Risk of Stochastic Systems for Temporal Logic Specifications

LARS LINDEMANN, University of Southern California, USA
LEJUN JIANG, Nuro, USA
NIKOLAI MATNI and GEORGE J. PAPPAS, University of Pennsylvania, USA

The wide availability of data coupled with the computational advances in artificial intelligence and machine learning promise to enable many future technologies such as autonomous driving. While there has been a variety of successful demonstrations of these technologies, critical system failures have repeatedly been reported. Even if rare, such system failures pose a serious barrier to adoption without a rigorous risk assessment. This article presents a framework for the *systematic and rigorous risk verification* of systems. We consider a wide range of system specifications formulated in signal temporal logic (STL) and model the system as a stochastic process, permitting discrete-time and continuous-time stochastic processes. We then define the STL robustness risk as *the risk of lacking robustness against failure*. This definition is motivated as system failures are often caused by missing robustness to modeling errors, system disturbances, and distribution shifts in the underlying data generating process. Within the definition, we permit general classes of risk measures and focus on tail risk measures such as the value-at-risk and the conditional value-at-risk. While the STL robustness risk is in general hard to compute, we propose the approximate STL robustness risk as a more tractable notion that upper bounds the STL robustness risk. We show how the approximate STL robustness risk can accurately be estimated from system trajectory data. For discrete-time stochastic processes, we show under which conditions the approximate STL robustness risk can even be computed exactly. We illustrate our verification algorithm in the autonomous driving simulator CARLA and show how a least risky controller can be selected among four neural network lane-keeping controllers for five meaningful system specifications.

CCS Concepts: • **Computer systems organization** → **Embedded and cyber-physical systems**; *Robotics*; • **Theory of computation** → **Logic**; *Logic and verification*; *Modal and temporal logics*; • **General and reference** → *Verification*; • **Mathematics of computing** → **Probability and statistics**; *Stochastic processes;*

Additional Key Words and Phrases: Risk-aware decision making, stochastic system verification, signal temporal logic, stochastic robustness

ACM Reference format:
Lars Lindemann, Lejun Jiang, Nikolai Matni, and George J. Pappas. 2023. Risk of Stochastic Systems for Temporal Logic Specifications. *ACM Trans. Embedd. Comput. Syst.* 22, 3, Article 54 (April 2023), 31 pages.
https://doi.org/10.1145/3580490

This research was supported by NSF award CPS-2038873, NSF CAREER award ECCS-2045834, and a Google Research Scholar award.

Authors' addresses: L. Lindemann, University of Southern California, 941 Bloom Walk, Los Angeles, CA, 90089; email: llindema@usc.edu; L. Jiang, Nuro, 1300 Terra Bella Ave, Mountain View, CA, 94043; email: lejunj@seas.upenn.edu; N. Matni and George J. Pappas, University of Pennsylvania, 200 South 33rd Street, Philadelphia, Pennsylvania, 19104; emails: {nmatni, pappasg}@seas.upenn.edu.

1539-9087/2023/04-ART54 $15.00
https://doi.org/10.1145/3580490

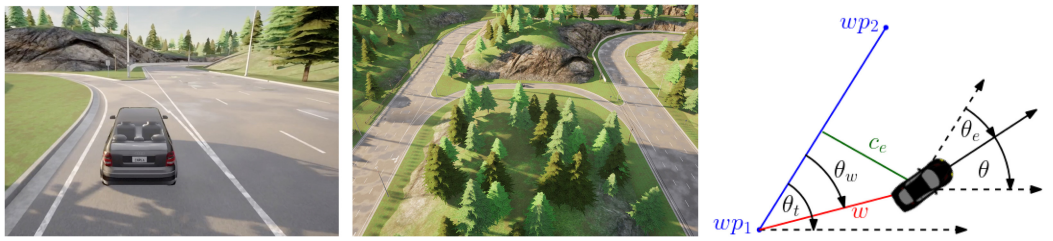


Fig. 1. Left: Simulation environment in the autonomous driving simulator CARLA. Middle: Double left turn on which we evaluate four trained neural network lane-keeping controllers. Right: Cross-track error c_e and orientation error θ_e used for risk verification of the neural network controllers.

1 INTRODUCTION

Over the next decade, large amounts of data will be generated and stored as devices that perceive and control the world become more affordable and available. Impressive demonstrations of data-driven and machine learning-enabled technologies exist already today, e.g., robotic manipulation [44], solving games [55, 74], and autonomous driving [19]. However, occasionally occurring system failures impede the use of these technologies particularly when system safety is a concern. For instance, neural networks, frequently used for perception and control in autonomous systems, are known to be fragile and non-robust [25, 77]. Especially the problem of long tails in training data distributions poses challenges, e.g., natural variations in weather and lighting conditions [61].

Moving forward, we expect that system failures appear less frequently due to advancing technologies—nonetheless, algorithms for the systematic and rigorous *risk verification* of such systems is needed. For instance, the National Transportation Safety Board emphasized in a statement in connection with an Uber accident from 2018 "the need for safety risk management requirements for testing automated vehicles on public roads" [12]. In this article, we show how to reason about the risk of systems that are modeled as stochastic processes. We consider a wide range of system specifications formulated in signal temporal logic (STL) [9, 51] and present a systematic way to quantify and compute the risk of a system lacking robustness against failure.

1.1 Related Work

Depending on the research disciplines and applications, risk can have various interpretations. While risk is often defined as a failure probability, it can also be understood in more general terms as a metric defined over a cost distribution, e.g., the expected value or the variance of a distribution. We focus on *tail risk measures* to capture the rare yet costly events of a distribution. In particular, we consider the value-at-risk (VaR), i.e., quantiles of a distribution, and the conditional value-at-risk (CVaR) [62, 63], i.e., the expected value over a quantile. Tail risk measures are more frequently being used in robotics and control applications where system safety is important [50].

Risk in control. Control design under risk objectives and constraints is increasingly been studied among control theorists, as machine learning components integrated into closed-loop systems cause stochastic system uncertainty. Oftentimes, the CVaR risk measure is used to capture risk due its convexity and the property of being an upper bound to the VaR. For instance, the authors in Reference [72] consider a stochastic optimal control problem with CVaR constraints over the distance to obstacles. Linear quadratic control under risk constraints was considered in Reference [81] to trade off risk and mean performance. A similar idea is followed for the risk constrained minimum mean squared error estimator in Reference [37]. Risk-aware model predictive control was considered in References [29, 76], while References [17, 73] present data-driven and distributionally

robust model predictive controllers. Risk-aware control barrier functions for safe control synthesis were proposed in Reference [2], while Reference [58] demonstrates the use of risk in sampling-based planning. We remark that we view these works to be orthogonal to our article, as we provide a data-driven framework for the risk assessment under complex temporal logic specifications, and we hope to inform future control design strategies.

Stochastic system verification. System verification has a long history in complementing and informing the control design process of systems, e.g., using model checking [6, 14]. When dealing with stochastic systems, system verification becomes computationally more challenging [40]. Statistical model checking has recently gained attention by relying on availability of data instead of computation [1, 22, 43, 85]. Another line of work considers stochastic barrier functions for safety verification of dynamical systems [32, 59]. The authors in References [33, 34] deal with the verification of stochastic dynamical systems during runtime. Motivated by the fragility and sensitivity of neural networks [25, 77], a special focus has recently been on verifying neural networks in open-loop [38, 75] and closed-loop [30]. We remark that our algorithms presented in this article permit verification of general classes of systems, including systems with neural networks, as long as we can obtain data, e.g., from a simulator. The guarantees obtained in these previous works are either worst-case guarantees or in terms of failure probabilities. Towards incorporating tail risk measures, the authors in References [15, 16] propose a risk-aware safety analysis framework using the CVaR. We are instead interested in system verification under more complex temporal logic specifications and risk.

Temporal logics. We use signal temporal logic to express a wide range of system specifications, e.g., surveillance ("visit regions A, B, and C every $10-60$ sec"), safety ("always between $5-25$ sec stay at least 1 m away from region D"), and many others. For deterministic signals, STL allows to calculate the robustness by which a signal satisfies an STL specification. Particularly, the authors in Reference [21] proposed the robustness degree as the maximal tube around a signal in which all signals satisfy the specification. The size of the tube consequently measures the robustness of this signal with respect to the specification. As the robustness degree is in general hard to calculate, the authors in Reference [21] proposed approximate yet easier to calculate robust semantics. Many forms of robust semantics have appeared, such as space and time robustness [18], the arithmetic-geometric mean robustness [53], the smooth cumulative robustness [28], averaged STL [3], and Reference [64], in which a connection with linear time-invariant filtering is established allowing to define various types of robust semantics.

For stochastic signals, the authors in References [35, 41, 45, 67, 80] propose notions of probabilistic signal temporal logic in which chance constraints over predicates are considered, while the Boolean and temporal operators of STL are not changed. Similarly, notions of risk signal temporal logic have recently appeared in References [46, 48, 69] by defining risk constraints over predicates while not changing the definitions of Boolean and temporal operators. In this article, we instead define risk over the whole STL specification. The work in Reference [23] considers the probability of an STL specification being satisfied instead of using chance or risk constraints over predicates. The authors in Reference [84] consider hyperproperties in STL, i.e., properties between multiple system executions. More with a control synthesis focus and for the less-expressive formalism of linear temporal logic, the authors in References [10, 42, 82] consider control over belief spaces, while the authors in Reference [27] consider probabilistic satisfaction over Markov decision processes. Complementary to these works, References [5, 60] propose techniques to infer STL specifications from data towards explaining the underlying data.

Risk verification with temporal logics. In this article, we quantify and compute the risk of lacking robustness against failure. We argue that the consideration of robustness in system

verification is crucial and are particularly motivated by the fact that system failures are often caused by missing robustness to modeling errors, system disturbances, and distribution shifts in the underlying data generating process. The authors in Reference [4] further highlight the importance of robustness in system verification. Probably closest to our article are the works in References [31, 70, 71] and References [7, 8]. In References [31, 70, 71], the authors combine data-driven and model-based verification techniques to obtain information about the satisfaction probability of a partially known system. The authors in References [7, 8] present a purely data-driven verification technique to estimate probabilities over robustness distributions of the system. Conceptually, our work differs in two directions. First, we consider general risk measures to be able to focus on the tails of the robustness distribution. We also show how to estimate the robustness risk from data with high confidence. Second, we use the robustness degree as defined in Reference [21] to obtain robustness distributions. This in fact allows us to obtain a precise geometric interpretation of risk. This article is based on our previous work [47]. We here permit continuous-time stochastic processes and the CVaR as a risk measure. We also show under which conditions the STL robustness risk can exactly be calculated, while presenting exhaustive simulations within the autonomous driving simulator CARLA [19].

1.2 Contributions and Article Outline

Our general goal is to analyze the robustness of stochastic processes and to quantify and compute the *risk of a system lacking robustness against system failure*. We make the following contributions:

- We consider discrete-time and continuous-time stochastic processes and show under which conditions the robust semantics and the robustness degree of STL are random variables. This enables us to define risk over these quantities.
- We define the STL robustness risk as the risk of a system lacking robustness against failure of an STL specification. The definition permits general classes of risk measures and has a precise geometric interpretation in terms of the size of permissible disturbances. We also define the approximate STL robustness risk as a computationally tractable upper bound of the STL robustness risk.
- For the VaR and the CVaR, we show how the approximate STL robustness risk can be estimated from system trajectory data. Importantly, no particular restriction on the distribution of the stochastic process has to be made. For discrete-time stochastic processes with a discrete state space, we show how the approximate STL robustness risk can even be computed exactly.
- We estimate the risk of four neural network lane-keeping controllers within the autonomous driving simulator CARLA. We show how to find the least risky controller.

In Section 2, we present background on signal temporal logic, stochastic processes, and risk measures. In Section 3, we define the STL robustness risk and the STL approximate robustness risk. Section 4 shows how the approximate STL robustness risk can be estimated from data, while Section 5 shows under which conditions it can be computed exactly. The simulation results within CARLA are presented in Section 6 followed by conclusions in Section 7.

2 BACKGROUND

We first provide background on signal temporal logic, stochastic processes, and risk measures.

2.1 Signal Temporal Logic

Signal temporal logic (STL) is based on deterministic signals $x : T \to \mathbb{R}^n$ where T denotes the time domain [51]. We particularly consider continuous time $T := \mathbb{R}$ (the set of real numbers) and

discrete time $T := \mathbb{Z}$ (the set of natural numbers). The atomic elements of STL are predicates that are functions $\mu : \mathbb{R}^n \to \mathbb{B}$ where $\mathbb{B} := \{\top, \bot\}$ is the set of Booleans consisting of the true and false elements $\top := 1$ and $\bot := -1$, respectively. Let us associate an observation map $O^\mu \subseteq \mathbb{R}^n$ with a predicate μ that indicates regions within the state space where the predicate μ is true, i.e.,

$$O^\mu := \mu^{-1}(\top),$$

where $\mu^{-1}(\top)$ denotes the inverse image of $\top$ under the function μ. We assume throughout the article that the sets O^μ and $O^{\neg\mu}$ are non-empty and measurable, which is a mild technical assumption. In other words, the sets O^μ and $O^{\neg\mu}$ are elements of the Borel σ-algebra $\mathcal{B}^n$ of $\mathbb{R}^n$.

Remark 1. For convenience, the predicate μ is often defined via a predicate function $h : \mathbb{R}^n \to \mathbb{R}$ as

$$\mu(\zeta) := \begin{cases} \top & \text{if } h(\zeta) \geq 0 \\ \bot & \text{otherwise} \end{cases}$$

for $\zeta \in \mathbb{R}^n$. In this case, we have $O^\mu = \{\zeta \in \mathbb{R}^n | h(\zeta) \geq 0\}$.

The syntax of STL, which recursively allows to formulate system specifications, is defined as

$$\phi \ ::= \ \top \mid \mu \mid \neg\phi \mid \phi' \wedge \phi'' \mid \phi' U_I \phi'' \mid \phi' \underline{U}_I \phi'', \tag{1}$$

where ϕ' and ϕ'' are STL formulas and where U_I is the future until operator with time interval $I \subseteq \mathbb{R}_{\geq 0}$, while $\underline{U}_I$ is the past until-operator. The Boolean operators $\neg$ and $\wedge$ encode negations and conjunctions, respectively. We say that an STL formula ϕ as in Equation (1) is bounded if the time interval I is restricted to be compact. Based on these elementary operators, we can define the set of operators

$$\begin{aligned} \phi' \vee \phi'' &:= \neg(\neg\phi' \wedge \neg\phi'') && \text{(disjunction operator)}, \\ F_I\phi &:= \top U_I \phi && \text{(future eventually operator)}, \\ \underline{F}_I\phi &:= \top \underline{U}_I \phi && \text{(past eventually operator)}, \\ G_I\phi &:= \neg F_I \neg\phi && \text{(future always operator)}, \\ \underline{G}_I\phi &:= \neg \underline{F}_I \neg\phi && \text{(past always operator)}. \end{aligned}$$

2.1.1 Semantics. To determine whether or not a signal $x : T \to \mathbb{R}^n$ satisfies an STL formula ϕ, we define the semantics of ϕ by means of the satisfaction function $\beta^\phi : \mathfrak{F}(T, \mathbb{R}^n) \times T \to \mathbb{B}$.[1] In particular, $\beta^\phi(x, t) = \top$ indicates that the signal x satisfies the formula ϕ at time t, while $\beta^\phi(x, t) = \bot$ indicates that x does not satisfy ϕ at time t. While the intuitive meanings of the Boolean operators $\neg$ ("not"), $\wedge$ ("and"), and $\vee$ ("or") are clear, we note that the future until operator $\phi' U_I \phi''$ encodes that ϕ' holds until ϕ'' holds. Specifically, $\beta^{\phi' U_I \phi''}(x, t) = \top$ means that ϕ' holds for all times after t (not necessarily at time t) until ϕ'' holds within the time interval $(t \oplus I) \cap T$.[2] Similarly, $\beta^{F_I\phi}(x, t) = \top$ encodes that ϕ holds eventually within $(t \oplus I) \cap T$, while $\beta^{G_I\phi}(x, t) = \top$ encodes that ϕ holds always within $(t \oplus I) \cap T$. For a formal definition of $\beta^\phi(x, t)$, we refer to Appendix A.

We are usually interested in the satisfaction function $\beta^\phi(x, 0)$, which determines the satisfaction of ϕ by x at time zero, the time at which we assume ϕ to be enabled. An STL formula ϕ is hence said to be satisfiable if $\exists x \in \mathfrak{F}(T, \mathbb{R}^n)$ such that $\beta^\phi(x, 0) = \top$. The following example is taken from Lindemann et al. [47] and used as a running example throughout the article:

[1] We use the notation $\mathfrak{F}(A, B)$ to denote the set of all measurable functions mapping from the domain A into the domain B, i.e., an element $f \in \mathfrak{F}(A, B)$ is a measurable function $f : A \to B$.

[2] We use the notation $\oplus$ and $\ominus$ to denote the Minkowski sum and the Minkowski difference, respectively.

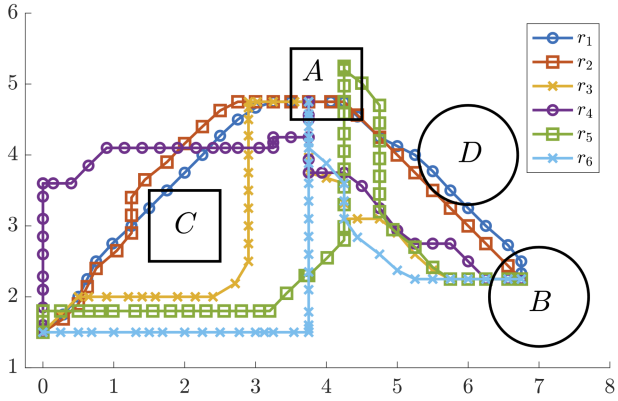


Fig. 2. The figure shows six potential robot trajectories r_1-r_6 and the four regions A, B, C, and D. The specification given in Equation (2) is violated by r_1 and satisfied by r_2-r_6. It can be seen that r_2 only marginally satisfies ϕ, while r_3-r_6 satisfy ϕ robustly.

Example 1. Consider a delivery robot that needs to perform two time-critical delivery tasks in regions A and B sequentially while avoiding areas C and D; see Figure 2. We consider the STL formula

$$\phi := G_{[0,3]}(\neg\mu_C \wedge \neg\mu_D) \wedge F_{[1,2]}(\mu_A \wedge F_{[0,1]}\mu_B), \tag{2}$$

where the regions A, B, C, and D are encoded by the predicates μ_A, μ_B, μ_C, and μ_D, respectively, that are defined below. Let the state $x(t) \in \mathbb{R}^{10}$ of the system at time t be

$$x(t) := [r(t) \quad a \quad b \quad c \quad d]^T$$

where $r(t)$ is the robot position at time t and where a, b, c, and d denote the center points of the regions A, B, C, and D that are defined as

$$a := [4 \quad 5]^T \quad b := [7 \quad 2]^T \quad c := [2 \quad 3]^T \quad d := [6 \quad 4].^T$$

The predicates μ_A, μ_B, μ_C, and μ_D are now defined by their observation maps

$$O^{\mu_A} := \{x \in \mathbb{R}^{10} | \|r - a\|_\infty \leq 0.5\},$$
$$O^{\mu_B} := \{x \in \mathbb{R}^{10} | \|r - b\|_2 \leq 0.7\},$$
$$O^{\mu_C} := \{x \in \mathbb{R}^{10} | \|r - c\|_\infty \leq 0.5\}, \tag{3}$$
$$O^{\mu_D} := \{x \in \mathbb{R}^{10} | \|r - d\|_2 \leq 0.7\}, \tag{4}$$

where $\|\cdot\|_2$ is the Euclidean and $\|\cdot\|_\infty$ is the infinity norm. In Figure 2, six different robot trajectories r_1-r_6 are shown. It can be seen that the signal x_1 that corresponds to r_1 violates ϕ, while x_2-x_6 satisfy ϕ, i.e., we have $\beta^\phi(x_1, 0) = \bot$ and $\beta^\phi(x_j, 0) = \top$ for all $j \in \{2, \ldots, 6\}$.

Remark 2. The operators U_I and $\underline{U}_I$ are the strict non-matching versions of the until operators. In particular, $\phi' U_I \phi''$ is: (1) strict, as it does not require ϕ' to hold at the current time t, and (2) non-matching, as it does not require that ϕ' and ϕ'' have to hold at the same time. When dealing with continuous-time stochastic systems later in this article, we replace the strict non-matching versions U_I and $\underline{U}_I$ by the non-strict matching versions that we denote by $\vec{U}_I$ and $\vec{\underline{U}}_I$; see Appendix A for their formal definitions. We note that STL with until operators U_I and $\underline{U}_I$ is more expressive than STL with $\vec{U}_I$ and $\vec{\underline{U}}_I$. When excluding Zeno-signals, there is, however, no difference between these two notions [24]. As one rarely encounters Zeno-signals, we argue that the

restriction to the non-strict matching version of the until operator for continuous-time stochastic processes is not restrictive in practice.

2.1.2 Robustness Degree. Importantly, one may also be interested in the quality of satisfaction and additionally ask how robustly the signal x satisfies the STL formula ϕ at time t. To answer this question, the authors in Fainekos and Pappas [21, Definition 7] define the *robustness degree* that we recall next in a slightly modified manner. If $\beta^\phi(x,t) = \top$, then the robustness degree quantifies how much the signal x can be perturbed by additive noise before changing the value of $\beta^\phi(x,t)$. Towards a formal definition, let us first define the set of signals that *violate* ϕ at time t as

$$\mathcal{L}^{\neg\phi}(t) := \{x \in \mathfrak{F}(T,\mathbb{R}^n) | \beta^{\neg\phi}(x,t) = \top\}.$$

To measure distances between signals, let us define the metric $\kappa : \mathfrak{F}(T,\mathbb{R}^n) \times \mathfrak{F}(T,\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ as

$$\kappa(x,x') := \sup_{t\in T} d\Big(x(t), x'(t)\Big),$$

where $\overline{\mathbb{R}}_{\geq 0} := \mathbb{R}_{\geq 0} \cup \{\infty\}$ is the set of nonnegative extended real numbers and where $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a metric assigning a distance in $\mathbb{R}^n$, e.g., the Euclidean norm. Throughout the article, we use the extended definitions of the supremum and infimum operators, e.g., $\sup \mathbb{R} = \infty$. Note that $\kappa(x,x')$ is the L_∞ norm of the signal $x - x'$ and measures the distance between the signals x and x'.

To set some general notation, for a metric space (S,κ) with metric κ, we denote by

$$\bar{\kappa}(x,S') := \inf_{x'\in S'} \kappa(x,x')$$

the distance of a point $x \in S$ to a nonempty set $S' \subseteq S$. Using this definition, the robustness degree $\mathrm{RD}^\phi : \mathfrak{F}(T,\mathbb{R}^n) \times T \to \overline{\mathbb{R}}_{\geq 0}$ is now defined via the metric κ as the distance of the signal x to the set of violating signals $\mathcal{L}^{\neg\phi}(t)$.

Definition 1 (Robustness Degree[3]). For a signal $x : T \to \mathbb{R}^n$ and an STL formula ϕ, the robustness degree $\mathrm{RD}^\phi(x,t)$ is defined as

$$\mathrm{RD}^\phi(x,t) := \bar{\kappa}\Big(x, \mathrm{cl}(\mathcal{L}^{\neg\phi}(t))\Big),$$

where $\mathrm{cl}(\mathcal{L}^{\neg\phi}(t))$ denotes the closure of the set $\mathcal{L}^{\neg\phi}(t)$.

By definition of the robustness degree, the following properties hold: If $\mathrm{RD}^\phi(x,t) > 0$, then $\beta^\phi(x,t) = \top$, i.e., the signal x satisfies ϕ at time t. It further follows that all signals $x' \in \mathfrak{F}(T,\mathbb{R}^n)$ with $\kappa(x,x') < \mathrm{RD}^\phi(x,t)$ are such that $\beta^\phi(x',t) = \top$. The robustness degree defines in fact a *robust neighborhood*, which is a set strictly containing x, so for all x' in this robust neighborhood we have $\beta^\phi(x,t) = \beta^\phi(x',t)$. Finally, note that $\mathrm{RD}^\phi(x,t) = 0$ may imply either $\beta^\phi(x,t) = \top$ or $\beta^\phi(x,t) = \bot$, i.e., the signal x either satisfies or violates ϕ at time t.

2.1.3 Robust Semantics. Note that it is in general difficult to calculate the robustness degree $\mathrm{RD}^\phi(x,t)$, as the set $\mathcal{L}^{\neg\phi}(t)$ is hard to calculate. The authors in Fainekos and Pappas [21] introduce the *robust semantics* $\rho^\phi : \mathfrak{F}(T,\mathbb{R}^n) \times T \to \overline{\mathbb{R}}$ as an alternative way of finding a robust neighborhood where $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ is, in direct analogy to $\overline{\mathbb{R}}_{\geq 0}$, the set of extended real numbers.

[3]The robustness degree in Fainekos and Pappas [21, Definition 7] is defined slightly differently by instead considering the signed distance of the signal x to the set of violating signals $\mathcal{L}^{\neg\phi}(t)$.

Definition 2 (Robust Semantics). For a signal $x : T \to \mathbb{R}^n$ and an STL formula ϕ, the robust semantics $\rho^\phi(x,t)$ are recursively defined as

$$\begin{aligned}
\rho^\top(x,t) &:= \infty,\\
\rho^\mu(x,t) &:= \begin{cases} \bar{d}\big(x(t), \mathrm{cl}(O^{\neg\mu})\big) & \text{if } x(t) \in O^\mu \\ -\bar{d}\big(x(t), \mathrm{cl}(O^\mu)\big) & \text{otherwise,} \end{cases}\\
\rho^{\neg\phi}(x,t) &:= -\rho^\phi(x,t),\\
\rho^{\phi'\wedge\phi''}(x,t) &:= \min(\rho^{\phi'}(x,t), \rho^{\phi''}(x,t)),\\
\rho^{\phi' U_I \phi''}(x,t) &:= \sup_{t''\in(t\oplus I)\cap T}\Big(\min\Big(\rho^{\phi''}(x,t''), \inf_{t'\in(t,t'')\cap T}\rho^{\phi'}(x,t')\Big)\Big),\\
\rho^{\phi' \underline{U}_I \phi''}(x,t) &:= \sup_{t''\in(t\ominus I)\cap T}\Big(\min\Big(\rho^{\phi''}(x,t''), \inf_{t'\in(t'',t)\cap T}\rho^{\phi'}(x,t')\Big)\Big).
\end{aligned}$$

Remark 3. With respect to Remark 2, the non-strict matching version of the until operators replace the open time intervals (t,t'') in Definition 2 by the closed time intervals $[t,t'']$ so

$$\begin{aligned}
\rho^{\phi' \vec{U}_I \phi''}(x,t) &:= \sup_{t''\in(t\oplus I)\cap T}\Big(\min\Big(\rho^{\phi''}(x,t''), \inf_{t'\in[t,t'']\cap T}\rho^{\phi'}(x,t')\Big)\Big),\\
\rho^{\phi' \vec{\underline{U}}_I \phi''}(x,t) &:= \sup_{t''\in(t\ominus I)\cap T}\Big(\min\Big(\rho^{\phi''}(x,t''), \inf_{t'\in[t'',t]\cap T}\rho^{\phi'}(x,t')\Big)\Big).
\end{aligned}$$

Importantly, by slight modification of Fainekos and Pappas [21, Theorem 28], we know that

$$\rho^\phi(x,t) \le \mathrm{RD}^\phi(x,t). \tag{5}$$

The robust semantics $\rho^\phi(x,t)$ hence provides a tractable under-approximation of the robustness degree $\mathrm{RD}^\phi(x,t)$. The robust semantics are sound in the sense that $\beta^\phi(x,t) = \top$ if $\rho^\phi(x,t) > 0$ and $\beta^\phi(x,t) = \bot$ if $\rho^\phi(x,t) < 0$ [21, Proposition 30].

Example 1 (continued). Consider again the trajectories shown in Figure 2. We obtain $\rho^\phi(x_1,0) = -0.15$, $\rho^\phi(x_2,0) = 0.01$, and $\rho^\phi(x_j,0) = 0.25$ for all $j \in \{3,\dots,6\}$. The reason for x_1 having negative robustness lies in r_1 intersecting with the region D. Marginal robustness of x_2 is explained as r_2 only marginally avoids the region D, while all other trajectories avoid the region D robustly.

2.2 Random Variables and Stochastic Processes

Instead of interpreting an STL specifications ϕ over deterministic signals, we will interpret ϕ over stochastic processes. Consider, therefore, the *probability space* $(\Omega, \mathcal{F}, P)$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra of Ω, and $P : \mathcal{F} \to [0,1]$ is a probability measure.

Let Z denote a real-valued *random vector*, i.e., a measurable function $Z : \Omega \to \mathbb{R}^n$. When $n = 1$, we say Z is a *random variable*. We refer to $Z(\omega)$ as a realization of the random vector Z where $\omega \in \Omega$. Since Z is a measurable function, a probability space can be defined for Z so probabilities can be assigned to events related to values of Z.[4] Consequently, a cumulative distribution function (CDF) $F_Z(z)$ can be defined for Z. Given a random vector Z, we can derive other random variables. Assume, for instance, a measurable function $g : \mathbb{R}^n \to \mathbb{R}$, then $g(Z(\omega))$ becomes a *derived random variable*, since function composition preserves measurability; see, e.g., Durrett [20] for more details.

[4]Particularly, this probability space is $(\mathbb{R}^n, \mathcal{B}^n, P_Z)$ where, for Borel sets $B \in \mathcal{B}^n$, the probability measure $P_Z : \mathcal{B}^n \to [0,1]$ is defined as $P_Z(B) := P(Z^{-1}(B))$, where $Z^{-1}(B) := \{\omega \in \Omega | Z(w) \in B\}$ is the inverse image of B under Z.

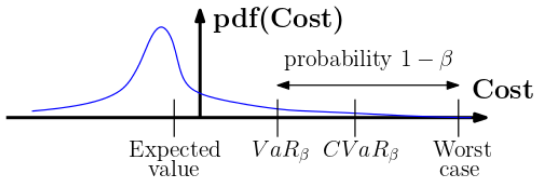


Fig. 3. Illustration of the expected value, the value-at-risk, and the conditional value-at-risk.

A *stochastic process* is a function $X : T \times \Omega \to \mathbb{R}^n$, where $X(t, \cdot)$ is a random vector for each fixed $t \in T$. A stochastic process can be viewed as a collection of random vectors $\{X(t, \cdot) | t \in T\}$ that are defined on a common probability space $(\Omega, \mathcal{F}, P)$ and that are indexed by T. For a fixed $\omega \in \Omega$, the function $X(\cdot, \omega)$ is a *realization* of the stochastic process. Another interpretation is that a stochastic process is a collection of deterministic functions of time $\{X(\cdot, \omega) | \omega \in \Omega\}$ that are indexed by Ω.

2.3 Risk Measures

A *risk measure* is a function $R : \mathfrak{F}(\Omega, \mathbb{R}) \to \mathbb{R}$ that maps from the set of real-valued random variables to the real numbers. In particular, we refer to the input of a risk measure R as the *cost random variable*, since typically a cost is associated with the input of R. Risk measures hence allow for a risk assessment in terms of such cost random variables.

In this article, we particularly use the expected value, the value-at-risk VaR_β, and the conditional value-at-risk $CVaR_\beta$ at risk level $\beta \in (0, 1)$, which are commonly used risk measures; see Figure 3. The VaR_β of a random variable $Z : \Omega \to \mathbb{R}$ is defined as

$$VaR_\beta(Z) := \inf\{\alpha \in \mathbb{R} | F_Z(\alpha) \geq \beta\},$$

i.e., the right $1 - \beta$ quantile of Z. The $CVaR_\beta$ of Z is defined as

$$CVaR_\beta(Z) := \inf_{\alpha \in \mathbb{R}} \left(\alpha + (1 - \beta)^{-1} E([Z - \alpha]^+)\right),$$

where $[Z - \alpha]^+ := \max(Z - \alpha, 0)$. When the CDF F_Z of Z is continuous, it holds that $CVaR_\beta(Z) := E(Z | Z \geq VaR_\beta(Z))$, i.e., $CVaR_\beta(Z)$ is the expected value of Z conditioned on the events where Z is greater or equal than $VaR_\beta(Z)$.

There are various desriable properties that a risk measure R may satisfy; see Majumdar and Pavone [50] for more information. We emphasize that our presented method is compatible with any monotone risk measure, where monotonicity of R is defined as follows:

- For two cost random variables $Z, Z' \in \mathfrak{F}(\Omega, \mathbb{R})$, the risk measure R is *monotone* if

$$Z(\omega) \leq Z'(\omega) \text{ for all } \omega \in \Omega \quad \Longrightarrow \quad R(Z) \leq R(Z').$$

The assumption of considering monotone risk measures is very mild, and both the value-at-risk $VaR_\beta(Z)$ and the conditional value-at-risk $CVaR_\beta(Z)$ as well as the expected value are monotone.

3 THE RISK OF LACKING ROBUSTNESS AGAINST FAILURE

We interpret STL formulas ϕ over stochastic processes X instead of deterministic signals x. It is, however, not immediately clear how to interpret the satisfaction of ϕ by X. One way is to argue about the probability of satisfaction; see, e.g., Farahani et al. [23], but probabilities provide no information about the risk and the robustness of X with respect to ϕ. In fact, some realizations of X may satisfy ϕ robustly, while some other realizations of X may satisfy ϕ only marginally or even violate ϕ. This observation leads us to the use of risk measures R to be able to argue about the risk of the stochastic process X lacking robustness against failure of the specification ϕ.

3.1 Measurability of Semantics, Robustness Degree, and Robust Semantics

To define the risk of a stochastic process X, we first need to show under which conditions the semantics $\beta^\phi(X,t)$, the robustness degree $\mathrm{RD}^\phi(X,t)$, and the robust semantics $\rho^\phi(X,t)$ are derived random variables. For discrete-time stochastic processes, no assumptions have to be made.

Theorem 1. *Let X be a discrete-time stochastic process, i.e., $T := \mathbb{Z}$. Let ϕ be an STL specification as in Equation (1). Then $\beta^\phi(X(\cdot,\omega),t)$, $RD^\phi(X(\cdot,\omega),t)$, and $\rho^\phi(X(\cdot,\omega),t)$ are measurable in ω for a fixed $t \in T$, i.e., $\beta^\phi(X,t)$, $RD^\phi(X,t)$, and $\rho^\phi(X,t)$ are random variables.*

For continuous-time stochastic processes, however, we have to impose additional technical assumptions. Particularly, we have to restrict the class of STL formulas in Equation (1) and make further assumptions on the stochastic process X.

Theorem 2. *Let X be a continuous-time stochastic process, i.e., $T := \mathbb{R}$. Let ϕ be a bounded STL specification as in Equation (1), but where the strict non-matching until operators U_I and $\underline{U}_I$ are replaced with the non-strict matching until operators $\vec{U}_I$ and $\underline{\vec{U}}_I$. Then $\beta^\phi(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$, i.e., $\beta^\phi(X,t)$ is a random variable. If $X(\cdot,\omega) : \Omega \to \mathfrak{F}(T,\mathbb{R}^n)$ is measurable,[5] then $RD^\phi(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$, i.e., $RD^\phi(X,t)$ is a random variable, and if additionally $X(\cdot,\omega)$ is a cadlag function[6] for each $\omega \in \Omega$, then $\rho^\phi(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$, i.e., $\rho^\phi(X,t)$ is a random variable.*[7]

Consequently, the probabilities $P(\beta^\phi(X,t) \in B)$, $P(\rho^\phi(X,t) \in B)$, and $P(\mathrm{RD}^\phi(X,t) \in B)$[8] are well defined for measurable sets B from the corresponding measurable spaces. This enables us to define the STL robustness risk in the next section.

Remark 4. We first note that the assumption of a bounded STL formula ϕ with the non-strict matching until operator is made for a technical reason. While the restriction to bounded formulas limits our expressivity to finite time specifications, the consideration of the non-strict matching until operator is not restrictive, as discussed in Remark 2. We remark that Bartocci et al. [8] showed measurability of $\rho^\phi(X(\cdot,\omega),t)$ under the assumption of a bounded STL specification ϕ with non-strict matching until operators, while we additionally show measurability of the semantics $\beta^\phi(X(\cdot,\omega),t)$ and the robustness degree $\mathrm{RD}^\phi(X(\cdot,\omega),t)$ without any additional continuity assumptions on X. Last, we recall that we do not need to assume that ϕ is bounded for a discrete-time stochastic process as per Theorem 1.

3.2 The STL Robustness Risk

One way of defining the risk associated with a stochastic process X is to consider the satisfaction function $\beta^\phi(X,t)$. However, not much information about the robustness of X can be inferred due to binary encoding of $\beta^\phi(X,t)$. Instead, we consider the risk of the stochastic process X lacking robustness against failure of the specification ϕ by considering the robustness degree $\mathrm{RD}^\phi(X,t)$.

Example 2. Consider an electric RC circuit consisting of a resistor with resistance $\mathcal{R}$ and a capacitor with capacitance $C := 1$. If the capacitor is initially charged with $V_0 := 5$, then the capacitor discharges its energy over time once the circuit is closed. In fact, the voltage over the capacitor is

[5]Here, we mean measurable with respect to the Borel σ-algebras induced by the Skorokhod metric; see Reference [8] for details.

[6]Cadlag functions are right continuous functions with left limits.

[7]The result for measurability of $\rho^\phi(X(\cdot,\omega),t)$ is mainly taken from Reference [8, Theorem 6].

[8]We use the shorthand notations $P(\beta^\phi(X,t) \in B)$, $P(\rho^\phi(X,t) \in B)$, and $P(\mathrm{RD}^\phi(X,t) \in B)$ instead of $P(\{\omega \in \Omega | \beta^\phi(X(\cdot,\omega),t) \in B\})$, $P(\{\omega \in \Omega | \rho^\phi(X(\cdot,\omega),t) \in B\})$, and $P(\{\omega \in \Omega | \mathrm{RD}^\phi(X(\cdot,\omega),t) \in B\})$, respectively.

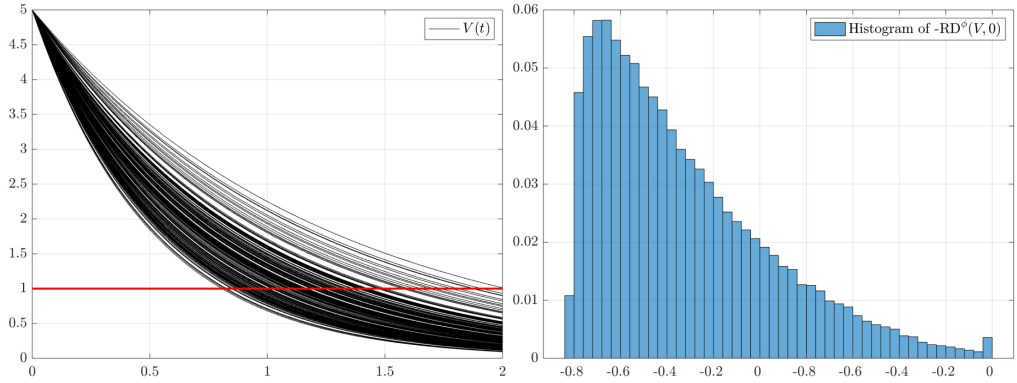


Fig. 4. Left: 200 realizations of the voltage $V(t)$ over the capacitor of an RC circuit. Right: Histogram of the negative robustness degree $-\mathrm{RD}^{\phi}(V, 0)$ of the specification $\phi := G_{[2,\infty)}(V \leq 1)$.

described by

$$V(t) = V_0 \exp(-\tau t),$$

where $\tau := 1/\mathcal{R}C$ is the time constant. Assume that the resistance is unknown and modeled as $\mathcal{R} := 0.5 + Z$, where Z is a random variable following a beta distribution with probability density function $f_Z(z) := \frac{1}{B(1.5,5)} z^{1.5-1}(1-z)^{5-1}$, where $B(1.5, 5)$ is the beta function with parameters 1.5 and 5. Consequently, the voltage V becomes a stochastic process of which we plot 200 realizations in Figure 4 (left). As a specification ϕ, we want that the voltage $V(t)$ drops below 1 after 2 s, i.e.,

$$\phi := G_{[2,\infty)}(V \leq 1).$$

In Figure 4 (right), we show the histogram of the negative robustness degree $-\mathrm{RD}^{\phi}(V, 0)$ for 100,000 realizations. To estimate the risk of the stochastic process X lacking robustness against failure of ϕ, we can now compose $-\mathrm{RD}^{\phi}(V, 0)$ with a risk measure R. For instance, the value-at-risk at level $\beta := 0.9$ is $VaR_{0.9}(-\mathrm{RD}^{\phi}(V, 0)) \approx -0.38$. Recall that $VaR_{0.9}(-\mathrm{RD}^{\phi}(V, 0))$ is the 0.1 quantile of $-\mathrm{RD}^{\phi}(V, 0)$. This means that with a probability of at least 0.9 the robustness degree is not smaller (i.e., greater) than $|VaR_{0.9}(-\mathrm{RD}^{\phi}(V, 0))| \approx 0.38$ or, in other words, that in at most 10% of the cases the robustness is smaller than 0.38. This information is useful, as it allows us to quantify how much uncertainty our system can handle, e.g., when we do not know the value of V_0 exactly.

The previous example motivates the following definition for the risk of the stochastic process X lacking robustness against failure of ϕ to which we refer as the STL robustness risk for brevity.

Definition 3 (STL Robustness Risk). Given an STL formula ϕ and a stochastic process $X : T \times \Omega \to \mathbb{R}^n$, the risk of X lacking robustness against failure of ϕ at time t is defined as

$$R(-\mathrm{RD}^{\phi}(X, t)).$$

We remark that a large positive value of $\mathrm{RD}^{\phi}(X(\cdot, \omega), t)$ for a realization $\omega \in \Omega$ indicates robust satisfaction of ϕ. Therefore, the negative robustness degree $-\mathrm{RD}^{\phi}(X, t)$ is the cost random variable that is chosen as the input for the risk measure R. This way, a large robustness degree results in a low cost. Finally, note that $R(-\mathrm{RD}^{\phi}(X', t)) \leq R(-\mathrm{RD}^{\phi}(X'', t))$ implies that the stochastic process X' is less risky than the stochastic process X'' with respect to the specification ϕ.

3.3 The Approximate STL Robustness Risk

Unfortunately, the STL robustness risk $R(-\mathrm{RD}^\phi(X,t))$ can in general not be calculated, as the robustness degree in Definition 1 is difficult to calculate. Instead, we will focus on $R(-\rho^\phi(X,t))$ using the robust semantics as an approximation of the STL robustness risk.

Definition 4 (Approximate STL Robustness Risk). Given an STL formula ϕ and a stochastic process $X : T \times \Omega \to \mathbb{R}^n$, the *approximate* risk of X lacking robustness against failure of ϕ at time t is defined as

$$R(-\rho^\phi(X,t)).$$

Fortunately, the approximate STL robustness risk $R(-\rho^\phi(X,t))$ over-approximates the STL robustness risk $R(-\mathrm{RD}^\phi(X,t))$ when R is a monotone risk measure, as shown next.

Theorem 3. *Let X be a stochastic process, ϕ be an STL specification as in Equation (1), and R be a monotone risk measure. Then it holds that*

$$R(-RD^\phi(X,t)) \leq R(-\rho^\phi(X,t)).$$

The previous result is important, as using $R(-\rho^\phi(X,t))$ instead of $R(-\mathrm{RD}^\phi(X,t))$ will not result in an optimistic risk assessment. Especially in safety-critical applications, it is desirable to be more risk-averse as opposed to being overly optimistic.

Sometimes one may be interested in scaling the robustness degree to associate a monetary cost with $\mathrm{RD}^\phi(X,t)$ to reflect the severity of events with low robustness. Let us for this purpose consider an increasing cost function $C : \mathbb{R} \to \mathbb{R}$.

Corollary 1. *Let X be a stochastic process, ϕ be an STL specification as in Equation (1), R be a monotone risk measure, and C be an increasing cost function. Then it holds that*

$$R(C(-RD^\phi(X,t))) \leq R(C(-\rho^\phi(X,t))).$$

4 DATA-DRIVEN ESTIMATION OF THE APPROXIMATE STL ROBUSTNESS RISK

In this section, we show how the approximate STL robustness risk $R(-\rho^\phi(X,t))$ can be estimated from data. We assume that we have observed N independent realizations of the stochastic process X, i.e., we know N realizations $X(\cdot,\omega^1),\dots,X(\cdot,\omega^N)$ where $\omega^1,\dots,\omega^N \in \Omega$ are drawn independently and according to the probability measure P. A practical example would be a simulator from which we can unroll trajectories $X(\cdot,\omega^i)$. For brevity, we denote $X(\cdot,\omega^1),\dots,X(\cdot,\omega^N)$ by $X^1,\dots,X^N$. In this way, one can think of $X^1,\dots,X^N$ as N independent copies of X. We emphasize that we do not need knowledge of the distribution of X. Our goal is to derive upper bounds of $R(-\rho^\phi(X,t))$ that hold with high probability. Let us, for convenience, first define the random variable

$$Z := -\rho^\phi(X,t).$$

For further convenience, let $Z^i := -\rho^\phi(X^i,t)$ and let us also define the tuple

$$\mathcal{Z} := (Z^1,\dots,Z^N).$$

We consider the value-at-risk $VaR_\beta(Z)$, the conditional value-at-risk $CVaR_\beta(Z)$, and the mean $E(Z)$. Particularly, we derive upper bounds $\overline{VaR}_\beta(\mathcal{Z},\delta)$, $\overline{CVaR}_\beta(\mathcal{Z},\delta)$, and $\overline{E}(\mathcal{Z},\delta)$ that hold with a probability of at least $1-\delta$. By Theorem 3 and Propositions 1, 2, and 3 (presented in the remainder), we then have computational algorithms to find tight upper bounds for the approximate STL

robustness risk and hence for the STL robustness risk, and it holds that with a probability of $1-\delta$

$$VaR_\beta(-\mathrm{RD}^\phi(X,t)) \le VaR_\beta(Z) \le \overline{VaR}_\beta(\mathcal{Z},\delta),$$
$$CVaR_\beta(-\mathrm{RD}^\phi(X,t)) \le CVaR_\beta(Z) \le \overline{CVaR}_\beta(\mathcal{Z},\delta),$$
$$E(-\mathrm{RD}^\phi(X,t)) \le E(Z) \le \overline{E}(\mathcal{Z},\delta).$$

4.1 Value-at-Risk (VaR)

For a risk level of $\beta \in (0,1)$, recall that the VaR of Z is given by

$$VaR_\beta(Z) := \inf\{\alpha \in \mathbb{R} | F_Z(\alpha) \ge \beta\},$$

where $F_Z(\alpha)$ denotes the CDF of Z. To estimate $F_Z(\alpha)$, we define the empirical CDF as

$$\widehat{F}(\alpha,\mathcal{Z}) := \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(Z^i \le \alpha),$$

where $\mathbb{I}$ denotes the indicator function defined as

$$\mathbb{I}(Z^i \le \alpha) := \begin{cases} 1 & \text{if } Z^i \le \alpha \\ 0 & \text{otherwise.} \end{cases}$$

Let now $\delta \in (0,1)$ be a probability threshold. Inspired by Szorenyi et al. [78], we calculate an upper bound of $VaR_\beta(Z)$ as

$$\overline{VaR}_\beta(\mathcal{Z},\delta) := \inf\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha,\mathcal{Z}) - \sqrt{\frac{\ln(2/\delta)}{2N}} \ge \beta\right\}$$

and a lower bound as

$$\underline{VaR}_\beta(\mathcal{Z},\delta) := \inf\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha,\mathcal{Z}) + \sqrt{\frac{\ln(2/\delta)}{2N}} \ge \beta\right\},$$

where we recall that $\inf \emptyset = \infty$ for $\emptyset$ being the empty set due to the extended definition of the infimum operator. We next show that $\overline{VaR}_\beta(\mathcal{Z},\delta)$ and $\underline{VaR}_\beta(\mathcal{Z},\delta)$ are upper and lower bounds of $VaR_\beta(Z)$, respectively, with a probability of at least $1-\delta$.

Proposition 1. *Assume that F_Z is continuous and let $\delta \in (0,1)$ be a probability threshold and $\beta \in (0,1)$ be a risk level. Let $\overline{VaR}_\beta(\mathcal{Z},\delta)$ and $\underline{VaR}_\beta(\mathcal{Z},\delta)$ be based on the data $\mathcal{Z}$. With a probability of at least $1-\delta$, it holds that*

$$\underline{VaR}_\beta(\mathcal{Z},\delta) \le VaR_\beta(Z) \le \overline{VaR}_\beta(\mathcal{Z},\delta).$$

We remark that Theorem 1 assumes that F_Z is continuous. If F_Z is not continuous, then one can derive upper and lower bounds by using order statistics following Nikolakakis et al. [57, Lemma 3].

4.2 Conditional Value-at-Risk (CVaR)

For a risk level of $\beta \in (0,1)$, recall that the CVaR of Z is given by

$$CVaR_\beta(Z) := \inf_{\alpha \in \mathbb{R}} (\alpha + (1-\beta)^{-1} E([Z-\alpha]^+)),$$

where $[Z-\alpha]^+ := \max(Z-\alpha, 0)$. For estimating $CVaR_\beta(Z)$ from data $\mathcal{Z}$, we focus here on the case where the random variable $\rho^\phi(X,t)$ (and hence Z) has bounded support for fixed t. In particular, we assume that $P(\rho^\phi(X,t) \in [a,b]) = 1$. Note that $\rho^\phi(X,t)$ has bounded support when the function ρ^ϕ is bounded, which can be achieved either by construction of ϕ or by clipping off ρ^ϕ outside the interval $[a,b]$ for some *a priori* chosen constants a and b, i.e., values outside this interval are

clipped to the end points a and b of the interval. We remark that clipping off ρ^ϕ is not restrictive in most practical applications, i.e., realizations of $\rho^\phi(X,t)$ that are larger than a sufficiently large value of $b > 0$ indicate robust satisfaction of ϕ and will not affect the risk associated with Z, while realizations of $\rho^\phi(X,t)$ smaller than $a < 0$ violate the specification ϕ already.[9] We will provide illustrative examples in our simulations in Section 6. This boundedness assumption enables us now to directly leverage results from Wang and Gao [83] to estimate upper and lower bounds of $CVaR_\beta(Z)$. Let us first define the empirical estimate of $CVaR_\beta(Z)$ as

$$\widehat{CVaR}_\beta(\mathcal{Z}) := \inf_{\alpha\in\mathbb{R}} \left(\alpha + (N(1-\beta))^{-1} \sum_{i=1}^{N} [Z^i - \alpha]^+ \right).$$

Based on Wang and Gao [83, Theorem 3.1], we can now calculate an upper bound of $CVaR_\beta(Z)$ as

$$\overline{CVaR}_\beta(\mathcal{Z},\delta) := \widehat{CVaR}_\beta(\mathcal{Z}) + \sqrt{\frac{5\ln(3/\delta)}{N(1-\beta)}}(b-a)$$

and a lower bound as

$$\underline{CVaR}_\beta(\mathcal{Z},\delta) := \widehat{CVaR}_\beta(\mathcal{Z}) - \sqrt{\frac{11\ln(3/\delta)}{N(1-\beta)}}(b-a).$$

We would like to highlight that the upper and lower bounds $\overline{CVaR}_\beta(\mathcal{Z},\delta)$ and $\underline{CVaR}_\beta(\mathcal{Z},\delta)$, respectively, become less accurate with larger values of $(b-a)$, which we can account for by increasing the number of observed trajectories N. The following proposition follows immediately from Wang and Gao [83, Theorem 3.1]:

Proposition 2. *Let $\delta \in (0,1)$ be a probability threshold and $\beta \in (0,1)$ be a risk level. Assume that $P(\rho^\phi(X,t) \in [a,b]) = 1$. Let $\overline{CVaR}_\beta(\mathcal{Z},\delta)$ and $\underline{CVaR}_\beta(\mathcal{Z},\delta)$ be based on the data $\mathcal{Z}$. With a probability of at least $1-\delta$, it holds that*

$$\underline{CVaR}_\beta(\mathcal{Z},\delta) \le CVaR_\beta(Z) \le \overline{CVaR}_\beta(\mathcal{Z},\delta).$$

Remark 5. The case where Z has unbounded support, but where Z is sub-Gaussian or sub-exponential has been considered in Bhat and L. A. [11], Brown [13], Kolla et al. [39], Mhammedi et al. [54], Thomas and Learned-Miller [79].

4.3 Mean

Define the empirical estimate of the mean $E(Z)$ as

$$\widehat{E}(\mathcal{Z}) := \frac{1}{N}\sum_{i=1}^{N} Z^i.$$

By the law of large numbers, $\widehat{E}(\mathcal{Z})$ converges to $E(Z)$ with probability one as N goes to infinity. For finite N and when again Z has bounded support, i.e., $P(Z \in [a,b]) = 1$, we can apply Hoeffding's inequality and calculate an upper $\overline{E}(\mathcal{Z},\delta)$ of the mean $E(Z)$ as

$$\overline{E}(\mathcal{Z},\delta) := \widehat{E}(\mathcal{Z}) + \sqrt{\frac{\ln(2/\delta)}{2N}}(b-a)$$

[9] In practice, it hence makes sense to select a negative value for a and to select b based on physical intuition that we may have—either from trajectories that we may have already observed or from domain knowledge, e.g., for a lane-keeping controller in autonomous driving, the value of $b = 1$ meter is a good robustness.

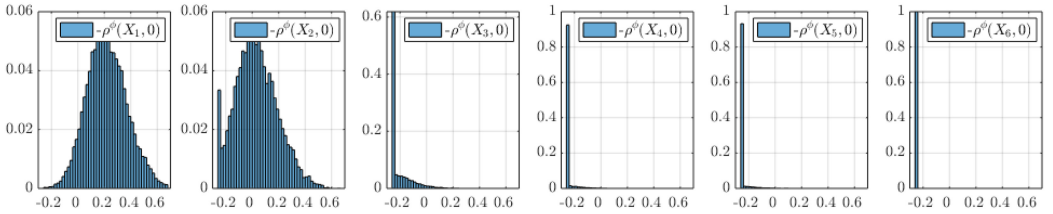


Fig. 5. Histogram of $-\text{RD}^{\phi}(X_j, 0)$ of the specification ϕ in (2) for robot trajectories $j \in \{1, \ldots, 6\}$.

and a lower bound as

$$\underline{E}(\mathcal{Z}, \delta) := \widehat{E}(\mathcal{Z}) - \sqrt{\frac{\ln(2/\delta)}{2N}}(b - a).$$

Similarly to the observation that we made for CVaR, note that the upper and lower bounds $\overline{E}(\mathcal{Z}, \delta)$ and $\underline{E}(\mathcal{Z}, \delta)$, respectively, become less accurate with increasing values of $(b-a)$ and more accurate with increasing N. We next show that we indeed obtain valid upper and lower bounds.

Proposition 3. *Let $\delta \in (0, 1)$ be a probability threshold. Assume that $P(\rho^{\phi}(X, t) \in [a, b]) = 1$. Let $\overline{E}(\mathcal{Z}, \delta)$ and $\underline{E}(\mathcal{Z}, \delta)$ be based on the data $\mathcal{Z}$. With a probability of at least $1 - \delta$, it holds that*

$$\underline{E}(\mathcal{Z}, \delta) \leq E(Z) \leq \overline{E}(\mathcal{Z}, \delta).$$

Example 1 (continued). We now modify Example 1 by considering that the regions C and D are not exactly known. Let c and d in Equations (3) and (4), respectively, be Gaussian random vectors as

$$c \sim \mathcal{N}\left(\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}0.2 & 0\\0 & 0.2\end{bmatrix}\right), \tag{6}$$

$$d \sim \mathcal{N}\left(\begin{bmatrix}6\\4\end{bmatrix}, \begin{bmatrix}0.2 & 0\\0 & 0.2\end{bmatrix}\right). \tag{7}$$

Consequently, the signals x_1-x_6 become stochastic processes denoted by X_1-X_6. Let now X_j^i denote the ith observed realization of X_j where $j \in \{1, \ldots, 6\}$. Our first goal is to estimate $VaR_{\beta}(Z)$ to compare the risk between the six robot trajectories r_1-r_6. We set $\delta := 0.01$ and $N := 15,000$.[10] The histograms of $-\rho^{\phi}(X_j)$ for each trajectory are shown in Figure 5. For different risk levels β, the resulting upper and lower bounds for the value-at-risk are shown in the next table.

j \ R	$\overline{VaR}_{0.9}$	$\overline{VaR}_{0.925}$	$\overline{VaR}_{0.95}$	$\overline{VaR}_{0.975}$	$\underline{VaR}_{0.9}$	$\underline{VaR}_{0.925}$	$\underline{VaR}_{0.95}$	$\underline{VaR}_{0.975}$
1	0.434	0.467	0.508	0.577	0.407	0.432	0.465	0.505
2	0.261	0.295	0.335	0.424	0.232	0.259	0.292	0.332
3	−0.075	−0.044	0.001	0.086	−0.1	−0.077	−0.046	−0.003
4	−0.25	−0.222	−0.177	−0.086	−0.25	−0.25	−0.225	−0.182
5	−0.249	−0.228	−0.18	−0.084	−0.249	−0.249	−0.23	−0.185
6	−0.249	−0.249	−0.249	−0.249	−0.249	−0.249	−0.249	−0.249

Across all β, it can be observed that the estimate $\overline{VaR}_{\beta}$ of VaR_{β} is relatively tight, as the difference $|\overline{VaR}_{\beta} - \underline{VaR}_{\beta}|$ between upper and lower bounds is small. The table indicates that trajectories r_1

[10] We can select smaller N at the cost of slightly more conservative estimates.

and r_2 are not favorable and are not robust. Recall that smaller risk values are favorable, as only negative values indicate actual robustness. Trajectory r_3 is better compared to trajectories r_1 and r_2, but worse than r_4-r_6 in terms of the approximate STL robustness risk of ϕ. For trajectories r_4-r_6, note that a $\beta = 0.9$ provides the information that the trajectories have roughly the same approximate STL robustness risk. However, once the risk level β is increased to 0.925, 0.95, and 0.975, it becomes clear that r_6 is preferable over r_4 and r_5. This matches with what one would expect by closer inspection of Figures 2 and 5.

We next estimate $CVaR_\beta(Z)$ and therefore restrict ρ^ϕ to lie within $[-0.5, 0.25]$ simply by clipping values that exceed this bound. This choice is motivated by our previous discussion in Section 4.2 and as ρ^ϕ is upper bounded by 0.25; see histograms in Figure 5. For different risk levels β, the resulting upper and lower bounds for the conditional value-at-risk are shown next.

j \ R	$\overline{CVaR}_{0.9}$	$\overline{CVaR}_{0.925}$	$\overline{CVaR}_{0.95}$	$\overline{CVaR}_{0.975}$	$\underline{CVaR}_{0.9}$	$\underline{CVaR}_{0.925}$	$\underline{CVaR}_{0.95}$	$\underline{CVaR}_{0.975}$
1	0.577	0.607	0.645	0.707	0.32	0.31	0.282	0.193
2	0.432	0.471	0.527	0.637	0.175	0.174	0.164	0.12
3	0.1	0.136	0.193	0.301	−0.16	−0.161	−0.17	−0.213
4	−0.078	−0.04	0.019	0.13	−0.335	−0.336	−0.344	−0.384
5	−0.08	−0.042	0.019	0.134	−0.337	−0.338	−0.344	−0.38
6	−0.146	−0.13	−0.103	−0.042	−0.403	−0.426	−0.466	−0.556

In general, the same observations regarding the ranking of $r_1 - r_6$ can be made based on the conditional value-at-risk. However, the risk levels are in general much higher, as $CVaR_\beta$ is more risk-sensitive than VaR_β. An important observation is that the estimates $\overline{CVaR}_\beta$ of $CVaR_\beta$ are not as tight as before for VaR_β, as the difference $|\overline{CVaR}_\beta - \underline{CVaR}_\beta|$ is larger, particularly for larger β due to the division by $1-\beta$ in the estimates of $\overline{CVaR}_\beta$ and $\underline{CVaR}_\beta$. For completeness, we also report the estimated mean of Z.

j \ R	$\overline{E}$	$\underline{E}$
1	0.227	0.207
2	0.043	0.023
3	−0.194	−0.214
4	−0.233	−0.253
5	−0.233	−0.253
6	−0.24	−0.26

5 EXACT COMPUTATION OF THE APPROXIMATE STL ROBUSTNESS RISK

In the previous section, we estimated the approximate STL robustness risk using observed realizations $X^1, \dots, X^N$ of the stochastic process X. In this section, we instead assume to know the distribution of X. There are two main challenges in computing the approximate STL robustness risk $R(-\rho^\phi(X,t))$ from the distribution of X. First, note that exact computation of $R(-\rho^\phi(X,t))$ requires knowledge of the CDF of $\rho^\phi(X,t)$. However, the CDF of $\rho^\phi(X,t)$ is in general not known and often hard to obtain analytically. Second, calculating $R(-\rho^\phi(X,t))$ may often involve solving high-dimensional integrals for which in most of the cases no closed-form expressions exists. For these reasons, we assume in this section that the STL formula ϕ is bounded and that $X : T \times \Omega \to \mathcal{X}$ is a discrete-time stochastic process, i.e., $T := \mathbb{Z}$, with a finite state space $\mathcal{X} \subseteq \mathbb{R}^n$ (i.e., the set $\mathcal{X}$ consists of a finite set of elements).

Recall that the time intervals I contained in a bounded STL formula ϕ are compact. The satisfaction of such an STL formula can hence be decided by finite signals. A bounded STL formula ϕ has a future formula length $L_f^\phi \in \mathbb{Z}$ and a past formula length $L_p^\phi \in \mathbb{Z}$. The future formula length L_f^ϕ can be calculated, similarly to Sadraddini and Belta [68], as

$$\begin{aligned} L_f^\top = L_f^\mu &:= 0 \\ L_f^{\neg\phi} &:= L_f^\phi \\ L_f^{\phi' \wedge \phi''} &:= \max(L_f^{\phi'}, L_f^{\phi''}) \\ L_f^{\phi' U_I \phi''} &:= \max\{I \cap \mathbb{Z}\} + \max(L_f^{\phi'}, L_f^{\phi''}) \\ L_f^{\phi' \underline{U}_I \phi''} &:= \max(L_f^{\phi'}, L_f^{\phi''}). \end{aligned}$$

The past formula length L_p^ϕ can be calculated similarly as

$$\begin{aligned} L_p^\top = L_p^\mu &:= 0 \\ L_p^{\neg\phi} &:= L_p^\phi \\ L_p^{\phi' \wedge \phi''} &:= \max(L_p^{\phi'}, L_p^{\phi''}) \\ L_p^{\phi' U_I \phi''} &:= \max(L_p^{\phi'}, L_p^{\phi''}) \\ L_p^{\phi' \underline{U}_I \phi''} &:= \max\{I \cap \mathbb{Z}\} + \max(L_p^{\phi'}, L_p^{\phi''}). \end{aligned}$$

A finite signal of length $L_f^\phi + L_p^\phi$ is now sufficient to determine if ϕ is satisfied at time t. In particular, information from the time interval $T_L := \{t - L_p^\phi, \dots, t, \dots, t + L_f^\phi\}$ is sufficient to determine if ϕ is satisfied at time t. Now, let $X : \Omega \times T_L \to \mathcal{X}$ be the discrete-time stochastic process under consideration where the state space $\mathcal{X} \subseteq \mathbb{R}^n$ is a finite set. Note that we can always obtain such a finite set $\mathcal{X}$ from a continuous state space by discretization. Let the probability mass function (PMF) $f_X(x)$ of X be given. The next result is stated without proof, as it follows immediately from the fact that T_L and $\mathcal{X}$, and consequently the set of signals $\mathfrak{F}(T_L, \mathcal{X})$ are finite sets.

Proposition 4. *Let ϕ be a bounded STL formula with future and past formula lengths L_f^ϕ and L_p^ϕ, respectively. Let $X : \Omega \times T_L \to \mathcal{X}$ be a discrete-time stochastic process with a finite state space $\mathcal{X}$. For $t \in \mathbb{R}$, we can calculate the PMF $f_Z(z)$ and the CDF $F_Z(z)$ of Z as*

$$\begin{aligned} f_Z(z) &= \sum_{x \in \mathfrak{F}(T_L, \mathcal{X})} \mathbb{I}(-\rho^\phi(x, t) = z) f_X(x), \\ F_Z(z) &= \sum_{x \in \mathfrak{F}(T_L, \mathcal{X})} \mathbb{I}(-\rho^\phi(x, t) \leq z) f_X(x). \end{aligned}$$

Note that $F_Z(z) = \sum_{z' \leq z} f_Z(z')$ holds as required. Having obtained the PMF $f_Z(z)$ and the CDF $F_Z(z)$ of Z, it is now straightforward to calculate $R(Z)$ for various risk measures R. Note, in particular, that Z is a discrete random variable so $f_Z(z)$ is discrete and $F_Z(z)$ is piecewise-continuous, hence simplifying the calculation of $R(Z)$, as no high-dimensional integrals need to be solved.

Example 1 (continued). Recall that c and d were assumed to be Gaussian distributed according to Equations (6) and (7), respectively. We first discretize the distributions of c and d; see Appendix G for details. From the PMFs f_c and f_d, we can now calculate the PMF $f_X(x)$ for any $x \in \mathfrak{F}(T_L, \mathbb{R}^6) \times \mathcal{C} \times \mathcal{D}$ where $\mathcal{C}$ and $\mathcal{D}$ are the discretized domains of c and d. We can hence

calculate $f_Z(z)$ according to Proposition 4. From this, the value at risk $VaR_\beta(Z)$ can be calculated, which is reported in the next table.

i \ R	$VaR_{0.9}$	$VaR_{0.925}$	$VaR_{0.95}$	$VaR_{0.975}$
1	0.403	0.429	0.461	0.509
2	0.225	0.255	0.29	0.348
3	−0.102	−0.067	−0.049	0.003
4	−0.249	−0.249	−0.222	−0.162
5	−0.25	−0.25	−0.222	−0.157
6	−0.249	−0.249	−0.249	−0.249

It can be seen that the STL robustness risks reported above closely resemble the sampling-based estimates $\overline{VaR}_\beta$ of VaR_β from Section 4.

6 SIMULATIONS: AUTONOMOUS DRIVING IN CARLA

We consider the verification of neural network-based lane-keeping controllers for lateral control in the autonomous driving simulator CARLA [19]; see Figure 1 (left). Lane-keeping in CARLA is achieved by tracking a set of predefined waypoints. For longitudinal control, a built-in PID controller is used to stabilize the car at 20 km/h. We particularly trained four different neural network controllers as detailed below. Our overall goal is to estimate and compare the risks of these four controllers for five different specifications during a double left turn; see Figure 1 (middle).

For the verification and comparison of these controllers, we are particularly interested in the cross-track error, which is a measure of the closest distance from the car to the path defined by the set of waypoints, as illustrated in Figure 1 (right). Formally, let wp_1 be the waypoint that is closest to the car and let wp_2 be the waypoint proceeding wp_1. Then the cross-track error is defined as $c_e := \|w\| \sin(\theta_w)$, where w is the vector pointing from wp_1 to the car and θ_w is the angle between w and the vector pointing from wp_1 to wp_2. We are also interested in the orientation error $\theta_e := \theta_t - \theta$ between the orientation of the reference path θ_t and the orientation of the car θ.

The state $x := (c_e, \theta_e, v, d, \dot{\theta}_t)$ of the car consists of the cross-track error c_e, the orientation error θ_e, the velocity v of the car, the internal state d of the longitudinal PID controller, and the rate $\dot{\theta}_t$ at which the orientation of the reference path changes. The control input for which we aim to learn and verify a lane-keeping controller is the steering angle u.

6.1 Training Neural Network Lane-keeping Controllers

We have trained four different neural network controllers. Two of these four controllers were obtained by using supervised imitation learning (IL) [65], while the other two controllers were obtained by learning control barrier functions (CBFs) from expert demonstrations [49].

To obtain two imitation learning controllers, we used a CARLA built-in PID controller u^* as an expert controller to collect expert trajectories, which are sequences of state and control input pairs. The first IL controller, denoted as IL_{full}, is trained using the full state x as an input to the neural network, while the control input u is the output. The second IL controller, denoted as $\text{IL}_{\text{partial}}$, is trained by only using partial state knowledge. In particular, only the cross-track error c_e, the orientation error θ_e, and the rate $\dot{\theta}_t$ at which the orientation of the path changes are used here as an input to the neural network. We used one-layer neural networks with 20 neurons per layer and ReLU activation functions and trained with the mean squared error as the loss function.

Remark 6. For simplicity, we did not attempt to address the distribution shift between the expert controller and the trained controller, e.g., by using DAGGER [66]. We remark that our primary goal lies in the verification and comparison of risk between controllers.

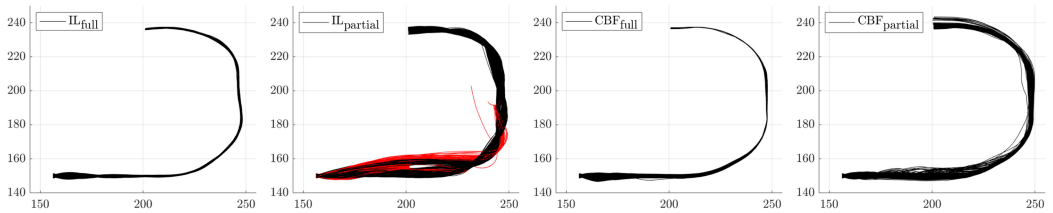


Fig. 6. Shown are 600 trajectories for each of the four controllers during the double left turn. Trajectories marked in red led to a collision with an obstacle.

To obtain the CBF-based controllers, we again used the expert controller u^* to get expert trajectories from which we learned robust control barrier functions following Lindemann et al. [49]. The first controller, denoted as CBF_{full}, uses again full state knowledge of x. The second controller, denoted as $\text{CBF}_{\text{partial}}$, estimates the cross-track error c_e from RGB dashboard camera images while assuming knowledge of the remaining states; see Lindemann et al. [49] for details. Both neural network controllers consist of two layers with 32 and 16 neurons and tanh activation functions.

6.2 Risk Verification and Comparison

For the risk verification and comparison of these four controllers, we tested each of them on the training course; see Figure 1 (middle). We uniformly sampled the initial position of the car in a range of $c_e \in [-1, 1]$ m and $\theta_e \in [-0.4, 0.4]$ rad and added normally distributed noise in a range of $[-0.1, 0.1]$ rad to the control input to simulate actuation noise so the car becomes a stochastic process X. We collected $N := 1{,}000$ trajectories for each controller, of which 600 are shown in Figure 6. From a visual inspection, we can already see that the controllers that use full state knowledge (IL_{full}, CBF_{full}) outperform the controllers that only use partial state knowledge ($\text{IL}_{\text{partial}}$, $\text{CBF}_{\text{partial}}$). Videos of each controller from five different initial conditions are provided under https://tinyurl.com/48xjf545.

To obtain a more formal assessment, we next estimate the risk of each controller with respect to: (1) the cross-track error over the whole trajectory, during steady state, and during the transient phase, (2) the responsiveness of the controller, and (3) the orientation error.

6.2.1 Cross-track Error. The specification that we look at here is that the cross-track error c_e should always be within the interval $[-2.25, 2.25]$, where 2.25 is a threshold that we selected based on the cross-track error induced by the expert controller u^*. In STL language, we have

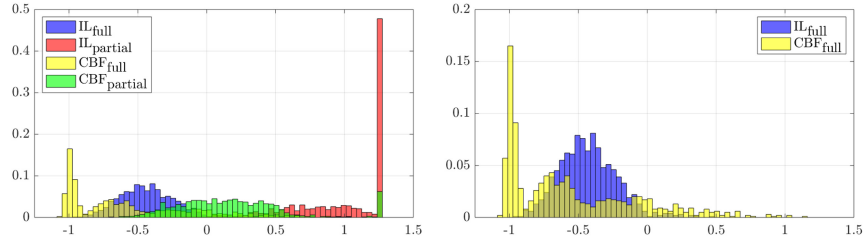
$$\phi_1 := G_{[0,\infty)}(|c_e| \leq 2.25).$$

We show the histograms of $\rho^{\phi_1}(X, 0)$ for each controller in Figure 7(a) (left).[11] We are particularly interested in the controllers IL_{full} and CBF_{full} and show their histograms isolated in Figure 7(a) (right) for better readability. Selecting $\delta := 0.01$, the estimates of $VaR_{0.85}$, $VaR_{0.95}$, $CVaR_{0.85}$, and E are reported in the table below. In the last column, we have additionally reported the empirical probability that the specification ϕ_1 is satisfied, which we calculate as

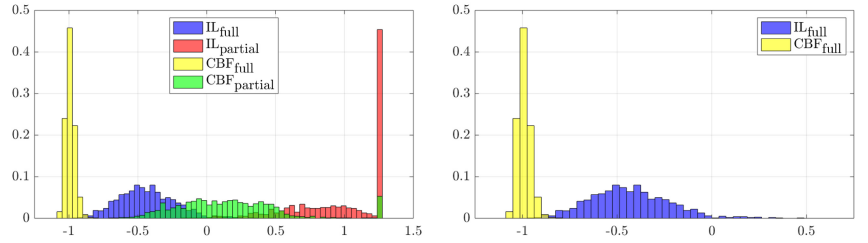
$$\#_{\phi_1} := \frac{\sum_{i=1}^{N} \mathbb{I}(\beta^{\phi_1}(X^i, 0) = \top)}{N}.$$

For each risk measure, we highlight the controller with the lowest risk in green.

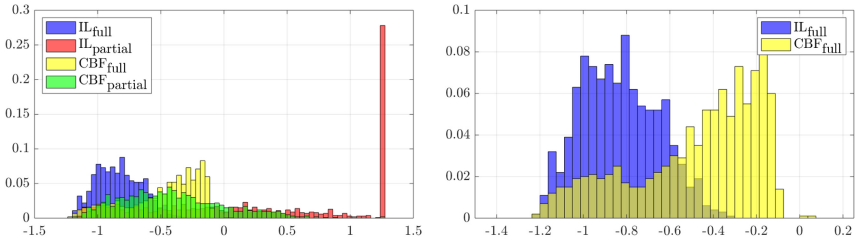
[11] We restrict ρ^{ϕ_1} to lie within the interval $[-1.25, 2.25]$, i.e., in this case, we clip the values of $\rho^{\phi_1}(X, 0) = \inf_{t \in \mathbb{Z}} 2.25 - |c_e(t)|$ to -1.25 if $\rho^{\phi_1}(X, 0) < -1.25$. In the remainder, we clip ρ^{ϕ_2}-ρ^{ϕ_5} in the same way for the specifications ϕ_2-ϕ_5.



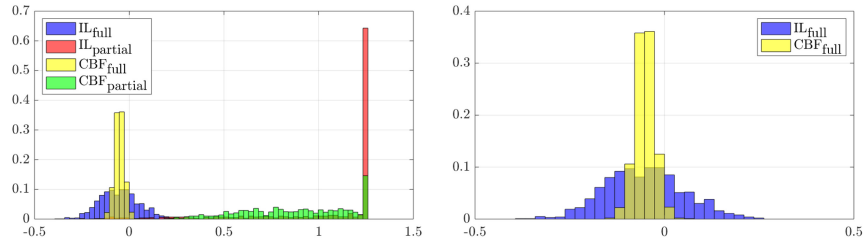
(a) $\phi_1 := G_{[0,\infty)}(|c_e| \leq 2.25)$



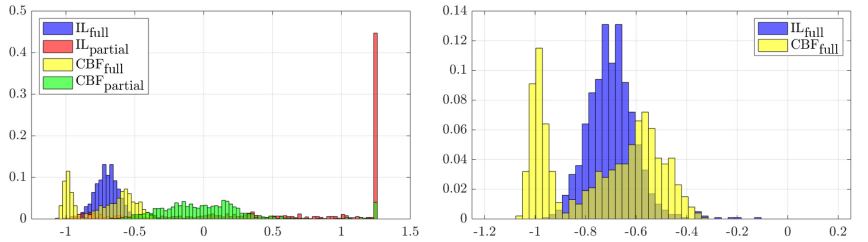
(b) $\phi_2 := G_{[10,\infty)}(|c_e| \leq 2.25)$



(c) $\phi_3 := F_{[0,5]}G_{[0,5]}(|c_e| \leq 1.25)$



(d) $\phi_4 := G_{[10,\infty)}\big((|c_e| \geq 1.25) \implies F_{[0,5]}G_{[0,5]}(|c_e| \leq 1.25)\big)$



(e) $\phi_5 := G_{[0,\infty)}\big((c_e \geq 1.25) \implies F_{[0,2]}G_{[0,1]}(\theta_e \leq 0) \wedge (c_e \leq -1.25) \implies F_{[0,2]}G_{[0,1]}(\theta_e \geq 0)\big)$

Fig. 7. Histograms of $-\rho^{\phi_i}(X, 0)$ for each controller for the specifications ϕ_1-ϕ_5.

u \ R	$\overline{VaR}_{0.85}$	$\overline{VaR}_{0.95}$	$\overline{CVaR}_{0.85}$	$\overline{E}$	$\underline{VaR}_{0.85}$	$\underline{VaR}_{0.95}$	$\underline{CVaR}_{0.85}$	$\underline{E}$	$\#_{\phi_1}$
IL_{full}	−0.168	0.462	1.436	−0.248	−0.258	−0.168	−2.354	−0.61	0.975
$\text{IL}_{\text{partial}}$	1.25	1.25	2.776	1.166	1.25	1.25	−1.014	0.806	0.005
CBF_{full}	0.135	1.125	1.818	−0.375	−0.125	0.105	−1.972	−0.736	0.863
$\text{CBF}_{\text{partial}}$	0.58	1.25	2.42	0.357	0.44	0.58	−1.37	−0.003	0.364

Based on these risk estimates, we make the following observations:

- As expected from the visual inspection of Figure 6, the controllers $\text{IL}_{\text{partial}}$ and $\text{CBF}_{\text{partial}}$ perform poorly. Among these two, $\text{CBF}_{\text{partial}}$ performs slightly better in terms of risk than $\text{IL}_{\text{partial}}$.
- The controllers IL_{full} and CBF_{full} perform better. The risk of CBF_{full} in terms of the expected value $\overline{E}$ is smaller than the risk of IL_{full}. Interestingly, the risk of IL_{full} in terms of the $\overline{VaR}_{0.85}$, $\overline{VaR}_{0.95}$, and $\overline{CVaR}_{0.85}$ is smaller than the risk of CBF_{full}. This is due to the long tail induced by CBF_{full}; see Figure 7(a) (right). We hence argue that IL_{full} is the better choice with respect to ϕ_1.
- The estimate $\overline{CVaR}_{0.85}$ of $CVaR_{0.85}$ is not tight and very conservative. The difference $|\overline{CVaR}_{0.85} - \underline{CVaR}_{0.85}|$ between the upper and lower bounds is large. To make this bound tighter, more data N is needed. We neglect the conditional value-at-risk in the remainder.
- In this case, it can be observed that a low empirical satisfaction probability $\#_{\phi_1}$ correlates with a high risk. We remark that this is not always the case, as risk considers characteristics of the right tail of the distribution $-\rho^{\phi_1}(X, 0)$, while satisfaction probabilities focus on the left tail of this distribution. This can be observed when we present the results for specification ϕ_5.

We formulate the hypothesis that the long tail of CBF_{full} that makes CBF_{full} more risky than IL_{full} is induced by the transient behavior. We analyze this hypothesis in detail in the remainder looking at the specifications ϕ_2 (steady-state) and ϕ_3 (transient phase).

6.2.2 Steady-state. In the previous section, we concluded that IL_{full} is the best controller for the specification ϕ_1, i.e., when considering the cross-track error c_e over the whole trajectory. We now study the steady-state behavior of each controller in terms of c_e and reveal that CBF_{full} is the least risky controller when only looking at the steady-state. Therefore, we check if the cross-track error c_e is always within the interval $[-2.25, 2.25]$ after 10 s by the specification

$$\phi_2 := G_{[10,\infty)}(|c_e| \leq 2.25).$$

We show the histograms of $\rho^{\phi_2}(X, 0)$ for each controller Figure 7(b) and report the risk estimates below.

u \ R	$\overline{VaR}_{0.85}$	$\overline{VaR}_{0.9}$	$\overline{VaR}_{0.95}$	$\overline{E}$	$\#_{\phi_2}$
IL_{full}	−0.168	−0.078	0.462	−0.254	0.975
$\text{IL}_{\text{partial}}$	1.25	1.25	1.25	1.153	0.005
CBF_{full}	−0.944	−0.924	−0.794	−0.81	1
$\text{CBF}_{\text{partial}}$	0.56	1.25	1.25	0.341	0.377

Based on these risk estimates, we make the following observations:

- We see that our stated hypothesis is true and observe that CBF_{full} now has the least risky behavior for all risk measures with respect to ϕ_2, i.e., during steady state.

- For CBF_{full}, we have $\overline{VaR}_{0.95}(-\rho^{\phi_2}(X,0)) = -0.794$. Consequently, for at most 5% of the realizations the robustness is less than 0.794.

6.2.3 Transient Phase. Complementary to the previous analysis, we now look at the transient behavior of the cross-track error c_e of each controller by imposing the specification

$$\phi_3 := F_{[0,5]}G_{[0,5]}(|c_e| \leq 1.25).$$

In other words, the specification ϕ_3 requires that eventually within the first 5 s the absolute value of the cross-track error falls below the threshold 1.25 for at least 5 s. We show the histogram of each controller in Figure 7(c) and report the corresponding risk estimates next.

u \ R	$\overline{VaR}_{0.85}$	$\overline{VaR}_{0.9}$	$\overline{VaR}_{0.95}$	$\overline{E}$	$\#_{\phi_3}$
IL_{full}	−0.584	−0.524	−0.324	−0.652	1
$\text{IL}_{\text{partial}}$	1.25	1.25	1.25	0.493	0.42
CBF_{full}	−0.157	−0.137	0.063	−0.297	0.998
$\text{CBF}_{\text{partial}}$	0.2	0.38	1.25	−0.221	0.83

For ϕ_3, we see a similar result as for ϕ_1 in the sense that IL_{full} is the least risky controller, but now clearly indicating that IL_{full} is the less risky controller across all risk measures. It is also worth pointing out that CBF_{full} and $\text{CBF}_{\text{partial}}$ have almost the same expected value, while $\overline{VaR}_{0.85}$, $\overline{VaR}_{0.9}$, and $\overline{VaR}_{0.95}$ indicate that CBF_{full} is less risky.

Summarizing the observations from ϕ_1, ϕ_2, and ϕ_3, IL_{full} is the least risky controller during the transient phase and CBF_{full} is the least risky controller during steady-state.

6.2.4 Responsiveness. So far, we focused on the cross-track error during steady-state and transient phase. We now analyze the responsiveness of the controllers when the cross-track error gets too large. We particularly analyze how responsive the controllers are in such situations and how quickly they can decrease the error again to an acceptable level. Let us therefore look at the specification

$$\phi_4 := G_{[10,\infty)}\Big((|c_e| \geq 1.25) \implies F_{[0,5]}G_{[0,5]}(|c_e| \leq 1.25)\Big).$$

In other words, whenever the cross-track error c_e leaves the interval $[-1.25, 1.25]$ after the transient phase has died out (approximately after 10 s), it should hold that within the next 5 s the cross-track error is again within the interval $[-1.25, 1.25]$ for at least 5 s. We show the histogram of each controller in Figure 7(d) and report the corresponding risk estimates below.

u \ R	$\overline{VaR}_{0.85}$	$\overline{VaR}_{0.9}$	$\overline{VaR}_{0.95}$	$\overline{E}$	$\#_{\phi_4}$
IL_{full}	0.088	0.128	0.248	0.127	0.703
$\text{IL}_{\text{partial}}$	1.25	1.25	1.25	1.226	0.026
CBF_{full}	−0.0152	−0.005	0.055	0.129	0.974
$\text{CBF}_{\text{partial}}$	1.25	1.25	1.25	1.054	0

The results are interesting in the sense that the risk of IL_{full} and CBF_{full} in terms of the expected value are almost identical, even slightly favoring IL_{full}, while the risk of CBF_{full} in terms of $\overline{VaR}_{0.85}$, $\overline{VaR}_{0.9}$, and $\overline{VaR}_{0.95}$ is much smaller.

6.2.5 Orientation Error. Let us now focus on the orientation error θ_e. In general, an orientation error is expected when either the orientation θ_t of the reference path changes or the car tries to reduce the cross-track error c_e by adjusting θ, e.g., when $|c_e| > 0$, we need $|\theta_e| > 0$ to reduce $|c_e|$

(see Figure 1). To analyze how well the orientation error is adjusted when the cross-track error leaves the interval $[-1.25, 1.25]$, we consider the specification

$$\phi_5 := G_{[0,\infty)}\Big((c_e \geq 1.25) \implies F_{[0,2]}G_{[0,1]}(\theta_e \leq 0) \wedge (c_e \leq -1.25) \implies F_{[0,2]}G_{[0,1]}(\theta_e \geq 0)\Big).$$

The specification ϕ_5 encodes that, whenever the cross-track error c_e leaves the interval $[-1.25, 1.25]$, the orientation error θ_e should, within 2 s, be such that the cross-track error decreases for at least 1 s. We show the histogram of each controller in Figure 7(r) and report the risk estimates below.

u \ R	$\overline{VaR}_{0.85}$	$\overline{VaR}_{0.9}$	$\overline{VaR}_{0.95}$	$\overline{E}$	$\#_{\phi_5}$
IL_{full}	-0.58	-0.54	-0.13	-0.517	1
$\text{IL}_{\text{partial}}$	1.25	1.25	1.25	0.762	0.247
CBF_{full}	-0.47	-0.44	-0.32	-0.553	1
$\text{CBF}_{\text{partial}}$	0.43	1.14	1.25	0.225	0.503

We can observe that the risk of IL_{full} is the lowest for $\overline{VaR}_{0.85}$ and $\overline{VaR}_{0.9}$, while the risks of IL_{full} and CBF_{full} are roughly equal for the expected value $\overline{E}$. However, the distribution induced by IL_{full} has a long tail, which is why the risk of CBF_{full} is the lowest for $\overline{VaR}_{0.95}$.

7 CONCLUSION

We defined the STL robustness risk to quantify the risk of a stochastic system lacking robustness against failure of an STL specification. The approximate STL robustness risk was defined as a computationally tractable upper bound of the STL robustness risk. It was shown how the approximate STL robustness risk is estimated from data for the value-at-risk and the conditional value-at-risk. We also provided conditions under which the approximate STL robustness risk can be computed exactly. Within the autonomous driving simulator CARLA, we trained four different neural network lane-keeping controllers and estimated their risk for five different STL system specifications.

APPENDICES

A SEMANTICS OF SIGNAL TEMPORAL LOGIC

The satisfaction function $\beta^\phi(x,t)$ determines whether or not the signal x satisfies the specification ϕ at time t. The definition of $\beta^\phi(x,t)$ follows recursively from the structure of ϕ as follows:

Definition 5 (STL Semantics). For a signal $x : T \to \mathbb{R}^n$ and an STL formula ϕ, the satisfaction function $\beta^\phi(x,t)$ is recursively defined as

$$\begin{aligned}
\beta^\top(x,t) &:= \top,\\
\beta^\mu(x,t) &:= \begin{cases} \top & \text{if } x(t) \in O^\mu \\ \bot & \text{otherwise,} \end{cases}\\
\beta^{\neg\phi}(x,t) &:= \neg\beta^\phi(x,t),\\
\beta^{\phi'\wedge\phi''}(x,t) &:= \min(\beta^{\phi'}(x,t), \beta^{\phi''}(x,t)),\\
\beta^{\phi' U_I \phi''}(x,t) &:= \sup_{t''\in(t\oplus I)\cap T} \Big(\min\Big(\beta^{\phi''}(x,t''), \inf_{t'\in(t,t'')\cap T} \beta^{\phi'}(x,t')\Big)\Big),\\
\beta^{\phi' \underline{U}_I \phi''}(x,t) &:= \sup_{t''\in(t\ominus I)\cap T} \Big(\min\Big(\beta^{\phi''}(x,t''), \inf_{t'\in(t'',t)\cap T} \beta^{\phi'}(x,t')\Big)\Big).
\end{aligned}$$

The semantics in Definition 5 use the strict non-matching versions U_I and $\underline{U}_I$ of the until operators. The non-strict matching versions of the until operator, in comparison, replace the open time intervals (t, t'') in Definition 5 by the closed time intervals $[t, t'']$ as follows:

$$\beta^{\phi' \vec{U}_I \phi''}(x,t) := \sup_{t'' \in (t \oplus I) \cap T} \Big(\min \Big(\beta^{\phi''}(x,t''), \inf_{t' \in [t,t''] \cap T} \beta^{\phi'}(x,t') \Big) \Big),$$

$$\beta^{\phi' \vec{\underline{U}}_I \phi''}(x,t) := \sup_{t'' \in (t \ominus I) \cap T} \Big(\min \Big(\beta^{\phi''}(x,t''), \inf_{t' \in [t'',t] \cap T} \beta^{\phi'}(x,t') \Big) \Big).$$

B PROOF OF THEOREM 1

We prove the statement of Theorem 1 first for the semantics $\beta^{\phi}(X,t)$, then for the robust semantics $\rho^{\phi}(X,t)$, and finally for the robustness degree $\mathrm{RD}^{\phi}(X,t)$.

B.1 Semantics $\beta^{\phi}(X,t)$

Let us define the power set of $\mathbb{B}$ as $2^{\mathbb{B}} := \{\emptyset, \top, \bot, \{\bot, \top\}\}$. Note that $2^{\mathbb{B}}$ is a σ-algebra of $\mathbb{B}$. To prove measurability of $\beta^{\phi}(X(\cdot,\omega),t)$ in ω for a fixed $t \in T$, we need to show that, for each $B \in 2^{\mathbb{B}}$, it holds that the inverse image of B under $\beta^{\phi}(X(\cdot,\omega),t)$ for a fixed $t \in T$ is contained within $\mathcal{F}$, i.e., that it holds that

$$\{\omega \in \Omega | \beta^{\phi}(X(\cdot,\omega),t) \in B\} \subseteq \mathcal{F}.$$

We show measurability of $\beta^{\phi}(X(\cdot,\omega),t)$ in ω for a fixed $t \in T$ inductively on the structure of ϕ.

$\top$: For $B \in 2^{\mathbb{B}}$, it trivially holds that $\{\omega \in \Omega | \beta^{\top}(X(\cdot,\omega),t) \in B\} \subseteq \mathcal{F}$, since $\beta^{\top}(X(\cdot,\omega),t) = \top$ for all $\omega \in \Omega$. This follows according to Definition 5 so $\{\omega \in \Omega | \beta^{\top}(X(\cdot,\omega),t) \in B\} = \emptyset \subseteq \mathcal{F}$ if $B \in \{\emptyset, \bot\}$ and $\{\omega \in \Omega | \beta^{\top}(X(\cdot,\omega),t) \in B\} = \Omega \subseteq \mathcal{F}$ otherwise.

μ: Let $1_{O^{\mu}} : \mathbb{R}^n \to \mathbb{B}$ be the indicator function of O^{μ} with $1_{O^{\mu}}(\zeta) := \top$ if $\zeta \in O^{\mu}$ and $1_{O^{\mu}}(\zeta) := \bot$ otherwise. According to Definition 5, we can now write $\beta^{\mu}(X(\cdot,\omega),t) = 1_{O^{\mu}}(X(t,\omega))$. Recall that O^{μ} is measurable and note that the indicator function of a measurable set is measurable again (see, e.g., Durrett [20, Chapter 1.2]). Since $X(t,\omega)$ is measurable in ω for a fixed $t \in T$ by definition, it follows that $1_{O^{\mu}}(X(t,\omega))$ and hence $\beta^{\mu}(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$. In other words, for $B \in 2^{\mathbb{B}}$, it follows that

$$\{\omega \in \Omega | \beta^{\mu}(X(\cdot,\omega),t) \in B\} = \{\omega \in \Omega | 1_{O^{\mu}}(X(t,\omega)) \in B\} \subseteq \mathcal{F}.$$

$\neg\phi$: By the induction assumption, $\beta^{\phi}(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$. Recall that $\mathcal{F}$ is a σ-algebra that is, by definition, closed under its complement so, for $B \in 2^{\mathbb{B}}$, it holds that

$$\{\omega \in \Omega | \beta^{\neg\phi}(X(\cdot,\omega),t) \in B\} = \Omega \setminus \{\omega \in \Omega | \beta^{\phi}(X(\cdot,\omega),t) \in B\} \subseteq \mathcal{F}.$$

$\phi' \wedge \phi''$: By the induction assumption, $\beta^{\phi'}(X(\cdot,\omega),t)$ and $\beta^{\phi''}(X(\cdot,\omega),t)$ are measurable in ω for a fixed $t \in T$. Hence, $\beta^{\phi' \wedge \phi''}(X(\cdot,\omega),t) = \min(\beta^{\phi'}(X(\cdot,\omega),t), \beta^{\phi''}(X(\cdot,\omega),t))$ is measurable in ω for a fixed $t \in T$, since the min operator of measurable functions is again a measurable function.

$\phi' U_I \phi''$ and $\phi' \underline{U}_I \phi''$: Recall the definition of the future until operator

$$\beta^{\phi' U_I \phi''}(X(\cdot,\omega),t) := \sup_{t'' \in (t \oplus I) \cap T} \Big(\min(\beta^{\phi''}(X(\cdot,\omega),t''), \inf_{t' \in (t,t'') \cap T} \beta^{\phi'}(X(\cdot,\omega),t')) \Big).$$

By the induction assumption, $\beta^{\phi'}(X(\cdot,\omega),t)$ and $\beta^{\phi''}(X(\cdot,\omega),t)$ are measurable in ω for a fixed $t \in T$. First note that $(t,t'') \cap T$ and $(t \oplus I) \cap T$ are countable sets, since $T = \mathbb{N}$. According to Guide [26, Theorem 4.27], the supremum and infimum operators over a countable number of measurable functions is again measurable. Consequently, the function $\beta^{\phi' U_I \phi''}(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$. The same reasoning applies to $\beta^{\phi' \underline{U}_I \phi''}(X(\cdot,\omega),t)$.

B.2 Robust Semantics $\rho^{\phi}(X,t)$

The proof for $\rho^{\phi}(X(\cdot,\omega),t)$ follows again inductively on the structure of ϕ, and the goal is to show that $\{\omega \in \Omega | \rho^{\phi}(X(\cdot,\omega),t) \in B\} \subseteq \mathcal{F}$ for each Borel set $B \in \mathcal{B}$. The difference here, compared to the proof for the semantics $\beta^{\phi}(X(\cdot,\omega),t)$ presented above, lies only in the way predicates μ are handled. Note first that we can write $\rho^{\mu}(X(\cdot,\omega),t)$ as

$$\begin{aligned}\rho^{\mu}(X(\cdot,\omega),t) = 0.5(1_{O^{\mu}}(X(t,\omega)) + 1)\bar{d}(X(t,\omega), \mathrm{cl}(O^{\neg\mu})) \\ + 0.5(1_{O^{\mu}}(X(t,\omega)) - 1)\bar{d}(X(t,\omega), \mathrm{cl}(O^{\mu})),\end{aligned} \tag{8}$$

where we recall that we interpret $\top := 1$ and $\bot = -1$. Since the composition of the indicator function with $X(t,\omega)$, i.e., $1_{O^{\mu}}(X(t,\omega))$, is measurable in ω for a fixed $t \in T$ as argued before, we only need to show that $\bar{d}(X(t,\omega), \mathrm{cl}(O^{\mu}))$ and $\bar{d}(X(t,\omega), \mathrm{cl}(O^{\neg\mu}))$ are measurable in ω for a fixed $t \in T$. This immediately follows, since $X(t,\omega)$ is measurable in ω for a fixed $t \in T$ by definition and, since the function $\bar{d}$ is continuous in its first argument, and hence measurable (see Guide [26, Corollary 4.26]), due to d being a metric defined on the set $\mathbb{R}^n$ (see, e.g., Munkres [56, Chapter 3]) so $\rho^{\mu}(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$.

B.3 Robustness Degree $\mathrm{RD}^{\phi}(X,t)$

For $\mathrm{RD}^{\phi}(X(\cdot,\omega),t)$, note that, for a fixed $t \in T$, the function RD^{ϕ} maps from the domain $\mathfrak{F}(T,\mathbb{R}^n)$ into the domain $\mathbb{R}$, while $X(\cdot,\omega)$ maps from the domain Ω into the domain $\mathfrak{F}(T,\mathbb{R}^n)$. Recall now that $\mathrm{RD}^{\phi}(X(\cdot,\omega),t) = \bar{\kappa}(X(\cdot,\omega), \mathrm{cl}(\mathcal{L}^{\phi}(t))) := \inf_{x^* \in \mathrm{cl}(\mathcal{L}^{\phi}(t))} \kappa(X(\cdot,\omega), x^*)$ and that κ is a metric defined on the set $\mathfrak{F}(T,\mathbb{R}^n)$, as argued in Fainekos and Pappas [21]. Therefore, it follows that the function $\bar{\kappa}$ is continuous in its first argument (see, e.g., Munkres [56, Chapter 3]), and hence measurable with respect to the Borel σ-algebra of $\mathfrak{F}(T,\mathbb{R}^n)$ (see, e.g., Guide [26, Corollary 4.26]). Consequently, the function $\mathrm{RD}^{\phi} : \mathfrak{F}(T,\mathbb{R}^n) \times T \to \mathbb{R}^n$ is measurable in its first argument for a fixed $t \in T$. As T is countable and X is a discrete-time stochastic process, it follows that $X(\cdot,\omega)$ is measurable with respect to the product σ-algebra of Borel σ-algebras $\mathcal{B}^n$, which is equivalent to the Borel σ-algebra of $\mathfrak{F}(T,\mathbb{R}^n)$ (see, e.g., Kallenberg [36, Lemma 1.2]). Since function composition preserves measurability, it holds that $\mathrm{RD}^{\phi}(X(\cdot,\omega),t)$ is measurable in ω for a fixed $t \in T$.

C PROOF OF THEOREM 2

We prove the statement of Theorem 2 first for the robustness degree $\mathrm{RD}^{\phi}(X,t)$, and finally for the semantics $\beta^{\phi}(X,t)$, then for the robust semantics $\rho^{\phi}(X,t)$.

C.1 Semantics $\beta^{\phi}(X,t)$

The proof again follows inductively on the structure of ϕ. The difference to the proof of Theorem 1 lies in the way the until operators are handled, which are now assumed to be the non-strict matching versions $\phi' \vec{U}_I \phi''$ and $\phi' \vec{\underline{U}}_I \phi''$. Note also that the time interval I is compact, as the formula ϕ is assumed to be bounded. The main idea is to show that infimum and supremum operators reduce to minimum and maximum operators that allow us to show measurability. Recall, therefore, the definition of the future until operator $\beta^{\phi' \vec{U}_I \phi''}(X(\cdot,\omega),t)$ as

$$\beta^{\phi' \vec{U}_I \phi''}(X(\cdot,\omega),t) := \sup_{t'' \in (t \oplus I) \cap T} \left(\min \left(\beta^{\phi''}(X(\cdot,\omega),t''), \inf_{t' \in [t,t'']\cap T} \beta^{\phi'}(X(\cdot,\omega),t') \right) \right).$$

We first show that the infimum operator in $\beta^{\phi' \vec{U}_I \phi''}(X(\cdot,\omega),t)$ reduces to a min operator. In particular, note now that $\inf_{t' \in [t,t'']\cap T} \beta^{\phi'}(X(\cdot,\omega),t')$ includes the compact time interval $[t,t''] \cap T$ instead of the open interval $(t,t'') \cap T$ due to the interpretation of the until operator as the non-strict matching version. It holds that the minimum of $\min_{t' \in [t,t'']\cap T} \beta^{\phi'}(X(\cdot,\omega),t')$ exists as

(1) the minimum is over the compact time interval $[t, t''] \cap T = [t, t'']$ (recall that $T = \mathbb{R}$), and
(2) the range of $\beta^{\phi'}(X(\cdot, \omega), t)$ is restricted to $\mathbb{B}$.

Consequently, the minimum corresponds to the infimum and it follows that

$$\inf_{t' \in [t,t'] \cap T} \beta^{\phi'}(X(\cdot, \omega), t') = \min_{t' \in [t,t''] \cap T} \beta^{\phi'}(X(\cdot, \omega), t').$$

Now, it holds that $\min_{t' \in [t,t''] \cap T} \beta^{\phi'}(X(\cdot, \omega), t')$ is equivalent to $\beta^{\phi'}(X(\cdot, \omega), t')$ for some $t' \in [t, t''] \cap T$. Since $\beta^{\phi'}(X(\cdot, \omega), t')$ is measurable in ω by the induction assumption, it follows that the function $\inf_{t' \in [t,t''] \cap T} \beta^{\phi'}(X(\cdot, \omega), t')$ is measurable in ω for a fixed $t \in T$. Note next that the supremum operator in $\beta^{\phi' \vec{U}_I \phi''}(X(\cdot, \omega), t)$ reduces to a max operator due to I being compact and following a similar argument as for the infimum operator. Measurability of $\beta^{\phi' \vec{U}_I \phi''}(X(\cdot, \omega), t)$ in ω for a fixed $t \in T$ then follows as in the proof of Theorem 1. The proof for $\beta^{\phi' \vec{\underline{U}}_I \phi''}(X(\cdot, \omega), t)$ follows similarly.

C.2 Robustness Degree $\mathrm{RD}^{\phi}(X, t)$

As shown in the proof of Theorem 1, the function $\mathrm{RD}^{\phi} : \mathfrak{F}(T, \mathbb{R}^n) \times T \to \mathbb{R}^n$ is continuous and hence Borel-measurable in its first argument for a fixed $t \in T$. By the assumption that $X(\cdot, \omega) : \Omega \to \mathfrak{F}(T, \mathbb{R}^n)$ is Borel-measurable, the result follows trivially.

C.3 Robust Semantics $\rho^{\phi}(X, t)$

The proof follows mainly from Reference [8, Theorem 6]. However, to apply this result, we need to show that the robust semantics $\rho^{\mu}(\zeta, t)$ of predicates μ are continuous in $\zeta \in \mathbb{R}^n$, where we recall that

$$\rho^{\mu}(\zeta, t) := \begin{cases} \bar{d}(\zeta, \mathrm{cl}(O^{\neg \mu})) & \text{if } \zeta \in O^{\mu} \\ -\bar{d}(\zeta, \mathrm{cl}(O^{\mu})) & \text{otherwise.} \end{cases}$$

Note that the functions $\bar{d}(\zeta, \mathrm{cl}(O^{\neg \mu}))$ and $\bar{d}(\zeta, \mathrm{cl}(O^{\mu}))$ are continuous in ζ. This follows due to Munkres [56, Chapter 3]. By definition, we have $\rho^{\mu}(\zeta, t) = 0$ if $\zeta \in \mathrm{bd}(O^{\mu})$, where $\mathrm{bd}(O^{\mu})$ denotes the boundary of O^{μ}. Note also that $\bar{d}(\zeta, \mathrm{cl}(O^{\neg \mu})) \to 0$ as $\zeta \to \mathrm{bd}(O^{\mu})$ as well as $-\bar{d}(\zeta, \mathrm{cl}(O^{\mu})) \to 0$ as $\zeta \to \mathrm{bd}(O^{\mu})$. It follows that $\rho^{\mu}(\zeta, t)$ is continuous in ζ. The assumption that $X(\cdot, \omega)$ is a cadlag function for each $\omega \in \Omega$ then enables us to apply Theorem 6 in Bartocci et al. [8].

D PROOF OF THEOREM 3

First note that $\rho^{\phi}(X(\cdot, \omega), t) \leq \mathrm{RD}^{\phi}(X(\cdot, \omega), t)$ for each realization $X(\cdot, \omega)$ of the stochastic process X with $\omega \in \Omega$ due to Equation (5). Consequently, we have that $-\mathrm{RD}^{\phi}(X(\cdot, \omega), t) \leq -\rho^{\phi}(X(\cdot, \omega), t)$ for all $\omega \in \Omega$. If R is now monotone, then it directly follows that $R(-\mathrm{RD}^{\phi}(X, t)) \leq R(-\rho^{\phi}(X, t))$.

E PROOF OF PROPOSITION 1

Let us assume that $X^1, \ldots, X^N$ are N independent copies of X. Consequently, all Z^i contained within $\mathcal{Z}$ are independent and identically distributed. We first recall the tight version of the Dvoretzky-Kiefer-Wolfowitz inequality as originally presented in Massart [52], which requires that F_Z is continuous.

Lemma 1. *Let $\widehat{F}(\alpha, \mathcal{Z})$ be based on the data $\mathcal{Z}$ consisting of $Z^1, \ldots, Z^N$, which are N independent copies of Z. Let $c > 0$ be a desired precision, then it holds that*

$$P\left(\sup_{\alpha} |\widehat{F}(\alpha, \mathcal{Z}) - F_Z(\alpha)| > c\right) \leq 2 \exp(-2Nc^2).$$

By setting $\delta := 2\exp(-2Nc^2)$ in Lemma 1, it holds with a probability of at least $1-\delta$ that

$$\sup_{\alpha} |\widehat{F}(\alpha, \mathcal{Z}) - F_Z(\alpha)| \leq \sqrt{\frac{\ln(2/\delta)}{2N}}.$$

With a probability of at least $1-\delta$, it now holds that

$$\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha, \mathcal{Z}) - \sqrt{\frac{\ln(2/\delta)}{2N}} \geq \beta\right\} \subseteq \{\alpha \in \mathbb{R} | F_Z(\alpha) \geq \beta\}$$

as well as

$$\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha, \mathcal{Z}) + \sqrt{\frac{\ln(2/\delta)}{2N}} \geq \beta\right\} \supseteq \{\alpha \in \mathbb{R} | F_Z(\alpha) \geq \beta\}.$$

Hence, it holds with a probability of at least $1-\delta$ that

$$\inf\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha, \mathcal{Z}) - \sqrt{\frac{\ln(2/\delta)}{2N}} \geq \beta\right\} \geq \inf\{\alpha \in \mathbb{R} | F_Z(\alpha) \geq \beta\}$$

as well as

$$\inf\left\{\alpha \in \mathbb{R} | \widehat{F}(\alpha, \mathcal{Z}) + \sqrt{\frac{\ln(2/\delta)}{2N}} \geq \beta\right\} \leq \inf\{\alpha \in \mathbb{R} | F_Z(\alpha) \geq \beta\}.$$

By the definition of $\underline{VaR}_\beta(\mathcal{Z}, \delta)$ and $\overline{VaR}_\beta(\mathcal{Z}, \delta)$, it holds with a probability of at least $1-\delta$ that

$$\underline{VaR}_\beta(\mathcal{Z}, \delta) \leq VaR_\beta(Z) \leq \overline{VaR}_\beta(\mathcal{Z}, \delta).$$

F PROOF OF PROPOSITION 3

Let us again assume that $X^1, \ldots, X^N$ are N independent copies of X. Consequently, all Z^i contained within $\mathcal{Z}$ are independent and identically distributed. Note first that $\widehat{E}(\mathcal{Z})$ is a random variable with the expected value according to

$$E(\widehat{E}(\mathcal{Z})) = \frac{1}{N}\sum_{i=1}^{N} E(Z_i) = \frac{1}{N}\sum_{i=1}^{N} E(Z) = E(Z).$$

For $c > 0$, we can now apply Hoeffding's inequality and obtain the concentration inequality

$$P\Big(|\widehat{E}(\mathcal{Z}) - E(Z)| \geq c\Big) \leq 2\exp\Big(-\frac{2Nc^2}{(b-a)^2}\Big).$$

By setting $\delta := 2\exp(-\frac{2Nc^2}{(b-a)^2})$, it holds with a probability of at least $1-\delta$ that

$$|\widehat{E}(\mathcal{Z}) - E(Z)| \leq \sqrt{\frac{\ln(2/\delta)(b-a)^2}{2N}}.$$

From this inequality, the result follows trivially.

G DISCRETIZATION OF c AND d IN EXAMPLE 1

To discretize the distributions of c and d in Equations (6) and (7), respectively, let $M := 32$ be the number of desired discretization steps and $\gamma := 0.55$ be a discretization bound. We uniformly discretize the interval $[-\gamma, \gamma]$ into M values $(s_1, \dots, s_M)$ where $s_m < s_{m+1}$. We additionally add $s_0 := 0$ and define $S := (s_0, s_1, \dots, s_M)$. We now assign a PMF $f_S(s_m)$ to each element $s_m \in S$ as

$$f_S(s_m) := \begin{cases} F_{\mathcal{N}}(s_m) & \text{if } s_m = s_1 \\ F_{\mathcal{N}}(s_m) - F_{\mathcal{N}}(s_{m-1}) & \text{if } s_1 < s_m < 0 \\ 2(F_{\mathcal{N}}(s_m) - F_{\mathcal{N}}(s_{m-1})) & \text{if } s_m = 0 \\ F_{\mathcal{N}}(s_{m+1}) - F_{\mathcal{N}}(s_m) & \text{if } 0 < s_m < s_M \\ 1 - F_{\mathcal{N}}(s_m) & \text{if } s_m = s_M, \end{cases}$$

where $F_{\mathcal{N}}(s)$ is the CDF of $\mathcal{N}(0, 0.2)$ (according to Equations (6) and (7)). We now assume, instead of Equations (6) and (7), that c and d take values in the sets

$$\mathcal{C} := 2 \oplus S \times 3 \oplus S$$
$$\mathcal{D} := 6 \oplus S \times 4 \oplus S,$$

where 2, 3, 6, and 4 are the mean values of c and d in Equations (6) and (7), respectively. Finally, we assume that the distributions of $c = \begin{bmatrix} c_1 & c_2 \end{bmatrix}^T$ and $d = \begin{bmatrix} d_1 & d_2 \end{bmatrix}^T$ are according to the PMFs $f_c(c) := f_S(c_1) f_S(c_2)$ and $f_d(d) := f_S(d_1) f_S(d_2)$, respectively.

REFERENCES

[1] Gul Agha and Karl Palmskog. 2018. A survey of statistical model checking. *ACM Trans. Model. Comput. Simul.* 28, 1 (2018), 1–39.

[2] Mohamadreza Ahmadi, Xiaobin Xiong, and Aaron D. Ames. 2022. Risk-averse control via CVaR barrier functions: Application to bipedal robot locomotion. *IEEE Contr. Syst. Lett.* 6 (2022), 878–883.

[3] Takumi Akazaki and Ichiro Hasuo. 2015. Time robustness in MTL and expressivity in hybrid system falsification. In *Proceedings of the International Conference on Computer-Aided Verification*. 356–374.

[4] Tzanis Anevlavis, Matthew Philippe, Daniel Neider, and Paulo Tabuada. 2022. Being correct is not enough: Efficient verification using robust linear temporal logic. *ACM Trans. Computat. Logic* 23, 2 (2022), 1–39.

[5] Nasim Baharisangari, Jean-Raphaël Gaglione, Daniel Neider, Ufuk Topcu, and Zhe Xu. 2021. Uncertainty-aware signal temporal logic inference. In *Software Verification*. Springer, 61–85.

[6] Christel Baier and Joost-Pieter Katoen. 2008. *Principles of Model Checking* (1st ed.). The MIT Press, Cambridge, MA.

[7] Ezio Bartocci, Luca Bortolussi, Laura Nenzi, and Guido Sanguinetti. 2013. On the robustness of temporal properties for stochastic models. In *Proceedings of the Workshop on Hybrid Systems and Biology*. 3–19.

[8] Ezio Bartocci, Luca Bortolussi, Laura Nenzi, and Guido Sanguinetti. 2015. System design of stochastic models using robustness of temporal properties. *Theoret. Comput. Sci.* 587 (2015), 3–25.

[9] Ezio Bartocci, Jyotirmoy Deshmukh, Alexandre Donzé, Georgios Fainekos, Oded Maler, Dejan Ničković, and Sriram Sankaranarayanan. 2018. Specification-based monitoring of cyber-physical systems: A survey on theory, tools and applications. In *Lectures on Runtime Verification*. Springer, 135–175.

[10] Suda Bharadwaj, Rayna Dimitrova, and Ufuk Topcu. 2018. Synthesis of surveillance strategies via belief abstraction. In *Proceedings of the Conference on Decision and Control*. 4159–4166.

[11] Sanjay P. Bhat and Prashanth L. A. 2019. Concentration of risk measures: A Wasserstein distance approach. *Proc. Conf. Neural Inf. Process. Syst.* 32 (2019), 11762–11771.

[12] National Transportation Safety Board. 2019. Collision between vehicle controlled by developmental automated driving system and pedestrian. *Highw. Accid. Rep. NTSB/HAR-19/03* (2019).

[13] David B. Brown. 2007. Large deviations bounds for estimating conditional value-at-risk. *Oper. Res. Lett.* 35, 6 (2007), 722–730.

[14] Christos G. Cassandras and Stephane Lafortune. 2009. *Introduction to Discrete Event Systems*. Springer Science & Business Media.

[15] Margaret P. Chapman, Jonathan Lacotte, Aviv Tamar, Donggun Lee, Kevin M. Smith, Victoria Cheng, Jaime F. Fisac, Susmit Jha, Marco Pavone, and Claire J. Tomlin. 2019. A risk-sensitive finite-time reachability approach for safety of stochastic dynamic systems. In *Proceedings of the American Control Conference*. 2958–2963.

[16] Margaret P. Chapman, Jonathan P. Lacotte, Kevin M. Smith, Insoon Yang, Yuxi Han, Marco Pavone, and Claire J. Tomlin. 2019. Risk-sensitive safety specifications for stochastic systems using conditional value-at-risk. *arXiv preprint arXiv:1909.09703* (2019).

[17] Jeremy Coulson, John Lygeros, and Florian Dörfler. 2021. Distributionally robust chance constrained data-enabled predictive control. *IEEE Trans. Automat. Contr.* 67, 7 (2021).

[18] Alexandre Donzé and Oded Maler. 2010. Robust satisfaction of temporal logic over real-valued signals. In *Proceedings of the Conference on Formal Modeling and Analysis of Timed Systems*. 92–106.

[19] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. CARLA: An open urban driving simulator. In *Proceedings of the Conference on Robot Learning*. 1–16.

[20] Rick Durrett. 2019. *Probability: Theory and Examples*. Vol. 49. Cambridge University Press.

[21] Georgios E. Fainekos and George J. Pappas. 2009. Robustness of temporal logic specifications for continuous-time signals. *Theoret. Comput. Sci.* 410, 42 (2009), 4262–4291.

[22] Chuchu Fan, Bolun Qi, Sayan Mitra, and Mahesh Viswanathan. 2017. DryVR: Data-driven verification and compositional reasoning for automotive systems. In *Proceedings of the International Conference on Computer Aided Verification*. 441–461.

[23] Samira S. Farahani, Rupak Majumdar, Vinayak S. Prabhu, and Sadegh Soudjani. 2018. Shrinking horizon model predictive control with signal temporal logic constraints under stochastic disturbances. *IEEE Trans. Automat. Contr.* 64, 8 (2018), 3324–3331.

[24] Carlo Alberto Furia and Matteo Rossi. 2007. On the expressiveness of MTL variants over dense time. In *Proceedings of the International Conference on Formal Modeling and Analysis of Timed Systems*. 163–178.

[25] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. 2014. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572* (2014).

[26] A Hitchhiker's Guide. 2006. *Infinite Dimensional Analysis*. Springer.

[27] Meng Guo and Michael M. Zavlanos. 2018. Probabilistic motion planning under temporal tasks and soft constraints. *IEEE Trans. Automat. Contr.* 63, 12 (2018), 4051–4066.

[28] Iman Haghighi, Noushin Mehdipour, Ezio Bartocci, and Calin Belta. 2019. Control from signal temporal logic specifications with smooth cumulative quantitative semantics. In *Proceedings of the Conference on Decision and Control*. 4361–4366.

[29] Eunjeong Hyeon, Youngki Kim, and Anna G. Stefanopoulou. 2020. Fast risk-sensitive model predictive control for systems with time-series forecasting uncertainties. In *Proceedings of the Conference on Decision and Control*. 2515–2520.

[30] Radoslav Ivanov, James Weimer, Rajeev Alur, George J. Pappas, and Insup Lee. 2019. Verisig: Verifying safety properties of hybrid systems with neural network controllers. In *Proceedings of the International Conference on Hybrid Systems: Computation and Control*. 169–178.

[31] John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. 2021. Formal verification of unknown dynamical systems via Gaussian process regression. *arXiv preprint arXiv:2201.00655* (2021).

[32] Pushpak Jagtap, Sadegh Soudjani, and Majid Zamani. 2018. Temporal logic verification of stochastic systems using barrier certificates. In *Proceedings of the International Symposium on Automated Technology for Verification and Analysis*. 177–193.

[33] Ashkan Jasour, Weiqiao Han, and Brian Williams. 2021. Real-time risk-bounded tube-based trajectory safety verification. In *Proceedings of the Conference on Decision and Control*. 4307–4313.

[34] Ashkan Jasour, Xin Huang, Allen Wang, and Brian C. Williams. 2021. Fast nonlinear risk assessment for autonomous vehicles using learned conditional probabilistic models of agent futures. *Auton. Robot* (2021), 1–14.

[35] Susmit Jha, Vasumathi Raman, Dorsa Sadigh, and Sanjit A. Seshia. 2018. Safe autonomy under perception uncertainty using chance-constrained temporal logic. *J. Automat. Reason.* 60, 1 (2018), 43–62.

[36] Olav Kallenberg. 1997. *Foundations of Modern Probability*. Vol. 2. Springer.

[37] Dionysios S. Kalogerias, Luiz F. O. Chamon, George J. Pappas, and Alejandro Ribeiro. 2020. Better safe than sorry: Risk-aware nonlinear Bayesian estimation. In *Proceedings of the Conference on Acoustics, Speech and Signal Processing*. 5480–5484.

[38] Guy Katz, Derek A. Huang, Duligur Ibeling, Kyle Julian, Christopher Lazarus, Rachel Lim, Parth Shah, Shantanu Thakoor, Haoze Wu, Aleksandar Zeljić, et al. 2019. The Marabou framework for verification and analysis of deep neural networks. In *Proceedings of the International Conference on Computer Aided Verification*. 443–452.

[39] Ravi Kumar Kolla, L. A. Prashanth, Sanjay P. Bhat, and Krishna Jagannathan. 2019. Concentration bounds for empirical conditional value-at-risk: The unbounded case. *Oper. Res. Lett.* 47, 1 (2019), 16–20.

[40] Marta Kwiatkowska, Gethin Norman, and David Parker. 2007. Stochastic model checking. In *Proceedings of the International School on Formal Methods for the Design of Computer, Communication and Software Systems*. 220–270.
[41] Panagiotis Kyriakis, Jyotirmoy V. Deshmukh, and Paul Bogdan. 2019. Specification mining and robust design under uncertainty: A stochastic temporal logic approach. *ACM Trans. Embed. Comput. Syst.* 18, 5s (2019), 1–21.
[42] Morteza Lahijanian, Sean B. Andersson, and Calin Belta. 2015. Formal verification and synthesis for discrete-time stochastic systems. *IEEE Trans. Automat. Contr.* 60, 8 (2015), 2031–2045.
[43] Axel Legay, Anna Lukina, Louis Marie Traonouez, Junxing Yang, Scott A. Smolka, and Radu Grosu. 2019. Statistical model checking. In *Computing and Software Science*. Springer, 478–504.
[44] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. 2016. End-to-end training of deep visuomotor policies. *J. Mach. Learn. Res.* 17, 1 (2016), 1334–1373.
[45] Jiwei Li, Pierluigi Nuzzo, Alberto Sangiovanni-Vincentelli, Yugeng Xi, and Dewei Li. 2017. Stochastic contracts for cyber-physical system design under probabilistic requirements. In *Proceedings of the International Conference on Formal Methods and Models for System Design*. 5–14.
[46] Xiao Li, Jonathan DeCastro, Cristian Ioan Vasile, Sertac Karaman, and Daniela Rus. 2022. Learning a risk-aware trajectory planner from demonstrations using logic monitor. In *Proceedings of the Conference on Robot Learning*. PMLR, 1326–1335.
[47] Lars Lindemann, Nikolai Matni, and George J. Pappas. 2021. STL robustness risk over discrete-time stochastic processes. In *Proceedings of the Conference on Decision and Control*. 1329–1335.
[48] Lars Lindemann, George J. Pappas, and Dimos V. Dimarogonas. 2021. Reactive and risk-aware control for signal temporal logic. *IEEE Trans. Automat. Contr.* 67, 10 (2021).
[49] Lars Lindemann, Alexander Robey, Lejun Jiang, Stephen Tu, and Nikolai Matni. 2021. Learning robust output control barrier functions from safe expert demonstrations. *arXiv preprint arXiv:2111.09971* (2021).
[50] Anirudha Majumdar and Marco Pavone. 2020. How should a robot assess risk? Towards an axiomatic theory of risk in robotics. In *Robotics Research*. Springer, 75–84.
[51] Oded Maler and Dejan Nickovic. 2004. Monitoring temporal properties of continuous signals. In *Proceedings of the Conference on Formal Techniques, Modelling and Analysis of Timed and Fault-tolerant Systems*. 152–166.
[52] Pascal Massart. 1990. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.* (1990), 1269–1283.
[53] Noushin Mehdipour, Cristian-Ioan Vasile, and Calin Belta. 2019. Arithmetic-geometric mean robustness for control from signal temporal logic specifications. In *Proceedings of the American Control Conference*. 1690–1695.
[54] Zakaria Mhammedi, Benjamin Guedj, and Robert C. Williamson. 2020. PAC-Bayesian bound for the conditional value at risk. *Proc. Conf. Adv. Neural Inf. Process. Syst.* 33 (2020), 17919–17930.
[55] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533.
[56] James R. Munkres. 2000. *Topology* (2nd ed.). Prentice Hall.
[57] Konstantinos E. Nikolakakis, Dionysios S. Kalogerias, Or Sheffet, and Anand D. Sarwate. 2021. Quantile multi-armed bandits: Optimal best-arm identification and a differentially private scheme. *IEEE J. Select. Areas Inf. Theor* 2, 2 (2021), 534–548.
[58] Truls Nyberg, Christian Pek, Laura Dal Col, Christoffer Norén, and Jana Tumova. 2021. Risk-aware motion planning for autonomous vehicles with safety specifications. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 1016–1023.
[59] Stephen Prajna, Ali Jadbabaie, and George J. Pappas. 2007. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Trans. Automat. Contr.* 52, 8 (2007), 1415–1428.
[60] Aniruddh G. Puranic, Jyotirmoy V. Deshmukh, and Stefanos Nikolaidis. 2021. Learning from demonstrations using signal temporal logic. In *Proceedings of the Conference on Robot Learning*.
[61] Alexander Robey, Hamed Hassani, and George J. Pappas. 2020. Model-based robust deep learning: Generalizing to natural, out-of-distribution data. *arXiv preprint arXiv:2005.10247* (2020).
[62] R. Tyrrell Rockafellar and Stanislav Uryasev. 2000. Optimization of conditional value-at-risk. *J. Risk* 2 (2000), 21–42.
[63] R. Tyrrell Rockafellar and Stanislav Uryasev. 2002. Conditional value-at-risk for general loss distributions. *J. Bank. Fin.* 26, 7 (2002), 1443–1471.
[64] Alena Rodionova, Ezio Bartocci, Dejan Nickovic, and Radu Grosu. 2016. Temporal logic as filtering. In *Proceedings of the International Conference on Hybrid Systems: Computation and Control*. 11–20.
[65] Stéphane Ross and Drew Bagnell. 2010. Efficient reductions for imitation learning. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*. 661–668.
[66] Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. 2011. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*. 627–635.

[67] Dorsa Sadigh and Ashish Kapoor. 2016. Safe control under uncertainty with probabilistic signal temporal logic. In *Proceedings of Robotics: Science and Systems XII*. AnnArbor, Michigan.
[68] Sadra Sadraddini and Calin Belta. 2015. Robust temporal logic model predictive control. In *Proceedings of the Conference on Communication, Control, and Computing*. 772–779.
[69] Sleiman Safaoui, Lars Lindemann, Dimos V. Dimarogonas, Iman Shames, and Tyler H. Summers. 2020. Control design for risk-based signal temporal logic specifications. *IEEE Contr. Syst. Lett.* 4, 4 (2020), 1000–1005.
[70] Ali Salamati, Sadegh Soudjani, and Majid Zamani. 2020. Data-driven verification under signal temporal logic constraints. *IFAC-PapersOnLine* 53, 2 (2020), 69–74.
[71] Ali Salamati, Sadegh Soudjani, and Majid Zamani. 2021. Data-driven verification of stochastic linear systems with signal temporal logic constraints. *Automatica* 131 (2021), 109781.
[72] Samantha Samuelson and Insoon Yang. 2018. Safety-aware optimal control of stochastic systems using conditional value-at-risk. In *Proceedings of the American Control Conference*. 6285–6290.
[73] Mathijs Schuurmans and Panagiotis Patrinos. 2020. Learning-based distributionally robust model predictive control of Markovian switching systems with guaranteed stability and recursive feasibility. In *Proceedings of the Conference on Decision and Control*. 4287–4292.
[74] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144.
[75] Gagandeep Singh, Timon Gehr, Markus Püschel, and Martin Vechev. 2019. An abstract domain for certifying neural networks. *Proc. ACM Program. Lang.* 3 (2019), 1–30.
[76] Sumeet Singh, Yinlam Chow, Anirudha Majumdar, and Marco Pavone. 2018. A framework for time-consistent, risk-sensitive model predictive control: Theory and algorithms. *IEEE Trans. Automat. Contr.* 64, 7 (2018), 2905–2912.
[77] Jiawei Su, Danilo Vasconcellos Vargas, and Kouichi Sakurai. 2019. One pixel attack for fooling deep neural networks. *IEEE Trans. Evolut. Computat.* 23, 5 (2019), 828–841.
[78] Balazs Szorenyi, Róbert Busa-Fekete, Paul Weng, and Eyke Hüllermeier. 2015. Qualitative multi-armed bandits: A quantile-based approach. In *Proceedings of the International Conference on Machine Learning*. 1660–1668.
[79] Philip Thomas and Erik Learned-Miller. 2019. Concentration inequalities for conditional value at risk. In *Proceedings of the International Conference on Machine Learning*. 6225–6233.
[80] Mattias Tiger and Fredrik Heintz. 2020. Incremental reasoning in probabilistic signal temporal logic. *Int. J. Approx. Reason.* 119 (2020), 325–352.
[81] Anastasios Tsiamis, Dionysios S. Kalogerias, Alejandro Ribeiro, and George J. Pappas. 2021. Linear quadratic control with risk constraints. *arXiv preprint arXiv:2112.07564* (2021).
[82] Cristian-Ioan Vasile, Kevin Leahy, Eric Cristofalo, Austin Jones, Mac Schwager, and Calin Belta. 2016. Control in belief space with temporal logic specifications. In *Proceedings of the Conference on Decision and Control*. 7419–7424.
[83] Ying Wang and Fuqing Gao. 2010. Deviation inequalities for an estimator of the conditional value-at-risk. *Oper. Res. Lett.* 38, 3 (2010), 236–239.
[84] Yu Wang, Mojtaba Zarei, Borzoo Bonakdarpour, and Miroslav Pajic. 2019. Statistical verification of hyperproperties for cyber-physical systems. *ACM Trans. Embed. Comput. Syst.* 18, 5s (2019), 1–23.
[85] Paolo Zuliani, André Platzer, and Edmund M. Clarke. 2010. Bayesian statistical model checking with application to simulink/stateflow verification. In *Proceedings of the International Conference on Hybrid Systems: Computation and Control*. 243–252.

Received 10 February 2022; revised 12 December 2022; accepted 9 January 2023