IMS Collections
Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz
Vol. 4 (2008) 137–153

DOI: 10.1214/074921708000000354

Determining the Optimal Control of Singular Stochastic Processes Using Linear Programming

Kurt Helmes[1] **and Richard H. Stockbridge**[2,*]

Humboldt-Universität zu Berlin and University of Wisconsin–Milwaukee

Abstract: This paper examines the numerical implementation of a linear programming (LP) formulation of stochastic control problems involving singular stochastic processes. The decision maker has the ability to influence a diffusion process through the selection of its drift rate (a control that acts absolutely continuously in time) and may also decide to instantaneously move the process to some other level (a singular control). The first goal of the paper is to show that linear programming provides a viable approach to solving singular control problems. A second goal is the determination of the absolutely continuous control from the LP results and is intimately tied to the particular numerical implementation. The original stochastic control problem is equivalent to an infinite-dimensional linear program in which the variables are measures on appropriate bounded regions. The implementation method replaces the LP formulation involving measures by one involving the moments of the measures. This moment approach does not directly provide the optimal control in feedback form of the current state. The second goal of this paper is to show that the feedback form of the optimal control can be obtained using sensitivity analysis.

1. Introduction

This paper examines a linear programming (LP) formulation for the long-term average cost of controlled stochastic processes. The processes under consideration have singular behavior (with respect to Lebesgue measure of time) that arises either from reflection or instantaneous jumping and which may include control decisions at the time of jumping. The use of linear programming to reformulate long-term average stochastic control problems began with Manne [17] in the context of a finite-state Markov chain in discrete time. This approach has been extended to general Markov processes in continuous time (lacking singular behavior) under a variety of optimality criteria in [2], [15] and [20]. The extension to include singular stochastic processes and control relies on an existence result given in [16]. In all of these LP formulations, the variables take the form of finite or probability measures, and as such, the problems are infinite-dimensional.

---

*This research has been supported in part by the U.S. National Security Agency under Grant Agreement Number H98230-05-1-0062. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation herein.

[1]Institut für Operations Research, Spandauer Str. 1, D-10178 Berlin, Germany, e-mail: helmes@wiwi.hu-berlin; url: www.wiwi.hu-berlin.de/or

[2]Department of Mathematical Sciences, Wisconsin 53201-0413, e-mail: stockbri@uwm.edu; url: www.uwm.edu/∼stockbri

AMS 2000 subject classifications: Primary 93E20, 90C05; secondary 60J60, 60J75, 93E04

Keywords and phrases: singular processes, singular and absolutely continuous control, linear programming, regime switching, repair model, stationary distributions

The LP approach has provided the foundation for analysis of *uncontrolled* stochastic processes by taking the control space to consist of a single value. Numerical implementation relies on a finite-dimensional approximation of the LP and has been shown to be effective in [7], [9] for exit time problems and [10] for steady-state analysis. Optimal stopping problems have also been solved using the LP methodology (see e.g. [3], [6], [11], [19]).

The papers [6]-[11] reformulate the LP in terms of the *moments* of the measures rather than in terms of the measures themselves. This reformulation must also include Hausdorff moment conditions, that is, a set of linear conditions which are necessary and sufficient for the infinite collection of variables to be the moments of some measure or measures on bounded regions. The finite-dimensional approximation truncates the number of moments and the Hausdorff conditions which thus allows points to be feasible that are not the initial terms of a moment sequence. The feasible set is therefore enlarged, implying that the optimal value of the approximating LPs provides an upper or lower bound (depending on the type of optimization) for the true optimal value.

The LP method has had only limited success so far in identifying optimal controls. Theoretically, an optimal control is obtained in relaxed feedback form from an optimal measure by taking the conditional distribution on the control space, given the state of the process. In practice, the selection of controls typically involves discretizing the control space (see e.g., [3], [18], [13]). This affects the reformulation by replacing measures on the product of the state space and control space by a finite collection of measures (one for each possible control value) on the state space alone. One difficulty with this discretization when using the moment reformulation is that the solution gives (pseudo-)moments of the measure corresponding to a value for the control and it is not transparent for which state values the control is active.

The first goal of this paper is to demonstrate that an analysis of the reduced cost coefficients associated with the non-basic variables in the LP determines an approximate optimal control directly from the LP solution. This method is especially effective when the optimal control is of bang-bang type. The second goal is to show that singular control problems can be solved using the LP methodology. We consider three examples of increasing levels of complexity to illustrate the methodology. These examples are presented in the following sections.

For a measurable space (S, Σ), $\mathcal{M}(S)$ denotes the collection of finite measures on (S, Σ) and $\mathcal{P}(S)$ is the subcollection of probability measures on (S, Σ).

2. Modified Bounded Follower Problem

The bounded follower problem of [1] considers a controlled process X which satisfies the stochastic differential equation

$$dX(t) = u(t)\,dt + \sigma\,dW(t), \tag{2.1}$$

in which W is a standard Brownian motion process, $\sigma > 0$ is constant and $u(t)$ is a non-anticipative process which is required to satisfy the hard constraints $u(t) \in [-1, 1]$, for all t. The objective of [1] is to minimize the long-term average second moment of X. The paper [8] modifies this problem by constraining X to remain in the interval $[0, 1]$. The constraints involve reflection at $\{0\}$ and a jump mechanism at $\{1\}$. Specifically, X is modelled as a solution of the patchwork martingale problem [14] in which the diffusion specified in (2.1) is active in the open interval $(0, 1)$, X sticks at $\{1\}$ for an exponential length of time (parameter λ) at which point it jumps to 0, and reflection occurs at $\{0\}$ by restricting the domain of the generator $Af(x,u) = uf'(x) + (\sigma^2/2)\, f''(x)$ to functions $f \in C^2[0,1]$ satisfying $f'(0) = 0$. The paper [8] demonstrates how to compare controls using a linear programming formulation of the problem and indicates numerical evidence of optimality.

The current paper extends the analysis of this model in two ways. The first is to allow *instantaneous* jumps when $X(t-) = 1$ along with the reflection at $\{0\}$. We initially formulate the processes to be considered as a quadruplet (X, Λ, L_0, N_1) which satisfies for each $f \in C^2[0,1]$

$$\begin{aligned} f(X(t)) - \int_0^t \int_{[-1,1]} Af(X(s),u)\,\Lambda_s(du)\,ds - \int_0^t B_0 f(X(s))\,dL_0(s) \\ - \int_0^t B_1 f(X(s-))\,dN_1(s). \end{aligned} \tag{2.2}$$

is a martingale, in which A is the generator above, Λ denotes a *relaxed* control process (for each s, Λ_s is a distribution on $[-1,1]$), $B_0 f(x) = f'(x)$, L_0 denotes the local time of X at $\{0\}$, $B_1 f(x) = f(0) - f(x)$ and N_1 denotes the process which counts the number of visits of X to $\{1\}$. Note, in particular, that the reflection of X at $\{0\}$ is captured through the integral term involving B_0 and so f is not required to satisfy the boundary condition $f'(0) = 0$. Also observe that the local time process L_0 and the counting process N_1 increase on sets of times which are *singular* with respect to Lebesgue measure of time.

The objective of the decision maker is to minimize the long-term average second moment

$$\limsup_{t\to\infty} t^{-1} E\left[\int_0^t X^2(s)\,ds\right]. \tag{2.3}$$

This criterion does not include any cost for using the control u so one would anticipate $u(t)$ taking only the extreme values, $u(t) \in \{-1,1\}$. This insight, however, is not assumed in determining the solution.

2.1. LP Formulation based on Measures

Let (X, Λ, L_0, N_1) satisfy (2.2). Then for each $t > 0$,

$$\begin{aligned} E[f(X(0))] = E\Bigg[f(X(t)) - \int_0^t \int_{[-1,1]} Af(X(s),u)\,\Lambda_s(du)\,ds \\ - \int_0^t B_0 f(X(s))\,dL_0(s) - \int_0^t B_1 f(X(s-))\,dN_1(s)\Bigg]. \end{aligned} \tag{2.4}$$

For $t > 0$ define the expected occupation measures (up to time t) μ^t on $[0,1]$, ν_0^t on $\{0\}$ and ν_1^t on $\{1\}$ by, for every $G \in \mathcal{B}([0,1]\times[-1,1])$,

$$\begin{aligned} \mu^t(G) &= t^{-1} E\left[\int_0^t \int_{[-1,1]} I_G(X(s),u)\,\Lambda_s(du)\,ds\right], \\ \nu_0^t(\{0\}) &= t^{-1} E[L_0(t)], \text{ and} \\ \nu_1^t(\{1\}) &= t^{-1} E[N_1(t)]. \end{aligned}$$

Since $[0,1]\times[-1,1]$ is compact, the collection $\{\mu^t : t > 0\}$ is relatively compact and hence there exist limits as $t \to \infty$. As a result, there will be corresponding limits of $\{\nu_0^t\}$ and $\{\nu_1^t\}$. Dividing by t and passing to the limit in (2.4) demonstrates that for each weak limit (μ, ν_0, ν_1) and for every $f \in C^2[0,1]$

$$\int Af(x,u)\,\mu(dx \times du) + \int B_0 f(x)\nu_0(dx) + \int B_1 f(x)\,\nu_1(dx) = 0. \tag{2.5}$$

The measure μ denotes the stationary distribution of (X, Λ) on $[0,1] \times [-1,1]$, ν_0 gives the expected long-term average amount of local time per unit of time, and ν_1 is the expected long-term average number of jumps per unit of time.

Theorem 1.7 in [16] shows that for each (μ, ν_0, ν_1) satisfying (2.5) there exists a stationary solution of (2.2) whose stationary distribution is given by μ and hence its objective function value (2.3) is given by $\int x^2 \mu(dx \times du)$. The relaxed control is given in feedback form as $\eta(X(s), \cdot)$, where η is a regular conditional distribution on $[-1,1]$ satisfying $\mu(dx \times du) = \eta(x, du)\mu(dx \times [-1,1])$. As a result, to any limiting (μ, ν_0, ν_1) arising from any control process Λ there is a stationary solution having the corresponding value. This observation indicates that optimizing over stationary processes is equivalent to optimizing over any solutions. Hence the control problem can be reformulated as an infinite-dimensional LP.

To simplify the expressions, for a measurable function g and a measure ν defined on a space (S, Σ), let $\langle g, \nu \rangle$ denote $\int g\, d\nu$. Then the linear programming formulation is

$$\text{LP1} \begin{cases} \text{Min.} & \langle x^2, \mu \rangle \\ \text{S.t.} & \langle Af, \mu \rangle + \langle B_0 f, \nu_0 \rangle + \langle B_1 f, \nu_1 \rangle = 0, \qquad \forall f \in C^2[0,1], \\ & \mu \in \mathcal{P}([0,1] \times [-1,1]), \\ & \nu_0 \in \mathcal{M}(\{0\}), \\ & \nu_1 \in \mathcal{M}(\{1\}). \end{cases}$$

Remark. An alternate approach to the solution of this stochastic control problem is to capture the singular behavior of the processes through restrictions on the domain of the generator A. Specifically, taking $f \in C^2[0,1]$ with $f'(0) = 0$ and $f(0) = f(1)$, the singular terms drop out of (2.2). Reflection of the process at $\{0\}$ is obtained by requiring the first condition and the instantaneous jump by the second condition. One is now able to solve the stochastic control problem by solving the Bellman equation in this restricted domain. The dynamic programming approach, however, becomes more complex for the other examples in this paper when the objective criterion includes a cost for the jumps which are dependent on a control variable.

2.2. LP Formulation based on Moments and Control Discretization

LP1 is the basis for the numerical solution of the control problem. Instead of allowing μ to be a probability measure on $[0,1] \times [-1,1]$, however, we discretize the set of controls to $U_k = \left\{u_j = \frac{j}{k} : j = -k, \ldots, k\right\}$ and require $\mu \in \mathcal{P}([0,1] \times U_k)$. This restriction reduces the number of feasible measures and thus provides an upper bound on LP1's optimal value. The discretization is naturally an approximation of the given problem. For all cases where the optimal control is of bang-bang type, the resulting error can, in principle, be made as small as possible by a proper choice of the discrete subset U_k. In this example (and also the next example), the restricted LP includes an optimal measure μ for LP1 and therefore solving the restricted LP provides a solution to LP1.

Define $\mu_j(\cdot) = \mu(\cdot \times \{\frac{j}{k}\})$ for $j = -k, \ldots, k$, realizing that each μ_j is a subprobability measure on $[0,1]$, with $\mu = \sum_j \mu_j$ being a probability measure. Also note that the "measures" ν_0 and ν_1 are actually point masses at $\{0\}$ and $\{1\}$, respectively.

Rather than work with the measures $\{\mu_j\}$ in the LP, we reformulate the problem again in terms of the moments, which completely determine the measures since they have support in the compact interval $[0,1]$. For each $j = -k, \ldots, k$ and for each $n \in \mathbb{N}$, define

$$m_j(n) = \int x^n \, \mu_j(dx). \tag{2.6}$$

Take $f(x) = x^n$ in (2.5) and abuse notation slightly by letting ν_0 and ν_1 denote the masses of the measures on the endpoints. Then LP1 takes the form

$$\text{LP2} \quad \begin{cases} \text{Min.} & \sum_j m_j(2) \\ \text{S.t.} & \sum_j \left[(nu_j) m_j(n-1) + \frac{n(n-1)\sigma^2}{2} m_j(n-2)\right] \\ & \qquad\qquad + n0^{n-1}\nu_0 + (0^n - 1)\nu_1 = 0, \quad \forall n \in \mathbb{N}, \\ & \sum_j m_j(0) = 1, \\ & m_j(n), \nu_0, \nu_1 \geq 0, \qquad \forall n \in \mathbb{N}. \end{cases}$$

In LP2, whenever the expression 0^0 appears, it is to be understood to equal 1.

The variables in LP2 are supposed to be the moments of measures defined on $[0,1]$; that is, we desire to have $m_j(n) = \langle x^n, \mu_j \rangle$ for some measure μ_j on $[0,1]$. The constraints in LP2, however, are not sufficient for $\{m_j(n) : n \in \mathbb{N}\}$ to be moments. Hausdorff [5] showed that necessary and sufficient conditions are provided by the set of linear inequalities obtained from the observation that for each $m, n \in \mathbb{N}$

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} \int_{[0,1]} x^{j+m} \, \nu(dx) = \int_{[0,1]} x^m (1-x)^n \, \nu(dx) \geq 0. \tag{2.7}$$

Adding (2.7) when $\nu = \mu_j$, $j = -k, \ldots, k$, to the constraint requirements of LP2 provides an equivalent LP formulation for the restricted LP1.

2.3. Finite-dimensional LPs and Numerical Results

The difficulty with this modified version of LP2 is that there are an infinite number of variables and a corresponding infinite number of constraints. To be computable, it is necessary to approximate LP2 by a finite-dimensional linear program.

One such approximation is obtained by restricting the number of moments to a finite collection, say $n = 0, 1, \ldots, M$, and limiting the constraints to those involving only the selected number of moments. A result, however, of this approximation is that the variables $\{m_j(n) : n = 0, \ldots, M\}$ are no longer guaranteed to correspond to the moments of a measure μ_j on $[0,1]$. The constraint requirements are relaxed and hence the set of feasible "pseudo-moments" is larger; that is, the feasible set of the approximating LP contains the zeroeth to Mth moments of the feasible measures of the amended LP2, but it contains other points which are not the initial terms of a moment sequence of some measure.

Now consider more carefully the constraints (2.7) when restricted to $j + m \leq M$. Each constraint defines a half-space and so the set of feasible finite sequences lies in a convex set defined by these half-spaces. This convex set is called the Hausdorff polytope. Helmes and Röhl [6] determine explicit formulas for the corner points of the Hausdorff polytope. A final modification to LP2 is therefore possible. Instead of imposing the finite Hausdorff conditions, characterize the Hausdorff polytope using convex combinations of the corner points. Thus the computable version of LP2 limits the number of variables to $M+1$ for each measure and only imposes those constraints which involve these variables, and then rewrites the variables as convex combinations of the corner points. The variables in this computable version are the convex coefficients $\{\lambda_j(n) : n = 0, \ldots, M; j = -k, \ldots, k\}$.

In addition to giving an explicit formula for the corner points, the paper [6] proves convergence of the approximating optimal solutions to an optimal solution of LP2 and, moreover, shows that the corner points can be identified with a measure that is a single point mass.

Table 1 displays a selection of values of the optimal convex coefficients $\lambda_j(n)$ corresponding to the extreme points of the Hausdorff polytope when $M = 60$. Notice that the solution only has positive weights on the corner points corresponding to the use of drift rates $\{\pm 1\}$, and that the weights correspond to $u = -1$ for the lower indices of the extreme points, whereas the higher indices have positive weights for $u = 1$. According to the results of [6], the extreme point having index n corresponds (asymptotically) to a point mass at $x = \frac{n}{M}$. Thus Table 1 tends to indicate that the control $u = -1$ is used for smaller values of x and at some point (between $\frac{40}{60}$ and $\frac{45}{60}$) the control switches to $u = 1$. The $\lambda_j(n)$ values do not provide a very accurate indication of the value of x where the switching occurs.

Sensitivity analysis of the LP can be utilized to obtain better accuracy for the switch point. The "reduced costs" are amounts by which the cost coefficients of each $\lambda_j(n)$ variable must change in order for the variable to become a basic variable; that is, should the cost coefficient change by the amount of the reduced cost for a variable $\lambda_j(n)$, then $\lambda_j(n)$ would be positive and be part of the basis for the solution. Table 2 displays the reduced costs for some of the values of n. First, notice that values of order 10^{-14} or 10^{-15} occur for those $\lambda_j(n)$ which have positive weights. These values should be understood to be numerically equivalent to 0, since the weights currently in the basis do not need to have any change in their cost coefficients in order to be basic variables.

To better distinguish the information contained in Table 2, it is helpful to scale the values by a factor of 100 and then round the values to the nearest integer. This scaling is displayed in Table 3. In contrast to the weights given in Table 1, a consistent pattern emerges with scaled reduced costs that indicates switching occurs close to index 43. Thus the control changes value from $u = -1$ to $u = 1$ when x is approximately $\frac{43}{60} \approx 0.71667$.

The numerical results depend, of course, on the choice of the highest moment. Table 4 displays the values of the optimal second moment, along with the values of the point masses p_0 and p_1 at $\{0\}$ and $\{1\}$, respectively, for a selection of values of M. The exact values can be obtained (see [12]) in which the switching location is the solution of a transcendental equation that is then used to determine the stationary density for the optimal process and hence the exact optimal value via integration. Numerical evaluation of the switch location yields $x = 0.70846$; the resulting objective function and masses are also provided in Table 4 for comparison purposes. Table 5 displays some significant scaled reduced costs for the case $M = 1024$, using a scale factor of 1000. These results indicate that the switch location lies between $x = \frac{725}{1024} \approx 0.70801$ and $x = \frac{726}{1024} \approx 0.70898$.

Table 1
Values of the weight variables $\lambda_j(n)$, $j = -3, \ldots, +3$; $\sigma = 1$, $M = 60$

index of extreme point	control indices j: j corresponds to $u = j/3$						
n	-3	-2	-1	0	1	2	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	0.175441	0	0	0	0	0	0
21	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0
29	0.033439	0	0	0	0	0	0
30	0.086116	0	0	0	0	0	0
31	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0
38	0.018578	0	0	0	0	0	0
39	0.036428	0	0	0	0	0	0
40	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0.022700
47	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0.001842
51	0	0	0	0	0	0	0.009488
52	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0.004646
55	0	0	0	0	0	0	0.000624
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

3. Regime Switching Model with Jumping Costs

The second model of this paper allows for changes in the regime of the diffusion along with control decisions to be made at the time the process hits $\{1\}$. The model contains two coordinate processes X and Y. The process Y, which tracks the regime, is a finite-state Markov chain having states $\mathbb{Y} = \{y_0, \ldots, y_l\}$ and transition rates given by a matrix $Q = (q_{yz})$. As in the modified bounded follower problem, the process X is a diffusion on the interior of $(0, 1)$, is reflected at $\{0\}$ and jumps instantaneously when $X(t-) = 1$. However, the coefficients of the diffusion now depend on the regime Y and in addition to selecting the drift rate, the decision maker also selects between several possible control actions when X hits $\{1\}$.

Let $0 = x_1 < \cdots < x_{k_1-1} < x_{k_1} = 1$ be points in the unit interval and let $\mathbb{V} = \{v_1, \ldots, v_{k_1}\}$ denote the possible singular controls. For $i < k_1$, selecting control v_i imposes an instantaneous jump to the target $\{x_i\}$ when the process hits $\{1\}$. The control v_{k_1} imposes a reflection on the process X at $\{1\}$. The absolutely continuous

TABLE 2
Reduced cost coefficients for $n = 31, \ldots, 49$; $M = 60$.

index of extreme point	control indices j: j corresponds to $u = j/3$					
n	-3	-2	-1	0	1	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
31	5.432×10^{-05}	0.02854	0.05703	0.08552	0.1140	0.1425
32	0.0001720	0.02782	0.05547	0.08311	0.1108	0.1384
33	0.0003364	0.02691	0.05348	0.08006	0.1066	0.1332
34	0.0004988	0.02575	0.05100	0.07626	0.1015	0.1268
35	0.0005846	0.02425	0.04792	0.07159	0.09526	0.1189
36	0.0005203	0.02233	0.04413	0.06594	0.08775	0.1096
37	0.0002872	0.01993	0.03958	0.05923	0.07887	0.09852
38	-5.551×10^{-15}	0.01717	0.03434	0.05150	0.06867	0.08584
39	-1.250×10^{-14}	0.01436	0.02871	0.04307	0.05742	0.07178
40	0.0009360	0.01213	0.02333	0.03453	0.04573	0.05693
41	0.003792	0.01149	0.01918	0.02688	0.03457	0.04227
42	0.009811	0.01367	0.01753	0.02139	0.02525	0.02910
43	0.02028	0.02000	0.01972	0.01944	0.01916	0.01888
44	0.03618	0.03151	0.02683	0.02215	0.01747	0.01279
45	0.05777	0.04848	0.03920	0.02992	0.02064	0.01136
46	0.08425	0.07020	0.05616	0.04212	0.02808	0.01404
47	0.1139	0.09497	0.07604	0.05711	0.03819	0.01926
48	0.1447	0.1207	0.09679	0.07286	0.04893	0.02500
49	0.1753	0.1462	0.1171	0.08807	0.05900	0.02993
⋮	⋮	⋮	⋮	⋮	⋮	⋮

and singular generators of the pair process (X, Y) are

$$\begin{aligned} Af(x,y,u) &= ub(y)f_x(x,y) + (1/2)\sigma^2(y)f_{xx}(x,y) + \sum_{z\in\mathbb{Y}} f(x,z)q_{yz}, \\ B_0 f(x,y) &= f_x(x,y), \\ B_1 f(x,y,v) &= -f_x(x,y)I_{\{v_{k_1}\}}(v) + \sum_{i=0,\ldots,k_1-1} [f(x_i,y) - f(x,y)]I_{\{v_i\}}(v). \end{aligned}$$

As in the previous example, u is again restricted to $[-1, 1]$ which means, in light of the term $b(y)$, that the decision maker is allowed to select different drift rates for the different regimes. The model includes the jump generator $\sum_z f(x,z)q_{yz}$ which implies that the regimes switch according to a Markov chain.

The processes under consideration form a sextuplet $(X, Y, \Lambda, \Psi, L_0, N_1)$, in which Ψ denotes a relaxed singular control process that chooses the values of v according to some probability measure, and satisfy the requirement that for each $f \in C^2([0,1] \times \mathbb{Y})$

$$\begin{aligned} f(X(t),Y(t)) &- \int_0^t \int_{[-1,1]} Af(X(s),Y(s),u)\Lambda_s(du)\,ds \\ &- \int_0^t B_0 f(X(s),Y(s))\,dL_0(s) \\ &- \int_0^t \int_{\mathbb{V}} B_1 f(X(s-)Y(s-),v)\Psi_s(dv)\,dN_1(s) \end{aligned} \tag{3.1}$$

is a martingale, in which Λ, Ł$_0$ and N_1 are the relaxed control process, local time process at $x = 0$ and counting process of visits to $x = 1$.

TABLE 3
Scaled reduced cost coefficients for $n = 30, \ldots, 50$; $M = 60$.

index of extreme point	control indices j: j corresponds to $u = j/3$						
n	-3	-2	-1	0	1	2	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	0	3	6	9	12	15	17
31	0	3	6	9	11	14	17
32	0	3	5	8	11	14	17
33	0	3	5	8	11	13	16
34	0	3	5	8	10	13	15
35	0	2	5	7	10	12	14
36	0	2	4	7	9	11	13
37	0	2	4	6	8	10	12
38	0	2	3	5	7	9	10
39	0	1	2	4	6	7	9
40	0	1	2	3	5	6	7
41	0	1	2	3	3	4	5
42	1	1	2	2	3	3	3
43	2	2	2	2	2	2	2
44	4	3	3	2	2	1	1
45	6	5	4	3	2	1	0
46	8	7	7	4	3	1	0
47	11	9	8	6	4	2	0
48	14	12	10	7	5	3	0
49	18	15	12	9	6	3	0
50	21	17	14	10	7	3	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

The objective of the decision maker is to minimize the expected long-term average cost

$$\limsup_{t\to\infty} t^{-1} E\left[\int_0^t c_{abs}(X(s), Y(s))\, ds + \int_0^t \int_{\mathbb{Y}} c_{sing}(Y(s-), v)\Psi_s(dv)\, dN_1(s)\right], \tag{3.2}$$

in which for illustrative purposes $c_{abs}(x, y) = c(y)x^2$, where $c(y)$ and $c_{sing}(y, v)$ denote the regime-dependent and/or decision-dependent coefficients for the cost rates. We point out that the cost structure has different costs for the different possible singular actions. The cost is higher for larger control actions. In our numerical examples, there is no cost for reflection at $\{1\}$ and the cost for jumping increases as the jump distance increases.

3.1. LP Formulation

As in Section 2.1, the stochastic control problem can be equivalently written in terms of the stationary distribution and the expected long-term average occupation measures at $\{0\}$ and on $\{1\} \times \mathbb{V}$. The infinite-dimensional LP is

$$\text{LP3} \begin{cases} \text{Min.} & \langle c_{abs}, \mu\rangle + \langle c_{sing}, \nu_1\rangle \\ \text{S.t.} & \langle Af, \mu\rangle + \langle B_0 f, \nu_0\rangle + \langle B_1 f, \nu_1\rangle = 0, \quad \forall f \in C^2([0,1]\times\mathbb{Y}), \\ & \mu \in \mathcal{P}([0,1]\times\mathbb{Y}\times[-1,1]), \\ & \nu_0 \in \mathcal{M}(\{0\}), \\ & \nu_1 \in \mathcal{M}(\{1\}\times\mathbb{V}). \end{cases}$$

TABLE 4
Objective function values and point masses as functions of M

M	objective value	p_0	p_1
16	0.10958	1.3796	0.6133
32	0.11117	1.4572	0.6377
64	0.11177	1.4640	0.6250
128	0.11193	1.5040	0.6317
256	0.11211	1.5337	0.6201
512	0.11218	1.5361	0.6276
1024	0.11225	1.5363	0.6287
exact	0.11260	1.5319	0.6194

The finite-dimensional approximation uses $f(x,y) = x^n I_{\{y_i\}}(y)$ in LP3, restricts n to the set $\{0, \ldots, M\}$, and employs the convex combination of the cornerpoints to characterize the feasible points in the Hausdorff polytope.

3.2. Numerical Results

To illustrate the success of the LP method for solving the stochastic control problem having both absolutely continuous and singular controls, we consider a particular set of parameters. In this example, there are two regimes ($\mathbb{Y} = \{0, 1\}$) and the decision maker can select from three singular control actions, so $k_1 = 3$ and $\mathbb{V} = \{1, 2, 3\}$. Control $v = 1$ requires the process X to jump to $x = 0$ when it hits $\{1\}$. Under $v = 2$, the process jumps to $x_2 = 0.5$, and the choice of $v = 3$ causes X to be reflected at $\{1\}$ so as to stay in the interval $[0, 1]$. The model parameters are given in Table 6. Notice, in particular, that when $y = 0$ the jumping costs are approximately the same, whereas the jumping cost to $\{0\}$ in state 1 is an order of magnitude larger than the cost for the process to be reset at $x = 0.5$. There is no cost for reflecting the process in either state. The selected diffusion coefficients and the switching rates are motivated by other studies ([21]). We also comment that since the optimal absolutely continuous control only takes values in $\{\pm 1\}$ (as evidenced in the previous example), we have limited our discrete choice of controls u to the set $\{-1, 0, 1\}$.

The scaled reduced cost coefficients for $M = 256$ are presented in Table 7. These numerical results indicate that the switch points for the absolutely continuous control should be located around $x = \frac{218}{256} \approx 0.852$ when $y = 0$ and near $x = \frac{237}{256} \approx 0.926$ for $y = 1$. Figure 1 displays, when $y = 0$, both the optimal u in feedback form as a function of the value of the driving force X and the optimal choice of singular control when X hits $\{1\}$. Similarly, Figure 2 displays the optimal values of u and v when $y = 1$. Since the cost for resetting to $\{0\}$ is low when $y = 0$, the decision maker makes this choice, but when $y = 1$ the cost for such a resetting is prohibitively expensive and the controller opts to reset X to $x = 0.5$. For these cost parameters, the cost for jumping is not significant enough for the decision maker to pick the reflection option; such choices are obtained when the costs for jumping are larger.

TABLE 5
Scaled reduced cost coefficients when $M = 1024$.

index of extreme point	control indices j: j corresponds to $u = j/3$						
n	-3	-2	-1	0	1	2	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
555	0	26	52	78	104	129	155
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
717	14	16	19	21	23	25	27
718	15	17	19	21	22	24	26
719	16	17	19	20	22	24	25
720	16	17	19	20	22	23	24
721	17	18	19	20	21	23	24
722	17	18	19	20	21	22	23
723	18	19	19	20	21	21	22
724	19	19	20	20	20	21	21
725	19	20	20	20	20	20	21
726	20	20	20	20	20	20	20
727	21	20	20	20	20	19	19
728	21	21	20	20	19	19	18
729	22	21	21	20	19	18	18
730	23	22	21	20	19	18	17
731	24	22	21	20	19	17	16
732	25	23	22	20	19	17	16
733	25	24	22	20	18	17	15
734	26	24	22	20	18	16	14
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
867	239	199	160	120	80	40	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

4. Repair Model

In our final example, the regime process Y represents the state of wear of a machine and X is a driving force for the level of deterioration. The levels of wear can be interpreted as moving from "new" $(x, y) = (0, 0)$ to "broken" $(x, y) = (1, 1)$, with several intermediate levels as well. In this framework, X can represent the fraction of deterioration at the current level. When X reaches $\{1\}$, Y instantaneously jumps up to the next level and X is instantaneously reset to $\{0\}$. In addition, a switching mechanism like in the previous example randomly makes the system jump from "newer" states to "older" states, with the implication that the rate matrix $Q = (q_{yz})$ is upper-triangular.

The decision maker influences the evolution of the paired process (X, Y) by choosing when to repair the machine. Thus when $Y(t-) = i$ for $0 \leq i < l$ and $X(t-) \in \mathbb{V}$, $\mathbb{V}$ a finite set of points in the open interval $(0, 1)$, the repair policy resets the driving process $X(t)$ to 0 at a cost which depends on the value at which the resetting is initiated. So the machine will be "better" after the repair but does not become "younger". Under this formulation, for levels of deterioration $X(t) < 1$ we only allow repair to the same level of wear. Should $X(t-) = 1$ when $Y(t-) = i < l$, the machine will be fixed, with $X(t) = 0$, but declared to have become "older," so $Y(t) = i + 1$. If $Y(t) = l$ we assume that repairs are no longer possible; should $X(t-) = 1$ when $Y(t-) = l$, the machine is declared to be "broken" and it is instantaneously replaced by a new machine, implying that the process has value $(X(t), Y(t)) = (0, 0)$.

TABLE 6
Model Parameters

Level	Drift	Diff.	$c_{abs}(x,y)$	$c_{sing}(y,v)$			Q	
y	$b(y)$	$\sigma(y)$	$c(y)$	$v=1$	$v=2$	$v=3$	0	1
0	1.5	0.44	1.5	0.05	0.06	0	–6.04	6.04
1	20.0	0.63	2.0	0.29	0.02	0	8.90	–8.90

TABLE 7
Scaled Reduced Cost Coefficients

index of extreme point	Regime $y=0$ control indices $u=j$		
n	-1	0	1
⋮	⋮	⋮	⋮
215	3	3	4
216	3	3	4
217	3	3	3
218	3	3	3
219	4	3	3
220	4	3	3
221	4	3	3
222	4	3	2
⋮	⋮	⋮	⋮

index of extreme point	Regime $y=1$ control indices $u=j$		
n	-1	0	1
⋮	⋮	⋮	⋮
233	16	20	24
234	17	20	23
235	18	20	22
236	19	20	21
237	19	20	20
238	20	20	19
239	21	20	18
240	22	20	17
⋮	⋮	⋮	⋮

The driving force X satisfies the stochastic differential equation

$$dX(t) = b(Y(t))\,dt + \sigma(Y(t))\,dW(t)$$

with $X(0) = 0$ and $Y(0) = 0$. Notice that this model does not include any explicit control on X and that the coefficients depend on the level of wear y.

From a modelling perspective, we briefly remark that since X is a diffusion process, the interpretation of this process as a “fraction of wear at the current level” has the implication that the deterioration of the machine can improve. One could replace the diffusion process by its running maximum so that the level of wear is monotone, as long as X is included in the model as a driving force so that the process is Markovian. Observe that the diffusion and its running maximum both hit a level within $\mathbb{V}$ at the same time so the repair mechanism would remain the same. The running maximum process increases singularly in time so would involve an additional singular generator along with an extra component (see e.g. [11] for a running maximum model). The model used in this section has the advantage of simplicity.

FIG 1. *Optimal Absolutely Continuous and Singular Control Policies in Regime $y = 0$.*

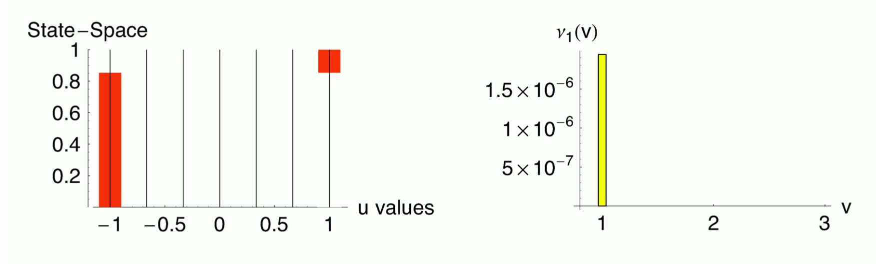
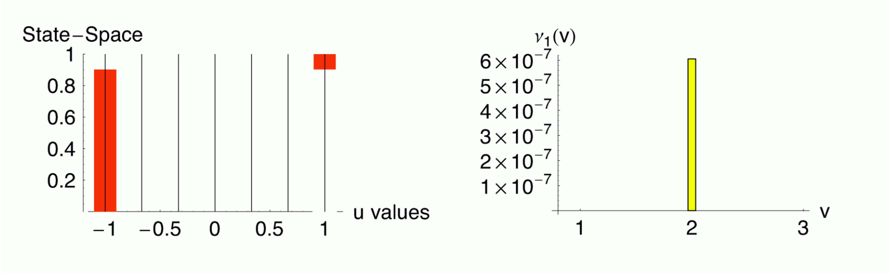


FIG 2. *Optimal Absolutely Continuous and Singular Control Policies in Regime $y = 1$.*



The generators for this repair model are

$$\begin{aligned}
Af(x,y) &= b(y)f_x(x,y) + (1/2)\sigma^2(y)f_{xx}(x,y) + \sum_{z\in\mathbb{Y}} f(x,z)q_{yz},\\
B_0 f(x,y) &= f_x(x,y),\\
B_1 f(x,y) &= \sum_{i=0}^{l-1}[f(0,y+1) - f(x,y)]I_{(1,i)}(x,y)\\
&\qquad\qquad + [f(0,0) - f(x,y)]I_{(1,l)}(x,y),\\
B_2 f(x,y,v) &= \sum_{i=0}^{l-1}[f(0,y) - f(x,y)]I_{(v,i)}(x,y).
\end{aligned}$$

A is the jump-diffusion operator for the driving force, B_0 captures the reflection of X at $\{0\}$, B_1 indicates that Y increases one level when X hits $\{1\}$, but resets when Y is at its maximum, and B_2 incorporates the control decisions. For each level $i < l$, the decision maker selects a position v at which repair occurs. Note that for this example, the singular controls are choices of $v \in \mathbb{V}$; in the most general case $\mathbb{V}$ could be the whole X-state space $[0, 1]$. The processes under consideration make

$$\begin{aligned}
(4.1)\qquad f(X(t),Y(t)) &- \int_0^t Af(X(s),Y(s))\,ds\\
&- \int_0^t B_0 f(X(s),Y(s))\,dL_0(s)\\
&- \int_0^t B_1 f(X(s-),Y(s-))\,dN_1(s)\\
&- \int_0^t \int_{\mathbb{V}} B_2 f(X(s-),Y(s-),v)\,\Psi_s(dv)\,dN_2(s)
\end{aligned}$$

a martingale for every $f \in C^2([0,1]\times\mathbb{Y})$, in which N_2 is the counting process which counts the number of repairs.

The cost criterion in which we are interested includes the cost of repairing or replacing the system and a cost associated with the second moment of the driving force, though with different coefficients for the different regimes so that higher levels of wear typically have higher costs. Let $c_{abs}(x,y) = c(y)x^2$ denote the running cost related to the position x, in which $c(y)$ allows for different cost rate factors for the different states of wear. Also let $c_1(x,y)$ denote the cost for replacement when the wear level is y; from the modelling, $x = 1$ when replacements occur. Finally, let $c_2(x,y,v)$ denote the cost for repairs. The objective is to minimize the long-term average cost given by

$$\begin{aligned}(4.2)\qquad \limsup_{t\to\infty} t^{-1}E\Bigg[\int_0^t & c_{abs}(X(s),Y(s))\,ds \\ &+ \int_0^t c_1(X(s-),Y(s-))\,dN_1(s) \\ &+ \int_0^t\int_{\mathbb{V}} c_2(X(s-),Y(s-),v)\Psi_s(dv)dN_2(s)\Bigg].\end{aligned}$$

For an example of a specific cost structure see Section 4.2.

4.1. LP Formulation

It is helpful to carefully define the occupation measures before displaying the LP formulation. For each $t > 0$, define the measures (on the appropriate Borel sets)

$$\begin{aligned}\mu^t(G) &= t^{-1}E\left[\int_0^t I_G(X(s),Y(s))\,ds\right],\\ \nu_0^t(\{(0,i)\}) &= t^{-1}E[I_{\{i\}}(Y(t))L_0(t)],\\ \nu_1^t(\{(1,i)\}) &= t^{-1}E[I_{\{i\}}(Y(t-))N_1(t)],\\ \nu_2^t(G) &= t^{-1}E\left[\int_0^t\int_{\mathbb{V}} I_G(X(s-),Y(s-),v)\,\Psi_s(dv)\,dN_2(s)\right].\end{aligned}$$

It is important to notice that though ν_2^t appears to be a measure on $[0,1]\times\mathbb{Y}\times\mathbb{V}$, N_2 only increases at times t such that $X(t-) \in \mathbb{V}$. As a result, ν_2^t only charges points (x,v) on the diagonal of $\mathbb{V}\times\mathbb{V}$. We can therefore simplify notation by taking ν_2^t to be a measure on $\mathbb{Y}\times\mathbb{V}$.

A similar tightness argument as in Section 2.1 implies existence of weak limits (μ,ν_0,ν_1,ν_2) of $\{(\mu^t,\nu_0^t,\nu_1^t,\nu_2^t) : t > 0\}$ as $t \to \infty$. As a result, (4.1) being a martingale implies

$$(4.3)\qquad \langle Af,\mu\rangle + \langle B_0 f,\nu_0\rangle + \langle B_1 f,\nu_1\rangle + \langle B_2 f,\nu_2\rangle = 0, \quad \forall f \in C^2([0,1]\times\mathbb{Y}).$$

Thus the equivalent infinite-dimensional LP formulation for this repair model is

$$\text{LP4}\quad\begin{cases}\text{Min.} & \langle c_{abs},\mu\rangle + \langle c_1,\nu_1\rangle + \langle c_2,\nu_2\rangle\\ \text{S.t.} & \langle Af,\mu\rangle + \langle B_0 f,\nu_0\rangle + \langle B_1 f,\nu_1\rangle + \langle B_2 f,\nu_2\rangle = 0,\\ & \qquad\qquad\qquad\qquad\qquad \forall f \in C^2([0,1]\times\mathbb{Y}),\\ & \mu \in \mathcal{P}([0,1]\times\mathbb{Y}),\\ & \nu_0 \in \mathcal{M}(\{0\}\times\mathbb{Y}),\\ & \nu_1 \in \mathcal{M}(\{1\}\times\mathbb{Y}),\\ & \nu_2 \in \mathcal{M}(\mathbb{Y}\times\mathbb{V}).\end{cases}$$

TABLE 8
Parameters for the repair model

Level	Drift	Diff.	$c_{abs}(x,y)$	$c_1(x,y)$	$c_2(y,v)$	Q		
y	$b(y)$	$\sigma(y)$	$c(y)$	$x=1$	$1 \leq v \leq 99$	0	1	2
0	0.7	0.44	1	40	10	–3	2	1
1	0.8	0.44	2	50	20	0	–1	1
2	0.9	0.44	3	100	30	0	0	0

The finite-dimensional approximation to LP4 is obtained using the test functions of the form $f(x,y) = x^n I_{\{i\}}(y)$ in (4.3), with $n = 0,\ldots,M$ and $i = 0,\ldots,l$. As before, this choice of functions results in conditions on the pseudo-moments associated with each measure (restricted to $[0,1]\times\{i\}$ for each i), and the Hausdorff polytope associated with each measure is characterized through convex coefficient weights $\lambda_i(n)$ on the corner points of the polytope.

4.2. Numerical Results

The numerical illustration in this section has three wear levels ($y = 0,1,2$) and allows the possibility of repair and/or replacement from the 99 values $x \in \mathbb{V} = \{\frac{n}{100} : n = 1,2,\ldots,99\}$. The other parameters are listed in Table 8. Since there is no absolutely continuous control for this example, it is only necessary to look at which locations x in each of state $y = 0$ and $y = 1$ repair occurs; recall that no repair is possible in state $y = 2$ so the only singular action is replacement. The masses of the measure ν_1 on $\{1\}\times\mathbb{Y}$ and the measure ν_2 on $\mathbb{Y}\times\mathbb{V}$ are displayed in Table 9.

TABLE 9
Masses of the singular measures ν_1 at $(1,y)$ and ν_2 at (y,v) where $v = \frac{n}{100}$

$\mathbb{V}\cup\{1\}$	State y		
n	$i=0$	$i=1$	$i=2$
⋮	⋮	⋮	⋮
26	0	0	0
27	0	0	0
28	0.647063	0	0
29	0	0	0
30	0	0	0
⋮	⋮	⋮	⋮
42	0	0	0
43	0	0	0
44	0	0.810149	0
45	0	0	0
46	0	0	0
⋮	⋮	⋮	⋮
98	0	0	0
99	0	0	0
100	0	0	0.526191

Notice, in particular, that repair occurs when $x = 0.28$ and the machine is "new" ($y = 0$), and at $x = 0.44$ when $y = 1$, and that the only mass when $y = 2$ is when $x = 1$ since repair is not allowed. The solution has a nice "cascading" structure in that the repair location for the higher level of wear is to the right of the wear location for the lower level, with the replacement being at the endpoint of the highest level of wear. Thus the random shocks increase the level so that the process is in a new position to the left of any place where singular control occurs. It should be noted that this structure is an artifact of the particular choice of parameters in the model; different choices of parameters lead to more complex repair policies.

References

[1] Beneš, V. E., Shepp, L. A. and Witsenhausen, H. S. (1980). Some solvable stochastic control problems. *Stochastics* **4** 39–83.

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

[3] Cho, M. J. and Stockbridge, R. H. (2002). Linear programming formulation for optimal stopping problems. *SIAM Journal of Control and Optimization* **40** 1965–1982.

[4] Decker, T. (2006). Die Charakterisierung des verallgemeinerten Dale-Polytops und ihre Verwendung in linearen Programmen zur Lösung von Austrittszeit-, Stopp- und anderen Optimierungsproblemen. Dissertation, Humboldt Universitaet, Berlin.

[5] Hausdorff, F. (1923). Momentprobleme für ein endliches Intervall. *Mathematische Zeitschrift* **16** 220–248.

[6] Helmes, K. and Röhl, S. (2008). A geometrical characterization of multidimensional Hausdorff polytopes with applications to exit time problems. *Math. Oper. Res.* **33** 315–326.

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

[8] Helmes, K. and Stockbridge, R. H. (2000). Numerical comparison of controls and verification of optimality for stochastic control problems. *Journal of Optimization Theory and Applications* **106** 107–127.

[9] Helmes, K. and Stockbridge, R. H. (2001). Numerical evaluation of resolvents and Laplace transforms of Markov processes. *Mathematical Methods of Operations Research* **53** 309–331.

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

[11] Helmes, K. and Stockbridge, R. H. (2007). Linear programming approach to the optimal stopping of stochastic processes. *Stochastics* **79** 309–335.

[12] Kaczmarek, P. (2006). Numerical analysis of a long-term average control problem. M.S. Thesis, University of Wisconsin–Milwaukee.

[13] Kaczmarek, P., Kent, S. T., Rus, G. A., Stockbridge, R. H. and Wade, B. A. (2007). Numerical solution of a long-term average control problem for singular stochastic processes. *Math. Oper. Res.* **66** 451–473.

[14] Kurtz, T. G. (1991). A control formulation for constrained Markov processes. *Mathematics of Random Media*. Lectures in Applied Mathematics **27**.

[15] Kurtz, T. G. and Stockbridge, R. H. (1998). Existence of Markov controls and characterization of optimal Markov controls. *SIAM Journal of Control and Optimization* **36** 609–653.

[16] Kurtz, T. G. and Stockbridge, R. H. (2001). Stationary solutions and forward equations for controlled and singular martingale problems. *Electronic Journal of Probability* **6** paper 14, 1–52.

[17] Manne, A. S. (1960). Linear programming and sequential decisions. *Management Science* **6** 259–267.

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

[19] Röhl. S. (2001). Ein linearer Programmierungsansatz zur Lösung von Stopp- und Steuerungsproblemen. Dissertation, Humboldt Universitaet, Berlin.

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

[21] Zhang, Q. (2001). Stock trading: An optimal selling rule. *SIAM Journal of Control and Optimization* **40** 64–87.