

# Stochastic Optimization in Asset & Liability Management: A Model for Non-Maturing Accounts<sup>1</sup>

Karl Frauendorfer  
*University of St. Gallen*  
*Institute of Operations Research*  
*St. Gallen, Switzerland*

Michael Schürle  
*University of St. Gallen*  
*Institute of Operations Research*  
*St. Gallen, Switzerland*

## Abstract

A multistage stochastic optimization model for the management of non-maturing account positions like savings deposits and variable-rate mortgages is introduced which takes the risks induced by uncertain future interest rates and customer behavior into account. Stochastic factors are discretized using the barycentric approximation technique. This generates two scenario trees whose associated deterministic equivalent programs provide exact upper and lower bounds to the original problem. Practical experience from the application in a major Swiss bank is reported.

**Keywords:** Stochastic programming, approximation, asset & liability management.

## 1 Introduction

Stochastic programming has received increasing attention from financial institutions recently since the shortcomings of traditional approaches that are widely used in practice came to light. For instance, in portfolio optimization the mean-variance

---

<sup>1</sup>Research for this paper was supported by the *Swiss National Science Foundation*, Grant No. 21-39'575.93.

framework due to Markowitz [48] captures the volatility and correlations among financial instruments but may generate solutions that do not seem like a reasonable mix to achieve the indicated risk and return (cf. [22]) and are highly sensitive to the input, i.e., expectations and covariances (cf. [3, 8]). Moreover, it does not take into account the possibility of future rebalancing transactions of the portfolio or additional in- and outflows of cash that occur during the planning period. In many problems in the field of asset and liability management (ALM), the increased volatility in financial markets since the 70s and the introduction of derivatives revealed severe deficiencies of popular portfolio immunization strategies which match the duration and, possibly, the convexity of both assets and liabilities (cf. [67]). These approaches hedge only against relatively small shifts in interest rates and are not appropriate to deal with the complex cash flow structures of new financial instruments (cf. [53]).

## 1.1 Stochastic optimization for financial decision making

For obvious reasons, stochastic optimization models seem to be a natural approach in order to address the requirements of a large number of financial planning problems (e.g., see [15, 49, 50] for an overview of applications, [4, 12, 14, 52] for general ALM model formulations or [23, 24, 36, 37, 40, 64, 65, 66, 67] for fixed-income problems). On one hand, the models allow the reflection of uncertainty in future prices, yields and exchange rates, volatilities etc. by generating scenarios of their possible future outcomes. These scenarios quantify the impact of changes in the underlying risk factors on the return of investment strategies or the deviation from a certain target position like an index, a benchmark portfolio, a liability etc.

On the other hand, a stochastic program reflects not only the dynamics of uncertain data but also of decisions more appropriately since transactions may take place at discrete points in time until the end of some predefined planning horizon. For all scenarios under consideration, a decision must be taken at each stage based on realizations of random data and earlier actions. This allows the correction of an initial policy, e.g., if it does not achieve the investment goal for certain scenarios.

In general, a stochastic optimization model yields a large-scale program since it has to include a high number of scenarios to reflect the entire universe of possible future outcomes of risk factors and cash flows. The first models for financial planning appearing in the 80s (cf. [44, 45]) could not meet this requirement due to limitations of the computational resources available at this time.

However, the dramatic improvement in powerful hardware as well as the development of efficient algorithms, in particular if they exploit the special structure and high sparsity inherent to stochastic programs, now provide the basis to solve problems where the number of scenarios is between some thousands and one million, depending on the problem structure and the system architecture (cf. [25]). Moreover, new theoretical models from the financial literature and related empirical evidence have sharpened the understanding of the dynamics of risk factors such as interest and ex-

change rates, prices of financial instruments etc. Both of these developments allow the modeling of complex problems more realistically. In this way, the fields of Finance and Optimization come closer together.

## 1.2 Review of current approaches

Meanwhile, a large number of stochastic optimization models has been introduced for various applications in Finance. Among them are the Russell-Yasuda-Kasai model due to Cariño et al. [4, 5], which can be seen as the first successful commercial application of multistage stochastic programming, the models for ALM and fixed-income portfolio management of Zenios et al. [36, 66, 67] and Dupačová et al. [18], or the multistage portfolio optimization models due to Dantzig and Infanger [11] and Steinbach [61], to mention just a few.

In general, one starts from assumptions about the distribution of risk factors which is typically continuous. Since scenarios are used as input for stochastic programs to describe the uncertainty, a discrete subset of possible realizations of the random data must be determined that is “representative” for the entire universe of future outcomes. Loosely speaking, this means that the solution of such an approximated problem comes “as close as possible” to the exact (but unknown) optimum of a problem with the true distribution.

Simulation is a widely exploited approach for the selection of scenarios and can be easily combined with decomposition methods (cf. [10, 38]). Since the solution depends heavily on the choice of the scenario set, a statistical analysis is essential to assess the accuracy and stability of the problem (cf. [16, 17]). Although the amount of computational effort is independent of the dimension of random data, Monte Carlo methods often suffer from a low convergence rate of the order  $1/\sqrt{s}$  as implied by the central limit theorem, where  $s$  is the sample size. Therefore, variance reduction techniques like importance sampling are helpful to reduce the error of estimates for the objective. For multistage problems, the *expected value of perfect information (EVPI)* (cf. [7, 13]) is a useful criterion for the selection of an enhanced set of representative scenarios. In any case, simulation based approaches provide only probabilistic error bounds.

*Approximation schemes* are based on partitioning the domain of random data into cells and using representative points within them (cf. [1, 19, 20, 26, 32, 43]). Exploiting certain properties of the stochastic program, mainly convexity of value functions, allows the determination of exact lower and upper bounds to the original problem. As a consequence, the error induced by the approximation can be quantified more precisely, and the accuracy may be improved by deliberately adding new scenarios. Careful control of this process is necessary since the number of scenarios grows exponentially with the dimension size and accuracy.

## 1.3 Contribution of this paper

In this paper, a stochastic optimization model for an application from the field of ALM is introduced which was developed in co-operation with a major Swiss bank for the management of so-called “non-maturing account” positions. This includes savings deposits as well as a special type of non-fixed mortgages which is common in Switzerland. Their characteristic feature is that there exists no contractual maturity on such products but bank customers are allowed to withdraw their investments or prepay their mortgages, respectively, at any point in time at no penalty. As a consequence, the volume of both positions fluctuates heavily as customers react to changes in the market environment, e.g., rising or falling yields, or the relative attractiveness of alternative investment opportunities.

Therefore, in the formulation of the stochastic program uncertainty affects not only the coefficients in the objective (future interest rates) but also on the right-hand-side of constraints (volume change). Moreover, both may be correlated to reflect a dependency between interest rates and volume. All these aspects can be addressed by the *barycentric approximation* technique introduced in the sequel to derive exact bounds for value functions corresponding to convex multistage stochastic programs. An approximation technique is preferred for this type of application since statistical analysis implies that the dynamics of interest rates can be described by at most three factors in order to explain more than 95 % of the variance. These factors control level, curvature, and steepness of the yield curve (cf. [46]). This enables keeping the problem size relatively moderate even in a multistage model.

The remainder of the paper is organized as follows: The next section introduces the formulation of the optimization model for non-maturing accounts. In section 3, the structural properties of convex multistage stochastic programs are outlined. Section 4 introduces the barycentric approximation scheme that is used to derive scenario trees associated with upper and lower bounds to the original problem. Numerical results for the approximation are also given. Section 5 reports practical experience from the application of the model compared to traditional approaches. The main results are summarized in section 6 together with an outlook to possible improvements of the model as well as future directions of research.

## 2 A model for uncertain maturities

### 2.1 Problem characteristics

As outlined above, non-maturing accounts can be characterized as follows: (1) There is no contractual maturity on these positions since bank customers are allowed to withdraw or repay their investments and credits at any point in time. (2) The customer rate is not indexed to certain interest rates or prices of traded instruments but adjustable to market conditions as a matter of policy. The most common examples

include some forms of savings accounts or non-fixed mortgages that are widespread in Europe and the U.S. The management of such account positions is a particularly ambitious task since these assets and liabilities are not only sensitive to changes in interest rates but have also embedded call or put options that may be exercised by the customer. For example, a homeowner has the option to prepay the outstanding balance of his mortgage and call the security.

It can be observed that customer behavior depends strongly on the current market environment. In case of variable-rate mortgages, changes in the total volume are positively correlated with interest rates. When the latter are low, there is a sharp drop in demand since customers switch to fixed-rate mortgages in order to hedge themselves against a future rise (*prepayment risk*). In case of savings deposits, the volume increases since their yields are relatively attractive when compared to alternative short-term securities, and even institutional investors like pension funds prefer these deposits instead of direct investments in the money market. This results in a negative correlation between interest rates and volume change. In such a situation, it is difficult for financial management to find a combination of fixed-income instruments that provides a sufficient margin and takes into account the risk that a significant portion of the deposits is withdrawn (*withdrawal risk*).

During a period of high interest rates, homeowners' demand for non-fixed mortgages rises significantly while investors shift their assets from variable-rate savings accounts to bonds with long maturities. As a consequence, the mortgages must be refinanced on the money and capital market at increased funding costs. Moreover, there is a political cap on the mortgage rate in Switzerland, and numerous banks were not able to refinance their mortgages at a positive margin at the beginning of the 90s. These difficulties caused a broad discussion among practitioners from the financial industry about the management of non-maturing account positions and the risks induced by the embedded options.

Clearly, the use of *duration matching* does not apply here since one cannot find adequate duration measures due to the volume fluctuations, beside other shortcomings of this concept. This has motivated the *replicating portfolio* approach which is based on the idea of mimicing the behavior of the target position in order to capture its characteristics. The objective is to find a portfolio of fixed-income securities whose return replicates the customer rate of the relevant asset or liability position plus a margin. Transaction costs remain low since liquid money market instruments and swaps are used that are held until maturity to avoid a rebalancing. Maturing funds are always renewed at the same maturity. Prepayment and withdrawal risks are implicitly taken into account as the volume of the replicating portfolio has to coincide with the volume of the target position at all points in time. The weights are determined through minimizing the tracking error for a historic sample period and remain constant over time.

By means of this approach, uncertain cash flows are transformed into (apparently) certain ones, allowing the bank to manage them like normal maturing accounts. These

replicating portfolios are implemented as passive investment and refinancing strategies. However, the question arises whether a dynamic policy with active reactions to changes in the market environment and customer behavior could increase the bank's profit. In particular, it remains to be clarified if the correlation between interest rates and volume can be exploited more appropriately to manage the inherent risks. Clearly, a stochastic optimization model is able to address most of these requirements.

## 2.2 Model formulation

For simplicity, only the problem of reinvesting savings accounts on the market is investigated here since the model for refinancing mortgages is equivalent and can be derived easily from it. The formulation of the optimization model is straightforward:  $\mathcal{D} = \{1, 2, \dots, D\}$  denotes a set of maturity dates for fixed-income securities held in the portfolio. Investment opportunities are given by a set of traded standard maturities  $\mathcal{D}^S \subset \mathcal{D}$ . Let  $\varphi_t^{d,+}$  be the discounted accrued interest payments for an investment of \$1 in maturity  $d \in \mathcal{D}^S$  at time  $t$ . The model has also the option to raise funds which are reinvested in addition to the total savings volume. The sum of interest payments for \$1 of such a short position is given by  $\varphi_t^{d,-}$ . Clearly, except for  $t = 0$  these coefficients depend on future interest rates and, hence, are uncertain.

The underlying interest rate model used here to describe the evolution of interest rates resembles the idea of *key rates* analogously to the duration model of Ho [39]. It is assumed that the yield curve can be segmented into  $K_t$  different sections where rates move in the same direction. These segments are represented by  $K_t$  key rates of different maturities whose dynamics can be described by correlated Brownian motions with (possibly time-dependent) drift. This results in normally distributed interest rate changes. Rates for the remaining maturities are interpolated.

The coefficients  $\varphi_t^{d,+}(\eta_t)$  and  $\varphi_t^{d,-}(\eta_t)$  for all  $d \in \mathcal{D}^S$  are functions of this  $K_t$ -dimensional Brownian motion  $(\eta_t; t = 1, \dots, T)$  in discrete time driving the evolution of the yield curve. Here,  $K_t = K = 3$  as a three factor model is sufficient to reflect a great variety of term structure movements. The functional relationship between  $\eta_t$  and  $\varphi_t^{d,+}, \varphi_t^{d,-}$  incorporates the sensitivity of interest rates subject to changes in the risk factors, transactions costs, a bid-ask spread as well as the discount mechanism. Only payments within the planning horizon are considered, i.e., those that are induced by an investment or borrowing in  $t \leq T$  but occur after  $T + 1$  are neglected. A formal specification of the relation between risk factors and coefficients in the objective is omitted here since the notation is rather cumbersome. Note that the current values of  $\eta_0$  can be derived from market observations.

At each point in time  $t = 0, \dots, T$ , decisions on the amount  $x_t^{d,+} \geq 0$  of long and  $x_t^{d,-} \geq 0$  of short positions in maturity  $d$  have to be made subject to budget constraints

$$\begin{aligned}
x_t^d - x_{t-1}^{d+1} - x_t^{d,+} + x_t^{d,-} &= 0 & t = 0, \dots, T; \forall d \in \mathcal{D}^S \\
x_t^d - x_{t-1}^{d+1} &= 0 & t = 0, \dots, T; \forall d \in \mathcal{D} \setminus \mathcal{D}^S.
\end{aligned}$$

The latter constraint ensures that the sum of all long and short positions  $x_t^d \in \mathbb{R}$  maturing after  $d$  periods is equal to the corresponding value in the previous period for non-traded maturity dates while the former corrects it by the new long and short sales in  $t$  for traded maturities. Note that  $x_{-1}^d$  indicates the amount of maturity  $d$  in the initial portfolio from decisions in the past. At time  $t$ , the portfolio has to match the total savings volume  $v_t \in \mathbb{R}$ :

$$v_t = \sum_{d \in \mathcal{D}} x_t^d \quad t = 0, \dots, T.$$

The total savings volume is given by its value in the previous period  $t - 1$ , corrected by the stochastic volume change  $\xi_t$  in  $t$ :

$$v_t = v_{t-1} + \xi_t \quad t = 1, \dots, T.$$

Again, the volume change is modeled by a Brownian motion with drift in discrete time ( $\xi_t; t = 1, \dots, T$ ) of dimension  $L_t = L = 1$ . It may be correlated with the components of the stochastic process  $\eta_t$  to reflect a relation between changes in interest rates and volume.

Raising short positions and reinvesting them in addition to the savings volume may be viewed as some speculative strategy. Therefore, short sales can be restricted to an amount equal to the sum of funds maturing at  $t + 1, \dots, t + m$ , where  $m > 0$  is defined by the decision maker ( $m = 1$  prohibits any borrowings):

$$\sum_{d \in \mathcal{D}^S} x_t^{d,+} \leq \xi_t + \sum_{d=1}^m x_{t-1}^d \quad t = 1, \dots, T.$$

Depending on the current interest rate curve and the amount that has to be reinvested, a situation might occur where the optimal investment strategy cannot be implemented due to liquidity restrictions in the Swiss market, in particular if the model finds a policy that is not broadly diversified over different maturities. This is addressed by imposing upper limits  $\ell_t^{d,+}, \ell_t^{d,-}$  for investments and borrowings:

$$\begin{aligned}
0 \leq x_t^{d,+} \leq \ell_t^{d,+} & \quad t = 0, \dots, T; \forall d \in \mathcal{D}^S \\
0 \leq x_t^{d,-} \leq \ell_t^{d,-} & \quad t = 0, \dots, T; \forall d \in \mathcal{D}^S.
\end{aligned}$$

The restrictions above must hold for all observations of  $\eta_t$  and  $\xi_t$ ,  $t = 1, \dots, T$ . Moreover, decisions  $x_t^{d,+}, x_t^{d,-}$  have to be made independent of future outcomes of  $\eta_{t+1}, \dots, \eta_T$  and  $\xi_{t+1}, \dots, \xi_T$  since these are unknown at time  $t$ . Hence, investment policies must not anticipate any information that becomes known in the future. This is incorporated in the optimization model by additional *nonanticipativity* constraints.

Finally, the objective is to maximize the expected present value of the income from all investments  $\varphi_t^{d,+} \cdot x_t^{d,+}$  minus the costs for borrowings  $\varphi_t^{d,-} \cdot x_t^{d,-}$  for all  $d \in \mathcal{D}^S$  over the planning horizon  $T$ . The expectation is taken with respect to the joint probability measure  $P$  of  $(\eta, \xi)$  associated with time  $t = 1, \dots, T$ , i.e.,  $\eta = (\eta_1, \dots, \eta_T)$ ,  $\xi = (\xi_1, \dots, \xi_T)$ . In the standard form of a minimization problem, the multistage stochastic program reads as:

$$\begin{aligned}
& \min \int \sum_{t=0}^T \sum_{d \in \mathcal{D}^S} (\varphi_t^{d,-}(\eta_t) \cdot x_t^{d,-} - \varphi_t^{d,+}(\eta_t) \cdot x_t^{d,+}) dP(\eta, \xi) \\
& \text{s.t. } x_t^d - x_{t-1}^{d+1} - x_t^{d,+} + x_t^{d,-} = 0 && t = 0, 1, \dots, T; \forall d \in \mathcal{D}^S \\
& x_t^d - x_{t-1}^{d+1} = 0 && t = 0, 1, \dots, T; \forall d \notin \mathcal{D}^S \\
& v_t - \sum_{d \in \mathcal{D}} x_t^d = 0 && t = 0, 1, \dots, T \\
& v_t - v_{t-1} = \xi_t && t = 1, 2, \dots, T \\
& \sum_{d \in \mathcal{D}^S} x_t^{d,+} - \sum_{d=1}^m x_{t-1}^d \leq \xi_t && t = 0, 1, \dots, T \\
& 0 \leq x_t^{d,+} \leq \ell_t^{d,+} && t = 0, 1, \dots, T; \forall d \in \mathcal{D}^S \\
& 0 \leq x_t^{d,-} \leq \ell_t^{d,-} && t = 0, 1, \dots, T; \forall d \in \mathcal{D}^S \\
& x_t^{d,+}, x_t^{d,-} \text{ nonanticipative} && t = 0, 1, \dots, T; \forall d \in \mathcal{D}^S \\
& v_t, x_t^d \in \mathbb{R} \text{ nonanticipative} && t = 0, 1, \dots, T; \forall d \in \mathcal{D}.
\end{aligned} \tag{1}$$

To evaluate the objective function for a particular policy, it is necessary to calculate an integral with respect to the measure  $P$ . At each stage  $t$ , the costs and profits induced by a decision can be quantified by a value function which is formally introduced in the next section. The latter is not given analytically, not even explicitly, but implicitly by the solutions of a multistage stochastic optimization problem with respect to the remaining stages  $t + 1, \dots, T$ . As a consequence, the integration in (1) cannot be performed analytically, and numerical methods are required.

A common approach is to approximate the continuous distribution of random data in the original stochastic program by a discrete one. More precisely, the stochastic evolution of interest rates and volume change is approximated by two *scenario trees*. This yields two other optimization problems where one deals with a sum in the objective function which can be easily calculated. To minimize the error induced by the discretization, it is exploited that the value functions at stage  $t$  associated with (1) are convex-concave saddle functions in  $(\eta_t, \xi_t)$ . The saddle property allows the determination of exact upper and lower bounds to the original problem and is discussed in the next section for a general formulation of multistage stochastic programs.

### 3 Multistage stochastic programs

#### 3.1 Formal description

Formally, the evolution of uncertain data over the planning horizon  $T$  in a multistage stochastic program can be described by a multi-dimensional stochastic process  $(\omega_t, t = 1, \dots, T)$  in discrete time on a common Borel space  $(\Omega, \mathcal{B}^M)$  with compact  $\Omega \subset \mathbb{R}^M$  (cf. [27, 28, 29]). Let  $P$  represent the (regular) joint probability measure of  $\omega := (\omega_1, \dots, \omega_T)$ . The associated conditional measure with respect to  $\omega_t$  is denoted  $P_t(\cdot | \omega^{t-1})$  for  $t = 1, \dots, T$ . For reasons of compactness,  $\omega^t := (\omega_1, \dots, \omega_t)$  represents the sequence of observations of  $\omega_t \in \Omega_t \subset \mathbb{R}^{M_t}$  up to time  $t$ , where  $\Omega_1 \times \dots \times \Omega_T = \Omega$ ,  $M_1 + \dots + M_T = M$ . Note that  $\omega_0$  denotes those data that are currently observed and, hence, deterministic.

At time  $t = 0$ , a decision  $u_0 \in \mathbb{R}^{n_0}$  is made without knowing  $\omega_t$  for the subsequent stages  $t = 1, \dots, T$ . After  $\omega_t$  was observed at time  $t > 0$ , the initial policy may be corrected by a new decision  $u_t \in \mathbb{R}^{n_t}$  based on the known history of observations  $\omega^t$  and decisions  $u^t := (u_0, u_1, \dots, u_t) \in \mathbb{R}^{n^t}$ ,  $n^t = n_0 + \dots + n_t$ . In particular,  $u_t$  has to be independent of future outcomes  $\omega_{t+1}, \dots, \omega_T$ . Therefore, the solution of the underlying stochastic optimization problem is a *recourse function* with the property

$$u(\omega) = (u_0, u_1(\omega^1), \dots, u_T(\omega^T)) \in \mathbb{R}^n, \quad n = n_0 + n_1 + \dots + n_T,$$

known as *nonanticipativity*. The initial decision  $u_0$  induces some (non-random) costs  $\rho_0$ . For the subsequent stages  $t = 1, \dots, T$ , the costs  $\rho_t(u^t, \omega^t)$  are determined by the sequence of earlier decisions  $u^t$  and realizations of  $\omega^t$ . The feasible set is assumed to be convex, compact, and non-empty for any  $\omega$ . Again, it depends on previous decisions and observations for  $t > 0$  and can be characterized by the system of inequalities

$$\begin{aligned} f_0(u_0) &\leq 0 \\ f_t(u^t, \omega^t) &\leq 0 \quad t = 1, \dots, T. \end{aligned} \tag{2}$$

$f_0(\cdot)$  and  $f_t(\cdot, \cdot)$  are vector-valued and  $\rho_0(\cdot)$  and  $\rho_t(\cdot, \cdot)$  are real-valued functions defined on the corresponding Euclidian spaces. Furthermore,  $\rho_t(\cdot, \cdot)$  are supposed to be convex in  $u^t$  for any random outcome  $\omega^t$ . The objective is to find a nonanticipative recourse function  $u(\cdot)$  that minimizes the expected total costs over the planning horizon and satisfies the constraints (2):

$$\begin{aligned} \min \quad & \left\{ \rho_0(u_0) + \int_{\Omega} [\sum_{t=1}^T \rho_t(u^t, \omega^t)] dP(\omega) \right\} \\ \text{s.t.} \quad & f_0(u_0) \leq 0 \\ & f_t(u^t, \omega^t) \leq 0, \quad t = 1, \dots, T, \\ & u(\cdot) \text{ nonanticipative.} \end{aligned} \tag{3}$$

The meaning of the last (nonanticipativity) constraint is: Let  $\omega^t$  and  $u^t$  satisfy  $f_0(u_0) \leq 0, f_1(u^1, \omega^1) \leq 0, \dots, f_t(u^t, \omega^t) \leq 0$ , then there always exists a sequence

$u_{t+1}, \dots, u_T$  for any  $(\omega_{t+1}, \dots, \omega_T)$ , so that  $u = (u_0, u_1, \dots, u_T)$  is feasible with respect to (2). As a consequence, there is always a feasible completion of the problem (3) for the remaining stages  $t + 1, \dots, T$  independent of the realizations of  $\omega_\tau$ ,  $\tau = t + 1, \dots, T$ , provided that the decisions  $u_\tau$  in  $\tau = 0, 1, \dots, t$  are feasible. This can be seen as a counterpart to the case of relatively complete recourse in two-stage stochastic programming (cf. [62]).

### 3.2 Saddle property of value functions

In order to distinguish those uncertain data that affect the objective from those influencing the constraints, the random vectors for  $t = 1, \dots, T$  are decomposed according to  $\omega_t = (\eta_t, \xi_t)$ , where  $\eta_t \in \Theta_t \subset \mathbb{R}^{K_t}$ ,  $\xi_t \in \Xi_t \subset \mathbb{R}^{L_t}$ ,  $\Omega_t = \Theta_t \times \Xi_t$ ,  $M_t = K_t + L_t$ . The functions defining the feasible set in the second line of (2) can then be written as  $f_1(u^1, \xi^1) \leq 0, \dots, f_T(u^T, \xi^T) \leq 0$ , and the costs are now of the form  $\rho_t(u^t, \eta^t)$ . In case such a decomposition of  $\omega_t$  into  $\eta_t$  and  $\xi_t$  is not obvious, one may augment the probability space (for details, see [27]). In order to write the problem without stating the constraints explicitly, the function

$$g_t(u^t, \eta^t, \xi^t) := \begin{cases} \rho_t(u^t, \eta^t) & \text{if } f_t(u^t, \xi^t) \leq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

is introduced. For the following analysis, it is useful to consider the dynamic version of problem (3) stated in terms of *recourse* or *value functions*. These can be obtained if the multistage program is written as a series of nested two-stage programs, starting in the last stage  $T$  with

$$\phi_T(u^{T-1}, \eta^T, \xi^T) = \min_{u_T \geq 0} g_T(u^{T-1}, u_T, \eta^T, \xi^T) \quad (4)$$

and then backwards for  $t = T - 1, \dots, 0$

$$\begin{aligned} \phi_t(u^{t-1}, \eta^t, \xi^t) &= \min_{u_t \geq 0} \{g_t(u^{t-1}, u_t, \eta^t, \xi^t) \\ &+ \int_{\Theta_{t+1} \times \Xi_{t+1}} \phi_{t+1}(u^{t-1}, u_t, \eta^t, \xi^t, \eta_{t+1}, \xi_{t+1}) dP(\eta_{t+1}, \xi_{t+1} | \eta^t, \xi^t)\}. \end{aligned} \quad (5)$$

Again,  $(\eta^0, \xi^0)$  are currently observed data, whereas  $u^{-1}$  represents decisions from the past. Here,  $u^t = (u^{t-1}, u_t)$  is decomposed in two parts to emphasize that decisions  $u^{t-1}$  were already made in the preceding stages and only  $u_t$  must be determined in (5). The optimal decision for the current stage represents a trade-off between the imminent costs  $\rho_t$  in  $t$  and the expected future costs for the remaining periods induced by  $u_t$ . According to [28], based on arguments in [54], to ensure that the problems (4) and (5) can be solved the following assumptions are required for  $t = 1, \dots, T$ :

- (i)  $\Theta_t \times \Xi_t$  is compact, convex, and covers the support of  $(\eta_t, \xi_t)$ .

(ii)  $\rho_t(u^t, \eta^t)$  is a continuous saddle function on  $\mathbb{R}^{n^t} \times \Theta^t$  which is convex in  $u^t$  and concave in  $\eta^t$ .

Note that this is satisfied, e.g., if  $\rho_t$  is bilinear.

(iii) The feasible sets  $\{u_0 | f_0(u_0) \leq 0\}$  and  $\{(u^t, \xi^t) | \xi \in \Xi, f_t(u^t, \xi^t) \leq 0\}$  are compact, convex subsets of  $\mathbb{R}^{n^t} \times \Xi^t$ .

In particular, this covers the case that the constraints can be written in the form

$$f_t(u^{t-1}, u_t, \xi^t) = d_t(u_t) - e_t(u^{t-1}, \xi^t)$$

with  $d_t$  convex and  $e_t$  linear affine.

As outlined before, approximation schemes are based on the convexity or, if uncertainty affects the objective and the right-hand-sides of constraints, saddle property of value functions which allows the derivation of lower and upper bounds based on the inequalities due to Jensen [42] and Edmundson-Madansky [21, 47]. Barycentric approximation is a generalization of these concepts for bounding the expectation of saddle functions in the case of dependent random variables which was introduced in [27] in the context of two-stage stochastic programming. It remains to clarify under which conditions it can be extended to the multistage case.

According to (i), the problem (4) for the final stage  $T$  is a convex optimization problem with parameters  $(u^{T-1}, \eta^T, \xi^T)$ . Assumption (ii) implies that the objective function  $g_T$  in (4) is a lower closed saddle function. Therefore, the corresponding value function  $\phi_T(u^T, \eta^T, \xi^T)$  is also a saddle function, convex in  $(x^{T-1}, \xi^T)$  and concave in  $\eta^T$ . In order to derive bounds for the expectation of (5), the saddle property of the value function in  $T$  must be “inherited” to the remaining stages  $T-1, \dots, 1$ . When calculating the expected recourse costs

$$E_t \phi_t(u^{t-2}, u_{t-1}, \eta^{t-1}, \xi^{t-1}) = \int_{\Theta_t \times \Xi_t} \phi_t(u^{t-2}, u_{t-1}, \eta^{t-1}, \xi^{t-1}, \eta_t, \xi_t) dP_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}), \quad (6)$$

the probability measure  $P_t$  depends on  $(\eta^{t-1}, \xi^{t-1})$ . As a consequence, the saddle property of  $\phi_t(x^{t-2}, x_{t-1}, \eta^{t-1}, \xi^{t-1}, \eta_t, \xi_t)$  is *not* inherited in general due to the integration with respect to  $P_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1})$ . However, if the distribution functions are of the form

$$Q_t((\eta_t, \xi_t) + H_t(\eta^{t-1}, \xi^{t-1})), \quad (7)$$

where  $Q_t$  is a regular distribution function over  $(\Theta_t \times \Xi_t, \mathcal{B}^{K_t+L_t})$  and  $H_t$  is a linear mapping, the integral in (6) can be written as

$$\int_{\Theta_t \times \Xi_t} \phi_t(u^{t-2}, u_{t-1}, \eta^{t-1}, \xi^{t-1}, (\eta_t, \xi_t) - H_t(\eta^{t-1}, \xi^{t-1})) dQ_t(\eta_t, \xi_t)$$

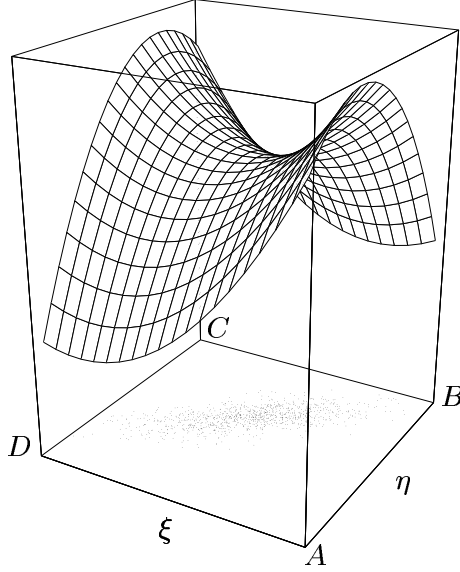


Figure 1: Saddle property of value functions

(note that the integration is now performed with respect to a measure independent of  $(\eta^{t-1}, \xi^{t-1})$ ). Then, it can be easily verified that the expectation in (6) is a saddle function on its domain  $\Theta^t \times \{(u^{t-1}, \xi^t) | \xi^t \in \Theta^t, f_\tau(u^\tau, \xi^\tau) \leq 0, \tau = 1, \dots, t-1\}$  which is *convex* in  $(u^{t-2}, u_{t-1}, \xi^{t-1})$  and *concave* in  $\eta^{t-1}$  (for details see [28]). Together with the convexity of  $g_t(u^t, \eta^t)$  implied by (ii), this results in the saddle property of the objective function of problem (5),

$$g_t(u^{t-1}, u_t, \eta^t, \xi^t) + E_{t+1} \phi_{t+1}(u^{t-1}, u_t, \eta^t, \xi^t, \eta_{t+1}, \xi_{t+1}). \quad (8)$$

Hence, for stages  $t = T-1, \dots, 1$  the value functions  $\phi_t(u^{t-1}, \eta^t, \xi^t)$  are lower closed saddle functions on their domain. An illustration is given in Figure 1 where  $\phi_t(u^{t-1}, \eta^{t-1}, \eta_t, \xi^{t-1}, \xi_t)$  is shown for  $(\eta_t, \xi_t) \in \Theta_t \times \Xi_t \subset \mathbb{R} \times \mathbb{R}$ , i.e., one-dimensional distributions for the coefficients in objective and right-hand-sides. Note that the value function quantifies the imminent costs for the current stage  $t$  plus the expected future costs provided that the subsequent decisions are optimal. Therefore, it cannot be represented analytically but is given implicitly for each  $(u^{t-1}, \eta^t, \xi^t)$  as the solution of a multistage stochastic program with respect to the remaining stages  $t+1, \dots, T$ .

### 3.3 Solvability of stochastic programs

To ensure that the multistage stochastic program can be solved, it is required that the corresponding value functions are continuous. A sufficient condition for this is that value functions are subdifferentiable which is given if the Slater condition holds: For the last stage  $T$ , there must be a point  $\hat{u}_T$  depending on  $(u^{T-1}, \xi^T)$  with  $f_T(u^{T-1}, \hat{u}_T, \xi^T) < 0$ . Calculating the expectation in (6) preserves continuity since the

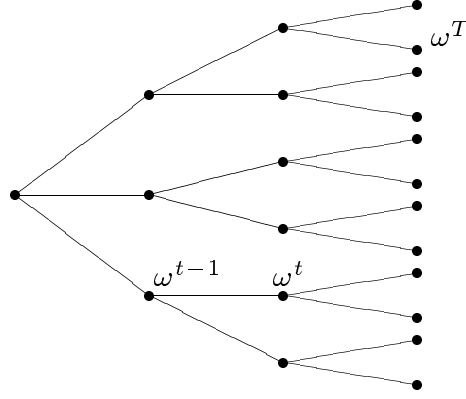


Figure 2: Illustration of a scenario tree

integration is done with respect to the compact region  $\Theta_t \times \Xi_t$ . To ensure continuity of the value functions for the preceding stages, again the existence of a Slater point is required which leads to the following additional assumption:

- (iv) For any  $\xi^t \in \Xi^t$  and any decision  $u^{t-1}$  that is feasible with respect to  $\xi^t$ , there exists some  $\hat{u}_t$  depending on  $(u^{t-1}, \xi^t)$  with  $f_t(u^{t-1}, \hat{u}_t, \xi^t) < 0$ .

This is known as *strict nonanticipativity* which can be seen as a stronger version of the nonanticipativity property introduced in 3.1 with regard to the constraint multifunction. There, it was required that  $\{u_t | f_t(u^{t-1}, u_t, \xi^{t-1}) \leq 0\} \neq \emptyset$ , i.e., the feasible set in  $t$  depending on  $u^{t-1}$  and  $\xi^t$  is non-empty. Now, in addition it has to contain inner points:  $\text{int}\{u_t | f_t(u^{t-1}, u_t, \xi^{t-1}) \leq 0\} \neq \emptyset$ . Together with the results from above concerning the inheritance of the saddle property, one obtains:

- (v) The expectation functionals  $E_t \phi_t(u^{t-1}, u_t, \eta^{t-1}, \xi^{t-1})$  are continuous saddle functions – *convex* in  $(u^{t-1}, u_t, \xi^{t-1})$  and *concave* in  $\eta^{t-1}$  with respect to their domain.

The numerical difficulty in solving a stochastic optimization problem of type (3) are the nested minimization and the multidimensional integration of the implicitly given value functions (6). As mentioned, approximation schemes partition the support of the original distribution in convex regions and use distinguished points within them. This is equivalent to a successive discretization of the conditional probability measure  $P_t(\cdot | \omega^{t-1})$  for  $t = 1, \dots, T$ , yielding *discrete* measures  $Q_t(\cdot | \omega^{t-1})$  with support  $\mathcal{A}_t(\omega^{t-1})$ . As a result, one obtains a *scenario tree*  $\mathcal{A}$  (see Figure 2) that can be formally defined as follows:

$$\mathcal{A} := \{(\eta, \xi) \in \Theta \times \Xi | (\eta_t, \xi_t) \in \mathcal{A}_t(\eta^{t-1}, \xi^{t-1}) \forall t > 0\}. \quad (9)$$

Each path in this tree represents a scenario for the evolution of  $\eta_t$  and  $\xi_t$  over the planning horizon  $T$ , and the associated probabilities are given by

$$q(\eta, \xi) := \prod_{t=1}^T q_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}). \quad (10)$$

Clearly, when the conditional probability measure is discrete in time and space, the stochastic two-stage program (5) has a characteristic block structure. As a consequence, the multistage problem (3) can be written as a mathematical program with dynamic block structure and high sparsity whose size depends on the number of scenarios within the tree. Powerful decomposition algorithms have been developed which exploit this special structure (see [2, 51, 55, 56, 57, 58, 59] for example).

## 4 Barycentric approximation

### 4.1 Discretization of distributions

In this section, it is shown how the original probability measure  $P_t(\cdot | \eta^{t-1}, \xi^{t-1})$  can be discretized using so-called *generalized barycenters*. These are calculated with respect to a cross-simplex (or briefly:  $\times$ -simplex), i.e., the Cartesian product of two simplices that cover the support of random data in the objective and the constraints, respectively. To this end, it is assumed that  $\Theta_t(\eta^{t-1}, \xi^{t-1}) \in \mathbb{R}^{K_t}$  and  $\Xi_t(\eta^{t-1}, \xi^{t-1}) \in \mathbb{R}^{L_t}$  are regular simplices covering the support of  $\eta_t$  and  $\xi_t$  (in the sequel, the dependency on previous observations may be omitted in the notation for simplicity). Their vertices are denoted

$$\begin{aligned} u_{\nu_t}(\eta^{t-1}, \xi^{t-1}) &= (u_{1,\nu_t}(\eta^{t-1}, \xi^{t-1}), \dots, u_{K_t,\nu_t}(\eta^{t-1}, \xi^{t-1}))' \in \Theta_t \\ v_{\mu_t}(\eta^{t-1}, \xi^{t-1}) &= (v_{1,\mu_t}(\eta^{t-1}, \xi^{t-1}), \dots, v_{L_t,\mu_t}(\eta^{t-1}, \xi^{t-1}))' \in \Xi_t. \end{aligned}$$

for  $\nu_t = 0, \dots, K_t$ ,  $\mu_t = 0, \dots, L_t$ . The *barycentric weights*

$$\lambda_t(\eta_t | \eta^{t-1}, \xi^{t-1}) = (\lambda_{t,0}(\eta_t | \eta^{t-1}, \xi^{t-1}), \dots, \lambda_{t,K_t}(\eta_t | \eta^{t-1}, \xi^{t-1}))'$$

of  $\eta_t$  with respect to  $\Theta_t(\eta^{t-1}, \xi^{t-1})$  are those nonnegative barycentric coordinates that allow the representation of  $\eta_t$  as a linear combination of the vertices  $u_{\nu_t}(\eta^{t-1}, \xi^{t-1})$  and sum up to one:

$$\begin{aligned} \lambda_{t,0} + \lambda_{t,1} + \dots + \lambda_{t,K_t} &= 1 \\ u_{t,0}\lambda_{t,0} + u_{t,1}\lambda_{t,1} + \dots + u_{t,K_t}\lambda_{t,K_t} &= \eta_t. \end{aligned}$$

Analogously, the barycentric weights

$$\tau_t(\xi_t | \eta^{t-1}, \xi^{t-1}) = (\tau_{t,0}(\xi_t | \eta^{t-1}, \xi^{t-1}), \dots, \tau_{t,L_t}(\xi_t | \eta^{t-1}, \xi^{t-1}))'$$

of  $\xi_t$  with respect to  $\Xi_t(\eta^{t-1}, \xi^{t-1})$  are defined. Briefly, they are given as the unique solution of the systems

$$U_t(\eta^{t-1}, \xi^{t-1}) \cdot \lambda_t = \begin{pmatrix} 1 \\ \eta_t \end{pmatrix} \quad \forall \eta_t \in \Theta_t(\eta^{t-1}, \xi^{t-1}) \quad (11)$$

$$V_t(\eta^{t-1}, \xi^{t-1}) \cdot \tau_t = \begin{pmatrix} 1 \\ \xi_t \end{pmatrix} \quad \forall \xi_t \in \Xi_t(\eta^{t-1}, \xi^{t-1}), \quad (12)$$

where  $U_t(\eta^{t-1}, \xi^{t-1})$  is a regular  $(K_t + 1) \times (K_t + 1)$ -matrix and  $V_t(\eta^{t-1}, \xi^{t-1})$  is a regular  $(L_t + 1) \times (L_t + 1)$ -matrix whose columns contain the vertices of  $\Theta_t$  and  $\Xi_t$ , respectively:

$$U_t(\eta^{t-1}, \xi^{t-1}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ u_0(\eta^{t-1}, \xi^{t-1}) & u_1(\eta^{t-1}, \xi^{t-1}) & \dots & u_{K_t}(\eta^{t-1}, \xi^{t-1}) \end{pmatrix}$$

$$V_t(\eta^{t-1}, \xi^{t-1}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ v_0(\eta^{t-1}, \xi^{t-1}) & v_1(\eta^{t-1}, \xi^{t-1}) & \dots & v_{L_t}(\eta^{t-1}, \xi^{t-1}) \end{pmatrix}.$$

Hence, the barycentric weights are obtained by inverting (11) and (12):

$$\lambda_t(\eta_t | \eta^{t-1}, \xi^{t-1}) = (U_t(\eta^{t-1}, \xi^{t-1}))^{-1} \cdot \begin{pmatrix} 1 \\ \eta_t \end{pmatrix} \quad (13)$$

$$\tau_t(\xi_t | \eta^{t-1}, \xi^{t-1}) = (V_t(\eta^{t-1}, \xi^{t-1}))^{-1} \cdot \begin{pmatrix} 1 \\ \xi_t \end{pmatrix}. \quad (14)$$

The key element of the approximation procedure is the determination of the generalized barycenters and corresponding probabilities. The probability measure  $P_t$  induces mass distributions  $\mathcal{M}_{\nu_t}$  on the  $L_t$ -dimensional simplices  $\{u_{\nu_t}\} \times \Xi_t$  with associated generalized barycenters

$$\xi_{\nu_t} = [\mathcal{M}_{\nu_t}(\{u_{\nu_t}\} \times \Xi_t)]^{-1} \cdot \sum_{\mu_t=0}^{L_t} v_{\mu_t} \int \lambda_{\mu_t}(\eta_t) \cdot \tau_{\nu_t}(\xi_t) dP_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}), \quad (15)$$

where

$$\mathcal{M}_{\nu_t}(\{u_{\nu_t}\} \times \Xi_t) = \int \tau_{\nu_t}(\xi_t) dP_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}) \quad (16)$$

is the mass assigned to  $(u_{\nu_t}, \xi_{\nu_t})$ . For  $\nu_t = 0, \dots, K_t$  these mass distributions add up to a conditional probability distribution. In this way, one obtains a discrete probability measure  $Q_t^l$  on  $\Theta_t \times \Xi_t$  when probability  $\mathcal{M}_{\nu_t}(\{u_{\nu_t}\} \times \Xi_t)$  is assigned to point  $(u_{\nu_t}, \xi_{\nu_t})$ . Analogously, the probability measure  $P_t$  induces mass distributions  $\mathcal{M}_{\mu_t}$  with generalized barycenters

$$\eta_{\mu_t} = [\mathcal{M}_{\mu_t}(\Theta_t \times \{v_{\mu_t}\})]^{-1} \cdot \sum_{\nu_t=0}^{K_t} u_{\nu_t} \int \lambda_{\mu_t}(\eta_t) \cdot \tau_{\nu_t}(\xi_t) dP_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}) \quad (17)$$

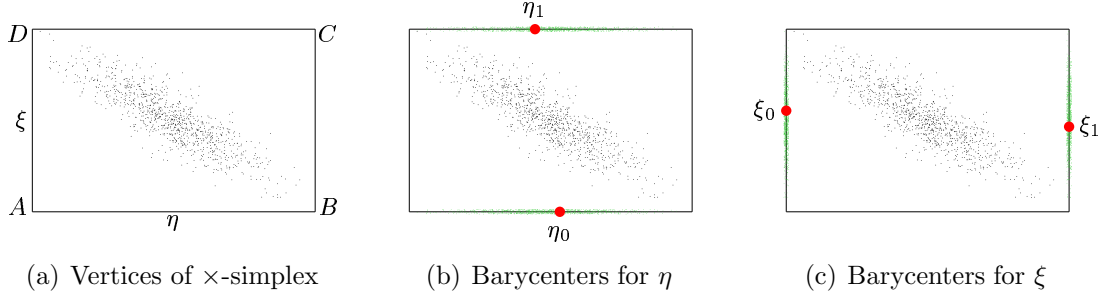


Figure 3: Determination of barycenters for a two-dimensional correlated distribution

on the  $K_t$ -dimensional simplices  $\Theta_t \times \{b_{\mu_t}\}$ . Again, for  $\mu_t = 0, \dots, L_t$  the mass

$$\mathcal{M}_{\mu_t}(\Theta_t \times \{v_{\mu_t}\}) = \int \lambda_{\mu_t}(\eta_t) dP_t(\eta_t, \xi_t | \eta^{t-1}, \xi^{t-1}) \quad (18)$$

is assigned to the points  $(\eta_{\mu_t}, v_{\mu_t})$ , and the mass distributions  $\mathcal{M}_{\mu_t}$  add up to a conditional probability distribution, yielding a discrete probability measure  $Q_t^u$  on  $\Theta_t \times \Xi_t$ . Note that the integrand  $\lambda_{\mu_t}(\eta_t) \cdot \tau_{\nu_t}(\xi_t)$  in (15) and (17) is a bilinear function in  $(\eta_t, \xi_t)$  since the barycentric weights  $\lambda_{\mu_t}$  and  $\tau_{\nu_t}$  are linear in their components. In this way, a discretization of the conditional probability measure  $P_t$  is derived. The two discrete measures  $Q^l$  and  $Q^u$  have support

$$\text{supp } Q_t^l = \left\{ (u_{\nu_t}(\eta^{t-1}, \xi^{t-1}), \xi_{\nu_t}(\eta^{t-1}, \xi^{t-1})) \mid \nu_t = 0, \dots, K_t \right\} \quad (19)$$

$$\text{supp } Q_t^u = \left\{ (\eta_{\mu_t}(\eta^{t-1}, \xi^{t-1}), v_{\mu_t}(\eta^{t-1}, \xi^{t-1})) \mid \mu_t = 0, \dots, L_t \right\}, \quad (20)$$

and according to equations (16) and (18), the corresponding probabilities are given by  $q_t^l(u_{\nu_t}, \xi_{\nu_t}) := \mathcal{M}_{\nu_t}(\{u_{\nu_t}\} \times \Xi_t)$  and  $q_t^u(\eta_{\mu_t}, v_{\mu_t}) := \mathcal{M}_{\mu_t}(\Theta_t \times \{u_{\nu_t}\})$ .

An illustration of the discretization is given in Figure 3 (a). The samples represent the joint distribution of  $\eta$  and  $\xi$  for  $K = L = 1$  (the time index is omitted for simplicity). Note that the sampling indicates (negative) correlation between the random data. Obviously, in the one-dimensional case the simplices covering the support of  $\eta$  and  $\xi$  are edges which results in a  $\times$ -simplex of rectangular shape. For instance, the edges  $AB$  and  $CD$  cover the support of  $\eta$  (the interest rate risk factor in the savings application under consideration), i.e.,  $A$  and  $D$  correspond to vertex  $u_0$  while  $B$  and  $C$  are equivalent to  $u_1$ . Analogously,  $AD$  and  $BC$  represent the domain of (the volume change)  $\xi$  and, hence, correspond to a simplex in  $\mathbb{R}$  with vertices  $v_0$  and  $v_1$ .

It can be seen from Figure 3 (b) and (c) that projecting the distribution mass onto  $AB$  and  $CD$ , taking into account the distance from each sample point to the edges, yields the barycenters  $\eta_0$  and  $\eta_1$ . On the other hand, the barycenters  $\xi_0$  and  $\xi_1$  are obtained from a projection of the mass onto  $AD$  and  $BC$ , respectively. In both

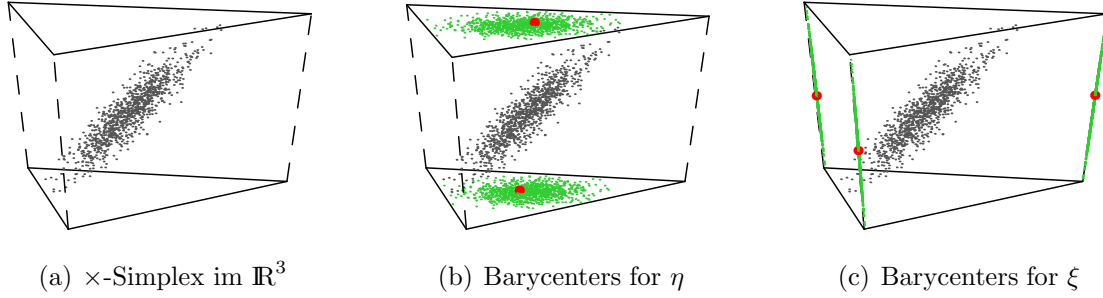


Figure 4: Simplicial coverage for  $K = 2$  und  $L = 1$

cases, the difference in the coordinates of the barycenters reflects the correlation of the original distribution. Another example is shown in Figure 4 for  $K = 2$  and  $L = 1$  where the support of the joint distribution of  $\eta$  and  $\xi$  is covered by a  $\times$ -simplex in  $\mathbb{R}^3$ . Again, the distribution mass is projected onto the simplices, taking into account the distance from each sample to the corresponding simplex. For each simplex, the barycenter is determined as the “center of gravity” of the projected mass, and its probability is equivalent to the proportion of the projection to the total mass.

## 4.2 Barycentric scenario trees

Applying the barycentric approximation technique introduced in the last subsection for the conditional distributions on all stages yields a stochastic process describing the evolution of random data under the new measures. As outlined in 3.3, any stochastic process which is discrete in both time and state can be represented as a scenario tree. From the measures  $Q^l$  and  $Q^u$ , two scenario trees can be constructed whose associated deterministic equivalent problems are lower and upper bounds to the original multistage stochastic program. For simplicity, the distinction between random data affecting the objective and the constraints is no longer maintained in the notation from now on. Using the notation introduced in (9), the support of the discretized distributions for  $t > 1$  is denoted

$$\begin{aligned}\mathcal{A}_t^l(\omega^{t-1}) &= \text{supp } Q_t^l(\cdot | \omega^{t-1}) \\ \mathcal{A}_t^u(\omega^{t-1}) &= \text{supp } Q_t^u(\cdot | \omega^{t-1}).\end{aligned}$$

Starting from currently observed data  $\mathcal{A}_0^l = \mathcal{A}_0^u = \{\omega_0\}$ , the two scenario trees are formally defined as

$$\begin{aligned}\mathcal{A}^l &= \{\hat{\omega}^{l,T} | \hat{\omega}_t^l \in \mathcal{A}_t^l(\hat{\omega}^{l,t-1}) \forall t > 0\} \\ \mathcal{A}^u &= \{\hat{\omega}^{u,T} | \hat{\omega}_t^u \in \mathcal{A}_t^u(\hat{\omega}^{u,t-1}) \forall t > 0\}.\end{aligned}$$

Here,  $\hat{\omega}^{l,t} := (\hat{\omega}_1^l, \dots, \hat{\omega}_t^l)$  and  $\hat{\omega}^{u,t} := (\hat{\omega}_1^u, \dots, \hat{\omega}_t^u)$  are paths from the root to a node of the scenario trees at stage  $t$  determined with the approximation technique introduced

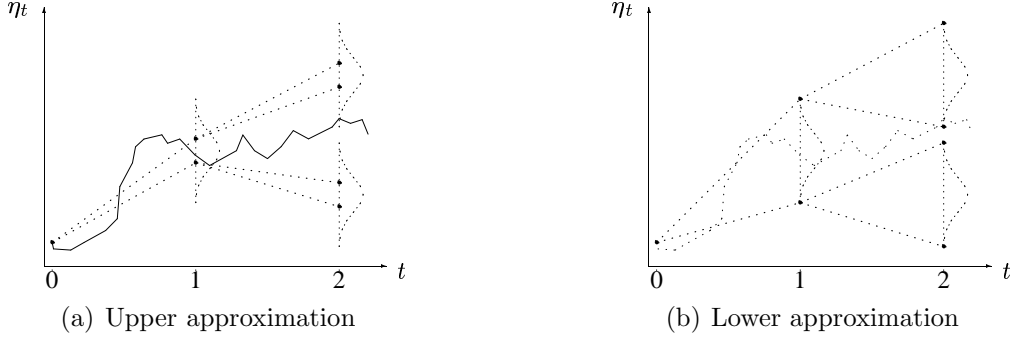


Figure 5: Evolution of risk factor (example) and barycentric scenario trees

above. The *barycentric scenarios* are constructed as follows: Let  $\hat{\omega}^{l,0} = \hat{\omega}^{u,0} = \omega^0$  be current observations and

$$\hat{\omega}_t^l = (u_{\nu_t}(\hat{\omega}^{l,t-1}), \xi_{\nu_t}(\hat{\omega}^{l,t-1})), \quad \nu_t = 0, \dots, K_t, \quad (21)$$

$$\hat{\omega}_t^u = (\eta_{\mu_t}(\hat{\omega}^{u,t-1}), v_{\mu_t}(\hat{\omega}^{u,t-1})), \quad \mu_t = 0, \dots, L_t \quad (22)$$

discrete outcomes at time  $t$  in the nodes of the lower and upper scenario tree, respectively. The corresponding path probabilities of a scenario  $\omega^T$  for each of the trees are given by

$$q^l(\hat{\omega}^T) = \prod_{t=1}^T q_t^l(\hat{\omega}_t^l | \hat{\omega}^{l,t-1}) \quad (23)$$

$$q^u(\hat{\omega}^T) = \prod_{t=1}^T q_t^u(\hat{\omega}_t^u | \hat{\omega}^{u,t-1}), \quad (24)$$

where the conditional probabilities  $q_t^l$  and  $q_t^u$  are derived from (16) and (18).

An illustration of the successive approximation can be found in Figure 5 where only the discretization of the one-dimensional risk factor  $\eta_t$  over the horizon  $T = 2$  is shown for simplicity. At time  $t$ , the value of  $\eta_0$  is known with certainty. For the subsequent stages, only the conditional distributions are given, indicated as density functions (as dotted lines). The upper scenario tree is obtained if the barycenters  $\eta_{\nu_t}$  for the risk factors in the objective are combined with the vertices  $v_{\mu_t}$  of the simplex covering the support of the random data on the right-hand-sides (not shown in Figure 5). Note that the barycenters may differ in coordinate which reflects a correlation between  $\eta_t$  and  $\xi_t$ , i.e., they deviate from the expected value  $E\eta_t$ . Furthermore, a conditional distribution depends on previous realizations, indicated by the different densities in  $t = 2$ . Analogously, in case of the lower scenario tree the vertices  $u_{\nu_t}$  of the simplex covering the support of the uncertain data in the objective are combined with the barycenters  $\xi_{\mu_t}$  for the stochastic right-hand-side coefficients.

### 4.3 Bounds for value functions

Recall the formulation of the original problem (3):

$$\min g_0(u_0) + \int_{\Omega} \left[ \sum_{t=1}^T g_t(u^t, \omega^t) \right] dP(\omega). \quad (25)$$

Replacing the probability measure  $P$  by its discrete approximations  $Q^l$  and  $Q^u$  yields the multistage stochastic programs

$$\psi_0 = \min g_0(u_0) + \int_{\Omega} \left[ \sum_{t=1}^T g_t(u^t, \omega^t) \right] dQ^l(\omega) \quad (26)$$

$$\Psi_0 = \min g_0(u_0) + \int_{\Omega} \left[ \sum_{t=1}^T g_t(u^t, \omega^t) \right] dQ^u(\omega). \quad (27)$$

It is shown in [28] that for the corresponding value functions

$$\begin{aligned} \psi_t(u^{t-1}, \omega^t) &= \min_{u_t} g_t(u^t, \omega^t) + \int_{\Omega_{t+1}} \psi_{t+1}(u^t, \omega^{t+1}) dQ_{t+1}^l(\omega_{t+1} | \omega^t) \\ \Psi_t(u^{t-1}, \omega^t) &= \min_{u_t} g_t(u^t, \omega^t) + \int_{\Omega_{t+1}} \Psi_{t+1}(u^t, \omega^{t+1}) dQ_{t+1}^u(\omega_{t+1} | \omega^t) \end{aligned}$$

for  $t = 0, \dots, T$  with  $\psi_{T+1}(\cdot) = \Psi_{T+1}(\cdot) := 0$ , the following relation holds:

$$\psi_t(u^{t-1}, \omega^t) \leq \phi_t(u^{t-1}, \omega^t) \leq \Psi_t(u^{t-1}, \omega^t). \quad (28)$$

Therefore, (26) is a lower and (27) an upper approximation to the original problem (3). The situation described by the inequalities in (28) is illustrated in Figure 6. Again, for simplicity only the case  $K = L = 1$  is considered and the time index omitted. For each stage  $t$ , the value function is supported by two bilinear functions. In particular, the supporting points for the minorant are the barycenters  $\xi_0$  and  $\xi_1$  while the majorant supports the value function in  $\eta_0$  and  $\eta_1$ . The bilinearity of  $\psi_t$  and  $\Psi_t$  can be explained by the fact that the integrand  $\lambda_{\mu_t}(\eta_t) \cdot \tau_{\nu_t}(\xi_t)$  in (15) and (17) is a bilinear function in  $(\eta_t, \xi_t)$ .

Solving the approximated problems (26) and (27) yields policies  $u^l := (u_0^l, \dots, u_T^l)$  and  $u^u := (u_0^u, \dots, u_T^u)$ , where  $u_t^l$  and  $u_t^u$  denote the decisions made after  $\hat{\omega}_t^l \in \mathcal{A}_t^l$  and  $\hat{\omega}_t^u \in \mathcal{A}_t^u$  are observed in  $t = 1, \dots, T$ . Clearly, lower and upper approximations of the value functions may result in different policies. This can lead to a situation where the accuracy of the approximation has to be improved, in particular if the case  $u_0^l \neq u_0^u$  occurs since for the decision maker only the decision at time  $t = 0$  is of interest. The policies in  $t > 0$  correspond to observations in the barycentric scenarios of the approximated problems and are not likely to be implemented.

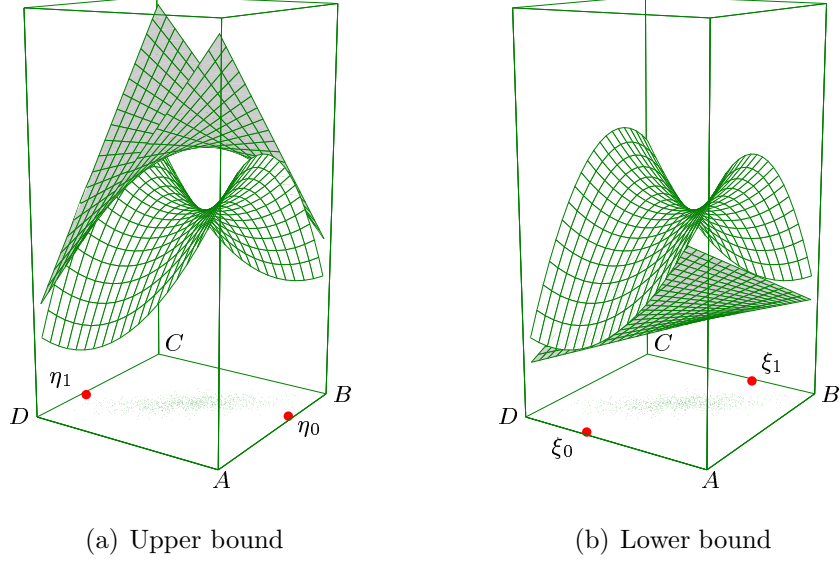


Figure 6: Bilinear approximations of value functions

For a *refinement* of the approximations (26) and/or (27), *epi-convergence* of the approximated value functions can be exploited, i.e.,  $\psi_t(\cdot)$  and  $\Psi_t(\cdot)$  converge to  $\phi_t(\cdot)$  if weak convergence of the discrete conditional probability measures  $Q_t^l(\cdot|\hat{\omega}^{l,t-1})$  and  $Q_t^u(\cdot|\hat{\omega}^{u,t-1})$  to  $P_t(\cdot|\hat{\omega}^{l,t-1})$  and  $P_t(\cdot|\hat{\omega}^{u,t-1})$ , respectively, is ensured for  $t = 1, \dots, T$  (see [28] for details). This requires the partition of the  $\times$ -simplices and the resulting sub- $\times$ -simplices until they become arbitrarily small with respect to the diameters.

The refinement procedure can be outlined as follows: Starting from an initial scenario tree  $\mathcal{A}$ , the  $\times$ -simplex  $\Omega_t(\omega^{t-1})$  covering the support of  $\omega^{t-1} \in \mathcal{A}_{t-1}(\omega^{t-2})$  is split with respect to either the  $\Theta_t$ - or the  $\Xi_t$ -component. Clearly, each of the resulting sub- $\times$ -simplices may be divided again, and barycenters may be derived as described in the previous sections. For the  $\ell_t(\omega^{t-1})$  sub- $\times$ -simplices corresponding to a partition of the support  $\text{supp}(\omega_t|\omega^{t-1})$ , the following conditions must be satisfied:

- (1)  $\bigcup_{i_t=1}^{\ell_t(\omega^{t-1})} \Omega_{t,i_t}(\omega^{t-1}) = \Omega_t(\omega^{t-1}) \supset \text{supp}(\omega_t|\omega^{t-1})$ ,
- (2)  $\Omega_{t,i_t}(\omega^{t-1}) \cap \Omega_{t,j_t}(\omega^{t-1}) = \emptyset$ ,  $i_t \neq j_t; i_t, j_t = 1, \dots, \ell_t(\omega^{t-1})$ ,
- (3)  $\Omega_{t,i_t}(\omega^{t-1})$  are regular  $\times$ -simplices for  $i_t = 1, \dots, \ell_t(\omega^{t-1})$ .

Each partition of a  $\times$ -simplex increases the number of scenarios  $|\mathcal{A}|$  and, hence, the computational complexity of the associated deterministic equivalent. As a consequence,  $\Omega_t$  cannot be divided arbitrarily often in practice. The total number of nodes

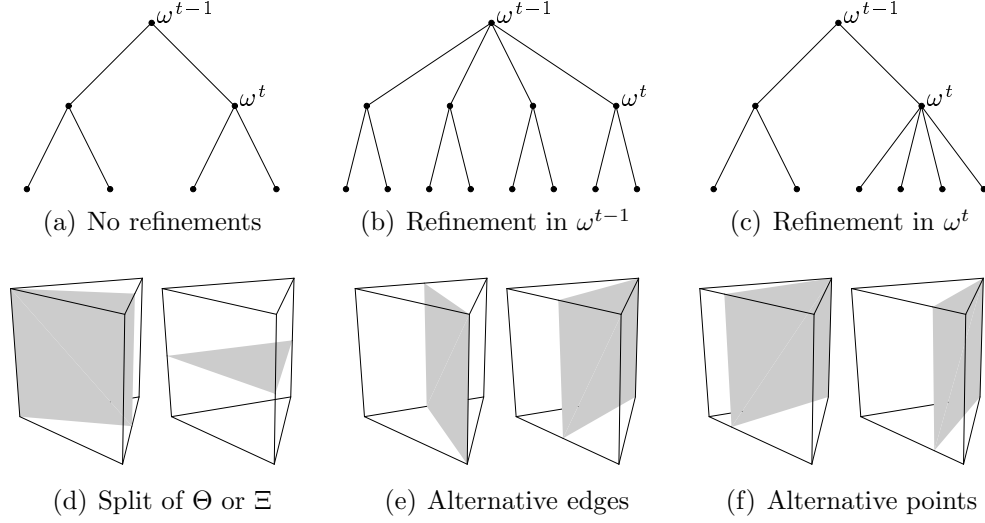


Figure 7: Possible refinements of scenario tree and  $\times$ -simplex

in the scenario trees at stage  $t = 1, \dots, T$  is given by

$$|\mathcal{A}^{l,t}| = \sum_{\hat{\omega}^{l,t-1} \in \mathcal{A}^{l,t-1}} (\ell_t(\hat{\omega}^{l,t-1}) - 1)(K_t + 1) \quad (29)$$

$$|\mathcal{A}^{u,t}| = \sum_{\hat{\omega}^{u,t-1} \in \mathcal{A}^{u,t-1}} (\ell_t(\hat{\omega}^{u,t-1}) - 1)(L_t + 1). \quad (30)$$

Therefore, the refinement process must be carefully monitored, particularly in the multistage case, in order to identify those nodes in the scenario trees where the largest approximation error

$$\epsilon_t(\omega^t) = \Psi_t(u^{t-1}, \omega^t) - \psi_t(u^{t-1}, \omega^t) \quad (31)$$

is observed. On the other hand, if  $\epsilon_t(\cdot) = 0$  for a certain node, the approximation of the value function  $\phi_t(\cdot)$  is exact and further refinements of the partition corresponding to this node will not improve the accuracy of the approximation. For an efficient implementation of refinement strategies, the following aspects must be considered:

- (1) In which node should the scenario tree be refined (i.e., what is the amount of the approximation error to refine the existing partition; this has an immediate impact to the number of scenarios and, hence, the problem size, see Figure 7 (a)–(c)),
- (2) does a division of  $\Omega_t$  with respect to  $\Theta_t$  or  $\Xi_t$  yield a higher accuracy (see Figure 7 (d)),
- (3) which is the edge where the simplex is split (see Figure 7 (e)) and

(4) where does this edge have to be divided (see Figure 7 (f))?

For a detailed discussion and solution techniques, see [33].

## 4.4 Computational results

To complete the introduction of the barycentric approximation technique, some numerical results are presented. Two- and three-dimensional Brownian motions in discrete time are used to model the evolution of key rates  $\eta_t$  and one-dimensional Brownian motions for the volume change  $\xi_t$  in order to illustrate the influence of correlations and the dimension size on the accuracy of the approximation with respect to different refinement strategies. In particular, the distribution of  $(\eta_t, \xi_t)$  at time  $t$  induced by these Brownian motions are independent of the realizations in  $t - 1$ . This case is covered by the general type of distribution functions (7) that is required to ensure the saddle property of value functions.

In the first case ‘2U’, two uncorrelated processes for key rates of maturity 1 and 12 months are considered with  $\sigma_1 = 0.179$  and  $\sigma_{12} = 0.125$ , rates for the remaining maturities are interpolated. Both risk factors are independent of the volume change  $\xi_t$  whose variance is given by  $\sigma_V = 341'056$ . The second case ‘2C’ uses the same volatility for interest rate and volume changes, together with covariances of  $\sigma_{1,12} = 0.117$  between both key rates as well as  $\sigma_{1,V} = 44.247$  and  $\sigma_{12,V} = 37.682$  between key rates and volume. Intuitively, such a model is able to reflect parallel and tilt movements of the yield curve.

Taking a third factor into account, e.g., the rate of an intermediate maturity, also allows the modeling of changes in the curvature of the term structure. In the last case ‘3C’, the 3 month rate with variance  $\sigma_3 = 0.141$  and covariances  $\sigma_{1,3} = 0.151$ ,  $\sigma_{3,12} = 0.121$ ,  $\sigma_{3,V} = 43.2379$  is considered in addition to the 1 and 12 month key rates. These parameter estimates were taken from the description of a collection of test problems for multistage stochastic programs in [30] and are derived from money market rates for the Swiss Franc. All drifts of the Brownian motions are equal to zero. Note that in the context of a real savings application, the correlation between interest rates and volume has a negative sign.

In section 3.2, it was assumed that the  $\times$ -simplex  $\Theta_t \times \Xi_t$  is compact and covers the support of  $(\eta_t, \xi_t)$ . However, the normal distributions at each point in time associated with the Brownian motions under consideration have unbounded support. In this case, one must ensure that  $P_t(\Theta_t \times \Xi_t | \eta^{t-1}, \xi^{t-1}) \geq 1 - \varepsilon$  for sufficiently small  $\varepsilon > 0$  and substitute  $P_t(\cdot | \eta^{t-1}, \xi^{t-1})$  by its normalized truncation (see [35]).

The procedure for the determination of a simplicial coverage is only conceptually outlined here. First, consider a  $K$ -dimensional standard normal distribution. A sphere with radius  $\delta$  around the origin contains a percentage of  $2\Phi(\delta) - 1$  of the total mass distribution, where  $\Phi$  denotes the c.d.f. This sphere can be covered by a simplex in  $\mathbb{R}^K$  with  $K + 1$  vertices. In the one-dimensional case, the simplex reduces to an interval  $[-\delta, \delta]$ , and for  $K = 2$  to a triangle with vertices  $u_0 = (-\sqrt{3}\delta, \delta)'$ ,  $u_1 =$

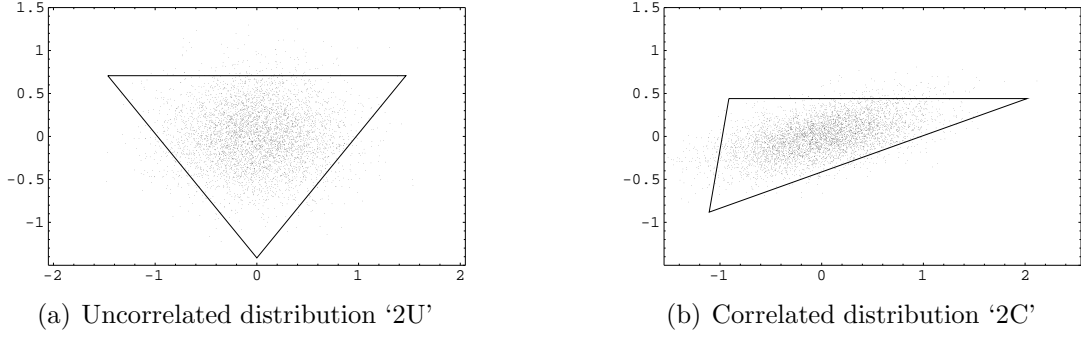


Figure 8: Simplicial coverage of two-dimensional distributions

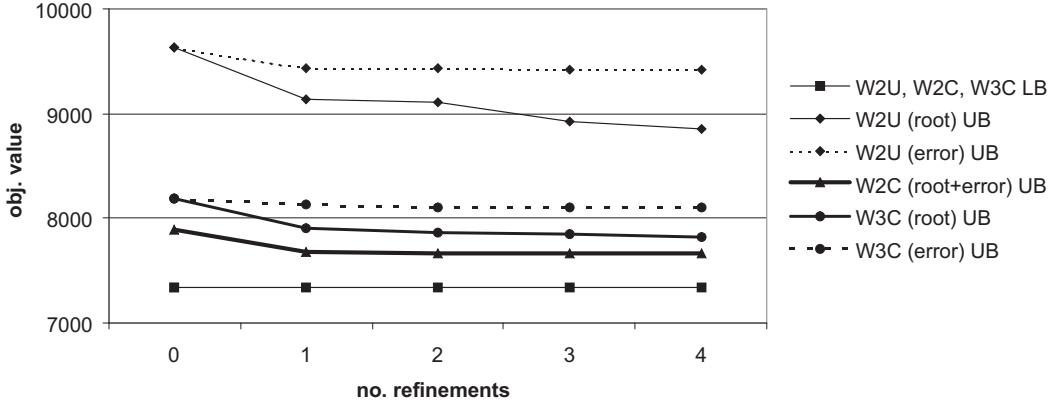


Figure 9: Objective values for different key rate distributions and refinement strategies

$(\sqrt{3}\delta, \delta)'$  and  $u_2 = (0, -2\delta)'$ . It is well known that a standard normally distributed random variable  $Z \in \mathbb{R}^K$  may be transformed into a  $\mathcal{N}(\mu, \Sigma)$ -distributed random variable  $Y \in \mathbb{R}^K$  using the lower triangular matrix of the Cholesky decomposition  $L$  of the covariance matrix  $\Sigma$ , i.e.,  $\Sigma = L \cdot L'$ . According to this rule, the vertices  $u_i^Y$  of the simplicial coverage for  $Y$  are given by

$$u_i^Y = \mu + L \cdot u_i^Z, \quad i = 0, \dots, K.$$

An example for the two-dimensional distributions ‘2U’ and ‘2C’ is shown in Figure 8 for  $\delta = 3$ , covering more than 99.7 % of the probability mass, i.e.,  $\varepsilon < 0.003$ . Note how the correlation between the risk factors in the second case influences the shape of the simplex which, loosely speaking, results in a “stretched” triangle. For a more thorough description of  $\times$ -simplicial coverages, together with formulae for simplices in higher dimensions and a discussion of accuracy estimates, see [31].

Figure 9 summarizes numerical results for the savings problem introduced in (1) with  $\mathcal{D}^S = \{1, 2, 3, 6, 12\}$  and  $m = 1$ . Since there are no short sales, to get positive values the objective function was multiplied by  $-1$  for convenience which yields a

maximization problem. In this case, the upper approximation is obtained from (26) while the lower is given by (27). The initial interest rates are 5.0 %, 5.1 %, 5.2 %, 5.5 % and 6.0 %. An amount of \$ 200'000 matures at all points in time. The graph shows the objective values of the upper and lower bounds depending on the number of refinements of the (sub-)simplices  $\Theta_t(\cdot|\eta^{t-1}, \xi^{t-1})$  with respect to the risk factors for a planning horizon of  $T = 4$ . Increasing the number of stages yields similar results which are not presented here. For a more detailed discussion, including additional models for the dynamics of interest rates (e.g., the well-known Vasicek arbitrage-free model [63] and a two-factor term structure model incorporating mean reversion), refer to the description of test problems in [30].

In case of all distributions under consideration, the lower bounds practically coincide. This can be explained by the fact that this approximation is formed by the vertices of the volume change  $\xi_t$  that remain the same in all cases and the barycenters of the risk factors  $\eta_t$ . The correlations used here are still too weak to cause more significant difference in their positions for the different distributions. Interestingly, the numerical accuracy of the uncorrelated distribution ‘2U’ is significantly worse than those of its correlated counterpart ‘2C’. Incorporating a third factor does not deteriorate the accuracy much for the correlated distribution ‘3C’ compared to the case of independent key rates. Moreover, partitioning the simplices in the root of the scenario trees always yields a much better improvement in the accuracy than splitting those  $\Theta_t$  corresponding to the nodes with the highest approximation error except for ‘2C’ which already exhibits a low gap between both approximations.

Another measure for the goodness of approximations in multistage stochastic programming is the local *expected value of perfect information (EVPI)* taken at the root node. It can be interpreted as the amount one would pay today to obtain perfect information about the uncertain future. In Figure 10, it is stated relative to the optimal objective value of the deterministic equivalent program corresponding to the upper approximation. Again, partitioning in nodes different from the root has only a slight impact on the refinement process, shown here for ‘2U’ only. Both correlated distributions exhibit a significantly higher relative EVPI than the latter which leads to the following – rather surprising – conclusion: Although the larger EVPI implies a higher degree of stochasticity, it becomes easier to approximate the stochastic programs by the barycentric deterministic equivalent problems the higher the correlations are.

Taking the graphical illustrations in Figure 8 into account, an intuitive explanation can be seen in the simplicial coverage of original distributions. The distortion of the simplex that results from the correlations can yield “more extreme” scenarios due to extreme positions of some vertices. Within importance sampling procedures for multistage stochastic programs, a high EVPI as a measure for uncertainty is often used as a criterion to increase the number of scenarios. Here, the EVPI may grow with the number of partitions although the accuracy of the approximation is getting better. In particular, the approximation error can be zero but the EVPI might still have a positive value for some nodes. This questions its suitability as a measure for

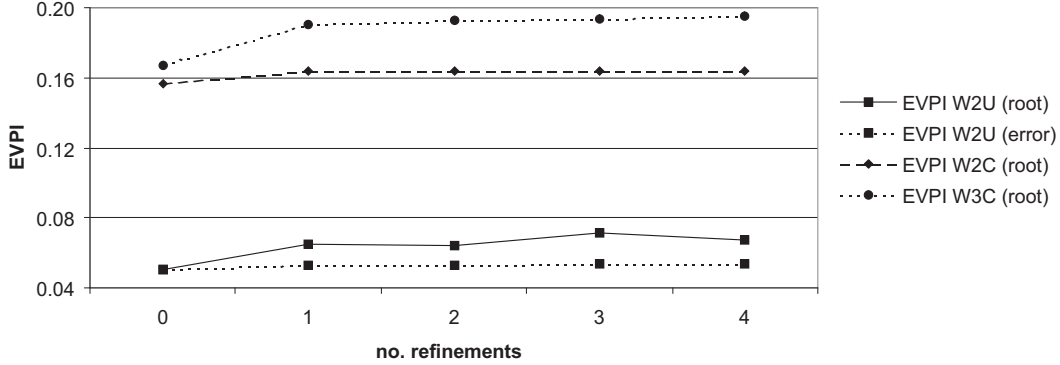


Figure 10: Number of refinements and relative EVPI

the control of the refinement process and underlines the advantage of quantifying the accuracy using exact error bounds.

## 4.5 Solution of deterministic equivalent problems

Barycentric scenario trees define multistage programs that minorize and majorize the given stochastic problem. It has been pointed out already that the corresponding deterministic equivalent programs are *large-scale problems* which limits the number of stages, i.e., the planning horizon  $T$ , the dimension  $M$  of the probability space  $\Omega$ , and/or the accuracy that can be achieved by partitioning the simplicial coverage. In most cases, a trade-off between these quantities must be found when a practical situation is modeled to ensure that the optimization problem can be solved by the available computational resources.

The number of nodes at each stage of the scenario trees was derived in equations (29) and (30). It can easily be seen how the number of partitions and the dimension of the underlying distributions for  $\eta_t$  and  $\xi_t$  influence the problem size that grows exponentially with the planning horizon  $T$  (*curse of dimensionality*). Let  $\sum_{t=0}^T |\mathcal{A}^t|$  be the total number of nodes in the (lower or upper) scenario tree  $\mathcal{A}$ . Omitting liquidity limits of type (5a) and (5b), the size of the deterministic equivalent programs corresponding the savings problem (1) is given by  $(D + 3) \cdot \sum_{t=0}^T |\mathcal{A}^t|$  constraints,  $(D + 2 \cdot |\mathcal{D}^S| + 1) \cdot \sum_{t=0}^T |\mathcal{A}^t|$  variables and  $[3(|\mathcal{D}^S| + D + 1) + m] \cdot \sum_{t=0}^T |\mathcal{A}^t|$  nonzeros (i.e., ‘ $\pm 1$ ’). Note that  $K = 3$  and  $L = 1$  and, hence, the deterministic equivalent corresponding to the upper tree is the larger of the two approximated problems.

The formulae given above also illustrate the block structure and sparsity of the constraint matrix that can be exploited by powerful decomposition algorithms. A detailed discussion of numerical aspects for alternative solution techniques can be found in [34] where also results are presented comparing Cplex, regularized decomposition (cf. [56, 59]), nested decomposition (cf. [2]), and MSLiP-OSL (cf. [9]) for a set of standard problems similar to (1).

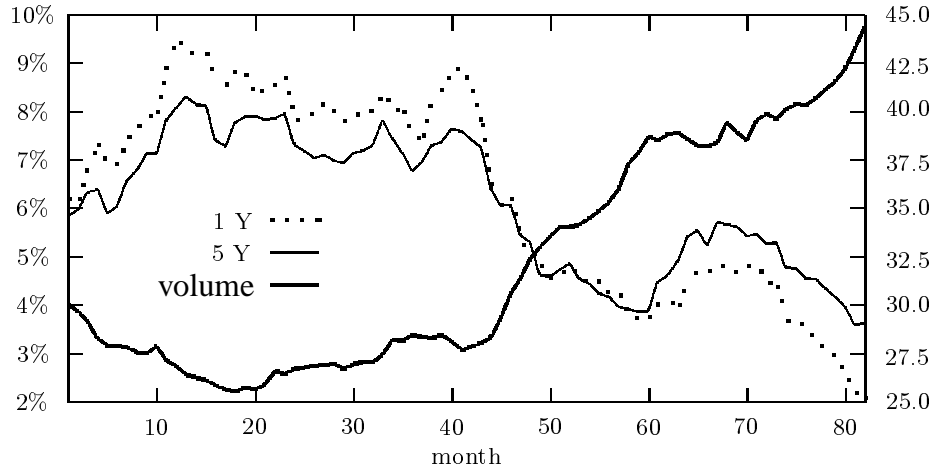


Figure 11: Interest rates (left) and savings volume (bio. CHF, right)

## 5 Reinvestment of savings accounts

The savings model (1) as well as an equivalent stochastic optimization model for refinancing variable-rate mortgages was developed and implemented in co-operation with a major Swiss bank. Before the bank decided to use them for the management of non-maturing account positions, a case study was conducted to assess their performance for the savings application and compare it with the replicating portfolio approach that was employed up to this time. The study was based on monthly interest data (money market and swap rates) for the Swiss Franc from March 1989 to April 1996 and the corresponding volume of a savings position (see Figure 11). The latter was taken from the aggregated savings volume of all financial institutions in Switzerland and the Principality of Liechtenstein as stated in the monthly reports of the Swiss National Bank and scaled to an initial volume of 30 billion.

Investment opportunities consisted of money market and swap positions with maturities of 1, 2, 3, 4, 5, 7, and 10 years. In order to take liquidity restrictions into account that come close to conditions on the Swiss interbank market, upper bonds were set to 500 mio. for maturities up to 5 years and 200 mio. for 7 and 10 years. For the initial portfolio composition, it was assumed that a constant mix policy had been implemented before consisting of 50 % two-year and 50 % five-year instruments, i.e., an amount of 875'000 matures in each of the first 24 months and 250'000 in each of the next 36 months.

It turned out that it is helpful to base the decisions not only on a single optimization where scenarios are determined with expected values of zero for the underlying multidimensional stochastic processes. Instead, a set of 11 stress scenarios characterized by various key rates drifts was defined which reflect shift and tilt movements of the yield curve. Performing the optimization for each of these expectations may result in different optimal investment policies and corresponding expected returns.

<i>scen.</i>	<i>key rate shifts</i>			<i>opt. sol.</i>	<i>objective values</i>		...
	<i>#1</i>	<i>#2</i>	<i>#3</i>		<i>P1</i>	<i>P2</i>	
0	--	--	--	P1	<b>668'734.7</b>	668'672.2	...
1	-	-	-	P2	890'117.8	<b>890'209.2</b>	...
2	-	--	-	P2	902'151.4	<b>902'156.8</b>	...
3	-	-	0	P1	<b>975'132.0</b>	975'051.5	...
4	-	0	+	P3	<b>1'174'434.0</b>	1'174'193.0	...
5	0	0	0	P4	<b>1'152'769.0</b>	1'152'613.0	...
6	+	0	-	P4	<b>1'379'315.0</b>	1'378'796.0	...
7	+	+	0	P4	<b>1'421'833.0</b>	1'421'339.0	...
8	++	+	0	P4	<b>1'648'060.0</b>	1'647'201.0	...
9	+	+	+	P4	<b>1'421'833.0</b>	1'421'339.0	...
10	++	++	++	P4	<b>1'691'181.0</b>	1'690'350.0	...

Table 1: Stress scenarios and risk analysis of different investment policies

This allows the assessment of the sensitivity of the solution with respect to different distribution assumptions and, in particular, to quantify the impact of “extremal events” that may be chosen by the decision maker.

An example for this set of stress scenarios is given in Table 1. Each ‘+’ and ‘-’ represents a decrease or increase in a key rate by one standard deviation. In general, the 11 optimization runs result in more than one solution. Then, the *first-stage decision* is fixed, and the optimization is repeated with respect to the remaining stages for all policies and all different key rate drifts. In other words, for the selected stress scenarios the consequences of suboptimal initial decisions and the costs for their correction in the subsequent periods are determined. This allows an analysis of the risk associated with the different solutions, taking into account non-anticipated shift and tilt movements of the term structure. In this way, one obtains a sort of profit and loss pattern which helps to identify dominant policies for the first-stage decision. For instance, in Table 1 policy P2 performs only slightly better than P1 for scenario 1 and 2 but is clearly inferior in all other cases. A pairwise comparison of the restricted solutions reveals the investment policy that is finally implemented.

The evolution of the margin between the return of the portfolio and the customer rate over the sample period is shown in Figure 12 for the dynamic policies determined by the stochastic optimization model with subsequent risk analysis and is compared to two constant mix strategies. The latter were determined by the replicating portfolio approach. In case of CM2, portfolio weights were derived under the assumption of a constant volume, i.e., it tracks only the customer rate and not the total volume. Note that this results in a lower and more volatile margin. According to Table 2, the average margin could be improved by 25 basis points compared to the better constant mix policy. In particular, the standard deviation of the margin was reduced significantly although volatility is not considered explicitly in the model’s objective. However, it is incorporated by the high number of scenarios in the tree of the stochastic program.

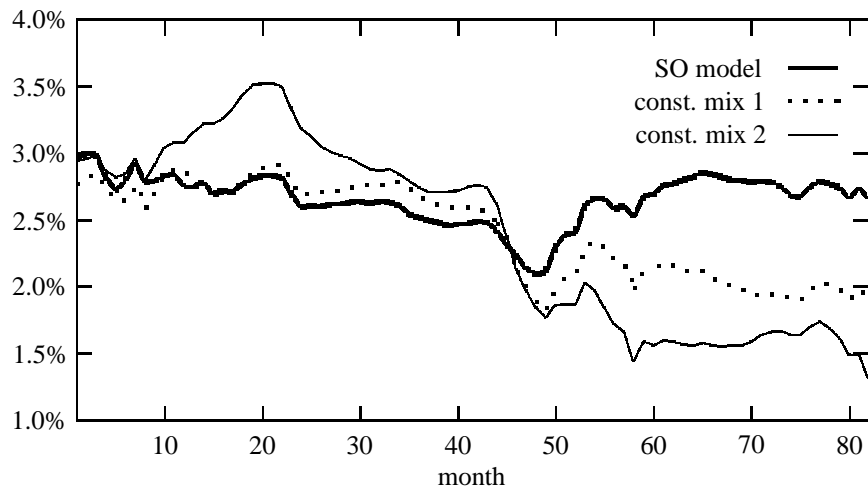


Figure 12: Evolution of margin for replicating portfolios and optimization model

Similar observations have been made for other stochastic optimization models (see [6] for example).

<i>policy</i>	<i>weights const. mix</i>			<i>margin [%]</i>	
	<i>1 y</i>	<i>2 yrs</i>	<i>5 yrs</i>	<i>mean</i>	<i>std. dev.</i>
SO model	–	–	–	2.659	0.188
const. mix 1	0.0	0.5	0.5	2.414	0.358
const. mix 2	0.35	0.35	0.3	2.399	0.696

Table 2: Comparison of dynamic policy with replicating portfolios

## 6 Conclusions and outlook

In this paper, a multistage stochastic optimization model was introduced for the management of non-maturing accounts like savings deposits. The solvability of the stochastic program suffers from the multidimensional integration and nested optimization of implicitly given value functions. In the convex case, which holds under certain restrictions for the underlying (conditional) probability distributions and the form of constraints, structural properties help to overcome these numerical difficulties. In particular, applying barycentric approximation yields distinguished scenario trees. The solution of the associated deterministic equivalent programs provides exact upper and lower bounds to the original problem. Moreover, the accuracy of the approximation may be quantified and improved by refinement techniques. An implementation of more efficient partitioning algorithms remains subject to future research.

The performance of the stochastic optimization model compared to traditional approaches encouraged a major Swiss bank to apply it for the management of their

non-maturing account positions. Two important conclusions can be drawn from the case study presented above: First, the results indicate that a dynamic policy is superior to a static one since portfolios selected with the stochastic program clearly outperform constant mix strategies. Second, the stochastic optimization model hedges against the various sources of uncertainty inherent to non-maturing accounts like interest rate and prepayment/withdrawal risk more appropriately than the replicating portfolio approach. In particular, it is able to take the dependencies between interest rates and volume for dynamic portfolio strategies into account.

Stochastic programming is well suited for a broad class of problems within asset and liability management (ALM) that are characterized by cross and/or serial correlations between risk factors, e.g., cash management in insurance companies where premium payments exhibit a seasonal behavior. Other types of risk (credit, currency etc.) may also be considered if appropriate. Clearly, the Brownian motions exploited here to model the evolution of interest rates and volume may be viewed as a rather simple approach to model the uncertainty of the relevant risk factors. More sophisticated models have been proposed in the financial literature and are currently under investigation. Schürle [60] examines alternative term structure models for the generation of interest rate scenarios and a trend-stationary process for evolution of savings deposits. In the latter case, using the risk factors driving the yield curve as explanatory variables allows for a good description of the aggregated savings volume published by the Swiss National Bank (see also [41] and the references herein for a discussion of alternative processes for the evolution of demand deposits). It is expected that the integration of more sophisticated models for the generation of interest rate and volume scenarios in the stochastic program will yield an additional improvement in the performance. However, different types of distribution functions for the random data might not preserve the inheritance of the saddle property of value functions.

The savings account model introduced here may be seen as a first step towards a general ALM model for a bank's complete balance sheet to optimize investment and refinancing decisions with respect to interest rate and volume risk. Additional constraints may be imposed to limit the absolute risk exposure of certain positions in order to comply with regulatory restrictions concerning capital requirements. Multistage stochastic programming helps to overcome many difficulties in modeling dynamic decision making. It can be seen as a versatile tool for financial planning problems under uncertainty.

## References

- [1] J.R. Birge and R.J.-B. Wets. Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Mathematical Programming Study*, 27:54-102, 1986.

- [2] J.R. Birge, C.J. Donohue, D.F. Holmes, and O.G. Svintsitski. A parallel implementation of the nested decomposition algorithm for multistage stochastic linear programs. *Mathematical Programming*, 75:327-352, 1995.
- [3] M. Britton-Jones. The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54:655-671, 1999.
- [4] D.R. Cariño, T. Kent, D.H. Myers, C. Stacy, M. Sylvanus, A.L. Turner, K. Watanabe, and W.T. Ziemba. The Russell-Yasuda Kasai model: An asset/liability model for a Japanese insurance company using multistage stochastic programming. *Interfaces*, 24:29-49, 1994.
- [5] D.R. Cariño, D.H. Myers, and W.T. Ziemba. Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46:450-462, 1998.
- [6] D.R. Cariño and W.T. Ziemba. Formulation of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46:433-449, 1998.
- [7] Z. Chen, G. Consigli, M.A.H. Dempster, and N. Hicks-Pedron. Towards sequential sampling algorithms for dynamic portfolio management. In: C. Zopounidis (ed.), *Operational Tools in the Management of Financial Risks*, pp. 197-211. Kluwer, 1998.
- [8] V.K. Chopra and W.T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 20:6-11, 1993.
- [9] G. Consigli and M.A.H. Dempster. Solving dynamic portfolio problems using stochastic programming. *Zeitschrift für Angewandte Mathematik und Mechanik (Supplement)*, 77:S 535-S 536, 1997.
- [10] G.B. Dantzig and G. Infanger. Large-scale stochastic linear programs: Importance sampling and Benders decomposition. Technical Report SOL 91-4, Stanford University, 1991.
- [11] G.B. Dantzig and G. Infanger. Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45:59-76, 1993.
- [12] R. Dembo. Scenario optimization. *Annals of Operations Research*, 30:63-80, 1991.
- [13] M.A.H. Dempster. The expected value of perfect information in the optimal evolution of stochastic problems. In: M. Arato, D. Vermes, and A.V. Balakrishnan (eds.), *Stochastic Differential Systems*, pp. 25-40. Springer, 1981.

- [14] C.L. Dert. *Asset liability management for pension funds: A multistage chance constraint programming approach*. Phd thesis, Erasmus University, Rotterdam, 1995.
- [15] J. Dupačová. Stochastic programming models in banking. Working paper, IIASA, Laxenburg, 1991.
- [16] J. Dupačová. Postoptimality for multistage stochastic linear programs. *Annals of Operations Research*, 56:65-78, 1995.
- [17] J. Dupačová. Scenario-based stochastic programs: Resistance with respect to sample. *Annals of Operations Research*, 64:21-38, 1996.
- [18] J. Dupačová, M. Bertocchi, and V. Moriggia. Postoptimality for scenario based financial planning models with an application to bond portfolio management. In: W.T. Ziemba and J.M. Mulvey (eds.), *World Wide Asset and Liability Modeling*, pp. 263-285. Cambridge University Press, Cambridge, 1998.
- [19] N.C.P. Edirisinghe. New second-order bounds on the expectation of saddle functions with applications to stochastic linear programming. *Operations Research*, 44:909-922, 1996.
- [20] N.C.P. Edirisinghe and W.T. Ziemba. Bounds for two-stage stochastic programs with fixed recourse. *Mathematics of Operations Research*, 19:292-313, 1994.
- [21] H.P. Edmundson. Bounds on the expectation of a convex function of a random variable. Technical Report 982, RAND Corporation, 1957.
- [22] D. Eichhorn, F. Gupta, and E. Stubbs. Using constraints to improve the robustness of asset allocation. *Journal of Portfolio Management*, 24:41-48, 1998.
- [23] S.-E. Fleten, K. Høyland, and S.W. Wallace. The performance of stochastic dynamic and fixed mix portfolio models. Working paper, Norwegian University of Science and Technology, Trondheim, 1998.
- [24] B. Forrest, K. Frauendorfer, and M. Schürle. A stochastic optimization model for the investment of savings account deposits. In: P. Kischka et al. (eds.), *Operations Research Proceedings 1997*, pp. 382-387, Springer, 1998.
- [25] E. Fragnière, J. Gondzio, and J.-P. Vial. A planning model with one million scenarios solved on an affordable parallel machine. Technical Report 1998.11, Logilab, University of Geneva, 1998.
- [26] K. Frauendorfer. Solving SLP recourse problems with arbitrary multivariate distributions: The dependent case. *Mathematics of Operations Research*, 13:377-394, 1988.

- [27] K. Frauendorfer. *Stochastic Two-Stage Programming*. Springer, 1992.
- [28] K. Frauendorfer. Multistage stochastic programming: Error analysis for the convex case. *ZOR – Mathematical Methods of Operations Research*, 39:93-122, 1994.
- [29] K. Frauendorfer. Barycentric scenario trees in convex multistage stochastic programming. *Mathematical Programming*, 75:277-293, 1996.
- [30] K. Frauendorfer and G. Haarbrücker. Test problems in stochastic multistage programming. *Optimization*, to appear.
- [31] K. Frauendorfer and F. Härtel. On the goodness of discretizing diffusion processes for stochastic programming. Working paper, Institute of Operations Research, University of St. Gallen, 1995.
- [32] K. Frauendorfer and P. Kall. A solution method for SLP recourse problems with arbitrary multivariate distributions: The independent case. *Problems of Control and Information Theory*, 17:177-205, 1988.
- [33] K. Frauendorfer and C. Marohn. Refinement issues in stochastic multistage linear programming. In: K. Marti and P. Kall (eds.), *Stochastic Programming Methods and Technical Applications (Proceedings of the 3rd GAMM/IFIP Workshop 1996)*, pp. 305-328, Springer, 1998.
- [34] K. Frauendorfer, C. Marohn, and M. Schürle. SG-portfolio test problems for stochastic multistage linear programming (II). Working paper, Institute of Operations Research, University of St. Gallen, 1997.
- [35] K. Frauendorfer and M. Schürle. Barycentric approximation of stochastic interest rate processes. In: W.T. Ziemba and J.M. Mulvey (eds.), *Worldwide Asset and Liability Modeling*, pp. 231-262. Cambridge University Press, Cambridge, 1998.
- [36] B. Golub, M. Holmer, R. McKendall, L. Pohlman, and S.A. Zenios. A stochastic programming model for money management. *European Journal of Operational Research*, 85:282-296, 1995.
- [37] T. Hakala. *A Stochastic Optimization Model for Multi-Currency Bond Portfolio Management*. Phd thesis, Helsinki School of Economics and Business Administration, 1996.
- [38] J.L. Higle and S. Sen. *Stochastic Decomposition – A Statistical Method for Large Scale Stochastic Linear Programming*. Kluwer, 1996.
- [39] T. Ho. Key rate durations: Measures of interest rate risks. *Journal of Fixed Income*, 2:29-44, 1992.

- [40] M.R. Holmer. The asset-liability management strategy system at Fannie Mae. *Interfaces*, 24:3-21, 1994.
- [41] R.A. Jarrow and D.R. van Deventer. The arbitrage-free valuation and hedging of demand deposits and credit card loans. *Journal of Banking and Finance*, 22:249-272, 1998.
- [42] J.L. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30:175-193, 1906.
- [43] P. Kall, A. Ruszczyński, and K. Frauendorfer. Approximation techniques in stochastic programming. In: Y. Ermoliev and R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization*, pp. 33-64. Springer, 1988.
- [44] J.G. Kallberg, R.W. White, and W.T. Ziemba. Short term financial planning under uncertainty. *Management Science*, 28:670-682, 1982.
- [45] M.I. Kusy and W.T. Ziemba. A bank asset and liability management model. *Operations Research*, 34:356-376, 1986.
- [46] R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1:54-61, 1991.
- [47] A. Madansky. Bounds on the expectation of a convex function of a multivariate random variable. *Annals of Mathematical Statistics*, 30:743-746, 1959.
- [48] H.M. Markowitz. *Portfolio Selection: Efficient Diversification of Investment*. Wiley, 1959.
- [49] J.M. Mulvey. Financial planning via multi-stage stochastic programs. In: J.R. Birge and K.G. Murty (eds.), *Mathematical Programming: State of the Art 1994*, pp. 151-171. University of Michigan, Ann Arbor, 1994.
- [50] J.M. Mulvey. Multi-stage financial planning systems. In: R.L. D'Ecclesia and S.A. Zenios (eds.), *Operations Research Models in Quantitative Finance*, pp. 18-35. Physica, 1994.
- [51] J.M. Mulvey and A. Ruszczyński. A new scenario decomposition method for large-scale stochastic optimization. *Operations Research*, 43:477-490, 1995.
- [52] J.M. Mulvey and A.E. Thorlacius. The Towers Perrin global capital market scenario generation system. In: W.T. Ziemba and J.M. Mulvey (eds.), *Worldwide Asset and Liability Modeling*, pp. 286-312. Cambridge University Press, 1998.
- [53] J.M. Mulvey and S.A. Zenios. Capturing the correlations of fixed-income instruments. *Management Science*, 40:1329-1342, 1994.

- [54] R. T. Rockafellar and R.J.-B. Wets. Nonanticipativity and  $L^1$ -martingales in stochastic optimization problems. *Mathematical Programming Study*, 6:170-187, 1976.
- [55] C.H. Rosa and A. Ruszczyński. On augmented Lagrangian decomposition methods for multistage stochastic programs. *Annals of Operations Research*, 64:289-309, 1996.
- [56] A. Ruszczyński. On the regularized decomposition method for stochastic programming problems. In: K. Marti and P. Kall (eds.), *Stochastic Programming: Numerical Techniques and Engineering Applications*, pp. 93-108, Springer, 1993.
- [57] A. Ruszczyński. Parallel decomposition of multistage stochastic programming problems. *Mathematical Programming*, 58:201-228, 1993.
- [58] A. Ruszczyński. Decomposition methods in stochastic programming. *Mathematical Programming*, 79:333-353, 1997.
- [59] A. Ruszczyński and A. Świątanowski. On the regularized decomposition method for two-stage stochastic linear problems. Working Paper WP-96-014, IIASA, Laxenburg, 1996.
- [60] M. Schürle. *Zinsmodelle in der stochastischen Optimierung*. Haupt, 1998.
- [61] M. Steinbach. Recursive direct algorithms for multistage stochastic programs in financial engineering. In: P. Kall and H.-J. Lüthi (eds.), *Operations Research Proceedings 1998*, pp. 241-250, Springer, 1999.
- [62] P. Varaiya and R.J.-B. Wets. Stochastic dynamic optimization - approaches and computation. Working Paper WP-88-87, IIASA, Laxenburg, 1988.
- [63] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177-188, 1977.
- [64] C. Vassiadou-Zeniou and S.A. Zenios. Robust optimization models for managing callable bond portfolios. *European Journal of Operational Research*, 91:264-273, 1996.
- [65] K.J. Worzel, C. Vassiadou-Zeniou, and S.A. Zenios. Integrated simulation and optimization models for tracking indices of fixed-income securities. *Operations Research*, 42:223-233, 1994.
- [66] S.A. Zenios. A model for portfolio management with mortgage-backed securities. *Annals of Operations Research*, 43:337-356, 1993.
- [67] S.A. Zenios. Asset/liability management under uncertainty for fixed-income securities. *Annals of Operations Research*, 59:77-97, 1995.