Computing Moments of the Exit Time Distribution for Markov Processes by Linear Programming *

Kurt Helmes

Institute of Operations Research, Humboldt University of Berlin,

Berlin, Germany

Stefan Röhl

Institute of Operations Research, Humboldt University of Berlin,

Berlin, Germany

Richard H. Stockbridge

Department of Statistics, University of Kentucky,

Lexington, Kentucky

February 23, 1999

*This research is partially supported by NSF under grant DMS 9803490.

Abstract

We provide a new approach to the numerical computation of moments of the exit time distribution of Markov processes. The method relies on a linear programming formulation of a process exiting from a bounded domain. The LP formulation characterizes the evolution of the process through the moments of the induced occupation measure and naturally provides upper and lower bounds for the exact values of the moments. The conditions the moments have to satisfy are derived directly from the generator of the Markov process and are not based on some approximation of the process. Excellent software is readily available since the computations involve finite dimensional linear programs.

The purpose of this paper is to provide a numerical method for computing the moments of the exit time distribution of Markov processes based on linear programming. This numerical approach has the advantage over other approaches that it provides upper bounds as well as lower bounds for the quantities of interest. Furthermore, excellent software is readily available.

We use a linear programming formulation for the evolution of Markov processes which involves solving a system of equations over measures on the states of the process. We characterize the set of feasible measures for the LP through linear conditions which their moments must satisfy. For numerical computation, the problem is reduced to finite dimensions by considering only a finite number of moments. An important aspect of this approach is that we are directly analyzing the stochastic processes and not using some approximating processes or some approximation of the generator of the process. To obtain higher moments of the exit time, however, we employ a truncation technique on the time dimension.

In this paper we consider the special case of exit problems for Markov processes. The mean exit times and the higher moments of the exit time distribution can be expressed in terms of the moments of the occupation measure. Using these expressions as the objective function in our LP problems and maximizing as well as minimizing the same problem we naturally obtain upper and lower bounds for these quantities. We illustrate the power and accuracy of the LP method by comparing the numerical results for the first and higher moments of the exit time with analytical ones. Specifically, we analyze the exit times of a class of interest rate processes which are useful in mathematical finance, of Bessel processes, of processes for a quality control model, of a two-dimensional Brownian motion process and of special multidimensional time-space processes, e.g. the Poisson and the Brownian motion processes. These examples are chosen because of the variety of the characteristics of the models, their importance in applications and to show some techniques which are useful in solving these types of problems. The models include equations whose coefficients are fractional as well as negative powers of the state.

Formulating the evolution of the Markov process through occupation measures developed from work by [23] for problems in the calculus of variations. The present approach is a special case of the more general situation of formulating stochastic control problems as linear programs over a space of stationary distributions studied in very general settings in [2], [12], [19], and [20]. The variable in the infinite-dimensional linear program is a measure on the product of the state and control spaces, and in the case of first-exit problems, this variable is augmented by a second measure on the exterior of the state space (see [12]).

The processes in this paper are *not* controlled. However, the results cited above are applicable to these processes by taking the control space to consist of a single element. It then follows that the variables are measures solely on the states of the process.

The fundamental idea of our numerical method is to start with this characterization of an optimal stochastic control problem as an infinite-dimensional linear program, approximate this by a finite-dimensional linear program and then compute. The constraints are generated by test functions via the martingale characterization of the processes (see expression (7)) and the associated basic adjoint equation (9); see also [7] for the martingale characterization. The collection of test functions used in this paper typically are the monomials which implies that the basic adjoint equation can be expressed in terms of the moments of the measures. This reduces the analysis from the space of measures to the space of moment sequences of the measures. Additional constraints are added to ensure that the sequences, in fact, are the moments of the measures.

The linear programming approach to stochastic control at this level of generality had its foundations in the existence of stationary processes corresponding to the measure which satisfies the basic adjoint equation. This was established in [19] and later extended to the more general adjoint equation for first exit problems in [12]. [11] also applied these techniques in uncontrolled settings for patchwork and constrained martingale problems and, in joint work with Dai, used this in the analysis of Brownian networks ([4]). The basic adjoint equation corresponding to the heavy traffic limit of queueing processes had previously been introduced by [9]. It has been extensively used to determine the stationary distributions for Brownian networks (see, for example, [3], [15] and [16] and the references therein).

The paper by [16] is particularly relevant to the discussion in the present paper. She uses moments in a linear programming approach to analyze the steady-state distribution for reflected Brownian motion. The use of the basic adjoint equation is essentially the same as in this paper. A distinction occurs between the papers in the additional constraints imposed which ensures that the sequences are moment sequences. Schwerer works with moments over unbounded regions and thus the necessary and sufficient conditions are expressed in terms of nonnegativity of certain quadratic forms or equivalently, in terms of nonnegativity of a collection of determinants involving the moments. While these countably many conditions are non-linear in nature Schwerer exploits the fact that they are equivalent to a set of uncountably many linear conditions involving the moments of a measure, i.e. a non-negativity condition for every finite dimensional vector of any dimension. The same moment inequalities as in [16] were used by Dawson [5, 6] to characterize the measure-valued solutions to stochastic differential equations in his work on the evolution of geostochastic systems.

In our paper, the processes evolve in a bounded region in the state space which allows the necessary and sufficient conditions on the moments to be expressed as countably many linear conditions, viz. the Hausdorff moment conditions. When time is included as a dimension of the state space in order to compute the higher moments of the exit time, we truncate the unbounded time interval so as to again use the linear conditions on the moments. The linear programs for these problems are large and highly structured. This structure means that matrix generating languages like AMPL or GAMS are ideal as interfaces with the linear programming software. Since for exit time problems many analytical results are known, these large LP problems may be useful in comparing present day LP solvers and in testing new LP methods.

This paper is organized as follows. We motivate the linear programming approach by considering, in section 1, the example of a Poisson process. This example contains most of the key ideas in a simple form. Then in section 2, we state the general linear programming formulation in more general terms and, in particular, identify the basic adjoint equation (9) and the Hausdorff moment conditions (11) which are at the center of this approach. Two formulations are given depending on whether the goal is to determine simply the mean exit time or higher moments as well. Beginning with section 3, we illustrate the accuracy of the linear programming approach with some particular examples for which formulas for the exact values are known. In section 3, we analyze the constant coefficient case of the Cox-Ingersoll-Ross interest rate model and provide extensive numerical comparisons of the LP approach for the mean exit time with exact values. This section also introduces the use of a change of measure to improve numerical accuracy. Section 4 analyzes the mean exit time of Bessel processes for various dimensions. Much of the approach for the interest rate model applies directly, though modifications are required in the change of measure and the basic adjoint equation. In section 5 the mean exit time of the Roberts-Shiryaev change detection model is examined. The mean exit time of a two-dimensional Brownian motion process from the unit square is considered in section 6, where it is also necessary to extend the Hausdorff moment conditions to the multidimensional setting. Higher moments of the exit time distribution are analyzed in sections 7 and 8. We return to the simple example of the Poisson process to introduce the ideas of including time as a component of the process in section 7 and the truncation of the time to a finite interval. In section 8, we evaluate the higher moments of the exit time of a Brownian motion process from the unit interval. We conclude the paper with a discussion of some of our computational experiences.

1 Motivating example

We introduce the basic idea underlying the linear programming approach to the computation of moments of the exit time distribution by considering one of the simplest examples. Let $N = (N_t : t \geq 0)$ be a Poisson process with intensity $\lambda > 0$ and let τ denote the first time at which $N_t = K$, for a fixed positive integer K. Partition the states of the process $\{0, 1, 2, \ldots\}$ into a bounded region $E_0 = \{0, 1, \ldots, K-1\}$ and its complement and view the time τ as the first time N exits E_0. Note that when N exits, it does so by hitting the boundary state K.

We illustrate the methodology by computing the mean exit time $E[\tau]$. The answer, of course, follows trivially without any computation from the observation that the process stays an average of $1/\lambda$ time units at each state and hence $E[\tau] = K/\lambda$.

Important to the linear programming formulation is the generator A of N. This operator acts on functions $f : \{0, 1, 2, \ldots\} \to \mathbb{R}$ and is given by

$$Af(x) = \lambda(f(x+1) - f(x)).$$

Also of importance is the observation that for each function f,

$$f(N_t) - f(N_0) - \int_0^t Af(N_s)ds \tag{1}$$

is a martingale, which implies that for each $t > 0$,

$$E\left[f(N_{t\wedge\tau}) - f(N_0) - \int_0^{t\wedge\tau} Af(N_s)ds\right] = 0.$$

Since $\tau < \infty$ a.s., letting $t \to \infty$ gives

$$E[f(N_\tau)] - E[f(N_0)] - E\left[\int_0^\tau Af(N_s)ds\right] = 0. \tag{2}$$

Now define the occupation measure μ_0 on the states $E_0 = \{0, 1, 2, \ldots, K-1\}$ such that, for each $i = 0, 1, \ldots, K-1$, and with $I_{\{i\}}$ the indicator function of $\{i\}$,

$$\mu_0(i) = E\left[\int_0^\tau I_{\{i\}}(N_s)ds\right]. \tag{3}$$

Then since $N_\tau = K$ and $N_0 = 0$, (2) can be written (after multiplication by -1) as

$$\sum_{i=0}^{K-1} \lambda(f(i+1) - f(i))\mu_0(i) + f(0) - f(K) = 0 \tag{4}$$

for each function f. This equation is an example of the more general equation, to be considered below,

$$\int_{E_0} Af(x)\mu_0(dx) + f(0) - \int_{E_0^c} f(x)\mu_1(dx) = 0, \tag{5}$$

where μ_0 is the occupation measure and μ_1 gives the distribution of the exit location when the process N exits. In this particular example, $N_\tau = K$ so the exit distribution μ_1 is a unit point mass on state K.

Observe that setting $\mu_0(i) = 1/\lambda$ for each $i = 0, 1, \ldots, K-1$ reduces the left hand side of (4) to a simple telescoping sum for each function f. This identifies μ_0 as an unnormalised uniform distribution on E_0. It then follows from (3) that

$$E[\tau] = E\left[\int_0^\tau 1ds\right] = E\left[\int_0^\tau \sum_{i=0}^{K-1} I_{\{i\}}(N_s)ds\right] = \sum_{i=0}^{K-1} \mu_0(i) = \frac{K}{\lambda}. \tag{6}$$

Thus, once the measure μ_0 has been determined, the mean exit time can readily be obtained from the 0th-moment of the measure μ_0.

An alternate approach to determining μ_0 is to choose the collection of functions $\mathcal{D}_0 = \{f_i(x) = x^i : i = 0, 1, \ldots, K-1\}$ and solve the resulting system of equations (4). This particular choice of collection $\mathcal{D}_0$ identifies the first $K-1$ moments of the measure μ_0. The solution can be obtained by treating the system of equations as the constraints of a linear program for some objective function. The solution procedure will then determine the unique feasible measure μ_0. Note that only a finite number of test functions are necessary since E_0 contains only K states. The selection of $\mathcal{D}_0 = \{f_i(x) = x^i : i = 0, 1, \ldots, K\}$ is not the only choice for such a collection. In fact, any K functions $\{f_1, \ldots, f_K\}$ for which the vectors

$$(f_1(0), f_1(1), \ldots, f_1(K-1)), \ldots, (f_K(0), f_K(1), \ldots, f_K(K-1))$$

are linearly independent can be used successfully.

For more general processes, however, it is usually not possible to solve equation (5) for the measure μ_0. Instead we choose functions f that can be used to determine the important quantity of interest and which uniquely characterize the measure. In this paper, we choose variants of the monomials $f(x) = x^m$ so that the general set of equalities (5) can be expressed in terms of the moments of the measures μ_0 and μ_1. The mean exit time is then obtained using the general form of (6).

2 The linear programming formulation

We consider a general process $Y = (Y_t : t \geq 0)$, with $Y_0 = y_0$ and generator A, and let $\mathcal{D}(A)$ denote the domain of the generator. We assume that Y takes values in $\mathbb{R}^d$ for some value of d. The states are partitioned into two regions: a bounded open region E_0 and its complement E_0^c. Let τ denote the first time Y hits E_0^c. We are interested in the distribution of τ.

This paper considers the task of computing the mean exit time $E[\tau]$ as well as the higher moments $E[\tau^m]$ of the exit time distribution. The formulation is simpler when the goal is solely to compute the mean exit time so we begin with this. We then extend the formulation by including time as a component of the process in order to determine the higher moments.

2.1 Mean exit times.

As in (1), the quantity

$$f(Y_t) - f(Y_0) - \int_0^t Af(Y_s)ds \tag{7}$$

is a martingale for each f in the domain $\mathcal{D}(A)$ and the same argument which established (2) shows that for all practical cases of interest

$$E[f(Y_\tau)] - E[f(Y_0)] - E\left[\int_0^\tau Af(Y_s)ds\right] = 0. \tag{8}$$

Define the occupation measure μ_0 and exit distribution μ_1 by

$$\mu_0(B) = E\left[\int_0^\tau I_B(Y_s)ds\right] \qquad \text{and} \qquad \mu_1(B) = P(Y_\tau \in B)$$

for sets B. It then follows that (8) can be written as

$$\int_{E_0} Af(y)\mu_0(dy) + f(y_0) - \int_{E_0^c} f(y)\mu_1(dy) = 0. \tag{9}$$

We refer to (9) as the *basic adjoint equation.* In [12], it is shown in a controlled setting that for each μ_0 and μ_1 satisfying (9) there is a process Y and a stopping time τ for which (8) is satisfied and τ is essentially the first exit time of Y. Thus (9) characterizes the occupation measure μ_0 and exit distribution μ_1 of processes having generator A.

For many processes of interest, such as diffusions and jump processes, the operator A is composed of differential operators and/or difference operators. This fact implies that for functions of the form $f(y) = y^k$, Af has the representation $Af(y) = \sum_{i=0}^n c_i y^i$ for some value n and coefficients c_i. (Refer to the following sections for specific examples.) This representation means that when $f(y) = y^k$ is used, (9) can be entirely written in terms of the moments $m_k = \int y^k \mu_0(dy)$ of μ_0 and the moments $b_k = \int y^k \mu_1(dy)$ of μ_1; for example with $f(y) = y^k$ and Af having the above representation, the basic adjoint equation (9) becomes

$$\sum_{i=0}^n c_i m_i + y_0^m - b_k = 0. \tag{10}$$

Furthermore, the moments characterize the measures so it is sufficient to work with the moments of the measures. Finally, as with the Poisson process of the previous section, the mean exit time is $E[\tau] = \int_0^\tau 1ds = \int y^0 \mu_0(dy) = m_0$.

At this point, we have rephrased the problem of determining the mean exit time to one of determining the moments of the measures μ_0 and μ_1 along with the requirement that these sequences satisfy the equations of the form (10) derived from (9).

Care must now be taken. The constraints are in terms of sequences $\{m_k\}$ and $\{b_k\}$ which satisfy (10). These conditions, however, are *not* sufficient to guarantee that the sequences are the moments of measures. It is therefore necessary to include additional constraints which characterize such sequences as moment sequences. For simplicity of exposition, we consider the specific case of $E_0 = (0,1) \subset \mathbb{R}$. The adjustments necessary for other intervals and for higher dimensions are given in the examples.

The additional constraints are the Hausdorff moment conditions which come from the observation that for any measure ν on [0,1]

$$\int y^k (1-y)^n \nu(dy) \geq 0, \qquad k = 0,1,2,\ldots; \quad n = 1,2,\ldots,$$

since the integrand is non-negative. These conditions become, for $k = 0,1,2,\ldots$ and $n = 1,2,3,\ldots$

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} m_{j+k} \geq 0. \tag{11}$$

Clearly it is necessary for the moments to satisfy the conditions (11). They are also sufficient to guarantee that the sequence corresponds to the moments of some measure ([8], p. 222). *The crucial property of these moment conditions is that they are linear* so they fit directly into the linear programming constraints. The reference [8] also discusses how (11) can be expressed as linear conditions involving iterative differences of the moment sequence. Both sets of conditions have been used in the examples with substantially the same results.

In summary, the goal of computing the mean exit time $E[\tau]$ has been reformulated as one of determining the value m_0 of the sequence $\{m_k\}$ of nonnegative values which, together with $\{b_k\}$, satisfies (10) and (11). Linear programming enters by choosing m_0 as the objective function and optimizing the criterion in order to find the feasible solution $\{m_k\}$ and $\{b_k\}$.

Again, we must be careful. The feasible sequences $\{m_k\}$ and $\{b_k\}$ are unique when the entire sequences are considered and therefore any objective function could be used to identify these sequences. However, in order to numerically compute these values, we must restrict consideration to a finite number of moments with the result that the set of feasible (finite) sequences is no longer necessarily unique. Since the true sequence of moments is included in the set of feasible points, we can use linear programming to provide bounds on the moments. By running the LP as a minimization, we determine a lower bound on the mean exit time. Similarly, by maximizing m_0, we obtain an upper bound on $E[\tau]$. In the examples, we use this technique to provide bounds. We also compare these bounds with exact values.

2.2 Higher moments

To determine the higher moments of the exit time of the process, it is necessary to augment a time component to the state and consider the time-space generator

$$\gamma(s)Af(y) + \gamma'(s)f(y) \tag{12}$$

where $f \in \mathcal{D}(A)$ and $\gamma \in \widehat{C}^1(\mathbb{R}^+)$ ($\widehat{C}$ denoting functions vanishing at ∞). The measures μ_0 and μ_1 are now defined by

$$\mu_0(B) = E\left[\int_0^\tau I_B(s, Y_s)\, ds\right] \qquad \text{and} \qquad \mu_1(B) = P((\tau, Y_\tau) \in B)$$

for subsets B of time-space. Note that μ_1 is the joint distribution of the exit time and exit location $(\tau, Y(\tau))$. Observe also that the nth-moment of the exit time is given by

$$E[\tau^n] = nE\left[\int_0^\tau s^{n-1}\, ds\right] = n\int_{\mathbb{R}^+ \times E_0} s^{n-1}\mu_0(ds \times dy) \tag{13}$$

and, in particular, the mean exit time is again the total mass of μ_0:

$$E[\tau] = \mu_0(\mathbb{R}^+ \times E_0). \tag{14}$$

The basic adjoint equation (9) now takes the form

$$\begin{aligned} 0 \;=\; & \int_{\mathbb{R}^+ \times E_0} [\gamma(s)Af(y) + \gamma'(s)f(y)]\, \mu_0(ds \times dy) + \gamma(0)f(y_0) \\ & - \int_{\mathbb{R}^+ \times E_0^c} \gamma(s)f(y)\mu_1(ds \times dy) \end{aligned} \tag{15}$$

for all $\gamma \in \widehat{C}^1(\mathbb{R}^+)$ and $f \in \mathcal{D}(A)$.

As in the case of the mean exit time, the measures μ_0 and μ_1 characterize processes Y having generator A up to the exit time τ. We wish to further characterize the measures μ_0 and μ_1 by their moments using analogs of the Hausdorff moment conditions (11). In order to do so, we truncate the unbounded time domain $\mathbb{R}^+$ to a finite interval $[0, T]$. Thus the process Y evolves in the bounded time-space region $[0, T) \times E_0$. The process exits when it first hits the complement: $[0, T] \times E_0^c$ or $\{T\} \times E_0$. The measure μ_0 has its support on $[0, T) \times E_0$, whereas μ_1 will have its support on the disjoint sets $[0, T] \times E_0^c$ and $\{T\} \times E_0$.

The idea is to choose test functions of the form $f_{m,n}(t, y) = (t/T)^m\, y^n$. The basic adjoint equation (15) can then be written in terms of the joint moments and the Hausdorff conditions can be extended to this new setting. We defer discussion of the details to sections 7 and 8.

Truncating the time dimension introduces an approximation into the method. Rather than the exit time τ, the analysis will evaluate the distribution of $\tau \wedge T$, the minimum of τ and T. For large values of T, the process Y will exit E_0 by time T with probability close to 1 so with high probability $\tau \wedge T = \tau$.

3 An interest rate process

3.1 The model

To begin with, we consider the process $(Y_t)_{t\geq 0}$ defined by

$$dY_t = (\alpha Y_t + \beta)dt + \sigma\sqrt{Y_t}dW_t, \qquad Y(0) = y_0 \in (r, 1), \tag{16}$$

where $\alpha \in \mathbb{R}$, $\beta, \sigma \geq 0$ are constants and $0 < r < 1$. Let τ denote the first exit time of the process $(Y_t)_{t\geq 0}$ from the interval $(r, 1)$. In this section, we are interested in computing the mean exit time $E[\tau]$. The process (16) is the constant coefficients case of the Cox-Ingersoll-Ross interest rate model, see, e.g. [1], [21], for details and further references.

In order to compute $E[\tau]$, we consider the basic adjoint relation (10) using the generator

$$Af(y) = \frac{\sigma^2}{2} y f''(y) + (\alpha y + \beta) f'(y)'$$

for f a twice continuously differentiable function, i.e. an element from the domain of A.

Observe that for this model, $E_0 = (r, 1)$ and $E_0^c = \{r\} \cup \{1\}$. Note that the probability measure μ_1 on E_0^c consists of point masses at $\{r\}$ and $\{1\}$ so $\mu_r + \mu_1 = 1$, where (slightly abusing notation) $\mu_r = \mu_1(\{r\})$ and $\mu_1 = \mu_1(\{1\})$. It is important to realize that μ_r and μ_1 give the probability that when the process exits it does so at r and 1, respectively.

For the collection $\mathcal{D}_0$, we shall use

$$\mathcal{D}_0 = \left\{ f(y) = \left(\frac{y-r}{1-r} \right)^n \right\}_{1 \leq n < \infty}.$$

The reason for choosing this set of functions is that after the obvious change of variables

$$x := \frac{y-r}{1-r}, \qquad r \leq y \leq 1,$$

$\mathcal{D}_0$ specifies the monomials on (0,1).

For $k = 0, 1, 2, \ldots$, define $m_k = \int_{(r,1)} (y-r)/(1-r))^k \mu_0(dy)$. Then for $f(y) = ((y-r)/(1-r))^n$, $n \geq 1$, after some straightforward calculations, the basic adjoint equation (10) becomes, for $n \geq 1$,

$$\begin{aligned} 0 &= n\alpha m_n + \frac{2n\alpha r + n(n-1)\sigma^2 + 2n\beta}{2(1-r)} m_{n-1} + \frac{n(n-1)r\sigma^2}{2(1-r)^2} m_{n-2} \\ &\quad + \left(\frac{y_0 - r}{1-r} \right)^n - \mu_1. \end{aligned} \tag{17}$$

We wish to introduce a slight modification to the above analysis which provides better numerical results. Observe that only μ_1 appears in (17) due to the fact that $f(r) = 0$ for each test function in $\mathcal{D}_0$. We add a test function to $\mathcal{D}_0$ which does not vanish at r so that the constraints derived from the basic adjoint relation includes a condition involving μ_r. We add the function $f(y) = \ln(y)$ which results in the constraint

$$0 = \int_{(r,1)} \left(\left(\beta - \frac{\sigma^2}{2} \right) + \alpha y \right) \frac{1}{y} \mu_0(dy) + \ln(y_0) - \ln(r)\mu_r. \tag{18}$$

This expression now poses a difficulty since the integrand has a term with y^{-1} that cannot be expressed using moments. We therefore introduce a change of measure by defining the measure ν on $(r, 1)$ having Radon-Nikodym derivative

$$\frac{d\nu}{d\mu_0}(y) = \frac{1}{y} \qquad y \in (r, 1). \tag{19}$$

Let $\tilde{m}_n = \int ((y-r)/(1-r))^n \, \nu(dy)$, then

$$\begin{aligned} m_n &= \int \left(\frac{y-r}{1-r}\right)^n \mu_0(dy) \\ &= \int \left(\frac{y-r}{1-r}\right)^n y\nu(dy) \\ &= (1-r)\tilde{m}_{n+1} + r\tilde{m}_n. \end{aligned}$$

Thus, the basic adjoint relations (18) and (17) can be expressed in terms of $\{\tilde{m}_n\}$ as:
For $f(y) = \ln(y)$,

$$0 = \left(\alpha r + \beta - \frac{\sigma^2}{2}\right)\tilde{m}_0 + \alpha(1-r)\tilde{m}_1 + \ln(y_0) - \ln(r)\mu_r$$

and for $f(y) = ((y-r)/(1-r))^n$,

$$\begin{aligned} 0 &= n\alpha(1-r)\tilde{m}_{n+1} + \left(2n\alpha r + \frac{n(n-1)\sigma^2}{2} + n\beta\right)\tilde{m}_n \\ &\quad + \left(\frac{n(n-1)\sigma^2 r + n\alpha r^2 + n\beta r}{(1-r)}\right)\tilde{m}_{n-1} + \left(\frac{n(n-1)\sigma^2 r^2}{2(1-r)^2}\right)\tilde{m}_{n-2} + \left(\frac{y_0 - r}{1-r}\right)^n - \mu_1. \end{aligned}$$

Finally, in order to numerically solve this linear program, we restrict analysis to a finite number of moments.

3.2 Analytical and numerical results

In this section, we exhibit the numerical values computed for the mean exit time of the interest rate model in the special case that $\beta = 0$ in Tables I - IV and for $\beta > 0$ in Tables V and VI. We also display the analytical formula when $\beta = 0$ for comparison purposes. When $\beta > 0$, the analytical formula involves an additional integration and thus becomes more complicated.

Proposition 1 *Consider the case $\beta = 0$. Let $\gamma = 2/\sigma^2$, $b = \alpha\gamma$ and*

$$C = \frac{-b\gamma \int_r^1 \ln(s) e^{bs}\, ds}{e^b - e^{br}}.$$

Then the mean exit time of the process (16) from the interval $(r, 1)$ is given by

$$E[\tau] = C\left[\frac{e^{b(1-y_0)} - 1}{b}\right] + \gamma e^{-by_0} \int_{y_0}^1 \ln(s) e^{bs}\, ds.$$

The proof runs along the same lines as for the mean exit time of the Bessel process which is given in the next section.

We now compare the LP results to the analytical results. The mean exit time formula in Proposition 1 has been computed numerically using Mathematica. The numerical evaluation is denoted in the tables by "NEF" which stands for "numerical evaluation of the formula.

We begin by varying the parameter α. Notice that the LP procedure provides tight bounds on the mean exit time and compares well to the exact value as determined by numerical evaluation of the analytical formula. It is also interesting to note that the LP procedure seems to work well for the case $\alpha = 10$ when the numerical evaluation fails to give a good value.

Table I

Bounds on $E[\tau]$ as a function of α ($r = 0.1$, $y_0 = 0.2$, $\sigma = 0.5$; $M = 64$)

α	LP – Simplex Algorithm				NEF
	primal algorithm		dual algorithm		
	lower bound	upper bound	lower bound	upper bound	
-1.0	0.427392	0.427888	0.427608	0.428098	0.427812
-0.5	0.632675	0.632975	0.632196	0.632943	0.632756
0.0	0.937386	0.938037	0.937456	0.938286	0.937707
0.5	1.04868	1.04925	1.0487	1.04906	1.04885
1.0	0.950042	0.950767	0.949977	0.950631	0.950471
2.0	0.70233	0.703849	0.694228	0.707175	0.70373
3.0	0.52267	0.5315	0.518636	0.535686	0.530723
10.0	0.153071	0.166396	0.156081	0.166416	$-1.80389 \cdot 10^{11}$

Table II compares the values as the location of the initial point of the process varies. These results are quite comparable.

Table II

Bounds on $E[\tau]$ as a function of y_0 ($r = 0.1$, $\alpha = 1.0$, $\sigma = 0.5$; $M = 64$)

y_0	LP – Simplex Algorithm				NEF
	primal algorithm		dual algorithm		
	lower bound	upper bound	lower bound	upper bound	
0.11	0.161416	0.162574	0.161775	0.162121	0.162061
0.2	0.950042	0.950767	0.949977	0.950631	0.950471
0.3	1.09844	1.10043	1.09929	1.09954	1.0994
0.4	0.983295	0.983441	0.983301	0.983725	0.983384
0.5	0.79221	0.796688	0.794403	0.794676	0.794431
0.7	0.421597	0.422227	0.421597	0.421666	0.421615
0.9	0.123583	0.123774	0.123578	0.123623	0.123597

In Table III, the diffusion coefficient varies and the results seem very good for the LP procedure and somewhat less satisfactory for the numerical evaluations of the analytical formula when $\sigma = 0.2$.

Table III

Bounds on $E[\tau]$ as a function of σ ($r = 0.1$, $y_0 = 0.2$, $\alpha = 1.0$; $M = 64$)

σ	LP – Simplex Algorithm				NEF
	primal algorithm		dual algorithm		
	lower bound	upper bound	lower bound	upper bound	
0.0	1.5351	1.60944	1.55723	1.60944	1.60944
0.2	1.62754	1.70275	1.62329	1.7027	6.0
0.5	0.950042	0.950767	0.949977	0.950631	0.950471
1.0	0.257998	0.258277	0.258128	0.258244	0.258208
2.0	0.0605077	0.0606367	0.0605235	0.0607046	0.0605481
5.0	0.00938572	0.00946213	0.00937848	0.00946491	0.00942956

Finally when $\beta = 0$, in Table IV we illustrate the dependence of the LP method on the choice of the number of moments. As can be seen, it is necessary to choose n reasonably large to obtain good accuracy using the LP method with the moments.

Table IV

$E[\tau]$ as a function of n ($r = 0.1$, $y_0 = 0.2$, $\alpha = 1.0$, $\sigma = 0.5$)

n	LP	NEF
5	0.395176	0.950471
10	0.880307	0.950471
15	0.93576	0.950471
20	0.944859	0.950471
25	0.948817	0.950471
31	0.950036	0.950471

Table V and Table VI give bounds on the mean exit time for two choices of α using various values of $\beta > 0$.

Table V

Bounds on $E[\tau]$ as a function of β for $\alpha = 1.0$ ($r = 0.1$, $y_0 = 0.2$, $\sigma = 0.5$; $M = 64$)

β	LP – Simplex Algorithm			
	primal algorithm		dual algorithm	
	lower bound	upper bound	lower bound	upper bound
0.0	0.950042	0.950767	0.949977	0.950631
0.2	0.936516	0.956607	0.944448	0.955779
0.5	0.731062	0.809656	0.74738	0.809656
1.0	0.501042	0.524712	0.509875	0.524713
2.0	0.304351	0.313318	0.300605	0.313318
5.0	0.134958	0.143425	0.13868	0.143425

Table VI

Bounds on $E[\tau]$ as a function of β for $\alpha = -1.0$ ($r = 0.1$, $y_0 = 0.2$, $\sigma = 0.5$; $M = 64$)

β	LP – Simplex Algorithm			
	primal algorithm		dual algorithm	
	lower bound	upper bound	lower bound	upper bound
0.0	0.427392	0.427888	0.427608	0.428098
0.2	0.954145	0.956959	0.953132	0.95775
0.5	3.37823	4.05991	3.40816	3.98639
1.0	1.56354	1.72669	1.57002	1.72669
2.0	0.541284	0.563649	0.544775	0.563649
5.0	0.16828	0.181594	0.175809	0.181594

4 The Bessel process

The Bessel processes provide a rich class of stochastic processes for testing our numerical method. Since many analytical results are known for this class of stochastic processes (see [14] and [22]), we can compare the numerical values with analytical ones.

Another reason for us to look at the Bessel process is that it provides an example where the drift coefficient is an unbounded function on the state space of the process, see equation (20) below.

4.1 The Model

Let $W = (W_t)_{t \geq 0}$ denote a d-dimensional Brownian motion process starting at $w_0 \in \mathbb{R}^d$. The Euclidean norm of W_t, $Y_t := \|W_t\|$, is called the d-dimensional Bessel process. It is well known that Y has a representation in terms of a one-dimensional Brownian motion process $B = (B_t)_{t \geq 0}$; that is, Y satisfies the stochastic differential equation (see [14])

$$dY_t = \left(\frac{d-1}{2}\right) \frac{1}{Y_t}\, dt + dB_t, \qquad Y_0 = \|W_0\| =: y_0. \tag{20}$$

This representation provides, among other things, the opportunity to define Bessel processes for fractional dimensions by allowing d to be any number greater than or equal to 1. In the sequel, we define $\alpha = (d-1)/2$. The generator of the Bessel process is thus

$$Af(y) = \frac{1}{2} f''(y) + \frac{\alpha}{y} f'(y), \qquad y > 0. \tag{21}$$

In this section, we are interested in the mean exit time of the Bessel process from an interval. By rescaling, if necessary, we can take the right endpoint of the interval to be 1 which will imply that the left endpoint is at some value r less than 1. Let τ denote the exit time of Y from an interval $(r, 1)$, where $0 < r < y_0 < 1$. We wish to compute the mean exit time $E[\tau]$.

As in the interest rate model, we choose for the class of test functions

$$\mathcal{D}_0 = \left\{ f(y) = \left(\frac{y-r}{1-r}\right)^n, n \geq 1 \right\} \cup \{f(y) = \ln(y)\}.$$

In place of the Radon-Nikodym transformation (19) we use the transformation

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

Then letting $m_n = \int ((y-r)/(1-r))^n \cdot 1/(y^2) \mu_0(dy)$, the basic adjoint relation (9) becomes for $f(y) = ((y-r)/(1-r))^n$, where $n \geq 1$,

$$0 \;=\; \left(\frac{n(n-1)}{2} + n\alpha\right) m_n + n \left(\frac{r}{1-r}\right) (\alpha + n - 1) m_{n-1}$$

$$+\frac{n(n-1)}{2}\left(\frac{r}{1-r}\right)^2 m_{n-2}+\left(\frac{y_0-r}{1-r}\right)^n-\mu_1$$

and for $f(y)=\ln(y)$,

$$0=\left(\alpha-\frac{1}{2}\right)m_0+\ln(y_0)+\ln(r)\,\mu_r.$$

4.2 Analytical and numerical results

The analytical formulas depend on the dimension d and are summarized in the following proposition.

Proposition 2 *The mean exit time of Y from the interval $(r,1)$ is*

$$E[\tau]=\frac{1}{d}(1-y_0^2)-\left(\frac{1-r^2}{d}\right)\frac{f(1)-f(y_0)}{f(1)-f(r)}\,, \tag{22}$$

where

$$f(y)=\begin{cases} y^{2-d}, & d\neq 2\\ \ln(y), & d=2.\end{cases} \tag{23}$$

Proof.

We first address the case $d\neq 2$ and consider the boundary value problem

$$\begin{aligned}\frac{1}{2}V''(y)+\frac{d-1}{2y}V'(y)+1=0, &\qquad 0<r\leq y\leq 1,\\ V(r)=V(1)=0. &\end{aligned} \tag{24}$$

We seek a solution of the form $V(y)=ay^2+by^n+c$. With this choice of V, we have

$$\begin{aligned}\frac{1}{2}V''(y)+\frac{d-1}{2y}V'(y)+1 &= \left(a+\frac{n(n-1)}{2}by^{n-2}\right)+\left(a(d-1)+\frac{n(d-1)}{2}by^{n-2}\right)+1\\ &= ad+1+\frac{n}{2}(n+d-2)by^{n-2}.\end{aligned}$$

The equation (24) requires $a=-1/d$ and $n=2-d$. The values of b and c are obtained from the boundary conditions $V(r)=V(1)=0$ yielding $b=(1-r^2)/[d(1-r^{2-d})]$ and $c=-b+1/d$. Thus

$$V(y)=\frac{1}{d}(1-y^2)-\left(\frac{1-r^2}{d}\right)\frac{1-y^{2-d}}{1-r^{2-d}}\,.$$

Letting Y be the solution of the stochastic differential equation (20), with $f = V$ it follows that for each $t > 0$

$$E\left[V(Y_{t\wedge\tau}) - V(Y_0) - \int_0^{t\wedge\tau} AV(Y_s)\,ds\right] = 0.$$

Using the fact that V satisfies (24), $\tau < \infty$ a.s. and letting $t \to \infty$

$$E\left[V(Y_0)\right] = E\left[\int_0^{\tau} 1 ds\right]$$

which gives (22) with f defined in (23) for $d \neq 2$.

For the case $d = 2$, the boundary value problem (24) takes the form

$$\begin{aligned} \frac{1}{2}V''(y) + \frac{1}{2y}V'(y) + 1 &= 0, \qquad 0 < r \leq y \leq 1, \\ V(r) = V(1) &= 0, \end{aligned}$$

and we seek a solution of the form $V(y) = ax^2 + b\ln(y) + c$. A similar analysis results in (22) with f given by (23) for $d = 2$. $\square$

Remark 1 *For $d = 1$ and $d = 3$, the formulas simplify to*

$$E[\tau] = (1 - y_0)(y_0 - r) = -y_0^2 + (1 + r)y_0 - r, \qquad d = 1$$

and

$$E[\tau] = \frac{1}{3}\left[1 - y_0^2 + r(r+1)\left(1 - \frac{1}{y_0}\right)\right], \qquad d = 3.$$

It should also be noted that the expression for $E[\tau]$ is continuous as a function of d at $d = 2$.

Table VII below lists a few representative values of $E[\tau]$ as a function of the (generalized) dimension of the Bessel process with $r = 0.1$ and $y_0 = 0.4$. The numbers are obtained running an implementation of the Hausdorff conditions using approximately $(M + 1)(M + 2)/2$ variables as the iterative differences, where M denotes the highest moment. We have taken $M = 64$. The program has been run using AMPL together with CPLEX as the LP solver. Note the excellent agreement of the LP results with the analytical ones.

Table VII

Bounds on $E[\tau]$ as a function of d ($r = 0.1$, $y_0 = 0.4$; $M = 64$)

Generalized Dimension d	LP – Simplex Algorithm				NEF
	primal algorithm		dual algorithm		
	lower bound	upper bound	lower bound	upper bound	
1	0.18	0.18	0.18	0.18	0.18
2	0.22302	0.22302	0.22302	0.22302	0.2230197
$\frac{e^2-3}{2}$	0.22691	0.227124	0.226888	0.227091	0.2269933
$\sqrt{5}$	0.227496	0.227629	0.227513	0.227914	0.2276065
e	0.228516	0.229261	0.227818	0.229698	0.2287905
3	0.223904	0.225631	0.223631	0.227126	0.225
5	0.163524	0.168	0.163004	0.168	0.16510135
10	0.0837364	0.084	0.0836852	0.084	0.0839985

We have also computed the mean exit times for the Bessel processes varying each of the other parameters; for example varying r while keeping y_0 and d fixed, and similarly for y_0. The results are similar to the corresponding results for the interest rate model.

5 Change-point detection

Rapid detection of changes in a system's behavior is important in quality control. Several procedures for detecting changes, such as CUSUM, EWMA and the Roberts-Shiryaev procedure, have been extensively discussed in the literature for discrete and continuous time models (see [13], [18]). All three procedures can be described in terms of one-dimensional diffusion processes and lead to exit time problems. The CUSUM and EWMA procedures result in diffusion processes with constant diffusion coefficients, whereas the Roberts-Shiryaev procedure involves a more interesting diffusion with state dependent diffusion coefficient. One important operating characteristic of all of these procedures is the exit time of the monitoring statistic.

5.1 The model

The basic continuous time model for change-point detection has the form

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

where Z is the state process and W is a standard Brownian motion process. In the simplest case the parameters ρ an σ are known constants. The unknown parameter θ marks the onset of a change in the drift of the process Z. The Roberts-Shiryaev procedure is defined to give an alarm at the time

$$\tau = \inf\{t > 0 : Y_t > H\}$$

where H is a specified limit, and where the monitoring statistic Y is defined as

$$Y_t = \int_0^t \exp\left[\frac{\rho}{\sigma}(Z_t - Z_s) - \frac{\rho^2}{2\sigma^2}(t-s)\right] ds.$$

Note that the integrand is the likelihood ratio, based on observations up to time t, of a change in drift at time s with respect to no drift. Due to the scaling properties of Brownian motion, we may assume, without loss of generality, that $(\rho^2)/(2\sigma^2) = 1$. Under the measure defined by the process Z for $\theta = 0$, Itô's formula implies that Y satisfies the stochastic differential equation

$$dY_t = (1 + 2Y_t)\, dt + \sqrt{2} Y_t\, dW_t, \qquad Y_0 = 0.$$

The generator of the process is

$$Af(y) = y^2 f''(y) + (1 + 2y) f'(y).$$

The set $E_0 = [0, H)$ and $E_0^c = [H, \infty)$. Observe that the process exits E_0 by hitting the point $\{H\}$. Thus $\mu_H = 1$, where $\mu_H = \mu_1(\{H\})$.

As test functions, we select $\mathcal{D}_0 = \{f(y) = (y/H)^n : n \geq 1\}$. Using the change of variables $y = Hx$, the basic adjoint relation becomes for $f(y) = (y/H)^n$,

$$0 \;=\; n(n+1)m_n + \frac{n}{H} m_{n-1} + \left(\frac{y_0}{H}\right)^n - \mu_H.$$

The mean exit time for the process Y is given by m_0.

5.2 Analytical and numerical results

A proof similar to that of Proposition 2 (cf. [18]) shows the following:

Proposition 3 *The mean exit time of* Y *from the interval* $(0, H)$ *is*

$$E[\tau] = e^{1/H} \int_{1/H}^{\infty} \frac{e^{-x}}{x} \, dx. \tag{25}$$

We compute $E[\tau]$ numerically using the LP approach and compare the results with numerical evaluation of (25).

Table VIII displays the numbers obtained from the LP approach and a naive numerical integration of (25) using MATLAB.

Table VIII

Bounds on the mean exit time of Y as a function of H

($M = 40$, for $1 \leq H \leq 3$; $M = 30$, for $4 \leq H \leq 10$)

H	LP – Simplex Algorithm				NEF
	primal algorithm		dual algorithm		
	lower bound	upper bound	lower bound	upper bound	
1	0.5962863986	0.596448402	0.5961816556	0.596448402	0.596348
2	0.9215313919	0.924721350	0.9206597148	0.9247213502	0.922911
3	1.152162073	1.164181386	1.150164531	1.164181386	1.156807
4	1.325581297	1.35676102	1.324911822	1.35676102	1.340887
5	1.466797809	1.519779467	1.469156691	1.519779467	1.493351
6	1.589625312	1.666436895	1.578947596	1.666436895	1.623802
7	1.681500163	1.822192892	1.681118009	1.822192892	1.737974
8	1.764363379	1.93783293	1.765960078	1.93783293	1.839588
9	1.847369069	2.06032195	1.841958313	2.06032195	1.931201
10	1.905816693	2.186281308	1.895024576	2.186281308	2.014652

6 Two-dimensional Brownian motion process

In this section we illustrate the LP method by looking at a multivariate example, the two-dimensional Brownian motion process in the unit square.

6.1 The model

Let $(X_s, Y_s) = (x_0 + W_s^1, y_0 + W_s^2)$ where $x_0, y_0 \in (0,1)$ and $W = (W_s^1, W_s^2)$ is a two-dimensional standard Brownian motion process. The generator of the process (X_s, Y_s) is given for $f \in \mathcal{D}(A) = \{f : f \text{ twice continuously differentiable on } \mathbb{R}^2\}$ by

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

The model considers the exit of the process from the bounded domain $E_0 = (0,1)^2$; the boundary consists of four parts $F_{\text{bottom}}{=}[0,1]{\times}\{0\}$, $F_{\text{top}}{=}[0,1]{\times}\{1\}$, $F_{\text{left}}{=}\{0\}\times(0,1)$ and $F_{\text{right}} = \{1\}\times(0,1)$. According to (9), 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 boundaries which together comprise the exit distribution μ_1 of the Brownian motion process from the unit square. Each boundary measure is characterized by its one-dimensional Hausdorff conditions (11) while the occupation measure on E_0 is characterized by the higher dimensional analogue (see [17])

$$\sum_{k=0}^{n}\sum_{l=0}^{m}\binom{m}{k}\binom{n}{l}(-1)^{k+l}\mu_{i+k,j+l} \geq 0 \qquad (26)$$

for all $m,n = 1,2,3,\ldots$ and $i,j = 0,1,2,\ldots$, where $\mu_{i,j} = \int_{(0,1)\times(0,1)} x^i y^j \mu_0(dx \times dy)$ denotes the i,jth moment of the occupation measure μ_0 defined on $(0,1)\times(0,1)$.

The collection of test functions we use is $\mathcal{D}_0 = \{f(x,y) = x^m y^n; m,n \geq 0\}$. The basic adjoint relations (9) can be expressed as follows. Let $\mu_{i,j}$ be as above, and $\mu_i^{(b)}$, $\mu_i^{(t)}$, $\mu_i^{(l)}$ and $\mu_i^{(r)}$ denote the ith moment of the measures on F_{bottom}, F_{top}, F_{left} and F_{right}, respectively. If we also define $0^0 = 1$, then for $m,n \geq 0$,

$$0 = \frac{m(m-1)}{2}\mu_{m-2,n} + \frac{n(n-1)}{2}\mu_{m,n-2} + x_0^m y_0^n - 0^n\mu_m^{(b)} - 1^n\mu_m^{(t)} - 0^m\mu_n^{(l)} - 1^m\mu_n^{(r)}.$$

Notice that when m and n both equal 0, this condition reduces to the requirement that the total mass of the four boundary measures is 1.

To these equality constraints we add the one-dimensional Hausdorff conditions (11) for each of the four sets of moments on the boundary and the two-dimensional Hausdorff conditions (26) on the moments of the measure in the interior in order to characterize the feasible moment sequences.

The mean exit time is given by $\mu_{0,0}$ (see (14)).

6.2 Analytical and numerical results

The mean exit time of a two-dimesional Brownian motion process from the square $(0, b)^2$ is proven in [10], p.71 to be (beware the misprint in the book)

$$E[\tau] = \frac{32b^2}{\pi^4} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{((2m-1)^2 + (2n-1)^2)^{-1}}{(2m-1)(2n-1)} \sin\left(\frac{(2m-1)\pi x_0}{b}\right) \sin\left(\frac{(2n-1)\pi y_0}{b}\right) .$$

The exact value of the mean exit time for the special case of the unit square can be computed from this formula to any desired level of accuracy. Table IX displays the mean exit times to the first six digits in the column labelled NEF. We have run the LP using various values for M; typically M is taken between 9 and 14. We used AMPL as a helpful interface, and the primal, dual and interior point method of CPLEX as the LP solvers.

Table IX reports the results when the primal method was used with $M = 11$. Due to the symmetry of the problem, we report the mean exit times for several initial positions only in the triangle below the main diagonal.

Table IX

Moments of the exit time distribution

y_0	primal algorithm					
	$x_0 = 0.5$			$x_0 = 0.8$		
	lower bound	upper bound	NEF	lower bound	upper bound	NEF
0.8				0.06872251	0.07018016	0.069294
0.7				0.08623314	0.08740675	0.086682
0.6				0.09594293	0.09690845	0.09634
0.5	0.14693254	0.14801697	0.14734	0.09905687	0.09994572	0.099396
0.4	0.14189854	0.14299119	0.14231	0.09593879	0.09690323	0.096311
0.3	0.12635959	0.12741458	0.12676	0.08622915	0.08743512	0.086682
0.2	0.09904923	0.10002026	0.099396	0.06871838	0.07018386	0.069294
0.1	0.05784276	0.05844282	0.058084	0.04114009	0.04265117	0.041761

Note that the exact values for the points $(0.8, 0.2)$ and $(0.8, 0.8)$ need to be the same. The LP results very well reflect this symmetry property of the problem.

7 Time-space Poisson process

In this section and the next one we illustrate how the LP method can be used to compute *higher moments* of the exit distribution of Markov processes. To obtain the higher moments, it is necessary to work with the time-space process in which the first component is the time and the second component is the location of the process.

We introduce the technique by returning to the simple example of the Poisson process having parameter λ. The time-space Poisson process (t, N_t) has a continuous time component and a discrete space component. This feature simplifies the LP computations by only involving a finite collection of one-dimensional measures (one for each state) and hence requires only the one-dimensional Hausdorff conditions.

7.1 The model

The generator A of the time-space process operates on test functions $f(t,x)$ for $t \geq 0$ and $x \in \{0,1,2,\ldots\}$ and takes the form

$$Af(t,x) = \frac{\partial f}{\partial t}(t,x) + \lambda(f(t,x+1) - f(t,x)),$$

where f is a continuously differentiable function in t and a sequence in x.

Let τ denote the exit time of the process N_t from the set $\{0,1,2,\ldots,K-1\}$ which coincides with the first hitting time of N_t at the value K. In order to characterize the moment sequences by the Hausdorff linear inequalities (11), we truncate the time dimension by selecting T large and consider $\widehat{\tau} = \tau \wedge T$. When T is large, the probability that the process N_t hits K before time T is close to 1 and so the distribution of $\widehat{\tau}$ will approximate that of τ. This truncation technique will, at the same time, provide numerical values for tail probabilities of the exit time.

The model now considers the exit of the process from the bounded domain $E_0 = (0,T) \times \{0,1,\ldots,K-1\}$. The boundary consists of the two sets, $F_{\text{top}} = \{T\} \times \{0,1,\ldots,K-1\}$ and $F_{\text{right}} = [0,T] \times \{K\}$. Thus the measure μ_0 on the interior is comprised of K one-dimensional measures on $[0,T)$ (one for each $[0,T) \times \{j\}$) and the boundary probability measure consists of a one-dimensional measure on $[0,T) \times \{K\}$ and a discrete measure on $\{T\} \times \{0,1,\ldots,K-1\}$.

For the collection of test functions we use

$$\mathcal{D}_0 = \left\{ f(t,x) = \left(\frac{t}{T}\right)^m I_{\{k\}}(x) : m = 0,1,2,\ldots,M,\ k = 0,1,\ldots,K \right\}$$

and define the mth moments in the time component by $\mu_m(k) = \int_0^T (t/T)^m\, I_{\{k\}}(x)\mu(dt{\times}dx)$ for $k = 0,1,\ldots,K-1$. We also define the discrete measure on $\{T\} \times \{0,1,\ldots,K-1\}$ by $\mu^{(t)}(k)$ and the moments of the measure on the right hand boundary $[0,T) \times \{K\}$ by $\mu_m^{(r)}$. With this notation, the basic adjoint relations (9) are, for $m = 0,1,\ldots M$,

$$\begin{aligned}
0 &= \frac{m}{T}\mu_{m-1}(0) - \lambda\mu_m(0) + 0^m - \mu^{(t)}(0) && \text{for } i = 0\\
0 &= \frac{m}{T}\mu_{m-1}(i) + \lambda\mu_m(i-1) - \lambda\mu_m(i) - \mu^{(t)}(i) && \text{for } 1 \leq i \leq K-1\\
0 &= \lambda\mu_m(K-1) - \mu_m^{(r)} && \text{for } i = K
\end{aligned}$$

To these constraints we add the one-dimensional Hausdorff conditions (11) on each sequence $\{\mu_0(k), \mu_1(k), \mu_2(k), \ldots\}$ for each $k = 0, 1, \ldots, K-1$ and on $\mu^{(r)}$. We also add nonnegativity constraints on $\mu^{(t)}$ and the probability constraint

$$1 = \sum_{i=0}^{K-1} \mu^{(t)}(i) + \mu_0^{(r)}.$$

The mth moment of the exit time distribution (see (13)) is given by

$$E[\tau^m] = mT^{m-1} \sum_{i=0}^{K-1} \mu_{m-1}(i).$$

7.2 Analytical and numerical results

For a Poisson process, the first hitting time τ of the state K has a Gamma distribution with parameters (K, λ). Thus explicit calculations show that the mth moment of this distribution is

$$E[\tau^m] = \binom{m+K-1}{m} \frac{m!}{\lambda^m}\,.$$

In the following table we give lower and upper bounds for the higher moments of the Poisson process. The first two columns are based on the LP program using only the first moment in the objective function. The next two columns show the results for the maximization and minimization problems using the individual moment in the objective function.

Table X

Poisson process on K steps ($K = 8$, $M = 20$, $T = 60$; $\lambda = 1$)

moments	LP – Simplex Algorithm				NEF
	objective: $E[\tau]$		objectives: $E[\tau^n]$		
	lower bound	upper bound	lower bound	upper bound	
1	8	8	8	8	8
2	72	72	72	72	72
3	719.999	720	719.938	720.002	720
4	7919.93	7920	7918.95	7920.03	7920
5	95035.8	95040	94966.4	95059.5	95040
6	1217140	1235520	1235500	1235880	1235520
7	16068900	17297300	17227400	17370800	17297300
8	239402000	262225000	230985000	259777000	259459000

8 Time-space Brownian motion process

We now demonstrate the technique for computing the higher moments of a diffusion process by considering the special case of one-dimensional Brownian motion. The difference between the Poisson process and the Brownian motion process is that it is no longer sufficient to work with one-dimensional measures; instead the time-space occupation measure is fully higher-dimensional.

8.1 The model

Let $Y_s := y_0 + W_s$, where $y_0 \in (0,1)$ and W is a one-dimensional Brownian motion process. As in the case of the Poisson process, we consider the time-space process $X = \{X_s = (s, Y_s) : s \geq 0\}$ and let τ denote the exit time of Y from the interval $(0,1)$.

The generator of X is given by (cf. (12))

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

In order to use the two-dimensional Hausdorff linear inequalities, we again truncate the time dimension by selecting T large and consider $\widehat{\tau} = \tau \wedge T$.

The model now considers the exit of the process from the bounded domain $E_0 = (0,T)\times(0,1)$; the boundary consists of three parts $F_{\text{bottom}} = [0,T]\times\{0\}$, $F_{\text{top}} = [0,T]\times\{1\}$ and $F_{\text{right}} = \{T\}\times(0,1)$. According to (9), we need to work with four measures: the occupation measure μ_0 of the process in E_0 , and measures on F_{bottom}, F_{top} and F_{right} which together comprise the exit distribution μ_1 of the time-space process on the boundary. Each boundary measure is characterized by its one-dimensional Hausdorff conditions (11) while the occupation measure on E_0 is characterized by the two-dimensional Hausdorff conditions, see inequalities (26). Let $\mu_{i,j} = \int_{(0,1)\times(0,1)} (t/T)^i\, y^j \mu_0(dt\times dy)$ denotes the i,jth moment of the occupation measure μ_0 defined on $(0,T)\times(0,1)$, subject to the scaling factor T in time.

The collection of test functions we use is $\mathcal{D}_0 = \{f(s,y) = (s/T)^m\, y^n; m,n \geq 0, m+n \geq 1\}$. The basic adjoint relations (9) can be expressed as follows. Let $\mu_{i,j}$ be as above, and $\mu_i^{(b)}$, $\mu_i^{(t)}$, and $\mu_i^{(r)}$ denote the ith moment (similarly scaled in time) of the measures on F_{bottom}, F_{top} and F_{right}, respectively. Then

$$
\begin{aligned}
0 &= \frac{m}{T}\,\mu_{m-1,n} + \frac{n(n-1)}{2}\,\mu_{m,n-2} - \mu_m^{(t)} - \mu_n^{(r)} && \text{for } m \geq 1, n \geq 2,\\
0 &= \frac{m}{T}\,\mu_{m-1,1} - \mu_m^{(t)} - \mu_n^{(r)} && \text{for } m \geq 1, n = 1,\\
0 &= \frac{m}{T}\,\mu_{m-1,0} - \mu_m^{(b)} - \mu_m^{(t)} - \mu_0^{(r)} && \text{for } m \geq 1, n = 0,\\
0 &= \frac{n(n-1)}{2}\mu_{0,n-2} + y_0^n - \mu_0^{(t)} - \mu_n^{(r)} && \text{for } m = 0, n \geq 1.
\end{aligned}
$$

To these constraints we add the one-dimensional Hausdorff conditions for the three sets of moments of the boundary measures, the two-dimensional Hausdorff conditions on the set of moments of the interior measure and the probability requirement

$$\mu_0^{(b)} + \mu_0^{(t)} + \mu_0^{(r)} = 1$$

(which can also be derived from the basic adjoint relation by taking $m = 0$ and $n = 0$ above) in order to characterize the feasible moment sequences.

The mth moment of the exit time distribution is given by $mT^{m-1}\mu_{m-1,0}$ (see (13)). Note the factor T^{m-1} adjusts for the scaling in time.

8.2 Analytical and numerical results

The moments of the exit time from $[0, 1]$ of a Brownian motion process satisfy the differential relation (see p.70 of [10])

$$f_k''(x) = -2kf_{k-1}(x), \quad f_k(0) = f_k(1) = 0$$

with $f_0(x) \equiv 1$, where $f_k(x)$ denotes the kth moment of the exit time of the Brownian motion process starting at x. Using this recursion formula, the functions f_k can be determined; for example,

$$\begin{aligned}
f_1(x) &= x - x^2 \\
f_2(x) &= \frac{1}{3}x - \frac{2}{3}x^3 + \frac{1}{3}x^4 \\
f_3(x) &= \frac{1}{5}x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{15}x^6 \\
f_8(x) &= \frac{929569}{2027025}x - \frac{43688}{57915}x^3 + \frac{2764}{7425}x^5 - \frac{248}{2835}x^7 + \frac{34}{2835}x^9 - \frac{8}{7425}x^{11} \\
&\quad + \frac{4}{57915}x^{13} - \frac{8}{2027025}x^{15} + \frac{1}{2027025}x^{16}
\end{aligned}$$

In Table XI we show the numerical results. We ran the LP using various values for M and T and tried several ad hoc scaling procedures. For simplicity we report the results when $M = 13$ and $T = 10$. We obtained the lower and upper bound by computing the minimum and maximum of the individual moment. We used AMPL as a helpful interface, CPLEX as the LP solver and employed a presolver. Notice that when only using $M = 13$ moments, the estimates are quite accurate.

Table XI

Moments of the exit time distribution

Moment	primal algorithm					
	$y_0 = 0.5$			$y_0 = 0.8$		
	lower bound	upper bound	exact	lower bound	upper bound	exact
0	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
1	0.250000	0.250000	0.250000	0.160000	0.160000	0.160000
2	0.104167	0.104167	0.104167	0.061867	0.061867	0.061867
3	0.063542	0.063542	0.063542	0.037393	0.037393	0.037393
4	0.051525	0.051526	0.051525	0.030290	0.030290	0.030290
5	0.052208	0.052211	0.052208	0.030469	0.030688	0.030688
6	0.063446	0.063510	0.063478	0.037293	0.037310	0.037312
7	0.089585	0.090057	0.090044	0.052708	0.052926	0.052926
8	0.141776	0.155844	0.145973	0.084223	0.11769	0.085801

Recall, the model involved truncating the time dimension to a bounded interval $[0, T]$. The LP then deterimined bounds on the moments of the truncated exit time $\widehat{\tau}$ which, as we see in Table XI, provides accurate bounds for the moments of the non-truncated exit time τ. In addition to providing these bounds, the LP also gives estimates of the tail probability that $\widehat{\tau} = T$. This probability is the mass $\mu_0^{(r)}$ of the measure on the right boundary $\{T\} \times (0, 1)$. By selecting $\mu_0^{(r)}$ as the objective function and performing both a minimization and a maximization, we can obtain bounds on this tail probability. Table XII illustrates this.

Table XII

Tail probabilities

T	lower bound	upper bound	exact
0.5	0.092102	0.113827	0.107977
1.0	0.007191	0.010172	0.009157

9 Concluding remarks

We have described a linear programming approach to computing the first and higher moments of the exit times of Markov processes and have illustrated this approach by considering some specific prototypical examples.

The examples, while being simple enough so that analytical formulas are known, were chosen for their potential applications as well as for characteristic features. The interest rate model provides an example of a diffusion process in which the diffusion coefficient is a fractional power of the state variable. The Bessel processes exemplify the case where the drift function is a negative power of the state, and the detection model illustrates the fact that prior information about the process can simplify the resulting linear program. The two-dimensional Brownian motion process and the time-space Brownian motion process are obviously the simplest of their kind, but the associated LP problem is generic for these higher dimensional processes.

We conclude by commenting on some of our computational experiences.

For one-dimensional problems, all LP solvers we used (CPLEX, NAG, SAS) handled problems using up to 35 moments quite well. Due to the way in which they deal with infeasibilities, some of the packages were even able to answer problems using more than 64 moments.

For two-dimensional space-time processes the numerical values of the first eight or nine moments of the exit time distribution were quite accurate when based on only 14 moments for each component.

The run times of the LPs for the various examples had a great deal of variability, depending on the dimension of the problem and the number of moments used. For the Bessel process example, the LPs took from ten to thirteen seconds to obtain a bound on the mean exit time. These were typical run times for the one-dimensional examples solely computing the mean exit time. For the two-dimensional Brownian motion process example, the LP took approximately twenty-four minutes and nearly 13,000 iterations to compute the lower bound when $(x_0, y_0) = (0.5, 0.3)$ in Table IX. For the higher moments of the Brownian motion process, the run times varied from approximately seventeen minutes and 21,000 iterations to sixty-two minutes and 62,000 iterations. These computations were run using CPLEX 5.0 and CPLEX 6.0 on an Ultra 10/300. In each case most of the iterations occurred during Phase 1 of the Simplex algorithm in which the program seeks a feasible solution as a starting point for Phase 2. Thus saving the optimal basis when parameters are changed that do not change the dimension of the problem could dramatically reduce the overall computing time.

The mathematically structured constraints of the linear programs (e.g. the Hausdorff inequalities and the difference equations for the basic adjoint relations) benefit greatly by employing matrix generator languages since the number of constraints in the resulting LP problems can easily run into the tens of thousands.

In light of the size of the LP problems for the exit time problems and the fact that many analytical results are known for these problems, these linear programs could be included in test banks to compare different LP solvers.

References

[1] M.W. Baxter and A.J.O. Rennie, *Financial Calculus,* Cambridge University Press, Cambridge (1996).

[2] A. G. Bhatt and V. S. Borkar, Occupation measures for controlled Markov processes: Characterization and optimality, *Ann. Probab.*, **24** (1996), pp. 1531–1562.

[3] J.G. Dai and J.M. Harrison, Steady-state analysis of RBM in a rectangle: Numerical methods and a queueing application, *Annals of Appl. Probab.¡* **1** (1991), pp. 16–35.

[4] J.G. Dai and T.G. Kurtz, The sufficiency of the basic adjoint relationship, unpublished manuscript (1991).

[5] D.A. Dawson, Geostochastic calculus, *Can. J. Stat.,* **6** (1978), pp. 143–168.

[6] D.A. Dawson, Qualitative behavior of geostochastic systems, *Stoch. Proc. Appl.,* **10** (1980), pp. 1–31.

[7] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York (1986).

[8] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, New York (1965).

[9] J.M. Harrison and M.I. Reiman, On the distribution of multidimensional reflected Brownian motion, *SIAM J. Appl. Math,* **41** (1981), pp. 345–361.

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

[11] T.G. Kurtz, Martingale problems for constrained Markov processes, in *Recent Advances in Stochastic Calculus*, J.S. Baras, V. Mirelli, eds., Springer–Verlag, New York, 1990.

[12] T.G. Kurtz and R.H. Stockbridge, Existence of Markov Controls and Characterization of Optimal Markov Controls, *SIAM J. Control. Optim.,* **36** (1998), pp. 609–653.

[13] M. Pollack and D. Siegmund, A diffusion process and its application to detecting a change in the drift of a Brownian motion process, *Biometrika*, **72** (1985), pp. 267–280.

[14] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 2nd ed., Springer–Verlag, Berlin, Heidelberg, New York, 1994.

[15] E. Schwerer, A linear programming approach to the steady-state analysis of reflected Brownian motion, Ph.D. thesis, Stanford University (1996).

[16] E. Schwerer, A linear programming approach to the steady-state analysis of reflected Brownian motion, *preprint* (1998).

[17] J. Shohat and J. Tamarkin, *The Problem of Moments* (1st ed.), American Mathematical Society, Providence, Rhode Island (1943).

[18] M.S. Srivastava and Y. Wu, Dynamic sampling plan in Shiryaev-Roberts procedure for detecting a change in the drift of Brownian motion, *Ann. Stat.*, **22** (1994), pp. 805–823.

[19] R.H. Stockbridge, Time-average control of martingale problems: Existence of a stationary solution, *Ann. Probab.*, **18** (1990), pp. 190–205.

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

[21] P. Wilmott, S. Howison and J. Dewynne, *The Mathematics of Financial Derivatives,* Cambridge University Press, Cambridge (1995).

[22] M. Yor, *Some Aspects of Brownian Motion. Part I: Some Special Functionals*, Birkhäuser, Basel, Boston, Berlin, 1992.

[23] L.C. Young, On approximation by polygons in the calculus of variations, *Proc. Royal Society, Series A*, **141** (1933), pp. 325–341.