Moment-SOS hierarchy and exit location of stochastic processes*

Didier Henrion[1,2], Mauricio Junca[3], Mauricio Velasco[4]

Draft of March 14, 2024

Abstract

The moment sum of squares (moment-SOS) hierarchy produces sequences of upper and lower bounds on functionals of the exit time location of a polynomial stochastic differential equation with polynomial constraints, at the price of solving semidefinite optimization problems of increasing size. In this note we use standard results from elliptic partial differential equation analysis to prove convergence of the bounds produced by the hierarchy. We also use elementary convex analysis to describe a super- and sub-solution interpretation dual to a linear formulation on occupation measures. Finally, we introduce a novel Christoffel-Darboux approach for the recovery of the exit location and occupation measures. The practical relevance of these results is illustrated with numerical examples.

1 Introduction

This paper deals with the numerical evaluation of functionals of solutions for nonlinear stochastic differential equations (SDE). Our approach consists of constructing a family of convex optimization problems (semidefinite programming problems, SDP) of increasing size whose solutions yield bounds on the given functional. This is an application of the so-called Lasserre or moment sum of squares (SOS) hierarchy [17, 9]. We are especially concerned about proving *convergence* of the bounds to the value of the functional.

The moment-SOS hierarchy was already developed and used in [19, 20] for obtaining bounds on SDEs coming from finance. However, only lower and upper bounds were obtained, and the question of convergence was left open.

[1]CNRS, LAAS, Université de Toulouse, France.

[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia.

[3]Department of Mathematics, Universidad de los Andes, Bogotá, Colombia.

[4]Departamento de Informática, Universidad Católica del Uruguay (UCU), Montevideo, Uruguay.

*M. Velasco was partially supported by Proyecto INV-2018-50-1392 (UAndes, Colombia) and Fondo Clemente Estable project FCE-1-2023-1-176172 (ANII, Uruguay). D. Henrion, M. Junca and M. Velasco were partially supported by ColCiencias Colombia-France cooperation Ecos Nord grant "Problemas de momentos en control y optimización" (C19M02).

A key step to construct the moment-SOS hierarchy is the reformulation of the original, typically nonlinear problem, as a linear problem on occupation measures. This linear reformulation is classical in Markov decision processes (MDP) [2, 16]. In order to prove convergence of the bounds obtained with the moment-SOS hierarchy, one has to prove that there is *no relaxation gap* between the original nonlinear problem and the linear problem on measures. This was already achieved in [2, 16] in the context of controlled MDP, but the proofs are lengthy and technical. Zero relaxation gap for optimal control of SDEs was proven in [3] with the help of viscosity solutions to Hamilton-Jacobi-Bellman partial differential equations (PDE).

In [13, 14], bounds on functionals of solutions of SDEs were obtained by a dual approach, seeking test functions satisfying inequalities. When the functions and the SDE coefficients are polynomial, the inequalities are replaced by SOS constraints and solved numerically with SDP. Our occupation measure formulation can be interpreted as a primal approach, from which the dual on test functions follows from elementary convex analysis arguments. More recently, a primal-dual moment-SOS hierarchy approach to optimal control of SDEs was followed in [12], as a stochastic counterpart of [11], and no relaxation gap was ensured by approximating the value function solving the dual HJB PDE.

In this paper, we focus on a specific class of SDE functional evaluation problems, namely the *exit location of an uncontrolled SDE*. The exit location is a random variable that can be characterized by its moments. As shown in [8], the exit time moments can be approximated numerically with occupation measures and linear programming (LP), with convergence guarantees based on the zero relaxation gap proof of [16].

Our contribution is as follows:

- we provide a new proof of the equivalence, or zero relaxation gap, between the infinite-dimensional linear formulation on occupation measures and the original nonlinear SDE; the proof, much shorter and simpler in our opinion than the MDP proofs of [2, 16] or the HJB proof of [3, 12], relies on standard results from elliptic PDE analysis;
- we describe a neat primal-dual linear formulation with no duality gap, allowing readily the application of the moment-SOS hierarchy.
- we propose a novel use of Christoffel-Darboux polynomial kernels [18] to approximate the densities of occupation measures and characterizing the exit location measure.

The paper is organized as follows. The exit location problem is defined in section 2. Its linear reformulation with occupation measures is described in section 3: we start with the primal formulation in subsection 3.1; our main result on zero relaxation gap is described and proved in subsection 3.2; the dual linear formulation is described in subsection 3.3. Application of the moment-SOS hierarchy and numerical examples are described in section 4. Section 5 describes how Christoffel-Darboux polynomial kernels can be used to approximate the densities of the occupation measures. Concluding remarks are gathered in section 6.

2 Exit location problem

Let $\mathscr{X}$ be a given open connected bounded set of $\mathbb{R}^n$ with smooth boundary $\partial\mathscr{X}$ and compact closure $\overline{\mathscr{X}} := \mathscr{X} \cup \partial\mathscr{X}$. Let $\mathbf{W}(t) = (W_k)_{k=1,\dots,m}$ denote the m-dimensional Brownian motion and let $\mathbf{X}(t)$ denote the solution of the stochastic differential equation (SDE)

$$d\mathbf{X} = \mathbf{b}(\mathbf{X})dt + \mathbf{B}(\mathbf{X})d\mathbf{W}, \quad \mathbf{X}(0) = x$$

starting at $x \in \mathscr{X}$ where drift functions $\mathbf{b} = (b_i)_{i=1,\dots,n} : \mathbb{R}^n \to \mathbb{R}^n$ and diffusion functions $\mathbf{B} = (b_{ij})_{i=1,\dots,n,\, j=1,\dots,m} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are given. We assume that $\mathbf{B}$ has full rank on $\mathscr{X}$, so that the matrix $\mathbf{A} = (a_{ij} := \frac{1}{2}\sum_{k=1}^m b_{ik}b_{jk})_{i,j=1,\dots,n} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is positive definite. Assume $\mathbf{b}$ and $\mathbf{B}$ are continuous on $\mathbb{R}^n$ and growing at most linearly outside of $\mathscr{X}$, so that by standard arguments [7, Chapter 5] there is a unique solution to the SDE, the stochastic process $\mathbf{X}(t)$.

Let $g : \partial\mathscr{X} \to \mathbb{R}$ be a given continuous function. We want to evaluate the function

$$E[g(\mathbf{X}(\tau_x))] \tag{1}$$

where τ_x is the first time $\mathbf{X}(.)$ hits $\partial\mathscr{X}$, see e.g. [7, Example 2, Section 6.2.1].

3 Linear reformulation

The generator of the stochastic process is the linear partial differential operator

$$-Lf := \sum_{i,j=1}^n a_{ij}\partial_i\partial_j f + \sum_{i=1}^n b_i\partial_i f$$

where ∂_i denotes the derivative with respect to the i-th variable. With this sign convention, and since the matrix $\mathbf{A}$ is positive definite, linear operator L is uniformly elliptic [6, Section 6.1.1].

Given a function $f \in C^2(\overline{\mathscr{X}})$, Itô's chain rule [7, Chapter 4] implies that

$$f(\mathbf{X}(\tau_x)) = f(\mathbf{X}(0)) - \int_0^{\tau_x} Lf ds + \int_0^{\tau_x} Df \cdot \mathbf{B}\, d\mathbf{W}$$

where

$$Df \cdot \mathbf{B}\, d\mathbf{W} = \sum_{k=1}^m \sum_{i=1}^n \partial_i f b_{ik}\, dW_k.$$

Taking the expected value yields Dynkin's formula [7, Section 6.1.3]:

$$E[f(\mathbf{X}(\tau_x))] = E[f(\mathbf{X}(0))] - E\left[\int_0^{\tau_x} Lf ds\right]$$

that we rearrange as follows

$$E[f(\mathbf{X}(\tau_x))] + E\left[\int_0^{\tau_x} Lf ds\right] = f(x). \tag{2}$$

Given $x \in \mathscr{X}$, define the expected occupation measure μ of the process $\mathbf{X}$ up to time τ_x, such that

$$\mu(\mathscr{A}) := E\left[\int_0^{\tau_x} I_{\mathscr{A}}(\mathbf{X}(s))ds\right]$$

for every set $\mathscr{A}$ in the Borel sigma algebra of $\mathscr{X}$, where $I_{\mathscr{A}}$ denotes the indicator function equal to one in $\mathscr{A}$ and zero outside. An equivalent analytic definition is

$$\langle f, \mu \rangle := E\left[\int_0^{\tau_x} f(\mathbf{X}(s))ds\right]$$

for any test function f, where

$$\langle f, \mu \rangle := \int f\mu$$

denotes the duality pairing of a continuous function f and a measure μ. Define the exit location measure ν as the law of $\mathbf{X}(\tau_x)$ i.e.

$$\nu(\partial\mathscr{B}) := E[I_{\partial\mathscr{B}}(\mathbf{X}(\tau_x))]$$

for every set $\mathscr{B}$ in the Borel sigma algebra of $\partial\mathscr{X}$. An equivalent analytic definition is

$$\langle f, \nu \rangle := E\left[f(\mathbf{X}(\tau_x))\right]$$

for every test function f. Then Dynkin's formula (2) becomes a linear partial differential equation on measures

$$\langle f, \nu \rangle + \langle Lf, \mu \rangle = f(x) \tag{3}$$

which can be written in the sense of distributions as

$$\nu + L'\mu = \delta_x \tag{4}$$

where L' is the linear operator adjoint to L and δ_x is the Dirac measure concentrated at x. This equation is called the Kolmogorov or Fokker-Planck equation.

So far we have assumed that the initial condition is deterministic. All the above developments generalize readily to the case when the initial condition $\mathbf{X}(0)$ in the SDE is a random variable whose law is a given probability measure ξ on $\mathscr{X}$.

The quantity (1) to be evaluated is now averaged with respect to the distribution of initial conditions:

$$v^*(\xi) := \int_{\mathscr{X}} E[g(\mathbf{X}(\tau_x))]d\xi(x) \tag{5}$$

and the Kolmogorov equation (4) becomes

$$\nu + L'\mu = \xi. \tag{6}$$

This general formulation allows us to recover the deterministic initial condition of starting at a given point $x \in \mathscr{X}$ by making the particular choice $\xi := \delta_x$.

3.1 Primal formulation

Following [19, 20], now define

$$v_{\min}(\xi) := \min_{\mu,\nu} \langle g, \nu \rangle \ \text{ s.t. } \ \nu + L'\mu = \xi \tag{7}$$

and

$$v_{\max}(\xi) := \max_{\mu,\nu} \langle g, \nu \rangle \ \text{ s.t. } \ \nu + L'\mu = \xi \tag{8}$$

which satisfy by construction

$$v_{\min}(\xi) \leq v^*(\xi) \leq v_{\max}(\xi).$$

Note that (7) and (8) are linear optimization problems over measures μ and ν supported on $\mathscr{X}$ and $\partial\mathscr{X}$ respectively.

3.2 No relaxation gap

Let us make the following regularity assumptions on the problem data, see [23, Sections 3 and 4] for definitions.

Assumption 1 *The boundary $\partial\mathscr{X}$ has class $C^{2,\alpha}$ and the coefficients a_{ij}, b_i have class $C^{0,\alpha}(\overline{\mathscr{X}})$ for some $\alpha \in (0,1)$.*

Theorem 1 *Under Assumption 1, there is no relaxation gap between the nonlinear function evaluation problem (5) and the linear optimization problems (7) and (8), i.e.*

$$v_{\min}(\xi) = v^*(\xi) = v_{\max}(\xi)$$

for each probability measure ξ on $\mathscr{X}$.

The proof of Theorem 1 is based on the following result.

Lemma 1 *Under Assumption 1, for each probability measure ξ on $\mathscr{X}$ there exists a unique exit location measure $\nu = \nu_\xi$ on $\partial\mathscr{X}$ and a unique expected occupation measure $\mu = \mu_\xi$ on $\mathscr{X}$ solving the Kolmogorov equation (6).*

Proof of Lemma 1: Let $\alpha \in (0,1)$ be the regularity exponent of Assumption 1. Given any function $p \in C^{2,\alpha}(\overline{\mathscr{X}})$, let $f_p \in C^{2,\alpha}(\overline{\mathscr{X}})$ be a solution to the Dirichlet boundary value problem

$$\begin{array}{rl} Lf = 0 & \text{in } \mathscr{X} \\ f = p & \text{on } \partial\mathscr{X}. \end{array} \tag{9}$$

Existence, uniqueness and regularity of the solution follows from [23, Theorem 4.5]. Integating f_p with respect to (6) yields

$$\langle p, \nu \rangle = \langle f_p, \xi \rangle. \tag{10}$$

Since $\overline{\mathscr{X}}$ is compact, the space $C(\overline{\mathscr{X}})$ is separable and by choosing countably many functions $p \in C^{2,\alpha}(\overline{\mathscr{X}})$ – e.g. polynomials – we can generate countably many linear relations (10) that uniquely specify the measure ν that we denote ν_ξ.

Uniqueness of the measure μ also holds. Indeed, if h ranges over a countable dense set of functions in $C^{2,\alpha}(\overline{\mathscr{X}})$ and the f_h are the unique solutions in $C^{2,\alpha}(\overline{\mathscr{X}})$ of the Poisson problem

$$\begin{aligned} Lf_h &= h \quad \text{in } \mathscr{X} \\ f_h &= 0 \quad \text{on } \partial\mathscr{X} \end{aligned}$$

then the Kolmogorov equation (6) yields the equality

$$\langle h, \mu \rangle = \langle f_h, \xi \rangle - \langle f_h, \nu \rangle = \langle f_h, \xi \rangle$$

and the moments of μ with respect to the functions in h are uniquely specified. By density, this implies the uniqueness of μ that we denote μ_ξ. □

Proof of Theorem 1: Let ν_ξ denote the exit location measure on $\partial\mathscr{X}$ solving the Kolmogorov equation (6). Notice that the objective function in problems (7) and (8) depends only on ν_ξ. It follows that $\langle g, \nu_x \rangle = v_{\min}(\xi) = v_{\max}(\xi)$.

Remark 1 *Existence and uniqueness of a regular solution to the Dirichlet problem* (9) *also follows from the uniformly ellipticity of* L *and weaker conditions on the data, as stated in [21, Theorem 9.2.14]. Uniqueness of the solution follows from [21, Theorem 9.2.13] under Hunt's condition and also under the assumption that all points of* $\partial\mathscr{X}$ *are regular, see [21, Definition 9.2.8]. Hunt's condition holds for all Itô diffusions whose diffusion coefficient matrix has a bounded inverse and whose drift coefficient satisfies the so-called Novikov condition. For our purpose however, we prefer to stick with stronger assumptions and the self-contained statement and proof of [23, Theorem 4.5].*

3.3 Duality

Now let us use elementary notions from convex duality to derive the dual problem to the minimization resp. maximization problem on measures. We show that admissible solutions to the dual problem are subsolutions resp. supersolutions to the boundary value PDE solved by the value function. In particular we show that the concept of supersolution (resp. subsolution) arises naturally from elementary duality theory.

Lemma 2 *The linear problem dual to* (7) *reads as follows*

$$\max_v \langle v, \xi \rangle \text{ s.t. } Lv \leq 0 \text{ in } \mathscr{X}, \ \ v \leq g \text{ on } \partial\mathscr{X} \tag{11}$$

where the maximization is with respect to functions $v \in C^2(\mathscr{X})$*. There is no duality gap, i.e. the value of* (11) *is equal to the value of* (7)*.*

Proof: Let us denote by $v \in C(\mathscr{X})$ the Lagrange multiplier corresponding to the equality constraint in primal problem (7), and build the Lagrangian $\ell(\mu, \nu, v) := \langle g, \nu \rangle + \langle v, \xi -$

$\nu - L'\mu\rangle = \langle g - v, \nu\rangle + \langle v, -L'\mu\rangle + \langle v, \xi\rangle = \langle g - v, \nu\rangle + \langle -Lv, \mu\rangle + \langle v, \xi\rangle$. The Lagrange dual function is then $\min_{\mu,\nu} \ell(\mu, \nu, v) = \langle v, \xi\rangle$ provided $v \leq g$ on $\partial\mathscr{X}$, the support of ν, and $Lv \leq 0$ on $\mathscr{X}$, the support of μ. The dual problem (11) then consists of maximizing the dual function subject to these inequality constraints. To prove that there is no duality gap, we use [1, Theorem IV.7.2] and the fact that the image through the linear map $(\langle g, \nu\rangle, \nu + L'\mu)$ of the cone of measures μ resp. ν supported on $\mathscr{X}$ resp. $\partial\mathscr{X}$ is nonempty and bounded in the metric inducing the weak-star topology on measures. □

As recalled in [7, Example 2, Section 6.B], the value function w is the solution of the boundary value problem

$$\begin{array}{ll} Lv = 0 & \text{in } \mathscr{X} \\ v = g & \text{on } \partial\mathscr{X}. \end{array} \tag{12}$$

Lemma 3 *Any admissible function v for linear problem (11) is a subsolution of boundary value problem (12), in the sense that $w \geq v$ on $\overline{\mathscr{X}}$.*

Proof: Let v be admissible for (11). Function $u := v - w$ is such that $Lu \leq 0$ in $\mathscr{X}$ and $u \leq 0$ on $\partial\mathscr{X}$. By the weak maximum principle [6, Theorem 1 page 327], if $Lu \leq 0$ in $\mathscr{X}$ then $\max_{\overline{\mathscr{X}}} u = \max_{\partial\mathscr{X}} u$. Since $u \leq 0$ on $\partial\mathscr{X}$, this implies that $u \leq 0$ and hence $w \geq v$ on $\overline{\mathscr{X}}$. □

Linear problem (11) selects a subsolution of maximal averaged value with respect to ξ. In particular, if $\xi = \delta_x$ for some $x \in \mathscr{X}$, an optimal subsolution touches the value function from below at x.

Similarly, the linear problem dual to (8) reads as follows

$$\min_v \langle v, \xi\rangle \text{ s.t. } Lv \geq 0 \text{ in } \mathscr{X}, \quad v \geq g \text{ on } \partial\mathscr{X} \tag{13}$$

where the maximization is with respect to functions $v \in C^2(\mathscr{X})$. There is no duality gap, i.e. the value of (13) is equal to the value of (8). Any admissible function v for linear problem (13) is a super-solution of boundary value problem (12), in the sense that $w \leq v$ on X and $\partial\mathscr{X}$. Linear problem (13) selects the supersolution of minimal averaged value with respect to ξ. In particular, if $\xi = \delta_x$ for some $x \in \mathscr{X}$, an optimal subsolution touches the value function from above at x.

4 Moment-SOS hierarchy and examples

If the SDE coefficients $\mathbf{b}$ and $\mathbf{B}$ and the functional g are semialgebraic[1] in a semialgebraic set $\mathscr{X}$, we can apply the moment-SOS hierarchy on (7) and (8) with convergence guarantees. In the primal, we obtain approximate moments (also called pseudo-moments or pseudo-expectations) of the occupation measures. They are not necessarily moments of the occupation measures as we are solving relaxations. In the SOS dual, we obtain polynomial sub- resp. super-solutions of increasing degrees of the boundary value problem (12). Each primal-dual problem is a semidefinite optimization problem.

[1]A function is semialgebraic if its graph is a semialgebraic set. A semialgebraic set is defined by a finite sequence of polynomial equations and inequalities.

4.1 Semidefinite relaxations

Let us briefly describe the construction of the moment-SOS hierarchy.

A bounded closed semialgebraic set $\mathscr{Z}$ of $\mathbb{R}^n$ can be written as the union of finitely many basic semialgebraic sets $\mathscr{Z}_i := \{z \in \mathbb{R}^n : p_{i,j}(z) \geq 0,\ j = 1, \ldots, m_i\}$, $i = 1, \ldots, m$, described by finitely many polynomials in a vector $p_i = (p_{i,j})_{j=1,\ldots,m_i}$. Note that polynomial equations can be modeled by two inequalities of reverse signs. Since $\mathscr{Z}$ is bounded, without loss of generality, for each $i = 1, \ldots, m$, one of the polynomials defining each $\mathscr{Z}_i$ can be chosen equal to $R^2 - \sum_{k=1}^{n} z_k^2$ for R sufficiently large, and for notational convenience we let $p_{i,0}(z) := 1$

There are several algebraic characterizations of the set of positive polynomials on $\mathscr{Z}_i$. To describe one such characterization, let r be a positive integer and for each $i = 1, \ldots, m$ define the truncated quadratic module of degree r generated by the polynomials in p_i defining $\mathscr{Z}_i$, denoted $Q(p_i)_r$, to be the set of polynomials which can be written as $\sum_{j=0}^{m_i} p_{i,j} s_{i,j}$ where the $s_{i,j}$ are sums of squares (SOS) of polynomials such that $2\mathrm{deg} s_{i,j} + \mathrm{deg} p_{i,j} \leq r$. Every polynomial in the Minkowski sum $\sum_{i=1}^{m} Q(p_i)_r$ is obviously nonnegative on $\mathscr{Z}$. Putinar's Positivstellensatz [22] is the much deeper statement that every polynomial strictly positive on $\mathscr{Z}$ lies in $\sum_{i=1}^{m} Q(p_i)_r$ for some r.

We will now describe a hierarchy of semidefinite optimization problems which depend on the degree r and provide us with upper and lower bounds on the value $v^*(\xi)$. We will show that as $r \to \infty$ these bounds converge to the value $v^*(\xi)$.

Let us assume that $\mathscr{X}$ is a bounded basic semi-algebraic set with definining polynomials p. Its boundary $\partial\mathscr{X}$ is then a union of finitely many bounded basic semi-algebraic sets $\mathscr{X}_i^\partial$ with defining polynomials p_i^∂, $i = 1, \ldots, m$. The primal (moment) problems are given by

$$
\begin{array}{rll}
p_r^{\min} := & \min & \sum_{i=1}^{m} \ell_{\nu_i}(g) \\
& \text{s.t.} & \ell_\mu(q) \geq 0, \quad \forall q \in Q(p)_r \\
& & \ell_{\nu_i}(q) \geq 0, \quad \forall q \in Q(p_i^\partial)_r \\
& & \ell_\mu(Lq) + \sum_{i=1}^{m} \ell_{\nu_i}(q) = \langle q, \zeta \rangle, \quad \forall q \in P_r
\end{array}
\tag{14}
$$

and

$$
\begin{array}{rll}
p_r^{\max} := & \max & \sum_{i=1}^{m} \ell_{\nu_i}(g) \\
& \text{s.t.} & \ell_\mu(q) \geq 0, \quad \forall q \in Q(p)_r \\
& & \ell_{\nu_i}(q) \geq 0, \quad \forall q \in Q(p_i^\partial)_r \\
& & \ell_\mu(Lq) + \sum_{i=1}^{m} \ell_{\nu_i}(q) = \langle q, \xi \rangle, \quad \forall q \in P_r
\end{array}
\tag{15}
$$

where the unknowns are linear operators $\ell_\mu, \ell_{\nu_1}, \ldots, \ell_{\nu_m}$ from P_r to $\mathbb{R}$, for P_r denoting the vector space of n-variate real polynomials of degree up to r. The dual (SOS) problems are given by

$$
\begin{array}{rll}
d_r^{\min} := & \max & \langle q, \xi \rangle \\
& \text{s.t.} & -Lq \in Q(p)_r \\
& & g - q \in \sum_{i=1}^{m} Q(p_i^\partial)_r
\end{array}
\tag{16}
$$

and

$$
\begin{array}{rll}
d_r^{\max} := & \min & \langle q, \xi \rangle \\
& \text{s.t.} & Lq \in Q(\mathscr{X})_r \\
& & q - g \in \sum_{i=1}^{m} Q(\mathscr{X}_i^\partial)_r
\end{array}
\tag{17}
$$

where the unknowns are polynomials $q \in P_r$.

Theorem 2 *Problems (14), (15), (16) and (17) are semidefinite programming problems. For each $r > 0$, it holds $d_r^{\min} \le p_r^{\min} \le v^*(\xi) \le p_r^{\max} \le d_r^{\max}$. Moreover $\lim_{r\to\infty}(d_r^{\max} - d_r^{\min}) = 0$.*

Proof: To show that problems (16) and (17) are semidefinite programming problems, just observe that a polynomial $p(z)$ of degree at most $2r$ is a sum of squares of polynomials if and only if there is a positive semidefinite symmetric matrix S such that $p(z) = b(z)'Sb(z)$ where $b(z)$ is a vector of polynomials spanning P_r. It follows that the truncated quadratic modules in (16) and (17) are projections of the semidefinite cone, and optimizing linear functions over them is an instance of semidefinite programming.

Problems (14) and (16) are in duality. This follows easily from computing the Lagrangian as in Lemma 2. The constraint in problem (14) that a linear operator is non-negative for all polynomials in a truncated module can be expressed as a linear matrix inequality, i.e. it forms a spectrahedron, a linear slice of the semidefinite cone. From weak duality we conclude that $p_r^{\min} \ge d_r^{\min}$. Similarly, problems (15) and (17) are in duality and $d_r^{\max} \ge p_r^{\max}$.

To prove that $p_r^{\min} \le v^*(\xi)$ observe that if (ν, μ) are measures satisfying $\nu + \mathcal{L}'\mu = \xi$ then defining $\ell_\mu(q) := \langle q, \mu\rangle$ and $\ell_{\nu_i}(q) := \langle q, \nu_i\rangle$ where $\nu = \sum_{i=1}^m \nu_i$ for each ν_i supported on $\mathscr{X}_i^\partial$, we obtain admissible linear operators for the primal problem (14). Coefficients of the linear operators are moments of the respective measures. It may however happen that linear operators admissible for problem (14) do not correspond to measures. Since we minimize a linear function on a possibly larger set, we obtain a lower bound. Similarly, we can prove that $v^*(\xi) \le p_r^{\max}$.

The most substantial claim is the convergence result. Given $\epsilon > 0$ we will show that there exists an integer r_ϵ such that $0 \le v^*(\xi) - d_r^{\min} \le \epsilon$ for $r \ge r_\epsilon$. Similar arguments imply an analogous convergence result for $d_r^{\max}$.

Recall from Section 3.3 that the unique function $f \in C^2(\mathscr{X})$ which satisfies $Lf = 0$ in $\mathscr{X}$ and $f = g$ on $\partial\mathscr{X}$ is the value function $f = w$ of the problem and it satisfies $v^*(\xi) = \langle w, \xi\rangle$. The proof proceeds by showing that this value function can be approximated by a sequence of elements v_n which are feasible for problem (16). Since operator L is uniformly elliptic, there is a polynomial u such that $-Lu > 0$ in $\mathscr{X}$, take e.g. a quadratic polynomial with sufficient large leading coefficients. By substracting a sufficiently large constant from u, we can also ensure that $u < g$ on $\partial\mathscr{X}$. Now let $\epsilon > 0$ be given and let $(w_n)_{n\in\mathbb{N}}$ be a sequence of polynomials which approximate w and its derivatives of order up to 2 uniformly on $\overline{\mathscr{X}}$. Let $\eta > 0$ be a real number with $\eta < \frac{\epsilon}{2\|u-w\|}$ and let $v_n := (1-\eta)w_n + \eta u$. Due to the uniformity of the convergence of w_n and the definition of v_n the following statements hold for all n sufficiently large: $\|w_n - w\| < \eta\|u - w\|$, $-Lv_n > 0$ in $\overline{\mathscr{X}}$ and $v_n < g$ on $\partial\mathscr{X}$. By Putinar's Positivstellensatz [22], for any such n there exists a degree r_n such that $-Lv_n \in Q(\mathscr{X})_{r_n}$ and $g - v_n \in Q(\partial\mathscr{X})_{r_n}$ and therefore v_n is feasible for problem (16) for $r = r_n$. Furthermore, for any such n we have $|\langle v_n - w, \xi\rangle| \le (1-\eta)\|w_n - w\| + \eta\|u - w\| \le 2\eta\|u - w\| < \epsilon$ so we conclude that $v^*(\xi) - d_{r_n}^{\min} \le \epsilon$ proving the claim since $\epsilon > 0$ was arbitrary.

In the above formulation of semidefinite programming problems (14) and (15), the moment constraints are represented by positivity conditions on the corresponding linear functionals. This is a basis independent formulation. Now if polynomials q are expressed in a given basis, the positivity conditions become explicit positive semidefinite conditions on symmetric matrices depending linearly on the moments, i.e. linear matrix inequalities. For example,

relaxation degree	2	4	6	8	10
lower bound	0.65000	0.92157	0.98118	0.99503	0.99827
upper bound	1.00000	1.00000	1.00000	1.00000	1.00000

Table 1: Scalar example - bounds for increasing relaxation degrees.

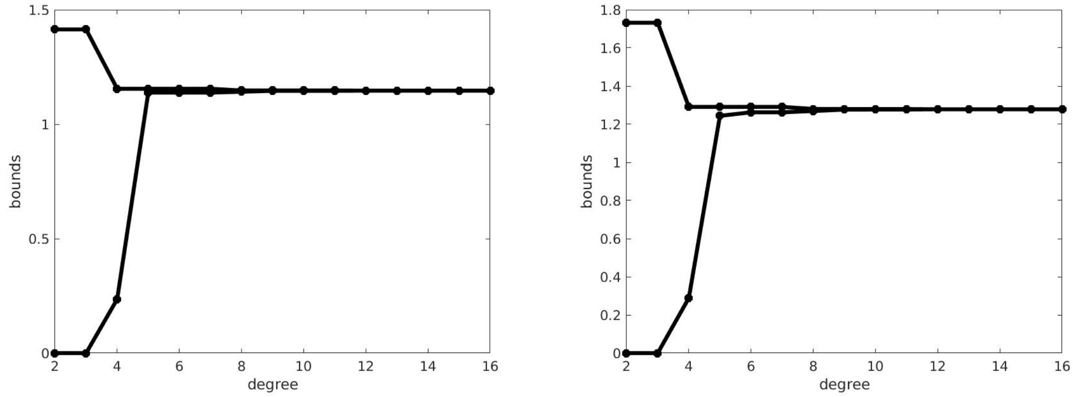

Figure 1: Lower and upper bounds on the functional for increasing relaxation degrees and dimension $n = 2$ (left) and $n = 3$ (right).

if monomials are used represent polynomials, the matrices have Hankel structure, see e.g. [17, Chapter 3] for explicit expressions. Software interfaces such as GloptiPoly [10] construct these matrices automatically. The resulting semidefinite relaxations are solved with the SDP solver in MOSEK [5].

4.2 Scalar example

Let us illustrate the application of the moment-SOS hierarchy with an elementary exit time problem (1) considered in [8, Example 5.1]. The SDE is $d\mathbf{X}(t) = (1 + 2\mathbf{X}(t))dt + \sqrt{2}\mathbf{X}(t)d\mathbf{W}(t)$ on the domain $\mathscr{X} := (0, 1)$ with initial condition $x = 1/2$. This process always exits at the point $\{1\}$, so $\nu = \delta_1$ solves (3) and the value of linear problems (7) and (8) is equal to $g(1)$. For the choice $g(z) = z^2$ we report in Table 1 the values of the lower and upper bounds obtained by solving the moment relaxations to problems (7) and (8), for increasing relaxation degrees. The degree 10 relaxation was solved in 0.15 seconds on our laptop. The corresponding GloptiPoly script in Matlab can be downloaded at homepages.laas.fr/henrion/software/exittime.m

4.3 Multivariate example

Consider problem (1) with $g(z) := \sum_{k=1}^n z_k^2$ for the n-dimensional Brownian motion $\mathbf{X}(t) = \mathbf{W}(t)$, i.e. $\mathbf{b} = 0$ and $\mathbf{B} = I_n$, in the convex semi-algebraic domain $\mathscr{X} := \{z \in \mathbb{R}^n : \sum_{k=1}^n z_k^4 \leq 1\}$ with initial condition $x = 0$.

On Figure 1 we plot for $n = 2$ and $n = 3$ the lower resp. upper bounds obtained by minimizing resp. maximizing the functional, for increasing relaxation degrees. We observe

dimension n	2	3	4	5	6	7	8
number of moments	73	249	705	1749	3927	8151	15873
CPU time (seconds)	0.15	0.59	1.9	5.6	19	51	195

Table 2: Computational burden for relaxation degree 8 and increasing dimensions.

a fast convergence of the bounds.

In Table 2 we report the number of moments as well as the computational time required to solve the moment relaxation of degree 8, for increasing values of the dimension n. For this relaxation degree the gap between the lower and upper bounds on the functional is less than 2%.

5 Approximating occupation and exit location densities

As described in the previous section, the moment-SOS hierarchy allows us to obtain estimates of the moments of our occupation and exit location measures, together with error bars on those estimates. For many applications however, it could be more useful to know the *densities* of the occupation and exit location measures with respect to some natural measures on $\mathscr{X}$ and $\partial\mathscr{X}$ respectively. In this section we outline a method, based on Christoffel-Darboux kernels [18], which allows us to construct sequences of functions converging to such densities. For simplicity we will state our results when $\mathscr{X}$ is the unit ball, but the same method can be applied more generally as discussed below.

5.1 Background on Christoffel-Darboux kernels

Let U be a compact topological space and let μ be a probability measure supported on U. The set $C(U)$ of continuous, real-vaued functions on U becomes a Hilbert space via the inner product $\langle f, g\rangle_U := E_\mu[fg] = \int_U f(u)g(u)d\mu(u)$. If $V \subseteq C(U)$ is a finite-dimensional vector space of functions then this space inherits a Hilbert space structure and in particular, for each $u \in U$ there exists a unique element $\phi_u \in V$ which reproduces the evaluation at u in the sense that the equality $\langle f, \phi_u\rangle_U = f(u)$ holds for every $f \in V$. Such operators determine the reproducing kernel function $K : V \times V \to \mathbb{R}$, defined via the formula $K(u, v) := \langle \phi_u, \phi_v\rangle_U$. This function satisfies, and is determined by, the following key reproducing property: $f(u) = \int_V K(u, v)f(v)d\mu(v) = \langle K(u, .), f\rangle_U$ valid for all $f \in V$ and all $u \in U$. The kernel is completely determined by the measure μ and the vector space V and we write $K_{\mu,V}$ instead of K when trying to make this dependence explicit. The following two formulas are very useful for computing and understanding the basic properties of reproducing kernels: 1) If $g_1, \ldots, g_m$ is any μ-orthonormal basis for the subspace V then $K_{\mu,V}(u, v) = \sum_{j=1}^{m} g_j(u)g_j(v)$ and furthermore 2) If $h_1, \ldots, h_m$ is any basis for the subspace V and M is the corresponding moment matrix given by $M_{ij} := \langle h_i, h_j\rangle_U$ then $K_{\mu,V}(u, v) = (h_1(u), \ldots, h_m(u))M^{-1}(h_1(v), \ldots, h_m(v))^T$. Finally a key quantity associated to a kernel is the Christoffel function $\Lambda_{\mu,V}(u) = K^{-1}_{\mu,V}(u, u)$, whose importance stems from the following variational characterization $\Lambda_{\mu,V}(u) := \inf_{f\in V}\langle f,f\rangle_U f^{-2}(u) = \inf_{f\in V}\{\langle f,f\rangle_U : f(u)=1\}$ which follows readily from the Cauchy-Schwartz inequality and the reproducing property.

In the special case when $U \subseteq \mathbb{R}^n$ and V_d is the subspace of polynomials of degree up to d in $\mathbb{R}^n$ restricted to U, the asymptotic behavior of the Christoffel function can often be used for density estimation [18]. A concrete set-up where these estimations can be carried out is described in [18, Theorem 5.3.4] as an extension of [15].

Theorem 3 *Suppose $U \subseteq \mathbb{R}^n$ is compact, $\omega \in C(U)$ is positive, λ is a probability measure supported on U and $d \mapsto N(d)$ is a polynomial such that $\lim_{d\to\infty}\sup_{u\in U}|N(d)\Lambda_{\lambda,V_d}(u) - \omega(u)| = 0$. If μ is a probability measure on U which is absolutely continuous with respect to λ, having a positive density $\rho \in C(U)$, then $\lim_{d\to\infty}\sup_{u\in U}\left|N(d)\Lambda_{\mu,V_d}(u) - \frac{\rho(u)}{\omega(u)}\right| = 0$.*

Example 1 *A basic example for which the hypotheses of Theorem 3 hold is the sphere $S^{n-1} \subseteq \mathbb{R}^n$ with its normalized area measure λ. In this case the invariance of the measure with respect to rotations implies that the function $K(u,u)$ is constant. The expression for $K(u,u)$ in terms of λ-orthonormal polynomials shows that the constant is the dimension of the space of polynomials restricted to the unit sphere. Since the sphere is defined by a single quadric, this dimension is given by the formula $N(d) := \binom{n+d}{d} - \binom{n+d-2}{d-2}$ so $N(d)\Lambda_{\lambda,V_d}(u) = 1$ for all $u \in U$ and we can use $w(u) = 1$ in the formula of Theorem 3 to conclude that for any measure μ which is absolutely continuous with respect to λ with a positive and continuous density ρ, this density can be estimated uniformly via the limit $\lim_{d\to\infty} N(d)\Lambda_{\mu,V_d}(u) = \rho(u)$.*

Example 2 *A similar result to Theorem 3 is available for the unit Euclidean ball $B^n \subseteq \mathbb{R}^n$. It is known [15, Theorem 1.3] that if $N(d) := \binom{n+d}{n}$ and λ is the measure on B^n with Lebesgue density $\omega(u) = 2\omega_n^{-1}(1-\|u\|_2^2)^{-1/2}$ where $\omega_n := 2\pi^{\frac{n+1}{2}}\Gamma^{-1}(\frac{n+1}{2})$ is the surface area of $S^{n-1} \subseteq \mathbb{R}^n$ then for any regular measure μ on B^n having a positive and continuous density ρ we have $\lim_{d\to\infty} N(d)\Lambda_{\mu,V_d}(u) = \frac{\rho(u)}{\omega(u)}$ where the equality holds uniformly on compact subsets of the interior of B^n.*

5.2 Computing occupation and exit location densities

To estimate the occupation density on the unit ball and the exit location density on its boundary we propose the following procedure. As before, let V_d denote the subspace of polynomials of degree up to d in $\mathbb{R}^n$. Let W_d be the vector space of functions obtained by restricting the elements of V_d to the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$.

Procedure 1 *Given a diffusion with generator $\mathcal{L}$ and initial location $\mathbf{X}(0) = x \in B^n \subseteq \mathbb{R}^n$, a total degree $d > 0$, an approximation degree $r > 0$ and an integer offset $s > 0$, carry out the following steps:*

1. *Fix a basis $g_1,\ldots,g_{n_g}$ for W_d. For each pair of elements g_i, g_j, let $g := g_i g_j$ and compute our estimate for $\langle g,\nu\rangle = \int g\nu$ as the average $\hat{A}_{ij} := (p_r^{\min} + p_r^{\max})/2$ of the values of the optimization problems (14) and (15).*

2. *Fix a basis $h_1, \dots, h_{n_h}$ for V_{d-s}. For each pair of elements h_i, h_j, find a polynomial g which solves the equation $Lg = h_i h_j$ and compute our estimate for $\langle g, \mu \rangle = \int g\mu$ as the shifted average $\hat{B}_{ij} = g(x) + (p_r^{\min} + p_r^{\max})/2$ of the values of the optimization problems (14) and (15).*

3. *Define our estimate for the exit location density on the unit sphere by*

$$\widehat{\sigma}(z) := \frac{\dim W_d}{(g_1(z), \dots, g_{n_g}(z))^T \hat{A}^{-1} (g_1(z), \dots, g_N(z))}$$

and our estimate for the occupation density on the unit ball by

$$\widehat{\beta}(z) = \frac{\dim V_{d-s}\, \omega(z)}{(h_1(z), \dots, h_{n_h}(z))^T \hat{B}^{-1} (h_1(z), \dots, h_{n_h}(z))}$$

with ω defined in Example 2.

Theorem 4 *If the following conditions hold*

1. *The exit location measure has a positive and continuous density σ with respect to the Lebesgue measure on the unit sphere S^{n-1}.*

2. *The occupation measure has a positive and continuous density β with respect to the Lebesgue measure on the unit ball B^n.*

3. *There exists a sequence $s(d)$ of positive integers with $d - s(d) \to \infty$ such that, for all sufficiently large d, the inclusion $LV_d \supseteq V_{d-s(d)}$ holds.*

4. *The chosen approximation degrees $r(d)$ grow sufficiently quickly so as to guarantee*

$$\lim_{d\to\infty} \sup_{g \in V_d} |p_{r(d)}^{\max}(g) - p_{r(d)}^{\min}(g)| \to 0$$

then the estimates $\widehat{\sigma}$ resp. $\widehat{\beta}$ defined in Procedure 1 converge uniformly to the true exit location density σ resp. occupation density β as $d \to \infty$.

Proof: If μ, ν denote the occupation measure and the exit location measure respectively and $\mathbf{X}(0) = x$ then the Kolmogorov equation implies that for any twice-differentiable function f we have

$$\langle f, \nu \rangle + \langle Lf, \mu \rangle = f(x)$$

If $LV_d \supseteq V_{d-s(d)}$ then this equation allows us to compute the polynomial moments of the occupation measure with respect to the functions in $V_{d-s(d)}$. More concretely, if $h \in V_{d-s(d)}$ then there exists $p \in V_d$, which can be easily computed with linear algebra, with $Lp = h$. As a result

$$\langle h, \mu \rangle = p(x) - \langle p, \nu \rangle$$

so estimates of the exit location moments $\langle p, \nu \rangle$ can be used to obtain certified estimates for the occupation moments as in Procedure 1. Such computations, which can become as good as we want by increasing the approximation degree $r(d)$, allow us to estimate the corresponding Christoffel functions as accurately as we wish. The result now follows from Theorem 3 and [15, Theorem 1.3]. □

Remark 2 *Procedure 1 can be carried out to estimate occupation densities for sets whose Equilibrium measure is well understood, playing the role of the density ω from Example 2. See [18, Section 4.2] for several such instances. Similarly, the assumption of Theorem 3 holds for several geometries beside spheres, see [18, p. 71] for some interesting instances.*

5.3 Some computational examples in the unit disc

In this section we illustrate the usefulness of Procedure 1 for estimating occupation densities in the unit ball and exit location densities on the unit sphere for some diffusions. The Julia code used for computing the examples in this Section is available via github at github.com/mauricio-velasco/Exit_location_density_estimation.git.

Example 3 *(Brownian motion in the unit disc). Let $W(t)$ be the two-dimensional brownian motion, let $X(t)$ be a solution of the SDE*

$$\begin{cases} X(0) = (0, 1/2) \\ dX(t) = dW(t) \text{ for } t \geq 0. \end{cases}$$

The generator of the process is $Lf = -\frac{1}{2}\Delta f$. We study the behavior of $X(t)$ on or before its exit time from the unit disc in Figure 2. Figure 2, parts (a) and (b) show five and one hundred approximate sample paths of the process. Figure 2, part (c) shows the unnormalized Christoffel-Darboux kernel estimates in degrees $r = 6, 8$ for the exit location density in the unit circle represented as the axis $-\pi/2 \leq \theta \leq 3\pi/2$. Figure 2, part (d) shows the Christoffel-Darboux kernel estimates in degree $r = 2$ for the occupation density of $X(t)$ before reaching the boundary in the unit disc.

Example 4 *(Brownian motion with drift in the unit disc). Let $W(t)$ be the two-dimensional brownian motion, let $X(t)$ be a solution of the SDE*

$$\begin{cases} (X_1(0), X_2(0)) = (0, 1/2) \\ \begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} dW(t). \end{cases}$$

The generator of the process is $Lf = -2\frac{\partial f}{\partial x_2} - \frac{1}{2}\Delta f$. We study the behavior of $X(t)$ on or before its exit time from the unit disc in Figure 3. Figure 3, parts (a) and (b) show five and one hundred approximate sample paths of the process. Figure 3, part (c) shows the unnormalized Christoffel-Darboux kernel estimates in degrees $r = 4, 6$ for the exit location density in the unit circle represented as the axis $-\pi/2 \leq \theta \leq 3\pi/2$. Figure 3, part (d) shows the Christoffel-Darboux kernel estimates in degree $r = 2$ for the occupation density of $X(t)$ before reaching the boundary of the unit disc.

Example 5 *(Square root Bessel process in the unit disc). Let $W(t)$ be the two-dimensional brownian motion, let $X(t)$ be a solution of the SDE*

$$\begin{cases} (X_1(0), X_2(0)) = (0, 1/2) \\ \begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & 2\sqrt{X_2(t)} \end{pmatrix} dW(t). \end{cases}$$

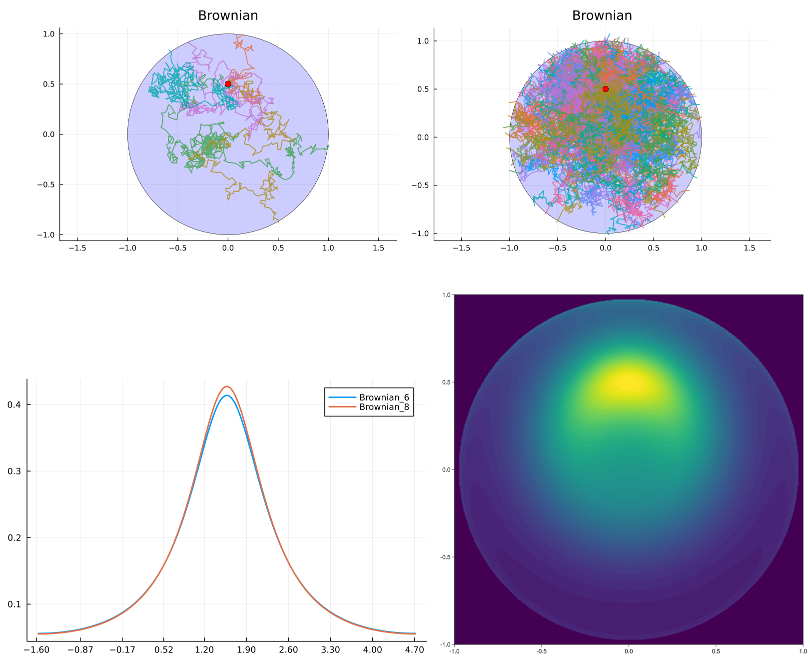

Figure 2: Example 3: brownian motion

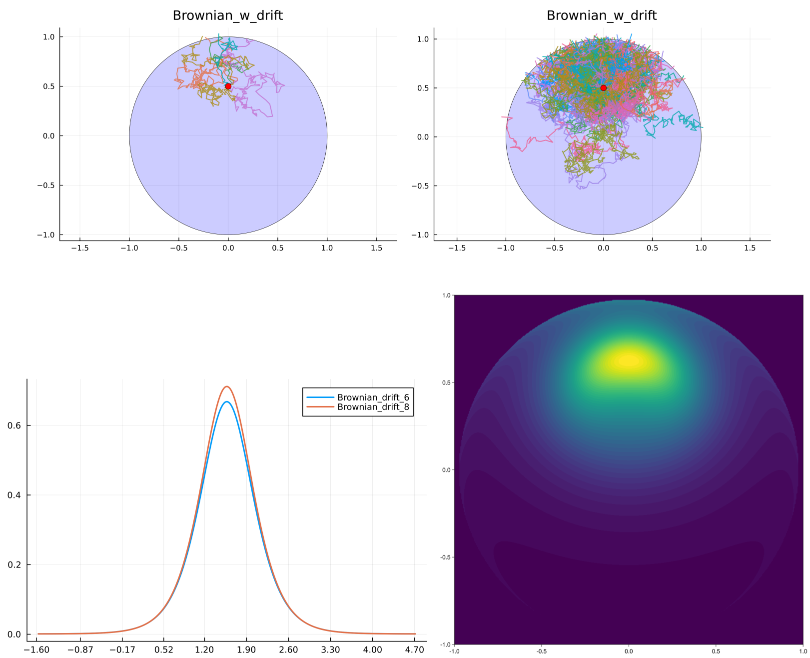

Figure 3: Example 4: brownian motion with upward drift

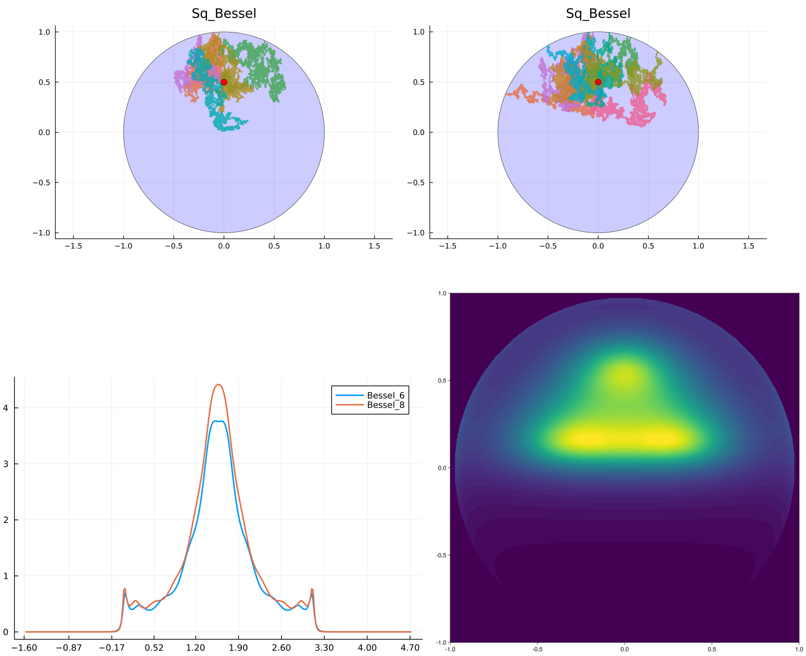

Figure 4: Example 5: square root Bessel process

The generator of the process is $Lf = -\frac{\partial f}{\partial x_2} - \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2} + 4x_2\frac{\partial^2 f}{\partial x_2^2}\right)$. *We study the behavior of* $X(t)$ *on or before its exit time from the unit disc in Figure 3. Figure 4, parts* (a) *and* (b) *show five and one hundred approximate sample paths of the process. The example is particularly interesting because, by the presence of the square root the process is forced to remain in the upper half of the disc producing a discontinuity in the exit location density. Although outside the domain of our theoretical guarantees, Figure 4, part* (c) *shows the unnormalized Christoffel-Darboux kernel estimates in degrees* $r = 4, 6$ *for the exit location density which seem to have the correct qualitative behavior. Extending our results to cover such densities will be topic of further work. Figure 4, part* (d) *shows the Christoffel-Darboux kernel estimates in degree* $r = 2$ *for the occupation density of* $X(t)$ *before reaching the boundary of the unit disc.*

6 Conclusion

Using elementary analytic arguments, we proved that there is no relaxation gap between the original problem and the linear problem on occupation measures in the special case of evaluating functionals of the exit time of stochastic processes on bounded domains. If the domain is basic semialgebraic and the SDE coefficients and the functional are polynomial or semi-algebraic, we can then readily apply the moment-SOS hierarchy with convergence guarantees. Tight bounds on the functionals can be obtained with off-the-shelf SDP solvers at a moderate cost.

Of particular practical interest are approximations to the moments of the exit time distribution, as studied in [8]. As we have shown, any sufficiently accurate method of moment estimation can be leveraged to recover the exit location and occupation densities allowing us to gain both a quantitative and qualitative understanding of diffusion processes on sufficiently well-behaved geometric domains.

As pointed to us by one of the reviewers of the article, it is natural to ask whether the linear operators arising in our moment relaxation satisfy a flat extension condition. When the ambient space dimension is at least two, both the occupation and exit location measures of a diffusion are typically supported on sets of positive measure (resp. positive measure in the boundary) and a flat extension condition never holds since this would imply that the corresponding measures are supported at points. This is one of the reasons why the Christoffel-Darboux kernel approach we introduce is not only useful but necessary.

We would also like to extend these techniques to optimal stopping times [4] and stochastic optimal control problems with expectation constraints [3], and to develop theoretical guarantees for success of the proposed Christoffel-Darboux kernel approach on general diffusions.

Acknowledgement

We are grateful to Milan Korda, Jean Bernard Lasserre, Matteo Tacchi, as well as anonymous reviewers, for their feedback on this work.

References

[1] A. Barvinok. A course in convexity. AMS, 2002.

[2] A. G. Bhatt, V. S. Borkar. Occupation measures for controlled Markov processes: characterization and optimality. Ann. Proba. 24:1531-1562, 1996.

[3] R. Buckdahn, D. Goreac, M. Quincampoix. Stochastic optimal control and linear programming approach. Appl. Math. Optim 63:257-276, 2011.

[4] M. J. Cho, R. H. Stockbridge. Linear programming formulation for optimal stopping problems. SIAM J. Control Optim. 40(6):1965-1982, 2002.

[5] J. Dahl. Semidefinite optimization using MOSEK. ISMP Berlin, 2012.

[6] L. C. Evans. Partial differential equations. AMS, 1998.

[7] L. C. Evans. An introduction to stochastic differential equations. AMS, 2013.

[8] K. Helmes, S. Röhl, R. H. Stockbridge. Computing moments of the exit time distribution for Markov processes by linear programming. Oper. Res. 49:516-530, 2001.

[9] D. Henrion, M. Korda, J. B. Lasserre. The moment-SOS hierarchy - Lectures in probability, statistics, computational geometry, control and nonlinear PDEs. World Scientific, 2020.

[10] D. Henrion, J. B. Lasserre, J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Meth. Soft. 24(4-5):761-779, 2009.

[11] D. Henrion, E. Pauwels. Linear conic optimization for nonlinear optimal control. Chapter 10 pp. 121-134 in S. Ahmed, M. Anjos, T. Terlaky (Editors). Advances and Trends in Optimization with Engineering Applications. SIAM, 2017.

[12] C. Jiao, R. Kawai. Computable primal and dual bounds for stochastic optimal control. SIAM J. Control. Optim. 58(6):3709-3733, 2020.

[13] K. Kashima, R. Kawai. On weak approximation of stochastic differential equations through hard bounds by mathematical programming. SIAM J. Sci. Comput. 35(1):A1-A21, 2013.

[14] R. Kawai. Explicit hard bounding functions for boundary value problems for elliptic partial differential equations. Comput. Math. Appl. 70:2822-2837, 2015.

[15] A. Kroó, D. S. Lubinsky. Christoffel functions and universality in the bulk for multivariate orthogonal polynomials. Canad. J. Math. 65(3):600-620, 2013.

[16] T. G. Kurtz, R. H. Stockbridge. Existence of Markov control and characterization of optimal Markov controls. SIAM J. Control Optim. 36:609-653, 1998.

[17] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, 2010.

[18] J. B. Lasserre, E. Pauwels, M. Putinar. The Christoffel-Darboux kernel for data analysis. Cambridge University Press, 2022.

[19] J. B. Lasserre, T. Priéto-Rumeau. SDP vs. LP relaxations for the moment approach in some performance evaluation problems. Stoch. Models 20:439-456, 2004.

[20] J. B. Lasserre, T. Priéto-Rumeau, M. Zervos. Pricing a class of exotic options via moments and SDP relaxations. Math. Finance 16:469-494, 2006.

[21] B. Øskendal. Stochastic differential equations - An introduction with applications. 6th edition with corrections. Springer, 2013.

[22] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. 42(3):969-984, 1993.

[23] J. Urbas. Lectures on second order linear partial differential equations. Proc. Center Math. Appl., Australian Nat. Univ., Workshop on Analysis and Geometry Part 1, 34:39-75, 1995.