldsymbol{F}_q \supseteq \bigcap_{j=1}^{n-f_c'-f_s-f_a} \boldsymbol{I}_p^{v_j} \cap \boldsymbol{I}_q^{v_j} = \boldsymbol{V} \tag{10}$$

for any subset $\boldsymbol{V} \in \boldsymbol{\mathcal{V}}_{pq}^{n-f_c'-f_s-f_a}$, so that $|\boldsymbol{F}_p \cap \boldsymbol{F}_q| \geq v^{n-f_c'-f_s-f_a}$ (distributed minimal intersection property),

(2) there are at least $n - \min\{f_c' + f_s', 2f_c - f_c'\} - 2f_s - 2f_a - f_a'$ pairs of non-faulty intervals $\{\boldsymbol{I}_p^{u_k}, \boldsymbol{I}_q^{u_k}\}$ with $\boldsymbol{I}_p^{u_k} \in \boldsymbol{\mathcal{J}}_p$ and $\boldsymbol{I}_q^{u_k} \in \boldsymbol{\mathcal{J}}_q$ such that $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$ is contained ($\subseteq$) in

$$\bigcap_{k=1}^{n-\min\{f_c'+f_s',2f_c-f_c'\}-2f_s-2f_a-f_a'} \boldsymbol{I}_p^{u_k} \sqcup \boldsymbol{I}_q^{u_k} \tag{11}$$

and hence

$$|\boldsymbol{F}_p \sqcup \boldsymbol{F}_q| \leq u^{n-\min\{f_c'+f_s',2f_c-f_c'\}-2f_s-2f_a-f_a'}.$$

Proof. First of all, we note that f_p gives indeed an upper bound on the number of intervals in $\boldsymbol{\mathcal{J}}_p$ that still may be faulty in presence of $o_p = f_c' + d_p' + e_p' \leq f_c' + f_s' + f_a' \leq f_c + f_s + f_a$ omissions, since $f_p = f_s + f_a$ if $o_p \leq f_c$, and $f_p = f_s + f_a - (o_p - f_c)$ otherwise (accounting for $o_p - f_c > 0$ symmetric/asymmetric faults that must have caused omissions at node p), hence

$$f_p \leq f_s + f_a. \tag{12}$$

Evidently, at least $n_p - f_p$ of the intervals in $\boldsymbol{\mathcal{J}}_p$ must be non-faulty. Rewriting the definition

$$\begin{aligned} n_p - f_p &= n - o_p - f_s - f_a + \max\{0, o_p - f_c\} \\ &= n - f_s - f_a + \max\{-o_p, -f_c\} \end{aligned} \tag{13}$$

and applying $\max\{0, x\} \geq x$ for any x, and the simple fact that $\max\{-o_p, -f_c\} \leq -f_c'$ since obviously $o_p \geq f_c'$ and $f_c' \leq f_c$, it follows easily that

$$\begin{aligned} n - f_c - f_s - f_a \leq n_p - f_p &\leq n - f_c' - f_s - f_a \\ &\leq n - f_s - f_a. \end{aligned}$$

Of course, analogous bounds hold for $n_q - f_q$.

Lemma 6 is applicable, and it follows that $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ are both accurate and satisfy the (local) minimal intersection property. That is, $\boldsymbol{F}_p$ contains any intersection of $n_p - f_p \leq n - f_c' - f_s - f_a$ non-faulty intervals present in $\boldsymbol{\mathcal{J}}_p$. If $\{v_j\}_{1 \leq j \leq n-f_c'-f_s-f_a}$ denotes any set of different indices of non-faulty pairs of intervals $\{\boldsymbol{I}_p^{v_j} \in \boldsymbol{\mathcal{I}}_p, \boldsymbol{I}_q^{v_j} \in \boldsymbol{\mathcal{I}}_q\}$ (of course also present in $\boldsymbol{\mathcal{J}}_p, \boldsymbol{\mathcal{J}}_q$), we thus have

$$\boldsymbol{W}_p = \bigcap_{j=1}^{n-f_c'-f_s-f_a} \boldsymbol{I}_p^{v_j} \subseteq \bigcap_{j=1}^{n_p-f_p} \boldsymbol{I}_p^{v_j} \subseteq \boldsymbol{F}_p$$

and, for the same set $\{v_j\}$, $\boldsymbol{W}_q = \bigcap_{j=1}^{n-f_c'-f_s-f_a} \boldsymbol{I}_q^{v_j} \subseteq \boldsymbol{F}_q$. By elementary set algebra, it thus follows that $\boldsymbol{V} = \boldsymbol{W}_p \cap \boldsymbol{W}_q \in \boldsymbol{\mathcal{V}}^{n-f_c'-f_s-f_a}$ satisfies eq. (10). Finally, $|\boldsymbol{F}_p \cap \boldsymbol{F}_q| \geq v^{n-f_c'-f_s-f_a}$ is a simple consequence of the definition of v^h as the maximum length of $\boldsymbol{V} \in \boldsymbol{\mathcal{V}}_{pq}^h$. This completes the proof of item (1).

For item (2), suppose that $g_{p,l}$ intervals belonging to a non-faulty pair of input intervals have a left edge smaller or equal than left($\boldsymbol{F}_p$), whereas $g_{q,r}$ intervals belonging to a non-faulty pair of input intervals have a right edge larger or equal than right($\boldsymbol{F}_q$). We must have

$$\begin{aligned} g_{p,l} &\geq n_p - f_p - (f_a' - e_p') - (s_{\text{left}}' - d_{p,\text{left}}') \\ &\geq n - f_s - f_a + \max\{-o_p, -f_c\} \\ &\qquad -f_a' - s_{\text{left}}' + d_{p,\text{left}}' + e_p' \\ g_{q,r} &\geq n_q - f_q - (f_a' - e_q') - (s_{\text{right}}' - d_{q,\text{right}}') \\ &\geq n - f_s - f_a + \max\{-o_q, -f_c\} \\ &\qquad -f_a' - s_{\text{right}}' + d_{q,\text{right}}' + e_q', \end{aligned}$$

where $s_{\text{left}}' + s_{\text{right}}' = f_s' \leq f_s$ are the number of symmetrically faulty pairs of intervals lying left resp. right of t, and $d_{p,\text{left}}' + d_{p,\text{right}}' = d_p'$, $d_{q,\text{left}}' + d_{q,\text{right}}' = d_q'$ denote the number of omissions among those at node p resp. q; the lower bounds follow immediately from (13).

However, we only have $g = n - f_c' - f_s' - f_a'$ different non-faulty pairs of intervals. Thus, the pigeonhole principle reveals that at least $Y = g_{p,l} + g_{q,r} - g$ given by

$$\begin{aligned} Y &\geq 2n + \max\{-o_p, -f_c\} + \max\{-o_q, -f_c\} \\ &\qquad -2f_s - 2f_a - 2f_a' - f_s' \\ &\qquad +d_{p,\text{left}}' + d_{q,\text{right}}' + e_p' + e_q' \\ &\qquad -n + f_c' + f_s' + f_a' \\ &\geq n + \max\{-f_c' - d_{p,\text{right}}', -f_c + d_{p,\text{left}}' + e_p'\} \\ &\qquad + \max\{-f_c' - d_{q,\text{left}}', -f_c + d_{q,\text{right}}' + e_q'\} \\ &\qquad +f_c' - 2f_s - 2f_a - f_a' \\ &\geq n + \max\{-2f_c' - f_s', -2f_c\} \\ &\qquad +f_c' - 2f_s - 2f_a - f_a' \\ &\geq n - \min\{f_c' + f_s', 2f_c - f_c'\} - 2f_s - 2f_a - f_a' \end{aligned}$$

of those must be the same. Abbreviating $\mu = \min\{f_c' + f_s', 2f_c - f_c'\}$, we can conclude that there are at least $n - \mu - 2f_s - 2f_a - f_a'$ pairs of accurate intervals, say, $\boldsymbol{I}_p^{b_1} \sqcup \boldsymbol{I}_q^{b_1}, \ldots, \boldsymbol{I}_p^{b_{n-\mu-2f_s-2f_a-f_a'}} \sqcup \boldsymbol{I}_q^{b_{n-\mu-2f_s-2f_a-f_a'}}$ with $\boldsymbol{I}_p^{b_i} \in \boldsymbol{\mathcal{J}}_p$ and $\boldsymbol{I}_q^{b_i} \in \boldsymbol{\mathcal{J}}_q$ such that $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$ is contained ($\subseteq$) in

$$\bigcap_{j=1}^{n-\mu-2f_s-2f_a-f_a'} \boldsymbol{I}_p^{b_j} \sqcup \boldsymbol{I}_q^{b_j} \in \boldsymbol{\mathcal{U}}_{pq}^{n-\mu-2f_s-2f_a-f_a'}, \tag{14}$$

which proves eq. (11). To complete the proof of Lemma 7, it only remains to justify $|\boldsymbol{F}_p \sqcup \boldsymbol{F}_q| \leq u^{n-\mu-2f_s-2f_a-f_a'}$, which is a simple consequence of the definition of u^h as the maximum length of $\boldsymbol{U} \in \boldsymbol{\mathcal{U}}_{pq}^h$. $\square$

Remarks:

1. Note carefully that Lemma 6 could also be used to deduce a "distributed application"-related result: Since $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ are both accurate and hence contain t, it follows from item (2) that $|\boldsymbol{F}_p \cup \boldsymbol{F}_q| \leq 2w^{n-2f-f_u}$. However, comparison with item (2) of Lemma 7 reveals that this result is essentially twice as large.
2. Our crash faults are more severe than the (system-wide consistently perceived) *benign faults* of [17], since it cannot be decided locally whether an omissive interval belongs to a crash fault or to an (inconsistent) receive omission. However, it is of course possible to "merge" crash and symmetric faults, in the sense that the former are counted in f'_s (resp. f_s) and $f'_c = f_c = 0$ (note that $n_p - f_p = n - f_s - f_a$ in this case). After all, we already accounted for symmetric/asymmetric faults involving empty intervals in the proof of Lemma 7.
3. Interpreting the accomplishments of Lemma 7 and the previous remark in terms of the usual fault-tolerance degree notion, it turns out that $n \geq \min\{f'_c + f'_s, 2f_c - f'_c\} + 2f_s + 2f_a + f'_a + 1$ nodes are required to guarantee that $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$ remains bounded by the length of the nc-union of at least one pair of non-faulty input intervals. Hence, as much as

   $\min\{f'_c + f'_s, 2f_c - f'_c\} + 1$ for f'_c crash faults,
   $2f_s + 1$ for $f'_s \leq f_s$ symmetric faults,
   $2f_a + f'_a + 1$ for $f'_a \leq f_a$ asymmetric faults

   nodes are required for tolerating faults of the given type.
4. It should be clear from the proof of Lemma 7 that the property that really pins down symmetric faults is the following one: If a symmetrically faulty interval $\boldsymbol{I}^s_q$ satisfies right$(\boldsymbol{I}^s_q) \geq$ right$(\boldsymbol{F}_q)$ (correctly accounted for in s_{right}), then its corresponding $\boldsymbol{I}^s_p$ must not have left$(\boldsymbol{I}^s_p) \leq$ left$(\boldsymbol{F}_p)$ (since it is not accounted for in s_{left}). This is the reason why $\boldsymbol{I}^s_p \neq 0$ being faulty and $\boldsymbol{I}^s_q \neq 0$ being non-faulty must be counted as an asymmetric fault in item (2) of Definition 3.
5. Comparison[2] of Lemma 7 with [12, Lem. 4] again reveals that $\boldsymbol{\mathcal{F}}$ has exactly the same worst-case performance as the optimal Marzullo function $\boldsymbol{\mathcal{M}}$.
6. The proof of Lemma 7 reveals the ultimate reason for using nc-unions $\sqcup$ instead of $\cup$ in the statement of item (2): It may be the case that, say, $\boldsymbol{F}_q \subseteq \boldsymbol{F}_p$, such that $\boldsymbol{F}_p$ would determine both left and right edge of $\boldsymbol{F}_p \cup \boldsymbol{F}_q$. By applying Lemma 6 with $n := n_p$, $f := f_p$, and $f'_u \leq f'_a$ (as well as $f_u \leq f_a$), we could show that there are at least $n_p - 2f_p - f'_u \geq n - \mu - 2f_s - 2f_a - f'_a$ non-faulty intervals $\boldsymbol{I}^{b_j}_p$ in $\boldsymbol{\mathcal{J}}_p$ the intersection of which majorizes $\boldsymbol{F}_p$. This does not imply, however, that all of those intervals appear in $\boldsymbol{\mathcal{J}}_q$ as well — just think of symmetric faults appearing non-faulty at p but omissive at q. Hence, we cannot claim that all the unions $\boldsymbol{I}^{b_j}_p \cup \boldsymbol{I}^{b_j}_q$ —the intersection of which would of course majorize $\boldsymbol{F}_q \cup \boldsymbol{F}_p$— involve non-fauly intervals only. Clearly, focussing upon $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q \subseteq \boldsymbol{F}_p \cup \boldsymbol{F}_q$ entirely avoids this difficulty.

The following lemma shows that Lemma 7 remains valid if a more severe fault comes out as a less severe one, and shows what happens if certain fault assumptions are violated. Note that crash faults are counted as symmetric ones here for simplicity.

Lemma 8 (Graceful Degradation). *Let $\boldsymbol{\mathcal{I}}_p = \{\boldsymbol{I}^1_p, \ldots, \boldsymbol{I}^n_p\}$ and $\boldsymbol{\mathcal{I}}_q = \{\boldsymbol{I}^1_q, \ldots, \boldsymbol{I}^n_q\}$ be two ordered sets of $n > f_s + f_a$, $f_s, f_a \geq 0$, compatible (or empty) accuracy intervals representing t, where $f'_s \leq f_s$ of the n pairs of intervals $\{\boldsymbol{I}^i_p, \boldsymbol{I}^i_q\}$ exhibit symmetric (or weaker) faults, $f'_a \leq f_a$ exhibit asymmetric (or weaker) faults, and the remaining ones are non-faulty. As in Lemma 7, define u^h resp. v^h to be the length of the largest intersection of $h \geq 1$ nc-unions ($\in \boldsymbol{\mathcal{U}}^h_{pq}$) resp. intersections ($\in \boldsymbol{\mathcal{V}}^h_{pq}$) of pairs of non-faulty intervals.*

Let $\boldsymbol{\mathcal{J}}_p = \{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_{n_p}\}$ be the set of $n_p = n - o_p$ non-empty intervals obtained from $\boldsymbol{\mathcal{I}}_p$ by discarding any of the o_p empty intervals caused by omissions. Using the upper bound $f_p = f_s + f_a - o_p$ on the number of intervals in $\boldsymbol{J}_p$ that (still) may be faulty in presence of o_p omissions, define

$$\boldsymbol{F}_p = \boldsymbol{\mathcal{F}}^{f_p}_{n_p}(\boldsymbol{\mathcal{J}}_p)$$
$$\boldsymbol{F}_q = \boldsymbol{\mathcal{F}}^{f_q}_{n_q}(\boldsymbol{\mathcal{J}}_q).$$

Then:

(1) Both $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ are accurate and

$$\boldsymbol{F}_p \cap \boldsymbol{F}_q \supseteq \bigcap_{j=1}^{n-f_s-f_a} \boldsymbol{I}^{v_j}_p \cap \boldsymbol{I}^{v_j}_q = \boldsymbol{V} \tag{15}$$

for any possible subset $\boldsymbol{V} \in \boldsymbol{\mathcal{V}}^{n-f_s-f_a}_{pq}$, so that $|\boldsymbol{F}_p \cap \boldsymbol{F}_q| \geq v^{n-f_s-f_a}$ (distributed minimal intersection property).

(2) There are at least $n - 2f_s - 2f_a - f'_a \geq n - 2f_s - 3f_a$ pairs of non-faulty intervals $\{\boldsymbol{I}^{u_k}_p, \boldsymbol{I}^{u_k}_q\}$ with $\boldsymbol{I}^{u_k}_p \in \boldsymbol{\mathcal{J}}_p$ and $\boldsymbol{I}^{u_k}_q \in \boldsymbol{\mathcal{J}}_q$ such that

$$\boldsymbol{F}_p \sqcup \boldsymbol{F}_q \subseteq \bigcap_{k=1}^{n-2f_s-2f_a-f'_a} \boldsymbol{I}^{u_k}_p \sqcup \boldsymbol{I}^{u_k}_q$$
$$\subseteq \bigcap_{k=1}^{n-2f_s-3f_a} \boldsymbol{I}^{u'_k}_p \sqcup \boldsymbol{I}^{u'_k}_q, \tag{16}$$

where the sequence $\{u'_k\}_{1 \leq k \leq n-2f_s-3f_a}$ is obtained from $\{u_k\}_{1 \leq k \leq n-2f_s-2f_a-f'_a}$ by discarding $f_a - f'_a$ arbitrary elements. Therefore, $|\boldsymbol{F}_p \sqcup \boldsymbol{F}_q| \leq u^{n-2f_s-2f_a-f'_a} \leq u^{n-2f_s-3f_a}$.

(3) Assume that the fault model is violated in the sense that $f' = f'_s + f'_a > f_s + f_a$ but still $n \geq 2f' + f'_u + 1$, where $f'_u \leq f'_a$ denotes the number of pairs of intervals that involve unbounded accuracy faults. If the violation of the fault model is not obvious, in the sense that $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ can be computed and are not empty due to $L > R$ in Definition 1, then there are $n - 2f' - f'_u$ non-faulty intervals $\boldsymbol{I}^{p_1}_p, \ldots, \boldsymbol{I}^{p_{n-2f'-f'_u}}_p$ in $\boldsymbol{\mathcal{J}}_p$ and $n - 2f' - f'_u$ non-faulty intervals $\boldsymbol{I}^{q_1}_q, \ldots, \boldsymbol{I}^{q_{n-2f'-f'_u}}_q$ in $\boldsymbol{\mathcal{J}}_q$ such that $\boldsymbol{F}_p \cup \boldsymbol{F}_q$ (and hence $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$) is contained ($\subseteq$) in

$$\left(\bigcap_{j=1}^{n-2f'-f'_u} \boldsymbol{I}^{p_j}_p\right) \cup \left(\bigcap_{j=1}^{n-2f'-f'_u} \boldsymbol{I}^{q_j}_q\right). \tag{17}$$

Hence, $|\boldsymbol{F}_p \cup \boldsymbol{F}_q| \leq w_p^{n-2f'-f'_u} + w_q^{n-2f'-f'_u}$, where w^h_p resp. w^h_q denote the length of the largest intersection of h accurate intervals in $\boldsymbol{\mathcal{I}}_p$ resp. $\boldsymbol{\mathcal{I}}_q$.

Nevertheless, $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ are not necessarily accurate and possibly not even intersecting; accurateness is guaranteed, however, if $f' \leq f_s + f_a$ but all f' faults are asymmetric ones.

Proof. Since crash faults are now considered as symmetric ones and hence accounted for in f'_s and f_s, see Remark 2 on Lemma 7, items (1) and (2) follow directly from adopting the results of Lemma 7 to $f'_c = f_c = 0$. Note that $n_p - f_p = n - f_s - f_a$ here. To confirm the assertions for asymmetric faults appearing as weaker ones, just consider the expressions supplied by Lemma 7 when temporarily setting $f_a := f_a - 1$ and $f_s := f_s + 1$.

To show item (3), we first note that we only have to consider the case where $f_s + f_a - o_p \geq 0$, since otherwise there would have been too many omissions to compute $\boldsymbol{F}_p$. Moreover, recalling that we assumed $\boldsymbol{F}_p \neq \emptyset$, we find

$$\boldsymbol{F}_p = \mathcal{F}_{n_p}^{f_s+f_a-o_p}(\mathcal{J}_p) \subseteq \mathcal{F}_{n_p}^{f'}(\mathcal{J}_p) \tag{18}$$

by item (1) of Lemma 3. Lemma 6 is now applicable to the right-hand side of eq. (18) and it follows by its item (2) that

$$\boldsymbol{F}_p \subseteq \bigcap_{j=1}^{n-2f'-f'_u} \boldsymbol{J}_p^{p_j}.$$

An analogous result holds for $\boldsymbol{F}_q$. Of course, the majorizing intersections for $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ involve non-faulty intervals only, hence are both accurate and thus intersecting. This justifies eq. (17) and the condition on $|\boldsymbol{F}_p \cup \boldsymbol{F}_q|$ given in the lemma. Note carefully, however, that this does not imply that $\boldsymbol{F}_p$ and $\boldsymbol{F}_q$ itself are accurate or even just intersecting! On the other hand, if $f' \leq f_s + f_a$, it follows from item (1) of Lemma 6 applied to the left-hand side of eq. (18) that $\boldsymbol{F}_p$ (and analogously $\boldsymbol{F}_q$) is accurate. □

Remarks:

1. It follows from item (3) of the above lemma that there are two possibilities in case of a violation of the fault assumptions: Either a node recognizes this fact because the result of $\mathcal{F}$ is empty, or the computed interval is not "too wrong". Obviously, this is some form of *graceful degradation* of $\mathcal{F}$'s performance.
2. Evidently, the worst situation with respect to the number of faults where one can hope to get a meaningful result is $n \geq 2f' + 1$. Item (3) of Lemma 8 can be used to deduce a result for this case as well: Setting $f'_u = 0$ and declaring any interval with an unbounded accuracy fault as being "non-faulty", we get from eq. (17) that $\boldsymbol{F}_p \cup \boldsymbol{F}_q$ lies in the union of the intersection of $n-2f'$ "non-faulty" intervals in $\mathcal{J}_p$ resp. $\mathcal{J}_q$.
3. Comparison[2] of Lemma 8 with [12, Lem. 5] shows that $\mathcal{F}$ again provides the same results as the Marzullo function $\mathcal{M}$. This finally justifies our claim that $\mathcal{F}$ and $\mathcal{M}$ have the same worst-case performance.

4 Conclusions

We presented and analyzed a novel Fault-Tolerant Interval (FTI) intersection function $\mathcal{F}$, which is optimal like the well-known Marzullo function $\mathcal{M}$ but satisfies a Lipschitz condition as well. The Lipschitz condition ensures that minor changes of the input intervals cause only minor changes of the result.

Our thorough analysis revealed that $\mathcal{F}$ has exactly the same worst-case performance as $\mathcal{M}$, although it may provide slightly sub-optimal results for non-worst-case input scenarios. For the local application case, we showed that the interval $\boldsymbol{F}$ provided by $\mathcal{F}$ on a single node lies within $n - 2f_n - 3f_u$ non-faulty input intervals, where f_n resp. f_u is the maximum number of non-accurate resp. excessively long intervals among the totally n input intervals. For the distributed application case, we showed that the non-commutative union of the results $\boldsymbol{F}_p \sqcup \boldsymbol{F}_q$ provided by $\mathcal{F}$ at two different nodes p and q lies within the intersection of $n - 2f_s - 3f_a$ nc-unions of corresponding non-faulty input intervals at p and q, where f_s resp. f_a gives the number of symmetric resp. asymmetric faults.

Therefore, $\mathcal{F}$ is a promising candidate for replacing the widespread usage of $\mathcal{M}$ in distributed applications. Some of our future work will be devoted to the investigation of its usefulness in our interval-based clock synchronization framework [16], where $\mathcal{F}$'s Lipschitz condition might prove particularly beneficial. This research also includes the usage of agreement protocols, which allow to reduce the inconsistency of $\mathcal{F}$'s input sets at different nodes.

Acknowledgements. We are grateful to Leslie Lamport and an anonymous referee for their stimulating comments on an earlier version of our manuscript. We would not have unveiled $\mathcal{F}$'s optimality without this feedback.

References

1. M. H. Azadmanesh and Roger M. Kieckhafer. New hybrid fault models for asynchronous approximate agreement. *IEEE Transactions on Computers*, 45(4):439–449, 1996.
2. R. R. Brooks and S. S. Iyengar. Robust distributed computing and sensing algorithms. *IEEE Computer*, pages 53–60, June 1996.
3. S.S. Iyengar, D.N. Jayashimha, and D. Nadig. A versatile architecture for the distributed sensor integration problem. *IEEE Transactions on Computers*, 43(2):175–185, February 1994.
4. D.N. Jayashimha, S.S. Iyengar, and R.L. Kashyap. Information integration and synchronization in distributed sensor networks. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5):1032–1043, September/October 1991.
5. Leslie Lamport. Synchronizing time servers. Technical Report 18, Digital System Research Center, 1987.
6. Jennifer Lundelius-Welch and Nancy A. Lynch. A new fault-tolerant algorithm for clock synchronization. *Information and Computation*, 77(1):1–36, 1988.
7. Keith A. Marzullo. *Maintaining the Time in a Distributed System: An Example of a Loosely-Coupled Distributed Service*. PhD dissertation, Stanford University, Department of Electrical Engineering, February 1984.
8. Keith A. Marzullo. Tolerating failures of continuous-valued sensors. *ACM Transactions on Computer Systems*, 8(4):284–304, 1990.
9. David L. Mills. Improved algorithms for synchronizing computer network clocks. *IEEE Transactions on Networks*, pages 245–254, June 1995.
10. OSF. *Introduction to OSF DCE*. Prentice Hall, Englewood Cliffs, NJ, 1992.

11. Ulrich Schmid. Synchronized Universal Time Coordinated for distributed real-time systems. *Control Engineering Practice*, 3(6):877–884, 1995. (Reprint from Proceedings 19th IFAC/IFIP Workshop on Real-Time Programming (WRTP'94), Lake Reichenau/Germany, 1994, p. 101–107.).
12. Ulrich Schmid. Orthogonal accuracy clock synchronization. *Chicago Journal of Theoretical Computer Science*, 3:1–77, 2000.
13. Klaus Schossmaier. An interval-based framework for clock rate synchronization algorithms. In *Proceedings 16th ACM Symposium on Principles of Distributed Computing*, pages 169–178, St. Barbara, USA, August 21–24, 1997.
14. Ulrich Schmid, Martin Horauer, and Nikolaus Kerö. How to distribute GPS-time over COTS-based LANs. In *Proceedings of the 31th IEEE Precise Time and Time Interval Systems and Application Meeting (PTTI'99)*, pages 545–560, Dana Point, California, December 1999.
15. Ulrich Schmid, Johann Klasek, Thomas Mandl, Herbert Nachtnebel, Gerhard R. Cadek, and Nikolaus Kerö. A Network Time Interface M-Module for distributing GPS-time over LANs. *J. Real-Time Systems*, 18(1):24–57, January 2000.
16. Ulrich Schmid and Klaus Schossmaier. Interval-based clock synchronization. *J. Real-Time Systems*, 12(2):173–228, March 1997.
17. Chris J. Walter, Patrick Lincoln, and Neeraj Suri. Formally verified on-line diagnosis. *IEEE Transactions on Software Engineering*, 23(11):684–721, November 1997.

Ulrich Schmid received his diploma and PhD degree in computer science and mathematics from the Technical University of Vienna. He has several years of experience in industrial electronics and embedded systems design and is now associate professor at the Department of Automation at TU Vienna. Dr. Schmid wrote numerous papers in the field of theoretical and technical computer science and received several best-paper awards and prices. His current research interests focus on distributed algorithms, real-time embedded systems and wireless computer communications. Ulrich Schmid is a member of the IEEE Computer Society and EATCS.

Klaus Schossmaier earned a Diploma (1991) and PhD (1998) degree in Computer Science from the TU Vienna. Under the Fulbrigh scholarship program he received a MSc (1994) degree in Computer Science from the University of Massachusetts at Amherst. His research interests cover distributed computing, embedded systems, and advanced operating systems. He has been working in project SynUTC at the TU Vienna and joined 1999 the ALICE data acquisition group at CERN in Geneva. Dr. Schossmaier is a member of the ACM.