SDP vs. LP relaxations for the moment approach in some performance evaluation problems

Jean-Bernard Lasserre\textsuperscript{1} and Tomás Prieto-Rumeau\textsuperscript{2}

\textsuperscript{1}Laboratoire d’Analyse et d’Architecture des Systèmes, Toulouse, France
\textsuperscript{2}Universidad Nacional de Educación a Distancia, Madrid, Spain

Abstract

Given a Markov process we are interested in the numerical computation of the moments of the exit time from a bounded domain. We use a moment approach which, together with appropriate semidefinite positivity moment conditions, yields a sequence of semidefinite programs (or SDP relaxations), depending on the number of moments considered, that provide a sequence of nonincreasing (resp. nondecreasing) upper (resp. lower) bounds. The results are compared to the linear Hausdorff moment conditions approach considered for the LP relaxations in [1]. The SDP relaxations are shown to be more general and more precise than the LP relaxations.

Key Words: Computational methods for Markov processes. Semidefinite relaxations. Linear relaxations
I. Introduction

This paper deals with a semidefinite programming (SDP) methodology for evaluating or approximating numerically moments of some exit time distributions for a class of Markov processes. This approach is compared with the linear programming (LP) approach introduced in [1].

**Background and related literature.** The evaluation of certain functionals of Markov processes can be formulated as an infinite-dimensional linear program over a space of distributions on the state space, see for instance [2] and [3]. For first-exit problems we need consider two measures, namely the expected occupation measure up to exit time, and the exit location distribution (see e.g. [4]). In many interesting cases, this LP formulation translates into a (simpler) infinite-dimensional LP with moments of the expected occupation and exit location measures as variables (see e.g. [1]). More precisely, the constraints of this new LP problem are derived from the martingale characterization of the process, which yields the so-called basic adjoint equations, see [4] and [5].

Thus we consider processes whose evolution can be reformulated through linear (martingale) constraints on the moments of the measures, as explained above. The key property of such processes is that their infinitesimal generator maps the polynomials into polynomials. In fact, many interesting processes do have this property: interest rates processes that are useful in mathematical finance, as the Cox-Ingersoll-Ross model that will be analyzed in this paper, the geometric Brownian motion (arising in the Black and Scholes model),
the Ornstein-Uhlenbeck process, or the Roberts-Shiryaev procedures for quality-control models. Examples of geostochastic models can be found in [6] and [7]. For the examples considered in this paper, an analytical formula for the moments of the exit time distribution is available, which allows us to check the accuracy of numerical approximations, and therefore, validate the SDP approach for more general performance evaluation problems.

For practical numerical computations, the equivalent infinite-dimensional LP problem (on the moment sequence) is relaxed by considering only finitely many moments. Finally, observe that, by its definition, the criterion is a linear function of the moments of the exit location distribution. Therefore, maximizing (resp. minimizing) such linear functions provides upper (resp. lower) bounds on the exact values of the moments of the exit time distribution. These bounds depend on the number of moments considered. This is the approach introduced in [1]. So the variables of the LP relaxations introduced in [1] must be interpreted as moments of some measures, and, therefore, additional moment constraints on those variables are needed to enforce that they should be moments of some measures.

In [1], these additional constraints are finite truncations of the infinitely many (one-dimensional and multi-dimensional) Hausdorff moment conditions, see [8] and [9]. As these Hausdorff conditions are also linear, the resulting optimization problems are standard LPs (whose size depends on the number of moments considered). In summary, a typical LP relaxation has two types of linear constraints: martingale constraints induced by the basic adjoint equations, and moment constraints induced by the Hausdorff moment conditions.

**Contribution.** In the SDP approach proposed in the present paper, one replaces the (linear) LP Hausdorff moment constraints of [1] with SDP moment constraints, that is, moment constraints stated in terms of semidefinite positivity of appropriate moment and localizing matrices. Therefore, the resulting optimization problems are now semidefinite programs (SDPs) instead of LPs. Those SDP relaxations are convex problems for which efficient software packages are available. For more details on Semidefinite Programming, the reader is referred to [10]. Thus, the SDP relaxations include the (lin-
ear) martingale constraints (as in [1]), and the new SDP moment constraints in lieu of the LP moment constraints in [1].

Finally, let us mention the recent contribution of Schwerer [11] who proposes LP relaxations for evaluating moments of the steady-state distribution of a reflected Brownian motion. Interestingly, the LP moment conditions are derived from (Stieltjes) SDP moment conditions (as they deal with measures on the positive orthant and its boundaries), so as to end up with a LP.

To compare the SDP and LP relaxations, we have considered the same stochastic models as those investigated in [1] via LP relaxations. Consequently, we can compare the precision of both procedures, the numerical problems that they may encounter, and the respective sizes of the LP and SDP programs needed to achieve bounds with same accuracy.

SDP and LP based relaxations procedures have been compared in several other optimization problems: in [12] for 0-1 programs, or in [13] for polynomial programming. In general, the SDP procedures turn out to be more efficient than the LP relaxations techniques. The main advantages and drawbacks of the SDP and LP relaxations are discussed, from a general point of view, in Section III.D of this paper.

In Section II we briefly introduce the martingale model and the associated basic adjoint equation. Both LP and SDP relaxations are introduced in Section III and are then compared in Section IV on some interesting models. A discussion follows.

II. The martingale model

We consider a Markov process \( Y = \{Y_t\}_{t \geq 0} \) with state space \( E \subseteq \mathbb{R}^d \), and with continuous sample paths. We assume that the state space is partitioned into two sets: \( E_0 \), which is open and bounded, and its complement: \( E_0^c = E - E_0 \). The stopping time \( \tau \) is defined as the first time that the process \( Y \) hits \( E_0^c \):

\[
\tau = \min\{t \geq 0 : Y_t \in E_0^c\}.
\]
We are interested in computing the mean exit time $E\tau$ or, more generally, the moments $E\tau^n$ of the distribution of $\tau$.

Let $A$ be the infinitesimal generator of $Y$, with domain $\mathcal{D}(A)$. Following [4], we say that $Y$ is a solution of the martingale problem if the process:

$$f(Y_t) - f(Y_0) - \int_0^t Af(Y_s)ds, \quad t \geq 0,$$

is a martingale for every $f \in \mathcal{D}(A)$. As a consequence of the martingale property (1), and if $E\tau < \infty$,

$$Ef(Y_\tau) - Ef(Y_0) - E\int_0^\tau Af(Y_s)ds = 0$$

for every $f \in \mathcal{D}(A)$. Let $\mu_0$ be the expected occupation measure up to time $\tau$, and let $\mu_1$ be the the exit location distribution, that is, for every measurable set $B \subset \mathbb{R}^d$,

$$\mu_0(B) := E\int_0^\tau I_{\{Y_s \in B\}}ds; \quad \mu_1(B) := Pr\{Y_\tau \in B\}.$$  

Notice that the measures $\mu_0$ and $\mu_1$ are supported on $E_0$ and $E_0^c$ (in fact $\partial E_0$), respectively. If $Y_0 = y_0 \in E_0$, then (2) becomes:

$$\int_{E_0^c} f(y)\mu_1(dy) - f(y_0) - \int_{E_0} Af(y)\mu_0(dy) = 0,$$

for every $f \in \mathcal{D}(A)$, which is called the basic adjoint equation. Conversely, (4) characterizes the measures $\mu_0$ and $\mu_1$ in (3) of a process having generator $A$; see e.g. [6] or [14].

Our goal is to compute an approximate value of the moments $E\tau^n$ using (4). We first relax condition (4) by replacing it with a condition stated now in terms of the moments of $\mu_0$ and $\mu_1$, as new variables. As we shall see in the next section, in many cases this new condition is linear in the variables. Further, we need to impose some conditions on these variables to enforce that they should be moments of some measures. This can be accomplished in two ways, which yield either LP or SDP relaxations.
III. LP and SDP relaxations

Let \( \{m_j\}_{j \in \mathbb{N}^d} \) and \( \{b_j\}_{j \in \mathbb{N}^d} \) be the moments of the measures \( \mu_0 \) and \( \mu_1 \) defined in (3), i.e.,

\[
m_j := \int_E y^j \mu_0(dy) \quad \text{and} \quad b_j := \int_E y^j \mu_1(dy),
\]

for \( j \in \mathbb{N}^d \), where \( y^j := y_1^{j_1} \cdots y_d^{j_d} \), assumed to be finite for every \( j \in \mathbb{N}^d \). Notice that the mean exit time \( E\tau \) reads

\[
E\tau = \int_E 1 \mu_0(dy) = m_0,
\]

the mass of the expected occupation measure \( \mu_0 \).

Usually, the operator \( A \) is a differential operator and, in many models of practical interest (e.g. the geometric Brownian motion or the Ornstein-Uhlenbeck process) if \( f \) is a polynomial then so is \( Af \). The models considered in [1], as those analyzed in this paper, fulfill this hypothesis. Notice that, in general, polynomials do not belong to the domain \( \mathcal{D}(A) \) of the infinitesimal generator, although under mild conditions the basic adjoint equation (4) also holds when \( f \) is a polynomial. As a consequence, in this particular case, when \( f \) is a monomial, say \( y \mapsto f(y) := y^k, k \in \mathbb{N}^d \), then \( y \mapsto Af(y) = \sum c_j(k)y^j \) for some scalars \( \{c_j(k)\} \), so that the basic adjoint equation (4) can be expressed in terms of the moments of \( \mu_0 \) and \( \mu_1 \). More precisely:

\[
b_k - y_k^0 - \sum_j c_j(k)m_j = 0.
\]

Condition (4), for every \( f \in \mathcal{D}(A) \), is thus relaxed by imposing (6) for every \( k \in \mathbb{N}^d \).

The main issue is now to find some conditions on the sequences \( \{m_j\} \) and \( \{b_j\} \) to enforce that they should be moment sequences of some measures. We consider two different such conditions, either via linear or semidefinite constraints.
A. LP moment constraints

The LP moment constraints are based on the Hausdorff moment conditions, that state necessary and sufficient conditions on a multi-index sequence \( \{m_j\} \) to be moments of a measure supported on a bounded polyhedron (or, polytope), see [9].

As an illustration, suppose that \( \mu \) is a measure supported on \([0, 1]\). Then for every \( k \geq 0 \) and \( n \geq 0 \):

\[
\int_0^1 y^k (1-y)^n \mu(dy) \geq 0. \tag{7}
\]

Let \( \{m_k\}_{k \geq 0} \) be the moments of \( \mu \). The previous inequality can be written:

\[
\sum_{r=0}^n (-1)^r \binom{n}{r} m_{k+r} \geq 0, \quad \text{for } k \geq 0 \text{ and } n \geq 0. \tag{8}
\]

The Hausdorff moment conditions (8) are necessary and sufficient for \( \{m_k\} \) to be the moments of a measure on \([0, 1]\), see e.g. [8, p. 222]. Observe that \( y \mapsto \binom{n+k}{k} y^k (1-y)^n \), for \( 0 \leq y \leq 1 \), cf. (7), is a Bernstein polynomial [8, p. 220].

Recalling (5) and noting that: the linear (martingale) constraints (6), together with the LP (moment) conditions (e.g. (8)) are a relaxation of the basic adjoint equation (4), it turns out that

\[
\max (\text{resp. } \min) \, m_0, \quad \text{subject to}
\]

\[
\begin{cases}
\cdot \ b_k - y_0^k - \sum_j c_j(k)m_j = 0, \quad \text{for } k \in \mathbb{N}^d, \\
\cdot \ \text{LP moment conditions on } \{m_j\}_{j \in \mathbb{N}^d} \\
\quad \text{(supported on } E_0), \\
\cdot \ \text{LP moment conditions on } \{b_j\}_{j \in \mathbb{N}^d} \\
\quad \text{(supported on } \partial E_0),
\end{cases} \tag{9}
\]

provides an upper (resp. lower) bound on \( E \tau \).

However, for computational purposes, we need to restrict ourselves to \textit{finitely many} moments, so that one only considers constraints in (9) involving the moments \( \{m_j\} \) and \( \{b_j\} \) with \( 0 \leq j \leq M \) (meaning that \( j = (j_1, \ldots, j_d) \) verifies \( j_1 + \cdots + j_d \leq M \)).
The issue of the determination of $M$ is very important from a computational point of view. On the one hand, the larger the degree $M$, the more accurate the approximations. But, on the other hand, the size of the problem grows polynomially with $M$; see (e) in the Discussion section, Section III.D. When only moments of order $0 \leq j \leq M$ are involved in the relaxation, we say that we are considering a relaxation with $M$ moments.

B. SDP moment constraints

In this section we provide an alternative derivation of moment conditions, namely the SDP moment conditions, which are based on the semidefinite positivity of some moment and localizing matrices defined below.

**Moment matrix.** Let $d$ be fixed and let:

$$
\{x^\alpha\} := 1, x_1, \ldots, x_d, x_1^2, x_1x_2, \ldots, x_1^k, x_1^{k-1}x_2, \ldots, x_d^k,
$$

be the usual basis of the space of real-valued polynomials in $d$ variables of degree at most $k$.

Let $\{m_\alpha\}$ be a multi-index sequence with $\alpha \in \mathbb{N}^d$ and define the moment matrix $M_k(m)$ with rows and columns indexed in the basis (10), as follows:

$$
\begin{bmatrix}
M_k(m)(1, j) = m_\alpha \\
M_k(m)(i, 1) = m_\beta
\end{bmatrix} \Rightarrow M_k(m)(i, j) = m_{\alpha + \beta}.
$$

For instance, when $d = 1$, $M_k(m)$ is the Hankel matrix

$$
H_k(m)(i, j) = m_{i+j-2} \quad i, j = 1, \ldots, k + 1,
$$

whereas when $d = 2$, $M_1(m)$ reads

$$
M_1(m) = \begin{pmatrix}
m_{00} & m_{10} & m_{01} \\
m_{10} & m_{20} & m_{11} \\
m_{01} & m_{11} & m_{02}
\end{pmatrix}.
$$

Let $\mu$ be a finite measure defined on the Borel sets of $\mathbb{R}^d$ and let $\{m_\alpha\}$ be the sequence of $\alpha$-moments of $\mu$ of order: $|\alpha| := \sum_i \alpha_i$, assumed
to be finite for all $\alpha \in \mathbb{N}^d$, that is:

$$m_\alpha := \int x^\alpha \mu(dx) < \infty, \quad \alpha \in \mathbb{N}^d.$$ 

For all $k \in \mathbb{N}$, the moment matrix $M_k(m)$ is positive semidefinite, (p.s.d., or also denoted $M_k(m) \succeq 0$). Indeed, for all polynomials $x \mapsto f(x)$ of degree at most $k$, with vector of coefficients $\overline{f} \equiv \{f_\alpha\}$ in the basis (10), we have

$$\langle \overline{f}, M_k(m) \overline{f} \rangle = \int f^2 \, d\mu \geq 0.$$ 

The converse is not true in general, that is, given a moment-like matrix $M_k(m) \succeq 0$, the $m_\alpha$’s are not necessarily the moments of some measure $\mu$ on $\mathbb{R}^d$.

**Localizing matrix.** Given a polynomial $q \in \mathbb{R}[x_1, \ldots, x_d]$, we define the localizing matrix $M_k(qm)$ relative to $q$ as follows. Let $\beta(i, j) \in \mathbb{N}^d$ denote the subscript of the entry $(i, j)$ of the matrix $M_k(m)$. If $\{q_\alpha\}$ is the vector of coefficients of $q$ in the basis (10), then the localizing matrix $M_k(qm)$ is defined by:

$$M_k(qm)(i, j) := \sum_\alpha q_\alpha m_{\beta(i, j)+\alpha}.$$ 

For example, if $x \mapsto q(x) := 1 - x_1^2 - x_2^3$, for $x \in \mathbb{R}^2$, then $M_1(qm)$ reads

$$
\begin{pmatrix}
1 - m_{20} - m_{02} & m_{10} - m_{30} - m_{12} & m_{01} - m_{21} - m_{03} \\
m_{10} - m_{30} - m_{12} & m_{20} - m_{40} - m_{22} & m_{11} - m_{31} - m_{13} \\
m_{01} - m_{21} - m_{03} & m_{11} - m_{31} - m_{13} & m_{02} - m_{22} - m_{04}
\end{pmatrix}.
$$

Consider the set $K := \{x \in \mathbb{R}^d : q(x) \geq 0\}$. If $\{y_\alpha\}$ are the moments of some measure $\mu$ supported on $K$ then $M_k(qy)$ is p.s.d. ($M_k(qy) \succeq 0$), because for all polynomials $x \mapsto f(x)$ of degree at most $k$, with vector of coefficients $\overline{f} \equiv \{f_\alpha\}$ in the basis (10),

$$\langle \overline{f}, M_k(qm) \overline{f} \rangle = \int f^2 q \, d\mu \geq 0,$$

Again, the converse is not true, that is, the condition $M_k(qm) \succeq 0$ is not sufficient to ensure that $m$ are the moments of some measure $\mu$ supported on $K$. 
In general, if $K$ is a semi-algebraic set defined by $r$ polynomial inequalities: $g_i(x) \geq 0$, for all $i = 1, \ldots, r$, the SDP conditions

$$M_k(m) \succeq 0; \quad M_k(gm) \succeq 0, \quad i = 1, \ldots, r,$$

are only necessary for $m$ to be moments of some measure $\mu$ supported on $K$. However, by a result in [16], if $K$ is compact and under some mild assumptions, the conditions (11) for all $k = 1, 2, \ldots$, are also sufficient.

Other sufficient conditions can be found in [17]. In particular, a sequence $(m_0, \ldots, m_{2k})$ is a truncated moment sequence of a measure supported on the interval $[a, b]$ (the truncated Hausdorff moment problem) if and only if $M_k(m) \succeq 0$ and $M_{k-1}(gm) \succeq 0$, where $x \mapsto g(x) := (b - x)(x - a)$; and it is a truncated moment sequence of a measure supported on the interval $[a, +\infty)$ (the truncated Stieltjes moment problem) if $M_k(m)$ and $M_{k-1}(gm)$ are positive definite, where $x \mapsto g(x) := x - a$.

**The SDP approach.** In the SDP approach to approximate the mean exit time $E\tau$, one considers the linear martingale constraints (6) and the moment constraints on the moment and localizing matrices associated with $\mu_0$ and $\mu_1$, i.e.,

- the sequence $\{m_j\}_{j \in \mathbb{N}^d}$ and the semi-algebraic set $E_0$, and
- the sequence $\{b_j\}_{j \in \mathbb{N}^d}$ and the semi-algebraic set $\partial E_0$,

instead of the (linear) Hausdorff conditions. Clearly, the linear martingale constraints along with the SDP moment constraints are a relaxation of the basic adjoint equation (4). Consequently, and as for the LP relaxations,

$$\max (\text{resp. } \min) \ m_0, \text{ subject to}$$

$$\left\{ \begin{array}{l}
\bullet \ b_k - y_k^0 - \sum_j c_j(k)m_j = 0, \text{ for } k \in \mathbb{N}^d, \\
\bullet \ \text{moment and localizing matrices for } \{m_j\}_{j \in \mathbb{N}^d} \text{ (supported on } E_0) \text{ are p.s.d.}, \\
\bullet \ \text{moment and localizing matrices for } \{b_j\}_{j \in \mathbb{N}^d} \text{ (supported on } \partial E_0) \text{ are p.s.d.}, \\
\end{array} \right. \quad (12)$$

provides upper (resp. lower) bounds on $E\tau$.

Notice that (12) is a semidefinite program, the SDP relaxation
analogue of the LP relaxation (9). Again, we need to restrict ourselves to \textit{finitely many} moments, and we will consider constraints in (12) involving the moments \(\{m_j\}\) and \(\{b_j\}\) with \(0 \leq j \leq M\).

C. Computing higher moments

In sections A and B we have analyzed the problem of approximating the mean exit time \(E \tau\). For higher order moments \(E \tau^n\), one needs to \textit{augment} the state space of the Markov process so as to include the time component. The new state space is \(\mathbb{R}^+ \times E\). The new infinitesimal generator \(\overline{A}\) of the time-space process is defined as:
\[
\overline{A}(\gamma, f) := \gamma Af + \gamma' f,
\]
where \(f \in \mathcal{D}(A)\) and \(\gamma\) is a \(C^1(\mathbb{R}^+)\) function vanishing at \(+\infty\), see [1, p. 519]. The expected occupation measure \(\mu_0\), and the expected joint time and exit location distribution, \(\mu_1\), are defined in a similar way:
\[
\mu_0(B) := \mathbb{E} \int_0^\tau I_{\{(s, Y_s) \in B\}} ds \quad \text{and} \quad \mu_1(B) := \Pr\{(\tau, Y_\tau) \in B\},
\]
for every \(\mathbb{R}^+ \times E\) measurable set \(B\). The basic adjoint equation is:
\[
\int_{\mathbb{R}^+ \times E_0} \gamma(s) f(y) \mu_1(ds \cdot dy) - \gamma(0) f(y_0) = \int_{\mathbb{R}^+ \times E_0} \overline{A}(\gamma, f)(s, y) \mu_0(ds \cdot dy)
\]
for every \(f \in \mathcal{D}(A)\) and \(\gamma\) a \(C^1(\mathbb{R}^+)\) function vanishing at \(+\infty\).

To obtain the analogues of constraints (6) derived from the basic adjoint equation, we now use the time-space monomials \((t, k) \mapsto t^k y^n\) as test functions. Under our previous assumption on the generator \(A\), it follows that the equations derived from (13) for time-space monomials are also linear functions of the moments of \(\mu_0\) and \(\mu_1\), exactly as before. The linear Hausdorff and p.s.d. moment conditions are also obtained with obvious adhoc modifications.

In conclusion, the constraints in both LP and SDP relaxations to compute the moments of the distribution of \(\tau\) for the \(\{(t, Y_t)\}_{t \geq 0}\) time-space process are the same as before, up to a state augment-
tation. Further, notice that the objective function to be maximized (resp. minimized) to obtain upper (resp. lower) bounds on the moments of $\tau$ is:

$$E \tau^n = n \int_{R^+ \times E_0} s^{n-1} \mu_0(ds \cdot dy) = n \cdot m_{n-1,0}. \quad (14)$$

In the LP relaxations one needs an additional approximation. Indeed, the time component is truncated at some large $T$, and one now considers the exit time from the (new) set $[0,T] \times E_0$ instead of $[0, +\infty) \times E_0$. This is because the LP Hausdorff moment conditions (e.g. (8)) are only valid for measures supported on a compact convex polyhedron (or, polytope). In contrast, in the SDP approach, necessary SDP moment conditions (e.g. (11)) can be derived for arbitrary semi-algebraic sets, not necessarily compact.

**D. Discussion**

Before proceeding to the numerical comparison between LP and SDP relaxation techniques, let us make some general comments on their potential advantages and drawbacks.

(a) The linear Hausdorff moment conditions may cause numerical problems due to the binomial coefficients involved in the constraints, see equation (8). In contrast, no such coefficient appears in the SDP constraints.

(b) The Hausdorff conditions are necessary and sufficient conditions for the $\{m_k\}$ to be moments of some measure (say on $[0,1]$) provided that the whole (infinite) sequence of conditions (8) is considered. A finite (even large) number of conditions does not suffice. In contrast, in the SDP approach (at least in the one-dimensional case), if the p.s.d. condition on the moment matrix is slightly enforced to definite positivity, then one obtains a sufficient condition for a finite sequence $\{m_k\}$ to be a moment sequence (see e.g. [17]).

(c) It is worth mentioning that in the general framework of global optimization, the LP relaxations may not converge. Moreover, even though they may be shown to converge asymptotically,
the convergence might not be finite, whereas finite convergence to the optimal value may be achieved for SDP relaxations. For a detailed discussion on a comparison between LP and SDP relaxations in polynomial programming, the interested reader is referred to [18].

(d) However, an advantage of LP relaxations is that there are powerful LP codes that may deal with (very) large size problems, whereas so far (and because SDP is a relatively recent technique), SDP software packages cannot handle large size problems yet.

(e) Observe that the number of variables involved in the LP problem (9) and in the SDP problem (12) is of order \( \binom{M+d}{d} \) when considering a truncation for \( M \) moments, which grows polynomially in \( M \) when \( d \) remains fixed.

Let us compare the number of constraints of the LP and SDP problems. The linear moment equalities obtained from the basic adjoint equation are the same for both LP and SDP relaxations. Consider now the LP and SDP moment conditions and, for better illustration, let us analyze the case of the polytope \([0,1]^d \subseteq \mathbb{R}^d\). For \( M \) moments, the corresponding LP relaxation requires \( O(M^2d) \) inequality constraints, whereas the corresponding SDP relaxation requires positive semidefiniteness of \( d+1 \) moment and localizing matrices of size \( O((M/2)^d) \). This results in a substantial reduction of the SDP approach computing time with respect to the LP approach computing time, as we shall see in the next section.

IV. Numerical results

In [1], the LP relaxations are used to compute the moments of the exit time distribution of many interesting models such as the Cox-Ingersoll-Ross interest rate model or the Bessel process, among others. These models are also interesting because the formula defining the moments of the exit time distribution can be evaluated numerically, which allows to check the accuracy of the relaxation techniques.
In this section, we present numerical results obtained for the models analyzed in [1], but now under the SDP approach. More precisely, we impose the same linear martingale constraints (6) as in [1], and then we replace the Hausdorff moment conditions with the SDP moment conditions. The bounds obtained by these two different approaches will be compared. An interesting issue is also to compare the accuracies of the two techniques as a function of the number \( M \) of moments of the measures \( \mu_0 \) and \( \mu_1 \) involved in the LP and SDP relaxations.

For the numerical implementation we have used SeDuMi 1.05 (see [19]), a SDP solver Matlab* add-on. We ran version 6.5 of Matlab on a Pentium 4 (1.60GHz and 256MB Ram) computer.

A. The Cox-Ingersoll-Ross interest rate model

Let the process \( Y \) be defined by:

\[
dY_t = (\alpha Y_t + \beta)dt + \sigma \sqrt{Y_t}dW_t,
\]

where \( Y_0 = y_0 \in (r,1) \) and \( \{W_t\} \) is a standard Brownian motion. The constants \( \alpha \in \mathbb{R}, \beta \geq 0, \sigma \geq 0 \) and \( r \in (0,1) \) are given. This is the Cox-Ingersoll-Ross (CIR) interest rate model with constant coefficients. For more details we refer to [20]. We are interested in the exit time of \( Y \) from the interval \((r,1)\). The generator \( A \) of \( Y \) is defined for twice continuously differentiable functions by:

\[
f \mapsto Af(y) := \frac{\sigma^2}{2} y f''(y) + (\alpha y + \beta) f'(y).
\]

Notice that when \( f \) is a polynomial then so is \( Af \). As pointed out in Section III, this allows us to introduce the linear (martingale) constraints (6). In this particular case, the expected occupation measure \( \mu_0 \) is supported on \((r,1)\) and the exit location distribution is concentrated on \( \{ r, 1 \} \). In [1], the authors rescale the measure \( \mu_0 \) to the interval \((0,1)\) and use as test functions in (4) not only the monomials but also the \( \ln \) function to derive an additional constraint from the

*Matlab is a trademark of The MathWorks, Inc.
basic adjoint equation. The results presented in this paper have been obtained using the same approach as in [1]. For an explicit expression of these equations and for an exact formula for $E\tau$, we refer to [1].

Table 1 shows the lower and upper bounds for $E\tau$ when $\beta = 0$, $\sigma = 0.5$, $r = 0.1$, $y_0 = 0.2$, and $\alpha$ takes different values. The results for the LP relaxations for $M = 64$ moments are taken from [1]. The bounds obtained from the SDP relaxations are shown when also $M = 64$, except for $\alpha = 2$ and $\alpha = 3$, where $M = 38$ and $M = 14$ moments are considered, respectively (displayed in italic characters in Table 1). The reason for this is that our solver ran into numerical problems in these two computations for $M = 64$. The ratio given in the last column corresponds to the length of the LP bounding interval divided by the length of the corresponding SDP bounding interval.

Table 2 is similar: we take $\alpha = 1$, $\sigma = 0.5$, $r = 0.1$ and $y_0 = 0.2$, while the parameter $\beta$ varies. The LP relaxation bounds, taken from [1], were obtained for $M = 64$ moments and the SDP relaxation bounds for $M = 12$ moments.

Finally, we compare the efficiency of both methods when the number of moments $M$ varies. The values of the parameters are $\alpha = 1$, $\beta = 0$, $\sigma = 0.5$, $r = 0.1$ and $y_0 = 0.2$. Let $\hat{x}_M$ be the lower bound of the LP relaxations and $\hat{y}_M$ the lower bound for the SDP relaxations when $M$ moments are considered. Table 3 shows, for some choices of $M$, the value of $(E\tau - \hat{x}_M)/(E\tau - \hat{y}_M)$, which is denoted by LP/SDP in the table.

The computing times to solve the SDP relaxations for the Cox-Ingersoll-Ross model ranged from 4 to 7 seconds.

Let us make some comments on the use of the $\ln$ function as a test function. As an illustration, for the case $\alpha = -0.5$ in Table 1, the bounding interval for $M = 64$ moments and without the $\ln$ function is [0.632667, 0.632922]. This interval is 1.2 times shorter than the corresponding LP interval with the $\ln$ function, whereas the SDP interval with the $\ln$ function is 16 times shorter than the LP interval with the $\ln$ function. Therefore, using the $\ln$ function considerably improves the bounding intervals.
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>lower bounds</th>
<th>upper bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP</td>
<td>SDP</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.427392</td>
<td>0.427803</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.632675</td>
<td>0.632745</td>
</tr>
<tr>
<td>0.0</td>
<td>0.937386</td>
<td>0.937695</td>
</tr>
<tr>
<td>0.5</td>
<td>1.048685</td>
<td>1.0488830</td>
</tr>
<tr>
<td>1.0</td>
<td>0.950042</td>
<td>0.950455</td>
</tr>
<tr>
<td>2.0</td>
<td>0.70233</td>
<td>0.703578</td>
</tr>
<tr>
<td>3.0</td>
<td>0.52267</td>
<td>0.529351</td>
</tr>
</tbody>
</table>

Table 1. CIR model. Bounds on $\mathbb{E} \tau$. LP: $M = 64$, SDP: $M = 64$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>lower bounds</th>
<th>upper bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP</td>
<td>SDP</td>
</tr>
<tr>
<td>0.0</td>
<td>0.950042</td>
<td>0.950295</td>
</tr>
<tr>
<td>0.2</td>
<td>0.936516</td>
<td>0.945262</td>
</tr>
<tr>
<td>0.5</td>
<td>0.731062</td>
<td>0.754963</td>
</tr>
<tr>
<td>1.0</td>
<td>0.501042</td>
<td>0.514956</td>
</tr>
<tr>
<td>2.0</td>
<td>0.304351</td>
<td>0.310647</td>
</tr>
<tr>
<td>5.0</td>
<td>0.134958</td>
<td>0.142242</td>
</tr>
</tbody>
</table>

Table 2. CIR model. Bounds on $\mathbb{E} \tau$. LP: $M = 64$, SDP: $M = 12$.

<table>
<thead>
<tr>
<th>$M$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP/SDP</td>
<td>1.8</td>
<td>49</td>
<td>384</td>
<td>289</td>
<td>94</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 3. CIR model. Ratio of the LP and SDP intervals length.
B. The Bessel process

The Bessel process of dimension \( d \geq 1 \) verifies the stochastic differential equation:

\[
dY_t = \frac{d - 1}{2Y_t} dt + dW_t,
\]

where \( Y_0 = y_0 \) and \( \{W_t\} \) is a standard Brownian motion. Its infinitesimal generator \( A \) is given by:

\[
f \mapsto Af(y) := \frac{1}{2} f''(y) + \frac{d - 1}{2y} f'(y).
\]

The stopping time \( \tau \) is the exit time of the Bessel process from the interval \((r, 1)\), where \( 0 < r < y_0 < 1 \). The test functions that we consider are the monomials and the \( \ln \) function, so that the linear constraints obtained from the basic adjoint equation are written as a function of the moments of the measure \( \nu \) that satisfies:

\[
\frac{d\nu}{d\mu_0}(y) = \frac{1}{y^2}.
\]

In Table 4 we compare the results obtained in [1] with the LP relaxations for \( M = 64 \) moments and the results obtained via the SDP relaxations for \( M = 12 \) moments. The last column of the table displays the exact value of \( E\tau \).

For the SDP approach, the running times for the Bessel process did not exceed 5 seconds, whereas the LP approach approximations took from 10 to 13 seconds computing time [1, p. 529].
C. Change-point detection

The change-point detection procedures, such as Roberts-Shiryaev or CUSUM, for example, deal with the problem of detecting changes in the behavior of a system. They have been analyzed for both discrete and continuous time models, see for instance [21]. For continuous time models, the change-point detection procedures can be represented as one-dimensional diffusions. The stochastic differential equation of the basic model is:

\[ dZ_t = \rho I_{\{t>\theta\}} dt + \sigma dW_t, \]

where the parameters \( \rho \) and \( \sigma \) in the above equation are known. When \( t > \theta \), which is an unknown parameter, there is a change in the drift of the process. The monitoring statistic \( Y_t \) based on the likelihood ratio, see [1], is a solution of:

\[ dY_t = (1 + 2Y_t) dt + \sqrt{2} Y_t dW_t, \]

where \( Y_0 = 0 \). The infinitesimal generator \( A \) of \( Y \) is:

\[ f \rightarrow Af(y) := y^2 f''(y) + (1 + 2y) f'(y). \]

In the Roberts-Shiryaev procedure, we are interested in the random time \( \tau \) which is defined as the first time that \( Y_t \) is greater than some \( H \), which is a given limit. In order to compute \( \mathbb{E} \tau \), notice that the occupation measure \( \mu_0 \) has its support on \([0, H]\) and that \( \mu_1 \) is a Dirac probability on \( \{H\} \). The test functions used in [1] are the monomials, after rescaling \( \mu_0 \) to the unit interval.

The bounds on the mean exit time are computed for several values of \( H \). In order to compare the LP and SDP relaxations, we have determined the number of moments needed in the SDP relaxations to reach the same accuracy as the LP relaxations for the examples shown in [1]. As an illustration, for \( 7 \leq H \leq 10 \), the SDP moment approach with \( M = 8 \) moments and the LP moment approach with \( M = 30 \) moments have the same accuracy.
D. Two-dimensional Brownian motion

So far, we have analyzed one-dimensional processes. In this section we compare the techniques of the LP relaxations and SDP relaxations for a two-dimensional process. More precisely, we are going to obtain an approximation, via the relaxations techniques, of the mean exit time from the open unit square: $E_0 = (0,1) \times (0,1)$, of a two-dimensional Brownian motion with initial value $(x_0, y_0) \in E_0$. We consider the process:

$$(X_t, Y_t) = (x_0 + W_{1,t}, y_0 + W_{2,t}),$$

where $W_{1,t}$ and $W_{2,t}$ are independent standard Brownian motions starting at $W_{1,0} = W_{2,0} = 0$. The infinitesimal generator $A$ of $(X,Y)$ is:

$$Af(x,y) := \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x,y) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(x,y),$$

for twice continuously differentiable functions $f : \mathbb{R}^2 \to \mathbb{R}$.

The occupation measure $\mu_0$ is concentrated on $E_0$ while the support of the exit location distribution $\mu_1$ is the boundary of the unit square. In order to implement the codes, we have decomposed $\mu_1$ as the sum of four one-dimensional measures on the unit interval, each one corresponding to one edge of the unit square, where the sum of their total masses is equal to one. The monomials $(x,y) \mapsto f(x,y) := x^m y^n$ are used in the basic adjoint equation to derive the linear (martingale) constraints on the moments of $\mu_0$ and $\mu_1$.

In the numerical results in [1], several starting points $(x_0, y_0)$ are considered and the upper and lower bounds on the mean exit time are shown for $M = 11$ moments. We have analyzed the same examples with the same number of moments: $M = 10$ (notice that due to the nature of the SDP relaxations, the number of moments is always an even number). The values shown in Table 5 are the lengths of the bounding intervals obtained for the LP relaxations divided by the lengths of the bounding intervals of the SDP relaxations. Notice that, for instance, the exact values for $(x_0, y_0) = (0.5,0.7)$ and $(x_0, y_0) = (0.5,0.3)$ need to be the same because of the symmetry of the problem. That is why the $x_0 = 0.5$ column in Table 5 is incom-
Table 5. Two-dimensional Brownian motion. Ratio LP/SDP of the intervals length.

For this example, the running times to solve the SDP problems were always below 8 seconds. The LP approach took 24 minutes using CPLEX 5.0 and 6.0 on an Ultra 10/300 [1, p. 529]. Indeed, back to (e) in the discussion of Section III.D, for this example both LP and SDP relaxations have 77 variables, but the LP moment conditions involve 1265 inequalities, whereas the SDP approach imposes that three matrices of dimension 21 and eight matrices of dimension 6 are p.s.d..

E. Time-space Poisson process

This example is concerned with the computation of higher moments of $\tau$, see Section III.C.

We consider a Poisson process $N$ with constant rate equal to $\lambda$. Though it is a one-dimensional process, in order to compute the moments of $\tau$ we must introduce the time component, so that we will deal with the two-dimensional process $Y = \{(t, N_t)\}_{t \geq 0}$. Notice that the state space of $Y$ is $E = [0, +\infty) \times \mathbb{Z}^+$, and that the infinitesimal generator of the process is written as a differential and difference operator:

$$f \mapsto Af(t, x) := \frac{\partial f}{\partial t}(t, x) + \lambda f(t, x + 1) - \lambda f(t, x),$$
for \((t, x) \in E\), and where \(f(\cdot, x) \in C^1[0, +\infty)\) for every \(x \in \mathbb{Z}^+\).

We define \(\tau\) as the time of the \(K\)-th arrival of the Poisson process. The time component is truncated at \(T\), so that \(\mu_0\) is supported on:

\[ E_0 = [0, T) \times \{0, \ldots, K - 1\}, \]

and \(\mu_1\) on:

\[ [0, T] \times \{K\} \cup \{T\} \times \{0, \ldots, K - 1\}. \quad (15) \]

In this case, as the state space has a continuous component and a discrete component, the measure \(\mu_0\) is decomposed into \(K\) one-dimensional measures. As in the preceeding example, the exit location distribution is also decomposed according to (15). We have considered as test functions the time-monomials for each

\[ k \in \{0, \ldots, K - 1\}. \]

Observe that the sample paths of the time-space Poisson process are not continuous. This is not a problem as long as the basic adjoint equation is satisfied, and provided that we can determine the support of the exit location measure. The bounds on the moments of \(\tau\) are obtained when the criterion given in (14) is maximized and minimized.

For the numerical examples, a Poisson process of rate \(\lambda = 1\) is considered. We want to obtain approximations of the moments \(E \tau^n\), for \(n = 4, 5, \ldots, 8\), where \(\tau\) is the first time that \(N\) reaches \(K = 8\). The time process is truncated at \(T = 60\). As in the previous table, Table 6 shows the quotient LP/SDP of the lengths of the bounding intervals on the moments of \(\tau\) computed for the same number of moments: \(M = 20\), for both LP and SDP relaxations.

Our running times ranged from 4 to 8 seconds.

It is also worth noting that if we do not truncate the time component, then the martingale moment conditions (6) yield a determinate system that completely specifies the moments of the expected occupation measure and the exit location probability measure. Thus, truncating the time component is not necessary from a theoretical point of view, but it allows to check the accuracy of the numerical methods.
Table 6. Time-space Poisson process. Ratio of the intervals length.

<table>
<thead>
<tr>
<th>n</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP/SDP</td>
<td>48</td>
<td>290</td>
<td>$10^6$</td>
<td>782</td>
<td>4000</td>
</tr>
</tbody>
</table>

V. Conclusion

For all the examples analyzed in this paper, the SDP relaxations methodology has been shown to be much more precise than the LP relaxations technique.

We have shown that for a similar precision, the LP relaxations need consider many more moments; for instance, for the Cox-Ingersoll-Ross model (Table 2) and the Bessel process (Table 4), SDP relaxations with $M = 12$ moments provide in every case tighter bounds than the LP relaxations with $M = 64$ moments, even with less computational effort (a reduction from 10 seconds to 5 seconds running time). The same happened in the change-point detection model analyzed in Section IV.C: SDP relaxations with $M = 8$ moments were as accurate as LP relaxations with $M = 30$ moments.

We have provided several examples for which the bounding intervals obtained with the SDP methodology are much shorter than those given by the LP technique, when the same number of moments is considered for both procedures. In Section IV.A, they are typically between 20 and 30 times shorter for $M = 64$ moments; see Table 1. For the two-dimensional Brownian motion analyzed in Section IV.D, the SDP intervals are around ten times shorter than the LP intervals; see Table 5. This ratio can also dramatically increase, as for the time-space Poisson process; see Table 6. When making comparisons for various values of $M$ the bounding intervals exhibit a similar behavior; see Table 3.

Also, the running times are lower for the SDP approach, even reaching a reduction in the computing time from 24 minutes to 8 seconds; see the example in Section IV.D.

Moreover, notice that a LP relaxation with $M = 64$ moments may cause numerical problems because of the binomial coefficients in-
involved in the constraints. This is certainly the reason why the bounds provided in [1] by the primal and the dual linear programs are different.

The solver we used could sometimes handle problems involving $M = 64$ moments (see Table 1), but when using more than $M = 40$ moments, the solver frequently ran into numerical problems. However, it is worth noting that this is because SDP solvers are nowadays less developed than their LP counterpart. In spite of this, the SDP relaxations method itself outperforms the LP relaxations approach, as it has been shown.

Concerning the computing times when using the SDP approach, these never exceeded 8 seconds, and they were not very much affected by the number of moments $M$. Indeed, our solver reached optimality quite quickly regardless of the problem size.

The theoretical convergence of the bounds of the SDP relaxations as $M \to +\infty$ holds provided that the measures $\mu_0$ and $\mu_1$ are completely determined by their moment sequences (general sufficient moment conditions may be found in [8, p.224]). For proofs of convergence in a similar context, we refer to [22].

To conclude, recall that the aim of this paper was to validate the SDP relaxations technique and compare it with the LP relaxations method for some concrete exit time problems. In view of the substantial gains in generality and precision, the results obtained with the SDP moment approach are very encouraging. Finally, it is worth noting that the moment approach is a very general technique that could be used in a wide variety of performance evaluation problems.

Acknowledgement. The authors wish to thank anonymous referees for their remarks and suggestions, that have helped us to improve the initial version of the paper.

References

1. Helmes, K.; Röhl, S.; Stockbridge, R.H. Computing moments of the exit time distribution for Markov processes by linear pro-
24

Lasserre and Prieto-Rumeau