# Peak Value-at-Risk Estimation for Stochastic Differential Equations using Occupation Measures

Jared Miller, Matteo Tacchi, Mario Sznaier, Ashkan Jasour

# Peak Value-at-Risk Estimation for Stochastic Differential Equations using Occupation Measures

Jared Miller[1], Matteo Tacchi[2], Mario Sznaier[1], Ashkan Jasour[3]

March 28, 2023

**Abstract**

This paper proposes an algorithm to upper-bound maximal quantile statistics of a state function over the course of a Stochastic Differential Equation (SDE) system execution. This chance-peak problem is posed as a nonconvex program aiming to maximize the Value-at-Risk (VaR) of a state function along SDE state distributions. The VaR problem is upper-bounded by an infinite-dimensional Second-Order Cone Program in occupation measures through the use of one-sided Cantelli or Vysochanskii-Petunin inequalities. These upper bounds on the true quantile statistics may be approximated from above by a sequence of Semidefinite Programs in increasing size using the moment-Sum-of-Squares hierarchy when all data is polynomial. Effectiveness of this approach is demonstrated on example stochastic polynomial dynamical systems.

**Acknowledgements**

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 in part supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States and by the International Student Exchange Program from AFOSR.
M. Tacchi was supported by the French company Réseau de Transport d'Électricité, as well as the Swiss National Science Foundation under the "NCCR Automation" grant n°51NF40_180545.

## 1 Introduction

This paper analyzes maximal $(1-\epsilon)$-quantile statistics of a state function $p(x)$ for Stochastic Differential Equation (SDE) trajectories evolving in a compact set $X$. An example of this type of quantile statistic for trajectory analysis is in establishing that there exists at least one time with a 1% chance of the aircraft exceeding a height of 100 meters. This task of quantile estimation is

[1]J. Miller, and M. Sznaier are with the Robust Systems Lab, ECE Department, Northeastern University, Boston, MA 02115. (e-mails: miller.jare@northeastern.edu, msznaier@coe.neu.edu).
[2]M. Tacchi is with Automatic Control Laboratory, EPFL, 1015 Lausanne, Switzerland, and 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]A. Jasour is with the Team 347T-Robotic Aerial Mobility, Jet Propulsion Lab, Pasadena, CA, 91109. (e-mail: jasour@jpl.caltech.edu).

related to peak and Value-at-Risk (VaR) estimation, and will also be referred to as the 'chance-peak' problem.

The $\epsilon$-VaR is the value at which there is an $\epsilon$-probability of exceedance [1]. Control and portfolio design typically aims to minimise the VaR. One specific VaR-upper-bounding coherent risk measure [2] that results in convex programs is the conditional VaR risk measure [3, 4]. The conditional VaR has been utilized for stochastic optimal control in [5], and for approximation of discrete-time risk-bounded sets using exponential and logarithmic inequalities with Markov Decision Processes in [6]. In contrast, the chance-peak approach upper-bounds maximum VaR of the continuous-time SDE state distribution of $x(t)$ across all times. We will solve this problem by maximizing the Cantelli [7] and Vysochanskij-Petunin (VP) [8] upper bounds for the VaR.

Chance constraints are an adjacent topic to VaR optimization, in which a probability inequality must hold as a hard constraint. Chance-constrained programs have a wide variety of application in control theory [9, 10, 11], and are generally intractable to solve explicitly. Approximation methods for chance constraints include the Cantelli [7] and VP [8] inequalities, and application of these tail-bounds in control include [12, 13]. The scenario approach for randomized constraint generation will converge in probability to the chance-constrained optimum, but carries a risk of failure and may require a large number of samples [14]. The moment-Sum of Squares (SOS) hierarchy of Semidefinite Programs (SDPs) will converge to the chance-constrained optimal solution under appropriate boundedness conditions [15].

The chance-peak problem is also related to a family of optimal stopping problems which can be solved using occupation measures. The work in [16] expressed optimal control problems of Ordinary Differential Equations (ODEs) as an infinite-dimensional Linear Program (LP) in an initial, terminal, and occupation measure. The peak estimation problem to maximize a state function $p(x)$ is an instance of optimal control with free terminal time and zero running cost. The work in [17] generalizes this LP to the stochastic case to find the maximum expectation of $p(x)$ when dynamics are phrased in terms of their infinitesimal generator (Feller process). Such LPs will converge to the true solution of the stopping problem under mild convergence, regularity, and well-posedness assumptions. The moment-SOS hierarchy of finite-dimensional SDPs will converge to the infinite-dimensional LP optimum if all problem data (e.g., dynamics, constraint sets) are polynomial-representable [18]. This convergent SDP approach has been used for optimal control [19], peak estimation [20] including compact-valued uncertainty [21], expectation-maximization of Lévy processes [22], and option pricing [23]. Other instances of the moment-SOS hierarchy used to solve stochastic safety problems include Barrier certificates [24], infinite-time averages [25], and Reach-Avoid sets [26].

The contributions of this paper are:

- An infinite-dimensional Second-Order Cone (SOCP) that upper-bounds the chance-peak program using VaR inequalities [7, 8]
- A convergent set of SDPs using the Moment-SOS hierarchy to the SOCP upper bound
- A verification of this approach on example polynomial SDE systems

This paper has the following structure: Section 2 gives an overview of notation, SDEs, and occupation measures. Section 3 upper-bounds the chance-peak problem using an infinite-dimensional SOCP in occupation measures. Section 4 reviews the moment-SOS hierarchy and presents a hierarchy of SDPs that approximate the infinte-dimensional chance-peak SOCP. Section 5 provides numerical examples of the chance-peak problem on ODE and SDE systems. Section 6 concludes the paper.

# 2 Preliminaries

## 2.1 Notation

The $n$-dimensional real Euclidean space is $\mathbb{R}^n$. The set of natural numbers is $\mathbb{N}$, and its subset of natural numbers between 1 and $N$ is $1..N$. The set of $n$-multi-indices is $\mathbb{N}^n$. The degree of a multi-index $\alpha \in \mathbb{N}^n$ is $|\alpha| = \sum_{i=1}^n \alpha_i$. A monomial is a term $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$ with degree $\deg x^\alpha = |\alpha|$. A polynomial $p(x) \in \mathbb{R}[x]$ may be uniquely represented in terms of multi-indices $\alpha$ and coefficients $p_\alpha$ as $p(x) = \sum_{\alpha \in \mathscr{J}} p_\alpha x^\alpha$ for some finite set $\mathscr{J} \in \mathbb{N}^n$. The degree of a vector of polynomials ($f \in (\mathbb{R}[x])^N$) is the maximum degree of any coordinate ($\deg f = \max_{i \in 1..N} \deg f_i$). The vector space of polynomials of degree at most $d$ is $\mathbb{R}[x]_{\leq d}$ and its dimension is $\binom{n+d}{d}$. The Second-Order Cone (SOC) (or Lorentz cone) is $Q^n = \{(s, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ : \ \kappa \geq \|s\|_2\}$, where $\|s\|_2 = (s_1^2 + \ldots + s_n^2)^{1/2}$ denotes the Euclidean norm.

The vector space of continuous functions over a topological space $X$ is $C(X)$, and its nonnegative subcone is $C_+(X)$. The topological dual of a Banach space $V$ is $V^*$. The cone of (nonnegative) Borel measures supported over $X$ is $\mathcal{M}_+(X)$ and the vector space of signed Borel measures supported on $X$ is $\mathcal{M}(X) = \mathcal{M}_+(X) - \mathcal{M}_+(X)$. When $X$ is compact, $C(X)$ and $\mathcal{M}(X)$ are topological dual spaces that have a duality product by Lebesgue integration: for $f \in C(X),\ \mu \in \mathcal{M}(X)$ the duality product is $\langle f, \mu \rangle = \int_X f(x) d\mu(x)$. This duality product also induces a duality pairing between $C_+(X)$ and $\mathcal{M}_+(X)$. As a slight abuse of notation, we extend this duality product to all Borel measurable functions $f$: $\langle f, \mu \rangle = \int_X f(x) d\mu(x)$. The set of $k$-times continuously differentiable functions over $X$ is $C^k(X)$.

The indicator function of a set $A \subseteq X$ is $I_A : X \to \{0, 1\}$, and has the values $I_A(x) = 0$ for $x \notin A$ and $I_A(x) = 1$ for $x \in A$. The measure of $A$ with respect to $\mu \in \mathcal{M}_+(X)$ is defined as $\mu(A) = \langle I_A, \mu \rangle$. The mass of a measure $\mu \in \mathcal{M}_+(X)$ is $\mu(X) = \langle 1, \mu \rangle$, and $\mu$ is a probability measure if this mass is 1. The support of a measure $\mu$ is the set of all points $x$ such that each open neighborhood $N_x$ of $x$ obeys $\mu(N_x) > 0$. The Dirac delta $\delta_{x'}$ supported only at the point $x'$ is a probability measure such that $\langle f, \delta_{x'} \rangle = f(x')$ for all $f \in C(X)$. Given two measures $\mu \in \mathcal{M}_+(X)$ and $\nu \in \mathcal{M}_+(Y)$, the product measure $\mu \otimes \nu$ is the unique measure that satisfies $\forall A \subseteq X,\ B \subseteq Y :\ (\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$. The pushforward of a function $Q : X \to Y$ along a measure $\mu(x)$ is $Q_\# \mu(y)$ and satisfies the relation $\forall g \in C(Y) :\ \langle g(y), Q_\# \mu(y) \rangle = \langle g(Q(x)), \mu(x) \rangle$.

The operator $\wedge$ will be used to denote the minimum of two quantities (stopping times) as $a \wedge b = \min(a, b)$. The adjoint of a linear operator $\mathcal{L} : X \to Y$ is $\mathcal{L}^\dagger : Y^* \to X^*$.

## 2.2 Probability Tail Bounds and Value-at-Risk

Let $\xi$ be a univariate probability measure $\xi(\omega) \in \mathcal{M}_+(\mathbb{R})$ for a coordinate $\omega \in \mathbb{R}$, with $\langle 1, \xi \rangle = 1$ and $|\langle \omega, \xi \rangle|, \langle \omega^2, \xi \rangle < \infty$ (finite first and second moments). In this paper, we define the $\epsilon$-VaR of $\xi$ as follows:

$$VaR_\epsilon(\xi) = \sup \{\lambda \in \mathbb{R} \mid \xi([\lambda, \infty)) \geq \epsilon\}. \tag{1}$$

Let $\sigma^2 = \langle \omega^2, \xi \rangle - \langle \omega, \xi \rangle^2$ be the variance of the probability distribution $\xi$.

The Cantelli bound for VaR is [7]

$$VaR_\epsilon(\xi) \leq \sigma\sqrt{1/(\epsilon)-1} + \langle \omega, \xi \rangle = VaR_\epsilon^{cant}(\xi). \tag{2a}$$

The VP bound for the VaR is [8]

$$VaR_\epsilon(\xi) \leq \sigma\sqrt{4/(9\epsilon)-1} + \langle \omega, \xi \rangle = VaR_\epsilon^{VP}(\xi). \tag{2b}$$

The Cantelli bound is applicable for any probability distribution $\xi(\omega)$ and value $\epsilon \in [0,1]$. The VP bound is sharper than the Cantelli bound, but is only valid when $\xi$ is unimodal and $\epsilon \leq 1/6$.

### 2.3 Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space with time-indexed filtration $\mathcal{F}_t$, $X \subset \mathbb{R}^n$ be a compact set, and $w$ be $n$-dimensional Wiener process. An Itô SDE with a drift function $f$ and diffusion function $g$ is [27]

$$dx = f(t,x)dt + g(t,x)dw. \tag{3}$$

In this paper, trajectories will start from an initial set $X_0 \subseteq X$ and will remain within $X$ in times $t \in [0,T]$ by virtue of stopping at the boundary $\partial X$. Define $\tau_X$ as a stopping time (random variable) corresponding to the time at which the process (3) starting from $X_0$ first touches the boundary $\partial X$ for the first time. A process of (3) starting from an initial condition $x(0) \in X_0$ in times $t \in [0,T]$ is

$$x(t) = x(0) + \int_{t=0}^{\tau_X \wedge T} f(t,x)dt + \int_{t=0}^{\tau_X \wedge T} g(t,x)dw. \tag{4}$$

Solutions of (4) are unique if there exists finite constants $C, D > 0$ such that for all $(t,x,x') \in [0,T] \times X^2$, the following Lipschitz and Growth conditions hold [28]:

$$\begin{aligned} D\|x - x'\|_2 &\geq \|f(t,x) - f(t,x')\|_2 + \|g(t,x) - g(t,x')\|_2 \\ C(1 + \|x\|_2) &\geq \|f(t,x)\|_2 + \|g(t,x)\|_2. \end{aligned} \tag{5}$$

The Lipschitz and Growth conditions will hold if $(f,g)$ are locally Lipschitz and the set $X$ is compact. Distributions of the densities of (4) may be computed by solving a Fokker-Planck equation with absorbing boundary conditions on $\partial X$ [29, 30].

The generator $\mathcal{L}$ associated with the SDE is a linear operator that satisfies $\forall v(t,x) \in C^2([0,T] \times X)$ [28]:

$$\mathcal{L}v(t,x) = \partial_t v + f(t,x) \cdot \nabla_x v + \frac{1}{2} g(t,x)^T \left(\nabla^2_{xx} v\right) g(t,x). \tag{6}$$

The $\nabla^2_{xx} v$ term arises from the Itô Lemma. Let $\tau$ be a stopping time adapted to the filtration, defined by $\tau = \tau_X \wedge T$. The occupation measure $\mu \in \mathcal{M}_+([0,T] \times X)$ corresponding to the stopping time $\tau$, initial distribution $\mu_0 \in \mathcal{M}_+(X_0)$, and dynamics (3) is $\forall A \subseteq [0,T],\ B \subseteq X$ is

$$\mu(A \times B) = \int_{X_0} \int_{t=0}^{\tau} I_{A \times B}\left((t, x(t \mid x_0))\right) dt\, d\mu_0(x_0). \tag{7}$$

The initial measure $\mu_0 \in \mathcal{M}_+(X_0)$, the occupation measure $\mu$ from (7), and the terminal measure $\mu_\tau \in \mathcal{M}_+([0,T] \times X)$ defined by following the SDE (3) from initial conditions $x_0 \sim \mu_0$ until the stopping time $\tau$, are all related by Dynkin's formula [31]

$$\langle v, \mu_\tau \rangle = \langle v(0,x), \mu_0(x) \rangle + \langle \mathcal{L}v, \mu \rangle \qquad \forall v \in C^2. \tag{8}$$

Dynkin's formula is an SDE generalization of the Liouville equation for ODEs. Equation (8) may be equivalently written in weak form as

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu. \tag{9}$$

An expectation-maximizing optimal stopping problem for the SDE in (3) with a reward function of $p(x)$ in the region $[0, T] \times X$, when starting at the initial condition $x(0) \sim \mu_0 \in \mathcal{M}_+(X_0)$, is $P^* = \sup \mathbb{E}_{\mu_0}[p(x(\tau))]$. The work in [17] presents an infinite-dimensional LP in measures to solve this stopping problem

$$\begin{aligned}
p^* = \ \sup \quad & \langle p, \mu_\tau \rangle && \text{(10a)}\\
& \mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu && \text{(10b)}\\
& \langle 1, \mu_0 \rangle = 1 && \text{(10c)}\\
& \mu, \mu_\tau \in \mathcal{M}_+([0, T] \times X) && \text{(10d)}\\
& \mu_0 \in \mathcal{M}_+(X_0). && \text{(10e)}
\end{aligned}$$

Any $\mu$ that is part of a feasible solution $(\mu, \mu_0, \mu_\tau)$ for (10b)-(10e) will be referred to as a *relaxed* occupation measure. Program (10) satisfies $p^* \geq P^*$, and tightness ($p^* = P^*$) is achieved under the assumptions of Lipschitz continuity and Growth (5), compactness of $[0, T] \times X$, and continuity of $p(x)$.

# 3 Peak Value-at-Risk Estimation

This section will present the chance-peak problem statement, and will also derive the infinite-dimensional SOCP to upper bound the chance-peak quantile statistic.

## 3.1 Problem Statement

Let $\epsilon \in [0, 1]$ be a value for the quantile statistic, $X$ be a compact set, $X_0 \subseteq X$ be a set of initial conditions, and (4) be the solution to an SDE evolving from $x(0) \in X_0$ that remains within $X$ until it stops. For a given initial probability distribution $\mu_0 \in \mathcal{M}_+(X_0)$, and for all $t \in [0, T]$, let $x(t)$ be the stochastic process of (4) at time $t$, and let $\mu_t \in \mathcal{M}_+(X)$ be its probability distribution (with $x(t)$ stopping at $\partial X$).

### 3.1.1 Assumptions

The following assumptions will be posed throughout this paper,

A1 The set $[0, T] \times X$ is compact and $X_0 \subseteq X$.

A2 The functions $(f, g)$ satisfy (5).

A3 The state function $p(x)$ is continuous on $X$.

A4 The initial measure $\mu_0 \in \mathcal{M}_+(X_0)$ is a given probability distribution ($\langle 1, \mu_0 \rangle = 1$).

#### 3.1.2 VaR Problem

**Problem 3.1.** *The chance-peak problem to find the $\epsilon$-VaR of $p(x)$ is*

$$P^* = \sup_{t^* \in [0,T]} VaR_\epsilon(p_\# \mu_{t^*}) \tag{11a}$$
$$dx = f(t,x)dt + g(t,x)dw \tag{11b}$$
$$\text{from } t = 0 \text{ until a stopping time of } \tau_X \wedge t^* \tag{11c}$$
$$x(0) \sim \mu_0. \tag{11d}$$

The pushforward $p_\# \mu_{t^*}$ from (11a) is the univariate probability distribution of $p(x)$ at the state distribution $x \sim \mu_{t^*}$.

#### 3.1.3 Tail-Bound Upper Bound

Let $r$ be the constant factor multiplying $\sigma$ in (2) such that

$$r^{cant} = \sqrt{1/(\epsilon) - 1} \qquad\qquad r^{VP} = \sqrt{4/(9\epsilon) - 1}. \tag{12}$$

It is further assumed that the VP-bound will only be used if its conditions are satisfied ($\epsilon \leq 1/6$, unimodal). The distribution of $p(x)$ with respect to the state distribution $\mu_{t^*}$ is univariate, for which the relation in (1) and the constants in (2) can be used to upper-bound on Problem 3.1. We will use the notation $\langle p^2, \mu_{t^*} \rangle$ to refer to $\langle p(x)^2, \mu_{t^*}(x) \rangle$.

**Problem 3.2.** *The tail-bound program that upper-bounds the chance-peak* (11) *with constant $r$ is*

$$P_r^* = \sup_{t^* \in [0,T]} r\sqrt{\langle p^2, \mu_{t^*} \rangle - \langle p, \mu_{t^*} \rangle^2} + \langle p, \mu_{t^*} \rangle \tag{13a}$$
$$dx = f(t,x)dt + g(t,x)dw \tag{13b}$$
$$\text{from } t = 0 \text{ until a stopping time of } \tau_X \wedge t^* \tag{13c}$$
$$x(0) \sim \mu_0. \tag{13d}$$

### 3.2 Nonlinear Measure Program

Problem 3.2 can be upper-bounded by an infinite-dimensional nonlinear program in a given initial probability distribution $\mu_0$, terminal measure $\mu_\tau$, and relaxed occupation measure $\mu$, using the generator $\mathcal{L}$ in (6) as

$$p_r^* = \sup r\sqrt{\langle p^2, \mu_\tau \rangle - \langle p, \mu_\tau \rangle^2} + \langle p, \mu_\tau \rangle \tag{14a}$$
$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \tag{14b}$$
$$\mu_\tau, \ \mu \in \mathcal{M}_+([0,T] \times X). \tag{14c}$$

**Theorem 3.3.** *Program 14 is an upper bound on* (13) *with $p_r^* \geq P_r^*$ under A1-A4.*

*Proof.* Let $t^*$ be a stopping time in $[0,T]$, and let $x_0 \in X_0$ be an initial condition. Measures $(\mu_0, \mu, \mu_\tau)$ that satisfy (14b) may be constructed from this $(t^*, x_0)$ by $\mu_{t^*}$ as the the state distribution of (4) at time $t^*$ given $\mu_0$, and $\mu$ as the occupation measure in (7) associated to this SDE trajectory with distribution $\mu_0$. Because the feasible set to constraint (14b) contains measures induced by all possible provided SDE trajectories starting from $\mu_0$, it holds that $p_r^* \geq P_r^*$. ∎

**Remark 1.** *The initial distribution* $\mu_0 \in \mathcal{M}_+(X_0)$ *may be optimized to find a supremal* $p_r^*$ *over all probability distributions in* $X_0$ *by adding* $\mu_0$ *as a variable and adding the constraint* $\langle \mu_0, 1 \rangle = 1$ *to* (14).

## 3.3 Measure Second-Order Cone Program

The nonlinear measure program (14) may be recast as an infinite-dimensional convex SOCP.

**Lemma 3.4.** *Let* $J_r(a, b) = r\sqrt{b - a^2} + a$ *be the objective* (14a) *with* $a = \langle p(x), \mu_\tau \rangle$ *and* $b = \langle p(x)^2, \mu_\tau \rangle$. *For any convex set* $C \in \mathbb{R} \times \mathbb{R}_+$ *with* $(a, b) \in C$, *the following pair of programs have the same optimal value (in which* $Q^3 = \{([s_1, s_2, s_3], \kappa) \in \mathbb{R}^3 \times \mathbb{R}_+ \mid \|s\|_2 \leq \kappa\}$ *is an SOC cone):*

$$\sup_{(a,b)\in C} \quad a + r\sqrt{b - a^2} \tag{15}$$

$$\sup_{(a,b)\in C,\ z\in\mathbb{R}} \quad a + rz : \ ([1 - b, 2z, 2a], 1 + b) \in Q^3. \tag{16}$$

*Proof.* The new variable $z$ is introduced under the constraint $\sqrt{b - a^2} \geq z$, implying that $z^2 + a^2 \leq b$. The SOCP equivalence follows from the power-representation of $\sqrt{b - a^2}$ from [32, 33], with the steps of

$$([1 - b, 2z, 2a], 1 + b) \in Q^3 \tag{17a}$$
$$(1 - b)^2 + 4(z^2 + a^2) \leq (1 + b)^2 \tag{17b}$$
$$(1 + b^2) - 2b + 4(z^2 + a^2) \leq (1 + b^2) + 2b \tag{17c}$$
$$4(z^2 + a^2) \leq 4b. \tag{17d}$$

□

**Theorem 3.5.** *An infinite-dimensional SOCP with the same optimal value and set of feasible solutions as* (14) *given* $\mu_0$ *is*

$$p_r^* = \sup \quad rz + \langle p, \mu_\tau \rangle \tag{18a}$$
$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \tag{18b}$$
$$u = [1 - \langle p^2, \mu_\tau \rangle,\ 2z,\ 2\langle p, \mu_\tau \rangle] \tag{18c}$$
$$(u, 1 + \langle p^2, \mu_\tau \rangle) \in Q^3 \tag{18d}$$
$$\mu,\ \mu_\tau \in \mathcal{M}_+([0, T] \times X), z \in \mathbb{R}, u \in \mathbb{R}^3. \tag{18e}$$

*Proof.* This results from an application of Lemma 3.4 to the objective term (14a). The optimization variables are now $(\mu_\tau, \mu, z, u)$. □

**Corollary 1.** *Program* (18) *is convex.*

*Proof.* The objective (18a) is affine in $(z, \mu_\tau)$. Constraints (18b)-(18e) are convex (affine for (18b) and SOC for (18d)), ensuring convexity of (18). □

**Remark 2.** *Problem* (18) *has an infinite-dimensional affine constraint in* (18b) *and a finite-dimensional SOC constraint in* (18d).

# 4 Finite Moment Program

This section will upper-bound (18) utilizing a converging hierarchy of SDPs of increasing size.

## 4.1 Review of Moment-SOS Hierarchy

All content from this subsection is referenced from [18]. For any multi-indexed sequence $\mathbf{m} = \{\mathbf{m}_\alpha\}_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, we define the Riesz functional $L_\mathbf{m}: \mathbb{R}[x] \to \mathbb{R}$ as follows:

$$p(x) = \sum_{\alpha\in\mathbb{N}^n} p_\alpha x^\alpha \quad \longmapsto \quad L_\mathbf{m} p = \sum_{\alpha\in\mathbb{N}^n} p_\alpha \mathbf{m}_\alpha. \tag{19a}$$

Let $\mu \in \mathcal{M}_+(X)$ be a measure. The $\alpha$-moment of $\mu$ for $\alpha \in \mathbb{N}^n$ is $\mathbf{m}_\alpha = \langle x^\alpha, \mu\rangle$. The collection of moments $\mathbf{m} = \{\mathbf{m}_\alpha\}_{\alpha\in\mathbb{N}^n}$ is a moment sequence and has the following property:

$$\forall p \in \mathbb{R}[x], \quad L_\mathbf{m} p = \langle p, \mu\rangle. \tag{19b}$$

A key result to build the moment-SOS hierarchy is the characterization of sequences $\mathbf{m}$ that correspond to moment sequences, i.e. such that (19b) holds for some $X \subset \mathbb{R}^n$ and some $\mu \in \mathcal{M}_+(X)$:

For $\mathbf{m} \in \mathbb{R}^{\mathbb{N}^n}$ and $h \in \mathbb{R}[x]$, we define the localizing bilinear functional $L_{h\mathbf{m}} : \mathbb{R}[x] \times \mathbb{R}[x] \longrightarrow \mathbb{R}$ by

$$L_{h\mathbf{m}} = (p, q) \longmapsto L_\mathbf{m}(hpq). \tag{20a}$$

Equipping $\mathbb{R}[x]$ with a linear basis $(e_i)_{i\in\mathbb{N}}$ (e.g. $e_i(x) = x^{\alpha_i}$ with $\{\alpha_i\}_{i\in\mathbb{N}} = \mathbb{N}^n$, an ordering of monomials such that $|\alpha_i| < |\alpha_j| \Rightarrow i < j$) yields an infinite size matrix representation of $L_{h\mathbf{m}}$, which we call the localizing matrix $\mathbb{M}[h\mathbf{m}] = (L_\mathbf{m}(h\, e_i\, e_j))_{i,j\in\mathbb{N}}$. For instance, if $h(x) = \sum_{\beta\in\mathbb{N}^n} h_\beta x^\beta$, using a basis of monomials with nondecreasing degrees yields, for all $i, j \in \mathbb{N}$

$$\mathbb{M}[h\mathbf{m}]_{i,j} = L_\mathbf{m}(hx^{\alpha_i}x^{\alpha_j}) = \sum_{\beta\in\mathbb{N}^n} h_\beta \mathbf{m}_{\alpha_i+\alpha_j+\beta}. \tag{20b}$$

A Basic Semialgebraic (BSA) set is a set defined by a finite number of bounded-degree inequality constraints such as $\mathbb{K} = \{x \mid h_k(x) \geq 0 : \; k = 1..N_c\}$. Assuming "ball constraints" $h_1(x) = 1$ and $h_{N_c}(x) = R - \|x\|_2^2$ (this can always be enforced if $\mathbb{K}$ is compact in $\mathbb{R}^n$, up to adding redundant constraints), $\mathbf{m} \in \mathbb{R}^{\mathbb{N}^n}$ has a representing measure $\mu \in \mathcal{M}_+(\mathbb{K})$ such that (19b) holds if, for all $k = 1..N_c$, the bilinear functional $L_{h_k\mathbf{m}}$ is positive semidefinite, i.e.

$$\forall p \in \mathbb{R}[x], k = 1..N_c, \quad L_\mathbf{m}(h_k p^2) \geq 0. \tag{21a}$$

or, equivalently,

$$\forall d \in \mathbb{N}, k = 1..N_c, \quad \mathbb{M}_d[h_k\mathbf{m}] \succeq 0. \tag{21b}$$

where $\mathbb{M}_d[h_k\mathbf{m}]$ is the top left block of size $\binom{n+d}{d}$ of $\mathbb{M}[h_k\mathbf{m}]$, which corresponds to the matrix representation of $L_{h\mathbf{m}}$ in the finite dimensional space $\mathbb{R}[x]_{\leq d}$.

For notational convenience, we define the block diagonal synthetic matrix

$$\mathbb{M}_d[\mathbb{K}\mathbf{m}] = \mathrm{diag}(\mathbb{M}_{d-\lceil d_k/2\rceil}[h_k\mathbf{m}])_{k=1..N_c} \tag{22}$$

where $d_k = \deg(h_k)$. This synthetic matrix has two important properties, deduced from (20b) and (21b):

- it exactly involves all the terms $\mathbf{m}_\alpha$ for $|\alpha| \leq 2d$
- (21) holds if and only if $\mathbb{M}_d[\mathbb{K}\mathbf{m}] \succeq 0$ for all $d \in \mathbb{N}$.

The process of increasing the degree $d \to \infty$ when posing Positive Semidefinite (PSD) constraints on $\mathbb{M}_d[\mathbb{K}\mathbf{m}]$ is called the moment-SOS hierarchy.

## 4.2 Moment Program

The following assumptions are required to utilize the moment-SOS hierarchy in approximating (18):

A5 The sets $X_0$ and $X$ are BSA sets with ball constraints.

A6 The functions $f(t,x), g(t,x)$ are polynomial vector fields and $p(x)$ is a polynomial.

Given an initial measure $\mu_0 \in \mathcal{M}_+(X_0)$, let $(\mathbf{m}, \mathbf{m}^\tau)$ be moment sequences corresponding to the measures $(\mu, \mu_\tau)$ respectively. For each monomial $x^\alpha t^\beta$ with $\alpha \in \mathbb{N}^n, \beta \in \mathbb{N}$, define the operator $\mathcal{D}_{\alpha\beta}(\mathbf{m}, \mathbf{m}^\tau)$ as the moment counterpart of the operator involved in Dynkin's formula (9)

$$\mathcal{D}_{\alpha\beta}(\mathbf{m}, \mathbf{m}^\tau) = \mathbf{m}^\tau_{\alpha\beta} - L_{\mathbf{m}}(\mathcal{L}(x^\alpha t^\beta)). \tag{23}$$

Define the dynamics degree $D$ as

$$D = d + \lceil \max(\deg f - 1, 2\deg g - 2)/2 \rceil. \tag{24}$$

so that for $(\alpha, \beta) \in \mathbb{N}^{n+1}$,

$$|\alpha| + \beta \leq 2d \Rightarrow \deg(\mathcal{L}(x^\alpha t^\beta)) \leq 2D.$$

**Problem 4.1.** *For $d \geq \deg(p)$, the order-$d$ moment problem that upper-bounds problem (18), given $\mu_0$ is*

$$\begin{aligned}
p^*_{r,d} = \max \quad & rz + L_{\mathbf{m}^\tau} p && \text{(25a)}\\
& z \in \mathbb{R}, \mathbf{m} \in \mathbb{R}^{\binom{2D+n+1}{n+1}}, \mathbf{m}^\tau \in \mathbb{R}^{\binom{2d+n+1}{n+1}} && \text{(25b)}\\
& \mathcal{D}_{\alpha\beta}(\mathbf{m}, \mathbf{m}^\tau) = \delta_{\beta 0}\langle x^\alpha, \mu_0\rangle \\
& \qquad \forall (\alpha, \beta) \in \mathbb{N}^{n+1} \quad s.t. \quad |\alpha| + \beta \leq 2d && \text{(25c)}\\
& s = [1 - L_{\mathbf{m}^\tau}(p^2),\ 2z,\ 2L_{\mathbf{m}^\tau} p] && \text{(25d)}\\
& (s, 1 + L_{\mathbf{m}^\tau}(p^2)) \in Q^3 && \text{(25e)}\\
& \mathbb{M}_d[([0,T] \times X)\mathbf{m}^\tau] \succeq 0 && \text{(25f)}\\
& \mathbb{M}_D[([0,T] \times X)\mathbf{m}] \succeq 0, && \text{(25g)}
\end{aligned}$$

where $\delta_{\beta 0}$ denotes the Kronecker symbol that is 1 if $\beta = 0$ and 0 otherwise. Note that constraint (25c) is a finite-dimensional truncation of the infinite-dimensional (18b).

The following boundedness result is required to ensure convergence:

**Lemma 4.2.** *All of $(\mu, \mu_\tau, z)$ are bounded in (18) under A1-A3.*

*Proof.* A sufficient condition for a measure to be bounded (in the sense that all moments are bounded) is that it has finite mass and is supported on a compact set. Compactness of $[0, T] \times X$ holds by A1. Assumption A4 imposes that $\langle 1, \mu_0 \rangle = 1$. By substituting $v(t, x) = 1$ (18b), it holds that $\langle 1, \mu_\tau \rangle = \langle 1, \mu_0 \rangle = 1$. Performing the same step with $v(t, x) = t$ yields $T \geq \langle t, \mu_\tau \rangle = \langle 1, \mu \rangle$. It therefore holds that $\langle p, \mu_\tau \rangle$ and $\langle p^2, \mu_\tau \rangle$ are bounded. The SOC constraint (18d) ensures that $z$ is finite, demonstrating that all variables are bounded. □

**Theorem 4.3.** *Under A1-A6, the optima in* (25) *will converge to* (18) *as* $\lim_{d\to\infty} p^*_{r,d} = p^*_r$.

*Proof.* This convergence will occur by Corollary 8 of [34] (when extending to the finite-dimensional SOC case) through Lemma 4.2. □

**Remark 3.** *The relation* $p^*_d \geq p^*_r \geq P^*_r$ *will still hold when* $[0, T] \times X$ *is noncompact (violating A1 and A5), but it may no longer occur that* $\lim_{d\to\infty} p^*_{r,d} = p^*_r$ *(the conditions Lemma 4.3 will no longer apply).*

### 4.3 Computational Complexity

In Problem (25), the computational complexity mostly depends on the number and size of the matrix blocks involved in LMI constraints (25f,25g), which in turn depend on the number and degrees of polynomial inequalities describing $X$ (the higher $d_k = \deg(h_k)$, the smaller $\mathbb{M}_{d-\lceil d_k/2\rceil}[h\mathbf{m}]$). At order-$d$, the maximum size of localizing matrices is $\binom{n+1+D}{D}$.

Problem (25) must be converted to SDP-standard form by introducing equality constraints between the entries of the moment matrices in order to utilize symmetric-cone Interior Point Methods (e.g., Mosek [35]). The per-iteration complexity of an SDP involving a single moment matrix of size $\binom{n+d}{d}$ scales as $n^{6d}$ [23]. The scaling of an SDP with multiple moment and localizing matrices generally depends on the maximal size of any PSD matrix. In our case, this size is at most $\binom{n+1+d}{d}$ with a scaling impact of $(n+1)^{6d}$. The complexity of using this chance-peak routine increases in a jointly polynomial manner with $d$ and $n$.

## 5 Numerical Examples

MATLAB (2022a) code to replicate experiments is available at https://github.com/Jarmill/chance_peak. Dependencies include Mosek [35] and YALMIP [36]. Monte Carlo (MC) sampling to empirically find VaR estimates is conducted over 50,000 SDE paths under antithetic sampling with a time spacing of $\Delta t = 10^{-3}$. All experiments contain a table of chance-peak bounds as well as solver-times to compute these bounds.

### 5.1 Two States

Example 1 of [37] is the following two-dimensional cubic polynomial SDE

$$dx = \begin{bmatrix} x_2 \\ -x_1 - x_2 - \frac{1}{2}x_1^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} dw. \tag{26}$$

This example performs chance-peak maximization of $p(x) = -x_2$ starting at the point (Dirac-delta initial measure $\mu_0$) $X_0 = [1, 1]$ with $X = [-1, 2] \times [-1, 1.5]$ and $T = 5$. Trajectories of (26) are

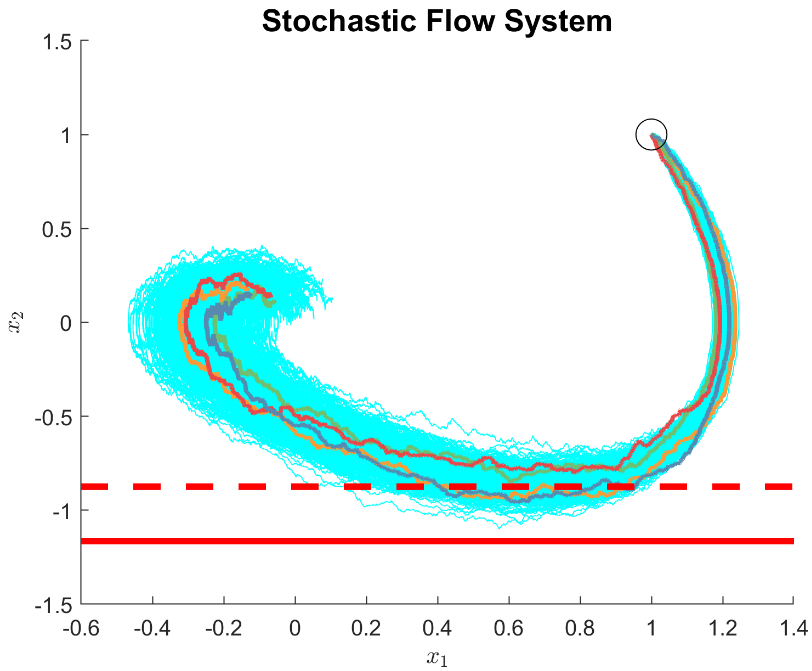


Figure 1: Trajectories of (26) with $\epsilon = 0.5$ (dashed red) and $\epsilon = 0.15$ (solid red) bounds

displayed in cyan in Figure 1 starting from the black-circle $X_0$, and four of these trajectories are marked in non-cyan colors. The $\epsilon = 0.5$ row of Table 1 displays the bounds on the mean distribution as solved through finite-degree SDP truncations of (10). The bounds at $\epsilon = \{0.15, 0.1, 0.05\}$ are obtained through the VP expression in (2b) and solving the SDPs obtained from (25). The dotted and solid red lines in Figure 1 are the $\epsilon = 0.5$ and $\epsilon = 0.15$ bounds respectively at order 5.

Table 1: Chance-Peak estimation of the Stochastic Flow System (26) to maximize $p(x) = -x_2$

| order | 2 | 3 | 4 | 5 | 6 | MC |
|---|---|---|---|---|---|---|
| $\epsilon = 0.5$ | 0.8817 | 0.8773 | 0.8747 | 0.8745 | 0.8744 | 0.8559 |
| $\epsilon = 0.15$ | 1.2210 | 1.1817 | 1.1689 | 1.1657 | 1.1642 | 0.9142 |
| $\epsilon = 0.1$ | 1.5009 | 1.4520 | 1.4361 | 1.4323 | 1.4303 | 0.9279 |
| $\epsilon = 0.05$ | 2.1306 | 2.0613 | 2.0387 | 2.0332 | 2.0305 | 0.9484 |

## 5.2 Three States

An SDE modification of the Twist system from [38] is,

$$dx = \begin{bmatrix} -2.5x_1 + x_2 - 0.5x_3 + 2x_1^3 + 2x_3^3 \\ -x_1 + 1.5x_2 + 0.5x_3 - 2x_2^3 - 2x_3^3 \\ 1.5x_1 + 2.5x_2 - 2x_3 - 2x_1^3 - 2x_2^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ 0.1 \end{bmatrix} dw. \tag{27}$$

Table 2: Solver time (seconds) to compute Table 1

| order | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $\epsilon = 0.5$ | 0.417 | 0.431 | 1.963 | 1.659 | 3.641 |
| $\epsilon = 0.15$ | 0.216 | 0.325 | 1.592 | 4.178 | 7.464 |
| $\epsilon = 0.1$ | 0.213 | 0.316 | 1.651 | 1.339 | 4.225 |
| $\epsilon = 0.05$ | 0.222 | 0.366 | 0.936 | 2.446 | 5.298 |

Table 3: Chance-Peak estimation of the Stochastic Twist System (27) to maximize $p(x) = x_3$

| order | 2 | 3 | 4 | 5 | 6 | MC |
|---|---|---|---|---|---|---|
| $\epsilon = 0.5$ | 0.9100 | 0.8312 | 0.8231 | 0.8211 | 0.8203 | 0.7206 |
| $\epsilon = 0.15$ | 1.1729 | 1.0813 | 1.0171 | 0.9891 | 0.9755 | 0.7685 |
| $\epsilon = 0.1$ | 1.4358 | 1.3180 | 1.2217 | 1.1859 | 1.1694 | 0.7801 |
| $\epsilon = 0.05$ | 2.0288 | 1.8497 | 1.6866 | 1.6299 | 1.5894 | 0.7970 |

This second example performs chance-peak maximization of $p(x) = x_3$ starting at the point $X_0 = [0.5, 0, 0]$ with $X = [-0.6, 0.6] \times [-1, 1] \times [-1, 1.5]$ and $T = 5$. VP bounds from solving the SDEs from (10) and (25) are recorded in Table 3 in the same manner as in Table 1. Figure 2 plots trajectories and bounds of (27) starting from the black-circle $X_0$ point, with four of these trajectories visibly distinguished. The solid red plane in Figure 2 is the $\epsilon = 0.15$ bound on $x_3$ at order 6, and the transluscent red plane is the $\epsilon = 0.5$ bound on $x_3$ (also at order 6).

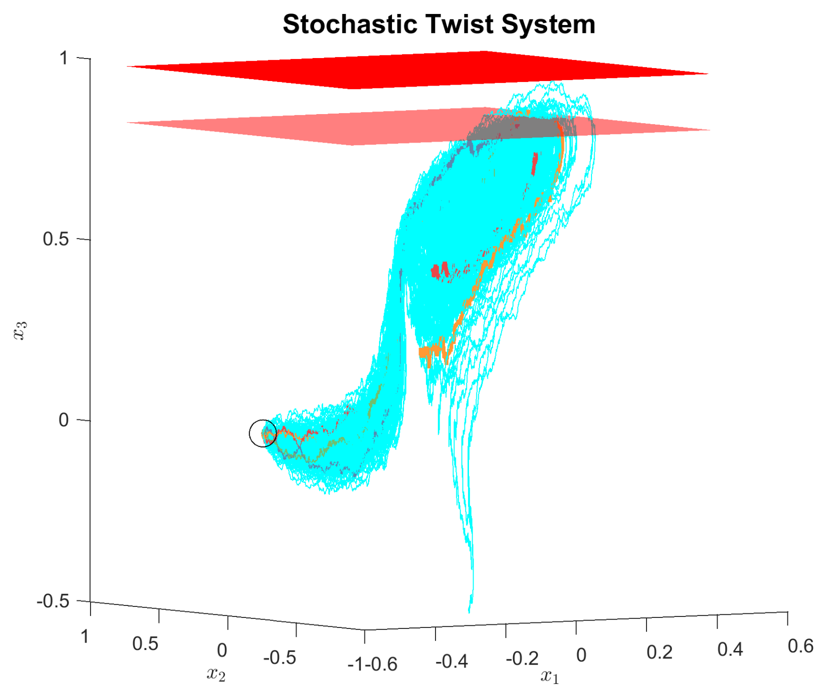


Figure 2: Trajectories of (27) with $\epsilon = 0.5$ (dashed red) and $\epsilon = 0.15$ (solid red) bounds

Table 4: Solver time (seconds) to compute Table 3

| order | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $\epsilon = 0.5$ | 0.526 | 1.880 | 5.213 | 28.269 | 81.736 |
| $\epsilon = 0.15$ | 0.397 | 1.117 | 3.876 | 23.302 | 132.539 |
| $\epsilon = 0.1$ | 0.210 | 1.080 | 6.313 | 27.829 | 116.733 |
| $\epsilon = 0.05$ | 0.185 | 1.352 | 4.999 | 20.863 | 137.833 |

# 6 Conclusion

This paper considered the chance-peak problem, which involved finding upper bounds on the quantiles of state functions $p(x)$ achieved by SDE systems. The true $(1-\epsilon)$-quantile statistic $P^*$ (11) is upper-bounded by the Cantelli/VP approximation $P_r^*$ (13), which in turn is upper bounded by an infinite-dimensional SOCP $p_r^*$ (18) and its moment-SOS finite-dimensional SDPs yielding $p_{r,d}^*$ with $\lim_{d\to\infty} p_{r,d}^* = p_d^*$. Each of these upper-bounds contribute valuable information towards the analysis of SDEs.

Future work includes finding conditions under which the measure-based upper-bounding does not add conservatism (e.g. cases where $p_r^* = P_r^*$), and utilizing higher-moment tail-probability inequalities to obtain closer estimates to the VaR [39]. Other work involves studying the duality structure of (18) with respect to the achievement of strong duality. The techniques introduced in this paper can also be extended to application domains such as distance estimation [38], Lévy processes [22], and exit-time statistics [19]. Another avenue involves developing stochastic optimal control strategies to minimize quantile statistics.

# References


[1] P. Jorion, Value at Risk. McGraw-Hill Professional Publishing, 2000.

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

[3] G. C. Pflug, "Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk," in Probabilistic constrained optimization. Springer, 2000, pp. 272–281.

[4] R. T. Rockafellar and S. Uryasev, "Conditional Value-at-Risk for General Loss Distributions," Journal of banking & finance, vol. 26, no. 7, pp. 1443–1471, 2002.

[5] C. W. Miller and I. Yang, "Optimal control of conditional value-at-risk in continuous time," SIAM Journal on Control and Optimization, vol. 55, no. 2, pp. 856–884, 2017.

[6] M. P. Chapman, R. Bonalli, K. M. Smith, I. Yang, M. Pavone, and C. J. Tomlin, "Risk-sensitive safety analysis using conditional value-at-risk," IEEE Transactions on Automatic Control, vol. 67, no. 12, pp. 6521–6536, 2021.

[7] F. P. Cantelli, "Sui confini della probabilita," in Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928, 1929, pp. 47–60.

[8] D. Vysochanskij and Y. I. Petunin, "Justification of the $3\sigma$ rule for unimodal distributions," Theory of Probability and Mathematical Statistics, vol. 21, no. 25-36, 1980.

[9] A. Charnes, W. W. Cooper, and G. H. Symonds, "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil," Management science, vol. 4, no. 3, pp. 235–263, 1958.

[10] C. M. Lagoa, X. Li, and M. Sznaier, "Probabilistically Constrained Linear Programs and Risk-Adjusted Controller Design," SIAM Journal on Optimization, vol. 15, no. 3, pp. 938–951, 2005.

[11] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, Robust Optimization. Princeton University Press, 2009, vol. 28.

[12] 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.

[13] W. Han, A. Jasour, and B. Williams, "Non-Gaussian Risk Bounded Trajectory Optimization for Stochastic Nonlinear Systems in Uncertain Environments," arXiv preprint arXiv:2203.03038, 2022.

[14] G. Calafiore and M. C. Campi, "Uncertain convex programs: randomized solutions and confidence levels," Mathematical Programming, vol. 102, no. 1, pp. 25–46, 2005.

[15] A. M. Jasour, N. S. Aybat, and C. M. Lagoa, "Semidefinite Programming For Chance Constrained Optimization Over Semialgebraic Sets," SIAM Journal on Optimization, vol. 25, no. 3, pp. 1411–1440, 2015.

[16] R. Lewis and R. Vinter, "Relaxation of Optimal Control Problems to Equivalent Convex Programs," Journal of Mathematical Analysis and Applications, vol. 74, no. 2, pp. 475–493, 1980.

[17] 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.

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

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

[20] G. Fantuzzi and D. Goluskin, "Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization," SIAM Journal on Applied Dynamical Systems, vol. 19, no. 3, pp. 1823–1864, 2020.

[21] J. Miller, D. Henrion, M. Sznaier, and M. Korda, "Peak Estimation for Uncertain and Switched Systems," in 2021 60th IEEE Conference on Decision and Control (CDC), 2021, pp. 3222–3228.

[22] K. Kashima and R. Kawai, "An optimization approach to weak approximation of Lévy-driven stochastic differential equations," in Perspectives in Mathematical System Theory, Control, and Signal Processing. Springer, 2010, pp. 263–272.

[23] J.-B. Lasserre, T. Prieto-Rumeau, and M. Zervos, "Pricing a class of exotic options via moments and SDP relaxations," Mathematical Finance, vol. 16, no. 3, pp. 469–494, 2006.

[24] 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.

[25] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko, "Bounds for Deterministic and Stochastic Dynamical Systems using Sum-of-Squares Optimization," SIAM Journal on Applied Dynamical Systems, vol. 15, no. 4, pp. 1962–1988, 2016.

[26] B. Xue, N. Zhan, and M. Fränzle, "Reach-Avoid Analysis for Stochastic Differential Equations," 2022.

[27] C. Gardiner, Stochastic methods. Springer Berlin, 2009, vol. 4.

[28] B. Øksendal, "Stochastic Differential Equations: An Introduction with Applications ," in Stochastic differential equations. Springer, 2003, pp. 65–84.

[29] R. Fürth, "Einige Untersuchungen über Brownsche Bewegung an einem Einzelteilchen," Annalen der Physik, vol. 358, no. 11, pp. 177–213, 1917.

[30] H. J. Hwang, J. Jang, and J. J. Velázquez, "The Fokker–Planck equation with absorbing boundary conditions," Archive for Rational Mechanics and Analysis, vol. 214, no. 1, pp. 183–233, 2014.

[31] E. B. Dynkin, "Markov Processes," in Markov processes. Springer, 1965, pp. 77–104.

[32] F. Alizadeh and D. Goldfarb, "Second-order cone programming," Mathematical programming, vol. 95, no. 1, pp. 3–51, 2003.

[33] J. Lofberg. (2009) Working with square roots. [Online]. Available: https://yalmip.github.io/squareroots

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

[35] M. ApS, The MOSEK optimization toolbox for MATLAB manual. Version 9.2., 2020. [Online]. Available: https://docs.mosek.com/9.2/toolbox/index.html

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

[37] 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.

[38] J. Miller and M. Sznaier, "Bounding the Distance of Closest Approach to Unsafe Sets with Occupation Measures," in 2022 61st IEEE Conference on Decision and Control (CDC), 2022.

[39] S. He, J. Zhang, and S. Zhang, "Bounding Probability of Small Deviation: A Fourth Moment Approach," Mathematics of Operations Research, vol. 35, no. 1, pp. 208–232, 2010.