# Disturbance Bounds for Signal Temporal Logic Task Satisfaction: A Dynamics Perspective

Prithvi Akella, *Graduate Student Member, IEEE*, and Aaron D. Ames, *Fellow, IEEE*

***Abstract*—This letter offers a novel approach to Test and Evaluation of pre-existing controllers from a control barrier function and dynamics perspective. More aptly, prior Test and Evaluation techniques tend to require *apriori* knowledge of a space of allowable disturbances. When these disturbances enter the dynamics linearly, however, our work determines a two-norm disturbance-bound rejectable by a system's controller *without requiring specific knowledge of these disturbances beforehand*. The authors posit that determination of such a disturbance bound offers a better understanding of the robustness with which a given controller achieves a specified task - as motivated through a simple, linear-system example. Additionally, we show that our resulting disturbance bound is accurate through simulation of 1000 randomized trials in which a Segway-controller pair successfully satisfies its specification despite randomized perturbations within our identified bound.**

***Index Terms*—Control barrier functions, signal temporal logic, uncertain systems.**

## I. Introduction

While there exist multiple temporal logic formalisms, two of increasing interest in the controls community are Linear Temporal Logic and Signal Temporal Logic [1]–[3]. This interest arises as these logical schemes offer succinct ways of expressing complex, desired behavior, while also providing necessary and sufficient criteria by which to determine if a system has achieved this behavior [3]–[6]. As a result, there has been significant work utilizing these logical formalisms to enforce satisfaction of these behavioral specifications [7]–[10]. Additionally, these formalisms and satisfaction criteria have also prompted the development of evaluation schemes to test a controllers ability to realize these desired system behaviors when experiencing environmental disturbances [11]–[15]. Finally, the authors note that there has also been significant work aimed at developing controllers that robustly reject these environmental disturbances, most recently with active disturbance rejection control [16]–[18].

However, this leads to a question we aim to explore in this letter. As mentioned prior, existing work in the Test and Evaluation community endeavors to test and evaluate a controller's ability to realize desired system specifications while subject to environmental disturbances. These procedures typically amount to an optimization problem over the feasible space of these disturbances, requiring identification of the allowable disturbances beforehand [19]. As such, the authors posit that it might be more fruitful were we to identify the level of disturbance that a given controller can reject as opposed to determining the worst-case disturbance from a given set. More accurately, can we use a system model and model-theoretic control techniques to identify a two-norm disturbance-bound that our controller can reject whilst still satisfying its incumbent specification?

*Our Contribution:* Our contribution is twofold. First, we construct two optimization problems that each generate two-norm disturbance-bounds rejectable by a system's controller while it steers its system to satisfy its specification. Each optimization problem focuses on a specific subset of Signal Temporal Logic, and we use their solutions to construct our system-level bound. Secondly, we show that our generated bound is accurate albeit conservative, as it depends on Lipschitz constants for the system dynamics and specification. Over 1000 simulated Segway runs where disturbances are sampled randomly from within our prescribed norm-bound, we show that the Segway-controller pair rejects disturbances within our identified bound and achieves its Signal Temporal Logic task. For context, the subset of STL tasks studied in this letter is consistent with prior works in the controls literature [9], [10], and we center our analysis on disturbances that enter the dynamics linearly.

*Organization:* Section II details some background material in Section II-A, motivates our problem in Section II-B, and formally states our problem in Section II-C. Then, Section III details our main contributions - the optimization problems determining two-norm disturbance-bounds rejectable by a system's controller. Finally, Section IV illustrates our results through a simulated Segway example.

## II. Problem Formulation

This section will detail some necessary background material - specifically Signal Temporal Logic and Control Barrier Functions. We will start with some notation.

*Notation:* $\|\cdot\|$ is the 2-norm over $\mathbb{R}^n$. $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$, $\mathbb{R}_{++} = \{x \in \mathbb{R} | x > 0\}$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz

Manuscript received September 14, 2021; revised November 18, 2021; accepted December 7, 2021. Date of publication December 21, 2021; date of current version December 30, 2021. This work was supported by the Air Force Office of Scientific Research Test and Evaluation Program under Grant FA9550-19-1-0302. Recommended by Senior Editor S. Tarbouriech. *(Corresponding author: Prithvi Akella.)*

The authors are with the Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: pakella@caltech.edu; ames@caltech.edu).

Digital Object Identifier 10.1109/LCSYS.2021.3137267

continuous if and only if $\exists\, L \in \mathbb{R}_+$ such that $|f(x) - f(z)| \leq L\|x - z\|$. A continuous function $\alpha \in \mathcal{K}_{e,\infty}$ if and only if $\alpha : (-\infty, \infty) \to \mathbb{R}$, $\alpha(0) = 0$, $r > s$ implies $\alpha(r) > \alpha(s)$, and $\lim_{r\to\infty} \alpha(r) = \infty$. For any continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}$ is a regular value if and only if $D_x h(x) \neq 0 \;\forall\; x$ s.t. $h(x) = a$. The space of all signals $\mathcal{S}^{\mathbb{R}^n} = \{s | s : [0, T] \to \mathbb{R}^n, \;\forall\; T > 0\}$ with $s$ a signal. $\|\cdot\|_{[a,b]}$ is an induced (semi)-norm over $\mathcal{S}^{\mathbb{R}^n}$ where $\|s\|_{[a,b]} = \max_{t\in[a,b]} \|s(t)\|$ for $s \in \mathcal{S}^{\mathbb{R}^n}$.

### A. Preliminaries

In this section, we will provide a brief description of Signal Temporal Logic and Control Barrier functions. Then, we will motivate the specific problem under study with an example.

*Signal Temporal Logic:* Signal Temporal Logic (STL) is a language by which rich, time-varying system behavior can be succinctly expressed. This language is based on predicates $\mu \in \mathcal{A}$ which are boolean-valued variables taking a truth value for each state $x \in \mathcal{X}$. Predicates $\mu$ and specifications $\psi$ are defined as follows, with "|" demarcating definitions:

$$\mu(x) = \text{True} \iff h_\mu(x) \geq 0, \;\; h_\mu : \mathcal{X} \to \mathbb{R},$$
$$\psi \triangleq \phi | \neg\psi | \psi_1 \vee \psi_2 | \psi_1 \wedge \psi_2 | \psi_1 \,\mathrm{U}_{[a,b]}\, \psi_2, \;\; \psi \in \mathbb{S}. \quad (1)$$

Here, $\psi_1, \psi_2$ are specifications themselves and $\mathbb{S}$ is the set of all STL specifications. We write $(s, t') \models \psi$ when a signal $s$ satisfies a specification $\psi$ for times $t \geq t'$, e.g., $(s, t) \models \psi_1 \,\mathrm{U}_{[a,b]}\, \psi_2$ implies that $\exists\, t^* \in [t+a, t+b]$ such that $(s, t') \models \psi_1 \;\forall\; t \leq t' \leq t^*$ and $(s, t^*) \models \psi_2$. To be brief, we will refrain from formally defining the satisfaction relation $\models$ for, as we will instead note that every STL specification $\psi$ has a robustness measure $\rho$ that is positive for signals $s$ that satisfy $\psi$ [2], [3], [6], [20].

*Definition 1:* A function $\rho : \mathcal{S}^{\mathbb{R}^n} \times \mathbb{R}_+ \to \mathbb{R}$ is a *robustness measure* for an STL specification $\psi$ if it satisfies the following equivalency: $\rho(s, t) \geq 0 \iff (s, t) \models \psi$.

Here, we note that while defining a robustness measure as per Definition 1 aligns with prior controls works [9], [10], [21] and our predicate definition as per equation (1), it is not the only way of defining such a measure, e.g., see [3, Definition 3] or [20, Sec. 2.3]. Finally, to simplify notation, two commonly used temporal logic operators will be produced here. The first is $\mathbf{F}_{[a,b]}\,\psi$ which reads as $\psi$ should be true *at some point in the future* for some time $t \in [a, b]$. The second is $\mathbf{G}_{[a,b]}\,\psi$ which reads as $\psi$ should be true *for all times* $t \in [a, b]$. In both cases, $b > a$.

$$\mathbf{F}_{[a,b]}\,\psi = \text{True}\,\mathrm{U}_{[a,b]}\,\psi, \quad \mathbf{G}_{[a,b]}\,\psi = \neg\big(\text{True}\,\mathrm{U}_{[a,b]}\,\neg\psi\big).$$

*Control Barrier Functions:* Originally inspired by their counterparts in optimization (see [22, Ch. 3]), control barrier functions are a modern control tool used to ensure safety in safety-critical systems that are control-affine, i.e.,

$$\dot{x} = f(x) + g(x)u, \quad x \in \mathcal{X} \subseteq, \mathbb{R}^n, \;\; u \in \mathcal{U} \subseteq \mathbb{R}^m. \quad (2)$$

We will assume we have a feedback controller $k(x)$ for (2), which results in the following closed-loop dynamics:

$$\dot{x} = f_{cl}(x) \triangleq f(x) + g(x)k(x), \;\; x \in \mathcal{X}. \quad (3)$$

Now, solutions to (3) may not exist for all time [23]. As such, we denote this interval of existence of solutions to (3) emanating from $x_0$ as $I(x_0) = [0, t_{\max}]$. We denote the corresponding solution as $\phi_t(x_0)$, where

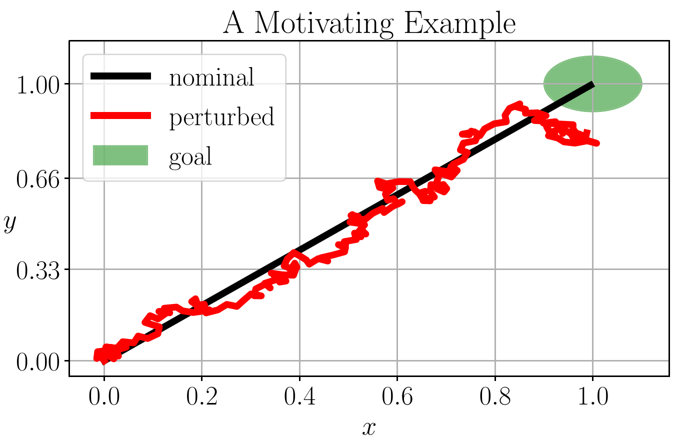

Fig. 1. The motivating example detailed in Section II-B for this letter's problem. For the closed-loop system shown, the undisturbed trajectory (black) satisfies its specification - reach the goal (green) within 2 seconds. However, disturbing the same system results in a trajectory (red) that fails to satisfy this specification. This phenomenon prompted the authors to ask the question - *can we determine the two-norm disturbance-bound that a controller can reject while steering a system to satisfy its specification?*

$$\dot{\phi}_t(x_0) = f_{cl}(\phi_t(x_0)), \quad \phi_0(x_0) = x_0. \quad (4)$$

Then, forward invariance is defined as follows.

*Definition 2:* The set $\mathcal{C} \subset \mathbb{R}^n$ is forward invariant with respect to the dynamical system (3) if $\forall\; x_0 \in \mathcal{C}$, $\phi_t(x_0) \in \mathcal{C} \;\forall\; t \in I(x_0)$, with $\phi_t(x_0)$ as per (4).

Control barrier functions then, are a tool used to ensure forward invariance of their 0-superlevel sets. Specifically, for a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$, define its 0-superlevel set $\mathcal{C}$ and boundary $\partial\mathcal{C}$ as follows:

$$\mathcal{C} = \{x \in \mathcal{X} | h(x) \geq 0\}, \;\; \partial\mathcal{C} = \{x \in \mathcal{X} | h(x) = 0\}. \quad (5)$$

Then, the definition of control barrier functions is as follows.

*Definition 3 (Adapted from [24, Definition 5]):* For the control-affine system (2), a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ with 0 a regular value is a *control barrier function* if $\exists\; \alpha \in \mathcal{K}_{e,\infty}$ such that $\forall\; x \in \mathcal{X}$,

$$\sup_{u\in\mathcal{U}} \left[ \dot{h}(x, u) \triangleq \frac{\partial h}{\partial x}(f(x) + g(x)u) \right] \geq -\alpha(h(x)).$$

This ends our brief overview of necessary topics. The next section motivates the specific problem under study.

### B. A Motivating Example

To better motivate our problem statement, we will provide a brief example. Consider the following single integrator system subject to an STL specification $\psi$ with associated robustness measure $\rho$ and with $g = [0.75, 0.75]^T$:

$$\dot{x} = u, \;\; x \in [-1, 1]^2, \;\; u \in [-0.5, 0.5]^2, \quad (6)$$
$$\mu_g(x) = \text{True} \iff \Big(h_\mu(x) \triangleq 0.1 - \|x - g\|_2\Big) \geq 0,$$
$$\psi = \mathbf{F}_{[0,2]}\,\mu_g, \;\; \rho(s, 0) \triangleq \max_{t\in[0,2]} \;\; h_\mu(s(t)). \quad (7)$$

It is fairly simple to construct a controller $U$ that ensures that $(\phi(\mathbf{0}), 0) \models \psi$, where $\phi(\mathbf{0}) \in \mathcal{S}^{\mathbb{R}^n}$ is the closed-loop solution of (6) and this controller $U$ starting from $x_0 = \mathbf{0}$. Figure 1 shows an example controller and resulting trajectory $\phi(\mathbf{0})$. Indeed, this controller also ensures that $\rho(\phi(\mathbf{0}), 0) = 0.09$, indicating that this controller robustly steers the system to satisfy $\psi$. However, if we introduce some disturbance to the system, as shown via the red trajectory in the same figure, the system fails to satisfy $\psi$. As a result, the controller is not as robust as once claimed. It is for this reason that we aim to develop techniques to discern the level of robustness - in a two-norm sense - that a controller can reject while still ensuring STL specification satisfaction. Such techniques would provide a better understanding of the efficacy of a controller in robustly realizing a required task. With this motivation in mind, we will formalize our problem statement.

### C. Problem Statement

We will start by mentioning two, separate systems - our nominal controlled system and its perturbed version.

$$\dot{x} = f(x) + g(x)U(x) \triangleq f_{cl}(x), x \in \mathcal{X}, \ \ U : \mathcal{X} \to \mathcal{U}, \qquad \text{(CL)}$$
$$\dot{x} = f_{cl}(x) + d, \ \ d \in \mathbb{R}^n. \qquad \text{(CL-d)}$$

For both closed-loop systems (CL) and (CL-d), we will assume $f, g, U$ are locally Lipschitz continuous. This implies $\forall\ x_0 \in \mathcal{X}$ that solutions $\phi(x_0)$ to (CL) and $\phi^d(x_0)$ to (CL-d) have nonzero intervals of existence $I(x_0)$ and $I^d(x_0)$ respectively [23]. Furthermore, we will denote $\phi(x_0) \in \mathcal{S}^{\mathbb{R}^n}$ to be the state trajectory signal and $\phi_t(x_0) \in \mathcal{X}$ to be the state at time $t$ as per equation (4).

We will also assume that this system is subject to an STL specification that is of the following form:

$$\begin{aligned} \omega &= \text{True}\,|\mu|\neg\mu|\omega_1 \wedge \omega_2, \\ \psi &= \mathbf{G}_{[a,b]}\,\omega|\,\mathbf{F}_{[a,b]}\,\omega|\omega_1\,\mathrm{U}_{[a,b]}\,\omega_2|\psi_1 \wedge \psi_2. \end{aligned} \qquad (8)$$

Additionally, we will make the following two assumptions about the predicate functions $h_\mu$ and the robustness measures $\rho$ used in our forthcoming analysis.

*Assumption 1:* The predicate functions $h_\mu$ as per equation (1) for each predicate $\mu$ are continuously differentiable.

*Assumption 2:* The robustness measure $\rho$ for our specification $\psi$ is partially Lipschitz continuous, i.e.,

$$\exists\ L, b \geq 0 \ \text{ s.t. } \ |\rho(s, 0) - \rho(z, 0)| \leq L\|s - z\|_{[0,b]},$$

where $\|\cdot\|_{[0,b]}$ is an induced (semi)-norm over $\mathcal{S}^{\mathbb{R}^n}$.

Here, we note that our restriction to this specific subclass of STL specifications aligns with prior work coupling Signal Temporal Logic and control barrier functions (see the examples in [9], [10]). We will also make one fairness assumption - that the intervals of existence for solutions to either system (CL) or (CL-d) are sufficiently large enough to permit analysis as to whether they satisfy their STL specification. We will also state one definition to formalize what we mean when we say a system satisfies a specification. Specifically, a system satisfies a specification over a given space $\mathcal{X}$ if all trajectories emanating from states $x \in \mathcal{X}$ satisfy $\psi$. A formal definition will follow.

*Definition 4:* We say (CL) satisfies a specification $\psi$ over the space $X$, i.e., (CL) $\models_X \psi$ iff, $\forall\ x \in X$, $(\phi(x), 0) \models \psi$.

Then our problem statement is as follows.

*Problem 1:* Let $\psi$ be a Signal Temporal Logic specification of the form in (8). Determine a space $X \subseteq \mathcal{X}$ and a disturbance bound $\delta_d$ such that (CL-d)$\models_X \psi\ \forall\ d$ s.t. $\|d\| \leq \delta_d$.

As part of our solution, we will require a notion of which predicates $\mu$ are "components" of $\omega$ as per equation (8). We will define this component set $P(\omega)$ as follows:

$$\begin{aligned} \mu \in \omega &\iff (\omega(x) = \text{True} \implies \mu(x) = \text{True}), \\ P(\omega) &= \{\mu \in \mathcal{A}|\mu \in \omega\}. \end{aligned} \qquad (9)$$

This results in the following Lemma.

*Lemma 1:* $\omega(x) \equiv (\wedge_{\mu \in P(\omega)}\ \mu(x))$.

*Proof:* Follows by definition of $\omega$ (8) and $P(\omega)$ (9). ■

We will now proceed to state our main contributions.

## III. Main Contribution

This section will be a series of optimization problems designed to identify spaces $X$ and norm bounds $\delta_d$ such that (CL-d)$\models_X \psi$ for any STL specification $\psi$ as per equation (8). Before formally stating our solution, however, we will provide a brief overview of our approach.

*Overarching Idea:* We will first connect Control Barrier Functions and the system's STL specification $\psi$ through the predicate functions $h_\mu$. More accurately, for a specific subset of specifications, we will construct an optimization problem identifying the maximum two-norm disturbance bound for which the system's controller $U$ still renders each $h_\mu$ a valid Control Barrier Function. Then for another subset of specifications, we will exploit Lipschitz continuity of the system dynamics and partial Lipschitz continuity of the robustness measure $\rho$ to generate a secondary disturbance bound. The minimum of these two bounds will be our final, albeit conservative, disturbance bound.

As such, we will start first with an optimization problem for specifications $\psi = \mathbf{G}_{[0,b]}\,\omega$, where we can connect satisfaction of this STL specification to a CBF-like condition. To facilitate its presentation, we will make the following definitions, with $\mathcal{C}_{h_\mu}$ as per equation (5):

$$\begin{aligned} &\psi = \mathbf{G}_{[0,b]}\,\omega, \quad \mathcal{C}_\omega = \mathcal{X} \cap_{\mu \in P(\omega)}\ \mathcal{C}_{h_\mu}. \\ &\xi(x, e, \mu) = \frac{\partial h_\mu}{\partial x}^T(x)f_{cl}(x) - \left\|\frac{\partial h_\mu}{\partial x}(x)\right\|e, \\ &\Delta(x, \mu, \alpha_\mu) = \left\{e \in \mathbb{R}|\xi(x, e, \mu) \geq -\alpha_\mu\big(h_\mu(x)\big)\right\}. \end{aligned} \qquad (10)$$

Then our proposed optimization problem determines an $\omega$-specific bound $\delta_d^0$ over $\mathcal{C}_\omega$ such that (CL-d)$\models_{\mathcal{C}_\omega} \psi$ i.e.,

$$\begin{aligned} \delta_d^0 = &\min_{x \in \mathcal{C}_\omega}\ \max_{e \in \mathbb{R}}\ e, \\ &\text{subject to } \ e \in \Delta(x, \mu, \alpha_\mu), \ \ \forall\ \mu \in P(\omega). \end{aligned} \qquad (11)$$

Here, we note that the inner maximization over $e \in \mathbb{R}$ is redundant when $|P(\omega)| = 1$. When $|P(\omega)| > 1$, however, the inner maximization allows us to cleanly identify the maximum possible disturbance bound $e$ such that all CBF conditions $\forall\ \mu \in P(\omega)$ are satisfied. Our theorem then states that if $\delta_d^0 \geq 0$, then (CL-d)$\models_{\mathcal{C}_\omega} \psi$ for all disturbances whose two norm is less than this maximum bound $\delta_d^0$. The formal statement of this theorem will follow.

*Theorem 1:* For equation (11), let each $\alpha_\mu \in \mathcal{K}_{e,\infty}$, let the specification $\psi$ and set $\mathcal{C}_\omega$ satisfy equation (10), and let each predicate function $h_\mu$ satisfy Assumption 1. Then,

$$\delta_d^0 \geq 0 \implies \text{(CL-d)} \models_{\mathcal{C}_\omega} \psi \ \forall\ d \ \text{ s.t. } \|d\| \leq \delta_d^0.$$

*Proof:* To start, for any $d$, Cauchy-Schwarz provides that

$$\frac{\partial h_\mu}{\partial x}^T(x)(f_{cl}(x) + d) \geq \frac{\partial h_\mu}{\partial x}^T(x)f_{cl}(x) - \left\|\frac{\partial h_\mu}{\partial x}(x)\right\|\|d\|.$$

Then for any $d$ such that $\|d\| \leq \delta_d^0$ we have that the derivative of $h_\mu$ with respect to the perturbed dynamics (CL-d) satisfies the following inequality as $\delta_d^0 \geq 0$:

$$\dot{h}_\mu(x, d) \geq -\alpha_\mu(h_\mu(x)), \ \forall\ \mu \in P(\omega), \ x \in \mathcal{C}_\omega.$$

Via Peano's Uniqueness Theorem [25, Th. 1.3.1] we know that $\dot{u} = -\alpha_\mu(u)$ has a unique solution $\forall\ u_0 \geq 0$ as $-\alpha_\mu$ is a continuous, non-increasing function in $u$. Using this uniqueness result in conjunction with a Comparison Lemma, [26, Lemma 3.4], allows us to state that

$$h_\mu\left(\phi_t^d(x_0)\right) \geq 0, \ \forall\ \mu \in P(\omega), \ x_0 \in \mathcal{C}_\omega, \ t \in I^d(x_0). \quad (12)$$

Here, we note that this chain of logic was also utilized in the proof for [24, Th. 1] as the proof for [26, Lemma 3.4] requires Lipschitz continuity of $\alpha_\mu$ to guarantee a unique solution (see [26, Appendix C.2]), and this is already provided for via Peano's Uniqueness Theorem. As a result, equation (12) implies that

$$h_\mu\left(\phi_t^d(x_0)\right) \geq 0, \ \forall\ \mu \in P(\omega), \ x_0 \in \mathcal{C}_\omega, \ t \in I^d(x_0).$$

By definition of $h_\mu$, our fairness assumption that $b \in I^d(x_0)$, Lemma 1 and equation (10), we have the following:

$$\text{(CL-d)} \models_{\mathcal{C}_\omega} \psi \ \forall\ d \ \text{ s.t. } \|d\| \leq \delta_d^0.$$

■

For the second set of optimization problems, we will require the Gronwall-Bellman Inequality.

*Theorem 2 (From [27, Th. 1.3.1]):* Let $u, f : J = [\alpha, \beta] \to \mathbb{R}_+$ be continuous over their domain, and let $n : J \to \mathbb{R}_+$ be continuous and non-decreasing. Then, $\forall\ t \in J$

$$u(t) \leq n(t) + \int_\alpha^t f(x)u(s)ds \implies u(t) \leq n(t)e^{\left(\int_\alpha^t f(s)ds\right)}.$$

This theorem allows us to establish the following lemma bounding the difference between solutions to dynamical systems (CL) and (CL-d).

*Lemma 2:* For both systems (CL) and (CL-d), let $f_{cl}$ be locally Lipschitz continuous with constant $L$ for some $x_0 \in \mathcal{X}$. Then, if $\forall\ d, \ \|d\| \leq \delta_d$,

$$\left\|\phi_t(x_0) - \phi_t^d(x_0)\right\| \leq \delta_d t e^{Lt}, \ \forall\ t \in I(x_0) \cap I^d(x_0).$$

*Proof:* This proof amounts to one application of Gronwall-Bellman's Inequality in Theorem 2. For the sake of brevity, we will refrain from proving this Lemma in full. However, the full proof can be found in [28]. ■

Our optimization problem for the remainder of the base specification types $\mathbf{G}_{[a,b]}\,\omega, \mathbf{F}_{[a,b]}\,\omega, \omega_1 \,\mathrm{U}_{[a,b]}\,\omega_2$ will make use of Lemma 2 and Assumption 2 to generate disturbance-bounds $\delta_d^1$ for the entire state space $\mathcal{X}$. More aptly, our setting is as follows, with "|" demarcating different specifications:

$$\psi = \mathbf{G}_{[a,b]}\,\omega | \mathbf{F}_{[a,b]}\,\omega | \omega_1 \,\mathrm{U}_{[a,b]}\,\omega_2, \quad (13)$$

$$\rho(s, 0) \geq 0 \iff (s, 0) \models \psi,$$

$$\Delta_d = \min_{x \in \mathcal{X}} \ \rho(\phi(x), 0). \quad (14)$$

Then our theorem identifying a disturbance-bound $\delta_d$ for specifications $\psi$ of the type in equation (13) is as follows.

*Theorem 3:* Let the closed-loop dynamics $f_{cl}$ be locally Lipschitz continuous with constant $L_f \ \forall\ x_0 \in \mathcal{X}$, let the specification $\psi$ be as per equation (13), and let the robustness measure $\rho$ also satisfy Assumption 2 with Lipschitz constant $L_\rho$ and time constant $b$. If $\Delta_d$ as in (14) is such that $\Delta_d \geq 0$,

$$\text{(CL-d)} \models_{\mathcal{X}} \psi \ \forall\ d \ \text{ s.t. } \|d\| \leq \frac{\Delta_d}{L_\rho b e^{L_f b}} \triangleq \delta_d^1. \quad (15)$$

*Proof:* For this proof, we will assume that our disturbances $d$ are such that $\|d\| \leq M$, and show $M = \delta_d^1$. Then, by local Lipschitz continuity of $f_{cl}$ and Lemma 2 $\forall\ x_0 \in \mathcal{X}$,

$$\left\|\phi_t(x_0) - \phi_t^d(x_0)\right\| \leq Mte^{L_f t}, \ \forall\ t \in I(x_0) \cap I^d(x_0).$$

Then as the robustness measure $\rho$ satisfies Assumption 2 with Lipschitz constant $L_\rho$ and time constant $b$, we have that $\forall\ x_0 \in \mathcal{X}$ and with $\|\cdot\|_{[0,b]}$ the induced signal norm,

$$\left|\rho(\phi(x_0), 0) - \rho\left(\phi^d(x_0), 0\right)\right| \leq L_\rho \left\|\phi(x_0) - \phi^d(x_0)\right\|_{[0,b]}.$$

Then, by definition of $\|\cdot\|_{[0,b]}$ and our fairness assumption that $b \in I(x_0) \cap I^d(x_0) \ \forall\ x_0 \in \mathcal{X}$, we have that

$$L_\rho \left\|\phi(x_0) - \phi^d(x_0)\right\|_{[0,b]} \leq L_\rho M b e^{L_f b}, \ \forall\ x_0 \in \mathcal{X}.$$

As a result, with $M = \delta_d^1 = \Delta_d/(L_\rho b e^{L_f b})$ we have that

$$\left|\rho(\phi(x_0), 0) - \rho\left(\phi^d(x_0), 0\right)\right| \leq \Delta_d, \ \forall\ x_0 \in \mathcal{X}.$$

By definition of $\Delta_d$ and $M$ and the above inequality holding $\forall\ x_0 \in \mathcal{X}$, we have that

$$\rho\left(\phi^d(x_0), 0\right) \geq 0, \ \forall\ x_0, d \ \text{ s.t. } x_0 \in \mathcal{X}, \ \|d\| \leq \delta_d^1.$$

Then the result follows by Definitions 1 and 4. ■

Now it remains to identify a composite disturbance-bound for specifications $\psi = \wedge_i \psi_i$ where each $\psi_i$ is one of the base specification forms already accounted for, i.e., $\mathbf{G}_{[0,b]}\,\omega, \mathbf{G}_{[a,b]}\,\omega, \mathbf{F}_{[a,b]}\,\omega$, or $\omega_1 \,\mathrm{U}_{[a,b]}\,\omega_2$. To do so, we will define an inclusion symbol for specifications.

$$\psi_i \in \psi \iff \psi = \wedge_i \psi_i, \ P^1(\psi) = \{\psi' | \psi' \in \psi\},$$

$$\text{e.g., for } \psi = \psi_1 \wedge (\psi_2 \wedge \psi_3), \ \psi_1, \psi_2, \psi_3 \in \psi.$$

This leads to the following lemma similar to Lemma 1.

*Lemma 3:* $(s, 0) \models \psi \iff (s, 0) \models \psi', \ \forall\ \psi' \in P^1(\psi)$.

*Proof:* This follows from the definition of the satisfaction relation $\models$ as defined in [6]. For the sake of brevity, the full proof can be found in [28]. ■

Then our final theorem determines a disturbance-bound $\delta_d$ for specifications $\psi = \wedge_i \psi_i$ where each $\psi_i$ is one of the base specification forms mentioned prior. We will first pose our optimization problem, then state our theorem. In what follows, $\delta_d^0$ is as per (11) and $\delta_d^1$ is as per (15).

$$\delta_d^T(\psi) = \min_{\psi_i \in P^1(\psi)} \begin{cases} \delta_d^0 & \text{if } \psi_i \text{ is as per (10)}, \\ \delta_d^1 & \text{else}, \end{cases} \tag{16}$$

$$\mathcal{C}_\psi = \mathcal{X} \bigcap_{\substack{\psi_i \in P^1(\psi) \text{ s.t.} \\ \psi_i \text{ as per (10)}}} \mathcal{C}_\omega \text{ as per (10).} \tag{17}$$

*Theorem 4:* Let the system's specification $\psi$ satisfy (8) and let the assumptions for Theorems 1 and 3 hold. If $\delta_d^T(\psi) \geq 0$ with $\delta_d^T(\psi)$ as per equation (16), then

$$\text{(CL-d)} \models_{\mathcal{C}_\psi} \psi \ \forall \ d \ \text{s.t.} \ \|d\| \leq \delta_d^T(\psi).$$

*Proof:* To start, we can assume without loss of generality that there exist zero or more specifications $\psi_i \in P^1(\psi)$ that are of the form in equation (10). By definition of $\delta_d^T(\psi)$ in equation (16), $\mathcal{C}_\psi$ in equation (17), and Theorem 1, we have for each such specification $\psi_i$ (should they exist),

$$\text{(CL-d)} \models_{\mathcal{C}_\psi} \psi_i \ \forall \ d \ \text{s.t.} \ \|d\| \leq \delta_d^T(\psi).$$

This follows as if we have two sets $A, B$ such that $A \subset B$, a system $S$, and a specification $\psi$, then by Definition 4,

$$S \models_B \psi \implies S \models_A \psi.$$

Then we can also assume without loss of generality that we have zero or more specifications $\psi_j \in P^1(\psi)$ such that $\psi_j$ are not of the form in equation (10). For each such $\psi_j$, by definition of $\delta_d^T(\psi)$, $\mathcal{C}_\psi$, and Theorem 3, we have that

$$\text{(CL-d)} \models_{\mathcal{C}_\psi} \psi_j \ \forall \ d \ \text{s.t.} \ \|d\| \leq \delta_d^T(\psi).$$

Then the result holds via Lemma 3. ■

This ends the series of optimization problems to determine our disturbance-bounds. We will now move to showcase these results through a simulated example on a Segway.

## IV. Simulated Examples

For our example, we aim to determine the robustness with which a Segway's LQR controller achieves two desired performance bounds. First, the Segway's pendulum angle is never to deviate too far from the vertical. Second, the Segway is to reach its goal - its state $\mathbf{x}$ should lie within a norm bounded ball around 0 - within two seconds. Mathematically this leads to the following setting:

$$h_1(\mathbf{x}) = 0.25 - \|\mathbf{x}\|, \ h_2(\mathbf{x}) = 10(0.3^2 - \theta^2) - 2\theta\dot{\theta}, \tag{18}$$

$$\begin{aligned} &\mu_i(x) \equiv (h_i(x) \geq 0), \quad \psi = \mathbf{F}_{[0,2]}\,\mu_1 \wedge \mathbf{G}_{[0,2]}\,\mu_2, \\ &\mathcal{X} \subset [-1, 1]^2 \times [-0.4, 0.4] \times [-1.5, 1.5], \\ &\mathbf{x} = [x, v, \theta, \dot{\theta}]^T \in \mathcal{X} \subset \mathbb{R}^4. \end{aligned} \tag{19}$$

Figure 3 shows the Segway setup and example LQR controller steering the Segway to satisfy this specification $\psi$.

To start, it is clear that both predicate functions $h_1, h_2$ in equation (18) satisfy Assumption 1. Indeed as both are Lipschitz continuous, so to are the associated robustness measures generated from these predicate functions Lipschitz continuous as well, which satisfies Assumption 2. As a result, we break our specification into two parts as required of Theorem 4 - $\psi_1 = \mathbf{F}_{[0,2]}\,\mu_1$ and $\psi_2 = \mathbf{G}_{[0,2]}\,\mu_2$. This resulted in a $\delta_d^0 = 0.89$ after utilizing Theorem 1 for $\psi_2$ and a $\Delta_d = 0.2$ after utilizing Theorem 3 for $\psi_1$.

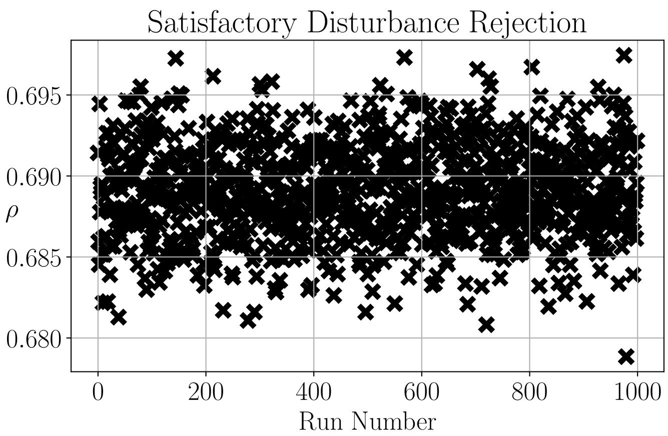

Fig. 2. The robustness measure for 1000 trials of the Segway detailed in Section IV when perturbed by randomly distributed disturbances whose two-norm is less than the upper bound calculated by Theorem 1, $\delta_d^0 = 0.89$. The robustness measure $\rho$ is for the specification $\psi_2 = \mathbf{G}_{[0,2]}\,\mu_2$ as per equation (7). In all cases, the system satisfies its specification as $\rho(\phi^d(x_0, 0)) \geq 0$. This success indicates that, with high probability, this Segway's LQR controller rejects disturbances whose norm $\|d\| \leq \delta_d^0$.

Figure 2 shows the results of 1000 randomized trials of the Segway undergoing disturbances $d$ such that $\|d\| \leq \delta_d^0 = 0.89$. As can be seen, the LQR controller realizes a positive robustness measure indicating that the system-controller pair can reject disturbances whose norm is under the bound we identify through our procedure. Additionally, under the assumption that our Segway's closed-loop dynamics $f_{cl}$ are Lipschitz continuous with constant $L_f \leq 1$ and knowing the associated robustness measure $\rho$ for $\mu_1$ as per (19) is Lipschitz continuous with $L_\rho = 1$, Theorem 3 provides a secondary disturbance-bound $\delta_d^1 = 0.01$. As per Theorem 4 this indicates that our Segway should satisfy its overall specification $\psi$ if its disturbance $d$ is such that $\|d\| \leq \delta_d^T(\psi) = 0.01$. Indeed the Segway does satisfy its specification after 1000 randomized runs when perturbed by normally distributed disturbances $d$ such that $\|d\| \leq 0.01$. One such run is shown in Figure 3.

*Remark on Conservative Bounds:* First, we note that assuming our Segway dynamics $f_{cl}$ to be Lipschitz continuous with $L_f \leq 1$ may not be true. However, $L_f > 1$ only decreases the resulting disturbance bound as shown in equation (15), and as we have shown that our bound $\delta_d^1 = 0.01$ is accurate, so to would any lower bound also be accurate. Second, we understand our resulting disturbance bound is conservative. This conservatism arises primarily through the optimization problem over signal trajectories in (14) and the growing Lipschitz signal tube proportional to $te^{Lt}$ as expressed in Lemma 2. As the time horizon for the future objective increases, the resulting disturbance bound will necessarily decrease due to this proportionality. This also implies that the system may be capable of rejecting disturbances whose two-norm is larger than our calculated bound. To resolve this issue, the ideal would be to develop an optimization problem over CBF-like conditions for the predicate functions $h_\mu$ for the specifications $\psi$ of the form in (13) - this is the subject of future work.

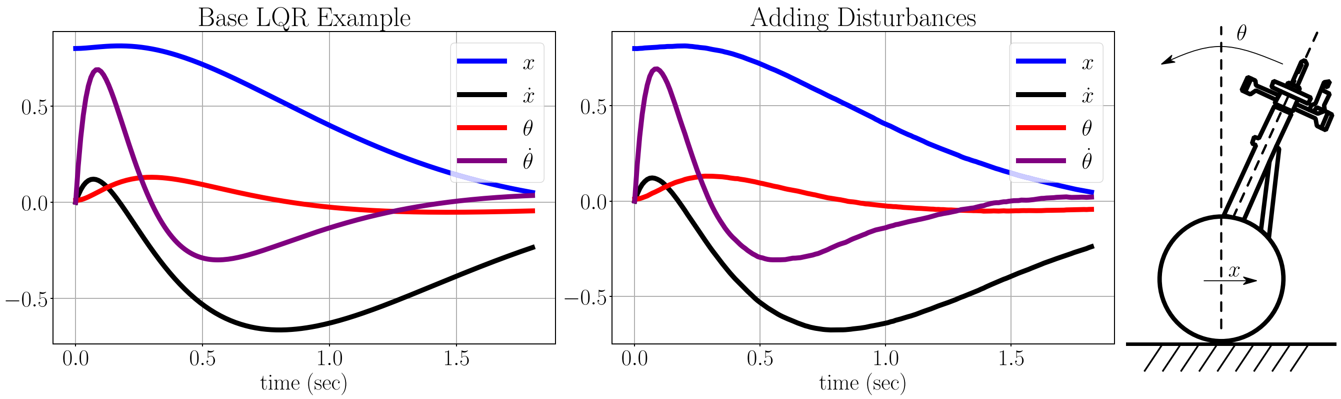

Fig. 3. Comparison of a Segway's LQR controller steering the nominal system (CL) to the zero point (left) and the disturbed system (CL-d) to the same zero point (center). The example Segway is illustrated to the far right. Notice that when the Segway undergoes disturbances whose norms are less than the max bound calculated via our procedure $\delta_d^T(\psi) = 0.01$, the specification $\psi$ in equation (18) is still satisfied.

## V. Conclusion

We constructed a series of optimization problems to determine a two-norm disturbance bound a system's controller can reject while satisfying its operational STL specification. We also showed our bounds were reasonable through a simulated example of a Segway perturbed by disturbances whose norm was less than our calculated bound. Future work aims to analyze the conservativeness of our generated bounds.

## References

[1] F. Baader, D. Calvanese, D. McGuinness, P. Patel-Schneider, and D. Nardi, *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[2] C. Baier and J.-P. Katoen, *Principles of Model Checking*. Cambridge, MA, USA: MIT Press, 2008.

[3] A. Donzé and O. Maler, "Robust satisfaction of temporal logic over real-valued signals," in *Proc. Int. Conf. Formal Model. Anal. Timed Syst.*, 2010, pp. 92–106.

[4] M. Y. Vardi, "An automata-theoretic approach to linear temporal logic," in *Logics for Concurrency*. New York, NY, USA: Springer, 1996, pp. 238–266.

[5] R. Gerth, D. Peled, M. Y. Vardi, and P. Wolper, "Simple on-the-fly automatic verification of linear temporal logic," in *Proc. Int. Conf. Protocol Specification Test. Verification*, 1995, pp. 3–18.

[6] O. Maler and D. Nickovic, "Monitoring temporal properties of continuous signals," in *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*. Heidelberg, Germany: Springer, 2004, pp. 152–166.

[7] E. M. Wolff, U. Topcu, and R. M. Murray, "Optimization-based trajectory generation with linear temporal logic specifications," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2014, pp. 5319–5325.

[8] T. Wongpiromsarn, U. Topcu, N. Ozay, H. Xu, and R. M. Murray, "Tulip: A software toolbox for receding horizon temporal logic planning," in *Proc. 14th Int. Conf. Hybrid Syst. Comput. Control*, 2011, pp. 313–314.

[9] L. Lindemann and D. V. Dimarogonas, "Control barrier functions for signal temporal logic tasks," *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 96–101, Jan. 2019.

[10] L. Lindemann and D. V. Dimarogonas, "Control barrier functions for multi-agent systems under conflicting local signal temporal logic tasks," *IEEE Control Syst. Lett.*, vol. 3, no. 3, pp. 757–762, Jul. 2019.

[11] M. Althoff and S. Lutz, "Automatic generation of safety-critical test scenarios for collision avoidance of road vehicles," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2018, pp. 1326–1333.

[12] Y. Annpureddy, C. Liu, G. Fainekos, and S. Sankaranarayanan, "S-TaLiRo: A tool for temporal logic falsification for hybrid systems," in *Proc. Int. Conf. Tools Algorithms Construct. Anal. Syst.*, 2011, pp. 254–257.

[13] C. E. Tuncali, T. P. Pavlic, and G. Fainekos, "Utilizing S-TaLiRo as an automatic test generation framework for autonomous vehicles," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, 2016, pp. 1470–1475.

[14] T. Dreossi *et al.*, "VerifAI: A toolkit for the formal design and analysis of artificial intelligence-based systems," in *Proc. Int. Conf. Comput. Aided Verification*, 2019, pp. 432–442.

[15] S. Ghosh, F. Berkenkamp, G. Ranade, S. Qadeer, and A. Kapoor, "Verifying controllers against adversarial examples with Bayesian optimization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2018, pp. 7306–7313.

[16] Y. Huang and W. Xue, "Active disturbance rejection control: Methodology and theoretical analysis," *ISA Trans.*, vol. 53, no. 4, pp. 963–976, 2014.

[17] Z. Gao, "Active disturbance rejection control: A paradigm shift in feedback control system design," in *Proc. Amer. Control Conf.*, 2006, p. 7.

[18] D. Sun, "Comments on active disturbance rejection control," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 3428–3429, Dec. 2007.

[19] A. Corso, R. J. Moss, M. Koren, R. Lee, and M. J. Kochenderfer, "A survey of algorithms for black-box safety validation," 2020, *arXiv:2005.02979*

[20] G. E. Fainekos and G. J. Pappas, "Robustness of temporal logic specifications for continuous-time signals," *Theor. Comput. Sci.*, vol. 410, no. 42, pp. 4262–4291, 2009.

[21] L. Lindemann and D. V. Dimarogonas, "Barrier function based collaborative control of multiple robots under signal temporal logic tasks," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 4, pp. 1916–1928, Dec. 2020.

[22] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Rev.*, vol. 44, no. 4, pp. 525–597, 2002.

[23] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems*. Heidelberg, Germany: Springer, 2006.

[24] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3861–3876, Aug. 2017.

[25] R. P. Agarwal, R. P. Agarwal, and V. Lakshmikantham, *Uniqueness and Nonuniqueness Criteria for Ordinary Differential Equations*, vol. 6. River Edge, NJ, USA: World Sci., 1993.

[26] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002. [Online]. Available: https://cds.cern.ch/record/1173048

[27] W. F. Ames and B. Pachpatte, *Inequalities for Differential and Integral Equations*, vol. 197. San Diego, CA, USA: Elsevier, 1997.

[28] P. Akella and A. D. Ames, "Disturbance bounds for signal temporal logic task satisfaction: A dynamics perspective," 2021, *arXiv:2110.12014*.