2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Risk-aware Spatio-temporal Logic Planning in Gaussian Belief Spaces

Matti Vahs, Christian Pek and Jana Tumova

***Abstract*— In many real-world robotic scenarios, we cannot assume exact knowledge about a robot's state due to unmodeled dynamics or noisy sensors. Planning in *belief space* addresses this problem by tightly coupling perception and planning modules to obtain trajectories that take into account the environment's stochasticity. However, existing works are often limited to tasks such as the classic reach-avoid problem and do not provide risk awareness. We propose a risk-aware planning strategy in belief space that minimizes the risk of violating a given specification and enables a robot to actively gather information about its state. We use Risk Signal Temporal Logic (RiSTL) as a specification language in belief space to express complex spatio-temporal missions including predicates over Gaussian beliefs. We synthesize trajectories for challenging scenarios that cannot be expressed through classical reach-avoid properties and show that risk-aware objectives improve the uncertainty reduction in a robot's belief.**

I. INTRODUCTION

Robots that operate in real world environments have to deal with uncertainty arising from unmodeled dynamics or imperfect sensor readings. The assumption of full state knowledge is often not justifiable for robots in the wild and state estimation pipelines need to be applied to obtain a robot's *belief*, i.e, a probability distribution over states, and planning needs to happen in belief space.

Traditionally, belief space planning (BSP) enables the robot to actively gain information about its state by explicitly penalizing uncertainties in the objective of the BSP problem [1]. For example, Fig. 1 shows a drone that uses noisy measurements from an antenna station for localization. The closer the drone is to the antenna, the more accurate the measurements get which can be expressed in a BSP problem to obtain the planned trajectory in red.

Suppose the drone is operating in a hazardous environment including safety-critical areas in which the robot might get damaged due to the intensity of, e.g., a radioactive source [2]. A common approach to handling safety constraints under uncertainty is to bound the probability of undesired random events through chance constraints [3]–[5]. The use of risk measures provides another comprehensive framework to quantify the outcome of such random events. Risk measures can distinguish the severity of violating the desired constraints and typically relate to one of the tail events. In contrast to chance constraints, which evaluate the probability of satisfying a given safety constraint, risk measures evaluate a certain outcome of the distribution of the safety constraint and are, thus, not saturated like chance constraints. The combination of risk measures and BSP is therefore of great interest since minimizing the risk of violating a safety constraint can lead to active information gathering behavior without explicitly penalizing uncertainties.

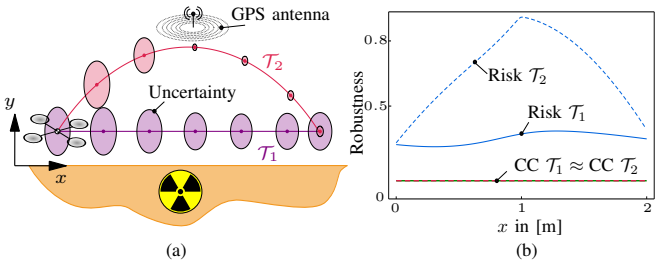

Fig. 1. (a) A drone with uncertain position is operating in a hazardous environment. Two planned trajectories $\mathcal{T}_1$ and $\mathcal{T}_2$ are shown where $\mathcal{T}_2$ is obtained through BSP. A safety constraint of the form $h := y \geq 0$ encodes that the drone must not enter the yellow area. (b) Comparison of the robustness of the safety constraint for a risk measure and chance constraints (CC), where values greater than zero indicate constraint satisfaction.

We are interested in risk-aware planning subject to complex spatio-temporal mission specifications, such as that a safety-critical area must not be entered *until* the localization uncertainty is sufficiently reduced or that a robot should move cautiously whenever it is uncertain about its position. The use of temporal logics as a specification language in belief spaces opens a variety of opportunities for encoding such desired behaviors. Existing works typically only reason about either spatio-temporal specifications [6]–[8] interpreted over known states, or spatial, but not temporal, aspects in belief space [1], [9], [10]. Only [11] and [12] consider temporal logics in belief spaces, but rely on discrete abstractions and do not consider aspects of risk. Our work considers continuous BSP problems with risk-aware objectives expressed in RiSTL and offers a solution to synthesize trajectories for specifications among three different dimensions: The *spatial*, the *temporal* and the *probabilistic* one.

A. Contributions

This work contributes to the fields of planning under uncertainty and control synthesis under temporal logic specifications. The proposed approach

1) allows designers to express complex spatial properties with desired risk levels for arbitrary convex polygons,
2) employs user-defined temporal logic specifications to provide risk-aware control synthesis in belief space,
3) directly minimizes the risk of violation while implicitly providing active information gathering behavior.

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. This research has been carried out as part of the Vinnova Competence Center for Trustworthy Edge Computing Systems and Applications at KTH Royal Institute of Technology.

The authors are with the Division of Robotics, Perception and Learning, School of Electrical Engineering and Computer Science , KTH Royal Institute of Technology, Stockholm, Sweden and also affiliated with Digital Futures. Mail addresses: {vahs, pek2, tumova} @kth.se

B. Related Work

Most relevant to our approach are classic works on Gaussian BSP, Signal Temporal Logic (STL) based synthesis and risk measures for quantification of severity of violations.

1) Gaussian BSP: In its most general form, a BSP problem can be formulated as a Partially Observable Markov Decision Process (POMDP) which is known to be difficult to solve [13], [14]. One common approach to make POMDPs tractable is to assume Gaussian parameterizations of the belief. The belief space is simplified to the space of possible mean vectors and covariance matrices, thus reducing the dimensionality significantly. Pioneering works on Gaussian BSP are based on stochastic optimal control and are able to solve reach-avoid problems under dynamic feasibility constraints [9], [10], [15]. Although these works are suitable for active information gathering tasks, they cannot be directly applied to complex temporal tasks.

2) STL Control Synthesis: A large body of work on solving planning problems under STL specifications exists [6], [16], [17] but only very few reason about state uncertainty. The control synthesis problem for STL specifications under uncertainty is addressed in [18], [19], yet exact state knowledge is assumed while the system is subject to stochastic disturbances. The works [11] and [12] are closest related to our approach which propose control methods in Gaussian belief space under temporal logic specifications. In contrast to these works, we do not rely on discrete transition systems but instead plan directly in the continuous belief space using a trajectory optimization scheme. Furthermore, as opposed to [11] we do not need to explicitly define information gathering objectives but they are considered implicitly by incorporating the risk of violating specifications.

In [20], the authors address the problem of temporal logic control of stochastic dynamical systems under consideration of risk. However, their approach is limited to linear systems and cannot be applied to Gaussian belief spaces. Thus, the problem of active information gathering cannot be addressed.

3) Risk Formulations: Measures of risk enable us to reason about specification violation in the presence of uncertainty. Chance constraints are widely used to account for the risk of violating a specification by assigning an upper bound on the probability of the occurrence of an undesired event, e.g. collision with an obstacle [3]–[5], [21], [22]. In [4], chance constraints are formulated for collision avoidance between multiple robots and ellipsoidal obstacles by linearization of collision constraints. We build upon this idea to formulate risk-aware stay-out and stay-in objectives. In contrast to [4], our approach can distinguish whether specifications are slightly or considerably violated.

Recently, coherent measures of risk have gained a lot of attention in the robotics community as these satisfy desirable axioms [23]. The greatest advantage of risk measures over chance constraints, however, is that they also quantify the severity of violation which we will further discuss in Section II-A. Coherent risk measures have been investigated in several robotics applications, including autonomous driving [24], reinforcement learning [25] and control under STL specifications [18]. To the authors' best knowledge, the use of risk measures such as the Value-at-Risk as quantification of specification violation in Gaussian BSP problems has not been studied before.

II. RISK-AWARE PLANNING PROBLEM

A. Risk Measures

Given a scalar random variable $y \in \mathcal{Y}$ with $p(y)$ denoting its probability density function (pdf), a risk measure is a function that assigns a real value to the random variable, i.e., $\mathcal{R} : \mathcal{Y} \mapsto \mathbb{R}$. A safety constraint can be expressed as a function h of the random variable y in a way such that $h \geq 0$ implies that the safety constraint is satisfied and vice versa, $h < 0$ implies that it is violated. Due to the stochasticity in y the function h is also a random variable with pdf $p(h)$. To evaluate tail events of a random variable, we use the Value-at-Risk (VaR) [26].

Definition 1. *The* VaR *of a scalar random variable $y \in \mathbb{R}$ with pdf $p(y)$ at level $\alpha \in (0, 1]$ is the $(1-\alpha)$ quantile, i.e.*

$$\mathcal{R}(y) = \mathrm{VaR}_\alpha(y) = \inf_{t \in \mathbb{R}} \{t \mid \Pr(y \leq t) \geq 1 - \alpha\} \quad (1)$$

The VaR is fundamentally different from the expected value $\mathbb{E}[h]$ since it considers the worst $1-\alpha$ quantile of the distribution. Such a tail-event risk measure is well suited for safety-critical applications since, as opposed to the expected value, it takes into account the variance of the distribution.

Example 1. *Consider the drone with an uncertain position in 2D in Fig. 1(a). We want to ensure that the robot does not enter a hazardous area, which can be expressed as safety constraint $h := y \geq 0$. Two different trajectories $\mathcal{T}_1$ and $\mathcal{T}_2$ are shown where $\mathcal{T}_2$ is obtained through BSP. In Fig. 1(b), the* VaR *is used to compare the robustness of a risk measure to a chance constraint formulation where values greater than zero indicate satisfaction of the safety constraint. It can be seen that chance constraints are not able to distinguish the two trajectories, whereas the risk robustness increases strongly, both, as an effect of increasing the distance y and the uncertainty reduction through BSP.*

The VaR can be used to formulate constraints that are qualitatively equivalent to chance constraints. Specifically, $\Pr[h < 0] \leq \alpha$ can be reformulated as $\mathrm{VaR}_\alpha(h) \geq 0$. However, VaR offers more gradient information regarding the extent to which the safety constraint is satisfied or violated. In contrast, the gradient quickly converges to zero for chance constraints. In turn, VaR is particularly useful for optimization-based motion planning frameworks where we reason about the severity of violations.

B. Risk Signal Temporal Logic (RiSTL)

RiSTL was first introduced in [18] and extends classic STL [27] with predicates that include uncertainty in the form of an external static random variable $\boldsymbol{X}$, namely risk predicates p^R and chance predicates p^C. A risk predicate is defined as

$$p^R(\boldsymbol{x}(t), \boldsymbol{X}) = \begin{cases} \top & \text{if } \mathcal{R}(-h(\boldsymbol{x}(t), \boldsymbol{X})) \leq \gamma \\ \bot & \text{otherwise} \end{cases} \quad (2)$$

where $\mathcal{R}(\cdot)$ is a suitable risk measure and γ denotes an acceptable risk level. Chance predicates are defined analogously with a predicate function $\Pr[h(\boldsymbol{x}(t), \boldsymbol{X}) \geq 0)] \geq \delta$.

RiSTL is interpreted over continuous deterministic signals $\boldsymbol{x}(t)$ and one of its major advantages is that it provides a measure of robustness. Robustness is a function to evaluate a RiSTL formula on $\boldsymbol{x}(t)$ that intuitively grows with a growing margin to violating the formula.

C. Problem Setting

We focus on mobile robot planning where the state of a robot cannot be observed directly. Let $\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{U} \subseteq \mathbb{R}^m$ and $\mathcal{Z} \subseteq \mathbb{R}^\ell$ be the state, control and measurement space, respectively. We assume we are given (potentially nonlinear) discrete-time stochastic motion and observation models of the form

$$\boldsymbol{x}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{u}_t) + \boldsymbol{w}_t, \qquad \boldsymbol{w}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}_t) \tag{3}$$

$$\boldsymbol{z}_t = \boldsymbol{\ell}(\boldsymbol{x}_t) + \boldsymbol{v}_t, \qquad \boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{V}_t) \tag{4}$$

where $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{u} \in \mathcal{U}$, $\boldsymbol{z} \in \mathcal{Z}$ and $\boldsymbol{w}, \boldsymbol{v}$ are the motion and observation noise, respectively. The robot operates in a 2D environment that includes convex polygon goal regions $\mathcal{G}$, obstacles $\mathcal{O}$ or any other regions of interest. We want to define risk-aware spatial specifications over these arbitrarily shaped convex polygon regions, such as that certain regions should be avoided or reached over time with high certainty.

Overall, we aim to solve spatio-temporal tasks for (3)-(4) under consideration of the risk of violating spatial specifications. To this end, we need to address two challenges: 1) to define robustness of a risk spatio-temporal logic over belief trajectories; and 2) to plan the robot's motion such that its mission can be satisfied to the maximal possible extent.

Problem 1. *Given dynamics as in Eq.* (3)-(4) *and a mission specification φ expressed in RiSTL over beliefs $\mathbf{b}$, find a belief trajectory $\mathcal{T} = (\boldsymbol{b}_0, \boldsymbol{u_0}), \cdots, (\boldsymbol{b}_N, \boldsymbol{u}_N)$ which maximizes the robustness of φ.*

III. RISK-AWARE SPATIO TEMPORAL PLANNING

To solve Problem 1, we address four different subproblems. First, we model the belief dynamics of the system given in (3)-(4). Second, we modify the syntax and the quantitative semantics of RiSTL such that we can express predicates over belief states. Third, we propose a method to express risk-aware stay-in and stay-out objectives for convex polygons to define desired spatial properties. These three steps allow us to formulate an optimization-based method to obtain belief trajectories in continuous space that maximize the robustness of the given specification.

A. Belief Dynamics

Ideally, we would use the Bayes filter to capture the exact belief evolution. However, exact belief calculations only exist in specialized cases [28]. Therefore, we consider a Gaussian parameterization of the state pdf which has been shown to be a useful assumption, especially if the true probability distribution is unimodal [28]. Thus, $\boldsymbol{x}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ follows a Gaussian distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^T \in \mathbb{R}^{n \times n}$. Specifically, the belief state $\boldsymbol{b} = [\boldsymbol{\mu}, \mathrm{vec}(\boldsymbol{\Sigma})]^T \in \mathbb{R}^{(n^2+3n)/2}$ is comprised of the mean vector and the upper triangular matrix of $\boldsymbol{\Sigma}$ written in vector form.

Given an initial belief $\boldsymbol{b}_0$, the belief state is recursively updated using the Extended Kalman Filter (EKF) [1]

$$\bar{\boldsymbol{\mu}}_{t+1} = \boldsymbol{f}(\boldsymbol{\mu}_t, \boldsymbol{u}_t) \tag{5}$$

$$\boldsymbol{\mu}_{t+1} = \bar{\boldsymbol{\mu}}_{t+1} + \boldsymbol{K}_t(\boldsymbol{z}_{t+1} - \boldsymbol{\ell}(\bar{\boldsymbol{\mu}}_{t+1})) \tag{6}$$

$$\boldsymbol{\Sigma}_{t+1} = (\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{H}_t)\boldsymbol{\Gamma}_t \tag{7}$$

where

$$\boldsymbol{K}_t = \boldsymbol{\Gamma}_t \boldsymbol{H}_t^T \left(\boldsymbol{H}_t \boldsymbol{\Gamma}_t \boldsymbol{H}_t^T + \boldsymbol{V}_t\right)^{-1} \tag{8}$$

$$\boldsymbol{\Gamma}_t = \boldsymbol{A}_t \boldsymbol{\Sigma}_t \boldsymbol{A}_t^T + \boldsymbol{W}_t, \tag{9}$$

$$\boldsymbol{A}_t = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{\mu}_t, \boldsymbol{u}_t), \qquad \boldsymbol{H}_t = \frac{\partial \boldsymbol{\ell}}{\partial \boldsymbol{x}}(\bar{\boldsymbol{\mu}}_{t+1}). \tag{10}$$

The resulting belief update can be written in compact form by combining (5)-(10) into a nonlinear dynamics equation. Note that the update of the mean vector in Eq. (6) depends on the obtained measurement at time $t+1$. Since future measurements are unknown at planning time t, the belief dynamics are stochastic.

Assumption 1. *We assume maximum likelihood observations (MLO) which states that future measurements are equal to the predicted ones $\boldsymbol{z}_{t+1} - \boldsymbol{h}(\bar{\boldsymbol{\mu}}_{t+1}, \boldsymbol{0}) = \boldsymbol{0}$.*

The relaxation of this assumption remains as a future work. Exploiting the MLO assumption yields the deterministic belief dynamics

$$\boldsymbol{b}_{t+1} = \boldsymbol{\Phi}(\boldsymbol{b}_t, \boldsymbol{u}_t) = \begin{bmatrix} \boldsymbol{f}(\boldsymbol{\mu}_t, \boldsymbol{u}_t) \\ \mathrm{vec}\left((\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{H}_t)\boldsymbol{\Gamma}_t\right) \end{bmatrix} \tag{11}$$

which can be used to predict how the belief evolves over time. In our approach, we leverage the deterministic belief dynamics as a constraint in a trajectory optimization problem to enable active information gathering behavior.

B. RiSTL in Belief Space

An important observation of RiSTL predicates as formulated in Eq. (2) is that they cannot be interpreted over belief states since static external uncertainty is assumed in the original formulation. In our problem the uncertainty is included in the state and changes over time.

We focus on planning problems over finite time horizons without explicit time bounds that still allow us to capture interesting spatio-temporal tasks and constraints (see Sec. IV for examples). The fragment of RiSTL we consider is inductively defined as

$$\varphi ::= \top \mid p^R \mid \neg\varphi \mid \varphi' \wedge \varphi'' \mid \varphi' U \varphi'' \tag{12}$$

where φ' and φ'' are RiSTL formulas and U is the until operator. Other common operators such as disjunctions or the temporal operators F (eventually) and G (always) can be derived from Eq. (12) [18].

We define our risk predicates over beliefs as

$$p^R(\boldsymbol{b}(t)) = \begin{cases} \top & \text{if } \mathcal{R}(-h(\boldsymbol{x}(t))) \leq \gamma \\ \bot & \text{otherwise} \end{cases} \tag{13}$$

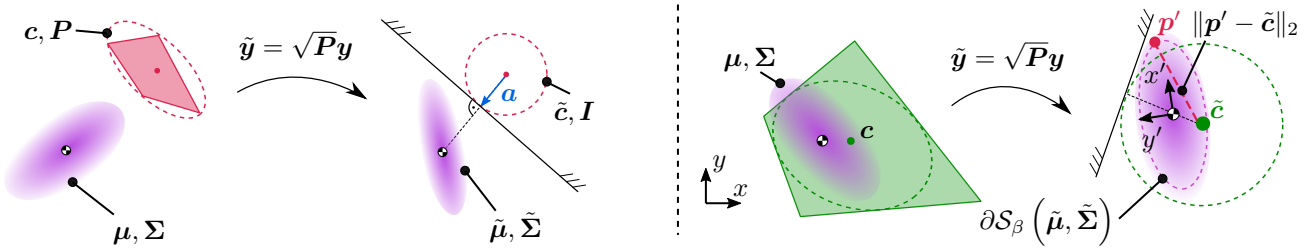

Fig. 2. Illustration of the risk calculations for stay-out (**left**) and stay-in (**right**) objectives based on ellipsoidal approximations. Transformed quantities are denoted by $\tilde{\cdot}$. Halfspaces are obtained by linearization around the transformed mean $\tilde{\boldsymbol{\mu}}$.

where $h(\boldsymbol{x}(t))$ is a predicate function with pdf $p(h)$ and, thus, $\mathcal{R}(-h(\boldsymbol{x}(t)))$ is a function of the belief $\boldsymbol{b}(t)$.

The main benefit of using RiSTL in belief space is that we can define the quantitative semantics (or robustness values) which enables us to reason about how much a given specification is satisfied or violated over a finite belief trajectory $\boldsymbol{b}_{t:t+k} = \boldsymbol{b}_t \boldsymbol{b}_{t+1} \cdots \boldsymbol{b}_{t+k}$.

Definition 2. *We define the quantitative semantics of a given formula ρ^{φ} on a belief trajectory $\boldsymbol{b}_{t:t+k}$ as*

$$
\begin{aligned}
\rho^{\top}(\boldsymbol{b}_{t:t+k}) &= \infty \\
\rho^{p^R}(\boldsymbol{b}_{t:t+k}) &= \gamma - \mathcal{R}(-h(\boldsymbol{b}_t)) \\
\rho^{\neg\varphi}(\boldsymbol{b}_{t:t+k}) &= -\rho^{\varphi}(\boldsymbol{b}_{t:t+k}) \\
\rho^{\varphi' \wedge \varphi''}(\boldsymbol{b}_{t:t+k}) &= \min\left(\rho^{\varphi'}(\boldsymbol{b}_{t:t+k}), \rho^{\varphi''}(\boldsymbol{b}_{t:t+k})\right) \\
\rho^{\varphi' \vee \varphi''}(\boldsymbol{b}_{t:t+k}) &= \max\left(\rho^{\varphi'}(\boldsymbol{b}_{t:t+k}), \rho^{\varphi''}(\boldsymbol{b}_{t:t+k})\right) \\
\rho^{\varphi' U \varphi''}(\boldsymbol{b}_{t:t+k}) &= \max_{t' \in [t,t+k]} \Big(\min\Big(\rho^{\varphi'}(\boldsymbol{b}_{t':t+k}), \\
&\qquad \min_{t'' \in [t,t']} \left(\rho^{\varphi''}(\boldsymbol{b}_{t'':t+k})\right)\Big)\Big) \\
\rho^{F\varphi}(\boldsymbol{b}_{t:t+k}) &= \max_{t' \in [t,t+k]} \left(\rho^{\varphi}(\boldsymbol{b}_{t':t+k})\right) \\
\rho^{G\varphi}(\boldsymbol{b}_{t:t+k}) &= \min_{t' \in [t,t+k]} \left(\rho^{\varphi}(\boldsymbol{b}_{t':t+k})\right).
\end{aligned}
$$

The quantitative semantics can be utilized to check whether a belief trajectory satisfies a given specification since $\rho^{\varphi}(\boldsymbol{b}_{t:t+k}) \geq 0 \Leftrightarrow \boldsymbol{b}_{t:t+k} \models \varphi$.

C. Risk Formulations for Polygon Regions

We can leverage the gradient of the robustness value in an optimization-based framework to find feasible solutions to Problem 1 by maximizing the robustness. To apply efficient gradient-based optimization techniques, we need to be able to express our risk predicates in Eq. (13) in closed form. However, a closed form expression to $\mathcal{R}(-h(\boldsymbol{x}(t)))$ exists only in rare cases when, e.g. h is linear. Therefore, this section focuses on deriving closed form risk predicates for stay-in and stay-out objectives for 2D convex polygons. Specifically, we want to find overapproximations for the risk of the signed distance $\mathrm{dist}(\boldsymbol{s}, \mathcal{P})$ between a point $\boldsymbol{s}$ and a polygon $\mathcal{P}$.

a) Stay-out objectives: We use minimum volume enclosing ellipsoids of the polygon region of interest $\mathcal{P}$ in the form of

$$
\mathcal{E} = \left\{ \boldsymbol{s} \in \mathbb{R}^2 \mid (\boldsymbol{s} - \boldsymbol{c})^T \boldsymbol{P} (\boldsymbol{s} - \boldsymbol{c}) \leq 1 \right\} \tag{14}
$$

where $\boldsymbol{P} = \boldsymbol{P}^T$ and $\boldsymbol{c}$ is the ellipsoid's center. To obtain a closed form expression for the risk of being inside an ellipsoid, we define the affine transformation $\tilde{\boldsymbol{y}} = \sqrt{\boldsymbol{P}}\boldsymbol{y}$ which maps the ellipsoid to a unit circle as illustrated in Fig. 2. We apply this transformation to the ellipsoid, $\tilde{\boldsymbol{c}} = \sqrt{\boldsymbol{P}}\boldsymbol{c}$, as well as the uncertain 2D position of the robot $\boldsymbol{s} = [p_x, p_y] \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$ obtained by marginalization of the belief $\boldsymbol{b}$ [28]. The transformed belief $\tilde{\boldsymbol{s}}$ is still Gaussian with mean and covariance

$$
\tilde{\boldsymbol{s}} \sim \mathcal{N}\left(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}\right) = \mathcal{N}\left(\sqrt{\boldsymbol{P}}\boldsymbol{\mu}_s, \sqrt{\boldsymbol{P}}^T \boldsymbol{\Sigma}_s \sqrt{\boldsymbol{P}}\right). \tag{15}
$$

To ensure that that the Gaussian distributed point $\boldsymbol{s}$ is outside the ellipsoid $\mathcal{E}$ (and, thus, outside $\mathcal{P}$), the distance in the transformed space should be greater than one which results in a stochastic predicate function of the form

$$
h^{\mathcal{P}}_{out}(\tilde{\boldsymbol{s}}) = \|\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{c}}\|_2 - 1. \tag{16}
$$

However, $p(h^{\mathcal{P}}_{out})$ is non-Gaussian due to the nonlinearity. Thus, it is generally difficult to obtain closed form expressions for risk measures. Therefore, we linearize the nonlinear predicate around the current transformed mean $\tilde{\boldsymbol{\mu}}$ which yields the linear stochastic predicate function

$$
\tilde{h}^{\mathcal{P}}_{out}(\tilde{\boldsymbol{s}}) = \boldsymbol{q}^T (\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{c}}) - 1 \tag{17}
$$

which represents a halfspace with $\boldsymbol{q} = (\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{c}}) / \|\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{c}}\|_2$, see Fig. 2. Consequently, the linearized predicate function $\tilde{h}^{\mathcal{P}}_{out}$ is also Gaussian distributed with pdf $p(\tilde{h}^{\mathcal{P}}_{out}) = \mathcal{N}(\mu_h, \sigma_h)$ where $\mu_h = \boldsymbol{q}^T (\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{c}}) - 1$ and $\sigma_h = \boldsymbol{q}^T \tilde{\boldsymbol{\Sigma}} \boldsymbol{q}$.

On the example of the VaR, we formulate the risk of a stay-out specification as follows. Given an ellipsoidal overapproximation $\mathcal{E}$ and a Gaussian distributed point $\boldsymbol{s}$, the VaR_α at level α can be obtained as

$$
\mathrm{VaR}_\alpha(-\tilde{h}^{\mathcal{P}}_{out}(\boldsymbol{s})) = \mu_h + \sqrt{2\sigma_h}\mathrm{erf}^{-1}(2\alpha - 1). \tag{18}
$$

Lemma 1. *The obtained risk $\mathcal{R}\left(-\tilde{h}^{\mathcal{P}}_{out}\right)$ for a stay-out objective is an overapproximation of the exact risk of $\mathcal{R}(-\mathrm{dist}(\boldsymbol{s}, \mathcal{P}))$.*

Proof sketch: Since $\mathcal{E}$ is a strict overapproximation of $\mathcal{P}$, i.e., $\mathcal{P} \subset \mathcal{E}$, and $\mathcal{E}$ is convex, it is ensured that the entire polygon $\mathcal{P}$ is inside the negative halfspace. Therefore, the linearized risk measure upper bounds the true risk.

b) Stay-in objectives: In contrast to stay-out objectives, we cannot upper bound the risk $\mathcal{R}(\mathrm{dist}(\boldsymbol{s}, \mathcal{P}))$ by a halfspace constraint for a stay-in objective. This is illustrated in the right of Fig. 2, where naively applying the same approach

yields the halfspace shown in black. Although the Gaussian belief satisfies the halfspace constraint, part of it is outside the transformed ellipse, as shown by $\boldsymbol{p}'$, and thus, violating the stay-in objective.

To overcome this issue, we propose a distance-based stay-in objective to ensure that a desired level of probability mass lies within the unit circle. The confidence set for a multivariate Gaussian that contains β % of the probability mass is given by

$$\mathcal{S}_\beta = \left\{ \boldsymbol{s} \in \mathbb{R}^2 \mid (\boldsymbol{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{s}-\boldsymbol{\mu}) \leq \chi_2^2(\beta) \right\} \quad (19)$$

where χ_2^2 is the chi-squared distribution with two degrees of freedom. Our idea is to 1) find the point $\boldsymbol{p}' \in \partial\mathcal{S}_\beta(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})$ on the boundary of $\mathcal{S}_\beta(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})$ with maximum distance to circle center $\tilde{\boldsymbol{c}}$, and 2) constrain it to be within the unit circle and thus respecting the stay-in objective.

To derive a closed form expression for $\boldsymbol{p}'$, we first compute the largest ellipse contained in a stay-in polygon $\mathcal{P}$ and use the same transformation as in Eq. (15) to map it to a unit circle. We obtain the normal form $\left({}^{x'}/_{a}\right)^2 + \left({}^{y'}/_{b}\right)^2 = 1$ of $\partial\mathcal{S}_\beta(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})$ through an eigendecomposition, i.e., $\chi_2^2(\beta)\tilde{\boldsymbol{\Sigma}} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T$, where $\boldsymbol{Q}$ is a rotation matrix and $\boldsymbol{\Lambda}$ is a diagonal matrix [4]. This ellipse is described in a local coordinate frame K' attached to its center $\tilde{\boldsymbol{\mu}}$, see Fig. 2. Note, that closed form expressions for a and b can be found in 2D.

To find $\boldsymbol{p}'$ with the maximum distance to $\tilde{\boldsymbol{c}}$, we first need to transform $\tilde{\boldsymbol{c}}$ to frame K' as $\tilde{\boldsymbol{c}}' = \boldsymbol{Q}^T(\tilde{\boldsymbol{c}} - \tilde{\boldsymbol{\mu}})$. We maximize $\|\boldsymbol{p}' - \tilde{\boldsymbol{c}}'\|_2$ by realizing that the dot product

$$\boldsymbol{t}(\boldsymbol{p}')^T (\boldsymbol{p}' - \tilde{\boldsymbol{c}}') = 0 \quad (20)$$

vanishes between the tangent $\boldsymbol{t}(\boldsymbol{p}')$ on the ellipse and the connecting line to $\boldsymbol{p}'$. Equation (20) results in a quartic equation which can also be solved analytically for $\boldsymbol{p}'$ using e.g. Ferrari's method [29].

Finally, we define the risk of a stay-in objective as

$$\tilde{h}_{in}^{\mathcal{P}}(\boldsymbol{b}) := \|\boldsymbol{p}' - \tilde{\boldsymbol{c}}'\|_2 - 1. \quad (21)$$

Equation (21) is our definition of the risk of a stay-in objective which is formally not the same as a risk predicate defined in Eq. (13) since we are not applying any risk measure to a random variable. However, our risk definition captures similar properties as stated in the following Lemma.

Lemma 2. *If $\tilde{h}_{in}^{\mathcal{P}} \geq 0$ for a stay-in objective with $\beta > 0$, then $\mathrm{VaR}_\beta(-\mathrm{dist}(\boldsymbol{s}, \mathcal{P})) \leq 0$.*

Proof sketch: Since the underapproximating ellipse is entirely contained in $\mathcal{P}$, and $\tilde{h}_{in}^{\mathcal{P}} \geq 0$ implies that the furthest point on $\partial\mathcal{S}_\beta$ is contained within the unit circle, it follows that at least β % of the probability mass is inside the polygon and, thus, $\mathrm{VaR}_\beta(-\mathrm{dist}(\boldsymbol{s}, \mathcal{P})) \leq 0$.

D. Trajectory Synthesis

The last ingredient to our proposed problem solution is a framework for generating belief trajectories $\boldsymbol{b}_{0:N}$ and the corresponding controls $\boldsymbol{u}_{0:N-1}$ that maximize the robustness of a given specification φ.

Given a belief space RiSTL specification that contains e.g. risk-aware stay-in or stay-out predicates as defined before, we propose a trajectory optimization scheme in which the satisfaction of the formula φ is maximized, reading

$$\begin{aligned} \max_{\boldsymbol{u}_{0:N-1}, \boldsymbol{b}_{0:N}} \quad & \rho^\varphi(\boldsymbol{b}_{0:N}) \\ \text{s.t.} \quad & \boldsymbol{b}_{t+1} = \boldsymbol{\Phi}(\boldsymbol{b}_t, \boldsymbol{u}_t) && \forall t = 0, \cdots, N-1 \\ & \boldsymbol{b}_t \in \mathcal{B}, \boldsymbol{u}_t \in \mathcal{U} && \forall t = 0, \cdots, N \\ & \boldsymbol{b}_0 = \boldsymbol{b}(t_0). \end{aligned} \quad (22)$$

If an optimized solution generates a belief trajectory such that $\rho^\varphi(\boldsymbol{b}_{0:N}) \geq 0$, then it is ensured that the specification is satisfied, i.e. $\boldsymbol{b}_{0:N} \models \varphi$.

The main issue with the optimization problem as formulated in Eq. (22) is that the objective function $\rho^\varphi(\boldsymbol{b}_{0:N})$ is non-convex and especially non-smooth due to Def. 2. The $\min(\cdot)$ and $\max(\cdot)$ operators can have a discontinuous gradient which is unsuitable for optimization-based approaches. For this reason, we leverage smooth approximations of the $\min(\cdot)$ and $\max(\cdot)$ operators such that an approximate but smooth robustness measure $\tilde{\rho}^\varphi$ can be obtained. Several approximations have been proposed [30]–[32] but we make use of the definition in [31]

$$\widetilde{\min_{i=1,\ldots,m}} x_i = -\frac{1}{\gamma} \ln\left(\sum_{i=1}^{n} \exp(-\gamma x_i)\right) \quad (23)$$

$$\widetilde{\max_{i=1,\ldots,m}} x_i = \frac{\sum_{i=1}^{n} x_i \exp(\eta x_i)}{\sum_{i=1}^{n} \exp(\eta x_i)} \quad (24)$$

where γ and η are tuning parameters. This choice of approximation is particularly useful as both approximations are strict underapproximations of the true operators. However, as $\gamma, \eta \to \infty$, the approximations approach the original operators. Replacing the original operators by the approximations in the quantitative semantics in Def. 2 yields the approximated robustness values $\tilde{\rho}^\varphi$ which are strictly less than the original robustness values. Consequently, it is guaranteed that $\tilde{\rho}^\varphi(\boldsymbol{b}_{0:N}) \geq 0$ implies $\rho^\varphi(\boldsymbol{b}_{0:N}) \geq 0$ and, thus, the specification is satisfied. Hence, we replace the original objective function in Eq. (22) by its smooth underapproximation, in order to solve the optimization problem by which we find a solution to Problem 1.

E. Discussion

Although the resulting problem is smooth, it is still a non-convex nonlinear program (NLP) but smooth optimization techniques such as interior-point methods can be applied. The existence of a solution, however, is not guaranteed and finding assumptions for completeness of the proposed framework remains future work.

Additionally, we introduce conservatism as a consequence of our risk-aware predicates since we are using ellipsoidal over- and underapproximations of polygon regions. However, if a satisfying solution is found, it is sound, and if a solution with negative robustness ρ^φ is found, it still might satisfy the true risk constraints which can be checked in a post-processing step.

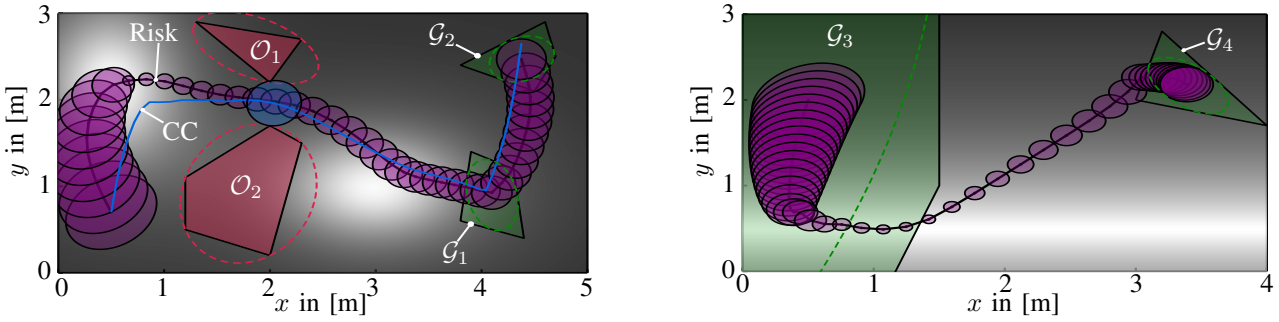

Fig. 3. Two exemplary trajectories in the light-dark domain where light colors represent areas with low measurement noise. Both trajectories satisfy the RiSTL specifications φ_1 (**left**) and φ_2 (**right**). The 90 % confidence ellipses about the robot's position are shown in purple, stay-in and stay-out regions are colored in green and red, respectively. The blue trajectory is shown for comparison where chance constraints (CC) are used for stay-out objectives.

IV. Simulation Study

We validate our approach in simulations for two different missions with increasing specification complexity. We show that maximizing a risk-aware robustness objective and active information gathering behavior can be achieved simultaneously with improved uncertainty reduction compared to chance constraints. Simulations are carried out on an Intel i9-10900K processor and 32 GB of RAM. The optimization problem is defined in Julia and solved with the open source nonlinear optimization software IPOPT.

We use a 2D system with state $\boldsymbol{x} = [p_x, p_y]^T$ and stochastic dynamics and observations

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \delta t\ \boldsymbol{u}_t + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma_w \boldsymbol{I}) \tag{25}$$

$$\boldsymbol{z}_t = \boldsymbol{x}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma_v(\boldsymbol{x}_t)\, \boldsymbol{I}), \tag{26}$$

where δt is the time step. In our simulation, the sensing noise $\boldsymbol{v}$ depends on the position of the robot in 2D space which is commonly known as the Light-dark problem [10]. In this problem, the robot can only localize itself accurately in light areas which may represent antenna stations sending out radio signals, as shown in Fig. 3.

The first problem is a risk-aware reach-avoid specification with two sequential goals which can be expressed as

$$\varphi_1 = \bigwedge_{i=1,2} G(\mathcal{R}(-\tilde{h}_{out}^{\mathcal{O}_i}) \leq \gamma) \wedge F(\tilde{h}_{in}^{\mathcal{G}_i} \geq 0.). \tag{27}$$

where $\mathcal{O}_i$ and $\mathcal{G}_i$ represent the obstacle and goal polygon regions, respectively. We apply the $\mathrm{VaR}_{0.1}$ as risk measure and set $\gamma = 0$. Further, we use an RRT* planner to generate a collision free initial guess for $\boldsymbol{\mu}_{0:N}$ to one of the goal regions. The optimized belief trajectory with a robustness value $\rho^{\varphi_1}(\boldsymbol{b}_{0:N}) = 0.122$ is shown in the left of Fig. 3. Active information gathering behavior can be observed although we do not explicitly formulate such objectives. This is due to the structure of our risk-aware predicate functions. Maximizing the robustness includes maximizing predicates which, in the case of stay-in and stay-out predicates, can be done in two ways. Either, the distance to a region is maximized/minimized (spatial robustness), or the uncertainty is reduced (active information gathering). Furthermore, for both reach-objectives it can be observed that the planned trajectory passes the center of the ellipsoids as the robustness of a stay-in objective is maximized at that point.

We also show a robustness maximizing trajectory in the case of chance constraints (CC) for stay-out objectives. The obtained robustness is greater than zero, indicating mission satisfaction, but active information gathering behavior is reduced since CCs are saturated. Consequently, the uncertainty at the critical narrow passage is 3.28 times higher than the one for the risk-aware trajectory. Here the uncertainty is quantified as $\mathrm{tr}(\boldsymbol{\Sigma})$ where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

Second, we investigate a problem with a more complex specification that cannot be expressed as classic reach-avoid property. In this scenario, we never want to leave an area $\mathcal{G}_3$ until the uncertainty is sufficiently low, always if the uncertainty is above a given threshold, the robot should move cautiously by slowing down and we eventually want to reach a goal region $\mathcal{G}_4$. This can be formally specified as

$$\begin{aligned}\varphi_2 = G(\mathrm{tr}(\boldsymbol{\Sigma}) \geq \varepsilon \implies \|\boldsymbol{u}\|_2 \leq u_{max}) \wedge \\ F(\tilde{h}_{in}^{\mathcal{G}_4} \geq 0) \wedge (\tilde{h}_{in}^{\mathcal{G}_3} \geq 0) U(\mathrm{tr}(\boldsymbol{\Sigma}) \leq \varepsilon),\end{aligned} \tag{28}$$

where u_{max} is an upper bound on the controls. The optimized trajectory with a robustness of $\rho^{\varphi_2} = 0.03$ is shown in the right of Fig. 3. We find a trajectory satisfying the given specification as it can, for example, be seen that the robot moves slowly in the presence of large uncertainties.

We found that the median planning time for a single trajectory optimization with horizon $N = 35$ takes $\bar{t}_1 = 5.2$ s and $\bar{t}_2 = 14.3$ s for specifications φ_1 and φ_2, respectively. The median was obtained over 10 simulation runs.

V. Conclusions and Future Work

Our work enables trajectory synthesis in which we can minimize risk objectives and achieve active information gathering simultaneously by combining BSP and risk measures. We allow for complex spatio-temporal mission specifications in belief space by defining RiSTL formulas over Gaussian beliefs. In experiments, we used our approach to synthesize trajectories for various tasks that include predicates over the state uncertainty and, thus, go beyond classic reach-avoid objectives.

So far, we try to solve the full nonlinear and non-convex optimization problem. In the future, we plan to formulate the problem in a way such that differential dynamic programming techniques can be applied which would increase the saclability to higher dimensional belief spaces and also make it possible to relax the MLO assumption.

REFERENCES

[1] J. Van Den Berg, S. Patil, and R. Alterovitz, "Motion planning under uncertainty using iterative local optimization in belief space," *The International Journal of Robotics Research*, vol. 31, no. 11, pp. 1263–1278, 2012.

[2] S. M. Lavalle and R. Sharma, "On motion planning in changing, partially predictable environments," *The International Journal of Robotics Research*, vol. 16, no. 6, pp. 775–805, 1997.

[3] F. Gavilan, R. Vazquez, and E. F. Camacho, "Chance-constrained model predictive control for spacecraft rendezvous with disturbance estimation," *Control Engineering Practice*, vol. 20, no. 2, pp. 111–122, 2012.

[4] H. Zhu and J. Alonso-Mora, "Chance-constrained collision avoidance for mavs in dynamic environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 776–783, 2019.

[5] L. Blackmore, M. Ono, and B. C. Williams, "Chance-constrained optimal path planning with obstacles," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1080–1094, 2011.

[6] V. Raman, A. Donzé, M. Maasoumy, R. M. Murray, A. Sangiovanni-Vincentelli, and S. A. Seshia, "Model predictive control with signal temporal logic specifications," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 81–87.

[7] I. Haghighi, N. Mehdipour, E. Bartocci, and C. Belta, "Control from signal temporal logic specifications with smooth cumulative quantitative semantics," in *IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 4361–4366.

[8] Y. V. Pant, H. Yin, M. Arcak, and S. A. Seshia, "Co-design of control and planning for multi-rotor uavs with signal temporal logic specifications," in *American Control Conference (ACC)*, 2021, pp. 4209–4216.

[9] J. Van Den Berg, P. Abbeel, and K. Goldberg, "LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 895–913, 2011.

[10] R. Platt, R. Tedrake, L. P. Kaelbling, and T. Lozano-Perez, "Belief space planning assuming maximum likelihood observations," in *Robotics: Science and Systems*, 2010.

[11] C.-I. Vasile, K. Leahy, E. Cristofalo, A. Jones, M. Schwager, and C. Belta, "Control in belief space with temporal logic specifications," in *IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 7419–7424.

[12] R. R. da Silva, V. Kurtz, and H. Lin, "Active perception and control from prstl specifications," *arXiv preprint arXiv:2111.02226*, 2021.

[13] Z. N. Sunberg and M. J. Kochenderfer, "Online algorithms for pomdps with continuous state, action, and observation spaces," in *Twenty-Eighth International Conference on Automated Planning and Scheduling*, 2018.

[14] N. Ye, A. Somani, D. Hsu, and W. S. Lee, "DESPOT: Online POMDP planning with regularization," *Journal of Artificial Intelligence Research*, vol. 58, pp. 231–266, 2017.

[15] M. P. Vitus and C. J. Tomlin, "Closed-loop belief space planning for linear, gaussian systems," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2152–2159.

[16] V. Raman, A. Donzé, D. Sadigh, R. M. Murray, and S. A. Seshia, "Reactive synthesis from signal temporal logic specifications," in *Proceedings of the 18th international conference on hybrid systems: Computation and control*, 2015, pp. 239–248.

[17] L. Lindemann and D. V. Dimarogonas, "Control barrier functions for signal temporal logic tasks," *IEEE control systems letters*, vol. 3, no. 1, pp. 96–101, 2018.

[18] S. Safaoui, L. Lindemann, D. V. Dimarogonas, I. Shames, and T. H. Summers, "Control design for risk-based signal temporal logic specifications," *IEEE Control Systems Letters*, vol. 4, no. 4, p. 1000–1005, 2020.

[19] D. Sadigh and A. Kapoor, "Safe control under uncertainty with probabilistic signal temporal logic," in *Proceedings of Robotics: Science and Systems XII*, June 2016.

[20] S. Safaoui, L. Lindemann, I. Shames, and T. H. Summers, "Risk-bounded temporal logic control of continuous-time stochastic systems," *arXiv preprint arXiv:2204.04310*, 2022.

[21] A. Jasour and C. Lagoa, "Convex chance constrained model predictive control," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 6204–6209.

[22] L. Blackmore, M. Ono, A. Bektassov, and B. C. Williams, "A probabilistic particle-control approximation of chance-constrained stochastic predictive control," *IEEE transactions on Robotics*, vol. 26, no. 3, pp. 502–517, 2010.

[23] A. Majumdar and M. Pavone, "How should a robot assess risk? towards an axiomatic theory of risk in robotics," in *Robotics Research*. Springer, 2020, pp. 75–84.

[24] T. Nyberg, C. Pek, L. Dal Col, C. Norén, and J. Tumova, "Risk-aware motion planning for autonomous vehicles with safety specifications," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, 2021, pp. 1016–1023.

[25] A. Majumdar, S. Singh, A. Mandlekar, and M. Pavone, "Risk-sensitive inverse reinforcement learning via coherent risk models." in *Robotics: Science and Systems*, vol. 16, 2017, p. 117.

[26] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.

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

[28] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. Cambridge, Mass.: MIT Press, 2005.

[29] Y. Yang and Z. Zhou, "An analytic solution to wahba's problem," *Aerospace Science and Technology*, vol. 30, no. 1, pp. 46–49, 2013.

[30] N. Mehdipour, C.-I. Vasile, and C. Belta, "Arithmetic-geometric mean robustness for control from signal temporal logic specifications," in *American Control Conference (ACC)*. IEEE, 2019, pp. 1690–1695.

[31] Y. Gilpin, V. Kurtz, and H. Lin, "A smooth robustness measure of signal temporal logic for symbolic control," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 241–246, 2020.

[32] Y. V. Pant, H. Abbas, and R. Mangharam, "Smooth operator: Control using the smooth robustness of temporal logic," in *IEEE Conference on Control Technology and Applications (CCTA)*, 2017, pp. 1235–1240.