KYBERNETIKA — VOLUME 44 (2008), NUMBER 2, PAGES 185 – 204

NUMERICAL STUDY OF DISCRETIZATIONS OF MULTISTAGE STOCHASTIC PROGRAMS

Petri Hilli and Teemu Pennanen

This paper presents a numerical study of a deterministic discretization procedure for multistage stochastic programs where the underlying stochastic process has a continuous probability distribution. The discretization procedure is based on quasi-Monte Carlo techniques originally developed for numerical multivariate integration. The solutions of the discretized problems are evaluated by statistical bounds obtained from random sample average approximations and out-of-sample simulations. In the numerical tests, the optimal values of the discretizations as well as their first-stage solutions approach those of the original infinite-dimensional problem as the discretizations are made finer.

Keywords: stochastic programming, discretization, integration quadratures, simulation

AMS Subject Classification: 90C15, 49M25, 90C25

1. INTRODUCTION

This paper is concerned with numerical solution of multistage stochastic programs where the underlying stochastic process has a continuous distribution. These are infinite-dimensional optimization problems that come up quite naturally in practical decision making. For example, in financial applications it is customary to model uncertain future development of asset prices by stochastic processes that have a continuous parametric distribution. Despite the large number of potential applications only few papers have been devoted to numerical solution of such models; see Olsen [13], Frauendorfer [5], Pflug [17], Shapiro [21, 22], Chiralaksanakul [2], Chiralaksanakul and Morton [3], Kuhn [10], Blomvall and Shapiro [1], Pennanen and Koivu [16], Pennanen [14, 15], Kall and Mayer [9], and Heitsch and Römisch [7]. This paper presents a numerical study of the discretization techniques proposed in [14, 15, 16].

We will consider the following stochastic programming model from Rockafellar and Wets [19, 20]. At each *stage* $k = 0, \dots, K$, the decision maker observes the value of a random variable ξ_k, and makes a decision x_k depending on the observed values of $\xi_0, \dots, \xi_k$. Each ξ_k takes values in a Borel subset Ξ_k of $\mathbb{R}^{d_k}$ and x_k is $\mathbb{R}^{n_k}$-valued. We assume that Ξ_0 is a singleton, so that ξ_0 and thus x_0 will be deterministic. The

vector $\xi = (\xi_0, \dots, \xi_K)$ will be modeled as a random variable in the probability space $(\Xi, \mathcal{F}, P)$, where $\Xi = \Xi_0 \times \cdots \times \Xi_K$, $\mathcal{F}$ is the Borel σ-field on Ξ and P is a probability measure on $(\Xi, \mathcal{F})$. A realization of ξ and a sequence $x = (x_0, \dots, x_K)$ of decisions will result in a cost given by a function $f : \mathbb{R}^n \times \Xi \to \mathbb{R} \cup \{+\infty\}$, where $n = n_0 + \ldots + n_K$. We will study *multistage stochastic programs* of the form

$$\underset{x \in \mathcal{N}(P)}{\text{minimize}} \quad \mathrm{E}^P f(x(\xi), \xi), \tag{SP(P)}$$

where E^P is the expectation operator, f is a *convex normal integrand* on $\mathbb{R}^n \times \Xi$ and

$$\mathcal{N}(P) = \big\{ x \in L^\infty(\Xi, \mathcal{F}, P; \mathbb{R}^n) \,|\, x \text{ contains an } (\mathcal{F}_k)_{k=0}^K\text{-adapted function} \big\}.$$

Here $(\mathcal{F}_k)_{k=0}^K$ is the filtration of σ-fields

$$\mathcal{F}_k := \big\{ B_k \times \Xi_{k+1} \times \cdots \times \Xi_K \mid B_k \in \mathcal{B}_k \big\},$$

where $\mathcal{B}_k$ is the Borel σ-field on $\Xi_0 \times \cdots \times \Xi_k$. A function $\tilde{x} = (\tilde{x}_0, \dots, \tilde{x}_K)$ is said to be *adapted* to $(\mathcal{F}_k)_{k=0}^K$ if for each k, $\tilde{x}_k$ is $\mathcal{F}_k$-measurable, or equivalently, only depends on the part of ξ that has been observed by stage k. We would like to emphasize that the function f is allowed to take on the value $+\infty$ so that various constraints (of the "almost surely-type") can be taken into account through infinite penalties.

When the stochastic process ξ is a random variable with an infinite sample space (as in most econometric models), (SP(P)) is an infinite-dimensional optimization problem whose solution requires discretization. One way to discretize it is to approximate the original measure P by a finitely supported measure of the form

$$P^\nu = \sum_{i \in I(\nu)} p^{\nu,i} \delta_{\xi^{\nu,i}},$$

where $I(\nu)$ is a finite index set, $\delta_{\xi^{\nu,i}}$ is the unit mass at a point $\xi^{\nu,i} \in \Xi$, and $p^{\nu,i} > 0$. Then $L^\infty(\Xi, \mathcal{F}, P^\nu; \mathbb{R}^n) \cong (\mathbb{R}^n)^{I(\nu)}$, and $(SP(P^\nu))$ can be written in the finite-dimensional form

$$\underset{x \in \mathcal{N}(P^\nu)}{\text{minimize}} \quad \sum_{i \in I(\nu)} p^{\nu,i} f(x(\xi^{\nu,i}), \xi^{\nu,i}), \tag{SP(P^\nu)}$$

where

$$\begin{aligned} \mathcal{N}(P^\nu) &= \big\{ x \in L^\infty(\Xi, \mathcal{F}, P^\nu; \mathbb{R}^n) \,|\, x \text{ contains an } (\mathcal{F}_k)_{k=0}^K\text{-adapted function} \big\} \\ &= \big\{ x \in L^\infty(\Xi, \mathcal{F}, P^\nu; \mathbb{R}^n) \,|\, x_k(\xi^{\nu,i}) = x_k(\xi^{\nu,j}) \;\text{ if }\; \xi_l^{\nu,i} = \xi_l^{\nu,j} \;\forall l = 0, \dots, k \big\}. \end{aligned}$$

This is a mathematical program which can in principle be solved numerically by general purpose solvers or special purpose algorithms designed to take advantage of problem structure.

A few methods for constructing approximations P^ν of P for purposes of multistage stochastic programming have been proposed. The best known, most widely studied and probably the most widely applied method is conditional sampling, where P^ν

is constructed by recursively sampling from the conditional distribution of ξ_k given $\xi_0, \ldots, \xi_{k-1}$; see [1, 2, 3, 21, 22]. In barycentric approximation, P^ν is constructed so that, under certain convexity properties with respect to the random variables, one obtains lower/upper approximations of the original problem; see [5, 10]. In the methods of [17] and [7], the general idea is to choose P^ν so as to approximate the original measure P as well as possible, in the sense of a metric in the space of probability measures. The same general idea is behind *quasi-Monte Carlo* (QMC) methods, which are a class of integration quadratures; see e. g. [11]. The main difference is in the choice of the metric and in that in QMC methods the original measure is the uniform distribution in the unit cube. For the metric used in QMC, there exist many techniques for constructing discrete measures that are close to the original one. To apply QMC methods, one has to write a given integral in terms of the uniform distribution, which can often be done by an appropriate change of variables.

We would like to emphasize that discretizations obtained with conditional sampling are *random* whereas those obtained with the methods of [5, 7, 10, 17] or with QMC are *deterministic*. This is a fundamental difference in that the first-stage solution obtained through deterministic discretization can be viewed as a *decision rule* that is fully determined by the original infinite dimensional problem and the chosen discretization procedure; see Section 3.2. This way, a deterministic discretization procedure gives a well-defined candidate solution for the original multistage stochastic program that can be evaluated e. g. by out-of-sample simulations.

QMC was first proposed for discretization of multistage stochastic programs in [16]. It was shown in [14, 15] that, under fairly general conditions, QMC produces consistent approximations in the sense that the optimal values of the approximations converge to that of the original problem and that the cluster points of the first stage solutions are optimal first stage solutions of the original problem. However, the results of [14, 15] are nonquantitative and, in particular, say nothing about the accuracy of $(SP(P^\nu))$ for a fixed ν. The purpose of this paper is to present a numerical study of the discretization procedures studied in [14, 15, 16]. We solve a sequence of refined discretizations numerically and compute statistical bounds for the optimal value of the original infinite-dimensional problem. The lower bound is obtained as a sample average of optimal values of random discretizations whereas the upper bound is obtained through out-of-sample evaluation of the strategies obtained through QMC discretizations. The results support the theoretical findings of [14, 15] namely that the optimal values of the discretizations converge to that of the original problem and that the cluster points of first stage solutions are optimal first stage solutions of the original problem.

The rest of this paper is organized as follows. Section 2 recalls the discretization procedure from [15, 16]. Section 3 describes the statistical lower and upper bounds used in the numerical tests. Section 4 describes the test problems and summarizes the computational results. The parameter values used in the test problems are given in the Appendix.

2. DISCRETIZATIONS

As in [15, 16], we assume that the underlying stochastic process ξ is driven by a time series model of the form

$$\xi_k = g_k(\xi_0, \dots, \xi_{k-1}, \omega_k) \quad \text{for } k = 1, \dots, K, \tag{1}$$

where ξ_0 is given, $\omega_1, \dots, \omega_K$ are mutually independent random variables, with ω_k uniformly distributed in the d_k-dimensional unit cube $(0,1)^{d_k}$, and $g_k : \Xi_0 \times \cdots \times \Xi_{k-1} \times (0,1)^{d_k} \to \Xi_k$ are Borel-measurable functions. It follows that $\xi = (\xi_0, \dots, \xi_K)$ is uniquely determined by $\omega = (\omega_1, \dots, \omega_K)$ so that

$$P = UG^{-1}, \tag{2}$$

where G denotes the mapping that sends ω to ξ and $U = U_1 \times \cdots \times U_K$, where U_k is the uniform distribution on $(0,1)^{d_k}$. Notation (2) means that $P(A) = U(G^{-1}(A))$ for every $A \in \mathcal{F}$, or equivalently, $\mathrm{E}^P \varphi(\xi) = \mathrm{E}^U \varphi(G(\omega))$ for any measurable function φ on Ξ.

Expression (2) suggests the following

Discretization procedure

1. For each $k = 1, \dots, K$, approximate U_k by a discrete measure U_k^ν;
2. Let $U^\nu = U_1^\nu \times \cdots \times U_K^\nu$ and $P^\nu = U^\nu G^{-1}$.

More concretely, if for $k = 1, \dots, K$

$$U_k^\nu = \sum_{i \in I_k(\nu)} p_k^{\nu,i} \delta_{\omega_k^{\nu,i}},$$

where $I_k(\nu)$ is a finite index set, then

$$U^\nu = \sum_{i \in I(\nu)} p^{\nu,i} \delta_{\omega^{\nu,i}},$$

where

$$\begin{aligned} I(\nu) &= \big\{(i_1, \dots, i_K) \mid i_k \in I_k(\nu)\big\}, \\ \omega^{\nu,i} &= \big(\omega_1^{\nu,i_1}, \dots, \omega_K^{\nu,i_K}\big), \\ p^{\nu,i} &= p_1^{\nu,i_1} \cdots p_K^{\nu,i_K}, \end{aligned}$$

and $P^\nu = U^\nu G^{-1}$ becomes

$$P^\nu = \sum_{i \in I(\nu)} p^{\nu,i} \delta_{\xi^{\nu,i}},$$

where $\xi^{\nu,i} = G(\omega^{\nu,i})$.

When the discrete measures U_k^ν are the empirical measures corresponding to a random sample of size $|I_k(\nu)|$, the above procedure is nothing but the well-known

conditional sampling procedure. It was shown in [14, 15] that if each measure U_k^ν converges weakly to the uniform distribution, as $\nu \to \infty$, then under rather mild conditions, problems $(SP(P^\nu))$ will be consistent discretizations of $(SP(P))$ in the sense that the optimal values of $(SP(P^\nu))$ converge to that of $(SP(P))$ and all cluster points of first-stage solutions of $(SP(P^\nu))$ are optimal first-stage solutions of $(SP(P))$.

The nontrivial parts in implementing the above procedure are the construction of the discrete measures U_k^ν that approximate the uniform distribution and the evaluation of $G(\omega^{\nu,i})$ for given $\omega^{\nu,i}$, that is, the evaluation of the functions (1) for $k = 1, \ldots, K$. As for U_k^ν, we will use *quasi-Monte Carlo* (QMC) methods, as proposed in [16]. In QMC, one sets $p_k^{\nu,i} = 1/|I_k(\nu)|$ and the points $\omega_k^{\nu,i}$ are chosen so as to minimize certain measure of uniformity. There exists a wide literature on such methods; see for example Niederreiter [11], Sloan and Joe [23] or Niederreiter and Talay [12].

Difficulty of evaluating $G(\omega^{\nu,i})$ depends on the form of the time series model at hand. For example, in (multivariate) ARMA, VEqC or GARCH models with Gaussian innovations this part is quite simple. For them, evaluation of g_k comes down to evaluating the inverse of the univariate Gaussian distribution function at the components of the the vectors ω_k^{ν,i_k}, forming the Cholesky factorization of a covariance matrix and performing few matrix-vector multiplications. These operations can be executed quite efficiently by publicly available subroutines.

3. STATISTICAL BOUNDS FOR THE OPTIMAL VALUE

3.1. Lower bounds

It is well-known that if one makes random discretizations by conditional sampling (i. e. taking U_k^ν in the above discretization procedure to be empirical measures corresponding to random samples), the corresponding optimal value v^ν of $(SP(P^\nu))$ is a random number whose expectation is less than that of $(SP(P))$; see Shapiro [21], Chiralaksanakul [2, Theorem 4] or Chiralaksanakul and Morton [3]. Moreover, v^ν converges almost surely to the optimal value of $(SP(P))$ as $\nu \to \infty$ provided the sample sizes $|I_k(\nu)|$ tend to infinity. These properties remain valid if instead of crude Monte Carlo sampling one uses non-iid sampling techniques such as Latin hyper cube sampling or antithetic sampling; see for example Glasserman [6, Chapter 4]. Such techniques can result in more efficient estimates in direct numerical integration and possibly tighter lower bounds in random discretization of stochastic programs.

One can estimate the expectation of v^ν by constructing N independent random discretizations and taking the average over the corresponding optimal values. The resulting average v_N^ν provides a statistical lower bound for the optimal value of $(SP(P))$. To our knowledge, the only numerical study of such estimates in case of multistage stochastic programs with continuous distributions has been reported in [1].

3.2. Upper bounds

For any $x \in \mathcal{N}(P)$ the expectation $\mathrm{E}^P f(x(\xi), \xi)$ is greater than or equal to the optimal value of $(SP(P))$. The expectation can be approximated by general techniques of numerical integration. Using Monte Carlo sampling one obtains random approximations whose expectations equal $\mathrm{E}^P f(x(\xi), \xi)$. Besides being a statistical upper bound for the optimal value of $(SP(P))$ the sample average gives also an unbiased estimator of the quality of the chosen strategy $x \in \mathcal{N}(P)$. Such a technique for evaluating a decision strategies is often called *out-of-sample* testing.

We propose to use strategies given by recursive use of the discretization procedure described in the Section 2. Other, computationally less demanding, methods for generating policies that can be used in computation of upper bounds have been proposed in [2, Section 3.3], [3] and [1].

A discrete measure $U^\nu = U_1^\nu \times \cdots \times U_K^\nu$ chosen in the discretization procedure of Section 2 gives rise to an adapted strategy $x^\nu \in \mathcal{N}(P)$ as follows. Given a scenario $\xi^i = (\xi_0^i, \ldots, \xi_K^i)$, define $x^\nu(\xi^i) = (x_0^\nu(\xi^i), \ldots, x_K^\nu(\xi^i))$ recursively for $k = 0, \ldots, K$ by letting $x_k^\nu(\xi^i)$ be the first-stage solution of the discretization of the multistage stochastic program

$$\underset{x \in \mathcal{N}_k(P_k)}{\text{minimize}} \quad \mathrm{E}^{P_k} f_k(x(\xi), \xi), \tag{$SP_k(P_k)$}$$

where P_k denotes the conditional probability distribution of $(\xi_{k+1}, \ldots, \xi_K)$ given $(\xi_0, \ldots, \xi_k) = (\xi_0^i, \ldots, \xi_k^i)$ and f_k is the normal integrand obtained from f by fixing the values of $\xi_0, \ldots, \xi_k$ and $x_0, \ldots, x_{k-1}$ to $\xi_0^i, \ldots, \xi_k^i$ and $x_0^i, \ldots, x_{k-1}^i$, respectively. Here, we discretize $(SP_k(P_k))$ with the procedure of Section 2 using the discrete measure $U_{k+1}^\nu \times \cdots \times U_K^\nu$. Note that the conditional distribution P_k is determined simply by fixing the values of $\xi_0, \ldots, \xi_k$ to $\xi_0^i, \ldots, \xi_k^i$ in (1). In this way, each discretization (choice of a QMC technique and the number of quadrature points per stage) gives rise to an adapted strategy $x^\nu \in \mathcal{N}(P)$ that can be evaluated numerically along a given scenario by solving a sequence of discretized multistage stochastic programs $(SP_k(P_k))$, $k = 0, \ldots, K$.

The out-of-sample test proceeds by sampling M scenarios $(\xi^i)_{i=1}^M$ of the stochastic process ξ. Then along each scenario $\xi^i = (\xi_0^i, \ldots, \xi_K^i)$, one solves discretizations of problems $(SP_k(P_k))$ for $k = 0, \ldots, K$ and records the value of $f(x^\nu(\xi^i), \xi^i)$. Again, one can use variance reduction techniques instead of crude Monte Carlo when constructing the scenarios $(\xi^i)_{i=1}^M$.

4. NUMERICAL TESTS

The discretization procedure of Section 2 was applied to four different multistage stochastic programs where the stochastic process ξ was driven by time series models of AR-, VEqC- and GARCH-type. Each of the four problems was discretized by choosing the discrete measures U_k^ν as empirical measures corresponding to quasi-Monte Carlo methods with equal number $|I_k(\nu)| = L$ of quadrature points for each period $k = 0, \ldots, K$. For a fixed L, we then solved the discretization, recorded the optimal value and computed lower and upper bounds. The lower bound was computed by making $N = 1000$ random discretizations with Monte Carlo with L

points per period. The upper bound was computed by making $M = 10000$ out-of-sample simulations with Monte Carlo and using quasi-Monte Carlo with L points per stage when solving the problems ($SP_k(P_k)$). We then increased the number L of quadrature points per stage and repeated the computations. The quasi-Monte Carlo method used in the experiments was the Sobol sequence [24]. We computed also a second set of lower and upper bounds by replacing Monte Carlo by antithetic sampling. In antithetic sampling, one half of the sample points are independently uniformly distributed in the unit cube and the second half is obtained by reflecting each sample point through the center of the cube.

The computations were performed on Red Hat Enterprise Linux WS v.4 operating system running on workstation with 64-bit 3.8GHz Intel Xeon processor and 8GB memory. The discretization procedure was implemented in C programming language using CBLAS and LAPACK libraries, Numerical recipes C-routines [18], Marsenne twister (www.math.sci.hiroshima-u.ac.jp/∼m-mat/MT/emt.html) for random number generation and Peter J. Acklam's algorithm for the inverse of the univariate normal distribution (home.online.no/∼pjacklam/notes/invnorm). The implementation applies to multivariate time series models of ARMA-, VEqC- and GARCH-type. It takes as input the parameters of the time series model, the initial values ξ_0 of the time series, the number of periods K and the number L of quadrature points used for each $k = 1, \ldots, K$ in the discretization U_k^ν of the uniform distribution U_k. The discretizations U_k^ν can be constructed either by crude Monte Carlo sampling, antithetic sampling or by quasi-Monte Carlo methods. The output is a description in an AMPL-format (see [4]) of the discretized measure P^ν in terms of the scenarios $\xi^{\nu,i}$ and the associated probabilities $p^{\nu,i}$. The discretized optimization problems were formulated in AMPL and solved with the primal-dual interior point solver of MOSEK (http://www.mosek.com). Instead of standard software, one could use (as e. g. in [1]) specialized software that employs the structure of tree based discretizations.

4.1. Problem 1: Swing option

The first test problem can be written as

$$\begin{aligned} \underset{x \in \mathcal{N}(P)}{\text{minimize}} \quad & \mathrm{E}^P \exp\left(-\rho \sum_{k=0}^{K} (S_k(\xi) - X) x_k(\xi)\right) \\ \text{subject to} \quad & \sum_{k=0}^{K} x_k(\xi) \le U, \\ & l \le x_k(\xi) \le u, \quad k = 0, \ldots, K, \\ & P\text{-a.s.} \end{aligned}$$

where ρ, X, U, l and u are positive constants and S is a real-valued stochastic price process. This models the problem of finding an optimal exercise strategy x for a "swing option" that gives the access to a total amount U of energy for a fixed unit price X over the life time $[0, K]$ of the option but restricts the usage x_k per period to lie in the interval $[l, u]$. It is assumed that at each stage, x_k will be immediately sold for the current market price S_k thus giving the revenue of $(S_k - X)x_k$. The

Table 1. Dimensions of decision and random variables in Problem 1. The numbers n_k and d_k give the dimensions of the decision and random variables, respectively, for period k.

k	0	1	2	3	4
n_k	1	1	1	1	1
d_k	0	1	1	1	1

Table 2. Dimensions and computation times (seconds) for discretizations of Problem 1.

L	Variables	Constraints	Discretization	AMPL	MOSEK
2	31	16	0.0	0.0	0.0
4	341	256	0.0	0.0	0.1
8	4681	4096	0.0	0.2	1.0
16	69905	65536	0.5	2.9	23.0
32	1082401	1048576	7.6	34.0	508.4

objective is to maximize the expected utility from cumulated wealth at the terminal stage K as measured by the exponential utility function with parameter ρ.

The price process S follows a geometric Brownian motion. That is, $S_k(\xi) = \exp(\xi_k)$, where $\xi = (\xi_k)_{k=0}^K$ follows a discrete time Brownian motion

$$\Delta\xi_k = \mu + \sigma\varepsilon_k,$$

where ε_k has standard normal distribution. All the parameters of the model are given in the Appendix.

In the numerical test, the number of stages was $K = 4$. Table 1 summarizes the dimensions of Problem 1. The number n_k is the dimension of the decision variable at stage k and d_k is the dimension of the random variable in period k (the one between stages $k-1$ and k). Table 2 gives the dimensions of its discretizations as well as computation times for increasing values of the discretization parameter L.

Figure 1 plots the optimal value of the discretizations together with the upper and lower bounds for increasing values of the discretization parameter L. As L is increased, the lower and upper bounds as well as the optimal value of the discretizations seem to converge towards a common value. This is in line with the conclusions of [15, Theorem 5], which says that the optimal values of discretizations converge to that of the original problem and that the cluster points of the first stage solutions are optimal first stage solutions of the original problem. The latter fact is reflected in the upper bound, which decreases as the discretizations are made finer. Indeed, the upper bound is a non-biased estimator of the objective value given by the strategy obtained by solving discretized problems stage by stage. Refining the discretization thus seems to yield better first stage solutions.

Evaluation of the upper bound with $M = 10000$ independent simulations and

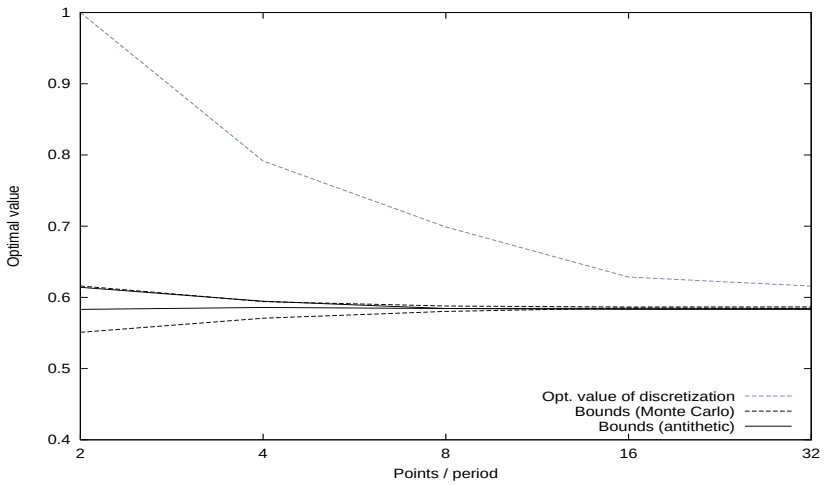

Fig. 1. Optimal values of discretizations and statistical bounds for Problem 1.

$L = 32$ quadrature points per period took approximately 31 hours. Evaluation of the lower bound with $N = 1000$ independent random discretizations and $L = 32$ quadrature points per period took approximately 152 hours.

The lower bound obtained with antithetic sampling is notably tighter than the one obtained with crude Monte Carlo. In the case of the upper bound, the difference is insignificant.

4.2. Problem 2: Portfolio optimization

The second test problem may be written as

$$\begin{aligned} \underset{h \in \mathcal{N}(P)}{\text{minimize}} \quad & \mathrm{E}^P \exp\left(-\rho \sum_{j \in J} R_{K,j} h_{K-1,j}\right) \\ \text{subject to} \quad & \sum_{j \in J} h_{0,j} \leq 1, \\ & \sum_{j \in J} h_{k,j} \leq \sum_{j \in J} R_{k,j} h_{k-1,j}, \quad k = 1, \dots K, \\ & h \geq 0, \\ & P\text{-a.s.} \end{aligned}$$

This models a portfolio optimization problem, where one is seeking for a nonnegative portfolio process $h = (h_k)_{k=0}^K$, where $h_k = (h_{k,j})_{j \in J}$ gives the amount of wealth invested in each asset $j \in J$ at the beginning of period k. The random number $R_{k,j}$ gives the return on an asset $j \in J$ over period k.

Table 3. Dimensions of decision and random variables in Problem 2.

k	0	1	2	3
n_k	10	10	10	0
d_k	0	10	10	10

Table 4. Dimensions and computation times (seconds for discretizations of Problem 2.

L	Variables	Constraints	Discretization	AMPL	MOSEK
2	70	7	0.0	0.0	0.0
4	210	21	0.0	0.0	0.1
8	730	73	0.0	0.0	0.2
16	2730	273	0.3	0.3	0.6
32	10570	1057	1.9	2.5	4.1
64	41610	4161	15.0	16.4	27.2

The set J contains 10 assets and the returns are given by $R_{k,j} = S_{k,j}/S_{k-1,j}$ where the price vector $S_k = (S_{k,j})_{j\in J}$ follows a 10-dimensional geometric Brownian motion. That is, $S_{k,j} = \exp(\xi_{k,j})$ where $\xi_k = (\xi_{k,j})_{j\in J}$ follows a 10-dimensional Brownian motion

$$\Delta\xi_k = \mu + \sigma\varepsilon_k,$$

where ε_k has 10-dimensional standard normal distribution. All the parameters of the model are given in the Appendix.

In the numerical test, the number of stages was $K = 3$. Table 3 summarizes the dimensions of Problem 2 and Table 4 gives the dimensions of its discretizations as well as computation times for increasing values of the discretization parameter L.

Figure 2 plots the optimal value of the discretizations together with the upper and lower bounds for increasing values of the discretization parameter L. Again, the lower and upper bounds converge towards each other while the optimal values of the discretizations approach the common value as the discretizations are made finer. The lower bound obtained with antithetic sampling is again much tighter than the one obtained with crude Monte Carlo.

Evaluation of the upper bound with $M = 10000$ independent simulations and $L = 64$ quadrature points per period took approximately 14 hours. Evaluation of the lower bound with $N = 1000$ independent random discretizations and $L = 64$ quadrature points per period took approximately 16 hours.

4.3. Problem 3: Optimal consumption

The third problem is a variation of the portfolio optimization problem where at each stage, one can withdraw some of the wealth for consumption. The goal is to maximize accumulated utility from consumption as measured by the negative power

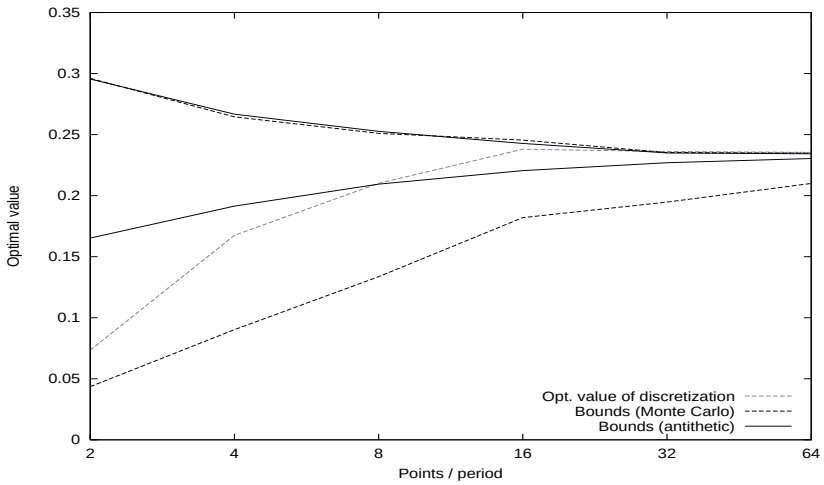

Fig. 2. Optimal values of discretizations and statistical bounds for Problem 2.

utility function. The problem is further complicated by adding transaction costs for buying and selling the assets. It can be written

$$
\begin{aligned}
\underset{(h,p,s,c)\in\mathcal{N}(P)}{\text{minimize}} \quad & \mathrm{E}^P \sum_{k=0}^{K} (c_k - \rho_1)^{-\rho_2} \\
\text{subject to} \quad & h_{0,j} \le h_j^0 + p_{0,j} - s_{0,j}, \\
& h_{k,j} \le R_{k,j} h_{k-1,j} + p_{k,j} - s_{k,j}, \quad k = 1, \ldots, K, \\
& c_k \le \sum_{j\in J} (1-\delta_j) s_{k,j} - \sum_{j\in J} (1+\delta_j) p_{k,j}, \quad k = 0, \ldots, K-1, \\
& c_K \le \sum_{j\in J} R_{k,j} h_{k-1,j}, \\
& p, s, h, c \ge 0, \\
& P\text{-a.s.}
\end{aligned}
$$

where h and R are as in the previous problem, parameter h_j^0 gives initial portfolio, $p_{k,j}$ and $s_{k,j}$ give the purchases and sales, respectively, of asset $j \in J$ at the beginning of period t, c_k gives the consumption over period t. In this problem, the decision strategy is the vector process consisting of h, p, s and c. The parameters δ_j give the proportional transaction costs for purchases and sales.

The set J contains three assets, two stock indices $SP500$ and NAS and a money market account $E3$. The returns can be expressed as

$$
R_{k,j} = \begin{cases} S_{k,j}/S_{k-1,j} & \text{if } j \in \{SP500, NAS\} \\ (1 + r_k/100)^{\frac{1}{4}} & \text{if } j = E3, \end{cases}
$$

Table 5. Dimensions of decision and random variables in Problem 3.

k	0	1	2	3
n_k	10	10	10	1
d_k	0	3	3	3

Table 6. Dimensions and computation times (seconds) for discretizations of Problem 3.

L	Variables	Constraints	Discretization	AMPL	MOSEK
2	78	36	0.0	0.0	0.0
4	274	148	0.0	0.0	0.0
8	1242	804	0.0	0.0	0.1
16	6826	5188	0.1	0.2	0.9
32	43338	36996	0.6	1.9	8.9
64	303754	278788	4.5	11.9	97.8

where S_j denotes the value of a stock index and r is a three-month interest rate. The vector of logarithms

$$\xi_k = (\ln S_{k,SP500}, \ln S_{k,NAS}, \ln r_k)$$

follows a VEqC-GARCH process

$$\Delta\xi_k = \mu_k + \sigma_k \varepsilon_k,$$

where

$$\mu_k = A\Delta\xi_{k-1} + \alpha(\beta^T \xi_{k-1} - \gamma) + c$$

and

$$\sigma_k^2 = C(\Delta\xi_{k-1} - \mu_{k-1})(\Delta\xi_{k-1} - \mu_{k-1})^T C^T + D\sigma_{k-l}^2 D^T + \Omega.$$

All the parameters of the model are given in the Appendix.

In the numerical test, the number of stages was $K = 4$. Table 5 summarizes the dimensions of Problem 3 and Table 6 gives the dimensions of its discretizations as well as computation times for increasing values of the discretization parameter L.

Figure 3 plots the optimal value of the discretizations together with the upper and lower bounds for increasing values of the discretization parameter L. Similar conclusions as for the earlier test problems apply except that here the improvement of the lower bound with antithetic sampling over crude Monte Carlo is even more striking.

Evaluation of the upper bound with $M = 10000$ independent simulations and $L = 64$ quadrature points per period took approximately 11 hours. Evaluation of the lower bound with $N = 1000$ independent random discretizations and $L = 64$ quadrature points per period took approximately 32 hours.

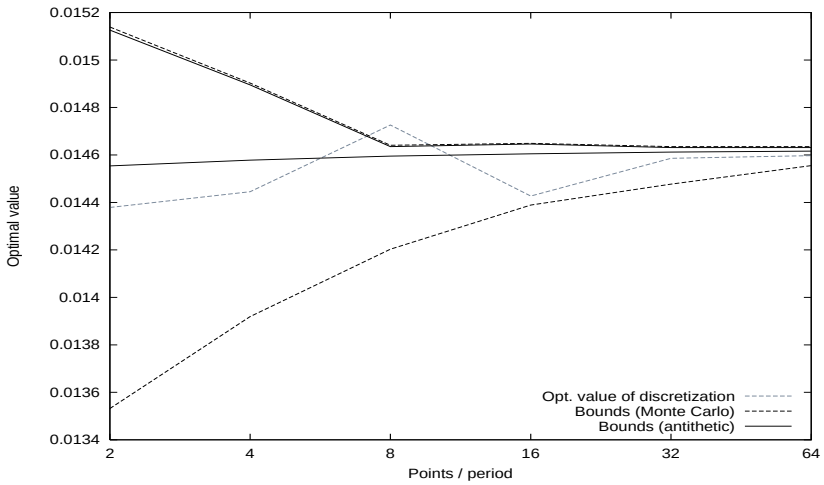

Fig. 3. Optimal values of discretizations and statistical bounds for Problem 3.

4.4. Problem 4: Asset liability management

The fourth problem is an asset liability management model developed originally for a Finnish pension fund; see [8]. Here the goal is to find an optimal portfolio strategy over a set $J = \{1, 2, 3, 4, 5\}$ of assets, with respect to nonlinear trading costs, portfolio constraints and certain solvency targets. The problem can be written as

$$
\begin{aligned}
\underset{(h,p,w,z)\in\mathcal{N}(P)}{\text{minimize}} \quad & \mathrm{E}^P\left[\sum_{k=0}^{K-1}\frac{1}{V_k}\left(w_k-\rho_1 z_{k,1}-\rho_2 z_{k,2}\right)+\frac{1}{V_K}\left(w_K-\rho_1 z_{K,1}\right)\right] \\
\text{subject to} \quad & h_{0,j} \le h_j^0 + p_{0,j}, \quad j\in J, \\
& h_{k,j} \le R_{k,j}h_{k-1,j} + p_{k,j}, \quad j\in J, \quad k=1,\dots,K \\
& C(p_k) \le \sum_{j\in J} D_{k,j}h_{k-1,j} + F_k, \quad k=0,\dots,K \\
& w_k \le \sum_{j\in J} h_j, \quad k=0,\dots,K-1 \\
& l_j w_k \le h_{k,j} \le u_j w_k, \quad j\in J, \quad k=0,\dots,K-1 \\
& z_{k,1} \ge V_k - w_k, \quad k=0,\dots,K \\
& z_{k,2} \ge V_k - w_k - m\cdot h_k + (h_k\cdot \Sigma h_k+\epsilon)^{\frac{1}{2}}, \quad k=1,\dots,K-1 \\
& h, z \ge 0, \\
& P\text{-a.s.}
\end{aligned}
\tag{3}
$$

where h and p are as in the previous problem, except that here p is allowed to be negative which corresponds to sales. The function

$$C(p) := p_1 + \sum_{j \in J \setminus \{1\}} \frac{\exp(\delta_j p_j) - 1}{\delta_j}$$

gives the total cost of purchases. Inequalities (3) describe proportional portfolio constraints. The following two constraints define "short fall" variables $z_{k,1}$ and $z_{k,2}$ that describe how much the total wealth w is short of target levels that depend on the value V of liabilities. The second target level depends also on the portfolio h according to a nonlinear formula specified by the supervisors in the Finnish pension scheme. The value of the liabilities V as well as the cash-flow F are stochastic processes whose development is described below. In this problem, the decision strategy is the vector process consisting of h, p, w, z_1 and z_2.

The stochastic factors R, D, F and V are driven by a quarterly time series model where as the decision stages occur every four quarters. The return variables are given by

$$R_{k,j} = \begin{cases} \prod_{t=t_{k-1}+1}^{t_k} (1 + r_t/100)^{\frac{1}{4}} & \text{if } j = 1, \\ \left(\frac{1+b_{t_k}/100}{1+b_{t_{k-1}}/100} \right)^{-5} & \text{if } j = 2, \\ S_{t_k,j}/S_{t_{k-1},j} & j \in \{3,4,5\}, \end{cases}$$

and the dividend variables by

$$D_k = \begin{cases} \sum_{t=t_{k-1}+1}^{t_k} b_t/400 & \text{if } j = 2, \\ \sum_{t=t_{k-1}+1}^{t_k} (S_{t_k,6}/S_{t_k,5} - 3)/400 & \text{if } j = 5, \\ 0 & j \in \{1,3,4\}. \end{cases}$$

Here $t_k := 4k$ gives the date of stage k in quarters. The values of F_0 and V_0 are fixed parameters whereas for $k = 1, \ldots, K$

$$F_k = q \sum_{t=t_{k-1}+1}^{t_k} W_t - e_k,$$
$$V_k = (1 + \bar{r})V_{k-1} + F_k,$$

where q, $\bar{r}$ and $(e_k)_{k=0}^K$ are parameters. The 8-dimensional vector

$$\xi_t = \begin{bmatrix} \ln r_t - \Delta \ln I_t \\ \ln b_t - \Delta \ln I_t \\ \ln S_{t,3} - \ln I_t \\ \ln S_{t,4} - \ln I_t \\ \ln S_{t,5} - \ln I_t \\ \ln S_{t,6} - \ln I_t \\ \ln W_t - \ln I_t \\ \ln I_t \end{bmatrix}$$

Table 7. Dimensions of decision and random variables in Problem 4.

k	0	1	2	3
n_k	13	13	13	2
d_k	0	32	32	32

Table 8. Dimensions and computation times (seconds) for discretizations of Problem 4.

L	Variables	Constraints	Discretization	AMPL	MOSEK
2	122	106	0.0	0.0	0.0
4	486	400	0.0	0.0	0.1
8	2558	1972	0.1	0.2	0.4
16	16110	11740	0.6	1.3	2.5
32	113102	79276	4.6	14.0	19.5
64	844686	578380	36.1	81.2	174.3

follows a VEqC-model $$\Delta\xi_t = \mu_t + \sigma\varepsilon_t,$$

where ε_t are iid standard normal,

$$\mu_t = A\Delta\xi_{t-1} + \alpha(\beta^T\xi_{t-1} - \gamma) + c$$

and the volatility matrix $\sigma \in \mathbb{R}^{8\times 8}$ is constant. All the parameters of the model are given in the Appendix.

In the numerical test, the number of stages was $K = 4$. Table 7 summarizes the dimensions of Problem 4 and Table 8 gives the dimensions of its discretizations as well as computation times for increasing values of the discretization parameter L.

Figure 4 plots the optimal value of the discretizations together with the upper and lower bounds for increasing values of the discretization parameter L. Again, one gets convergence of the bounds and the optimal values of the discretizations with the lower bound obtained with antithetic sampling being much tighter than the one obtained with crude Monte Carlo.

Evaluation of the upper bound with $M = 10000$ independent simulations and $L = 64$ quadrature points per period took approximately 74 hours. Evaluation of the lower bound with $N = 1000$ independent random discretizations and $L = 64$ quadrature points per period took approximately 81 hours.

APPENDIX: PARAMETERS FOR THE TEST PROBLEMS

Problem 1. Parameters for the objective and constraints are as follows

Parameter	ρ	X	U	l	u
Value	1	1	2	0	1

.

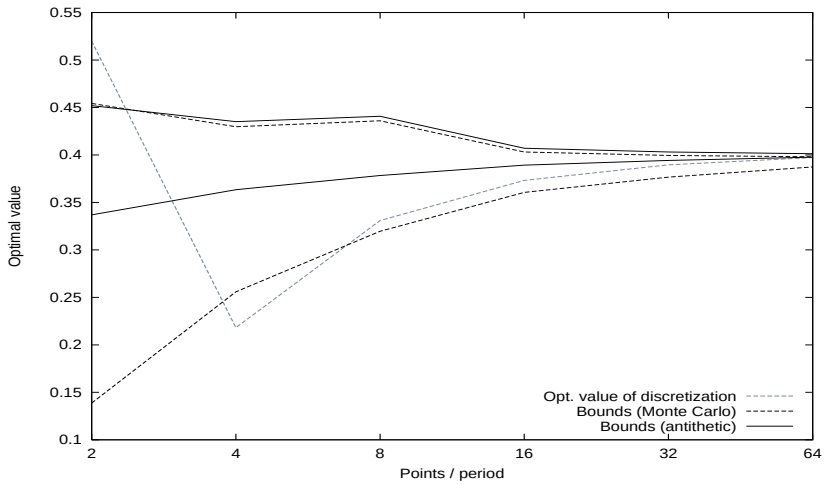

Fig. 4. Optimal values of discretizations and statistical bounds for Problem 4.

Parameters for the time series model are

$$\xi_0 = \ln 1,$$
$$\mu = 0,$$
$$\sigma^2 = 0.2.$$

Problem 2. The parameter in the objective is

$$\rho = 1.$$

Parameters for the time series model are

$$\xi_0 = \begin{pmatrix} 4.16 & 3.81 & 3.54 & 4.15 & 3.30 & 2.99 & 3.44 & 4.42 & 4.24 & 2.96 \end{pmatrix}^T,$$

$$\mu = \begin{pmatrix} 0.084 & 0.071 & 0.041 & 0.087 & 0.071 & 0.09 & 0.089 & 0.080 & 0.083 & 0.072 \end{pmatrix}^T,$$

$$\sigma^2 = 10^{-2} \begin{pmatrix} 12.90 & 3.96 & 1.01 & 5.17 & 3.71 & 3.23 & 4.41 & 1.05 & 5.64 & 4.48 \\ 3.96 & 11.83 & 0.73 & 3.53 & 2.98 & 3.28 & 3.56 & 1.29 & 4.66 & 4.68 \\ 1.01 & 0.73 & 4.92 & 0.70 & 0.75 & 0.56 & 0.67 & 0.21 & 1.47 & 0.55 \\ 5.17 & 3.53 & 0.70 & 19.15 & 4.07 & 2.40 & 4.51 & 1.39 & 6.36 & 4.38 \\ 3.71 & 2.98 & 0.75 & 4.07 & 11.25 & 4.88 & 3.23 & 1.22 & 4.59 & 5.00 \\ 3.23 & 3.28 & 0.56 & 2.40 & 4.88 & 18.06 & 3.10 & 1.43 & 4.91 & 7.30 \\ 4.41 & 3.56 & 0.67 & 4.51 & 3.23 & 3.10 & 12.22 & 1.16 & 3.85 & 3.88 \\ 1.05 & 1.29 & 0.21 & 1.39 & 1.22 & 1.43 & 1.16 & 9.70 & 1.42 & 1.78 \\ 5.64 & 4.66 & 1.47 & 6.36 & 4.59 & 4.91 & 3.85 & 1.42 & 20.60 & 6.81 \\ 4.48 & 4.68 & 0.55 & 4.38 & 5.00 & 7.30 & 3.88 & 1.78 & 6.81 & 18.77 \end{pmatrix}.$$

Problem 3. Parameters for the objective and constraints are

Parameter	ρ_1	ρ_2	δ_{SP500}	δ_{NAS}	δ_{E3}	h^0_{SP500}	h^0_{NAS}	h^0_{E3}
Value	-1	10	0.005	0.005	0	1	1	1

Parameters for the time series model are

$$\xi_{-1} = \begin{pmatrix} 7.016233 & 7.547945 & 0.582216 \end{pmatrix}^T,$$

$$\xi_0 = \begin{pmatrix} 7.030150 & 7.588319 & 0.765468 \end{pmatrix}^T,$$

$$\mu_0 = \begin{pmatrix} 0.005000 & 0.007622 & 0.041185 \end{pmatrix}^T,$$

$$A = 10^{-2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9.5920 & 0 \\ 0 & 7.7350 & 55.0770 \end{pmatrix}$$

$$\alpha = \begin{pmatrix} 0 & 0 & -0.005770 \end{pmatrix}^T, \qquad \beta = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T, \quad \gamma = \begin{pmatrix} 0 & 0 & 1.300000 \end{pmatrix}^T,$$

$$c = \begin{pmatrix} 0.005000 & 0.004520 & -0.000387 \end{pmatrix}^T,$$

$$C = 10^{-2} \begin{pmatrix} 37.9891 & 0 & 0 \\ 0 & 33.0872 & 0 \\ 0 & 0 & 50.2267 \end{pmatrix},$$

$$D = 10^{-2} \begin{pmatrix} 86.7300 & 0 & 0 \\ 0 & 91.9705 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\Omega = 10^{-2} \begin{pmatrix} 0.0200 & 0.0188 & 0.0165 \\ 0.0188 & 0.0206 & 0.0373 \\ 0.0165 & 0.0373 & 0.1782 \end{pmatrix},$$

$$\sigma_0^2 = 10^{-2} \begin{pmatrix} 0.1050 & 0.1467 & 0.0163 \\ 0.1467 & 0.2797 & 0.0368 \\ 0.0163 & 0.0368 & 0.1783 \end{pmatrix}.$$

Problem 4. Parameters for the objective and constraints are

Parameter	ρ_1	ρ_2	h^0_1	h^0_2	h^0_3	h^0_4	h^0_5	δ_2	δ_3	δ_4	δ_5
Value	10	2	20	65	40	60	50	0.0002	0.002	0.002	0.02

Parameter	l_1	l_2	l_3	l_4	l_5	u_1	u_2	u_3	u_4	u_5
Value	0.01	0	0	0	0	1	1	0.5	1	0.45

$$m = \begin{pmatrix} -0.001944 & -0.011664 & -0.120528 & -0.120528 & -0.071928 \end{pmatrix}^T,$$

$$\Sigma = 10^{-3} \begin{pmatrix} 1.270210 & -0.444573 & -2.718249 & -2.718249 & 0.000000 \\ 0.444573 & 15.560068 & 9.513870 & 9.513870 & -6.668600 \\ -2.718249 & 9.513870 & 581.705188 & 581.705188 & 122.321184 \\ -2.718249 & 9.513870 & 581.705188 & 581.705188 & 122.321184 \\ 0.000000 & -6.668600 & 122.321184 & 122.321184 & 285.797160 \end{pmatrix}.$$

$$\epsilon = 0.01.$$

Parameters for F and V are

Parameter	V_0	F_0	q	$\bar{r}$	e_0	e_1	e_2	e_3
Value	215	25.948	0.153097	0.045	27.90	29.11	30.18	31.55

Parameters for the time series model are

$$\xi_{-1} = \begin{pmatrix} 0.353101 & 1.150000 & 1.723809 & 1.002339 & 0.101330 & 1.911712 & 4.879275 & 0.249151 \end{pmatrix}^T,$$

$$\xi_0 = \begin{pmatrix} 0.313223 & 1.200000 & 1.811451 & 1.083672 & 0.102742 & 1.914840 & 4.878644 & 0.254137 \end{pmatrix}^T,$$

$$A = 10^{-2} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 70.3070 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 75.0830 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 98.4630 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 85.3790 \end{pmatrix},$$

$$\alpha = 10^{-2} \begin{pmatrix} -6.9950 & 0 & 0 \\ 0 & -20.3310 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4.4625 \\ 0 & 0 & 0 \\ 0 & 0.0475 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\beta = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\gamma = \begin{pmatrix} -0.213706 & 0.223144 & 1.945910 \end{pmatrix}^T,$$

$$c = \begin{pmatrix} 0.014949 & -0.045367 & 0 & 0 & 0.086836 & 0 & 0.000106 & 0 \end{pmatrix}^T,$$

$$\sigma^2 = 10^{-4} \begin{pmatrix} 2459.3665 & 2474.6456 & -44.8775 & -88.3291 & 14.5393 & 6.6564 & -6.3217 & 7.2190 \\ 2474.6456 & 2591.4208 & -18.6423 & -74.1652 & 20.4939 & 6.2654 & -6.4979 & 7.5418 \\ -44.8775 & -18.6423 & 253.0327 & 122.0457 & 5.1302 & -0.3554 & 0.1016 & -0.0265 \\ -88.3291 & -74.1652 & 122.0457 & 94.3365 & 2.8128 & 0.0872 & 0.2266 & -0.0933 \\ 14.5393 & 20.4939 & 5.1302 & 2.8128 & 3.5865 & 0.1731 & -0.0452 & 0.0106 \\ 6.6564 & 6.2654 & -0.3554 & 0.0872 & 0.1731 & 0.1939 & -0.0174 & 0.0046 \\ -6.3217 & -6.4979 & 0.1016 & 0.2266 & -0.0452 & -0.0174 & 0.0167 & -0.0198 \\ 7.2190 & 7.5418 & -0.0265 & -0.0933 & 0.0106 & 0.0046 & -0.0198 & 0.0618 \end{pmatrix}.$$

ACKNOWLEDGEMENT

The work of the second author was supported by Finnish Academy under contract no. 3385. We would like to thank Matti Koivu for fruitful discussions and for implementing essential parts of the discretization routine.

(Received October 31, 2007.)

REFERENCES

[1] J. Blomvall and A. Shapiro: Solving multistage asset investment problems by the sample average approximation method. Math. Program. *108* (2006), 571–595.
[2] A. Chiralaksanakul: Monte Carlo Methods for Multi-stage Stochastic Programs. PhD. Thesis, University of Texas at Austin, 2003.
[3] A. Chiralaksanakul and D. Morton: Assessing policy quality in multi-stage stochastic programming. Stochastic Programming E-Print Series, 2004.
[4] R. Fourer, D. M. Gay, and B.W. Kernighan: AMPL: A Modeling Language for Mathematical Programming. Second edition. Thomson, Canada 2002.
[5] K. Frauendorfer: Barycentric scenario trees in convex multistage stochastic programming. Math. Programming *75* (1996), 277–293.
[6] P. Glasserman: Monte Carlo Methods in Financial Engineering. Springer-Verlag, New York 2004.
[7] H. Heitsch and W. Römisch: Scenario tree modelling for multistage stochastic programs. Math. Programming. To appear.
[8] P. Hilli, M. Koivu, T. Pennanen, and A. Ranne: A stochastic programming model for asset liability management of a Finnish pension company. Ann. Oper. Res. *152* (2007), 115–139.
[9] P. Kall and J. Mayer: Stochastic Linear Programming. Springer-Verlag, New York 2005.
[10] D. Kuhn: Generalized Bounds for Convex Multistage Stochastic Programs. Springer-Verlag, Berlin 2005.
[11] H. Niederreiter: Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia 1992.
[12] H. Niederreiter and D. Talay: Monte Carlo and Quasi-Monte Carlo Methods 2004. Springer-Verlag, Berlin 2006.
[13] P. Olsen: Discretizations of multistage stochastic programming problems. Math. Programming Stud. *6* (1976), 111–124.
[14] T. Pennanen: Epi-convergent discretizations of multistage stochastic programs. Math. Oper. Res. *30* (2005), 245–256.

[15] T. Pennanen: Epi-convergent discretizations of multistage stochastic programs via integration quadratures. Math. Programming, Series B. To appear.
[16] T. Pennanen and M. Koivu: Integration quadratures in discretization of stochastic programs. SPEPS E-Print Series, 2002.
[17] G. Pflug: Scenario tree generation for multiperiod financial optimization by optimal discretization. Math. Programming *89* (2001), 251–271.
[18] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery: Numerical Recipes in C. Cambridge University Press, Cambridge 1992.
[19] R. T. Rockafellar and R. J.-B. Wets: Continuous versus measurable recourse in N-stage stochastic programming. J. Math. Anal. Appl. *48* (1974), 836–859.
[20] R. T. Rockafellar and R. J.-B. Wets: Nonanticipativity and L^1-martingales in stochastic optimization problems. Math. Programming Stud. (1976), 170–187.
[21] A. Shapiro: Inference of statistical bounds for multistage stochastic programming problems. Math. Methods Oper. Res. *58* (2003), 57–68.
[22] A. Shapiro: On complexity of multistage stochastic programs. Oper. Res. Lett. *34* (2006), 1–8.
[23] I. H. Sloan and S. Joe: Lattice Methods for Multiple Integration. The Clarendon Press Oxford University Press, New York 1994.
[24] I. M. Sobol': The distribution of points in a cube and the approximate evaluation of integrals. U.S.S.R. Computational Math. And Math. Phys. (1967), 86–112.

Petri Hilli and Teemu Pennanen, Department of Business Technology, Helsinki School of Economics, PL 1210, 00101 Helsinki. Finland.
e-mails: petri.hilli@qsa.fi, teemu.pennanen@hse.fi