

Monitoring Real-Time Systems under Parametric Delay[★]

Martin Fränzle¹, Thomas M. Grosen², Kim G. Larsen², and Martin Zimmermann²

¹ Carl von Ossietzky Universität, Oldenburg, Germany martin.fraenzle@uol.de

² Aalborg University, Aalborg, Denmark {tmgr,kgl,mzi}@cs.aau.dk

Abstract. Online monitoring of embedded real-time systems can be achieved by reduction of an adequate property language, like Metric Interval Temporal Logic, to timed automata and symbolic execution of the resulting automata on the trace observed from the system. This direct construction however only is faithful if observation of the trace is immediate in the sense that the monitor can assign exact time stamps to the actions it observes, which is rarely true in practice due to the substantial and fluctuating parametric delays introduced by the circuitry connecting the observed system to its monitoring device. We present a purely zone-based online monitoring procedure and its implementation which handle such parametric delays exactly without recurrence to costly verification procedures for parametric timed automata.

Keywords: Monitoring · Timing uncertainty · MITL · Timed Automata.

1 Introduction

Online monitoring is an important tool to ensure functional correctness of safety-critical systems. It analyses the execution traces observed from the system during its runtime by determining in real-time whether the observed traces satisfy the system’s specification. Continuous online monitoring consequently is concerned with unbounded time horizons, unlike offline monitoring where a fixed finite trace is analysed after the execution has terminated. Hence, specifications for online monitoring are typically defined over infinite traces, the most significant approach being temporal logics. As specifications often include real-time requirements, e.g., “every request is answered within 10 milliseconds”, we focus here on metric-time temporal logics over timed words. More precisely, we consider Metric Interval Temporal Logic (MITL) [2], which offers a good balance between expressiveness and algorithmic properties. For example, the request-response specification above is expressed by the MITL formula $G_{\geq 0}(\mathbf{req} \rightarrow F_{\leq 10}\mathbf{resp})$.

While the specifications classify infinite traces, the traces observed online and to be checked against the specification remain finite. Nevertheless, one can still return verdicts: for example, *every* infinite extension of a finite trace with some request that is not answered within 10 milliseconds violates the request-response specification above. Hence, violation of the specification is already witnessed by such a finite trace. Dually, consider the specification “system calibration is completed within 500 milliseconds”, expressed by the formula $F_{\leq 500}\mathbf{cc}$ with the proposition $\mathbf{cc}$ representing the completion of calibration. Every infinite extension of a finite trace on which the calibration is completed within 400 milliseconds satisfies the specification. Hence, satisfaction of the specification is already witnessed by such a finite trace. However, there are also traces and specifications for which no verdict can currently be drawn, like in the situation where no calibration has been observed yet at current time of 350 milliseconds. As usual, we capture these three situations with the three verdicts $\top$ (satisfaction for every extension), $\perp$ (violation for every extension), and $?$ (inconclusive).

Online monitoring can be achieved by compiling the MITL specification into an equivalent timed Büchi automaton and then symbolically executing the automaton on the observed trace of the system [7,12].

[★] M. Fränzle has been funded by Deutsche Forschungsgemeinschaft under grant FR 2715/5-1 and by the State of Lower Saxony, Zukunftslabor Mobilität.

T.M. Grosen, K.G. Larsen, and M. Zimmermann have been supported by DIREC - Digital Research Centre Denmark.

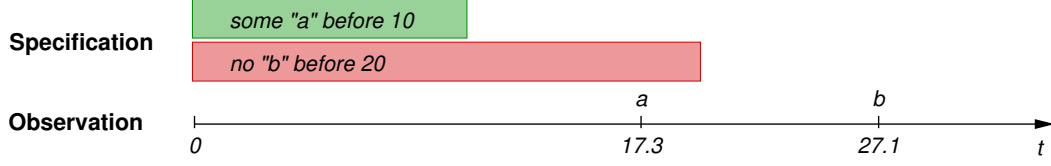


Fig. 1: Monitoring under observation delay: at time $t = 27.1$ we can conclusively decide that the MITL property $F_{[0,10]}a \wedge G_{[0,20]}\neg b$ is violated irrespective of the latency of the observation channel, provided its jitter is less than 0.2.

However, this approach is correct only if the actions of the monitored system can be observed immediately by the monitor. In practice, there is usually a communication delay between the system and the monitor. This delay is induced by various types of circuitry at their interfaces, like technical sensors, conversion between analog and digital signals, and communication networks forwarding signals to the monitor.

Consequently, the monitored system and the symbolic execution are no longer synchronized but deviate by a delay, for which only bounds, yet not exact values tend to be known. Even then, one can still provide meaningful verdicts, see Fig. 1: The specification $F_{\leq 10}a \wedge G_{\leq 20}\neg b$ expresses that an a occurs within 10 milliseconds and no b occurs within 20 milliseconds. The observed trace shows the first a at 17.3 milliseconds and the first b at 27.1 milliseconds. This observation only is consistent with satisfaction of the constraint $F_{\leq 10}a$ if a 's observation delay exceeds 7.3 milliseconds, while satisfaction of $G_{\leq 20}\neg b$ requires a delay of at most 7.1 milliseconds for b . When unknown communication delays vary by a jitter of less than 0.2 milliseconds, the specification definitely is violated.

Our Contribution. We follow the approach described in McGraw-Hill's Encyclopedia of Networking and Telecommunications [15] where a communication delay consists of a constant part (latency) and varying part (jitter). Based on previous work by Grosen et al. [12] on online monitoring of MITL specifications without delay via timed Büchi automata, we present a symbolic MITL monitoring algorithm that provides exact verdicts under unknown delay consisting of parametric (i.e. unknown within bounds) latency and jitter. While an unknown delay is a timing parameter, our construction avoids the semidecidability [3] of analysis for parameterized timed automata, and instead uses only classical clock zones [9].

In addition, our approach has the advantage that it is even more informative than typical monitoring algorithms, which only return a verdict in $\{\top, \perp, ?\}$. Recall the example specification $F_{\leq 10}a \wedge G_{\leq 20}\neg b$. Now, consider a trace with an a occurring after 17.3 milliseconds and a b after 27.5 milliseconds. Using reasoning as above, one can see that the specification is satisfied when the delay is in the interval $[7.3, 7.5)$, while for all other delays it is violated. Our algorithm, for which we also provide a prototype implementation and experimental evaluation, computes the set of potential delays under which the specification is satisfied as well as the set of potential delays under which the specification is violated.

All proofs omitted due to space restrictions can be found in the appendix.

Related Work. Our automata-based monitoring of finite traces against specifications over infinite words using the three verdicts $\{\top, \perp, ?\}$ follows the seminal work of Bauer et al. [7], who presented monitoring algorithms for LTL and timed LTL. Their algorithm for timed LTL is based on clock regions [1], while we follow the approach of Grosen et al. [12] and use clock zones [9], whose performance is an order of magnitude faster. Also, they translated timed LTL into event-clock automata, which are less expressive than the timed Büchi automata (TBA) used both by Grosen et al. [12] and here.

As our algorithms work with TBA, we also support MITL specifications, as these can be compiled into TBA. The monitoring problem for MITL (without delay) has been investigated before. Baldor et al. showed how to construct a monitor for dense-time MITL formulas by constructing a tree of timed transducers [5]. Ho et al. split unbounded and bounded parts of MITL formulas for monitoring, using traditional LTL monitoring for the unbounded parts and permitting a simpler construction for the (finite-word) bounded parts [13].

There is also a large body of work on monitoring with finite-word semantics. Roşu et al. focussed on discrete-time finite-word MTL [16], while Basin et al. proposed algorithms for monitoring real-time finite-word properties [6] and compared the differences between different time models. André et al. consider monitoring finite logs of parameterized timed and hybrid systems [19]. Finally, Ulus et al. described monitoring timed regular expressions over finite words using unions of two-dimensional zones [17,18].

2 Preliminaries

The set of natural numbers (excluding zero) is $\mathbb{N}$, we define $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, the set of rational numbers is $\mathbb{Q}$, the set of non-negative rational numbers is $\mathbb{Q}_{\geq 0}$, the set of real numbers is $\mathbb{R}$, and the set of non-negative real numbers is $\mathbb{R}_{\geq 0}$. The powerset of a set S is denoted by 2^S .

Timed Words. A timed word over a finite alphabet Σ is a pair $\rho = (\sigma, \tau)$ where σ is a nonempty word over Σ and τ is a sequence of non-decreasing non-negative real numbers of the same length as σ . Timed words may be finite or infinite; in the latter case, we require $\limsup \tau = \infty$, i.e., time diverges. The set of finite timed words is denoted by $T\Sigma^*$ and the set of infinite timed words by $T\Sigma^\omega$. We also represent a timed word as a sequence of pairs $(\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots$. If $\rho = (\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots, (\sigma_n, \tau_n)$ is a finite timed word, we denote by $\tau(\rho)$ the total time duration of ρ , i.e., τ_n .

If $\rho_1 = (\sigma_1^1, \tau_1^1), \dots, (\sigma_n^1, \tau_n^1)$ is a finite timed word, $\rho_2 = (\sigma_1^2, \tau_1^2), (\sigma_2^2, \tau_2^2), \dots$ is a finite or infinite timed word, and $t \in \mathbb{Q}_{\geq 0}$ then the timed word concatenation $\rho_1 \cdot t \rho_2$ is defined iff $\tau(\rho_1) \leq t$. Then, $\rho_1 \cdot t \rho_2 = (\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots$ such that

$$\sigma_i = \begin{cases} \sigma_i^1 & \text{iff } i \leq n \\ \sigma_{i-n}^2 & \text{else} \end{cases} \quad \text{and} \quad \tau_i = \begin{cases} \tau_i^1 & \text{iff } i \leq n \\ \tau_{i-n}^2 + t & \text{else.} \end{cases}$$

Timed Automata. A timed Büchi automaton (TBA) $\mathcal{A} = (Q, Q_0, \Sigma, C, \Delta, \mathcal{F})$ consists of a finite alphabet Σ , a finite set Q of locations, a set $Q_0 \subseteq Q$ of initial locations, a finite set C of clocks, a finite set $\Delta \subseteq Q \times Q \times \Sigma \times 2^C \times G(C)$ of transitions with $G(C)$ being the set of clock constraints over C , and a set $F \subseteq Q$ of accepting locations. A transition (q, q', a, λ, g) is an edge from q to q' on input symbol a , where λ is the set of clocks to reset and g is a clock constraint over C . A clock constraint is a conjunction of atomic constraints of the form $x \sim n$, where x is a clock, $n \in \mathbb{N}_0$, and $\sim \in \{<, \leq, =, \geq, >\}$. A state of $\mathcal{A}$ is a pair (q, v) where q is a location in Q and $v: C \rightarrow \mathbb{R}_{\geq 0}$ is a valuation mapping clocks to their values. For any $d \in \mathbb{R}_{\geq 0}$, $v + d$ is the valuation $x \mapsto v(x) + d$.

A run of $\mathcal{A}$ from a state (q_0, v_0) over a timed word $(\sigma_1, \tau_1)(\sigma_2, \tau_2) \dots$ is a sequence of steps $(q_0, v_0) \xrightarrow{(\sigma_1, \tau_1)} (q_1, v_1) \xrightarrow{(\sigma_2, \tau_2)} (q_2, v_2) \xrightarrow{(\sigma_3, \tau_3)} \dots$ where for all $i \geq 1$ there is a transition $(q_{i-1}, q_i, \sigma_i, \lambda_i, g_i)$ such that $v_i(x) = 0$ for all x in λ_i and $v_i(x) = v_{i-1}(x) + (\tau_i - \tau_{i-1})$ otherwise, and g_i is satisfied by the valuation $v_{i-1} + (\tau_i - \tau_{i-1})$. Here, we use $\tau_0 = 0$. Given a run r , we denote the set of locations visited infinitely many times by r as $\text{Inf}(r)$. A run r of $\mathcal{A}$ is accepting if $\text{Inf}(r) \cap F \neq \emptyset$. The language of $\mathcal{A}$ from a starting state (q, v) , denoted $L(\mathcal{A}, (q, v))$, is the set of all infinite timed words with an accepting run in $\mathcal{A}$ starting from (q, v) . We define the language of $\mathcal{A}$, written $L(\mathcal{A})$, to be $\bigcup_q L(\mathcal{A}, (q, v_0))$, where q ranges over Q_0 and where $v_0(x) = 0$ for all $x \in C$.

Metric Interval Temporal Logic. We use Temporal interval logic (MITL) to express properties to be monitored; these are subsequently translated into equivalent TBA which we use in our monitoring algorithm.

The syntax of MITL formulas over a finite alphabet Σ is defined as $\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid X_I\varphi \mid \varphi U_I\varphi$ where $p \in \Sigma$ and I ranges over non-singular intervals over $\mathbb{R}_{\geq 0}$ with endpoints in $\mathbb{N} \cup \{\infty\}$. We write $\sim n$ for $I = \{d \in \mathbb{R}_{\geq 0} \mid d \sim n\}$ for $\sim \in \{<, \leq, \geq, >\}$ and $n \in \mathbb{N}$.

The satisfaction relation $\rho, i \models \varphi$ is defined for infinite timed words $\rho = (\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots$, positions $i \geq 1$, and MITL formulas φ :

$$- \rho, i \models p \text{ iff } p = \sigma_i.$$

- $\rho, i \models \neg\varphi$ iff $\rho, i \not\models \varphi$.
- $\rho, i \models \varphi \vee \psi$ if $\rho, i \models \varphi$ or $\rho, i \models \psi$.
- $\rho, i \models X_I\varphi$ iff $\rho, (i+1) \models \varphi$ and $\tau_{i+1} - \tau_i \in I$.
- $\rho, i \models \varphi U_I\psi$ iff there exists $k \geq i$ s.t. $\rho, k \models \psi$, $\tau_k - \tau_i \in I$, and $\rho, j \models \varphi$ for all $i \leq j < k$.

We write $\rho \models \varphi$ whenever $\rho, 1 \models \varphi$. We also define the standard syntactic sugar: $\mathbf{true} = p \vee \neg p$, $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$, $F_I\varphi = \mathbf{true} U_I\varphi$, and $G_I\varphi = \neg F_I\neg\varphi$. The language $L(\varphi)$ of an MITL formula φ is the set of all $\rho \in T\Sigma^\omega$ such that $\rho \models \varphi$.

Theorem 1 ([2,10]). *For each MITL formula φ there exists a TBA $\mathcal{A}$ with $L(\varphi) = L(\mathcal{A})$.*

Fig. 2 illustrates Theorem 1 by providing TBA's for the formula $F_{[0,10]}a \wedge G_{[0,20]}\neg b$ from the introduction and its negation.

3 Monitoring under Delayed Observation

According to McGraw-Hill's Encyclopedia of Networking and Telecommunications [15], a communication delay consists of a constant part (latency) and varying part (jitter). We describe the delay as a pair $(\delta, \varepsilon) \in \mathbb{R}_{\geq 0}^2$ where δ is the constant latency for all signals and ε is the bound on the jitter. Thus, all signals from the system are delayed within $[\delta, \delta + \varepsilon]$ before they arrive at the monitor.

In the simplest case, our obligation is to monitor violation of an MITL specification φ by a system while observing the events through a channel *Chan* featuring a constant, yet unknown (up to given, but maybe trivial, lower bound $l \in \mathbb{R}_{\geq 0}$ and upper bound $u \in \mathbb{R}_{\geq 0}$) transportation latency $\delta \in [l, u]$ and a varying jitter bounded by $\varepsilon \in \mathbb{R}_{\geq 0}$. Fig. 1 shows an example of a property and an observation that conclusively violates the specification at time 27.1, even if the channel latency $\delta \in [0, \infty[$ is unknown, as long as the jitter is bounded by 0.2.

We begin by formalizing the concept of observation, where the occurrence of observed events is constrained by a set $\mathcal{D}$ capturing known bounds on the delay.

Definition 1. *A delay set $\mathcal{D}$ is a nonempty subset of $\mathbb{R}_{\geq 0}^2$. A $\mathcal{D}$ -observation is a finite timed word $\rho^* = (\sigma_1^*, \tau_1^*), \dots, (\sigma_m^*, \tau_m^*)$ with $\tau_1^* \geq \delta$ for some $(\delta, \varepsilon) \in \mathcal{D}$.*

As the ground-truth occurrence times of events in the system cannot be determined exactly from their delayed copies that the monitor receives through the communication channel, we have to consider all ground-truth timed words that the particular observation is consistent with, as follows.³

Definition 2 (Consistency). *A finite timed word $\rho = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ is consistent with a $\{(\delta, \varepsilon)\}$ -observation $\rho^* = (\sigma_1^*, \tau_1^*), \dots, (\sigma_m^*, \tau_m^*)$ at observation time $t \in \mathbb{R}_{\geq 0}$ under latency δ and jitter ε iff*

³ Note that we simplify our definitions by assuming that jitter does not change the order of observations. Under the assumption that only a bounded number of events can be generated by the system in each unit of time, it is possible to take "overtaking" of events into account, by looking at all consistent permutations. We refrain from doing so for the sake of clarity.

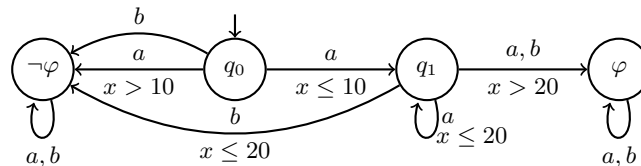


Fig. 2: An automaton for the language of the property $\varphi = F_{[0,10]}a \wedge G_{[0,20]}\neg b$ and its negation: If location φ is accepting then it accepts $L(\varphi)$, if location $\neg\varphi$ is accepting then it accepts $L(\neg\varphi)$.

1. $\tau_i \leq t$ for each $i \in \{1, \dots, n\}$ and $\tau_j^* \leq t$ for each $j \in \{1, \dots, m\}$,
2. $n \geq m$, and $\sigma_i = \sigma_i^*$ and $\tau_i + \delta - \tau_i^* \in [0, \varepsilon]$ for all $i \in \{1, \dots, m\}$, and
3. if $n > m$ then $\tau_{m+1} + \delta + \varepsilon \geq t$.

We denote the set of timed words ρ that are consistent with an $\{(\delta, \varepsilon)\}$ -observation ρ^* at observation time t under latency δ and jitter ε by $GT_{\delta, \varepsilon}(\rho^*, t)$. Then, we define $GT_{\mathcal{D}}(\rho^*, t) = \bigcup_{(\delta, \varepsilon) \in \mathcal{D}} GT_{\delta, \varepsilon}(\rho^*, t)$.

$GT_{\mathcal{D}}(\rho^*, t)$ thus collects the possible ground-truths that are consistent with the observation ρ^* when the time elapsed since the system has started is t , and the delay (δ, ε) is within the set $\mathcal{D}$. Note that $GT_{\mathcal{D}}(\rho^*, t)$ is always nonempty, if ρ^* is a $\mathcal{D}$ -observation and $t \geq \tau(\rho^*)$.

A monitor obviously ought to supply a verdict iff that verdict applies across *all possible* ground-truth timed words that the observed word explains.

Definition 3 (Monitor verdicts under delay). Given a language $L \subseteq T\Sigma^\omega$, a set of possible observation delays $\mathcal{D}$, a $\mathcal{D}$ -observation $\rho^* \in T\Sigma^*$, and an observation time $t \geq \tau(\rho^*)$, the function $\mathcal{V}_{\mathcal{D}}: 2^{T\Sigma^\omega} \rightarrow T\Sigma^* \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{B}_3$ evaluates to the verdict

$$\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \begin{cases} \top & \text{if } \rho \cdot_t \mu \in L \text{ for all } \rho \in GT_{\mathcal{D}}(\rho^*, t) \text{ and all } \mu \in T\Sigma^\omega, \\ \perp & \text{if } \rho \cdot_t \mu \notin L \text{ for all } \rho \in GT_{\mathcal{D}}(\rho^*, t) \text{ and all } \mu \in T\Sigma^\omega, \\ ? & \text{otherwise.} \end{cases}$$

$\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t)$ is undefined when $t < \tau(\rho^*)$.

Example 1. Consider the property $\varphi = F_{[0,10]}a \wedge G_{[0,20]} \neg b$ and observed word $\rho^* = (a, 17.3), (b, 27.1)$ shown in Fig. 1, time point $t = 27.1$, and set of delays $\mathcal{D} = \{(\delta, \varepsilon) \mid \varepsilon = 0.2\}$. As the jitter is bounded by 0.2, in all ground truths either a occurred after time point 10, or b occurred before time point 20. Thus, all extensions of all possible ground truths satisfy $\neg\varphi$, i.e., $\mathcal{V}_{\mathcal{D}}(L(\varphi))(\rho^*, t) = \perp$.

Note that for the special case of $\mathcal{D} = \{(0, 0)\}$ we cover classical (i.e., delay-free) monitoring [12]. Before we turn our attention to computing $\mathcal{V}$ we study some properties of our definition. First, let us note that the ability to make firm verdicts increases with increased certainty of the observation channel delay.

Lemma 1. Let $L \subseteq T\Sigma^\omega$, $\rho^* \in T\Sigma^*$, let $\mathcal{D} \subseteq \mathcal{D}'$ be delay sets, let ρ^* be a $\mathcal{D}$ -observation, and let $t \geq \tau(\rho^*)$. Then, $\mathcal{V}_{\mathcal{D}'}(L)(\rho^*, t) = \top$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$ and $\mathcal{V}_{\mathcal{D}'}(L)(\rho^*, t) = \perp$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$.

As a refinement of the verdict function in Definition 3, one may provide information about the delays (δ, ε) that can explain an observation. Given $L \subseteq T\Sigma^\omega$, a finite timed word $\rho^* \in T\Sigma^*$, and $t \geq \tau(\rho^*)$, the set of delays $\Delta(L, \rho^*, t)$ that are consistent with the observation ρ^* at t is defined as

$$\Delta(L, \rho^*, t) = \{(\delta, \varepsilon) \mid \exists \rho \in GT_{\delta, \varepsilon}(\rho^*, t) \exists \mu \in T\Sigma^\omega \text{ s.t. } \rho \cdot_t \mu \in L\}.$$

We denote by $\Delta_{\mathcal{D}}(L, \rho^*, t)$ the set $\Delta(L, \rho^*, t) \cap \mathcal{D}$. Now we can characterize the conclusive monitoring verdicts via these delay sets.

Lemma 2. Given $L \subseteq T\Sigma^\omega$, a set $\mathcal{D}$ of delays, a $\mathcal{D}$ -observation $\rho^* \in T\Sigma^*$, and $t \geq \tau(\rho^*)$, we have

1. $\Delta_{\mathcal{D}}(L, \rho^*, t) = \emptyset$ iff $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$, and
2. $\Delta_{\mathcal{D}}(\overline{L}, \rho^*, t) = \emptyset$ iff $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$.

But even in the case when both delay-sets are nonempty (i.e., the verdict is $?$) we can still provide useful information in terms of the sets $\Delta(L, \rho^*, t)$ and $\Delta(\overline{L}, \rho^*, t)$ of consistent delays. In particular, the set of consistent delays is non-increasing during observations. To explain this, we first define a notion of extensions for finite timed words (w.r.t. a current time instants t, t'). Formally, let $\rho = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ and $\rho' = (\sigma'_1, \tau'_1), \dots, (\sigma'_{n'}, \tau'_{n'})$ be two finite timed words. Also let $t, t' \in \mathbb{R}_{\geq 0}$. Then, we define $(\rho, t) \sqsubseteq (\rho', t')$, if $n \leq n'$, $\sigma_i = \sigma'_i$ and $\tau_i = \tau'_i$ for all $i \leq n$, and either $n = n'$ and $t \leq t'$ or $n < n'$ and $t \leq \tau'_{n+1}$. By extending the observations, we (potentially) reduce the set of consistent delays.

Lemma 3. *Let $(\rho_1^*, t_1) \sqsubseteq (\rho_2^*, t_2)$ for finite timed words ρ_1^*, ρ_2^* and $t_1 \geq \tau(\rho_1^*)$ and $t_2 \geq \tau(\rho_2^*)$. Then, $\Delta(L, \rho_1^*, t_1) \supseteq \Delta(L, \rho_2^*, t_2)$.*

Another interesting point is that in some cases, no extension of the observed word will provide a definitive verdict.

Example 2. Consider the language $L(F_{\leq 10}a)$, the observation $\rho^* = (a, 15)$, and the set $\mathcal{D} = \{(\delta, 0) \mid \delta \in [0, 10]\}$ of delays. For any given $t \geq \tau(\rho^*)$ the set of consistent delays are $\Delta_{\mathcal{D}}(L, \rho^*, t) = \{(\delta, 0) \mid \delta \in [5, 10]\}$ and $\Delta_{\mathcal{D}}(\bar{L}, \rho^*, t) = \{(\delta, 0) \mid \delta \in [0, 5]\}$, i.e., both sets of consistent delays are a strict subset of $\mathcal{D}$. Further, due to Lemma 3, this will be the case, no matter what observations occur in the future. Then, the verdict is $?$, even if additional observations occur.

The following lemma formalizes this: as soon as the set of consistent delays w.r.t. $L(\bar{L})$ is no longer equal to $\mathcal{D}$, then the verdict cannot be $\top$ ($\perp$) anymore.

Lemma 4. *Let $L \subseteq T\Sigma^\omega$, $\mathcal{D}$ be a set of delays, and $\rho^* = (\sigma_1^*, \tau_1^*), \dots, (\sigma_m^*, \tau_m^*)$ a nonempty $\mathcal{D}$ -observation. Then, for all $t > \tau(\rho^*)$*

1. $\Delta_{\mathcal{D}}(L, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ implies there is no $\rho_1^* \in T\Sigma^*$ such that $\mathcal{V}_{\mathcal{D}}(L)(\rho^* \cdot_t \rho_1^*, t') = \top$ for any $t' \geq t + \tau(\rho_1^*)$, and
2. $\Delta_{\mathcal{D}}(\bar{L}, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ implies there is no $\rho_1^* \in T\Sigma^*$ such that $\mathcal{V}_{\mathcal{D}}(L)(\rho^* \cdot_t \rho_1^*, t') = \perp$ for any $t' \geq t + \tau(\rho_1^*)$.

Note that $\Delta_{\mathcal{D}}(L, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ and $\Delta_{\mathcal{D}}(\bar{L}, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ can both be true simultaneously (as in Example 2). In this situation, we will under no future observation reach a conclusive verdict.

4 Towards an Implementation

Typically, monitoring algorithms rely on automata-based techniques. To this end, first the specification and its complement are translated into suitable automata. Then one computes the set of states reachable by processing the observation and then checks whether from one these states the automaton can still accept an infinite continuation. If this is the case for both automata, then the verdict is $?$, if it is only the case for the automaton for the specification, then the verdict is $\top$, and vice versa for the complement automaton and $\perp$.

We want to follow the same blueprint, but we need to make adjustments to handle delay. Intuitively, we need to compute all states that are reachable by possible ground-truths of a given observation. However, a ground-truth may contain more events than the observation, as some events may not yet have been observed due to delay. This complicates the construction of the set of reachable states, as an unbounded number of events may not yet have been observed.

In the definition of $GT_{\mathcal{D}}$ there is an implicit universal quantification over all possible sequences of such events that have not yet been observed. We exploit the fact that the verdicts are defined with respect to all possible extensions μ of a possible ground-truth (i.e., also a universal quantification over the μ 's) to “move” the universal quantification over events that have not yet been observed to the universal quantification of the extension μ . Then, a possible ground-truth has exactly the same number of events as the observation. We begin defining this restricted notion of possible ground-truth by strengthening Definition 2.

Definition 4 (EL-Consistency). *A finite timed word $\rho = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ is EL-consistent with a $\{(\delta, \varepsilon)\}$ -observation $\rho^* = (\sigma_1^*, \tau_1^*), \dots, (\sigma_m^*, \tau_m^*)$ at observation time $t \in \mathbb{R}_{\geq 0}$ under latency δ and jitter ε iff ρ is consistent with ρ^* at t under δ and ε and $m = n$. We denote the set of timed words ρ that are EL-consistent with an observed timed word ρ^* at observation time t under latency δ and jitter ε by $GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$ and define $GT_{\mathcal{D}}^{\text{el}}(\rho^*, t) = \bigcup_{(\delta, \varepsilon) \in \mathcal{D}} GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$.*

Now, we present the revised verdict function using only EL ground-truths. Note that moving the unobserved events from the possible ground-truth ρ to the extension μ requires changing the time instant at which we concatenate the ground-truth and the extension: $t - (\delta + \varepsilon)$ is the earliest time point at which an event can occur that may not yet have been observed at time t . Due to jitter however, there might also be events after $t - (\delta + \varepsilon)$ that have been observed, which are in the possible ground-truth ρ : the last such event happened at time $\tau(\rho)$. Hence, we need to concatenate at time point $\max(\tau(\rho), t - (\delta + \varepsilon))$.

Definition 5 (Monitor verdicts under delay – EL version). *Given $L \subseteq T\Sigma^\omega$, a set $\mathcal{D}$ of delays, a $\mathcal{D}$ -observation $\rho^* \in T\Sigma^*$, and $t \geq \tau(\rho^*)$, the function $\mathcal{V}_\mathcal{D}^{\text{el}}: 2^{T\Sigma^\omega} \rightarrow T\Sigma^* \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{B}_3$ evaluates to the verdict*

$$\mathcal{V}_\mathcal{D}^{\text{el}}(L)(\rho^*, t) = \begin{cases} \top & \text{if } \rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \mu \in L \text{ for all } (\delta, \varepsilon) \in \mathcal{D}, \\ & \text{all } \rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t) \text{ and all } \mu \in T\Sigma^\omega, \\ \perp & \text{if } \rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \mu \notin L \text{ for all } (\delta, \varepsilon) \in \mathcal{D}, \\ & \text{all } \rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t) \text{ and all } \mu \in T\Sigma^\omega, \\ ? & \text{otherwise.} \end{cases}$$

$\mathcal{V}_\mathcal{D}^{\text{el}}(L)(\rho^*, t)$ is undefined when $t < \tau(\rho^*)$.

Next, we show that both verdict functions coincide.

Lemma 5. $\mathcal{V}_\mathcal{D}^{\text{el}}(L)(\rho^*, t) = \mathcal{V}_\mathcal{D}(L)(\rho^*, t)$ for all $L \subseteq T\Sigma^\omega$, all sets $\mathcal{D}$ of delays, all $\mathcal{D}$ -observations ρ^* , and all $t \geq \tau(\rho^*)$.

Next, we show that we can indeed make the definition of $\mathcal{V}^{\text{el}}$ effective using automata-theoretic constructions. First, we formally capture the set of states that can be reached by processing the possible EL ground-truths of an observation. Let $\mathcal{A}$ be a TBA. We write $(q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}} (q_n, v_n)$ for a finite timed word $\rho = (\sigma, \tau) \in T\Sigma^*$ to denote the existence of a finite sequence of states $(q_0, v_0) \xrightarrow{(\sigma_1, \tau_1)} (q_1, v_1) \xrightarrow{(\sigma_2, \tau_2)} \dots \xrightarrow{(\sigma_n, \tau_n)} (q_n, v_n)$ of $\mathcal{A}$ where for all $1 \leq i \leq n$ there is a transition $(q_{i-1}, q_i, \sigma_i, \lambda_i, g_i)$ of $\mathcal{A}$ such that $v_i(x) = 0$ for all x in λ_i and $v_{i-1}(x) + (t_i - t_{i-1})$ otherwise, and g is satisfied by the valuation $v_{i-1} + (t_i - t_{i-1})$, where we use $t_0 = 0$. Given a TBA $\mathcal{A}$, a set $\mathcal{D}$ of delays, a finite observed timed word $\rho \in T\Sigma^*$, and $t \geq \tau(\rho)$, we define

$$\begin{aligned} \mathcal{R}_\mathcal{A}^\mathcal{D}(\rho^*, t) = \{ & (q, v + \max(0, (t - (\tau(\rho) + \delta + \varepsilon)))) \mid (q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}} (q, v) \text{ where} \\ & (q_0, v_0) \text{ with } q_0 \in Q_0, v_0(x) = 0 \text{ for all } c \in C, \text{ and} \\ & \rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t) \text{ for some } (\delta, \varepsilon) \in \mathcal{D} \}. \end{aligned}$$

We call this the reach-set of ρ in $\mathcal{A}$ at t w.r.t. $\mathcal{D}$.

Next, we define the set of states of a TBA from where it is possible to reach an accepting location infinitely many times in the future, i.e., those states from which an accepting run is possible. This is useful, because if processing a finite timed word leads to such a state, then the timed word can be extended to an infinite one in the language of the automaton, a notion that underlies the definitions of the verdict functions. Given a TBA $\mathcal{A} = (Q, Q_0, \Sigma, C, \Delta, \mathcal{F})$, the set of states with nonempty language is

$$S_\mathcal{A}^{ne} = \{(q, v) \mid q \in Q, v \in C \rightarrow \mathbb{R}_{\geq 0} \text{ s.t. } L(\mathcal{A}, (q, v)) \neq \emptyset\}.$$

The set $S_\mathcal{A}^{ne}$ can be computed using a zone-based fixpoint algorithm [12]. Using these definitions, we can give an *effective* definition of the verdict functions, which we show to be equivalent to the previous definitions and implementable.

In the following definition, $\mathbb{A}$ denotes the set of all TBA.

Definition 6 (Monitoring TBA). *Given a TBA $\mathcal{A}$, a complement automaton $\overline{\mathcal{A}}$ (i.e., with $L(\overline{\mathcal{A}}) = T\Sigma^\omega \setminus L(\mathcal{A})$), a set $\mathcal{D}$ of delays, a $\mathcal{D}$ -observation $\rho \in T\Sigma^*$, and $t \geq \tau(\rho)$, $\mathcal{M}_\mathcal{D}: \mathbb{A} \times \mathbb{A} \rightarrow T\Sigma^* \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{B}_3$ computes the verdict*

$$\mathcal{M}_\mathcal{D}(\mathcal{A}, \overline{\mathcal{A}})(\rho^*, t) = \begin{cases} \top & \text{if } \mathcal{R}_\mathcal{A}^\mathcal{D}(\rho^*, t) \cap S_\mathcal{A}^{ne} = \emptyset, \\ \perp & \text{if } \mathcal{R}_\mathcal{A}^\mathcal{D}(\rho^*, t) \cap S_\mathcal{A}^{ne} \neq \emptyset, \\ ? & \text{otherwise.} \end{cases}$$

$\mathcal{M}_{\mathcal{D}}(\mathcal{A}, \bar{\mathcal{A}})(\rho^*, t)$ is undefined if $t < \tau(\rho)$.

Next we show that this automata-based definition of monitoring is equal to the verdict functions defined above.

Theorem 2. $\mathcal{M}_{\mathcal{D}}(\mathcal{A}, \bar{\mathcal{A}})(\rho^*, t) = \mathcal{V}_{\mathcal{D}}^{\text{el}}(L(\mathcal{A}))(\rho^*, t)$ for all sets $\mathcal{D}$ of delays, all TBA $\mathcal{A}$ (and complement automata $\bar{\mathcal{A}}$), all $\mathcal{D}$ -observations ρ^* , and all $t \geq \tau(\rho^*)$.

Recall that $S_{\mathcal{A}}^{\text{ne}}$ can be computed for any given TBA $\mathcal{A}$. Therefore, in the next section, we show how to calculate $\mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t)$ for a given TBA $\mathcal{A}$, set $\mathcal{D}$ of delays, observation ρ^* , and time point t using a zone-based algorithm. This will then allow us to compute verdicts effectively.

5 A Zone-Based Monitoring Algorithm

In this section, we demonstrate how to compute the reach-set of ρ in $\mathcal{A}$ at t w.r.t. $\mathcal{D}$. So far we have developed the theory with observations, latency, and jitter being reals. Now, we are concerned with algorithms and thus assume all these quantities to be rationals. For the monitoring algorithm, we use – as standard in analysing timed automata models – symbolic states being pairs (q, Z) of locations and zones. A zone is a finite conjunction of constraints of the form $x \sim t$ and $x - x' \sim t$ for clocks x, x' , constants $t \in \mathbb{Q}_{\geq 0}$, and $\sim \in \{<, \leq, =, \geq, >\}$. Given two zones Z and Z' over a set C of clocks, and a set of clocks $\lambda \subseteq C$, we define the following operations on zones (which can be efficiently implemented using the DBM data-structure [9]):

- $Z[\lambda] = \{v \mid \exists v' \models Z \text{ s.t. } v(x) = 0 \text{ if } x \in \lambda, \text{ otherwise } v(x) = v'(x)\}$
- $Z^{\nearrow} = \{v \mid \exists v' \models Z \text{ s.t. } v = v' + d \text{ for some } d \in \mathbb{R}_{\geq 0}\}$
- $Z \wedge Z' = \{v \mid v \models Z \text{ and } v \models Z'\}$, and

We can use these functions to compute the successor states after an input. Given a TBA $\mathcal{A} = (Q, Q_0, \Sigma, C, \Delta, \mathcal{F})$, a symbolic state (q, Z) , and a letter $a \in \Sigma$, we define $\text{Post}((q, Z), a) = \{(q', Z') \mid (q, q', a, \lambda, g) \in \Delta, Z' = (Z^{\nearrow} \wedge g)[\lambda]\}$, as the set of states one can reach by taking an a -transition at some point in the future from (q, Z) . Using Post we can compute the successor states of a timed input $(a, \tau) \in \Sigma \times \mathbb{Q}_{\geq 0}$ by extending the zones with an additional clock *time* just recording time since system start. The successors of a symbolic state is $\text{Succ}((q, Z), (a, \tau)) = \{(q', Z') \mid (q', Z'') \in \text{Post}((q, Z), a), Z' = Z'' \wedge \text{time} = \tau\}$ and the successors of a set of symbolic states S is $\text{Succ}(S, (a, \tau)) = \bigcup_{(q', Z') \in S} \text{Succ}((q', Z'), (a, \tau))$.

In handling delayed observations, we assume that the delay set $\mathcal{D}$ consists of pairs (δ, ε) where the latency δ is bounded by an interval $[l, u] \subseteq \mathbb{R}_{\geq 0}$ for given $l, u \in \mathbb{Q}_{\geq 0}$, and that the jitter is bounded by a given $\varepsilon \in \mathbb{Q}_{\geq 0}$. To represent the latency and thereby be able to reason about and indirectly store the latency bounds, we add a clock *etime* representing the “expected” real time that an event generated just now could be observed by the monitor after having been delayed according to the latency. This allows us to

1. represent the actual latency as $\text{etime} - \text{time}$,
2. represent the initial knowledge about latencies by initializing $\text{etime} - \text{time}$ to the initially known bounds on latency, namely $\text{etime} - \text{time} \in [l, u]$ by setting time to 0 and constraining etime to $[l, u]$,
3. refine our knowledge about the actual latency when observing an event (σ^*, τ^*) by then setting etime to a value in $[\tau^* - \varepsilon, \tau^*]$.

Consequently, we change the initial zones to include the latency bounds l and u as the differences between the clocks *etime* and *time*. This way, *etime* represents the expected time an event is observed (at the monitor) and *time* represents the actual time the event happened (at the system). The aforementioned refinement (see Item 3 above and Fig. 3) then permits to deduce actual latency ranges consistent with the specification (or its negation) from observation times of events.

In detail, this refinement of the $\text{etime} - \text{time}$ relation works as follows. Given a TBA $\mathcal{A}$ extended with the clocks *time* and *etime*, and an observation $(\sigma, \tau^*) \in \Sigma \times \mathbb{Q}_{\geq 0}$, the successors of (q, Z) are

$$\begin{aligned} \text{Succ}_d((q, Z), (\sigma, \tau^*)) &= \{(q', Z') \mid (q', Z'') \in \text{Post}((q, Z), \sigma), \\ &\quad Z' = Z'' \wedge \text{etime} \leq \tau^* \wedge \text{etime} \geq \tau^* - \varepsilon\} \end{aligned}$$

and the successors $\text{Succ}_d(S, (\sigma, \tau^*))$ of a set of symbolic states S is equal to $\cup_{(q, Z) \in S} \text{Succ}_d((q, Z), (\sigma, \tau^*))$.

The monitoring algorithm will essentially apply Succ_d repeatedly to update the reach-set. The initial reach-set is given by the following zone Z_0^d requiring all ordinary clocks of the TBA $\mathcal{A}$ to be zero and with time and etime satisfying $\text{etime} - \text{time} \in [l, u]$. That is

$$Z_0^d \equiv \underbrace{\text{etime} - \text{time} \leq u \wedge \text{time} - \text{etime} \leq -l}_{\text{etime} - \text{time} \in [l, u]} \wedge \underbrace{\bigwedge_{x \in C \cup \{\text{time}\}} x = 0}_{x_1, \dots, x_{|C|} = 0, \text{time} = 0}.$$

Given a fixed jitter bound ε , we can now compute the reach-set after a sequence of observations under delay, where the latency is bounded in $[l, u]$.

Theorem 3. *Given a TBA $\mathcal{A}$, a delay set $\mathcal{D} = \{(\delta, \varepsilon) \mid \delta \in [l, u]\}$ with $l, u, \varepsilon \in \mathbb{Q}_{\geq 0}$, a $\mathcal{D}$ -observation $\rho^* = (\sigma_1, \tau_1^*), \dots, (\sigma_n, \tau_n^*)$, and $t \in \mathbb{Q}_{\geq 0}$ with $t \geq \tau_n^*$, the reach-set of ρ^* can be computed as*

$$\mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \{(q', Z') \mid (q', Z'') \in S_n, Z' = Z'' \wedge \text{etime} = t - \varepsilon\}$$

where $S_0 = \{(q_0, Z_0^d) \mid q_0 \in Q_0\}$ and $S_i = \text{Succ}_d(S_{i-1}, (\sigma_i, \tau_i^*))$ for $i \in \{1, \dots, n\}$.

This theorem allows us to implement a monitoring algorithm by computing the reach-sets and intersecting them with the set of nonempty language states.

The observation of events may lead to refinement of the difference between time and etime as depicted in Fig. 3.

Lemma 6. *Given $\mathcal{A}$, $\mathcal{D}$, ρ^* , t , and S_n as in Theorem 3, we can compute the set of consistent delays by looking at the bounds on $\text{etime} - \text{time}$:*

$$\Delta_{\mathcal{D}}(L(\mathcal{A}), \rho^*, t) = \{(\delta, \varepsilon) \in \mathcal{D} \mid S \models \text{etime} - \text{time} = \delta\}.$$

This information can be used to decorate the $?$ verdict, so that we can report a set of bounds on the latency for which we would provide a $\top$ or $\perp$ verdict.

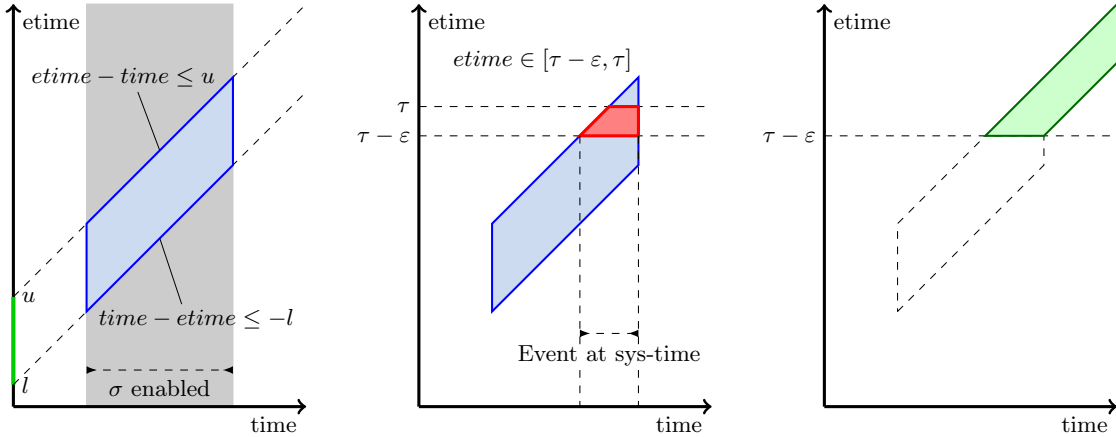


Fig. 3: Illustration of a single zone in the Succ_d computation (only time-etime plane depicted). Left: initial zone (in green) is diagonally extrapolated for time passage and then intersected with the guard of an edge. Middle: observing event σ at time τ . By restricting etime to $[\tau - \varepsilon, \tau]$, the clock time is restricted to when the event could have occurred at the system. Right: computing the future zone we see that the bound on $\text{time} - \text{etime}$ is now stricter and thus the bounds for the consistent latencies are refined.

Example 3. Let us show an example of our algorithm for monitoring under delayed observation. Consider the property $\varphi = F_{[0,10]}a \wedge G_{[0,20]}$ from Fig. 1. The TBA's accepting $L(\varphi)$ and $L(\neg\varphi)$ are shown in Fig. 2. The nonempty language states for $\mathcal{A}_\varphi$ and $\mathcal{A}_{\neg\varphi}$ are $S_{\mathcal{A}_\varphi}^{ne} = \{(q_0, \{x \leq 10\}), (q_1, \emptyset), (\varphi, \emptyset)\}$ and $S_{\mathcal{A}_{\neg\varphi}}^{ne} = \{(q_0, \emptyset), (q_1, \{x \leq 20\}), (\neg\varphi, \emptyset)\}$. Let us assume the latency is between 0 and 10, and the jitter is bounded by 0.2. Now we compute the reach-sets S_0 (initial), S_1 (after $(a, 17.3)$), and S_2 (after $(b, 27.5)$) as

$$\begin{aligned} S_0 &= \{(q_0, \{x = 0, etime \leq 10, etime - x \in [0, 10]\})\}, \\ S_1 &= \{(q_1, \{x \in [7.1, 10], etime \in [17.1, 17.3], etime - x \in [7.1, 10]\}), \\ &\quad (\neg\varphi, \{x \in [10, 17.3], etime \in [17.1, 17.3], etime - x \leq 7.3\})\}, \text{ and} \\ S_2 &= \{(\varphi, \{x \in]20, 20.4], etime \in [27.3, 27.5[, etime - x \in [7.1, 7.5]\}), \\ &\quad (\neg\varphi, \{x \in [17.3, 27.5], etime \in [27.3, 27.5], etime - x \in [0, 10]\})\}. \end{aligned}$$

Note that we omit the clock time and only look at x and $etime$ since $time$ and x always have the same constraints.

All reach-sets intersect with both sets of nonempty language states; thus, the verdict is $?$. However, we can refine this verdict with knowledge about the consistent delays that change after each observation. The jitter bound is fixed at 0.2, but the bounds on the latency can be found in the clock constraints on the difference between $etime$ and x . For $\perp$, the latency range remains $[0, 10]$ in all reach sets. For $\top$, the consistent latency range is $[0, 10]$ in S_0 , $[7.1, 10]$ in S_1 , and it is $[7.1, 7.5[$ in S_2 . This means that if the latency is outside $[7.1, 7.5[$, then the verdict is $\perp$.

On the other hand, for the observation $\rho^* = (a, 17.3), (b, 27.1)$ from Example 1 (and using the same latency and jitter bounds as above), we compute the reach-sets S_0 , S_1 , and S'_2 where

$$S'_2 = \{(\neg\varphi, \{x \in [16.9, 27.1], etime \in [26.9, 27.1], etime - x \in [0, 10]\})\}.$$

As S'_2 has an empty intersection with $S_{\mathcal{A}_\varphi}^{ne}$, the verdict is $\perp$.

6 Implementation

We implemented the methods described in this paper in the tool MONITAAL⁴ written in C++. This includes the Difference Bounded Matrix data structure to handle clock-zones, parsing property automata modelled in UPPAAL, computing the set of nonempty language states, computing the reach-sets in an online fashion over an observed word based on latency and jitter bounds in $[0, \infty[$, providing verdicts $\top$, $\perp$ or $?$ and latency bounds consistent with $\top$ and $\perp$.

We demonstrate MONITAAL on a trace generated by the gear controller model from [14]. The trace is generated in UPPAAL [8] and the monitored properties are six response properties derived from error locations in the model. We show results with and without delay consisting of a latency in $[0, 100]$ jitter bounded by 5. The results in Table 1 show the number of observations, length (in time) of the observed word, the over running time, the maximal response time (the time it takes to process a single observation and return the next verdict), and the maximal number of stored symbolic states. Under delay we see that the state storage grows linearly with the number of observations, which in turn results in a growing overhead when computing the reach-sets after each observation. The reason for this is the uncertainty of the delay increases the size of the reach-set. Nevertheless, the maximal response time is in all cases less than 1.2 ms. In the delay-free case, the memory usage is constant.

7 Conclusion

We have introduced a zone-based algorithm realizing optimal (in the sense of being anticipating [7]) online operational monitoring of embedded real-time systems when the communication between the monitor and the

⁴ <https://github.com/DEIS-Tools/MoniTAal>

Table 1: Results for simultaneously monitoring six response properties over a trace generated by the gear controller model from [14].

# Observ.	$\tau(\rho)$	Time (ms)		Max. resp. time (μs)		# Symbolic States	
		Delay	Delay-free	Delay	Delay-free	Delay	Delay-free
1000	63112	76	42	182	143	39	12
2000	124028	223	84	228	188	57	12
3000	184743	418	116	382	86	78	12
4000	244717	691	154	410	81	109	12
5000	306015	1037	198	571	281	135	12
6000	366814	1463	237	680	163	167	12
7000	438799	1973	278	767	175	192	12
8000	501070	2554	314	986	193	215	12
9000	563296	3212	357	1159	175	238	12
10000	624530	3929	384	1099	109	266	12

system is subject to unknown (up to bounds) delay. This situation is rather typical in practice as observations are mediated by sensors, may involve conversion between analog and digital, or pass communication networks and consequently are indirect in general, leading to delays and inexact time-stamping. Our constructions thus fill a gap in the pre-existing theories for monitoring hard real-time systems, which tend to assume full and exact temporal observability by immediate coupling or, equivalently, perfect synchrony between systems and their monitors.

A notable point of our construction is that it applies a reduction to simple timed automata and is purely zone-based despite the unknown communication delay being a timing parameter. The construction thus not only avoids the complexities of property analysis for parameterized timed automata [4], but also provides an instance of monitoring under uncertainty where the underlying arithmetic constraint systems remain of fixed dimensionality (namely the number of clocks in the property automata plus two for monitoring) despite their history dependence. This is in stark contrast to direct constraint encodings growing linearly over history length as in [11].

References

1. Alur, R., Dill, D.L.: A theory of timed automata. *Theoretical Computer Science* **126**(2), 183–235 (1994). [https://doi.org/https://doi.org/10.1016/0304-3975\(94\)90010-8](https://doi.org/https://doi.org/10.1016/0304-3975(94)90010-8)
2. Alur, R., Feder, T., Henzinger, T.A.: The benefits of relaxing punctuality. *Journal of the ACM* **43**(1) (01 1996)
3. Alur, R., Henzinger, T.A., Vardi, M.Y.: Parametric real-time reasoning. In: Kosaraju, S.R., Johnson, D.S., Aggarwal, A. (eds.) *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing*, May 16–18, 1993, San Diego, CA, USA. pp. 592–601. ACM (1993). <https://doi.org/10.1145/167088.167242>, <https://doi.org/10.1145/167088.167242>
4. André, É., Lime, D., Roux, O.H.: Reachability and liveness in parametric timed automata. *Log. Methods Comput. Sci.* **18**(1) (2022). [https://doi.org/10.46298/lmcs-18\(1:31\)2022](https://doi.org/10.46298/lmcs-18(1:31)2022)
5. Baldor, K., Niu, J.: Monitoring dense-time, continuous-semantics, metric temporal logic. In: *Runtime Verification*. pp. 245–259. Springer Berlin Heidelberg (2013). https://doi.org/10.1007/978-3-642-35632-2_24
6. Basin, D., Klaedtke, F., Zălinescu, E.: Algorithms for monitoring real-time properties. In: *RV*. pp. 260–275. Springer (2012). https://doi.org/10.1007/978-3-642-29860-8_20
7. Bauer, A., Leucker, M., Schallhart, C.: Monitoring of real-time properties. In: Arun-Kumar, S., Garg, N. (eds.) *FSTTCS 2006: Foundations of Software Technology and Theoretical Computer Science*. pp. 260–272. Springer, Berlin, Heidelberg (2006). https://doi.org/10.1007/11944836_25
8. Behrmann, G., David, A., Larsen, K.G.: A tutorial on uppaal. In: Bernardo, M., Corradini, F. (eds.) *Formal Methods for the Design of Real-Time Systems, International School on Formal Methods for the Design of Computer, Communication and Software Systems, SFM-RT 2004*, Bertinoro, Italy, September 13–18, 2004, Revised

- Lectures. LNCS, vol. 3185, pp. 200–236. Springer (2004). https://doi.org/10.1007/978-3-540-30080-9_7, https://doi.org/10.1007/978-3-540-30080-9_7
9. Bengtsson, J., Yi, W.: Timed automata: Semantics, algorithms and tools. In: Lectures on Concurrency and Petri Nets, Advances in Petri Nets. LNCS, vol. 3098, pp. 87–124. Springer (2003). https://doi.org/10.1007/978-3-540-27755-2_3
 10. Brihaye, T., Geeraerts, G., Ho, H.M., Monmege, B.: MightyL: A compositional translation from MITL to timed automata. In: Computer Aided Verification. pp. 421–440. Springer (2017). https://doi.org/10.1007/978-3-319-63387-9_21
 11. Finkbeiner, B., Fränzle, M., Kohn, F., Kröger, P.: A truly robust signal temporal logic: Monitoring safety properties of interacting cyber-physical systems under uncertain observation. *Algorithms* **15**(4), 126 (2022). <https://doi.org/10.3390/a15040126>
 12. Grosen, T.M., Kauffman, S., Larsen, K.G., Zimmermann, M.: Monitoring timed properties (revisited). In: Bogomolov, S., Parker, D. (eds.) Formal Modeling and Analysis of Timed Systems - 20th International Conference, FORMATS 2022, Warsaw, Poland, September 13-15, 2022, Proceedings. LNCS, vol. 13465, pp. 43–62. Springer (2022). https://doi.org/10.1007/978-3-031-15839-1_3
 13. Ho, H.M., Ouaknine, J., Worrell, J.: Online monitoring of metric temporal logic. In: Runtime Verification. pp. 178–192. Springer (2014). https://doi.org/10.1007/978-3-319-11164-3_15
 14. Lindahl, M., Pettersson, P., Yi, W.: Formal Design and Analysis of a Gearbox Controller. *Springer International Journal of Software Tools for Technology Transfer (STTT)* **3**(3), 353–368 (2001)
 15. Sheldon, T.: McGraw-Hill’s Encyclopedia of Networking and Telecommunications. McGraw-Hill Professional (2001)
 16. Thati, P., Rosu, G.: Monitoring algorithms for metric temporal logic specifications. In: Havelund, K., Rosu, G. (eds.) Proceedings of the Fourth Workshop on Runtime Verification, RV@ETAPS 2004, Barcelona, Spain, April 3, 2004. Electronic Notes in Theoretical Computer Science, vol. 113, pp. 145–162. Elsevier (2004). <https://doi.org/10.1016/J.ENTCS.2004.01.029>
 17. Ulus, D., Ferrère, T., Asarin, E., Maler, O.: Timed pattern matching. In: Formal Modeling and Analysis of Timed Systems. pp. 222–236. Springer (2014). https://doi.org/10.1007/978-3-319-10512-3_16
 18. Ulus, D., Ferrère, T., Asarin, E., Maler, O.: Online timed pattern matching using derivatives. In: Tools and Algorithms for the Construction and Analysis of Systems. pp. 736–751. Springer (2016). https://doi.org/10.1007/978-3-662-49674-9_47
 19. Waga, M., André, É., Hasuo, I.: Model-bounded monitoring of hybrid systems. *ACM Trans. Cyber Phys. Syst.* **6**(4), 30:1–30:26 (2022). <https://doi.org/10.1145/3529095>, <https://doi.org/10.1145/3529095>

A Proofs Omitted in the Main Part

Throughout the appendix, we often need to “shift” a timed word in the sense that we add or subtract a $d \in \mathbb{R}_{\geq 0}$ to each time point of ρ . In the latter case, we need to be careful to ensure that the time points in the shifted word are still nonnegative. For the sake of readability, let us introduce some notation for these operations. Given a (finite or infinite) timed word $\rho = (\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots$ and such a $d \in \mathbb{R}_{\geq 0}$,

- let $\rho + d$ denote the timed word $(\sigma_1, \tau_1 + d), (\sigma_2, \tau_2 + d), \dots$, and
- if $d \leq \tau_1$, let $\rho - d$ denote the timed word $(\sigma_1, \tau_1 - d), (\sigma_2, \tau_2 - d), \dots$. This is well-defined, as we require that τ_1 (and therefore each τ_i) is at least d , so we never obtain negative time points in $\rho - d$.

The following properties follow directly from the definition of timed concatenation and will be applied in the proofs below.

Remark 1. Let $\rho = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n) \in T\Sigma^*$, let $\mu = (\sigma'_1, \tau'_1), (\sigma'_2, \tau'_2), \dots \in T\Sigma^\omega$, and let $t \leq \tau(\rho)$.

1. Let $t' \in \mathbb{R}_{\geq 0}$ be such that $t - t' \geq \tau(\rho)$. Then

$$\rho \cdot_t \mu = \rho \cdot_{t-t'} (\mu + t').$$

2. Let $0 \leq t' \leq t$, let $n' \in \{0, 1, \dots, n\}$ be such that $\tau_{n'} \leq t' \leq \tau_{n'+1}$ (were we use $t_0 = -\infty$ to allow $n' = 0$ and $\tau_{n+1} = \infty$ to allow $n' = n$) and define $\rho_1 = (\sigma_1, \tau_1), \dots, (\sigma_{n'}, \tau_{n'})$ as well as $\rho_2 = (\sigma_{n'+1}, \tau_{n'+1}), \dots, (\sigma_n, \tau_n)$. Then

$$\rho \cdot_t \mu = \rho_1 \cdot_{t'} ((\rho_2 - t') \cdot_{t-t'} \mu).$$

3. Let $t' \geq 0$ be such that $\tau(\rho) \leq t - t'$, let $n' \geq 1$ be such that $\tau_{n'}' \leq t' \leq \tau_{n'+1}'$ (this is well-defined due to time-divergence), and define $\rho' = (\sigma'_1, \tau'_1), \dots, (\sigma'_{n'}, \tau'_{n'})$ as well as $\mu' = (\sigma'_{n'+1}, \tau'_{n'+1}), (\sigma'_{n'+2}, \tau'_{n'+2}), \dots$. Then

$$\rho \cdot_{t-t'} \mu = (\rho \cdot_{t-t'} \rho') \cdot_t (\mu' - t').$$

Furthermore, the following properties about consistent words will be useful in the proofs below.

Lemma 7.

1. Let $(\rho_1^*, t_1) \sqsubseteq (\rho_2^*, t_2)$, let $\rho_2 = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n) \in GT_{\delta, \varepsilon}(\rho_2^*, t_2)$, let $n' \in \{0, 1, \dots, n\}$ be such that $\tau_{n'} \leq t_1 \leq \tau_{n'+1}$ (were we use $t_0 = -\infty$ to allow $n' = 0$ and $\tau_{n+1} = \infty$ to allow $n' = n$), and define $\rho_1 = (\sigma_1, \tau_1), \dots, (\sigma_{n'}, \tau_{n'})$. Then $\rho_1 \in GT_{\delta, \varepsilon}(\rho_1^*, t_1)$.
2. Let $\rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$, and let ρ' be a finite timed word with $\tau(\rho') \leq t - \max(\tau(\rho), t - (\delta + \varepsilon))$. Then, $\rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \rho' \in GT_{\delta, \varepsilon}(\rho^*, t)$.
3. Let $\rho \in GT_{\delta, \varepsilon}(\rho^*, t)$, let ρ have m letters, and let ρ' be the prefix of ρ with m letters. Then, $\rho' \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$.

Proof. 1.) We need to show that ρ_1 is consistent with ρ_1^* at t_1 under δ and ε . The first requirement of the definition of consistency follows from $\tau(\rho_1^*) \leq t_1$ and the choice of n' (which implies $\tau(\rho_1) \leq t_1$). The second requirement follows from the fact that ρ_2 is consistent with ρ_2^* at t_2 under δ and ε and the fact that ρ_1 is a prefix of ρ_2 and ρ_1^* is a prefix of ρ_2^* . Finally, consider the third requirement and assume it is violated, i.e., let ρ_1^* have m letters and assume ρ_1 has at least $m + 1$ letters such that the time point τ_{m+1} of the $(m + 1)$ -st letter of ρ_1 satisfies $\tau_{m+1} + \delta + \varepsilon < t_1$. Then, as ρ_2 is consistent with ρ_2^* at t_2 under δ and ε , we obtain a contradiction. Either ρ_2^* has also m letters. In this case, $\tau_{m+1} + \delta + \varepsilon < t_1 \leq t_2$ implies that the third requirement of the definition of consistency is violated for ρ_2 and ρ_2^* . Otherwise, ρ_2^* has at least $m + 1$ letters. In this case $\tau_{m+1} + \delta + \varepsilon < t_1 \leq t_2$ implies that the second requirement of the definition of consistency is violated for $i = m + 1$.

2.) We have to show that $\rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \rho'$ is consistent with ρ^* at t under δ and ε . This follows directly from the fact that all events in ρ' have time points (in $\rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \rho'$) in the interval $[t - (\delta + \varepsilon), t]$ and are therefore covered by the third requirement of the definition of consistency.

3.) We need to show that ρ' is EL-consistent with ρ^* at t under δ and ε . By definition, ρ' has the same length as ρ^* and the first two requirements of the definition of consistency are satisfied, as ρ is consistent with ρ^* at t under δ and ε and ρ' is a prefix of ρ . Hence, it is EL-consistent, as the third requirement only refers to ground-truth that have more letters than the observation. $\square$

Now we are ready to present the proofs omitted in the main part.

Proof of Lemma 1

Recall that we need to show $\mathcal{V}_{\mathcal{D}'}(L)(\rho^*, t) = \top$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$ and that $\mathcal{V}_{\mathcal{D}'}(L)(\rho^*, t) = \perp$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$ for $\mathcal{D} \subseteq \mathcal{D}'$.

Proof. Note that $\mathcal{D} \subseteq \mathcal{D}'$ implies $GT_{\mathcal{D}}(\rho^*, t) \subseteq GT_{\mathcal{D}'}(\rho^*, t)$. Thus, the universal quantification over possible ground-truths ρ in the first two cases of the definition of $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t)$ ranges over a subset of the possible ground-truths that are considered for $\mathcal{V}_{\mathcal{D}'}(L)(\rho^*, t)$. Hence, the result follows. $\square$

Proof of Lemma 2

Recall that we need to show

1. $\Delta_{\mathcal{D}}(L, \rho^*, t) = \emptyset$ iff $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$, and
2. $\Delta_{\mathcal{D}}(\bar{L}, \rho^*, t) = \emptyset$ iff $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$.

Proof. We have

$$\begin{aligned} \Delta_{\mathcal{D}}(L, \rho^*, t) &= \emptyset \\ \Leftrightarrow \rho \cdot_t \mu &\in \bar{L} \text{ for all } (\delta, \varepsilon) \in \mathcal{D}, \text{ all } \rho \in GT_{\delta, \varepsilon}(\rho^*, t), \text{ and all } \mu \in T\Sigma^\omega \\ \Leftrightarrow \rho \cdot_t \mu &\in \bar{L} \text{ for all } \rho \in GT_{\mathcal{D}}(\rho^*, t) \text{ and all } \mu \in T\Sigma^\omega \\ \Leftrightarrow \mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) &= \perp. \end{aligned}$$

The second claim is obtained by a dual argument (swapping $\perp$ with $\top$ and L with $\bar{L}$). $\square$

Proof of Lemma 3

Recall that we need to show $\Delta(L, \rho_1^*, t_1) \supseteq \Delta(L, \rho_2^*, t_2)$ for all $(\rho_1^*, t_1) \sqsubseteq (\rho_2^*, t_2)$.

Proof. Let $(\delta, \varepsilon) \in \Delta(L, \rho_2^*, t_2)$, i.e., there exists $\rho_2 \in GT_{\delta, \varepsilon}(\rho_2^*, t_2)$ and a $\mu_2 \in T\Sigma^\omega$ such that $\rho_2 \cdot_{t_2} \mu_2 \in L$.

Let $\rho_2 = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ and let n' be maximal with $\tau_{n'} \leq t_1$. Then, Lemma 7.1 yields

$$\rho_1 = (\sigma_1, \tau_1), \dots, (\sigma_{n'}, \tau_{n'}) \in GT_{\delta, \varepsilon}(\rho_1^*, t_1)$$

and an application of Remark 1.2 yields

$$\rho_1 \cdot_{t_1} \left([((\sigma_{n'+1}, \tau_{n'+1}), \dots, (\sigma_n, \tau_n)) - t_1] \cdot_{t_2-t_1} \mu \right) = \rho_2 \cdot_{t_2} \mu_2 \in L.$$

This implies $(\delta, \varepsilon) \in \Delta(L, \rho_1^*, t_1)$. $\square$

Proof of Lemma 4

Recall that we need to show

1. $\Delta_{\mathcal{D}}(L, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ implies there is no $\rho_1^* \in T\Sigma^*$ such that $\mathcal{V}_{\mathcal{D}}(L)(\rho^* \cdot_t \rho_1^*, t') = \top$ for any $t' \geq t + \tau(\rho_1^*)$, and
2. $\Delta_{\mathcal{D}}(\bar{L}, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$ implies there is no $\rho_1^* \in T\Sigma^*$ such that $\mathcal{V}_{\mathcal{D}}(L)(\rho^* \cdot_t \rho_1^*, t') = \perp$ for any $t' \geq t + \tau(\rho_1^*)$.

for all $t \geq \tau(\rho^*)$.

Proof. Let $\Delta_{\mathcal{D}}(L, \rho^*, t) \subsetneq \mathcal{D} \cap \{(\delta, \varepsilon) \mid \delta \leq \tau_1^*\}$, i.e., there is a $(\delta, \varepsilon) \in \mathcal{D}$ with $\delta \leq \tau_1^*$ and $(\delta, \varepsilon) \notin \Delta_{\mathcal{D}}(L, \rho^*, t)$. Thus, by definition, for all $\rho \in GT_{\delta, \varepsilon}(\rho^*, t)$ and all $\mu \in T\Sigma^\omega$ we have $\rho \cdot_t \mu \in \bar{L}$ ($\dagger$).

Now, let $\rho_1^* \in T\Sigma^*$ and assume, towards a contradiction, we have $\mathcal{V}_{\mathcal{D}}(L)(\rho^* \cdot_t \rho_1^*, t') = \top$ for $t' \geq t + \tau(\rho_1^*)$, i.e., for all $\rho' \in GT_{\mathcal{D}}(\rho^* \cdot_t \rho_1^*, t')$ and all $\mu \in T\Sigma^\omega$ we have $\rho' \cdot_{t'} \mu \in L$ ($\dagger\dagger$). As $GT_{\delta, \varepsilon}(\rho^* \cdot_t \rho_1^*, t')$ is nonempty, let us fix one such $\rho' \in GT_{\delta, \varepsilon}(\rho^* \cdot_t \rho_1^*, t')$. Also, let us fix a $\mu \in T\Sigma^\omega$.

Note that we have $(\rho^*, t) \sqsubseteq (\rho^* \cdot_t \rho_1^*, t')$. So, let $\rho' = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ and let n' be maximal with $\tau_{n'} \leq t$. Then, Lemma 7.1 yields

$$\rho'_1 = (\sigma_1, \tau_1), \dots, (\sigma_{n'}, \tau_{n'}) \in GT_{\delta, \varepsilon}(\rho^*, t)$$

and application of Remark 1.2 yields

$$\rho' \cdot_{t'} \mu = \rho'_1 \cdot_t \underbrace{\left([((\sigma_{n'+1}, \tau_{n'+1}), \dots, (\sigma_n, \tau_n)) - t] \cdot_{t'-t} \mu \right)}_{=\mu'}.$$

This yields the desired contradiction, as $\rho' \cdot_t \mu$ is in L (see $\dagger\dagger$) while $\rho'_1 \cdot_t \mu'$ is not in L (see $\dagger$).

The second claim is proven by a dual argument (swapping L with $\bar{L}$ and $\top$ with $\perp$). $\square$

Proof of Lemma 5

Recall that we need to show $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \mathcal{V}_{\mathcal{D}}(L)(\rho^*, t)$.

Proof. Let $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$. We show $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \top$ by proving that we have $\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu' \in L$ for all $\rho' \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$ for some $(\delta, \varepsilon) \in \mathcal{D}$ and all $\mu' = (\sigma_1, \tau_1), (\sigma_2, \tau_2), \dots \in T\Sigma^\omega$.

First, consider the case where $\tau(\rho') < t - (\delta + \varepsilon)$. Let n be maximal with $\tau_n \leq \delta + \varepsilon$ (this is well-defined due to time-divergence), let $\rho'_1 = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n)$ and $\mu'_2 = ((\sigma_{n+1}, \tau_{n+1}), (\sigma_{n+2}, \tau_{n+2}), \dots) - (\delta + \varepsilon)$. Note that μ'_2 is well-defined as τ_{n+1} is, by the choice of n , greater than $\delta + \varepsilon$. Then, Lemma 7.2 yields $\rho' \cdot_{t - (\delta + \varepsilon)} \rho'_1$ is in $GT_{\delta, \varepsilon}(\rho^*, t)$ and Remark 1.3 yields

$$\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu' = \rho' \cdot_{t - (\delta + \varepsilon)} \mu' = (\rho' \cdot_{t - (\delta + \varepsilon)} \rho'_1) \cdot_t \mu'_2.$$

Therefore, $\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu'$ is the concatenation of the possible ground-truth $(\rho' \cdot_{t - (\delta + \varepsilon)} \rho'_1)$ of ρ^* and the suffix μ'_2 . As we have $\rho \cdot_t \mu \in L$ for all $\rho \in GT_{\mathcal{D}}(\rho^*, t)$ and all $\mu \in T\Sigma^\omega$ (due to $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$), we conclude $\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu' \in L$ as required.

Now, consider the case where $\tau(\rho') \geq t - (\delta + \varepsilon)$. Note that we have $t - \tau(\rho') \geq 0$ due to $\rho' \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$. Hence, let n be maximal with $\tau_n \leq t - \tau(\rho')$ (again, this is well-defined due to time-divergence), let $\rho'_1 = (\sigma_1, \tau_1) \cdots (\sigma_n, \tau_n)$ and $\mu'_2 = ((\sigma_{n+1}, \tau_{n+1}), (\sigma_{n+2}, \tau_{n+2}), \dots) - (t - \tau(\rho'))$. Then, Lemma 7.2 yields $\rho' \cdot_{\tau(\rho')} \rho'_1$ is in $GT_{\delta, \varepsilon}(\rho^*, t)$ and Remark 1.3 yields

$$\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu' = \rho' \cdot_{\tau(\rho')} \mu' = (\rho' \cdot_{\tau(\rho')} \rho'_1) \cdot_t \mu'_2.$$

As $\rho' \cdot_{\max(\tau(\rho'), t - (\delta + \varepsilon))} \mu'$ is the concatenation of a possible ground-truth of ρ^* and an arbitrary suffix, it is again, as required, in L .

Using a dual argument (i.e., swapping $\top$ with $\perp$ and L with $\bar{L}$), we can show that $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$ implies $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \perp$.

Now, we show that $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \top$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$. A dual argument again shows that $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \perp$ implies $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \perp$. This will then complete our proof, as both functions only have three elements in their codomain and we have shown that two of them have the same preimage w.r.t. both functions.

So, let $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \top$. We show $\mathcal{V}_{\mathcal{D}}(L)(\rho^*, t) = \top$ by showing $\rho \cdot_t \mu \in L$ for all $\rho = (\sigma_1, \tau_1), \dots, (\sigma_n, \tau_n) \in GT_{\mathcal{D}}(\rho^*, t)$ and all $\mu = (\sigma'_1, \tau'_1), (\sigma'_2, \tau'_2), \dots \in T\Sigma^\omega$. By definition, there is a $(\delta, \varepsilon) \in \mathcal{D}$ such that $\rho \in GT_{\delta, \varepsilon}(\rho^*, t)$.

Let ρ^* have m letters. If $m = n$, then we also have $\rho \in GT_{\mathcal{D}}^{\text{el}}(\rho^*, t)$. We consider two cases: If $\tau(\rho) < t - (\delta + \varepsilon)$, then an application of Remark 1.1 yields

$$\rho \cdot_t \mu = \rho \cdot_{t - (\delta + \varepsilon)} (\mu + (\delta + \varepsilon)) = \rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} (\mu + (\delta + \varepsilon)),$$

and if $\tau(\rho) \geq t - (\delta + \varepsilon)$, then an application of Remark 1.1 yields

$$\rho \cdot_t \mu = \rho \cdot_{\tau(\rho)} (\mu + (t - \tau(\rho))) = \rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} (\mu + (t - \tau(\rho))),$$

where $t - \tau(\rho)$ is nonnegative by definition of consistency. Hence, in both cases, $\rho \cdot_t \mu$ is the concatenation of a possible EL ground-truth of ρ^* and an arbitrary suffix. Due to $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L)(\rho^*, t) = \top$, all concatenations $\rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \mu$ for $(\delta, \varepsilon) \in \mathcal{D}$, $\rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$, and $\mu \in T\Sigma^\omega$ are in L , which yields $\rho \cdot_t \mu \in L$.

It remains to consider the case where $n > m$, which we again split into two subcases. But first let us define $\rho_1 = (\sigma_1, \tau_1) \cdots (\sigma_m, \tau_m)$ as well as $\rho_2 = (\sigma_{m+1}, \tau_{m+1}) \cdots (\sigma_n, \tau_n)$. Lemma 7.3 yields $\rho_1 \in GT_{\mathcal{D}}^{\text{el}}(\rho^*, t)$.

First, consider the subcase where $\tau(\rho_1) < t - (\delta + \varepsilon)$. By definition of consistency, $n > m$ implies $\tau_{m+1} + \delta + \varepsilon \geq t$, and thus $\tau_{m+1} \geq t - (\delta + \varepsilon)$ ($\dagger$). Hence, an application of Remark 1.2 yields

$$\begin{aligned} \rho \cdot_t \mu &= \rho_1 \cdot_{t - (\delta + \varepsilon)} [(\rho_2 - (t - (\delta + \varepsilon))) \cdot_{\delta + \varepsilon} \mu] \\ &= \rho_1 \cdot_{\max(\tau(\rho_1), t - (\delta + \varepsilon))} [(\rho_2 - (t - (\delta + \varepsilon))) \cdot_{\delta + \varepsilon} \mu] \end{aligned}$$

Note that the first time point of ρ_2 , τ_{m+1} , is at least $(t - (\delta + \varepsilon))$ as required, as $\tau_{m+1} \geq t - (\delta + \varepsilon)$ (see $\dagger$). Hence, $\rho \cdot_t \mu$ is the concatenation of a possible EL ground-truth of ρ^* and an arbitrary suffix and therefore in L .

Finally, consider the subcase where $\tau(\rho_1) \geq t - (\delta + \varepsilon)$. Then, an application of Remark 1.2 yields

$$\begin{aligned} \rho \cdot_t \mu &= \rho_1 \cdot_{\tau(\rho_1)} [(\rho_2 - \tau(\rho_1)) \cdot_{t - \tau(\rho_1)} \mu] \\ &= \rho_1 \cdot_{\max(\tau(\rho_1), t - (\delta + \varepsilon))} [(\rho_2 - \tau(\rho_1)) \cdot_{t - \tau(\rho_1)} \mu]. \end{aligned}$$

Again, this is well-defined as we have $\tau_{m+1} \geq \tau(\rho_1)$ (as τ_{m+1} is the next time instant after $\tau_m = \tau(\rho_1)$ in ρ) and as $t \geq \tau(\rho_1)$ by the definition of consistency. Hence, $\rho \cdot_t \mu$ is again the concatenation of a possible EL ground-truth of ρ^* and an arbitrary suffix and therefore in L . $\square$

Proof of Theorem 2

Recall that we need to show $\mathcal{M}_{\mathcal{D}}(\mathcal{A}, \bar{\mathcal{A}})(\rho^*, t) = \mathcal{V}_{\mathcal{D}}^{\text{el}}(L(\mathcal{A}))(\rho^*, t)$.

Proof. We will show that $\mathcal{R}_{\mathcal{A}'}^{\mathcal{D}}(\rho^*, t) \cap S_{\mathcal{A}'}^{ne}$ is nonempty iff there exists a $\rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$ and a $\mu \in T\Sigma^\omega$ with $\rho \cdot_{\max(\tau(\rho), t - (\delta + \varepsilon))} \mu \in L(\mathcal{A}')$ for any TBA $\mathcal{A}'$. Then we obtain

- $\mathcal{M}_{\mathcal{D}}(\mathcal{A}, \bar{\mathcal{A}})(\rho^*, t) = \top$ iff $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L(\mathcal{A}))(\rho^*, t) = \top$ by instantiating the equivalence for $\mathcal{A}' = \bar{\mathcal{A}}$, and
- $\mathcal{M}_{\mathcal{D}}(\mathcal{A}, \bar{\mathcal{A}})(\rho^*, t) = \perp$ iff $\mathcal{V}_{\mathcal{D}}^{\text{el}}(L(\mathcal{A}))(\rho^*, t) = \perp$ by instantiating the equivalence for $\mathcal{A}' = \mathcal{A}$.

This completes the proof, as both functions only have three elements in their codomain and we have shown that two of them have the same preimage w.r.t. both functions.

So, let $\mathcal{R}_{\mathcal{A}'}^{\mathcal{D}}(\rho^*, t) \cap S_{\mathcal{A}'}^{ne} \neq \emptyset$. Then, by definition, there is a state (q, v') of $\mathcal{A}'$ such that

- $(q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}'} (q, v)$ for some initial state (q_0, v_0) of $\mathcal{A}'$, some $\rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$ for some $(\delta, \varepsilon) \in \mathcal{D}$, and $v' = v + \max(0, (t - (\tau(\rho) + \delta + \varepsilon)))$, and
- and there is an accepting infinite run of $\mathcal{A}'$ starting in (q, v') that processes some $\mu \in T\Sigma^\omega$.

These two runs can be combined into an accepting run of $\mathcal{A}'$ that starts in (q_0, v_0) and processes

$$\rho \cdot \tau(\rho) + \max(0, (t - (\tau(\rho) + \delta + \varepsilon))) \mu = \rho \cdot \max(\tau(\rho), (t - (\delta + \varepsilon))) \mu,$$

which implies that it is in $L(\mathcal{A}')$ as required.

Conversely, let there be a $\rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t)$ and a $\mu \in T\Sigma^\omega$ with

$$\mu \cdot \max(\tau(\rho), (t - (\delta + \varepsilon))) \mu \in L(\mathcal{A}').$$

Then, there exists an accepting run of $\mathcal{A}'$ starting in some initial state (q_0, v_0) that processes $\rho \cdot \max(\tau(\rho), (t - (\delta + \varepsilon))) \mu$. This run can be split into

- $(q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}'} (q, v)$ for some state (q, v) of $\mathcal{A}'$ and
- an accepting infinite run of $\mathcal{A}'$ starting in (q, v') that processes μ , where

$$v' = v + \max(0, (t - (\tau(\rho) + \delta + \varepsilon))).$$

Hence, $(q, v') \in \mathcal{R}_{\mathcal{A}'}^{\mathcal{D}}(\rho^*, t) \cap S_{\mathcal{A}'}^{ne}$, which is therefore, as required, nonempty. $\square$

Proof of Theorem 3

Below we give a sketch of a proof. Let $\rho^* = (\sigma_1^*, \tau_1^*), \dots, (\sigma_n^*, \tau_n^*)$ be an observed timed word. We want to show that

$$\mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \{(q', Z') \mid (q', Z'') \in S_n, Z' = Z''^{\nearrow} \wedge etime = t - \varepsilon\} \quad (1)$$

where $\mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t)$ is defined by

$$\begin{aligned} \mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \{ & (q, v + \max(0, (t - (\tau(\rho) + \delta + \varepsilon)))) \mid (q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}} (q, v) \text{ where} \\ & (q_0, v_0) \text{ with } q_0 \in Q_0, v_0(x) = 0 \text{ for all } c \in C, \text{ and} \\ & \rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t) \text{ for some } (\delta, \varepsilon) \in \mathcal{D}\}. \end{aligned}$$

in the setting where $\mathcal{D} = \{(\delta, \varepsilon) \mid \delta \in [l, u]\}$ for given $l, u, \varepsilon \in \mathbb{Q}_{\geq 0}$. Also, recall that we have defined $S_0 = \{(q_0, Z_0^d) \mid q_0 \in Q_0\}$ and $S_i = \text{Succ}_d(S_{i-1}, (\sigma_i, \tau_i^*))$ for $i \in \{1, \dots, n\}$.

Proof. Given the form of $\mathcal{D}$ we can rewrite the definition of the reach-set to

$$\begin{aligned} \mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \{ & (q, v + \max(0, (t - (\tau(\rho) + \delta + \varepsilon)))) \mid (q_0, v_0) \xrightarrow{\rho}_{\mathcal{A}} (q, v) \text{ where} \\ & (q_0, v_0) \text{ with } q_0 \in Q_0, v_0(x) = 0 \text{ for all } c \in C, \text{ and} \\ & \rho \in GT_{\delta, \varepsilon}^{\text{el}}(\rho^*, t) \text{ for some } \delta \in [l, u]\}. \end{aligned}$$

Now, extending the transition relation $\xrightarrow{\rho}_{\mathcal{A}}$ to clock valuations over the extended set of clocks $C \cup \{time, etime\}$, we may further reformulate the reach-set as follows:

$$\begin{aligned} \mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \{ & (q_n, v^* + \max(0, (t - (v_n^*(etime) + \varepsilon)))) \mid (q_0, v_0^*) \xrightarrow{\rho}_{\mathcal{A}} (q_n, v_n^*) \text{ where} \\ & (q_0, v_0^*) \text{ with } q_0 \in Q_0, v_0^*(x) = 0 \text{ for all } c \in C, \text{ and} \\ & v_0^*(etime) - v_0^*(time) \in [l, u] \text{ and} \\ & v_i^*(etime) \leq \tau_i^* \wedge v_i^*(etime) \geq \tau_i^* - \varepsilon \wedge \sigma_i = \sigma_i^* \text{ for } i \in \{0, \dots, n\}\}. \end{aligned}$$

A key observation for the correctness of the above reformulation, is that the extended clocks *etime* and *time* are *not* modified by the TBA $\mathcal{A}$. That is $v_i^*(etime) - v_i^*(time) = v_0^*(etime) - v_0^*(time) \in [l, u]$ for all $i = \{0, \dots, n\}$.

Now let $\mathcal{R}_{\mathcal{A}}^{\mathcal{D},j}(\rho^*, t)$ for $j \in \{0, \dots, n\}$ be defined as follows:

$$\begin{aligned} \mathcal{R}_{\mathcal{A}}^{\mathcal{D},j}(\rho^*, t) = \{ & (q_j, v_j^*) \mid (q_0, v_0^*) \xrightarrow{\rho}_{\mathcal{A}} (q_j, v_j^*) \text{ where} \\ & (q_0, v_0^*) \text{ with } q_0 \in Q_0, v_0^*(x) = 0 \text{ for all } c \in C, \text{ and} \\ & v_0^*(etime) - v_0^*(time) \in [l, u] \text{ and} \\ & v_i^*(etime) \leq \tau_i^* \wedge v_i^*(etime) \geq \tau_i^* - \varepsilon \wedge \sigma_i = \sigma_i^* \text{ for } i \in \{0, \dots, j\} \}. \end{aligned}$$

Then clearly $\mathcal{R}_{\mathcal{A}}^{\mathcal{D},0}(\rho^*, t) = \{(q_0, Z_0^d) \mid q_0 \in Q_0\} = S_0$ and $\mathcal{R}_{\mathcal{A}}^{\mathcal{D},j}(\rho^*, t) = \text{Succ}_d(\mathcal{R}_{\mathcal{A}}^{\mathcal{D},j-1}(\rho^*, t), (\sigma_i, \tau_i))$ for $j \in \{1, \dots, n\}$, using standard arguments for the correctness of symbolic exploration of timed automata, e.g. [9]. Finally, as $(t - (v_n^*(etime) + \varepsilon)) = ((t - \varepsilon) - v_n^*(etime))$ it follows that $\mathcal{R}_{\mathcal{A}}^{\mathcal{D}}(\rho^*, t) = \mathcal{R}_{\mathcal{A}}^{\mathcal{D},n}(\rho^*, t) \nearrow \wedge etime = t - \varepsilon$. $\square$

B More Details on the Implementation

MoniTAal requires, as input, two automata accepting the complement language of the other, a series of observations, and optionally bounds on latency and jitter. MoniTAal can be run as a binary where the automata are parsed from a UPPAAL xml file while the observations are parsed as text from a file or standard input. MoniTAal can also be used as a C++ library.

In Listing 1.1 we give a short demonstration going through Example 3 using MoniTAal. In the output we see that the consistent latency for the $\top$ verdict (POSITIVE) verdict tightens to $[71, 100]$ and then $[71, 75]$ while the final verdict is $?$ (INCONCLUSIVE). Note that we multiply all values (in observations and automata) by 10 in order to use integer, rather than rational, time points.

```

1 $ ./MoniTAal-bin -p positive delay.xml -n negative delay.xml -v --latencyl 0 --latencyu 10 --jitter 2
2 Input: @[173, 173] a
3
4 Verdict: INCONCLUSIVE
5 Positive:
6 Consistent latencies: {[71,100]}
7 Jitter bound: 2
8 Negative:
9 Consistent latencies: {[0,100]}
10 Jitter bound: 2
11
12 Input: @[271, 271] b
13
14 Verdict: INCONCLUSIVE
15 Positive:
16 Consistent latencies: {[71,75]}
17 Jitter bound: 2
18 Negative:
19 Consistent latencies: {[0,100]}
20 Jitter bound: 2

```

Listing 1.1: Demonstration of MoniTAal over Example 3.