TOWARDS SEQUENTIAL SAMPLING ALGORITHMS FOR DYNAMIC PORTFOLIO MANAGEMENT

Z.Chen, G.Consigli, M.A.H.Dempster, N.Hicks-Pedrón.
Finance Research Group, Judge Institute of Management Studies, University of Cambridge

Abstract. This paper describes in detail the computations required to generate and solve large scale strategic financial portfolio management problems by sequential importance sampling methods. Data and model generation processes are emphasized and expected value of perfect information importance sampling criteria under current development outlined.

Keywords: Dynamic stochastic programming, datapath generation, expected value of perfect information, sequential sampling.

1. Introduction

Dynamic stochastic programming (*DSP*) formulations are particularly suitable for the solution of strategic portfolio management problems requiring the consideration of a large set of state variables. By focussing on a limited set of decision stages, they allow the characterization of portfolios with a rich set of investment and liability classes [8]. At each stage the portfolio manager (of a financial institution, insurance company, industrial conglomerate, etc.) takes a decision – in the form of a portfolio allocation – in the face of uncertainty typically generated by the random behaviour of market prices.

The solution of this dynamic decision problem depends crucially on the stochastic process model adopted to describe the behaviour of the random variables relevant to each stage of the problem. The random behaviour of the rates of return in stock portfolios was first identified by Markowitz [27] as the main source of uncertainty for the definition of an optimal investment decision in static stock portfolio problems. In modern applications stock prices provide only one possible source of risk for the portfolio manager [8, 28]; in many cases the randomness of short and long term interest rates, exchange rates, and other possible factors, needs to be considered for a correct representation of the problem. The definition of a stochastic, possibly multidimensional, financial data process and its inclusion in the generation of a stochastic optimization problem for numerical solution represents an important and controversial aspect of applied stochastic programming techniques for portfolio management [20, 24, 8, 28].

In §2 we consider a set of related issues concerning *scenario generation* in stochastic programming models when arbitrary underlying models of uncertainty are considered. To this end we introduce a distinction between a random vector *data process*, representing a primary source of uncertainty, and a random *coefficient process*, dependent on the former, which is problem-dependent and whose behaviour generates the specific information upon which the portfolio manager bases his strategy.

This distinction clarifies the important interaction between scenario generation and the subsequent *problem solution* when an *importance sampling* criterion based on the *Expected Value of Perfect Information* (*EVPI*) process [14, 10, 16, 9] is introduced, as discussed in §3. The properties of the EVPI process allow the selection of an enhanced set of *relevant representative* data paths in a sequential sampling refinement of an original stochastic optimization problem.

In our applications we consider a constrained stochastic optimization problem in the form of a *dynamic recourse problem* (*DRP*) (*cf.* Dempster [13], Ermoliev and Wets [19]) whose canonical formulation is given by (bold characters denote random elements)

$$\max_{x_1 \in \mathbb{R}^{n_1}} \{ f_1(x_1) + \mathbb{E}_{\boldsymbol{\xi}_2} [\max_{\mathbf{x}_2} (\mathbf{f}_2(\mathbf{x}_2) + \ldots + \mathbb{E}_{\boldsymbol{\xi}_T | \boldsymbol{\xi}^{T-1}} [\max_{\mathbf{x}_T} \mathbf{f}_T(\mathbf{x}_T)])]\}$$

$$\begin{array}{llllllll}
\text{s.t.} & & & & & & & \\
A_1 x_1 & & & & & = & b_1 & \\
\mathbf{B}_2 x_1 & + & \mathbf{A}_2 \mathbf{x}_2 & & & = & \mathbf{b}_2 & a.s. \\
 & & \mathbf{B}_3 \mathbf{x}_2 & + & \mathbf{A}_3 \mathbf{x}_3 & = & \mathbf{b}_3 & a.s. \\
 & & & \ldots & & & \ldots & \\
 & & & \mathbf{B}_T \mathbf{x}_{T-1} & + \ \mathbf{A}_T \mathbf{x}_T & = & \mathbf{b}_T & a.s.
\end{array}$$

$$l_1 \leq x_1 \leq u_1$$

$$\mathbf{l}_t \leq \mathbf{x}_t \leq \mathbf{u}_t \ \ a.s., \quad t = 2, \ldots, T \ . \tag{1}$$

In (1) the constraint region is appropriately represented by a set of linear constraints representing financial as well as strategic and regulatory constraints [8]. The process $\boldsymbol{\xi}$ in $(\Omega^\xi, \mathcal{F}^\xi, P^\xi)$ is typically defined as a discrete, possibly autocorrelated, vector process with sample space Ω^ξ. The filtration $\mathcal{F}_t^\xi := \sigma\{\boldsymbol{\xi}^t\}$ generated in Ω^ξ by the history $\xi^t := (\xi_1, \ldots, \xi_t)$ of the random process $\boldsymbol{\xi}$ at time t defines the information set available to the decision maker at the different stages of the problem. In financial planning problems the process $\boldsymbol{\xi}$ is considered to be a function, $\boldsymbol{\xi}_t := \xi(\boldsymbol{\omega}_t)$, of a process $\boldsymbol{\omega}$ defined in a different probability space $(\Omega, \mathcal{F}^\omega, P^\omega)$. We refer to $\boldsymbol{\xi}$ as the *coefficient process* and

$\boldsymbol{\omega}$ as the *data process* of the problem.

The objective in (1) is defined through a sequence of *nested* optimization problems corresponding to the different stages. Each decision $x_t \in X_t$ is required to be feasible with respect to a sequence of stage-dependent constraints: $A_1 \in I\!R^{m_1 \times n_1}$ and $b_1 \in I\!R^{m_1}$ define deterministic constraints on the first stage decision x_1, while, for $t = 2, \ldots, T$, $\mathbf{A}_t : \Omega^\xi \rightarrow I\!R^{m_t \times n_t}$, $\mathbf{B}_t : \Omega^\xi \rightarrow I\!R^{m_t \times n_{t-1}}$ and $\mathbf{b}_t : \Omega^\xi \rightarrow I\!R^{m_t}$ define stochastic constraint regions for the recourse decisions $\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_T$. $I\!E_{\boldsymbol{\xi}_T|\boldsymbol{\xi}^{T-1}}$ denotes conditional expectation of the state $\boldsymbol{\xi}_T$ of the coefficient process $\boldsymbol{\xi}$ with respect to the history $\boldsymbol{\xi}^{T-1}$. At each stage previous decisions affect remaining optimization problems through the stochastic matrices $\mathbf{B}_t$, $t = 2, \ldots, T$.

The sequence of random events and decisions is given in Figure 1.

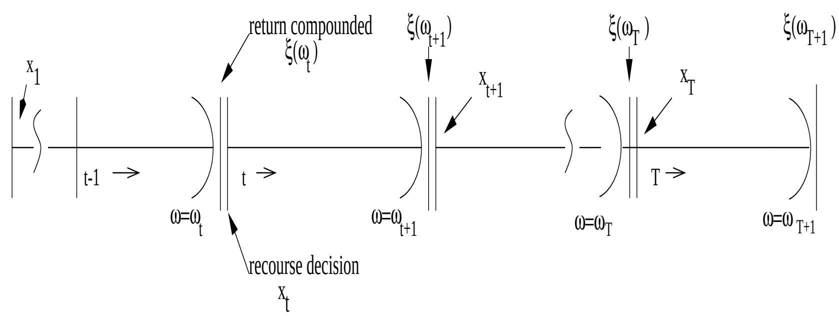


Figure 1. Sequence of decisions and random events in dynamic stochastic programming

The decision process $\mathbf{x} := \{\mathbf{x}_t\}_{t=1}^T$ is required to be strictly *adapted* or *nonanticipative*, i.e. $\mathbf{x}_t = \{\mathbf{x}_t | \mathcal{F}_t^\xi\}$ *a.s.*, with respect to the filtration $\mathcal{F}_t^\xi$ generated by the process. This condition can be imposed in the model implicitly [8, 7] or explicitly, leading to a stochastic program in *split-variable* form [14, 3].

Dynamic portfolio problems are easily formulated as a DRP. Applications of this approach can be found in Bradley and Crane [4], Lane and Hutchinson [26], Kusy and Ziemba [25], Dempster and Ireland [15], Mulvey and Vladimirou [29], Zenios [33], Cariño *et al.* [6]. The *CALM* model (Dempster [8]) has been formulated as a linearly constrained mixed integer stochastic programming problem and adopted for the formulation of a 10 year pension fund asset and liability problem – the *Watson* model [8] – with uncertainty generated according to Wilkie's autoregressive model [32] – and a 20 year asset allocation problem – the *FRC* model – with uncertainty generated according to the extended Brennan and Schwartz model [5]. The CALM-FRC model has been developed for a Frank Russell Company sponsored project and is defined with three asset classes: consol bonds, stocks and bank deposits, and an underlying four dimensional Itô process for the short and long interest rates, the dividend yield and the stock price. Due to its simpler associated data generator, it is being used as the reference model for the algorithm development described briefly in the final section.

2. Specification of the data process for dynamic stochastic programmes

The distinction introduced in §1 between the processes $\boldsymbol{\xi}$ and $\boldsymbol{\omega}$ is both conceptual and methodological. Unlike $\boldsymbol{\xi}$ which is constructed as a *discrete time* path-dependent process in accordance with the DRP formulation of the problem, the data process $\boldsymbol{\omega}$ in $(\Omega, \mathcal{F}^{\omega}, P^{\omega})$ may be given different characterizations, all referring to a conceptually underlying *continuous time* process. This is the sense in which we refer to *dynamic* – in contrast to *multistage* – recourse problems. In [5, 17, 20, 28] $\boldsymbol{\omega}$ is an element of the class of real-valued diffusion processes with time set $[1, \tau]$, $\tau < \infty$ and uncertainty is generated by a (multivariate) Wiener process $\mathbf{W}_t$.

In [8], following Wilkie [32], $\boldsymbol{\omega}$ belongs to the class of autoregressive processes of the j-th order with continuous state space and discrete time set $T = \{1, 2, \ldots, \tau\}$, with random behaviour induced by disturbances $\mathbf{e}_t \sim \mathcal{N}(0, \sigma^2(\omega))$ and depending on the financial variable (e.g. the long interest rate) we may have an autoregressive equation up to the third order in the model. In Zenios [33] and Klaassen [24] $\boldsymbol{\omega}$ is a discrete state binomial process.

All these cases may be described in a form suitable for simulation purposes as

$$\omega_{n_t+1} - \omega_{n_t} = \mu(\omega)\Delta n_t + \sigma(\omega)\varepsilon_{n_t} \quad , \tag{2}$$

for $t = 1, \ldots, T$ and $n_t = 1, \ldots, N_t$ up to $N_T - 1$, where $\mu(\omega)$ defines the *drift* of the process, $\sigma(\omega)$ its *volatility*, $\Delta n_t := (n_t + 1) - n_t$, and each stage of (1) refers to N_t subperiods.

For N_t sufficiently large and $\varepsilon_{n_t} \sim \mathcal{N}(0, \Delta n_t)$, (2) describes the discrete version of a diffusion process driven by Wiener noise [20, 28]. For smaller N_t and ε_{n_t} with arbitrary probability distribution, we typically have autoregressive models for long term allocation problems or binomial or trinomial models.

The different discretization schemes are all made consistent with a (DRP) characterization of the decision problem by introducing a *compound return* function defined by

$$\Gamma_t := \Pi_{n_t=1}^{(N_t-1)}(1+\omega_{n_t}) - 1 \qquad t = 1, \ldots, T\ , \tag{3}$$

which gives at the end of period t the return of a monetary unit invested at the beginning of the period; where for each t each ending cash position is carried over to the position at the first index of the next period.

The history ω^t of the data process, for $t = 1, 2, \ldots, T$, enters a specific portfolio problem as parameters $\xi_t = \xi(\omega^t)$, required for the recourse decision x_t, as described in Figure 1. The decision maker is assumed to follow the behaviour of the price processes over time, while recourse decisions are allowed only at the end of every period consistent with the nonanticipativity requirement. Inhomogeneous time stages are easily accomodated in this framework and alternative stochastic models as described above can be adopted as individual inputs for the definition of the discrete vector process $\boldsymbol{\omega} := \{\boldsymbol{\omega}_t\}_{t=1}^T$.

The specification of $\boldsymbol{\omega}_t$ is the output of a data generator *datagen*, in terms of a set of random functions with coefficient estimates for its mean and volatility functions, and a random number generator of the type described briefly in §2.2.

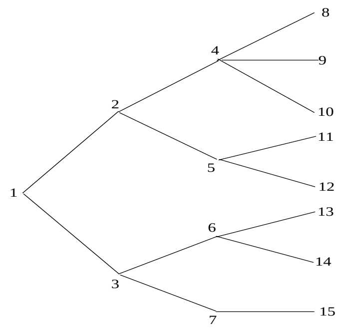
Datagen takes as inputs the initial state of the process together with a *nodal partition matrix* identifying the associated tree structure. It is interfaced with the generator of the random coefficients of the problem – the scenario generator *scengen* – needed for the definition of the stochastic program for numerical solution. *Scengen* takes as input the complete data process specification along the scenario tree and generates, as output, the scenario-dependent coefficients required by the mathematical formulation of the problem.

2.1. Datapath generation

We consider in this section an iterative procedure, interfaced with the data simulator, for the correct generation of data paths in the form of a scenario tree.

The definition of a *scenario tree nodal partition matrix* as a two-dimensional array, with number of rows equal to the number of scenarios of the problem and number of columns equal to the number of stages, is at the core of the conditional simulator. The matrix identifies uniquely the tree structure for the associated stochastic program and is used by the data generator in order to derive the states of the data process in conditional mode, and by the model generator *STOCHGEN* [10, 8] for the definition of the corresponding *SMPS* files [1] necessary for the numerical solution of the problem [8].

Figure 2 provides an example of the matrix specification associated with an arbitrary tree structure.



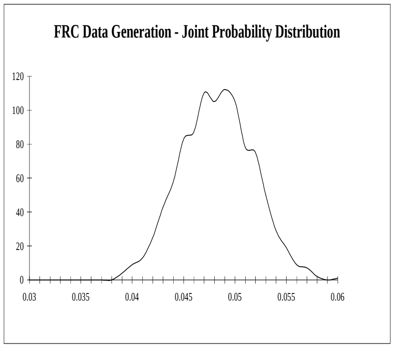
$$M := \begin{bmatrix} 1 & 2 & 4 & 8 \\ 1 & 2 & 4 & 9 \\ 1 & 2 & 4 & 10 \\ 1 & 2 & 5 & 11 \\ 1 & 2 & 5 & 12 \\ 1 & 3 & 6 & 13 \\ 1 & 3 & 6 & 14 \\ 1 & 3 & 7 & 15 \end{bmatrix}$$

Figure 2. Definition of the nodal partition matrix

Following the matrix order, conditional simulations are run, compound annual rates of return computed and initial conditions passed consistently along the tree. Consistent with the time partition of the planning horizon, the generator runs over the N_t subperiods for $t = 1, \ldots, T$ and the final state ω_{N_t} (see eq. 2) for one simulation is adopted as initial state for the following run. The nodal partition matrix allows both the conditional run of the data generator – with one run for every increment in the matrix entries, columnwise – and the consistent updating of the initial seeds – rowwise. The stage-oriented nodal labeling order is convenient in view of the sequential generation-solution procedure described in §3.

In the FRC problem below a simulator has been constructed for a 4-dimensional diffusion system driven by a Wiener process with state variables representing stock return, short and long interest rate and dividend yield processes and based on estimated instantaneous mean vector and correlation matrix [5]. In this case the Wiener noise was generated by implementing a *congruential* method based on the Park and Miller *minimal standard method* [30] for the generation of normal unit deviates $\nu \sim \mathcal{N}(0,1)$ and applying the transformation $W_t = \nu dt$ leading to $W_t \sim \mathcal{N}(0, dt)$.

The set of data paths generated by *datagen* permits estimation of the joint probability distribution of the return at the horizon of a portfolio which initially invests equal amounts in each asset. Based on $1,024$ data paths an estimated joint probability distribution generated at the horizon by the multidimensional conditional generator for the FRC and the Watson problems is displayed in Figure 3. In the case of the Watson problem the generation of the normal random variates from Wilkie's model was based on Marsaglia's *polar method* [32].



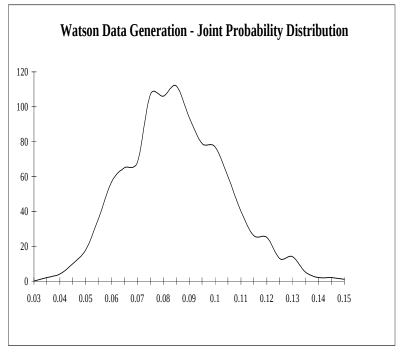


Figure 3. FRC and Watson problems empirical balanced portfolio return probability distributions generated by *datagen*.

The accuracy of the sequential procedure described in §3 relies on the possibility of an unbiased approximation of the form displayed in Figure 3 of the continuous probability density generated by the random process underlying the optimization problem. This density, as shown in Figure 3, is in general not consistent with the usual assumptions of normality or log-normality made in finance theory.

2.2. Coefficient process specification

The distinction between the data process $\boldsymbol{\omega}$ in $(\Omega, \mathcal{F}^\omega, P^\omega)$, $\quad \boldsymbol{\omega} := \{\omega_{n_t} : n_t = 1, \ldots, N_t,\ t = 1, \ldots, T\}$ and the corresponding coefficient process $\boldsymbol{\xi}$ in $(\Omega^\xi, \mathcal{F}^\xi, P^\xi)$, $\quad \boldsymbol{\xi} := \{\xi_t : t = 1, 2, \ldots, T\}$ with $\boldsymbol{\xi}_t := \xi(\boldsymbol{\omega}^t)$ is motivated by the following considerations in formalizing financial planning problems:-

- Recent DRP formulations of asset and liability models have adopted a characterization of uncertainty based on complete market arbitrage free models of interest rates and price processes developed and well-established in the financial literature [2, 23, 32]. These, however, explain only in part the risk embedded in financial positions of investors operating worldwide and across different markets [8, 28].

- The specification of an optimal policy in recourse models generally relies on the definition of complex hierarchical forecasting and model generation systems [28, 8] for the definition of a set of (coefficient) scenarios derived in a cascade structure from the specification of a set of underlying *core* random processes possibly defined in different probability spaces.

The filtration $\mathcal{F}_t^\xi$ generated for $t = 1, 2, \ldots, T$ by the histories of the coefficient process $\boldsymbol{\xi}$ contains the information necessary for the solution of the corresponding dynamic portfolio problem. In general $\mathcal{F}_t^\omega \subset \mathcal{F}_t^\xi$.

Important examples of data and coefficient process specification and generation within portfolio management tools are the two instances of the CALM model – Watson and FRC, the Towers-Perrin model of Mulvey [28], the general asset and liability model of Klaassen [24], the MBS model of Zenios [33] and the Yasuda-Kasai model of the Frank Russell Company [7]. All these applications require the derivation from the relevant data generator of a large set of coefficients which are needed for the mathematical specification of the problem.

This step, which generally results in the definition of *ad hoc*, problem dependent, valuation criteria has an impact on the properties of the stochastic program finally generated [9].

In (1) the process $\boldsymbol{\xi}$ is defined by $\boldsymbol{\xi}_t := (\boldsymbol{\zeta}_t, \mathbf{A}_t, \mathbf{B}_t, \mathbf{b}_t)$, with $\boldsymbol{\zeta}_t$ denoting a random parameter in the objective functional given by $f_t(\zeta_t, x_t)$. The specification of the random coefficient matrices $\mathbf{A}_t, \mathbf{B}_t$ and $\mathbf{b}_t$ in $(\Omega^\xi, \mathcal{F}^\xi, P^\xi)$ refers to the generation of the complete information structure necessary for the solution of the portfolio problem.

The steps required by conditional scenario generation may be briefly summarized as:-

- Initially the number of scenarios and stages, with associated stage discretization $n_t = 1, \ldots, N_t$, for $t = 1, \ldots, T$, are defined.

- The nodal partition matrix is then specified in order to define the complete tree structure for the problem (note that one simulation here corresponds to one complete data path along the event tree). Then recursively:-

  - the vector of initial conditions is defined and *datagen* is run, travelling the tree forward from the root node to the terminal node;
  - for each such simulation the compounded returns are computed;
  - the simulations are associated with the stages according to the nodal partition matrix and the complete set of conditional data paths specified;
  - for every trajectory of the data process, *scengen* is run, the corresponding set of model coefficients defined and

- a set of scenarios are generated and interfaced with a matrix generator (e.g. MODLER [22]).

Given the model formulation and the generation of the model coefficients, the resulting stochastic programming problem is now defined in standard input format for numerical solution [1]. In our system the SMPS format is generated using STOCHGEN [10].

3. Information flows and the resolution of uncertainty

We consider a stochastic programming system for the solution of financial planning problems based on:-

1. The representation of the decision problem in dynamic recourse form with implicit or explicit characterization of the nonanticipativity condition [14, 9, 18].

2. A data path simulator for an underlying continuous data vector process $\boldsymbol{\omega}$ in $(\Omega, \mathcal{F}^{\omega}, P^{\omega})$ representing the core uncertainty of the portfolio allocation problem.

3. A scenario generator for the specification of the vector stochastic process $\boldsymbol{\xi}$ in $(\Omega^{\xi}, \mathcal{F}^{\xi}, P^{\xi})$ defining the coefficients of the model and interfaced with this simulator.

4. Generation of the SMPS format, for which we use the STOCHGEN library [11] incorporating Greenberg's MODLER [22], required for the numerical solution of the stochastic programming problem.

5. The solution of the problem either by a *primal-dual interior point* (IP) method (CPLEX 4.0, 1996) [12], or by *nested Benders decomposition* (MSLiP-OSL, Version 8.3, 1995) [21, 31].

We intend to show here how the phases 2, 3, 4 and 5 are integrated, based on the valuation of the information generated by the coefficient data process, when the sample space approximation of this process is sequentially refined using estimates of the EVPI process(es) below.

Consider problem (1) in the more compact *dynamic programming* representation which takes advantage of the Markov structure exhibited by the set of constraints. For each $t = 1, \ldots, T$ we have the set of *nodal* problems

$$\begin{aligned} \pi_t(\boldsymbol{\xi}^t) \;:=\; & \max_{x_t \in X_t} \mathbb{E}\{f_t(\boldsymbol{\xi}^t, x^{t-1}, \mathbf{x}_t) + v_{t+1}(\boldsymbol{\xi}^t, \mathbf{x}^t) \mid \mathcal{F}_t^{\xi}\} \\ & \text{s.t.} \\ & \mathbf{B}_t x_{t-1} \;+\; \mathbf{A}_t \mathbf{x}_t = \mathbf{b}_t \;\; a.s., \end{aligned} \tag{4}$$

where v_{t+1} expresses the optimal expected value for the remaining optimization problem for the stages from $t+1$ to T. At the horizon $v_{T+1}(\boldsymbol{\xi}^T, \mathbf{x}^T) := 0$. In (4) the dependence of the decision vector $\mathbf{x}_t$ on the filtration $\mathcal{F}_t^\xi$ is expressed explicitly.

The *expected value of perfect information* (EVPI) process [14, 10, 16] is defined by

$$\eta_t(\boldsymbol{\xi}^t) := \phi_t(\boldsymbol{\xi}^t) - \pi_t(\boldsymbol{\xi}^t) \quad t = 1, 2, \ldots, T, \tag{5}$$

where $\phi_t(\boldsymbol{\xi}^t)$ corresponds to the set of *distribution* problems associated with the relaxation of the nonanticipativity condition to the case of *perfect foresight*

$$\phi_t(\boldsymbol{\xi}^t) := I\!E[\max_{x_t \in X_t} \{ f_t(\boldsymbol{\xi}^t, x^{t-1}, \mathbf{x}_t) + v_{t+1}(\boldsymbol{\xi}^t, \mathbf{x}^t) \mid \mathcal{F}_T^\xi \}] \,. \tag{6}$$

Based on the behaviour of $\boldsymbol{\eta}_t$ we can both assess the level of *stochasticity* of the DRP problem [8, 10] and define a sampling procedure for the selection of a sample set of *relevant representative* data paths in a sequential procedure. From the definition of the EVPI process we have at the horizon, by construction, $\boldsymbol{\eta}_{T+1} := 0$. For the properties of the $\boldsymbol{\eta}$ process which justifies its adoption as an importance sampling criterion for selection of a sample set of *objective-relevant* sample paths we refer to [14, 10, 9, 16]. Of particular importance is the characterization of the process as a nonnegative *supermartingale* [14] which reflects the nonnegative and increasing value associated with early resolution of uncertainty.

This property has two impacts useful in defining a sampling procedure: when the EVPI value is zero at one node in the tree, say ξ^t, it will remain null in all descendant nodes. Furthermore, if $\eta_t(\xi^t) = 0$ for some ξ^t, then there is a decision x_t optimal at t for all subsequent nodes. The future uncertainty is thus *irrelevant* and the local problem can be replaced by a deterministic problem.

The same properties of the EVPI process are shared by the *marginal EVPI*, δ-EVPI, or *shadow price of information* process [14] defined by the dual variables of the stochastic programming problem associated with the nonanticipativity constraints of the model in *split variable* form. Unlike (4) we now consider an *explicit* characterization of the nonanticipativity condition in *conditional expectation form*:

$$x_t(\boldsymbol{\xi}^t) = \sum_{\hat{\boldsymbol{\xi}}^t} p(\hat{\boldsymbol{\xi}}^t) x_t(\hat{\boldsymbol{\xi}}^t) \quad t = 1, 2, \ldots, T, \tag{7}$$

where $\hat{\boldsymbol{\xi}}^t$ denotes, at each stage t, the set of scenarios descending from the current node $\boldsymbol{\xi}^t$. Accordingly $p(\hat{\boldsymbol{\xi}}^t)$ denotes the probability of each such scenario occurring conditional on the fact that the process is in state ξ^t at time t. Definition (7) is referred to as the nonanticipativity condition in *conditional expectation projection* form (*cf.* [14]).

The nonanticipativity condition (7) leads to the specification, for $t = 1, 2, \ldots, T$, of a sequence of stochastic dynamic programs in the form

$$\begin{array}{l} \max_{x_t \in X_t} \mathbb{E}\{f_t(\boldsymbol{\xi}^t, x^{t-1}, \mathbf{x}^t) + v_{t+1}(\boldsymbol{\xi}^t, \mathbf{x}^t) \mid \mathcal{F}_t^{\xi}\} \\ \text{s.t.} \\ \mathbf{B}_t x_{t-1} + \mathbf{A}_t \mathbf{x}_t = \mathbf{b}_t \;\; a.s. \\ (I_t - \mathbf{\Pi}_t)\mathbf{x}_t = 0 \;\; a.s. \, . \end{array} \tag{8}$$

The programme (8) has associated Lagrangean given by

$$\begin{array}{rl} \mathcal{L}(x_t, y_t', \rho_t') := & \mathbb{E}\{[f_t(\boldsymbol{\xi}^t, x^{t-1}, \mathbf{x}^t) + v_{t+1}(\boldsymbol{\xi}^t, \mathbf{x}^t)] + \\ & \mathbf{y}_t'(\mathbf{B}_t x_{t-1} + \mathbf{A}_t \mathbf{x}_t - \mathbf{b}_t) + \boldsymbol{\rho}_t'(I_t - \mathbf{\Pi}_t)\mathbf{x}_t \mid \mathcal{F}_t^{\xi}\} \, . \end{array} \tag{9}$$

The *marginal EVPI process* $\boldsymbol{\rho} := \{\boldsymbol{\rho}_t\}_{t=1}^{T+1}$ is thus the dual process associated with the nonanticipativity condition in conditional expectation form. At the optimum the δ-EVPI coefficients provide a measure of the value generated by a perturbation of the constraint. Unlike the full EVPI process, the marginal process is defined at every node of the tree up to and *including* the last stage. This property makes the criterion suitable for the solution of two stage problems by δ-EVPI sampling. At present the estimation of the δ-EVPI process requires the generation and solution of the complete deterministic equivalent problem [8, 3] with explicit nonanticipativity constraints.

We are now in position to sketch a sequential procedure based on the solution of the stochastic optimization problem with either the MSLiP-OSL solver [31] or the Cplex IP solver [12]. The two solvers are interfaced respectively with the EVPI sampling algorithm developed by Dempster and Corvera-Poiré [10, 11] and the δ-EVPI sampling algorithm currently under development.

Based on the EVPI information, the sampling procedure allows the sequential refinement of an original tree structure according to the procedure outlined in Table 1.

In both sampling procedures the *permanence* after resampling of the nodal EVPI values in the neighbourhood of 0 leads to a deterministic optimization problem over the remaining periods up to the horizon.

Each iteration with either importance sampling criterion requires: the generation of the data paths for the data process, the derivation of the coefficient scenarios, the definition of the standard input SMPS format and the solution by nested Benders decomposition or the IP

Table I. EVPI-Sampling Algorithm

define	number of iterations in the algorithm: J
define	initial scenario tree structure: T_1
	The Algorithm
	$j = 1$
	while $j \leq J$
1.	**construct** a tree T_j based on EVPI information
2.	**solve** problem T_j and compute its nodal EVPI
3.	**if** EVPI *near* 0, **resample**
4.	**else if** EVPI *near* 0 after resampling
	take one sample scenario
5.	**else if** EVPI $>$ 0, increase branching at the node
6.	$j = j + 1$
	CONTINUE

method including the current estimates of the nodal EVPI values. Sequential refinement of the previous tree structure is based on an analysis of the current EVPI process – full or marginal – and the definition of a new nodal partition matrix that allows *datagen* to run again, as described in the inner loop of Figure 4.

The adoption of the full, as opposed to the marginal, EVPI sampling criterion has been previously reported [10, 8, 16]. Results have been presented in the case of a sampling procedure *independent* of the phase of scenario generation considered in §2.

4. Conclusions and further research

The sequential procedure outlined in Figure 4 calls for a few final remarks.

The system under development relies on the definition of a master program that calls at every iteration of the sampling procedure the subroutines for the data process generation – *datagen*, the coefficient process generation – *scengen*, the model generator – *STOCHGEN* and the solver, analyzes the EVPI estimates and derives the nodal partition matrix for the next iteration. The same framework is adopted for the use of the marginal EVPI importance sampling criterion derived from the solution of the problem with an IP method.

The efficiency of the sequential solution procedure relies heavily on

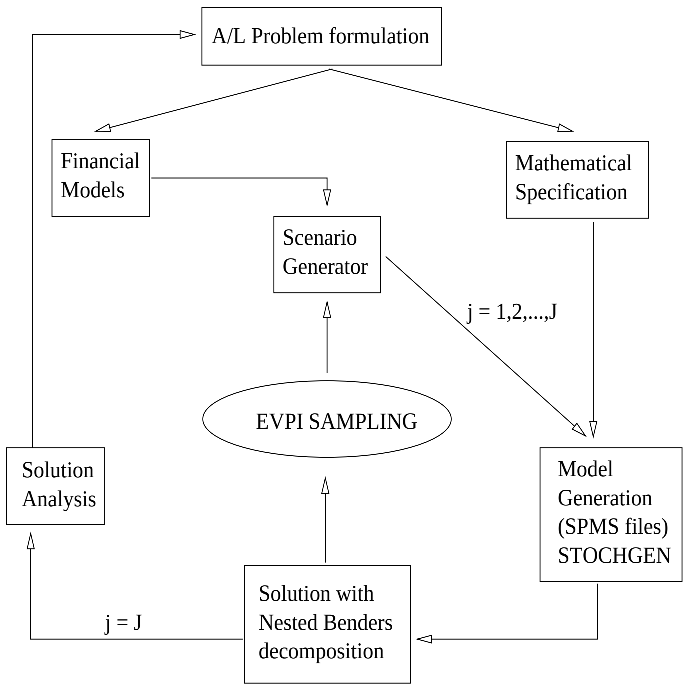


Figure 4. EVPI based sequential solution procedure

the speed and accuracy of the model generation. This step is currently based on MODLER [22] which was not originally designed for sequential matrix generation. We will shortly be in position to integrate the recursive MPS generator AIMS into our system with a very positive impact on the speed and efficiency of the sequential solution procedure.

In previous work [9, 16] we have established the accuracy of the EVPI sampling rule as a criterion for the approximation of large scale stochastic problems with an EVPI-based selection of scenarios sampled from a pregenerated finite population. In this paper the sampling framework has been extended to a dynamic procedure in which the sample of the random process generating the uncertainty in a portfolio allocation problem is associated with an increasingly representative stochastic sample problem.

Acknowledgements

Research partially supported through contract “HPC-Finance” of the INCO ’95 (no 951139) project funded by Directorate General III (Industry) of the European Commission, the UK EPSRC and the FECIT Laboratory of Fujitsu Systems (Europe) Limited. Partial support was also provided by the “HPC-Finance” partner institutions: Universi-

ties of Bergamo (IT), Cambridge (UK), Calabria (IT), Charles (CZ), Cyprus (CY), Erasmus (ND), Technion (IL) and the "Centro per il Calcolo Parallelo e i Supercalcolatori (IT)".

References

1. J.R. Birge, M.A.H. Dempster, H.I. Gassmann, E.A. Gunn, A.J. King and S. Wallace. A standard input format for multiperiod stochastic linear programs. Mathematical Programming Society, *Committee on Algorithms Newsletter* **17** (1987) 1-20.
2. F. Black, E. Derman and W. Toy. A one factor model of interest rates and its application to Treasury bond options. *Financial Analysts Journal*, Jan-Feb (1990) 33-41.
3. A. Berger, J.M. Mulvey, E. Rothberg and R. Vanderbei. Solving multistage stochastic using tree dissection. Statistics and Operations Research Research Report, Princeton University, Princeton, NJ (1995).
4. S.P. Bradley and D.B. Crane. A dynamic model for bond portfolio management. *Management Sciences* **19.2** (1972) 139-151.
5. M.J. Brennan, E.S. Schwartz and R. Lagnado (1996). Strategic Asset Allocation. *Journal of Economic Dynamics and Control*, forthcoming.
6. D. Cariño, T. Kent, D. 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** (1994) 24-49.
7. D. Cariño, D.H. Myers and W.T. Ziemba. Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. Research Report, Frank Russell Company, Tacoma, Washington, May (1995).
8. G. Consigli and M.A.H. Dempster. Dynamic stochastic programming for asset-liability management. To appear in *Annals of Operations Research*. Proceedings of APMOD95 Conference, Brunel University of West London (1996).
9. G.Consigli and M.A.H. Dempster. Solving dynamic portfolio problems using stochastic programming. To appear in *Zeitschrift fur Angewandte Mathematik und Mechanik*. Proceedings of the GAMM96 Conference, Charles University, Prague, May (1996).
10. X. Corvera-Poiré. Model Generation and Sampling Algorithms for Dynamic Stochastic Programming. PhD Thesis, Dept. of Mathematics, Univ. of Essex, U.K. (1995).
11. X. Corvera-Poiré. STOCHGEN User's Manual. Dept. of Mathematics, Univ. of Essex, U.K. (1995).
12. Cplex Optimization, Inc. *Using the Cplex Callable Library, Version 4.0* Incline Village NE, USA (1996).
13. M.A.H. Dempster. Stochastic programming: An introduction. In M.A.H. Dempster, ed. *Stochastic Programming*. Academic Press, London (1980) 3–59.
14. M.A.H. Dempster. On stochastic programming: II. Dynamic problems under risk. *Stochastics* **25** (1988) 15–42.
15. M.A.H. Dempster and A. Ireland. Object oriented model integration in a financial decision support system. *Decision Support Systems* **7** (1991) 329–340.
16. M.A.H. Dempster and R.T. Thompson. EVPI-Based importance sampling solution procedures for multistage stochastic linear programmes on parallel MIMD

architectures. To appear in *Annals of Operations Research.* Proceedings of the POC96 Conference, Versailles (1996).

17. M.A.H. Dempster. The CALM-FRC Model. Internal Document. Finance Research Group, Judge Institute of Management Studies, University of Cambridge, U.K. (1996).
18. J. Dupacova. Multistage stochastic programs: The state-of-the-art and selected bibliography. *Kybernetica* **31** (1995) 151–174.
19. Yu. Ermoliev and R.J-B. Wets, eds. *Numerical Techniques for Stochastic Optimization.* Springer-Verlag, Berlin (1988).
20. K. Frauendorfer, C. Marohn and M. Schurle. SG-Portfolio Test problems for stochastic multistage linear programming, Institute of OR, Univ. of St. Gallen, Switzerland (1995).
21. H.I. Gassmann. MSLIP: a computer code for the multi-stage stochastic linear programming problem. *Mathematical Programming* **47** (1990) 407-423.
22. H.J. Greenberg. A Primer for MODLER: Modelling by Object-Driven Linear Elemental Relations. Mathematics Department, University of Colorado at Denver (1995).
23. T.S.Y. Ho and S.B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance* **41** (1986) 1011–1029.
24. P. Klaassen. Stochastic Programming Models for Interest-Rate Risk Management. PhD Thesis, Sloan School of Management, M.I.T., Cambridge, Massachusetts, May (1994). Published as IFSRC Discussion Paper.
25. M.I. Kusy and W.T. Ziemba. A Bank Asset and Liability Management Model. *Operations Research* **34** (1986) 356–376.
26. M. Lane and P. Hutchinson. A model for managing a certificate of deposit portfolio under uncertainty. In M.A.H. Dempster, ed. *Stochastic Programming.* Academic Press, London (1980) 473–493.
27. H.M. Markowitz. Portfolio Selection. *Journal of Finance* **7** (1952) 77–91.
28. J.M. Mulvey. Generating Scenarios for The Towers Perrin Investment System. *Interfaces* **26.2** (1996) 1–15.
29. J.M. Mulvey and H. Vladimirou. Stochastic Network Optimization Models for Investment Planning. *Annals of Operations Research* **20** (1989) 187–217.
30. S.K. Park and K.W. Miller. *Communications of the ACM* **31** (1988) 1192–1201.
31. R.T. Thompson. MSLiP-OSL 8.3 User's Guide. Judge Institute of Management Studies, University of Cambridge, U.K. (1997).
32. A.D. Wilkie. More on a stochastic asset model for actuarial use. Institute of Actuaries, London (1995).
33. S.A.Zenios. Asset-Liability management under uncertainty for fixed-income securities. *Annals of Operations Research* **59** (1995) 77–97.