Automatica 94 (2018) 94–101

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# A dynamic game approach to distributionally robust safety specifications for stochastic systems☆

Insoon Yang

*Department of Electrical and Computer Engineering, Automation and Systems Research Institute, Seoul National University, Seoul, 08826, South Korea*

ARTICLE INFO

*Article history:*
Received 30 January 2017
Received in revised form 26 October 2017
Accepted 14 March 2018

*Keywords:*
Stochastic control
Stochastic systems
Verification
Safety analysis
Reachability
Game theory
Duality
Dynamic programming

ABSTRACT

This paper presents a new safety specification method that is robust against errors in the probability distribution of disturbances. Our proposed distributionally robust safe policy maximizes the probability of a system remaining in a desired set for all times, subject to the worst possible disturbance distribution in an ambiguity set. We propose a dynamic game formulation of constructing such policies and identify conditions under which a non-randomized Markov policy is optimal. Based on this existence result, we develop a practical design approach to safety-oriented stochastic controllers with limited information about disturbance distributions. However, an associated Bellman equation involves infinite-dimensional minimax optimization problems since the disturbance distribution may have a continuous density. To alleviate computational issues, we propose a duality-based reformulation method that converts the infinite-dimensional minimax problem into a semi-infinite program that can be solved using existing convergent algorithms. We prove that there is no duality gap, and that this approach thus preserves optimality. The results of numerical tests confirm that the proposed method is robust against distributional errors in disturbances, while a standard stochastic safety verification tool is not.



## 1. Introduction

For safety-critical systems subject to uncertain disturbances, reachability-based safety specification techniques have been used to compute the reachable sets and safe sets, which allow one to verify that a system is evolving within a safe range of operation and to synthesize controllers to satisfy safety constraints (e.g., Althoff, Le Guernic, and Krogh, 2011; Bertsekas and Rhodes, 1971; Cardaliaguet, Quincampoix, and Saint-Pierre, 1999; Chen, Herbert, Vashishtha, Bansal, and Tomlin, 2016; Ghaemi and Del Vecchio, 2014; Girard, 2005; Kurzhanski and Varaiya, 2002; Lygeros, Tomlin, and Sastry, 1999; Margellos and Lygeros, 2011; Mitchell, Bayen, and Tomlin, 2005; Rakovic, Kerrigan, Mayne, and Lygeros, 2006). These methods assume that disturbances lie in a compact set, and thus require information only about the support of disturbances. However, these techniques often produce conservative results as no additional information about uncertain disturbances is used. These deterministic methods are a natural choice when the data of disturbances are not continuously collected, and thus a reliable stochastic model is unavailable for them. Advances in sensing, communication, and computing technologies as well as statistical learning and estimation tools make it possible to shift this paradigm; sensors, data storage and computing infrastructure can now provide data to help estimate disturbance distributions. Stochastic reachability analysis tools are based on the assumption that the probability distribution of disturbances is available and can be used to reduce the conservativeness of their deterministic counterpart. However, this assumption is often restrictive in practice because obtaining an accurate distribution requires large-scale high-resolution sensor measurements over a long training period or multiple periods. Furthermore, the accuracy of the distribution obtained by computational methods is often unreliable as it is subject to the quality of the collected data, statistical learning or filtering methods, and prior knowledge. Thus, probabilistic safety specification tools can lead to the design of an unreliable controller that may violate safety constraints.

This paper aims to bridge the gap between the two methods by proposing a *distributionally robust safety specification* tool. Our approach assumes that the distribution of disturbances is not fully known but lies in a so-called *ambiguity set* of probability distributions. The proposed *distributionally robust safe policy* maximizes the probability of a system remaining within a desired set for all times subject to the worst possible disturbance distribution in the ambiguity set. Therefore, the probabilistic safe set of the closed-loop system is robust against distributional errors within

☆ Supported in part by NSF under ECCS 1708906 and CNS 1657100, and in part by Research Resettlement Fund for the new faculty of Seoul National University (0668-20180071). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Gurdal Arslan under the direction of Editor Ian R. Petersen.

*E-mail address:* insoonyang@snu.ac.kr.

https://doi.org/10.1016/j.automatica.2018.04.022
0005-1098/

the ambiguity set. We propose a dynamic game formulation of constructing distributionally robust safe policies and safe sets (in Section 2). Specifically, it is a two-player zero-sum dynamic game in which Player I selects a policy by which the controller can maximize the probability of safety, while (fictitious) Player II determines a strategy for the probability distribution of disturbances to minimize the same probability. Player II's action space is generally infinite dimensional since the disturbances may have a continuous density function. Therefore, an associated Bellman equation involves infinite-dimensional optimization problems that are computationally challenging. Furthermore, the existence of an optimal control policy is not guaranteed.

The contributions of this work are threefold. First, we characterize conditions for the existence and optimality of a non-randomized Markov control policy for Player I (in Section 3). This characterization helps greatly reduce the control strategy space we need to search for because it is enough to restrict our attention to non-randomized Markov policies. Second, we develop a design approach to a safety-oriented controller with limited information about disturbance distributions. This control method can be used to minimize another cost function while guaranteeing that the probability for a system being safe for all remaining stages is greater than or equal to a pre-specified threshold. Third, we propose a duality-based reformulation method for the Bellman equation in cases with moment uncertainty (in Section 4). We show that there is no duality gap in the inner minimization problem of the Bellman equation, which is an infinite-dimensional minimax problem. Using the strong duality result, we reformulate the Bellman equation as a semi-infinite program without sacrificing optimality. This reformulation alleviates the computational issue arising from the infinite dimensionality of the original Bellman equation because the reformulated Bellman equation can be solved via existing convergent algorithms for semi-infinite programs. The distributional robustness of the proposed tool is illustrated through examples (in Section 5).

We summarize related studies as follows. A probabilistic reachability tool using a Markov chain approximation has been proposed in Hu, Prandini, and Sastry (2005) and Prandini and Hu (2006). In Prajna, Jadbabaie, and Pappas (2007), barrier certificates are employed to calculate an upper bound of the probability that a system will reach a target set. Additionally, Mitchell and Templeton (2005) propose a toolbox that supports expectation-based reachability problems by extending the celebrated Hamilton–Jacobi–Isaacs reachability analysis (Mitchell et al., 2005). For discrete-time stochastic hybrid systems, an elegant dynamic programming approach has been proposed to compute the maximal probability of safety (Abate, Prandini, Lygeros, & Sastry, 2008). This method has been extended to stochastic reach–avoid problems (Summers & Lygeros, 2010), stochastic hybrid games (Ding, Kamgarpour, Summers, Abate, Lygeros, & Tomlin, 2013), and partially observable hybrid systems (Lesser & Oishi, 2014). However, all the aforementioned methods are based on the possibly restrictive assumption that the probability distribution of disturbances is completely known. This paper is also closely related to *distributionally robust control*; it minimizes the worst-case cost, assuming that the probability distribution of uncertain variables lies within an ambiguity set of distributions. A distributionally robust Markov decision process (MDP) formulation has recently been developed while focusing on finite-state, finite-action MDPs (Xu & Mannor, 2012; Yang, 2017a; Yu & Xu, 2016). For cases with moment uncertainty, Van Parys, Kuhn, Goulart, & Morari (2016) investigate linear feedback strategies in linear–quadratic settings with risk constraints and propose a semidefinite programming approach. We extend the theory of distributionally robust control to the case of continuous state spaces and apply it to safety specifications.

We use the following notation throughout the paper. Given a Borel space $X$, $\mathcal{B}(X)$ and $\mathcal{P}(X)$ represent its Borel $\sigma$-algebra and the set of Borel probability measures on $X$, respectively. The set $\mathbb{S}^l_+$ denotes the space of $l\times l$ symmetric positive semidefinite matrices. We also let $\mathcal{T} := \{0, 1, \ldots, T-1\}$ and $\bar{\mathcal{T}} := \{0, 1, \ldots, T\}$.

## 2. Distributionally robust safe sets and policies

Consider a discrete-time stochastic system of the form

$$x_{t+1} = f(x_t, u_t, w_t) \quad \forall t \in \mathcal{T}, \quad x_0 = \boldsymbol{x}, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^m$ is the control input, $w_t \in \mathbb{R}^l$ is the stochastic disturbance, and $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^n$ is a measurable function. We assume that the disturbance process $\{w_t\}_{t=0}^{T-1}$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and that $w_s$ and $w_t$ are independent for any $s \neq t$. As mentioned in Section 1, it is often difficult to obtain full information about the probability distribution $\mu_t$ of $w_t$. To mathematically model distributional ambiguity, we assume that $\mu_t$ is not fully known but contained in a so-called *ambiguity set* of distributions, denoted by $\mathbb{D}_t \subseteq \mathcal{P}(\mathbb{R}^l)$.

We now briefly discuss admissible control and disturbance distribution strategies. Let $H_t$ be the set of histories up to stage $t$, whose element takes the form $h_t := (x_0, u_0, w_0, \ldots, x_{t-1}, u_{t-1}, w_{t-1}, x_t)$.[1] The set of admissible control strategies is chosen as $\Pi := \{\pi := (\pi_0, \ldots, \pi_{T-1}) \mid \pi_t(\mathbb{U}(x_t)|h_t) = 1 \ \forall h_t \in H_t\}$, where $\pi_t$ is a stochastic kernel from $H_t$ to $\mathbb{R}^m$ and $\mathbb{U}(x_t)$ is the set of admissible actions given state $x_t$. Note that this strategy space is sufficiently broad to contain randomized non-Markov policies. Considering an adversarial player who chooses the disturbance's probability distribution $\mu_t$, the set of admissible disturbance distribution strategies is defined as $\Gamma := \{\gamma := (\gamma_0, \ldots, \gamma_{T-1}) \mid \gamma_t(\mathbb{D}_t|h^e_t) = 1 \ \forall h^e_t \in H^e_t\}$, where $H^e_t$ is a set of extended histories up to stage $t$, whose element is of the form $h^e_t := (x_0, u_0, w_0, \mu_0, \ldots, x_{t-1}, u_{t-1}, w_{t-1}, \mu_{t-1}, x_t, u_t)$. Note that the distributional constraints in the ambiguity set $\mathbb{D}_t$ is encoded in the strategy space $\Gamma$.

### 2.1. Distributionally robust safety specifications

Our goal is to compute the worst-case probability of a system remaining in a desired set for all times when the distribution of $w_t$ is not fully known but lies within an ambiguity set, $\mathbb{D}_t$. To formulate a concrete safety specification problem, we consider a desired set $A$ for safety, which is an arbitrary compact Borel set in the state space $\mathbb{R}^n$. We also introduce the following definitions:

**Definition 1** (*Probability of Safety*)**.** We define the probability that the system (1) is safe for all $t \in \bar{\mathcal{T}}$ given the strategy pair $(\pi, \gamma)$ and the initial value $\boldsymbol{x}$ as

$$P^{\text{safe}}_{\boldsymbol{x}}(\pi, \gamma; A) := \mathbb{P}^{\pi,\gamma}\{x_t \in A \ \ \forall t \in \bar{\mathcal{T}} \mid x_0 = \boldsymbol{x}\},$$

which we call the *probability of safety* for $A$. We also define the *probabilistic safe set* with probability $\alpha$ under $(\pi, \gamma)$ as $S_\alpha(\pi, \gamma; A) := \{\boldsymbol{x} \in \mathbb{R}^n \mid P^{\text{safe}}_{\boldsymbol{x}}(\pi, \gamma; A) \geq \alpha\}$.

This set contains all the initial states such that the probability that the system stays in the set $A$ is greater than or equal to $\alpha$ given the strategy pair $(\pi, \gamma)$. This definition generalizes the probabilistic safe set introduced in Abate et al. (2008) to the case

[1] To make it practically sound, we assume that the controller is unable to observe the disturbance's probability distribution. However, due to Theorem 1, all of our results and analyses remain valid even when $h_t$ includes $(\mu_0, \ldots, \mu_{t-1})$.