A Geometrical Characterization of Multi-dimensional Hausdorff Polytopes with Applications to Exit Time Problems

Kurt Helmes

Institute for Operations Research, Humboldt University of Berlin, Berlin, Germany,

helmes@wiwi.hu-berlin.de

Stefan Röhl

Vorarlberg University of Applied Sciences, Dornbirn, Austria, stefan.roehl@fhv.at

Abstract. We present a formula for the corner points of the multi-dimensional Hausdorff polytopes and show how this result can be used to improve linear programming models for computing, e. g., moments of exit time distributions of diffusion processes. Specifically, we compute the mean exit time of two-dimensional Brownian motion from the unit square, as well as higher moments of the exit time of time-space Brownian motion, i.e. the two-dimensional process composed of a one-dimensional Wiener process and the time component, from a rectangle. The corner point formula is complemented by a convergence result which provides the analytical underpinning of the numerical method which we use.

Key Words. Linear programming, Hausdorff polytopes, moment problems, Brownian motion, numerical methods for exit time problems.

Subject Classification (AMS 2000). *Primary* 52B12. *Secondary* 90C08, 44A60, 60J65, 60J70.

1 Introduction

In articles by Cho (2000), Cho and Stockbridge (2002), Helmes (1999), (2002), Helmes et al. (2001), Helmes and Stockbridge (2000), (2001), (2003),Mendiondo and Stockbridge (1998), and Röhl (2001) numerical methods for the computational analysis of exit time problems, invariant distributions of diffusions and optimal stopping and control problems have been proposed which are based on a linear programming approach to these kind of problems. The formulation of infinite dimensional linear programs for such problems is an extension of work by Manne (1960) who initiated the formulation of stochastic control problems as linear programs over a space of stationary distributions for the long term average control of finite state Markov chains, see Hernandez-Lerma et al. (1991) for details and additional references. The generalization of the LP-formulation for continuous time, general state and control spaces, and different objective functions has been established by Stockbridge (1990), Kurtz and Stockbridge (1998), (1999), and Bhatt and Borkar (1996).

The basic idea of the LP-approach to the analysis of controlled and uncontrolled Markov processes is to formulate such problems as linear programs over a space of stationary distributions. Specifically, the variables in these infinite dimensional linear programs are measures on the product of the state and control spaces and in the case of exit problems, each such variable is augmented by a second measure on the exterior of the state space. These variables are constrained by equations involving the generator of the Markov process and a family of test functions. Different numerical methods are determined by a judious choice of a finite set of test functions combined with a selection of a finite number of variables and/or restrictions imposed on the support of the occupation measure and the exterior measure. Such choices determine approximations of the infinite dimensional optimization problem by finite dimensional ones.

One class of approximating problems exploits the characterization of measures on bounded intervals by their moments and the identification of moment sequences by a countable

family of linear inequality conditions. Hausdorff (1921) and (1923) formulated these inequalities for measures on the interval $[0, 1]$. We therefore call these inequality conditions the Hausdorff conditions, see Section 2. Hildebrandt and Schoenberg (1933) generalized these results to the multi-dimensional case, i.e. to measures with support in $[0, 1]^d$, $d \geq 1$. For some applications of moment theory see Ang et al. (2002). Using but a finite number of these inequalities to partially describe the feasible set of finite dimensional LPs, cf. Section 3, leads us to study the geometry of what we call d-dimensional Hausdorff polytopes. Specifically, we are interested in formulae for the vertices of these polytopes since, as will be illustrated by numerical examples, such formulae enhance the accuracy of the numerical methods to which we referred above. The geometry of moment spaces for the one-dimensional case was first considered in detail in the paper by Karlin and Shapley (1953). To approximate moment sequences they introduced special simplices defined as the convex hull of specific points. In Section 2 we prove that the d-dimensional Hausdorff polytopes are in fact extensions to higher dimensions of the simplices described by Karlin and Shapley.

Since, for dimension 1, the Hausdorff polytopes of order n contain the first $n + 1$ components of all moment sequences we work with an *outer* approximation of the projection of moment sequences onto $\mathbb{R}^{n+1}$. This approximation greatly differs from the *inner* approximation of this set by cyclic polytopes, cf. Ziegler (1995) and also Karlin and Shapley (1953). Using an outer approximation ensures that no restrictive assumptions on the support of the occupation measures of Markov processes to be analyzed need to be made, cf. Section 3. As will be seen in Section 2 these ideas can be generalized to higher dimensions.

The fact that we work with an outer approximation of the projection of moment sequences onto $\mathbb{R}^{n+1}$ is actually one of the advantages of this particular LP-approximation when computing, for instance, mean exit time problems. By solving two finite dimensional LP-problems, a maximization problem and a corresponding minimization one, cf. Section 3 for details, we naturally obtain numerical upper bounds and lower bounds for the quantity

of interest. Moreover, our convergence result guarantees that in the limit these bounds coincide with the exact value.

An alternative to the finite dimensional linear programs referred to above was proposed by Lasserre and Rumeau (2004). Instead of LPs they suggest using semidefinite programs (SDP-method), cf. also Schwerer (2001). Recently, Riechert (2004) compared the SDP-method with the original LP-approach of Helmes, Röhl, and Stockbridge (2001) as well as with the modified LP-technique which is based on the formulae for the corner points of the Hausdorff polytopes to be presented in Section 2. Using some of the examples that had also been analyzed in Helmes, Röhl, and Stockbridge (2001) and Lasserre and Rumeau (2004) as test cases we demonstrate that by using the corner point formulas the modified LP-method yields much better results than the original method as advertised in Helmes, Röhl, and Stockbridge (2001), and that the new technique yields the same degree of accuracy as the SDP-method.

The classical method for analyzing exit time problems is solving an associated PDE, see Section 3 for details. Both methods have advantages and disadvantages. While the PDE-approach simultaneously provides mean exit times for all initial values, the LP-approach requires solving a new linear program for each starting point. On the other hand, the LP-method provides numerical upper and lower bounds simultaneously for a large number of moments of the exit time distribution. Thus unlike in the case of the PDEs where one has to solve different partial differential equations for each moment and often for a very fine mesh, it is only necessary to solve *one* LP to be close to the exact values of several moments.

The LP-method can also be used to analyze exit time problems on general regions. To this end, triangulation techniques are combined with a geometrical characterization of Dale polytopes. While Hausdorff Polytopes are associated with measures whose support is contained in a hypercube, Dale Polytopes are associated with measures defined on the unit triangle $S^d = \{x \in [0,1]^d \mid \sum_{i=1}^{d} x_i \leq 1\}$. Necessary and sufficient conditions

for a doubly indexed sequence to be the sequence of (joint) moments of a measure on the two-dimensional triangle S^2 have been given by Dale (1987); for extensions of Dale's result see Helmes and Stockbridge (2003) and, especially, Stockbridge (2003). As for Hausdorff polytopes one can also derive formulae for the vertices of Dale polytopes as well as generalized Dales polytopes, cf. Decker (2006).

This paper is organized as follows. Section 2 presents our characterization results on Hausdorff polytopes. We introduce convenient notation in Sections 2 which should facilitate reading the proofs. The most important results of the manuscript are the formulae for the corner points of the Hausdorff polytopes and the convergence result, cf. Theorems 2.4 and 2.6. The connection between the vertices of these polytopes and Dirac measures on $[0,1]^d$, given in terms of iterated differences, is described by Proposition 2.5, and by the remark which follows this Proposition. Theorems 2.6 and 3.1 are fundamental convergence results which provide one justification for the LP-technique to be presented in Section 3. This section includes numerical illustrations and briefly recaps the formulation of the LP-approach for these examples.

2 The multi-dimensional Hausdorff polytope

For multi-indices $j = (j_1, \ldots, j_d) \in \mathbb{Z}_+^d$, $n = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d$, and a vector $x \in \mathbb{R}^d$ we use the following abbreviating notation:

$$
(1) \qquad \binom{n}{j} := \prod_{i=1}^{d} \binom{n_i}{j_i}, \qquad x^j := \prod_{i=1}^{d} x_i^{j_i}, \quad \text{and} \quad \sum_{j=0}^{n} := \sum_{j_1=0}^{n_1} \cdots \sum_{j_d=0}^{n_d}.
$$

For a given real valued function u on $E^d = [0,1]^d$ and multi-index $n \in \mathbb{Z}_+^d$ we call the polynomial

$$
(2) \qquad B_{n,u}(x) := \sum_{j=0}^{n} u_j^{(n)} \binom{n}{j} x^j (1-x)^{n-j}, \qquad \text{for } x \in E^d,
$$

the Bernstein Polynomial of degree n corresponding to u where $u_j^{(n)} = u(\frac{j}{n}) := u(\frac{j_1}{n_1}, \ldots, \frac{j_d}{n_d})$ and $\mathbf{1} = (1, \ldots, 1)$ is the main diagonal vector in E^d (cf. Knill (1997)).

5

For any finite or infinite multi-indexed sequence $\{x_n\}$ we define the differences

$$(\triangle_i^1 x)_{(n_1,\ldots,n_d)} := x_{(n_1,\ldots,n_i,\ldots,n_d)} - x_{(n_1,\ldots,n_i+1,\ldots,n_d)}, \quad i = 1,\ldots,d,$$

for all indices for which the right hand side is well defined. Note that we are using backward differences and therefore follow the same sign convention as Dale (1987) and Knill (1997), in contrast to the classical notation used by Feller (1971). This convention avoids unwanted factors in some of the expressions below. Using the notation

$$(\triangle_1^0 \ldots \triangle_d^0 x)_{(n_1,\ldots,n_d)} := x_{(n_1,\ldots,n_d)},$$

we define the iterated differences of higher order as follows:

(3) $\quad (\triangle_1^{j_1} \ldots \triangle_i^{j_i+1} \ldots \triangle_d^{j_d} x)_{(n_1,\ldots,n_d)} := (\triangle_i^1 (\triangle_1^{j_1} \ldots \triangle_i^{j_i} \ldots \triangle_d^{j_d} x))_{(n_1,\ldots,n_d)} =$

$$(\triangle_1^{j_1} \ldots \triangle_i^{j_i} \ldots \triangle_d^{j_d} x)_{(n_1,\ldots,n_i,\ldots,n_d)} - (\triangle_1^{j_1} \ldots \triangle_i^{j_i} \ldots \triangle_d^{j_d} x)_{(n_1,\ldots,n_i+1,\ldots,n_d)}.$$

For such higher order differences we use the abbreviating notation

$$\triangle^j x_n := (\triangle_1^{j_1} \ldots \triangle_d^{j_d} x)_{(n_1,\ldots,n_d)},$$

where $j = (j_1,\ldots,j_d)$, and $n = (n_1,\ldots,n_d)$ are multi-indices.

Let T_i, $1 \le i \le d$, denotes the shift operator applied to the i-th coordinate, i.e.

$$(T_i x)_n = x_{n_1,\ldots,n_{i-1},n_i+1,n_{i+1},\ldots,n_d}.$$

Then we can write the operator $\triangle^j$ as a product of differences of simple commuting operators, viz.

$$\triangle^j = \prod_{i=1}^{d} (\triangle^{\mathbf{0}} - T_i)^{j_i}.$$

Let X be a random variable on $E^d = [0,1]^d$ distributed according to a distribution function F having moments $\underline{\mu} = \{\mu_j\}_{j \in \mathbb{Z}_+^d}$, i.e.

$$\mu_j = \mu_{(j_1,\ldots,j_d)} = \int_{E^d} x_1^{j_1} x_2^{j_2} \cdots x_d^{j_d} \, dF(x_1,\ldots,x_d).$$

By induction over the sum of the components of k and m we see that for all $k, m \in \mathbb{Z}_+^d$ the following equation holds

(4)
$$(\triangle^m \underline{\mu})_k = \int_{E^d} x^k (\mathbf{1} - x)^m \, dF(x).$$

Integrating Equation (2) with respect to F we obtain the identity,

$$\mathbf{E}_F B_{n,u} = \sum_{j=0}^{n} u_j^{(n)} \binom{n}{j} (\triangle^{n-j} \underline{\mu})_j.$$

We define

$$p_j^{(n)} := \binom{n}{j} (\triangle^{n-j} \underline{\mu})_j \qquad \text{for} \quad j = (j_1, \ldots, j_d), \; 0 \le j_i \le n_i.$$

It follows from Equation (4) that $p_j^{(n)} \ge 0$ for all $j \le n$. Using Definition (2) with the function $u(x) \equiv 1$ we obtain $\mathbf{E}_F B_{n,u} = 1$, and hence

$$\sum_{j=0}^{n} p_j^{(n)} = 1.$$

So we may interpret the vector $\{p_j^{(n)}\}_{j=0}^n$ as a discrete distribution function $F^{(n)}$ with jumps on the set of points $\{(\frac{j}{n})\}_{j=0}^n = \{(\frac{j_1}{n_1}, \ldots, \frac{j_d}{n_d})\}_{j=0}^n$. Here the notation $\{.\}_{j=0}^n$ means $\{ . \mid 0 \le j_1 \le n_1, \ldots, 0 \le j_d \le n_d \}$.

The following d-dimensional Hausdorff conditions, see (5) and (6) below, are straightforward generalizations of the conditions for the one- and two-dimensional cases.

Theorem 2.1 *A multi-indexed sequence $\underline{\mu} = \{\mu_j\}_{j \ge 0}$ of real numbers is a sequence of moments of a random variable on $E^d = [0, 1]^d$ iff*

(5)
$$(\triangle^m \underline{\mu})_k \; \ge \; 0$$

or, equivalently,

(6)
$$\sum_{s=0}^{m} (-\mathbf{1})^s \binom{m}{s} \mu_{k+s} \; \ge \; 0$$

for all multi-indices $k, m \in \mathbb{Z}_+^d$.

For the proof of Theorem 2.1 we refer the reader to Hildebrandt and Schoenberg (1933), Shohat and Tamarkin (1943) p. 9 ff or Knill (1997).

Theorem 2.1 suggest the following definitions.

Definition 2.2

1. *For each element* $z \in \mathbb{R}^{\tilde{n}}$, $\tilde{n} = (n_1 + 1) \cdot \ldots \cdot (n_d + 1)$, *we define the linear transformation* $R^{(n)} : \mathbb{R}^{\tilde{n}} \to \mathbb{R}^{\tilde{n}}$ *by*

$$\left(R^{(n)}(z)\right)_j := \binom{n}{j} (\triangle^{n-j} z)_j \quad \text{for} \quad j = (j_1, \ldots, j_d), \; 0 \leq j_i \leq n_i.$$

2. *Let* $\tilde{\mathcal{H}}_n^d \subset \mathbb{R}^{\tilde{n}}$ *be the set of all arrays* $\{z_j\}_{j=0}^n$ *fulfilling the Hausdorff conditions up to order* n, *i.e. we require the Inequalities (5) or (6) to hold only for multi-indices* k *and* m *which satisfy* $k + m \leq n$. *The set* $\tilde{\mathcal{H}}_n^d$ *is called the d-dimensional Hausdorff Polygon of order* n.

3. *For* $c > 0$ *we define the set*

$$\mathcal{H}_{n,c}^d := \left\{ z \in \tilde{\mathcal{H}}_n^d \mid z_{(0,\ldots,0)} = c \right\}.$$

   The set $\mathcal{H}_n^d := \mathcal{H}_{n,1}^d$ *is called the d-dimensional Hausdorff polytope of order* n.

4. *Let* $\mathcal{K}_n^d(\partial E^d) := \left\{ \sum_\varphi \nu^\varphi \; \middle| \; \nu^\varphi \in \mathcal{H}_{n,c_\varphi}^{d-1}, \; \sum_\varphi c_\varphi = 1, \; \varphi \text{ runs over all } (n-1)\text{-dimensional facets of } E^d \right\}.$

Remark: The transformation $R^{(n)}$ is an extension of the mapping $\{\mu_j\}_{j=0}^n \mapsto p_j^{(n)}$, see above, where $\{\mu_j\}_{j=0}^n$ is the truncated sequence of moments of a finite measure on $[0, 1]^d$.

Lemma 2.3 *For a finite sequence* $z = \{z_j\}_{0 \leq j \leq n}$, $n = (n_1, \ldots, n_d)$, *the Hausdorff conditions up to order* n *are equivalent to the following reduced number of conditions (7) or (8):*

(7) $$(\triangle^{n-k} z)_k \geq 0,$$

or, equivalently,

(8) $$\sum_{i=0}^{n-k}(-\mathbf{1})^i\binom{n-k}{i} z_{k+i} \geq 0$$

for all multi-indices k with $0 \leq k \leq n$ component-wise.

PROOF. These conditions are obviously necessary, cf. (5) and (6). To see that Condition (7) or Condition (8) are also sufficient we observe that the defining equation for iterated differences imply

(9) $$(\triangle^{m-e_i}z)_j = (\triangle^m z)_j + (\triangle^{m-e_i}z)_{j+e_i}$$

for all m, j with $m+j \leq n$, and each component $i \in \{1, \ldots, d\}$, where e_i denotes the i-th unit vector in $\mathbb{R}^d$ and $m_i \geq 1$. Using the shift operator T_i Equation (9) can be written as

$$\triangle^{m-e_i} = \triangle^m + T_i \triangle^{m-e_i} .$$

Next, starting with $m = n - k$ and $j = k$ we know that by assumption both terms on the right hand side of Equation (9) are nonnegative. So we obtain the inequalities

$$(\triangle^{n-k-e_i}z)_k \geq 0$$

for all $i \in \{1, \ldots, d\}$. Repeating this kind of reasoning we show by iteration that component-wise $(\triangle^m z)_j \geq 0$ holds for all multi-indices m, j with $m + j \leq n$. $\square$

The following representation theorem is the main contribution of Section 2.

Theorem 2.4 *The Hausdorff polytope $\mathcal{H}_n^d$ has $\tilde{n} := \prod_{i=1}^{d}(n_i+1)$ corner points. The corner point $z^{(k;n)}$ associated with the multi-index $k = (k_1, \ldots, k_d)$, $0 \leq k_i \leq n_i$, $i = 1, \ldots, d$, has the following coordinates (which are arranged according to the lexicographic order)*

$$z^{(k;n)} = \left(z^{(k;n)}_{(0,\ldots,0,0)}, \ldots, z^{(k;n)}_{(0,\ldots,0,n_d)}, \ldots, z^{(k;n)}_{(n_1,\ldots,n_{d-1},0)}, \ldots, z^{(k;n)}_{(n_1,\ldots,n_{d-1},n_d)} \right),$$

where

(10) $$z^{(k;n)}_j = \prod_{i=1}^{d} \binom{n_i}{k_i}^{-1}\binom{n_i-j_i}{k_i-j_i} = \binom{n}{k}^{-1}\binom{n-j}{k-j};$$

for $t < 0$ we define the binomial coefficient $\binom{s}{t} = 0$ for all $s \in \mathbb{Z}_+$. In the sequel we shall drop the index n whenever feasible and write $z^{(k)}$ instead of $z^{(k;n)}$.

9

PROOF. A finite sequence $z = \{z_j\}_{0 \leq j \leq n}$ is an element of $\mathcal{H}_n^d$ if and only if (i) $z_{(0,\ldots,0)} = 1$ and (ii), cf. Lemma 2.3, the $\tilde{n} = \prod_{i=1}^{d}(n_i+1)$ inequalities

$$(11) \qquad \sum_{j=0}^{n-l}(-\mathbf{1})^j \binom{n-l}{j} z_{l+j} \geq 0$$

are satisfied for all multi-indices l with $0 \leq l \leq n$. In order to show that the set of vectors $z^{(k)}$, see (10), is the set of all corner points of $\mathcal{H}_n^d$ we need to prove that, first of all, $z_{(0,\ldots,0)}^{(k)} = 1$. Moreover, since the vectors $z^{(k)}$ are linearly independent, we need to show that for all vectors $z^{(k)} = (1, z_{(0,\ldots,0,1)}^{(k)}, \ldots, z_{(n_1,\ldots,n_d)}^{(k)}) \in \{1\} \times \mathbb{R}^{\tilde{n}-1}$, $\tilde{n} - 1$ of the Inequalities (11) become equalities while the remaining inequality is a strict inequality.

Let $k \leq n$ be given. The equation $z_0^{(k)} = 1$ trivially holds.

Next we consider three cases for indices ℓ, $0 \leq \ell \leq n$. Case 1 is to assume the multi-index to satisfy $l_i > k_i$ for at least one $i \in \{1, \ldots, d\}$. By Definition (10) we know that for such indices $z_l^{(k)} = 0$. Since the sum on the left hand side of Inequality (11) only includes terms containing the factors $z_{l+j}^{(k)}$, $0 \leq j \leq n - l$, we know that the left hand side of Inequality (11) actually equals the right hand side.

If $l = k$, our case 2, the left hand side of (11) reduces to

$$\sum_{j=0}^{n-k}(-\mathbf{1})^j \binom{n-k}{j} z_{k+j}^{(k)} = (-\mathbf{1})^0 \binom{n-k}{0} z_k^{(k)} = \binom{n}{k}^{-1},$$

which is obviously a positive number.

Finally it remains to consider the last case, $l \leq k$, $l \neq k$. It follows by the first case that

$$\sum_{j=0}^{n-l}(-\mathbf{1})^j \binom{n-l}{j} z_{l+j}^{(k)} = \sum_{j=0}^{k-l}(-\mathbf{1})^j \binom{n-l}{j} z_{l+j}^{(k)}.$$

By Definition (10) we obtain

$$\sum_{j=0}^{k-l}(-\mathbf{1})^j \binom{n-l}{j} z_{l+j}^{(k)} = \binom{n}{k}^{-1} \sum_{j=0}^{k-l}(-\mathbf{1})^j \binom{n-l}{j}\binom{n-l-j}{k-l-j}.$$

For one-dimensional binomial coefficients we know that

$$\binom{r}{s}\binom{r-s}{t-s} = \binom{t}{s}\binom{r}{t};$$

this identity easily extends to the multi-dimensional case by forming products on both sides. Hence, since $l \neq k$, we also obtain

$$
\begin{aligned}
\sum_{j=0}^{k-l}(-\mathbf{1})^j \binom{n-l}{j} z_{l+j}^{(k)} &= \binom{n}{k}^{-1} \binom{n-l}{k-l} \sum_{j=0}^{k-l}(-\mathbf{1})^j \binom{k-l}{j} \\
&= \binom{n}{k}^{-1} \binom{n-l}{k-l} (\mathbf{1}-\mathbf{1})^{k-l} = 0. \qquad \square
\end{aligned}
$$

Proposition 2.5 *Let n and k be given multi-indices satisfying $k \leq n$ component-wise. The transformation $R^{(n)}$ applied to the corner points $z^{(k)}$ of the Hausdorff polytope $\mathcal{H}_n^d$ yields the set of unit vectors, i. e. $j = (j_1, \ldots, j_d) \leq n$,*

$$
R^{(n)}(z^{(k)})_j = \begin{cases} 1 & \text{for} \quad j = k \\ 0 & \text{otherwise.} \end{cases}
$$

PROOF. The Hausdorff polytope $\mathcal{H}_n^d$ is characterized by $\tilde{n}$ inequality conditions. For each corner point $z^{(k)}$ exactly $\tilde{n}-1$ of these conditions are active. We know, cf. proof of Theorem 2.4, that $\sum_{j=0}^{n} R^{(n)}(z^{(k)})_j = 1$ and that for $\tilde{n}-1$ different multi-indices $j \leq n$

$$
(\triangle^{n-j} z^{(k)})_j = 0.
$$

Thus

$$
R^{(n)}(z^{(k)})_j = \binom{n}{j}(\triangle^{n-j} z^{(k)})_j = 0
$$

for all but one multi-index $m \leq n$. For this particular index m the equation

$$
R^{(n)}(z^{(k)})_m = 1
$$

holds. But

$$
\begin{aligned}
R^{(n)}(z^{(k)})_k &= \binom{n}{k}(\triangle^{n-k} z^{(k)})_k \\
&= \binom{n}{k} \sum_{s=0}^{n-k}(-\mathbf{1})^s \binom{n-k}{s} z_{k+s}^{(k)} \\
&= \binom{n}{k}\binom{n-k}{0}\binom{n}{k}^{-1}\binom{n-k}{0} \\
&= 1,
\end{aligned}
$$

since, by definition of $z^{(k)}$, all other terms of this sum disappear. Hence, $m = k$ and $R^{(n)}(z^{(k)})$ is a unit vector. $\qquad \square$

So we can think of the image of $R^{(n)}$ of the corner point $z^{(k)}$ as being the Dirac measure at the point $(\frac{k_1}{n_1}, \ldots, \frac{k_d}{n_d})$, and we can think of the range $R^{(n)}(\mathcal{H}_n^d)$ as the set of all discrete probability measures on the set of points $\{(\frac{j}{n})\}_{j=0}^n$.

The following abstract convergence result together with Theorem 3.1 is the fundamental analytical underpinning of the numerical computations in Section 3. Note that each Hausdorff polytope $\mathcal{H}_n^d$ can be naturally embedded into the unit ball B_1^d of ℓ_∞^d, the space of all bounded infinite sequences on $\mathbb{N}_0^d$ equipped with the sup-norm. We simply extend each (finite) sequence of $\mathcal{H}_n^d$ to an infinite one by assigning zero values to all coordinates greater than n. With a view towards the assumption of the theorem which follows we recall that every sequence in B_1^d has at least one weak-$*$-convergent subsequence since ℓ_∞^d is a separable Banach space and B_1^d is weak-$*$-compact, i.e. compact with respect to the $\sigma(\ell_1^d, \ell_\infty^d)$-topology.

Theorem 2.6 *Let* $\left\{\mu^{(n)}\right\}_{n \geq 0}$ *be a sequence of elements with* $\mu^{(n)} \in \mathcal{H}_n^d$, $n \in \mathbb{N}_0^d$. *Then the limit* μ *of any weak-$*$-convergent subsequence* $\left\{\mu^{(n_s)}\right\}_{s \in \mathbb{N}}$, $n_s \to \infty$, *is a moment sequence of a measure defined on* $[0, 1]^d$.

PROOF. Let $\left\{\mu^{(n_s)}\right\}_{s \in \mathbb{N}}$ converge to μ in the weak-$*$-sense. Each Hausdorff condition, i.e. $(\Delta^m \mu)_k \geq 0$, $k, m \in \mathbb{Z}_+^d$, can be equivalently expressed in terms of the value of the natural pairing of μ with a finite sequence h, when h is considered to be an element of ℓ_1^d, cf. remark below. Thus, the assumption $\langle h, \mu^{(n_s)} \rangle \geq 0$, for s large enough, together with the convergence of $\left\{\mu^{(n_s)}\right\}_{s \in \mathbb{N}}$ imply that μ satisfies all Hausdorff conditions. By Theorem 2.1 μ is a moment sequence. $\qquad \square$

Remark. The representation of $(\Delta^m \mu)_k$ as the scalar product of $\mu \in \ell_\infty^d$ and an element $h^{(k,m)} \in \ell_1^d$ is given as follows:

For $k, m \in \mathbb{Z}_+^d$ define

$$
h_i^{(k,m)} = \begin{cases} (-\mathbf{1})^s \binom{m}{s} & \text{if } i = k + s, \ 0 \le s \le m \\ \\ 0 & \text{else}. \end{cases}
$$

Then, for $\mu \in B_1^d$,

$$
\langle h^{(k,m)}, \mu \rangle = \sum_{i=0}^{\infty} h_i^{(k,m)} \mu_i = \sum_{s=0}^{m} (-\mathbf{1})^s \binom{m}{s} \mu_{k+s} = (\Delta^m \mu)_k,
$$

cf. Theorem 2.1.

A specific example of the convergence result Theorem 2.6 is provided by the sequence of corner points $z^{(k;n)}$, where $k, n \in \mathbb{Z}_+^d$ such that $\frac{k}{n} \to t \in [0,1]^d$. For this example and for more general ones it is shown in Decker (2006) that $z^{(k;n)}$ converges in the weak-$*$-sense to δ_t, the Dirac measure at point t.

3 Numerical examples

The examples to be considered in this section are uncontrolled diffusion processes with smooth coefficients and which evolve on the bounded domain E^d up to the time when they hit the boundary. To make the exposition self-contained, we briefly recap the formulation of the LP-approach to exit time problems of such processes; see Kurtz and Stockbridge (1998) for an exposition of the general case of controlled and uncontrolled Markov processes evolving on general state spaces.

Let $X = \{X_t\}_{t \ge 0}$, $X_0 = x_0$, be a diffusion process on Ω_0, where Ω_0 denotes the interior of $\Omega = E^d$. Note, the case when the bounded domain is a general rectangular hypersolid $\prod_{i=1}^{d} [a_i, b_i]$ can be reduced to the set E^d by a linear coordinate transformation, i.e. $[a_i, b_i] \ni x_i \mapsto \frac{x_i - a_i}{b_i - a_i} \in [0,1]$. We take the generator A of the diffusion X to be $A = \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{d} b_j(x) \frac{\partial}{\partial x_j}$ and will restrict A to $\mathcal{C}_0^2(\mathbb{R})$, the set of twice continuously differentiable functions with compact support. We assume the coefficients $a_{ij}(x)$

and $b_j(x)$ of A to be such that the martingale problem for A and any starting point x_0 is well posed on $\mathbb{R}^d$, cf. Ethier and Kurtz (1986) Chapters 4 and 8. Note that these assumptions are satisfied for all the examples to be considered below and that for every test function f which we shall consider $Af_{|\Omega_0}$ will be bounded. Moreover, whenever we choose a monomial as test function f we assume the monomial to be modified outside of $\mathbf{E}^d$ in such a way that the modification belongs to $\mathcal{C}_0^2$. We shall denote the boundary of Ω by Ω_1. Let τ denote the first time X hits Ω_1. Since we are interested in the distribution of τ we shall explain how to compute upper and lower bounds for the moments of the random variable τ. Since the formulation of the LP-approach for the mean exit time is simpler than the formulation for higher moments we begin with the simpler case; afterwards we extend the formulation to higher moments.

The basic fact which underlies the LP-approach is that for each $f \in \mathcal{C}_0^2(\mathbb{R}^d)$,

$$f(X_t) - f(X_0) - \int\limits_0^t Af(X_s)\,ds$$

is a martingale and, if τ has finite expectation, the martingale property implies the equation

$$(12) \qquad \mathbf{E}[f(X_\tau)] - \mathbf{E}[f(X_0)] - \mathbf{E}\left[\int\limits_0^\tau Af(X_s)ds\right] = 0.$$

Note, the assumptions made actually imply that Equation (12) holds for every $f \in \mathcal{C}^2$, cf. Øksendal (2003) Section 7.4.

Define the expected occupation measure μ_0 and the exit distribution μ_1 by

$$\mu_0(B) = \mathbf{E}\left[\int\limits_0^\tau I_B(X_s)\,ds\right] \qquad \text{and} \qquad \mu_1(B) = P(X_\tau \in B)$$

for Borel sets B. It then follows, since $X_0 = x_0$, that Equation (12) can be written as

$$(13) \quad \langle f, \mu_1 \rangle := \int\limits_{\Omega_1} f(x)\,\mu_1(dx) = f(x_0) + \int\limits_{\Omega_0} Af(x)\,\mu_0(dx) =: f(x_0) + \langle Af, \mu_0 \rangle.$$

Furthermore, we assume that for each starting point in Ω_0 the first exit time from Ω_0 has finite expectation.

We formally define the adjoint operator A^*, $\mu \mapsto A^*\mu$, by

$$\langle f, A^*\mu \rangle := \langle Af, \mu \rangle,$$

for all $f \in \mathcal{C}_0^2$, and use the shorthand writing

(14) $$\mu_1 - \delta_{x_0} - A^*\mu_0 = 0,$$

where δ_{x_0} denotes the Dirac measure at x_0, to express Equation (13). The results in Kurtz and Stockbridge (1998) imply that for each μ_0 and μ_1 which satisfy Equation (14) – to be understood in the sense of (13) – there is a process X and an exit time τ for which Equation (12) holds. Thus Equation (14) characterizes the occupation measure μ_0 and the exit distribution μ_1 of a Markov process starting at x_0 and having generator A. Assuming that the constant function $\mathbb{1}(x) \equiv 1$ is an element of the range of A, Equation (14) implies that the mean exit time can be described as the value of two infinite dimensional linear programs which, since the corresponding martingale problem is well posed, have but *one* feasible solution (μ_0, μ_1), viz.

$$\mathbf{E}_{x_0}[\tau] = \inf_{\mu_0, \mu_1 \geq 0} \left\{ \int_{\Omega_0} \mu_0(dx) \ \middle| \ \mu_1(\Omega_1) = 1 \text{ and Equation (13) holds for all } f \in \mathcal{C}_0^2 \right\}$$

$$= \sup_{\mu_0, \mu_1 \geq 0} \left\{ \int_{\Omega_0} \mu_0(dx) \ \middle| \ \mu_1(\Omega_1) = 1 \text{ and Equation (13) holds for all } f \in \mathcal{C}_0^2 \right\}.$$

More generally, if R, l resp., are bounded measurable functions on Ω_1, Ω_0 resp., using the shorthand writing (14), we may write

$$\mathbf{E}_{x_0} \left[R(X_\tau) + \int_0^\tau l(X_s)\, ds \right] = \inf_{\mu_0, \mu_1 \geq 0} \{ \langle R, \mu_1 \rangle + \langle l, \mu_0 \rangle \mid \langle \mathbb{1}, \mu_1 \rangle = 1, \ \mu_1 - A^*\mu_0 = \delta_{x_0} \}$$

(15) $$= \sup_{\mu_0, \mu_1 \geq 0} \{ \langle R, \mu_1 \rangle + \langle l, \mu_0 \rangle \mid \langle \mathbb{1}, \mu_1 \rangle = 1, \ \mu_1 - A^*\mu_0 = \delta_{x_0} \}.$$

It needs to be stressed that for exit time problems of *well posed* martingale problems the linear programs (15) are actually an artifact since the constraints uniquely determine the feasible pair (μ_0, μ_1). While this is true for the infinite dimensional problems (15) this is no longer the case for the finite dimensional approximating LPs, see below.

Next, if R and l are polynomials, i.e.

$$R(x) \; = \; \sum_{j=0}^{n_R} r_j x^j, \qquad l(y) \; = \; \sum_{j=0}^{n_l} \lambda_j y^j$$

for $x \in \Omega_1$, $y \in \Omega_0$ and multi-indices n_R and n_l, and A maps monomials onto polynomials, e.g. for $f(x) = x^k$, k a multi-index, there exist n_k and coefficients c_{k_j}, $0 \le j \le n_k$, such that

$$(16) \qquad Af(x) \; = \; \sum_{j=0}^{n_k} c_{k_j} x^j,$$

then the infinite dimensional LPs (15) are equivalent to infinite dimensional LPs whose variables are moment sequences.

Using but finite sequences and a finite number of the characterizing inequalities as constraints will determine finite dimensional linear programs whose values provide upper and lower bounds on the expected value which is to be computed. For instance, assuming R and l to be polynomial functions restricted to $[0,1]^d$, A the generator of a diffusion with polynomial coefficients, $\Omega_0 = (0,1)^d$, $\Omega_1 = [0,1]^d \backslash (0,1)^d$ and choosing the Hausdorff conditions up to a given order n the following finite dimensional LPs provide bounds on the quantity of interest, i.e.

$$(17) \qquad \underline{\varphi}_n \; \le \; \mathbf{E}_{x_0}\left[R(X_\tau) + \int_0^\tau l(X_s)\,ds \right] \; \le \; \bar{\varphi}_n,$$

where

$$(18) \qquad \bar{\varphi}_n \; := \; \max_{z^{(0)},\, z^{(1)}} \left\{ \sum_{j=0}^{n_R} r_j z_j^{(1)} + \sum_{j=0}^{n_l} \lambda_j z_j^{(0)} \,\middle|\, z^{(0)} \in \tilde{\mathcal{H}}_n^d, \; z^{(1)} \in \mathcal{K}_n^d, \right.$$
$$\left. \text{and for all } k \text{ such that } n_k \le n : z_k^{(1)} - \sum_{j=0}^{n_k} c_{k_j} z_j^{(0)} = x_0^k \right\}$$

and

$$(19) \qquad \underline{\varphi}_n \; := \; \min_{z^{(0)},\, z^{(1)}} \left\{ \sum_{j=0}^{n_R} r_j z_j^{(1)} + \sum_{j=0}^{n_l} \lambda_j z_j^{(0)} \,\middle|\, z^{(0)} \in \tilde{\mathcal{H}}_n^d, \; z^{(1)} \in \mathcal{K}_n^d, \right.$$
$$\left. \text{and for all } k \text{ such that } n_k \le n : z_k^{(1)} - \sum_{j=0}^{n_k} c_{k_j} z_j^{(0)} = x_0^k \right\}.$$

Note, each vector $z^{(1)} \in \mathcal{K}_n^d$ is the convex combination of scaled elements of Hausdorff polytopes associated with $(d-1)$-dimensional hypercubes. Since for exit time problems the total mass of the corresponding occupation measure is typically different from one, the vectors $z^{(0)}$ are also scaled versions of elements – this time – of d-dimensional Hausdorff polytopes. For the two-dimensional example below, see Example 3.2, our LP-code incorporates the corner point formulas for $\mathcal{H}_n^2$ only once, while we use the one-dimensional formulas four times (for each side of the unit square). Thus, using the cornerpoint representation of the Hausdorff polytopes instead of the set of characterizing Inequalities (7) dramatically decreases the size of the linear programs (18) and (19).

The following result shows that under natural assumptions the values $\underline{\varphi}_n$ and $\bar{\varphi}_n$ converge to the value of the corresponding infinite dimensional linear program, see Equation (15). By construction the bounded sequence $\{\underline{\varphi}_n\}_n$, $\{\bar{\varphi}_n\}_n$ resp., is monotone increasing, decreasing resp., and therefore both sequences naturally have a limit.

Theorem 3.1 *Assume Ω_0 to be the interior of E^d and let Ω_1 denote the boundary of Ω_0. Let A be the generator of a diffusion process with polynomial coefficients, cf. Equation (16), such that the corresponding martingale problem has a unique solution on a ball which contains $\Omega = \Omega_0 \cup \Omega_1$. Assume that for every initial value of the diffusion the associated exit time from Ω_0 has finite expectation. Let $(\underline{\nu}_n, \underline{\mu}_n)$, $(\bar{\nu}_n, \bar{\mu}_n)$ resp., be a pair of optimal programs of the minimization problem (19), maximization problem (18) resp., with corresponding value $\underline{\varphi}_n$, $\bar{\varphi}_n$ resp. Then*

$$\lim_{n\to\infty} \underline{\varphi}_n = \lim_{n\to\infty} \bar{\varphi}_n,$$

and the common limit is the value of the infinite dimensional linear program (15). Moreover, there are subsequences $\{\underline{\nu}_{n_k}, \underline{\mu}_{n_k}\}_k$ and $\{\bar{\nu}_{n_k}, \bar{\mu}_{n_k}\}_k$ which converge in the weak-$*$-sense to moment sequences $(\vec{\nu}, \vec{\mu})$; the pair of measures (ν^*, μ^*) which is associated with $(\vec{\nu}, \vec{\mu})$ is an optimal program of the LP (15).

PROOF. It will suffice to only consider, for instance, the maximization problem since the

same arguments apply to the minimizing LP. According to Theorem 2.6 the sequences $\{\bar{\nu}_n\}_n$ and $\{\bar{\mu}_n\}_n$ determine measures ν and μ such that for all polynomial functions p

(20) $$\langle p, \nu \rangle - \langle p, \delta_{x_o} \rangle - \langle Ap, \mu \rangle = 0.$$

Since each function f on $\mathbb{R}^d$ which, along with its derivatives up to the K-th order, $K \in \mathbb{N}_0$, is continuous on Ω, may be uniformly approximated by polynomials in such a way that the derivatives of f up to the K-th order are also approximated uniformly by the corresponding derivatives of the polynomials, see Courant and Hilbert (1937), Chapter II.4, and since A has polynomial coefficients Equation (20) holds for all $f \in \mathcal{C}_0^2$. Since our assumptions imply that there is a unique pair of measures (ν, μ) which satisfy Equation (15) and, since

$$\bar{\varphi}_n = \sum_{j=0}^{n_R} r_j \bar{\nu}_n^{(j)} + \sum_{j=0}^{n_\ell} \lambda_j \bar{\mu}_n^{(j)} \quad \rightarrow \quad \langle R, \nu \rangle + \langle \ell, \mu \rangle$$

if $n \to \infty$, the assertions follow. $\qquad \square$

The numerical examples below will demonstrate that for reasonably large values of n the difference between $\bar{\varphi}_n$ and $\underline{\varphi}_n$ will typically be small. In earlier publications, e.g. Helmes, Röhl, and Stockbridge (2001), we have implemented the requirements $z^{(0)} \in \tilde{\mathcal{H}}_n^d$ and $z^{(1)} \in \mathcal{K}_n^d$ using Inequalities (7) or (8) instead of the corner point formulas; actually, to increase numerical stability we usually used the recursive definition of iterated differences, i.e. (3), and (7). Such implementations result in large programs, whose size and run times restrict the values of n. For typical two-dimensional problems, e. g., the computation of the mean exit time of a two-dimensional Brownian motion from a square, cf. Example 3.2 below, we took $n = (M, M)$, $M \leq 14$. But larger values of M provide better bounds on $\mathbf{E}[\tau]$. Larger values of M become a possibility when the corner point formula, cf. Section 2, is used. The following examples will show that by using the corner point formulas and larger values of M the resulting LPs not only provide better bounds than the old programs do, but also require fewer iterations.

Finally, to compute bounds for higher moments of the exit times of such processes we need

to augment the time coordinate to the state and consider the time-space generator $\tilde{A}$,

$$\tilde{A}f(t,x) = \frac{\partial f}{\partial t}(t,x) + Af(t,x),$$

acting on functions f depending on t and x. The measures μ_0 and μ_1 are now defined by

$$\mu_0(\Gamma) = \mathbf{E}_{x_0}\left[\int_0^\tau I_\Gamma(s, X_s)ds\right] \quad \text{and} \quad \mu_1(\Gamma) = P_{x_0}\left((\tau, X_\tau) \in \Gamma\right)$$

for subsets Γ of $[0, \infty) \times \Omega$. The extension of the fundamental Equation (13) takes the form

$$(21) \quad \int_{\mathbb{R}^+ \times \Omega_1} f(s,x)\mu_0(ds \times dx) - f(t_0, x_0) - \int_{\mathbb{R}^+ \times \Omega_0}\left[\frac{\partial f}{\partial t}(s, X_s) + Af(s, X_s)\right] = 0$$

for all functions f such that $f(t, \cdot) \in \mathcal{C}_0^2$, and $f(\cdot, x)$ is continuously differentiable in t and vanishes at ∞. Note that the n-th moment of the exit time is given by

$$\mathbf{E}_{x_0}[\tau^n] = n \int_{\mathbb{R}^+ \times \Omega_0} s^{n-1}\mu_0(ds \times dx).$$

It follows from the general theory (see Kurtz and Stockbridge (1998)) that, as in the case of the mean exit time, measures μ_0 and μ_1 which satisfy (21) characterize processes having generator A up to the exit time τ. But while the process $\{X_t\}_{t \geq 0}$ evolves on a bounded set the process $\{t, X_t\}_{t \geq 0}$ does not. To mimic the LP-approach for X, i.e. exploit the characterization of μ_0 and μ_1 by their moments, we truncate the unbounded time domain $\mathbb{R}^+$ to a finite interval $[0, T]$. This introduces an additional approximation into the method in all those cases where $\tau \neq \tau \wedge T$. Rather than the exit time τ, the analysis will evaluate the distribution of $\tau \wedge T$. But, by the assumption that $\mathbf{E}[\tau]$ is finite, for large values of T the difference between the exit time and the truncated exit time will be zero with high probability. Moreover, there are many problems for which there is a natural bound T. In at least those cases an analogue of the convergence result 3.1 holds. For the general case of unbounded domains the convergence question is still under investigation.

Example 3.2 *Two-dimensional Brownian motion on the unit square:*

Let $(X_s, Y_s) = \left(x_0 + W_s^{(1)}, y_0 + W_s^{(2)} \right)$, where $x_0, y_0 \in (0, 1)$ and $W = \left(W_s^{(1)}, W_s^{(2)} \right)$ is a two-dimensional standard Brownian motion process. The generator of the process (X_s, Y_s) is given for $f \in \mathcal{C}_o^2$ by

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

In this example we consider the exit time of the process from the bounded domain $\Omega_0 = (0, 1) \times (0, 1)$; the boundary consists of four parts $\Gamma_{(b)} = [0, 1] \times \{0\}$, $\Gamma_{(t)} = [0, 1] \times \{1\}$, $\Gamma_{(\ell)} = \{0\} \times (0, 1)$ and $\Gamma_{(r)} = \{1\} \times (0, 1)$. Therefore, we need to work with five measures: the occupation measure μ_0 of the process in the open unit square, and measures on the four boundary parts which together comprise the exit distribution μ_1 of the Brownian motion process from the unit square. We associate with each boundary measure elements of the one-dimensional Hausdorff polytope of order M, while the occupation measure is associated with elements of the two-dimensional Hausdorff Polygon, i.e. $\tilde{\mathcal{H}}^2_{(M,M)}$.

Using the corner point formula (10) the constraints of the associated finite dimensional LP-problems can be expressed as follows. Let $z = \{z_{ij}\}_{i,j=0}^{M} \in \tilde{\mathcal{H}}^2_{(M,M)}$ be associated with the occupation measure μ_0 on Ω_0 and let $z^{(b)} = \{z_i^{(b)}\}_{i=0}^{M}$, $z^{(t)}$, $z^{(\ell)}$, $z^{(r)}$, being associated with the exit distribution on $\Gamma_{(b)}$, $\Gamma_{(t)}$, $\Gamma_{(\ell)}$ and $\Gamma_{(r)}$, be elements of $\mathcal{H}_M^1$. Then for $0 \leq m, n \leq M$,

$$
\begin{aligned}
0 \;=\; & \frac{m(m-1)}{2} z_{m-2,n} + \frac{n(n-1)}{2} z_{m,n-2} \\
& + x_0^m y_0^n - 0^n z_m^{(b)} - 1^n z_m^{(t)} - 0^m z_n^{(\ell)} - 1^m z_n^{(r)},
\end{aligned}
$$

and all the z-vectors are convex combinations of the corner points described in Section 2. Note, that we define $0^0 = 1$.

Table 1 reports our results on the mean exit time for a sample of initial values. Because of the symmetry of the problem we only report the values for initial values in the first quarter of the unit square. The numbers in rows labeled LP1 are those which we reported in Helmes, Röhl, and Stockbridge (2001). There we used the Hausdorff conditions expressed

in terms of the Inequalities (6). This restricted our choice of M; typically, M was chosen between 11 and 14 resulting in LPs which required approximately 13000 simplex iterations and ~ 25 minutes on an Ultra Sun 10/300 using AMPL/CPLEX 6.5. In rows labeled LP2 we give the numbers when using the same software version but employing the corner point formulas. They allow us to choose $M \sim 36$ which increases the accuracy of the values while actually requiring fewer iterations, ~ 10000 simplex iterations, and less time on the same machine. In rows labeled LP3 we report the values when using the corner point formulas and CPLEX 9.1. The numbers in the column "exact value" are based on the formula for the mean exit time of Brownian motion from a square, cf. Knight (1981) or Helmes, Röhl, and Stockbridge (2001). In Table 1 we also include the average of the two bounds since this value provides a good estimate for the quantity of interest. Moreover the last two columns show

(i) how much more accurate the modified LP-model is compared to the original one, and

(ii) that the modified LP-model and the SDP-method guarantee roughly the same accuracy. The entries in both columns are the quotients of the differences of the bounds when computed either by LP1 and LP2, or by LP3 and Lasserre's SDP-method. Numbers larger than one, smaller than one, resp., indicate that the method mentioned first is worse, better resp., than the other one.

Example 3.3 *Higher moments for the one-dimensional Brownian motion on [0,1]:*

Let $Z = \{t, X_t\}_{t \geq 0}$, where $\{X_t\}_{t \geq 0}$ denotes one-dimensional Brownian motion starting at $x_0 = 0.8$. We consider the process X on $[0, 1]$ up to the time $T = 10$; a truncation level larger than 10 does not provide better numerical results. Let τ denote the first time when $X_\tau = 0$ or $X_\tau = 1$ or $X_\tau = T$. The generator A of Z is given by

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

21

y_0		lower bound	upper bound	mean	exact value	ratio LP1/LP2	ratio LP3/SDP
0.5	LP1	0.146933	0.148017	0.147475			
	LP2	0.147290	0.147455	0.147372		6.6	
	LP3	0.147308	0.147408	0.147358	0.147340		0.94
0.4	LP1	0.141899	0.142991	0.142445			
	LP2	0.142244	0.142404	0.142324		6.8	
	LP3	0.142268	0.142373	0.142321	0.142310		1.00
0.3	LP1	0.126360	0.127415	0.126887			
	LP2	0.126696	0.126886	0.126791		5.6	
	LP3	0.126717	0.126826	0.126771	0.126760		1.06
0.2	LP1	0.099049	0.100020	0.099534			
	LP2	0.099343	0.099505	0.099424		6.0	
	LP3	0.099365	0.099457	0.099411	0.099396		0.99
0.1	LP1	0.057843	0.058443	0.058143			
	LP2	0.058054	0.058141	0.058097		6.9	
	LP3	0.058064	0.058124	0.058094	0.058084		0.95

Table 1: Bounds on the mean exit time of two-dimensional Brownian motion from the unit square as a function of y_0 based on different LP formulations; $x_0 = 0.5$, $M = 36$.

The dynamics of the process Z (up to the time of stopping) and the geometry of the state space determine the following equality constraints, where again $\tilde{z} \in \tilde{\mathcal{H}}^2_{(M,M)}$ and $z^{(b)}$, $z^{(t)}$ and $z^{(r)} \in \tilde{\mathcal{H}}^1_M$:

$$0 = \frac{m}{T}\tilde{z}_{m-1,n} + \frac{1}{2}n(n-1)\tilde{z}_{m,n-2} + 0^m x_0^n - 0^n z_m^{(b)} - z_m^{(t)} - z_n^{(r)},$$

for $0 \leq m, n \leq M$.

Table 2 illustrates the dependence of the accuracy of the bounds on higher moments of τ on the parameter M. We report bounds for moments up to order 5 for $M = 10$ and

$M = 20$ using the corner point formulas and running CPLEX 9.1.

moments of the exit time	M	lower bound	upper bound	mean	exact value
1st	10	0.160000	0.160000	0.160000	
	20	0.160000	0.160000	0.160000	0.160000
2nd	10	0.061865	0.061867	0.061866	
	20	0.061867	0.061867	0.061867	0.061867
3rd	10	0.037370	0.037393	0.037382	
	20	0.037393	0.037393	0.037393	0.037393
4th	10	0.029974	0.030290	0.030132	
	20	0.030290	0.030290	0.030290	0.030290
5th	10	0.026649	0.030688	0.028668	
	20	0.030688	0.030688	0.030688	0.030688

Table 2: Higher moments of the exit time of time-space Brownian motion from the rectangle $[0, T] \times [0, 1]$ for $T = 10$; $M = 10$ and $M = 20$; $(t_0, x_0) = (0, 0.8)$.

References

Ang, D. D., R. Gorenflo, V. K. Le, and D. D. Trong (2002). *Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction*, Volume 1792 of *Lecture Notes in Math.* Springer–Verlag, Berlin.

Bhatt, A. G. and V. S. Borkar (1996). Occupation measures for controlled Markov processes: Characterisation and optimality. *Ann. Prob. 24*(3), 1531–1562.

Cho, M. J. (2000). *Linear Programming Formulation for Optimal Stopping.* Ph. D. thesis, The Graduate School, University of Kentucky, Lexington.

Cho, M. J. and R. H. Stockbridge (2002). Linear programming formulation for optimal stopping problems. *SIAM J. Control Optim. 40*, 1965–1982.

Courant, R. and D. Hilbert (1937). *Methods of Mathematical Physics, Vol. 1*. Wiley Classics Edition (1989). New York: John Wiley & Sons.

Dale, A. I. (1987). Two-dimensional moment problems. *Math. Sci. 12*, 21–29.

Decker, T. (2006). *Die Eckpunkte des allgemeinen Dale-Polytopes und ihre Anwendungen in linearen Programmen zur Bestimmung von Austrittszeiten*. Ph. D. thesis, Humboldt-Universität zu Berlin, Berlin, Germany.

Ethier, S. N. and T. G. Kurtz (1986). *Markov Processes. Characterization and Convergence*. New York: John Wiley & Sons.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications* (2nd ed.), Volume 2. New York: John Wiley & Sons.

Hausdorff, F. (1921). Summationsmethoden und Momentenfolgen. I. & II. *Math. Z. 9*, 74–109, 280–299.

Hausdorff, F. (1923). Momentenprobleme für ein endliches Intervall. *Math. Z. 16*, 220–248.

Helmes, K. (1999). Computing moments of the exit distribution for diffusion processes using linear programming. In P. Kall and J.-H. Lüthi (Eds.), *Oper. Res. Proc. 1998*, Berlin, pp. 231–240. Springer–Verlag.

Helmes, K. (2002). Numerical methods for optimal stopping using linear and non-linear programming. In B. Pasik-Duncan (Ed.), *Proceedings of a Workshop "Stochastic Theory and Control"*, Lecture Notes in Control and Information Sciences, Berlin, pp. 185–202. Springer–Verlag.

Helmes, K., S. Röhl, and R. H. Stockbridge (2001). Computing moments of the exit time distribution for Markov processes by linear programming. *Oper. Res. 49*, 516–530.

Helmes, K. and R. H. Stockbridge (2000). Numerical comparison of controls and verification of optimality for stochastic control problems. *J. Optim. Theory Appl. 106*,

107–127.

Helmes, K. and R. H. Stockbridge (2001). Numerical evaluation of resolvents and Laplace transforms of Markov processes. *Math. Methods Oper. Res. 53*, 309–331.

Helmes, K. and R. H. Stockbridge (2003). Extension of Dale's moment conditions with application to the Wright-Fisher model. *Stoch. Models 19*(2), 255–267.

Hernandez-Lerma, O., J. C. Hennet, and J. B. Lasserre (1991). Average cost Markov decision processes: Optimality conditions. *J. Math. Anal. Appl. 158*, 396–406.

Hildebrandt, T. H. and I. J. Schoenberg (1933). On linear functional operations and the moment problem for a finite interval in one or several dimensions. *Ann. of Math. 34*, 317–328.

Karlin, S. and L. S. Shapley (1953). Geometry of moment spaces. *Mem. Amer. Math. Soc. 12*, 1–91.

Knight, F. B. (1981). *Essentials of Brownian Motion and Diffusion*. Mathematical Surveys No. 18. Providence, RI: American Mathematical Society.

Knill, O. (1997). On Hausdorff's moment problem in higher dimensions. Preprint, `http://abel.math.harvard.edu/~knill/preprints/stability.ps` viewed 04/02/2007.

Kurtz, T. G. and R. H. Stockbridge (1998). Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim. 36*, 609–653.

Kurtz, T. G. and R. H. Stockbridge (1999). Martingale problems and linear programs for singular control. In *Thirty-Seventh Annual Allerton Conference on Communication, Control, and Computing (Monticello, Ill.)*, Urbana-Champaign, Ill., pp. 11–20. Univ. Illinois.

Lasserre, J. B. and T. P. Rumeau (2004). SDP vs. LP relaxations for the moment approach in some performance evaluation problems. *Stoch. Models 20*, 439–456.

Manne, A. S. (1960). Linear programming and sequential decisions. *Management Sci. 6*, 259–267.

Mendiondo, M. S. and R. H. Stockbridge (1998). Approximation of infinite-dimensional linear programming problems which arise in stochastic control. *SIAM J. Control Optim. 36*, 1448–1472.

Øksendal, B. (2003). *Stochastic Differential Equations* (6th ed.). Springer-Verlag.

Riechert, S. (2004). Ein Vergleich der SDP- mit der LP-Relaxation für die Berechnung optimaler Stoppzeiten. Master's thesis, Humboldt-Universität zu Berlin, Berlin, Germany.

Röhl, S. (2001). *Ein linearer Programmierungsansatz zur Lösung von Stopp- und Steuerungsproblemen.* Ph. D. thesis, Humboldt-Universität zu Berlin, Berlin, Germany.

Schwerer, E. (2001). A linear programming approach to the steady-state analysis of reflected Brownian motion. *Stoch. Models 17*, 341–368.

Shohat, J. A. and J. D. Tamarkin (1943). *The Problem of Moments* (1st ed.). Providence, Rhode Island: American Mathematical Society.

Stockbridge, R. H. (1990). Time-average control of martingale problems: A linear programming formulation. *Ann. Probab. 18*, 206–217.

Stockbridge, R. H. (2003). The problem of moments on a polytope and other bounded regions. *J. Math. Anal. Appl. 285*, 356–375.

Ziegler, G. M. (1995). *Lectures on Polytopes*, Volume 152 of *Graduade Texts in Mathematics.* New York: Springer–Verlag.