Efficient Parametric Identification for STL

Alexey Bakhirkin  
Université Grenoble Alpes, VERIMAG  
F-38000 Grenoble, France

Thomas Ferrère  
IST Austria  
A-3400 Klosterneuburg, Austria

Oded Maler  
Université Grenoble Alpes, VERIMAG  
CNRS, VERIMAG  
F-38000 Grenoble, France

ABSTRACT

We describe a new algorithm for the parametric identification problem for signal temporal logic (STL), stated as follows. Given a dense-time real-valued signal $w$ and a parameterized temporal logic formula $\phi$, compute the subset of the parameter space that renders the formula satisfied by the signal. Unlike previous solutions, which were based on search in the parameter space or quantifier elimination, our procedure works recursively on $\phi$ and computes the evolution over time of the set of valid parameter assignments. This procedure is similar to that of monitoring or computing the robustness of $\phi$ relative to $w$. Our implementation and experiments demonstrate that this approach can work well in practice.

CCS CONCEPTS

• Computing methodologies → Simulation evaluation;  
• Theory of computation → Modal and temporal logics;

ACM Reference Format:

ACKNOWLEDGMENTS

This work was partially supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement nr. 306595 "STATOR", and by the Austrian Science Fund (FWF) under grant S11402-N23 "RiSE / SHiNE".

1 INTRODUCTION

Signal temporal logic (STL) [33, 35] is an extension of temporal logic designed to handle real-valued dense-time signals which gained a lot of popularity in recent years as a rigorous and expressive formalism to describe behaviors of continuous and hybrid systems in various domains such as analog circuits [28], systems and synthetic biology [8, 17, 40], biomedical systems [13, 14] and cyber-physical control systems [15, 19, 30, 36, 37]. The reader is referred to [9] for an introduction and a survey of applications.

The major use of STL is in monitoring: a signal $w$, which is the output of a simulator or a sequence of measurements from a real system, is observed, and a monitoring procedure then checks whether it satisfies an STL formula $\phi$, a fact denoted as $w \models \phi$. The procedure for checking satisfaction works along two dimensions, one is related to the parse tree of the formula and one to time. For each sub-formula $\phi'$ a satisfaction signal is computed whose value at $t$ indicates whether $w$ satisfies $\phi'$ from $t$. The computation of satisfaction of a formula like $\diamond [a, b] \phi$ (eventually $\phi$ within $r \in [a, b]$ time) at $t$ is propagated backwards from the satisfaction of $\phi$ at the interval $[t + a, t + b]$.

The inverse problem of parametric identification for STL has been introduced in [6]. It uses PSTL, a parametric version of STL, admitting formulas where some of the constants in the numerical predicates and quantitative timing operators are replaced by parameters taken from $P = \{p_1, \ldots, p_k\}$ ranging over some parameter space $V \subseteq \mathbb{R}^k$. Each selection of parameter values $v \in V$ transforms a PSTL formula $\phi$ into an STL formula $\phi[v]$ that might be satisfied or not by a given signal $w$. The problem solved in [6] is to compute the validity domain of $\phi$ relative to $w$, that is, the set of parameter valuations $v$ such that $w \models \phi[v]$. This result provides an enabling technology for applying learning and data-mining techniques to observed behaviors of cyber-physical systems [10, 23, 25–27, 39, 41].

The major result of [6] states that when signals, presented as a sequence of time-stamped values, are interpreted as piecewise-linear, the validity domains are semi-linear sets. Two approaches were proposed to compute them. The first was based on quantifier elimination in linear arithmetic, for a formula whose size is linear in the number of sampling points of the signal. The other approach was approximate, based on conducting search over the parameter space. Under certain reasonable assumptions of monotonicity in parameter influence, the latter problem is equivalent to approximating the Pareto front in multi-criteria optimization, e.g. [31].

In this paper we propose an alternative approach which resembles the way monitoring is done for qualitative (satisfaction) and quantitative (robustness) semantics. Robustness was defined and computed for STL in [18, 20] following [21, 22] and [38] and its relation to parametric identification is worth discussing. The robustness $\rho(\phi, w)$ is a real number which is positive if and only if $w$ satisfies $\phi$. Moreover, all signals whose pointwise distance from $w$ is less than $\rho(\phi, w)$ have the same satisfaction status as $w$.

Consider the constraint $x \geq 0$ satisfied by some signal $w$ in which variable $x$ has value $c$. The robustness of this satisfaction is defined as $\rho(x \geq 0, w) = c$. On the other hand, the validity domain of the PSTL formula $x \geq p$ is $D(x \geq p, w) = \{v : v \leq c\}$. Consider now the formula $x \geq 0 \lor y \geq 0$ which depends on two variables. According to the common definition of robustness we will have $\rho(x \geq 0 \lor y \geq 0, w) = \max\{\rho(x \geq 0, w), \rho(y \geq 0, w)\}$.
a one-dimensional object that mixes the tolerances associated with the two variables. On the other hand, the validity domain for the PSTL formula $x \geq p \lor y \geq q$ is a two-dimensional object providing more refined information concerning the possible deformations of the signal that do not change satisfaction status.

In signal temporal logic, the above quantities are not constants but signals themselves. The major contribution of the paper is a new procedure to compute validity domains that follows the computation of satisfaction [33] and robustness [18] signals by propagating them as function of time, from sub-formulas to formulas. With each formula $\phi$ and a signal $w$ we associate a parametric validity signal whose value at $t$ indicates the set of parameter valuations $v$ such that $(w, t) \models \phi[v]$. The crucial component is to compute the validity domain of $\psi_{[a, b]} \phi$ from that of $\phi$. Like the robustness computation in [18], this involves aggregating values of a signal over a shifting window. However, we are dealing with a multi-dimensional partially-ordered parameter space, and the validity domains are typically Pareto-like sets.

Technically, we consider PSTL formulas with space parameters and signals which are interpreted using a piecewise-constant interpolation. In this special case, we show that time can be partitioned into finitely many intervals and the validity domain in each interval is a finite union of rectangles. We implemented the procedure for computing validity signals and demonstrate its performance on rather long signals. Potential extension to piecewise-linear signals and to timing parameters is discussed at the end of the paper.

2 PARAMETRIC SIGNAL TEMPORAL LOGIC

Parametric Signal Temporal Logic (PSTL) [6] extends the logic STL [33] with parameters. Itself, STL allows specifying properties of Boolean and real-valued signals, using atomic formulas of the form $x \geq c$, and temporal formulas of the form $\psi_{[a, b]} \phi$. Additionally, the logic provides Boolean connectives, and the temporal until operator. For the sake of simplicity we omit Boolean variables from the syntax. This is without loss of generality, Boolean variables can be seen as special real-valued variables taking their values in $\{0, 1\}$. We also only consider the case of closed timing intervals, other form of intervals do not pose additional difficulty.

A signal $w$ is a function $\mathbb{N} \rightarrow \mathbb{R}^n$ which we call time domain. The value $d$ is the duration of the signal, and we denote it by $|w|$. The value of signal $w$ at time $t \in T$ is denoted $w[t] \in \mathbb{R}^n$. Signal values are accessed by variables from the set $X = \{x_1, \ldots, x_n\}$. The projection of $w$ onto some variable $x \in X$ is denoted $w_x$. Using these conventions, $w_x[t]$ denotes the value of variable $x$ at time $t$ given in signal $w$.

Let $P = \{p_1, \ldots, p_k\}$ be the set of parameters. A parameter valuation is a vector $v$ that assigns a value to every parameter. We assume that parameter valuations range over a parameter space $V \subseteq \mathbb{R}^k$. The value of parameter $p \in P$ in a valuation $v \in V$ is denoted $v_p$. We use often logical notation to describe sets of parameter valuations. For example, we write $p_1 \geq 1 \land p_2 \leq 2$ to denote the set of parameter valuations $\{v \in V : v_{p_1} \geq 1 \land v_{p_2} \leq 2\}$.

Definition 2.1 (PSTL Syntax). Formulas $\phi$ of PSTL are described by the following grammar:

$$\phi ::= \text{true} \mid x \geq c \mid x \geq c \mid x \geq p \mid x \geq p \mid \neg \phi \mid \phi_1 \lor \phi_2 \mid \psi_{[a, b]} \phi \mid \phi_1 \mathcal{U} \phi_2$$

where $x \in X$ is a signal variable, $c \in \mathbb{R}$ is a constant, $p \in P$ is a parameter, and $0 \leq a \leq b \in \mathbb{R}$ are positive constants.

We also make use of standard abbreviations $x < c \equiv \neg(x \geq c)$, $\phi \land \psi \equiv \neg(\neg \phi \lor \neg \psi)$, and common temporal eventually and (timed) always defined as follows: $\phi \equiv \text{true} \lor \phi$, $\phi \equiv \neg \neg \phi$, and $\exists \phi \equiv \psi_{[0, a]} \phi \equiv \psi_{[a, b]} \neg \phi$.

The timed until operators $\mathcal{U}_{[a, b]}$ and $\mathcal{U}_{[a, \infty]}$ are often considered primitive, but can also be derived from the untimed until and bounded eventually as follows [33]:

$$\phi \mathcal{U}_{[a, b]} \psi \equiv \psi_{/[a, a]} \phi \mathcal{U} \psi$$

$$\phi \mathcal{U}_{[a, \infty]} \psi \equiv \psi_{/[a, b]} \phi \land (\psi \mathcal{U}_{[a, \infty]} \psi)$$

Finally, we introduce the release operator $\mathcal{R}$, as the dual of until:

$$\phi \mathcal{R} \psi \equiv \neg(\neg \phi \mathcal{U} \neg \psi)$$

STL can be defined as the subset of PSTL formulas free of parameters. A parameter valuation $v \in V$ transforms a PSTL formula $\phi$ into the STL formula denoted $\phi[v]$ obtained by replacing in $\phi$ every parameter $p$ with its value $v_p \in \mathbb{R}$.

We now recall the semantics of STL from [33].

Definition 2.2 (STL Semantics). The satisfaction of STL formula $\phi$ by signal $w$ at time $t \in T$, denoted by $(w, t) \models \phi$, is defined inductively as follows:

$$(w, t) \models \text{true}$$

$$(w, t) \models x \geq c \quad \text{iff} \quad w_x[t] \geq c$$

$$(w, t) \models x \geq c \quad \text{iff} \quad w_x[t] \leq c$$

$$(w, t) \models \neg \phi \quad \text{iff} \quad (w, t) \not\models \phi$$

$$(w, t) \models \phi_1 \lor \phi_2 \quad \text{iff} \quad (w, t) \models \phi_1 \text{ or } (w, t) \models \phi_2$$

$$(w, t) \models \psi_{[a, b]} \phi \quad \text{iff} \quad \exists t' \in t \circ [a, b], (w, t') \models \phi$$

$$(w, t) \models \phi_1 \mathcal{U} \phi_2 \quad \text{iff} \quad \exists t' \in [t, \infty), (w, t') \models \phi_2 \text{ and } \forall t'' \in (t, t'), (w, t'') \models \phi_1$$

We say that $w$ satisfies $\phi$, written $w \models \phi$, when $(w, 0) \models \phi$.

For simplicity, we use non-strict until semantics, meaning that $\psi$ is allowed to hold at the current point in time, rather than strictly in future. See [24] for a discussion of strict versus non-strict semantics.

The introduction of parameters enables to consider several variants of the same formula, with different constants. This is captured in the notion of validity domain.

Definition 2.3 (Validity Domain). The validity domain of a PSTL formula $\phi$ relative to some signal trace $w$, denoted by $D(\phi, w)$, is the set of parameter valuations for which the formula is satisfied:

$$D(\phi, w) = \{v \in V : w \models \phi[v]\}$$
Our method explicitly computes the validity domain of the formula in a bottom-up style of computation. Starting from the validity domains of atomic formulas, which derive directly from the input signals, the validity domain of non-atomic formulas is assembled by Boolean and temporal combinations of validity domains of its subformulas. In the case of temporal operators, we require the validity domain of the subformulas at future time points.

**Definition 2.4 (Parametric Validity Signal).** Given a PSL formula \( \varphi \), a signal \( w \), the parametric validity signal denoted \( d(\varphi, w) : T \rightarrow \wp(V) \) is defined as follows:

\[
d(\varphi, w)[t] = \{ v \in V : (w, t) \models \varphi[v] \}
\]

The value of the validity signal at time 0 gives the validity domain:

\[
D(\varphi, w) = d(\varphi, w)[0].
\]

**Lemma 2.5 (Inductive Characterization).** For a PSL formula, a signal \( w \), and a time point \( t \in T \) we have

\[
d(x \leq c, w)[t] = \{ w_k[t] \leq c \text{ then } V \text{ else } \\}
\]

\[
d(x \geq c, w)[t] = \{ w_k[t] \geq c \text{ then } V \text{ else } \\}
\]

\[
d(x \leq p, w)[t] = \{ v \in V : w_k[t] \leq v_p \}
\]

\[
d(x \geq p, w)[t] = \{ v \in V : w_k[t] \geq v_p \}
\]

\[
d(\lnot \varphi, w)[t] = V \setminus d(\varphi, w)[t]
\]

\[
d(\varphi \lor \psi, w)[t] = d(\varphi, w)[t] \cup d(\psi, w)[t]
\]

\[
d(\varphi \land \psi, w)[t] = \bigcup_{t' \in \Theta[a,b]} d(\varphi, w)[t']
\]

\[
d(\varphi \mathcal{U} \psi, w)[t] = \bigcup_{t' \in \Theta[a,b]} \left( d(\varphi, w)[t'] \cap \bigcap_{t'' \in \Theta(t', t'')} d(\psi, w)[t'''] \right)
\]

where \( x \in X \), \( c \in \mathbb{R} \), and \( p \in P \).

When the satisfaction of the formula is monotonic in each of its parameters, one may equivalently talk about the set of tightest parameter valuations such that the formula is satisfied by the signal.

We now introduce the notion of polarity of a parameter \( p \in P \) in \( \varphi \), intending to assign positive polarity to \( p \) if \( \varphi \) is easier to satisfy as the value of \( p \) increases, and negative polarity if \( \varphi \) is harder to satisfy as \( p \) increases. In general, formula satisfaction may not be monotonic in \( p \), and \( p \) may not have a defined polarity.

**Definition 2.6.** The polarity set \( \pi(\varphi, p) \subseteq \{-1, 1\} \) of a parameter \( p \in P \) in a formula \( \varphi \) is defined by induction as follows:

\[
\pi(\text{true}, p) = \pi(x \leq c, p) = \pi(x \geq c, p) = \{-1, 1\}
\]

\[
\pi(x \leq p', p) = \{1\} \text{ if } p' = p, \text{ otherwise } \{-1, 1\}
\]

\[
\pi(x \geq p', p) = \{-1\} \text{ if } p' = p, \text{ otherwise } \{-1, 1\}
\]

\[
\pi(\lnot \varphi, p) = \{-i : i \in \pi(\varphi, p)\}
\]

\[
\pi(\varphi \lor \psi, p) = \pi(\varphi, p) \cup \pi(\psi, p)
\]

\[
\pi(\varphi \land \psi, p) = \pi(\varphi, p) \cap \pi(\psi, p)
\]

\[
\pi(\varphi \mathcal{U} \psi, p) = \pi(\varphi, p) \cap \pi(\psi, p)
\]

Intuitively, 1 and -1 denote positive and negative polarity respectively. The value \( \varnothing \) means that different subformulas of \( \varphi \) give different polarity to \( p \), and thus \( p \) does not have consistent polarity within. The value \( \{-1, 1\} \) means that \( \varphi \) does not depend on \( p \).

In this work, we restrict ourselves to formulas whose parameters have consistent polarity, which is a common case. The satisfaction of such formulas is monotonic with respect to parameter values and this allows us to characterize signals in terms of the tightest parameter values with which a given formula is satisfied. Restricting to consistent polarity does not incur a loss of generality, in the following sense. Given a formula \( \varphi \) with an inconsistent parameter \( p \), we can replace its negative occurrences with a fresh parameter \( p' \), and keep the positive occurrences untouched. Then we can intersect the validity domain of this new formula with the plane \( p = p' \) to obtain the validity domain of the original formula.

In addition to assuming consistent polarity, we assume that the formula is given to our algorithm in negation normal form. That is, our input language is described by the following grammar:

\[
\varphi ::= \text{true} \mid \text{false} \mid x < c \mid x > c \mid x < p \mid x > p \mid \varphi \lor \varphi_1 \mid \varphi_1 \land \varphi_2 \mid \varphi_1 \mathcal{U} \varphi_2 \mid \varphi_1 \mathcal{R} \varphi_2
\]

where \( < \in \{<, \leq\} \) is a comparison operator, and the meaning of \( x, c, p, a \), and \( b \) is as before. Every formula \( \varphi \) has an equivalent in negation normal form, produced by rewriting \( \varphi \) using standard arithmetic and logical rules and the temporal equivalences:

\[
\lnot(\varphi \mathcal{U} \psi) \iff \lnot\varphi \mathcal{R} \lnot\psi
\]

Finally, we restrict ourselves to computing the topological closure of the validity domain. One can also see this as computing the boundary of the validity domain disregarding the fact whether the boundary lies inside or outside of it. Syntactically, taking the topological closure of the domain is equivalent to replacing strict inequalities in the formula with their non-strict versions. We assume that our formulas are in negation normal form and only contain non-strict inequalities \( x \leq c, x \geq c, x \leq p, x \geq p \).

### 3 COMPUTING THE VALIDITY SIGNALS

In this section, we study the case of piecewise-constant input signals, which we represent as finite sequences of time intervals mapped to values. We show how to compute the validity signals of formulas by induction on their structure.

**Definition 3.1 (Piecewise-Constant Signal Representation).**

A piecewise-constant signal \( w \) is represented as a sequence of intervals \( [t_{i-1}, t_i] \) mapped to values \( w^j \):

\[
((t_0, t_1) \mapsto w^1, (t_1, t_2) \mapsto w^2, \ldots, (t_{n-1}, t_n) \mapsto w^n)
\]

where \( t_0 = 0, t_n = |w| \) is the duration of \( w \), and \( w^j \) is the value of \( w[t] \) when \( t \in [t_{i-1}, t_i] \). We say that \( n \) is the length of the signal. We call segment a mapping of an interval to a constant value, written \( [t_{i-1}, t_i) \mapsto w^j \).

For a piecewise-constant input signal, the validity signal is also piecewise-constant.

**Theorem 3.2.** For a formula \( \varphi \) and piecewise-constant signal \( w \),

1. \( d(\varphi, w) \) is a piecewise-constant signal;
2. \( d(\varphi, w)[t] \) is finite union of rectangles in \( V \) for all \( t \in T \).

The rest of the section can be seen as a constructive proof of this.

For the Boolean constants and atomic comparisons, we can directly build the piecewise-constant validity signals, where the validity domains are rectangles. For the logical and temporal operators,
we combine the validity signals of the subformulas using a finite number of set unions and intersections. We process the validity signals of the subformulas piece by piece, and concatenate the resulting pieces together. We now introduce additional notation for this purpose. Let $w = (t_0, t_1) \mapsto w^1; \ldots; (t_{n-1}, t_n) \mapsto w^n$ be a signal and let $s = (t, t') \mapsto w'$ be a new segment. Then, the validity domains for atomic constraints are computed in the same duration. We have $s$ may be represented as the sequence $\langle s \rangle = \langle s(n), s(n-1), \ldots, s(1), s(0) \rangle$ to denote the result of appending. When $t_0 > 0$ (i.e., $w$ is not a fully constructed signal) and $t' = t_0$, we may prepend the new segment to $w$. We write $\langle s; t \rangle$ to denote the result of appending. We write $\langle \rangle$ to denote the empty signal, that has no segments and has duration 0. The empty signal can be appended or prepended to any signal $s$, which leaves it unaffected, that is, $\langle \langle \rangle; s \rangle = \langle s; \rangle \rangle = s$.

### 3.1 Overall Algorithm

The algorithm proceeds by induction on the formula structure. For Boolean constant and atomic comparisons, we compute the validity signal directly, by iterating over the segments of the input signal. For the logical and temporal operators, we first compute the full validity signals of the subformulas. Finally, the validity domain of a formula is given by the value of its validity signal at time 0.

Below we explain how to compute the validity signal for every type of formula. For every operator, the computation step implements the corresponding relation of Lemma 2.5. The algorithms in this section view validity domains as plain set of parameter valuations (i.e., subsets of $V$) without any internal structure, and combine them with set union and set intersection. We will later discuss the properties of validity domains of consistent-polarity formulas and how these domains can be efficiently represented.

#### 3.1.1 Boolean Constants

For the Boolean constants, the validity signals are the constant functions:

$$d(\text{true}, w) = \{(0, |w|) \mapsto \top\}$$

$$d(\text{false}, w) = \{(0, |w|) \mapsto \bot\}$$

#### 3.1.2 Atomic Comparisons

Let the component $x$ of the input signal $w$ be represented as the sequence $\langle (t_1, t) \mapsto w^i \rangle _{i=1..n}$. Then, the validity domains for atomic constraints are computed in the following way. For the comparisons with constants $x \leq c$ and $x \geq c$, we have

$$d(x \leq c, w) = \langle (t_1, t) \mapsto \top \rangle$$

$$d(x \geq c, w) = \langle (t_1, t) \mapsto \bot \rangle$$

For the comparison with parameters $x \leq p$ and $x \geq p$, we have

$$d(x \leq p, w) = \langle (t_1, t) \mapsto \{v \in V : v < p \} \rangle$$

$$d(x \geq p, w) = \langle (t_1, t) \mapsto \{v \in V : v > p \} \rangle$$

#### 3.1.3 Disjunction and Conjunction

The validity signals $d(\varphi \lor \psi, w)$ and $d(\varphi \land \psi, w)$ are produced by combining the validity signals $d(\varphi, w)$ and $d(\psi, w)$ point-wise with set union and set intersection respectively. In Fig. 1, we show the function $\text{Combine}$ that produces such a combination for two piecewise-constant validity signals of the same duration. We have

$$d(\varphi \lor \psi, w) = \text{Combine}(d(\varphi, w), d(\psi, w))$$

$$d(\varphi \land \psi, w) = \text{Combine}(d(\varphi, w), d(\psi, w))$$

### Figure 1: Combining two validity signals $d^\varphi$ and $d^\psi$ with a given function $f$.

#### 3.1.4 Bounded Eventually and Always

Given a time point $t$, the value of the validity signal $d(\varphi(a, b), w)$ at time $t$ is the union of values of $d(\varphi, w)$ on the interval $[t + a, \min(t + b, |w|)]$. We call this interval the forward cone of $t$ w.r.t. $[a, b]$. Conversely, we observe that given a time point $t'$, for every $t \in [\min(0, t' - b), t' - a]$ we have $d(\varphi(a, b), w)[t] \subseteq d(\varphi, w)[t']$. The interval $[\min(0, t' - b), t' - a]$ is called a backward cone of $t'$. This definition can be lifted to an interval $[t, t')$ for which the $[a, b]$-backward cone is defined as $[t, t') \cap [a, b] = [\max(0, t - b), t' - a]$. This operation is based on Minkowski difference (over a bounded time domain) and we denote it accordingly. Now, if the piecewise-constant representation of $d(\varphi, w)$ contains a segment $\langle (t_i, t_{i+1}) \mapsto d^\varphi \rangle$ then for every $t \in [t_i, t_{i+1}] \cap [a, b]$ we have $d(\varphi(a, b), w)[t] \subseteq d^\varphi$. This leads to the idea that we can compute $d(\varphi(a, b), w)$ by backshifting: for every segment in the piecewise-constant representation of $d(\varphi, w)$ we produce its backward cone and then combine the backward cones with set union.

Let us consider an example, shown in Fig. 2. At the top of the figure, we show the input validity signal that talks about a single parameter $p$. It maps the interval $[0, 3]$ to the set $\{v : v < p \geq 3\}$, and $[3, 6]$ to $\{v : v < 6\}$, to $\{v : v > 4\}$, to $\{v : v < 4\}$, and $[17, 20]$ to $\{v : v < 20\}$. Below, we show the backward cones of these intervals w.r.t. the time window $[1, 8]$. For this time window, the backward cone of $[0, 3]$ is $[0, 2]$, the backward cone of $[9, 17]$ is $[1, 16]$ and so on. At the bottom of the figure we show the result of combining the backward cones with set union. Now, how no backward cones overlap the interval $[19, 20]$ and thus the result maps this interval to the empty set.

The algorithm $\text{Backshift}$ in Figures 3–5 is designed to combine the backward cones efficiently. The combination is done with set union for the $\text{eventually}$ operator and with set intersection for always. For a number of initial segments of $d(\varphi, w)$, their backward cones start at time 0 (in the example in Fig. 2, these are segments $[0, 3]$, $[3, 6]$, and $[6, 9]$). We can produce the intermediate result of combining just these backward cones by scanning $d(\varphi, w)$ backwards and computing the running union or intersection. This is done by the function $\text{BackshiftInit}$ in Fig. 4. The remaining segments of $d(\varphi, w)$ are combined by scanning them forward and calling the function $\text{BackshiftAdd}$ (in Fig. 5) for every backward cone.
Algorithm BackshiftAdd scans the intermediate result of backshifting res backwards and combines the validity domains of the existing segments \( d_i \) and the validity domain of the new segment \( d' \). An important property is that in the intermediate result of backshifting res, segments that overlap with the incoming backward cone \([τ, τ']\) are always arranged in descending order when combining with set union and in ascending order when combining with intersection. This means that BackshiftAdd can stop scanning res when it encounters a validity domain which is a superset of \( d' \). We return to this point later when discussing complexity, in Section 5.

Finally, we get

\[
\begin{align*}
    d'(\{a,b\}, \varphi, w) &= \text{Backshift}(d(\varphi, w), a, b, \cup, \varnothing) \\
    d(\varnothing, \{a,b\}, \varphi, w) &= \text{Backshift}(d(\varphi, w), a, b, \cap, T)
\end{align*}
\]

3.1.5 Until and Release. The computation of the validity domain of \( τ' U τ \) is better explained in the case when \( d(\varphi, w) \) and \( d(ψ, w) \) are represented using the same sequence of intervals. That is, \( d(\varphi, w) = (\{t_{i-1}, t_i\} \mapsto d_i^φ)_{i=1..n} \) and \( d(ψ, w) = (\{t_{i-1}, t_i\} \mapsto d_i^ψ)_{i=1..n} \) and thus \( d(φ U ψ, w) = (\{t_{i-1}, t_i\} \mapsto d_i^φ U d_i^ψ)_{i=1..n} \). Then, the validity domains \( d_i^φ U d_i^ψ \) can be inductively defined as follows. To satisfy \( φ U ψ \) at time \( t \in \{t_{i-1}, t_i\} \) during the last time interval, we have to satisfy \( ψ \). That is, \( d_{res} = d_i^ψ \). For \( 1 \leq i < n \), to satisfy \( φ U ψ \) at time \( t \in \{t_{i-1}, t_i\} \), we have to satisfy \( ψ \); or we have to satisfy \( φ U ψ \) during the following time interval \( \{t_{i}, t_{i+1}\} \) and also satisfy \( φ \) on \( \{t_{i-1}, t_i\} \). That is, for \( 1 \leq i < n \), we have \( d_{res} = d_i^φ U d_i^ψ \).

The function Until in Fig. 6 generalizes this inductive definition for the case when \( d(\varphi, w) \) and \( d(ψ, w) \) are represented using different sequences of intervals. It is also parameterized with two
function Until(dφ, dψ, f, g, ⊥f)

let dφ = \{ (t_{i-1}, t_i) \mapsto dφ^i \}_{i=1..n}

let dψ = \{ (t_{j-1}, t_j) \mapsto dψ^j \}_{j=1..m}

res ← \{ \}, i ← n, j ← m, d_{out} ← ⊥f

while true do

res ← \{ max(t_{i-1}, t_{j-1}), min(t_i, t_j) \} \mapsto f(dψ^j, d_{out}^i), res

t_end ← max(t_i, t_j)

if i = 1 ∧ j = 1 then break

if t_{j-1} > t_{i-1} then

\text{d}_{out} ← f(d_{out}, g(dφ^i, dψ^j)), j ← j - 1

else if t_{i-1} > t_{j-1} then

\text{d}_{out} ← g(d_{out}, dφ^{i-1}), i ← i - 1

else

\text{d}_{out} ← g(f(dφ^i, d_{out}), dφ^{i-1}), i ← i - 1, j ← j - 1

return res

Figure 6: Computing the validity signal of _until_ and _release_.
Here dφ and dψ are the validity signals of the subformulas; (f, g, ⊥f) are (∪, ∩, ⊥) for _until_ and (∪, ∩, V) for _release_.

For the formulas with consistent polarity, we can actually make all the validity domains upward-closed, if for a negative parameter p, we interpret v_p as the opposite of the value of p. This is equivalent to replacing in a formula every negative parameter p with a positive parameter −p (as done in [6]). This allows to represent validity domains as sets of their support points and also to restate Theorem 3.2 to talk about finite unions of upward-closed rectangles.

3.2.2 Distributing Temporal Operators over Boolean. Before running the computation, by default, we rewrite the input formula by distributing eventually over disjunction and always over conjunction, using the equivalences:

\begin{align*}
\phi \lor \psi & \equiv \phi \lor \psi \lor \phi \psi \\
\phi \land \psi & \equiv \phi \land \psi \land \phi \psi \\
\square \phi & \equiv \square \phi \lor \square \phi \psi \\
\lozenge \phi & \equiv \lozenge \phi \lor \square \phi \psi
\end{align*}

It is often the case that \(\phi\) and \(\psi\) have different sets of parameters, thus \(d(\phi, \psi)\) and \(d(\psi, \phi)\) take values of smaller size and dimension than \(d(\phi \lor \psi, \psi)\) and usually have shorter sequences of incomparable values (we discuss this further in Section 5). In some of our experiments, this rewriting reduced the run time of backshifting up to a factor of 10.

3.2.3 Merging Adjacent Segments. In our implementation of the algorithms (Combine, Backshift, Until), we maintain the invariant that in the representation of a piecewise-constant signal there are no two adjacent intervals that map to the same value. For clarity of presentation though, we do not show this in the pseudo-code of the algorithms. Maintaining this invariant does not come for free. In some cases, we can rely on the properties of underlying operations (e.g., the implementation of set union can detect the cases when the result is equal to one of the operands), but in other cases, we have to perform additional set comparisons explicitly. Maintaining this invariant can greatly reduce the length of validity signals for _until_ and _eventually/always_ over wide temporal windows. In our experiments, applying a temporal operator to a signal with tens of thousands of segments could produce a validity signal with just about a hundred of segments.

4 EXAMPLES

In this section, we illustrate the outcome of the algorithm as applied to the sample signal shown in Fig. 7. Its components, x and y, are square waves that alternate between two stable values: 0 and 2, but for a short period of time (for 1 unit) can take the transient value 1. They can be viewed as coarse quantizations of some periodic signals and the reader can imagine how they can be refined in space and time.

Figure 7: Example of signal.
4.1 Range of Values

Perhaps the simplest use of PSTL is to find the bounds on the value of a signal \( x \). For that, we can use the formula \( \square(x \leq p_1 \land x \geq p_2) \). Our procedure finds the smallest possible value of \( p_1 \) and the largest possible value of \( p_2 \) that renders the formula true, which gives for this signal the validity domain \((p_1 \geq 2 \land p_2 \leq 0)\) at time 0, meaning that the value of \( x \) lies between 0 and 2.

4.2 Enumeration of Values

Consider now the formula \( \Diamond(x \leq p_1 \land x \geq p_2) \). For the example signal, the validity domain of this formula at time 0 is \((p_1 \geq 0 \land p_2 \leq 0) \lor (p_1 \geq 1 \land p_2 \leq 1) \lor (p_1 \geq 2 \land p_2 \leq 2)\). This validity domain is shown in Fig. 8, under the convention of negating the values of negative-polarity parameter \( p_2 \). It actually enumerates all possible values of \( x \). Applying this formula to a sampled analog signal would yield a fast-growing set of incomparable rectangles. This is expensive to compute yet the outcome is not very informative.

![Figure 8: Validity domain at time 0 for the signal in Fig. 7 and the formula \( \Diamond(x \leq p_1 \land x \geq p_2) \).](image)

4.3 Common Threshold

We can use the formula \( \square(x \geq p \lor y \geq p) \) to find the common threshold, s.t. at all times at least one signal component is above it. For our example the validity domain of this formula at time 0 is \((p \leq 1)\), meaning that at all times \( x \) is above 1 or \( y \) is above 1.

4.4 High/Low Values

As a less simple example, we consider an analog signal \( x \) whose value is interpreted as Boolean 1 when above some threshold \( x_{hi} \), as Boolean 0 when below some threshold \( x_{lo} \), and considered transient otherwise, during a rising or a falling edge. Our example signal is a simple instance of this case, where \( x_{hi} = 2 \) and \( x_{lo} = 0 \). We can use PSTL to find the values of \( x_{hi} \) and \( x_{lo} \), e.g., as follows. Let us assume we know the maximum duration of a rising or a falling edge \( t_{edge} \) and the minimum amount of time \( t_{stab} \) that the signal will spend in a well-defined Boolean state after an edge. Then, we can use the formula \( \square[0, t_{edge} + t_{stab}](\square[0, t_{stab}]x \leq p_1) \lor (\square[0, t_{stab}]x \geq p_2) \). For our example signal in Fig. 7, the validity domain of this formula at time 0 is the set \((p_1 \geq 2) \lor (p_1 \geq 0 \land p_2 \leq 2) \lor (p_2 \leq 0)\). In particular, we are interested in the rectangle that has the form \( x_{hi} \), \( p_1 \geq 2 \) \( x_{lo} \), \( x_{hi} \), i.e., the rectangle \( (p_1 \geq 0 \land p_2 \leq 2) \), which allows to conclude that \( x_{hi} = 2 \) and \( x_{lo} = 0 \), as expected. The other two rectangles in the validity domain are not relevant for our question. In practice, depending on the signal, we may have to apply the outer \textit{always} with the time window \([0, |t_{edge} - t_{stab}|] \). If the signal ends in a transient state, the final time points will fail to satisfy the \textit{eventually} subformula, and we will want to exclude them from the computation.

5 PERFORMANCE

In this section we give a preliminary performance evaluation of our algorithm from both theoretical and empirical perspectives, the latter based on our implementation of the identification procedure in OCaml, available at [GitHub](https://gitlab.com/abakhirkin/stlpi).

Our backshifting algorithm can actually be seen as a modification of Lemire’s algorithm [32] for computing minima and maxima over a shifting window, which in particular was first used for robustness computation in [18]. The novel feature of our algorithm is that we work in a partially-ordered parameter space and take the union and intersection of validity domains (represented by Pareto sets of minimal supporting points) instead of applying min and max.

During backshifting, the tail of the intermediate result \( res \) is arranged in descending order (when \( f \) is set union, and in ascending order when \( f \) is set intersection) and plays the role of Lemire’s queue. Thus, when the values of the validity signal (i.e., validity domains) that is being backshifted are totally ordered, backshifting performs \( O(n) \) set operations (union, intersection), where \( n \) is the length of the signal. This is, e.g., the case when the validity domains have one dimension. Until and backshifting with a large upper bound (when \textit{BackshiftInit} does all the work) take \( O(n) \) set operations regardless of the structure of validity domains.

When some validity domains are incomparable, backshifting performs \( O(mn) \) set operations where \( m \) is the maximum number of incomparable elements that fall within the same backshifting window. The value of \( m \) depends on the width of the backshifting window, but also comes from some property of the input signal. In favorable cases, it may be the period of a periodic signal or the width of raising and falling edges; in the worst case this may be the number of distinct values that the signal takes (see Section 4 for an example).

With our current implementation, the worst-case complexity of operations on sets of support points is \( O(l^2) \) where \( l \) is the number of points in a set. For dimensions 1 and 2, this can be improved to \( O(l) \) by storing the points in a sorted array, and there also exist tree structures suitable for higher dimensions. We do not have a good intuition into how \( l \) is connected to the length of the signal, \( n \). It is easy to construct examples, where \( l \) is proportional to the length of the signal, but it is unclear whether it can grow faster. In this work we focus on formulas and signals for which \( l \) is small and does not depend too much on factors such as timing constants in the formula, the length of the signal and the presence of noise.

5.1 Artificially Generated Signals

In the first group of experiments, we evaluate the implementation using artificially generated signals of different length, and we have two kinds of signals. Signals of the first kind, referred as \( w_{sincos} \), have two components: a sine wave \( w_\text{sinc}[t] = \sin(\frac{2\pi t}{500}) \) and a cosine wave \( w_\text{cos}[t] = \cos(\frac{2\pi t}{500}) \). Both are sampled with a step of one unit time and thus have a period of 500 samples. Signals of the second...
kind, referred as \( w_{\text{square}} \), are randomly generated square waves with noise and non-zero duration of the rising and falling edges. The average period is 1000 units, and again we take one sample per time unit. A fragment of this signal is shown in Fig. 9. We generate three versions of each signal with varying length: 10 000, 100 000, and 1 million samples.

![Figure 9: Fragment of a generated square wave with noise that we use in the evaluation.](image)

Using these signals, we run parameter identification for a number of formulas and report the results (run time of the identification procedure in seconds and, space permitting, the validity domain at time 0 for the signal with 1 million samples) in Table 1. Time figures were obtained on a PC with a Core i7-3630QM CPU and 8GB RAM.

In this evaluation we are interested in signals and formulas for which the validity domains have a small number of rectangles that does not depend too much on the temporal windows in the formula, the length of the signal, presence of noise and so on; in particular, we wish to avoid validity domains that enumerate distinct signal values. The only formula in our evaluation which is ill-behaved in this sense is \( \phi_6 \); its validity domain depends on the presence of noise and also on the length of the signal.

The goal of our implementation was to efficiently perform parameter identification for formulas with small validity domains, and we think that we achieved this goal. For most example formulas, it takes a few seconds to run parameter identification for a signal with 1 million samples, and the run time grows linearly with the length of the signal or close to that.

We evaluate multiple versions of the formulas \( \phi_5 \) and \( \phi_6 \), with different temporal windows. Formulas \( \phi_{5,1}, \phi_{5,2} \) and \( \phi_{6,1}, \phi_{6,2} \) demonstrate the effect of changing the window of an outermost temporal operator in a formula with a small validity domain. We can observe that for shorter signals it is more efficient to backshift with a higher upper bound, since in the implementation BackshiftInit is more efficient than repeated application of BackshiftAdd. Formula \( \phi_{6,3} \) demonstrates how changing the window of a nested temporal operator can change the formula from being well-behaved to ill-behaved for a given signal. There appears to be an interaction between the upper bounds of eventually and the signal period (250 corresponds to the half-period, 125 is the quarter-period).

### 5.2 Airplane Pitch Control Model

We also evaluate the implementation using the data produced by the Simulink model of an airplane longitudinal flight control system. The aim of this system is to control the pitch orientation of the airplane. We assume the control is biased with some fixed offset, and the expected property of the system is that within a delay of 15 units the response signal of the system stabilizes close to the command signal, and the stable state is kept for at least 30 units. The model is driven with a command signal that is piecewise-constant with constant periods lasting 200 units (long enough for the response to stabilize), and varying command values in the range \([-1.0, 1.0]\). In Fig. 10 we show a fragment of the difference signal between command and response. In Table 1, we refer to this signal as \( w_{\text{pitch}} \). In the evaluation, we measure the bounds on the difference between command and response in a stable state. We use the fact that response always stabilizes above command and employ the following formula: \( \square [0.45, 0.66] (w_{\text{cmd}} - w_{\text{resp}} \geq p_1 \land x_{\text{resp}} - x_{\text{cmd}} \leq p_2) \). The validity domain of this formula at time 0 is \((p_1 \leq -0.066 \land p_2 \geq 0.066)\), which means that response stabilizes within the range of 0.066 from command signal.

### 6 RELATED WORK

The problem of parameter estimation/synthesis in system models [15] as well as in properties such as those expressed in PSTL, is a crucial problem in the design and analysis of systems and is implemented in tools such as S-Taliro [5] and Breach [16]. The industrial-size case studies of [27] and [36] demonstrate the practical relevance of PSTL, and its value in conjunction with other methods.

In general, the problem that we solve in this paper can be seen as a case of learning from positive examples since we observe only behaviors which are possible. In other contexts, such as those of [13] and [11, 30], the traces can be classified as normal or abnormal and the problem becomes that of learning from both positive and negative examples, a case of supervised learning.

The works of [26], of [25, 42], and of [29] are most related to ours. The problem they study can be stated as follows: given a system model \( M \) and a parameterized temporal formula \( \varphi[p] \), find the validity domain of \( \varphi \) relative to \( M \), that is, the set of parameter valuations \( v \) such that \( w \models [v] \). Here the system model is viewed as a black box, that can produce from some input \( u \in U \) a simulation trace \( w = M(u) \). The input space \( U \) is typically large or uncountable, and such methods are approximate. They explore the boundary of \( D(\varphi, M) = \bigcap_{w \in M(U)} D(\varphi, w) \) by search in the combined space \( U \times V \) of system inputs and formula parameters.
Our work builds up on that of [6] to solve the problem, for a single simulation or execution trace of the system. In fact [6] considered two techniques, one exact and based on translation to a quantified formula in linear arithmetic, and one approximate, based on search in the parameter space. In this work, we solve the problem in an exact manner using signal-processing computations in the style of [18, 33], avoiding reliance on a (costly) quantifier elimination routine. The exhaustive computation of the validity domain is similar to that of [12], who study the problem of monitoring STL*, a variant of STL with freeze quantification [4]. There the parameter space is over subsets of $\mathbb{R}^k$ for $k$ frozen variables, but could be rephrased over $\mathbb{R}^k$, something we intend to explore further.

With or without parameters, the main application of STL monitoring is found in falsification. This problem, dual of verification, attempts not to prove that the system $M$ is correct under all inputs $u \in U$, but simply to find a faulty execution $w = M(u)$, without any formal guarantees that it will be found. The most effective technique, as illustrated in [2, 34], turns the falsification problem into the following optimization problem: “minimize $\rho(p, w)$ subject to $w = M(u), u \in U$.” The robustness value $\rho$ is expected to be continuous in $u$, and by definition is $w \not= \phi$ when $\rho(p, w) < 0$. This technique was indeed used in [25, 27, 29, 42] precisely for the purpose of temporal logic parameter exploration.

The choice of minimization algorithm in the above is crucial, and several alternatives have been explored, see [1, 39] for instance. It can be argued with [3, 26] that the choice of the cost function is equally important. An undesirable behavior of the robustness value as cost function, is due to its absolute-norm semantics, which selects the value of the safest signal variable (the furthest away from violation) as the one to optimize. We believe that parametrizations of STL formulas provide a way to define better behaved (smoother) cost functions. For this, observe that the robustness $\rho(p, w)$ can be recovered as the tightest parameter assignments of PSTL formula $\phi[p]$ where $p$ is a unique parameter replacing all constants, and $\phi'$ is the positive normal form of $\phi$. Using several parameters (for several signal variables) could provide a cost function with non-zero derivative in more than one variable.

### 7 CONCLUSIONS AND FUTURE WORK

In this work we presented a novel algorithm for parametric identification for STL and applied it to formulas with space parameters and piecewise-constant (PC) signals where the validity domains are unions of rectangles. We have shown that in many cases, a prototype implementation of our algorithm can compute the validity domain for signals with millions of samples. Hence our method provides a viable alternative to the previously used algorithms based on quantifier elimination or search in the parameter-space, and will allow the derivation of compact representations of systems based on observable behaviors.

The immediate direction for future work is improving the implementation, in particular the representation of validity domains. Now, we store validity domains as arrays of support points, and the complexity of disjunction and conjunction is quadratic in the size of the support set. More efficient representations of such sets are known in the context of maintaining a Pareto front, but adapting the data structures and algorithms to our setting is still a challenge.

There are two immediate extensions of our work. If we use a piecewise-linear (PL) interpolation for signals, we will not have rectangular validity domains in each intervals but polytopes that depend on $t$. The computational trade-offs between using PL and PC signals should be investigated. On one hand, rectangles are easier to manipulate but on the other, PC signals will require a denser sampling than PL signals to achieve the same approximation level with respect to the underlying continuous signal. In general the influence of sampling rates of the same signal on the obtained validity domain should be studied.

The more interesting and urgent extension is to include timing parameters where, naturally, validity domains will explicitly depend on time, which changes continuously and does not immediately fit in the piecewise-constant setting. We are confident, though, that this can be done and that the validity domains associated with timing parameters and PC signals are more restricted types of polyhedral sets than the semi-linear sets associated with PL signals. It should be noted that mixing space and time parameters will often
lead to many incomparable points. For example, the validity domain for the PSTL formula $0(p_1) \leq p_2$ relative to a decaying signal consists of a continuum of incomparable $(p_1, p_2)$ points. Another question is whether our approach can be translated to other timed formalisms such as signal regular expressions [7].

Parametric validity domains can be viewed as a multi-dimensional generalization of robustness which we suggest will provide finer information concerning the robustness associated with subformulas and with different parts of the signal. Some work is needed to gain better insights on the precise relation between the two and its implications for robustness-guided falsification algorithms.

Finally, in addition to providing a succinct representation of observed behaviors, we can view PSTL formulas as a new type of feature extractors, functions that map high-dimensional objects such as signals into low-dimensional objects, in our case sets of tightest values in the validity domain. Once mapped into this space, the signals can be subject to various learning and clustering algorithms as suggested recently in [41].

REFERENCES


[27] Xiaojian Jing, Alexandre Donzé, Jöystormy V. Deshmukh, and Sanjit A. Seshia. 2016. Mining Requirements from Closed-loop Control Models. In FATES.