Unsafe Probabilities and Risk Contours for Stochastic Processes using Convex Optimization

Jared Miller, Matteo Tacchi, Didier Henrion, Mario Sznaier

Unsafe Probabilities and Risk Contours for Stochastic Processes using Convex Optimization

Jared Miller[1], Matteo Tacchi[2], Didier Henrion[3], Mario Sznaier[4]

January 2, 2024

Abstract

This paper proposes an algorithm to calculate the maximal probability of unsafety with respect to trajectories of a stochastic process and a hazard set. The unsafe probability estimation problem is cast as a primal-dual pair of infinite-dimensional linear programs in occupation measures and continuous functions. This convex relaxation is nonconservative (to the true probability of unsafety) under compactness and regularity conditions in dynamics. The continuous-function linear program is linked to existing probability-certifying barrier certificates of safety. Risk contours for initial conditions of the stochastic process may be generated by suitably modifying the objective of the continuous-function program, forming an interpretable and visual representation of stochastic safety for test initial conditions. All infinite-dimensional linear programs are truncated to finite dimension by the Moment-Sum-of-Squares hierarchy of semidefinite programs. Unsafe-probability estimation and risk contours are generated for example stochastic processes.

1 Introduction

This work performs risk analysis of stochastic processes by computing the probability that trajectories will reach an unsafe set X_u. Trajectories evolve over a maximal time horizon of T in a state space $X \subseteq \mathbb{R}^n$, beginning in an initial set $X_0 \subseteq X$ and possibly entering the unsafe set $X_u \subseteq X$. We consider stochastic processes described by a generator $\mathcal{L}$ (e.g., a Stochastic Differential Equation (SDE)), for which X_t is the time-dependent state probability distribution of the process starting from X_0 and terminating upon exit from X (with exit time distribution $\tau_X = \inf\{t : X_t \in \partial X\}$). The worst-case probability of unsafety for $\mathcal{L}$-trajectories starting at a point $x_0 \in X$ within a time horizon of $[t_0, T]$ is:

$$P^*(t_0, x_0) = \sup_{t^* \in [t_0, T]} \quad \mathrm{Prob}_{X_{t^*}}[x \in X_u] \tag{1a}$$

$$x(t) \text{ follows } \mathcal{L} \quad \forall t \in [t_0, \min(t^*, \tau_X)] \tag{1b}$$

$$x(0) = x_0. \tag{1c}$$

[1] J. Miller is with the Automatic Control Laboratory (IfA), Department of Information Technology and Electrical Engineering (D-ITET), ETH Zürich, Physikstrasse 3, 8092, Zürich, Switzerland (e-mail: jarmiller@control.ee.ethz.ch).

[2] M. Tacchi is with Univ. Grenoble Alpes, CNRS, Grenoble INP (Institute of Engineering Univ. Grenoble Alpes), GIPSA-lab, 38000 Grenoble, France. (e-mail: matteo.tacchi@gipsa-lab.fr)

[3] D. Henrion is with the Polynomial Optimization group of LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France; and the Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic. (henrion@laas.fr)

[4] M. Sznaier is with the Robust Systems Lab, ECE Department, Northeastern University, Boston, MA 02115. (e-mail: msznaier@coe.neu.edu).

J. Miller and M. Sznaier were partially supported by NSF grants CNS–1646121, ECCS–1808381 and CNS–2038493, AFOSR grant FA9550-19-1-0005, and ONR grant N00014-21-1-2431. J. Miller was partially supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

The stopping time $\min(t^*, \tau_X)$ in (1b) also ensures that trajectories $x(t)$ remain in X for all relevant t. The worst-case probability of unsafety for $\mathcal{L}$-trajectories over an initial set $X_0 \subseteq X$ is:

$$\begin{aligned} P^*(t_0, X_0) &= \sup_{x_0 \in X_0} P^*(t_0, x_0) && \text{(2a)} \\ &= \sup_{t^* \in [0,T],\ x_0 \in X_0} \quad \mathrm{Prob}_{X_{t^*}}[x \in X_u] && \text{(2b)} \\ & \qquad x(t) \text{ follows } \mathcal{L} \quad \forall t \in [t_0, \min(t^*, \tau_X)] && \text{(2c)} \\ & \qquad x(0) = x_0. && \text{(2d)} \end{aligned}$$

The initial set of X_0 is safe if $P^*(t_0, X_0) = 0$ and is unsafe if $P^*(t_0, X_0) = 1$. Any other value $P^*(t_0, X_0) \in (0, 1)$ returns a maximal probability of unsafety of the SDE.

The map $(t_0, x_0) \mapsto P^*(t_0, x_0)$ can be interpreted as a risk function in trajectory planning, in which level-sets of P^* are risk contours with constant unsafe probability. This paper will also develop a piecewise-polynomial sequence of risk functions that will converge from above in an L_1 sense to the optimal map P^*. Superlevel sets of any one of these risk-function-proxies can be used to identify unsafe regions.

Safety verification of stochastic systems includes the tasks of stochastic reachability [1], reach-avoid analysis [2, 3], ruin [4], and performance checking [5].

Infinite-dimensional convex Linear Program (LP) formulations of generically nonconvex and nonlinear stochastic trajectory problems can be used to analyze and control system executions, with no conservatism added under appropriate compactness and regularity conditions. The infinite-dimensional LP must be truncated into finite-dimensional programs in order to obtain numerical solutions. Algorithms to perform this truncation for generic stochastic processes include gridding [6, 7], radial basis function selection [8], and random sampling [9].

Barrier functions [10] provide a level-set certificates of invariance with probability $(1 - \epsilon)$. Control barrier functions [11–13] may also be generated to form controllers with stochastic guarantees of safety.

In the specific case where the stochastic process and all problem data have a rational representation, the moment-Sum of Squares (SOS) hierarchy [14] can be employed to truncate the LP into a hierarchy of Semidefinite Programs (SDPs) in increasing size, parameterized by the polynomial degree. Application of the moment-SOS hierarchy for stochastic analysis and control includes regional verification [15], reach-avoid estimation [16], reach-avoid control [17], region of attraction estimation [18], chance-constrained trajectory planning via moment propagation [19], exit-time estimation [20], and (conditional) value-at-risk upper-bounding [5].

This paper continues a research direction involving safety quantification of nonlinear systems. The safety of a set X_0 in (2) is quantified by its worst-case probability of unsafety. Other such quantities include the expected time a trajectory spends in the unsafe set [21], constraint violations representing an expanded unsafe set [22], the distance of closest approach [23], value-at-risk of state functions [5], and the perturbation intensity necessary to crash [24].

The two most relevant prior works to our paper are [25] and [26]. The work in [25] presents infinite-dimensional LPs to generate superlevel-based outer approximations of probability-p-safe sets of an SDE for a fixed initial distribution. Over the course of their derivations, they present a nonconservative formulation in Lemma 2 that is similar to our approach. Our problem evaluates safety of a specific initial set X_0 with respect to a free initial distribution, performs risk contour analysis, and extends the framework towards more general stochastic processes. The work in [26] synthesizes risk contours for optimization problems with distributional uncertainty (with an emphasis on obstacles and path planning), and our paper extends this work by analyzing safety of stochastic dynamical systems.

This paper has the following structure: Section 2 introduces preliminaries such as notation, stochastic processes, and SOS proofs of nonnegativity. Section 3 poses infinite-dimensional convex LPs to compute the worst-case probability of unsafety of a stochastic process w.r.t. an initial set X_0. Section 4 modifies the worst-case unsafety LPs to create risk contours that upper bound the true risk map $P^*(t_0, x_0)$. Section 5 truncates the unsafety LPs into a converging sequence of finite-dimensional SDPs using the moment-SOS hierarchy. Section 6 demonstrates the computation of unsafe probabilities and risk contours on example stochastic process with polynomial dynamics. Section 7 concludes the paper.

2 Preliminaries

2.1 Notation

The subset of natural numbers between a and b is $a..b \subset \mathbb{N}$. The minimum of two quantities will be denoted as $a \wedge b = \min(a, b)$. The set of polynomials with real-valued coefficients in an indeterminate $x \in \mathbb{R}^n$ is $\mathbb{R}[x]$. Every polynomial $c \in \mathbb{R}[x]$ may be uniquely described by a finite sum over multi-indices $\alpha \in \mathbb{N}^n$ by $c(x) = \sum_\alpha c_\alpha x^\alpha$. The degree of a polynomial $\deg c$ is equal to the largest exponent-sum $\sum_{i=1}^n \alpha_i$ such that $c_\alpha \neq 0$. The set of polynomials in x with degree at most $2k$ is $\mathbb{R}[x]_{\leq 2k}$.

The set of continuous functions over a set S is $C(S)$. Its subcone of nonnegative continuous functions is $C_+(S)$. Given a product set $S \times H$, the set of functions that are once-continuously differentiable in the first variable s and are twice-continuously differentiable in the second variable h is $C^{1,2}(S \times H)$. The set of signed Borel measures supported in S is $\mathcal{M}(S)$, and its subset of nonnegative Borel measures supported in S is $\mathcal{M}_+(S)$. The support of a measure $\mu \in \mathcal{M}_+(\mathbb{R}^n)$ is the smallest closed set of points s such that every open neighborhood $N_\epsilon(s)$ has positive measure $\mu(N_\epsilon(s)) > 0$. The sets $C_+(S)$ and $\mathcal{M}_+(S)$ are in topological duality when S is compact, and they admit a duality pairing $\langle \cdot, \cdot \rangle$ by Lebesgue integration: $\forall f \in C_+(S),\ \mu \in \mathcal{M}_+(S) : \langle f, \mu \rangle = \int_S f(s) d\mu(s)$. This duality pairing will be extended into a bilinear pairing between $C(S)$ and $\mathcal{M}_+(S)$. The μ-measure of a set $A \subseteq S$ may also be expressed as the pairing of μ with the indicator function $I_A :\ \mu(A) = \langle I_A, \mu \rangle = \inf\{\langle w, \mu \rangle \mid w \in C(S), w \geq I_A\}$. The mass of μ is $\mu(S) = \langle 1, \mu \rangle$, and μ is a probability distribution if this mass is 1. A vital example of a probability distribution is the Dirac delta $\delta_{s'}$ supported only at s', obeying the point-evaluation pairing rule $\forall f \in C(S) :\ \langle f, \delta_{s'} \rangle = f(s')$. The Lebesgue (volume) distribution over a space S is λ_S with $\forall f \in C(S) :\ \langle f, \lambda_S \rangle = \int_S f(s) ds$. The product measure between $\mu \in \mathcal{M}_+(S_1)$ and $\nu \in \mathcal{M}_+(S_2)$ is the unique measure $\mu \otimes \nu$ satisfying $\forall A_1 \times A_2 \subseteq S_1 \times S_2 :$ $\mu \otimes \nu(A_1 \times A_2) = \mu(A_1)\nu(A_2)$.

2.2 Stochastic Processes

A stochastic process is a time-indexed set of random variables $\{X_t\}$ that are related together through system dynamics [27] (pushforward of an initial distribution along flow maps). Properties of the stochastic process will be analyzed in terms of test functions in an appropriate set $\mathcal{C}$ (such as $\mathcal{C} = C([0, T] \times X)$. The expectation of a test function $v(s, x)$ at the time t according to the distribution X_t is $\mathbb{E}[v(t, x) \mid X_t]$. Letting $\Delta t > 0$ be a time step, the **generator** $\mathcal{L}$ of a stochastic process satisfies (for all appropriate test functions $v \in \mathrm{dom}\mathcal{L}$)

$$\mathcal{L}_{\Delta t} v(t, x) = \lim_{\Delta t' \to \Delta t} \frac{\mathbb{E}[v(t + \Delta t', x) \mid X_{t+\Delta t'}] - v(t, x)}{\Delta t'}. \tag{3}$$

We will express the domain of $\mathcal{L}$ as $\mathcal{C} = \mathrm{dom}\mathcal{L}$, such that $\mathcal{C}$ is a subset of the preimage of continuous functions under $\mathcal{L}$. The generator for a discrete-time stochastic process is $\mathcal{L}_{\Delta t}$ with $\Delta t > 0$, which is defined w.r.t. the test function class $\mathcal{C} = C([0, T] \times X)$. For a discrete-time law following $x_{t+\Delta t} = f(t, x_t, \lambda_t)$ in which the time-varying parameter $\lambda_t \in \Lambda$ has a probability distribution of $\xi(\lambda_t)$, the associated generator satisfies

$$\mathcal{L}_{\Delta t} v(t, x) = (\Delta t^{-1}) \int_\Lambda v\left(t + \Delta t, f(t, x, \lambda)\right) d\xi(\lambda) - v(t, x). \tag{4}$$

The generator for a continuous-time (Feller) stochastic process is $\mathcal{L}_{\Delta t = 0}$ for the class $\mathcal{C} = C^{1,2}([0, T] \times X)$ [28]. The dynamical behavior of an Itô SDE has a description

$$dx = f(t, x)dt + g(t, x)dW, \tag{5}$$

in which W is the Wiener process. The generator of the SDE in (5) is

$$\begin{aligned} \mathcal{L}_0 v(t, x) =& \partial_t v(t, x) + f(t, x) \cdot \nabla_x v(t, x) \\ &+ g(t, x)^T \nabla^2_{xx} v(t, x) g(t, x)/2. \end{aligned} \tag{6}$$

A sequence of random variables $\{Y_t\}$ is a *martingale* if $\mathbb{E}[Y_{t+\delta t} \mid \{Y_{t'}\}_{t' \leq t}] = Y_t$ [27] (in the sense of generalized conditional expectation). The generator $\mathcal{L}$ solves a martingale problem for all possible time steps $s > 0$ and test functions $v \in \mathcal{C}$ [29]. The martingale problem in discrete-time (with $s \in \mathbb{N}$) is

$$\begin{aligned} 0 = \mathbb{E}[v(t+s,x) \mid X_{t+s}] - \mathbb{E}[v(t,x) \mid X_t] \\ - \textstyle\sum_{s'=t}^{t+s} \mathbb{E}[\mathcal{L}_{\Delta t} v(t,x) \mid X_{s'}], \end{aligned} \tag{7}$$

and in continuous time (with $s \in \mathbb{R}_{\geq 0}$) is

$$\begin{aligned} 0 = \mathbb{E}[v(t+s,x) \mid X_{t+s}] - \mathbb{E}[v(t,x) \mid X_t] \\ - \textstyle\int_{s'=t}^{t+s} \mathbb{E}[\mathcal{L}_0 v(t,x) \mid X_{s'}]. \end{aligned} \tag{8}$$

For the rest of this paper, we will refer to the generator $\mathcal{L}$ interchangeably with its stochastic process.

The martingale problem can be described using the theory of occupation measures. Given an initial condition (point) $x_0 \in X_0 \subseteq X$ and a time $t \in [0,T]$, we refer to $x(t \mid x_0)$ as the trajectory of $\{X_t\}$ with generator $\mathcal{L}$ starting at the initial distribution $\delta_{x=x_0}$. The **occupation measure** $\mu \in \mathcal{M}_+([0,T] \times X)$ and **terminal measure** $\mu_\tau \in \mathcal{M}_+([0,T] \times X)$ of the stochastic process $\{X_t\}$ from a starting time t_0 up to a stopping time t^* satisfies $\forall A \subseteq [t_0,T], B \subseteq X$:

$$\mu(A \times B) = \int_{X_0} \int_{t=t_0}^{t^*} I_{A\times B}\left(t, x(t \mid x_0)\right) dt \, d\mu_0(x_0) \tag{9a}$$

$$\mu_\tau(A \times B) = \int_{X_0} I_{A\times B}\left(t^*, x(t^* \mid x_0)\right) \, d\mu_0(x_0). \tag{9b}$$

The measures in (9) may also be defined with respect to a probability distribution of stopping times $t^* \in \mathcal{M}_+([0,T])$.

A tuple of measures $(\mu_0, \mu, \mu_\tau) \in \mathcal{M}_+(X_0) \times \mathcal{M}_+([0,T] \times X)^2$ obeys the martingale relations (7) or (8) of the stochastic process with generator $\mathcal{L}$

$$\forall v \in \mathcal{C}: \qquad \langle v, \mu_\tau \rangle = \langle v(t_0, \cdot), \mu_0 \rangle + \langle \mathcal{L} v, \mu \rangle. \tag{10}$$

Relation (10) is called Liouville equation for Ordinary Differential Equations (ODEs) and is referred to as Dynkin's equation in the context of SDEs. We will also equivalently express (10) in a shorthand form by

$$\mu_\tau = \delta_{t_0} \otimes \mu_0 + \mathcal{L}^\dagger \mu. \tag{11}$$

2.3 Sum-of-Squares Methods

Let $c \in \mathbb{R}[x]_{\leq 2k}$ be a polynomial. The polynomial c is nonnegative if $\forall x \in \mathbb{R}^n : c(x) \geq 0$. Verification of polynomial nonnegativity is generically an NP-hard task. One tractable method to construct nonnegative polynomials is through SOS representations. The polynomial c is SOS if there exists N polynomials $q_j \in \mathbb{R}[x]$ such that $c(x) = \sum_{j=1}^N q_j^2(x)$. The cone of SOS polynomials is $\Sigma[x]$, and its subset of SOS polynomials with degree $\leq 2k$ is $\Sigma[x]_{\leq 2k}$. The set of SOS polynomials equals the cone of nonnegative polynomials only when $n = 1, 2k = 2$, or $(n = 2, 2k = 4)$ [30]. SOS polynomials become vanishingly small in the set of nonnegative polynomials as n and d rise [31].

SOS representations up to fixed degree can be conducted by formulating SDPs. To each polynomial $c \in \mathbb{R}[x]_{\leq 2k}$, there exists a (typically nonunique) associated polynomial vector $v_k(x) \in (\mathbb{R}[x]_{\leq d})^s$ and a symmetric *Gram* matrix $Q \in \mathbb{S}^s$ such that $c(x) = v_k(x)^T Q v_k(x)$. One such choice of vectors $v_k(x)$ is the set of monomials in x between degrees $0..k$, in which case $s = \binom{n+k}{k}$. A polynomial c is SOS if its Gram matrix Q is also Positive Semidefinite (PSD) [32].

A Basic Semialgebraic (BSA) set $\mathbb{K}$ is a set described by a finite number of bounded-degree polynomial inequality (N_g) and equality constraints (N_h):

$$\mathbb{K} = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, \ h_j(x) = 0, \forall i \in 1..N_g, j \in 1..N_h\}. \tag{12}$$

The class of Weighted Sum of Squares (WSOS) polynomials over $\mathbb{K}$ in (12) is the set $\Sigma[\mathbb{K}]$ of polynomials c that admit a description of

$$c(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x) \tag{13a}$$

$$\sigma_0, \sigma_i \in \Sigma[x], \quad \phi_j \in \mathbb{R}[x]. \tag{13b}$$

The set $\mathbb{K}$ is compact if there exists a finite $R > 0$ such that $\{x \mid \|x\|_2^2 \leq R\} \supseteq \mathbb{K}$. The set is additionally Archimedean if $R - \|x\|_2^2 \in \Sigma[\mathbb{K}]$. Not every compact set is Archimedean [33], but if an R verifying compactness of $\mathbb{K}$ is known, then the constraint $R - \|x\|_2^2 \geq 0$ can be added to the description of $\mathbb{K}$ to render $\mathbb{K}$ Archimedian. Every compact polytope and ellipsoid is also Archimedean [14].

The Putinar Positivestellensatz states that every positive polynomial over $\mathbb{K}$ is also a member of $\Sigma[\mathbb{K}]$ when $\mathbb{K}$ is Archimedean [34]. However, the polynomial degrees of the multipliers (σ, ϕ) needed to certify nonnegativity of such a positive c could be exponential in n and k [35].

3 Unsafe Linear Programs

This section will present LPs to analyze safety of a stochastic process with generator $\mathcal{L}$.

3.1 Assumptions

We will posit the following assumptions throughout this paper:

A1 The state-sets X_0, X_u, X are all compact.

A2 The time interval $[t_0, T]$ is compact.

A3 Trajectories stop upon the first exit from X $(\tau_X \wedge T)$.

A4 The test function set $\mathcal{C} = \mathrm{dom}(\mathcal{L})$ satisfies $\mathcal{C} \subseteq C([t_0, T] \times X)$ with $1 \in \mathcal{C}$ and $\mathcal{L}1 = 0$.

A5 The set $\mathcal{C}$ separates points and is multiplicatively closed.

A6 There exists a countable set $\{v_k\} \in \mathcal{C}$ such that $\forall v \in \mathcal{C} : (v, \mathcal{L}v)$ is contained in the bounded pointwise closure of the linear span of $\{(v_k, \mathcal{L}v_k)\}$.

Assumptions A4-A6 are the same as in Condition 1 of [6]. We note that generators for Markov, SDE, and Lévy processes evolving in a compact set satisfy A1-A6.

3.2 Measure Program

We will pose a convex but infinite-dimensional LP in order to upper-bound program (2). Such an LP will be posed in terms of a relaxed occupation measure (μ_0, μ_τ, μ).

Theorem 3.1. *The following LP will upper-bound* (2) *with* $p^*(t_0, X_0) \geq P^*(t_0, X_0)$ *for each* $X_0 \subseteq X$

$$p^*(t_0, X_0) = \sup \quad \langle I_{X_u}, \mu_\tau \rangle \tag{14a}$$
$$\mu_\tau = \delta_{t_0} \otimes \mu_0 + \mathcal{L}^\dagger \mu \tag{14b}$$
$$\langle 1, \mu_0 \rangle = 1 \tag{14c}$$
$$\mu_0 \in \mathcal{M}_+(X_0) \tag{14d}$$
$$\mu, \ \mu_\tau \in \mathcal{M}_+([t_0, T] \times X). \tag{14e}$$

Proof. Let $x_0 \in X_0$ be an initial condition and $t^* \in [t_0, T]$ be a stopping time. A set of measures (μ_0, μ, μ_τ) may be constructed to fulfill constraints (14b)-(14e). To begin, the initial measure μ_0 may be chosen as the Dirac delta $\delta_{x=x_0}$. The measure μ can be set to the occupation measure of the stochastic trajectory $x(t \mid x_0)$ between the initial time of t_0 and the stopping time of $t^* \wedge \tau_X$. The terminal measure $\mu_\tau(t, x)$ may then be picked as the the time-state distribution at time t^* (in which trajectories are stopped according to $t^* \wedge \tau_X$). Because there exists a mapping from every feasible point (t^*, x_0) of (2) to a set of measures in (16b)-(16f), it holds that $p^*_\tau(t_0, X_0) \geq P^*(t_0, X_0)$. □

The function $I_{X_u}(x)$ in the objective (14a) is a discontinuous function of x. We can reformulate (14) as an LP where all terms are continuous in (t, x) (under A1-A6) through the use of the following restriction lemma:

Lemma 3.2 (Theorem 3.1 of [36]). *Let S_1, S_2 be spaces with $S_1 \subseteq S_2$, and let $\xi \in \mathcal{M}_+(S_2)$ be a measure. The restriction of ξ to S_1 is uniquely given by the measure ν^* solving the following optimization problem*

$$Q^* = \sup_{\nu, \hat{\nu}} \langle 1, \nu \rangle \tag{15a}$$

$$\xi = \nu + \hat{\nu} \tag{15b}$$

$$\nu \in \mathcal{M}_+(S_1) \tag{15c}$$

$$\hat{\nu} \in \mathcal{M}_+(S_2). \tag{15d}$$

*The objective Q^* will satisfy $Q^* = \langle 1, \nu^* \rangle = \langle I_{S_1}, \xi \rangle$. The measure $\hat{\nu}^*$ will be supported on **closure**$(S_2 \setminus S_1)$.*

We will define a peak measure $\mu_p \in \mathcal{M}_+([t_0, T] \times X_u)$ and a complement measure $\mu_c \in \mathcal{M}_+([t_0, T] \times X)$ in order to split the terminal measure μ_τ as $\mu_\tau = \mu_p + \mu_c$. This decomposition (through Lemma 3.2) will allow for the definition of an LP with continuous coefficients in (t, x). This continuity provides strong duality (under A1-A6) and convergence of the moment-SOS hierarchy to $P^*(t_0, X_0)$ from (2) (under A1-A6 and assumptions on polynomial structure). The LP in which μ_τ is decomposed as $\mu_p + \mu_c$ is:

Theorem 3.3. *The following LP has the same objective value $p^*(t_0, X_0)$ as in (2):*

$$p^*(t_0, X_0) = \sup \quad \langle 1, \mu_p \rangle \tag{16a}$$

$$\mu_p + \mu_c = \delta_{t_0} \otimes \mu_0 + \mathcal{L}^\dagger \mu \tag{16b}$$

$$\langle 1, \mu_0 \rangle = 1 \tag{16c}$$

$$\mu_0 \in \mathcal{M}_+(X_0) \tag{16d}$$

$$\mu, \ \mu_c \in \mathcal{M}_+([t_0, T] \times X) \tag{16e}$$

$$\mu_p \in \mathcal{M}_+([t_0, T] \times X_u). \tag{16f}$$

Proof. **1)** Let $(\mu_0, \mu_p, \mu_c, \mu)$ be a feasible point to (16b)-(16f). Then $(\mu_0, \mu_p + \mu_c, \mu)$ is a feasible point to (14b)-(14e).

2) Let (μ_0, μ_τ, μ) be a feasible point to the constraints (14b)-(14e). Unique measures μ_p, μ_c may be defined according to the restriction Lemma 3.2 with $S_1 = [t_0, T] \times X_u$ and $S_2 = [t_0, T] \times X$ by solving

$$\sup \langle 1, \mu_\tau \rangle \tag{17a}$$

$$\mu_\tau = \mu_p + \mu_c \tag{17b}$$

$$\mu_c \in \mathcal{M}_+([t_0, T] \times X) \tag{17c}$$

$$\mu_p \in \mathcal{M}_+([t_0, T] \times X_u), \tag{17d}$$

with $\langle I_{X_u}, \mu_\tau \rangle = \langle 1, \mu_p \rangle$.

There exists a mapping relating each feasible point of (14) to (16) and vice versa while keeping the same objective, which proves that the two objectives are equal.

□

Theorem 3.4. *Under assumptions A1-A6, the objectives of (16) and (1b) are equal ($p^*(t_0, X_0) = P^*(t_0, X_0)$).*

Proof. By Theorem 3.1 of [6] (and under Assumptions A1-A6), every tuple (μ_0, μ, μ_τ) with $\mu_0 \in \mathcal{M}_+([t_0, T] \times X)$ (with $\langle 1, \mu_0 \rangle = 1$) and $\mu, \mu_\tau \in \mathcal{M}_+([t_0, T] \times X)$ satisfying a Martingale relation (11) is supported on the graph of a stochastic process (1b).

The tuple $(\mu_0, \mu, \mu_p + \mu_c)$ from the Martingale relation (11) is supported on the graph of a stochastic process. The measures μ_p and μ_c can be separated according to Lemma 3.2. This concludes the no-relaxation-gap proof. □

3.3 Function Program

Theorem 3.5. *An LP in terms of an auxiliary function v that forms a weak dual ($p^*(t_0, X_0) \geq d^*(t_0, X_0)$) to* (16) *is*

$$\begin{aligned}
d^*(t_0, X_0) = \inf_{\gamma \in \mathbb{R}} \quad & \gamma && && \text{(18a)}\\
& \gamma \geq v(t_0, x) && \forall x \in X_0 && \text{(18b)}\\
& \mathcal{L}v(t, x) \leq 0 && \forall (t, x) \in [t_0, T] \times X && \text{(18c)}\\
& v(t, x) \geq 0 && \forall (t, x) \in [t_0, T] \times X && \text{(18d)}\\
& v(t, x) \geq 1 && \forall (t, x) \in [t_0, T] \times X_u && \text{(18e)}\\
& v \in \mathcal{C}. && && \text{(18f)}
\end{aligned}$$

Strong duality ($p^*(t_0, X_0) = d^*(t_0, X_0)$) will hold under assumptions A1-A4.

Proof. See Appendix A. □

Remark 1. *Constraints* (18b)-(18f) *have the same form as a stochastic (time-dependent) barrier function at a fixed probability level γ (from [37]). Our work involves application of γ as an optimization variable, as well as a proof of no relaxation gap in Theorem 3.4 under A1-A6.*

4 Risk Contour Linear Programs

The LPs in (16) and (18) determine the worst-case probability of unsafety among stochastic trajectories starting from an initial set X_0. The optimal value of (16) and (18) is therefore equal to 1 when $X_0 = X$ and X_u is nonempty (because then $\varnothing \neq X_u \subset X_0$ and there exists trajectories starting in the unsafe set). This section will form a continuous-function LP similar to (18), whose function values evaluated at a given initial point X_0 form a upper-bound on the probability of unsafety when starting at x_0. These upper-bounds will be proven to converge in an L_1 sense over X to the true risk function P^*. Such upper-bounds on unsafety allow for interpretation of risk when performing path planning tasks.

Let $\mu_0 \in \mathcal{M}_+(X)$ be a fixed probability distribution ($\langle 1, \mu_0 \rangle = 1$). The μ_0-averaged probability of unsafety is

$$M^*(t_0, \mu_0) = \int_X P^*(t_0, x_0) \; d\mu_0(x_0). \tag{19}$$

One specific case we will focus on in the examples is where μ_0 is the the uniform distribution over X ($\mu_0 = \lambda_X / \text{vol}(X)$).

4.1 Measure Program

Corollary 1. *An LP to upper-bound* (2) *with $m^*(t_0, \mu_0) \geq M^*(t_0, \mu_0)$ in general, and $m^*(t_0, \mu_0) = M^*(t_0, \mu_0)$ under A1-A6, is*

$$\begin{aligned}
m^*(t_0, \mu_0) = \sup \quad & \langle 1, \mu_p \rangle && \text{(20a)}\\
& \mu_p + \mu_c = \delta_{t_0} \otimes \mu_0 + \mathcal{L}^\dagger \mu && \text{(20b)}\\
& \mu, \; \mu_c \in \mathcal{M}_+([t_0, T] \times X) && \text{(20c)}\\
& \mu_p \in \mathcal{M}_+([t_0, T] \times X_u). && \text{(20d)}
\end{aligned}$$

Proof. These optimality properties follow from arguments used in Theorem 3.3 (upper-bound by a measure construction starting from the distribution μ_0) and Theorem 3.4 (restriction measures). The measures $(\mu, \mu_p + \mu_c)$ are supported on the graph of the stochastic process $\mathcal{L}$ starting from μ_0 for every feasible (μ, μ_p, μ_c). □

4.2 Function Program

For given $v \in \mathcal{C}$ and $\mu_0 \in \mathcal{M}_+(X)$, let $v(t_0, \bullet)$ be shorthand notation for the function $x \mapsto v(t_0, x)$ such that $\langle v(t_0, \bullet), \mu_0 \rangle = \langle v(t_0, x), \mu_0(x) \rangle$. We will modify terms (18a)-(18b) in order to form an LP penalized over the whole space X with an objective integrating against the weighting distribution μ_0:

$$\begin{aligned}
J^*(t_0, \mu_0) = &\inf \langle v(t_0, \bullet), \mu_0 \rangle && && \text{(21a)}\\
&\mathcal{L}v(t, x) \leq 0 && \forall (t, x) \in [t_0, T] \times X && \text{(21b)}\\
&v(t, x) \geq 0 && \forall (t, x) \in [t_0, T] \times X && \text{(21c)}\\
&v(t, x) \geq 1 && \forall (t, x) \in [t_0, T] \times X_u && \text{(21d)}\\
&v \in \mathcal{C}. && && \text{(21e)}
\end{aligned}$$

Lemma 4.1. *The objective* (21a) *from* (21) *is upper-bounded by* $J^*(t_0, \mu_0) \leq 1$.

Proof. The function $v(t, x) = 1$ is feasible for all constraints (21b)-(21e). Given that μ_0 is a probability measure, the objective in (21a) will satisfy $J^* \leq \langle 1, \mu_0 \rangle = 1$. □

We now show that any solution of (21b)-(21e) returns an upper-bound of the probability of unsafety.

Theorem 4.2. *Let v be feasible for* (21b)-(21e) *and let assumptions A1-A6 hold. Then*

$$v(t_0, x_0) \geq P(t_0, x_0) \tag{22}$$

for every initial condition $x_0 \in X$.

Proof. See Appendix B. □

Remark 2. *Level sets of any $v(t, x)$ solving* (21b)-(21e) *can therefore be used to upper-bound the risk of stochastic execution for $\mathcal{L}$ when starting at any $(t, x_0) \in [t_0, T] \times X$.*

The notation $p(t_0, x_0)$ will be subsequently used to denote the function obtained by solving (16) as $x_0 \mapsto p(t_0, X_0 = \{x_0\})$.

Theorem 4.3. *Programs* (20) *and* (21) *obey strong duality $m^*(t_0, \mu_0) = J^*(t_0, \mu_0)$ under A1-A6.*

Proof. This strong duality holds by slight modification of the method used in Theorem 3.5 and Appendix A. □

Theorem 4.4. *Let $\{v_k\}_{k\geq 1}$ be a sequence of solutions to* (21b)-(21e) *with $\langle v_k(t_0, \bullet), \mu_0 \rangle \to J^*(t_0, \mu_0)$ (under A1-A6). Then v_k converges in $L_1(\mu_0)$ to p^* as in*

$$\int_X \Big(v_k(t_0, x_0) - p^*(t_0, x_0) \Big)\, d\mu_0(x_0) \underset{k\to\infty}{\longrightarrow} 0. \tag{23}$$

Proof. See Appendix C. □

5 Unsafe Semidefinite Programs

We will approximate programs (18) and (21) through the moment-SOS hierarchy of SDPs, as reviewed in Section 2.3. In order to apply SOS methods towards convergent approximation, we require additional assumptions:

A7 The sets X_0, X_u, X are all Archimedean BSA sets.

A8 The generator $\mathcal{L}$ is closed under polynomials ($v(t, x) \in \mathbb{R}[t, x] \implies \mathcal{L}v(t, x) \in \mathbb{R}[t, x]$).

Given a degree k and a generator $\mathcal{L}$ obeying Assumption A8, we define the dynamics degree $\tilde{k}$ as

$$\tilde{k} = \lceil \max_{v \in \mathbb{R}_{\leq 2k}[t,x]} \deg(\mathcal{L}v(t, x))/2 \rceil. \tag{24}$$

5.1 Unsafe-Probability SDP

The degree-k SOS tightening of program (18) is

$$d_k^*(t_0, X_0) = \inf_{\gamma \in \mathbb{R}} \quad \gamma \tag{25a}$$

$$\gamma - v(t_0, x) \in \Sigma[X_0]_{\leq 2k} \tag{25b}$$

$$-\mathcal{L}v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2\tilde{k}} \tag{25c}$$

$$v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2k} \tag{25d}$$

$$v(t, x) - 1 \in \Sigma[[0, T] \times X_u]_{\leq 2k}. \tag{25e}$$

Remark 3. *Constraint* (25d) *absorbs the polynomial restriction of* $v \in \mathbb{R}[t, x]_{\leq 2k}$.

We require the following Lemma to ensure convergence of (25) to (18) as $k \to \infty$.

Lemma 5.1. *Under Assumptions A1-A6, all feasible measures* $\boldsymbol{\mu} = (\mu_0, \mu, \mu_c, \mu_\tau)$ *in* (16) *are bounded.*

Proof. Boundedness of nonnegative measures will be proven by the sufficient condition of compact support and finite mass. Compact support of measures in $\boldsymbol{\mu}$ is ensured by A1. The initial measure μ_0 is a probability distribution ($\langle 1, \mu_0 \rangle = 1$) by constraint (16c). The sum $\mu_c + \mu_\tau$ likewise has mass 1 by the Liouville constraint (16b), when passing in the test function $v = 1$. Given that μ_c and μ_τ are each nonnegative measures, it holds that they both have finite mass (upper-bounded by 1). Lastly, assignment of $v = t$ to (16b) with $\mathcal{L}t = 1$ results in $\langle 1, \mu \rangle = \langle t, \mu_c + \mu_\tau \rangle \leq T$. All measures have bounded masses and compact supports, and therefore are bounded. □

Theorem 5.2. *Under assumptions A1-A8, program* (25) *will converge to* (18) *with* $\lim_{k\to\infty} d_k^*(t_0, X_0) = d^*(t_0, X_0)$.

Proof. This convergence will hold by Corollary 8 of [38], because all sets are Archimedean (A7), dynamics are polynomial (A8), measures are bounded (Lemma 5.1), and the objective value of (25) is finite (bounded below by 0). □

5.2 Risk SDP

The degree-k SOS tightening for the risk-contour program in (21) is

$$J_k^*(t_0, \mu_0) = \inf \int_X v(t_0, x) d\mu_0(x) \tag{26a}$$

$$-\mathcal{L}v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2\tilde{k}} \tag{26b}$$

$$v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2k} \tag{26c}$$

$$v(t, x) - 1 \in \Sigma[[0, T] \times X_u]_{\leq 2k}. \tag{26d}$$

Remark 4. *The function* $v(t, x) = 1$ *is feasible for constraints* (26b)-(26d) *at every degree* $k \in \mathbb{N}$. *As a result, the objective in* (26) *is always upper-bounded by* $J_k^*(t_0, \mu_0) \leq 1$.

Theorem 5.3. *Under assumptions A1-A8, the risk contour program* (26) *will converge to* (21) *with* (18) *with* $\lim_{k\to\infty} J_k^* = J^*$.

Proof. The proof of this theorem follows the same steps as in Theorem 5.2 with respect to the measure program in (20). □

Let $v_k(t, x) \in \Sigma[[0, T] \times X]_{\leq 2k}$ be the solution to (26) at degree-k, and let I_u be the 0/1 indicator function on the unsafe set X_u. Then the probability of unsafety when starting at a point $x_0 \in X$ is upper-bounded by

$$q_{1:k}(x) = \min(1, \min_{k' \in 1..k} v_{k'}(t_0, x)). \tag{27}$$

Corollary 2. *The sequence of functions* $q_{1:k}(x)$ *in increasing* k *will converge in measure* μ_0 *to* $v(t_0, x)$.

Proof. This corollary follows from Theorem 4.4, in which the sequence $\{v_{k'}\}$ is used to approximate v^*. The minimization among all k' in (27) further sharpens the estimate of v^*. □

5.3 Computational Complexity

We will quantify computational complexity of the degree-k tightenings of (25) and (26) by the size of the maximal Gram matrices involved in their SOS program. In the typical case where $\tilde{k} > k$ (only violated under A1-A8 when $\mathcal{L}$ maps every polynomial to a constant), the largest Gram matrix will occur in the Lie constraints (25c) and (26b). The Lie constraints each have $n+1$ variables (t, x), so the Gram matrix size when using the monomial basis is $\binom{n+1+\tilde{k}}{\tilde{k}}$. All other constraints have a lower degree (k rather than $\tilde{k}$), or are posed only over the n variables x. The complexity of using an interior-point method to solve the SOS programs will therefore scale based on $O((n+1)^{6d})$ for fixed d or $O(d^{(n+1)})$ for fixed n [14, 39].

6 Numerical Examples

MATLAB (2021a) code to reproduce all examples is available at https://github.com/jarmill/prob_unsafe. All programs are modeled using YALMIP [40] and solved using Mosek 10 [41]. All examples will involve an initial time of $t_0 = 0$.

6.1 Two-Dimensional Cubic SDE

Our first demonstration analyzes safety of a cubic polynomial SDE from Example 1 of [10]:

$$dx = \begin{bmatrix} x_2 \\ -x_1 - x_2 - \frac{1}{2}x_1^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} dW. \tag{28}$$

Safety of (28) is evaluated within the state space of $X = [-2, 2]^2$ until a time horizon of $T = 5$. The unsafe set is a moon-shaped region $X_u = \{x \in \mathbb{R}^2 \mid 0.9083^2 \leq (x_1+0.5006)^2+(x_2+0.2902)^2, 0.5^2 \geq (x_1-0.2)^2+x_2^2\}$. The initial set X_0 is a circle of radius R_0 and center $[0.85; -0.75]$. Figure 2a plots trajectories of (28) starting from X_0 (magenta region with $R_0 = 0.2$). Note how some sampled trajectories touch the right corner of the red half-circle X_u (and are therefore unsafe).

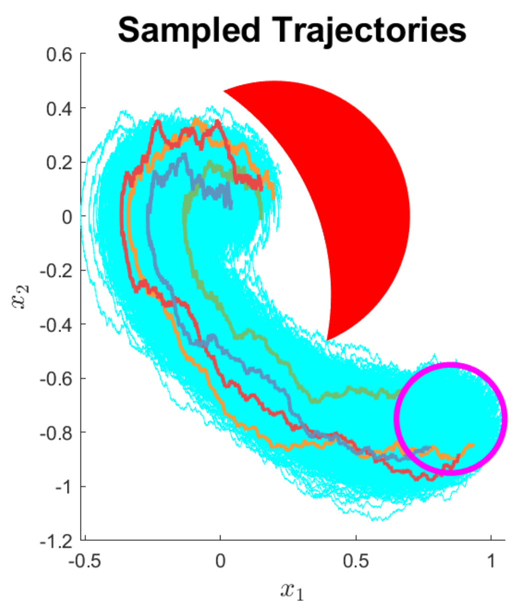


Figure 1: Trajectories of SDE (28) (cyan) initialized in a disk X_0 (purple boundary), hitting an unsafe region X_u (red moon).

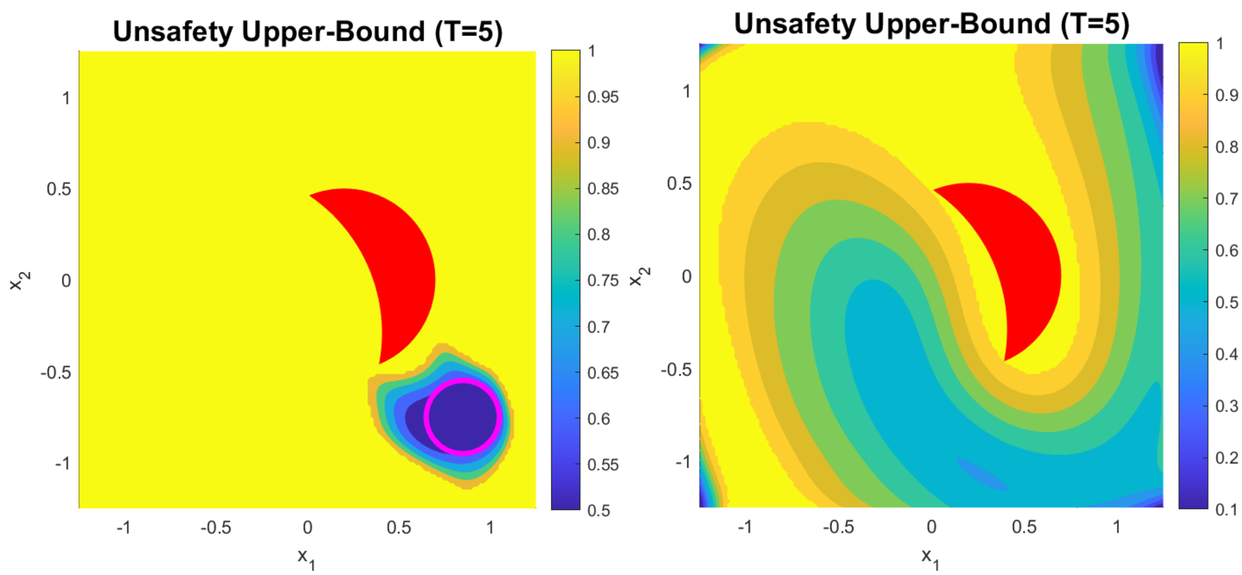
Table 1 reports the probability of unsafety computed by program (25) for a circular initial set of radius R_0.

Table 1: Unsafe probability upper-bounds for system (28)

order	1	2	3	4	5	6
$R_0 = 0$	1	0.9442	0.6072	0.4805	0.4068	0.3696
$R_0 = 0.2$	1	0.9736	0.7943	0.7091	0.6390	0.6132

Figure 2a plots the unsafety upper-bounding function $\min(1, v(0, x))$ at $T = 5$ acquired by solving (25) at order k =6 with $R_0 = 0.2$. The unsafe set X_u is the red moon. The magenta circle is the boundary of the initial set X_0. Note how the probability estimate is sharper in the region surrounding X_0, as compared to the sea of Prob = 1 saturation away from X_0.

In contrast, Figure 2b solves the order-6 SOS tightening of program (21) to produce a risk map v valid in $X = [-2, 2]^2$. This risk map results in $v(0, [0.85; -0.75]) = 0.4366$, which is looser than the order-6 probability bound of 0.3696 from Table 1. However, Figure 2b produces an interpretable visualization of risk across the entire set X.



(a) solution of (25)

(b) solution of (26)

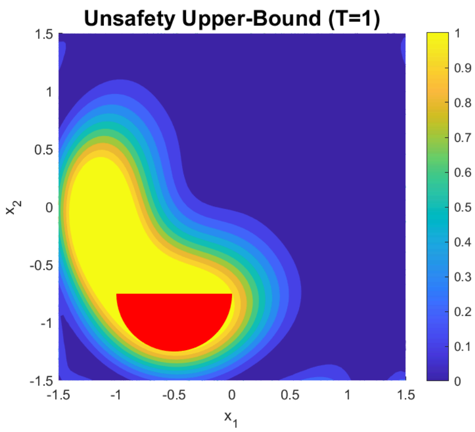
Figure 2: Risk levels of the unsafety upper-bound function $v(0, x)$ at $t = T = 5$ and order k =6 for SDE (28) initialized in a disk X_0 (purple boundary) with unsafe region X_u (red moon).

6.2 Adjustment of Time Horizons

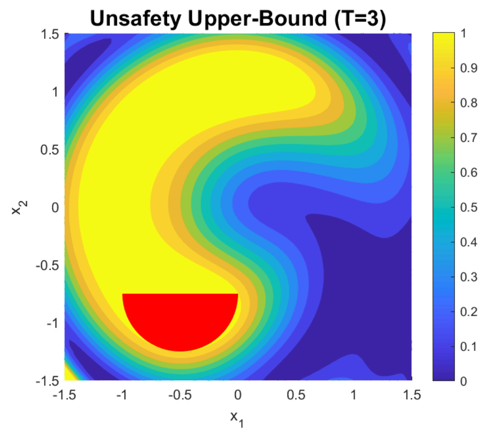
In the second experiment, we consider a linear SDE with dynamics

$$dx = [-x_2; x_1]dt + [0; 0.1]dW \tag{29}$$

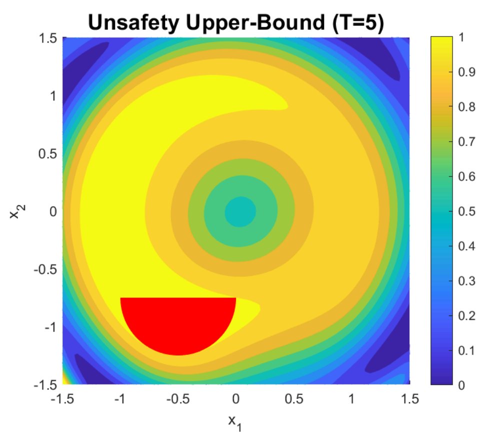
The unsafe set for this example is a half-circular region with $X_u = \{x \mid 0.5^2 \geq (x + 0.5)^2 + (x_2 + 0.75)^2, x_2 \geq -0.75\}$. The plots in Figure 3 are obtained by solving (26) at order $k = 6$ for time horizons of $T \in \{1, 3, 5\}$ and a state set of $X = [-2, 2]^2$. As the time horizon T increases, so too does the area of X with high probabilities of unsafety. The growth of the high-probability regions occurs approximately in a clockwise pattern from X_u, matching the interpretation of (29) as simple harmonic motion perturbed by diffusion in the vertical coordinate.



(a) $T = 1$



(b) $T = 3$



(c) $T = 5$

Figure 3: Risk contours for the linear SDE in (29) in increasing time horizon T w.r.t. unsafe region X_u (red half-disk)

6.3 Discrete-time example

This subsection will focus on a discrete-time stochastic process involving a time-step of $\tau = 1$. The law of this stochastic process with parameter $\lambda \in \mathbb{R}$ is

$$x_+ = \begin{bmatrix} -0.3x_1 + 0.8x_2 + x_1 x_2 \lambda/4 \\ -0.9x_1 - -0.1x_2 - 0.2x_1^2 + \lambda/40 \end{bmatrix}, \tag{30}$$

in which λ is i.i.d. sampled according to a unit-normal distribution at each time ($\lambda_t \sim \mathcal{N}(0, 1)$).

We evaluate the safety of (30) with respect to the state set $X = [-1.5, 1.5]^2$, the time horizon of $T = 10$, and the half-circle unsafe set of $X_u = \{x \mid 0.4^2 \geq (x_1 - 0.8)^2 + (x_2 - 0.2)^2,\ x_1 + x_2 \geq 1\}$. Stochastic trajectories of (30) evolve starting at an circular initial set with radius R_0 and center $[-1; 0]$. Figure 4 plots 5,000 sampled trajectories (blue dots) with respect to the magenta circular initial set and the red half-circle unsafe set. Some of the sampled trajectory points fall inside the red half-circle, corresponding to unsafety when beginning in X_0.

Table 2 reports probabilities of unsafety for (30) found by solving (25) at initial radii $R_0 = 0$ and $R_0 = 0.4$. To improve numerical conditioning, we normalize the time-steps from $(\tau, T) = (1, 10)$ to $(\tau, T) = (0.1, 1)$ without affecting the autonomous dynamics in (30).

Figure 5a plots risk contours found by solving (25) at order 6 for $R_0 = 0.4$. Figure 5b plots risk contours $v(0, x)$ acquired from (26) at degree k =6 and time $t = T = 10$. The returned risk map has an evaluation of $v(0, [-1; 0]) = 0.4915$, as compared to the (25) estimate of 7.052×10^{-4} at $R_0 = 0$.

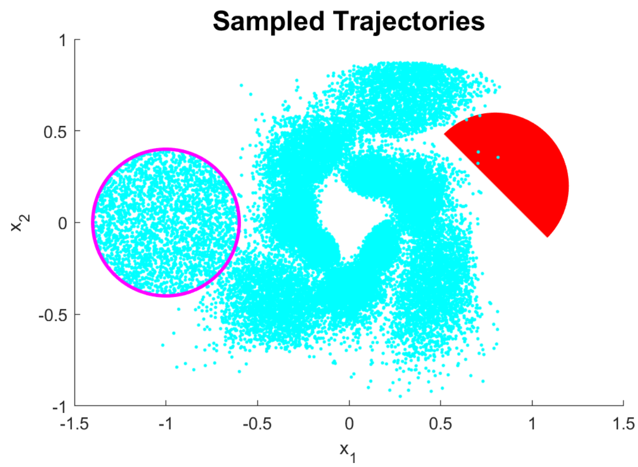
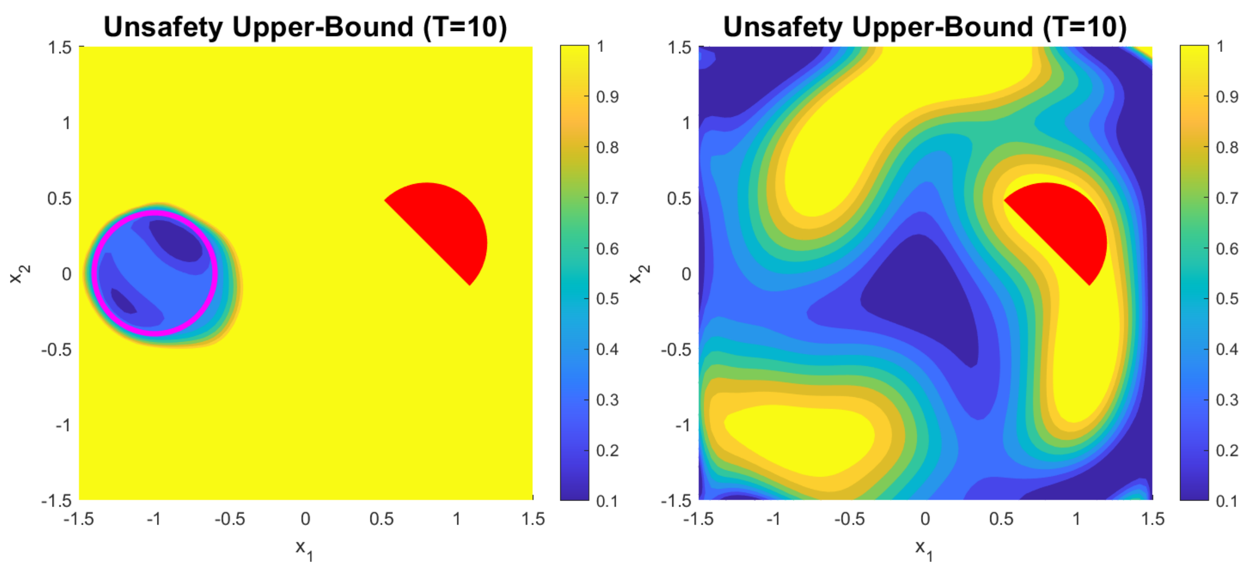


Figure 4: Trajectories of discrete-time (30) (cyan) initialized in a disk X_0 (purple boundary), hitting an unsafe region X_u (red half disk).

Table 2: Unsafe probability upper-bounds for system (30)

order	1	2	3	4	5	6
$R_0 = 0$	1.00	1.00	0.1569	0.0103	1.871e-3	7.052e-4
$R_0 = 0.4$	1.00	1.00	0.9801	0.7054	0.5225	0.4017



(a) solution of (25)

(b) solution of (26)

Figure 5: Risk levels of unsafety upper-bound function $v(0, x)$ at $t = T = 10$ and order k =6 for SDE (30) initialized in a disk X_0 (purple boundary) with unsafe region X_u (red moon).

7 Conclusion

This paper presents a method to analyze the probability of unsafety for stochastic processes by forming LPs in occupation measures. These LPs are upper-bounded by the moment-SOS hierarchy of SDPs, yielding a convergent sequence of bounds to the true probability of unsafety. Modification of the objective leads to the development of visually interpretable risk contours for use in analysis and motion planning. The presented method can be used for any stochastic process satisfying assumptions A1-A8 (compactness, described by generators, having a polynomial description).

Future work involves finding redundant constraints to refine and improve the unsafe-probability bounds. Other work includes extension to Conditional Value-at-Risk-minimizing stochastic control methods.

Acknowledgements

The authors would like to thank Roy S. Smith, Niklas Schmid, and the Automatic Control Laboratory at ETH Zürich for their advice and support.

A Strong Duality

This appendix will prove strong duality between (16) and (18) by using arguments from Theorem 2.6 of [42].

We collect the variables in (16) and (18) respectively into,

$$\boldsymbol{\mu} = (\mu_0, \mu, \mu_c, \mu_p) \qquad\qquad \boldsymbol{\ell} = (\gamma, v). \tag{31}$$

Variable spaces related to $\boldsymbol{\mu}$ are,

$$\begin{aligned} \mathcal{X}' &= C(X_0) \times C([0,T] \times X)^2 \times C([0,T] \times X_u) \\ \mathcal{X} &= \mathcal{M}(X_0) \times \mathcal{M}([0,T] \times X)^2 \times \mathcal{M}([0,T] \times X_u), \end{aligned} \tag{32}$$

with nonnegative subcones of

$$\begin{aligned} \mathcal{X}'_+ &= C_+(X_0) \times C_+([0,T] \times X)^2 \times C_+([0,T] \times X_u) \\ \mathcal{X}_+ &= \mathcal{M}_+(X_0) \times \mathcal{M}_+([0,T] \times X)^2 \times \mathcal{M}_+([0,T] \times X_u). \end{aligned} \tag{33}$$

Corresponding spaces to $\boldsymbol{\ell}$ are

$$\mathcal{Y}' = \mathcal{C} \times \mathbb{R} \tag{34}$$

$$\mathcal{Y} = \mathcal{C}' \times 0. \tag{35}$$

Following the notation of [42], we write that $\mathcal{Y}_+ = \{0_{\mathcal{Y}}\}$ and $\mathcal{Y}'_+ = \mathcal{Y}'$. We note the containments of $\boldsymbol{\mu} \in \mathcal{X}_+,\ \boldsymbol{\ell} \in \mathcal{Y}',$ and also that $(\mathcal{X}, \mathcal{X}')$ form a pair of topological dual spaces under Assumption A1. The topologies for the spaces $\mathcal{X}$ and $\mathcal{Y}'$ are the weak-* topology and sup-norm-bounded weak topology respectively.

An adjoint pair of affine maps $\mathcal{A}$ and $\mathcal{A}'$ may be defined as

$$\begin{aligned} \mathcal{A}(\boldsymbol{\mu}) &= [\mu_p + \mu_c - \mathcal{L}^\dagger \mu - \delta_{t_0} \otimes \mu_0,\ \langle 1, \mu_0 \rangle] \\ \mathcal{A}^*(\boldsymbol{\ell}) &= [\gamma - v(t_0, \bullet),\ -\mathcal{L}_v,\ v,\ v], \end{aligned} \tag{36}$$

with vectors describing the cost and constraint terms as

$$\mathbf{b} = [0,\ 1] \tag{37a}$$

$$\mathbf{c} = [0,\ 0,\ 0,\ 1]. \tag{37b}$$

Pairings with the vectors in (37) satisfy

$$\langle \mathbf{c}, \boldsymbol{\mu} \rangle = \langle 1, \mu \rangle \tag{38a}$$

$$\langle \boldsymbol{\ell}, \mathbf{b} \rangle = \gamma. \tag{38b}$$

Problem (16) may be expressed as

$$p^*(t_0, X_0) = \sup_{\boldsymbol{\mu} \in \mathcal{X}_+} \langle \mathbf{c}, \boldsymbol{\mu} \rangle \qquad\qquad \mathbf{b} - \mathcal{A}(\boldsymbol{\mu}) \in \mathcal{Y}_+. \tag{39}$$

Similarly, the function LP in (18) can be expressed as

$$d^*(t_0, X_0) = \inf_{\boldsymbol{\ell} \in \mathcal{Y}'_+} \langle \boldsymbol{\ell}, \mathbf{b} \rangle \qquad\qquad \mathcal{A}'(\boldsymbol{\ell}) - \mathbf{c} \in \mathcal{X}_+. \tag{40}$$

Sufficient conditions to prove strong duality between (39) and (40) (by Theorem 2.6 of [42]) are:

R1 All feasible measures $\boldsymbol{\mu} \in \mathcal{X}_+$ with $b - \mathcal{A}(\boldsymbol{\mu}) \in \mathcal{Y}_+$ are bounded, and there exists such a feasible $\boldsymbol{\mu}$.

R2 All functions used to define $\mathbf{c}$, $\mathbf{b}$, $\mathcal{A}$ are continuous.

Boundedness of measures in R1 is proven in Lemma 5.1, and feasibility of a measure solution is demonstrated by the construction process used in the proof of Theorem 3.3. For R2, we note that both $\mathbf{c}$ and $\mathbf{b}$ in (37) are constant and are therefore continuous. Additionally, the set $\mathcal{C}$ is specifically chosen as the preimage of continuous functions under $\mathcal{L}$. R1 and R2 are both satisfied under assumptions A1-A6, thus proving strong duality.

B Superlevel Approximation of Risk Map

This appendix provides the proof of Theorem 4.2.

B.1 Supermartingale Property

We first introduce the notion of supermartingales, and then note that any feasible v for (21b)-(21e) is a supermartingale.

Definition B.1. *A process $\{Y_t\}$ is a supermartingale if $\mathbb{E}[Y_{t+\Delta t} \mid \{Y_t\}] \leq Y_t$ [29].*

Proposition B.1. *Any v that satisfies* (21b) *and* (21e) *is a supermartingale for the process $\mathcal{L}$ [28].*

Proof. Let $\{X_t\}$ be the state-dependent distribution for a trajectory of the stochastic process $\mathcal{L}$ in (1b), respecting exit time $t^* \wedge \tau_X$. This set of probability distributions satisfies the Martingale property of (7) (discrete-time) or (8) (continuous-time) as appropriate. Given that $\mathcal{L}v \leq 0$ from (21), it holds that $\mathbb{E}[\mathcal{L}_0 v(t,x) \mid X_{s'}] \leq 0$ for every stopping time s' adapted to $t^* \wedge \tau_X$. The supermartingale relation $\mathbb{E}[v(t+s,x) \mid X_{t+s}] \leq v(t,x)$ therefore holds for v. $\square$

The following lemma uses the previous supermartingale criterion from Proposition B.1 in combination with Doob's supermartingale inequality to provide a probability bound on v.

Lemma B.2 (based on Lemma 6 of [10])**.** *Let $v(t,x) \in \mathcal{C}$ be a nonnegative function over $[t_0, T] \times X$ and also form a supermartingale with respect to $\mathcal{L}$ and X_t. For a value $\lambda \geq 0$, an initial point $x_0 \in X$, and a $\mathcal{L}$-trajectory $x(t)$ starting from x_0 at time t_0, the following inequality holds:*

$$Prob_{\{X_t\}}\left[\sup_{t\in[t_0,T]} v(t,x(t)) \geq \lambda\right] \leq v(t_0,x_0)/\lambda. \tag{41}$$

Lemma B.2 modifies Lemma 6 of [10] by stopping at time T and allowing for time-dependent functions $v(t,x)$.

B.2 Superlevel Property

Theorem 4.2 can now be proven based on arguments from Theorem 7 of [10]. By constraint (21d), the unsafe set X_u is inside the 1-superlevel set of $v(t,x)$ at every time $t \in [t_0, T]$. This superlevel relation implies that

$$\begin{aligned} P(t_0,x_0) &= \sup_{t^*\in[t_0,T]} \mathrm{Prob}_{\{X_{t^*}\}}\left[x(t^*) \in X_u \mid x_0\right] \qquad (42)\\ &\leq \mathrm{Prob}_{\{X_t\}}\left[\sup_{t\in[t_0,T]} v(t,x(t)) \geq 1 \mid x_0\right]. \end{aligned}$$

Through Lemma B.2 with $\lambda = 1$, it holds that

$$\mathrm{Prob}_{\{X_t\}}\left[\sup_{t\in[t_0,T]} v(t,x(t)) \geq 1 \mid x_0\right] \leq v(t_0,x_0), \tag{43}$$

thus proving the theorem.

C Risk Contour Convergence

This appendix will provide a proof of Theorem 4.4.

We first notice that

$$\begin{aligned} 0 \leq p^*(t_0, x_0) &= \sup\{\langle 1, \mu_p\rangle \mid \mu_p + \mu_c = \delta_{(t_0,x_0)} + \mathcal{L}^\dagger \mu\} \\ &= \inf\{v(t_0, x_0) \mid \mathcal{L}v \leq 0 \ ; \ v \geq I_{X_u}\} \end{aligned}$$

by applying Theorem 4.2 (using A1-A6) to the case where $\mu_0 = \delta_{x_0}$ (to have strong duality and thus access the function interpretation of the problem). We remark that our v_k is feasible for the dual problem on functions (21), so by optimality of $p^*(t_0, x_0)$ for each $x_0 \in X_0$, one has $v_k(t_0, x_0) \geq p^*(t_0, x_0)$. Moreover, as $J^*(t_0, \mu_0) \leq 1 < \infty$ by Lemma 4.1, each $v_k(t_0, \bullet)$ is μ_0-integrable, and thus the function $v^\star(t_0, \bullet)$ is μ_0-integrable as well. It remains to prove that their difference $v^k(t_0, \bullet) - p^*(t_0, \bullet)$ ultimately vanishes in a μ_0-sense. This vanishing difference is proven by contradiction: suppose that there exists a $\eta > 0$ such that for all $k \in \mathbb{N}$, there is a $\varphi(k) \geq k$ with $\langle v_{\varphi(k)}(t_0, \bullet) - p^*(t_0, \bullet), \mu_0\rangle \geq \eta$ (which is the negation of (23)). Then, one has

$$\langle p^*(t_0, \bullet), \mu_0\rangle \leq \langle v_{\varphi(k)}(t_0, \bullet), \mu_0\rangle - \eta \underset{k\to\infty}{\longrightarrow} J^*(t_0, \mu_0) - \eta.$$

However, as we already noticed, $p^*(t_0, x_0)$ is feasible for (21), so this is a contradiction with optimality of the value $J^*(t_0, \mu_0)$ (given that $J^* - \eta < J^*$). Such a contradiction proves (23), and therefore proves Theorem 4.4.

References

[1] M. Prandini, J. Hu, C. Cassandras, and J. Lygeros, "Stochastic reachability: Theory and numerical approximation," *Stochastic hybrid systems, Automation and Control Engineering Series*, vol. 24, pp. 107–138, 2006.

[2] P. Mohajerin Esfahani, D. Chatterjee, and J. Lygeros, "The stochastic reach-avoid problem and set characterization for diffusions," *Automatica*, vol. 70, pp. 43–56, 2016.

[3] N. Schmid and J. Lygeros, "Probabilistic reachability and invariance computation of stochastic systems using linear programming," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 11 229–11 234, 2023, 22nd IFAC World Congress.

[4] S. Asmussen and H. Albrecher, *Ruin Probabilities*. World scientific, 2010, vol. 14.

[5] J. Miller, M. Tacchi, M. Sznaier, and A. Jasour, "Peak Value-at-Risk Estimation for Stochastic Processes using Occupation Measures," 2023, arXiv:2303.16064.

[6] M. J. Cho and R. H. Stockbridge, "Linear Programming Formulation for Optimal Stopping Problems," *SIAM J. Control Optim.*, vol. 40, no. 6, pp. 1965–1982, 2002.

[7] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, 2008.

[8] N. Kariotoglou, M. Kamgarpour, T. H. Summers, and J. Lygeros, "The Linear Programming Approach to Reach-Avoid Problems for Markov Decision Processes," *Journal of Artificial Intelligence Research*, vol. 60, pp. 263–285, 2017.

[9] V. A. Huynh, S. Karaman, and E. Frazzoli, "An incremental sampling-based algorithm for stochastic optimal control," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 2865–2872.

[10] S. Prajna, A. Jadbabaie, and G. J. Pappas, "Stochastic Safety Verification Using Barrier Certificates," in *2004 43rd IEEE conference on decision and control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 1. IEEE, 2004, pp. 929–934.

[11] A. Clark, "Control barrier functions for stochastic systems," *Automatica*, vol. 130, p. 109688, 2021.

[12] A. Salamati, A. Lavaei, S. Soudjani, and M. Zamani, "Data-driven safety verification of stochastic systems via barrier certificates," *IFAC-PapersOnLine*, vol. 54, no. 5, pp. 7–12, 2021.

[13] C. Santoyo, M. Dutreix, and S. Coogan, "A barrier function approach to finite-time stochastic system verification and control," *Automatica*, vol. 125, p. 109439, 2021.

[14] J. B. Lasserre, *Moments, Positive Polynomials And Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.

[15] J. Steinhardt and R. Tedrake, "Finite-time Regional Verification of Stochastic Nonlinear System," *The International Journal of Robotics Research*, vol. 31, no. 7, pp. 901–923, 2012.

[16] B. Xue, N. Zhan, and M. Fränzle, "Reach-avoid analysis for stochastic differential equations," *arXiv:2208.10752*, 2022.

[17] D. Drzajic, N. Kariotoglou, M. Kamgarpour, and J. Lygeros, "A Semidefinite Programming Approach to Control Synthesis for Stochastic Reach-Avoid Problems." in *ARCH@ CPSWeek*, 2016, pp. 134–143.

[18] E. Ahbe, A. Iannelli, and R. S. Smith, "Region of attraction analysis of nonlinear stochastic systems using polynomial chaos expansion," *Automatica*, vol. 122, p. 109187, 2020.

[19] A. Wang, A. Jasour, and B. C. Williams, "Non-Gaussian Chance-Constrained Trajectory Planning for Autonomous Vehicles Under Agent Uncertainty," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6041–6048, 2020.

[20] D. Henrion, M. Junca, and M. Velasco, "Moment-SOS hierarchy and exit time of stochastic processes," *arXiv preprint arXiv:2101.06009*, 2021.

[21] X. Chen, S. Chen, and V. M. Preciado, "Safety Verification of Nonlinear Polynomial System via Occupation Measures," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 1159–1164.

[22] J. Miller, D. Henrion, and M. Sznaier, "Peak Estimation Recovery and Safety Analysis," *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 1982–1987, 2021.

[23] J. Miller and M. Sznaier, "Bounding the Distance to Unsafe Sets with Convex Optimization," *IEEE Transactions on Automatic Control*, pp. 1–15, 2023.

[24] ——, "Quantifying the Safety of Trajectories using Peak-Minimizing Control," 2023, arXiv:2303.11896.

[25] C. Sloth and R. Wisniewski, "Safety Analysis of Stochastic Dynamical Systems," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 62–67, 2015, analysis and Design of Hybrid Systems ADHS.

[26] A. M. Jasour and B. C. Williams, "Risk Contours Map for Risk Bounded Motion Planning under Perception Uncertainties," in *Robotics: Science and Systems*, 2019.

[27] D. W. Stroock and S. S. Varadhan, *Multidimensional Diffusion Processes.* Springer Science & Business Media, 1997, vol. 233.

[28] L. C. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales: Volume 1, foundations.* Cambridge university press, 2000, vol. 1.

[29] D. Williams, *Probability with Martingales.* Cambridge university press, 1991.

[30] D. Hilbert, "Über die Darstellung definiter Formen als Summe von Formenquadraten," *Mathematische Annalen*, vol. 32, no. 3, pp. 342–350, 1888.

[31] G. Blekherman, "There are Significantly More Nonnegative Polynomials than Sums of Squares," *Israel Journal of Mathematics*, vol. 153, no. 1, pp. 355–380, 2006.

[32] M.-D. Choi, T. Y. Lam, and B. Reznick, "Sums of squares of real polynomials," in *Proceedings of Symposia in Pure mathematics*, vol. 58. American Mathematical Society, 1995, pp. 103–126.

[33] J. Cimprič, M. Marshall, and T. Netzer, "Closures of quadratic modules," *Israel Journal of Mathematics*, vol. 183, no. 1, pp. 445–474, 2011.

[34] M. Putinar, "Positive Polynomials on Compact Semi-algebraic Sets," *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.

[35] J. Nie and M. Schweighofer, "On the complexity of Putinar's Positivstellensatz," *Journal of Complexity*, vol. 23, no. 1, pp. 135–150, 2007.

[36] D. Henrion, J. B. Lasserre, and C. Savorgnan, "Approximate Volume and Integration for Basic Semialgebraic Sets," *SIAM review*, vol. 51, no. 4, pp. 722–743, 2009.

[37] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A Framework for Worst-Case and Stochastic Safety Verification Using Barrier Certificates," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415–1428, 2007.

[38] M. Tacchi, "Convergence of Lasserre's hierarchy: the general case," *Optimization Letters*, vol. 16, no. 3, pp. 1015–1033, 2022.

[39] J. Miller, T. Dai, and M. Sznaier, "Data-Driven Superstabilizing Control of Error-in-Variables Discrete-Time Linear Systems," in *2022 61st IEEE Conference on Decision and Control (CDC)*, 2022, pp. 4924–4929.

[40] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB," in *ICRA (IEEE Cat. No.04CH37508)*, 2004, pp. 284–289.

[41] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 10.1.*, 2023.

[42] M. Tacchi, "Moment-sos hierarchy for large scale set approximation. application to power systems transient stability analysis," Ph.D. dissertation, Toulouse, INSA, 2021.